Published in final edited form as:

Curr Opin Biotechnol. 2011 December ; 22(6): 800–808. doi:10.1016/j.copbio.2011.04.013.

Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation

Aram Chang¹, Shanteri Singh², George N. Phillips Jr.¹, and Jon S. Thorson^{2,*}

¹Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

²Pharmaceutical Sciences Division, School of Pharmacy, Center for Natural Products Research, University of Wisconsin-Madison, 777 Highland Avenue, Madison, Wisconsin 53705-2222, USA

Abstract

Glycosyltransferases (GTs) are ubiquitous in nature and are required for the transfer of sugars to a variety of important biomolecules. This essential enzyme family has been a focus of attention from both the perspective of a potential drug target as well as a catalyst for the development of vaccines, biopharmaceuticals and small molecule therapeutics. This review attempts to consolidate the emerging lessons from Leloir (nucleotide-dependent) GT structural biology studies and recent applications of these fundamentals toward rational engineering of glycosylation catalysts.

I. Introduction

Glycosyltransferases (GTs) mediate the regio- and stereospecific transfer of sugars from various sugar donors to a variety of important biomolecules including glycans, lipids, peptides, and small molecules [1–3]. Understanding GT function, mechanism and structure has contributed to both the development of GT inhibitors as drugs to target pathogens and human disease, as well as the use of GTs to glycodiversify molecules in the search of drug leads [4–7]. Great advances toward understanding GT genomics, catalysis and structure have occurred over the last decade [1,3,7–10]. This review emphasizes the recent advances in Leloir (nucleotide-dependent) GT structural biology and its impact on delineating and engineering GT reaction specificity.

II. Global folds

Currently, there are over 65,000 GT sequences within the Carbohydrate-Active enZyme database (CAZy, <http://www.cazy.org/GlycosylTransferases.html>) comprising 89 families based upon sequence similarity [11]. To date, there exist 105 GT structures in the pdb (GT-A: 41, GT-B: 57, GT-C: 2, unclassified: 6), representing 36 of the 89 CAZy GT families, most of which adopt one of two predominate structural folds (GT-A or GT-B). Sometimes described as two closely abutting domains [1], the GT-A fold is a single domain fold composed of a seven stranded β -sheet core flanked by α -helices and a small antiparallel β -sheet ($\beta 4'$ and $\beta 7'$) bridged via a signature metal-coordinating DXD motif containing loop L4 (Figures 1a and 1c) [12]. The GT-B fold is comprised of two distinct N-terminal and C-

© 2011 Elsevier Ltd. All rights reserved.

*To whom correspondence should be addressed. jsthorson@pharmacy.wisc.edu; phillips@biochem.wisc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

terminal Rossmann-like domains of six or seven parallel β -sheet linked to α -helices, connected by a linker region and an interdomain cleft [1] (Figures 1b and 1d). Both the common GT-A or GT-B core can be further amended with additional domains for oligomerization [13] or membrane insertion [14–16]. In some cases, there is domain conservation within a species as exemplified by PSPG domain predominant in plant GTs [17]. Multivariate sequence analysis also predicts the existence of both GT-C and GT-D folds [18] and two recent structures have emerged of putative GT-C members - STT3 oligosaccharyltransferase and PglB [19,20].

The binding of metal ion and sugar donor to the GT-A fold is accompanied by local conformational changes of a flexible loop L3, (Figures 1a and 1c) as exemplified by loop rearrangements observed in human GTA and GTB [21,22], followed by additional flexible loop movement (loop L7') upon the acceptor binding [12,23]. In contrast, GT-B folds often exhibit a global domain movement upon substrate binding, differing in the type and degree of motion in a catalyst-specific manner. This open to closed conformational transition (~ 10 to 12° shift) [24], typically brings acceptor and donor into close proximity and is also accompanied by several loop movements (N3, N5, and C1) (Figures 1b and 1d). The mycothiol GT MshA, which shows a $\sim 97^\circ$ rotational reorientation of the N-terminal domain relative to the C-terminal domain [25], stands as a notable GT-B exception in this regard.

III. Sugar nucleotide binding

The GT-A fold sugar donor binding site is consistently found in the central β -sheet region that also contains the DXD signature in L4 (Figures 1a and 1c). In both folds, hydrogen-bonding and π -stacking interactions predominate in base recognition, as recently highlighted in the UDP-Gal-utilizing β -1,4-Gal-T7 (GT-A) [26] and the macrolide GT OleD (GT-B) [24]. While the residue at the end of β 1 impacts GT-A nucleotide specificity via hydrogen bonding (Figures 1a and 1c), the conformation and length of the C3 loop (Figures 1b and 1d) stands as a key GT-B determinant for purine versus pyrimidine binding [27,28]. In some cases, π -stacking can be substituted by hydrophobic stacking interactions as recently illustrated in the chondroitin polymerase (GT-A, PDB ID: 2Z86) [29] and phosphatidylinositol mannosyltransferase (GT-B, PDB ID: 2GEJ) [27]. Binding of the core sugar nucleotide diphosphate is enabled by the DXD-coordinated cation (GT-A) or a positive charged residue such as His (OleD [24], VvGT1 [30]), Lys (WsaF [31]) or Arg (PimA [27], PimB [32]) (GT-B). While the residue directly adjacent to the strictly conserved E/D-Q (GT-B loop C5, Figures 1b and 1d) has been implicated to influence GT-B sugar C2-C4 specificity (e.g., Glu352/Gln353 in OleI, Asp330/Gln331 in OleD [24], Asp398/Gln399 in glucuronosyltransferase, 2B7 [15]), the GT-A fold lacks sugar 'signatures'. In both folds, interactions with the donor sugar are often minimal and dominated by simple steric occlusion. Consistent with this, GTs are often sugar indiscriminant [4,33] and can even transfer non-carbohydrates such as glycerol (TagF) [34].

IV. Acceptor binding and regiospecificity

Given the vast structural diversity of acceptors, it is not surprising that the sequences and structures of the acceptor binding regions are divergent. In the GT-A fold, the central β -sheet core of the C-terminus is topologically conserved, but significant variation in C-terminal β -strands (α/β) (GnT-I [35]), α -helices (LgtC [36] and Lgt1 [23]) and/or loops (GTA, GTB [37], β -1,4-Gal-T1 [12] and β -1,4-Gal-T7 [26]) contribute to acceptor specificity and regiospecificity. Rearrangements of the flexible loop L7' (Figures 1a and 1c), induced by metal ion and donor binding, completes the formation of acceptor molecule binding site [12].

The N-terminal domain is an important feature in GT-B acceptor binding. Loops N3 and N5 (Figures 1b and 1d), which often undergo a disordered to ordered transition in the presence of ligand, and the neighboring N3 and N5 α -helices form an internal hydrophobic cavity (OleI, OleD [24], GtfD [38]) or solvent exposed pocket (GtfA [39]). In contrast, plant flavone GT (VvGT1 and UGT78G1 [30,40]) and the recent CalG1 and CalG3 structures (Figure 2) reveal a more equal engagement of both domains in acceptor coordination (N3 helices and C1, C5 loops) in a unique ‘bi-domain’ binding event. A comparison of the ‘bi-domain’ calicheamicin binding GTs (CalG1/CalG3) to the N-terminal domain calicheamicin-binding GTs (CalG2/CalG4), reveals the presence or absence of a cavity between the N3 and N5 α -helices as the key determinant for single or bi-domain binding, respectively. In addition to impacting acceptor specificity, these distinctions also influence reaction regiospecificity (e.g., the vancomycin GtfA and GtfD [38,39]; the calicheamicin CalG1-G4, Figure 2) and, to a lesser extent, sugar nucleotide specificity (e.g., OleD versus OleI).

V. Stereospecificity and nucleophile specificity

GT reactions are also classified based upon inversion or retention of anomeric configuration (Figure 3a) but it is important to note that the given GT structural fold does not correlate to reaction stereospecificity. For GT-A-catalyzed inversion, a conserved catalytic Asp or Glu (at the end of α_6 , Figures 1a and 1c) mediates the S_N2 direct displacement. In the GT-B fold, a conserved N α 1 catalytic residue (Asp [38,39] or His [24,30], Figures 1b and 1d, Figure 3b) located in the interdomain cleft mediates the displacement reaction. The mechanism for retention remains somewhat controversial [1] and support for both a double displacement [41,42] or an S_Ni -like mechanism [43] exists. From structural studies, the lack of suitably positioned nucleophiles for double displacement in retaining GT-A (Lgt1 [24]) or GT-B folds (WsaF [31], MshA [25], PimA [27], and PimB [32]) also lends indirect support for the S_Ni -like mechanism. The residues on the β -face of donor molecule binding site have been implicated in substrate recognition, orientation, and reaction stereospecificity in both retaining GT-A and GT-B folds, but general structural signatures are lacking [1].

GTs glycosylate a variety of acceptor nucleophiles including oxygen, nitrogen, sulfur and carbon and, with the exception of perhaps those involved in C-glycosylation, most GTs contain the fundamental machinery for heteroatom (*O*-, *N*-, and *S*-) glycosylation (Figure 3b). In most, if not all, cases where a perceived nucleophile specificity has been reported, the reactivity of the nucleophile (not that of the GT catalytic machinery or structural fold) is the predominant factor influencing observed differences in substrate turnover. Consistent with this, *O*-, *N*-, and *S*-glycosylation by GT-B folds (e.g. OleD [4,44] and *Arabidopsis thaliana* UGT72B1 [45]) has been reported. The exception, C-glycosylation (exemplified by the urdamycin UrdGT2 [46]), catalyzes nucleophilic aromatic substitution which is aided by an acceptor's ortho and/or para directing hydroxyl groups [47,48]. Based upon the structure, UrdGT2 Asp137 has been implicated in deprotonation of the ortho-directing C9 acceptor hydroxyl as the activating step in C-glycosylation (Figure 3b). Recently, two Cys-specific *S*-glycosyltransferases have been reported (SunS [49] and GccA [50]) and, based upon sequence analysis, these *S*-GTs may adopt novel folds.

VI. GT Engineering

Sequence- or structure-guided GT engineering includes positional mutagenesis and the construction of chimeric GTs. Success in the latter case, which includes plant flavonoid, and bacterial aminoglycoside, glycopeptide, and angucycline GTs [45,51–56], typically requires a very high degree of sequence/structural homology between the GTs to be combined. This section focuses upon the engineering fundamentals of donor, aglycon and nucleophile

specificity as well as the impact of GT proficiency upon engineering ‘new’ activities in the context of both the GT-A and GT-B fold. A summary of the representative GT engineering examples highlighted within this section is also provided in Table I.

Given the lack of global sugar ‘signatures’, few, if any, general rules exist for engineering sugar donor specificity of either the GT-A or GT-B fold and notably, the modifications that have achieved altered donor specificity are not exclusively found within the donor-binding site. Among recent GT-A examples (Figure 4a), a Y289L mutation of galactosyltransferase (β -1,4-GalT1) enabled the transfer of both GalNAc and Gal at comparable rates and mutation of R228K of β -1,4-GalT1 showed 15-fold improved glucosyltransferase (GlcT) activity at the expense of GalT activity [12]. Studies with a β -1,3-glucuronosyltransferase (β -1,3-GlcAT) created a mutant, H308R, with broadened donor specificity toward Glc, Man and GlcNAc [57]. Among recent GT-B examples, a single R140W mutation changed the glucuronosyltransferase (GlcAT) VvGT5 to a bifunctional GlcT/GalT [58] and replacement of Gln373 with histidine of the bifunctional GlcT/GalT VvGT6 resulted in loss of VvGT6 activity toward UDP-Glc without substantial loss of GalT activity. This study also identified Gln373 and Pro19 to play important roles in the bifunctional specificity of the enzyme [58,59] (Figure 4b).

Given the divergent nature of acceptors and acceptor binding, general rules are also lacking for engineering acceptor specificity of either the GT-A or GT-B fold and again, the modifications that have achieved altered acceptor specificity are not exclusively found within the acceptor-binding site. Studies to alter acceptor specificity of the GT-A fold are lacking, possibly limited due to the requirement of donor binding for the formation of the acceptor binding site. Among recent GT-B examples, structure-guided mutagenesis of the legume flavonoid GT, UGT71G1, identified two mutations (F148V and Y202A, Figure 4b) which drastically changed the regioselectivity of quercetin glycosylation [60]. In another GT-B example, domain swapping studies between the landomycin LndGT1 and LanGT1 (Figure 4c), revealed that the iterative function of LanGT1 is controlled by up to 20 amino acids dispersed throughout the catalyst [54]. Similarly, studies of chimeric GTs identified amino acid region 125–239 of plant GT, UGT71E1 as crucial for the regioselectivity of *trans*-resveratrol glycosylation [53].

Engineering aglycon nucleophile specificity has also been accomplished in the GT-B fold. Toward this goal, Brazier-Hicks et. al performed domain shuffling between two plant flavonoid GTs (UGT72B1, a bifunctional, *N/O*-GT, and *Bn*UGT, an *O*-GT) and found that only chimeras containing residues 259–370 from UGT72B1 could support *N*-glucosylation. Mutation at just two positions, D312N/F315Y in *Bn*UGT were ultimately sufficient for *N*-GT, while the reciprocal mutations in UGT72B1 (N312D-Y315F) decreased the *N*-GT function of UGT72B1 [45].

Engineering either acceptor or donor binding sites can also improve rates and/or catalytic efficiency. Representative examples include mannosylglycerate synthase (GT-A fold, donor site mutation, 72-fold increase in k_{cat}) [61], *Medicago truncatula* UGT85H2 (GT-B fold, donor or acceptor site mutations, up to 54-fold increase in the catalytic efficiency, Figure 4b) [62] and macrolide GT OleD (GT-B fold, donor and acceptor site mutations, up to 60-fold increase in the catalytic efficiency) [33]. Importantly, these increases in GT proficiency also often correlate to increases in GT promiscuity. Perhaps the best illustration of this phenomenon derives from the directed evolution of OleD (Figure 4d) which led to a series of more proficient variants capable of: *i*) utilizing 12 ‘new’ sugar donors, *ii*) utilizing >50 ‘new’ acceptors (most of which are drugs or drug-like molecules), *iii*) the glycosylation of a wide range of nucleophiles (including alcohols, amines, thiols, oximes, hydrazines, hydrazides, *N*-hydroxyamides, *O*-substituted oxyamines, and carboxylic acids), and *iv*)

iterative glycosylation (to enable the formation of di- and trisaccharides) [4,33,63]. Thus, as GT engineering and evolution efforts continue to advance at a rapid pace, it will be important for those in the field to carefully distinguish true specificity changes from ‘new’ activities simply deriving from such global overall proficiency improvements.

Acknowledgments

The unpublished work presented was supported by NIH (AI52218, CA84374 and GM074901) and the Laura and Edward Kremers Chair in Natural Products Chemistry (J.S.T.). We thank Prof. Gavin Williams (NCSU Dept. Chem.) for helpful discussion and review of this manuscript and also Craig Bingman and other members of the Center for Eukaryotic Structural Genomics.

References

1. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem.* 2008; 77:521–555. [PubMed: 18518825] •• This is an outstanding in-depth review on the fundamentals of glycosyltransferase catalysis
2. Weadge, JT.; Palcic, MM. Glycosyltransferases, Chemistry of. In: Begley, TP., editor. Wiley Encyclopedia of Chemical Biology. Vol. vol 2. 2009. p. 198-211.
3. Williams GJ, Thorson JS. Natural product glycosyltransferases: properties and applications. *Adv Enzymol Relat Areas Mol Biol.* 2009; 76:55–119. [PubMed: 18990828]
4. Gantt RW, Goff RD, Williams GJ, Thorson JS. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew Chem Int Ed Engl.* 2008; 47:8889–8892. [PubMed: 18924204] •• The dramatic promiscuity of an evolved OleD mutant was probed, culminating in the demonstrated ability of this mutant to glycosylate >70 drugs and drug-like compounds, act upon a variety of unusual nucleophiles (including amines and thiols), and catalyze iterative glycosylation.
5. Williams GJ, Yang J, Zhang C, Thorson JS. Recombinant *Escherichia coli* prototype strains for *in vivo* glycorandomization. *ACS Chem Biol.* 2010; 6:95–100. [PubMed: 20886903]
6. Thibodeaux CJ, Melancon CE, Liu HW. Unusual sugar biosynthesis and natural product glycodiversification. *Nature.* 2007; 446:1008–1016. [PubMed: 17460661]
7. Harle J, Bechthold A. Chapter 12. The power of glycosyltransferases to generate bioactive natural compounds. *Methods Enzymol.* 2009; 458:309–333. [PubMed: 19374988]
8. Hansen SF, Bettler E, Rinnan A, Engelsen SB, Breton C. Exploring genomes for glycosyltransferases. *Mol Biosyst.* 2010; 6:1773–1781. [PubMed: 20556308] • This review highlights recent efforts toward plant genome mining for novel GTs and the state-of-the-art bioinformatic tools for preliminary sequence-based annotation.
9. Erb A, Weiss H, Harle J, Bechthold A. A bacterial glycosyltransferase gene toolbox: generation and applications. *Phytochemistry.* 2009; 70:1812–1821. [PubMed: 19559449] • This review highlights recent efforts toward bacterial genome mining for novel GTs and their application toward the *in vivo* engineering of novel natural product glycosides.
10. Roychoudhury R, Pohl NL. New structures, chemical functions, and inhibitors for glycosyltransferases. *Curr Opin Chem Biol.* 2010; 14:168–173. [PubMed: 20129812]
11. Cantarel BL, Coutinho PM, 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–D238. [PubMed: 18838391]
12. Qasba PK, Ramakrishnan B, Boeggeman E. Structure and function of beta -1,4-galactosyltransferase. *Curr Drug Targets.* 2008; 9:292–309. [PubMed: 18393823]
13. Baskaran S, Roach PJ, DePaoli-Roach AA, Hurley TD. Structural basis for glucose-6-phosphate activation of glycogen synthase. *Proc Natl Acad Sci U S A.* 2010; 107:17563–17568. [PubMed: 20876143] • The structural elucidation of yeast glycogen synthase revealed a unique regulatory GT-B helical insertion essential for the formation of an unusual tetrameric structure (basal state) which also serves as a key to activation of the catalyst (via structural rearrangement upon substrate binding or phosphorylation).
14. Rao FV, Rich JR, Rakic B, Buddai S, Schwartz MF, Johnson K, Bowe C, Wakarchuk WW, DeFrees S, Withers SG, et al. Structural insight into mammalian sialyltransferases. *Nature*

Structural & Molecular Biology. 2009; 16:1186–1188. •• Crystal structures of mammalian sialyltransferase (porcine ST3Gal-I), in both apo and ligand-bound form, revealed important insights relevant to mechanism and mechanism-based inhibitor design.

15. Miley MJ, Zielinska AK, Keenan JE, Bratton SM, Radominska-Pandya A, Redinbo MR. Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *J Mol Biol.* 2007; 369:498–511. [PubMed: 17442341]
16. Lazarus MB, Nam Y, Jiang J, Sliz P, Walker S. Structure of human *O*-GlcNAc transferase and its complex with a peptide substrate. *Nature.* 2011; 469:564–567. [PubMed: 21240259] •• Crystal structures of the mammalian nutrient sensor *O*-GlcNAc transferase, in ligand-bound form (UDP and peptide), revealed important insights relevant to protein recognition as well as the presence of tetratricopeptide repeat units and an intervening domain between the canonical GT-B fold N-terminal and C-terminal domains.
17. Osmani SA, Bak S, Moller BL. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry.* 2009; 70:325–347. [PubMed: 19217634]
18. Rosen ML, Edman M, Sjoström M, Wieslander A. Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *J Biol Chem.* 2004; 279:38683–38692. [PubMed: 15148316]
19. Igura M, Maita N, Kamishikiryo 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. [PubMed: 18046457]
20. Maita N, Nyirenda J, Igura M, Kamishikiryo J, Kohda D. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J Biol Chem.* 2010; 285:4941–4950. [PubMed: 20007322]
21. Pesnot T, Jorgensen R, Palcic MM, Wagner GK. Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. *Nat Chem Biol.* 2010; 6:321–323. [PubMed: 20364127] • This study highlights a novel base-modified UDP-Gal to function as an efficient GT inhibitor. Structural elucidation of an inhibitor-bound complex revealed the substrate analog to lock the target GT in a catalytically inactive conformation and suggests this unique mode of GT inhibition to be generally applicable.
22. Schuman B, Persson M, Landry RC, Polakowski R, Weadge JT, Seto NO, Borisova SN, Palcic MM, Evans SV. Cysteine-to-serine mutants dramatically reorder the active site of human ABO(H) blood group B glycosyltransferase without affecting activity: structural insights into cooperative substrate binding. *J Mol Biol.* 2010; 402:399–411. [PubMed: 20655926]
23. Lu W, Du J, Stahl M, Tzivelekidis T, Belyi Y, Gerhardt S, Aktories K, Einsle O. Structural basis of the action of glucosyltransferase Lgt1 from *Legionella pneumophila*. *J Mol Biol.* 2010; 396:321–331. [PubMed: 19941871]
24. Bolam DN, Roberts S, Proctor MR, Turkenburg JP, Dodson EJ, Martinez-Fleites C, Yang M, Davis BG, Davies GJ, Gilbert HJ. The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc Natl Acad Sci U S A.* 2007; 104:5336–5341. [PubMed: 17376874] • Structural elucidation of macrolide-inactivating GTs OleD and OleI revealed both the macrolide structure (erythromycin versus oleandomycin) and elements of the corresponding GT to contribute to each macrolide adopting distinct conformations in complex with the catalyst.
25. Vetting MW, Frantom PA, Blanchard JS. Structural and enzymatic analysis of MshA from *Corynebacterium glutamicum*: substrate-assisted catalysis. *J Biol Chem.* 2008; 283:15834–15844. [PubMed: 18390549]
26. Ramakrishnan B, Qasba PK. Crystal structure of the catalytic domain of *Drosophila* beta-1,4-galactosyltransferase-7. *J Biol Chem.* 2010; 285:15619–15626. [PubMed: 20236943]
27. Guerin ME, Kordulakova J, Schaeffer F, Svetlikova Z, Buschiazzi A, Giganti D, Gicquel B, Mikusova K, Jackson M, Alzari PM. Molecular recognition and interfacial catalysis by the essential phosphatidylinositol mannosyltransferase PimA from *Mycobacterium*. *J Biol Chem.* 2007; 282:20705–20714. [PubMed: 17510062]
28. Sheng F, Yep A, Feng L, Preiss J, Geiger JH. Oligosaccharide binding in *Escherichia coli* glycogen synthase. *Biochemistry.* 2009; 48:10089–10097. [PubMed: 19761218]

29. Osawa T, Sugiura N, Shimada H, Hirooka R, Tsuji A, Shirakawa T, Fukuyama K, Kimura M, Kimata K, Kakuta Y. Crystal structure of chondroitin polymerase from *Escherichia coli* K4. *Biochemical and Biophysical Research Communications*. 2009; 378:10–14. [PubMed: 18771653]
30. Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J*. 2006; 25:1396–1405. [PubMed: 16482224]
31. Steiner K, Hagelueken G, Messner P, Schaffer C, Naismith JH. Structural basis of substrate binding in WsaF, a rhamnosyltransferase from *Geobacillus stearothermophilus*. *J Mol Biol*. 2010; 397:436–447. [PubMed: 20097205]
32. Batt SM, Jabeen T, Mishra AK, Veerapen N, Krumbach K, Eggeling L, Besra GS, Futterer K. Acceptor substrate discrimination in phosphatidyl-myo-inositol mannoside synthesis: structural and mutational analysis of mannosyltransferase *Corynebacterium glutamicum* PimB'. *J Biol Chem*. 2010; 285:37741–37752. [PubMed: 20843801]
33. Williams GJ, Zhang C, Thorson JS. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat Chem Biol*. 2007; 3:657–662. [PubMed: 17828251]
•• The first application of the directed evolution of a GT-B fold glycosyltransferase (OleD) highlighted prodigy with remarkable improvements in GT promiscuity and promiscuity.
34. Lovering AL, Lin LY, Sewell EW, Spreter T, Brown ED, Strynadka NC. Structure of the bacterial teichoic acid polymerase TagF provides insights into membrane association and catalysis. *Nat Struct Mol Biol*. 2010; 17:582–589. [PubMed: 20400947]
35. Unligil UM, Zhou S, Yuwaraj S, Sarkar M, Schachter H, Rini JM. X-ray crystal structure of rabbit *N*-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J*. 2000; 19:5269–5280. [PubMed: 11032794]
36. Persson K, Ly HD, Dieckelmann M, Wakarchuk WW, Withers SG, Strynadka NC. Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol*. 2001; 8:166–175. [PubMed: 11175908]
37. Letts JA, Rose NL, Fang YR, Barry CH, Borisova SN, Seto NO, Palcic MM, Evans SV. Differential recognition of the type I and II H antigen acceptors by the human ABO(H) blood group A and B glycosyltransferases. *J Biol Chem*. 2006; 281:3625–3632. [PubMed: 16326711]
38. Mulichak AM, Lu W, Losey HC, Walsh CT, Garavito RM. Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry*. 2004; 43:5170–5180. [PubMed: 15122882]
39. Mulichak AM, Losey HC, Lu W, Wawrzak Z, Walsh CT, Garavito RM. Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc Natl Acad Sci U S A*. 2003; 100:9238–9243. [PubMed: 12874381]
40. Modolo LV, Li L, Pan H, Blount JW, Dixon RA, Wang X. Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. *J Mol Biol*. 2009; 392:1292–1302. [PubMed: 19683002]
41. Soya N, Fang Y, Palcic MM, Klassen JS. Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology*. 2010
42. Lairson LL, Chiu CP, Ly HD, He S, Wakarchuk WW, Strynadka NC, Withers SG. Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. *J Biol Chem*. 2004; 279:28339–28344. [PubMed: 15075344]
43. Errey JC, Lee SS, Gibson RP, Fleites CM, Barry CS, Jung PMJ, O'Sullivan AC, Davis BG, Davies GJ. Mechanistic insight into enzymatic glycosyl transfer with retention of configuration through analysis of glycomimetic inhibitors. *Angewandte Chemie International Edition*. 2010; 49:1234–1237.
44. Yang M, Proctor MR, Bolam DN, Errey JC, Field RA, Gilbert HJ, Davis BG. Probing the breadth of macrolide glycosyltransferases: In vitro remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *Journal of the American Chemical Society*. 2005; 127:9336–9337. [PubMed: 15984838]
45. Brazier-Hicks M, Offen WA, Gershtater MC, Revett TJ, Lim EK, Bowles DJ, Davies GJ, Edwards R. Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc Natl Acad Sci U S A*. 2007; 104:20238–20243. [PubMed:

18077347] •• Crystal structures of various ligand-bound forms of the bifunctional *O/N*-GT UGT72B1, in conjunction with domain shuffling, highlighted specific modifications which could improve the *N*-GT activity.

46. Mittler M, Bechthold A, Schulz GE. Structure and action of the C-C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J Mol Biol.* 2007; 372:67–76. [PubMed: 17640665] •• The first reported structure of a C-glycosyltransferase, and a corresponding substrate-bound model, implicated Asp137 as the key catalytic residue for aglycon activation.
47. Fischbach MA, Lin H, Liu DR, Walsh CT. In vitro characterization of IroB, a pathogen-associated C-glycosyltransferase. *Proc Natl Acad Sci U S A.* 2005; 102:571–576. [PubMed: 15598734]
48. Brazier-Hicks M, Evans KM, Gershtater MC, Puschmann H, Steel PG, Edwards R. The C-glycosylation of flavonoids in cereals. *Journal of Biological Chemistry.* 2009; 284:17926–17934. [PubMed: 19411659]
49. Oman TJ, Boettcher JM, Wang H, Okalibe XN, van der Donk WA. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nat Chem Biol.* 2011; 7:78–80. [PubMed: 21196935] • The characterization of the first Cys S-GT (sublancin SunS) highlights a relatively permissive S-GT that may serve as a new tool for general protein glycosylation.
50. Stepper J, Shastri S, Loo TS, Preston JC, Novak P, Man P, Moore CH, Havlicek V, Patchett ML, Norris GE. Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *Febs Letters.* 2011; 585:645–650. [PubMed: 21251913]
51. Truman AW, Dias MV, Wu S, Blundell TL, Huang F, Spencer JB. Chimeric glycosyltransferases for the generation of hybrid glycopeptides. *Chem Biol.* 2009; 16:676–685. [PubMed: 19549605]
52. Park SH, Park HY, Sohng JK, Lee HC, Liou K, Yoon YJ, Kim BG. Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnol Bioeng.* 2009; 102:988–994. [PubMed: 18985617] • A chimeric GT between the N-terminal domain of kanamycin KanF and the C-terminal domain of vancomycin GtfE revealed a functional domain-swapped kanamycin GT with broadened sugar promiscuity reminiscent of that displayed by GtfE.
53. Hansen EH, Osmani SA, Kristensen C, Moller BL, Hansen J. Substrate specificities of family 1 UGTs gained by domain swapping. *Phytochemistry.* 2009; 70:473–482. [PubMed: 19261311]
54. Krauth C, Fedoryshyn M, Schleberger C, Luzhetskyy A, Bechthold A. Engineering a function into a glycosyltransferase. *Chemistry & Biology.* 2009; 16:28–35. [PubMed: 19171303]
55. Hoffmeister D, Ichinose K, Bechthold A. Two sequence elements of glycosyltransferases involved in urdamycin biosynthesis are responsible for substrate specificity and enzymatic activity. *Chemistry & Biology.* 2001; 8:557–567. [PubMed: 11410375]
56. Hoffmeister D, Wilkinson B, Foster G, Sidebottom PJ, Ichinose K, Bechthold A. Engineered urdamycin glycosyltransferases are broadened and altered in substrate specificity. *Chemistry & Biology.* 2002; 9:287–295. [PubMed: 11927254]
57. Magdalou J, Fournel-Gigleux S, Ouzzine M. Insights on membrane topology and structure/function of UDP-glucuronosyltransferases. *Drug Metabolism Reviews.* 2010; 42:159–166. [PubMed: 19807219]
58. Ono E, Homma Y, Horikawa M, Kunikane-Doi S, Imai H, Takahashi S, Kawai Y, Ishiguro M, Fukui Y, Nakayama T. Functional differentiation of the glycosyltransferases that contribute to the chemical diversity of bioactive flavonol glycosides in grapevines (*Vitis vinifera*). *Plant Cell.* 2010; 22:2856–2871. [PubMed: 20693356]
59. Osmani SA, Bak S, Imbert A, Olsen CE, Moller BL. Catalytic key amino acids and udp-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1: Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. *Plant Physiology.* 2008; 148:1295–1308. [PubMed: 18829982]
60. He XZ, Wang XQ, Dixon RA. Mutational analysis of the *Medicago* glycosyltransferase UGT71G1 reveals residues that control regioselectivity for (iso)flavonoid glycosylation. *Journal of Biological Chemistry.* 2006; 281:34441–34447. [PubMed: 16982612]

61. Nielsen MM, Suits MD, Yang M, Barry CS, Martinez-Fleites C, Tailford LE, Flint JE, Dumon C, Davis BG, Gilbert HJ, et al. Substrate and metal-ion promiscuity in mannosylglycerate synthase. *J Biol Chem.* 2011
62. Modolo LV, Escamilla-Trevino LL, Dixon RA, Wang X. Single amino acid mutations of *Medicago* glycosyltransferase UGT85H2 enhance activity and impart reversibility. *FEBS Lett.* 2009; 583:2131–2135. [PubMed: 19500551]
63. Williams GJ, Goff RD, Zhang CS, Thorson JS. Optimizing glycosyltransferase specificity via "Hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chemistry & Biology.* 2008; 15:393–401. [PubMed: 18420146]
64. Modolo LV, Li L, Dixon RA, Wang X. Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of isoflavonoids. *J. Mol. Biol.* 2009; 392:1292–1302. [PubMed: 19683002]
65. Noguchi A, Horikawa M, Fukui Y, Fukuchi-Mizutani M, Iuchi-Okada A, Ishiguro M, Kiso Y, Nakayama T, Ono E. Local differentiation of sugar donor specificity of flavonoid glycosyltransferase in *Lamiales*. *Plant Cell.* 2009; 21:1556–1572. [PubMed: 19454730]
66. Ramos A, Olano C, Braña AF, Méndez C, Salas JA. Modulation of deoxysugar transfer by the elloramycin glycosyltransferase ElmGT through site-directed mutagenesis. *J Bacteriol.* 2009; 191:2871–2875. [PubMed: 19233921]

Figure 1. Representative glycosyltransferase global folds. **(a)** GT-A fold: human blood group GTA with UDP (Stick model) and Mn²⁺ ion with DXD signature highlighted (PDB ID: 3IOJ). **(b)** GT-B fold: OleD, oleandomycin glycosyltransferase with UDP and erythromycin (Stick model) with E/D-Q signature highlighted (PDB ID: 2IYF). **(c)** Secondary structure topology of GT-A fold. **(d)** Secondary structure topology of GT-B fold. Variable regions and domains for certain GTs (transmembrane domain for GT-A; tetratricopeptide unit, intervening domain for GT-B) are indicated by shaded regions. Secondary structure elements are color coded as green (helix) and magenta (β-sheet) and the N-terminal and C-terminal domains are labeled ‘N’ and ‘C’, respectively.

Figure 2.

Structures of the four calicheamicin GTs (CalG1-G4) and the corresponding calicheamicin glycosidic bonds that they form (center). CalG1 (PDB ID: 3OTH, cyan, upper left), G2 (PDB ID: 3IAA, magenta, lower left), G3 (PDB ID: 3OTI, green, upper right), and G4 (PDB ID: 3IA7, orange, lower right). Bound substrates are shown in sphere (TDP: red, CLM: slate) and the N-terminal and C-terminal domains are labeled 'N' and 'C', respectively. CalG2 utilizes the predominate N-terminal domain binding mode while CalG3 and G1 utilize a more equivalent 'bi-domain' binding mode.

(a)

Inverting mechanism **Retaining mechanism (double displacement)** **Retaining mechanism (S_Ni like mechanism)**

(b)

O-GT **N-GT** **C-GT** **CalG3** **UGT72B1** **UrdGT2****Figure 3.**

(a) GT reaction mechanisms that lead to the ultimate inversion or retention of the anomeric stereocenter upon glycosyltransfer. (b) Comparison of the active sites of a representative GT-B O-GT, N-GT, and C-GT. Substrates and catalytic residues are illustrated as sticks. (CalG3, O-GT, PDB ID: 3OTI; UGT72B1, a bifunctional O/N-GT, PDB ID: 2VCE; UrdGT2, C-GT, PDB ID: 2P6P).

Figure 4.

Representative examples of engineering donor and acceptor specificities, iterative function and proficiency with residues invoking desired changes upon mutation highlighted. **(a)** GT-A donor specificity, β -1,4-GalT1, PDB ID: 1FR8. **(b)** GT-B donor and acceptor specificity, OleD, PDB ID: 2IYF. Highlighted residues are structurally equivalent to those in other GT-B fold GTs (OleD residues G19, V136, L170, P186, A242 and Q331 correspond to P19 of VvGT6, F148 of UGT71G1, V200 of UGT85H2, Y202 of UGT71G1, I305 of UGT85H2 and Q373 of VvGT6, respectively). **(c)** GT-B iterative function, OleD, PDB ID: 2IYF. Residues highlighted represent the structurally equivalent region/residues identified to be important for iterative function by domain swapping experiments between LanGT1 and LndGT1. **(d)** GT-B proficiency and expanded donor/acceptor promiscuity, OleD, PDB ID: 2IYF. The residues highlighted indicate amino acid residues identified to alter function via directed evolution. For panels b-d, the GT-B fold N-terminal and C-terminal domains are labeled 'N' and 'C', respectively.

Table I

A summary of the representative GT-engineering studies highlighted.

glycosyltransferase	mutation(s)	engineering outcome	reference(s)
BnUGT (<i>O</i> -GT)	D312N/F315Y	new <i>N</i> -GT activity led to bifunctional <i>N</i> , <i>O</i> -GT	45
VvGT5 (GlcAT)	R140W	changed to a bifunctional GlcT/GalT	58
VvGT6 (bifunctional GlcT/GalT)	Q373H	loss of GlcT activity with gain of GalT activity	58
VvGT6 (bifunctional GlcT/GalT)	P19T	inversion of the GlcT:GalT activity	58
UGT71G1	F148V/Y202A	Altered the quercetin glycosylation regiospecificity	60
UGT85H2	I305T	improved catalytic efficiency	62
UGT85H2	V200E	improved catalytic efficiency and revealed reversibility	62
UGT78G1	E192A	reverse reaction was compromised	64
UGT88D7 (GlcAT)	S127T/R350W	specificity switch from GlcAT toward GlcT activity	65
ElmGT	L309I, L309V / N312T	altered donor specificity	66
OleD	P67TS132F/A242V/Q268V	improved proficiency and broadened donor and acceptor specificity	4, 33,63
bovine β -1,4-GalT1	Y289L	new GlcNAcT activity led to bifunctional GlcNAcT /GalT	12
β -1,4-GalT1	R228K	15-fold improved GlcT	12
β -1,3-GlcAT	H308R	broadened donor specificity	57
mannosylglycerate synthase	Q66A	improved rate	61