Molecular Cloning and Characterization of Two Major Endoglucanases from *Penicillium decumbens*

Wei, Xiao-Min¹, Yu-Qi Qin³*, and Yin-Bo Qu¹,²*

¹State Key Laboratory of Microbial Technology, and ²National Glycoengineering Research Center; ³School of Pharmaceutical Science Shandong University, 27 Shanda South Road, Jinan, Shandong 250100, P. R. China

Received: April 28, 2009 / Revised: August 5, 2009 / Accepted: September 21, 2009

Two major endoglucanase genes (*cel7B* and *cel5A*) were cloned from *Penicillium decumbens* 114-2 using the method of modified thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR). The result of Southern blotting suggested that *P. decumbens* has a single copy of the *cel5A* gene and a single copy of the *cel7B* gene in its chromosomal DNA. The expression levels of *cel5A* and *cel7B* were determined by means of real-time quantitative PCR, suggesting that the two genes were coordinately expressed, and repressed by glucose and induced by cellulose. Both endoglucanase genes were expressed in *Saccharomyces cerevisiae* and the recombinant proteins were purified. The recombinant Cel7B and Cel5A were both optimally active at 60°C and pH 4.0. The recombinant Cel7B showed more than 8-fold, 30-fold, and 5-fold higher enzyme activities toward carboxymethyl cellulose, barley β-glucan, and PASC, respectively, in comparison with that of Cel5A. However, their activities toward pNPC and Avicel showed minor differences. The results suggested that Cel7B is a strict endoglucanase, whereas Cel5A showed processivity because of its relative higher ability to hydrolyze the crystal cellulose.

**Keywords:** *Penicillium decumbens*, Cel7B, Cel5A, processivity

Cellulose, one of most abundant biopolymers on Earth, can be efficiently degraded through the synergistic action of cellulolytic enzymes [12]. Those enzymes are divided into three major groups based on their substrate specificity [11]: endoglucanases (E.C. 3.2.1.4), cellubiohydrolases (E.C. 3.2.1.91), and β-glucosidases (E.C. 3.2.1.21). Fungi and bacteria are the main organisms of cellulose degradation and have evolved complex, multicomponent enzyme systems to mediate conversion of cellulose to glucose [23]. Since fungi produce a large amount of extracellular cellulases, they have been widely studied and utilized for industrial application in the food, animal feed, paper and pulp, textile, and chemical industries [2, 5, 31]. Among the cellulolytic fungi, *Trichoderma reesei* has been widely studied because of its high capability of cellulase secretion. As another important cellulase-producing fungus, *Penicillium* species showing the ability of producing high activities of cellulase and hemicellulase have been described, and they displayed relatively higher β-glucosidase activity than that of *Trichoderma* spp. [6, 14, 20]. *P. echinulatum* has been identified as a potential candidate for cellulase production because of its secreting capacity is almost equivalent to the best fungal strains reported in the literature [9, 10]. However, there was little information about the single component of cellulase of genus *Penicillium*, especially the information about the endoglucanase.

*P. decumbens* 114-2 is a fast-growing cellulolytic fungus that was isolated from soil [25] and its catabolic repression-resistant mutants have been used industrially for biomass hydrolysis [21, 26]. As per our previous report, its cellulase is composed of endoglucanase, cellubiohydrolase, and β-glucosidase, based on zymogram analysis following polyacrylamide gel electrophoresis [29]. In this paper, we focused on the isolation and characterization of two major endoglucanase (Cel7B and Cel5A) genes, and also described the expression of two endoglucanase genes in *Saccharomyces cerevisiae* and some properties of these enzymes.

**Materials and Methods**

**Strains, Plasmids, and Culture Conditions**

*P. decumbens* 114-2 was cultured on wheat bran extract agar (10% wheat bran extract, 2% agar) plates at 30°C for 4 days, and then the
spores were inoculated into the cellulose-inducing medium. The medium was modified Mandel's solution containing (per liter) 3 g KH₂PO₄, 2.6 g NaNO₃, 0.5 g MgSO₄·7H₂O, 0.5 g CaCl₂, 0.5 g urea, 7.5 mg FeSO₄·7H₂O, 2.5 mg MnSO₄·H₂O, 3.6 mg ZnSO₄·7H₂O, 3.7 mg CoCl₂·6H₂O, 1 g peptone with 1% microcrystalline cellulose, and 2% wheat bran as carbon source. After growth for 2 days at 30°C with shaking, mycelia were harvested by filtration and used for DNA and RNA isolation.

*Escherichia coli* DH5α and the plasmid pMD18-T vector (TaKaRa, Dalian, China) were used for general DNA manipulation and DNA sequencing. *S. cerevisiae* H158 [MATa leu2-3 leu2-112 ura3-52 trp1-289 his4-519 prb1 cir+] was used as the host and the plasmid pJ401 (ura3, PGK promoter, and PGK terminator) was used as the expression vector. Yeast cells were cultured as previously described [24].

### Cloning and Sequencing of the Full-Length Genes by TAIL--PCR

According to the conserved amino acid sequences of Cel7B and Cel5A from known fungi including *T. reesei*, *Hunicola insolens*, *Macrophomina phaseolina*, *Fusarium oxysporum*, *Phanerochaete chrysosporium*, *Aspergillus fumigatus*, and *Irpex lacteus*, two pairs degenerate primers, Endol and EndoI (Table 1), were designed to amplify the fragments of the cel7B and cel5A genes, respectively. Genomic DNA of *P. decumbens* was used as the template. To amplify the 5'-end and 3'-end of the cel7B and cel5A genes, primer sets for TAIL--PCR were designed using the oligonucleotides described below (Table 1).

#### Table 1. Oligonucleotides used in the studies.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’–3’)*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primers for degenerate PCR</strong></td>
<td></td>
</tr>
<tr>
<td>Endol-F</td>
<td>CTBCCNTGYGGNRTWRYAYGG</td>
</tr>
<tr>
<td>Endol-R</td>
<td>GTNWDCCRATRDCCTCCARCG</td>
</tr>
<tr>
<td>EndoI-F</td>
<td>AAYATYTTYCGYTTYBCC</td>
</tr>
<tr>
<td>EndoI-R</td>
<td>TCNSWRTCVAGRATCTRTG</td>
</tr>
<tr>
<td><strong>Primers for TAIL--PCR</strong></td>
<td></td>
</tr>
<tr>
<td>TAIL-AD1</td>
<td>TGWGNAGWANCAGA</td>
</tr>
<tr>
<td>TAIL-AD2</td>
<td>AGWGNAGWANCAWAGG</td>
</tr>
<tr>
<td>TAIL-AD3</td>
<td>STGNASTNCTNGC</td>
</tr>
<tr>
<td>TAIL-AD4</td>
<td>NTGASTWTSWGWT</td>
</tr>
<tr>
<td>TAIL-AD5</td>
<td>NGTCGASWGANAWGAA</td>
</tr>
<tr>
<td>TAIL-AD6</td>
<td>WGTGNAGWANCAGA</td>
</tr>
<tr>
<td>TAIL-AD7</td>
<td>WTCGNTCTWANTACT</td>
</tr>
<tr>
<td><strong>Primers for real-time RT-PCR</strong></td>
<td></td>
</tr>
<tr>
<td>act-F</td>
<td>CTCATCCAGGCCGTTTCTG</td>
</tr>
<tr>
<td>act-R</td>
<td>CATGAGTGATGGCAGGAC</td>
</tr>
<tr>
<td>cel5A-F</td>
<td>AACGCGACGCTTCAAA</td>
</tr>
<tr>
<td>cel5A-R</td>
<td>CAGCAGGCTCATTCAAGGTA</td>
</tr>
<tr>
<td>cel7B-F</td>
<td>AACCTGGAAGAACAGGC</td>
</tr>
<tr>
<td>cel7B-R</td>
<td>CCTTGTCACAGTCATCGGAG</td>
</tr>
</tbody>
</table>

*For degenerate primers, the following abbreviations are used: N: A, T, G, C; B: T, G, C; D: A, T, G; V: A, G, C; Y: C, T; S: C, G; W: A, T; and R: A, G.*

Southern Blot Analyses

Chromosomal DNA from *P. decumbens* 114-2 was digested to completion overnight with *EcoRI*, *BamHI*, *XhoI*, *XbaI*, and *HindIII* (Fermentas, Lithuania), separated by agarose gel electrophoresis, and transferred to Hybond-N nylon membranes (Amersham, U.S.A.). DNA labeling and detection were done by using a DIG-High prime DNA labeling and detection starter kit (Roche, Germany), respectively, following the manufacturer's protocol.

Quantification of Gene Transcripts

To determine endoglucanase genes expression, glucose (2%) and microcrystalline cellulose (2%) served as the repression and induction carbon source in the liquid medium, respectively. The inoculated media were incubated for 41 h at 30°C with shaking (180 rpm). The mycelia harvested from the culture were used to extract mRNA. First-strand cDNA was synthesized using a PrimeScript RT reagent kit (TaKaRa, Dalian, China). PCR amplification was performed using a SYBR Premix Ex Taq kit (TaKaRa, Dalian, China) with primer sets for act, cel5A and cel7B (Table 1). Based on the data from quantitative PCR, the copy number of gene expression was calculated by comparing with a standard curve of each gene, respectively. The transcript number of the actin gene was quantified as an internal standard.

Expression, Purification, Deglycosylation, and Activity Staining of Recombinant Endoglucanases

The gel-purified PCR fragment of the full-length endoglucanase genes with its own signal peptide sequence was digested with *EcoRI* and *XhoI* and ligated into pAJ401. The expression plasmids were then transformed into *S. cerevisiae* H158. Purification and activity staining of the recombinant Cel7B and Cel5A from *S. cerevisiae* H158 were performed as previously described [24]. To determine whether the recombinant endoglucanases were N-linked glycosylated, the recombinant Cel7B and Cel5A were treated with endoglycosidase H (New England Biolabs, U.S.A.), respectively, following the manufacturer’s protocol. The molecular mass and purity of recombinant Cel7B and Cel5A were analyzed with the ImageQuant TL software (Amersham Biosciences Corp, U.S.A.).

### Enzyme Assay

The enzyme activities toward several substrates were determined at 60°C by measuring the amount of reducing sugar equivalents. For CMC, barley β-glucan, and xylan (Sigma, U.S.A.), the concentration of substrates was 5 g/l and the hydrolysis time was 30 min. For PASC (phosphoric acid swollen cellulose) and Avicel (Sigma, U.S.A.), the concentration of substrates was 5 g/l and the hydrolysis time was 24 h. Reducing sugar was analyzed with the Somogyi–Nelson method [22, 27]. One unit of the enzyme activity was defined as the amount of enzyme releasing 1 μmol of glucose equivalent per
minute. Optimum pH and pH stability for purified enzymes were determined on CMC. The buffers used were 50 mM citrate buffer (pH 3), 50 mM acetate buffer (pH 4–6), and 50 mM phosphate buffer (pH 7–8). Optimum temperature and thermal stability were assayed by incubating the enzymes at the temperatures in 50 mM acetate buffer (pH 4.0), followed by measuring the residual activity at different time points.

RESULTS AND DISCUSSION

Characterization of the Endoglucanase Genes

The structure of the cel5A and cel7B. The gene cel7B (GenBank EU339127) has an open reading frame of 1,425 bp, encoding a protein with 474 amino acid residues. Comparison of the deduced Cel7B amino acid sequence from P. decumbens with those available on the NCBI database reveals identity values of 75%, 75%, 71%, and 59%, respectively, with the Cel7B from Neosartorya fischeri (GenBank XP_001257357), A. fumigatus (GenBank XP_747897), A. terreus (GenBank XP_001217291), and T. reesei (GenBank AAX28897). An alignment of the deduced polypeptide sequence showed that the modular structure was conserved, with an N-terminal CD (aa 22–401) belonging to glycosyl-hydrolase family 7 via a proline/serine/threonine rich linker (aa 402–441) region to the C-terminal fungal cellulose binding domain (CBD) (aa 442–474). Analysis by the signalP method [4] predicted that Cel7B has an N-terminal secretion signal peptide with a cleavage point between residues 21 and 22. The putative residues of the nucleophile and proton acid catalysts for Cel7B are Glu216 and Glu221, respectively.

The gene cel5A (GenBank EU315320) has an open reading frame of 1,236 bp, encoding a putative protein of 411 amino acid residues. NCBI BLASTP search revealed that the amino acid sequence deduced from P. decumbens Cel5A was similar to those of the Cel5A from P. janthinellum (77%, GenBank CAA61740), M. phaseolina (62%, GenBank AAB03889), T. reesei (58%, GenBank ABA64553), and P. chrysosporium (44%, GenBank AAU12275). Compared with Cel7B, Cel5A has a family 5 catalytic domain at the C-terminus (aa 100–411) and a fungal CBD at the N-terminus (aa 22–57). The two putative catalytic residues were identified to be Glu342 and Glu232 as the nucleophile and the proton acid catalysts, respectively. The signal sequence was also determined by the signalP method, and the result suggested that the amino acid residues 1–21 were signal peptides.

Gene copy number of cel5A and cel7B. Southern blotting was performed to identify whether the gene is present in only one or multiple copies in the P. decumbens genome (Fig. 1A and 1B). A single hybridizing band was present in all digestions. The results suggest that both cel7B and cel5A are a single copy in the chromosomal DNA of P. decumbens, which is the same as that reported in T. reesei [30]. In contrast, both P. chrysosporium and P. janthinellum contain multiple copies of cellulase genes in their chromosomal DNA [8, 17].

Expression levels of cel5A and cel7B. Expression levels of cel5A and cel7B in cultures containing different carbon sources were quantified by means of real-time quantitative
PCR. As shown in Fig. 2, in the glucose medium, the expression levels of cel5A and cel7B were clearly lower, and the expression levels were obviously increased when mycelia were cultivated in cellulose medium. The transcript number of cel5A and cel7B increased 58 times and 108 times, respectively. These results indicate that cel5A and cel7B are coordinately expressed and repressed by glucose and induced by cellulose. In T. reesei and Aspergillus spp., cellulase genes were coordinately expressed and regulated at the transcriptional level [13, 28]. Several transcriptional factors were reported to regulate the expression of cellulase genes in T. reesei and Aspergillus spp. [1, 28], and we also cloned the gene homologs of these factors in P. decumbens such as ACEI, XlnR, and CreA (GenBank EU239662; GenBank EU855737; GenBank EU239661). It is therefore likely that, as in T. reesei and Aspergillus spp., these transcriptional factors also regulate the expression of cellulase genes in P. decumbens.

Characterization of the Two Endoglucanases

Molecular masses of recombinant Cel7B and Cel5A.
The apparent molecular masses of Cel7B and Cel5A estimated by SDS–PAGE were about 79 kDa (Fig. 3A) and 70 kDa (Fig. 3B), respectively. The observed molecular masses of Cel7B and Cel5A were higher than the calculated molecular masses (47 kDa and 42 kDa) from the deduced amino acids sequence. Activity staining showed that it had the detectable CMCase activities regardless of their larger molecular masses. As most fungal cellulases are glycoproteins, the difference may be caused by the over-glycosylation of the recombinant endoglucanase produced by S. cerevisiae [18]. Cel7B contains three putative N-glycosylation sites (Asn97, Asn124, and Asn202) and several possible O-glycosylation sites mostly at the linker region. Cel5A has only one putative N-glycosylation site (Asn278) and also has O-glycosylation sites at the linker region. After endoglycosidase H treatment, the recombinant Cel7B decreased in molecular mass from 79 to 71 kDa (Fig. 3A)
and retained about 70% of its original activity (data not shown). However, Cel5A was not cleaved by endoglycosidase H (Fig. 3B), suggesting that the recombinant Cel5A was not N-linked glycosylated. It is likely that the recombinant Cel5A was heavy O-linked glycosylated, just as the reported endoglucanase in *Syncephalastrum racemosum* [32].

**Effects of temperature and pH on enzyme activity.** The optimal temperature for the reaction of Cel7B and Cel5A was 60°C and the optimal pH of both enzymes was pH 4. Both enzymes were stable at pH between 3 to 8 at 4°C for 16 h. At temperatures between 30°C to 50°C, Cel7B and Cel5A showed similarly thermal stability; both enzymes retained more than 90% of the maximal activity after incubation for 3 h. When incubated at 60°C for 1 h, Cel7B retained more than 90% of the initial activity, and Cel5A retained only 80% of the initial activity (data not shown).

**The difference between Cel7B and Cel5A: A strict endoglucanase and a processive endoglucanase, respectively.** The purified recombinant Cel7B and Cel5A were assayed for their activities toward different substrates (Table 2). Both enzymes showed relatively high activity toward CMC and low activity toward PASC and Avicel, suggesting that Cel7B and Cel5A are both endoglucanases. The result also showed that Cel7B and Cel5A belong to the group of “nonspecific” EGs, since both enzymes hydrolyzed not only cellulose, but also the β-1,4 linkages in xylan and barley β-glucan. It has been reported that Cel7B and Cel5A of *T. reesei* are able to hydrolyze CMC, barley β-glucan, xylan, PASC, and Avicel [16, 31]. The hydrolytic characteristic of Cel7B and Cel5A of *P. decumbens* are therefore in good agreement with that of *T. reesei*.

The recombinant Cel7B hydrolyzed most substrates more efficiently than Cel5A. The catalytic efficiency of Cel7B was 8 times higher than that of Cel5A toward CMC, 30 times higher toward β-glucan, and 5 times higher toward PASC. However, although the activities of Cel5A toward CMC, barley β-glucan, and PASC were much lower than those of Cel7B, its activity toward Avicel had only minor difference with that of Cel7B, and it had a higher activity toward pNPC. We also analyzed the products of recombinant Cel5A hydrolyzing Avicel. The main product was cellobiose together with traced glucose (data not shown), which are not the expected products of a typical endoglucanase activity on a solid substrate. According to the most widely accepted hypothesis, classical endoglucanases nick the cellulose internally, thus disrupting its crystallinity and generating new free ends in the polymer [3]. However, it has been reported that some organisms produce processive endoglucanases, which cleave cellulose internally but also release soluble oligosaccharides before detaching from the polysaccharide [7, 33, 34]. The result showed that Cel7B is a strict endoglucanase, whereas Cel5A shows processivity because of its relative higher ability to hydrolyze the crystal cellulose.

As an endoglucanase, the recombinant Cel7B showed relatively high specific activity toward amorphous cellulose among the fungal endoglucanases found so far. In the well-studied fungi *T. reesei* and *H. insolens*, Cel5A was more efficient than other endoglucanases when hydrolyzing amorphous cellulose [15, 16, 31]. However, our result showed that Cel7B from *P. decumbens* was more efficient than Cel5A, and this was different from that of *T. reesei* and *H. insolens*. The recombinant Cel7B was more active against barley β-glucan than CMC; a likely reason for this was that CMC was highly substituted with methoxy side chains, which may interfere with the enzyme activity [15].

This is the first report of cloning of the endoglucanase genes from *P. decumbens*. Like most fungi, *P. decumbens* also produces two major endoglucanases belonging to GH family 7 and family 5, respectively. However, the recombinant Cel5A and Cel7B showed different enzymatic properties, which may lead to some differences between the roles of the two endoglucanases in hydrolyzing cellulose. Our results indicate that cel5A and cel7B are coordinately expressed and can be repressed and induced together. Further investigation of the induction and repression mechanisms would be important for improving the fungus for the production of cellulas.

**Table 2. Substrate specificity of the recombinant endoglucanases.**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Specific activity (U/mg of protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cel7B</td>
</tr>
<tr>
<td>CMC</td>
<td>411.3±14.5</td>
</tr>
<tr>
<td>PASC</td>
<td>1.1±0.03</td>
</tr>
<tr>
<td>Avicel</td>
<td>0.15±0.02</td>
</tr>
<tr>
<td>pNPG</td>
<td>0</td>
</tr>
<tr>
<td>pNPC</td>
<td>2.5±0.1</td>
</tr>
<tr>
<td>Xylan</td>
<td>9.0±0.4</td>
</tr>
<tr>
<td>β-Glucan</td>
<td>582.1±13.9</td>
</tr>
</tbody>
</table>


The enzyme activities for pNPC and pNPG were determined as the enzyme necessary to release 1 μmol of *p*-nitrophenol per minute.

**Acknowledgments**

This work was supported by grants from the National High Technology (863) Program of China (No. 2006AA020201) and the National Natural Science Foundation (30900028) of the People’s Republic of China.

**References**


