Structural and mutational analysis of glycoside hydrolase family 1 Br2 β-glucosidase derived from bovine rumen metagenome

Wilaiwan Kaenying [a], Takayoshi Tagami [b], Eukote Suwan [c], Chariwat Pitsanuwong [d], Sinchai Chomngam [e], Masayuki Okuyama [b], Palangpon Kongsaeree [e], Atsuo Kimura [b], Prachumporn T. Kongsaeree [a,*]

[a] *Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*
[b] *Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*
[c] *Department of Veterinary Technology, Faculty of Veterinary Technology, Kasetsart University, Bangkok 10900, Thailand*
[d] *Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok 10300, Thailand*
[e] *Department of Chemistry and Center for Excellence in Protein and Enzyme Technology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand*

ARTICLE INFO

ABSTRACT

Ruminant animals rely on the activities of β-glucosidases from residential microbes to convert feed fibers into glucose for further metabolic uses. In this report, we determined the structures of Br2, which is a glycoside hydrolase family 1 β-glucosidase from the bovine rumen metagenome. Br2 folds into a classical $(\beta/\alpha)_8$-TIM barrel domain but displays unique structural features at loop $\beta5{\rightarrow}\alpha5$ and α-helix 5, resulting in different positive subsites from those of other GH1 enzymes. Br2 exhibited the highest specificity toward laminaritriose, suggesting its involvement in β-glucan hydrolysis in digested feed. We then substituted the residues at subsites +1 and + 2 of Br2 with those of *Halothermothrix orenii* β-glucosidase. The C170E and C221T mutations provided favorable interactions with glucooligosaccharide substrates at subsite +2, while the A219N mutation probably improved the substrate preference for cellobiose and gentiobiose relative to laminaribiose at subsite +1. The N407Y mutation increased the affinity toward cellooligosaccharides. These results give further insights into the molecular determinants responsible for substrate specificity in GH1 β-glucosidases and may provide a basis for future enzyme engineering applications.

1. Introduction

Glycoside hydrolase family 1 (GH1) consists of a group of enzymes that hydrolyze the glycosidic linkage of disaccharides or glycosides, such as β-glucosidases (E.C. 3.2.1.21), β-galactosidases (E.C. 3.2.1.23), β-mannosidase (E.C. 3.2.1.25), β-glucuronidase (E.C. 3.2.1.31), β-D-fucosidase (E.C. 3.2.1.38), 6-phospho-β-galactosidase (E.C. 3.2.1.85), 6-phospho-β-glucosidase (E.C. 3.2.1.86), and myrosinase (E.C. 3.2.1.147) [1]. Among these enzymes, β-glucosidases play an essential role in cellulolysis as they hydrolyze cellobiose to glucose; in addition, β-glucosidases participate in the hydrolysis of various glucosides involved in physiological functions, such as plant−microbe and plant−insect interactions, hormone activation, defense mechanisms, degradation of glycolipids and exogenous

glucosides, cell wall catabolism, and lignification [2]. β-Glucosidases are present in all organisms, including the residential microbes in the bovine rumen [3]. Since the structure of the cyanogenic β-glucosidase from *Trifolium repens* was first elucidated in 1995 (PDB code 1CBG) [4], all members of GH1 have been found to exhibit a common $(\beta/\alpha)_8$-barrel fold. The catalytic acid/base and nucleophile in GH1 enzymes are the highly conserved glutamate residues within the NEP and YITENG motifs, respectively, located approximately 5.5 Å apart [5].

Metagenomics is an important molecular biology tool used to investigate the metagenome extracted from specified samples, which bypasses conventional isolation and cultivation. Through this technique, we can study and utilize novel enzymes from microbes, which are otherwise unculturable [6]. Most metagenomic β-glucosidases with known structures are GH1 enzymes from environmental samples, such as Td2F2 from a compost metagenome and MeBglD2 from a soil metagenomic library [7,8]. Most reported β-glucosidases from rumen metagenomes belong to the GH3 family [9–11], but only the structures of a GH3 broad-specificity β-glucosidase/β-xylosidase from cow rumen are available [12]. On the other hand, only two GH1 enzymes from the ruminal metagenome have been characterized thus far, which are Br2 and PersiBGL1 from bovine and sheep rumens, respectively [3,13], but their three-dimensional structures have yet to be reported.

Ruminal animals rely on the activity of cellulolytic enzymes from ruminal microorganisms for the degradation of cellulosic materials into glucose. These include endo-β-1,4-glucanases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91), and β-glucosidases [11]. We previously employed a metagenomic approach to identify a 1338-bp gene from a bovine rumen sample, which encodes a novel 445-amino acid GH1 Br2 β-glucosidase [3]. The residues Glu163 and Glu350 were identified as the acid/base catalyst and nucleophile, respectively, by amino acid sequence alignment with other GH1 β-glucosidases (Supplementary Fig. S1). The recombinant Br2 showed an apparent subunit molecular weight of 53 kDa, and optimal activity at pH 6.5 and 40 °C. Interestingly, Br2 lacks the signal peptide for extracellular secretion and exhibited very low efficiency toward cellobiose, suggesting that Br2 was probably not involved in cellulolysis as we initially assumed [3]. Therefore, in this study, we aimed to further characterize Br2 by determining its three-dimensional structure and the amino acid residues responsible for substrate specificities toward various glucooligosaccharides. This report provides better knowledge on the structure–function relationships in Br2 that may assist further protein engineering efforts to modulate the enzyme properties.

2. Materials and methods

2.1. Construction and production of Br2 mutants

The catalytic residues (Glu163 and Glu350) and those likely involved in substrate binding in Br2 (GenBank accession number APM84071) were targeted for site-directed mutagenesis. PCR-based site-directed mutagenesis reactions [14] were performed using sense and antisense primers (Supplementary Table S1) and the pET15b-Br2 plasmid as a template [3]. The sequences of Br2 mutants were confirmed by DNA sequencing before transformation into *E. coli* BL21(DE3) by electroporation. The transformants were grown on LB agar containing 50 µg mL^{-1} ampicillin.

All wild-type and mutant forms of Br2 were expressed and purified from their respective *E. coli* BL21(DE3) clones as described previously [3].

2.2. Crystallization and structure determination

The screening solutions for protein crystallization were prepared according to Hampton Research Crystal Screen (HR2-110; Aliso Viejo, CA, USA). The wild-type and mutant enzymes (10 mg/mL) were crystallized by the hanging drop vapor diffusion method at 4 °C for ∼ 1 week. Crystals were obtained in 0.1 M sodium acetate trihydrate, pH 7.4, 1.6 M ammonium sulfate. The crystals were flash-cooled after soaking in crystallization conditions containing 20 % glycerol and subjected to X-ray crystallography at the SPXF beamline BL13C1 of the National Synchrotron Radiation Research Center (Hsinchu, Taiwan) at a wavelength of 0.975 Å. Datasets were processed with the XDS package [15]. The structure of the E350G mutant was solved by the molecular replacement method with AutoMR in the PHENIX suite [16] using the structure of *Tm*GH1 from *Thermotoga maritima* (PDB code 1W3J; 49 % sequence identity to Br2) [17] as a search model. The structures of wild-type Br2 and the E163Q mutant were determined by the molecular replacement method as described above using the structure of the E350G mutant as a search model. Interactive model correction and refinement were performed with several cycles of Coot [18], PHENIX software suite, and REFMAC5 of the CCP4 suite [19]. The molecular interfaces were analyzed using the PDBe PISA server [20]. Structural images were prepared with PyMOL (Schrödinger, LLC; New York, NY, USA).

2.3. Hydrolysis of oligosaccharides

The relative activity of Br2 toward various oligosaccharides was determined by reacting the wild-type Br2 with 1 mM oligosaccharides in 0.1 M sodium phosphate, pH 6.5, at 40 °C for 5 min. The reactions were stopped by boiling the solutions for 5 min. The released glucose was reacted with 2 mg/mL ABTS and glucose oxidase reagent (Sigma–Aldrich, St. Louis, MO, USA) at 37 °C for 15 min. The amount of glucose was determined by measuring the absorbance at 400 nm and comparing the results with its standard curve.

The mode of activity of Br2 was determined by reacting 1 µg of the wild-type Br2 with 20 mM laminaripentaose in 20 mM sodium phosphate, pH 6.5, at 40 °C for 120 min. At different time intervals, an aliquot of 4 µL was taken and spotted on a TLC plate (Silica gel 60 F254; Merck, Darmstadt, Germany), which was developed twice in a mixture of 2:1:1 (v/v) 1-butanol:acetic acid:water. The spots were visualized by soaking the TLC plate in 20 % sulfuric acid in ethanol and heating it until spots were observed [3].

2.4. Kinetic analysis

The kinetic reactions of the wild type and mutants were performed against various concentrations of *p*-nitrophenyl *β*-D-gluco-pyranoside (*p*NP-Glc) and glucooligosaccharides in 0.1 M sodium phosphate, pH 6.5, at 40 °C for 10 min. Then, 1.33 M sodium carbonate was added to stop the reactions with *p*NP-Glc, and the amount of released *p*-nitrophenol was determined by measuring the absorbance at 400 nm and comparing it with a standard curve. The reactions with glucooligosaccharides were stopped by boiling the solutions for 5 min. The released glucose was determined as described above. Kinetic parameters (K_m, k_{cat} and k_{cat}/K_m) were calculated by nonlinear regression of the Michaelis–Menten equation with KaleidaGraph (Synergy Software; Reading, PA, USA).

3. Results and discussion

3.1. The overall structure of Br2

The structures of the wild-type Br2, the acid/base mutant (E163Q) and the nucleophile mutant (E350G) were solved at 2.00 Å, 2.10 Å and 1.62 Å resolutions, respectively (Table 1). Each structure contains four molecules in an asymmetric unit (Fig. 1A), which is consistent with the apparent native molecular weight (189 kDa) estimated by size exclusion chromatography (data not shown), suggesting a tetrameric protein in solution. Their quaternary structures, as well as the subunits, are almost identical, with root-mean-square deviations of ~0.25 Å over ~1790 C_α atoms. The subunits interact through two different types of interfaces, which are between AB and CD, and between AC and BD (Supplementary Tables S2–S3). The buried surface areas of the AC and BD interfaces are approximately 2-fold larger and contain approximately 5-fold more hydrogen bonds with few salt bridges compared to the AB and CD interfaces. However, the solvation free energy ($\Delta^i G$) of the AB and CD interfaces was approximately 2-fold more negative than those of the AC and BD interfaces, suggesting that the AB and CD interfaces contain mainly hydrophobic contacts, and are more stable than the AC and BD interfaces [20].

The subunit of the wild-type Br2 displays a common $(\beta/\alpha)_8$ TIM-barrel fold (Fig. 1B), which is similar to other GH1 enzymes [4,17]. Glycerol, which was used as a cryoprotectant, is present in the active site pocket of the wild-type Br2. Glu163 and Glu350, which are the general catalytic acid/base and nucleophile residues, respectively, are located at the ends of *β*-strands 4 and 7, respectively. The overall structure of Br2 was superimposed with selected GH1 *β*-glucosidases, namely the Rice_Bglu1 E176Q mutant from *Oryza sativa*, *Ho*BGLA from *Halothermothrix orenii*, and Td2F2 from the compost metagenome (Fig. 1C; Supplementary Fig. S2) [7,21,22]. The Br2

Table 1

Data collection and refinement statistics of the wild-type and mutant forms of Br2.

Crystal	Wild-type	E163Q	E350G
PDB code	8J3M	8J5L	8J5M
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, b, c (Å)	105.7, 113.9, 181.9	105.5, 113.4, 180.5	105.3, 113.6, 180.9
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Molecules per asymmetric unit	4	4	4
Resolution (Å)	50–2.00 (2.12–2.00)	50–2.10 (2.22–2.10)	50–1.62 (1.72–1.62)
R_{sym} (%)	7.3 (72.5)	9.7 (89.8)	6.6 (72.7)
$I/\sigma(I)\rangle$	12.03 (1.76)	10.25 (1.55)	13.18 (1.96)
Completeness (%)	91.4 (97.1)	90.3 (96.4)	99.2 (97.1)
Redundancy	4.6 (4.3)	4.9 (4.8)	4.9 (4.8)
$CC_{1/2}$ (%)	99.8 (64.5)	99.7 (62.5)	99.9 (74.5)
Refinement			
Resolution (Å)	2.00	2.10	1.62
No. reflections	133,972	115,104	269,390
R_{work}/R_{free}	0.233/0.265	0.227/0.261	0.161/0.188
No. atoms			
Protein	14,482	14,439	14,504
Ligand/ion	74	99	201
Water	132	338	2107
B-factors			
Protein	43.20	48.71	23.31
Ligand/ion	66.01	71.29	56.44
Water	32.66	38.50	34.73
R.m.s. deviation values from ideal			
Bond lengths (Å)	0.006	0.013	0.012
Bond angles (°)	1.16	1.64	1.61
Ramachandran plot analysis			
Favored region (%)	96.69	96.07	97.54
Allowed region (%)	3.31	3.93	2.46
Outlier region (%)	0	0	0

Note: Highest resolution shell is shown in parenthesis. One crystal was used for the data collection in each structure analysis.
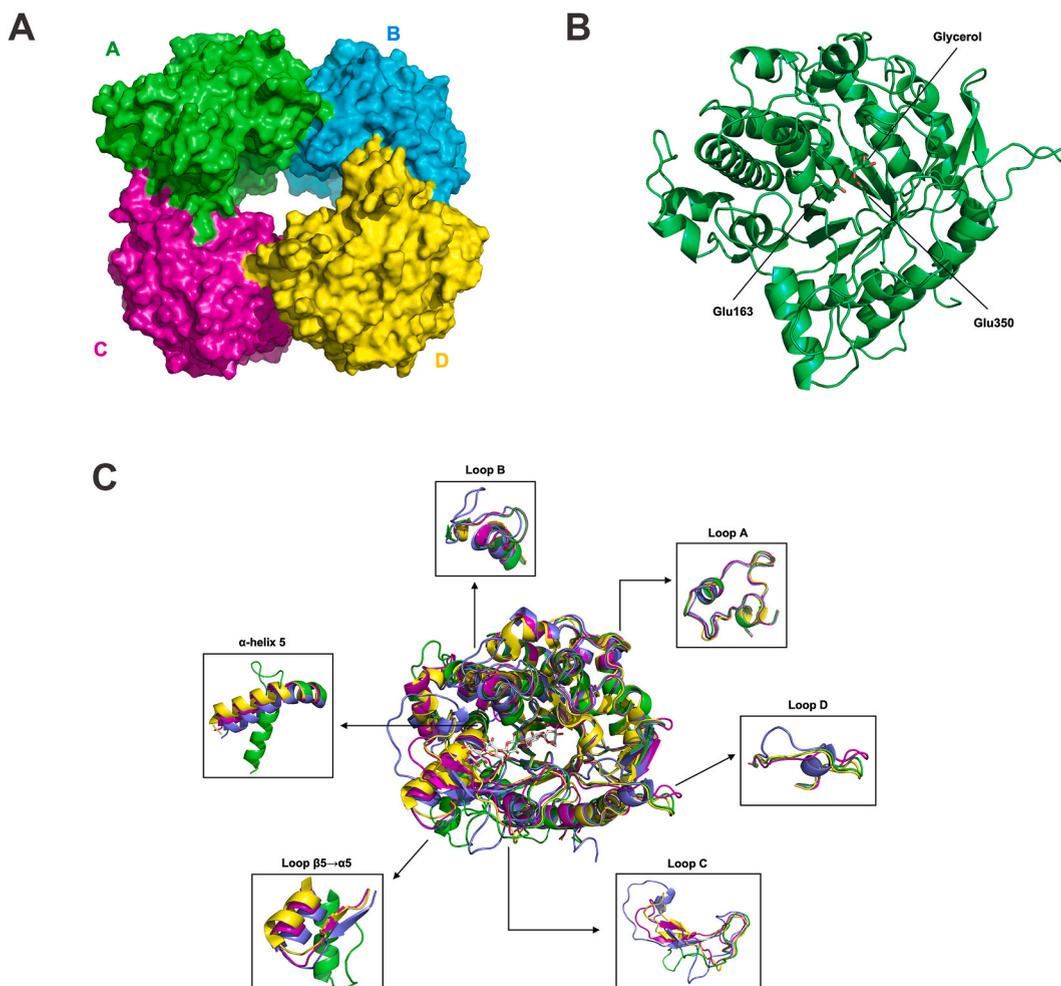
Fig. 1. The structure of the wild-type Br2. (**A**) The tetrameric structure of Br2. (**B**) The $(\beta/\alpha)_8$ structure of Br2 subunit. The side chains of Glu163 and Glu350, and glycerol are shown as stick models. (**C**) Superimposition of the $(\beta/\alpha)_8$ structures of the wild-type Br2 (PDB code 8J3M, green), the Rice_Bglu1 E176Q mutant in complex with cellopentaose (PDB code 3F5K, purple), *Ho*BGLA as the 2-deoxy-2-fluorogloucopyranosyl-enzyme intermediate (PDB code 4PTW, pink), and Td2F2 (PDB code 3WH5, yellow). The cellopentaose moiety in the Rice_BGlu1 E176Q mutant is shown in grey. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

structure is similar to these GH1 structures with root-mean-square deviations for the C_α atoms of 2.6–3.0 Å for approximately 450 residues; however, notable differences are observed in three out of four loop regions surrounding the active site, as defined previously [23]. Loop A (loop $\beta1\rightarrow\alpha1$) of Br2 is similar to the other three structures. Loops B and D (loops $\beta4\rightarrow\alpha4$ and $\beta7\rightarrow\alpha7$, respectively) of Br2 are similar to those of *Ho*BGLA and Td2F2, while Rice_BGlu1 contains a longer loop B and a short helix in loop D. In contrast, a large difference is found in loop C (loop $\beta6\rightarrow\alpha6$), in which Br2 contains only a long random coil, and *Ho*BGLA and Td2F2 harbor the β-turn-β motif, while Rice_BGlu1 contains a long coil and a short helix connecting the anti-parallel β-strands. More deviations are found in loop $\beta5\rightarrow\alpha5$ and α-helix 5. Although the three β-glucosidases possess a β-strand at loop $\beta5\rightarrow\alpha5$, Br2 contains a coil that directly connects β-strand 5 and α-helix 5. The α-helix 5 of most GH1 enzymes is approximately 25 residues long and bent in the middle, but the α-helix 5 of Br2 is divided into two helices by a unique 10-residue insertion (Ser242-Thr251) of a random coil structure.

3.2. Comparison between the active sites of Br2 and other GH1 structures

The amino acid residues in the active site of Br2 were identified by inference from the active sites of other GH1 β-glucosidases containing the complexed ligand. Among these, we were interested in the structures of the Rice_BGlu1 E176Q mutant in complex with cellopentaose and the covalent intermediate of *Ho*BGLA and 2-deoxy-2-fluoro-D-glucose (Fig. 2; Supplementary Fig. S1) [21,22]. The residues for glycone binding at subsite −1 of Br2 include Gln17, His118, Trp119, Asn162, Glu163 (acid/base catalyst), Tyr299, Glu350 (nucleophile), Glu404, and Trp405, which are all conserved among GH1 β-glucosidases except that Tyr131 is present in Rice_BGlu1 instead of Trp in other enzymes [23]. In the structure of the Rice_BGlu1 E176Q mutant, the residues at subsite −1 form direct hydrogen bonds to the nonreducing-end glucosyl unit of the complexed cellopentaose, while the residues at subsites +1 to +4 are diverse and
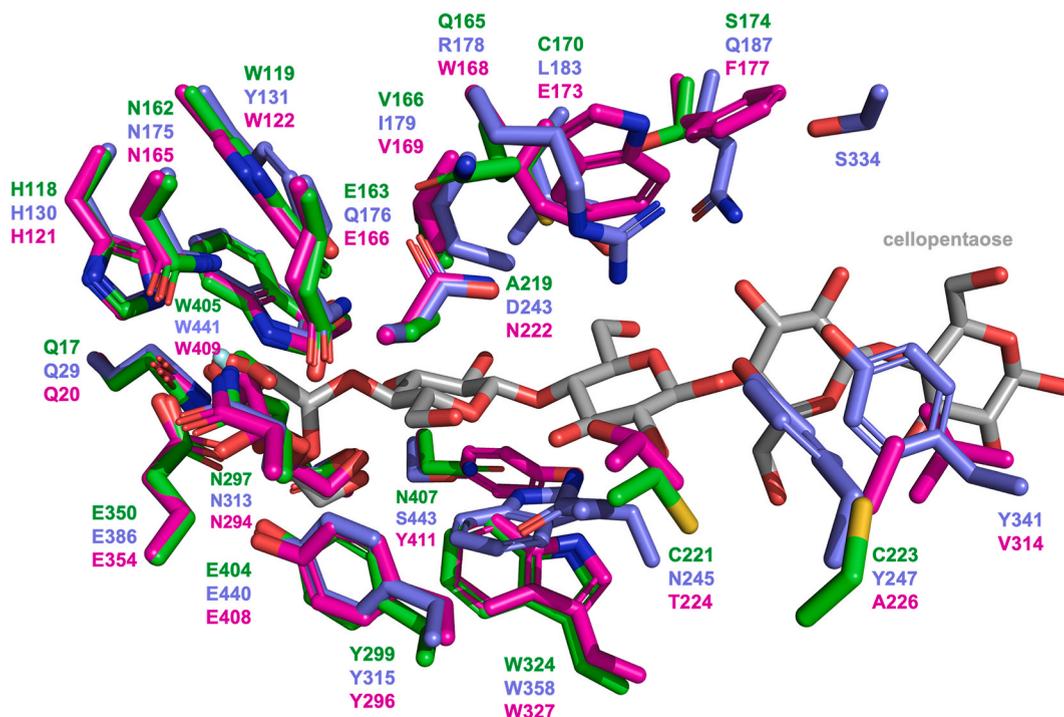
Fig. 2. Superimposition of the active site residues of the wild-type Br2 (PDB code 8J3M, green), the Rice_BGlu1 E176Q mutant in complex with cellopentaose (PDB code 3F5K, purple), and *Ho*BGLA as the 2-deoxy-2-fluoroglucopyranosyl-enzyme intermediate (PDB code 4PTW, pink). The cellopentaose moiety in the Rice_BGlu1 E176Q mutant is shown in grey. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

bind to the second to fifth glucosyl residues of cellopentaose via water-mediated hydrogen bonds [22]. These residues at subsites +1 to +2 of Rice_BGlu1 (Arg178, Ile179, Leu183, Asp243, Asn245, Asn313, Trp358 and Ser443) correspond to Gln165, Val166, Cys170, Ala219, Cys221, Asn297, Trp324 and Asn407 in Br2. Remarkably, the Rice_BGlu1 E176Q mutant contains Gln187, Tyr247, Ser334 and Tyr341 at subsites +3 to +4. In comparison, Br2 harbors only Ser174 and Cys223 at the two former positions, while *Ho*BGLA contains Phe177 and Ala226 at the two former positions as well as Val314 at the last position. The long binding pocket of Rice_Bglu1 concurred with its tendency to bind long cellooligosaccharides [22], while the relatively short binding pocket of Br2 may not support its hydrolytic activities toward the same substrates.

3.3. Hydrolysis of oligosaccharides

We next examined the relative hydrolytic activity of the wild-type Br2 toward glucooligosaccharides of various degrees of polymerization (Supplementary Table S4). Among the glucodisaccharides, Br2 showed the highest velocity toward laminaribiose (β-1,3 linkage), followed by sophorose (β-1,2 linkage), cellobiose (β-1,4 linkage) and gentiobiose (β-1,6 linkage). Among the glucooligosaccharides, Br2 displayed higher activities toward laminarioligosaccharides (β-1,3 linkage) than cellooligosaccharides (β-1,4 linkage), with the highest preference toward laminaritriose. However, Br2 showed low relative activity toward laminarin, which is a β-1,3-glucopolysaccharide. These results agree with the presence of the substrate binding residues at subsites +1 to +2 of Br2 and only two residues at the distal end of the binding pocket (Fig. 2). The time-course hydrolysis of laminaripentaose yielded glucose and laminaritetraose as initial hydrolysis products (Supplementary Fig. S3). Based on its relative activity and lack of signal peptide [3], Br2 appears to be an exo-acting β-1,3-glucosidase, possibly involved in the intracellular degradation of laminarioligosaccharides from digested β-1,3-glucans rather than β-1,4-cellulose [3,24].

3.4. Kinetic analysis

The previously reported kinetic parameters of *Ho*BGLA and the wild-type Br2 toward cellobiose revealed that both enzymes showed similar K_m values (25.4 and 24.3 mM, respectively), but *Ho*BGLA exhibited a 120-fold higher k_{cat} value (366 and 3 s^{-1}, respectively), resulting in an approximately 110-fold higher k_{cat}/K_m value (14.4 and 0.13 mM^{-1}s^{-1}, respectively) than that of Br2 [3,22]. We were interested in the roles of amino acid residues at the positive subsites, which could be responsible for these catalytic differences but have received much less attention than those at subsite -1. Thus, five residues located at subsites +1 to +2 of Br2, namely Gln165, Cys170, Ala219, Cys221 and Asn407, were individually replaced with the corresponding residues of *Ho*BGLA, namely Trp168, Glu173, Asn222,

Thr224 and Tyr411, respectively (Fig. 2), generating the Q165W, C170E, A219N, C221T and N407Y mutants, respectively. Unfortunately, the Q165W mutant could not be characterized, as it was present in the insoluble fraction after cell lysis. The other four Br2 mutants were successfully expressed and purified. Among the glucooligosaccharides tested, the wild-type Br2 showed the highest catalytic efficiency toward laminaritriose, and higher efficiency toward laminarioligosaccharides than cellooligosaccharides with the same degree of polymerization (Table 2). These data agree well with the relative hydrolytic activities (Supplementary Table S4) and further support the likely role of Br2 in the degradation of β-1,3-glucans in animal feed.

Cys170 of Br2 corresponds to Leu183 of Rice_Bglu1, Phe198 of *Zea mays* β-glucosidase Glu1 (ZmGlu1) and Val196 of *Sorghum bicolor* dhurrinase-1 (SbDhr1), all of which were key binding residues at subsite +1 (Fig. 2; Supplementary Fig. S1) [23,25,26], while the N189F mutation in *Dalbergia cochinchinensis* dalcochinase (TRDC) resulted in improved binding to long-chain alkyl alcohols [27, 28]. Cys170 of Br2 is located at the equivalent position to Glu173 of *Ho*BGLA, which was modeled to form hydrogen bonds to the second galactosyl unit of β-D-galactopyranosyl-(1 → 6)-lactose [21]. Except for *p*NP-Glc and laminaribiose, the C170E mutant of Br2 showed approximately 1.5- to 2-fold higher efficiency toward all substrates compared with that of the wild-type enzyme (Table 2). This mutant exhibited approximately 4- and 5-fold higher efficiency toward laminaritriose and cellotriose than their respective disaccharides. These data suggested that the C170E mutation in Br2 likely contributed to the binding of glucooligosaccharide substrates at subsite +2.

Ala219 of Br2 corresponds to the highly conserved Asn residue in several GH1 β-glucosidases (Supplementary Fig. S1). Both Asn223 of Td2F2 and Asn227 of MeBglD2 are important for binding to an acceptor sugar at subsite +1 [7,8], and Asn259 of SbDhr1 binds to the cyano group of dhurrin via a water molecule [26]. In addition, the N227A and N227D mutations in MeBglD2 reduced the activity toward cellobiose [8]. On the other hand, the A219N mutant of Br2 showed higher K_m values and lower k_{cat} values for all substrates tested compared with those of the wild-type enzyme (Table 2), implying the substrate recognition manner of Br2 at subsite +1 is different from Td2F2 and MeGglD2. Nonetheless, this mutation resulted in 50- and 9-fold lower hydrolytic efficiency toward cellobiose and gentiobiose, respectively, relative to laminaribiose, whereas the wild-type enzyme exhibited as much as 350-fold reductions in the corresponding values. Therefore, the A219N substitution in Br2 may have a positive effect on the substrate preference at subsite +1 for

Table 2
Kinetic parameters of the wild-type and mutant forms of Br2 toward *p*NP-Glc and glucooligosaccharides.

Substrate	*p*NP-Glc			Sophorose			Gentiobiose		
Parameters	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})
Wild-type	0.55 ± 0.04	88.5 ± 1.2	160 ± 11	2.68 ± 0.16	65.7 ± 1.4	24.5 ± 1.5	35.8 ± 10.9	4.15 ± 1.26	0.12 ± 0.05
C170E	0.43 ± 0.02	40.2 ± 0.4	93.8 ± 5.2	2.33 ± 0.13	122 ± 2	52.3 ± 3.0	35.9 ± 6.1	5.81 ± 0.98	0.16 ± 0.04
A219N	1.03 ± 0.07	21.4 ± 0.3	20.8 ± 1.4	14.1 ± 6.3	5.97 ± 2.66	0.42 ± 0.27	56.3 ± 25.4	2.69 ± 1.21	0.05 ± 0.03
C221T	0.66 ± 0.07	31.1 ± 0.9	47.0 ± 5.3	2.23 ± 0.15	111 ± 3	49.9 ± 3.6	33.2 ± 6.4	3.88 ± 0.74	0.12 ± 0.03
N407Y	1.03 ± 0.04	27.5 ± 0.2	26.7 ± 1.0	3.01 ± 0.18	63.2 ± 1.5	21.0 ± 1.4	37.4 ± 8.3	2.91 ± 0.64	0.08 ± 0.02
Substrate	**Laminaribiose**			**Laminaritriose**			**Laminaritetraose**		
Parameters	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})
Wild-type	1.29 ± 0.08	54.5 ± 1.0	42.2 ± 2.6	0.90 ± 0.09	76.9 ± 2.5	85.4 ± 8.8	2.56 ± 0.41	23.8 ± 1.7	9.31 ± 1.65
C170E	0.99 ± 0.07	34.8 ± 0.76	35.1 ± 2.7	1.37 ± 0.06	181 ± 3	132 ± 6	3.86 ± 0.49	59.4 ± 3.8	15.4 ± 2.2
A219N	11.7 ± 4.2	5.03 ± 1.81	0.43 ± 0.22	6.55 ± 1.45	8.58 ± 1.87	1.31 ± 0.41	13.8 ± 3.8	0.75 ± 0.21	0.05 ± 0.02
C221T	1.70 ± 0.15	47.7 ± 1.5	28.1 ± 2.6	0.38 ± 0.02	145 ± 2	382 ± 23	3.56 ± 0.31	66.3 ± 2.8	18.6 ± 1.8
N407Y	1.39 ± 0.09	32.1 ± 0.8	23.1 ± 1.7	1.58 ± 0.11	71.7 ± 1.7	45.4 ± 3.2	4.99 ± 0.46	30.3 ± 1.6	6.06 ± 0.64
Substrate	**Cellobiose**			**Cellotriose**			**Cellotetraose**		
Parameters	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1}mM^{-1})
Wild-type	24.3 ± 9.8	3.14 ± 1.26	0.13 ± 0.07	6.73 ± 0.32	5.73 ± 0.11	0.85 ± 0.04	3.87 ± 0.16	5.89 ± 0.12	1.52 ± 0.07
C170E	17.4 ± 3.3	4.03 ± 0.76	0.23 ± 0.06	4.86 ± 0.64	5.70 ± 0.31	1.17 ± 0.17	6.05 ± 1.26	15.0 ± 3.1	2.48 ± 0.73
A219N	37.1 ± 9.3	0.32 ± 0.08	0.009 ± 0.003	25.5 ± 3.7	0.18 ± 0.03	0.007 ± 0.002	13.1 ± 3.2	0.09 ± 0.02	0.007 ± 0.002
C221T	20.6 ± 3.1	3.84 ± 0.58	0.19 ± 0.04	4.72 ± 0.56	8.09 ± 0.36	1.72 ± 0.22	3.46 ± 0.64	13.8 ± 1.3	4.00 ± 0.82
N407Y	19.7 ± 4.3	2.00 ± 0.44	0.10 ± 0.03	3.57 ± 0.33	2.35 ± 0.07	0.66 ± 0.06	1.83 ± 0.09	0.73 ± 0.01	0.40 ± 0.02

cellobiose and gentiobiose relative to laminaribiose.

Cys221 of Br2 occurs at the same position as Asn245 of Rice_Bglu1, Asn229 of MeBglD2 and Phe261 of SbDhr1, which are involved in substrate binding at the positive subsites (Fig. 2; Supplementary Fig. S1) [8,22,26]. Mutations at the equivalent position altered the properties of subsite +1 of GH1 β-glucosidases, such as the N229A and N229D mutations in MeBglD2 [8] and the N245V mutation in Rice_Bglu1 [23]. Compared with the wild-type enzyme, the C221T mutant of Br2 exhibited decreased efficiency toward pNP-Glc and laminaribiose but up to 4-fold increased efficiency toward all other substrates except gentiobiose. This mutation resulted in 14- and 9-fold higher efficiency toward laminaritriose and cellotriose than their respective disaccharides (Table 2). Thus, the C221T mutation in Br2 could also contribute to the interactions with glucooligosaccharide substrates at subsite +2, similar to the C170E mutation mentioned above.

The position of Asn407 of Br2 is highly variable in GH1 β-glucosidases, such as Ala647 of ZmGlu1, Ser463 of SbDhr1, and Arg413 of MeBglD2, and hence, it has been proposed to be responsible for substrate specificity (Supplementary Fig. S1) [8,25,26]. Compared with the wild-type enzyme, the N407Y mutant of Br2 exhibited similar kinetic parameters toward most glucooligosaccharide substrates (Table 2). However, its K_m values toward cellooligosaccharides were lower, while those toward laminarioligosaccharide were higher compared with those of the wild-type enzyme. Additionally, the N407Y mutation led to 6-fold higher efficiency toward cellotriose than toward cellobiose. Thus, the N407Y substitution in Br2 may offer a favorable binding site for cellooligosaccharides at the positive subsites.

In conclusion, we determined the structure of Br2, which, to our knowledge, is the first structure of rumen metagenomic GH1 β-glucosidase. Br2 displays the canonical $(\beta/\alpha)_8$-barrel structure, but the length and conformation of loop $\beta5 \rightarrow \alpha5$ are notably different and a 10-residue coil is inserted in the middle of the distorted $\alpha5$. The residues for glycone binding at subsite −1 of Br2 are similar to those of other GH1 β-glucosidases, but there are many differences at the positive subsites. Br2 showed a much higher preference toward laminarioligosaccharides than cellooligosaccharides, with the highest catalytic efficiency toward laminaritriose; thus, Br2 may play a role in the hydrolysis of β-1,3-glucans rather than β-1,4-cellulose. The residues at subsites +1 and + 2 of Br2 were individually substituted with the corresponding residues of HoBGLA, which exhibited high catalytic efficiency toward cellobiose. The C170E and C221T mutations increased the catalytic efficiency toward laminaritriose and cellotriose compared with their respective disaccharides. The A219N mutation appeared to improve the substrate preference for cellobiose and gentiobiose relative to laminaribiose, while the N407Y mutation increased the binding affinity toward cellooligosaccharides. Together, our results highlighted the roles of amino acid residues at the positive subsites, which contributed to the catalytic prowess of Br2.

Data availability statement

Data associated with this study has been deposited at Protein Data Bank under the accession numbers 8J3M (wild-type), 8J5L (E163Q) and 8J5M (E350G).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e21923.

References

[1] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2014) D490−D495.

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

[3] E. Suwan, S. Arthornthurasuk, P.T. Kongsaeree, A metagenomic approach to discover a novel *β*-glucosidase from bovine rumens, Pure Appl. Chem. 89 (2017) 941–950.

[4] T. Barrett, C.G. Suresh, S.P. Tolley, E.J. Dodson, M.A. Hughes, The crystal structure of a cyanogenic *β*-glucosidase from white clover, a family 1 glycosyl hydrolase, Structure 3 (1995) 951–960.

[5] C.S. Rye, S.G. Withers, Glycosidase mechanisms, Curr. Opin. Chem. Biol. 4 (2000) 573–580.

[6] D. Cowan, Q. Meyer, W. Stafford, S. Muyanga, R. Cameron, P. Wittwer, Metagenomic gene discovery: past, present and future, Trends Biotechnol. 23 (2005) 321–329.

[7] T. Matsuzawa, T. Jo, T. Uchiyama, J.A. Manninen, T. Arakawa, K. Miyazaki, S. Fushinobu, K. Yaoi, Crystal structure and identification of a key amino acid for glucose tolerance, substrate specificity, and transglycosylation activity of metagenomic *β*-glucosidase Td2F2, FEBS J. 283 (2016) 2340–2353.

[8] T. Matsuzawa, M. Watanabe, Y. Nakamichi, H. Akita, K. Yaoi, Crystal structure of metagenomic *β*-glycosidase MeBglD2 in complex with various saccharides, Appl. Microbiol. Biotechnol. 106 (2022) 4539–4551.

[9] L. Bao, Q. Huang, L. Chang, Q. Sun, J. Zhou, H. Lu, Cloning and characterization of two *β*-glucosidase/xylosidase enzymes from yak rumen metagenome, Appl. Biochem. Biotechnol. 166 (2012) 72–86.

[10] 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, Microbial *β*-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail, Biotechnol. Biofuels 5 (2012) 73.

[11] Y. Li, N. Liu, H. Yang, F. Zhao, Y. Yu, Y. Tian, X. Lu, Cloning and characterization of a new *β*-glucosidase from a metagenomic library of rumen of cattle feeding with *Miscanthus sinensis*, BMC Biotechnol. 14 (2014) 1–9.

[12] M. Ramírez-Escudero, M.V. Del Pozo, J. Marín-Navarro, B. González, P.N. Golyshin, J. Polaina, M. Ferrer, J. Sanz-Aparicio, Structural and functional characterization of a ruminal β-glycosidase defines a novel subfamily of glycoside hydrolase family 3 with permuted domain topology, J. Biol. Chem. 291 (2016) 24200–24214.

[13] S. Ariaeenejad, S. Nooshi-Nedamani, M. Rahban, K. Kavousi, A.G. Pirbalooti, S. Mirghaderi, M. Mohammadi, M. Mirzaei, G.H. Salekdeh, A novel high glucose-tolerant *β*-glucosidase: targeted computational approach for metagenomic screening, Front. Bioeng. Biotechnol. 8 (2020) 813.

[14] J. Braman, C. Papworth, A. Greener, Site-directed mutagenesis using double-stranded plasmid DNA templates, Methods Mol. Biol. 57 (1996) 31–44.

[15] W. Kabsch, XDS, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.

[16] P.V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, T.C. Terwilliger, A. Urzhumtsev, P.H. Zwart, P.D. Adams, Towards automated crystallographic structure refinement with phenix.refine, Acta Crystallogr. D Biol. Crystallogr. 68 (2012) 352–367.

[17] T.M. Gloster, J.M. Macdonald, C.A. Tarling, R.V. Stick, S.G. Withers, G.J. Davies, Structural, thermodynamic, and kinetic analyses of tetrahydrooxazine-derived inhibitors bound to *β*-glucosidases, J. Biol. Chem. 279 (2004) 49236–49242.

[18] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501.

[19] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 355–367.

[20] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, J. Mol. Biol. 372 (2007) 774–797.

[21] N. Hassan, T.H. Nguyen, M. Intanon, L.D. Kori, B.K.C. Patel, D. Haltrich, C. Divne, T.C. Tan, Biochemical and structural characterization of a thermostable *β*-glucosidase from *Halothermothrix orenii* for galacto-oligosaccharide synthesis, Appl. Microbiol. Biotechnol. 99 (2015) 1731–1744.

[22] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, J. Svasti, J.R. Ketudat Cairns, The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase, J. Struct. Biol. 173 (2011) 169–179.

[23] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, W. Oonanant, D.R. Bevan, A. Esen, C.J. Chen, R. Opassiri, J. Svasti, J.R. Ketudat Cairns, Structural insights into rice BGlu1 β-glucosidase oligosaccharide hydrolysis and transglycosylation, J. Mol. Biol. 377 (2008) 1200, 1125.

[24] A. Nakashima, K. Yamada, O. Iwata, R. Sugimoto, K. Atsuji, T. Okawa, N. Ishibashi-Ohgo, K. Suzuki, *β*-Glucan in foods and its physiological functions, J. Nutr. Sci. Vitaminol. 64 (2018) 8–17.

[25] M. Czjzek, M. Cicek, V. Zamboni, D.R. Bevan, B. Henrissat, A. Esen, The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes, Proc. Natl. Acad. Sci. USA 97 (2000) 13555–13560.

[26] L. Verdoucq, J. Moriniére, D.R. Bevan, A. Esen, A. Vasella, B. Henrissat, M. Czjzek, Structural determinants of substrate specificity in family 1 *β*-glucosidases: novel insights from the crystal structure of sorghum dhurrinase-1, a plant *β*-glucosidase with strict specificity, in complex with its natural substrate, J. Biol. Chem. 279 (2004) 31796–31803.

[27] N. Tongtubtim, P. Thenchartanan, K. Ratananikom, K. Choengpanya, J. Svasti, P.T. Kongsaeree, Multiple mutations in the aglycone binding pocket to convert the substrate specificity of dalcochinase to linamarase, Biochem. Biophys. Res. Commun. 504 (2018) 647–653.

[28] P. Thenchartanan, P. Pitchayatanakorn, P. Wattana-Amorn, A. Ardá, J. Svasti, J. Jiménez-Barbero, P.T. Kongsaeree, Synthesis of long-chain alkyl glucosides via reverse hydrolysis reactions catalyzed by an engineered *β*-glucosidase, Enzyme Microb. Technol. 140 (2020), 109591.