Cloning and Transcriptional Analysis of Exocellulase 1 Gene from Irpex lacteus

Journal of Tokyo University of Fisheries, Vol. 87, pp. 39–44, 2001

Cloning and Transcriptional Analysis of Exocellulase I Gene from *Irpex lacteus*

Naoko Hamada*[1] and Ken Hirohashi*[2]

(Received August 31, 2000)

Abstract: The *cex1* gene encoding exocellulase I (Ex-1) of the white rot fungi, *Irpex lacteus* DSM 1183, was cloned and sequenced. The *cex1* gene consisted of an open reading frame of 1,569 nucleotides encoding a protein of 523 amino acids which was interrupted by two putative introns. The deduced amino acid sequence showed that Ex-1 consisted of a catalytic domain and a cellulose-binding domain (CBD) separated by a serine-rich linker region. The Ex-1 catalytic domain was found to belong to cellulase family C, and showed the highest similarity with cellobiohyrolase I (CBHI) from *Phanerochaete chrysosporium*. In addition to the TATA-like sequence and CAAT motifs, putative CREA binding sites were observed in 5′ upstream of the Ex-1 coding region. Northern hybridization analysis revealed that expression of the *cex1* gene was induced by various cellulosic substrates and repressed by glucose, fructose, and lactose.

***Key words*:** *Cloning*, *PCR*, *Transcriptional analysis*, *Cellulase*, *White rot fungi*, *Irpex lacteus*

Introduction

The enzymatic hydrolysis of cellulose takes place under the action of a cellulase complex containing at least three types of glucanases, namely, endo-type cellulase [1,4-β-D-glucan 4-glucanohydrolase, EC 3.2.1.4], exo-type cellulase [1,4-β-D-glucan cellobiohydrolase, EC 3.2.1.91] and β-glucosidase [or cellobiase, EC 3.2.1.21]. These enzymes together comprise a system to convert cellulose to glucose efficiently.

In plants, cellulose is usually found in association with both lignin and hemicelluloses as a component of lignocellulose. Unlike *Trichoderma reesei*, the best studied fungal cellulolytic organism, the white rot fungi, e.g., *Phanerochaete chrysosporium* and *Irpex lacteus* are able to degrade all of these components of lignocellulose. This explains why these organisms, *P. chrysosporium* and *I. lacteus* are expected to be used for bioremediation. In fact, *P. chrysosporium* is able to mineralize, at least partially and sometimes completely, a variety of persistent environmental pollutants. This is achieved through the action of a large number of extracellular enzymes, including two distinct classes of peroxidases believed to degrade lignin, a number of xylanases and several cellulases. Genes coding for both types of peroxidases and two cellulases (cellobiohydrolase I and cellobiohydrolase II) from *P. chrysosporium* have been described[1,2]. On the other hand, only few studies on *I. lacteus* were reported and the possibility of bioremediation using this organism has not been investigated. The elucidation of the properties of extracellular enzymes and the genes coding them will contribute toward practical application of this organism in bioremediation.

In the cellulase system of *I. lacteus*, two types of cellulases, exocellulase I (Ex-1), endocellulase I (En-1) and two xylanases (Xylanase I and Xylanase III) were purified and their properties were characterized[3–5]. Recently, we have isolated and characterized three exocellulase genes, *cel1*[6], *cel2*[7] and *cel3*[8], from *I. lacteus* MC-2. Here we describe the cloning of a gene encoding a major exocellulase (Ex-1) of *I. lacteus* DSM 1183, a different strain of the same organism for the purpose of considering the diversity of the exocellulase gene. We also report on the analysis of Ex-1 transcription induced by various cellulosic substrates.

Materials and Methods

Substrates

Carboxymethyl cellulose (CMC) was purchased from Wako Pure Chemical Industries, Ltd. Avicel, a microcrystalline cellulose powder (Art. 2331), was purchased from E. Merck (Dermstadt, Germany). Cotton was provided by Daiwa Shizai (Tokyo). Phosphoric acid-swollen cotton was prepared as described by Tanaka *et al*[9]. Bacterial cellulose powder was donated by BPR Co., Ltd.

*1 Department of Food Science and Technology, Tokyo University of Fisheries, 5–7, Konan 4-chome, Minato-ku, Tokyo 108–8477, Japan.

*2 Biomate Co., Ltd., 7–20, Higashikasai 6-chome, Edogawa-ku, Tokyo 134–0084, Japan.

Organism and Culture Conditions

I. lacteus DSM 1183 was grown at 25°C with rotary shaking at 170 rpm, in 100 m*l* of modified Mandel's medium[10] (1% cellulosic substrate, 0.1% polypeptone, 0.14% NH_4SO_4, 0.2% KH_2PO_4, 0.03% urea, 0.03% $MgSO_4 \cdot 7H_2O$, 0.03% $CaCl_2$, 0.1% (v/v) Tween 80, 0.05% $FeSO_4 \cdot 7H_2O$, 0.016% $MnSO_{(4-6)}$ H_2O, 0.014% $ZnSO_4$ and 0.02% $CoCl_2$; pH 6.0).

Peptide Sequencing

Ex-1 extracted from *I. lacteus* was purified by successive column chromatographies using Bio Gel P-100 (Bio Rad Laboratories, CA, USA), DEAE Sephadex A-50 (Pharmacia Biotech Inc., Uppasala, Sweden), DEAE Sepharose CL-6B (Pharmacia Biotech Inc.) and Butyl-Toyopearl 650 M (Tosoh Co., Tokyo). The protein sample was cleaved with V8 protease (Wako Pure Chemical Industries Ltd.). An aliquot (500 μl) of 1 mg/m*l* Ex-1 dissolved in 50 mM Na_2HPO_4-citric acid buffer (pH 7.8) was mixed with 10 μl of 1 mg/m*l* V8 protease solution. After incubation at 37°C for 30 min, 20 μl of the reaction mixture was added to 10 μl of 1 M Na_2HPO_4-citric acid buffer (pH 3.0) to terminate the digestion. Subsequently, the generated polypeptides were subjected to SDS-PAGE, and the separated polypeptides were electroblotted onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a semi-dry blotting apparatus (Bio-Rad). The N-terminal amino acid sequence of peptide fragments obtained by V8 protease digestion was determined using a protein sequencer (Shimadzu PSQ-1, Shimadzu Co., Ltd., Kyoto).

PCR Cloning

I. lacteus cells were grown in Avicel medium for 5 days and then total RNAs were extracted with a RNeasy Plant Mini kit (Qiagen K.K, Tokyo). A reverse transcription-mediated PCR (RT-PCR) was performed to amplify the first-strand cDNA that was synthesized from total RNA (5 μg) using *NotI* $d(T)_{18}$ primer (10 pmol). An aliquot (1/32) of the mixture was used as a template for PCR with 300 pmol of degenerate primers synthesized based on the amino acid sequences. The reaction was performed through 3 cycles of 1 min at 95°C, 1 min at 72°C, 1 min at 37°C, 35 cycles of 1 min at 95°C, 1 min at 72°C, 1 min at 50°C, and 5 min at 72°C. PCR products were separated by 3% polyacrylamide gel electrophoresis. The DNA fragment was extracted and subcloned in pUC119 (Takara Shuzo Co., Ltd., Otsu).

Construction and Screening of Genomic DNA Library

Chromosomal DNA was prepared by the method of Garber and Yoder[11] and partially digested with *Sau* 3AI. The DNA fragments were fractionated by sucrose density-gradient centrifugation, and fragments of about 9–23 kb were recovered. The resulting DNA fragments were ligated to the dephosphorylated *Bam*H I site of the λ DASH II vector (STRATAGENE, Heidelberg, Germany). The packaged recombinant phages were used to infect *Escherichia coli* XL1-Blue MRF to construct a genomic library. The genomic library was screened by plaque hybridization using the ^{32}P-labeled PCR fragment as a probe. Hybridization was performed at 42°C for 16 h in a solution containing 50% (v/v) formamide, 2×SSC, 10% dextran sulfate, 1% SDS, 50% formamide and 100 μg/m*l* denatured fragmented salmon sperm DNA. The membranes were washed with 2×SSC–0.1% SDS at room temperature for 10 min. After autoradiography, positive plaques were isolated.

DNA Sequencing and Sequence Analysis

DNA sequencing of both strands was performed by the dideoxy nucleotide chain-termination method using a dye terminator DNA sequencing kit (Perkin Elmer, Foster, CA, USA) with an autosequencer (ABI: model 373 A DNA Sequencing system). Data was analyzed using GENETYX (version 9.0) software. Homology searches were performed using the BLAST program.

Northern Hybridization Analysis

Total RNAs were fractionated by 1.2% agarose gel electrophoresis in the presence of formaldehyde, and transferred to a Gene Screen Plus (Dupont, Boston, MA, USA) membrane. The membrane was prehybridized with 6×SSC, 5×Denhardt's solution, 0.1% SDS, and 50% formamide for 1 h at 42°C. Then, a ^{32}P-labeled probe, synthesized using a Multiprime DNA labeling kit (Amersham, Pharmacia Biotech Inc.), and denatured salmon sperm DNA were added, followed by hybridization at 42°C for 16 h. The membrane was washed with 2×SSC–1% SDS at 42°C for 10 min and then with 0.2×SSC–1% SDS at 42°C for 10 min.

Results and Discussion

Determination of Amino Acid Sequence

We attempted to determine the amino terminal region of the Ex-1 protein. From the preliminary sequencing, the amino terminus of the purified protein seemed to be modified. Therefore, the protein sample was cleaved with V8 protease, obtaining four major fragments [V8 I (31 kDa), V8 II (27 kDa), V8 III (24 kDa) and V8 IV (22 kDa)]. The N-terminal sequences of V8 I and V8 III fragments originated from Ex-1 were identical. The amino terminus of the V8 II fragment seemed to be

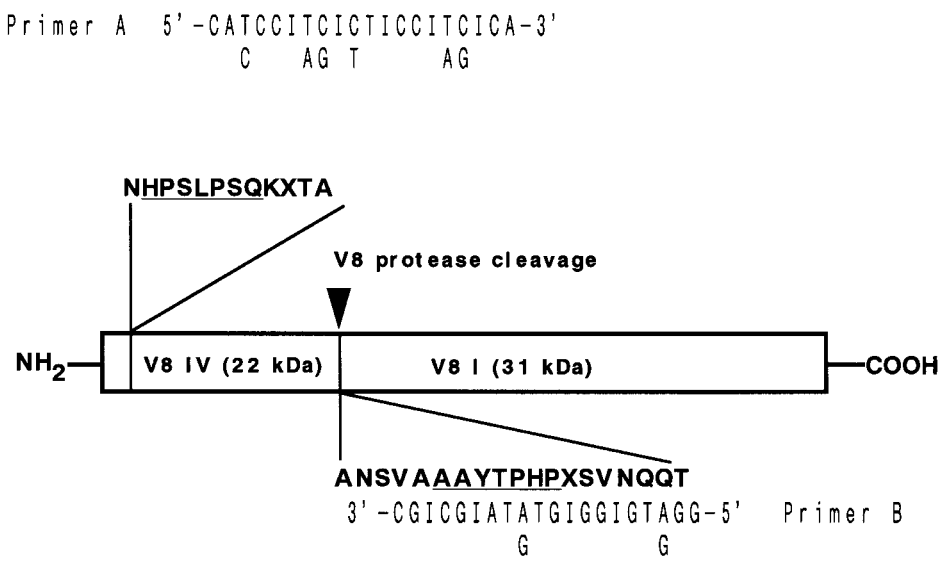

Figure 1. Partial amino acid sequences of Ex-1 and primer design based on the predicted location of two V8 fragments. A set of primer was designed based on the underlined amino acid sequences. The letter I in the primer sequence denotes inosine.

blocked. The amino terminal amino acid sequence of V8 I fragment (therefore, that of V8 III fragment also) was found to be ANSVAAAYTPHPXSVNQQT and that of V8 IV was NHPSLPSQKXTA (X: not identified).

Cloning of Cex Gene

Based on the determined amino acid sequences, two degenerate primers (A and B) were synthesized (Fig. 1). As a result of RT-PCR, a specific product of 600 bp in size was obtained. This 600 bp amplified fragment was used as a hybridization probe for genomic DNA library screening, and the nucleotide sequences of the isolated genomic DNA fragments were determined.

The determined nucleotide sequence contains a single open reading frame (ORF) (Fig. 2). The amino acid sequence deduced from this ORF includes both amino acid sequences determined from the purified Ex-1 protein. Ex-1 ORF consisted of 1,569 bp and encoded a polypetide of 523 amino acids. Two putative introns (56 bp and 61 bp) which interrupted the coding region were present. These two introns contained the 5′ and 3′ border sequences and the putative lariat sequences which were typical for introns in genes of filamentous fungi. Moreover, the two introns occurred at the same position as those of CBHI from *P. chrysosporium*[1]. Comparison of the deduced amino acid sequence of Ex-1 with other homologous fungal cellulases suggested that the precursor of Ex-1 contained the signal sequence of 18 amino acids in its amino terminal. The secretory precursor was thought to be processed at a specific cleavage site between Gly 18 and Gln 19 residues, resulting in the formation of a mature enzyme composed of 505 amino acids with a molecular weight of 52,640. Considering the sugar content of Ex-1 (2.0%), this calculated value was in good agreement with the experimental value of 53,000 determined by SDS-PAGE.

Comparison of the Amino Acid Sequence of Ex-1 with Other Fungal Cellulases

Comparison of the amino acid sequence of Ex-1 with available sequences in the SWISS-PROT databases revealed that the sequence of Ex-1 had high homology with those of *P. chrysosporium* CBHI (sequence identity: 70%)[1], *Agaricus bisporus* CEL2 (60%)[12], *A. aculeatus* CBHI (55%)[13] and *T. reesei* CBHI (51%)[14]. All of these cellulases belong to cellulase family C of the glycosyl hydrolases, indicating that *I. lacteus* Ex-1 is also a member of cellulase family C[15]. The putative two glutamic acid residues, considering the catalytic center, and twenty-two cysteine residues characteristic of CBHI-type enzymes were conserved perfectly in *I. lacteus* Ex-1. The fact that these residues were conserved in *I. lacteus* Ex-1 also supported the finding that *I. lacteus* Ex-1 was one of the CBHI-type cellulases. The C-terminal domain of *I. lacteus* Ex-1 showed a high similarity with CBDs of family I. The aromatic moieties of three conserved Tyr side chains (Tyr 5, Tyr 31, and Tyr 32 in *T. reesei* (CBHI-CBD) and the side chains of three amino acids (Gly 7,

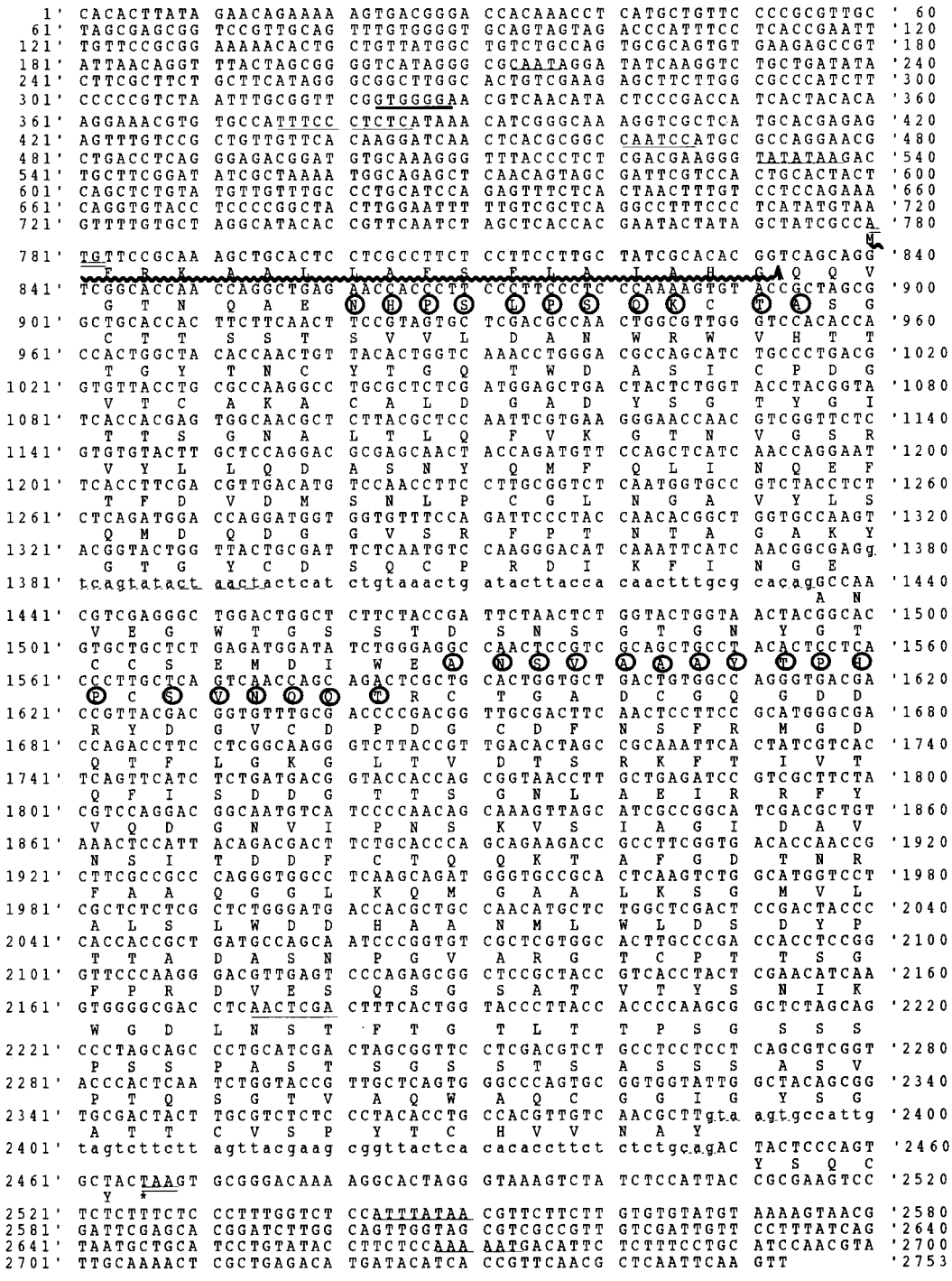

Figure 2. Nucleotide sequence of the *cex1* gene and deduced amino acid sequence of Ex-1. ATG at position +780 is the start codon of the *cex1* gene. The deduced amino acid sequence is represented by one-letter amino acid symbols shown below the DNA sequence. A predicted TATA box, CAAT motofs, a C/T-rich region and the sequences homologous to polyadenylation signal are shown by underlines. The sequence homologous to the *A. nidulans* CREA binding consensus site is shown by a bold underline. The putative recognition site of the signal peptidase is shown by an arrowhead in the amino acid sequence. Wavy underline in the amino acid sequence indicated the putative signal sequences. Putative *N* glycosylation site (Asn-Xaa-ser/Thr) is boxed. Two putative introns are shown by lower-case letters. The consensus sequences typical of introns in genes of filamentous fungi are indicated by dotted underlines. The amino acids identified by Edman degradation are circled.

Asn 29, and Gln 34 in *T. reesei* (CBHI-CBD) were also strictly conserved except for the replacement (Tyr 5→Trp 5) in the CBD of *I. lacteus* Ex-1. One putative *N*-glycosylation site, Asn^{446}-Thr-Thr, was detected at the C terminus of the catalytic domain. The deduced amino acid sequence showed that Ex-1 had a modular structure consisting of a catalytic domain of 449 aa and a C-terminal 36-aa cellulose-binding domain (CBD) separated by a serine-rich linker region of 38 aa.

Regulation of Cex1 Gene Expression

To gain insight into the regulation of *cex1* gene expression in *I. lacteus*, transcriptional analysis was carried out by northern hybridization. The *Sac*I-*Pst*I fragment was used for northern hybridization. To confirm that the loading of each RNA sample was equal, the membrane was rehybridized with a *T. reesei* actin probe, pTrA[16]. When various cellulosic substrates that differed in their degree of crystallinity were added to *I. lacteus* cultures grown on the medium containing glucose, *cex1* mRNA was detected after 3 days and accumulated thereafter. The amount of *cex1* transcrips was almost equal among cultures using various cellulosic substrates (Fig. 3A lanes 1 to 5). *Cex1* mRNA was not detected in the control medium without any cellulosic substrates (Fig. 3A lane 6). The effect of other carbon sources was investigated, and the results are shown in Fig. 3B. For this purpose, various sugars were added to Avicel-grown *I. lacteus* cultured for 5 days. After 40 h of cultivation, the mycelia were harvested for RNA analysis. *Cex1* mRNA was undetectable after addition of fructose, glucose and lactose, whereas glycerol and mannitol did not affect *cex1* mRNA accumulation.

Analysis of the 5′-noncoding region reveals the presence of putative regulatory elements in the expression of the *cex1* gene. In addition to the TATA-like sequence, CT-rich sequence and CAAT motifs, the 5′-(G/C)-PyGGGG-3′ sequence identified as a carbon catabolite repressor binding site in the *Aspergillus nidulans* CREA[17] repressors was found. Northern blot analysis indicated that *cex1* was differentially regulated by carbon sources at the transcription level, which would be expected to contain a catabolite-responsive site recognized by transcription elements in the promoter region[17]. A large amount of *I. lacteus* Ex-1 produced indicated that it was expressed under the control of a strong promoter. In the *I. lacteus cex1* gene, a CREA-binding consensus sequence was sandwiched between two CAAT boxes. The same sandwiched sequence was also found in the *Aspergillus aculeatus cbhI* gene, coding for a major cellobiohydrolase, CBHI[13].

Northern blot analysis in this study indicated that two

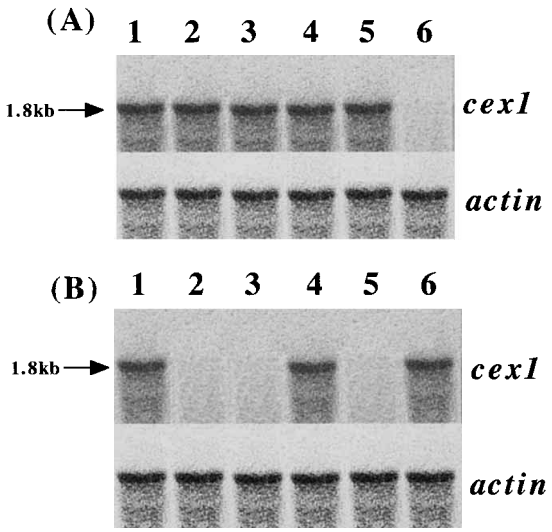

Figure 3. Northern analysis of *cex1* gene transcript.
(A) Mycelium grown for 3 days in glucose medium was aseptically washed with distilled water and transferred to various cellulosic medium. Mycelium was harvested after 5 days and analyzed for the presence of *cex1* transcripts. After hydridization with the *Sac*I-*Pst*I *cex1* fragment as a probe, the filters were rehybridized with a *Trichoderma reesei* actin probe. Lane 1: CMC, lane 2: Avicel, lane 3: Bacterial cellulose, lane 4: cotton, lane 5: phosphoric acid-swollen cotton. lane 6: control (without cellulosic substrates).
(B) Various carbon sources were added in 5 days Avicel-grown *I. lacteus* cultures to a final concentration of 1% (w/v). Lane 1: control (Avicel medium), lane 2: fructose, lane 3: glucose, lane 4: glycerol, lane 5: lactose, lane 6: mannitol. Mycelia were harvested 40 h after addition of various carbon sources and RNAs were analyzed by northern hybridizaion.

different control mechanisms regulated *cex1* gene expression: induction by cellulose and repression by glucose, fructose or lactose. Glycerol and mannitol were carbon sources which did not participate in catabolite repression. These two mechanisms were functional in another exocellulase gene from *I. lacteus* MC-2[6–8]. These results suggest that cellulase synthesis in *I. lacteus* is controlled at the transcription level.

References

1) P. Sims, C. James, and P. Brada: Identification, molecular cloning and characterization of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exocellobiohydrolase I gene from *Trichoderma reesei. Gene*, 74, 411–422 (1988).
2) C. A. M. Tempelaars, P. R. J. Birch, P. F. G. Sims, and P.

Brada: Isolation, characterization, and analysis of the expression of the *cbhII* gene of *Phanerochaete chrysosporium. Appl. Environ. Microbiol*., 60, 4387–4393 (1994).
3) T. Kanda, S. Nakakubo, K. Wakabayashi, and K. Nisizawa: Purification and properties of an exo-cellulase of avicelase type from a wood-rotting fungus, *Irpex lacteus* (*Polyporus tulipiferae*). *J. Biochem*., 84, 1217–1226 (1978).
4) T. Kanda, K. Wakabayashi, and K. Nisizawa: Purification and properties of a lower-molecular-weight endo-cellulase from *Irpex lacteus* (*Polyporus tulipiferae*). *J. Biochem.*, 87, 1625–1634 (1980).
5) T. Kanda, Y. Amano, and K. Nisizawa: Purification and properties of two endo-1,4-β-xylanases from *Irpex lacteus* (*Polyporus tulipiferae*). *J. Biochem.*, 98, 1545–1554 (1985).
6) N. Hamada, R. Okumura, N. Fuse, R. Kodaira, M. Shimosaka, T. Kanda, and M. Okazaki: Isolation and transcriptional analysis of a cellulase gene (*cel1*) from the basidiomycete *Irpex lacteus. J. Biosci. Bioeng*., 87, 97–102 (1999).
7) N. Hamada, K. Ishikawa, N. Fuse, R. Kodaira, M. Shimosaka, Y. Amano, T. Kanda, and M. Okazaki: Purification, characterization and gene analysis of exo-cellulase II (Ex-2) from the basidiomycete *Irpex lacteus. J. Biosci. Bioeng.*, 87, 432–441 (1999).
8) N. Hamada, N. Fuse, M. Shimosaka, R. Kodaira, Y. Amano, T. Kanda, and M. Okazaki: Cloning and characterization of a new exo-cellulase gene, *cel3*, in *Irpex lacteus.* FEMS. *Microbiol. Lett.*, 172, 231–237 (1999).
9) M. Tanaka, M. Taniguchi, T. Morita, R. Matsuno, and T. Kamikubo: Effect of chemical treatment on solubilization of crystalline cellulose and cellulosic wastes with P*ellicularia filamentosa* cellulase. *J. Ferment. Technol*., 57, 186–190 (1979).
10) M. Mandel and E. T. Reesei: Induction of cellulase in *Trichoderma viride* as influenced by carbon source and metals. *J. Bacteriol*., 73, 269–278 (1957).
11) R. C. Garbe and YO. C. oder: Isolation of DNA from filamentous fungi and separation into nuclear, mitochondrial, ribosomal, and plasmid components. *Anal. Biochem.*, 135, 416–422 (1983).
12) E. Yagüe, M. Mehak-Zunic, L. Morgan, D. A. Wood, and C. F. Thurston: Expression of CEL2 and CEL4, two proteins from *Agaricus bisporus* with similarity to fungal cellobiohydrolase I and β-mannanase, respectively, is regulated by the carbon source. *Microbiology*, 143, 239–244 (1997).
13) G. Takada, T. Kawaguchi, J. Sumitani, and M. Arai: Cloning, nucleotide sequence, and transcriptional analysis of *Aspergillus aculeatus* No. F-50 cellobiohydrolase I (*cbhI*) gene. *J. Ferment. Bioeng*., 85, 1–9 (1998).
14) S. Shoemaker, V. Schweickart, M. Landner, D. Gelfand, S. Kwok, K. Myambo, and M. A. Innis: Molecular cloning of exo-hydrolases I derived from *Trichoderma reesei* strain L27. *Bio/Technol*., 1, 691–696 (1983).
15) B. Henrissat, and A. Bairoch: New families in the classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J*., 293, 781–788 (1993).
16) E. Jr. Matheucci, F. Henrique-Silva, S. El-Gogary, C. H. Rossini, A. Leite, J. E. Vera, J. C. Urioste, O. Crivellaro, and H. El-Dorry: Structure, organization and promoter expression of the actin-encoding gene in *Trichoderma reesei*. *Gene*, 161, 103–106 (1995).
17) P. Kulmburg, M. Mathiew, C. Dowzer, J. Kelly, and B. Felenbok: Specific binding sites in the *alcR* and *alcA* promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans. Mol. Microbiol*., 7, 847–857 (1993).

Irpex lacteus 由来のエキソ型セルラーゼ遺伝子のクローニングと転写解析

濱田奈保子[*1]・広橋　憲[*2]
([*1]東京水産大学大学院水産学研究科, [*2]バイオメイト(株))

白色腐朽菌 *Irpex lacteus* のエキソ型セルラーゼ I (Ex-1)**をコードする遺伝子 *cex1* を** *I. lacteus* DSM 1183 **株より単離し，その塩基配列を決定した。*cex1* 遺伝子は** 2 **つのイントロンと推定される配列によって分断される** 1,569 bp **の** ORF **よりなり，**523 **アミノ酸をコードしていた。アミノ酸配列より推定された** Ex-1 **タンパクの構造は触媒ドメインとセルロース吸着ドメインをセリン残基に富むリンカーが連結した構造であった。その触媒ドメインはセルラーゼファミリー** C **に属しており，*Phanerochaete chrysosporium* 由来のセロビオヒドラーゼ** I **と最も高い相同性が見られた。コード領域の** 5′ **上流には** TATA **様配列，**CAAT **配列，さらに** CREA **結合サイトが存在していた。ノーザン分析の結果から *cex1* 遺伝子の発現は種々のセルロース基質によって誘導され，グルコース，フルクトース，ラクトースによって抑制されることがわかった。**

キーワード：クローニング, PCR, **転写解析**, **セルラーゼ，白色腐朽菌，** *Irpex lacteus*