



# Bifidobacterial GH146 $\beta$ -L-arabinofuranosidase for the removal of $\beta$ 1,3-L-arabinofuranosides on plant glycans

Kiyotaka Fujita<sup>1,2</sup> · Hanako Tsunomachi<sup>1</sup> · Pan Lixia<sup>3,4</sup> · Shun Maruyama<sup>3</sup> · Masayuki Miyake<sup>3</sup> · Aimi Dakeshita<sup>1</sup> · Kanefumi Kitahara<sup>1,2</sup> · Katsunori Tanaka<sup>5,6</sup> · Yukishige Ito<sup>5,7</sup> · Akihiro Ishiwata<sup>5</sup> · Shinya Fushinobu<sup>3,8</sup>

Received: 16 November 2023 / Revised: 29 December 2023 / Accepted: 14 January 2024  
© The Author(s) 2024

## Abstract

L-Arabinofuranosides with  $\beta$ -linkages are present in several plant molecules, such as arabinogalactan proteins (AGPs), extensin, arabinan, and rhamnogalacturonan-II. We previously characterized a  $\beta$ -L-arabinofuranosidase from *Bifidobacterium longum* subsp. *longum* JCM 1217, Bll1HypBA1, which was found to belong to the glycoside hydrolase (GH) family 127. This strain encodes two GH127 genes and two GH146 genes. In the present study, we characterized a GH146  $\beta$ -L-arabinofuranosidase, Bll3HypBA1 (BLLJ\_1848), which was found to constitute a gene cluster with AGP-degrading enzymes. This recombinant enzyme degraded AGPs and arabinan, which contain Araf- $\beta$ 1,3-Araf structures. In addition, the recombinant enzyme hydrolyzed oligosaccharides containing Araf- $\beta$ 1,3-Araf structures but not those containing Araf- $\beta$ 1,2-Araf and Araf- $\beta$ 1,5-Araf structures. The crystal structures of Bll3HypBA1 were determined at resolutions up to 1.7 Å. The monomeric structure of Bll3HypBA1 comprised a catalytic  $(\alpha/\alpha)_6$  barrel and two  $\beta$ -sandwich domains. A hairpin structure with two  $\beta$ -strands was observed in Bll3HypBA1, to extend from a  $\beta$ -sandwich domain and partially cover the active site. The active site contains a Zn<sup>2+</sup> ion coordinated by Cys<sub>3</sub>-Glu and exhibits structural conservation of the GH127 cysteine glycosidase Bll1HypBA1. This is the first study to report on a  $\beta$ 1,3-specific  $\beta$ -L-arabinofuranosidase.

## Key points

- $\beta$ 1,3-L-Arabinofuranose residues are present in arabinogalactan proteins and arabinans as a terminal sugar.
- $\beta$ -L-Arabinofuranosidases are widely present in intestinal bacteria.
- Bll3HypBA1 is the first enzyme characterized as a  $\beta$ 1,3-linkage-specific  $\beta$ -L-arabinofuranosidase.

**Keywords**  $\beta$ -L-Arabinofuranosidase · *Bifidobacterium longum* subsp. *longum* · Arabinogalactan-protein · Glycoside hydrolase

✉ Kiyotaka Fujita  
k4022897@kadai.jp

✉ Shinya Fushinobu  
asfushi@mail.ecc.u-tokyo.ac.jp

<sup>1</sup> Faculty of Agriculture, Kagoshima University, 1-21-24 Korimoto, Kagoshima, Kagoshima 890-0065, Japan

<sup>2</sup> The United Graduate School of Agricultural Sciences, Kagoshima University, 1-21-24 Korimoto, Kagoshima, Kagoshima 890-0065, Japan

<sup>3</sup> Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-Ku, Tokyo 113-8657, Japan

<sup>4</sup> National Key Laboratory of Non-food Biomass Energy Technology, Guangxi Key Laboratory of Marine Natural

Products and Combinatorial Biosynthesis Chemistry, Guangxi Academy of Sciences, Nanning 530007, China

<sup>5</sup> RIKEN, Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

<sup>6</sup> Department of Chemical Science and Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-Ku, Tokyo 152-8552, Japan

<sup>7</sup> Graduate School of Science, Osaka University, 1-1 Machikaneyama-Cho, Toyonaka, Osaka 560-0043, Japan

<sup>8</sup> Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-Ku, Tokyo 113-8657, Japan

## Introduction

L-Arabinose occurs in four distinct forms in plant polysaccharides:  $\alpha$ - and  $\beta$ -L-arabinofuranose (Araf) and L-arabinopyranose (Arap). Although  $\beta$ -Araf occurs in significantly less amount than  $\alpha$ -Araf in plant carbohydrate polymers, it is present as  $\beta$ -L-arabinooligosaccharide ( $\beta$ -AOS) chains in extensin (e.g., Araf- $\beta$ 1,2-Araf- $\beta$ 1,2-Araf- $\beta$ 1-hydroxyproline (Hyp); Ara<sub>3</sub>-Hyp).  $\beta$ -Araf is also present as a terminal sugar in plant polysaccharides, e.g., terminal  $\beta$ 1,2- and  $\beta$ 1,5-Araf are present in the complicated sugar chain of rhamnogalacturonan-II (RG-II) in pectin (Pellerin et al. 1996). Moreover, terminal Araf- $\beta$ 1,3-Araf- $\alpha$ 1 structures are found in several plant polysaccharides, including arabinan from sugar beet (Wefers et al. 2017) and quinoa seeds (Wefers et al. 2014), arabinoxylan from green leaves of Lauraceae (Das et al. 2013), arabinoxyloglucan from tomato cultured cells (York et al. 1996), and arabinogalactan proteins (AGPs) from rice anthers (Kawaguchi et al. 1996).  $\beta$ -Arap is also present as a terminal sugar in plant polysaccharides, including type-II arabinogalactan (AG) from larch (Ponder and Richards 1997), AGPs from wheat (Tryfona et al. 2010), and gum arabic (Tischer et al. 2002).  $\alpha$ -Arap is found only as a constituent sugar of RG-II in plant polysaccharides (Pellerin et al. 1996).

Pfam DUF1680 (PF07944), which was renamed to Glyco\_hydro\_127, contains 24,000 proteins that are distributed among 7482 bacterial, fungal, and plant species. HypBA1 (Bll1HypBA1; encoded by BLLJ\_0211) is the first characterized GH127  $\beta$ -L-arabinofuranosidase (EC 3.2.1.185) (Fujita et al. 2014b). Bll1HypBA1 from *Bifidobacterium longum* subsp. *longum* JCM 1217 degrades Araf- $\beta$ 1,2-Ara released from Ara<sub>3</sub>-Hyp by GH121  $\beta$ -L-arabinobiosidase (HypBA2: BLLJ\_0212) (Fujita et al. 2011). Crystallographic studies of Bll1HypBA1 have revealed that a cysteine residue serves as the catalytic nucleophile (Ishiwata et al. 2022b; Ito et al. 2014; Maruyama et al. 2022; McGregor et al. 2021). In addition to GH137 and GH142  $\beta$ -L-arabinofuranosidases for RG-II degradation in *Bacteroides thetaiotaomicron* (Ndeh et al. 2017), GH146 was established after characterization of BT0349 as an  $\beta$ -L-arabinofuranosidase for arabinan-derived arabinotetraose and  $\beta$ 1,2-L-arabinobiose (Luis et al. 2018). We also characterized a GH146 member, XCV2724, from *Xanthomonas euvesicatoria* (XeHypBA1) (Nakamura et al. 2018). This enzyme hydrolyzed Araf- $\beta$ 1-Hyp predominately and partially for Araf- $\beta$ 1,2-Ara (Ishiwata et al. 2023). In addition to two GH127 members (Bll1HypBA1 and BLLJ\_1826), *B. longum* subsp. *longum* JCM 1217 encodes two GH146 members (BLLJ\_1848 and BLLJ\_0089). BLLJ\_0089 was recently characterized as a  $\beta$ -L-arabinofuranosidase (Bll4HypBA1) for Araf- $\beta$ 1-linked

Hyp on the backbone of Hyp-rich glycoprotein (HRGP) (Ishiwata et al. 2023). At present, the three-dimensional structure of only one enzyme (BT0349) in the GH146 family has been reported (Luis et al. 2018; McGregor et al. 2021). In the present study, we characterized the recombinant BLLJ\_1848 as a  $\beta$ -1,3-specific GH146  $\beta$ -L-arabinofuranosidase (Bll3HypBA1) using natural and synthetic substrates. In addition, we performed X-ray crystallography of Bll3HypBA1.

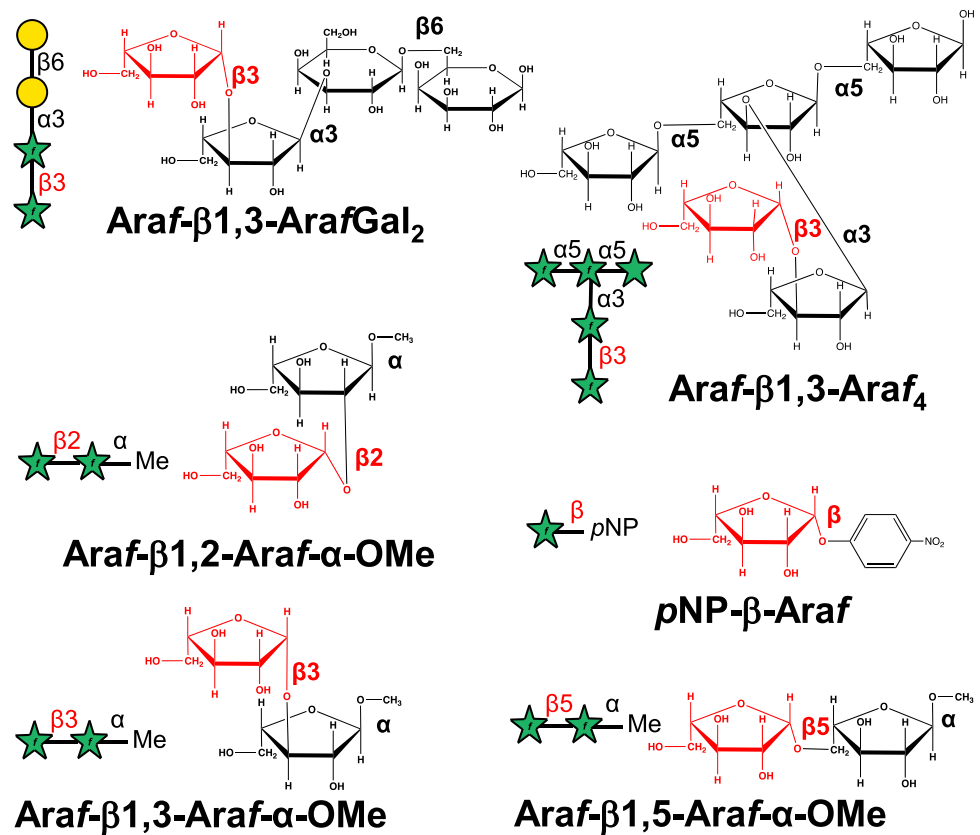
## Materials and methods

### Substrates

Larch wood AG was purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). Gum arabic from *Acacia senegal* was purchased from Sigma-Aldrich (St. Louis, MO, USA). *p*-Nitrophenyl (*p*NP)  $\beta$ -L-arabinofuranoside was synthesized as described in a previous study (Kaeothip et al. 2013a). Young rice panicles, cultivar Nikomaru, were harvested from experimental farms at the Faculty of Agriculture, Kagoshima University (Kagoshima, Japan). According to a previous study (Kawaguchi et al. 1996), L-Araf- $\beta$ 1,3-L-Araf- $\alpha$ 1,3-Gal- $\beta$ 1,6-Gal (Araf- $\beta$ 1,3-ArafGal<sub>2</sub>) and AGP from rice anthers were prepared from young rice panicles. L-Araf- $\alpha$ 1,5-[L-Araf- $\beta$ 1,3-L-Araf- $\alpha$ 1,3-]-L-Araf- $\alpha$ 1,5-L-Ara (Araf- $\beta$ 1,3-Araf<sub>4</sub>) from quinoa seed arabinan was prepared through endo-1,5- $\alpha$ -L-arabinanase treatment as described by Wefers et al. (Wefers et al. 2014). These oligosaccharides were further purified using HPLC on a Cosmosil PBr column (Nacalai Tesque Inc., Kyoto, Japan) as described previously (Sasaki et al. 2021). The fractions containing oligosaccharides were collected and analyzed using MALDI-TOF MS (Bruker Daltonics, Leipzig, Germany).

L-Araf- $\beta$ 1,2-L-Araf- $\alpha$ 1-OMe (Araf- $\beta$ 1,2-Araf- $\alpha$ -OMe), L-Araf- $\beta$ 1,3-L-Araf- $\alpha$ 1-OMe (Araf- $\beta$ 1,3-Araf- $\alpha$ -OMe), and L-Araf- $\beta$ 1,5-L-Araf- $\alpha$ 1-OMe (Araf- $\beta$ 1,5-Araf- $\alpha$ -OMe) were previously prepared stereospecifically (Ishiwata et al. 2022a) via 1,2-*cis* selective  $\beta$ -arabinofuranosylation using 2-naphthylmethyl ether-mediated intramolecular aglycon delivery (Ishiwata and Ito 2011; Ishiwata et al. 2014, 2008; Kaeothip et al. 2013b).  $\beta$ -L-Arabinooligosaccharides were prepared according to previous reports (Fujita et al. 2014b). Other chemicals were purchased from Fujifilm Wako Pure Chemical Industries Ltd. (Osaka, Japan). Araf- $\beta$ 1,3-ArafGal<sub>2</sub> was labeled with *p*-aminobenzoic ethyl ester (ABEE) as described by Wang (Wang et al. 1984). The ABEE derivative was purified on a Cosmosil PBr column with a linear gradient of CH<sub>3</sub>CN/water from

**Fig. 1** Chemical and schematic structures of the substrates used in this study.  $\beta$ -L-Araf-structure is highlighted in red. Symbols: D-galactose, yellow circle; L-arabinose, green star



**Table 1** Enzymes used in this study

Enzyme name (locus tag)	Family	Activity
Bll1HypBA1 (BLLJ_0211)	GH127	$\beta$ -L-Arabinofuranosidase
Bll2HypBA1 (BLLJ_1826)	GH127	$\beta$ -L-Arabinofuranosidase
Bll3HypBA1 (BLLJ_1848)	GH146	$\beta$ -L-Arabinofuranosidase
Bll4HypBA1 (BLLJ_0089)	GH146	$\beta$ -L-Arabinofuranosidase
BIArafA (BLLJ_1854)	GH43_22	$\alpha$ -L-Arabinofuranosidase
Bll3Gal (BLLJ_1840)	GH43_24	Exo- $\beta$ -1,3-galactanase
AnAFASE	GH51	$\alpha$ -L-Arabinofuranosidase
AnEARAB	GH43	Endo-1,5- $\alpha$ -L-arabinanase

0:100 to 70:30 (v/v) for 30 min at a constant flow rate of 4.7 mL/min at 30 °C. Elution was monitored using a fluorescence detector (FP-2020, JASCO, Japan) with Ex/Em of 305/360 nm. Figure 1 depicts the chemical structures of the substrates used in this study.

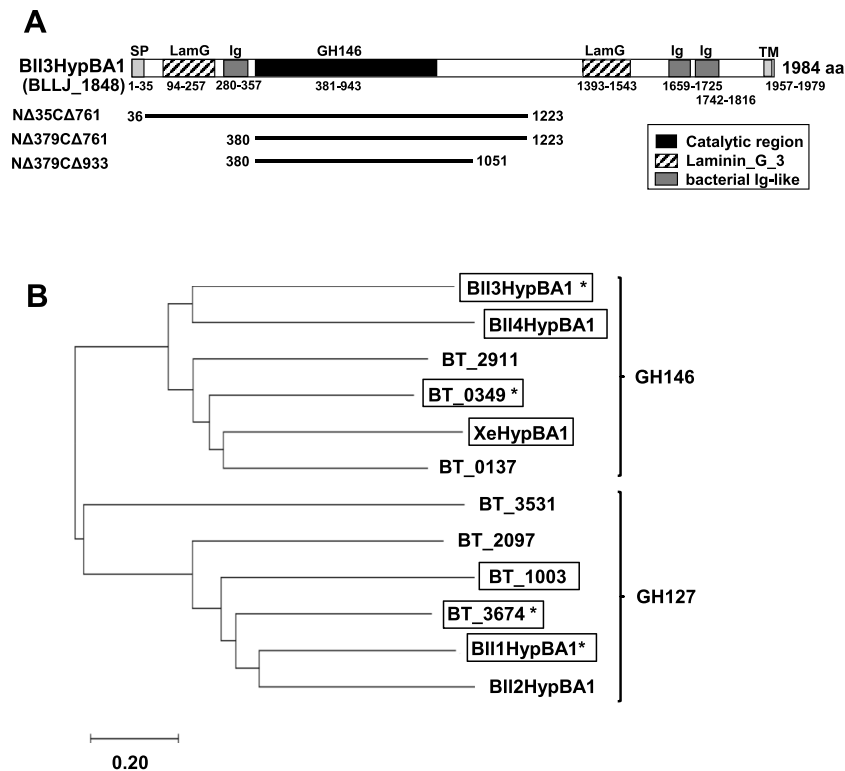
## Enzymes

$\beta$ -L-Arabinofuranosidases Bll1HypBA1 (Fujita et al. 2014b), Bll2HypBA1, and Bll4HypBA1 (Ishiwata et al. 2023) enzymes were prepared according to previous reports. GH43\_24 exo- $\beta$ -1,3-galactanase Bll3Gal (Fujita et al. 2014a) and GH43\_22  $\alpha$ -L-arabinofuranosidase

BIArafA (Fujita et al. 2019a) were prepared as described previously. Endo-1,5- $\alpha$ -L-arabinanase (AnEARAB) and  $\alpha$ -L-arabinofuranosidase (AnAFASE) from *Aspergillus niger* were purchased from Megazyme (Bray, Ireland). Table 1 shows the enzyme list used in this study.

## Expression and purification of recombinant Bll3HypBA1

The genomic DNA of *B. longum* subsp. *longum* JCM 1217 was subjected to PCR amplification of the Bll3HypBA1 gene. The forward (5'-AGGAGATATACCATGGCAGAA GTGGACTCCAGCA-3') and reverse (5'-TGCTCGAGT GCGGCCGCGTTCTGCATGCGCAC-3') primers were designed using nucleotides 106–124 and 3655–3671 from the Bll3HypBA1 gene. The underlined texts indicate complementary nucleotides to the template. The PCR amplification of Bll3HypBA1-N $\Delta$ 35C $\Delta$ 761, which encodes amino acids (aa) 36–1223, was designed to eliminate the N-terminal signal peptide and C-terminal domains (Fig. 2A). The amplicon was cloned into the pET-23d vector (Novagen, Madison, WI, USA) using the In-Fusion HD Cloning Kit (Clontech Laboratories Inc., Palo Alto, CA, USA). The KOD-plus mutagenesis kit (Toyobo Co., Ltd., Osaka, Japan) and the primers listed in Table S1 were used to generate the Bll3HypBA1 deletion mutants. The primers N $\Delta$  reverse



**Fig. 2** Sequence features of BII3HypBA1 and GH127/146  $\beta$ -L-arabinofuranosidases. **A** Domain structure of BII3HypBA1. Domain structures were predicted using SignalP5.0 (<https://services.healthtech.dtu.dk/services/SignalP-5.0/>) and InterPro (<https://www.ebi.ac.uk/interpro/>) servers. The domains were indicated as follows: signal peptide (SP), laminin\_G\_3 (LamG), bacterial Ig-like domain (Ig), and transmembrane region (TM). Here, the lines indicate the expression regions of the BII3HypBA1 recombinant proteins. **B** The phylogenetic tree of GH146 and GH127  $\beta$ -L-arabinofuranosidase members in *Bacteroides thetaiotaomicron* VPI-5482, *X. euvesicatoria*, and *B. longum* subsp. *longum* JCM 1217. The phylogenetic tree was constructed with the MEGAX software. Enzymes char-

acterized previously or in this study are enclosed in the box. Here, the asterisk (\*) indicates enzymes for which crystallographic studies have been performed. Almost all enzymes were characterized as  $\beta$ -L-arabinofuranosidase, except for BT1003 (acetic acid hydrolase). GenBank accession numbers are shown alongside the characterized enzymatic names and/or protein locus tags as follows: BII1HypBA1/BLLJ\_0211 (BAJ65881), BII2HypBA1/BLLJ\_1826 (BAJ67491), BII4HypBA1/BLLJ\_0089 (BAJ65759), XeHypBA1/XCV2724 (CAJ24403), BT\_0349 (AAO75456), BT\_2911 (AAO78017), BT\_0137 (AAO75244), BT3531 (AAO78637), BT2097 (AAO77204), BT1003 (AAO76110), BT3674 (AAO78779)

and NΔ379 forward were used to construct BII3HypBA1-NΔ379CΔ761 (aa: 380–1223), whereas the primers CΔ forward and CΔ933 reverse were used to construct BII3HypBA1-NΔ379CΔ933 (aa: 380–1051) (Fig. 2A). The constructed plasmids were sequenced and transformed into *E. coli* BL21 ( $\lambda$ DE3) cells, which were then grown at 20 °C using the Overnight Express Autoinduction System (Novagen). The His-tagged proteins were extracted using BugBuster protein extraction reagent (Novagen), purified on a TALON metal affinity resin (Clontech Laboratories Inc.), and then desalted and concentrated using an ultrafiltration membrane (10-kDa cut-off; Millipore Co., Billerica, MA, USA). The sequence of the BII3HypBA1-NΔ35CΔ761 gene was deposited in the DDBJ database under the accession number LC765464.

### Thin-layer chromatography (TLC), high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD), and HPLC

For TLC, silica gel 60 aluminum plates (Merck KGaA, Darmstadt, Germany) were used with a 7:1:2 (v/v/v) of 1-propanol/EtOH/water solvent mixture and then visualized by spraying orcinol-sulfate on the plates (Holmes and O'Brien 1979). For HPAEC-PAD, oligosaccharides were analyzed using a CarboPac PA-1 column (Dionex Corp., Sunnyvale, CA, USA) as described previously (Sasaki et al. 2021). For HPLC, ABEE-labeled Araf- $\beta$ 1,3-Araf/Gal<sub>2</sub> and reaction products were analyzed by a Cosmosil Sugar-D (Nacalai Tesque Inc.) as described previously (Ishiwata et al. 2022b).

## pH and temperature dependence of BII3HypBA1 activity

The hydrolytic activity of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 was assayed using Ara $\beta$ -1,3-Ara $\beta$ Gal<sub>2</sub>-ABEE as the substrate and 50 mM sodium acetate (pH 3.5–6.0) and 50 mM sodium phosphate (pH 6.5–8.0) as buffers. The substrate (25  $\mu$ M) was incubated with 0.05  $\mu$ g/ml of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 in 40  $\mu$ L of each buffer (50 mM) at 40 °C for 20 min. The reaction was stopped by adding 10  $\mu$ L of 5% trichloroacetic acid (TCA) and then analyzed via HPLC. Next, the effect of temperature on enzyme activity was examined using 50 mM sodium acetate buffer (pH 5.5) at 25–60 °C.

## Substrate specificities of BII3HypBA1 toward polysaccharides

The hydrolytic activity of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 was analyzed using sugar beet arabinan, larch AG, and gum arabic AGP as polysaccharide substrates. These substrates (1.0%) were incubated with 0.1  $\mu$ g/mL of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761, BII1HypBA1, BII2HypBA1, or BII4HypBA1 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 40 °C for 16 h. Following incubation, the released L-arabinose was analyzed using TLC, as described previously. For the comparison of the specific activity of BII3HypBA1 with or without N-terminal LamG domain, the substrates (1.0%) were incubated with 0.1  $\mu$ g/ml of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 and BII3HypBA1-N $\Delta$ 379C $\Delta$ 761 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 40 °C for 2 h. The reaction was terminated by boiling, and the liberated L-arabinose was analyzed using HPAEC-PAD.

## Combination reactions of BII3HypBA1 and AGP degradative enzymes

BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 (0.1  $\mu$ g/ml) reacted with Ara $\beta$ -1,3-Ara $\beta$ Gal<sub>2</sub> and Ara $\beta$ -1,3-Ara $\beta$ <sub>4</sub> with or without BIAra $\beta$ A/AnAFASE at 40 °C for 16 h. The reaction mixtures were analyzed via HPAEC-PAD and MALDI-TOF MS. To determine the mode of action of rice AGP, the following reactions involving BII3HypBA1-N $\Delta$ 35C $\Delta$ 761, BIAra $\beta$ A, and BII1,3Gal were conducted: 1.0% rice AGP were incubated with enzymes (0.1  $\mu$ g/mL each) in a 50 mM sodium phosphate buffer (pH 6.0) at 37 °C for 16 h. The reaction products were analyzed via HPAEC-PAD.

## Substrate specificities of BII3HypBA1 toward oligosaccharides

Next, the substrate specificities of BII3HypBA1 were analyzed using synthetic and natural substrates. Since the

region following residue 1052 in the N $\Delta$ 379C $\Delta$ 761 was not affected to structure of the catalytic region, BII3HypBA1-N $\Delta$ 379C $\Delta$ 761 or C $\Delta$ 933 were used for the analysis of substrate specificities. Synthetic Ara $\beta$ -1,2-Ara $\beta$ - $\alpha$ -OMe, Ara $\beta$ -1,3-Ara $\beta$ - $\alpha$ -OMe, and Ara $\beta$ -1,5-Ara $\beta$ - $\alpha$ -OMe were incubated with 0.1  $\mu$ g/ml of the BII3HypBA1-N $\Delta$ 379C $\Delta$ 933 or BII1HypBA1 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 37 °C for 16 h. The reaction products were analyzed using TLC as previously described. To determine specific activities, natural oligosaccharides containing terminal Ara $\beta$ -1,3 and Ara $\beta$ -1,2 and synthetic Ara $\beta$ - $\beta$ -pNP and Ara $\beta$ - $\beta$ -OMe were used. Ara $\beta$ -1,3-Ara $\beta$ Gal<sub>2</sub> (0.15 mM) and Ara $\beta$ -1,3-Ara $\beta$ <sub>4</sub> (0.15 mM) were incubated with 0.01  $\mu$ g/mL of BII3HypBA1-N $\Delta$ 379C $\Delta$ 761 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 40 °C for 20 min. Ara $\beta$ - $\beta$ -pNP (1.0 mM) was incubated with 0.1  $\mu$ g/mL of BII3HypBA1-N $\Delta$ 379C $\Delta$ 761 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 40 °C for 2 h. A final concentration of 0.35 mM Ara $\beta$ -1,2-Ara, 0.40 mM Ara $\beta$ -OMe, 0.11 mM Ara $\beta$ - $\beta$ -Hyp (*cis*), 0.11 mM Ara $\beta$ - $\beta$ -Hyp (*trans*), 0.13 mM Ara $\beta$ -1,2-Ara $\beta$ -Hyp, and 1.0 mM Ara $\beta$ -1,2-Ara $\beta$ -1,2-Ara $\beta$ -Hyp were incubated with 12  $\mu$ g/ml of the BII3HypBA1-N $\Delta$ 379C $\Delta$ 761 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) at 40 °C for 16 h. The reactions were terminated by adding 10  $\mu$ L of 5% TCA and subsequently analyzed using HPAEC-PAD. One unit of enzyme activity was defined as the amount of enzyme required to produce 1  $\mu$ mol L-arabinose per minute.

## Transglycosylation activities

Transglycosylation reactions were performed using Ara $\beta$ 1,3-AraGal<sub>2</sub> as the donor and the 1-alkanols as acceptors. The substrate was incubated with 0.30  $\mu$ g/mL of BII3HypBA1-N $\Delta$ 35C $\Delta$ 761 in 40  $\mu$ L of 50 mM sodium acetate buffer (pH 5.5) containing 5% methanol as the acceptor. After 1 h of incubation at 37 °C, the reaction was terminated by boiling. For ethanol and 1-propanol, the reactions were terminated after 3 h. TLC was used for confirmation of the reaction products as described above.

## Protein purification for crystallization

For crystallization, selenomethionine (SeMet)-labeled and native proteins were expressed in *E. coli* BL21-CodonPlus (DE3)-RP-X and BL21-CodonPlus (DE3)-RIL (Agilent Technologies, Santa Clara, CA, USA), respectively. For the expression of the SeMet-labeled protein, the transformants were grown in Se-Met core medium supplemented with 10 g/L D-glucose, 250 mg/L MgSO<sub>4</sub>•7H<sub>2</sub>O, 4.2 mg/L FeSO<sub>4</sub>•7H<sub>2</sub>O, 10 ml/L Kao and Michayluk Vitamin Solution (Sigma-Aldrich), and 25 mg/L seleno-L-methionine. For the expression of the native protein, lysogeny broth was used. Transformants were grown at 37 °C for 2 h in a medium



supplemented with 50 µg/mL chloramphenicol and 100 µg/mL ampicillin. Protein expression was induced by adding 0.1 mM isopropyl-β-D-thiogalactopyranoside and 0.5 mM ZnSO<sub>4</sub> to the medium, and cells were further cultured at 18 °C for 24 h. Cells were harvested via centrifugation, suspended in 20 mM Tris-HCl (pH 7.5) and 300 mM NaCl (buffer A), and later disrupted via sonication; the supernatant was then purified via serial column chromatography. Ni-immobilized metal affinity chromatography was conducted using cOmplete™ His-Tag Purification Resin (Sigma-Aldrich) with wash and elution steps of 20 mM and 300 mM imidazole in buffer A, respectively. The buffer of the eluted protein sample was exchanged with 20 mM Tris-HCl (pH 7.5) (buffer B) using Vivaspin Turbo 15 (50-kDa cut-off; Sartorius Stedim Biotech, Göttingen, Germany). The protein sample was applied to a HiTrap Q HP column (5 mL) (Cytiva, Marlborough, MA, USA) equilibrated with buffer B and eluted with a linear gradient of 0 to 500 mM NaCl. Then, the sample was concentrated using Vivaspin Turbo 15 MWCO 50,000, and the solution was changed to 20 mM Tris-HCl (pH 7.5) and 150 mM NaCl (buffer C). Subsequently, gel filtration chromatography was conducted using a HiLoad 16/60 Superdex 200 pg column (Cytiva) equilibrated with buffer C at a flow rate of 1 mL/min. The purified protein was concentrated again using Vivaspin Turbo 15, and the solution was changed to buffer B. Protein concentrations were determined using a BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA, USA) with bovine serum albumin as the standard.

## Crystallography

The protein crystals were grown at 20 °C using the hanging drop vapor-diffusion method by combining the protein solution with an equal volume of a reservoir solution. SeMet-labeled Bll3HypBA1-NΔ379CΔ761 was crystallized using a 17 mg/mL protein solution and a reservoir solution containing 12% PEG 8000, 0.2 M MgCl<sub>2</sub>, and 0.1 M Tris-HCl (pH 8.5). The native Bll3HypBA1-NΔ379CΔ761 protein was crystallized using a protein solution (10 mg/mL) and a reservoir solution containing 15% PEG 8000, 0.2 M MgCl<sub>2</sub>, and 0.1 M Tris-HCl (pH 8.5). The crystals were cryoprotected in a reservoir solution supplemented with 25% glycerol. The native Bll3HypBA1-NΔ379CΔ933 protein was crystallized using a protein solution (14 mg/mL) and a reservoir solution containing 45% MPD, 0.2 M ammonium acetate, and 0.1 M HEPES-NaOH (pH 7.5). The crystal was soaked in a reservoir solution supplemented with 100 mM L-arabinose before cryocooling, but no electron density for L-arabinose was detected in the crystal structure. The crystals were flash-cooled by dipping into liquid nitrogen. X-ray diffraction data were collected at 100 K on beamlines at the Photon Factory of the High Energy Accelerator Research

Organization (KEK, Tsukuba, Japan). Preliminary diffraction data were collected at SPring-8 (Hyogo, Japan). The datasets were processed using XDS (Kabsch 2010) and Aimless (Evans and Murshudov 2013). Phase determination and automated model building were performed for the data of SeMet-labeled crystal using the AutoSol pipeline of PHENIX (Adams et al. 2010). Manual model rebuilding and crystallographic refinement were performed using Coot (Emsley et al. 2010) and Refmac5 (Murshudov et al. 2011). Molecular graphic images were prepared using PyMOL (Schrödinger, LLC, New York, NY, USA).

## Site-directed mutagenesis and activity assay

Site-directed mutants of Bll3HypBA1 were constructed using the PrimeSTAR mutagenesis basal kit (TaKaRa Bio, Ohtsu, Japan) with the primers shown in Table S1 and the pET23d-Bll3HypBA1-NΔ379CΔ933 plasmid as the template. These mutant enzymes were expressed and purified using the same procedures for Bll3HypBA1-NΔ379CΔ933. We also constructed a C805A mutant; however, it did not express in *E. coli*. The activity of the wild-type enzyme (NΔ379CΔ933) and mutants was assayed via TLC using Araf-β1,3-Araf-α-OMe as the substrate. The assay solution consisted of 12.5 mM Araf-β1,3-Araf-α-OMe, 0.0125 mg/mL enzyme, 10 mM dithiothreitol, and 20 mM sodium acetate buffer (pH 4.5). After incubation at 37 °C for 30 min, the reaction was stopped by heating the sample at 95 °C for 10 min. Then, 2 µL of the sample was spotted on a TLC plate (TLC Silica gel 60 F<sub>254</sub>, Merck) and developed using a 2:1:1 solvent mixture (v/v/v) of ethyl acetate/acetic acid/water. The separated sugars were visualized by spraying and heating with phosphomolybdic acid in ethanol on the plates. For mutants at Zn-coordinating and catalytic residues (E694Q, E694A, E723Q, E723A, C725S, C725A, C804S, C805S, and C805A) and E694Q/C804S double mutant, we determined specific activities via HPLC using Araf-β1,3-ArafGal<sub>2</sub>-ABEE as described above.

## Results

### Sequence features of Bll3HypBA1

Bll3HypBA1 (BLLJ\_1848) contains a GH146 catalytic domain, a signal peptide, two laminin\_G\_3 (LamG) domains, which belong to the concanavalin A-like lectin/glucanase superfamily, three bacterial Ig-like domains, and a transmembrane domain (Fig. 2A). Furthermore, these features are shared with other cell surface anchoring glycosidases of *B. longum* subsp. *longum* JCM 1217, such as β-L-arabinobiosidase HypBA2 (BLLJ\_0212) (Fujita et al. 2011), GH43 α-L-arabinofuranosidases HypAA

(BLLJ\_0213), BI<sub>Ara</sub>F (BLLJ\_1854) (Fujita et al. 2019a), BI<sub>Ara</sub>F<sub>B</sub> (BLLJ\_1853), BI<sub>Ara</sub>F<sub>C</sub> (BLLJ\_1852) (Komeno et al. 2022), BI<sub>Ara</sub>F<sub>D</sub> (BLLJ\_1851) (Komeno et al. 2022), BI<sub>Ara</sub>F<sub>E</sub> (BLLJ\_1850) (Sasaki et al. 2022), GH43\_24 exo- $\beta$ -1,3-galactanase BI<sub>1,3</sub>Gal (BLLJ\_1840) (Fujita et al. 2014a), and GH30\_5 exo- $\beta$ -1,6-galactobiohydrolase BI<sub>1,6</sub>Gal (BLLJ\_1841) (Fujita et al. 2019a). The aa sequence of the catalytic domain of BI<sub>3</sub>HypBA1 exhibits 30% identity with the other GH146 paralog BI<sub>4</sub>HypBA1, a  $\beta$ -L-arabinofuranosidase for Ara<sub>f</sub>- $\beta$ 1-linked Hyp on the HRGP backbone (Ishiwata et al. 2023). In addition, BI<sub>3</sub>HypBA1 exhibits 31% and 28% identities with the other characterized GH146  $\beta$ -L-arabinofuranosidases BT0349 from *Bacteroides thetaiotaomicron* (Luis et al. 2018) and XeHypBA1 from *X. euvesicatoria* (Nakamura et al. 2018), respectively. Furthermore, the enzymes GH146 and GH127 are separated in the phylogenetic tree (Fig. 2B), and BI<sub>3</sub>HypBA1 shows only 23% and 19% identities with the GH127  $\beta$ -L-arabinofuranosidase paralogs BI<sub>1</sub>HypBA1 and BI<sub>2</sub>HypBA1 (BLLJ\_1826), respectively. The alignment of the GH127 and GH146 catalytic domains revealed that the Zn-coordinating residues and catalytic residues are conserved in BI<sub>3</sub>HypBA1 (Fig. S1).

### Production of recombinant BI<sub>3</sub>HypBA1 protein

The recombinant BI<sub>3</sub>HypBA1-N $\Delta$ 35C $\Delta$ 761 protein without the N-terminal signal peptide and the C-terminal domains (Fig. 2A) was detected as a soluble protein only at a low expression level under induction condition at 20 °C. The purified recombinant BI<sub>3</sub>HypBA1-N $\Delta$ 35C $\Delta$ 761 protein migrated as a single band on SDS-PAGE, which was in agreement with the theoretical molecular mass of 129,964 Da (Fig. S2). To obtain a clone suitable for X-ray crystallography and biochemical analysis, we constructed BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 and BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 933 without the LamG region. These clones were expressed as soluble proteins at 30 °C and purified as a single band in agreement with the theoretical molecular mass of 94,145 Da for BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 and 75,016 Da for BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 933 (Fig. S2). The estimated molecular masses by gel filtration chromatography were 91 kDa and 76 kDa for BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 and BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 933, respectively, suggesting that both protein constructs are monomeric in solution (data not shown).

### Substrate specificity and general properties of the recombinant BI<sub>3</sub>HypBA1 protein

The substrate specificity of BI<sub>3</sub>HypBA1 was compared with that of the other three paralogs of *B. longum* subsp. *longum* JCM 1217 using some natural polysaccharides as substrates.

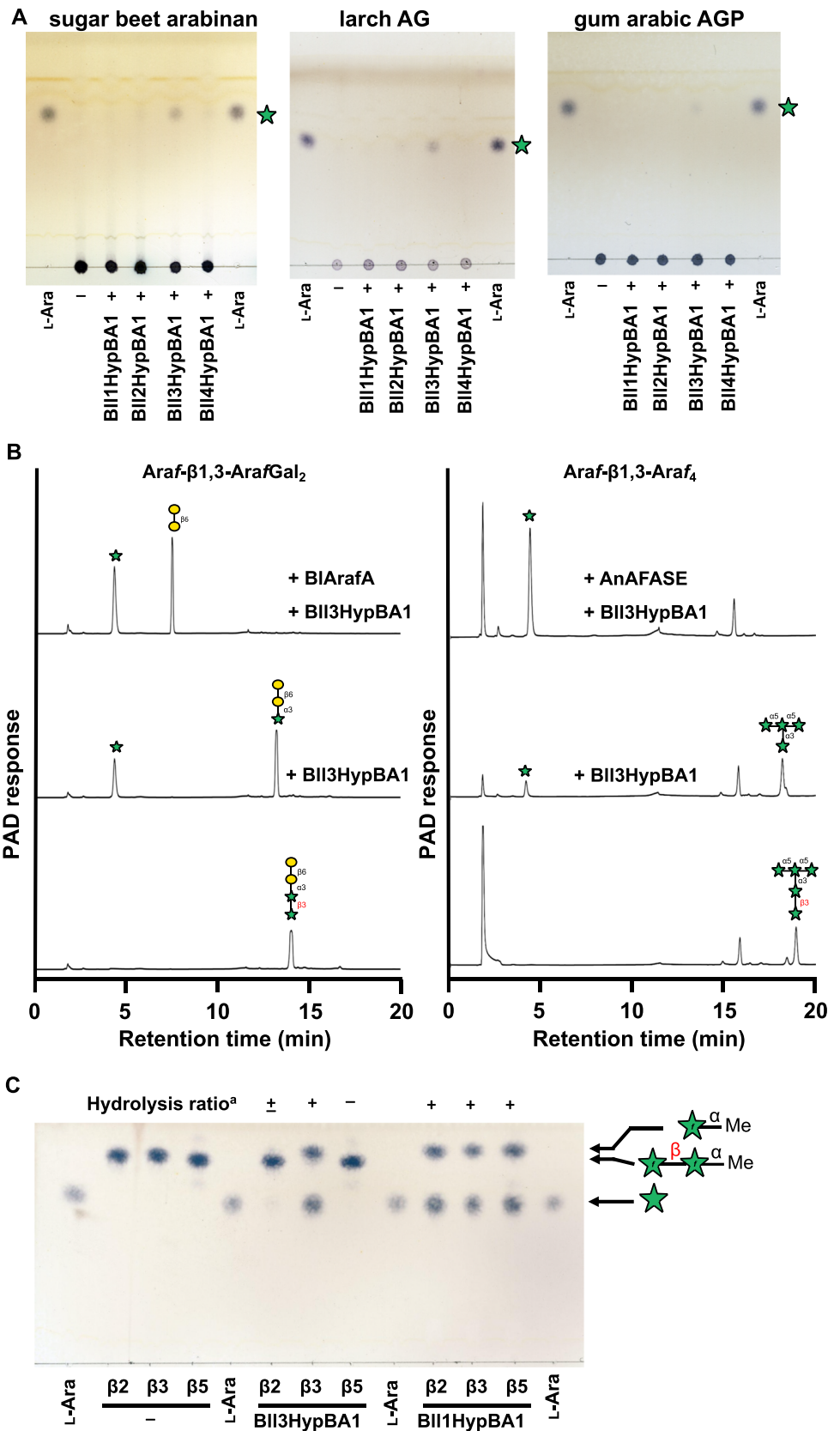
As shown in Fig. 3A, BI<sub>3</sub>HypBA1-N $\Delta$ 35C $\Delta$ 761 released L-arabinose from sugar beet arabinan, larch AG, and gum arabic AGP, whereas other paralogs in *B. longum* subsp. *longum* JCM 1217 could not. This result suggests that BI<sub>3</sub>HypBA1 has a specific ability to degrade the terminal  $\beta$ -L-Ara<sub>f</sub> moieties of these polysaccharides. The L-arabinose content after the overnight reaction of BI<sub>3</sub>HypBA1 was 0.32% for sugar beet arabinan, 0.39% for larch AG, and 0.20% for gum arabic AGP. This result indicates that a small amount of terminal  $\beta$ -L-Ara<sub>f</sub> residues is present in AGPs and arabinans. BI<sub>3</sub>HypBA1 also released L-arabinose from Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub> (Fig. 3B left). The mass of the degraded Ara<sub>f</sub>Gal<sub>2</sub> was  $m/z$  474.40 (calc.  $m/z$  474.16), which was a reduction in the mass of L-arabinose (MW 132) from Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub> ( $m/z$  606.48; calc.  $m/z$  606.20). Subsequently, the released Ara<sub>f</sub>Gal<sub>2</sub> was further hydrolyzed to  $\beta$ 1,6-Gal<sub>2</sub> by  $\alpha$ -L-arabinofuranosidase BI<sub>Ara</sub>F. BI<sub>3</sub>HypBA1 also released L-arabinose from Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub><sub>4</sub> ( $m/z$  678.77; calc.  $m/z$  678.22), and the hydrolysate Ara<sub>f</sub><sub>4</sub> was further degraded by *A. niger*  $\alpha$ -L-arabinofuranosidase AnAFASE (Fig. 3B, right). The specific activity of BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 for Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub> was two times higher than that for Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub><sub>4</sub> (Table 2). In a comparison of the specific activities of larch AG, sugar beet arabinan, and gum arabic AGP, BI<sub>3</sub>HypBA1-N $\Delta$ 35C $\Delta$ 761 with N-terminal LamG was 5.8-, 1.7-, and 1.1-fold higher than BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 without N-terminal LamG, respectively (Table S2). Alternatively, BI<sub>3</sub>HypBA1-N $\Delta$ 35C $\Delta$ 761 was 1.6-fold lower than BI<sub>3</sub>HypBA1-N $\Delta$ 379C $\Delta$ 761 for Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub>.

Furthermore, synthetic arabinosyl disaccharides containing terminal  $\beta$ 1,2-,  $\beta$ 1,3-, and  $\beta$ 1,5-Ara<sub>f</sub> structures were used for a detailed evaluation of the linkage specificity of this enzyme (Ishiwata et al. 2022a). BI<sub>3</sub>HypBA1 released L-arabinose from Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>- $\alpha$ -OMe and slightly from Ara<sub>f</sub>- $\beta$ 1,2-Ara<sub>f</sub>- $\alpha$ -OMe but not from Ara<sub>f</sub>- $\beta$ 1,5-Ara<sub>f</sub>- $\alpha$ -OMe (Fig. 3C). Conversely, BI<sub>1</sub>HypBA1 completely hydrolyzed these substrates. BI<sub>3</sub>HypBA1 also exhibited < 1% degradative activity for Ara<sub>f</sub>- $\beta$ -pNP, Ara<sub>f</sub>- $\beta$ -OMe, Ara<sub>f</sub>- $\beta$ -Hyp glycosides, and  $\beta$ -AOSs containing Ara<sub>f</sub>- $\beta$ 1,2-linkage compared with Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub> containing Ara<sub>f</sub>- $\beta$ 1,3-linkage (Table 2). These results showed that BI<sub>3</sub>HypBA1 is a  $\beta$ -L-arabinofuranosidase specific for Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub> structures.

The optimal temperature and pH for Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub>-ABEE were 50 °C and 5.5, respectively (Fig. S3). Transglycosylation products were detected via TLC when 5% methanol, ethanol, and 1-propanol were used as acceptors (Fig. S4). These R<sub>f</sub> values of TLC spots were similar to previous ones for BI<sub>1</sub>HypBA1 (Fujita et al. 2014b). This result indicates that BI<sub>3</sub>HypBA1 is the same anomer-retaining enzyme as other GH127/146 enzymes.

The rice anther AGP was predicted to contain Ara<sub>f</sub>- $\beta$ 1,3-Ara<sub>f</sub>Gal<sub>2</sub> side chains, as depicted in the schematic in Fig. 4's right column. To clarify the importance of BI<sub>3</sub>HypBA1 in

**Fig. 3** Substrate specificities of BII3HypBA1. **A** TLC analysis using sugar beet arabinan, larch AG, and gum arabic AGP as substrates. The substrates were incubated in the absence (–) or presence (+) of enzymes (BII1HypBA1, BII2HypBA1, BII3HypBA1-NΔ35CΔ761, or BII4HypBA1) at 37 °C for 16 h. **B** HPAEC-PAD analysis using Araf-β1,3-ArafGal<sub>2</sub> and Araf-β1,3-Araf<sub>4</sub> as substrates. The substrates were incubated with BII3HypBA1-NΔ379CΔ933 or BII1HypBA1 at 37 °C for 16 h. The hydrolysate was further incubated with α-L-arabinofuranosidases BI ArafA or AnAFASE. **C** TLC analysis using Araf-β1,2-Araf-α-OMe (β2), Araf-β1,3-Araf-α-OMe (β3), and Araf-β1,5-Araf-α-OMe (β5) in the absence (–) or presence of BII3HypBA1-NΔ379CΔ933 or BII1HypBA1 at 37 °C for 16 h. <sup>a</sup>The substrate was (+) completely hydrolyzed, (±) partially hydrolyzed, and (–) not hydrolyzed





**Table 2** Substrate specificity of Bll3HypBA1 toward oligosaccharides containing terminal  $\beta$ -Araf structure

Substrates	Conc. (mM)	Specific activity <sup>a</sup> (units/mg)	Relative activity <sup>b</sup> (%)
Araf- $\beta$ 1,3-ArafGal <sub>2</sub>	0.15	13	100
Araf- $\beta$ 1,3-Araf <sub>4</sub>	0.15	6.9	51
Araf- $\beta$ -pNP	1.0	0.11	0.82
Araf- $\beta$ -OMe	0.40	0.0027	0.021
Araf- $\beta$ -Hyp( <i>cis</i> )	0.11	0.00027	<0.01
Araf- $\beta$ -Hyp( <i>trans</i> )	0.11	0.0071	0.053
Araf- $\beta$ 1,2-Ara	0.35	0.011	0.079
Araf- $\beta$ 1,2-Araf- $\beta$ 1,2-Araf- $\beta$ -Hyp	1.0	0.00011	<0.01
Araf- $\beta$ 1,2-Araf- $\beta$ -Hyp	0.13	0.00043	<0.01

<sup>a</sup>Bll3HypBA1-N $\Delta$ 379C $\Delta$ 761 was used for the enzymatic reactions as described in Materials and Methods

<sup>b</sup>Relative activity was expressed as the percentage of the activity toward Araf- $\beta$ 1,3-ArafGal<sub>2</sub>

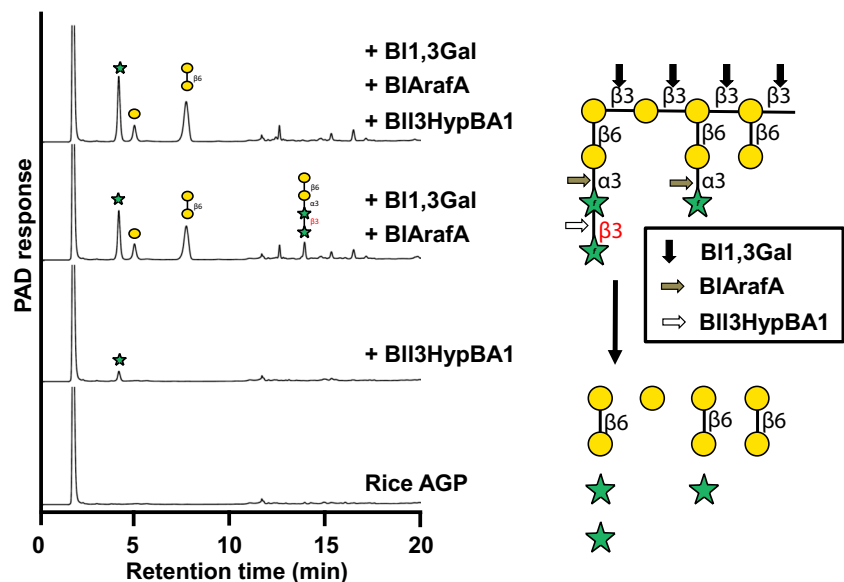
the degradation of AGPs, a combination reaction with AGP-degrading enzymes was performed. The reaction of Bll3HypBA1 alone produced a small amount of L-arabinose, whereas the reaction of Bll1,3Gal and BI ArafA without Bll3HypBA1 produced Araf- $\beta$ 1,3-ArafGal<sub>2</sub> in addition to L-arabinose, galactose, and  $\beta$ 1,6-Gal<sub>2</sub> (Fig. 4). In contrast, Araf- $\beta$ 1,3-ArafGal<sub>2</sub> was completely degraded to L-arabinose and  $\beta$ 1,6-Gal<sub>2</sub> in a combination reaction involving three enzymes. The releasing  $\beta$ 1,6-Gal<sub>2</sub> is probably degraded to galactose by intracellular GH42  $\beta$ -galactosidase (BLLJ\_0443) in *B. longum* subsp. *longum* JCM 1217 (Fujita et al. 2019b). These findings suggest that Bll3HypBA1 is an effective cell surface anchoring enzyme for the degradation of AGPs.

## Crystal structure of Bll3HypBA1

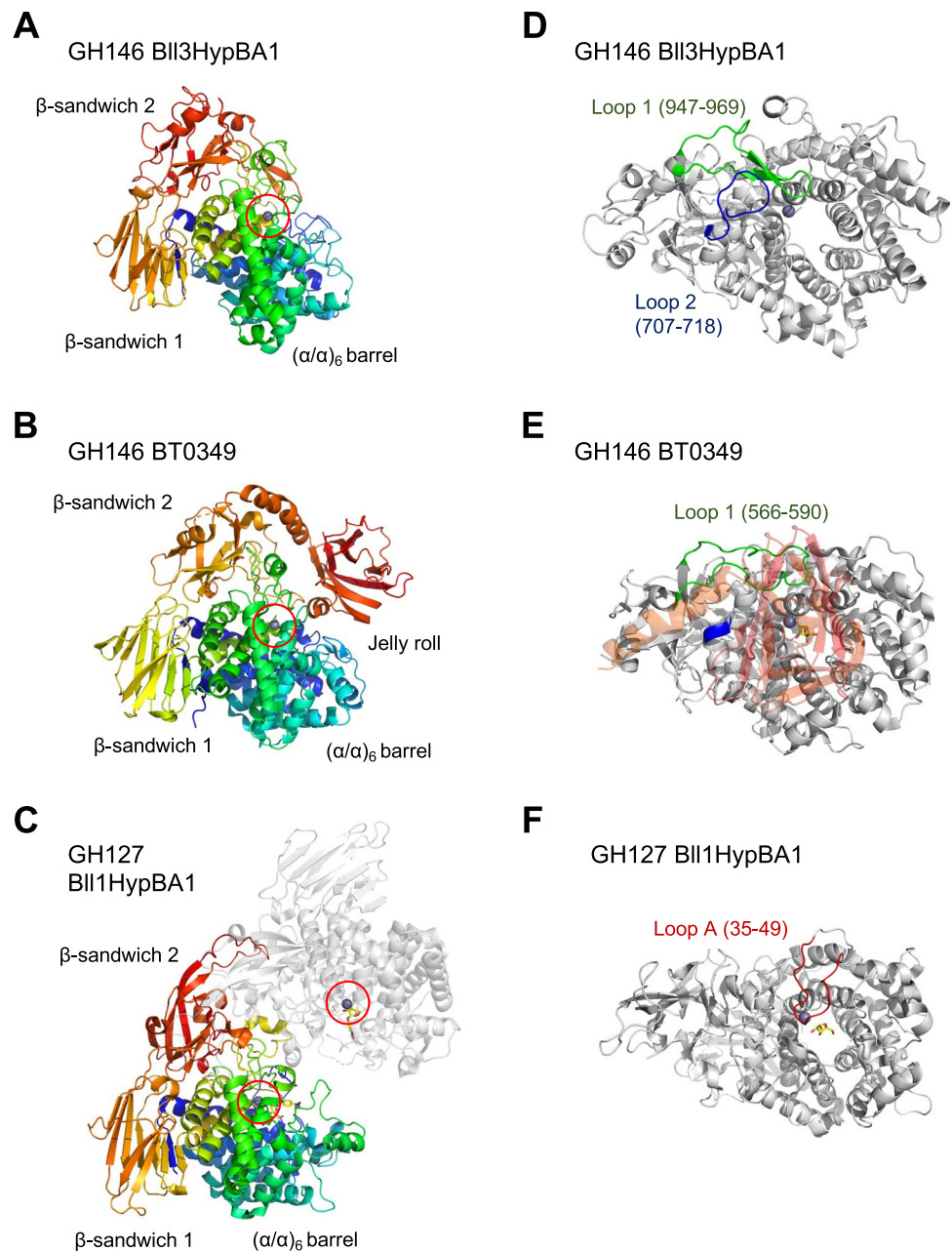
The crystal structure of Bll3HypBA1-N $\Delta$ 379C $\Delta$ 761 (aa: 380–1223) was determined at a resolution of 1.75 Å by the single-wavelength anomalous dispersion using a SeMet-derivative (Table S3). The N $\Delta$ 379C $\Delta$ 761 structure contained one molecule per asymmetric unit, and residues M379–V1051 were modeled. In the active site of the N $\Delta$ 379C $\Delta$ 761 structure, a Zn atom coordinated by E723, C725, C804, and C805 was clearly observed, and a Tris molecule was bound to the L-arabinose-binding site (sub-site – 1) (Fig. S5A and B). A crystallographic anomalous scattering analysis was performed to identify the metal atom (Table S3). A Bijvoet difference density map of the data collected at 1.280 Å wavelength revealed a prominent peak at the metal site, whereas a difference map of the data collected at 1.300 Å wavelength, which is above the absorption edge of Zn (1.2837 Å), showed a significant decrease in the anomalous scattering ability (Fig. S5C). This result confirmed that the metal ion bound at the active site is Zn<sup>2+</sup>.

Since the region following residue 1052 in the N $\Delta$ 379C $\Delta$ 761 structure was disordered, we also crystallized the Bll3HypBA1-N $\Delta$ 379C $\Delta$ 933 (aa: 380–1051) construct. The crystal structure of N $\Delta$ 379C $\Delta$ 933 was determined at a resolution of 1.70 Å (Table S3). The N $\Delta$ 379C $\Delta$ 933 structure contained two molecules per asymmetric unit, and residues from M379 to E1050 were modeled for both chains (A and B). The three molecules, including one in N $\Delta$ 379C $\Delta$ 761 and two in N $\Delta$ 379C $\Delta$ 933, had almost identical main chain structures, as the root mean square deviation (RMSD) values for the C $\alpha$  atoms between all chain pairs were less than 0.24 Å. The Zn-coordination site structure of N $\Delta$ 379C $\Delta$ 933 was also almost identical to that of N $\Delta$ 379C $\Delta$ 761 (Fig. S5D), but no Tris molecule was identified in the active site.

**Fig. 4** Stepwise hydrolysis of rice AGP by bifidobacterial AGP degradative enzymes. Rice AGP was incubated with a combination of Bll1,3Gal, BI ArafA, and Bll3HypBA1-N $\Delta$ 35C $\Delta$ 761. A schematic drawing is shown in the right column. The cleavage sites of enzymes: Bll1,3Gal (black arrow), BI ArafA (gray arrow), and Bll3HypBA1 (white arrow)



**Fig. 5** Overall and the catalytic domain structures of the GH146 and GH127 enzymes. **A–C** Domain structures of **A** GH146 Bll3HypBA1 and **B** GH146 BT0349 and dimer structure of **C** GH127 Bll1HypBA1. A ribbon presentation with a rainbow color (blue to red from the N- to C-terminus) is shown. Red circles indicate the Zn atoms at the active site (gray spheres). **D–F** Top views of the catalytic domain of **D** GH146 Bll3HypBA1, **E** GH146 BT0349, and **F** GH127 Bll1HypBA1. In **E**, an  $\alpha$ -helix linker and the C-terminal jelly roll domain are shown in transparent color (orange to red)



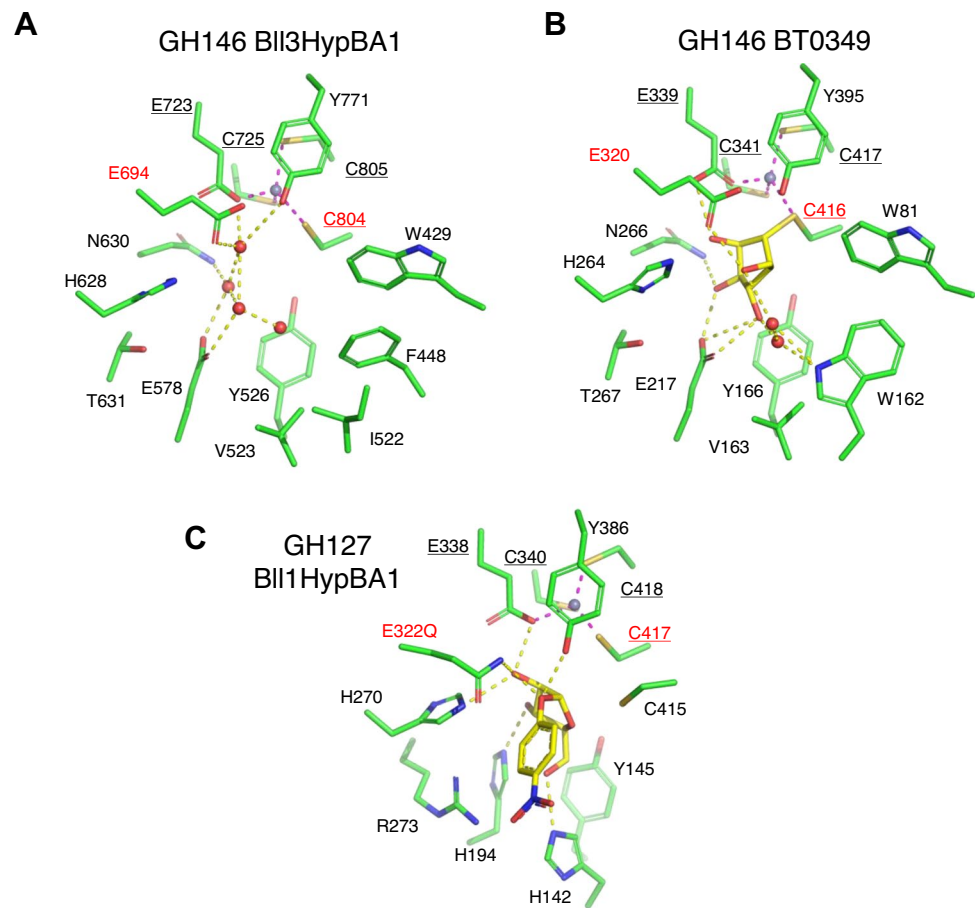
Therefore, we primarily focused on the chain A structure of N $\Delta$ 379C $\Delta$ 933.

The crystal structure of GH146 Bll3HypBA1 consists of three domains: catalytic ( $\alpha/\alpha$ )<sub>6</sub> barrel (aa: 379–820),  $\beta$ -sandwich 1 (aa: 821–920), and  $\beta$ -sandwich 2 (aa: 921–1051) (Fig. 5A). The GH146 family member BT0349 has an  $\alpha$ -helical linker and a jelly roll domain following the  $\beta$ -sandwich 2 domain (Fig. 5B). Although the C-terminal jelly roll domain of BT0349 is structurally similar to carbohydrate-binding module (CBM) family 35, the function of this domain remains unknown (Luis et al. 2018). While the GH146 enzymes (Bll3HypBA1 and BT0349) are monomeric in solution, GH127 Bll1HypBA1 forms a homodimer by

interactions between the  $\beta$ -sandwich 2 domains (Fig. 5C). Extended structural elements in the  $\beta$ -sandwich 2 domain of the GH146 enzymes (a loop with two  $\beta$ -strands in Bll3HypBA1 and the jelly roll domain in BT0349) interact with the catalytic domain (Fig. 5A and B).

Figure 5D, E, and F depict the top views of the catalytic domain of GH146 and GH127 enzymes. In Bll3HypBA1, loop 1 in  $\beta$ -sandwich 2 (aa: 947–969) partially covers the active site, whereas loop 2 (aa: 707–718) supports loop 1 (Fig. 5D). The C-terminal jelly roll domain of BT0349 shields its active site (Fig. 5E, transparent red). Loop 1 (aa: 566–590) is distant from the active site in BT0349, whereas loop 2 is absent. In GH127 Bll1HypBA1, the active site is

**Fig. 6** Active site structures of the GH146 and GH127 enzymes. **A** GH146 Bll3HypBA1, **B** GH146 BT0349, and **C** GH127 Bll1HypBA1. The label characters of the Zn-coordinating residues are underlined. The catalytic residues (nucleophile and acid/base catalyst) are indicated by labels of red characters



covered by a long loop A (aa: 35–49) in the  $(\alpha/\alpha)_6$  barrel domain (Fig. 5F) (Ito et al. 2014).

Figure 6 depicts the active site structures of the GH146 and GH127 enzymes. The Zn atom is coordinated by E723, C725, C804, and C805 in Bll3HypBA1 (Fig. 6A). The Zn-coordination structure of Bll3HypBA1 is identical to that of BT0349 and Bll1HypBA1 (Fig. 6B and C). R273, H194, H142, and Y145 form the substrate binding site of GH127 Bll1HypBA1 (Fig. 6C), whereas these residues are completely different in GH146 enzymes. F448 and I522 occupy the active site in Bll3HypBA1 (Fig. 6A), whereas W162 occupies this position in BT0349 (Fig. 6B). Therefore, these residues may determine the  $\beta$ -linkage specificity of Bll3HypBA1.

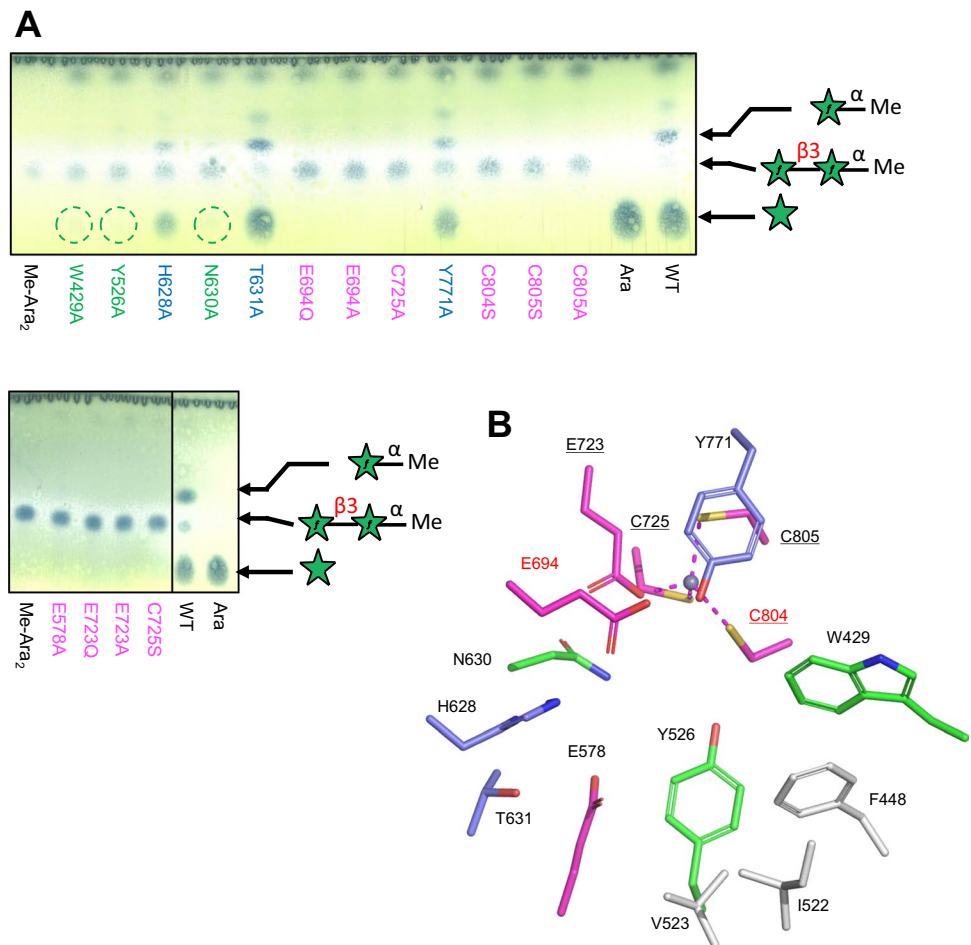
### Mutational analysis

Next, using site-directed mutagenesis, we determined the significance of active site residues on enzyme activity (Fig. 7). Substitutions were made at the Zn-coordinating residues (E723, C725, C804, and C805), catalytic residues (E694 and C804), and residues near the catalytic center (W429, Y526, E578, H628, N630, T631, and Y771). The activity of the mutants was analyzed via

TLC using Araf- $\beta$ 1,3-Araf- $\alpha$ -OMe as the substrate. The mutant enzymes at the Zn-coordinating and catalytic residues (E694Q, E694A, E723Q, E723A, C725S, C725A, C804S, C805S, and C805A) completely lost their activity (Fig. 7A). HPLC analysis also confirmed the importance of these residues (Table S4). For other active site residues, E578A also lost its activity, whereas W429A, Y526A, and N630A exhibited slight spots of L-arabinose release on TLC, indicating significantly diminished activity (Fig. 7A). H628A, T631A, and Y771A exhibited residual spots of the substrate Araf- $\beta$ 1,3-Araf- $\alpha$ -OMe but clear spots of L-arabinose after reaction at 37 °C for 30 min, indicating that these mutants retained their activity.

The result of the mutational analysis is shown in Fig. 7B by color codes. Residues without activity, residues with slight activity, and residues with significant activity in the mutant enzymes are depicted in magenta, green, and blue, respectively. The Zn-coordinating three Cys and one Glu residue and the acid/base catalyst residue were essential for activity. In the active site, E578 mutants lacked activity. W429 and E578 may participate in substrate recognition at subsite – 1 (Fig. 6A). N630 and Y526 play significant roles in the activity as well. H628, T631, and Y771 are involved in substrate recognition but are not required

**Fig. 7** Mutational analysis of Bll3HypBA1. **A** TLC analysis of the wild-type Bll3HypBA1- $\Delta$ 379C $\Delta$ 933 and its mutants using Araf- $\beta$ 1,3-Araf- $\alpha$ -OMe at 37 °C for 30 min. **B** Active site residues colored by the effects of mutation: magenta, no activity detected for mutants; green, weak activity detected by TLC; blue, activity detected; white, not examined



for activity. These nonessential residues may recognize the reducing end side of Araf- $\beta$ 1,3-Araf at subsite + 1.

## Discussion

Previously, we characterized AGP-degrading enzymes in *B. longum* subsp. *longum*. The degradative enzymes are encoded by an AGP degradation gene cluster containing GH43\_24 Bll1,3Gal (BLLJ\_1840) and GH30\_5 Bll1,6Gal (BLLJ\_1841) and a GH43 gene cluster containing five GH43  $\alpha$ -L-arabinofuranosidases. GH43\_22 BIArafA (BLLJ\_1854) acts on the  $\alpha$ 1,3-linked Araf of AGP (Fujita et al. 2019a), GH43\_22 BIArafB (BLLJ\_1853) acts on the  $\alpha$ 1,5-linked Araf of the arabinan backbone, and GH43\_27 BIArafC (BLLJ\_1852) acts on the  $\alpha$ 1,2- and  $\alpha$ 1,3-linked Araf on the arabinan side chains (Komeno et al. 2019). Recently, BIArafD (BLLJ\_1851) and BIArafE (BLLJ\_1850) with tandem GH43 domains have been characterized. The GH43\_UC (uncharacterized subfamily of GH43) domain of BIArafD acts on the  $\alpha$ 1,2-linked Araf of the  $\alpha$ 1,2- and  $\alpha$ 1,3-Araf doubly substituted arabinoxylan, whereas the GH43\_26 BIArafD domain acts on the

arabinan backbone (Komeno et al. 2022). In contrast, the GH43\_22 BIArafE acts on the  $\alpha$ 1,3-linked Araf of the  $\alpha$ 1,3- and  $\alpha$ 1,4-Araf doubly substituted gum arabic AGP, whereas GH43\_34 BIArafE acts on the remaining  $\alpha$ 1,4-linked Araf (Sasaki et al. 2022). A mixture of Bll1,3Gal, Bll1,6Gal, and BIArafA had a synergistic effect on the degradation of larch AG but only released 3.3% of the polysaccharide's sugar component (Fujita et al. 2019a). A GH127  $\beta$ -L-arabinofuranosidase, BT3674, in *Bacteroides thetaiotaomicron* showed a synergistic effect with exo- $\beta$ -1,3-galactanases for the degradation of larch AG (Cartmell et al. 2018). Notably, Bll3HypBA1 (BLLJ\_1848) is flanked by the AGP degradation gene cluster and GH43 gene cluster. In addition to Bll3HypBA1, all AGP degradation enzymes and GH43  $\alpha$ -L-arabinofuranosidases contain C-terminal membrane anchoring regions, indicating that these enzymes act synergistically to degrade AGPs and arabinans on the bifidobacterial cell surface. In contrast, the AGP degradation gene cluster is conserved in nearly all *B. longum* subsp. *longum* strains, whereas the Bll3HypBA1 ortholog (> 99% identity) was conserved in only 31 strains (5.2%) in 600 *B. longum* and *B. longum* subsp. *longum* strains in the NCBI genome database. In addition,



the orthologous proteins (> 78% identity) are not conserved in other bifidobacterial species except for strains of *B. aesculapii* and *B. primatum*, which were isolated from the feces of common marmosets and cotton-top tamarins, respectively. These findings indicate that the Bll3HypBA1 ortholog is not a universal gene in bifidobacteria.

Bll3HypBA1 contains two independent LamG domain modules. Upon comparing the reactivities of Bll3HypBA1-N $\Delta$ 35C $\Delta$ 761 with N-terminal LamG and Bll3HypBA1-N $\Delta$ 379C $\Delta$ 761 without LamG, the former were more reactive for polysaccharides than the latter (Table S2). Notably, the N-terminal LamG domain was conserved in two adjacent AGP degradative enzymes (Bll<sub>3</sub>Gal and Bll<sub>6</sub>Gal) and four GH43  $\alpha$ -L-arabinofuranosidases (BlArafB-E) in *B. longum* subsp. *longum* JCM 1217. CBMs in bifidobacteria are predicted to serve as substrate and/or hydrolysate docking stations (van den Broek et al. 2008). The N-terminal LamG domain may aid in the degradation of AGPs and arabinans for bifidobacterial cell surface anchoring enzymes.

In the present study, the crystal structure of Bll3HypBA1 was determined as the second three-dimensional structure of the GH146 member following BT0349. BT0349 has a C-terminal jelly roll domain that covers the active site, and this enzyme does not appear to bind polymer substrates (Fig. 5B). BT0349 is present in the PUL of rhamnogalacturonan-I (RG-I), which comprises multiple GHs and is anticipated to act on oligosaccharides cleaved by other GHs. In contrast, Bll3HypBA1 can act independently on rice AGP (Fig. 4). This is consistent with the structural characteristics of Bll3HypBA1, which contains no large structural element covering the active site. The Zn atom of Bll3HypBA1 is coordinated by Cys<sub>3</sub>-Glu residues, and GH127 and GH146 enzymes share this structural characteristic (Cartmell et al. 2018; Ito et al. 2014; Luis et al. 2018; Maruyama et al. 2022; McGregor et al. 2021). In our previous studies, we elucidated the reaction mechanism of GH127 Bll1HypBA1 using synthetic inhibitors that were specifically designed for cysteine glycosidase (Ishiwata et al. 2022b; Maruyama et al. 2022). It has been demonstrated that the coordination of cysteine to Zn<sup>2+</sup> is necessary for deglycosylation from the thioglycosyl intermediate, which is energetically unfavorable in the absence of metal coordination (McGregor et al. 2021). Consequently, this zinc coordination structure is shared by clan GH-P (GH127 and GH146) with cysteine as the catalytic nucleophilic residue. In addition, we performed a mutational analysis of the active site residues of Bll3HypBA1 and demonstrated the significance of these residues for the enzyme activity toward  $\beta$ 1,3-linked Araf disaccharide. In addition to Zn-coordinating and catalytic residues, E578 was demonstrated to be essential for activity. Several other active site residues have been shown to contribute significantly to the enzyme activity, providing a structural basis for the strict recognition toward the Araf- $\beta$ 1,3-linkages.

Although it is known that gum arabic AGP from *Acacia senegal* is partially substituted by Gal- $\alpha$ 1,3- or Araf- $\beta$ 1,3- at the Araf- $\alpha$ 1,3-terminal sugar (Sasaki et al. 2021), the presence of Araf- $\beta$ 1,3 substitution has never been reported. Bifidobacterial GH39 3-O- $\alpha$ -D-galactosyl- $\alpha$ -L-arabinofuranosidase has been previously characterized for the release of Gal- $\alpha$ 1,3-Araf disaccharide and 3-O- $\beta$ -L-arabinopyranosyl- $\alpha$ -L-arabinofuranosidase for the release of Araf- $\beta$ 1,3-Araf disaccharide on AGP (Sasaki et al. 2021, 2023). The presence of  $\beta$ -Araf in gum arabic AGP and larch AG has been demonstrated, but the linkage positions remain unclear (Cartmell et al. 2018). In addition to quinoa arabinan (Wefers et al. 2014) and rice AGP (Kawaguchi et al. 1996), we found Araf- $\beta$ 1,3 substitution in larch AG, gum arabic AGP, and sugar beet arabinan. These results indicated that AGPs and arabinans may contain terminal Araf- $\beta$ 1,3-linkages. Since partially substituted sugars prevent enzymatic degradation of these polysaccharides by bacteria, the presence of Bll3HypBA1 and orthologous enzymes is advantageous for some *Bifidobacterium* and *Bacteroides* species with AGP-degrading ability. Jones et al. found GH39  $\alpha$ -L-( $\beta$ -1,2)-arabinofuranobiosidase from rumen fungi; this enzyme releases Araf- $\beta$ 1,2-Araf disaccharide on sugar beet arabinan (Jones et al. 2017). To the best of our knowledge, Araf- $\beta$ 1,3-Araf disaccharide-releasing enzyme has not been discovered. Bll3HypBA1 is useful for glycan structure analysis, oligosaccharide preparation, and Araf- $\beta$ 1,3 content measurement due to its specificity for the Araf- $\beta$ 1,3-Araf structure.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00253-024-13014-8>.

**Acknowledgements** We thank Tomohide Shimotashiro (Kagoshima University, Japan) for the kind gifts of the young panicles of rice, Dr. Takatoshi Arakawa for helping experimental design, and the staff of Photon Factory and SPring-8 for the X-ray data collection.

**Author contribution** KF, AI, and SF conceived the project and designed the experiments. HT, PL, SM, MM, AD, KK, KT, YI, and AI conducted experiments. KF, AI, and SF drafted and edited the manuscript. All authors have reviewed the data and contributed to the completion of the manuscript.

**Funding** Open access funding provided by Kagoshima University. This work was supported in part by JSPS KAKENHI Grant-in-Aid for Scientific Research (24580144 to K.F.; JP19H00929 to A.I.; JP18K05345 to A.I.; 15H02443 to S.F., K.F., and A.I.; 19H00929 to S.F., K.F., and A.I.; and JP 22H02196 to Y.I.) and Grant-in-Aid for Specially Promoted Research (JP16H06290 to Y.I. and A.I.).

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethical approval** This article does not contain any studies with animals performed by any of the authors.



**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221. <https://doi.org/10.1107/S0907444909052925>
- Cartmell A, Muñoz-Muñoz J, Briggs JA, Ndeh DA, Lowe EC, Baslé 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>
- Das D, Maiti S, Maiti TK, Islam SS (2013) A new arabinoxylan from green leaves of *Litsea glutinosa* (Lauraceae): structural and biological studies. *Carbohydr Polym* 92:1243–1248. <https://doi.org/10.1016/j.carbpol.2012.10.052>
- Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>
- Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69:1204–1214. <https://doi.org/10.1107/S0907444913000061>
- Fujita K, Sakamoto S, Ono Y, Wakao M, Suda Y, Kitahara K, Suganuma T (2011) Molecular cloning and characterization of a  $\beta$ -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>
- Fujita K, Sakaguchi T, Sakamoto A, Shimokawa M, Kitahara K (2014a) *Bifidobacterium longum* subsp. *longum* exo- $\beta$ -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>
- Fujita K, Takashi Y, Obuchi E, Kitahara K, Suganuma T (2014b) Characterization of a novel  $\beta$ -L-arabinofuranosidase in *Bifidobacterium longum*: functional elucidation of a DUF1680 protein family member. *J Biol Chem* 289:5240–5249. <https://doi.org/10.1074/jbc.M111.248690>
- Fujita K, Sakamoto A, Kaneko S, Kotake T, Tsumuraya Y, Kitahara K (2019a) 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>
- Fujita K, Sasaki Y, Kitahara K (2019b) Degradation of plant arabinogalactan proteins by intestinal bacteria: characteristics and functions of the enzymes involved. *Appl Microbiol Biotechnol* 103:7451–7457. <https://doi.org/10.1007/s00253-019-10049-0>
- 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)
- Ishiwata A, Ito Y (2011) Synthesis of docosasaccharide arabinan motif of mycobacterial cell wall. *J Am Chem Soc* 133:2275–2291. <https://doi.org/10.1021/ja109932t>
- Ishiwata A, Munemura Y, Ito Y (2008) NAP ether mediated intramolecular aglycon delivery: a unified strategy for 1,2-*cis*-glycosylation. *Eur J Org Chem* 2008:4250–4263. <https://doi.org/10.1002/ejoc.200800249>
- Ishiwata A, Kaeothip S, Takeda Y, Ito Y (2014) Synthesis of the highly glycosylated hydrophilic motif of extensins. *Angew Chem Int Ed Engl* 53:9812–9816. <https://doi.org/10.1002/anie.201404904>
- Ishiwata A, Fujita K, Fushinobu S, Tanaka K, Ito Y (2022a) Synthesis of naturally occurring  $\beta$ -L-arabinofuranosyl-L-arabinofuranoside structures towards the substrate specificity evaluation of  $\beta$ -L-arabinofuranosidase. *Bioorg Med Chem* 68:116849. <https://doi.org/10.1016/j.bmc.2022.116849>
- Ishiwata A, Narita S, Kimura K, Tanaka K, Fujita K, Fushinobu S, Ito Y (2022b) Mechanism-based inhibition of GH127/146 cysteine glycosidases by stereospecifically functionalized L-arabinofuranosides. *Bioorg Med Chem* 75:117054. <https://doi.org/10.1016/j.bmc.2022.117054>
- Ishiwata A, Tsunomachi H, Kameyama K, Sophon K, Nakamura M, Kitahara K, Tanaka K, Ito Y, Fujita K (2023) Bifidobacterial GH146  $\beta$ -L-arabinofuranosidase (Bif4HypBA1) as the last enzyme for the complete removal of oligoarabinofuranosides from hydroxyproline-rich glycoproteins. *ChemBioChem* 24:e202200637. <https://doi.org/10.1002/cbic.202200637>
- Ito T, Saikawa K, Kim S, Fujita K, Ishiwata A, Kaeothip S, Arakawa T, Wakagi T, Beckham GT, Ito Y, Fushinobu S (2014) Crystal structure of glycoside hydrolase family 127  $\beta$ -L-arabinofuranosidase from *Bifidobacterium longum*. *Biochem Biophys Res Commun* 447:32–37. <https://doi.org/10.1016/j.bbrc.2014.03.096>
- Jones DR, Uddin MS, Gruninger RJ, Pham TTM, Thomas D, Boraston AB, Briggs J, Pluvinage 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>
- Kabsch W (2010) Xds *Acta Crystallogr D Biol Crystallogr* 66:125–132. <https://doi.org/10.1107/S0907444909047337>
- Kaeothip S, Ishiwata A, Ito T, Fushinobu S, Fujita K, Ito Y (2013a) Preparation of *p*-nitrophenyl  $\beta$ -L-arabinofuranoside as a substrate of  $\beta$ -L-arabinofuranosidase. *Carbohydr Res* 382:95–100. <https://doi.org/10.1016/j.carres.2013.10.005>
- Kaeothip S, Ishiwata A, Ito Y (2013b) Stereoselective synthesis of *Arabidopsis* CLAVATA3 (CLV3) glycopeptide, unique protein post-translational modifications of secreted peptide hormone in plant. *Org Biomol Chem* 11:5892–5907. <https://doi.org/10.1039/C3OB41212A>
- Kawaguchi K, Shibuya N, Ishii T (1996) A novel tetrasaccharide, with a structure similar to the terminal sequence of an arabinogalactan-protein, accumulates in rice anthers in a stage-specific manner. *Plant J* 9:777–785. <https://doi.org/10.1046/j.1365-313X.1996.9060777.x>
- Komono M, Hayamizu H, Fujita K, Ashida H (2019) Two novel  $\alpha$ -L-arabinofuranosidases from *Bifidobacterium longum* subsp. *longum* belonging to glycoside hydrolase family 43 cooperatively degrade

- arabinan. *Appl Environ Microbiol* 85:e02582-e2618. <https://doi.org/10.1128/AEM.02582-18>
- Komeno M, Yoshihara Y, Kawasaki J, Nabeshima W, Maeda K, Sasaki Y, Fujita K, Ashida H (2022) Two  $\alpha$ -L-arabinofuranosidases from *Bifidobacterium longum* subsp. *longum* are involved in arabinoxylan utilization. *Appl Microbiol Biotechnol* 106:1957–1965. <https://doi.org/10.1007/s00253-022-11845-x>
- Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schüchel J, Venditto I, Ralet M-C, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nat Microbiol* 3:210–219. <https://doi.org/10.1038/s41564-017-0079-1>
- Maruyama S, Sawano K, Amaki S, Suzuki T, Narita S, Kimura K, Arakawa T, Yamada C, Ito Y, Dohmae N, Fujita K, Ishiwata A, Fushinobu S (2022) Substrate complex structure, active site labeling and catalytic role of the zinc ion in cysteine glycosidase. *Glycobiology* 32:171–180. <https://doi.org/10.1093/glycob/cwab103>
- McGregor NGS, Coines J, Borlandelli V, Amaki S, Artola M, Nin-Hill A, Linzel D, Yamada C, Arakawa T, Ishiwata A, Ito Y, van der Marel GA, Codee JDC, Fushinobu S, Overkleef HS, Rovira C, Davies GJ (2021) Cysteine nucleophiles in glycosidase catalysis: application of a covalent  $\beta$ -L-arabinofuranosidase inhibitor. *Angew Chem Int Ed Engl* 60:5754–5758. <https://doi.org/10.1002/anie.202013920>
- Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) *REFMAC5* for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367. <https://doi.org/10.1107/S0907444911001314>
- Nakamura M, Yasukawa Y, Furusawa A, Fuchiwaki T, Honda T, Okamura Y, Fujita K, Iwai H (2018) Functional characterization of unique enzymes in *Xanthomonas euvesicatoria* related to degradation of arabinofurano-oligosaccharides on hydroxyproline-rich glycoproteins. *PLoS ONE* 13:e0201982. <https://doi.org/10.1371/journal.pone.0201982>
- Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffet F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neill MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet MC, Martens EC, Henrissat B, Gilbert HJ (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 544:65–70. <https://doi.org/10.1038/nature21725>
- Pellerin P, Doco T, Vidal S, Williams P, Brillouet JM, O'Neill MA (1996) Structural characterization of red wine rhamnogalacturonan II. *Carbohydr Res* 290:183–197
- 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)
- Sasaki Y, Horigome A, Odamaki T, Xiao JZ, Ishiwata A, Ito Y, Kitahara K, Fujita K (2021) Novel 3-O- $\alpha$ -D-galactosyl- $\alpha$ -L-arabinofuranosidase for the assimilation of gum arabic arabinogalactan protein in *Bifidobacterium longum* subsp. *longum*. *Appl Environ Microbiol* 87:e02690-e2720. <https://doi.org/10.1128/AEM.02690-20>
- Sasaki Y, Komeno M, Ishiwata A, Horigome A, Odamaki T, Xiao J-Z, Tanaka K, Ito Y, Kitahara K, Ashida H, Fujita K, Atomi H (2022) Mechanism of cooperative degradation of gum arabic arabinogalactan protein by *Bifidobacterium longum* surface enzymes. *Appl Environ Microbiol* 88:e02187-e2221. <https://doi.org/10.1128/aem.02187-21>
- Sasaki Y, Yanagita M, Hashiguchi M, Horigome A, Xiao J-Z, Odamaki T, Kitahara K, Fujita K (2023) Assimilation of arabinogalactan side chains with novel 3-O- $\beta$ -L-arabinopyranosyl- $\alpha$ -L-arabinofuranosidase in *Bifidobacterium pseudocatenulatum*. *Microbiome Res Rep* 2:12. <https://doi.org/10.20517/mrr.2023.08>
- 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)
- Tryfona T, Liang H-C, 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>
- van den Broek LA, Hinz SW, Beldman G, Vincken JP, Voragen AG (2008) *Bifidobacterium* carbohydrases—their role in breakdown and synthesis of (potential) prebiotics. *Mol Nutr Food Res* 52:146–163. <https://doi.org/10.1002/mnfr.200700121>
- Wang WT, LeDonne NC Jr, Ackerman B, Sweeley CC (1984) Structural characterization of oligosaccharides by high-performance liquid chromatography, fast-atom bombardment-mass spectrometry, and exoglycosidase digestion. *Anal Biochem* 141:366–381. [https://doi.org/10.1016/0003-2697\(84\)90057-5](https://doi.org/10.1016/0003-2697(84)90057-5)
- Wefers D, Tyl CE, Bunzel M (2014) Novel arabinan and galactan oligosaccharides from dicotyledonous plants. *Front Chem* 2:1–9. <https://doi.org/10.3389/fchem.2014.00100>
- Wefers D, Dong J, Abdel-Hamid AM, Paul HM, Pereira GV, Han Y, Dodd D, Baskaran R, Mayer B, Mackie RI, Cann I (2017) Enzymatic mechanism for arabinan degradation and transport in the thermophilic bacterium *Caldanaerobius polysaccharolyticus*. *Appl Environ Microbiol* 83:e00794-e817. <https://doi.org/10.1128/AEM.00794-17>
- York WS, Kumar Kolli VS, Orlando R, Albersheim P, Darvill AG (1996) The structures of arabinoxyloglucans produced by solanaceous plants. *Carbohydr Res* 285:99–128. [https://doi.org/10.1016/S0008-6215\(96\)90176-7](https://doi.org/10.1016/S0008-6215(96)90176-7)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.