Supporting Information

Targeted Covalent Inhibition of O-GlcNAc Transferase in Cells

Matthew Worth,[‡a] Chia-Wei Hu,[‡b] Hao Li,[b] Dacheng Fan,[b] Arielis Estevez,[b] Dongsheng Zhu,[b]

Ao Wang[b] and Jiaoyang Jiang*[b]

a. Department of Chemistry, University of Wisconsin–Madison, Madison, Wisconsin 53705, USA
b. Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin–Madison, Madison, Wisconsin 53705, USA. E-mail: jiaoyang.jiang@wisc.edu
‡ These authors contributed equally to this work

I. Supporting Information Figures

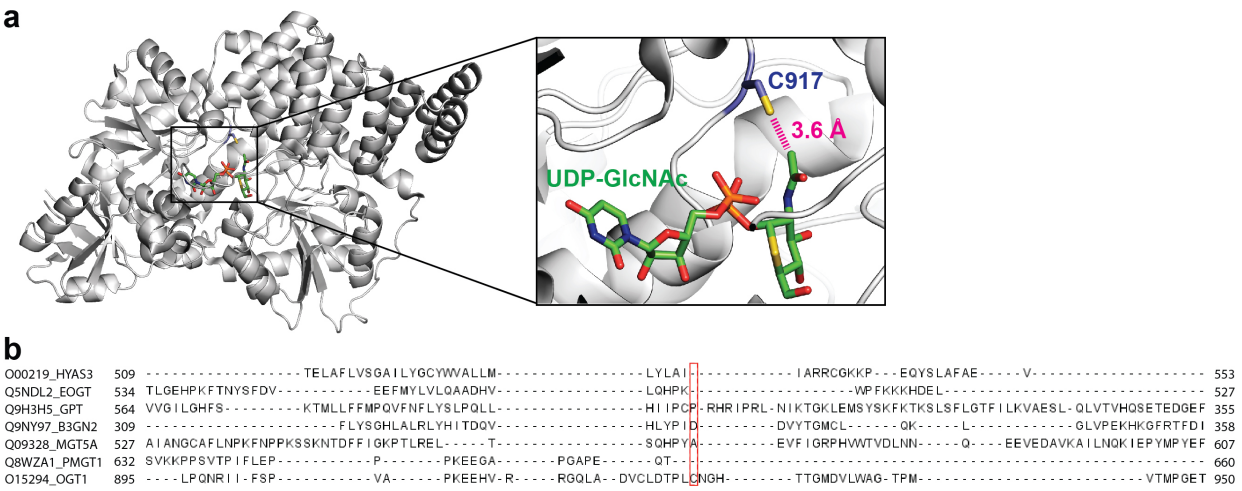


Figure S1. Structure-based rational design of ES compounds as targeted covalent inhibitors of OGT. (a) Based on the crystal structure of OGT:UDP-GlcNAc (PDB 4GZ5), the thiol of C917 residue of OGT is in close proximity (3.6 Å) to the *N*-acetyl group of UDP-GlcNAc.[1] (b) Sequence alignment of seven representative human glycosyltransferases using UDP-GlcNAc as sugar donor shows that C917 of OGT is rarely found in related glycosyltransferases. The C917 residue of OGT and its alignment with other glycosyltransferases are indicated in a red box.

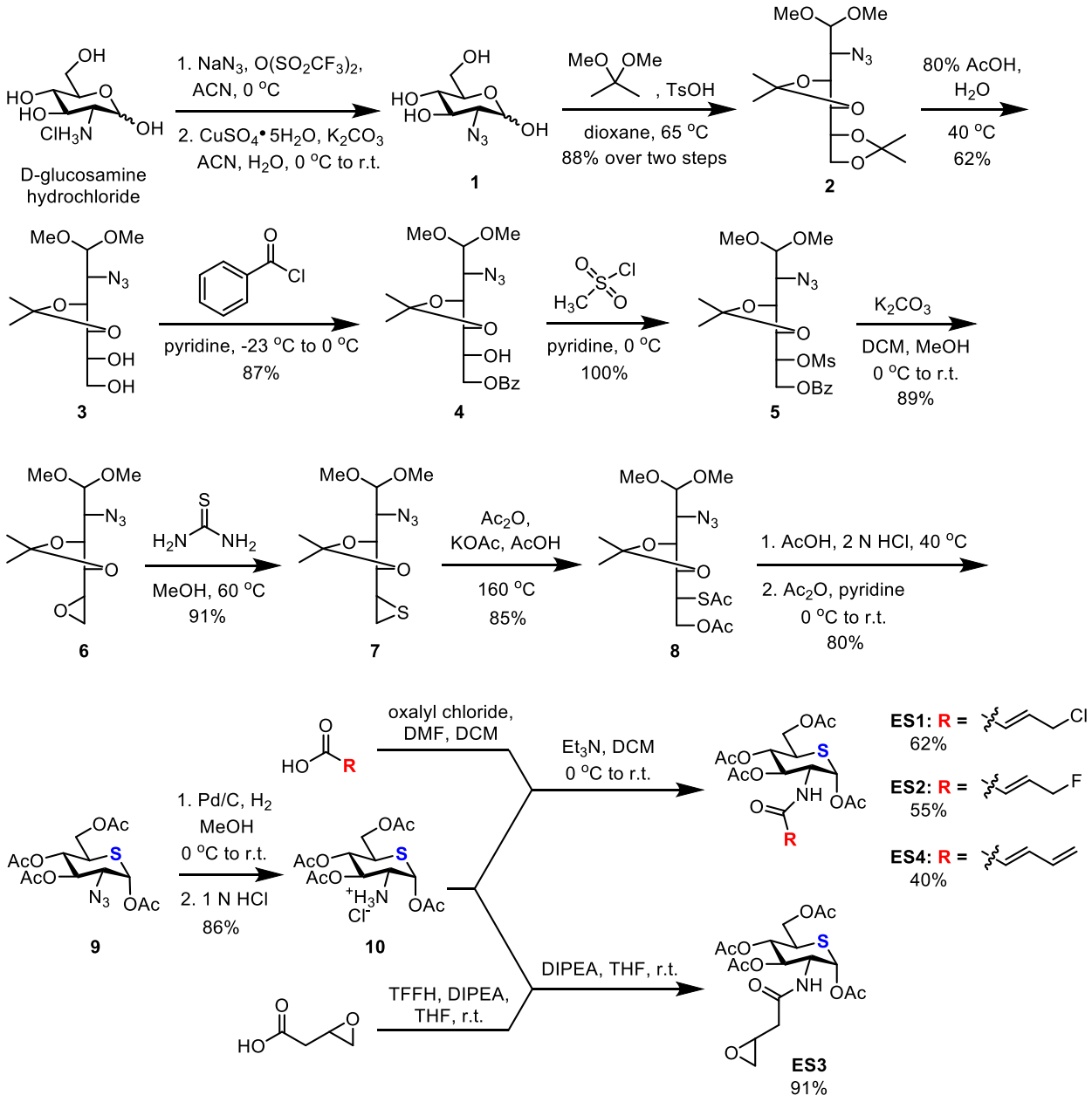


Figure S2. Synthetic scheme of cell-permeable inhibitors **ES1**-**ES4**.

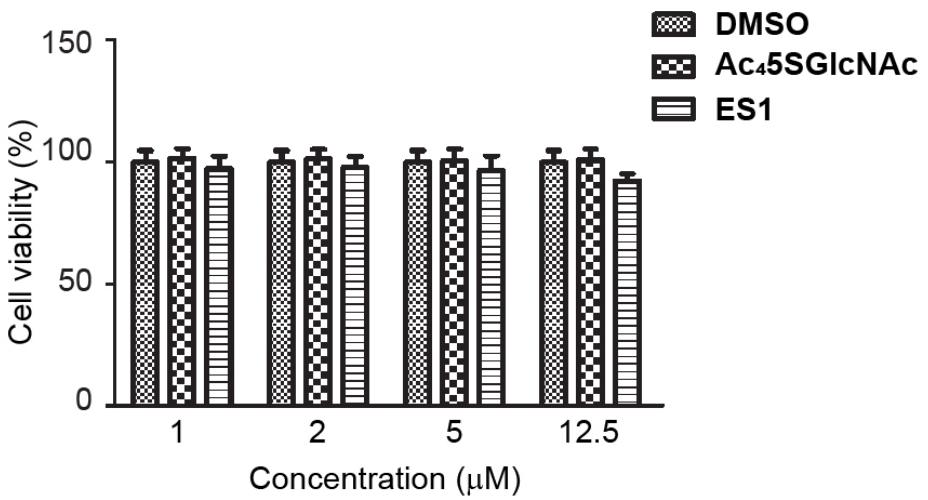


Figure S3. Cytotoxicity evaluation of **ES1** in MCF7 cells. Cell viability of MCF7 cells were measured after 13 h of pre-treatment with **ES1**, $Ac_4$5SGlcNAc, or DMSO and 8 h of post-incubation in inhibitor-free medium. Data represent the average values ± S.D. from five different wells.

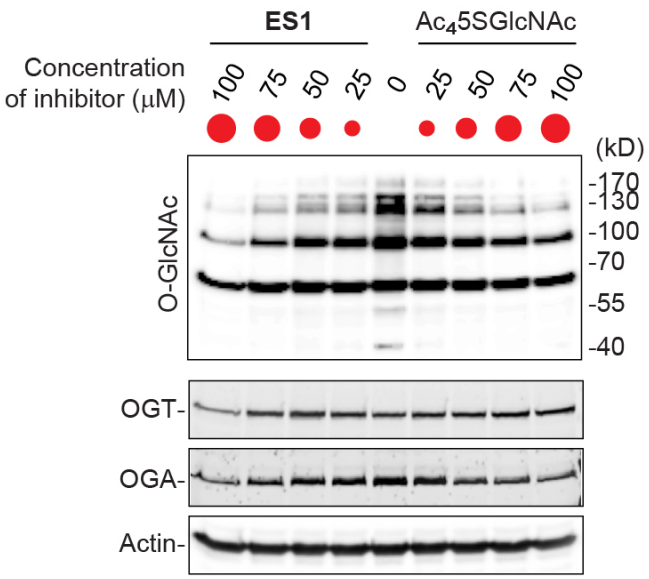


Figure S4. The global O-GlcNAc profile of COS-7 cells treated with **ES1** or $Ac_4$5SGlcNAc under pre-treatment (13 h) and post-incubation (8 h) condition. Protein O-GlcNAcylation was detected with RL2 antibody.

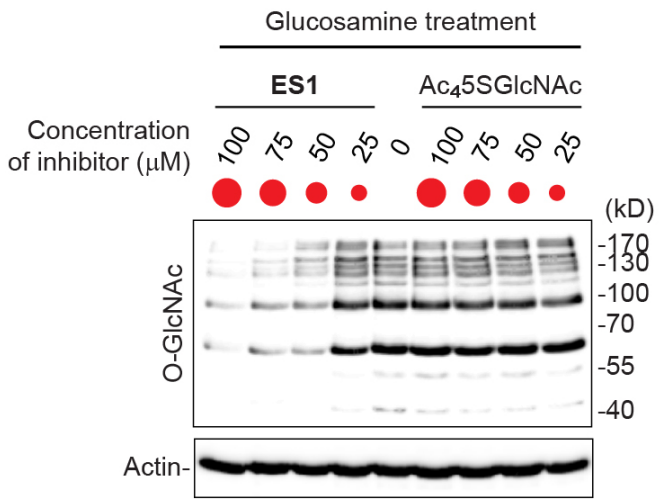


Figure S5. The global O-GlcNAc profile of COS-7 cells with inhibitor and glucosamine treatment. Cells were pre-treated with **ES1** or $Ac_4$5SGlcNAc for 13 h followed by 8 h of post-incubation in inhibitor-free medium. Glucosamine (10 mM) was added to the cells during post-incubation stage five hours before harvest. Protein O-GlcNAcylation was detected with RL2 antibody.

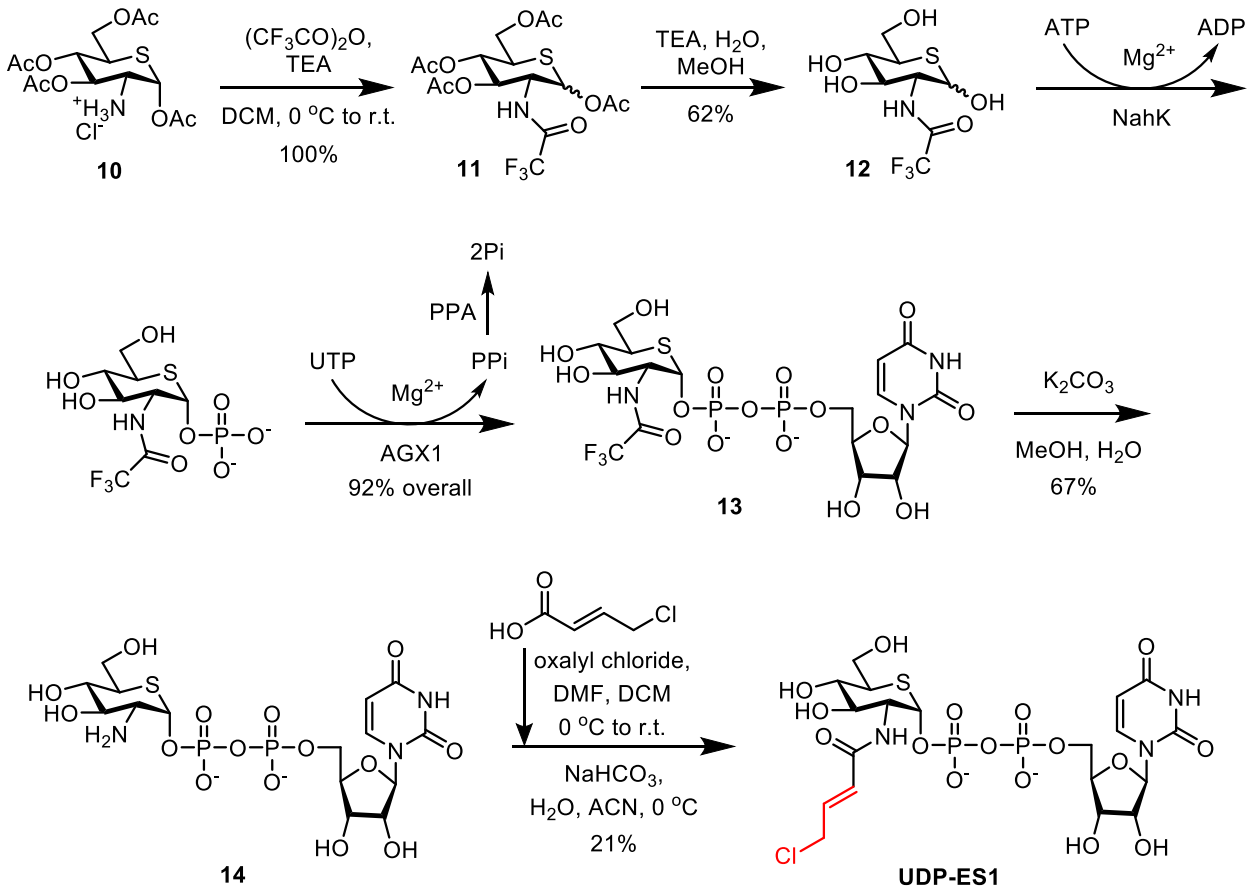


Figure S6. Synthetic scheme of **UDP-ES1**.

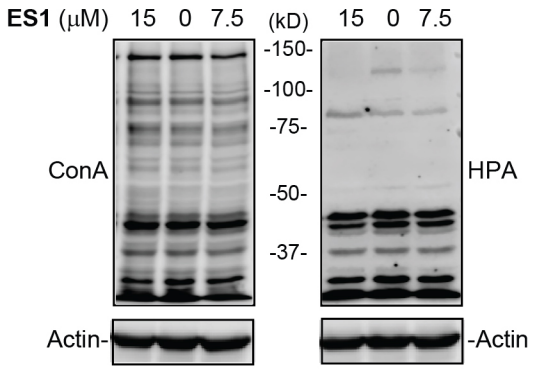


Figure S7. Lectin blots demonstrate the expression of glycans in MCF7 cells treated with **ES1**. Cells were treated with **ES1** at the indicated concentrations for 24 h. Lectin ConA and HPA were used to detect the N-linked and O-linked glycans, respectively.

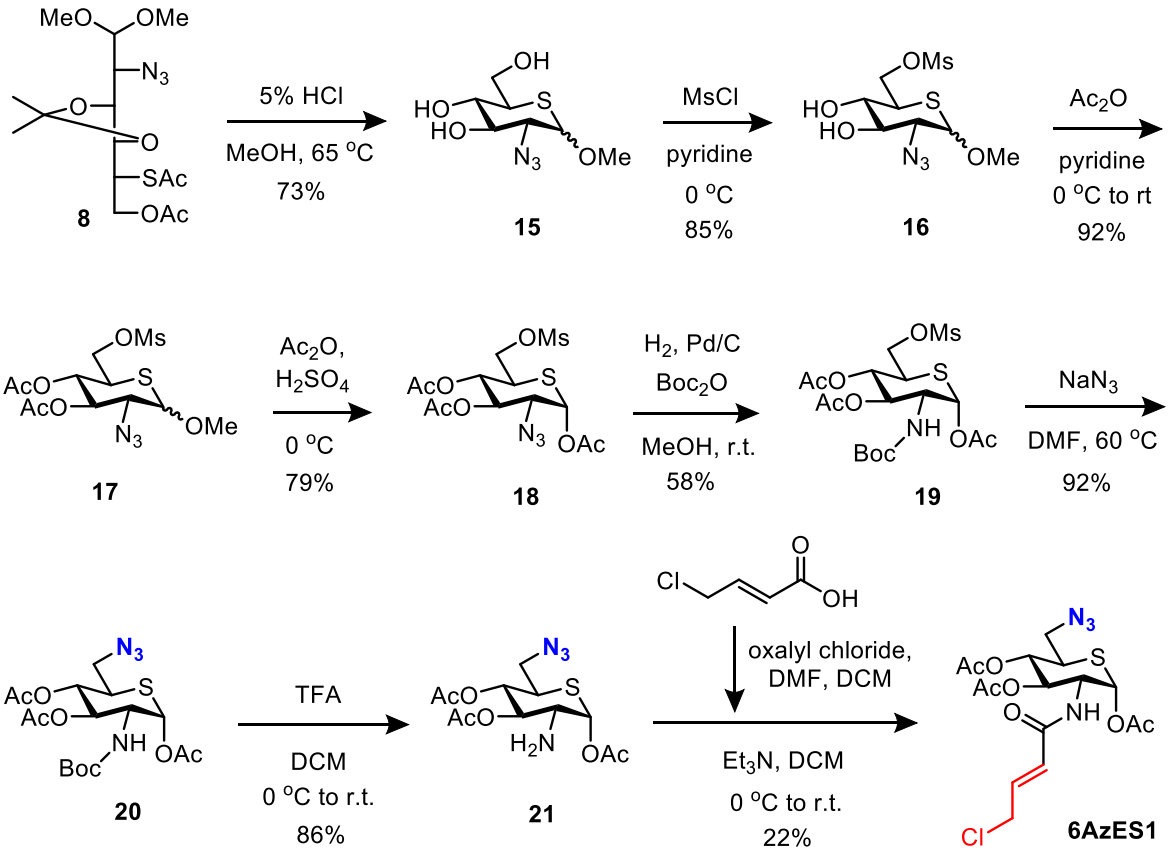


Figure S8. Synthetic scheme of **6AzES1**.

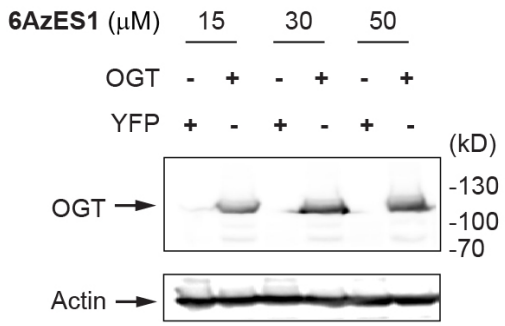


Figure S9. Validation of exogenous expression of OGT in **6AzES1**-treated HEK293 cells. Cells were transfected with flag-OGT (or YFP as a control) for 48 h followed by **6AzES1** treatment at indicated concentration for 24 h. Flag-OGT detected by anti-flag antibody corresponds to the major **6AzES1**-labeling band shown in the in-gel fluorescence of Figure 4b.

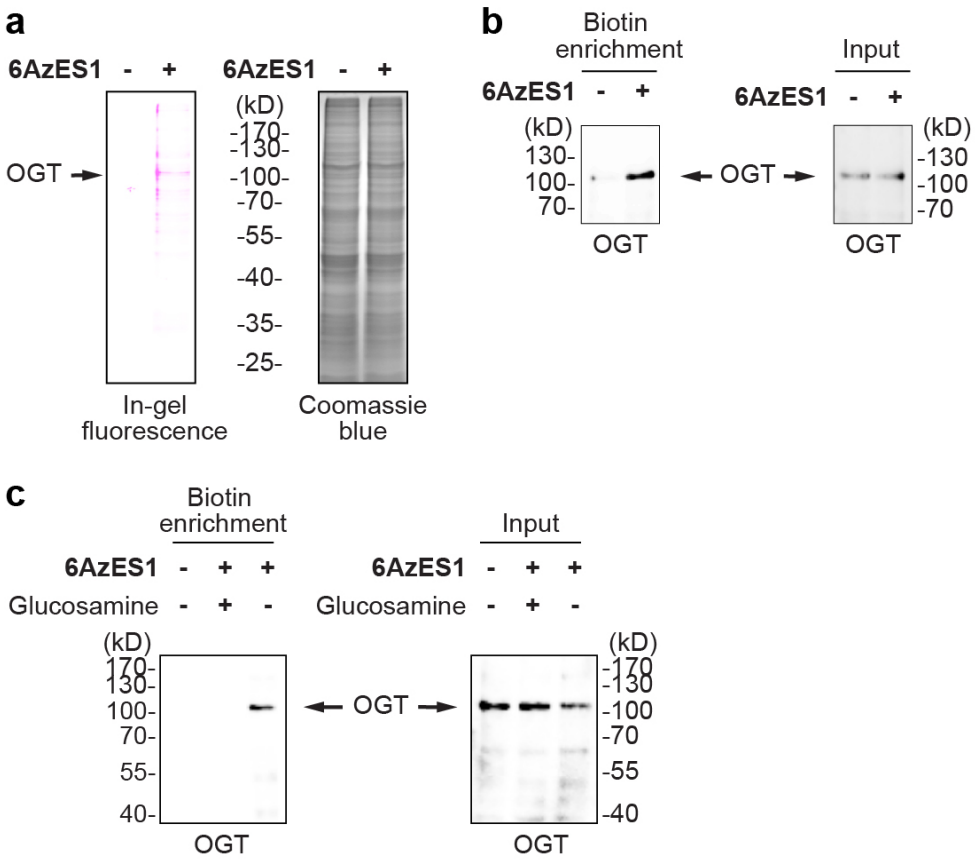


Figure S10. 6AzES1 targets endogenous OGT in cells. (a) HEK293 cells were treated with **6AzES1** (25 μM) for 24 h. The **6AzES1**-labeled proteins were coupled with fluorescent alkyne and detected by in-gel fluorescence. Coomassie Blue staining of the same gel shows the protein loading amount. (b) The **6AzES1**-modified proteins from the same batch of sample in (a) were conjugated with alkyne-biotin and enriched by NeutrAvidin. The OGT protein identity was validated by Western blot using anti-OGT antibody. (c) To evaluate the on-target engagement of **6AzES1**, glucosamine (15 mM) was added eight hours before and during **6AzES1** treatment. **6AzES1**-modified proteins were enriched and detected similarly as described in (b).

II. Supporting Information Tables

Table S1. Data collection and refinement statistics (molecular replacement).

	OGT:UDP-ES1:CKII
Data collection	
Space group	*F*222
Cell dimensions	
a, *b*, *c* (Å)	137.7, 151.0, 198.9
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	50.0–2.55 (2.64–2.55) *
R_{merge}	9.4 (51.2)
$I / \sigma I$	16 (2.5)
$CC_{1/2}$	99.2 (93.2)
Completeness (%)	99.8 (99.9)
Redundancy	13.3 (12.1)
Refinement	
Resolution (Å)	50.0–2.55
No. reflections	34,179
R_{work} / R_{free}	16.7 / 21.1
No. atoms	
Protein	5,563
Ligand	41
Water	191
B-factors	40.01
Protein	40.21
Ligand	29.68
Water	36.38
R. m. s. deviation	
Bond length (Å)	0.008
Bond angle (°)	0.968

*Each structure was determined from one crystal.
*Values in parentheses are for highest-resolution shell.

Table S2. List of primers used in this study.

Primer	**Sequence (5´-3´)**
Nahk_fwd	ACCCCATATGAACAACACCAATGAAGCCCTG
Nahk_rev	TGACCTCGAGCTTGGTCGTCTCCATGACGTCG
AGX1_fwd	ACTGCATATGAACATTAATGACCTCAAACTC
AGX1_rev	CGCGGATCCTCAAATACCATTTTTCACCAGC
OGT_C917S_fwd	CTGGACACCCCGCTGAGCAACGGTCACACC
OGT_C917S_rev	GGTGTGACCGTTGCTCAGCGGGGTGTCCAG
ncOGT_fwd	CGGGGTACCATGGCGTCTTCCGTGGGCA
ncOGT_rev	CGCGGGCCCCTACTTGTCATCGTCATCCTTGTAGTCGATGTCATG
NUP62_fwd	GGGGTACCATGAGCGGCTTTAATTTT
NUP62_rev	GCTCTAGAGTCAAAGGTGATCCGGAA

III. Supporting Information Methods

General Chemical Synthesis

All reagents were purchased from Sigma-Aldrich, MP Biomedicals, Alfa Aesar, TCI, or Thermo Fisher Scientific. Unless otherwise stated, all reactions were performed in flame-dried, round-bottomed flasks fitted with rubber septa under nitrogen atmosphere. Reagent and solvent concentrations presented as "%" were defined as (v/v) unless otherwise indicated. Reactions were monitored with thin layer chromatography (TLC) unless otherwise stated. TLC was conducted on silica gel plates (0.25 mm, 60 Å pore size) with fluorescent indicator on glass (254 nm) and visualized by ceric ammonium molybdate (CAM), basic $KMnO_4$, or UV light. Nuclear magnetic resonance (NMR) spectra were obtained on Varian UI400, UI500 or Bruker AV400 spectrometers. Spectra were recorded at 400 or 500 MHz for ^{1}H NMR, 101 or 126 MHz for ^{13}C NMR, 162 MHz for ^{31}P NMR, and 376 or 470 MHz for ^{19}F NMR. Chemical shifts are expressed in parts per million (ppm, δ scale) and referenced to $CDCl_3$ or D_2O solvent peaks. Data for ^{1}H NMR spectra are reported as follows: chemical shift (ppm), multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; ddd, doublet of doublet of doublets; dt, doublet of triplets; td, triplet of doublets; br s, broad singlet), coupling constant (Hz), and integration. ^{13}C NMR spectra are expressed as chemical shifts, followed by multiplicity and coupling constant (Hz) where appropriate. ^{31}P NMR spectra are represented as follows: chemical shifts, multiplicity, and coupling constant (Hz). ^{19}F NMR spectra are represented as chemical shift. High-resolution mass spectra (HRMS) were obtained from Q-TOF Maxis 4G (Bruker). Analysis of low-resolution MS (LRMS) and compound purity was performed on 6120 Quadrupole (Agilent) coupled with 1290 UHPLC (Agilent). $Ac_4$5SGlcNAc was synthesized as previously described.[2]

Synthesis of ES1-4, UDP-ES1, and 6AzES1

2-azido-2-deoxy-D-glucose (**1**)[3]: NaN_3 (1.45 g, 22.2 mmol) was suspended in dry ACN (26 mL) and cooled to 0 °C. Trifluoromethane sulfonic anhydride (2.85 mL, 16.9 mmol) was slowly added over 20 minutes and then stirred at 0 °C for 5 h. A solution of D-glucosamine-hydrochloride (3.01 g, 13.9 mmol), Cu(II)SO_4·5H_2O (35 mg, 0.139 mmol), K_2CO_3 (3.84 g, 27.8 mmol) and water (15 mL) was prepared. The azide solution was added through a glass wool filter to the glucosamine solution at 0 °C. The mixture was stirred on ice for 1 h, then at room temperature (r.t.) for 19 h, and the solution was concentrated *in vacuo*. The product was purified by column chromatography on silica gel (DCM/MeOH, 2:1) and dried down to a white solid (yield determined after second step due to over-quantitative yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.35 (d, J = 3.6 Hz, 0.4H), 4.71 (d, J = 8.2 Hz, 0.6H), 3.92─3.72 (m, 3H), 3.54─3.44 (m, 2.4H), 3.28 (dd, J = 9.6, 8.2 Hz, 0.6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 95.39, 91.49, 76.31, 74.67, 71.90, 71.78, 70.11, 69.79, 67.21, 63.87, 60.92, 60.77.

2-azido-2-deoxy-3,4,5,6-di-*O*-isopropylidene-aldehydo-D-glucose dimethyl acetal (**2**): Compound **1** (2.8 g, 13.9 mmol), dry 1,4-dioxane (31 mL), 2,2-dimethoxypropane (28 mL, 228 mmol), and *p*-toluene sulphonic acid (600 mg, 3.5 mmol) were combined in a dry flask and stirred at 65 °C for 6 h. The conversion of substrate was monitored by ^{1}H NMR. The flask was cooled to r.t. and the mixture was concentrated *in vacuo*. The residue was taken up in 60 mL EtOAc and washed with saturated $NaHCO_3$ (60 mL) and water (60 mL). The organic layer was dried over Na_2SO_4. The crude product was purified by column chromatography on silica gel (hexanes/EtOAc, 98:2) and dried down to give **2** as a thick amber oil (4.07 g, 88% yield over two steps). ^{1}H NMR

(500 MHz, $CDCl_3$) δ 4.60 (d, J = 8.2 Hz, 1H), 4.15─4.11 (m, 2H), 4.03 (m, 1H), 3.97 (dd, J = 8.5, 4.4 Hz, 1H), 3.94 (t, J = 8.4 Hz, 1H), 3.51 (m, 1H), 3.50 (s, 3H), 3.46 (s, 3H), 1.44 (s, 3H), 1.40 (s, 3H), 1.36 (s, 3H), 1.32 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 110.07, 109.93, 103.93, 78.80, 77.46, 77.32, 68.06, 61.65, 55.01, 54.85, 27.24, 26.85, 26.75, 25.36. LRMS: *m/z* 304.1 $[M-2N+H]^+$.

2-azido-2-deoxy-3,4-*O*-isopropylidene-aldehydo-D-glucose dimethyl acetal (**3**): Compound **2** (4.07 g, 12.3 mmol) and 80% acetic acid in water (43 mL) were combined in a flask and stirred at 40 °C for 7 h. The mixture was concentrated *in vacuo* and purified by column chromatography on silica gel (hexanes/EtOAc, 1:1) to give **3** as a thick oil (2.20 g, 62% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.62 (d, J = 7.9 Hz, 1H), 4.22 (dd, J = 7.9, 1.9 Hz, 1H), 4.02 (t, J = 7.6 Hz, 1H), 3.83 (dd, J = 11.0, 3.2 Hz, 1H), 3.74 (m, 1H), 3.68 (dd, J = 11.0, 5.5 Hz, 1H), 3.51 (dd, J = 8.0, 1.8 Hz, 1H), 3.48 (s, 3H), 3.46 (s, 3H), 2.77 (br s, 1H), 2.28 (br s, 1H), 1.45 (s, 3H), 1.37 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 110.09, 103.70, 78.14, 77.17, 73.05, 64.17, 61.69, 55.10, 54.41, 27.34, 26.82. LRMS: *m/z* 314.1 $[M+Na]^+$.

2-azido-6-*O*-benzoyl-2-deoxy-3,4-*O*-isopropylidene-aldehydo-D-glucose dimethyl acetal (**4**): Compound **3** (3.79 g, 13.0 mmol) was dissolved in dry pyridine (45 mL) and cooled to -23 °C in a dry ice/CCl_4 bath. Benzoyl chloride (1.6 mL, 13.78 mmol) was added, and the solution was stirred at -23 °C for 10 minutes and then in a NaCl/ice bath at -12–0 °C for 5 h. The reaction mixture was concentrated *in vacuo* and purified by column chromatography on silica gel (hexanes/EtOAc, 3:1) to give **4** as a thick oil (4.47 g, 87% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.05 (dd, J = 8.4, 1.2 Hz, 2H), 7.58 (tt, J = 7.4, 1.2 Hz, 1H), 7.45 (t, J = 7.8 Hz, 2H), 4.64 (d, J = 8.0 Hz, 1H), 4.62 (dd, J = 11.9, 2.8 Hz, 1H), 4.42 (dd, J = 11.9, 6.6 Hz, 1H), 4.28 (dd, J = 7.7, 1.8 Hz, 1H), 4.08 (t, J = 7.8 Hz, 1H), 4.00 (ddd, J = 8.3, 6.5, 2.8 Hz, 1H), 3.53 (dd, J = 8.0, 1.9 Hz, 1H), 3.46 (s, 3H), 3.44 (s, 3H), 2.88 (br s, 1H), 1.47 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.47, 133.55, 129.92, 129.67, 128.65, 110.44, 103.51, 78.66, 76.30, 72.73, 67.45, 61.74, 54.84, 54.28, 27.34, 26.86. LRMS: *m/z* 418.0 $[M+Na]^+$.

2-azido-6-*O*-benzoyl-2-deoxy-3,4-*O*-isopropylidene-5-*O*-mesyl-aldehydo-D-glucose dimethyl acetal (**5**): Compound **4** (256 mg, 0.65 mmol) was dissolved in pyridine (3 mL) and cooled to 0 °C. Methanesulfonyl chloride (360 µL, 4.55 mmol) was added dropwise to the solution and then stirred at 0 °C for 3 h. The reaction mixture was concentrated *in vacuo*. The residue was dissolved in chloroform (20 mL), washed with 2 N HCl (2x 20 mL), and then saturated $NaHCO_3$ (20 mL). The aqueous phases were back extracted with chloroform. The combined organic phase was concentrated to give **5** as a colorless oil with no further purification (306 mg, 100% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.07 (d, J = 8.4 Hz, 2H), 7.59 (tt, J = 7.4, 1.3 Hz, 1H), 7.46 (t, J = 7.7 Hz, 2H), 5.07 (td, J = 6.6, 2.8 Hz, 1H), 4.78 (dd, J = 12.6, 2.9 Hz, 1H), 4.62 (d, J = 7.6 Hz, 1H), 4.47 (dd, J = 12.6, 6.8 Hz, 1H), 4.37─4.32 (m, 2H), 3.48 (s, 3H), 3.47 (s, 3H), 3.42 (dd, J = 7.6, 1.9 Hz, 1H), 3.08 (s, 3H), 1.48 (s, 3H), 1.43 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.15, 133.61, 129.94, 129.45, 128.74, 111.17, 103.68, 79.26, 77.74, 75.27, 63.65, 61.68, 55.51, 54.37, 39.18, 27.19, 26.84. LRMS: *m/z* 496.0 $[M+Na]^+$.

5,6-anhydro-2-azido-2-deoxy-3,4-*O*-isopropylidene-aldehydo-D-glucose dimethyl acetal (**6**): Compound **5** (264 mg, 0.56 mmol) was dissolved in a mixture of DCM (4 mL) and MeOH (4 mL), and cooled to 0 °C. Potassium carbonate (360 mg, 2.80 mmol) was added to the solution as a solid and the reaction was warmed to r.t. and stirred for 5 h. Diethyl ether was added into the solution and the mixture was filtered through glass wool. The filtrate was washed with brine (30 mL) and

the aqueous phase was extracted with DCM three times. The combined organic phase was dried over Na_2SO_4 and concentrated *in vacuo* to get the crude product as a colorless oil. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 4:1), and dried down to give **6** as a colorless oil (136 mg, 89% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.59 (d, J = 7.4 Hz, 1H), 4.19 (dd, J = 8.3, 2.6 Hz, 1H), 3.97 (dd, J = 8.3, 5.0 Hz, 1H), 3.49 (s, 3H), 3.48 (s, 3H), 3.21 (dd, J = 7.4, 2.5 Hz, 1H), 3.03 (ddd, J = 5.0, 4.1, 2.6 Hz, 1H), 2.80 (dd, J = 5.2, 4.1 Hz, 1H), 2.71 (dd, J = 5.2, 2.6 Hz, 1H), 1.44 (s, 3H), 1.39 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 110.58, 104.07, 77.55, 76.52, 61.52, 56.24, 54.68, 51.02, 44.05, 26.94, 26.84.

2-azido-2,5,6-trideoxy-5,6-epithio-3,4-*O*-isopropylidene-aldehydo-D-glucose dimethyl acetal (**7**): Compound **6** (2.14 g, 7.83 mmol) was dissolved in MeOH (40 mL) and thiourea (1.80 mg, 23.6 mmol) was added. The reaction mixture was heated to 60 °C and stirred for 3 h. The reaction mixture was concentrated *in vacuo*, purified by column chromatography on silica gel (hexanes/EtOAc, 8:1), and dried down to give **7** as a colorless oil (2.05 g, 91% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.61 (d, J = 8.0 Hz, 1H), 4.23 (dd, J = 8.0, 1.7 Hz, 1H), 3.50 (s, 3H), 3.48 (s, 3H), 3.46─3.44 (m, 1H), 3.38 (dd, J = 8.0, 1.7 Hz, 1H), 2.89 (ddd, J = 8.7, 6.0, 5.1 Hz, 1H), 2.58 (dd, J = 6.0, 1.5 Hz, 1H), 2.31 (dd, J = 5.1, 1.5 Hz, 1H), 1.46 (s, 3H), 1.45 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 110.49, 103.77, 82.25, 79.39, 61.54, 55.37, 54.53, 33.77, 27.25, 26.83, 23.77. LRMS: *m/z* 262.1 $[M\text{-}2N+H]^+$.

6-*O*-acetyl-5-*S*-acetyl-2-azido-2-deoxy-3,4-*O*-isopropylidene-5-thio-aldehydo-D-glucose dimethyl acetal (**8**): Compound **7** (120 mg, 0.41 mmol) was dissolved in a mixture of acetic anhydride (3.8 mL) and acetic acid (0.6 mL), and potassium acetate (186 mg, 2.07 mmol) was added to the solution. The reaction mixture was heated to 160 °C and stirred for 10 h. The reaction was concentrated *in vacuo* and dissolved in chloroform (10 mL), then washed with 2 N HCl (10 mL) followed by saturated $NaHCO_3$ (10 mL). The aqueous phases were back extracted with chloroform. The combined organic phase was dried over Na_2SO_4 and concentrated *in vacuo* to get the crude product as a colorless oil. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 8:1), and dried down to give **8** as a colorless oil (138 mg, 85% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.57 (d, J = 8.0 Hz, 1H), 4.38 (dd, J = 11.5, 4.3 Hz, 1H), 4.32 (dd, J = 11.5, 5.0 Hz, 1H), 4.19 (dd, J = 8.5, 7.4 Hz, 1H), 4.09 (dd, J = 7.4, 1.8 Hz, 1H), 3.92 (ddd, J = 8.6, 4.9, 4.3 Hz, 1H), 3.49 (s, 3H), 3.44 (s, 3H), 3.36 (dd, J = 8.0, 1.8 Hz, 1H), 2.36 (s, 3H), 2.07 (s, 3H), 1.45 (s, 3H), 1.42 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 193.72, 170.72, 110.59, 103.96, 78.31, 76.45, 64.01, 62.57, 55.83, 54.62, 44.96, 30.80, 27.27, 26.91, 20.93.

1,3,4,6-tetra-*O*-acetyl-2-azido-2-deoxy-5-thio-α-D-glucose (**9**): Compound **8** (336 mg, 0.86 mmol) was dissolved in a mixture of acetic acid (7.5 mL) and 2 N HCl aqueous solution (2.5 mL). The reaction mixture was heated to 40 °C and stirred for 17 h. The reaction was concentrated *in vacuo* and the residue was dissolved in pyridine (6 mL). Acetic anhydride (2 mL) was then added dropwise at 0 °C. The reaction mixture was warmed to r.t. and stirred for 21 h and concentrated *in vacuo*. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 85:15) and dried down to give **9** as a colorless oil (268 mg, 80% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.11 (d, J = 3.1 Hz, 1H), 5.40 (ddd, J = 10.1, 9.5, 0.4 Hz, 1H), 5.32 (dd, J = 10.7, 9.5 Hz, 1H), 4.40 (ddd, J = 12.2, 4.9, 0.4 Hz, 1H), 4.03 (dd, J = 12.2, 3.1 Hz, 1H), 3.89 (dd, J = 10.2, 3.1 Hz, 1H), 3.56 (ddd, J = 10.7, 4.9, 3.0 Hz, 1H), 2.19 (s, 3H), 2.11 (s, 3H), 2.07 (s, 3H), 2.06 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.58, 169.85, 169.65, 168.81, 71.89, 71.87, 71.85, 65.03, 61.04, 39.96, 21.11, 20.73, 20.72, 20.64.

1,3,4,6-tetra-*O*-acetyl-2-amino-2-deoxy-5-thio-α-D-glucose hydrochloride (**10**): Compound **9** (95 mg, 0.24 mmol) was dissolved in MeOH (2.5 mL) and Pd/C (10% loading, 30 mg) was added under N_2. The reaction was then stirred under 1 atm of H_2 at r.t. for 1 h. HCl aqueous solution (1 N, 27 μL) was added at 0 °C. The reaction mixture was filtered through a pad of Celite® and the filtrate was concentrated to give **10** as a white solid (84 mg, 86% yield). ^{1}H NMR (400 MHz, CD_3OD) δ 6.05 (d, J = 3.1 Hz, 1H), 5.33 (t, J = 10.1 Hz, 1H), 5.24 (dd, J = 10.7, 9.5 Hz, 1H), 4.39 (dd, J = 12.4, 4.6 Hz, 1H), 4.05 (dd, J = 10.6, 3.1 Hz, 1H), 3.97 (dd, J = 12.3, 2.9 Hz, 1H), 3.65 (ddd, J = 10.6, 4.3, 2.9 Hz, 1H), 2.16 (s, 3H), 2.04 (s, 3H), 1.98 (s, 3H), 1.95 (s, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 171.95, 171.49, 171.10, 170.10, 72.94, 72.01, 71.89, 61.94, 56.82, 40.79, 20.78, 20.73, 20.46, 20.44. LRMS: *m/z* 386.0 $[M+Na]^+$.

1,3,4,6-tetra-*O*-acetyl-2-[[(2*E*)-1-oxo-4-chloro-2-buten-1-yl]amino]-2-deoxy-5-thio-α-D-glucose (**ES1**): (*E*)-4-chloro-2-butenoic acid (synthesized as previously described[4])(24 mg, 0.20 mmol), dry DCM (1 mL), and one drop of dry DMF were combined and cooled to 0 °C. Oxalyl chloride (28 μL, 0.32 mmol) was added dropwise. The solution was stirred at 0 °C for 20 minutes and then at r.t. for 2.5 h. The reaction mixture was concentrated *in vacuo* and the residue was cooled to 0 °C. A solution of **10** (16 mg, 0.040 mmol), Et_3N (6 μL, 0.040 mmol), and dry DCM (1.0 mL) was prepared, and slowly added to the residue. The mixture was stirred for 10 minutes at 0 °C and then 2 h at r.t. The solution was concentrated *in vacuo*, taken up in EtOAc (8 mL), and washed with saturated $NaHCO_3$ (8 mL) and water (8 mL). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo*. The crude product was purified by column chromatography on silica gel (hexanes/EtOAc, 2:1), and dried down to give **ES1** as a white solid (11.6 mg, 62% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.90 (dt, J = 14.9, 5.7 Hz, 1H), 5.98 (m, 2H), 5.87 (d, J = 8.8 Hz, 1H), 5.40 (dd, J = 10.8, 9.6 Hz, 1H), 5.22 (dd, J = 10.9, 9.6 Hz, 1H), 4.71 (ddd, J = 11.1, 8.8, 3.1 Hz, 1H), 4.36 (dd, J = 12.1, 5.0 Hz, 1H), 4.15 (dd, J = 5.6 Hz, 1.5 Hz, 2H), 4.05 (dd, J = 12.1, 3.1 Hz, 1H), 3.49 (ddd, J = 10.8, 4.8, 3.3 Hz, 1H), 2.19 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 2.02 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.99, 170.71, 169.28, 168.80, 164.17, 139.85, 125.04, 72.79, 71.90, 71.57, 61.23, 55.51, 42.78, 39.91, 21.26, 20.80, 20.78, 20.66. Theoretical m/z calculated for $C_{18}H_{24}ClNNaO_9S$ $[M+Na]^+$: 488.0753. HRMS found: 488.0754. 95% purity by UHPLC.

1,3,4,6-tetra-*O*-acetyl-2-[[(2*E*)-1-oxo-4-fluoro-2-buten-1-yl]amino]-2-deoxy-5-thio-D-glucose (**ES2**): **ES2** was synthesized according to the procedure for **ES1**, using (*E*)-4-fluoro-2-butenoic acid in the first step (6.3 mg, 55% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.89 (ddt, J = 24.5, 15.4, 3.4 Hz, 1H), 5.98 (m, 2H), 5.86 (d, J = 8.7 Hz, 1H), 5.41 (t, J = 10.2 Hz, 1H), 5.23 (t, J = 10.2 Hz, 1H), 5.05 (ddd, J = 46.2, 3.4, 2.2 Hz, 2H), 4.72 (ddd, J = 11.2, 8.9, 3.1 Hz, 1H), 4.37 (dd, J = 12.1, 4.9 Hz, 1H), 4.06 (dd, J = 12.1, 3.0 Hz, 1H), 3.51 (ddd, J = 10.6, 4.6, 3.5 Hz, 1H), 2.19 (s, 3H), 2.08 (s, 3H), 2.05 (s, 3H), 2.03 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.93, 170.72, 169.30, 168.83, 164.25, 139.78 (d, J = 15.3 Hz), 121.97 (d, J = 10.1 Hz), 81.34 (d, J = 171.9 Hz), 72.78, 71.87, 71.58, 61.25, 55.51, 39.92, 21.24, 20.79, 20.78, 20.66. ^{19}F NMR (376 MHz) δ -125.65. Theoretical m/z calculated for $C_{18}H_{24}FNNaO_9S$ $[M+Na]^+$: 472.1048. HRMS found: 472.1049. 95% purity by UHPLC.

1,3,4,6-tetra-*O*-acetyl-2-[[1-oxo-3,4-epoxybutane-1-yl]amino]-2-deoxy-5-thio-D-glucose (**ES3**): Oxiraneacetic acid (22 mg, 0.20 mmol), dry THF (1 mL), N,N,N′,N′-tetramethylfluoroformamidinium hexafluorophosphate (TFFH, 66 mg, 0.25 mmol), and DIPEA (43.5 μL, 0.25 mmol) were combined and stirred at r.t. for 2.5 h, and then compound **10** (10 mg, 0.025 mmol) was added. After another 8 h stirring at r.t. the solution was partially concentrated *in*

vacuo, leaving a small amount of liquid in the flask, taken up with EtOAc (6 mL), and washed with water (6 mL). The organic layer was dried over Na_2SO_4 and concentrated. The crude product was purified by column chromatography on silica gel (hexanes/EtOAc, 1:2) (10.2 mg, 91% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.18 (d, J = 9.1 Hz, 0.5H), 6.11 (d, J = 8.8 Hz, 0.5H), 6.03 (d, J = 3.1 Hz, 0.5H), 5.96 (d, J = 3.1 Hz, 0.5H), 5.38 (t, J = 10.4 Hz, 1H), 5.23 (m, 1H), 4.64 (m, 1H), 4.37 (dd, J = 11.8, 5.2 Hz, 1H), 4.05 (dd, J = 12.1, 3.1 Hz, 1H), 3.51 (m, 1H), 3.13 (m, 1H), 2.84 (m, 1H), 2.65 (m, 1H), 2.55 (dd, J = 4.7, 2.6 Hz, 0.5H), 2.51 (dd, J = 4.6, 2.7 Hz, 0.5H), 2.20 (s, 3H), 2.08 (s, 3H), 2.05 (s, 3H), 2.05 (s, 3H). ^{13}C NMR (126 MHz, CD_3OD) δ 173.03, 172.66, 172.39, 172.15, 171.83, 171.76, 171.26, 170.80, 73.61, 73.58, 73.52 (2C), 72.81, 72.71, 62.38, 56.14, 56.07, 49.80, 49.74, 47.31 (2C), 40.87, 40.82, 40.03, 39.92, 20.60, 20.58, 20.53 (2C), 20.49 (2C). Theoretical m/z calculated for $C_{18}H_{25}NNaO_{10}S$ $[M+Na]^+$: 470.1091. HRMS found: 470.1081.

1,3,4,6-tetra-*O*-acetyl-2-[[(2*E*)-1-oxo-2,4-pentadiene-1-yl]amino]-2-deoxy-5-thio-D-glucose (**ES4**): Compound **ES4** was synthesized according to the procedure for **ES1** using 2,4-pentadienoic acid in place of (*E*)-4-chloro-2-butenoic acid (9.3 mg, 40% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.18 (dd, J = 15.0, 11.0 Hz, 1H), 6.39 (dt, J = 16.9, 10.5 Hz, 1H), 5.97 (d, J = 2.9 Hz, 1H), 5.76 (m, 2H), 5.59 (d, J = 16.9 Hz, 1H), 5.47 (d, J = 10.2 Hz, 1H), 5.41 (t, J = 10.2 Hz, 1H), 5.22 (t, J = 10.2 Hz, 1H), 4.73 (ddd, J = 11.0, 8.9, 3.0 Hz, 1H), 4.36 (dd, J = 12.1, 4.9 Hz, 1H), 4.05 (dd, J = 12.1, 3.0 Hz, 1H), 3.49 (m, 1H), 2.18 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 2.01 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.02, 170.73, 169.28, 168.79, 165.30, 142.83, 134.56, 125.65, 123.30, 72.98, 71.97, 71.60, 61.27, 55.43, 39.93, 21.26, 20.81, 20.68. Theoretical m/z calculated for $C_{19}H_{25}NNaO_9S$ $[M+Na]^+$: 466.1142. HRMS found: 466.1152. 95% purity by UHPLC.

1,3,4,6-tetra-*O*-acetyl-2-deoxy-2-trifluoroacetamido-5-thio-α-D-glucose (**11**): Compound **10** (78 mg, 0.20 mmol) was dissolved in DCM (2 mL) and triethylamine (TEA, 60 µL, 0.43 mmol) was added. Trifluoroacetic anhydride was added dropwise at 0 °C and the reaction mixture was warmed to r.t. overnight. The mixture was diluted with DCM (10 mL) and washed with saturated $NaHCO_3$ (10 mL). The aqueous phase was extracted with DCM twice. The combined organic phase was dried over Na_2SO_4 and concentrated to give **11** as a pale yellow oil (89 mg, 100% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.62 (d, J = 8.4 Hz, 1H), 6.01 (d, J = 3.1 Hz, 1H), 5.41 (dd, J = 10.8, 9.6 Hz, 1H), 5.25 (dd, J = 10.8, 9.7 Hz, 1H), 4.63 (ddd, J = 11.3, 8.7, 3.1 Hz, 1H), 4.38 (dd, J = 12.2, 4.9 Hz, 1H), 4.06 (dd, J = 12.2, 3.1 Hz, 1H), 3.52 (ddd, J = 10.8, 4.8, 3.2 Hz, 1H), 2.21 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.84, 170.64, 169.24, 168.68, 157.02 (q, J = 38.2 Hz), 115.50 (q, J = 287.9 Hz), 71.97, 71.41, 71.16, 60.99, 55.96, 39.98, 21.08, 20.74, 20.60, 20.50. ^{19}F NMR (470 MHz, $CDCl_3$) δ -75.96. Theoretical m/z calculated for $C_{16}H_{20}F_3NNaO_9S$ $[M+Na]^+$: 482.0709. HRMS found: 482.0703.

2-deoxy-2-trifluoroacetamido-5-thio-D-glucose (**12**): Compound **11** (113 mg, 0.21 mmol) was dissolved in a mixture of MeOH:H_2O:TEA (5:2:1, 2 mL). The reaction mixture was stirred at r.t. for 2 h and concentrated *in vacuo*. The product was purified by column chromatography on silica gel (DCM/MeOH, 9:1) and dried down to give **12** as a white solid (38 mg, 62% yield). ^{1}H NMR (400 MHz, D_2O) δ 5.08 (d, J = 2.8 Hz, 1H), 4.25 (dd, J = 10.6, 2.8 Hz, 1H), 3.97 (dd, J = 11.9, 5.6 Hz, 1H), 3.91 (dd, J = 12.0, 3.4 Hz, 1H), 3.85 (dd, J = 10.4, 9.2 Hz, 1H), 3.72 (dd, J = 10.5, 9.0 Hz, 1H), 3.33 (m, 1H). ^{13}C NMR (101 MHz, D_2O) δ 158.96 (q, J = 37.9 Hz), 115.92 (q, J = 286.0 Hz), 74.08, 71.33, 71.18, 60.22, 58.89, 43.25. ^{19}F NMR (470 MHz, D_2O) δ -75.24. LRMS: *m/z* 290.0 $[M-H]^-$.

Chemoenzymatic synthesis of UDP-5SGlcNTFA (**13**): The reaction was set up with 100 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 12 mM ATP, 12 mM UTP, 10 mM **12**, 0.5 mg/mL NahK (protein purification is described in *biochemical experiments*), 0.75 mg/mL AGX1 (protein purification is described in *biochemical experiments*), and 1 U/mL pyrophosphatase (Sigma #I1643-100UN) in a total volume of 0.5 mL. The mixture was incubated at 37 °C for 36 h and the reaction was concentrated *in vacuo*. The product was purified by column chromatography on silica gel (EtOAc/MeOH/H_2O, 4:2:1) and dried down to give **13** as a white solid (3.3 mg, 92% yield). ^{1}H NMR (400 MHz, D_2O) δ 7.98 (d, J = 8.1 Hz, 1H), 6.00─5.96 (m, 2H), 5.44 (dd, J = 7.8, 2.5 Hz, 1H), 4.40─4.37 (m, 2H), 4.33 (dt, J = 10.5, 2.6 Hz, 1H), 4.31─4.19 (m, 3H), 4.00 (dd, J = 12.0, 5.3 Hz, 1H), 3.93─3.87 (m, 2H), 3.76 (m, 1H), 3.45─3.41 (m, 1H). Theoretical m/z calculated for $C_{17}H_{22}F_3N_3NaO_{16}P_2S$ $[M-2H+Na]^-$: 698.0051. HRMS found: 698.0046.

Uridine 5′-diphospho-2-amino-2-deoxy-5-thio-α-D-glucose (**14**): **13** (5.6 mg, 8.3 μmol) was dissolved in a mixture of MeOH:H_2O (2:1, 0.9 mL) and K_2CO_3 (50 mg, 0.36 mmol) was added. The reaction mixture was stirred at r.t. overnight. The mixture was concentrated *in vacuo* and purified by column chromatography on silica gel (EtOAc/MeOH/H_2O, 1:1:1). The product was dried down to give **14** as a powder (3.0 mg, 67% yield). ^{1}H NMR (400 MHz, D_2O) δ 7.94 (d, J = 8.1 Hz, 1H), 6.04─5.92 (m, 2H), 5.43 (dd, J = 7.5, 2.7 Hz, 1H), 4.37 (dd, J = 3.3, 1.0 Hz, 2H), 4.28 (m, 1H), 4.24 (m, 1H), 3.97 (dd, J = 12.1, 5.3 Hz, 1H), 3.87 (dd, J = 12.0, 3.1 Hz, 1H), 3.70─3.61 (m, 2H), 3.41─3.34 (m, 2H), 3.23 (dt, J = 9.9, 3.0 Hz, 1H).

Uridine 5′-diphospho-2-deoxy-2-[[(2*E*)-1-oxo-4-chloro-2-buten-1-yl]amino]-5-thio-α-D-glucose (**UDP-ES1**): (*E*)-4-chloro-2-butenoic acid (22.2 mg, 0.19 mmol), dry DCM (1 mL), and dry DMF (two drops) were combined and cooled to 0 °C. Oxalyl chloride (26 μL, 0.30 mmol) was then added to the reaction mixture and stirred at 0 °C for 20 minutes followed by 3 h at r.t. A solution of **14** (7 mg, 0.012 mmol) and $NaHCO_3$ (125 mg, 1.5 mmol) in ACN:H_2O (1:1, 1 mL) was prepared and cooled to 0 °C. The acid chloride reaction mixture was partially concentrated *in vacuo* leaving a small amount of liquid, and the yellow residue was taken up in dry ACN (0.3 mL) and dropped into the stirred **14** mixture. The reaction was then stirred at 0 °C for 4 h, concentrated *in vacuo*, taken up in water (5 mL), and washed with DCM (5 mL). The product was purified by column chromatography on silica gel (EtOAc/MeOH/H_2O, 7:2:1) and dried down to give **UDP-ES1** as a white solid (1.7 mg, 21% yield). ^{1}H NMR (400 MHz, D_2O) δ 7.93 (d, J = 8.0 Hz, 1H), 6.89─6.82 (m, 1H), 6.40 (d, J = 12.0 Hz, 1H), 5.98 (d, J = 4.0 Hz, 1H), 5.94 (d, J = 8.0 Hz, 1H), 5.36 (d, J = 4.0 Hz, 1H), 4.38─4.33 (m, 2H), 4.31─4.28 (m, 4H), 4.23─4.18 (m, 2H), 4.01─3.97 (m, 1H), 3.92─3.90 (m, 1H), 3.82─3.72 (m, 1H), 3.43─3.39 (m, 1H). ^{13}C NMR (126 MHz, D_2O) δ 167.72, 166.44, 151.96, 141.73, 139.68, 125.45, 102.76, 88.64, 83.21, 76.91, 73.97, 73.68, 72.29, 69.74, 65.13, 60.01, 57.93, 43.76, 43.00. ^{31}P NMR (162 MHz, D_2O) δ -11.24 (d, J = 22.2 Hz), -12.97 (d, J = 22.2 Hz). Theoretical m/z calculated for $C_{19}H_{26}ClN_3O_{16}P_2S$ $[M-2H]^{2-}$: 340.5104. HRMS found: 340.5074.

Methyl 2-azido-2-deoxy-5-thio-D-glucose (**15**): Compound **8** (500 mg, 1.28 mmol) was dissolved in dry MeOH (7 mL) and concentrated HCl (0.63 mL, 7.6 mmol) was added. The solution was then stirred at 65 °C for 4.5 h. The reaction mixture was concentrated *in vacuo* and purified by column chromatography on silica gel (DCM/MeOH, 98:2). The product was dried down to give **15** as a brown residue (220 mg, 73% yield). ^{1}H NMR (500 MHz, D_2O) δ 4.82 (d, J = 2.9 Hz, 1H), 4.00─3.91 (m, 2H), 3.87 (dd, J = 10.0, 2.9 Hz, 1H), 3.78 (t, J = 9.4 Hz, 1H), 3.72

(t, J = 9.6 Hz, 1H), 3.50 (s, 3H), 3.12 (ddd, J = 9.6, 5.5, 3.4 Hz, 1H). ^{13}C NMR (126 MHz, D_2O) δ 82.13, 73.74, 73.25, 67.79, 60.04, 56.07, 42.97. LRMS: *m/z* 258.0 $[M+Na]^+$.

Methyl 2-azido-2-deoxy-6-*O*-mesyl-5-thio-D-glucose (**16**): Compound **15** (75 mg, 0.32 mmol) was dissolved in dry pyridine (1.5 mL) and cooled to 0 °C. Methanesulfonyl chloride (55 µL, 0.70 mmol) was added dropwise and stirred at 0 °C for 2 h. The mixture was concentrated *in vacuo*, purified by column chromatography on silica gel (hexanes/EtOAc, 1:1), and dried down to give **16** as a light brown residue (85 mg, 85% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.74 (dd, J = 11.0, 4.3 Hz, 1H), 4.58 (d, J = 2.8 Hz, 1H), 4.36 (dd, J = 11.0, 2.5 Hz, 1H), 3.90 (dd, J = 10.0, 8.8 Hz, 1H), 3.76 (dd, J = 10.4, 8.9 Hz, 1H), 3.59 (dd, J = 10.2, 2.8 Hz, 1H), 3.45 (s, 3H), 3.23 (m, 1H), 3.08 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 83.12, 73.57, 72.77, 67.80, 67.10, 56.76, 40.31, 37.52. LRMS: *m/z* 335.9 $[M+Na]^+$.

Methyl 3,4-di-*O*-acetyl-2-azido-2-deoxy-6-*O*-mesyl-5-thio-D-glucose (**17**): Compound **16** (180 mg, 0.57 mmol) was dissolved in dry pyridine and cooled to 0 °C. Acetic anhydride (1.1 mL, 12 mmol) was added and the reaction mixture was stirred for five minutes at 0 °C and then at r.t. for 10 h. The mixture was concentrated *in vacuo*, diluted with DCM (10 mL), and washed with 2 N HCl (10 mL) and saturated $NaHCO_3$ (10 mL). The organic layer was dried over Na_2SO_4, purified by column chromatography on silica gel (hexanes/EtOAc, 3:1), and dried down to give **17** as a residue (209 mg, 92% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.44 (dd, J = 10.4, 9.4 Hz, 1H), 5.26 (dd, J = 10.7, 9.3 Hz, 1H), 4.66 (d, J = 2.8 Hz, 1H), 4.42 (dd, J = 10.7, 4.6 Hz, 1H), 4.20 (dd, J = 10.8, 3.1 Hz, 1H), 3.66 (dd, J = 10.4, 2.8 Hz, 1H), 3.50 (m, 1H), 3.47 (s, 3H), 3.02 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.86, 169.71, 82.95, 71.90, 71.41, 65.74, 65.51, 56.81, 38.56, 37.72, 20.67, 20.64. Theoretical m/z calculated for $C_{12}H_{19}N_3NaO_8S_2$ $[M+Na]^+$: 420.0511. HRMS found: 420.0502.

1,3,4-tri-*O*-acetyl-2-azido-2-deoxy-6-*O*-mesyl-5-thio-α-D-glucose (**18**): Compound **17** (209 mg, 0.53 mmol) was cooled to 0 °C in a flask, and acetic anhydride (3 mL) and concentrated sulfuric acid (52 µL) were added. The solution was stirred at 0 °C for 4 h. Sodium bicarbonate was then added to neutralize the reaction. The mixture was concentrated *in vacuo*, taken up in DCM, and filtered through glass wool. It was purified by column chromatography on silica gel (hexanes/EtOAc, 2:1). The product was dried down to give **18** as a white amorphous solid (177 mg, 79% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.13 (d, J = 3.0 Hz, 1H), 5.42 (t, J = 9.9 Hz, 1H), 5.32 (t, J = 10.2 Hz, 1H), 4.44 (dd, J = 10.9, 4.8 Hz, 1H), 4.21 (dd, J = 10.9, 3.0 Hz, 1H), 3.89 (dd, J = 10.3, 3.1 Hz, 1H), 3.63 (ddd, J = 10.7, 4.4, 3.2 Hz, 1H), 3.04 (s, 3H), 2.20 (s, 3H), 2.11 (s, 3H), 2.09 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.79 (2C), 168.68, 71.72 (2C), 71.51, 64.94, 64.89, 39.90, 37.81, 21.07, 20.70, 20.66. Theoretical m/z calculated for $C_{13}H_{19}N_3NaO_9S_2$ $[M+Na]^+$: 448.0460. HRMS found: 448.0451.

1,3,4-tri-*O*-acetyl-2-(tert-butyloxycarbonylamino)-2-deoxy-6-*O*-mesyl-5-thio-α-D-glucose (**19**): Compound **18** (177 mg, 0.42 mmol), Pd/C (10 wt. %, 187 mg, 0.18 mmol) and dry MeOH (3 mL) were combined in a flask at r.t., and then flushed with H_2. A solution of di-tert-butyl dicarbonate (163 mg, 0.74 mmol) in MeOH (1 mL) was then added to the reaction mixture and stirred at r.t. for 3 h. The mixture was filtered through glass wool and concentrated *in vacuo*. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 2:1) and dried down to give **19** as a white powder (121 mg, 58% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.96 (d, J = 2.4 Hz, 1H), 5.34 (t, J = 10.0 Hz, 1H), 5.17 (t, J = 10.0 Hz, 1H), 4.74 (d, J = 9.5 Hz, 1H),

4.40─4.30 (m, 2H), 4.22 (dd, J = 10.8, 3.3 Hz, 1H), 3.56 (dt, J = 10.3, 4.4 Hz, 1H), 3.04 (s, 3H), 2.20 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 1.41 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.09, 169.56, 168.84, 154.76, 80.71, 72.98, 71.77, 71.64, 65.40, 56.51, 39.90, 37.79, 28.34, 21.22, 20.71, 20.70. LRMS: *m/z* 522.0 $[M+Na]^+$.

1,3,4-tri-*O*-acetyl-6-azido-2-(tert-butyloxycarbonylamino)-2,6-dideoxy-5-thio-α-D-glucose (**20**): Compound **19** (150 mg, 0.30 mmol), dry DMF (2 mL), and sodium azide (60 mg, 0.90 mmol) were combined at r.t. and stirred at 60 ºC for 4 h. The solution was cooled to r.t. and partially concentrated *in vacuo*, leaving a small amount of liquid. The mixture was diluted with 15 mL DCM, washed twice with 15 mL H_2O and dried over Na_2SO_4. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 3:1), and dried down to give **20** as a yellow residue (124 mg, 92% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.95 (d, J = 2.8 Hz, 1H), 5.27 (t, J = 10.0 Hz, 1H), 5.15 (t, J = 10.0 Hz, 1H), 4.73 (d, J = 9.5 Hz, 1H), 4.31 (td, J = 10.5, 2.6 Hz, 1H), 3.56─3.50 (m, 1H), 3.44─3.34 (m, 2H), 2.18 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H), 1.40 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.15, 169.47, 168.87, 154.77, 80.65, 73.03, 73.02, 71.72, 56.47, 50.74, 40.65, 28.34, 21.23, 20.73, 20.72. LRMS: *m/z* 469.0 $[M+Na]^+$.

1,3,4-tri-*O*-acetyl-6-azido-2-amino-2,6-dideoxy-5-thio-α-D-glucose trifluoroacetic acid salt (**21**): Compound **20** (43 mg, 0.096 mmol) was dissolved in DCM (0.8 mL) and cooled to 0 ºC. Trifluoroacetic acid (TFA, 0.8 mL) was added and stirred at r.t. for 45 minutes. The reaction was then concentrated and dried *in vacuo* to give **21** as a light brown powder, which was used without further purification (38 mg, 86% yield). ^{1}H NMR (400 MHz, D_2O) δ 6.19 (d, J = 3.1 Hz, 1H), 5.53 (t, J = 10.0 Hz, 1H), 5.33 (t, J = 10.0 Hz, 1H), 4.26 (dd, J = 10.7, 3.1 Hz, 1H), 3.78 (dt, J = 10.7, 4.5 Hz, 1H), 3.70 (m, 2H), 2.26 (s, 3H), 2.18 (s, 3H), 2.17 (s, 3H). ^{13}C NMR (101 MHz, D_2O) δ 172.64, 172.49, 171.40, 72.70, 71.16, 70.55, 54.75, 49.25, 39.44, 20.25, 20.11, 19.98. Theoretical m/z calculated for $C_{12}H_{18}N_4NaO_6S$ $[M+Na]^+$: 369.0845. HRMS found: 369.0840.

1,3,4-tri-*O*-acetyl-2-[[(2*E*)-1-oxo-4-chloro-2-buten-1-yl]amino]-6-azido-2,6-dideoxy-5-thio-α-D-glucose (**6AzES1**): (*E*)-4-chloro-2-butenoic acid (20 mg, 0.16 mmol), DCM (1 mL) and DMF (1 drop) were combined in a flask and cooled to 0 ºC. Oxalyl chloride (22 µL, 0.26 mmol) was added dropwise and stirred at 0 ºC for 10 minutes, then at r.t. for 2.5 h. The solution was partially concentrated *in vacuo*, leaving a small amount of liquid, and cooled to 0 ºC. A solution of **21** (11 mg, 0.033 mmol) and Et_3N (11 µL, 0.082 mmol) in DCM (1 mL) was slowly added to the acid chloride mixture and stirred at 0 ºC for 10 minutes, then at r.t. for 1 h. The mixture was diluted with DCM (6 mL), washed with water (6 mL), and dried over Na_2SO_4. The product was purified by column chromatography on silica gel (hexanes/EtOAc, 3:1) and dried down to give **6AzES1** as a residue (3.2 mg, 22% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.91 (dt, J = 14.9, 5.6 Hz, 1H), 5.97 (m, 2H), 5.83 (d, J = 8.7 Hz, 1H), 5.34 (t, J = 9.9 Hz, 1H), 5.22 (t, J = 10.2 Hz, 1H), 4.70 (ddd, J = 11.6, 8.9, 3.1 Hz, 1H), 4.16 (dd, J = 5.6, 1.5 Hz, 2H), 3.55 (dd, J = 12.6, 3.9 Hz, 1H), 3.47─3.38 (m, 2H), 2.19 (s, 3H), 2.09 (s, 3H), 2.03 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.00, 169.42, 168.74, 164.16, 139.93, 125.00, 72.74, 72.64, 71.65, 55.45, 50.73, 42.78, 40.70, 21.25, 20.81, 20.74. Theoretical m/z calculated for $C_{16}H_{21}ClN_4NaO_7S$ $[M+Na]^+$: 471.0717. HRMS found: 471.0710. 95% purity by UHPLC.

Biochemical and Mass Spectrometric Experiments

Cloning, expression, and purification of one-pot enzymes Nahk_ATCC15697 and AGX1. Nahk_ATCC15697 was amplified from genomic DNA of *Bifidobacterium longum* subsp. *infantis* ATCC#15697 (ref. 5) using the primers listed in Table S2 and cloned into pET28a(+) vector to generate an N-terminal His_6-tagged fusion protein. AGX1 (ref. 6) was similarly cloned from human AGX1 gene (Harvard PlasmID #Hscd00329319) using the primers in Table S2 into pET28a(+) vector to generate an N-terminal His_6-tagged fusion protein. *E. coli* BL21(DE3) transformant containing the plasmid of pET28a-Nahk_ATCC15697 or pET28a-AGX1 was grown in LB medium with kanamycin sulfate at 37 °C until OD_{600} reached 0.8. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM to induce protein expression at 16 °C overnight. The cells were harvested by centrifugation and the cell pellet was resuspended in TBS buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl) supplemented with 1 mM PMSF and lysed by ultra-high-pressure cell disrupter (Emulsiflex-C5, ATA Scientific). The lysate was clarified by centrifugation and the supernatant was purified on Ni-NTA column with TBS as the equilibration buffer. Ni-NTA beads were purchased from Qiagen. The desired protein was eluted in TBS buffer containing 250 mM imidazole (pH 8.0) and 0.5 mM tris(hydroxypropyl)phosphine (THP) and further purified by size-exclusion chromatography (Superdex 200 increase 10/300, GE Healthcare) using the same buffer. The fractions containing the purified enzyme were concentrated, and the final product was stored at -80 °C in small aliquots. The purity of proteins was examined by SDS-PAGE gel.

NUP62 protein expression and purification. Human NUP62 expression plasmid in pET21a vector (a kind gift from Dr. Suzanne Walker's lab) was transformed into *E. coli* BL21(DE3) competent cells and expressed as a fusion protein containing a C-terminal His_6-tag. The transformant was grown at 37 °C in LB medium. After OD_{600} reached 0.4, the culture was then induced with 0.2 mM IPTG for 3 h. Cells were pelleted and resuspended in lysis buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 0.5 M NaCl, and 1 mM PMSF). The cell suspension was lysed with ultra-high-pressure cell disrupter followed by centrifugation at 20,500 g (13,096 r.p.m.) at 4 °C for 30 minutes. The collected pellets containing the inclusion body of NUP62 were then washed twice with 30 mL detergent solution (1.1 M urea, 2% Triton X-100) and resuspended in 30 mL denaturing buffer (8 M urea, 50 mM Tris-HCl pH 8.0, 1 mM EDTA, 2 mM DTT) and incubated at r.t. to dissolve the inclusion body. The lysate was then centrifuged at 16,100 g (11,606 r.p.m.) at 20 °C for 20 minutes to remove the unbroken pellets. The supernatant was centrifuged at 50,000 g (20,453 r.p.m.) at 20 °C for 15 minutes to further clear up the lysate. Proteins were dialyzed against 2 M urea in 20 mM Tris-HCl pH 7.5 for 4 h followed by 20 mM Tris-HCl pH 7.5 overnight at 4 °C. Following dialysis, samples were centrifuged at 16,100 g at 4 °C for 20 minutes. The supernatant containing NUP62 was concentrated and stored at -80 °C until use. The purity of NUP62 protein was examined by SDS-PAGE gel. All the centrifugation protocols of lysates were conducted using the SS-34 rotor (Sorvall).

OGT purification. The full-length ncOGT and $OGT_{4.5}$ expression plasmids[7] in pET24b vector were kind gifts from Dr. Suzanne Walker's lab. Each plasmid was transformed into *E. coli* BL21(DE3) competent cells, and the transformant was grown at 37 °C in LB medium. After OD_{600} reached 1.0, the culture was induced overnight by 0.2 mM IPTG at 16 °C. The cells were pelleted, re-suspended in TBS buffer supplemented with 1 mM PMSF, and lysed with ultra-high-pressure

cell disrupter. After centrifugation, the supernatant was subjected to Ni-NTA column for affinity purification with TBS as the equilibration buffer. The recombinant protein was subsequently eluted with TBS buffer containing 250 mM imidazole (pH 8.0) and 0.5 mM THP. The purified proteins were concentrated and stored in small aliquots at -80 °C.

Mutagenesis. $OGT_{4.5}$–C917S mutant was generated using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's instructions. $OGT_{4.5}$ was used as the DNA template with the primers listed in Table S2. The DNA sequence was verified by sequencing. The mutant protein was expressed and purified similarly as $OGT_{4.5}$ as mentioned above.

Radiolabeled assay to evaluate the time-dependent inhibition of OGT by UDP-ES1. Purified OGT (1 μM) was pre-incubated with the indicated concentration of **UDP-ES1** at r.t. for 2.5 or 5 minutes. The mixture was then diluted 50-fold with the reaction buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5 mM THP) containing OGT substrates NUP62 (10 μM) and UDP-^{3}H-GlcNAc (100 μM, specific activity 0.092 Ci/mmol). UDP-^{3}H-GlcNAc was purchased from Perkin Elmer. The reactions were incubated at 37 °C for 0.5 h and then quenched by transferring the samples onto the nitrocellulose membrane and air-dried. The membranes were washed in PBS buffer four times for 5 minutes each. The radioactivity on each membrane was counted by Tri-Carb 2900TR liquid scintillation analyzer (Perkin Elmer). A reaction in the absence of OGT and another reaction without **UDP-ES1** were served as controls. Experiments were performed in triplicates. Data were analyzed using GraphPad Prism v5 (GraphPad Software, Inc.). The $t_{1/2}$ of each concentration of inhibitor was determined from the plot of Log (% OGT activity) versus preincubation time. Based on the linear equation of each inhibitor concentration $y_1 = a_1x_1 + b_1$, plot y_1 as Log50 and solve for each individual x_1 as the $t_{1/2}$. Therefore, $t_{1/2} = (\text{Log50} - b_1)/a_1$. With all of the $t_{1/2}$ calculated, the plot of $t_{1/2}$ versus 1/[inhibitor] was constructed. K_I and k_{inact} were subsequently derived from the linear equation of this plot $y_2 = a_2x_2 + b_2$. $K_I = a_2/b_2$, and $k_{inact} = 0.693/b_2$. k_{inact}/K_I can be calculated accordingly.

Intact protein mass spectrometry. To detect the covalent modification of **UDP-ES1** on OGT, purified OGT (10 μM wild-type $OGT_{4.5}$ or C917S mutant) was incubated in a 25 μL reaction containing 70 μM of **UDP-ES1** in the reaction buffer (10 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM THP, 200 U/mL CIP alkaline phosphatase, and 30 mM $MgCl_2$) at r.t. for 2 h followed by C8 StageTip desalting.[8] Samples were SpeedVac dried and re-dissolved in 20 μL of 0.1% formic acid. MS analysis of intact protein was performed on ESI-Q-TOF Maxis II (Bruker) coupled with an ACQUITY UPLC (Waters). Samples were loaded onto a 500 μm × 200 mm PLRP column (PLRP-S, 10 μm, 1,000 Å, Varian). The injection volume was 7 μL with a flow rate of 300 nL/min. The mobile phases consisted of 0.1% TFA (solvent A) and 0.1% TFA in ACN (solvent B). LC program: 5% B for 5 minutes, 5–40% B for 10 minutes, 40–70% B for 20 minutes, 70–95% B for 10 minutes, 95% B for 5 minutes, and 95–5% B for 1 minute. MS analysis was operated in positive mode. End plate offset and capillary voltages were set at 400 and 4,000 V, respectively. The nebulizer was set at 0.5 bar. The interface heater temperature was set at 220 °C with the dry gas flow rate 4.0 L/min. Data were acquired using one full MS scan (m/z 500–3,000) with the scan rate at 1 Hz. LC-MS data were processed and analyzed using Compass Data Analysis software (version 4.3, Bruker). Deconvolution of the mass spectra was performed by maximum entropy algorithm and the parameters were as follows: mass range 50,000–100,000 Da, auto data point spacing, and the

resolving power of 8,000. The increased mass resulting from **UDP-ES1** modification was determined by subtracting the deconvoluted protein mass from the unmodified protein.

nanoLC-MS/MS analysis of UDP-ES1-modified peptides. To detect **UDP-ES1**-labeled OGT peptides, purified full-length OGT protein (70 μM) was incubated in a 10 μL reaction containing 490 μM of **UDP-ES1** in the reaction buffer (10 mM Tris-HCl, pH 8.0, 75 mM NaCl, 0.5 mM THP, 200 U/mL CIP alkaline phosphatase, and 30 mM $MgCl_2$) at r.t. for 2 h. The reaction was stopped by adding 10 mM DTT and 8 M urea and diluted to 0.5 μg/μL in denature buffer (8 M urea and 50 mM triethylamine bicarbonate). After reduction by 10 mM DTT for 30 minutes at r.t. and carbidomethylation with 55 mM iodoacetamide in the dark for 30 minutes at r.t., alkylated proteins were seven-fold diluted using 25 mM triethylammonium bicarbonate and then digested by trypsin/Lys-C mix (protein:protease = 40:1) (Promega) at r.t. for 20 h. Formic acid was added to the sample to reach final 0.5% for enzyme inactivation and sample acidification. Peptides were SpeedVac dried after SDB-XC StageTip desalting.

Peptide samples were dissolved in 0.1% formic acid for LC-MS/MS analysis on an Orbitrap Q-Exactive (Thermo Scientific) equipped with a nanoAcaquity UPLC system (Waters). Peptides (0.25 μg) were loaded onto a 75 μm × 15 cm 1.7 μm BEH C18 column at a flow rate of 300 nL/minute. Mobile phase A consisted of 0.1% formic acid, and solvent B was 0.1% formic acid in ACN. A linear gradient of 0–4% B for 0.1 minute, 4–35% B for 30 minutes, 35–75% B for 0.1 minute, 75% B for 9.5 minutes, 75–95% B for 0.1 minute, 95% B for 9.5 minutes, 95–0% B for 0.5 minute, and 0% B for 9.5 minutes was employed throughout this study. Mass spectra from full scans were acquired in a data-dependent mode (*m/z* 200–2,000). The resolution of survey scan was set to 17,500 at *m/z* 400 with an automated gain control (AGC) value of 10^6. The top 10 most-intense precursor ions were selected from the MS scan for subsequent higher energy collisional dissociation (HCD, normalized collision energy 30 eV) MS/MS scan. Peptide identification was performed by MaxQuant (v1.5.2.5)[9] against a composite target-decoy protein sequence database containing Uniprot database (release 2017_09, subset human, 20,237 protein entries) (UniProt Consortium 2014). The search criteria used in this study include trypsin specificity allowing up to 2 missed cleavages, and variable modifications of **UDP-ES1** ($C_{19}H_{27}N_3O_{16}P_2S$) on Cys, **ES1** glycosylation ($C_{10}H_{14}ClNO_4S$) on Ser/Thr, carbamidomethyl on Cys, and oxidation on Met. The precursor mass tolerance and the fragment ion tolerance were set at ± 10 ppm and ± 0.6 Da, respectively. Peptide was considered identified based on the posterior error probability with a false discovery rate of 1%. The spectra of peptides were manually inspected.

Crystallization of OGT:UDP-ES1:CKII complex. Purified $OGT_{4.5}$ was pre-incubated with **UDP-ES1** for 1 h at 4 ˚C before incubating with 2 mM CKII peptide (YPGGSTPVSSANMM, synthesized by Biomatik) for another 1 h at 4 ˚C. Seeds were generated from $OGT_{4.5}$ crystals grown in 1.3 M DL-malic acid and 0.1 M Bis-Tris propane, pH 6.4. The crystal used for X-ray diffraction was generated by mixing 2 μL protein solution with 1 μL reservoir solution and 0.5 μL seeds, then equilibrated against 200 μL reservoir solution using hanging-drop vapor-diffusion method at 20 ˚C. Crystals were obtained in the reservoir solution containing 1.45 M potassium phosphate dibasic, 10 mM EDTA, and 1% xylitol. The crystals were transferred into cryoprotectant solution containing 1.45 M potassium phosphate dibasic, 10 mM EDTA, and 27% xylitol, before being flash frozen in liquid nitrogen for storage.

X-ray data collection and structure determination. All X-ray data were collected on the Life Sciences Collaborative Access Team (LS-CAT) beamline 21-ID-D at the Advanced Photon Source, Argonne National Laboratory, IL. The wavelength for data collection was 1.078 Å. Data sets were processed using the HKL2000 package.[10] The crystal belongs to the space group of F222 and contains one molecule per asymmetric unit, with the cell dimensions: a = 137.7 Å; b = 151.0 Å; c = 198.9 Å; α = β = γ = 90˚. The structure was solved by molecular replacement, using OGT as a search model (PDB 3PE3).[7] Iterative model building was performed in COOT[11] and refinement was completed in PHENIX.[12] Final statistics were summarized in Table S1. All structural figures were prepared using PyMOL (DeLano Scientific, http://www.pymol.org/). The coordinate and structure factor have been deposited with PDB code 6E37. To evaluate the possibility of sugar transfer from **UDP-ES1** to CKII peptide during crystallization, we collected X-ray datasets from 10 different OGT crystals. Following data processing, we did not detect any electron density for potentially glycosylated peptide residue in these data. In contrast, all of these X-ray structures demonstrated clear electron density bridging OGT C917 side chain and the electrophilic group of **UDP-ES1**, which is consistent with its covalent inhibition mechanism. This observation strongly suggests that the possibility of **UDP-ES1** being used as a sugar donor for OGT is negligible.

Cell culture and inhibitor treatment. MCF7, COS-7, and HEK293 cells were cultured in DMEM media (Corning) supplemented with 10% fetal bovine serum (FBS) (Sigma) and 0.8% penicillin-streptomycin. Cells were maintained in 5% CO_2 at 37 ºC and were seeded 20-24 h prior to inhibitor treatment at 50-65% confluence. Upon treatment, media was replaced with low glucose DMEM medium (1 g/L glucose, 5% FBS) containing $Ac_4$5SGlcNAc, **ES1**-**ES4**, BZX2, or DMSO at given concentration (final 1% DMSO). Cells were incubated for 13 h unless otherwise indicated. For post-incubation, medium was replaced with the inhibitor-free low glucose DMEM medium after inhibitor treatment and cells were incubated for another 8 h. For glucosamine treatment, 10 or 15 mM glucosamine was added to the COS-7 or MCF7 cells after 3 h of post-incubation, respectively. Cells were harvested by scrapping, washed by PBS, pelleted at 150 g (1,274 r.p.m., F241.5P rotor, Beckman Coulter) at 4 °C for 5 minutes, and stored at -80 °C until use.

Cell viability assay. MCF7 cells were seeded at 2,200 cells per well in a 96-well plate 23 h prior to inhibitor treatment. The cells reached 60% confluence upon treatment. $Ac_4$5SGlcNAc, **ES1**, or DMSO was added at varied concentrations in low glucose DMEM medium (1 g/L glucose, 5 % FBS) and pre-treated for 13 h. The medium was then replaced with inhibitor-free low glucose medium for 8 h post-incubation. The cell viability was measured by adding 50 μg/mL of resazurin[13] to each well and incubated for 12 h at 37 °C. The optical density was detected using an excitation wavelength of 560 nm and an emission wavelength of 590 nm in a microplate reader (Synergy H1 Hybrid, Biotek). Results from five different wells of each condition were used for quantitation. Data were analyzed using GraphPad Prism v5 and shown as mean values with error bars representing ± S.D. Statistical significance was determined using Student's *t*-test.

Western blot and lectin blot. Pellet was lysed in RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 % NP40, 0.2 mM EDTA, 0.1 % SDS, 0.5% sodium deoxycholate, 10 μM thiamet G (synthesized as previously described[14]) and protease inhibitor cocktail (Sigma) by incubating on ice for 50 minutes with occasional vortex. Cell debris was removed by centrifugation at 16,000 g (13,159 r.p.m., F241.5P rotor, Beckman Coulter) at 4 °C for 15 minutes. Protein concentration was measured by BCA protein assay kit (Thermo Scientific). Cell lysates were separated by SDS-PAGE gel and transferred onto a nitrocellulose membrane (Life Technology). The membrane was

then blocked with 0.9% bovine serum albumin followed by primary antibody hybridization overnight at 4 °C. For O-GlcNAc detection, anti-O-GlcNAc antibody RL2 (Life Tech #MA1072) or CTD 110.6 (Sigma #MABS1254) was used. To detect OGT, OGA, actin, and *c*-myc tagged, and flag-tagged proteins, the blot was hybridized with anti-OGT (Proteintech #11576-2-AP or SinoBiological #102091-T36), anti-OGA (G-12, Santa Cruz #sc-376429), anti-actin-HRP antibody (Proteintech #HRP-60008), anti-*c*-myc (BD Bioscience #BDB551102), or anti-flag (M2, Sigma #F1804). Secondary antibodies including anti-mouse-lgG-HRP (Life Tech #16066), anti-mouse-IgG- spectra700 (Azure #S1015), anti-rabbit-IgG-spectra800 (Azure #S1021), or anti-mouse-IgM-HRP (Life Tech #62-6820) was applied to the blot with 1 h incubation at r.t. For glycan detection, the blot was incubated with ConA-HRP (EY laboratory, #H-1104-1) or HPA-HRP (EY laboratory, #H-3601-1). The blot was developed by ECL Substrate (Azure or Bio-Rad) if needed. Signal was detected on Odyssey Fc imager (LI-COR Bioscience) or C600 imager (Azure Biosystem) using chemiluminescent or near-infrared mode. The relative quantitation of global O-GlcNAc level was conducted by the ImageStudio Lite software v5.2 (LI-COR). The overall intensity of detected O-GlcNAc bands was normalized to the corresponding actin loading control in each sample.

Transfection of OGT, YFP, NUP62, and mEGF1-36. Human ncOGT with C-terminal 3xFlag-tag was PCR-amplified from pDONR221-OGT (Harvard PlasmID #HsCD0004564) and subcloned into pcDNA5/FRT/TO vector between the KpnI and ApaI restriction sites. The plasmid of pcDNA5/FRT/TO-3xFlag-YFP was purchased from Addgene (#40998). Human NUP62 fused with C-terminal *c*-Myc-His_6-tag was PCR-amplified from pET21a-NUP62 and subcloned into pcDNA3.1/myc-His A vector between KpnI and XbaI restriction sites. The used primers were listed in Table S2. The plasmid of pSectag2/Hygro C/mouse Notch1 EGF1-36-*c*-myc-His_6/IRES-EGFP was a kind gift from Dr. Robert Haltiwanger's lab.[15] For transfection, HEK293 cells were seeded in a 6-well plate in DMEM medium with 10% FBS to reach 70% confluence. The plasmid DNA (3.5 μg for mEGF1-36, 2.5 μg for OGT and NUP62, and 1 μg for YFP) was delivered into the cells in each well using calcium phosphate method as previously described.[16]

6AzES1 treatment. After 48-58 h of transfection, the medium was replaced with low-glucose DMEM medium (0.5 g/L glucose, 5% FBS) containing **6AzES1** at varied concentrations (final 1% DMSO) for 13 or 24 h as indicated. For post-incubation, medium was replaced with **6AzES1**-free low-glucose DMEM medium after inhibitor treatment followed by another 8 h of incubation. To detect the endogenous proteins that are modified by **6AzES1** in cells, HEK293 cells were seeded at 65% confluence for 22 h. **6AzES1** (25 μM) was added to the cells with indicated treatment time. To compete with OGT labeling, glucosamine (15 mM) was added to the medium eight hours before and during **6AzES1** treatment. Cells were harvested by scrapping, snap frozen in liquid N_2, and stored at -80 °C until use.

Click chemistry and in-gel fluorescence scanning to detect 6AzES1-modified proteins in cells. To obtain whole cell lysates, cell pellet was lysed in buffer containing 50 mM TEA pH 7.4, 150 mM NaCl, and 1% NP40. For detecting the endogenous proteins modified by **6AzES1** in nucleus, cell pellet was first resuspended and incubated in buffer containing 50 mM TEA pH 7.4, 10 mM KCl, and 0.5% NP40 for 15 minutes followed by centrifugation at 10,000 g (11,600 r.p.m., F241.5P rotor, Beckman Coulter) for 15 minutes. After the removal of supernatant, the remaining pellet was lysed in the nuclear lysis buffer (50 mM TEA pH 7.4, 225 mM NaCl, 0.2% SDC, and 1% NP40). Cell debris was removed by centrifugation at 16,000 g (13,159 r.p.m., F241.5P rotor)

for 15 minutes. Thiamet G (10 μM) and protease inhibitor cocktail were added during lysis. Protein concentration was measured by BCA protein assay and was adjusted to 1 μg/μL using 50 mM TEA (pH 7.4) and 0.5% NP40. Lysates (30 or 35 μg) from each condition were used for click chemistry reaction. The reagent mixture (1 mM $CuSO_4$, 5 mM tris-hydroxypropyltriazolylmethylamine (THPTA), 50 μM fluor 488-alkyne (Sigma) and 7.5 mM sodium ascorbate) was freshly prepared and immediately added to each sample. Click chemistry reaction was performed at r.t. for 15 minutes or 1 h in the dark. Proteins were precipitated in MeOH overnight at -80 °C followed by centrifugation at 16,000 g (F241.5P rotor) at 4 °C for 10 minutes to pellet the proteins. The pellet was washed by MeOH twice and then re-dissolved in 4% SDS. Protein samples were separated on SDS-PAGE gel and detected by in-gel fluorescence scanning and Coomassie Blue staining on C600 imager (Azure Biosystem).

Immunoprecipitation. After 55 h of mEGF1-36 or NUP62 transfection, cells were treated with **ES1** or DMSO in low glucose DMEM medium (1 g/L glucose, 5% FBS) for another 17 h, and harvested as described above. The collected cells were frozen in liquid N_2 and stored at -80 °C until use. Pellet was lysed in the buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% NP40, 0.1% sodium deoxycholate, 10 μM thiamet G, and protease inhibitor cocktail by incubating on ice for 50 minutes with occasional vortex. Cell debris was removed by centrifugation at 14,000 g (12,309 r.p.m., F241.5P rotor, Beckman Coulter) at 4 °C for 15 minutes. Protein concentration was measured as described above. Cell lysates (1 and 1.3 mg) of were diluted to 1 and 1.3 mg/mL with PBS for NUP62 and mEGF1-36 pull-down, respectively. The lysates were then incubated with anti-c-myc agarose (Life Tech) for 18 h at 4 °C. The resins were washed three times with 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% NP40 followed by one wash with 20 mM Tris-HCl (pH 7.5), 150 mM NaCl. Bound proteins were eluted by two-fold SDS loading buffer at 35 °C. The eluted proteins were analyzed by Western Blot and the relative O-GlcNAc changes were quantified as described above. Three independent experiments were conducted. Data were analyzed using GraphPad Prism v5.

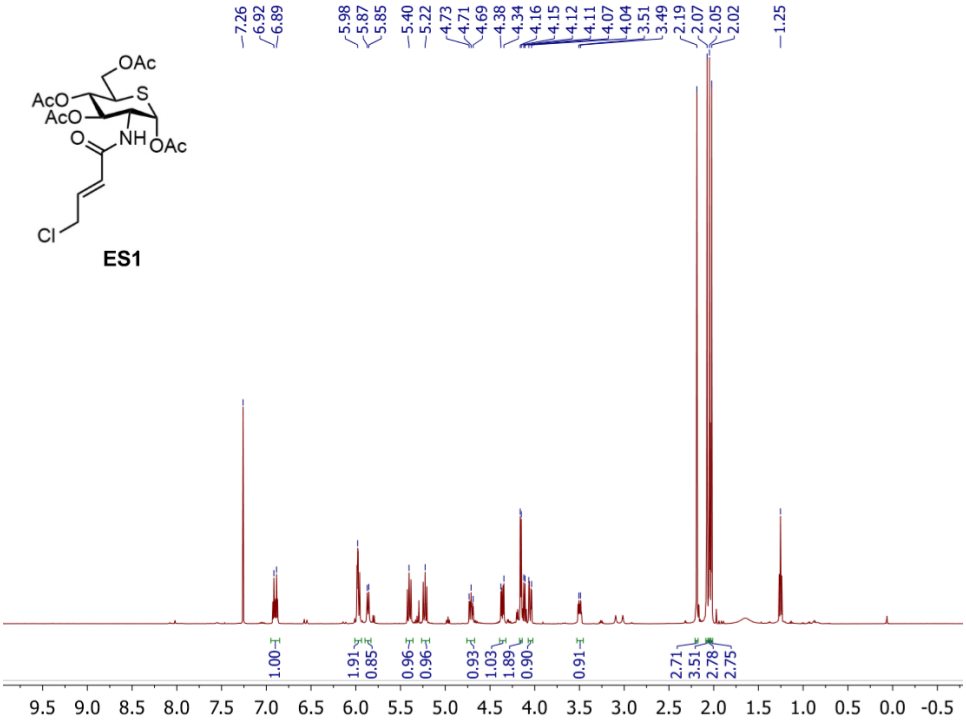
Enrichment of biotinylated 6AzES1-modified proteins. Cells were lysed in buffer containing 50 mM TEA pH 7.4, 10 mM KCl, and 0.5% NP40 followed by centrifugation at 10,000 g (F241.5P rotor) for 20 min. To obtain the nuclear extracts, the remaining pellet was lysed in the lysis buffer containing 50 mM TEA pH 7.4, 225 mM NaCl, 0.2% SDC, and 1% NP40. Cell debris was removed by centrifugation at 16,000 g (F241.5P rotor) for 15 minutes. Thiamet G (10 μM) and protease inhibitor cocktail were added during lysis. Protein concentration was determined as mentioned above. Click chemistry of 300 μg lysates was conducted as described above using 100 μM biotin-PEG4 alkyne (Click Chemistry Tools) followed by MeOH precipitation overnight. Protein pellet was re-solubilized using 50 μL of 8 M urea/0.1 M triethylamine bicarbonate and diluted to 500 μL using 0.5% NP40/PBS. The protein solution was then incubated with NeutrAvidin agarose (Life Tech) for two to three hours at r.t. The resins were washed five times with 0.5% NP40/PBS. Bound biotinylated proteins were eluted by boiling in two-fold SDS loading buffer for five minutes twice. The eluents were analyzed by Western Blot.

IV. References

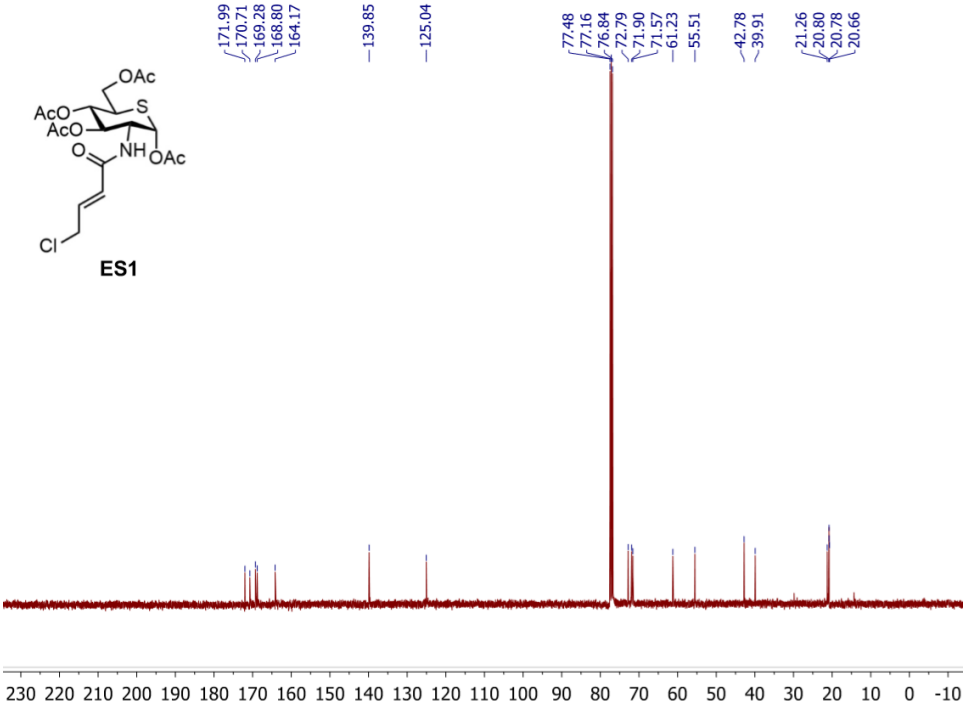
1. M. B. Lazarus, J. Jiang, T. M. Gloster, W. F. Zandberg, G. E. Whitworth, D. J. Vocadlo and S. Walker, *Nat. Chem. Biol.*, 2012, **8**, 966.
2. T. M. Gloster, W. F. Zandberg, J. E. Heinonen, D. L. Shen, L. H. Deng and D. J. Vocadlo, *Nat. Chem. Biol.*, 2011, **7**, 174.
3. R.-B. Yan, F. Yang, Y. Wu, L.-H. Zhang and X.-S. Ye, *Tetrahedron Lett.*, 2005, **46**, 8993.
4. J. A. Tunge and S. R. Mellegaard, *Org. Lett.*, 2004, **6**, 1205.
5. Y. Li, H. Yu, Y. Chen, K. Lau, L. Cai, H. Cao, V. K. Tiwari, J. Qu, V. Thon, P. G. Wang and X. Chen, *Molecules*, 2011, **16**, 6396.
6. W. Guan, L. Cai and P. G. Wang, *Chem. - Eur. J.*, 2010, **16**, 13343.
7. M. B. Lazarus, Y. S. Nam, J. Jiang, P. Sliz and S. Walker, *Nature*, 2011, **469**, 564–567.
8. J. Rappsilber, M. Mann and Y. Ishihama, *Nat. Protoc.*, 2007, **2**, 1896.
9. S. Tyanova, T. Temu and J. Cox, *Nat. Protoc.*, 2016, **11**, 2301.
10. Z. Otwinowski and W. Minor, *Methods Enzymol.*, 1997, **276**, 307.
11. P. Emsley and K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.*, 2004, **60**, 2126.
12. P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, *Acta Crystallogr. D Biol. Crystallogr.*, 2002, **58**, 1948.
13. G. R. Nakayama, M. C. Caton, M. P. Nova, Z. Parandoosh, *J. Immunol. Methods*., 1997, **204**, 205.
14. S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies and D. J. Vocadlo, *Nat. Chem. Biol.*, 2008, **4**, 483.
15. L. Shao, D. J. Moloney and R. Haltiwanger, *J. Biol. Chem.*, 2003, **278**, 7775.
16. M. Wigler, S. Silverstein, L.-S. Lee, A. Pellicer, Y.-C. Cheng and R. Axel, *Cell*, 1977, **11**, 223.

NMR spectra of ES1-ES4, UDP-ES1, and 6AzES1.

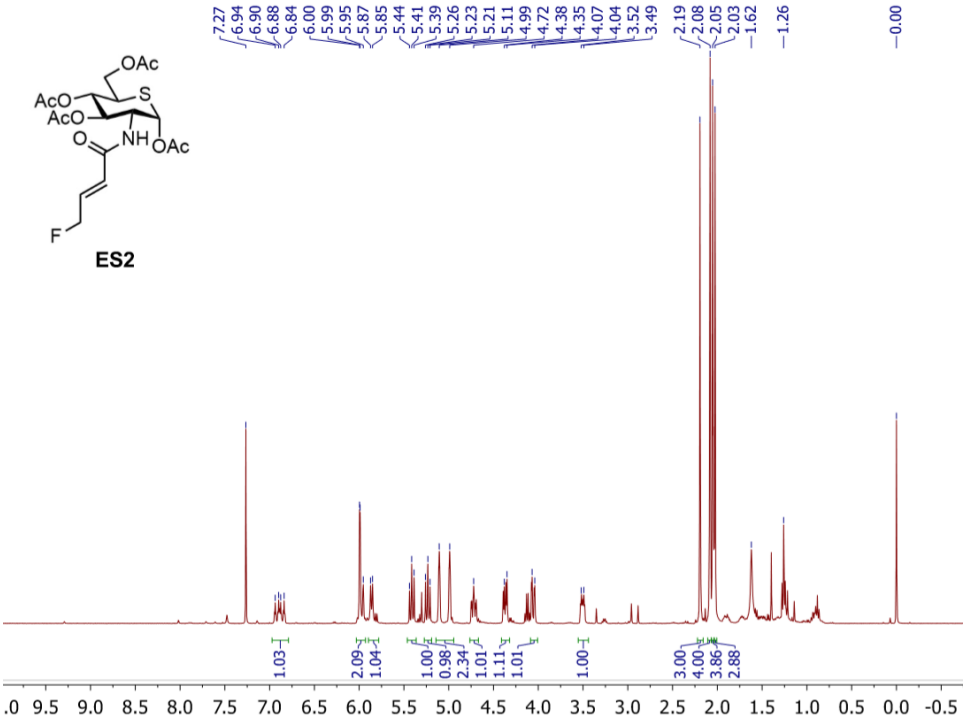
^{1}H NMR spectrum of **ES1**



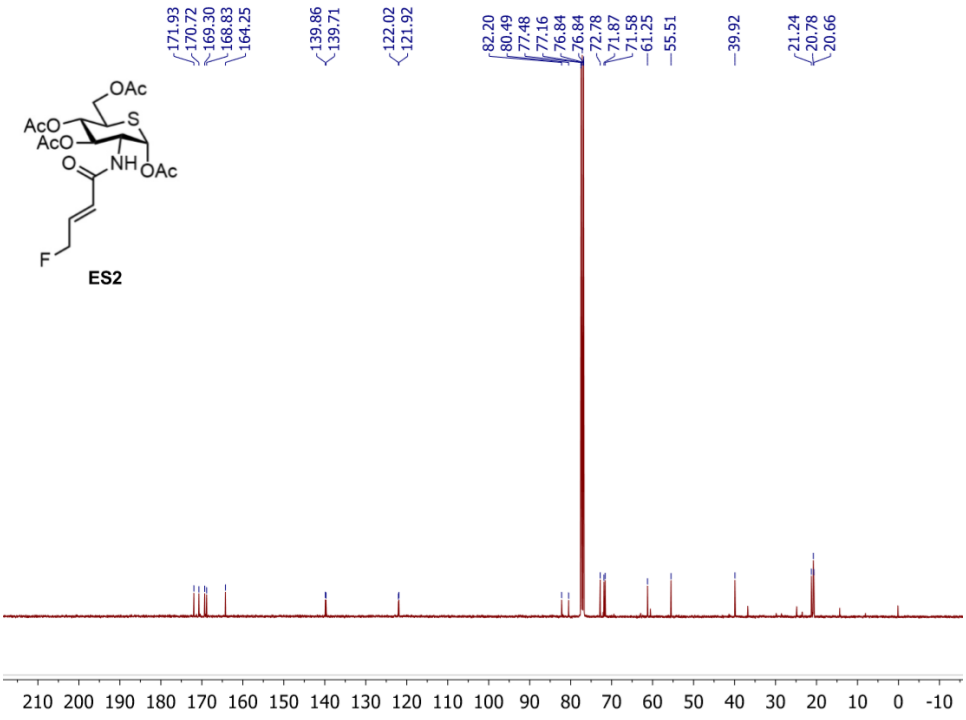
^{13}C NMR spectrum of **ES1**



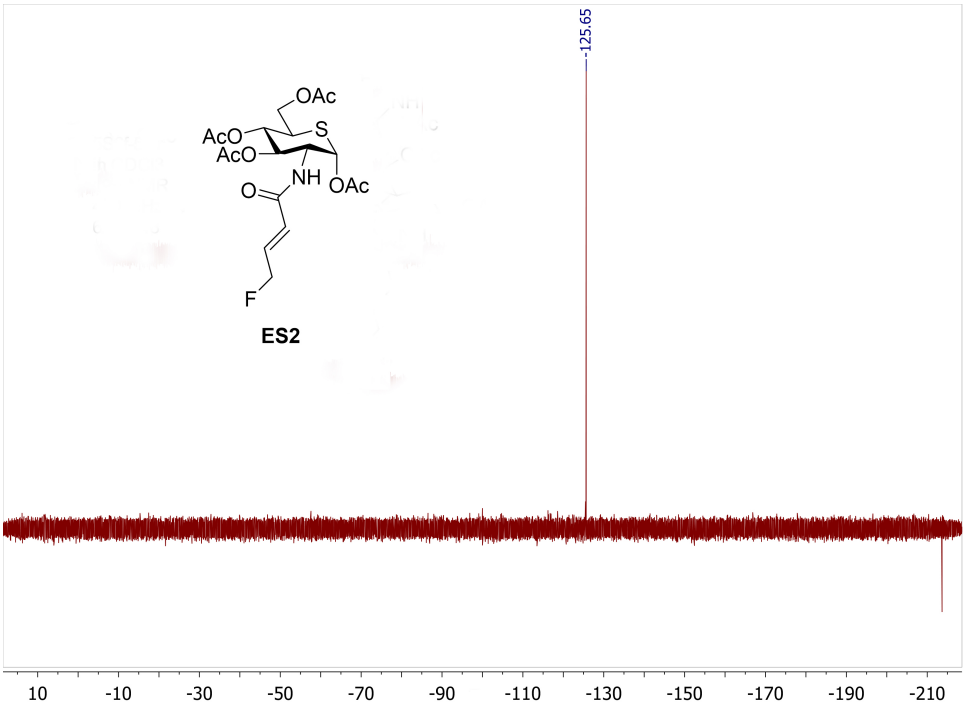
^{1}H NMR spectrum of **ES2**



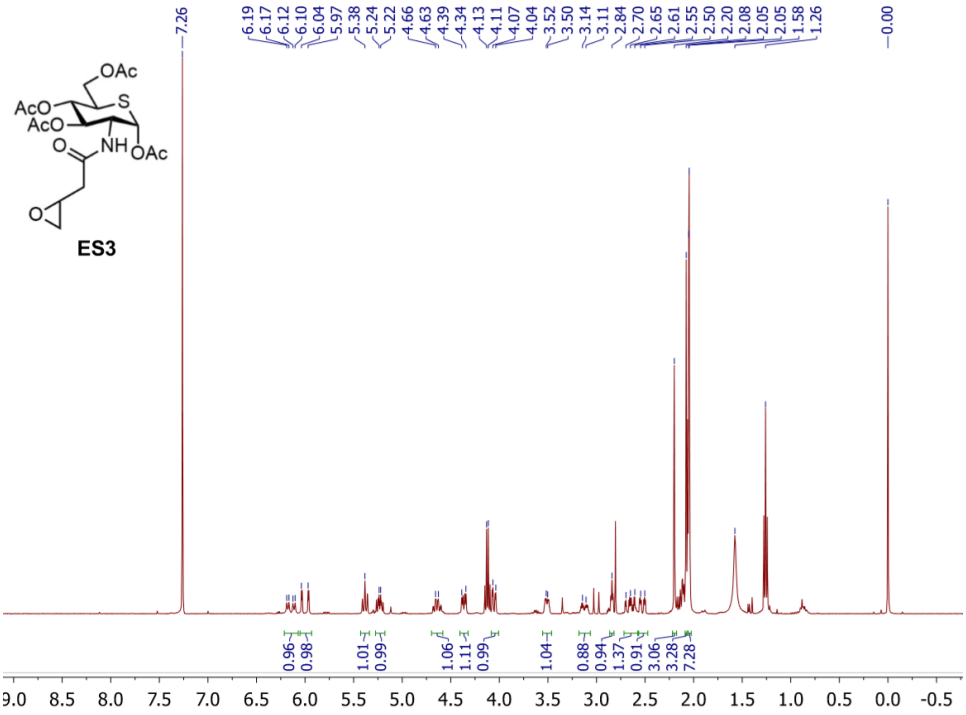
^{13}C NMR spectrum of **ES2**



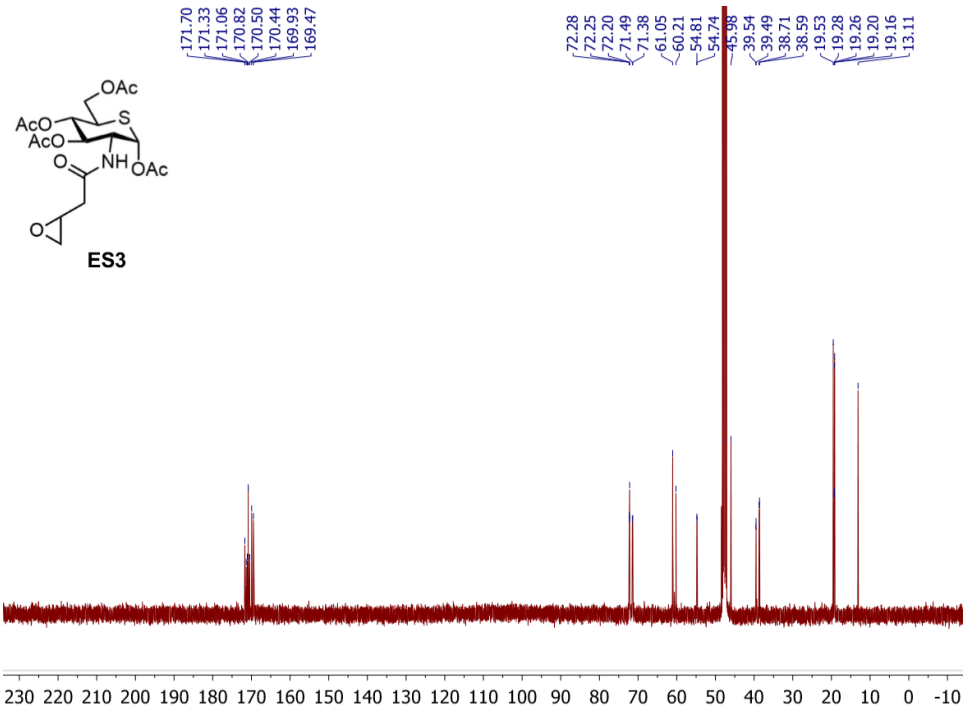
^{19}F NMR spectrum of **ES2**



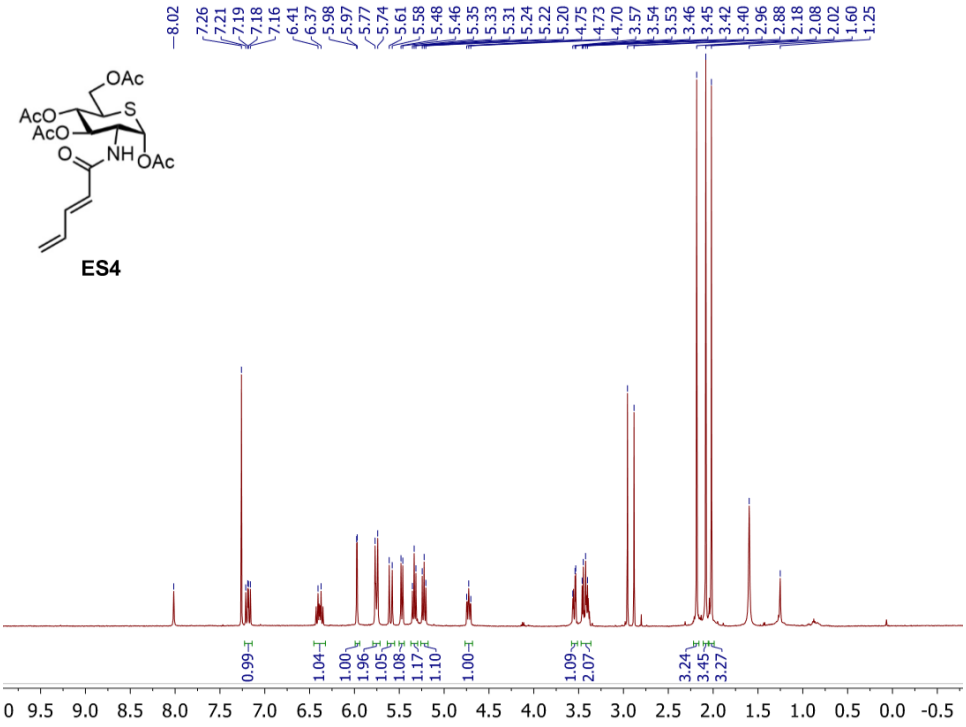
^{1}H NMR spectrum of **ES3**



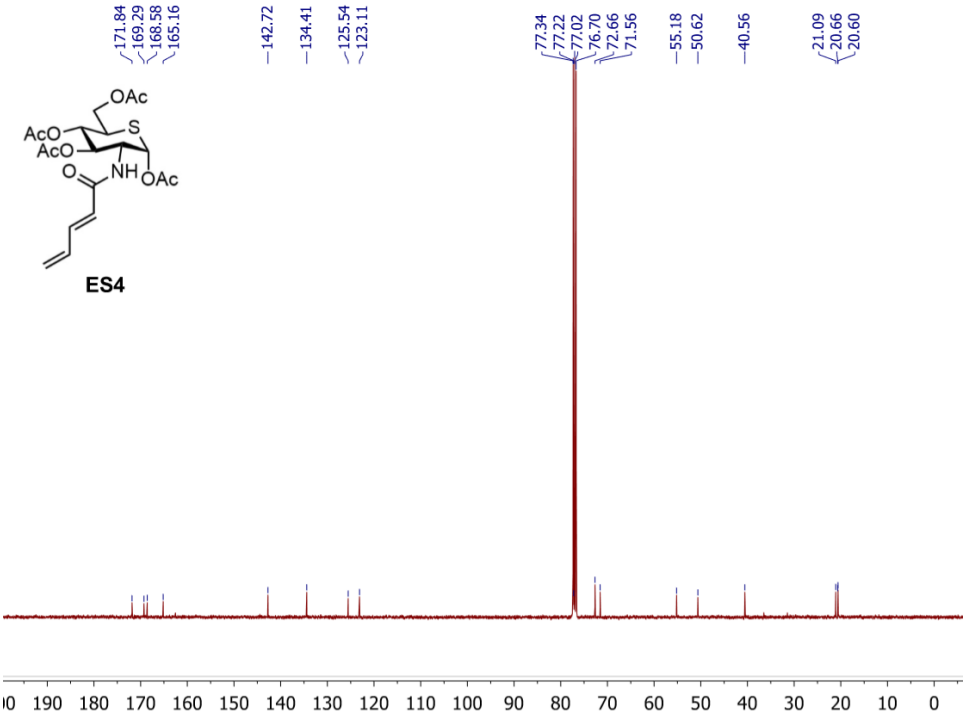
^{13}C NMR spectrum of **ES3**



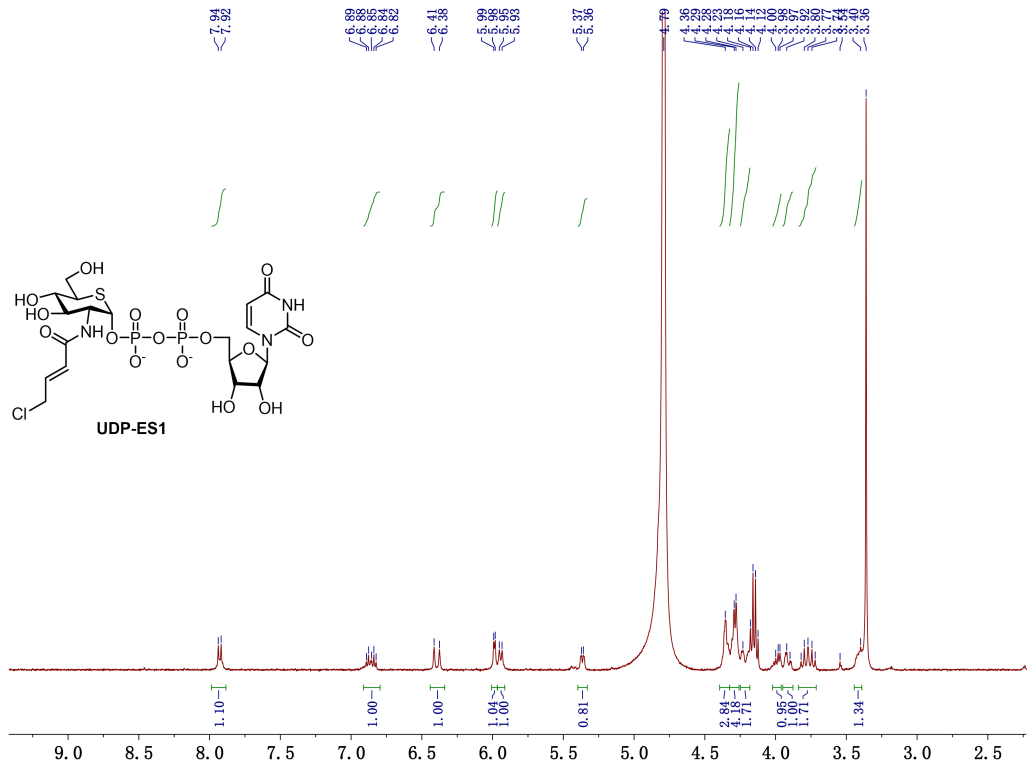
^{1}H NMR spectrum of **ES4**



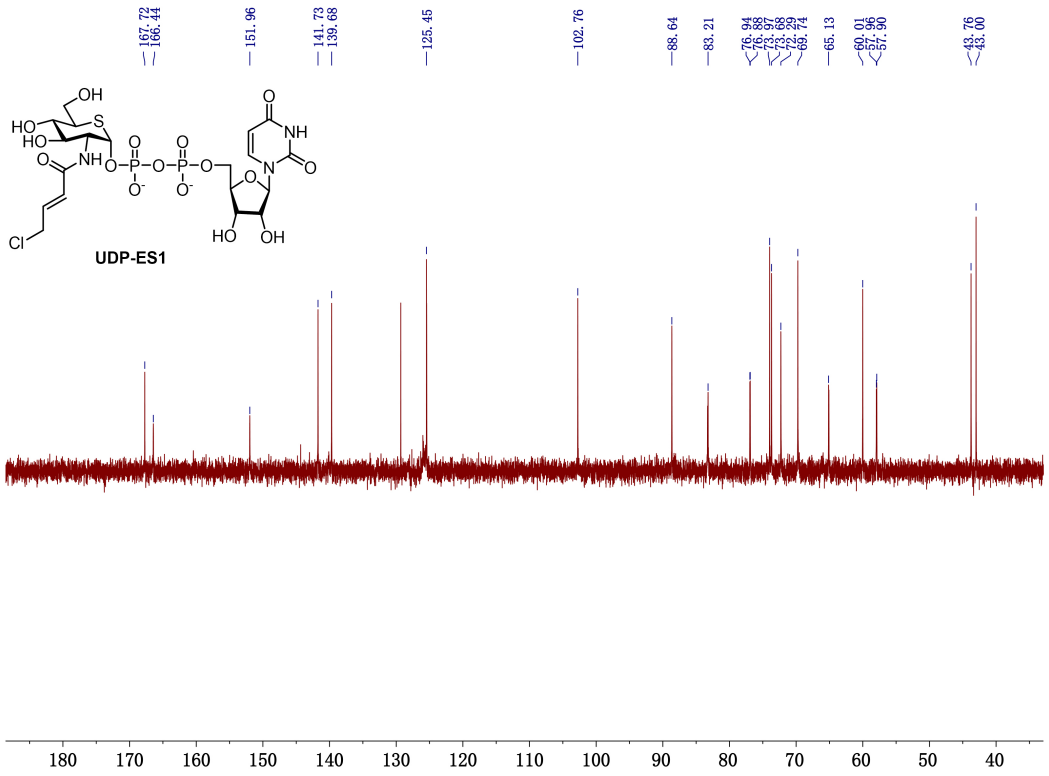
^{13}C NMR spectrum of **ES4**



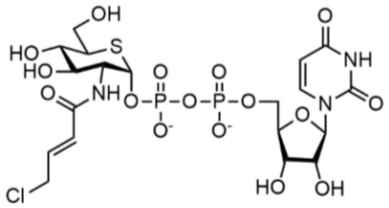
^{1}H NMR spectrum of **UDP-ES1**



^{13}C NMR spectrum of **UDP-ES1**

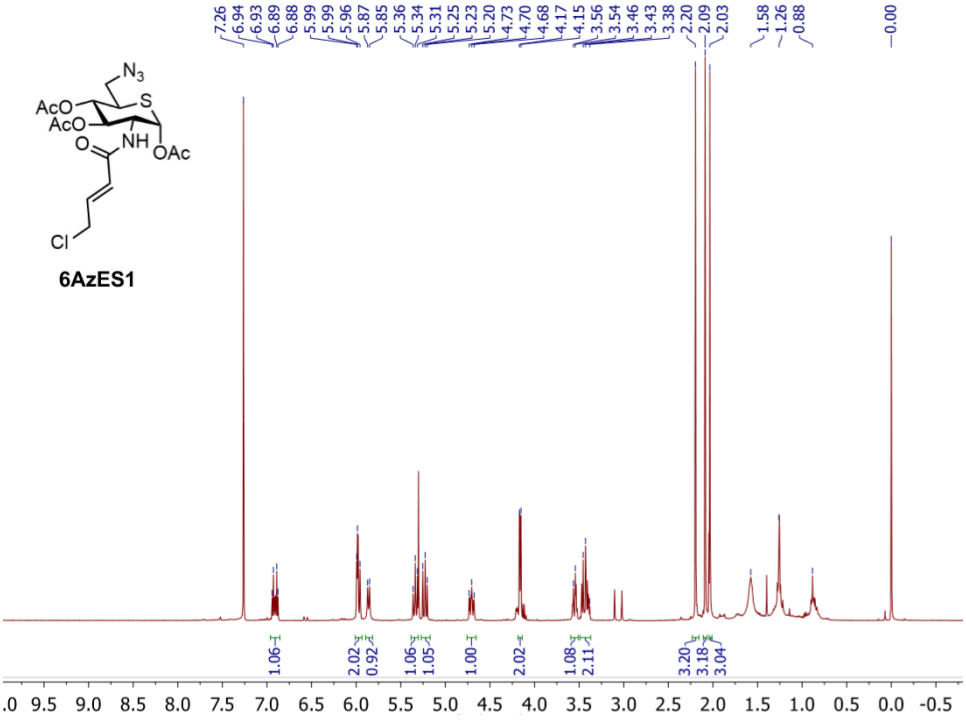


^{31}P NMR spectrum of **UDP-ES1**

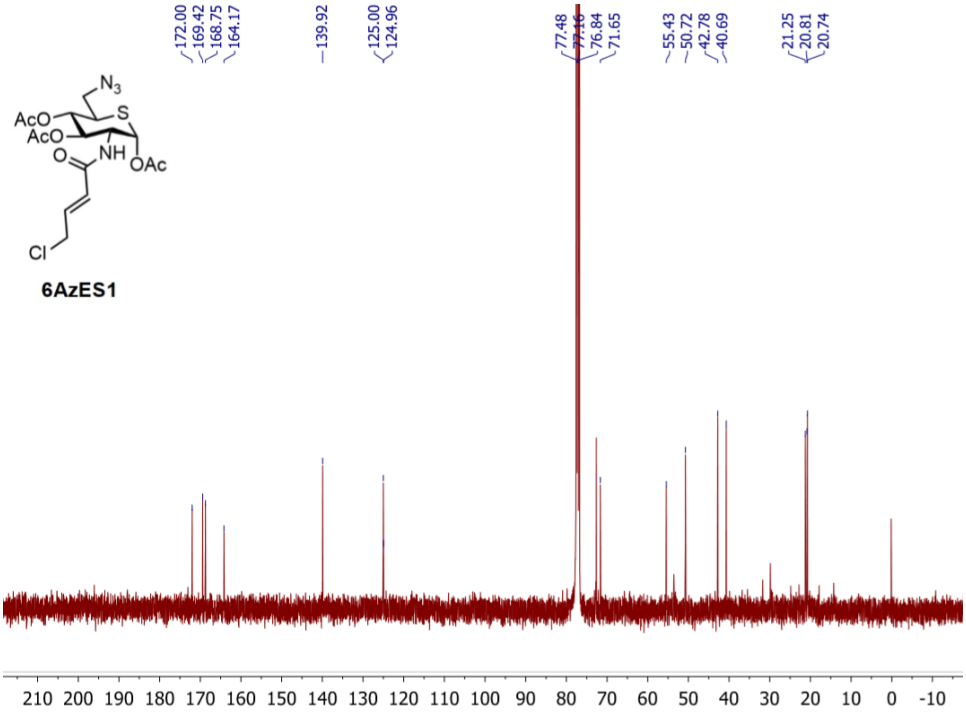


UDP-ES1

^{1}H NMR spectrum of **6AzES1**

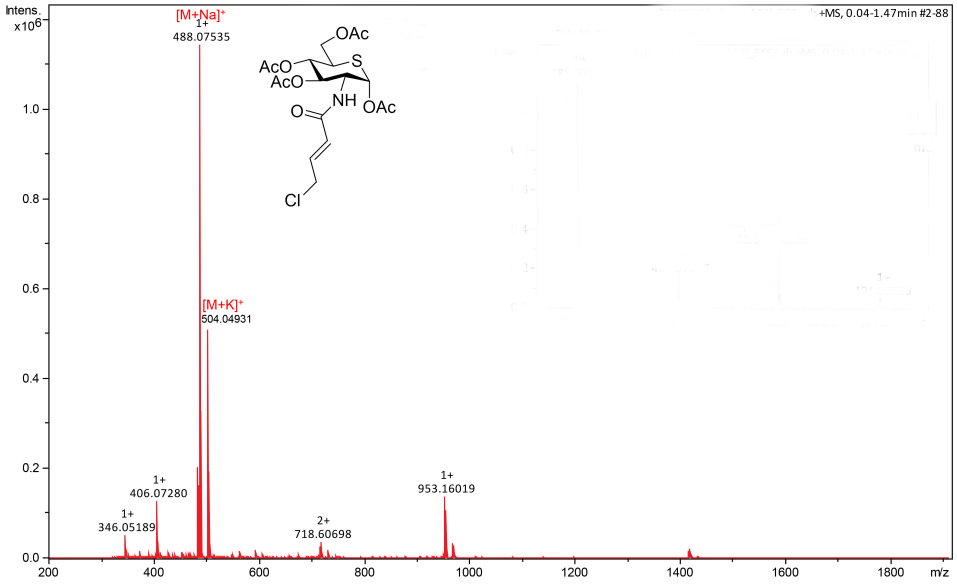


^{13}C NMR spectrum of **6AzES1**

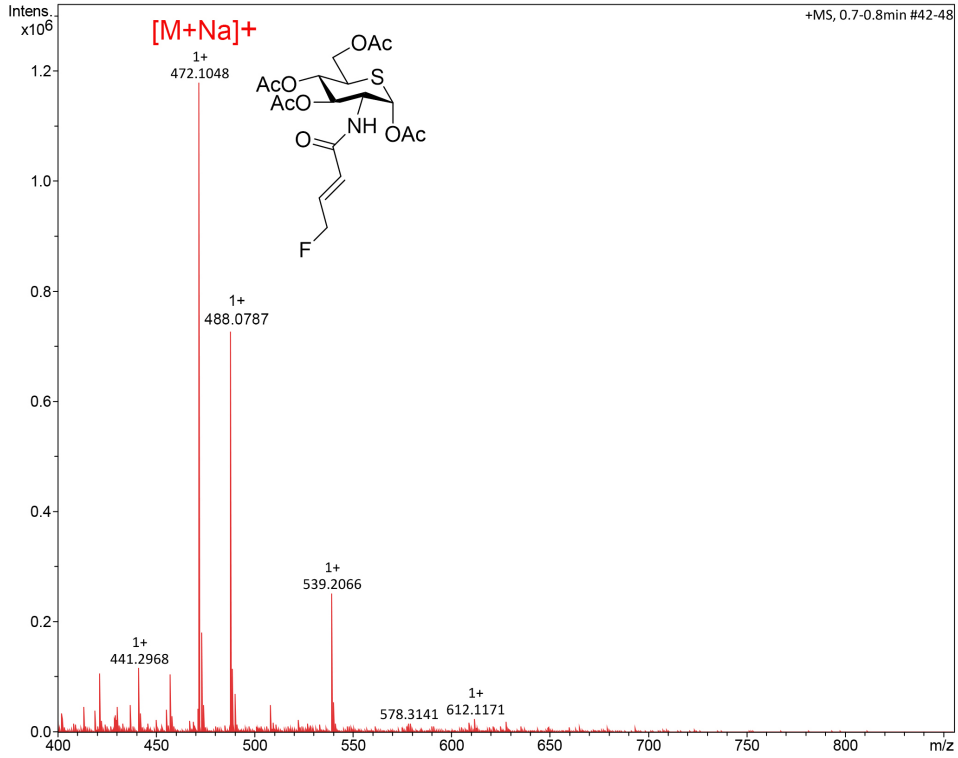


HRMS spectra of ES1-ES4, UDP-ES1, and 6AzES1.

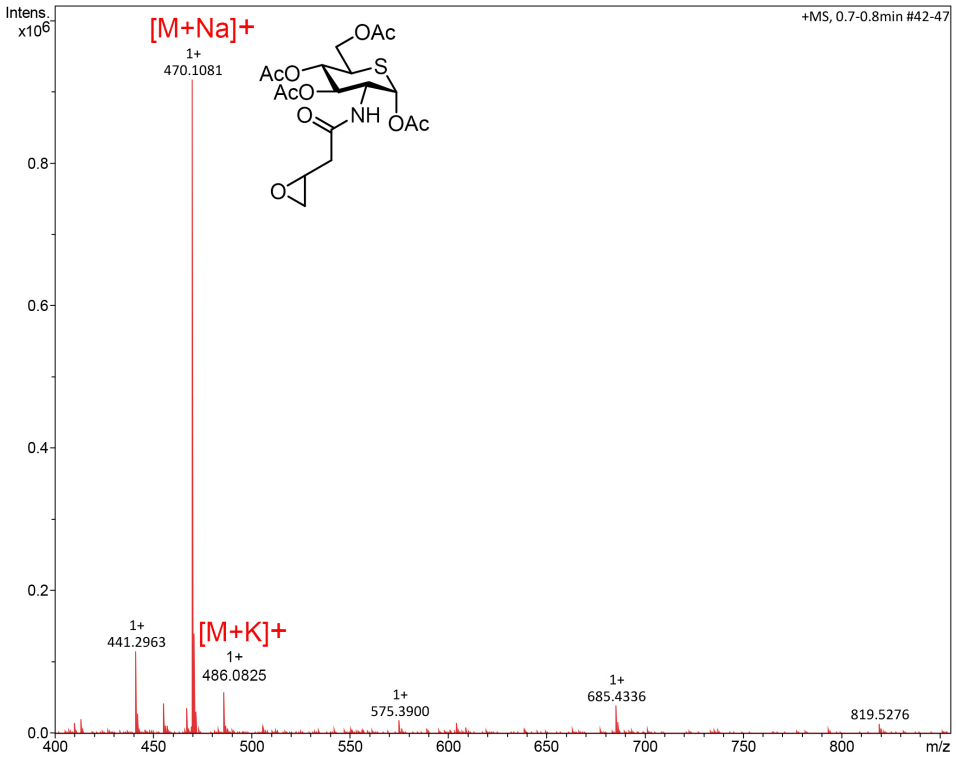
MS spectrum of **ES1**. **ES1** $[M+Na]^+$ = 488.0758; $[M+K]^+$ = 504.0497



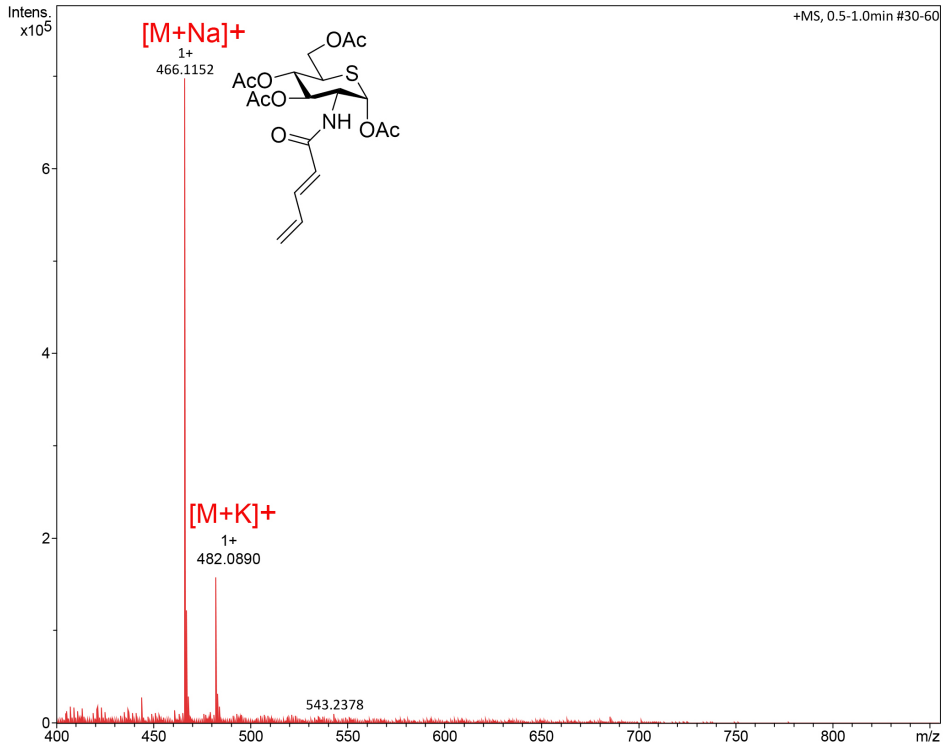
MS spectrum of **ES2**. **ES2** $[M+Na]^+$ = 472.1053; $[M+K]^+$ = 488.0793



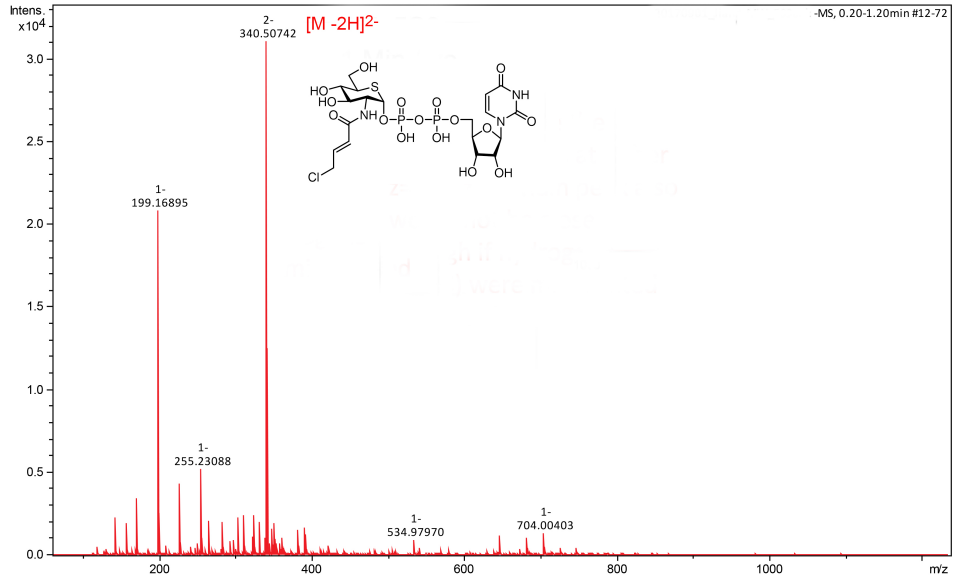
MS spectrum of **ES3**. **ES3** $[M+Na]^+ = 470.1081$; $[M+K]^+ = 486.0836$



MS spectrum of **ES4**. **ES4** $[M+Na]^+ = 466.1148$; $[M+K]^+ = 482.0887$



MS spectrum of **UDP-ES1**. **UDP-ES1** $[M-2H]^{2-}$ = 340.5104



MS spectrum of **6AzES1**. **6AzES1** $[M+Na]^+$ = 471.0717; $[M+K]^+$ = 487.0457