Cloning and expression of *Lentinula edodes* cellobiohydrolase gene in *E. coli* and characterization of the recombinant enzyme

Sabira TAIPAKOVA¹, Bauyrzhan SMAILOV¹, Gulshan STANBEKOVA² and Amangeldy BISSENBAEV¹,*

¹Department of Genetics and Molecular Biology, Faculty of Biology, al-Farabi Kazakh National University, 050038, Almaty, al-Farabi 71, Kazakhstan
²Institute of Molecular Biology and Biochemistry, Dosmukhamedov 86, 480012, Almaty, Kazakhstan
(* author for correspondence; Amangeldy.Bisenbaev@kaznu.kz)

Received: 11 March 2011; Accepted: 10 June 2011

Abstract

A gene encoding cellobiohydrolase CEL7A was successfully isolated from the *L. edodes* mushroom strain N127 using RT-PCR. The deduced amino acid sequence encoded by cel7A showed high homology with the sequence of glycoside hydrolase family 7. To confirm the gene sequence encoding the CEL7A the cloned gene was expressed in *E. coli*. For the first time the cel7A gene from the *L. edodes* was expressed in *E. coli* and characterized. The recombinant CEL7A has the ability to hydrolyze Avicel, Filter paper, p-Nitrophenyl β-D-lactopyranoside (pNP-Lac) and p-Nitrophenyl β-D-cellobioside (pNP-Cel). The activity of the cloned enzyme towards carboxymethylcellulose (CMC) is much lower. It showed an optimal working condition at 50 °C and pH 7.

Keywords: Cellulose, cellobiohydrolase, *Lentinula edodes*, gene expression, enzyme activity

Introduction

Cellulose, consisting of glucose units linked together by β-1,4-glycosidic bonds, is the most abundant carbohydrate in the biosphere. An estimated rate of cellulose synthesis is approximately 4 x 10⁷ tons per year. For a long-range solution for problems of energy, chemicals, and food, cellulose is the most promising renewable carbon source that is available in large quantities (Murai *et al.*, 1998).

Full hydrolysis of cellulose requires synergistic action of three major types of enzymatic activity: 1,4-β-D-endoglucanases (EC 3.1.2.4), 1,4-β-D-cellobiohydrolases (CBH; EC 3.1.2.91) and β-glucosidases (EC 3.2.1.21). CBHs hydrolyze...
crystalline cellulose by initiating their action from the ends of the cellulose chains and producing primarily cellobiose. Endoglucanases catalyze hydrolysis internally in the cellulose chain and attack the amorphous regions in cellulose, providing new chain ends for CBHs. β-glucosidases hydrolyze cellobiose to glucose. These enzymes are collectively known as cellulases and act in a synergistic manner to facilitate complete cleavage of the cellulose β-1,4-glycosidic bonds (Teeri, 1997).

CBHs are key components in the multi-enzyme cellulose complexes. Most of them belong to glycosyl hydrolase families 6 or 7. Cellulolytic fungus generally produces two different CBHs, CBHI and CBHII. These two types of enzymes which are classified based on sequence identity and can achieve complete, although slow, solubilization of cellulose crystals even without help of endoglucanases (Teeri, 1997).

Lentinula edodes, commonly referred to as the Shiitake mushroom, is the most popular and economically important edible mushroom in the world because of its taste, nutritional and medicinal properties (Mizuno, 1995). L. edodes has traditionally been grown on freshly cut logs (Leatham, 1985). It is a white rot fungus that decays for nutrients all the major polymers (cellulose, hemicelluloses, and lignin) found in wood lignocelluloses (Lee et al., 2001). Thus, L. edodes produces a wide variety of enzymes that may have high activities against cellulosic biomass.

Several papers reported that L. edodes has a strong ability to produce thermostable cellulases, including CBH (Leatham, 1985; Lee et al., 2001; Pereira Júnior et al., 2003). L. edodes produces at least two forms of cellbiohydrolase (CEL7A and CEL6B). Genes encoding each of these enzymes were cloned from L. edodes grown on a wood substrate using a PCR based strategy with degenerate primers directed at the cellulose-binding domain (Lee et al., 2001). On the basis of nucleotide sequence analysis it was shown that cel7A encodes a 516-amino acid (aa) protein that belongs to glycosyl hydrolase family 7 and has sequence similarities to CBHI genes from other fungi. Whereas cel6B gene encodes a 444 aa protein that belongs to glycosyl hydrolase family 6 and has sequence similarities to CBHII genes from other fungi. However, both fungal CBHs genes have not been previously expressed and characterized.

In this paper we report the cloning of a L. edodes cel7A gene and its successful expression in a heterologous host, E. coli.

Materials and methods

Materials

Restriction enzymes, T4 DNA ligase and Taq DNA polymerase were purchased from Fermentas Life Sciences (St. Leon-Rot, Germany). Plasmid DNA was isolated using High Pure Plasmid Isolation Kit and DNA fragments were purified using Agarose Gel DNA Extraction Kit (Roche Diagnostics GmbH, Mannheim, Germany). All other chemicals and reagents were analytical grade and were supplied by Sigma-Aldrich Corp (St. Louis, MO USA). The expression vector pET11d was from Invitrogen, USA.

Strains and culture conditions

L. edodes strain N127 was obtained from collection of Kazakh National Agrarian University (Almaty, Kazakhstan). L. edodes strain N127 was cultured at room temperature on poplar sawdust blocks. From sawdust blocks the mycelium was harvested with a scalpel. The mycelium samples were pooled and frozen in liquid nitrogen and stored at -80°C until further processing. Escherichia coli DH5α were used as a host for plasmid propagation. E.coli Rosetta (DE3) (Invitrogen, USA) was used as a host for expression of the recombinant CEL7A.

Total RNA extraction and RT-PCR

Total RNA extraction of L.edodes and RT-PCR was performed as described (Sambrook et al., 1989). Total RNA isolated as described above was used as a template for RT reactions. Using the sequence of the L. edodes cel7A cDNA available in the GenBank database (GenBank accession number AF411250), the forward primer, cel7A Dir: 5'-GATCACCATGGTCCGAACAGCAGCTCTCCT-3' and the reverse primer, cel7A Rev: 5'-CTAGGATCCCTACAACATTGACTGTAGTAAGG-3' were designed and used for RT-PCR. The underlined bases in primers are the sites of restriction enzymes NcoI and BamHI, respectively.
**Expression of L. edodes cel7A gene in E.coli**

**Construction of expression vector and protein expression**

The PCR product and pET11d vector were double-digested by NcoI and BamHI at 37°C according to the protocol. The digested products were ligated by T4 ligase at 4°C for overnight. The recombinant vector pET-11d/cel7A was transformed into E.coli Rosetta (DE3) competent cells and the transformants were selected on LB plates with ampicillin. The different resistant transformants were picked and inoculated into 20 mL LB medium with 200 μg/ml ampicillin and grown at 37°C in a shaking incubator (150 rpm) for 12 to 16 h. Then, the incubation was continued in 1L LB medium. When the culture grown at 30°C reached an OD600=0.4–0.6, isopropyl-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM to induce expression. After incubation cells were harvested by centrifugation at 6000 rpm for 5 min at 4°C. Cells were resuspended in storage buffer containing 20 mM HEPES-KOH pH 7.6, 40 mM NaCl. The cells were lysed by successive cycles of freezing (liquid nitrogen) and thawing by sonication for 20 s at half power (14 mA) until a clear cell lysate was obtained. This mix was used as crude protein extracts or alternatively the sonicated mixture was centrifuged at 14 000 g for 10 min to remove cell debris, and this cleared supernatant fluid was used as crude proteins extracts. The crude protein extracts and cellular fractions regarded as crude enzyme extracts were used for identifying CEL7A expression by SDS-PAGE, Western blotting and enzyme analysis.

**Preparation of anti-CEL7A antibodies and Western blotting**

The enzyme sample (preincubated with Complete EDTA-free Protease Inhibitor Cocktail) was electrophoresed on a 10% (w/v) polyacrylamide gel containing 0.1% (w/v) SDS. The gel was stained with Coomassie Brilliant Blue R-250 and the major protein band was excised from the gel and used as antigen. The gel piece was homogenized with liquid nitrogen and emulsified in Freund’s complete adjuvant (Sigma) and injected subcutaneously into a rabbit. The same amount of booster injections was given every two weeks using incomplete Freund’s adjuvant. The rabbit was bled before the first injection (preimmune serum) and then one week after the last booster to obtain immune serum (anti-CEL7A antibodies). Western blots were performed by standard protocols using 1:400 anti-CEL7A polyclonal antibody diluted in blocking buffer and 1:10000 horseradish peroxidase-conjugated secondary antibody.

**Mass spectrometry**

The target bands on SDS-PAGE gels were excised and subjected to in-gel digestion with trypsin followed by peptide mass fingerprinting by matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF) using a MALDI-TOF-TOF spectrometer, Ultraflex II (Bruker). Spectra from all experiments were converted to DTA files and merged to facilitate database searching using the Mascot search algorithm v2.1 (Matrix Science, Boston, MA) against the non-redundant protein sequences of GenBank (National Center for Biotechnology Information).

**Protein and enzyme assays**

The protein content was determined by Bradford method (Bradford, 1976) Enzyme activity was assayed with Avicel, filter paper, carboxymethylcellulose (CMC) as a substrate by measuring the amount of reducing sugar according to the Nelson-Somogyi method (Somogyi, 1952; Nelson, 1944). The reaction mixture consisted of 50 mM sodium phosphate buffer (pH 7.0), 1% (w/v) of CMC or Avicel or 50 mg of filter paper (1x 6 cm) and an appropriate amount of enzyme. The generated reducing sugar was quantified at 540 nm and using D-glucose as a standard for spectrophotometric estimation. One unit of cellobiohydrolase activity was defined as the amount of enzyme that releases 1 μM of reducing sugar per minute per mg of total protein under the assay conditions. When p-Nitrophenyl β-D-lactopyranoside (pNP-Lac) and p-Nitrophenyl β-D-cellobioside (pNP-Cel) were used as a substrate, assays were performed in 500 μl reaction volume in 50 mM sodium phosphate buffer (pH 7.0) containing 8 mM of pNP-Cel or pNP-Lac with appropriate amounts of enzyme. The reaction was terminated by the addition of 2 ml of 1 M Na2CO3 after 1 h incubation. One unit of enzyme activity was defined as the amount of the enzyme that produced the equivalence of 1 μM pNP at the optimal conditions in 1 min per 1 mg of total protein.
Results

To isolate a gene coding CEL7A, we designed two oligonucleotide primers based on the previously published sequence (Lee et al., 2001). First, PCR was done using genomic DNA of *L. edodes* N127 as a template. PCR product showed a clear band about 2000 bp (Figure 1B).

In the following experiments we amplified *cel7A* cDNA by reverse transcriptase (RT) PCR starting from total RNA. High quality RNA was isolated from mycelium of *L. edodes* grown on the poplar sawdust using the CTAB-based extraction method. RNA examined by electrophoresis on 0.8% agarose gel showed two bands corresponding to 18S and 25S rRNA (Figure 1A), that little or no RNA degradation occurred during the isolation. The A260/A280 and A260/A230 absorbance ratios were 1.9 and 2.0, respectively. This indicates low contamination by protein substances and secondary metabolites in the isolated RNA samples. In general, the RNA obtained was of high quality and integrity. Total RNA prepared as above was subjected to RT-PCR with the same primers. A single DNA band about 1551 bp in size was amplified (Figure 1B). The 1551 bp DNA fragment was ligated into pET11d vector and propagated in *E.coli*. The clones were sequenced from both directions.

Determination of the nucleotide sequence revealed complete coincidence with the nucleotide sequence of the *cel7A* gene *L. edodes* strain Stamets CS-2, that was published earlier (Lee et al., 2001).

The blastp results in Table 1 showed that the deduced amino acid sequence of 516 amino acids is identical with *L. edodes* cellobiohydrolase I amino acid sequence and showed high homology to other fungal CBHI enzymes belonging to glycosyl hydrolases family 7: 99% identity to cellulase of *Irpex lacteus*, 96% to cellobiohydrolase I of *Schizophyllum commune* and 95% to cellobiohydrolase I–II of *Volvariella volvacea*.

Results in this study show that amino acid sequence deduced from the nucleotide sequence of *cel7A* is homologous to other CBHs. However, until now the *cel7A* gene of *L. edodes* cellobiohydrolase has not been expressed in bacterial and yeast systems. Therefore, it was needed to be determined if it really codes a CBH.

In order to confirm that the *cel7A* gene encodes a cellobiohydrolase we used *E. coli* Rosetta (DE3) expression system. The synthesized by RT-PCR 1551bp fragment was digested with restriction endonucleases *NcoI* and *BamHI* and cloned in the pET11d plasmid with retention of the reading frame. The inserted fragment was tested for the absence of mutations by sequencing. The resulting

![Figure 1. A) Extraction of total RNA from *L. edodes* N127, B) PCR amplification of *cel7A* gene. M: marker (bp), Lane1: Genomic DNA as template, Lane2: total RNA as template.](image-url)
plasmid pET11d/cel7A was then used to transform *E. coli* Rosetta (DE3).

**Table 1. The significant matches of blastp analysis for cDNA of *L.edodes* N127**

<table>
<thead>
<tr>
<th>Genbank accession</th>
<th>Organism and gene</th>
<th>Score</th>
<th>E value</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAK95563.1</td>
<td>Lentinula edodes cellulase CEL7A mRNA</td>
<td>1043</td>
<td>1043</td>
<td>100%</td>
</tr>
<tr>
<td>BAA76365.1</td>
<td>Irpex lacteus cellulase mRNA</td>
<td>639</td>
<td>639</td>
<td>99%</td>
</tr>
<tr>
<td>AAX55505.1</td>
<td>Schizophyllum commune H4-8 glycoside hydrolase family 7 protein</td>
<td>586</td>
<td>625</td>
<td>96%</td>
</tr>
<tr>
<td>AAT64007.1</td>
<td>Volvariella volvacea cellobiohydrolase I-II mRNA</td>
<td>647</td>
<td>681</td>
<td>95%</td>
</tr>
</tbody>
</table>

The transformation of competent *E.coli* cells with the ligation products of the pET11d vector and the *L. edodes* cel7A gene amplicon yielded over 47 colonies, among which 4 individual clones were selected. All the selected colonies were assayed by PCR and restriction analysis for the presence of the recombinant plasmids carrying the cel7A gene. Plasmid DNAs were isolated and purified according to the protocol of High Pure Plasmid Isolation Kit. Plasmid DNAs were digested by NcoI and BamHI, followed by agarose gel electrophoresis. From Figure 2 we can see that plasmid DNAs was cut into two fragments, one about 1551 bp corresponding to cel7A gene, and the other about 5674 bp corresponding to the vector. We amplified the cel7A gene by PCR, using transformant plasmid DNA as a template and gene specific primers. The fragment detected by agarose gel electrophoresis corresponded in length to the cloned cel7A gene (Figure 3). Detection of the PCR product corresponding in length to the cloned cel7A gene sequence indicated that the analyzed colonies contained plasmids carrying the relevant gene. As a result of screening, we identified four colonies carrying the recombinant plasmids.
Synthesis of CEL7A protein was assessed in recombinant E. coli. Cell lysate of an E. coli expressing the cel7A gene was analyzed by SDS-PAGE and Western blotting using polyclonal anti-CEL7A antibodies.

After induction with IPTG for 4-12 hours the cells were lysed and the protein samples were prepared for SDS-PAGE by boiling in 2x sample buffer. 53.5kDa and 49 kDa protein bands from crude extracts of cells harboring the plasmid pET11b/cel7A were shown on SDS-PAGE (Figure 4A), whereas no similar protein bands from crude extracts of cells harboring the plasmid pET11d were detected (Figure 4A). In Figure 4B, the SDS-PAGE gel was transferred to a PVDF membrane and probed with anti-CEL7A polyclonal antibody.

Western blot analysis revealed a major protein band of 53.5-kDa specific to CEL7A in the crude extracts of the cells harboring the plasmid pET11d/cel7A, but not pET11d (Figure 4B). To validate the identity of these proteins to CEL7A, the polypeptides were excised from the gel, subjected to trypsin digestion and MALDI-TOF-MS analysis. MASCOT search results revealed a top score of 659 for CBH1, where probability based mowse score > 44 are significant and indicate identity or extensive homology (p<0.05).

To compare the amino acid sequence of the putative recombinant protein examined by MALDI-TOF with those of other CBH, the NCBI BLAST database was searched. The recombinant protein showed high homology with enzymes belonging to glycoside hydrolase family 7A (data not shown).

The effect of different pHs on CEL7A activity was determined with pNPL as a substrate. The crude extract of E. coli harboring pET11d/cel7A after 12 h induction with IPTG was incubated at 50 °C for 1 h at different pHs (sodium acetate buffer pH 4-6, sodium phosphate buffer pH 6-7 and glycine buffer pH 9). The result (Figure 5A) showed that the recombinant enzyme had an optimal activity at pH 7. The temperature profile of this enzyme was determined by incubating the reactions in 0.05 M sodium phosphate buffer (pH 7.0) for 1 h at different temperatures (30–80 °C). The optimal temperature of the enzyme was 50 °C (Figure 5 B).

**Figure 4.** Expression of L. edodes CEL7A in E. coli. A) SDS-PAGE of crude extracts of E. coli Rosetta (DE3) M: Marker (kDa); 1: Crude extract of E. coli harboring control plasmid pET11d; 2: Crude extract of E. coli harboring pET11d/cel7A after 4 h induction with IPTG; 3: Crude extract of E. coli harboring pET11d/cel7A after 12 h induction with IPTG. B) Western blotting of crude extracts of E. coli Rosetta (DE3) M:Marker (kDa); 1:Crude extract of E. coli harboring control plasmid pET11d; 2: Crude extract of E. coli harboring pET11d/cel7A after 12 h induction with IPTG
Expression of *L. edodes* cel7A gene in *E. coli*

**Figure 5.** Effects of pH and temperature on the activities of recombinant CEL7A. A) pH profile was determined by incubating the enzyme at 50\(^\circ\)C for 1 h at varying pHs (sodium acetate buffer pH 4-6, sodium phosphate buffer pH 6-7 and glycine buffer pH 9). B) Temperature profile was determined by incubating the enzyme in 0.05M sodium phosphate buffer pH 7 for 1 h at different temperatures.

Crude extract of an *E. coli* expressing the *cel7A* gene was examined for its ability to hydrolyze various cellulosic substrates at pH 7 and 50\(^\circ\)C (Table 2). It could hydrolyze Avicel, Filter paper, pNP-Lac and pNP-Cell, but activity towards carboxymethylcellulose (CMC) was much lower compared to Avicel and Filter paper. These results strongly indicate that CEL7A has cellobiohydrolase activity.

**Table 2.** Activities of the recombinant enzyme toward different cellulosic substrates at pH 7.0 and 50\(^\circ\)C

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Activity (U/mg total protein)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avicel</td>
<td>21.4 ± 1.4</td>
</tr>
<tr>
<td>Filter paper</td>
<td>23.2 ± 0.3</td>
</tr>
<tr>
<td>CMC</td>
<td>2.9 ± 0.5</td>
</tr>
<tr>
<td>pNP-Lac</td>
<td>4.1 ± 0.08</td>
</tr>
<tr>
<td>pNP-Cell</td>
<td>3.8 ± 0.1</td>
</tr>
</tbody>
</table>
Discussion

CBH enzymes are key components in fungal cellulase systems, and their functional activity is critical for consolidated bioprocessing for bioethanol production. For example, CBHs make up to 80% of the total mass for the *T. reesei* system, and CBH plays a particularly important role, making up 60% of the total mass (van Zyl et al., 2007). A number of cellulase genes from bacteria and fungus have been cloned and expressed in yeast or *E. coli* (Qiao et al., 2004; Hong et al., 2001, Hong et al., 2003). However fungal and bacterial endoglucanase (EG) production in recombinant strains was more successful than CBH production (van Zyl et al., 2007). This is not surprising considering that EG enzymes usually have specific activities of 2 to 3 orders higher in magnitude on synthetic and amorphous cellulose substrates, such as carboxymethylcellulose (CMC), in comparison to CBHs. Thus it is easier to measure the presence of even small amounts of heterologous EG compared to CBHs.

Expression of fungal cellulases in *E. coli* as non-glycosylated forms has been attempted with limited success. Only a few cellulose genes, such as *cbhl* and *egl3* of the well-studied fungus *T. reesei*, were expressed in *E. coli*, but the productivity of cellulases was very low (Ekino et al., 1999).

In this study, we have successfully expressed the *cel7A* gene of *L. edodes* in *E. coli* under the control of the T7 promoter. SDS–PAGE and Western blotting analysis showed that CEL7A constitutes a major protein produced in *E. coli*, but the productivity of cellulases was very low (Ekino et al., 1999).

As shown in Figure 4A plasmid vector pET11d/cel7A expresses two additional bands which in SDS-PAGE migrate as 53 kDa and 49 kDa proteins. Since the difference in molecular weight between these two proteins is approximately 5 kDa we may suggest that the CEL7A protein was partially degraded possibly during cell lysis. The cel7A enzyme contains 430 aa catalytic domain and 57 aa C-terminal cellulose-binding domain (CBD) and serine/threonine-rich linker sequence (Lee et al., 2001). We may speculate that the proteolytic degradation of 57 aa C-terminal CBD and linker sequence results in truncated 49 kDa protein.

Indeed, western blot analysis confirms the presence of low molecular weight CEL7A proteins which migrate below the full-length recombinant protein in SDS-PAGE suggesting proteolytic degradation of CEL7A during expression and/or cell lysis. Furthermore we hypothesize that the linker region of *E. coli*-expressed CEL7A might be more susceptible to proteolysis as compared to glycosylated form of this protein expressed in yeast cells (unpublished observation). It seems likely that glycosylated linker region in native CEL7A of *L. edodes* protects the enzyme from proteolytic degradation nevertheless further investigations are required.

The substrate specificities indicated that CEL7A contains cellbiohydrolase activity since it could hydrolyze crystalline cellulose (Avicel, Filter paper). Activity of cloned enzyme toward CMC was much lower than to Avicel and filter paper. These data are consistent with the established results that fungal CBH proteins have very limited action on substituted cellulose such as CMC or hydroxyethylcellulose (Basiria and Mishra, 1989; Kanokratana et al., 2002). Since most of the fungal cellulases are glycoproteins, the modification of the native protein by glycosylation has been reported to play an important role in synthesis, secretion, and stability of extracellular cellulases (Basiria and Mishra, 1989). Here we have demonstrated that CEL7A synthesized in *E. coli* without glycosylation has an enzymatical activity. This indicates that glycosylation is not necessary for enzymatic activity of CEL7A.

In summary, a gene encoding CEL7A was successfully isolated from *L. edodes* strain N127 using RT-PCR technique. The determined nucleotide sequence is completely corresponded to the nucleotide sequence of the *cel7A* gene of *L. edodes* strain Stamets CS-2 which was described previously (Lee et al., 2001). The enzyme was successfully expressed in *E. coli* in an active form. Enzymatic properties of CEL7A were also determined. The optimal temperature for enzyme activity was 50°C and the optimal pH was 7.0.

Acknowledgement

This work was supported by grant from National Centre for Biotechnology of the Republic of Kazakhstan.
Expression of *L. edodes* cel7A gene in *E. coli*

References


