0273-2289/96/5758—0389$7.25

Cloning and Expression of Full-Length *Trichoderma reesei* Cellobiohydrolase I cDNAs in *Escherichia coli*

ROBERT A. LAYMON, WILLIAM S. ADNEY, ALI MOHAGHEGHI, MICHAEL E. HIMMEL, AND STEVEN R. THOMAS*

Applied Biological Sciences Branch, Alternative Fuels Division, National Renewable Energy Laboratory, 1617 Cole Blvd., Golden, CO 80401

ABSTRACT

The process of converting lignocellulosic biomass to ethanol via fermentation depends on developing economic sources of cellulases. *Trichoderma reesei* cellobiohydrolase (CBH) I is a key enzyme in the fungal cellulase system; however, specific process application requirements make modification of the enzyme by site-directed mutagenesis (SDM) an attractive goal. To undertake SDM investigations, an efficient, cellulase-free host is required. To test the potential of *Escherichia coli* as a host, *T. reesei* CBH I cDNA was expressed in *E. coli* strain GI 724 as a C-terminal fusion to thermostable thioredoxin protein. Full-length expression of CBH I was subsequently verified by molecular weight, Western blot analysis, and activity on soluble substrates.

Index entries: Cellobiohydrolase I (CBH I); *T. reesei;* cellulases; fusion proteins.

INTRODUCTION

During the simultaneous saccharification and fermentation (SSF) process, ethanol is produced by concurrent saccharification and yeast or bacterial fermentation that utilizes the glucose produced by hydrolyzing the cellulosic content of biomass. The cost of cellulase enzymes and the efficiency of the saccharification process are of great interest not only for separate hydrolysis and fermentation (SHF) applications *(1)*, but also for SSF processes *(2,3)*.

Cellulase enzymes act synergistically on cellulose to produce cellobiose and glucose by cleaving the β-*[1,4]*-D-glucosidic linkages in the polymer chain. CBH I (EC 3.2.1.91; β-*[1,4]*-D-glucan cellobiohydrolase) is a glycoprotein with a polypeptide chain of about 52,214 Dalton and recently has been shown to cleave cellobiose exclusively from the reducing terminus of cellulose chains *(4)*. The importance of this component enzyme of the cellulase system is reflected by the observation

*Author to whom all correspondence and reprint requests should be addressed.

that CBH I and CBH II compose approx 80% of the *Trichoderma reesei* cellulase complex, with CBH I representing approx 75% of the cellobiohydrolase activity *(5)*. β-*(1,4)*-D-Exoglucanases probably perform most of the hydrolytic events during the degradation of cellulose *(6,7)*. Indeed, the major purpose of the endoglucanases seems to be the production of new reducing and nonreducing ends at which cellobiohydrolases can engage. CBH I activity is absolutely crucial to the efficient and complete saccharification of crystalline cellulose by *T. reesei* or synthetic cellulase preparations *(8,9)*.

The first report of successful cloning of *T. reesei* CBH I in yeast was in 1984 from Shoemaker *(10)*. Shoemaker used a yeast-expression plasmid constructed with a phosphoglycerokinase promotor to introduce a functional CBH I gene into *Saccharomyces cerevisiae*. The recombinant (r)CBH I produced in these studies exhibited a specific activity on amorphous cellulose of about 50% of the native enzyme and was heavily *N*-glycosylated. These sugars could be removed by treatment with endoglycosidase H or heat *(10)*. In 1987, van Arsdell et al. *(11)* and Penttilä *(12)* also independently reported the expression of *T. reesei* CBH I in *S. cerevisiae*. The yeast-produced rCBH I from van Arsdell's studies was about 60% as active as the native fungal enzyme (substrate not specified). Consistent with earlier work *(10)*, the rCBH I recovered from both the intracellular and extracellular fractions of yeast reported by Penttilä was a highly glycosylated enzyme (>200 kDa) that showed some activity against phosphoric acid-swollen cellulose *(13)*. Barnett and Shoemaker also noted the overglycosylation of *T. reesei* cellulases expressed in another filamentous fungus, *Aspergillus* sp. *(14)*. Overglycosylated cellulases, however, are undesirable for protein engineering studies, because they may be expected to display nonnative activity characteristics by virtue of their altered physical size and ionic properties *(15)*. This study provides a progress report on our efforts to express CBH I in *E. coli* using a thioredoxin fusion protein and, in doing so, examine this approach as a means of providing active unglycosylated rCBH I for subsequent enzyme engineering studies.

MATERIALS AND METHODS

Isolation of mRNA

Total RNA was isolated from 10 mL of frozen *T. reesei* RUT-C30 (ATCC 56765) cells that were grown at 28°C according to Tangnu et al. *(16)* with 1% w/v Solka-floc as the carbon source and lactose (0.1 wt%) induction. The cells were ground to a fine powder using a mortar and pestle under liquid nitrogen. The powder was then added to denaturing solution (Stratagene, La Jolla, CA: mRNA Isolation Kit) without β-mercaptoethanol. The slurry was then extracted two times with 10 m*M* Tris, pH 8.0, buffer containing 1 m*M* EDTA (TE)-equilibrated phenol and extracted one time with chloroform to remove proteins. (All solutions and glassware are RNase-free, and the techniques used follow commonly practiced RNA extraction and isolation procedures.) The aqueous phase was transferred to a 50-mL conical tube and the RNA precipitated with 8*M* LiCl to a final concentration of 2*M*. The RNA was precipitated overnight at 4°C, then centrifuged (12,000*g*, 10 min, 4°C), and the pellet resuspended in 5 mL diethylpyrocarbonate (DEPC)-treated water. The RNA was transferred to a new conical tube, and 3*M* sodium acetate (pH 5.5) was added to a final concentration of 0.3*M* followed by the addition of 3 vol of 100% ethanol (EtOH)

to precipitate the RNA. The EtOH precipitate was stored for 1 h at –80°C, and the RNA was recovered by centrifugation (as above) and resuspended in 10 mL of room temperature elution buffer (Stratagene, La Jolla, CA).

Construction of a *T. reesei* cDNA Library

A cDNA library was constructed using 10 μg of poly(A)$^+$ mRNA from *T. reesei* as described in the ZAP express cDNA synthesis protocol. Original clones are λ phage clones, which were screened by plaque hybridization on lawns of XL1-Blue MRF′ *E. coli* host cells (Stratagene, ZAP Express cDNA Synthesis Kit).

Isolation of Full-Length CBH I Clones from the *T. reesei* cDNA Library

Oligonucleotide primers (Macromolecular Resources, Fort Collins, CO) specific for the *T. reesei* CBH I coding sequence (5'>ATG TAT CGG AAG TTG GCC GTC ATC<3'; 5'>TTT ACA GGC ACT GAG AGT AGT AAG<3'; *17*) were used to synthesize a digoxigenin (DIG)-labeled polymerase chain reaction (PCR) probe specific for the CBH I gene template (i.e., DNA extracted from the *T. reesei* cDNA library). Labeling with DIG and PCR amplification was performed according to the method described by Perkin Elmer (Norwalk, CT; GeneAmp PCR Reagent Kit with AmpliTaq DNA Polymerase), except that the dTTP was partially replaced with DIG-11-dUTP (Boehringer Mannheim, Indianapolis, IN). Plaque hybridization was performed using the Genius System (Boehringer Mannheim, Genius System User's Guide for Filter Hybridization, Version 2. 0, and the Genius Nonradioactive Nucleic Acid Labeling and Detection Kit). Plasmid (phagemid) DNA from selected clones was recovered using the in vivo excision protocol of the ExAssist/XLOLR system (Stratagene).

Construction of the CBH I Expression Vector

The expression vector pTrxFus (Invitrogen, San Diego, CA) was used to clone and express the rCBH I enzyme in the *E. coli* strain GI 724 as a C-terminal fusion to the thermostable thioredoxin protein. Purification of rCBH I consisted of the isolation of inclusion bodies, followed by their solubilization, renaturation *(18)*, and enterokinase digestion (Invitrogen: ThioFusion Expression System, Version 1.1). Restriction enzymes were obtained from New England Biolabs (Beverly, MA) or Boehringer Mannheim (Indianapolis, IN), and used according to manufacturer's specifications.

Isolation and Purification of Inclusion Bodies (IB) from pTF-CBHI

Inclusion bodies (IB) were produced from an 8-mL culture of pTF-CBH I cells induced after 4 h of incubation. The cells were centrifuged, suspended in 4 mL of TE buffer, and lysed by successive sonication and freeze/thaw cycles. The IBs were separated by centrifugation, and the pellet washed three times with TE buffer. IBs were solubilized overnight at 4°C by the addition of 8*M* urea in TE buffer with 10 m*M* dithiothreitol (DTT). The recombinant CBH I was then renatured by slowly removing the urea by diafiltration against 50 m*M* Tris, pH 8.0, buffer containing 5 m*M* $CaCl_2$ using a 500-mL Amicon stirred cell equipped with a 10,000 mol-wt cutoff (PM10) membrane for a period of 2 d at 4°C.

CBH I Enzyme Assay

The activity of the recombinant enzyme was determined by hydrolyzing the fluorogenic substrate analogs 4-methlyumbelliferyl-β-D-lactopyranoside (MUL) and 4-methlyumbelliferyl-β-D-cellobioside (MUC) (Sigma, St. Louis, MO) at 50°C. Enzyme activity on MUL/MUC was determined by adding a 0.1-mL enzyme aliquot to 0.9 mL of 50 m*M* acetate, pH 5.0, buffer containing 0.5 m*M* MUL/MUC solution. Successive, timed 50-μL samples were removed over a period of 60 min, and the reaction stopped by the addition of 0. 1 mL of 150 m*M* glycine-NaOH, pH 10.3 in a Falcon 96-well, flat-bottom microtiter plate. Fluorescence of 4-methylumbelliferone (MU) was calculated from a series of standards measured with a fluorometer equipped with a 360-nm excitation filter and a 460-nm emission filter (Model 7620; Cambridge Technologies, Cambridge, MA). One unit of activity was defined as that amount of enzyme that releases 1.0 μmol of MU/min at 50°C in the described assay conditions.

RESULTS

T. reesei Poly(A)$^+$ mRNA Isolation and cDNA Library Construction

T. reesei poly (A)$^+$ mRNA was used to construct a cDNA library as described in Materials and Methods. The library was amplified and screened using a digoxigenin-labeled probe. Of the approx 50,000 plaques/plate, an average of 20 positive clones were observed. Positive clones were.isolated after a second screening, and phagemids were then excised (Stratagene). DNA from four of the positive clones was purified and cut with several restriction enzymes to verify identity of the clones. Of these clones, three gave the expected restriction digestion patterns as determined from previously published DNA sequence data *(17)*.

Construction of the Thioredoxin-CBH I Expression System

The strategy for constructing the CBH I expression vector began with sequencing both the 5' and 3' ends of the cDNA (Sequenase Version 2.0 Sequencing Kit, United States Biochemical, Cleveland, OH) of two positive phagemid clones (pB210 and pB291). It was determined, based on the aforementioned DNA sequence data, that these cDNA clones contained the entire CBH I coding sequence, including the signal peptide and 5'-upstream sequences (data not shown). In Fig. 1, the 6.4-kb phagemid, pB210, is shown and includes restriction sites found in the insert DNA and the location of the two primers (P1, P2) used in a subsequent PCR reaction employed in the construction of the expression vector. The sequences of P1 and P2 were as follows:

P1: 5'>gag aga <u>ccc ggg</u> (*XmaI*) CAG TGG CCT GCA TCT C<3'

P2: 5'>aga gag <u>tct aga</u> (*XbaI*) CAA TGG AGA GGC TGT TAC CCG<3'

The upper-case nucleotides in these two oligonucleotides represent those bases that are homologous to the CBH I clone and are used to prime the 270-bp PCR reaction. The lower-case nucleotides were added to facilitate cloning of the PCR product.

This 270-bp fragment was designed to include a unique *XmaI* restriction site at the 5' end of the fragment and a unique *XbaI* site at the 3' end of the DNA fragment,

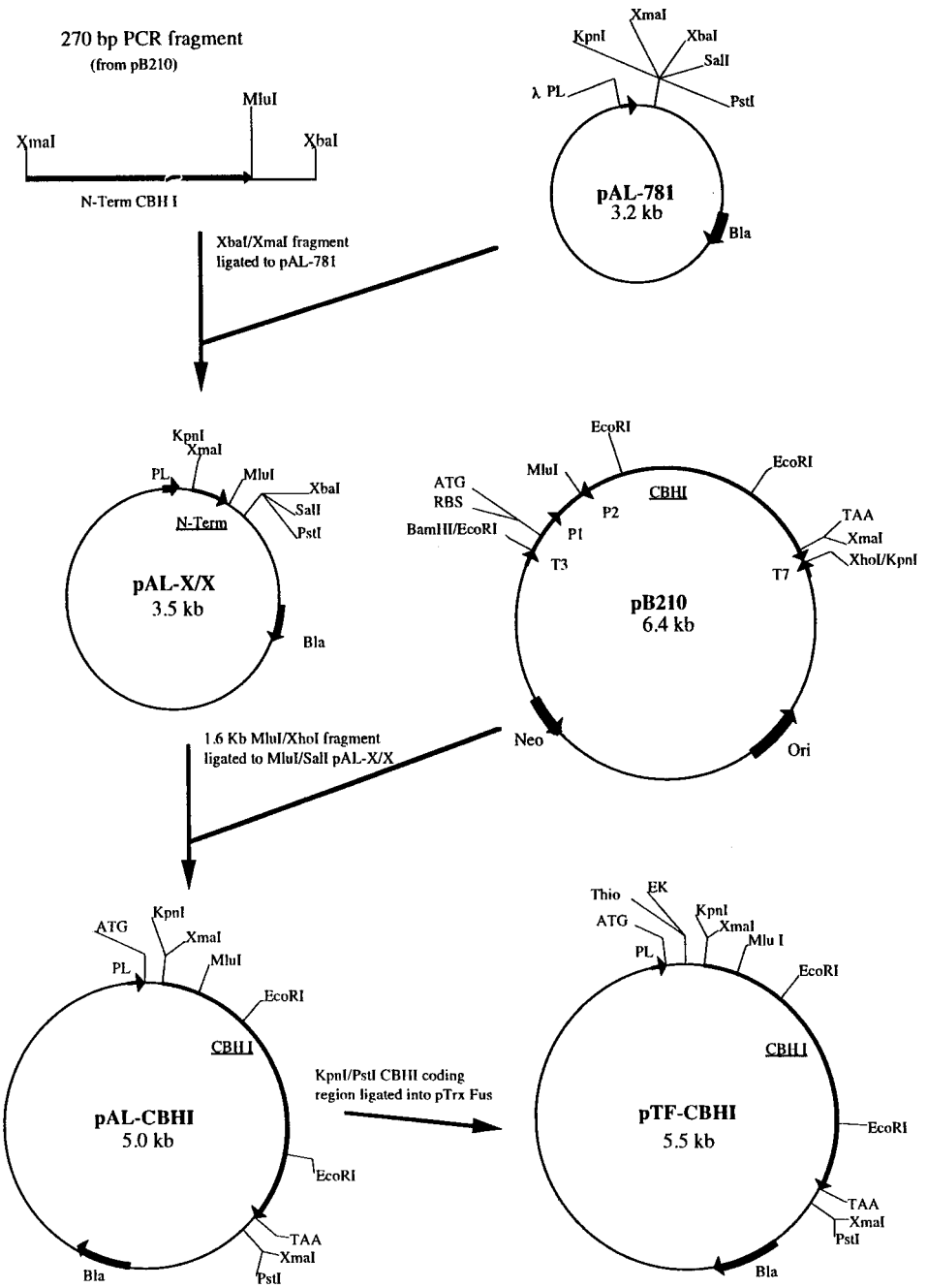

Fig. 1. The 270-bp DNA *Xma*I/*Xba*I fragment was synthesized by PCR using pB210 as a template and ligated into pAL-781, the parent of the pTrxFus expression vector. This small fragment contains the nucleotide coding region for the N-terminal 77 amino acids of CBH I and is inserted in frame with the thioredoxin/enterokinase coding sequence of the vector. The remaining 1.6-kb CBH I coding region was removed from pB210 using *Mlu*I and *Xho*I, and ligated into pAL-X/X cut with *Mlu*I and *Sal*I. The entire coding sequence was then removed from pAL-CBHI with *Kpn*I and *Pst*I and ligated into the same cut pTrxFus vector. The phagemid pB210, isolated from a *T. reesei* λ ZAP library, contains a cDNA encoding CBH I. P1 and P2 represent primers used for the PCR reaction. RBS, ATG, and TAA are the ribosome binding site, initiation codon, and termination codon, respectively. T3 and T7 represent the T3 and T7 promoters, respectively.

which were later used to facilitate construction of expression vectors. This 270-bp PCR fragment encodes the mature N-terminal 77 amino acids of CBH I and was cloned into the *Xma*I/*Xba*I sites in pAL-781, the parent plasmid of pTrxFus (Invitrogen) (*see* Fig. 1).

The phagemid clone, pB210, was cleaved with *Mlu*I and *Xho*I to liberate a 1.6-kb fragment of the *cbh1* cDNA (Fig. 1). This fragment was ligated into pAL-X/X cut with *Mlu*I and *Sal*I to produce pAL-CBHI. The entire CBH I gene was then excised from pAL-CBH I with *Kpn*I and *Pst*I and ligated into pTrxFus (Invitrogen). This construct, pTF-CBHI, was then used to transform *E. coli* GI 724.

Expression of the CBH I Gene Product in *E. coli*

After inducing the cells with tryptophan and a shift in temperature from 30 to 37°C, hourly samples were taken and the cells prepared for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by boiling in 2X sample buffer. As shown in Fig. 2A, 4 h after induction, the rCBH I fusion protein migrates as a prominent band with a calculated mol wt of 68 kDa. Following treatment with enterokinase and analysis using SDS-PAGE, the size of rCBH I was reduced by approx 12 kDa (data not shown).

In Fig. 2B, the SDS-PAGE gel was transferred to a nitrocellulose membrane and probed with an MAb to native *T. reesei* CBH I protein *(19)*. The time points t = 0 and t = 1 h showed no detectable amount of rCBH I, whereas the samples taken at 2 and 4 h after induction reacted strongly with the CBH I antibody. Figure 2B shows anti-CBH I reactive species of approx 68, 58, and 35 kDa. It is likely that the 58- and 35-kDa mol wt forms represent in vivo protease digestion products of rCBH I.

Phase-contrast microscopy indicated that a large fraction of induced cells contained inclusion bodies. This insoluble protein was purified, solubilized, and renatured according to methods described by Marston and Hartley *(18)*.

Activity Levels of the Recombinant *T. reesei* CBH I Protein

Activity of the rCBH I at 50°C on soluble substrates, MUL and MUC, was confirmed using the methods described in the section on CBH I Enzyme Assay. The renatured rCBH I recovered from purified inclusion bodies, both with and without enterokinase treatment, was found to release 1.1 and 4.3 μmol MU/min/mg crude IB protein, from MUL and MUC, respectively. After consideration for the approx 10% rCBH I content of IB protein, these values lie in the same order of magnitude as specific activities found in this study for MUL and MUC hydrolysis using purified native CBH I, i. e., 18.0 and 1.84 μmol MU/min/mg protein, respectively.

DISCUSSION

Several full-length cDNAs encoding the *T. reesei* CBH I protein have been expressed in *E. coli*. The expression system depends on a construct, pTrxFus, which utilizes a tryptophan-inducible λ P_L promoter to drive expression of a thioredoxin-CBH I fusion protein. Unfortunately, the expression product is mostly localized in insoluble inclusion bodies, which must be purified and solubilized before either enterokinase cleavage or refolding. This result is like a problem identified by Teeri when attempting to express CBH I as either an *E. coli* β-galactosidase or λCro peptide-fusion protein in *E. coli (20)*. Their CBH I-fusion proteins, however, showed activity only on a soluble substrate (4-methylumbelliferyl-β-D-lactoside). The poor

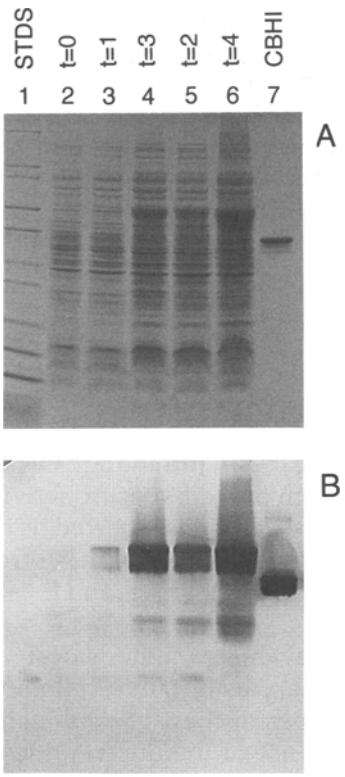

Fig. 2. **(A)** SDS-PAGE of CBH I overexpressed in *E. coli* transformed with pTF-CBHI. Lane 1 was loaded with Novex (San Diego, CA) Mark 12 mol-wt standards: (200, 116, 97, 66, 55.4, 36.5, 31, 21.5, 14.4, and 6 kDa). Cells were grown to early log phase, and then induced with tryptophan at a final concentration of 0.1 mg/mL. After induction, samples were taken every hour, for 4 h (Lanes 2–6) to determine the optimal induction time. The samples were lysed in standard lysis buffer (Invitrogen), boiled for 5 min, and centrifuged before loading. Lane 7 is a purified CBH I-positive control (56 kDa). Times (t) are shown in hours since induction. The gel was stained with Coomassie blue. The overexpressed protein migrates at approx 68 kDa on a 4–20% gradient gel. **(B)** Western blot from a duplicate SDS-PAGE gel loaded as the samples were loaded on the Coomassie-stained gel shown above. The probe is a mouse MAb prepared against purified CBH I.

survivability of nonfused CBH I and other fungal cellulases in *E. coli* cytosol has been documented and is attributed to proteolytic digestion of unprotected, soluble foreign proteins *(20)*. In the present study, the N-terminal thioredoxin was removed from a fraction of the fusion protein sample by digestion with enterokinase (data not shown). Catalytically active enzyme was obtained in reasonable yields following enterokinase digestion; moreover, enzyme recovered from inclusion bodies without subsequent enterokinase treatment was also catalytically active.

Considering that the apparent molecular weight of the *E. coli* derived rCBH I before protease treatment (68 kDa) includes the 12-kDa fusion peptide and that the native CBH I peptide (mol wt = 52.2 kDa *[21]*) is *O*-linked to 6 kDa of carbohydrate content *(21)*, the estimate of rCBH I molecular weight by SDS-PAGE is consistent with the nonglycosylated fusion protein. Whether the differences in migration on SDS-PAGE observed for the native and recombinant enzymes are caused by differ-

Table 1
Comparison of Native and Recombinant CBH I Proteins

	Native CBH I[a]	rCBH I
Molecular weight (SDS-PAGE shown in Fig. 2; kDa)	56	68
Molecular weight of CBH I polypeptide (inferred from this study; kDa)	50	56
Molecular weight of CBH I from DNA sequence (*21*); kDa	52. 2	52. 2
Glycosylation (*21*)	6	None

[a]Purified according to the procedure of Shoemaker et al. (*22*).

ences in glycosylation patterns or in incomplete/incorrect refolding and renaturation that accompany the purification of the rCBH I protein is unclear. However, because yeast-derived rCBH I treated with endoglycosidase H (*10*) and rCBH I expressed in *E. coli* (*21*) retain much of their specific activity on soluble substrates, the refolding steps in the present study are suspected as the major culprit leading to the altered electrophoretic behavior of enterokinase-treated, thioredoxin-fused rCBH I. The characteristics of the purified, native, and *E. coli*-derived CBH I product are in Table 1.

Future experiments should evaluate methods designed to fractionate differently folded forms of rCBH I (perhaps using a thio-cellobioside affinity chromatography), so that the specific activities of these recombinant species, instead of the mixture, can be determined. Because of the uncertainties concerning the tertiary structure of rCBH I produced as a fusion protein in *E. coli* (demonstrated by SDS-PAGE behavior), this expression system is probably not ideal for testing CBH I site-directed mutagenesis concepts.

ACKNOWLEDGMENTS

This work was funded by the Biochemical Conversion Element of the US Department of Energy Biofuels Energy Systems Program.

REFERENCES

1. Wright, J. D., Power, A. J., and Douglas, L. J. (1986), *Biotechnol. Bioeng. Symp.* **17,** 285–302.
2. Wright, J. D., Wyman, C. E., and Grohmann, K. (1988), *Applied Biochem. Biotechnol.* **17,** 75–90.
3. Wilke, C. R., Yan, R. D., and von Stockar, U. (1976), *Biotechnol. Bioeng. Symp.* **6,** 55.
4. Teeri, T. T., Koivula, A., Reinikainen, T., Ruohonen, L., and Srisodsuk, M. (1994), Hydrolysis of crystalline cellulose by native and engineered *Trichoderma reesei* Cellulases, Proceedings of the Symposium on Enzymic Degradation of Insoluble Polysaccharides, The 1994 Annual American Chemical Society Meeting, San Diego, CA.
5. Enari, T.-M. (1983), in *Microbial Enzymes and Biotechnology*, Fogarty, W. M., ed., Applied Science, London, pp. 183–223.
6. Eveleigh, D. E. (1987), *Phil. Trans. Royal Soc. Lond.* **A321,** 435–447.
7. Fägerstam, L. G. and Pettersson, L. G. (1980), *FEBS Lett.* **119,** 97.
8. Irwin, D. C., Spezio, M., Walker, L. P., and Wilson, D. B. (1993), *Biotechnol. Bioeng.* **42,** 1002–1013.
9. Thomas, S. R., Laymon, R. A., Chou, Y.-C., Tucker, M. P., Vinzant, T. B., Adney, W. S., Baker, J. O., Nieves, R. A., Mielenz, J. R., and Himmel, M. E. (1995), in *Advances in the Bioconversion of Lignocellulosics*, Saddler, J. N., and Penner, M. N., eds., ACS Series 618, American Chemical Society, Washington, DC, in press.

10. Shoemaker, S. P. (1984), in *The Cellulase System of* Trichoderma reesei: Trichoderma *Strain Improvement and Expression of* Trichoderma *Cellulases in Yeast*, Online, Pinner, UK, pp. 593–600.
11. van Arsdell, J. N., Kwok, S., Schweickart, V. L., Ladner, M. B., Gelfand, D. H., and Innis, M. A. (1987), *Bio/Technology* **5**, 60–64.
12. Penttilä, M. E. (1987), *Construction and Characterization of Cellulolytic Yeast*, Publications No. 39, VTT, Biotechnical Laboratory, Espoo, Finland.
13. Penttilä, M. E., Andre, L., Lehtovaara, P., Bailey, M., Teeri, T. T., and Knowles, J. K. C. (1988), *Gene* **63**, 103–112.
14. Barnett, C. and Shoemaker, S. P. (1987), in *FEMS Symposium, Biochemistry and Genetics of Cellulose Degradation*, Paris, France, pp. 2–18.
15. Teeri, T. T., Jones, A., Kraulis, P., Rouvinen, J., Penttila, M., Harkki, A., Nevalainen, H., Vanhanen, S., Saloheimo, M., and Knowles, J. K. C. (1990), in Trichoderma reesei *Cellulases*, Kubicek, C. P., Eveleigh, D. E., Esterbauer, H., Steiner, W., and Kubicek-Pranz, E. M., eds., Royal Society of Chemistry, London, UK, pp. 156–167.
16. Tangnu, S. K., Blanch, H. W., and Wilke, C. R. (1981), *Biotechnol. Bioeng.* **23**, 1837–1849.
17. Shoemaker, S., Schweikart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K., and Innis, M. (1983), *Bio/Technology* **1**, 691–696.
18. Marston, F. A. O. and Hartley, D. L. (1990), *Methods Enzymol.* **182**, 264–276.
19. Nieves, R. A., Himmel, M. E., and Ellis, R. P. (1990), *Appl. Biochem. Biotechnol.* **24/25**, 397–406.
20. Teeri, T. T. (1987), Publications No. 38, VTT, Biotechnical Laboratory, Espoo, Finland.
21. Salovuori, I. (1987), Construction and Characterization of Cellulolytic Yeast, Publications No. 34, VTT, Biotechnical Laboratory, Espoo, Finland.
22. Shoemaker, S., Watt, K., Tsitovsky, G., and Cox, R. (1983), *Bio/Technology* **1**, 687–689.