Comparison of novel GH 68 levansucrases of levan-overproducing *Gluconobacter* species

Frank Jakob, Daniel Meißner, Rudi F. Vogel

Lehrstuhl für Technische Mikrobiologie, Technische Universität München, Freising, Germany

Abstract

Gluconobacter species are capable of incomplete oxidations which are exploited in food biotechnology. Levans isolated from exopolysaccharide (EPS)-overproducing *Gluconobacter* species are promising functional compounds for food applications. Fructan production strongly depends on the corresponding fructosyltransferases (Ftfs), which catalyze the formation of these polymers from sucrose. Therefore, we characterized novel Ftfs from three EPS-overproducing food-grade strains, *i.e. Gluconobacter sp.* TMW 2.767 and *Gluconobacter sp.* TMW 2.1191 isolated from water kefir, and *Gluconobacter cerinus* DSM 9533T isolated from cherries. Several PCR techniques, including degenerate gradient temperature PCR, modified and standard inverse PCR, modified site-finding PCR and modified single primer PCR, were used to finally detect complete open reading frames coding for Ftfs. The prospective *ftf*-gene sequences were heterologously expressed in *Escherichia coli* Top 10. *E. coli* transformants harboring one of the three different *ftf*-genes produced polysaccharides from sucrose in contrast to the *E. coli* wildtype. Each of the heterologously expressed proteins encoded a levansucrase, catalyzing the formation of β-(2→6)-linked fructose polymers, which corresponded to our previous analyses about the chemical nature of the isolated polymers formed by these *Gluconobacter* strains. Structurally, these enzymes belong to the glycoside hydrolase 68 family (GH 68), sharing the typical modular topology of levansucrases from gram-negative bacteria. In conclusion, we could identify novel active levansucrases, which can be used for *ex situ* (enzymatic catalyses) or *in situ* (fermentation) production of functional fructan polymers by *Gluconobacter* strains in food and other applications.

Introduction

Most microbial fructosyltransferases (Ftfs), such as levansucrases and inolusucrases, belong to the glycoside hydrolase (GH) 68 family according to the similarity of their amino acid sequences.[1] Whereas levansucrases catalyze the formation of β-(2→6)-linked β-D-fructofuranosyl polymers, inolusucrases catalyze the formation of β-(2→1)-linked β-D-fructofuranosyl polymers. The major substrate for fructosyltransferases (Ftfs) is sucrose. As a first step, the disaccharide is bound to the active center of the corresponding enzyme and cleaved, which results in the release of glucose. The remaining bound fructose moiety can then be transferred to the growing polymer chain with a terminal sucrose molecule as starting precursor for polysaccharide formation. Alternatively, water can serve as acceptor and reach the catalytic center of the corresponding Ftf after cleavage of donor sucrose, which results in the release of initially bound fructose and the complete hydrolysis of sucrose.[2] Fructanpolymers produced by microbial levansucrases can have a degree of polymerization (DP) of up to 10^6 fructose molecules with occasionally occurring β-(2→1)-linked branches.[3] Additionally, these enzymes can catalyze the formation of short chain fructooligosaccharides (FOS, *e.g.* kestose, nystose) with a DP ranging from 3-10 carbohydrate moieties. Whereas Ftfs of gram-negative bacteria are usually secreted in the environment acting as typical extracellular enzymes, the C-terminal domains of Ftfs of gram-positive bacteria contain motifs involved in cell-wall-anchoring.[4] The production of exopolysaccharides can protect microorganisms against environmental stress factors such as desiccation and osmotic pressure.[5] Furthermore, they can be assimilated as nutrients, used as electron acceptors if the inevitably produced remaining sugar monomer is fructose, used to form biofilms or serve as virulence factors during colonization of the host.[3]

Strains of some species of acetic acid bacteria (AAB) have been shown to produce levan[6-10] or to possess genes coding for levansucrases.[11,12] One of the most extensively studied levansucrases is expressed by the N_2-fixing endophyte *Gluconacetobacter (Ga.) diazotrophicus*. Besides resolving the crystal structure of this enzyme,[13] determining its enzyme kinetics[14] and, furthermore, describing biological functions for this species,[5] the use of this enzyme for producing levans and FOS, which can be used as calorie-free, non-cariogenic sweeteners and prebiotics in functional foods,[15] was patented.[16]

We recently demonstrated that the addition of high molecular levans produced by *Gluconobacter sp.* TMW 2.767 and *Gluconobacter cerinus* DSM 9533T to breads caused an increased volume, softer crumb and retarded staling.[17] This is the first reported practical application for fructans isolated from AAB. EPSs were isolated from these strains because of their capability of synthesizing large amounts of these polysaccharides from sucrose. Whereas strains of different *Gluconobacter* species are already applied in several (food) biotechnological processes, which are constantly being optimized to get higher product yields of low molecular weight fermentation metabolites,[18] the production of functional polysaccharides by *Gluconobacter* strains for *e.g.* food applications is not exploited. In contrast, homopolysaccharides such as dextrans and levans from lactobacilli have been shown to be promising functional compounds for bakery products.[19,20] Moreover, inulins are already used as supplements or macronutrient substitutes in different foods while increasing dietary fiber contents and exhibiting positive nutritional properties (mainly bifidogenic effects).[21] Beyond that, new EPS types, which can either be isolated with lower costs or be produced *in situ* in high amounts from food-grade strains, have the potential of replacing commonly used food additives such as xanthan and guar gum via *in situ* production enabling clean label products.[22]

Since fructan formation strongly depends on the corresponding enzyme catalyzing its production, the aim of this study was to identify novel levansucrases of levan overproducing *Gluconobacter* strains. For this purpose, we first selected strains capable of synthesizing high amounts of fructan (two isolates from water kefir and *G. cerinus* DSM 9533T) and afterwards detected open reading frames coding for their putative levansucrases using different PCR approaches.

Materials and Methods

Strains, media and growth conditions

Gluconobacter sp. TMW 2.767 and TMW 2.1191 (isolated from water kefir;[23] unpublished data), *G. cerinus* DSM 9533T (isolated from cherries) and *G. oxydans* 621H (DSM 2343) were cultivated aerobically for at least 24 h at 30°C in modified, liquid gluconate media (20 g/L sodium gluconate, 3 g/L yeast extract, 2 g/L peptone, 3 g/L glycerol, 10 g/L mannitol, pH adjusted to 6.0). For solid media 20 g/L agar, for induction of EPS production 80 g/L sucrose were added. *Escherichia coli* strains DH5α (containing either pBADMycHis (Invitrogen) or pBluescript II SK (Stratagene) vectors) and Top 10 (Invitrogen, for heterologous expression) generally were cultivated aerobically at 37°C in lysogeny broth (LB) medium. To select plasmid harboring *E. coli* cells LB media were supplemented with 100 μg/mL ampicillin (Amp100). For induction of polysaccharide synthesis by recombinant *E. coli* cells. 60 g/L sucrose and 2 g/L arabinose (induction) were added to solid LB Amp100 media.

General molecular techniques

Genomic DNA was isolated following the instructions of the E.Z.N.A. Bacterial DNA Kit (Omega Bio-tek, Norcross, USA). *E. coli* plasmid DNA was purified using the peqGOLD Plasmid Miniprep Kit (peqlab, Erlangen, Germany). Restriction endonuclease digestions, dephosphorylation with shrimp alkaline phosphatase (SAP) and ligations with T4 DNA ligase were performed as recommended by the suppliers (Fermentas, St. Leon-Roth, Germany). Unknown DNA sequences were amplified using Taq-DNA-Polymerase (Qbiogene, USA). For amplification of known DNA sequences, as well as 16S rRNA genes, the proof-reading Phusion High-Fidelity DNA Polymerase (Finnzymes, Finland) was used. When specific primer sets were put in the reaction mixture, PCR cycle conditions were routinely performed according to the manufacturers' instructions. Oligonucleotide primers for PCR reactions were

obtained from MWG Biotech AG (Ebersberg, Germany). Potentially complex amplification products resulting from degenerate or modified single-primer PCR were sequenced by Sequiserve (Vaterstetten, Germany). All other PCR products were sequenced by GATC Biotech (Konstanz, Germany). Preparative DNA isolations from agarose gels were performed with the peqGOLD Gel extraction Kit (Peqlab, Erlangen, Germany), PCR products were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA).

Identification of levansucrase encoding genes

Primers used for PCR reactions are listed in Table 1. An identification scheme for initially unknown genes, which includes used primers, methods and restriction endonucleases, is shown in Figure 1. Open reading frames (ORFs) were identified as described below.

Core sequence detection

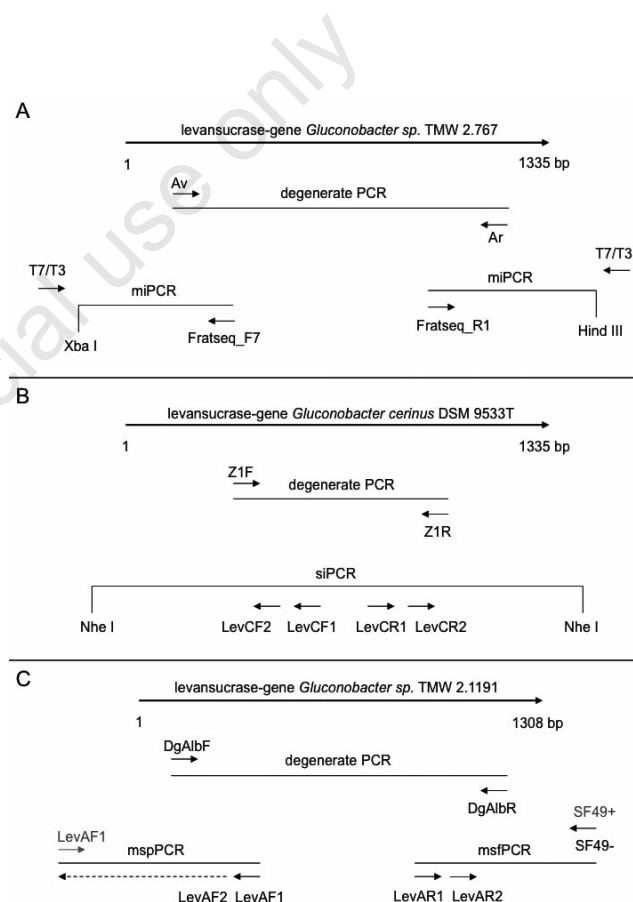To identify core sequences of the unknown *ftf*-genes, degenerate

Figure 1. Schematic representation of the methods used to obtain unknown ORFs coding for levansucrases of (A) *Gluconobacter sp.* **TMW 2.767, (B)** *Gluconobacter cerinus* **DSM 9533T and (C)** *Gluconobacter sp.* **TMW 2.1191. ORFs were identified in the chronological order A→C. An initial alignment including the amino acid sequences of levansucrases of** *Zymomonas mobilis* **(AAA27696),** *Gluconobacter oxydans* **621H (AAW60646) and** *Gluconacetobacter xylinus* **(BAA93720) was generated to create degenerate primers Av and Ar. Further degenerate primers were constructed involving newly identified sequences to enhance the possibility for positive amplicons (3.2). Gray primer names and arrows indicate unspecific binding. Primer LevAF2 was solely used as sequencing primer (dotted arrow). The used restriction endonucleases are named below vertical lines.**

PCR reactions were performed using a gradient cycler (Eppendorf, Hamburg, Germany), which allowed optimization of primer annealing using different annealing temperatures for multiple samples containing equal reaction mixtures. Typical PCR conditions were as follows: initial denaturation at 94°C for 3 min; 35 cycles: denaturation at 94°C for 30 s, annealing for 1 min±30 sec (temperatures were chosen in 1°C steps approximately in the range of the melting point T_m (primer) - 10°C; time was varied if necessary according to the results), and extension at 72°C (duration adapted to the expected size of the PCR product); final elongation at 72°C for 10 min. Generally, reaction mixtures were prepared according to the suppliers' instructions. Clear bands of the expected size were cut out of the gels, following preparative DNA extraction and subsequent sequencing as described above.

Identification of parts flanking the identified core regions

Based on core sequence information, specific inverse primers (5'→ 3', Table 1) were designed targeting the direction of the missing part of the corresponding *ftf*-gene. For standard inverse PCR (siPCR), genomic DNA was digested with a restriction endonuclease, which could not cut in the known sequence. DNA subsequently was religated and used as template for siPCR. In a modified inverse PCR (miPCR) approach, digested DNA was cloned into plasmid pBluescript II SK, which was ini-tially treated with the same restriction enzyme. The resulting constructs were afterwards used as template for miPCR using inverse and plasmid primers (T7/T3), which flank the vector's multiple cloning site. Modified single primer PCR (mspPCR) was performed using solely one inverse primer in a temperature gradient PCR reaction. PCR products of equal size, that showed a constantly increasing intensity towards the primer's T_m after staining, were preparatively purified and sequenced with a second inverse primer, which again binds specifically in the expected amplicon. For modified site-finding PCR (msfPCR)[24] random primer SF49+ (containing a 6×N anchor) and an inverse primer were used for an initial PCR reaction. The resulting product mixture served as template for a second PCR reaction with random primer SF49- (without anchor) and a second inverse primer. Positive PCR amplicons were subsequently sequenced and annotated to the corresponding core sequences.

Cloning and heterologous expression of complete Ftfs

Newly identified, putative *ftf*-gene sequences and the levansucrase gene of the genome sequenced strain *G. oxydans* 621 H were cloned into expression vectors of the pBADMycHis series (Invitrogen, Darmstadt, Germany), which possess a C-terminal 6×histidine tag for further purification experiments. Primers for amplification of the corresponding *ftf*-genes contained 5'-overhangs with restriction sites (underlined in Table 1), allowing in-frame cloning into the respective multiple cloning site. Restriction, SAP treatment and ligation were performed according to the

Table 1. Primers used for gene identifications and cloning procedures (restriction sites are underlined).

Strain(s)	Primer	Sequence (5'→ 3')	Use
G. sp. TMW 2.767			
	Av	GCNGAYGCNATGAAR	degenerate PCR
	Ar	NGTNGGNGCNAGNGTNCC	degenerate PCR
	Fratseq_F7	CACATCGTTGGACATGACC	miPCR
	Fratseq_R1	CCCTGATGGTGTGTATGG	miPCR
	T7	GTAATACGACTCACTATAGGGC	miPCR
	T3	AATTAACCCTCACTAAAGGG	miPCR
	FtffFratBAD_F	GCCCATGGCTAATGCTATTTCCAGCCGAATTC	PCR + cloning
	FtfFratBAD_R	GCTCTAGAAGGGCGCGAACGTCATAG	PCR + cloning
G. cerinus DSM 9533T			
	Z1F	ACRACGACVATGCCG	degenerate PCR
	Z1R	CSGTCTGGTCATTSACGC	degenerate PCR
	LevCF1	CGCAGCGCACCACTGTCCCA	siPCR
	LevCR1	CTGGCCCGGTGTGTCTCCGG	siPCR
	LevCF2	AAATCTTGGCGTTGGTCTTA	siPCR
	LevCR2	CTTGCACCGACGGTTCGTCT	siPCR
	FtfCerBAD_F	GCCTCGAGTAATGCTATTTCCTCCCAGTC	PCR + cloning
	FtfCerBAD_R	GCTAAGCTTGGCTCTGACGTCATACGCC	PCR + cloning
G. sp. TMW 2.1191			
	DgAlbF	GATCCGACGACGACMATG	degenerate PCR
	DgAlbR	CCGVTCACCRTCSAGAAC	degenerate PCR
	LevAF1	TGGCATCTGCCAGAATG	mspPCR
	LevAR1	CATGTCGTCTTCCAGGAC	msfPCR
	LevAF2	GTAGGACGGAACAGACTCG	sequencing
	LevAR2	CTGACCTATCTGTTCACGATC	msfPCR
	SF49+	GCACTGCTATGCTTACTGGNNNNNN	msfPCR
	SF49-	GCACTGCTATGCTTACTGG	msfPCR
	FtfAlbBAD_F	GCCTCGAGTAACGTGGTTTCCAGCACAC	PCR + cloning
	FtfAlbBAD_R	GCAAGCTTGGAACGCTTGTCAAAGGCC	PCR + cloning
G. oxydans 621 H			
	FtfOxBAD_F	GCCCATGGCTAACGCTGTTTCCAGCACG	PCR + cloning
	FtfOxBAD_R	GCTCTAGAAGGGAACGCTTGTCCCAGG	PCR + cloning
TMW 2.767, DSM 9533T, TMW 2.1191			
	616V	AGAGTTTGATYMTGGCTCAG	16S PCR
	630R	CAKAAAGGAGGTGATCC	16S PCR

manufacturers' instructions. The constructs were transformed into competent *E. coli* Top10, cells that were generated by the rubidium chloride method, via heat shock transformation. Cells harboring plasmids were selectively cultivated on LB Amp100 media and subsequently transferred to LB Amp100 media supplemented with 6% (w/v) sucrose and 0.2% (w/v) arabinose. In this way, positive *E. coli* clones expressing the corresponding levansucrase as an active enzyme could be selected due to their ability to produce mucous polysaccharide in contrast to the wildtype.

The correct *ftf*-gene insertion and sequence of the insert were further proven by control PCR using plasmid primers flanking the multiple cloning site and by subsequent sequencing.

Computational analysis

Gene sequences were edited and analyzed in terms of ORFs, dyad symmetries and energy calculations for potential hairpin structures using Clone Manager 5 (Sci Ed Central, Cary, USA). Alignments were created using clustalw (EMBL-EBI, Cambridge, UK) and graphically optimized with ESPript.[25] Phylogenetic trees were calculated with BioNumerics 6.50 (Applied Maths, Belgium). Protein structures were predicted using the I-TASSER server.[26] Newly identified 16S rRNA and levansucrase gene sequences were deposited at the European Nucleotide Archive (ENA). The corresponding accession numbers are shown in Figures 2 and 3, respectively.

Results

Selection of fructan producers

Three effective fructan producers were detected among 14 *Gluconobacter* strains of our actual strain collection according to their capability to form mucous colony morphologies on sucrose containing agar plates. Thirteen of these *Gluconobacter* strains have already been shown to produce fructans in different amounts.[17] Figure 4 illustrates the capability of the selected *Gluconobacter* strains to produce mucous polysaccharides. In comparison with *G. oxydans* 621H (weak producer), the *Gluconobacter* strains selected for this study show distinctly different colony morphologies due to the formation of higher amounts of levan when cultivated on sucrose-containing media.

Identification of novel levansucrase encoding genes

The approaches used to identify three novel levansucrase encoding genes are illustrated in Figure 1. In a first step, degenerate primers Av and Ar were constructed from conserved motifs ADAMK and GTLAPT located in the N- and C-terminal domain of known levansucrases of *G. oxydans*, *Gluconacetobacter xylinus* and *Zymomonas mobilis*, respectively. In this way, an expected 1000 bp core fragment of *G. sp.* TMW 2.767 levansucrase was obtained. Finally, two miPCR approaches yielded a 1955 bp chromosomal fragment, which included a complete ORF (1335 bp), putatively coding for a levansucrase in *G. sp.* TMW 2.767 (Figure 1A). Using primers Av and Ar, no positive PCR products could be amplified for *G. cerinus* DSM 9533T. Therefore, an alignment including the levansucrase gene sequences of *Ga. xylinus* (AB034152), *G. oxydans* 621H (NC_006677) and *G. sp.* TMW 2.767 was calculated to identify conserved regions in known levansucrases of AAB. In this way, Z1F and Z1R were created, which amplified the expected core fragment of approximate 750 bp in size. Finally, SiPCR was performed to detect a complete ORF coding for a levansucrase in *G. cerinus* DSM 9533T (Figure 1B). Furthermore, dyad symmetries were detected downstream from the levansucrase genes of *G. sp.* TMW 2.767 and *G. cerinus* DSM 9533T (54 nt and 51 nt, respectively), possibly being involved in hairpin formation (17.9 kcal mol^{-1} and 18.6 kcal mol^{-1}, respectively) and thus possibly acting as rho-independent transcription terminators.
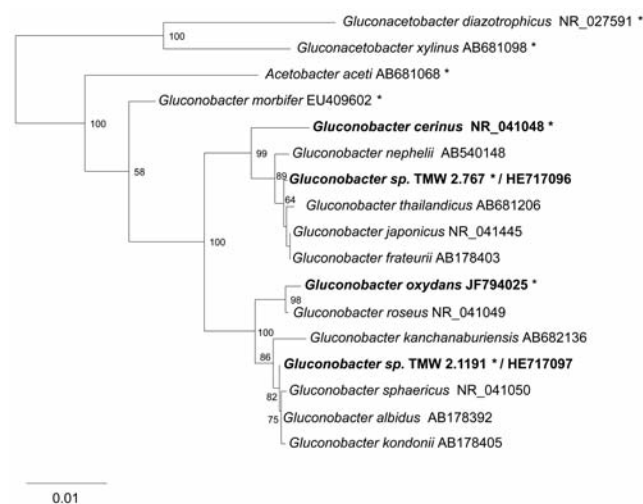

Figure 2. Phylogenetic tree calculated on the basis of 16S rRNA gene sequences using the neighbor-joining method. The dendrogram includes all type strains of the genus *Gluconobacter* and AAB strains that possess levansucrase-encoding genes (indicated by asterisks). The scale-bar indicates number of changes per nucleotide. Strains investigated in this study are highlighted in bold. Numbers following species names represent the accession numbers of the corresponding 16S rRNA genes as deposited in GenBank. Bootstrap values indicated at nodes are derived from 1000 replications.
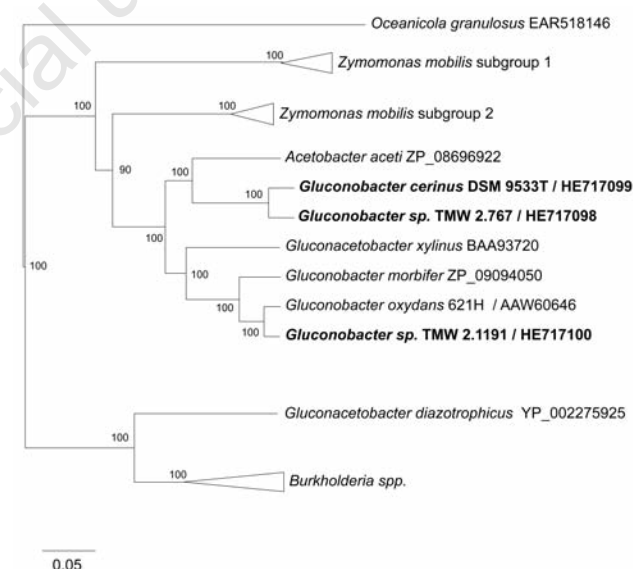

Figure 3. Phylogenetic tree including all currently identified levansucrases from AAB species and their next related enzymes according to Velazquez-Hernandez *et al.*[3] Similarities were calculated using the neighbor-joining method. Numbers following species names represent the accession numbers of the corresponding proteins as deposited in GenBank. Newly identified Ftfs are highlighted in bold. The scale bar represents number of changes per amino acid. The levansucrase protein of *Oceanicola granulosus* (EAR51814) was used as an outgroup. Bootstrap values indicated at nodes are derived from 1000 replications. Compressed branches are composed of the following accession numbers: *Zymomonas mobilis* subgroup 1 (YP_004662017, AAC36942, YP_003225996), *Zymomonas mobilis* subgroup 2 (YP_004662018, BAA04475, AAA27695, AAA27702), *Burkholderia spp.* (ZP_02904976, ZP_02379902, ZP_02885256, ACC75109, ZP_02365208, ZP_0250513, ZP_02385398, ZP_02371500, YP_001117258, ZP_02893677, ZP_02906975, ZP_04941833).

These potential hairpins are composed of a twelve base pair stem with a loop of 5 unpaired bases followed by 5 uracil moieties (*G. sp.* TMW 2.767) and a thirteen base pair stem with a loop of 8 unpaired bases followed by 7 uracil moieties (*G. cerinus* DSM 9533T). To enhance the possibility of amplifying quickly a core-sequence of the unknown *ftf*-gene of *G. sp.* TMW 2.1191, primers DgAlbF and DgAlbR were deduced from an alignment of known levansucrase gene sequences of the *Gluconobacter* strains *G. oxydans* 621H, *G. sp.* TMW 2.767 and *G. cerinus* DSM 9533T. Using this primer set, just one PCR approach was sufficient to obtain the expected core amplicon (~ 1100 bp) of a putative levansucrase in *G. sp.* TMW 2.1191. Since suitable restriction sites and endonucleases were rarely available, mspPCR and msfPCR were performed to finally detect the missing terminal regions of this gene (Figure 1C). Due to the lack of downstream sequence information, no potential transcription terminators could be predicted as described for *G. sp.* TMW 2.767 and *G. cerinus* DSM 9533T. Additional sequence information derived from the flanking regions of the predicted levansucrase genes revealed no homologies to proteins of known function using the Blast-Function of GenBank.

Phylogenetic comparison of EPS-producing *Gluconobacter* strains and their identified levansucrases

For species identification, the isolates were subjected to sequencing of their 16S rRNA genes. The phylogenetic position of these strains based on 16S rRNA gene analysis is presented in Figure 2; this also corroborates the finding of Yamada and Yukphan,[27] that *Gluconobacter* species can phylogenetically be divided into the *G. cerinus* and *G. oxydans* group. Recently, the insect-pathogenic species *G. morbifer* was isolated and described as new species[28] taking a different position to all other *Gluconobacter* species (Figure 2). *G. sp.* TMW 2.767 shares 99% sequence similarity (based on 1355 bp) with the type strains of *G. frateurii* and *G. japonicus*, which again possess identical 16S rRNA genes. *G. sp.* TMW 2.1191 shares 100% sequence similarity (based on 1358 bp) to the type strain of *G. albidus*, which forms white colonies as included in its species description.[29] However, *G. sp.* TMW 2.1191 forms pink-colored colonies (Figure 4A) like the type strain of *G. roseus*.[30]

A similar phylogenetic grouping to the one based on 16S rRNA gene

analysis is obtained when comparing the newly identified Ftf-proteins shown in Figure 3. Levansucrases of *G. sp.* TMW 2.767 and *G. cerinus* DSM 9533T (91% positives in 445 amino acids) as well as levansucrases of *G. sp.* TMW 2.1191 and *G. oxydans* 621H (94% positives in 436 amino acids) (Figure 5) can phylogenetically be grouped in different subclusters (Figure 3). Therefore, the levansucrases investigated in this study show the highest similarity to those of their closest related species. Together with levansucrases of *G. morbifer*, *G. xylinus* and *Acetobacter aceti* and those from ethanologenic *Zymomonas (Z.) mobilis species* (α-proteobacteria, *Sphingomonadeceae*), they form an independent cluster, as also described by Velazquez-Hernandez *et al.*[3] Species similarity is not inevitably associated with levansucrase similarity, as shown for the levansucrase of *Ga. diazotrophicus,* which forms an independent cluster to those originating from different *Burkholderia* species (β-proteobacteria) (Figure 3). In contrast, its next related species *Ga. xylinus* (based on 16S rRNA gene analysis, Figure 2) possesses a levansucrase, which is more similar to those of *Gluconobacter* species (Figure 3). In conclusion, the newly identified proteins can be classified as Ftfs belonging to the GH 68 family based on their phylogenetic grouping and sequence similarity.

Structural features and heterologous expression of the identified levansucrases

The levansucrases of the *Gluconobacter* species investigated in this study are structurally composed of several domains, which share identical or similar sequences. To demonstrate similarities *versus* differences, their protein sequences are compared in Figure 5 and shown as red and yellow blocks, respectively. Moreover, no predictable signal peptides could be detected, whereas the conserved motif WTIADA(M/L) is located in the N-terminal domain (Figure 5). Both features have been described as typical features of levansucrases from gram-negative bacteria.[3] No stabilizing disulphide-bridges as described for levansucrase (LsdA) of *Ga. diazotrophicus*[13] could be observed as predicted by I-TASSER.[26] The respective predicted molecular weights are 47.8 kDa (*G. sp.* TMW 2.1191), 47.9 kDa (*G. oxydans* 621H), 49.2 (*G. cerinus* DSM 9533T) and 49.5 kDa (*G. sp.* TMW 2.767). These differences in size are mainly due to missing amino acids in the N-terminal domain of Ftfs from *G. sp.* TMW 2.1191 and *G. oxydans* 621H. In this domain, the highly similar lev-
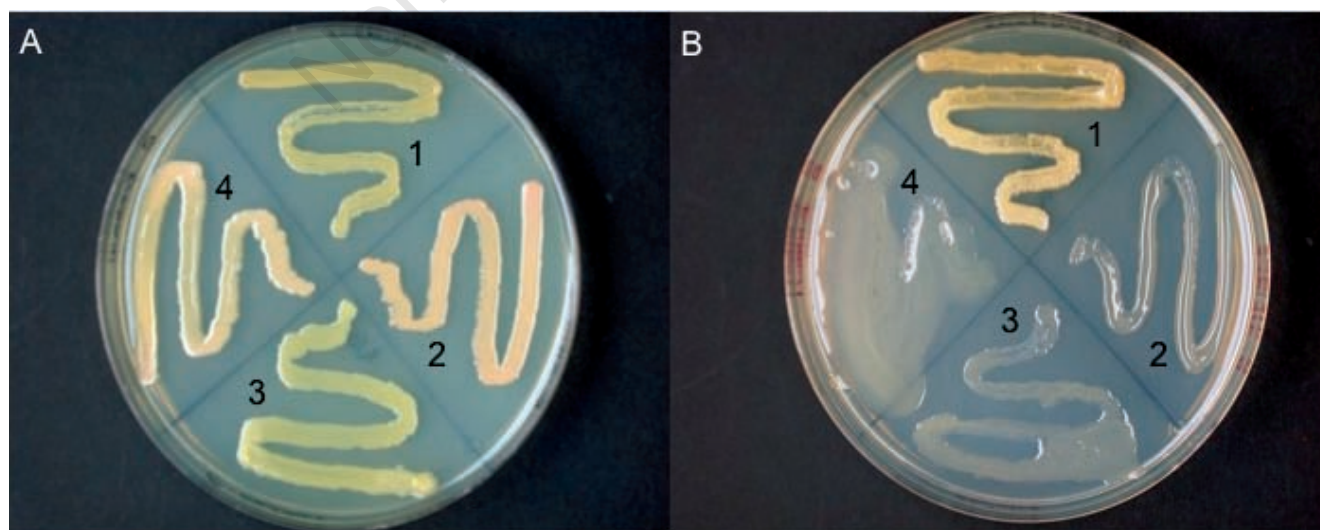

Figure 4. Morphologies of (1) *Gluconobacter oxydans* 621H, (2) *Gluconobacter sp.* TMW 2.1191, (3) *Gluconobacter cerinus* DSM 9533T and (4) *Gluconobacter sp.* TMW 2.767 on modified gluconate media without (A) or with (B) supplemented sucrose (8% w/v).

ansucrases of *G. sp.* TMW 2.767 and *G. cerinus* DSM 9533T (91%) exhibit their strongest sequence differences (only 5 positives in amino acids 6-24, Figure 5). The 11 motifs shown in Figure 5 at some points exhibit slight differences in amino acids compared to the consensus sequences described by Velazquez-Hernandez *et al.*[3] This may indicate that further amino acids could be involved in the catalysis of levan formation. Key amino acids Asp[52/62], Asp[205/215] and Glu[293/303] involved in substrate binding (Figure 5) are also highly conserved in these proteins. Figure 6 demonstrates the capability of recombinant *Escherichia coli* Top10 cells, which harbor the newly identified levansucrase genes, to form polysaccharide from sucrose. Heterologous expression of the corresponding *ftf*-genes in *E. coli* can lead to recombinant strains being able to produce mucous substances from sucrose. However, insertion of the correspon-

ding gene into the used expression vector and subsequent transformation into *E. coli* cells not necessarily yielded recombinant *E. coli* cells capable of producing mucous polysaccharides. Only about 10% of transformants, which harbored expression vectors containing *ftf*-genes as proven by control PCR, exhibited this macroscopic observable feature. Replica plating of all EPS producing *E. coli* mutants on LB media without sucrose yielded their typically wild-type colony morphologies, respectively. The nucleotide sequences of these genes contain a few minor codons which are rarely used by *E. coli*: *G. sp.* TMW 2.767 (1×CGA, 2×CCC); *G. cerinus* DSM 9533T (3×CGA, 3×CCC); *G. sp.* TMW 2.1191 (3×CCC); *G. oxydans* 621H (1×AGG, 3×CCC).

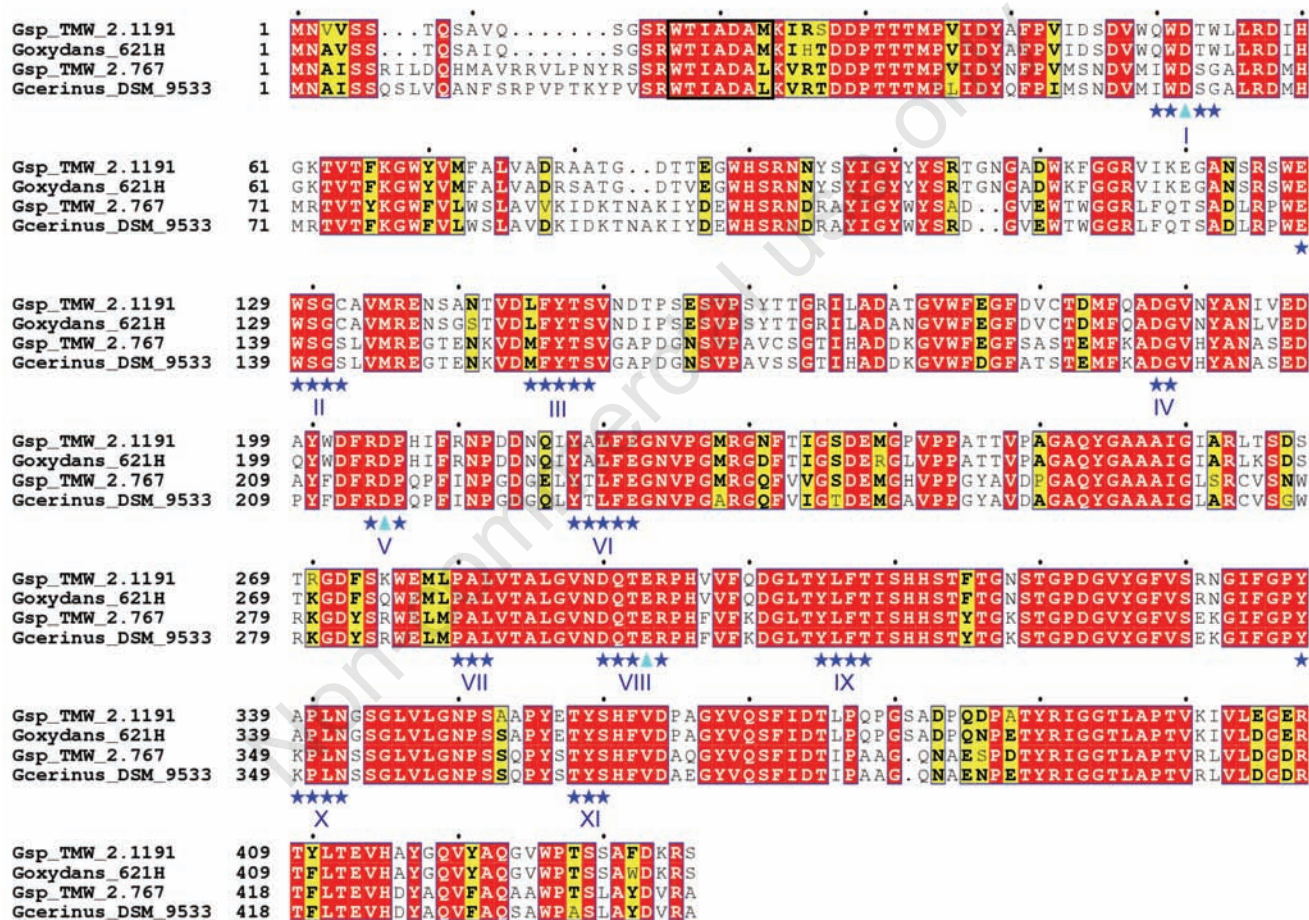These genes can, therefore, be heterologously expressed as active proteins by *E. coli* cells.

Figure 5. Alignment of newly identified levansucrases of *Gluconobacter sp.* TMW 2.1191, *Gluconobacter sp.* TMW 2.767 and *Gluconobacter cerinus* DSM 9533T including the levansucrase sequence obtained from the genome sequenced strain *Gluconobacter oxydans* 621H (accession number AAW60646). Red colored groups represent highly conserved, identical sequence patterns; yellow colored groups indicate well conserved residues within a group, which again are symbolized in bold. Blue asterisks indicate conserved motifs in the catalytic domains of fructosyltransferases belonging to the GH 68 family according to Velazquez-Hernandez *et al.*[3] Meng and Fütterer,[39] Martinez-Fleites *et al.*[13] and van Hijum *et al.*,[40] which are (if identified) functionally designated as (I) catalytic nucleophile, (II) sucrose box 1, (III) sucrose box 2; involved in acceptor recognition, (IV) involved in acceptor recognition, (V) catalytic center; transition state stabilizing; involved in sucrose hydrolysis and fructan polymerization, (VI) involved in sucrose hydrolysis+transfructosylation, (VII) conserved among GH 68 family proteins, (VIII) catalytic center+acid-base catalyst, (IX) conserved in microbial fructosyltransferases, (X) conserved in the GH 68 family and (XI) orientating the catalytic nucleophile for transfructosylation. Light blue triangles symbolize acidic residues Asp[52/62], Asp[205/215] and Glu[293/303], which are supposed to act as catalytic nucleophile, transition state stabilizer and general acid-base catalyst, respectively, and are strictly conserved among members of the glycoside hydrolase families 32, 43, 62 and 68.[1] The bolded box indicates a motif, which is conserved in the N-terminal domain of Ftfs from gram-negative bacteria.[3]

Discussion

Recently discovered *Gluconobacter* strains have increased the previously limited number of known species to a total of twelve type strains. Currently, phylogenetic grouping on the basis of 16S rRNA gene analysis can be used to discriminate species belonging to the *G. oxydans* and *G. cerinus* group.[27] However, species within these groups are hard to distinguish from each other by using the 16S rRNA gene as sole phylogenetic marker as corroborated in this work. *G. sp.* TMW 2.767 shares 99% sequence similarity of the 16S rRNA gene sequences with the type strains of *G. frateurii* (data not shown) and *G. japonicus*.[31] However, the currently deposited sequences of these type strains are identical. As our isolate differs from the (shared) 16S rRNA gene sequence of these closest relatives, *G. sp.* TMW 2.767 may constitute a new species. Another example of the difficult phylogenetic grouping of *Gluconobacter* strains is seen in *G. sp.* TMW 2.1191. This strain shares an identical 16S rRNA gene with the type strain of *G. albidus* but cannot be clearly assigned to this species because of its phenotypically characteristic pink colonies. The characteristic to form *white* colonies is included in the species description of *G. albidus* and constitutes its designation.[29] Therefore, this strain may also constitute a new species.

A PCR based approach was chosen to identify unknown levansucrase gene sequences of the investigated levan-overproducing *Gluconobacter* strains. This strategy was proven to be useful, since at least some levansucrase sequences of related organisms were initially available in public databases. Newly identified levansucrase sequences could subsequently be used to create further degenerate primers, thereby targeting more specifically against conserved levansucrase gene regions of AAB species. Moreover, primers Z1F/Z1R and DgAlbF/DgAlbR were deduced from alignments, which included nucleotide instead of amino acid sequences because of possible similar codon usages among AAB strains, to possibly enhance specific

Figure 6. Recombinant *Escherichia coli* Top10 cells cultivated on LB media supplemented with 6 % (w/v) sucrose and 0.2 % (w/v) arabinose. (1) *E. coli* + empty pBAD vector, (2) *E. coli*+Ftf of *G. sp.* TMW 2.767, (3) *E. coli* + Ftf of *G. cerinus* DSM 9533T, (4) *E. coli*+Ftf of *G. sp.* TMW 2.1191, (5) *E. coli*+Ftf of *G. oxydans* 621H.

binding during degenerate PCR reactions. Whereas the first core-sequence detection (*G. sp.* TMW 2.767) required extensive optimization of PCR conditions due to relatively unspecific binding of degenerate primers Av and Ar, a partial fragment of the last identified levansucrase gene (*G. sp.* TMW 2.1191) could be amplified performing just one PCR approach. Therefore, these newly identified nucleotide sequences might help to quickly obtain parts of potential similar levansucrase genes of further levan-producing *Gluconobacter* strains.

Formation of fructans from sucrose appears to be widespread among the genus *Gluconobacter* as we already demonstrated when screening a total of 13 different *Gluconobacter* strains for EPS synthesis to find strong producers.[17] *G. sp.* TMW 2.767, *G. cerinus* DSM 9533T and *G. sp.* TMW 2.1191 were initially selected for their ability to produce high amounts of levan. The *in silico* translated levansucrase sequences of these three strains can be phylogentically grouped in a separate cluster which includes further levansucrases from other AAB species and those from different strains of *Z. mobilis*. Interestingly, it appears that strains from the genera *Gluconobacter*, *Acetobacter* and *Zymomonas*, which naturally occur in sugar and alcohol-rich environments,[32,33] possess related levansucrases. On the other hand, the N$_2$-fixing endophyte *Ga. diazotrophicus*, which is symbiotic with different plants,[34] expresses a levansucrase with greater similarity to those of *Burkholderia* strains, many species of which are found in the rhizosphere of plants acting as N$_2$-fixers.[35] This observation may indicate horizontal gene transfer between species living in similar ecological niches or even structural adaptations of the enzymes to their surrounding environment due to being extracellularly active.

The levansucrase of *Ga. diazotrophicus* has recently been shown to be involved in biofilm formation and to prevent desiccation.[5] Furthermore, it was speculated, that high amounts of biofilm-forming levan are needed to effectively prevent desiccation, which again results from osmotic stress. *Gluconobacter* species, which predominantly live in sugar-/ethanol rich environments should be confronted with high osmotic stress and, therefore, be able to protect themselves against this. Therefore, an extensive formation of EPSs could help to prevent desiccation. Moreover, it has been demonstrated that osmotically active substances like NaCl strongly influence the microstructure of the related extracellular levansucrase of *Z. mobilis*.[36] At high NaCl concentrations, this enzyme forms fibrils, which distinctly promote the formation of high molecular weight levan. As our levansucrases share high sequence homologies to that of *Z. mobilis* and the producing organisms share the environment, osmotic induction of these enzymes, which might result in the formation of high molecular weight fructans, may be a way to survive in these habitats.

The predicted ORFs coding for levansucrases of the strains investigated in this study could be successfully expressed in *E. coli* showing high activity in phenotypic assays. In this way, approximately 10% of *E. coli* transformants harboring a *ftf*-gene adapted macroscopic observable levan production, which could be explained by spontaneous secretions of the proteins to the extracellular space. As expected, intracellular accumulation of high-molecular weight fructan would yield bursting *E. coli* cells, since levansucrases such as sacB from *Bacillus subtilis* can be used as selection markers due to this *toxic* feature.[37] Besides this, some of these recombinant *E. coli* transformants could possibly cope with the given codon usage of the cloned levansucrase-genes more effectively, at which only a few minor codons rarely used by *E. coli* are present in these genes. Whereas these differences among the cloned *E. coli* mutants are not easily explained, this simple phenotypic screening system allows an easy selection of *E. coli* transformants capable of expressing active levansucrases. The correctness of these identified ORFs coding for levansucrases is further confirmed by the detection of transcription terminators downstream of the respective codons (*G. sp.* TMW 2.767 and *G.*

cerinus DSM 9533T) and additionally by the conserved motif WTIA-DA(M/L), which is located in the N-terminal domain of Ftfs from gram-negative bacteria. LsdA from *Ga. diazotrophicus* exhibits the unique feature among gram-negative bacteria to possess a signal peptide of 30 aa, which is cleaved off during transport to the outside of the cell via a type II secretion pathway.[38] Though, the conserved motif WTRADAL of LsdA remains in the active protein indicating an essential function for levansucrases of gram-negative bacteria. The possibility for a secretion-dependent cleavage of parts of the N-terminal domains of the identified levansucrases is, therefore, rather low (Figure 5), which again indicates that the complete translated ORFs of the newly identified levansucrases and of *G. oxydans* 621H represent the fully active enzymes, respectively. Additionally, 11 conserved motifs present in all Ftfs of microbial origin could be detected, which are thought to be involved in enzymatic catalysis as shown by Meng and Fütterer,[39] Martinez-Fleites *et al.*[13] and van Hijum *et al.*[40]

In conclusion, we report the identification of three novel GH 68 levansucrases from *Gluconobacter* strains, which are able to synthesize high amounts of levan from sucrose. The identified proteins offer new application possibilities and their expression and functionality may be further optimized to produce functional fructopolymers by food-grade *Gluconobacter* strains. The newly identified genes can be heterologously expressed in *E. coli* as active proteins.

References

1. Pons T, Naumoff DG, Martinez-Fleites C, Hernandez L. Three acidic residues are at the active site of a β-propeller architecture in glycoside hydrolase families 32, 43, 62, and 68. Proteins 2004;54:424-32.

2. Ozimek LK, Kralj S, van der Maarel MJEC, Dijkhuizen L. The levansucrase and inulosucrase enzymes of Lactobacillus reuteri 121 catalyse processive and non-processive transglycosylation reactions. Microbiol 2006;152:1187-96.

3. Velazquez-Hernandez ML, Baizabal-Aguirre VM, Bravo-Patin A, et al. Microbial fructosyltransferases and the role of fructans. J Appl Microbiol 2009;106:1763-78.

4. Korakli M, Vogel RF. Structure/function relationship of homopolysaccharide producing glycansucrases and therapeutic potential of their synthesised glycans. Appl Microbiol Biotechnol 2006;71:790-803.

5. Velazquez-Hernandez ML, Baizabal-Aguirre VM, Cruz-Vazquez F, et al. Gluconacetobacter diazotrophicus levansucrase is involved in tolerance to NaCl, sucrose and desiccation, and in biofilm formation. Arch Microbiol 2011;193:137-49.

6. Arrieta J, Hernandez L, Coego A, et al. Molecular characterization of the levansucrase gene from the endophytic sugarcane bacterium Acetobacter diazotrophicus SRT4. Microbiol 1996;142:1077-85.

7. Lisdiyanti P, Kawasaki H, Widyastuti Y, et al. Kozakia baliensis gen. nov., sp. nov., a novel acetic acid bacterium in the a-proteobacteria. Int J Sys Evol Microbiol 2002;52:813-8.

8. Kato N, Mizuno M, Nakai Y, et al. Structural analysis of the water-soluble carbohydrate from Asaia bogorensis by NMR spectroscopy. J Appl Glycosci 2007;54:231-3.

9. Kornmann H, Duboc P, Marison I, von Stockar U. Influence of nutritional factors on the nature, yield, and composition of exopolysaccharides produced by Gluconacetobacter xylinus I-2281. Appl Env Microbiol 2003;69:6091-8.

10. Paul A, Samaddar N, Dutta D, et al. Mercuric ion stabilizes levansucrase secreted by Acetobacter nitrogenifigens strain RG1(T). Protein J 2011;30:262-72.

11. Sakurai K, Arai H, Ishii M, Igarashi Y. Transcriptome response to different carbon sources in Acetobacter aceti. Microbiol 2011;157:899-910.

12. Prust C, Hoffmeister M, Liesegang H, et al. Complete genome sequence of the acetic acid bacterium Gluconobacter oxydans. Nature Biotechnol 2005;23:195-200.

13. Martinez-Fleites C, Ortiz-Lombardia M, Pons T, et al. Crystal structure of levansucrase from the gram-negative bacterium Gluconacetobacter diazotrophicus. Biochem J 2005;390:19-27.

14. Hernandez L, Arrieta JG, Menendez C, et al. Isolation and enzymic properties of levansucrase secreted by Acetobacter diazotrophicus SRT4, a bacterium associated with sugar cane. Biochem J 1995;309:113-8.

15. Maiorano AE, Piccoli RM, da Silva ES, de Andrade Rodrigues MF. Microbial production of fructosyltransferases for synthesis of prebiotics. Biotechnol Lett 2008;30:1867-77.

16. Arrieta Sosa J, Soza GS-H, Hernandez Garcia L, Gonzalez AC. Fructosyltransferase enzyme, method for its production and DNA encoding the enzyme. US Patent 08/362,232, 1997.

17. Jakob F, Steger S, Vogel RF. Influence of novel fructans produced by selected acetic acid bacteria on the volume and texture of wheat breads. Eur Food Res Technol 2012;234:493-9.

18. Macauley S, McNeil B, Harvey LM. The genus Gluconobacter and its applications in biotechnology. Crit Rev Biotechnol 2001;1:1-25.

19. Van Geel-Schutten GH. Use of a polysaccharide as bread improver. Patent no. WO 2006/062410 A1, 2006.

20. Vincent S, Brandt M, Cavadini C, et al. Levan-producing Lactobacillus strain and method of preparing human or pet products using the same. US Patent 6,932,991 B2, 2005.

21. Coussement PAA. Inulin and oligofructose: safe intakes and legal status. J Nutr 1999;129:1412-7.

22. Tieking M, Korakli M, Ehrmann MA, Vogel RF. In situ production of exopolysaccharides during sourdough fermentation by cereal and intestinal isolates of lactic acid bacteria. Appl Environ Microbiol 2003;69:945-52.

23. Gulitz A, Stadie J, Wenning M, et al. The microbial diversity of water kefir. Int J Food Microbiol 2011;151:284-8.

24. Tan G, Gao Y, Shi M, et al. Site-Finding PCR. A simple and efficient method for chromosome walking. Nucleic Acids Res 2005;33;e122.

25. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999;15:305-8.

26. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Prot 2010;5:725-38.

27. Yamada Y, Yukphan P. Genera and species in acetic acid bacteria. Int J Food Microbiol 2008;125:15-24.

28. Roh SW, Nam YD, Chang HW, et al. Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in Drosophila melanogaster. Appl Env Microbiol 2008;74:6171-7.

29. Yukphan P, Takahashi M, Potacharoen W, et al. Gluconobacter albidus (ex Kondo and Ameyama 1958) sp. nov., nom. rev., an acetic acid bacterium in the alpha-proteobacteria. J Gen Appl Microbiol 2004;50:235-42.

30. Malimas T, Yukphan P, Takahashi M, et al. Gluconobacter roseus (ex Asai 1935) sp. nov., nom. rev., a pink-colored acetic acid bacterium in the alphaproteobacteria. J Gen Appl Microbiol 2008; 54:119-25.

31. Malimas T, Yukphan P, Takahashi M, et al. Gluconobacter japonicus sp. nov., an acetic acid bacterium in the alphaproteobacteria. Int J Syst Evol Microbiol 2009;59:466-71.

32. De Ley J, Gillis M, Swings J. Family VI. Acetobacteraceae. In: JG

Holt JG (ed.) Bergey's Manual of Systematic Bacteriology, vol. 1. Williams and Wilkins, Baltimore, 1984, pp 267-78.

33. Swings J, De Ley J. The biology of Zymomonas. Microbiol Mol Biol Rev 1977;41:1.

34. Saravanan VS, Madhaiyan S, Osborne J, et al. Ecological occurrence of Gluconacetobacter diazotrophicus and nitrogen-fixing Acetobacteraceae members: their possible role in plant-growth promotion. Microbial Ecol 2008;55:130-40.

35. Coenye T, Vandamme P. Diversity and significance of Burkholderia species occupying diverse ecological niches. Env Microbiol 2003;5:719-29.

36. Goldman D, Lavid N, Schwartz A, et al. Two active forms of Zymomonas mobilis levansucrase. J Biol Chem 2008;47:32209-17.

37. Schweizer HP. Alielic exchange in Pseudomonas aeruginosa using novel ColE1-type vectors and a family of cassettes containing a portable oriT and the counter-selectable Bacillus subtilis sacB marker. Mol Microbiol 1992;6:1195-204.

38. Arrieta JG, Sotolongo M, Menendez C, et al. A type II protein secretory pathway required for levansucrase secretion by Gluconacetobacter diazotrophicus. J Bacteriol 2004;186:5031-9.

39. Meng G, Fütterer K. Structural framework of fructosyl transfer in Bacillus subtilis levansucrase. Nat Struct Biol 2003;10:935-41.

40. Van Hijum SAFT, Kralj S, Ozimek K, et al. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. Microbiol Mol Biol Rev 2006;70:157-76.