
# Purification and characterization of a novel GH1 beta-glucosidase from Jeotgalibacillus malaysiensis

Kok Jun Liew, Lily Lim, Hui Ying Woo, Kok-Gan Chan, Mohd Shahir Shamsir, Kian Mau Goh

PII: S0141-8130(18)30602-0
DOI: doi:10.1016/j.ijbiomac.2018.04.156
Reference: BIOMAC 9564

To appear in:

Received date: 6 February 2018
Revised date: 24 April 2018
Accepted date: 28 April 2018

Please cite this article as: Kok Jun Liew, Lily Lim, Hui Ying Woo, Kok-Gan Chan, Mohd Shahir Shamsir, Kian Mau Goh , Purification and characterization of a novel GH1 beta-glucosidase from Jeotgalibacillus malaysiensis. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Biomac(2017), doi:10.1016/j.ijbiomac.2018.04.156

# Purification and characterization of a novel GH1 beta-glucosidase from *Jeotgalibacillus malaysiensis*

Kok Jun Liew [a], Lily Lim[a], Hui Ying Woo[a], Kok-Gan Chan [b,c], Mohd Shahir Shamsir [a], Kian Mau Goh [a,*]

[a] Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
[b] Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Vice Chancellor Office, Jiangsu University, Zhenjiang 212013, PR China

* Corresponding author.
*Email address:* gohkianmau@utm.my (K.M. Goh)

**Abstract**

Beta-glucosidase (BGL) is an important industrial enzyme for food, waste and biofuel processing. *Jeotgalibacillus* is an understudied halophilic genus, and no beta-glucosidase from this genus has been reported. A novel beta-glucosidase gene (1344 bp) from *J. malaysiensis* DSM 28777$^T$ was cloned and expressed in *E. coli*. The recombinant protein, referred to as BglD5, consists of a total 447 amino acids. BglD5 purified using a Ni-NTA column has an apparent molecular mass of 52 kDa. It achieved the highest activity at pH 7 and 65°C. The activity and stability were increased when $CaCl_2$ was supplemented to the enzyme. The enzyme efficiently hydrolyzed salicin and (1→4)-beta-glycosidic linkages such as in cellobiose, cellotriose, cellotetraose, cellopentose, and cellohexanose. Similar to many BGLs, BglD5 was not active towards polysaccharides such as Avicel, carboxymethyl cellulose, Sigmacell cellulose 101, alpha-cellulose and xylan. When BglD5 blended with Cellic® Ctec2, the total sugars saccharified from oil palm empty fruit bunches (OPEFB) was enhanced by 4.5%. Based on sequence signatures and tree analyses, BglD5 belongs to the Glycoside Hydrolase family 1. This enzyme is a novel beta-glucosidase attributable to its relatively low sequence similarity with currently known beta-glucosidases, where the closest characterized enzyme is the DT-Bgl from *Anoxybacillus* sp. DT3-1.

## 1. Introduction

Lignocellulose feedstock such as agricultural and forest residues are rich in carbohydrates, and could thus serve as potential raw materials for large-scale bioethanol production [1]. Cellulases, i.e exoglucanase, endoglucanase and beta-glucosidase are glycoside hydrolase (GH) enzymes that depolymerize cellulose into fermentable glucose for value-added bioethanol [2]. Among the aforementioned list, beta-glucosidase (EC 3.2.1.21, synonym gentiobiase; cellobiase; β-D-glucosidase) plays a key role in the biomass deconstruction pipeline, as it reduces cellobiose (two glucose units) or short chain cello-oligosaccharides to glucose.

A total of 153 GH families are currently listed on the Carbohydrate Active enzyme (CAZy) website (http://www.cazy.org). Beta-glucosidases (BGLs) are categorized in families GH1, GH3, GH5, GH9, GH30, and GH116. A notable exception is that GH9 enzymes adopt an inverting mechanism, while BGLs in other families use the retaining mechanism [3, 4]. Besides, BGLs are usually classified into three groups according to their substrate specificity: (i) aryl-beta-glucosidases, which acts on aryl beta-glucosides; (ii) cellobiases, which hydrolyze oligosaccharides only; and (iii) wide range beta-glucosidases, which exhibit broad activities on various types of substrates [5]. BGLs are important enzymes in various biotechnological processes, including wine, food, and saccharification of lignocellulosic materials for biofuels [6-8]. A recent work suggests that BGLs from *Bacillus licheniformis* could potentially be used to degrade polysaccharide bioflocculant [9]. Detail reviews regarding beta-glucosidase are available elsewhere [10-14].

*Jeotgalibacillus* is a mesophilic and halophilic genus. *J. malaysiensis* sp. nov. (strain D5$^T$=DSM 28777$^T$=KCTC33550$^T$) was isolated from a sandy beach in Johor state, Malaysia [15]. The genome and transcriptome analyses of this bacterium were performed earlier by our team [16]. A total of 25 GHs and six carbohydrate esterase genes were annotated in the genome of this strain. *J. malaysiensis* does not harbor any gene for GH3, GH5, GH9, GH30, and GH116 BGL. However, a GH1 gene was present in its genome. Beta-glucosidase from this genus has not been characterized, and we hypothesize that this enzyme has important implications towards the biotechnological industry. This report describes the biochemical properties of this unexplored BGL.

## 2. Methods

### *2.1 Gene cloning of beta-glucosidase BglD5*

The stock culture of *J. malaysiensis* [15] was cultured on Marine Agar (Pronadisa Laboratorios Conda, Spain) at 50°C for two days. The genomic DNA was isolated using a Wizard Genomic DNA Purification Kit (Promega, USA) according to the manufacturer's instructions. Specific primers for *bglD5* gene amplification were designed according to locus-tag JMA_34740 annotated in a completed genome sequencing project (DDBJ/EMBL/GenBank: CP009416) [17]. Gene amplification was performed using KAPA HiFi HotStart ReadyMix (KAPA Biosystems, USA) with forward primer (5'- GCCGGATCCATGAGAAAGTTTCCTGAACATTTC G -3') and reverse primer (5'- GCGGCTCGAGAG TATATTGTTGTTGAATCTTTAACC C -3'), and the underlined sequences were restriction sites for *BamH*I and *Xho*I, respectively. The gene was cloned into pET28a and maintained in *E. coli* BL21 (DE3).

*2.2 Expression and purification of BglD5*

The recombinant *E. coli* BL21 (DE3) with pET28a-*bglD5* gene was cultured overnight on Luria-Bertani (LB) medium, supplemented with 50 μg/mL kanamycin at 37°C with 200rpm shaking. Next, two milliliters of the overnight culture were inoculated into 200 mL of fresh LB/Kanamycin broth, and the fresh culture was grown in the same shaking incubator. When the culture's $OD_{600nm}$ reached between 0.5~0.6, protein expression was induced by the addition of 1.0 mM isopropyl-β-D-thiogalactopyranoside into the culture, and growth continued for another 3 hours at 37°C with 200rpm shaking. The cultures were then harvested by centrifugation at 5000 × g for 10 minutes. The crude enzyme was extracted from the harvested cell using a B-PER Direct Bacterial Protein Extraction Kit (Thermo Scientific, USA) according to the supplier's protocol. The crude enzyme was then subjected to purification, which was conducted on a ÄKTA Start chromatography system (GE Healthcare, Sweden) connected with a Qiagen Ni-NTA Superflow column. The column resin was equilibrated with a binding buffer (20 mM sodium phosphate buffer, pH 7.4, containing 500 mM NaCl and 60 mM imidazole), and the bound proteins were eluted with the elution buffer (20 mM sodium phosphate buffer, pH 7.4, containing 500 mM NaCl and 350 mM imidazole). The purified BglD5 was dialyzed against 100 mM sodium phosphate buffer (pH 7.4) at 4°C overnight, prior to further analyses.

*2.3 Enzyme assay*

The enzyme activity of purified BglD5 was determined using substrate p-nitrophenyl-beta -D-glucopyranoside (pNPG). The purified BglD5 (0.05 mL) was mixed with 5 mM of pNPG (0.5 mL) in a 100 mM sodium phosphate buffer (pH 7.4). After being incubated at 65°C for 15 minutes, the enzyme reaction was stopped by adding 0.5 mL 1M sodium hydroxide (NaOH). The absorbance was detected by using a spectrophotometer at 405 nm. One unit (U) of enzyme activity is defined as the amount of enzyme that catalyzed the formation of 1 μmol pNP per minute under the conditions of the assay. Enzyme assays were performed in at least triplicates for each analysis.

*2.4 Characterization of BglD5*

*2.4.1 Effects of temperature and pH on BglD5 activity*

The optimum temperature for BglD5 was determined by incubating the enzyme reactions at various temperatures ranging from 20-90°C for 15 minutes. The effects of pH on purified BglD5 was determined by conducting the reactions in various buffers, with pH ranging from 2-11. The buffers employed were glycine-HCl buffer (pH 2-3), sodium acetate buffer (pH 4-5.5), sodium phosphate buffer (pH 6-7.5), Tris-HCl buffer (pH 8-9) and carbonate-bicarbonate buffer (pH 10-11). The relative activity of each of the reactions was calculated by using enzyme activity at the optimum temperature or pH as reference. Thermostability of BglD5 (with and without 5 mM $CaCl_2$) was conducted by incubating the purified enzyme at 65°C, without the addition of pNPG. At a specific time interval, the incubated enzyme was withdrawn and subjected to enzyme reaction, and the residual activity of the incubated enzyme was measured by using an enzyme activity at time zero as reference.

*2.4.2 Effects of glucose on BglD5 activity*

The purified BglD5 was subjected to glucose inhibition evaluation by reacting pNPG substrates in the presence of glucose (concentration ranging from 0 mM to 2500 mM). The relative activity of each reaction was calculated against the reaction tube with the highest activity.

*2.4.3 Effects of metal ions and chemical reagents on BglD5 activity*

The enzyme activity of BglD5 was evaluated by conducting the standard enzyme reactions with supplementation of 5 mM calcium chloride, sodium chloride, potassium chloride, magnesium chloride, iron (III) chloride, nickel (II) chloride, cobalt (II) chloride, ammonium chloride, zinc sulfate, manganese sulfate, copper (II) sulfate, rubidium chloride, strondium chloride, and barium chloride. Besides, enzymatic reactions were also tested with 5% (*v/v*) urea, sodium dodecyl sulfate (SDS), Ethylenediaminetetraacetic acid (EDTA), Tween-20, Tween-80, Triton X-100, and dimethyl sulfoxide (DMSO). The residual activity of BglD5 in each reaction tube was calculated with using enzyme reactions without any salts or chemical reagents as reference.

*2.4.4 Enzyme kinetic*

The enzyme kinetic parameters of the purified BglD5, for instance, Michaelis-Menten constant ($K_m$), maximum velocity ($V_{max}$), and turnover number ($k_{cat}$) were assessed by measuring the hydrolysis rate of pNPG at various concentrations (0.05-3.0 mM) at 65°C for 15min in 100 mM sodium phosphate buffer (pH 7). The data obtained was used for Michaelis-Menten plotting.

*2.4.5 HPLC analyses*

Substrate specificity of BglD5 was determined by incubating 100 µL purified enzyme (0.2 U) in various 1% (w/v) substrates, for instance, cellodextrins (cellobiose – cellohexaose), sucrose, salicin, maltose, lactose, gentibiose, alpha-cellulose, xylan, Avicel, Sigmacell cellulose and carboxymethyl cellulose. All the reactions were conducted at 65°C, and pH 7 for 24 hours, and the reactions were stopped by boiling for 5 min. Upon stopping the enzyme reaction, the samples were analyzed in an Agilent 1260 Infinity high performance liquid chromatography (HPLC), coupled with an Agilent 385-Evaporative Light Scattering Detector (Agilent Technologies, USA) with a Rezex RPM-Monosaccharide $Pb^{2+}$ column (Phenomenex Inc, USA) connected. Unless specified, enzyme activity was measured by quantifying the amount of glucose released, and the relative hydrolysis rate of each reaction was calculated by using cellobiose reaction tube as reference.

In an independent experiment, the hydrolysis of oil palm empty fruit bunches (OPEFB) was carried out for one hour at 65°C and pH 7. The raw OPEFB biomass was obtained from a local oil palm mill, and then washed, dried, and cut to the size of about 0.1 mm ~ 5.0 mm. It was then autoclaved prior to being used. The following enzymatic hydrolyses were setup in which 1% (v/v) of OPEFB in a total volume of 5 mL was reacted with (i) 875 U BglD5; (ii) 250 µL Cellic® Ctec2 (Novozymes, Denmark); and (iii) 875 U BglD5+ 250 µL Cellic® Ctec2. The total sugar production was measured using the HPLC.

*2.4.6 Bioinformatic analyses of BglD5*

Sequences of beta-glucosidases from 70 different species were retrieved using the Uniprot database. Phylogenetic relationships of BglD5 with those BGLs sequences were inferred using the Maximum-Parsimony (MP) methods with the help of Molecular Evolutionary Genetic Analysis (MEGA 6) program [18]. The SPRINT database that collected five motifs of BGLs was used as reference to determine the location of motifs in BglD5 [19]. A Weblogo representation was generated using an online generator [20]. The crystal structures of proteins were obtained from PDB database.

## 3. Results

*3.1 Characterization of purified BglD5*

Our team sequenced and discussed the genomes of a several type strains of *Jeotgalibacillus* [16, 17, 21, 22]. In the genome of *J. malaysiensis*, we identified the presence of a gene (1344 bp) that encodes a beta-glucosidase (BglD5) with 447 amino acids (Uniprot ID: A0A0B5ARU7). This gene was cloned and expressed in *E. coli*. In order to eliminate the effect of other factors which may be presented in the crude enzyme preparation, and may thus interfere with the BglD5 function, the purification of the recombinant protein was carried out by a single step immobilized metal affinity chromatography (IMAC) using Ni-NTA, according to the method previously reported [23]. According to Fig. 1, the SDS-PAGE analysis results indicate the high purity of BglD5 with the molecular weight of approximately 52kDa (calculated based on protein sequence).

**Fig 1.** SDS-PAGE analysis of purified BglD5. Lane M: Benchmark™ Protein Ladder (Thermo Scientific, USA); lane 1: cell lysate from recombinant *E. coli* BL21 (DE3); lane 2: flowthrough (uncaptured) fraction from purification column; lane 3: purified BglD5. A 5 µL of each sample was loaded into each well and the gel was stained by Imperial™ Protein Stain (Thermo Scientific, USA).

The enzymatic and physiological characterization of the recombinant enzyme elucidated that recombinant BglD5 demonstrates optimal activity at a temperature of 65°C (Fig. 2a) and maintains approximately 95% of the maximal activity at 70°C. BglD5 also shows high activity at pH in the range of 6 ~ 7.5 (85 ~ 100% of the maximal activity), with the highest at pH 7 (Fig. 2b). Several buffers (pH 7.0) were tested to examine the buffer that was better for BglD5 activity. The optimal buffer was the sodium phosphate buffer, but the activity in the potassium phosphate buffer was similarly good. BglD5 also worked well in citrate-phosphate, HEPES and MOPS buffers, with an approximately 8% lower activity compared to that of the sodium phosphate buffer. Retardation in the activity of BglD5 in the Tris-HCl buffer is likely due to the inhibition effect caused by Tris that could bind to the active site cavity [24-26]. Fig 2c showed that BglD5 was activated by the addition of glucose at low concentration, the enzyme activity gradually increased from 0 mM to 250 mM of glucose. Yet, it was able to tolerate glucose concentration up to 2500 mM, retaining about 80% of its maximal activity. Using pNPG as a substrate and a reaction conducted at optimum temperature and pH, BglD5 indicated the $K_m$ value of 0.50 ± 0.02 mM, a $V_{max}$ of 39.48 ± 0.63 U/mg, and a $k_{cat}$ value of 33.93 ± 0.54$s^{-1}$, using a Michaelis-Menten function (Fig. 2e). A comparison of biochemical properties and kinetic performance of BglD5 with some other BGLs is shown in Table 1.

Table 2 summarizes the overall effects of metal ions and detergents on the activity of BglD5. With the supplementation of 5 mM calcium chloride, the relative activity of BglD5 increased to 121%. Besides, calcium chloride has a positive effect on thermostability (Fig. 2d). A slight increment of activity was also observed when magnesium chloride was introduced into the reaction (Table 2), while iron chloride, nickel chloride, cobalt chloride, zinc sulfate, manganese sulfate, and copper sulfate inhibited BglD5 activity.

The substrate specificity of BglD5 was determined by reacting the enzyme with various substrates (10 mg/mL each) (Table 3). The enzyme efficiently hydrolyzed substrates with (1→4)-beta-glycosidic bond. For an example, cellobiose was effectively hydrolyzed, and yielded glucose. The effeciency of BglD5 against longer-chains, for instance cellotriose, cellotetraose and cellopentose, was equally good. Although the cleavage of cellohexanose (with 6 glucose units) was possible, the relative hydrolysis activity was significantly reduced; thus suggesting that BglD5 is more suitable to cleave short-length cellodextrins. The enzyme was unable to hydrolyze sucrose and maltose that consist of alpha-glycosidic bonds. Besides, BglD5 was not active against polysaccharides such as Avicel, carboxymethyl cellulose, sigmacell cellulose, alpha-cellulose and xylan. This result is in line with previous findings, which also indicate that long chains of polysaccharides hinder the activity of BGLs due to substrates complexity [23, 27]. The enzyme exhibited a high activity against salicin (117.5% relative activity to cellobiose). This finding is in agreement with the report by Riou, et al. [28] since salicin is typically a good substrate for BGLs. Biomass OPEFB is a very common agricultural waste generated in Malaysia by the palm oil industry. Due to the highly complex structure of the raw OPEFB, BglD5 alone was unable to saccharify the non-chemical treated OPEFB. However, the addition of BglD5 into a commercial enzyme cocktail (Cellic® Ctec2) was able to successfully increase the total sugar production from OPEFB by approximately 4.5% (Appendix A1).

**Fig 2.** Enzymatic characterization of recombinant BglD5. (**a**) The optimal temperature of BglD5. Reactions were conducted in 100 mM sodium phosphate buffer (pH 7.4) at various temperatures for 15 min. (**b**) The optimal pH of BglD5. Reactions were conducted at 65°C for 15 min in various buffers, pH ranging from 2 to 11. (**c**) Effects of glucose (0 to 2500 mM) on BglD5 activity towards pNPG substrate. The reactions were conducted at 65°C and pH 7 for 15 min. (**d**) Thermostability of BglD5. After preincubation of BglD5 at 65°C (supplemented either with or without calcium chloride), the enzyme was withdrawn and subjected for enzyme assay, which was conducted at 65°C and pH 7 for 15 min. (**e**) Nonlinear Michaelis-Menten plot of the purified BglD5. The reactions were conducted at 65°C and pH 7 for 15 min, with pNPG concentration ranging from 0.05 to 3.00 mM.

**Table 1.** Properties of GH1 beta-glucosidase from various microorganism

**Table 2.** Effects of metal ions and chemical reagents to the activity of BglD5

**Table 3.** Hydrolysis of various substrates by purified BglD5

*3.2 Bioinformatic analyses of beta-glucosidase BglD5*

BglD5 is an intracellular enzyme in the wild type of *J. malaysiensis* due to the lack of a signal peptide. BglD5 is a member of GH1, as the protein sequence contained a GH1-specific domain. It shared the closest sequence similarity of 55.0-56.2% to BGLs from *Anoxybacillus* spp., followed by 54.3% to *Caldanaerobius fijiensis*, 51.0-54.1% to *Alicyclobacillus*, 52.3-53.7% to *Thermoanaerobacter*, 52.8% to *Pontibacillus halophilus*, 52.3% to *Tumebacillus flagellatus*, 52.3% to *Roseiflexus castenholzii*, 51.8-52.0% to *Orenia* spp., and 51.8% to

*Halobacteroides halobius*. The majority of the aforementioned homologous BGLs were annotated in whole-genome sequencing projects, and have not been biochemically characterized, with the exception of DT-Bgl from *Anoxybacillus* sp. DT3-1, which was previously reported by our team [23].

The differences between the motifs of BglD5 to 70 GH1 beta-glucosidases were generated using a Weblogo representation (Fig 3). It appeared that the motifs of BglD5 were slightly different at certain positions. The putative active sites and important substrate binding sites for BglD5 were predicted based on the well-investigated BGL sequence of fungi *Oryza sativa* subsp. *japonica* (Q7XKV4) [26]. Residue Glu 162 was identified as the acid base catalyst, while residue Glu 348 corresponded to the nucleophile. BglD5 is a genuine GH1 beta-glucosidase with the presence of dual glutamic acid as the catalytic site. The locations of five motifs in BglD5 were deduced by comparison to the BGL sequence of *Thermotoga maritima* (Q08638) (Fig 4).

A maximum-Parsimony phylogenetic tree was constructed to gain insight on the relationship among GH1 beta-glucosidases (Fig 5). BglD5 formed a cluster with BGLs from thermophilic *Anoxybacillus* spp. Interestingly, BglD5 shared a low similarity to GH1 beta-glucosidases identified in the genomes of *J. soli* (A0A0C2RQE4) and *J. campisalis* (A0A0C2W2Y2) (39-44%) [21, 22]; despite them belonging to the same genus. In addition, BGLs from *J. soli* and *J. campisalis* fall into other groups in the tree (Fig. 5). It is thus concluded that BglD5 is a novel beta-glucosidase that is closer to thermostable or thermophilic sources.

**Fig 3.** The consensus motif and signature of BglD5 and other 70 GH1 beta-glucosidases. Weblogo representation was generated using an online generator [20].

**Fig 4.** Multiple sequence alignment of BglD5 with GH1 beta-glucosidases from different microorganisms. Abbreviations: Jeotgali- *J. malaysiensis* BglD5 (this study), Anoxybacil: *Anoxybacillus* sp. DT3-1 (Uniprot id: M5QUM2), Thermoanae: *Thermoanaerobacter brockii* (Q60026) [12], Thermotoga: *Thermotoga maritima* (Q08638, PDB id 5N6S) [47], Paenibacil: *Paenibacillus polymyxa* (P22505, PDB id 2O9P) [48], Oryza: *Oryza sativa* subsp. *japonica* (Q7XKV4, PDB id 3PTK, 3PTM, and 3PTQ) [26], Trichoderma: *Trichoderma harzianum* (A0A127SA86, PDB id 5BWF) [31] and Humicola: *Humicola insolens* (A0A076JRL8, PDB id 4MDO) [25]. The active site is shown in the box. The underlined sequences are the locations of motifs and GH1 N-terminal signature. Amino acids shown with a yellow (subsite -1), blue (subsite +1), and green (subsite +2) background are the locations of subsites in the structure of 2O9P or 3PTQ. Red background residues functioned as gatekeepers in 4MDO. The grey colour background for BGL Thermoanae represents highly conserved residues that may be important for glucose tolerant characteristics [12].

**Fig 5.** Phylogenetic analysis of BglD5 with other BGLs. The Maximum-Parsimony (MP) tree was constructed using MEGA 6 program to elucidate the relationship of BglD5 and other enzymes across bacteria, archaea, and eukaryote.

## 4. Discussion

Beta-glucosidase is an enzyme that hydrolyzes glycosidic bonds to release non-reducing terminal glucosyl residues from glycosides and oligosaccharides [3]. In the past, detailed investigations have been performed for fungi BGLs [31, 35, 36]. Fungi such as *T. harzianum* has multiple genes encoded for GH1 and GH3 BGLs. The expression of GH1 BGL genes under biomass degradation was significantly higher compared to GH3 genes [31], and two of the highly expressed GH1 BGLs were previously discussed [49]. Due to the merit of thermostability, BGLs from hyper- and thermophiles were cloned, expressed and biochemically investigated [8, 33, 38]. According to Pei, et al. [8], thermostable BGLs could potentially facilitate the rates of the cellulose hydrolysis in biomass degradation. In addition, by using thermostable BGLs in cellulose hydrolysis, the product formation could be increased, while microbial contamination could be decreased, since the enzymatic hydrolysis process could be carried out at a higher temperature. Some BGLs from marine sources were also reported [37, 40, 50]. GH1 BGLs from various sources were characterized, and demonstrated diverse biochemical properties (Table 1).

*Jeotgalibacillus* is a rarely explored halophile. *J. malaysiensis* is gram-stain-positive, endospore-forming, rod-shaped, that optimally grew at a pH of 7 to 8 and 37°C [15]. To date, BGL from this genus has not been biochemically characterized. Despite the wild type *J. malaysiensis* being a mesophilic bacterium, BglD5 exhibited a relatively high enzymatic optimum temperature of 65°C, comparable to BGLs found in thermophilic counterparts such as *Anoxybacillus* sp. DT3-1 [23] and *Thermoanaerobacterium* spp. [8, 40]. BglD5 remained active at a pH range of 6 – 7.5, and this characteristic is similar to BGLs found in *R. flavipes* [34] and *F. islandicum* [38], whereby the enzymes were also highly active at a neutral pH.

The activity and stability of BglD5 was enhanced in the presence of calcium chloride. BglD5 had an activity half-life of 70 min at 65°C, in the presence of $CaCl_2$. Therefore, thermostable BglD5 is an attractive candidate to be applied in cellulase cocktails, in order to improve the biomass hydrolysis. The addition of $CaCl_2$ has a positive effect on other enzymes as well, for instance, BGLs obtained from *R. flavipes*, *Exiguobacterium* sp. DAU5 and *Weissella cibaria* 37 [34, 42, 43]. However, $CaCl_2$ did not improve the activity of BGLs obtained from *B. licheniformis*, *P. roxburghii*, *H. insolens* RP86, *Clostridium cellulovorans*, *Trichorderma reesei* and *Neotermes*

*koshunensis* [9, 24, 30, 35]. The binding pocket for $Ca^{2+}$ for GHs such as endoglucanase, α-amylase, pullulanase, and CGTase have been well-studied [51-55]. There is lack of studies in the literature on the $Ca^{2+}$ binding site for BGLs. In one particular study, a $Ca^{2+}$ ligand was identified in the crystal structure of GH1 BGL from *Thermotoga maritima* [56]. This particular structure consists of 4 chains (PDB id 2WC4). However, $Ca^{2+}$ bound only to chain-C interacted by a single amino acid, Glu329 (*T. maritima* numbering). It is likely that the interaction between $Ca^{2+}$ and Glu329 was unspecific. In another structure of *T. maritima* BGL (PDB id 2J7B), the $Ca^{2+}$ interacted with Asp278, Ser281, and Glu282 [57]. Nevertheless, these residues were not conserved for BglD5. It is thus at this point that the $Ca^{2+}$ binding pocket for BglD5 remains to be determined.

A slight increment in BglD5 enzyme activity was also observed when magnesium chloride was added to the reaction tube. A similar effect was previously reported, where $MgCl_2$ enhanced the BGLs from *Cellulomonasmicrobium cellulans* and *T. thermosaccharolyticum* [8, 29]. Similar to that of the $Ca^{2+}$ cation, the binding site for the $Mg^{2+}$ cation was not emphasized in any earlier studies. The protein structure of GH1 BGLs that were found to have an $Mg^{2+}$ binding site are 3VKK and 2ZOX, both of which originate from humans [58]. The physiological relevance of $Mg^{2+}$ to BGL is yet to be confirmed.

To date, over 330 BGL structures of various GH families are available at the PDB database. GH1 BGLs consist of a typical TIM-barrel fold similar to BGL from *P. polymyxa* (PDB id 2O9P), *H. orenii* (4PTX) or *T. harzianum* (5BWF), and the active site pocket is located near the C-terminal region of the $(\alpha/\beta)_8$ barrel [31]. Based on template 4PTX, the predicted structure of BglD5 has a similar overall structure to the template (Appendix A2). The sequences of GH1 BGLs are very diverse. Among the examined 70 sequences (Appendix A3), only 17 residues are completely identical at designated positions. Almost all GH1 BGLs had a NEP amino acid stretch (161-163 BglD5 numbering), where the Glu is the catalytic sites (acid/base). Interestingly, enzymes identified in some *Bacillus* spp. and *J. soli* (A0A0C2RQE, A0A0Q6HVV2, D5E2L2, A0A0M1NP20, and A0A098FCN7; formed a group in Fig. 5) had the stretch NET (Threonine instead of Proline) following the acid/base catalytic site. Besides, stretch TENG (347-350 BglD5 numbering, where the Glu is among the catalytic sites (nucleophile), is another conserved region in GH1 BGLs from bacteria, archaea and eukaryotes. Exceptionally, vegetative weed *Arabidopsis thaliana* (AAB64244.1) had a sequence of MENG.

SPRINT cataloged the presence of five motifs for BGLs [19]. Nucleophile Glu348 was placed in motif 2. Motifs 1, 4, and 5 consist of several important substrate interacting residues (Fig. 3). These five motifs are very diversifying in amino acid selection (Fig. 4). Interestingly, we notice the presence of a highly conserved region (BglD5: $^{11}$GTATSSFQI$^{19}$) near the N-terminal of GH1 beta-glucosidase that was not curated in the SPRINT database. ScanProsite predicted that this conserved region is Glycosyl Hydrolases family 1 N-terminal signature (Fig. 3 and 4). Based on earlier findings related to BGL obtained from *O. sativa subsp. japonica*, a residue Gln (Q29) in protrude to the direction of glycon residue and formed hydrogen bond around subsite -1 [26]. It is thus possible that N-terminal signature may be important for substrate recognition or catalytic reaction.

As mentioned previously, the BGL superfamily was grouped into four GH families. Pei, et al. [8] proposed 5 clades for BGLs affiliated to GH1 and GH3. Accordingly, the report suggested that Clade I is represented by GH1 enzymes from mesophilic bacteria; Clade II is grouped by GH1 BGL from fungi; Clade III by GH3 BGL from bacteria; Clade IV by GH3 enzyme from fungi; and Clade V contains a mixture of GH1 BGL from thermophilic bacteria and *Bacillus*. As the analysis applied limited datasets, it may not comprehensively provide clear insight. Besides, the proposed grouping may be confusing, as the existing clade-groupings for GH3 beta-glucosidase [59] have not been considered in the analysis [8]. In fact, CAZy has already classified families GH1, GH5, and GH30 to the Clan GH-A [3]. Therefore, classification suggested by Pei, et al. [8] should be relooked, since, indeed, there is a need to propose more comprehensive groupings. In this work, we constructed a Maximum Parsimony tree using 70 sequences of genuine GH1 (Criteria: (i) contains 2 highly conserved Glu as catalytic residues, and (ii) consisted of GH1 domain). The tree (Fig. 5) is robust, as the placement of taxa is similar to when a Neighbour-Joining tree was used (not shown). BglD5 was clustered together with the BGLs from thermophilic *Anoxybacillus* spp. Interestingly, two other uncharacterized BGLs from *J. soli* (closer to *Bacillus megaterium* beta-glucosidase) and *J. campisalis* (closer to *Bacillus halodurans*) were located far away from BglD5, suggesting that BglD5 had evolved more rapidly compared to other counterparts. BGLs from hyperthermophiles, in particular, archaea, were clustered at the base of the tree, while enzymes from eukaryotic sources (fungi, termites, plants, and silkworms) formed a distinctive group. The rest of the trees were mainly BGL sequences from bacteria. For the current tree, we could only suggest a relative relationship among these enzymes, and it is not possible to propose defined clades, since BGLs (in particular, bacteria origin) were not grouped according to taxonomy, optimum growth temperature, or salinity requirements. Instead of sequence classification, functional classification summarized by a review article [13] may be more appropriate. According to the authors, BGLs are categorized as class I (aryl beta-glucosidases), class II (true cellobiases), and class III (broad substrate specificity enzymes), and we agree that this functional classification is important to differentiate BGLs. For the group that consists of BglD5 from *J. malaysiensis*, DT-Bgl from *Anoxybacillus* sp. DT3-1, and other uncharacterized BGLs from *Anoxybacillus* spp. (Fig. 5), these BGLs can be classified as class III.

Similar to BGLs isolated from *B. subtilis* and *T. thermarum* DSM 5069T [32, 33], BglD5 exhibited glucose-tolerant characteristics. BglD5 tolerated up to 2500 mM glucose and retained about 80% of its maximum activity, yet the tolerance level is far lower to that of DT-Bgl from *Anoxybacillus* sp. DT3-1, which was previously reported by our team [23]. It has been reported that GH1 BGLs are 10- to 1000- fold more glucose-tolerant compared to GH3 BGLs, due to the architecture of active site pocket [10, 60]. GH1 enzymes harbor a deeper and narrower active site at $(\alpha/\beta)_8$ barrel, while GH3 BGLs have relatively shallower cavity where glucose easily binds at the catalytic pocket, and results in a competitive inhibition with real substrates [25]. In fact, the location of the active site for GH1 and GH3 enzymes is different. In GH3 BGLs, the catalytic center is located at the interface between $(\alpha/\beta)_8$ barrel and $(\alpha/\beta)_6$ domain [59]. Collectively, GH1 BGLs are more suitable to be used for the final step of cellulose saccharification to form fermentable glucose. Mariano, et al. [12] performed a bioinformatic analysis of 23 protein sequences of glucose-tolerant BGLs. Collectively, the team suggested that 11 residues, namely, H121, W122, N166, E167, N297, Y299, E355, W402, E409, W410, and F418 (based on the numbering of *T. brockii* beta-glucosidase Q60026), may be crucial for glucose tolerant behaviors (Fig. 4). Through protein engineering of two BGLs, Yang, et al. [61] have elucidated that glucose tolerance is determined by more than a single site. This finding complied with the findings of work [62], where a mutant with three replacement sites (W174C/A404V/L441F) exhibited greater thermostability, glucose tolerance and activity. In another report, Liu, et al. [50] suggest that the replacement of H184 with Phe significantly increased glucose tolerance for a metagenome-derived enzyme [61]. The shape and electrostatic properties of the deep active-site entrance, and the +2 subsite, in particular, Trp168 and Leu173 (*H. insolens* A0A076JRL8 numbering), determined the degree of tolerance to glucose in GH1 members [25]. Besides, for *H. insolens* BGL, three phenylalanine residues (Phe325, Phe333 and Phe348), along with another four residues (highlighted in red in Fig. 4), functioned as gatekeepers, and demarcated a restricted entrance to the active site [25]. Nevertheless, these residues are not conserved in BglD5. In a separate finding, residual adjacent to subsite +2 may also regulate glucose tolerance, as reported by mutagenesis H183F (equivalent to 176 in BglD5 numbering) in a marine metagenome derived BGL [50].

## 5. Conclusions

In conclusion, the *bglD5* gene from a marine bacterium *J. malaysiensis* coding for beta-glucosidase was successfully expressed in *E. coli*, and it was found out that $CaCl_2$ enhanced the activity of the heterologous enzyme. BglD5 demonstrates its advantageous application in the saccharification process of cellulosic materials, attributable to good thermostability and low glucose inhibition. The enzyme is classified as class III BGL.

## Conflicts of Interest

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## Acknowledgments

This work was supported by Universiti Teknologi Malaysia Research Grants (Grant No. 16H89, & 14H67) awarded to M.S. Shamsir and K.M. Goh. The authors gratefully acknowledge the financial support provided by grants from JBK, NRE, University of Malaya (GA001-2016, GA002-2016) awarded to Kok-Gan Chan, FASc . K.J. Liew thanks the Zamalah Program of Universiti Teknologi Malaysia for scholarship support.

## Appendix A. Supplementary Data

Please refer to the link provided for supplementary data.

Appendix A1 Comparison of sugar production from OPEFB, treatment with (i) BglD5, (ii) Cellic® Ctec2, and (iii) BglD5+Cellic® Ctec2 (https://figshare.com/s/a53a86b97be931b6df53)

Appendix A2 3D model structure of BglD5 predicted by I-TASSER with its secondary structure labelled in pink – alpha-helix; yellow – beta-sheet; and grey – loops. (https://figshare.com/s/e2e7c49b2d126fce75dc)

Appendix A3 Multiple sequence alignment of 70 selected GH1 beta-glucosidases. (https://figshare.com/s/984a967f2a261dd7e1ea)

## References

[1] S. Banerjee, S. Mudliar, R. Sen, B. Giri, D. Satpute, T. Chakrabarti, R. Pandey, Commercializing lignocellulosic bioethanol: technology bottlenecks and possible remedies, Biofuels Bioprod. Bioref. 4 (2010) 77–93.

[2] Y. Han, H. Chen, Characterization of β-glucosidase from corn stover and its application in simultaneous saccharification and fermentation, Bioresour. Technol. 99 (2008) 6081–6087.

[3] J.R.K. Cairns, A. Esen, β-Glucosidases, Cell. Mol. Life Sci. 67 (2010) 3389–3405.

[4] A. White, D.R. Rose, Mechanism of catalysis by retaining β-glycosyl hydrolases, Curr. Opin. Struct. Biol. 7 (1997) 645–651.

[5] Y. Guo, Q. Yan, Y. Yang, S. Yang, Y. Liu, Z. Jiang, Expression and characterization of a novel β-glucosidase, with transglycosylation and exo-β-1, 3-glucanase activities, from *Rhizomucor miehei*, Food Chem. 175 (2015) 431–438.

[6] M.A. Villena, J.F.U. Iranzo, A.I.B. Pérez, β-Glucosidase activity in wine yeasts: application in enology, Enzyme Microb. Technol. 40 (2007) 420–425.

[7] G.P. Maitan-Alfenas, G.d.A. Lorena, M.N. de Almeida, E.M. Visser, S.T. de Rezende, V.M. Guimarães, Hydrolysis of soybean isoflavones by *Debaryomyces hansenii* UFV-1 immobilised cells and free β-glucosidase, Food Chem. 146 (2014) 429–436.

[8] J. Pei, Q. Pang, L. Zhao, S. Fan, H. Shi, *Thermoanaerobacterium thermosaccharolyticum* β-glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose, Biotechnol. Biofuels 5 (2012) 31.

[9] Z. Chen, T. Meng, Z. Li, P. Liu, Y. Wang, N. He, D. Liang, Characterization of a beta-glucosidase from *Bacillus licheniformis* and its effect on bioflocculant degradation, AMB Express 7 (2017) 197.

[10] R.R. Singhania, A. Patel, A. Pandey, E. Gnasounou, Genetic modification: a tool for enhancing beta-glucosidase production for biofuel application, Bioresour. Technol. (2017).

[11] R.R. Singhania, A.K. Patel, R.K. Sukumaran, C. Larroche, A. Pandey, Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production, Bioresour. Technol. 127 (2013) 500–507.

[12] D. Mariano, C. Leite, L. Santos, L. Marins, K. Machado, A. Werhli, L. Lima, R. de Melo-Minardi, Characterization of glucose-tolerant β-glucosidases used in biofuel production under the bioinformatics perspective: a systematic review, Genet. Mol. Res. 16 (2017).

[13] P. Tiwari, B. Misra, N.S. Sangwan, β-Glucosidases from the fungus *Trichoderma*: an efficient cellulase machinery in biotechnological applications, BioMed Res. Int. 2013 (2013) e203735.

[14] Y. Bhatia, S. Mishra, V. Bisaria, Microbial β-glucosidases: cloning, properties, and applications, Crit. Rev. Biotechnol. 22 (2002) 375–407.

[15] A.S. Yaakop, K.-G. Chan, R. Ee, U.M. Kahar, W.C. Kon, K.M. Goh, Isolation of *Jeotgalibacillus malaysiensis* sp. nov. from a sandy beach, and emended description of the genus *Jeotgalibacillus*, Int. J. Syst. Evol. Microbiol. 65 (2015) 2215–2221.

[16] K.M. Goh, K.-G. Chan, A.S. Yaakop, R. Ee, Complete genome of *Jeotgalibacillus malaysiensis* D5(T) consisting of a chromosome and a circular megaplasmid, J. Biotechnol. 204 (2015) 13–14.

[17] A.S. Yaakop, K.-G. Chan, R. Ee, Y.L. Lim, S.-K. Lee, F.A. Manan, K.M. Goh, Characterization of the mechanism of prolonged adaptation to osmotic stress of *Jeotgalibacillus malaysiensis* via genome and transcriptome sequencing analyses, Sci. Rep. 6 (2016) 33660.

[18] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729.

[19] T.K. Attwood, P. Bradley, D.R. Flower, A. Gaulton, N. Maudling, A.L. Mitchell, G. Moulton, A. Nordle, K. Paine, P. Taylor, PRINTS and its automatic supplement, prePRINTS, Nucleic Acids Res. 31 (2003) 400–402.

[20] G.E. Crooks, G. Hon, J.-M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.

[21] A.S. Yaakop, K.-G. Chan, H.M. Gan, K.M. Goh, Draft genome of *Jeotgalibacillus campisalis* SF-57(T), a moderate halophilic bacterium isolated from marine saltern, Mar. Genomics 23 (2015) 59–60.

[22] K.M. Goh, K.-G. Chan, A.S. Yaakop, C.S. Chan, R. Ee, W.-S. Tan, H.M. Gan, Draft genome sequence of *Jeotgalibacillus soli* DSM 23228, a bacterium isolated from alkaline sandy soil, Genome Announc. 3 (2015) e00512–15.

[23] C.S. Chan, L.L. Sin, K.-G. Chan, M.S. Shamsir, F.A. Manan, R.K. Sani, K.M. Goh, Characterization of a glucose-tolerant β-glucosidase from *Anoxybacillus* sp. DT3-1, Biotechnol. Biofuels 9 (2016) 174.

[24] W.-Y. Jeng, N.-C. Wang, M.-H. Lin, C.-T. Lin, Y.-C. Liaw, W.-J. Chang, C.-I. Liu, P.-H. Liang, A.H.-J. Wang, Structural and functional analysis of three β-glucosidases from bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis*, J. Struct. Biol. 173 (2011) 46–56.

[25] P.O. de Giuseppe, T.d.A. Souza, F.H.M. Souza, L.M. Zanphorlin, C.B. Machado, R.J. Ward, J.A. Jorge, R.d.P.M. Furriel, M.T. Murakami, Structural basis for glucose tolerance in GH1 β-glucosidases, Acta Crystallogr. D Biol. Crystallogr. 70 (2014) 1631–1639.

[26] S. Sansenya, R. Opassiri, B. Kuaprasert, C.-J. Chen, J.R.K. Cairns, The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing β-glucosidase with significant thioglucohydrolase activity, Arch. Biochem. Biophys. 510 (2011) 62–72.

[27] S. Paavilainen, J. Hellman, T. Korpela, Purification, characterization, gene cloning, and sequencing of a new beta-glucosidase from *Bacillus circulans* subsp. *alkalophilus*, Appl. Environ. Microbiol. 59 (1993) 927–932.

[28] C. Riou, J.-M. Salmon, M.-J. Vallier, Z. Günata, P. Barre, Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus oryzae*, Appl. Environ. Microbiol. 64 (1998) 3607–3614.

[29] Y. Yuan, F. Xu, J. Yao, Y. Hu, J. Wang, T. Zhao, Y. Zhou, J. Gao, Cloning, expression and biochemical characterization of a GH1 β-glucosidase from *Cellulosimicrobium cellulans*, Biocatal. Biotransformation (2017) 1–10.

[30] B. Kar, P. Verma, G.K. Patel, A.K. Sharma, Molecular cloning, characterization and in silico analysis of a thermostable β-glucosidase enzyme from *Putranjiva roxburghii* with a significant activity for cellobiose, Phytochemistry 140 (2017) 151–165.

[31] C.A. Santos, L.M. Zanphorlin, A. Crucello, C.C. Tonoli, R. Ruller, M.A. Horta, M.T. Murakami, A.P. Souza, Crystal structure and biochemical characterization of the recombinant ThBgl, a GH1 β-glucosidase overexpressed in *Trichoderma harzianum* under biomass degradation conditions, Biotechnol. Biofuels 9 (2016) 71.

[32] S. Chamoli, P. Kumar, N.K. Navani, A.K. Verma, Secretory expression, characterization and docking study of glucose-tolerant β-glucosidase from *B. subtilis*, Int. J. Biol. Macromol. 85 (2016) 425–433.

[33] L. Zhao, J. Xie, X. Zhang, F. Cao, J. Pei, Overexpression and characterization of a glucose-tolerant β-glucosidase from *Thermotoga thermarum* DSM 5069T with high catalytic efficiency of ginsenoside Rb1 to Rd, J. Mol. Catal. B Enzym. 95 (2013) 62–69.

[34] M.E. Scharf, E.S. Kovaleva, S. Jadhao, J.H. Campbell, G.W. Buchman, D.G. Boucias, Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite *Reticulitermes flavipes*, Insect Biochem. Mol. Biol. 40 (2010) 611–620.

[35] F.H.M. Souza, L.P. Meleiro, C.B. Machado, A.L.R.L. Zimbardi, R.F. Maldonado, T.A.C.B. Souza, D.C. Masui, M.T. Murakami, J.A. Jorge, R.J. Ward, Gene cloning, expression and biochemical characterization of a glucose-and xylose-stimulated β-glucosidase from *Humicola insolens* RP86, J. Mol. Catal. B Enzym. 106 (2014) 1–10.

[36] B. Benoliel, M.J. Poças-Fonseca, F.A.G. Torres, L.M.P. de Moraes, Expression of a glucose-tolerant β-glucosidase from *Humicola grisea* var. *thermoidea* in *Saccharomyces cerevisiae*, Applied Biochem. Biotechnol. 160 (2010) 2036–2044.

[37] Z. Fang, W. Fang, J. Liu, Y. Hong, H. Peng, X. Zhang, B. Sun, Y. Xiao, Cloning and characterization of a β-glucosidase from marine microbial metagenome with excellent glucose tolerance, J. Microbiol. Biotechnol. 20 (2010) 1351–1358.

[38] D. Jabbour, B. Klippel, G. Antranikian, A novel thermostable and glucose-tolerant β-glucosidase from *Fervidobacterium islandicum*, Applied Biochem. Biotechnol. 93 (2012) 1947–1956.

[39] J. Lu, L. Du, Y. Wei, Y. Hu, R. Huang, Expression and characterization of a novel highly glucose-tolerant β-glucosidase from a soil metagenome, Acta Biochim. Biophys. Sin. 45 (2013) 664–673.

[40] F. Yang, X. Yang, Z. Li, C. Du, J. Wang, S. Li, Overexpression and characterization of a glucose-tolerant β-glucosidase from *T. aotearoense* with high specific activity for cellobiose, Applied Biochem. Biotechnol. 99 (2015) 8903–8915.

[41] P. Ramachandran, M.K. Tiwari, R.K. Singh, J.-R. Haw, M. Jeya, J.-K. Lee, Cloning and characterization of a putative β-glucosidase (NfBGL595) from *Neosartorya fischeri*, Process Biochem. 47 (2012) 99–105.

[42] J. Chang, I.-H. Park, Y.-S. Lee, S.-C. Ahn, Y. Zhou, Y.-L. Choi, Cloning, expression, and characterization of [beta]-glucosidase from *Exiguobacterium* sp. DAU5 and transglycosylation activity, Biotechnol. Bioprocess Eng. 16 (2011) 97.

[43] K.W. Lee, N.S. Han, J.H. Kim, Purification and characterization of beta-glucosidase from *Weissella cibaria* 37, J. Microbiol. Biotechnol. 22 (2012) 1705–1713.

[44] N.A. Spiridonov, D.B. Wilson, Cloning and biochemical characterization of BglC, a β-glucosidase from the cellulolytic actinomycete *Thermobifida fusca*, Curr. Microbiol. 42 (2001) 295–301.

[45] S. Naz, N. Ikram, M. Rajoka, S. Sadaf, M. Akhtar, Enhanced production and characterization of a [beta]-glucosidase from *Bacillus halodurans* expressed in *Escherichia coli*, Biochemistry 75 (2010) 513.

[46] C. Schröder, S. Elleuche, S. Blank, G. Antranikian, Characterization of a heat-active archaeal β-glucosidase from a hydrothermal spring metagenome, Enzyme Microb. Technol. 57 (2014) 48–54.

[47] T.J. Beenakker, D.P. Wander, W.A. Offen, M. Artola, L. Raich, M.J. Ferraz, K.-Y. Li, J.H. Houben, E.R. van Rijssel, T. Hansen, Carba-cyclophellitols are neutral retaining-glucosidase inhibitors, J. Am. Chem. Soc. 139 (2017) 6534–6537.

[48] P. Isorna, J. Polaina, L. Latorre-García, F.J. Cañada, B. González, J. Sanz-Aparicio, Crystal structures of *Paenibacillus polymyxa* β-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases, J. Mol. Biol. 371 (2007) 1204–1218.

[49] R.N. Florindo, V.P. Souza, H.S. Mutti, C. Camilo, L.R. Manzine, S.R. Marana, I. Polikarpov, A.S. Nascimento, Structural insights into β-glucosidase transglycosylation based on biochemical, structural and computational analysis of two GH1 enzymes from *Trichoderma harzianum*, N. Biotechnol. 40 (2018) 218–227.

[50] J. Liu, X. Zhang, Z. Fang, W. Fang, H. Peng, Y. Xiao, The 184th residue of β-glucosidase Bgl1B plays an important role in glucose tolerance, J. Biosci. Bioeng. 112 (2011) 447–450.

[51] S. Chauvaux, P. Beguin, J.-P. Aubert, K.M. Bhat, L.A. Gow, T.M. Wood, A. Bairoch, Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanase D, Biochem. J. 265 (1990) 261–265.

[52] K.P. Chai, N.F.B. Othman, A.-H. Teh, K.L. Ho, K.-G. Chan, M.S. Shamsir, K.M. Goh, C.L. Ng, Crystal structure of *Anoxybacillus* α-amylase provides insights into maltose binding of a new glycosyl hydrolase subclass, Sci. Rep. 6 (2016) 23126.

[53] E. Boel, L. Brady, A.M. Brzozowski, Z. Derewenda, G.G. Dodson, V.J. Jensen, S.B. Petersen, H. Swift, L. Thim, H.F. Woldike, Calcium binding in α-amylases: an x-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus*, Biochemistry 29 (1990) 6244–6249.

[54] J. Xu, F. Ren, C.H. Huang, Y. Zheng, J. Zhen, H. Sun, T.P. Ko, M. He, C.C. Chen, H.C. Chan, Functional and structural studies of pullulanase from *Anoxybacillus* sp. LM18‐11, Proteins: Structure, Function, and Bioinformatics 82 (2014) 1685–1693.

[55] C. Klein, G.E. Schulz, Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution, J. Mol. Biol. 217 (1991) 737–750.

[56] M. Aguilar-Moncayo, T.M. Gloster, J.P. Turkenburg, M.I. García-Moreno, C.O. Mellet, G.J. Davies, J.M.G. Fernández, Glycosidase inhibition by ring-modified castanospermine analogues: tackling enzyme selectivity by inhibitor tailoring, Org. Biomol. Chem. 7 (2009) 2738–2747.

[57] T.M. Gloster, P. Meloncelli, R.V. Stick, D. Zechel, A. Vasella, G.J. Davies, Glycosidase inhibition: an assessment of the binding of 18 putative transition-state mimics, J. Am. Chem. Soc. 129 (2007) 2345–2354.

[58] J. Noguchi, Y. Hayashi, Y. Baba, N. Okino, M. Kimura, M. Ito, Y. Kakuta, Crystal structure of the covalent intermediate of human cytosolic β-glucosidase, Biochem. Biophys. Res. Commun. 374 (2008) 549–552.

[59] A.J. Harvey, M. Hrmova, R. De Gori, J.N. Varghese, G.B. Fincher, Comparative modeling of the three‐dimensional structures of family 3 glycoside hydrolases, Proteins: Struct. Funct. Bioinf. 41 (2000) 257–269.

[60] L.P. Meleiro, J.C.S. Salgado, R.F. Maldonado, S. Carli, L.A.B. Moraes, R.J. Ward, J.A. Jorge, R.P.M. Furriel, Engineering the GH1 β-glucosidase from *Humicola insolens*: Insights on the stimulation of activity by glucose and xylose, PLoS One 12 (2017) e0188254.

[61] Y. Yang, X. Zhang, Q. Yin, W. Fang, Z. Fang, X. Wang, X. Zhang, Y. Xiao, A mechanism of glucose tolerance and stimulation of GH1 β-glucosidases, Sci. Rep. 5 (2015).

[62] L.-C. Cao, Z.-J. Wang, G.-H. Ren, W. Kong, L. Li, W. Xie, Y.-H. Liu, Engineering a novel glucose-tolerant β-glucosidase as supplementation to enhance the hydrolysis of sugarcane bagasse at high glucose concentration, Biotechnol. Biofuels 8 (2015) 202.

**List of Figures**

**Fig 1.** SDS-PAGE analysis of purified BglD5. Lane M: Benchmark™ Protein Ladder (Thermo Scientific, USA); lane 1: cell lysate from recombinant *E. coli* BL21 (DE3); lane 2: flowthrough (uncaptured) fraction from purification column; lane 3: purified BglD5. A 5 µL of each sample was loaded into each well and the gel was stained by Imperial™ Protein Stain (Thermo Scientific, USA).

**Fig 2.** Enzymatic characterization of recombinant BglD5. (**a**) The optimal temperature of BglD5. Reactions were conducted in 100 mM sodium phosphate buffer (pH 7.4) at various temperatures for 15 min. (**b**) The optimal pH of BglD5. Reactions were conducted at 65°C for 15 min in various buffers, pH ranging from 2 to 11. (**c**) Effects of glucose (0 to 2500 mM) on BglD5 activity towards pNPG substrate. The reactions were conducted at 65°C and pH 7 for 15 min. (**d**) Thermostability of BglD5. After preincubation of BglD5 at 65°C (supplemented either with or without calcium chloride), the enzyme was withdrawn and subjected for enzyme assay, which was conducted at 65°C and pH 7 for 15 min. (**e**) Nonlinear Michaelis-Menten plot of the purified BglD5. The reactions were conducted at 65°C and pH 7 for 15 min, with pNPG concentration ranging from 0.05 to 3.00 mM.

**Fig 3.** The consensus motif and signature of BglD5 and other 70 GH1 beta-glucosidases. Weblogo representation was generated using an online generator [20].

**Fig 4.** Multiple sequence alignment of BglD5 with GH1 beta-glucosidases from different microorganisms. Abbreviations: Jeotgali- *J. malaysiensis* BglD5 (this study), Anoxybacil: *Anoxybacillus* sp. DT3-1 (Uniprot id: M5QUM2), Thermoanae: *Thermoanaerobacter brockii* (Q60026) [12], Thermotoga: *Thermotoga maritima* (Q08638, PDB id 5N6S) [47], Paenibacil: *Paenibacillus polymyxa* (P22505, PDB id 2O9P) [48], Oryza: *Oryza sativa* subsp. *japonica* (Q7XKV4, PDB id 3PTK, 3PTM, and 3PTQ) [26], Trichoderma: *Trichoderma harzianum* (A0A127SA86, PDB id 5BWF) [31] and Humicola: *Humicola insolens* (A0A076JRL8, PDB id 4MDO) [25]. The active site is shown in the box. The underlined sequences are the locations of motifs and GH1 N-terminal signature. Amino acids shown with a yellow (subsite -1), blue (subsite +1), and green (subsite +2) background are the locations of subsites in the structure of 2O9P or 3PTQ. Red background residues functioned as gatekeepers in 4MDO. The grey colour background for BGL Thermoanae represents highly conserved residues that may be important for glucose tolerant characteristics [12].

**Fig 5.** Phylogenetic analysis of BglD5 with other BGLs. The Maximum-Parsimony (MP) tree was constructed using MEGA 6 program to elucidate the relationship of BglD5 and other enzymes across bacteria, archaea, and eukaryote.

**List of Tables**

**Table 1.** Properties of GH1 beta-glucosidase from various microorganism

**Table 2.** Effects of metal ions and chemical reagents to the activity of BglD5

**Table 3.** Hydrolysis of various substrates by purified BglD5

**Table 1.** Properties of GH1 beta-glucosidase from various microorganism

| Source | MW (kDa) | Opt. temp. °C | Opt. pH | Thermostability | Km (mM)[a] | Vmax (U/mg)[b] | Ref |
|---|---|---|---|---|---|---|---|
| *J. malaysiensis* (this study) | 52 | 65 | 7 | $t_{1/2}$ 65°C = 35 min (w/o $CaCl_2$) $t_{1/2}$ 65°C = 70 min (with $CaCl_2$) | 0.50 | 39.48 | this study |
| *Anoxybacillus* sp. DT3-1 | 53 | 70 | 8.5 | $t_{1/2}$ 60°C = 24 hours | 0.22 | 923.7 | [23] |
| *Cellulosimicrobium cellulans* | 57 | 55 | 6.0 | $t_{1/2}$ 45°C = 25 min $t_{1/2}$ 50°C = 3 min | 0.36 | 4.09 | [29] |
| *Putranjiva roxburghii* | 61 | 65 | 4.6 | Native: Retained 67% activity after incubation at 70°C for 1 hour Recombinant: Retained 35% activity after incubation at 70°C for 1 hour | Native: 0.53 (pNPG) 0.81 (cellobiose) Recombinant: 0.58 (pNPG) 0.87 (cellobiose) | Native: 0.181 µkat/mg (pNPG) 0.122 µkat/mg (cellobiose) Recombinant: 0.157 µkat/mg (pNPG) 0.105 µkat/mg (cellobiose) | [30] |
| *Trichoderma harzianum* | 54 | 40 | 6 | CD thermal-induced unfolding, Tm = 49°C | 0.97 (pNPG) 1.22 (cellobiose) | 29.3 (pNPG) 10.4 (cellobiose) | [31] |
| *Bacillus subtilis* | 53 | 60 | 6 | $t_{1/2}$ 70°C = 45 min | 0.82 | 55.55 | [32] |
| *Thermoanaerobacterium thermosaccharolyticum* DSM 571 | 52 | 70 | 6.4 | $t_{1/2}$ 68°C = 1 hour | 0.62 (pNPG) 7.9 (cellobiose) | 64 (pNPG) 120 (cellobiose) | [8] |
| *Thermotoga thermarum* DSM 5069T | 55 | 90 | 4.8 | $t_{1/2}$ 90°C = 2 hours $t_{1/2}$ 95°C = 50 min | 0.59 | 142 | [33] |
| *Reticulitermes flavipes* | 56 | <40 | 7.0 | $t_{1/2}$ 45°C = 20 min (w/o $CaCl_2$) $t_{1/2}$ 45°C = 30 min (with $CaCl_2$) | 1.66 (pNPG) 1.44 (cellobiose) | 22.92 (pNPG) 278 (cellobiose) | [34] |
| *Humicola insolens* RP86 | 56 | 60 | 5-7 | $t_{1/2}$ 55°C = 50 min | 0.20 mmol/L (pNPG) 0.38 mmol/L (cellobiose) | 36.4 (pNPG) 183.4 (cellobiose) | [35] |
| *Humicola grisea var. thermoidea* | 57 | 40 | 6 | Retained 80% activity after incubation at 40°C for 1 hour | 0.16 | 6.72 µmol/min | [36] |
| Marine microbial | 51 | 40 | 6.5 | Retained 70% activity after | 0.39 (pNPG) | 50.7 (pNPG) | [37] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| metagenome | | | | incubation at pH 7.5 and 15°C for 1 hour | 20.4 (cellobiose) | 15.5 (cellobiose) | |
| *Fervidobacterium islandicum* | 53 | 90 | 6-7 | $t_{1/2}$ 90°C = 25 min<br>$t_{1/2}$ 100°C = 15 min | n.a | n.a | [38] |
| soil metagenome | 52 | 50 | 6 | n.a | 2.09 | 183.9 | [39] |
| *Thermoanaerobacterium aotearoense* P8G3#4 | 46 | 60 | 6 | $t_{1/2}$ 55°C = 90min | 0.66 (pNPG)<br>25.45 (cellobiose) | 25.45 (pNPG)<br>740.5 (cellobiose) | [40] |
| *Neosartorya fischeri* NRRL181 | 60 | 40 | 6 | $t_{1/2}$ 37°C = 480 min<br>$t_{1/2}$ 40°C = 60 min<br>$t_{1/2}$ 50°C = 17 min<br>$t_{1/2}$ 60°C = 8 min | 2.8 | 1693 | [41] |
| *Exiguobacterium* sp. DAU5 | 52 | 45 | 7 | $t_{1/2}$ 40°C = 60 min | 2.33 | 31.6 | [42] |
| *Weissella cibaria* 37 | 50 | 45 | 5.5 | $t_{1/2}$ 45°C = 30min | 0.92 µmol/min | 0.92 µmol/min | [43] |
| *Thermobifida fusca* | 53 | 50 | 7 | Stable at 60°C and rapidly inactivated at 65°C | 0.34 (cellobiose) | 29 (cellobiose) | [44] |
| *Bacillus halodurans* | 51 | 45 | 8 | Retained 80% activity after incubation at 45°C for 1 hour | 4 | n.a | [45] |
| Hot spring metagenome (archaea) | 57 | 90 | 6.5 | $t_{1/2}$ 70°C = 14 hours | 0.8 | 0.022 µmol/min | [46] |

[a, b] Unless specified, the data were obtained with standard assay using pNPG as substrate.

**Table 2.** Effects of metal ions and chemical reagents to the activity of BglD5

| **Metal ions/Chemical reagents (5 mM or 5% *v/v*)** | **Relative activity (%)** |
|---|---|
| Reference | 100 |
| Calcium Chloride | 121 |
| Sodium Chloride | 100 |
| Potassium Chloride | 104 |
| Magnesium Chloride | 109 |
| Iron (III) Chloride | 6 |
| Nickel (II) Chloride | 3 |
| Cobalt (II) Chloride | 4 |
| Ammonium Chloride | 95 |
| Zinc Sulfate | 7 |
| Manganese Sulfate | 7 |
| Copper (II) Sulfate | 0 |
| Rubidium Chloride | 77 |
| Strondium Chloride | 95 |
| Barium Chloride | 61 |
| Urea | 85 |
| SDS | 2 |
| EDTA | 4 |
| Tween-20 (5% *v/v*) | 75 |
| Tween-80 (5% *v/v*) | 47 |
| Triton X-100 (5% *v/v*) | 70 |
| DMSO (5% *v/v*) | 84 |

**Table 3.** Hydrolysis of various substrates by purified BglD5

| Substrates | Relative hydrolysis ability (%) |
|---|---|
| Cellobiose | 100 |
| Cellotriose | 124 |
| Cellotetraose | 125 |
| Cellopentose | 125 |
| Cellohexanose | 85 |
| Salicin | 118 |
| Gentibiose | 54 |
| Lactose | 34 |
| Sucrose | 0 |
| Maltose | 0 |
| Avicel | 0 |
| Carboxymethyl cellulose | 0 |
| Sigmacell cellulose | 0 |
| Alpha-cellulose | 0 |
| Xylan | 0 |

Highlights

- BglD5 is the first discovered beta-glucosidase from halophilic genus *Jeotgalibacilus.*
- BglD5 belongs to the Glycoside Hydrolase family 1 together with a lately discovery beta-glucosidase from *Anoxybacillus.*
- BglD5 is a class III (broad substrate specificity enzymes) beta-glucosidase

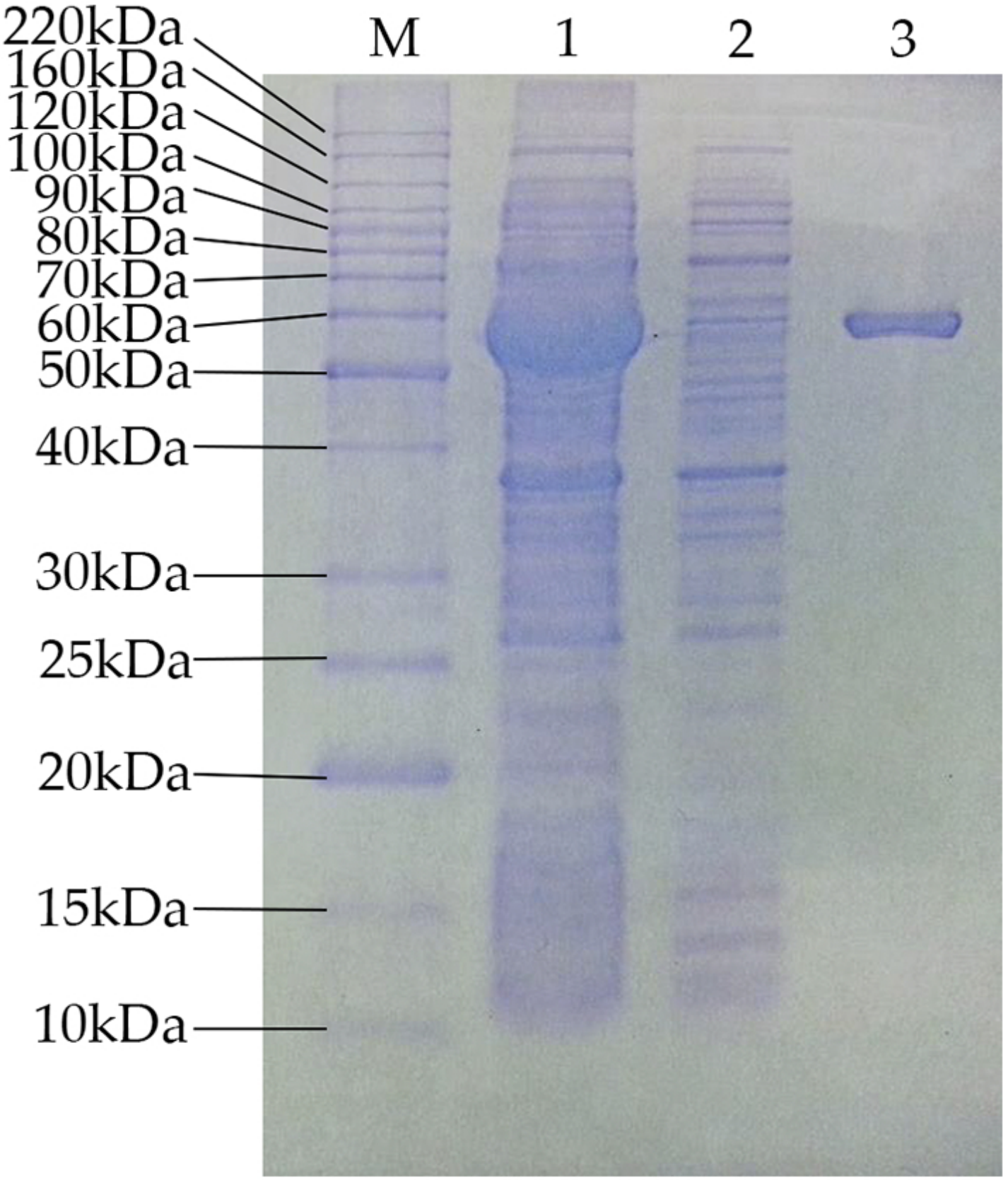
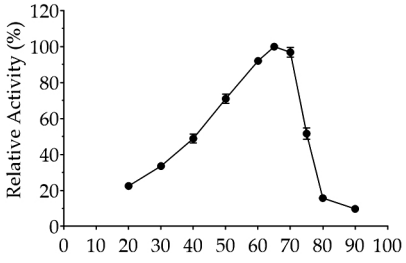
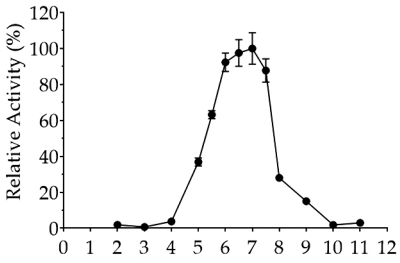

220kDa
160kDa
120kDa
100kDa
90kDa
80kDa
70kDa
60kDa
50kDa
40kDa
30kDa
25kDa
20kDa
15kDa
10kDa
M
1
2
3

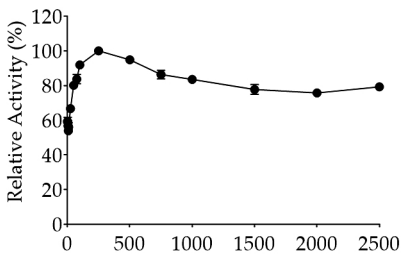
Figure 1



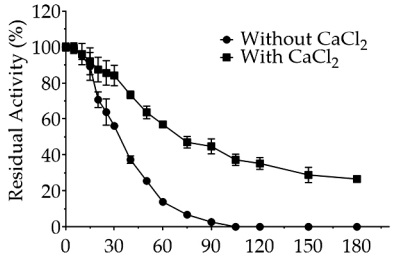

(a)



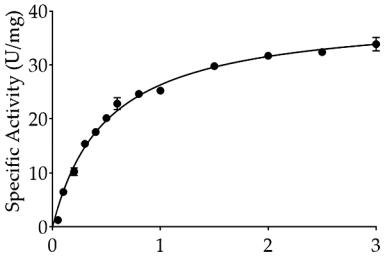

(b)




(c)




(d)




(e)

Figure 2


GH 1 N-terminal signature
G T A T S S F Q I E
motif 1
A V D F L G I N Y Y S I S - Y L
motif 2
I Y V T E N G A A
motif 3
D P E R V Q Y L Q E H L
motif 4
D V D L R G Y F A W S F L D N F E W
motif 5
R F G L V Y V D F - E T Q Q

Figure 3

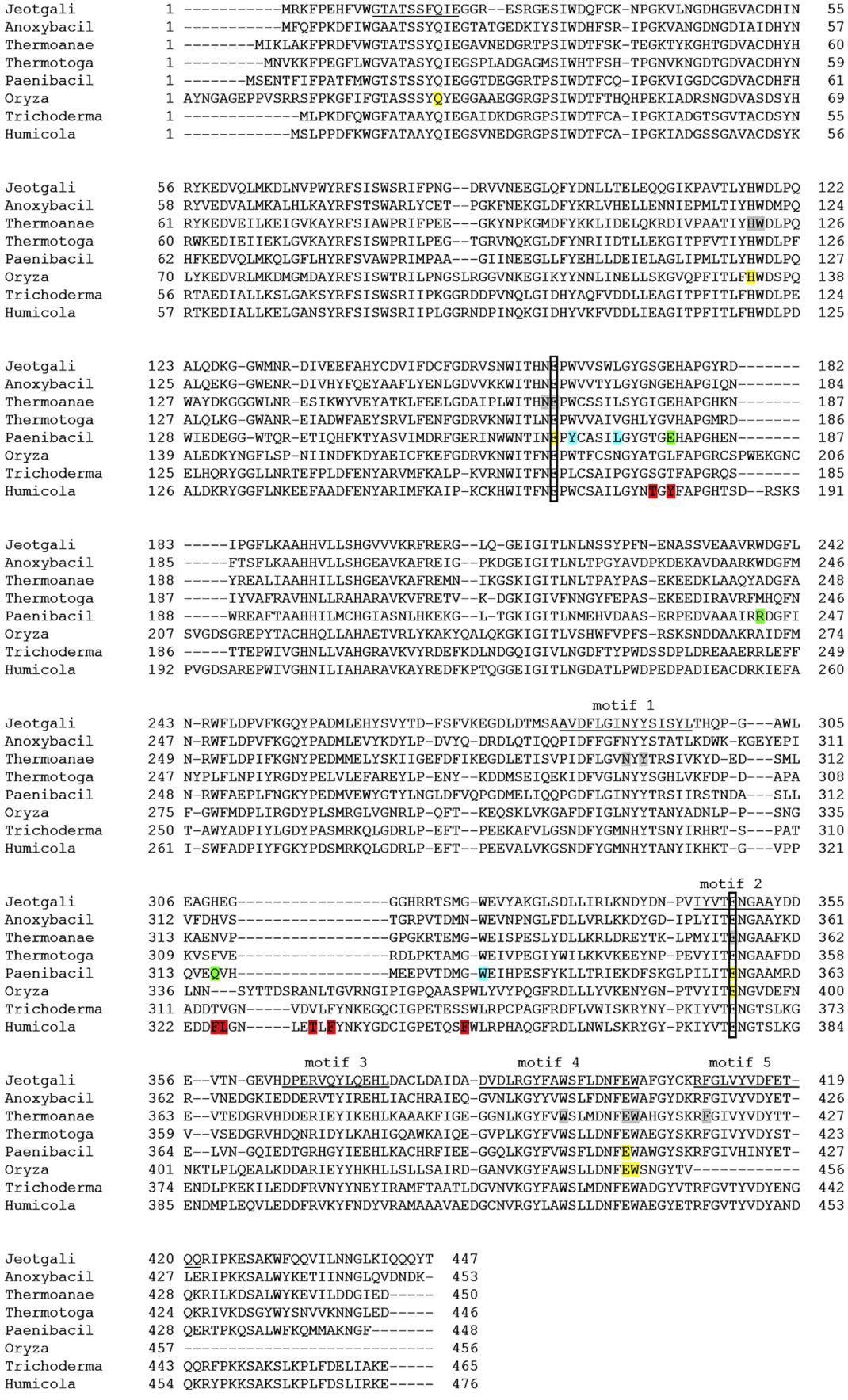

GH1 N-terminal signature
motif 1
motif 2
motif 3
motif 4
motif 5

Figure 4

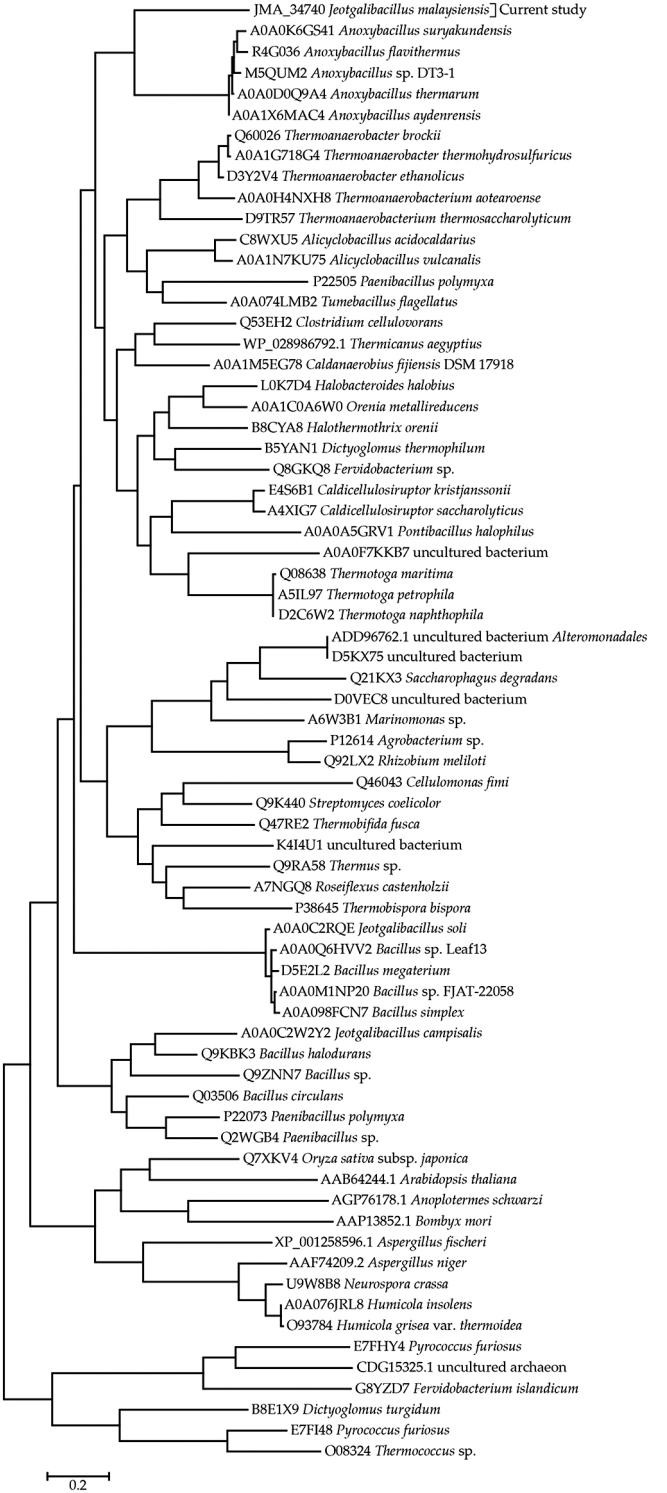

JMA_34740 Jeotgalibacillus malaysiensis] Current study
A0A0K6GS41 Anoxybacillus suryakundensis
R4G036 Anoxybacillus flavithermus
M5QUM2 Anoxybacillus sp. DT3-1
A0A0D0Q9A4 Anoxybacillus thermarum
A0A1X6MAC4 Anoxybacillus aydenrensis
Q60026 Thermoanaerobacter brockii
A0A1G718G4 Thermoanaerobacter thermohydrosulfuricus
D3Y2V4 Thermoanaerobacter ethanolicus
A0A0H4NXH8 Thermoanaerobacterium aotearoense
D9TR57 Thermoanaerobacterium thermosaccharolyticum
C8WXU5 Alicyclobacillus acidocaldarius
A0A1N7KU75 Alicyclobacillus vulcanalis
P22505 Paenibacillus polymyxa
A0A074LMB2 Tumebacillus flagellatus
Q53EH2 Clostridium cellulovorans
WP_028986792.1 Thermicanus aegyptius
A0A1M5EG78 Caldanaerobius fijiensis DSM 17918
L0K7D4 Halobacteroides halobius
A0A1C0A6W0 Orenia metallireducens
B8CYA8 Halothermothrix orenii
B5YAN1 Dictyoglomus thermophilum
Q8GKQ8 Fervidobacterium sp.
E4S6B1 Caldicellulosiruptor kristjanssonii
A4XIG7 Caldicellulosiruptor saccharolyticus
A0A0A5GRV1 Pontibacillus halophilus
A0A0F7KKB7 uncultured bacterium
Q08638 Thermotoga maritima
A5IL97 Thermotoga petrophila
D2C6W2 Thermotoga naphthophila
ADD96762.1 uncultured bacterium Alteromonadales
D5KX75 uncultured bacterium
Q21KX3 Saccharophagus degradans
D0VEC8 uncultured bacterium
A6W3B1 Marinomonas sp.
P12614 Agrobacterium sp.
Q92LX2 Rhizobium meliloti
Q46043 Cellulomonas fimi
Q9K440 Streptomyces coelicolor
Q47RE2 Thermobifida fusca
K4I4U1 uncultured bacterium
Q9RA58 Thermus sp.
A7NGQ8 Roseiflexus castenholzii
P38645 Thermobispora bispora
A0A0C2RQE Jeotgalibacillus soli
A0A0Q6HVV2 Bacillus sp. Leaf13
D5E2L2 Bacillus megaterium
A0A0M1NP20 Bacillus sp. FJAT-22058
A0A098FCN7 Bacillus simplex
A0A0C2W2Y2 Jeotgalibacillus campisalis
Q9KBK3 Bacillus halodurans
Q9ZNN7 Bacillus sp.
Q03506 Bacillus circulans
P22073 Paenibacillus polymyxa
Q2WGB4 Paenibacillus sp.
Q7XKV4 Oryza sativa subsp. japonica
AAB64244.1 Arabidopsis thaliana
AGP76178.1 Anoplotermes schwarzi
AAP13852.1 Bombyx mori
XP_001258596.1 Aspergillus fischeri
AAF74209.2 Aspergillus niger
U9W8B8 Neurospora crassa
A0A076JRL8 Humicola insolens
O93784 Humicola grisea var. thermoidea
E7FHY4 Pyrococcus furiosus
CDG15325.1 uncultured archaeon
G8YZD7 Fervidobacterium islandicum
B8E1X9 Dictyoglomus turgidum
E7FI48 Pyrococcus furiosus
O08324 Thermococcus sp.
0.2

Figure 5