Novel 3-O- α -D-Galactosyl- α -L-Arabinofuranosidase for the Assimilation of Gum Arabic Arabinogalactan Protein in *Bifidobacterium longum* subsp. *longum*

Yuki Sasaki,^a Ayako Horigome,^b Toshitaka Odamaki,^b Jin-Zhong Xiao,^b Akihiro Ishiwata,^c Yukishige Ito,^{c,d} Kanefumi Kitahara,^{a,e} Kiyotaka Fujita^{a,e}

^aThe United Graduate School of Agricultural Sciences, Kagoshima University, Kagoshima, Kagoshima, Japan

^bNext Generation Science Institute, Morinaga Milk Industry Co., Ltd., Zama, Kanagawa, Japan

^cRIKEN Cluster for Pioneering Research, Wako, Saitama, Japan

^dGraduate School of Science, Osaka University, Toyonaka, Osaka, Japan

^eFaculty of Agriculture, Kagoshima University, Kagoshima, Kagoshima, Japan

ABSTRACT Gum arabic arabinogalactan (AG) protein (AGP) is a unique dietary fiber that is degraded and assimilated by only specific strains of *Bifidobacterium longum* subsp. *longum*. Here, we identified a novel 3-O- α -D-galactosyl- α -L-arabinofuranosidase (GAfase) from *B. longum* JCM7052 and classified it into glycoside hydrolase family 39 (GH39). GAfase released α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from larch AGP, and the α -D-Galp-(1 $\rightarrow$ 3)-L-Ara release activity was found to be 594-fold higher than that of β -L-Arap-(1 $\rightarrow$ 3)-L-Ara. The GAfase gene was part of a gene cluster that included genes encoding a GH36 α -galactosidase candidate and ABC transporters for the assimilation of the released α -D-Galp-(1 $\rightarrow$ 3)-L-Ara in *B. longum*. Notably, when α -D-Galp-(1 $\rightarrow$ 3)-L-Ara was removed from gum arabic AGP, it was assimilated by both *B. longum* JCM7052 and the nonassimilative *B. longum* JCM1217, suggesting that the removal of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP by GAfase permitted the cooperative action with type II AG degradative enzymes in *B. longum*. The present study provides new insight into the mechanism of gum arabic AGP degradation in *B. longum*.

IMPORTANCE Bifidobacteria harbor numerous carbohydrate-active enzymes that degrade several dietary fibers in the gastrointestinal tract. *B. longum* JCM7052 is known to exhibit the ability to assimilate gum arabic AGP, but the key enzyme involved in the degradation of gum arabic AGP remains unidentified. Here, we cloned and characterized a GH39 3-O- α -D-galactosyl- α -L-arabinofuranosidase (GAfase) from *B. longum* JCM7052. The enzyme was responsible for the release of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP. The presence of a gene cluster including the GAfase gene is specifically observed in gum arabic AGP assimilative strains. However, GAfase carrier strains may affect GAfase noncarrier strains that express other type II AG degradative enzymes. These findings provide insights into the bifidogenic effect of gum arabic AGP.

KEYWORDS arabinogalactan protein, bifidobacteria, dietary fiber, glycosidase, gum arabic, gut microbiota

Bifidobacteria are natural inhabitants of the human gut, and the engraftment and abundance of these bacteria in the human intestine are associated with the health of their host (1). In the human intestine, the predominant *Bifidobacterium* species varies with host age (2, 3); *Bifidobacterium breve*, *B. longum* subsp. *infantis*, and *B. bifidum* are frequently observed in the infant intestine, whereas *B. adolescentis*, *B. catenulatum*, and *B. pseudocatenulatum* have been detected in the adult intestine. Conversely, *B. longum* subsp. *longum* has

Citation Sasaki Y, Horigome A, Odamaki T, Xiao J-Z, Ishiwata A, Ito Y, Kitahara K, Fujita K. 2021.

Novel 3-O- α -D-galactosyl- α -L-arabinofuranosidase for the assimilation of gum arabic arabinogalactan protein in *Bifidobacterium longum* subsp. *longum*. Appl Environ Microbiol 87:e02690-20. <https://doi.org/10.1128/AEM.02690-20>.

Editor Haruyuki Atomi, Kyoto University

Copyright © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Kiyotaka Fujita, k4022897@kadai.jp.

Received 3 November 2020

Accepted 19 February 2021

Accepted manuscript posted online 5 March 2021

Published 27 April 2021

FIG 1 Schematic structures of AGPs and oligosaccharides used in this study. The oligosaccharides that featured the α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf- structure (S3-GA and S5-GA) and β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf- structure (S3-AA) were purified from free oligosaccharides in gum arabic reagent. Arrows indicate the cleavage sites for Gafase.

been detected in hosts across a wider age range, from infancy to adult (4). This age-related variation in distribution may occur due to environmental factors, such as a shift in the main nutrient in diet (5, 6), diminished function of the host immune system, and difference in the composition of the host mucosa (7, 8). The ability of bacteria to survive in the human intestine was recently shown to be dependent on the presence of carbohydrate-active enzymes that act on diet- and/or host-derived glycans (9–11).

Arabinogalactan (AG) proteins (AGPs) are complex proteoglycans that are widely distributed in plant species and play a crucial role in various physiological functions, such as specialization and growth (12). AGPs are composed mainly of type II AG chains, which feature a branched structure comprising a β 1,3-galactan chain backbone and β 1,6-galactan side chains. The complexity of the β 1,6-galactan side chain differs according to plant species. For example, gum arabic AGP is heavily modified by L-arabinofuranose, L-arabinopyranose, L-rhamnose, L-fucose, glucuronic acid, and 4-O-methyl-glucuronic acid (13, 14), whereas larch AGP is simply modified by L-arabinofuranose and L-arabinopyranose (15) (Fig. 1).

AGPs are found in normal human diet, such as in vegetables, fruits, and grains (16, 17), and gum arabic AGP is commonly used as a food additive for emulsification and stabilization (18). Additionally, gum arabic AGP has been reported to increase the concentration of bifidobacteria in the descending colon, based on the results of studies conducted using a simulator of the human intestinal microbial ecosystem (19, 20) and *in vivo* experiments in human

subjects (21). Notably, *in vitro* assimilation tests performed using isolated *Bifidobacterium* species revealed that certain *B. longum* and *B. adolescentis* strains can proliferate on gum arabic AGP (22, 23), and *B. longum* JCM7052 and JCM7053 were recognized to exhibit the ability to assimilate gum arabic AGP (24). However, the key carbohydrate-active enzyme involved in the degradation of gum arabic AGP remains unidentified.

In previous studies, we have characterized type II AG degradative enzymes in *B. longum* JCM1217 (29, 40). Larch AGP was degraded by the synergistic action of glycoside hydrolase (GH) family 43_24 exo- β -1,3-galactanase (BI1,3Gal), GH30_5 exo- β -1,6-galactobiohydrolase (BI1,6Gal), and GH43_22 α -L-arabinofuranosidase (BIArA) on the surface of the bacterial cell. In contrast, JCM1217 was found to be a nonassimilative strain for gum arabic AGP, and the aforementioned enzymes showed little activity toward gum arabic AGP. Therefore, we predicted the presence of a novel enzyme for the degradation of gum arabic AGP in specific *B. longum* strains that can assimilate gum arabic AGP. In this study, we compared gum arabic AGP assimilative strains and nonassimilative strains by performing *in vitro* assimilation tests and comparative genome analysis and identified a key enzyme that plays a crucial role in gum arabic AGP assimilation in *B. longum*. We also propose here a rational strategy for the assimilation of gum arabic AGP in *B. longum*.

RESULTS

***In vitro* assimilation test of AGPs and bacterial enzyme activity assays.** First, we tested the ability of 12 *B. longum* strains to grow in medium containing larch AGP or gum arabic AGP as the sole carbon source. The assimilation abilities of the strains were assessed by measuring the increase in 600-nm absorbance (Fig. 2A) and the decline in pH (data not shown) of the medium after 48 h. All *B. longum* strains grew on larch AGP, whereas only *B. longum* JCM7052 and MCC00300 strains exhibited notable growth on gum arabic AGP. Subsequently, we assayed oligosaccharide-releasing activities from gum arabic AGP by using bacterial cells of the 12 *B. longum* strains grown in larch AGP and the two strains grown in gum arabic AGP (Fig. 2B and C). Intriguingly, a distinct peak (peak X) with a retention time of 6 min was observed only in the case of the strains capable of assimilating gum arabic AGP. Moreover, the peak X-releasing activity was not detected in the culture supernatant of *B. longum* JCM7052 (data not shown), suggesting that the releasing enzyme was a protein anchored to the bacterial cell wall.

To determine the sugar composition of peak X, we performed acid hydrolysis (Fig. 3A) and enzymatic degradation with GH27 α -D-galactopyranosidase/ β -L-arabinopyranosidase BAD_1525 from *B. adolescentis* (25) (Fig. 3B). In both cases, peak X was degraded into L-arabinose and galactose at a 1:1 molar ratio. 3-O- α -D-Galactopyranosyl-L-arabinose [α -D-Galp-(1 $\rightarrow$ 3)-L-Ara] has previously been isolated from gum arabic AGP through partial acid hydrolysis (26), with the retention time corresponding to that observed for peak X (Fig. 3B). These results suggested that peak X represented a disaccharide, α -D-Galp-(1 $\rightarrow$ 3)-L-Ara, attached to the end of the side chains of gum arabic AGP.

Genome comparative analysis. To identify the gene cluster involved in the assimilation of gum arabic AGP, we analyzed the genome sequences of 10 *B. longum* Morinaga Culture Collection (MCC) strains. The general features of each newly determined *B. longum* genome are presented in Table S1 in the supplemental material. Comparative analysis of the genomes of the 12 *B. longum* strains revealed a gene cluster (BLGA_00290 to BLGA_00350) in *B. longum* JCM7052 that was conserved in the gum arabic AGP assimilative strain MCC00300 but not in other, nonassimilative strains (Fig. 4A; Fig. S1). Thus, we hypothesized that this candidate gene cluster may endow *B. longum* with assimilation ability. The gene cluster comprised genes encoding two glycoside hydrolases (BLGA_00330 and BLGA_00340), a Lac-I-type transcriptional regulator (BLGA_00290), and ABC transporter components, including an extracellular solute-binding protein (BLGA_00300), two ABC transporter permeases (BLGA_00310 and BLGA_00320), and an ATP-binding protein (BLGA_00350). Furthermore, the candidate gene cluster in JCM7052 showed 100% identity at 100% coverage with that in MCC00300 (MCC00300_11160 to MCC00300_11100). BLGA_00330 exhibited 47.0% sequence identity at 84% coverage with a previously characterized GH36

FIG 2 *In vitro* assimilation test and sugar-releasing activity toward gum arabic AGP in *B. longum*. (A) Growth of *B. longum* strains was monitored in MRS medium containing larch AGP (gray bars) or gum arabic AGP (black bars) as the sole carbon source. The absorbance of growth medium was calculated by subtracting the absorbance of the initial medium. (B and C) Gum arabic AGP was incubated with bacterial cell fractions of *B. longum* strains grown in larch AGP (B) and gum arabic AGP (C). Reaction products were analyzed by HPAEC-PAD performed using a PA-1 column. Arrows indicate released sugar (peak X).

α -galactosidase from *Streptomyces* sp. strain S27 (ACCC41168) (GenBank accession no. [ACN78885.1](#)) (27). The gene cluster was replaced with several types of genes in the nonassimilative strains (Fig. S1).

BLGA_00340 (GenBank accession no. [BBV22623.1](#)), which was annotated as a GH39 member, contained a putative signal peptide, a catalytic domain, two carbohydrate-

FIG 3 Analysis of the liberated peak X by using acid hydrolysis (A) and enzymatic hydrolysis (B). (A) Purified peak X (bottom) was analyzed by performing acid hydrolysis treatment (middle). The sugar composition was analyzed through HPAEC-PAD by using a PA-10 column. L-Fucose (Fuc), L-rhamnose (Rha), L-arabinose (Ara), and galactose (Gal) were used as standards (top). (B) Peak X and a standard preparation of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara were incubated with (+) or without (-) GH27 α -D-galactopyranosidase/ β -L-arabinopyranosidase BAD_1525 and then analyzed by performing HPAEC-PAD with a PA-1 column.

binding module (CBM) 35 domains, a predicted CBM, a bacterial immunoglobulin (Ig)-like domain, and a transmembrane domain (Fig. 4B). Notably, the catalytic region of BLGA_00340 shared a low level of sequence identity (<25%) with the catalytic regions of other characterized enzymes: GH39 α -L-(β -1,2)-arabinofuranobiosidase (NF2152) and D-galacto-(α 1,2)-L-arabinosidase (NF2523) from rumen anaerobic fungi and GH39 β -xylosidases (Fig. 4C). Phylogenetic analysis of BLGA_00340 revealed a slightly closer relationship of BLGA_00340 with NF2152 and NF2523 than GH39 β -xylosidases. The homologous genes were also sporadically observed in other adult gut-associated bifidobacterial species, such as *B. adolescentis* (sequence identity, 63.8%), *B. catenulatum* (63.8%), and *B. pseudocatenulatum* (60.0%) (Fig. 4C).

Only one of the 10 MCC strains possessed the candidate gene (Fig. S1). Furthermore, we found that BLGA_00340 homologous genes were present in 4.42% of other *B. longum* MCC strains ($n = 113$) isolated from fecal samples of Japanese individuals belonging to a broad age group (28). In total, BLGA_00340 homologous genes were conserved in 6.84% of *B. longum* strains ($n = 307$) in the NCBI database. Moreover, BLGA_00340 homologous genes were conserved in 7.02% of *B. adolescentis* strains ($n = 57$), 12.5% of *B. catenulatum* strains ($n = 8$), and 23.4% of *B. pseudocatenulatum* strains ($n = 77$) in the NCBI database.

Preparation of recombinant GAfase. Recombinant BLGA_00340 protein (named 3-O- α -D-galactosyl- α -L-arabinofuranosidase [GAfase]) was designed without the N-terminal signal peptide and the C-terminal 162 amino acids (aa) (C Δ 162), including the bacterial Ig-like domain and the transmembrane domain. The recombinant molecule was highly expressed as a soluble protein at 25°C and was purified using a C-terminal His tag. The purified recombinant GAfase was found to migrate as a single band with an apparent molecular mass of 102 kDa on SDS-PAGE, which corresponded to its calculated molecular mass of 101,871 Da (Fig. S2).

Polysaccharide-degradative activity of recombinant GAfase. We tested the ability of gum arabic AGP, larch AGP, sugar beet arabinan, and wheat arabinoxylan to serve as potential substrates for GAfase (Fig. 5). Recombinant GAfase mainly released α -D-Galp-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP, which corresponded with peak X released

FIG 4 Gene clusters and phylogenetic relationships of BLGA_00340 (GAfase gene). (A) A gene cluster for assimilation of gum arabic AGP in *B. longum* JCM7052. The gene cluster was compared with the corresponding cluster in *B. longum* JCM1217. The arrowheads (with names below) indicate genes annotated in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The glycoside hydrolases in the Carbohydrate-Active enZymes (CAZy) database are shown inside the red arrowheads. Light gray bars indicate orthologous regions (>99% identity). (B) Domain structure of GAfase. Domain structures were predicted using SignalP4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) and InterPro (<http://www.ebi.ac.uk/interpro/>) servers. The overall structure comprises the signal peptide (SP), the GH39 catalytic domain, two CBM35 domains, a galactose-binding domain-like region, an Ig-like domain (Ig), and the transmembrane (TM) domain. (C) Phylogenetic tree of GAfase and GH39 family members. The homologous proteins were derived based on the hypothesized protein from the *Bifidobacterium* species and the characterized proteins GH39 D-galacto-(α -1,2)-L-arabinosidase [G(1,2)Afase], α -L-(β -1,2)-L-arabinofuranobiosidase [A(1,2)Afase], and β -xylosidases. (Continued on next page)

FIG 5 HPAEC-PAD analysis of GAfase reaction with polysaccharides. Gum arabic AGP (bottom), larch AGP (middle), and sugar beet arabinan (top) were incubated with (+) or without (-) GAfase at 37°C for 16 h.

by the bacterial enzyme. The enzyme also released 3-*O*- β -L-arabinopyranosyl-L-arabinose [β -L-Arap-(1 $\rightarrow$ 3)-L-Ara] from gum arabic AGP and larch AGP. The β -L-Arap-(1 $\rightarrow$ 3)-L-Ara structure was inferred based on sugar composition analysis, enzymatic treatment, and a previous study relevant to AGP structures (29). A small amount of β -L-Arap-(1 $\rightarrow$ 3)-L-Ara was also detected from sugar beet arabinan but not from wheat arabinoxylan (data not shown). The relative activity of GAfase was 10.2% for larch AGP and 0.198% for sugar beet arabinan compared to gum arabic AGP (Table 1). The optimal temperature and pH for GAfase activity were 45 to 50°C and 4.5, respectively, when gum arabic AGP was used as the substrate (Fig. S3).

To quantify the residues of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara present in polysaccharides, AGPs and sugar beet arabinan were incubated for 3 days with an excess amount of GAfase. The amounts of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara released from gum arabic AGP were 9.36% (wt/wt) and 1.87% (wt/wt), respectively; in the case of larch AGP, the amount of β -L-Arap-(1 $\rightarrow$ 3)-L-Ara released was 1.24% (wt/wt). In addition, the amount of β -L-Arap-(1 $\rightarrow$ 3)-L-Ara released from arabinan was found to be only 0.153% (wt/wt).

Transglycosylation activity of recombinant GAfase. Recombinant GAfase was allowed to react with gum arabic AGP in the presence of 30% methanol, ethanol, or 1-propanol, and transglycosylation products were detected using thin-layer chromatography (TLC) (Fig. 6A). The purified transglycosylation product from the reaction in 30% methanol was then hydrolyzed to α -D-Galp-(1 $\rightarrow$ 3)-L-Ara by the GAfase treatment (Fig. 6B). Furthermore, we identified the structure of the transglycosylation product by performing nuclear magnetic resonance (NMR) analysis, including ^1H , ^{13}C , and basic two-dimensional NMR (2D-NMR) techniques (Fig. 6C and D and Table 2). The observed multiplicity of the anomeric proton signal (singlet at 4.98 ppm) was consistent with methyl α -L-arabinofuranoside (30), which indicated that the methanol was transferred in the

FIG 4 Legend (Continued)

The phylogenetic tree was constructed using the neighbor-joining method and the aligned sequences. For construction of the tree, the program MUSCLE was implemented in MEGA7 software (60). GAfase is indicated by a dashed-line box. The characterized enzymatic activities or locus tags are shown alongside the abbreviated names of the organisms as follows: B. Cat_BBCT_0495, *Bifidobacterium catenulatum* BBCT_0495 (GenBank accession no. [BAR01463.1](#)); B. Pse_BAD12_RS08860, *B. pseudocatenulatum* BAD12_RS08860 ([WP_065438160.1](#)); B. Lon_GAfase, *B. longum* GAfase; B. Ado_B5789_0044, *B. adolescentis* B5789_0044 ([WP_080863062.1](#)); N. Fro_G(1,2)Afase, *Neocallimastix frontalis* NF2523 D-galacto-(α -1,2)-L-arabinosidase ([ASF57709.1](#)); N. Fro_NF2215, *N. frontalis* NF2215 ([ASF57708.1](#)); P. Rhi_2455, *Piromyces rhizinflatus* PR2455 ([ASF57710.1](#)); N. Fro_A(1,2)Afase, *N. frontalis* NF2152 α -L-(β -1,2)-L-arabinofuranobiosidase ([ASF57707.1](#)); A. Cav_B1,4Xyl, *Aeromonas caviae* β -1,4-xylosidase ([BAA95685.1](#)); C. Owe_B1,4Xyl, *Caldicellulosiruptor owensensis* β -1,4-xylosidase ([ADQ03734.1](#)); C. Sac_BXyl, *Caldicellulosiruptor saccharolyticus* β -xylosidase ([ABP67986.1](#)); T. Sac_BXyl, *Thermoanaerobacterium saccharolyticum* B6A-RI ([AAA27369.1](#)); B. Hal_BXyl, *Bacillus halodurans* β -xylosidase ([BAB04787.1](#)); G. sp_BXyl, *Geobacillus* sp. β -xylosidase ([WP_020755811.1](#)); G. Ste_BXyl, *Geobacillus stearothermophilus* β -xylosidase ([ABI49941.1](#)). Arrows to the right of the phylogenetic tree indicate cleavage sites of disaccharide-releasing enzymes.

TABLE 1 Substrate specificity of GAfase toward polysaccharides

Polysaccharide	Sp act (U/mg)	Relative activity (%) ^a
Gum arabic AGP	13.3	100
Larch AGP	1.35	10.2
Sugar beet arabinan	0.0263	0.198
Wheat arabinoxylan	Trace	Trace

^aRelative activity is expressed as the percentage of activity toward gum arabic AGP.

α -anomeric form. Moreover, the presence of the enhanced cross-peaks, GalpH1 (δ 5.02)/ArafC3 (δ 83.65), ArafH3 (δ 3.98)/GalpC1 (δ 82.92), and ArafH1 (δ 4.96)/ArafC4 (δ 82.92), in the heteronuclear multiple bond coherence (HMBC) spectra led us to unambiguously determine the linkage of the galactopyranoside unit to C-3 of the arabinofuranoside unit. Thus, the transglycosylation product was identified as α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe (Me, methyl), indicating that GAfase is a retaining glycoside hydrolase. Furthermore, the transglycosylation product derived from larch AGP under the 30% methanol environment was also identified as β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-OMe by using ^1H and ^{13}C NMR (Fig. S4 and Table S2).

Cleavage site of GAfase in gum arabic AGP. To identify the cleavage site of GAfase in gum arabic AGP, we performed ^1H NMR analysis of hydroxyproline (Hyp)-linked type II AG (Hyp-AG) prepared from gum arabic AGP (Fig. 7). The chemical shifts of the anomeric proton were assigned based on a comparison with a previous report (31). We also assigned [β -L-Arap-(1 $\rightarrow$)] at 5.56 ppm and [α -D-Galp-(1 $\rightarrow$)] at 5.59 ppm based on the transglycosylation products β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-OMe and α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe. The GAfase treatment of Hyp-AG resulted in the disappearance of two ^1H peaks, at 5.85 and 5.59 ppm, corresponding to [$\rightarrow$ 3]- α -L-Araf-(1 $\rightarrow$)] and [α -D-Galp-(1 $\rightarrow$)], respectively. These NMR analyses revealed that GAfase released α -D-Galp-(1 $\rightarrow$ 3)-Ara from the nonreducing termini of β -(1 $\rightarrow$ 6)-galactosyl side chains of the gum arabic AGP structure. Furthermore, the ^1H peak at 5.94 ppm was observed to be markedly decreased concomitant with an increase at 5.81 ppm. The chemical shift of the terminal α -L-Araf-(1 $\rightarrow$ 4)- was likely affected by the removal of α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-(1 $\rightarrow$ 3)- by GAfase. The shift in peaks was also observed upon the removal of the terminal α -D-Galp-(1 $\rightarrow$ 3)- structure of gum arabic AGP (31).

Substrate specificity of recombinant GAfase. To investigate GAfase reactivity toward oligosaccharides, we purified the oligosaccharides that featured the α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf- structure (S3-GA and S5-GA) and β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf- structure (S3-AA) (Fig. 1) from free oligosaccharides contained in gum arabic reagent. GAfase released α -D-Galp-(1 $\rightarrow$ 3)-L-Ara from S3-GA and S5-GA and released β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from S3-AA (Fig. 8). The peaks of S3-GA and S3-AA shifted to the same retention time as that for S3, which comprised α -L-Araf-(1 $\rightarrow$ 4)- β -D-Galp-(1 $\rightarrow$ 6)-Gal, and peaks for S5-GA shifted to S5, possessing the α -L-Rha-(1 $\rightarrow$ 4)-GlcA structure. S5-GA and S3-GA were substrates that were more susceptible for GAfase than S3-AA (Table 3). The substrate preference of GAfase for α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe relative to β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-OMe was also indicated by the results of kinetic analysis (Table 4). The k_{cat}/K_m value determined for α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe was 594-fold higher than that for β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-OMe. Moreover, GAfase exhibited no reactivity toward *p*-nitrophenyl (*p*NP)- α -Araf or any other *p*NP substrate tested (Fig. S5). These results suggested that GAfase was an *exo*-acting α -L-arabinofuranosidase that recognizes the α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-(1 $\rightarrow$ 3)- structure rather than the β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-(1 $\rightarrow$ 3)- structure. Based on the reactivity, we named the enzyme 3-*O*- α -D-galactosyl- α -L-arabinofuranosidase (GAfase).

Site-directed mutagenesis and deletion mutagenesis. GH39 α -L-(β -1,2)-arabinofuranobiosidase (NF2152), D-galacto-(α -1,2)-L-arabinosidase (NF2523), and GAfase have similar features that hydrolyze the α -L-arabinofuranoside linkage of disaccharides. Glu-155 and Glu-254 in NF2152 were predicted as the acid base catalytic residue and nucleophile, respectively (32). The catalytic and several other critical amino acid residues present in the active site of NF2152 were also conserved in GAfase (Fig. S6A). We evaluated

FIG 6 Transglycosylation activity of GAfase in the presence of 1-alkanols. (A) Thin-layer chromatography analysis of reaction products. GAfase was incubated with gum arabic AGP in the absence of 1-alkanols (lane 2) or in the presence of 30% methanol (lane 3), ethanol (lane 4), or 1-propanol (lane 5) at 37°C for 16 h. Lane 1, α -D-Galp-(1 $\rightarrow$ 3)-L-Ara standard. (B) Purified α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe was incubated with (+) or without (-) GAfase at 37°C for 16 h. α -D-Galp-(1 $\rightarrow$ 3)-L-Ara was used as the standard. (C) Chemical structure of α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe. (D) ^1H and ^{13}C NMR spectra of α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe in D_2O at RT. The chemical shifts are listed in Table 2. Abbreviations: GA, α -D-Galp-(1 $\rightarrow$ 3)-L-Ara; Me, methyl; Et, ethyl; Pr, propyl.

the importance of the corresponding residues of GAfase by using site-directed mutagenesis to construct Gln substitutions of GAfase at Glu-194 and Glu-321. The E194Q and E321Q mutant enzymes exhibited no enzymatic activity toward gum arabic AGP (Fig. S6B), suggesting that these amino acid residues were catalytic residues in GAfase. We also constructed the truncated derivatives lacking the C-terminal domain, C Δ 355 (positions 37 to 781, 745 aa), C Δ 477 (positions 37 to 659, 623 aa), and C Δ 630 (positions 37 to 506, 470 aa) to evaluate the importance of the CBM domains. These deletion mutants were expressed

TABLE 2 Assignment of signal in ^1H and ^{13}C NMR spectra of $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-L-Araf-OMe}$ in D_2O at 400 MHz at RT^a

Unit and signal		Value for:					
Galp		C-1	C-2	C-3	C-4	C-5	C-6a C-6b
^1H		5.02	3.82	4.03	4.01	4.04	3.77 3.77
^3J		d 4.1	dd 10.5, 4.1	dd 10.5, 2.6	d 2.6	t 6.4	d 6.4 d 6.4
^{13}C		99.28	68.23	69.18	69.25	71.46	61.19
^1J		175					
Araf		C-1	C-2	C-3	C-4	C-5a C-5b	OMe
^1H		4.96	4.28	3.98	4.22	3.82 3.76	3.44
^3J		s	dd 2.8, 1.4	dd 6.0, 2.8	dt 5.6, 3.2	dd 12.8, 3.2 dd 12.8, 5.6	
^{13}C		108.67 ^e	79.52	83.65 ^b	82.92 ^{c,d}	61.35	55.0 ^d
^1J		176					

^a ^1H , 400 MHz; ^{13}C , 100 MHz; Me, methyl; s, singlet; d, doublet; t, triplet; dd, doublet of doublets; dt, doublet of triplets.

^bHMBC cross peak of enhancement at 83.65 (C^{Araf3}) by irradiation at 5.02 ($\text{C}^{\text{Galp1-H}}$) was observed.

^cHMBC cross peak of enhancement at 82.92 (C^{Galp1}) by irradiation at 3.98 ($\text{C}^{\text{Araf3-H}}$) was observed.

^dHMBC cross peaks of enhancement at 55.00 (MeO) and 82.92 (C^{Araf4}) by irradiation at 4.96 ($\text{C}^{\text{Araf1-H}}$) were observed.

^eHMBC cross peak of enhancement at 108.67 (C^{Araf1}) by irradiation at 3.44 (MeO) was observed.

as soluble proteins but exhibited no enzymatic activity toward gum arabic AGP and $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-L-Araf-OMe}$ (data not shown).

In vitro assimilation test conducted using $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-L-Ara-free}$ gum arabic AGP and combination reactions of type II AG degradative enzymes. *In vitro* assimilation tests were performed on *B. longum* JCM1217 and JCM7052 by using 1.0% $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-L-Ara-free}$ gum arabic AGP and intact gum arabic AGP. Interestingly, JCM1217, which showed no ability to assimilate intact gum arabic AGP, exhibited a growth rate equivalent to that of JCM7052 in $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-L-Ara-free}$ gum arabic AGP (Fig. 9A). To further investigate the effect of GAfase acting together with type II AG degradative enzymes, $\alpha\text{-D-Galp-(1}\rightarrow\text{3)-L-Ara-free}$ gum arabic AGP was allowed to react with JCM1217 and JCM7052 cells grown in modified de Man-Rogosa-Sharpe (MRS) medium containing larch AGP (Fig. 9B). The released oligosaccharides corresponding to peaks of S3 and S5, the degradation products of S3GA and S5GA, respectively, were detected in the reaction products of the JCM1217 cell. To further elucidate the combination effect of GAfase with type II AG degradative enzymes (BI1,3Gal, BI1,6Gal, and BIArafA), gum arabic AGP was incubated with the three type II AG degradative enzymes in the presence or absence of GAfase (Fig. 9C). Notably, no distinctive peaks of oligosaccharides were detected in the absence of GAfase, whereas the peaks of S3 and S5 were detected in the combined presence of GAfase and the three type II AG degradative enzymes. Based on the substrate specificities of the enzymes, these oligosac-

FIG 7 ^1H NMR spectra of Hyp-AG treated without (A) or with (B) GAfase. Linkages shown are in reference to corresponding chemical shifts assigned previously (31). NA, nonassigned peak.

FIG 8 HPAEC-PAD analysis of GAfase reaction with oligosaccharides derived from gum arabic AGP. S3-GA, S5-GA, and S3-AA were incubated with (+) or without (-) GAfase at 37°C for 16 h.

charides were likely produced by the combined reaction of Bl1,3Gal and GAfase. This finding suggested that the removal of terminal α -D-Galp-(1 $\rightarrow$ 3)-L-Ara of gum arabic AGP by GAfase promoted the removal of the steric hindrance of Bl1,3Gal.

In vitro assimilation test of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara. To evaluate the assimilation ability of the reaction products of GAfase, α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara, *B. longum* JCM1217 and JCM7052 were cultivated for 48 h in MRS medium containing 0.2% α -D-Galp-(1 $\rightarrow$ 3)-L-Ara or β -L-Arap-(1 $\rightarrow$ 3)-L-Ara as the sole carbon source. *B. longum* JCM7052 exhibited moderate growth on α -D-Galp-(1 $\rightarrow$ 3)-L-Ara after 48 h but not on β -L-Arap-(1 $\rightarrow$ 3)-L-Ara (Fig. S7A). In contrast, JCM1217 exhibited no remarkable growth on either α -D-Galp-(1 $\rightarrow$ 3)-L-Ara or β -L-Arap-(1 $\rightarrow$ 3)-L-Ara. The analysis of the culture supernatant further supported these results (Fig. S7B). The peak of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara was diminished only in JCM7052, not in JCM1217, whereas the peak of β -L-Arap-(1 $\rightarrow$ 3)-L-Ara was not decreased in either case.

DISCUSSION

In this study, we characterized a newly identified GAfase in *B. longum* that releases α -D-Galp-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP. The enzyme was classified as a retaining glycoside hydrolase belonging to the GH39 family. However, the bifidobacterial GH39 enzymes, including the newly identified GAfase, exhibited a low level of amino acid sequence identity (<25%) to the other characterized GH39 β -xylosidases and rumen fungal enzymes and were located in the distant clade among GH39 members. This finding suggested that the proteins should comprise a new subfamily of GH39. Hydrolysis by retaining enzymes proceeds through a double-displacement mechanism involving two catalytic residues, with either glutamate or aspartate typically being used as a catalytic amino acid residue (33). Gln substitutions of GAfase Glu-194 and Glu-321, corresponding to the catalytic amino acid residues in GH39 D-galacto-(α -1,2)-L-arabinosidase, resulted in a loss of enzymatic activity. Hence, these residues were considered the catalytic residues in GAfase. The α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-releasing activity of GAfase was 594-fold higher than the β -L-Arap-(1 $\rightarrow$ 3)-L-Ara-releasing activity, and GAfase showed no activity toward any tested pNP substrate, including pNP- α -Araf. These results suggested that GAfase was an exo-acting enzyme that strictly recognizes

TABLE 3 Substrate specificity of GAfase toward oligosaccharides

Oligosaccharide	Sp act (U/mg)	Relative activity (%) ^c
S3-GA ^a	2.85	61.9
S5-GA ^a	4.61	100
S3-AA ^b	0.0197	0.426

^aThe substrate concentration was 5 μ M.

^bThe substrate concentration was 10 μ M.

^cRelative activity is expressed as the percentage of activity toward S5-GA.

TABLE 4 Kinetic parameters of GAFase on transglycosylation products

Product	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)
α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe	0.633 ± 0.01	199 ± 3.60	315
β -L-Arap-(1 $\rightarrow$ 3)- α -L-Araf-OMe	22.8 ± 1.39	12.1 ± 0.349	0.530

α -L-Araf in the -1 subsite of GAFase and that the presence of α -D-Galp or β -L-Arap corresponding with the -2 subsite may be essential for the enzymatic activity. β -L-Arap shares structural similarity with α -D-Galp in the orientation of the hydroxyl groups at C-1, C-2, C-3, and C-4 but does not contain a C-5 hydroxyl methyl group. Because of this similarity, β -L-arabinopyranosidase and α -D-galactosidase frequently exhibit cross-reactivity with α -D-Galp and β -L-Arap (25, 34, 35). Therefore, GAFase predominantly releases α -D-Galp-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP and secondarily releases β -L-Arap-(1 $\rightarrow$ 3)-L-Ara. We also compared the amino acid residues corresponding to the -1 and -2 subsites of four GH39 enzymes characterized as α -L-(β -1,2)-arabinofuranobiosidase or D-galacto-(α -1,2)-L-arabinosidase with the residues in GAFase (see Fig. S6A in the supplemental material) (32). The amino acid residues present in the -2 subsite of GAFase differed from those in D-galacto-(α -1,2)-L-arabinosidase, an α -D-Galp-(1 $\rightarrow$ 2)-L-Ara-releasing enzyme. Furthermore, the amino acid residue was also different from that in a GAFase homolog from *B. pseudocatenulatum*. Thus, the amino acid residue in the -2 subsite of GAFase may be optimized for α -1,3-linked Galp. However, further investigation based on crystal structure analysis is required to determine the substrate recognition mechanism of GAFase.

Considering the constitution of genes surrounding the GAFase gene, we proposed the following model for the assimilation of gum arabic AGP in *B. longum* JCM7052 (Fig. 10). First, GAFase (BLGA_00340) releases α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from gum arabic AGP on the surface of the bacterial cell. The released α -D-Galp-(1 $\rightarrow$ 3)-L-Ara is then internalized into the cell through ABC transport systems (BLGA_00300, BLGA_00310, BLGA_00320, and BLGA_00350), after which intracellular GH36 α -D-galactosidase (BLGA_00330) hydrolyzes the α -D-Galp-(1 $\rightarrow$ 3)-L-Ara into galactose and L-arabinose. These monosaccharides are finally catabolized through a bifid shunt, and acetic acid and lactic acid are produced (36). We also confirmed that *B. longum* JCM7052 used α -D-Galp-(1 $\rightarrow$ 3)-L-Ara but not β -L-Arap-(1 $\rightarrow$ 3)-L-Ara as a carbon source (Fig. S7), suggesting that α -D-Galp-(1 $\rightarrow$ 3)-L-Ara was degraded by GH36 α -D-galactosidase (BLGA_00330). In contrast, we have no results directly indicating that ABC transport systems (BLGA_00300, BLGA_00310, BLGA_00320, and BLGA_00350) transport α -D-Galp-(1 $\rightarrow$ 3)-L-Ara. Future studies focusing on transcriptional analysis of genes in this cluster and functional analysis and targeted gene mutagenesis of the solute-binding protein candidate (BLGL_00300) are needed.

GAFase contains two CBM35 domains and one galactose-binding domain-like superfamily domain in the C-terminal region. To date, CBM35 and CBM13 have been shown to bind to β 1,3-galactan main chains lacking side chains but not to intact gum arabic AGP (37–39). For instance, CBM35 in exo- β -1,3-galactanase (previously classified as CBM6) from *Phanerochaete chrysosporium* (GenBank accession no. [BAD98241.1](#)) was reported to be critical for recognizing two continuous β 1,3-linked galactopyranosyl residues (37). The third domain from the N terminus (amino acid residues 791 to 966) displays 34.1% amino acid sequence identity at 35% coverage to the closest functionally characterized CBM35 domain from *Phanerochaete chrysosporium*. The first CBM35 domain (amino acid residues 512 to 657) displays 28.2% amino acid identity at 41% coverage to the CBM35 domain from *Paenibacillus* sp. (GenBank accession no. [ACS99969.1](#)). The second CBM35 domain (amino acid residues 665 to 790) displays 21.5% amino acid sequence identity at 41% coverage to the CBM35 domain from *Aspergillus nidulans* FGSC A4 (GenBank accession no. [ABF50881.1](#)). Although we expressed C-terminal deletion mutants of GAFase in the soluble form, the mutation completely abolished enzymatic activity toward gum arabic AGP and α -D-Galp-(1 $\rightarrow$ 3)- α -L-Araf-OMe (data not shown). Thus,

FIG 9 *In vitro* assimilation test performed using α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP. (A) Growth profiles of *B. longum* JCM1217 and JCM7052 on gum arabic AGP (left) and α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP (right). Error bars indicate standard deviations ($n=3$). (B) Intact gum arabic AGP and α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP were incubated with bacterial cell fractions of *B. longum* JCM1217 or JCM7052 grown in larch AGP. Reaction products were analyzed by performing HPAEC-PAD with a PA-1 column. (C) HPAEC-PAD analysis of the combination reactions of Gafase and type II AG degradative enzymes (B1,3Gal, B1,6Gal, and B1AraF) on gum arabic AGP. Gum arabic AGP was incubated with type II AG degradative enzymes, Gafase, or a mixture of type II AG degradative enzymes and Gafase. S3GA and S5GA mixture and the Gafase reaction products were used as standards.

the C-terminal CBM domains are likely essential for catalytic activity of Gafase. Based on our previous reports, the catalytic module of modular enzymes can function independently (40, 41). On the other hand, enzymatic activity of truncated derivatives lacking the CBM domain was lost in an arabinoxylan-specific xylanase, CtXyl5A, from *Clostridium thermocellum*, which is composed of catalytic module CtGH5 and CtCBM6, and β -(1,4)-mannanase PaMan26A from *Podospora anserina*, which is composed of GH26 and CBM35 (42, 43). The linker region of these enzymes interacts with the catalytic module and the CBM. Similarly, the catalytic module of Gafase may also be closely associated with CBM.

We have previously characterized the type II AG degradative enzymes in *B. longum*

FIG 9 (Continued)

JCM1217 that hydrolyzed β 1,3-galactan main chains and β 1,6-galactan side chains of larch AGP (29, 40). However, these enzymes were hardly able to act on intact gum arabic AGP because its structure is more complex than that of other plant-derived AGPs (40). In this study, we showed that *B. longum* JCM1217 exhibited a growth rate equivalent to that of JCM7052 in medium containing α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP (Fig. 9A). Furthermore, we reported that type II AG degradative enzymes in JCM1217 acted on α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP (Fig. 9B). JCM7052 also conserved type II AG degradative enzymes, such as BI1,3Gal (BLGA_18370), BI1,6Gal (BLGA_18380), BIARafA (BLGA_18470), BIARafB (BLGA_18460), and BIARafE (BLGA_18450) (Fig. 10). These results suggested that the presence of the GAfase gene in *B. longum* is a critical factor for growth on gum arabic AGP. GAfase activity releases α -D-Galp-(1 $\rightarrow$ 3)-L-Ara as a favorable carbon source for gum arabic AGP assimilative strains and also markedly influences exo- β 1,3-galactanase reactivity in gum arabic AGP nonassimilative strains by facilitating access to the β 1,3-galactan main chain of gum arabic AGP. In contrast, to the best of our knowledge, α -L-rhamnosidase and β -glucuronidase candidates were not conserved in *B. longum* strains. Therefore, oligosaccharides containing L-rhamnose and glucuronic acid may be utilized by other intestinal bacteria.

To date, a few published studies have investigated the link between the presence of carbohydrate-active enzymes and bacterial growth on gum arabic AGP. For example, an extracellular endo- β 1,3-galactanase in *Bacteroides cellulosilyticus* and *Bacteroides caccae* was reported to play a critical role in bacterial proliferation in medium containing gum arabic AGP as the sole carbon source (44). In this study, we showed that GAfase plays a crucial role in gum arabic AGP assimilation in *B. longum* by using an *in vitro* assimilation test and performing comparative genome analyses between gum arabic AGP assimilative strains and nonassimilative strains. This gene-trait matching approach was applied to identify a key gene cluster involved in carbohydrate assimilation. The agreement between the phenotype and genotype in this study may enable us to determine how specific bifidobacteria in the human intestine increase by intake of gum arabic AGP. The symbiotic relationship of GAfase with other bacteria that have degrading enzymes and other factors needs to be evaluated in the future.

Several enzymes can be used for AGP structural analysis, such as exo- β 1,3-

FIG 10 Schematic model of gum arabic AGP metabolic pathway in *B. longum* JCM7052.

galactanase, endo- β -1,3-galactanase, endo- β -1,6-galactanase, α -L-rhamnosidase, α -L-arabinopyranosidase, β -L-arabinopyranosidase, β -L-arabinofuranosidase, 4-O-methyl- α -glucuronidase, and α -galactosidase (31). Alongside these enzymes, GAfase can serve as a promising tool for the structural analysis of gum arabic AGP, and GAfase may help to elucidate how AGPs are fermented by human gut microbes.

GAfase treatment revealed that the gum arabic AGP content of α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara was 9.36% and 1.87% (wt/wt), respectively. The results of a methylation analysis in a prior study showed that gum arabic AGP contained 12 mol % of terminal Gal, 1 mol% of terminal Arap, and 16 mol% of $\rightarrow$ 3-L-Araf 1 $\rightarrow$ (45). The α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara units are considered the only structures which have $\rightarrow$ 3-L-Araf1 $\rightarrow$ in gum arabic AGP. Therefore, these disaccharide units are considered to be contained at $\sim$ 16 mol% with α -D-Galp-(1 $\rightarrow$ 3)-L-Ara and β -L-Arap-(1 $\rightarrow$ 3)-L-Ara present in a roughly 12:1 molar ratio. Furthermore, the results of prior

structural analyses of gum arabic AGP have also supported the experimental values reported in this study (46, 47). To the best of our knowledge, the α -D-Galp-(1 $\rightarrow$ 3)-L-Ara structure has been identified only in gum arabic AGP, whereas β -L-Arap-(1 $\rightarrow$ 3)-L-Ara has been detected in larch AGP (48) and wheat AGP (49). In this study, we found that GAfase exhibited weak activity to release β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from sugar beet arabinan, which has not been reported to contain this disaccharide unit (32, 50). It is likely that these commercially available plant polysaccharides contain only a few AGPs as contamination, because AGPs are covalently attached to pectin and xylan in the plant cell wall (51). In naturally occurring AGPs, the β -L-Arap-(1 $\rightarrow$ 3)-L-Ara structure is more likely to be present than the α -D-Galp-(1 $\rightarrow$ 3)-L-Ara structure. *B. pseudocatenulatum* harbors the highest percentage of GAfase homologous genes (23.4%) among bifidobacterial species. The GAfase homolog in *B. pseudocatenulatum* holds the potential for 3-O- β -L-arabinopyranosyl- α -L-arabinofuranosidase activity, which releases β -L-Arap-(1 $\rightarrow$ 3)-L-Ara from AGPs, since the candidate gene encoding β -L-arabinopyranosidase was flanked by the GAfase homolog. To provide a comprehensive understanding of the AGP-degrading mechanisms, characterization of the bifidobacterial GAfase homologs, structural analysis of AGPs, and AGP fermentation experiments using a combination of bifidobacteria are required.

In vitro fermentation analyses performed using human intestinal bacteria revealed that gum arabic AGP was fermented in *B. longum* and *B. adolescentis* but not in *B. pseudocatenulatum* (22). Furthermore, gum arabic AGP increased bifidobacteria in human subjects (21) and human fecal batch culture (52) and selectively increased *B. longum* in a human gastrointestinal simulator (19). Although intake of gum arabic AGP was reported to increase bifidobacteria, GAfase carrier strains accounted for only 6.84% of the *B. longum* strains ($n = 307$) based on the NCBI database, as described above. Since α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP was assimilated by a gum arabic AGP non-assimilative *B. longum* strain, similar cooperative action may be achieved through gum arabic AGP assimilative bacteria in the gut microbiota, such as *Bacteroides caccae* and *Bacteroides cellulosilyticus*. GAfase could thus serve as a promising tool for discriminating between responder and nonresponder subjects for the intake of gum arabic AGP and help reveal the basis for the individual differences in the effects of prebiotics.

MATERIALS AND METHODS

Materials. Gum arabic from *Acacia senegal* and nearly all pNP substrates were obtained from Sigma-Aldrich (St. Louis, MO, USA). pNP- β -Araf was synthesized as described previously (53). Larch AGP was purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). Sugar beet arabinan and wheat arabinoxylan were purchased from Megazyme (Wicklow, Ireland). These polysaccharides were used after purification through ethanol precipitation. The oligosaccharides S3-GA, S5-GA, and S3-AA (Fig. 1) were prepared from the supernatant of the ethanol precipitation of gum arabic reagent, as described by Tischer et al. (54). Briefly, the oligosaccharides were collected and evaporated to dryness. The dried samples were dissolved in water and separated via gel filtration chromatography on a Bio-Gel P-2 column (dimensions $[\varphi]$, 25 by 830 mm; Bio-Rad Laboratories, Hercules, CA, USA) equilibrated with water. Subsequently, the oligosaccharides were purified using a combination of an activated carbon column (Autoprep FiberAC; Showa Denko, Tokyo, Japan), high-pressure liquid chromatography on a Cosmosil Sugar-D column (φ , 4.6 by 250 mm; Nacalai Tesque Inc., Kyoto, Japan), and a Cosmosil PBr column (φ , 4.6 by 250 mm; Nacalai Tesque Inc.). Hyp-AG of gum arabic AGP was prepared by means of an alkali treatment performed using barium hydroxide, as described previously (50, 55). The dried sample was dissolved in water and separated by gel filtration chromatography on a Sephadex G50 column equilibrated with 100 mM ammonium acetate buffer (pH 6.8). Partially dearabinosylated gum arabic AGP was prepared using mild acid degradation, as described by Ling et al. (56). α -D-Galp-(1 $\rightarrow$ 3)-L-Ara-free gum arabic AGP was prepared using the membrane-enclosed enzymatic catalysis (MEEC) approach (57) to avoid inhibition of GAfase. Gum arabic AGP (1%) was incubated with GAfase (0.120 mg/ml) at 37°C in a dialysis membrane (size 36; Wako Pure Chemical Industries, Ltd., Osaka, Japan) and dialyzed against 50 mM sodium acetate buffer (pH 5.5). The polysaccharide was then precipitated with ethanol. All other chemicals used in the study were obtained from Wako Pure Chemical Industries, Ltd.

***In vitro* assimilation test of AGPs and assays of bacterial enzyme activity.** The following 10 strains of *B. longum* used in *in vitro* assimilation tests were obtained from stock cultures maintained in the MCC (Morinaga Milk Industry Co., Ltd., Zama, Japan): MCC00055, MCC00198, MCC00215, MCC00231, MCC00256, MCC00300, MCC00316, MCC00325, MCC00328, and MCC00353. *B. longum* JCM7052 and JCM1217 were obtained from the Japan Collection of Microorganisms (RIKEN Bioresource Centre, Ibaraki, Japan). The strains were precultured at 37°C under anaerobic conditions, using an AnaeroPack system (Mitsubishi Gas Chemical, Tokyo, Japan), in MRS broth containing 0.05% L-cysteine hydrochloride. The precultured cells were inoculated into MRS medium containing 1.0% gum arabic AGP or larch AGP as the sole carbon source, and the cells

were then cultured for 48 h at 37°C under anaerobic conditions. Each culture medium was evaluated by measuring the absorbance at 600 nm. To assay bacterial enzyme activity, cell pellets were washed with 5 mM sodium acetate buffer (pH 5.0) and stored at –20°C. Subsequently, the freeze-thawed pellets were resuspended in 2.0 ml of 50 mM sodium acetate (pH 5.0) containing 1.0% gum arabic AGP and incubated at 37°C for 3 days. The assay supernatants were then analyzed by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using a CarboPac PA-1 column (φ , 4 by 250 mm; Dionex Corp., Sunnyvale, CA, USA), which was eluted at a flow rate of 1.0 ml/min by using the following gradient: 0 to 5 min, 100% eluent A (0.1 M NaOH); 5 to 30 min, 0% to 100% eluent B (0.5 M sodium acetate in 0.1 M NaOH); and 30 to 35 min, 100% eluent B.

Analysis of peak X. Acid hydrolysis was performed by incubating the purified peak X with 2 M trifluoroacetic acid at 120°C for 1 h, after which the reaction mixture was allowed to cool to room temperature (RT) and then dried using a SpeedVac. The sample was twice dried and washed with water, dissolved using an appropriate amount of water for HPAEC-PAD analysis, and applied onto a CarboPac PA-10 column (φ , 4 by 250 mm; Dionex Corp., Sunnyvale, USA), which was eluted at a flow rate of 1.0 ml/min by isocratic elution (18 mM NaOH). For enzymatic hydrolysis, GH27 α -D-galactopyranosidase/ β -L-arabinopyranosidase BAD_1525 from *B. adolescentis* was prepared as described previously (25). The purified peak X and α -D-Galp-(1 $\rightarrow$ 3)-L-Ara isolated using partial acid degradation were incubated with BAD_1525 in 40 μ l of 50 mM sodium acetate buffer (pH 5.0) for 16 h at 37°C, after which the reaction products were analyzed by HPAEC-PAD by using the PA-1 column as previously mentioned.

Genome sequencing and bioinformatics analysis. Genomic DNA of *B. longum* strains was extracted using the DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA), and libraries were constructed using a Nextera XT DNA sample preparation kit (Illumina, Inc., San Diego, CA, USA) according to the manufacturer's protocol. Paired-end sequencing was performed using an Illumina MiSeq instrument and the MiSeq v3 reagent kit (Illumina, Inc.), which achieved at least 100-fold coverage. Quality trimming and *de novo* assembly of the raw paired-end reads were performed by using the CLC Genomics Workbench software package (v8.0; Qiagen) with default settings, except for contig length (minimum contig length, 500 bp). Open reading frames were predicted and annotated using the DDBJ Fast Annotation and Submission Tool with default settings (58).

Expression and purification of recombinant Gafase. The genomic DNA of *B. longum* JCM7052 (GenBank accession no. [AP022379](#)) was used for PCR amplification of the BLGA_00340 (Gafase) gene. The forward (5'-AGGAGATATACCATGGAGGAGATGCCGAG-3') and reverse (5'-GTGGTGTGCTCGAGACCGCTACGGAGTCGATC-3') primers were designed from nucleotides 109 to 124 and 2904 to 2922, respectively. The underlined letters represent the nucleotides complementary to the template. The amplicon was cloned into a pET-23d vector (Novagen, Madison, WI, USA) by using an In-Fusion HD cloning kit (Clontech Laboratories, Inc., Palo Alto, CA, USA). The resulting plasmid was used to transform *Escherichia coli* BL21 (λ DE3) cells (Genlantis, San Diego, CA, USA), which were grown at 25°C by using the Overnight Express autoinduction system (Novagen). The cell culture was centrifuged, the resultant pellet was resuspended in the BugBuster protein-extraction reagent (Novagen), and the C-terminal His-tagged Gafase protein was purified using a column containing Talon metal-affinity resin (Clontech). The 5 and 10 mM imidazole fractions containing Gafase were desalted and concentrated using an ultrafiltration membrane (10-kDa-molecular-weight cutoff [MWCO]; Millipore Co., Billerica, MA, USA).

Enzyme assays. The reactivity of the recombinant enzyme was analyzed using the following substrates: polysaccharide substrates (gum arabic AGP, larch AGP, sugar beet arabinan, and wheat arabinoxylan), purified oligosaccharides from gum arabic AGP (S3-GA, S5-GA, and S3-AA), and pNP substrates (pNP- α -L-Araf, pNP- α -L-Arap, pNP- α -Galp, pNP- α -Xylp, pNP- α -Glc, pNP- β -Araf, pNP- β -Arap, pNP- β -Galp, pNP- β -Xylp, pNP- β -Glc, and pNP- β -GlcA). The hydrolytic activity toward each substrate was analyzed as follows: the polysaccharides (final concentration, 0.1%) were incubated with 47.3 μ g/ml Gafase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5), the purified oligosaccharides (final concentration, 5 to 10 μ M) were incubated with 0.421 μ g/ml Gafase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5), and the pNP substrates (final concentration, 5 mM) were incubated with 1.47 μ g/ml Gafase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5). After incubation at 37°C for 16 h, the reaction mixtures were analyzed using HPAEC-PAD as mentioned above or by thin-layer chromatography (TLC). In the TLC analysis, reaction products were spotted on Silica Gel 60 aluminum plates (Merck, Darmstadt, Germany) with a 7:1:2 (vol/vol/vol) 1-propanol-ethanol-water solvent mixture, and the separated sugars were visualized by spraying an orcinol-sulfate reagent on the plates (59).

For NMR analysis of the Hyp-AG sugar chains of gum arabic AGP, Hyp-AG was incubated in the presence or absence of 1.18 μ g/ml Gafase at 37°C for 72 h in 1 ml of 50 mM sodium acetate buffer (pH 5.5) in a dialysis membrane tube (6- to 8-kDa MWCO), according to the membrane-enclosed enzymatic catalysis approach. The reaction was terminated by boiling for 3 min, and the Gafase-treated Hyp-AG was purified using 80% ethanol precipitation. ^1H NMR spectra were recorded, in D_2O at 80°C, at 400 MHz on a JEOL ECX-400 spectrometer, and chemical shifts were referenced to the internal semiheavy water (HOD) signal (4.80 ppm).

Optimal pH and temperature. The optimal pH of enzyme activity was determined using gum arabic AGP as the substrate between pH 3.5 and 7.0 by using 50 mM sodium acetate buffer (pH 3.5 to 6.0) and 50 mM MES buffer (pH 5.5 to 7.0). The optimal temperature of enzyme activity was determined using 50 mM sodium acetate buffer (pH 5.5) at 30 to 60°C. The reactions were stopped by adding 5% trichloroacetic acid (TCA), and the samples were analyzed using HPAEC-PAD.

Transglycosylation activity. Transglycosylation reactions were performed using gum arabic AGP as the donor and 1-alkanols as the acceptors; 1.0% gum arabic AGP was incubated at 40°C for 2 h with 0.295 μ g/ml Gafase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5) in the presence of 30% methanol, ethanol, or 1-propanol as the acceptor. Subsequently, the reaction products were analyzed using TLC.

TABLE 5 Primers for site-directed and deletion mutagenesis

Primer name	Oligonucleotide sequence ^a
GAfase_E194Qmut_for	5'-AGCCGGATGGCGGCAACTGGTACGCT-3'
GAfase_E194Qmut_rev	5'-GGTTGAACGGCATGAATACAATATTGTC-3'
GAfase_E321Qmut_for	5'-C ^u CAATTCGGCGAACTGCGTGACATGTC-3'
GAfase_E321Qmut_rev	5'-GAGATGTTGATGGCACGCGGCGTGATG-3'
CA_Foward_Primer	5'-CTCGAGCACCACCACCACTG-3'
CA355rev (37–781)	5'-CGCACGCGCGGTGGTGACACTCGCCA-3'
CA477rev (37–659)	5'-GGAGACCTGGTAGAGCAGCAGTTGTC-3'
CA630rev (37–506)	5'-ATCCTGATGAGGCGTCACCACAAGCTG-3'

^aThe positions of the mutated sequences are underlined.

For NMR analysis, the transglycosylation product α -D-Galp-(1→3)- α -L-Araf-OMe was synthesized by incubating 1% gum arabic AGP at 40°C for 16 h with 0.295 μ g/ml GAfase in 40 ml of 50 mM sodium acetate buffer (pH 5.5) with 40% methanol as the acceptor. The supernatant after ethanol precipitation was purified using an activated carbon column. Similarly, β -L-Arap-(1→3)- α -L-Araf-OMe was prepared by using the same procedure with larch AGP as the donor substrate. ¹H and ¹³C NMR spectra were measured, in D₂O at RT, at 400 MHz on a JEOL ECX-400 spectrometer.

Substrate specificity. The substrate specificity toward polysaccharides was assayed by incubating the substrates (1.0%) with 1.47 μ g/ml GAfase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5) at 40°C for an appropriate reaction time. For analyzing substrate specificity toward oligosaccharides, 5 μ M S3-GA, 5 μ M S5-GA, and 10 μ M S3-AA were incubated with 0.295 $\times 10^{-1}$ μ g/ml GAfase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5) at 45°C for an appropriate reaction time. The reactions were terminated by adding 10 μ l of 5% TCA, and the products were analyzed using HPAEC-PAD, as described in the preceding paragraphs. To examine substrate specificity toward transglycosylation products, 0.1 mM α -D-Galp-(1→3)- α -L-Araf-OMe and β -L-Arap-(1→3)- α -L-Araf-OMe were incubated with 0.295 μ g/ml GAfase in 40 μ l of 50 mM sodium acetate buffer (pH 5.5) at 45°C for an appropriate reaction time, and after the reactions were stopped by adding 10 μ l of 5% TCA, the products were analyzed using HPAEC-PAD. One unit of enzyme activity was defined as the amount of enzyme required to produce 1 μ mol of α -D-Galp-(1→3)-L-Ara or β -L-Arap-(1→3)-L-Ara per min.

Kinetic analysis. GAfase kinetic parameters were determined by using 0.08125 to 5.0 mM α -D-Galp-(1→3)- α -L-Araf-OMe and 5.0 to 40 mM β -L-Arap-(1→3)-L-Araf-OMe as substrates after purification, as described above. In the case of α -D-Galp-(1→3)- α -L-Araf-OMe, an 80- μ l reaction mixture containing 50 mM sodium acetate buffer (pH 4.5) and 0.059 μ g/ml GAfase was incubated at 45°C for 20 min, whereas for β -L-Arap-(1→3)-L-Araf-OMe, 11.8 μ g/ml GAfase was used and the same procedure was followed. The reactions were terminated by adding 20 μ l of 5% TCA, and the products were analyzed using HPAEC-PAD. The concentrations of the reaction products were calculated based on the peak areas.

Site-directed mutagenesis and deletion mutagenesis. A KOD-Plus mutagenesis kit (Toyobo Co., Ltd., Osaka, Japan) was used to introduce amino acid substitutions or deletions into GAfase by using specific primers (Table 5). After confirmation of the mutations through DNA sequencing, the mutant enzymes were expressed and purified using the same procedure as that used for the wild-type enzyme.

In vitro assimilation test of the reaction products of GAfase and assays of bacterial enzyme activity. Precultured *B. longum* JCM1217 and JCM7052 cells were inoculated into MRS medium containing 1.0% α -D-Galp-(1→3)-L-Ara-free gum arabic AGP, intact gum arabic AGP, 0.2% α -D-Galp-(1→3)-L-Ara, or β -L-Arap-(1→3)-L-Ara as the sole carbon source, with the culture conditions being the same as in the assimilation test of AGPs described above. The 600-nm absorbance of α -D-Galp-(1→3)-L-Ara-free gum arabic AGP and intact gum arabic AGP was monitored every ~3 h for 48 h and for α -D-Galp-(1→3)-L-Ara and β -L-Arap-(1→3)-L-Ara after 48 h. After 48 h of cultivation, the culture media were centrifuged and the supernatants were analyzed by HPAEC-PAD with a PA-1 column.

The freeze-thawed cell pellets of *B. longum* JCM1217 and JCM7052 grown on larch AGP were incubated with 1.0% gum arabic AGP and α -D-Galp-(1→3)-L-Ara-free gum arabic AGP in 50 mM sodium acetate buffer at 37°C for 16 h, and the reaction products were analyzed using HPAEC-PAD.

Combination reactions of GAfase and type II AG degradative enzymes. To assess combination effects of GAfase and type II AG degradative enzymes, 0.05% gum arabic AGP was incubated with BI1,3Gal, BI1,6Gal, and BIArafA in 100 μ l of 50 mM sodium acetate buffer (pH 5.0) at 37°C for 16 h in the presence or absence of GAfase. The reactions were stopped by boiling for 3 min, and products were analyzed by performing HPAEC-PAD with a PA-1 column.

Data availability. The data have been deposited with links to accession numbers [PRJDB9197](#) (GenBank accession number [AP022379.1](#)) and [PRJDB10586](#) (GenBank accession numbers [BNGW000000000.1](#), [BNGX000000000.1](#), [BNGY000000000.1](#), [BNGZ000000000.1](#), [BNHA000000000.1](#), [BNHB000000000.1](#), [BNHC000000000.1](#), [BNHD000000000.1](#), [BNHE000000000.1](#), and [BNHF000000000.1](#)) in the DDBJ BioProject database.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 2 MB.

ACKNOWLEDGMENTS

We thank Shinya Fushinobu for helpful discussions.

This work was supported in part by JSPS KAKENHI grants-in-aid, grant numbers 19J20806 (to Y.S.), 19K05816 (to K.F.), 18K05345 (to A.I.), and 16H06290 (to Y.I. and A.I.).

REFERENCES

- Mitsuoka T. 1990. Bifidobacteria and their role in human health. *J Ind Microbiol* 6:263–267. <https://doi.org/10.1007/BF01575871>.
- Mitsuoka T. 1984. Taxonomy and ecology of bifidobacteria. *Bifidobacteria Microflora* 3:11–28. https://doi.org/10.12938/bifidus1982.3.1_11.
- Kato K, Odamaki T, Mitsuyama E, Sugahara H, Xiao JZ, Osawa R. 2017. Age-related changes in the composition of gut *Bifidobacterium* species. *Curr Microbiol* 74:987–995. <https://doi.org/10.1007/s00284-017-1272-4>.
- Turroni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, Kerr C, Hourihane J, Murray D, Fuligni F, Gueimonde M, Margolles A, De Bellis G, O'Toole PW, van Sinderen D, Marchesi JR, Ventura M. 2012. Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* 7:e36957. <https://doi.org/10.1371/journal.pone.0036957>.
- Milani C, Lugli GA, Duranti S, Turroni F, Mancabelli L, Ferrario C, Mangifesta M, Hevia A, Viappiani A, Scholz M, Arioli S, Sanchez B, Lane J, Ward DV, Hickey R, Mora D, Segata N, Margolles A, van Sinderen D, Ventura M. 2015. Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. *Sci Rep* 5:15782. <https://doi.org/10.1038/srep15782>.
- Odamaki T, Horigome A, Sugahara H, Hashikura N, Minami J, Xiao JZ, Abe F. 2015. Comparative genomics revealed genetic diversity and species/strain-level differences in carbohydrate metabolism of three probiotic bifidobacterial species. *Int J Genomics* 2015:567809. <https://doi.org/10.1155/2015/567809>.
- Ouwehand AC, Isolauri E, Kirjavainen PV, Salminen SJ. 1999. Adhesion of four *Bifidobacterium* strains to human intestinal mucus from subjects in different age groups. *FEMS Microbiol Lett* 172:61–64. <https://doi.org/10.1111/j.1574-6968.1999.tb13450.x>.
- He F, Ouwehand AC, Isolauri E, Hosoda M, Benno Y, Salminen S. 2001. Differences in composition and mucosal adhesion of bifidobacteria isolated from healthy adults and healthy seniors. *Curr Microbiol* 43:351–354. <https://doi.org/10.1007/s002840010315>.
- Shepherd ES, DeLoache WC, Pruss KM, Whitaker WR, Sonnenburg JL. 2018. An exclusive metabolic niche enables strain engraftment in the gut microbiota. *Nature* 557:434–438. <https://doi.org/10.1038/s41586-018-0092-4>.
- Milani C, Turroni F, Duranti S, Lugli GA, Mancabelli L, Ferrario C, van Sinderen D, Ventura M. 2016. Genomics of the genus *Bifidobacterium* reveals species-specific adaptation to the glycan-rich gut environment. *Appl Environ Microbiol* 82:980–991. <https://doi.org/10.1128/AEM.03500-15>.
- Arbolea S, Bottacini F, O'Connell-Motherway M, Ryan CA, Ross RP, van Sinderen D, Stanton C. 2018. Gene-trait matching across the *Bifidobacterium longum* pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. *BMC Genomics* 19:33. <https://doi.org/10.1186/s12864-017-4388-9>.
- Seifert GJ, Roberts K. 2007. The biology of arabinogalactan proteins. *Annu Rev Plant Biol* 58:137–161. <https://doi.org/10.1146/annurev.arplant.58.032806.103801>.
- Street CA, Anderson DM. 1983. Refinement of structures previously proposed for gum arabic and other *Acacia* gum exudates. *Talanta* 30:887–893. [https://doi.org/10.1016/0039-9140\(83\)80206-9](https://doi.org/10.1016/0039-9140(83)80206-9).
- Aalbers F, Turkenburg JP, Davies GJ, Dijkhuizen L, Lammerts van Bueren A. 2015. Structural and functional characterization of a novel family GH115 4-O-methyl- α -glucuronidase with specificity for decorated arabinogalactans. *J Mol Biol* 427:3935–3946. <https://doi.org/10.1016/j.jmb.2015.07.006>.
- Goellner EM, Utermohlen J, Kramer R, Classen B. 2011. Structure of arabinogalactan from *Larix laricina* and its reactivity with antibodies directed against type-II-arabinogalactans. *Carbohydr Polym* 86:1739–1744. <https://doi.org/10.1016/j.carbpol.2011.07.006>.
- Jermyn M, Yeow YM, Woods E. 1975. A class of lectins present in the tissues of seed plants. *Functional Plant Biol* 2:501–531. <https://doi.org/10.1071/PP9750501>.
- Ichinose H, Kotake T, Tsumuraya Y, Kaneko S. 2008. Arabinogalactan-proteins degrading enzymes. *J Appl Glycosci* 55:149–155. <https://doi.org/10.5458/jag.55.149>.
- Patel S, Goyal A. 2015. Applications of natural polymer gum arabic: a review. *Int J Food Properties* 18:986–998. <https://doi.org/10.1080/10942912.2013.809541>.
- Terpend K, Possemiers S, Daguet D, Marzorati M. 2013. Arabinogalactan and fructo-oligosaccharides have a different fermentation profile in the Simulator of the Human Intestinal Microbial Ecosystem (SHIME). *Environ Microbiol Rep* 5:595–603. <https://doi.org/10.1111/1758-2229.12056>.
- Daguet D, Pinheiro I, Verhelst A, Possemiers S, Marzorati M. 2016. Arabinogalactan and fructooligosaccharides improve the gut barrier function in distinct areas of the colon in the Simulator of the Human Intestinal Microbial Ecosystem. *J Funct Foods* 20:369–379. <https://doi.org/10.1016/j.jff.2015.11.005>.
- Calame W, Weseler AR, Viebke C, Flynn C, Siemensma AD. 2008. Gum arabic establishes prebiotic functionality in healthy human volunteers in a dose-dependent manner. *Br J Nutr* 100:1269–1275. <https://doi.org/10.1017/S0007114508981447>.
- Crociani F, Alessandrini A, Mucci MM, Biavati B. 1994. Degradation of complex carbohydrates by *Bifidobacterium* spp. *Int J Food Microbiol* 24:199–210. [https://doi.org/10.1016/0168-1605\(94\)90119-8](https://doi.org/10.1016/0168-1605(94)90119-8).
- Salys AA, West SE, Vercellotti JR, Wilkins TD. 1977. Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* 34:529–533. <https://doi.org/10.1128/AEM.34.5.529-533.1977>.
- Saishin N, Yamamoto I. 2009. α -Galactosidase purified from *Bifidobacterium longum* JCM 7052 grown on gum arabic. *J Biol Macromol* 9:71–80.
- Sasaki Y, Togo N, Kitahara K, Fujita K. 2018. Characterization of a GH36 β -L-arabinopyranosidase in *Bifidobacterium adolescentis*. *J Appl Glycosci* 65:23–30. https://doi.org/10.5458/jag.jag.2018_001.
- Aspinall GO, Knebl MC. 1986. The location of α -D-galactopyranose residues in gum arabic. *Carbohydr Res* 157:257–260. [https://doi.org/10.1016/0008-6215\(86\)85074-1](https://doi.org/10.1016/0008-6215(86)85074-1).
- Cao Y, Yuan T, Shi P, Luo H, Li N, Meng K, Bai Y, Yang P, Zhou Z, Zhang Z, Yao B. 2010. Properties of a novel α -galactosidase from *Streptomyces* sp. S27 and its potential for soybean processing. *Enzyme Microb Technol* 47:305–312. <https://doi.org/10.1016/j.enzmictec.2010.09.007>.
- Odamaki T, Bottacini F, Kato K, Mitsuyama E, Yoshida K, Horigome A, Xiao JZ, van Sinderen D. 2018. Genomic diversity and distribution of *Bifidobacterium longum* subsp. *longum* across the human lifespan. *Sci Rep* 8:85. <https://doi.org/10.1038/s41598-017-18391-x>.
- Fujita K, Sakamoto A, Kaneko S, Kotake T, Tsumuraya Y, Kitahara K. 2019. Degradative enzymes for type II arabinogalactan side chains in *Bifidobacterium longum* subsp. *longum*. *Appl Microbiol Biotechnol* 103:1299–1310. <https://doi.org/10.1007/s00253-018-9566-4>.
- Pitson SM, Voragen AG, Beldman G. 1996. Stereochemical course of hydrolysis catalyzed by arabinofuranosyl hydrolases. *FEBS Lett* 398:7–11. [https://doi.org/10.1016/S0014-5793\(96\)01153-2](https://doi.org/10.1016/S0014-5793(96)01153-2).
- Maruta A, Yamane M, Matsubara M, Suzuki S, Nakazawa M, Ueda M, Sakamoto T. 2017. A novel α -galactosidase from *Fusarium oxysporum* and its application in determining the structure of the gum arabic side chain. *Enzyme Microb Technol* 103:25–33. <https://doi.org/10.1016/j.enzmictec.2017.04.006>.
- Jones DR, Uddin MS, Gruninger RJ, Pham TTM, Thomas D, Boraston AB, Briggs J, Pluvinaige B, McAllister TA, Forster RJ, Tsang A, Selinger LB, Abbott DW. 2017. Discovery and characterization of family 39 glycoside hydrolases from rumen anaerobic fungi with polyspecific activity on rare arabinosyl substrates. *J Biol Chem* 292:12606–12620. <https://doi.org/10.1074/jbc.M117.789008>.
- McCarter JD, Withers GS. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol* 4:885–892. [https://doi.org/10.1016/0959-440X\(94\)90271-2](https://doi.org/10.1016/0959-440X(94)90271-2).
- Ichinose H, Fujimoto Z, Honda M, Harazono K, Nishimoto Y, Uzura A, Kaneko S. 2009. A β -L-arabinopyranosidase from *Streptomyces avermitilis* is a novel member of glycoside hydrolase family 27. *J Biol Chem* 284:25097–25106. <https://doi.org/10.1074/jbc.M109.022723>.
- Shimokawa M, Kitahara K, Fujita K. 2015. Characterization of a β -L-arabinopyranosidase from *Bifidobacterium longum* subsp. *longum*. *J Appl Glycosci* 62:1–6. https://doi.org/10.5458/jag.jag.2014_006.

36. Liu D, Wang S, Xu B, Guo Y, Zhao J, Liu W, Sun Z, Shao C, Wei X, Jiang Z, Wang X, Liu F, Wang J, Huang L, Hu D, He X, Riedel CU, Yuan J. 2011. Proteomics analysis of *Bifidobacterium longum* NCC2705 growing on glucose, fructose, mannose, xylose, ribose, and galactose. *Proteomics* 11:2628–2638. <https://doi.org/10.1002/pmic.201100035>.
37. Ichinose H, Yoshida M, Kotake T, Kuno A, Igarashi K, Tsumuraya Y, Samejima M, Hirabayashi J, Kobayashi H, Kaneko S. 2005. An exo- β -1,3-galactanase having a novel β -1,3-galactan-binding module from *Phanerochaete chrysosporium*. *J Biol Chem* 280:25820–25829. <https://doi.org/10.1074/jbc.M501024200>.
38. Ichinose H, Kuno A, Kotake T, Yoshida M, Sakka K, Hirabayashi J, Tsumuraya Y, Kaneko S. 2006. Characterization of an exo- β -1,3-galactanase from *Clostridium thermocellum*. *Appl Environ Microbiol* 72:3515–3523. <https://doi.org/10.1128/AEM.72.5.3515-3523.2006>.
39. Sakamoto T, Ishimaru M. 2013. Peculiarities and applications of galactanolytic enzymes that act on type I and II arabinogalactans. *Appl Microbiol Biotechnol* 97:5201–5213. <https://doi.org/10.1007/s00253-013-4946-2>.
40. Fujita K, Sakaguchi T, Sakamoto A, Shimokawa M, Kitahara K. 2014. *Bifidobacterium longum* subsp. *longum* exo- β -1,3-galactanase, an enzyme for the degradation of type II arabinogalactan. *Appl Environ Microbiol* 80:4577–4584. <https://doi.org/10.1128/AEM.00802-14>.
41. Saito K, Viborg AH, Sakamoto S, Arakawa T, Yamada C, Fujita K, Fushinobu S. 2020. Crystal structure of β -L-arabinobiosidase belonging to glycoside hydrolase family 121. *PLoS One* 15:e0231513. <https://doi.org/10.1371/journal.pone.0231513>.
42. Correia MA, Mazumder K, Bras JL, Firbank SJ, Zhu Y, Lewis RJ, York WS, Fontes CM, Gilbert HJ. 2011. Structure and function of an arabinoxylan-specific xylanase. *J Biol Chem* 286:22510–22520. <https://doi.org/10.1074/jbc.M110.217315>.
43. Couturier M, Roussel A, Rosengren A, Leone P, Stalbrand H, Berrin JG. 2013. Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β -(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. *J Biol Chem* 288:14624–14635. <https://doi.org/10.1074/jbc.M113.459438>.
44. Cartmell A, Munoz-Munoz J, Briggs JA, Ndeh DA, Lowe EC, Basle A, Terrapon N, Stott K, Heunis T, Gray J, Yu L, Dupree P, Fernandes PZ, Shah S, Williams SJ, Labourel A, Trost M, Henrissat B, Gilbert HJ. 2018. A surface endogalactanase in *Bacteroides thetaiotaomicron* confers keystone status for arabinogalactan degradation. *Nat Microbiol* 3:1314–1326. <https://doi.org/10.1038/s41564-018-0258-8>.
45. Sims IM, Furneaux RH. 2003. Structure of the exudate gum from *Meryta sinclairii*. *Carbohydr Polym* 52:423–431. [https://doi.org/10.1016/S0144-8617\(02\)00326-0](https://doi.org/10.1016/S0144-8617(02)00326-0).
46. Goodrum LJ, Patel A, Leykam JF, Kieliszewski MJ. 2000. Gum arabic glycoprotein contains glycomodules of both extensin and arabinogalactan-glycoproteins. *Phytochemistry* 54:99–106. [https://doi.org/10.1016/S0031-9422\(00\)00043-1](https://doi.org/10.1016/S0031-9422(00)00043-1).
47. Nie SP, Wang C, Cui SW, Wang Q, Xie MY, Phillips GO. 2013. A further amendment to the classical core structure of gum arabic (*Acacia senegal*). *Food Hydrocoll* 31:42–48. <https://doi.org/10.1016/j.foodhyd.2012.09.014>.
48. Ponder GR, Richards GN. 1997. Arabinogalactan from Western larch. Part III: alkaline degradation revisited, with novel conclusions on molecular structure. *Carbohydr Polym* 34:251–261. [https://doi.org/10.1016/S0144-8617\(97\)00099-4](https://doi.org/10.1016/S0144-8617(97)00099-4).
49. Tryfona T, Liang HC, Kotake T, Kaneko S, Marsh J, Ichinose H, Lovegrove A, Tsumuraya Y, Shewry PR, Stephens E, Dupree P. 2010. Carbohydrate structural analysis of wheat flour arabinogalactan protein. *Carbohydr Res* 345:2648–2656. <https://doi.org/10.1016/j.carres.2010.09.018>.
50. Fujita K, Sakamoto S, Ono Y, Wakao M, Suda Y, Kitahara K, Suganuma T. 2011. Molecular cloning and characterization of a β -L-arabinobiosidase in *Bifidobacterium longum* that belongs to a novel glycoside hydrolase family. *J Biol Chem* 286:5143–5150. <https://doi.org/10.1074/jbc.M110.190512>.
51. Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D. 2013. An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25:270–287. <https://doi.org/10.1105/tpc.112.107334>.
52. Alarifi S, Bell A, Walton G. 2018. *In vitro* fermentation of gum acacia—impact on the faecal microbiota. *Int J Food Sci Nutr* 69:696–704. <https://doi.org/10.1080/09637486.2017.1404970>.
53. Kaeothip S, Ishiwata A, Ito T, Fushinobu S, Fujita K, Ito Y. 2013. Preparation of p-nitrophenyl β -L-arabinofuranoside as a substrate of β -L-arabinofuranosidase. *Carbohydr Res* 382:95–100. <https://doi.org/10.1016/j.carres.2013.10.005>.
54. Tischer CA, Gorin PAJ, Iacomini M. 2002. The free reducing oligosaccharides of gum arabic: aids for structural assignments in the polysaccharide. *Carbohydr Polym* 47:151–158. [https://doi.org/10.1016/S0144-8617\(01\)00173-4](https://doi.org/10.1016/S0144-8617(01)00173-4).
55. Mahendran T, Williams PA, Phillips GO, Al-Assaf S, Baldwin TC. 2008. New insights into the structural characteristics of the arabinogalactan-protein (AGP) fraction of gum arabic. *J Agric Food Chem* 56:9269–9276. <https://doi.org/10.1021/jf800849a>.
56. Ling NXY, Pettolino F, Liao ML, Bacic A. 2009. Preparation of a new chromogenic substrate to assay for β -galactanases that hydrolyse type II arabinogalactans. *Carbohydr Res* 344:1941–1946. <https://doi.org/10.1016/j.carres.2009.07.014>.
57. Bednarski MD, Chenault HK, Simon ES, Whitesides GM. 1987. Membrane-enclosed enzymic catalysis (MEEC): a useful, practical new method for the manipulation of enzymes in organic synthesis. *J Am Chem Soc* 109:1283–1285. <https://doi.org/10.1021/ja00238a067>.
58. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.
59. Holmes EW, O'Brien JS. 1979. Separation of glycoprotein-derived oligosaccharides by thin-layer chromatography. *Anal Biochem* 93:167–170. [https://doi.org/10.1016/S0003-2697\(79\)80131-1](https://doi.org/10.1016/S0003-2697(79)80131-1).
60. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.