Characterization of 4,6-α-glucanotransferase enzymes and their functional role in Lactobacillus reuteri
Bai, Yuxiang

Important note: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG Research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 6

Crystal structure of the 4,6-α-glucanotransferase GtfB: Evidence that dietary changes triggered the evolution of glucansucrases from α-amylases

Yuxiang Bai\textsuperscript{1,3}, Joana Gangoiti-Muñecas\textsuperscript{1}, Bauke W. Dijkstra\textsuperscript{2}, Lubbert Dijkhuiizen\textsuperscript{1}, Tjaard Pijning\textsuperscript{2,*}

\textsuperscript{1} Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, The Netherlands

\textsuperscript{2} Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

\textsuperscript{3} The State Key Laboratory of Food Science and Technology, School of Food Science and Technology, Jiangnan University, Wuxi 214122, China

\textit{In preparation for submission}
Abstract

Dental caries is caused by oral bacteria adhering themselves to the tooth enamel, promoted by the formation of biofilms. A major constituent of biofilms are extracellular polysaccharides (EPS) synthesized by glucansucrase (GS) enzymes from sucrose (yielding α-glucans). Since sucrose is a relative recent addition to the human diet, the evolutionary origin of GS is intriguing and may help to understand the origin of caries. GS are classified in glycoside hydrolase family 70 (GH70); it has been proposed that they evolved from GH13 α-amylases. If true, evolutionary events would have had to involve changes in reaction- and substrate specificity; however the details of such changes, and which factors would trigger them, are unclear. Interestingly, a GH70 subfamily with 4,6-α-glucanotransferase (4,6-α-GTase) activity “intermediate” of α-amylases and GS has been discovered. In this study we provide 3D structural, genomic and phylogenetic data supporting the evolution of α-amylases to GS via 4,6-α-GTases. The crystal structure of the 4,6-α-GTase GtfB-ΔNΔV from Lactobacillus reuteri 121 reveals how the α-amylase substrate binding scaffold is retained, but different enough to favor transglycosylation instead of degradation. It also explains how GS could have evolved further towards sucrose utilization instead of starch. In vivo, 4,6-α-GTase and GS gene products within the same Lactobacillus strain complement each other in EPS synthesis, depending on the available carbon source. We propose that the evolution of GS in oral bacteria was triggered by dietary changes, in order to expand the reaction specificity and fermentable substrate specificity.
Introduction

Dental caries, a bacterial infection of the tooth enamel, poses a worldwide health problem (1, 2). It is caused by plaque-forming bacteria that ferment dietary carbohydrates in the oral cavity, and, by doing so, produce acids that dissolve calcium phosphate in the enamel. Adhesion of the oral bacteria to the tooth enamel is thought to be promoted by the formation of biofilms, which is facilitated by the synthesis of extracellular polysaccharides (EPS) (3-6). EPS are synthesized by glucansucrase (GS) enzymes, from sucrose (7-9), a widely accepted but relative recent addition to the human diet (yielding α-glucans). In this light, the evolutionary origin of GS from oral bacteria such as Lactobacillus and Streptococcus species, has raised interest, as it may help to understand the origin of dental caries. Notably, GS, belonging to the glycoside hydrolase family 70 (GH70; www.cazy.org (10), share sequence homology with α-amylases of GH13, which are also widely distributed in oral bacteria. The core domains of GS and α-amylases have similar folds; however the GS have a circularly permuted (β/α)₈ barrel in the catalytic domain, and contain auxiliary domains which are absent in α-amylases (11, 12). Moreover, oral α-amylases act on starch-derived oligosaccharides (OS) instead of sucrose. Thus, if GS have evolved from α-amylases, a change in reaction type specificity (from substrate degradation to polymerization) and a change in substrate specificity (from starch to sucrose) would be two key evolutionary achievements. Vujičić-Žagar et al. (12) proposed that intermediate/precursor enzymes attained circular permutation by duplication of a common ancestor α-amylase gene, and that auxiliary domains originate from insertion events; however, little is known about how GS evolved to be able to use sucrose for the synthesis of EPS, and which factors would drive such evolutionary pathways.

Recently we identified 4,6-α-glucanotransferase enzymes (4,6-α-GTases) in Lactobacillus species. These enzymes belong to GH70 and share 45-50% sequence identity with GS, but phylogenetically clearly form a small subfamily, with notable differences occurring near the active site and in loop lengths (13, 14). Biochemical characterization of 4,6-α-GTases showed that in vitro they do not act on sucrose, but instead, like α-amylases, process starch-derived OS; on the other hand their product specificity does resemble the GS in that they synthesize α-glucan oligo-/polysaccharides with a high degree of α1→6 linkages (13-17).
These 4,6-α-GTases thus have an activity with characteristics from both α-amylases and GS. Whether this activity reflects the role of these enzymes in vivo, remains to be established.

Here we provide 3D structural, genomic and phylogenetic data supporting an evolutionary pathway in which α-amylases evolved to GS via 4,6-α-GTase-like intermediates, with the latter two sharing a common ancestor closely related to subfamily 5 of GH13 α-amylases. The unique active site architecture of the 4,6-α-glucanotransferase GtfB-ΔNΔV from L. reuteri 121 retains an α-amylase-like scaffold, with modifications that explain the preference for transglycosylation and synthesis of polymer products. The 4,6-α-GTase architecture is also different from that of GS and shows which structural changes in the latter are related to shifting towards sucrose utilization. In vivo, depending on the available carbohydrate source (starch-derived OS or sucrose), GS GtfA and 4,6-α-GTase GtfB of Lactobacillus reuteri strain 121 both synthesize α-glucans with α1→4 and α1→6 linkages. We propose that oral bacteria such as Lactobacilli have evolved to expand their reaction specificity and fermentable substrate specificity, driven by changes in carbohydrate uptake in the oral cavity.

Materials and Methods

Cloning, truncation and site-directed mutagenesis, expression, purification and characterization

The cloning, mutagenesis and truncation strategy of gtfB-ΔN and gtfB-ΔNΔV are described in SI text. Expression and purification of gtfB-ΔN was performed as described by Bai et al. (15); expression and purification of GtfB-ΔNΔV and its mutants followed a similar procedure with a modification as shown below. The single mutant Y1521A and double mutant W790A/W811A were constructed to study the possible role of aromatic stacking interactions of remote sugar binding sites; mutation T920W aimed at blocking donor subsites -2/-3 in the tunnel. Finally, mutants K1128A/N/H/W and Y1055G were constructed to investigate the importance of these residues near subsites -1/+1 and subsite +2, respectively.

After His-tag affinity and anion exchange chromatography, the proteins were purified by gel filtration (Superdex 200 10/300 GL) in sodium acetate buffer (25
mM, pH 5.0, 150 mM NaCl, 1 mM CaCl₂). Dynamic light scattering experiments of wild-type GtfB-ΔNAV were performed on a Wyatt NanoStar (Wyatt Technology, Santa Barbara, US) at 293 K.

**Crystallization and soaking experiments**

Attempts to crystallize GtfB and GtfB-ΔN were unsuccessful. For GtfB-ΔNAV, hanging drop vapour diffusion crystallization experiments were setup at 293 K using drops containing 1.5 μl of protein solution (4-5 mg/ml GtfB-ΔNAV, 25 mM HAc/NaAc buffer, pH 5.0, 0.15 M NaCl, 1 mM CaCl₂) mixed with 1.5 μl precipitant solution (17-20% (w/v) PEG 3350, 0.1 M (NH₄)₂SO₄, 0.1 M NaCl) equilibrated against 0.5 ml of precipitant solution. To avoid excessive nucleation, a microseeding protocol was used in which 0.3 μl of diluted seed stock was introduced to hanging drops containing 1.5 μl of protein solution (as above) mixed with 1.5 μl of a solution containing 31% (w/v) PEG 3350, 0.1 M (NH₄)₂SO₄, 0.1 M NaCl). Crystals of GtfB-ΔNAV and its mutants tended to grow in clusters, and were cryoprotected in stabilizing solution containing 35% (w/v) PEG 3350, 0.1 M (NH₄)₂SO₄, 25 mM BisTris-HCl buffer pH 5.5, 25 mM HAc/NaAc buffer pH 5.0, 1 mM CaCl₂ (supplemented with 10% (v/v) glycerol for the wild-type apo data set). For soaking experiments, wild-type GtfB-ΔNAV crystals were soaked in stabilizing solution supplemented with 50 mM maltohexaoxe (4-10 minutes); GtfB-ΔNAV D1015N crystals were soaked in stabilizing solution supplemented with 100 mM maltopentaose (8 min).

**Data collection**

Data sets for wild-type apo and D1015N/maltopentaose were collected at the European Synchrotron Radiation Facility (ESRF) at beamlines ID29 and ID23-1, respectively, at 100 K; details and processing statistics are given in Table S4. All data sets were processed with XDS (18) and AIMLESS (19). Crystals of GtfB-ΔNAV and its mutants belong to space group C 2 and tended to show high mosaicity.

**Structure determination**

The structure of GtfB-ΔNAV was determined by molecular replacement using PHASER (20), using domains A, B, C and IV of the crystal structure of *L. reuteri* 180 Gtf180-ΔN (PDB 3KLK) (12) as a search model. The asymmetric unit
Chapter 6

contained two molecules of GtfB-ΔNΔV (Fig. 1A). Rigid-body and restrained refinement was performed using REFMAC5 (21) within the CCP4 package (22). Cycles of restrained refinement were alternated with manual building in COOT (23); non-crystallographic symmetry (NCS) restraints were used throughout refinement. The last 11 or 10 C-terminal residues (including the His6-tag) were not visible in electron density. Carbohydrate ligands were included only in the last stages of refinement; when the ligands had comparable binding modes in both protein molecules, we only described the situation for protein chain A. Final models were validated using MolProbity (24). PDBeFold (25) was used to analyze structural homology to other GH13 and GH70 enzymes. Structure figures were prepared with PyMOL (Schrödinger, LLC).

**Enzymatic assays**

Enzyme activities of wild-type GtfB-ΔNΔV and mutants derived were measured based on an assay developed by Bai et al. (15). Final products were analyzed by TLC, 1H NMR and HPAEC. Reaction of GtfB-ΔN (and maltase) with a blocked substrate (BPNPG7) was done to investigate the endo-specificity of GtfB-ΔNΔV. Details of the reactions and the respective product analysis are given in SI text.

**Modeling and docking**

Structural superpositions between GtfB-ΔNΔV and related enzymes were done in PyMOL by fitting the Cα and Cβ atoms of the corresponding three catalytic residues. The following glucansucrase structures were used: *Lactobacillus reuteri* 180 Gtf180-ΔN complexed with sucrose or with maltose (PDBs: 3HZ3 and 3KLL) (12); *Lactobacillus reuteri* 121 GtfA-ΔN (PDB: 4AYG) (26); *Streptococcus mutans* Gtf-SI (PDB: 3AIE) (27). The structure of *Bacillus halmapalus* α-amylase complexed with a pseudononasaccharide (PDB: 2GJP) (28) and *B. subtilis* α-amylase E208Q complexed with maltopentaose (PDB 1BAG) (29) served as models for α-amylases.

The structure of the *Bacillus subtilis* α-amylase maltopentaose complex (PDB 1BAG) (29) served as a template to model donor substrate binding, to extend the donor substrate observed in the GtfB-ΔNΔV - G5 complex towards acceptor subsites. The structure of *Bacillus circulans* strain 251 CGTase covalent glucosyl-enzyme intermediate (PDB: 1CXL) (36) was used to construct a model
of the glucosyl-enzyme intermediate in GtfB-ΔNΔV. Subsequently, a model for binding of the acceptor substrate panose (O-α-D-glucopyranosyl-(1,6)-O-α-D-glucopyranosyl-(1,4)-D-glucose) was constructed by placing the trisaccharide in subsites +1 to +3, guided by hydrogen bond interactions and available docking space. Further details of modeling procedures are given in SI text.

Genomic mapping of glucansucrase and 4,6-α-glucanotransferase

The locations of the genes encoding glucansucrase and 4,6-α-glucanotransferase proteins from Lactobacillus reuteri strains were obtained based on the graphics data in NCBI database.

Sequence alignments and phylogenetic analyses

MUSCLE (30) was used to align the amino acid sequences of full-length glucansucrases and 4,6-α-glucanotransferases, and included 4,6-α-glucanotransferases GtfB of Lactobacillus reuteri 121 (Uniprot entry Q5SBM0), GtfML4 of Lactobacillus reuteri ML1 (Q5SBN1), GtfW of L. reuteri DSM 20016 (A5VL73), and glucansucrases Gtf180 L. reuteri 180 (Q5SBN3), GtfA of L. reuteri 121 (Q5SBL9), GtfO of L. reuteri ATCC 55730 (Q4JLC7), GtfML1 of L. reuteri ML1 (Q5SBN0), Gtf-SI of Streptococcus mutans ATCC 700610 (P13470) and DSRE of Leuconostoc mesenteroides NRRL B-1299 (Q8G9Q2). The α-amylases of Bacillus licheniformis (P06278), B. halmapalus (O82839), Geobacillus stearothermophilus (P06279) and B. subtilis (P00691) were aligned separately. Alignment of homology loops A1, A2 and B was done using ESPript (31), adapted by structural superposition of the 4,6-α-glucanotransferase GtfB-ΔNΔV (this study), the glucansucrase Gtf180-ΔN (PDB 3KLK) (12) and the α-amylase from Bacillus licheniformis (PDB 1BLI) (32).

Following alignment of all characterized 4,6-α-GTases and GS, a phylogenetic was performed by the maximum likelihood method using MEGA version 6 (33).

A second phylogenetic analysis included all bacterial GH13 and GH70 enzymes; for the GH70 enzymes, domains A and B were first rearranged to ‘correct’ for the circular permutation of the (β/α)₈ barrel (BₕAₕ⋯AᵥBᵥ→AᵥBᵥBᵥAₕ, where the subscript denotes either the N- or C-terminal segment). The rearranged GH70 sequences were then added to the alignment of all characterized GH13 members classified into subfamilies, followed by phylogenetic analysis.
Homologues of full-length GtfB were searched using Protein-BLAST in the NCBI (http://www.ncbi.nlm.nih.gov) and UniProtKB (http://www.uniprot.org) databases.

**In vivo experiments**

Growth conditions of plate-grown *Lactobacillus reuteri* 35-5 are described in SI text. The extracellular polysaccharides (EPS) produced on agar plates were extracted and analyzed by monosaccharide analysis using HPAEC, size determination using SEC, and linkage determination using NMR. Details of the procedures are described in SI text.

**Results**

**Overall structure**

The 4,6-α-glucanotransferase GtfB from *L. reuteri* 121 could only be successfully crystallized after deletion of the N-terminal domain and domain V. The 1.80 Å resolution crystal structure of the resulting construct (GtfB-ΔNΔV, residues Met-762-1619-His6) shows the presence of four domains (A, B, C and IV; Fig. 1A). In the asymmetric unit, the two molecules can be superimposed almost perfectly with an r.m.s.d. of 0.27 Å; most of the crystal contacts between them are made by domains C, A and B lying adjacent to each other in opposite directions. In solution, the estimated molecular mass of 85 kDa observed in dynamic light scattering experiments suggests that GtfB-ΔNΔV is monomeric. The core domains A (catalytic), B and C of GtfB-ΔNΔV are also found in GH13 α-amylases, while the auxiliary domain IV is only found in GH70 glucansucrases. The other auxiliary domains V and N of *L. reuteri* 121 GtfB precede domain IV and likely extend away from the rest of the protein. Like in glucansucrases, the polypeptide chain follows a “U-course” forming first the N-terminal halves of domains IV, B and A, then domain C, and then the C-terminal halves of domains A, B and IV (Fig. 1A). In the catalytic domain A, the (β/α)₈ barrel is circularly permuted as in other GH70 enzymes; this permutation leads to a different order of the four homology regions (II-III-IV-I) than in GH13 α-amylases but keeps their relative spatial positioning. At the interface of domains A and B, a deep pocket is observed which is surrounded by residues strictly conserved in GH13 and GH70 enzymes, including the proposed catalytic residues (see below). By homology,
this pocket would harbor subsites -1 and +1 (nomenclature according to Davies_1997) (34) where catalysis takes place.

Fig. 1. (A) Overall structure of GtfB-ΔNΔV showing the two molecules in the asymmetric unit. From the N-terminal residue (Nt) the polypeptide forms the N-terminal segments of domains IV (yellow), B (green) and A (blue), then domain C (purple) and the C-terminal segments of domains A, B and IV towards the C-terminal residue (Ct). The location of the catalytic site at the interface of domains A and B is indicated. (B) Loop arrangement around the active site of GH13 α-amylases (left), 4,6-α-glucanotransferases (middle) and GH70 glucansucrases (right), at the border of the catalytic domain A (blue) and domain B (green), looking down the (β/α)8 barrel. In this view, acceptor substrate subsites are positioned left of the catalytic residues (shown in stick representation), and donor substrate subsites to their right. Loop A2 (red) is conserved in all three enzyme classes, but only in glucansucrases it forms a wall adjacent to the active site. Loop A1 (purple) and loop B (brown) are present only in 4,6-α-glucanotransferases and glucansucrases, but they are longer in the former, where they form a flap covering the space adjacent to the active site, forming a tunnel. In glucansucrases these two loops are shorter and leave the active site surrounding much more open. Structures used are Bacillus licheniformis α-amylase (PDB 1BLI) (32), Lactobacillus reuteri 121 GtfB-ΔNΔV (this work) and Lactobacillus reuteri 180 Gtf180-ΔN (PDB 3KLK) (12).
Active site architecture – comparison with α-amylases and glucansucrases

GtfB-ΔNΔV shows structural homology with both GH13 α-amylases and GH70 glucansucrases, with differences in the arrangement of loops around the active site (Fig. 1B). Regarding α-amylases, the structurally closest homologue of GtfB-ΔNΔV is the α-amylase from *B. licheniformis* (PDB 1BLI (32)); superposition of domains A+B results in an r.m.s.d. of 2.10 Å. Although in GtfB-ΔNΔV the (β/α)₈ barrel is circularly permuted and domain IV (together with other auxiliary domains) is inserted into domain B, a long groove similar to α-amylases resides at the interface of domains A and B. This groove bends at the position of the deep pocket; differences are mainly found in the loops surrounding this groove. First, GtfB-ΔNΔV has an extra helix-loop-helix subdomain between strand β4 and helix α5 of the domain A (β/α)₈ barrel; the loop connecting the helices (residues 1139-1151 in GtfB-ΔNΔV, hereafter referred to as “loop A1”) forms a large protrusion towards domain B, which is absent in α-amylases. Second, a long loop connecting the N-terminal segments of domains IV and B (residues 905-924, “loop B”) folds over the groove to contact the first helix of the helix-loop-helix subdomain with its tip (residues 917-918); notably, this loop has no equivalent in α-amylases. Third, the loop connecting strand β7 and helix α8 of the (β/α)₈ barrel (“loop A2”, residues 1430-1440), is slightly longer than in α-amylases and takes a different course, creating a slight protrusion adjacent to loop A1.

Comparison with the structurally closest glucansucrase homologue of GtfB-ΔNΔV, Gtf180-ΔN (from *L. reuteri* 180, PDB 3KLK) (12), shows that domains A+B can be superimposed with an r.m.s.d. of 0.84 Å; thus their overall structure is very similar (more so than for α-amylases). Nevertheless, differences between GtfB-ΔNΔV and Gtf180-ΔN are again observed in the three loops described before. Notably, loops A1 and loop B are longer in GtfB-ΔNΔV than in Gtf180-ΔN; together with a 20-30° rotation of the helix-loop-helix subdomain, this results in a much less open arrangement near the active site with respect to GS. In addition, loop A2 is 5 residues shorter in GtfB-ΔNΔV than in GS and therefore protrudes less from the protein surface.

Together, the above mentioned structural features of GtfB-ΔNΔV result in a unique active site architecture. While the “base” of the long α-amylase like groove is maintained, its different loop arrangement creates a tunnel which is
covered by loops A1 and B, and lined by loop A2. Superposition with α-amylases indicates that the tunnel would harbor donor subsites -2 and -3 (see also below). In contrast to the situation in GS, access to these donor subsites is not restricted, because of a shorter loop A2. Sequence alignment of the regions containing loops A1, A2 and B (Fig. S2B) shows that other 4,6-α-GTases likely have the same unique loop arrangement.

At the amino acid residue level, six of the seven strictly conserved GH13 residues surrounding subsites -1 and +1 also found in GtfB-ΔNΔV (the three proposed catalytic residues D1015, E1053 and D1125 plus residues R1013, H1124 and D1479); the seventh one is replaced by a glutamine residue (Q1484) like in glucansucrases (Fig. 2A and S1A). The arrangement of residues around subsites -1 and +1 thus is very similar to both that of GH13 α-amylases and GH70 GS, but also reveals a few differences, mainly near subsite +1. First, a tyrosine residue (Y1055) replaces the tryptophan residue found in almost all GS (and α-amylases); furthermore, the region Q1126-R1127-K1128-N1129 following the transition state stabilizing residue (D1125) is different from both corresponding regions in GS and in α-amylases (Figs. S1 and S2).

**Soaking experiments and modeling**

We performed several soaking experiments using either crystals of the wild-type enzyme or of the inactive mutant D1015N, and a range of glucooligosaccharides (DP2-DP7) with different α1→4 and/or α1→6 linkage patterns, as well as acarbose. Ligand electron density was only observed in donor subsites, with the reducing end glucosyl unit always occupying subsite -1 or -2. The donor subsites are bordered by several loops (from domain A and B) which, with respect to the apo structure, did not show large movements upon ligand binding (positional differences <0.5 Å). In contrast to donor subsites, acceptor subsites (+1, +2 etc.) never revealed enough electron density to model a bound carbohydrate. Although it seems that there is enough space for a maltooligosaccharide chain to occupy acceptor subsites, we noticed that access to these sites may be hindered by symmetry-related neighbouring protein molecules. Below, the soaking experiments with maltopentaose and maltohexaose are described.
Fig. 2. Active site of 4,6-α-glucanotransferase *L. reuteri* 121 GtfB-ΔNΔV; molecule A of the asymmetric unit is shown with domains A, B and IV in blue, green and yellow, respectively. (A) Apo GtfB-ΔNΔV showing residues surrounding the cleavage site; catalytic residues are labeled in bold. A glycerol molecule bound in the active site is shown with yellow carbon atoms. Dashed black lines indicate the tunnel structure. (B) Structure of the GtfB-ΔNΔV D1015N - maltopentaose complex with the pentasaccharide bound in subsites -1 to -5, and a maltose near remote subsite I (W790/W811 in domain IV) in the extension of the donor binding groove. The catalytic residues are indicated. (C) Structure of GtfB-ΔNΔV soaked with maltohexaose; the transglycosylation pentasaccharide product (6'-α-D-glucosyl-maltotetraose) bound in subsites -2 to -6 has an α1→6 linked glucosyl unit at the non-reducing end (indicated by the asterisk).

Soaking of maltopentaose (G5) in crystals of the inactive mutant GtfB-ΔNΔV D1015N revealed the pentasaccharide bound in subsites -1 to -5 in both molecules of the asymmetric unit (Figs. 2B and S1B). Comparison with ligand-bound α-amylase structures such as those from *B. halmapalus* (PDB: 2GJP) (28)
and *B. licheniformis* (PDB: 1BAG) (32) shows that G5 has a similar helical binding mode in GtfB-ΔNΔV. Notably, subsites -2 and -3 are shielded from the solvent by loops A1 and B. Most of the ligand-protein interactions are observed in subsites -1, -2 and -3; in particular, the reducing end glucosyl unit bound in subsite -1 has numerous hydrogen bond interactions with the conserved GH13/GH70 residues. In other donor subsites, residues from loops A1 and A2 (subsites -2, -3) and from loops in domain B (subsites -4, -5) provide most of the direct or water-mediated hydrogen bond interactions to the bound substrate.

In wild-type GtfB-ΔNΔV crystals soaked with maltohexaose (G6) we observed electron density for an isomalto/maltooligosaccharide occupying subsites -2 to -6 (Figs. 2C and S1C), adopting a similar conformation as G5. Interestingly, the non-reducing end glucosyl unit of the oligosaccharide bound in subsite -6 is α1→6-linked; it has a stacking interaction with Y1521, a residue which is conserved in all putative 4,6-α-GTases. In subsite -1, remaining positive electron density was visible, but not enough to convincingly model a complete glucosyl unit.

The soaking experiment with G5 revealed three more binding sites for carbohydrates. The first remote site is formed by two tryptophan residues in domain IV (W790 and W811) whose aromatic side chains form a hydrophobic stacking platform for the placed disaccharide maltose (Fig. 2B); remaining positive electron density suggested that the actual bound ligand is longer but too flexible to be fully observable. The bound maltose mostly stacks with W811 and lies more or less in the extension of the donor tunnel at a distance of about 16 Å from subsite -6, corresponding to 2 or 3 glucosyl units. Notably, the tryptophan pair is unique to and conserved in 4,6-α-GTases; in most GS only one of the two tryptophan residues is conserved, while α-amylases lack both due to the absence of domain IV. Two more remote sites reside in domain A, near the loop connecting β6 to α7 and the loop connecting β7 to α8, and near helices α2 and α3; the maltose and a maltotriose ligand placed in electron density respectively only make few interactions with protein residues. Since the mentioned secondary structure elements are either non-conserved or semi-conserved in α-amylases and glucansucrases, these remote sites were not considered further.
Chapter 6

Modeling

We modeled donor substrate binding across the cleavage site by superposition with the maltopentaose bound in subsites -3 to +2 in *B. subtilis* α-amylase (PDB 1BAG) (35). Only minor adjustments of the glycosidic torsion angles were needed to model the glucosyl moieties in positive subsites (+1 and +2) without clashes in GtfB-ΔNΔV (Fig. 3A). Thus, maltooligosaccharides likely can bind in GtfB-ΔNΔV in a similar way as observed in α-amylases. Residues that may interact with sugar units at subsites +1 and +2 are N1019, E1053 (the catalytic acid/base), H1056, and possibly K1128. Hydrophobic interactions may be provided by Y1055 and L971. We also modeled acceptor binding by first constructing a model for a covalent glucosyl-enzyme intermediate at residue D1015, based on the experimentally observed structure of a CGTase covalent intermediate (36). In agreement with their high conservation of residues surrounding subsite -1, the GtfB-ΔNΔV covalent intermediate model is very similar. The model for panose (O-α-D-glucopyranosyl-(1,6)-O-α-D-glucopyranosyl-(1,4)-D-glucose) bound in this covalent glucosyl-enzyme intermediate model of GtfB shows the trisaccharide in subsites +1 to +3, with the non-reducing end glucosyl unit at subsite +1 in a productive orientation that is different from the donor substrate model. Residues possibly involved in hydrogen bond interactions at subsites +1 and +2 include N1019, E1053 (the catalytic acid/base), H1056, N1126 and Y1079 while residues Y1055 and L971 may have hydrophobic interactions. The glucosyl moiety at subsite +3 would be only loosely bound.

Reaction specificity

To test for endolytic activity, GtfB-ΔN was incubated with a blocked 4-nitrophenyl-α-maltoheptaoside substrate (BPNPG7) in the presence of a maltase. This resulted in the release of *p*-nitrophenol as indicated by an increase in 405 nm absorption. In contrast, incubation with α-glucosidase alone did not result in such an increase (Fig. S3).
Fig. 3. (A) Model for donor substrate binding across the cleavage site: maltoheptaose (yellow carbon atoms) bound in subsites -5 to +2, based on the maltopentaose-bound crystal structure. Putative hydrogen bond interactions with the sugar units in subsites -1, +1 and +2 are indicated as red dashed lines; residue Y1055 may provide a hydrophobic stacking interactions at subsite +2. At the reducing end (indicated by an asterisk), there is space for longer donor substrates. (B) View (from a different angle) of the model for acceptor substrate binding; panose (yellow carbon atoms) is bound in subsites +1 to +3 of the covalent glucosyl-enzyme intermediate model of GtfB-ΔNΔV (cyan carbon atoms). In this orientation, the O6 atom of the non-reducing end glucosyl unit of panose can attack the C1 atom of the covalent intermediate (indicated by the arrow) to form an α1→6 glycosidic linkage. (C) Proposed reaction pathways of 4,6-α-GTases. R = reducing end, NR = non-reducing end; four of the six donor subsites are indicated, of which subsites -2 and -3 are in the tunnel. Newly formed bonds are indicated in red. In (a), donor binding involves a single subsite, or multiple ones (grey units); after the first transglycosylation step, the product can either diffuse away to give products of type 1 or 2, or shift into the tunnel to undergo a next reaction cycle (b), eventually giving products of type 3 or 4. Products of type 2 and 4 can become acceptors for subsequent reaction cycles. Eventually, most of the products will have isomaltooligosaccharide segments at their non-reducing end are obtained (5).
Genomic analysis

Fig. S8 shows the genomic location mapping of fully or partially sequenced 4,6-\(\alpha\)-glucanotransferases in their host bacteria. Using the \textit{L. reuteri} 121 GtfB sequence as query to search the NCBI database, 28 putative 4,6-\(\alpha\)-GTases (identity >50\%) were identified almost exclusively in \textit{Lactobacillus} strains.

In several \textit{Lactobacillus} strains, the 4,6-\(\alpha\)-GTase gene coexists with a GS gene in the same genome. Moreover, downstream or upstream of the 4,6-\(\alpha\)-GTase or GS gene, transposase encoding genes were observed. Especially for 4,6-\(\alpha\)-GTase genes, almost all are accompanied by a transposase encoding gene on either or both sides.

Phylogenetic analysis

A phylogenetic tree using the full-length sequences of characterized GH70 enzymes (Fig. S9), including 4,6-\(\alpha\)-GTases as well as GS, reveals that GS from Lactobacilli are homologous to those from \textit{Streptococcus}, \textit{Leuconostoc} and \textit{Weisella} species. Notably, the characterized 4,6-\(\alpha\)-GTases from \textit{L. reuteri} strains are more homologous to their GS counterparts (in the same organism) than to the GS from other species. This is supported by the results from a Protein-BLAST search of the NCBI database using the \textit{L. reuteri} 121 GtfB sequence; e.g. the highest homology scores are obtained for the glucansucrase GtfA from the same strain rather than for GS from other strains (data not shown; similar results were obtained in a \textit{vice versa} analysis using the \textit{L. reuteri} 121 GtfA sequence). The phylogenetic tree also suggests that \textit{L. reuteri} 4,6-\(\alpha\)-GTase and GS genes originated from a common ancestor gene.

A phylogenetic analysis using the catalytic cores (domains A and B) of bacterial GH13 enzymes, shows the clustering of homologues sequences into different subfamilies of GH13, in accordance with their different reaction specificities (Fig. S10). Including the rearranged catalytic cores (B-A-(C)-A-B→A-B-A) of GH70 GS and 4,6-\(\alpha\)-GTase enzymes in this analysis reveals that these two functionalities have a common ancestor, which is closely related to maltooligosaccharide-processing \(\alpha\)-amylases from GH13 subfamily 5, but phylogenetically much more distant from members of GH13 subfamilies 4 and 18 active on sucrose.
DISCUSSION

Unique structural features of GtfB-ΔNΔV

The crystal structure of GtfB-ΔNΔV is the first representative of the 4,6-α-glucanotransferase (4,6-α-GTase) subfamily within GH70. As expected from sequence alignments, the overall domain order and arrangement of GtfB-ΔNΔV resembles that of the GH70 glucansucrases (GS), with a circular permutation of the (βα)_8 barrel in the catalytic domain, and a similar spatial arrangement of the catalytic triad (residues D1015, E1053 and D1125). Surprisingly however, while in GS sugar binding subsites beyond -1 are blocked by the loop connecting β7 and α8 (12), in GtfB-ΔNΔV this loop is shorter, uncovering multiple donor subsites much like in GH13 α-amylases (37), while subsites -2 and -3 are occluded from the solvent by two long loops from domain A and B. Together the three loops, which seem to be specific for 4,6-α-GTase, create a tunnel-like structure capable of holding maltooligosaccharides, as exemplified by the structures of the complexes with maltopentaose and with 6⁴-α-D-glucosyl-maltotetraose. Thus, 4,6-α-GTases structurally have characteristics both from GH70 glucansucrases as well as from GH13 α-amylases.

Mechanism and mode of action of 4,6-α-GTases

Knowledge of the GtfB-ΔNΔV crystal structure and its complexes combined with the results of blocked substrate and acceptor reactions prompt us to propose an extended reaction scheme with a second mode of action (Fig. 3C). Notably, while the reaction scheme proposed by Leemhuis et al. (14) described an exo-α1→4 glycosidase activity and a concomitant transfer of single glucosyl units, we propose an additional endo-α1→4 glycosidase activity with transfer of multiple glucosyl units, supported by the finding of the ‘donor tunnel’ with subsites and by the observed release of 4-nitrophenol from BPNPG7. The extended scheme better explains the observed substrate and product specificity of GtfB and other 4,6-α-GTases (17).

The first step in the 4,6-α-GTase reaction is donor substrate binding across the cleavage site; modeling a maltooligosaccharide in GtfB by superposition with an α-amylase showed that the binding modes likely are very similar (Fig. 3A) such that maltooligosaccharides can enter the tunnel and occupy multiple donor
subsites, thus contributing to the shared substrate specificity between 4,6-α-GTases and α-amylases. On the other hand, differences in some of the residues surrounding subsite +1 between 4,6-α-GTases and GS (Fig. S1A) may explain why 4,6-α-GTases do not use sucrose as donor substrate like glucansucrases. For example, in the Gtf180-ΔN - sucrose complex, a conserved tryptophan (W1065) and a conserved glutamine (Q1140) residue provide direct hydrogen bonds to the fructosyl moiety of the substrate in subsite +1. Instead, at the equivalent positions, GtfB-ΔNAV and other 4,6-α-GTases have a tyrosine (Y1055), unable to provide an equivalent hydrogen bond, and a lysine (K1128), exchanging a positively charged side chain by a neutral one. Moreover, differences in residues of loops from domain B that line the active site result in a somewhat wider pocket in GtfB-ΔNAV. Together, although there is space for sucrose to bind (not shown), the shape and properties of subsite +1 likely result in a too low affinity for sucrose to be a substrate. Instead, maltooligosaccharides are the preferred donor substrates, and the fact that our soaking experiments did not reveal occupied acceptor subsites but only occupied donor subsites suggests that the latter have a higher affinity. Combined with the observation of endoglycolylytic activity of GtfB-ΔN with BPNPG7 (Fig. S3), we conclude that donor substrate binding can involve multiple donor subsites (Fig. 3C, grey squares). After donor substrate binding, cleavage of the α1→4 glycosidic bond and formation of the covalent intermediate occurs in the same way as in α-amylases, yielding covalent glucosyl- but also oligoglucosyl-enzyme intermediates. Thus, endolytic cleavage is a significant activity of GtfB. Regarding the transglycosylation step, linkage type specificity must originate from residues forming acceptor subsites. For example, in the glucansucrase Gtf180-ΔN - maltose complex (12), resembling α1→6 specificity, residues providing hydrogen bonds to the acceptor are mainly from homology regions II, III and IV (domain A); the importance of these residues in linkage specificity has been shown in several mutation studies (38-40). To investigate acceptor binding in GtfB-ΔNAV (Fig. 3B), we used panose because it is a better acceptor than maltotriose or isomaltotriose. The modeling results show that this trisaccharide can indeed bind in a productive orientation in subsites +1 to +3 for the formation of an α1→6 glycosidic linkage. Residues E1053 (homology region III, the catalytic base) and N1019 (homology region II, known to affect transglycosylation) seem critical for the orientation of the glucosyl unit in subsite +1, while other residues (N1129 from homology region IV, and Y1079) may help
orient the acceptor in subsites +2 and +3. Residue Y1055 (conserved in 4,6-α-GTases) may provide a non-specific hydrophobic stacking platform for acceptors at subsite +2, similar to the corresponding tryptophan in glucansucrases; the loss of activity (especially transglycosylation) in mutant Y1055G supports this hypothesis. The partially occluded active site of 4,6-α-GTases explains why these enzymes are less hydrolytic than α-amylases with a fully accessible binding groove. Finally, the presence of the tunnel-forming loops A1 and B near the cleavage/transglycosylation site prevents the 4,6-α-GTases to cleave or form branched oligosaccharide products, as opposed to glucansucrases.

After transglycosylation, in the initially proposed mechanism, products having a (single) α1→6 linked glucosyl unit bond at the non-reducing end would diffuse out of the active site via the acceptor side (Fig. 4, panel a). Based on the obtained crystal structures and observed reaction specificity of GtfB-ΔN(ΔV), we propose that such products alternatively can shift into the tunnel by one or more positions, and become donor substrates for the next reaction cycle (Fig. 4, panel B). Evidence for this comes from the pentasaccharide bound in subsites -2 to -6 observed in the G6 soaking experiment with active GtfB-ΔNΔV, representing a transglycosylation product with an α1→6 linked glucosyl unit at the non-reducing end, shifted by 5 positions, showing that movement of transglycosylation products in the donor direction is possible. Accordingly, mixed isomalto-/maltooligosaccharides, also representing transglycosylation products, are processed by GtfB-ΔN (Fig. S4); this suggests they occupy multiple donor subsites and can be cleaved in an endolytic fashion. The observation of bound saccharides at remote site I (W790/W811) in almost all our soaking experiments with other malto- and isomaltooligosaccharides (data not shown) leads us to speculate that hydrophobic stacking interactions at about 16 Å distance from donor subsite -6 provide extra affinity in the extension of the donor groove/tunnel. At subsite -6, residue Y1521 plays a similar role. The fact that mutants W790A/W811A and Y1521 have only a partly reduced enzyme activity (Table S2) suggests that affinity at remote donor subsites is functional but not essential (for transglycosylation). The importance of the tunnel for transglycosylation is further supported by the observation that mutant T920W has a heavily impaired transglycosylation activity, likely because the large aromatic side chain blocks its access. Finally, the tunnel feature explains why GtfB-ΔN selectively modifies
unbranched segments of starch while leaving the core structure intact. This specificity provides interesting opportunities for the industrial application of 4,6-α-GTase enzymes in starch-processing.

**Genomic and phylogenetic analysis of 4,6-α-GTase genes**

The occurrence of homologous genes in a species is believed to be the result of either horizontal gene transfer (orthologs) or gene duplication (paralogs). During evolution, such events are proposed to be driven by changes in habitat such as the availability of fermentable enzyme substrates (41). Upon orthologous or paralogous acquisition, the genes may be partly modified to adapt to the new environment, and afterwards may further evolve independently (42, 43). Earlier studies showed evidence that both horizontal gene transfer and gene duplication events have occurred for GH70 glucansucrases (GS) of oral *Streptococcus* and *Leuconostoc* species (29, 44, 45). The results of our genomic and phylogenetic analysis suggest that the same phenomena are involved in the evolution of 4,6-α-GTase genes in Lactobacilli. First, in most of the *Lactobacillus* species, the 4,6-α-GTase gene is flanked by a transposase gene on either or both sides (Fig. S8), as was observed earlier for several of their GS genes (8). Transposases are involved in horizontal gene transfer, and have been associated with intra- and interspecies genomic variation of oral bacteria such as Streptococci (46, 47). Interestingly, Hoshino *et al.* (41) stated that GS genes from Streptococci were acquired from Lactobacilli *via* transposase-mediated gene transfer, when these species encountered each other in fermented food. Thus, the presence of 4,6-α-GTase-adjacent transposase genes suggests that also 4,6-α-GTase genes may have undergone horizontal gene transfer during their evolution from one species to another. Second, we observed tandem-arranged GH70 genes encoding a 4,6-α-GTase and a GS in several of the *Lactobacillus* strains (Fig. S8); their high sequence and length similarity clearly suggest they are the result of gene duplications. This is further supported by the fact that nearby the 4,6-α-GTase / GS gene pairs, transposases genes are found which mediate gene duplication. Indeed, tandem GH70 genes are found in many bacteria, suggesting they readily duplicate. Together, we conclude that the evolution of 4,6-α-GTase and GS genes in lactic acid bacteria involved both horizontal gene transfer and gene duplication, leading to two closely related subfamilies with different reaction specificities.
The high sequence similarity between 4,6-α-GTase and GS gene pairs within a genome also is reflected in the results of a phylogenetic analysis of GH70 enzymes in lactic acid bacteria (Fig. S9). Earlier phylogenetic studies only included the conserved catalytic regions (13, 41) and suggested that the 4,6-α-GTase cluster constitutes a separate subfamily independent of glucansucrases. In contrast, our analysis of the full-length genes shows that 4,6-α-GTase genes are closely related to their GS tandem counterpart, more than to GS genes from other species. Moreover, in accordance with the evidence for gene duplication events described above, the GH70 4,6-α-GTase and GS genes clearly share a common ancestor. In a second phylogenetic analysis using rearranged catalytic cores of GH13 and GH70 enzymes (Fig. S10), we found that this common ancestor is closely related to α-amylases from GH13 subfamily 5 (GH13_5) rather than to sucrose-acting enzymes from subfamilies 4 and 18. Interestingly, enzymes from this subfamily are mainly found in Streptococcus, Bacillus and Exigobacterium species, correlating with the occurrence of GH70 enzymes (see CAZy database) (10). Enzymes from GH13_5 process maltooligosaccharides; combining this observation with the present-day substrate specificities of GH70 4,6-α-GTases and GS (maltooligosaccharides and sucrose, respectively), we propose that the common ancestor of GH70 4,6-α-GTases and GS processed maltooligosaccharides. The ability to utilize sucrose would then have been acquired later (in glucansucrases). Our conclusions from the genomic and phylogenetic analyses are supported by the crystal structure of GtfB-ΔNΔV. The domain organization of GtfB-ΔNΔV is shared with glucansucrase GtfA from the same organism, supporting that tandem 4,6-α-GTase and GS pairs within a single genome arose from gene duplication, while the active site architecture of GtfB-ΔNΔV resembles more the GH13 α-amylases. Thus, the 4,6-α-GTase subfamily can be regarded as an evolutionary intermediate between GH13 maltooligosaccharide-processing α-amylases and GH70 glucansucrases.

**Functional role of 4,6-α-GTases**

From the above described bioinformatic analyses we speculated how GH70 genes may have evolved. To further underpin this, we investigated EPS production in vivo in a Lactobacillus strain constitutively expressing both a 4,6-α-glucanotransferase and a glucansucrase. Provided with maltodextrins or sucrose as carbon source, L. reuteri 35-5 utilizes either its 4,6-α-glucanotransferase GtfB
or its glucansucrase GtfA enzyme for the synthesis of α-glucans with (α1→4 and α1→6) linkages (chapter 5). The extracellular polymers, present in the slimy halos around the colonies, also possess α1→4 and α1→6 linkages and have a similar size distribution. In short, both L. reuteri enzymes synthesize a related type of EPS, from different substrates.

Although the in vivo functional roles of glucansucrases in lactic acid bacteria remain to be understood, many glucansucrases have been proven to contribute to biofilm formation in vivo by producing α-glucans. Importantly, for Streptococcus mutans, which is considered the principle pathogen of dental caries, it was shown that the α-glucan produced by its glucansucrase facilitates the adherence of the bacterium to the tooth surface (4). In Lactobacillus reuteri, gene knockout results demonstrated that, without glucansucrase, biofilm formation is less effective (5). Notably, these studies were based on the supply of sucrose. Recently, we also show that in Lactobacillus reuteri 121 (strain 35-5) the uptake of starch, the most abundant dietary carbohydrate, also results in α-glucan production, and that the responsible enzyme is its 4,6-α-glucanotransferase GtfB (Bai et al., submitted). The α-glucans produced by L. reuteri on agar plates with sucrose or starch are structurally similar to that produced by its glucansucrase (GtfA). Therefore, also 4,6-α-GTases have great potential to contribute to biofilm formation by utilizing starch instead of sucrose. This finding is also in accordance with the above proposed ideas about gene evolution; both the 4,6-α-GTase and GS produce related α-glucans, but from different substrates, due to adaptive evolution driven by substrate variation.

**Evolutionary relationships between α-amylases, 4,6-α-glucanotransferases and glucansucrases**

Based on the structural observations in GtfB-ΔNΔV, and the genomic and phylogenetic analyses of 4,6-α-glucanotransferases we propose the following hypothesis for the evolutionary relation of this GH70 subfamily to GH13 α-amylases and GH70 glucansucrases. α-Amylases from GH13 subfamily 5 mainly degrade starch and starch-derived malto-oligosaccharides. They possess a long and fully solvent-accessible substrate binding groove with multiple donor subsites, due to the relatively short loops adjacent to this groove. As a consequence, these enzymes are mainly hydrolytic and produce glucose and/or short oligosaccharides.
Evolution of GH13_5 enzymes, involving domain insertion and circular permutation, leads to a common ancestor of family GH70 that acquired transglycosylation activity with maltooligosaccharides to produce polysaccharides instead. In order to achieve such different reaction specificity (from hydrolysis to polymerization), some of the loops (A1 and B) surrounding the donor half of the binding groove elongated, and partially shielded donor subsites from the solvent. This prevented diffusion of oligosaccharide reaction intermediates out of the binding groove, while keeping them available for the subsequent acceptor reaction (transglycosylation) and favoring a more processive mode of action. Whether they still had the 4-transglycosylation specificity observed in α-amylases, or already possessed the 6-transglycosylation specificity observed in 4,6-α-glucanotransferases is unknown. Evolution towards the glucansucrase subfamily, possessing a different substrate specificity (sucrose instead of starch), resulted from the elongation of another loop (A2) that blocked donor subsites beyond -1. Concurrently, loops A1 and B were no longer needed to shield the donor half of the binding groove, and could become shorter. Also, slight changes around subsite +1 likely were needed to prefer interaction with the fructosyl moiety in the donor reaction.

Fig. 4. Evolutionary relationships between α-amylases, 4,6-α-glucanotransferases and glucansucrases and the role of loops A1, A2 and B. The triangle indicates the catalytic site for cleavage and transglycosylation. The α-amylases that process maltooligosaccharides (white circles represent glucose units) have an open binding groove with multiple donor (-1 - -4) and acceptor (+1 - +3) subsites, lacking loops A1 and B, and are unable to synthesize polymers. The 4,6-α-glucanotransferases acquired this capability via elongation of loops B and A1 forming a tunnel at the donor half of the groove, allowing them to better capture the intermediate for subsequent transglycosylation and thus shift their product specificity towards synthesis of α-glucan polymers. Subsequently, a shift in substrate specificity was obtained by elongation of loop A2 which blocks donor subsites beyond -1; loops A1 and B became shorter; in this way, glucansucrases acquired the capability to use sucrose (grey circle represents fructose unit) for the synthesis of α-glucan polymers.
Conclusions

The crystal structure of GtfB-ΔNΔV is the first representative of the GH70 4,6-α-glucanotransferase subfamily, with an overall domain organization resembling GH70 glucansucrases, and an active site architecture reminiscent of GH13 α-amylases. A tunnel-like binding groove allows for a dual mode of action for 4,6-α-GTases involving both exo-specific and endo-specific cleavage/transfer reactions, and explains why GtfB only processes the outer (amylose) branches of starch substrates. The GtfB-ΔNΔV structure provides a starting point for (rational) mutation of 4,6-α-GTases to alter enzyme specificity, e.g. in starch-processing applications aiming at the production of prebiotic dietary fibers from a cheap substrate, using a GRAS bacterial strain. The GtfB-ΔNΔV crystal structure clearly supports our previous proposal that the 4,6-α-GTase subfamily is an evolutionary intermediate between GH70 and GH13 enzymes. The structural observations support our phylogenetic and genomic analyses, revealing the close relation between 4,6-α-GTases and GS, and suggesting that gene duplication as well as horizontal gene transfer events occurred during evolution. Loop mutations near the active site of a common ancestor likely lead to either an α1→6 specific transglycosylation activity (in 4,6-α-GTase) or to a different substrate specificity (in GS). We propose that the evolution of bacterial species such as Lactobacillus, using 4,6-α-glucanotransferases and glucansucrases to produce α-glucan exopolysaccharides with α1→4 and α1→6 linkages, was driven by changes in the availability of fermentable sugars in the human diet.
SI Materials and Methods

Construction of truncated GtfB-ΔNΔV

Truncation of GtfB lacking both the N-terminal variable domain and domain V was based on the alignment of the *Lactobacillus reuteri* 121 GtfB with glucansucrase sequences and on the crystal structures of Gtf180-ΔN glucansucrase (PDB entry 3KLK) from *Lactobacillus reuteri* 180 and DSR-E-ΔN (PDB entry 3TTO) from *Leuconostoc mesenteroides* NRRL B-1299. The gtfB gene fragment encoding GtfB (Uniprot entry Q5SBM0) amino acids 761–1619 was amplified by PCR using High Fidelity PCR enzyme mix (Thermo-Scientific, Landsmeer, The Netherlands) with pET15b-gtfB as template and the primers CHisFor-dNdVgtfB

\[
5'\text{-GATGCATCCATGGGACCAGGTACTTGGGAAAATATG}
\]

3' and CHisRev-dNdVgtfB

\[
5'\text{-CCTCCTTTCTAGATCTATTAGTGATGGGTGATGGTTGTTAGT}
\]

TTAATGAAAATTGCAGTTGG-3'. A nucleotide sequence encoding a 6×His-tag was fused in-frame to the 3' end of the gtfB-ΔN gene, using the reverse primer. The resulting PCR product was digested with NcoI and BglII, and was ligated into the corresponding site of pET15b. The constructed plasmid was transformed into *Escherichia coli* DH5α. The construct extracted from *E. coli* DH5α was confirmed by nucleotide sequencing (GATC Biotech, Glasgow, UK) followed by transforming into *Escherichia coli* BL21 star (DE3).

Site-directed mutagenesis of GtfB-ΔNΔV

The plasmid constructed above was used as the template for site-directed mutagenesis mutagenesis. Mutations were introduced by PCR with primers (see Table S1) using the Quik-Change site-directed mutagenesis kit (Stratagene. La Jolla, CA). The primers are purchased from Sigma-Aldrich. The PCR product was cleaned up with the PCR cleaning up kit (Sigma-Aldrich) followed by transformation into *Escherichia coli* BL21 star (DE3). Mutation of the plasmid was confirmed by gene sequencing in GATC Biotech (Glasgow, UK).

Substrate endo specificity

GtfB-ΔNΔV (1 μM) was incubated with Blocked 4-Nitrophenyl-α-Maltoheptaoside (BPNPG7, 10 mM) and α-glucosidase (yeast maltase,
Megazyme, 20 U/ml), at pH 6.0 (25 mM sodium citrate buffer) and 40 °C for 5 h; release of the 4-nitrophenyl group was determined by following the absorbance change at 405 nm.

GtfBΔN (1 μM) was incubated with 6\(^3\)-α-D-Glucosyl-maltotriosyl-maltriose (10 mM) in sodium acetate buffer (25 mM, pH 5.0, 1 mM CaCl\(_2\)) at 40 °C for 24 h. Reaction products were analyzed by HPAEC (see below). Maltooligosaccharides (G1-G8) and 6\(^3\)-α-D-glucosyl-maltriose were applied as references.

**Acceptor specificity**

GtfB-ΔNΔV (60 nM) was incubated with amylose (0.25%, w/v) in sodium acetate buffer (25 mM, pH 5.0, 1 mM CaCl\(_2\)) at 40 °C, in the presence or absence of acceptor substrates maltotriose, isomaltotriose and panose (10 mM). Reaction rates were determined using an amylose-iodine assay; transglycosylation factors were calculated as described by Bai et al. (2015).

**Characterization of mutants**

Enzymatic characterization of the mutants and the analysis of their products are shown in Table S2 and Fig. S6. Mutation of the catalytic nucleophile D1015 and of the general acid/base E1053 resulted in complete inactivation of the enzyme; activity was below the detection level. Enzyme activities of the single mutants T920W and Y1521A (tunnel), K1128A/N/H/W (subsite -1/+1), Y1055G (subsite +2) and of the double mutant W790A/W811A were determined using amylose V as a substrate; products were analyzed by TLC, \(^1\)H NMR and HPAEC.

Compared to wild-type GtfB-ΔNΔV (Fig. S6A), mutation of the aromatic residues at positions 790, 811 and 1521 to an alanine resulted in a 30-40% loss of total activity; however linkage specificity and product spectra were virtually unaffected.

Mutant T920W retained only 17% of total activity. The almost complete absence of α1→6 linkages in products indicates that this mutant has impaired transglycosylation activity and is virtually hydrolytic, accumulating glucose as the main product over time (Fig. S6B).
Removing the aromatic side chain in the Y1055G mutant results in a 60% loss of total activity and a 10-fold decrease in the relative amount of α1→6 linkages in its products. Compared to wild-type GtfB-ΔNΔV, this mutant released less free glucose over time, but hydrolyzed amylose to form maltooligosaccharides (Fig. S6C).

Mutation of K1128 to A, N, H or W resulted in elevated relative amounts of α1→4 linkages in the products and free glucose; accompanied by a slight decrease (K1128H) or marked increase (K1128A/N/W) of the total activity. All mutants produced similar products, acting endolytically and mainly converting amylose to maltooligosaccharides without forming transglycosylation products. The K1128W mutant produced lower amounts of glucose, and released maltooligosaccharides of different DP (Fig. S6D).

**Modeling**

*Donor substrate modeling.* The structure of the GtfB-ΔNΔV with maltopentaose (G5) bound in subsites -1 to -5 was superposed on that of *Bacillus subtilis* α-amylase complexed with G5 in subsites -3 to +2 (PDB 1BAG) (35), using their catalytic triads to improve the superposition; as a result, the glucosyl moiety at subsite -1 superimposed almost perfectly. The G5 in GtfB-ΔNΔV was then extended in acceptor subsites +1 and +2 by connecting two glucosyl moieties, and adjusting the glycosidic torsion angles such as to avoid clashes with protein residues. The result is shown in Fig. 3A.

*Acceptor substrate modeling.* The apo GtfB-ΔNΔV crystal structure was superposed on that of the *Bacillus circulans* strain 251 CGTase covalent maltotriosyl-enzyme intermediate (PDB: 1CXL) (36); a covalently linked glucosyl moiety was then ‘attached’ to the OD2 atom of the catalytic nucleophile D1015. A slight rotation of the OD2-C1 bond was applied to optimize hydrogen bond possibilities of the glucosyl unit with surrounding protein residues. Subsequently, the trisaccharide panose was modeled in acceptor subsites +1 to +3 guided by the following requirements. First, the O6 of the non-reducing end glucosyl unit should point towards the anomeric C1 atom of the covalent glucosyl-enzyme intermediate, at a reasonable distance for attack. Second, a hydrogen bond interaction to the same O6 should be provided by the catalytic base E1053; this interaction is required for the abstraction of the O6 proton in
order to activate the acceptor for the transglycosylation to take place. Furthermore, hydrogen bond interactions likely are provided to the non-reducing end glucosyl unit by residue N1019, as mutation of this residue is known to decrease the relative $\alpha1\rightarrow6$ transglycosylation activity (data not shown). The other two glucosyl units of panose were then placed such that, by adjusting glycosidic torsion angles, clashes with protein residues are avoided and enough space for longer acceptor substrates is available (at the reducing end). The result is shown in Fig. 3B.

**In vivo growth conditions**

*Lactobacillus reuteri* 35-5 was subcultured from stocks stored at -80 °C in 10 ml of sugar-free MRS medium (liquid culture or agar plate) supplemented with 1% (w/v) glucose. The fresh cultures were inoculated anaerobically in modified sugar-free MRS-agar plate with sucrose (5%, w/v), or maltodextrins (dextrose equivalent (DE) = 13-17, Sigma-Aldrich, US, 5%, w/v). One liter MRS-agar medium contained 10 g bactopeptone, 4 g yeast extract, 5 g sodium acetate, 2 g tri-ammonium citrate, 0.2 g MgSO$_4$$\cdot$7H$_2$O, 0.05 g MnSO$_4$$\cdot$7H$_2$O; and 1 ml Tween 80 and 15 g/l agar.

Growth of *L. reuteri* 35-5 on either sucrose or maltodextrins results in the formation of transparent colonies surrounded by a slimy halo (Fig. S7A). Previously, it has been shown that the EPS formed by *L. reuteri* 35-5 from sucrose is a reuteran containing both $\alpha1\rightarrow6$ and $\alpha1\rightarrow4$ linkages (49). Monosaccharide analysis (data not shown) and NMR analysis (Fig. S7B) showed that the EPS derived from the maltodextrin-grown cultures also has with $\alpha1\rightarrow6$ and $\alpha1\rightarrow4$ linkages (61 versus 39%, respectively).

In-gel PAS staining of enzymes (Fig. S7D) (Bai et al., submitted) revealed that, when incubated with sucrose substrate, the constitutively expressed glucansucrase GtfA (199 kDa) is associated with reuteran synthesis, in agreement with earlier studies (48). The minor band at ~120 kDa may represent GtfA-$\Delta$N, which is known to synthesise a reuteran very similar in size distribution and linkage type to that produced by the full-length enzyme (49). In case of incubation of the gel with maltodextrins, the band corresponding to GtfA is absent in the PAS stained gel, while a band corresponding to GtfB (179 kDa) is stained, indicating that not the glucansucrase GtfA, but the 4,6-$\alpha$-glucanotransferase GtfB is associated with
α-glucan synthesis when *L. reuteri* 35-5 is supplied with maltodextrins. SEC analysis of the EPS isolated from the plated cultures (Fig. S7C) reveals that the synthesized reuterans are approximately of the same size (35-40 MDa). When incubated with a mixture of sucrose and maltodextrins, both bands corresponding to GtfA and GtfB are associated with α-glucansynthesis. Notably, genomic data suggest that the only carbohydrate acting enzymes larger than 150 kDa are GtfA and GtfB (Gangoiti et al., in preparation).

**Extraction of extracellular polysaccharides (EPS)**

The extracellular polysaccharides (EPS) were washed out by 10 ml water from MRS-agar plates that had been incubated for 72 h at 37 °C with different carbon sources. Two volumes of cold ethanol (-20 °C) were added to the dissolved samples and kept at 4 °C overnight. The precipitate was harvested by centrifugation (12 000 ×g, 30 min, 4 °C) and re-dissolved in 1 volume of dd H₂O. The EPS were re-precipitated with 2 volumes of ethanol at 4 °C overnight. After centrifugation, the harvested precipitate was dissolved in water, and then dialyzed (10 kDa MWCO, Thermo Scientific) against water for 48 h with changes of dd H₂O in each 12 h (50). The EPS from followed by ethanol precipitation as described above, and analyzed by 1D 1H NMR (see below).

**Monosaccharide analysis**

EPS samples (0.5 mg) were dissolved in 200 µl Milli-Q water in a glass reaction tube. An equal volume of 4 M trifluoroacetic acid (TFA) was added, followed by hydrolysis for 4 h at 100 °C. Samples were dried under a flow of dry nitrogen and subsequently dissolved in 100 µl isopropanol and dried again by evaporation under dry nitrogen. Samples were dissolved in 1 ml dimethyl sulfoxide (DMSO) and used for HPAEC analysis (see below).

**Size exclusion chromatography (SEC)**

DMSO-LiBr (0.05M) was prepared by stirring for 3 h at room temperature followed by degassing for 15 min using an ultrasonic cleaner (Branson 1510, Branson, Danbury, CT). Samples were dissolved at a concentration of 4 mg/ml in DMSO-LiBr by overnight rotation at room temperature, followed by 30 min heating in an oven at 80 °C, obtaining clear sample solutions. The samples were cooled to room temperature and filtered through a 0.45-µm Millex PTFE
membrane (Millipore Corporation, Billerica, MA). The SEC system set-up (Agilent Technologies 1260 Infinity) from PSS (Mainz, Germany) consisted of an isocratic pump, auto sampler, an online degasser, an inline 0.2 μm filter, a refractive index detector (G1362A 1260 RID Agilent Technologies), viscometer (ETA-2010 PSS, Mainz) and MALLS (SLD 7000 PSS, Mainz). WinGPC Unity software (PSS, Mainz) was used for data processing. Samples (100 µl) were injected into a PFG guard column using an autosampler at a flow rate of 0.5 ml/min and DMSO-LiBr as eluent. The separation was done by three PFG-SEC columns with porosities of 100, 300 and 4000 Å. The columns were held at 80°C, the refractive index detector at 45 °C and the viscometer was thermostatted at 60 °C. A standard pullulan kit (PSS, Mainz, Germany) with molecular masses from 342 to 805000 Da was used. The specific RI increment value dn/dc was measured by PSS and is 0.072 (obtained from PSS company).

Nuclear magnetic resonance (NMR) spectroscopy
EPS samples isolated from *L. reuteri* 35-5 agar plate were analyzed by NMR. Resolution-enhanced 1D 500-MHz ¹H NMR spectra were recorded in D₂O (99.9 atm% D, Cambridge Isotope Laboratories, Inc.) on a Varian Inova Spectrometer (NMR Center, University of Groningen) at probe temperatures of 300 K. Spectra were recorded with 16k complex points, using a 5000 Hz spectral width. Data are expressed relative to internal acetone (δ¹H 2.225) Prior to analysis, samples were exchanged twice with D₂O with intermediate lyophilization and then dissolved in 0.6 ml D₂O. Spectra were processed using MestReNova 5.3 software (Mestrelabs Research SL, Santiago de Compostella, Spain), using Whittaker Smoother baseline correction and zero filling to 32k complex points.

High-performance anion-exchange chromatography (HPAEC)
Samples were injected onto a 4×250 nm CarboPac PA-1 column connected to a Dionex DX500 workstation (Dionex). Samples were run with a gradient of 30-600 mM NaAc in 100 mM NaOH (1 ml/min), and detected by an ED40 pulsed amperometric detector. A mixture with known concentrations of glucose, isomaltose, isomaltotriose, maltose, panose, maltotriose, maltotetraose, maltopentaose, maltohexaose and maltoheptaose was used as reference.
Thin-layer chromatography (TLC)
The TLC silica gel 60F254 plates (Merck) were run with butanol/acetic acid/water (2:1:1, v/v/v) as solvent. After running for 6 h, TLC plates were developed with 10% (v/v) H$_2$SO$_4$ and 2 g/l orcinol in methanol and then heated by oven at 100 °C for 30 min. Glucose, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose, and maltoheptaose were used as markers.
Table S1. Oligonucleotides used for site-directed mutagenesis of gtfB-ΔNΔV

<table>
<thead>
<tr>
<th>Mutation in GtfB-ΔNΔV protein</th>
<th>Sequence (5’→3’)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>W790A</td>
<td>GGTACTTAAGTTATACTGATGCTATGTCCCTTATGGCACAAG TTGTGGCCATAAGGACGATACGCATCGATATTAAGTAAAC</td>
</tr>
<tr>
<td>W811</td>
<td>GGTACAAAAACACTGCAATGGATGCGGCTCCATTACTGATGTATATTGTG AACATACACATCGTAATGGACGCGCATCCATTGCGGTTTGTACC</td>
</tr>
<tr>
<td>T920W</td>
<td>CGACCAGATAAAGTGGATGGATTTGATAGTGCACAGTC GACATTGATACCTATCCAATCCATCCACTTTTATCTGGTCA</td>
</tr>
<tr>
<td>D1015N</td>
<td>GTTTCGCCGATTTAATTGCTGCTGATA TATCAGCAGCATTAACTCGGAAACC</td>
</tr>
<tr>
<td>Y1055G</td>
<td>GTTATAATGAGGGCGGTCATTCTGGGTGCTGCTCGAGCAGCACCAGAATGGACCCTCCATTAAC</td>
</tr>
<tr>
<td>K1128A</td>
<td>CTAAACATGATCAACGAGCGAATTGTGATTAATAG CTATACAAATATCCATCGGATTGATGTTAG</td>
</tr>
<tr>
<td>K1128N</td>
<td>CTAAACATGATCAACGAAACATTTGATTAATAG CTATACAAATATTGGTTCATGGATGTTAG</td>
</tr>
<tr>
<td>K1128H</td>
<td>CTAAACATGATCAACGACAAACATTTGATTAATAG CTATACAAATATTGGTTCATGGATGTTAG</td>
</tr>
<tr>
<td>K1128W</td>
<td>CTAAACATGATCAACGATGGAATTTGATTAATAG CTATACAAATATTCCATCGGATTGATGTTAG</td>
</tr>
<tr>
<td>Y1521A</td>
<td>GCAAATCAAATTTTATTTCAGTCAACATGGGCGGAAATGG CCATTCCGCCACCAGTTGAGGTCCGAAAATAATTGATGTTG</td>
</tr>
</tbody>
</table>
Table S2. Effects of mutations on the reaction and product specificity of *L. reuteri* 121 GtfB-ΔNΔV.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Enzyme activity (U mg⁻¹)</th>
<th>Product mixture linkages (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>total b</td>
</tr>
<tr>
<td>wild-type</td>
<td>2.9 ± 0.17</td>
<td>0.37 ± 0.02</td>
</tr>
<tr>
<td>W790A/W811A</td>
<td>1.7 ± 0.02</td>
<td>0.27 ± 0.06</td>
</tr>
<tr>
<td>Y1521A</td>
<td>2.1 ± 0.17</td>
<td>0.36 ± 0.01</td>
</tr>
<tr>
<td>T920W</td>
<td>0.5 ± 0.01</td>
<td>0.27 ± 0.01</td>
</tr>
<tr>
<td>Y1055G</td>
<td>1.2 ± 0.15</td>
<td>0.08 ± 0.03</td>
</tr>
<tr>
<td>K1128A</td>
<td>5.4 ± 0.47</td>
<td>0.48 ± 0.01</td>
</tr>
<tr>
<td>K1128N</td>
<td>6.5 ± 0.09</td>
<td>0.68 ± 0.01</td>
</tr>
<tr>
<td>K1128H</td>
<td>2.1 ± 0.11</td>
<td>0.26 ± 0.01</td>
</tr>
<tr>
<td>K1128W</td>
<td>3.3 ± 0.08</td>
<td>0.11 ± 0.08</td>
</tr>
</tbody>
</table>

a Total and hydrolytic activities were determined using 0.125% w/vol amylose V as substrate. Assays were performed at 37 °C in 25 mM sodium acetate buffer, pH 4.7 containing 1 mM CaCl₂.
b Total activities were measured by the iodine-staining assay.
c Hydrolytic activities were determined by following the release of glucose in time using the GOPOD kit (15).
d The percentages of glucose, α₁→4 and α₁→6 glycosidic linkages were determined by ¹H NMR analysis of the product mixtures synthesized by incubations of amylose V (0.6% w/v) with the different enzymes (95 mU ml⁻¹).
Fig. S1. (A) Comparison of *L. reuteri* 121 GtfB-ΔNΔV (this work; molecule A of the asymmetric unit is shown, left) with the *L. reuteri* 180 Gtf180-ΔN - sucrose complex (right) (12). Blue represents domain A and green represents domain B; the tunnel in GtfB-ΔNΔV is indicated with dashed lines. Catalytic residues are labeled bold. In Gtf180-ΔN, the fructosyl moiety in subsite +1 has hydrogen bond interactions with residues Q1140 and W1065 and, via a water-mediated hydrogen bond network, with residues N1411 and D1458. In GtfB-ΔNΔV, the corresponding residues are K1128, Y1055, L1390, but the fourth residue is absent; moreover, the side chains of Y1055 and L1390 are unable to make hydrogen bonds. Finally, differences in residues of domain B cause the active site pocket to be wider in GtfB-ΔNΔV. (B) GtfB-ΔNΔV D1015N - maltopentaose complex (stereo view) showing direct and water-mediated hydrogen bond interactions of the pentasaccharide with protein residues in donor subsites -1 to -5. Catalytic residues are labeled bold; residues from loops A1, A2 and B are labeled purple, red and brown, respectively. (C) Stereo view of the pentasaccharide observed in the GtfB-ΔNΔV - G6 soak showing hydrogen bond interactions, in donor subsites -2 to -6.
Fig. S2. Sequence alignment of 4,6-α-glucanotransferase (4,6-α-GTase), glucansucrases (GS) and α-amylases (α-Amy); the latter were aligned separately. Alignments were made using ESPript v3.0 (31). Numbering corresponds to the sequences of LrGtfB and BIAMY. Top: conserved motifs I-IV. The seven conserved α-amylase superfamily residues are indicated by triangles (▲), including the catalytic residues (NU = nucleophile, A/B = general acid/base, TS stab = transition state stabilizing residue). Other residues near subsites -1 and +1 are indicated by stars (*). Bottom: loop regions around the active site: loops A1, A2 and B (grey shading); α-amylases were aligned only for loop A2 since these enzymes do not have an equivalent to loops A1 and B. LrGtfB = *Lactobacillus reuteri* 121 GtfB, LrGtfML4 = *L. reuteri* ML1 GtfML4, LrGtfW = *L. reuteri* DSM 20016 GtfW, LrGtf180 = *L. reuteri* 180 Gtf180, LrGtfA = *L. reuteri* 121 GtfA, LrGtfO = *L. reuteri* ATCC 55730 GtfO, LrGtfML1 = *L. reuteri* ML1 GtfML1, SmGtf-SI = *Streptococcus mutans* Gtf-SI, LmDSRE = *Leuconostoc mesenteroides* NRRL B-1299 DSRE, BIAMY = *Bacillus licheniformis* α-amylase, BhAMY = *Bacillus halmapalus* α-amylase, GsAMY = *Geobacillus stearothermophilus* α-amylase, BsAMY = *Bacillus subtilis* α-amylase.
Fig. S3. Release of $p$-nitrophenol from BPNPG7 (structure shown in inset) incubated with α-glucosidase (red) or α-glucosidase + GtfB-ΔN (blue) as measured by the absorbance change at 405 nm. The exo-acting yeast maltase cannot cleave the substrate; in contrast, with the addition of GtfB-ΔNΔV, $p$-nitrophenol is released. This indicates that GtfB-ΔN cleaves $\alpha1\rightarrow4$ linkages in an endo fashion, releasing the blocked end of the substrate; subsequently the maltase hydrolyzes the intermediate products until it reaches the non-reducing end to release $p$-nitrophenol.
Fig. S4. HPAEC analysis of 6'3-α-D-glucosyl-maltotriosyl-maltrose (10 mM) processed by GtfB-ΔN (1 μM) in sodium acetate buffer (25 mM, pH 5.0, 1 mM CaCl₂) at 40 °C for 24 h (black line). Glucose (G1), fructose (F), isomaltose (G2), sucrose (S), maltose (G2), panose (P), maltose (G2), maltotriose (G3), maltotetraose (G4), maltopentaose (G5), maltohexaose (G6), maltoheptaose (G7), maltooctaose (G8), and 6'3-α-D-glucosyl-maltotriose were applied as references (colored lines).
Fig. S5. TLC analysis of the product mixtures obtained from the incubations of amylose V (0.6% w/vol) with 95 mU ml⁻¹ of wild-type GtfB-ΔNΔV, and its mutants: Lane 1, wild-type GtfB-ΔNΔV; lane 2, W790A-W811A; lane 3, Y1521A; lane 4, T920W; lane 5, Y1055G; lane 6, K1128A; lane 7, K1128N; lane 8, K1128H; and lane 9, K1128W. The reaction mixtures were incubated at 37 °C and pH 4.7 during 48 h. Std, standards; G1, glucose; G2, maltose; G3, maltotriose; G4, maltotetraose; G5, maltopentaose; G6, maltohexaose; G7, maltoheptaose; Pol, polymer.
Fig. S6. HPAEC-PAD profiles of the oligosaccharide mixtures formed upon the incubation of amylose V with (A) GtfB-ΔNΔV, (B) T920W, (C) Y1055G, and (D) K1128W for t = 10 min, 30 min, 24 h (pH 4.7, 37 °C). The identity of peaks was assigned using commercial oligosaccharide standards (G1, glucose; G2-G7, maltose to maltoheptaose; iso-G2, isomaltose; iso-G3, isomaltotriose; Pa, panose).
Fig. S7. *In vivo* EPS formation by *L. reuteri* 35-5. (A) Growth of *L. reuteri* 35-5 on MRS agar supplemented with different carbon sources, forming slimy or non-slimy colonies. (B) 1D $^1$H NMR analysis of the EPS extracted from the slimy colonies on agar plates shows that their composition is similar with respect to the glycosidic linkage types present ($\alpha1\rightarrow4$ and $\alpha1\rightarrow6$). (C) SEC analysis of the EPS extracted from the slimy colonies on agar plates. (D) Periodic Acid-Schiff (PAS) stained SDS-PAGE gels (adapted from Bai et al., submitted) of whole cell proteins of *L. reuteri* 35-5 grown in the presence of either sucrose, sucrose plus maltodextrins or maltodextrins, revealing the association of the glucansucrase GtfA with sucrose-derived EPS and of the 4,6-$\alpha$-glucanotransferase GtfB with maltodextrins-derived EPS.
Fig. S8. Genomic mapping of 4,6-α-glucanotransferase (4,6-α-GTase) and glucansucrase genes. Included are gtf180 from *Lactobacillus reuteri* 180, gtfB and gtfA from *L. reuteri* 121, gtf106B and gtf106A from *L. reuteri* TMW1.106, gtfML4 and gtfML1 from *L. reuteri* ML1, 4,6-α-GTase and GS genes from *L. reuteri* TMW1.656, and 4,6-α-GTase and GS genes from *L. salivarius* GI-24, other putative 4,6-α-GTase genes from *L. delbrueckii*, *L. plantarum*, *L. sanfranciscensis*, *L. mucosae*, and *Pediococcus pentosaceus*. 
Fig. S9. Unrooted phylogenetic tree of characterized 4,6-α-glucanotransferases (green box) and glucansucrases from *Lactobacillus reuteri* strains (pink box) and other lactic acid bacteria. Alignments and dendrogram construction were carried out using full length protein sequences (including N-terminus, domain V and IV) with MEGA version 6, using the maximum likelihood method. The bar corresponds to a genetic distance of 0.1 substitution per position (10% amino acid sequence difference).
**Fig. S10.** Phylogenetic tree of family GH13 and GH70. The bacterial protein sequences were all characterized based on CAZY database ([www.cazy.org](http://www.cazy.org)) (10). Domains A and B of GH70 members were extracted and rearranged according to the (non-circularly permuted) GH13 domain arrangement. The sequences classified into different subfamilies as represented by different colors.
Acknowledgements

The study was financially supported by the Chinese Scholarship Council (to YB), the University of Groningen and by the TKI Agri&Food program as coordinated by the Carbohydrate Competence Center (CCC-ABC; www.cccresearch.nl).

Reference


characterization of the glucan and fructan exopolysaccharides synthesized by the *Lactobacillus reuteri* wild-type strain and by mutant strains. Appl Environ Microb 65:3008-3014.
