Supplementary Information

O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis

Marianne Schimpl†[1], Xiaowei Zheng†[1], Vladimir S. Borodkin†[1], 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] and D.M.F. van Aalten[1,*]

[1] College of Life Sciences, University of Dundee, Dundee, UK

[2] Louisiana State University, Baton Rouge, LA, USA

†These authors contributed equally to this work

*Correspondence to: dmfvanaalten@dundee.ac.uk

Supplementary Results

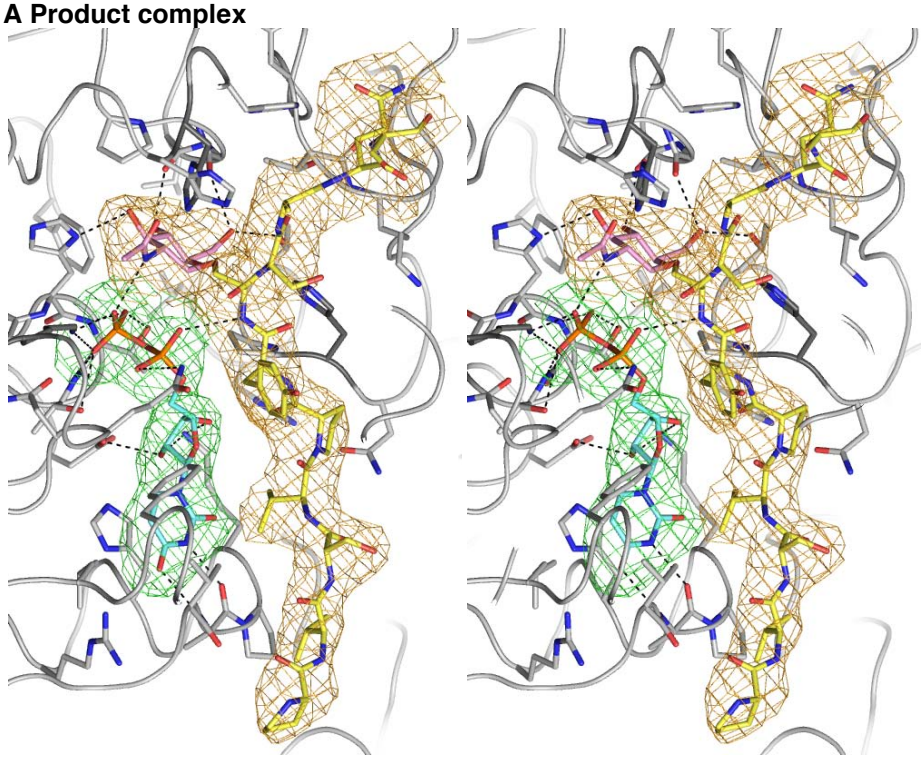
Supplementary Table 1: Crystallographic data collection and structure refinement statistics

	Product complex	Michaelis complex
	hOGT+UDP+gTAB1tide	hOGT+UDP-5S-GlcNAc+aaTAB1tide
Data collection		
Space group	*P*321	*P*321
Cell dimensions		
a = *b*, *c* (Å)	274.2, 142.2	275.4, 142.6
Resolution (Å)	30—3.15 (3.32—3.15) *	40—3.30 (3.39—3.30) *
R_{merge}	0.113 (0.481)	0.131 (0.496)
$I / \sigma I$	6.7 (2.2)	7.7 (1.7)
Completeness (%)	99.7 (99.9)	98.2 (96.9)
Redundancy	3.6 (3.3)	2.2 (2.2)
Refinement		
Resolution (Å)	30—3.15	40—3.30
No. reflections	105108	91074
R_{work} / R_{free}	17.4 / 20.4	22.7 / 27.2
No. atoms		
Protein	22052	22052
Nucleotide(sugar)	96	152
(Glyco)peptide	364	364
B-factors		
Protein	60.32	46.68
Nucleotide(sugar)	62.6	46.5
(Glyco)peptide	89.9	64.31
R.m.s. deviations		
Bond lengths (Å)	0.017	0.010
Bond angles (°)	2.04	1.31

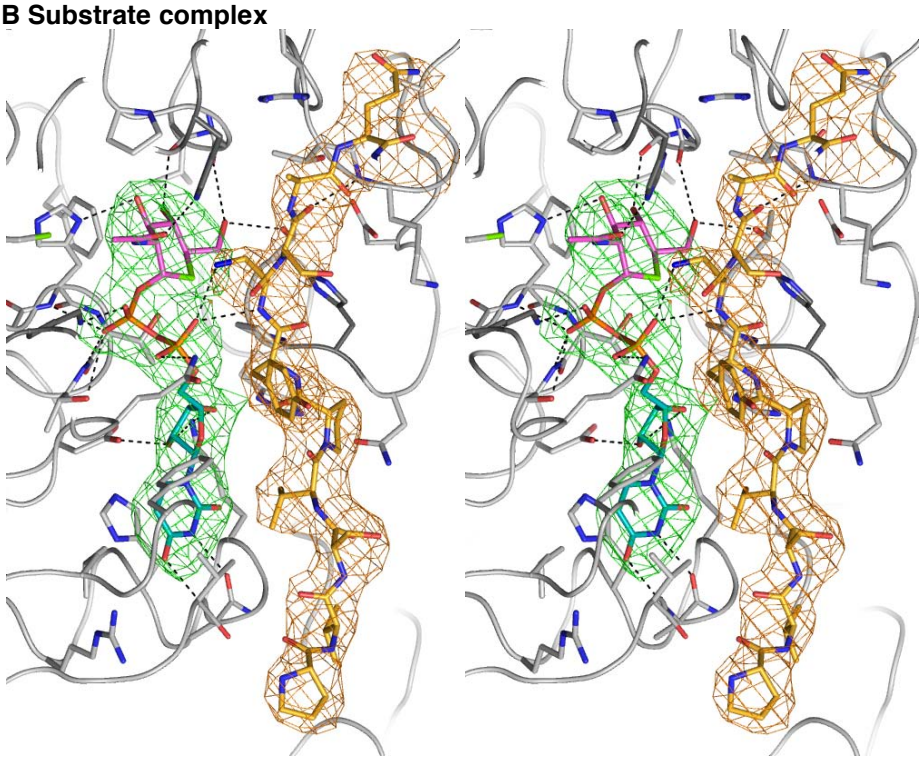
*Highest-resolution shell is shown in parentheses.

Supplementary Figure 1: Stereo image of hOGT with full length peptide ligands showing unbiased F_o-F_c electron density after four-fold averaging (2.5 σ)

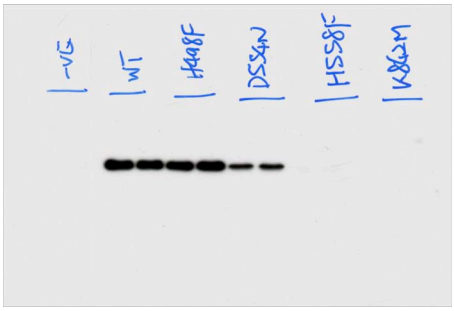
A Product complex



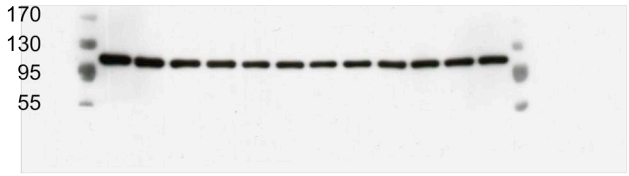
B Substrate complex



Supplementary Figure 2: Activity of hOGT point mutants in an *in vitro* O-GlcNAcylation assay of TAB1 protein (full size blots). O-GlcNAc was detected by immunoblotting with a pan-O-GlcNAc antibody (RL-2).



blot: anti-O-GlcNAc (RL2)



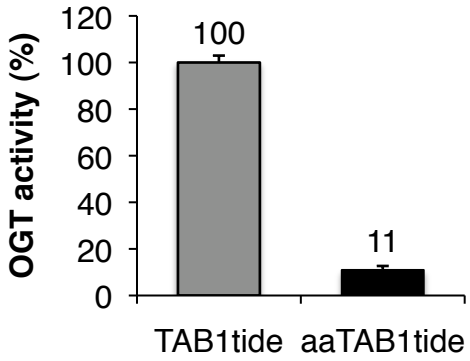
blot: anti-OGT

blot : anti-TAB1

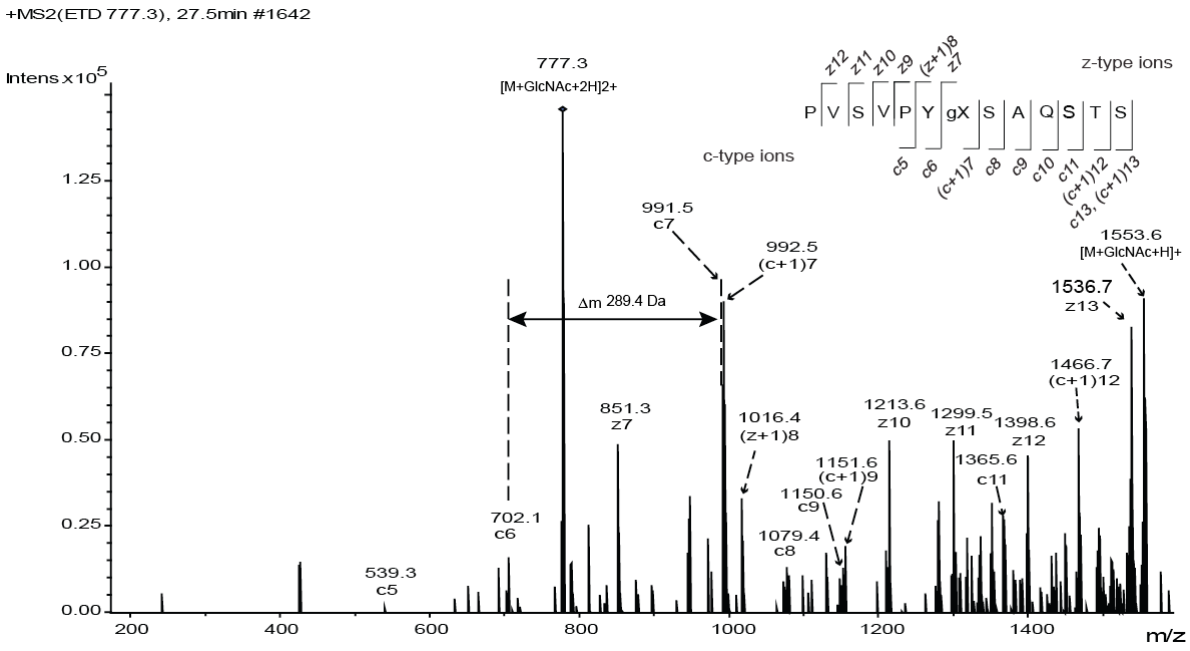
Supplementary Figure 3: Activity of hOGT on TAB1tide and aaTAB1tide

For crystallographic studies, the acceptor serine of the TAB1tide was replaced by an aminoalanine. **A** Glycosyl transfer to N-terminally biotinylated TAB1tide and aaTAB1tide was determined in a scintillation proximity assay as described in Materials and Methods. **B** *In vitro* O-GlcNAc modified aaTAB1tide (MW = 1551.75 Da) containing a HexNAc (+203.1Da) was detected by ETD-MS/MS as the doubly charged precursor ion $[M+HexNAc+2H]^{2+}$ with an m/z ratio of 777.3. Using fragmentation by ETD-MS/MS, we observed a wide number of z and c type ions including a mass difference of 289.4 Da between c6 (702.1) and c7 (991.5), representing the artificial aminoalanine plus GlcNAc, and assigning the site of glycosylation to the aminoalanine in position 7 with high confidence.

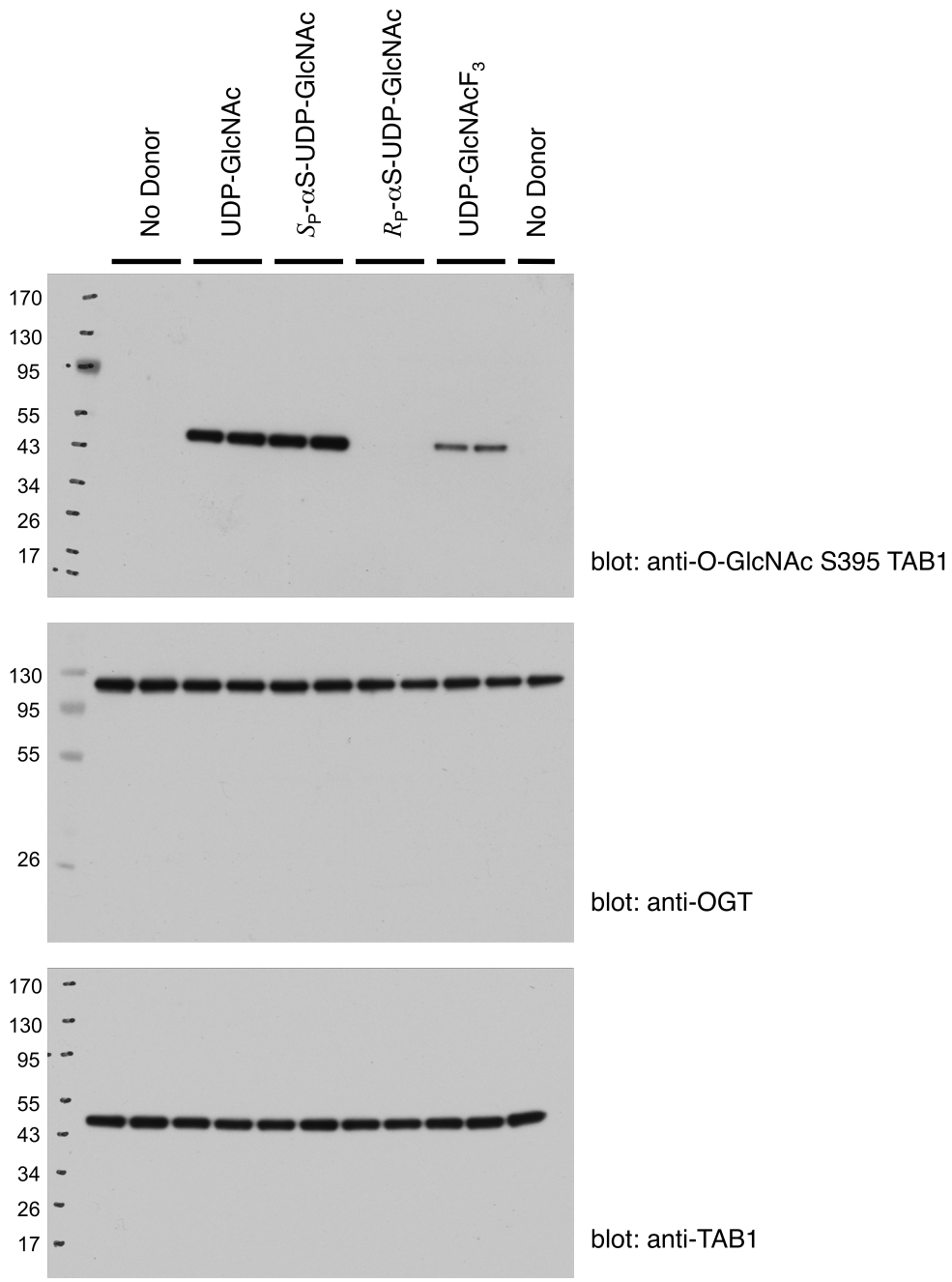
A



B



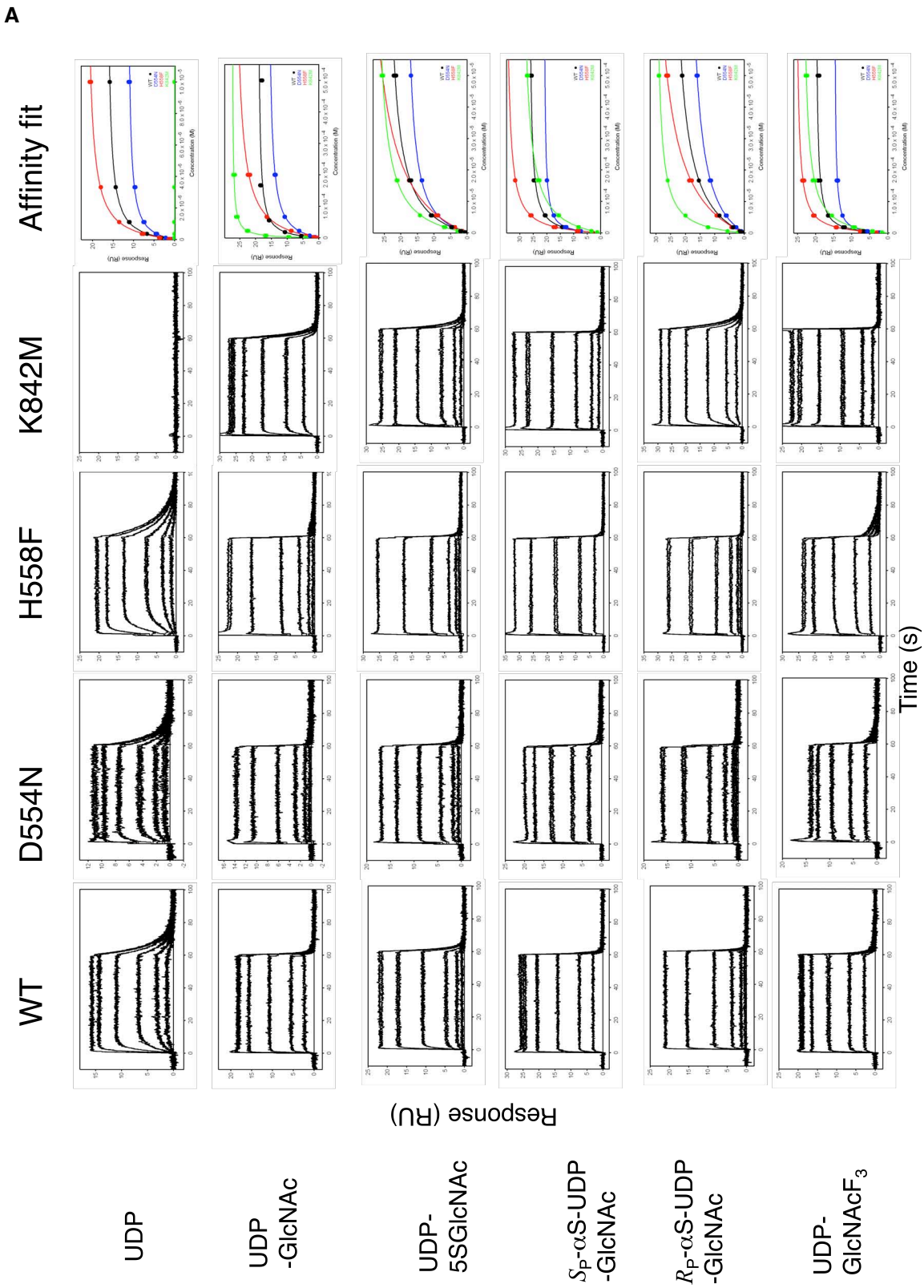
Supplementary Figure 4: *In vitro* O-GlcNAcylation assay using mechanism-inspired UDP-GlcNAc analogs (full size blots). O-GlcNAcylation of TAB1 by hOGT (312—1031) was detected by immunoblotting with a site-specific TAB1 anti-O-GlcNAc S395 antibody.



Supplementary Figure 5: Surface Plasmon Resonance (SPR) and biolayer interferometry sensograms

A. Sensograms for binding of UDP, UDP-GlcNAc, UDP-5S-GlcNAc (**1**), UDP-GlcNAcF$_3$ (**2**), S_P-αS-UDP-GlcNAc (**3**) and R_P-αS-UDP-GlcNAc (**4**) to hOGT: WT, D554N, H558F and K842M. Compounds were injected in duplicates and with highest concentrations starting at 10 μM (UDP binding to WT, D554N, H558F and K842M), 50 μM (UDP-5S-GlcNAc (**1**) and R_P-αS-UDP-GlcNAc (**4**) binding to WT, D554N, H558F and K842M), 200 μM (UDP-GlcNAc binding to D554N, H558F and K842M) and 500 μM (UDP-GlcNAc binding to WT and S_P-αS-UDP-GlcNAc (**3**) and UDP-GlcNAcF$_3$ (**2**) binding to WT, D554N, H558F and K842M). Equilibrium affinity fits are summarized in overlays for each compound binding to all hOGT mutants in the final column. Each mutant is represented in a different color.

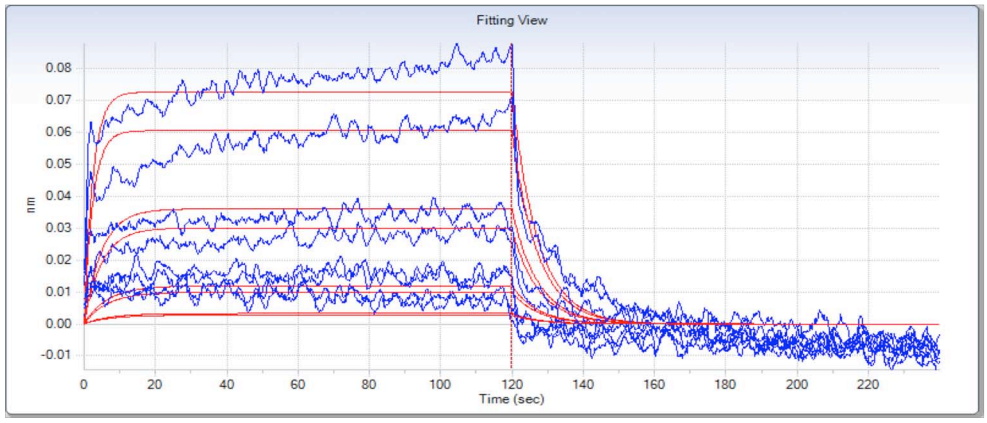
B. Determination of K_d values for TAB1tide binding to hOGT in the presence of UDP-GlcNAc and R_P-αS-UDP-GlcNAc (**4**) using biolayer interferometry. The top panel shows binding profiles for a concentration series of TAB1tide (2.5 mM, 0.625 mM, 0.156 mM, 0.039 mM) interacting with hOGT in the presence of saturating concentrations of UDP-GlcNAc and R_P-αS-UDP-GlcNAc (**4**). Two independent experiments were carried out for each concentration series. The bottom panel shows equilibrium affinity fits for each experiment. For experimental details see Supplementary Methods.


A
WT
D554N
H558F
K842M
Affinity fit
UDP
UDP -GlcNAc
UDP- 5SGlcNAc
S_P-αS-UDP -GlcNAc
R_P-αS-UDP -GlcNAc
UDP- GlcNAcF$_3$
Response (RU)
Time (s)

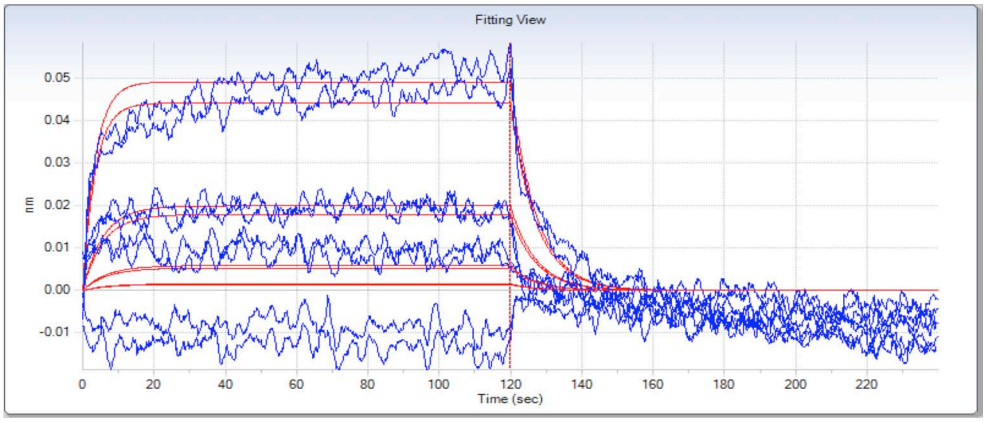
B

	TAB1tide K_D (mM)
UDP-GlcNAc	1.3 ± 0.3
R_P-αS-UDP-GlcNAc	2.4 ± 0.4

Sensogram

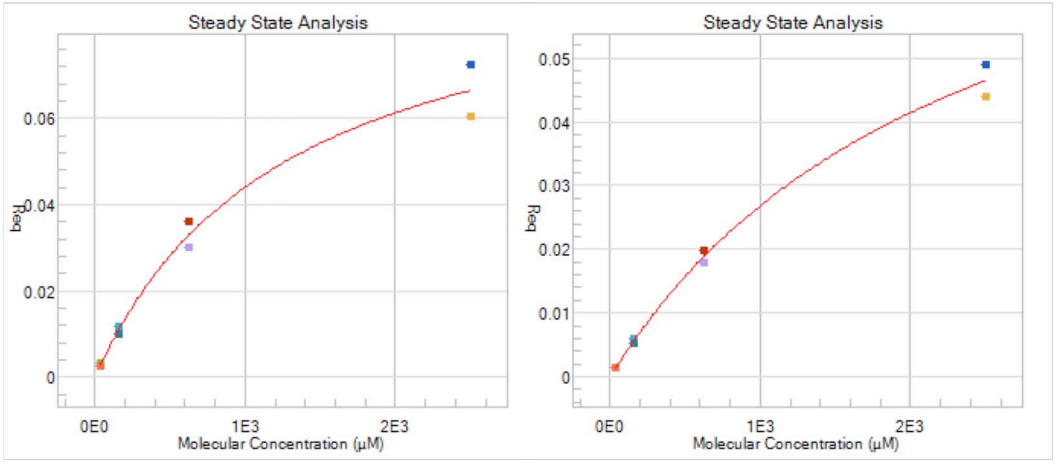


UDP-GlcNAc



R_P-αS-UDP-GlcNAc

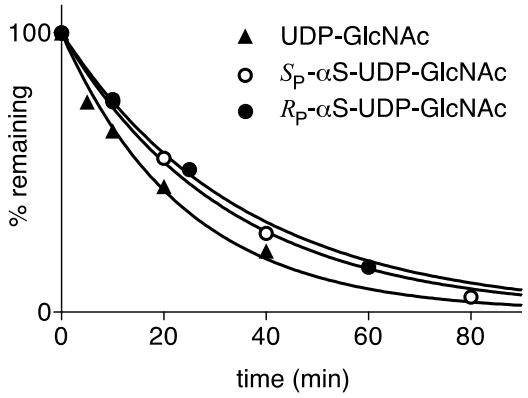
Affinity Fit



UDP-GlcNAc R_P-αS-UDP-GlcNAc

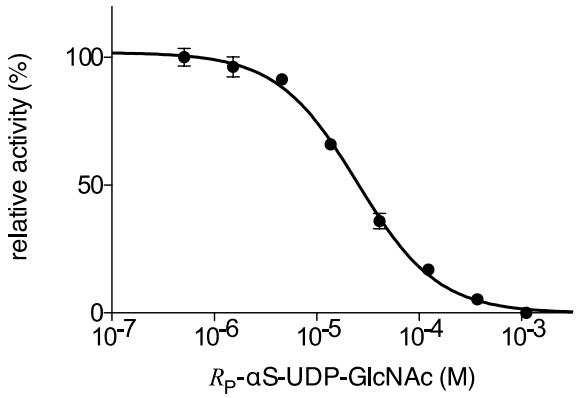
Supplementary Figure 6: Solvolysis of phosphorothioate analogues of UDP-GlcNAc

The rates of non-enzymatic solvolysis in aqueous solution were determined for UDP-GlcNAc, S_P-αS-UDP-GlcNAc (**3**) and R_P-αS-UDP-GlcNAc (**4**). Reactions were performed in 0.1 M glycine-HCl buffer pH 2.5 at 80 °C, and analyzed by ion-pair reversed-phase HPLC. Solvolysis was observed to follow first-order reaction kinetics, with rate constants of k = 2.49 h^{-1} for UDP-GlcNAc, k = 1.86 h^{-1} for S_P-αS-UDP-GlcNAc and k = 1.70 h^{-1} for R_P-αS-UDP-GlcNAc.



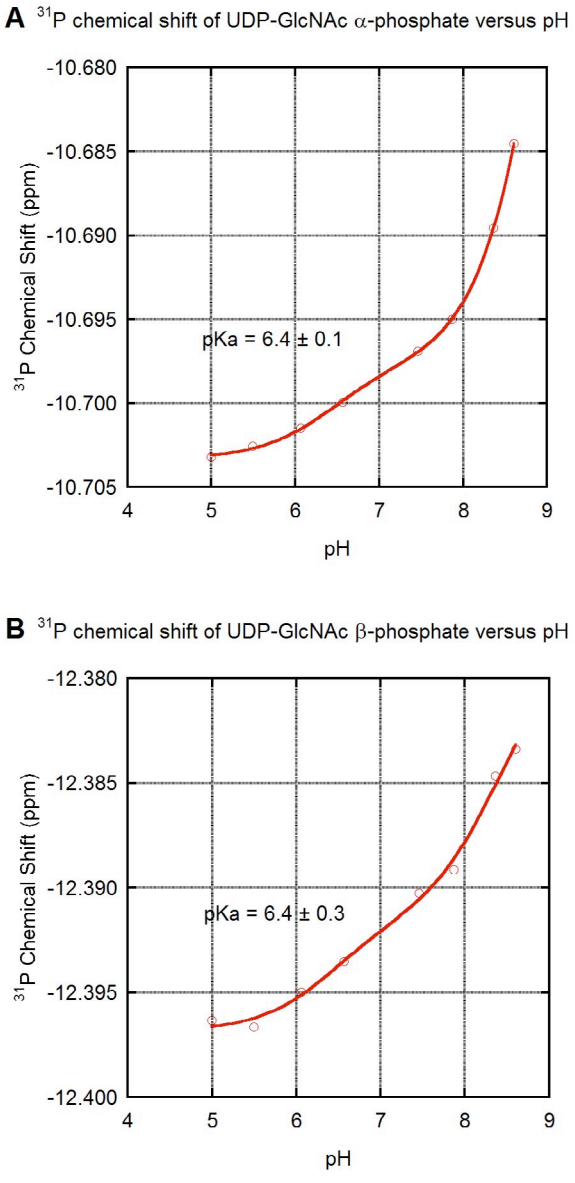
Supplementary Figure 7: Inhibition of hOGT by the donor analog R_P-αS-UDP-GlcNAc

Activity of hOGT on biotinylated TAB1tide was determined radiometrically as described in the Materials and Methods. Measurements were performed in triplicate (error bars represent the s.e.m.) and values were fitted to the standard equation for dose-dependent inhibition in GraphPad Prism 5.0. R_P-αS-UDP-GlcNAc (**4**) inhibited hOGT with an IC_{50} of 25 μM.



Supplementary Figure 8: Determination of the apparent pK_a values of the α- and β-phosphates of UDP-GlcNAc by ^{31}P NMR

The apparent pK_a values for the α- and β-phosphates of UDP-GlcNAc were determined by measuring the ^{31}P NMR chemical shift as the pH of the sample was varied from 5.0 to 8.5 (open circles). The chemical shift curves were fit to determine the pK_a values (details in the Material and Methods section. The best fit (red line), pK_a values, and fit errors are shown on each graph. The Pearson's correlation coefficients for the α- and β-phosphate curve fit are 0.9999 and 0.9976, respectively.



Supplementary Methods

Molecular Cloning of hOGT Expression Constructs

An truncated construct encoding hOGT 312-1031 was generated from cDNA. Oligonucleotide primers (fwd: 5'-CTGGAAGTTCTGTTCCAGGGGCCCGGATCCTGTCCCACCCATGCAGACTCTCTGAATAAC and rev: 5'-GCGGCCGCTTATGCTGACTCAGTTTATTCAACAGGCTTAATCATGTGGTC) were used to amplify a PCR product intended to remove the CTG before the *Bam*HI site in pGEX6P1 as well as the region encoding the first 311 amino acids of hOGT. The reverse primer introduced a stop codon after residue 1031. The PCR product was inserted into the pGEX6P1 plasmid by restriction free cloning[1]. Point mutations were introduced by site-directed mutagenesis and verified by DNA sequencing.

hOGT Protein Expression and Purification

hOGT was expressed in ArcticExpress *E. coli* (Stratagene) by inoculating 1 l of LB medium with 5 ml of overnight culture grown in the presence of 50 μg/l ampicillin and 20 μg/l gentamycin and allowing it to grow at 37 °C until OD_{600} reached 1.0. The culture was cold shocked at 4 °C for 1 h followed by induction with 125 μM IPTG and incubation at 12 degrees for 48 h. The culture was harvested by centrifugation at 3,000 *g* for 30 min and the pellet washed with TBS buffer (25 mM Tris-Cl pH 7.5, 150 mM NaCl). The pellet was then resuspended in lysis buffer (25 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF and 5 μM leupeptin) and lysed using a continuous flow cell disruptor (Constant Systems) at 30 kpsi. The lysate was centrifuged at 40,000 *g*, the supernatant filtered through 0.2 μm cellulose acetate filters and incubated with glutathione sepharose 4B beads (GE Healthcare) for 2 h at 4 °C on a rotating platform. The beads were then washed with washing buffer (25 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM DTT) and incubated with PreScission Protease (200 μg protease per ml of beads) at 4 °C for 16 h. The supernatant containing the cleaved protein was diluted to 25 mM NaCl in Tris-Cl, pH 8.5 before IEX on Q sepharose eluting with a linear gradient of 0.025—0.5 M NaCl. Fractions selected were pooled and buffer-exchanged into 25 mM Tris-Cl pH 7.5, 150 mM NaCl and 1 mM THP. The proteins were snap-frozen at 1 g/l concentration in the same buffer containing 10 % glycerol for future use.

Liquid Chromatography-Tandem Mass Spectrometry

For the site mapping analysis of enzymatically glycosylated aaTAB1tide, i*n vitro* glycosylation reactions contained 1 mg peptide, 5 mM UDP-GlcNAc and 1 μg OGT in 100 μl of 50 mM sodium bicarbonate. After 2 h incubation at 37 °C, OGT enzyme was removed using a centrifugal concentrator with 10 kDa molecular weight cut-off. The sample was evaporated to dryness and reconstituted in 0.1 % formic acid, before separation on a PepMap C18 column (75 μm i.d. x 15 cm x 2 μm) (LC Packings/Dionex) with a 4–60% acetonitrile gradient at a flow rate of 300 nl/min. A nESI source (Proxeon) was used for nESI-MS/MS, coupling the LC system to a 3D high capacity ion trap mass spectrometer with ETD capabilities (amaZon ETD; Bruker Daltonics). The mass spectrometry method used was based on alternating CID/ETD measurements on "Data Dependent mode", where up to three most abundant ions were selected per every full scan cycle for MS/MS with 30 s of dynamic exclusion time. Reagent ion ICC (Ion Charge Control) target was set to 500,000, maximum emission current 4 μA and ionization energy of 75 eV.

Data Analysis. Raw mass spectrometry data were processed in DataAnalysis 4.0 and BioTool 3.2 SR1 software packages (Bruker Daltonic). (N-term) Acetylation (+42.011 Da), (C-term) amidation (+0.984 Da) and GlcNAc (X) (+203.0794 Da) all were used as fixed modifications for ETD-MS/MS data analysis. The aminoalanine was integrated to Biotools automated data analysis using Sequence Editor Software (Bruker Daltonic). Subsequently the O-GlcNAc site mapping data were manually inspected and validated. Annotated spectra with O-GlcNAc site assignment supplied in supplemental figure S3B.

Biolayer interferometry measurements

Peptide binding affinity to hOGT in the presence of donor substrate was measured using the OCTET® RED384 system (ForteBio). hOGT was biotinylated by mixing with amine-binding biotin (Pierce) in 1:1 molar ratio. Biotin-hOGT was immobilised onto Superstreptavidin biosensors followed by blocking with 10 μg/ml biocytin . A second set of Superstreptavidin biosensors without hOGT coupling was blocked with biocytin to act as a control for non-specific binding of TAB1tide to the sensor surface. The probes were washed and equilibrated in buffer (25 mM Tris-Cl, pH 7.5, 150 mM NaCl, 1 mM DTT) containing UDP-

GlcNAc or R_P-αS-UDP-GlcNAc at saturating concentration of 500 mM. The biosensors then sampled a dilution series (4 points, 4-fold dilutions) of TAB1tide starting from a top concentration of 2.5 mM. Each experiment was run using two biosensors and at temperature 25 °C. The data were processed with ForteBio Data Analysis v.7.0. Double referencing was used to correct for instrument drift and non-specific interactions between TAB1tide and the biosensor surface. To obtain the dissociation equilibrium constant (K_d) the sensograms were fitted to a global Langmuir 1:1 model using 3 concentrations.

pK_a determination using ^{31}P NMR

Samples of 1 mM UDP-GlcNAc were prepared in 25 mM sodium phosphate buffer at pH values ranging from 5.2—8.5. 10 % D_2O was added to provide a signal for locking. The pH of the samples in 5 mm NMR tubes was measured using a calibrated micro pH probe. ^{31}P NMR spectra were collected on a Varian DPX 400 MHz instrument with 128 averaged scans, ^{1}H decoupling, coaxial, external reference (85 % H_3PO_4, 0 ppm), and 3 Hz exponential line broadening. Each phosphate displays a doublet peak due to J-coupling between the phosphorus nuclei. The average chemical shift for each doublet was used to plot against the sample pH (Supplementary Figure 8). The chemical shift curves were fit using Maple 15 software to the following equation,

(1) $\delta + \Delta\delta_1 /(1+10^{(pKa1 - pH)}) + \Delta\delta_2/(1+10^{(pKa2 - pH)})$

where δ is the chemical shift of the protonated phosphate; $\Delta\delta_1$ and $\Delta\delta_2$ are the chemical shift differences between the protonated phosphate and the deprotonated phosphate and deprotonated uracil, respectively; and pK_{a1} and pK_{a2} are the acid dissociation constants for phosphate and uracil deprotonation, respectively. The error determined for each phosphate pK_a is the fit error.

Chemical synthesis of UDP-GlcNAc analogs

Bis (9H-fluoren-9-ylmethyl) 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-a-D-glucopyranosyl phosphate (**6**): To a solution of *2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-5-thio-α-D-glucopyranose* (**5**)[2] (0.254 g; 0.7 mmol) and bis(9H-fluoren-9-ylmethyl)-diisopropylamidophosphite[3] (0.887 g, 2.43 mmol) in MeCN (10 mL) 0.45 M

solution of 1*H*-tetrazole in MeCN (6.2 mL, 2.8 mmol) was added *via* cannula at 0 °C. The reaction was removed from the cooling bath and stirred for 2 h at RT. The reaction was cooled down to 0 °C and 5.5 M solution of *t*-BuOOH in octane (0.4 mL, 2.5 mmol) was added *via* cannula and the reaction was stirred for 1 h at 0 °C. The reaction was removed from the cooling bath and additionally stirred for 45 min at RT. The reaction was diluted with DCM, the solids were filtered off on a *Celite* pad and the filtrate was concentrated. The residue was purified by flash chromatography in [PE-DCM 4:1]-Me_2CO 5—40 % to give 0.512 g (0.64 mmol; 91 %) of the target product **6** as amorphous solid.

$[\alpha]_D$ = +77.7° c 1.0 $CHCl_3$

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 – 7.71 (m, 4H, Arom.), 7.62 – 7.58 (m, 1H, Arom.), 7.58 – 7.56 (m, 1H, Arom.), 7.54 – 7.51 (m, 1H, Arom.), 7.47 – 7.25 (m, 10H, Arom.), 6.03 (d, *J* = 9.1 Hz, 1H, NHAc), 5.29 (dd, *J* = 10.9, 9.6 Hz, 1H, H-4), 5.11 – 5.07 (m, 2H, H-1, H-3), 4.49 (ddt, *J* = 12.0, 9.2, 2.8, 2.8 Hz, 1H, H-2), 4.40 – 4.28 (m, 4H, 2×FlCH_2), 4.24 (dd, *J* = 12.1, 5.1 Hz, 1H, H-6b), 4.18 (t, *J* = 6.5, 6.5 Hz, 1H, 9-H-Fl), 4.12 (t, *J* = 6.3, 6.3 Hz, 1H, 9-H-Fl′), 3.82 (dd, *J* = 12.1, 3.0 Hz, 1H, H-6a), 3.21 (ddd, *J* = 10.9, 5.0, 3.1 Hz, 1H, H-5), 2.06 (s, 3H), 2.02 (s, 3H, Ac), 1.94 (s, 3H, Ac), 1.73 (s, 3H, NHAc).

^{13}C NMR (126 MHz, $CDCl_3$) δ 170.94, 170.38, 169.70, 169.09, 142.81, 142.78, 142.75, 142.61, 141.41, 141.38, 141.34, 128.13, 128.07, 127.35, 127.29, 127.24, 125.03, 125.00, 124.79, 120.22, 120.16, 120.11, 78.90 (d, $J_{C,P}$ = 8.0 Hz, C-1), 71.35 (C-3/C-4), 71.25 (C-3/C-4), 69.82 (d, $J_{C,P}$ = 5.7 Hz, FlCH_2), 69.54 (d, $J_{C,P}$ = 5.5 Hz, FlCH'_2), 60.86 (C-6), 55.99 (d, $J_{C,P}$ = 6.2 Hz, C-2), 47.87 (9-CH-Fl′), 47.81 (9-CH-Fl), 39.37 (C-5), 22.82 (AcNH), 20.53 (2×s, 2×Ac), 20.42 (Ac).

^{31}P NMR (202 MHz, $CDCl_3$) δ -2.73.

HRMS-TOF (+): m/z = 800.2307, expected 800.2294 $[M+H]^+$

Uridine 5′-(2-acetamido-2-deoxy-5-thio-α-D-glucopyranosyl)diphosphate bis ammonium salt (**1**): A solution of (**6**) (0.168 g, 0.21 mmol) in a mixture of DCM (4 mL) and Et_3N (1 mL) was kept at RT for 36 h, concentrated and dried in vacuum. A solution of uridine 2′,3′-di-O-acetyl-5′-O-(N,N-di-isopropylamino-O-cyanoethyl)phosphoramidite **7**[4] (0.122 g, 0.23 mmol) in MeCN (5 mL) was added drop-wise to a

suspension of the above residue and 4,5-dicyanoimidazole (DCI; 0.049 g, 0.42 mmol) in MeCN (5 mL) at RT. The reaction mixture was stirred for 45 min. The reaction was cooled to 0 °C and 5.5 M solution of t-BuOOH (0.15 mL, 0.84 mmol) was added. The reaction was removed from the ice-bath and stirred for 45 min at RT. Finally, DBU (0.156 mL, 1.05 mmol) was added to the reaction. After additional 45 min the reaction was concentrated. The residue was mixed with MeOH-H_2O-Et_3N 5:2:1 (8 mL) and stirred at RT for 16 h. The reaction mixture was concentrated and partitioned between water (10 mL) and $CHCl_3$ (10 mL). The layers were separated by centrifugation at 4000 rpm for 15 min at 4 °C. The organic layer was additionally extracted with water (5mL); the layers were separated by centrifugation. The combined aqueous layer was extracted with $CHCl_3$ (5 mL); the layers were separated by centrifugation. The aqueous layer was concentrated with addition of *n*-BuOH to give an oily residue. This was purified by size exclusion chromatography (Bio-Gel P2 fine; column 2.6×100 cm; flow rate 0.4 ml/min; elution with 0.25 M NH_4HCO_3). The fractions containing the product were pooled and evaporated with addition of *n*-BuOH. The residue was further purified by anion exchange chromatography on 25×150 mm Q FF sepharose column (flow rate 10 mL/min, linear gradient 0—100 % 0.25 M NH_4HCO_3 in 15 CV). The fractions containing the product were pooled, concentrated and freeze-dried to give 0.08 g (0.12 mmol, 58 % yield) of the product **1** as fluffy white substance.

^{1}H NMR (500 MHz, D_2O) δ 7.86 (d, J = 8.1 Hz, 1H, H-6), 5.88 (d, J = 4.0 Hz, 1H, H-1′), 5.87 (d, J = 8.1 Hz, 1H, H-5), 5.21 (d, J = 7.6 Hz, 1H, H-1′′), 4.26 (m, 2H, H-2′, H-3′), 4.20 – 4.12 (m, 2H, H-4′, H-5′b), 4.12 – 4.05 (m, 2H, H-2′′, H-5′a), 3.87 (dd, J = 12.0, 5.5 Hz, 1H, H-6′′b), 3.78 (dd, J = 12.0, 2.7 Hz, 1H, H-6′′a), 3.69 – 3.57 (m, 2H, H-3′′, H-4′′), 3.27 (m, 1H, H-5′′), 1.97 (s, 3H, AcNH).^{13}C NMR (126 MHz, D_2O) δ 174.24, 166.19, 151.76, 141.61 (C-5), 102.61 (C-6), 88.45 (C-1′), 83.15 (d, $J_{C,P}$ = 9.1 Hz, C-4′), 76.62 (d, $J_{C,P}$ = 7.2 Hz, C-1′′), 73.74 and 73.61 (C-2′/C-3′, C-4′′), 72.00 (C-3′′), 69.60 (C-2′/C-3′), 64.92 (d, $J_{C,P}$ = 5.1 Hz, C-5′), 59.91 (C-6′′), 57.74 (d, $J_{C,P}$ = 7.4 Hz, C-2′′), 43.59 (C-5′′), 22.06 (NHAc).

^{31}P NMR (202 MHz, D_2O) δ -11.39 (d, J = 21.1 Hz), -13.01 (d, J = 21.1 Hz).

HRMS-TOF (+): m/z = 624.0659, expected 624.0666 $[M+H]^+$

Uridine 5′-(2-deoxy-2-trifluoroacetamido-α-D-glucopyranosyl)diphosphate dilithium salt (**2**)[5]: A solution of dibenzyl 2-deoxy-2-trifluoroacetamido-α-*D*-glucopyranosyl phosphate **8**[6] (0.17 g, 0.26 mmol) in MeOH (5 mL) was hydrogenated over Pd/C catalyst under slight H_2 overpressure for 2 h. The reaction was filtered through a Celite pad, Et_3N (0.1 mL) was added to the filtrate and the reaction was concentrated. 2′,3′-Di-O-acetyl-5′-O-(N,N-di-isopropylamino-O-cyanoethyl) phosphoramidite **7** (0.185 g, 0.35 mmol) and the above residue was co-evaporated with 5 mL MeCN, briefly dried in vacuum and re-dissolved in a fresh portion of MeCN (10 mL). To this solution 4,5-dicyanoimidazole (DCI; 0.083 g, 0.7 mmol) was added in one portion at RT and the reaction mixture was stirred for 45 min. A stock 5.5 M solution of t-BuOOH (0.16 mL, 0.85 mmol) was added to the reaction at 0 °C and the reaction was stirred for 45 min. Finally, DBU (0.164 mL, 1.1 mmol) was added and, after additional 30 min the reaction was concentrated.

The residue was dissolved in a stock solution of guanidine free base (12.5 mL, 1.25 mmol) [prepared from guanidinium hydrochloride (0.12 g, 1.25 mmol) by neutralization with freshly prepared solution of sodium methoxide (MeONa-MeOH, 0.1 M, 12.5 mL)]. The resulting slightly milky solution was kept for 70 min at RT and neutralized by addition of AcOH (0.06 mL). The reaction mixture was concentrated, diluted with water and extracted with $CHCl_3$. The mixture was centrifuged at 4000 rpm for 15 min at 4 °C. The clear layers were separated; the organic layer was extracted with water and centrifuged once more. The combined aqueous layer was concentrated with addition of *n*-BuOH. The residual syrup was applied onto a column (50 mL) of Dowex 1X4-200 (Cl^-) in 0.0003 M HCl (pH 3.5) and eluted with a step gradient of LiCl 0.1 to 0.8 M. Fractions containing product were pooled, and concentrated to a minimal volume with addition of *n*-BuOH. The residual syrup was diluted with water and desalted by size exclusion chromatography (Bio-Gel P2 fine; column 2.6×100 cm; flow rate 0.4 ml/min; elution with water). Fractions containing the product were pooled and concentrated with addition of *n*-BuOH. Finally, the material was passed through the size exclusion column again using the same conditions to remove the remaining salt to give (0.12 mmol, 46 %) of the target product **2** as fluffy white substance after freeze drying.

^{1}H NMR (500 MHz, D_2O) δ 7.84 (d, *J* = 8.1 Hz, 1H, H-5), 5.86 (d, *J* = 4.5 Hz, 1H, H-1′), 5.84 (d, *J* = 8.2 Hz, 1H, H-6), 5.49 (dd, *J* = 7.1, 3.3 Hz, 1H, H-1′′), 4.30 – 4.20 (m, 2H, H-2′, H-3′), 4.18 – 4.14 (m, 1H, H-4′), 4.11 (ddd, *J* = 11.7, 4.5, 2.5 Hz, 1H, H-5′b), 4.06 (ddd, *J* = 11.8, 5.6, 3.0 Hz, 1H, H-5′a), 3.98 (dt, *J* = 10.6,

2.9, 2.9 Hz, 1H, H-2′′), 3.87 – 3.82 (m, 1H, H-5′′), 3.82 (dd, *J* = 10.6, 9.2 Hz, 1H, H-3′′), 3.76 (dd, *J* = 12.5, 2.3 Hz, 1H, H-6′′b), 3.70 (dd, *J* = 12.5, 4.2 Hz, 1H, H-6′′a), 3.46 (t, *J* = 9.7, 9.7 Hz, 1H, H-4′′).

^{13}C NMR (75 MHz, D_2O) δ 166.26, 159.25 (t, $J_{C,F}$ = 37.7 Hz , $CF_3\underline{C}ONH$), 151.86, 141.72 (C-5), 115.82 (q, $J_{C,F}$ = 286.7 Hz, $\underline{C}F_3CONH$), 102.71 (C-6), 93.80 (d, $J_{C,P}$ = 6.1 Hz, C-1′′), 88.52 (C-1′), 83.29 (d, $J_{C,P}$ = 9.0 Hz, C-4′), 73.82 (C-2′/C-3′), 73.07 (C-5′′), 70.35 (C-3′′), 69.71 and 69.59 (C-2′/C-3′, C-4′′), 64.97 (d, $J_{C,P}$ = 5.3 Hz, C-5′), 60.31 (C-6′′), 54.45 (d, $J_{C,P}$ = 8.5 Hz, C-2′′).

^{31}P NMR (202 MHz, D_2O) δ -11.43 (d, *J* = 20.2 Hz), -13.29 (d, *J* = 20.2 Hz).

^{19}F NMR (471 MHz, D_2O) δ -75.03.

Uridine 5′-(2-acetamido-2-deoxy-α-D-glucopyranosyl)(1(S)-thiodiphosphate) bis triethylammonium salt (**3**) and *uridine 5′-(2-acetamido-2-deoxy-α-D-glucopyranosyl)(1(R)-thiodiphosphate) bis ammonium salt* (**4**): A solution of dibenzyl 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-*a-D*-glucopyranosyl phosphate (**9**)[7] (0.135 g, 0.22 mmol) and Et_3N (0.07 mL) in MeOH (2 mL) was stirred under slight H_2 overpressure with 10% Pd/C catalyst (0.05 g) for 1 h at RT. The reaction was filtered through a pad of Celite, concentrated and dried in vacuum.

A solution of 2′,3′-Di-O-acetyl-5′-O-(N,N-di-isopropylamino-O-cyanoethyl) phosphoramidite (**7**) (0.145 g, 0.275 mmol) in MeCN (5 mL) was added dropwise to a suspension of the above residue and 4,5-dicyanoimidazole (0.065 g, 0.55 mmol) in MeCN (5 mL) at RT. The reaction mixture was stirred for 45 min. Solid 3H-1,2-benzodithiol-3-one 1,1-dioxide[8] (0.067 g, 0.33 mmol) was added in one portion; the reaction was further stirred for 5 min. Finally, DBU (0.098 mL, 0.5 mmol) was added to the reaction; after additional 45 min the reaction was concentrated. The residue was mixed with MeOH-H_2O-Et_3N 5:2:1 (8 mL) and stirred at RT for 16 h. The reaction mixture was concentrated and partitioned between water (10 mL) and $CHCl_3$ (10 mL). The layers were separated by centrifugation at 4000 rpm for 15 min at 4 °C. The organic layer was additionally extracted with water (5 mL); the layers were separated by centrifugation. The combined aqueous layer was extracted with $CHCl_3$ (5mL); the layers were separated by centrifugation. The aqueous layer was concentrated with addition of *n*-BuOH to give an oily residue. This was purified by size exclusion chromatography (Bio-Gel P2 fine; column 2.6×100 cm; flow rate 0.4 ml/min;

elution with 0.25 M NH_4HCO_3). The fractions containing the product were pooled, concentrated with addition of *n*-BuOH and freeze-dried to give 0.132 g (~ 0.2 mmol) of the product as lightly yellow solid foam.

^{31}P NMR (202 MHz, D_2O) δ 43.26 (d, *J* = 29.3 Hz), 43.17 (d, *J* = 27.6 Hz), -14.04 (d, *J* = 27.6 Hz), -14.06 (d, *J* = 29.3Hz).

HRMS-TOF (+): m/z = 624.0662, expected 624.0666 $[M+H]^+$

R_P and S_P isomers of αS-UDP-GlcNAc were separated by ion-pair reverse-phase HPLC[9] on Waters XBridge C18 Peptide separation technology 19×100 column (flow rate 24 mL/min) using a linear gradient (2 to 30 % in 15 min) of eluent B (50 mM phosphate/2.5 mM TBAHS in 50 % aqueous MeCN; pH 6.2) in eluent A (50 mM phosphate/2.5 mM TBAHS in water; pH 6.2). Retention time for the earlier eluting (S_P) isomer **3** was 6.19 min, and 7 min for the later eluting (R_P) isomer **4** (Fig. S7C). Fractions containing the products were pooled and concentrated with addition of *n*-BuOH. The residue was desalted by anion exchange chromatography on 25×150 mm Q FF sepharose column (flow rate 10 mL/min) using a linear gradient (0 to 0.4 M in 15 min) of NH_4HCO_3. Final polishing was achieved by size exclusion chromatography (Bio-Gel P2 fine; column 2.6×100 cm; flow rate 0.4 ml/min); elution with 0.25 M NH_4HCO_3.

Uridine 5'-(2-acetamido-2-deoxy-α-D-glucopyranosyl)(1(S)-thiodiphosphate) bis ammonium salt (**3**):

^{1}H NMR (500 MHz, D_2O) δ 7.92 (d, *J* = 8.2 Hz, 1H), 5.85 (d, *J* = 3.9 Hz, 1H), 5.83 (d, *J* = 8.1 Hz, 1H), 5.40 (dd, *J* = 6.9, 3.3 Hz, 1H), 4.24 (d, *J* = 2.7 Hz, 2H), 4.16 (s, 1H), 4.12 (d, *J* = 7.3 Hz, 2H), 3.88 – 3.79 (m, 2H), 3.77 – 3.71 (m, 1H), 3.71 – 3.64 (m, 2H), 3.42 (t, *J* = 9.6, 9.6 Hz, 1H), 1.95 (s, 3H).

^{31}P NMR (202 MHz, D_2O) δ 43.25 (d, *J* = 27 Hz), -13.99 (d, *J* = 27 Hz).

^{13}C NMR (126 MHz, D_2O) d 174.7, 166.5, 152, 141.8, 102.6, 94.4 (d, $J_{C,P}$ = 5.9, C-1′′), 88.4 (C-1′), 83 (d, $J_{C,P}$ = 9.5, C-4′), 73.8 (C-2′/C-3′), 72.9 (C-5′′), 70.8 (C-3′′), 69.7 (C-2′/C-3′), 69.4 (C-4′′), 64.9 (d, $J_{C,P}$ = 6.0, C-5′), 60.2 (C-6′′), 53.7 (d, $J_{C,P}$ = 8.8, C-2′′), 22.2.

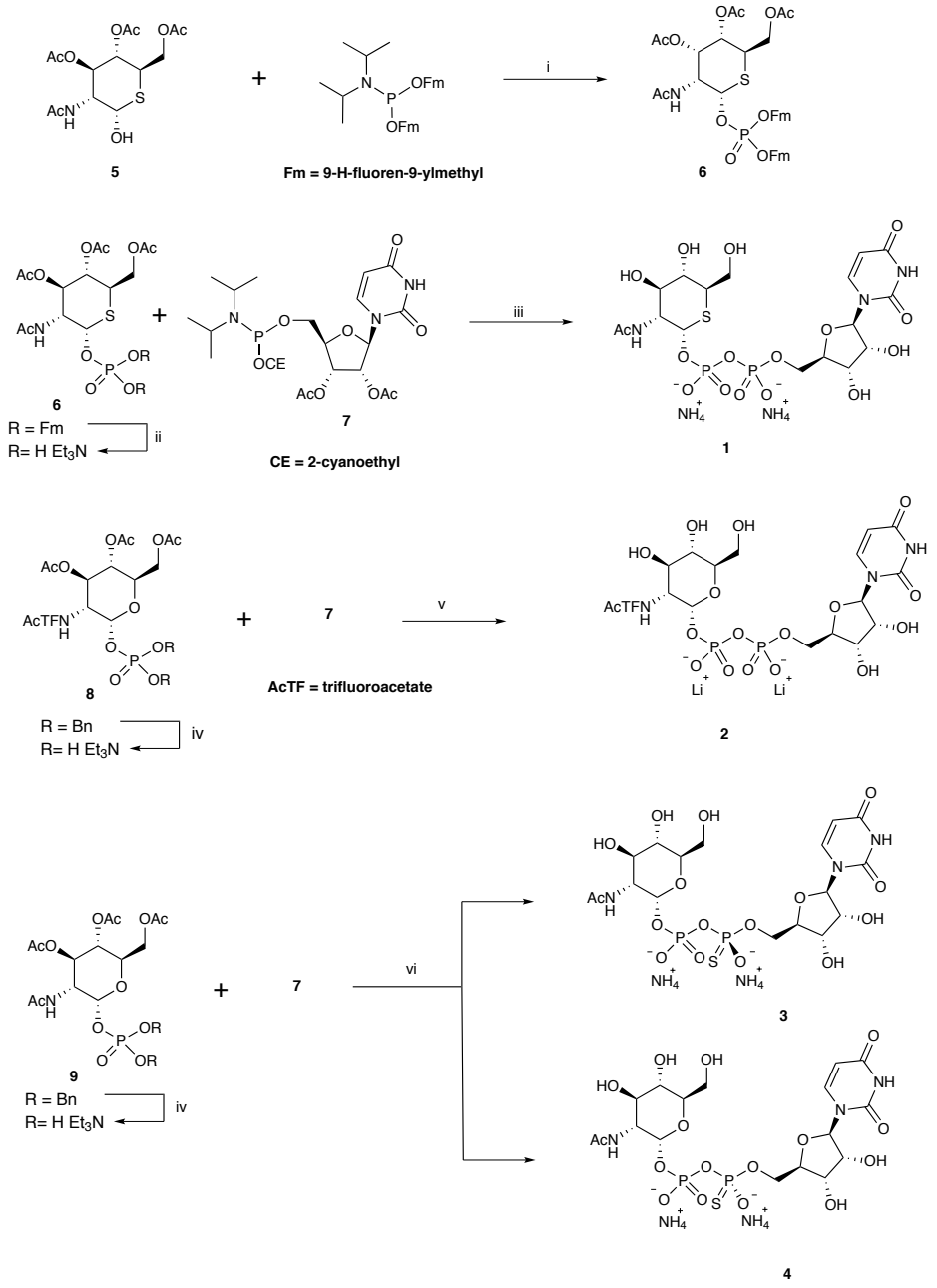
Uridine 5'-(2-acetamido-2-deoxy-α-D-glucopyranosyl)(1(R)-thiodiphosphate) bis ammonium salt (**4**):

^{1}H NMR (500 MHz, D_2O) δ 7.88 (d, *J* = 7.9 Hz, 1H), 5.83 (d, *J* = 3.0 Hz, 1H), 5.82 (d, *J* = 8.0 Hz, 1H), 5.43 – 5.38 (m, 1H), 4.23 (m, 2H), 4.16 (m, 2H), 4.09 (m, 1H), 3.82 (t, *J* = 12.2, 12.2 Hz, 2H), 3.75 – 3.62 (m, 3H), 3.41 (t, *J* = 9.2, 9.2 Hz, 1H), 1.95 (s, 3H).

^{31}P NMR (202 MHz, D_2O) δ 43.32 (d, *J* = 28.3 Hz), -14.01 (d, *J* = 28.3 Hz).

^{13}C NMR (126 MHz, D_2O) δ 174.7, 166.2, 151.8, 141.8, 102.6, 94.4 (d, $J_{C,P}$ = 6.1, C-1′′), 88.3 (C-1′), 83.1 (d, $J_{C,P}$ = 9.7, C-4′), 73.8 (C-2′/C-3′), 72.9 (C-5′′), 70.8 (C-3′′), 69.7 (C-2′/C-3′), 69.4 (C-4′′), 65.1 (d, $J_{C,P}$ = 5.8, C-5′), 60.2 (C-6′′), 53.6 (d, $J_{C,P}$ = 8.9, C-2′′), 22.3.

Supplementary Scheme 1: Reaction flow diagram for the synthesis of UDP-GlcNAc analogs



Reagents and conditions:

i) (a) 1*H*-tetrazole, MeCN, 2 h, RT; (b) t-BuOOH, 45 min, 0 °C, 91 %;
ii) DCM, 20 % Et_3N, 36 h, RT;
iii) (a) 4,5-dicyanoimidazole, MeCN, 45 min, RT; (b) t-BuOOH, 0 °C to RT, 45 min; (c) DBU, 45 min, RT; (d) MeOH:H_2O:Et_3N (5:2:1), 16 h, RT, 58 %;
iv) H_2 20 psi, Pd-C 10 %, MeOH, 1h, RT;
v) (a) 4,5-dicyanoimidazole, MeCN, 45 min, RT; (b) t-BuOOH, 0 °C to RT, 45 min; (c) DBU, 45 min, RT; (d) guanidine free base, MeOH, 70 min, RT, 46 %.
vi) (a) 4,5-dicyanoimidazole, MeCN, 45 min, RT; (b) 3H-1,2-benzodithiol-3-one 1,1-dioxide, 5 min, RT; (c) DBU, 45 min, RT; (d) MeOH:H_2O:Et_3N (5:2:1), 16 h, RT, about 100 % mixture of R_P, S_P isomers.

Enzymatic synthesis of S_P-αS-UDP-GlcNAc

S_P-αS-UDP-GlcNAc (**3**) was produced by stereoselective enzymatic synthesis from the diastereometrically pure R_P isomer of Uridine-5'-1-(thiotriphosphate) (R_P-UTP-αS, Biolog, Germany; > 95 % purity according to the manufacturer) and GlcNAc-1-phosphate as described by Zhao *et al.*[10] *Aspergillus fumigatus* UDP-*N*-acetylhexosamine pyrophosphorylase (*Af*UAP1) was a kind gift from Dr. Olawale G. Raimi, Lagos State University, Nigeria. 15 mM GlcNAc-1P, 10 mM R_P-UTP-α-S and 1 mg/ml *Af*UAP1 were incubated in 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 2 mM DTT and 10 % glycerol for 4 h at 37 °C.

Supplementary Figure 9: Partial ^{1}H NMR spectra of synthetic and enzymatically produced αS-UDP-GlcNAc analogs. Area of the most characteristic difference between S_P-αS-UDP-GlcNAc **3** and R_P-αS-UDP-GlcNAc **4** is highlighted. Overlaid partial 1H NMR spectra of enzymatically produced compound **3** (a); synthetic compound **3** (b); synthetic compound **4** (c).

Area of the most characteristic difference between **3** and **4** is highlighted. Additional differences in the spectra are caused by differences in experimental conditions (sample concentrations and purity, buffer composition) due to limiting (<0.5 mg) quantities of enzymatically produced compound **3** being available. In particular, the AB multiplet in the leftmost part of the spectrum which corresponds to the H-2 and H-3 protons of the ribose, is a non-first order spectrum and therefore particularly dependent on experimental conditions.

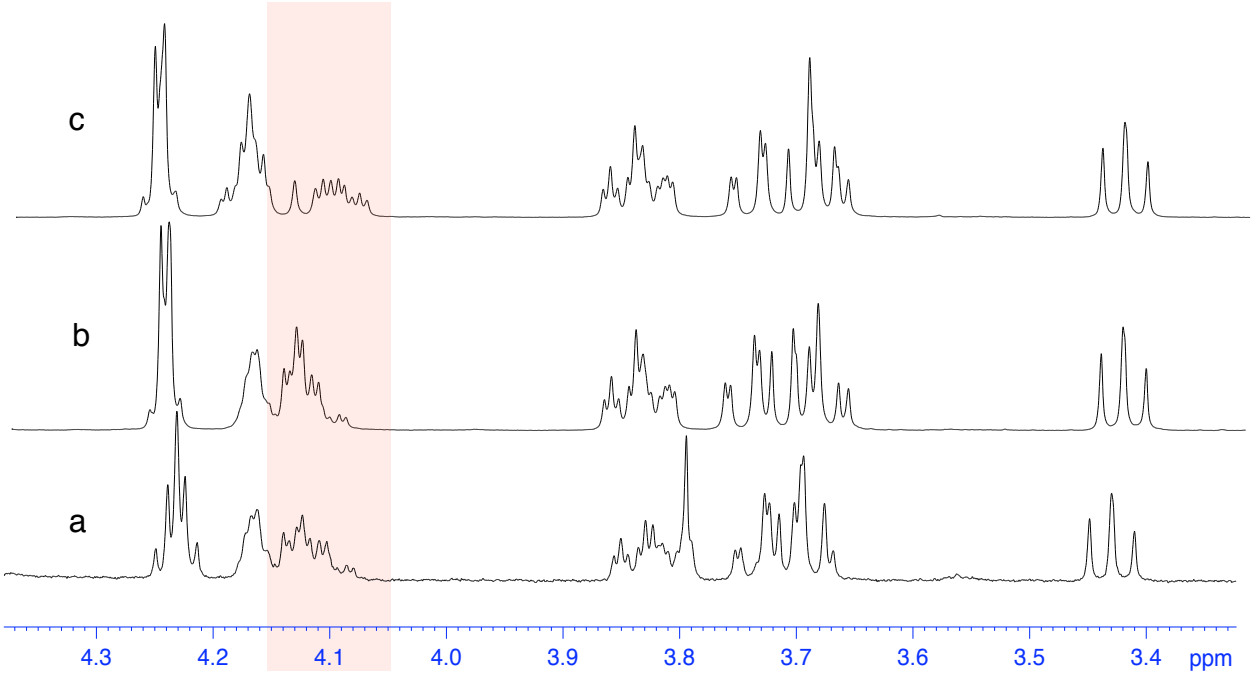
Conversely, the highlighted multiplets in (a), (b) and (c), which correspond to the pair of H-5 protons of the ribose, appear as a first order spectrum, less prone to concentration / buffer dependent variation. They are matching well for (a) and (b) and present a characteristically different appearance in (c), which is consistent with their proximity to the center of chirality (the phosphorous) – which is the key issue of consideration in these spectra.

Finally, the central and the rightmost parts of the spectra, which represent the protons of N-acetyl-glucosamine, look almost identical in all three spectra, with some impurities visible in spectrum (a).

A. Enzymatically produced S_P-αS-UDP-GlcNAc (**3**).

B. Synthetic compound **3**.

C. Synthetic compound **4**.

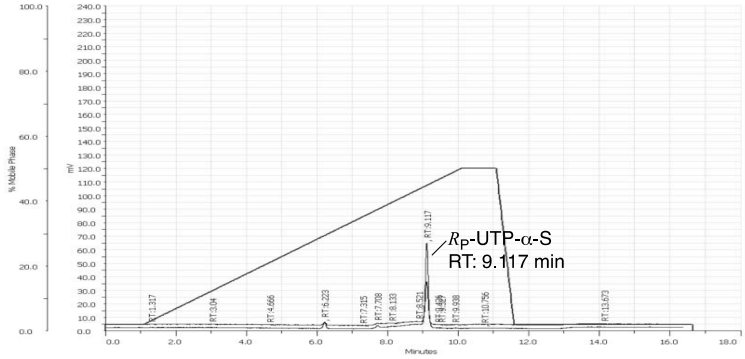


Supplementary Figure 10: Determination of the stereochemistry of αS-UDP-GlcNAc diastereoisomers in reference to the product of (stereospecific) enzymatic synthesis

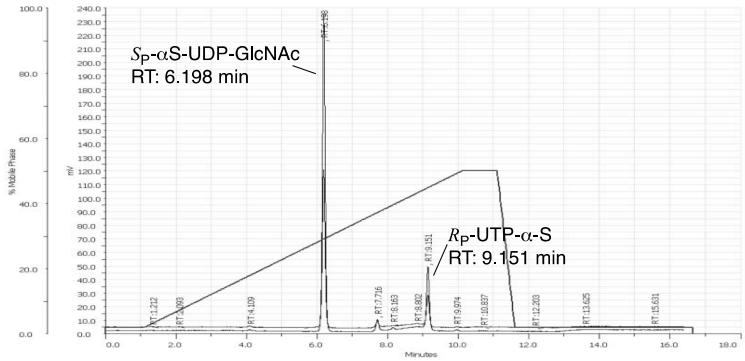
The chemically synthesized diastereoisomers of α-phosphorothioate analogs of UDP-GlcNAc were separable by ion-pair reversed-phase HPLC following a procedure described in [9]. The configuration at the α-phosphorus atom was unambiguously established by comparison of the retention times of the synthetic compounds with the product of enzymatic pyrophosphorylation catalyzed by *Af*UAP1, starting from optically pure R_P-UTP-α-S: R_P-UTP-α-S + GlcNAc-1-P -------> S_P-αS-UDP-GlcNAc (**3**) + PP_i.

Chromatograms showing the UV absorption at 214 / 220 nm are shown below.

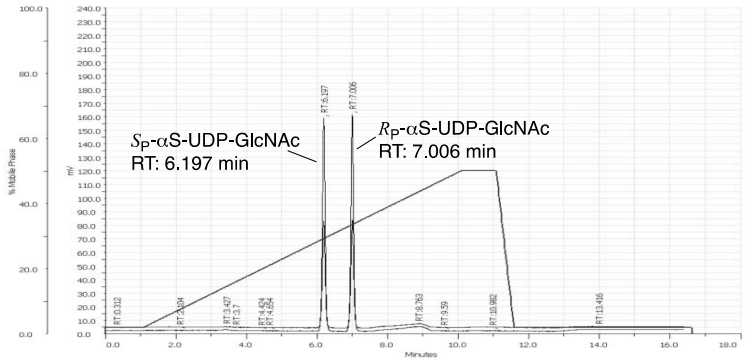
A Enzymatic reaction substrate R_P-UTP-α-S



B Enzymatic reaction products S_P-αS-UDP-GlcNAc and R_P-UTP-α-S



C Chemical synthesis products S_P and and R_P isomers of αS-UDP-GlcNAc

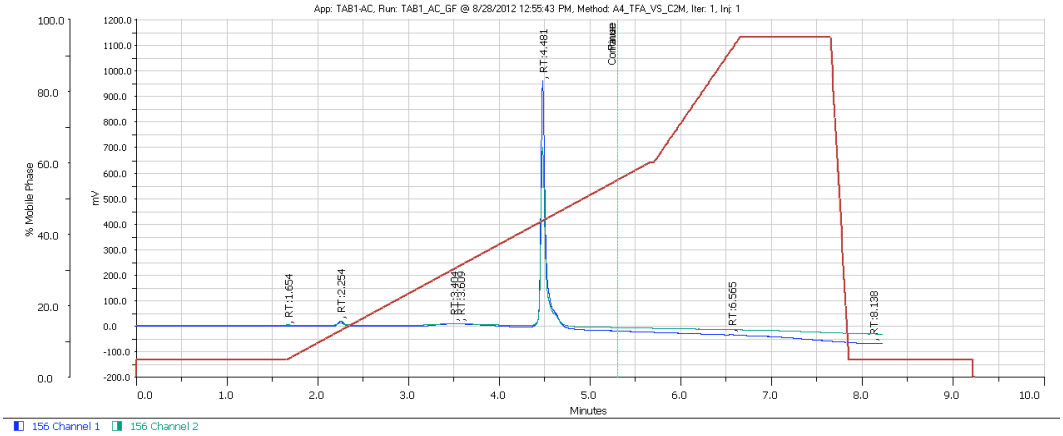


Peptide Synthesis

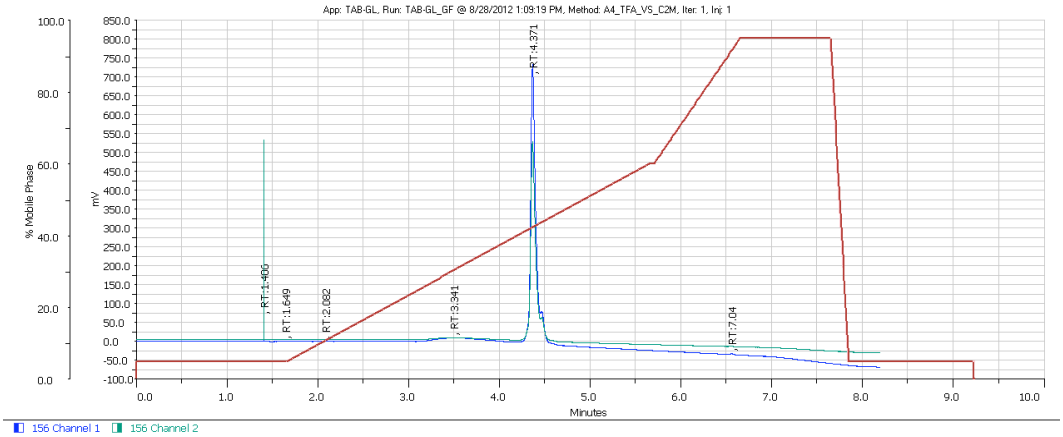
Microwave-assisted solid phase peptide synthesis was performed with a CEM Liberty automated peptide synthesizer on Rink amide MBHA resin (Novabiochem) using standard Fmoc chemistry protocols. All peptides were purified by reverse-phase HPLC. N-terminal biotin tag was introduced by adding a 13-atom PEG spacer, otherwise, all peptides were N-terminally acetylated. For the O-GlcNAc peptide (gTAB1tide), a 3,4,6-triacetyl-O-GlcNAc-Fmoc-SerOH building block was synthesized as described previously[11]. Purity was > 90 % for all peptides (see Supplementary Figure 11).

Supplementary Figure 11: HPLC profiles and MS information for synthetic peptides

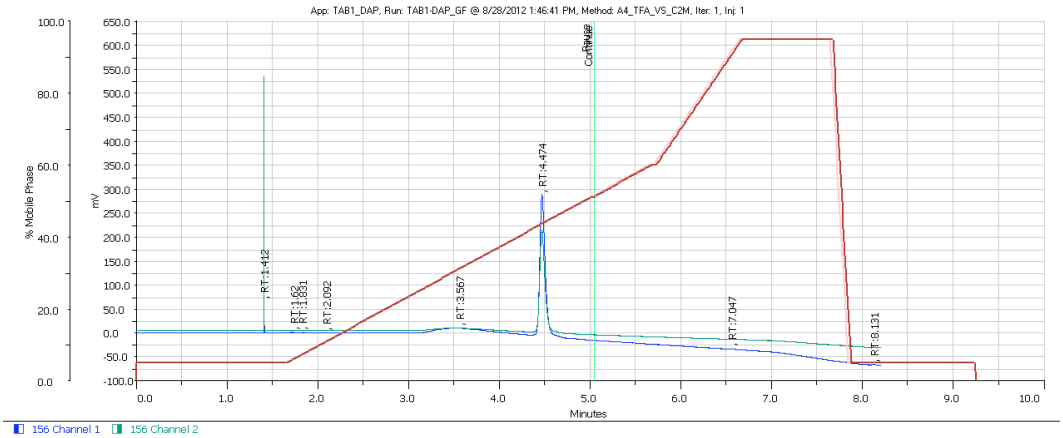
TAB1tide (Ac-Pro-Val-Ser-Val-Pro-Tyr-Ser-Ser-Ala-Gln-Ser-Thr-Ser-NH_2)
Measured $[M+H]^+$ = 1350.6603 Expected for $C_{58}H_{92}N_{15}O_{22}$ = 1350.6536 Error -5.0 ppm



aaTAB1tide (Ac-Pro-Val-Ser-Val-Pro-Tyr-AA-Ser-Ala-Gln-Ser-Thr-Ser-NH_2) AA = aminoalanine
Measured $[M+H]^+$ = 1349.6774 Expected for $C_{58}H_{93}N_{16}O_{21}$ = 1349.6696 Error -5.8 ppm



gTAB1tide (Ac-Pro-Val-Ser-Val-Pro-Tyr-[GlcNAc-Ser]-Ser-Ala-Gln-Ser-Thr-Ser-NH_2)
Measured $[M+H]^+$ = 1553.7428 Expected for $C_{66}H_{105}N_{16}O_{27}$ = 1553.7330 Error -6.4 ppm



References

1. van den Ent, F. & Löwe, J. RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67-74 (2006).

2. Gloster, T.M. et al. Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells. *Nat. Chem. Biol.* **7**, 174-181 (2011).

3. Bialy, L. & Waldmann, H. Total synthesis and biological evaluation of the protein phosphatase 2A inhibitor cytostatin and analogues. *Chemistry - A European Journal* **10**, 2759-2780 (2004).

4. Gold, H. et al. Synthesis of sugar nucleotides by application of phosphoramidites. *J. Org. Chem.* **73**, 9458-9460 (2008).

5. 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).

6. Gross, B.J., Kraybill, B.C. & Walker, S. Discovery of O-GlcNAc transferase inhibitors. *J. Am. Chem. Soc.* **127**, 14588-14589 (2005).

7. Sim, M.M., Kondo, H. & Wong, C.H. Synthesis and use of glycosyl phosphites: An effective route to glycosyl phosphates, sugar nucleotides, and glycosides. *J. Am. Chem. Soc.* **115**, 2260-2267 (1993).

8. Iyer, R.P., Phillips, L.R., Egan, W., Regan, J.B. & Beaucage, S.L. The automated synthesis of sulfur-containing oligodeoxyribonucleotides using 3H-1,2-benzodithiol-3-one 1,1-dioxide as a sulfur-transfer reagent. *J. Org. Chem.* **55**, 4693-4699 (1990).

9. 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).

10. Zhao, G., Guan, W., Cai, L. & Wang, P.G. Enzymatic route to preparative-scale synthesis of UDP-GlcNAc/GalNAc, their analogues and GDP-fucose. *Nat. Protoc.* **5**, 636-646 (2010).

11. Salvador, L.A., Elofsson, M. & Kihlberg, J. Preparation of building blocks for glycopeptide synthesis by glycosylation of Fmoc amino acids having unprotected carboxyl groups. *Tetrahedron* **51**, 5643-5656 (1995).