Jurnal Teknologi, 49(F) Dis. 2008: 367–381

CLONING OF *Aspergillus Niger BglA* AND EXPRESSION OF RECOMBINANT β-GLUCOSIDASE IN METHYLOTROPHIC YEAST *Pichia Pastoris*

SHAZILAH KAMARUDDIN[1], FARAH DIBA ABU BAKAR[2], ROSLI MD ILLIAS[3], AMIR RABU[4], MAMOT SAID[5], OSMAN HASSAN[6] & ABDUL MUNIR ABDUL MURAD[7*]

Abstract. Full length cDNA of *bglA* gene encoding *Aspergillus niger* ATCC10574 β-glucosidase was isolated and sequenced. The cDNA has a length of 2583 bp which encodes a polypeptide of 860 amino acid residues with predicted pI value of 4.6 and molecular weight of 93 kDa. Amino acid analysis of BGLA from four different isolates of *A. niger*, isolates ATCC10574, ATCC1015, B1 and CBS513.88, detected a total of 29 amino acids differences. The degree of differences varies between different variants, from 0.46% up to 2.9%. Around 34% of these differences were located in β-glucosidase two conserved domains, the glycosyl hydrolase family 3 N-terminal and the C-terminal domains. Both of the domains are important for the catalytic activity of the enzyme and these differences might contribute to different biophysical and biochemical enzyme properties. Heterologous expression of BGLA in methylotrophic yeast, *Pichia pastoris* has been carried out using methanol as inducer resulting in the production of recombinant protein with molecular weight around 90 kDa. β-glucosidase activity was detected from the culture filtrate using UV-stimulated fluorescence of cleaved fluorescence substrate, 4-methylumbelliferyl-β-D-glucopyranoside (MUGlc). The specific activity of the crude recombinant enzyme for cellobiose hydrolysis was 18 U/mg.

Keywords: *Aspergillus niger*; β-glucosidase; *Pichia pastoris*; heterologous expression; glycosyl hydrolase

Abstrak. Jujukan penuh cDNA gen *bglA* yang mengkodkan enzim β-glucosidase *Aspergillus niger* ATCC10574 telah dipencil dan dijujuk. cDNA ini bersaiz 2583 pb dan mengkod 860 asid amino dengan anggaran nilai pI 4.6 dan berat molekul 93 kDa. Analisa jujukan asid amino BGLA daripada empat pencilan *A.niger* yang berbeza, iaitu pencilan ATCC1057, ATCC1015, B1 dan CBS513.88 menunjukkan terdapat sebanyak 29 perbezaan asid amino. Darjah perbezaan antara varian adalah diantara 0.46% hingga 2.9%. Sebanyak 34% daripada perbezaan tersebut terletak di dalam dua domain terpulihara β-glukosidase, iaitu domain glikosil hidrolase famili 3 terminal-N dan domain glikosil hidrolase famili 3 terminal-C. Kedua-dua domain ini penting untuk aktiviti pemangkinan enzim dan perbezaan ini berkemungkinan menghasilkan enzim yang mempunyai

[1,2,4&7]School of Bioscience and Biotechnology, Faculty Science and Technology, Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia
[5&6]Department of Bioprocess Engineering, Faculty of Chemical and Natural Resources Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor Bahru, Malaysia
[7] School of Chemistry and Food Technology, Faculty Science and Technology,Universiti Kebangsaan Malaysia, Bangi, 43600 Selangor, Malaysia
[*] Corresponding author: Tel: +60389215696, Fax: +60389252698. Email: munir@pkrisc.cc.ukm.my

sifat biofizik dan biokimia yang berbeza. Pengekspresan secara heterologus cDNA BGLA dalam sistem pengekspressan yis metilotrofik, *Pichia pastoris*, menggunakan metanol sebagai bahan pengaruh telah dijalankan dan kehadiran protein rekombinan dengan berat molekul 90 kDa telah dikenal pasti. Aktiviti β-glukosidase di dalam medium pertumbuhan dikesan dengan pemotongan substrat berpendarfluor, 4-methylumbelliferyl-β-D-glucopyranoside (MUGlc). Aktiviti spesifik enzim rekombinan kasar untuk hidrolisis selobiosa adalah 18 U/mg.

Kata kunci: *Aspergillus niger*; β-glukosidase; *pichia pastoris*; pengekspressan heterologus; glikosil hidrolase

1.0 INTRODUCTION

Cellulose, a polymer of glucose with β-1,4 linkages, is the most abundant biomass on earth. After occurrence of several oil crises, ethanol became a focus as an energy-source substitute for petroleum and the production of ethanol from cellulosic biomass has been widely explored [1]. The complete degradation of cellulose requires the synergistic action of three main cellulolytic enzymes, namely endoglucanase (EC 3.2.1.4), cellobiohydrolase (EC 3.2.1.91) and β-glucosidase (EC 3.2.1.21) [2]. Endoglucanases randomly attack the internal chain of the cellulose to produce cello-oligosaccharides while cellobiohydrolases catalyze the hydrolysis of crystalline cellulose from the ends of the cellulose chains to produce cellobiose. These sugars are further hydrolyze by β-glucosidase to produce glucose. The combination actions of these three enzymes are essential for the conversion of cellulosic materials into sugars, which can then be further converted to other chemicals such as ethanol or hydrocarbon molecules.

The role of β-glucosidase in cellulose degradation is important not only for the conversion of cellobiose and cello-oligosaccharides into glucose but also in the production of inducers for cellulases enzyme production [3]. Fowler and Brown [4] showed that deletion of *Trichoderma reesei bgl1* gene, which encodes the extracellular β-glucosidase, resulted in the decrease of endoglucanase activities and a lag in the transcription of genes such as *cbh1*, *cbh2*, *egl1*, and *egl3* that encode for cellobiohydrolase and endoglucanase. These observations suggested that β-glucosidase may be partially responsible for induction of enzymes for cellulose degradation. Subsequently *bgl2* gene from *T. reesei*, which encodes β-glucosidase, has been expressed in *Escherichia coli* and the recombinant enzyme was shown to produce sophorose from glucose via transglycosylation [5]. Sophorose was known as one of the potent inducer for cellulase enzyme system in *Trichoderma* [5].

Several fungal species such as members of the genus *Aspergillus*, *Trichoderma* and *Phanerochaete* have been extensively investigated for their β-glucosidase production. Among these fungi, *Aspergillus* species are known to be the strong producer of β-glucosidases and *Aspergillus niger* has been identified as one of the most efficient producer of β-glucosidase [6]. However, different isolates of *A. niger* was reported to express β-glucosidase of different properties, including differences in catalytic activity, thermostability and optimum pH [6]. These differences could be

due to variation in the amino acid sequences of β-glucosidase produce by different strains which may rose through evolution resulted from their adaptation to different niches. Hence, it is essential to investigate the biochemical and biophysical properties of this enzyme from different *A. niger* isolates as they might have unique features and may be useful for different industrial applications. To understand these differences and to allow protein engineering work to be performed to improve the properties of this enzyme, it is important to clone and sequence gene encoding this enzyme from different isolates and produce recombinant enzyme in an expression system that will allow high yield enzyme production but employ simple separation and purification protocols.

This study aims to clone and analyze a gene encoding for β-glucosidase of *A. niger* strain ATCC10574. In addition, we investigate the suitability of yeast, *Pichia pastoris*, to produce soluble and active recombinant β-glucosidase, using methanol as inducer. The production of recombinant protein in this methylotrophic yeast, which utilize cheap carbon source as an inducer, will provide an alternative expression system for efficient production of recombinant β-glucosidase.

2.0 MATERIALS AND METHODS

2.1 Fungi and Culture Conditions

Aspergillus niger ATCC10574 was obtained from the American Type Culture Collections (ATCC), USA. The fungus was cultured in cellulosic medium containing 1% Avicel (Fluka, Switzerland), 0.14% $(NH_4)_2SO_4$, 0.2% KH_2PO_4, 0.03% urea, 0.03% $CaCl_2.2H_2O$, 0.03% $MgSO_4.7H_20$, 0.1% peptone, 1% yeast extract, 0.1% Tween 80, and 1 mL/L trace element solution, containing 18 mM $FeSO_4.7H,O$, 6.6 mM $MnSO_4$, 4.8 mM $ZnSO_4.7H,O$, and 15 mM $COCl_2$ [7]. The medium was autoclaved and inoculated with 10^6/mL spores of *A.niger*. Plate assay to monitor β-glucosidase activity was performed using method described by Dan *et al.* [6]. Approximately 5 μL of culture filtrate were placed onto 1% agar plates containing fluorescent substrate, 0.5 mM 4-methylumbeliferyl-β-D glucopyranoside (MuGlc) (Sigma, USA). Activity of β-glucosidase enzyme was detected after 24 hours of fungal cultivation. For RNA extraction, fungal mycelia were harvested after 48 hours growth by filtration with filter paper, frozen with liquid nitrogen and stored in –80 °C until further usage.

2.2 Total RNA Extraction and cDNA Amplification

A total of 2 g of frozen mycelia were grounded to fine powder using RNAse free mortar and pestle. Total RNA was extracted using TRIzol® Reagent according to manufacturer instruction (Invitrogen, USA). The RNA was used as template for the synthesis of the first strand cDNA using Superscript III First Strand cDNA Synthesis Kit and oligo dT primer according to the manufacturer instruction (Invitrogen, USA).

Subsequently, the cDNA was used as template for PCR reaction. Specific primers were designed based on DNA sequence of *bglA* obtained from *A. niger*. Genome Database available at the Joint Genome Institute (JGI), USA (http://genome.jgi-psf.org/Aspni5/Aspni5.home.html). The sequence of forward primer was 5'-ATGAGGTTCACTTTGATCGAGGC-3' and reverse primer was 5'-TTTAGTGAACAGTAGGCAGAGACGC-3'. PCR was performed using *KOD Hot Start* DNA Polymerase Kit (Novagen, USA) and the following amplification protocol was performed: initial denaturation at 94 °C for 2 min followed by 30 cycles of 94 °C for 20s, 60 °C for 30s and 72 °C for 90s. The amplified cDNA was then cloned into pGEMT-Easy vector (Promega, USA) and sequenced using SP6 and T7 universal primers.

2.3 Construction of an Expression Vector

The plasmid containing *bglA* cDNA was used as template for the amplification of full length cDNA but without the signal peptide. Restriction enzyme sites were also added during the amplification process to assist cloning of the gene into expression vector. The following primers were designed and used to amplify the DNA fragment: Forward primer 5'- ATCGATTGATGAATTGGCCTAC-3' and reverse primer 5'-GTGATTCTAGATTGTGAACAGTAGGC-3'. The forward primer contained a *Cla*1 while the reverse primer contained *Xba*1 restriction sites. The PCR product of *bglA* cDNA was cloned into pGEMT-Easy vector (Promega, USA), digested with *Cla*1 and *Xba*1 and ligated into pPICZαC vector (Invitrogen, USA). The vector obtained was designated as BGLA- pPICZαC.

2.4 Screening for Positive *Pichia pastoris* Transformants

The BGLA-pPICZαC vector was linearized with *Bst*X1 restriction enzyme and transformed into *Pichia pastoris* strain X-33 which is a wild-type *Pichia* strain and useful for selection on Zeocin. Transformants that grow on plate containing 100 µg Zeocin were screened for Mut phenotype and putative multi-copy integrant transformants as described by manufacturer (Invitrogen, USA). Mut phenotypes were screened by growing the *P. pastoris* transformants onto minimal methanol agar plates and incubated at 30 °C for three days. This procedure was performed to screen for transformants that can utilize methanol as sole carbon source and able to express the gene of interest. Transformants that can grow well on this medium were classified as Mut^+ phenotype. Screening for multiple integrant transformants were carried out by plating transformants on YPD plate (1% yeast extract, 2% peptone, 2% dextrose and 2% agar) with increasing concentration of Zeocin, from 100 µg/mL up to 2000 µg/mL. The integration of *bglA* cDNA into the genome of *P. pastoris* was confirmed via PCR using 5'AOX1 primer: 5'- GACTGGTTCCAATTGACAAGC-3' and 3'AOX1 primer: 5'-GCAAATGGCATTCTGACATCC-3'.

2.5 Expression of *bglA* cDNA in *P. pastoris*

One of the positive *P. pastoris* transformant was selected and used to inoculate 50 mL of BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate pH6, 1.34% yeast nitrogen base without amino acids, 0.004 % biotin and 1% glycerol) in 250 mL Erlenmeyer flask. Culture was growth at 30 °C, and shake at 250 rpm until OD_{600} reached 2 – 6, in which cells were in the log-phase growth. Cells were harvested by centrifugation at 3000 × g for 5 minutes at room temperature, resuspended in BMMY medium [1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6), 1.34% yeast nitrogen base without amino acids, 0.004 % biotin and 0.5% methanol] to an OD_{600} of 1.0 and cultivated for 4 days at 30 °C, and shake at 250 rpm. Absolute methanol was added to a final concentration of 3% every 24 hours to maintain induction. The culture supernatant was collected and concentrated using Amicon Ultra Centrifugal Filter Devices (Milipore, USA). Secreted protein was analyzed using SDS-PAGE. Protein activity was screened using UV-stimulated fluorescence substrate, MuGlc (Sigma, USA). Approximately 10 μL of supernatant were plated on 1% agar plates containing 0.5 mM MuGlc. The plate was incubated at 50 °C for 1 hour and then illuminated with long UV light. An intense fluorescence was an indicator of β-glucosidase activity.

2.6 Enzyme Assay

Total protein concentration was measured using Bradford method [8]. All enzyme activities were assayed in 100 mM potassium phosphate buffer pH 6 at 30 °C. Cellobiose hydrolyzing activity was performed by monitoring the release of glucose from cellobiose. A total of 50 μg of crude recombinant enzyme was incubated with 50 mM cellobiose for 30 minutes. The concentration of glucose released was determined using dinitrosalicyclic acid (DNS) assay [9] with glucose as the standard. One unit of enzyme activity is defined as the amount of protein that produces 1 μmole of glucose per minute under the standard assay condition.

3.0 RESULTS AND DISCUSSIONS

3.1 *Bgl*1 Activity in *A.niger* Culture Medium

In this study *A. niger* isolate ATCC10574 was used. The ability of this strain to produce β-glucosidase was determined by culturing fungal mycelia in media containing Avicel as sole carbon source and testing culture filtrate for β-glucosidase activity. The presence of β-glucosidase was monitored using fluorescent substrate, 4-methylumbeliferyl-β-D glucopyranoside (MUGlc). An intense fluorescent was an indicative of β-glucosidase activity. Figure 1 shows the result of plate assay which indicate that this isolate was able to produce β-glucosidase. However, the flourescent zone produced was smaller than the zone produced by commercial β-glucosidase.

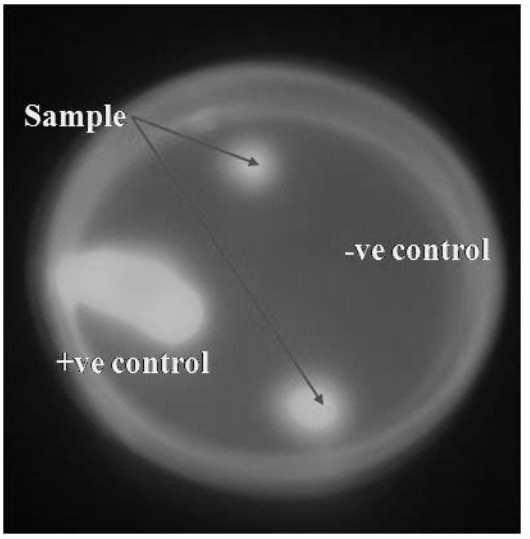


Figure 1 β-glucosidase activity of *A. niger* ATCC10574 culture medium. Total proteins were obtained from 24 hours cultivation medium. Positive control was commercial β-glucosidase, Novozyme188 (Novozyme, Denmark)

This could be due to difference in concentration of β-glucosidase presence in the culture filtrate as compared to the commercially prepared enzymes.

3.2 Cloning and Sequencing of Full Length *bgl*1 cDNA

The *bglA* cDNA of isolate ATCC10574 was cloned based on the sequence of *bglA* gene of *A. niger* strain ATCC1015 available at the Joint Genome Institute (JGI) *A. niger* Genome Database. *A. niger* strain ATCC1015 was the reference strain used in JGI *A. niger* genome sequencing project. Reverse transcriptase-polymerase chain reaction was performed to isolate the full length cDNA using RNA extracted from cells grown in Avicel. Subsequently a single PCR amplicon with the size of approximately 2.6 kb was amplified, cloned and sequenced to completion. Figure 2 shows the full *bglA* cDNA sequence and its corresponding amino acids. The cDNA has a length of 2583 bp which encodes a polypeptide of 860 amino acid. The predicted pI value of the protein is 4.6 with estimated molecular weight of 93 kDa. A conserved signal peptide sequence with the size of 19 amino acids (MRFTLIEAVALTAVSLASA) was identified at the N-terminal end of the protein (Figure 2).

Protein domain analyses indicate that the BGLA contains conserved domain which belongs to glycosyl hydrolase family 3. Amino acid alignment with selected members of family 3 glycosyl hydrolase from several fungi showed that BGLA possess one of the important conserved residues, asparagine at position 280 (Asn280) (Figure 3).

atgaggttcactttgatcgaggcggtggctctgactgccgtctcgctggccagcgctgat
MRFTLIEAVALTAVSLASAD
gaattggcctactcccctccgtattacccctcccctgggccaatggccagggtgactgg
E L A Y S P P Y Y P S P W A N G Q G D W
gcggaagcataccagcgcgctgttgatatcgtctcgcagatgacattggctgagaaggtc
A E A Y Q R A V D I V S Q M T L A E K V
aatttgactacgggaactggatgggaattggaattatgtgttggtcagactggaggtgtt
N L T T G T G W E L E L C V G Q T G G V
ccccgattgggaattccgggaatgtgtgcacaggatagccctctgggtgttcgtgactcc
P R L G I P G M C A Q D S P L G V R D S
gactacaactctgcgttccccgccggtgtcaacgtggccgcaacctgggacaagaatctg
D Y N S A F P A G V N V A A T W D K N L
gcttacctgcgtggccaggctatgggtcaggagtttagtgacaagggtgctgatatccaa
A Y L R G Q A M G Q E F S D K G A D I Q
ttgggtccagctgccggccctctcggtagaagtcccgacggcggtcgtaactgggagggc
L G P A A G P L G R S P D G G R N W E G
ttctcccccgacccggccctcagtggtgtgctctttgcagagacaatcaagggtattcag
F S P D P A L S G V L F A E T I K G I Q
gatgctggtgtggttgcaacggctaagcactacatcgcctacgagcaggagcatttccgt
D A G V V A T A K H Y I A Y E Q E H F R
caggcgcctgaagctcaaggctacggattcaatattaccgagagtggaagcgcgaacctc
Q A P E A Q G Y G F N I T E S G S A N L
gacgataagactatgcatgagctgtacctctggccccttcgcggatgccatccgtgcaggt
D D K T M H E L Y L W P F A D A I R A G
gccggtgctgtgatgtgctcgtacaaccagatcaacaacagctatggctgccagaacagc
A G A V M C S Y N Q I N N S Y G C Q N S
tacactctgaacaagctgctcaaggctgagctgggtttccagggctttgtcatgagtgat
Y T L N K L L K A E L G F Q G F V M S D
tgggcggctcaccatgccggtgtgagtggtgctttggcgggattggacatgtctatgccg
W A A H H A G V S G A L A G L D M S M P
ggagacgtcgattacgacagtggcacgtcttactggggtaccaacttgaccatcagtgtg
G D V D Y D S G T S Y W G T N L T I S V
ctcaacgggacggtgccccaatggcgtgttgatgacatggctgtccgcatcatggccgcc
L N G T V P Q W R V D D M A V R I M A A
tactacaaggtcggccgtgaccgtctgtggactcctcccaacttcagctcatggaccaga
Y Y K V G R D R L W T P P N F S S W T R
gatgaatacggcttcaagtactactatgtctcggagggaccgtatgagaaggtcaaccag
D E Y G F K Y Y Y V S E G P Y E K V N Q
ttcgtgaatgtgcaacgcaaccatagcgagttgatccgccgtattggagcagacagcacg
F V N V Q R N H S E L I R R I G A D S T
gtgctcctcaagaacgatggcgctcttcccttgactggaaaggagcgcttggtcgcccctt
V L L K N D G A L P L T G K E R L V A L
atcggagaagatgcgggttccaatccttatggtgccaacggctgcagtgaccgtgggtgc
I G E D A G S N P Y G A N G C S D R G C
gacaatgggacattggcgatgggctggggaagtggcactgccaactttccctacttggtg
D N G T L A M G W G S G T A N F P Y L V
acccccgagcaggccatctcgaacgaggtgctcaagaacaagaatggcgtattcactgcg
T P E Q A I S N E V L K N K N G V F T A

accgataactgggctattgatcagattgaggcgcttgctaagaccgccagtgtctctctt
T D N W A I D Q I E A L A K T A S V S L
gtctttgtcaacgccgactctggtgagggttatatcaatgtcgacggaaacctgggtgac
V F V N A D S G E G Y I N V D G N L G D
cgcaggaacctgaccctgtggaggaacggcgacaatgtgatcaaggctgctgctagcaac
R R N L T L W R N G D N V I K A A A S N
tgcaacaacacgatcgttattattcactctgtcggcccagtcttggttaacgagtggtac
C N N T I V I I H S V G P V L V N E W Y
gacaaccccaatgttaccgctattctctggggtggtcttcccggtcaggagtctggcaac
D N P N V T A I L W G G L P G Q E S G N
tccctcgccgacgtgctctacggccgtttcaaccccggtgccaagtcgcccttcacctgg
S L A D V L Y G R F N P G A K S P F T W
ggcaaaactcgtgaggcctaccaagattacttgtacaccgagcccaacaacggcaacgga
G K T R E A Y Q D Y L Y T E P N N G N G
gcgccccaggaagacttcgtcgagggcgtcttcattgactaccgcggatttgacaagcgc
A P Q E D F V E G V F I D Y R G F D K R
aacgagactcctatctatgagttcggctatggtccgagctacaccaccttcaactactcg
N E T P I Y E F G Y G P S Y T T F N Y S
aaccttcaggtggaggttctgagcgcccctgcgtacgagcctgcttcgggcgagactgag
N L Q V E V L S A P A Y E P A S G E T E
gcagcgccgactttcggagaggtcggaaatgcgtcggattacctctaccccgatggactg
A A P T F G E V G N A S D Y L Y P D G L
cagagaatcaccaagttcatctacccctggctcaacagtaccgatcttgaggcgtcttct
Q R I T K F I Y P W L N S T D L E A S S
ggggatgctagctatgggcaggatgcctcagactatcttcccgagggagccaccgatggc
G D A S Y G Q D A S D Y L P E G A T D G
tctgcgcaaccgatcctgcctgccggtggtggtgctggcggcaaccctcgcctgtacgac
S A Q P I L P A G G G A G G N P R L Y D
gagctcatccgcgtgaccgtgactatcaagaacaccggcaagattgcgggtgatgaagtt
E L I R V T V T I K N T G K I A G D E V
cctcaactgtatgtttctcttggcggccctaacgaacccaagatcgtgctgcgtcaattc
P Q L Y V S L G G P N E P K I V L R Q F
gagcgtatcacgctgcagccgtcggaagagacgcagtggagcacgactctgacgcgccgt
E R I T L Q P S E E T Q W S T T L T R R
gaccttgcgaactggaatgttgagacgcaggactgggagattacgtcgtatcccaagatg
D L A N W N V E T Q D W E I T S Y P K
Mgtgtttgtcggaagctcctcgcggaagctgccgctccgggcgtctctgcctactgttcac
V F V G S S S R K L P L R A S L P T V H

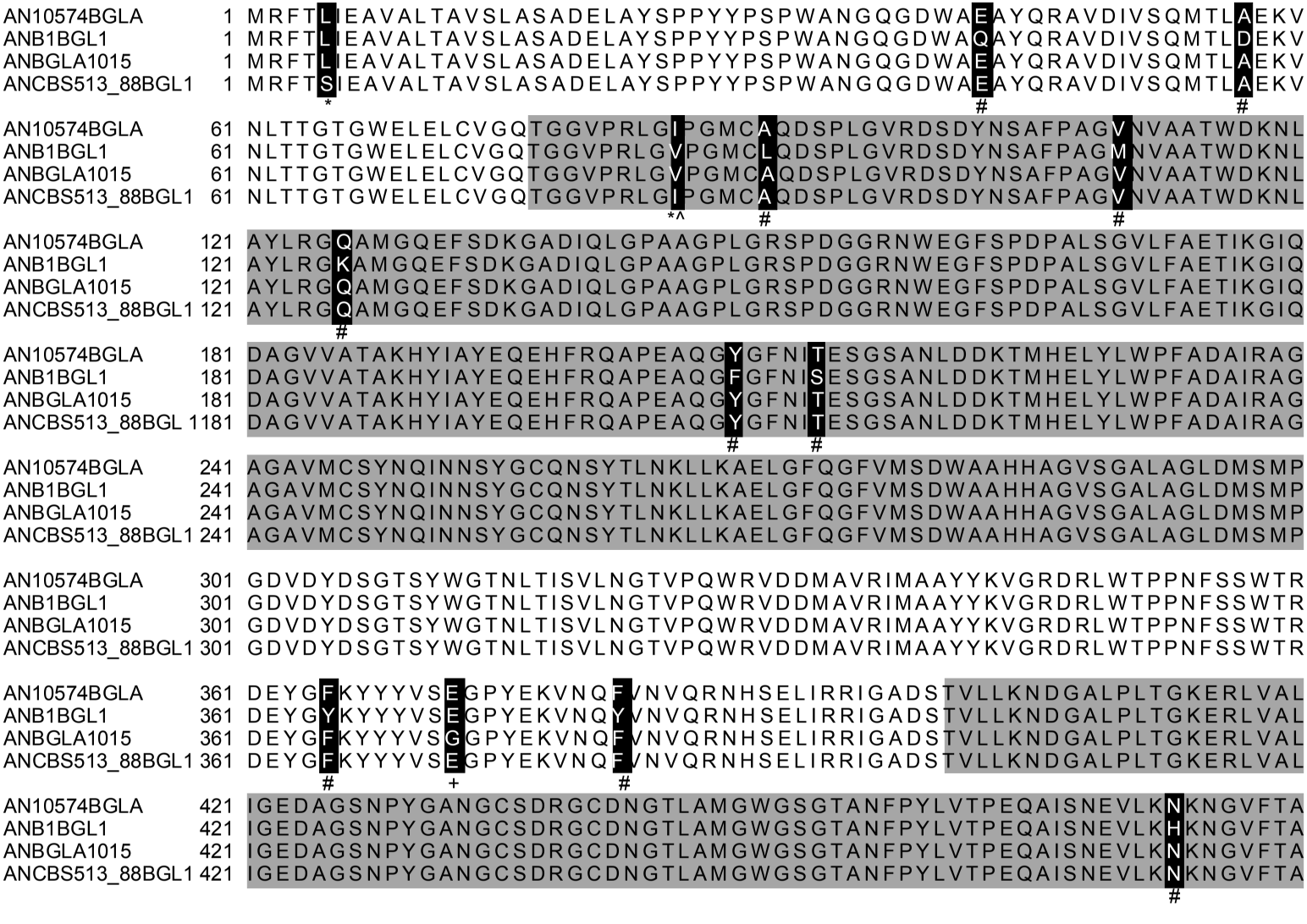
Figure 2 *bglA* cDNA and its deduced amino acid sequences. The signal peptide is indicated by underlined, bolded and *italicised*

This residue has been identified as one of the important residue for members of family 3 glycosyl hydrolase. Dan *et al.* [6] showed that this residue acted as the catalytic nucleophile within the sequence of β-glucosidase. Ly and Withers [10] demonstrated that mutation of this residue in family 3 glycosyl hydrolase resulted in total loss of enzymatic activity of the protein.

Figure 3 Alignment of *A. niger* BGLA amino acids with other glycosyl hydrolase family 3. Position of Asp-281 was indicated with asterisk. Other conserved residues were indicated in box. ASPACBGL1 represents *Aspergillus aculeatus* BGL1 (accession no. P48825); SACFBGL1 and SACFIBGL2 represent *Saccharomycopsis fibuligera* BGL1 and BGLII respectively (accession no. P22506 and P22507); CLOTHBGLB represents *Clostridium thermocellum* BGLB (accession no. P14002); HANANBGGLS represents *Hansenula anomala* BGLS (accession no. P06835); KLUMABGLS represents *Kluyveromyces marxianus* BGLS (accession no. P07337) ANBGLA represents *A. niger* ATCC10574 BGLA (this study) and ANBGL1 represents *A. niger* BI BGL1 (accession no: AJ132386)

3.3 Analysis of Amino Acids Sequence Between Different *A. niger* Isolates

To identify differences between BGLA amino acid sequence from different *A. niger* isolates, we compared our sequence to three published *A. niger* β-glucosidase sequences. Two of these sequences were generated through genome sequencing project of two reference strains, *A. niger* ATCC1015 and *A. niger* CBS513.88 [11, 12] while the third sequence was obtained from work described by Dan *et al.* using *A. niger* isolate B1 [6]. All isolates produce BGLA with the same size, 860 amino acids. However, alignment analysis detected a total of 29 amino acids differences between the four BGLA sequences with the degree of differences varies, from four amino acids (0.46%) up to 25 amino acids (2.9%) (Figure 4). Around 34% of these differences were located in two β-glucosidase conserved domains, the glycosyl hydrolase family 3 N-terminal and C-terminal domains (Figure 4). Around 34% of these differences were located in β-glucosidase two conserved domains, the glycosyl hydrolase family 3 N-terminal and C-terminal domains (Figure 4). These domains are important for the catalytic activity of the enzyme and involve in binding of the enzyme to beta-glucan [13]. These differences might generate different biophysical and biochemical properties of β-glucosidase enzyme from different isolates. Thus, cloning and expression of β-glucosidase from different isolates may result in discovery of β-glucosidase enzyme with high activity and improved properties.


AN10574BGLA 1 MRFTLIEAVALTAVSLASADELAYSPPYYPSPWANGQGDWAEAYQRAVDIVSQMTLAEKV
ANB1BGL1 1 MRFTLIEAVALTAVSLASADELAYSPPYYPSPWANGQGDWAQAYQRAVDIVSQMTLDEKV
ANBGLA1015 1 MRFTLIEAVALTAVSLASADELAYSPPYYPSPWANGQGDWAEAYQRAVDIVSQMTLAEKV
ANCBS513_88BGL1 1 MRFTSIEAVALTAVSLASADELAYSPPYYPSPWANGQGDWAEAYQRAVDIVSQMTLAEKV
AN10574BGLA 61 NLTTGTGWELELCVGQTGGVPRLGIPGMCAQDSPLGVRDSDYNSAFPAGVNVAATWDKNL
ANB1BGL1 61 NLTTGTGWELELCVGQTGGVPRLGVPGMCLQDSPLGVRDSDYNSAFPAGMNVAATWDKNL
ANBGLA1015 61 NLTTGTGWELELCVGQTGGVPRLGVPGMCAQDSPLGVRDSDYNSAFPAGVNVAATWDKNL
ANCBS513_88BGL1 61 NLTTGTGWELELCVGQTGGVPRLGIPGMCAQDSPLGVRDSDYNSAFPAGVNVAATWDKNL
AN10574BGLA 121 AYLRGQAMGQEFSDKGADIQLGPAAGPLGRSPDGGRNWEGFSPDPALSGVLFAETIKGIQ
ANB1BGL1 121 AYLRGKAMGQEFSDKGADIQLGPAAGPLGRSPDGGRNWEGFSPDPALSGVLFAETIKGIQ
ANBGLA1015 121 AYLRGQAMGQEFSDKGADIQLGPAAGPLGRSPDGGRNWEGFSPDPALSGVLFAETIKGIQ
ANCBS513_88BGL1 121 AYLRGQAMGQEFSDKGADIQLGPAAGPLGRSPDGGRNWEGFSPDPALSGVLFAETIKGIQ
AN10574BGLA 181 DAGVVATAKHYIAYEQEHFRQAPEAQGYGFNITESGSANLDDKTMHELYLWPFADAIRAG
ANB1BGL1 181 DAGVVATAKHYIAYEQEHFRQAPEAQGFGFNISESGSANLDDKTMHELYLWPFADAIRAG
ANBGLA1015 181 DAGVVATAKHYIAYEQEHFRQAPEAQGYGFNITESGSANLDDKTMHELYLWPFADAIRAG
ANCBS513_88BGL 1181 DAGVVATAKHYIAYEQEHFRQAPEAQGYGFNITESGSANLDDKTMHELYLWPFADAIRAG
AN10574BGLA 241 AGAVMCSYNQINNSYGCQNSYTLNKLLKAELGFQGFVMSDWAAHHAGVSGALAGLDMSMP
ANB1BGL1 241 AGAVMCSYNQINNSYGCQNSYTLNKLLKAELGFQGFVMSDWAAHHAGVSGALAGLDMSMP
ANBGLA1015 241 AGAVMCSYNQINNSYGCQNSYTLNKLLKAELGFQGFVMSDWAAHHAGVSGALAGLDMSMP
ANCBS513_88BGL1 241 AGAVMCSYNQINNSYGCQNSYTLNKLLKAELGFQGFVMSDWAAHHAGVSGALAGLDMSMP
AN10574BGLA 301 GDVDYDSGTSYWGTNLTISVLNGTVPQWRVDDMAVRIMAAYYKVGRDRLWTPPNFSSWTR
ANB1BGL1 301 GDVDYDSGTSYWGTNLTISVLNGTVPQWRVDDMAVRIMAAYYKVGRDRLWTPPNFSSWTR
ANBGLA1015 301 GDVDYDSGTSYWGTNLTISVLNGTVPQWRVDDMAVRIMAAYYKVGRDRLWTPPNFSSWTR
ANCBS513_88BGL1 301 GDVDYDSGTSYWGTNLTISVLNGTVPQWRVDDMAVRIMAAYYKVGRDRLWTPPNFSSWTR
AN10574BGLA 361 DEYGFKYYYVSEGPYEKVNQFVNVQRNHSELIRRIGADSTVLLKNDGALPLTGKERLVAL
ANB1BGL1 361 DEYGYKYYYVSEGPYEKVNQYVNVQRNHSELIRRIGADSTVLLKNDGALPLTGKERLVAL
ANBGLA1015 361 DEYGFKYYYVSGGPYEKVNQFVNVQRNHSELIRRIGADSTVLLKNDGALPLTGKERLVAL
ANCBS513_88BGL1 361 DEYGFKYYYVSEGPYEKVNQFVNVQRNHSELIRRIGADSTVLLKNDGALPLTGKERLVAL
AN10574BGLA 421 IGEDAGSNPYGANGCSDRGCDNGTLAMGWGSGTANFPYLVTPEQAISNEVLKNKNGVFTA
ANB1BGL1 421 IGEDAGSNPYGANGCSDRGCDNGTLAMGWGSGTANFPYLVTPEQAISNEVLKHKNGVFTA
ANBGLA1015 421 IGEDAGSNPYGANGCSDRGCDNGTLAMGWGSGTANFPYLVTPEQAISNEVLKNKNGVFTA
ANCBS513_88BGL1 421 IGEDAGSNPYGANGCSDRGCDNGTLAMGWGSGTANFPYLVTPEQAISNEVLKNKNGVFTA

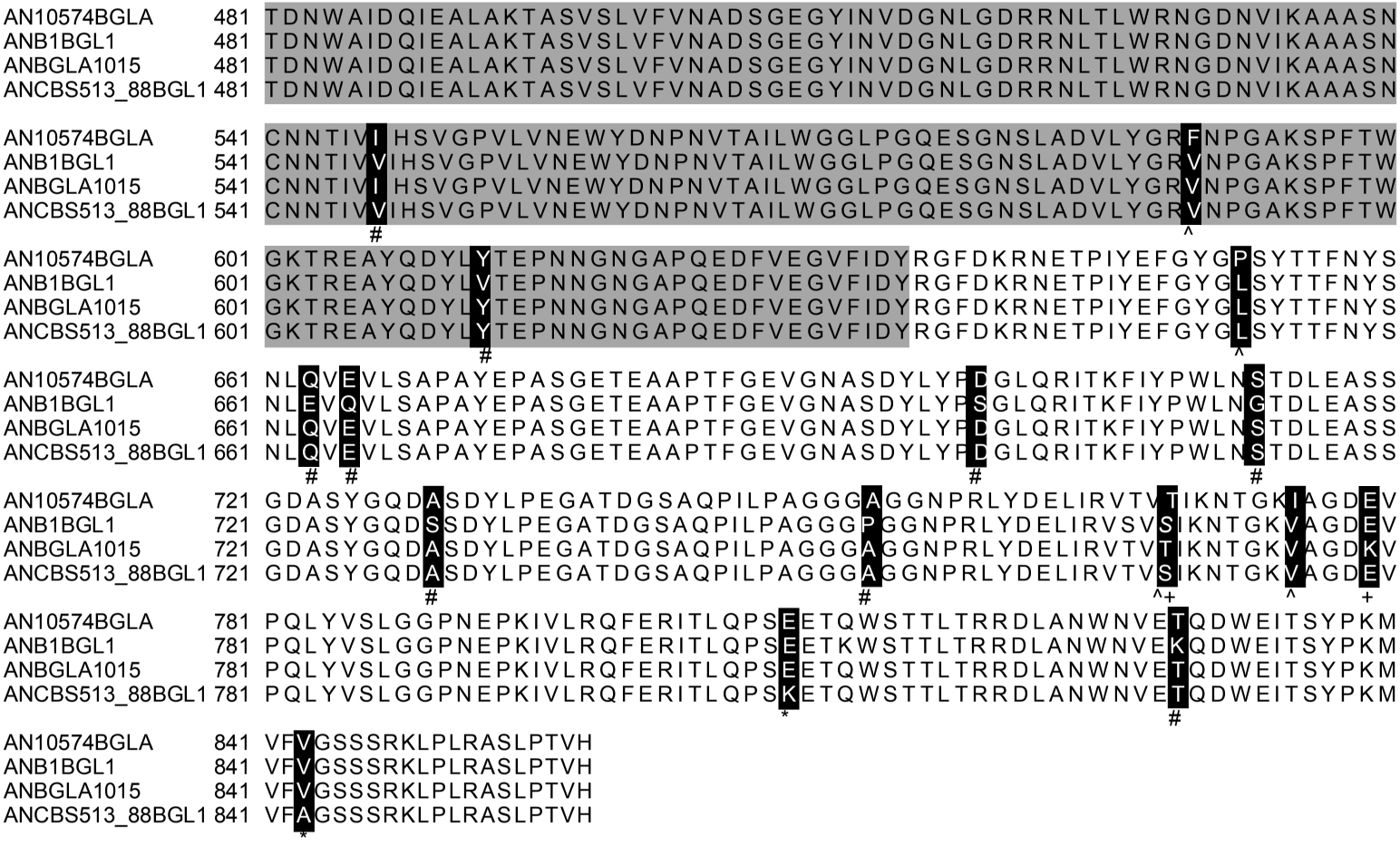


Figure 4 Amino acid alignment between BGLA from different *A. niger* isolates. AnCBS513.88 refer to *A. niger* strain CBS513.88 [11], ANBGLA1015 refer to *A. niger* strain ATCC1015 [12] ANATCC10574 refer to *A. niger* strain ATCC10574 (this study) and ANB1BGL1 refer to *A. niger* strain B1[6]. Symbol * indicates different amino acid detected only on CBS513.88. Symbol # indicates different amino acid detected only on isolate B1. Symbol + indicates different amino acids detected only on ATCC1015. Symbol ^ indicates different amino acid detected only on ATCC10574. Symbol ^* indicates the same amino acid that detected both on ATCC10574 and CBS513.88 but differ from other isolates. Symbol ^+ indicates the same amino acid detected both on ATCC10574 and ATCC1015 but differ from other isolates

3.4 Construction of Expression Vector and Screening for Positive Transformants

The pPICZαC expression vector carrying cDNA of *bglA* downstream of the promoter of alcohol oxidase gene (*AOX1*) and α-factor signal peptide has been constructed and designated as BGLA- pPICZαC. The expression vector also contains polyhistidine tag at the C-terminal of the expression cassette, which will assist in purification of the recombinant protein later. The linearized expression vector was transformed into the genome of *P. pastoris* strain X-33 and plated on YPD plate containing 100 μg/mL Zeocin. Approximately 50 colonies were observed to grow on the YPD plate after 2 days. All transformants were selected and screened further on YPD plate containing higher concentration of Zeocin (up to 2000 μg/mL). This procedure should select transformants carried multi-copy integrants of the expression cassette in their genome. A total of 20 colonies were able to grow on plate containing up to 1500 μg/mL and no colonies was detected on plate containing 2000 μg/mL Zeocin. Subsequently ten colonies were selected for PCR screening using 5'AOX1 and 3'AOX1 primers to amplify the full length *bglA* cDNA which has been integrated into *P.pastoris* genome. Figure 5 shows eight of the ten colonies were detected to have *bglA* cDNA integrated into their genome. This was represented by the 3.1 kb amplicon that was successfully

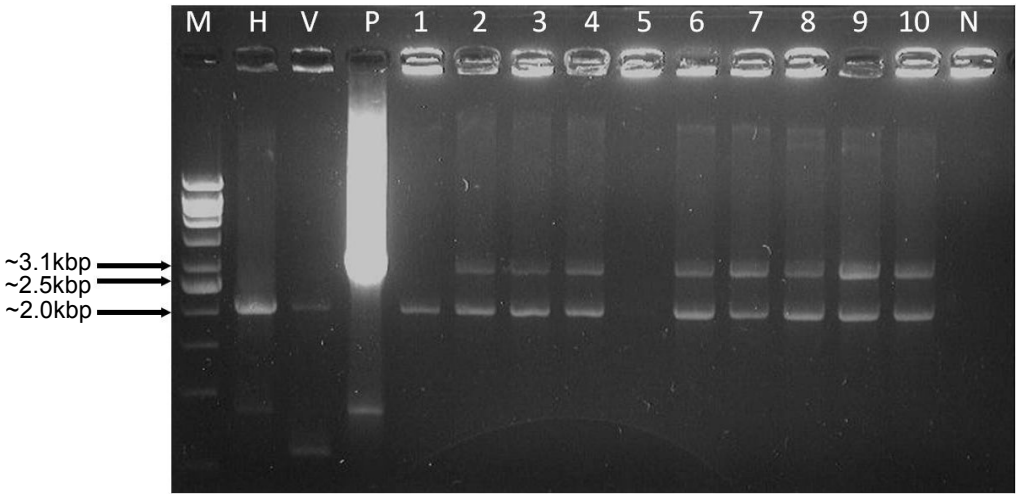


Figure 5 PCR amplification to screen for *P. pastoris* transformants carried *bglA* cDNA in their genome. Lane M: 1 kb DNA Ladder (Vivantis, Malaysia); H: untransformed *P. pastoris*; V: *P. pastoris* transformed with pPICZαC alone; P: pPICZαC vector containing *bglA* cDNA in DH5α as positive control; 1-10: *P. pastoris* transformants and N: negative control. Positive transformants that contained *bglA* cDNA integrated into their genome produced two PCR amplicons with the size of 2.2 kb and 3.1 kb. The 2.2 kb amplicon represents the amplified AOX1 gene in *P. pastoris* genome while the 3.1 kb amplicon represents *bglA* cDNA integrated downstream of *AOX1* promoter in *P. pastoris* genome

amplified using the 5'AOX1 and 3'AOX1 primers. One of the clones was then selected for the expression of recombinant β-glucosidase.

3.5 Expression of rBGLA in *P. pastoris* and Enzyme Activity Assay

Recombinant beta-glucosidase was successfully expressed in *P. pastoris* using 3% final concentration of methanol as an inducer. Methanol was used as inducer because *bglA* cDNA was cloned downstream of alcohol oxidase (*AOX1*) promoter in pPICZαC vector. In *P. pastoris*, the *AOX1* promoter is tightly regulated and induced by the presence of methanol in the growth media. The culture filtrate was collected on day four after inductin with methanol, concentrated and checked for the presence of recombinant protein via SDS-PAGE analysis (Figure 6). Coomassie blue staining of the gel showed the presence of recombinant protein with molecular weight around 90 kDa in the culture filtrate of *P. pastoris* transformant that had been induced with methanol (Figure 6). The protein size agrees with the predicted molecular weight from the deduced amino acid sequence. The β-glucosidase activity was detected by UV-stimulated fluorescence of cleaved 4-methylumbelliferyl-β-D-glucopyranoside (Figure 7). An intense fluorescence was detected and this result suggested that the recombinant enzyme was successfully expressed and secreted out by the yeast. Enzyme assay of the crude recombinant β-glucosidase showed 18 U/mg of specific activity towards cellobiose degradation. This indicates that an active recombinant β-glucosidase was successfully expressed extracellularly in *P. pastoris* culture medium.

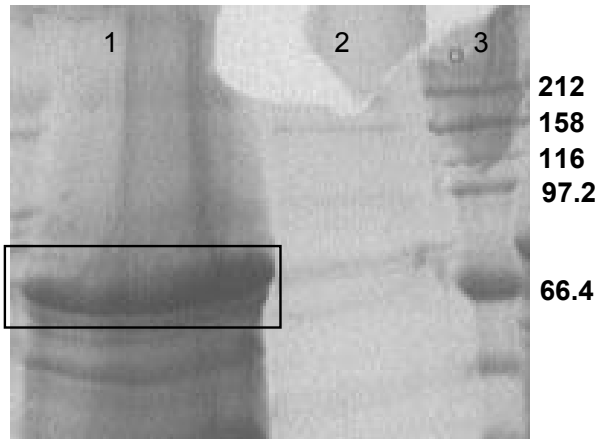


Figure 6 SDS-PAGE of recombinant β-glucosidase on 10% polyacrylamide gel. Lane 1: protein from *P. pastoris* transformant induced with 3% methanol; lane 2: protein from *P. pastoris* transformants uninduced with methanol; lane 3: Protein Marker Broad Range (NEB, UK)

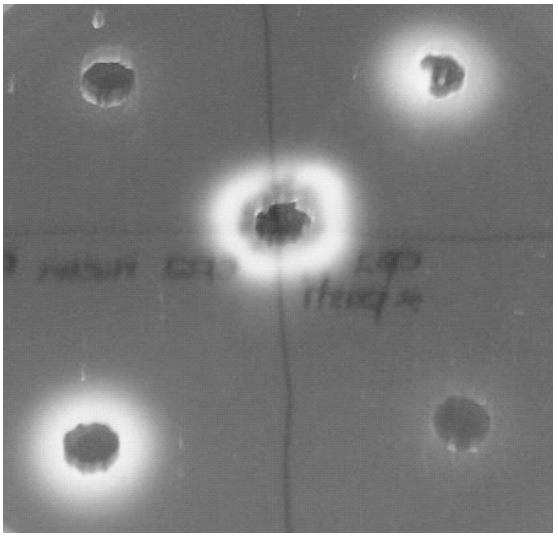


Figure 7 Activity screening of rBGLA secreted out into the growth medium. Well 1&3: rBGLA; well 2: commercial β-glucosidase, Novozyme 188 (Novozymes, Denmark); well 4 : negative control

4.0 CONCLUSION

In this work we cloned *A. niger* ATCC10574 *bglA* cDNA and analyzed its corresponding amino acids. We showed that variation between amino acids sequence of BGLA between different *A. niger* isolates existed and this difference varies from one isolate to another. The differences in the amino acid sequence could be one of the factors that contribute towards the difference in the properties of the enzyme. We also expressed an active recombinant β-glucosidase enzyme using *P. pastoris* expression system. The system developed should be able to assist our future work on the characterization of the enzyme from different isolates and for enzyme engineering work to improve the properties of the enzyme.

ACKNOWLEDGEMENT

The authors would like to thank the financial support from the National Biotechnology Directorate, MOSTI for the financial support through grant no. 07-05-16-MGI-GMB12. Shazilah Kamaruddin is supported by National Science Fellowship, MOSTI.

REFERENCES

[1] Zhang, Y. H. P. 2008. Reviving the Carbohydrate Economy via Multi-product Lignocellulose Biorefineries. *Journal of Industrial Microbiology and Biotechnology*. 35: 367-375.

[2] Lynd, L. R., P. J. Weimer, W. H. van Zyl and I. S. Pretorius. 2002. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*. 66(3): 506-577.

[3] Penshin, A. and J. M. S. Mathur. 1999. Purification and Characterization of β-glucosidase from *Aspergillus niger* Strain 322. *Letters in Applied Microbiology.* 28: 401-404.

[4] Fowler, T. and R. D. Brown. 1992. The *bgl1* Gene Encoding Extracellular β-glucosidase from *Trichoderma reesei* is Required for Rapid Induction of the Cellulase Complex. *Molecular Microbiology*. 6: 3225-3235.

[5] Saloheimo, M., J. Kuja-Panula, E. Ylosmaki, M. Ward and M. Penttila. 2002. Enzymatic Properties and Intracellular Localization of the Novel *Trichoderma Reesei* β-glucosidase BGLII (cel1A). *Applied and Environmental Microbiology*. 68: 4546-4553.

[6] Dan, S., I. Marton, M. Dekel, B. A. Bravdo, S. He, G. S. Withers and O. Shoseyov. 2000. Cloning, Expression, Characterization and Nucleophile Identification of Family 3, *Aspergillus niger* β-glucosidase. *Journal of Biological Chemistry.* 275: 4973-4980.

[7] Hong, J., H. Tamaki, S. Akiba, K. Yamamoto and H. Kumagai. 2001. Cloning of a Gene Encoding a Highly Stable Endo-b-1,4-glucanase from *Aspergillus niger* and its Expression in Yeast. *Journal of Bioscience and Bioengineering*. 92: 434-441.

[8] Bradford, M. M**.** 1976. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein using the Principle of Protein-dye Binding. *Analytical Biochemistry*. 72: 248-252.

[9] Miller, G. L. 1959. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Analytical Chemistry*. 31: 426-428.

[10] Ly, H. and S. G. Withers. 1999. Mutagenesis of Glycosidases. *Annual Review of Biochemistry*. 68: 487-522.

[11] Pel, H. J., J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann, P. J. Schaap, G. Turner, R. P. de Vries, R. Albang and K. Albermann. *et al*. 2007. Genome Sequencing and Analysis of the Versatile Cell Factory *Aspergillus niger* CBS 513.88. *Nature Biotechnology*. 25(2): 221-231.

[12] Baker, S. E. 2006. *Aspergillus niger* Genomics: Past, Present and into the Future. *Medical Mycology*. 44: S17-S21

[13] Varghese, J. N., Hrmova, M. and G. B., Fincher. 1999. Three-dimensional Structure of a Barley Beta-D-glucan Exohydrolase, a Family 3 Glycosyl Hydrolase. *Structure*. 7(2): 179-90.