Conference paper

Eukote Suwan, Siriphan Arthornthurasuk and Prachumporn T. Kongsaeree*

A metagenomic approach to discover a novel β-glucosidase from bovine rumens

DOI 10.1515/pac-2016-0924

Abstract: β-Glucosidases play an important role in biomass degradation as they hydrolyze cellobiose to glucose in a final step of cellulolysis. In particular, ruminant animals rely on β-glucosidases from rumen microorganisms for conversion of plant cellulosic materials into glucose. In this study, we are interested in characterization of a novel β-glucosidase from rumen microorganisms. However, most rumen microorganisms are obligate anaerobes, which require special cultivation conditions. Presently, the metagenomic techniques, which enable isolation and characterization of microbial genes directly from environmental samples, have been applied to overcome these problems. In this study, the sequence-based screening approach was successfully applied to identify a novel β-glucosidase gene, *Br2*, from a bovine rumen metagenomic sample. A 1338-bp complete coding sequence of *Br2* encodes a 51-kDa GH1 β-glucosidase of 445 amino acid residues with 59 % sequence identity to a β-glucosidase from *Cellulosilyticum ruminicola* JCM 14822. The recombinantly expressed Br2 exhibited an optimal activity at pH 6.5 and 40 °C, reflecting its rumen bacterial origin, and relatively higher catalytic efficiencies toward glucoside and fucoside substrates than other glycosides, similar to many previously reported bacterial β-glucosidases. Our sequence-based screening approach can be applied to identify other genes of interest from environmental samples.

Keywords: β-glucosidase; bovine; ICS-28; metagenome; rumen.

Introduction

Cellulose is the most abundant biomass on Earth as it is a major component of plant cell wall. It is a long polysaccharide chain of glucose units linked together by β-1,4 glycosidic bonds. Cellulose is now recognized as an alternative source of glucose (rather than starch or sugar cane molasses) which is a starting material for production of bioethanol and bio-based chemicals [1]. Many microorganisms, including bacteria and fungi in soils, protozoa in termite gut, and microbial communities in ruminant digestive tract, produce cellulolytic enzymes, thus enabling them to utilize cellulose as their carbon source. Cellulolytic enzymes can be catagorized into three groups, namely (1) endoglucanases (EC 3.2.1.4), which randomly hydrolyze internal cellulose chains to short chain polysaccharides; (2) exoglucanases, including cellodextrinases (EC 3.2.1.74) and cellobiohydrolases (EC 3.2.1.91), which convert short chain polysaccharides to oligosaccharides and cellobiose;

Article note: A collection of invited papers based on presentations at the XXVIII International Carbohydrate Symposium (ICS-28), New Orleans, July 17–21, 2016.

***Corresponding author: Prachumporn T. Kongsaeree,** Interdisciplinary Graduate Program in Genetic Engineering, Faculty of Graduate School, Kasetsart University, Bangkok 10900, Thailand; Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand; and Center for Advanced Studies in Tropical Natural Resources, NRU-KU, Kasetsart University, Bangkok 10900, Thailand, e-mail: fscippt@ku.ac.th
Eukote Suwan: Interdisciplinary Graduate Program in Genetic Engineering, Faculty of Graduate School, Kasetsart University, Bangkok 10900, Thailand
Siriphan Arthornthurasuk: Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

and (3) β-glucosidases (EC 3.2.1.21), which degrade oligosaccharide and cellobiose from the non-reducing end to release glucose monomer units [2].

There has been a strong interest in β-glucosidases, partly due to its prominent role in glucose production in the final step of cellulolysis. To date, β-glucosidases are classified according to the amino acid sequence similarity into glycoside hydrolase (GH) families 1, 3, 5, 9, 30, 116 and non-classified (CAZy: http://www.cazy.org/) [3]. GH families 1, 5 and 30 are grouped into clan GH-A, which is the largest clan and exhibits a $(\beta/\alpha)_8$ structure. GH9 exhibits an $(\alpha/\alpha)_6$ structure, while GH3 has two domains, which are a $(\beta/\alpha)_8$ structure followed by an α/β sandwich structure. β-Glucosidases perform catalysis via two mechanisms (retaining and inverting) using two catalytic residues (acid/base and nucleophile). They are found and play various important roles in all living organisms, and their applications are diverse [4]. Especially, β-glucosidases can be applied in the degradation of cellulosic biomass to hydrolyze cellobiose to glucose, thus reducing the amount of cellobiose that could inhibit the activity of cellulolytic enzymes [2].

Ruminant animals are a group of herbivore animals, such as cow, buffalo, deer, sheep, etc., which contain four stomach chambers, namely rumen, reticulum, omasum and abomasum. These animals normally feed on plant materials, which are mainly composed of cellulose. Rumen is a site for degradation and fermentation of ingested feed by residential microorganisms. The diversity of rumen microbial eco-system was found to include bacteria (10^{10}–10^{11} cells/mL, including more than 50 genera), ciliate protozoa (10^4–10^6/mL, including 25 genera), anaerobic fungi (10^3–10^5 zoospores/mL, including five genera) and bacteriophages (10^8–10^9/mL) [5]. Glucose from cellulose degradation is converted to pyruvate, which is a precursor for synthesis of volatile fatty acids, including acetic acid, propionic acid and butyric acid, while CO_2 and methane are obtained as by-products. The volatile fatty acids are then absorbed by rumen epithelial cells for use as a source of energy or are converted to other biomolecules [6]. Degradation of cellulosic biomass into glucose is accomplished via activities of numerous cellulolytic enzymes produced by these rumen microorganisms [7, 8]. Therefore, it is of our interest to identify and characterize these rumen microbial cellulolytic enzymes. However, these microorganisms are anaerobes, which complicate the traditional cultivation-based methods. Therefore, a metagenomics approach has been used to access a wealth of genomic information from these unculturable microorganisms [9].

Metagenomics is a culture-independent method, involving genetic material extraction, library construction and screening (based on nucleotide sequences or activities). As such, it has been designed to over-come cultivation limitations, thus enabling utilization of genetic materials from natural resources [9]. This method has been successfully applied to identify many novel enzymes, including β-glucosidases, from various sources, such as soil, marine, hot spring, termite and rumen [10–15]. However, only a few rumen microbial β-glucosidases have been reported and characterized by using this approach [8, 16, 17].

In this study, we were interested in identification and characterization of a novel β-glucosidase from rumen microorganisms by using the sequence-based metagenomic approach. The consensus primers of GH1 β-glucosidases were used to amplify a partial β-glucosidase sequence from the rumen metagenomic libraries, which was further extended by using a genome walking method to cover a full-length β-glucosidase gene, *Br2*. Here we have provided a complete molecular and enzymatic characterization of the recombinantly expressed Br2. In addition, we have demonstrated that our sequence-based approach is applicable to the discovery of many more novel genes without the need of high-throughput function-based screening facility.

Results and discussion

The search results from the National Center for Biotechnology Information database for the β-glucosidase sequences from rumen microorganism revealed 118 GH1 and 186 GH3 sequences from 16 to 20 species, respectively (Supplementary Table 1). β-Glucosidases sequences in GH families 5, 9, 30 and 116 were not found. Most of these sequences belonged to bacteria in Phylum Firmicutes, which were dominated by

Erysipelotrichaceae bacterium (22.22 %) for GH1, and *Butyrivibrio fibrisolvens* (22.04 %) and *Lachnospiraceae bacterium* (22.04 %) for GH3.

Multiple amino acid sequence alignment and phylogenetic tree analysis were performed (Supplementary Fig. 1). In GH1, a total of 111 full-length amino acid sequences shared 22–76 % sequence identities. Conserved domain analysis found that all of GH1 sequences contained domains that are conserved with either the BglB superfamily (β-glucosidase/6-phospho-β-glucosidase/β-galactosidase, GenBank Accession no. COG2723) or the Glyco_hydro_1 super family (glycosyl hydrolase family 1, GenBank Accession no. Cl23725). In GH3, a total of 183 full-length amino acid sequences shared 20–76 % sequence identities, all of which contained a Glyco_hydro_3 superfamily domain (glycosyl hydrolase family 3 N-terminal domain, GenBank Accession no. Cl07971). About 60 % of them showed conservation with the BglX multi-domain family (periplasmic β-glucosidase and related glycosidases; GenBank Accession no. COG1472), with (40 %) or without (20 %) the presence of the Glyco_hydro_3_C multi-domain family (glycosyl hydrolase family 3 C-terminal domain, GenBank Accession no. pfam01915) and the Fn3-like domain (fibronectin type III-like domain, GenBank Accession no. pfam14310). The other 26 % of the GH3 sequences contained the Fn3-like domain, and either the PRK15098 domain (β-D-glucoside glucohydrolase, GenBank accession no. PRK15098) or the PLN03080 domain (probable β-xylosidase, GenBank accession no. PLN03080).

These full-length amino acid sequences of GH1 and GH3 were used to compute their physical and chemical parameters. The length, molecular weight and theoretical pI values of GH1 β-glucosidases were between 417 and 503 residues, 48–58 kDa and pI of 4.8–6, respectively. On the other hand, GH3 β-glucosidases were separately computed according to their conserved domain groups. For the first group, their length, molecular weight and theoretical pI values were between 341 and 823 residues, 36–92 kDa and pI of 4.2–9.5, respectively. For the second GH3 group, their parameters are 635–1067 residues, 71–116 kDa and pI of 4.1–8.6, respectively.

The degenerate primers designed from the conserved regions of the GH1 enzymes were used to screen for a novel β-glucosidase sequence from metagenomic DNA samples extracted from bovine rumen fluid. DNA sequencing results of the 873-bp PCR product revealed a partial sequence of a β-glucosidase gene, *Br1* (Supplementary Fig. 2), with 56 % amino acid sequence identity to a β-glucosidase from *Cellulosilyticum ruminicola* JCM 14822 (GenBank Accession no. ACZ98664) (Supplementary Fig. 3), which is an anaerobic,

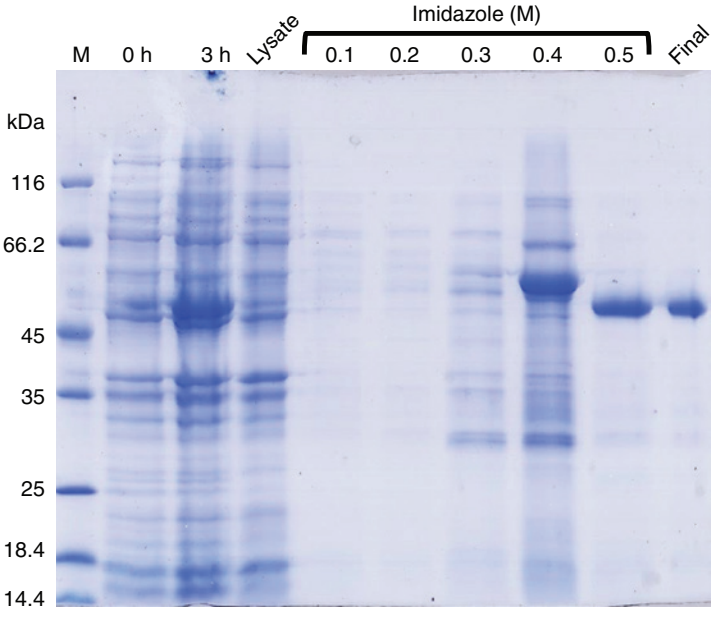

Fig. 1: Ten percent SDS-PAGE analysis of the purification of the recombinant Br2. The protein samples are indicated above the lanes. Lane M is protein size markers. Br2 was eluted from the Ni^{2+}-NTA column using a stepwise gradient of 0.1–0.5 M of imidazole as indicated.

gram-negative, mesophilic, cellulolytic bacterium previous isolated from yak rumen [18]. The deduced amino acid sequence of Br1 contained the conserved motifs unique to GH1 β-glucosidases, NEP and I/VTENG, with the glutamate residues acting as an acid/base catalyst and a nucleophile, respectively. So, it is likely that the sequence of *Br1* was part of a gene for a novel β-glucosidase from rumen microorganism. So, the genome walking method was applied to obtain the complete coding sequence of this novel β-glucosidase, *Br2*, which was 1,338 bp long (GenBank Accession no. KU166869). After translation to protein, the sequence of Br2 was 445 residues long, with an expected size of 51 kDa and a theoretical pI of 5.6, but without a signal peptide for extracellular secretion. Br2 is a GH1 β-glucosidase with 59 % sequence identity to the same *C. ruminicola* β-glucosidase, as well as 48–56 % sequence identity to a number of GH1 β-glucosidases from anaerobic bacteria commonly found in mammalian gastrointestinal tracts (Supplementary Fig. 3).

Subsequently, the sequence of Br2 was cloned into pET15b expression vector in frame with the coding sequence of the N-terminal poly-histidine tag. Recombinant expression of Br2 yielded a single protein band of about 53 kDa. While the majority of Br2 was present in the insoluble fraction (Supplementary Fig. 4), about 100 units of β-glucosidase activity were detected in the cell lysate (from 1 L culture). Br2 was then purified from cell lysate to homogeneity via immobilized metal affinity chromatography (Fig. 1), which could be detected by a polyclonal antibody against a polyhistidine tag (Fig. 2). The final yield of Br2 was 1 mg and 96 units from 1 L culture, with about a 140-fold increase in purity.

The optimal temperature and the optimal pH for Br2 was found to be 40 °C and pH 6.5, respectively (Fig. 3a and b), which are consistent with the optimal growth conditions for rumen bacteria and for *C. ruminicola* [5, 19]. Br2 showed thermal and pH stabilities between 5 and 35 °C and pH 6.0–9.0, respectively, with at least 70 % activity remaining after incubation at these temperatures or pHs for 30 min or 16 h, respectively (Fig. 3a and b). At 70 °C and above, Br2 was completely denatured and lost all activity. The time-course hydrolysis of cellotetraose by Br2, as monitored by TLC, revealed glucose and cellotriose as initial hydrolysis products (Fig. 4), indicating its exo-acting action, which releases only a single sugar residue from the non-reducing end of the substrate, in agreement with its classification as a GH1 enzyme.

Kinetic analysis of Br2 was performed at 40 °C and pH 6.5, which was its optimum reaction condition. The reported k_{cat} values were calculated per subunit of 53 kDa (Table 1). Br2 showed high catalytic efficiencies (k_{cat}/K_m) toward *p*-nitrophenyl-β-D-glucoside and *p*-nitrophenyl-β-D-fucoside substrates, but relatively low efficiencies (≤2 % relative to that of *p*-nitrophenyl-β-D-glucoside) toward most other aryl glucoside substrates, similar to the kinetic properties reported for many bacterial GH1 β-glucosidases [20–22]. Interestingly,

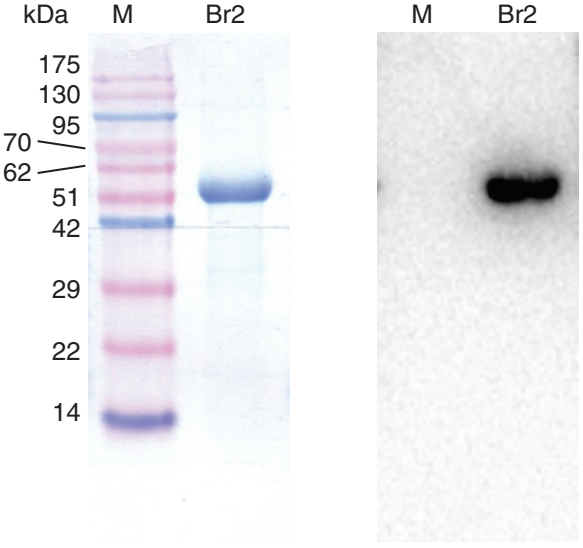

Fig. 2: Ten percent SDS-PAGE analysis (left panel) and western blot (right panel) of the purified recombinant Br2. Lane M is protein size markers.

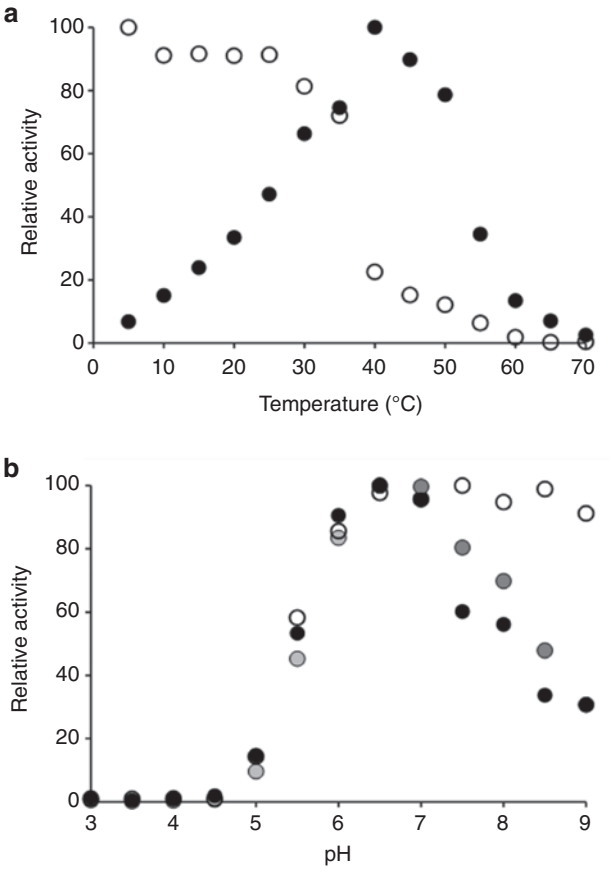

Fig. 3: The effects of temperature (a) and pH (b) on enzymatic properties of Br2. The optimal temperature and pH are shown as filled circles, while the temperature and pH stabilities are shown as open circles. The buffers (0.1 M) used in the determination of the optimal pH were sodium acetate, pH 3.0–6.5 (light grey), sodium phosphate, pH 6.5–9.0 (dark grey), and McIlvaine buffer, pH 3.0–9.0 (black).

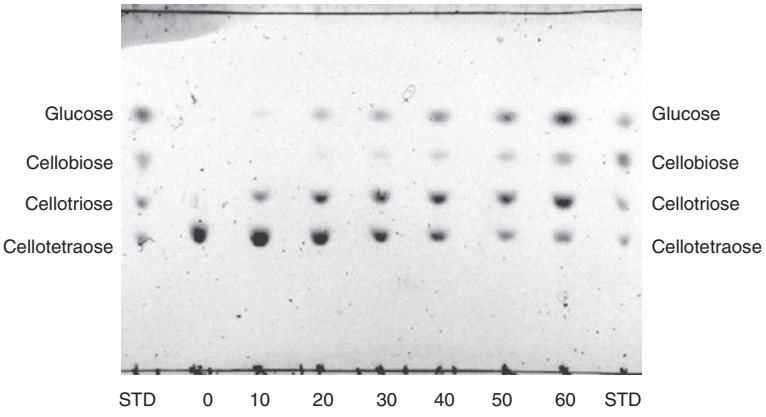

Fig. 4: TLC analysis of the hydrolysis of cellotetraose by Br2. STD are 10 nmol of cellooligosaccharide standards. The incubation times (in min) between cellotetraose and Br2 are indicated below the lanes.

Br2 exhibited a high k_{cat} value toward xylobiose, but its K_m value was also high, suggesting its low affinity. However, the efficiency toward cellobiose was very low (about 0.1 % relative to that of *p*-nitrophenyl-β-D-glucoside), due to its low k_{cat} and moderately high K_m values, suggesting that Br2 was probably not involved in

Table 1: Kinetic parameters of Br2.

Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})	Relative k_{cat}/K_m (%)
p-Nitrophenyl-β-D-glucoside[a]	0.552 ± 0.037	88.5 ± 1.2	160 ± 11	100
p-Nitrophenyl-β-D-fucoside[a]	0.319 ± 0.026	74.7 ± 1.4	234 ± 20	146
p-Nitrophenyl-α-L-arabinoside[a]	0.712 ± 0.067	2.15 ± 0.06	3.02 ± 0.29	1.9
p-Nitrophenyl-β-D-xyloside[a]	1.83 ± 0.17	2.13 ± 0.06	1.16 ± 0.11	0.7
p-Nitrophenyl-β-D-galactoside[b]	6.81 ± 1.24	88.9 ± 16.1	13.1 ± 3.3	8.2
Cellobiose[b]	24.3 ± 9.8	3.14 ± 1.26	0.129 ± 0.074	0.1
Xylobiose[b]	118 ± 49	294 ± 121	2.49 ± 1.45	1.6

[a]Data obtained from the Michaelis–Menten equation; [b]data obtained from the Lineweaver–Burk equation.

cellulolysis, consistent with its lack of signal peptide for extracellular secretion. In comparison, four recombinantly expressed GH3 β-glucosidases cloned from bovine rumen metagnome, showed about 0.1–2% kinetic efficiencies toward cellobiose relative to that of *p*-nitrophenyl-β-D-glucoside. Nonetheless, the most active β-glucosidase could complement saccharification activity of the commercial cellulase [16].

A previous study on the fibrolytic activities of *C. ruminicola* when grown on various substrates found that *C. ruminicola* expressed a range of both cellulytic and hemicellulytic enzymes, especially endoglucanase (52–1279 mU/mg), xylanase (568–7000 mU/mg) and mannanase (40–510 mU/mg), indicating its role in feed digestion in the yak rumen. However, β-glucosidase activities were relatively minor, being higher in filter paper culture (44.2 mU/mg) than in corn cob (1.4 mU/mg) and in xylan (0.4 mU/mg) [18]. Unfortunately, the β-glucosidase in *C. ruminicola* was not purified and characterized, so comparison with the activities of our Br2 was not possible.

In conclusion, we have successfully applied the sequence-based metagenomic approach to discover a novel β-glucosidase gene, *Br2*, from bovine rumen. The complete coding sequence of Br2 was delineated and its enzymatic properties characterized. So, the data presented here gave comprehensive information on the molecular and enzymatic properties of a newly described β-glucosidase from bovine rumen metegenome. Furthermore, the sequence-based screening approach used in this study provides an alternative to the function-based screening of a large metagenomics library, which usually requires an expensive high-throughput facility. So, our approach can be applied to identify other novel genes from metagenome samples, depending on the properties of the amino acid sequences used in multiple sequence alignment for the design of the degenerate primers.

Experimental section

Metagenome DNA extraction

Three types of rumen samples, namely fluid, content and endothelium, were collected from cow rumens at a slaughterhouse in Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom, Thailand. The metagenomic DNA was extracted from the rumen samples as described previously [23]. The extracted metagenomics DNA was verified by using agarose gel electrophoresis and stored at –20 °C until use.

Amplification of the rumen metagenomic DNA

The amino acid sequences of rumen microorganism β-glucosidases were searched from the National Center for Biotechnology Information database during February–March, 2015. All amino acid sequences of rumen microorganism β-glucosidases were classified according to their conserved domains by using CD-Search, which uses a search algorithm to compare a query protein sequence against the NCBI's conserved domain

database (available on: http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [24]. The full-length amino acid sequences of each GH were aligned by using Clustal Omega, which is based on seeded guide trees and profile hidden Markov model methods, and the phylogenic tree was generated by using the internal neighbor joining algorithm and the default parameters (available on: http://www.ebi.ac.uk/Tools/msa/clustalo/) [25, 26]. The length, molecular weight and theoretical pI of the amino acid sequences were computed using ProtParam, which is a primary structure analysis based on the properties of amino acids that compose the given protein sequences (available on: http://web.expasy.org/protparam/) [27]. The conserved nucleotide regions were verified according to conserved amino acid regions of each GH family. The degenerate primers were designed according to conserved nucleotide regions.

The partial fragment of a β-glucosidase gene from rumen microorganisms, *Br1*, was obtained by amplification of the rumen metagenomic DNA samples by using the degenerate primers (Forward primer: 5′ ACD MTK TAY CAY TGG GAY CTN CCV CA 3′; Reverse primer: 5′ CCA YTC RAA RTT RTC VAD VAR NGA CCA 3′). The amplification reactions were made up of 0.4 μM of forward and reverse primers, 0.24 mM dNTP, 1.5 mM $MgCl_2$, 1.25 U Maximo *Taq* DNA polymerase with buffer (GeneON, Ludwigshafen, Germany), and 100 ng of rumen metagenomic DNA template. PCR was performed at 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 55 °C for 30 s, 72 °C for 3 min, and 72 °C for 10 min. The PCR product was verified by 1 % agarose gel electrophoresis, and purified, before ligation with pGEM®-T Easy vector (Promega, Madison, WI, USA), and transformation into *E. coli* DH5α on LB-agar supplemented with 50 μg/mL ampicillin. The clones containing an insert of an expected size were subjected to DNA sequencing (Macrogen, Seoul, South Korea). The translation of the *Br1* nucleotide sequence into the protein sequence was performed by using FGENESB, which is a suite of programs, based on Markov chain models of coding regions, and translation and termination sites, for annotation of bacterial genomes and bacterial community sequences (available on: http://linux1.softberry.com/) [28]. The nucleotide and translated amino acid sequences of *Br1* were compared with β-glucosidase sequences in GenBank by using BLAST, which is based on alignment of regions of local similarity between sequences and calculates the statistical significance of matches (available on: https://blast.ncbi.nlm.nih.gov/Blast.cgi) [29]. The conserved domains were predicted by using CD-Search program.

The complete coding sequence of β-glucosidase gene from rumen microorganism, *Br2*, was obtained by performing genome walking using the Universal GenomeWalker™ 2.0 kit (Clontech Laboratories, Mountain View, CA, USA) according to manufacturer's instruction. The *Br1* sequence was used to design two pairs of gene specific primers (GSP1_forward: 5′ GCA ATT GCA GAT GGT GTG GAT GTA GAA GGT 3′; GSP2_forward: 5′ ATG CCT GGT CCC TGT TGG ATA ATT TTG AGT 3′; GSP1_reverse: 5′ TGA TAA ATA CCC ACA GCC AAC AAA AAC CTG 3′; GSP2_reverse: 5′ TTT GAG ATG AAG GGC ATG TGG CAG GTC CCA 3′). Briefly, the rumen metagenomic DNA samples were digested with four different blunt-end restriction enzymes, namely *Dra*I, *Eco*RV, *Pvu*II and *Stu*I. The digested metagenomic DNA was ligated with the genome walker adaptor to construct four genome walker libraries. These libraries were used as template to perform primary PCR with adaptor primer 1 and GSP1. The primary PCR products was used as template for secondary (nested) PCR with adaptor primer 2 and GSP2. The positive bands of secondary PCR were subjected to DNA sequencing. The sequencing results were assembled using CAP3, which uses base quality values to compute overlapped reads, construct multiple sequence alignments, and generate consensus sequences (available on: http://doua.prabi.fr/software/cap3) [30].

Expression and purification

The full-length *Br2* gene was amplified from the rumen metagenomic DNA samples by using a pair of 5′ and 3′ gene specific primers (5′GSP: 5′ GAG GAT TCT CAT ATG GGG TTT CCA AAA G 3′; 3′GSP: 5′ CAG TAC TGG ATC CTA CAA CAG ATC ACC G 3′), containing the *Nde*I and *Bam*HI sites (underlined), respectively. The PCR composition and condition were as described above. The *Br2* PCR product was then inserted into pET15b expression vector at the *Nde*I-*Bam*HI sites, and transformed into *E. coli* DH5α on LB-agar supplemented with 50 μg/mL ampicillin. The nucleotide sequence of *Br2* in pET15b was confirmed by DNA sequencing, before

transformation of the recombinant plasmid into *E. coli* BL21(DE3) on LB-agar supplemented with 50 μg/mL ampicillin.

A single colony of *E. coli* BL21(DE3) containing the *Br2* sequence was cultured in LB broth with 100 μg/mL ampicillin at 37 °C, 200 rpm, overnight. Then, 1/20 volume of starter culture was inoculated into 2TY containing 100 μg/mL ampicillin, and grown at 37 °C, 200 rpm, until OD_{600} reached 0.6. The protein expression was induced by adding 0.3 mM IPTG and further incubated at 37 °C for 3 h. The cells were harvested by centrifugation, re-suspended in ENZhance Lysis Buffer (BIOTEC, Pathum Thani, Thailand) and lysed by sonication. The supernatant was collected by centrifugation at 10 000 rpm for 30 min at 4 °C. Br2 was purified from total cell lysate by using Ni^{2+}-NTA affinity chromatography (GE Healthcare, Uppsala, Sweden) and eluted with stepwise gradient of 0.1–0.5 M of imidazole in 50 mM sodium phosphate, 500 mM NaCl, pH 8.0. The protein samples were analyzed by 10 % sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Western blot analysis and chemiluminescent detection was done with Pierce 6x-His epitope tag antibody (Thermo Scientific, Rockford, IL, USA), horseradish peroxidase-conjugated polyclonal rabbit anti-mouse immunoglobulins (Dako, Glostrup, Denmark) and Clarity Western ECL substrate (Bio-Rad, Hercules, CA, USA). Protein concentration was determined by using Bio-Rad Protein Assay Reagent Kit (Bio-Rad, Hercules, CA, USA) and compared to a standard curve of bovine serum albumin. The β-glucosidase activity of Br2 during purification procedures was assayed against 1 mM *p*-nitrophenyl-β-D-glucoside in 0.1 M sodium phosphate, pH 6.5, at 40 °C for 30 min. The reaction was stopped by 2 M sodium carbonate. The absorbance of *p*-nitrophenol product was measured at 400 nm, and compared with the *p*-nitrophenol standard curve.

Enzyme characterization

The β-glucosidase activity of the purified Br2 was assayed against 15 mM *p*-nitrophenyl-β-D-glucoside in 0.1 M sodium phosphate, pH 6.5, at 40 °C for 5 min. The reaction was stopped by 2 M sodium carbonate. The absorbance of *p*-nitrophenol product was measured at 405 nm, and compared with the *p*-nitrophenol standard curve. The effects of temperature and pH on enzymatic properties of Br2 were determined as follows. The optimal temperature was tested in 0.1 M sodium phosphate, pH 6.5, for 5 min, at 5–70 °C. The optimal pH was tested in 0.1 M sodium acetate, pH 3.0–6.5, 0.1 M sodium phosphate, pH 6.5–9.0, and 0.1 M McIlvaine buffer, pH 3.0–9.0, for 5 min, at its optimal temperature. The temperature stability was performed by pre-incubating Br2 at 5–70 °C for 30 min, while the pH stability was performed by pre-incubating Br2 in 0.1 M McIlvaine buffer, pH 3.0–9.0, at 4 °C for 16 h, before measuring the enzyme activity at its optimal temperature and pH.

To determine whether Br2 exhibited an endo- or exo-type of activity, 1.5 μg of the purified Br2 was incubated with 10 mM cellotetraose in 20 mM sodium phosphate, pH 6.5, at 40 °C for 60 min. At different time intervals, 4 μL of the reaction was taken and was spotted on the TLC plate (Silica gel 60 F254, Merck, Darmstadt, Germany), which was developed twice in n-butanol:acetic acid:water, 2:1:1 by volume. The spots were visualized by soaking in 20 % sulfuric acid in ethanol, and heating on the hotplate until the spots were seen [31].

Kinetic parameters of the purified Br2

The purified Br2 was assayed against various concentrations of a range of *p*-nitrophenyl glycoside substrates, as well as cellobiose and xylobiose, in 0.1 M sodium phosphate, pH 6.5, at 40 °C for 5 min, under normal aerobic conditions. The reactions with *p*-nitrophenyl glycosides were stopped by adding 2 M sodium carbonate, and the absorbance of *p*-nitrophenol product was measured at 400 nm, and compared with the *p*-nitrophenol standard curve. The reactions with cellobiose were stopped by boiling for 5 min, and the released glucose was reacted with a glucose oxidase/peroxidase reagent (Sigma-Aldrich, St. Louis, MO, USA)

at 37 °C for 15 min. The amount of glucose was determined by reading the absorbance at 405 nm, and comparing it to a standard curve of glucose. The reactions with xylobiose were stopped by reacting with 1 % dinitrosalicylic acid, then heating to 95 °C for 5 min and cooling on ice. The amount of xylose was determined by reading the absorbance at 540 nm, and comparing it to a standard curve of xylose. Kinetic parameters were calculated from the Michaelis–Menten equation using KaleidaGraph program (Synergy Software). For substrates with large K_m values (compared with substrate concentrations), saturation of the initial velocity could not be observed such that the Michaelis–Menten parameters could not be determined accurately. In these cases, the Lineweaver–Burk plots were used instead, in order to obtain better estimates of the K_m and V_{max} values.

Acknowledgements: This project was financially supported by Kasetsart University Research and Development Institute (V-T(D) 13.55), the Higher Education Research Promotion and National Research University Project of Thailand, and the Graduate School, Kasetsart University. E.S. is a recipient of the PhD scholarship from the Graduate School, Kasetsart University.

References

[1] W. Vermerris, A. Abril. *Curr. Opin. Biotech.* **32**, 104 (2015).
[2] L. R. Lynd, P. J. Weimer, W. H. van Zyl, I. S. Pretorius. *Microbiol. Mol. Biol. Rev.* **66**, 506 (2002).
[3] V. Lombard, H. G. Ramulu, E. Drula, P. M. Coutinho, B. Henrissat. *Nucl. Acids Res.* **42**, D490 (2014).
[4] J. R. Ketudat Cairns, A. Esen. *Cell. Mol. Life Sci.* **67**, 3389 (2010).
[5] D. N. Kamra. *Curr. Sci.* **89**, 124 (2005).
[6] R. L. Baldwin, M. J. Allison. *J. Anim. Sci.* **57**, 461 (1983).
[7] P. J. Weimer. *Crit. Rev. Biotechnol.* **12**, 189 (1992).
[8] M. Hess, A. Sczyrba, R. Egan, T.-W. Kim, H. Chokhawala, G. Schroth, S. Luo, D. S. Clark, F. Chen, T. Zhang, R. I. Mackie, L. A. Pennacchio, S. G. Tringe, A. Visel, T. Woyke, Z. Wang, E. M. Rubin. *Science* **331**, 463 (2011).
[9] D. Cowan, Q. Meyer, W. Stafford, S. Muyanga, R. Cameron, P. Wittwer. *Trends Biotechnol.* **23**, 321 (2005).
[10] C. Jiang, S.-X. Li, F.-F. Luo, K. Jin, Q. Wang, Z.-Y. Hao, L.-L. Wu, G.-C. Zhao, G.-F. Ma, P.-H. Shen, X.-L. Tang, B. Wu. *Bioresource Technol.* **102**, 3272 (2011).
[11] G. Li, Y. Jiang, X.-J. Fan, Y.-H. Liu. *Bioresource Technol.* **123**, 15 (2012).
[12] M. O. A. Sommer, G. M. Church, G. Dantas. *Mol. Syst. Biol.* **6**, 360 (2010).
[13] S.-J. Kim, C.-M. Lee, M.-Y. Kim, Y.-S. Yeo, S.-H. Yoon, H.-C. Kang, B.-S. Koo. *J. Microbiol. Biotechnol.* **17**, 905 (2007).
[14] Z. Fang, W. Fang, J. Liu, Y. Hong, H. Peng, X. Zhang, B. Sun, Y. Xiao. *J. Microbiol. Biotechnol.* **20**, 1351 (2010).
[15] J. C. Bergmann, O. Y. A. Costa, J. M. Gladden, S. Singer, R. Heins, P. D'haeseleer, B. A. Simmons, B. F. Quirino. *FEMS Microbiol. Lett.* **351**, 147 (2014).
[16] M. V. Del Pozo, L. Fernández-Arrojo, J. Gil-Martínez, A. Montesinos, T. N. Chernikova, T. Y. Nechitaylo, A. Waliszek, M. Tortajada, A. Rojas, S. A. Huws, O. V. Golyshina, C. J. Newbold, J. Polaina, M. Ferrer, P. N. Golyshin. *Biotechnol. Biofuels* **5**, 73 (2012).
[17] L. Bao, Q. Huang, L. Chang, Q. Sun, J. Zhou, H. Lu. *Appl. Biochem. Biotechnol.* **166**, 72 (2012).
[18] S. Cai, J. Li, F. Z. Hu, K. Zhang, Y. Luo, B. Janto, R. Boissy, G. Ehrlich, X. Dong. *Appl. Environ. Microbiol.* **76**, 3818 (2010).
[19] S. Cai, X. Dong. *Int. J. Syst. Evol. Microbiol.* **60**, 845 (2010).
[20] K. M. Harada, K. Tanaka, Y. Fukuda, W. Hashimoto, K. Murata. *Arch. Microbiol.* **184**, 215 (2005).
[21] J. B. Kempton, S. G. Withers. *Biochemistry* **31**, 9961 (1992).
[22] W. Hashimoto, H. Miki, H. Nankai, N. Sato, S. Kawai, K. Murata. *Arch. Biochem. Biophys.* **360**, 1 (1998).
[23] J. Zhou, M. A. Bruns, J. M. Tiedje. *Appl. Environ. Microbiol.* **62**, 316 (1996).
[24] A. Marchler-Bauer, M. K. Derbyshire, N. R. Gonzales, S. Lu, F. Chitsaz, L. Y. Geer, R. C. Geer, J. He, M. Gwadz, D. I. Hurwitz, C. J. Lanczycki, F. Lu, G. H. Marchler, J. S. Song, N. Thanki, Z. Wang, R. A. Yamashita, D. Zhang, C. Zheng, S. H. Bryant. *Nucl. Acids Res.* **43**, D222 (2015).
[25] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, R. Lopez. *Nucl. Acids Res.* **41**, W597 (2013).
[26] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins. *Mol. Syst. Biol.* **7**, 539 (2011).
[27] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch. "Protein identification and analysis tools on the ExPASy server", in *The Proteomics Protocols Handbook*, chapter 52, J. M. Walker (Ed.), pp. 571–607, Humana Press, New York (2005).

[28] V. Solovyev, A. Salamov. "Automatic annotation of microbial genomes and metagenomic sequences", in *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*, chapter 4, R. W. Li (Ed.), pp. 61–78, Nova Science Publishers, New York (2011).
[29] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman. *J. Mol. Biol.* **215**, 403 (1990).
[30] X. Huang, A. Madan. *Genome Res.* **9**, 868 (1999).
[31] K. Choengpanya, S. Arthornthurasuk, P. Wattana-amorn, W. T. Huang, W. Plengmueankae, Y.-K. Li, P. T. Kongsaeree. *Protein Expres. Purif.* **115**, 132 (2015).

Supplemental Material: The online version of this article (DOI: 10.1515/pac-2016-0924) offers supplementary material, available to authorized users.