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

Ernesto Yagüe,† Milena Mehak-Zunic, Louis Morgan, David A. Wood and Christopher F. Thurston

Author for correspondence: Ernesto Yagüe. Tel: +44 171 955 4083; Fax: +44 171 403 8883.

e-mail e.yague@umds.ac.uk

1 Division of Life Sciences, King's College London, Campden Hill Road, London W8 7AH, UK
2 Amycel/Sapawnmate, 777 Maher Court, PO Box 189, Watsonville, CA 95076-9014, USA
3 Plant Pathology and Microbiology Department, Horticulture Research International, Wellesbourne, Warwick CV35 9EF, UK

Two new cellulose-growth specific (cel) cDNAs, *cel*2 and *cel*4, have been isolated from an *Agaricus bisporus* cDNA expression library by immunoscreening with an *A. bisporus* anti-‘endoglucanase’ antibody. The deduced amino acid sequences showed that both *CEL2* and *CEL4* proteins have a modular structure consisting of a fungal-type cellulose-binding domain (CBD) and a catalytic domain separated by a linker region rich in Pro, Ser and Thr. The *CEL2* and *CEL4* catalytic domains were homologous to fungal cellbiohydrolases (CBH) in family 7 and to fungal mannanases in family 5 of the glycosyl hydrolases, respectively. A previously isolated cDNA derived from a constitutive gene was also sequenced. The deduced amino acid sequence corresponded to 5-aminolaevulinic acid synthase (ALA), the first enzyme in the haem biosynthetic pathway, and was most similar to other fungal ALAs. RNA analysis showed that the expression of *cel*2 and *cel*4 genes was induced by cellulose and repressed by glucose, fructose and lactose. The soluble cellulose derivative CM-cellulose induced mRNA accumulation for *cel*1 but not *cel*2, *cel*3 or *cel*4. Mannitol, maltose, sorbitol and glycerol decreased *cel*2 and *cel*4 mRNA levels to different extents. *cel*1, *cel*2, *cel*3 and *cel*4 mRNAs all disappeared after the addition of glucose with apparent half-lives of less than 20 min. Whether *cel* mRNAs have short half-lives or glucose affects the stability of *cel* transcripts remains to be investigated.

**Keywords:** *Agaricus bisporus*, cellbiohydrolase, mannanase, gene regulation, 5-aminolaevulinic acid synthase

**INTRODUCTION**

*Agaricus bisporus*, the edible mushroom, produces a complete cellulase system that enables its growth on crystalline cellulose as the sole carbon source (Manning & Wood, 1983; Wood et al., 1988). We have undertaken the study of cellulase production by *A. bisporus* from a molecular perspective with the aim of understanding its regulation both in minimal media and during development of the sporophore. Two genes (*CEL1* and *CEL3*) and their products (*CEL1* and *CEL3*) have been characterized to date (Raguz et al., 1992; Chow et al., 1994; Armesilla et al., 1994; Yagüe et al., 1996). *CEL1* and *CEL3* proteins have a modular structure (Fig. 1) like many other cellulases (Béguin & Aubert, 1994). Recombinant CEL3 has been expressed in *Saccharomyces cerevisiae* as an active protein able to hydrolyse several cellulolytic substrates (Chow et al., 1994). However, CEL1, which has no homologue in the sequence databases, has no endoglucanase (EG) activity (measured as Congo red staining), but it binds to crystalline cellulose (Armesilla et al., 1994). In addition, recombinant CEL1 expressed in *S. cerevisiae* does not decrease the viscosity of CM-cellulose solutions, release soluble sugars from...
Avicel or filter paper, or have any turbidimetric effect on Avicel suspensions, nor does it produce any disrupting effect on cotton cellulose fibrils (this laboratory, unpublished observations).

Cellulase production by A. bisporus is inducible by cellulose and repressed by easily metabolizable carbon sources (Manning & Wood, 1983; Wood et al., 1988). Induced expression of both cel1 and cel3 in minimal medium occurs at least in part by an increase in transcription of the genes, as measured in nuclear run-on assays (Yagüe et al., 1994; Chow et al., 1994).

In this paper we present the isolation and characterization of two new A. bisporus cel cDNAs, the characterization of a constitutive cDNA, and the regulation of expression of the corresponding genes by the carbon source in minimal medium.

METHODS

Organism and culture conditions. A. bisporus D649 mycelium was grown in 50 ml of defined minimal medium (Manning & Wood, 1983) with Whatman CC41 microcrystalline cellulose (0.5 g l\(^{-1}\)), or with CM-cellulose, D-fructose, D-glucose or glycerol (each 1.0 g l\(^{-1}\)). The transfer experiments were carried out by filtering the culture medium through a nylon mesh, frozen in liquid nitrogen, and kept at -70 °C. Addition of carbon sources to previously grown cultures was made in 5 ml minimal medium. After the specified cultivation times mycelium was harvested by filtering through a nylon mesh, frozen in liquid nitrogen, and kept at -70 °C.

Recombinant DNA techniques. A previously described cDNA library from cellulose-grown mycelium in lambda ZAP was immunoscreened with an A. bisporus anti-'EG' antiserum (diluted 1:200) (Raguz et al., 1992) and a horseradish-peroxidase-linked donkey anti-rabbit Ig second antibody (diluted 1:2500) (Sigma). Immunodetection was carried out using the ECL detection kit (Amersham) following the manufacturer's instructions. Positive immunocrossreacting plaques were isolated and clones obtained by in vivo excision and rescued using ExAssist helper phage (Stratagene). DNA fragments were subcloned into M13mp18 and sequenced with Sequenase (Amersham). Clone pScr300 was sequenced mainly as double-stranded template in an ABI automatic sequencer (Molecular Medicine Unit, King's College London) following the manufacturer's instructions. Sequence analysis was performed with the programs in the GCG package (Genetics Computer Group, 1991). Total RNA isolation and Northern analysis were carried out essentially as previously described (Raguz et al., 1992).

RESULTS

Isolation of cel2 and cel4 cDNAs

Approximately 170000 recombinant plaques from an A. bisporus cDNA library in lambda ZAP were immunoscreened with an A. bisporus anti-'EG' antibody (Raguz et al., 1992). Thirty putative positive plaques were selected in a first screen, of which 13 were clearly positive in a second screen. The recombinant cDNAs were rescued as inserts in pBluescript. One of these clones hybridized to cel1 cDNA (Raguz et al., 1992) and two to cel3 cDNA (Chow et al., 1994); these were discarded. The remaining 10 were used to make radio-labelled probes and hybridized to RNA isolated from mycelium growing on minimal medium with either crystalline cellulose or glucose as the sole carbon source (results not shown). Seven clones were further discarded since five were not derived from cellulose-growth specific mRNAs and two gave ambiguous results. The remaining three clones, pEYC521, pEYC522 and pEYC1200, hybridized to mRNAs synthesized under cellulose-growth specific conditions of approximately 1.4 kb (pEYC521 and pEYC522) and 1.2 kb (pEYC1200).

Cross-hybridization experiments indicated that both pEYC521 and pEYC522 were derived from the same mRNA and these clones also hybridized to a PCR fragment obtained by using primer based on sequence from cbh1 of Trichoderma reesei (results not shown). Clones pEYC521 and pEYC1200 were used as probes to re-screen the cDNA library and isolate approximately 20 clones hybridizing to each cDNA. The longest clones were further analysed and sequenced.

The architecture of the CEL2 and CEL4 proteins indicates a multidomain structure

Deduced amino acid sequences from the open reading frames found in cel2 (pEYC521 and derivatives) and cel4 (pEYC1200 and derivatives) cDNAs showed that the encoded proteins had a modular structure (Fig. 1). Both showed a signal peptide at the N-terminus typical of the proteins to be secreted to the extracellular medium. A linker rich in Pro, Ser and Thr separated the fungal cellulose-binding domain (CBD) from the catalytic domain. As in other fungal cellulose-binding proteins, the CBD can be located either at the C-terminus (CEL1 and CEL2) or at the N-terminus (CEL3 and CEL4) of the protein.

Glycosyl hydrolases have been classified into several families according to amino acid sequence similarities (Henrissat, 1991; Henrissat & Bairoch, 1993). CEL2 belongs to family 7, a family presently containing only fungal enzymes with either EG or cellobiohydrolase (CBH) activity. Sequence comparison using the PILEUP program showed that CEL2 aligned better with the CBHs than with the EGs (results not shown). All the CBHs show four distinct regions that are deleted from EGs. These regions form extended loops over the catalytic cleft of CBHs, explaining the differences in substrate specificity and catalytic properties between EGs and CBHs (Rouvinen et al., 1990; Divne et al., 1994). Within the CBHs, CEL2 showed the strongest homology with CBH-I-4 from Phanerochaete chrysosporium (67.1% identical amino acids) and the weakest with CBH-I-2 from the same organism (57.3% identity) (Fig. 2).

The catalytic domain of CEL4 belongs to family 5, by far the most extensive family of glycosyl hydrolases (Henrissat & Bairoch, 1993). This family includes members from both aerobic and anaerobic bacteria, and
Regulation of cellulose-induced genes in Agaricus

<table>
<thead>
<tr>
<th>Name</th>
<th>Architecture</th>
<th>T. reesei homologue</th>
<th>Mol. mass (kDa)</th>
<th>Function</th>
<th>Family</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEL1</td>
<td>-</td>
<td>-</td>
<td>30.5</td>
<td>49.8</td>
<td>CELLULOSE BINDING NOT CELLULASE</td>
<td>-</td>
</tr>
<tr>
<td>CEL2</td>
<td>CBHI</td>
<td>52.4</td>
<td>52.4</td>
<td>52.0</td>
<td>CELLULASE</td>
<td>7</td>
</tr>
<tr>
<td>CEL3</td>
<td>CBHII</td>
<td>44.1</td>
<td>44.9</td>
<td>44.9</td>
<td>CELLULASE</td>
<td>6</td>
</tr>
<tr>
<td>CEL4</td>
<td>MANI</td>
<td>44.9</td>
<td>100 aa</td>
<td>MANANASE</td>
<td>5</td>
<td>250095</td>
</tr>
</tbody>
</table>

Fig. 1. Characteristics of the CEL proteins from A. bisporus. Proteins are depicted from the N- (left) to the C-terminus (right). Dotted domains at the N-terminus represent the signal peptides and black domains the Pro/Ser/Thr-rich linker regions that separate the small CBDs from the bigger catalytic domains. Molecular masses were deduced from the corresponding open reading frame and do not include the signal peptide. In vivo molecular masses have been determined by SDS-PAGE. The functions of CEL1 and CEL3 were derived from experimental results whereas those of CEL2 and CEL4 are deduced from sequence similarity. Family numbers follow the classification of glycosyl hydrolases (Henrissat & Bairoch, 1993).

Fig. 2. Dendrogram showing the relatedness of A. bisporus CEL2 with CBHs belonging to family 7 of glycosyl hydrolases. Catalytic domain sequences were aligned by using the program PILEUP and the output was used to create a matrix of pairwise distances (expressed as substitutions per 100 amino acids) between the aligned sequences with the program DISTANCES. The dendrogram was created with the program GROWTREE from the distance matrix. A.b., A. bisporus; H.g., Humicola grisea; N.c., Neurospora crassa; P.c., Phanerochaete chrysosporium; P.j., Penicillium janthinellum; T.k., Trichoderma koningii; T.r., T. reesei; T.v., T. viride. Database accession numbers are indicated.

fungi, with EG, mannanase, or 1,3-β-exoglucanase activity. CEL4 showed the strongest homology with the mannanases from Aspergillus aculeatus and T. reesei: 43.2% and 41.8% identical amino acid residues, respectively (Fig. 3). Homologies with EGs, 1,3-β-exoglucanases and bacterial mannanases were between 16-1 and 21.7% (data not shown). Interestingly, the T. reesei mannanase also has a CBD, but in contrast to CEL4 it is at the C-terminal end (Stålbrand et al., 1995).

Hem1: a gene whose expression is independent of the carbon source

We have sequenced a full-length cDNA, pSRc300 (EMBL accession number Z50096), derived from a gene whose mRNA synthesis (Yagüe et al., 1994) and steady-state levels (Fig. 3b in Raguz et al., 1992) are independent of the carbon source used for growth of A. bisporus. The deduced protein sequence showed a high degree of similarity with ALA (results not shown), the first enzyme in the haem biosynthetic pathway. The highest scores were obtained with the fungal ALAs: the A. bisporus, Aspergillus nidulans and Saccharomyces cerevisiae enzymes share 38.9% identical residues. The Asp. nidulans ALA was the most similar to the A. bisporus enzyme, sharing 321 identical amino acids out of 648 and 621, respectively (data not shown).

Regulation of hem1 gene expression has been well studied in mammals (Melefors et al., 1993; Conboy et al., 1992; Yamamoto et al., 1988). In S. cerevisiae, which can switch metabolism between aerobic and anaerobic conditions, hem1 gene expression is regulated by transcriptional activation and repression (Keng & Güarente, 1987); however, in aerobic fungi, the product of the gene is likely to be essential and therefore to be expressed constitutively. Steady-state levels of the hem1 mRNA from Asp. nidulans do not change in response to
variations in oxygen tension, carbon source or heat shock (Bradshaw et al., 1993), and in A. bisporus both the synthetic rate and steady-state levels of hem1 mRNA are independent of the carbon source (Raguz et al., 1992; Yagüe et al., 1994). Consequently, we have used the hem1 mRNA as an internal control for the analysis of cel mRNAs levels described below.

Steady-state levels of cel2 and cel4 mRNAs are regulated by the carbon source

To gain insight into the regulation of cel2 and cel4 gene expression we performed similar experiments to those previously described with cel1 and cel3 (Yagüe et al., 1994; Chow et al., 1994). When crystalline cellulose was added to A. bisporus cultures grown on fructose, cel2 and cel4 mRNAs were detected only after 11 d (when the fructose concentration in the culture had decreased to 1 μg ml⁻¹; data not shown) and accumulated thereafter (Fig. 4a). A similar decrease in fructose concentration in the control experiment, to which only minimal medium was added, did not trigger accumulation of cel2 and cel4 mRNAs. When the fructose-grown mycelium was washed prior to transfer to cellulose medium, cel2 and cel4 mRNAs were detectable after 3 d, although significant accumulation was evident only after 6 d (Fig. 4b). These results are therefore similar for the four cel genes studied so far, and indicate that the presence of crystalline cellulose is necessary, though not sufficient, to trigger the expression of the genes. However, of the four cel genes, cel1 is unique in that induced mRNA levels are found during growth on the soluble cellulose derivative CM-cellulose albeit at a lower level than that observed on crystalline cellulose (Yagüe et al., 1994). Differences between the accumulation of cel mRNAs induced by crystalline cellulose were observed when other sugars were added to the culture medium. Like cel3 (Chow et al., 1994), cel2 and cel4 mRNAs were undetectable 24 h after the addition of glucose, fructose or lactose (Fig. 4c). When the effect of glucose addition on cel mRNAs steady-state levels was monitored after shorter incubation times, all cel mRNAs disappeared with apparent half-lives of less than 20 min (data not shown). Mannitol and maltose led to a greater decrease of the cel2 transcript than glycerol and sorbitol. By contrast, the four carbon sources decreased the amount of cel4 transcript to a similar extent (Fig. 4c). In contrast, with the exception of glycerol, all the sugars tested rendered cellulose-induced cel1 mRNA undetectable under the same conditions (Yagüe et al., 1994). None of the cel mRNAs were detectable in glycerol-grown mycelium in the absence of cellulose (results not shown; Yagüe et al., 1994; Chow et al., 1994).

DISCUSSION

After a thorough screening of a cDNA expression library with an A. bisporus anti-'EG' antibody, the cDNAs for two CBHs (CEL2 and CEL3), one putative mannanase (CEL4) and a cellulose-binding protein of as yet unknown catalytic activity have been isolated (Fig. 1). Although originally described as an anti-EG antibody (Wood & Thurston, 1991), further analysis has demonstrated that the A. bisporus culture fraction used to raise this antiserum had, in addition to EG, β-glucosidase, CBH (measured as 4-methylumbelliferone 4-β-d-cellobiopyranoside hydrolysis) and crystalline-cellulose-degrading activities (this laboratory, unpublished observations). In addition, this antibody was unable to recognize a glutathione-S-transferase-CEL1 fusion protein (Armesilla et al., 1994) in which the linker and CBD had been deleted (results not shown). Because of the high antigenicity of the linker and CBD, polyclonal antibodies raised against T. reesei EGs recognize CBHs, and vice versa (Aho & Paloheimo, 1990). In all probability,
therefore, the *A. bisporus* anti-EG’ antibody recognizes preferentially any protein with a CBD. The reason why no EG has yet been detected in the cDNA expression library remains uncertain, although several explanations can be envisaged. The EG cDNAs could be under-represented in the library due to low levels of their transcripts in the sample used to construct the library. Alternatively, the most abundant *A. bisporus* EGs may not possess a CBD, as is the case for some other fungal EGs (Ooi *et al.*, 1990).

The CEL4 catalytic core belongs to family 5 of the glycosyl hydrolases, the most extensive family to date. Its members can be grouped in different subfamilies having different catalytic properties, such as EG, mannanase or 1,3-β-glucanase activity (Henrissat & Bairoch, 1993). By sequence comparison, CEL4 is most closely related to the mannanase subfamily (Fig. 3). However, substrate specificity of family 5 glycosyl hydrolases can be relatively broad. For example, CelH of *Clostridium thermocellum* is able to hydrolyse CM-cellulose, xylan and barley β-glucan (Yagüe *et al.*, 1990). The chemically determined N-terminal sequence of some EG fractions from *A. bisporus* (Wood & Thurston, 1991) matches well with the deduced N-terminus of CEL4 (data not shown). Therefore, at present it cannot be ruled out that CEL4 may have some EG activity, although such activity has not been documented in the *T. reesei* mannanases characterized so far (Arisan-Atac *et al.*, 1993; Stålbrand *et al.*, 1993, 1995).

Accumulation of the four *A. bisporus* cel mRNAs induced by crystalline cellulose follows a very similar pattern (Fig. 4; Yagüe *et al.*, 1994; Chow *et al.*, 1994). We have demonstrated previously by nuclear run-on analysis that this accumulation is, in part, the result of an activation of transcription for the *cel1* and *cel3* genes (Yagüe *et al.*, 1994; Chow *et al.*, 1994), and this is also likely for *cel2* and *cel4*. Whether an additional regulation takes place at the level of mRNA degradation remains to be determined. In yeast, glucose is known to affect the stability of transcripts for maltase and the iron-protein subunit of succinate dehydrogenase, among others (Federoff *et al.*, 1983; Lombardo *et al.*, 1992). In *A. bisporus*, cel mRNAs decay rapidly after the onset of repression. However, demonstration that regulation occurs at the level of mRNA degradation would require comparison of the kinetics of cel mRNA decay under inducing and repressing conditions. Such data may be obtained by pulse-chase experiments (Federoff *et al.*, 1983), by using temperature-sensitive mutants of RNA polymerase II (Parker *et al.*, 1991), or by blocking transcription with appropriate inhibitors. The latter would probably be the option of choice to study mRNA turnover in *A. bisporus*.

**ACKNOWLEDGEMENTS**

This work was supported by the Agricultural Systems Directorate of BBSRC (United Kingdom).

**REFERENCES**


Received 18 June 1996; revised 16 August 1996; accepted 27 August 1996.