Chapter 6

Summary and concluding remarks
Summary

Fructans are homopolysaccharides with a single type of glycopyranosyl unit, i.e. fructose, and are synthesized by microorganisms and plants. Microbial fructans are exopolysaccharides (EPS), as they are produced outside the microbial cells. Microorganisms synthesize two types of fructans, called inulins and levans, using sucrose as substrate. Inulins and levans are mostly linear polymers, with fructose units linked through β(2-1) and β(2-6) bonds, respectively. The fructans synthesized from sucrose contain a glucose molecule at one end and are designated as GF\(_n\), where G refers to the terminal glucose unit, F refers to the fructose units and n is the number of fructosyl units per fructan molecule and is called degree of polymerization (DP). Short chain oligosaccharides, termed fructo-oligosaccharides (FOS), are also produced by several microorganisms (1,2,27) (see also Chapters 2 and 3). Fructan polymers may render protection to microbial cells in their natural environment, against desiccation, phagocytosis and phage attack, antibiotics or toxic compounds (e.g. toxic metal ions, sulfur dioxide, ethanol), predation by protozoans, and certain stresses such as osmotic stress and freezing and drying (4,22,29). They also may play a role in adhesion of cells to solid surfaces and biofilm formation, and facilitate interactions between bacteria and their hosts (24,31).

Microorganisms possess two distinct types of enzymes, levansucrase and inulosucrase, collectively called fructansucrases (FSs) or fructosyltransferases (FTFs), to polymerize the fructose moiety of their substrate sucrose into levan or inulin type fructans, respectively. Levansucrase genes/enzymes have been characterized from a wide variety of microorganisms including Gram-positive and Gram-negative bacteria, while inulosucrase genes/enzymes have only been found in a few species of lactic acid bacteria (LAB), namely in Streptococcus mutans (21), Leuconostoc citreum CW28 (17), Lactobacillus reuteri 121 (27), Lactobacillus reuteri TMW 1.106 (23), Lactobacillus johnsonii NCC 533 (Chapter 2), Lactobacillus gasseri (Chapter 3), and in one species of non-lactic acid Gram-positive bacteria, i.e. Bacillus sp. (30).

The inulosucrase (E.C.2.4.1.9) and levansucrase (E.C.2.4.1.10) enzymes belong to Glycoside Hydrolase family 68 (GH 68) of carbohydrate active enzymes (www.eazy.org)(3). The family GH 68 is grouped together with family GH 32 in clan GH-J on the basis of common protein folds. 3D structures have been reported for six family GH32 proteins (reviewed by (13)), and (7,28). At the start of these studies, 3D structures had been resolved
for two family GH 68 proteins, the SacB levansucrase of the Gram-positive bacterium *B. subtilis* (15,16) and the LsdA levansucrase of the Gram-negative bacterium *Gluconacetobacter diazotrophicus* (14). The overall 3D structures of the two levansucrase proteins display a rare five-fold β-propeller topology, with their active sites located in a deep, negatively charged pocket surrounded by several conserved residues. The fructosyl unit of sucrose is bound at the bottom of this pocket, with three absolutely conserved acidic residues Asp86, Asp247, and Glu342 (SacB numbering) approaching it. The catalytic residues of LsdA are perfectly superimposable with their equivalents in the *B. subtilis* levansucrase, only a greater variation is observed in the surface loops (14).

All known levansucrases and inulosucrases from lactic acid bacteria share high amino acid sequence similarity (>60%) and have similar overall domain organization with four distinct regions: (i) a signal peptide, (ii) an N-terminal variable domain, (iii) a catalytic core, (iv) a C-terminal domain, in some cases containing a cell-wall anchoring domain (12,26,29). The signal sequence of about 36 amino acids targets the protein for secretion, while no specific function has yet been assigned to the N-terminal domain. The catalytic core consists of about 450-500 amino acids and is the most extensively studied region of FTF enzymes. This domain contains the three important catalytic residues corresponding to Asp86 (nucleophile), Asp247 (transition state stabilizer) and Glu342 (general acid) in the *B. subtilis* levansucrase. The catalytic domain also contains several amino acid residues that modulate the specificity and the efficiency of the transfructosylation process in *Bacillus subtilis* levansucrase (5), residues critical for the polymer versus oligosaccharide product ratio of the *Bacillus megaterium* levansucrase (10), or residues important in maintaining the nucleophile position in the active site and with an important role in acceptor substrate specificity (18). The C-terminal domain is variable in length and no clear function has yet been assigned to it. However, this domain contains a cell-wall-binding sub-domain consisting of (i) a spacer region (50 to 125 amino acid residues) rich in Pro/Gly and / or Thr/Ser residues, (ii) an LPXTG/LPKAG sortase recognition motif, (iii) a stretch of hydrophobic amino acids containing PXX repeats (12,26).

Depending on the nature of the acceptor substrates used, bacterial FTFs catalyze two types of reactions i.e. transglycosylation and hydrolysis. FTF activity takes place by a ping-pong type of reaction mechanism (6). In the first step, FTFs cleave the glycosidic bond
between glucose and fructose residues in their substrate sucrose, forming a covalent enzyme-fructosyl complex (http://www.cazypedia.org/index.php/Glycoside_Hydrolase_Family_68) (20). In the transglycosylation reaction, the fructosyl unit is subsequently coupled to the fructosyl moiety of sucrose (FOS formation), or to a growing fructan chain (polymer formation). In the hydrolysis reaction, the fructosyl unit is coupled to a water molecule producing fructose and glucose.

In spite of having high similarity in their amino acid sequences, the same catalytic amino acid residues and a similar domain organization, the levansucrase and inulosucrase enzymes still retain important differences that need to be addressed to understand their exact mode of catalysis. For instance, differences in their structural features that determine their specificity for synthesis of either \(\beta(2-6)\) or \(\beta(2-1)\) linkages and the sizes of fructans produced, production of mainly polymer and kestose (GF\(_2\)) by levansucrases but a wide range of FOS and polymer by inulosucrases, and a relatively low transglycosylation/hydrolysis ratio for levansucrases as compared to inulosucrases. At the start of this study, only a few inulosucrases genes/enzymes had been characterized, and the available FTF amino acid sequence information was not sufficient for a straightforward targeting of residues potentially responsible for functional differences between levansucrase and inulosucrase enzymes. Earlier studies had mainly focused on determining the roles of amino acids that are conserved in both levansucrases and inulosucrases. Consequently, the limited knowledge available for inulosucrase enzymes and lack of 3D structural information hampered efforts to understand the differences between these two closely related enzymes producing different products. Characterization of more inulosucrase sequences, and/or 3D structural information for inulosucrase proteins, appeared to be necessary in order to make progress in such protein structure-function studies.

This PhD thesis focused on (i) characterization of \(ftf/FTF\) genes/enzymes from one \textit{Lb. johnsonii} and three \textit{Lb. gasseri} strains, (ii) elucidation of high resolution 3D structures of the inulosucrase enzyme (InuJ) from \textit{Lb. johnsonii} NCC533 (in collaboration with Tjaard Pijning and Bauke Dijkstra, Biophysical Chemistry, University of Groningen), (iii) alignment of levansucrase and inulosucrase protein sequences from different sources to identify amino acid residues that are highly conserved in inulosucrases but are different in levansucrases. Subsequent replacement of these targeted amino acids in an inulosucrase with
Summary

those in levansucrase at the same location and analysis of the effects on enzyme properties of the mutant proteins.

Characterization of fructansucrase genes/enzymes from lactobacilli (Chapters 2 and 3)

Various probiotic lactobacilli were evaluated for fructan synthesis and the presence of \textit{fff} genes. This allowed identification and characterization of three inulosucrase genes and enzymes from four different strains of lactobacilli. Studies with \textit{Lb. johnsonii} NCC 533 provided the first example of synthesis of inulin and FOS \textit{in situ} by cells grown on sucrose supplemented MRS medium (2)(Chapter 2). Subsequently, three different \textit{Lb. gasseri} strains from the DSMZ culture collection, including the type strain, were studied and considerable heterogeneity was found among their \textit{fff}/FTF genes/enzymes and the fructan products synthesized (1)(Chapter 3). The data also provided important novel insights in the variation in properties occurring among these closely related strains: \textit{Lb. gasseri} DSM 20604 produced inulin and a broad range of FOS \textit{in situ}, while \textit{Lb. gasseri} DSM 20077 produced levan and GF$_2$ type FOS. The \textit{Lb. gasseri} DSM 20604 strain is unique in its distribution of FOS synthesized, ranging from DP2 to DP13, which has not been observed before in other bacteria. On the contrary, the type strain \textit{Lb. gasseri} DSM 20243 did not produce any oligosaccharide or fructan polymer. Further investigations showed that this strain harbors an \textit{fff} gene (designated here as \textit{inuGA}) with a premature stop codon rendering it inactive (see below).

To obtain higher expression levels, truncated versions (without signal peptide, part of the N-terminal variable region and part of the C-terminal domain) of the \textit{fff} genes were amplified, cloned in the pET15b vector and expressed in \textit{E. coli}. In case of \textit{inuGA}, the premature stop codon was replaced with a glutamine residue prior to cloning and expression, yielding an active recombinant enzyme (InuGA-RM). The truncation strategy resulted in very high expression levels in \textit{E. coli}, e.g. providing about 25 mg of His-tag purified InuJ protein from 1 L culture, which also facilitated the crystallization attempts for elucidation of its 3D structure (Chapter 4). Incubation of the purified recombinant enzymes with sucrose and subsequent $^{13}$C NMR analysis of the purified polymer products confirmed that \textit{Lb. johnsonii} harbors an inulosucrase (InuJ); \textit{Lb. gasseri} strain DSM 20243 an inulosucrase
InuGA); strain DSM 20604 an inulosucrase (InuGB) ; and DSM 20077 a levansucrase (LevG). Moreover, the data showed that each of these strains carries a single \textit{fft} gene.

The inulosucrase proteins of \textit{Lb. johnsonii}, \textit{Lb. gasseri} DSM 20243 and DSM 20604 cluster closely in a phylogenetic tree of all bacterial FTF proteins (Chapter 1, Fig. 2), indicating their relatedness. Moreover, these inulosucrase proteins (excluding InuGB), together with LevG, share a common cell wall anchoring motif LPKAG, which is different from that in FTFs of other lactic acid bacteria that contain an LPQTG motif instead. Presence of an incomplete cell wall anchoring domain and lack of the LPXTG/LPKAG motif raised questions about the cellular location of the \textit{Lb. gasseri} 20604 InuGB inulosucrase protein. The impact was clearly indicated by the tendency of the enzyme to be secreted into the growth medium, in contrast to the \textit{Lb. gasseri} 20077 levansucrase, which was exclusively detected associated with the cell surface (Chapter 3, Table 4).

An interesting feature of the \textit{Lb. gasseri} strains is that they were unable to ferment raffinose, whereas their respective recombinant enzymes used raffinose as a substrate (Chapter 3, Fig. 4a). However, all three strains showed copious growth on MRS-raffinose medium supplemented with sucrose even at lower concentration (50 mM), indicating that the growth was not inhibited by raffinose. Possibly, the cells required sucrose for the induction of FTF proteins.

In agreement with the proven role in other FTF enzymes, calcium ions were also found to be essential for optimum enzyme activity and stability of the recombinant FTF enzymes from all the lactobacilli used in these studies. Another common feature of the recombinant inulosucrase enzymes used in this study is that the total and transfructosylation activities of the enzymes could not be saturated with substrate, not even at higher substrate concentrations. Consequently, these reactions did not obey Michaelis-Menten kinetics. The non-Michaelian behavior of InuJ may be attributed to the oligosaccharides that were initially formed in the reaction, and which may act as better acceptors than the growing polymer chain (Chapter 2, Fig. 3). The kinetic data, in accordance with available reports, show that levansucrases are relatively more hydrolytic as compared to inulosucrases.

The amino acid sequences of FTFs from these lactobacilli were aligned with those of other available inulo- and levansucrases from Gram-positive bacteria to identify residues as targets for mutations (see below; Chapter 5).
Elucidation of the 3D structure of inulosucrase (InuJ) from *Lb. johnsonii* NCC 533 (Chapter 4)

The 3D structure of the *Lb. johnsonii* inulosucrase (InuJ) was resolved to 1.75 Å resolution, and that of its inactive mutant D272N to 2.69 Å (Chapter 4). This first inulosucrase structure shows the typical five-fold β-propeller topology of the catalytic domain present in levansucrases (family GH68). The central funnel formed by the five-bladed β-propeller has high structural similarity with the known levansucrase structures (SacB and LsdA). Surprisingly, the native InuJ subsite -1 is in fact identical to that of SacB and LsdA with the three catalytic residues (D272, D425, E524) assuming the same relative positions and orientations (Chapter 4, Figure 5). This homology at the -1 subsite is conceivable since it has the same function in both types of FTFs, namely to cleave the glycosidic bond of sucrose, and to subsequently transfer the fructosyl moiety to an acceptor substrate. The fact that sucrose can adopt a similar orientation in InuJ as observed in SacB, and the similarities in the -1 and +1 subsites, show that binding and cleavage of the substrate in inulosucrase and levansucrase proteins is likely to be very similar.

The orientation of the incoming acceptor substrate will ultimately determine the type of glycosidic bond made in the product. Functional differences therefore can be expected at the +2 and/or higher acceptor substrate binding subsites. Indeed, some differences were observed between the acceptor subsite architecture of InuJ and levansucrases (Chapter 4, Figure 4, Table 2). Compared to SacB, the shift of InuJ R623, the presence of R545 replacing a lysine (of SacB), and the absence of some residues that provide water-mediated contacts at the +2 subsite of SacB, give rise to an acceptor binding pocket with different properties. However, sequence alignments show that R623 and R545 are not exclusively present in inulosucrases but also are conserved in several levansucrases. This indicates that other structural features must be important for determining the product specificity. While the highest similarity was observed in the β-propeller blades of InuJ and SacB, structural differences were mainly found in some other structural elements. For example, helix α7 of InuJ, located in the loop between β-strands 4B and 4C and stacking parallel to helix α3, is absent in SacB and LsdA.

Several levansucrases (8,9,11,25) have been shown to produce 1-kestose from sucrose, which is a trisaccharide with two β(2-1)-linked fructose units. It has been
hypothesized that levansucrase produces 1-kestose as a minor product without further transfer of fructosyl residue onto it, while 6-kestose would be the major product leading to the formation of levan. However, *Bacillus megaterium* levansucrase is the only FTF for which 6-kestose has been characterized as one of the products synthesized. Interestingly, the *B. megaterium* levansucrase data clearly shows the consumption of 1-kestose and apparent accumulation of 6-kestose (10). Another study showed that the levansucrase of *Lb. reuteri* 121 uses 1-kestose to produce longer chain FOS (19). Indeed, the 3D structure of InuJ suggests that the initial step in inulosucrase and levansucrase reactions is similar. Consequently, it can be speculated that the main structural differences determining the product specificity of inulo- and levansucrases is related to the +2, +3 or higher subsites, which determine the orientation of the incoming acceptor substrate for inulin or levan synthesis. Therefore, in order to explain the details of the differences in product specificity of inulo- and levansucrases, 3D structures with bound acceptors of GF$_2$ or GF$_3$ type may give more insight.

**Studies on Inulosucrase Mutants (Chapter 5)**

Only a few inulosucrase proteins were known before the work reported in this thesis was initiated. Characterization of the additional inulosucrases from *Lb. johnsonii* (Chapter 2) and *Lb. gasseri* strains DSM 20243 and 20604 (Chapter 3) facilitated a more detailed comparison of levansucrase and inulosucrase proteins for the determination of the functional importance of amino acid residues that are well conserved in inulosucrases but not in levansucrases. The Inu and Lev proteins of *Lb. reuteri* 121 were selected for comparison as they both are produced by the same organism and are closely related (86% similarity within 768 amino acids) at the amino acid sequence level. As these studies were initiated before the elucidation of 3D structure of InuJ (Chapter 4), mutation targets in Inu were selected solely on the basis of alignments of levansucrase and inulosucrase sequences of lactic acid bacteria. However, the results also were interpreted on the basis of the InuJ 3D structure, which became available at a later time. The point mutations were introduced at positions highly conserved in inulosucrases but different/variable in levansucrases (Chapter 5, Table 1). Also mutants with multiple residue changes (GM1-GM4) were constructed in view of the location
of these amino acids in the active site funnel. None of these residues have been mutated before.

Site-directed mutagenesis of the targeted residues in the *Lb. reuteri* 121 Inu protein, followed by purification and biochemical characterization of the 19 mutant proteins, revealed that several of these residues have clear functional roles in inulosucrase reaction and product specificity. For instance, mutating residue G416, present at the rim of the active site pocket in loop 415-423 (compared with the *Lb. johnsonii* NC 533 InuJ 3D structure, Chapter 4), into glutamate increased the hydrolytic activity 2-fold, without significantly changing transglycosylation activity. The multiple mutant GM4 (T413K, K415R, G416E, A425P, S442N, W486L, P516L), which also contains three residues from the above mentioned loop, synthesized 1-kestose only, though at low efficiency. Mutation A538S, located behind the general acid/base, increased enzyme activity 2-3 fold. The InuJ 3D structure showed that InuJ contains a unique helix α7, absent in the levansucrases SacB and LsdA (Chapter 4). Mutations in residues located in the region of blade 4, including the unique helix α7, had a clear effect on transglycosylation activity and the FOS product profile. Particularly, mutations in N543, which forms a hydrogen bond to the adjacent arginine R544 (present near subsites +1 and/or +2), reduced synthesis of FOS compared to wild type.

**Conclusions and future prospects**

Lactic acid bacteria are present in our food and they contribute to a healthy microflora of human mucous membranes. In the work presented here, we characterized three inulosucrase (and one levansucrase) enzymes from human associated lactobacilli. Only three inulosucrase enzymes had been characterized before the work reported in this thesis was initiated. Our *Lb. johnsonii* and *Lb. gasseri* DSM 20604 studies provided the first time evidence that cells of these lactobacilli convert sucrose in inulin and inulin FOS *in situ*, highlighting their probiotic potential. Some fundamental features of these bacteria were also brought to light in these studies. For example, the presence of a premature stop codon in the *Lb. gasseri* 20243 inulosucrase gene, and the abrupt termination (before the cell wall anchoring motif) of the inulosucrase gene of *Lb. gasseri* DSM 20604, indicated the occurrence of frequent genetic modifications in the ecological niches of these lactobacilli.
The impact of these genetic disorders was clearly reflected in their physiological functions. Strain DSM 20243 was unable to produce any fructan and FOS products while DSM 20604 produced predominantly non-cell wall associated inulosucrase. In vivo restoration of normal $ftf$ gene functions by eliminating premature stop codons can further confirm these physiological traits.

The inulosucrase protein sequences that had become available were used in alignments to identify amino acids that are well conserved in inulosucrases but not in levansucrases. These residues were targets for site-directed mutagenesis studies aiming to determine their functional roles. The results of these mutagenesis studies provided a better fundamental understanding of differences between inulosucrase/levansucrase enzymes, and are important for the construction of inulosucrase mutants with higher transglycosylation specificity, higher catalytic rates, and different FOS/polymer size distribution. Finally, these studies resulted in elucidation of the first 3D structure of an inulosucrase protein (InuJ) (in collaboration with Tjaard Pijning and Bauke Dijkstra, Biophysical Chemistry). Although the -1 subsite of InuJ was virtually identical to that of $B. subtilis$ SacB, structural differences were observed in certain other elements. Future work may succeed in elucidating the structural basis for the observed differences in product specificity ($\beta$(2-1) or $\beta$(2-6) linkages) of inulosucrase and levansucrase enzymes, Especially 3D structures with bound acceptor substrates (e.g. kestose or nystose) may give further insights. Such experiments are in progress, aiming to determine the structures of such complexes.
References


