Glycobiology vol. 24 no. 2 pp. 208–216, 2014
doi:10.1093/glycob/cwt104
Advance Access publication on November 23, 2013

Distinct substrate specificities of three glycoside hydrolase family 42 β-galactosidases from *Bifidobacterium longum* subsp. *infantis* ATCC 15697

Alexander H Viborg[2,5], Takane Katayama[4], Maher Abou Hachem[2], Mathias CF Andersen[3], Mamoru Nishimoto[5], Mads H Clausen[3], Tadasu Urashima[6], Birte Svensson[1,2], and Motomitsu Kitaoka[1,5]

[2]Enzyme and Protein Chemistry, Department of Systems Biology; [3]Center for Nanomedicine and Theranostics, Department of Chemistry, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark; [4]Ishikawa Prefectural University, Nonoichi, Ishikawa 921-8836, Japan; [5]National Food Research Institute, National Agriculture and Food Research Organization, 2-1-12, Kannondai, Tsukuba, Ibaraki 305-8642, Japan; and [6]Graduate School of Food Hygiene, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080-8555, Japan

Received on October 3, 2013; revised on November 14, 2013; accepted on November 15, 2013

Glycoside hydrolase family 42 (GH42) includes β-galactosidases catalyzing the release of galactose (Gal) from the non-reducing end of different β-D-galactosides. Health-promoting probiotic bifidobacteria, which are important members of the human gastrointestinal tract microbiota, produce GH42 enzymes enabling utilization of β-galactosides exerting prebiotic effects. However, insight into the specificity of individual GH42 enzymes with respect to substrate monosaccharide composition, glycosidic linkage and degree of polymerization is lagging. Kinetic analysis of natural and synthetic substrates resembling various milk and plant galactooligosaccharides distinguishes the three GH42 members, Bga42A, Bga42B and Bga42C, encoded by the probiotic *B. longum* subsp. *infantis* ATCC 15697 and revealed the glycosyl residue at subsite +1 and its linkage to the terminal Gal at subsite −1 to be key specificity determinants. Bga42A thus prefers the β1-3-galactosidic linkage from human milk and other β1-3- and β1-6-galactosides with glucose or Gal situated at subsite +1. In contrast, Bga42B very efficiently hydrolyses 4-galactosyllactose (Galβ1-4Galβ1-4Glc) as well as 4-galactobiose (Galβ1-4Gal) and 4-galactotriose (Galβ1-4Galβ1-4Gal). The specificity of Bga42C resembles that of Bga42B, but the activity was one order of magnitude lower. Based on enzyme kinetics, gene organization and phylogenetic analyses, Bga42C is proposed to act in the metabolism of arabinogalactan-derived oligosaccharides. The distinct kinetic signatures of the three GH42 enzymes correlate to unique sequence motifs denoting specific clades in a GH42 phylogenetic tree providing novel insight into GH42 subspecificities. Overall, the data illustrate the metabolic adaptation of bifidobacteria to the β-galactoside-rich gut niche and emphasize the importance and diversity of β-galactoside metabolism in probiotic bifidobacteria.

Keywords: galactooligosaccharides / phylogenetic tree / prebiotics / probiotics / sequence motifs

Introduction

Probiotic bacteria constitute a physiologically important group of the human gut microbiota exposed to a range of dietary β-galactosides (Roberfroid et al. 2010) from human milk (Han et al. 2011), cow's milk (Urashima et al. 2001) and derived from plant primary cell wall pectic polysaccharides, e.g. rhamnogalacturonan I (Ridley et al. 2001) and arabinogalactans I and II (Vincken et al. 2003).

Prebiotic non-digestible oligosaccharides stimulate the growth of probiotic bacteria in the gut, especially bifidobacteria (Scott et al. 2013) and are important for human health (van den Broek et al. 2008). Galactooligosaccharides (GOS) are prebiotics that selectively enhance the growth of probiotic bacteria (Boehm and Stahl 2007) and commercial GOS products contain isomeric oligosaccharides consisting of repeats of β1-3/4/6-linked galactosyl residues with different degree of polymerization (DP) ranging from 2 to 10 (Coulier et al. 2009). Studies in humans show that bifidobacteria strains are enriched in the gastrointestinal tract (GIT) by intake of GOS (Davis et al. 2011).

Similar to numerous bacteria of the human gut microbiota, bifidobacteria encode glycoside hydrolases (GHs), which allow utilization of different oligosaccharides with prebiotic properties (Pokusaeva et al. 2011). Glycoside hydrolase family 42 (GH42) thus comprises β-galactosidases and belongs to clan A of retaining GHs catalyzing release of β-D-galactose (Gal) from the non-reducing end of β-galactosides (Cantarel et al. 2009). GH42 enzymes occur in bacteria from diverse habitats such as soil (Veith et al. 2004), human GIT (Sela et al. 2008) and hot springs (Moore et al. 1994; Di Lauro et al.

[1]To whom correspondence should be addressed: Tel: +45-4525-2740; Fax: +45-4588-6307; e-mail: bis@bio.dtu.dk (B.S.); Tel: +81-29-838-8071; Fax: +81-29-838-7321; email: mkitaoka@affrc.go.jp (M.K.)

2008) and degrade β-galactosides from various natural sources, e.g. human milk oligosaccharides (HMOs) and plant pectic galactan (Hinz et al. 2004; Yoshida et al. 2012; Tabachnikov and Shoham 2013).

Genomes of probiotic bifidobacteria often encode several GH42 enzymes (Sela et al. 2008) emphasizing the central role of β-galactoside metabolism in the GIT microbiota and suggesting that GH42 possesses distinct subspecificities. *Bifidobacterium longum* has a large number of genes related to oligosaccharide metabolism (Schell et al. 2002) and the probiotic *B. longum* subsp. *infantis* ATCC 15697 gained increasing attention due to its ability to utilize HMOs (Sela et al. 2008; Asakuma et al. 2011) and other β-galactosides including prebiotic GOS (Garrido et al. 2013). We recently reported that the GH42 enzyme Bga42A of *B. longum* subsp. *infantis* ATCC 15697 hydrolyses lacto-*N*-tetraose from human milk with high efficiency (Yoshida et al. 2012). Knowledge is limited, however, on the connection between enhanced growth of certain bifidobacterial strains and the specificity of GH42 β-galactosidases with respect to substrate DP, bond type and monosaccharide composition (Barboza et al. 2009; O'Connell Motherway et al. 2011). In the present study, substrate specificity and kinetic properties of all three intracellular GH42 β-galactosidases (Bga42A, Bga42B and Bga42C) from *B. longum* subsp. *infantis* ATCC 15697 were described by using a series of β-galactosides as substrates, some of which contain glucose (Glc) and *N*-acetylglucosamine (GlcNAc). The three enzymes are shown to have complementary substrate specificity for different mammalian milk oligosaccharides and β-galactosides and segregate in a phylogenetic analysis that reflects taxonomy and enzymatic function.

Results

Biochemical and bioinformatic evidence for subfamily classification in GH42

We recently revealed a distinct substrate specificity of the three intracellular GH42 β-galactosidases Bga42A, Bga42B and Bga42C of the probiotic *B. longum* subsp. *infantis* ATCC 15697 (Yoshida et al. 2012). Referring to a substrate-binding subsite nomenclature of increasing negative numbers toward the non-reducing end from the point of hydrolysis and increasing positive numbers toward the reducing end (Davies et al. 1997), GH42 was previously shown to recognize only non-reducing end β-D-galactosyl residues at subsite −1 as in *para*-nitrophenyl β-D-galactopyranoside (Holmes et al. 1997; Hung et al. 2001; Hu et al. 2007; Hildebrandt et al. 2009; Lee et al. 2011). This agrees with structural information from both of the two reported GH42 crystal structures (Hidaka et al. 2002; Maksimainen et al. 2012), where steric hindrance of an equatorial C4-OH of a D-glucosyl residue is suggested. The residues at subsite −1 forming hydrogen bonds to β-D-Gal (R111, N149, E150, E307, E355 and H358 in Bcα-β-Gal; Maksimainen et al. 2012; PDB ID: 3TTY, marked with stars in Supplementary data, Figure S1) and the aromatic residues at subsite +1 (Y277 and W315; Maksimainen et al. 2012) are proposed invariant in the 61 GH42 sequences from bifidobacteria with the exception of group 1 (see below), where H358 is substituted by a tryptophan.

A phylogenetic tree with four major groups was generated from the multiple sequence alignment (Supplementary data, Figure S1) of the 61 bifidobacterial GH42 sequences present in Carbohydrate-Active enZYmes (CAZy; Cantarel et al. 2009) to identify the clustering of subgroups within GH42. Based on peptide conservation scores, it was possible to assign distinct sequence motifs associated with nine residues (marked by filled circles, Supplementary data, Figure S1) to the four different clusters in the bifidobacterial phylogenetic tree (data not shown). These distinct sequence motifs were used as queries to extract additional sequences from other taxonomic groups. An alignment was made from all sequences harboring the identified active site motif sequences and used to calculate a phylogenetic tree (Figure 1A). Additionally, the phylogenetic tree was constructed as a phylogram to provide insight into the relationships between the different groups (Figure 1B). The location of the particular sequence motif region is in a loop that flanks one side of the active site cleft and contains the invariant R111 (in Bcα-β-Gal, Maksimainen et al. 2012; PDB ID: 3TTY, Figure 2A and B) that is hydrogen bonding with C4-OH of the β-galactosyl ring at subsite −1. The sequence motif characteristic of group 1 is found throughout the bacterial kingdom (Supplementary data, Table S1). The motif distinguishing subgroup 4.1 is predominantly found in various bifidobacteria and lactobacilli from the gut niche (19 organisms found), whereas that for subgroup 3.1 is a strict signature of the *B. longum* group.

Bga42A specificity is governed by subsite +1

Among the substrates tested, Bga42A most efficiently hydrolyses 3-galactosylglucose (Galβ1-3Glc) ($k_{cat}/K_m = 1671 \pm 150$ s^{-1} mM^{-1}) followed by 3-galactobiosyllactose (Galβ1-3Galβ1-3Galβ1-4Glc), 3-galactobiose (Galβ1-3Gal) and 6-galactobiose (Galβ1-6Gal) (Table I). Lacto-*N*-tetraose (Galβ1-3GlcNAcβ1-3Galβ1-4Glc) is also a relatively good substrate for which interactions at subsites +2 and +3 probably may play a role in recognition when compared with the poorly binding lacto-*N*-biose I (Galβ1-3GlcNAc) that has 10-fold higher K_m (Yoshida et al. 2012). The 2-*N*-acetyl group is accommodated by Bga42A in both lacto-*N*-biose I and lacto-*N*-tetraose, however, when this group is lacking K_m is reduced as found for the disaccharide 3-galactosylglucose that is the best substrate for Bga42A and has 560-fold higher k_{cat}/K_m when compared with lacto-*N*-biose I (Yoshida et al. 2012). Furthermore, 3-galactobiosyllactose and 3-galactotriosyllactose (Galβ1-3Galβ1-3Galβ1-3Galβ1-4Glc) are hydrolyzed with higher than and comparable efficiency to lacto-*N*-tetraose, respectively, but less efficiently than 3-galactosylglucose. Notably, Bga42A hydrolyses 3-galactobiose, 4-galactobiose (Galβ1-4Gal) and 6-galactobiose with k_{cat}/K_m values of 136–422 s^{-1} mM^{-1}, i.e. similar to the efficiency found for lacto-*N*-tetraose ($k_{cat}/K_m = 120$ s^{-1} mM^{-1}; Yoshida et al. 2012). This suggests the ability of Bga42A to accept the bulky GlcNAc of lacto-*N*-tetraose at subsite +1 as well as similarly good accommodation of all the different structural isomers of galactobiose at subsite +1. Additionally, K_m values for 3-galactobiose ($K_m = 2.7$ mM^{-1}) and 3-galactosylglucose ($K_m = 0.38$ mM^{-1}) reflect improved affinity elicited by a sugar unit with equatorial C4-OH at subsite +1, consistent with the importance of this subsite in substrate recognition by Bga42A.

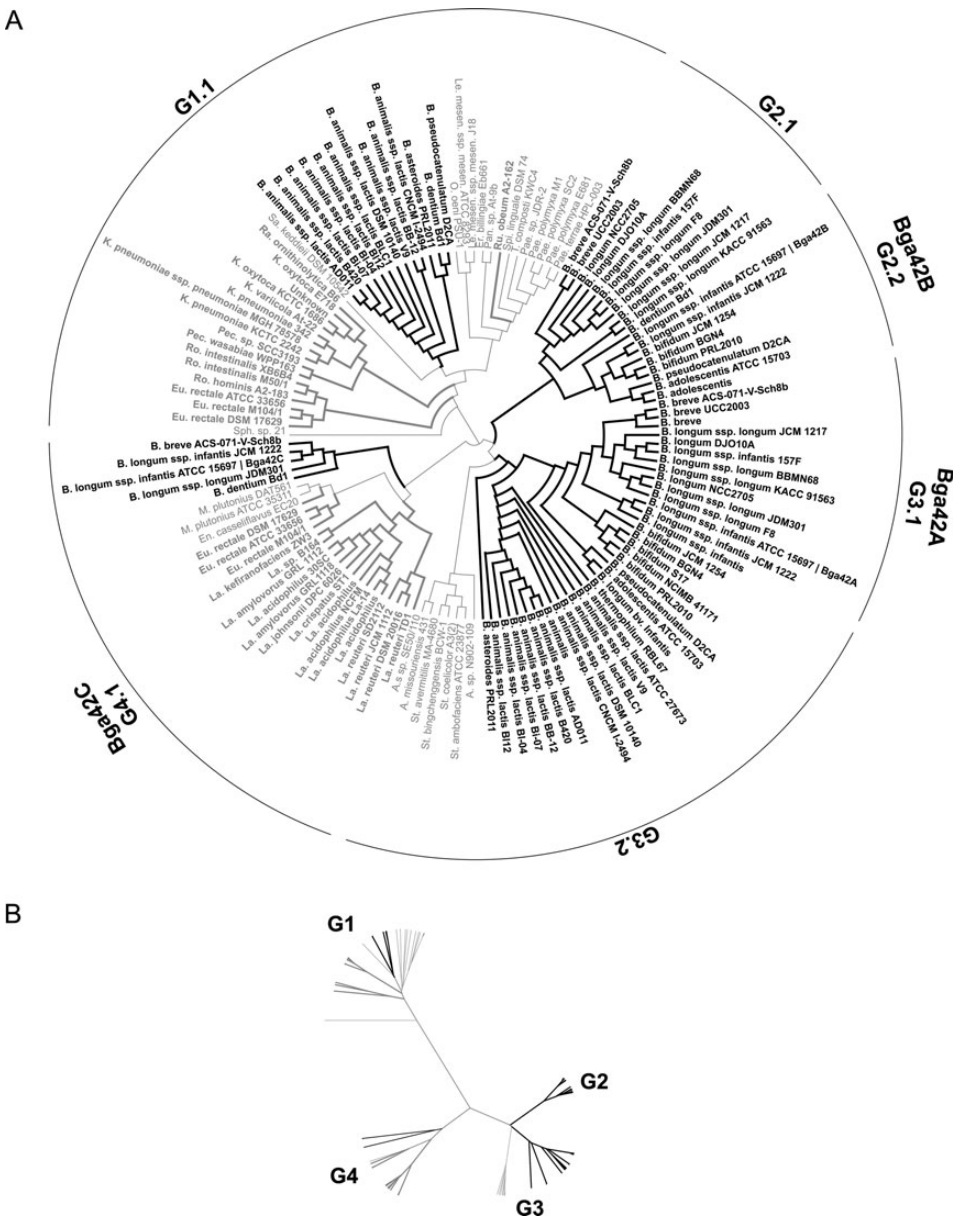

Fig. 1. Phylogenetic tree comprising the 61 non-redundant GH42 sequences from bifidobacteria. (**A**) Cladogram indicating the major groups and subgroups. Subtrees populated by bifidobacteria are in black, whereas branches with other taxonomic groups are in grey, and organisms from the human intestinal niche as identified in the human microbiome project (NIH HMP Working Group et al. 2009) are marked by bold. The three GH42 enzymes in this study, Bga42A, Bga42B and Bga42C are denoted in the subgroups to which they belong. The identified major groups are associated with a distinct sequence motif. G1.1: YG[X]R[Q|E][X]M [D|N]I, G2.1: AG[X]R[Q|E][X]W[Q|R]P, G2.2: AG[X]R[Q|E][X]WSP, G3.1: PG[X]R[Q|E][X]WRA, G3.2: PG[X]R[Q|E][X]WRP and G4.1: FG[X]RH[X][A|F] CP. The genera are abbreviated as, *A.*: *Actinoplanes*, *B.*: *Bifidobacterium*, *En.*: *Enterococcus*, *Er.*: *Erwinia*, *Eu.*: *Eubacterium*, *K.*: *Klebsiella*, *La.*: *Lactobacillus*, *Le.*: *Leuconostoc*, *M.*: *Melissococcus*, *O.*: *Oenococcus*, *Pae.*: *Paenibacillus*, *Pan.*: *Pantoea*, *Pec.*: *Pectobacterium*, *Ra.*: *Raoultella*, *Ro.*: *Roseburia*, *Ru.*: *Ruminococcus*, *Sa.*: *Sanguibacter*, *Sph.*: *Sphingobacterium*, *Spi.*: *Spirosoma*, *St.*: *Streptomyces* and *T.*: *Thermobacillus*, additionally both *mesenteroides* species and *mesenteroides* subspecies are abbreviated *mesen.* All organisms including GenBank accessions are listed in Supplementary data, Table S1. (**B**) Phylogram of the GH42 sequences based on the above sequence motifs (Figure 1A and Supplementary data, Table S1) including 61 non-redundant bifidobacteria, providing insight into the relationships between the different groups.

Bga42B requires Gal at subsite +1

Bga42B is specific for β1-4-galactosides as it recognizes Gal at subsite +1 and hydrolyzes 4-galactobiose with high efficiency ($k_{cat}/K_m = 1697\ s^{-1}\ mM^{-1}$), but has extremely low activity toward lactose (Galβ1-4Glc) ($k_{cat}/K_m = 0.19\ s^{-1}\ mM^{-1}$; Yoshida et al. 2012) due to 100-fold increase in K_m and 90-fold decrease in k_{cat}. Thus, the linkage involving the axial C4-OH of Gal at subsite +1 is crucial for the activity of Bga42B as illustrated by the 8900-fold lower catalytic efficiency when this linkage is substituted to the equatorial C4-OH of Glc that

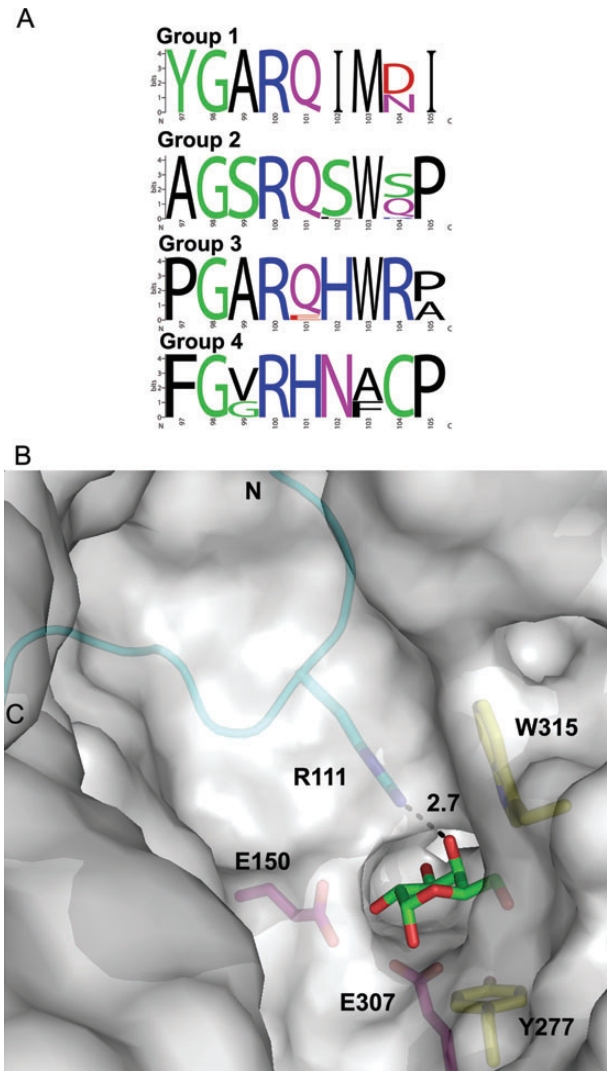

Fig. 2. Distinct sequences motifs flanking the active site in bifidobacteria GH42 enzymes. (**A**) Distinct sequence motifs directly associated with the individual specificities of Bga42A, Bga42B and Bga42C and are connected to the active site through R111 that contributes to the recognition of the non-reducing end β-Gal at subsite −1 through a hydrogen bond to the C4 hydroxyl group as observed in crystal structures (Hidaka et al. 2002; Maksimainen et al. 2012). (**B**) Crystal structure of the Bca-β-gal active site (numbering according to Maksimainen et al. 2012, PDB: 3TTY), highlighting (in cyan) the location of the loop that forms distinct sequence motifs (shown in panel A) associated with substrate specificity. The catalytic acid/base (E150) and nucleophile (E307) are colored in purple and the aromatic residues W315 and Y277 flanking the active site are shown in yellow.

changes the positioning of the pyranose ring at subsite +1. Remarkably, only a modest increase in catalytic efficiency is found toward the trisaccharides 4-galactotriose (Galβ1-4Galβ1-4Gal) and 4-galactosyllactose (Galβ1-4Galβ1-4Glc), suggesting that no recognition occurs beyond subsite +1 for Bga42B. Furthermore, the catalytic efficiency is extremely low for β1-6 linkages, whereas a more modest 10-fold decrease was seen for the β1-3 linked 3-galactosylglucose when compared with 4-galactobiose (Table I). An analysis of the genes surrounding Bga42B gene homologs (groups 2.1 and 2.2; Figure 1A) showed that a GH53 endo-galactanase encoding gene is found adjacent to the gene encoding GH42 in subgroup 2.1, but not in subgroup 2.2.

Bga42C is a putative exo-β1-4-galactosidase featuring in metabolism of arabinogalactan fragments

Bga42C has essentially the same k_{cat} for hydrolysis of 4-galactosyllactose and 4-galactobiose as Bga42B, however, with ∼15-fold higher K_m values. Bga42B and Bga42C thus share substrate preferences. Although Bga42C has ∼10-fold lower k_{cat}/K_m than Bga42B, its catalytic efficiency (k_{cat}/K_m) for the best substrates is comparable with that of Bga42A acting on 4-galactobiose and several other substrates. In a recent study, Bga42C was seen to prefer 4-galactobiose, but this substrate was not tested for Bga42B (Garrido et al. 2013). Bga42C has reasonable activity and affinity ($K_m = 0.90$ mM) toward 3-galactosylglucose, which is the preferred substrate of Bga42A, but Bga42C and Bga42A are distinguished by their specificity profiles (Table I), acting on distinct linkages to the terminal Gal at subsite −1. The Bga42C encoding gene and homologs thereof in *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* are consistently co-localized with genes encoding a GH43 α-l-arabinofuranosidase (EC 3.2.1.55), which catalyzes the removal of arabinosyl decorations from galactan main chains in the pectic polysaccharide arabinogalactan. This conserved gene organization within subgroup 4.1 in *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* suggests that Bga42C is likely to feature in the metabolism of arabinogalactan fragments.

Discussion

GOS are important commercial prebiotics able to modulate the composition of the microbiota of the human GIT and promote bifidobacterial growth in infants and adults (O'Connell Motherway et al. 2013). The three GH42 β-galactosidases Bga42A, Bga42B and Bga42C of *B. longum* subsp. *infantis* ATCC 15697 show distinct kinetic properties toward a series of milk and synthetic β-galactosides substrates (Table I) highlighting the diversity of β-galactosides utilization provided by probiotic bifidobacteria. While the activity on lactose, i.e. the most abundant β-galactoside oligosaccharide in human milk, by all three Bga42 enzymes is low ($k_{cat}/K_m = 0.08–6.1$ s^{-1} mM^{-1}; Yoshida et al. 2012), and e.g. the GH42 β-Gal II from *B. adolescentis* DSM 20083 was demonstrated to have no activity on lactose (Van Laere et al. 2000), it seems that some GH42 evolved to utilize different β-galactosides than lactose. In fact, both *B. longum* subsp. *infantis* ATCC 15697 and *B. adolescentis* DSM 20083 hydrolyze lactose through GH2 β-galactosidases (Van Laere et al. 2000; Yoshida et al. 2012). The catalytic efficiencies (k_{cat}/K_m) of the three enzymes of *B. longum* subsp. *infantis* ATCC 15697 span 5 orders of magnitude from 0.1 to 2745 mM^{-1} s^{-1} and data for a wide range of substrates including structural isomers attribute different subspecificities to Bga42A, Bga42B and Bga42C and identify key structural elements for the distinctive molecular recognition. A bioinformatic approach links certain enzymatic characteristics with defined groups of GH42 obtained by phylogenetic analysis (Figure 1).

Table I. Kinetic parameters of GH42 Bga42A, Bga42B and Bga42C from *B. longum* subsp. *infantis* ATCC 15697

Enzyme	Bga42A (Blon_2016)			Bga42B (Blon_2123)			Bga42C (Blon_2416)		
Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
4-galactobiose Galβ1-4Gal	3.7 ± 0.3	500 ± 14	136 ± 10	0.44 ± 0.04	746 ± 20	1697 ± 127	7.1 ± 0.5	635 ± 18	89 ± 4
Lactose Galβ1-4Glc[c]	16 ± 2	97 ± 3	6.1 ± 0.5	44 ± 10	8.6 ± 0.1	0.19 ± 0.00			0.08 ± 0.00
4-galactosyllactose Galβ1-4Galβ1-4Glc	18 ± 2	673 ± 48	37 ± 2	0.36 ± 0.03	755 ± 16	2098 ± 134	5.7 ± 1.1	676 ± 46	118 ± 16
4-galactotriose Galβ1-4Galβ1-4Gal	17 ± 1	126 ± 6	7.4 ± 0.3	0.26 ± 0.03	725 ± 21	2745 ± 289	3.7 ± 0.6	394 ± 17	107 ± 12
4-galactotetraose Galβ1-4Galβ1-4Galβ1-4Gal			0.8 ± 0.1	1.7 ± 0.2	795 ± 21	481 ± 52	14 ± 2	621 ± 45	45 ± 3
6-galactobiose Galβ1-6Gal	1.4 ± 0.2	578 ± 12	407 ± 40	9.0 ± 0.4	0.89 ± 0.02	0.10 ± 0.00	12 ± 4	2.0 ± 0.3	0.17 ± 0.03
6-galactotetraose Galβ1-6Galβ1-6Galβ1-6Gal	2.9 ± 0.3	338 ± 11	118 ± 11			n.d.			n.d.
6-galactosyllactose Galβ1-6Galβ1-4Glc			<LNT[a]			n.d.[a]			n.d.[a]
3-galactobiose Galβ1-3Gal	2.7 ± 0.3	1126 ± 47	422 ± 27	4.0 ± 0.4	297 ± 15	73 ± 4	4.4 ± 0.6	343 ± 24	78 ± 6
3-galactosylglucose Galβ1-3Glc	0.38 ± 0.04	628 ± 15	1671 ± 150	0.68 ± 0.06	119 ± 3	174 ± 11	0.90 ± 0.05	65 ± 0.1	73 ± 3
Lacto-N-*biose I (LNB)* Galβ1-3GlcNAc[c]	28 ± 2	86 ± 3	3.0 ± 0.2			<0.002			n.d.
3-galactosyllactose Galβ1-3Galβ1-4Glc			>>LNT[a]			n.d.			n.d.
3-galactobiosyllactose Galβ1-3 Galβ1-3Galβ1-4Glc[b]	0.83 ± 0.31	703 ± 162	842 ± 124			n.d.			n.d.
Lacto-N-*tetraose (LNT)* Galβ1-3GlcNAcβ1-3Galβ1-4Glc[N]	2.2 ± 0.3	524 ± 42	238 ± 10			<0.0004			n.d.
3-galactotriosyllactose Galβ1-3Galβ1-3Galβ1-3Galβ1-4Glc[O]	2.8 ± 0.4	304 ± 23	107 ± 8			n.d.			n.d.

Substrate specificities of GH42 Bga42A, Bga42B and Bga42C from *B. longum* subsp. *infantis* ATCC 15697. n.d., not detected.
[a]Activity level was assessed by TLC and compared with the hydrolysis of lacto-*N*-tetraose by Bga42A.
[b]Substrate inhibition by 3-galactobiosyllactose observed for Bga42A with $K_i = 2.1 \pm 0.8$ mM.
[c]From Yoshida et al. 2012 (measured at 30°C).

Table II. Bifidobacterium strains found in subgroups 2.1 and 2.2

Subgroup 2.1 Organisms (Gene Accession)	Subgroup 2.2 Organisms (Gene Accession)
B. breve ACS-071-V-Sch8b (AEF26610.1)	*B. adolescentis* DSM 20083 (AAR24113.1)
B. breve UCC2003 (ABE95118.1)	*B. adolescentis* ATCC 15703 (BAF40182.1)
B. longum NCC2705 (AAN24101.1)	*B. bifidum* JCM 1254 (BAI94828.1)
B. longum DJO10A (ACD98451.1)	*B. bifidum* PRL2010 (ADP35297.1)
B. longum subsp. *infantis* 157F (BAJ70496.1)	*B. dentium* Bd1 (ADB10488.1)
B. longum subsp. *longum* JCM 1217 (BAJ66022.1)	*B. longum* subsp. *infantis* JCM 1222 (BAJ69777.1)
B. longum subsp. *longum* KACC 91563 (AEI98104.1)	*B. longum* subsp. *infantis* 157F (BAJ70804.1)
B. longum subsp. *longum* BBMN68 (ADQ01982.1)	*B. longum* subsp. *infantis* ATCC 15697 (ACJ53184.1)
B. longum subsp. *longum* F8 (CBK70970.1)	
B. longum subsp. *longum* JDM301 (ADG99893.1)	

Subgroups 2.1 and 2.2 are solely populated by bifidobacteria species in a phylogenetic tree generated for the ~600 putative GH42 enzymes found in CAZy (not shown). The bifidobacteria species in these two subgroups are listed. *Bifidobacterium* (*B.*).

Bga42A is able to hydrolyze 3-galactosylglucose with higher efficiency than lacto-*N*-tetraose, which suggests that *B. longum* subsp. *infantis* ATCC 15697 can utilize both HMO type I, i.e. lacto-*N*-tetraose, and 3-galactobiosyllactose, through Bga42A. Interestingly, Bga42A is not included within a gene HMO cluster (Yoshida et al. 2012). One rationale for the broad specificity of Bga42A can be an evolutionary advantage in the transition from a strictly human milk-based diet in infants (HMO type I) to more broad galactoside intake from milk of other mammals and from plants after weaning (Urashima et al. 2012).

The genes in the β-galactoside metabolism are often organized together in functional regulons in probiotic lactobacilli (Andersen et al. 2011) and bifidobacteria (Andersen et al. 2013), which motivated analysis of the surrounding genes of Bga42B gene homologs (subgroups 2.1 and 2.2; Figure 1A). Clusters involving eight genes downstream and eight genes upstream the Bga42B homologs from ~350 bacteria were considered. Group 2 GH42 sequences were found to segregate in two main clades. Bga42B is in subgroup 2.2 together with BgalII, the only reported GH42 from *B. adolescentis* DSM20083 (*B. adolescentis* ATCC 15703; Van Laere et al. 2000). This enzyme has a specificity profile similar to Bga42B with essentially equally high efficiency toward 4-galactotriose ($k_{cat}/K_m = 2745\ s^{-1}\ mM^{-1}$) and 4-galactosyllactose ($k_{cat}/K_m = 2098\ s^{-1}\ mM^{-1}$) (Table I), although the apparent K_m for BgalII is generally 10-fold higher compared with Bga42B. Remarkably, subgroup 2.1 contains all the 10 bifidobacteria strains where a GH42 gene is adjacent to a GH53 gene encoding endo-galactanase that confers plant galatan utilization. Such clustering of GH42 and GH53 genes is also noted in soil niche Actinobacteria (e.g. *Streptomyces*) and Bacillaceae, e.g. *Geobacillus stearothermophilus* T-6 and *Bacillus subtilis* subsp. *subtilis* str. 168 (Shipkowski and Brenchley, 2006), but these enzymes show poor homology to group 2 bifidobacterial counterparts (Table II) that also cluster together in a phylogenetic analysis of all 587 GH42 sequences currently found in CAZy (not shown).

Bifidobacterium longum subsp. *infantis* 157F encodes two GH42 enzymes from subgroups 2.1 and 2.2, respectively. Notably, the likely in vivo substrates for Bga42B are hydrolysis products of galactans acquired through cross-feeding from other bacteria (Scott et al. 2013) as well as 4-galactosyllactose present in commercially produced GOS and bovine milk (Nakajima et al. 2006; Coulier et al. 2009; Marino et al. 2011). The 4-galactosyllactose concentration is comparable with those of 3- and 6-galactosyllactose in bovine colostrum (Urashima et al. 2013). The clear preference for 4-galactosyllactose of Bga42B and the phylogenetic analysis suggests an evolutionary scenario where GH42 sequences in bifidobacterial subgroup 2.1 evolved from ancestral enzymes, which together with GH53 extracellular endo-galactanases, conferred galactan utilization in soil-adapted bacteria. During adaptation to the gut niche, the presence of different oligomeric β-galactosides from mammalian milk and plant origin produced by the other taxa plausibly drove the evolution of subgroup 2.2 enzymes and the loss of the GH53 genes, whereas a few strains retained the ancestral gene organization and ability to utilize galactan as represented by enzymes in subgroup 2.1. Avoiding competition with specialized primary degraders that dominate polysaccharides depolymerization in the gut, e.g. members of the dominant commensal *Bacteroides* (Koropatkin et al. 2012), provides a plausible explanation for this evolutionary trajectory as these organisms are highly specialized and are less efficient in capture of smaller substrates (Dodd et al. 2011). This is also consistent with the observation that *B. longum* subsp. *infantis* (found in subgroup 2.2) exhibits more genomic potential to utilize HMO, whereas *B. longum* subsp. *longum* (found in subgroup 2.1) display higher genomic potential for utilization of plant-derived oligosaccharides. Such adaptation to the gut is highlighted in the specialization of *B. longum* subsp. *infantis* ATCC 15697 and related strains that target milk oligosaccharides from both humans and other mammals as supported by the data for Bga42A (subgroup 3.1) and Bga42B, respectively. This hypothesis is in accordance with the previous suggestion that bovine milk oligosaccharides can be used in milk products as a prebiotic for humans (Gopal and Gill 2000).

The genes surrounding Bga42A, Bga42B and Bga42C encoding genes also reveal large differences. Interestingly, *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* that have a GH42 with the sequence signature of subgroup 4.2 also contain a conserved GH43 α-l-arabinofuranosidase (EC 3.2.1.55) in the particular gene cluster (data not shown). It is therefore proposed that likely substrates for Bga42C are arabinogalactan-derived oligosaccharides, as the β1-4-Gal*p* arabinogalactan backbone is frequently branched with l-Ara*f*

(Hinz et al. 2005), which are acquired through cross-feeding from other bacteria (Scott et al. 2013) due to the lack of extracellular galactanases in *B. longum* subsp. *infantis* ATCC 15697.

The Bga42A, Bga42B and Bga42C from *B. longum* subsp. *infantis* possess clearly different specificity profiles governed by variation in substrate recognition at subsite +1. The present comparative kinetic analysis of a range of β-galactosides furthermore indicates only minor contribution to the specificity of these GH42 enzymes from interactions beyond subsite +1. Thus, the key substrate structural element for molecular recognition in GH42 is the glycosyl residue accommodated at subsite +1 and the glycosidic linkage to the galactosyl residue at subsite −1. The subspecificities can be assigned distinct sequence motifs in GH42 as apparent also in the phylogenetic tree. Bga42B has highest activity toward 4-galactosyllactose and a proposed name of this enzyme is therefore 4-galactosyllactose β1-4-galactosidase. The data clearly show a diverse β-galactoside metabolism within *B. longum* subsp. *infantis* and among probiotic bifidobacteria, which is an important evolutionary metabolic adaptation of bifidobacteria to the β-galactoside-rich gut niche.

Materials and methods

Bioinformatics and molecular graphics

The conserved domain Glyco_hydro_42 (pfam02449; Marchler-Bauer et al. 2013) of all 68 bifidobacteria GH42 was extracted from CAZy (Cantarel et al. 2009), and additional 54 full-length sequences from bifidobacteria were extracted from GenBank using Glyco_hydro_42 (pfam02449; Marchler-Bauer et al. 2013). Redundant sequences were removed, resulting in 61 full-length GH42 sequences. The sequence motifs derived from the phylogenetic analysis of 61 non-redundant GH42 sequences from bifidobacteria were submitted to BLASTP and the organisms encoding GH42 enzyme with a 100% positive match were extracted. All multiple sequence alignments were performed with the program MUSCLE v3.8 (Edgar 2004) using default settings. A phylogenetic tree was constructed using the ClustalW2 phylogeny (Larkin et al. 2007) with default settings and visualized in Dendroscope 3.0 (Huson and Scornavacca 2012). Analysis of surrounding genes relied on the automatic annotation by sequence similarity of the eight downstream and eight upstream genes from the respective GH42 β-galactosidase. A model showing the location of nine residues associated with substrate specificity was created using the crystal structure of Bcα-β-Gal (Maksimainen et al. 2012; PDB ID: 3TTY) and explored using PyMOL (Schrödinger, LLC, New York).

Substrates

D-Gal and 4-galactosyllactose were purchased from Wako Pure Chemical Industries (Osaka, Japan); 6-galactobiose was from Dextra Laboratories (Reading, UK); 4-galactobiose from Megazyme (Wicklow, Ireland) and 6-galactosyllactose from Carbosynth (Berkshire, UK). 3-Galactobiose and 3-galactosylglucose were synthesized as described (Nakajima et al. 2009). 4-Galactotriose, 4-galactotetraose and 6-galactotetraose were chemically synthesized (MCFA and MHC, unpublished). 3-Galactosyllactose, 3-galactobiosyl lactose and 3-galactotriosyllactose were purified essentially as described (Messer et al. 1980; Collins et al. 1981) from Tammar Wallaby and Wombat milk (generous gifts of Dr Michael Messer, School of Molecular and Microbial Biosciences, University of Sydney, Australia).

Enzyme preparation

Bifidobacterium longum subsp. *infantis* ATCC 15697 Bga42A, Bga42B and Bga42C (locus tags, blon_2016, blon_2123, and blon_2416) were produced recombinantly with C-terminal His_6-tag (Yoshida et al. 2012) and purified by Ni^{2+}-nitrilotriacetic acid affinity chromatography followed by size exclusion chromatography on a Superdex 200 10/300 GL column (GE Healthcare Bio-Sciences, Uppsala, Sweden). Protein concentration was determined spectrophotometrically using ε_M at 280 nm calculated (Wilkins et al. 1999) for Bga42A, Bga42B and Bga42C to 172,270, 172,270 and 172,240 M^{-1} cm^{-1}, respectively. Specific activities of Bga42A, Bga42B and Bga42C toward lactose were verified to be as reported (Yoshida et al. 2012).

Enzyme assay and substrate specificity analysis

Bga42A, Bga42B and Bga42C (60–65 nM) were incubated with 2 mM 3-galactosyllactose, 4-galactosyllactose or 6-galactosyllactose in 50 mM sodium phosphate pH 6.5 at 30° C for 20 and 60 min. At appropriate time intervals, 1 µl aliquots was spotted onto thin-layer chromatography (TLC) silica gel plates, developed (mobile phase 80:20 acetonitrile:water), dried, dipped in sulfuric acid-methanol and tarred.

A continuous coupled assay using Gal mutarotase, galactokinase, UDP-Glc hexose-1-phosphate uridylyltransferase, phosphoglucomutase and Glc 6-phosphate dehydrogenase (Nihira et al. 2007) was performed on 6-galactobiose, 6-galactotetraose, 3-galactobiose, 3-galactobiosyllactose, 3-galactotriosyllactose, 4-galactobiose, 4-galactotriose, 4-galactosyllactose or 4-galactotetraose (0.2–20 mM) in 40 mM sodium citrate, pH 6.5, 0.05% Triton X-100 (total volume: 50 µL) at 37°C. Reactions were initiated by addition of pre-incubated diluted enzyme (final 0.008–18 nM) at 37°C into 384-well plates and released Gal was monitored every 30 s for up to 90 min by colorimetric quantification of α-D-Gal 1-phosphate as described (Nihira et al. 2007). The kinetic parameters k_{cat} and K_m were calculated by curve fitting the Michaelis–Menten equation to initial rate data using GraphPad Prism 5.0 (GraphPad Software). When saturation was not observed, k_{cat}/K_m was determined from linear rates of product formation at appropriate substrate concentrations.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

The Technical University of Denmark is acknowledged for two PhD stipends (to AHV and MCFA). The Idella Foundation and the Scandinavia-Japan Sasakawa Foundation are thanked for travel support (to AHV). This work was supported in part by a grant from Science and Technology Research Promotion

Program for Agriculture, Forestry, Fisheries and Food Industry, Japan. The employment of TK is supported by the Institute for Fermentation, Osaka (IFO). The Danish Strategic Research Council is acknowledged for support of the GlycAct project (MHC and MCFA).

Acknowledgements

We are grateful to Dr Michael Messer (School of Molecular and Microbial Biosciences, University of Sydney, Australia) for the generous gift of mammalian milk samples containing 3-galactosyllactose, 3-galactobiosyllactose and 3-galactotriosyllactose.

Conflict of interest

None declared.

Abbreviations

CAZy, Carbohydrate-Active enZYmes; DP, degree of polymerization; Gal, galactose; GH, glycoside hydrolase; GIT, gastrointestinal tract; Glc, glucose; GlcNAc, N-acetylglucosamine; GOS, galactooligosaccharide; HMO, human milk oligosaccharide; TLC, thin-layer chromatography.

References

Andersen JM, Barrangou R, Abou Hachem M, Lahtinen S, Goh YJ, Svensson B, Klaenhammer TR. 2011. Transcriptional and functional analysis of galactooligosaccharide uptake by *lac*S in *Lactobacillus acidophilus*. *Proc Natl Acad Sci USA*. 108:17785–17790.

Andersen JM, Barrangou R, Abou Hachem M, Lahtinen SJ, Goh YJ, Svensson B, Klaenhammer T. 2013. Transcriptional analysis of oligosaccharide utilization by *Bifidobacterium lactis* Bl-04. *BMC Genomics*. 14:312.

Asakuma S, Hatakeyama E, Urashima T, Yoshida E, Katayama T, Yamamoto K, Kumagai H, Ashida H, Hirose J, Kitaoka M. 2011. Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem*. 286:34583–34592.

Barboza M, Sela DA, Pirim C, Locascio RG, Freeman SL, German JB, Mills DA, Lebrilla CB. 2009. Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Appl Environ Microbiol*. 75:7319–7325.

Boehm G, Stahl B. 2007. Oligosaccharides from milk. *J Nutr*. 137:847S–849S.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res*. 37:D233–D238.

Collins JG, Bradbury JH, Trifonoff E, Messer M. 1981. Structures of four new oligosaccharides from marsupial milk, determined mainly by 13C-n.m.r. spectroscopy. *Carbohydr Res*. 92:136–140.

Coulier L, Timmermans J, Bas R, Van Den Dool R, Haaksman I, Klarenbeek B, Slaghek T, Van Dongen W. 2009. In-depth characterization of prebiotic galacto-oligosaccharides by a combination of analytical techniques. *J Agric Food Chem*. 57:8488–8495.

Davies GJ, Wilson KS, Henrissat B. 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J*. 321(Pt 2):557–559.

Davis LMG, Martinez I, Walter J, Goin C, Hutkins RW. 2011. Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans. *PLoS One*. 6:e25200.

Di Lauro B, Strazzulli A, Perugino G, La Cara F, Bedini E, Corsaro MM, Rossi M, Moracci M. 2008. Isolation and characterization of a new family 42 β-galactosidase from the thermoacidophilic bacterium *Alicyclobacillus acidocaldarius*: Identification of the active site residues. *Biochim Biophys Acta*. 1784:292–301.

Dodd D, Mackie RI, Cann IK. 2011. Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. *Mol Microbiol*. 79:292–304.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.

Garrido D, Ruiz-Moyano S, Jimenez-Espinoza R, Eom HJ, Block DE, Mills DA. 2013. Utilization of galactooligosaccharides by *Bifidobacterium longum* subsp. *infantis* isolates. *Food Microbiol*. 33:262–270.

Gopal PK, Gill HS. 2000. Oligosaccharides and glycoconjugates in bovine milk and colostrum. *Br J Nutr*. 84(Suppl 1):S69–S74.

Han NS, Kim TJ, Park YC, Kim J, Seo JH. 2011. Biotechnological production of human milk oligosaccharides. *Biotechnol Adv*. 30:1268–1278.

Hidaka M, Fushinobu S, Ohtsu N, Motoshima H, Matsuzawa H, Shoun H, Wakagi T. 2002. Trimeric crystal structure of the glycoside hydrolase family 42 β-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose. *J Mol Biol*. 322:79–91.

Hildebrandt P, Wanarska M, Kur J. 2009. A new cold-adapted β-d-galactosidase from the Antarctic *Arthrobacter* sp. 32c - gene cloning, overexpression, purification and properties. *BMC Microbiol*. 9:151.

Hinz SW, van den Broek LA, Beldman G, Vincken JP, Voragen AG. 2004. β-Galactosidase from *Bifidobacterium adolescentis* DSM20083 prefers β(1,4)-galactosides over lactose. *Appl Microbiol Biotechnol*. 66:276–284.

Hinz SW, Verhoef R, Schols HA, Vincken JP, Voragen AG. 2005. Type I arabinogalactan contains β-d-Gal*p*-(1-3)-β-d-Gal*p* structural elements. *Carbohydr Res*. 340:2135–2143.

Holmes ML, Scopes RK, Moritz RL, Simpson RJ, Englert C, Pfeifer F, Dyall-Smith ML. 1997. Purification and analysis of an extremely halophilic β-galactosidase from *Haloferax alicantei*. *Biochim Biophys Acta*. 1337:276–286.

Hu JM, Li H, Cao LX, Wu PC, Zhang CT, Sang SL, Zhang XY, Chen MJ, Lu JQ, Liu YH. 2007. Molecular cloning and characterization of the gene encoding cold-active β-galactosidase from a psychrotrophic and halotolerant *Planococcus* sp. L4. *J Agric Food Chem*. 55:2217–2224.

Hung MN, Xia Z, Hu NT, Lee BH. 2001. Molecular and biochemical analysis of two β-galactosidases from *Bifidobacterium infantis* HL96. *Appl Environ Microbiol*. 67:4256–4263.

Huson DH, Scornavacca C. 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst Biol*. 61:1061–1067.

Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol*. 10:323–335.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*. 23:2947–2948.

Lee JH, Kim YS, Yeom SJ, Oh DK. 2011. Characterization of a glycoside hydrolase family 42 β-galactosidase from *Deinococcus geothermalis*. *Biotechnol Lett*. 33:577–583.

Maksimainen M, Paavilainen S, Hakulinen N, Rouvinen J. 2012. Structural analysis, enzymatic characterization, and catalytic mechanisms of β-galactosidase from *Bacillus circulans* sp. *alkalophilus*. *FEBS J*. 279:1788–1798.

Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, et al. 2013. CDD: Conserved domains and protein three-dimensional structure. *Nucleic Acids Res*. 41:D348–D352.

Marino K, Lane JA, Abrahams JL, Struwe WB, Harvey DJ, Marotta M, Hickey RM, Rudd PM. 2011. Method for milk oligosaccharide profiling by 2-aminobenzamide labeling and hydrophilic interaction chromatography. *Glycobiology*. 21:1317–1330.

Messer M, Trifonoff E, Stern W, Collins JG, Bradbury JH. 1980. Structure of a marsupial-mild trisaccharide. *Carbohydr Res*. 83:327–334.

Moore JB, Markiewicz P, Miller JH. 1994. Identification and sequencing of the *Thermotoga maritima lacZ* gene, part of a divergently transcribed operon. *Gene*. 147:101–106.

Nakajima K, Kinoshita M, Matsushita N, Urashima T, Suzuki M, Suzuki A, Kakehi K. 2006. Capillary affinity electrophoresis using lectins for the analysis of milk oligosaccharide structure and its application to bovine colostrum oligosaccharides. *Anal Biochem*. 348:105–114.

Nakajima M, Nishimoto M, Kitaoka M. 2009. Characterization of three β-galactoside phosphorylases from *Clostridium phytofermentans*: Discovery of d-galactosyl-β1-4-l-rhamnose phosphorylase. *J Biol Chem*. 284:19220–19227.

NIH HMP Working GroupPeterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, et al. 2009. The NIH Human Microbiome Project. *Genome Res*. 19: 2317–2323.

Nihira T, Nakajima M, Inoue K, Nishimoto M, Kitaoka M. 2007. Colorimetric quantification of α-d-galactose 1-phosphate. *Anal Biochem*. 371:259–261.

O'Connell Motherway M, Fitzgerald GF, van Sinderen D. 2011. Metabolism of a plant derived galactose-containing polysaccharide by *Bifidobacterium breve* UCC2003. *Microb Biotechnol*. 4:403–416.

O'Connell Motherway M, Kinsella M, Fitzgerald GF, van Sinderen D. 2013. Transcriptional and functional characterization of genetic elements involved in galacto-oligosaccharide utilization by *Bifidobacterium breve* UCC2003. *Microb Biotechnol*. 6:67–79.

Pokusaeva K, Fitzgerald GF, van Sinderen D. 2011. Carbohydrate metabolism in Bifidobacteria. *Genes Nutr*. 6:285–306.

Ridley BL, O'Neill MA, Mohnen D. 2001. Pectins: Structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry*. 57:929–967.

Roberfroid M, Gibson GR, Hoyles L, McCartney AL, Rastall R, Rowland I, Wolvers D, Watzl B, Szajewska H, Stahl B, et al. 2010. Prebiotic effects: Metabolic and health benefits. *Br J Nutr*. 104(Suppl 2):S1–S63.

Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, Zwahlen MC, Desiere F, Bork P, Delley M, et al. 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci USA*. 99:14422–14427.

Scott KP, Gratz SW, Sheridan PO, Flint HJ, Duncan SH. 2013. The influence of diet on the gut microbiota. *Pharmacol Res*. 69:52–60.

Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, Lapidus A, Rokhsar DS, Lebrilla CB, German JB, et al. 2008. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci USA*. 105:18964–18969.

Shipkowski S, Brenchley JE. 2006. Bioinformatic, genetic, and biochemical evidence that some glycoside hydrolase family 42 β-galactosidases are arabinogalactan type I oligomer hydrolases. *Appl Environ Microbiol*. 72:7730–7738.

Tabachnikov O, Shoham Y. 2013. Functional characterization of the galactan utilization system of *Geobacillus stearothermophilus*. *FEBS J*. 80:950–964.

Urashima T, Asakuma S, Leo F, Fukuda K, Messer M, Oftedal OT. 2012. The predominance of type I oligosaccharides is a feature specific to human breast milk. *Adv Nutr*. 3:473S–482S.

Urashima T, Saito T, Nakamura T, Messer M. 2001. Oligosaccharides of milk and colostrum in non-human mammals. *Glycoconj J*. 18:357–371.

Urashima T, Taufik E, Fukuda K, Asakuma S. 2013. Recent advances in studies on milk oligosaccharides of cows and other domestic farm animals. *Biosci Biotechnol Biochem*. 77:455–466.

van den Broek LAM, Hinz SWA, Beldman G, Vincken J, Voragen AGJ. 2008. Bifidobacterium carbohydrases-their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res*. 52:146–163.

Van Laere KM, Abee T, Schols HA, Beldman G, Voragen AG. 2000. Characterization of a novel β-galactosidase from *Bifidobacterium adolescentis* DSM 20083 active towards transgalactooligosaccharides. *Appl Environ Microbiol*. 66:1379–1384.

Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, Baumer S, Henne A, Liesegang H, Merkl R, et al. 2004. The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential. *J Mol Microbiol Biotechnol*. 7:204–211.

Vincken JP, Schols HA, Oomen RJ, McCann MC, Ulvskov P, Voragen AG, Visser RG. 2003. If homogalacturonan were a side chain of rhamnogalacturonan I. Implications for cell wall architecture. *Plant Physiol*. 132:1781–1789.

Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 112:531–552.

Yoshida E, Sakurama H, Kiyoharan M, Nakajima M, Kitaoka M, Ashida H, Hirose J, Katayama T, Yamamoto K, 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.