



HAL
open science

Functional analysis of family GH36 α -galactosidases from *Ruminococcus gnavus* E1: insights into the metabolism of a plant oligosaccharide by a human gut symbiont

M. Cervera-Tison, L. E. Tailford, C. Fuell, L. Bruel, G. Sulzenbacher, Bernard Henrissat, Jean-Guy Berrin, M. Fons, T. Giardina, N. Juge

► To cite this version:

M. Cervera-Tison, L. E. Tailford, C. Fuell, L. Bruel, G. Sulzenbacher, et al.. Functional analysis of family GH36 α -galactosidases from *Ruminococcus gnavus* E1: insights into the metabolism of a plant oligosaccharide by a human gut symbiont. *Applied and Environmental Microbiology*, 2012, 78 (21), pp.7720-7732. 10.1128/AEM.01350-12 . hal-01268089

HAL Id: hal-01268089

<https://hal.science/hal-01268089>

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Functional Analysis of Family GH36 α -Galactosidases from *Ruminococcus gnavus* E1: Insights into the Metabolism of a Plant Oligosaccharide by a Human Gut Symbiont

M. Cervera-Tison, L. E. Tailford, C. Fuell, L. Bruel, G. Sulzenbacher, B. Henrissat, J. G. Berrin, M. Fons, T. Giardina and N. Juge

Appl. Environ. Microbiol. 2012, 78(21):7720. DOI: 10.1128/AEM.01350-12.

Published Ahead of Print 24 August 2012.

Updated information and services can be found at:
<http://aem.asm.org/content/78/21/7720>

SUPPLEMENTAL MATERIAL

These include:

[Supplemental material](#)

REFERENCES

This article cites 60 articles, 25 of which can be accessed free at: <http://aem.asm.org/content/78/21/7720#ref-list-1>

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](#)

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Functional Analysis of Family GH36 α -Galactosidases from *Ruminococcus gnavus* E1: Insights into the Metabolism of a Plant Oligosaccharide by a Human Gut Symbiont

M. Cervera-Tison,^{a,b} L. E. Tailford,^a C. Fuell,^a L. Bruel,^b G. Sulzenbacher,^c B. Henrissat,^c J. G. Berrin,^d M. Fons,^b T. Giardina,^b and N. Juge^a

Institute of Food Research, The Gut Health and Food Safety Institute Strategic Programme, Norwich Research Park, Norwich, United Kingdom^a; Campus Scientifique de Saint Jérôme, Aix-Marseille Université, ISM2/Biosciences UMR CNRS 7313, Marseille, France^b; Architecture et Fonction des Macromolécules Biologiques UMR CNRS 7257, Université Aix-Marseille, Marseille, France^c; and INRA, UMR1163, Biotechnologie des Champignons Filamenteux, INRA, ESIL, Marseille, France^d

Ruminococcus gnavus belongs to the 57 most common species present in 90% of individuals. Previously, we identified an α -galactosidase (Aga1) belonging to glycoside hydrolase (GH) family 36 from *R. gnavus* E1 (M. Aguilera, H. Rakotoarivonina, A. Brutus, T. Giardina, G. Simon, and M. Fons, Res. Microbiol. 163:14–21, 2012). Here, we identified a novel GH36-encoding gene from the same strain and termed it *aga2*. Although *aga1* showed a very simple genetic organization, *aga2* is part of an operon of unique structure, including genes putatively encoding a regulator, a GH13, two phosphotransferase system (PTS) sequences, and a GH32, probably involved in extracellular and intracellular sucrose assimilation. The 727-amino-acid (aa) deduced Aga2 protein shares approximately 45% identity with Aga1. Both Aga1 and Aga2 expressed in *Escherichia coli* showed strict specificity for α -linked galactose. Both enzymes were active on natural substrates such as melibiose, raffinose, and stachyose. Aga1 and Aga2 occurred as homotetramers in solution, as shown by analytical ultracentrifugation. Modeling of Aga1 and Aga2 identified key amino acids which may be involved in substrate specificity and stabilization of the α -linked galactoside substrates within the active site. Furthermore, Aga1 and Aga2 were both able to perform transglycosylation reactions with α -(1,6) regioselectivity, leading to the formation of product structures up to [Hex]₁₂ and [Hex]₈, respectively. We suggest that Aga1 and Aga2 play essential roles in the metabolism of dietary oligosaccharides and could be used for the design of galacto-oligosaccharide (GOS) prebiotics, known to selectively modulate the beneficial gut microbiota.

The human gut is colonized by a complex, diverse, and dynamic community of microbes that continuously interact with the host (30). The majority belongs to only four bacterial divisions, *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria*, whereas other minor taxonomic divisions are quite diverse (19, 43, 62). Several ecological studies have shown that microbial symbionts have adapted to maximize metabolic access to a wide variety of dietary and host-derived carbohydrates (glycans), and competition for these nutrients is considered a major factor shaping the structure-function of the microbiota (33). Recently, a metagenomic analysis of gut microbial communities in humans proposed three predominant variants, or “enterotypes,” dominated by *Bacteroides*, *Prevotella*, and *Ruminococcus* (3). A controlled-feeding study showed that enterotype partitioning associates with long-term diets (61). Furthermore, the ability to selectively use prebiotics carbohydrates, ranging from oligosaccharides to polysaccharides, provides a competitive advantage over other bacteria in this ecosystem (28). These studies highlight the importance of understanding precisely how nutrient metabolism serves to maintain a symbiotic relationship between gut bacteria and the host. The genomes of gut bacteria encode a wide array of carbohydrate-active enzymes (CAZymes) that degrade dietary complex carbohydrates that cannot be hydrolyzed by host enzymes and ferment these otherwise indigestible glycans into short-chain fatty acids that are in turn used by the host (34). These enzymes are emerging both as key players to define molecular mechanisms of gut adaptation and as potentially exploitable biological catalysts.

Ruminococcus gnavus is a Gram-positive anaerobic bacterium, belonging to the *Firmicutes* division, *Clostridia* class and XIVa

cluster, and *Lachnospiraceae* family (35). A recent molecular inventory revealed that *R. gnavus* is widely distributed among individuals and is represented in the 57 most common species present in $\geq 90\%$ of individuals and can thus be considered a model organism to study bacterial adaptation to the gastrointestinal tract (42). *R. gnavus* E1, a strain isolated from a healthy adult, is known to grow on melibiose [α -D-Galp-(1 \rightarrow 6)-D-Glcp] and raffinose [α -D-Galp-(1 \rightarrow 6)- α -D-Glcp-(1 \leftrightarrow 2) β -D-Fru] as a sole carbon and energy source (1). Raffinose (Raf) and the related disaccharide sucrose (Suc) [α -D-Glcp-(1 \leftrightarrow 2) β -D-Fru] are the most abundant soluble carbohydrates in plant tissues (55). α -D-Galactosidases (EC 3.2.1.22) are exo-acting glycoside hydrolases (GH) that cleave α -linked galactose residues from carbohydrates commonly found in legumes and seeds. These enzymes occur widely in microorganisms, plants, and animals. However, no α -(1,6)-galactosidase activity is encoded by the human intestine mucosa, and α -galactosides are exclusively fermented by microbial enzymes. Based on their sequence similarities, α -galactosidases have been classified into families GH4, GH27, GH36, GH57, GH97, and GH110 of the carbohydrate active enzyme (CAZy) database (www.cazy.org)

Received 30 April 2012 Accepted 21 August 2012

Published ahead of print 24 August 2012

Address correspondence to N. Juge, nathalie.juge@ifr.ac.uk

Supplemental material for this article may be found at <http://aem.asm.org/>.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01350-12

(14). Two putative GH36 α -galactosidase genes, *aga1* and *agaSK*, were previously identified in the *R. gnavus* E1 genome (1, 9). AgaSK was recently characterized as a bifunctional enzyme composed of a GH36 α -galactosidase and a kinase domain able to phosphorylate sucrose provided by raffinose hydrolysis, highlighting a putative novel glycolytic pathway in bacteria for sucrose assimilation. The crystal structure of the AgaSK galactosidase domain was determined in the apo form and in complex with the product (9). Here, we identified a novel GH36 α -galactosidase termed Aga2 with strict specificity for α -linked galactose and carried out a thorough comparative analysis of enzymatic and structural properties of Aga1 and Aga2. Analysis of the α -galactosidase clusters and enzyme properties shed new light on the raffinose and melibiose utilization pathways in *R. gnavus* E1.

MATERIALS AND METHODS

Materials, plasmids, and strains. *R. gnavus* E1 was originally isolated from the feces of a healthy human (44). *Escherichia coli* strain DH5 α was used for DNA manipulation, and strain BL21(DE3) pLys was used for protein expression. High-purity salt-free oligonucleotides were from Sigma (St. Louis, MO), pOPINF expression vector was from the Oxford Protein Production Facility, and restriction endonucleases were from New England BioLabs (Ipswich, MA) and Invitrogen (Life Technologies, New York, NY). DNA-modifying enzymes were from Promega (Madison, WI) and Sigma (St. Louis, MO). Galactose (Gal), glucose (Glc), melibiose (Mel), raffinose (Raf), stachyose (Sta), locust bean gum (LBG), guar gum (GG), *para*-nitrophenyl- α -D-galactopyranoside (pNPGal), *ortho*-nitrophenyl- β -D-galactopyranoside (oNPGal), *para*-nitrophenyl- α -D-glucopyranoside (pNPGlc), and *para*-nitrophenyl- α -D-xylopyranoside (pNPXyl) were from Sigma.

Media and growth conditions. *R. gnavus* E1 was grown in an anaerobic cabinet at 37°C in brain heart infusion (BHI; Difco Laboratories, Detroit, MI) broth supplemented with 5 g · liter⁻¹ of yeast extract (Difco Laboratories) and 5 mg · liter⁻¹ of hemin (Sigma). *E. coli* strains were grown at 37°C with shaking at 200 rpm in Luria-Bertani (LB) medium supplemented with ampicillin (Amp) or carbenicillin (Carb) (50 μ g · ml⁻¹). Solid medium was obtained by adding agarose (15 g · liter⁻¹) to each medium.

Cecal content collection. Fisher axenic rats came from the ANAXEM platform (INRA, Jouy-en-Josas, France) and were reared as previously described (15). Three Fisher axenic rats were inoculated with around 10⁹ cells of the E1 strain (0.5 ml of fresh culture) by the intragastric route. After 1 week, individual fecal samples were collected and bacterial counts were estimated. Then, the animals were sacrificed, and the cecal contents were collected. Animal experiments were performed according to the guidelines of the French Ethics Committee.

Genomic DNA extraction. For the isolation of *R. gnavus* E1-chromosomal DNA, cells from a 2-ml overnight culture were harvested by centrifugation (10,000 \times g, 5 min, 4°C). The cell pellet was resuspended in 200 μ l of TES buffer (10 mM Tris, 1 mM EDTA, sucrose [0.1 mM]; pH 8.0) supplemented with lysozyme (20 mg · ml⁻¹) and incubated for 15 min at 37°C. Sodium dodecyl sulfate (SDS) (250 μ l of a 20% solution) was added, and the cell suspension was incubated for 10 min at 50°C until a clear lysate was obtained. After addition of 250 μ l of phenol and centrifugation (10,000 \times g, 5 min, 4°C), the aqueous phase was treated with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) and subsequently with chloroform-isoamyl alcohol only (24:1). After precipitation with ice-cold ethanol (70%) and incubation for 20 min at -20°C, the genomic DNA was resuspended in 100 μ l of TE buffer (10 mM Tris, 1 mM EDTA; pH 8.0)–RNase A (20 μ g · ml⁻¹). The solution was incubated for 15 min at 37°C and the DNA stored at -20°C.

Sequence analysis of *R. gnavus* E1 α -galactosidases. Sequencing of the genome of *R. gnavus* E1 was performed by Genoscope (Evry, France). The genomic organization of *R. gnavus* E1 *aga2* was determined using the

MaGe platform (Magnifying Genomes, Microbial Genome Annotation System; Genoscope, Laboratoire de Génomique Comparative) (56) at www.genoscope.cns.fr/aggc/website/. Putative transcriptional terminators were predicted *in silico* using the RNAfold program (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>). Amino acid sequences were acquired from the CAZy (Carbohydrate-Active Enzymes) database (www.cazy.org) or from the NCBI website (www.ncbi.nlm.nih.gov/pubmed/). Protein and nucleotide sequence homologies were searched using the blastp and blastn software, respectively (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Multalign (<http://multalin.toulouse.inra.fr/multalin/>) was used for the multiple alignment of protein sequences, and the alignments were submitted as Fasta files to Geneiousbasic v.4.6.4 software for the design of phylogenetic trees.

***aga2* cluster organization.** RNA extraction was carried out from aliquots (200 mg) of frozen cecal content from rats monoassociated with *R. gnavus* E1 as previously described (18). After purification using an RNeasy kit (Qiagen) and treatment with DNase I (Invitrogen, Carlsbad, CA), total RNAs (150 ng) were reverse transcribed using a SuperScript VILO cDNA synthesis kit from Invitrogen following the manufacturer's instructions. In order to check if the different open reading frames (ORFs) constituting the *aga2* cluster belonged to the same transcriptional unit, PCR amplifications were carried out using primer pairs designed to amplify each intergenic region (*sucR* and *scrA1*, *scrA1* and *malE1*, *malE1* and *aga2*, *aga2* and *scrA2*, and *scrA2* and *sacA*) as listed in Table S1 in the supplemental material. The PCR fragments were then separated by electrophoresis on agarose gel (1% [wt/vol]).

***aga1* and *aga2* cloning.** The α -galactosidase genes were amplified by PCR from *R. gnavus* E1 genomic DNA with 1.25 U of DNA polymerase (Prime star Hot Start DNA polymerase; TaKaRa) and 0.2 mM deoxynucleoside triphosphate (dNTP). In each PCR, primer pair Aga1IFFFor and Rev and primer pair Aga2IFFFor and Rev for *aga1* and *aga2*, respectively (see Table S1 in the supplemental material), were used at a final concentration of 20 pmol in 50- μ l volumes for 30 cycles of denaturation (1 min, 98°C), annealing (1.5 min, 48°C), and extension (3 min, 60°C) in a Mastercycler gradient thermocycler (Eppendorf, Hamburg, Germany). The PCR products were checked by agarose gel analysis and purified using a Wizard SV Gel and PCR Clean-Up System (Promega, France). The PCR products were combined with the pOPINF expression vector, following the manufacturer's instructions (In-Fusion; BD-Clontech). The recombinant vectors were checked by restriction analysis and double-stranded DNA sequencing (Beckman Coulter Genomics, Takeley, United Kingdom).

Expression and purification of recombinant proteins. *E. coli* strain BL21 (*pLys*) was transformed with recombinant vectors, and bacterial clones were grown at 37°C in Amp- or Carb-supplemented LB media with shaking at 200 rpm. Expression of *aga1* and *aga2* was induced with 2 mM isopropyl- β -D-thiogalactopyranoside (IPTG; Sigma) when the optical density (OD) reached 0.6. After induction at 37°C for 4 h, the cells were formed into pellets by centrifugation at 8,500 rpm for 30 min at 4°C and stored at -80°C until use. Bacterial lysis was performed by mechanical cell disruption. The cell pellet was resuspended in binding buffer (20 mM sodium phosphate buffer [pH 7.4], 0.5 M NaCl, 40 mM imidazole) with protease inhibitors (SigmaFAST Protease Inhibitor Cocktail Tablets EDTA-Free; Sigma) at 5 ml · g⁻¹ of the cell pellet and was processed twice in a cell disruptor (TS Series Constant Systems disrupter; CellD, Sauveterre, France) using the One-Shot system at 139 kPa. Cells debris was separated from the soluble fraction by centrifugation at 4°C and 10,000 \times g for 20 min. The supernatant was loaded onto a 5-ml HiTrap nickel-nitrilotriacetic acid (Ni-NTA) column at a flow rate of 5 ml · min⁻¹ (GE Healthcare, Fairfield, CT) and eluted with a gradient of elution buffer (20 mM sodium phosphate buffer [pH 7.4], 0.5 M NaCl, 500 mM imidazole) at a flow rate of 5 ml · min⁻¹. Recombinant Aga1 and Aga2 were purified by immobilized metal affinity chromatography (IMAC) using an Akta express system (GE Healthcare). All buffers contained protease inhibitors (SigmaFAST Protease Inhibitor Cocktail Tablets EDTA-Free).

Protein assays and electrophoresis. Protein concentrations were determined by spectrophotometry at 280 nm using extinction coefficients of $\epsilon = 136,030 \text{ M}^{-1} \cdot \text{cm}^{-1}$ and $\epsilon = 118,610 \text{ M}^{-1} \cdot \text{cm}^{-1}$ for Aga1 and Aga2, respectively, as determined by ProtParam (www.expasy.ch/tools/protparam.html). SDS-PAGE was performed using a 10% (wt/vol) polyacrylamide gel (Bio-Rad, Marnes-la-Coquette, France) and a Bio-Rad Precision Protein Plus electrophoresis calibration kit or using a NuPage Novex Bis-Tris gel (Invitrogen) (4% to 12%) and a Promega broad-range-electrophoresis calibration kit. Native isoelectric focusing (IEF) was carried out using a Bio-Rad vertical system or a Model 111 Mini IEF Cell horizontal system with the Bio-Rad IEF standards for Aga1 or Aga2, respectively, and ampholyte carriers in a pH range of 3.0 to 8.0 (Sigma). Proteins were visualized by the use of Coomassie blue stain, a Novex Colloidal blue stain kit (Invitrogen), or IEF gel staining solution (Bio-Rad).

AUC. Analytical ultracentrifugation (AUC) experiments were performed at the University of East Anglia (Norwich, United Kingdom) using a Beckman XLI analytical ultracentrifuge equipped with absorbance optics. The partial specific volume of each of the recombinant protein was calculated from the amino acid sequence using SEDNTERP (www.rasmb.bbri.org/). Aga1 was tested alone or in an equimolar mixture with Glc, Gal, or Mel in the presence of 50 mM HEPES (pH 7.4); Aga2 was tested in the presence of 50 mM HEPES (pH 7.0). The proteins were diluted to an A_{280} value of 0.5. Sedimentation equilibrium experiments were performed at 20°C and 9,000 rpm. The concentration profiles during AUC were measured at 280 nm. Scans were recorded every 4 h to determine when protein samples had reached equilibrium in the centrifuge. At equilibrium, 10 scans were recorded for each sample. The program Ultrascan II (B. Demeler, The University of Texas Health Science Center, San Antonio, TX) was used to fit the obtained sedimentation equilibrium profiles.

α -Galactosidase activity assays. α -Galactosidase activity was measured using the synthetic substrate pNPGal. The enzyme (57 nM for Aga1 and 117 nM for Aga2) was incubated with 2 mM pNPGal in McIlvaine's buffer (citric acid [0.1 M], Na_2HPO_4 [0.2 M], pH 6.0 for Aga1 and pH 5.5 for Aga2) in microtiter wells (200 μl), and the increase in absorbance at 405 nm was monitored for 10 min at 37°C in a microplate reader (KRL test [Spiral patent]; Kirial International). When appropriate, the assay was proportionally scaled up to 1 ml. One unit of enzyme activity was defined as the amount of protein that released 1 μmol of pNP/min at 37°C and pH 6.0 or pH 5.5. The extinction coefficients for the pNP under those conditions were $1,582.72 \text{ M}^{-1} \cdot \text{cm}^{-1}$ at pH 6.0 and $577.07 \text{ M}^{-1} \cdot \text{cm}^{-1}$ at pH 5.5. The optimal pH was determined on pNPGal (20 mM) in McIlvaine's buffer in a pH range of 3.0 to 8.0. The optimal temperature was determined at temperatures ranging from 20 to 70°C. For determination of the apparent Michaelis-Menten constants, the initial velocities of the enzymes were measured at 37°C in McIlvaine's buffer at pH 6.0 and pH 5.5 for Aga1 and Aga2, respectively, with pNPGal concentrations ranging from 2.5 mM to 100 mM. The kinetic parameters were estimated using nonlinear regression with GraphPad Prism software (GraphPad Prism version 3.00 for Windows 95; GraphPad Software, San Diego, CA). The sensitivity of the α -galactosidases to metals was assayed using pNPGal at 37°C and pH 6.0 and pH 5.5 for Aga1 and Aga2, respectively, in the absence or presence of metal ions and EDTA. Metal salts were dissolved in water to final concentrations ranging from 10 to 20 mM. In a microtiter plate, McIlvaine's buffer at pH 6.0 or pH 5.5 was mixed with Aga1 (57 nM) or Aga2 (117 nM), with or without prior addition of 1 mM EDTA (pH 8.6), and incubated for 10 min at 37°C. Metal-containing solutions (3 mM) were added in each well, and the mixture was incubated at 37°C for 10 min. The reaction was initiated by the addition of 2 mM pNPGal (total volume of 200 μl), and the increase in absorbance was monitored by spectrophotometry at 410 nm over a 5-min period. Inhibition assays of the α -galactosidases were carried out using pNPGal at 37°C and pH 6.0 and pH 5.5 for Aga1 and Aga2, respectively. The enzymes (Aga1 at 57 nM and Aga2 at 117 nM) were incubated for 30 min at 37°C with 10 μM miglitol, 1-deoxy-

jirimycin (DNJ), and deoxygalactonojirimycin (Gal-DNJ). The inhibition constant (K_i) was determined by incubating 57 nM or 117 nM Aga1 or Aga2, respectively, with various concentrations of Gal-DNJ (0.01 to 10 μM for Aga1 and 0 to 15 μM for Aga2) at different pNPGal concentrations (0.25 to 10 mM) at 37°C for 30 min.

In order to assess the substrate specificity of Aga1 and Aga2, the recombinant enzymes were tested against *ortho*-nitrophenyl- β -D-galactopyranoside (oNPGal), *para*-nitrophenyl- α -D-glucopyranoside (pNPGlc), and *para*-nitrophenyl- α -D-xylopyranoside (pNPXyl) under the conditions reported above for pNPGal. Activity assays on plant-based substrates, Mel, Raf, Sta, LBG, and GG, were performed using a Lactose/D-Galactose (rapid) kit (Megazyme). Briefly, Aga1 (114 nM) and Aga2 (234 nM) were incubated with 12 mM Mel, Raf, and Sta or 5 $\text{mg} \cdot \text{ml}^{-1}$ of LBG or GG in McIlvaine's buffer at pH 6.0 or pH 5.5, respectively (200 μl), and incubated for various times (0 to 10 min for Mel, Raf, and Sta; 0 to 300 min for LBG and GG) at 37°C. The reaction was terminated by boiling for 5 min. The reaction mixtures were centrifuged for 3 min at 14,500 rpm and cooled. The amount of Gal released was calculated relative to a Gal standard curve (0 to 10 μg). One unit of enzyme activity was defined as the amount of protein that released 1 μmol of Gal/min at 37°C and pH 6.0 or pH 5.5. The hydrolysis of Raf by both enzymes was also analyzed by high-performance anion exchange chromatography-pulsed amperometric detection (HPAEC-PAD) using these conditions.

Transglycosylation reactions. Self-condensation reactions were carried out by incubating Aga1 (57 nM) with Mel at concentrations of 11.5%, 30.5%, 49.5%, and 68.5% (wt/vol). The mixture was incubated for various time periods up to 150 h at 37°C. Assays were stopped by rapid freezing on dry ice. The transglycosylation reactions were carried out using the pNPGal donor at 40 mM with various acceptors: Gal, Glc, mannose (Man), xylose (Xyl), Mel, lactose (Lac), Suc, and Raf at 400 mM. Donor and acceptors were preincubated for 10 min at 37°C. The reaction was initiated by the addition of Aga1 or Aga2 (57 nM for both) and the mixture incubated at 37°C for 25.0 h or 36.5 h, respectively. Reactions were stopped by rapid freezing or by boiling for 20 min (and the reaction mixtures were then stored frozen). Transglycosylation products were further analyzed by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF)/TOF-mass spectrometry (TOF-MS), gas chromatography-mass spectrometry (GC-MS), and HPAEC-PAD.

HPAEC-PAD. The hydrolysis or transglycosylation reactions were carried out using appropriate enzyme and substrate concentrations in McIlvaine's buffer at pH 6.0 and pH 5.5 for Aga1 and Aga2, respectively. Samples were defrosted and appropriately diluted prior injection onto an HPAEC-PAD system. Reaction products were analyzed by HPAEC-PAD (ICS3000; Dionex, Camberley, United Kingdom) equipped with a guard column, a CarboPac PA-100 analytical column (250 by 4 mm), and a CarboPac PA1 analytical column (250 by 4 mm) in series. Elution (0.5 $\text{ml} \cdot \text{min}^{-1}$) was carried out using a 68-min gradient program from 96% A (10 mM NaOH) and 4% B (100 mM NaOH) to 80% A and 20% B, followed by a two-step regeneration program from 80% A and 20% B to 100% B and from 100% B to 100% C (200 mM NaOH).

CD. Circular dichroism (CD) spectra were recorded using a JASCO J-700 spectropolarimeter under the following conditions: 20 nm/min scan speed, bandwidth of 1 nm, response time of 2 s, 5 points/nm, and 4 accumulations. Far-UV spectra (260 to 180 nm) were recorded in a 0.1-mm- or 0.5-mm-path-length cell. The spectropolarimeter was calibrated using camphorsulfonic acid. Analysis was carried out using the SELCON and CONTIN algorithms via the DichroWeb interface (<http://dichroweb.cryst.bbk.ac.uk>). Far-UV CD was represented as molar CD, based on the average molecular mass of an amino acid of $113 \text{ g} \cdot \text{mol}^{-1}$. Molar CD was calculated according to the following equation: (average molecular weight $\times 30 \cdot 10^{-6}$) / [concentration (in milligrams per milliliter) \times path length (in centimeters)].

Homology modeling. Homology models for Aga1 and Aga2 were generated with the program COOT (21) using the crystal structure of the α -galactosidase domain of AgaSK, AgaSK-tru (Protein Data Bank [PDB] acces-

sion no. 2YFN), as the template. Briefly, amino acids of the AgaSK-tru model were replaced automatically by residues of the target sequence, short insertions and deletions were taken care of manually, and the stereochemistry was subsequently idealized. A 14-residue extension at the C terminus of Aga1 with respect to AgaSK-tru was not modeled. The quality of the models was assessed by the QMEAN server (6) (<http://swissmodel.expasy.org/qmean/cgi/index.cgi>). Models of Aga1 and Aga2 tetramers were produced using the AgaSK-tru tetrameric assembly as the template. Interactions of residues of the active site of Aga1 and Aga2 with raffinose derived from a superposition with the structure of *Saccharomyces cerevisiae* α -galactosidase 1, Mel1, in complex with raffinose (PDB accession no. 3LRM) (23), were defined with a hydrogen bonding cutoff of 3.25 Å. Figures were drawn with the program PyMOL (Molecular Graphics System, version 1.5.0.1; Schrödinger, LLC) (www.pymol.org/).

MALDI-TOF/TOF-MS. The oligosaccharides were permethylated according to the method described in reference 41 and analyzed on an UltraFlex MALDI-TOF/TOF mass spectrometer (MS) [Bruker Daltonics (United Kingdom) Ltd., Coventry, United Kingdom] in the positive-ion mode using a nitrogen laser ($\lambda = 337$ nm). Samples were cocrystallized at 1:1 on a stainless steel target with a saturated solution of 2,5-dihydroxybenzoic acid in 30% acetonitrile–0.1% trifluoroacetic acid (TFA).

GC-MS. For monosaccharide analysis, the oligosaccharides were hydrolyzed with trifluoroacetic acid (TFA) and derivatized into deuterated alditol acetates (AA) as previously described (41). Linkage analysis was carried out by derivatizing the permethylated oligosaccharides into deuterated partially methylated alditol acetates (PMAA) using the same method as for the AA. Both the AA and the PMAA were then analyzed using a Thermo Trace MS^{plus} GC-MS system (Thermo Fisher Scientific Inc., Waltham, MA) with Xcalibur software. The monosaccharide derivatives were separated using a ZB-5MS column (Phenomenex, Macclesfield, United Kingdom) (30 m by 0.25 mm; 0.25 μ m pore size) with helium as the carrier gas at 1 ml min⁻¹. The injection of 1 μ l of sample was made at 110°C, and the reaction was performed for 2 min at 110°C, followed by an increase to 320°C at 6°C min⁻¹ and 10 min at 320°C. The instrument was used in split mode with 15 ml min⁻¹ at 200°C. MS data were obtained by using the instrument in EI mode with a scan time of 0.4 s for a mass range of 50 to 700 nm. The GC-MS data were analyzed using ACD/SpecManager version 10.02 (Advanced Chemistry Development Inc., Toronto, Canada).

Nucleotide sequence accession numbers. The gene sequences for the RUGNE_61259, RUGNE_61256, RUGNE_61257, RUGNE_61261, RUGNE_61260, and RUGNE_61258 loci in the *aga2* operon have been deposited in the GenBank database under accession numbers FO203360, FO203357, FO203358, FO203359, FO203361, and FO203362, respectively.

RESULTS

Genetic organization and amino acid sequence analysis of the *aga2* cluster. In order to gain novel insights into raffinose and meliobiose (Mel) utilization by *R. gnavus* E1 and the potential link with the sucrose metabolism pathway, we searched for other potential α -galactosidase-encoding genes in the genome of *R. gnavus* E1. No GH4, GH27, GH57, GH97, or GH110 sequences were identified. However, in addition to *aga1* and *agaSK*, we found one candidate open reading frame (ORF) containing a potential GH36 domain and designated it *aga2*. This gene encodes a putative 727-amino-acid (aa) polypeptide containing the following putative sequences and domains: an N-terminal domain (aa 1 to 302), a GH36 α -galactosidase catalytic domain (aa 303 to 634), and a C-terminal domain (aa 635 to 727). *Aga2* showed 45% and 40% identities in amino acid sequences with *Aga1* (NCBI ACL13770) and *AgaSK* (NCBI CCA61959), respectively. We used the MaGe platform to analyze the genetic environment of the *aga2* gene, showing that *aga2* was part of a 9,836-kb cluster with predicted

functions in transcriptional regulation, glycoside import, and hydrolysis (Fig. 1A). The first ORF, *sucR*, encodes a putative transcriptional regulator of the LacI-GalR family, which is often found associated with sucrose utilization genes (45). The primary structure of *SucR* showed 30.7% identity with that of the catabolite control protein A (CcpA) and 29.1% identity with that of the sucrose operon repressor (ScrR) from *Staphylococcus xylosum* (27, 31). Two genes, *scrA1* and *scrA2*, showed significant homologies with genes encoding sucrose-specific phosphotransferase system (PTS) permeases (enzyme II [EII] complexes). *ScrA1* showed 40% amino acid identity with enzyme II^{Suc} of the PTS (*ScrA*) from *Streptococcus mutans* (50) and 37.5% identity with the sucrose-specific enzyme II (EII) BCA component (*ScrA*) from *Pediococcus pentosaceus* ATCC 27745 (39), whereas *ScrA2* displayed 52.5% amino acid identity with the sucrose-specific EIIBC component (*ScrA*) from *S. xylosum* (57) and 52.8% identity with the sucrose-specific EIIBC component (*ScrA*) from *Vibrio alginolyticus* (8). The essential phosphorylation site (37) and the active site turn (26) on the EIIB component could be identified in both *ScrA1* and *ScrA2*, while the histidine protein (HPr) interaction site, the phosphorylation site (59), the active site (11), and the glycerol kinase (GK) interaction site (22, 46) were identified in the EIIC component of *ScrA1* (not shown). Located between *scrA1* and *aga2*, the *malE1* gene codes for a putative GH13 enzyme with around 55% and 55.5% identity with oligo-1,6-glucosidases (MalL) from *Bacillus thermoglucosidasius* and *Bacillus halodurans* (EC 3.2.1.10) (52, 60), respectively, and 56% identity with a glucan-1,6- α -glucosidase (DexB) from *Streptococcus pneumoniae* (EC 3.2.1.70) (54). Downstream from *scrA2*, the *sacA* gene codes for a putative GH32 enzyme sharing 37.5% identity with a sucrose-6-phosphate hydrolase (*SacA*) (EC 3.2.1.26) from *B. subtilis* (24) and 38.8% identity with a β -fructofuranosidase or raffinose invertase (*RafD*) (EC 3.2.1.26) from *E. coli* (4). To get more insight into the organization of the *aga2* cluster, reverse transcriptase PCR (RT-PCR) amplification was carried out using total RNAs extracted from the cecum of rats monoassociated with *R. gnavus* E1 using primers designed to cover each intergenic region in the *aga2* cluster. Each amplicon gave a size of between 500 and 600 bp, indicating that, *in vivo*, all genes belonged to the same transcriptional unit (Fig. 1B).

A phylogenetic analysis of the *R. gnavus* E1 GH36 α -galactosidase protein sequences showed that they all fall into the *Firmicutes* cluster (data not shown) and that *Aga1* and *AgaSK* associate closely with the two putative α -galactosidases from *R. gnavus* ATCC 29149 (Fig. 2). Analysis of synteny of the two strains showed that *R. gnavus* ATCC 29149 lacks *malE1*-, *aga2*-, and *scrA2*-like genes, whereas ORFs showing significant identity with the other genes of the *aga2* cluster seem to be conserved (see Fig. S1 in the supplemental material). There was no noticeable difference in percent G+C (G+C%) between the fragment harboring *treC*, *aga2*, and *scrA2* and the other genes of the *aga2* cluster or the rest of the chromosome (average G+C content in *R. gnavus* E1 is 42.8% and in *R. gnavus* ATCC 29149 is 42.9%).

Alignment of the sequences of *R. gnavus* E1 α -galactosidases with functionally characterized α -galactosidases from the GH36 family, for which structural information is available, led to the classification of *Aga1* and *Aga2* in subgroup I of the GH36 family based on amino acid homology (not shown) (25). Moreover, the presence in *Aga1* and *Aga2* of a G-x-x-L-x-x-x-G motif found in

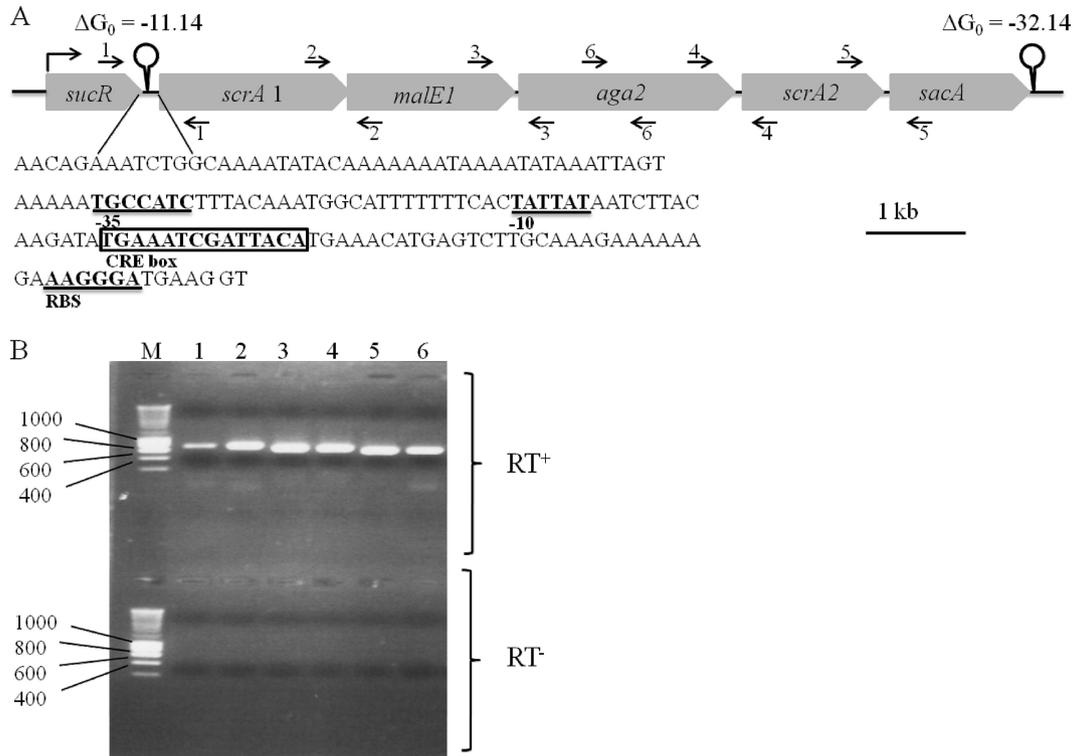


FIG 1 Transcriptional organization of the *aga2* operon in *R. gnavus* E1. (A) Schematic diagram of the *aga2* cluster. *sucR* (RUGNEv3_61261) is annotated as a CcpA-like regulator, *scrA1* (RUGNEv3_61260) as a sucrose-specific Glc-like PTS EIIBC, *malE1* (RUGNEv3_61259) as a GH13 α -amylase/trehalose_{tre}C, *aga2* (RUGNEv3_61258) as a GH36 α -galactosidase, *scrA2* (RUGNEv3_61257) as a sucrose-specific PTS EIIBC, and *sacA* (RUGNEv3_61256) as a GH32 sucrose-6-phosphate hydrolase. The general promoter of the *aga2* cluster, as determined by RT-PCR, is marked by an arrow. Circles above thick vertical lines indicate potential stem-loop structures that might act as rho-independent transcriptional terminators. The free energy of the thermodynamic ensemble is given on top, expressed as kcal \cdot mol⁻¹. The inset shows the DNA sequence of the catabolite-repression-responding promoter located upstream of *PTS1*. The putative -35 and -10 regions and ribosome-binding site (RBS) are underlined. The catabolite-responsive element (CRE) box is indicated with a black box. (B) Agarose gel electrophoresis of PCR products obtained following RT-PCR of total RNA extracted from gnotobiotic rats associated with *R. gnavus* E1 was performed using primers spanning the intergenic regions of the *aga2* cluster (RT⁺). In addition, one pair of primers was designed to amplify an internal region of the gene encoding Aga2, as a positive control. The negative controls (RT⁻) were without reverse transcriptase; the absence of PCR fragments confirmed that the products were amplified from RNA and not from potential contamination by genomic DNA. Lanes 1, 2, 3, 4, 5, and 6 correspond to the *sucR-scrA1*, *scrA1-malE1*, *malE1-aga2*, *aga2-scrA2*, and *scrA2-sacA* intergenic regions and an internal region of *aga2*, respectively. The positions of the primers are shown in panel A. M, DNA ladder size marker (with increments indicated in base pairs).

97% of the subgroup I sequences (25) strongly suggests that Aga1 and Aga2 associate into tetrameric assemblies.

Cloning, expression, and characterization of Aga1 and Aga2 from *R. gnavus* E1. In order to further assess the proposed role for GH36 α -galactosidases in *R. gnavus* E1 carbohydrate metabolism, Aga2 and Aga1 were heterologously expressed in *E. coli* and the recombinant enzymes biochemically characterized. The 2,184-bp *aga2* gene encodes a protein of a predicted molecular mass of 83.5 kDa and a theoretical pI of 5.52. The *aga1* gene was previously identified (1); it encodes a protein of 743 aa with a predicted molecular mass of 85.2 kDa and a theoretical pI of 5.47. The genes encoding Aga1 and Aga2 were amplified from the genomic DNA of *R. gnavus* E1 using specific primers (see Table S1 in the supplemental material) and cloned into the pOPINF vector using In-Fusion (IF) technology (7). Recombinant Aga1 and Aga2 were produced in *E. coli* BL21 as N-terminal (His)₆-tagged proteins. After purification by IMAC, a major band of 85 kDa was observed on SDS-PAGE corresponding to each recombinant protein, in agreement with their predicted molecular masses (see Fig. S2A in the supplemental material). IEF revealed a major band of pI 6.5 for recombinant Aga1 and a band of pI 6.2 for recombinant Aga2 (see

Fig. S2B in the supplemental material). The circular dichroism spectra obtained for each recombinant protein were similar and in agreement with the secondary-structure predictions based on the amino acid sequences, suggesting a correct folding of the recombinant enzymes (not shown). In the absence of a crystal structure, homology models for Aga1 and Aga2, based on the crystal structure of the α -galactosidase domain of AgaSK, AgaSK-tru (PDB accession no. 2YFN) (9), were produced. Model quality raw scores as determined by the model quality evaluation server QMEAN (6) were 0.713 and 0.738 for the models of Aga1 and Aga2, respectively, and root mean square deviation (RMSD) values for overlays with the crystal structure of AgaSK-tru were 0.207 Å for Aga1 and 0.262 Å for Aga2. The monomeric structures of the Aga1 and Aga2 models show the typical GH36 organization in three domains: the N-terminal domain composed of β -sheets, the catalytic domain presenting the characteristic (α/β)₈ fold of the clan D enzymes, and the C-terminal domain constituted of an antiparallel β -sheet (Fig. 3A). AgaSK has been observed to form a tetramer both in solution and in the crystal structure (9). Aga1 and Aga2 also exhibit multimeric behavior in solution, as shown by analytical ultracentrifugation (AUC). In each sample, there was little

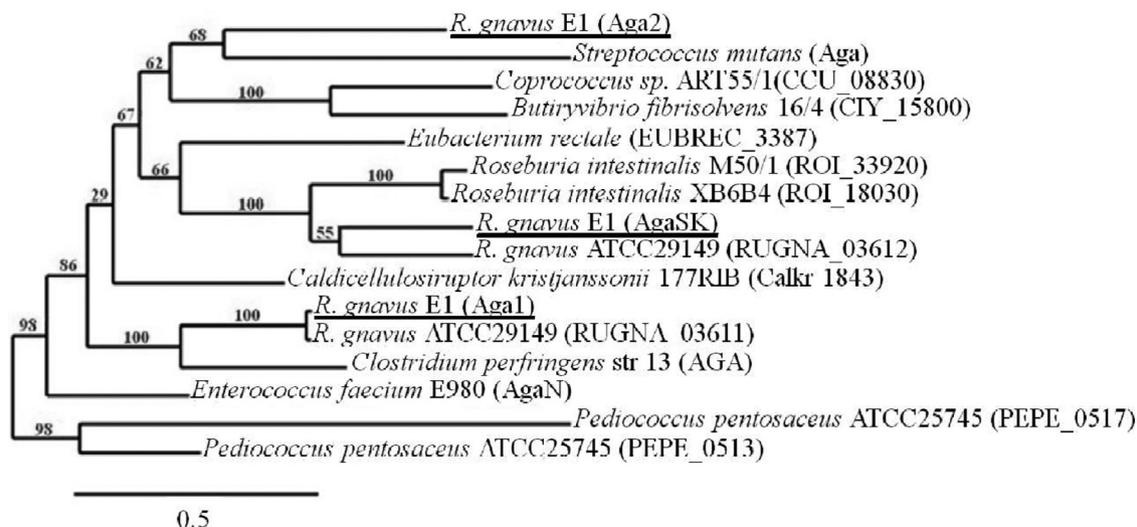


FIG 2 Unrooted phylogenetic tree of GH36 α -galactosidases from *Firmicutes*. *R. gnavus* E1 α -galactosidases Aga1, Aga2, and AgaSK are underlined. Percent bootstrap values are indicated. str, strain.

evidence of heterogeneity, confirming the stability of the enzymes (see the top panel of Fig. S3 in the supplemental material). Aga1 and Aga2 gave values of 335 kDa and 314 kDa, respectively (see Fig. S3A and B in the supplemental material), corresponding to 4 monomers. Moreover, addition of Gal, Glc, or Mel to the Aga1 samples did not alter the densitometric profile (data not shown). This, together with the bioinformatics analysis (see above), confirmed the tetrameric nature of the recombinant enzymes, and models of tetrameric assemblies of Aga1 and Aga2 were thus generated using the crystal structure of the AgaSK-tru tetramer as a guideline (Fig. 3B).

The general enzymatic properties of Aga1 and Aga2 were determined using *p*NPGal as the substrate. Aga1 showed over 70% activity in a pH range of pH 5.0 to pH 7.0, with an optimum pH of 6.0 (Fig. 4A). More than 50% of the activity was observed between 55 and 85°C, with a maximum seen at 75°C (Fig. 4B). Recombinant Aga2 conserved more than 50% of the activity in a range from pH 4.5 to pH 6.0, with a maximum at pH 5.5 (Fig. 4A). The enzyme displayed an optimal temperature of 40°C, and enzyme activity remained above 50% between 30°C and 45°C (Fig. 4B). The Ca^{2+} , Na^+ , Ni^{2+} , Mg^{2+} , and Zn^{2+} cations did not affect enzyme activity, whereas Cu^{2+} reduced the activity of both enzymes to 50% and Ag^+ and Hg^{2+} completely inactivated the enzymes (Table S2), suggesting the presence of a sulfhydryl (or thiol) group near the catalytic site. Furthermore, EDTA had no effect on the enzymatic activities of Aga1 and Aga2 at pH 5.5 to 6 in the presence or absence of metal ions, suggesting that metal cofactors are not required for α -galactosidase activity, although the possibility of the presence of tightly bound ions in the active site may not be excluded under these conditions.

The kinetic parameters of recombinant Aga1 and Aga2 were determined at 37°C and pH 6.0 and pH 5.5, respectively, using *p*NPGal as the substrate. Recombinant Aga1 displayed a higher catalytic efficiency (k_{cat}/K_m of 943.96) than recombinant Aga2 (k_{cat}/K_m of 129.61) (Table 1). Aga1 catalyzed the hydrolysis of the substrate with a rate (k_{cat}) of 196.06 s^{-1} and an apparent affinity constant (K_m) of 0.21 mM compared to 84.2 s^{-1} and 0.40 mM, respectively, for Aga2. The Hill equation was applied in order to

assess whether another substrate binding site was present in each enzyme (with V_{max} fixed to the Michaelis-Menten V_{max}). Using this method, no enzymatic cooperation or additional substrate binding sites could be identified ($n = \sim 1$) in either of the Aga1 and Aga2 recombinant forms (Table 1).

Substrate specificity of Aga1 and Aga2. In order to assess the substrate specificity of Aga1 and Aga2, the recombinant enzymes were tested against *o*NPGal, *p*NPGlc, and *p*NPXyl (Fig. 5A). No activity was detected using these synthetic substrates. The absence of activity on *p*NPXyl confirmed the strict specificity of the recombinant enzymes for the hexose conformation for the glycoside extension, while the absence of activity on *p*NPGlc confirmed the specificity of the enzymes for the axial orientation of the hydroxyl group on the C4 of the hexose. The lack of activity on *o*NPGal further demonstrated the specificity of the enzymes for the α -linkage between Gal and the aglycon.

Tested against natural substrates, both recombinant Aga1 and Aga2 displayed the same activity profile: Mel > Raf \sim Sta. These substrates consist of galactose units α -(1,6) linked to glucose (Mel), sucrose (Raf), or raffinose (Sta). Aga1 was more active than Aga2 on plant-based substrates, with specific activity (SA) on Mel, Raf, and Sta of 42.3 ± 3.4 , 8.6 ± 0.4 , and $7.5 \pm 2.7 \text{ U} \cdot \text{mg}^{-1}$ and 10.1 ± 1.1 , 4.5 ± 0.4 , and $4.3 \pm 0.8 \text{ U} \cdot \text{mg}^{-1}$ for Aga1 and Aga2, respectively (Fig. 5B). No detectable activity could be measured on LBG and on GG. The product profile resulting from the hydrolysis of Raf was analyzed by HPAEC-PAD (Fig. 5C) with Gal and Suc produced from the cleavage of Gal from Raf. Interestingly, both Aga1 and Aga2 were also able to catalyze the hydrolysis of Gal α 1,3Gal, as shown by HPAEC-PAD, with the release of Gal in the reaction assays (see Fig. S4 in the supplemental material).

To further characterize Aga1 and Aga2 specificity, the activity of the recombinant enzymes was assessed in the presence of chemical glycosidase inhibitors using *p*NPGal as the substrate at 37°C with pH 6.0 and 5.5, respectively. Galactodeoxyojirimycin (Gal-DNJ) was the only compound which strongly inhibited both Aga1 and Aga2, by 97.6% and 60.6%, respectively, whereas miglitol and 1-deoxyojirimycin (DNJ) had no inhibitory effect. Gal-DNJ is derived from galactose, in contrast to miglitol and DNJ, which are

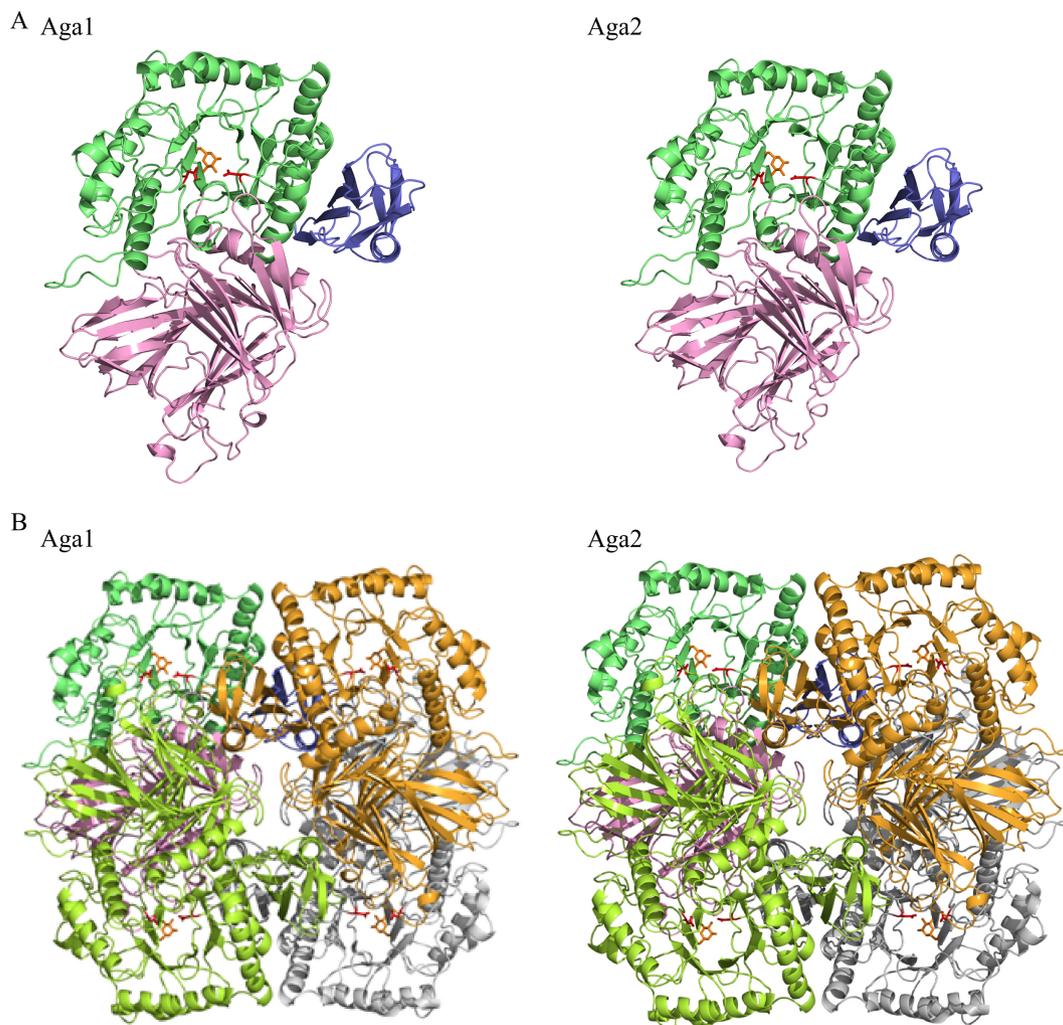


FIG 3 Homology models of *R. gnavus* E1 Aga1 and Aga2. (A) Cartoon representation of the monomeric structure models of Aga1 and Aga2, with the N-terminal domains in pink, the catalytic (β/α)₈ barrels in green, and the C-terminal domains in slate blue. Gal and catalytic residues are shown as orange and red sticks, respectively. (B) Cartoon representation of the tetrameric structure models. One monomer is oriented and colored as described for panel A, and the other monomers are colored in yellow, orange, and gray. Galactose and catalytic residues are shown as orange and red sticks, respectively.

both derived from glucose, confirming the strict specificity of Aga1 and Aga2 for the axial orientation of the hydroxyl group on C4. Aga1 was the more sensitive to Gal-DNJ, with a K_i of $0.39 \pm 0.03 \mu\text{M}$ compared to a K_i of $1.31 \pm 0.08 \mu\text{M}$ for Aga2 (not shown).

Sequence alignment of Aga1 and Aga2 with AgaSK-tru allowed the unambiguous identification of residues involved in catalysis (not shown). The role of the nucleophile could be assigned to D479 in Aga1 and D477 in Aga2, while the acid/base residue is predicted to be D549 and D547 in Aga1 and Aga2, respectively (Fig. 6). In order to identify the residues which may be involved in substrate specificity, the three-dimensional (3D) models of Aga1 and Aga2 (see above) were superimposed onto a model of AgaSK-tru in complex with the trisaccharide raffinose, derived from a superposition of the crystal structure of the AgaSK-tru-galactose complex with the structure of family GH27 *S. cerevisiae* α -galactosidase 1, Mel1, in complex with raffinose (PDB accession no. 3LRM) (23). In both the Aga1 and Aga2 models, numerous interactions with residues lining the active site stabilize galactose within subsite -1, numbered in accordance with

the nomenclature (16) (Fig. 6). In the model of Aga1, the OH2 hydroxyl interacts with C527, D549, and the main chain nitrogen of G530; the OH3 hydroxyl interacts with K477, which also makes a hydrogen bond with OH4, together with D367 and W412, whereas OH6 interacts with D368 and R444. Residues W200 and S529 presumably participate in substrate recognition in subsite +1, as well as H54, originating from a neighboring subunit in the tetramer. Putative hydrogen bonding partners for a sugar in Aga1 subsite +2 are Y341 and R444 (Fig. 6A). In the model of Aga2, the galactose residue of the substrate is stabilized in subsite -1 by interactions of OH2 with C525, G528, and D547; OH3 interacts with K475, which in turn interacts with OH4 together with D365 and W410; OH6 interacts with D366 and R442. Residues probably defining subsite +1 are H198, R442, and S527, together with two residues originating from a neighboring subunit in the tetramer, namely, Y51 and D53. Residues Y339 and R442 may play a role in stabilization of the substrate in Aga2 subsite +2 (Fig. 6B). By analogy to the crystal structures of AgaSK-tru and *Lactobacillus acidophilus* strain NCFM α -galactosidase MelA (25) in complex with galactose, the strong hydrogen interactions of the

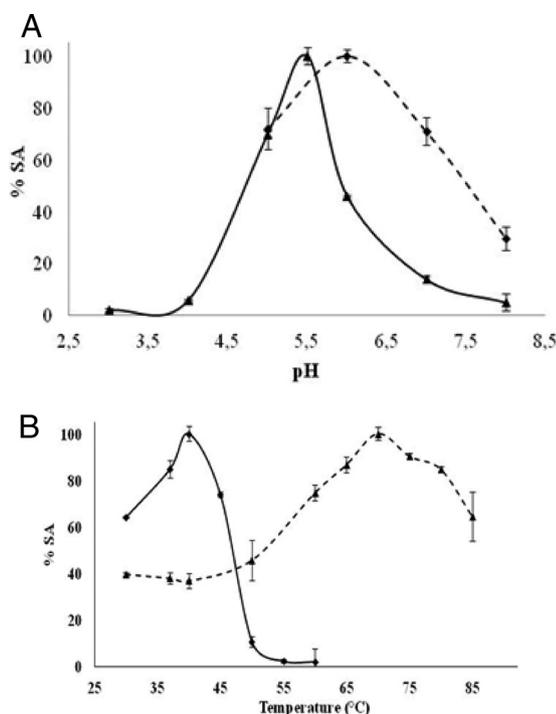


FIG 4 Effect of pH (A) and temperature (B) on the α -galactosidase-specific activity (SA) of recombinant *R. gnavus* E1 Aga1 (dashed lines and rhombuses) and Aga2 (solid lines and triangles). The data represent the means \pm standard deviations of the results of at least three independent assays.

OH4 with three neighboring amino acids, and the fact that an OH4 in equatorial configuration would make a steric clash with an adjacent Trp residue, explains the strict specificity of Aga1 and Aga2 for hexose with axial OH4.

Transglycosylation and self-condensation activities of Aga1 and Aga2. Aga1 and Aga2 belong to family GH36, and consequently they are expected to perform hydrolysis by the activity of a retaining mechanism (13, 29). Besides hydrolytic activity, some retaining glycosidases show transglycosylation activity, in which the glycone moieties are transferred to appropriate acceptors with hydroxyl groups.

We first examined the self-condensation activity of Aga1 and Aga2 by incubating Mel at a high concentration (68.5% or 2 M) with the recombinant enzymes for up to 150 h at 37°C. The products of the Aga1 reaction were analyzed by HPAEC-PAD. Initially, Mel was hydrolyzed to Gal and Glc, as expected (Fig. 7A to C); after longer incubation times, additional products were formed (Fig. 7D and E), and by 150 h there were 6 new peaks, in addition to those corresponding to Mel, Gal, and Glc, which appeared at retention times of 35.18 (P1), 47.58 (P2), 77.82 (P3), 52.11 (P4), 56.80 (P5), and 62.63 (P6) min (Fig. 7F), whereas no peak other than that corresponding to Mel appeared in the control experiment performed without the enzyme (Fig. 7B). The formation of

the new products was in agreement with the concomitant decrease in the Mel peak area and the increase in the product peak areas, with P1, P2, and P3 being the major products formed. The optimal transglycosylation yields were obtained after 96 h at 37°C, with no further increase at up to 150 h (not shown). To gain further structural information on the transglycosylation products, Aga1- and Aga2-catalyzed reactions after 96 h were analyzed by MALDI-TOF/TOF and GC-MS, with the results indicating the formation of α -(1,6)-linked [Hex]₄ to [Hex]₇ oligosaccharides corresponding to molecular ion peaks m/z = 885.4 for [Hex]₄, m/z = 1,089.5 for [Hex]₅, m/z = 1,293.6 for [Hex]₆, and m/z = 1,497.8 for [Hex]₇ in the MALDI-TOF analysis. The linear structure of the oligosaccharides was confirmed in the tandem MS (MS²) fragmentation with MALDI-TOF/TOF. The linkage analysis performed with GC-MS showed the 6-linked Gal as the main product, indicating the presence of 1,6-linked monosaccharides in the oligosaccharides before preparation of the samples for the analysis (Table 2; also see Fig. S5 in the supplemental material).

Next, we examined the ability of Aga1 and Aga2 to catalyze transglycosylation reactions by incubating the recombinant enzymes with pNPGal as the donor and various sugar acceptors such as Man, Gal, Glc, Mel, Suc, Lac, and Raf. Aga1 could transglycosylate all hexoses tested and α -(1 \rightarrow 6) linked oligosaccharides (Raf and Mel) but could not transglycosylate Xyl pentose or the β -linked Suc or Lac sugars (Table 2). The three monosaccharides were shown to be good acceptors for both Aga1 and Aga2, resulting in transglycosylation products of up to [Hex]₄ with Man and products of up to [Hex]₇ with Gal. Interestingly, both enzymes were able to iteratively add residues to Glc, resulting in oligosaccharide structures of up to [Hex]₁₂. Among the disaccharides tested, only Mel was a good acceptor for Aga1, with product structures of up to [Hex]₇, thus confirming the results obtained in self-condensation assays with Mel. Structures of up to [Hex]₅ were obtained when Aga2 was used with this substrate. Products of up to [Hex]₅ and [Hex]₄ were detected when the trisaccharide Raf was used as the acceptor with Aga1 and Aga2, respectively. Taken together, these results strongly suggest that Aga1 and Aga2 are able to catalyze the synthesis of novel galacto-oligosaccharides (GOS) by transglycosylation.

DISCUSSION

Microbial GH36 α -galactosidases play an important role in the degradation of dietary plant α -galactosides by the human microbiota. *R. gnavus* is one of the species of bacteria frequently found in the gastrointestinal tract of healthy individuals (42). Specific expression of α -galactosidase was detected in the gut of gnotobiotic mice associated with *R. gnavus* E1, and the genes encoding Aga1 and AgaSK were previously reported (1, 9). Here, we identified another GH36 gene, *aga2*, which is expressed *in vivo* and may contribute to *R. gnavus* E1 α -galactosidase activity. In contrast to the simple genetic organization reported for *aga1*, consisting of one AraC-XylS regulator and of the gene encoding the α -galactosidase on the opposite frame (1), the gene encoding GH36 Aga2 is

TABLE 1 Kinetic parameters of Aga1 and Aga2^a

Galactosidase	V_{\max} (mmol \cdot min ⁻¹)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ \cdot mM ⁻¹)	n
Aga1	137.79 \pm 2.96	0.21 \pm 0.03	196.06	943.96	0.88 \pm 0.12
Aga2	915.7 \pm 8.54	0.40 \pm 0.17	84.2	129.61	1.07 \pm 0.08

^a Data represent calculated values \pm standard deviations. n , number of substrate binding sites as calculated using the Hill equation.

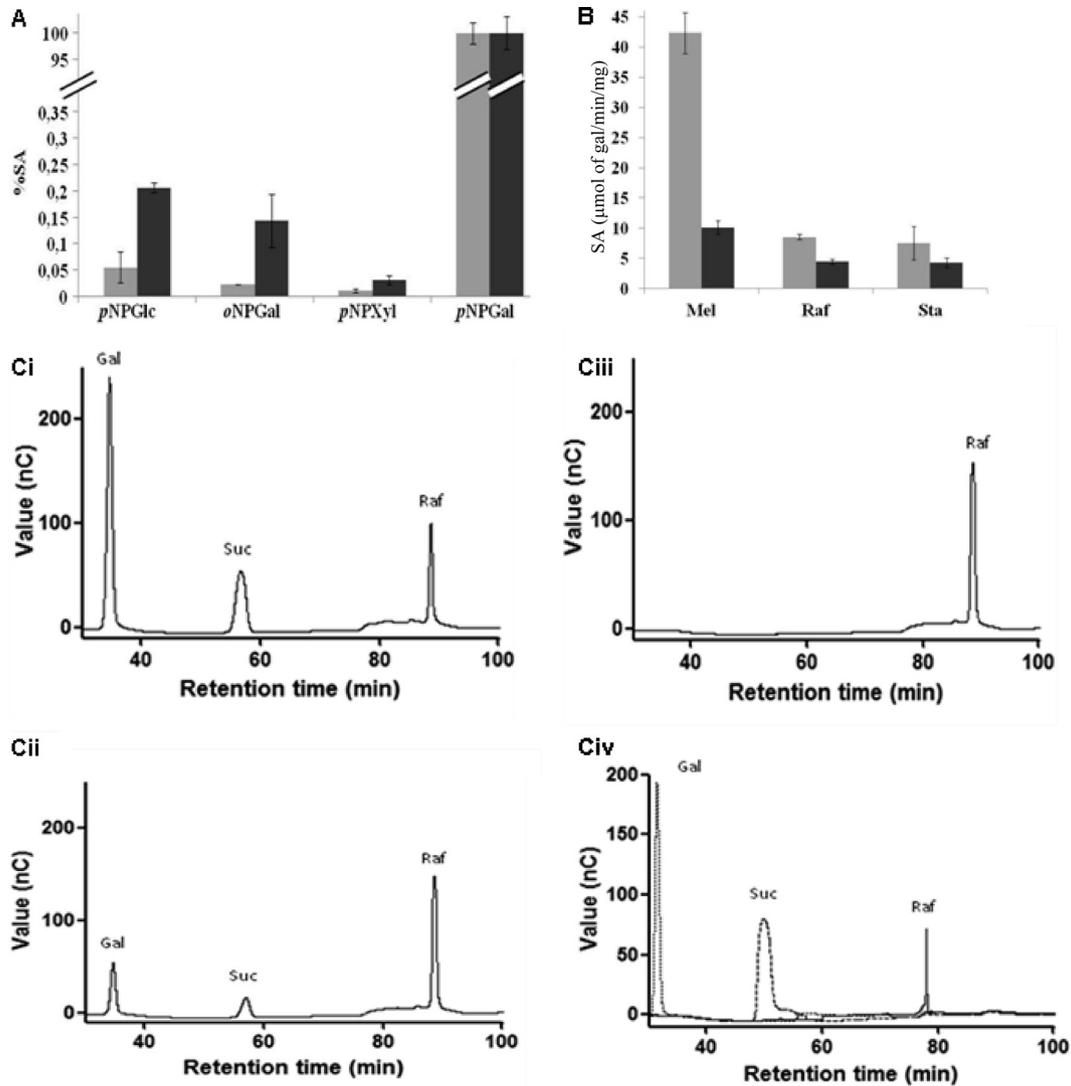


FIG 5 Substrate specificities of *R. gnavus* E1 Aga1 and Aga2. (A) Hydrolysis of *pNP*-pyranosides by *R. gnavus* E1 Aga1 (light gray) and Aga2 (dark gray). The recombinant enzymes were incubated with different sugars linked to *pNP*, and the cleavage of the linkage was estimated as the amount of *pNP* released in the reaction mixture. (B) Hydrolysis of natural substrates by *R. gnavus* E1 Aga1 (light gray) and Aga2 (dark gray). The specific activity was determined by quantifying the amounts of Gal released during incubation with the recombinant enzymes. (C) HPAEC-PAD analysis of hydrolysis products of Raf (12 mM) incubated for 4 h with Aga1 (114 nM; panel Ci) or Aga2 (234 nM; panel Cii) or a no-enzyme control (panel Ciii). All samples were diluted 1/10 prior to injection. Panel Cii shows an overlay of 0.5 mM standards of Raf (solid line), Suc (dashed line), and Gal (dotted line). nC, nanocoloumbs.

part of a cluster of 6 genes involved in sugar transport and metabolism, including ORFs coding for the putative regulator LacI, two sucrose-specific permeases, ScrA1 and -2, and two other glycoside hydrolases, GH13 and GH32 (see Fig. S1 in the supplemental material). In low-G+C content, Gram-positive bacteria, global regulation is implemented by the catabolite control protein, CcpA, a transcriptional regulator of the LacI-GalR family (40). RNA analysis confirmed that all six genes of the *R. gnavus* E1 *aga2* operon were cotranscribed from the promoter located upstream of *lacI* in gnotobiotic mice *in vivo*. A putative additional promoter responding to catabolite repression and embedding a consensual catabolite-responsive element (CRE) box was found in the *R. gnavus* E1 *aga2* cluster, downstream of *lacI*, suggesting that the genes within this operon respond to carbon catabolite repression. All three putative glycoside hydrolases, GH13, GH32, and GH36, are predicted to be cytosolic. The presence of complete putative sucrose-

specific PTS enzyme II complexes suggests that the *R. gnavus* E1 *aga2* operon is involved in extracellular sucrose utilization (see Fig. S6 in the supplemental material). For the majority of Gram-positive bacteria, the sucrose-dependent PTS is the predominant mechanism that facilitates sucrose uptake (5). This system is composed of two energy-coupling proteins, enzyme I and HPr (heat-stable, histidine-phosphorylatable protein), and several sugar-specific enzyme II proteins, which typically consist of three protein domains, EIIA, EIIB, and EIIC (48). The organization of the EII domains differs between bacterial species and may consist of a single fused protein or different fused and unfused domains. The translocation of the specific sugar through the membrane is facilitated by the integral membrane EIIC^{sugar} porter. However, it is the complex which catalyzes the concomitant transport and phosphorylation of its sugar substrates, resulting in an intracellular pool of phosphorylated carbohydrate (48). Here, we showed

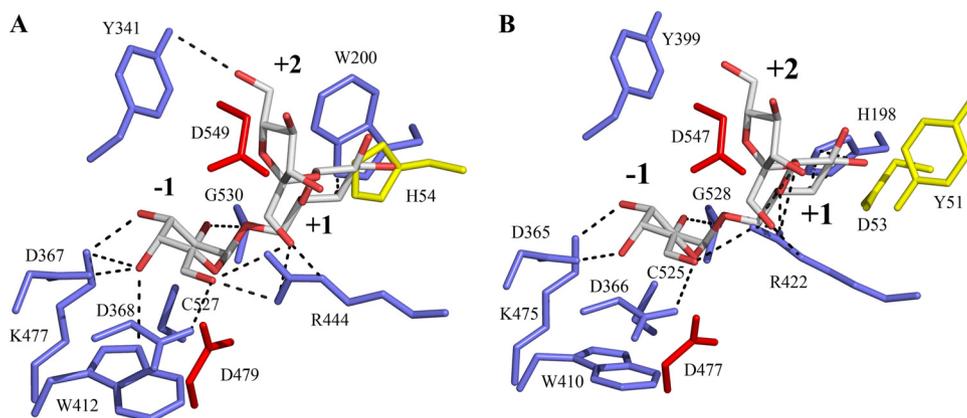


FIG 6 Close-up view of *R. gnavus* E1 Aga1 (A) and Aga2 (B) active sites. Raffinose, derived from a superposition of the models with *S. cerevisiae* α -galactosidase 1, Mel1, complexed with Raf (PDB accession no. 3lrm), is shown as sticks with gray carbons and red oxygens. Residues interacting with the substrate are shown as blue sticks, and residues originating from a neighboring subunit are depicted as yellow sticks. Catalytic residues are shown in red. Hydrogen bonding interactions with a distance threshold of 3.25 Å are shown as dashed lines, and substrate binding subsites are indicated in bold.

the presence of two putative PTS EII^{suc} in the *aga2* cluster, a complete sucrose-specific PTS1 permease that includes IIA, IIB, and IIC, and a PTS2 that includes IIB and IIC. Homologues of enzyme I and HPr are present elsewhere in the *R. gnavus* E1 genome. Enzyme IIC proteins that lack a cognate IIA and/or IIB domain or protein were considered to be incomplete but in the Glc family, and several PTS porters have the IIB and IIC domains fused but lack their own IIA domain and use the IIA^{Glc} protein of another system instead (53). Although results of the investigations conducted so far have indicated that it is restricted to glucose transport, it is tempting to speculate that such a mechanism also occurs in *R. gnavus* E1 for PTS-dependent sucrose transport. These sucrose-specific transport systems are generally clustered with the catabolic and regulatory genes in various arrangements in different bacteria, resulting in intracellular accumulation of sucrose-6P and further metabolism via a GH32 sucrose-6P hydrolase. In organisms that do not possess active PTS, non-PTS sugar permeases facilitate sucrose accumulation without chemical modification (49) and sucrose hydrolysis relies on the presence of a GH32 invertase or a GH13 sucrose phosphorylase for hydrolysis into fructose and glucose or glucose-1P, respectively. Generally, sucrose phosphorylases and sucrose hydrolases are found in different clusters within bacterial genomes (45). The presence of a putative GH32 sucrose-6P hydrolase enzyme in the *R. gnavus* E1 *aga2* operon is consistent with a cluster dedicated to extracellular sucrose utilization where the sucrose is transported into the cells as sucrose-6-P via the PTS-dependent sucrose system and subsequently hydrolyzed by the GH32 enzyme to yield glucose-6-P and fructose (see Fig. S6 in the supplemental material). Such genetic organization is reminiscent of sucrose-utilization loci in lactic acid bacteria (LAB) (36) such as *P. pentosaceus* ATCC 25745, *Lactobacillus sakei* subsp. *sakei* 23K, and *Lactobacillus plantarum* JDM1 (see Fig. S1 in the supplemental material). Since the *R. gnavus* E1 GH32 enzyme also shares 29% identity with the functionally characterized β -fructofuranosidases (*cscA*) from *Bifidobacterium longum* NCC2705 and *Bifidobacterium breve* UCC2003 (32, 47), it is also possible that this enzyme plays a role in the metabolism of intracellular sucrose transported unmodified via a non-PTS permease or generated by the breakdown of more complex carbohydrate to yield glucose and fructose (see Fig. S6 in the sup-

plemental material). In *R. gnavus* E1, GH36 Aga2 may provide an additional route for the cells to accumulate unmodified sucrose by intracellularly releasing sucrose and galactose from raffinose uptake (see Fig. S6 in the supplemental material), as previously suggested for AgaSK (9). The *R. gnavus* E1 *agaSK* operon includes a putative transcriptional regulator, a potential raffinose ABC transporter, and a putative GH13 sucrose phosphorylase (9). It has been suggested that, due to its bifunctional activity, AgaSK is able to release sucrose-6P from raffinose hydrolysis in *R. gnavus* E1 whereas sucrose, which escapes the kinase activity of AgaSK, is probably cleaved by the GH13 sucrose phosphorylase, encoded by the gene adjacent to *agaSK* (see Fig. S1 in the supplemental material). In the present work, we identified a putative GH32 sucrose-6P hydrolase within the *aga2* operon which would be able to further metabolize sucrose-6P into glucose-6P, thus complementing the proposed novel glycolytic pathway for intracellular sucrose assimilation (see Fig. S6 in the supplemental material).

Taken together, these studies suggest that in *R. gnavus* E1, raffinose is transported within the cytoplasm by an ABC transporter and subsequently hydrolyzed by the α -galactosidase activity of AgaSK, Aga1, and/or Aga2 into galactose and sucrose. Intracellular sucrose can then be either further phosphorylated into sucrose-6-P by AgaSK and cleaved by the *aga2* operon-specific GH32 into glucose-6-P and fructose or phosphorylated by the GH13 sucrose phosphorylase found in the *agaSK* operon into glucose-1P and fructose (see Fig. S6 in the supplemental material). Glucose-1P would be subsequently transformed by a phosphoglucomutase (Pgm) into glucose-6P. In both mechanisms, fructose is afterward phosphorylated by a fructokinase (FruK). Genes with significant homologies to Pgm and FruK have been found elsewhere in the *R. gnavus* E1 genome.

It is believed that bacterial sucrose-utilization systems have developed by modular evolution, where individual genes sharing a common origin have been independently associated into regulons and operons in different bacterial strains (12). Part of the *aga2* operon is absent in the genome of *R. gnavus* ATCC 29149, and our bioinformatics analysis suggests that *scrA2* has been added to the *R. gnavus* E1 genome together with *malE1* and *aga2* in the existing sucrose operon between the *scrA1* and *sacA* borders, resulting in a sucrose-raffinose utilization cluster. The *malE1* sequence belongs

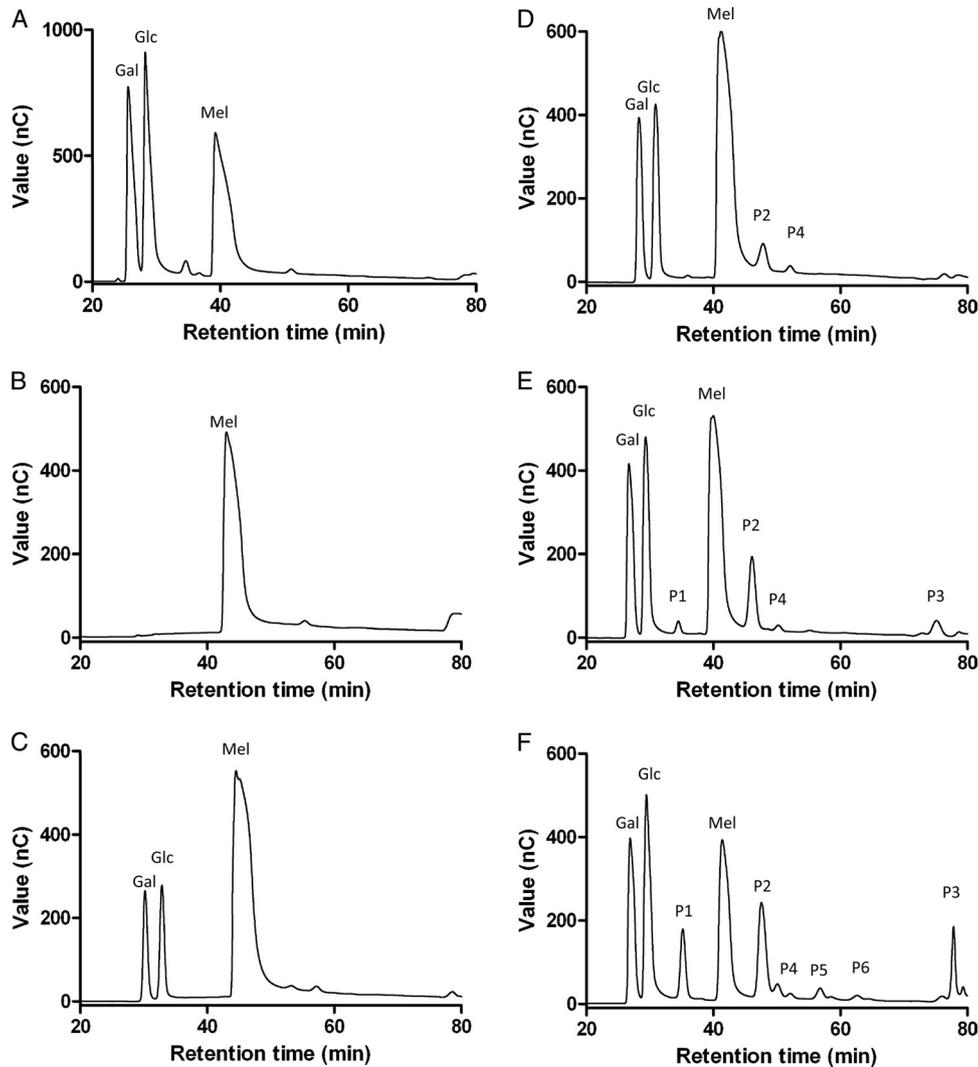


FIG 7 HPAEC-PAD analysis of self-condensation products of melibiose (2 M) incubated with Aga1 (15.7 nM). (A) Standards Glc, Gal, and Mel at 500 μ M; (B) no-enzyme control after 150 h; (C) reaction after 1 h; (D) reaction after 10 h; (E) reaction after 24 h; (F) reaction after 150 h.

to GH13_31 (i.e., subfamily 31 of family GH13) (14). Both *P. pentosaceus* and *L. sakei* sucrose clusters have an *agl* gene, encoding an α -glucosidase, downstream of the sucrose regulator gene (39) (see Fig. S1 in the supplemental material). The *P. pentosaceus* and *L. sakei* α -glucosidases share around 49% and 56% identity with the GH13 found in the *aga2* cluster, showing evidence for gene acquisition in existing sucrose operons, although the functionality and specificity of these enzymes remain to be biochemically confirmed. Indeed, it has been reported that an *S. xyloso* α -glucosidase essential for maltose-maltotriose catabolism conferred sucrose activity to *E. coli* and may be implicated in sucrose-6P hydrolysis of *S. xyloso* (20).

In order to confirm the suggested role of Aga2 in the *R. gnavus* E1 cluster, we functionally characterized Aga2 together with Aga1 and showed that the recombinant enzymes were active with respect to Raf, in agreement with their role in the proposed sucrose utilization pathway in *R. gnavus* E1 (see Fig. S6 in the supplemental material). Functional redundancy is a common evolutionary response in bacteria exposed to different environmental pressures

and is considered an important feature of bacterial adaptation to the gastrointestinal tract (42). Both Aga2 and Aga1 were active against other natural substrates such as Sta and Mel but not on LBG and GG, thus classifying these enzymes in the GH36 group I, encompassing α -galactosidases preferentially active on oligosaccharides, in contrast with group II, which includes enzymes preferentially active on polysaccharides (25), as is also in line with our bioinformatics analysis. The release of products from Raf and Mel was analyzed by HPAEC-PAD, and in both cases the terminal Gal was cleaved, leaving Suc, but not Fru, in the case of Raf and leaving Glc in the case of Mel. This is in accordance with the enzyme specificity; indeed, the product profile obtained for Aga1 and Aga2 is similar to that obtained for the hydrolysis of Raf and Mel by AgaSK (9). Differences between Aga1 and Aga2 in catalytic efficiency can be attributed both to a lower turnover and to a weaker substrate binding for Aga2. The strict specificity of the enzymes with respect to α -linked galactose was further demonstrated by their absence of hydrolysis of *o*NPGal, *p*NPGlc, and *p*NPXyl. Modeling of Aga1 and Aga2 identified key amino acid residues

TABLE 2 Self-condensation and transglycosylation products for Aga1 and Aga2 reactions as analyzed by MALDI-TOF/LIFT mass spectrometry

Product category and sugar	Enzyme ^a			
	Aga1		Aga2	
	[Hex] _{max}	<i>m/z</i>	[Hex] _{max}	<i>m/z</i>
Self-condensation ^b				
Mel	[Hex] ₇	1,497.8	[Hex] ₇	1,497.8
Transglycosylation ^c				
Mel	[Hex] ₇	1,497.8	[Hex] ₆	1,293.6
Gal	[Hex] ₅	1,089.5	[Hex] ₇	1,497.8
Glc	[Hex] ₁₂	2,519.2	[Hex] ₈	1,701.9
Man	[Hex] ₄	885.4	[Hex] ₄	885.4
Lac	NP	NA	NT	NA
Suc	NP	NA	NT	NA
Raf	[Hex] ₅	1,089.5	[Hex] ₄	885.4
Xyl	NP	NA	NT	NA

^a [Hex]_n, maximum size of oligosaccharide detected in mass spectrometry; NP, no product; NT, not tested; NA, not applicable.

^b Data represent the results of a 96-h melibiose self-condensation reaction.

^c Data represent the results of a 25-h transglycosylation reaction performed using various sugar acceptors.

which may be involved in substrate specificity and stabilization of the α -linked galactoside substrates within the active site. Some family GH36 members exhibit α -*N*-acetylgalactosaminidase activity (10), but the tight interactions of the substrate with Aga1 and Aga2 do not allow for exocyclic substitutions on the galactose moiety. Aga1 and Aga2 occur as homotetramers in solution, as shown by AUC results, in line with their classification in GH36 subgroup I, in which the majority of the members adopt a tetrameric form. Tetrameric assemblies based on the structure of AgaSK could be constructed for Aga1 and Aga2 and revealed that the active site pocket, similarly to that of *Lactobacillus brevis* and *L. acidophilus* NCFM α -galactosidases (25), is forged by the tight association of adjacent monomers, suggesting cooperation between subunits, although this needs to be further confirmed. Family GH36 members carry out hydrolysis with a net retention of the anomeric configuration, which is why enzymes from this family have been shown to possess transglycosylation activity (38). This property is essential for the synthesis of oligosaccharide structures which cannot be synthesized by classical chemistry (51). Aga1 and Aga2 showed the ability to transglycosylate monosaccharides and α (1 \rightarrow 6)-linked oligosaccharides. Strikingly, high-molecular-weight (up to [Hex]₁₂) GOS could be synthesized using glucose as a sugar acceptor, which is unique within GH36 α -galactosidases. These oligosaccharides have the potential to act as prebiotics by selectively stimulating the growth and/or activity of the beneficial gut microbiota (2, 58). The properties of GOS that occur *in vivo* are caused by selective fermentation as well as by competitive interactions within the intestinal environment (17). Defining the molecular basis underlying the interrelationship between dietary carbohydrates and their assimilation by human gut symbionts is essential to enhance our ability to manipulate the properties of the microbiota.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Biotechnology and Biological Sciences Research Council (BBSRC). This research was partly funded

by the BBSRC Institute Strategic Programme IFR/08/1 (Integrated Biology of the Gastrointestinal Tract), the Ministère de l'Enseignement Supérieur et de la Recherche Scientifique, and the Alliance program.

We thank Tom Clarke (University of East Anglia, Norwich, United Kingdom) and Fiona Husband (Institute of Food Research) for their assistance with AUC and CD analysis, respectively, and Ange Pujol (University of Aix Marseille) for his help with the bioinformatics analysis. Further, we thank Mark Philo (Institute of Food Research, Norwich, United Kingdom) and Mike Naldrett (John Innes Centre, Norwich, United Kingdom) for technical support with the MS analysis. We are indebted to Sophie Rabot for providing access to the ANAXEM platform (INRA, Jouy-en-Josas, France) and to Chantal Bridonneau and Pascal Guillaume for their skillful technical assistance with the axenic rats.

REFERENCES

- Aguilera M, et al. 2012. Aga1, the first alpha-galactosidase from the human bacteria *Ruminococcus gnavus* E1, efficiently transcribed in gut conditions. *Res. Microbiol.* 163:14–21.
- Andersen JM, et al. 2011. Transcriptional and functional analysis of galactooligosaccharide uptake by lacS in *Lactobacillus acidophilus*. *Proc. Natl. Acad. Sci. U. S. A.* 108:17785–17790.
- Arumugam M, et al. 2011. Enterotypes of the human gut microbiome. *Nature* 473:174–180.
- Aslanidis C, Schmid K, Schmitt R. 1989. Nucleotide sequences and operon structure of plasmid-borne genes mediating uptake and utilization of raffinose in *Escherichia coli*. *J. Bacteriol.* 171:6753–6763.
- Barabote RD, Saier MH, Jr. 2005. Comparative genomic analyses of the bacterial phosphotransferase system. *Microbiol. Mol. Biol. Rev.* 69:608–634.
- Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27:343–350.
- Berrow NS, et al. 2007. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* 35:e45. doi:10.1093/nar/gkm047.
- Blatch GL, Scholle RR, Woods DR. 1990. Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrose uptake-encoding region. *Gene* 95:17–23.
- Bruel L, et al. 2011. α -Galactosidase/sucrose kinase (AgaSK), a novel bifunctional enzyme from the human microbiome coupling galactosidase and kinase activities. *J. Biol. Chem.* 286:40814–40823.
- Calcutt MJ, Hsieh HY, Chapman LF, Smith DS. 2002. Identification, molecular cloning and expression of an alpha-N-acetylgalactosaminidase gene from *Clostridium perfringens*. *FEBS Microbiol. Lett.* 214:77–80.
- Chen Y, Case DA, Reizer J, Saier MH, Jr, Wright PE. 1998. High-resolution solution structure of *Bacillus subtilis* IIAGlc. *Proteins* 31:258–270.
- Comas I, González-Candelas F, Zúñiga M. 2008. Unraveling the evolutionary history of the phosphoryl-transfer chain of the phosphoenolpyruvate:phosphotransferase system through phylogenetic analyses and genome context. *BMC Evol. Biol.* 8:147. doi:10.1186/1471-2148-8-147.
- Comfort DA, et al. 2007. Biochemical analysis of *Thermotoga maritima* GH36 alpha-galactosidase (TmGalA) confirms the mechanistic commonality of clan GH-D glycoside hydrolases. *Biochemistry* 46:3319–3330.
- Coutinho PM, Henrissat B. 1999. Carbohydrate-active enzymes: an integrated database approach, p 3–12. In Gilbert HJ, Davies G, Henrissat B, Svensson B (ed), *Recent advances in carbohydrate bioengineering*. The Royal Society of Chemistry, Cambridge, United Kingdom.
- Crost EH, et al. 2011. Ruminococcin C, a new anti-*Clostridium perfringens* bacteriocin produced in the gut by the commensal bacterium *Ruminococcus gnavus* E1. *Biochimie* 93:1487–1494.
- Davies GJ, Wilson KS, Henrissat B. 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* 321:557–559.
- Davis LM, Martínez I, Walter J, Goin C, Hutkins RW. 2011. Barcoded pyrosequencing reveals that consumption of galactooligosaccharides results in a highly specific bifidogenic response in humans. *PLoS One* 6:e25200. doi:10.1371/journal.pone.0025200.
- Doré J, Sghir A, Hannequart-Gramet G, Corthier G, Pochart P. 1998. Design and evaluation of a 16S rRNA-targeted oligonucleotide probe for specific detection and quantitation of human faecal Bacteroides populations. *Syst. Appl. Microbiol.* 21:65–71.

19. Eckburg PB, et al. 2005. Diversity of the human intestinal microbial flora. *Science* 308:1635–1638.
20. Egeter O, Brückner R. 1995. Characterization of a genetic locus essential for maltose-maltotriose utilization in *Staphylococcus xylosus*. *J. Bacteriol.* 177:2408–2415.
21. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501.
22. Feese M, Pettigrew DW, Meadow ND, Roseman S, Remington SJ. 1994. Cation-promoted association of a regulatory and target protein is controlled by protein phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* 91:3544–3548.
23. Fernández-Leiro R, Pereira-Rodríguez A, Cerdán ME, Becerra M, Sanz-Aparicio J. 2010. Structural analysis of *Saccharomyces cerevisiae* alpha-galactosidase and its complexes with natural substrates reveals new insights into substrate specificity of GH27 glycosidases. *J. Biol. Chem.* 285:28020–28033.
24. Fouet A, Klier A, Rapoport G. 1986. Nucleotide sequence of the sucrose gene of *Bacillus subtilis*. *Gene* 45:221–225.
25. Fredslund F, et al. 2011. Crystal structure of alpha-galactosidase from *Lactobacillus acidophilus* NCFM: insight into tetramer formation and substrate binding. *J. Mol. Biol.* 412:466–480.
26. Gemmecker G, et al. 1997. Glucose transporter of *Escherichia coli*: NMR characterization of the phosphocysteine form of the IIB(Glc) domain and its binding interface with the IIA(Glc) subunit. *Biochemistry* 36:7408–7417.
27. Gering M, Bruckner R. 1996. Transcriptional regulation of the sucrose gene of *Staphylococcus xylosus* by the repressor ScrR. *J. Bacteriol.* 178:462–469.
28. Gibson GR, Roberfroid MB. 1995. Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. *J. Nutr.* 125:1401–1412.
29. Henrissat B, Davies G. 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7:637–644.
30. Hooper LV, Midtvedt T, Gordon JI. 2002. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu. Rev. Nutr.* 22:283–307.
31. Jankovic I, Egeter O, Brückner R. 2001. Analysis of catabolite control protein A-dependent repression in *Staphylococcus xylosus* by a genomic reporter gene system. *J. Bacteriol.* 183:580–586.
32. Kullin B, Abratt VR, Reid SJ. 2006. A functional analysis of the *Bifidobacterium longum* cscA and scrP genes in sucrose utilization. *Appl. Microbiol. Biotechnol.* 72:975–981.
33. Ley RE, et al. 2008. Evolution of mammals and their gut microbes. *Science* 320:1647–1651.
34. Lozupone CA, et al. 2008. The convergence of carbohydrate active gene repertoires in human gut microbes. *Proc. Natl. Acad. Sci. U. S. A.* 105:15076–15081.
35. Ludwig W, Schleifer KH, Whitman WB. 2009. Revised road map to the phylum Firmicutes, p 1–13. In De Vos P, et al. (ed), *Bergey's manual of systematic bacteriology*, 2nd ed, vol 3 (the Firmicutes). Springer-Verlag, New York, NY.
36. Makarova K, et al. 2006. Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 103:15611–15616.
37. Meins M, et al. 1993. Cysteine phosphorylation of the glucose transporter of *Escherichia coli*. *J. Biol. Chem.* 268:11604–11609.
38. Nakai H, et al. 2010. *Aspergillus nidulans* alpha-galactosidase of glycoside hydrolase family 36 catalyses the formation of alpha-galacto-oligosaccharides by transglycosylation. *FEBS J.* 277:3538–3551.
39. Naumov DG, Livshits VA. 2001. Molecular structure of the locus for sucrose utilization by *Lactobacillus plantarum*: comparison with *Pediococcus pentosaceus*. *Mol. Biol. (Mosk.)* 35:19–27. (In Russian.)
40. Nguyen CC, Saier MH, Jr. 1995. Phylogenetic, structural and functional analyses of the LacI-GalR family of bacterial transcription factors. *FEBS Lett.* 377:98–102.
41. Oxley D, Currie G, Bacic A. 2004. Analysis of carbohydrates from glycoproteins, p 579–636. In Simpson RJ (ed), *Purifying proteins for proteomics: a laboratory manual*, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
42. Qin J, et al. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59–65.
43. Rajilić-Stojanović M, Smidt H, de Vos WM. 2007. Diversity of the human gastrointestinal tract microbiota revisited. *Environ. Microbiol.* 9:2125–2136.
44. Ramare F, et al. 1993. Trypsin-dependent production of an antibacterial substance by a human *Peptostreptococcus* strain in gnotobiotic rats and in vitro. *Appl. Environ. Microbiol.* 59:2876–2883.
45. Reid SJ, Abratt VR. 2005. Sucrose utilisation in bacteria: genetic organisation and regulation. *Appl. Microbiol. Biotechnol.* 67:312–321.
46. Robillard GT, Broos J. 1999. Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system. *Biochim. Biophys. Acta* 1422:73–104.
47. Ryan SM, Fitzgerald GF, van Sinderen D. 2005. Transcriptional regulation and characterization of a novel beta-fructofuranosidase-encoding gene from *Bifidobacterium breve* UCC2003. *Appl. Environ. Microbiol.* 71:3475–3482.
48. Saier MH, Fagan MJ, Hoischen C, Reizer J. 1993. Transport mechanisms, p 133–156. In Sonenshein AL, Hoch JA, Losick R (ed), *Bacillus subtilis* and other Gram-positive bacteria: biochemistry, physiology, and molecular genetics. American Society for Microbiology, Washington, DC.
49. Saier MH. 1998. Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea and eukarya. *Adv. Microb. Physiol.* 40:81–136.
50. Sato Y, Poy F, Jacobson GR, Kuramitsu HK. 1989. Characterization and sequence analysis of the scrA gene encoding enzyme IIScr of the *Streptococcus mutans* phosphoenolpyruvate-dependent sucrose phosphotransferase system. *J. Bacteriol.* 171:263–271.
51. Shaikh FA, Withers SG. 2008. Teaching old enzymes new tricks: engineering and evolution of glycosidases and glycosyl transferases for improved glycoside synthesis. *Biochem. Cell Biol.* 86:169–177.
52. Takami H, et al. 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res.* 28:4317–4331.
53. Tchieu JH, Norris V, Edwards JS, Saier MH, Jr. 2001. The complete phosphotransferase system in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* 3:329–346.
54. Tettelin H, et al. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* 293:498–506.
55. Trugo LC, Farahand A, Cabral L. 1995. Oligosaccharide distribution in Brazilian soya bean cultivars. *Food Chem.* 52:385–387.
56. Vallenet D, et al. 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. *Database (Oxford)* 2009:bap021. doi:10.1093/database/bap021.
57. Wagner E, Götz F, Brückner R. 1993. Cloning and characterization of the scrA gene encoding the sucrose-specific enzyme II of the phosphotransferase system from *Staphylococcus xylosus*. *Mol. Gen. Genet.* 241:33–41.
58. Walton GE, et al. 2012. A randomised crossover study investigating the effects of galacto-oligosaccharides on the faecal microbiota in men and women over 50 years of age. *Br. J. Nutr.* 107:1466–1475.
59. Wang G, et al. 2000. Solution structure of the phosphoryl transfer complex between the signal transducing proteins HPr and IIA(glucose) of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system. *EMBO J.* 19:5635–5649.
60. Watanabe K, Chishiro K, Kitamura K, Suzuki Y. 1991. Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from *Bacillus thermoglucosidasi* KP1006. *J. Biol. Chem.* 266:24287–24294.
61. Wu GD, et al. 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334:105–108.
62. Yatsunenkov T, et al. 2012. Human gut microbiome viewed across age and geography. *Nature* 486:222–227.