university of groningen

University Medical Center Groningen

University of Groningen

Structural and functional investigations of Lactobacillus reuteri glucansucrase

Vujicic Zagar, Andreja

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2007

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Vujicic Zagar, A. (2007). *Structural and functional investigations of Lactobacillus reuteri glucansucrase: ... with crystallographic studies on an α-amylase and a prolyl endoprotease from Aspergillus niger*. s.n.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 16-05-2025

Chapter 1

General introduction to glucansucrases

Andreja Vujičić Žagar, Slavko Kralj, Tjaard Pijning, Lubbert Dijkhuizen and Bauke W. Dijkstra

Introduction

It is believed that already in prehistoric times humans used to chew sugarcane to extract its sweet juice. The first evidence of sugar production from sugarcane dates back to about 500 years BC and it took place in India, from where it spread to China and Persia. Centuries later, after conquering Persia, Arabs learnt how to produce sugar and about AD 600 they introduced it into Europe. For about 1000 years it remained a luxury product as it was produced in only small quantities. Only after the large scale production of sugarcane, which became possible because of the importation of slaves from Africa to work on sugarcane plantations in the New World, sugar did start to be commonly used as a sweetener (Kiple and Ornelas, 2000).

Today, the European Union (EU) is the world's second largest sugar producer with ~ 22 000 million tons produced in 2005/06. In the EU and the USA sugar is mainly produced from sugar beets (*Beta* spp.), while in tropical and sub-tropical countries (such as Brazil and India, the first and the third world's largest sugar producers, respectively) it is produced from sugarcane (*Saccharum* spp.) (Sugar Illovo Ltd., 2006). Sugar is mostly used in the food industry, for example in the production of sweets (candies, chocolate), jams, jellies, juices, in the bakery and dairy industries and as a fermentation additive in the production of alcohol-containing drinks. It is also used in non-food industries, such as the textile and construction industries, in leather tanning, and it is used as an ingredient in printer's ink (Canadian Sugar Institute, 2007).

"Sugar" is the common name for sucrose (O-α-D-glucopyranosyl-(1$\rightarrow$2)-β-D-fructofuranoside), a disaccharide formed from a glucose and a fructose residue. In 1861 Louis Pasteur inferred that the jellifying and thickening texture of sugarcane and beet syrups was of microbial origin (Pasteur, 1861). In the 1880s lactic acid bacteria of the species *Leuconostoc mesenteroides* were identified as the organism that produces a large polysaccharide from sucrose (an α-glucan named dextran) that was responsible for the thickening and jellifying properties of the syrups (Crescenzi, 1995). The finding that dextran is produced by a glucansucrase enzyme took place only in the 1950s (Koepsell *et al.*, 1953). It was then also reported that in the presence of certain acceptor sugars, such as the glucose disaccharides maltose and isomaltose, oligosaccharide

synthesis takes place at the expense of polymer production (Koepsell *et al.*, 1953). Later on, other suitable molecules were identified that could also be used as acceptors by glucansucrases leading to formation of various *α*-glyco-conjugates (see below) (Meulenbeld and Hartmans, 2000; Richard *et al.*, 2003; Seo *et al.*, 2005). Partial characterisation of a glucansucrase isolated from *Leuconostoc mesenteroides* NRRL B-1355 was reported by Robyt *et al.* in late 1982 (Côté and Robyt, 1982). Since then, numerous publications on glucansucrases have appeared in the literature (Monchois *et al.*, 1999b; Moulis *et al.*, 2006; van Hijum *et al.*, 2006).

The products of the glucansucrases (both the *α*-glucans and *α*-glyco-conjugates) as well as their derivatives have vast potential applications in industry. *α*-Glucans may be used in the food industry in the production of *e.g.* dairy (due to their thickening and water-binding properties), baked products and beverages. *α*-Glucans and *α*-glyco-oligosaccharides can also be used as prebiotics[1] and bioactive agents (*e.g.* as immunomodulators, anti-ulcer agents, cholesterol-lowering agents). Chemically or physically modified *α*-glucans may be used as metal-complexing compounds, emulsifiers, detergents, bleaching assistants, strengthening additives, and in waste-water purification (Van Geel-Schutten, 2005). In addition, it was recently discovered that α-glucans produced by *Lactobacillus reuteri* 180 have anti-corrosive properties (Van Geel-Schutten, 2005). Bearing in mind that the annual corrosion costs in the industry in the USA alone amount to 276 billion dollars (Koch *et al.*, 2002) and that common anti-corrosive treatments make use of heavy metal (chrome, nickel) containing alloys (Hou *et al.*, 2000), use of an environment-friendly alternative for corrosion prevention or delay such as the above mentioned *α*-glucans would have serious impact on infrastructure, transportation and other industries where metal constructions are employed (Koch *et al.*, 2002).

The importance of the sugar industry in the world's economy is evident from the number of treaties and agreements between countries which regulate and/or control its production and trade (Economic Partnership Agreements between African, Caribbean and Pacific countries and the

[1] Prebiotics are non-digestible food additives that stimulate growth of beneficial intestinal bacteria (Gibson and Roberfroid, 1995).

European Union (Bugnot, 2005), and The General Agreement on Tariffs and Trade, Uruguay Round (GATT Digital Library: 1947-1994, 2006)). Enzymatic production of industrially important *α*-glucans, *α*-glyco-conjugates and their derivatives starting from a cheap substrate such as sucrose may have a significant influence on the sugar industry in the future, as sucrose would not be used only as a food/non-food additive but also as an important raw material.

Lactobacillus reuteri

The gastrointestinal tract of vertebrates is inhabited by many bacteria that together with other autochthonous microflora play an important role in the physiology of these animals (Walter *et al.*, 2003). *Lactobacillus reuteri* is such a bacterium, which was isolated from gastrointestinal tracts of birds and mammals, including humans (Bath *et al.*, 2005; Roos and Jonsson, 2002).

This Gram-positive lactic acid bacterium is known for the production of reuterin (Talarico *et al.*, 1988) and reutericyclin (Ganzle *et al.*, 2000), two broad-spectrum water-soluble antimicrobial compounds. The beneficial effects of *L. reuteri* have been reported against gastrointestinal pathogens like *Helicobacter pylori* (Mukai *et al.*, 2002), common food poisoning bacteria from *Salmonella* spp. (Alak *et al.*, 1997), infectious *Staphylococcus aureus* (Laughton *et al.*, 2006; Vesterlund *et al.*, 2006) and others (Alak *et al.*, 1997; Shornikova *et al.*, 1997). Also the capacity to stimulate the immune system (Christensen *et al.*, 2002) as well as hypocholesterolemic effects of *L. reuteri* have been reported (Taranto *et al.*, 1998). *L. reuteri* has also been reported to have an inhibitory effect on the growth of the major cariogenic microorganisms such as *Streptococcus mutans* (Nikawa *et al.*, 2004). Other studies suggest that certain *L. reuteri* strains may prevent urinary tract infections in females (Reid and Bruce, 2006). Due to these health benefits, *L. reuteri* is generally recognized as safe (GRAS status) and therefore the bacteria as well as its enzymes and products may safely be used in food/feed and pharmaceutical applications.

The ability of *L. reuteri* to adhere to the mucus layer of epithelial cells of the host is of crucial importance, because it allows the bacterium to colonize and persist in the gastrointestinal tract. Lactic acid bacteria (including *L. reuteri*) like many other microorganisms produce a layer of extracellular

polysaccharides (or exopolysaccharides, EPS), which surrounds the bacteria (Crescenzi, 1995). Unlike the cell-surface associated structures (such as capsular polysaccharides), which remain attached to the bacterial cell, EPS are secreted into the environment where they constitute the main part of a slimy and sticky biofilm layer, in which the bacteria are imbedded (see Fig. 1) (Sutherland, 2001). In general, bacterial biofilms are often found associated to implanted medical prosthetic devices (*e.g.* catheters and cardiac pacemakers) and, due to their resistance to antimicrobial agents, are involved in many chronic bacterial infections in humans (Costerton *et al.*, 1987; Costerton *et al.*, 1999). The EPS (more precisely, homopolysaccharides, see below) of such a bacterial biofilm produced by the lactic acid bacteria *Streptococcus mutans* and several *Lactobacillus* species are responsible for the attachment of the bacteria and food debris to tooth enamel and thus play an important role in the formation of dental plaque and cariogenesis (Byun *et al.*, 2004; Hamada and Slade, 1980; Kreth *et al.*, 2005; Loesche, 1986; Tsumori *et al.*, 1997) .

The role of this EPS layer in bacterial adhesion to biological surfaces and in colonization has also been reported (Colby *et al.*, 1995; Costerton *et al.*, 1987). Moreover, it is generally believed that the EPS also play a role in protecting bacterial cells against desiccation, engulfment by phagocytes or predatory microorganisms (Cerning, 1990) and in serving as a reserve of carbohydrates (van Hijum *et al.*, 2006).

Figure 1. *L. reuteri* 121 bacteria growing on agar supplemented with 5% sucrose. The produced exopolysaccharides form a slimy, sticky layer that surrounds the colonies (Reproduced with permission from the thesis by Gerritdina Hendrika van Geel-Schutten, University of Groningen, 2000).

Extracellular polysaccharides (EPS) of lactic acid bacteria

Based on their composition and pathway of biosynthesis EPS are divided in two classes, heteropolysaccharides and homopolysaccharides (van Hijum *et al*., 2006).

Heteropolysaccharides (HePS) are composed of repeating units of three to eight residues long oligosaccharides (Welman and Maddox, 2003). These oligosaccharides are composed of monosaccharides such as galactose, glucose, rhamnose, fructose, mannose and galactosamine (van der Berg *et al*., 1993) and present low structural similarity (De Vuyst, 2001).

The molecular mass of HePS varies from 4.0×10^4 to 6.0×10^6 Da (Ruas-Madiedo, 2002). The repeating units of HePS are synthesized in the cytoplasm from sugar nucleotides by specific glycosyltransferases (Boels, 2001; Welman and Maddox, 2003). These glycosyltransferases (GTFs) are classified in the Glycosyl Transferase section of the Carbohydrate-Active Enzymes database (at http://afmb.cnrs-mrs.fr/CAZY/) (Henrissat and Davies, 1997). They catalyse the transfer of sugar moieties from activated sugar nucleotides to acceptor molecules and are distinct from the glucansucrases (GS), which are also sometimes called glycosyltransferases, GTFs. These latter enzymes have been classified in the Glycoside Hydrolase section of the Carbohydrate-Active Enzymes database in Glycoside Hydrolase family 70 (see below) and they do not require activated sugar moieties for polysaccharide synthesis. The repeating units of HePS are synthesized on a lipophilic carrier molecule attached to the cytoplasmic membrane (van Kranenburg, 1999). How these repeating units are assembled into the polymer and how they are secreted is not fully understood (Welman and Maddox, 2003). Nevertheless, a mechanism involving a “flippase” that would flip the anchored repeating unit form the cytoplasmic to the periplasmic side of the membrane has been proposed (Laws, 2001). Genes encoding the HePS synthesis are found both on plasmids and in the chromosome and are organized in clusters (Welman and Maddox, 2003). These gene clusters consist of genes coding for glycosyltransferases, genes coding for proteins involved in the biosynthesis of repeating units, and genes for polymerisation, export and regulation (Jolly *et al.*, 2002). Examples of HePS used in the industry are xanthan and gellan, which are used as rheological agents in tertiary oil recovery (Crescenzi, 1995) and as a food additive (E418), respectively.

Homopolysaccharides (HoPS) contain only one type of monosaccharide, which can be either a D-glucose or a D-fructose (van Hijum *et al.*, 2006). The respective polymers are therefore named glucans and fructans. In contrast to HeSP synthesis, which requires a relatively large number of enzymes and nucleotide-activated sugars, HoPS are synthesized by the action of a single enzyme. This enzyme uses sucrose as a substrate and the synthesis occurs outside the cell (van Hijum *et al.*, 2006). Fructans are fructose polymers synthesized from sucrose by the action of fructansucrase enzymes. Two types of fructans can be distinguished, levan and inulin, which are synthesized by levansucrases and inulosucrases, respectively. These two fructans differ in the glycosidic bond between the fructose units: levan is a mainly $\beta(2\rightarrow6)$ linked polymer with $\beta(2\rightarrow1)$ linked branches, whereas inulin has mainly $\beta(2\rightarrow1)$ linkages with $\beta(2\rightarrow6)$ linked branches. The molecular mass of the fructans produced by lactic acid bacteria ranges from 2×10^4 to 50×10^6 Da (van Hijum *et al.*, 2006). Levans from *Streptococcus* spp. are reported to be involved in the formation of dental plaque. Other fructans and fructo-oligosaccharides are used for example in the food industry as prebiotics (Parker and Creamer, 1971; van Hijum *et al.*, 2006). Unlike glucansucrases (see below) fructansucrase enzymes have been identified in both Gram-positive and Gram-negative bacteria (van Hijum *et al.*, 2006).

GH70 family glucansucrases

The main topic of this thesis is on glucansucrases (GS or GTF) (EC 2.4.5.1), which are relatively large extracellular enzymes (~160 kDa) secreted by lactic acid bacteria. More precisely, they have been found solely in lactic acid bacteria (LAB) of the genera *Streptococcus*, *Leuconostoc*, *Weissella*, *Lactobacillus* and *Oenococcus* (Monchois *et al*., 1999b; van Hijum *et al*., 2006).

Genes encoding glucansucrases are constitutively expressed in *Streptococcus* and *Lactobacillus* spp., while in *Leuconostoc* spp. they are facultatively expressed (Kralj, 2004; Robyt, 1995; van Geel-Schutten *et al.*, 1999). Based on their amino acid sequences, which are 30 to 80% identical, glucansucrases have been classified in Glycoside Hydrolase family 70 (GH70) (Henrissat and Davies, 1997).

The enzymes cleave the glycosidic bond of sucrose, and transfer the monosaccharide glucose to either a growing glucan chain (polymerisation reaction) or to a low molecular mass acceptor substrate such as maltose or isomaltose (acceptor reaction). In addition, they may transfer the glucose to a water molecule (sucrose hydrolysis) (Monchois *et al.*, 1999b).

Four different GS enzymes have been distinguished based on the glucans that they synthesize. **Dextransucrases** (DSR) are mainly found in *Leuconostoc* spp. and make *α*-glucans with mainly *α*(1→6) linkages called dextran (from dextrarotatory) (Crescenzi, 1995; Monchois *et al.*, 1999b). Dextrans are known for their thickening and jellifying properties, they were used in medicine as plasma extenders and in micosurgery, and are nowadays widely used as chromatographic column material (Crescenzi, 1995; Groenwall and Ingelman, 1948; Jallali, 2003). **Mutansucrases** are mainly found in *Streptococcus* spp. and make water-insoluble mutan, a mainly *α*(1→3) linked glucose polymer (Guggenheim, 1970). They play an important role in the formation of dental plaque and induction of dental caries (Balakrishnan *et al.*, 2000). **Alternansucrase** (ASR), reported only for the *Leuconostoc mesenteroides* NRRL B-1355 strain, alternates formation of *α*(1→3) and *α*(1→6) bonds, and therefore the synthesized *α*-glucan is named alternan (Argüello-Morales *et al.*, 2000). **Reuteransucrases**, GTFA and GTFO, isolated from *Lactobacillus reuteri* strains, make reuteran, a glucose polymer linked by *α*(1→4) and *α*(1→6) glucosidic bonds (Kralj *et al.*, 2002; Kralj *et al.*, 2004b; van Geel-Schutten *et al.*, 1999). A rather unique GH70 glucansucrase is the *L. mesenteroides* NRRL B-1299 dextransucrase DSRE, which is the only glucansucrase reported to date to posses two catalytic domains, CD1 and CD2 (Bozonnet *et al.*, 2002). According to Fabre *et al.* (2005) the CD2 domain (residues 1980-2835) of DSRE is not a glucansucrase *senso stricto* even though its catalytic domain shows "~44% identity" to the catalytic domains of GH70 glucansucrases, because it is incapable of synthesizing the *α*-glucan polymer from sucrose. Instead, it synthesizes *α*(1→2) bonds in the presence of *α*(1→6) gluco-oligosaccharides and is thus a "*α*(1→2) branching enzyme". The CD1 domain of DSRE is specific for *α*(1→6) bond formation according to the same authors, and thus is a typical dextransucrase.

Repeated units	Consensus sequence
"RDV"	**R**(P/N)**DV**-x_{12}-**SGF**-x_{19-22}-**R**(Y/F/D)**S**
Motif T	TDD**KA**(A/T)**TT**A(A/D)**T**S
Motif S	PA(A/T)**DKA**VD**TT**P(A/T)**T**
"TTQ"	TTTQN(A/T)(P/A)NN(S/G)N(D/G)PQS
CW_binding_1[1]	NG**W**IKDNGN**WYY**FDSD**G**KM
A	WYYFNxDGQAATGLQTIDGQTVFDDNGxQVG
B	VNGKTYYFGSDGTAQTQANPKGQTFKDGSVLRFYNLEGQYVSGSGWY
C	GKIFFDPDSGEVVKNRFV
D	GGVKNADGTYSKY
YG	NDGYYFxxxGxxH^Ox(G/N)X$H^O$$H^O$$H^O$
ASR-repeats	DGx_4APY
"KYQ"	AVK(T/A)A(K/Q)(A/T)(Q/K)(L/V)(A/N)K(T/A)KAQ(I/V)(A/T)KYQ- KALKKAKTTKAK(A/T)QARK(S/N)LKKA(E/N)(T/S)S(F/L)(S/T)KA

Table 1. Sequence repeats found in the N- and the C-terminal ends of GH70 glucansucrases. Sequence repeats "RDV" to "TTQ" are found at the N-terminus while repeats A to "KYQ" are found at the C-terminus of the polypeptide chain. In some GH70 GSs the A repeats may also be found at the C-terminus, see text. Legend: in bold = conserved residue, x = non-conserved residue, H^O = hydrophobic residue. [1]Pfam database. For references see text.

Acceptor reaction. In addition to glucose polymers, GH70 glucansucrases are also known for the production of *α*-glyco-conjugates, which are produced by glycosylation of suitable acceptor molecules, such as the disaccharides maltose and isomaltose (Monchois *et al.*, 1999b; van Hijum *et al.*, 2006), alkyl-glucosides (Lovell *et al.*, 2003), and other molecules (salicin, salicyl alcohol, phenol (Seo *et al.*, 2005), and catechol, 4-methylcatechol and 3-methoxycatechol (Meulenbeld and Hartmans, 2000)). Production of glycosylated compounds may find application in the pharmaceutical and cosmetic industries (Cosmetics Solabia Group, 2007; Schweizer and Hindsgaul, 1999).

Putative domain organization of GH70 glucansucrases

Analysis of the amino acid sequences of GH70 glucansucrases led to the hypothesis that these proteins have three putative domains: *1)* an N-terminal variable domain (~120-700 residues), preceded by a secretion peptide (36-40 residues), *2)* a conserved catalytic domain (~900 residues), and *3)* a C-terminal domain (~300-400 residues) (Monchois *et al.*, 1999b; van Hijum *et al.*, 2006).

The N-terminal variable domain. This domain is variable in both length and composition (van Hijum *et al.*, 2006). Different repeating units not previously described in the Pfam (Bateman *et al.*, 2004) and PROSITE (Hulo *et al.*, 2006) databases were identified in its primary structure. The "RDV" motif was found in *L. reuteri* 121 GTFA, *L. reuteri* 180 GTF180, and others (Kralj *et al.*, 2004a);

motif I		
GS180	1503	A**D**WVPD**Q**
amylosucrase	181	V**D**FIFN**H**
TAKA-amylase	116	V**D**VVAN**H**
CGTase	134	I**D**FAPN**H**
consensus seq.		X**D**XXXN**H**

motif II		
GS180	1021	GI**R**V**D**AVDN
amylosucrase	282	IL**R**M**D**AVAF
TAKA-amylase	202	GL**R**I**D**TVKH
CGTase	225	GI**R**M**D**AVKH
consensus seq.		GX**R**X**D**XXZZ

motif III		
GS180	1059	INIL**E**DWG
amylosucrase	324	FFKS**E**AIV
TAKA-amylase	226	YCIG**E**VLD
CGTase	253	FTFG**E**WFL
consensus seq.		XXXX**E**ZZZ

motif IV		
GS180	1131	FVRA**HD**
amylosucrase	388	YVRS**HD**
TAKA-amylase	292	FVEN**HD**
CGTase	323	FIDN**HD**
consensus seq.		XXBB**HD**

Table 2. Conserved sequence motifs of the catalytic domain of the GH70 family *L. reuteri* GS180 glucansucrase aligned with the GH13 family N. polysaccharea amylosucrase, *A. oryzae/niger* (TAKA) α-amylase and *Bacillus circulans* CGTase. The consensus sequence of GH13 family enzymes is indicated at the bottom (legend: X=usually hydrophobic, B=usually hydrophilic, Z=important for specificity, modified from MacGregor *et al.* . The strictly conserved GH13 residues are depicted in bold. Residues of the catalytic triad are enlarged.

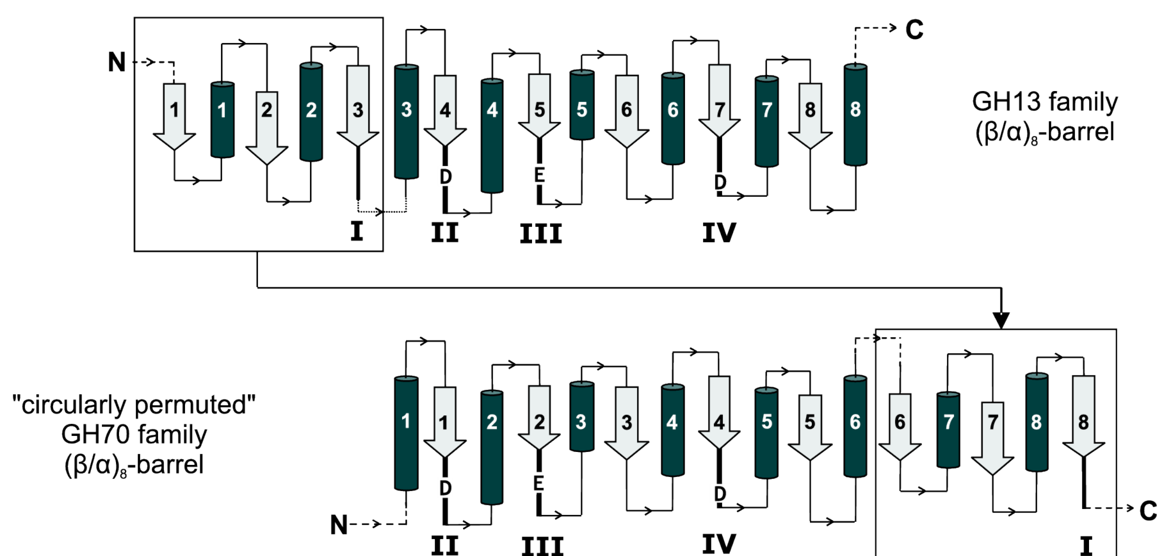


Figure 2. Order of the $(\beta/\alpha)_8$-barrel elements in the GH13 and GH70 family enzymes. Positions of the four conserved sequence motifs (I-IV) found in the α-amylase superfamily are shown.

a "motif-T" was found in *L. mesenteroides* DSRT (Funane *et al.*, 2000), a "motif-S" in *L. mesenteroides* DSRE (Bozonnet *et al.*, 2002), while a "TTQ" motif was identified in *Lactobacillus parabuchneri* 33 GTF33 (Kralj *et al.*, 2004a). The A-repeats, which are usually found at the C-terminus of the polypeptide chain in many GH70 glucansucrases (Monchois *et al.*, 1999b), are found also at the N-terminus in *L. mesenteroides* ASR, DSRS, and DSRB (Janeček *et al.*, 2000) (see Table 1 for consensus and reported sequences of the above mentioned sequence motifs). The role of the N domain and its repeats remains unclear to date. Its deletion in *L. reuteri* GTFA and GTF33 had only minor effects on the glucan product (Kralj *et al.*, 2004a), while deletion of the same domain in the *S. sobrinus* glucansucrase had no effect on the enzyme's activity (Abo *et al.*, 1991). Similar results were also found for *S. downei* GTFI (Monchois *et al.*, 1999a) and *L. mesenteroides* DSRS (Moulis *et al.*, 2006). One member of the GH70 family, the *L. mesenteroides* DSRA dextransucrase completely lacks this N-terminal domain (Monchois *et al.*, 1996). Thus, its function is apparently not essential.

The putative catalytic domain of GH70 glucansucrases contains stretches of amino acid residues that show high sequence similarity to four short amino acid sequence motifs (I-IV) identified in the catalytic $(\beta/\alpha)_8$-barrel domain of *α*-amylase family enzymes and amylomaltases classified in Glycoside Hydrolase families 13 (GH13) and 77 (GH77), respectively, in the CAZy database (http://afmb.cnrs-mrs.fr/CAZY/). The $(\beta/\alpha)_8$-barrel fold, in which eight parallel β-strands are surrounded by eight α-helices, was first found in triosephosphate isomerase (isolated from chicken muscle) and is also known as the TIM barrel fold (Banner *et al.*, 1975). The $(\beta/\alpha)_8$-barrel fold is one of the most ubiquitous folds and it is found in all kingdoms of life. Proteins with a $(\beta/\alpha)_8$-barrel fold are almost always enzymes involved in energy and molecular metabolism (Nagano *et al.*, 2002).

Several strictly conserved amino acid residues found in the four sequence motifs (see Table 2) are involved in catalysis and substrate binding. Therefore, the three enzyme families (GH13, GH70 and GH77) are believed to share the same catalytic mechanism and have been grouped together in the *α*-amylase superfamily or GH-H clan (Henrissat and Davies, 1997; MacGregor *et al.*, 2001). A detailed description of the GH13 family catalytic mechanism is given below.

Circularly permuted (β/α)8-barrel of GH70 glucansucrases. The order of the four conserved sequence motifs in the primary structure of GH70 glucansucrases is different from the one found in the GH13 and GH77 families due to a "circular permutation" event that must have occurred during evolution of the GH70 glucansucrases (MacGregor *et al.*, 1996). In GH13/GH77 families the $(\beta/\alpha)_8$-barrel starts with strand β1, followed by helix α1, β2α2, etc. and it ends with strand β8, followed by helix α8 strand β8 (β1α1, β2α2, … β8α8 arrangement, Fig. 2). The GH70 "circularly permuted" $(\beta/\alpha)_8$-barrel starts with an α-helix which corresponds to helix α3 of the GH13/GH77 barrel and ends with strand β3 (α3, β4α4,…β8α8, β1α1, β2α2, β3 arrangement, Fig. 2). The four conserved motifs in the GH13/GH77 families are found at the C-terminal ends of β-strands β3 (region I), β4 (region II), β5 (region III) and β7 (region IV). Instead, in GH70 glucansucrases, sequence motif I is found C-terminally of motifs II, III and IV. That is, sequence motifs I, II, III and IV in GH70 GSs are located at the C-terminal ends of β-strands β8 (motif I), β1 (motif II), β2 (motif III), and β4 (motif IV). These β-strands are equivalent to strands β3, β4, β5, and β7 of the GH13 family $(\beta/\alpha)_8$-barrel, respectively. A hypothetical evolutionary pathway of gene rearrangements that might have occurred during evolution of GH70 glucansucrases yielding in the "circularly permuted" $(\beta/\alpha)_8$-barrel is given in Chapter 2.

The C-terminal domain (~300-400 residues) follows the conserved putative catalytic domain in all GH70 glucansucrases except in *L. mesenteroides* DSRE, where the additional catalytic domain follows the C-terminal domain (Bozonnet *et al.*, 2002). The C-terminal domain is characterized by the presence of tandem sequence repeats that vary in composition, length and number, and that are specific for each glucansucrase enzyme (Monchois *et al.*, 1999b). Sequence repeats A, B, C, D (Monchois *et al.*, 1999b), N (Monchois *et al.*, 1998), and YG (Giffard and Jacques, 1994) have been defined for GH70 glucansucrases, see Table 1. The sequence repeats A and C are found in *Streptococcus* and *Leuconostoc* species (Monchois *et al.*, 1999b). Repeats B are found in *S. downei* GTFI (Ferretti *et al.*, 1987), the D repeats have solely been identified in *Streptococcus salivarius* ATCC 25975, while the N-repeats have so far been only identified in *L. mesenteroides* DSRS (Monchois *et al.*, 1998). Other sequence repeats unique to *L. mesenteroides* ASR (Argüello-Morales *et al.*, 2000) and *Lb. parabuchneri* GTF33 (Kralj *et al.*, 2004a), respectively have also been identified,

see Table 1. Within the A, B, C and D repeats YG-repeats could be “discerned”. They have been described for streptococcal GSs (Giffard and Jacques, 1994) and later on also for several *Lactobacillus* GSs (Kralj *et al.*, 2004a; Kralj *et al.*, 2002).

The 33 residues long A repeats resemble cell wall (CW) binding repeats (~20 residues long, CW_binding _1, Pfam database (Bateman *et al.*, 2004)). Both repeats consist of conserved aromatic and glycine residues. According to Janeček *et al.* (2000) one and a half CW binding repeat would correspond to one A repeat, but according to Shah *et al.* (2004) A repeats are different from the CW binding repeats. The CW binding repeats are found in *Clostridium difficile* and *Clostridium sordelli* toxins (Moncrief and Wilkins, 2000; Wren, 1991), in choline-binding proteins of *Streptococcus pneumoniae* and its bacteriophages (García *et al.*, 1988), as well as in other surface-associated proteins (Kremer *et al.*, 1999; Makino *et al.*, 1998; Sanchez-Beato *et al.*, 1995). The C-terminal repetitive domain (CRD) of *Clostridium* toxins binds carbohydrates in the gut epithelium (Greco *et al.*, 2006; Moncrief and Wilkins, 2000), while choline-binding proteins bind choline-containing teichoic acids found in *S. pneumoniae* cell walls (García *et al.*, 1988). A description of the 3D structure of the *C. difficile* CRD in complex with a trisaccharide (Greco *et al.*, 2006) as well as of the 3D structure of the *S. pneumoniae* ChBD in complex with choline (Fernández-Tornero *et al.*, 2001) will be given in Chapter 2. It is important to mention that the CRD sequence repeats consist of both short (21-30 residues) and long (50 residues) repeats, whereas the ChBD contains only short repeats (20 residues). These differences are responsible for the different binding specificities of the ChBD and CRD domains (for more details see Chapter 2, p. 67). Due to the similarity of the A and CW repeats, *Streptococcus downei* GTFI was tested for the ability to bind choline molecules, but no evidence for such binding was obtained (Shah *et al.*, 2004). However, this could be expected since the repeating units at the C-terminus of GTFI (A-A-C-A-C-A-C-B-A-C-B-A-C (Ferretti *et al.*, 1987)) are different in both length and composition from the CW tandem repeats found in the *S. pneumoniae* ChBD.

The C-terminal domains of the glucansucrases from oral streptococci which contain A, B and C repeats (Monchois *et al.*, 1999b) bind glucans (*e.g.* dextran) and were thus called glucan binding domains (GBD) (Lis *et al.*, 1995; Shah *et al.*, 2004). GBDs are also found in proteins from oral

streptococci that do not have glucansucrase activity (Wren, 1991). The glucan binding ability of the C-terminal domain of oral streptococcal glucansucrases is believed to be responsible for the sucrose-dependent attachment of the bacteria to tooth surfaces (Hamada and Slade, 1980; Loesche, 1986). Involvement of the C-terminal domain of oral streptococcal GSs in glucan binding is well documented (Kingston *et al.*, 2002; Konishi *et al.*, 1999; Lis *et al.*, 1995). However, the term glucan binding domain is widely used also for the C-terminal domains of other GH70 glucansucrases (Monchois *et al.*, 1999b) even though there is no clear evidence that these domains are generally involved in glucan binding (Shah *et al.*, 2004).

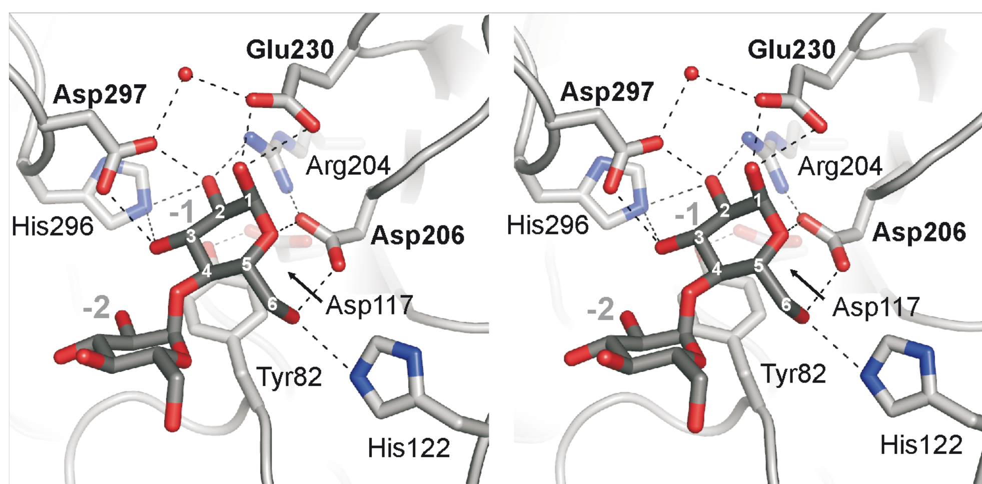
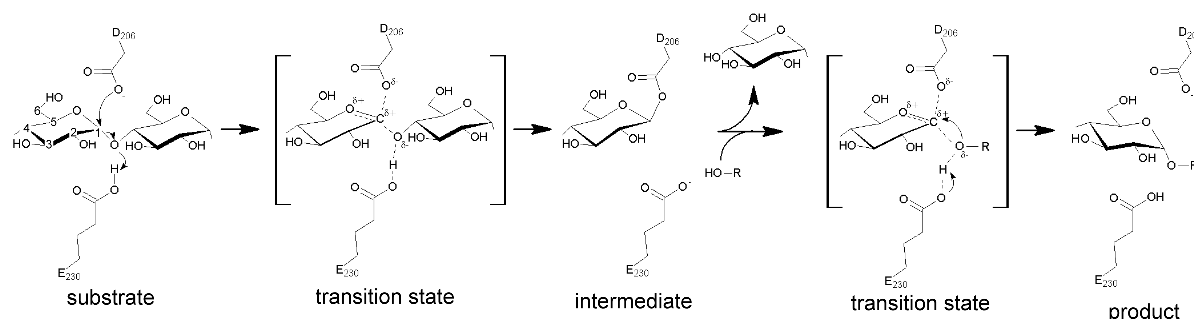


Figure 3. Stereo-view of the active site of *Aspergillus niger* α-amylase in complex with maltose (Vujičić-Žagar and Dijkstra, 2006). The seven strictly conserved amino acid residues of sequence motifs I-IV are shown, together with a conserved tyrosine residue. The catalytic triad residues are labelled in bold letters. Hydrogen bonding interactions (up to H-bond length of 3.5 Å) are depicted as dashed lines.



Scheme 1. General reaction mechanism of α-amylase superfamily enzymes.

Catalytic mechanism of α-amylase superfamily enzymes

The GH13 family is the largest of the 109 GH families currently tabulated in the CAZy database (http://www.cazy.org/). The first 3D structure of a GH13 family member was that of the *α*-amylase from the fungus *Aspergillus oryzae*, which had been isolated from "Takadiastase Sankyo" and was therefore named TAKA-amylase. This 3 Å structure was published in 1980 (Matsuura *et al.*, 1980). Being the first structurally characterized GH13 family enzyme the TAKA-amylase structure has often been used as the representative for the entire family (Kuriki and Imanaka, 1999). In Chapter 4 a 1.8 Å resolution crystal structure is described of an *Aspergillus niger α*-amylase, which shows 100% sequence identity to TAKA-amylase in complex with maltose, the shortest substrate of this enzyme (Nitta *et al.*, 1971). In addition, a 1.6 Å resolution structure of a novel orthorhombic crystal form of this enzyme is presented in that chapter.

α-Amylases (1,4-*α*-D-glucan glucanohydrolase, EC 3.2.1.1) catalyse the hydrolysis of *α*(1→4) glucosidic bonds in starch and related oligo- and polysaccharides. The active site of GH13 *α*-amylases is positioned at the bottom of a substrate binding cleft formed by the C-terminal ends of β-strands β4, β5, β7 and a conserved tyrosine residue which in the primary structure of GH13 enzymes is located N-terminally to the β4 strand. Cleavage of the *α*(1→4)-glycosidic bond occurs between subsites -1 and +1 via a double displacement mechanism involving a covalent glucosyl-enzyme intermediate at subsite -1 and with retention of the *α*-anomeric configuration of the product upon hydrolysis, see below (Koshland and Clarke, 1953; Uitdehaag *et al.*, 1999b). Three acidic amino acids participate in the hydrolysis reaction: Asp206 (TAKA-amylase numbering) is the nucleophile, Glu230 acts as the catalytic acid/base, while Asp297 is involved in stabilization of the oxocarbenium ion-like transition state (Matsuura *et al.*, 1984; Uitdehaag *et al.*, 1999b). These three catalytic residues are located in the conserved sequence motifs II, III and IV, respectively (Table 2). In total six strictly conserved residues are found in sequence motifs I-IV; together with an additionally conserved tyrosine residue they are all involved in the catalytic cleavage of the scissile glucosidic bond and transition state stabilization, Fig. 3 (MacGregor *et al.*, 2001).

The general reaction mechanism of GH13 family enzymes, which was first proposed by Koshland in late 1953 (Koshland, 1953) is depicted in Scheme 1. Upon substrate binding, in the first reaction step the glutamic acid residue (in the uncharged, protonated state) acts as an acid protonating the glycosidic oxygen of the scissile bond. At the same time the carboxylic oxygen of the nucleophilic aspartic acid residue attacks the C1 carbon of the substrate's glucosyl residue bound at subsite -1. This leads to the formation of a covalently bound *β*-glucosyl enzyme intermediate, which reaction is believed to proceed via an oxocarbenium ion-like transition state. Then, the protonated sugar moiety bound at the +1 subsite is released. The second part of the reaction is the reverse of the first part (double displacement). The now negatively-charged glutamate residue acts as a base deprotonating a hydroxyl group of an incoming sugar or a water molecule, which molecule then acts as the nucleophile attacking the C1 carbon of the covalently bound glucosyl unit. This leads to the collapse of the glucosyl-enzyme intermediate and the release of product. The reaction occurs thus with overall retention of the *α*-anomeric configuration of the -1 glucosyl moiety.

Detailed structural analysis of the GH13 family member *Bacillus circulans* 251 cyclodextrin glycosyltransferase (CGTase) in complex with an intact substrate (maltononaose) and covalently bound intermediate (4-deoxyl-maltotriosyl-CGTase) solved at 2.1 and 1.8 Å resolution, respectively, revealed the following details regarding the reaction mechanism of GH13 enzymes (Uitdehaag *et al.*, 1999b; Uitdehaag *et al.*, 2002).

Substrate binding. An *α*(1→4) linked glucose polymer binds in the active site of an GH13 enzyme with its non-reducing end bound at subsite –*n*, and extends to the +*n* subsites. Upon binding to subsite -1 the -1 glucose ring conformation changes from the minimum energy $^4C_1$ chair conformation to a distorted half-chair conformation ($^2H_3$), in which the C2-C1-O5-C5 torsion angle has flattened by about 20° towards the transition state conformation (0°). This distortion is caused mainly by the transition state stabilizer Asp297, TAKA-amylase numbering (or Asp1136, GS180 numbering) and His122 (Gln1509) (Uitdehaag *et al.*, 1999b). The Asp297 carboxylate group binds to the C2 and C3 hydroxyl groups of the -1 glucose, while His122 is hydrogen-bonded to its C6-hydroxyl group, Fig. 3 (Brzozowski and Davies, 1997; Kuriki and Imanaka, 1999; Strokopytov *et al.*, 1995; Uitdehaag *et*

al., 1999b; Vujičić-Žagar and Dijkstra, 2006). In a complex of TAKA-amylase with the inhibitor acarbose the acid/base Glu230 is at hydrogen bonding distance to inhibitor's nitrogen atom mimicking the glycosidic oxygen of the scissile bond (Brzozowski and Davies, 1997) and the same is observed in the *B. circulans* CGTase-acarbose complex (Strokopytov *et al.*, 1995) as well as in the CGTase (E257Q/D229N)-maltononaose complex (Uitdehaag *et al.*, 1999b). Thus, all crystal structures are compatible with a function of Glu230 (Glu1063) as an acid protonating the scissile glycosidic oxygen in the first reaction step upon substrate binding. At the same time one of the carboxylic oxygen atoms of the nucleophilic aspartic acid Asp206 (Asp1025) is at an appropriate distance and orientation with its *syn* lone pair electrons pointing towards the anomeric C1 carbon of the -1 bound glycosyl moiety of the substrate to perform a nucleophilic attack on that atom (Uitdehaag *et al.*, 1999b). This leads to the formation of the *β*-glucosyl enzyme intermediate, which is believed to proceed *via* an oxocarbenium ion-like transition state. The second Asp206 carboxylate oxygen interacts with its *anti*-oriented lone pair orbital with another strictly conserved residue, Arg204 (Arg1023), Fig 3. The latter residue is therefore believed to keep the nucleophilic Asp206 in a correct orientation for the nucleophilic attack and to ensure that the nucleophilic Asp206 stays deprotonated (Uitdehaag *et al.*, 1999b).

In the uncomplexed CGTase structure the acid/base Glu230 (Glu257 in CGTase) makes a direct H-bond to the transition state stabilizer residue Asp297 (Asp328 in CGTase) (Strokopytov *et al.*, 1995; Uitdehaag *et al.*, 1999b). It has been suggested that Asp328 might plays a role in increasing the p*Ka* of the acid/base Glu257, ensuring its protonated state even at higher pHs (Strokopytov *et al.*, 1995). However, in the *A. niger* apo- and maltose-bound *α*-amylase structures the acid/base Glu230 is at ~3.8 Å distance from the transition state stabilizer Asp297. Instead of a direct H-bond Glu230 makes short water-mediated H-bonds (both 2.7 Å long) to Asp297. The same is observed in the TAKA-amylase-acarbose complex (Brzozowski and Davies, 1997).

Transition state. In the oxocarbenium-ion-like transition state the glucose reaction centre (C2-C1-O5-C5 atoms) at subsite -1 is planar. The bond between the O5 and C1 atoms has a double bond character with the positive charge localized mainly on the O5 atom. The glycosidic oxygen atom is almost protonated and the bond between the glycosidic oxygen atom and the C1 atom is almost

broken (Sinnott, 1990; Tanaka *et al.*, 1994; Uitdehaag *et al.*, 2002). The amino acid residues that contribute to the stabilization of the transition state are Asp297 (Asp1136, GS180 numbering), which binds the C2 and C3 glucosyl hydroxyl groups, His296 (His1135) and Arg204 (Arg1023), which contact the C2 hydroxyl group, and His122 (Gln1509), which makes a strong H-bond to the C6 hydroxyl group (Nakamura *et al.*, 1993; Søgaard *et al.*, 1993; Tanaka *et al.*, 1994). His296 (His1135) is reported to make H-bonds not only to C2, but also to the C3 hydroxyl group in a number of other GH13 *α*-amylases (Brzozowski and Davies, 1997; Qian *et al.*, 1994; Søgaard *et al.*, 1993; Vujičić-Žagar and Dijkstra, 2006). The importance of the interactions with the C2-hydroxyl group for transition state stabilization is well documented (Braun *et al.*, 1995; McCarter *et al.*, 1992). Binding to the C2-hydroxyl group contributes “at least” 33.5 kJ/mol to transition state stabilization, while binding to the C3-, C4- and C6- hydroxyl groups contributes to a much lesser extent (~16.7 kJ/mol) (McCarter *et al.*, 1992). This is explained by the fact that the C2-hydroxyl group is the one that undergoes the most significant structural rearrangements between the ground state and the transition state (McCarter *et al.*, 1992). The strong H-bonds to the C2-hydroxyl group helps the -1 glucose ring distortion towards the planar transition state conformation (in particular by Asp297 (Asp1136)), as well as to reduce the electronegativity of the glucose O2 atom, which would otherwise destabilize the positively charged oxocarbenium ion-like transition state (Uitdehaag *et al.*, 1999b). The strong H-bond with one of the carboxylate oxygen atoms of Asp297 (Asp1136) negatively polarizes the glucose O2 atom (by moving the hydroxyl proton from O2 towards Asp297). The resulting partial negative charge on O2 is then stabilized by interactions with His296 (His1135) and Arg204 (Arg1023). This explains the presence of two basic residues in subsite -1, while one would expect that the positively charged transition state is generally stabilized by a negative potential in the active site (Uitdehaag *et al.*, 1999b).

Intermediate. The first 3D structure of a covalent glycosyl-enzyme intermediate of a GH13 family enzyme was that of 4-deoxy-maltotriosyl-CGTase, solved at 1.8 Å resolution (Uitdehaag *et al.*, 1999b). Recently, also the crystal structures of *Neisseria polysaccharea* amylosucrase (AS) at 2.2 Å resolution with a covalently bound glucosyl moiety was reported (Jensen *et al.*, 2004) and that of the GH77 *Thermus thermophilus* amylomaltase with a covalently bound acarbose-derived trisaccharide

(Barends *et al.*, 2007). The conformation of the glucose ring in these covalent intermediate structures is an undistorted $^4C_1$ chair conformation, indicating that subsite -1 is stereo-chemically complementary to the intermediate (Jensen *et al.*, 2004; Uitdehaag *et al.*, 1999b). Compared to substrate binding at the same subsite, the hydrogen bonding geometries of the intermediate to Arg227 (TAKA-amylase numbering; Arg1023, GS180 numbering) and His327 (His1135) residues improve, as well as the hydrophobic stacking interaction to Tyr100 (Tyr1465) (Uitdehaag *et al.*, 1999b). In CGTase a shift of 2 Å of the glucose intermediate is accompanied by a shift of the transition state stabilizer Asp297 (Asp1136), maintaining the hydrogen bonds of the Asp297 carboxylate oxygen atoms to the C2 and C3 hydroxyl groups. In contrast, the hydrogen bond of the O6 hydroxyl group to His140 (His122 in TAKA-amylase) is lost in CGTase (Uitdehaag *et al.*, 1999b), while in amylosucrase the corresponding residue (His187 in AS) is still H-bonded to the C6 hydroxyl group (Jensen *et al.*, 2004). In addition, improved stacking interactions with Tyr82 (Y1465) might further stabilize the intermediate. Movement of the corresponding tyrosine residue in the CGTase covalent intermediate (Tyr100) by ~1.5 Å is believed to pull the -1 glucose out of the substrate position (Uitdehaag *et al.*, 1999b).

Asp117 (Asp1504) is the only amino acid of the seven strictly conserved residues of the α-amylase family which does not make direct contacts with substrate at subsite -1. Instead, it is at short H-bonding distance (2.5 Å in the *A. niger* maltose-α-amylase structure) to the conserved Tyr82 (Tyr1465), which has hydrophobic stacking interactions with the glucose residue bound at subsite -1, Fig. 3.

Stereochemical complementarity between the intermediate and subsite -1 could be important to stabilize the intermediate to prevent reformation of the substrate since the carboxylate group of the nucleophilic aspartate is a much better leaving group than the C4-hydroxyl group of the +1 glucose moiety (Uitdehaag *et al.*, 1999b). A relatively stable intermediate might be necessary to allow the +1 (to +*n*) bound substrate to leave and an acceptor (another sugar or a water molecule) to enter the active site (Uitdehaag *et al.*, 1999b). Therefore, the driving force for the formation of the covalent glucosyl-enzyme intermediate starting from the substrate would come from the relaxation of the high-energy distorted chair conformation of the substrate, the improvement of the H-bonding network to

His296 (His1136) and Arg204 (Arg1023), and the improved stacking interactions with the conserved Tyr82 (Tyr1465) residue (Uitdehaag *et al.*, 1999b).

GH13 family glucansucrase - *Neisseria polysaccharea* amylosucrase

Neisseria polysaccharea amylosucrase (AS) (E.C. 2.4.1.4) is a GH13 family glucansucrase, which catalyses the transfer of a glucosyl moiety from sucrose to an acceptor molecule with formation of exclusively $\alpha(1 \rightarrow 4)$ glucosidic bonds (Potocki de Montalk *et al.*, 2000b). In this way an amylose-like insoluble glucose polymer is synthesized, and also malto-oligosaccharides and sucrose isoforms can be formed, while sucrose hydrolysis is a minor side reaction (Potocki de Montalk *et al.*, 2000b). AS is a 72 kDa enzyme and thus significantly smaller than GH70 glucansucrases. The in vivo function of the enzyme is believed to be the extension of glycogen-like oligosaccharides, since a 98-fold increase in k_{cat} is observed in the presence of such oligosaccharides (Albenne *et al.*, 2007; Potocki de Montalk *et al.*, 2000a). The oligosaccharide extension occurs at the expense of oligosaccharide formation and sucrose hydrolysis (Potocki de Montalk *et al.*, 2000a). Extensive structural investigations of the full-length *N. polysaccharea* AS (Skov *et al.*, 2001) in complexes with D-glucose and sucrose (Mirza *et al.*, 2001; Skov *et al.*, 2006), maltoheptaose (Skov *et al.*, 2002; Skov *et al.*, 2006), and a covalent glucosyl-enzyme intermediate (Jensen *et al.*, 2004) revealed that the reaction mechanism of AS is the same as that described for other GH13 family members (see above). GH13 amylosucrases have also been identified in other bacteria such as *Deinococcus radiodurans* (Pizzut-Serin *et al.*, 2005) and *Neisseria meningitis* (Jensen *et al.*, 2004).

Overall structure. The crystal structure of AS revealed that the polypeptide chain is folded into five distinct domains. An α-helical domain N is located at the N-terminus. It is followed by the catalytic domain A, which is folded into a $(\beta/\alpha)_8$-barrel. Domains B and B' are inserted between strand β3 and helix α3, and strand β7 and helix α7 of the $(\beta/\alpha)_8$-barrel, respectively, while domain C, which has an eight-stranded β-sandwich fold, is located at the C-terminus of the polypeptide chain (Skov *et al.*, 2001). Domains A, B and C are found in other GH13 family members such as TAKA-amylase

(Brzozowski and Davies, 1997), CGTase (Lawson *et al.*, 1994), and oligo-1,6-glucosidase (Watanabe *et al.*, 1997), but domains N and B' are specific to AS. No Ca^{2+} ions were found to bind to AS, which is in contrast to the GH70 glucansucrases, which are Ca^{2+}-dependent enzymes (see Chapter 2 for details). Domain B in *N. polysaccharea* AS consists of two short antiparallel β-sheets, which are sandwiched between two α-helices. This domain, which contributes to substrate binding (Van der Veen *et al.*, 2000), varies considerably in both sequence and length among the GH13 family members (Janeček, 1997). In contrast, domain C, which is built up of β-strands in a Greek-key motif (Richardson, 1977), is in general structurally conserved in the GH13 family. For this domain a role in raw starch binding has been reported for *B. circulans* 251 CGTase (Penninga *et al.*, 1996). Sugar binding to domain C has also been reported for AS (Skov *et al.*, 2002). The AS specific domain B' consists of three short α-helices and a short β-sheet; it is believed to assist in elongation of glycogen branches *via* conformational movements, see below (Albenne *et al.*, 2007).

The active site and reaction mechanism. The active site of AS is located at the bottom of a deep and narrow pocket (Mirza *et al.*, 2001). Sucrose binds with its glucosyl moiety in subsite -1 and with the fructosyl moiety in subsite +1. Co-crystallization experiments of the AS(E328Q) mutant with maltoheptaose enabled mapping of subsites +2 to +5, and also revealed the presence of two other binding sites located at the surface of the enzyme in domains B' and C (Skov *et al.*, 2002). In contrast to the endo-acting enzymes from the GH13 family, such as TAKA-amylase, in which the active site is a long cleft containing subsites mapped from -3 to +5 (Brzozowski and Davies, 1997; Vujičić-Žagar and Dijkstra, 2006), the active site of AS is a pocket closed by a salt-bridge, which does not extend beyond subsite -1 (for more details see Chapter 2). The AS E328Q mutant in complexes with sucrose (Mirza *et al.*, 2001) and a covalently bound glucosyl enzyme intermediate (Jensen *et al.*, 2004) show that the amino acid residues involved in binding the glucosyl moiety of sucrose (subsite -1) and in cleavage of the sucrose glycosidic bond are strictly conserved, as expected for a GH13 family enzyme, Table 2. From this it can be inferred that the reaction catalyzed by AS proceeds via a covalently bound *β*-glucosyl-enzyme intermediate utilizing a double displacement mechanism with retention of the *α*-anomeric configuration as described above.

Upon sucrose cleavage and formation of the intermediate, the fructosyl moiety should leave the active site making place for the non-reducing end of the acceptor molecule (a glycogen branch). After the transfer reaction the now glucosylated acceptor molecule leaves the active site and another sucrose can enter. How this entering and leaving process proceeds has been analyzed by “robotics calculations”, molecular modelling and site-directed mutagenesis experiments (Albenne *et al.*, 2007). Back and forth movements of a flexible loop in domain B’ (residues 443-449) would move the glucosylated acceptor in and out of the active site. The “out” movement should free subsites -1 and +1 to allow a sucrose molecule to enter, while the “in” movement would bring the acceptor close to the catalytic centre once the glycosyl-enzyme intermediate has been formed and fructose has been released from the active site (Albenne *et al.*, 2007). The B’ sugar binding site is proposed to be an “anchoring platform” that captures the acceptor polymer (glycogen) and directs it towards the active site. Thus, a semi-processive elongation of the glycogen by *N. polysaccharea* AS is proposed (Albenne *et al.*, 2007).

Control of hydrolysis reaction. According to Albenne *et al.* (2007) binding of glycogen in proximity to the active site results in exclusion of water molecules explaining the absence of sucrose hydrolysis in the presence of glycogen. Additionally, as recently described for GH77 amylomaltases (Barends *et al.*, 2007), the covalent intermediate of AS (Jensen *et al.*, 2004) is protected from hydrolysis in three different ways. The -1 glucose is bound in its low energy $^4C_1$ chair conformation, the acid/base Glu328 is not productively positioned to activate a water molecule, and the ester bond between the nucleophile (Asp286) and the glucose ring is almost perpendicular making the C1 atom less accessible to an incoming nucleophile. Thus, this GH13 family glucansucrase protects its covalent intermediate in the same way as the related GH77 family enzymes.

Reaction mechanism of GH70 glucansucrases

Based on the sequence similarity to enzymes of the GH13 *α*-amylase family, which includes a glucansucrase (*N. polysaccharea* AS) reaction mechanism similar to that of GH13 enzymes has been proposed for GH70 glucansucrases. In these latter enzymes the reaction is also supposed to proceed

via a double displacement mechanism, involving a *β*-glucosyl-enzyme intermediate, retaining the *α*-anomeric configuration of the substrate in the product (Moulis *et al.*, 2006; van Hijum *et al.*, 2006). However, how GH70 glucansucrases transfer the covalently linked glucosyl moiety to an acceptor during the second half of the reaction and, in particular to which end of the acceptor molecule, has been a matter of debate for several decades. Altogether, two catalytic mechanisms were proposed. The first one is analogous to the general mechanism of GH13 enzymes as described above, in which the covalently linked glucosyl moiety is transferred to the non-reducing end of the growing glucan chain (Monchois *et al.*, 1999b). The other proposed mechanism involves two catalytic centres and transfers the glucosyl-moiety to the reducing end of a polymer (Robyt, 1995; Robyt *et al.*, 1974). Nowadays, the first reaction mechanism has found wider acceptance. Supporting evidence has come from the findings of Mooser *et al.* (1989) who trapped a single *β*-glucosyl-enzyme intermediate of GH70 *Streptoccocus sobrinus* glucansucrase and from the crystal structure of the *β*-glucosyl-enzyme intermediate of the GH13 glucansucrase *N. polysaccharea* AS (Jensen *et al.*, 2004). Moulis *et al.* (2006) have recently reported detailed biochemical and mutagenesis experiments of two GH70 enzymes, the alternansucrase (ASR) and dextransucrase (DSRS) from *Leuconoctoc mesenteroides*, which also support a single active site and a reaction mechanism with elongation of the polymer chain at the non-reducing end.

Nevertheless, the product specificity determinants of GH70 glucansucrases, which determine the different physical and rheological properties of the products (Yang, 2000), remain unknown. The product specificity is determined in the second step of the reaction (after the *β*-glucosyl-enzyme intermediate has been formed) and it depends on the position and orientation of the (glucose) acceptor that binds at subsite +1. Thus, depending on whether the C2, C3, C4 or C6 hydroxyl group of the acceptor glucose points towards the catalytic centre different glucosidic bonds ($\alpha(1\rightarrow2)$, $\alpha(1\rightarrow3)$, $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ bonds, respectively) will be formed.

Physicochemical properties of the product, such as viscosity, adhesiveness, resistance to breakdown, etc. depend on the type of glucosidic bonds linking consecutive glucose units, the degree and type of branching, the length of the branched chains, and the glucan molecular mass, all of them

specific for a particular GH70 enzyme (Monchois *et al.*, 1999b; van Hijum *et al.*, 2006). Thus, understanding the glucosidic bond specificity of GH70 glucansucrases, as well as the mechanisms that control and/or regulate the elongation of the polymer chain is the aim of the current research on glucansucrase enzymes.

Conclusions

GH70 glucansucrases are relatively large (~160 kDa) extracellular enzymes produced by lactic acid bacteria. Using sucrose as a substrate they synthesize high molecular mass glucose polymers, called *α*-glucans. Additionally, in the presence of suitable acceptor molecules (such as *e.g.* maltose) *α*-glyco-conjugates are synthesized at the expense of the polymeric product. Due to the broad product specificity among the GH70 glucansucrases their products have markedly different properties, which make them interesting compounds for various industrial applications. To date, the lack of any structural information has hampered in-depth understanding of the product specificity determinants of GH70 glucansucrases.

Scope of this thesis

This thesis is divided in two parts. The first part (Part I) includes chapters 1, 2, 3, and 4, while the second part (Part II) consists of chapter 5.

The main topic of chapters 1-3 is a GH70 family glucansucrase from *Lactobacillus reuteri* 180, named GS180. The aim of this research was to structurally characterise an N-terminally truncated (-ΔN), fully active GS180-ΔN in the apo-form and in complex with substrate and acceptor molecules in order to enable investigations of the enzyme's product specificity. In chapter 4 the 3D structure of an *Aspergillus niger α*-amylase-maltose complex (*A. niger α*-amylase shows 100% sequence identity to TAKA-amylase) in a novel crystal form is reported. In the second part of the thesis, consisting of chapter 5 crystallization experiments and preliminary X-ray data analysis of a prolyl endoprotease from *Aspergillus niger* are described.

In **Chapter 1** general introduction to glucansucrase enzymes is given.

In **Chapter 2** the three-dimensional structure of the N-terminally truncated (-ΔN), fully active 117 kDa glucansucrase GS180-ΔN from *Lactobacillus reuteri* 180 is presented. The structure was solved using experimental phase information to 2.3 Å resolution obtained from a three-wavelength anomalous diffraction data set (MAD) collected around the Se edge. The phase information was combined with 1.7 Å resolution native data and the model was refined at high resolution. Furthermore, 3D structures of the inactive (active site mutant D1025N) GS180-ΔN in complex with sucrose (substrate), and wild-type GS180-ΔN in complex with maltose (acceptor) determined at 2.3 and 2.0 Å resolution, respectively, give insights into the product specificity determinants of the enzyme.

In **Chapter 3** two additional crystal forms (“open” and “closed”) of *L. reuteri* GS180-ΔN are presented, which reveal conformational flexibility of the enzyme. In the “open” crystal form a BIS-TRIS-propane molecule from the buffer is bound in the active site.

In **Chapter 4** purification, crystallization and structure determination of a hitherto not described monoclinic crystal form of the *A. niger* α-amylase in complex with maltose, the shortest chain length substrate of this enzyme, is reported at 1.8 Å resolution. Furthermore, a 1.6 Å new orthorhombic crystal form is presented ($P2_12_12$).

In **Chapter 5** an introduction is given to prolyl endoprotease (EndoPro), an industrially important and poorly studied enzyme secreted by the food grade fungus *A. niger*. Crystallization trials for both the glycosyated and deglycosylated enzyme are described, and a preliminary analysis of X-ray diffraction data to 2.8 Å resolution collected for the fully glycosylated enzyme is presented.