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 inulosucrases, 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, inulosucrases 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 106 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 levan6-10 or to possess genes coding for levansucrases.11,12 One of the most extensively studied levansucrases is expressed by the N2-fixing endophyte Gluconacetobacter (Ga.) diazotrophicus. Besides resolving the crystal structure of this enzyme,13 determining its enzyme kinetics14 and, furthermore, describing biological functions for this species,15 the use of this enzyme for producing levan and FOS, which can be used as calorie-free, non-cariogenic sweeteners and prebiotics in functional foods,16 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 producing *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 DH5\(_\alpha\) (containing either pBADMyehIS (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 \(\mu\)g/mL ampicillin (Amp100). For induction of polysaccharide synthesis by recombinant *E. coli* cells LB medium were supplemented with 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 (Qiobio, 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 *lrf*-genes, degenerate

![Figure 1](image-url)
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 Tm (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 initially 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 (msPCR) 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 Tm 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) 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 levanurase 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>
<thead>
<tr>
<th>Strain(s)</th>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>G. sp. TMW 2.767</em></td>
<td>hν</td>
<td>GCGNGGCGATGAA</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>Ar</td>
<td>NGTNGNGGCGAGGTTG</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>Fratseq_F7</td>
<td>CAGGCTGAGTCCGACAGG</td>
<td>miPCR</td>
</tr>
<tr>
<td></td>
<td>Fratseq_R1</td>
<td>GCCGGCGAGTCCCGGGG</td>
<td>miPCR</td>
</tr>
<tr>
<td></td>
<td>T7</td>
<td>GAGCATCGCTATGAGG</td>
<td>miPCR</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>AATTAACCCTCACTAAAGG</td>
<td>miPCR</td>
</tr>
<tr>
<td></td>
<td>FtfFratBAD_F</td>
<td>GCCCATGCTAATCGTTTCCAGCGA</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td></td>
<td>FtfFratBAD_R</td>
<td>GCTCTAGAAGGCGCCGAGTCATAG</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td><em>G. cerinus DSM 9533T</em></td>
<td>Z1F</td>
<td>ACRCACGAVAGCCG</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>Z1R</td>
<td>CGSTCTGCTCATGACCGC</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>LevCF1</td>
<td>CGCCGGCCGCGCGCGCC</td>
<td>siPCR</td>
</tr>
<tr>
<td></td>
<td>LevCR1</td>
<td>CGCTGCGCGCGCGCGCGCG</td>
<td>siPCR</td>
</tr>
<tr>
<td></td>
<td>LevCF2</td>
<td>AACTGGCATGCGGCGTATTC</td>
<td>siPCR</td>
</tr>
<tr>
<td></td>
<td>LevCR2</td>
<td>CGCGCGCGCGCTGGGCGTGG</td>
<td>siPCR</td>
</tr>
<tr>
<td></td>
<td>FtfCerBAD_F</td>
<td>GGCTCATGCAATCTATTCATCGTCC</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td></td>
<td>FtfCerBAD_R</td>
<td>GCTCTAGAAGGCGCCGAGTCATAG</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td><em>G. sp. TMW 2.1191</em></td>
<td>DgAlbF</td>
<td>GATCCCGGCGAGCACMATG</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>DgAlbR</td>
<td>GGCT CGCCGCGCGCGCAAGA</td>
<td>degenerate PCR</td>
</tr>
<tr>
<td></td>
<td>LevAF1</td>
<td>TGGCTCGTCCCTCGGAGTG</td>
<td>nsPCR</td>
</tr>
<tr>
<td></td>
<td>LevAR1</td>
<td>GATGCTCGTCCCTCGGAGTG</td>
<td>nsPCR</td>
</tr>
<tr>
<td></td>
<td>LevAF2</td>
<td>AATGACCGGACAGAAGATG</td>
<td>sequencing</td>
</tr>
<tr>
<td></td>
<td>LevAR2</td>
<td>CGCTGCGCGCGCTCGACGCG</td>
<td>nsPCR</td>
</tr>
<tr>
<td></td>
<td>SF49+</td>
<td>GCACTGCTATGCTATGCTATG</td>
<td>nsPCR</td>
</tr>
<tr>
<td></td>
<td>SF49-</td>
<td>GCACTGCTATGCTATGCTATG</td>
<td>nsPCR</td>
</tr>
<tr>
<td></td>
<td>FtAlbBAD_F</td>
<td>GGGCGTATGCGGCGGCGGCGG</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td></td>
<td>FtAlbBAD_R</td>
<td>GCAAGCTGGGCACGCGCTGGGCC</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td><em>G. oxydans 621 H</em></td>
<td>FtOxBAD_F</td>
<td>GGGCGTATGCGGCGGCGGCGG</td>
<td>PCR + cloning</td>
</tr>
<tr>
<td></td>
<td>FtOxBAD_R</td>
<td>GCAAGCTGGGCACGCGCTGGGCC</td>
<td>PCR + cloning</td>
</tr>
</tbody>
</table>

Table 1. Primers used for gene identifications and cloning procedures (restriction sites are underlined).
manufacturers’ instructions. The constructs were transformed into competent Escherichia 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 wild-type.

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. Phylogenetic trees were calculated with BioNumerics 6.50 (Applied Maths, Belgium). Protein structures were predicted using the I-TASSER server. 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. 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 of approximate 750 bp in size. Finally, SiPCR was performed to detect a 1955 bp chromosomal fragment, which included a complete ORF coding for a levansucrase in _G. cerinus_ DSM 9533T. Therefore, an alignment including the levansucrase gene sequences of _G. xylinus_ (AB034152), _G. oxydans_ 621H (NC_0016677) 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. 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 granulosa_ (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, AAC56942, YP_003225996), _Zymomonas mobilis_ subgroup 2 (YP_004662018, BAA9475, AAA27702, AAA27702) and _Burkholderia_ spp. (ZP_02904976, ZP_02379902, ZP_02885256, ACC75109, ZP_02365208, ZP_0250513, ZP_02385398, ZP_02371500, ZP_02365208, ZP_02371500, YP_001117258, ZP_02885256, ZP_02904976, 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 fif-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. frateurii and G. japonicus, which again possess identical 16S rRNA genes. G. sp. TMW 2.1191 shares 100% 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 Fif-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. morbiler, G. xylinus and Acetobacter aceti and those from ethanelogenic Zymomonas (Z.) mobilis species (α-proteobacteria, Sphingomonadeceae), they form an independent cluster, as also described by Velazquez-Hernandez et al.1 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 Fifs 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(ML) 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 (LdA) of Ga. diazotrophicus12 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 Fifs from G. sp. TMW 2.1191 and G. oxydans 621H. In this domain, the highly similar lev-
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. This may indicate that further amino acids could be involved in the catalysis of levan formation. Key amino acids Asp^{282}, Asp^{302} and Glu^{293} 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 corresponding 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. Meng and Füttterer, Martinez-Fleites et al. and van Hijum et al., which are (if identified) functionally designated as (I) catalytic nucleophile, (II) sucrase box 1, (III) sucrase 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^{302} and Glu^{293} 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. The bolded box indicates a motif, which is conserved in the N-terminal domain of Ftfs from gram-negative bacteria.
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. 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*. However, the currently deposited sequences of these type strains are identical. As our isolate differs from the (shared) 16S rRNA phenotypically characteristic pink colonies. The characteristic to form *G. albidus* this strain shares an identical 16S rRNA gene with the type strain of *G. oxydans* and *G. cerinus* 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. *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 levans. The in silico translated levansucrase sequences of these three strains can be phylogenetically 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, possess related levansucrases. On the other hand, the *N*-fixing endophyte *Ga. diazotrophicus*, which is symbiotic with different plants, 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*-fixers. 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. Furthermore, it was speculated, that high amounts of biofilm-forming levans 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*. At high NaCl concentrations, this enzyme forms fibrils, which distinctly promote the formation of high molecular weight levans. 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. 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. sp.* TMW 2.1191).
cerinus DSM 9533T) and additionally by the conserved motif WTRADAL of LsdA remains in the active protein indicating an essential function for levansucrases of 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.28 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 Pfams of microbial origin could be detected, which are thought to be involved in enzymatic catalysis as shown by Meng and Fütterer,29 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