FEMS Microbiology Letters 122 (1994) 303–308

Published by Elsevier

FEMSLE 06189

Transcriptional regulation of the *Trichoderma longibrachiatum egl1* gene

R. González [1], J.A. Pérez-González, L. González-Candelas and D. Ramón *

Departamento de Biotecnología, Instituto de Agroquímica y Tecnología de los Alimentos (CSIC), Jaime Roig 11, 46010-Valencia, Spain

(Received 29 July 1994; accepted 31 July 1994)

Abstract: Transcription of the *Trichoderma longibrachiatum egl1* gene is induced in the presence of lactose and β-methylglucoside and repressed by glucose. A DNA fragment containing 722 bp upstream of the ATG codon has been sequenced. The gene has two major transcription start points (20 and 24 nucleotides upstream from the ATG codon) and several transcription termination points (located in a region around 130 nt downstream of the stop codon). Two 6-mer sequences (5'-CTGGAG-3') separated by 16 bp are present in the *egl1* gene promoter. These sequences match the *Aspergillus nidulans* consensus CreA binding site and might be implicated in carbon catabolite repression of *egl1* transcription.

Key words: *Trichoderma longibrachiatum*; *egl1*; β-(1,4)-Endoglucanase; Transcriptional regulation; Carbon catabolite repression; CreA

Introduction

Cellulose is the most abundant naturally occurring organic polymer. This plant polysaccharide is used as a carbon source by numerous and diverse microorganisms which secrete cellulolytic enzymes. Filamentous fungi are good producers of cellulolytic activities and the genes encoding these proteins have been cloned from several fungal species, mainly *Trichoderma reesei* [1]. While the structure and function of these genes have been dealt with in detail, little work has been dedicated to the study of their regulation at the molecular level. In *T. reesei* the regulation of cellulase formation occurs at the level of transcription [2,3], is induced by cellulose and repressed by glucose [2].

The mesophilic deuteromycete *Trichoderma longibrachiatum* produces a complete set of cellulases which act synergistically to hydrolyse crystalline cellulose to glucose. Recently, we have cloned and sequenced the *T. longibrachiatum egl1* gene encoding β-(1,4)-endoglucanase activity [4]. This gene is interrupted by two introns and encodes a protein very similar to its *T. reesei* counterpart. The cDNA of this gene has been expressed in industrial wine yeast strains, producing a wine with an increased fruity aroma [5]. Using a

* Corresponding author. Tel.: (+34-6) 369 0800; Fax: (+34-6) 393 0001; e-mail: DRAMON@IATA.CSIC.ES

[1] *Present address:* Institut de Génétique et Microbiologie, URA 1354, Université Paris-Sud, Bâtiment 409, 91405-Orsay Cedex, France.

SSDI 0378-1097(94)00341-6

transformation system based on the acquisition of hygromycin B resistance, we have constructed *T. longibrachiatum* multicopy transformants for the *egl1* gene, which produce increased endoglucanase levels [6]. In the present work we analyse the transcriptional regulation of this gene.

Materials and Methods

Strains and culture conditions

Escherichia coli DH5α F^-, *endA1, hsdR17, gyrA96, thi1, recA1, supE44,* Δ*lacU169* (ϕ*80 lacZ*Δ*M15*) was used as the recipient strain in cloning experiments, and was grown as described in Sambrook et al. [7].

T. longibrachiatum CECT2606 strain was used throughout this work. Induction conditions for endoglucanase production were as follows. 1-l Erlenmeyer flasks containing 250 ml of *Trichoderma* minimal medium [8], omitting urea and using glucose as the only carbon source, were inoculated with 5×10^5 spores ml^{-1}. After 24 h of incubation at 30°C in an orbital shaker at 200 rpm, the mycelium was recovered by filtration and resuspended at a concentration of 20 mg

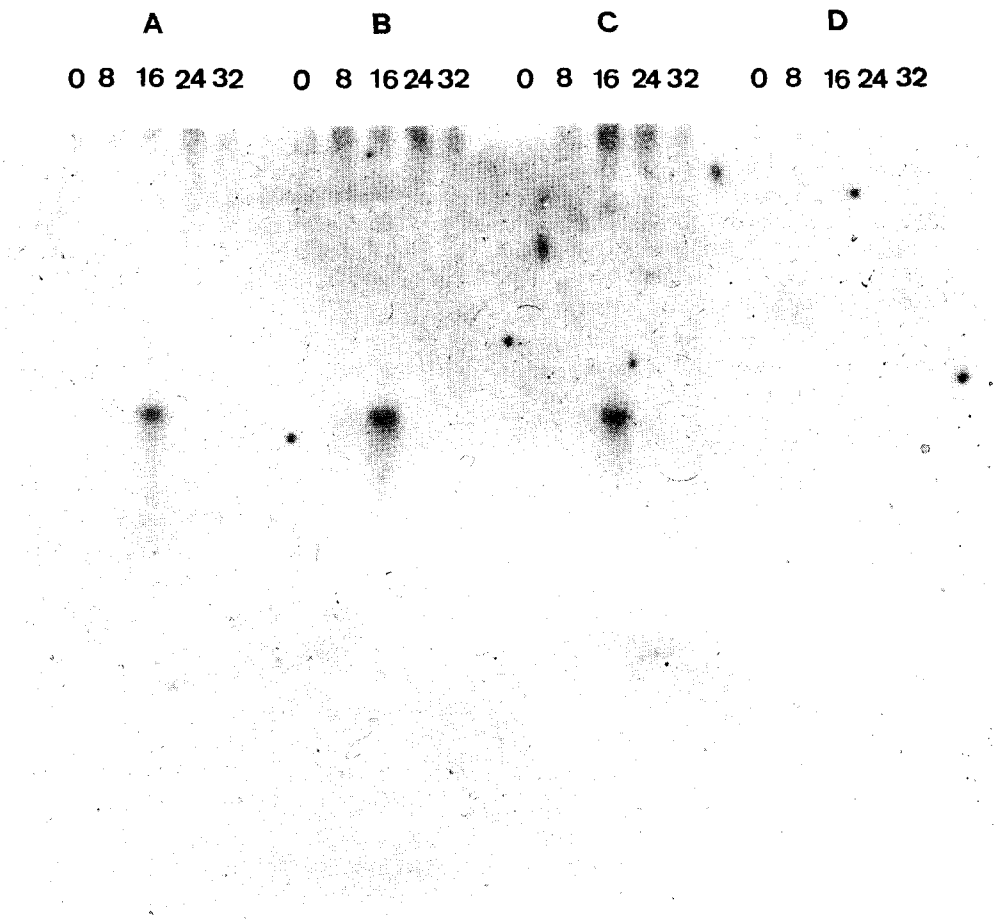

Fig. 1. Effect of carbon sources on the level of *T. longibrachiatum egl1* mRNA. Aliquots containing 50 μg of total RNA were electrophoresed, transferred to a Nylon membrane, and hybridised using the *T. longibrachiatum egl1* cDNA as a probe. Lanes A–C, RNA obtained at different times from three different *T. longibrachiatum* transfers to medium containing lactose and β-methylglucoside as inducers. Lane D, RNA from mycelia obtained from a similar transfer culture in which glucose 1% (w/v) was added 14 h after the transfer. Numbers indicate time (h) after the transfer.

ml^{-1} (wet weight) in 30 ml of induction medium [9] containing 0.05% (w/v) of each of the inducers, lactose and β-methylglucoside (Sigma Co., USA). The samples were incubated for an additional period under the same conditions. For repression experiments, glucose at 1% (w/v) final concentration was added 14 h after the transfer.

DNA manipulations

DNA was sequenced using the dideoxy chain termination protocol [10]. Sequence data appear in the GenBank database under accession number U08857.

RNA manipulations

Total RNA isolation and Northern analysis were carried out as described previously [4]. To determine the transcription start point, a primer corresponding to nucleotides + 7 to + 35 in the *egl1* sequence [4] was annealed to mRNA obtained under inducing conditions and extended with *Tth* polymerase (Boehringer Mannheim, FRG). The 3′ end of the transcript was mapped using a simplification of the rapid amplification of cDNA ends method [11]. mRNA was obtained from mycelia grown under inducing conditions and was reverse transcribed and amplified using the same protocol and oligonucleotides described previously [4]. The resulting cDNA population was digested with *Bam*HI and *Hin*dIII and cloned in pUC18 digested with the same enzymes. Ten different plasmids containing cDNA inserts were sequenced at their 3′ ends to determine the transcription termination points.

Results and Discussion

Transcription of the egl1 gene

Transcription of the *T. longibrachiatum egl1* gene is induced in the presence of 1% (w/v) carboxymethylcellulose as the only carbon source [4]. Several repetitions of this experiment showed that the lag time for the appearance of endoglucanase activity in the culture filtrates varied considerably from one experiment to the other. In addition, the detection of the *egl1* transcript in these culture conditions was not repeatable. We decided to use transfer cultures in order to obtain more reproducible induction conditions. Different inducers and combinations of inducers were tested (not shown) and the most reproducible results were obtained using the induction conditions described in Materials and Methods (i.e. lactose and β-methylglucoside as inducers). In a triplicate experiment under these conditions, endoglucanase activity appeared in the media 16 h after transfer and the maximum level of activity was obtained 8 h later. Total RNA was isolated from mycelia harvested at different times after transfer and the *egl1* transcript detected by Northern analysis using a 1.5-kb *Bam*HI–*Hin*dIII fragment containing the *egl1* cDNA [4] as a probe. As seen in Fig. 1 lanes A–C, the *egl1* transcript was detectable only 16 h after the transfer in the three separate fermentations. Interestingly, under these culture conditions the period for the expression of the *egl1* gene is very short.

Sequence of the egl1 gene promoter and location of the transcriptional start and termination points

The *T. longibrachiatum egl1* gene was originally cloned in a 1.8-kb *Sau*3A fragment that only contained 58 bp upstream of the ATG codon [4]. To obtain more information about this promoter, both strands of a 775-bp *Sph*I–*Hpa*II fragment which includes 722 bp upstream of the ATG codon were sequenced (Fig. 2A). The transcription start point of the *egl1* gene was mapped by primer extension analysis. Extension products corresponding to major transcriptional start points at the thymidine in position −20 and the cytidine at position −24 were observed. As in many fungal genes, the region immediately upstream of this nucleotide is pyrimidine-rich in the coding strand (Fig. 2A). No clear TATA or CAAT boxes were detected upstream of the transcriptional start points but similar results have been described for other filamentous fungal promoters containing pyrimidine rich sequences [12,13]. The translation initiation point has a G at position −3, which is consistent with the general pattern for fungal genes [12].

The 3′ end of the *egl1* transcript was mapped using the PCR technique described in Materials and Methods. There are multiple termination

A

-722 *GCATGCCGCG* GGTGCAGCGA GCCAAAGAAG CAAAATGTGT CGCCGGACCG
-672 TAGTTGGTAT CAAAAGGGGG AACTATTACT TACCTAGATC GGTAGGTATC
-622 TAGCCAGGGG AAATGGATCA AGGTTGCCAC TGACTGTTTC CCCGTCATGC
-572 AATCTTGTAT AGAGCCTGAC CAATCAGGCT AAATCATTGT GTATACTGAG
-522 CCTAGGAAGA CCTGAATTTT GTCCTTGTCG TTCTCTATTA TCCAATTATC
-472 CGTTGATTCG AGCGATGATA CAGTATGATT TTTGCCGAGA TTTTGCTAAA
-422 GGTAGTATCG ATGAGGTTGC AGGATTTAAA TCGAGATAAC GGCTCTAAAC
-372 AGCAGCTAGT AGAAGGAACC CGGACCTCTG CTTCAATACC TTTCATTTTC
-322 CCATTATCCC CTTGCAGCAG AAATACAACG GTAGCATCAG GTCTGTGGTT
-272 GACGGCAAAA AGCATGGCTC TGCTGAGATG CTCCAGGTCA ACCTATCGGG
-222 TCCTGGAGGC TATTTCCAAA GTCTCATAGC CGGACAGAAC AACGGAAGTC
-172 GGCATCTTTT CGACTCGAAC CGATCAGATG GGTTGTCACA TCGAGGCAGA
-122 CGTATGCTGT GCTGAAATCC GACCTTAAAA AGAGGAAGAC TGCAACTTCA
-72 TCTAGCCGTC ATCAGATCGT TCGCCAAGCA CCTTCCCCCC TTAACCTGCT
-22 TGTCCTCCAT AGTTGTCAAG AA*ATG*→

B

2296 *TAG*AGCTTGA CCTTGTGTCT AGCCTGTCCG GACGTGGGTA CTATCATACG
2346 ATGCGGACAA GCAGGGAAAT CGTAGACATT GGGCTGCATA TCAGGAGCAG
2396 AGCATGCTAT ATTGTATCTA CGTTAGCAAA AAGAAAAACA AATGAAAAGG
2446 AAGCTACCTA TCCAGCACTG TACCGAAAAA GCTCATCATG CGTCGGATGG
2496 GTTCACCTGA *TC*

Fig. 2. Nucleotide sequence of the promoter (**A**) and terminator (**B**) of the *T. longibrachiatum egl1* gene. Letters in italics represent restriction endonuclease sites (*Sph*I in (**A**) and *Sau*3A in (**B**)). Letters in bold italics indicate the ATG start (**A**) and the TAG stop (**B**) codons. Bold underlined letters represent the transcriptional start points (**A**) and the transcriptional stop points (**B**). Putative CreA binding sites are double underlined in (**A**).

points (Fig. 2B), of which that corresponding to the A in position + 2429 was the major one (6 out of the 10 clones analysed). There is no consensus AATAAA polyadenylation sequence [12] between the TAG stop codon and the termination points.

Carbon catabolite repression of egl1 expression

In order to demonstrate that the synthesis of EGL1 is under carbon catabolite repression (CCR), a culture was transferred to inducing medium, 14 h after which the transfer culture was supplemented with 1% (w/v) glucose. Endoglucanase activity was not detected in this transfer medium (results not shown). Mycelia harvested at different times after the transfer were used to isolate RNA and the presence of the *egl1* transcript was analyzed as described above. As seen in Fig. 1 (lane D), transcription of the *egl1* gene is clearly repressed by glucose, indicating that CCR of the synthesis of EGL1 is exercised at the level of transcription. Similar CCR transcriptional control has been demonstrated for the *Agaricus bisporus cel1* [14] and *T. reesei cbh1* [2] genes.

In *Aspergillus nidulans* most mutations affecting CCR map to a single gene named *creA* [15]. The *creA* gene has been cloned [16] and sequenced [17], indicating that the *creA* gene product is a DNA-binding protein which acts as a transcriptional repressor. Moreover, the consensus nucleotide sequence for CreA binding has been suggested as 5′-SYGGRG-3′, but not all the possible sites derived from this consensus do in fact bind CreA and the binding of at least some sequences of the form 5′-SYGGAG-3′ is context-dependent in *A. nidulans* as they only bind when S is a G and the site is preceded by AT-rich stretches [18]. It may be noteworthy that in all the cases in which CreA binding in vivo has been demonstrated, two closely spaced CreA binding sites are present [18–20]. Very recently, Espeso and Peñalva [21] demonstrated that CreA binding by some atypical non-consensus sequences is possible if a second, closely linked CreA binding site is present. Two divergently oriented *A. nidulans* CreA consensus sequences in the form 5′-CYGGAG-3′ appear separated by 16 bp in the *T. longibrachiatum egl1* gene promoter (see Fig. 2A). Although these sites have been shown not to bind CreA in vitro in the *prn* [18] and the *alcR* [20] *A. nidulans* promoters, the conservation of these sites in the *T. reesei egl1* gene promoter [22] and the observed CCR of *egl1* transcription could be indicative of an in vivo function. Assessment of the role of these sites and their sequence context may provide information about the putative CreA *Trichoderma* spp. sites.

The results presented here demonstrate that, under our experimental conditions, the synthesis of *T. longibrachiatum* EGL1 endoglucanase is regulated at the level of transcription by at least two mechanisms: induction by the combination of

lactose and β-methylglucoside and repression by glucose which can override the induction.

Acknowledgements

This work has been supported by a grant from the Comisión Interministerial de Ciencia y Tecnología of the Spanish Government (CICYT ALI90-0842). R.G. is the recipient of a fellowship from the Ministerio de Educación y Ciencia. Thanks are due to A.P. MacCabe for critical reading of the manuscript.

References

1 Coughlan, M.P. (1991) Mechanism of cellulose degradation by fungi. Animal Feed Sci. Technol. 32, 77–100.
2 El-Gogary, S., Leite, A., Crivellaro, O., Eveleigh, D.E. and El-Dorry, H. (1989) Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. Proc. Natl. Acad. Sci. USA 86, 6138–6141.
3 Messner, R. and Kubicek, C.P. (1991) Carbon source control of cellobiohydrolase I and II formation by *Trichoderma reesei*. Appl. Environ. Microbiol. 57, 630–635.
4 González, R., Ramón, D. and Pérez-González, J.A. (1992) Cloning, sequence analysis and yeast expression of the *egl1* gene from *Trichoderma longibrachiatum*. Appl. Microbiol. Biotechnol. 38, 370–375.
5 Pérez-González, J.A., González, R., Querol, A., Sendra, J. and Ramón, D. (1993) Construction of a recombinant wine yeast strain expressing a β-(1,4)-endoglucanase activity and its use in microvinification experiments. Appl. Environ. Microbiol. 59, 2801–2806.
6 Sánchez-Torres, P., González, R., Pérez-González, J.A., González-Candelas, L. and Ramón, D. (1994) Development of a transformation system for *Trichoderma longibrachiatum* and its use to construct multicopy transformants for the *egl1* gene. Appl. Microbiol. Biotechnol. 41, 440–446.
7 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
8 Sternberg, D. (1976) Production of cellulases by *Trichoderma*. Biotechnol. Bioeng. Symp. 6, 35–53.
9 Sternberg, D. and Mandels, G.R. (1979) Induction of cellulolytic enzymes in *Trichoderma reesei* by sophorose. J. Bacteriol. 139, 761–769.
10 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.
11 Frohman, M.A., Dush, M.K. and Martin, G.R. (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.
12 Gurr, S.J., Unkles, S.E. and Kinghorn, J.R. (1987) The structure and organization of nuclear genes of filamentous fungi. In: Gene Structure in Eukaryotic Microbes (Kinghorn, J.R., Ed.), pp. 93–139. IRL Press, Oxford.
13 Davis, M.A. and Hynes, M.J. (1991) Regulatory circuits in *Aspergillus nidulans*. In: More Gene Manipulations in Fungi (Bennet, J.W. and Lasure, L.L., Eds.), pp. 151–189. Academic Press, New York, NY.
14 Yagüe, E., Wood, D.A. and Thurston, C.F. (1994) Regulation of transcription of the *cel1* gene in *Agaricus bisporus*. Mol. Microbiol. 12, 41–47.
15 Bailey, C. and Arst, H.N. Jr. (1975) Carbon catabolite repression in *Aspergillus nidulans*. Eur. J. Biochem. 51, 573–577.
16 Dowzer, C.E.A. and Kelly, J.M. (1989) Cloning of the *creA* gene from *Aspergillus nidulans*: a gene involved in carbon catabolite repression. Curr. Genet. 15, 457–459.
17 Dowzer, C.E.A. and Kelly, J.M. (1991) Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. Mol. Cell. Biol. 11, 5701–5709.
18 Cubero, B. and Scazzocchio, C. (1994) Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. EMBO J. 13, 407–415.
19 Hintz, W.E. and Lagosky, P.A. (1993) A glucose-derepressed promoter for expression of heterologous products in the filamentous fungus *Aspergillus nidulans*. Bio/Technology 11, 815–818.
20 Kulmburg, P., Mathieu, M., Dowzer C., Kelly J. and Felenbok, B. (1993) Specific binding sites in the *alcR* and *alcA* promoters of the ethanol regulon for the CREA suppressor mediating carbon catabolite repression in *Aspergillus nidulans*. Mol. Microbiol. 7, 847–857.
21 Espeso, E.A. and Peñalva, M.A. (1994) In vitro binding of the two-finger repressor CreA to several consensus and non-consensus sites at the *ipnA* upstream region is context dependent. FEBS Lett. 342, 43–48.
22 Penttillä, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R. and Knowles, J.K.C. (1986) Homology between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. Gene 45, 253–263.