

Identification of the Gene Encoding the Major Cellobiohydrolase of the White Rot Fungus *Phanerochaete chrysosporium*

AMBER VANDEN WYMELENBERG,¹ SARAH COVERT,^{2†} AND DANIEL CULLEN^{1*}

*Institute for Microbial and Biochemical Technology, Forest Products Laboratory, Forest Service,
U.S. Department of Agriculture, One Gifford Pinchot Drive, Madison, Wisconsin 53705-2398,¹ and
Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin 53705²*

Received 29 April 1993/Accepted 12 July 1993

Previous studies have shown that the cellobiohydrolases of the white rot basidiomycete *Phanerochaete chrysosporium* are encoded by a family of structurally related genes. In this investigation, we identified and sequenced the most highly transcribed gene, *cbh1-4*. Evidence suggests that in this fungus the dominant isozyme, CBH1, is encoded by *cbh1-4*.

The white rot fungus *Phanerochaete chrysosporium* has been widely studied as a model organism for lignocellulose degradation (3, 6). In submerged culture, *P. chrysosporium* secretes a complex array of degradative enzymes, including three classes of hydrolytic cellulases: endoglucanases, cellobiohydrolases (CBH), and β -glucosidases. Multiple CBH isozymes have been characterized (11). One isozyme, CBH1, constitutes 12% of the total extracellular protein (11).

Various studies have shown that the number, structure, organization, and regulation of *P. chrysosporium* CBH genes are unlike previously described cellulolytic systems (1, 2, 9). Six *cbh*-like genes have been cloned and characterized, although the individual translational products of these genes have not been identified. Three genes, designated *cbh1-1*, *cbh1-2*, and *cbh1-3*, were shown to be clustered within a 25-kb region. Competitive reverse transcription-polymerase chain reaction (RT-PCR) techniques showed that *cbh1-1* and *cbh1-2* were constitutively expressed at low levels (1). Transcript levels of *cbh1-3* were approximately 1,000-fold higher than those of *cbh1-1* and *cbh1-2* under cellulose induction in submerged culture (2).

Multiple alignments of partial DNA sequences clearly showed that *cbh1-3*, *cbh1-4*, *cbh1-5*, and *cbh1-6* represent a highly homologous subfamily distinct from *cbh1-1* and *cbh1-2* (1). Pairwise comparisons of deduced amino acid sequences within this subfamily show approximately 85 to 95% identity (1). Investigation of the transcriptional regulation of these genes is lacking, in large part because of the difficulties in distinguishing such closely related transcripts.

In this study, relative levels of *cbh1-3*, *cbh1-4*, *cbh1-5*, and *cbh1-6* transcripts under cellulose induction were determined using competitive RT-PCR techniques (5). The strategy involved first-strand cDNA synthesis with an oligo(dT) primer followed by PCR amplification with primers specific for each gene (Table 1). Gene-specific primers were designed to amplify the same intron-containing region of all four genes. Dilutions of competitive genomic templates were added to the PCR reaction mixtures. The initial concentration of a specific cDNA was then determined by estimating the dilution points at which the concentrations of the target

cDNA and competitive template were equivalent. In each case, cosmid clones (1) served as competitive templates. By this method, cDNA levels were estimated to be approximately 0.01, 0.5, 0.02, and 0.05 ng for *cbh1-3*, *cbh1-4*, *cbh1-5*, and *cbh1-6*, respectively (Fig. 1). Previous studies had shown that *cbh1-3* transcripts are far more abundant than *cbh1-1* and *cbh1-2* transcripts under cellulose induction (2). Thus, *cbh1-4* is the most abundant transcript under these culture conditions.

The complete nucleotide sequence of a genomic *cbh1-4* clone was determined (Fig. 2), and introns were confirmed by comparison with the cDNA sequence. Comparison with *cbh1-3* (2) showed that the positions and approximate lengths of the two introns were conserved, but the internal sequences diverge substantially. The possibility of additional introns in the 5' untranslated region cannot be ruled out by this analysis.

The deduced amino acid sequence of the *cbh1-4* gene product is similar to other CBH gene products, especially that of *cbh1-3* (9). Overall, the two gene products are 79.2% identical, and most of sequence divergence occurs near the carboxy terminus in a region rich in Ser, Thr, and Pro residues. This region, often referred to as the hinge, sepa-

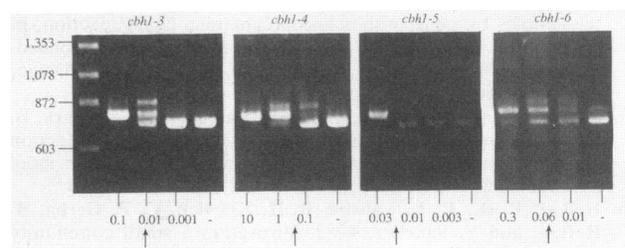


FIG. 1. Competitive PCR quantitation of mRNAs of *P. chrysosporium* *cbh1-3* through *cbh1-6*. Total RNA was extracted from cellulose-grown cultures as previously described (2). PCR primers are shown in Table 1. PCR reaction mixtures received competitive cosmid templates as indicated in nanograms; -, reaction mixtures received no competitive template. RT-PCR reaction conditions were as previously described (10). Vertical arrows indicate estimated equivalence between cDNA and competitive targets. Sizes of molecular weight markers are shown in the left margin.

* Corresponding author.

† Present address: Department of Genetics, University of Georgia, Athens, GA 30602.

TABLE 1. Primers for PCR amplification of *P. chrysosporium* CBH genes

Gene ^a	5' primer ^b	3' primer ^{b,c}
<i>cbh1-3</i>	5'-AAGATCGTGCTCGACGCGAAActg.....ctgAAGCCCTTCACCGTTGTCACc-3'	
<i>cbh1-4</i>	5'-AAGGTCGTCTCGACTCGAAActg.....CTCCAAGCCCTTCACCGTcgcac-3'	
<i>cbh1-5</i>	5'-ACCGTCGTGCTCGACTCGAAActg.....CAGCAAGCCCTTCACCGTcgcacg-3'	
<i>cbh1-6</i>	5'-AAGATCGTGCTCGACGCGAAActg.....cagcAAGCCCTTCACCGTcgcacg-3'	

^a Based on previously reported sequences: *cbh1-3* (9), *cbh1-5* and *cbh1-6* (1).
^b Dotted lines represent approximately 750 bp of intervening sequence in genomic clones. Capital letters designate oligonucleotide sequences.
^c For 3' primers, the reverse complement oligonucleotide was used in PCR.

rates the catalytic domain from the cellulose-binding domain common to microbial cellulases (7). Typical of an extracellular enzyme, a putative secretion signal can be identified by the net positive charge and hydrophobic core of the first 18 amino acids. Thus, the predicted molecular mass and pI of the mature peptide would be 52 kDa and 4.4, respectively. These values are in reasonable agreement with experimentally determined properties of the glycosylated CBH1 isozyme (molecular mass, 60 kDa; pI, 3.8) (11). Finally, the deduced sequence is identical to CBH1 peptide sequences (Fig. 2) (1). The dominance of the *cbh1-4* transcript, the predicted physical characteristics of the protein, and the precise matching of the deduced and experimentally determined sequences provide strong evidence that *cbh1-4* encodes the major CBH isozyme.

Nucleotide sequence accession number. The GenBank and EMBL nucleotide sequence accession number is L22656.

This work was supported in part by USDA grants 90-37291-5724 and 92-37103-7980.

REFERENCES

- Covert, S., J. Bolduc, and D. Cullen. 1992. Genomic organization of a cellulase gene family in *Phanerochaete chrysosporium*. *Curr. Genet.* 22:407-413.
- Covert, S., A. Vanden Wymelenberg, and D. Cullen. 1992. Structure, organization and transcription of a cellobiohydrolase gene cluster from *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* 58:2168-2175.
- Eriksson, K.-E. L., R. A. Blanchette, and P. Ander. 1990. Microbial and enzymatic degradation of wood and wood components. Springer Series in Wood Science. Springer-Verlag, Berlin.
- Frohman, M. A., M. K. Dush, and G. R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85:8998-9002.

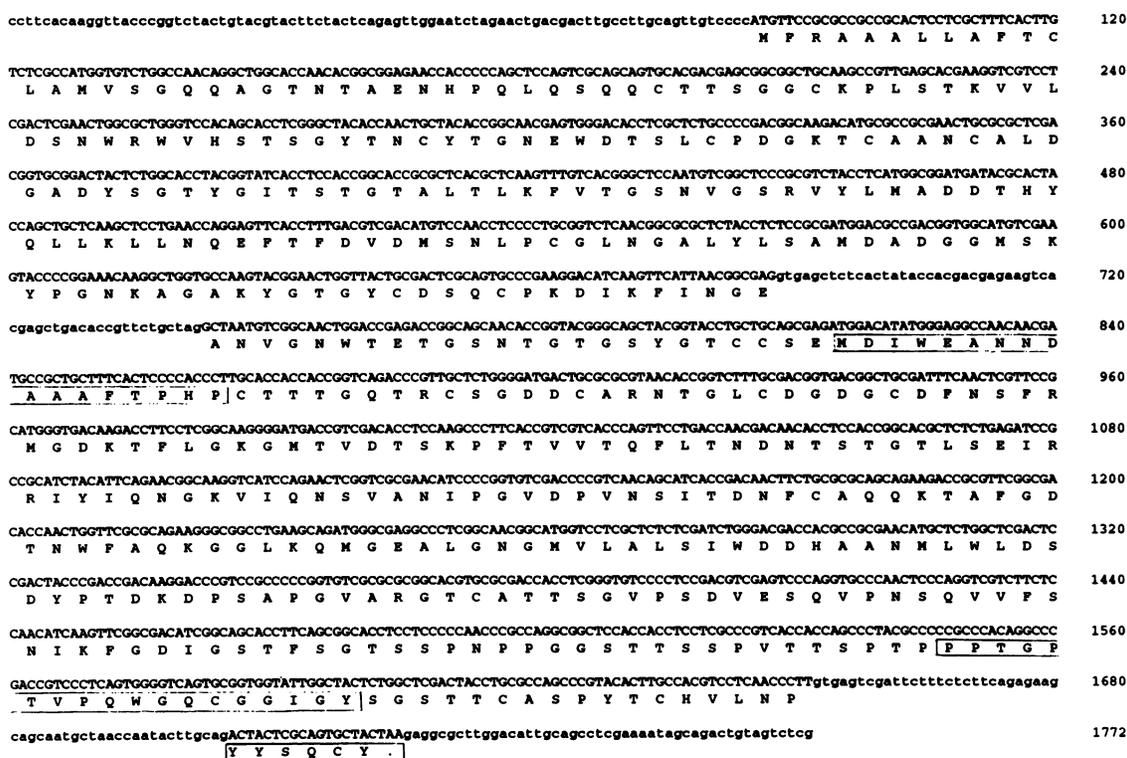


FIG. 2. Nucleotide and deduced amino acid sequences of *P. chrysosporium cbh1-4*. Lowercase lettering indicates presumed untranslated sequence. Intron positions were verified from the cDNA sequence. The cDNA was amplified by RT-PCR (4) with oligo(dT) and 5'-AAGGTCGTCTCGACTCGAA-3' primers. Boxed amino acid sequences are identical to experimentally determined peptide sequences of isozyme CBH1 (11). Double-stranded templates were sequenced by the dideoxy chain termination method (8). Approximately 550 bp of 5' untranslated sequence of *cbh1-4* was previously reported (1).

5. Gilliland, G., S. Perrin, and H. Bunn. 1990. Competitive PCR for quantitation of mRNA, p. 60-69. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. Academic Press, Inc., New York.
6. Kirk, T. K., and R. L. Farrell. 1987. Enzymatic "combustion": the microbial degradation of lignin. *Annu. Rev. Microbiol.* **41**:465-505.
7. Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. *Trends Biotechnol.* **5**:255-261.
8. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
9. Sims, P., C. James, and P. Broda. 1988. The identification, molecular cloning and characterization of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exo-cellobiohydrolase I gene from *Trichoderma reesei*. *Gene* **74**:411-422.
10. Stewart, P., P. Kersten, A. Vanden Wymelenberg, J. Gaskell, and D. Cullen. 1992. The lignin peroxidase gene family of *Phanerochaete chrysosporium*: complex regulation by carbon and nitrogen limitation and identification of a second dimorphic chromosome. *J. Bacteriol.* **174**:5036-5042.
11. Uzcatogui, E., A. Ruiz, R. Montesino, G. Johansson, and G. Pettersson. 1991. The 1,4- β -D-glucan cellobiohydrolase from *Phanerochaete chrysosporium*. I. A system of synergistically acting enzymes homologous to *Trichoderma reesei*. *J. Biotechnol.* **19**:271-286.