



Effect of *Pichia pastoris* host strain on the properties of recombinant *Aspergillus niger* endoglucanase, EglB

Shazilah Kamaruddin¹, Nor Muhammad Mahadi², Rosli Md Illias³, Osman Hassan⁴, Suhaila Sulaiman⁵, William Broughton⁶, Izwan Bharudin¹, Farah Diba Abu Bakar^{1*}, Abdul Munir Abdul Murad¹

¹School of Biosciences & Biotechnology, Faculty of Science & Technology, Universiti Kebangsaan Malaysia, Bangi 43600 Selangor, Malaysia.

²Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

³Department of Bioprocess Engineering, Faculty of Chemical and Natural Resources Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia.

⁴School of Chemistry & Food Technology, Faculty of Science & Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia.

⁵Felda Global Ventures Research and Development Sdn Bhd, FGV Innovation Centre (Biotechnology), Bandar Enstek, 71760 Negeri Sembilan, Malaysia.

⁶BAM Federal Institute for Materials Research and Testing, Unter den Eichen 87, 12205 Berlin, Germany.
Email: fabyff@ukm.edu.my

ABSTRACT

Aims: The methylotrophic yeast *Pichia pastoris* is widely used to express foreign proteins fused to secretion signals. As the effect of the expression host on the final protein product is unclear, we compared the properties of an endoglucanase (*eglB* of *Aspergillus niger*) expressed in two different *P. pastoris* strains.

Methodology and results: Full-length cDNA encoding endoglucanase of *A. niger* strain ATCC10574 was isolated and expressed in *P. pastoris* X33 (the methanol utilisation plus phenotype, Mut⁺) and *P. pastoris* GS115 (slow methanol utilisation, Mut^s). EglB-GS115 showed the highest activity and stability at 60 °C while EglB-X33 was most active at 50 °C. EglB-X33 was active towards other substrates such as arabinogalactan, guar gum and locust bean gum besides its specific substrate, carboxymethyl cellulose (CMC). However, EglB-GS115 was only active on CMC. The affinity of EglB-X33 towards CMC ($K_m = 7.5$ mg/mL and specific activity 658 U/mg) was higher than that of EglB-GS115 ($K_m = 11.57$ mg/mL, specific activity 144 U/mg).

Conclusion, significance and impact of study: Although *eglB* was cloned in the same expression vector (pPICZαC), two different characteristics of enzymes were recovered from the supernatant of the different hosts. Thus, expression of recombinant enzyme in different *P. pastoris* strains greatly affects the physical structure and biochemical properties of the enzyme.

Keywords: cellulase, endoglucanase, glycosylation, methanol utilisation phenotype, recombinant enzyme

INTRODUCTION

Pichia pastoris is widely used for high-level protein production of complex proteins in industry and research (Krainer *et al.*, 2012). Methanol serves both as a carbon source for *P. pastoris* and as an inducer of the alcohol oxidases AOX1 and AOX2. Alcohol oxidase catalyses the conversion of methanol to formaldehyde and hydrogen peroxide (Cereghino *et al.*, 2000). Since AOXs catalyse the first step in methanol utilisation, they can accumulate up to 35% of the total protein in cells grown on limited methanol (Sreekrishna *et al.*, 1997) and for this reason, the strongest promoter that modulates alcohol oxidase synthesis is widely used to drive heterologous protein expression in *P. pastoris*.

The wild-type *P. pastoris* strain which is a methanol utilisation plus (Mut⁺) phenotype, grows well on methanol. Loss of *aox1*, and thus loss of most of the cell's alcohol oxidase activity, results in strains that are phenotypically slow in methanol utilisation (Mut^s) (Sreekrishna *et al.*, 1997). Mut^s cells therefore grows poorly on methanol containing media. The methanol-utilisation minus phenotype, Mut⁻ in which both *aox*-genes have been rendered inactive however, has totally lost the ability to grow on methanol.

The *Pichia* strain X-33 is a wild-type while GS115 strain has a mutation in the histidinol dehydrogenase gene (*his4*), which prevents it from synthesising histidine. Transformation of *P. pastoris* strains GS115 or X-33 with linear constructs favours single cross-over recombination

*Corresponding author

at the *aox1* locus. Most of the resulting transformants will therefore be Mut⁺ but *aox1* sequences in the plasmid can force recombination at the 3'-*aox1* terminus. This causes disruption of the native *aox1*-gene and creates Mut^S transformants (Cereghino *et al.*, 2000).

Expression of heterologous proteins in different *P. pastoris* strains affects expression levels and may lead to enhanced expression of targeted proteins. Since the *aox1* promoter is highly inducible, most researchers use Mut⁺ hosts (Macaulay-Patrick *et al.*, 2005). As an example, expression level of the peroxidase of *Coprinus cinereus* in a Mut⁺ host yielded three-fold higher as compared to a Mut^S strain (Kim *et al.*, 2009). Nevertheless, Mut^S strains have other advantages (Ascacio-Martinez and Barrera-Saldana, 2004; Guo *et al.*, 2008; Krainer *et al.*, 2012). For example, the recombinant growth hormone (rCGH) from the canine, *Canis familiaris* was expressed at 40 and 15 µg.mL⁻¹ for constructs in Mut^S and Mut⁺ hosts, respectively (Ascacio-Martinez and Barrera-Saldana, 2004). Thus, it is difficult to predict which strain/isolate will be suitable for expressing a particular protein.

To the best of our knowledge, the effect of different *P. pastoris* strains on the properties of the enzymes produced in them is not known. To address this question, we cloned the cDNA of *eglB* from *Aspergillus niger* strain ATCC10574 and produced a his-tag recombinant endoglucanase (EC 3.1.3.2) in *P. pastoris* strains GS115 and X33. Although the endoglucanase EglB is derived from a mesophilic fungus, *A. niger*, it has a relatively high optimum temperature (70 °C) and is stable for up to 3 h at 60 °C retaining over 90% of its original activity Li *et al.* (2012). In addition, EglB remained active in the pH 3 to 10 range with a tendency towards increased stability under more alkaline conditions. For all these reasons it is widely used to degrade biomass as well as in the industries such as textiles and paper and pulp industry. Thus expression of this protein in various hosts might affect their physical and biochemical properties. The aim of this study is to determine the influence of different *P. pastoris* strains on the biochemical properties of the endoglucanase, EglB produced.

MATERIALS AND METHODS

Fungi and culture conditions

Aspergillus niger ATCC10574 was obtained from the American Type Culture Collection (ATCC) (ATCC, Manassas, VA, USA) and was cultured in Mendels medium as described by Hong *et al.* (2001) with 1% (w/v) Avicel (Fluka Sigma-Aldrich, Basel, Switzerland) as inducer. The medium was autoclaved and inoculated with 10⁶ spores/mL of *A. niger*. Fungal mycelia were harvested after 48 h of growth by filtration through filter paper, frozen with liquid nitrogen and stored at 80 °C until further use.

Amplification and cloning of endoglucanase cDNA

Total RNA was extracted according to Oh *et al.* (2009). The RNA was used as a template for the synthesis of the first cDNA strand using a Superscript III First Strand cDNA Synthesis Kit (Invitrogen Life Technologies, Grand Island, NY, USA) and an oligo-dT primer in accordance to the manufacturer's procedure.

To obtain the full-length *eglB* cDNA, specific primers (Forward primer: 5'GTTGATCGAGCAGTCGTAGC3'; Reverse primer: 5'TCAGCGACAACTGCCTAC3') were designed and synthesised according to the DNA sequence of *eglB* from the *A. niger* Genome Database of the Joint Genome Institute (JGI), Walnut Creek, CA, USA (<http://genome.jgi.psf.org/Aspni5/Aspni5.home.html>). PCR was performed using a KOD Hot Start DNA Polymerase Kit (Merck KGaA, Darmstadt, Germany). The amplified cDNA was cloned into a pGEMT-Easy vector (Promega, Madison, WI, USA) and sequenced.

Sequence analysis

The full-length sequence obtained was translated into an amino acid sequence and a similarity search for the amino acid sequence against public databases was carried out using Basic Local Alignment Search Tool (BLAST) (<http://www.ncbi/BLAST>) (Altschul *et al.*, 1997). The NetNGlyc 1.0 server (<http://www.cbs.dtu.dk/services/NetNGlyc>) and OGPET 1.0 (<http://ogpet.utep.edu/OGPET>) predicted the *N*-glycosylation and *O*-glycosylation sites.

Homology modelling

Based on the structure of *Thermoascus aurantiacus* endoglucanase 1H1N (DOI:10.2210/pdb1h1n/pdb) from the Protein Data Bank (PDB), a three-dimensional homology model of EglB was generated using MODELLER9v11 (Webb *et al.*, 2014). The sequence identity between EglB and endoglucanase from *T. aurantiacus* was 72%. The quality of the final model was checked using PROCHECK (Laskowski *et al.*, 1993).

Molecular cloning and expression in *P. pastoris*

Construction of expression vectors was performed using *eglB* cDNA as the template. A fragment encoding mature EglB (without the signal peptide) was amplified using a 5'-primer containing a *Cla*I site (5' ATCGATTGATGAATTGGCCTAC 3') while the 3'-primer contained a *Kpn*I site (5'GTGATTCTAGATTGTGAACAGTAGGC 3') and subsequently, this amplified fragment was cloned into pPICZαC vector (Invitrogen Life Technologies, Grand Island, NY, USA). The resulting vector obtained was designated EglB-pPICZαC.

The EglB-pPICZαC vector was linearised and then used to transform *P. pastoris* strains X-33 and GS115 by electroporation (Electroporator 2510, Eppendorf, Hauppauge, NY, USA). Transformants obtained were

screened for expression by identifying putative multi-copy recombinants on plates of YPD containing 1% (w/v) yeast extract, 2% (w/v) dextrose, 2% (w/v) peptone and 2% (w/v) agar, with an increasing concentration of Zeocin® (1000 µg/mL, 1,500 µg/mL, 2000 µg/mL). Subsequently, transformants were tested for the Mut phenotype by growing on Minimal Methanol plates. The endoglucanase gene integration into the *P. pastoris* genome was validated by PCR amplification with the primers, 5'AOX1 (5'GACTGGTCCAATTGACAAGC3') and 3'AOX1 (5'GCAAATGGCATTCTGACATCC 3').

Positive transformants were used to inoculate 100 mL of BMGY medium (buffered glycerol-complex medium, Invitrogen) in 1 L Erlenmeyer flasks. Only GS115 transformants were supplemented with 4×10^{-3} % (w/v). The cells were cultivated with rotary shaking at 30 °C and 250 rpm until the OD₆₀₀ reached 2-6, indicating log-phase growth. Cells to be harvested were centrifuged at 3000 × g for 5 min in room temperature. Cell pellets were resuspended in 20 mL BMMY medium (buffered methanol-complex medium, Invitrogen) supplemented with 4×10^{-3} % (w/v) histidine (for GS115 transformants) and cultivated for 3 days. A final concentration of 3% methanol (v/v) was added to the yeast cultures every 24 h to maintain induction. Pooled culture supernatants were concentrated using Amicon Ultra Centrifugal Filter Devices (Merck Millipore, Billerica, Massachusetts, USA). Secreted proteins were analysed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

Purification of recombinant EglB

Recombinant endoglucanase B proteins fused to a 6x histidine tag were subjected to Ni-NTA Sepharose column purification using an AKTA Prime System (Amersham, Buckinghamshire, UK). The target protein fractions were pooled and analysed by 12% SDS-PAGE.

Biochemical characterisation of EglB

The amount of recombinant EglB produced was measured according to the Bradford Method (Bradford, 1976). Enzymes were assayed using carboxymethyl cellulose (CMC) (Sigma Aldrich, St. Louis, MO, USA) as the substrate. The CMC case standard assay was performed by adding approximately 1 µg purified EglB to 2.5% (w/v) CMC in 50 mM sodium acetate buffer pH 4 and incubated at 50 °C for 30 min. Dinitrosalicylic acid (DNS) assay was employed to determine the reducing sugars released (Miller, 1959). Glucose was used as a standard, and the liberation of glucose was measured at 540 nm. Enzyme activity of one unit was defined as the amount of enzyme that produced 1 µmole of glucose in one minute using the standard assay conditions.

To determine the optimum temperature of EglB, standard assays as described previously were carried out for 30 min at 20, 30, 40, 50, 60, 70, and 80 °C. Stability of EglB in various temperatures was assessed by measuring the residual activity after pre-incubation of the enzyme in the absence of substrate at various temperatures

between 50 to 70 °C for 30 min. Subsequently, the residual activity of the enzyme was assayed by the standard assay method as described previously. The optimum pH was analysed by running the standard assay over a range of pH values from 3 to 8, while pH stability of the enzyme was determined in pH 3 - 8 at 50 °C for 30 min; residual activity was recorded as described previously. The buffers used were sodium acetate (pH 3 to 5) and potassium phosphate (pH 6 to 8).

Michaelis-Menten constants (K_m) and maximum velocities of substrate hydrolysis (V_{max}) were calculated from Lineweaver-Burk plots. The reactions were performed by incubating 0.1 mL of purified EglB with 0.9 mL 50 mM sodium citrate buffer (pH 4.0) with various substrates over a range of concentrations at 50 °C for 30 min. The substrates used were Avicel (Sigma Aldrich), xylan (Thomson Biotech (S) Pte. Ltd., Xiamen, Fujian, China) pNP-β-D-cellobioside, (Sigma Aldrich), pNP-β-D-glucopyranoside (Sigma Aldrich), arabinogalactan (Sigma Aldrich), LBG (Sigma Aldrich), guar gum (Rama Industries, Ahmedabad, Gujarat, India) and CMC. Enzyme activities, catalytic constants (K_{cat}) and specificity constants (K_{cat}/K_m) were calculated using standard procedures. All experiments were performed with three technical replicates and two biological replicates.

RESULTS

Isolation and sequence analysis of *A. niger* *eglB* cDNA

The *eglB* cDNA comprised 996 bp, encoding a polypeptide of 331 amino acid residues and had an estimated molecular mass of 34.8 kDa. BLAST analyses showed that the EglB amino acid sequence was 100% identical to that of EglB of *A. niger* strain BCRC31494 (Accession number: ACT68011.1) (Li *et al.*, 2012).

EglB and the endoglucanase of *Thermoascus aurantiacus* share 72% identity and for this reason the *T. aurantiacus* endoglucanase 1H1N (DOI:10.2210/pdb1h1n/pdb) was used to build a model of EglB (Figure 1) (Van-Petegem *et al.*, 2002). This model was verified by PROCHECK and Ramachandran Plot analyses which showed that 91% of the residues were in the most favoured regions. Three predicted N-glycosylation sites were found on the protein (Figure 1), two (N100 and N211) at the surface while N38 was located inside the barrel, next to the two catalytic residues. Glycosyl hydrolase family 5, possesses two invariant glutamic acid residues that correspond to a nucleophile and proton donor (Wang *et al.*, 1993) both of which were present in EglB (Figure 1). The Glu-160 residue is the potential catalytic proton donor, while Glu-226 is the potential nucleophile in the displacement reaction. Several other important residues surrounded the N38 sites including Trp-299, a residue implicated in substrate-binding and His-120 which contacts the sugar residue in subsite-1 (Dominguez *et al.*, 1996; Van-Petegem *et al.*, 2002). A further two residues (His-224

and Tyr-226) near N38 that interact with the nucleophile were also present.

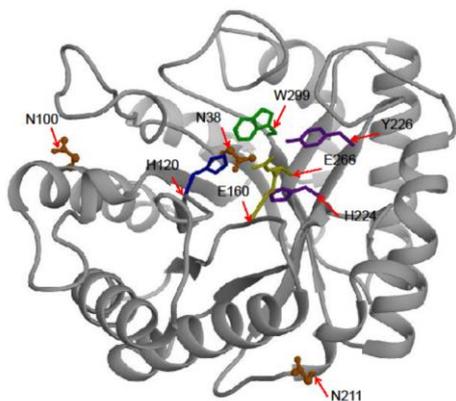


Figure 1: Homology model of EglB (generated using the MODELLER 9.11 programme) showing all secondary structure elements with arrows representing important amino acids. The two catalytic residues labelled E160 and E266 and three N-glycosylation sites predicted are labelled as N38, N100 and N211. Aromatic residues possibly involved in the binding of substrates are also shown and labelled as H224, Y226, H120 and W299.

Expression and purification of EglB

Partially purified samples containing EglB of both *P. pastoris* strains X-33 and GS115 revealed a recombinant protein of ~46 kDa for EglB-X33 and ~58 kDa for EglB-GS115 (Figure 2). These values are slightly higher than that predicted for EglB with a histidine tag (38 kDa). Most likely this was due to the N-glycosylation, something that was predicted by *in silico* analyses of the EglB amino acid sequence. No O-glycosylation sites were found.

Biochemical properties of EglB

Purified EglB-X33 and EglB-GS115 exhibited maximal activity on CMC at pH 4.0 (Figure 3a). Activity declined rapidly to 40% in alkaline pH at pH 8.0, a characteristic of most fungal cellulases that often limits their application under neutral or alkaline conditions (Wang *et al.*, 2005). Nevertheless, both endoglucanases were very stable over a wide range of pHs; stability even increased towards alkaline pH (Figure 3b). Stability was marked between pH 5.0 and 8.0. Both endoglucanases retained more than 80% of the maximum activity at pH 8.0 after incubation at 50 °C for 30 min and lost almost 80% of its maximal activity at pH 4.0 under the same conditions.

The optimum temperature of EglB-X33 for hydrolysis of CMC was 50 °C (at pH 4.0) (Figure 3c) with over 50% of maximal activity at 70 °C. EglB-X33 was relatively stable up to 60 °C (Figure 3d) retaining 80% of its original activity but was completely inactive at 70 °C. In contrast,

the optimum temperature of EglB-GS115 for CMC hydrolysis was 60 °C (pH 4.0) (Figure 3b) and the enzyme had a broader optimum temperature range, with over 50% of maximal activity at 80 °C. EglB-GS115 was 100% stable up to 60 °C (Figure 3b) but only retained 20% activity at 70 °C.

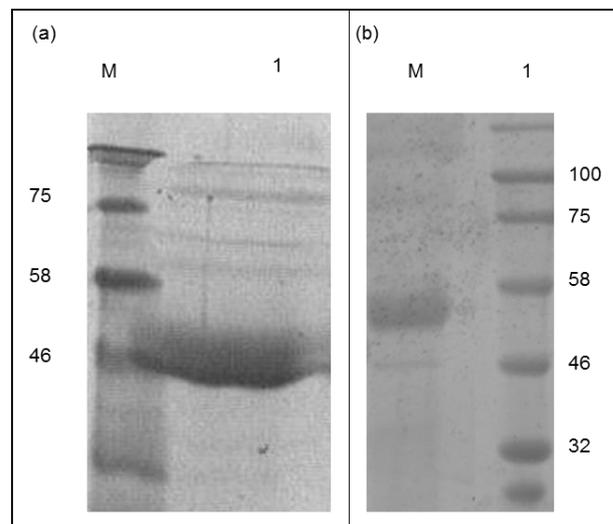


Figure 2: SDS-PAGE of recombinant EglB endoglucanase on 12% polyacrylamide gels. a) Lane 1: M, Protein Marker Broad Range (NEB, Ipswich, Massachusetts, USA), 2: Purified protein of *P. pastoris* X-33/EglB_pPICZαC. b) Lane 1: Purified protein of *P. pastoris* GS115/EglB_pPICZαC, 2: M, Protein Marker Broad Range (NEB, Ipswich, Massachusetts, USA).

Kinetic parameters of recombinant EglB

EglB-X33 had significant activity on four soluble substrates: arabinogalactan, CMC, guar gum and LBG but were highest on CMC (specific activity - 658 U/mg) followed by arabinogalactan, guar gum and LBG (344 U/mg, 46 U/mg and 39 U/mg respectively) (Table 1). However, the enzyme was inactive on insoluble or synthetic substrates such as Avicel™, pNP-β-D-cellobioside, pNP-β-D-glucopyranoside and xylan. Interestingly, EglB-GS115 was specific to CMC (specific activity 144 U/mg) being inactive on all other soluble and insoluble substrates.

Specificity of enzymes (= the specificity constant) is often expressed in terms of the K_{cat}/K_m (catalytic efficiency), a useful index for comparing the relative rates of activity on alternative, competing substrates (Pham *et al.*, 2011). EglB-X33 had broad substrate specificity and was able to hydrolyse both β-(1,4)-linked glucose and β-(1,4)-linked mannose residues, as well as β-(1,3)-linked galactan. Yet the K_{cat}/K_m of EglB-X33 was highest towards arabinogalactan ($136 \text{ s}^{-1} \cdot \text{mg mL}^{-1}$) rather than CMC ($56 \text{ s}^{-1} \cdot \text{mg mL}^{-1}$) which gave the highest specific activity. EglB-GS115 that was only able to hydrolyse CMC, had a K_{cat}/K_m of $11.3 \text{ s}^{-1} \cdot \text{mg mL}^{-1}$.

Table 1: Kinetic parameters of EglB.

Enzymes	Substrate/Kinetics parameters	V_{max} ($\mu\text{mole}/\text{min}$)	K_m (mg/mL)	K_{cat}	K_{cat}/K_m	Specific activity (U/mg)
EglB-X33	CMC	0.0461	7.50	417.38	55.65	658
	Arabinogalactan	0.0241	1.60	218.65	136.66	344
	Locust bean gum	0.0027	0.30	24.72	82.39	38.9
	Guar gum	0.0032	0.49	29.13	59.46	45.9
	Xylan	n.d				
	Avicel	n.d				
	pNP- β -D-cellobioside	n.d				
	pNP- β -D-glucoopyranoside	n.d				
EglB-GS115	CMC	1.44	11.57	130	11.3	144
	Arabinogalactan	n.d				
	Locust bean gum	n.d				
	Guar gum	n.d				
	Xylan	n.d				
	Avicel	n.d				
	pNP- β -D-cellobioside	n.d				
	pNP- β -D-glucoopyranoside	n.d				

n.d.: not detectable

Table 2: Biochemical properties of EglB-X33 compared to other EglB in various hosts.

Enzymes/properties	EglB-Wild Type	EglB_GS115_Li	EglB-GS115	EglB-X33
Origin	<i>A. niger</i> ATCC10577	<i>A. niger</i> BCRC31494	<i>A. niger</i> ATCC10574	<i>A. niger</i> ATCC10574
Protein mass (without deglycosylation) kDA	43	51	~58	~46
Expression host	Native	<i>P. pastoris</i> GS115 (MutS)	<i>P. pastoris</i> GS115 (MutS)	<i>P. pastoris</i> X33 (Mut+)
Expression vector	-	pPICZ α C	pPICZ α C	pPICZ α C
Optimum temperature	60 °C	70 °C	60 °C	50 °C
Temperature stability (relative activity > 70 %)	20 - 50 °C	30 - 60 °C	50 - 60 °C	50 - 60 °C
K_m towards CMC	4.7 mg/mL	135 mg/mL	11.57 mg/mL	7.5 mg/mL
Specific activity (U.mg ⁻¹)	24.6	0.76	144	658
Reference	(Vidmar <i>et al.</i> , 1984)	(Li <i>et al.</i> , 2012)	This study	This study

1 U: The amount of enzyme releasing 1 μmole of glucose per minute under standard condition

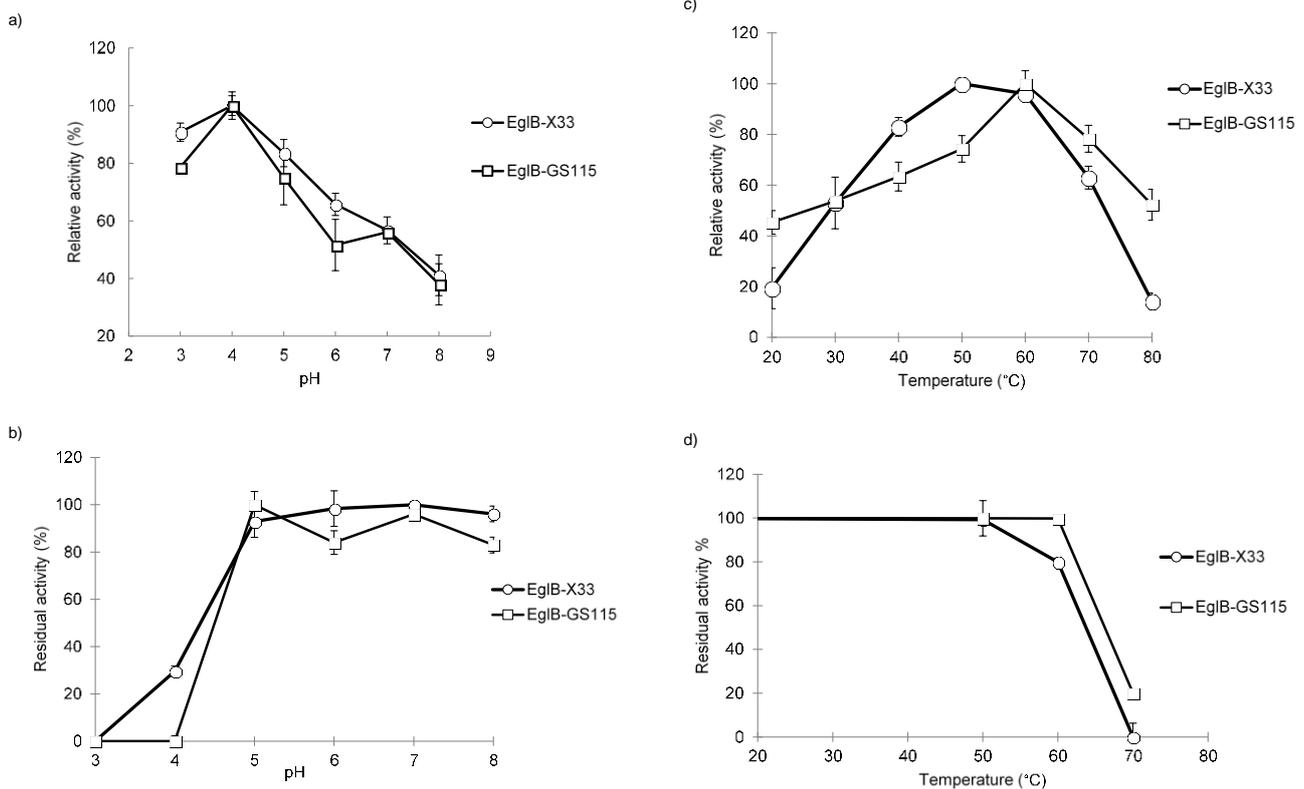


Figure 3: Characterisation of EglB in terms of: (a) optimum pH for maximal activity. The highest activities taken as 100% relative activity that were equivalent to 658 U/mg for EglB-X33 and 144 U/mg for EglB-GS115 respectively; b) stability of EglB at various pH values; (c) optimum temperature for maximal activity. The highest activities taken as 100% relative activity that were equivalent to 658 U/mg for EglB-X33 and 144 U/mg⁻¹ for EglB-GS115 respectively, and; (d) stability at various temperatures.

Comparison of EglB produced in *P. pastoris* strain X33 and strain GS115

We compared the biochemical properties of recombinant EglB (EglB-X33; Mut⁺ and EglB-GS115; Mut^S) with that produced naturally in *A. niger* (EglB-wild type) along with EglB expressed in *P. pastoris* GS115; Mut^S by (Li *et al.*, 2012) (EglB_GS115_Li) (Table 2). Variation was abundant including protein size that differed slightly between all endoglucanases. EglB-GS115_Li was the largest protein (58 kDa) while EglB-X33, EglB-GS115 and EglB-wild type were 46, 51 and 43 kDa respectively. The optimum temperature of EglB-X33 was lower than that of EglB-wild type and EglB-GS115 in this study with an optimum of 50 °C. Optimum temperatures of EglB-GS115 and EglB-wild type were both 60 °C while EglB_GS115_Li was 70 °C. K_m values towards CMC varied even more. EglB-X33 had a K_m of 7.5 mg/mL similar to that EglB-wild type (4.7 mg/mL) (Vidmar *et al.*, 1984). Our EglB-GS115, along with the one reported by Li *et al.* (2012) (EglB_GS115_Li) had a very low affinities towards CMC (K_m values of 11.6 and 135 mg/mL respectively). There are also variations seen in specific activities. EglB-X33 had the highest specific activity (658 U/mg) while EglB-

GS115 was 144 U/mg, EglB-wild type 25 U/mg and EglB_GS115_Li 0.8 U/mg (Li *et al.*, 2012). These differences clearly show that expression of the same protein in different strains and phenotype has large effects on the biochemical properties of the enzyme.

DISCUSSION

Glycosyl hydrolase family 5 includes cellulases, mannanases and xylanases. Biely and Tenkanen (1999) showed that all *T. reesei* endoglucanases are able to hydrolyse glucomannan. Vlasenko *et al.* (2009) examined family 5 endoglucanases phylogenetically and suggested that endoglucanase and mannanase activities might originate from an ancestral protein that used mannan as a substrate.

We posed the question – is the amino acid sequence of an endoglucanase the sole determinant of substrate specificity? In this study, we report the characterization of EglB, an endoglucanase from *A. niger* expressed in *P. pastoris* strain X33 (EglB-X33) and strain GS115 (EglB-GS115). This enzyme was previously expressed and characterised in *P. pastoris* strain GS115 by Li *et al.* (2012) (EglB_GS115_Li). By using identical coding

sequences but expressing the constructs in various *P. pastoris* cytoplasmic backgrounds, we expected only minimal differences between the recombinant enzymes. However, detailed study of the recombinant enzymes revealed substantial differences in activity, substrate affinity, optimum temperature and stability. EglB-X33 can degrade locust bean gum, a hemicellulose substrate besides CMC which is a normal substrate for endoglucanases while EglB-GS115 in this study was restricted to hydrolyse only CMC. Since both EglB-X33 and EglB-GS115 share the same gene sequence, the differences in activity, substrate affinity, optimum temperature and stability might be due to the post-translational modifications in the host. Thus, it can also be postulated that the expression in different Mut phenotype strains is the main factor that contributes to the differences between these two enzymes. Furthermore, the methanol utilisation phenotype (Mut) affected growth rate probably led to different glycosylation patterns. Mut^s strains grew more slowly, suggesting that post-translational processes may also take longer while Mut⁺ strains are characterised by a higher growth rate than Mut^s. Computer predictions suggested that EglB in this study has three N-glycosylation sites and when produced in strain GS115 (slow methanol utilisation) resulted in a molecular mass of ~58 kDa but expression in strain X-33 gave a molecular mass of only ~46 kDa. We suggest that the length of the oligosaccharide chain at the N-glycosylation site varied because glycosylation is not a template-driven/proof-reading process. Thus, factors such as the expression host, development and differentiation of the host-cells that are all dependent on culture conditions including media composition (Bretthauer *et al.*, 1999; Guo *et al.*, 2008) modulate the molecular properties of the expressed protein. Different growth rates also might lead to different glycosylation patterns, thereby producing different glycoproteins. The occurrence of glycosylation in some proteins has been suggested to have a little effect on its activity (Han *et al.*, 1999; Hamilton *et al.*, 2007), but glycosylation on some proteins showed tremendous effect on its function such as changing its role from inhibitor to a substrate (Mochizuki *et al.*, 2001). Thus the differences in glycosylation patterns in EglB-X33 and EglB-GS115 may have resulted in different lengths and moieties of carbohydrate chains that subsequently affect the affinity of both recombinant proteins towards CMC as well as their activity. EglB-X33 demonstrates higher activity towards CMC (658 U/mg) compared to EglB-GS115 (144 U/mg). EglB-X33 also has a stronger affinity towards CMC as compared to EglB-GS115 in which the K_m value is 7.5 mgmL⁻¹ and 11.57 mg/mL respectively. It can be postulated that glycosylation in the slower growing Mut^s strain is more complete since it takes a longer time to fulfil the post-translational modification process. Therefore, the Mut^s strain might produce a longer carbohydrate chain that makes the EglB enzyme more stable at a higher temperature as compared to when the enzyme was expressed in the Mut⁺ strain. However, more carbohydrate chains produced near the loop of binding site and as well as the active site might affect its affinity

and activity towards CMC. The EglB produced in the Mut⁺ strain (EglB-X33) might be less glycosylated at this particular site, therefore the enzyme becomes less stable at a higher temperature but has strong affinity and activity towards CMC. As shown in Figure 1, a predicted glycosylation site was identified at the asparagine residue (N38). This position is near the substrate-binding site which is between $\beta 6$ and $\alpha 6$ (Van-Petegem *et al.*, 2002). This glycosylation site is also near two catalytic residues at the active site of this protein. Therefore, the addition of a carbohydrate chain at this position might interrupt the binding of the substrate with this enzyme and also affect its catalytic activity.

EglB-X33 has a lower optimum temperature and stability compared to the EglB-GS115. The optimal temperature of EglB-X33 for CMC hydrolysis was achieved at 50 °C in sodium acetate buffer pH 4.0 (Figure 3b). The enzyme shows a broader optimum temperature range, having 96% of its maximal optimum activity at 60 °C and 63% at 70 °C. A thermostability assay of EglB-X33 revealed that it is relatively stable up to 60 °C (Figure 3b). EglB-GS115 has maximal activity at 60 °C and possesses high thermotolerance having 78% of its maximal activity at 70 °C and more than 50% activity at 80 °C. In terms of thermostability, EglB-X33 completely lost its activity after 30 min incubation at 70 °C, while EglB-GS115 retained at least 20% of its maximal activity after 30 min incubation at 70 °C.

In general, proteins that have been glycosylated might display differences in stability even though there are no major changes in their structure. It has been suggested that a single carbohydrate is sufficient to stabilise the region surrounding the glycosylation site (Imperiali and O'Connor, 1999) and N-glycosylation has been given the most attention since it often contribute to protein function (Hamilton *et al.*, 2007). EglB was predicted to have three N-glycosylation sites; where two of them are positioned at the surface of the enzyme. Some amino acid residues at the surface of the enzyme have been proven to play an important role on the stability of the enzyme (Wang *et al.*, 2005). The rate of post-translational modification process might lead to differences in the length of the carbohydrate chain or carbohydrate moiety that eventually will affect the optimum temperature and stability of the protein. Furthermore, the thermostability in endoglucanases is also fold specific (Yennamalli *et al.*, 2011). Since glycosylation can affect the secondary structure of a protein, it might have an ability to direct protein folding, destabilise the unfolded state or enhance the reversibility of the unfolding process by increasing the protein solubility or discouraging the formation of aggregates. Therefore, the modification of secondary structures might also enhance the overall stability of the protein as well as enable it to perform a new function at the affected site (Imperiali and O'Connor, 1999).

CONCLUSION

In this study, we demonstrate the effect of different *P. pastoris* strains and phenotype towards the biochemical

properties of EglB. In future experiments, we will carry out functional and mutational analysis particularly at the glycosylation site to examine the role of different cytoplasmic backgrounds (especially the Mut-phenotype) on glycosylation and other post-translational modifications on the activity and conformation of EglB.

ACKNOWLEDGEMENTS

This work was supported by the National Biotechnology Directorate, MOSTI, Malaysia through research grant 07-05-16-MGI-GMB12.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.
- Ascacio-Martinez, J. and Barrera-Saldana, H. (2004). Production and secretion of biologically active recombinant canine growth hormone by *Pichia pastoris*. *Gene* **340**, 261-266.
- Biely, P. and Tenkanen, M. (1999). Enzymology of hemicellulose degradation. In: *Trichoderma and Gliocladium*. Enzymes, Biological Control and Commercial Applications. Kubicek, C. P. and Harman, G. E. (eds.). Taylor and Francis, London. pp. 25-47.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254.
- Bretthauer, R. K. and Castellino, F. J. (1999). Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnology Applied Biochemistry* **30**(Pt 3), 193-200.
- Cereghino, J. L. and Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Review* **24**, 45-66.
- Domínguez, R., Souchon, H., Lascombe, M. B. and Alzari P, M. (1996). The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. *Journal of Molecular Biology* **257**, 1042-1051.
- Guo, M., Hang, H., Zhu, T., Zhuang, Y., Chu, J. and Zhang, S. (2008). Effect of glycosylation on biochemical characterization of recombinant phytase expressed in *Pichia pastoris*. *Enzyme and Microbial Technology* **42**, 340-345.
- Hamilton, S. R. and Gerngross, T. U. (2007). Glycosylation engineering in yeast: The advent of fully humanized yeast. *Current Opinion in Chemical Biology* **18**, 387-392.
- Han, Y and Lei, X. G. (1999). Role of glycosylation in the functional expression of an *Aspergillus niger* phytase (phyA) in *Pichia pastoris*. *Archives of Biochemistry and Biophysics* **364**, 83-90.
- Hong, J., Tamaki, H., Akiba, S., Yamamoto, K. and Kumagai, H. (2001). Cloning of a gene encoding a highly stable endo- β -1,4-glucanase from *Aspergillus niger* and its expression in yeast. *Journal of Bioscience and Bioengineering* **92**, 434-441.
- Imperiali, B. and O'Connor, S. E. (1999). Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Current Opinion in Chemical Biology* **3**, 643-649.
- Kim, S. J., Lee, J. A., Kim, Y. H. and Song, B. K. (2009). Optimization of the functional expression of *Coprinus cinereus* peroxidase in *Pichia pastoris* by varying the host and promoter. *Journal of Microbiology and Biotechnology* **19**, 966-971.
- Krainer, F., Dietzsch, C., Hajek, T., Herwig, C., Spadiut, O. and Glieder, A. (2012). Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway. *Microbial Cell Factories* **11**(1), 22.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**, 283-291.
- Li, C. H., Wang, H. R. and Yan, T. R. (2012). Cloning, purification, and characterization of a heat-and alkaline-stable Endoglucanase B from *Aspergillus niger* BCRC31494. *Molecules* **17**, 9774-9789.
- Macauley-Patrick, S., Fazenda, M. L., McNeil, B. and Harvey, L. M. (2005). Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* **22**, 249-270.
- Mochizuki, S., Hamato, N., Hirose, M., Miyano, K., Otani, W., Kameyama, S., Kuwae, S., Tokuyama, T and Ohi, H. (2001). Expression and characterization of recombinant human antithrombin III in *Pichia pastoris*. *Protein Expression and Purification* **23**(1), 55-65.
- Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry* **31**, 426-428.
- Oh, S. S. L., Abu Bakar, F. D., Adnan, A. M., Mahadi, N. M., Hassan, O. and Abdul Murad, M. A. (2009). Isolation and characterization of glyceraldehydes-3-phosphate dehydrogenase gene of *Trichoderma virens* UKM1. *Biotechnology* **8**, 194-203.
- Pham, T. H., Quyen, D. T., Nghiem, N. M. and Vu, T. D. (2011). Cloning, expression, purification, and properties of an endoglucanase gene (glycosyl hydrolase family 12) from *Aspergillus niger* VTCC-F021 in *Pichia pastoris*. *Journal of Microbiology and Biotechnology* **21**(10), 1012-1020.
- Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J. T., Smith, P. L., Wierschke, J. D., Subramaniam, A. and Birkenberger, L. A. (1997). Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190**, 55-62.
- Van-Petegem, F., Vandenberghe, I., Bhat, M. K. and Van-Beeumen, J. (2002). Atomic resolution structure

- of the major endoglucanase from *Thermoascus aurantiacus*. *Biochemical and Biophysical Research Communications* **296**, 161-166.
- Vidmar, S., Turk, V. and Kregar, I. (1984).** Cellulolytic complex of *Aspergillus niger* under conditions for citric acid production. Isolation and characterization of two β -(1,4)-glucan hydrolases. *Applied Microbiology and Biotechnology* **20**, 326-330.
- Vlasenko, E., Schülein, M., Cherry, J. and Xu, F. (2009).** Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. *Bioresource Technology* **101**, 2405-2411.
- Wang, Q., Tull, D., Meinke, A., Gilkes, N. R., Warren, R. A., Aebersold, R. and Withers, S. G. (1993).** Glu280 is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-beta-1,4-glucanase. *Journal of Biological Chemistry* **268**, 14096-14102.
- Wang, T., Liu, X., Yu, Q., Zhang, X., Qu, Y., Gao, P. and Wang, T. (2005).** Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. *Biomolecular Engineering* **22**, 89-94.
- Webb, B. and Sali, A. (2014).** Protein structure modeling with MODELLER. *In: Protein Structure Prediction*. Springer, New York. pp. 1-15.
- Yennamalli, R., Rader, A., Wolt, J. and Sen, T. (2011).** Thermostability in endoglucanases is fold-specific. *BMC Structural Biology* **11**, 10.