APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 1995, p. 3741–3744 Vol. 61, No. 10
0099-2240/95/$04.00+0

NOTES

Substrate-Dependent Differential Splicing of Introns in the Regions Encoding the Cellulose Binding Domains of Two Exocellobiohydrolase I-Like Genes in *Phanerochaete chrysosporium*

PAUL R. J. BIRCH, PAUL F. G. SIMS,* AND PAUL BRODA

Department of Biochemistry and Applied Molecular Biology, University of Manchester Institute of Science and Technology, Manchester M60 1QD, United Kingdom

Received 3 April 1995/Accepted 27 July 1995

Recently, we have shown differential splicing of an intron in the *cbhI.2* gene of *Phanerochaete chrysosporium* ME446; this intron lies within the region of the gene encoding the cellulose binding domain (P. F. G. Sims, M. S. Soares-Felipe, Q. Wang, M. E. Gent, C. Tempelaars, and P. Broda, Mol. Microbiol. 12:209–216, 1994). Here, we show that such differential splicing occurs in the *cbhI.1* gene of this fungus as well as in the *cbhI.2* gene and that this phenomenon is substrate dependent. Avicel elicits the synthesis of both classes of mRNA transcripts from both of these genes. In contrast, carboxymethyl cellulose predominantly elicits the synthesis of fully spliced transcripts from both genes. Such differential splicing might allow this fungus to regulate the specificities of substrate binding for these cellulases.

Many bacterial and fungal species produce a number of different cellulases, which act synergistically to effect the cleavage of the β-1,4 bonds of the cellulose chain. The best-characterized cellulolytic fungus, *Trichoderma reesei*, contains genes encoding at least five different activities, two exocellobiohydrolases (CBHI and CBHII) and three endoglucanases (EGI, EGII, and EGV) (11, 13, 14). Each of these proteins is encoded by a single gene and displays a common structural organization, consisting of catalytic and cellulose binding domains (CBD) separated by a region termed the hinge (9, 11).

In nature, cellulose is commonly associated with lignin and hemicellulose to form lignocellulose. Unlike *T. reesei*, the white rot fungus *Phanerochaete chrysosporium* is able to degrade lignin as well as the carbohydrate moieties of lignocellulose (10). These fungi also differ with respect to their cellulolytic systems. *P. chrysosporium* (*Sporotrichum pulverulentum*) has been reported to secrete as many as five proteins with endoglucanase activity (8) and at least three with cellobiohydrolase activity (20). CBHII is encoded by a single gene (19), but unlike that of *T. reesei*, the genomes of *P. chrysosporium* strains contain multiple nonallelic *cbhI*-like sequences (5, 6, 16, 17). However, sequences encoding proteins similar to the *T. reesei* endoglucanases have not been found; we have therefore proposed that in *P. chrysosporium* ME446, variant *cbhI*-like genes encode proteins with endoglucanase activity (17).

The coding regions of both *cbhI.1* and *cbhI.2* of *P. chrysosporium* are interrupted by two introns (17). In the case of *cbhI.2*, cDNA clones in which the 3′ intron was not excised were isolated (17). Translation through this intron changes the reading frame and the location of the translation termination signal of the C-terminal coding region; this would result in a protein with a significantly different CBD and thus, presumably, an altered cellulose binding affinity. The equivalent intron in *cbhI.1* also lacks an in-frame stop codon (16). However, in this case, any differential splicing of the intron would not alter the C terminus of the encoded protein; instead, it would result in a 17-amino-acid insertion in the CBD.

In this report, we show that the intron within the 3′ region of the *cbhI.1* gene, like that of the *cbhI.2* gene, is differentially spliced. We then show that with both genes, the formation of such unspliced mRNA transcripts is substrate dependent.

P. chrysosporium ME446 was maintained on slopes of 2% (wt/vol) malt extract. The culture media for expression studies have been described previously (19). The carbon source used (0.2% [wt/vol]) was either glucose, Avicel (microcrystalline cellulose), ball-milled straw (BMS; lignocellulose), carboxymethyl cellulose (CMC; amorphous cellulose), or cellobiose. The protocols for RNA extraction, poly(A)$^+$ mRNA purification, cDNA synthesis with oligo(dT) primers, and PCR amplification have been described previously (19). The criteria used for the design of gene-specific primer pairs for the *cbhI.1*, *cbhI.2*, and *trpC* genes of *P. chrysosporium* were those of Brooks et al. (4). Details of the primer pairs used are listed in Table 1, and their annealing sites within the relevant genes are illustrated in Fig. 1A. The PCR product derived from the *trpC* gene has previously been used as a control for successful cDNA synthesis and for semiquantitative estimates of genomic DNAs (gDNAs) in such preparations (3) (see below also).

Differential splicing of both *cbhI.1* and *cbhI.2* is substrate dependent. PCR has previously been used to analyze the expression of closely related lignin peroxidase (4) and cellulase genes (19, 21). The expression of the *cbhI.1* and *cbhI.2* genes is coordinately regulated (3, 19), whereas the *cbhII* gene is inde-

* Corresponding author. Mailing address: Department of Biochemistry and Applied Molecular Biology, University of Manchester Institute of Science and Technology, P.O. Box 88, Manchester M60 1QD, United Kingdom.

TABLE 1. Gene-specific sequences of upstream and downstream primers for PCR amplification

Gene[a]	Primer sequence	Name
cbhI.1, u overall	ACA ATG TTC CGC ACT GCT ACT T	G36
cbhI.1, d overall	AGG GTG CCC GCG GAG GTG CC	G32
cbhI.1, u partially spliced	GCA TGG CCA AGT ACC CGA CC	M16
cbhI.1, d partially spliced	ATG AGG AGG TAC TCA CAT GGG	H13
cbhI.2, u overall	CAC TCC TCG CAT TCA CTT GTC T	G34
cbhI.2, d overall	CTG CCG GTC TCG GTC CAG TTG C	G35
cbhI.2, u partially spliced	GCA TGT CGA AGT ACC CCG A	M15
cbhI.2, d partially spliced	GGT CAG CAT TGC TGC TTC TC	G22
trpC, u	CAC GGG CAT CGT GAC GGA TAC	
trpC, d	TGG GTC TTG AGT GTG TAG TGG	

[a] u, upstream; d, downstream.

pendently regulated (19). Previous analyses of *cbhI.1* and *cbhI.2* have employed PCR primers spanning the 5′ intron that interrupts the coding region. Therefore, such investigations of overall expression have not discriminated between mRNA transcripts in which the intron interrupting the 3′ region of either gene is differentially processed, as has been detected by analyses of cDNA clones of *cbhI.2* (17). Here, we have used PCR and additional primers to determine whether the 3′ introns of both *cbhI* genes are differentially spliced and which nutrient conditions elicit this phenomenon.

PCR primer pairs for each gene were designed so that the 3′ primers (H13 for *cbhI.1* and G22 for *cbhI.2* [Fig. 1A]) contain sequences complementary and specific to regions of nucleotide sequence within the corresponding 3′ introns. Hence, these primers can anneal only to cDNA sequences containing the appropriate intron or to gDNA. In our hands, even the most rigorously purified mRNA preparations also contain gDNA; the sensitivity of PCR is such that this gDNA could interfere with the identification and quantitation of mRNA-derived cDNA. To allow simple size differentiations of the amplification products from the cDNA and gDNA versions of each gene, 5′ primers were designed so that, in each case, the resulting PCR products spanned the 5′ intron interrupting the coding region of that gene (M16 for *cbhI.1* and M15 for *cbhI.2* [Fig. 1A]).

To test the specificities of the *cbhI.1* and *cbhI.2* primer pairs, amplification products derived from total gDNA (Fig. 1B, lane 2) were excised from a gel and submitted to restriction analysis. In each instance, the restriction pattern was that expected of the corresponding gene. The specificities of primer pairs were further confirmed by attempted PCR amplifications with clones 3E2D (containing fully spliced *cbhI.1* cDNA [17]) and pBSC4 (*cbhI.2* cDNA derived from an mRNA transcript containing the 3′ intron [17]) as templates. As expected, a PCR amplification product was generated from the pBSC4 template only when *cbhI.2*-specific primers were used. Furthermore, *cbhI.1*-specific primers failed to generate a PCR product from either pBSC4 or 3E2D (Fig. 1B).

cDNA populations were prepared from the mycelia of cultures grown for 4 days in minimal media containing the different carbon sources listed above. PCR with primer pairs specific for partially spliced templates (M16 and H13 or M15 and G22) was then performed on these cDNA populations (Fig. 1B). Only amplification products of the expected sizes were generated from gDNA templates in the case of mycelia grown on glucose, cellobiose, or CMC; thus, no unspliced cDNAs from either gene were detected in these cDNA populations. This had been expected for glucose, as the expression of both genes is repressed by growth on this substrate (3, 17, 19). However, both cellobiose and CMC elicit overall expression of both genes (19) (Fig. 2). We infer from the absence of unspliced cDNA forms in these cases that *cbhI.1* and *cbhI.2* gene expression on cellobiose and CMC yields only mRNA transcripts lacking the 3′ intron.

In contrast, for both genes PCR analyses of cDNA samples prepared from BMS- and Avicel-grown cultures generated products of the expected sizes from partially spliced cDNA templates as well as from gDNA templates. Previously, we have reported the occurrence of *cbhI.2* transcripts containing the 3′ intron in a cDNA library prepared from mycelia grown on BMS (17). However, the results presented here are the first evidence that the *cbhI.1* gene is also differentially spliced when strain ME446 is grown on this substrate; of three such cDNA clones studied previously, all of them lacked the 3′ intron (17). We have also reported that Avicel and BMS elicit similar patterns of expression of both genes (3). The presence of

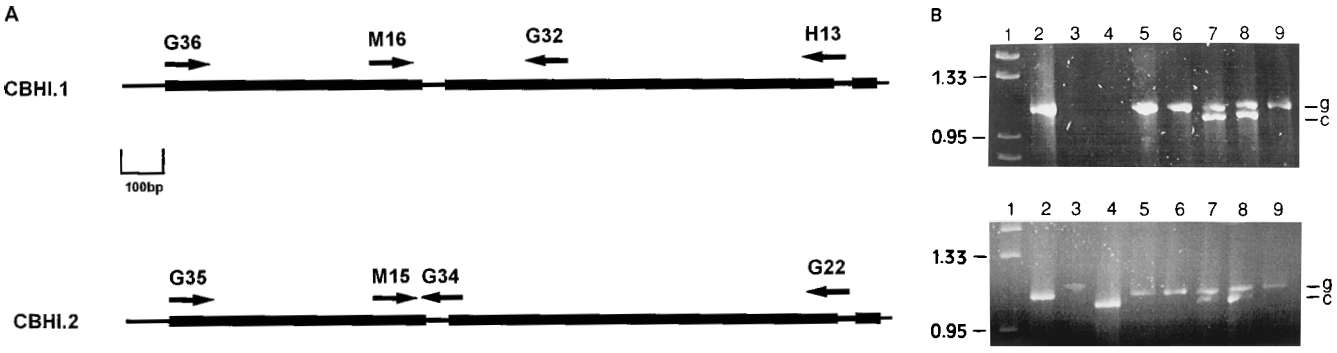

FIG. 1. (A) Diagrams of the *cbhI.1* and *cbhI.2* genomic sequences. Exons, thick lines; introns and flanking sequences, thin lines. The end opposite the point of each arrow indicates the position at which that PCR primer anneals to its target gene. The identity of each oligonucleotide is indicated above the corresponding arrow. (B) PCR amplification products generated from 50 ng of cDNA with primer pairs specific for *cbhI.1* (primers M16 and H13; upper) and *cbhI.2* (primers M15 and G22; lower). The templates used were *P. chrysosporium* gDNA (lanes 2), cloned *cbhI.1* cDNA sequence of 3E2D (lanes 3), cloned *cbhI.2* cDNA sequence of pBSC4 (lanes 4), and cDNA populations prepared from mycelia grown on glucose (lanes 5), cellobiose (lanes 6), BMS (lanes 7), Avicel (lanes 8), and CMC (lanes 9). Lambda DNA digested with *Eco*RI-*Hin*dIII was used as a size marker (lanes 1), and the sizes (in kilobases) of the marker fragments closest in size to the amplification products are on the left. The positions of amplified gDNA (g) and cDNA (c) products are indicated on the right.

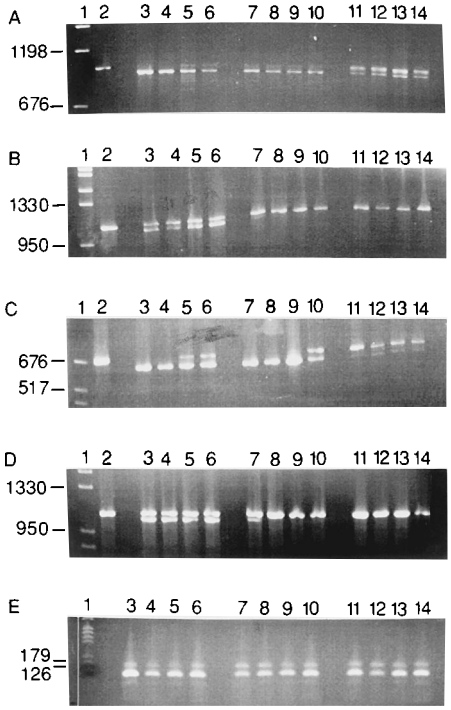

FIG. 2. PCR amplification products generated from 50 ng of cDNA with primers to detect overall *cbhI.1* gene expression (A) (primers G32 and G36), *cbhI.1* sequences containing the 3′ intron (B) (primers M16 and H13), *cbhI.2* overall expression (C) (primers G34 and G35), *cbhI.2* sequences containing the 3′ intron (D) (primers M15 and G22), and *trpC* expression (E). The templates used were total gDNA (lanes 2 [except for panel E]) and cDNA populations prepared from mycelia grown on Avicel (lanes 3 to 6), CMC (lanes 7 to 10), and cellobiose (lanes 11 to 14) for 2 to 5 days. Lanes 1, size markers of either lambda DNA digested with *Hin*dIII-*Eco*RI (B and D) or pGEM DNA (Promega; A, C, and E). The sizes of marker DNA (in base pairs) closest in size to the amplification products are on the left.

unspliced cDNA forms of both genes in cDNA samples from mycelia grown on these substrates is consistent with this observation.

The amplification product of the *cbhI.1* gene, generated from the Avicel-grown cDNA sample, was partially sequenced from an internal primer (G32 [Fig. 1A]) by using an automated sequencer (model 373a; Applied Biosystems). The sequence precisely matched that of *cbhI.1*, confirming that it had been derived from this gene.

Comparisons of the amounts of mRNA transcripts containing the 3′ intron to those of total mRNA for the *cbhI.1* and *cbhI.2* genes. For each gene, we compared the levels of expression of the unspliced mRNA species with the levels of overall expression of both classes of mRNAs in mycelia grown for 2 to 5 days on Avicel, CMC, or cellobiose. Whereas Avicel has been used as a substrate for the synthesis of cellobiohydrolases by *P. chrysosporium* (20), CMC is regarded as a substrate for endoglucanases, and cellobiose has been proposed as an inducer of cellulolytic activity in this fungus (7).

We have previously shown that cDNA samples of the type prepared for this work contain detectable amounts of gDNA (3). The ratios of PCR products derived from gDNA and cDNA *trpC* templates were similar for all of the cDNA preparations analyzed in this work (Fig. 2E). This was expected, as the *trpC* gene of *P. chrysosporium* is known to be constitutively expressed (15), and confirms that the levels of gDNA present in different cDNA preparations are relatively constant (3). In this work, we have sought to compare different transcripts from a single gene within individual cDNA samples. The target cDNA and competitor gDNA differ by only a small intron, and both of them amplify with equal efficiency (results not shown). Since the *trpC* controls show that the proportion of competitor gDNA within all of the cDNA samples is relatively constant, gDNA can be exploited as an internal standard for competitive PCR, allowing semiquantitative interpretation of the results (3). The categories employed are high (amplified cDNA only, indicating that gDNA is outcompeted by cDNA), low (amplified cDNA and gDNA are present in approximately equal amounts), or undetectable (amplified gDNA only) within a single cDNA sample for each PCR.

As observed previously (3, 19), there were high levels of overall expression of the *cbhI.1* and *cbhI.2* genes at early time points after growth on either Avicel or CMC (Fig. 2A and C). However, the nature of the transcripts produced from these genes depends on the substrate. With Avicel at each time point and for both *cbhI.1* and *cbhI.2*, cDNA sequences containing the 3′ intron were amplified (Fig. 2B and D, lanes 3 to 6). However, in all cases, the abundance of this mRNA transcript was less than that of total mRNA for this gene (Fig. 2; compare lanes 3 to 6 in panels A through D); this suggests that with Avicel, both spliced and unspliced classes of mRNA transcripts are formed from both genes. In contrast, after growth on CMC, the unspliced mRNA transcript of each gene was detectable only at a very low level and only on day 2 (Fig. 2B and D, lanes 7). We conclude that with CMC over the period studied, fully spliced mRNA is by far the predominant class of mRNA transcript from both *cbhI.1* and *cbhI.2*.

An alternative explanation for the detection of such differential splicing is that it is an artifact arising from very different overall levels of expression of the cellulase genes on Avicel compared with those on CMC. With both substrates, cDNA predominates over gDNA since cDNA almost completely outcompetes the gDNA used as the internal standard (Fig. 2A and C, lanes 3 to 6 and 7 to 10). Thus, it is possible that the absolute amount of mRNA present (and, therefore, of cDNA formed) after growth on one substrate is much greater than that formed on the other. If the partially spliced transcript is formed as a constant fraction of overall mRNA production, its detection only after growth on Avicel could merely be the result of more transcription on this substrate. To test this hypothesis, we compared the overall abundances of *cbhI.1* mRNA transcripts after growth on Avicel and CMC by competitive PCR. Equal amounts of cDNA (50 ng) prepared after 4 days of mycelial growth were submitted to PCR with primers spanning the 5′ internal intron. Increasing, known amounts of the plasmid construct pBS+1, which contains a *cbhI.1* gDNA insert (16), were added as the competitor. The addition of 10^{-6} μg of

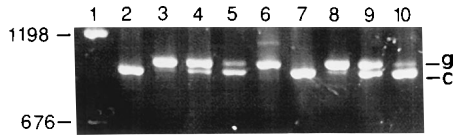

FIG. 3. Competitive PCR analysis of the levels of *cbhI.1* gene expression after growth on either Avicel or CMC. PCR amplification was performed with 50 ng of cDNA prepared after growth on Avicel (lanes 2 to 5) or CMC (lanes 7 to 10). The following amounts of clone pBS+1 were added to cDNA preparations as the competitive template in PCR: 10^{-5} μg (lanes 3 and 8), 10^{-6} μg (lanes 4 and 9), and 10^{-7} μg (lanes 5 and 10). PCR amplification of pBS+1 as a control is shown in lane 6. Lane 1 contains pGEM marker DNA (Promega), with sizes (in base pairs) given on the left. The positions of amplified gDNA (g) and cDNA (c) products are indicated on the right.

plasmid to cDNA prepared after growth on either Avicel or CMC yielded approximately equal amounts of amplified cDNA and gDNA (Fig. 3). Therefore, the levels of *cbhI.1* gene expression elicited by these two substrates were similar. This excludes the hypothesis being tested, and we conclude that differential splicing is indeed substrate dependent.

The expression of the *cbhI* genes on cellobiose was similar to that observed on CMC, since cDNA sequences containing the 3′ intron were not detected for either gene. We infer from the low levels of overall expression of these genes that in this case, only fully spliced cDNA sequences are formed.

Our results are consistent with other evidence that these two genes are coordinately regulated (3, 19). In *P. chrysosporium* K3, the proteins encoded by the equivalents of these genes, CBHI and CBH62, act synergistically. Although both enzymes are active against microcrystalline cellulose, CBH62 is regarded as an endo-type activity (20). The terminal regions of these two proteins suggest that CBHI is the translation product of fully spliced mRNA transcripts of the *cbhI.2*-like gene. In contrast, from present information the CBH62 protein could be derived from either fully spliced or unspliced *cbhI.1* transcripts, since read-through of the intron in the 3′ region does not alter the reading frame and thus the C terminus of the protein is unchanged. No protein sequences have yet been reported to confirm that the unspliced mRNA transcripts described here are translated into active cellulases. However, we can speculate on why such mRNA transcripts are produced and on why their occurrence is substrate dependent.

Since the differential splicing of *cbhI.1* and *cbhI.2* alters the amino acid sequences within the CBD of their protein products, this may be a mechanism by which this fungus can alter the substrate specificities of these cellulases. Changes in protein sequences brought about by proteolytic processing of the terminal regions of fungal cellulases in culture supernatants have previously been proposed as another mechanism for the regulation and adjustment of substrate specificities (11). It is interesting that fungal glucoamylases, which hydrolyze another important polysaccharide (starch), have a domain structure similar to that of cellulases. In the case of the glucoamylase gene of *Aspergillus niger*, an intron within the region of the gene encoding the C-terminal binding domain is also differentially processed, leading to the translation of two distinct glucoamylase enzymes (1, 2). We believe that this gene and the *cox-5* gene of *Neurospora crassa* (12) are the only other reported instances of differential gene splicing in fungi.

Both cDNA classes of each of these genes are being cloned into an heterologous expression system to analyze the mechanism of action and to investigate possible differences in substrate affinities among the four enzymes apparently encoded by the two *P. chrysosporium cbhI* genes. Until we have such information, it is difficult to rationalize the observed differences in intron splicing and to classify such gene products as either endo- or exocellulases. In any case, the latter may not be meaningful since a recent report has suggested that all four of the major *T. reesei* cellulases are able to hydrolyze internal glucosidic bonds (18). As a result, none of them can be referred to as a true exocellobiohydrolase.

REFERENCES

1. **Boel, E., M. T. Hansen, I. Hjort, I. Hoegh, and N. P. Fiil.** 1984. Two different types of intervening sequence in the glucoamylase gene from *Aspergillus niger*. EMBO J. **3:**1581–1585.
2. **Boel, E., I. Hjort, B. Svensson, F. Norris, K. E. Norris, and N. P. Fiil.** 1984. Glucoamylases G1 and G2 from *Aspergillus niger* are synthesised from two different but closely related mRNAs. EMBO J. **3:**1097–1102.
3. **Broda, P., P. R. J. Birch, P. R. Brooks, and P. F. G. Sims.** 1995. PCR-mediated analysis of lignocellulolytic gene transcription by *Phanerochaete chrysosporium*; substrate-dependent differential expression within gene families. Appl. Environ. Microbiol. **61:**2358–2364.
4. **Brooks, P., P. F. G. Sims, and P. Broda.** 1993. Isozyme specific polymerase chain reaction analysis of differential gene expression: a general method applied to lignin peroxidase genes of *Phanerochaete chrysosporium*. Bio/Technology **11:**830–834.
5. **Covert, S. F., J. Bolduc, and D. Cullen.** 1992. Genomic organisation of a cellulase gene family in *Phanerochaete chrysosporium*. Curr. Genet. **22:**407–413.
6. **Covert, S. F., 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.
7. **Eriksson, K.-E., and S. G. Hamp.** 1978. Regulation of endo-1,4-β-glucanase production in *Sporotrichum pulverulentum*. Eur. J. Biochem. **90:**183–190.
8. **Eriksson, K.-E., and B. Pettersson.** 1975. Extracellular enzyme system utilised by the fungus *Sporotrichum pulverulentum* (*Chrysosporium lignorum*) for the breakdown of cellulose. Eur. J. Biochem. **51:**193–206.
9. **Gilkes, N. R., B. Henrissat, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Domains in microbial β-1,4-glycanases: sequence conservation, function, and enzyme families. Microbiol. Rev. **55:**303–315.
10. **Gold, M. H., and M. Alic.** 1993. Molecular biology of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Microbiol. Rev. **57:**605–622.
11. **Knowles, J., P. Lehtovaara, and T. Teeri.** 1987. Cellulase families and their genes. Trends Biotechnol. **5:**255–261.
12. **Sachs, M. S., and C. Yanofsky.** 1991. Developmental expression of genes involved in conidiation and amino acid biosynthesis in *Neurospora crassa*. Dev. Biol. **148:**177–182.
13. **Saloheimo, A., B. Henrissat, A.-M. Hoffren, O. Teleman, and M. Pentilla.** 1994. A novel, small endoglucanase gene, *egl5*, from *Trichoderma reesei* isolated by expression in yeast. Mol. Microbiol. **13:**219–228.
14. **Saloheimo, M., P. Lehtovaara, M. Penttila, T. Teeri, J. Stahlberg, G. Johansson, G. Pettersson, M. Claeyssens, P. Tomme, and J. K. C. Knowles.** 1988. EGIII, a new endoglucanase from *Trichoderma reesei*: the characterisation of both gene and enzyme. Gene **63:**11–21.
15. **Schrank, A., C. A. M. Tempelaars, P. F. G. Sims, S. G. Oliver, and P. Broda.** 1991. The *trpC* gene of *Phanerochaete chrysosporium* is unique in containing an intron but nevertheless maintains the order of functional domains seen in other fungi. Mol. Microbiol. **5:**467–476.
16. **Sims, P., C. James, and P. Broda.** 1988. The identification, molecular cloning and characterisation of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exo-cellobiohydrolase I gene from *Trichoderma reesei*. Gene **74:**411–422.
17. **Sims, P. F. G., M. S. Soares-Felipe, Q. Wang, M. E. Gent, C. Tempelaars, and P. Broda.** 1994. Differential expression of multiple exo-cellobiohydrolase I-like genes in the lignin-degrading fungus *Phanerochaete chrysosporium*. Mol. Microbiol. **12:**209–216.
18. **Stahlberg, J., G. Johansson, and G. Pettersson.** 1993. *Trichoderma reesei* has no true exo-cellulase: all intact and truncated cellulases produce new reducing end groups on cellulose. Biochim. Biophys. Acta **1157:**107–113.
19. **Tempelaars, C. A. M., P. R. J. Birch, P. F. G. Sims, and P. Broda.** 1994. Isolation, characterization, and analysis of the expression of the *cbhII* gene of *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. **60:**4387–4393.
20. **Uzcategui, E., A. Ruiz, R. Montesino, G. Johansson, and G. Pettersson.** 1991. The 1,4-β-D-glucan cellobiohydrolases from *Phanerochaete chrysosporium*. I. A system of synergistically acting enzymes homologous to *Trichoderma reesei*. J. Biotechnol. **19:**271–286.
21. **Vanden Wymelenberg, A., S. Covert, and D. Cullen.** 1993. Identification of the gene encoding the major cellobiohydrolase of the white rot fungus *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. **59:**3492–3494.