BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Characterization of the catalytic domains of *Trichoderma reesei* endoglucanase I, II, and III, expressed in *Escherichia coli*

Hikaru Nakazawa · Katsunori Okada ·
Ryota Kobayashi · Tetsuya Kubota · Tomoko Onodera ·
Nobuhiro Ochiai · Naoki Omata · Wataru Ogasawara ·
Hirofumi Okada · Yasushi Morikawa

Abstract The genes encoding the catalytic domains (CD) of the three endoglucanases (EG I; Cel7B, EG II; Cel5A, and EG III; Cel12A) from *Trichoderma reesei* QM9414 were expressed in *Escherichia coli* strains Rosetta-gami B (DE3) pLacI or Origami B (DE3) pLacI and were found to produce functional intracellular proteins. Protein production by the three endoglucanase transformants was evaluated as a function of growth temperature. Maximal productivity of EG I-CD at 15°C, EG II-CD at 20°C and EG III at 37°C resulted in yields of 6.9, 72, and 50 mg/l, respectively. The endoglucanases were purified using a simple purification method based on removing *E. coli* proteins by isoelectric point precipitation. Specific activity toward carboxymethyl cellulose was found to be 65, 49, and 15 U/mg for EG I-CD, EG II-CD, and EG III, respectively. EG II-CD was able to cleave 1,3–1,4-β-D-glucan and soluble cellulose derivatives. EG III was found to be active against cellulose, 1,3–1,4-β-D-glucan and xyloglucan, while EG I-CD was active against cellulose, 1,3–1,4-β-D-glucan, xyloglucan, xylan, and mannan.

H. Nakazawa · K. Okada · R. Kobayashi · T. Kubota ·
T. Onodera · N. Ochiai · N. Omata · W. Ogasawara · H. Okada ·
Y. Morikawa (✉)
Department of Bioengineering,
Nagaoka University of Technology,
1603-1 Kamitomioka,
Nagaoka 940-21, Japan
e-mail: yasushi@vos.nagaokaut.ac.jp

Introduction

Cellulose, which is a major constituent of plant cell walls, is the most abundant form of biomass on earth and can be converted into bioethanol and other chemical products via fermentation following hydrolysis to fermentable sugars. Cellulose is a linear homopolymer of D-glucose linked by β-1,4 bonds with a crystalline structure stabilized by intermolecular and intramolecular hydrogen bonds.

Cellulase, consisting of cellobiohydrolase (CBH), endo-glucanase (EG), and β-glucosidase, hydrolyzes the β-1,4-D-glycosidic bonds of cellulose. CBH releases cellobiose from either the reducing or non-reducing ends of the cellulose chain, whereas EG randomly degrades the internal bonds of the amorphous region of cellulose. Cellulolytic organisms are widely distributed among plants, animals, bacteria, and fungi. The filamentous fungus *Trichoderma reesei* is known as a potent cellulase producer because of its remarkable ability to secrete cellulolytic enzymes. As of this writing, it has been reported that the fungus possesses two CBH genes, *cbh1–2*, and eight EG genes, *egl1–8* (Foreman et al. 2003), and that CBH I–II and EG I–VI are secreted proteins. The reason is not so far cleared why it produces many kinds of EG to degrade cellulosic biomass. Although EG I (Cel7B), EG II (Cel5A), and EG III (Cel12A) are well characterized relatively (Biely et al.1991; Macarrón et al. 1993; Bailey et al. 1993; Rose et al. 2002; Hui et al. 2002; Van Arsdell et al. 1987; Aho 1991; Saloheimo et al. 1998; Karlsson et al. 2002a, b; Sandgren et al. 2003; Yuan et al. 2001) and three-dimensional structures for EG I and EG III have also been determined (Kleywegt et al. 1997; Sandgren et al. 2001), the comparison among each EGs, especially that for substrate specificity, have been poorly investigated in spite of the key to understand the role of the enzymes. Use of the

native enzymes purified from *T. reesei* cannot strictly demonstrate their characteristics due to the possibility to contaminate other cellulase constituent. A better understanding of the *T. reesei* cellulose degradation system is expected to be achieved by expressing the individual enzymes in a cellulase non-producing host (e.g. *Eschericha coli*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*). Indeed, expression of *T. reesei* cellulases in these hosts has been evaluated over the past 20 years (Rose et al. 2002; Van Arsdell et al. 1987; Aho 1991, 1996; Robert et al. 1996; Boer et al. 2000; Okada et al. 1998a, b; Penttilä et al. 1987, 1988; Saloheimo et al. 1997; Takashima et al. 1998; William et al. 2003). In some cases, heterologous expression resulted in low yields, proteolytic degradation of the expressed enzymes, formation of inclusion bodies, and high levels of glycosylation. Species of *Aspergillus* have also been used as possible hosts because they produce high amounts of heterologous proteins and possess transcriptional, translational, and post-translational mechanisms similar to those in *T. reesei* (Rose and Van Zyl 2002; Takashima et al. 1998; William et al. 2003).

Recently, improvement of individual cellulase enzymes has been attempted by protein engineering (rational protein engineering or directed evolution) in order to enhance cellulose degradation ability in *T. reesei* (Teter et al. 2004; Day 2003; Wang et al. 2005). It is difficult to obtain the desired enzyme by rational protein engineering because enzyme structure–function relationships are poorly understood. Directed evolution following random mutagenesis appears to be a better approach not dependent on an understanding of structure–function relationships. The use of *Escherichia coli* as a host for directed evolution of cellulase has obvious advantages. *E. coli* is amenable to molecular genetic analysis, grows rapidly, and is transformed at high frequency. To our knowledge, expression of *T. reesei* cellulases in *E. coli* has only been reported for CBH I and EG III. It was reported that slight CBH I activity was detectable when its catalytic core was expressed, but that full-length CBH I was not expressed (William et al. 2003). We reported that EG III was slightly expressed as an

active enzyme in the soluble fraction (4.4 mg/l) in *E. coli*, but most of the recombinant protein was found in an inclusion body (Okada et al. 1998a, b). Refolding of the protein recovered from the inclusion body followed by purification was achieved (Okada et al. 2000). Several endoglucanases from other *Trichoderma* species were also expressed in *E. coli*. Il and coworkers showed that *Trichoderma viride* EG I was expressed in *E. coli* as an inclusion body and that refolding of the recombinant EG I was successful (Il et al. 1999). EG II (C4endoII) from *Trichoderma* sp. C-4 was also expressed in *E. coli*, although the amount of enzyme produced was not described and the enzyme was not purified (Sul et al. 2004).

In this study, we describe efforts to express the catalytic domains of EG I and EG II (EG I-CD and EG II-CD) from *T. reesei* and to increase expression of EG III as active forms in *E. coli*. Enzymatic properties of purified EG I-CD and EG II-CD were characterized and compared to those of EG III.

Materials and methods

Strains, growth media, and vectors

E. coli DH5α was used as the cloning host, and *E. coli* BL21 (DE3), RosettaBlue (DE3) pLacI, Rosetta (DE3) pLacI, Rosetta 2 (DE3) pLacI, Origami B (DE3) pLacI, Origami 2 (DE3) pLacI, Rosetta-gami B (DE3) pLacI, and Rosetta-gami 2 (DE3) pLacI (Novagen) were used as expression hosts for cellulases. Strains were cultivated in Luria–Bertani (LB) medium [1% tryptone, 0.5% yeast extract, and 1% sodium chloride (*w/v*)] or 2× yeast extract tryptone (YT) medium [1.6% tryptone, 1% yeast extract, and 0.5% sodium chloride (*w/v*)], supplemented with suitable antibiotics. Cloning vector pT7Blue T-vector and expression vector pET22b and pET15b were purchased from Novagen (Table 1). We selected either pET22b or pET15b as the expression vector to use appropriate restriction site for insertion of the *egl* genes.

Table 1 Vectors used in this study

Name	Description	Reference
pT7Blueegl1	pT7Blue T-vector with *egl1* cDNA	This study
pT7Blueegl2	pT7Blue T-vector with *egl2* cDNA	This study
pAGmegl3	pAG9–3 with egl3 cDNA except for DNA coding signal sequence	Okada et al. (2000)
pET22b	Expression vector (T7promoter-T7terminater)	Novagen
pET15b	Expression vector (T7promoter-T7terminater)	Novagen
pET15bs	pET15b removing *Bgl*II/*Acc*I 3096bp fragment	This study
pETegl1-cd	pET22b with *egl1-cd*	This study
pETegl2-cd	pET15bs with *egl2-cd*	This study
pETegl3	pET22b with *egl3*	This study

Construction of expression plasmids

Full-length *egl1* and *egl2* reverse transcription polymerase chain reaction (RT-PCR) products derived from a *T. reesei* QM9414 first-strand complementary DNA (cDNA) library were isolated and ligated to pT7Blue T-vector using the same methods as described by Okada et al. (1998a, b). pT7Blueegl1 and pT7Blueegl2 were used as templates for amplifying insert cDNA for ligation into the expression vector (Table 1). To remove sequences encoding the signal peptide, the linker region and C-terminal cellulose binding domain (CBD) in EG I and to allow directional insertion of *egl1* cDNA into *Nde*I and *Hin*dIII sites of the pET expression vector, PCR was performed with a sense primer, 5′-*CATATG*CAGCAACCGGGTACCAGCACCC-3′ and an antisense primer, 5′-*GAAGCTT*ACGCAGTCGAGTTCG TAGTAGACCCAATG-3′. Amplified fragments were digested with *Nde*I and *Hin*dIII and ligated with pET22b digested with the same enzymes to create pETegl1. In the case of *egl2*, sequences encoding the signal peptide, linker region, and N-terminal CBD were removed by performing a PCR to amplify 306 bp of 5′ *egl2* coding region with the sense primer, 5′-**CCATGG**GGGTCCGATTTGCCGGCGT TAAC-3′ and the antisense primer 5′-CATAATTGTG GAT**GTCGAC**GATGCAGTATGCGC-3′. Amplified fragments were digested with *Nco*I and *Sal*I and ligated to 714 bp of the 3′ *egl2* coding region from pT7Blueegl2. The entire fragment encoding the catalytic domain of EG II (EG II-CD) was subcloned into the *Nco*I and *Hin*dIII sites of pET15bs to generate pETegl2. Expression vector pAGmegl3 (Okada et al. 2000) was used as a template to amplify *egl3* cDNA. To allow the directional import of *egl3* cDNA into the *Nde*I and *Hin*dIII sites of the pET expression vector pET22b, PCR was performed with the sense primer, 5′-CG**CATATG**GCA CAAACCAGCTGTGACC-3′, and the antisense primer, 5′-GAGTGCGGCCGC**AAGCTT**-3′. The PCR product was cleaved by *Nde*I and *Hin*dIII and ligated into *Nde*I- and *Hin*dIII-digested pET22b to create pETegl3. Bold letters in the above primer sequences indicate added restriction enzyme sites.

DNA sequencing

PCR products were sequenced using a GenomeLab™Dye Terminator Cycle Sequencing Quick Start kit with CEQ™2000XL DNA sequencer (Beckman coulter). Accession number of the *egl1*, *egl2*, and *egl3* sequences in GenBank is M15665, M19373, and AB003694, respectively.

Enzyme production and purification

Enzymes were produced in the *E. coli* Rosetta-gami B (DE3) pLac I host strain (Novagen). Overnight pre-cultures of each transformant (2.5 ml) were inoculated into 50 ml of LB medium (*egl1-cd*) or 2× YT medium (*egl2-cd*) containing 100 μg/ml ampicillin, 15 μg/ml kanamycin, 30 μg/ml chloramphenicol, and 12.5 μg/ml tetracycline at 15°C (*egl1-cd*), 20°C (*egl2-cd*), and 37°C *egl3* and 120 rpm. Protein expression was induced by adding 1 mM of isopropyl-β-D-thio-galactopyranoside (IPTG), when the OD_{600} reached 0.4–1.0. After induction, the transformant was grown under the same conditions for 48 h (*egl1-cd*), 24 h (*egl2-cd*), and 6 h (*egl3*). Cells were collected and gently disrupted by use of the BugBuster™ Protein Extraction Reagent (Novagen) according to the manufacturer's instructions. Two volumes of 100 mM sodium acetate buffer (pH 4.0) was added to remove *E. coli*-derived proteins. After 1 day, the cell suspension was centrifuged at 16,000×*g* for 20 min. For EG II-CD and EG III, the supernatant containing the purified enzyme was dialyzed by ultrafiltration (PBGC membrane, Millipore) with 10 mM acetate buffer adjusted individually to pH values where the enzymes were stable. For EG I-CD, because residual *E. coli* proteins in the supernatant were abundant due to low expression of EG I-CD, they were concentrated and precipitated using ultrafiltration (PBGC membrane). After removing the pellet by centrifugation, the purified EG I-CD solution was buffer-exchanged with 10 mM acetate buffer (pH 5.0) by ultrafiltration. The purified enzymes were freeze-dried and stored at −20°C. They are stable for at least 2 months.

Enzyme assays and protein analysis

A plate assay was performed by modifying the method described previously (Farkas et al. 1985). *E. coli* was grown on LB plates (pH 7.0) for 1 day at 37°C. Plates were overlaid with 5 ml of a mixture of 0.1% carboxymethyl cellulose (CMC), 0.6% agar, 1 mM IPTG, and 50 mM sodium acetate buffer (at the optimal pH for each enzyme) and were incubated for 24 h to induce enzyme production at 22°C. The enzyme reaction was carried out by incubating the plates for 1 h at 50°C. Plates were then exposed to 1 M Tris–HCl buffer (pH 7.5), after which the buffer was replaced with 0.1% Congo red–0.2% sodium hydrate solution. After incubation for 30 min, the plates were washed with 1 M NaCl to reveal clear zones against a red background developed by hydrolysis of CMC. Plates were rinsed with 5% acetic acid in order to delineate the zones of clearing as much as possible. Recombinant enzyme activity was determined by measuring reducing sugar using the Somogyi–Nelson method (Somogyi et al. 1952; Nelson et al. 1984). The reaction was performed in 50 mM sodium acetate buffer (pH 5.0) at 50°C. pH optima, pH stability, and thermostability were also determined using CMC as substrate. For the determination of pH optima, the
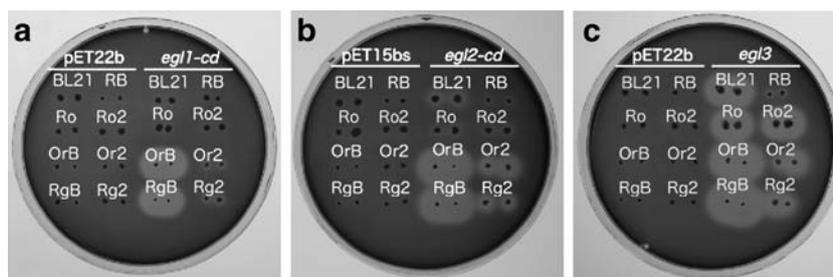
Fig. 1 Detection of endoglucanase activity in *E. coli* transformants by use of the halo assay: EG I-CD (**a**), EG II-CD (**b**), and EG III (**c**). After the transformants were grown on an LB plate for 1 day at 37°C, top agar containing of 0.1% CMC and 1 mM IPTG was added. After induction for 24 h at room temperature (22°C), the halo assay was performed using Congo-red staining. *BL21* BL21 (DE3), *RB* Rosetta blue (DE3) pLacI, *Ro* Rosetta (DE3) pLacI, *Ro2* Rosetta 2 (DE3) pLacI, *OrB* Origami B (DE3) pLacI, *Or2* Origami 2 (DE3) pLacI, *RgB* Rosetta-gami B (DE3) pLacI, *Rg2* Rosetta-gami 2 (DE3) pLacI. Hosts carrying the empty expression vectors pET22b or pET15bs were used as controls

following buffers were used: pH 3–4, 50 mM citrate-sodium hydrate buffer; pH 4–6, 50 mM sodium acetate buffer; and pH 6–8, 50 mM sodium phosphate buffer. pH stability was assessed by measuring residual activity after pre-incubation of the enzyme in the absence of substrate at 50°C for 30 min in 50 mM buffer at the appropriate pH. For measuring thermostability, the enzyme was incubated in 50 mM sodium acetate buffer (pH 5.5) at various temperatures for 15, 30, 45, and 60 min. Substrate specificity was determined using the following substrates: 1.0% CMC (50–200 cps, Sigma), hydroxyethyl cellulose (HEC, 200–300 cps, Nacalai tesque), birchwood xylan (Sigma), tamarind xyloglucan (a gift from Dainippon Sumitomo Pharma and Dr. Yaoi of Advanced Industrial Science and Technology, AIST), barley-β-glucan (1,3–1,4-β-D-glucan, Fulka), lichenan from *Cetraria islandica* (1,3–1,4-β-D-glucan, Sigma), laminaran from *Eisenia bicyclis* (1,3–1,6-

β-D-glucan, Tokyo Chemical Industry), 0.5% Avicel (Funakoshi), and 0.25% phosphoric-acid swollen cellulose (PSC) prepared as described previously (Walseth 1952). One unit of activity was defined as the amount of enzyme that produced 1 μmol of reducing sugars per minute in glucose equivalents. Activity toward *p*-nitrophenyl (*p*NP) derivatives was determined by measuring *p*NP released from the following substrates, *p*NP-β-D-glucopyranoside (*p*NP-G, Sigma), *p*NP-cellobioside (*p*NP-G_2, Sigma), and *p*NP-cellotrioside (*p*NP-G_3, Tronto Research Chemicals). One unit of activity was defined as the amount of enzyme that produced 1 μmol of *p*NP per minute from the substrate. Activity toward azrine-crosslinked (AZCL)-mannan (Megazyme) was determined as described by the manufacturer. Protein concentration was determined by the Lowry method using bovine serum albumin as a standard. The enzymes were assayed at least in duplicate. The data presented in the
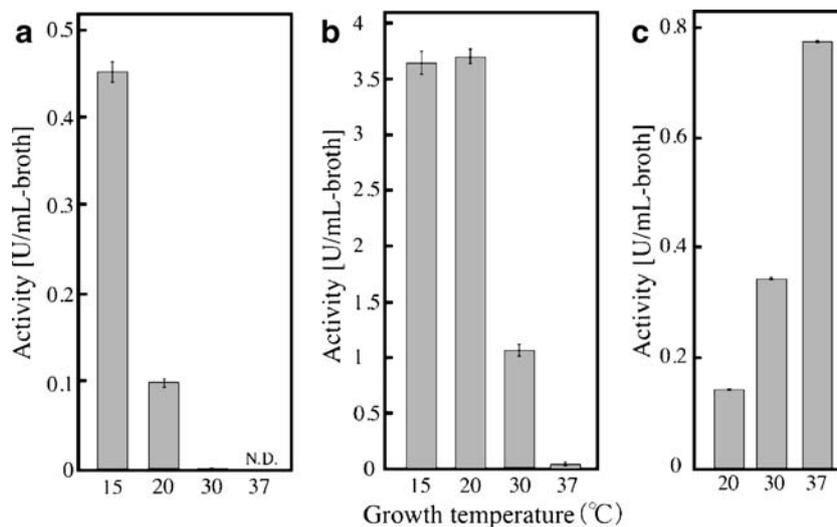

Fig. 2 Effect of growth temperature on endoglucanase production using Rosetta-gami B (DE3) pLacI strain as a host. **a** EG I-CD; **b** EG II-CD; **c** EG III. *E. coli* transformants were grown in 2× YT broth at each temperature. When OD_{600} reached 0.4–1.0, the cultures were induced by addition of 1 mM IPTG. Following induction, cultures were harvested after 6 h at 37°C, 12 h at 30°C, 24 h at 20°C, and 48 h at 15°C. Activity in cell lysates was assayed using CMC as substrate. *N.D.* Not detected

Table 2 Purification of EG I-CD, EG II-CD, and EG III

Purification step	Protein (mg)	Activity (U)	Yield [%]	Specific activity (U/mg)	Purification factor (fold)
EG I-CD					
Cell lysate	300	301	100	1.0	1.0
Soluble fraction at pH 4.0[a]	13	284	94	23	23
Ultra filtration	4.0	259	86	65	65
EG II-CD					
Cell lysate	900	3510	100	3.9	1.0
Soluble fraction at pH 4.0[a]	68	3340	95	49	13
EG III					
Cell lysate	250	1340	100	5.3	1.0
Soluble fraction at pH 4.0[a]	71	1100	82	15	2.8

[a] The soluble fraction at pH 4.0 is the supernatant obtained after centrifugation of the cell lysate mixed with two volumes of 100 mM acetate buffer (pH 4.0).

tables are the mean values. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (1970). Proteins were stained with Coomassie blue. Molecular-mass markers used for SDS-PAGE contained phosphorylase B (97 kDa), albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa), and α-lactalbumin (14.4 kDa; LMW marker kit number 17-0446-01, GE Healthcare).

Results

### Expression of recombinant EG I-CD, EG II-CD, and EG III

The *egl1-cd* and *egl2-cd* genes lacking signal sequences, CBDs, and linkers were subcloned into expression vectors, pET22b for *egl1-cd* and pET15bs for *egl2-cd*. The *egl3* gene was subcloned from the *E. coli* expression vector pAGmegl3 into pET22b. Intracellular expression was tested in the *E. coli* hosts BL21 (DE3), RosettaBlue (DE3) pLacI, Rosetta (DE3) pLacI, Rosetta 2 (DE3) pLacI, Origami B (DE3) pLacI, Origami 2 (DE3) pLacI, Rosetta-gami B (DE3) pLacI, and Rosetta-gami 2 (DE3) pLacI (Novagen). Plate assay based on Congo-red staining is sensitive than the assay based on measurement of reducing sugar. For EG

III produced in cytoplasm of *E. coli*, we previously observed extremely a little enzyme activity to leak outside of the cell in plate assay. Therefore, because the success or failure of the active enzyme expression is confirmed easily, plate assay is employed. A plate assay for recombinant EGs was performed using Congo-red staining, with 0.1% CMC as the substrate when the transformants were grown at room temperature of 22°C (Fig. 1). The largest zone of clearing representing the highest amount of CMCase activity was observed in the Origami B (DE3) pLacI and Rosetta-gami B (DE3) pLacI strains. Based on this assay, B strain derivatives exhibited greater activity than K-12 strain derivatives. The origami strains among them produced the largest zone of clearing. Activity in the soluble fraction of the transformants cultivated using Rosetta-gami B (DE3) pLacI as the host was measured using 1% CMC as the substrate at various cultivation temperatures (Fig. 2). The optimal temperatures for maximal productivity of EG I-CD, EG II-CD, and EG III were 15°C (0.45 U/ml broth), 20°C (3.7 U/ml broth), and 37°C (0.76 U/ml broth), respectively (Fig. 2).

Characterization of the recombinant enzymes

Three endoglucanases expressed in *E. coli* Rosetta-gami B (DE3) pLacI were purified by precipitation of *E. coli*

Fig. 3 SDS-PAGE of EG I-CD (**a**), EG II-CD (**b**), and EG III (**c**). *Lane M* Low molecular mass marker; *lane 1* 5 μl cell lysate; *lane 2* 5 μl insoluble fractions at pH 4.0; *lane 3* 5 μl soluble fractions at pH 4.0; *lane 4* 2 μg of the purified EG I-CD. The *arrow* indicates the position of recombinant EG I-CD, EG II-CD and EG III
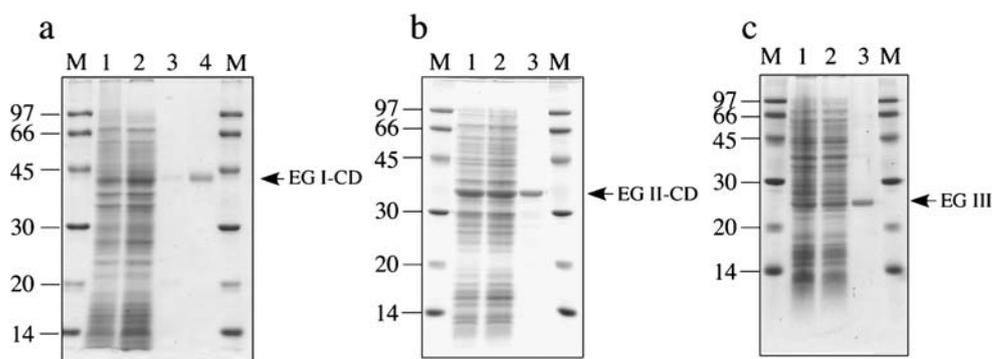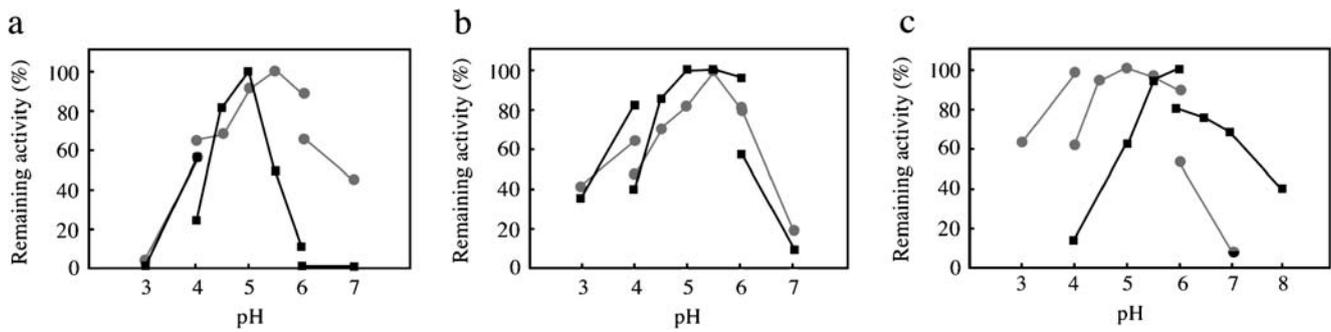
Fig. 4 Enzymatic activity and stability of EG I-CD (**a**), EG II-CD (**b**), and EG III (**c**) at various pH values. *Squares*, pH optima. *Circles* pH stability. pH 3–4, 50 mM citrate-sodium hydrate buffer; pH 4–6, 50 mM acetate buffer; pH 6–7, 50 mM phosphate buffer. To assess pH stability, the enzyme (100 μg/ml) was held at 50°C for 30 min in various buffers prior to addition of substrate. The activity determined prior to incubation was taken as 100%

proteins at pH 4.0. The specific activity of the purified recombinant enzymes toward CMC was estimated to be 65, 49, and 15 U/mg protein in EG I-CD, EG II-CD, and EG III, respectively (Table 2). SDS-PAGE analysis indicated that the molecular mass of purified EG I-CD and EG II-CD was 40 and 35 kDa, respectively, in good agreement with the molecular masses predicted from the amino acid sequences (Fig. 3). The molecular mass of EG III was estimated to be 25 kDa by SDS-PAGE, while the calculated mass is 23 kDa. This result is in agreement with previous reports (Okada et al. 1998a, b, 2000). The pH activity and pH stability profiles for the three EGs were determined at 50°C and from pH 3.0 to 8.0 with CMC as the substrate (Fig. 4). Maximal activity was observed at pH 5.0 for EG I-CD and at pH 5.5 for EG II-CD. The pH optimum for EG III was found to be 6.0, differing from our previously published value of 5.5 (Okada et al. 1998a, b). EG II-CD was active over a broad pH range (4.5–6.0) compared to EG I-CD (pH 4.5–5.0) and EG III (pH 5.5–6.0). EG I-CD and EG II-CD were stable at pH values between 5.0 and 6.0 for 30 min at 50°C (80% remaining activity). EG III retained more than 80% activity at pH 4.5–6.0 after a 30-min incubation at 50°C. Residual activity for these EGs

toward CMC was measured when the enzymes were incubated at temperatures ranging from 4°C to 70°C (Fig. 5). EG I-CD was found to be stable for 60 min at 40°C, but the activity decreased to 70% when incubated for 60 min at 50°C. EG II-CD was stable at 50°C for 60 min, but its activity decreased to 40% after incubation for 15 min at 70°C and to 60% after 60 min at 60°C. EG III was found to be stable for 60 min at 40°C, but its activity decreased to 90% when incubated for 60 min at 50°C. The purified EGs were assayed toward various polymeric substrates (Table 3). The three EGs show a low activity against PSC because of insoluble substrate. Avicelase activity was not detected for EG I or EG II, whereas EG III exhibited minor activity in spite of consisting solely of a catalytic domain. EG I-CD exhibited the broadest specificity among them with a preference for CMC, lichenan, and xyloglucan. Its activity was low toward xylan and mannan. EG II-CD has a preference for lichenan and barley-β-glucan, equivalent activity toward HEC and CMC, and was not active toward other substrates. EG III well cleaved CMC and lichenan, and activity toward HEC, barley-β-glucan, and xyloglucan was about 60–70% compared to that toward CMC. No EG was found to have laminarinase activity (1,3–1,6 glycosidic bond
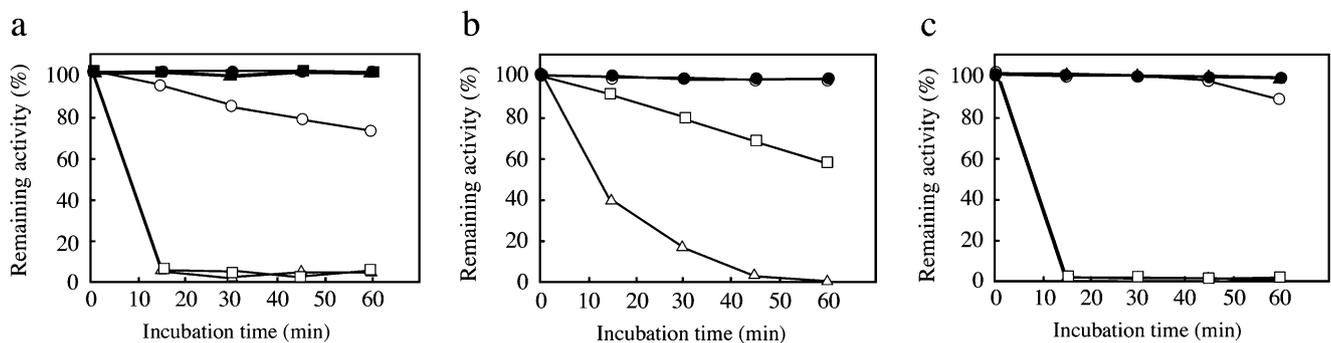

Fig. 5 Thermal stability of EG I-CD (**a**) and EG II-CD (**b**) and EG III (**c**) at various temperatures. *Closed circles* 4°C, *closed squares* 30°C, *closed triangles* 40°C, *open circles* 50°C, *open squares* 60°C, *open triangles* 70°C. The enzyme (100 μg/ml) was held at various temperatures in the absence of substrate for up to 60 min in a buffer at a pH where the enzyme was stable. The activity determined prior to incubation was taken as 100%

Table 3 Hydrolysis activity against polymeric carbohydrates

Substrate	Relative activity (%)		
	EG I-CD	EG II-CD	EG III
Carboxymethyl cellulose (CMC)	100 (65 U/mg)[a]	100 (49 U/mg)[a]	100 (15 U/mg)[a]
Hydroxyethyl cellulose (HEC)	38	83	68
PSC	10	10	12
Avicel	N.D.	N.D.	<1(0.15)
Barley β-glucan	42	90	60
Lichenan	77	113	103
Tamarind xyloglucan	75	N.D.	64
Birchwood xylan	6	N.D.	N.D.
Laminaran	N.D.	N.D.	N.D.
AZCL-mannan	4	<1(0.6)	N.D.

[a] Activity was shown relatively as 100% on CMCase activity

N.D. Not detected

degrading activity). Three EGs are thought to degrade substrates containing β-1,4-glycosidic linkages. Striking differences in specificity among the three EGs was observed with respect to xylanase and xyloglucanase activities. EG II failed to digest any of the substrates, EG III was active toward xyloglucan, and EG I had activity against all substrates. With respect to the *p*NP derivatives, EG I-CD preferentially liberated *p*NP from *p*NP-cellobioside, while both EG II-CD and EG III did so from *p*NP-cellotrioside (Table 4). None of the EGs were able to digest *p*NP-glucoside.

Discussion

In this study, we successfully expressed the catalytic domains of three *T. reesei* endoglucanases in *E. coli* using pET expression vectors and were able to use a plate assay based on Congo-red staining to select transformants. Because the signal sequences were removed for heterologous expression in *E. coli*, the extracellular activity observed in plate assays was unexpected. Although the reason for the extracellular activity is unclear, the size of the zone of clearing around the colonies correlated with activity measured in lysates. For example, large zones of clearing around the Rosetta-gami B and Origami B transformants correlated with greater activity in lysates than that

Table 4 Activity against synthetic substrates

Substrate	Specific activity (mU/mg-protein)		
	EG I-CD	EG II-CD	EG III
*p*NP-glucoside	0.1	0.2	N.D.
*p*NP-cellobioside	490	13	3.4
*p*NP-cellotrioside	43	270	66

N.D. Not detected

from Rosetta Blue transformants showing the small zone of clearing (data not shown). Growing the B strain-derived transformants at low temperature was found to increase activity, whereas growing the K-12 strain-derived transformants at low temperature did not. It is unclear why the B-strain derivative was a better host for *T. reesei* cellulase production than the K-12-strain derivatives. Origami strain series deficient in both thioredoxin reductase and glutathione reductase also produced large zones of clearing. The lack of the two reductases has been reported to be beneficial for intracellular disulfide bond formation in *E. coli*. Furthermore, the use of the origami strain has frequently been reported to have yielded good expression of fungal enzymes for which functional expression has been problematic (Liu et al. 2006; Lorenzo et al. 2005). *T. reesei* EG I-CD, EG II-CD, and EG III have been reported to form 8, 0, and 2 disulfide bonds, respectively. Regardless of the number of disulfide bonds, functional expression of all enzymes was found to be enhanced in the Origami host. Growth at low temperature of *T. reesei* EG I-CD and EG II-CD transformants also resulted in a significant increase in functional expression using the Rosetta-gami B strain, although a considerable amount of inactive enzyme was associated with the inclusion body (data not shown). Furthermore, in order to increase yields of functional *T. reesei* proteins from *E. coli*, a simple purification method was developed. The method was based on the differential precipitation of *E. coli* proteins at pH 4.0 that does not precipitate functional *T. reesei* EG. The differential precipitation method may be first reported, as far as we know, except for a patent (Gehant 2003). Thus, recombinant EGs from *E. coli* were easily separated from native *E. coli* proteins, yielding single bands by SDS-PAGE. These methods [Rosetta-gami B (DE3) pLacI as recombinant host, cultivation at low temperature, and differential precipitation at acidic pH] are expected to greatly facilitate

development of improved *T. reesei* cellulases, through possible use of high-throughput screening of candidate enzymes obtained by directed evolution without need for refolding an inclusion body.

On the basis of the specific activities of the purified enzymes, yields of functional EG I-CD, EG II-CD, and EG III in *E. coli* were estimated to be 6.9, 72, and 50 mg/l broth, respectively. The yield of active EG III was 11 times greater than that of our previous report (Okada et al. 2000). Recombinant EG I-CD had reduced specific activity (65 U/mg) and lower thermostability compared to full-length EG I (104 U/mg) expressed in *A. oryzae* (unpublished data). However, no changes in substrate specificity were observed except for reduced activity toward crystalline cellulose due to lack of CBD, suggesting that catalysis remained the same. A similar result was reported for a full-length EG I from *T. viride* expressed and refolded as a fusion protein in *E. coli* that had reduced specific activity (67.8%) but no changes in transglycosidase activity relative to native *T. viride* EG I (Il et al. 1999). Il et al. speculated that *T. viride* EG I produced in *E. coli* might be destabilized either by non-glycosylation or by fusion. In the present study, EG I-CD was expressed not only as a non-fusion protein but also as a non-glycosylated protein. Thus, its instability might reflect its non-glycosylated state. EG II-CD had the same pH optima and thermal stability as full-length EG II expressed in *S. pombe*, but its pH stability changed slightly from 5.0 to 5.5 (unpublished data). In the present study, the optimal pH for EG III shifted from 5.5 reported by Okada et al. to 6.0. The present results suggest that expression of the three recombinant EGs in *E. coli* provides sufficient activity to serve as a baseline for enzyme improvement. Our evaluation of substrate specificity indicates that EG I have the broadest specificity among the three EGs, whereas EG II possesses the narrowest specificity. EG II degraded amorphous cellulose and β-1,3–1,4-glucan but could not digest hemicelluloses and laminaran. Laminaran is comprised of β-1,3- and β-1,6-glycosidic bonds. Therefore, EG II appears to cleave the β-1,4-linkages in β-1,3–1,4 glucan. EG II was also unable to degrade xyloglucan. Xyloglucan consists of a backbone of β-1,4-linked glucose residues, which are substituted with α-1,6-linked xylose residues. EG II seems to cleave the unsubstituted β-1,4-glucosidic linkages, suggesting that the enzyme is a cellulose-specific EG. Furthermore, EG III degraded xyloglucan in addition to the EG II substrates. Thus, EG III could also cleave substituted β-1,4-glucosidic linkages. EG I was active against all tested substrates except for laminaran. Because the xylan backbone is twisted at torsion angles of 120°, EG I may have a relaxed specificity to recognize a polymer with variable secondary structure. The ability of the enzymes to hydrolyze the *p*NP derivatives was consistent with EG I having four subsites and EG II having five as

previously reported (Biely et al. 1991; Macarrón et al. 1993; Sandgren et al. 2001). Based on currently available information, the distinct roles that EG I and EG II play in cellulose degradation in vivo are still unclear.

To our knowledge, this is the first report describing an expression system for *T. reesei* EG I-CD, EG II-CD, and EG III in a prokaryotic host. Such a system is a key prerequisite for high-throughput screening to improve EGs for bioethanol production or synthesis of bioactive products using transglycosylation activity. *E. coli* as an expression host has major advantages compared to other hosts including *S. cerevisiae*. In addition to lower cultivation costs and more facile gene manipulation, the number of clones that can be screened daily is much higher for *E. coli* than for other hosts. We believe that the present study is an important contribution to the efforts that will be needed to obtain significantly improved EGs by directed evolution or other means.

References

Aho S (1991) Structural and functional analysis of *Trichoderma reesei* endoglucanase I expressed in yeast *Saccharomyces cerevisiae*. FEBS Lett 291:45–49

Aho S, Arffman A, Korhola M (1996) *Saccharomyces cerevisiae* mutants selected for increased production of *Trichoderma reesei* cellulases. Appl Microbiol Biotechnol 46:36–45

Bailey MJ, Siika-Aho M, Valkeajärvi A, Penttilä ME (1993) Hydrolytic properties of two cellulases of *Trichoderma reesei* expressed in yeast. Biotechnol Appl Biochem 17:65–76

Biely P, Vršanska M, Claeyssens M (1991) The endo-1,4-b-glucanase I from *Trichoderma reesei*. Eur J Biochem 200:157–163

Boer H, Teeri TT, Koivula A (2000) Characterization of *Trichoderma reesei* cellobiohydrolase Cel7A secreted from *Pichia pastoris* using two different promoters. Biotechnol Bioeng 69:486–494

Day AG (Genencor International IU) (2003) Novel variant *Hypocrea jecorina* CBH I cellulases, their production with recombinant cells, and their uses. Patent 2,004,016,760

Farkas V, Liskova M, Biely P (1985) Novel media for detection of microbial producers of cellulase and xylanase. FEMS Microbiol Lett 28:137–140

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJM, Yao J, Ward M (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem 278:31988–31997

Gehant RL (Genentech Inc) (2003) Process for protein extraction. Patent 6,967,241

Hui JPM, White TC, Thibault P (2002) Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanase I and II from *Trichoderma reesei*. Glycobiol 12:837–849

Il K, Ekino K, Goto M, Furukawa K (1999) Heterologous expression and characterization of endoglucanase I (EGI) from *Trichoderma viride* HK-75. Biosci Biotechnol Biochem 63:1714–1720

Karlsson J, Momcilovic D, Wittgren B, Schülein M, Tjerneld F, Brinkmalm G (2002a) Enzymatic degradation of carboxymethyl cellulose hydrolyzed by the endoglucanases Cel5A, Cel7B, and Cel45A from *Humicola insolens* and Cel7B, Cel12A and Cel45Acore from *Trichoderma reesei*. Biopolymers 63:32–40

Karlsson J, Siika-Aho M, Tenkanen M, Tjerneld F (2002b) Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EGV) of *Trichoderma reesei*. J Biotechnol 99:63–78

Kleywegt GJ, Zou J, Divne C, Davies GJ, Sinning I, Ståhlberg J, Reinikainen T, Srisodsuk M, Teeri TT, Alwyn Jones T (1997) The crystal structure of the catalytic core domain of endoglucanase I from at 3.6 Å resolution, and a comparison with related enzymes. J Mol Biol 272:383–397

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Liu D, Schmid RD, Rusnak M (2006) Functional expression of *Candida antarctica* lipase B in the *Escherichia coli* cytoplasm-a screening system for a frequently used biocatalyst. Appl Microbiol Biotechnol 72:1024–1032

Lorenzo LD, Aurelio H, Michael H, Uwe BT (2005) Heterologous production of forms of *Rhizopus oryzae* lipase in *Escherichia coli*. Appl Environ Microbial 17:8974–8977

Macarrón R, Acebel C, Castillón MP, Domínguez JM, Mata LDL, Pettersson G, Tomme P, Claeyssens M (1993) Mode of action of endoglucanase III from *Trichoderma reesei*. Biochem J 289:867–873

Nelson NJ (1984) A photometric adaptation of the Somogyi method for determination of glucose. J Biol Chem 375:375–380

Okada H, Sekiya T, Yokoyama K, Tohda H, Kumagai H, Morikawa Y (1998a) Efficient secretion of *Trichoderma reesei* cellobiohydrolase II in *Schizosaccharomyces pombe* and characterization of its products. Appl Microbiol Biotechnol 49:301–308

Okada H, Tada K, Sekiya T, Yokoyama K, Takahashi A, Tohda H, Kumagai H, Morikawa Y (1998b) Molecular characterization and heterologous expression of the gene encoding a low-molecular-mass endoglucanase from *Trichoderma reesei* QM9414. Appl Environ Microbial 64:555–563

Okada H, Mori K, Tada K, Nogawa M, Morikawa Y (2000) Identification of active site carboxylic residues in *Trichoderma reesei* endoglucanase Cel12A by site-directed mutagenesis. J Mol Cat B, Enzymatic 10:249–255

Penttilä ME, André L, Saloheimo M, Lehtovaara P, Knowles JKC (1987) Expression of two *Trichoderma reesei* endoglucanases in the yeast *Saccharomyces cerevisiae*. Yeast 3:175–185

Pentillä ME, André L, Lehtovaara P, Bailey M, Teeri TT, Knowles JKC (1988) Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. Gene 63:103–113

Robert AL, William SA, Ali M, Michael EH, Steven RT (1996) Cloning and expression of full-length *Trichoderma reesei* cellobiohydrolase I cDNAs in *Escherichia coli*. Appl Biochemi Biotechnol 57/58:389–397

Rose SH, Van Zyl WH (2002) Constitutive expression of the *Trichoderma reesei* b-1,4-xylanase gene (xyn2) and the b-1,4-endoglucanase gene (*egl*) in *Aspergillus niger* in molasses and defined glucose media. Appl Microbiol Biotechnol 58:461–468

Saloheimo M, Nakari-Setälä T, Tenkanen M, Penttilä M (1997) cDNA cloning of a *Trichoderma reesei* cellulase and demonstration of endoglucanase activity by expression in yeast. Eur J Biochem 249:584–591

Saloheimo M, Lehtovaara P, Penttilä M, Teeri TT, Ståhlberg J, Johansson G, Pettersson G, Claeyssens M, Tomme P, Knowles JKC (1998) EG III, a new endoglucanase from *Trichoderma reesei*: the characterization of both gene and enzyme. Gene 63:11–21

Sandgren M, Shaw A, Ropp TH, Wu S, Bott R, Cameron AD, Ståhlberg J, Mitchinson C, Alwyn jones T (2001) The X-ray crystal structure of the *Trichoderma reesei* family12 endoglucanase 3, Cel12A, at 1.9 Å resolution. J Mol Biol 308:295–310

Sandgren M, Gualfetti PJ, Shau A, Gross LS, Saldajeno M, Day AG, Jones TA, Mitchinson C (2003) Comparison of family 12 glycoside hydrolases and recruited substitutions important for thermal stability. Protein Science 12:848–860

Somogyi MJ (1952) Notes on sugar determination. J Biol Chem 195:19–23

Sul O, Kim J, Park S, Son Y, Park B, Chung DK, Jeong C, Han I (2004) Characterization and molecular cloning of a novel endoglucanase from *Trichoderma* sp. C-4. Appl Microbiol Biotechnol 66:63–70

Takashima S, Iikura H, Nakamura A, Hidaka M, Masaki H, Uozumi T (1998) Overproduction of recombinant *Trichoderma reesei* cellulase by *Aspergillus oryzae* and their enzymatic properties. J Biotechnol 65:163–171

Teter S, Cherry J, Ward C, Jones A, Harris P, Yi J (Novozymes Biotech IU) (2004) Variants of cellobiohydrolase I from *Trichoderma reesei* with improved properties. Patent 20050048619

Van Arsdell JN, Kwok S, Schweickart VL, Ladner MB, Gelfand DH, Innis MA (1987) Cloning, characterization, and expression in *Saccharomyces cerevisiae* of endoglucanase I from *Trichoderma reesei*. Bio/Technology 5:60–64

Walseth CS (1952) Occurrence of cellulases in enzyme preparations from microorganisms. TAPPI 35:228–233

Wang T, Liu X, Yu Q, Zhang X, Qu Y, Gao P, Wang T (2005) Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. Biomole Eng 22:89–94

William SA, Yat-chen C, Stephen RD, Shi-you D, John OB, Glenna K, Todd BV, Michael EH (2003) Heterologous expression of *Trichoderma reesei* 1,4-b-D-Glucan cellobiohydrolase (Cel7A). Am Chem Soc Chapter 23:403–437

Yuan S, Wu Y, Cosgrove DJ (2001) A fungal endoglucanase with plant cell wall extension activity. Plant Physiol 127:324–333