

Cloning and characterization of *cbhII* gene from *Trichoderma parceramosum* and its expression in *Pichia pastoris*

Saber Zahri, Mohammad Reza Zamani*, Mostafa Motallebi, Mehdi Sadeghi

National Institute for Genetic Engineering and Biotechnology (NIGEB), P.O. Box 14155-6343, Tehran, I.R. Iran

Abstract

The genomic and cDNA clones encoding cellobiohydrolase II (CBHII) have been isolated and sequenced from a native Iranian isolate of *Trichoderma parceramosum*, a high cellulolytic enzymes producer isolate. This represents the first report of *cbhII* gene from this organism. Comparison of genomic and cDNA sequences indicates this gene contains three short introns and also an open reading frame coding for a protein of 470 amino acids. The deduced amino acid sequence includes a putative signal peptide, cellulose binding domain, linker region, and catalytic domain. Homology between this sequence and other reported *Trichoderma* CBHII proteins and also structural prediction of this enzyme are discussed. The coding sequence of *cbhII* gene was cloned in pPIC9 expression vector and expressed in *Pichia pastoris* GS115. The expression was confirmed by Northern dot blot, RT-PCR and enzyme activity staining.

Keywords: *Trichoderma parceramosum*; *Cellobiohydrolase II*; Fungus; cDNA Sequence

INTRODUCTION

Much of the organic material on the surface of the earth is stored in the form of cellulose. Cellulose is a simple, linear polymer built up from glycosyl units connected by $\beta(1-4)$ linkages. These linear chains can vary in length, and often consist of many thousands of units (Beguin and Aubert, 1994). Within the biosphere,

there is an enormous of cellulose through large-scale production and degradation. Though chemically simple, cellulose is physically complex with both crystalline and amorphous regions, and a number of different enzymes are required for its efficient hydrolysis (Tomme *et al.*, 1988). Cellulose is degraded in nature by the concerted action of several synergistically functioning enzymes. Depending on their mode of action, cellulolytic enzymes fall into one of two main groups, endoglucanase or cellobiohydrolase (Beguin and Aubert, 1994). The complete degradation of cellulose to glucose requires the action of at least three types of enzymes: endo- β -1,4 -glucanase, exo- β -1,4 -glucanase (cellobiohydrolase) and β -glucosidase (Beguin and Lemaire, 1996). The most abundant enzymes are two cellobiohydrolase, Cel7A and Cel6A, also called CBHI and CBHII, respectively. These are also the most efficient enzymes on highly crystalline cellulose (Koivula *et al.*, 2002).

Sequence determination and functional studies have shown that most fungal cellulases have a bifunctional organization in which a catalytic core domain and a cellulose-binding domain are connected by a flexible glycosylated linker peptide (Van Tilbeurgh *et al.*, 1986; Tomme *et al.*, 1988). Unlike CBHI, CBHII prefers to act on the cellulose chain from the nonreducing end (Nutt *et al.*, 1998) and classified in the Cel6A family of the glycohydrolases (Henrissat *et al.*, 1996).

The methylotrophic yeast *Pichia pastoris* has been developed as a widely used host organism for recombinant protein production (Cereghino and Cereg,

*Correspondence to: Mohammad Reza Zamani, Ph.D.

Tel: +98 21 44580363; Fax: + 98 21 44580399

E-mail: mrzamani97@yahoo.com

2000). As a single cell microorganism, it grows rapidly on inexpensive substrates and are easily manipulated. As an eukaryote, it can often correctly fold and process other eukaryotic proteins. *P. pastoris* has several key advantages over other yeast expression systems (Higgins and Cregg, 1998). It has a tightly regulated and highly efficient methanol induced alcohol oxidase (AOX1) promoter. This promoter stimulates high levels of transcription in cells cultured on methanol but is strongly repressed in cells cultured on other carbon sources (Tschoop *et al.*, 1987; Cregg *et al.*, 1989). The other important advantage is the *P. pastoris* readily grows to unusually high-cell densities (130 g^l⁻¹ dry cell weight) in continuous culture (Cregg and Madden, 1987). This is possible because *P. pastoris* strongly favors a respiratory mode of metabolism over a fermentative mode (Lin Cereghino *et al.*, 2001, 2002). As a result, *P. pastoris* does not have a tendency to produce ethanol, an inhibitor of cell growth at high concentrations. Third advantage of *P. pastoris* for certain recombinant proteins are its capacity for secreting large amounts of these proteins into the culture medium. Since *P. pastoris* secretes only low levels of its own proteins, secretion provides an effective method to separate recombinant proteins from the bulk of intracellular host proteins and other cellular materials (Romanos, 1995). Finally, unlike the expression from plasmid, the heterologous genes of interest are stably integrated into *P. pastoris* genome via homologous recombination.

In this study, we have cloned and sequenced the *cbhII* gene (genomic DNA and cDNA) from a high producer cellulolytic enzymes isolate of *T. parceramosum* to characterize its gene structure for the first time. We described structural prediction of CBHII based on comparative protein modeling. We also report the heterologous expression of CBHII in active form by *P. pastoris*.

MATERIALS AND METHODS

Microorganism, plasmid vector, and culture conditions: *Trichoderma parceramosum* (Seyed Asli *et al.*, 2004) was identified in our laboratory. The stock culture was stored on agar (1.5%) slant of MY medium (2% malt extract, 0.2% yeast extract, 1% maltose)

For chromosomal DNA preparation, the spores were inoculated in the liquid MY medium and grown for 2 days at 30°C with shaking (100 rpm), and mycelia were harvested by filtration (Takashima *et al.*, 1998). *Escherichia coli* DH5 α as a host and pBluescript SK(+)(Stratagene, La Jolla, CA) as a vector used for routine cloning. *E. coli* were cultured in LB medium (1% trypton, 0.5% yeast extract, and 1% NaCl). *P. pastoris* GS115 (his4) and pPIC9 vector have been used for heterologous expression of CBHII (Invitrogen BV, Groningen, The Nederland).

P. pastoris GS115 was grown in YPD (1% yeast extract, 2% peptone, and 2% dextrose), YPG (1% yeast extract, 2% peptone, and 1% glycerol), YPM (1% yeast extract, 2% peptone, and 0.5% methanol), BMG (a buffered minimal glycerol consist of: 100 mM potassium phosphate, 1.34% YNB 4 \times 10⁻⁵ % biotin, and 1% glycerol), BMM (100 mM potassium phosphate, 1.34% YNB 4 \times 10⁻⁵ % biotin, and 0.5% methanol), BMGY (100 mM potassium phosphate, 1.34% YNB 4 \times 10⁻⁵ % biotin, 1% yeast extract and 1% glycerol) or BMMY(100 mM potassium phosphate, 1.34% YNB 4 \times 10⁻⁵ % biotin, 1% yeast extract and 0.5% methanol).

Genomic DNA PCR amplification: For amplification of cellobiohydrolase II (*cbhII*) gene from *T. parceramosum*, the two tailed primers, CF22 and CB22 (Table 1), were designed based on sequence similarity of existing *cbhII* genes present in the database. To facilitate subsequent cloning of the PCR-derived frag-

Table 1. Oligonucleotides (primers) used in this study.

Name	Oligonucleotides Sequence
CF22	5'-GCGGAATTCATGATTGTCGGCATTCTCACC-3'
CB22	5'-GCGGGATCCTTACAGGAACGATGGGTTTGC-3'
CP3	5'-CAGTTTGTGGTGTATGACTT-3'
CP4	5'-CTGGCTTGACCCAGACAAA-3'
C2EF	5'-GCGGAATTC AAGCTTGCTCAAGCGTCTGGG-3'
C2EB	5'-CCC GCGGCCGCTTACAGGAACGATGGGTTTGC-3'
18sU	5'-CTTCGGGATCGGAGTAATGATTAA-3'
18sR	5'-GCCCAGAACATCTAAGGGCATCACAGA-3'
5' OX1	5'- GACTGGTTCCAATTGACAAGC-3'
3' AOX1	5'- GCAAATGGCATTCTGACATCC-3'

ments, *EcoRI* and *BamHI* restriction sites (bolded) were added to the 5'-end of CF22 and CB22 primers, respectively.

Fungal chromosomal DNA was prepared as described by Bao Sun *et al.* (2002). Amplification of the DNA fragments encoding *T. parceramosum cbhII* gene was performed using polymerase chain reaction (PCR). To identify and confirm the PCR products, two primers, CP3 and CP4 (Chen *et al.*, 1987), were synthesized and used for nested PCR. PCR reactions contained 2.5 units of *Pfu* DNA polymerase (Fermentas), 1X buffer, 200 μ M of each deoxynucleotide triphosphate, 2 μ M $MgSO_4$ and 0.5 μ M primers. Reaction conditions for PCR amplification were 94°C for 1min, 54°C for 1 min, and 72°C for 1.5 min, for 30 cycles followed by a final extension of 10 min. PCR products were separated by electrophoresis on a 1% agarose gel and purified by High pure PCR product purification Kit (Roche). The purified fragments were cloned into pBluescript SK(+).

RNA isolation and cDNA synthesis: RNA was isolated as described by Jain (2004). For RNA isolation from *T. parceramosum*, it was grown in 250 ml Shacked flasks at 28°C and 200 rpm in 50 ml medium which contained 0.14% $(NH_4)_2SO_4$, 0.2% KH_2PO_4 , 0.03% $CaCl_2 \cdot 2H_2O$, 0.03% $MgSO_4 \cdot 7H_2O$, 0.1% Tween 80, and 1 ml/l trace elements solution, contained 18 mM $FeSO_4 \cdot 7H_2O$, 6.6 mM $MnSO_4$, 4.8 mM $ZnSO_4 \cdot 7H_2O$, and 15 mM $CoCl_2$ (Takashima *et al.*, 1998). Avicel (2%) used as a carbon source and inducer. Cells were harvested after 25h of growth and frozen in liquid nitrogen. Frozen mycelium was ground and suspended in 5 volumes of guanidine isothiocyanate, 0.5% Na-lauryl sarcosinate, 25 mM Na.citrate pH 7.0 and 0.1 M β -mercaptoethanol.

The messenger RNA was purified by chromatography on oligo(dT)- cellulose (Sambrook and Russell, 2001). Poly(A)⁺ RNA eluted from the column was concentrated by ethanol precipitation, resuspended in sterilized dH_2O , and used freshly.

The cDNA synthesis using the poly(A)⁺ RNA was carried out by Revert Aid™ First Strand cDNA Synthesis Kit (Fermentas). The reaction volume was 50 μ l and contained: 5 μ g of poly(A)⁺ RNA, 1 mg of oligo(dT)₁₈, 20 units of RNase Block Ribonuclease Inhibitor, 1 X buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 10 mM dithiothreitol, 3 mM $MgCl_2$) 500 μ M of each dNTP, and 200 units reverse transcriptase. The RNA was denatured at 70°C, cooled slowly at room temperature to allow the annealing of primers before it was added to the reaction mixture. The reaction mixture was incubated at 42°C for 1h and then incubated

at 70°C for an additional 5 min. The cDNA from the reaction was kept at -70°C and used for a PCR reaction with specific primers. DNA amplification was carried out in a PCR reaction as described above.

The resulting 1.4-kb cDNA gene fragment was recovered and purified from a 1.2% (w/v) agarose gel using the High pure PCR product purification Kit and cloned into pBluescript SK(+) cloning vector.

General DNA procedures: Plasmid DNA preparation and electrophoresis of DNA fragments were performed by routine procedures (Sambrook and Russell, 2001). Enzymatic treatments of DNA molecules were carried out as recommended by manufacture (Fermentas). DNA fragment was recovered after electrophoresis on a 1% low melting temperature agarose gel and bacteria were transformed by the $CaCl_2$ method (Sambrook and Russell, 2001).

Construction of expression vector: Full length cDNA encoding the CBHII, without native signal sequence, was amplified with specific primers (C2EF and C2EB) using *Pfu* polymerase (Fermentas) and cloned into pPIC9 at *NotI/EcoRI* restriction sites and sequenced by a commercial service (Seqlab, Germani). The *Saccharomyces cerevisiae* α -factor secretion signal was used for the secretion of the recombinant enzyme. The construct was designated as pSZ5.

Transformation of *P. pastoris* and screening of recombinant colonies: *P. pastoris* strain GS115 was transformed by electroporation (BioRad-Gene PulserII) using 0.2 cm cuvettes (1.5 kv for 8 ms) with 20 μ g DNA fragment of *cbhII* cassette. In pSZ5 the *cbhII* cassette is flanking by *BglII* sites (Fig. 5). pSZ5 was digested by *BglII* and the fragment containing *cbhII* gene was purified from the gel and used for electroporation. After electroporation, 1ml of ice-cold 1M sorbitol solution was immediately added to the cuvettes, and the cuvette contents were transferred to 1.5 ml tube. After incubation at 30°C for 1h without shaking, cells were spread on minimal dextrose (MD) plates and incubated at 28°C for 2 days.

Before transformation, pSZ5 was linearized with *BglII*. *CbhII* gene is, flanked at its 5' end by AOX1 promoter and at its 3' end by the *his 4* gene and the AOX1 terminator sequence. The same cassette without an insert (*cbhII* gene) was also transformed into *P. pastoris* and used as a control.

His⁺ transformant colonies were picked from MD plates and transferred to minimal methanol (MM) plates. Colonies were spotted onto MM-Agar plates

supplemented with 0.5% Carboxy methyl cellulose (CMC) and incubated at 30°C for three days, followed by incubation at 45°C for 3h. The colonies were removed from surface of the media and were stained with 1% Congo red. Cellulase producing transformants were identified by the presence of a faint colorless halo around the *Pichia* colonies on the plates, resulting from CBHII activity.

Time-course studies of CBHII expression: A subculture of transformed *P. pastoris* desired clone was grown overnight in YPD medium (10 ml) in a 100 ml flask and 0.5 ml of this subculture was used to inoculate media containing glycerol (100 ml). After the OD₆₀₀ value reached three to four, cells were harvested by centrifugation at 3000×g for 5 min at room temperature and washed two times with quarter strength ringer solution, resuspended in 20 ml media containing methanol to induce expression. Methanol was added daily to a final concentration of 0.5% for 6 day to compensate the medium evaporation. CBHII expression was studied in three sets of media (YPG/YPM, BMGY/BMMY and BMG/BMM), temperatures (30, 25 and 22°C) and pH (4, 5 and 6). Every 24h, 0.5 ml of culture was removed from flasks, yeast cells pelleted down by centrifugation. The supernatant concentrated 10 times by ultrafiltration using Millipore membrane (with cut-off 30 kDa) and CBHII activity in supernatants was determined (Laymon *et al.*, 1996).

Northern dot blot and RT-PCR analysis: RNA was isolated from induced yeast cells after 48h. The cells were frozen in liquid nitrogen and ground in a mortar together with 1 g acid washed glass beads until a floury consistency was reached. Total RNA was extracted by the guanidinium thiocyanate method, (Sambrook and Russell, 2001). The mRNA was purified from total RNA using High Pure mRNA Purification Kit (Roch).

The Dig labeled probe specific for *cbhII* gene was synthesized by PCR amplification using CP3 and CP4 primers, Dig labeled dNTP (Roch) and pSZ4 as template and the 697 bp fragment isolated from the gel. 5 to 10 µg of the purified mRNA was transferred to Hybond N nylon membrane using Dot Blotter, and backed at 80°C for 2h. Hybridization was performed by standard methods (Sambrook and Russell, 2001) and developed according to Dig-labeling Kit manufacturer's instruction (Roche).

Sequence analysis and structure prediction: CBHII sequences from different *Trichoderma* sp. were retrieved from uniprot database (<http://www.pir.uniprot.org>).

Multiple sequence alignment was generated using Clustal W (<http://www.ebi.ac.uk/ClustalW>). A three dimensional (3D) model of *T. parceramosum* CBHII protein was predicted based on the known 3D structure of *T. reesei* CBHII molecule (Koivula *et al.*, 1996).

RESULTS

In order to amplify *cbhII* gene from *T. parceramosum*, a high cellulase enzymes producer, we aligned a set of known *cbhII* gene sequences from the GenBank database at the National Center for Biotechnology Information (NCBI). Based on the sequence similarity, the specific oligonucleotide primers (CF22 and CB22) were designed and the genomic and cDNA copies of *cbhII* were amplified from chromosomal and first-strand cDNA, respectively. A specific band of 1582 bp was amplified from *T. parceramosum* chromosomal DNA and confirmed by nested PCR using CP3 and CP4 primers (Fig. 1). The restriction pattern analysis using *XhoI*, *PstI*, *BclI* and *XbaI* enzymes did not reveal any similarity to the *T. reesei* PCR product, using CF22 and CB22 primers (data not shown). Following amplification of genomic DNA from *T. parceramosum*, DNA fragment was cloned in pBluescript SK(+) and designated pSZ3. The cloned fragment (1582 bp) was sequenced and submitted to the NCBI data base under accession No. AY651786.

Under conditions described in material and methods, cDNA fragment about 1.4 kb was obtained. This PCR product was isolated, ligated to the pBluescript SK(+) vector and designated pSZ4. Comparison of the cDNA sequence with the genomic *cbhII* sequence

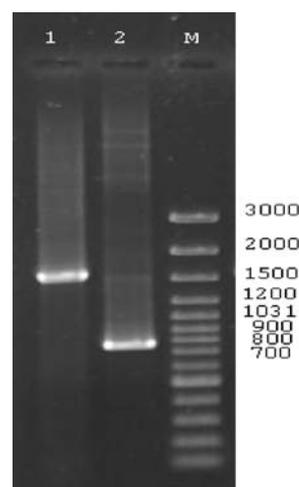


Figure 1. PCR amplification of *cbhII* gene. Lane1: *T. parceramosum* *cbhII* gene using CF22 and CB22 primers; Lane 2: nested PCR using CP3 and CP4 primers, M: DNA size marker.

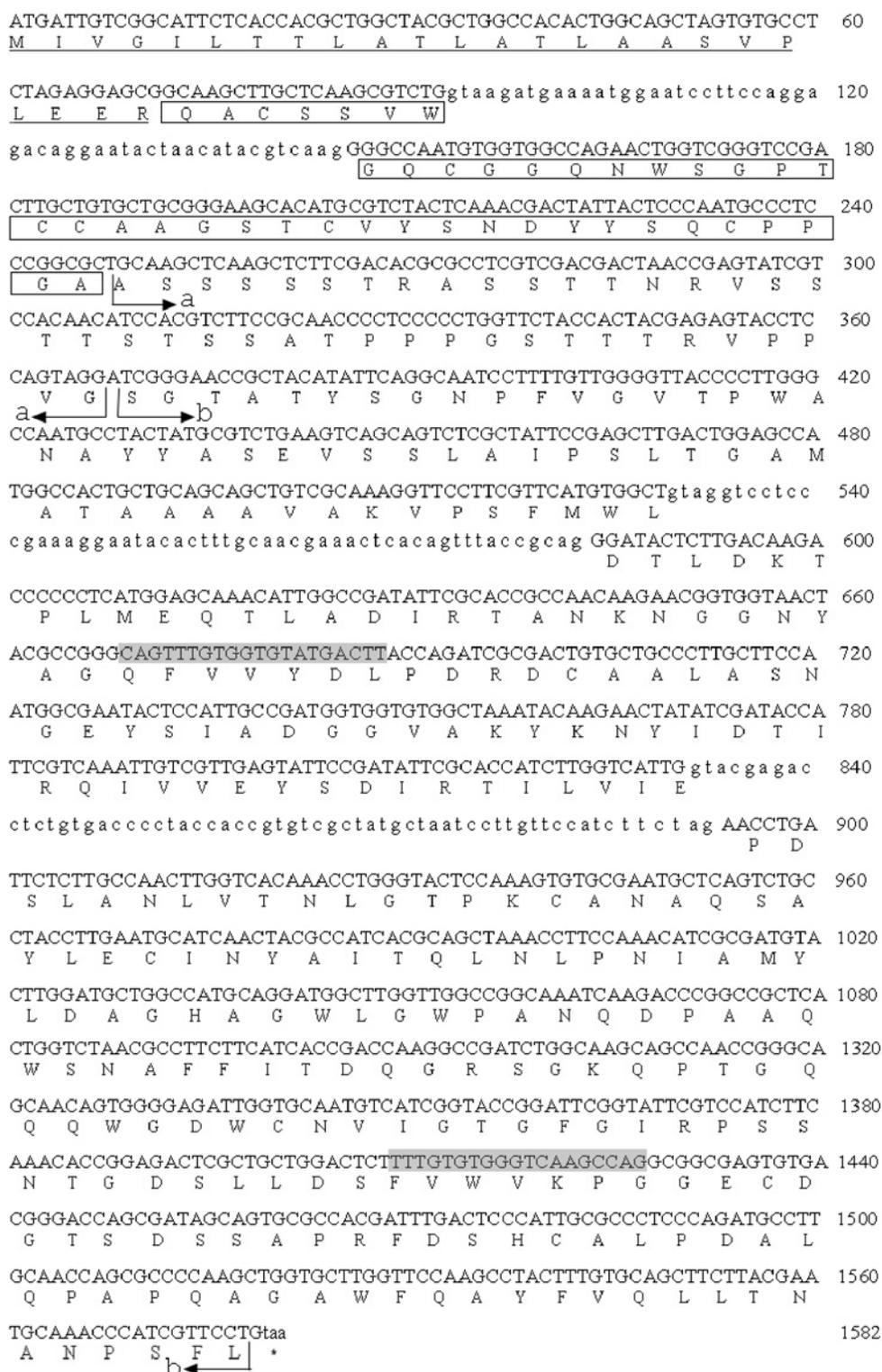


Figure 2. Nucleotide and deduced amino acid sequences of *T. parceramosum cbhl1* gene. The intron sequences are presented with small character. The signal sequence is underlined. Stop codon is indicated by an asterisk. The fungal type cellulose-binding domain (fCBD) is boxed. The linker region is shown by arrow "a". The catalytic core domain is shown by arrow "b". Shaded regions correspond to primers CP3 and CP4 used for nested PCR.

showed that this gene contains three introns, 53, 54 and 62 bp in length (Fig. 2). The cDNA contains an open reading frame encoding a protein of 470 amino acids (Fig. 2).

Multiple alignment of the deduced amino acid sequence with fungal CBHII proteins was performed with the CLUSTAL W program. When *T. parceramosum* CBHII was compared with the previously reported CBHII sequences of *T. koningi* (AF315681), *T. reesei* (M16190), and *T. viride* (AY368688), most of the residues are conserved and only 14 positions have been changed (Fig. 3). Pair wise alignment shows very high homology between the *T. parceramosum* CBHII and other fungal CBHII proteins, *T. koningi* (97.2%), *T. reesei* (97%), and *T. viride* (97.2%) (Fig. 4), while the homology at nucleotide sequence level showed to be 85.9%, 85.9% and 88.4%, respectively, most of the nucleotide differences were located in the introns.

Construction of the vector for CBHII expression in *P. pastoris* GS115: In this investigation we have used the methylotrophic yeast, *P. pastoris*, as the host for the expression and secretion of recombinant CBHII from *T. parceramosum*. The plasmid pSZ5 contained the native *S. cerevisiae* α -factor secretion signal upstream of the coding sequence of CBHII together with the termination signal and the AOX1 promoter (Fig. 5). For transformation of *cbhII* cassette into *P. pastoris* GS115, pSZ5 was digested by *Bgl*II and the fragment containing *cbhII* gene was purified from the gel and used for electroporation. Integration of *cbhII* gene into the yeast genome was confirmed by PCR using gene-specific and AOX1 promoter primers (Fig. 6). 63 transformants were screened out by functional assay after 2 days incubation by methanol. All colonies were transferred to minimal dextrose (MD) media and a replica plate in a minimal methanol (MM) media

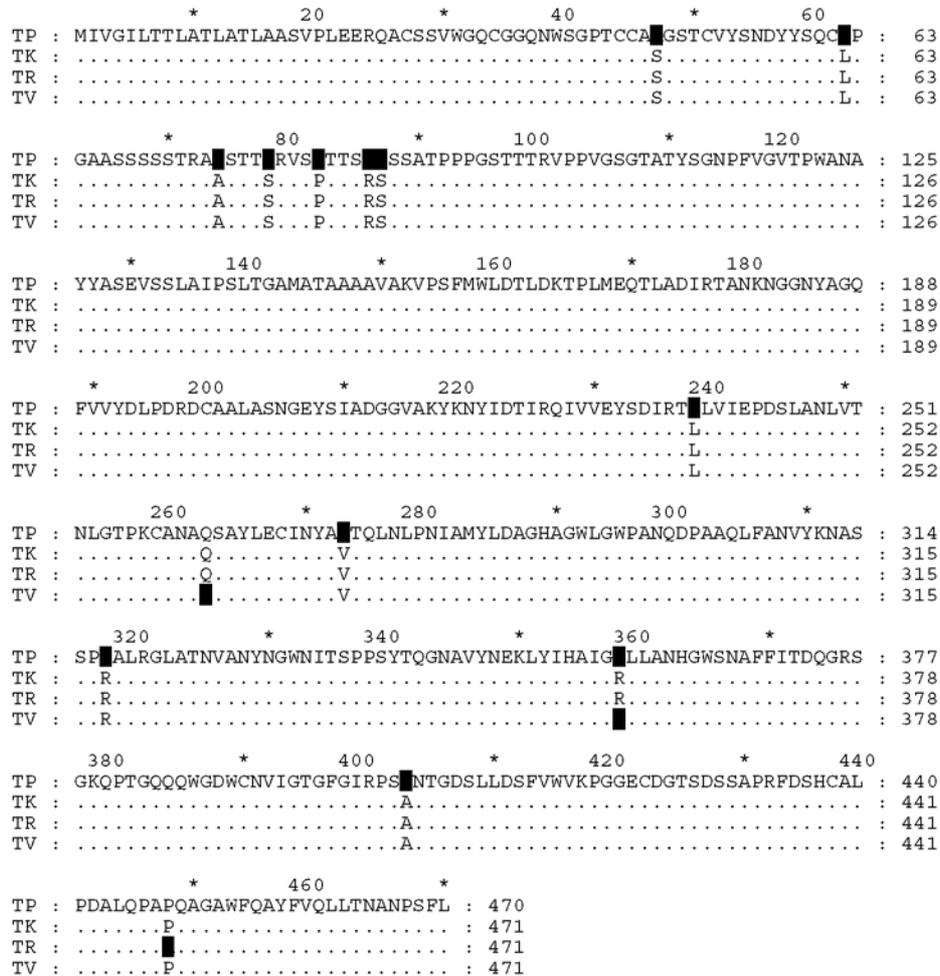


Figure 3. Sequence alignment with CLUSTAL W of the deduced amino acid sequences of *T. parceramosum* CBHII (TP) with those of *T. koningi* (TK), *T. reesei* (TR), and *T. viride* (TV). Single-letter amino acid is used.

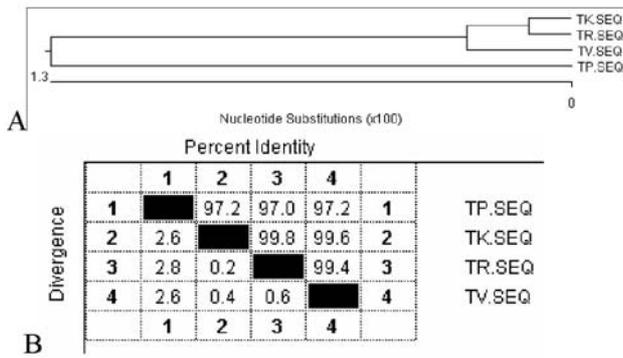


Figure 4. A) Phylogenetic tree and sequence distances of the CBHII amino acid sequences comparing *T. parceramosum* (TP) CBHII with that of *T. koningi* (TK), *T. reesei* (TR) and *T. viride*(TV). Figures show the phylogenetic relationships and B) percentage of sequence similarity and divergence between amino acid of *T. parceramosum* (TP) CBHII and *T. koningi* (TK), *T. reesei* (TR) and *T. viride*(TV) generated from multiple alignment.

supplemented with 0.5% carboxy methyl cellulose (MM-CMC) was prepared. Three colonies were selected by staining the MM-CMC plates with Congo red.

Expression of cellobiohydrolase II in *P. pastoris*: An attempt was made to express CBHII in baffled shake flasks in different media (containing glycerol or methanol), pH (4, 5 and 6) and temperatures (22, 25 and 30°C). Protein expression was initiated by checking carbon source from glycerol to methanol using YPG to YPM, BMG to BMM and BMGY to BMMY. The pH affects proteases activity thus finding suitable pH is important. Temperature value probably affects the efficiency of correct folding and secretion. The supernatants were collected and screened for CBHII

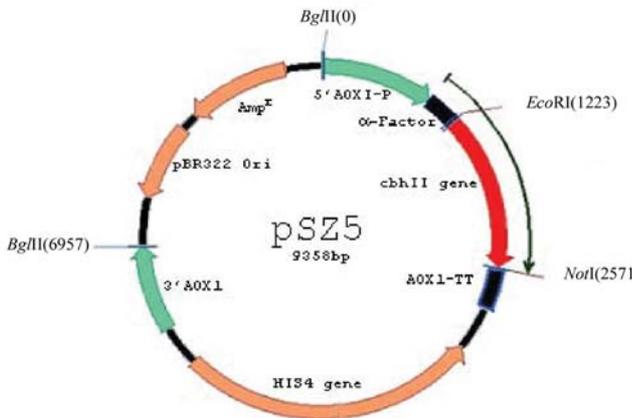


Figure 5. Plasmid construction for the expression of cellobiohydrolase II in *P. pastoris*. The *cbhII* cDNA was cloned into pPIC9 downstream of the native *S. cerevisiae* α - factor secretion signal to create pSZ5 plasmid.

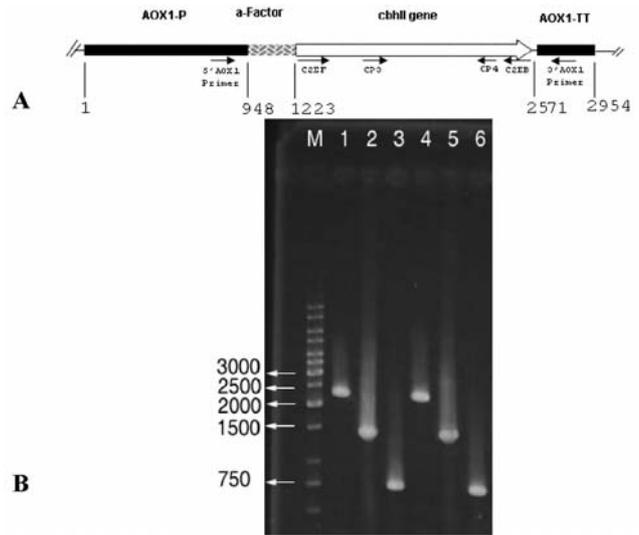


Figure 6. A) Position of primers on integrated cassette in *P. pastoris* genome. B) Confirmation of integration of *cbhII* cDNA in *P. pastoris* genome by PCR using 5'AOX1/3'AOX1 primers (lanes 1 and 4) Nested PCR using C2EF/C2EB (lanes 2 and 5), Nested PCR using C2EF/C2EB (lanes 3 and 6), M) DNA size marker.

activity every 24h. Enzyme activity was detected in BMG /BMM medium at pH 5 and 22°C culture condition and induction after 6 days (Fig. 7). However only low enzyme activity was detected in culture supernatants. Northern dot blot and RT-PCR analysis revealed transcription of the gene (Fig. 8).

Structure prediction of CBHII: Sequence of *T. parceramosum* exhibited 98.3% identity to Cel6A catalytic domain of *T. reesei* (PDB code: 1CB2). Figure 9 illustrates the observed sequence identity by showing the alignment of these two protein sequences. This allowed us to use the Cel6A as suitable template for protein 3D structure modeling. The predicted structure was found to have an RMS fit at 0.23Å for all

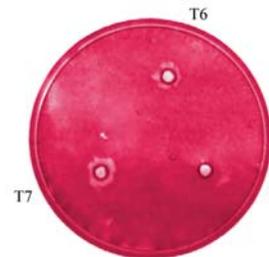


Figure 7. Activity staining of transformant *P. pastoris* culture supernatant. Enzyme activity detected as clear zone, T6: Transformant culture supernatant after 6 days, T7: Transformant culture supernatant after 7 days, C: Negative Control (transformant with pPIC9 vector).

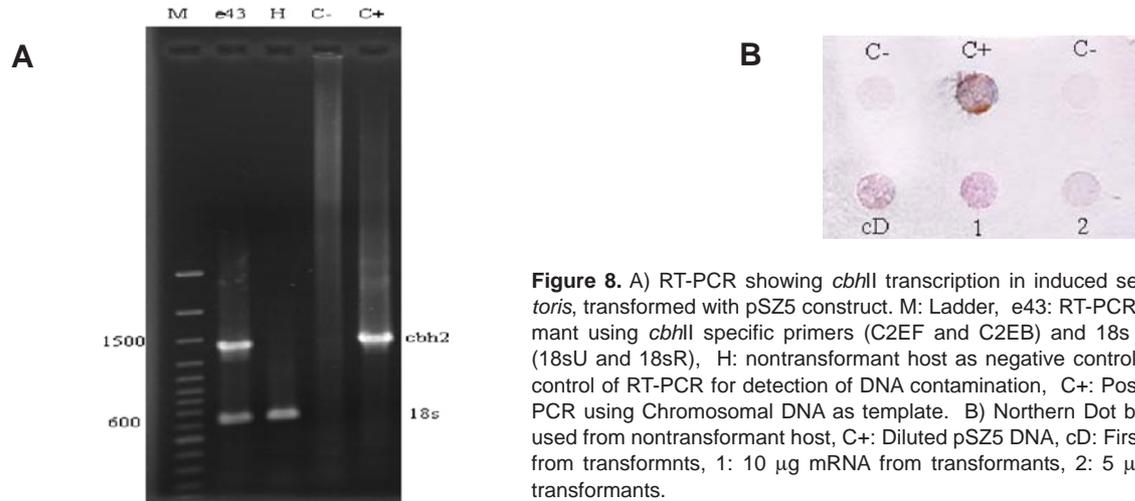


Figure 8. A) RT-PCR showing *cbhII* transcription in induced selected *P. pastoris*, transformed with pSZ5 construct. M: Ladder, e43: RT-PCR from transformant using *cbhII* specific primers (C2EF and C2EB) and 18s rRNA primers (18sU and 18sR), H: nontransformant host as negative control, C-: Negative control of RT-PCR for detection of DNA contamination, C+: Positive control of PCR using Chromosomal DNA as template. B) Northern Dot blot. C-: mRNA used from nontransformant host, C+: Diluted pSZ5 DNA, cD: First strand cDNA from transformants, 1: 10 µg mRNA from transformants, 2: 5 µg mRNA from transformants.

backbone atoms. This structure alignment shows functional similarity between two proteins (Fig. 10). Comparison of amino acid sequences in linker region of *T. parceramosum* CBHII with *T. reesei* shows amino acid differences in this region which increases the putative glycosylation sites in *T. parceramosum* CBHII protein (<http://www.cbs.dtu.dk/services/NetOGlyc>).

Analysis of CBHII enzyme from *T. reesei* based on crystallographic studies demonstrated three key residues of D199, D245 (proton donor) and D425

(nucleophile) in the active site, and four disulfide bounds which have main role in its 3D conformation (Rouvinen *et al.*, 1990 and Koivula *et al.*, 1996). Sequence similarity of CBHII from *T. reesei*, *T. parceramosum* and other related and unrelated fungi showed this active site residues and disulfide bound sites have been conserved (Fig. 11), suggesting a similar role of action in all cellobiohydralases II from different genus and species.

		EE	HHHHHHHHHH	HHH	HHHHHHHHHH	EEE	HHHH	
TP	: 165	SGTATYSGNPFVGVTPUANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKT						224
1CB2	: 1	SGTATYSGNPFVGVTPUANAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKT						60
		HHHHHHHHHHHHHHHH	EEEEEE		SSS S		HHHHHHHHHHHH	
TP	: 225	PLMEQTLADIRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNIYIDTI						284
1CB2	: 61	PLMEQTLADIRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNIYIDTI						120
		HHHHHH	EEEE	HHHH	HHHHHH	HHHHHHHHHHHHHH	EEEE	
TP	: 285	RQIVVEYSDIRITLLVIEPDSLNLVITNLGTPKCANAQSAYLECINYAITQLNLPNIAMYL						344
1CB2	: 121	RQIVVEYSDIRITLLVIEPDSLNLVITNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYL						180
		E	HHHH	HHHHHHHHHHHH	EEEE		HHH	
TP	: 345	DAGHAGWLGWPANQDPAALFANVYKNASSPSALRGLATNVANYNGWNITSPPSYTOGNA						404
1CB2	: 181	DAGHAGWLGWPANQDPAALFANVYKNASSPRALRGLATNVANYNGWNITSPPSYTOGNA						240
		HHHHHHHH	HHHHH	EEEE				
TP	: 405	VYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDUCNVIGTGFGRIPSSN						464
1CB2	: 241	VYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDUCNVIGTGFGRIPSAN						300
		EEE		HHH		HHHHHHHH		
TP	: 465	TGDSLLDSFVWVKPGGECGTSDSSAPRFDSHCALPDALQPAPQAGAMFQAYFVQLLTNA						524
1CB2	: 301	TGDSLLDSFVWVKPGGECGTSDSSAPRFDSHCALPDALQPAPQAGAMFQAYFVQLLTNA						360
TP	: 525	NPSFL						529
1CB2	: 361	NPSFL						365

Figure 9. Alignment of the catalytic domain of CBHII from *T. reesei* and deduced amino acid sequence of the catalytic domain of *T. parceramosum* CBHII by BLAST2. H and S represent Helix and b-Strand respectively.

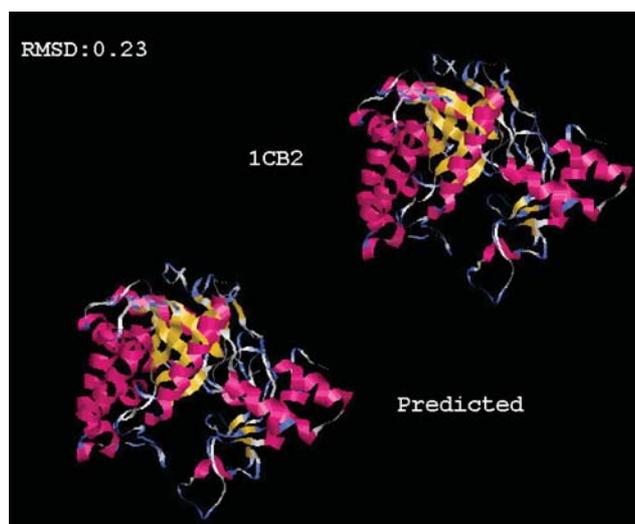


Figure 10. Structural prediction of CBHII using SWISS-MODEL, Predicted) *T. parceramosum*, 1CB2) *T. reesei*

DISCUSSION

A highly producer of cellulase enzymes Iranian isolate of *T. parceramosum* was used for amplification of the genomic and cDNA copies of *cbhII* gene. DNA sequence information confirmed showed high homology to the previously reported *cbhII* sequences of *T. koningi* (Zhu *et al.*, 2000a), *T. reesei* (Zhu *et al.*, 2000b), and *T. viride* (Liu *et al.*, 2003) has been cloned.

Analysis of the *cbhII* sequence demonstrated that it contains three small introns which also been reported in other *cbhII* fungal genes (Chen *et al.*, 1987; Murray *et al.*, 2003). The consensus sequences, GT on the 5' end and AG on the 3' end for each intron of the *cbhII* gene are also observed. The coding region of *cbhII* from *T. parceramosum* cods for a polypeptide of 470 amino acids, the first 24 residues of which form a putative signal peptide by similarity (Fagerestam and Petterson, 1980). The calculated size of the predicted product is 49 kDa. Thus, the mature secreted protein would have a calculated molecular mass of 47 kDa. The mature protein starts with a cellulose binding domain (41 amino acids), followed by a proline/serine/threonine rich 41 amino acids (from 66 to 106, based on the amino acid sequence alignment between CBHII protein from *T. parceramosum* and three other mentioned *Trichoderma* species) as linker and a 365 amino acids (107 to 470) (Fig. 2) as catalytic domain (Koivula *et al.*, 1996).

Expression of recombinant cellulases has previously been reported in bacteria (Laymond *et al.*, 1996), the yeast *S. cerevisiae* (Penttila *et al.*, 1988), and in fil-

amentous fungi (Takashima *et al.*, 1998). However, although the transformed cellulase gene was expressed and translated in *E. coli*, the enzyme product was not secreted but remained as inclusion bodies in cytoplasm (Laymond *et al.*, 1996), whereas recombinant cellulase secreted by *S. cerevisiae* was hyper-glycosylated resulting in reduced substrate-binding capacity and catalytic activity (Penttila *et al.*, 1988). We used *Pichia pastoris* for expression of CBHII protein.

Since the secretion of recombinant proteins is dependent upon the type of signal sequence adopted (Romanos *et al.*, 1992), the construct (pSZ5) contained the native *S. cerevisiae* α -factor secretion signal upstream of the coding sequence of CBHII together with the termination signal and the AOX1 promoter was used. Fusion of the target gene to the *S. cerevisiae* α -factor signal peptide directs efficient secretion of the target protein to the culture medium, which will facilitate the subsequent study of recombinant protein. Previous reports showed that the AOX1 promoter stimulates high levels of transcription in cells cultured in methanol containing medium and repressed in cells cultured on other carbon sources (Tschopp *et al.*, 1987; Cregg *et al.*, 1989).

In optimum conditions (pH=5, Temperature 22°C, and 6 days of incubation) for expression of CBHII in *P. pastoris* only low level of enzyme activity was detected. Although it has been reported that *P. pastoris* represents an appropriate host for the heterologous expression of functional fungal cellulases (Ding *et al.*, 2002), however low level of expression of recombinant proteins in *P. pastoris* may be the result of potential impediments for protein secretion including the codon usage of the expressed gene (Outchkourov *et al.*, 2002), copy number of the gene (Vassileva *et al.*, 2001), the efficiency and strength of promoters (Sears *et al.*, 1998), translation signals (Cavener and Ray, 1991) signal peptides (Zaror *et al.*, 2000 and Raemaekers *et al.*, 1999), processing and folding in the endoplasmic reticulum and Golgi (Kowalski *et al.*, 1998), environmental factors of expression (Villatte *et al.*, 2001), extracellular secretion (Rossini *et al.*, 1993), and protein turnover by proteolysis (Cregg *et al.*, 2000). Therefore, it remains to be studied which of these control points for enhancement of CBHII production exist in the secretory machinery of *P. pastoris*.

Structural prediction of CBHII based on comparative protein modeling was elaborated online using SWISS-MODEL and Swiss-pdbViewer (Guex and Peitsch, 1997). The fit between 3D structure model was evaluated in Swiss-pdbViewer by calculating the Root Mean Square (RMS) deviation. The predicted structure was found to have an RMS fit at 0.23Å for all

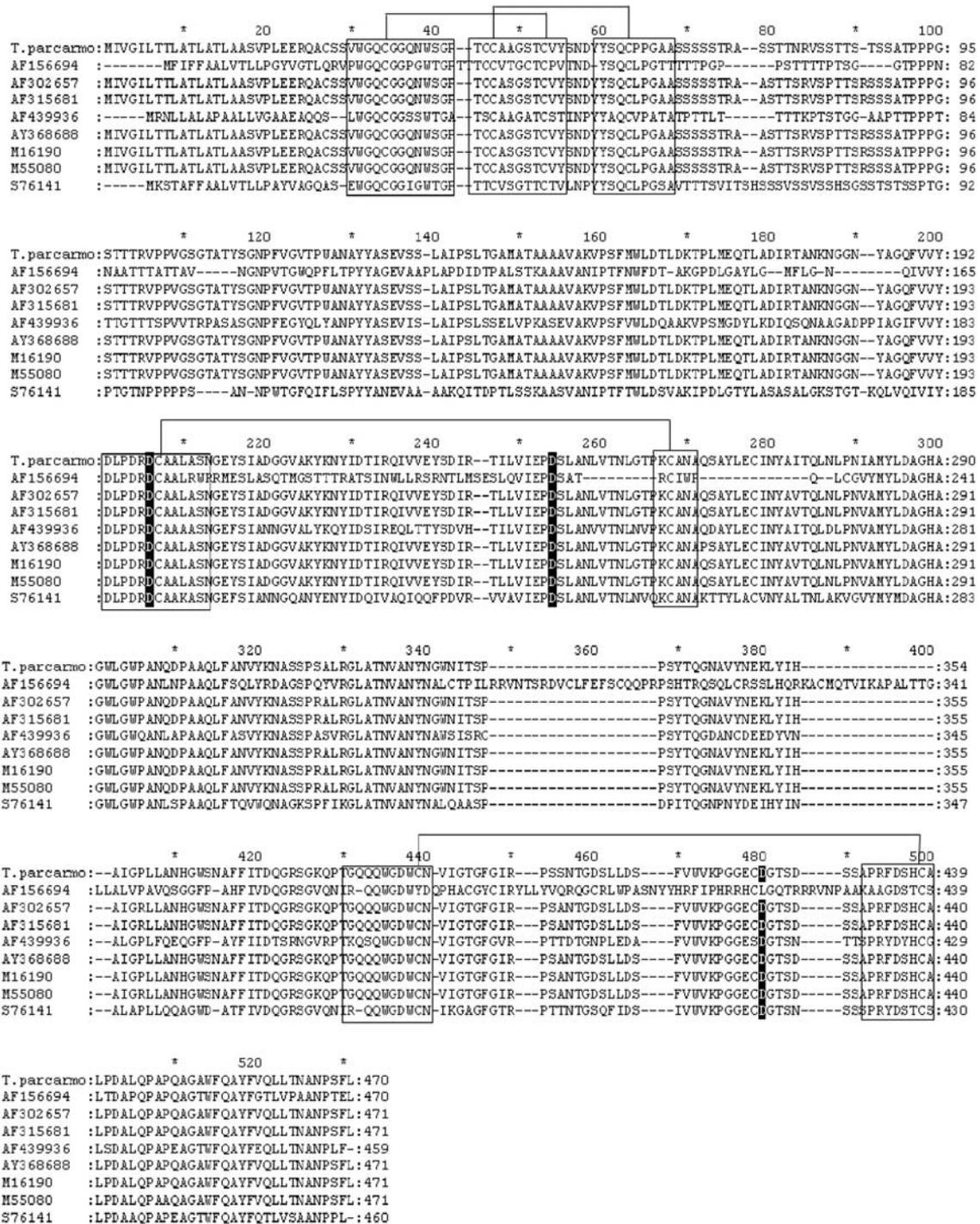


Figure 11. Alignment of CBHII protein sequence from different sources. Cystein residues which involve in disulfide bonds [at the conserved regions (Boxed)], were shown by linked lines. Three key residues (D) in active site were highlighted. (AF156694, *Volvariella volva*; AF302657, *Hypocrea jecorina*; AF315681, *T. koningi*; AF439936, *Talaromyces emersoni*; AY368688, *T. viride*; M16190, *T. reesei*; M55080, *T. reesei*; S76141, *Phanerochaete chrysosporium*).

backbone atoms. This structure alignment shows functional similarity between two proteins

The role of glycosylation is related to the secretion of these cellulolytic enzymes, to provide sufficient spatial separation between the core and cellulose binding domains, and to protect the linker peptide against proteolysis (Clarke, 1997; Hui *et al.*, 2001). Linker region of *T. parceramosum* CBHII shows to have higher putative glycosylation sites in comparison to *T. reesei* CBHII. We thus predict the structure of *T. parceramosum* CBHII with some functional changes in comparison to the same enzyme from *T. reesei*.

References

- Beguín P, Aubert J (1994). The biological degradation of cellulose. *FEMS Microbiol Rev.* 13: 25-58.
- Beguín P, Lemaire M (1996). The cellulosome: An exocellular, multi-protein complex specialized in cellulose degradation. *Crit Rev Biochem Mol Biol.* 31: 201-236.
- Cavener DR, Ray SC (1991). Eukaryotic start and stop translation sites. *Nucleic Acids Res.* 19: 3185-3192.
- Cereghino JL, Cregg JM (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev.* 24 45-66.
- Chen CM, Gritzali M, Stafford DW (1987). Nucleotide sequence and deduced primary structure of cellobiohydrolase II from *Trichoderma reesei*. *Biotechnology.* 5: 274-278.
- Clarke AJ (1997). Biodegradation of cellulose. In: *Enzymology and Biotechnology*. Technomic publishing, Pennsylvania, pp: 55.
- Cregg JM, Cereghino JL, Shi J, Higgins DR (2000). Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol.* 16: 23-52.
- Cregg JM, Madden KR (1987). *Biological Research on Industrial Yeasts*. Volume II, pp 1-18, CRC Press. Boca Raton.
- Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman CA (1989). Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. *Mol Cell Biol.* 9: 1316-1323.
- Ding SJ, Gea W, Buswell JA (2002). Secretion, purification and characterization of a recombinant *Volvariella volvacea* endoglucanase expressed in the yeast *Pichia pastoris*. *Enz Microbial Technol.* 31: 621-626.
- Fagerestam LG, Petterson LG (1980). The 1,4- β - Glucan cellobiohydrolases of *Trichoderma reesei* QM9414, A new type of cellulolytic synergism. *FEBS Lett.* 119: 97-100.
- Guex N, Peitsch MC (1997). SWISS-MODEL and Swiss-pdbViewer: An environment for comparative protein modeling. *Electrophoresis.* 18: 2714-2723
- Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G (1996). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci USA.* 93: 5674.
- Higgins DR, Cregg JM (1998). Introduction to *Pichia pastoris*. In: Higgins, DR, Cregg, JM (Eds.), *Methods in Molecular Biology: Pichia Protocols*. Humana Press, Totowa, NJ, pp. 1-15.
- Hui JPM, Lanthier P, White TC, McHugh SG, Yaguchi M, Roy R, Thibault P (2001). Characterization of cellobiohydrolase I (Cel7A) glycoforms from extracts of *Trichoderma reesei* using capillary isoelectric focusing and electrospray mass spectrometry. *J Chromatogr B.* 752: 349-368.
- Jain AK (2004). Cloning and structural analysis of a cDNA clone encoding glycinin (*Gly-1*) seed storage protein of peanut. *Electronic J Biotechnol.* 7: 224-234
- Koivula A, Reinikainen T, Ruohonen L, Valkeajarvi A, Claeysens M, Teleman O, Kleywegt GJ, Szardenings M, Rouvinen J, Jones TA, Teeri TT (1996). The active site of *Trichoderma reesei* cellobiohydrolase II: the role of tyrosine 169. *Protein Eng.* 9: 691-699.
- Koivula A, Ruohonen L, Wohlfahrt G, Reinikainen T, Teeri TT, Piens K, Claeysens M, Weber B, Vasella A, Becker D, Sinnott ML, Zou J, Kleywegt G, Szardenings M, Stahlberg J, Jones TA (2002). The Active Site of the Cellobiohydrolase Cel6A from *Trichoderma reesei*: The Roles of the Aspartic acids D221 and D175. *J Am Chem Soc.* 124: 10015-10024.
- Kowalski JM, Parekh RN, Mao J, Wittrup KD (1998). Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control. *J Biol Chem.* 273: 19453-19458.
- Laymon RA, Adney WS, Mohagheghi A, Himmel ME, Thomas SR (1996). Cloning and expression of full-length *Trichoderma reesei* cellobiohydrolase I cDNA in *Escherichia coli*. *Appl Biochem Biotechnol.* 57/58: 389-97.
- Lin Cereghino GP, Lin Cereghino J, Ilgen C, Cregg JM (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr Opin Biotechnol.* 13: 329-332.
- Lin Cereghino GP, Sunga AJ, Lin Cereghino J, Cregg JM (2001). Expression of foreign genes in the yeast *Pichia pastoris*. In: Setlow JK (Ed.). *Genetic Engineering*. 157-170.
- Liu BD, Yang Q, Zhou Q, Song JZ (2003). Cloning and Sequence Analysis of cellobiohydrolase II (*cbhII*) gene from *Trichoderma viride* CICC 13038, Department of Life Science and Engineering, Harbin Institute of Technology, Dong Da Zhi, Harbin, Heilongjiang 150001, PR. China.
- Murray PG, Collins CM, Grassick A, Tuohy MG (2003). Molecular cloning, transcriptional, and expression analysis of the first cellulase gene (*cbh2*), encoding cellobiohydrolase II, from the moderately thermophilic fungus *Talaromyces emersonii* and structure prediction of the gene product. *Biochem Biophys Res Commun.* 301: 280-6.
- Nutt A, Sild V, Pettersson G, Johansson G (1998). Progress curves: A mean for functional classification of cellulases. *Eur J Biochem.* 258: 200-206.
- Outchkourov NS, Stiekema WJ, Jongsma MA (2002). Optimization of the expression of equistatin in *Pichia pastoris*. *Protein Expr Purif.* 24: 18-24
- Penttila M, Andre L, Lehtovaara P, Bailey M, Teeri TT, Knowles J (1988). Efficient secretion of two cellobiohydrolases by *Saccharomyces cerevisiae*. *Gene.* 63: 103-12.
- Raemaekers RJ, de Muro L, Gatehouse JA, Fordham-Skelton AP (1999). Functional phytohemagglutinin (PHA) and Galanthus nivalis agglutinin (GNA) expressed in *Pichia pastoris* correct Nterminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *Eur J Biochem.* 265: 394-403.
- Romanos MA, Scorer CA, Clare JJ (1992). Foreign gene expression in yeast. A review: *Yeast.* 8: 423-88.
- Romanos M (1995). Advances in the use of *Pichia pastoris* for high-level expression. *Curr Opin Biotechnol.* 6: 527- 533.
- Rossini D, Porro D, Brambilla L, Venturini M, Ranzi BM, Vanoni M, Alberghina L (1993). In *Saccharomyces cerevisiae*, protein secretion into the growth medium depends on environmental

- factors. *Yeast*, 9: 77-84.
- Rouvinen J, Bergfors T, Teeri TT, Knowles JKC, Jones TA (1990). Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science*, 249: 380-386.
- Sambrook J, Russell DW (2001). *Molecular Cloning*. Cold Spring Harbor: New York.
- Sears IB, O'Connor J, Rossanese OW, Glick BS (1998). A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast*, 14: 783-790.
- Seyed Asli N., Zamani MR, Motallebi M (2004). Study of chitinolytic enzyme production in *Trichoderma* isolates. *Iranian J Biol*. 17: 227-246.
- Sun CB, Kong QL, Xu WS (2002). Efficient transformation of *Penicillium chrysogenum* mediated by *Agrobacterium tumefaciens* LBA4404 for cloning of *Vitreoscilla hemoglobin* gene. *E J Biotechnol*. 5: 29-41.
- Takashima S, Iikura H, Nakamura A, Hidaka M, Masaki H, Uozumi T (1998). Overproduction of recombinant *Trichoderma reesei* cellulase by *Aspergillus oryzae* and their enzymatic properties. *J Biotechnol*. 65: 163-171.
- Tomme P, Van Tilbeurgh H, Pettersson G, Van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeysens M (1988). Study of cellulolytic system of *Trichoderma reesei* QM 9414: Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem*. 170: 575-581.
- Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR (1987). Expression of the *lacZ* gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acid Res*. 15: 3859-3876.
- Van Tilbeurgh H, Tomme P, Claeysens M, Bhikhabhai R, Peterson G (1986). Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. *FEBS Lett*. 204: 223-227.
- Vassileva A, Chugh DA, Swaminathan S, Khanna N (2001). Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic Yeast *Pichia pastoris*. *Protein Expr Purif*. 21: 71-80.
- Villatte F, Hussein AS, Bachmann TT, Schmid RD (2001). Expression level of heterologous proteins in *Pichia pastoris* is influenced by flask design. *Appl Microbiol Biotechnol*. 55: 463-465.
- Zaror I, Crawford K, Bishop RJ, Innis MA (2000). *Pichia* secretary leader for protein expression, *US Patent No.* 6,107,157.
- Zhu L, Dong Z, Yu W (2000a). *Cloning and analysis of the cbhII gene*. Institute of Microbiology: Beijing 100080, China.
- Zhu L, Dong Z, Yu W (2000b). *Cloning and sequence analysis of the cellobiohydrolase II gene (cbh2) from Trichoderma koningii K*. Institute of Microbiology: Beijing 100080, China.