

## A Rhamnogalacturonan Lyase in the *Clostridium cellulolyticum* Cellulosome

Sandrine Pagès,<sup>1,2\*</sup> Odile Valette,<sup>1</sup> Laetitia Abdou,<sup>1,2</sup>  
Anne Bélaïch,<sup>1</sup> and Jean-Pierre Bélaïch<sup>1,2</sup>

Laboratoire de Bioénergétique et Ingénierie des Protéines, IBSM, Centre National de la Recherche Scientifique,<sup>1</sup> and Université de Provence,<sup>2</sup> Marseille, France

Received 10 March 2003/Accepted 28 May 2003

*Clostridium cellulolyticum* secretes large multienzymatic complexes with plant cell wall-degrading activities named cellulosomes. Most of the genes encoding cellulosomal components are located in a large gene cluster: *cipC-cel48F-cel8C-cel9G-cel9E-orfX-cel9H-cel9J-man5K-cel9M*. Downstream of the *cel9M* gene, a new open reading frame was discovered and named *rgl11Y*. Amino acid sequence analysis indicates that this gene encodes a multidomain pectinase, Rgl11Y, containing an N-terminal signal sequence, a catalytic domain belonging to family 11 of the polysaccharide lyases, and a C-terminal dockerin domain. The present report describes the biochemical characterization of a recombinant form of Rgl11Y. Rgl11Y cleaves the  $\alpha$ -L-Rhap-(1 $\rightarrow$ 4)- $\alpha$ -D-GalpA glycosidic bond in the backbone of rhamnogalacturonan I (RGI) via a  $\beta$ -elimination mechanism. Its specific activity on potato pectic galactan and rhamnogalacturonan was found to be 28 and 3.6 IU/mg, respectively, indicating that Rgl11Y requires galactan decoration of the RGI backbone. The optimal pH of Rgl11Y is 8.5 and calcium is required for its activity. Rgl11Y was shown to be incorporated in the *C. cellulolyticum* cellulosome through a typical cohesin-dockerin interaction. Rgl11Y from *C. cellulolyticum* is the first cellulosomal rhamnogalacturonase characterized.

Pectin is the most complex class of plant cell wall polysaccharides. Three pectic polysaccharides—namely, homogalacturonan (HG), rhamnogalacturonan I (RGI), and RGII—have been isolated from primary cell walls (35). HG, also called the smooth region, is a linear chain of  $\alpha$ -(1,4)-linked D-galactopyranosyluronic acid (GalpA) residues. The carboxyl groups of some GalpA are methyl esterified and some residues are O-acetylated. The RGI, or hairy region, is composed of alternating L-rhamnopyranosyl (L-Rhap) and GalpA residues. Some GalpA residues of the disaccharide repeating unit [ $\rightarrow$ 4)- $\alpha$ -D-GalpA-(1 $\rightarrow$ 2)- $\alpha$ -L-Rhap-(1 $\rightarrow$ )] are O-acetylated (9, 35). “Hairy” refers to the many side chains of neutral polymers like arabinans, galactans, and arabinogalactans, which are attached to the rhamnose residues. Indeed, branched and linear oligosaccharides composed predominantly of  $\alpha$ -L-arabinofuranosyl (L-Araf) and  $\beta$ -D-galactopyranosyl residues (D-Galp) are linked to some Rhap residues. The RGII backbone is related to HG [linear chain of  $\alpha$ -(1,4)-linked D-GalpA residues]. Four structurally different oligosaccharide side chains, comprising at least 11 different monosaccharides, including rare sugars such as apiose and aceric acid (10, 35), are linked to the GalpA residues of the backbone.

In primary plant cell walls, cellulose, hemicellulose (mannan, xylan, and xyloglucan, etc.), and pectin form a polysaccharidic extracellular matrix. In order to decompose plant cell wall polysaccharides, saprophytic and phytopathogenic microorganisms secrete a range of cellulases, hemicellulases, and

pectinases. Some saprophytic clostridia such as *Clostridium thermocellum*, *C. cellulovorans*, and *C. cellulolyticum* secrete plant cell wall degrading enzymes assembled in multisubunit systems called cellulosomes (2, 4, 8, 15, 28, 36). The enzymes incorporated in the cellulosomes are mainly cellulases and hemicellulases (6, 11, 17, 18, 19, 22, 23, 25, 27, 39). So far, only one cellulosomal pectinase has been characterized (37), namely, *C. cellulovorans* pectate lyase Pel4A, which is an HG-degrading enzyme belonging to family 4 of the polysaccharide lyases according to the classification established by Coutinho and Henrissat (<http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html>). Polysaccharide lyases (EC 4.2.2.) are enzymes that cleave polysaccharide chains via a  $\beta$ -elimination mechanism, resulting in the formation of a double bond at the nonreducing end of the product. Pectate lyases catalyze the  $\beta$ -eliminative cleavage of  $\alpha$ -(1,4)-glycosidic bonds between GalpA residues in pectate (a low methylesterified form of pectin), and generate  $\Delta$ 4,5 unsaturated GalpA as the product. Such enzymes have not been identified in the *C. cellulolyticum* cellulosomes, which contain several cellulases (4, 15) and a mannanase bound to a scaffolding protein CipC. Many of the latter enzymes have been extensively characterized (3, 4, 12, 14, 16, 34) as well as the scaffolding protein CipC (30, 31). CipC interacts with cellulose via its cellulose-binding module (CBM), allowing the anchorage of the cellulosome onto its preferential substrate. The genes encoding the cellulosomal components of *C. cellulolyticum* are organized in a 26-kb cluster. This large cluster, *cipC-cel48F-cel8C-cel9G-cel9E-orfX-cel9H-cel9J-man5K-cel9M* (1, 3), contains the genes encoding the scaffolding protein CipC, endo- and processive cellulases, mannanase, and a cohesin-containing protein of unknown function called OrfXp. Three isolated genes encoding cellulosomal cellulases CelA and CelD (12) and the noncellulosomal

\* Corresponding author. Mailing address: Centre National de la Recherche Scientifique, IBSM, UPR 9036, Laboratoire de Bioénergétique et Ingénierie des Protéines, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France. Phone: 33 4 91 16 45 47. Fax: 33 4 91 71 33 21. E-mail: pages@ibsm.cnrs-mrs.fr.

cellulase Cell were found elsewhere on the *C. cellulolyticum* chromosome (C. Tardif, personal communication).

The present report provides data indicating that a new gene *rgl11Y* (formerly named *orfY*), located downstream *cel9M* in the *cel* gene cluster of *C. cellulolyticum* encodes a cellulosomal endo-acting rhamnogalacturonan lyase, called RglY, belonging to family 11 of the polysaccharide lyases. Rgl11Y is the first cellulosomal rhamnogalacturonan lyase discovered to date.

#### MATERIALS AND METHODS

**Bacterial strains and plasmids.** *C. cellulolyticum* ATCC 35319 was used as the source of genomic DNA and for the production of cellulosomes. *Escherichia coli* DH5 $\alpha$  (Bethesda Research Laboratories) and *E. coli* BL21(DE3) (Novagen) were used as the host for pET-22b(+) (Novagen) derivative expression vector pETYhp3. *E. coli* was grown at 37°C in Luria-Bertani (LB) medium supplemented with ampicillin (100 or 200  $\mu$ g/ml) when required.

***rgl11Y* gene cloning and DNA sequencing.** *rgl11Y* was located downstream of the *cel9M* gene in the large *C. cellulolyticum* gene cluster by inverse PCR from the total chromosomal DNA of *C. cellulolyticum* digested by *Pst*I, with oligonucleotides 4060 (5'-GAC GGT GAA ATC AAT GCT CTT GAC CTT-3') and 214457 (5'-TTG CGT AAT CGA GAG CAT CTA CTA GCT TCC-3'). This way, the 3' region of the *cel9M* gene and the 5' region of the *rgl11Y* gene was sequenced. Restriction map of this DNA fragment revealed a *Hind*III site located 700 bp upstream of the *Pst*I site used for the inverse PCR. To obtain the 3' region of the *rgl11Y* gene the total chromosomal DNA of *C. cellulolyticum* was digested by *Hind*III, and used for another inverse PCR. This second PCR was performed with oligonucleotides 422299 (5'-TTG TCC TGT TCT GTC ATC ACG G-3') and 421467 (5'-GGG TTT TGA AAT GTG GTC ATC AG-3') located between the restriction sites *Hind*III and *Pst*I, and the 3' region of *rgl11Y* and the 5' region of the *cel5N* gene (the last gene of the 26-kb cluster) were sequenced this way. The accession number for this sequence is AF 316823.

The region of the *rgl11Y* gene that encodes the mature form of Rgl11Y was thus amplified by PCR with primers Y1 (5'-aattccatagACGGCAGCATCGGCACGC-3') and Y7 (5'-gggctcgagCGAAAGCAAACCTTGCTTTTAG-3'), generating a *Nde*I and an *Xho*I (underlined) site upstream and downstream of the Rgl11Y coding sequence, respectively. A 2-kb DNA fragment was amplified.

Since the *rgl11Y* gene contains an *Nde*I site, two steps were necessary to clone the 2-kb DNA fragment amplified by PCR into vector pET-22b(+). The PCR fragment was first digested by *Nde*I and *Hind*III, and the resulting 0.9-kb *Nde*I-*Hind*III fragment was cloned into *Nde*I-*Hind*III-linearized pET-22b(+), generating plasmid pETYa. Secondly, the 2-kb DNA fragment amplified by PCR was digested by *Hind*III and *Xho*I and the resulting 1.1-kb fragment was cloned into *Hind*III-*Xho*I-linearized pETYa.

The resulting plasmid, named pETYhp3, contains the gene encoding the mature form of Rgl11Y cloned in frame with 6 histidine codons. The inserted sequence was sequenced (Genome Express, Grenoble, France) to verify the lack of unwanted mutations and used to transform *E. coli* BL21(DE3).

**Rgl11Y production and purification.** (i) **Production and purification from inclusion bodies.** *E. coli* BL21(DE3) cells containing the recombinant plasmid pETYhp3 were grown in 50 ml of LB medium supplemented with ampicillin (100  $\mu$ g/ml) at 37°C, 200 rpm, during 4 h. This preculture (4 ml) was used to inoculate 0.5 liter of LB medium (containing ampicillin 100 mg and 10 ml of glycerol 60%), which was incubated at 37°C up to an optical density of 0.8 at 600 nm. Cells were cooled to 15°C and IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) was then added to a final concentration of 100  $\mu$ M, and the culture was continued for 16 h at 180 rpm. The cells were collected by centrifugation (10 min at 5,000  $\times$  g) and resuspended in 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub>. DNase I (5  $\mu$ g/ml) was added to the crude extract, and cells were broken with a French press. After centrifugation, the pellet containing inclusion bodies was incubated for 30 min at room temperature with 8 M urea and then centrifuged (10 min at 5,000  $\times$  g).

The solubilized inclusion bodies were loaded onto 5 ml of Ni-CAMHC resin (Sigma) equilibrated in 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub>. The resin was washed with 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub>, and the recombinant protein Rgl11Y was eluted with 20 mM imidazole-30 mM Tris-HCl (pH 8.0). The eluate was dialyzed against 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub> and concentrated by ultrafiltration in an Amicon cell (membrane cutoff, 10 kDa). Purified Rgl11Y was used to produce polyclonal antibodies.

(ii) **Production of soluble Rgl11Y and purification.** *E. coli* BL21(DE3) cells containing recombinant plasmid pETYhp3 were grown in 50 ml of LB medium

with ampicillin (100  $\mu$ g/ml) at 37°C, 200 rpm, during 4 h. This preculture (4 ml) was used to inoculate 0.5 liter of LB medium containing ampicillin (200  $\mu$ g/ml), sorbitol (1 M), and betaine (2.5 mM), which was incubated at 37°C to an optical density of 0.6 at 600 nm. After cooling to 18°C, IPTG was added to the culture to a final concentration of 100  $\mu$ M. After 16 h of growth at 18°C, 180 rpm, the cells were collected by centrifugation (10 min at 5,000  $\times$  g), resuspended in cold 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub>, and washed three times with the same buffer to remove sorbitol. DNase I (5  $\mu$ g/ml) was added to the crude extract and cells were broken with a French press. After centrifugation, the supernatant containing soluble recombinant Rgl11Y was loaded onto 5 ml of Ni-CAMHC resin (Sigma-Aldrich Co.) equilibrated with 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub>.

The resin was first washed with 30 mM Tris-HCl (pH 8.0) containing 1 mM CaCl<sub>2</sub> and recombinant Rgl11Y was eluted with 20 mM imidazole-30 mM Tris-HCl (pH 8.0). The eluate was dialyzed against 30 mM Tris-HCl (pH 8.0)-1 mM CaCl<sub>2</sub> and concentrated as previously described. The pure Rgl11Y obtained by this procedure was used for the kinetic assays. The Rgl11Y concentration was estimated spectrophotometrically at 280 nm in 6 M guanidine hydrochloride, based on the known amino acid composition of recombinant Rgl11Y.

**SDS-PAGE.** Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with precast 12.5% polyacrylamide gels using a Phast-System apparatus (Amersham Pharmacia Biotech) or polyacrylamide gels (8 or 10%) (ProSieve 50 Gel solution for protein analysis; TEBU) using a Vertical Electrophoresis System apparatus (Pharmacia Biotech).

**Production of antibodies raised against the recombinant Rgl11Y.** Polyclonal antibodies against Rgl11Y were raised in rabbits by subcutaneous injection of the protein purified from inclusion bodies. Antisera were stored at -20°C with 0.3% Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>. Antisera raised against Rgl11Y was preadsorbed with an *E. coli* BL21 (DE3)[pET22b(+)] extract.

**Cellulosome production, purification and immunodetection of Rgl11Y.** Cellulosomes were purified from a cellulose-grown culture supernatant of *C. cellulolyticum* and used to perform Western blot experiments as previously described (15).

**Rgl11Y-dockerin/cohesin interaction.** Interaction of Rgl11Y with cohesin was examined by using the biotinylated miniCipC1 probe as previously described (15, 32).

**Enzymatic assays.** Rhamnogalacturonan lyase activity was assayed spectrophotometrically by measuring, at regular time intervals, the formation of  $\Delta$ 4,5 unsaturated galacturonic acid at 235 nm. The activity was calculated using the known molar extinction coefficient of 5,200 M<sup>-1</sup> cm<sup>-1</sup> (25). The substrates used were potato pectic galactan (PPG) and rhamnogalacturonan from Megazyme, and polygalacturonic acid and pectin from citrus from Sigma-Aldrich.

The specific activity of Rgl11Y was determined as follows. Five milliliters of substrate (0.5%, wt/vol) in 50 mM Tris-HCl (pH 8.5) containing 1 mM CaCl<sub>2</sub> were incubated at 37°C with 5  $\mu$ g of recombinant Rgl11Y. At regular time intervals aliquots were removed and the enzymatic reaction was stopped by adding 3 volumes of 50 mM HCl. Rgl11Y activity was expressed in international units (IU). One unit of enzyme activity correspond to the release of 1  $\mu$ mol of product per min. For the determination of the rhamnogalacturonan lyase activity of the cellulosome, 60  $\mu$ g of cellulosomes was used.

For the determination of the optimum pH the following buffers were used: 50 mM morpholinepropanesulfonic acid (MOPS), pH 6.0 and 6.5; 50 mM Tris-HCl, pH 7.0, 7.5, 8.0, 8.5, and 9.0; and 50 mM CAPS [3-(cyclohexylamino)-2-hydroxypropane-1-sulfonic acid], pH 9.5, 10.0, and 10.5. All buffers contained 1 mM CaCl<sub>2</sub> and the substrate concentration was of 0.5% (wt/vol).

$K_m$  and  $k_{cat}$  values were determined in standard assay conditions used for the determination of the specific activity. Substrate concentrations ranged from 0.02 to 1% (wt/vol).

The influence of metal ions on the activity of Rgl11Y was determined after a 30 min pretreatment of the enzyme with 5 mM EDTA, at 4°C. The activity then was measured as previously described, using 50 mM Tris-HCl (pH 8.5) buffer supplemented with 5, 10, and 25 mM concentrations of CaCl<sub>2</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, LiCl<sub>2</sub>, and NiCl<sub>2</sub>.

## RESULTS

**Nucleotide sequence and deduced amino acid sequence of *rgl11Y*.** A new gene, *rgl11Y* (formerly named *orfY*), was identified downstream of the *cel9M* gene by chromosome walking in the large *C. cellulolyticum* gene cluster. *rgl11Y* consists of 2,028-nucleotides encoding a 676-amino-acid protein. A putative ri-

1 . VRKKIFMLLLALLQVALLTIPQQLVTAASARQMEYLDRGIVAVKVSNGVVFVSWRMLGTDSTSI AFNLYRDGKK  
 2 . ARR - EYLDRGVVALPSGSGIYIGWRMLGDDPANIGFHVYRNGTR

VNTPITSGTNYVDPTGTASSTYYVCPVKDGVELAKSESTAVWGQNYLSLP - - - - LQIPAGGTTKSGEYYTYS  
 ITSSPITNSTNYFDASGSSSAAYTVRPVNVNGVEQAANPAKATW - - - - SNPYWLVLNLRNPAGGTTPSGEAYSYP

NDCSVGDLGDGQYEIILKWDPSNAKDNSQSGYTGNYVLDAYKLNKFLWRIDLGRNIRAGAHYTQFMVYDLGD  
 NDLSVGDLGDGQYDIIVKWEPSNAKDNSQSGYTGNYVLDAYRLNLSQLWRIDLGRNIRAGAHYTQFIVYDLGD

GKAEVACKTADGTDGIGKVI GNANADYT - - NSNGYILSGPEYLTFNGETGAALT TTDYDPPRGTVSSWGDSYG  
 GKAEVAMKTAPGTRDAAGNVLGGSNANYRLRNASGYVLSGAEYLTFNGQTVAMATTDYVPARGTVSSWGDNYG

NRVDRFLACIAYLDGVHPSLVMCRGYYTRAVLAAYNWRDGKLT KVWTFDSKDSGNSAYEGQGNHNSVGDVSDG  
 NRVDRFLAGVAYLDGQRPSLVMARGYYTRAVVVAWDWRNGLSRRWTFDSNSGGNSAAAGQGAHSLTIGDVDQDG

MDEIVYGACAI DHNGRGLYTTKLNHG DAMHLSMDPGRPGLEVWQCHETGSNASSGEY - - - - - RDARTGQLI W  
 RDEIVYGAATINDNGTLLYSTGLGHGDALH LGDFNPNRPGLEVFMVHE - - - SPSS - - YGQHGIEMHDARTGAIW

GLAGTSDTGRGLALDIDPRYLGFEMWSSSDGVYNVNGTKVSTTKP - SMNFGVWWDGDLNRELLDGT KLDKWDYV  
 SVQGGD IGRGVTMDVDP RYPGNESWASTG - GLYSATGQQITTSKPASINFGVWWDADLLREQLNNTMIDKWNYS

NSTPMRLLTP - - - ADCASNSTKANPCL SADI LGDWREEVIWR TTDNKYLRIYTTAVTSNRIY TLMHDPLYRL  
 NNTSSRLLTAYNYGA - - ASNNGTKATPGLSADLFGDWREEVIWR HDNNSQLLI FSTTAVTTHKLR TLMHDTQYRT

**SIAWQNVAYNQPPHTGFYLG DGMSQAPT PNIYIAKPASSFLPGDVNNDGSVDALDFAIVK KYLLGQSTELDTGKA**  
**AIAWQNVAYNQPPHPSFFLGD**

DMNSDGDINALDLALLKASLLS

FIG. 1. Amino acid sequence of *C. cellulolyticum* Rgl11Y (sequence 1) and alignment with the catalytic domain of *P. cellulosa* Rgl11A (sequence 2). Identical amino acids are shown in boldface type. Rgl11Y putative signal sequence is shown in italics, and the putative dockerin domain is underlined.

bosome-binding site, AGGAGG, is present 7 bp upstream of the translational start codon GTG. This translational start codon is located 106 bp downstream of the *cel9M* gene stop codon.

The analysis of the deduced amino acid sequence of *rgl11Y*, revealed a multidomain protein with an N-terminal signal sequence, containing a putative cleavage site located between alanines 27 and 28. A 594-amino-acid putative catalytic domain and a 55-amino-acid putative dockerin domain characteristic of *C. cellulolyticum* cellulases follow this signal peptide (Fig. 1). The putative catalytic domain of Rgl11Y presents extensive sequence similarity with family 11 polysaccharide lyases of the carbohydrate-active enzyme database (<http://afmb.cnrs-mrs.fr/CAZY>) (Table 1). Among the polysaccharide lyase family 11 members, only *Pseudomonas cellulosa* Rgl11A has been charac-

TABLE 1. Sequence identity between Rgl11Y from *C. cellulolyticum* and family 11 polysaccharide lyase proteins

Family 11 PL protein <sup>a</sup>	Identity (%) with Rgl11Y from <i>C. cellulolyticum</i>
Rgl11A from <i>Pseudomonas cellulosa</i> .....	61
YesW from <i>Bacillus subtilis</i> .....	64
YesX from <i>Bacillus subtilis</i> .....	60
PSP from <i>Streptomyces coelicolor</i> .....	59
ORF BH1116 from <i>Bacillus halodurans</i> .....	58
ORF BH1114 from <i>Bacillus halodurans</i> .....	54
ORF6 from <i>Caldicellulosiruptor</i> sp. ....	65
Unknown ORF from <i>Polyangium cellulosum</i> .....	51

<sup>a</sup> Abbreviations: PL, polysaccharide lyase; PSP, putative secreted protein; ORF, open reading frame.

terized (25). This protein is a multidomain rhamnogalacturonan lyase containing an N-terminal signal sequence, a 596-amino-acid catalytic domain, a 94-amino-acid domain exhibiting extensive identity with fibronectin type 3 modules found in a range of glycoside hydrolases and pectate lyases, and a C-terminal 90 residue CBM from family CBM2. Rgl11A is an endo-acting rhamnogalacturonan lyase that cleaves  $\alpha$ -L-Rhap-(1 $\rightarrow$ 4)- $\alpha$ -D-GalpA bonds in the backbone of deesterified RGI.

**Production and purification of Rgl11Y.** The DNA sequence encoding the full-length mature protein was cloned into the *E. coli* high-level expression vector pET-22b(+) that uses the bacteriophage T7 promoter to direct the expression of the target gene in the cytoplasm. The resulting plasmid, pETYhp3, was used to transform the expression strain.

When expressed in *E. coli*, foreign proteins often accumulate as insoluble aggregates of improperly folded proteins called inclusion bodies. Under standard growth and standard induction conditions (Novagen protocol), *E. coli* transformed with pETYhp3 produces recombinant Rgl11Y in inactive and insoluble aggregates. Isolated inclusion bodies were solubilized by 8 M urea, purified and refolded in an appropriate buffer (see Materials and Methods). The pure renatured Rgl11Y, which was found to be inactive on all tested substrates (PPG and rhamnogalacturonan), was used to prepare rabbit antibodies.

Various conditions have been tested in order to obtain soluble proteins, including induction between 15 and 20°C, the use of lower concentrations of IPTG (10 to 100  $\mu$ M), and/or the use of a specific culture medium (5). Soluble and active recombinant Rgl11Y was produced only by growing and inducing the cells under osmotic stress in the presence of sorbitol (1 M) and betaine (2.5 mM) (5, 20, 21). Approximately 50% of recombinant Rgl11Y was produced in the soluble fraction of *E. coli* cell extracts by using 100  $\mu$ M IPTG and an induction temperature of 18°C. Rgl11Y was purified by a one-step procedure on a nickel-nitrilotriacetic acid column. The recombinant protein has an apparent molecular mass of 75 kDa on SDS-PAGE (Fig. 2), in good agreement with its theoretical mass (72,054 Da).

**Catalytic properties of Rgl11Y.** To investigate the catalytic properties of the recombinant enzyme, the activity of Rgl11Y was tested on four different substrates (polygalacturonic acid, pectin from citrus, PPG, and rhamnogalacturonan) by measuring the increase of absorbance at 235 nm. In a standard assay, the specific activity of Rgl11Y was found to be of 28 and 3.6 IU/mg, respectively, using PPG and rhamnogalacturonan as substrates. No activity against polygalacturonic acid and pectin from citrus could be detected, indicating that Rgl11Y is not an HG-degrading enzyme. Rgl11Y cleaves specifically the backbone of RGI in pectin and thus is a rhamnogalacturonan lyase. It is interesting to compare the composition of the two substrates (rhamnogalacturonan and PPG) in order to understand the substrate preference of Rgl11Y. Rhamnogalacturonan is prepared from soybean fibers by exhaustive hydrolysis with a series of pectinolytic and cellulolytic enzymes, resulting in an RGI backbone of alternating rhamnose and galacturonic acid residues. Other sugars, such as galactose (2.5%) and arabinose (12%), are present and constitute the RGI side chains. In contrast, PPG is prepared by alkaline extraction of potato fibers, followed by a starch-removing enzymatic treatment. PPG consists of short RGI backbone chains, highly substituted

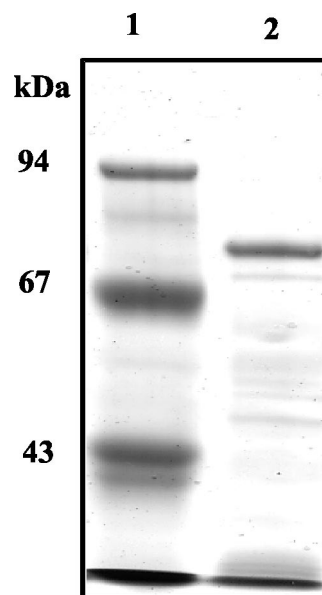


FIG. 2. SDS-PAGE analysis of purified recombinant Rgl11Y. Recombinant Rgl11Y was recovered from the soluble fraction of *E. coli* cell extracts. After a one-step purification process on nickel-nitrilotriacetic acid column, the Rgl11Y-containing protein fraction was submitted to SDS-PAGE (8% acrylamide) and stained with Coomassie brilliant blue. Lane 1, molecular mass markers; lane 2, recombinant form of Rgl11Y purified by Ni<sup>2+</sup>-affinity chromatography.

by galactose (82%) (Megazyme technical data). RGI is a poorly substituted substrate compared to PPG. Our results show that Rgl11Y preferentially degrades PPG such as Rgl11A from *P. cellulosa* (25) indicating that galactan side chains are probably required for Rgl11Y activity. High-performance liquid chromatography analysis of the products released by Rgl11Y and measurement of enzyme activity against purified oligosaccharides will be performed, in the future, to characterize the mode of action of Rgl11Y.

The enzymes that depolymerize the backbone of rhamnogalacturonan utilize two distinct cleavage mechanisms, either hydrolytic, or  $\beta$ -eliminative. The method used to evaluate Rgl11Y activity provides experimental evidence that Rgl11Y is a rhamnogalacturonan lyase (25). Indeed, lyases catalyze the  $\beta$ -eliminative cleavage of glycosidic bonds with the production of  $\Delta$ 4,5 unsaturated sugar, which can be followed spectrophotometrically at 235 nm.

The kinetic parameters of the recombinant enzyme,  $k_{cat}$  and  $K_m$  were found to be 34 s<sup>-1</sup> and 1 g/liter, respectively, on PPG. The pH optimum for Rgl11Y activity using PPG as a substrate was found to be 8.5. At the optimum pH for most of the *C. cellulolyticum* cellulases (pH 6 to 7), Rgl11Y displayed only 4 to 13% of its maximal activity. At pH 10, Rgl11Y was completely inactive.

To address the influence of metal ions on the catalytic activity of Rgl11Y, the enzyme was subjected to an EDTA treatment, which completely abolished activity against PPG. Incubation of the inactivated enzyme with zinc, magnesium, manganese, lithium or nickel did not restore activity. Only the addition of 10 mM of calcium was able to restore 80% of the Rgl11Y activity obtained in standard assay conditions.



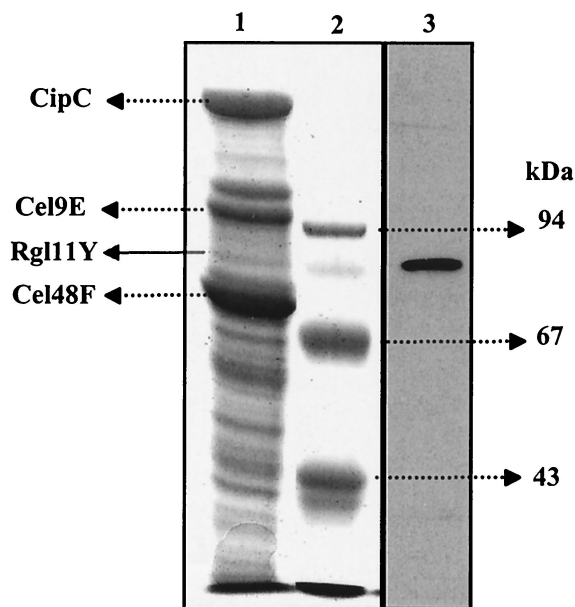


FIG. 3. Identification of Rgl11Y as a component of the *C. cellulolyticum* cellulosome. The cellulosomal components purified from a cellulose-grown were separated by SDS-PAGE (10% acrylamide), transferred to a nitrocellulose membrane, and incubated with antibodies raised against Rgl11Y. Lane 1, SDS-PAGE (10% acrylamide) of the *C. cellulolyticum* cellulosome proteins (the proteins corresponding to the three major polypeptides of the cellulosomes [scaffolding protein CipC and cellulases Cel9E and Cel48F] and Rgl11Y are indicated by arrows on the left); lane 2, molecular mass markers; lane 3, Western blot analysis of the *C. cellulolyticum* cellulosomes with antibodies raised against Rgl11Y.

**Rgl11Y, an enzymatic subunit of the cellulosomes.** The presence of Rgl11Y in the cellulosomal fraction purified from a *C. cellulolyticum* cellulose culture was monitored by a Western blotting experiment, as previously described (15) using antibodies raised against Rgl11Y. Cellulosome components were separated into subunits by SDS-PAGE (Fig. 3, lane 1), transferred on a nitrocellulose membrane and incubated with anti-Rgl11Y antibodies. The results obtained (Fig. 3, lane 3) suggest that Rgl11Y is an enzymatic subunit of the cellulosomes with an estimated molecular mass of 84 kDa. The difference between the estimated molecular mass of the cellulosomal and recombinant form of Rgl11Y, 84 and 75 kDa, respectively, is probably attributable to glycosylation (15, 31).

**Cohesin-dockerin interaction.** The existence of a C-terminal dockerin domain in Rgl11Y suggests that incorporation of the enzyme in the cellulosome is due to a cohesin-dockerin interaction, as previously demonstrated for most of the *C. cellulolyticum* cellulosomal cellulases (13, 15, 26, 30, 31, 32, 33).

The possible involvement of a cohesin-dockerin interaction in the association of Rgl11Y with the scaffolding protein CipC was studied as follows. Recombinant Rgl11Y was subjected to SDS-PAGE, blotted onto nitrocellulose and then incubated with a biotinylated-miniCipC1, a recombinant form of the scaffolding protein CipC containing the first cohesin domain (15, 32). Data obtained (Fig. 4) provide experimental evidence that the incorporation of Rgl11Y into the *C. cellulolyticum* cellulosome is mediated by a typical cohesin-dockerin association.

**A rhamnogalacturonan lyase activity in the cellulosome.** The cellulosomes of *C. cellulolyticum* exhibit rhamnogalacturonan lyase activity. The specific activity determined against PPG (0.5%) in 50 mM Tris-HCl-CaCl<sub>2</sub> 1 mM buffer (pH 8.5) was found to be 0.25 IU/mg.

Based on the specific activity of Rgl11Y and assuming that Rgl11Y is the only cellulosomal rhamnogalacturonan lyase, one can estimate that Rgl11Y represents less than 1% of the total cellulosomal proteins. This result is in agreement with the cellulosomal Coomassie-blue-stained Rgl11Y (Fig. 3, lane 1), which appears as a minor cellulosomal protein in the culture conditions used here to grow *C. cellulolyticum*.

**DISCUSSION**

The results presented here provide experimental evidence of the presence of a rhamnogalacturonan lyase in the cellulosomes of *C. cellulolyticum*.

To date, only two lyases have been found in cellulosomes, namely, Rgl11Y described here and Pel4A from *C. cellulovorans*, a multidomain pectate lyase belonging to family 4 of the polysaccharide lyases (37). Although these two enzymes share the same  $\beta$ -elimination mechanism, they degrade two distinct regions in the pectin molecule. Whereas Rgl11Y digests the RGI backbone, Pel4A cleaves HG chains in pectate (a low methylesterified form of pectin). Rgl11Y from *C. cellulolyticum* therefore provides the first example of a cellulosomal rhamnogalacturonase. Its incorporation into cellulosomes is mediated by a typical cohesin-dockerin interaction.

Cellulosomal enzymes of *C. thermocellum* and *C. cellulovorans* are able to catalyze the degradation of a wide range of substrates including cellulose, xylan, mannan (in *C. thermocellum* and *C. cellulovorans*), chitin (*C. thermocellum*), and pectin (*C. cellulovorans*) (6, 11, 17, 18, 19, 22, 23, 24, 39). In contrast, *C. cellulolyticum* cellulosomal enzymes characterized to date

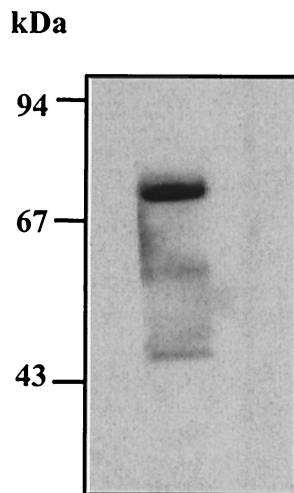


FIG. 4. Cohesin-dockerin interaction assay. Recombinant Rgl11Y was recovered from the soluble fraction of *E. coli* cell extracts, purified by Ni<sup>2+</sup>-affinity chromatography, loaded on SDS-10% acrylamide gels, transferred to a nitrocellulose membrane, and probed with a cohesin-containing protein (miniCipC1) labeled by biotinylation (32). Detection was performed using streptavidin-peroxidase conjugate. Molecular mass markers are indicated by arrows on the left.

are mainly cellulases, although xylanases (27) and a mannanase (A. Bélaïch, unpublished data) are also present. The present report indicates that *C. cellulolyticum* is also able to secrete a range of plant cell wall polysaccharide-degrading enzymes including pectinases, similar to *C. thermocellum* and *C. cellulovorans*.

The specific activity of Rgl11Y was found to be 28 IU/mg on PPG, whereas it was only 3.6 IU/mg on rhamnogalacturonan. Sequence similarity places Rgl11Y in family 11 of the polysaccharide lyases. In this sequence-based family, only Rgl11Y and Rgl11A from *P. cellulosa* have been extensively characterized (25). Rgl11Y and Rgl11A are comparable in terms of substrate specificity, catalytic properties ( $k_{\text{cat}}$ , 34 and 65 s<sup>-1</sup>, respectively;  $K_m$ , 1 and 8.5 mg/ml, respectively) (25), pH optimum (8.5 and 9.5, respectively), and Ca<sup>2+</sup> requirement. Galactan side chains seem to be required for the rhamnogalacturonan lyase activity. McKie et al. (25) have shown that Rgl11A from *P. cellulosa* cleaved an oligosaccharide comprising the backbone Rhap<sub>1</sub>-GalpA<sub>1</sub>-Rhap<sub>2</sub>-GalpA<sub>2</sub>-Rhap<sub>3</sub>-GalpA<sub>3</sub> in which Rhap<sub>1</sub> and Rhap<sub>2</sub> are substituted by galactose residues. Moreover it has been shown that removal of galactan side chains from RGI backbone decreased the catalytic efficiency of a rhamnogalacturonan lyase (family 4 of the polysaccharide lyases) from *Aspergillus aculeatus* (29). According our data Rgl11Y from *C. cellulolyticum* may have the same galactan requirement.

It is interesting to compare the modular organization of Rgl11Y and Rgl11A. In addition to its catalytic domain, Rgl11A contains a domain exhibiting significant sequence similarity to fibronectin type 3 modules, and a C-terminal CBM. The presence of CBMs in pectinases is rare in the microbial world. Rgl11A from *P. cellulosa* and the pectate lyase Pel10A from the same bacterium constitute two of the few examples of CBM-containing pectinases (7, 25). Rgl11A and Pel10A from *P. cellulosa* could interact with cellulose fibers via their CBMs, and would thus be in close proximity of their target substrates, namely, the main chain of the hairy and smooth regions in the pectin molecule (7, 25). Pel4A from *C. cellulovorans* also contains a CBM (37). Although the *C. cellulolyticum* Rgl11Y polypeptide does not contain a CBM, the enzyme is probably brought in contact to cellulose via the cellulosomal scaffolding protein, which contains a CBM. This would bring Rgl11A close to its preferential substrate, which is tightly associated to cellulose microfibrils in the primary plant cell wall. It is likely that this could render rhamnogalacturonan degradation easier.

The optimum pH of most of the *C. cellulolyticum* cellulases is around 6.5, whereas that of Rgl11Y is 8.5. The optimum pH value of Rgl11Y is similar to those of *P. cellulosa* Rgl11A (pH 9.5) and Pel10A (pH 10) and *C. cellulovorans* Pel4A (pH 8.0) (25, 37). It has been suggested that the pH value of some plant tissues changes during microbial attack. This could indicate that the degradation of plant cell wall polysaccharides occurs sequentially according to the pH of the plant tissues.

Cellulosomal gene clusters have been cloned and sequenced in *C. cellulolyticum* and *C. cellulovorans*. In *C. cellulovorans* the 22-kb cluster *cbpA-exg48S-eng9H-eng9K-hbpA-eng9L-man5A-eng9M-eng9N* is flanked at the 5' and 3' ends by genes that do not encode cellulosomal components. Unlike the gene encoding *C. cellulolyticum* Rgl11Y, the gene encoding *C. cellulovorans* Pel4A is not included in the cluster (37, 38). The *pel4A*

gene is located in a 3.7-kb *EcoRI-HindIII* DNA fragment containing genes *engY* and *helA*, which encode an endoglucanase and a DNA helicase, respectively.

To date 12 genes encoding *C. cellulolyticum* cellulosomal components have been found in the large cluster *cipC-cel48F-cel8C-cel9G-cel9E-orfX-cel9H-cel9J-man5K-cel9M-rgl11Y-cel5N*. The last identified gene, *cel5N*, located at the 3'-extremity of the cluster encodes a putative family 5 glycoside hydrolase containing a typical C-terminal dockerin domain (A. Bélaïch, unpublished data). In total, the cluster is 26 kb long, and so far is the largest *cel* cluster identified in cellulosome-producing clostridia. The region flanking the 3' extremity of the cluster has not been investigated in detail yet. The sequencing of approximately 2 kb (H. Maamar, personal communication) upstream of the *cipC* gene reveals that *cipC* is the first gene of the *cel* cluster.

#### ACKNOWLEDGMENTS

This work was funded in part by grants from the Centre National de la Recherche Scientifique, the Université de Provence, the Région Provence-Alpes-Côte d'Azur, and the Conseil Général des Bouches du Rhône.

We thank H. Maamar for providing purified cellulosomes from *C. cellulolyticum*. C. Giacalone and G. Sugier are thanked for their help during the cloning of the *rgl11Y* gene. We are indebted to P. de Philip, C. Tardif, H.-P. Fierobe, and B. Henrissat for critically reading the manuscript and D. Henrissat for correcting the English manuscript.

#### REFERENCES

- Bagnara-Tardif, C., C. Gaudin, A. Bélaïch, P. Hoest, T. Citard, and J. P. Bélaïch. 1992. Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. *Gene* **119**:17-28.
- Bayer, E. A., L. J. Shimon, Y. Shoham, and R. Lamed. 1998. Cellulosomes-structure and ultrastructure. *J. Struct. Biol.* **124**:221-234.
- Bélaïch, A., G. Parsiegla, L. Gal, C. Villard, R. Haser, and J. P. Bélaïch. 2002. Cel9M, a new family 9 cellulase of the *Clostridium cellulolyticum* cellulosome. *J. Bacteriol.* **184**:1378-1384.
- Bélaïch, J. P., C. Tardif, A. Bélaïch, and C. Gaudin. 1997. The cellulolytic system of *Clostridium cellulolyticum*. *J. Biotechnol.* **57**:3-14.
- Blackwell, J. R., and R. Horgan. 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Lett.* **295**:10-12.
- Blum, D. L., I. A. Kataeva, X. L. Li, and L. G. Ljungdahl. 2000. Feruloyl esterase activity of the *Clostridium thermocellum* cellulosome can be attributed to previously unknown domains of XynY and XynZ. *J. Bacteriol.* **182**:1346-1351.
- Brown, I. E., M. H. Mallen, S. J. Charnock, G. J. Davies, and G. W. Black. 2001. Pectate lyase 10A from *Pseudomonas cellulosa* is a modular enzyme containing a family 2a carbohydrate binding module. *Biochem. J.* **355**:155-165.
- Doi, R. H., and Y. Tamaru. 2001. The *Clostridium cellulovorans* cellulosome: an enzyme complex with plant cell wall degrading activity. *Chem. Rec.* **1**:24-32.
- Edashige, Y., and T. Ishii. 1997. Rhamnogalacturonan I from xylem differentiating zones of *Cryptomeria japonica*. *Carbohydr. Res.* **304**:357-365.
- Edashige, Y., and T. Ishii. 1998. Rhamnogalacturonan II from cell walls of *Cryptomeria japonica*. *Phytochemistry* **49**:681-690.
- Fernandes, A. C., C. M. Fontes, H. J. Gilbert, G. P. Hazlewood, T. H. Fernandes, and L. M. Ferreira. 1999. Homologous xylanases from *Clostridium thermocellum*: evidence for bi-functional activity, synergism between xylanase catalytic modules and the presence of xylan-binding domains in enzyme complexes. *Biochem. J.* **342**:105-110.
- Fierobe, H. P., C. Bagnara-Tardif, C. Gaudin, F. Guerlesquin, P. Sauve, A. Bélaïch, and J. P. Bélaïch. 1993. Purification and characterization of endoglucanase C from *Clostridium cellulolyticum*. Catalytic comparison with endoglucanase A. *Eur. J. Biochem.* **217**:557-565.
- Fierobe, H. P., S. Pagès, A. Bélaïch, S. Champ, D. Lexa, and J. P. Bélaïch. 1999. Cellulosome from *Clostridium cellulolyticum*: molecular study of the Dockerin/Cohesin interaction. *Biochemistry* **38**:12822-12832.
- Gal, L., C. Gaudin, A. Bélaïch, S. Pagès, C. Tardif, and J. P. Bélaïch. 1997. CelG from *Clostridium cellulolyticum*: a multidomain endoglucanase acting efficiently on crystalline cellulose. *J. Bacteriol.* **179**:6595-6601.
- Gal, L., S. Pagès, C. Gaudin, A. Bélaïch, C. Reverbel-Leroy, C. Tardif, and J. P. Bélaïch. 1997. Characterization of the cellulolytic complex (cellulo-

- some) produced by *Clostridium cellulolyticum*. Appl. Environ. Microbiol. **63**:903–909.
16. Gaudin, C., A. Bélaïch, S. Champ, and J. P. Bélaïch. 2000. CelE, a multidomain cellulase from *Clostridium cellulolyticum*: a key enzyme in the cellulosome? J. Bacteriol. **182**:1910–1915.
  17. Halstead, J. R., P. E. Vercoe, H. J. Gilbert, K. Davidson, and G. P. Hazlewood. 1999. A family 26 mannanase produced by *Clostridium thermocellum* as a component of the cellulosome contains a domain which is conserved in mannanases from anaerobic fungi. Microbiology **145**:3101–3108.
  18. Hayashi, H., K. I. Takagi, M. Fukumura, T. Kimura, S. Karita, K. Sakka, and K. Ohmiya. 1997. Sequence of *xynC* and properties of XynC, a major component of the *Clostridium thermocellum* cellulosome. J. Bacteriol. **179**:4246–4253.
  19. Hayashi, H., M. Takehara, T. Hattori, T. Kimura, S. Karita, K. Sakka, and K. Ohmiya. 1999. Nucleotide sequences of two contiguous and highly homologous xylanase genes *xynA* and *xynB* and characterization of XynA from *Clostridium thermocellum*. Appl. Microbiol. Biotechnol. **51**:348–357.
  20. Kempf, B., and E. Bremer. 1998. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. Arch. Microbiol. **170**:319–330.
  21. Kilstrup, M., S. Jacobsen, K. Hammer, and F. K. Vogensen. 1997. Induction of heat shock proteins DnaK, GroEL, and GroES by salt stress in *Lactococcus lactis*. Appl. Environ. Microbiol. **63**:1826–1837.
  22. Kim, H., K. H. Jung, and M. Y. Pack. 2000. Molecular characterization of *xynX*, a gene encoding a multidomain xylanase with a thermostabilizing domain from *Clostridium thermocellum*. Appl. Microbiol. Biotechnol. **54**:521–527.
  23. Kosugi, A., K. Murashima, and R. H. Doi. 2001. Characterization of xylanolytic enzymes in *Clostridium cellulovorans*: expression of xylanase activity dependent on growth substrates. J. Bacteriol. **183**:7037–7043.
  24. Kurokawa, J., E. Hemjinda, T. Arai, S. Karita, T. Kimura, K. Sakka, and K. Ohmiya. 2001. Sequence of the *Clostridium thermocellum* mannanase gene *man26B* and characterization of the translated product. Biosci. Biotechnol. Biochem. **65**:548–554.
  25. McKie, V. A., J. P. Vincken, A. G. Voragen, L. A. van den Broek, E. Stimson, and H. J. Gilbert. 2001. A new family of rhamnogalacturonan lyases contains an enzyme that binds to cellulose. Biochem. J. **355**:167–177.
  26. Mechaly, A., H. P. Fierobe, A. Bélaïch, J. P. Bélaïch, R. Lamed, Y. Shoham, and E. A. Bayer. 2001. Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition. J. Biol. Chem. **276**:9883–9888.
  27. Mohand-Oussaid, O., S. Payot, E. Guedon, E. Gelhaye, A. Youyou, and H. Petitdemange. 1999. The extracellular xylan degradative system in *Clostridium cellulolyticum* cultivated on xylan: evidence for cell-free cellulosome production. J. Bacteriol. **181**:4035–4040.
  28. Murashima, K., A. Kosugi, and R. H. Doi. 2002. Determination of subunit composition of *Clostridium cellulovorans* cellulosomes that degrade plant cell walls. Appl. Environ. Microbiol. **68**:1610–1615.
  29. Mutter, M., I. J. Colquhoun, G. Beldman, H. A. Schols, E. J. Bax, and A. G. J. Voragen. 1998. Characterization of recombinant rhamnogalacturonan  $\alpha$ -l-rhamnopyranosyl-(1, 4)- $\alpha$ -D-galactopyranosyluronide lyase from *Aspergillus aculeatus*. Plant Physiol. **117**:141–152.
  30. Pagès, S., A. Bélaïch, J. P. Bélaïch, E. Morag, R. Lamed, Y. Shoham, and E. A. Bayer. 1997. Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. Proteins **29**:517–527.
  31. Pagès, S., A. Bélaïch, H. P. Fierobe, C. Tardif, C. Gaudin, and J. P. Bélaïch. 1999. Sequence analysis of scaffolding protein CipC and ORFXp, a new cohesin-containing protein in *Clostridium cellulolyticum*: comparison of various cohesin domains and subcellular localization of ORFXp. J. Bacteriol. **181**:1801–1810.
  32. Pagès, S., A. Bélaïch, C. Tardif, C. Reverbel-Leroy, C. Gaudin, and J. P. Bélaïch. 1996. Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. J. Bacteriol. **178**:2279–2286.
  33. Park, J. S., Y. Matano, and R. H. Doi. 2001. Cohesin-dockerin interactions of cellulosomal subunits of *Clostridium cellulovorans*. J. Bacteriol. **183**:5431–5435.
  34. Reverbel-Leroy, C., S. Pagès, A. Bélaïch, J. P. Bélaïch, and C. Tardif. 1997. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. J. Bacteriol. **179**:46–52.
  35. Ridley, B. L., M. A. O'Neill, and D. Mohnen. 2001. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry **57**:929–967.
  36. Schwarz, W. H. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Appl. Microbiol. Biotechnol. **56**:634–649.
  37. Tamaru, Y., and R. H. Doi. 2001. Pectate lyase A, an enzymatic subunit of the *Clostridium cellulovorans* cellulosome. Proc. Natl. Acad. Sci. USA **98**:4125–4129.
  38. 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.
  39. Zverlov, V., K. P. Fuchs, and W. H. Schwarz. 2002. Chi18A, the endochitinase in the cellulosome of the thermophilic, cellulolytic bacterium *Clostridium thermocellum*. Appl. Environ. Microbiol. **68**:3176–3179.