BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

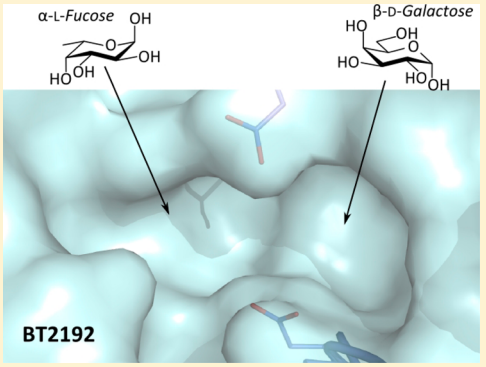
Unraveling the Substrate Recognition Mechanism and Specificity of the Unusual Glycosyl Hydrolase Family 29 BT2192 from *Bacteroides thetaiotaomicron*

Laure Guillotin, Pierre Lafite, and Richard Daniellou*

Université Orléans, CNRS, ICOA, UMR 7311, F-45067 Orleans, France

Supporting Information

ABSTRACT: Glycosyl hydrolase (GH) family 29 (CAZy database) consists of retaining α-L-fucosidases. We have identified BT2192, a protein from *Bacteroides thetaiotaomicron*, as the first GH29 representative exhibiting both weak α-L-fucosidase and β-D-galactosidase activities. Determination and analysis of X-ray structures of BT2192 in complex with β-D-galactoside competitive inhibitors showed a new binding mode different from that of known GH29 enzymes. Three point mutations, specific to BT2192, prevent the canonical GH29 substrate α-L-fucose from binding efficiently to the fucosidase-like active site relative to other GH29 enzymes. β-D-Galactoside analogues bind and interact in a second pocket, which is not visible in other reported GH29 structures. Molecular simulations helped in the assessment of the flexibility of both substrates in their respective pocket. Hydrolysis of the fucosyl moiety from the putative natural substrates like 3-fucosyllactose or Lewis^X antigen would be mainly due to the efficient interactions with the galactosyl moiety, in the second binding site, located more than 6−7 Å apart.



Glycosyl hydrolases (GHs, EC 3.2.1.xx) are widespread enzymes that catalyze the hydrolysis of glycosidic bonds. To take into account the structural diversity of carbohydrates and enzymes,[1] a classification of GHs based on their sequence similarity was adopted.[2] More than 130 families of GHs have been identified and are available through the CAZy database (http://www.cazy.org).[3] In addition, two main mechanisms of glycosidic bond hydrolysis by GHs have been characterized, involving either inversion or retention of the configuration of the anomeric carbon.[4] In particular, two side-chain residues, generally carboxylates, mediate the reaction in the retaining mechanism, leading to the formation of a covalent glycosyl−enzyme intermediate. The first carboxylate residue acts as a general acid/base, whereas the second is the catalytic nucleophile.

Bacteroides thetaiotaomicron is one of the dominant bacteria in the human gut. Its genome contains genes encoding many carbohydrate-processing proteins, which give the bacteria a great ability to degrade nondigestible polysaccharides from plants.[5] More than 200 GHs have been predicted to be encoded by the *B. thetaiotaomicron* genome; among them, nine have been classified as part of the GH29 family.

The GH29 family consists of retaining α-L-fucosidases; several individual enzymes have been characterized. Five X-ray structures of GH29 enzymes have been deposited in the Protein Data Bank (PDB), including proteins from bacteria *Thermotoga maritima* (TM0306, PDB entry 1HL8[6]), *Bifidobacterium longum* (Blon_2336, PDB entry 3MO4[7]), and *B. thetaiotaomicron* (BT2970, PDB entry 2WVS;[8] BT2192, PDB entry 3EYP; and BT3798, PDB entry 3GZA). The latter two were the result of structural genomics programs and were originally deposited as "putative α-L-fucosidases", based on their sequence similarity and features [up to 30% (45%) identical (homologous) sequence], and the presence of conserved catalytic residues. Other GH29 enzymes were characterized and further identified as α-L-fucosidases in many organisms, including archaea,[9] plants,[10,11] and animals.[12,13] Site-directed mutagenesis, coupled with mechanism-based inactivation and chemical rescue by N_3, led to the identification of catalytic carboxylate side-chain residues in several GH29 representatives.[9,14−16] A global identification of the catalytic carboxylates was recently reported for the GH29 family.[17] GH29 enzymes can be classified in two subgroups, according to their substrate specificity and sequence homology. Group GH29-A contains enzymes that have low substrate specificity.[17] The second group, GH29-B, includes GH29 enzymes that are more specific for α(1→3/4) fucosidic linkages.[18,19]

In this study, we have focused on the BT2192 protein, the X-ray structure of which was previously reported (PDB entry 3EYP), and BT2192 was recently classified as part of the GH29-B group.[20] This 53 kDa protein has been shown to be a dimer in solution.[20] We have demonstrated that although BT2192 can hydrolyze α(1→3/4) fucosidic linkage-containing compounds, it also exhibited weak β-D-galactosidase activity,

Received: July 17, 2013
Revised: February 12, 2014
Published: February 14, 2014

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447−1455

which was unexpected for GH29 enzymes. Moreover, crystallographic structures of BT2192 in complex with β-D-galactosidase competitive inhibitors revealed an unexpected binding site for β-D-galactosides, as well as identification of new catalytic carboxylate residues, distinct from those of canonical GH29 α-L-fucosidases. Molecular modeling simulations highlighted the importance of a branched galactosyl residue in substrate recognition for α-L-fucosidase activity, thus leading to a better understanding of GHs from this family.

EXPERIMENTAL PROCEDURES

Materials. All *p*-nitrophenyl monosaccharides were purchased from Carbosynth (Compton, U.K.). All chemicals used were of analytical reagent grade. Lewis^X trisaccharide was a generous gift from Dr. Duverger (Glycodiag, Orléans, France). The pSGX3 expression plasmid expressing BT2192 from *B. thetaiotaomicron* (Protein Structure Initiative clone 11092k3BCt5p1) was a generous gift from L. Stead (Albert Einstein College of Medicine, Bronx, NY). This plasmid (pSGX3-BT2192) encodes an N-terminal His_6 tag fused to the entire coding region of the BT2192 gene.

Expression and Purification of BT2192. *Escherichia coli* Rosetta(DE3) cells transformed with pSGX3-BT2192 were grown in selective LB medium at 37 °C and 220 rpm until the OD_{600} reached 0.6. The culture was then induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranose (IPTG) and incubated overnight at 30 °C and 220 rpm. Cells were then harvested by centrifugation and lysed with freeze−thaw cycles. After centrifugation, the supernatant was further used for IMAC purification. The Ni-NTA column was equilibrated with 50 mM Tris-HCl (pH 8) and 200 mM NaCl, and the bound proteins were eluted using an imidazole gradient (10 to 500 mM). For crystallization, the tagged protein was further purified via size exclusion chromatography [Superdex 200GL, elution buffer consisting of 50 mM Tris-HCl (pH 8) and 200 mM NaCl].

Enzyme Activity Assays. BT2192 activity was assayed at 37 °C in 200 μL reaction mixtures containing substrate (0.01−10 mM), 16 μg of enzyme, and Tris buffer (200 mM, pH 7.0). Residual spontaneous hydrolysis of the substrate was assessed on a sample containing H_2O instead of enzyme and was used for the normalization of data. For *p*-nitrophenol (pNP)-containing substrates, after reaction for 30 min, 100 μL of 1 M sodium carbonate was added, and the amount of pNP produced was quantified by absorbance measurement at 405 nm. For oligosaccharide substrates, activity was assessed after incubation for 3 h at 37 °C by thin layer chromatography[19,20] on silica, using a butanol/acetic acid/water mixture (2:1:1) as the developing solvent, and visualized using orcinol stain [5% orcinol in a H_2SO_4 (5%)/EtOH solution]. Similar assays were performed with a pH range of 4−10 and reaction temperatures of 25, 37, 45, 60, and 80 °C to determine the optimal enzymatic reaction conditions. Inhibition was assayed under identical conditions, adding inhibitor (0.01−5 mM in water) prior to the substrate. The reaction mixture was incubated for 30 min at 37 °C, and the resulting activity was measured every 5 min at 405 nm. All kinetic parameters were calculated by fitting saturation curves (as a mean of triplicate measurements) with standard Michaelis−Menten and inhibition equations,[21] using Prism 6 (GraphPad). The Arrhenius equation[21] was used to calculate the activation energy for substrate hydrolysis by BT2192.

Sequence Analysis. The BT2192 peptidic sequence was retrieved from Uniprot (http://www.uniprot.org; accession number Q8A5P6) and was blasted using the implemented Blast software in Uniprot against Uniprot and Swiss Prot Databases. Sequence alignments between GH29 enzymes were performed using Clustal Omega on the EBI server [22,23].

Crystallization, Data Collection, and Model Refinement. BT2192 crystals were grown at 20 °C by sitting-drop vapor diffusion in 2 μL drops containing an equal amount of pure protein (20 mg/mL) and 2.5 M ammonium citrate. For BT2192 in a complex with either IPTG or oNPTG, crystals were soaked for 30 min in mother liquor containing 250 or 5 mM ligand, respectively. The crystals were cryoprotected upon addition of 20% glycerol to the mother liquor and flash-cooled in liquid nitrogen. Data were collected on beamline Proxima 1 at the Soleil Synchrotron Facility and reduced and scaled with the X-ray Detector Software suite (XDS).[24] The structures were determined by molecular replacement using Molrep[25] and the ligand-free structure of BT2192 as the template (PDB entry 3EYP). Refinement and model building were conducted using Phenix[26] and COOT.[27] Model quality was assessed at every refinement step using Molprobity.[28] X-ray structures of BT2192 in complex with IPTG and oNPTG were deposited to the Protein DataBank (PDB codes 4OUE and 4OZO). Data and final refinement statistics are listed in Table 1. Root-mean-square deviation (rmsd) calculations for structural comparisons and figure generation were performed using PyMOL (http://

Table 1. Data Collection and Refinement Statistics for BT2192 Structures in Complex with IPTG and oNPTG

	IPTG-bound	oNPTG-bound
space group	$I2_12_12_1$	$I2_12_12_1$
unit cell dimensions	a = 126 Å, b = 122 Å, c = 160 Å	a = 125 Å, b = 122 Å, c = 160 Å
	$\alpha = \beta = \gamma = 90°$	$\alpha = \beta = \gamma = 90°$
resolution (outer shell) (Å)	6.50−2.35 (2.45−2.35)	7.0−2.60 (2.50−2.60)
total no. of reflections	381914	281551
no. of unique reflections	51368	38008
completeness (%)	100.0 (100.0)	100.0 (99.9)
redundancy	7.4 (7.5)	7.4 (7.4)
R_{merge}	0.104 (0.627)	0.138 (0.713)
$I/\sigma I$	17.7 (4.4)	13.8 (3.6)
R_{work}	0.16	0.17
R_{free}	0.20	0.22
rmsd from ideal geometry		
bonds (Å)	0.007	0.007
bond angles (deg)	1.018	1.072
no. of atoms	7789	7569
protein	7319	7300
solvent	416	203
ligand	54	66
overall B factor (Å^2)	22.9	18.7
protein	22.5	18.6
main-chain	22.3	17.9
side-chain	23.5	19.5
solvent	28.8	19.1
ligand	27.2	23.3
Ramachandran statistics (favored and allowed) (%)	99.9	98.6

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447−1455

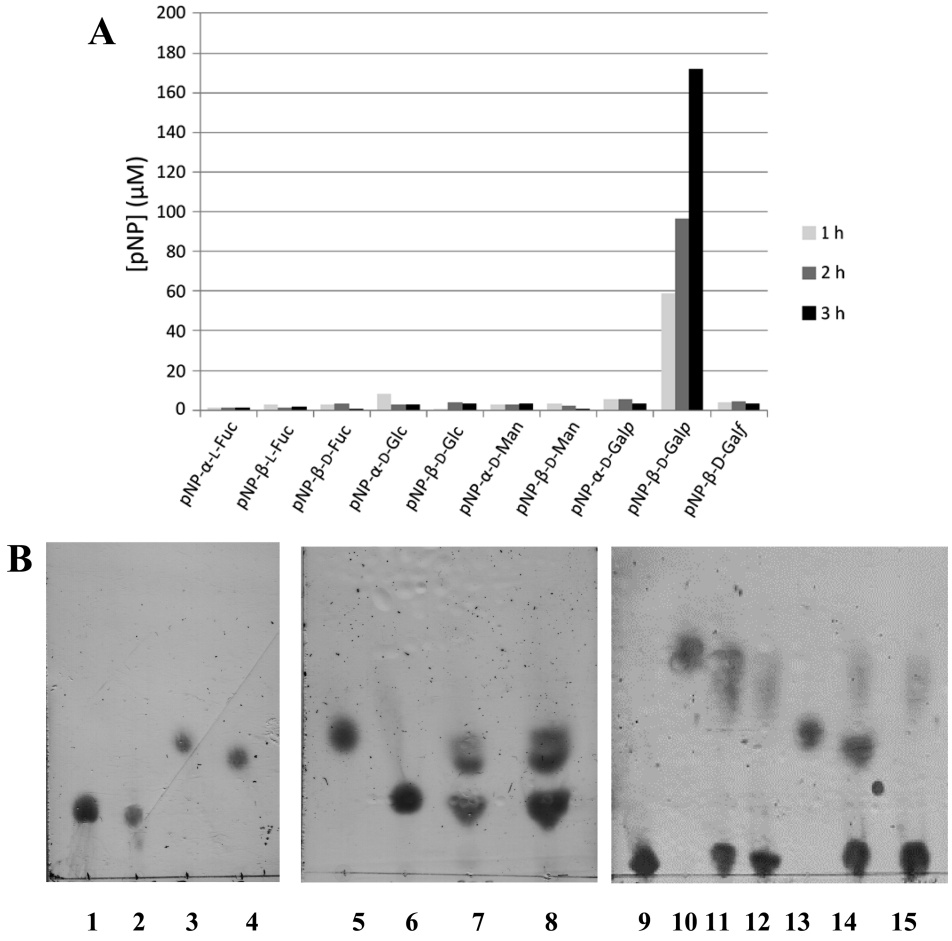


Figure 1. (A) pNP formation upon incubation of BT2192 with different pNP-*O*-monosaccharides. Substrates (1 mM) were incubated for the indicated time at 37 °C with 30 μM BT2192. (B) TLC analysis of BT2192-catalyzed oligosaccharides hydrolysis: lane 1, Lac; lane 2, Lac and BT2192 (22 μM, 3 h, pH 7, 37 °C); lane 3, D-Gal; lane 4, D-Glc; lane 5, D-Gal; lane 6, Lac; lane 7, pNP-Gal, Lac, and BT2192 (22 μM, 3 h, pH 7, 37 °C); lane 8, Lac and D-Gal; lane 9, Lewis^X; lane 10, L-Fuc; lane 11, Lewis^X and BT2192 (22 μM, 3 h, pH 7, 37 °C); lane 12, co-elution of the contents of lanes 10 and 15; lane 13, D-Gal; lane 14, co-elution of the contents of lanes 13 and 15; lane 15, Lewis^X and BT2192 (22 μM, 3 h, pH 7, 37 °C).

www.pymol.org). Structural alignments of GH29 catalytic sites with the BT2192 structure (in complex with IPTG or oNPTG) were performed using the following truncated sequences: residues 35−343 for BT2192, residues 21−350 for PDB entry 1HL8, residues 52−376 for PDB entry 2WVU, residues 30−359 for PDB entry 3GZA, and residues 22−349 for PDB entry 3MO4.

Computational Simulations. NAMD[29] was used to perform all molecular dynamics (MD) simulations of 3-fucosyllactose, Lewis^X, pNP-β-D-galactose, and pNP-α-L-fucose ligands in complex with BT2192. Topology and parameter files for oligosaccharides were obtained from Glycam[30,31] and the *tleap* program from AmberTools 13.[32] 3-Fucosyllactose and Lewis^X were aligned and constrained in the BT2192 active site, using BT2192 in complex with IPTG (this work) and BT2970 in complex with pNP-α-L-fucopyranose (PDB entry 2WVU) for the galactose (pocket G) and fucose (pocket F) moiety, respectively. Complexes of pNP-Gal and pNP-Fuc with BT2192 were also prepared using these X-ray structures, in the corresponding pockets (pocket G for pNP-Gal, pocket F for pNP-Fuc, and pocket F for pNP-Gal). The protein−ligand complexes were then equilibrated with several cycles of minimization (steepest descent, 10000 steps) and MD simulations (200 K, 20 ps) with the protein backbone, galactose, and fucose atoms restrained (*in vacuo*, with a relative dielectric constant set to 80 to mimic the water solvation). Then, 10 MD simulations were run independently for 0.5 ns for each ligand, without any restraints at 200 K. For postprocessing analysis, final models were minimized (steepest descent, 10000 steps) and individual rmsd values were calculated for all protein and ligand atoms over the 10 individual snapshots. In addition, 1000 randomly chosen snapshots of the 10 MD simulations were extracted to perform rmsd calculations, which gave results similar to those obtained with minimized models.

RESULTS

Substrate Specificity of BT2192. As reported previously, BT2192 belongs to subgroup GH29-B, which is constituted by GH29 enzymes that are specific of the α(1→3/4) fucosidic linkage.[17] BT2192 cannot hydrolyze the non-natural substrate pNP-Fuc, which is commonly used to assess fucosidase activity[20] (Figure 1A). However, substrates containing an α(1→3/4) fucosidic linkage, such as 3-fucosyllactose [Galβ1→4(Fucα1→3)Glc] or Lewis^X antigen, are weakly hydrolyzed in the presence of BT2192[20] (Figures 1B and 2). As these activities were dramatically lower than those of other GH29 enzymes, the selectivity of BT2192 for other carbohydrates was assessed using other pNP monosaccharides. This class of compounds is usually used to assay glycosidase activity and specificity, as the pNP group acts a a good leaving group, and

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447−1455

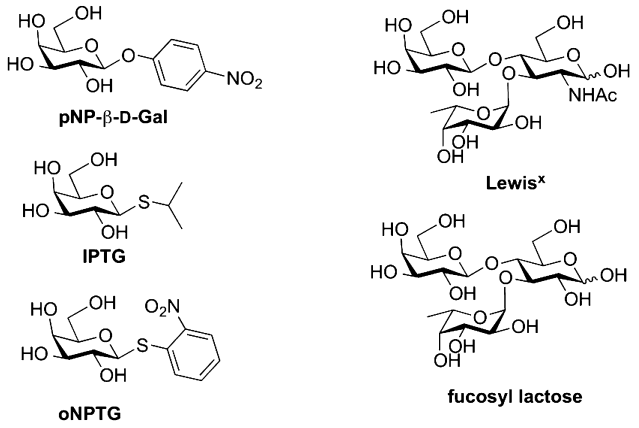


Figure 2. Structures of substrates and inhibitors of BT-2192.

can be easily quantified by spectrophotometry. Eight compounds did not exhibit detectable activity after incubation for 3 h, including pNP-β-L-fucopyranose, pNP-β-D-fucopyranose, pNP-α-D-glucopyranose, pNP-β-D-glucopyranose, pNP-α-D-mannopyranose, pNP-β-D-mannopyranose, pNP-α-D-galactopyranose, and pNP-β-D-galactofuranose (Figure 1A). Unexpectedly, pNP-β-D-galactopyranose (pNP-Gal) appeared to be the only compound hydrolyzed by BT2192. This specificity is closely related to the stereochemistry of the carbohydrate. As pNP-α-L-fucopyranose and pNP-β-D-glucopyranose are not hydrolyzed by BT2192, one can hypothesize that the presence of a 6-OH and the C4 axial epimer are required for efficient catalysis. Moreover, only the β anomer of pNP-D-galactopyranose is hydrolyzed by BT2192. Lactose (Lac) was also tested as a substrate and was not hydrolyzed by purified BT2192 (Figure 1B, lane 2). This lack of lactose hydrolysis activity is a good indicator of the absence of the remaining *E. coli* endogenous β-galactosidase (*lacZ*) in the BT2192 purified samples, as lactose is also a substrate for *lacZ*.[33] The presence of Lac did not prevent the hydrolysis of pNP-Gal at an equimolar concentration (lane 7). Finally, as shown in Figure 1B, BT2192 β-D-galactosidase activity is hardly detectable when using Lewis^X antigen as the substrate (lanes 11 and 15). This indicates that when an oligosaccharide bearing both a strained α-L-fucose and a branched β-D-galactose is assayed as a substrate, only α-L-fucosidase activity can be detected (unlike β-D-galactosidase activity).

Characterization of the β-D-Galactosidase Activity of BT2192. Initial optimization of the temperature for efficient pNP-Gal hydrolysis was performed, which produced an optimal temperature of 45 °C (Figure S1 of the Supporting Information). However, to prevent excessive nonenzymatic hydrolysis of the substrate, a lower reaction temperature of 37 °C was used thereafter for the kinetic analysis and pH dependence of BT2192 activity. As the GH retaining mechanism involves two carboxylate residues, we assessed the pH dependence of the β-D-galactosidase reaction and found that BT2192 β-D-galactosidase activity was at a maximum in pH 7 buffer. Figure 3A represents the typical curve obtained for kinetic hydrolysis of pNP-Gal by BT2192. A fit of these data to the Michaelis–Menten equation gave a k_{cat} of 0.41 ± 0.02 min^{-1} and a K_M of 147 ± 23 μM. The K_M value is in the range of those reported for other β-D-galactosidases,[34−36] indicating that BT2192 exhibits a good affinity for pNP-Gal, yet the k_{cat} value is very small compared to generally observed GH catalytic constants on the order of magnitude of inverse seconds (http://www.brenda-enzymes.org). This low k_{cat} value is in contrast with the low activation energy of 3.5 kcal/mol, similar to reported values for other galactosidases.[37,38]

To assess the good affinity of pNP-Gal for BT2192, we tested two nonhydrolyzable substrate analogues as inhibitors. *o*-Nitrophenyl β-D-1-thiogalactopyranose (oNPTG) competitively inhibited BT2192 (Figure 3B), with an inhibition constant (K_I) of 147 ± 59 μM. IPTG was also demonstrated to be an inhibitor (K_I > 500 μM). Indeed as $K_I/K_M \leq 1$, these two compounds are weak inhibitors of pNP-Gal hydrolysis activity. Thus, the β-D-galactopyranosyl moiety is well recognized in the BT2192 active site, and structural studies were then performed to identify the interactions of the substrate with BT2192.

Structural Characterization of BT2192 in a Complex with IPTG and oNPTG. The X-ray structure of BT2192 was previously obtained by the New York Structural Genomics Research Consortium (http://www.sbkb.org) (PDB entry 3EYP). However, no ligand binding analyses were performed. Therefore, we determined the crystal structure of BT2192 in complex with IPTG (PDB entry 4OUE) and with oNPTG (PDB entry 4OZO). Ligand-free BT2192 crystals were soaked in ligand solutions (250 mM for IPTG and 5 mM for oNPTG), and structures were refined to 2.35 and 2.8 Å, respectively, using the reported ligand-free protein as model for molecular replacement. BT2192 consists of two domains (Figure 4). The catalytic N-terminal domain exhibits the canonical TIM-barrel $(\beta/\alpha)_8$ fold that was reported for other GH29 enzymes, and the C-terminal domain is a β-sandwich domain (Figure 4). The

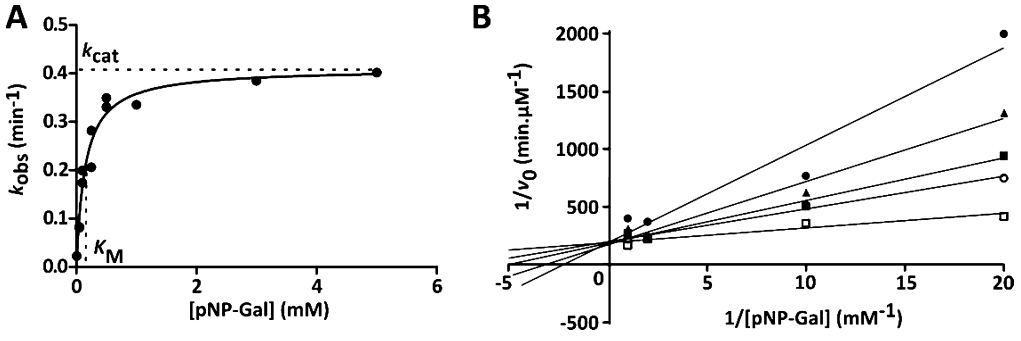


Figure 3. (A) Typical kinetic data for hydrolysis of pNP-Gal by BT2192. (B) Lineweaver–Burk plot for oNPTG competitive inhibition of hydrolysis of pNP-Gal by BT2192. Initial reaction rates (v_0) were determined in the absence (□) or presence of 0.1 (●), 0.25 (■), 0.5 (▲), or 1 mM (●) oNPTG.

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447−1455

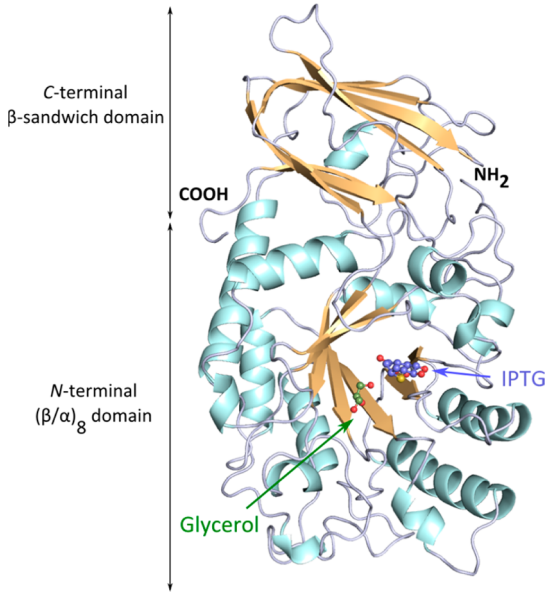


Figure 4. Overall structure of BT2192 in complex with IPTG and glycerol. Helices and sheets are rendered as cartoons and colored cyan and orange, respectively. IPTG and glycerol are depicted as ball and sticks, with carbon, oxygen, and sulfur atoms colored blue (green for glycerol), red, and yellow, respectively. N- and C-termini and structural domains are indicated.

catalytic domain can be structurally aligned with *T. maritima* (PDB entry 1HL8), *B. thetaiotaomicron* (PDB entries 3GZA and 2WVU), and *Bi. longum* (PDB entry 3MO4), with calculated rmsd values for Cα atoms between BT2192 and those structures of 2.4, 1.0, 1.2, and 0.9 Å, respectively. This structural alignment between GH29 structures allowed us to identify the two catalytic residues (Asp188 and Glu254) that have been identified and studied in other GH29 α-L-fucosidases.[9,14−17]

Electron density in both ligand-bound structures allowed the modeling of IPTG and oNPTG (Figure 5). As these two compounds are competitive inhibitors of β-D-galactosidase activity, their binding site is expected to be similar to that of the substrate pNP-Gal. Moreover, β-D-galactosides ($^4C_1$) do not have the same conformation as α-L-fucosides ($^1C_4$); thus, a second binding pocket for β-D-galactosides would allow efficient binding of these compounds (oNPTG has a K_I value of 150 μM). The binding site for β-D-1-galactopyranosyl compounds in BT2192 is located in a pocket (named hereafter the G pocket) different from the one that was expected for α-L-fucose (the F pocket), in comparison with GH29-containing substrate structures (Figure 5).[20] The two pockets are visible in BT2192−ligand structures, although glycerol is bound to the F pocket according to structural alignment with GH29 (Figure 6). The cryoprotectant used for freeze-drying of the BT2192 crystal was glycerol, which explains the presence of this ligand. In the previously reported BT2192 structure, glycerol was also bound to the G site. Nucleophile and general acid/base

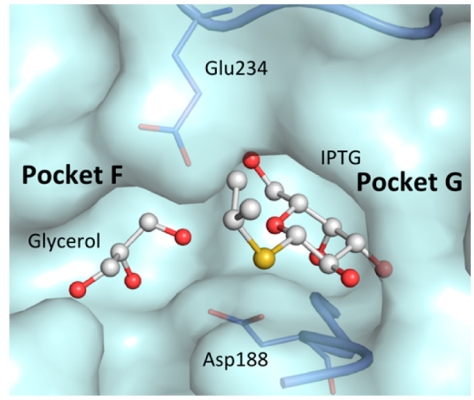


Figure 6. Binding pockets of glycerol and IPTG in the BT2192 active site. The van der Waals surface of BT2192 protein atoms is rendered and colored light green. GH29-like catalytic residues (Glu234 and Asp288) are rendered as dark blue sticks and delimit the two binding pockets for glycerol and IPTG. Ligands are depicted as white balls and sticks.

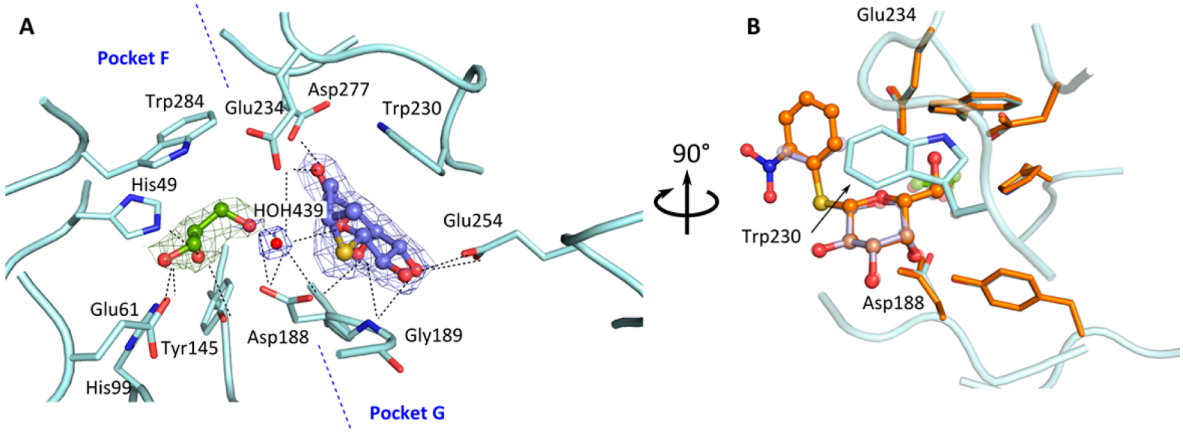


Figure 5. (A) Interactions among BT2192 active site residues, glycerol, and IPTG. Residues interacting with ligands are rendered as sky-blue sticks. Binding pockets F and G, as indicated in the text, are denoted. Glycerol, IPTG, and the interacting water molecule (HOH439) are depicted as balls and sticks. Carbons of IPTG and glycerol are colored blue and green, respectively. For all compounds, nitrogen, oxygen, and sulfur atoms are colored dark blue, red, and yellow, respectively. IPTG, glycerol, and the binding water are shown with the corresponding $2F_o - F_c$ omit map density contoured at 1.5σ in light blue, green, and dark blue, respectively. (B) Overlay of structures of BT2192−IPTG and BT2192−oNPTG complexes. BT2192 bound with IPTG is rendered and colored as in panel A. All carbon atoms of the BT2192−oNPTG complex are colored orange, and nitrogen, oxygen, and sulfur atoms are colored dark blue, red, and yellow, respectively. For the sake of clarity, the structure was rotated 90° and Glu254 was removed.

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447−1455

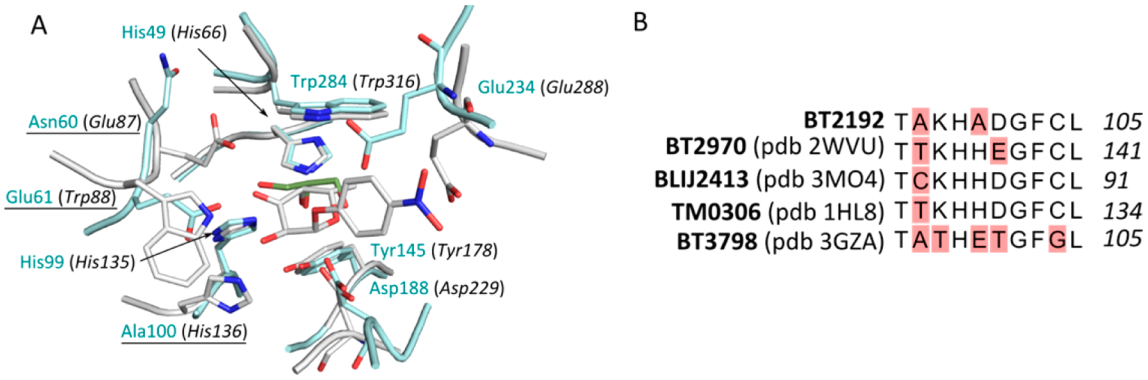


Figure 7. (A) Structural overlay of the glycerol binding site of BT2192 and another GH29 fucosidase bound to pNP-α-L-fucopyranose (PDB entry 2WVU, gene BT2970). Residues interacting with ligands are depicted as sticks and colored cyan for BT2192 and light gray for BT2970. Their name and sequence position are indicated and similarly colored. Nonhomologous residues found in BT2192 are underlined. Glycerol bound to BT2192 is colored green, and pNP-α-L-fucopyranose bound to BT2970 is colored light gray. (B) Peptidic sequence alignment of residues surrounding Ala100 of BT2192 with other GH29 PDB structures. Gene database identifications are given in bold, and PDB entries are given in parentheses. Nonhomologous residues are highlighted in pink, and the numbering refers to each sequence.

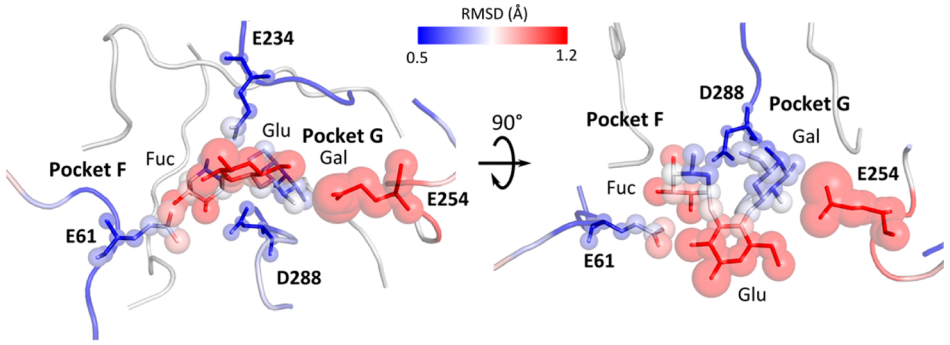


Figure 8. Flexibility of 3-fucosyllactose modeled in the BT2192 active site. Ligand and interacting carboxylate residues are represented as sticks. Surrounding residues are depicted as cartoons. The right panel is rotated 90° along the *y*-axis. For the sake of clarity, hydrogen atoms are not shown, and the peptidic region around E234 is not depicted in the right panel. The rmsd values calculated during MD simulations are rendered as red to blue scaled coloring of atoms, as indicated by the legend. For the sake of clarity, transparent spheres are also superimposed on the stick depiction, with a radius defined as the rmsd value calculated for each atom.

catalytic residues (Asp188 and Glu234, respectively) can be identified using sequence and structural alignment with GH29 fucosidases and delimit the separation between the two binding pockets. The recognition of the pyranosyl moiety of IPTG and oNPTG by BT2192 occurs though many stabilizing interactions (Figure 5), mainly with the four surrounding carboxylic acid side-chain residues: Asp188, Asp277, Glu234, and Glu254. These four residues form hydrogen bonds with the hydroxyl groups of the β-D-1-galactosyl derivative, either directly or via a water molecule (between Asp188 and the ligand O5 atom). Other hydrogen bonds between O3 and O4 and the amide nitrogen of Gly189 are visible. Trp230, located parallel with the hexose ring, also stabilizes the carbohydrate with hydrophobic interactions. Glycerol bound in the F pocket adopts a conformation mimicking half of the fucosyl ring as seen in other GH29–fucose structures (Figure 7). Several interactions are conserved between glycerol and identical residues in all GH29 structures: O1 forms a hydrogen bond with Asp188, O2 interacts with Tyr145 and His99, and O3 is hydrogen-bonded to His99. Considering structural and sequence alignment with GH29 enzymes that were extensively studied, Asp188 and Glu234 were expected to be the catalytic nucleophile and general acid/base, respectively (Figure S4 of the Supporting Information). However, the orientation of the β-D-1-galactopyranosyl moiety in the G pocket prevents Asp188 from attacking *anti* to the anomeric leaving group (Figure 5). Two other carboxylate residues, Glu234 and Glu254, are approximately equidistant from the anomeric carbon of β-D-1-galactopyranosyl (4.5 and 4.2 Å, respectively) and can be considered candidate nucleophiles. However, Glu254 is more likely to be the reactive nucleophile as the Glu234 carboxylate is located in the hexose plane and appears to be unable to form the required glycosyl–enzyme intermediate. Alternatively, a water molecule designated HOH439 visible in the highest-resolution structure, IPTG-bound BT2192, could be the second nucleophile attacking the intermediate to release the D-galactose. This water molecule could be deprotonated by either Asp188 or Glu234 (distance of 3.4 or 3.2 Å, respectively).

Conformational Mobility of Ligands in the BT2192 Active Site. To assess the flexibility of ligands bound to the active site and the reactivity of fucosidase and galactosidase of BT2192, 3-fucosyl lactose and Lewis^{X} [20] were modeled in the enzyme active site, and molecular dynamics were used to assess ligand flexibility, as reported previously for other GHs.[39,40] The galactosyl moiety of each substrate was originally placed according to the X-ray structures of galactosides bound to BT2192 reported in this work (in the G pocket). The fucosyl moiety was oriented in the F pocket, taking into account previously reported structures of GH29–fucoside complexes.[8]

dx.doi.org/10.1021/bi400951q | *Biochemistry* 2014, 53, 1447–1455

After several cycles of energy minimization and molecular dynamics to equilibrate the complexes, 10 independent MD simulations lasting 0.5 ns were conducted for both substrates. For each atom, the rmsd was calculated over the 10 minimized models, and values obtained for fucosyllactose−BT2192 complexes are presented in Figure 8. For Lewis^X, similar rmsd values were calculated. The galactosyl moiety of the substrate is the most anchored to the G binding site. Indeed, during the MD simulations, this part of the ligand appears to be the most rigid (rmsd ∼ 0.5 Å). All the interactions described in the X-ray structures of the galactoside−BT2192 complex are maintained in all MD simulations, indicating a strong binding of galactose in its pocket. Unlike the galactosyl residue, the central glucosyl part of the ligand is the most mobile (rmsd ∼ 1−1.2 Å), as no specific interactions could be seen with BT2192 residues. Most interestingly, half of the fucoside moiety exhibits a high rmsd. Indeed, C2, C3, and C4 atoms and corresponding hydroxyls are not stable during the simulations. This study of rmsd evolution during MD simulations highlights the strong anchoring of the substrate in the BT2192 active site through the galactosyl part of the molecule, whereas the fucosyl moiety is flexible in the BT2192 active site. Another relevant aspect is the high mobility of Glu254 that is predicted to act as the nucleophile for β-D-galactosidase activity. This high mobility may partly explain the absence of galactosidase activity when the trisaccharide substrate is present in the active site.

Similar docking and MD simulations of pNP-Fuc and pNP-Gal in pockets F and G, respectively, indicated that these two compounds can bind simultaneously to BT2192 and that the two active sites F and G may act independently.

DISCUSSION

The GH29 family consists of α-L-fucosidase enzymes,[7,9,11,12,14,16] which can be classified into two groups according to their substrate specificity.[19] In this study, we have confirmed a previous report that BT2192 belonged to the GH29-B family.[20] This family includes enzymes that do not possess pNP-α-L-fucopyranose hydrolase activity but can hydrolyze α(1→3/4) fucosidic linkage-containing compounds. We have demonstrated that BT2192 exhibits weak β-D-galactosidase activity when using pNP-Gal as a substrate, which has not been reported for other GH29 enzymes. This activity could not be attributed to endogenous *E. coli* galactosidase, as no lactose hydrolysis could be observed on the purified BT2192 sample. The absence of a galactose hydrolysis product when using lactose or Lewis^X antigen as substrates indicates that BT2192 needs an aglycone that is a good leaving group such as *p*-nitrophenoxyl, and that hydrolysis cannot occur when the galactose forms a glycosidic bond with another oligosaccharide. Although the weak β-D-galactosidase activity observed cannot be linked to a biological function of BT2192, the interaction of a branched galactose is crucial for efficient fucosidase activity.

X-ray crystallography revealed that BT2192 exhibits two binding pockets, which we call F and G, and the former is structurally identical to other GH29 fucosidase active sites. However, this binding site lacks three amino acid residues that are found in almost all other GH29 structures (Figure 7) and are involved in fucose recognition. For instance, Asn60 in BT2192 is found to be a glutamate, which forms hydrogen bonds with O3 of α-L-fucose. In the BT2192 X-ray structure, the side chain of Asn60 is flipped and cannot form such stabilizing interactions. A second residue, Glu61, is replaced with tryptophan in other GH29 enzymes and interacts with the fucose O2 atom. Although Glu61 also interacts with the glycerol bound in the F pocket in BT2192, the change in the physicochemical properties of a negatively charged (glutamate) versus an aromatic residue (tryptophan) will lead to a dramatic change in the substrate recognition interactions involving this residue. Finally, the third residue not found in other GH29 enzymes is Ala100. In other characterized GH29 enzymes, this residue is a histidine and interacts strongly with the O2 atom of fucose through a direct hydrogen bond. Sequence alignment of BT2192 with all characterized GH29 fucosidases extracted from the CAZy database shows that only BT2192 exhibits these particular mutations, especially the missing histidine at position 100, which seems to be crucial for fucose recognition[6,15] (Figure 7B and Figure S2 of the Supporting Information).

The binding site for the β-D-galactosyl moiety of IPTG and oNPTG was found to be in a second cavity in the BT2192 active site. In *T. maritima* fucosidase, mutations close to this pocket led to the modified activity and specificity of the acceptor during transfucosylation.[41] Structural and sequential comparisons among all members of the GH29 family indicate that the G pocket is not conserved in the GH29 family. The nature of the amino acids, as well as the three-dimensional relationship of the peptide backbone narrowing this pocket, is divergent. For BT2192, the G binding site can efficiently accommodate and bind the β-D-galactopyranosyl moiety. These interactions are confirmed *in vitro* by the submillimolar K_M and K_I values determined for β-D-galactosyl-containing compounds. The complexed structures also indicated that the slow reaction rate calculated for pNP-Gal hydrolysis could be explained by the positioning of the catalytic residues. Alignment of BT2192 with GH29 enzymes showed that the two carboxylates Asp188 and Glu234 are expected to be the catalytic amino acids for the fucosidase activity.[17] However, the spatial orientation of IPTG and oNPTG clearly demonstrates that these two residues cannot act as a nucleophile and a general acid/base for pNP-Gal hydrolysis. The identity of the residues involved in β-D-galactosidase activity is still being investigated. A third residue, Glu254, is more likely to be the nucleophilic residue and attack in a position *anti* to the leaving group. The orientation of this residue still does not favor an efficient catalysis, as the nucleophile is not properly aligned with the leaving group.[4] Among all GH29 structures, BT2192 is the only one exhibiting a carboxylate side-chain residue in this pocket. Because of the great diversity in sequence and structure on this part of the protein in the GH29 family, there is no clear evidence that BT2192 is the only GH29 representative having such a residue in this position. Moreover, no structural or sequential alignment with β-D-galactosidases could be achieved with the binding site of the β-D-galactopyranose moiety.

B. thetaiotaomicron is peculiar in that more than 200 of its genes encode GH enzymes. Among them, nine encode GH29 enzymes. To the best of our knowledge, only two of these proteins have been characterized as α-L-fucosidases[8] whose expression was reported to be up- or downregulated in all transcriptomic analyses of the *Bacteroides* genome.[42−44] On the contrary, BT2192 expression has not been detected in such analyses (and the same is true of its homologues in *Bacillus subtilis*), indicating that this gene is not expressed constitutively.

MD simulations suggest that the bound galactosyl moiety is fixed, while fucose appears to be mobile, due to the loss of several strong interactions like the hydrogen bond with Ala100 (a highly conserved histidine in all other GH29 enzymes).

Thus, the structural feature needed for substrate recognition is mainly the presence of the branched galactose in the substrate oligosaccharide, which was predicted previously.[20] Unlike the binding of galactose to the G site, binding of fucose to the F active site is weaker, and its contribution to substrate binding may be small, although BT2192 is classified as an α-L-fucosidase. Indeed, Sakurama and co-workers could not measure a proper K_M value for 3-fucosyllactose hydrolysis, with substrate concentration of ≤20 mM.[20]

To conclude, this study highlights the critical role of all the pockets constituting the binding site in glycosyl hydrolase specificity. Indeed, BT2192 can act as an α-L-fucosidase only when a substrate bearing a branched β-D-galactose can bind to a second enzyme binding site, thus sufficiently stabilizing the enzyme−substrate complex to allow efficient catalysis.

■ ASSOCIATED CONTENT

Supporting Information

pH and temperature dependence of pNP-Gal hydrolysis (Figure S1) and multiple-sequence alignment of BT2192 with GH29 family fucosidases (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: richard.daniellou@univ-orleans.fr. Telephone: +33 2 38 49 49 78.

Funding

This work was supported by Région Centre.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. L. Stead for the generous gift of the pSGX3-BT2192 plasmid and Dr. Coste for his help on crystallographic studies. We are grateful to the staff of the Soleil Proxima 1 beamline (Gif/Yvette, France) for their help in the collection of diffraction data. We also thank Pr. David Palmer (University of Saskatchewan, Saskatoon, SK) for thoroughly reading and correcting the manuscript.

■ ABBREVIATIONS

GH, glycosyl hydrolase; IPTG, isopropyl β-D-1-thiogalactopyranose; Lac, lactose; MD, molecular dynamics; oNPTG, *o*-nitrophenyl β-D-1-thiogalactopyranose; pNP, *p*-nitrophenol; pNP-Gal, *p*-nitrophenyl β-D-1-galactopyranose.

■ REFERENCES

(1) Lafite, P., and Daniellou, R. (2012) Rare and unusual glycosylation of peptides and proteins. *Nat. Prod. Rep. 29*, 729−738.

(2) Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J. 280* (Part 2), 309−316.

(3) Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res. 37*, D233−D238.

(4) Rye, C. S., and Withers, S. G. (2000) Glycosidase mechanisms. *Curr. Opin. Chem. Biol. 4*, 573−580.

(5) Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V., and Gordon, J. I. (2003) A Genomic View of the Human-*Bacteroides thetaiotaomicron* Symbiosis. *Science 299*, 2074−2076.

(6) Sulzenbacher, G., Bignon, C., Nishimura, T., Tarling, C. A., Withers, S. G., Henrissat, B., and Bourne, Y. (2004) Crystal Structure of *Thermotoga maritima* α-L-Fucosidase: Insights into the catalytic mechanism and the molecular basis for fucosidosis. *J. Biol. Chem. 279*, 13119−13128.

(7) Sela, D. A., Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H. J., Joachimiak, A., Lebrilla, C. B., and Mills, D. A. (2012) *Bifidobacterium longum* subsp. *infantis* ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides. *Appl. Environ. Microbiol. 78*, 795−803.

(8) Lammerts van Bueren, A., Ardèvol, A., Fayers-Kerr, J., Luo, B., Zhang, Y., Sollogoub, M., Blériot, Y., Rovira, C., and Davies, G. J. (2010) Analysis of the Reaction Coordinate of α-L-Fucosidases: A Combined Structural and Quantum Mechanical Approach. *J. Am. Chem. Soc. 132*, 1804−1806.

(9) Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M., and Moracci, M. (2003) Identification of the Catalytic Nucleophile of the Family 29 α-L-Fucosidase from *Sulfolobus solfataricus* via Chemical Rescue of an Inactive Mutant. *Biochemistry 42*, 9525−9531.

(10) de la Torre, F., Sampedro, J., Zarra, I., and Revilla, G. (2002) AtFXG1, an *Arabidopsis* Gene Encoding α-L-Fucosidase Active against Fucosylated Xyloglucan Oligosaccharides. *Plant Physiol. 128*, 247−255.

(11) Zeleny, R., Leonard, R., Dorfner, G., Dalik, T., Kolarich, D., and Altmann, F. (2006) Molecular cloning and characterization of a plant α1,3/4-fucosidase based on sequence tags from almond fucosidase I. *Phytochemistry 67*, 641−648.

(12) Berteau, O., McCort, I., Goasdoué, N., Tissot, B., and Daniel, R. (2002) Characterization of a new α-L-fucosidase isolated from the marine mollusk *Pecten maximus* that catalyzes the hydrolysis of α-L-fucose from algal fucoidan (*Ascophyllum nodosum*). *Glycobiology 12*, 273−282.

(13) Fisher, K. J., and Aronson, N. N., Jr. (1989) Isolation and sequence analysis of a cDNA encoding rat liver α-L-fucosidase. *Biochem. J. 264*, 695−701.

(14) Tarling, C. A., He, S., Sulzenbacher, G., Bignon, C., Bourne, Y., Henrissat, B., and Withers, S. G. (2003) Identification of the Catalytic Nucleophile of the Family 29 α-L-Fucosidase from *Thermotoga maritima* through Trapping of a Covalent Glycosyl-Enzyme Intermediate and Mutagenesis. *J. Biol. Chem. 278*, 47394−47399.

(15) Cobucci-Ponzano, B., Mazzone, M., Rossi, M., and Moracci, M. (2005) Probing the Catalytically Essential Residues of the α-L-Fucosidase from the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. *Biochemistry 44*, 6331−6342.

(16) Liu, S.-W., Chen, C.-S., Chang, S.-S., Mong, K.-K. T., Lin, C.-H., Chang, C.-W., Tang, C. Y., and Li, Y.-K. (2009) Identification of Essential Residues of Human α-L-Fucosidase and Tests of Its Mechanism. *Biochemistry 48*, 110−120.

(17) Shaikh, F. A., Lammerts van Bueren, A., Davies, G. J., and Withers, S. G. (2013) Identifying the Catalytic Acid/Base in GH29 α-L-Fucosidase Subfamilies. *Biochemistry 52*, 5857−5864.

(18) Ashida, H., Miyake, A., Kiyohara, M., Wada, J., Yoshida, E., Kumagai, H., Katayama, T., and Yamamoto, K. (2009) Two distinct α-L-fucosidases from *Bifidobacterium bifidum* are essential for the utilization of fucosylated milk oligosaccharides and glycoconjugates. *Glycobiology 19*, 1010−1017.

(19) Yoshida, E., Sakurama, H., Kiyohara, M., Nakajima, M., Kitaoka, M., Ashida, H., Hirose, J., Katayama, T., Yamamoto, K., and Kumagai, H. (2012) *Bifidobacterium longum* subsp. *infantis* uses two different β-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. *Glycobiology 22*, 361−368.

(20) Sakurama, H., Tsutsumi, E., Ashida, H., Katayama, T., Yamamoto, K., and Kumagai, H. (2012) Differences in the substrate specificities and active-site structures of two α-L-fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. *Biosci., Biotechnol., Biochem. 76*, 1022−1024.

(21) Cornish-Bowden, A. (2012) *Fundamentals of Enzyme Kinetics*, 4th ed., Wiley-Blackwell, Weinheim, Germany.

(22) Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res. 38*, W695−W699.
(23) Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol. 7*, 539.
(24) Kabsch, W. (2010) XDS. *Acta Crystallogr. D66*, 125−132.
(25) Vagin, A., and Teplyakov, A. (1997) MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Crystallogr. 30*, 1022−1025.
(26) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.
(27) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D66*, 486−501.
(28) Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D66*, 12−21.
(29) Wesener, D. A., May, J. F., Huffman, E. M., and Kiessling, L. L. (2013) UDP-Galactopyranose Mutase in Nematodes. *Biochemistry 52*, 4391−4398.
(30) Horler, R. S. P., Müller, A., Williamson, D. C., Potts, J. R., Wilson, K. S., and Thomas, G. H. (2009) Furanose-specific Sugar Transport: Characterization of a bacterial galactofuranose-binding protein. *J. Biol. Chem. 284*, 31156−31163.
(31) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(32) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(33) Lederberg, J. (1950) The β-D-galactosidase of *Escherishia coli*, Strain K-12. *J. Bacteriol. 60*, 381−392.
(34) Martino, S., Tiribuzi, R., Tortori, A., Conti, D., Visigalli, I., Lattanzi, A., Biffi, A., Gritti, A., and Orlacchio, A. (2009) Specific determination of β-galactocerebrosidase activity via competitive inhibition of β-galactosidase. *Clin. Chem. 55*, 541−548.
(35) Juers, D. H., Rob, B., Dugdale, M. L., Rahimzadeh, N., Giang, C., Lee, M., Matthews, B. W., and Huber, R. E. (2009) Direct and indirect roles of His-418 in metal binding and in the activity of β-galactosidase (*E. coli*). *Protein Sci. 18*, 1281−1292.
(36) Hall, J. C., Jacobetz, D. R., LaMarche, M. D., Kochins, J. G., and Tubbs, C. E. (1997) Purification, characterization, and expression of rat epididymal β-D-galactosidase. *Biochem. Mol. Biol. Int. 42*, 443−451.
(37) Kishore, D., and Kayastha, A. M. (2012) A β-galactosidase from chick pea (*Cicer arietinum*) seeds: Its purification, biochemical properties and industrial applications. *Food Chem. 134*, 1113−1122.
(38) Craig, D. B., and Chase, L. N. (2012) Arrhenius plot for a reaction catalyzed by a single molecule of β-galactosidase. *Anal. Chem. 84*, 2044−2047.
(39) Dunstan, M. S., Barkauskaite, E., Lafite, P., Knezevic, C. E., Brassington, A., Ahel, M., Hergenrother, P. J., Leys, D., and Ahel, I. (2012) Structure and mechanism of a canonical poly(ADP-ribose) glycohydrolase. *Nat. Commun. 3*, 878.
(40) Slade, D., Dunstan, M. S., Barkauskaite, E., Weston, R., Lafite, P., Dixon, N., Ahel, M., Leys, D., and Ahel, I. (2011) The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature 477*, 616−620.
(41) Osanjo, G., Dion, M., Drone, J., Solleux, C., Tran, V., Rabiller, C., and Tellier, C. (2007) Directed Evolution of the α-L-Fucosidase from *Thermotoga maritima* into an α-L-Transfucosidase. *Biochemistry 46*, 1022−1033.
(42) Bjursell, M. K., Martens, E. C., and Gordon, J. I. (2006) Functional Genomic and Metabolic Studies of the Adaptations of a Prominent Adult Human Gut Symbiont, *Bacteroides thetaiotaomicron*, to the Suckling Period. *J. Biol. Chem. 281*, 36269−36279.
(43) Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C.-H., Westover, B. P., Weatherford, J., Buhler, J. D., and Gordon, J. I. (2005) Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. *Science 307*, 1955−1959.
(44) Sund, C. J., Rocha, E. R., Tzianabos, A. O., Wells, W. G., Gee, J. M., Reott, M. A., O'Rourke, D. P., and Smith, C. J. (2008) The *Bacteroides fragilis* transcriptome response to oxygen and H_2O_2: The role of OxyR and its effect on survival and virulence. *Mol. Microbiol. 67*, 129−142.