THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 17, pp. 11795–11806, April 26, 2013

Crystal Structures of a Glycoside Hydrolase Family 20 Lacto-*N*-biosidase from *Bifidobacterium bifidum*[*] [S]

Received for publication, September 15, 2012, and in revised form, March 11, 2013 Published, JBC Papers in Press, March 11, 2013, DOI 10.1074/jbc.M112.420109

Tasuku Ito[‡], Takane Katayama[§], Mitchell Hattie[¶1], Haruko Sakurama[§2], Jun Wada[‖3], Ryuichiro Suzuki[‡4], Hisashi Ashida[**], Takayoshi Wakagi[‡], Kenji Yamamoto[§], Keith A. Stubbs[¶5], and Shinya Fushinobu[‡6]

From the [‡]Department of Biotechnology, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, the [§]Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Nonoichi, Ishikawa 921-8836, Japan, the [¶]School of Chemistry and Biochemistry, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia, the [‖]Graduate School of Biostudies, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan, and the [**]Department of Science and Technology on Food Safety, Faculty of Biology-oriented Science and Technology, Kinki University, Kinokawa, Wakayama 649-6493, Japan

Background: Infant gut-associated bifidobacteria possess lacto-*N*-biosidase, which releases lacto-*N*-biose I (LNB) from human milk oligosaccharides.
Results: The crystal structures of lacto-*N*-biosidase complexed with LNB and LNB-thiazoline were determined.
Conclusion: The intermediate analog complex allows the proposal of a conformational reaction coordinate.
Significance: The structures of a key enzyme in the colonization of human commensal bacteria provided its structural basis and insight into the development of inhibitors.

Human milk oligosaccharides contain a large variety of oligosaccharides, of which lacto-*N*-biose I (Gal-β1,3-GlcNAc; LNB) predominates as a major core structure. A unique metabolic pathway specific for LNB has recently been identified in the human commensal bifidobacteria. Several strains of infant gut-associated bifidobacteria possess lacto-*N*-biosidase, a membrane-anchored extracellular enzyme, that liberates LNB from the nonreducing end of human milk oligosaccharides and plays a key role in the metabolic pathway of these compounds. Lacto-*N*-biosidase belongs to the glycoside hydrolase family 20, and its reaction proceeds via a substrate-assisted catalytic mechanism. Several crystal structures of GH20 β-*N*-acetylhexosaminidases, which release monosaccharide GlcNAc from its substrate, have been determined, but to date, a structure of lacto-*N*-biosidase is unknown. Here, we have determined the first three-dimensional structures of lacto-*N*-biosidase from *Bifidobacterium bifidum* JCM1254 in complex with LNB and LNB-thiazoline (Gal-β1,3-GlcNAc-thiazoline) at 1.8-Å resolution. Lacto-*N*-biosidase consists of three domains, and the C-terminal domain has a unique β-trefoil-like fold. Compared with other β-*N*-acetylhexosaminidases, lacto-*N*-biosidase has a wide substrate-binding pocket with a −2 subsite specific for β-1,3-linked Gal, and the residues responsible for Gal recognition were identified. The bound ligands are recognized by extensive hydrogen bonds at all of their hydroxyls consistent with the enzyme's strict substrate specificity for the LNB moiety. The GlcNAc sugar ring of LNB is in a distorted conformation near ^4E, whereas that of LNB-thiazoline is in a $^4C_1$ conformation. A possible conformational pathway for the lacto-*N*-biosidase reaction is discussed.

Human milk contains various oligosaccharides collectively termed human milk oligosaccharides (HMOs)[7] (1, 2). HMOs function as prebiotics, promoting the growth of bifidobacteria in the gastrointestinal tracts of breastfed infants, which in turn promotes optimal health (3, 4). Most HMOs contain a lactose moiety (Gal-β1,4-Glc) at their reducing end, which is elongated by β1,3-linked lacto-*N*-biose I (Gal-β1,3-GlcNAc; LNB, to give type I HMOs) and/or β1,3/6-linked *N*-acetyllactosamine (Gal-β1,4-GlcNAc; LacNAc, to give type II HMOs). A unique feature of the composition of HMOs is the predominance of type I over type II oligosaccharides. Indeed, such a composition has not been observed in milk oligosaccharides from other mammals, including anthropoids (5). Further elongation of

* This work was supported in part by the Program for the Promotion of Basic Research Activities for Innovative Bioscience, Japan (to S. F., K. Y., and T. K.), by Japan Society for the Promotion of Science KAKENHI Grant 24380053 (to S. F. and H. A.), and by the Australian Research Council. Part of this work was performed under the Priority Program for Disaster-affected Quantum Beam Facilities Proposals PF 2011G080 and SPring-8 2011A1908.
[S] This article contains supplemental Tables S1 and S2, Figs. S1–S4, and additional references.
The atomic coordinates and structure factors (codes 4H04 and 4JAW) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Supported by an Australian Postgraduate Award from the University of Western Australia and a Jean Rogerson Postgraduate Scholarship.
[2] Present address: Graduate School of Agriculture, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto, 606-8502, Japan.
[3] Present address: Kyoto Municipal Institute of Industrial Technology and Culture, Chudoji Awatacho, Shimogyo-ku, Kyoto 600-8815, Japan.
[4] Present address: Dept. of Biological Production, Faculty of Bioresource Sciences, Akita Prefectural University, Akita 010-0195, Japan.
[5] Recipient of funding from the Australian Research Council.
[6] To whom correspondence should be addressed. Tel.: 81-3-5841-5151; Fax: 81-3-5841-5151; E-mail: asfushi@mail.ecc.u-tokyo.ac.jp.

[7] The abbreviations used are: HMO, human milk oligosaccharide; LNB, lacto-*N*-biose I, (Gal-β1,3-GlcNAc); LacNAc, *N*-acetyllactosamine, (Gal-β1,4-GlcNAc); LNFP I, lacto-*N*-fucopentaose I (Fuc-α1,2-Gal-β1,3-GlcNAc-β1,3-Gal-β1,4-Glc); LNDFH I, lacto-*N*-difucohexaose I (Fuc-α1,2-Gal-β1,3-(Fuc-α1,4-)GlcNAc-β1,3-Gal-β1,4-Glc); LNT, lacto-*N*-tetraose (Gal-β1,3-GlcNAc-β1,3-Gal-β1,4-Glc); GNB, galacto-*N*-biose (Gal-β1,3-GalNAc); LNBase, lacto-*N*-biosidase; *Bb*LNBase, LNBase from *B. bifidum* JCM1254; GH, glycoside hydrolase; CBM, carbohydrate-binding module; β-HexNAcase, β-*N*-acetylhexosaminidase; *p*NP, *p*-nitrophenyl; PA, pyridylamino; r.m.s.d., root mean square deviation; *Sp*Hex, β-HexNAcase from *Streptomyces plicatus*; PUGNAc, *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)-amino-*N*-phenylcarbamate.

these core structures is made by the addition of fucose and sialic acid residues via α1,2/3/4- and α2,3/6-linkages, respectively. HMOs are composed of more than 130 different oligosaccharide structures that account for 2% of the solid components of dried human milk. Of note is that four oligosaccharides constitute 25–33% of total HMOs (6) as follows: 2′-fucosyl-lactose (Fuc-α1,2-Gal-β1,4-Glc); lacto-*N*-fucopentaose I (Fuc-α1,2-Gal-β1,3-GlcNAc-β1,3-Gal-β1,4-Glc; LNFP I); lacto-*N*-difucohexaose I (Fuc-α1,2-Gal-β1,3-(Fuc-α1,4-)GlcNAc-β1,3-Gal-β1,4-Glc; LNDFH I), and lacto-*N*-tetraose (Gal-β1,3-GlcNAc-β1,3-Gal-β1,4-Glc; LNT). Three of these oligosaccharides (LNFP I, LNDFH I, and LNT) contain an LNB unit, highlighting the importance of this component in these systems.

In 2005, a novel metabolic pathway specific to LNB and galacto-*N*-biose (Gal-β1,3-GalNAc; GNB) was uncovered in bifidobacteria (7) and was termed the GNB/LNB pathway (8). Considering the living environment of bifidobacteria (the gastrointestinal tract of infants), LNB and GNB are hypothesized to originate from HMOs and intestinal mucin glycoproteins, respectively. Proteins related to this pathway have been actively investigated, and crystallographic analyses (9) have uncovered the structures of an extracellular endo-α-*N*-acetylgalactosaminidase that releases GNB from mucin-type *O*-glycans (10), a solute-binding protein of an ABC transporter specific to GNB and LNB (11), and an intracellular phosphorylase that cleaves both GNB and LNB (12). In addition, it has been demonstrated that LNB is the "bifidus factor" that selectively promotes growth of infant gut-associated bifidobacteria (13, 14), but interestingly, a free form of LNB has not been found in HMOs (1).

Lacto-*N*-biosidase (LNBase, EC 3.2.1.140), which liberates LNB from the nonreducing end of oligosaccharides, was first found in the soil actinomycete *Streptomyces* sp. 142 (15, 16). Since then, LNBase activity has also been found in several strains of *Bifidobacterium bifidum* and *Bifidobacterium longum* (17), which are particularly predominant in the intestines of infants (18). Bifidobacterial LNBase is a membrane-anchored extracellular enzyme, which suggests that it may play a key role in the excision of LNB from HMOs to supply LNB to the GNB/LNB pathway. Recently, the gene for LNBase from *B. bifidum* JCM1254 (*Bb*LNBase) was cloned, and its recombinant protein was characterized in detail. *Bb*LNBase consists of 1,112 amino acids and contains a signal sequence, a glycoside hydrolase (GH) family 20 domain, a carbohydrate-binding module (CBM) family 32 domain, a bacterial Ig-like domain, and a transmembrane region (Fig. 1*A*). The enzyme's activity favored LNT as a substrate to produce LNB and lactose, and it was found that it could not act on oligosaccharides where the LNB unit is modified with fucose. In addition, it was found that *Bb*LNBase specifically releases LNB at the nonreducing end of type I oligosaccharides but does not hydrolyze type II oligosaccharides. Therefore, a potential use for this enzyme is in identifying type I structures in glycoconjugates. Intriguingly, most of the cancer-associated oligosaccharide antigens (including sialyl Le^a, sialyl Le^x, and their derivatives) have a core structure containing type I or type II chains (19).

LNBases are members of the GH20 family of glycoside hydrolases (20) along with β-*N*-acetylhexosaminidases (β-HexNAcases); however, they exhibit very low sequence homology. For example, *Bb*LNBase exhibits less than 24% amino sequence identity to all the characterized β-HexNAcases. GH20 enzymes cleave the glycosidic linkage at the reducing end of GlcNAc via a retaining substrate-assisted catalytic mechanism in which the 2-acetamido group of the substrate acts as the catalytic nucleophile (21–23). Whereas β-HexNAcases release a monosaccharide (GlcNAc), LNBase releases a disaccharide (LNB), implying that the latter has an extended −2 subsite. The crystal structures of multiple GH20 β-HexNAcases have been reported to date (21, 24–32); however, the three-dimensional structure of LNBase is not yet available.

Here, we report two crystal structures of *Bb*LNBase complexed with LNB and LNB-thiazoline, a potent inhibitor of *Bb*LNBase (K_i = 125 ± 8 nM at pH 4.5) (33). Using the sugar ring conformations of LNB and LNB-thiazoline molecules found in the active site, the reaction mechanism and possible conformational changes of the substrate are discussed.

EXPERIMENTAL PROCEDURES

Protein Production and Purification—The overexpression vector for N-terminally His_6-tagged *Bb*LNBase (residues 41–663) was constructed by inserting the PCR-amplified fragment of the *lnbB* gene (17) into the NdeI and EcoRI sites of the pET28b plasmid (*kan*^{r}; Novagen, Madison, WI). The primers used were 5′-gggaattccatatggggtacagtgccacggctccc-3′ and 5′-ccggaattctcagtcgctgaccaggtcag-5′ (restriction sites are underlined). The plasmid was introduced into *Escherichia coli* BL21 CodonPlus (DE3)-RIL (Stratagene, La Jolla, CA) for native protein expression. For selenomethionine-labeled protein expression, an NcoI-XhoI fragment of the pET28b-based expression plasmid was inserted into the pET19b plasmid (*amp*^{r}), which was subsequently introduced into *E. coli* BL21 CodonPlus (DE3)-RIL-X (*kan*^{r}; Stratagene). The transformants were cultured in Luria-Bertani medium (native protein) or LeMaster medium (selenomethionine-labeled protein) containing 100 mg/liter kanamycin or ampicillin and 20 mg/liter chloramphenicol at 25 °C for 20 h. Isopropyl 1-thio-β-D-galactopyranoside was added to a final concentration of 1.0 mM to induce protein expression. Following an additional incubation at 25 °C for 20 h, the cells were harvested by centrifugation and suspended in 50 mM HEPES-NaOH (pH 7.5). Cell extracts were obtained by sonication followed by centrifugation to remove cell debris. The protein was purified to homogeneity by sequential column chromatography involving nickel-nitrilotriacetic acid superflow (Qiagen, Hilden, Germany), Mono Q 10/100 GL, and Superdex 200 pg 16/60 column chromatography (GE Healthcare). The protein concentration was determined using the BCA protein assay kit (Thermo Fisher Scientific) with bovine serum albumin as a standard.

Crystallography—Crystals of both native and selenomethionine-labeled *Bb*LNBase complexed with LNB were obtained at 20 °C using the sitting drop vapor diffusion method, by mixing 0.5 μl of protein solution containing 7 mg/ml *Bb*LNBase and 10 mM LNB with an equal volume of reservoir solution (0.2 M potassium sodium tartrate tetrahydrate, 0.1 M sodium citrate

(pH 5.6), and 2.0 M ammonium sulfate). Crystals of *Bb*LNBase complexed with LNB-thiazoline (synthesized as described previously (33)) were obtained in a similar manner, except that the concentration of LNB-thiazoline used was 0.1 mM. Diffraction data were collected at 100 K using a charge-coupled device camera on beamline BL17A at the Photon Factory of the High Energy Accelerator Research Organization (KEK, Tsukuba, Japan) and processed using HKL2000 (34). The selenomethionine sites were determined using SnB (35), and the initial phases were calculated using Solve/Resolve (36). Automated model building was performed using ARP/wARP (37). Manual model rebuilding and refinement was achieved using Coot (38) and REFMAC5 (39). The statistics for data collection and refinement are provided in Table 1. Molecular graphic images were prepared using PyMOL (DeLano Scientific, Palo Alto, CA).

Enzymatic Characterization of Wild-type and Mutant LNBases—C-terminally His_6-tagged proteins (residues 35 to 1064) were used for kinetic analysis. Purification and expression of wild-type enzyme were carried out as described previously (17). H263F, D320A, D320N, and Y419F mutants were constructed by using the QuikChange site-directed mutagenesis method (Stratagene) with the plasmid pET23b-*lnbB* as the template. The following primers and their complementary strands were used: 5′-aactccccgggcttcatgaacgtctgg-3′ (H263F), 5′-cacatgggcgccgcggagtacatgatc-3′ (D320A), 5′-cacatgggcgccaacgagtacatgatc-3′ (D320N), and 5′-acccaggccctgttctggtccgttcg-3′ (Y419F). The entire sequence used for later manipulation was sequenced to check that no base change other than those designed had occurred. The mutants were expressed and purified using a similar procedure as described for the wild-type enzyme.

Activity measurements of the enzymes were conducted using *p*-nitrophenyl (*p*NP)-LNB (Sigma) as a substrate. The assay mixture (50 μl) contained substrate dissolved in 50 mM McIlvaine buffer (pH 4.5) and enzymes (13 nM for WT, 1.3 μM for H263F and Y419F, and 6.3 μM for D320A and D320N). The substrate concentrations were varied from 0.3- to 2-fold of the K_m values of the respective enzymes. The reaction was carried out at 30 °C and stopped by adding an equal volume of 1 M Na_2CO_3. The amounts of liberated 4-nitrophenolate were determined by measuring the absorbance at 400 nm.

The hydrolysis of LNT (Dextra Laboratory, Reading, UK) by wild-type enzyme was monitored by high performance anion-exchange chromatography with a CarboPac PA1 column, followed by pulsed amperometric detection (Dionex ICS3000). The elution was performed by a linear gradient of 0–0.5 M sodium acetate in 125 mM NaOH at a flow rate of 1 ml/min for 30 min. The kinetic parameters were calculated by curve-fitting the experimental data using the Michaelis-Menten equation in KaleidaGraph 4.0 (Synergy Software).

RESULTS AND DISCUSSION

Structural Determination of BbLNBase—We first constructed various deletion mutants to determine a minimal region that retained activity toward the substrate *p*NP-LNB (supplemental Table S1 and supplemental Fig. S1). Constructs 37–663, 41–663, and 46–663 (numbers indicate the residues) exhibited virtually the same activity as the full-length construct (31–1064, which has only the signal and membrane anchor regions deleted). Thus, these three constructs were selected for crystallization screening, but only diffraction-quality crystals were obtained with the 41–663 construct in the presence of LNB. The crystal structure of *Bb*LNBase was determined by the single-wavelength anomalous dispersion method using a selenomethionine derivative. Subsequently, we determined the crystal structure of native (nonlabeled) protein crystals (complexed with LNB and LNB-thiazoline) both at 1.8-Å resolution (Table 1 and Fig. 1*B*). The crystal contains two molecules in the asymmetric unit, and the final model contains residues from Ser-30 to Ser-662 of the A chain and from Ser-30 to Val-661 of the B chain. The protein has an N-terminal His_6 tag containing 21 amino acid residues derived from the pET-28b vector (MGSSHHHHHHSSGLVPRGSHM), in which 11 amino acids (SSGLVPRGSHM, residues 30–40) were visible in the electron density map. A region of the artificial tag sequence (Ser-30 to Arg-36) contributes to the packing of the dimer in the asymmetric unit (supplemental Fig. S2). The molecular masses of the 41–663 construct of *Bb*LNBase as deduced from the amino acid sequence, estimated by SDS-PAGE and calibrated gel filtration chromatography, were 71.8, 72.4, and 64.9 kDa, respectively (data not shown), suggesting that it is monomeric in solution. The root mean square deviations (r.m.s.d.) for the Cα atoms of all pairs of the four molecules (two chains in the two crystal structures) are less than 0.5 Å. The two ligand molecules bound to chains A and B are also virtually identical (r.m.s.d. = 0.047 and 0.069 Å for LNB and LNB-thiazoline, respectively). Hence, our subsequent descriptions will refer to chain A, unless otherwise noted.

Overall Structure—The *Bb*LNBase monomer consists of three domains as follows: an N-terminal domain (N-domain, residues 41–178); a catalytic $(\beta/\alpha)_8$ barrel domain (barrel domain, residues 179–496), and a C-terminal domain (C-domain, residues 497–662) (Fig. 1*B*). The N-domain has an α/β topology with a seven-stranded β-sheet exposed to the surface, and two α-helices buried in the interface with the subsequent barrel domain. The N-domain and the barrel domain correspond to the two conserved domains of typical GH20 β-HexNAcases (21, 24, 25, 27–32). The N-domain and the barrel domain of *Bb*LNBase are structurally similar to the domains I and II, respectively, of β-HexNAcase from *Streptomyces plicatus* (*Sp*Hex) (Fig. 1*C*). The N-domain is conserved in most GH20 enzymes, although its function remains unknown.

The C-domain, however, is not common to GH20 enzymes. Two examples of GH20 β-HexNAcases that have a distinct domain in the C-terminal region of the catalytic barrel domain are the β-HexNAcases from *Serratia marcescens* (chitobiase) and *Streptococcus gordonii* (GcnA), which have a small (67 amino acids) immunoglobulin-like β-sandwich domain (domain IV) and a large (227 amino acids) α-helical domain (domain III), respectively (24, 28). In the case of the C-domain of *Bb*LNBase, there is no resemblance to these GH20 C-domain structures. The C-domain of *Bb*LNBase is located on the side of the barrel domain that is farthest from the active site. A deletion of only five residues at the C terminus of this domain (construct 37–658) significantly reduced the catalytic activity,

TABLE 1
Data collection and refinement statistics

Dataset	Selenomethionine-labeled peak	LNB	LNB-thiazoline
Protein Data Bank entry		4H04	4JAW
Data collection statistics			
Wavelength	0.97898 Å	1.00000 Å	1.00000 Å
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell	a = 117.3 Å	a = 116.8 Å	a = 116.5 Å
	b = 131.6 Å	b = 131.0 Å	b = 131.6 Å
	c = 104.6 Å	c = 104.4 Å	c = 104.8 Å
Resolution[a]	50.0 to 2.20 Å (2.28 to 20 Å)	50.0 to 1.80 Å (1.86 to 1.80 Å)	50.0 to 1.80 Å (1.83 to 1.80 Å)
Total reflections	607,378	1,107,867	1,106,951
Unique reflections	158,438	148,717	150,089
Completeness[a]	100% (99.9%)	100% (100%)	100% (100%)
Redundancy[a]	3.8% (3.8%)	7.4% (7.4%)	7.4% (7.3%)
Mean $I/\sigma(I)$[a]	14.22% (3.41%)	27.64% (4.67%)	23.97% (4.85%)
R_{merge}[a]	10.0% (31.2%)	7.9% (28.1%)	10.2% (48.5%)
Refinement statistics			
Resolution range		47.51–1.80 Å	43.61–1.80 Å
No. of reflections		141,195	141,665
R-factor/R_{free}		15.5/19.1%	18.2/21.3%
r.m.s.d. from ideal values			
Bond lengths		0.028 Å	0.027 Å
Bond angles		2.503°	2.365°
Average B-factor			
Protein (chain A/B)		16.0/15.7 Å^2	11.6/11.3 Å^2
Ligand (chain A/B)		11.5/10.1 Å^2	11.6/9.12 Å^2
Water		26.6 Å^2	12.3 Å^2
SO_4^{2-}		32.3 Å^2	23.8 Å^2
Ramachandran plot[b]			
Favored (chain A/B)		97.8/97.3%	98.3/98.4%
Allowed (chain A/B)		1.7/2.4%	1.7/1.4%
Outlier (chain A/B)		0.5/0.3%	0.0/0.2%

[a] Values in parentheses are for the highest resolution shell.
[b] Data were calculated by RAMPAGE (71).

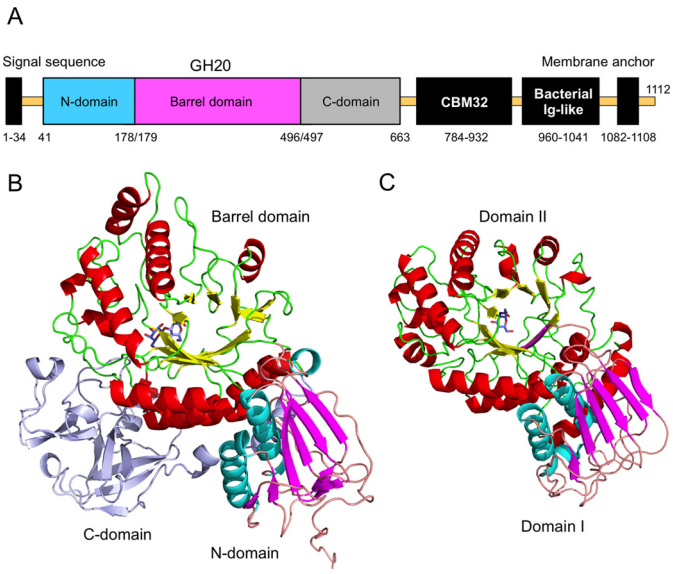

FIGURE 1. *A*, domain structure of *Bb*LNBase. *B*, overall structure of *Bb*LNBase (residues 41–663) complexed with LNB. *C*, overall structure of the wild-type *Sp*Hex complexed with GlcNAc (Protein Data Bank code 1M01). *B* and *C*, α-helices and β-strands in the catalytic $(\alpha/\beta)_8$ barrel domain are shown in *red* and *yellow*, respectively, and α-helices and β-strands in the N-terminal domain are shown in *cyan* and *magenta*, respectively. Bound ligands are shown as *blue sticks*.

and further deletions caused complete inactivation (supplemental Table S1). These results indicate that the C-domain of *Bb*LNBase is essential for protein stability and catalytic activity. Of interest is that the C-domain possesses a broken β-trefoil fold. According to a structural similarity search using the DALI server (40), the C-domain is slightly similar to an R-type lectin (MOA) from *Marasmius oreades* (CBM13) (Z-score = 11.0, r.m.s.d. = 2.0 Å for 102 Cα atoms), which also has a typical β-trefoil fold (41). The β-trefoil fold of MOA is composed of three subdomains that are similar in structure and are assembled at a 3-fold axis (supplemental Fig. S3*C*), and each subdomain contains four-stranded β-hairpin turns (β1–β4) (supplemental Fig. S3*E*). The C-domain of *Bb*LNBase can be divided into an α-subdomain (residues 522–550), β-subdomain (residues 551–615), and γ-subdomain (residues 616–662) (supplemental Fig. S3, *A* and *B*), with each subdomain having at least one additional α-helix. The β-subdomain retains the typical topology of the β-trefoil fold with a complete set of four β-strands (supplemental Fig. S3*D*). However, the α- and γ-subdomains lack one or two β-strand(s), and the C-domain is largely broken at the interface between these subdomains. In the β-subdomain, a disulfide bond is formed between Cys-564 (within β2) and Cys-589 (loop after β3). Interestingly, these structural features (a disulfide bond in the β-subdomain and an additional α-helix after β3) are also present in a CBM13-related arabinose-binding domain (CBM42) of the fungal GH54 α-L-arabinofuranosidase (42). The CBM42 domain also partially lacks the 3-fold symmetry, and its α-subdomain lacks the arabinose-binding site. However, the C-domain of *Bb*LNBase appears to be more broken than CBM42.

Active Site of BbLNBase—We observed a clear electron density for ligands at the center of the barrel domain (Fig. 2, *A* and *B*). The pyranose ring of the GlcNAc residue, in LNB, is in the ^4E conformation and that of GlcNAc-thiazoline, in LNB-thiazoline, is in the $^4C_1$ conformation (Fig. 2, *C* and *D*, discussed below). For LNB, GlcNAc and Gal are bound to the −1 and −2 subsites, respectively, with all the hydroxyl groups of LNB forming hydrogen bonds with the surrounding amino acids.

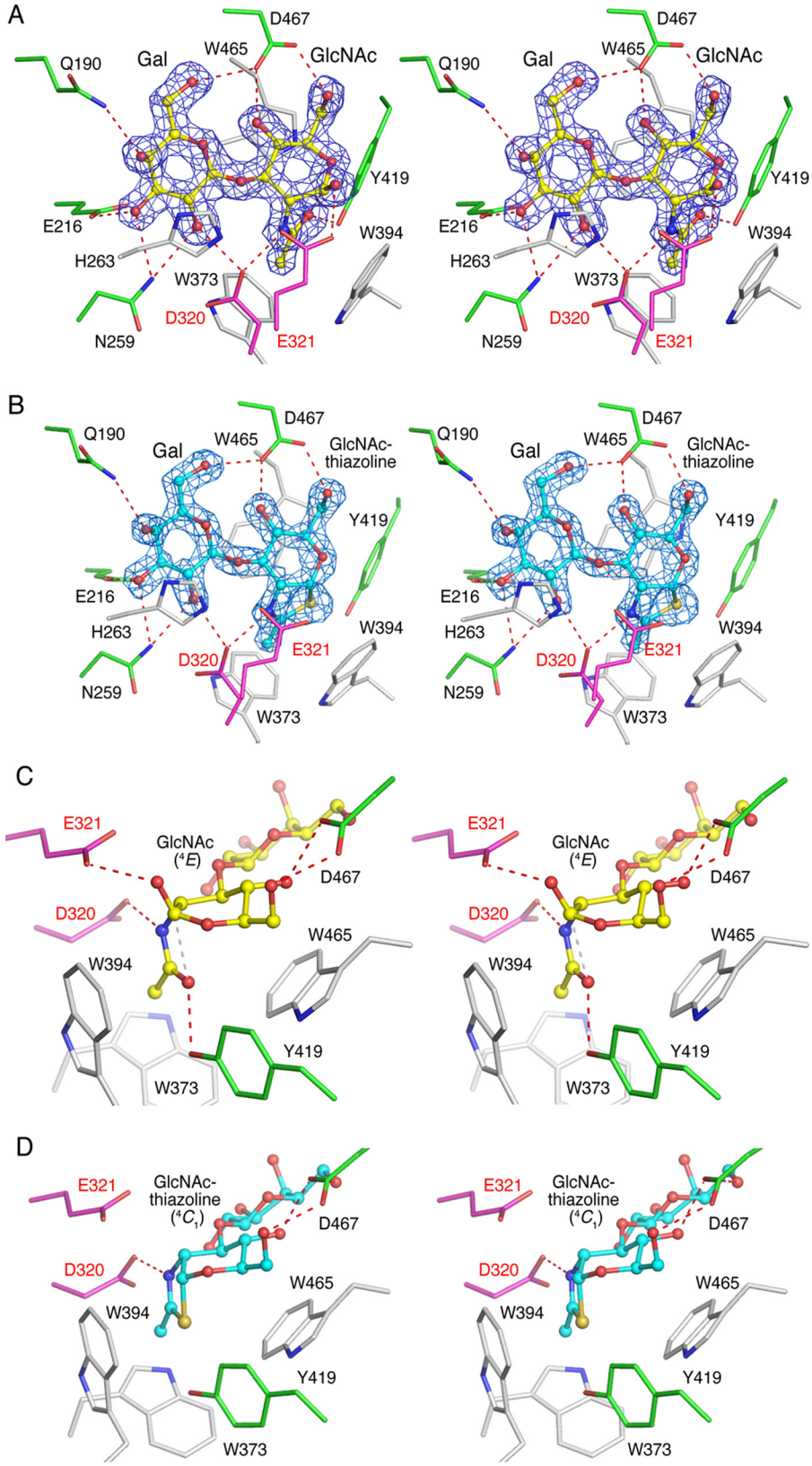

FIGURE 2. **Stereoviews of LNB (*A* and *C*, *yellow*) and LNB-thiazoline (*B* and *D*, *cyan*) bound to the active site of *Bb*LNBase.** Catalytic residues, residues forming hydrogen bonds, and residues forming hydrophobic interactions are shown in *magenta*, *green*, and *white*, respectively. Hydrogen bonds are shown as *red dashed lines*. *A* and *B*, $|F_o| - |F_c|$ omit electron density map (contoured at 4σ) and interactions with protein atoms. *C* and *D*, GlcNAc and GlcNAc-thiazoline sugar ring and surrounding residues.

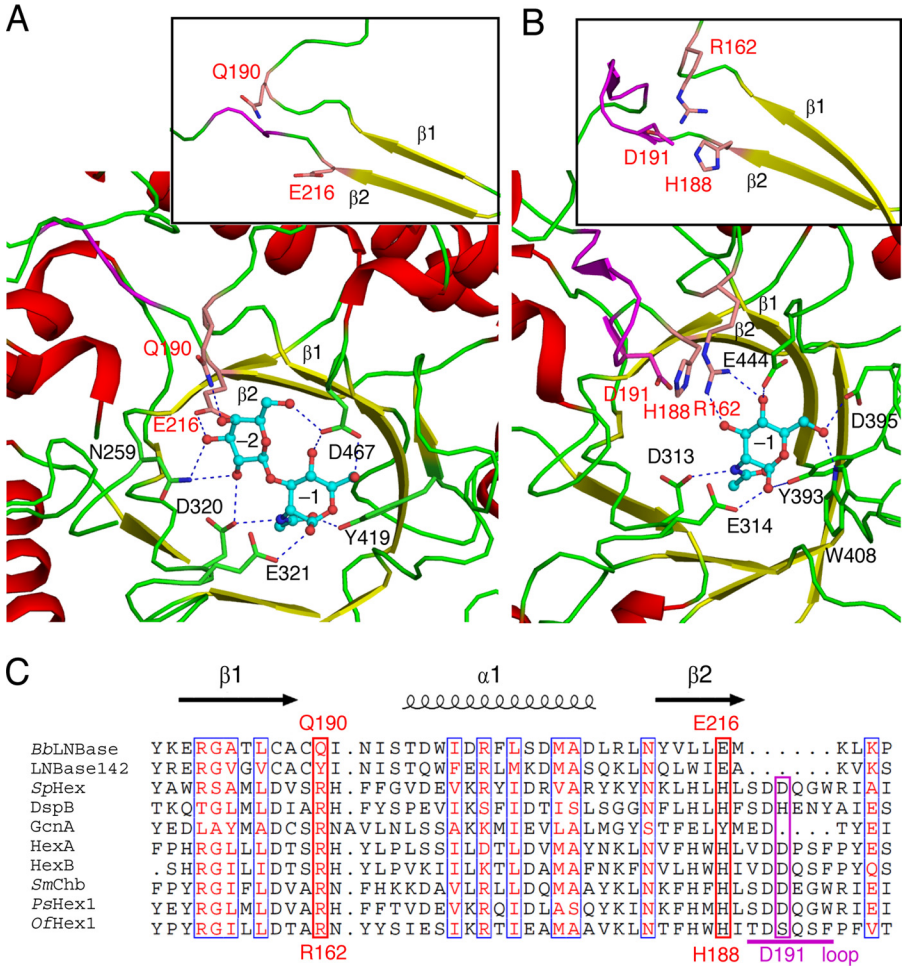

FIGURE 3. **Comparison of the active sites of GH20 enzymes.** *A* and *B*, active sites of *Bb*LNBase (*A*) and *Sp*Hex (*B*), respectively. *Inset*, loop regions following β1 and β2 that determine the presence or absence of −2 subsite. *C*, amino acid sequence alignment of the loop regions of GH20 enzymes (two LNBases and eight β-HexNAcases). Residues important for the presence or absence of −2 subsites are labeled with *red letters*. Six-amino acid loop insertion of β-HexNAcases is indicated by a *magenta bar*. *LNBase142*, *Streptomyces* sp. 142 LNBase; *DspB*, *Actinobacillus actinomycetemcomitans* dispersin B; *GcnA*, *S. gordonii* GcnA; *HexA*, human lysosomal β-HexNAcase A; *HexB,* human lysosomal β-HexNAcase B; *SmChb*, *S. marcescens* chitobiase; *PsHex1*, *Paenibacillus* sp. TS12 Hex1; *OfHex1*, *O. furnacalis* β-HexNAcase.

Such extensive recognitions give evidence for the strict substrate specificity observed for this enzyme (17) as demonstrated for *p*NP-β-GNB, the 4-epimer of *p*NP-β-LNB, that shows highly reduced activity. In the *Bb*LNBase structure, the C4-hydroxyl group of GlcNAc of LNB forms a hydrogen bond with the side chain of Asp-467 (Fig. 2), and this amino acid would block GalNAc-containing substrates from binding such as GNB. In accordance with other biochemical studies, *Bb*LNBase does not hydrolyze fucosylated substrates such as pyridylamino (PA)-LNFP I and PA-LNFP II (17), which can be rationalized as the active site of *Bb*LNBase, and it has no space for the fucosyl group at the C2-hydroxyl of Gal (blocked by Asn-259 and Asp-320).

The substrate-binding site of *Bb*LNBase was compared with that of *Sp*Hex, which is one of the most well studied GH20 β-HexNAcases (Fig. 3) (21–23). Amino acids surrounding the GlcNAc at −1 subsite (Asp-320, Glu-321, Tyr-419, and Asp-467 in *Bb*LNBase) are highly conserved with *Sp*Hex and among GH20 enzymes (supplemental Fig. S4). The two catalytic residues of GH20, Asp-320 (polarizing residue) and Glu-321 (acid/base catalytic residue), form hydrogen bonds with the amide nitrogen of the 2-acetamido group and the O^1-hydroxyl, respectively. Tyr-419 is a highly conserved residue in GH20 enzymes, and its side-chain hydroxyl group forms a hydrogen bond with the carbonyl oxygen atom of the 2-acetamido group. Asp-467 forms bifurcated hydrogen bonds with the O^4- and O^6-hydroxyl groups of the GlcNAc residue. The corresponding residue is often substituted with Glu in β-HexNAcases (Glu-444 in *Sp*Hex), which forms a hydrogen bond with the O^4-hydroxyl group alone. The hydrogen bond with Asp-467 fixes the O^6-hydroxyl group in a *gauche-gauche* orientation in *Bb*LNBase (Fig. 3*A*), whereas the O^6-hydroxyl group of GlcNAc in *Sp*Hex is in the *gauche-trans* orientation due to hydrogen bonds with Asp-395 and Trp-408 (Fig. 3*B*). In addition, His-263, Trp-373, Trp-394, and Trp-465 form hydrophobic interactions with the substrate in the active site (Fig. 2) and are highly conserved in GH20 (supplemental Fig. S4).

To investigate the importance of these residues in the −1 subsite, we constructed site-directed mutants D320A, D320N, Y419F, and H263F (Table 2). Mutations at the catalytically important Asp-320 residue (D320A and D320N) did not affect the K_m values, but they exhibited significantly reduced k_{cat} val-

ues, respectively, which is consistent with those of the corresponding mutants of *Sp*Hex (D313A and D313N) (22), human HexB (D354N) (43), and *Paenibacillus* sp. TS12 Hex1 (D321N) (29). In the case of Y419F, the K_m and k_{cat} values were both significantly reduced compared with the wild-type enzyme. The mutation at His-263 also showed a reduced k_{cat} value, which is consistent with this residue being critical in the hydrophobic interaction with GlcNAc.

As expected from the disaccharide-releasing characteristics of *Bb*LNBase, a −2 subsite specific for β-1,3-linked Gal is clearly defined in the crystal structure (Fig. 3*A*). The side chains of Gln-190 and Glu-216 form hydrogen bonds with the O^4- and O^3-hydroxyl groups of Gal, respectively. Gln-190 and Glu-216 in *Bb*LNBase are replaced by Arg-162 and His-188, respectively, in *Sp*Hex (Fig. 3*C*) and jointly block the space for a potential −2 subsite in this enzyme (Fig. 3*B*). Furthermore, in most β-HexNAcases, a 6-amino acid loop insertion is present just after the His-188 residue (Fig. 3*C*), and a relatively conserved Asp residue (Asp-191 in *Sp*Hex) located in the loop occupies a position corresponding to the −2 subsite (Gal binding site) of *Bb*LNBase (Fig. 3*B*). Comparison of the molecular surfaces of *Bb*LNBase and *Sp*Hex (Fig. 4) illustrates pockets that are suitable in size for binding disaccharide and monosaccharide units, respectively. However, we were unable to determine the positive subsites in the *Bb*LNBase structure. Attempts at co-crystallization and soaking experiments with lactose, Gal, or Glc were undertaken, but no electron densities for these sugars were found (data not shown). *Bb*LNBase has been previously shown to efficiently hydrolyze PA-LNT as well as *p*NP-β-LNB (17). However, the enzymatic activity against LNT, the natural form of a major HMO component, has not been fully studied. Thus, we measured the kinetic parameters for hydrolysis of LNT by *Bb*LNBase. The K_m, k_{cat}, and k_{cat}/K_m values were determined to be 626 ± 23 μM, 42.1 ± 0.8 s^{-1}, and 67.2 ± 1.9 s^{-1} mM^{-1}, respectively. The K_m and k_{cat} values were higher compared with those for *p*NP-LNB (Table 2), demonstrating that this enzyme is an *exo*-acting enzyme and does not have positive subsites specific for sugar moieties. In addition, the entrance of the substrate-binding pocket of *Bb*LNBase appears to be wider than that of *Sp*Hex (Fig. 4). This is in agreement with the finding that LNBase from *Streptomyces* sp. 142 releases LNB from the nonreducing end of various oligosaccharides, including a large triantennary sugar chain (15, 16).

TABLE 2
Kinetic parameters of wild-type and mutants of *Bb*LNBase
Activities of pNP-LNB hydrolysis were measured in 50 mM McIlvaine buffer (pH 4.5) at 30 °C.

Enzyme	K_m	k_{cat}	k_{cat}/K_m
	$\times 1^{-6}$ M	s^{-1}	$\times 10^3$ s^{-1}/M
Wild type	99 ± 11	15 ± 1	160 ± 10
H263F	120 ± 20	0.15 ± 0.01	1.3 ± 0.1
D320A	130 ± 20	0.0068 ± 0.0005	0.052 ± 0.006
D320N	380 ± 50	0.011 ± 0.001	0.028 ± 0.002
Y419F	–[a]	0.086[b]	–

[a] The K_m value was too low to be determined.
[b] The $v/[E]_0$ value at 100 μM substrate.

Reaction Mechanism and Conformational Changes of the Substrate—In the widely accepted mechanism of GH20 (Fig. 5), which is one of substrate-assisted catalysis, the 2-acetamido group of a substrate is polarized and oriented by a deprotonated Asp residue that is located at the neighboring position of the acid/base Glu residue (44). The oxygen atom of the carbonyl group acts as the nucleophile, which attacks the anomeric carbon, resulting in the formation of an oxazoline or oxazolinium ion intermediate. The protonated acid/base Glu residue facilitates bond cleavage by providing a proton to the glycosidic oxygen. In the second step of the reaction, the acid/base Glu activates a water molecule, thereby facilitating its nucleophilic attack at the anomeric carbon. During the course of this reaction, two oxocarbenium ion-like transition states are thought to exist one on each side of the oxazoline or oxazolinium ion intermediate.

When investigating the catalytic mechanism of carbohydrate-active enzymes, the dynamic behavior of a substrate sugar molecule during the catalytic route in the active site is of particular interest. In this work, we observed distorted GlcNAc sugars of LNB in the −1 subsite of the *Bb*LNBase structure (Fig. 2*C*), whereas the GlcNAc-thiazoline group of the LNB-thiazoline complex is in a $^4C_1$ conformation (Fig. 2*D*). Therefore, to gain insight into the conformations of the sugar at binding, we analyzed the sugar ring conformations in detail using the Cremer-Pople system, which is widely used to define the puckering conformations of a pyranose (45). The Cremer-Pople parameters of GlcNAc of LNB in chains A and B were $\phi = 245.9°$, $\theta = 60.1°$, and $Q = 0.595$, and $\phi = 252.6°$, $\theta = 56.7°$, and $Q = 0.616$, respectively, indicating that they are in a conformation close to ^4E (ideally, $\phi = 240°$ and $\theta = 54.7°$). On the other hand, the parameters of GlcNAc-thiazoline of LNB-thiazoline in chains A and B were $\phi = 237.6°$, $\theta = 13.2°$, and $Q = 0.547$, and $\phi =$

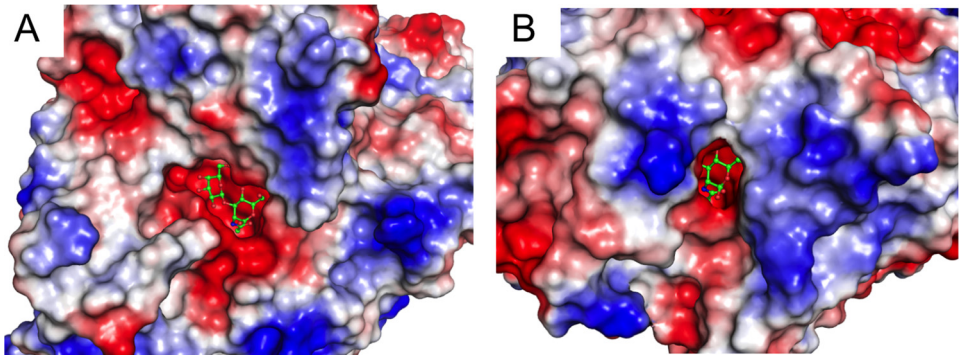

FIGURE 4. **Molecular surfaces of *Bb*LNBase (*A*) and *Sp*Hex (*B*) showing the substrate binding pockets.** According to the electrostatic potential, the surface is colored *blue* for positive and *red* for negative charges. LNB and GlcNAc bound to the pocket are shown as *green sticks*.

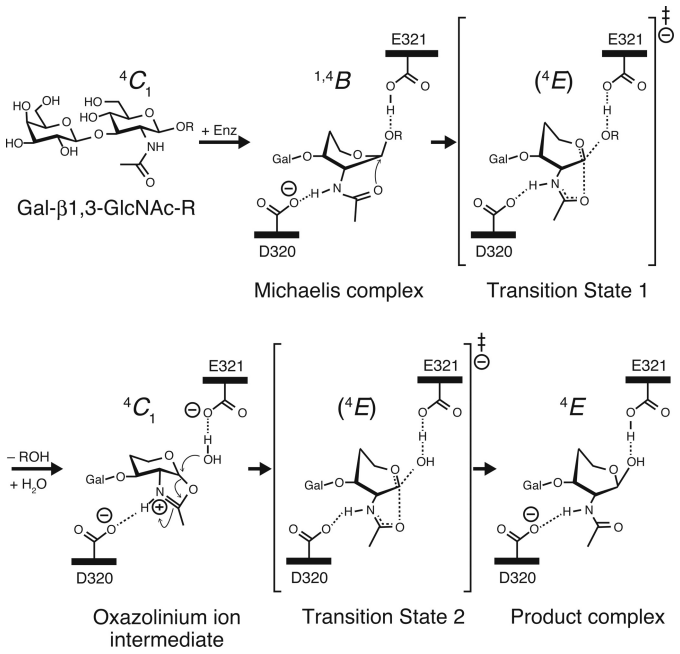

FIGURE 5. **Proposed reaction pathway of *Bb*LNBase and possible conformations of the GlcNAc sugar ring.** *R* = a leaving group.

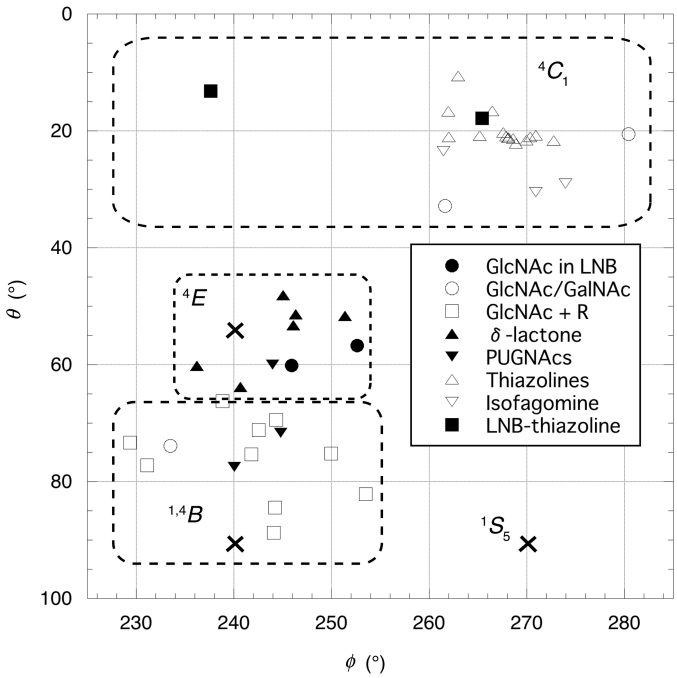

FIGURE 6. **Distribution map of sugar ring conformations observed in the −1 subsite of GH20 enzymes as analyzed by Cremer-Pople angle parameters.** *Cross symbols* indicate ideal positions of $^{1,4}B$ ($\phi = 240°$, $\theta = 90°$), ^4E ($\phi = 240°$, $\theta = 54.7°$), and $^1S_5$ ($\phi = 270°$, $\theta = 90°$) conformations. *Open squares* (*GlcNAc + R*) include chitobiose, GlcNAc-β1,2-Man, GMMG, NGAB2, and TMG-chitotriomycin. *Closed downward triangles* (*PUGNAcs*) include PUGNAc and Gal-PUGNAc. *Open upward triangles* (thiazolines) include GlcNAc-thiazoline and GalNAc-thiazoline. *Open downward triangles* (isofagomine) indicate GalNAc-isofagomine. *Closed squares* indicate LNB-thiazoline. Details of the sugar compound names and their parameters are listed in supplemental Table S2.

265.4°, $\theta = 17.9°$, and $Q = 0.597$, respectively, indicating that they are in a conformation close to $^4C_1$.

The 2-acetamido group of GlcNAc in LNB is nestled in among three aromatic residues (Trp-373, Trp-394, and Trp-465) and fixed by two hydrogen bonds with Asp-320 and Tyr-419 (Fig. 2*C*). The 2-acetamido group at its fixed position elevates the C1 atom of GlcNAc, and the hydrogen bond between Asp-467 and O^4-hydroxyl fixes the C4 atom of GlcNAc in an elevated position. These interactions potentially distort the conformation of GlcNAc to ^4E, which is not intrinsically stable as evidenced by crystallographic analysis and molecular simulation (46, 47).

Currently, various substrates (*e.g.* chitobiose), inhibitors (*e.g.* GlcNAc-thiazoline), and reaction products (*e.g.* GlcNAc) of β-HexNAcases have been observed in the crystal structures of other GH20 enzymes (supplemental Table S2) (21, 22, 24, 25, 27–31, 48–52). In an effort to gain insight into the global binding of these ligands against GH20 enzymes, their Cremer-Pople parameters (ϕ and θ) were evaluated and plotted (Fig. 6) together with those of the GlcNAc sugars of LNB and LNB-thiazoline in *Bb*LNBase.

GlcNAc-thiazoline and GalNAc-thiazoline are specific inhibitors of GHs that utilize the substrate-assisted mechanism (53, 54) because they structurally resemble the oxazoline or oxazolinium ion intermediate. A recent study using *ab initio* molecular dynamic simulations predicts that the oxazoline or oxazolinium ion intermediate is distorted to a conformational region of $^4H_3/^4E/^4H_5$ (44). However, in GH20 β-HexNAcases, the thiazoline inhibitors always adopt conformations close to $^4C_1$ ($\theta < 23°$) (21, 25, 27, 52). Our crystallographic result also indicates that the LNB-thiazoline in LNBase also adopt a $^4C_1$ conformation (Fig. 6). These ample crystallographic data indicate that the oxazolinium ion intermediate adopts the $^4C_1$ conformation (21).

Various substrates that carry sugar moieties at positive subsites have also been observed in crystal structures of GH20 β-HexNAcases. Complexes of a chitobiase (β-HexNAcase) from *S. marcescens* with chitobiose (di-*N*-acetyl-D-glucosamine) were trapped using the mutants D537A and E540A (catalytic residues) (48) or by a simple soaking method (24). The complex structures of the N- and C-terminal modules (GH20A and GH20B) of the large multimodular β-HexNAcase StrH from *Streptococcus pneumoniae* with disaccharide (in GH20A), tetrasaccharide (in GH20B), and bisected glycan heptasaccharide (in GH20B) substrates were reported using their acid/base residue mutants (E361Q of GH20A and E805Q of GH20B) (31). The GlcNAc moieties of these substrates bound to the −1 subsite were all in conformations close to a $^{1,4}B$ conformation ($229° < \phi < 254°$ and $\theta > 66°$; ideally $\phi = 240°$ and $\theta = 90°$). Moreover, the complex structure of TMG-chitotriomycin, a linear tetrasaccharide inhibitor (55), bound to the β-HexNAcase from *Ostrinia furnacalis* was reported (30) with the *N*,*N*,*N*-trimethyl-D-glucosamine group at the nonreducing end of TMG-chitotriomycin also adopting a $^{1,4}B$-like conformation ($\phi = 244.2°$ and $\theta = 84.5°$). Therefore, the substrate of the Michaelis complex for GH20 enzymes is thought to adopt a $^{1,4}B$ conformation.

Transition state analogs that have an sp^2-hybridized planar anomeric carbon were also used to study β-HexNAcases and related enzymes as these compounds are found to be potent inhibitors of these enzymes. In the complex structure of human lysosomal β-HexNAcase B, 2-acetamido-2-deoxy-D-glucono-

1,5-lactone (δ-lactone) was bound in conformations close to ^4E ($236° < \phi < 253°$ and $48° < \theta < 64°$) (50). *O*-(2-Acetamido-2-deoxy-D-glucopyranosylidene)-amino-*N*-phenylcarbamate (PUGNAc) is a potent inhibitor of a wide range of β-HexNAcases that utilize the classical double displacement mechanism (*e.g.* GH3) as well as the substrate-assisted mechanism used by GH20 enzymes. In β-HexNAcases from *O. furnacalis* and *Paenibacillus* sp. TS12 (Hex1), PUGNAc adopts conformations ranging from ^4E to $^{1,4}B$ ($240° < \phi < 245°$ and $60° < \theta < 78°$) (29, 51). Our results strongly support the hypothesis that the conformation at the transition states of GH20 enzymes is nearly ^4E, which is in compliance with the requirement of coplanarity at the C2, C1, O5, and C5 pyranose ring atoms to attain the oxocarbenium ion-like transition state. ^4E-Like transition states are also predicted for other GH enzymes that utilize the substrate-assisted mechanism, namely a GH18 chitinase (by a QM/MM modeling study) (56), and a GH84 *O*-GlcNAcase (57).

According to previous knowledge regarding GH20 β-HexNAcases and the present crystal structures, we propose a conformational itinerary pathway for the *Bb*LNBase-catalyzed reaction that obeys the "principle of least nuclear motion" (21, 58): $^{1,4}B$ (Michaelis complex)-^4E (transition state 1)-$^4C_1$ (oxazolinium ion intermediate)-^4E (transition state 2)-^4E (product complex) (Fig. 5). Of interest is the ^4E product complex observed in this study. In GH20 β-HexNAcases, several complexes observed with reaction products (GlcNAc or GalNAc) have been reported. However, the ring conformations and binding modes of these molecules differ depending on the conditions of complex formation. In the structures of β-HexNAcase (Hex1) from *Paenibacillus* sp. TS12, GlcNAc is in a $^4C_1$ conformation, whereas GalNAc is in a $^{1,4}B$ conformation (29). In *Sp*Hex, GlcNAc is bound in a $^4C_1$ conformation with the wild-type enzyme, but it adopts the alternative conformations of $^{1,4}B$ and $^4C_1$ in the D313A mutant (22). In the D313N mutant, the bound GlcNAc is in a $^4C_1$ conformation but is tilted and dropped from its normal position in the wild-type complex. In *Bb*LNBase, we speculate that there are two major factors that fix LNB with its GlcNAc moiety in a ^4E conformation. First, the Gal moiety bound to −2 subsite acts as an anchor to lock in place the LNB disaccharide in the active site. This anchor is lacking in other GH20 β-HexNAcases as they only have confined space in the active site to accommodate a GlcNAc/GalNAc residue. Second, the bifurcated hydrogen bond between Asp-467 and the O4/O6-hydroxyls of the GlcNAc moiety strongly fix it in the ^4E and *gauche-gauche* conformations. These factors are unique to LNBases and have to date not been shown in β-HexNAcases.

C-terminal Domains—In the crystal structure of *Bb*LNBase, we observed a novel β-trefoil-like domain (C-domain) that is required for protein stability. One of our most striking findings was that the C-domain has several features similar to those of the arabinose-binding domain CBM42. However, there is currently no evidence indicating that the C-domain in *Bb*LNBase has carbohydrate-binding ability. We attempted co-crystallization and soaking experiments with various sugars such as lactose, Gal, or Glc at high concentrations (up to 1 M), but we could not observe any extra electron density in this domain (data not shown). In a study examining β-HexNAcases from *Paenibacillus* sp. TS12, the C-terminal region of one of the two β-HexNAcases (residues 503–978 of Hex1) was suggested to help the enzyme in its interaction with glycosphingolipid substrates in the absence of detergent (29). However, the three-dimensional structure of this domain remains unknown because only a truncated structure of Hex1 is available (deletion mutant 1–502), and there is no amino acid sequence homology between the C-terminal region of Hex1 and the C-domain of *Bb*LNBase. The full-length *Bb*LNBase has a CBM32 domain (residues 784–932) in a region adjacent to the C-domain (Fig. 1*A*). CBM32 domains of GHs from enteric bacteria have been shown to recognize a terminal Gal or GalNAc, or disaccharide motifs such as LacNAc and GlcNAc-α1,4-Gal (59). Therefore, the CBM32 domain of *Bb*LNBase may also function in binding β-1,3-linked Gal, LNB, or other moieties of HMOs.

Biological Implications—In the type I HMO-degradation system of *B. bifidum* JCM1254, *Bb*LNBase plays a central role in releasing LNB from HMOs (9, 60). However, the assistance of other enzymes that remove modifying sugars such as fucose or sialic acid is required. Many extracellular glycosidases involved in HMO degradation have been found in *B. bifidum* JCM1254. GH95 1,2-α-L-fucosidase (AfcA) (61, 62), GH29 1,3–1,4-α-L-fucosidase (AfcB) (63, 64), and GH33 exo-α-sialidases (SiaBb1 and SiaBb2) (65) all share roles in the complete degradation of α1,2- and α1,3/4-fucosylated and sialylated HMOs. In this study, we provide a structural basis for the substrate specificity of *Bb*LNBase. This enzyme cannot accommodate any modified LNB moieties because of a substrate binding pocket that is constrained to allow only LNB binding. This feature is suitable for the specificity of the solute-binding protein of the GNB/LNB transporter (GL-BP) in the GNB/LNB pathway, which specifically binds unmodified LNB disaccharide (11). GL-BP binds LNB and GNB with low K_d values (< 0.09 μM), whereas LNT, a major HMO tetrasaccharide containing LNB, exhibits a significantly higher K_d value (11 μM). Following uptake, LNB is subsequently metabolized by GNB/LNB phosphorylase and other enzymes, including *N*-acetylhexosamine kinase, UDP-glucose hexose-1-phosphate uridylyltransferase, and UDP-glucose 4-epimerase (8). The positive subsites of *Bb*LNBase are wider and appear to be capable of accommodating various groups, suggesting that this enzyme acts on various type I HMOs after the action of fucosidases and sialidases. Moreover, *B. bifidum* JCM1254 has three extracellular enzymes, including one GH2 β-galactosidase (BbgIII) and two GH20 β-HexNAcases (BbhI and BbhII) (66). Two of these enzymes (BbgIII and BbhI) are suggested to play essential roles in degrading type II HMOs because they catalyze the complete hydrolysis of lacto-*N*-neotetraose (Gal-β1,4-GlcNAc-β1,3-Gal-β1,4-Glc) to monosaccharides. BbgIII hydrolyzes β-1,4- and β-1,6-linked Gal but not β-1,3-linked Gal in LNB and LNT, suggesting that it is also involved in the complete degradation of type I HMOs after the action of LNBase (*e.g.* cleavage of lactose released from LNT).

In contrast to *B. bifidum* JCM1254, *B. longum* subsp. *infantis* ATCC15697 lacks the LNBase gene. A genomic analysis of *B. longum* subsp. *infantis* identified a unique gene cluster, the HMO cluster, containing various intracellular GHs that lack signal sequences (67, 68). A GH29 fucosidase, a GH95 fucosidase, a GH33 sialidase, a GH20 β-HexNAcase, a GH2 β-galac-

tosidase, and at least four putative sugar transporters were found in the HMO cluster. The GH2 β-galactosidase is specific for lactose and type II HMOs (69). Furthermore, the intracellular GH42 β-galactosidase Bga42A, which is distant from the HMO cluster, is highly specific for LNT and functions as the sole β-galactosidase acting on type I HMOs. In the HMO degradation system of this strain, it was suggested that low molecular weight HMOs (degree of polymerization ranging from 3 to 8) are directly imported into the cells by transporters and subsequently cleaved by intracellular GHs (70). This analysis as well as recent research on commensal bacteria is unveiling the different HMO consumption modes of bifidobacterial strains. Extracellular bifidobacterial LNBases may also aid in the HMO consumption of other bifidobacterial strains that lack endogenous LNBases. Importantly, in this study, we identified residues that are important for β-1,3-linked Gal recognition at the −2 subsite (*viz.* Gln-190 and Glu-216). Conservation of these residues is a critical marker that helps to predict putative LNBase genes. A BLAST search against protein databases indicates that putative LNBase genes are present in many bacteria, mainly in the Actinobacteridae subclass. The presence (or absence) of LNBase in the genome of microbes in an infant's gastrointestinal tract will provide important information that will advance our understanding of how these microbes metabolize HMOs.

Acknowledgments—We thank Drs. M. Nishimoto and M. Kitaoka (National Food and Research Institute, National Agriculture and Food Research Organization, Japan) for providing LNB and for helpful discussion. We thank the staff of the Photon Factory and SPring-8 for the x-ray data collection. M. H. and K. A. S. thank the Centre for Microscopy, Characterization, and Analysis at the University of Western Australia, which is supported by University, State, and Federal Government funding.

REFERENCES

1. Urashima, T., Fukuda, K., Kitaoka, M., Terabayashi, T., and Kobata, A. (eds) (2011) *Milk Oligosaccharides*, Nova Science Publishers, New York
2. Urashima, T., Asakuma, S., Leo, F., Fukuda, K., Messer, M., and Oftedal, O. T. (2012) The predominance of type I oligosaccharides is a feature specific to human breast milk. *Adv. Nutr.* **3,** 473S-482S
3. Zivkovic, A. M., German, J. B., Lebrilla, C. B., and Mills, D. A. (2011) Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 4653–4658
4. Venema, K. (2012) Intestinal fermentation of lactose and prebiotic lactose derivatives, including human milk oligosaccharides. *Int. J. Dairy* **22,** 123–140
5. Urashima, T., Odaka, G., Asakuma, S., Uemura, Y., Goto, K., Senda, A., Saito, T., Fukuda, K., Messer, M., and Oftedal, O. T. (2009) Chemical characterization of oligosaccharides in chimpanzee, bonobo, gorilla, orangutan, and siamang milk or colostrum. *Glycobiology* **19,** 499–508
6. Asakuma, S., Urashima, T., Akahori, M., Obayashi, H., Nakamura, T., Kimura, K., Watanabe, Y., Arai, I., and Sanai, Y. (2008) Variation of major neutral oligosaccharides levels in human colostrum. *Eur. J. Clin. Nutr.* **62,** 488–494
7. Kitaoka, M., Tian, J., and Nishimoto, M. (2005) Novel putative galactose operon involving lacto-*N*-biose phosphorylase in *Bifidobacterium longum. Appl. Environ. Microbiol.* **71,** 3158–3162
8. Nishimoto, M., and Kitaoka, M. (2007) Identification of *N*-acetylhexosamine 1-kinase in the complete lacto-*N*-biose I/galacto-*N*-biose metabolic pathway in *Bifidobacterium longum. Appl. Environ. Microbiol.* **73,** 6444–6449
9. Fushinobu, S. (2010) Unique sugar metabolic pathways of bifidobacteria. *Biosci. Biotechnol. Biochem.* **74,** 2374–2384
10. Suzuki, R., Katayama, T., Kitaoka, M., Kumagai, H., Wakagi, T., Shoun, H., Ashida, H., Yamamoto, K., and Fushinobu, S. (2009) Crystallographic and mutational analyses of substrate recognition of endo-α-*N*-acetylgalactosaminidase from *Bifidobacterium longum. J. Biochem.* **146,** 389–398
11. Suzuki, R., Wada, J., Katayama, T., Fushinobu, S., Wakagi, T., Shoun, H., Sugimoto, H., Tanaka, A., Kumagai, H., Ashida, H., Kitaoka, M., and Yamamoto, K. (2008) Structural and thermodynamic analyses of solute-binding protein from *Bifidobacterium longum* specific for core 1 disaccharide and lacto-*N*-biose I. *J. Biol. Chem.* **283,** 13165–13173
12. Hidaka, M., Nishimoto, M., Kitaoka, M., Wakagi, T., Shoun, H., and Fushinobu, S. (2009) The crystal structure of galacto-*N*-biose/lacto-*N*-biose I phosphorylase: a large deformation of a TIM barrel scaffold. *J. Biol. Chem.* **284,** 7273–7283
13. Kiyohara, M., Tachizawa, A., Nishimoto, M., Kitaoka, M., Ashida, H., and Yamamoto, K. (2009) Prebiotic effect of lacto-*N*-biose I on bifidobacterial growth. *Biosci. Biotechnol. Biochem.* **73,** 1175–1179
14. Xiao, J. Z., Takahashi, S., Nishimoto, M., Odamaki, T., Yaeshima, T., Iwatsuki, K., and Kitaoka, M. (2010) Distribution of *in vitro* fermentation ability of lacto-*N*-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains. *Appl. Environ. Microbiol.* **76,** 54–59
15. Sano, M., Hayakawa, K., and Kato, I. (1992) An enzyme releasing lacto-*N*-biose from oligosaccharides. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 8512–8516
16. Sano, M., Hayakawa, K., and Kato, I. (1993) Purification and characterization of an enzyme releasing lacto-*N*-biose from oligosaccharides with type 1 chain. *J. Biol. Chem.* **268,** 18560–18566
17. Wada, J., Ando, T., Kiyohara, M., Ashida, H., Kitaoka, M., Yamaguchi, M., Kumagai, H., Katayama, T., and Yamamoto, K. (2008) *Bifidobacterium bifidum* lacto-*N*-biosidase, a critical enzyme for the degradation of human milk oligosaccharides with a type 1 structure *Appl. Environ. Microbiol.* **74,** 3996–4004
18. Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., Kerr, C., Hourihane, J., Murray, D., Fuligni, F., Gueimonde, M., Margolles, A., De Bellis, G., O'Toole, P. W., van Sinderen, D., Marchesi, J. R., and Ventura, M. (2012) Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* **7,** e36957
19. Magnani, J. L., Steplewski, Z., Koprowski, H., and Ginsburg, V. (1983) Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19–9 in the sera of patients as a mucin. *Cancer Res.* **43,** 5489–5492
20. 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
21. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G., and James, M. N. (2001) Crystallographic evidence for substrate-assisted catalysis in a bacterial β-hexosaminidase. *J. Biol. Chem.* **276,** 10330–10337
22. Williams, S. J., Mark, B. L., Vocadlo, D. J., James, M. N., and Withers, S. G. (2002) Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by orienting the 2-acetamido group and stabilizing the transition state. *J. Biol. Chem.* **277,** 40055–40065
23. Vocadlo, D. J., and Withers, S. G. (2005) Detailed comparative analysis of the catalytic mechanisms of β-*N*-acetylglucosaminidases from families 3 and 20 of glycoside hydrolases. *Biochemistry* **44,** 12809–12818
24. Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S., and Vorgias, C. E. (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat. Struct. Biol.* **3,** 638–648
25. Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S., and James, M. N. (2003) Crystal structure of human β-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease. *J. Mol. Biol.* **327,** 1093–1109
26. Ramasubbu, N., Thomas, L. M., Ragunath, C., and Kaplan, J. B. (2005) Structural analysis of dispersin B, a biofilm-releasing glycoside hydrolase from the periodontopathogen *Actinobacillus actinomycetemcomitans. J. Mol. Biol.* **349,** 475–486

27. Lemieux, M. J., Mark, B. L., Cherney, M. M., Withers, S. G., Mahuran, D. J., and James, M. N. (2006) Crystallographic structure of human β-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *J. Mol. Biol.* **359,** 913–929
28. Langley, D. B., Harty, D. W., Jacques, N. A., Hunter, N., Guss, J. M., and Collyer, C. A. (2008) Structure of *N*-acetyl-β-D-glucosaminidase (GcnA) from the endocarditis pathogen *Streptococcus gordonii* and its complex with the mechanism-based inhibitor NAG-thiazoline. *J. Mol. Biol.* **377,** 104–116
29. Sumida, T., Ishii, R., Yanagisawa, T., Yokoyama, S., and Ito, M. (2009) Molecular cloning and crystal structural analysis of a novel β-*N*-acetylhexosaminidase from *Paenibacillus* sp. TS12 capable of degrading glycosphingolipids. *J. Mol. Biol.* **392,** 87–99
30. Liu, T., Zhang, H., Liu, F., Wu, Q., Shen, X., and Yang, Q. (2011) Structural determinants of an insect β-*N*-acetyl-D-hexosaminidase specialized as a chitinolytic enzyme. *J. Biol. Chem.* **286,** 4049–4058
31. Pluvinage, B., Higgins, M. A., Abbott, D. W., Robb, C., Dalia, A. B., Deng, L., Weiser, J. N., Parsons, T. B., Fairbanks, A. J., Vocadlo, D. J., and Boraston, A. B. (2011) Inhibition of the pneumococcal virulence factor StrH and molecular insights into *N*-glycan recognition and hydrolysis. *Structure* **19,** 1603–1614
32. Jiang, Y. L., Yu, W. L., Zhang, J. W., Frolet, C., Di Guilmi, A. M., Zhou, C. Z., Vernet, T., and Chen, Y. (2011) Structural basis for the substrate specificity of a novel β-*N*-acetylhexosaminidase StrH protein from *Streptococcus pneumoniae* R6. *J. Biol. Chem.* **286,** 43004–43012
33. Hattie, M., Debowski, A. W., and Stubbs, K. A. (2012) Development of tools to study lacto-*N*-biosidase: an important enzyme involved in the breakdown of human milk oligosaccharides. *ChemBioChem* **13,** 1128–1131
34. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
35. Rappleye, J., Innus, M., Weeks, C. M., and Miller, R. (2002) SnB version 2.2: an example of crystallographic multiprocessing. *J. Appl. Crystallogr.* **35,** 374–376
36. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55,** 849–861
37. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6,** 458–463
38. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
39. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
40. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233,** 123–138
41. Grahn, E., Askarieh, G., Holmner, A., Tateno, H., Winter, H. C., Goldstein, I. J., and Krengel, U. (2007) Crystal structure of the *Marasmius oreades* mushroom lectin in complex with a xenotransplantation epitope. *J. Mol. Biol.* **369,** 710–721
42. Miyanaga, A., Koseki, T., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2004) Crystal structure of a family 54 α-L-arabinofuranosidase reveals a novel carbohydrate-binding module that can bind arabinose. *J. Biol. Chem.* **279,** 44907–44914
43. Hou, Y., Vocadlo, D. J., Leung, A., Withers, S. G., and Mahuran, D. (2001) Characterization of the Glu and Asp residues in the active site of human β-hexosaminidase B. *Biochemistry* **40,** 2201–2209
44. Greig, I. R., Zahariev, F., and Withers, S. G. (2008) Elucidating the nature of the *Streptomyces plicatus* β-hexosaminidase-bound intermediate using ab initio molecular dynamics simulations. *J. Am. Chem. Soc.* **130,** 17620–17628
45. Cremer, D., and Pople, J. A. (1975) A general definition of ring puckering coordinates. *J. Am. Chem. Soc.* **97,** 1354–1358
46. Wada, M., Kobayashi, K., Nishimoto, M., Kitaoka, M., and Noguchi, K. (2009) 2-Acetamido-2-deoxy-3-*O*-β-D-galactopyranosyl-D-glucose dihydrate. *Acta Crystallogr. E* **65,** 1781–1782
47. Sattelle, B. M., and Almond, A. (2011) Is *N*-acetyl-D-glucosamine a rigid $^4C_1$ chair? *Glycobiology* **21,** 1651–1662
48. Prag, G., Papanikolau, Y., Tavlas, G., Vorgias, C. E., Petratos, K., and Oppenheim, A. B. (2000) Structures of chitobiase mutants complexed with the substrate Di-*N*-acetyl-D-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J. Mol. Biol.* **300,** 611–617
49. Mark, B. L., Vocadlo, D. J., Zhao, D., Knapp, S., Withers, S. G., and James, M. N. (2001) Biochemical and structural assessment of the 1-*N*-azasugar GalNAc-isofagomine as a potent family 20 β-*N*-acetylhexosaminidase inhibitor. *J. Biol. Chem.* **276,** 42131–42137
50. Maier, T., Strater, N., Schuette, C. G., Klingenstein, R., Sandhoff, K., and Saenger, W. (2003) The x-ray crystal structure of human β-hexosaminidase B provides new insights into Sandhoff disease. *J. Mol. Biol.* **328,** 669–681
51. Liu, T., Zhang, H., Liu, F., Chen, L., Shen, X., and Yang, Q. (2011) Active-pocket size differentiating insectile from bacterial chitinolytic β-*N*-acetyl-D-hexosaminidases. *Biochem. J.* **438,** 467–474
52. Sumida, T., Stubbs, K. A., Ito, M., and Yokoyama, S. (2012) Gaining insight into the inhibition of glycoside hydrolase family 20 exo-β-*N*-acetylhexosaminidases using a structural approach. *Org. Biomol. Chem.* **10,** 2607–2612
53. Knapp, S., Vocadlo, D. J., Gao, Z., Kirk, B., Lou, J., and Withers, S. G. (1996) NAG-thiazoline, an *N*-acetyl-β-hexosaminidase inhibitor that implicates acetamido participation. *J. Am. Chem. Soc.* **118,** 6804–6805
54. Amorelli, B., Yang, C., Rempel, B., Withers, S. G., and Knapp, S. (2008) *N*-Acetylhexosaminidase inhibitory properties of C-1 homologated GlcNAc- and GalNAc-thiazolines. *Bioorg. Med. Chem. Lett.* **18,** 2944–2947
55. Usuki, H., Nitoda, T., Ichikawa, M., Yamaji, N., Iwashita, T., Komura, H., and Kanzaki, H. (2008) TMG-chitotriomycin, an enzyme inhibitor specific for insect and fungal β-*N*-acetylglucosaminidases, produced by actinomycete *Streptomyces anulatus* NBRC 13369. *J. Am. Chem. Soc.* **130,** 4146–4152
56. Jitonnom, J., Lee, V. S., Nimmanpipug, P., Rowlands, H. A., and Mulholland, A. J. (2011) Quantum mechanics/molecular mechanics modeling of substrate-assisted catalysis in family 18 chitinases: conformational changes and the role of Asp-142 in catalysis in ChiB. *Biochemistry* **50,** 4697–4711
57. Rao, F. V., Dorfmueller, H. C., Villa, F., Allwood, M., Eggleston, I. M., and van Aalten, D. M. (2006) Structural insights into the mechanism and inhibition of eukaryotic *O*-GlcNAc hydrolysis. *EMBO J.* **25,** 1569–1578
58. Sinnott, M. L. (1984) On the antiperiplanar lone pair hypothesis and its application to catalysis by glycosidases. *Biochem. J.* **224,** 817–821
59. Ficko-Blean, E., and Boraston, A. B. (2012) Insights into the recognition of the human glycome by microbial carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* **22,** 570–577
60. Asakuma, S., Hatakeyama, E., Urashima, T., Yoshida, E., Katayama, T., Yamamoto, K., Kumagai, H., Ashida, H., Hirose, J., and Kitaoka, M. (2011) Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J. Biol. Chem.* **286,** 34583–34592
61. Katayama, T., Sakuma, A., Kimura, T., Makimura, Y., Hiratake, J., Sakata, K., Yamanoi, T., Kumagai, H., and Yamamoto, K. (2004) Molecular cloning and characterization of *Bifidobacterium bifidum* 1,2-α-L-fucosidase (AfcA), a novel inverting glycosidase (glycoside hydrolase family 95). *J. Bacteriol.* **186,** 4885–4893
62. Nagae, M., Tsuchiya, A., Katayama, T., Yamamoto, K., Wakatsuki, S., and Kato, R. (2007) Structural basis of the catalytic reaction mechanism of novel 1,2-α-L-fucosidase from *Bifidobacterium bifidum*. *J. Biol. Chem.* **282,** 18497–18509
63. 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
64. Sakurama, H., Fushinobu, S., Hidaka, M., Yoshida, E., Honda, Y., Ashida, H., Kitaoka, M., Kumagai, H., Yamamoto, K., and Katayama, T. (2012) 1,3–1,4-α-L-Fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. *J. Biol. Chem.* **287,** 16709–16719
65. Kiyohara, M., Tanigawa, K., Chaiwangsri, T., Katayama, T., Ashida, H.,

and Yamamoto, K. (2011) An exo-α-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates. *Glycobiology* **21,** 437–447
66. Miwa, M., Horimoto, T., Kiyohara, M., Katayama, T., Kitaoka, M., Ashida, H., and Yamamoto, K. (2010) Cooperation of β-galactosidase and β-*N*-acetylhexosaminidase from bifidobacteria in assimilation of human milk oligosaccharides with type 2 structure. *Glycobiology* **20,** 1402–1409
67. Sela, D. A., Chapman, J., Adeuya, A., Kim, J. H., Chen, F., Whitehead, T. R., Lapidus, A., Rokhsar, D. S., Lebrilla, C. B., German, J. B., Price, N. P., Richardson, P. M., and Mills, D. A. (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 18964–18969
68. Sela, D. A., and Mills, D. A. (2010) Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides. *Trends Microbiol.* **18,** 298–307
69. 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
70. Barboza, M., Sela, D. A., Pirim, C., Locascio, R. G., Freeman, S. L., German, J. B., Mills, D. A., and Lebrilla, C. B. (2009) Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Appl. Environ. Microbiol.* **75,** 7319–7325
71. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: φ, ψ, and Cβ deviation. *Proteins* **50,** 437–450