



# Structure–function relationships of family GH70 glucansucrase and 4,6- $\alpha$ -glucanotransferase enzymes, and their evolutionary relationships with family GH13 enzymes

Xiangfeng Meng<sup>1</sup> · Joana Gangoiti<sup>1</sup> · Yuxiang Bai<sup>1</sup> · Tjaard Pijning<sup>2</sup> · Sander S. Van Leeuwen<sup>1</sup> · Lubbert Dijkhuizen<sup>1</sup>

Received: 21 April 2016 / Accepted: 22 April 2016 / Published online: 7 May 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Lactic acid bacteria (LAB) are known to produce large amounts of  $\alpha$ -glucan exopolysaccharides. Family GH70 glucansucrase (GS) enzymes catalyze the synthesis of these  $\alpha$ -glucans from sucrose. The elucidation of the crystal structures of representative GS enzymes has advanced our understanding of their reaction mechanism, especially structural features determining their linkage specificity. In addition, with the increase of genome sequencing, more and more GS enzymes are identified and characterized. Together, such knowledge may promote the synthesis of  $\alpha$ -glucans with desired structures and properties from sucrose. In the meantime, two new GH70 subfamilies (GTFB- and GTFC-like) have been identified as 4,6- $\alpha$ -glucanotransferases (4,6- $\alpha$ -GTs) that represent novel evolutionary intermediates between the family GH13 and “classical GH70 enzymes”. These enzymes are not active on sucrose; instead, they use ( $\alpha$ 1  $\rightarrow$  4) glucans (i.e. malto-oligosaccharides and starch) as substrates to synthesize novel  $\alpha$ -glucans by introducing linear chains of ( $\alpha$ 1  $\rightarrow$  6) linkages. All these GH70 enzymes are very interesting biocatalysts and hold strong potential for

applications in the food, medicine and cosmetic industries. In this review, we summarize the microbiological distribution and the structure–function relationships of family GH70 enzymes, introduce the two newly identified GH70 subfamilies, and discuss evolutionary relationships between family GH70 and GH13 enzymes.

**Keywords** GH70 · GH13 · Glucansucrase · 4,6- $\alpha$ -Glucanotransferase · Structure–function · Evolution

## Abbreviations

EPS	Extracellular polysaccharides
GRAS	Generally regarded as safe
GS	Glucansucrase
4,6- $\alpha$ -GTs	4,6- $\alpha$ -Glucanotransferases
HMM	High-molecular-mass
HPAEC-PAD	High-performance anion-exchange chromatography equipped with an ED40 pulsed amperometric detection (PAD) system
IMMO	Isomalto-/malto-oligosaccharide
IMMP	Isomalto-/malto-polysaccharide
LAB	Lactic acid bacteria
LMM	Low-molecular-mass

**Electronic supplementary material** The online version of this article (doi:10.1007/s00018-016-2245-7) contains supplementary material, which is available to authorized users.

✉ Lubbert Dijkhuizen  
l.dijkhuizen@rug.nl

<sup>1</sup> Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747, AG, Groningen, The Netherlands

<sup>2</sup> Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, Nijenborgh 7, 9747, AG, Groningen, The Netherlands

## Introduction

Lactic acid bacteria (LAB) are known to produce lactic acid from sugar metabolism and have been used for the production of traditional dairy products since ancient times. Nowadays, LAB are exploited as starter cultures to improve the preservation, nutritional values, taste and mouthfeel of fermented foods [1]. Various strains are also

proven probiotics due to their beneficial effects on human health [2]. LAB have the “generally regarded as safe” (GRAS) status, and are known to produce extracellular polysaccharides (EPS) [3–6]. These EPS are likely involved in the protection of LAB from harsh environmental conditions such as desiccation, osmotic stress and the presence of antimicrobial factors [6–8]. Some EPS facilitate the adhesion of bacterial cells to surfaces and, thus, help LAB to colonize different environments through biofilm formation [6–8]. The EPS produced by *Streptococcus mutans* are important pathogenic factors in dental caries [9, 10]. These EPS have also found valuable applications in food and medicine, and in the cosmetic industries [3–8]. As food supplements, EPS of LAB are explored as texturizers, emulsifiers and viscosifiers and they also hold potential health beneficial effects as dietary fiber and prebiotics [5, 7, 8]. Depending on their composition, EPS are divided into two groups: heteropolysaccharides and homopolysaccharides [11]. Heteropolysaccharides of LAB contain different types of monosaccharides (e.g., glucose, galactose and rhamnose), while homopolysaccharides of LAB consist of only one type of monosaccharide (glucose or fructose). The biosynthesis of heteropolysaccharides is complex and requires the combined action of a large number of proteins including enzymes, transporters and regulators [11, 12]. Generally, Leloir glycosyltransferase enzymes are involved that require expensive nucleotide-activated sugars (e.g., UDP-glucose) [11, 12]. On the contrary, homopolysaccharides are generally synthesized from sucrose using a single glucansucrase (GS) or fructansucrase enzyme [13].

The GS enzymes of LAB belong to glycoside hydrolase family 70 and catalyze the synthesis of  $\alpha$ -glucan homopolysaccharides from sucrose [11, 14–16]. Recently, two new GH70 subfamilies have been established, including the GTFB-like 4,6- $\alpha$ -glucanotransferases (GTFB-like 4,6- $\alpha$ -GTs) and GTFC-like 4,6- $\alpha$ -glucanotransferases (GTFC-like 4,6- $\alpha$ -GTs) [17, 18]. They both synthesize novel  $\alpha$ -glucans, but use ( $\alpha$ 1  $\rightarrow$  4) linked glucans (i.e. malto-oligosaccharides and amylose) as substrates, and are inactive with sucrose.  $\alpha$ -Glucan products synthesized by family GH70 enzymes (GSs, GTFB-like 4,6- $\alpha$ -GTs and GTFC-like 4,6- $\alpha$ -GTs) differ in size, type of linkages, and degree of branching, hence in their physico-chemical properties [13, 16]. This review surveys the recent developments in family GH70 enzyme research with emphasis on their microbiological distribution, the structure–function relationships of GSs and the 2 newly identified GH70 subfamilies, and the evolutionary relationships between family GH70 and GH13 enzymes.

## Microbiological distribution of GH70 GS enzymes

Sucrose, also known as table sugar, is one of the most abundant carbohydrates consumed in our daily life. It is a disaccharide with the formula  $C_{12}H_{22}O_{11}$ , consisting of the monosaccharides glucose (D-glucopyranose) and fructose (D-fructofuranose), linked by an ( $\alpha$ 1 $\leftrightarrow$  $\beta$ 2) glycosidic linkage. Early in 1861, Pasteur found a microorganism-derived substance being responsible for the gelification of sugar-cane syrups and it was named “dextran” [19, 20]. Van Tiehem isolated this microorganism and named it as *Leuconostoc mesenteroides* in 1878 [19, 20]. An enzyme from the cell-free supernatant was found to be responsible for the synthesis of dextran [21, 22]. Now, dextran is defined as a homopolysaccharide which is composed of D-glucose residues with ( $\alpha$ 1  $\rightarrow$  6) linkages in the main chain and different degrees of ( $\alpha$ 1  $\rightarrow$  2) or ( $\alpha$ 1  $\rightarrow$  3) branched glucosyl units. The GS enzyme that synthesizes dextran is named accordingly as dextransucrase (EC 2.4.1.5).

GSs are exclusively found in LAB, such as *Leuconostoc*, *Streptococcus*, *Lactobacillus* and *Weissella* [15]. GSs are extracellular enzymes and depending on the particular bacterial source, they are produced either as cell wall-attached or free enzymes in culture fluids, or both [23–25]. A diversity of GSs has been characterized from various LAB and were found to produce  $\alpha$ -glucans with all the possible glycosidic linkages [( $\alpha$ 1  $\rightarrow$  2), ( $\alpha$ 1  $\rightarrow$  3), ( $\alpha$ 1  $\rightarrow$  4) and ( $\alpha$ 1  $\rightarrow$  6)], each enzyme with its own linkage specificity. With the fast development of genome sequencing, the number of GSs annotated is rapidly increasing. By Nov. 2015, 264 GSs had been annotated, including 57 characterized in the family GH70 of Carbohydrate-Active Enzymes Database (CAZy; see <http://www.cazy.org>). GSs are mainly found within the genera *Leuconostoc* (64 of 264), *Streptococcus* (154 of 264), and *Lactobacillus* (23 of 264). Some GSs are also present in other LAB, i.e. *Weissella* (22 of 264) and *Oenococcus* (1 of 264). Some LAB strains produce more than one GS enzyme. For example, *Streptococcus mutans*, which is the main pathogen responsible for dental caries, produces three distinct GSs (GTFB, GTFC and GTFD) (Table 1) [26, 27]. Six different GSs (DSRA, DSRB, DSRE, DSR-DP, DSR-M and BRS-A) are found in the genome of *Leuconostoc citreum* NRRL B-1299 NRRL B-1299 (originally *Leuconostoc mesenteroides* NRRL B-1299) (Table 1) [28]. Several other strains, i.e. *Leuconostoc mesenteroides* NRRL B-1355 (Table 1) [29] and *Streptococcus sobrinus* were also found to contain multiple GSs [30]. These multiple GS enzymes generally display different product (linkage) specificity (Table 1), but it has remained unclear whether they have different physiological roles. Detailed transcriptomic analysis may reveal their individual in vivo roles. Recently,

**Table 1** Examples of GS enzymes characterized from different LAB and the glycosidic linkage composition of their  $\alpha$ -glucan products

Species	Strains	Enzymes	Glucans	Linkage composition (%)				References	
				( $\alpha 1 \rightarrow 6$ )	( $\alpha 1 \rightarrow 3$ )	( $\alpha 1 \rightarrow 4$ )	( $\alpha 1 \rightarrow 2$ )		
<i>Leuconostoc</i>	<i>mesenteroides</i> NRRL B-512F	DSRS	Dextran	95	5			[138]	
		<i>citreum</i> NRRL B-1299 <sup>a</sup>	DSRA	Dextran	85	15			[38]
			DSRB	Dextran	95	5			[39]
			DSRE	Dextran	81	10	3	5	[37]
			DSR-DP	Dextran	100				[28]
			DSR-M	Dextran	100				[28]
	BRS-A <sup>b</sup>	( $\alpha 1 \rightarrow 2$ )				34	[28]		
	<i>mesenteroides</i> NRRL B-1355	ASR	Alternan	57	43			[40]	
	<i>citreum</i> B/110-1-2	DSRF	Dextran	93	6	1		[42]	
<i>mesenteroides</i> NRRL B-1118	DSRI	Mutan	50	50			[43]		
<i>Streptococcus</i>	<i>mutans</i> GS5	GTFB	Mutan	12	88			[49]	
		GTFC	Mutan	15	85			[50]	
		GTFD	Dextran	70	30			[26]	
	<i>oralis</i>	GTFR	Dextran	86	14			[91]	
	<i>downei</i> Mfe 28	GTF-S	Dextran	90	10			[53]	
		GTF-I	Mutan	12	88			[139]	
<i>Lactobacillus</i>	<i>reuteri</i> 121	GTFA	Reuteran	42		58		[58]	
	<i>reuteri</i> ATCC 55730	GTFO	Reuteran	21		79		[60]	
	<i>reuteri</i> 180	GTF180	Dextran	69	31			[57]	
	<i>reuteri</i> MLI	GTFMLI	Mutan	35	65			[57]	
	<i>sakei</i> Kg15	GTFKg15	Dextran	90	10			[57]	
	<i>fermentum</i> Kg3	GTFKg3	Dextran	92	8			[57]	
	<i>parabuchneri</i> 33	GTF33	Dextran	81	19			[57]	
<i>Weissella</i>	<i>confusa</i> 39-2	DSRC39-2	Dextran	97	3			[62]	
	<i>cibaria</i>	DSRWC	Dextran	100				[63]	

<sup>a</sup> *Leuconostoc mesenteroides* NRRL B-1299 has been reclassified as *Leuconostoc citreum* NRRL B-1299 [28]

<sup>b</sup> With sucrose as sole substrate, BRS-A does not catalyze polymer synthesis. In the presence of linear dextran (33 mM) as acceptor substrate and 146 mM sucrose as donor substrate; 34 % ( $\alpha 1 \rightarrow 2$ ) linkages were found in the product mixture [28]

many  $\alpha$ -glucan-producing LAB strains have been isolated from fermented food or sugar syrups and were found to possess GSs. In a recent study, a total of thirty LAB from French traditional sourdoughs have been screened for the diversity of exopolysaccharides produced from sucrose [23]. These LAB are mainly *Leuconostoc* and *Weissella* strains. They were found to produce glucans with various glycosidic linkages [( $\alpha 1 \rightarrow 2$ ), ( $\alpha 1 \rightarrow 3$ ) and ( $\alpha 1 \rightarrow 6$ )] and the presence of GS-encoding genes was confirmed [23].

### GS enzymes from the genus *Leuconostoc*

*Leuconostoc* is most often found in fermented food. GS enzymes are widespread in *Leuconostoc* and the expression of GS from *Leuconostoc* is generally induced by sucrose [31]. Using chemical mutagenesis, mutants (*L. mesenteroides* B-512FMC, B-742CA, B-742CB, B-1142C,

B-1299C, B-1355CA and B-1355CB), constitutively expressing GSs, were obtained from wild-type strains *L. mesenteroides* NRRL B-512FM, B-742, B-1142, B-1299 and B-1355 [31–33]. These mutant strains produce glucans with identical structures to those of the wild-type organisms. Most of the GSs from *Leuconostoc* strains produce dextran with mainly ( $\alpha 1 \rightarrow 6$ ) linkages and minor ( $\alpha 1 \rightarrow 3$ ) branching linkages (Table 1). The dextran produced by DSRS from *L. mesenteroides* NRRL B-512F has been studied most and is widely used in medicine, food and cosmetic industry [34]. Other  $\alpha$ -glucans with different structures are also produced by *Leuconostoc* bacteria. For instance, DSRE from *L. citreum* NRRL B-1299 is a unique enzyme that synthesizes dextran with ( $\alpha 1 \rightarrow 2$ ) branching linkages (Table 1) [35]. The molecular characterization of this enzyme showed the presence of two catalytic domains (CD1 and CD2), separated by a central glucan-binding domain [35]. Biochemical studies showed that CD1 catalyzed the synthesis of the glucan main chain with

predominantly ( $\alpha 1 \rightarrow 6$ ) linkages, whereas CD2 formed ( $\alpha 1 \rightarrow 2$ ) branch points on the ( $\alpha 1 \rightarrow 6$ ) main chain [35–37]. *L. citreum* NRRL B-1299 also produces two additional GSs (DSRA, DSRB). DSRA synthesizes an  $\alpha$ -glucan with both ( $\alpha 1 \rightarrow 6$ ) and ( $\alpha 1 \rightarrow 3$ ) linkages (Table 1) [38], while the  $\alpha$ -glucan produced by DSRB contains larger amounts of ( $\alpha 1 \rightarrow 6$ ) linkages (Table 1) [39]. The genome sequence analysis of *L. citreum* NRRL B-1299 strain revealed the presence of three more GSs (DSR-DP, BRS-A and DSR-M) [28]. DSR-DP and DSR-M mainly catalyze the synthesis of  $\alpha$ -glucans with ( $\alpha 1 \rightarrow 6$ ) linkages, while BRS-A introduces ( $\alpha 1 \rightarrow 2$ ) branching linkages [28]. Interestingly, *L. mesenteroides* NRRL B-1355 produces a GS (ASR) (Table 1) [29, 40], synthesizing a glucan with alternating ( $\alpha 1 \rightarrow 6$ ) and ( $\alpha 1 \rightarrow 3$ ) linkages [40]. This distinct  $\alpha$ -glucan has been named alternan and its enzyme was designated alternansucrase (ASR, EC 2.4.1.140). Recently, more  $\alpha$ -glucan-producing *Leuconostoc* bacteria and corresponding GSs have been isolated from various sources. An isolate from sugarcane juice was identified as *Leuconostoc citreum* B/110-1-2 encoding a novel dextransucrase (DSRF) [41]. DSRF synthesizes a dextran with 93 % ( $\alpha 1 \rightarrow 6$ ) linkages, 6 % ( $\alpha 1 \rightarrow 3$ ) linkages and 1 % ( $\alpha 1 \rightarrow 4$ ) linkages (Table 1) [42]. DSRI of *L. mesenteroides* NRRL B-1118 was recently characterized and shown to produce an insoluble  $\alpha$ -glucan with approximately 50 % ( $\alpha 1 \rightarrow 6$ ) and 50 % ( $\alpha 1 \rightarrow 3$ ) linkages [43]. In a recent report, four and five putative GS genes were identified in the genome sequences of *L. citreum* LBAE-E16 and LBAE-E16, respectively [44].

### GS enzymes from the genus *Streptococcus*

*Streptococcus* strains, especially *S. mutans*, have been recognized as the major dental caries pathogenic bacteria [9, 45]. Dental caries is generally initiated by biofilm formation involving extracellular polysaccharides produced by microorganisms [9, 10]. One of the main components of this biofilm is  $\alpha$ -glucan (10–20 % dry weight of biofilm) [46, 47]. The biofilms also trap other bacteria and food debris. Once established, the bacteria in the biofilm ferment sugars and produce acids that cause dental caries [45, 48]. The GSs from *Streptococcus* are constitutively expressed [31]. *S. mutans* produces three distinct GSs (GTFB, GTFC and GTFD) (Table 1) [26, 27, 49, 50]. GTFB (formerly known as GTF-I) and GTFC (GTF-SI) synthesize water-insoluble glucans with large amounts of ( $\alpha 1 \rightarrow 3$ ) linkages [designated as mutan and its corresponding enzyme as mutansucrase (EC 2.4.1.125)], while GTFD (GTF-S) catalyzes the synthesis of water-soluble glucan with mainly ( $\alpha 1 \rightarrow 6$ ) linkages [26, 27]. It has been reported that the inactivation of any of the three enzymes resulted in a

decrease of smooth-surface carious lesions in the specific-pathogen-free rat model system [51]. In another study, it was demonstrated that GTFB and GTFC play an important role in cellular adherence to smooth surfaces [52]. However, deletion of the *gtfD* gene only slightly affected *S. mutans* in its cellular adherence. Using GTF-deficient *S. mutans* mutants, it was shown that the presence of all 3 GSs at optimum ratio was important for sucrose-dependent adherence [27]. *S. sobrinus* 6715, which is involved in dental caries as well, also contains multiple GSs, producing soluble and insoluble glucans [30]. *Streptococcus downei* Mfe28 also produces two GSs GTF-S and GTF-I (Table 1), being responsible for the synthesis of soluble glucan and non-soluble glucan, respectively [53]. Considering the importance of GSs in the process of dental caries, it has been suggested that specific inhibitors of GS enzymes may be effective for preventing dental caries [9, 27]. However, these specific inhibitors should not inhibit human GH13 enzymes which share high similarity with GH70 enzymes and are essential for the digestion of our carbohydrate food intake.

### GS enzymes from the genus *Lactobacillus*

*Lactobacillus* strains are widely spread in nature and have been used for food application for ages. Some species, e.g., *Lactobacillus reuteri* strains, are considered as probiotic strains due to their beneficial effects for human health [54]. *Lactobacillus* strains were found to produce  $\alpha$ -glucans by novel GSs. In an early study, a total of 182 *Lactobacillus* strains were screened for EPS production with sucrose-rich medium; 60 of them were found to produce EPS (glucan or fructan), of which 17 produced large amounts (more than 100 mg/L) [55]. The GSs from *Lactobacillus* are expressed constitutively [13, 25, 56]. Later on, the genes encoding these GSs were cloned and the enzymes were biochemically characterized [56–58]. GTFA from *L. reuteri* 121 synthesizes an  $\alpha$ -glucan with 58 % ( $\alpha 1 \rightarrow 4$ ) linkages and 42 % ( $\alpha 1 \rightarrow 6$ ) linkages (Table 1) [56].  $\alpha$ -Glucan containing large amounts of ( $\alpha 1 \rightarrow 4$ ) linkages is referred to as reuteran and its corresponding enzyme as reuteransucrase (EC 2.4.1.-). Structural analysis of reuteran produced by GTFA revealed a large amount of alternating ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages [59]. GTFD of the probiotic strain *L. reuteri* ATCC 55730 represents another reuteran-producing GS and its reuteran has an even larger amount of ( $\alpha 1 \rightarrow 4$ ) linkages (~80 %) compared to GTFA (Table 1) and a smaller amount of ( $\alpha 1 \rightarrow 6$ ) linkages (~20 %) [60]. A variety of other GSs producing dextran (GTF180, GTFKg15, GTFKg3 and GTF33) and mutan (GTFML1) were also identified in the genus *Lactobacillus* (Table 1) [57]. For example, GTF180 from *L. reuteri* 180 produces



an  $\alpha$ -glucan with 69 % ( $\alpha 1 \rightarrow 6$ ) linkages and 31 % ( $\alpha 1 \rightarrow 3$ ) linkages, while GTFML1 synthesizes an  $\alpha$ -glucan with large amounts of ( $\alpha 1 \rightarrow 3$ ) linkages ( $\sim 70$  %) and ( $\alpha 1 \rightarrow 6$ ) linkages ( $\sim 30$  %).

## GS enzymes from other lactic acid bacteria

Apart from *Leuconostoc*, *Streptococcus* and *Lactobacillus*, other LAB (i.e. *Weissella*) also contain genes encoding GS enzymes. *Weissella*, originally included in the genus *Leuconostoc*, also produces EPS from sucrose. Most of the  $\alpha$ -glucans produced by *Weissella* strains are dextrans with large amounts of ( $\alpha 1 \rightarrow 6$ ) linkages (more than 90 %) (Table 1). *Weissella confusa* 39-2 and *Weissella cibaria* LBAE K39 were isolated from wheat sourdoughs and found to produce linear dextrans [61]. The GS (DRSC39-2) from *W. confusa* 39-2 has been characterized [62]. *W. cibaria* isolated from human saliva was found to encode a GS (DSRWC), synthesizing a soluble glucan with large amounts of ( $\alpha 1 \rightarrow 6$ ) linkages [63]. Also a putative GS gene from *Oenococcus oeni* PSU-1 was identified by genome sequencing [64]. This enzyme and its product remain to be characterized.

## Structures of family GH70 GS enzymes

### Domain organization of family GH70 GS enzymes

Based on sequence similarity, GS enzymes are classified as members of glycoside hydrolase family 70 (GH70) in the CAZy database. Prior to the availability of crystal structures of GSs, their catalytic mechanism was mainly explored by comparative studies with GH13 enzymes, which are structurally, evolutionarily and mechanistically related to GH70 family enzymes. GSs are large proteins with an average molecular weight of  $\sim 160$  kDa. Amino acid sequence analysis showed that all GS proteins have the same domain organization, with only a few exceptions. The amino acid sequences of different GSs contain four different regions [57, 65, 66] (Fig. 1): (a) signal peptide (SP), (b) N-terminal variable region (VR), containing different amino acid repeat units [35, 56, 57, 67, 68], (c) conserved catalytic domain (CD), comprising the catalytic  $(\beta/\alpha)_8$  barrel and (d) C-terminal glucan-binding



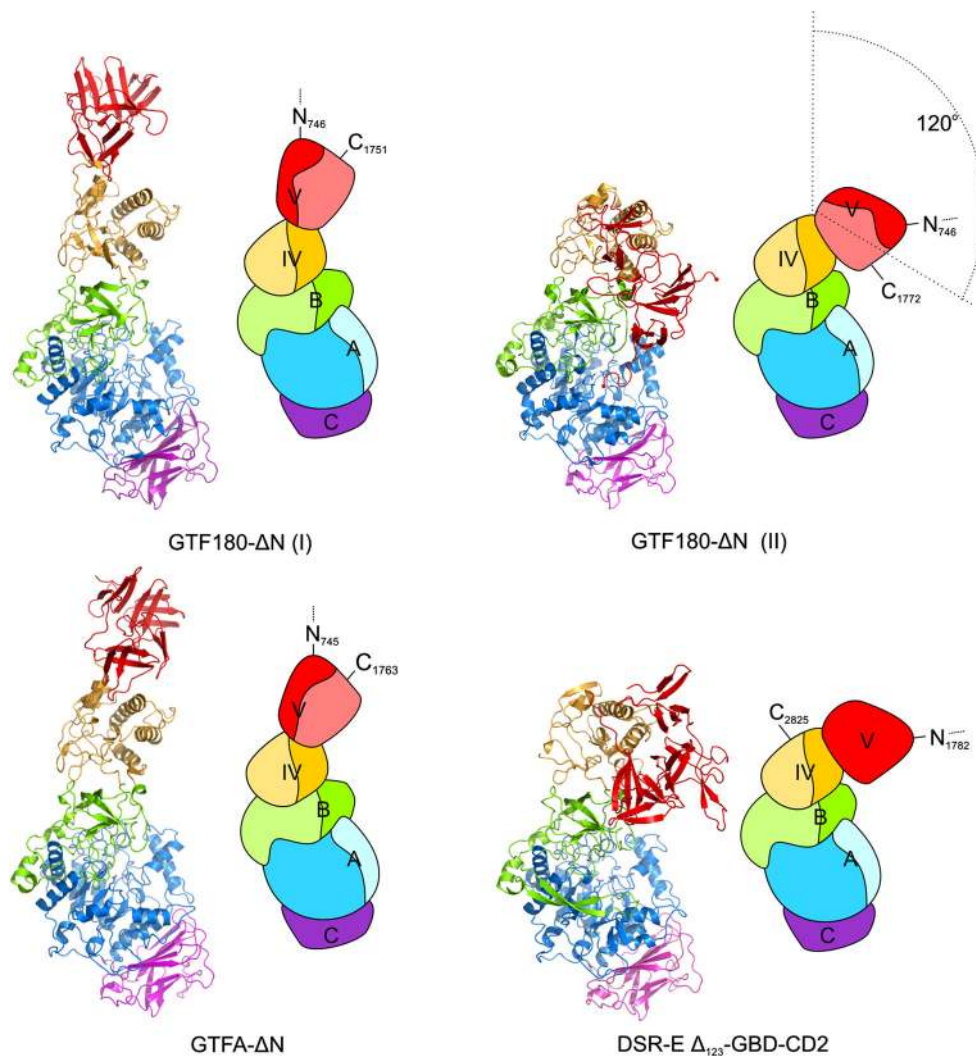
**Fig. 1** General primary structure of GS proteins from lactic acid bacteria. SP signal peptide, VR variable region, CD catalytic domain, GBD glucan binding domain

domain (GBD), also consisting of different amino acid repeats and proposed to be involved in  $\alpha$ -glucan binding [39, 56, 57, 65, 69, 70].

Over the years, large efforts have been made to generate crystals of GS proteins suitable for X-ray diffractions studies, but only recently the first 3D structures of GS proteins have been solved [71]. To date, the crystal structures of four GH70 GSs (GTF180- $\Delta$ N from *L. reuteri* 180, GTFA- $\Delta$ N from *L. reuteri* 121, GTF-SI (amino acid residues 244–1163) from *S. mutans* and  $\Delta$ N<sub>123</sub>-GBD-CD2 of DSR-E from *L. citreum* NRRL B-1299) have been determined (Fig. 2), revealing a common domain organization [36, 71–73]. In all cases, a truncated enzyme (i.e. lacking the N-terminal variable region, but retaining full activity) was used for crystallization. Rather surprisingly, the 3D structures of these GSs revealed a novel domain organization (Fig. 2), different from the previous prediction based on primary sequence alignments (Fig. 1). The polypeptide chains of the truncated GSs follow a U-shape path to form five domains (A, B, C, IV and V) (Fig. 2) [71]. Except for domain C, each of the four domains is formed by two discontinuous polypeptide chains from both the N- and C-termini. The A, B, and C domains form the catalytic core and are also found in GH13 enzymes (e.g.,  $\alpha$ -amylase). In contrast, domains IV and V are unique to GH70 GSs.

Domain A contains the circularly permuted  $(\beta/\alpha)_8$  barrel in contrast to the  $(\beta/\alpha)_8$  barrel in GH13 enzymes (Fig. 3) [74]. The  $(\beta/\alpha)_8$  barrel is characterized by the presence of 8  $\beta$ -strands ( $\beta 1$ – $\beta 8$ ) residing in the core of the enzyme alternating with 8  $\alpha$ -helices ( $\alpha 1$ – $\alpha 8$ ) surrounding the  $\beta$ -strands (Fig. 3). From N-terminus to C-terminus, the circularly permuted  $(\beta/\alpha)_8$  barrel of GH70 GSs starts with the  $\alpha$ -helix that corresponds with  $\alpha 3$  of family GH13 enzymes with the sequence of N terminus- $\alpha 3$ - $\beta 4$ - $\alpha 4$ - $\beta 5$ - $\alpha 5$ - $\beta 6$ - $\alpha 6$ - $\beta 7$ - $\alpha 7$ - $\beta 8$ - $\alpha 8$ - $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\alpha 2$ - $\beta 3$ -C-terminus (Fig. 3a and 3c) [74]. The four conserved amino acid sequence motifs (I to IV) of the GH13 family enzymes are also present in GH70 GSs (Fig. 4) [56, 75]. Due to the circularly permuted structure of GS enzymes, their conserved motif I is located C-terminal of motifs II to IV (Fig. 3). Six of the seven conserved residues from motifs I to IV in family GH13 are also present in family GH70 (Fig. 4) [13]. Only the His134 ( $\alpha$ -amylase of *Bacillus licheniformis* numbering) in family GH13 is replaced by Gln (Gln1509, GTF180 numbering) in family GH70 GSs [76]. Specifically, the three catalytic site residues (the nucleophile Asp1025, acid/base catalyst Glu1063 and transition state stabilizer Asp1136, GTF180 numbering) are located in the loops following  $\beta$ -strands  $\beta 4$ ,  $\beta 5$  and  $\beta 7$ , in the conserved sequence motifs II, III and IV, respectively (Fig. 3) [71]. The mechanism of glycosidic linkage cleavage by GH70 GS enzymes is similar to that of family GH13  $\alpha$ -amylase enzymes (see below).

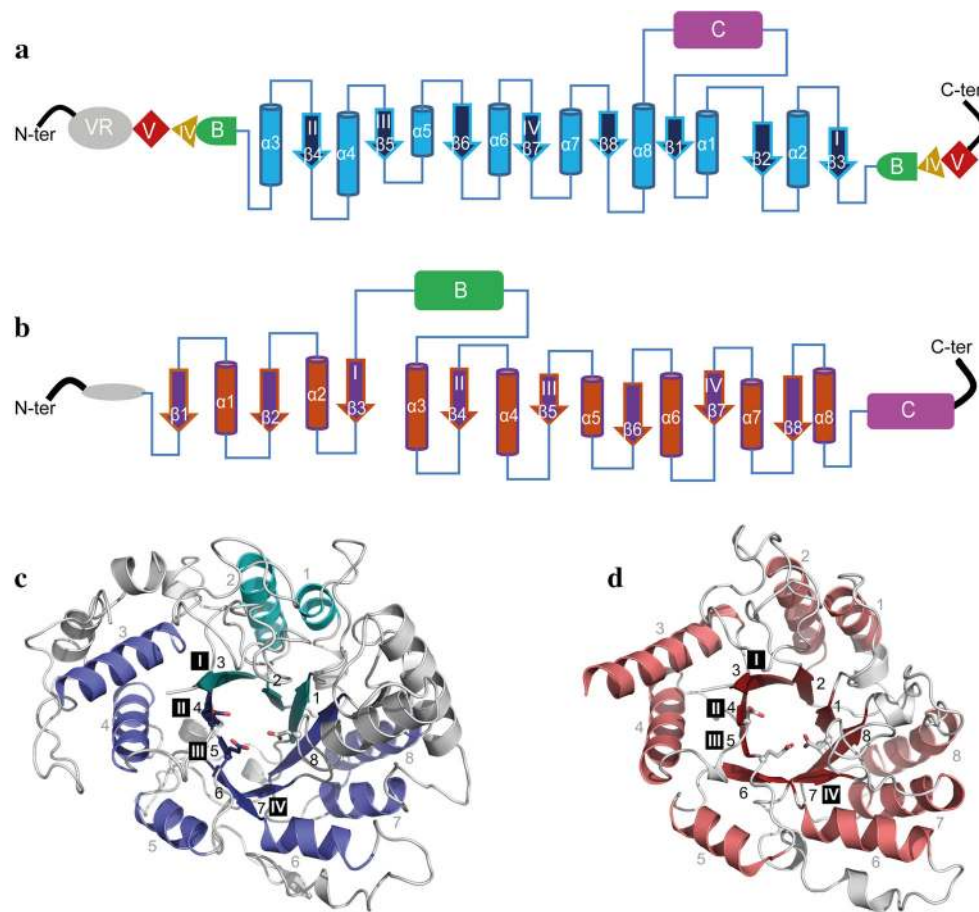
**Fig. 2** Three-dimensional structures and schematic domain organization of GSs from family GH70. Different domains are colored in blue (A), green (B), magenta (C), yellow (IV) and red (V). Crystal structures [GTF180- $\Delta$ N I (PDB: 3KLK, 1.65 Å), GTF180- $\Delta$ N II (PDB: 4AYG, 2.0 Å), GTFA- $\Delta$ N (PDB: 4AMC, 3.60 Å) and DSR-E  $\Delta$ N<sub>123</sub>-GBD-CD2 (PDB: 3TTQ, 1.90 Å)] are shown. This figure has been adapted from [15]



Two discontinuous inserts with a large stretch of amino acids between  $\beta$ -sheet 3 and  $\alpha$ -helix 3 from the N- and C-termini form a separate domain B next to domain A (Figs. 2, 3). The position of domain B in GH70 GS enzymes is similar to that in family GH13 enzymes. The active site of GSs is located in a pocket-shaped cavity and lies at the interface of domain A and domain B [71]. Domain B contains several amino acids (L938, L940, A978 and L981) for shaping the substrate/acceptor binding sites, which may contribute to the enzyme product specificity. [71]. Domain C of GH70 GSs, located at the bottom of the U-shape (Figs. 2, 3), contains an eight-stranded  $\beta$ -sheet with a Greek key motif similar to that of domain C in family GH13 enzymes [71, 77]. It is the only domain in GH70 GS enzymes that is formed by a continuous polypeptide segment. The function of domain C remains unclear, although it is widely distributed within the GH13 and GH70 families.

In addition to domains A, B and C, GH70 GSs have two extra domains (IV and V) attached to the catalytic core.

Domain IV lies between domain B and domain V. The structure of domain IV is novel and it has no similarity to any other known protein structure and only occurs in GH70 enzymes [15, 71]. Domain V is located adjacent to domain IV. It contains several sequence repeats which have been shown to be involved in glucan binding [13, 67, 69, 70, 78]. Structural analysis of domain V revealed the presence of a consensus  $\beta$ -solenoid fold with multiple copies [15, 71]. The precise roles of domains IV and V have remained unknown. It has been proposed that domain IV acts as a hinge that facilitates the growth of glucan chain by bringing the glucan chain bound to domain V toward and away from the catalytic site, but no experimental evidence is available yet [72]. Comparison of crystal structures of different GS enzymes revealed a positional variability of domain V (Fig. 2). For example, compared to the crystal structure of GTF180- $\Delta$ N, domain V of GTFA- $\Delta$ N showed a shift of about 20 Å with respect to the other domains [73]. Surprisingly, domain V of  $\Delta$ N<sub>123</sub>-GBD-CD2 of DSRE is located in a completely different position adjacent to the



**Fig. 3** Topology diagrams models of family GH70 GSs with a circularly permuted  $(\beta/\alpha)_8$  barrel (**a**) and the family GH13  $\alpha$ -amylase  $(\beta/\alpha)_8$  barrel (**b**). Cylinders represent  $\alpha$ -helices and arrows represent  $\beta$ -strands. The equivalent  $\alpha$ -helices and  $\beta$ -strands in GH70 GSs and GH13  $\alpha$ -amylases are numbered the same. The different domains in GH70 and GH13 enzymes are indicated. Domain C of GH70 GSs is inserted between  $\alpha$ -helix 8 and  $\beta$ -strand 1 while that of GH13 family  $\alpha$ -amylase locates C-terminally of the  $(\beta/\alpha)_8$  barrel. Domain B of GH13  $\alpha$ -amylases is inserted between  $\beta$ -strand 3 and  $\alpha$ -helix 3 while that of GH70 GSs is formed by two discontinuous polypeptide segments from both the N- and C-termini. The same is true for domains IV and V of GH70 GSs. A variable region (VR) is present in the N-terminus of GH70 GSs. The four conserved sequence

motifs (I–IV) which are located in  $\beta$ -strands 3, 4, 5, and 7, respectively, and are shared between family GH70 GSs and GH13 enzymes, are indicated within the  $\beta$ -strand. The structure of the catalytic domain in the GH70 GSs representative GTF180- $\Delta$ N (**c**, PDB: 3KLLK) of *L. reuteri* 180 and in the GH13 representative  $\alpha$ -amylase of *Bacillus licheniformis* (**d**, PDB: 1BPL). The  $(\beta/\alpha)_8$  barrel is colored for a better representation.  $\alpha$ -Helices and  $\beta$ -strands are numbered, and the conserved sequence motifs (I–IV) are indicated at the corresponding  $\beta$ -strand. The circularly permuted  $(\beta/\alpha)_8$  barrel of GH70 GSs is formed by two separate polypeptide segments (N-terminal parts colored in deep blue and C-terminal parts colored in cyan), which is caused by the insertion of domain C

catalytic core, adopting a similar fold as domain V of GTF180- $\Delta$ N [36]. A B-factor analysis of domain V in the different crystal structures showed that its average value is higher than the other four domains, which indicates that it is more flexible [79]. Recently, the flexibility of domain V was also demonstrated by the elucidation of a new crystal form of GTF180- $\Delta$ N with a  $120^\circ$  rotation at a hinge located between domains IV and V (Fig. 2), further supported by the observation of positional flexibility of domain V in solution [79]. Truncation of domain V from GTF180- $\Delta$ N did not have significant effects on its linkage specificity

[80]. However, higher amounts of oligosaccharides (size  $\sim 2$  kDa) were produced at the expense of polysaccharide production [80]. This provided direct evidence for the involvement of domain V in polysaccharide synthesis.

Although the crystal structure of a full-length GS is not available yet, small angle X-ray scattering studies have showed that the N-terminal variable region ( $\sim 700$  amino acids) extends further away from domain V [79]. As a result, the overall shape of GTF180 showed an almost symmetric boomerang-like molecular shape with the bend point located between domains IV and V.



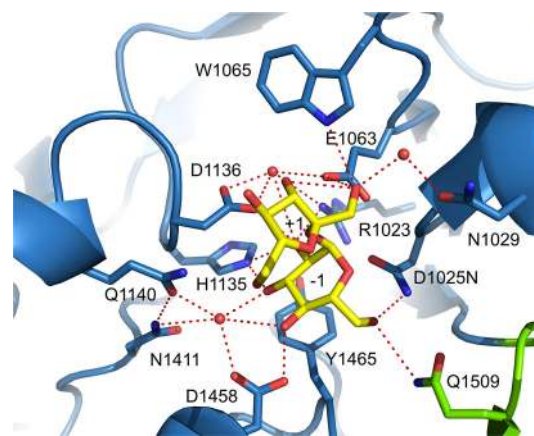
Bacterial strain	NCBI accession numbers	Motif I		Motif II		Motif III		Motif IV	
		1	2	3	4	5	6	7	
<b>a</b>									
<i>Lactobacillus reuteri</i> 180 (GTF180)	AAU08001.1	1503	ADWVPDQ	1021	GIRVDAVDNVD	1058	HINILEDWGW	1131	FVRAHDSNAQDQIRO
<i>Leuconostoc mesenteroides</i> NRRL B-512F (DSRS)	AAD10952.1	1023	ADWVPDQ	547	GIRVDAVDNVD	584	HLSILEDWSDH	657	FVRAHSEVQTVIAQ
<i>Streptococcus oralis</i> ATCC10557 (GTFR)	BAA95201.1	1003	ADWVPDQ	512	GVRVDAVDNVD	549	HLSILEAWSDN	622	FVRAHDSVQTVIAD
<i>Lactobacillus reuteri</i> 121 (GTFA)	AAU08015.1	1508	ADWVPDQ	1020	SVRVDAVDNVD	1056	HINILEDWNSH	1128	FVRAHDNNSQDQIQN
<i>Lactobacillus reuteri</i> ATCC 55730 (GTF0)	AAV86923.1	1508	ADWVPDQ	1020	SVRVDAVDNVD	1056	HINILEDWNSH	1128	FVRAHDNNSQDQIQN
<i>Streptococcus mutans</i> SI (GTFC)	BAA26114.1	954	ADWVPDQ	473	SIRVDAVDNVD	510	HLSILEAWSYN	583	FVRAHDSVQDLIRD
<i>Leuconostoc mesenteroides</i> NRRL B-1355 (ASR)	CAB65910.2	1168	ADWVPDQ	631	GIRVDAVDNVD	668	HLSILEDWNGK	762	FVRAHSDYAQDPIRK
<i>Leuconostoc mesenteroides</i> NRRL B-1299 (DSRECD2)	CAD22883.1	2688	ADVVDNQ	2206	SIRVDAVDNVD	2243	HISLVEAGLDA	2317	IIHAHDKGVQEKVGA
<b>b</b>									
<i>Geobacillus stearothermophilus</i> (BSTA)	AAA22235.2	134	ADVVDNQ	264	GFRVDAVKHIK	293	LFTVGEYYSYD	360	FVDNHDTEPGQALQS
<i>Bacillus licheniformis</i> (BLA)	WP_025807921.1	128	GDVVINH	256	GFRVDAVKHIK	285	MFTVAEYQWQD	352	FVDNHDTPQGSLES
<i>Bacillus</i> sp. 707	AAA22231.1	134	GDVVINH	265	GFRVDAVKHIK	294	MFAVAEYQWQD	362	FVDNHDSPPEEALAS
<i>Halotheothrix orenii</i> (AmyB)	ACL70573.1	251	FDAVLNH	370	GFRVDAVKHID	399	VFFVGEAWVED	466	FVDNHDTRDEGSYT
<b>c</b>									
<i>Lactobacillus reuteri</i> 121 (GTFB)	AAU08014.2	1478	EDIVMNQ	1011	GFRVDAADNID	1048	HLSYNEGYHSG	1120	FVTNHDQR-KNLINR
<i>Lactobacillus reuteri</i> ML1 (ML4)	AAU08003.2	1479	EDIVMNQ	1012	GFRVDAADNID	1049	HLSYNEGYHSG	1121	FVTNHDQR-KNLINR
<i>Lactobacillus salivarius</i> GJ-24	EGM52218.1	1485	VDIVMNQ	1026	GFRVDAADNID	1063	HLSYNEGYRSG	1134	YVTNHDQR-ANLING
<i>Pediococcus pentosaceus</i> IE-3	CCG90643.1	841	EDIVMNQ	380	GFRVDAADNID	417	HLSYNEGYHSG	489	FVTNHDQR-KNLINS
<i>Lactobacillus reuteri</i> DSM 20016 (GTFW)	ABQ83597.1	1215	EDLVMNQ	748	GFRVDAADNID	785	HLSYNEGYHSG	858	FVTNHDQR-KNVIHQ
<i>Lactobacillus fermentum</i> ATCC 14931	EEI21226.1	862	EDLVMNQ	406	GFRVDAADNID	443	HLSYNEGYHSG	512	FVTNHDQR-KNLINQ
<b>d</b>									
<i>Exiguobacterium sibiricum</i> 255-15 (GTFC)	ACB62096.1	138	MDLVPNQ	403	GFRVDAASHYD	433	HLSYIESYKSE	504	FVNNHDQE-KNRVNO
<i>Exiguobacterium undae</i>	WP_028105602.1	138	MDLVPNQ	403	GFRVDAASHYD	433	HLSYIESYKSE	504	FVNNHDQE-KNRVNO
<i>Exiguobacterium antarcticum</i> B7	AFS71545.1	138	MDLVPNQ	403	GFRVDAASHYD	433	HLSYIESYKSE	504	FVNNHDQE-KNRVNO
<i>Exiguobacterium acetylicum</i>	WP_029342707.1	138	MDLVPNQ	403	GFRVDAASHYD	433	YLSYIESYKTE	503	FVNNHDQE-KNRVNO
<i>Bacillus kribbensis</i>	WP_035322188.1	130	EDLVPNQ	397	GFRVDAASHYD	429	HLSYIESYSNV	491	FVNNHDQE-KNRVNN
<i>Bacillus coagulans</i> DSM1	AJH79253.1	128	EDLVPNQ	394	GFRVDAASHYD	426	HLSYIESYQSA	497	FVNNHDQE-KNRVNK

**Fig. 4** Amino acid sequence alignment of conserved sequence motifs I, II, III, and IV in the catalytic domains of GH70 GSs (a), GH13  $\alpha$ -amylases (b), (putative) GTFB-like 4,6- $\alpha$ -GTs (c) and (putative) GTFC-like 4,6- $\alpha$ -GTs (d). The seven strictly conserved amino acid residues in GH13 enzymes (indicated by the numbers 1–7 above the sequences) are also highlighted. Six of them are also conserved in the

GH70 GSs, novel GH70 GTFB-like 4,6- $\alpha$ -GTs and GTFC-like 4,6- $\alpha$ -GTs. Amino acids that constitute the catalytic triad are highlighted in red. Residues forming acceptor binding subsites –1, +1 and +2 in GTF180- $\Delta$ N are shown lightly shaded. Symbols NU = nucleophile, A/B = general acid/base, TS = transition state stabilizer

## Catalytic mechanism of GS enzymes

Similar to family GH13 enzymes, the  $\alpha$ -retaining double displacement reaction mechanism is used by GH70 GSs [13, 65, 77]. This two-step mechanism involves three catalytic residues, the nucleophile and the acid/base catalysts as well as the transition state stabilizing residue (Fig. 4). In the first step, the ( $\alpha$ 1 $\leftrightarrow$  $\beta$ 2) glycosidic linkage of sucrose is cleaved by the attack of the nucleophile with the formation of a  $\beta$ -glucosyl-enzyme covalent intermediate. This glucosyl-enzyme intermediate is stabilized by the transition state stabilizing residue. The acid/base catalyst protonates the fructosyl moiety, resulting in release of fructose. In the subsequent step, the glucosyl moiety is transferred to the non-reducing end of an acceptor with retention of the  $\alpha$ -anomeric configuration. Repeating this reaction cycle results in the synthesis of gluco-oligosaccharides and  $\alpha$ -glucan polysaccharides from sucrose. The crystal structure of GTF180- $\Delta$ N validates that GSs use the same set of amino acids to catalyze the reaction as family GH13 enzymes [71, 77]. The crystal structure of the inactive mutant GTF180- $\Delta$ N D1025 N bound with sucrose revealed that the seven strictly conserved residues at the active site,



**Fig. 5** Sucrose (donor substrate) binding site –1 and +1 in the crystal structure of the GTF180- $\Delta$ N mutant D1025N sucrose complex (PDB: 3HZ3) [71]. Sucrose is shown with yellow carbon atoms. Residues from domain A (blue) and B (green) surrounding the –1 and +1 subsites are indicated. Hydrogen bonds are shown as dashed lines

six also employed by GH13 enzymes, make similar interactions with the –1 glucosyl moiety of sucrose (Figs. 4, 5). The glucosyl moiety binds at the –1 site by the conserved interactions with R1023, D1025, H1135, D1136, E1063,



Y1465 and Q1509, of which R1023, D1025, H1135, D1136 and Q1509 make direct H-bonds to glucosyl hydroxyl groups. Residues D1025 and E1063 are oriented towards the glycosidic oxygen and the anomeric C1 atom of the glucosyl moiety of sucrose. Residue D1025 acts as the nucleophilic residue which attacks the anomeric C1 carbon of the glucosyl unit of sucrose to form a  $\beta$ -glucosyl-enzyme covalent intermediate, stabilized by the transition state stabilizer (D1136), while residue E1063 is the acid/base catalyst donating a proton to facilitate the release of fructose and deprotonating the acceptor molecule to activate it. Residue Y1465 located at the bottom of subsite  $-1$  provides hydrophobic interactions with the glucosyl moiety of sucrose. The conserved residue D1504 is outside of the catalytic center, making a H-bond to the conserved Y1465 hydroxyl group. Furthermore, subsite  $-1$  is lined by residues Q1140 and N1411, which form hydrogen bonds with the C3 hydroxyl group of the glucosyl moiety through a water molecule. This results in a pocket-shaped active site which can only accommodate one glucosyl moiety, demonstrating that GH70 GSs are able to transfer only one glucosyl unit in each reaction cycle. In contrast, a cleft-shaped active site with contiguous multiple donor subsites is present in family GH13  $\alpha$ -amylases. In GTF180- $\Delta$ N, the fructosyl moiety binds more loosely at the  $+1$  subsite, having interactions with residues E1063, W1065, N1029, D1136 and Q1140 (Fig. 5). The C3 hydroxyl group of the fructosyl moiety is involved in a hydrogen bond network with the side chains of W1065, D1136 and E1063, while N1029 and Q1140 make a hydrogen bond with the C1 and

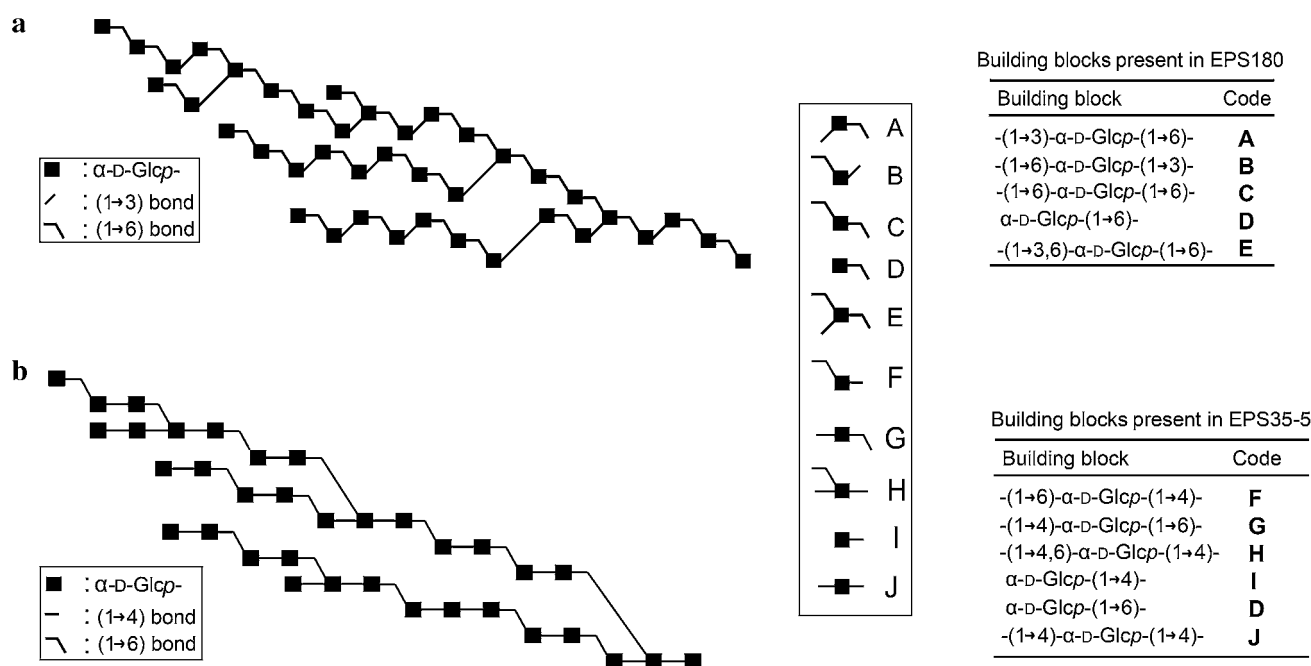
C6 hydroxyl groups, respectively. In addition, two hydrophobic residues (L981 and L982) from domain B make Van der Waals' interactions with the fructosyl moiety [71].

### Mechanism of $\alpha$ -glucan synthesis by family GH70 GS enzymes

Despite the availability of crystal structures, the mechanism of glucan synthesis by GS is not fully understood yet, especially regarding the initiation of  $\alpha$ -glucan synthesis and the synthesis process, the mode of action (processive versus non-processive), the formation of branches and the linkage specificity.

#### Initiation of $\alpha$ -glucan synthesis

An intriguing question is how  $\alpha$ -glucan synthesis by GSs is initiated, starting from sucrose. In other words, how does the  $\alpha$ -glucan chain grow during the synthesis process? The structures of the GTFA and GTF180  $\alpha$ -glucan polysaccharides have been examined by 1D and 2D NMR spectroscopy, together with methylation analysis. Additionally, structural analysis has been performed on isolated oligosaccharides obtained by enzymatic hydrolysis, partial acid hydrolysis and Smith degradation of polysaccharides [59, 81–83]. Composite models of the  $\alpha$ -glucans produced by wild-type and mutant GSs were constructed by combining all the information derived from the above-

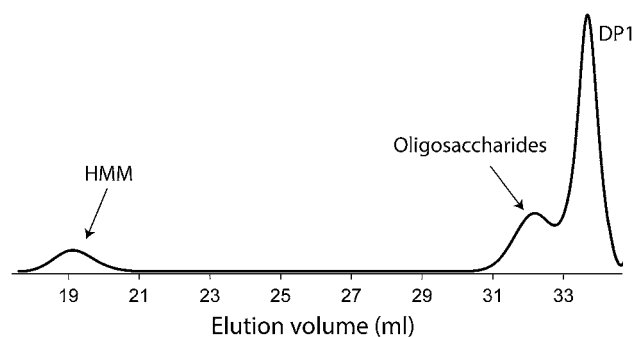


**Fig. 6** Visual representation of the composite models of the  $\alpha$ -glucan polysaccharides produced by GTF180 (a) [82] and by GTFA (b) [59]

mentioned analysis (Fig. 6). These models provide valuable information about the structures of these  $\alpha$ -glucans and the structure–function relationships of GTFA and GTF180 enzymes. Since GSs are able to transfer only one glucose unit in each reaction cycle, it is expected that  $\alpha$ -glucans are synthesized by step-wise addition of one glucosyl unit to the growing glucan chains [13, 71]. Characterization of the oligosaccharides formed in the early phase of the reaction in time, therefore, would provide important information about the synthesis process. For that purpose, oligosaccharides formed by GTFA during the initial phase of the reaction were isolated and structurally characterized [84]. The results showed that the main oligosaccharides were sucrose elongated with alternating ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages [84]. The abundance of these linkages is also shown in the composite model of the  $\alpha$ -glucan synthesized by GTFA (Fig. 6). This demonstrated that polysaccharide synthesis starts with the transfer of glucosyl units to the non-reducing glucose end of sucrose with the formation of alternating ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages. Subsequently, once the oligosaccharides reach a certain degree of polymerization (DP), branch points are formed and are further elongated with alternating ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages. As a result, a branched  $\alpha$ -glucan with large amounts of alternating ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages is synthesized. This study provides direct evidence that sucrose is used as a primer by GTFA in the synthesis of  $\alpha$ -glucan. The low-molecular mass (DP 20–30) and linear dextran produced by GTF-S3 from *S. sobrinus* were also found to be terminated with a sucrose moiety [85]. Moulis et al. also reported that DSRS of *L. mesenteroides* NRRL B-512F and alternansucrase of *L. mesenteroides* NRRL B-1355 used sucrose as an initiator of polysaccharide synthesis [65]. However, DSRS of *L. mesenteroides* NRRL B-512F also formed a series of isomalto-oligosaccharides (acceptor products starting from glucose) reaching DP higher than 25, while oligosaccharides with sucrose at the end did not exceed a DP higher than 12 [65]. In this work, both sucrose and glucose were proposed as initial acceptor for polysaccharide synthesis, but the latter was preferred [65]. However, whether this could be extended to other GSs is not known yet.

### GH70 GSs mode of action: processive versus non-processive

There has been a controversy about the mode of glucan chain elongation (processive versus non-processive). Previously, GH70 GSs were found to synthesize high-molecular weight polysaccharides during the early phase of reaction without the detection of intermediate oligosaccharides [86]. Consequently, GH70 GSs were assumed to act processively in the synthesis of  $\alpha$ -glucan polysaccharides. Using a more



**Fig. 7** Molecular mass distribution of the products generated by incubating GTF180- $\Delta$ N (1.0 U/ml) with 0.1 M sucrose at 37 °C and pH 4.5. The HPSEC profile corresponds to the product mixture obtained after sucrose depletion. *HMM* high-molecular-mass polysaccharides, *DP1* monosaccharides (mainly fructose). This figure has been adapted from [99]

sensitive method (high-performance anion-exchange chromatography equipped with an ED40 pulsed amperometric detection system, HPAEC-PAD), oligosaccharides were detected later on, which implied that GH70 GSs act non-processively as well [65]. Size-exclusion chromatography analysis showed that product mixtures formed by incubation of sucrose with GS enzymes generally contain high-molecular-mass (HMM) glucan and low-molecular-mass (LMM) oligosaccharides (Fig. 7) [65]; no intermediate size  $\alpha$ -glucan products were detected. Kinetic analysis of polysaccharide synthesis by DSRS of *L. mesenteroides* NRRL B-512F revealed that HMM dextrans reached the maximum size after only 23 % of sucrose consumption with the simultaneous detection of oligosaccharides [65]. The polysaccharide synthesized by GTFA was also found to reach its maximum size in a relatively short time, with the simultaneous detection of oligosaccharides [84]. The reuteran polysaccharide size did not increase further, not even with the availability of excess sucrose [87]. The detection of HMM polysaccharide with a maximum size indicates a processive mode, while the detection of oligosaccharides points at non-processive mode. Therefore, taking into account all the information, Moulis et al. proposed a semi-processive mechanism of polymerization for GH70 GSs [65]. In the initial phase of the reaction, GH70 GSs catalyze the synthesis of oligosaccharides in a non-processive mode; when the oligosaccharides reach a certain size, polysaccharide synthesis proceeds in a processive mode. The structural basis for processive polysaccharide synthesis and non-processive oligosaccharide synthesis is proposed to be located in the repeat units in the C- and N-termini, which have been shown to be involved in glucan binding [56, 67, 88]. Mutants of DSRS with truncated A repeats (WYYFN XDGQAATGLQTIDGQTVFDDNGXQVG) were less efficient in polysaccharide synthesis compared with the wild-type enzyme [65]. This suggests that these repeats facilitate

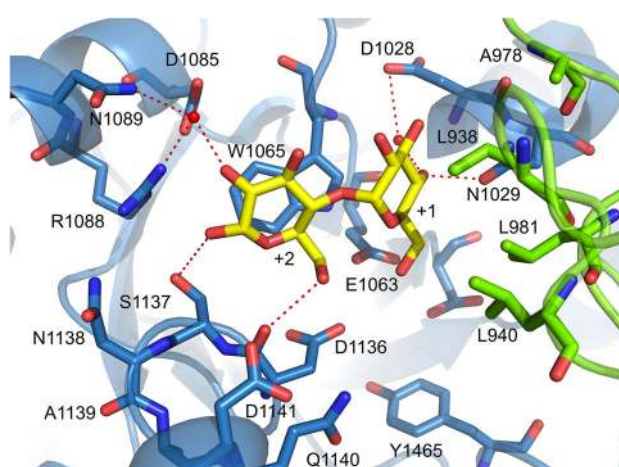
polysaccharide synthesis by anchoring the growing polysaccharide chain close to the active site. With the availability of the crystal structures, it became clear that the sequence repeats present in domain V formed a modular  $\beta$ -solenoid fold and domain V was proposed to play a role in carbohydrate binding [71, 72]. Indeed, truncation of domain V from GTF180- $\Delta$ N heavily impaired polysaccharide synthesis and increased the oligosaccharide synthesis [80]. Interestingly, mutation of residues around the acceptor binding sites partially restored the polysaccharide synthesis of GTF180- $\Delta$ N $\Delta$ V [80]. Furthermore, several mutants of GSs, targeting amino acid residues close to the acceptor binding site, have been reported to produce different ratios of polysaccharide, oligosaccharide and glucose from sucrose [65, 89–92]. For instance, S628 (a residue C-terminal to the transition state stabilizer) mutants of GTFR (S628D and S628R) from *S. oralis* abolished polysaccharide synthesis and only produced short chain oligosaccharides [91]. Moulis et al. also reported that mutations of the residues C-terminal to the transition state stabilizer of DSRS from *L. mesenteroides* NRRL B-512F and alternansucrase from *L. mesenteroides* NRRL B-1355 abolished or reduced polysaccharide synthesis [65]. The polysaccharide synthesis increased significantly (about 2 fold) by mutation of residues around the acceptor substrate binding site (L940) of GTF180- $\Delta$ N [92]. These residues locate at acceptor substrate binding site and mutations at these subsites may affect the affinity of the enzyme with the growing glucan chain, resulting in decrease or increase of polysaccharide synthesis. These results suggest that the structural basis for processive polysaccharide synthesis lies in both domain V, and the acceptor substrate binding sites, representing remote and close binding sites for glucan chains, respectively. The glucan chain may also bind at the protein surface located in between these sites. Elucidation of the crystal structures of GH70 GSs with bound larger size glucan chains in future studies may reveal more information about such intermediate sites. However, previous attempts have so far been unsuccessful, indicating that obtaining such crystal structures is not trivial and may require other strategies.

It has also been shown that the reaction conditions, especially sucrose concentration, have significant effect on the product distribution of GSs. Kim et al. showed that the amount of HMM dextran produced by dextransucrase from *L. mesenteroides* B-512F/MCM decreased while the amount of LMM dextran increased with increasing sucrose concentrations [93]. Dextrans of different molecular mass were synthesized at controlled sucrose concentrations, enzyme concentrations and reaction temperatures by dextransucrase from *L. mesenteroides* B-512F/MC [94]. The ratio of oligosaccharide synthesis versus polysaccharide synthesis of GTFA from *L. reuteri* 121 is directly proportional to the concentration of sucrose. However, the sizes of the

polysaccharides produced at different sucrose concentrations were identical [87]. At present, it remains unknown what determines the molecular size of the  $\alpha$ -glucan polysaccharides synthesized by GS enzymes.

### Formation of branches

Most of the  $\alpha$ -glucans synthesized by GSs are branched, to varying degrees [13, 15, 16]. Most GSs do not require an extra enzyme for the formation of branches in their  $\alpha$ -glucan products. The mechanism for forming branch points remains unclear. Robyt and Taniguchi proposed that the formation of branched glucosyl units is through the acceptor reaction of GSs [95]. Site-directed mutagenesis has identified several amino acid residues that are involved in the formation of branched glucosyl units [96]. Mutation of three residues (S1137:N1138:A1139) following the transition state stabilizer residue (D1136) in GTF180 resulted in several mutants which synthesized  $\alpha$ -glucans with a higher degree of branches [81]. These residues are located close to the +2 glucosyl unit of maltose in the crystal structure of GTF180- $\Delta$ N in complex with maltose (Fig. 8) [71]. Residues D1085, R1088 and N1089 from  $\alpha$ -helix 4 are located at the other side of the +2 glucosyl unit of maltose and they all make an indirect hydrogen bond with the +2 glucosyl unit of maltose through the same water molecule (Fig. 8). Multiple and single mutations in these residues resulted in mutants producing hyper-branched  $\alpha$ -glucans (15–22 % branching) compared to that of wild-type (13 %) [97]. Single mutation studies showed that D1085 and R1088, but not N1089, are responsible for the increase in branched glucosyl units. Irague et al. also



**Fig. 8** Maltose binding sites +1 and +2 in the crystal structure of the GTF180- $\Delta$ N maltose complex (PDB: 3KLL) [71]. Maltose is shown with yellow carbon atoms. Residues from domain A (blue) and B (green) surrounding the +1 and +2 subsites are indicated. Hydrogen bonds are shown as dashed lines

reported that mutations in the corresponding residues of DSRS of *L. mesenteroides* NRRL B-512F (D460, H463 and T464) increased the proportion of ( $\alpha 1 \rightarrow 3$ ) linkages [98]. Mutational studies of A978 and D1028 in GTF180 identified the involvement of a groove preceding the +1 subsite in the formation of branched glucosyl units [99]. Partially blocking the groove by mutating these residues to amino acid residues with larger side chains reduced the amount of branching linkages in the  $\alpha$ -glucans synthesized. To summarize, the proportion of branched points in the  $\alpha$ -glucans produced by GSs could be manipulated by mutations close to the acceptor substrate binding subsites. Thus, GS enzymes indeed may form branching linkages through the acceptor reaction, involving dissociation of the growing glucan chain from the acceptor binding subsites and subsequent rebinding in a different way that allows the formation of branch points. However, the question, when and where branching linkages are formed, remains unanswered.

In addition, ( $\alpha 1 \rightarrow 2$ ) branching GH70 sucrases have been reported [28, 35, 37]. With only sucrose as substrate, they are unable to polymerize glucosyl units and mainly catalyze the hydrolysis of sucrose and the formation of leucrose [28, 35, 37]. However, using sucrose as donor substrate and dextran as acceptor substrate, they catalyze the synthesis of ( $\alpha 1 \rightarrow 2$ ) branching linkages onto the dextran acceptor substrate [28, 35, 37]. Interestingly, ( $\alpha 1 \rightarrow 3$ ) branching GH70 sucrases have also been reported recently [100]. Structure–function relationship studies of these branching sucrases will provide valuable information about the mechanism of branching linkage formation.

### Linkage specificity determination of family GH70 GS enzymes

Although the active site of GSs is highly conserved, they produce  $\alpha$ -glucans with different structures especially regarding glycosidic linkages [13, 16, 96]. These various  $\alpha$ -glucans possess different physico-chemical properties such as molecular mass, solubility and viscosity. For instance, the solubility of mutan containing mainly ( $\alpha 1 \rightarrow 3$ ) linkages is generally low, while dextran with predominantly ( $\alpha 1 \rightarrow 6$ ) linkages is more soluble. Due to their ability to produce a diverse range of  $\alpha$ -glucans, GSs have attracted interest for industrial applications in food, medicine, cosmetics, etc. [7, 34]. Dextran produced by the DSRS from *L. mesenteroides* NRRL B-512F is extensively applied as gelling, viscosifying and emulsifying agent in the food industry. Bakery products with dextran have improved softness and increased volume [101]. Dextran is also applied as size-exclusion chromatography material in

research, and as a plasma expander in medicine [20, 102, 103]. Various  $\alpha$ -glucans have been shown to possess anti-corrosion activity possibly by forming biofilms on the surface of steel [55, 104, 105]. Moreover,  $\alpha$ -glucans and oligosaccharides formed by GSs have potential prebiotic activities and, therefore, can be used to stimulate growth of beneficial intestinal bacteria such as *Bifidobacterium* and *Lactobacillus* [106]. Linear and ( $\alpha 1 \rightarrow 2$ )-branched dextran produced by the crude GS enzyme preparation from *L. citreum* NRRL B-1299 were reported to increase *Bifidobacterium* populations in vitro [107]. Monsan et al. reported that the addition of  $\alpha$ -glucan to animal feed improved the weight gain of pig and broilers [108]. Therefore, the large variety of  $\alpha$ -glucans and oligosaccharides with different structures holds great potential for industrial applications. A more detailed understanding of the linkage specificity of GSs may allow production of tailor-made  $\alpha$ -glucans with desired properties.

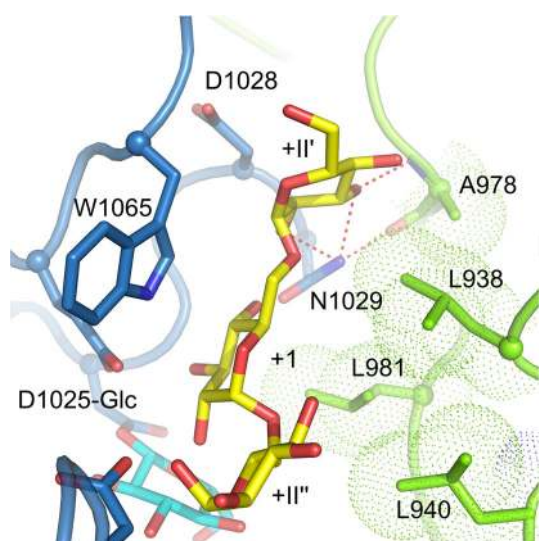
The linkage specificity of GS enzymes appears to be determined by only a small number of amino acids [56, 90, 96]. As discussed above, all GSs initiate the reaction cycle by formation of a covalent C1-linked glucosyl-enzyme intermediate [15]. Consequently, it has been proposed that it is the way in which acceptor substrates bind at acceptor binding subsites that determine the linkage specificity of a GS enzyme, revealing the importance of amino acid residues at these sites [56, 65, 90, 96]. Prior to the availability of crystal structures, putative regions involved in acceptor binding site have been identified by alignment with family GH13 enzymes in view of their high sequence similarity [65, 89, 90]. These include residues C-terminal to the catalytic nucleophile D1025 (GTF180 numbering, conserved sequence motif II, Fig. 4), residues C-terminal to the acid base catalyst E1063 (motif III, Fig. 4) and residues C-terminal to the transition state stabilizer D1136 (motif IV, Fig. 4). These motifs display amino acid variations in different GSs [15]. Indeed, mutagenesis studies targeting these residues confirmed their importance for the product specificity of GSs. GTFA mutants P1026 V and I1029 V (motif II, GTFA numbering) displayed different product spectra with sucrose as substrate yielding higher levels of isomaltose and leucrose production [90]. Mutants A1066 N and H1065S:A1066S of GTFA (region III, GTFA numbering) produced similar products as wild type, but displayed lower activity [90]. In several GH70 GSs, residues located C-terminal to the catalytic transition state stabilizer (N1134:N1135:S1136: Q1137:D1138, motif IV, GTFA numbering) have been found to be critical for linkage specificity [65, 81, 90, 91]. Combined mutations in the tripeptide (N1134:N1135:S1136, motif IV, GTFA numbering) following the transition state stabilizing residue (D1133) in GTFA from *L. reuteri* 121 shifted GTFA linkage specificity from mainly ( $\alpha 1 \rightarrow 4$ ) linkages to



( $\alpha 1 \rightarrow 6$ ) linkages, indicating their involvement in acceptor substrate binding and hence linkage specificity determination [90]. Further mutation studies showed that, among these three residues, residue N1134 plays a major role in linkage specificity determination [89]. Similarly, mutations in the corresponding tripeptide in GTFR from *S. sobris* [91], GTF180 from *L. reuteri* 180 [81], and DSRS from *L. mesenteroides* NRRL B-512F [65] altered their linkage specificity as well. Mutation of the residues near the transition state stabilizer in GTFR (R628G:V630I:D717A) resulted in an increase of ( $\alpha 1 \rightarrow 3$ ) linkages in the polysaccharide produced [91]. Surprisingly, combining the mutations in motifs II and IV in GTF180, mutant V1027P:S1137 N:A1139S introduced 12 % ( $\alpha 1 \rightarrow 4$ ) linkages (not present in the wild-type) in the  $\alpha$ -glucan produced [83]. Mutants of the fourth residue located at C-terminal transition state stabilizer of GTF180 (Q1140A and Q1140H) produced  $\alpha$ -glucans with a higher percentage of ( $\alpha 1 \rightarrow 6$ ) linkages. The GTF180 Q1140E mutant also produced  $\alpha$ -glucan with 3 % ( $\alpha 1 \rightarrow 4$ ) linkages [81]. In addition, the fifth residue following the C-terminal transition state stabilizer was also shown to be involved in linkage specificity determination. Random mutagenesis of D569 in GTF-I of *S. downei* showed that mutations at this position affected the structure of the  $\alpha$ -glucan and the size of the synthesized oligosaccharides [109]. Mutations of the fifth residue following the transition state stabilizer in DSRI of *L. mesenteroides* NRRL B-1118 also affected the linkage composition of the products, with mostly increased ( $\alpha 1 \rightarrow 3$ ) linkages [110]. These results demonstrate that the involvement of residues located C-terminal to the transition state stabilizing residue in acceptor substrate binding is a general feature of GH70 GSs.

Elucidation of the crystal structure of GTF180- $\Delta$ N with a bound maltose acceptor substrate provided structural explanations for the effects of mutations observed previously [36, 71, 72] and confirmed that the amino acid residues at the acceptor binding site are critical for linkage specificity. Four bound maltose molecules (M1, M2, M3 and M4) were revealed by soaking GTF180- $\Delta$ N crystals with maltose [71]. However, the residues forming the binding sites for M2, M3 and M4 are not conserved within family GH70 GS enzymes and only have nonspecific interactions with GTF180- $\Delta$ N amino acid residues. Maltose M1 binds at the acceptor binding sites +1 and +2 with its C6 hydroxyl group of non-reducing end glucosyl moiety pointing towards the sucrose binding pocket. This C6 hydroxyl group activated by the acid/base catalyst (E1063) attacks the C1 of the glucosyl-enzyme intermediate, resulting in the formation of an ( $\alpha 1 \rightarrow 6$ ) linkage (Fig. 8). Thus, the binding mode of M1 explains how an ( $\alpha 1 \rightarrow 6$ ) linkage is formed with maltose as acceptor substrate

(Fig. 8). At subsite +1, N1029 makes direct and indirect hydrogen bonds with the +1 C4 and C3 hydroxyl group, respectively; D1028 forms an indirect hydrogen bond with the +1 C4 hydroxyl group; residues from domain B (L938, L940, A978 and L981) shape the groove near the +1 subsite. These residues have rarely been targeted for mutagenesis studies prior to the availability of crystal structures [96]. In recent site-directed mutagenesis studies, these residues, especially residues from domain B, were found to be critical for linkage specificity and to display different roles [92, 99]. The L940 mutants synthesized  $\alpha$ -glucans with larger amounts of ( $\alpha 1 \rightarrow 6$ ) linkages [92]. Surprisingly, the L940 W mutant produced linear  $\alpha$ -glucan with only ( $\alpha 1 \rightarrow 6$ ) linkages; the synthesis of ( $\alpha 1 \rightarrow 3$ ) linkages was abolished completely. Docking studies with isomaltotriose showed that this tryptophan blocks a groove, preventing the reducing end of isomaltotriose to occupy the space observed in the wild-type [92]. Consequently, the C3 hydroxyl group of the non-reducing end glucose unit is too far away to attack the C1 of the glucosyl-enzyme intermediate. Instead, the C6 hydroxyl group is within the distance for forming ( $\alpha 1 \rightarrow 6$ ) linkages. This highlights the critical importance of the groove, where L940 is located, for ( $\alpha 1 \rightarrow 3$ ) linkage synthesis in GTF180. Residue A978 was found to be involved in branch points formation. Mutations of A978 to residues with a larger side chain (Leu, Pro, Phe and Tyr) reduced the branch formation in the  $\alpha$ -glucans produced, while mutations to small residues (Gly and Ser) had no significant effects [99]. All D1028 mutants had increased amounts of ( $\alpha 1 \rightarrow 6$ ) linkages in their  $\alpha$ -glucan products. The branching linkages were also influenced by D1028 mutations. Again, mutations to residues with a large side chain (Tyr and Trp) reduced the number of branches [99]. Docking studies showed that both A978 and D1028 are involved in shaping the groove above the +1 subsite, the space of which is required for acceptor substrate binding for formation of branching linkages (Fig. 9). Mutation of these two residues to bulky residues apparently partially blocked the groove, resulting in a decrease in branch point formation. L938 mutants were also shown to produce  $\alpha$ -glucan with mostly increased ( $\alpha 1 \rightarrow 6$ ) linkages and thus also involved in linkage specificity. Mutations in N1029 resulted in an increase of ( $\alpha 1 \rightarrow 3$ ) linkages in the  $\alpha$ -glucan products [99]. At acceptor binding site +2, residues S1137, N1138, A1139, Q1140, D1141, which have been shown to be important for linkage specificity, are located at one side of the +2 glucosyl unit (Fig. 8). It is worth to note that only S1137 has a direct hydrogen bond with the +2 C1 hydroxyl group. Indeed, S1137 was found to be the main determinant for linkage specificity in previous mutation studies [65, 81, 89–91]. This confirms the predicted involvement of these residues at this acceptor binding site. At the other side of



**Fig. 9** View of a docked isomaltotriose (yellow carbon atoms) in the active site of a modeled GTF180- $\Delta$ N glucosyl-enzyme intermediate (cyan atoms) [71], at the interface of domain A (blue) and B (green). The trisaccharide occupies subsites +1, +II' and +II''; the latter two subsites are different from subsite +2 observed in the maltose complex

the +2 glucosyl unit, residues D1085, R1088 and N1089 all possess an indirect hydrogen bond with the +2 C2 hydroxyl group through the same water molecule (Fig. 8). Single and combinatorial mutations studies showed that these residues are also involved in linkage specificity determination [97]. Mutations at these three residues, especially D1085 and R1088, introduced extra ( $\alpha$ 1  $\rightarrow$  4) linkages in the  $\alpha$ -glucans produced. Remarkably, some mutants also produce hyper-branched  $\alpha$ -glucans with 15–22 % branched glucosyl units. Single mutational studies showed that residues D1085 and R1088, but not N1089, were responsible for the increase in branching linkages. Another combinatorial mutagenesis study, targeting the same residues in DSRS, also demonstrated their importance [98]. In addition, residue W1065 has hydrophobic stacking interactions with both the +1 and +2 glucosyl moiety of maltose [71]. In the complex of GTF180- $\Delta$ N D1025 N with sucrose, this residue also has a direct hydrogen bond with the C1 hydroxyl group of the fructosyl moiety [71]. Mutating W491 of GTFI from *Streptococcus mutans* (W1065 in GTF180- $\Delta$ N) to either glycine or alanine resulted in an enzyme devoid of detectable activity [111], demonstrating its essential role in the catalytic activities. Similarly, mutating W1065 of GTF180- $\Delta$ N to non-aromatic amino acid residues heavily impaired the enzyme activity and abolished the synthesis polysaccharides (unpublished results). Mutant W1065F still retained the ability to synthesize polysaccharides with altered linkage composition. By now, only acceptor binding sites +1 and +2 have been mapped out and no crystal structure

of a GH70 GS proteins with a bound higher molecular mass acceptor substrate is available; it remains unclear whether amino acid residues at further acceptor binding sites are determinants for linkage specificity. These remote acceptor binding sites are difficult to identify by sequence alignment or to predict based on a crystal structure, because they may not be conserved.

In some cases, it has been shown that the next linkage type is also determined by the previous linkage formed. One example is the alternansucrase (ASR) from *L. mesenteroides* NRRL B-1355, which catalyzes the synthesis of alternan with alternating ( $\alpha$ 1  $\rightarrow$  6) and ( $\alpha$ 1  $\rightarrow$  3) linkages. This suggests that an acceptor substrate with an ( $\alpha$ 1  $\rightarrow$  6) linkage between the +1 and +2 subsites favors the formation of an ( $\alpha$ 1  $\rightarrow$  3) linkage in the next reaction cycle; in turn, an ( $\alpha$ 1  $\rightarrow$  3) linkage at this position induces the formation of an ( $\alpha$ 1  $\rightarrow$  6) linkage. It has been proposed that the non-reducing end of the acceptor substrate is not well stabilized at the +1 subsite, while the glucosyl residue at the +2 site of ASR is more stabilized due to the stacking interaction with W675 and/or Y768 (Fig. 4) [65]. Residue Y768 is unique for ASR and has been proposed to provide a 2nd stacking platform for a glucosyl residue at the +2 subsite. Thus, the linkage between the +1 and +2 glucosyl residues of a bound acceptor substrate probably determines that the +2 glucosyl residue stacks with either residue W675 or Y768, resulting in the accessibility of the C3 or C6 hydroxyl group of the non-reducing end glucosyl residue to the glucosyl-enzyme intermediate. Mutant Y768S:D769E:A770 V of the alternansucrase of *L. mesenteroides* NRRL B-1355 was unable to synthesize alternan and produced more oligo-dextrans from sucrose and maltose [65]. Another example is GTFA from *L. reuteri* 121 [59]; characterization of the oligosaccharides initially produced by GTFA revealed that the most prominent products were oligosaccharides with alternating ( $\alpha$ 1  $\rightarrow$  6) and ( $\alpha$ 1  $\rightarrow$  4) linkages [84]. However, GTFO from *L. reuteri* ATCC 55730, which shares high similarity with GTFA, synthesizes a reuteran with a high amount of ( $\alpha$ 1  $\rightarrow$  4) linkages instead of alternating ( $\alpha$ 1  $\rightarrow$  6) and ( $\alpha$ 1  $\rightarrow$  4) linkages. The determinants of the different linkage specificity of GTFA and GTFO are currently under investigation. GTF180 from *L. reuteri* 180 synthesizes an  $\alpha$ -glucan with ( $\alpha$ 1  $\rightarrow$  6) and ( $\alpha$ 1  $\rightarrow$  3) linkages [57]. Structural analysis of this  $\alpha$ -glucan showed that it is built up with different lengths of isomalto-oligosaccharides, interconnected by single ( $\alpha$ 1  $\rightarrow$  3) linkages [82]. All (-)- $\alpha$ -D-Glcp-(1  $\rightarrow$  3)- units were found to be 6-substituted and no consecutive ( $\alpha$ 1  $\rightarrow$  3) linkages were found [82]. Taken together, these examples show that the linkage specificity of GH70 GSs is also depending on the linkage present between the +1 and +2 acceptor substrate binding site. However, unlike the situation in ASR, no 2nd aromatic

residue was found at subsite +2 in GTFA and GTF180; multiple mechanisms of alternating linkage formation thus occur.

To conclude, these mutagenesis studies showed that the linkage specificity of GH70 GSs is determined by an interplay of different amino acid residues from both domains A and B, shaping the acceptor binding sites. The specific interactions between the acceptor substrate and its acceptor binding sites determine which hydroxyl group of the non-reducing end glucosyl moiety of an acceptor substrate is capable of attacking the glucosyl-enzyme intermediate to form the next  $\alpha$ -glycosidic linkage. Therefore, it is not one but several amino acid residues of GH70 GSs that determine their enzyme linkage specificity. This explains why, even with different amino acid residues at a certain position, different GSs still may have similar linkage specificities; and with the same amino acid residues at a certain position, they catalyze the synthesis of different linkages. Although the contribution of different residues complicates the rationalization of specificity in GSs, 3D structure-guided mutagenesis is an effective approach for changing the linkage specificity of GS enzymes and producing novel  $\alpha$ -glucans. Together, the various mutagenesis studies, supported by 3D structural observations, contribute to a better understanding of the GH70 GS linkage specificity.

### Biochemical properties of GTFB- and GTFC-like 4,6 $\alpha$ -glucanotransferase enzymes

Although sucrose remains the canonical D-glucosyl donor substrate for GSs, some GS enzymes also can use oligosaccharides as donor/acceptor substrates [112]. In this so-called disproportionation reaction, the glucosyl group is transferred from one saccharide donor to an identical or similar saccharide acceptor substrate [112]. Binder et al. found that dextranase from *L. mesenteroides* NRRL B-512F and GTF-S from *S. mutans* 6715 have disproportionation activity with isomalto-oligosaccharides, malto-oligosaccharides and panose.

### GTFB-like 4,6 $\alpha$ -glucanotransferase enzymes

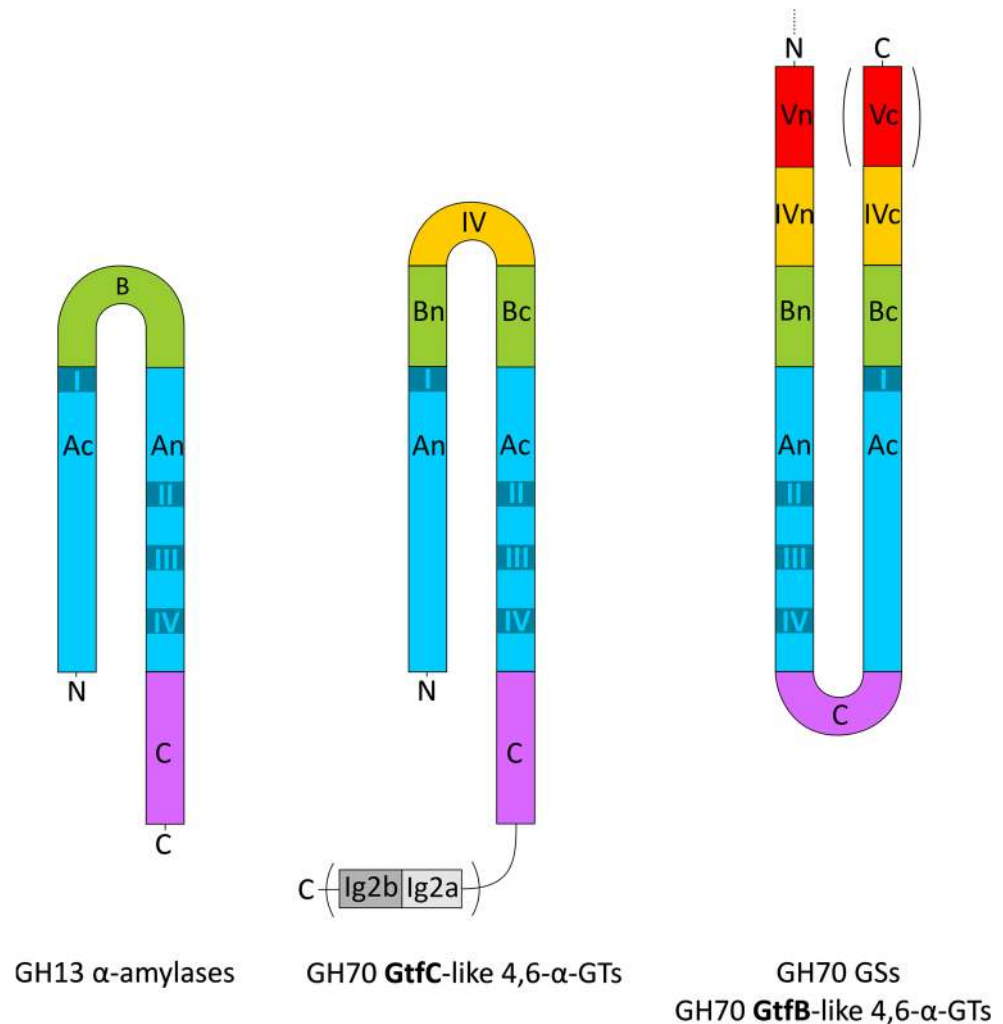
This disproportionation activity of GSs has been overlooked for decades due to its relatively low activity compared to the glucosyl transfer from sucrose. However, it was found that *Lactobacillus reuteri* 121 possessed a GS-like enzyme which was inactive on sucrose, and instead used malto-oligosaccharides as donor and acceptor substrates [18]. The gene encoding this enzyme designated as GTFB, is located upstream the gene encoding the GTFA

that catalyses the synthesis of reuteran with ( $\alpha 1 \rightarrow 4$ ) and ( $\alpha 1 \rightarrow 6$ ) linkages from sucrose [58]. GTFB predominantly cleaves an ( $\alpha 1 \rightarrow 4$ ) glycosidic linkage from the non-reducing end of the donor substrate [( $\alpha 1 \rightarrow 4$ )-glucan] and transfers the cleaved glucosyl unit to the non-reducing end of another ( $\alpha 1 \rightarrow 4$ )-glucan acceptor substrate, forming mainly ( $\alpha 1 \rightarrow 6$ ) linkages. Products formed with an ( $\alpha 1 \rightarrow 6$ ) linkage at the non-reducing end become better acceptor substrates and are further elongated in a linear manner with ( $\alpha 1 \rightarrow 6$ ) linked glucosyl units. This results in the formation of isomalto/malto-oligosaccharide and polysaccharide mixtures with increasing percentages of ( $\alpha 1 \rightarrow 6$ ) linkages [113, 114]. Two more GTFB-like enzymes (GTFW from *L. reuteri* DSM 20016 and GTFML4 from *L. reuteri* ML1) were characterized later on [18, 114]. In view of their clearly distinct reaction specificity, these enzymes have been designated as 4,6- $\alpha$ -glucanotransferases (4,6- $\alpha$ -GTs) (EC 2.4.1.-). Besides these 3 biochemically characterized GTFB-like 4,6- $\alpha$ -GTs (GTFB, GTFW and GTFML4), 46 putative GTFB-type 4,6- $\alpha$ -GTs are currently found in the GenBank database. Except for three from *Pediococcus* strains, they are all found within the genus *Lactobacillus*.

Interestingly, GTFB converts amylose and high amylose starches into isomalto-/malto-polysaccharides (IMMP) with high percentages of ( $\alpha 1 \rightarrow 6$ ) linkages [115]. For example, incubation of the amylose V of potato with GTFB yielded a linear product with 91 % of ( $\alpha 1 \rightarrow 6$ ) linkages. The degree of branching in starch substrates negatively correlates with the amount of ( $\alpha 1 \rightarrow 6$ ) linkages in the produced IMMP; it was, therefore, postulated that the GTFB enzyme cannot get across the branch points and, thus, cannot act on internal ( $\alpha 1 \rightarrow 4$ ) linkages [115]. This was further confirmed by the fact that when using debranched starches as substrate, GTFB introduced a higher percentage of ( $\alpha 1 \rightarrow 6$ ) linkages in the IMMP product [115]. Structural and size analysis showed that the IMMP constituted an entirely novel class of  $\alpha$ -glucans, e.g., different from the known dextrans and isomalto-oligosaccharides. These IMMP are proven soluble dietary fibers since the segments rich in ( $\alpha 1 \rightarrow 6$ ) linkages largely pass the human small intestine and enter into the large intestine [115].

The GTFB-like 4,6- $\alpha$ -GT enzymes show about 50 % amino acid sequence identity with GH70 GSs and clearly belong to family GH70 [18, 114]. Primary structure analysis revealed that GTFB-like 4,6- $\alpha$ -GTs, like GH70 GSs, have the same domain organization in that domains A, B, C and IV are made up from discontinuous N- and C-terminal stretches of the polypeptide chain (Fig. 10). It should be noted that, while in most GH70 GSs domain V consists of both N- and C-terminal polypeptide segments, domain V of GTFB is smaller since it only consists of an N-terminal

**Fig. 10** Schematic representation of the domain organization of GH13  $\alpha$ -amylases and GH70 4,6- $\alpha$ -glucanotransferases and GH70 GSs. Domains A, B, C, IV, and V are colored in blue, green, magenta, yellow, and red, respectively. Ig2-like domains are indicated in gray. Brackets indicate that the Ig2-like domains are not identified in all GtfC-like enzymes, and that the C-terminal half of domain V is not found in all GtfB-like 4,6- $\alpha$ -glucanotransferases [17]. This figure has been adapted from [17]



polypeptide segment. The catalytic domain of GTFB-like 4,6- $\alpha$ -GTs also consists of a circularly permuted  $(\beta/\alpha)_8$  barrel with conserved sequence motifs I to IV as in family GH70 GS enzymes (Fig. 10). The six conserved residues in conserved sequence motifs I to IV including the three catalytic residues (the nucleophile D1025, the acid/base catalyst E1063 and the transition state stabilizer D1136, GTF180 numbering) in GH70 GSs and family GH13 enzymes are also present in GTFB-like 4,6- $\alpha$ -GTs (Fig. 4). Similar to the GH70 GSs, the seventh conserved residue (His134,  $\alpha$ -amylase of *Bacillus licheniformis* numbering) of family GH13 enzymes is also replaced by a Gln residue in GTFB-like 4,6- $\alpha$ -GTs. Besides the similarities, a large number of amino acid residues in motifs I to IV of GTFB-like 4,6- $\alpha$ -GTs show variations as compared to GH70 GSs, especially for the residues contributing to the -1, +1, and +2 donor/acceptor binding subsites in motifs I to IV (Fig. 4). For example, amino acid residues downstream of the acid/base catalyst are completely different in GH70 GSs and GTFB-like 4,6- $\alpha$ -GTs (Fig. 4). Specifically,

residue W1065 (GTF180 numbering) which is highly conserved in GH70 GSs (except for DSRECD2) is replaced by a Tyr residue in GTFB-like 4,6- $\alpha$ -GTs. Residues located C-terminal of the transition state stabilizing residue in conserved sequence motif IV, part of acceptor subsite +2 in GH70 GSs and critical for their linkage specificity, are also very different in GTFB-like 4,6- $\alpha$ -GTs. These differences in amino acid residues may contribute to the distinct linkage specificity of GTFB-like 4,6- $\alpha$ -GTs. Further mutational studies may allow elucidation of the functional roles of these amino acid residues in determining the substrate specificity and product linkage specificity of GTFB-like 4,6- $\alpha$ -GTs.

### GTFC-like 4,6- $\alpha$ -glucanotransferase enzymes

Recently, we have reported the identification of a second novel GH70 subfamily with 4,6- $\alpha$ -glucanotransferase activity on maltodextrins and starch (designated as GTFC-



like 4,6- $\alpha$ -GTs), present in *Exiguobacterium* and *Bacillus* strains. Both *Exiguobacterium* and *Bacillus* are genera within the class Bacilli and members of the low GC phylum of Firmicutes, but they do not belong to the lactic acid bacteria. The *Exiguobacterium sibiricum* 255-15 GTFC was selected as the first representative member of this group and characterized in detail [17]. *E. sibiricum* 255-15 is a free-living, psychrotrophic, nonsporulating, Gram-positive bacterium that was isolated from a 3 million year old Siberian permafrost core [116, 117]. This bacterium was found to produce exopolysaccharides, proposed to be the result of the activity of a putative GH70 enzyme, later designated as GTFC, identified in its genome [116, 117]. Analysis of sequence-based relatives of the *E. sibiricum* GTFC, using BLAST, revealed that the closest homologs of these GTFC-like 4,6- $\alpha$ -GTs are GTFB-like proteins (with statistically significant E values), even though they share only 30 % of sequence identity. Similar to the GTFB-like 4,6- $\alpha$ -GTs, GTFC catalyzes cleavage of ( $\alpha$ 1  $\rightarrow$  4) glycosidic linkages and synthesis of consecutive ( $\alpha$ 1  $\rightarrow$  6) linkages. GTFC differs from GTFB in converting amylose/starch substrates into Isomalto-/malto-oligosaccharides (IMMO), instead of the (modified) polymers (IMMP) synthesized by GTFB [115]. The similar reactions catalyzed by these two enzyme types correlate well with the highly conserved motifs I to IV, particularly in residues forming the acceptor binding subsites (Fig. 4). Similar to GTFB-like proteins, a Tyr residue in GTFC-like 4,6- $\alpha$ -GTs proteins replaces the subsite +1/+2 Trp residue conserved in almost all GH70 GSs (W1065 in GTF180- $\Delta$ N). The amino acids following the putative transition state stabilizer residue in GTFC homologs are very different from those in GH70 GSs (Fig. 4). These amino acid residues clearly share higher similarity with GTFB-like 4,6- $\alpha$ -GTs. GTFB- and GTFC-like 4,6- $\alpha$ -GTs all have an one amino acid gap at position 1139 (GTF180 numbering) and share residues Gln, Lys and Asn at positions 1137, 1140 and 1141, respectively. However, at position 1138, GTFB-like 4,6- $\alpha$ -GTs all have an Arg residue, while in GTFC-like 4,6- $\alpha$ -GTs a Glu residue is present. Most importantly, GTFC-like 4,6- $\alpha$ -GTs differ from GTFB-like 4,6- $\alpha$ -GTs and GH70 GSs by a non-circular permutation of conserved sequence motifs I to IV (Fig. 10). In GTFC-like 4,6- $\alpha$ -GTs, the order of these conserved sequence motifs is identical to that in GH13 (I-II-III-IV). Thus, the domain organization of the catalytic core (domains A, B, and C) of these GTFC-like 4,6- $\alpha$ -GTs resembles that of GH13  $\alpha$ -amylases, lacking the circular permutation of the ( $\beta/\alpha$ )<sub>8</sub> barrel (Fig. 10). A striking feature of these GTFC-like 4,6- $\alpha$ -GTs, with respect to GH13 amylases, is that they possess an extra contiguous domain IV interrupting domain B (Fig. 10). Domain IV is exclusively found in family GH70 enzymes and thus far its function has remained unknown. In

addition, several (but not all) GTFC-like proteins contain one or two Ig2-like domains of unknown function, whereas they lack the variable N-terminal domain and domain V typically present in GH70 members. In view of the clear sequence similarity between GTFC-like 4,6- $\alpha$ -GTs and other GH70 enzymes, it was found more appropriate to classify these GTFC-like proteins as a new subfamily in family GH70 than to establish an entirely new family. They thus constitute the first subfamily GH70 enzymes that lack the circular permutation of the ( $\beta/\alpha$ )<sub>8</sub> barrel that hitherto had been considered as a very characteristic difference between families GH13 and GH70 [17, 74].

### Evolutionary relationships between family GH70 and GH13 members

The GH13 enzymes, degrading or modifying mainly starch-like substrates, constitute one of the largest Glycoside Hydrolase family at present in the CAZy database ([www.cazy.org](http://www.cazy.org)) and are found in a very wide range of organisms from all kingdoms [118, 119].  $\alpha$ -Amylase enzymes, catalyzing hydrolysis of internal ( $\alpha$ 1  $\rightarrow$  4) glycosidic bonds in starch and related substrates, are the main representatives of family GH13. Despite the fact that starch and sucrose acting GH13 and GH70 enzymes differ in their overall activities, there is no doubt that they share a common ancestor [74, 120]. Members of family GH13 and GH70 share a ( $\beta/\alpha$ )<sub>8</sub> barrel domain and employ a similar catalytic mechanism involving a covalent glucosyl intermediate and the retention of the  $\alpha$ -configuration in their products [121]. Structurally, the catalytic core of GH70 GSs, composed of A, B and C domains, was found to be similar and can be superimposed on the A, B and C domains present in  $\alpha$ -amylases [71]. Moreover, the GH13 and GH70 protein sequences exhibit 4–7 conserved sequence motifs that can be used as sequence fingerprints for the individual enzyme specificities [122].

GTFB-like 4,6- $\alpha$ -GTs constitute a first GH70 subfamily, unable to act on sucrose, but active on starch and maltodextrins, cleaving ( $\alpha$ 1  $\rightarrow$  4) linkages and synthesizing consecutive ( $\alpha$ 1  $\rightarrow$  6) linkages [18, 113, 114]. Considering the structural similarity of GTFB-like 4,6- $\alpha$ -GTs with GH70 GSs and their reaction similarity with family GH13 using ( $\alpha$ 1  $\rightarrow$  4) glucans as substrates, GTFB-like 4,6- $\alpha$ -GTs were proposed to be an evolutionary intermediate of family GH13  $\alpha$ -amylases and GH70 GSs [18, 114]. The characterization of GTFC of *Exiguobacterium sibiricum* 255-15 showed that it has a similar activity as GTFB-like 4,6- $\alpha$ -GTs, but, like GH13 family enzymes, lacks a permuted ( $\beta/\alpha$ )<sub>8</sub> barrel.

The BLASTp analysis of the *E. sibiricum* GTFC 4,6- $\alpha$ -GT revealed that the closest homologs of these GTFC type

of proteins are GTFB-like 4,6- $\alpha$ -GTs, whereas the next hits were (putative) GSs followed by (putative) GH13  $\alpha$ -amylases. The evolutionary relationships between all the biochemically characterized GSs as well as the (putative) GH70 4,6- $\alpha$ -GTs and (putative) GH13  $\alpha$ -amylases identified by this BLASTp search are depicted in the phylogenetic tree constructed based on the alignment of the complete sequences (Fig. 11). GH70 proteins are divided into three clearly separated clusters in the phylogenetic tree: GH70 GSs, GTFB-like and GTFC-like 4,6- $\alpha$ -GTs. Phylogenetically, GTFC-like 4,6- $\alpha$ -GTs are more closely related to GTFB-like 4,6- $\alpha$ -GTs, but are positioned between family GH70 enzymes and GH13  $\alpha$ -amylases, a fact that is also reflected in their non-permuted GH13 type of domain architecture. The position of the GTFB-like 4,6- $\alpha$ -GTs, in between GH70 GSs and GTFC-like 4,6- $\alpha$ -GTs, may be related to their high sequence identity to GH70 GSs. Most of the biochemically characterized family GH70 GSs cluster in clades in the phylogenetic tree, reflecting their bacterial hosts, but with a few exceptions (Fig. 11). Most GSs of *Leuconostoc* produce dextran types [mainly ( $\alpha 1 \rightarrow 6$ )] of polysaccharides and they form three separate clusters through the tree. The *Leuconostoc* alternansucrases and ( $\alpha 1 \rightarrow 2$ ) branching sucrases form separate clades from the main cluster of *Leuconostoc* GSs, reflecting their different product specificity. GSs from *Streptococcus* form two main clades, which produce soluble (dextran) and insoluble (mutan) polysaccharides, respectively. The characterized GSs of *Weissella* are closely related and produce dextran types of polysaccharides. Also *Lactobacillus* derived GSs cluster together; however, they produce different types of polysaccharides (reuteran, dextran and mutan). It is worth to note that within GH70 GSs, the ( $\alpha 1 \rightarrow 2$ ) branching GS is phylogenetically most close to GTFB-like 4,6- $\alpha$ -GTs. The reason for this is not clear.

It has been proposed that GH70 proteins evolved from a precursor GH13  $\alpha$ -amylase rather than vice versa, since the latter are widespread through different taxonomic groups and occur even in Archaea [18]. It appears more likely that amyolytic activity emerged first, and that later in time transglycosidase activity arose by optimization of the acceptor binding subsites that allowed the recognition and utilization of the products formed by the enzyme as acceptor substrates. The elucidation of the crystal structure of GTF180 GS prompted Vujičić et al., 2010 to propose that GSs emerged from an ancestor  $\alpha$ -amylase by an evolutionary pathway based on the “permutation per duplication model” [123]. Based on this pathway, the circularly permuted domain A, discontinuous domain B, and contiguous domain C characteristic for GSs are the result of the duplication, fusion and partial truncation of an ancestor GH13  $\alpha$ -amylase gene. However, it remained unclear whether these sequences of gene rearrangements

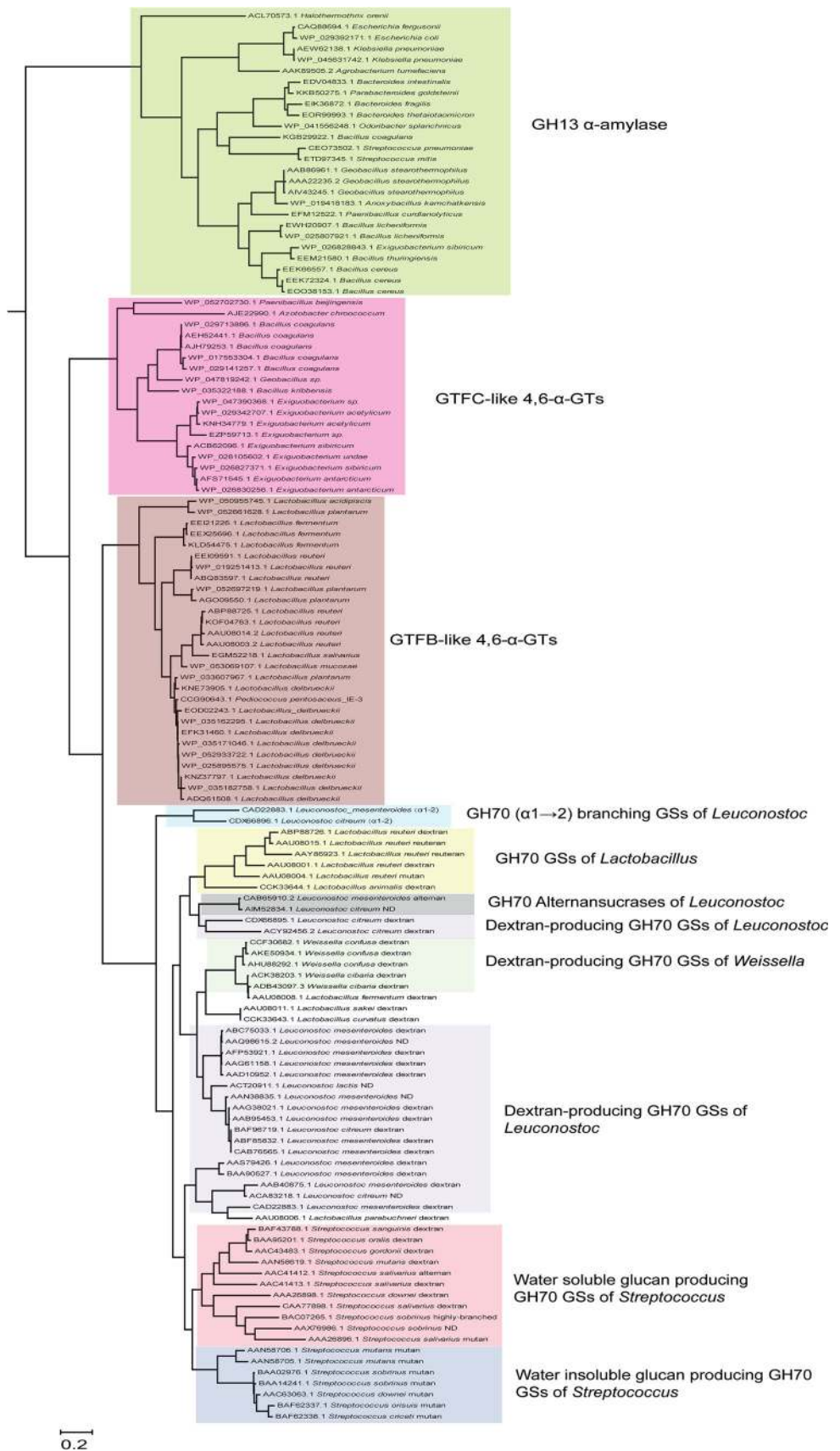
that resulted in the circular permutation occurred before or after the insertion of domain IV into the precursor gene. The existence of GTFC-like 4,6- $\alpha$ -GTs displaying an  $\alpha$ -amylase GH13-like fold but with domain IV inserted in domain B sheds light on the evolutionary history of family GH70 proteins (Fig. 12). First, insertion of domain IV may have led to the “intermediate” GTFC subfamily. Later, the unusual “U” fold domain organization of GS and GTFB-like 4,6- $\alpha$ -GTs evolved from a GTFC-like 4,6- $\alpha$ -GTs intermediate via the permutation-by-duplication mechanism, followed by insertion into the gene encoding domains V and N. In parallel, some but not all of these GTFC-like 4,6- $\alpha$ -GTs acquired Ig2-like domains at their C-terminus. While the above pathway describes the evolutionary changes in domain organization of GH13 and GH70 enzymes, it should be noted that the reaction specificity of the respective enzymes also has changed from one (sub)family to another. As of today (November, 2015), we do not know the structural details related to these changes, since no crystal structures have been reported yet for GTFB- or GTFC-like 4,6- $\alpha$ -GTs.

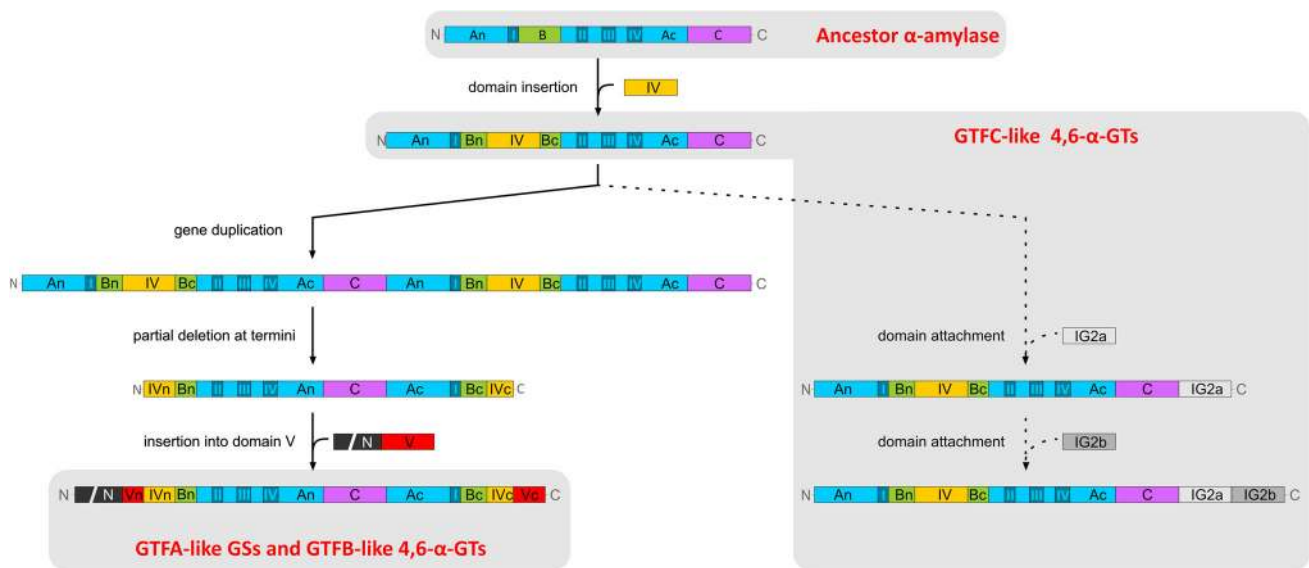
## Future perspectives

The number of putative GH70 GSs is increasing with the number of genome sequences available. At present, GSs are exclusively found in LAB, mainly from *Leuconostoc*, *Lactobacillus*, *Streptococcus*, and *Weissella* [15, 96]. However, only a relatively small number of these GSs have been biochemically characterized [15]. Only a single GS-encoding gene has been identified in the genus *Oenococcus* (*Oenococcus oeni* PSU-1); this enzyme has not been characterized yet and at present it remains unknown whether it has special properties compared to GSs found in the other LAB [124]. It also has remained unclear why GSs are only found in LAB. This may reflect a special feature of LAB, for instance their strong association with surfaces, present in biofilms. The question remains whether the increasing genome sequence data also will result in identification of GSs in non-LAB.

The number of GTFB-like 4,6- $\alpha$ -GTs is also increasing with the availability of genome sequence data. Very recently, the genome sequences of 213 *Lactobacillus* strains and associated genera have been sequenced and analyzed through comparative genomics [125]. This genome sequencing initiative has contributed to further increase the number of available GTFB-like 4,6- $\alpha$ -GTs protein sequences. As of November 2015, the GTFB-like 4,6- $\alpha$ -GT subfamily contains 46 proteins. Most of these GTFB-like 4,6- $\alpha$ -GTs are found in *Lactobacillus* (except a 3 putative GH70 enzymes identified in *Pediococcus* strains). Only 3 GTFB-like 4,6- $\alpha$ -GTs have been

**Fig. 11** Phylogenetic tree analysis of GH70 and GH13 proteins. The evolutionary tree is based on the alignment of the complete sequences of all biochemically characterized GH70 GSs, (putative) GH70 GTFB-like and GTFC-like 4,6- $\alpha$ -GTs and (putative) GH13  $\alpha$ -amylases identified by a BLAST search using the *E. sibiricum* GTFC protein as query sequence. The evolutionary history was inferred using the Maximum Likelihood method based on the JTT matrix-based model. The *bar* represents a genetic distance of 0.2 substitution per position. Each sequence is labeled with its Genbank accession number and bacterial origin. In the case of the biochemically characterized GH70 GSs, their  $\alpha$ -glucan polysaccharide products are also shown. Details of the aligned sequences are shown in Table S1





**Fig. 12** Proposed evolutionary pathway based on the “permutation per duplication model” leading to the unusual domain organization of GH70 GSs and GTFB-like 4,6- $\alpha$ -GTs. Insertion of domain IV into domain B of the ancestor  $\alpha$ -amylase led to the formation of the “intermediate” GTFC subfamily. Formation of the GTFC “intermediate” would be followed by the circular permutation (gene duplication and partial terminal deletions), and domain insertion

events leading to GH70 GSs and GTFB-like 4,6- $\alpha$ -GTs. Some of the GTFC-like 4,6- $\alpha$ -GTs members acquired Ig2-like domains (route indicated by dotted arrows). Sequence segments forming domains A, B, C, IV, and V are colored as before in blue, green, magenta, yellow, and red, respectively. Ig2-like domains are indicated in gray. This figure has been taken from [17]

biochemically characterized showing a similar substrate specificity and product linkage specificity [cleaving ( $\alpha 1 \rightarrow 4$ ) and synthesizing consecutive ( $\alpha 1 \rightarrow 6$ ) linkages] [18, 114]. GSs are capable of synthesizing all four possible types of  $\alpha$ -glycosidic linkages in  $\alpha$ -glucans [96]. The biochemical characterization of additional GTFB-like proteins may result in identification of novel enzymes, displaying different product linkage specificities. Such biochemical data will also contribute to a better understanding of the GTFB structure–function relationships. The same is valid for GTFC-like 4,6- $\alpha$ -GTs, which only have been found in different Gram-positive bacteria (i.e. *Exiguobacterium* and *Bacillus*). Only a single GTFC-like 4,6- $\alpha$ -GT has been biochemically characterized [17] and further cloning and biochemical characterization will provide more detailed insights into their substrate and product linkage specificity. GTFB-like 4,6- $\alpha$ -GTs and GTFC-like 4,6- $\alpha$ -GTs represent the evolutionary intermediates between family GH70 GSs and family GH13 enzymes. Further genome sequencing is likely to result in identification of novel putative GH13-GH70 intermediates representing additional GH70 subfamilies; their detailed biochemical characterization may provide detailed insights into the evolutionary events that have occurred.

Previous mutagenesis studies [65, 81, 89–91, 96, 109, 110, 126], elucidation of the crystal structures of GH70 GS enzymes synthesizing different glycosidic linkages [36, 71–73, 79] and recent site-directed mutagenesis studies

based on these 3D structures [92, 98, 99] contribute to a better understanding of GH70 GS structure–function relationships, especially regarding linkage specificity. These studies showed that the linkage specificity of GH70 GSs is determined by the interplay of residues from both domains A and B, forming acceptor substrate binding subsites. By now, only acceptor substrate binding subsites +1 and +2 of maltose have been studied in detail [71]. The availability of three-dimensional GS structures with a bound longer carbohydrate substrate, and/or a bound branched oligosaccharide, would greatly support unraveling additional acceptor substrate binding sites. With the presently available 3D structural information, combinatorial mutagenesis (targeting several residues) allows diversification of the glycosidic linkage composition of the  $\alpha$ -glucans produced. In a recent study of dextransucrase DSRS of *L. mesenteroides* NRRL B-512F, guided by the homologous GTF180- $\Delta$ N crystal structure, several residues forming acceptor substrate binding sites were targeted for combinatorial mutagenesis [98, 127]. Characterization of the mutants obtained revealed synthesis of a range of novel  $\alpha$ -glucans with an altered relative amount of ( $\alpha 1 \rightarrow 3$ ) linkages (3–20 %) [98, 127]. A directed evolution approach combined with a high-throughput screening facility for specific linkage types represents an alternative approach to study linkage specificity [128]. Such studies are likely capable of identifying hotspot regions that are located more distantly from the acceptor substrate binding sites, which



may allow fine tuning of the acceptor substrate binding sites [129]. Biochemical and structural characterization of mutant GS enzymes and their products synthesized will generate a detailed understanding of the structure–function relationships of GS enzymes. Such detailed insights may allow rational construction of GS mutants with acceptor substrate binding subsites providing desired physico-chemical microenvironments that produce tailor-made  $\alpha$ -glucans, i.e. with desired linkage type distributions, degree of branching, and molecular mass. Currently, the determinants of product size of GH70 GSs have remained largely unknown. An effective way to control the sizes of the  $\alpha$ -glucans synthesized by GSs is required for future industrial applications. Also the (thermo)stability of GSs needs to be improved; the availability of protein crystal structures facilitates a computational design approach to achieve this [130].

The donor substrate specificity of GH70 GSs is limited to sucrose, resulting in the transfer of only a glucosyl moiety. Various sucrose analogs have been synthesized, in which the glucosyl unit of sucrose is replaced by alternative glycosyl moieties [131–133]. The sucrose analogs  $\alpha$ -D-xylopyranosyl- $\beta$ -D-fructofuranoside (Xyl-Fru),  $\alpha$ -D-mannopyranosyl- $\beta$ -D-fructofuranoside (Man-Fru),  $\alpha$ -D-galactopyranosyl- $\beta$ -D-fructofuranoside (Gal-Fru) and  $\alpha$ -D-fucosylpyranosyl- $\beta$ -D-fructofuranoside (Fuc-Fru) have been synthesized by transferring the fructosyl unit of sucrose to monosaccharides acceptor substrates using fructansucrase enzymes [133–136]. The use of sucrose analogs as donor substrates by GS enzymes would allow transfer of a wider range of monosaccharides for oligosaccharide and glycoconjugate synthesis. GSs will need to be engineered to improve their ability to transfer such different glycosyl moieties. In a recent study, GTFA of *L. reuteri* 121 was reported to use  $\alpha$ -D-allopyranosyl- $\beta$ -D-fructofuranoside as donor substrate and to transfer an allose unit to several acceptor substrates [137]. The use of sucrose analogs as donor substrates provides a promising opportunity to extend the glycodiversity of GS products.

GTFB-like 4,6- $\alpha$ -GTs show the same domain organization and share a circularly permuted ( $\beta/\alpha$ )<sub>8</sub> barrel with GH70 GSs [18, 114]; however, they show clearly different substrate specificity and product linkage specificity [18, 114]. Identification of the determinants for these characteristics is an interesting topic for future research. Elucidation of the 3D structure of GTFB-like 4,6- $\alpha$ -GTs and mutational studies will provide a firm basis for such studies. GTFC-like 4,6- $\alpha$ -GTs differ from GTFB-like 4,6- $\alpha$ -GTs and GH70 GSs by a non-permuted ( $\beta/\alpha$ )<sub>8</sub> barrel. Elucidation of the 3D structure of GTFC-like 4,6- $\alpha$ -GTs and mutational studies may also serve to elucidate features that determine their substrate and product linkage

specificity. The crystal structures of GH70 GSs showed that these proteins follow a U-shape course to form different domains; the same is probably true for GTFB-like 4,6- $\alpha$ -GTs while the overall folding of GTFC-like 4,6- $\alpha$ -GTs may be similar to that of family GH13 enzymes and different from those of GSs and GTFB-like 4,6- $\alpha$ -GTs.

In recent years, this field has advanced by the characterization of novel  $\alpha$ -glucans containing different (ratios of) glycosidic linkages and various degrees of branching, produced from sucrose by newly derived GH70 GS enzymes and their mutant derivatives, but also by the identification and characterization of GH70 subfamily GTFB-like 4,6- $\alpha$ -GTs and GTFC-like 4,6- $\alpha$ -GTs, using maltodextrin/starch as substrates. The physico-chemical properties of these  $\alpha$ -glucans remain to be investigated to elucidate their structure–function relationships. Industrial application processes eventually may be developed based on the knowledge obtained in such studies.

In view of recent developments, characterization of new (sub)family GH70 enzymes, elucidation of their crystal structure, (rational) enzyme engineering and substrate engineering (sucrose analogs), the prospects of producing tailor-made  $\alpha$ -glucans are promising. This further enhances the applicability of GH70 enzymes as biocatalysts, and their products in food and medicine, and in the cosmetic industry. Moreover, their structural and functional relationships with family GH13 enzymes provide a very interesting showcase for the analysis of protein evolution processes in nature.

**Acknowledgments** This work was financially supported by the University of Groningen (to LD, SSVL, TP), the China Scholarship Council (to XM), the TKI Agri&Food program as coordinated by the Carbohydrate Competence Center (CCC-ABC; <http://www.cccresearch.nl>) (to YB, JG) and by EU project NOVOSIDES FP7-KBBE-2010-4-265854 (to LD).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Leroy F, De Vuyst L (2004) Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci Technol* 15(2):67–78. doi:10.1016/j.tifs.2003.09.004
2. Kumari A, Catanzaro R, Marotta F (2011) Clinical importance of lactic acid bacteria: a short review. *Acta Biomed* 82(3):177–180
3. Galle S, Arendt EK (2014) Exopolysaccharides from sourdough lactic acid bacteria. *Crit Rev Food Sci Nutr* 54(7):891–901. doi:10.1080/10408398.2011.617474

4. Jolly L, Vincent SJ, Duboc P, Neeser JR (2002) Exploiting exopolysaccharides from lactic acid bacteria. *Antonie Van Leeuwenhoek* 82(1–4):367–374
5. Ryan PM, Ross RP, Fitzgerald GF, Caplice NM, Stanton C (2015) Sugar-coated: exopolysaccharide producing lactic acid bacteria for food and human health applications. *Food Funct* 6(3):679–693. doi:10.1039/C4fo00529e
6. Welman AD, Maddox IS (2003) Exopolysaccharides from lactic acid bacteria: perspectives and challenges. *Trends Biotechnol* 21(6):269–274. doi:10.1016/S0167-7799(03)00107-0
7. Torino MI, Font de Valdez G, Mozzi F (2015) Biopolymers from lactic acid bacteria. Novel applications in foods and beverages. *Front Microbiol* 6:834. doi:10.3389/fmicb.2015.00834
8. Zannini E, Waters DM, Coffey A, Arendt EK (2015) Production, properties, and industrial food application of lactic acid bacteria-derived exopolysaccharides. *Appl Microbiol Biotechnol*. doi:10.1007/s00253-015-7172-2
9. Bowen WH, Koo H (2011) Biology of *Streptococcus mutans*-derived glucosyltransferases: role in extracellular matrix formation of cariogenic biofilms. *Caries Res* 45(1):69–86. doi:10.1159/000324598
10. Gibbons RJ (1968) Formation and significance of bacterial polysaccharides in caries etiology. *Caries Res* 2(2):164–171. doi:10.1159/000259554
11. Laws A, Gu Y, Marshall V (2001) Biosynthesis, characterisation, and design of bacterial exopolysaccharides from lactic acid bacteria. *Biotechnol Adv* 19(8):597–625. doi:10.1016/S0734-9750(01)00084-2
12. De Vuyst L, Degeest B (1999) Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiol Rev* 23(2):153–177
13. van Hijum SA, Kralj S, Ozimek LK, Dijkhuizen L, van Geel-Schutten IGH (2006) Structure–function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. *Microbiol Mol Biol Rev* 70(1):157–176. doi:10.1128/MMBR.70.1.157-176.2006
14. Korakli M, Vogel RF (2006) Structure/function relationship of homopolysaccharide producing glycansucrases and therapeutic potential of their synthesised glycans. *Appl Microbiol Biotechnol* 71(6):790–803. doi:10.1007/s00253-006-0469-4
15. Leemhuis H, Pijning T, Dobruchowska JM, van Leeuwen SS, Kralj S, Dijkstra BW, Dijkhuizen L (2013) Glucansucrases: three-dimensional structures, reactions, mechanism,  $\alpha$ -glucan analysis and their implications in biotechnology and food applications. *J Biotechnol* 163(2):250–272. doi:10.1016/j.jbiotec.2012.06.037
16. Monchois V, Willemot R-M, Monsan P (1999) Glucansucrases: mechanism of action and structure–function relationships. *FEMS Microbiol Rev* 23(2):131–151. doi:10.1111/j.1574-6976.1999.tb00394.x
17. Gangoit J, Pijning T, Dijkhuizen L (2015) Biochemical characterization of the *Exiguobacterium sibiricum* 255-15 GtfC enzyme representing a novel glycoside hydrolase 70 subfamily of 4,6- $\alpha$ -glucanotransferase enzymes. *Appl Environ Microbiol*. doi:10.1128/AEM.03420-15
18. Kralj S, Grijpstra P, van Leeuwen SS, Leemhuis H, Dobruchowska JM, van der Kaaij RM, Malik A, Oetari A, Kamerling JP, Dijkhuizen L (2011) 4,6- $\alpha$ -glucanotransferase, a novel enzyme that structurally and functionally provides an evolutionary link between glycoside hydrolase enzyme families 13 and 70. *Appl Environ Microbiol* 77(22):8154–8163. doi:10.1128/AEM.05735-11
19. Monsan P, Bozonnet S, Albenne C, Joucla G, Willemot R-M, Remaud-Siméon M (2001) Homopolysaccharides from lactic acid bacteria. *Int Dairy J* 11(9):675–685. doi:10.1016/S0958-6946(01)00113-3
20. Naessens M, Cerdobbel A, Soetaert W, Vandamme EJ (2005) *Leuconostoc* dextransucrase and dextran: production, properties and applications. *J Chem Technol Biotechnol* 80(8):845–860. doi:10.1002/Jctb.1322
21. Hehre EJ (1941) Production from sucrose of a sero-logically reactive polysaccharide by a sterile bacterial extract. *Science* 93(2410):237–238. doi:10.1126/science.93.2410.237
22. Hehre EJ, Sugg JY (1942) Serologically reactive polysaccharides produced through the action of bacterial enzymes: Dextran of *Leuconostoc mesenteroides* from sucrose. *J Exp Med* 75(3):339–353
23. Bounaix MS, Gabriel V, Morel S, Robert H, Rabier P, Remaud-Siméon M, Gabriel B, Fontagné-Faucher C (2009) Biodiversity of exopolysaccharides produced from sucrose by sourdough lactic acid bacteria. *J Agric Food Chem* 57(22):10889–10897. doi:10.1021/jf902068t
24. López-Munguía A, Pelenc V, Remaud-Siméon M, Biton J, Michel JM, Lang C, Paul F, Monsan P (1993) Production and purification of alternansucrase, a glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355 for the synthesis of oligoalternans. *Enzyme Microb Technol* 15(1):77–85. doi:10.1016/0141-0229(93)90120-Q
25. Van Geel-Schutten IGH, Faber EJ, Smit E, Bonting K, Smith MR, Ten Brink B, Kamerling JP, Vliegthart JFG, Dijkhuizen L (1999) Biochemical and structural characterization of the glucan and fructan exopolysaccharides synthesized by the *Lactobacillus reuteri* wild-type strain and by mutant strains. *Appl Environ Microbiol* 65(7):3008–3014
26. Hanada N, Kuramitsu HK (1989) Isolation and characterization of the *Streptococcus mutans* *gtfD* gene, coding for primer-dependent soluble glucan synthesis. *Infect Immun* 57(7):2079–2085
27. Ooshima T, Matsumura M, Hoshino T, Kawabata S, Sobue S, Fujiwara T (2001) Contributions of three glucosyltransferases to sucrose-dependent adherence of *Streptococcus mutans*. *J Dent Res* 80(7):1672–1677. doi:10.1177/00220345010800071401
28. Passerini D, Vuillemin M, Ufarte L, Morel S, Loux V, Fontagné-Faucher C, Monsan P, Remaud-Siméon M, Moulis C (2015) Inventory of the GH70 enzymes encoded by *Leuconostoc citreum* NRRL B-1299 - identification of three novel  $\alpha$ -transglucosylases. *FEBS J* 282(11):2115–2130. doi:10.1111/febs.13261
29. Argüello-Morales MA, Remaud-Siméon M, Pizzut S, Sarcabal P, Willemot R-M, Monsan P (2000) Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from *Leuconostoc mesenteroides* NRRL B-1355. *FEMS Microbiol Lett* 182(1):81–85. doi:10.1111/j.1574-6968.2000.tb08878.x
30. Walker GJ, Cheetham NWH, Taylor C, Pearce BJ, Slodki ME (1990) Productivity of 4  $\alpha$ -D-glucosyltransferases released by *Streptococcus sobrinus* under defined conditions in continuous culture. *Carbohydr Polym* 13(4):399–421. doi:10.1016/0144-8617(90)90039-U
31. Kim D, Robyt JF (1994) Production and selection of mutants of *Leuconostoc mesenteroides* constitutive for glucansucrases. *Enzyme Microb Technol* 16(8):659–664. doi:10.1016/0141-0229(94)90086-8
32. Kim D, Robyt JF (1995) Dextransucrase constitutive mutants of *Leuconostoc mesenteroides* B-1299. *Enzyme Microb Technol* 17(12):1050–1056. doi:10.1016/0141-0229(95)00039-9
33. Kim D, Robyt JF (1995) Production, selection, and characteristics of mutants of *Leuconostoc mesenteroides* B-742 constitutive for dextransucrases. *Enzyme Microb Technol* 17(8):689–695. doi:10.1016/0141-0229(94)00021-1
34. Badel S, Bernardi T, Michaud P (2011) New perspectives for Lactobacilli exopolysaccharides. *Biotechnol Adv* 29(1):54–66. doi:10.1016/j.biotechadv.2010.08.011

35. Bozonnet S, Dols-Laffargue M, Fabre E, Pizzut S, Remaud-Siméon M, Monsan P, Willemot R-M (2002) Molecular characterization of DSR-E, an  $\alpha$ -1,2 linkage-synthesizing dextranucrase with two catalytic domains. *J Bacteriol* 184(20):5753–5761. doi:10.1128/JB.184.20.5753-5761.2002
36. Brison Y, Pijning T, Malbert Y, Fabre E, Mourey L, Morel S, Potocki-Véronèse G, Monsan P, Tranier S, Remaud-Siméon M, Dijkstra BW (2012) Functional and structural characterization of  $\alpha$ -(1  $\rightarrow$  2) branching sucrose derived from DSR-E glucansucrase. *J Biol Chem* 287(11):7915–7924. doi:10.1074/jbc.M111.305078
37. Fabre E, Bozonnet S, Arcache A, Willemot R-M, Vignon M, Monsan P, Remaud-Siméon M (2005) Role of the two catalytic domains of DSR-E dextranucrase and their involvement in the formation of highly  $\alpha$ -1,2 branched dextran. *J Bacteriol* 187(1):296–303. doi:10.1128/JB.187.1.296-303.2005
38. Monchois V, Willemot R-M, Remaud-Siméon M, Croux C, Monsan P (1996) Cloning and sequencing of a gene coding for a novel dextranucrase from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only  $\alpha$  (1  $\rightarrow$  6) and  $\alpha$  (1  $\rightarrow$  3) linkages. *Gene* 182(1–2):23–32. doi:10.1016/S0378-1119(96)00443-X
39. Monchois V, Remaud-Siméon M, Monsan P, Willemot R-M (1998) Cloning and sequencing of a gene coding for an extracellular dextranucrase (DSRB) from *Leuconostoc mesenteroides* NRRL B-1299 synthesizing only a  $\alpha$ (1  $\rightarrow$  6) glucan. *FEMS Microbiol Lett* 159(2):307–315. doi:10.1111/j.1574-6968.1998.tb12876.x
40. Côté GL, Robyt JF (1982) Isolation and partial characterization of an extracellular glucansucrase from *Leuconostoc mesenteroides* NRRL B-1355 that synthesizes an alternating (1  $\rightarrow$  6)/(1  $\rightarrow$  3)- $\alpha$ -D-glucan. *Carbohydr Res* 101(1):57–74. doi:10.1016/S0008-6215(00)80795-8
41. Vidal RF, Martinez A, Moulis C, Escalier P, Morel S, Remaud-Siméon M, Monsan P (2011) A novel dextranucrase is produced by *Leuconostoc citreum* strain B/110-1-2: an isolate used for the industrial production of dextran and dextran derivatives. *J Ind Microbiol Biotechnol* 38(9):1499–1506. doi:10.1007/s10295-010-0936-x
42. Fraga Vidal R, Moulis C, Escalier P, Remaud-Siméon M, Monsan P (2011) Isolation of a gene from *Leuconostoc citreum* B/110-1-2 encoding a novel dextranucrase enzyme. *Curr Microbiol* 62(4):1260–1266. doi:10.1007/s00284-010-9851-7
43. Côté GL, Skory CD (2012) Cloning, expression, and characterization of an insoluble glucan-producing glucansucrase from *Leuconostoc mesenteroides* NRRL B-1118. *Appl Microbiol Biotechnol* 93(6):2387–2394. doi:10.1007/s00253-011-3562-2
44. Amari M, Valérie G, Robert H, Morel S, Moulis C, Gabriel B, Remaud-Siméon M, Fontagné-Faucher C (2015) Overview of the glucansucrase equipment of *Leuconostoc citreum* LBAE-E16 and LBAE-C11, two strains isolated from sourdough. *FEMS Microbiol Lett* 362(1):1–8. doi:10.1093/femsle/fnu024
45. Hamada S, Slade HD (1980) Biology, immunology, and cariogenicity of *Streptococcus mutans*. *Microbiol Rev* 44(2):331–384
46. Mayhall CW, Butler WT (1976) The carbohydrate composition of experimental salivary pellicles. *J Oral Pathol* 5(6):358–370. doi:10.1111/j.1600-0714.1976.tb01787.x
47. Emilson CG, Nilsson B, Bowen WH (1984) Carbohydrate composition of dental plaque from primates with irradiation caries. *J Oral Pathol* 13(3):213–220. doi:10.1111/j.1600-0714.1984.tb01419.x
48. Hata S, Mayanagi H (2003) Acid diffusion through extracellular polysaccharides produced by various mutants of *Streptococcus mutans*. *Arch Oral Biol* 48(6):431–438. doi:10.1016/S0003-9969(03)0032-3
49. Shiroza T, Ueda S, Kuramitsu HK (1987) Sequence analysis of the *gtfB* gene from *Streptococcus mutans*. *J Bacteriol* 169(9):4263–4270
50. Ueda S, Shiroza T, Kuramitsu HK (1988) Sequence analysis of the *gtfC* gene from *Streptococcus mutans* GS-5. *Gene* 69(1):101–109. doi:10.1016/0378-1119(88)90382-4
51. Yamashita Y, Bowen WH, Burne RA, Kuramitsu HK (1993) Role of the *Streptococcus mutans* *gtf* genes in caries induction in the specific-pathogen-free rat model. *Infect Immun* 61(9):3811–3817
52. Tsumori H, Kuramitsu H (1997) The role of the *Streptococcus mutans* glucosyltransferases in the sucrose-dependent attachment to smooth surfaces: essential role of the GtfC enzyme. *Oral Microbiol Immunol* 12(5):274–280. doi:10.1111/j.1399-302X.1997.tb00391.x
53. Gilmore KS, Russell RRB, Ferretti JJ (1990) Analysis of the *Streptococcus downei* *gtfS* gene, which specifies a glucosyltransferase that synthesizes soluble glucans. *Infect Immun* 58(8):2452–2458
54. Valeur N, Engel P, Carbajal N, Connolly E, Ladefoged K (2004) Colonization and immunomodulation by *Lactobacillus reuteri* ATCC 55730 in the human gastrointestinal tract. *Appl Environ Microbiol* 70(2):1176–1181. doi:10.1128/AEM.70.2.1176-1181.2004
55. van Geel-Schutten IGH, Flesch F, ten Brink B, Smith MR, Dijkhuizen L (1998) Screening and characterization of *Lactobacillus* strains producing large amounts of exopolysaccharides. *Appl Microbiol Biotechnol* 50(6):697–703. doi:10.1007/s002530051353
56. Kralj S, van Geel-Schutten IGH, van der Maarel MJEC, Dijkhuizen L (2004) Biochemical and molecular characterization of *Lactobacillus reuteri* 121 reuteransucrase. *Microbiology* 150(Pt 7):2099–2112. doi:10.1099/mic.0.27105-0
57. Kralj S, van Geel-Schutten IGH, Dondorff MMG, Kirsanovs S, van der Maarel MJEC, Dijkhuizen L (2004) Glucan synthesis in the genus *Lactobacillus*: isolation and characterization of glucansucrase genes, enzymes and glucan products from six different strains. *Microbiology* 150(Pt 11):3681–3690. doi:10.1099/mic.0.27321-0
58. Kralj S, van Geel-Schutten IGH, Rahaoui H, Leer RJ, Faber EJ, van der Maarel MJEC, Dijkhuizen L (2002) Molecular characterization of a novel glucosyltransferase from *Lactobacillus reuteri* strain 121 synthesizing a unique, highly branched glucan with  $\alpha$ -(1  $\rightarrow$  4) and  $\alpha$ -(1  $\rightarrow$  6) glucosidic bonds. *Appl Environ Microbiol* 68(9):4283–4291. doi:10.1128/AEM.68.9.4283-4291.2002
59. van Leeuwen SS, Kralj S, van Geel-Schutten IGH, Gerwig GJ, Dijkhuizen L, Kamerling JP (2008) Structural analysis of the  $\alpha$ -D-glucan (EPS35-5) produced by the *Lactobacillus reuteri* strain 35-5 glucansucrase GTFA enzyme. *Carbohydr Res* 343(7):1251–1265. doi:10.1016/j.carres.2008.01.044
60. Kralj S, Stripling E, Sanders P, van Geel-Schutten IGH, Dijkhuizen L (2005) Highly hydrolytic reuteransucrase from probiotic *Lactobacillus reuteri* strain ATCC 55730. *Appl Environ Microbiol* 71(7):3942–3950. doi:10.1128/AEM.71.7.3942-3950.2005
61. Bounaix MS, Robert H, Gabriel V, Morel S, Remaud-Siméon M, Gabriel B, Fontagné-Faucher C (2010) Characterization of dextran-producing *Weissella* strains isolated from sourdoughs and evidence of constitutive dextranucrase expression. *FEMS Microbiol Lett* 311(1):18–26. doi:10.1111/j.1574-6968.2010.02067.x
62. Amari M, Arango LF, Gabriel V, Robert H, Morel S, Moulis C, Gabriel B, Remaud-Siméon M, Fontagné-Faucher C (2013) Characterization of a novel dextranucrase from *Weissella confusa* isolated from sourdough. *Appl Microbiol Biotechnol* 97(12):5413–5422. doi:10.1007/s00253-012-4447-8
63. Kang HK, Oh JS, Kim D (2009) Molecular characterization and expression analysis of the glucansucrase DSRWC from *Weissella cibaria* synthesizing a  $\alpha$ (1  $\rightarrow$  6) glucan. *FEMS Microbiol Lett* 292(1):33–41. doi:10.1111/j.1574-6968.2008.01460.x



64. Dimopoulou M, Vuillemin M, Campbell-Sills H, Lucas PM, Ballestra P, Miot-Sertier C, Favier M, Coulon J, Moine V, Doco T, Roques M, Williams P, Petrel M, Gontier E, Moulis C, Remaud-Siméon M, Dols-Lafargue M (2014) Exopolysaccharide (EPS) synthesis by *Oenococcus oeni*: from genes to phenotypes. *PLoS One* 9(6):e98898. doi:10.1371/journal.pone.0098898
65. Moulis C, Joucla G, Harrison D, Fabre E, Potocki-Véronèse G, Monsan P, Remaud-Siméon M (2006) Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *J Biol Chem* 281(42):31254–31267. doi:10.1074/jbc.M604850200
66. Monchois V, Reverte A, Remaud-Siméon M, Monsan P, Willemot R-M (1998) Effect of *Leuconostoc mesenteroides* NRRL B-512F dextranase carboxy-terminal deletions on dextran and oligosaccharide synthesis. *Appl Environ Microbiol* 64(5):1644–1649
67. Janecek S, Svensson B, Russell RRB (2000) Location of repeat elements in glucansucrases of *Leuconostoc* and *Streptococcus* species. *FEMS Microbiol Lett* 192(1):53–57. doi:10.1111/j.1574-6968.2000.tb09358.x
68. Funane K, Mizuno K, Takahara H, Kobayashi M (2000) Gene encoding a dextranase-like protein in *Leuconostoc mesenteroides* NRRL B-512F. *Biosci Biotechnol Biochem* 64(1):29–38. doi:10.1271/bbb.64.29
69. Giffard PM, Jacques NA (1994) Definition of a fundamental repeating unit in *Streptococcal* glucosyltransferase glucan-binding regions and related sequences. *J Dent Res* 73(6):1133–1141. doi:10.1177/00220345940730060201
70. Kingston KB, Allen DM, Jacques NA (2002) Role of the C-terminal YG repeats of the primer-dependent streptococcal glucosyltransferase, GtfJ, in binding to dextran and mutan. *Microbiology* 148(Pt 2):549–558. doi:10.1099/00221287-148-2-549
71. Vujičić-Žagar A, Pijning T, Kralj S, López CA, Eeuwema W, Dijkhuizen L, Dijkstra BW (2010) Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. *Proc Natl Acad Sci USA* 107(50):21406–21411. doi:10.1073/pnas.1007531107
72. Ito K, Ito S, Shimamura T, Weyand S, Kawarasaki Y, Misaka T, Abe K, Kobayashi T, Cameron AD, Iwata S (2011) Crystal structure of glucansucrase from the dental caries pathogen *Streptococcus mutans*. *J Mol Biol* 408(2):177–186. doi:10.1016/j.jmb.2011.02.028
73. Pijning T, Vujičić-Žagar A, Kralj S, Dijkhuizen L, Dijkstra BW (2012) Structure of the  $\alpha$ -1,6/ $\alpha$ -1,4-specific glucansucrase GTFA from *Lactobacillus reuteri* 121. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68(12):1448–1454. doi:10.1107/S1744309112044168
74. MacGregor EA, Jespersen HM, Svensson B (1996) A circularly permuted  $\alpha$ -amylase-type  $\alpha/\beta$ -barrel structure in glucan-synthesizing glucosyltransferases. *FEBS Lett* 378(3):263–266. doi:10.1016/0014-5793(95)01428-4
75. Svensson B (1994) Protein engineering in the  $\alpha$ -amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol Biol* 25(2):141–157. doi:10.1007/BF00023233
76. Monchois V, Vignon M, Escalier PC, Svensson B, Russell RRB (2000) Involvement of Gln937 of *Streptococcus downei* GTF-I glucansucrase in transition-state stabilization. *Eur J Biochem* 267(13):4127–4136. doi:10.1046/j.1432-1327.2000.01448.x
77. Uitdehaag JCM, Mosi R, Kalk KH, van der Veen BA, Dijkhuizen L, Withers SG, Dijkstra BW (1999) X-ray structures along the reaction pathway of cyclodextrin glucosyltransferase elucidate catalysis in the  $\alpha$ -amylase family. *Nat Struct Biol* 6(5):432–436. doi:10.1038/8235
78. Lis M, Shiroza T, Kuramitsu HK (1995) Role of the C-terminal direct repeating units of the *Streptococcus mutans* glucosyltransferase-S in glucan binding. *Appl Environ Microbiol* 61(5):2040–2042
79. Pijning T, Vujičić-Žagar A, Kralj S, Dijkhuizen L, Dijkstra BW (2014) Flexibility of truncated and full-length glucansucrase GTF180 enzymes from *Lactobacillus reuteri* 180. *FEBS J* 281(9):2159–2171. doi:10.1111/febs.12769
80. Meng X, Dobruchowska JM, Pijning T, Gerwig GJ, Kamerling JP, Dijkhuizen L (2015) Truncation of domain V of the multidomain glucansucrase GTF180 of *Lactobacillus reuteri* 180 heavily impairs its polysaccharide-synthesizing ability. *Appl Microbiol Biotechnol* 99(14):5885–5894. doi:10.1007/s00253-014-6361-8
81. van Leeuwen SS, Kralj S, Eeuwema W, Gerwig GJ, Dijkhuizen L, Kamerling JP (2009) Structural characterization of bioengineered  $\alpha$ -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180. *Biomacromolecules* 10(3):580–588. doi:10.1021/bm801240r
82. van Leeuwen SS, Kralj S, van Geel-Schutten IGH, Gerwig GJ, Dijkhuizen L, Kamerling JP (2008) Structural analysis of the  $\alpha$ -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme. *Carbohydr Res* 343(7):1237–1250. doi:10.1016/j.carres.2008.01.042
83. van Leeuwen SS, Kralj S, Gerwig GJ, Dijkhuizen L, Kamerling JP (2008) Structural analysis of bioengineered  $\alpha$ -D-glucan produced by a triple mutant of the Glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180: generation of ( $\alpha$ 1  $\rightarrow$  4) linkages in a native (1  $\rightarrow$  3)/(1  $\rightarrow$  6)- $\alpha$ -D-glucan. *Biomacromolecules* 9(8):2251–2258. doi:10.1021/bm800410w
84. Dobruchowska JM, Meng X, Leemhuis H, Gerwig GJ, Dijkhuizen L, Kamerling JP (2013) Gluco-oligomers initially formed by the reuteransucrase enzyme of *Lactobacillus reuteri* 121 incubated with sucrose and malto-oligosaccharides. *Glycobiology* 23(9):1084–1096. doi:10.1093/glycob/cwt048
85. Cheetham NWH, Slodki ME, Walker GJ (1991) Structure of the linear, low-molecular-weight dextran synthesized by a D-glucosyltransferase (Gtf-S3) of *Streptococcus sobrinus*. *Carbohydr Polym* 16(4):341–353. doi:10.1016/0144-8617(91)90053-F
86. Ebert KH, Schenk G (1968) Mechanisms of biopolymer growth: the formation of dextran and levan. *Adv Enzymol Relat Areas Mol Biol* 30:179–221. doi:10.1002/9780470122754.ch4
87. Meng X, Dobruchowska JM, Gerwig GJ, Kamerling JP, Dijkhuizen L (2015) Synthesis of oligo- and polysaccharides by *Lactobacillus reuteri* 121 reuteransucrase at high concentrations of sucrose. *Carbohydr Res* 414:85–92. doi:10.1016/j.carres.2015.07.011
88. Shah DSH, Joucla G, Remaud-Siméon M, Russell RRB (2004) Conserved repeat motifs and glucan binding by glucansucrases of oral Streptococci and *Leuconostoc mesenteroides*. *J Bacteriol* 186(24):8301–8308. doi:10.1128/JB.186.24.8301-8308.2004
89. Kralj S, Eeuwema W, Eckhardt TH, Dijkhuizen L (2006) Role of asparagine 1134 in glucosidic bond and transglycosylation specificity of reuteransucrase from *Lactobacillus reuteri* 121. *FEBS J* 273(16):3735–3742. doi:10.1111/j.1742-4658.2006.05376.x
90. Kralj S, van Geel-Schutten IGH, Faber EJ, van der Maarel MJE, Dijkhuizen L (2005) Rational transformation of *Lactobacillus reuteri* 121 reuteransucrase into a dextranase. *Biochemistry* 44(25):9206–9216. doi:10.1021/bi050447q
91. Hellmuth H, Wittrock S, Kralj S, Dijkhuizen L, Hofer B, Seibel J (2008) Engineering the glucansucrase GTFR enzyme reaction and glycosidic bond specificity: toward tailor-made polymer and oligosaccharide products. *Biochemistry* 47(25):6678–6684. doi:10.1021/bi800563r
92. Meng X, Dobruchowska JM, Pijning T, López CA, Kamerling JP, Dijkhuizen L (2014) Residue Leu<sup>940</sup> has a crucial role in the linkage and reaction specificity of the glucansucrase GTF180 of the probiotic bacterium *Lactobacillus reuteri* 180. *J Biol Chem* 289(47):32773–32782. doi:10.1074/jbc.M114.602524



93. Kim D, Robyt JF, Lee S-Y, Lee J-H, Kim Y-M (2003) Dextran molecular size and degree of branching as a function of sucrose concentration, pH, and temperature of reaction of *Leuconostoc mesenteroides* B-512FMC dextranase. *Carbohydr Res* 338(11):1183–1189. doi:10.1016/S0008-6215(03)00148-4
94. Falconer DJ, Mukerjea R, Robyt JF (2011) Biosynthesis of dextrans with different molecular weights by selecting the concentration of *Leuconostoc mesenteroides* B-512FMC dextranase, the sucrose concentration, and the temperature. *Carbohydr Res* 346(2):280–284. doi:10.1016/j.carres.2010.10.024
95. Robyt JF, Taniguchi H (1976) The mechanism of dextranase action. Biosynthesis of branch linkages by acceptor reactions with dextran. *Arch Biochem Biophys* 174(1):129–135. doi:10.1016/0003-9861(76)90331-3
96. Leemhuis H, Pijning T, Dobruchowska JM, Dijkstra BW, Dijkhuizen L (2012) Glycosidic bond specificity of glucansucrases: on the role of acceptor substrate binding residues. *Biocatal Biotransform* 30(3):366–376. doi:10.3109/10242422.2012.676301
97. Meng X, Dobruchowska JM, Pijning T, Gerwig GJ, Dijkhuizen L (2016) Synthesis of new hyperbranched  $\alpha$ -glucans from sucrose by *Lactobacillus reuteri* 180 glucansucrase mutants. *J Agric Food Chem*. doi:10.1021/acs.jafc.5b05161
98. Irague R, Tarquis L, André I, Moulis C, Morel S, Monsan P, Potocki-Véronèse G, Remaud-Siméon M (2013) Combinatorial engineering of dextranase specificity. *PLoS One* 8(10):e77837. doi:10.1371/journal.pone.0077837
99. Meng X, Pijning T, Dobruchowska JM, Gerwig GJ, Dijkhuizen L (2015) Characterization of the functional roles of amino acid residues in acceptor binding subsite +1 in the active site of the glucansucrase GTF180 enzyme of *Lactobacillus reuteri* 180. *J Biol Chem*. doi:10.1074/jbc.M115.687558
100. Vuillemin M, Clavierie M, Brison Y, Séverac E, Bondy P, Morel S, Monsan P, Moulis C, Remaud-Siméon M (2016) Characterization of the first  $\alpha$ -(1  $\rightarrow$  3) branching sucrases of GH70 family. *J Biol Chem*. doi:10.1074/jbc.M115.688044
101. Galle S, Schwab C, Dal Bello F, Coffey A, Gänzle MG, Arendt EK (2012) Influence of in situ synthesized exopolysaccharides on the quality of gluten-free sorghum sourdough bread. *Int J Food Microbiol* 155(3):105–112. doi:10.1016/j.ijfoodmicro.2012.01.009
102. Spence JH, Higgins AR, Kimmel JR (1952) The clinical effects of dextran as a plasma volume extender. *US Armed Forces Med J* 3(4):515–518
103. Kim D, Day D-F (1994) A new process for the production of clinical dextran by mixed-culture fermentation of *Lipomyces starkeyi* and *Leuconostoc mesenteroides*. *Enzyme Microb Technol* 16(10):844–848. doi:10.1016/0141-0229(94)90058-2
104. Finkinstadt VL, Côté GL, Willett JL (2011) Corrosion protection of low-carbon steel using exopolysaccharide coatings from *Leuconostoc mesenteroides*. *Biotechnol Lett* 33(6):1093–1100. doi:10.1007/s10529-011-0539-2
105. Scheerder J, Breur R, Slaghek T, Holtman W, Vennik M, Ferrari G (2012) Exopolysaccharides (EPS) as anti-corrosive additives for coatings. *Prog Org Coat* 75(3):224–230. doi:10.1016/j.porgcoat.2012.05.003
106. Olano-Martin E, Mountzouris KC, Gibson GR, Rastall RA (2000) *In vitro* fermentability of dextran, oligodextran and maltodextrin by human gut bacteria. *Br J Nutr* 83(3):247–255. doi:10.1017/S0007114500000325
107. Sarbini SR, Kolida S, Naeye T, Einerhand A, Brison Y, Remaud-Siméon M, Monsan P, Gibson GR, Rastall RA (2011) *In vitro* fermentation of linear and  $\alpha$ -1,2-branched dextrans by the human fecal microbiota. *Appl Environ Microbiol* 77(15):5307–5315. doi:10.1128/AEM.02568-10
108. Monsan P, Paul F, Auriol D (1995) New developments in the application of enzymes to synthesis reactions. Peptides and oligosaccharides. *Ann NY Acad Sci* 750:357–363. doi:10.1111/j.1749-6632.1995.tb19980.x
109. Monchois V, Vignon M, Russell RRB (2000) Mutagenesis of Asp-569 of glucosyltransferase I glucansucrase modulates glucan and oligosaccharide synthesis. *Appl Environ Microbiol* 66(5):1923–1927. doi:10.1128/AEM.66.5.1923-1927.2000
110. Côté GL, Skory CD (2014) Effects of mutations at threonine-654 on the insoluble glucan synthesized by *Leuconostoc mesenteroides* NRRL B-1118 glucansucrase. *Appl Microbiol Biotechnol* 98(15):6651–6658. doi:10.1007/s00253-014-5622-x
111. Tsumori H, Minami T, Kuramitsu HK (1997) Identification of essential amino acids in the *Streptococcus mutans* glucosyltransferases. *J Bacteriol* 179(11):3391–3396
112. Binder TP, Côté GL, Robyt JF (1983) Disproportionation reactions catalyzed by *Leuconostoc* and *Streptococcus* glucansucrases. *Carbohydr Res* 124(2):275–286. doi:10.1016/0008-6215(83)88463-8
113. Dobruchowska JM, Gerwig GJ, Kralj S, Grijpstra P, Leemhuis H, Dijkhuizen L, Kamerling JP (2012) Structural characterization of linear isomalto-/malto-oligomer products synthesized by the novel GTFB 4,6- $\alpha$ -glucanotransferase enzyme from *Lactobacillus reuteri* 121. *Glycobiology* 22(4):517–528. doi:10.1093/glycob/cwr167
114. Leemhuis H, Dijkman WP, Dobruchowska JM, Pijning T, Grijpstra P, Kralj S, Kamerling JP, Dijkhuizen L (2013) 4,6- $\alpha$ -Glucanotransferase activity occurs more widespread in *Lactobacillus* strains and constitutes a separate GH70 subfamily. *Appl Microbiol Biotechnol* 97(1):181–193. doi:10.1007/s00253-012-3943-1
115. Leemhuis H, Dobruchowska JM, Ebbelaar M, Faber F, Buwalda PL, van der Maarel MJ, Kamerling JP, Dijkhuizen L (2014) Isomalto/malto-polysaccharide, a novel soluble dietary fiber made via enzymatic conversion of starch. *J Agric Food Chem* 62(49):12034–12044. doi:10.1021/jf503970a
116. Rodrigues DF, Goris J, Vishnivetskaya T, Gilichinsky D, Thomashow MF, Tiedje JM (2006) Characterization of *Exiguobacterium* isolates from the siberian permafrost. Description of *Exiguobacterium sibiricum* sp. nov. *Extremophiles* 10(4):285–294. doi:10.1007/s00792-005-0497-5
117. Rodrigues DF, Ivanova N, He Z, Huebner M, Zhou J, Tiedje JM (2008) Architecture of thermal adaptation in an *Exiguobacterium sibiricum* strain isolated from 3 million year old permafrost: a genome and transcriptome approach. *BMC Genom* 9:547. doi:10.1186/1471-2164-9-547
118. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42 (Database issue):D490–D495. doi:10.1093/nar/gkt1178
119. van der Maarel MJEC, van der Veen B, Uitdehaag JCM, Leemhuis H, Dijkhuizen L (2002) Properties and applications of starch-converting enzymes of the  $\alpha$ -amylase family. *J Biotechnol* 94(2):137–155. doi:10.1016/S0168-1656(01)00407-2
120. Janecek S, Svensson B, MacGregor EA (2014)  $\alpha$ -Amylase: an enzyme specificity found in various families of glycoside hydrolases. *Cell Mol Life Sci* 71(7):1149–1170. doi:10.1007/s00018-013-1388-z
121. MacGregor EA, Janecek S, Svensson B (2001) Relationship of sequence and structure to specificity in the  $\alpha$ -amylase family of enzymes. *Biochim Biophys Acta* 1546(1):1–20. doi:10.1016/S0167-4838(00)00302-2
122. Janecek S (2002) How many conserved sequence regions are there in the  $\alpha$ -amylase family? *Biologia* 57:29–41
123. Peisajovich SG, Rockah L, Tawfik DS (2006) Evolution of new protein topologies through multistep gene rearrangements. *Nat Genet* 38(2):168–174. doi:10.1038/ng1717

124. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103(42):15611–15616. doi:[10.1073/pnas.0607117103](https://doi.org/10.1073/pnas.0607117103)
125. Sun ZH, Harris HMB, McCann A, Guo CY, Argimon S, Zhang WY, Yang XW, Jeffery IB, Cooney JC, Kagawa TF, Liu WJ, Song YQ, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Ross RP, Yang RF, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui YJ, Zhang HP, O'Toole PW (2015) Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nature communications*. doi:[10.1038/Ncomms9322](https://doi.org/10.1038/Ncomms9322)
126. Funane K, Ishii T, Ono H, Kobayashi M (2005) Changes in linkage pattern of glucan products induced by substitution of Lys residues in the dextranucrase. *FEBS Lett* 579(21):4739–4745. doi:[10.1016/j.febslet.2005.07.050](https://doi.org/10.1016/j.febslet.2005.07.050)
127. Irague R, Rolland-Sabaté A, Tarquis L, Doublie JL, Moulis C, Monsan P, Remaud-Siméon M, Potocki-Véronèse G, Buléon A (2012) Structure and property engineering of  $\alpha$ -D-glucans synthesized by dextranucrase mutants. *Biomacromolecules* 13(1):187–195. doi:[10.1021/bm201453r](https://doi.org/10.1021/bm201453r)
128. Irague R, Massou S, Moulis C, Saurel O, Milon A, Monsan P, Remaud-Siméon M, Portais JC, Potocki-Véronèse G (2011) NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Anal Chem* 83(4):1202–1206. doi:[10.1021/ac1032148](https://doi.org/10.1021/ac1032148)
129. Reetz MT (2011) Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. *Angew Chem* 50(1):138–174. doi:[10.1002/anie.201000826](https://doi.org/10.1002/anie.201000826)
130. Wijma HJ, Floor RJ, Jekel PA, Baker D, Marrink SJ, Janssen DB (2014) Computationally designed libraries for rapid enzyme stabilization. *Protein Eng Des Sel* 27(2):49–58. doi:[10.1093/protein/gzt061](https://doi.org/10.1093/protein/gzt061)
131. Daude D, Remaud-Siméon M, Andre I (2012) Sucrose analogs: an attractive (bio)source for glycodiversification. *Nat Prod Rep* 29(9):945–960. doi:[10.1039/C2np20054f](https://doi.org/10.1039/C2np20054f)
132. Seibel J, Jordening HJ, Buchholz K (2010) Extending synthetic routes for oligosaccharides by enzyme, substrate and reaction engineering. *Adv Biochem Eng Biot* 120:163–193. doi:[10.1007/10\\_2009\\_54](https://doi.org/10.1007/10_2009_54)
133. Seibel J, Moraru R, Gotze S, Buchholz K, Na'ammieh S, Pawlowski A, Hecht HJ (2006) Synthesis of sucrose analogues and the mechanism of action of *Bacillus subtilis* fructosyltransferase (levansucrase). *Carbohydr Res* 341(14):2335–2349. doi:[10.1016/j.carres.2006.07.001](https://doi.org/10.1016/j.carres.2006.07.001)
134. Baciú IE, Jordening HJ, Seibel J, Buchholz K (2005) Investigations of the transfructosylation reaction by fructosyltransferase from *B. subtilis* NCIMB 11871 for the synthesis of the sucrose analogue galactosyl-fructoside. *J Biotechnol* 116(4):347–357. doi:[10.1016/j.jbiotec.2004.10.019](https://doi.org/10.1016/j.jbiotec.2004.10.019)
135. Cheetham PSJ, Hacking AJ, Vlitos M (1989) Synthesis of novel disaccharides by a newly isolated fructosyl transferase from *Bacillus subtilis*. *Enzyme Microb Technol* 11(4):212–219. doi:[10.1016/0141-0229\(89\)90095-1](https://doi.org/10.1016/0141-0229(89)90095-1)
136. Seibel J, Moraru R, Gotze S (2005) Biocatalytic and chemical investigations in the synthesis of sucrose analogues. *Tetrahedron* 61(30):7081–7086. doi:[10.1016/j.tet.2005.05.063](https://doi.org/10.1016/j.tet.2005.05.063)
137. Timm M, Gohl J, Kraus M, Kralj S, Hellmuth H, Beine R, Buchholz K, Dijkhuizen L, Seibel J (2013) An unconventional glycosyl transfer reaction: glucansucrase GTFA functions as an allosyltransferase enzyme. *ChemBioChem* 14(18):2423–2426. doi:[10.1002/cbic.201300392](https://doi.org/10.1002/cbic.201300392)
138. Monchois V, Remaud-Siméon M, Russell RRB, Monsan P, Willemot R-M (1997) Characterization of *Leuconostoc mesenteroides* NRRL B-512F dextranucrase (DSRS) and identification of amino-acid residues playing a key role in enzyme activity. *Appl Microbiol Biotechnol* 48(4):465–472. doi:[10.1007/s002530051081](https://doi.org/10.1007/s002530051081)
139. Ferretti JJ, Gilpin ML, Russell RRB (1987) Nucleotide sequence of a glucosyltransferase gene from *Streptococcus sobrinus* MFe28. *J Bacteriol* 169(9):4271–4278