GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Recombinant expression and characterization of an endoglucanase III (cel12a) from *Trichoderma harzianum* (Hypocreaceae) in the yeast *Pichia pastoris*

W.C. Generoso[1], W. Malagó-Jr.[1], N. Pereira Jr.[2] and F. Henrique-Silva[1]

[1]Laboratório de Biologia Molecular, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, SP, Brasil
[2]Laboratórios de Desenvolvimento de Bioprocessos, Departamento de Engenharia Bioquímica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

Corresponding author: F. Henrique-Silva
E-mail: dfhs@ufscar.br

Genet. Mol. Res. 11 (2): 1544-1557 (2012)
Received February 7, 2012
Accepted April 25, 2012
Published May 21, 2012
DOI http://dx.doi.org/10.4238/2012.May.21.11

ABSTRACT. Filamentous fungi from the genus *Trichoderma* have been widely investigated due to their considerable production of important biotechnological enzymes. Previous studies have demonstrated that the *T. harzianum* strain IOC-3844 has a high degree of cellulolytic activity. After excluding the native signal peptide, the open reading frame of the *T. harzianum* endoglucanase III enzyme was cloned in the expression vector pPICZαA, enabling protein secretion to the culture medium. The recombinant plasmid was used to transform *Pichia pastoris*. Recombinant expression in the selected clone yielded 300 mg pure enzyme per liter of induced medium. The recombinant enzyme proved to be active in a qualitative analysis using Congo red. A quantitative assay, using dinitrosalicylic acid, revealed a high degree of activity at pH 5.5 and around 48°C. This information contributes to our understanding of the cellulolytic repertory of *T. harzianum* and the

determination of a set of enzymes that can be incorporated into mixes for second-generation ethanol production.

Key words: *Trichoderma harzianum*; Second-generation ethanol; *Pichia pastoris*

INTRODUCTION

Plant biomass has emerged as a feasible source for biofuel production due to its relatively low cost and large-scale feedstock procurement (Lynd et al., 2002, 2008). As Brazil is the main producer of sugarcane worldwide, sugarcane bagasse is the country's most abundant source of plant residual biomass and is mainly destined for the production of electricity in boilers (Pessoa-Jr et al., 2005). This bagasse contains about 50% cellulose, 25% hemicellulose and 25% lignin, which can be used as feedstock for different organisms, resulting in value-added products or essentially more sugar for ethanol fermentation, which is important to the sustainability of energy consumption, especially when considering the limited land resources and noncompetitive land use (Pandey et al., 2000; Lynd et al., 2008).

The depolymerization of polysaccharides is the key step in ethanol production from lignocellulosic biomass. Known since 1819, acid treatment is the precursor method of biomass hydrolysis and involves either concentrated acids at low temperatures or diluted acids at high temperatures (Galbe and Zacchi, 2002). The main disadvantage of acid hydrolysis is the generation of fermentation inhibitors (Galbe and Zacchi, 2002). The use of microbial enzymes is an attractive alternative for cellulose depolymerization due to the specificity and efficiency of the enzymes as well as the non-inhibitor feature and rapid hydrolysis (Arantes and Saddler, 2010; Banerjee et al., 2010).

The cellulolytic complex basically involves β-1,4-endoglucanases, which internally cleave the cellulose molecule, cellobiohydrolases, which break down cellulose into cellobiose, and β-glycosidases, which hydrolyze cellobiose into glucose (Foreman et al., 2003). A number of different microorganisms produce cellulases, among which the most often employed are the filamentous fungi of the genus *Trichoderma* (Lynd et al., 2002; Schuster and Schmoll, 2010). *Trichoderma reesei* (*Hypocrea jecorina*) is known to be an effective cellulase producer and is currently the main source of enzymatic cocktails for the degradation of biomass (Schuster and Schmoll, 2010). On the other hand, strains of *Trichoderma harzianum* (*Hypocrea lixii*) are seldom reported as cellulase producers and are more frequently known as control agents against fungal pathogens (Arantes and Saddler, 2010; Banerjee et al., 2010). Nevertheless, recent studies have revealed the potential of this filamentous fungus for large-scale cellulase production and industrial applications (Ahmed et al., 2009; de Castro et al., 2010a,b). In studying the *T. harzianum* IOC-3844 strain, de Castro and collaborators (2010a) found a well-balanced enzymatic complex comparable to commercial cocktails and with greater endoglucanase content, making this strain a promising producer of cellulolytic complex for the hydrolysis of sugarcane bagasse.

The present paper describes the cloning of an endoglucanase from *T. harzianum* IOC-3844 and its heterologous production in the yeast *Pichia pastoris*. This enzyme exhibits a high degree of similarity with endoglucanase III (cel12a) from *T. reesei*. Endoglucanase III (EGIII) enzymes have a low molecular mass and lack the cellulose binding domain (CBD) (Okada

et al., 1998; Henriksson et al., 1999; Nakazawa et al., 2009). These enzymes are also able to degrade amorphous cellulose and have a compact β-jellyroll fold (Okada et al., 1998; Henriksson et al., 1999; Sandgren et al., 2003, 2005; Nakazawa et al., 2009).

The aim of the present study was to determine the proprieties of EGIII from *T. harzianum* IOC-3844 expressed in *P. pastoris* with regard to biochemical and enzyme parameters.

MATERIAL AND METHODS

Strains and media

Escherichia coli DH5α (Invitrogen, USA) was used for the general propagation of plasmids. Bacterial cells were grown with agitation at 37°C in a low-salt Luria-Bertani medium containing Zeocin™ for a final concentration of 25 µg/mL (EasySelect Pichia Expression kit, 2001). *P. pastoris* KM71H(Mut^S) (Invitrogen) was used as the expression host. Yeast cultures were maintained on a yeast extract-peptone-dextrose (YEPD) medium. The media for growth and induction were BMGY (buffered glycerol-complex medium) and with BMMY (buffered methanol-complex medium), respectively, both at pH 6.5 (EasySelect Pichia Expression kit, 2001).

Cloning of endoglucanase ORF and transformation in *P. pastoris*

EGIII cDNA was obtained from a clone of a cDNA library constructed from *T. harzianum* IOC-3844 induced by microcrystalline cellulose (Malagó-Jr. W, Santos-Silva LK, Pereira N Jr. and Henrique-Silva F, unpublished results). This clone was found by comparison using the BLASTx (Altschul et al., 1990) local alignment program and had its open reading frame (ORF) entirely sequenced.

The EGIII ORF was cloned into pPICZαA, excluding the putative signal peptide. The mature ORF was amplified with the forward primer 5'-CGGAATTCCAGACCAGCTGCGAA C-3' and reverse primer 5'-CGCGTCGACGTTGATAGATGCGGTCC-3' (underlined bases indicate introduced restriction sites).

Both the amplicon and pPICZαA plasmid were digested with *Eco*RI and *Sal*I and ligated, creating a recombinant plasmid containing the mature EGIII ORF flanked by the secretion signal peptide (α-factor) at the N-terminal and 6xHisTag at the C-terminal. The pPICZ vectors also contain an inducible promoter from the alcohol oxidase gene AOX1, which is induced by methanol.

Before *P. pastoris* transformation, the recombinant plasmid was linearized with *Pme*I endonuclease and then introduced into the yeast by electroporation (1.5 kV, 25 µF, 200 Ω), following the method described by Creeg (2007). Transformants were cultivated on solid YEPD with 1 M sorbitol and 100 µg/mL Zeocin. Recombinants were identified by polymerase chain reaction using the α-factor and AOX 3'-primers.

Expression, purification and identification of recombinant enzyme

The transformant yeasts were screened for protein induction in 24-well plates, as described by Boettner et al. (2002). A recombinant yeast was selected for enzyme production and purification.

Expression induction for purification was carried out following the manual of the *P. pastoris* expression kit (EasySelect Pichia Expression kit, 2001). A single colony of the recombinant yeast was grown overnight in 5 mL BMGY at 30°C with shaking at 250 rpm. The culture was used to inoculate 500 mL fresh BMGY in a baffled shake flask and was grown at 30°C and at 250 rpm, for 24 h ($OD_{600\,nm}$ = 6). The content was centrifuged at 1500 *g*, for 5 min. The supernatant was discarded and the cells were resuspended in 100 mL BMMY containing 1% methanol. The culture was maintained at 30°C and at 250 rpm, with the daily addition of 0.75% methanol to maintain induction until the best production time. The supernatants were recovered and analyzed using SDS-PAGE (Sambrook and Russell, 2001) and as well as through the qualitative identification of enzyme activity using the Congo red method (Teather and Wood, 1982), with modifications. For such, 20 µL of each supernatant was dispensed on plates containing 1% carboxymethylcellulose (CMC) and 2% agar and incubated for 20 min at 30°C. The plate was then stained with 1% Congo red solution for about 30 min and destained with 1 M sodium chloride until the appearance of degradation halos.

Recombinant enzyme purification from the supernatant was performed by affinity chromatography using a 5-mL silica resin column containing nickel (Ni-NTA Superflow resin, Qiagen, USA). The binding buffer employed consisted of 10 mM Tris base, 50 mM sodium phosphate and 100 mM sodium chloride at pH 8.0. The elution buffer was the same as the binding buffer, except for the addition of imidazole ranging from 5 to 250 mM. The polyhistidine-tagged protein was eluted with a 2-fold resin volume of elution buffer at each imidazole concentration used. The fractions containing purified protein were pooled and dialyzed in membranes of 14,000 MW (Pierce, USA) three times for 4 h at 4°C in 2 L 50 mM Tris-HCl buffer, pH 7.0. The enzyme solution was sterilized by filtration through a 0.22-µm membrane, quantified using the BCA Protein Assay kit (Thermo Scientific, USA) and stored at -20°C.

For the identification of EGIII from *T. harzianum* IOC-3844, the purified protein and supernatant of the fungus were separated on 12% acrylamide gels and were used for Western blotting. *T. harzianum* was induced with 1% microcrystalline cellulose for 36 h, following the method described by Mandels et al. (1962). For Western blotting, SDS-PAGE-separated proteins were blotted onto a polyvinylidene fluoride (PVDF) membrane (Pierce) using the Mini-Trans-Blot Module (Mini-Protean Tetra Cell, Bio-Rad, USA) for 2 h in transfer buffer [20 mM Tris base, 55 mM glycine and 16.5% (v/v) methanol]. The separated proteins were detected with the primary anti-ThEGIII antibody and the secondary anti-mouse IgG-alkaline phosphatase-conjugated antibody (Sigma, Germany), both diluted to 1:10,000 in Tris-buffered saline (TBS, pH 7.4). The BCIP/NBT-Blue Liquid Substrate System (Sigma) was used for detecting alkaline phosphatase.

Anti-Th-EGIII polyclonal antibodies were produced in 20-day-old Swiss albino mice (*Mus musculus*) subjected to two injections of 100 µg purified recombinant *T. harzianum* EGIII. The first immunization was performed using 100 µL protein solution combined with 100 µL Freund's complete adjuvant (Sigma). The second inoculation was performed after 45 days, with 100 µL protein conjugated with 100 µL incomplete Freund's adjuvant (Sigma). The mice were sacrificed after 10 days; blood was collected and centrifuged at 15,700 *g* and at 4°C, for 5 min to obtain the serum containing the antibodies.

CMCase activity assays

CMCase activity was determined using the 3,5-dinitrosalicylic acid method (Miller,

1959) based on the quantification of the reducing sugar content, using glucose as the standard for the calibration curves. All assays were performed in triplicate with 2% low-viscosity CMC (50 to 200 cps, Sigma).

Briefly, 45 µL 2% CMC in 100 mM McIlvaine's buffer, pH 5.5, was incubated in a Peltier bath at 42°C for 15 min; 5 µL enzyme solution was then added to the reaction. The mixture was incubated for the enzyme reaction and the reaction was stopped after 12 min by adding 100 µL dinitrosalicylic acid solution. The mixture was boiled for 5 min, and the reducing sugars were measured at 540 nm with a spectrophotometer (Victor³, Perkin-Elmer, USA).

The optimal temperature for EGIII was determined by altering the reaction temperature within a range of 20° to 50°C in 100 mM McIlvaine's buffer, pH 5.5. The optimal temperature was then employed to determine the optimal pH, altering the reaction buffer within a range from 2.5 to 7.0. For such, the following buffers were used: 100 mM sodium-citrate buffer (pH 2.5 to 4.0), 100 mM McIlvaine's buffer (pH 4.5 to 6.0), 100 mM sodium phosphate buffer (pH 6.5 to 7.0) and 100 mM Tris-HCl buffer (pH 7.5 to 8.0).

Thermostability was analyzed through the measurement of CMCase activity following pre-incubation of the purified enzyme for 1 h in the absence of substrate at temperatures ranging from 0° to 65°C. The thermal resistance of the enzyme was determined by monitoring the residual activity every 24 h of pre-incubation at the last stable temperature (40°C). Residual activity was measured at optimal pH and temperature.

K_{cat} and K_m values were derived from a non-linear Michaelis-Menten curve of enzyme activity, measured by varying CMC concentrations from 3 to 32.4 g/L at the optimal pH and temperature.

All experimental data and graphs were generated with the Prism 5 statistical program (GraphPad, USA). Values are reported as the mean and standard deviation.

RESULTS

Sequence analysis of EGIII in *P. pastoris*

EGIII cDNA is deposited in the GenBank database (accession No. JQ923478) and contains an ORF of 705 bp, which encodes a protein with 235 amino acids. *T. harzianum* EGIII has 81% similarity with the peptide sequence of EGIII from *T. reesei*, 57% similarity with that from *Aspergillus fumigatus* and 47% similarity with that from *Humicola grisea* (for alignments, see Supplementary material). The SignalP program (Bendtsen et al., 2004) predicts a putative signal peptide of 16 amino acids, likely cleaved in the AVA*QT sequence. The predicted peptide sequence has 56.25% similarity with EGIII from *T. reesei*. No N-glycosylation site was predicted using the NetNGlyc program (http://www.cbs.dtu.dk/services/NetNGlyc/).

Expression, purification and identification of EGIII

Several recombinants were analyzed in a small-scale expression assay regarding the production of EGIII (data not shown). The recombinant with the best expression was employed for the production and purification of the recombinant enzyme (Figure 1A). The recombinant enzyme, called rThEGIII, was secreted in an active form by the yeast (Figure 1B). The purification yielded 300 mg enzyme per liter of induced medium.

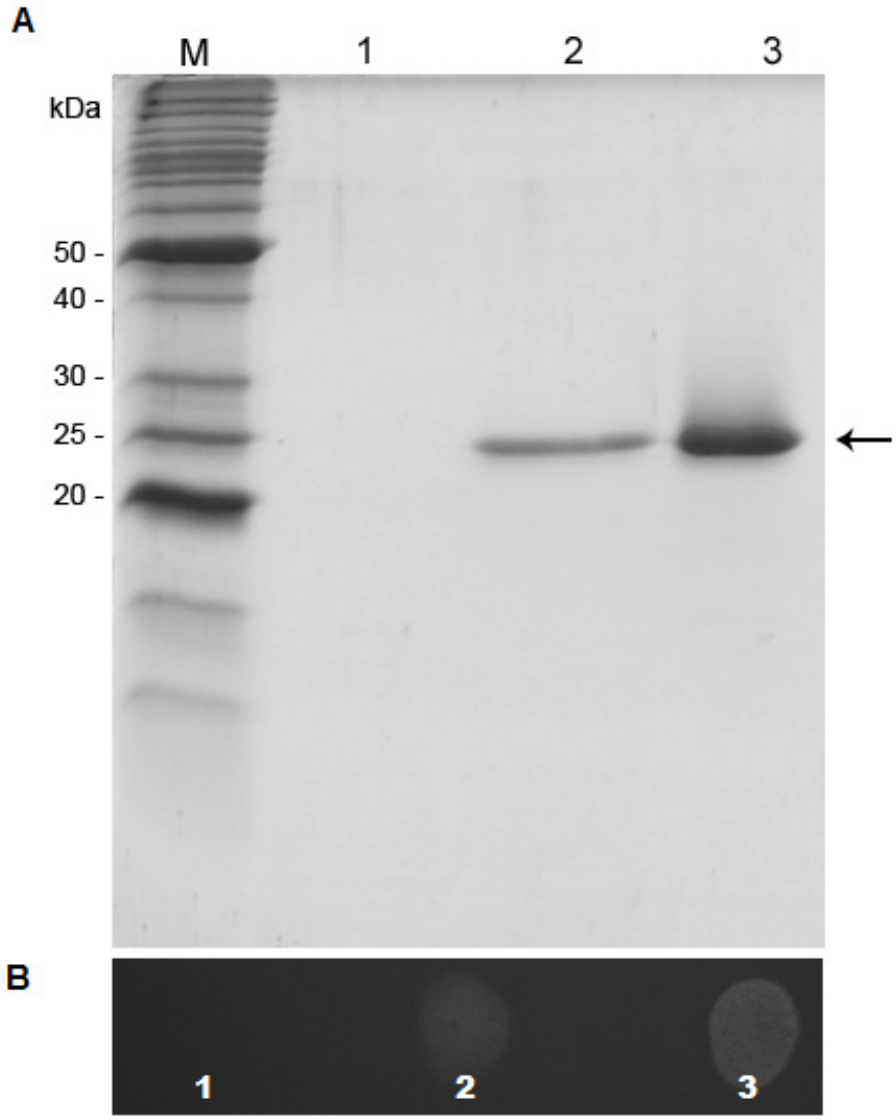


Figure 1. SDS-PAGE analysis of rThEGIII (**A**) and rThEGIII Congo red assay with carboxymethylcellulose (**B**). *Lane M* = BenchMark molecular mass ladder (Invitrogen); *lane 1* = supernatant of recombinant *Pichia pastoris* before induction with methanol; *lane 2* = supernatant of recombinant P. pastoris induced by 1% methanol for 24 h; *lane 3* = rThEGIII purified from supernatant. The arrow indicates the recombinant protein.

The purified recombinant enzyme was used to produce polyclonal antibodies, which were used against the supernatant of *T. harzianum* IOC-3844 induced with 1% microcrystalline cellulose. The molecular mass of the enzyme detected in the fungal supernatant was very similar to that of the recombinant enzyme produced in *P. pastoris* (Figure 2).

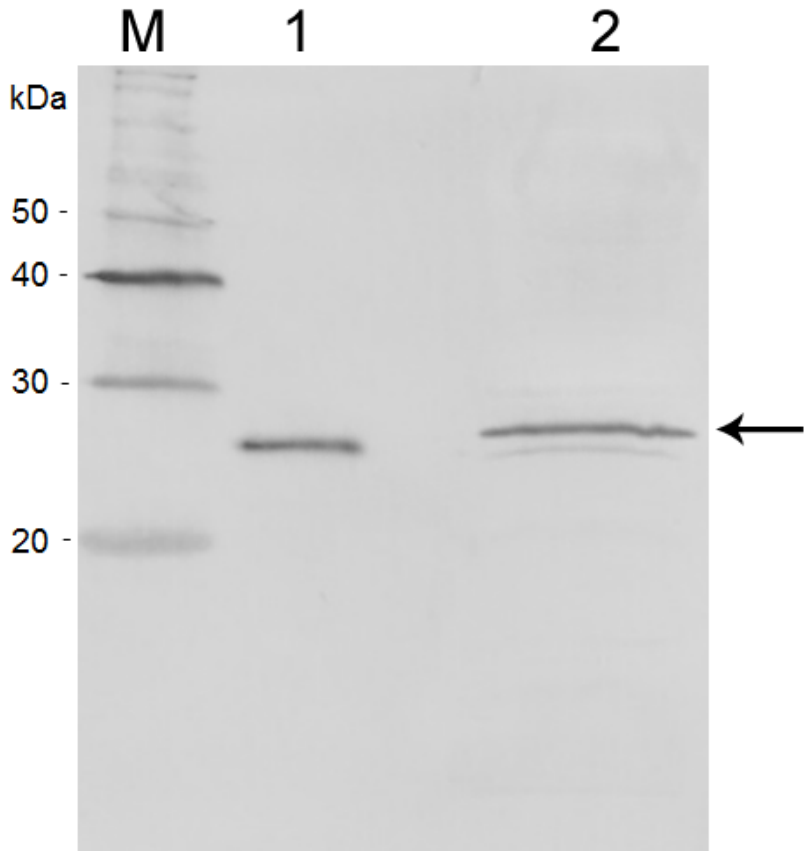


Figure 2. Western blotting analysis of EGIII from *Trichoderma harzianum* IOC-3844. *Lane M* = MagicMark XP Western molecular mass ladder (Invitrogen); *lane 1* = purified rThEGIII; *lane 2* = supernatant of *T. harzianum* induced with 1% microcrystalline cellulose for 36 h. The arrow indicates the recombinant protein.

Characterization of biochemical parameters

For the hydrolytic assays, the enzyme concentration of 0.4 µM was established from previously obtained curves of different enzyme dilutions and reaction times (data not shown). The optimal temperature for rThEGIII was estimated by varying the temperature assays from 21° to 58°C. The best consumption of CMC was achieved at 48.2°C (Figure 3A). Optimal pH was determined within a range of 2.5 to 8.0. The previously determined optimal temperature (48.2°C) was used for the optimal pH tests. The greatest rThEGIII activity was seen at pH 5.5 (Figure 3B).

A thermostability curve was obtained after 1 h of incubation at temperatures ranging from 0° to 65°C. The enzyme was stable up at temperatures up to 40°C (Figure 4A). Considering the decrease in activity over 40°C, enzymatic thermal resistance was evaluated at this temperature for different time intervals. The rThEGIII enzyme exhibited at least 50% activity up to the third day of pre-incubation (Figure 4B).

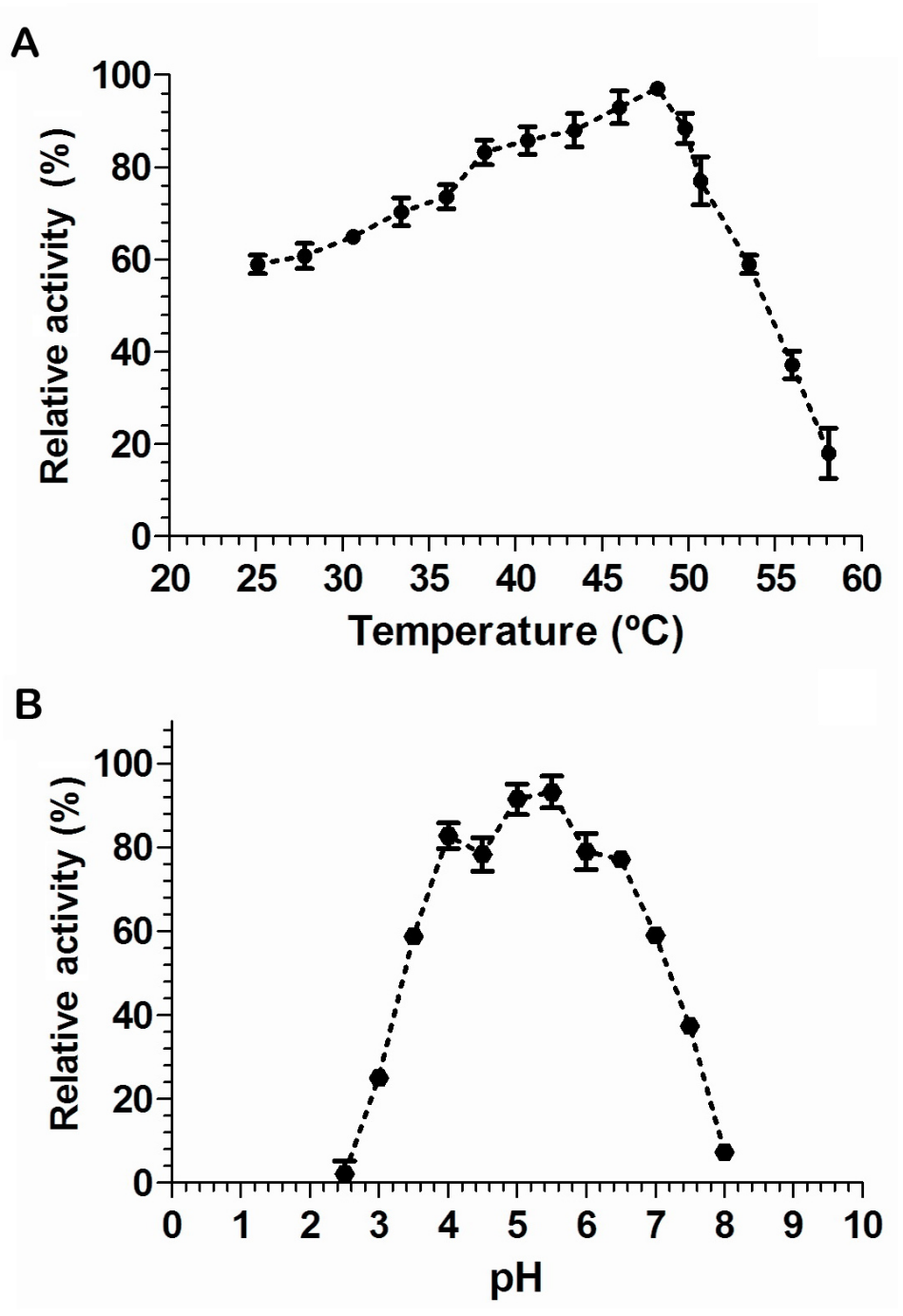


Figure 3. Effect of temperature (**A**) and pH (**B**) on activity of purified rThEGIII (see Methods for reactions and conditions).

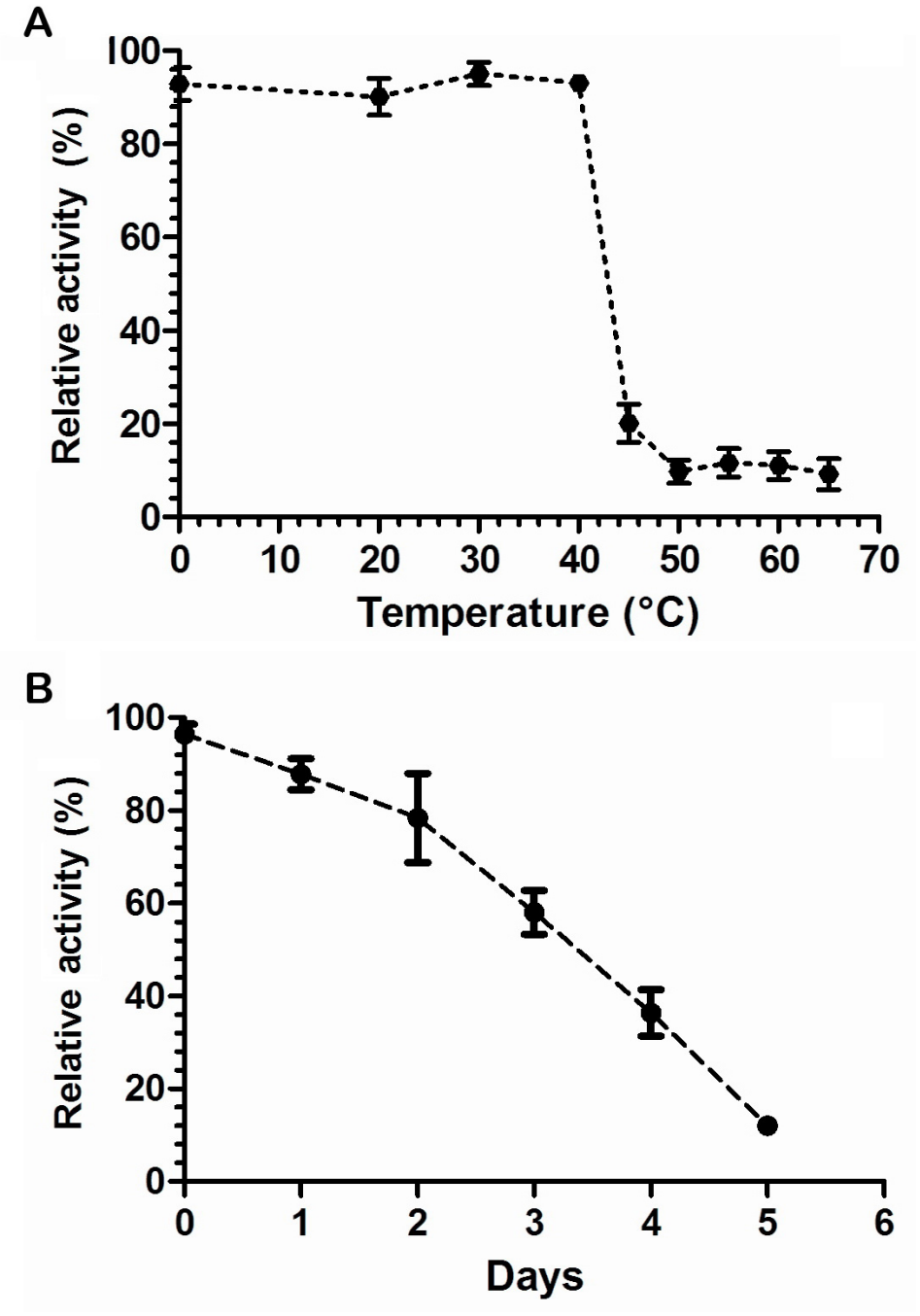


Figure 4. Thermostability of rThEGIII after 1 h of incubation (**A**) and thermal resistance at 40°C (**B**) (see Methods for reactions and conditions).

Characterization of enzyme parameters

For the kinetic characterization of rThEGIII, the CMC concentration ranged from 0.3 to 32.4 g/L. The Michaelis-Menten curve fit revealed a K_m of 21.35 g/L and a turnover number of 1880 min^{-1}. These data are displayed in Figure 5 and Table 1.

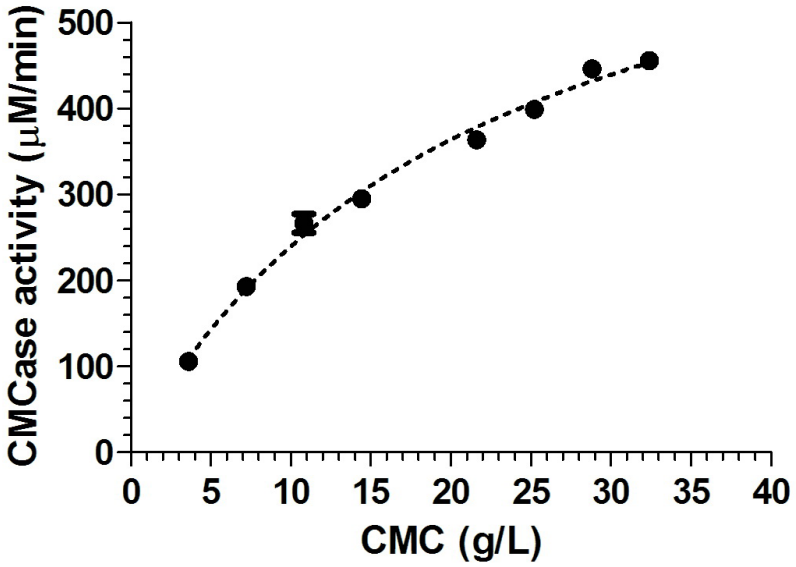


Figure 5. Michaelis-Menten curve fit of rate of reduction sugar generated (μM/min) versus carboxymethylcellulose (CMC) concentration (g/L).

Table 1. Kinetic enzymatic parameters of rThEGIII.

	K_m (g/L)	K_{cat} (min^{-1})	K_{cat}/K_m ($L \cdot g^{-1} \cdot min^{-1}$)
Value	21.35	1880	88.06
Standard deviation	± 1.87	± 82	

DISCUSSION

There is a growing body of research on the isolation and enhancement of enzymes related to the degradation of plant biomass for the development of feasible methods regarding the production of second-generation ethanol. In the present study, a β-1,4-endoglucanase cDNA from the *T. harzianum* strain IOC-3844 was isolated and heterologously expressed in the yeast *P. pastoris*. The recombinant enzyme was characterized considering its possible application in the process of cellulose depolymerization.

The β-1,4-endoglucanase studied has a high degree of similarity with EGIII from *T.*

reesei. EGIII from *T. harzianum* belongs to glycosyl hydrolase family 12 and this enzyme can therefore also be described as cel12a (Foreman et al., 2003). Glycosyl hydrolase family 12 comprises hydrolytic enzymes that cleave glycosidic bonds between two carbohydrates or a carbohydrate from a non-carbohydrate moiety. The cellulases from this family have no CBD and do not easily access crystalline cellulose (Macarron et al., 1993; Okada et al., 1998). However, these enzymes contribute to the breakdown of biomass by hydrolyzing amorphous cellulose and acting synergistically with expansin-like proteins (Arantes and Saddler, 2010).

In the present study, the *P. pastoris* system was employed for the heterologous expression of the identified *T. harzianum* EGIII. The use of this yeast host is recommended primarily due to its simplicity and proximal enzymatic machinery in comparison with *T. harzianum*. Thus, the recombinant enzyme could have similar characteristics as the native one. Another advantage in using *P. pastoris* is the low amount of native-secreted proteins (Cereghino and Cregg, 2000), as illustrated in Figure 1A. For the industrial use of rThEGIII, the purification step could be unnecessary. However, due to the low degree of production of native cellulolytic enzymes of *P. pastoris*, the purification step was necessary for the characterization of the recombinant enzyme.

A large amount of rThEGIII was produced after 24 h of methanol induction. In approximately 48 h, about 300 mg purified rThEGIII can be obtained from 1 L of induced medium. This high yield, which can be further increased in industrial fermenters, is important for the competitiveness of biomass ethanol.

The estimated molecular mass for rThEGIII in SDS-PAGE was about 25 kDa, which is very similar to the predicted mass of 24.6 kDa (Figure 1A). This low molecular mass is also reported for cel12a cellulases and could be related to the absence of CBD and N-glycosylation sites, which are also characteristics of cel12a, unlike other endoglucanases with a high molecular mass (Kwon et al., 1999; Sandgren et al., 2005).

The optimal pH and temperature for rThEGIII were 5.5 and 48.2°C, respectively. This optimal pH is quite similar to that reported for EGIII from some filamentous fungi, such as *T. reesei*, *T. koningii*, *Hypocrea schweinitzii* and *H. grisea* (Karlsson et al., 2002; Sandgren et al., 2005; Wang et al., 2005). On the other hand, the optimal temperature for rThEGIII is not similar to that for EGIII from other filamentous fungi. The optimal temperature for EGIII from *H. schweinitzii* and *Gliocladium roseum* is about 47°C, whereas the optimal temperature for *T. reesei*, *T. koningii* and *Fusarium javanicum* is about 55°C, and finally nearly 70°C is reported for *H. grisea* (Karlsson et al., 2002; Lynd et al., 2002; Sandgren et al., 2005; Nakazawa et al., 2009).

Analyzing the crude extract of *T. harzianum* IOC-3844, de Castro et al. (2010a) found a high degree of endoglucanase activity at temperatures around 52°C and pH 5.1. However, at about 48°C (the optimal temperature of rThEGIII), high endoglucanase activity was also detected in the fungal extract (de Castro et al., 2010a). This difference between the crude extract and the purified rThEGIII could be explained by the existence of other endoglucanases and the synergy among them within the fungal culture supernatant. Analyzing an endoglucanase from *T. harzianum* ETS 323 with molecular mass similar to that of rThEGIII, Liu et al. (2010) found an optimal temperature of 50°C, which is very similar to that of rThEGIII.

These optimal features are interesting, since, under suitable conditions of simultaneous saccharification and fermentation, an enzyme with an optimal temperature closer to 37°C and pH 5.5 seems to be more appropriate (Vasquez et al., 2007). Even with a lower degree of

thermostability than the EGIII from *T. reesei*, which displays a high degree of activity until 50°C (Nakazawa et al., 2009), rThEGIII demonstrated stability at temperatures close to the optimal, lasting several days with fairly good activity. Thus, EGIII may be a promising enzyme for the purpose of simultaneous saccharification and fermentation, since this process requires hydrolysis under mild conditions.

Compared to the kinetic parameters reported by Macarron et al. (1993) and Nakazawa et al. (2009) for CMCase activity, the K_{cat} value obtained here is higher in comparison to EGIII from *T. reesei*, with the K_m being twice as high. The K_m is similar to that found for endoglucanase activity in the crude extract of *T. harzianum* IOC-3844 (19 g/L) and the endoglucanase from *T. harzianum* ETS 323 (23 g/L) (de Castro et al., 2010a; Liu et al., 2010). This lower specificity in comparison to *T. reesei* does not preclude the use of this enzyme in saccharification, since a high cellulose concentration is used in this process.

In conclusion, the heterologous production of the β-1,4-endoglucanase III (cel12a) from *T. harzianum* was successfully performed in *P. pastoris*, yielding large amounts of enzyme. The use of yeasts for the production of recombinant enzymes for enzymatic cocktails can be of great value. The primary reason is that there is a low expression of native extracellular proteins by *P. pastoris*, but also, the industrial system for ethanol production is already well suited for the use of yeasts. These results indicate that this enzyme may be a feasible component of commercial enzyme cocktails for both separate and simultaneous hydrolysis and fermentation, mainly due to a high turnover rate. However, further investigations are required to measure the synergy between this enzyme and others from the cellulolytic complex and to evaluate its use on an industrial scale.

ACKNOWLEDGMENTS

Research supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

REFERENCES

Ahmed S, Bashir A, Saleem H, Saadia M, et al. (2009). Production and purification of cellulose-degrading enzymes from a filamentous fungus *Trichoderma harzianum*. *Pak. J. Bot.* 41: 1411-1419.

Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.

Arantes V and Saddler JN (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3: 4.

Banerjee G, Scott-Craig JS and Walton JD (2010). Improving enzymes for biomass conversion: a basic research perspective. *Bioenerg. Res.* 3: 82-92.

Bendtsen JD, Nielsen H, von Heijne G and Brunak S (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.

Boettner M, Prinz B, Holz C, Stahl U, et al. (2002). High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *J. Biotechnol.* 99: 51-62.

Cereghino JL and Cregg JM (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24: 45-66.

Corpet F (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*. 16: 10881-10890.

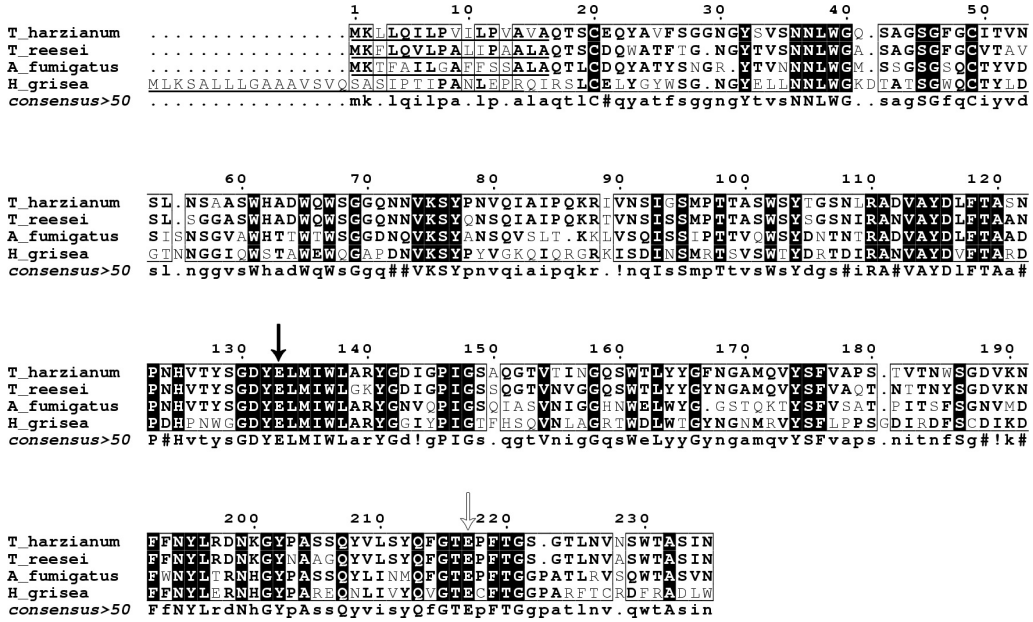
Cregg JM (2007). *Pichia* Protocols: Methods in Molecular Biology. Humana Press, Totowa.

de Castro AM, Ferreira MC, da Cruz JC, Pedro KC, et al. (2010a). High-yield endoglucanase production by *Trichoderma harzianum* IOC-3844 cultivated in pretreated sugarcane mill byproduct. *Enzyme Res.* 2010: 854526.

de Castro AM, Pedro KC, da Cruz JC, Ferreira MC, et al. (2010b). *Trichoderma harzianum* IOC-4038: A promising strain for the production of a cellulolytic complex with significant beta-glucosidase activity from sugarcane bagasse

cellulignin. *Appl. Biochem. Biotechnol.* 162: 2111-2122.
EasySelect™ Pichia Expression Kit (2001). A Manual of Methods for Expression of Recombinant Proteins Using pPICZ and pPICZα in *Pichia pastoris*. Version G, Invitrogen, Carlsbad.
Foreman PK, Brown D, Dankmeyer L, Dean R, et al. (2003). Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J. Biol. Chem.* 278: 31988-31997.
Galbe M and Zacchi G (2002). A review of the production of ethanol from softwood. *Appl. Microbiol. Biotechnol.* 59: 618-628.
Henriksson G, Nutt A, Henriksson H, Pettersson B, et al. (1999). Endoglucanase 28 (Cel12A), a new *Phanerochaete chrysosporium* cellulase. *Eur. J. Biochem.* 259: 88-95.
Karlsson J, Siika-aho M, Tenkanen M and Tjerneld F (2002). Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of *Trichoderma reesei*. *J. Biotechnol.* 99: 63-78.
Kwon I, Ekino K, Goto M and Furukawa K (1999). Heterologous expression and characterization of endoglucanase I (EGI) from *Trichoderma viride* HK-75. *Biosci. Biotechnol. Biochem.* 63: 1714-1720.
Liu SY, Shibu MA, Jhan HJ, Lo CT, et al. (2010). Purification and characterization of novel glucanases from *Trichoderma harzianum* ETS 323. *J. Agric. Food Chem.* 58: 10309-10314.
Lynd LR, Weimer PJ, van Zyl WH and Pretorius IS (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 66: 506-577.
Lynd LR, Laser MS, Bransby D, Dale BE, et al. (2008). How biotech can transform biofuels. *Nat. Biotechnol.* 26: 169-172.
Macarron R, Acebal C, Castillon MP, Dominguez JM, et al. (1993). Mode of action of endoglucanase III from *Trichoderma reesei*. *Biochem J.* 289: 867-873.
Mandels M, Parrish FW and Reese ET (1962). Sophorose as an inducer of cellulase in *Trichoderma viride*. *J. Bacteriol.* 83: 400-408.
Miller GL (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31: 426-428.
Nakazawa H, Okada K, Onodera T, Ogasawara W, et al. (2009). Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. *Appl. Microbiol. Biotechnol.* 83: 649-657.
Okada H, Tada K, Sekiya T, Yokoyama K, et al. (1998). Molecular characterization and heterologous expression of the gene encoding a low-molecular-mass endoglucanase from *Trichoderma reesei* QM9414. *Appl. Environ. Microbiol.* 64: 555-563.
Pandey A, Soccol CR, Nigam P and Soccol VT (2000). Biotechnological potential of agro-industrial residues. I: sugarcane bagasse. *Bioresour. Technol.* 74: 69-80.
Pessoa-Jr A, Roberto IC, Menossi M, dos Santos RR, et al. (2005). Perspectives on bioenergy and biotechnology in Brazil. *Appl. Biochem. Biotechnol.* 121-124: 59-70.
Sambrook J and Russell DW (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.
Sandgren M, Gualfetti PJ, Paech C, Paech S, et al. (2003). The *Humicola grisea* Cel12A enzyme structure at 1.2 A resolution and the impact of its free cysteine residues on thermal stability. *Protein Sci.* 12: 2782-2793.
Sandgren M, Stahlberg J and Mitchinson C (2005). Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Prog. Biophys. Mol. Biol.* 89: 246-291.
Schuster A and Schmoll M (2010). Biology and biotechnology of *Trichoderma*. *Appl. Microbiol. Biotechnol.* 87: 787-799.
Teather RM and Wood PJ (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Appl. Environ. Microbiol.* 43: 777-780.
Vasquez MP, da Silva JN, de Souza MBJ and Pereira N Jr (2007). Enzymatic hydrolysis optimization to ethanol production by simultaneous saccharification and fermentation. *Appl. Biochem. Biotechnol.* 137-140: 141-153.
Wang T, Liu X, Yu Q, Zhang X, et al. (2005). Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. *Biomol. Eng.* 22: 89-94.

SUPPLEMENTARY MATERIAL



Supplementary Figure 1. Sequence alignment of glycosyl hydrolase family 12 amino acid sequences from filamentous fungi. Underlined sequences indicate putative signal peptide. Black narrow indicates nucleophile residue and open narrow the acid-base residue. The protein sequence GenBank accession codes were: *Trichoderma harzianum* IOC-3844 (this study), *Trichoderma reesei* QM9414 (BAA20140), *Aspergillus fumigatus* Af293 (XP_748895), and *Humicola grisea* (AAM77714). The sequences were obtained from GenBank, and the alignment conducted with the MultAlin software (Corpet, 1988).