Cloning of a Gene Encoding a Putative Xylanase with a Cellulose-Binding Domain from *Humicola grisea*

Hiroshi Ikura, Shou Takashima, Akira Nakamura, Haruhiko Masaki, and Takeshi Uozumi

Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113; Japan

Received March 5, 1997

We have isolated a genomic clone of a putative xylanase gene (xyn1) from *Humicola grisea* by using the DNA fragment encoding a cellulose-binding domain of *H. grisea* cellobiohydrolase I as a probe. The translation product of the xyn1 gene predicts a xylanase of 429 amino acids in length, with a cellulose-binding domain in the C-terminus.

Key words: cellulose-binding domain; *Humicola grisea*; xylanase

Many fungal cellulases, in particular, cellobiohydrodolases (CBH, exo-1,4-β-glucanases, EC 3.2.1.91) and endoglucanases (EGL, endo-1,4-β-glucanase, EC 3.2.1.4) have characteristic domain structures consisting of a catalytic domain, a cellulose-binding domain, and a hinge region between these two domains, which is rich in serine, threonine, and proline residues.11 Deletion of the cellulose-binding domain from *Trichoderma reesei* CBHI and CBHII greatly reduced the enzymatic activity toward crystalline cellulose,22 indicating that the tight binding to cellulose mediated by the cellulose-binding domain is necessary for an efficient hydrolysis of crystalline cellulose. The cellulose-binding domains of fungal cellulases contain 30-40 amino acid residues. All the fungal cellulose-binding domains show high sequence similarity, of which some amino acid residues are highly conserved. It is known that some of fungal xylanases (XYN, EC 3.2.1.8) also have characteristic domain structures like cellulases, consisting of a catalytic domain, a cellulose-binding domain (but not a xylan-binding domain), and a hinge region.

Among the cellulolytic fungi, the genus *Humicola* has been known to produce several kinds of thermostable cellulases,4,5 and several cellulase genes have been cloned from *H. grisea*60 and *H. insolens*.31 Some of *Humicola* cellulases also have a cellulose-binding domain that shows high sequence similarity with other fungal cellulose-binding domains. It seems probable that using a DNA fragment encoding a cellulose-binding domain as a probe, novel genes encoding enzymes with a cellulose-binding domain could be cloned. Therefore, we tried to clone a novel gene encoding an enzyme with a cellulose-binding domain, from the genomic DNA of *H. grisea* var. thermoidea IF09854.4,60

A 218-bp *Aor*5HI–EcoRI fragment encoding a cellulose-binding domain of *H. grisea* CBHI was prepared from the cloned cbhl gene.60 The DNA fragment was radiolabeled with [32P]dCTP, using a random primed DNA labeling kit (Boehringer Mannheim), and used as a probe for detection of DNA fragments with a sequence similar to that encoding the cellulose-binding domain. Southern hybridization71 with this probe DNA was done to the chromosomal DNA of *H. grisea* digested withSplI, according to the standard procedure except that we washed the filter at 37°C instead of 65°C to get a low stringency. We found that several DNA fragments were hybridized with this probe DNA other than that corresponding to the cbhl gene, of which a 6.9-kb fragment showed a strong signal. Therefore we cloned with DNA fragment by the colony hybridization technique.72 Restriction and partial sequence analyses of this clone suggested that a part of it, 4.4-kb HindIII–SphI fragment contained a gene encoding an enzyme with a cellulose-binding domain, and this fragment was sequenced on both strands by the dyeoxy sequencing method (Fig. 1).81

When compared with the amino acid sequences in data bases, we found the deduced amino acid sequence showed a certain similarity to several xylanases rather than cellulases. However, the reading frame was interrupted at two positions. Taken together with the consensus sequence of fungal introns,80 we estimated the locations and lengths of two introns as nucleotide positions 337–408 and 636–709. The predicted translation product of the coding region would have 429 amino acids with a molecular mass of 47,016 Da. The overall translation product sequence sequence similarity with xylanases belonging to cellulase family F (corresponding to glycosyl hydrolase family 10)10, 40.1% identity with *Streptomyces lividans* xylanase A,11 39.2% identity with *Magnaporthe grisea* xylanase 33,12 and 33.3% identity with *Thermotoga* sp. xylanase A.13 The typical hinge and cellulose-binding domain are also found in C-terminal of the translation product. The amino acid sequence of the cellulose-binding domain shows 72.2% identity with that of *H. grisea* CBHII, and 60.0% identity with both that of *T. reesei* CBHI14 and *T. reesei* EGLI (Fig. 2).15 Therefore, the cloned gene seemed to encode a xylanase with a cellulose-binding domain, which belongs to cellulase family F, and we designated it as the xyn1 gene.

The frequencies of T, C, A, and G at the third letter of codons of the xyn1 gene are 27T, 271C, 1A, and 130G, respectively. In the xyn1 gene, cytosine and guanine are preferred at the third letter of codons, while there is a strong bias against the use of adenine at the third letter of codons. This bias in codon usage is also observed in some cellulase genes of *H. grisea*,6,10 suggesting this bias may be related to the rate of expression of each gene. The GC content of the xyn1 gene is 66.3%, which is higher than those of the *H. grisea* cbhl gene (63.4%), eglI gene (64.9%), and egl2 gene (59.5%).6,10

In the 5′-upstream sequence of the xyn1 gene that was analyzed, a putative TATA box and putative CAAT motifs are found. There are also three 6-bp sites which have an identical or similar sequence to the consensus sequence for binding catabolite repressors, *Aspergillus nidulans* CREA and *T. reesei* CRE1 (5′-SYGGGRG-3′; S=G or C, Y=C or T, R=G or A).17,11 It may be possible that some of these sequences are related to carbon catabolite repression mediated by a catabolite repressor of *H. grisea*.

Genomic Southern hybridization analysis with the cloned xyn1
gene as a probe showed a single band on any restriction digests of H. grisea chromosomal DNA. This indicates that the *xyn1* gene is a single copy gene (Fig. 3a).

To investigate whether the *H. grisea xyn1* gene is functional, mycelia were grown for 36 h at 37°C in the presence of 1% glucose, 1% xylose, or 1% Avicel (microcrystalline cellulose)
Fig. 3. Hybridization Analysis of the H. grisea xyn1 Gene.
(a) Southern hybridization analysis of the H. grisea xyn1 gene. H. grisea chromosomal DNA digested with EcoRI (lane 1), HindIII (lane 2), and SphI (lane 3).
(b) Northern hybridization analysis of the xyn1 transcript. RNA was extracted from mycelia grown on 1% glucose (lane 1), 1% xylan (lane 2), or 1% Avicel (lane 3) as a main carbon source.

as a main carbon source. After mycelia were harvested and ground in liquid nitrogen, total RNA was extracted by using a QuickPrep total RNA extraction kit (Pharmacia). Approximately 7.5 μg of total RNA was fractionated by formaldehyde-agarose gel electrophoresis, transferred to a nylon membrane, and Northern hybridization was done with a radiolabelled fragment of the xyn1 gene as a probe. The result showed that the xyn1 gene was transcribed when mycelia were grown on xylan or Avicel as a main carbon source, but was not transcribed when mycelia were grown on glucose as a main carbon source (Fig. 3b). It seemed that xylan is more inducible than Avicel for the expression of the xyn1 gene. It has been known that some enzymes belonging to cellulase family F have broad substrate specificities and hydrolyze both cellulose substrate and xylan.\(^{39}\) The xyn1 gene is inducible by xylan and Avicel, although the substantial inducer seems to be degradation product of these substrates. So, it is suggested that the translation product of the xyn1 gene may hydrolyze some cellulose substrate as well as xylan.

We also tried to express the xyn1 gene in Aspergillus oryzae using the amyB promoter,\(^{16}\) but failed for reasons we do not understand.

In this study, we have cloned a novel xylanase gene from H. grisea, using high similarity observed among the fungal cellulose-binding domains. There are also many enzymes which have characteristic domain structures like fungal cellulases. So this cloning strategy seems applicable for the cloning of genes encoding an enzyme with such characteristic domain structures.

References