

Polysaccharide-synthesizing Glycosyltransferases and Carbohydrate Binding Modules: the case of Starch Synthase III

Diego F. Gomez-Casati^{1,2*}, Mariana Martín¹ and María V. Busi^{1,2}

¹Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI-CONICET), Universidad Nacional de Rosario, Suipacha 531, 2000, Rosario, Argentina; ²IIB- Universidad Nacional de General San Martín (UNSAM), 25 de Mayo y Francia, 1650, San Martín, Buenos Aires, Argentina

Abstract: Glycosyltransferases (GTs) are a ubiquitous group of enzymes that catalyze the transfer of sugar moieties from activated donor molecules to specific acceptor molecules, forming glycosidic bonds. Nucleotide-sugars, lipid phosphate sugars and phosphate sugars can act as activated donor molecules while acceptor substrates involve carbohydrates, proteins, lipids, DNA and also, numerous small molecules (i. e. antibiotics, flavonols, steroids). GTs enzyme families are very ancient. They are founded in all the three domains of life and display three different folds (named GT-A, GT-B and GT-C) which are a variant of a common α/β scaffold. In addition, several GTs contain an associated non-catalytic carbohydrate binding module (CBM) that could be critical for enzyme activity.

This work reviews the current knowledge on the GTs structures and functions and highlights those enzymes that contain CBMs, particularly starch binding domains (SBDs). In addition, we also focus on *A. thaliana* starch synthase III enzyme, from the GT-5 family. This protein has a GT-B fold, and contains in its N-terminal region three in tandem SBDs, which are essential for the regulation of enzyme activity.

Keywords: Carbohydrate binding modules, enzymes, glycosyltransferases, starch, starch binding domains, starch synthase.

INTRODUCTION

Structure and classification of glycosyltransferases

Glycosyltransferases (GTs) are defined as the enzymes that transfer an activated glycosyl donor, with a phosphate leaving group, to an acceptor molecule [1]. The most common sugar substrates are nucleotide sugars such as uridine nucleotides (i.e. UDPGlc, UDPGal, UDPGalNAc); adenine nucleotides (ADPGlc); and citidine and guanidine nucleotides such as CDPGlc and GDPMan, however, GTs could also use other donors such as nucleosides monophosphates and lipid phosphate sugars. Nucleotide sugars GTs are often referred to as “Leloir enzymes” in honor to L.F. Leloir, who discovered the sugar nucleotides and their role in the biosynthesis of carbohydrates [1, 2].

The acceptor substrates could be a wide range of biomolecules, including lipids, proteins and oligosaccharides, thus synthesizing glycolipids, glycoproteins and glycoconjugates. Based on the type of sugar that they transfer, GTs are grouped into several families (i.e. glucosyltransferases, galactosyltransferases, sialyltransferases, etc.).

In addition, GTs can also be classified as either retaining or inverting enzymes according to the anomeric configuration of the substrates and reaction products. This stereochemically classification is the result of the different mecha-

nisms used by the two groups or enzymes [3]. Retaining GTs operate via an internal return S_Ni-like mechanism with the retention of stereochemistry at the anomeric reaction center of the donor substrate; whereas inverting GTs most likely follow a single displacement mechanism where the acceptor substrate performs a nucleophilic attack at carbon C-1 of the sugar donor [1].

In the case of GTs, the enzymes within the same family catalyze the reaction with the same stereoselectivity. However, despite the fact that many GTs recognize identical donor or acceptor substrates, there is a very low sequence identity among the enzymes from different families [4].

It has been described that GTs adopt two different folds named GT-A and GT-B (Fig. 1) [5, 6]. The GT-A fold comprises two dissimilar domains, one for nucleotide binding and the other for the binding of the acceptor molecule, and was described for example for the SpsA protein, a nucleotide-diphosphosugar transferase from *Bacillus subtilis* involved in the synthesis of the spore coat and belonging to the GT-2 family (Fig. 1A) [7, 8]. The GT-B fold was described initially for the bacteriophage T4 beta-glucosyltransferase (EC 2.4.1.27) which catalyses the transfer of glucose from UDPGlc to cytosine bases in T4 duplex DNA [9]. Furthermore, the GT-B fold is present in other families such as the GT-3 and GT-5 families including about 90 glycogen or starch synthases from different organisms. The first protein whose structure was described in this group of GTs was the glycogen synthase from *Agrobacterium tumefaciens* (Fig. 1B) [10]. It was reported that the overall fold and the architecture of the active site of the protein are remarkably similar to those of glycogen phosphorylase [10].

*Address correspondence to this author at the Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI-CONICET), Universidad Nacional de Rosario, Suipacha 531, 2000, Rosario, Argentina; Tel: +54 341 437-1955; Fax: + 54 341 437-0044; E-mail: gomezcasati@cefobi-conicet.gov.ar

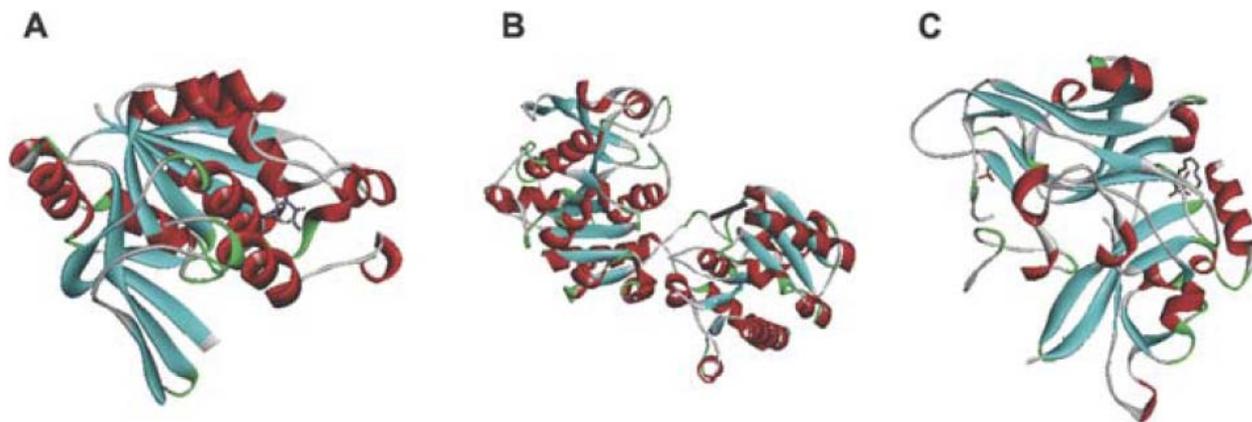


Figure 1. Ribbon structures showing the folds described for GTs: (A) GT-A fold represented by the SpsA protein, a nucleotide-diphosphosugar transferase from *Bacillus subtilis* (PDB ID 1H7L); (B) GT-B fold, represented by the glycogen synthase from *Agrobacterium tumefaciens* (PDB ID 1RZU); (C) GT-C fold, from EmbC, the arabinosyltransferase from *Mycobacterium tuberculosis* (PDB ID 3PTY). α helices are colored in red, whereas β strands are represented in cyan.

to those of glycogen phosphorylase [10]. Nevertheless, the overall fold of the protein does not dictate the stereochemistry of the reaction that it catalyzes. Coutinho *et al.* [11] described several examples where both, inverting and retaining glycosyltransferases belongs to the GT-A and GT-B fold classes.

A third fold was also described, the GT-C fold which is a predicted protein topology for integral membrane glycosyltransferases (Fig. 1) [12]. However, this fold was not experimentally verified until the characterization of the 3D structure of a soluble C-terminal domain of an oligosaccharyltransferase STT3 from *Pyrococcus furiosus* [13]. Recently it has also been described other GT-C type GT essential for the synthesis of arabinogalactan and lipoarabinogalactan polymers in *Mycobacterium tuberculosis*, the EmbC arabinosyltransferase (PDB ID 3PTY, Fig. 1C) [14]. The co-crystallization of EmbC with a di-arabinoside analogue suggest that this protein contains two different carbohydrate binding sites, associated with subdomains I and II, respectively. Furthermore, the subdomain II functions as a carbohydrate-binding module (CBM) [14].

To date, GTs have been classified into 94 distinct families based in sequence data, available at the Carbohydrate Active Enzymes database server (CAZy, <http://www.cazy.org/>) [15]. The last family, 94, was recently defined after the biochemical characterization of the GumI protein, a non-processive β -mannosyltransferase that uses GDPMan as glycosyl donor substrate, essential for xanthan polysaccharide synthesis [16].

However, GT families 36, 46 and 86 were recently extinguished. The report on the evolutionary, structural and mechanistic relationship of the chitobiose phosphorylase with GHs of the clan GH-L provided insights to merge of GT family 36 to the glycosyl hydrolase family 94 [17]. The lack of consistent data for the biochemical characterization of GT-46 members led to the extinction of such family while members of GT-86 were merged to GT-83 (CAZY, <http://www.cazy.org/>).

The first 3D structures described for GTs were those from a glucanotransferase, the cyclodextrin GT from *Bacillus circulans* (EC 2.4.1.19, PDB code 1CGT), in 1991 [18, 19]; and the β -GT from bacteriophage T4 mentioned above (EC 2.4.1.27, PDB code 1BGT) [9]. The β -GT belongs to the GT-63 family, whereas the cyclodextrin GT was classified into the glycoside hydrolase 13 family (GH-13) in the CAZY server. While only few structures of GTs were described about 20 years ago, at the present it have been reported the 3D structure of 124 proteins belonging to 40 different GT families (Table 1). In addition, there are 1457 structure hits in the protein data bank (<http://www.rcsb.org/>) corresponding to GTs (EC 2.4.x.x). Most of the 3D structures (about 1437) were solved by X-ray diffraction, whereas few structures were determined by solution NMR.

Carbohydrate Binding Modules

Most of the enzymes that degrade, modify, or create glycosidic bonds share a common modular organization. Such disposition usually includes a catalytic domain typical for each enzyme and one or more carbohydrate binding domain (CBM) connected by a loosely-structured chain. CBMs were initially classified as cellulose binding domains (CBDs), based on the initial discovery of several modules with the ability to bind this polysaccharide [20, 21]. Nonetheless, there are other types of modules different from these that can bind carbohydrates different from cellulose. These findings created the need for a new classification of these peptides using a more inclusive terminology. Thus, a CBM is nowadays defined as the contiguous amino acid sequence in an active enzyme involved in carbohydrate metabolism, with the ability to bind carbohydrates [22]. CBMs contain between 30 and 200 amino acids and may exist as simple, double or triple repeats in a protein (there are rare instances of independent putative CBMs). These domains can occupy N or C-terminal positions and occasionally can be found in the middle of the polypeptide chain.

CBMs have been found in glycoside hydrolases, glycosyltransferases, polysaccharide lyases, as well as carbohyd-

Table 1. Structures Reported for Polysaccharide-synthesizing Glycosyltransferases (Until September 2012*)

Family (total structures)	Protein	PDB code	Reference
GT-1 (22)	GtfA <i>Amycolatopsis orientalis</i>	1PN3	[51]
GT-1 (22)	GtfA <i>Amycolatopsis orientalis</i>	1PN3	[51]
GT-2 (6)	SpsA <i>Bacillus subtilis</i>	1H7L	[8]
GT-3 (1)	glycogen synthase 2 (Gsy2, <i>S. cerevisiae</i>)	3NB0	[52]
GT-4 (15)	sucrose synthase (Sus1, <i>A. thaliana</i>)	3S27	[53]
GT-5 (3)	glycogen synthase 1 (GlgA, <i>A. tumefaciens</i>)	1RZU	[10]
GT-6 (3)	α -1,3 galactosyltransferase (GGTA1, <i>Bos taurus</i>)	1FG5	[54]
GT-7 (3)	β -1,4 galactosyltransferase-7 (b4Galt-7, <i>D. melanogaster</i>)	3LW6	[55]
GT-8 (4)	glycogenin (Gyg1, Homo sapiens)	3Q4S	[56]
GT-9 (4)	ADP heptose LPS heptosyltransferase II (WaaF, <i>E. coli</i>)	1PSW	http://www.rcsb.org
GT-10 (1)	α -1,3 fucosyltransferase (FucT, <i>H. Pylori</i>)	2NZW	http://www.rcsb.org
GT-13 (1)	<i>N</i> -acetylglucosaminyltransferase I (GnT1, rabbit)	1FO8	[57]
GT-14 (1)	2 beta1,6- <i>N</i> -acetylglucosaminyltransferase (C2GnT-L, <i>M. musculus</i>)	2GAK	[58]
GT-15 (1)	α -1,2-mannosyltransferase (Kre2, <i>S. cerevisiae</i>)	1S4N	[59]
GT-20 (2)	α -trehalose-phosphate synthase (OtsA, <i>E. coli</i>)	1GZ5	[60]
GT-23 (2)	α -1,6- <i>L</i> -fucosyltransferase (NodZ, <i>Bradyrhizobium</i> sp.)	2HHC	[61]
GT-27 (3)	UDP-GalNAc:polypeptide α - <i>N</i> -acetyl galactosaminyltransferase-2 (<i>H. sapiens</i>)	2FFU	[28]
GT-28 (1)	MurG, <i>E. coli</i>	1F0K	[62]
GT-29 (1)	α -2,3-sialyltransferase (ST3Gal I, <i>Sus scrofa</i>)	2WML	[63]
GT-30 (1)	Membrane GT (WaaA, <i>A. aeolicus</i>)	2XCI	[64]
GT-31 (1)	<i>O</i> -fucosylpeptide β -1,3-GlcNAc transferase (<i>M. musculus</i>)	2J0A	[65]
GT-35 (6)	glycogen phpsphorylase (PygM, rabbit muscle)	1A8I	[66]
GT-41 (3)	UDP-Glc: protein <i>N</i> - β -glucosyltransferase (HMW1C, <i>Actinobacillus pleuropneumoniae</i>)	3Q3E	[67]
GT-42 (2)	α -2,3-sialyltransferase (Cst-I, <i>C. jejuni</i>)	2P2V	[68]
GT-43 (3)	beta1,3-glucuronyltransferase I (GlcAT-I, <i>H. sapiens</i>)	1FGG	[69]
GT-44 (4)	Toxin A (<i>Clostridium difficile</i>)	3SRZ	[70]
GT-51 (6)	Penicillin binding protein (<i>A. baumannii</i>)	3UDF	[71]
GT-52 (1)	β -galactosamide α -2,3-sialyltransferase (NST, <i>N. meningitidis</i>)	2YK4	[72]
GT-53 (1)	α -1,5-arabinosyltransferase C (EmbC, <i>M. tuberculosis</i>)	3PTY	[14]
GT-55 (2)	α -mannosyltransferase (MPG, <i>Pyrococcus horikoshii</i>)	2ZU7	http://www.rcsb.org
GT-63 (1)	DNA β -glucosyltransferase (BGT, Bacteriophage T4)	1BGT	[9]
GT-64 (1)	α - <i>N</i> -acetylhexosaminyltransferase (ExtI2, <i>M. musculus</i>)	1OMX	[73]
GT-65 (1)	GDP-Fuc: protein <i>O</i> -fucosyltransferase 1 (PoFut1, <i>C. elegans</i>)	3ZY2	[74]
GT-66 (4)	oligosaccharyltransferase (AglB, <i>Archaeoglobus fulgidus</i>)	3VGP	[75]
GT-68 (1)	GDP-L-Fucose: protein <i>O</i> -fucosyltransferase 2 (PoFUT2, <i>H. sapiens</i>)	4AP5	[76]
GT-70 (1)	β -1,2-glucuronosyltransferase (GumK, <i>Xanthomonas campestris</i>)	2HY7	[77]
GT-72 (1)	DNA α -glucosyltransferase (AGT, Bacteriophage T4)	1XV5	[78]

(Table 1) Comparison...

Family (total structures)	Protein	PDB code	Reference
GT-78 (1)	Mannosylglycerate synthase (Mgs, <i>Rhodothermus marinus</i>)	2BO4	[79]
GT-80 (4)	α -2,3/2,6-sialyltransferase 1 (St1, <i>Pasteurella multocida</i>)	2IHJ	[80]
GT-81 (3)	glucosyl-3-phosphoglycerate synthase (GpgS, <i>M. tuberculosis</i>)	3E25	[81]
GT-88 (2)	elongation factor 1A α -glucosyltransferase (Lgt1, <i>Legionella pneumophila</i>)	2WZG	[82]

*One representative member of each family is described

rate esterases. The CBMs can increase the rate of catalysis by approximating the catalytic domain of the protein to its substrate or act as a scaffolding domain that organizes the catalytic subunits into a cohesive multi-enzymatic complex [22]. Thus, they have three general roles with respect to the function of their cognate catalytic modules: (i) a proximity effect, (ii) a targeting function and (iii) a disruptive function [23]. Currently, more than 300 putative sequences in over 50 different species have been identified and the binding domains have been classified in 64 families, based on their amino acid sequence, substrate binding specificity and structure at the CAZY server (<http://www.cazy.org/>) [22-26].

It is important to note that several 3-D structures of members from at least 23 representative families have been obtained, in some cases as protein-ligand complexes. Data from these structures shed light on the underlying mechanism for each interaction between a CBM and its substrate. CBMs from different families are structurally similar and their ability to bind carbohydrates can be attributed, at least partially, to several aromatic residues which assemble into a hydrophobic surface [22].

It is interesting to note that glycosyltransferases, the main topic of this review, are represented in 4 of the 64 CBM families already mentioned and are listed in Table 2. There are several polypeptide N-acetylgalactosaminyltransferases in the CBM13 family from different organisms including *C. elegans*, *B. Taurus*, *D. melanogaster*, *T. gondii* and *H. sapiens*. These GTs belong to GT-2 family and catalyze the transfer of an N-acetyl-D-galactosamine residue to a serine or threonine residue to glycosylate the receptor protein [27]. The x-ray crystal structure of the human enzyme (ppGalNAcT-2 (hT2)) is available and shows the dynamic association between the catalytic and lectin domains where CBMs are located [28].

The CBM20 family is one of the protein families where the starch binding function of these modules has been demonstrated in several cases (<http://www.cazy.org/>). This family include different GTs from Archea, such as an α -cyclomaltodextrin glycosyltransferase [29]; several β -cyclodextrin glycosyltransferases from bacteria, which cyclizes part of the α -1,4 glucan chain [30, 31] and two glucanotransferases, essential for maltose metabolism in plants and probably, playing a role in freezing tolerance [32].

CBM48 family is represented by GT from bacteria and eucaryota. All examples involve branching enzymes that catalyzes the formation of α -1,6-glucosidic bonds in bacterial glycogen and plant starch as well as in the synthesis of glycogen in mammals. One of the most characterized mem-

bers of this family is the starch branching enzyme I from *O. sativa* which its 3D structure was recently solved [33] (see also Table 2).

The last family is the CBM53 family, which currently has about 20 characterized entries in the CAZY database. All these entries correspond to starch synthase III (SSIII) belonging to higher plant species (*Arabidopsis thaliana*, *Zea mays*, *Solanum tuberosum*, *Phaseolus vulgaris*, *Vigna unguiculata* and *Oriza sativa*) and *Chlamydomonas reinhardtii*. It has been described that SSIII proteins contains three in tandem N-terminal Starch Binding Domains (SBDs) [34, 35].

The SBDs have acquired the evolutionary advantage to break the structure of their substrate when compared to the CBD, due to the presence of two binding sites for the polysaccharide [36]. So far, no CBD structure containing such binding sites has been reported. In view of this, such domains are unable to break down the polysaccharide structure as efficiently as the SBD does [36, 37]. Thus, SBDs are involved in the attachment of the catalytic domain to its polymeric substrate as well as in the disruption of the surface of the starch granule [38, 39].

Currently, SBDs are categorized into the ten following CBM families: (i) CBM20, i.e., the C-terminal SBD from *Aspergillus niger* glucoamylase; (ii) CBM21, located at the N-terminal domain in amylases; (iii) CBM25, containing one (i.e., β -amylase from *B. circulans*) or two (i.e., *Bacillus* sp. α -amylase) modules; (iv) CBM26, mostly organized in tandem repeats (i.e., C-terminal domains from *Lactobacillus manihottivorans* α -amylase); (v) CBM34, present in the N-terminal domains of neopullulanase, maltogenic amylase, and cyclomaltodextrinase; (vi) CBM41, N-terminal SBDs, present mostly in bacterial pullulanases; (vii) CBM45, originating from eukaryotic proteins from the plant kingdom (i.e., N-terminal modules of α -amylases and α -glucan water dikinases); (viii) CBM48, modules with glycogen-binding function (including SBD from the GH13 pullulanase and regulatory domains of mammalian AMP-activated protein kinase); (ix) CBM53, SBD modules from SSIII; and, (x) CBM58, represented by *Bacteroides thetaiotaomicron* SusG protein that has been shown to bind maltoheptose [15, 24, 25, 40]. However, CBM20 and CBM21 families are proposed to form a common clan [23]. Analyses of three-dimensional structures from many of these families showed a characteristic SBD fold of a distorted β -barrel composed of several β -strands and with one or, more frequently, two distinct binding sites that exhibit site-dependent modes of carbohydrate binding [25, 26, 41-43].

quence identity among the enzymes from different families, all enzymes remarkably converge in only three folds named GT-A, GT-B and GT-C. Hence, other properties must contribute to the huge functional plasticity of GTs; such as a large structural variability in the acceptor binding domain, the presence of CBMs, that could be involved in the modulation of the GT catalytic activity, as well as the possibility of supramolecular organization of GTs within the cell.

ACKNOWLEDGEMENTS

This work was supported by grants from Biotechnology Program from Universidad Nacional de General San Martín (UNSAM) (PROG07F/2-2007), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, PIP 00237) and Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 2010 – 0543 and PICT 2010 – 0069). MVB, MM and DGC are research members from CONICET.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ABBREVIATIONS

GT	=	(glycosyltransferase)
CBM	=	(carbohydrate binding module)
SBD	=	(starch binding domain)
SSIII	=	(starch synthase III)

REFERENCES

- Lairson, L.L.; Henrissat, B.; Davies, G.J.; Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.*, **2008**, *77*, 521-555.
- Caputto, R.; Leloir, L.F.; Cardini, C.E.; Paladini, A.C. Isolation of the coenzyme of the galactose phosphate-glucose phosphate transformation. *J. Biol. Chem.*, **1950**, *184*, 333-350.
- Sinnott, M.L. Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev.*, **1990**, *90*, 1171-1202.
- Breton, C.; Mucha, J.; Jeanneau, C. Structural and functional features of glycosyltransferases. *Biochimie*, **2001**, *83*, 713-718.
- Unligil, U.M.; Rini, J.M. Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.*, **2000**, *10*, 510-517.
- Bourme, Y.; Henrissat, B. Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.*, **2001**, *11*, 593-600.
- Charnock, S.J.; Davies, G.J. Structure of the nucleotide-diphosphogluco transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry*, **1999**, *38*, 6380-6385.
- Tarbouriech, N.; Charnock, S.J.; Davies, G.J. Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J. Mol. Biol.*, **2001**, *314*, 655-661.
- Vrielink, A.; Ruger, W.; Driessen, H.P.; Freemont, P.S. Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.*, **1994**, *13*, 3413-3422.
- Buschiazio, A.; Ugalde, J.E.; Guerin, M.E.; Shepard, W.; Ugalde, R.A.; Alzari, P.M. Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *EMBO J.*, **2004**, *23*, 3196-3205.
- Coutinho, P.M.; Deleury, E.; Davies, G.J.; Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.*, **2003**, *328*, 307-317.
- Liu, J.; Mushegian, A. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.*, **2003**, *12*, 1418-1431.
- Igura, M.; Maita, N.; Kamishikiryu, J.; Yamada, M.; Obita, T.; Maenaka, K.; Kohda, D. Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.*, **2008**, *27*, 234-243.
- Alderwick, L.J.; Lloyd, G.S.; Ghadbane, H.; May, J.W.; Bhatt, A.; Eggeling, L.; Futterer, K.; Besra, G.S. The C-terminal domain of the Arabinosyltransferase *Mycobacterium tuberculosis* EmbC is a lectin-like carbohydrate binding module. *PLoS Pathog.*, **2011**, *7*, e1001299.
- Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.*, **2009**, *37*, D233-238.
- Salinas, S.R.; Bianco, M.I.; Barreras, M.; Ielpi, L. Expression, purification and biochemical characterization of GumI, a monotopic membrane GDP-mannose:glycolipid 4-β-D-mannosyltransferase from *Xanthomonas campestris* pv. *campestris*. *Glycobiology*, **2011**, *21*, 903-913.
- Hidaka, M.; Honda, Y.; Kitaoka, M.; Nirasawa, S.; Hayashi, K.; Wakagi, T.; Shoun, H.; Fushinobu, S. Chitobiose phosphorylase from *Vibrio proteolyticus*, a member of glycosyl transferase family 36, has a clan GH-L-like (α/α) barrel fold. *Structure*, **2004**, *12*, 937-947.
- Klein, C.; Schulz, G.E. Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution. *J. Mol. Biol.*, **1991**, *217*, 737-750.
- Klein, C.; Hollender, J.; Bender, H.; Schulz, G.E. Catalytic center of cyclodextrin glycosyltransferase derived from X-ray structure analysis combined with site-directed mutagenesis. *Biochemistry*, **1992**, *31*, 8740-8746.
- Tomme, P.; Van Tilbeurgh, H.; Pettersson, G.; Van Damme, J.; Vandekerckhove, J.; Knowles, J.; Teeri, T.; Claeysens, M. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur. J. Biochem.*, **1988**, *170*, 575-581.
- Gilkes, N.R.; Warren, R.A.; Miller, Jr. R.C.; Kilburn, D.G. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J. Biol. Chem.*, **1988**, *263*, 10401-10407.
- Shoseyov, O.; Shani, Z.; Levy, I. Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 283-295.
- Machovic, M.; Svensson, B.; MacGregor, E.A.; Janecek, S. A new clan of CBM families based on bioinformatics of starch-binding domains from families CBM20 and CBM21. *FEBS J.*, **2005**, *272*, 5497-54513.
- Machovic, M.; Janecek, S. Starch-binding domains in the post-genome era. *Cell Mol. Life Sci.*, **2006**, *63*, 2710-2724.
- Machovic, M.; Janecek, S. The evolution of putative starch-binding domains. *FEBS Lett.*, **2006**, *580*, 6349-6356.
- Boraston, A.B.; Bolam, D.N.; Gilbert, H.J.; Davies, G.J. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.*, **2004**, *382*, 769-781.
- Hagen, F.K.; Nehrke, K. cDNA cloning and expression of a family of UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase sequence homologs from *Caenorhabditis elegans*. *J. Biol. Chem.*, **1998**, *273*, 8268-8277.
- Fritz, T.A.; Raman, J.; Tabak, L.A. Dynamic association between the catalytic and lectin domains of human UDP-GalNAc: polypeptide alpha-N-acetylgalactosaminyltransferase-2. *J. Biol. Chem.*, **2006**, *281*, 8613-8619.
- Rashid, N.; Cornista, J.; Ezaki, S.; Fukui, T.; Atomi, H.; Imanaka, T. Characterization of an archaeal cyclodextrin glucanotransferase with a novel C-terminal domain. *J. Bacteriol.*, **2002**, *184*, 777-784.
- Penninga, D.; van der Veen, B.A.; Knegetel, R.M.; van Hijum, S.A.; Rozeboom, H.J.; Kalk, K.H.; Dijkstra, B.W.; Dijkhuizen, L. The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J. Biol. Chem.*, **1996**, *271*, 32777-32784.
- Costa, H.; del Canto, S.; Ferrarotti, S.; de Jimenez Bonino, M.B. Structure-function relationship in cyclodextrin glycosyltransferase from *Bacillus circulans* DF 9R. *Carbohydr. Res.*, **2009**, *344*, 74-79.

- [32] Chia, T.; Thorncroft, D.; Chapple, A.; Messerli, G.; Chen, J.; Zeeman, S.C.; Smith, S.M.; Smith, A.M. A cytosolic glucosyltransferase is required for conversion of starch to sucrose in Arabidopsis leaves at night. *Plant J.*, **2004**, *37*, 853-863.
- [33] Noguchi, J.; Chaen, K.; Vu, N. T.; Akasaka, T.; Shimada, H.; Nakashima, T.; Nishi, A.; Satoh, H.; Omori, T.; Kakuta, Y.; Kimura, M. Crystal structure of the branching enzyme I (BEI) from *Oryza sativa* L with implications for catalysis and substrate binding. *Glycobiology*, **2011**, *21*, 1108-1116.
- [34] Palopoli, N.; Busi, M.V.; Fornasari, M.S.; Gomez-Casati, D.; Ugalde, R.; Parisi, G. Starch-synthase III family encodes a tandem of three starch-binding domains. *Proteins*, **2006**, *65*, 27-31.
- [35] Valdez, H.A.; Busi, M.V.; Wayllace, N.Z.; Parisi, G.; Ugalde, R.A.; Gomez-Casati, D.F. Role of the N-terminal starch-binding domains in the kinetic properties of starch synthase III from *Arabidopsis thaliana*. *Biochemistry*, **2008**, *47*, 3026-3032.
- [36] Southall, S.M.; Simpson, P.J.; Gilbert, H.J.; Williamson, G.; Williamson, M.P. The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS Lett.*, **1999**, *447*, 58-60.
- [37] Tormo, J.; Lamed, R.; Chirino, A.J.; Morag, E.; Bayer, E.A.; Shoham, Y.; Steitz, T.A. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.*, **1996**, *15*, 5739-5751.
- [38] Rodriguez-Sanoja, R.; Oviedo, N.; Sanchez, S. Microbial starch-binding domain. *Curr. Opin. Microbiol.*, **2005**, *8*, 260-267.
- [39] Sorimachi, K.; Le Gal-Coeffet, M.F.; Williamson, G.; Archer, D.B.; Williamson, M.P. Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to beta-cyclodextrin. *Structure*, **1997**, *5*, 647-661.
- [40] Janecek, S.; Svensson, B.; MacGregor, E.A. Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb. Technol.*, **2011**, *49*, 429-440.
- [41] Lawson, C.L.; van Montfort, R.; Strokopytov, B.; Rozeboom, H.J.; Kalk, K.H.; de Vries, G.E.; Penninga, D.; Dijkhuizen, L.; Dijkstra, B.W. Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a maltose-dependent crystal form. *J. Mol. Biol.*, **1994**, *236*, 590-600.
- [42] Hashimoto, H. Recent structural studies of carbohydrate-binding modules. *Cell Mol. Life Sci.*, **2006**, *63*, 2954-2967.
- [43] Guillen, D.; Sanchez, S.; Rodriguez-Sanoja, R. Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.*, **2010**, *85*, 1241-1249.
- [44] Zhang, X.; Myers, A.M.; James, M.G. Mutations affecting starch synthase III in Arabidopsis alter leaf starch structure and increase the rate of starch synthesis. *Plant Physiol.*, **2005**, *138*, 663-674.
- [45] Zeeman, S.C.; Kossmann, J.; Smith, A.M. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu. Rev. Plant Biol.*, **2010**, *61*, 209-234.
- [46] Szydowski, N.; Ragel, P.; Raynaud, S.; Lucas, M.M.; Roldan, I.; Montero, M.; Munoz, F.J.; Ovecka, M.; Bahaji, A.; Planchot, V.; Pozueta-Romero, J.; D'Hulst, C.; Merida, A. Starch granule initiation in Arabidopsis requires the presence of either class IV or class III starch synthases. *Plant Cell*, **2009**, *21*, 2443-2457.
- [47] Li, Z.; Mouille, G.; Kosar-Hashemi, B.; Rahman, S.; Clarke, B.; Gale, K.R.; Appels, R.; Morell, M.K. The structure and expression of the wheat starch synthase III gene. Motifs in the expressed gene define the lineage of the starch synthase III gene family. *Plant Physiol.*, **2000**, *123*, 613-624.
- [48] Busi, M.V.; Palopoli, N.; Valdez, H.A.; Fornasari, M.S.; Wayllace, N.Z.; Gomez-Casati, D.F.; Parisi, G.; Ugalde, R.A. Functional and structural characterization of the catalytic domain of the starch synthase III from *Arabidopsis thaliana*. *Proteins*, **2008**, *70*, 31-40.
- [49] Wayllace, N.Z.; Valdez, H.A.; Ugalde, R.A.; Busi, M.V.; Gomez-Casati, D.F. The starch-binding capacity of the noncatalytic SBD2 region and the interaction between the N- and C-terminal domains are involved in the modulation of the activity of starch synthase III from *Arabidopsis thaliana*. *FEBS J.*, **2010**, *277*, 428-440.
- [50] Valdez, H.A.; Peralta, D.A.; Wayllace, N.Z.; Grisolia, M.J.; Gomez-Casati, D.F.; Busi, M.V. Preferential binding of SBD from *Arabidopsis thaliana* SSIII to polysaccharides. Study of amino acid residues involved. *Starch/Stärke*, **2011**, *63*, 451-460.
- [51] Mulichak, A.M.; Losey, H.C.; Lu, W.; Wawrzak, Z.; Walsh, C.T.; Garavito, R.M. Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*, 9238-9243.
- [52] Baskaran, S.; Roach, P.J.; DePaoli-Roach, A.A.; Hurley, T.D. Structural basis for glucose-6-phosphate activation of glycogen synthase. *Proc. Natl. Acad. Sci. USA*, **2010**, *107*, 17563-17568.
- [53] Zheng, Y.; Anderson, S.; Zhang, Y.; Garavito, R.M. The structure of sucrose synthase-1 from *Arabidopsis thaliana* and its functional implications. *J. Biol. Chem.*, **2011**, *286*, 36108-36118.
- [54] Gastinel, L.N.; Bignon, C.; Misra, A.K.; Hindsgaul, O.; Shaper, J.H.; Joziase, D.H. Bovine α 1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. *EMBO J.*, **2001**, *20*, 638-649.
- [55] Ramakrishnan, B.; Qasba, P.K. Crystal structure of the catalytic domain of Drosophila β 1,4-Galactosyltransferase-7. *J. Biol. Chem.*, **2010**, *285*, 15619-15626.
- [56] Chaikuad, A.; Froese, D.S.; Berridge, G.; von Delft, F.; Oppermann, U.; Yue, W.W. Conformational plasticity of glycogenin and its maltosaccharide substrate during glycogen biogenesis. *Proc. Natl. Acad. Sci. USA*, **2011**, *108*, 21028-21033.
- [57] Unligil, U.M.; Zhou, S.; Yuwaraj, S.; Sarkar, M.; Schachter, H.; Rini, J.M. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J.*, **2000**, *19*, 5269-5280.
- [58] Pak, J.E.; Armoux, P.; Zhou, S.; Sivarajah, P.; Satkunarajah, M.; Xing, X.; Rini, J.M. X-ray crystal structure of leukocyte type core 2 β 1,6-N-acetylglucosaminyltransferase. Evidence for a convergence of metal ion-independent glycosyltransferase mechanism. *J. Biol. Chem.*, **2006**, *281*, 26693-26701.
- [59] Lobsanov, Y.D.; Romero, P.A.; Sleno, B.; Yu, B.; Yip, P.; Herscovics, A.; Howell, P.L. Structure of Kre2p/Mnt1p: a yeast α 1,2-mannosyltransferase involved in mannoprotein biosynthesis. *J. Biol. Chem.*, **2004**, *279*, 17921-17931.
- [60] Gibson, R.P.; Turkenburg, J.P.; Charnock, S.J.; Lloyd, R.; Davies, G.J. Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chem. Biol.*, **2002**, *9*, 1337-1346.
- [61] Brzezinski, K.; Stepkowski, T.; Panjkar, S.; Bujacz, G.; Jaskolski, M. High-resolution structure of NodZ fucosyltransferase involved in the biosynthesis of the nodulation factor. *Acta Biochim. Pol.*, **2007**, *54*, 537-549.
- [62] Ha, S.; Walker, D.; Shi, Y.; Walker, S. The 1.9 Å crystal structure of Escherichia coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Sci.*, **2000**, *9*, 1045-1052.
- [63] Rao, F.V.; Rich, J.R.; Rakic, B.; Buddai, S.; Schwartz, M.F.; Johnson, K.; Bowe, C.; Wakarchuk, W.W.; Defrees, S.; Withers, S.G.; Strynadka, N.C. Structural insight into mammalian sialyltransferases. *Nat. Struct. Mol. Biol.*, **2009**, *16*, 1186-1188.
- [64] Schmidt, H.; Hansen, G.; Singh, S.; Hanuszkiewicz, A.; Lindner, B.; Fukase, K.; Woodard, R.W.; Holst, O.; Hilgenfeld, R.; Mamat, U.; Mesters, J.R. Structural and mechanistic analysis of the membrane-embedded glycosyltransferase WaaA required for lipopolysaccharide synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **2012**, *109*, 6253-6258.
- [65] Jinek, M.; Chen, Y.W.; Clausen, H.; Cohen, S.M.; Conti, E. Structural insights into the Notch-modifying glycosyltransferase Fringe. *Nat. Struct. Mol. Biol.*, **2006**, *13*, 945-956.
- [66] Gregoriou, M.; Noble, M.E.; Watson, K.A.; Garman, E.F.; Krulle, T.M.; de la Fuente, C.; Fleet, G.W.; Oikonomakos, N.G.; Johnson, L.N. The structure of a glycogen phosphorylase glucopyranose spirohydantoin complex at 1.8 Å resolution and 100 K: the role of the water structure and its contribution to binding. *Protein Sci.*, **1998**, *7*, 915-927.
- [67] Kawai, F.; Grass, S.; Kim, Y.; Choi, K.J.; St Geme, J.W.; Yeo, H.J. Structural insights into the glycosyltransferase activity of the *Actinobacillus pleuropneumoniae* HMWIC-like protein. *J. Biol. Chem.*, **2011**, *286*, 38546-38557.
- [68] Chiu, C.P.; Lairson, L.L.; Gilbert, M.; Wakarchuk, W.W.; Withers, S.G.; Strynadka, N.C. Structural analysis of the α -2,3-sialyltransferase Cst-I from *Campylobacter jejuni* in apo and substrate-analogue bound forms. *Biochemistry*, **2007**, *46*, 7196-7204.
- [69] Pedersen, L.C.; Tsuchida, K.; Kitagawa, H.; Sugahara, K.; Darden, T.A.; Negishi, M. Heparan/chondroitin sulfate biosynthesis. Structure and mechanism of human glucuronyltransferase I. *J. Biol. Chem.*, **2000**, *275*, 34580-34585.
- [70] Pruitt, R.N.; Chumbler, N.M.; Rutherford, S.A.; Farrow, M.A.; Friedman, D.B.; Spiller, B. and Lacy, D.B. Structural determinants

- of *Clostridium difficile* toxin A glucosyltransferase activity. *J. Biol. Chem.*, **2012**, 287, 8013-8020.
- [71] Han, S.; Caspers, N.; Zaniewski, R.P.; Lacey, B.M.; Tomaras, A.P.; Feng, X.; Geoghegan, K.F.; Shanmugasundaram, V. Distinctive attributes of beta-lactam target proteins in *Acinetobacter baumannii* relevant to development of new antibiotics. *J. Am. Chem. Soc.*, **2011**, 133, 20536-20545.
- [72] Lin, L.Y.; Rakic, B.; Chiu, C.P.; Lameignere, E.; Wakarchuk, W.W.; Withers, S.G.; Strynadka, N.C. Structure and mechanism of the lipooligosaccharide sialyltransferase from *Neisseria meningitidis*. *J. Biol. Chem.*, **2011**, 286, 37237-37248.
- [73] Pedersen, L.C.; Dong, J.; Taniguchi, F.; Kitagawa, H.; Krahn, J.M.; Pedersen, L.G.; Sugahara, K.; Negishi, M. Crystal structure of an alpha 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. *J. Biol. Chem.*, **2003**, 278, 14420-14428.
- [74] Lira-Navarrete, E.; Valero-Gonzalez, J.; Villanueva, R.; Martinez-Julvez, M.; Tejero, T.; Merino, P.; Panjikar, S.; Hurtado-Guerrero, R. Structural insights into the mechanism of protein O-fucosylation. *PLoS ONE*, **2011**, 6, e25365.
- [75] Matsumoto, S.; Igura, M.; Nyirenda, J.; Matsumoto, M.; Yuzawa, S.; Noda, N.; Inagaki, F.; Kohda, D. Crystal structure of the C-terminal globular domain of oligosaccharyltransferase from *Archaeoglobus fulgidus* at 1.75 Å resolution. *Biochemistry*, **2012**, 51, 4157-4166.
- [76] Chen, C.I.; Keusch, J.J.; Klein, D.; Hess, D.; Hofsteenge, J.; Gut, H. Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation. *EMBO J.*, **2012**, 31, 3183-3197.
- [77] Barreras, M.; Salinas, S.R.; Abdian, P.L.; Kampel, M.A.; Ielpi, L. Structure and mechanism of GumK, a membrane-associated glucuronosyltransferase. *J. Biol. Chem.*, **2008**, 283, 25027-25035.
- [78] Lariviere, L.; Sommer, N.; Morera, S. Structural evidence of a passive base-flipping mechanism for AGT, an unusual GT-B glycosyltransferase. *J. Mol. Biol.*, **2005**, 352, 139-150.
- [79] Flint, J.; Taylor, E.; Yang, M.; Bolam, D.N.; Tailford, L.E.; Martinez-Fleites, C.; Dodson, E.J.; Davis, B.G.; Gilbert, H.J.; Davies, G.J. Structural dissection and high-throughput screening of mannosylglycerate synthase. *Nat. Struct. Mol. Biol.*, **2005**, 12, 608-614.
- [80] Ni, L.; Chokhawala, H.A.; Cao, H.; Henning, R.; Ng, L.; Huang, S.; Yu, H.; Chen, X.; Fisher, A.J. Crystal structures of *Pasteurella multocida* sialyltransferase complexes with acceptor and donor analogues reveal substrate binding sites and catalytic mechanism. *Biochemistry*, **2007**, 46, 6288-6298.
- [81] Pereira, P.J.; Empadinhas, N.; Albuquerque, L.; Sa-Moura, B.; da Costa, M.S.; Macedo-Ribeiro, S. *Mycobacterium tuberculosis* glucosyl-3-phosphoglycerate synthase: structure of a key enzyme in methylglucose lipopolysaccharide biosynthesis. *PLoS ONE*, **2008**, 3, e3748.
- [82] Hurtado-Guerrero, R.; Zusman, T.; Pathak, S.; Ibrahim, A.F.; Shepherd, S.; Prescott, A.; Segal, G.; van Aalten, D.M. Molecular mechanism of elongation factor 1A inhibition by a *Legionella pneumophila* glycosyltransferase. *Biochem. J.*, **2010**, 426, 281-292.