Molecular Characterization of a Novel Glucosyltransferase from Lactobacillus reuteri Strain 121 Synthesizing a Unique, Highly Branched Glucan with α-(1→4) and α-(1→6) Glucosidic Bonds

S. Kralj,1,2 G. H. van Geel-Schutten,1,3 H. Rahaoui,4 R. J. Leer,3 E. J. Faber,5 M. J. E. C. van der Maare,1,6 and L. Dijkhuizen1,2,*

Centre for Carbohydrate Bioengineering, TNORUG,1 and Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Haren,2 Department of Applied Microbiology and Gene Technology3 and Department of Food Microbiology TNO Nutrition and Food Research,4 Zeist, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, Utrecht,2 and Carbohydrate Technology Department, TNO Nutrition and Food Research, Groningen, The Netherlands

Received 24 January 2002/Accepted 8 June 2002

Lactobacillus reuteri strain 121 produces a unique, highly branched, soluble glucan in which the majority of the linkages are of the α-(1→4) glucosidic type. The glucan also contains α-(1→6)-linked glucosyl units and 4,6-disubstituted α-glucosyl units at the branching points. Using degenerate primers, based on the amino acid sequences of conserved regions from known glucosyltransferases (gft) genes from lactic acid bacteria, the L. reuteri strain 121 glucosyltransferase gene (gftA) was isolated. The gftA open reading frame (ORF) was 5,343 bp, and it encodes a protein of 1,781 amino acids with a deduced Mr of 198,637. The deduced amino acid sequence of GTFA revealed clear similarities with other glucosyltransferases. GTFA has a relatively large variable N-terminal domain (702 amino acids) with five unique repeats and a relatively short C-terminal domain (267 amino acids). The gftA gene was expressed in Escherichia coli, yielding an active GTFA enzyme.

With respect to binding type and size distribution, the recombinant GTFA enzyme and the L. reuteri strain 121 culture supernatants synthesized identical glucan polymers. Furthermore, the deduced amino acid sequence of the gftA ORF and the N-terminal amino acid sequence of the glucosyltransferase isolated from culture supernatants of L. reuteri strain 121 were the same. GTFA is thus responsible for the synthesis of the unique glucan polymer in L. reuteri strain 121. This is the first report on the molecular characterization of a glucosyltransferase from a Lactobacillus strain.

Many lactic acid bacteria employ large extracellular enzymes, glucosyltransferases (GTFs) (EC 2.4.1.5; common name, glucansucrases), for the synthesis of high-molecular-weight α-glucans from sucrose. Moreover, low-molecular-weight oligosaccharides are produced in the presence of suitable acceptor molecules.

The GTF enzymes of oral streptococci and the dextransucrases (DSRs) and alternansucrases (ASRs) from Leuconostoc mesenteroides strains have been studied in the most detail. All GTFs from lactic acid bacteria share a common structure and are composed of four distinct domains: their N-terminal end starts with (i) a signal peptide of 32 to 34 amino acids, which is followed by (ii) a highly variable stretch of 123 to 129 amino acids, (iii) a highly conserved catalytic or sucrose binding domain of about 1,000 amino acids, and (iv) a C-terminal glucan binding domain of about 500 amino acids, which is composed of a series of tandem repeats (30).

Amino acid sequence comparisons revealed that GTFs possess a (β/α)n barrel structure similar to that of glycoside hydrolases of family 13. This family includes, for instance, α-amylase and cyclodextrin glycosyltransferase (CGTase) (43). The core of the proteins belonging to this family comprises eight β-sheets alternated with eight α-helices. In GTFs, however, this (β/α)n barrel structure is circularly permuted (12, 27). Therefore, GTFs are classified in family 70 of glycoside hydrolases (http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html). Some GTF enzymes, such as dextransucrase from L. mesenteroides strain B-512F, catalyze the formation of linear glucans containing mostly α-(1→6) linkages (dextran). Other types synthesize dextrans with α-(1→2), α-(1→3), or α-(1→4) branches (16, 30). ASR from L. mesenteroides strain NRRL B-1355 synthesizes a glucan with alternating α-(1→6) and α-(1→3) glucosidic bonds (3). Recently, a GTF of L. mesenteroides strain NRRL B-1355 that synthesizes a glucan containing α-(1→2) glucosidic linkages has been characterized (39).

There are few reports, however, about glucan synthesis in lactobacilli (13, 21, 37, 44, 45). A biochemical and molecular characterization of the enzyme(s) responsible for glucan synthesis in lactobacilli has not been reported.

In previous studies we have isolated a strain of Lactobacillus reuteri that is capable of producing both a fructan and a glucan. Depending on the carbon source in the culture medium, the GTF responsible for the synthesis of this polysaccharide ma-
its glucan product. The **gtfA** medium with sucrose (100 g liter⁻¹) or in MRS-s medium (i.e., MRS Food Research, Zeist, The Netherlands) were cultivated anaerobically at 37°C and 35°C (both obtained from the culture collection of TNO Nutrition and Food Research, Zeist, The Netherlands) was used for expression of the **gtfA** gene in *Lactobacillus* strains containing recombinant plasmids. E. coli strain 121 total DNA was isolated according to the method in reference 32. Plasmid DNA of *L. reuteri* strain 121 was isolated using a Wizard Plus SV plasmid extraction kit (Promega). Molecular techniques. General procedures for cloning, E. coli transformations, DNA manipulations, and agarose gel electrophoresis were carried out as described previously (35). Restriction endonuclease digestions and ligations with T4 DNA ligase were performed as recommended by the enzyme suppliers (New England Biolabs, Beverly, Mass.; Roche Biochemicals, Basel, Switzerland). Primers were obtained from Eurogentec, Seraing, Belgium. Sequencing was performed according to the method in reference 41. DNA was amplified by PCR on a DNA Thermal Cycler 480 (Perkin-Elmer, Boston, Mass.) using ampliTaq DNA polymerase (Promega) and a modiﬁcation of the method in reference 36. DNA was ampliﬁed by PCR using a Qiagen (Hilden, Germany) gel extraction kit following the instructions of the supplier.

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, media, and growth conditions.** *L. reuteri* strains 121 and 35-5 (both obtained from the culture collection of TNO Nutrition and Food Research, Zeist, The Netherlands) were cultivated anaerobically at 37°C in MRS medium (Difco, Franklin Lakes, N.J.) (9) or in MRS-s medium (i.e., MRS medium with sucrose [100 g liter⁻¹] instead of glucose [20 g liter⁻¹]) and 30°C in LB medium (44, 45). This work describes the molecular characterization of a *Lactobacillus* GTF gene (**gtfA**), expression of **gtfA** in *Escherichia coli*, and the characterization of its glucan product. The **gtfA** gene encodes a novel type of GTF, synthesizing a highly branched glucan with a unique structure containing α-(1→4) and α-(1→6) linkages.

**Isolation of DNA.** *L. reuteri* strain 121 total DNA was isolated according to the method in reference 32. Plasmid DNA of *L. reuteri* strain 121 was isolated using a modification of the methods described previously (2, 7). Prewarmed (37°C) fresh MRS broth (10 ml) was inoculated with 200 μl of an overnight culture and incubated for 2.5 h at 37°C. Cells were harvested by centrifugation and washed with 2 ml of sterile STE buffer (0.1 M NaCl, 10 mM Tris-HCl, 1 mM EDTA [pH 8]). After centrifugation, the pellet was resuspended in 380 μl of solution I (0.5 M sucrose, 50 mM Tris-HCl, 1 mM EDTA [pH 8]) containing lysozyme (2 mg/ml; Sigma, St. Louis, Mo.) and 6.6 U of mutanolysin (Sigma). After 1.5 h of incubation at 37°C, 50 μl of solution II (0.25 M EDTA, 50 mM Tris-HCl [pH 8]) and 30 μl of solution III (20% sodium dodecyl sulfate, 50 mM Tris-HCl, 20 mM EDTA [pH 8]) were added and the suspension was mixed. NaOH (30 μl of 3 M solution) was added, which was followed by 50 μl of 2 M Tris-HCl and 72 μl of 5 N NaCl. After extraction with equal volumes of phenol and chloroform, the DNA was precipitated with ethanol as described previously (35).

Plasmid DNA of *E. coli* was isolated using the alkaline lysis method (6) or with a Wizard Plus SV plasmid extraction kit (Promega).

**Strains, plasmids, media, and growth conditions.** E. coli strains containing recombinant plasmids were cultivated in LB medium (22), and *E. coli* TOP 10 (Invitrogen, Carlsbad, Calif.) were used as hosts for cloning purposes. Plasmids pCR2.1-TOPO (Invitrogen), pCR-XL-TOPO (Invitrogen), and pEMBL8 (11) were used for cloning of the **gtfA** gene in *E. coli* strains. Plasmid pET15b (Novagen, Madison, Wis.) was used for expression of the **gtfA** gene in *E. coli* BL21 Star(DE3) (Invitrogen). *E. coli* strains were grown aerobically at 37°C in LB medium (4). *E. coli* strains containing recombinant plasmids were cultivated in LB medium with the appropriate antibiotic (ampicillin [100 μg ml⁻¹] or kanamycin [50 μg ml⁻¹]). Agar plates were made by adding 1.5% agar to the LB medium; when appropriate, X-Gal (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) (20 μg ml⁻¹) was added.

**Identification and nucleotide sequence analysis of the GTF gene.** An overview of the isolation strategy of the **gtfA** gene is given in Fig. 1. The first part of the GTF gene was isolated by amplification of chromosomal DNA of *L. reuteri* strain 121.
FIG. 2. Overview of primers and restriction sites used for cloning of the gtfA gene in the expression vector pET15B (Novagen).

with PCR using degenerate primers (GTPr1, 5′-GAYAAKWSNAASYNRT NGTNSARGC-3′, and GTPr2, 5′-ANRTCNCTRTARTANACNGK-3′ [Y = T or C; K = G or T; W = A or T; S = C or G; R = A or G; N = A, C, G, or T]) based on conserved amino acid sequences present in the catalytic core, as deduced from the GTF genes of Streptococcus downei (gfs), Streptococcus mutans (gfsC), S. downei (gfs), Streptococcus salivarius (gfsK and gfsM), and dsrA of L. mesenteroides (14, 18, 20, 31, 38, 40). An amplification product with the expected size of about 660 bp was obtained and cloned in E. coli JM109 using the pCR2.1 vector. Analysis of its nucleotide sequence (659 bp [Fig. 1A]) confirmed its GTF identity. The 659-bp amplified fragment was used to design the primers GTPr1 (5′-GGGCTTTCAGATGCAACTAATCGTTGGGG-3′) and GTPr2 (5′-CATATGCGGTTCGTTGCATTCCATGAAAC-3′). This information was used to design the diverging primers GTPr7 (5′-ACATGCTGGATTCC-3′) and GTPr8 (5′-CCTATTTGAGACAGCATCTGAGTCCATACC-3′), containing the corresponding sites of pBluescript II SK(−) and GTPr13 (5′-GATGGTGATG-3′ [in boldface type] and BamHI (i) sites) and GTPr14 (5′-GAYAAKWSNAAKSYNRT AATGATAAGCGCCATACAG-3′) was designed to screen a partial gene library of S. mutans GS5 (AAA85889), and GTXC of L. mesenteroides NRRL B-1355 (CAA77898), GTFB of S. sobrinus OMZ176 (CAB65910), and GTXF of L. mesenteroides NRRL B-1299 (AAB40875), ASR of L. mesenteroides NRRL B-1355 (CAB65910), ISRT of L. mesenteroides NRRL B-512F (BAA09527), and GTFA of L. reuteri strain 121; (ii) GTFs of Streptococcus sobrinus OMZ176 (BAA14241), GTFI of S. downei Mfc28 (AABA2096), GTFIs of S. sobrinus OMZ176 (BAA02976), GTFB of S. mutans GSS (AA88588), and GTXD of S. mutans GS5 (AA88589); and (iii) GTFG of Streptococcus gordoni (AAC34838), GTXJ of Streptococcus oralis ATCC 10557 (BAA95201), GTFS of S. sobrinus ATCC 25975 (AAC41413), GTXH of S. salivarius ATCC 25973 (AAC41412), GTXN of S. salivarius ATCC 25975 (AAC05165), GTFF of S. mutans GSS (AA26895), GTFE of S. salivarius ATCC 25975 (CAG78900), GTFT of S. sobrinus OMZ176 (D13928), and GTXG of S. salivarius ATCC 25975 (CAG77898). Amino acid sequences were aligned first within each group. The complete alignment was performed by aligning groups i to iii with each other. Tree construction was performed using TreeCon 1.3b (no correction for distance estimation, 10 bootstrap samples, and the neighbor-joining algorithm were used).

Preparation of E. coli cell extracts. Cells of E. coli BL21 Star(DE3) harboring p15gtf were harvested by centrifugation (10 min at 4°C at 10,000 × g) after 16 h of growth. The pellet was washed with 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2 and 1% (vol/vol) Tween 80, and the suspension was centrifuged again (10 min at 4°C at 10,000 × g). Pelleted cells were resuspended in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2, 1% (vol/vol) Tween 80, and 5 mM β-mercaptoethanol. Cells were broken by sonication (seven 15-s pulses of 7 μ at 30-s intervals). Cell debris and intact cells were removed by centrifugation for 20 min at 4°C at 10,000 × g, and the resulting cell extract (supernatant) was used in the enzyme assays.

Enzyme assays. Using E. coli cell extracts or L. reuteri strain 121 grown on MRS-s culture supernatant as the source of enzyme, GTF activity was measured by determining the release of fructose from sucrose at 37°C in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2, and 1% (vol/vol) Tween 80, and the suspension was centrifuged again (10 min at 4°C at 10,000 × g). Pelleted cells were resuspended in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2, 1% (vol/vol) Tween 80, and 5 mM β-mercaptoethanol. Cells were broken by sonication (seven 15-s pulses of 7 μ at 30-s intervals). Cell debris and intact cells were removed by centrifugation for 20 min at 4°C at 10,000 × g, and the resulting cell extract (supernatant) was used in the enzyme assays.

Sample preparation. Supernatants of cultures of L. reuteri strain 121 grown on MRS-s culture supernatant as the source of enzyme, GTF activity was measured by determining the release of fructose from sucrose at 37°C in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2, and 1% (vol/vol) Tween 80, and the suspension was centrifuged again (10 min at 4°C at 10,000 × g). Pelleted cells were resuspended in 50 mM sodium acetate buffer, pH 5.5, containing 1 mM CaCl2, 1% (vol/vol) Tween 80, and 5 mM β-mercaptoethanol. Cells were broken by sonication (seven 15-s pulses of 7 μ at 30-s intervals). Cell debris and intact cells were removed by centrifugation for 20 min at 4°C at 10,000 × g, and the resulting cell extract (supernatant) was used in the enzyme assays.

(ii) GTF activity staining. After separation in the second dimension, GTF activity was identified as described previously (44), using periodic acid-Schiff staining. The same gel was silver stained (Genotech, St. Louis, Mo.) to identify other proteins on the gel, thereby facilitating the identification of the GTF spot on the blot (see below).

Amino acid sequencing. A second gel obtained after 2D-PAGE, identical to the GTF activity stained gel, was blotted according to the method in reference book. 
26 onto a polyvinylidene difluoride membrane (Roche Biochemicals) with a semidyry electrobolter (Ancos, Hoesjy, Denmark). The polyvinylidene difluoride membrane was stained with Coomassie brilliant blue R-250. The N-terminal amino acid sequence of the protein band which corresponded with the activity band obtained by periodic acid-Schiff staining was determined on a Perkin-Elmer ABI 476A automated sequencer using Edman degradation (NAPS Protein Sequencing and Peptide Mapping Laboratory, Vancouver, Canada).

Characterization of the glucans produced. Glucans were produced by incubating the enzyme preparations overnight, using the conditions described above under “Enzyme assays.” Glucans produced by _L. reuteri_ strain 121 and strain 35-5 (producing only the wild-type glucan and not the fructan) and glucans produced with the GTF expressed in _E. coli_ were isolated by precipitation with ethanol (44). Nuclear magnetic resonance (NMR) spectroscopy and methylation analysis were performed as described earlier (44). The molecular weights of the glucans were determined by high-performance size exclusion chromatography (HPSEC) coupled online with a multianlge laser light scattering (MALLS) and a differential refractive index detection (Schambeck SDF). The HPSEC system consisted of an isocratic pump, an injection valve, a guard column, and a set of two SEC columns in series (Shodex SB806MHQ column and TSK gel 6000PW). A Dawn-DSP-F (Wyatt Technology, St. Barbara, Calif.) laser photometer (HeNe [\(\lambda = 632.8 \text{ nm}\] equipped with a K5 flow cell, thermostatted by a Peltier heating system, was used as a MALLS detector. Samples were filtered through a 0.45-μm-pore-size filter (MILLLEX), and the injection volume was 220 μL Na_2SO_4 (0.1 M) was used as eluent at a flow rate of 0.8 ml min⁻¹. Pullulan and dextran samples with _M_w_ ranging from 4 \times 10^5 to 2 \times 10^6 were used as standards. Determinations were performed in duplicate.

**Nucleotide accession number.** The nucleotide sequence of _gfA_ has been assigned accession no. AX306822 by GenBank.

## RESULTS

**Isolation and nucleotide sequence analysis of the putative _L. reuteri_ strain 121 GTF gene.** Based on sequence homology between conserved regions located in the catalytic core of different _gf_ genes of gram-positive bacteria, degenerate primers were designed and used for PCR with chromosomal DNA of _L. reuteri_ strain 121 as the template. A single fragment of 659 bp (Fig. 1) was obtained, and sequence analysis confirmed its _gf_ identity. Southern hybridization of chromosomal DNA of _L. reuteri_ strain 121 with this 659-bp amplified PCR fragment, followed by washing under nonstringent conditions, revealed one hybridizing fragment. Plasmid DNA of _L. reuteri_ did not show hybridization.

In subsequent steps a total of 9,310 bp was obtained and sequenced (Fig. 1). One complete open reading frame (ORF) and two partial ORFs were located on this compiled sequence: ORF1 (5,343 bp [Fig. 1]), which encodes a putative GTF (GTFA); ORF2 (2,403 bp), upstream of ORF1; and ORF3 (786 bp), downstream of _gfA_. The deduced amino acid sequence of ORF2 showed homology with GTFs, whereas the deduced amino acid sequence of ORF3 did not show significant homology to any protein present in databases. The _gfA_ gene encodes a putative protein of 1,781 amino acids, with a deduced molecular weight of 198,637 and a pI of 5.04. It was preceded by a putative ribosomal binding site (GAAGGAG), localized 6 bp upstream from the ATG start codon. According to the consensus promoter sequences described previously for lactobacilli (34), a potential promoter sequence in the sequence upstream of _gfA_ (42 bp from the start codon) could be identified, with a −35 sequence (TTGAAA) separated by 19 bp from a −10 sequence (TATAAT).

Two inverted repeats were located 61 bp downstream from the _gfA_ termination codon. These repeats could form a stem (20 bp)-loop (12 nucleotides) secondary structure with a ΔG value of −20.4 kcal mol⁻¹, followed by a series of thymidine residues, suggesting a rho-independent transcription termination signal (5).

**Amino acid sequence alignments of _L. reuteri_ GTFA with other GTFs.** Alignment of the amino acid sequence of GTFA with other GTFs, performed with BLAST (1), revealed clear similarities. The highest similarity (46% identity and 59% similarity within 1,261 amino acids) at the amino acid level was found with ASR from _L. mesenteroides_ NRRL B-1355. The putative protein structure of GTFA was similar to that of other GTFs containing (i) an N-terminal signal sequence of 38 amino acids, (ii) a relatively large variable N-terminal domain of 702 amino acids, (iii) a catalytic domain of 774 amino acids, and (iv) a C-terminal domain of 267 amino acids (Fig. 3).

The deduced N-terminal amino acid sequence of GTFA contained a putative secretion peptide with a predicted signal peptidase cleavage site between amino acids 38 and 39 (http://www.cbs.dtu.dk/services/SignalP/). To confirm this cleavage site, the _L. reuteri_ strain 121 enzyme was purified by 2D-PAGE and subjected to N-terminal sequence analysis. The first 13 amino acids were identified as DQOVOQASTLQDO; except for the 10th residue the sequence was identical to that of the deduced amino acid sequence following the predicted cleavage site. 2D-PAGE experiments also confirmed the predicted _M_ₐ and pl of GTFA. Within the deduced N-terminal variable region of GTFA, a series of five repeating units was found. These repeating units, designated RDV, were on average 41 amino acids long, separated on average by 71 amino acids (Fig. 4). These repeats have never been seen in other GTF enzymes.
and showed no significant homology to any protein motifs present in the databases.

The putative catalytic domain of GTFA showed high similarity (about 45% identity and 60% similarity) to other known streptococcal and Leuconostoc GTF enzymes. However, not all of the conserved amino acids found in the other GTFs were found in the deduced amino acid sequence of GTFA. Particularly in the region downstream of Asp\(^{1024}\), 3 out of 10 conserved amino acids were not found in GTFA (Fig. 5). One of the conserved amino acid substitutions in this region of GTFA, Ile\(^{1029}\), was also found in amylosucrase, a GTF from Neisseria polysaccharea synthesizing an \(\alpha\)-(1→4) glucan (10).

The relatively short C-terminal domain of GTFA contains four YG-repeating units according to the definition of Giffard and Jacques (19) and seven YG-repeating units which are less conserved (Fig. 6).

**Dendrogram.** Construction of a dendrogram (Fig. 7), based on the complete amino acid sequences of different GTFs of lactic acid bacteria, revealed that GTFA of \(L.\) reuteri strain 121 is distinct from other GTFs known. Again, GTFA is most closely related to the ASR of \(L.\) mesenteroides NRRL B-1355 (3).

**Analysis of the glucans produced by \(L.\) reuteri and \(E.\) coli containing p15gtf.** Extracts of \(E.\) coli(p15gtf) cells and supernatants of sucrose-grown cultures of \(L.\) reuteri incubated with sucrose both produced high-molecular-weight glucans. Using HPSEC-MALLS, the average molecular weight of the glucan produced by \(L.\) reuteri strain 35-5 was determined to be \(4 \times 10^7\) (±5% SEM), whereas that of the glucan produced by \(E.\) coli harboring p15gtf was \(8 \times 10^7\) (±5% SEM).

The \(^1\)H NMR spectra of the glucans produced by the recombinant GTF present in \(E.\) coli and the GTF enzyme in \(L.\) reuteri strain 35-5, producing only the glucan, were virtually identical (Fig. 8). Comparison of both \(^1\)H NMR spectra with that of potato starch (17) showed that both glucans consist of \(\alpha\)-(1→4)- and \(\alpha\)-(1→6)-linked glucopyranosyl units. The identical nature of the glucans was confirmed by methylation analysis (Table 1).

**DISCUSSION**

This work reports the molecular characterization of the first \(Lactobacillus\) gene (gtfA) encoding a GTF enzyme (GTFA). A detailed analysis showed that GTFA produces a unique soluble glucan in which the majority of the linkages are of the \(\alpha\)-(1→4) glucosidic type. The glucan also contains \(\alpha\)-(1→6)-linked glucosyl units and 4,6-disubstituted \(\alpha\)-glucosyl units at the branch points. Expression of gtfA in \(E.\) coli yielded an active GTF synthesizing the same highly branched soluble glucan.

GTFA showed high similarity with streptococcal and Leuconostoc GTFs. Similar to other GTFs, GTFA contains an N-terminal signal sequence, a variable N-terminal domain, a catalytic core, and a C-terminal domain. Striking features of GTFA are its relatively large variable domain (702 amino acids), which contains five distinct unique repeats (RDV repeats) (Fig. 4), and its relatively short C-terminal domain (267 amino acids). ASR of \(L.\) mesenteroides NRRL B-1355 also possesses three N-terminal repeats (24), but these do not show homology to the N-terminal repeats found in GTFA. The exact function of the variable domain is unknown. The variable domain of GTFI from \(S.\) downei Mfc28 contains no repeats and is five times smaller than the GTFA domain. Deletion of the GTFI variable domain yielded a mutant enzyme which retained function (28).

Based on alignments with other GTFs from lactic acid bacteria (30), the putative catalytic residues in \(L.\) reuteri strain 121 GTFAs are Asp\(^{1024}\), Glu\(^{1061}\), and Asp\(^{1133}\). The putative calcium-binding site is Asp\(^{996}\), and the putative chloride binding site is Arg\(^{1022}\). Addition of \(Ca^{2+}\) increases enzyme activity and stability (data not shown). Five residues may play a role in the
binding of acceptor molecules and the transfer of the glucosyl residue. These are GTFA residues Asp1027, Asn1028, Asp1062, and Trp1063 (Fig. 5). The fifth amino acid, possibly playing a role in acceptor binding or transfer of the glucosyl residue, a Ser in other GTFs (except for ASR of L. mesenteroides and DSRA of L. mesenteroides), was replaced by Asn1064 in GTFA.

The C-terminal domain of GTFA, which consists of 267 amino acids, is shorter than corresponding domains in other GTFs (~500 amino acids). The C-terminal domain of streptococcal and Leuconostoc GTFs consist of a series of different tandem repeats, which have been divided into four classes: A, B, C, and D repeats. These repeats exhibit high similarity to the repeats found in the glucan binding protein from S. mutans as well as the ligand binding domains in Clostridium difficile toxin A and the lysins from Streptococcus pneumoniae (18, 47).

DSRS from L. mesenteroides NRRL B-512F also contains, in addition to A and C repeats, N repeats, which have not been identified in streptococcal GTFs. ASR from L. mesenteroides NRRL B-1355 contains a single A repeat and distinct short repeats D(X)4APY (24). Within the A, B, C, and D repeats, a repeating unit designated YG can be distinguished (19). The repeats in the C-terminal domain of other GTFs were not found in GTFA. Instead, four YG repeating units and seven less-conserved YG repeats could be identified (Fig. 6).

The highest overall homology of GTFA at the amino acid level was found with ASR from L. mesenteroides NRRL B-1355 (3), which is responsible for the synthesis of an alternan with 50% (1-3)6 and 50% (1-3)3 linkages, and with DSRS from L. mesenteroides NRRL B-512F (29, 46), which synthesizes a dextran with 95% (1-3)6 and 5% (1-3)3 linkages. Homology of GTFA with other GTFs was highest in the highly conserved putative catalytic domain, which had roughly the same size and structure as the corresponding domains of other GTFs. However, not all the conserved residues were found in the L. reuteri strain 121 GTFA. Relatively many differences with amino acids conserved in other GTFs were found directly downstream of the putative catalytic Asp1024 (Fig. 5). This region constitutes the α/β-barrel 4 of the enzymes of family 13.
of glycoside hydrolases (27). The domain directly downstream of the catalytic Asp^{1024} contains the conserved amino acids Asp-Ala-Val-Asp-Asn in other GTFs. In CGTase these residues constitute part of the acceptor binding site (residues Asp^{229}-Ala-Val-Lys-His^{233} in *Bacillus circulans* 251 CGTase), responsible for the stereospecific positioning of the molecule accepting the glucosyl unit (25). The structure of this acceptor site determines the type of glucosidic bond formed (41). In the corresponding region of GTFA, Pro^{1026} is found in a position where a conserved Val is found in other GTFs (Fig. 5). Compared with Val, the presence of Pro causes a more rigid protein structure, which may have a direct effect on the type of glucosidic bonds formed in the glucan synthesized by the enzyme. The presence of the Pro^{1026} residue could therefore be part of the explanation for the unique structure of the glucan with \( (1\rightarrow3) \) and \( (1\rightarrow6) \) bonds produced by GTFA. The conserved Val is present in amylosucrase (Fig. 5), a GTF synthesizing \( (1\rightarrow4) \) bonds. However, immediately downstream of this Val, the conserved Asp-Asn residues are replaced by Ala-Phe (10). The following amino acid in amylosucrase is Ile, which is also present at that position in GTFA, whereas in other GTFs a conserved Val is found (Fig. 5). This also suggests that the above-mentioned region downstream of the catalytic Asp^{1024} may be of influence on the type of bonds being formed. Therefore, Pro^{1026} and Ile^{1029} of GTFA are likely targets for site-directed mutagenesis experiments.

The partial ORF upstream of *gtfA* (ORF2 [Fig. 1]) may encode a second GTF enzyme in *L. reuteri* strain 121. However, the N-terminal amino acid sequence of GTF purified from culture supernatants of *L. reuteri* strain 121 was the same as the deduced N-terminal amino acid sequence of the *gtfA* gene, and

![A](image1.png)

![B](image2.png)

**FIG. 8.** 500-MHz \(^1\)H NMR spectra of the glucan produced by *L. reuteri* strain 35-5 GTFA present in culture supernatant (A) and by *E. coli* GTFA cell extracts (B), recorded in D\(_2\)O at 80°C. Chemical shifts are given in parts per million relative to the signal of internal acetone (\( \delta = 2.225 \)).

<table>
<thead>
<tr>
<th>Type of glucosyl unit</th>
<th>% Methylation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>L. reuteri</em> strain 121</td>
</tr>
<tr>
<td>Glep-(1→4)Glep-(1→)</td>
<td>24</td>
</tr>
<tr>
<td>Glep-(1→4)Glep-(1→2)</td>
<td>42</td>
</tr>
<tr>
<td>Glep-(1→4)Glep-(1→7)</td>
<td>22</td>
</tr>
<tr>
<td>Glep-(1→4)Glep-(1→11)</td>
<td>12</td>
</tr>
</tbody>
</table>
the $M_1$ and $p$ of the purified enzyme were the same as those predicted from the nucleotide sequence of the gfa4 gene. Furthermore, the $^1H$ NMR spectra of the glucans produced by the L. reuteri GTF-IA present in culture supernatant and by the E. coli GTF-A in cell extracts were virtually identical (Fig. 8). This, combined with the results of the methylation (Table 1) and the molecular weight determinations of the glucans, shows that the E. coli GTF-A and the L. reuteri enzyme present in culture supernatants synthesize the same glucan with a unique structure: a highly branched glucan containing $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ bonds. Therefore, it is concluded that the gfa4 gene encodes the active GTF of L. reuteri strain 121.

In the future, ORF2 upstream of gfa4 will be characterized in further detail and GTF-A will be characterized molecularly and biochemically.

ACKNOWLEDGMENTS

S.K. and G.H.V.-G.-S. contributed equally to this work.

We thank Jos van der Vossen, Egbert Smit, Dick van den Berg, and Peter Terpstra for help with the isolation of gfa4; Isabel Capron for the HPSEC-MALLS analyses; Joost Uitdehaag, Bart van der Veen, and Hans Leemhuis for discussions concerning the structure-function relationships of family 13 glycoside hydrolases; and Sacha van Hijn for his contribution in the construction of the dendrogram.

REFERENCES


