COMPARISON OF THE EXPRESSION IN *Saccharomyces cerevisiae* OF ENDOGLUCANASE II FROM *Trichoderma reesei* AND ENDOGLUCANASE I FROM *Aspergillus aculeatus*

Weina Zhang,[a,b] Cheng Liu,[a,c] Guoqiang Wang,[a,b] Yuanyuan Ma,[a] Kun Zhang,[a,c] Shaolan Zou,[a,*] and Minhua Zhang [a,c]

Two distinct expression cassettes were synthesized by overlapping PCR for expressing the endoglucanase I gene (*egl*1) from *Aspergillus aculeatus* and the endoglucanase II gene (*egl*2) from *Trichoderma reesei* in a *Saccharomyces cerevisiae* host. One contained the anchored sequence from the *S. cerevisiae cwp*2 gene, while the other did not. The low and high copy number plasmids YCplac33 and YEplac195 were used. The enzymatic activities and viscosity changes in the YP-CMC medium varied between the eight recombinant yeast strains produced, and the greatest values were obtained with the YE-TrEII' strain, which had an activity of 347.7 U/g dry cell weight (DCW) and viscosity at 12 h of 4.7% of the initial control value, respectively; YE-TrEII' was YEplac195-based and contained *T. reesei egl*2 and no Cwp2 sequence. Strains YC-AaEI and YC-TrEII showed the lowest enzyme activitiy (80.5 and 30.4 U/g DCW, respectively) and viscosity changes at 12 h (20.5 and 26.2% of the initial control viscosity, respectively), which were YCplac33-based and contained the Cwp2 sequence. The results showed that gene copy number was the most significant factor to influence the expression of endoglucanases in *S. cerevisiae*, and the existence of Cwp2 sequence led to decreased enzymatic level and viscosity-reducing performance, while it was shown not to realize efficient surface display of these two endoglucanases.

Keywords: Endoglucanase; Aspergillus aculeatus; Trichoderma reesei; Expression in yeast; Enzymatic activity; Viscosity

*Contact information: a: Tianjin R&D Center for Petrochemical Technology, Tianjin University, Tianjin 300072, China; b: School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China; c: Key Laboratory for Green Chemical Technology of Ministry of Education, Tianjin University, Tianjin 300072, China; *Corresponding author: slzhou@tju.edu.cn*

INTRODUCTION

The recent emergence of biofuels as a potential renewable alternative to fossil fuels necessitates a low-cost process for converting biomass into ethanol. Although *Saccharomyces cerevisiae* has a long commercial history in brewing, winemaking, baking, and ethanol production, it is not cellulolytic. The engineering of *S. cerevisiae* to express a heterologous cellulase system, thus enabling cellulose utilization, would help to reduce production costs associated with the production of ethanol from biomass (Lynd *et al.* 2002, 2005; van Zyl *et al.* 2007; La Grange *et al.* 2010).

The major requirement for *S. cerevisiae* as a cellulolytic yeast is to ensure sufficient expression and production of extracellular saccharolytic enzymes (van Zyl *et*

al. 2007). Much previous work has been performed in *S. cerevisiae* regarding the single or multiple expression of the three main classes of cellulases required to degrade crystalline cellulose: endoglucanases (EGs, EC3.2.1.4), cellobiohydrolases (CBHs, EC3.2.1.91), and beta-glucosidases (BGLs, EC3.2.1.21) (Lynd *et al.* 2002, 2005; van Zyl *et al.* 2007; La Grange *et al.* 2010; Qin *et al.* 2008; Yamada *et al.* 2010; Kuroda and Ueda 2011). Earlier studies have shown that many factors affect the expression levels of cellulases in *S. cerevisiae*, including the type of enzymatic activity, the origin of the gene, gene copy number, and the promoter and secretion signal sequences used. In addition, yeast surface display has developed as an important method for expressing recombinant proteins (Kuroda and Ueda 2011; van der Vaart *et al.* 1997), which necessitates the use of anchored peptide sequences; this is a factor that also affects cellulase expression in *S. cerevisiae*. Moreover, it is important that enzymes are expressed at concentrations that provide functionality but that do not exert too much unnecessary stress on the cellulolytic yeast (van Zyl *et al.* 2007). To obtain this, it may be necessary to synergize all of the aforementioned factors.

Being one of the most important cellulase components, EGs act on the amorphous regions of cellulose and yield cellobiose and cello-oligosaccharides as hydrolysis products (La Grange *et al.* 2010). EGs are further shown to be the key enzymes that rapidly reduce the viscosity during enzymatic hydrolysis of lignocellulosic biomass (Szijarto *et al.* 2011). The important sources of EGs are *Trichoderma reesei* and *Aspergillus aculeatus*, which are the widely studied and industrially important cellulolytic fungi (Lynd *et al.* 2002; van Zyl *et al.* 2007). *T. reesei* produces at least four endoglucanases, and endoglucanase II is one of the most abundant endoglucanases from *T. reesei* and also has the highest catalytic efficiency (Qin *et al.* 2008). *A. aculeatus* produces at least three endoglucanases, and endoglucanase I (formerly called FI-CMCase) is the most abundant enzyme of the cellulase components produced by *A. aculeatus*, playing an important role in cellulose hydrolysis (Minamiguchi *et al.* 1995; Murao *et al.* 1988). These two endoglucanases have been expressed separately in *S. cerevisiae*, but their expression levels and performance in this system have not been compared directly (Lynd *et al.* 2002; van Zyl *et al.* 2007; Qin *et al.* 2008; Yamada *et al.* 2010).

Thus, in the present study, the effects of gene source, anchored peptide, and copy number on the expression of endoglucanases in *S. cerevisiae* were investigated. Two endoglucanase genes were selected, namely *egl*2 that encodes endoglucanase II in *T. reesei* (Tr *egl*2) (Saloheimo *et al.* 1988) and *egl*1 that encodes endoglucanase I in *A. aculeatus* (Aa *egl*1) (Ooi *et al.* 1990b), and a comparative study of their expression in *S. cerevisiae* was performed with *S. cerevisiae* Cwp2 anchor sequence or not and using a low- or high copy number plasmid. To date, there have been very few studies of this type.

EXPERIMENTAL

Strains, Plasmids, Media, and Growth Condition

The microbial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* was used for recombinant DNA manipulation. *S. cerevisiae* W303-1A

was used as the parent strain for expressing the endoglucanases, and the filamentous fungi *A. aculeatus* and *T. reesei* were the sources of the endoglucanase genes.

Table 1. Microbial Strains and Plasmids Used in this Study

Strains or plasmids	Essential properties	Source or reference
S. cerevisiae W303-1A	*MAT*a *ade*2 *trp*1 *his*3 *can*1 *ura*3 *leu*2	In our lab
A. aculeatus	Wild-type	CICC 2193
T. reesei	Hypersecreting mutant of wild-type RUT-C30	ATCC 56765
E. coli Top10	F⁻ *mcr*A Δ(*mrr-hsd* RMS-*mcr*BC) ϕ80 *lacZ*□ΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen
pGEM-T easy	Amp^r	Promega
YCplac33	Amp^r *URA*3, low copy number in *S. cerevisiae*	ATCC 87586 Gietz *et al*.1988
YEplac195	Amp^r *URA*3, high copy number in *S. cerevisiae*	ATCC 87589 Gietz *et al*.1988
pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	pGEM-T easy vector, open reading frame of beta-glucosidase I gene from *A. aculeatus*	In our lab

E. coli was grown at 37°C in Luria-Bertani medium containing 10 g/L tryptone, 5 g/L yeast extract, and 10 g/L sodium chloride. When needed, 100 mg/L ampicillin was added to this medium. *S. cerevisiae* W303-1A was cultured aerobically at 30°C in YPD medium (10 g/L yeast extract, 20 g/L tryptone, and 20 g/L glucose as the sole carbon source). *S. cerevisiae* W303-1A transformants were selected and maintained on CMG agar (6.7 g/L yeast nitrogen base without amino acids, 18 g/L agar, 20 g/L glucose, and the appropriate amino acid and nucleic acid supplements). *A. aculeatus* and *T. reesei* were cultured at 30°C in malt juice medium, of which the sugar degree was 5 Baume.

DNA Manipulation, Plasmid Construction, and Yeast Transformation

Standard molecular genetic techniques were used for the nucleic acid manipulations (Sambrook and Russell 2001). The primers used and plasmids constructed are listed in Table 2 and Appendix Table 1, respectively.

Primers P1 (5'-AAGACTGACATCCTCCGCTTG-3') and P2 (5'-CTACTGTAC GCTGGCA-3') were used to isolate the Aa *egl*1 gene and were designed based on sequences EMBL: D00546 (from genomic DNA) and EMBL: X52525 (from mRNA), which are the only entries reported for the Aa *egl*1 gene (Ooi *et al*. 1990a,b). The 1.2 kb polymerase chain reaction (PCR) product from the genomic DNA of *A. aculeatus* was introduced into the pGEM-T easy vector (Promega) to obtain the pAa1196 plasmid (Table 2 and Appendix Fig. 1a). The cloning fragment in pAa1196 was sequenced and analyzed. Then, two expression cassettes of Aa *egl*1, P*tpi-xyn*2s-Aa *EGI*-T*adh*I, and P*tpi-xyn*2s-Aa *EGI-cwp*2-T*adh*I, were prepared by overlapping PCR using the primers P3 to P14, and pAa1196 and pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I as the templates (Table 1 and Appendix Table 1). The resultant 1.86 kb and 2.06 kb PCR fragments were digested with *Pst*I and *Sca*I and ligated into the *Pst*I- and *Sma*I-digested YCplac33 and YEplac195 vectors (Table 1), respectively; the resulting four plasmids are shown in Table 2 and Appendix Figs. 1b and 1c.

Table 2. Recombinant Plasmids and *S. cerevisiae* Strains Constructed in this Study

Recombinant plasmids	Plasmid and corresponding *S. cerevisiae* W303-1A transformants designation[a]
pGEM-T easy-Aa *EGI* 1196 bp (containing introns)	pAa1196
YCplac33-P*tpi-xyn*2s-Aa *EGI-cwp*2-T*adh*I	YC-AaEI
YCplac33-P*tpi-xyn*2s-Aa *EGI*-T*adh*I	YC-AaEI'
YEplac195-P*tpi-xyn*2s-Aa *EGI-cwp*2-T*adh*I	YE-AaEI
YEplac195-P*tpi-xyn*2s-Aa *EGI*-T*adh*I	YE-AaEI'
YCplac33-P*tpi-xyn*2s-Tr *EGII-cwp*2-T*adh*I	YC-TrEII
YCplac33-P*tpi-xyn*2s-Tr *EGII*-T*adh*I	YC-TrEII'
YEplac195-P*tpi-xyn*2s-Tr *EGII-cwp*2-T*adh*I	YE-TrEII
YEplac195-P*tpi-xyn*2s-Tr *EGII*-T*adh*I	YE-TrEII'

[a] The control *S. cerevisiae* strains W303-1A (YCplac33) and W303-1A (YEplac195) were named YC and YE, respectively.

The regulatory elements used in the aforementioned expression cassettes included the anchored peptide-encoding sequence (named *cwp*2) that was 207 bp at the 3' end of *S. cerevisiae* cell wall protein 2 (van der Vaart *et al.* 1997), the TPI1 promoter (P*tpi*) and the ADH1 terminator (T*adh*I) from *S. cerevisiae*, and the secretion signal of the *xyn*2 gene from *T. reesei* (*xyn*2s).

Using methods similar to those mentioned above, the two expression cassettes of Tr *egl*2 were prepared using primers P15 to P24 and with genomic DNA from *T. reesei* and pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I as the templates (Table 1 and Appendix Table 1). Then, four recombinant plasmids were constructed (Table 2). The sequence information of Tr *egl*2 gene used here for primers design was from EMBL: M19373 (Saloheimo *et al*. 1988).

Transformation of the plasmids into *S. cerevisiae* was performed according to the lithium acetate method (Gietz *et al.* 1995), and *URA*3 was used as the selective marker. For their cellulose-degrading ability, the yeast transformants containing endoglucanase activity were screened using the Congo Red halo test. After incubating the yeasts on CMG-1% CMC (sodium carboxymethylcellulose, degree of substitution 0.7) agar plates at 30°C for 2 days, colonies were washed off the plate with water and any remaining CMC was stained with 0.1% Congo Red before destaining with 1% (w/v) sodium chloride (Den Haan *et al.* 2007).

Enzyme Assay and SDS-PAGE

Endoglucanase activity was measured according to the dinitrosalicylic acid (DNS) method. The composition of the DNS reagent solution was 10 g/L 3,5-dinitrosalicylic acid, 10 g/L NaOH, 200 g/L potassium sodium tartrate, 2 g/L phenol, and 5 g/L Na_2SO_3. Yeast strains were cultured aerobically in selective CMG medium at 30°C for 24 h. The supernatant and cells of the resultant cultures were separated by centrifugation before the cells were washed twice with distilled water and finally resuspended in distilled water. The supernatant and resuspended cells were tested for activity. The enzyme assays were started by adding 100 µL of test solution or suspension to 400 µL of citric acid buffer (50 mM, pH 5.0) containing 10 g/L CMC. After incubating for 30 min at 50°C, 600 µL of DNS reagent was added, and then the solution was vortexed quickly

and boiled for 10 min. The sample was cooled to room temperature and centrifuged before the absorbance value of each supernatant was read at 540 nm. One unit of enzyme activity was defined as the amount of enzyme required to produce 1 µmol of reducing sugar per min. Glucose was used as the standard. Results were expressed as activity units per g dry cell weight (DCW) (U/g DCW). When used for studying the effects of pH and temperature, the pH or temperature values varied, but the other operating conditions remained as above. The aforementioned supernatants and cell pellets were also used for SDS-PAGE analysis in 10% polyacrylamide (samples were first mixed with 5× loading buffer at a volume ratio of 4:1, boiled at 100°C for 10 min and then centrifuged to remove the debris), which was followed by staining with Coomassie Brilliant Blue G-250.

Viscosity Measurements in YP-CMC Media

After being precultivated in CMG medium for 24 h, the recombinant *S. cerevisiae* W303-1A strains were grown aerobically in fresh medium for 24 h at 30°C. The resultant cells were collected by centrifugation, and these were washed twice with distilled water and then inoculated into 1 L shake flasks containing 200 mL of YP-1% CMC medium (10 g/L yeast extract, 20 g/L tryptone, and 10 g/L CMC). The initial optical density at 660 nm (OD_{660}) of the medium was adjusted to 1. These cultures were allowed to grow anaerobically at 30°C and 190 rpm for viscosity measurements. Viscosity measurements were performed on a Stress Tech Senior rheometer (Rheometric Scientific, USA) according to the method described by Den Haan *et al.* (2007).

RESULTS AND DISCUSSION

Cloning of Aa *egl*1 Gene, Construction of Recombinant Plasmids, and Recombinant *S. cerevisiae* W303-1A Strains Expressing the Endoglucanases

A preliminary experiment demonstrated that there was a relatively significant difference of the Aa *egl*1 gene sequences between *A. aculeatus* CICC2193 used in this study and *A. aculeatus* No. F-50 strain reported by sequences EMBL: D00546 and EMBL: X52525 (Ooi *et al.* 1990a,b). Thus it was necessary to first isolate and identify the Aa *egl*1 gene of *A. aculeatus* CICC2193. The sequence of the PCR fragment in recombinant plasmid pAa1196 has been submitted to GenBank (accession no. JQ581513). Sequence alignment analysis indicated that the nucleotide sequence of the 1196 bp fragment corresponded to nucleotides 25-1216 of EMBL: D00546 (85% homology), while the 666 bp nucleotide sequence of the three exons corresponded to nucleotides 98-763 of EMBL: X52525 (93% homology). The nucleotide sequence of the second intron in Aa *egl*1 was 84 bp, while this was 87 bp in the *A. aculeatus* No. F-50 strain (EMBL: X52525) (Ooi *et al.* 1990a).

The sequence data for the expression cassette fragments in the recombinant plasmids confirmed that none of the expression cassettes contained mutations.

Eight recombinant expression vectors (Table 2) and two control plasmids, YCplac33 and YEplac195 (Table 1), were transformed into *S. cerevisiae* W303-1A. The

resulting strains are summarized in Table 2. For each of two genes (Tr *egl*2 or Aa *egl*1), there were four recombinant strains, two of which were YCplac33-based (YC-) and the other YEplac195-based (YE-), or, contained Cwp2 sequence (AaEI or TrEII) and the other contained no (AaEI' or TrEII'). In the Congo Red halo test, transparent halos were observed around all the transformants but not the control strains (data not shown), indicating that the endoglucanase genes were expressed successfully.

Activity Measurement and SDS-PAGE of Recombinant *S. cerevisiae* Cultures

While *S. cerevisiae* is engineered to be cellulolytic and thus enabled to directly produce ethanol from biomass, it would be more suitable to be evaluated and used as a whole cell catalyst. Thus, in this work, it was the whole cell performance, not the purified recombinant enzyme property, that was evaluated in terms of growth, enzymatic activities, and viscosity changes.

Figure 1a shows that the growth curves of all the recombinant strains in CMG media were similar. Thus, it seems that expression of the fused proteins did not impose an evident metabolic burden on the cells in CMG medium. Figure 1b shows that the supernatant enzyme activity varied between the strains and was of the following order (from greatest to lowest): YE-TrEII' (347.7 ± 4.2 U/g DCW) > YE-AaEI' (230.49 ± 5.4 U/g DCW) > YE-AaEI > YE-TrEII > YC-AaEI' > YC-TrEII' > YC-AaEI (80.55 ± 0.8 U/g DCW) > YC-TrEII (30.38 ± 7.8 U/g DCW) > YC. Surprisingly, it seemed that the cwp2-anchored peptides did not function as cell-surface displaying endoglucanases, as might have been expected, because the eight recombinant strains each showed very low cell activities (3.2~8.0 U/g DCW) but relatively high supernatant activities (Fig. 1b). The ratio of cell activities to total activities was 2 to 5 % for four strains expressing Cwp2 fusion proteins. In addition, the cell debris prepared from cell pellets with zymolyase also showed low activities similar to cell pellets.

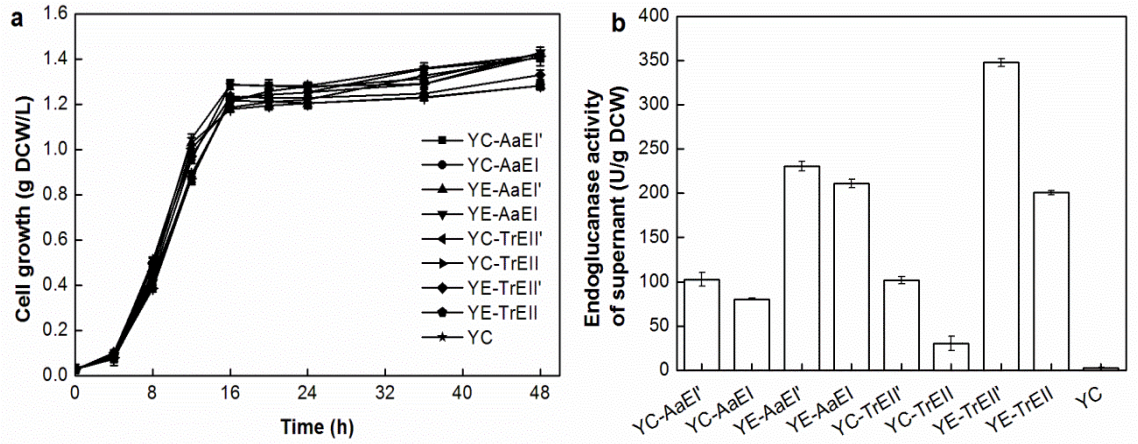

Fig. 1. The growth curve and endoglucanase activity of the recombinant yeast strains. (**a**) Time course of recombinant cell growth in CMG aerobically; (**b**) Extracellular endoglucanase activity of recombinant strains growing in CMG aerobically for 24 h. Each experiment was done in triplicate.

SDS-PAGE of the supernatants preliminarily showed the existence of several recombinant protein bands, which were indicated by the arrows (Fig. 2). At greater

sample concentration, a ~27 kD band was visualized for YC-AaEI' (lane 1), and a ~76 kD band was observed for YC-TrEII (lane 6), while no similar bands could be seen for the control YC or YE samples. Thus, SDS-PAGE analysis indicated that the eight strains each expressed and secreted endoglucanases into the medium and that the approximate protein quantities of secreted endoglucanases for most of strains was in general agreement with the corresponding level of enzyme activity (Figs. 1b and 2).

An earlier investigation (Minamiguchi *et al.* 1995) reported that the molecular weight of native and secreted *A. aculeatus* endoglucanase I was 24 kD, but when expressed in *S. cerevisiae*, it was 24 kD (non-glycosylated) or 28 kD (glycosylated). Meanwhile, the molecular weight of native and secreted *T. reesei* endoglucanase II is reported as 48 kD, but it was 57 kD when expressed in *S. cerevisiae* (Qin *et al.* 2008). Thus, the results in this present study are in good agreement with these earlier reports (Minamiguchi *et al.* 1995; Qin *et al.* 2008).

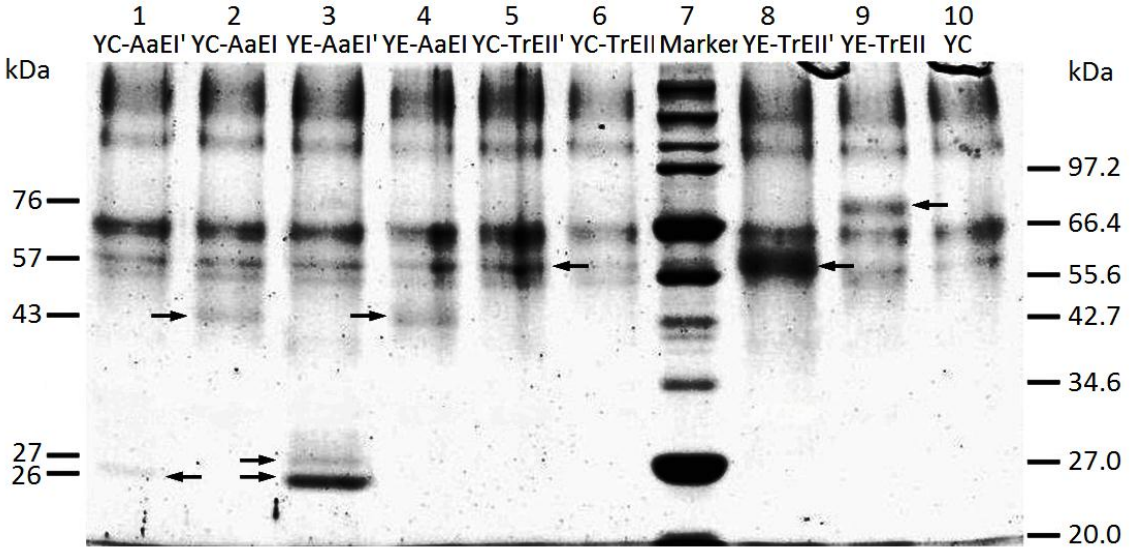

Fig. 2. SDS-PAGE analysis of culture supernatants. Supernatants were from cultures growing in CMG aerobically for 24 h. Arrows point to the protein bands of which the size is indicated on the *left* distinctly different from ones of Lane 10 (control YC) and the approximate size of those protein bands is indicated on the *left*. The size of the protein molecular weight marker bands is indicated on the *right*. The cell number corresponding to each sample loading of nine samples were kept constant as far as possible.

It was noteworthy that there existed ~43 and ~76 kD bands for four strains, YC-AaEI, YE-AaEI and YC-TrEII, YE-TrEII. They should be the Cwp2-fusion enzymes, but they were to varying extents glycosylated and abnormally secreted into supernatants. The approximate proportional relation between the protein amounts of target bands and the supernatant activities implied that the Cwp2 anchor peptide did not significantly affect the specific activity of the secreted Cwp2-fusion enzyme (Figs. 1b and 2).

On the other hand, SDS-PAGE analysis of cell pellets seemed not to demonstrate expected differences in the electrophoretic patterns between the eight recombinant strains and the samples from the control YC and YE strains. So did for the spheroplasts and cell debris prepared from the cell pellets with zymolyase. Nevertheless, this seems to be in

agreement with the enzyme activity data from the cell pellets and cell debris. Of course, it should not be overlooked that the proteins from the cell lysates were complicated, making differences difficult to identify.

The effects of pH and temperature on the culture supernatant activities of YE-AaEI' and YE-TrEII' were first investigated. Figure 3 shows that the culture supernatant from YE-AaEI' had maximal activity at pH 5.0 or 60°C, and this was the same for endoglucanase I from the other recombinant strains expressing Aa *egl*1 (data not shown). These observations contrasted with previous reports (Minamiguchi *et al.* 1995; Murao *et al.* 1988), where it was found that the optimum pH and temperature values for endoglucanase I from *A. aculeatus* were pH 5 and 50°C when expressed in *S. cerevisiae* (Minamiguchi *et al.* 1995) and pH 4.5 and 50°C (Murao *et al.* 1988) or pH 5 and 50°C (Minamiguchi *et al.* 1995) when expressed in the native host. Figure 3 shows that culture supernatant from YE-TrEII' had maximal activity at pH 5.0 or 70°C, and these values were identical in the other recombinant strains expressing Tr *egl*2 (data not shown). The pH optimum is generally consistent with data reported by Qin *et al.* (2008), but the temperature effect curve of *T. reesei* endoglucanase II expressed in *S. cerevisiae* has not been reported previously (Lynd *et al.* 2002; van Zyl *et al.* 2007; La Grange *et al.* 2010).

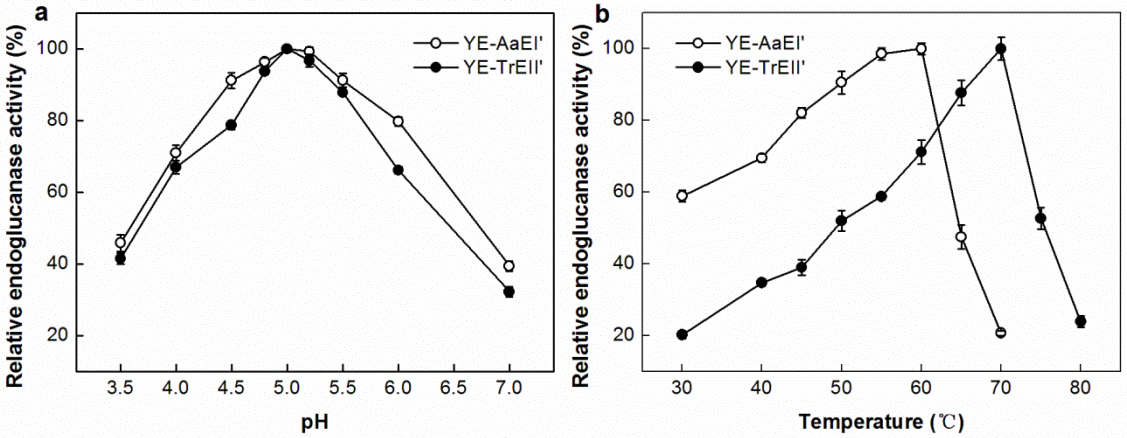

Fig. 3. Effect of pH and temperature on recombinant endoglucanase activity. (**a**) Effect of pH on recombinant endoglucanase activity; (**b**) Effect of temperature on recombinant endoglucanase activity. Cultures were grown in CMG medium at 30°C for 24 h, and supernatants were obtained by centrifugation as the crude free enzymes. For AaEGI, the absolute activity corresponding to 100% was 230.5 U/g DCW at pH 5.0 and 254.3 U/g DCW at 60°C; for TrEGII, the absolute activity corresponding to 100% was 347.7 U/g DCW at pH 5.0 and 668.7 U/g DCW at 70°C. Three independent experiments were done for each data point.

Viscosity Changes in YP-CMC Media

It is well known that the cost of processing lignocellulosic biomass can be reduced by increasing the concentration of insoluble solids in various process operations by reducing the energy requirements associated with water removal; however, higher solids concentration increases the apparent viscosity of biomass slurries, which makes mixing and conveying operations more challenging (Szijarto *et al.* 2011). The cited authors had reported that EGs are the key enzymes to rapidly reduce the viscosity of lignocellulose substrate. Here, the typical hydrocolloid CMC was selected as substitution for biomass and used for comparing the enzymatic liquefaction (reduction of viscosity)

performance of eight strains. Figure 4 shows that the viscosity of the YP-CMC medium changed significantly at 12 h, but then the extents of viscosity change during 24 h varied between the different strains. Nevertheless, the orders of the strains from greatest to the lowest values were identical to the order determined for the supernatant activities (Fig. 1b). The viscosity of the control YC culture showed slight variation, and the viscosities of the YE-TrEII', YE-AaEI', YC-AaEI, and YC-TrEII cultures at 12 h were 4.7, 6.0, 20.5, and 26.2% of the initial values, respectively (Fig. 4). Thus, the changes in viscosity are much faster than reported previously (Den Haan *et al.* 2007), but the use of a different substrate (CMC versus phosphoric acid swollen cellulose) may explain this observation (Den Haan *et al.* 2007). The OD_{660} values of the eight strains at 48 h varied slightly and the maximum observed was 1.44 ± 0.05 AU for the YE-TrEII' strain.

Therefore, the enzymatic activity and viscosity change data above indicated that among the three factors investigated, gene copy number is the most significant to influence the expression of endoglucanases in *S. cerevisiae*, since all of the YEplac195-based recombinant strains surpassed the values observed for the YCplac33-based strains. The second most significant factor is Cwp2 anchor sequence, because yeasts carrying no anchor sequence always show better performance than ones carrying anchor sequence under the same conditions. Lastly, the differences between *T. reesei* endoglucanase II and *A. aculeatus* endoglucanase I expression were only evident if the YEplac195 vector and no anchor sequence (YE-TrEII' and YE-AaEI') or the YCplac33 vector and anchor sequence (YC-AaEI and YC-TrEII) were used simultaneously.

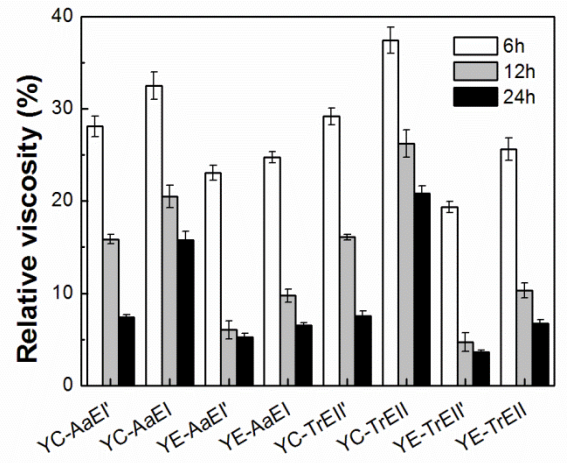

Fig. 4. Viscosity changes in YP-CMC (10 g/L CMC) media of the anaerobic cultures. Viscosity measurements were done over a shear rate range of 2-200 s^{-1} and at the temperature of 30°C. The average viscosities of the culture media containing yeast cells were referred to a percentage of the average viscosity of YP-CMC fresh medium, which was 102 mPa s. The viscosity of the control YC culture was (95 ± 1.7)%, (92 ± 2.1)%, and (90 ± 1.4)% at 6 h, 12 h, and 24 h, respectively. Each experiment was done in triplicate.

In this study, the cell activities of recombinant strains and the function or effect of the *cwp*2 anchoring peptide warrant further investigation and discussion. The key problem is that the enzymatic activity, SDS-PAGE, and viscosity change data above indicated that Cwp2-fusion proteins were expressed but were not effectively displayed on the cell surface by Cwp2. It is well known that "the characteristics of carrier protein,

passenger protein and host cell, and fusion method all affect the efficiency of surface display of proteins" (Lee *et al*. 2003). On the other hand, in enzymatic reactions using surface display system, the accessibility of fusion protein to the substrate, especially polymeric substrate such as starch and cellulose, is very important and influenced by the length of the fusion protein as well as the anchoring domain (Sato *et al*. 2002). In this case, the Cwp2 used was 69 amino acid residues. It had been shown to be one of the best candidates for surface display of foreign genes in *S. cerevisiae* and successfully used for display expression of some foreign genes (van der Vaart *et al*. 1997; van Rooyen *et al*. 2005; Liu *et al*.2010; Kuroda and Ueda 2011). Further, Cwp2 had also been used for cell-surface expression of three beta-glucosidases from different sources in our lab and the immobilization ratio was all above 90%. Comparative analysis of the similarities and differences between two cases (endoglucanases expression and glucosidases expression) reminded us that it may be the enzyme itself (passenger protein) and substrate used (CMC for endoglucanase and cellobiose for glucosidase), rather than the Cwp2p, host cell and vectors, that is the first and main cause for the low displaying efficiency, because the latter used in two cases was exactly the same. Considering the length of ORFs of Tr *egl*2 and Aa *egl*1 gene were 1194 bp and 666 bp, respectively, far lower than that of glucosidase-encoding genes (about 2.5 kb), we attempted to first optimize the fusion method. One approach was to insert the $(G_4S)_3$ sequence, a widely used flexible linker (Liu *et al*. 2010), between the target gene and Cwp2 gene, and the other was to replace the Cwp2 sequence with the AGα1 anchored sequence, which was 320 amino acid residues and had widely been used to display various peptides and proteins in *S. cerevisiae* (van der Vaart *et al*. 1997; van Rooyen *et al*. 2005; Kuroda and Ueda 2011). Activity measurement of cultures of eight new strains produced showed that the anchored ratio increased 5 to 10% for four strains expressing linker-Cwp2 fusion proteins and 20 to 35% for four strains expressing AGα1 fusion proteins, and was 2 to 4 times and 10 to 15 times as high as the level of previous strains, respectively. But, at the same time, the total activity for four strains expressing AGα1 fusion proteins all decreased and was only 20 to 50% of previous values (for linker-Cwp2-containing strains, the total activity kept constant or slightly elevated)

Therefore, to obtain efficient surface display of Tr *egl*2 and Aa *egl*1 gene in *S. cerevisiae*, there is still a need to further investigate and optimize the expressing system. However, the data above might have implied that fusion cellulase, free or anchored, would not help to improve the whole cell performance of engineered cellulolytic yeast strain in biomass utilization, possibly due to decreased activity and/ or limited accessi-bility of fusion protein to lignocellulosic substrate compared to free cellulase.

Although it seems that secreted heterologous endoglucanases from each strain were to varying extents glycosylated or hyperglycosylated (Fig. 2), this did not negatively influence enzyme activity (Figs. 1b and 4). Such observations are in agreement with previous reports (Lynd *et al*. 2002; van Zyl *et al*. 2007).

In preliminary experiments, the transcriptional level of promoters for *TPI*1, *PDC*1, *ADH*1, and *PGK*1 were compared in media containing glucose or xylose, and it was observed that the *TPI*1 promoter gave most effective results. Thus, the *TPI*1 promoter was selected as the promoter for expressing the endoglucanases in this present study. The *TPI*1 promoter is used only rarely, and it is much more common to use the *PGK*1, *PDC*1,

*ADH*1, and *ENO*1 promoters (Lynd *et al.* 2002; van Zyl *et al.* 2007; La Grange *et al.* 2010).

To the best of our knowledge, this is the first report that has investigated the comparison of the expression of Tr *egl*2 and Aa *egl*1 gene and the effect of Cwp2 anchor sequence and copy number on their expression in *S. cerevisiae*. The results presented herein will help in the construction of further cellulolytic yeasts that express these extracellular saccharolytic enzymes at high levels, which will ultimately play an important role in the production of ethanol from biomass.

CONCLUSIONS

1. The effect of gene source (Tr *egl*2 or Aa *egl*1), Cwp2 anchor sequence, and low- or high-copy number vector (YCplac33 or YEplac195) on the expression of endoglucanases in *S. cerevisiae* was investigated. Thus, eight endoglucanase-expressing *S. cerevisiae* strains were successfully constructed, and a comparison of their enzymatic levels and viscosity-reducing performance was made.
2. The results showed that enzymatic activities and viscosity changes in the YP-CMC medium of the YEplac195-based recombinant strains surpassed the values observed for the YCplac33-based strains, and yeasts carrying no Cwp2 anchor sequence always showed better performance than ones carrying Cwp2 anchor sequence under the same conditions. In addition, only when the YEplac195 vector and no Cwp2 anchor sequence were used simultaneously was Tr *egl*2-containing strain shown to be better than the Aa *egl*1-containing strain. Thus, YE-TrEII' was shown to be the best endoglucanase-expressing *S. cerevisiae* strain, which had activity of 347.7 U/g DCW and viscosity at 12 h of 4.7 % of the initial control value.
3. The endoglucanase I gene of *A. aculeatus* CICC2193 was isolated, and the 666 bp nucleotide sequence of the three exons was identified by sequence alignment analysis. Its function was further confirmed by its efficient expression in *S. cerevisiae*.

ACKNOWLEDGMENTS

We would like to thank professor, Dr. P. Ma for his help. This study was supported by the National Natural Science Foundation of China (No. 30900033 and No. 30600369).

REFERENCES CITED

Den Haan, R., Rose, S. H., Lynd, L. R., and van Zyl, W. H. (2007). "Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*," *Metab. Eng.* 9, 87-94.

Gietz, R. D., and Sugino, A. (1988). "New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites," *Gene* 74, 527-534.

Gietz, R. D., Schiestl, R. H., Willems, A. R., and Woods, R. A. (1995). "Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure," *Yeast* 11(4), 355-360.

Kuroda, K., and Ueda, M. (2011). "Cell surface engineering of yeast for applications in white biotechnology," *Biotechnol. Lett.* 33(1), 1-9.

La Grange, D. C., Den Haan, R., and van Zyl, W. H. (2010). "Engineering cellulolytic ability into bioprocessing organisms," *Appl. Microbiol. Biotechnol.* 87(4), 1195-1208.

Lee, S.Y., Choi, J. H., and Xu, Z. H. (2003). "Microbial cell-surface display," *Trends Biotechnol.* 21(1), 45-52.

Liu, W. S., Zhao, H. Y., Jia, B., Xu, L., and Yan, Y. J. (2010). "Surface display of active lipase in *Saccharomyces cerevisiae* using Cwp2 as an anchor protein," *Biotechnology Letters*. 32(2), 255-260.

Lynd, L. R., van Zyl, W. H., McBride, J. E., and Laser, M. (2005). "Consolidated bioprocessing of cellulosic biomass: an update," *Curr. Opin. Biotechnol.* 16(5), 577-583.

Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002). "Microbial cellulose utilization: Fundamentals and biotechnology," *Microbiol. Mol. Biol. Rev.* 66(3), 506-577.

Minamiguchi, K., Ooi, T., Kawaguchi, T., Okada, H., Murao, S., and Aria, M. (1995). "Secretive expressing of the *Aspergillus aculeatus* cellulose (FI-CMCase) by *Saccharomyces cerevisiae*," *J. Ferment. Bioeng.* 79(4), 363-366.

Murao, S., Sakamoto, R., and Aria, M. (1988). "Cellulases of *Aspergillus aculeatus*," *Method. Enzymol.* 160, 274-299.

Ooi, T., Shinmyo, A., Okada, H., Hara, S., Ikenaka, T., Murao, S., and Arai, M. (1990a). "Cloning and sequence analysis of a cDNA for cellulase (FI-CMCase) from *Aspergillus aculeatus*," *Curr. Genet.* 18(3), 217-222.

Ooi, T., Shinmyo, A., Okada, H., Murao, S., Kawaguchi, T., and Arai, M. (1990b). "Complete nucleotide sequence of a gene coding for *Aspergillus aculeatus* cellulase (FI-CMCase)," *Nucleic Acids Res.* 18(19), 5884.

Qin, Y. Q., Wei, X. M., Liu, X. M., Wang, T. H., and Qu, Y. B. (2008). "Purification and characterization of recombinant endoglucanase of *Trichoderma reesei* expressed in *Saccharomyces cerevisiae* with higher glycosylation and stability," *Protein Expres. Purif.* 58, 162-167.

Saloheimo, M., Lehtovaara, P., Penttilae, M., Teeri, T. T., Stahlberg, J., Johansson, Pettersson, G., Claeyssens, M., Tomme, P., and Knowles, J. K. C. (1988). "EGIII, a new endoglucanase from *Trichoderma reesei*: The characterization of both gene and enzyme," *Gene* 63(1), 11-21.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sato, N., Matsumoto, T., Ueda, M., Tanaka, A., Fukuda, H., and Kondo, A. (2002). "Long anchor using Flo1 protein enhances reactivity of cell surface-displayed glucoamylase to polymer substrates," *Appl. Microbiol. Biotechnol.* 60, 469-474.

Szijarto, N., Siika-Aho, M., Sontag-Strohm, T., and Viikari, L. (2011). "Liquefaction of hydrothermally pretreated wheat straw at high-solids content by purified *Trichoderma* enzymes," *Bioresource Technol.* 102(2), 1968-1974.

van Der Vaart, J. M., te Biesebeke, R., Chapman, J. W., Toschka, H. Y., Klis, F. M., and Verrips, C. T. (1997). "Comparison of cell wall proteins of *Saccharomyces cerevisiae* as anchors for cell surface expression of heterologous proteins," *Appl. Environ. Microbiol.* 63(2), 615-620.

van Rooyen, R., Hahn-Hagerdal, B., La Grange, D. C., and van Zyl, W. H. (2005). "Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains," *J. Biotechnol.* 120(3), 284-295.

van Zyl, W. H., Lynd, L. R., Den Haan, R., and McBride, J. E. (2007). "Consolidated bioprocessing for bioethanol production using *Saccharomyces cerevisiae*," *Adv. Biochem. Engin. Biotechnol.* 108, 205-235.

Yamada, R., Taniguchi, N., Tanaka, T., Ogino, C., Fukuda, H., and Kondo, A. (2010). "Cocktail δ-integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains," *Microb. Cell Fact.* 9(32), 2-8.

Article submitted: April 14, 2012; Peer review completed: May 19, 2012; Revised version received and accepted: July 5, 2012; Published: July 13, 2012.

APPENDIX

Table 1. Expression Elements, Origin, PCR Primers Used for their Amplification

Expression elements (size)	Template	Primer name	Primer sequence (The restriction sites included in the primers are indicated in underline bold face)
To construct gene expressing cassette: **P*tpi-xyn*2s-Aa *EGI*-T*adh*I**, 1858 bp			
P*tpi*[a]-*xyn*2s[b], 833 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I[c]	P3	TGACCACATAT***GGCCGCCTAGGCC***TCGAGCT GCAGTAACGGGAGCGTAATGG (*Nde*I/*Avr*II/*Sfi*I/*Xho*I/*Pst*I)
		P4	GAGTCTGTTGGCGCTTCTCCACAGCC
exon1, 385 bp	pAa1196[d]	P5	GGAGAAGCGCCAACAGACTCAACTCT
		P6	ACGAGCAAGCCAGATCATGAGCTCAT
exon2, 241 bp	pAa1196	P7	TCATGATCTGGCTTGCTCGTTACGGC
		P8	AATTGGAGCGTGATGAGGTACTGGCT
exon3, 100 bp	pAa1196	P9	TACCTCATCACGCTCCAATTCGGCAC
		P10	AGAAATTCGCCTACTGTACGCTGGCA
T*adh*I[e], 379 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	P11	CGTACAGTAGGCGAATTTCTTATGAT
		P12	GTTGTACCAT***GGCCTAGGCGGCC***GACGATAG TCAGTACTATCCGTGTGGAAGAAC (*Sca*I/*Sfi*I/*Avr*II/*Nco*I)
To construct gene expressing cassette: **P*tpi-xyn*2s-Aa *EGI-cwp*2-T*adh*I**, 2062 bp			
P*tpi-xyn*2s-Aa *EGI,* 1499 b*p*	P*tpi-xyn*2s-Aa *EGI*-T*adh*I	P13[f]	TTTGAGAAATCTGTACGCTGGCAGAC
*cwp*2[g]-T*adh*I, 586 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	P14[h]	CAGCGTACAGATTTCTCAAATCACTG
To construct gene expressing cassette: **P*tpi-xyn*2s-Tr *EGII*-T*adh*I**, 2362 bp			
P*tpi-xyn*2s, 823 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	P15	AGTTTGAAGCTTGGCCGCCTAGGCCTAACGG GAGCGTAA (*Hin*dIII/*Avr*II/*Sfi*I)
		P16	CAGTCTGCTGGCGCTTCTCCACAGCC
exon1, 286 bp	*T. reesei*	P17	GGAGAAGCGCCAGCAGACTGTCTGGG
		P18	GCAAGTGCCATCTGTGGTACAGCCAA
exon2, 948 bp	*T. reesei*	P19	GTACCACAGATGGCACTTGCGTTACC
		P20	AGAAATTCGCCTACTTTCTTGCGAGA
T*adh*I, 365 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	P21 P22	AAGAAAGTAGGCGAATTTCTTATGAT AGCCCAGAATTCGGCCTAGGCGGCCATCCGT GTGGAAGAAC(*Sfi*I/*Avr*II/*Eco*RI)
To construct gene expressing cassette: **P*tpi-xyn*2s-Tr *EGII-cwp*2-T*adh*I**, 2566 bp			
P*tpi-xyn*2s-Tr *EGII* , 1313 bp	P*tpi-xyn*2s-Tr *EGII*-T*adh*I	P23[i]	TTTGAGAAATCTTTCTTGCGAGACAC
*cwp*2-T*adh*I, 572 bp	pGEM-P*tpi-xyn*2s-Aa *BG-cwp*2-T*adh*I	P24[j]	CGCAAGAAAGATTTCTCAAATCACTG

[a] P*tpi* was the *TPI*1 promoter from *S. cerevisiae*; [b] *xyn*2s was the secretion signal gene from *T. reesei*; [c] This plasmid was shown in Table 1; [d] This plasmid was shown in Table 2; [e] T*adh*I was the *ADH*1 terminator from *S. cerevisiae*; [f] P13 was used as lower primer with P3 as upper primer; [g] *cwp*2 was the anchored peptide-encoding sequence that was 207 bp at 3' end of *S. cerevisiae* cell wall protein 2; [h] P14 was used as upper primer with P12 as lower primer; [i] P23 was used as lower primer with P15 as upper primer; [j] P24 was used as upper primer with P22 as lower primer.

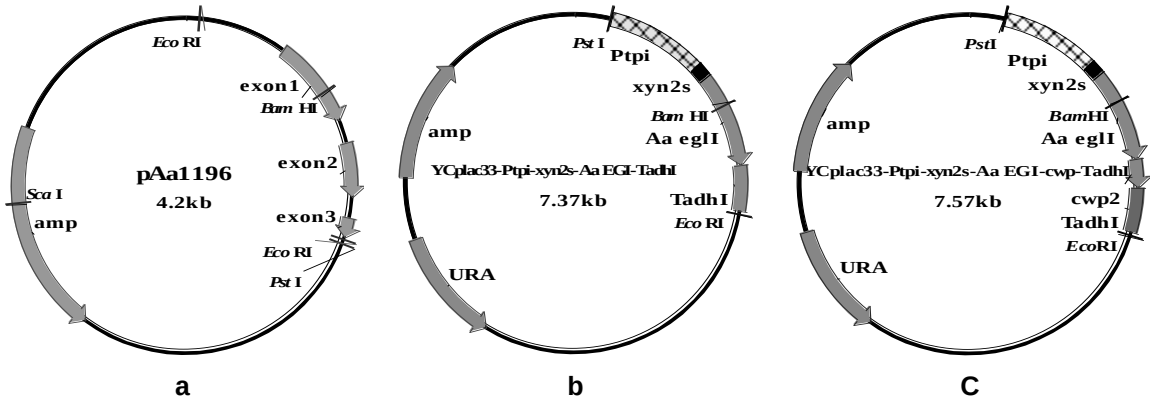

Fig. 1. Diagram illustrating three plasmids constructed in this study. (**a**) Physical maps of pAa1196 for cloning the full length of Aa *eg*1; the sequences of three exons and two introns are shown in GenBank accession no. JQ581513, respectively; (**b**) Physical maps of expression plasmid YCplac33-P*tpi-xyn*2s-Aa *EGI*-T*adh*I for Aa EGI secretory expression by no anchored sequence *cwp*2 used; (**c**) Physical maps of expression plasmid YCplac33-P*tpi-xyn*2s-Aa *EGI-cwp*2-T*adh*I for yeast cell surface displaying Aa EGI by the anchored sequence *cwp*2 used.

Supplementary data 1. Gene sequence (GenBank: JQ581513)

```
   1 AAGACTGACA TCCTCCGCTT GAATCACATT ATTCGGCACG GGCTGATCGC
  51 CTCCGTACCC AAATCTCCGA CGGGTGGGCC CAAAGGGTCC AAGATCCTCA
 101 TCTTTGATAC ACTAGCTTCG GACTCGATTT GCCAAATGTG CCAATTGTCT
 151 CGTACTTCTG GCAGTGGACA TGCATCGCGT TCAAGCTGTG TCTCCTGCTG
 201 TAGTCTGCAT ATTCAAGCTA TAAATTCGGC TCATAACCAC CATCTCAGAA
 251 GCCAGCATCT CATCCCCCAT CAAACATCAC AGCAAGCATT TCCCATCATC
 301 GAAAATGAAG GCCTTCTACT TCCTCGCATC TCTCGCCGGT GCCGCTGTCG
          Secretion signal sequence
 351 CCCAACAGAC TCAACTCTGC GATCAATACG CTACCTACAC CGGTAGCGTG
        Exon 1
 401 TACACCATCA ACAACAATCT GTGGGGCAAG GACGCTGGCT CCGGCTCGCA
 451 ATGCACCACC GTGAAATCCG CCTCCAGCGC GGGCACATCT TGGTCGACGA
 501 AATGGAACTG GTCCGGCGGA GAGAACTCGG TTAAGAGCTA CGCCAACTCC
 551 GGCTTGAGCT TTAACAAGAA GCTTGTTAGC CAGATCAGCC GGATCCCCAC
 601 CGCCGCCCAG TGGAGCTATG ACAACACGGG CATCCGCGCG GACGTGGCCT
 651 ACGATCTCTT CACCGCGGCG GATATCAACC ATGTCACCTG GAGTGGGGAC
 701 TATGAGCTCA TGATCTGGTA TGTACCCTCC TGTGGAAACC AAGCGCGCCC
                          Intron1
 751 AGAAAGCTCA GCGGCATGCG AACCGAGCCC ACCATCTCAC CCCTCTTTTC
 801 CGTTAAAATA GGCTTGCTCG TTACGGCGGC GTCCAGCCCC TAGGGTCCAA
                    Exon2
 851 GATCGCGACC GCCACGGTCG AGGGCCAGAC CTGGGAGCTG TGGTACGGCG
 901 TCAACGGCGC GCAGAAGACC TACAGCTTTG TGGCTCCGAC CCCGATCACC
 951 TCGTTTCAGG GCGACGTCAA CGATTTCTTC AAGTACCTGA CGCAGAACCA
1001 CGGCTTCCCC GCCAGTAGCC AGTACCTCAT CAGTAAGCCA TTTCAACCTC
                                                  Intron2
1051 TCCTCCCTCC AGTTTCCCGA GCACTGTCTG TGGGAGAGCT TTGTTCTGAT
1101 GGAAATTACC TGCTAGCGCT CCAATTCGGC ACCGAGCCTT TCACGGGTGG
                          Exon 3
1151 TCCCGCCACG CTCACTGTCT CGGACTGGTC TGCCAGCGTA CAGTAG 1196
```