J. Microbiol. Biotechnol. (2005), **15**(2), 302–309

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

© The Korean Society for Microbiology and Biotechnology

Cloning of *celC*, Third Cellulase Gene, from *Pectobacterium carotovorum* subsp. *carotovorum* LY34 and its Comparison to Those of *Pectobacterium* sp.

LIM, WOO JIN[1], SUNG KEE RYU[1], SANG RYEOL PARK[3], MIN KEUN KIM[1], CHANG LONG AN[1], SU YOUNG HONG[1], EUN CHULE SHIN[1], JONG YEOUL LEE[3], YONG PYO LIM[4], AND HAN DAE YUN[1,2]*

[1]*Division of Applied Life Science and* [2]*Research Institute of Life Science, Gyeongsang National University, Chinju 660-701, Korea*
[3]*National Institute of Agricultural Biotechnology, RDA, Suwon 441-707, Korea*
[4]*Department of Horticulture, Chungnam National University, Daejeon 305-764, Korea*

Received: May 11, 2004
Accepted: September 16, 2004

Abstract Phytopathogenic *Pectobacterium carotovorum* subsp. *carotovorum* (*Pcc*) LY34 secretes multiple isozymes of the plant cell wall degrading enzyme endoglucanases. We have cloned a third *cel* gene encoding CMCase from *Pcc* LY34. The structural organization of the *celC* gene (AY188753) consisted of an open reading frame (ORF) of 1,116 bp encoding 371 amino acid residues with a signal peptide of 22 amino acids within the NH_2-terminal region of pre-CelC. The predicted amino acid sequence of CelC was similar to that of *Pectobacterium chrysanthemi* Cel8Y (AF282321). The CelC has the conserved region of the glycoside hydrolase family 8. The apparent molecular mass of CelC was calculated to be 39 kDa by CMC-SDS-PAGE. The cellulase-minus mutant of *Pcc* LY34 was as virulent as the wild-type in pathogenicity tests on tubers of potato. The results suggest that the CelC of *Pcc* LY34 is a minor factor for the pathogenesis of soft-rot.

Key words: *Pectobacterium carotovorum* subsp. *carotovorum* LY34, *celC* gene, glycoside hydrolase family 8, soft-rot

Pectobacterium carotovorum subsp. *carotovorum* (*Pcc*), previously classified as *Erwinia carotovora* subsp. *carotovora* (*Ecc*), is one of the pathogenic entrobacteria causing soft-rot disease in plants. This bacterium is capable of degrading macromolecules that compose the structure of plant cell wall and middle lamellae, resulting in maceration of the plant tissue. This macerating capacity of the organism comes from the secretion of a set of extracellular enzymes that include both pectinolytic as well as cellulolytic enzymes. These enzymes occur in multiple enzymatic forms: The bacterial synthesis of isozymes may ensure more efficient degradation of polysaccharides present in the plant cell wall or provide more advantageous regulatory strategy to the bacteria [3].

The enzymes of *Pectobacterium chrysanthemi* (*Pch*) and *Pcc*, such as pectinases, proteases, and cellulases, exist in multiple enzymatic forms. As many as 8 *Ewinia chrysanthemi* (*Ech*) 3937 pectate lyases have been characterized; PelA, PelB, PelC, PelD, PelE, PelI, PelL, and PelZ [25, 29]. So far, most of the attention has been focused on studies of the pectic enzymes, and relatively little work has been done to characterize other secreted enzymes such as cellulase and hemicellulase. In *Pectobacterium* sp., 2 cellulases, referred to as CelZ and CelY, have been extensively characterized in *Pch* 3937 [2, 4, 10]. No similarity was found between the predicted amino acid sequences of CelZ (family 5) and CelY (family 8). CelS (family 12) of strain SCC3193 was the first cellulase gene from *Pcc* to be cloned and characterized [27]. A second group of *Pcc* cellulases, containing CelV (family 5) of strain SCRI193 [7] and CelVI (family 5) of strain SCC3193 [19], was investigated. CelA (family 5) and CelB (family 12) of *Pcc* LY34 strain [24] and Cel5Z (family 5) and Cel8Y (family 8) of *Pch* PY35 strain [6,22] have previously been isolated in our laboratory.

It is of interest to compare the structures of multigenic genes, especially the *cel* genes. The occurrence of multiple isozymes has complicated the biochemical and genetic analyses of individual isozymes. The very similar molecular weights of cellulases cause serious problems in purifying them to homogeneity. Therefore, the purity of those cellulases that have previously been prepared is highly unreliable, and additional sensitive cellulase measurement techniques are required. CMC-SDS-PAGE procedure was applied to the study of *Pectobacterium* sp. [23, 24]. Activity staining after CMC-SDS-PAGE appeared to be a

*Corresponding author
Phone: 82-55-751-5469; Fax: 82-55-757-0178;
E-mail: hdyun@nongae.gsnu.ac.kr

better procedure for the characterization of the activities of these enzymes. This technique entails enzyme activity and high resolution of proteins based on their size in the presence of SDS.

The *Pcc* LY34, used in the present study, was originally isolated from Chinese cabbage tissue with soft-rot symptoms. By the CMC-SDS-PAGE method, it was possible to detect at least five different cellulases in the wild-type *Pcc* LY34 strain [17], suggesting that a complex system of cellulolytic enzymes exists in this strain. In this report, we describe the cloning of a third β-1,4-D-glucanohydrolase (EC 3.2.1.4; endoglucanase) from *Pcc* LY34. The genetic map and the complete nucleotide sequence of the CMCase gene, named *celC*, are presented. The deduced amino acid sequence and the CMCase activity of the protein were compared to CelA and CelB from *Pcc* LY34 and other endoglucanases.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

E. coli XL1-blue cells and recombinant *E. coli* cells were cultured in LB containing appropriate antibiotics (ampicillin, 50 μg/ml; kanamycin, 50 μg/ml; tetracycline, 10 μg/ml).

Extracellular Cellulase Activity Assay and Purification

To screen cellulase activity in *E. coli* harboring the cloned cellulase gene, bacterial colonies were grown on a cellulase activity indicator medium [LB agar plates containing appropriate antibiotics and 15% (v/v) Cellomix (RNA, Suwon, Korea)]. After growth at 37°C for 24 h, positive clones for extracellular cellulase activity were found to be surrounded by a light blue halo against a blue background. For the purification of the enzyme, *E. coli* XL1-Blue harboring pRY200 was grown at 37°C. After centrifugation of the culture at 10,000 ×*g* for 10 min at 4°C, the supernatant was precipitated with 70% ammonium sulfate. After the precipitant was dialyzed against 10 mM Tris-HCl (pH 7.5), the preparation was subjected to Q-sepharose fractionation in the presence of urea. The samples showing good specific activity were dialyzed against 10 mM Tris-HCl (pH 7.5) and loaded onto a phenyl-sepharose column. One unit of the enzyme activity was defined as the amount of the enzyme to liberate 1 μmol of reducing sugar per min at 40°C. β-Glucosidase activity was assayed by measuring the production of *p*NP from *p*-nitrophenyl-D-glucopyranoside (*p*NPG) as previously described [16, 18]. Avicel (Merck), used as the indicator substrate for any hydrolyzing activity of crystalline cellulose, was incubated with 4 mU of the enzyme in 1.5 ml of buffer for 60 min. The extracellular cellulase activity in the recombinant DNA clones was measured by using dinitrosalicylic acid, as previously described [13-15, 20]. The concentration of CMC (low viscosity, Sigma) used as substrate was varied between 1 to 10 mg per ml. Protein concentrations were determined by the method of Bradford [5].

Recombinant DNA Techniques

Standard procedures for agarose gel electrophoresis, purification of DNA from agarose gels, DNA ligation, restriction endonuclease digestion, and other cloning-related techniques were as described by Sambrook and Russel [28]. Nucleotide sequencing was done by the dideoxy-chain termination method, using the PRISM Ready Reaction Dye terminator/primer cycle sequencing kit (Perkin-Elmer Corp., Norwalk, CN, U.S.A.). Nucleotide sequence data

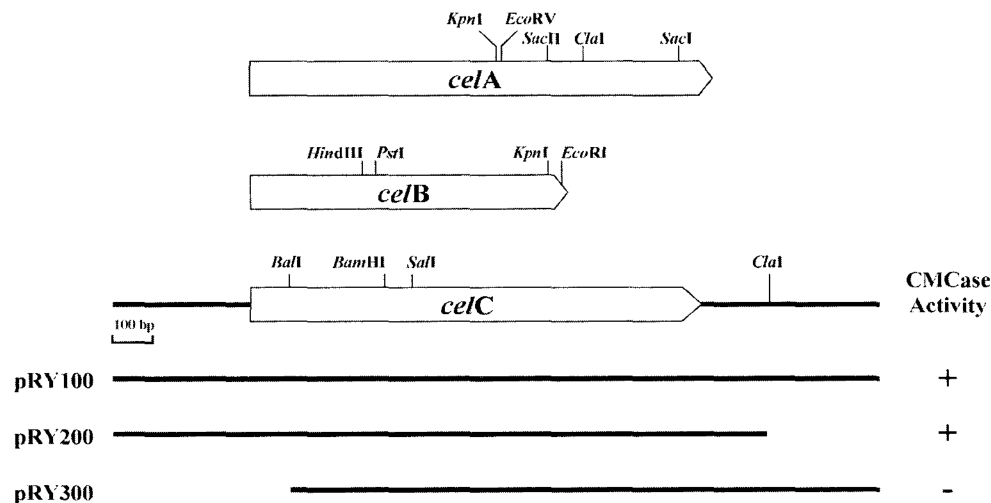

Fig. 1. Restriction map of the *celA*, *celB*, and *celC* (pRY100) genes from *Pectobacterium carotovorum* subsp. *carotovorum* LY34.
The cleavage sites of restriction enzymes *BalI*, *BamHI*, *SalI*, *ClaI*, *BstXI*, and *SacI* are shown in pRY100. pRY100 was constructed by cloning a 1.9-kb *Sau3AI* fragment of *Pcc* LY34 DNA into the *BamHI* site of pBluescript II SK+ vector. pRY200 (*celC*) was derived by subcloning of the 1.7-kb fragment into the corresponding sites of pBluescript II SK+. (+) and (-) represent the result of the test for CMC-degrading activity.

reported are available in the GenBank database under the accession number AY188753.

Site-Directed Mutagenesis

The *cel*C gene (1.9-kb), encoding the endo-1,4-β-D-glucanase from *Pcc* LY34, was used for *in vitro* site-directed mutations. Site-directed *in vitro* mutations were used for E57A with 27-mer synthetic oligonucleotide primers; 5'-ATCACCACGTCCGCAGGGCAAAGCTAC-3' (forward) and 5'-GTAGCTTTGCCCTGCGGACGT-GGTGAT-3' (reverse), for D118A with 28-mer primers 5'-CAACTCCGCTTCCGCTGCCGACCTGTGG-3' (forward) and 5'-CCACAGGTCGGCAGCGGAAGCGGAGTT-3' (reverse), and for D245A with 28-mer 5'-CATCAGCAG-CTACGCCGCCATCCGCGTC-3' (forward) and 5'-GAC-GCGGATGGCGGCGTAGCTGCTGATG-3' (reverse). Fifty μl of reaction mixture contained 1 μl of the pRY100 DNA (Fig. 1), 10 pmol each of primer, 1 μl of 2 mM dNTP mixture, and 5 μl of 10× *Pfu* DNA polymerase buffer containing 20 mM $MgSO_4$ and 2.5 U of cloned *Pfu* DNA polymerase [purchased from Stratagene (La Jolla, CA, U.S.A.)]. PCR products were incubated on ice for 5 min, and 1 μl of *Dpn*I restriction enzyme (10 U/μl) was then added, and the reaction mixture was gently and thoroughly mixed and incubated at 37°C for 1 h. Following *in vitro* mutation, the *Dpn*I-treated plasmids were transformed to *E. coli* XL1-blue according to the manufacturer's specifications on the site-directed mutagenesis kit (Stratagene, La Jolla, CA, U.S.A.).

Activity Staining on CMC-SDS-PAGE

CMC-SDS-PAGE (carboxymethylcellulose-sodium dodecyl sulfate-polyacrylamide gel electrophoresis) was performed by the method of Ryu *et al.* [26]. *E. coli* XL1-blue cells harboring the *cel*C gene were cultured at 37°C for 24 h in LB medium supplemented with 0.1% (w/v) CMC. Whole cell extracts from these cultures were prepared by sonication (three times for 30 sec each at 4°C). The cell extracts and supernatants were mixed at 1:1 (v/v) with sample buffer (62 mM Tris-HCl, pH 6.8, 10% glycerol, 0.025% bromophenol blue, 5% β-mercaptoethanol, and 2% SDS), and the mixtures were heated at 95°C for 3 min. The protein samples (30 μg) were electrophoresed.

RESULTS

Isolation and Restriction Map of the *cel*C Gene

The genomic DNA of *Pcc* LY34 was partially digested with *Sau*3AI. Two- to three-kb fragments were then ligated into the *Bam*HI site of pBluescript II SK+. Among 2,000 transformants, one positive colony surrounded by a light blue ring, that indicates CMCase activity, was isolated. A 1.9-kb fragment (pRY100) of *Pcc* LY34 genomic DNA contained a *cel* gene. The size of the inserted DNA and the orientation of restriction cleavage sites were determined. The inserted DNA of pRY100 contains the restriction sites for *Bal*I, *Bam*HI, *Sal*I, *Cla*I, *Bst*XI, and *Sac*I. By subsequent subcloning, a 1.7-kb fragment (pRY200) was defined and designated *cel*C, since this fragment was sufficient for cellulase activity. The restriction map of this clone was clearly different from those of the previously characterized *cel*A and *cel*B genes (Fig. 1). The activity of *E. coli* harboring the gene was lower than that of CelA of the *Pcc* LY34, as evidenced by the plate assay (Fig. 2).

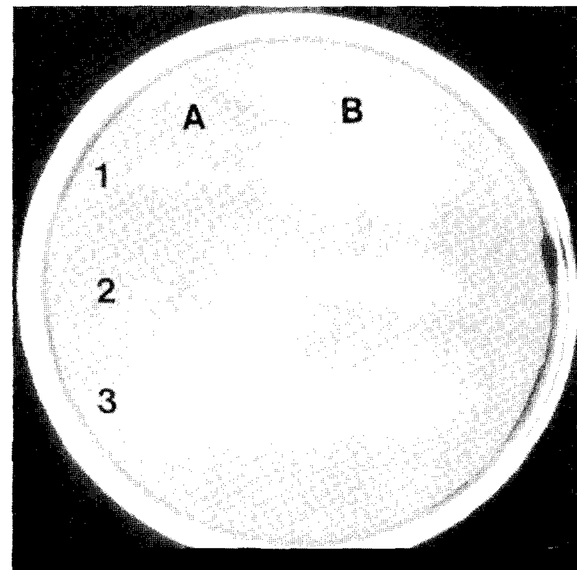

Fig. 2. Detection of CMCase-positive clones by the agar diffusion method.
The cells were incubated at 37°C for 1 day. 1A, *E. coli* XL1-blue harboring pBluescript II SK+ for negative control; 2A, pRY300; 3A, *E. coli* XL1-blue harboring *cel5Z* of *Pch* PY35 for the positive control [22]; 1B, *E. coli* XL1-blue harboring *cel*A of *Pcc* LY34 for the positive control [24]; 2B, *E. coli* XL1-blue harboring *cel*B of *Pcc* LY34 for the positive control [23]; 3B, *E. coli* XL1-blue harboring *celC* of *Pcc* LY34.

Nucleotide Sequence of the *cel*C Gene

The 1.9-kb inserted fragment in pRY100 was sequenced the using the dideoxy chain-termination method. It contained one complete open reading frame (ORF). Figure 3 depicts the *cel*C structural gene with its flanking regions. The ORF contains 1,116 nucleotides encoding a protein of 371 amino acid residues with a predicted molecular mass of 41,580 Da. The ATG initiation codon at nucleotide position 347 is preceded by a putative Shine-Dalgarno sequence, GGGA. The ORF ends with the TAA stop codon at position 1,459. The first 22 amino acids of CelC separated in front of Ala23 have typical features of the prokaryotic signal peptide within the NH_2-terminal region of pre-CelC (Fig. 3, http://www.cbs.dtu.dk/services/SignalP/). This sequence is likely to function in *E. coli* XL1-blue in the

```
   1  TCTCCGTATCACGACCAGCGTAGCGTTGTCGCCCTGCTGGCCGATAGCCCACAGGGCTATACGCTGCTCAATAAC
  76  GCGCTGATCGACAGCGAAAAAAGAGCCTCGCTGTTCGGTTCCGTTTCCGTCATCCGCGAATCGGGTATTAATAAT
 151  CTCCGCGTTGGAGACATTTATTACGTCGGCCATCTGCCGTGGTGGGAACGCATCTGGCACGCATTGGCGCAGCAT
 226  CCCGTCTGGCTAGCCATCATCTCGACGCTTACCGTCGTTATTGTTGCCTGGCTGCTGTGGCGTGGCCTGAAATTC
                                                     -35                      -10
 301  TTCAGTCGCCGTCGCCTGTCGCCGGATGAAAGGGACTAACGCACCATGCCACGCGTGCTGCACTACCTGATCCCC
   1                                      SD           M  P  R  V  L  H  Y  L  I  P

 376  ACGTTGCTGTGGCTATGGGCTTCCCAGGCCACCGCCGCCGTCTGCGACTGGCCAGCCTGGGAGCAGTACAAACAG
  11   T  L  L  W  L  W  A  S  Q  A  T  A  A  V  C  D  W  P  A  W  E  Q  Y  K  Q

 451  CATTACATCAGTGCGGAAGGACGGGTGATTGATACCTCTACCCCCAACAAAATCACCACGTCCGAAGGGCAAAGC
  36   H  Y  I  S  A  E  G  R  V  I  D  T  S  T  P  N  K  I  T  T  S  E  G  Q  S
                                                                        ▲
 526  TACGCCATGTTCTTTGCTCTGGTCGCCAACGATCGCGTGATGTTTGATCGGCTGCTGCAATGGACGGAAAACAAC
  61   Y  A  M  F  F  A  L  V  A  N  D  R  V  M  F  D  R  L  L  Q  W  T  E  N  N

 601  CTGTCCGCAGGCGATTTACGTGCGCATCTGCCCGCCTGGCTGTGGGGAGAAAACAAAGATAAGCAGTGGACGGTG
  86   L  S  A  G  D  L  R  A  H  L  P  A  W  L  W  G  E  N  K  D  K  Q  W  T  V

 676  CTGGATCCCAACTCCGCTTCCGATGCCGACCTGTGGATCGCCTACAGCCTGCTTGAGGCTGGTCGACTGTGGAAA
 111   L  D  P  N  S [A  S  D  A  D  L  W  I  A  Y  S  L  L  E  A  G  R  L  W] K
                             ▲
 751  GAGGCGCGCTATCAGACAGTTGGCACCGCGTTGCTCGCCCGTATCGCCAAAGAAGAGGTCGTCAATATTCCAGGA
 136   E  A  R  Y  Q  T  V  G  T  A  L  L  A  R  I  A  K  E  E  V  V  N  I  P  G

 826  CTGGGCGTGATGTTGCTGCCCGGCAAAGTAGGCTTCGCAGAGAAAGAGAGCTGGCGATTAAATCCCAGTTACCTT
 161   L  G  V  M  L  L  P  G  K  V  G  F  A  E  K  E  S  W  R  L  N  P  S  Y  L

 901  CCGCCACAGCTGCTGGCTCGCTTTGCGCCACTGAGCGAGACGTGGAAAGCGATGCAGCGCACCACGCAACGCCTG
 186   P  P  Q  L  L  A  R  F  A  P  L  S  E  T  W  K  A  M  Q  R  T  T  Q  R  L

 976  CTGTTGGAAACCGCGCCAAAAGGCTTCTCGCCTGATTGGGTCATCTGGCAAAAAGATAAGGGCTGGCAGCCTGAT
 211   L  L  E  T  A  P  K  G  F  S  P  D  W  V  I  W  Q  K  D  K  G  W  Q  P  D

1051  ACCACCAAACCCAACATCGGCAGCTACGACGCCATCCGCGTCTATCTGTGGGCAGGGATGATGGCCGACAGCAGC
 236   T  T  K  P  N  I  G  S  Y  D  A  I  R  V  Y  L  W  A  G  M  M  A  D  S  S
                                ▲
1126  AAAGGGAAAACCGATCTGATCAAACAGTTTCAGCCAATGGTTCAGCAGACGATAAAGCAAGGTCTGCCGCCCGAA
 261   K  G  K  T  D  L  I  K  Q  F  Q  P  M  V  Q  Q  T  I  K  Q  G  L  P  P  E

1201  AAAACCGACACCGCGACAGGCACCGTCACTGGACAGGGATCGGTTGGGTTTTCCGCTTCGCTGCTTCCGATGCTT
 286   K  T  D  T  A  T  G  T  V  T  G  Q  G  S  V  G  F  S  A  S  L  L  P  M  L

1276  TCTCGTCAGTCGGATGCACTGGCTACCCAACGACAGCGGCTGGCCGACAACCCTCCGGGGGATGATGCGTATTTC
 311   S  R  Q  S  D  A  L  A  T  Q  R  Q  R  L  A  D  N  P  P  G  D  D  A  Y  F

1351  TCCGCCTCTCTGACGCTCTTTGGTCAGGGATGGGATCAGAAGCGCTATCGCTTCACTTCACAAGGCCAACTTTTA
 336   S  A  S  L  T  L  F  G  Q  G  W  D  Q  K  R  Y  R  F  T  S  Q  G  Q  L  L

1426  CCGTCTCGGGGCAGCCAATGCATAACAACACCGTAAACTGGCTGCGCTTCCTCCCGTTATTACTGGTTGCTGCGC
 361   P  S  R  G  S  Q  C  I  T  T  P  *

1501  CGCAGGCATACAGTGCGGAAATCGCATCGCCAGAGCAATTTTTGATGGAGCAGGTGCGTCTGGGAGAAGCCAGCA
1576  ACAAAGACGACCTCGTGCGCCAATCGCTTTATCGACTGGAGCTTATCGATCCAAACAACCCCGAGGTTATCGCCG
1651  CCCGATTGCGGCTGGCGCTGCGTCAAGGCGATCAGGCCCAGGCTCGTCAGCAACTGGACAAGCTGAAGGCGGTGG
1726  CTCCCGACTCTGCCACCTACCATCAGTCAGCCACCACGCTGGCACTGACGCAGGATGCCGCTCGCCAGCAATTAC
1801  AGCAGGCGCGCTTGCTGTCTACCGCCGGGCGCTATGCTGAAGCCAAAGTGCAGTACGATGCGCTCTTTCACGGCG
```

Fig. 3. Nucleotide and deduced amino acid sequence of the *cel*C gene of *Pcc* LY34 with its flanking region.

The putative Shine-Dalgarno (SD) box is indicated. The - 10 region and - 35 region are underlined. The underline indicates signal peptide. The conserved region of the glycoside hydrolase family 8 is boxed. Arrowheads indicate the residues exchanged by site-directed mutagenesis.

export of CelC to the periplasm. We examined whether the *cel*C gene comes from the chromosome of *Pcc* LY34 by using PCR and DNA sequencing. An internal primer set from the ORF region of *cel*C, 5'-GCACCACGCAAC-GCCTGCTGTTGG-3' (forward) and 5'-GCACGAGGTC-GTCTTTGTTGCTGG-3' (reverse) was used to amplify the corresponding region from the genomic DNA of Pcc LY34 with PCR. The PCR product was electrophoresed on an agarose gel. A gel slice containing the corresponding 0.6-kb fragment was cut and extracted with a gel extraction kit (NucleoGen, Seoul, Korea). The extracted DNA was analyzed with an automated DNA sequencer. The result of DNA sequencing demonstrated that the *cel*C gene came from the original organism (AY188753).

Amino Acid Sequence Similarities Between CelC and Other Cellulases

The amino acid sequences of *Pcc* LY34 CelC were shown to have 27%, 61%, and 28% identity with those of *Pch* PY35 Cel8Y [6], *E. coli* BcsC [23], and *Cellulomonas uda* Cel

Table 1. Pairwise similarity between CMCase amino acid sequences[a] of *Pectobacterium carotovorum* subsp. *carotovorum* and *Pectobacterium chrysanthemi.*

CMCase amino acid sequences[b]	Similarity (%) with CMCase amino acid sequences:				
	1	2	3	4	5
1. CelC	100.0	8.4	5.0	17.5	30.9
2. CelA		100.0	20.3	34.4	7.4
3. CelB			100.0	11.7	5.0
4. Cel5Z				100.0	11.7
5. Cel8Y					100.0

[a]Calculated with CLUSTAL W and the PAM250 residue weight table.

[b]The sequences are from the following sources: CelC, from *Pectobacterium carotovorum* subsp. *carotovorum* LY34 (this study); CelA, *Pectobacterium carotovorum* subsp. *carotovorum* LY34 (AF025768); CelB, *Pectobacterium carotovorum* subsp. *carotovorum* LY34 (AF025769); Cel5Z, *Pectobacterium chrysanthemi* PY35 (AF208495); Cel8Y, *Pectobacterium chrysanthemi* PY35 (AF282321).

[21], respectively. Alignment of the amino acid sequences of CelA, CelB, and CelC of *Pcc* LY34 showed that these three cellulases are clearly distinct, although they came from the same strain. The deduced amino acid sequence alignments of CelA, CelB, and CelC of *Pcc* LY34 were very similar to Cel5Z of *Pch* PY35 (family 8), Cel8Y of *Pch* PY35 (family 5), and CelA of *Rhodothermus marinus* (family 12) [11], respectively. Consequently, the Pcc LY34 appears to harbor cellulases belonging to three different families (Table 1). Therefore 35 different gene sequences known in the cellulase family were retrieved from the SwissProt and GenBank databases via an NCBI Blast search, and a phylogenetic tree of cellulase proteins was constructed by the PC/GENE CLUSTAL method. As seen in Fig. 4, the phylogenetic tree showed that the CelC protein was very close to the BcsC of *E. coli* K12 (P37651), and that the topology of the evolutionary tree with regard to the CelC protein was quite similar to that of the glycoside hydrolase family 8. The amino acid similarity and phylogenetic analysis of the *cel*C gene strongly suggest that CelC can be assigned to family 8, as seen in Fig. 4.

Identification and Characterization of *cel*C Gene Product

For the identification of the *cel*C gene product, CMC-SDS-PAGE described by Ryu *et al.* [26] was performed. An active protein band had apparent molecular mass of about 39,000 Da. Since the predicted *cel*C gene product consisted of 349 amino acids with estimated molecular mass of 39,064 Da, the size of the protein identified on the zymogram corresponded well with the predicted size (Fig. 5 and Table 2), and those of CelA and CelB of *Pcc* LY34 previously reported for positive control [23, 24]. By the CMC-SDS-PAGE method, we earlier observed the presence of five cellulases in the wild-type *Pcc* LY34 strain [17],

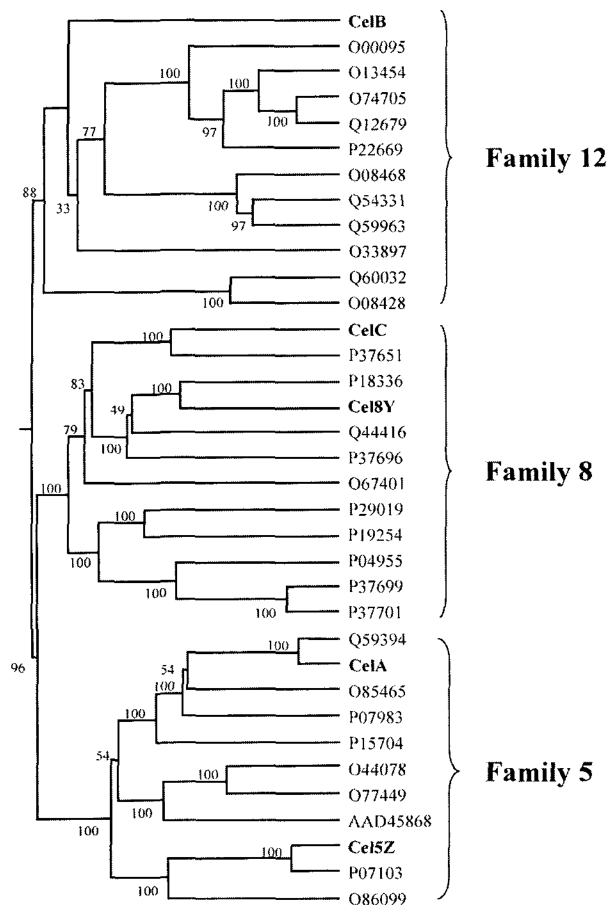

Fig. 4. Phylogenetic tree, showing evolutionary relatedness and levels of homology among the CMCase amino acid sequences.
The estimated genetic distance between sequences is proportional to the lengths of the horizontal lines connecting one sequence to another. Bootstrap values of the major branch points are shown; they represent the number of times the group consisting of the species to the right of that branch occurred among 100 trees. The sequences are of the following proteins: CelB, *Pectobacterium carotovorum* subsp. *carotovorum* LY34 (AF025769); O00095, *Trichoderma reesei*; O13454, *Aspergillus oryzae*; O74705, *Aspergillus niger*; Q12679, *Aspergillus kawachii*; P22669, *Aspergillus aculeatus*; O08468, *Streptomyces halstedii*; Q54331, *Streptomyces lividans*; Q59963, *Streptomyces rochei* A2; O33897, *Rhodothermus marinus*; Q60032, *Thermotoga maritima*; O08428, *Thermotoga neapolitana*; CelC, *Pectobacterium carotovorum* subsp. *carotovorum* LY34 (AY188753); P37651, *Escherichia coli* K12; P18336, *Cellulomonas uda*; Cel8Y, *Pectobacterium chrysanthemi* PY35 (AF232821); Q44416, *Agrobacterium tumefaciens*; P37696, *Acetobacter xylinus* ATCC 23769; O67401, *Aquifex aeolicus*; P29019, *Bacillus* sp. KSM-330; P19254, *Bacillus circulans* WL-12; P04955, *Clostridium thermocellum* NCIB 10682; P37699, *Clostridium cellulolyticum*; P37701, *Clostridium josui*; Q59394, *Erwinia carotovora atroseptica*; CelA, *Pectobacterium carotovorum* subsp. *carotovorum* LY34; O85465, *Bacillus agaradherans*; P07983, *Bacillus subtilis* DLG; P15704, *Clostridium acetobutylicum*; O44078, *Globodera rostochiensis*; O77449, *Heterodera glycines*; AAD45868, *Meloidogyne incognita*; Cel5Z, *Pectobacterium chrysanthemi* PY35 (AF208495); P07103, *Erwinia chrysanthemi* 3937; O86099, *Alteromonas haloplanktis*.

suggesting that a complex system of cellulolytic enzymes exists in *Pcc* LY34, where two cellulases have previously been demonstrated.

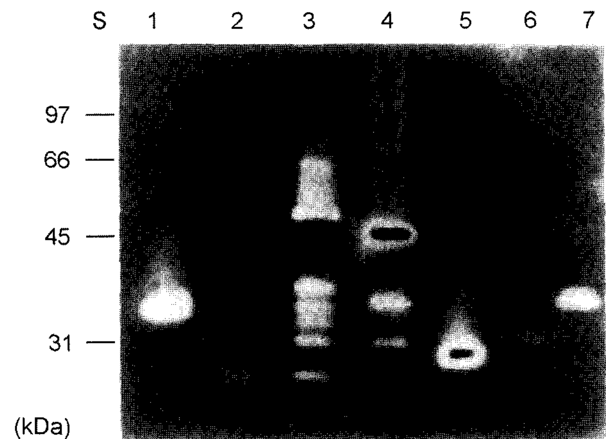

Fig. 5. Detection of CMCase activities of *Pectobacterium* sp., CelC, and other *cel* gene products of *Pectobacterium* sp. by CMC-SDS-PAGE.
The sonicated extract of cells for the intracellular cellulase and the supernatant of the culture for the extracellular cellulase after centrifugation were loaded on a CMC-SDS PA gel. After electrophoresis and protein reaction, cellulase activities were detected by staining with Congo red and HCl solution. Lane 1, *E. coli* XL1-blue harboring *cel*C of *Pcc* LY34; lane 2, *E. coli* XL1-blue harboring *cel*B of *Pcc* LY34; lane 3, the sonicated extract of cells for the intracellular cellulase plus the supernatant of the culture for the extracellular cellulase *Pcc* LY34; lane 4, cell extract of *Pch* PY35; lane 5, the supernatant of culture of *Pch* PY35; lane 6, *E. coli* XL1-blue harboring *cel*8Y of *Pch* PY35; lane 7, *E. coli* XL1-blue harboring *cel*5Z of *Pch* PY35; and lane S, the standard molecular size marker. Molecular weight markers used were phosphorylase b (97,400), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400).

The enzyme was purified by the method described in Materials and Methods, and the purified enzyme was analyzed by SDS-PAGE. The result showed a single protein band of about 39 kDa, in agreement with that deduced from the nucleotide sequence of the *cel*C gene in pRY200. The maximal activity was observed at pH 7.0 and 40°C. The activity of the purified enzyme toward *p*NPG and Avicel was only negligible. The utilization of α-cellulose fibers (C8002; Sigma) and micro-crystalline Avicel in M9 medium by *cel*C harboring *E. coli* cells was also tested by microscopically examining the cell growth. However, there was no evidence of an enzymatic activity on either carbohydrate after 2, 5, and 10 days of incubation. Although the precise role of CelC in cell wall degradation is not yet clearly understood, the enzyme has been classified as a CMCase type that does not appear to attack insoluble or crystalline forms of cellulose.

The predicted primary structure of the *Pcc* LY34 CelC protein shows regions of similarity to *Clostridium thermocellum* endoglucanase CelA [1, 9] and family 8 endoglucanases, which included the catalytic amino acids. In the CelA of *C. thermocellum*, the positions of the catalytic residues have been identified by high-resolution X-ray crystallography: The carboxylate group of Glu95 is hydrogen-bonded to the

Table 2. Comparison of molecular weights of the pre- and the processed forms of CelA, CelB, CelC, Cel5Z, and Cel8Y of *Pectobacterium carotovorum* subsp. *carotovorum* and *Pectobacterium chrysanthemi.*

	Molecular mass of protein form (Da)/number of amino acids[a]				Family member	Accession number	Ref.
	Precusor M.W.[b]	Signal peptide M.W.[c]	Calculated M.W.[d]	Apparent M.W.[e]			
CelA	42,003/387	3,495/31	38,525/336	39,000±500/-	5	AF025768	[24]
CelB	29,890/264	3,921/36	25,988/228	26,000±500/-	12	AF025769	[23]
CelC	48,210/430	2,652/23	45,576/407	46,000±500/-	8	AY188753	This work
Cel5Z	46,473/426	4,450/41	42,041/385	42,000±500/-	5	AF208495	[22]
Cel8Y	37,627/332	2,755/23	34,872/309	35,000±500/-	8	AF282321	[6]

[a]Molecular weight was calculated with the PC/GENE program.
[b]Protein of the primary gene product before signal modification.
[c]Hydrophobicity analysis of the signal peptide was determined with the PC/GENE program.
[d]Protein after post-translation modification of the primary gene product.
[e]Actual protein electrophoresed by using CMC-SDS-PAGE.

β-1,4-linking oxygen, and has been assigned the role of proton donor. The carboxylate group of Asp152 or Asp278, being close to the scissile glycosidic bond, is the likely candidate for the general base catalyst in the hydrolytic reaction. The corresponding counterparts to these residues also appear in CelC of *Pcc* LY34 as Glu57, Asp118, and Asp245. For studying protein structure-function relationships, the enzyme activity was assayed for the single-point mutants of E57A, D118A, and D245A; we replaced all three possible candidates with Ala. We sequenced all the mutant plasmids and confirmed that no error other than the intended mutation was introduced. The activity was not detectable in all the mutants (E57A, D118A, and D245A). The result on the mutant enzymes indicate that Glu57, Asp118, and Asp245 are important for CelC, and that those residues are well conserved in all the endoglucanases of family 8.

DISCUSSION

The current concept is that cellulolytic enzymes released by phytopathogens are involved in the enzymatic hydrolysis of plant cell walls. It is generally believed that these enzymes serve as cell wall modifying enzymes, since their actions may also render other polysaccharide components in the cell wall more susceptible to hydrolysis. A study of the exact roles of these enzymes is therefore important to the understanding of the mechanism of host-parasite interaction in the disease cycle.

Using the CMC-SDS-PAGE method, we previously demonstrated the presence of at least five different cellulases in the wild-type *Pcc* LY34 strain [17] (Fig. 5). Since *Pcc* LY34 is also a convenient strain from a genetic, physiological, and pathological point of view, we chose it for studying the isozyme system. We earlier isolated and characterized *cel*A and *cel*B of *Pcc* LY34 [24]. This study showed the new endoglucanase gene from *Pcc* LY34 as the third endoglucanase, *cel*C. Comparison of the amino acid sequence of CelC with a range of cellulases discovered only limited sequence similarities. On the basis of comparison, 482 sequences corresponding to 52 EC entries were classified to 45 families [12] and CelC was classified into the endoglucanase family 8. Earlier, CelA and CelB showed similarities to cellulases of family 5 and family 12, respectively [24]. Amino acid homology has placed the *cel* genes of *Pch* and *Pcc* into three families (Fig. 4). Three distinct families of *cel* genes (family 5, family 8, and family 12) are extracellular endoglucanases. Two *cel* genes of *Pch* PY35 belong to family 5 (*cel*5Z) and family 8 (*cel*8Y), while two genes of *Pcc* LY34 belong to family 5 (*cel*A) and family 12 (*cel*B). Common *cel* genes between *Pch* PY35 and *Pcc* LY34 are family 5 and family 8 (Fig. 4).

The apparent molecular size of the processed form has been estimated by CMC-SDS-PAGE to be approximately 39 kDa. The difference between the observed molecular size and the calculated one may reflect the existence of the signal peptide, suggesting that the signal peptide of CelC was cleaved between res-22 alanine and res-23 alanine during the secretion of pre-CelC. The mature CelC was also purified and the amino acid sequence of the NH_2-terminal end was determined (data not shown), and the sequence of the first 22 amino acid residues was found to be removed, similar to the predicted sequence of pre-CelC polypeptide beginning with the 23rd alanine residue (Fig. 3).

In nature, cellulose is degraded by concerted action of a number of bacterial and fungal organisms. Degradation of cellulose poses an interesting problem, not only from the biotechnological point of view, but also from the view point of basic research, since the insoluble nature of crystalline cellulose constitutes a challenge for its enzymatic hydrolysis. However, the low activity of cellulase enzymes and the resulting cost of hydrolysis represent major barriers for the use of lignocellulose [8]. Recent study by Zhou and Ingram [30] has identified that *Pch* produces a battery of hydrolases and lyases, which are very effective in the maceration of plant cell walls. Although two endoglucanases (CelZ and CelY) are produced, CelZ represents approximately

95% of the total carboxymethyl cellulase activity. They have examined the effectiveness of CelY and CelZ individually as well as in combination of both enzymes, using carboxymethyl cellulose and amorphous cellulose (acid-swollen cellulose) as substrates. Full synergy was observed by sequential hydrolysis of CMC, provided that CelY was used as the first enzyme. Considering these results, it is possible that the synergistic hydrolysis of cellulose by three endoglucanases (CelA, CelB, and CelC) of *Pcc* LY34 was to enhance the effect of the combination.

The question of why *Pectobacterium* has acquired multiple genes encoding β-1,4-D-endoglucanases provokes considerable interest, particularly since cellulose is chemically homogeneous and theoretically a natural substrate for endoglucanase. It is quite possible that multiple isozymes may allow for more a complex regulatory strategy and degradation of cellulose in diverse environments, or simply provide higher levels of CMCase activity through an increased gene dosage. An understanding of the exact roles of these enzymes is necessary to understand the mechanisms of the plant-pathogen interaction in the disease cycle. The present study supports the concept of a complex system of cellulolytic enzymes. Further study is expected to find the reason for multiple cell wall degrading enzymes, and the cloning and characterization of more *Pcc* LY34 *cel* genes would be helpful to an answer for this.

Acknowledgments

This work was supported by a grant from BioGreen 21 Program and by ARPC Program, Ministry of Agriculture and Forestry, Republic of Korea. W.J.L. is the recipient of a BK21 fellowship from the Ministry of Education (2004). C.L.A. is the recipient of a fellowship from the Cooperative Research Fund for the Development of Useful Nature Resources and Restoration of Degraded Ecosystems in Northeast Asia Regions.

REFERENCES

1. Alzari, P. M., H. Souchon, and R. Dominguez. 1996. The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. *Structure* **4:** 265-275.
2. Aymeric, J. L., A. Guiseppi, M. C. Pascal, and M. Chippaux. 1988. Mapping and regulation of the *cel* genes in *Erwinia chrysanthemi*. *Mol. Gen. Genet.* **211:** 95-101.
3. Beguin, P. 1990. Molecular biology of cellulose degradation. *Annu. Rev. Microbiol.* **44:** 219-248.
4. Boyer, M. H., B. Cami, A. Koutoujansky, J. P. Chambost, C. Frixon, and J. Cattaneo. 1987. Isolation of the gene encoding the major endoglucanase of *Erwinia chrysanthemi*; homology between *celZ* genes of two strains of *Erwinia chrysanthemi*. *FEMS Microbiol. Lett.* **41:** 351-356.
5. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72:** 248-254.
6. Cho, S. J., S. R. Park, M. K. Kim, W. J. Lim, S. K. Ryu, C. L. An, S. Y. Hong, H. Kim, and H. D. Yun. 2002. Cloning of *cel8Y* gene from *Pectobacterium chrysanthemi* PY35 and its comparison to *cel* genes of soft-rot *Pectobacterium*. *Mol. Cells* **13:** 28-34.
7. Cooper, V. J. C. and G. P. C. Salmond. 1993. Molecular analysis of the major cellulase (CelV) of *Erwinia carotovora*: Evidence for an evolutionary mix-and-match of enzyme domains. *Mol. Gen. Genet.* **241:** 341-350.
8. Dale, B. E. 1999. Biobased industrial products: Bioprocess engineering when cost really counts. *Biotechnol. Prog.* **15:** 775-776.
9. Guerin, D. M. A., M. B. Lascombe, M. Costtabel, H. Souchon, V. Lamzin, P. Beguin, and P. M. Alzari. 2002. Atomic (0.94 Å) resolution structure of an inverting glycosidase in complex with substrate. *J. Mol. Biol.* **316:** 1061-1069.
10. Guiseppi, A., J. L. Aymeric, B. Cami, F. Barras, and N. Creuzet. 1991. Sequence analysis of the cellulase-encoding *celY* gene of *Erwinia chrysanthemi*: A possible case of interspecies gene transfer. *Gene* **106:** 109-114.
11. Halldorsdottir, S., E. T. Thorolfsdottir, R. Spilliaert, M. Johansson, S. H. Thorbjarnardottir, A. Palsdottir, G. O. Hreggvidsson, J. K. Kristjansson, O. Holst, and G. Eggertsson. 1998. Cloning, sequencing and overexpression of a *Rhodothermus marinus* gene encoding a thermostable cellulase of glycosyl hydrolase family 12. *Appl. Microbiol. Biotechnol.* **49:** 277-284.
12. Henrissat, B. and A. Bairoch. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293:** 781-788.
13. Hong, I. P., H. K. Jang, S. Y. Lee, and S. G. Choi. 2003. Cloning and characterization of a bifunctional cellulase-chitosanase gene from *Bacillus licheniformis* NBL420. *J. Microbiol. Biotechnol.* **13:** 35-42.
14. Johnvesly, B., S. Virupakshi, G. N. P. Ramalingam, and G. R. Naik. 2002. Cellulase-free thermostable alkaline xylanase from thermophilic and alkalophilic *Bacillus* sp. JB-99. *J. Microbiol. Biotechnol.* **12:** 153-156.
15. Kim, K. C., S. S. Yoo, Y. A. Oh, and S. J. Kim. 2003. Isolation and characteristics of *Trichoderma harzianum* FJ1 producing cellulases and xylanase. *J. Microbiol. Biotechnol.* **13:** 1-8.
16. Kim, K. J., Y. J. Yang, and J. K. Kim. 2002. Production of α-glucosidase inhibitor by β-glucosidase inhibitor-producing *Bacillus lentimorbus* B-6. *J. Microbiol. Biotechnol.* **12:** 895-900.
17. Lim, S. T., Y. W. Park, S. J. Cho, and H. D. Yun. 1997. Phytopathogenicity of *Erwinia carotovora* subsp. *carotovora* LY34 and production of CMCase isozymes. *Kor. J. Appl. Microbiol. Biotechnol.* **25:** 468-476.

18. Lymar, E. S., B. Li, and V. Renganathan. 1995. Purification and cellulose-degrading cultures of *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **61:** 2976-2980.
19. Mae, A. R., R. Heikinheimo, and E. T. Palva. 1995. Structure and regulation of the *Erwinia carotovora* SCC3193 cellulase gene *cel*V1 and the role of cellulase in phytopathogenicity. *Mol. Gen. Genet.* **247:** 17-26.
20. Miller, G. L., R. Blum, W. E. Glennon, and A. L. Burton. 1960. Measurement of carboxylmethylcellulose activity. *Anal. Biochem.* **2:** 127-132.
21. Nakamura, K., N. Misawa, and K. Kitamura. 1986. Sequence of a cellulase gene of *Cellulomonas uda* CB4. *J. Biotechnol.* **4:** 247-254.
22. Park, S. R., M. K. Kim, J. O. Kim, D. W. Bae, S. J. Cho, Y. U. Cho, and H. D. Yun. 2000. Characterization of *Erwinia chrysanthemi* PY35 *cel* and *pel* genes existing in tandem and rapid identification of their gene products. *Biochem. Biophys. Res. Commun.* **268:** 420-425.
23. Park, W. Y., S. T. Lim, and H. D. Yun. 1998. Cloning and characterization of a CMCase gene, *cel*B, of *Erwinia carotovora* subsp. *carotovora* LY34 and its comparison to *cel*A. *Mol. Cells* **8:** 280-285.
24. Park, Y. W., S. T. Lim, S. J. Cho, and H. D. Yun. 1997. Characterization of *Erwinia carotovora* subsp. *carotovora* LY34 endo-1,4-β-glucanase genes and rapid identification of their gene products. *Biochem. Biophys. Res. Commun.* **241:** 636-641.
25. Pissavin, C., J. Robert-Baudouy, and N. Hugouvieux-Cotte-Pattat. 1996. Regulation of *pel*Z, a gene of the *pel*BC cluster encoding a new pectate lyase in *Erwinia chrysanthemi* 3937. *J. Bacteriol.* **178:** 7187-7196.
26. Ryu, S. K., S. R. Park, W. J. Lim, M. K. Kim, C. L. An, S. J. Cho, Y. W. Park, B. K. Kim, H. Kim, and H. D. Yun. 2001. Cloning and characterization of the *cel*9A gene of marine bacterium *Pseudomonas* sp. SK38 causing green spot rot of *Porphyra dentata*. *Appl. Microbiol. Biotechnol.* **57:** 138-145.
27. Saarilahti, H. T., B. Henrissat, and E. T. Palva. 1990. CelS: A novel endoglucanase identified from *Erwinia carotovora* subsp. *carotovora*. *Gene* **90:** 9-14.
28. Sambrook, J. and D. W. Russel. 2001. *Molecular Cloning: A Laboratory Manual*, 3rd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, U.S.A.
29. Shevchik, V. E., G. Condemine, N. Hugouvieux-Cotte-Pattat, and J. Robert-Baudouy. 1996. Characterization of pectin methylesterase B, an outer membrane lipoprotein of *Erwinia chrysanthemi* 3937. *Mol. Microbiol.* **19:** 445-466.
30. Zhou, S. and L. O. Ingram. 2000. Synergistic hydrolysis of carboxymethyl cellulose and acid-swollen cellulose by two endoglucanases (CelZ and CelY) from *Erwinia chrysanthemi*. *Appl. Environ. Microbiol.* **67:** 6-14.