

Sweet secrets of synthesis

Gideon J. Davies

The crystal structure of a 'retaining' glycosyltransferase provides the first view of α -glycosidic bond synthesis using activated nucleotide sugars.

These are exciting times for glycobiology. Simply in terms of quantity, the formation of glycosides such as di-, oligo- and polysaccharides and various glycoconjugates, is the most important biological reaction on earth. Yet despite their profound importance, the enzymes involved in this process — nucleotide-sugar dependent glycosyltransferases — have until recently remained one of the great uncharted areas of structural and mechanistic enzymology. On page 166 of this issue of *Nature Structural Biology*, however, Persson and colleagues¹ reveal the structure of an enzyme that makes α -glycosidic linkages using activated α -linked nucleotide-sugar donors — a so-called 'retaining' glycosyltransferase. The α -1,4-galactosyltransferase C (lgtC) from *Neisseria meningitidis* provides our first view of how nature synthesizes glycosidic linkages with retention of anomeric configuration. The structure determination and analysis involved the harnessing of specifically synthesized UDP-sugar and oligosaccharide analogs, site-directed mutagenesis and kinetics. This work has important ramifications for researchers in glycobiology, especially those in the quest for new therapeutic agents.

Over the past decade, our understanding of the structures and mechanisms involved in the hydrolysis of glycosides has blossomed (for reviews see refs 2,3). Yet knowledge of the enzymes and mechanisms involved in the synthesis of di-, oligo- and polysaccharides has remained elusive. This reflects both the difficulty in expressing these enzymes, which are frequently membrane-bound, and the problems in characterizing enzymes whose substrates are, potentially, enormously complex and extremely challenging to synthesize (for example, a hexasaccharide has 10^{12} possible isomers).

Sequences and families of glycosyltransferases

At the sequence level there is now a large number of open reading frames that correspond to glycosyltransferases. The sequence family classification that forms

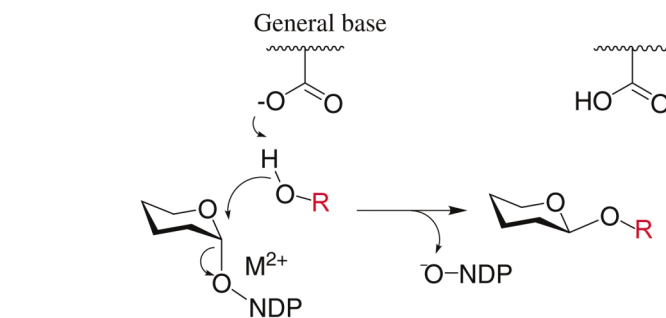


Fig. 1 Reaction mechanism for an inverting NDP-sugar glycosyltransferase. The reaction involves a single displacement in which an activated sugar donor is attacked by the acceptor leading to inversion of anomeric configuration. A Brønsted base aids nucleophilic attack by deprotonating the acceptor species and often a divalent metal ion assists leaving group departure. The recent rabbit N-acetylglucosamine transferase¹¹ and human glucuronyltransferase¹² structures provide fine views of sugar donor and acceptor species on inverting enzymes, respectively.

the foundation of research into glycoside hydrolases has now been extended to the activated-sugar dependent glycosyltransferases and currently reveals 49 potential families (available at the carbohydrate-active enzymes server⁴, CAZY, <http://afmb.cnrs-mrs.fr/~pedro/CAZY/>). Assigning a function to these ORFs remains tricky. A handful of amino acid substitutions may change the specificity of a transferase, and some enzymes even change their bond specificity depending on the nature of the substrates involved. Furthermore, like many of the enzymes acting on carbohydrates, glycosyltransferases are frequently multi-modular, in that they are often composed of independent domains within the same peptide chain. This poses additional problems for correct genomic annotation⁵. The big challenge for structural glycobiologists, therefore, is to dissect the structure and function of the enzymes in these 49 families. Put more simply, how does nature make glycosidic bonds? The structure reported by Persson and colleagues¹ goes a long way toward answering some of the unresolved issues.

The enzymatic synthesis of glycosides

The enzymatic synthesis of glycosidic bonds in nature involves several different

strategies. Glycoside hydrolases, which may normally be considered degradative enzymes, may, under favorable conditions, act 'in reverse' to synthesize oligo- and polysaccharides (for example see ref. 6). This strategy is limited because most reactions take place in aqueous solution and hydrolytic (degradative) processes are usually favored. The vast majority of glycosidic bonds in nature are, therefore, instead synthesized using 'activated' sugar precursors. Such compounds possess good leaving-groups which drive the reaction in favor of synthesis. Typically the leaving group is a lipid-phosphate sugar or a nucleotide-sugar, such as the UDP-galactose utilized by lgtC.

As with hydrolytic enzymes, we may consider two general classes of glycosyltransferases based upon whether they invert or retain the anomeric configuration. The cyclized or 'ring' forms of sugars possess what is termed an 'anomeric' carbon. In the linear form of a sugar this atom is achiral, but upon cyclization it acquires chirality and may display one of two stereochemistries termed α or β . Enzymes that perform nucleophilic substitution at this special anomeric center, including glycoside hydrolases and glycosyltransferases, therefore form two general classes based upon whether they invert or retain the stereochemistry at this center.

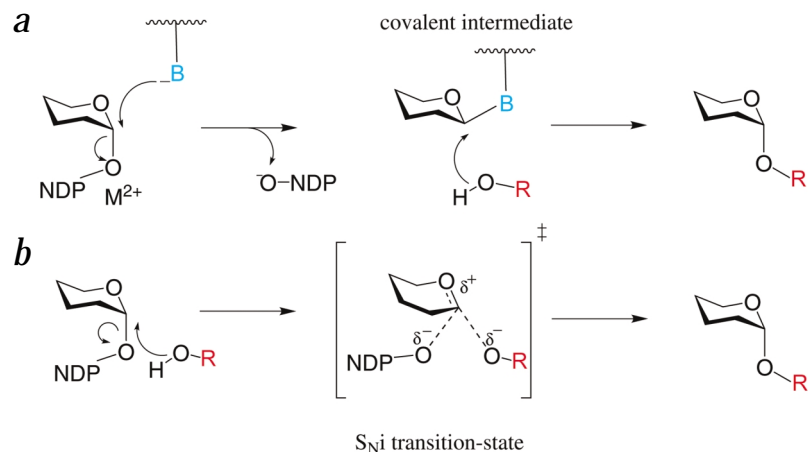


Fig. 2 Possible mechanisms for a retaining NDP-sugar glycosyltransferase. **a**, A reaction that occurs, by analogy with retaining glycoside hydrolases, *via* a covalent intermediate. Persson and colleagues¹ consider both enzymatic and acceptor-derived intermediates. **b**, An unusual S_Ni -like 'internal return' transition state in which both departing and attacking groups exist on the same side of the sugar ring.

In the case of nucleotide-sugar dependent glycosyltransferases the nucleotide-sugar donor is invariably an α -linked species. Inversion therefore gives rise to a β -linked product (that is, one whose anomeric configuration is inverted relative to the substrate) while retention of anomeric configuration gives rise to an α -linked product whose stereochemistry is identical to that of the starting sugar donor. At the mechanistic level it is these latter, retaining, enzymes that have generated the most intrigue. The inversion mechanism may easily be envisaged (Fig. 1) but the intimate secrets of retaining enzymes are less easily revealed. Furthermore, while the last few months have witnessed the determination of a number of structures of enzymes involved in glycosidic bond formation^{7–12} all these have been of the conceptually easier inverting category.

The breakthrough made by the Persson *et al.*¹ is that they have determined the first structure of LgtC, a configuration-retaining glycosyltransferase involved in the synthesis of the major glycolipid on the cell surface of *N. meningitidis*. LgtC is a membrane-associated protein that catalyzes the addition of galactose, from

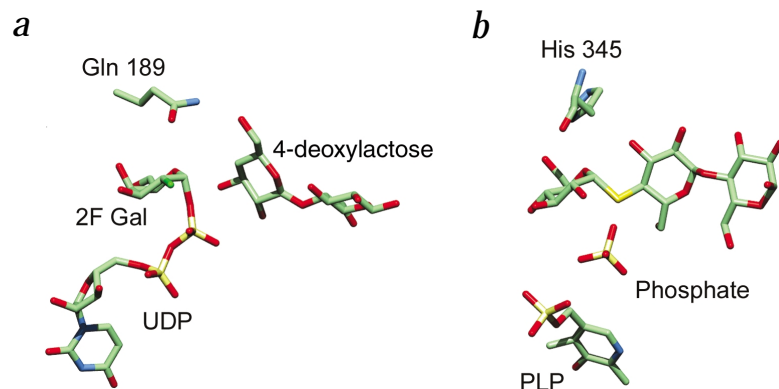
UDP-galactose, to the terminal lactose moiety on the lipopolysaccharide. The structure was determined in complex with both nucleotide-sugar donor and sugar acceptor analogs, which are themselves no mean feat of organic synthesis.

As with the inverting transferase structures (for review see ref. 13) the LgtC structure forms an α/β protein with an N-terminal nucleotide-binding fold in which the inert UDP-sugar analog binds in a shallow cleft. The C-terminal region forms the binding site for the acceptor species, which has been visualized through the use of an acceptor analog, 4-deoxylactose. Together, the positions of the UDP-2F sugar and the acceptor analog begin to shed light upon the catalytic mechanism of these enzymes.

The most important question concerning retaining glycosyltransferases is to determine what chemical mechanism is utilized for glycosidic bond formation. The mechanism for glycosidic bond hydrolysis, by glycoside hydrolases, with net retention of anomeric configuration is now very well-understood. Hydrolysis by this class of enzymes occurs in a double-displacement reaction *via* the formation and subsequent hydrolysis of a covalent

intermediate species. Yet while retaining glycoside hydrolases are well-characterized, the analogous synthetic reaction carried out by the activated-sugar dependent glycosyltransferases remains unclear. The analogy with glycosidases suggests a covalent intermediate, but such a proposal lacks any experimental evidence. In such a double-displacement mechanism (Fig. 2a) one would expect a nucleophile to lie close to the correct (β) face of the UDP-sugar donor. In the quest for a solution Persson and colleagues¹ mutated and analyzed all such potential catalytic nucleophiles in the vicinity of the UDP-sugar, yet no mutation generated an enzyme with activity reduced enough to suggest deletion of a catalytically essential nucleophile. The authors then considered a novel mechanism in which a sugar hydroxyl from the lactosyl acceptor acts as a nucleophile in a two-step reaction; again the results with specifically synthesized sugar analogs were inconclusive. A further possibility proposed by the authors is that the enzyme utilizes an extremely unusual S_Ni 'internal return' mechanism in which nucleophilic attack by the acceptor and leaving-group departure occur simultaneously on the same face of the sugar (Fig. 2b). While such a mechanism is not

Fig. 3 The active center of LgtC is reminiscent of glycogen and maltodextrin phosphorylases. Comparison of **a**, a pseudo-substrate complex of LgtC with a UDP 2-fluorogalactoside and 4-deoxylactose (PDB code 1GA8) and **b**, a pseudo-product complex of maltodextrin phosphorylase with the pyridoxal phosphate cofactor, phosphate plus a thioligosaccharide (1QM5)²¹. The unusual folded-back conformation of the α -linked UDP-Gal is similar to the glucose and phosphate positions in maltodextrin phosphorylase. In both cases a carbonyl group sits $\sim 3.3\text{Å}$ 'above' the β -face of the anomeric carbon. In LgtC this carbonyl is provided by Gln 189 and in maltodextrin phosphorylase it is provided by the main chain carbonyl of His 345.





without chemical precedent, it would be extremely unusual. It has, however, also previously been proposed for glycogen phosphorylase¹⁸, an enzyme that bears a great deal of mechanistic similarity to lgtC.

Similarities with glycogen phosphorylase

Indeed, one of the extremely revealing features of the lgtC structure is this similarity with glycogen phosphorylase. Glycogen phosphorylase itself may also be considered a retaining glycosyltransferase (albeit not one that uses nucleotide-sugars), since it utilizes glucose-1-phosphate as the activated sugar donor to synthesize α -1,4 glycosidic bonds in a reversible manner. As with the retaining nucleotide-sugar dependent transferases, its mechanism remains unclear. A mechanism has been proposed based upon crystal structures that involves a long-lived oxocarbenium ion intermediate, but this mechanism is not widely believed by chemists^{18,19}. As with lgtC, a covalent reaction intermediate would seem most likely but again none has ever been identified. It has even been proposed that a main chain carboxylate functions as a nucleophile or that the enzyme uses the unusual S_Ni mechanism described above.

There are great parallels between the way the UDP-sugar binds to lgtC, in an unusual folded-back conformation, and the binding of the pyridoxal 5'-phosphate (PLP) cofactor, glucose and phosphate in glycogen phosphorylase which again hints

that these two enzymes may share similar reaction mechanisms (Fig. 3). The PLP group of phosphorylase mimics the UMP moiety of the UDP-Gal in lgtC while the Glc-1-phosphate of phosphorylase lies similarly positioned to the galactosyl-phosphate moiety of the UDP-sugar in lgtC. Furthermore, the main chain carbonyl postulated as a nucleophile for phosphorylase (from His 345 in the case of *E. coli* maltodextrin phosphorylase), overlays the side chain carbonyl of Gln 189 in lgtC. While both are positioned close to the β -face of the sugar donor, mutation of Gln 189 in lgtC suggests that it does not play the role of an enzymatic nucleophile. The true catalytic mechanism of both glycogen phosphorylase and lgtC therefore remains a mystery.

Glycosyltransferases are central to all synthetic processes involving carbohydrates. They are important drug targets in the fight against cancer as well as bacterial, viral and fungal infections because they often perform organism and cell-specific processes. They also offer enormous potential for the chemoenzymatic synthesis of oligosaccharide-based therapeutic agents²⁰. Knowledge of their three-dimensional structures and reaction mechanisms is essential for the exploitation of these enzymes as catalysts and drug targets. The lgtC structure provides our first glimpse of a retaining glycosyltransferase and offers preliminary illumination of this complex area. Yet, while the lgtC structure, skillfully analyzed with both donor and substrate

analogs, gives us the first view of the synthetic apparatus for α -linked sugars and reveals the location of appropriate chemical functions, the exact mechanism of catalysis remains unclear. Nature has not revealed all her sweet secrets just yet.

Gideon Davies is at the Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, UK. email: davies@ysbl.york.ac.uk

1. Persson, K. *et al. Nature Struct. Biol.* **8**, 166–175 (2001).
2. Zechel, D.L. & Withers, S.G. *Acc. Chem. Rev.* **33**, 11–18 (2000).
3. Davies, G. & Henrissat, B. *Structure* **3**, 853–859 (1995).
4. Coutinho, P.M. & Henrissat, B. In *Recent Advances in Carbohydrate Engineering* (eds Gilbert, H.J., Davies, G.J., Svensson, B. & Henrissat, B.) 3–12 (Royal Society of Chemistry, Cambridge, UK; 1999).
5. Henrissat, B. & Davies, G.J. *Plant Physiol.* **124**, 1515–1519 (2000).
6. Uitdehaag, J.C.M. *et al. Nature Struct. Biol.* **6**, 432–436 (1999).
7. Vrieling, A. *et al. EMBO J.* **13**, 3413–3422 (1994).
8. Charnock, S. J. & Davies, G.J. *Biochemistry* **38**, 6380–6385 (1999).
9. Gastinel, L.N. *et al. EMBO J.* **18**, 3546–57 (1999).
10. Ha, S. *et al. Protein Sci.* **9**, 1045–1052 (2000).
11. Unligil, U.M. *et al. EMBO J.* **19**, 5269–5280 (2000).
12. Pedersen, L.C. *et al. J. Biol. Chem.* **275**, 34580–34585 (2000).
13. Unligil, U.M. & Rini, J.M. *Curr. Opin. Struct. Biol.* **10**, 510–517 (2000).
14. White, A. *et al. Nature Struct. Biol.* **3**, 149–154 (1996).
15. Davies, G.J. *et al. Biochemistry* **37**, 11707–11713 (1998).
16. Kirby, A.J. *Nature Struct. Biol.* **3**, 107–108 (1996).
17. Davies, G.J. & Wilson, K.S. *Nature Struct. Biol.* **6**, 406–408 (1999).
18. Sinnott, M.L. *Chem. Rev.* **90**, 1171–1202 (1990).
19. Davies, G.J. *et al.* In *Comprehensive biological catalysis* vol. 1 (ed. Sinnott, M.L.) 119–209 (Academic Press, London, UK; 1997).
20. Sears, P. & Wong, C.H. *Proc. Natl. Acad. Sci. USA* **93**, 12086–93 (1996).
21. Watson, K.A. *et al. EMBO J.* **18**, 4619–4632 (1999).