O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis

Marianne Schimpl[1,3], Xiaowei Zheng[1,3], Vladimir S Borodkin[1,3], David E Blair[1], Andrew T Ferenbach[1], Alexander W Schüttelkopf[1], Iva Navratilova[1], Tonia Aristotelous[1], Osama Albarbarawi[1], David A Robinson[1], Megan A Macnaughtan[2] & Daan M F van Aalten[1]*

Protein O-GlcNAcylation is an essential post-translational modification on hundreds of intracellular proteins in metazoa, catalyzed by O-linked β-N-acetylglucosamine (O-GlcNAc) transferase (OGT) using unknown mechanisms of transfer and substrate recognition. Through crystallographic snapshots and mechanism-inspired chemical probes, we define how human OGT recognizes the sugar donor and acceptor peptide and uses a new catalytic mechanism of glycosyl transfer, involving the sugar donor α-phosphate as the catalytic base as well as an essential lysine. This mechanism seems to be a unique evolutionary solution to the spatial constraints imposed by a bulky protein acceptor substrate and explains the unexpected specificity of a recently reported metabolic OGT inhibitor.

Modification of hundreds of intracellular proteins with O-GlcNAc in metazoa has been shown to affect protein stability, subcellular localization, phosphorylation and ubiquitination[1,2]. Dysregulation of cellular O-GlcNAc has been implicated in diabetes, neurodegenerative disease and cancer[3]. The addition of a single O-GlcNAc moiety to specific serines or threonines on nucleocytoplasmic proteins is accomplished by the enzyme uridine diphospho-N-acetylglucosamine:polypeptide β-N-acetylglucosaminyltransferase, also known as O-GlcNAc transferase or OGT. Deletion or knockdown of OGT has demonstrated its essential role in early embryogenesis[4]. The enzyme is part of the glycosyltransferase superfamily, a large group of enzymes that can be categorized into different families depending on their donor substrate specificity, metal dependence, structural fold and the stereochemical outcome (inverting or retaining) of the reactions they catalyze[5]. OGT is a metal-independent enzyme that uses the nucleotide sugar UDP-GlcNAc as the donor substrate to transfer β-linked O-GlcNAc on the serine or threonine of more than 1,000 acceptor proteins in the human cell. The reaction proceeds with inversion of stereochemistry and formation of UDP as the second product. Inverting glycosyltransferases, such as OGT, are thought to use an S_N2-type direct displacement catalytic mechanism (reviewed in ref. 6), which is believed to require deprotonation of the acceptor by a catalytic base at some point along the reaction coordinate. The identity of the catalytic residues is uncertain for many glycosyltransferases owing to difficulties in obtaining structures of ordered complexes with donors and/or acceptor substrates. OGT is composed of a glycosyltransferase type B (GT-B) catalytic domain and thirteen N-terminal tetratricopeptide repeats (TPRs) that are thought to be required for recognition of large protein substrates but are not essential for glycosyl transfer onto acceptor peptides[7]. How OGT targets specific sites on a limited subset of intracellular proteins and how it catalyzes O-GlcNAc transfer is not currently understood. Although the structures of bacterial OGT orthologs and human OGT (hOGT) have recently been reported[8–10], the lack of structures of complete ternary product and substrate complexes, specifically the absence of the sugar moiety from the hOGT structures, has limited the mechanistic insights gained from these studies.

Recently, several cell-penetrant OGT inhibitors have been reported that are invaluable tools for cell biological studies investigating the function of O-GlcNAc. A dicarbamate compound identified from a high-throughput small-molecule screen was shown to be a covalent inhibitor, targeting a conserved lysine (Lys842) and cysteine (Cys917) in the active site[11]. An alternative strategy has exploited a metabolic precursor, tetra-acetyl-5S-GlcNAc, which, in the form of the free sugar, is a substrate of the UDP-GlcNAc biosynthetic pathway, resulting in intracellular synthesis of the actual inhibitor UDP-5S-GlcNAc (**1**; **Fig. 1a**)[12]. Remarkably, the latter seems to specifically inhibit OGT, despite many other glycosyltransferases using UDP-GlcNAc as the sugar donor, suggesting OGT may use a mechanism different from that of other glycosyltransferases.

Here we report what are to our knowledge the first structures of ternary product and substrate complexes of hOGT with sugar donor analogs and synthetic peptides corresponding to the O-GlcNAc site from the innate immunity signaling protein TAB1. These structures suggest that OGT substrate recognition is centered on the −3 to +3 subsites of the acceptor peptide and reveal an unusual conformation of the sugar donor analog that brings the acceptor serine hydroxyl in close proximity to the α-phosphate of the donor substrate, UDP-GlcNAc. Together with *in vitro* glycosylation experiments using individual stereoisomers of the α-phosphorothioate analog of UDP-GlcNAc and site-directed mutagenesis, these data delineate a new catalytic mechanism for inverting glycosyl transfer.

RESULTS

hOGT product complex defines peptide-binding mode

Previous attempts at determining the structure of complexes between hOGT and its donor substrate UDP-GlcNAc have resulted in hydrolysis of the substrate[10]. To ascertain the precise position of GlcNAc in the active site and gain insight into OGT substrate recognition, we used a synthetic O-GlcNAcylated peptide (gTAB1tide; **Fig. 1a**) derived from the regulatory O-GlcNAc site on the innate immunity signaling protein TAB1 (ref. 13) to trap the enzyme–glycopeptide product complex. We crystallized hOGT in complex with UDP and then soaked the crystals with gTAB1tide before data collection. Electron density maps at 3.15 Å, improved by four-fold

[1]College of Life Sciences, University of Dundee, Dundee, UK. [2]Department of Chemistry, Louisiana State University, Baton Rouge, Louisiana, USA. [3]These authors contributed equally to this work. *e-mail: dmfvanaalten@dundee.ac.uk
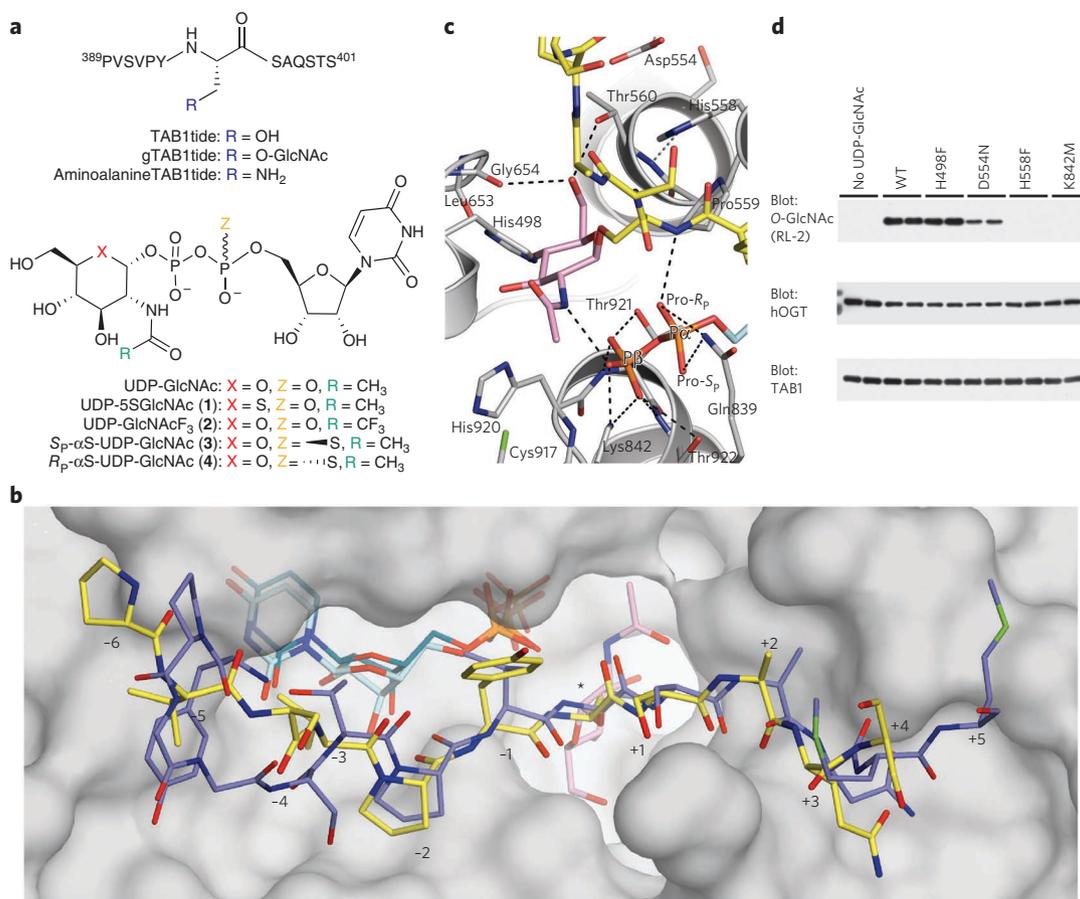
Figure 1 | The structure of a ternary hOGT product complex gives insights into the peptide-binding mode and participation of active site residues in *O*-GlcNAc transfer. (a) Chemical structures of substrate analogs: peptides based on the *O*-GlcNAc site (Ser395) from TAB1 protein and donor analogs derived from UDP-GlcNAc. (**b**) Substrate-binding groove of hOGT. Surface representation of the hOGT active site with reaction products gTAB1tide and UDP shown as sticks (the peptide, sugar and UDP are shown with yellow, pink and turquoise carbon atoms, respectively). The previously reported[10] complex with the unmodified CKII peptide and uridine diphosphate (PDB code 3PE4) is shown as sticks with blue and dark green carbon atoms, respectively. The position of the modified serine is marked by an asterisk, and subsites −6 to +5 are numbered. (**c**) hOGT product complex. Close-up view of the active site of the complex between hOGT, gTAB1tide and UDP. The enzyme is shown in cartoon representation with side chains shown as sticks with gray carbon atoms. Colors for the glycopeptide and UDP are as in **b**. The α and β phosphates of UDP are labeled with Pα and Pβ, respectively. Hydrogen bonds are depicted by dashed lines. (**d**) Activity of hOGT point mutants in an *in vitro O*-GlcNAcylation assay of TAB1 protein. *O*-GlcNAc was detected by immunoblotting with a pan-*O*-GlcNAc antibody (RL-2). WT, wild type. Experiments were performed in duplicate, as shown. Full-size blots are shown in **Supplementary Figure 2**.

noncrystallographic averaging, revealed unambiguous density for UDP and gTAB1tide (**Supplementary Results**). UDP adopts the same conformation as that observed in the recently reported hOGT–UDP–peptide complex (maximum atom shift = 1.0 Å), tethered by interactions with nine residues that are all conserved in metazoan OGTs. The peptide shows ordered density for the −6 to +4 subsites (**Supplementary Fig. 1a**) and adopts a backbone conformation in the −3 to +3 subsites, similar to the previously reported complex of hOGT with a CKII-derived substrate peptide[10] (VPYSSAQ for gTAB1tide, TPVSSAN for CKII; **Fig. 1b**), suggesting that hOGT may impose structural and/or sequence constraints on the acceptor peptide and that hOGT specificity at the peptide-sequence level may be worth exploring.

The enzyme active site does not harbor the catalytic base

In the gTAB1tide–hOGT complex, the β-linked sugar, adopting the $^4C_1$ chair conformation, projects into a conserved pocket, where it is tethered by Thr560, His920, Leu653 and Gly654 (**Fig. 1c**). The methyl group of the *N*-acetyl moiety points between the GT-B catalytic core and the TPR repeats into a pocket formed by Cys917, Met501 and Leu502, explaining why hOGT can tolerate a diverse

array of UDP-GlcNAc analogs bearing bulky amido substituents, including azido derivatives that are widely used with click chemistry to identify and enrich for *O*-GlcNAc proteins[14,15] (**Fig. 1b,c**). The identity of the OGT catalytic base that is thought to activate the serine or threonine in acceptor proteins has been the subject of a number of studies[8–10] that propose either of two histidines, His498 or His558, as candidates. Inspection of the OGT–UDP–glycopeptide product complex described here reveals that both of these histidines are positioned >4.5 Å from the acceptor hydroxyl and lack interacting residues that would support either of them acting as a catalytic base (**Fig. 1c**). Indeed, when we mutated His498, most recently proposed as the catalytic base[10], to phenylalanine, the enzyme retains activity (**Fig. 1d**). Notably, a phenylalanine is found in an equivalent position in the *Xanthomonas campestris* putative OGT[8]. The other candidate for the catalytic base, His558, is sandwiched between Pro559 and Asp554, with the carboxylate of Asp554 stacking with the His558 imidazole side chain, suggesting that Asp554 and His558 may form a catalytic dyad. However, whereas mutation of His558 renders the enzyme inactive, mutation of Asp554 does not abrogate catalysis, suggesting the effects of the His558 mutation are presumably due to structural reasons (**Fig. 1c,d**). Furthermore,

Table 1 | Binding of sugar nucleotides to hOGT.

	Wild type	**D554N**	**H558F**	**K842M**
UDP	0.54 ± 0.01	0.60 ± 0.03	0.7 ± 0.1	No binding
UDP-GlcNAc	16.1 ± 0.1	32.1 ± 0.3	42.3 ± 0.4	4.7 ± 0.1
UDP-5S-GlcNac	7.5 ± 0.1	6.2 ± 0.1	14.6 ± 0.1	5.7 ± 0.1
S_P-αS-UDP-GlcNAc	16.2 ± 0.1	10.8 ± 0.1	19.6 ± 0.1	50.0 ± 0.1
R_P-αS-UDP-GlcNAc	11.3 ± 0.1	11.5 ± 0.1	14.0 ± 0.1	2.8 ± 0.1
UDP-GlcNAcF$_3$	12.1 ± 0.1	9.5 ± 0.1	13.8 ± 0.1	29.3 ± 0.1

Binding affinity (K_d ± s.d.) of UDP, UDP-sugars and the α-phosphorothioate analogs of UDP-GlcNAc as determined by SPR (sensograms are shown in **Supplementary Fig. 5**). Values shown are in μM.

Figure 2 | The unusual conformation of the sugar nucleotide in the hOGT pseudo-Michaelis complex suggests substrate-assisted catalysis.
(**a**) hOGT pseudo-Michaelis complex. Close-up view of the active site of hOGT (gray carbons) in complex with the donor analog and inhibitor UDP-5S-GlcNAc (pink carbons for the sugar, turquoise carbons for the uridine moiety) and the synthetic peptide aaTAB1tide (acceptor serine is replaced with an aminoalanine, yellow carbons). Hydrogen bonds are depicted by dashed lines. The α and β phosphates of UDP-GlcNAc are labeled with Pα and Pβ, respectively. (**b**) Comparison of sugar nucleotide conformations from complexes with GT-B family enzymes. The structures of donor substrates in complex with active GT-B family enzymes deposited in the PDB database are shown superimposed on the sugar ring. The donor substrate from the hOGT complex is shown with pink carbons for the sugar and turquoise carbons in the uridine moiety. Sugar nucleotide coordinates were obtained from crystal structures of the following enzymes and are depicted in increasingly lighter shades of gray: inverting enzymes MurG (PDB code 1NLM), UGT71G1 (2ACW), UGT72B1 (2VCE) and VvGT1 (2C1Z) and retaining enzymes AGT (1Y6F), OtsA (1UQU and 1UQT) and WaaG (2IW1). (**c**) Selected torsion angles of the donor substrates of active glycosyltransferases belonging to the GT-B family (see **b**) plotted in a two-dimensional graph. (**d**) *In vitro* O-GlcNAcylation assay using mechanism-inspired UDP-GlcNAc analogs. O-GlcNAcylation of TAB1 by hOGT (residues 312–1031) was detected by immunoblotting with a site-specific TAB1 O-GlcNAc Ser395–specific antibody. Experiments were performed in duplicate, as shown. Full-size blots are shown in **Supplementary Figure 4**.

the imidazole side chain of His558 accepts a hydrogen bond from a protein backbone amide on (deprotonated) Nδ (**Fig. 1c**), implying that the imidazole Nε, facing the acceptor serine, will be protonated at neutral pH, which is not compatible with a role as a general base.

A Michaelis complex suggests substrate-assisted catalysis

To our surprise, it seems that none of the enzyme side chains closest to the acceptor serine can act as a catalytic base. To uncover the identity of the catalytic base, we endeavored to trap a complex with intact substrates by reducing the rate of enzymatic turnover *in crystallo* using two artificial substrate analogs: UDP-5S-GlcNAc (**1**), a recently

reported hOGT donor substrate analog inhibitor[12], and the amino-alanine derivative of the TAB1 acceptor peptide (aaTAB1tide), where the serine hydroxyl is replaced with a primary amine (**Fig. 1a**; chemical syntheses are described in **Supplementary Methods**). Notably, although turnover is reduced, OGT can use both UDP-5S-GlcNAc (reported reduction of k_{cat} by a factor of 14 compared to UDP-GlcNAc[12]) and aaTAB1tide as substrates for glycosyl transfer onto the aminoalanine (**Supplementary Fig. 3**). The observed activity with these pseudo-substrates suggests that they have catalytically competent binding modes and that their binding modes resemble those of the natural substrates. hOGT was crystallized in complex with UDP-5S-GlcNAc (**1**) and soaked with aaTAB1tide. Synchrotron data were collected to 3.3 Å, and electron density maps (improved by four-fold noncrystallographic averaging) revealed unambiguous density of a pseudo-Michaelis complex (**Fig. 2a** and **Supplementary Fig. 1b**). The overall conformation of the enzyme is almost unchanged (r.m.s. deviation on 698 Cα atoms = 0.3 Å; maximum atomic shift of any active site residue after overall superposition = 0.3 Å). The UDP moiety of UDP-5S-GlcNAc adopts a conformation similar to that of UDP in the product complex (r.m.s. deviation = 0.2 Å; maximum atomic shift = 1.0 Å), whereas the sugar is tilted away from the acceptor compared to its position in the product complex (angle of rotation = ~30°) and is tethered by His920, Leu653, Gly654 and Thr560 on the O3, O4 and O6 hydroxyls (an interpolation illustrating the atomic shifts between the ternary substrate and product complexes is shown in **Supplementary Movie 1**).

Notably, the observed conformation of the donor substrate is remarkably different from previously reported structures of GT-B donor complexes (**Fig. 2b,c**). Binding in the OGT active site seems to induce a 'back-bent' UDP-GlcNAc conformation, characterized by unusual torsion angles of the pyrophosphate, which positions the sugar directly opposite the α-phosphate (**Fig. 2a–c**). This unusual donor conformation brings the α-phosphate nonbonding (pro-R_P) oxygen to within 2.8 Å of the aminoalanine amino group mimicking the serine hydroxyl, suggesting that they would form a hydrogen bond. Concomitantly, the carbonyl group of the GlcNAc N-acetyl group approaches the serine analog to within 2.9 Å (**Fig. 2a**). In fact, these two substrate moieties approach the serine analog side chain more closely than any atom on the OGT enzyme itself. The acceptor seems to be positioned for nucleophilic attack with in-line displacement on the sugar anomeric carbon (angle between the nucleophile, anomeric carbon and leaving group = 151°), yet the side chains of His498 and His558 remain >4.5 Å away from the acceptor serine, as in the product complex (**Figs. 1c** and **2a**). Thus, inspection of this pseudo-Michaelis complex leads to the tentative identification of two nonenzymic functional groups residing on the substrate as candidates for the elusive catalytic base. We therefore hypothesized that GlcNAc transfer catalyzed by OGT proceeds without involvement of traditional enzymic general bases such as aspartate or histidine. This hypothesis was tested by devising derivatives of the substrate,

a

b
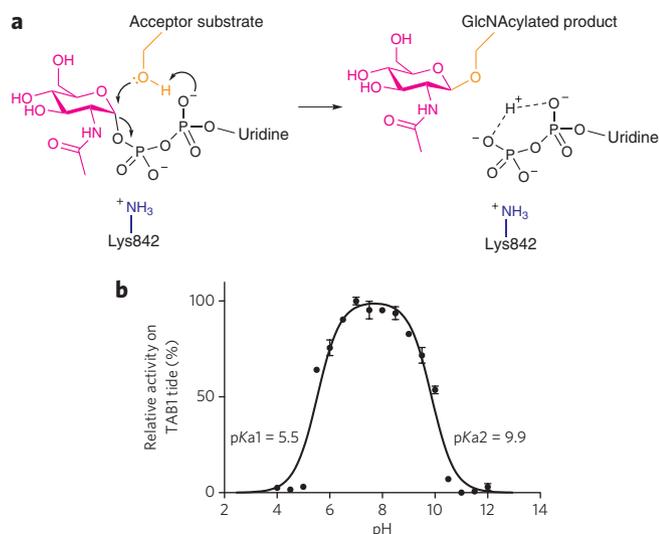
Figure 3 | Proposed catalytic mechanism and pH activity profile of hOGT.
(**a**) Schematic representation of the proposed catalytic mechanism of hOGT, showing substrate-assisted catalysis involving the sugar donor phosphates. (**b**) Activity of wild-type hOGT on N-terminally biotinylated TAB1tide in phosphate buffer with pH 4–12 was measured by scintillation proximity assay. Data points show the mean ± s.e.m. of three observations.

UDP-GlcNAc, analogous to the common practice of investigating enzymic bases by site-directed mutagenesis.

Mechanism-inspired donor analogs identify catalytic base

To investigate the possibility of a mechanism involving substrate-assisted catalysis, we prepared a number of UDP-GlcNAc derivatives. The two moieties considered as possible candidates for the catalytic base were the carbonyl oxygen of the N-acetyl group and the pro-R_P oxygen of the α-phosphate because of their proximity to the acceptor serine analog in the pseudo-Michaelis complex (**Fig. 2a**). Although the carbonyl oxygen of the N-acetyl group has been shown to act as the catalytic nucleophile in O-GlcNAc hydrolysis[16], it is not likely to act as a general base catalyst in O-GlcNAc transfer, as the extremely low pK_a (~−0.5) of the conjugated acid makes it an unlikely proton acceptor. However, pK_a values of titratable groups can be perturbed to a considerable extent in the active sites of enzymes, so we prepared the previously described N-trifluoroacetyl UDP-GlcNAc derivative[17] (**2**, UDP-GlcNAcF$_3$; **Fig. 1a**) with the aim of affecting the electronegativity of the N-acetyl carbonyl moiety to formally exclude it as the catalytic base. We found that UDP-GlcNAcF$_3$ bound the enzyme with a K_d similar to that of UDP-GlcNAc and proved to be a functional donor substrate for hOGT (**Table 1** and **Fig. 2d**), which excludes a substantial catalytic role of the N-acetyl group. To investigate the potential role of the UDP-GlcNAc α-phosphate as the catalytic base, we designed a pair of diastereomeric phosphorothioate analogs of UDP-GlcNAc in which either of the nonbonding oxygens (pro-S or pro-R) of the α-phosphate was replaced with a sulfur atom (**3** and **4**, respectively; **Fig. 1a**). Both diastereomers were able to form a complex with the enzyme with K_d values similar to that of UDP-GlcNAc (**Table 1**) and underwent nonenzymatic hydrolysis at rates similar to that of UDP-GlcNAc (**Supplementary Fig. 6**). The S_P diastereomer (**3**) was found to be a functional donor in an O-GlcNAc transfer reaction (**Fig. 2d**). Remarkably, the R_P diastereomer (**4**), where the sulfur replaces the oxygen pointing toward the acceptor serine (**Fig. 2a**), was not a substrate (**Fig. 2d**) and indeed inhibited the reaction with an half-maximum inhibitory concentration (IC$_{50}$) of 25 μM (**Supplementary Fig. 7**) but did not affect binding of the acceptor peptide (**Supplementary Fig. 5**). These data

suggest that the unusual conformation of UDP-GlcNAc positions the α-phosphate to act as the proton acceptor in hOGT-catalyzed O-GlcNAc transfer (**Fig. 3a**). The pK_a of UDP-GlcNAc in solution is thought to be approximately 2, but our measurements suggest that the pK_a may be much higher (**Supplementary Fig. 8**). Additionally, pK_a perturbations by 3–4 pH units are common in the active site of enzymes[18], and the acidic leg of a pH-activity profile of hOGT shows a pK_a of 5.5 (**Fig. 3b**). Inspection of the substrate-product complexes also reveals that the α-phosphate lacks any interactions with positively charged side chains, which are likely to increase its pK_a. Conversely, the β-phosphate of UDP-GlcNAc is located in a perfectly shaped oxyanion hole, constructed by three elements: hydrogen bonds from the backbone amides of His920, Thr921 and Thr922; alignment with an α-helical electrostatic dipole; and interaction with the evolutionarily conserved Lys842 (**Figs. 1c** and **2a**), an arrangement suited to stabilize the unusual conformation of the substrate that is required for correct positioning of the α-phosphate opposite the acceptor serine (**Fig. 2a**) as well as stabilizing the negative charge developing on the leaving group, UDP. It is possible that this developing negative charge on the β-phosphate further contributes to the α-phosphate acting as initial proton acceptor, with the β-phosphate being the terminal proton acceptor (**Fig. 3a**). As expected, the K842M mutation, removing the positive charge near the β-phosphate, abrogates enzymatic activity (**Fig. 1d**) as well as binding of the reaction product, UDP (**Table 1**). With a theoretical pK_a of 10.5, deprotonation of this lysine may represent the basic leg of the pH-activity profile of hOGT (**Fig. 3b**).

Finally, the proposed mechanism explains the unexpected observation that UDP-5S-GlcNAc is a potent and specific inhibitor of hOGT but not of any other human GlcNAc transferases[12]. The smaller C5-S-C1 bond angle, longer C5-S and S-C1 bonds and larger van der Waals radius of the sulfur (**Fig. 2**) would perturb the trajectory of proton transfer from the acceptor serine hydroxyl onto the UDP α-phosphate, a process that so far seems to be unique to hOGT.

DISCUSSION

OGT is an essential metazoan glycosyltransferase that targets specific sites on a large number of protein substrates, using UDP-GlcNAc as the donor substrate. We report crystal structures of ternary substrate and product complexes, providing snapshots of hOGT catalysis. These data show that hOGT binds peptide substrates in a defined orientation and with similar conformations near the site of O-GlcNAc transfer. Together with synthetic probes, which are diastereomeric phosphorothioate derivatives of UDP-GlcNAc, these structures define a new mechanism of inverting glycosyl transfer, where the catalytic base is not provided by the enzyme but by the α-phosphate on the donor substrate itself. This differs from other Leloir-type glycosyltransferases, where this function is generally performed by a side chain carboxylate or imidazole (reviewed in ref. 6). Participation of substrate phosphates in enzymatic catalysis is an established concept for farnesyl pyrophosphate synthases[19], and the cumulative evidence of multiple crystal structures of *Clostridium* spp. glucosylating toxins has led to the suggestion that the β-phosphate of UDP-glucose is the base in these inverting GT-A family enzymes[20]. The main impediment to a mechanism involving pyrophosphodiesters as proton acceptors is the weak basicity of such moieties (pK_a ~2 in solution). The general assumption about inverting glycosyl transfer is that it proceeds sequentially: a catalytic base deprotonates the acceptor, which launches a nucleophilic attack onto the anomeric carbon, with in-line displacement of the leaving group, resulting in inversion of stereochemistry. Yet there is little direct proof for this sequence of events, and enzymic bases are unlikely to achieve deprotonation of a serine side chain into an alcoholate (pK_a ~16). S$_N$2 reactions can be asynchronous or concerted events, and, depending on transition-state poise, the strength of the base may not be limiting. Similarly, retaining glycosyltransfer

is thought to proceed via a highly dissociative transition state[6,21]. Recently, a catalytic role for the β-phosphate has been proposed for the GDP-fucose protein O-fucosyltransferase-1 (POFUT1)[22], an inverting GT-B family enzyme, although a recent study reporting the structure of a glycosyltransferase from the same family, POFUT2, proposes an enzyme side chain as the catalytic base[23]. Markedly, the substrates for POFUT1 are specific serine residues on EGF repeats of the Notch receptor, and the clostridial toxins glucosylate Rho GTPases; like OGT, these enzymes act on protein substrates. The involvement of the UDP-GlcNAc α-phosphate in catalysis may be what enables OGT to act on both serine and threonine acceptors, as it provides an evolutionary solution to the simultaneous spatial constraints imposed by a large peptide acceptor, the requirement for recognition of the N-acetyl group on the UDP-GlcNAc donor and the need to accommodate the additional γ-methyl group on threonines. The mechanistic insights into hOGT catalysis are invaluable for the design of drug-like inhibitors to facilitate further research into the cell biological role of O-GlcNAc and the targeting of human diseases such as diabetes and cancer. It is noteworthy that as a part of our mechanistic investigations, we have now identified a hOGT point mutant, hOGT[K842M], that seems to be unaffected in substrate binding but is catalytically inactive and would be useful for cell biological dissection of the role of OGT as an enzyme versus having a scaffolding function through its TPR repeats.

METHODS

Protein crystallography. Expression and purification details of human OGT (312–1031) are in **Supplementary Methods**. Protein was crystallized in complex with donor substrate or product, and crystals were soaked with peptide or glycopeptide before freezing. Vapor diffusion crystallization experiments with hanging drops containing 1 μl protein (100 μM in a buffer of 10 mM Tris-Cl pH 8.5, 50 mM NaCl, 0.5 mM tris(hydroxypropyl)phosphine and 1 mM UDP or UDP-5S-GlcNAc) and 0.6 μl reservoir solution (1.45 M potassium phosphate, 10 mM EDTA, 1% (w/v) xylitol) gave bar-shaped crystals with maximum dimensions of 0.1 mm × 0.1 mm × 0.4 mm after 3–4 d at 20 °C. These were transferred to a drop of reservoir solution containing 2 mM of the peptide or glycopeptide (Acetyl (Ac)-PVSVPYS(-β-O-GlcNAc)SAQSTS-NH₂) for 30 min, and then were cryoprotected (1.45 M potassium phosphate, 10 mM EDTA, 27% (w/v) xylitol) and flash frozen. Data were collected at the European Synchrotron Radiation Facility at 100 K and wavelengths of 0.939 Å and 0.873 Å on beamlines ID14-4 and ID23-2, respectively. Crystals belonged to space group P321 and contained four molecules per asymmetric unit. The structure was solved by molecular replacement using the A chain of the structure deposited under Protein Data Bank (PDB) code 3PE3 as the search model. Model building was performed in Coot, and various programs of the CCP4 suite[24,25] were used for structure refinement. Ligand topologies were calculated using PRODRG[26]. Data collection and refinement statistics are given in **Supplementary Table 1**.

In vitro glycosylation of hTAB1. We incubated 1 mg ml⁻¹ of purified hTAB1 protein (resides 7–402)[27] with 0.05 mg ml⁻¹ of purified hOGT in 100 mM potassium phosphate pH 7.5 containing 1 mM UDP-GlcNAc and 1 mM dithiothreitol (DTT). The reaction was allowed to proceed for 1.5 h at 37 °C and was stopped by addition of SDS loading buffer and heating to 95 °C. Samples were subjected to SDS-PAGE and transferred to polyvinylidene fluoride membrane, followed by immunological detection of OGT (DM-17, Sigma), TAB1 and O-GlcNAc (RL-2, Abcam) with primary and secondary antibodies diluted 1:5,000 in Tris-buffered saline–Tween containing 3% (w/v) BSA. Site-specific antibody against O-GlcNAc-TAB1 Ser395 was raised in rabbit against a keyhole limpet hemocyanin–conjugated glycopeptide[13]. The TAB1-specific antibody[28] was obtained from the Division of Signal Transduction and Therapy, University of Dundee.

hOGT activity measurements by scintillation proximity assay. Radiometric detection of OGT activity on peptide substrates was achieved through scintillation proximity technology (PerkinElmer). Assays were conducted in 20-μl format in 384-well polypropylene plates. Reactions contained 200 nM hOGT (residues 312–1031), 2 μM biotinylated substrate peptides and 500 nM UDP-GlcNAc with 0.3 Ci mmol⁻¹ UDP-[³H]GlcNAc as a radioactive tracer, 100 mM potassium phosphate, 1 mM DTT and 0.2 mg ml⁻¹ BSA. Reactions were stopped by addition of 40 μl of 0.75 M phosphoric acid and then were transferred to a streptavidin-coated FlashPlate-384 (PerkinElmer) for detection on a TopCount NXT microplate luminescence counter (PerkinElmer).

Surface plasmon resonance. Surface plasmon resonance (SPR) measurements were collected using a Biacore T100 instrument (GE Healthcare). Streptavidin

was immobilized on a CM5 sensor chip (GE Healthcare) using the standard amine-coupling method. A running buffer containing 10 mM HEPES and 150 mM NaCl, pH 7.4 was used for immobilization. hOGT was biotinylated by mixing hOGT with amine-binding biotin (Pierce) in a 1:1 molar ratio. The chip surface was primed with running buffer (25 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT and 0.05% Tween 20), and biotinylated hOGT protein (residues 312–1031) was captured on the streptavidin surface. All compounds were injected in duplicates with highest concentrations ranging from 10–500 μM depending on affinity, followed by a 1:3 dilution series. Association was measured for 1 min, and dissociation was measured for 2 min. All experiments were performed at a flow rate of 30 μl min⁻¹ and at 25 °C. All data were referenced for blocked streptavidin surface and blank injections of buffer. Srubber 2 (BioLogic Software) was used to process and analyze data. Affinities were calculated using a 1:1 equilibrium binding fit. Sensograms and curve fit are shown in **Supplementary Figure 5**.

Synthesis of UDP-GlcNAc analogs. UDP-5S-GlcNAc, αS-UDP-GlcNAc and UDP-GlcNAcF₃ were synthesized using the recently published pyrophosphoryl bond-forming strategy[29] (a reaction flow diagram is in **Supplementary Scheme 1**). Briefly, the reaction of 3,4,6-tri-O-Ac-5-S-GlcNAc-1-phosphate or 3,4,6-tri-O-Ac-GlcNAcF₃-1-phosphate with 2′,3′-di-O-acetyl-5′-O-(N,N-diisopropylamino-O-cyanoethyl)phosphoramidite in the presence of 4,5-dicyano imidazole as anucleophilic catalyst resulted in the formation of intermediate phosphite-phosphate anhydrides, which were oxidized with anhydrous tert-butyl hydroperoxide to yield the required sugar nucleotide analogs after global deacetylation with triethylamine-methanol-water or guanidine, respectively. To synthesize αS-UDP-GlcNAc, we elaborated on the method to allow for the synthesis of phosphorothioates by substituting the oxidization step by sulfurization. The intermediate phosphite-phosphate anhydride obtained from 3,4,6-tri-O-Ac-GlcNAc-1-phosphate and 2′,3′-di-O-acetyl-5′-O-(N,N-diisopropylamino-O-cyanoethyl) phosphoramidite was treated with Beaucage reagent and deprotected to give a 1:1 diastereomeric mixture of αS-UDP-GlcNAc derivatives. The ³¹P-NMR spectra of the product after size-exclusion purification showed two pairs of characteristic doublets (δ 43.26 (d, $J_{P\alpha,P\beta}$ = 29.3 Hz), 43.17 (d, $J_{P\alpha',P\beta'}$ = 27.6 Hz), −14.04 (d, $J_{P\alpha',P\beta'}$ = 27.6 Hz) and −14.06 (d, $J_{P\alpha,P\beta}$ = 29.3 Hz), proving the formation of the phosphorothioate derivative. A detailed description of the synthetic procedure is in **Supplementary Methods**.

The mixture of R_P and S_P isomers of αS-UDP-GlcNAc was separated by ion-pair reverse-phase HPLC[30] on a Waters XBridge C18 Peptide separation technology 19 × 100 column (flow rate 24 ml min⁻¹) using a 15-min linear gradient (1–15% acetonitrile in 50 mM phosphate and 2.5 mM tetrabutylammonium hydrogen sulfate, pH 6.2). Retention time for the earlier-eluting isomer (S_P) was 6.19 min and 7 min for the later-eluting isomer (R_P). Products were desalted by anion-exchange chromatography on a 25 mm × 150 mm Q FF Sepharose column (flow rate 10 ml min⁻¹) using a linear gradient (0 M to 0.4 M in 15 min) of NH₄HCO₃. Final polishing was achieved by size-exclusion chromatography (Bio-Gel P2 fine; column 2.6 cm × 100 cm; flow rate 0.4 ml min⁻¹) in 0.25 M NH₄HCO₃.

The configuration at the α-phosphorus atom was unambiguously established by direct comparison of the retention times of the synthetic compounds with stereodefined product of enzymatic synthesis (**Supplementary Methods**).

References

1. Hart, G.W., Housley, M.P. & Slawson, C. Cycling of O-linked β-N-acetylglucosamine on nucleocytoplasmic proteins. Nature **446**, 1017–1022 (2007).
2. Fujiki, R. et al. GlcNAcylation of histone H2B facilitates its monoubiquitination. Nature **480**, 557–560 (2011).
3. Hart, G.W., Slawson, C., Ramirez-Correa, G. & Lagerlof, O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem. **80**, 825–858 (2011).
4. Love, D.C., Krause, M.W. & Hanover, J.A. O-GlcNAc cycling: emerging roles in development and epigenetics. Semin. Cell Dev. Biol. **21**, 646–654 (2010).
5. Cantarel, B.L. et al. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. **37**, D233–D238 (2009).
6. Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. **77**, 521–555 (2008).
7. Iyer, S.P. & Hart, G.W. Roles of the tetratricopeptide repeat domain in O-GlcNAc transferase targeting and protein substrate specificity. J. Biol. Chem. **278**, 24608–24616 (2003).
8. Clarke, A.J. et al. Structural insights into mechanism and specificity of O-GlcNAc transferase. EMBO J. **27**, 2780–2788 (2008).

9. Martinez-Fleites, C. *et al.* Structure of an O-GlcNAc transferase homolog provides insight into intracellular glycosylation. *Nat. Struct. Mol. Biol.* **15**, 764–765 (2008).
10. Lazarus, M.B., Nam, Y., Jiang, J., Sliz, P. & Walker, S. Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature* **469**, 564–567 (2011).
11. Jiang, J., Lazarus, M.B., Pasquina, L., Sliz, P. & Walker, S. A neutral diphosphate mimic crosslinks the active site of human O-GlcNAc transferase. *Nat. Chem. Biol.* **8**, 72–77 (2011).
12. Gloster, T.M. *et al.* Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **7**, 174–181 (2011).
13. Pathak, S. *et al.* O-GlcNAcylation of TAB1 modulates TAK1-mediated cytokine release. *EMBO J.* **31**, 1394–1404 (2012).
14. Wang, Z. *et al.* Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Sci. Signal.* **3**, ra2 (2010).
15. Vocadlo, D.J., Hang, H.C., Kim, E.J., Hanover, J.A. & Bertozzi, C.R. A chemical approach for identifying *O*-GlcNAc–modified proteins in cells. *Proc. Natl. Acad. Sci. USA* **100**, 9116–9121 (2003).
16. Macauley, M.S., Whitworth, G.E., Debowski, A.W., Chin, D. & Vocadlo, D.J. O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J. Biol. Chem.* **280**, 25313–25322 (2005).
17. Sala, R.F., MacKinnon, S.L., Palcic, M.M. & Tanner, M.E. UDP-*N*-trifluoroacetylglucosamine as an alternative substrate in N-acetylglucosaminyltransferase reactions. *Carbohydr. Res.* **306**, 127–136 (1998).
18. Harris, T.K. & Turner, G.J. Structural basis of perturbed pK_a values of catalytic groups in enzyme active sites. *IUBMB Life* **53**, 85–98 (2002).
19. Hosfield, D.J. *et al.* Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *J. Biol. Chem.* **279**, 8526–8529 (2004).
20. Ziegler, M.O., Jank, T., Aktories, K. & Schulz, G.E. Conformational changes and reaction of clostridial glycosylating toxins. *J. Mol. Biol.* **377**, 1346–1356 (2008).
21. Lee, S.S. *et al.* Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* **7**, 631–638 (2011).
22. Lira-Navarrete, E. *et al.* Structural insights into the mechanism of protein O-fucosylation. *PLoS ONE* **6**, e25365 (2011).
23. Chen, C.I. *et al.* Structure of human POFUT2: insights into thrombospondin type 1 repeat fold and O-fucosylation. *EMBO J.* **31**, 3183–3197 (2012).
24. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
25. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
26. Schüttelkopf, A.W. & van Aalten, D.M.F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363 (2004).
27. Conner, S.H. *et al.* TAK1-binding protein 1 is a pseudophosphatase. *Biochem. J.* **399**, 427–434 (2006).
28. Cheung, P.C., Campbell, D.G., Nebreda, A.R. & Cohen, P. Feedback control of the protein kinase TAK1 by SAPK2a/p38α. *EMBO J.* **22**, 5793–5805 (2003).
29. Gold, H. *et al.* Synthesis of sugar nucleotides by application of phosphoramidites. *J. Org. Chem.* **73**, 9458–9460 (2008).
30. Meynial, I., Paquet, V. & Combes, D. Simultaneous separation of nucleotides and nucleotide sugars using an ion-pair reversed-phase HPLC: application for assaying glycosyltransferase activity. *Anal. Chem.* **67**, 1627–1631 (1995).

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information and chemical compound information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to D.M.F.v.A.