

Supporting Information

An Isotope-Coded Photocleavable Probe for Quantitative Profiling of Protein O-GlcNAcylation

Jingchao Li, Zhonghua Li, Xiaotao Duan, Ke Qin, liuyi Dang, Shisheng Sun, Li Cai,
Linda C. Hsieh-Wilson, liming Wu*, and Wen Yi*

To whom correspondence should be addressed. E-mail: wlm@zju.edu.cn,
wyi@zju.edu.cn

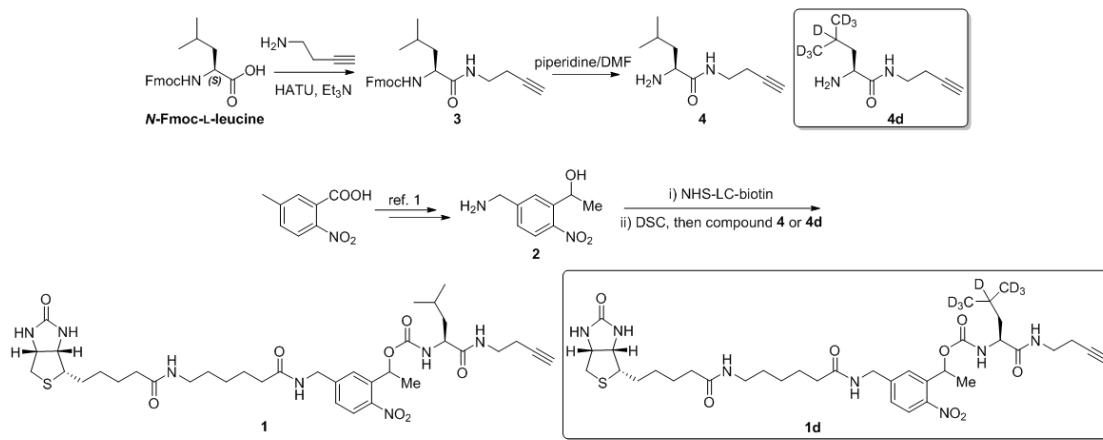
This file includes:

SI Materials and Methods

SI Figures and Tables

Materials. All chemical reagents were obtained from commercial suppliers, and used without further purification unless otherwise stated. NHS-LC-biotin and O-GlcNAcylated model peptide TAPT(gS)TIAPG were purchased from Thermo Scientific. OGA inhibitor ThiaMet-G was obtained from Tocris Bioscience. Site-directed mutagenesis kit was purchased from Agilent Technologies. All other reagents were obtained from Sigma-Aldrich.

Synthesis of the isotope-labeled photocleavable probes **1** ($^1\text{H}_7$) and **1d** ($^2\text{D}_7$).



1-(2-nitro-5-((6-((5-((3aS,4S,6aR)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamido)hexanamido)methyl)phenyl)ethyl ((S)-1-(but-3-yn-1-ylamino)-4-methyl-1-oxopentan-2-yl)carbamate (1**). 1-(5-(Aminomethyl)-2-nitrophenyl)ethanol (**2**) was synthesized as previously described in literature¹. To a solution of NHS-LC-biotin (250 mg, 0.55 mmol) in DMF (8 mL), compound **2** (131 mg, 0.67 mmol) was added followed by triethylamine (115 μ L, 0.82 mmol). The reaction was stirred at room temperature for overnight. The solvent was removed under reduced pressure, and the residue was purified by flash silica gel chromatography (CH₂Cl₂/CH₃OH 10:1) to afford the conjugate (206 mg, 70%). m/z calcd for C₂₅H₃₈N₅O₆S [M + H]⁺: 536.2537. MS found: 536.2. The conjugate (200 mg, 0.37 mmol) was redissolved in DMF (10 mL) and *N,N'*-disuccinimidyl carbonate (143 mg, 0.56 mmol) and triethylamine (155 μ L, 1.11 mmol) were added. After stirring for overnight, **L-leucinylbuty-nylamide (4)** (337 mg,**

1.85 mmol) was added and the reaction mixture was stirred at room temperature for another 24 h. Upon completion, the solvent was removed under reduced pressure and the residue was purified by flash silica gel chromatography ($\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$ 10:1) to afford the target compound (**1**) (150 mg, 54%). ^1H NMR (400 MHz, $\text{DMSO}-d_6$), δ (ppm): 8.43 (m, 1H), 8.03-7.94 (m, 2H), 7.74 (t, $J = 5.6$ Hz, 1H), 7.58 (d, $J = 20.8$ Hz, 1H), 7.45 (m, 1H), 7.36 (d, $J = 8.4$ Hz, 1H), 6.41 (s, 1H), 6.35 (s, 1H), 6.02-5.96 (m, 1H), 4.39-4.29 (m, 3H), 4.12 (m, 1H), 3.84 (m, 1H), 3.17-3.00 (m, 5H), 2.81 (m, 2H), 2.58 (d, $J = 13.2$ Hz, 1H), 2.18 (m, 4H), 2.04 (m, 2H), 1.68-1.22 (m, 18 H), 0.85 (m, 5H), 0.70 (m, 1H). m/z calcd for $\text{C}_{36}\text{H}_{54}\text{N}_7\text{O}_8\text{S} [\text{M} + \text{H}]^+$: 744.3749. MS found: 744.3 (M + H). The deuterated compound **1d** was synthesized following the same procedure described above using the deuterated leucine derivative (**4d**). ^1H NMR (400 MHz, $\text{DMSO}-d_6$), δ (ppm): 8.42 (m, 1H), 8.02 (m, 1H), 7.95 (d, $J = 8.0$ Hz, 1H), 7.74 (m, 1H), 7.55 (s, 1H), 7.43 (d, $J = 8.4$ Hz, 1H), 7.36 (d, $J = 8.4$ Hz, 1H), 6.41 (s, 1H), 6.35 (s, 1H), 5.97 (m, 1H), 4.35-4.28 (m, 3H), 4.14-4.11 (m, 1H), 3.83 (m, 1H), 3.17 (m, 2H), 3.10 (m, 2H), 3.01 (m, 2H), 2.80 (m, 2H), 2.57 (d, $J = 13.2$ Hz, 1H), 2.26 (m, 2H), 2.17 (m, 2H), 2.06-2.02 (m, 2H), 1.58-1.22 (m, 16H). m/z calcd for $\text{C}_{36}\text{H}_{47}\text{D}_7\text{N}_7\text{O}_8\text{S} [\text{M} + \text{H}]^+$: 751.4188. MS found: 751.4 (M + H).

N-Fluorenylmethyloxycarbonyl-L-leucinylbuty-nylamide (3). 3-Butynylamine (95% purity, 0.52 ml, 6.0 mmol) was added to *N*-Fmoc-L-leucine (1414 mg, 4.0 mmol) and HATU (1673 mg, 4.4 mmol) in DMF (12 mL) at room temperature under nitrogen. The solution was cooled on ice and triethylamine (1.67 mL, 12 mmol) was added, the reaction mixture was then allowed to reach room temperature and stirred for 24 h. Upon completion, the reaction mixture was diluted with EtOAc (100 ml), washed with 0.1 M HCl (2 × 30 ml), saturated NaHCO_3 (30 ml), brine (50 ml), dried over anhydrous Na_2SO_4 and the solvents were removed under reduced pressure. The residue was purified by flash silica gel chromatography (hexane/EtOAc 1:1). **N-Fluorenylmethyloxycarbonyl-L-leucinylbuty-nylamide (3)** was afforded as a white solid (880 mg, 2.18 mmol, 54%). ^1H NMR (400 MHz, CDCl_3), δ (ppm): 7.77 (d, $J = 7.6$ Hz, 2H), 7.58 (d, $J = 7.6$ Hz, 2H), 7.41 (m, 2H), 7.32 (m, 2H), 6.27 (brs, 1H), 5.17 (d, $J = 7.6$ Hz, 1H), 4.43 (m, 2H), 4.21 (m, 2H), 3.40 (m, 2H), 2.39 (brs, 2H), 1.96 (s, 1H), 1.63 (m, 3H), 0.94 (s, 6H).

L-Leucinylbuty-nylamide (4). A 20% (v/v) solution of piperidine in DMF (20 mL) was added to a residue of (**3**) (780 mg, 1.93 mmol). After 10 min, a precipitate formed and 3

mL DMF was added to the mixture. After 30 minutes, volatiles were removed reduced pressure. H₂O was added to the residue and the mixture was stirred vigorously and then stored in the refrigerator overnight. The precipitate was removed via filtration, and washed with cold H₂O. The aqueous layer was extracted CH₂Cl₂ (3 × 50 mL) and EtOAc (3 × 50 mL). The combined organic layers were dried over Na₂SO₄ and concentrated to provide leucine derivative (**4**) 240 mg (1.32 mmol, 68%). ¹H NMR (400 MHz, CDCl₃), δ (ppm): 7.60 (brs, 1H), 3.41-3.35 (m, 3H), 2.40-2.36 (m, 2H), 1.99 (t, *J* = 2.8 Hz, 1H), 1.70-1.67 (m, 2H), 1.47 (s, 2H), 1.33 (t, 1H), 0.91 (m, 6H). The deuterated leucine derivative **4d** was synthesized following the same procedure described above.

Labeling and quantification of model O-GlcNAcylated peptide. 20 μg of O-GlcNAcylated peptide (H-Thr-Ala-Pro-Thr-(O-GlcNAc)Ser-Thr-Ile-Ala-Pro-Gly-OH) was dissolved in 20 mM HEPES pH 7.9, and labeled using the Click-iT O-GlcNAc Enzymatic Labeling System (Invitrogen) for 2 h at 37°C. The reaction mixture was then filtered through a 2000 Da MWCO spin filter and injected into a reverse-phase HPLC C18 column (Spherisorb; 5μm, 250 x 4.6 mm). Agilent 1260 HPLC system was used to generate the following HPLC gradient: 12-18% B in 10 min, 18-40% B in 3 min, 40-55% B in 5 min (A: 0.1% Trifluoroacetic acid, B: Acetonitrile) with a flow rate of 1 mL/min. Then the product was separated by HPLC and lyophilized. The GalNAz-labeled O-GlcNAc peptide was divided equally into two portions, which were conjugated with ¹H₇ coded photocleavable probe and ²D₇ coded photocleavable probe respectively. The CuAAC (the copper(I)-catalyzed azide alkyne cycloaddition, 0.1 mM isotopic coded probe, 0.25 mM CuSO₄, 0.5 mM BTTAA ligand, 2.5 mM sodium ascorbate) was proceeded for 2 h at room temperature in the dark. Then the isotope-labeled peptide mixtures were combined and purified by HPLC using the conditions described above, and lyophilized. The purified isotope-labeled O-GlcNAc peptides were dissolved in the binding buffer (100 mM phosphate buffer, pH 7.2, 150 mM NaCl), and incubated with streptavidin resins (Pierce) with end-to-end rotation for 2 h at room temperature in the dark. Resins were washed with 1 mL of Milli-Q water for 5 times, 70% methanol twice, and then resuspended in 70% methanol. The bound glycopeptides were released from resins by UV (365 nm) irradiation for 1 h. The supernatant was collected and analyzed by

HPLC and further quantified by electron-transfer/higher-energy collision dissociation (EThcD) based MS/MS as described below.

Cell culture and lysate preparation. The cell lines 293T, MCF-7, HepG2, and A549 were obtained from ATCC and cultured according to ATCC protocols. Adriamycin-resistant MCF-7 cells, sorafenib-resistant HepG2 cells, and docetaxel-resistant A549 cells were obtained from GeneChem LLC, China, and independently verified using laboratory procedures. After cells were collected, cells were washed with cold PBS for three times, and lysed on ice with RIPA lysis buffer (50 mM Tris, 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% sodium deoxycholate, 0.1% (w/v) SDS, pH 7.4) containing EDTA-free protease inhibitor cocktail (Roche) and 10 µM ThiaMet-G. After sonication, cell lysates were centrifuged at 15,000 g for 15 min at 4 °C. The supernatant was collected and the total protein concentration was measured by the BCA protein assay kit (Beyotime biotechnology).

Generation of gene expression vectors and mutants. The coding sequences of ZYX, ZHX3 (cDNA clones obtained from Genescript) were cloned into p3XFlag-CMV-7.1 vector (Sigma-Aldrich) to form Flag-ZYX-WT (with the forward primer, 5'-ccaaagcttatggcgcccccccggccgtct-3', and the reverse primer, 5'-cgggatcctcaggctggcttagcagt-3'), Flag-ZHX3-WT (with the forward primer, 5'-ccaaagcttatggccagcaagagggaaatccacca-3', and the reverse primer, 5'-cgggatcctcagtcgtttcgagctgacgtcca-3') plasmids. Flag-ZYX-S169A (with the forward primer, 5'-caaagccgggtggcatctggatatgtgccccac-3', and the reverse primer, 5'-catatccagatgccaccgggtttgaaaggatc-3'), Flag-ZHX3-S240A (with the forward primer, 5'-gttccagtcgcccaggcatctgcagctctgc-3', and the reverse primer, 5'-gatgcctggcgactggaactgccccattgatg-3') mutants were generated using QuikChange II Site Directed Mutagenesis Kit (Agilent Technologies).

Western-blotting analysis of O-GlcNAcylated proteins. Cell lysates were subjected to the GalT enzymatic labeling reaction, and further conjugated with an alkyne-biotin compound as per the Click-iT Protein Analysis Detection Kit protocol (Invitrogen).

Biotinylated lysates were precipitated, resolubilized in 1% SDS, and neutralized with an equal volume of neutralization buffer (6% NP-40, 100 mM Na₂HPO₄, 150 mM NaCl). Lysates were then incubated with streptavidin resin (Pierce) with end-to-end rotation at 4 °C overnight. Resin was then washed 5 times with 1 mL of low salt buffer (100 mM Na₂HPO₄, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 0.5% sodium deoxycholate) and 5 times with 1 mL of high salt buffer (100 mM Na₂HPO₄, 500 mM NaCl, 0.2% Triton X-100). Biotinylated proteins were eluted by boiling the resin in 50 mM Tris-HCl pH 6.8, 2.5% SDS, 100 mM DTT, 10% glycerol and 20 mM biotin for 10 min. The eluents were then resolved on a 4-12% SDS-PAGE gel, transferred to Immobilon-FL PVDF membrane (Millipore), and immunoblotted with the indicated antibodies.

Quantification of protein O-GlcNAcylation in 293T cell lysates. 293T cell treated with DMSO or TMG (50 µM for 12 hours) were harvested as above-mentioned. 4 mg of total protein mixtures from each sample were chemoenzymatic labeled as above. To eliminate the impact of N-glycan, PNGase F (New England Biolabs) was added in the labeling buffer. The labeled proteins were dissolved in 1% SDS (in 50 mM Tris-HCl pH 8.0), and conjugated with ¹H₇ labeled and ²D₇ labeled probes, respectively. The reaction mixture was incubated for 2 h at room temperature with gentle agitation in the dark. Then the two portions (1:1) were mixed together, and precipitated using methanol and chloroform as described in the Click-iT Protein Analysis Detection Kit. The labeled proteins were solubilized in 800 µL 8M urea, diluted to 10-fold volume with 50 mM NH₄HCO₃, pH 8.1, and further digested with sequencing grade modified trypsin (Promega) at a ratio of 1:100 trypsin : protein (wt/wt) for 16 h at 37 °C in the dark. The reaction was quenched with 0.4% TFA, 2% acetonitrile added, desalted with Sep-Pac Vac tC18 Cartridges (Waters), and dried in a vacuum centrifuge. Peptides were dissolved in binding buffer (100 mM phosphate buffer, pH 7.2, 150 mM NaCl), and incubated with streptavidin resins with end-to-end rotation for 2 h at room temperature in the dark. Resins were washed with 1 mL of binding buffer for 10 times, Milli-Q water for 6 times, 20% methanol twice, 70% methanol once, and then resuspended in 70% methanol. The captured peptides were released from resins by UV irradiation (365 nm) for 1 h. The

supernatant was collected, dried in a vacuum centrifuge, and then analyzed by CID and ETD based LC-MS/MS as described below.

LC-MS/MS analysis. The labeled peptides were desalted with C18 Zip-Tips (Merck) and then loaded on an in-house packed capillary reverse-phase C18 column (15 cm length, 100 mM ID x 360 mM OD, 3 mM particle size, 100 Å pore diameter, Phenomenex) connected to a Thermo Easy-nLC1000 HPLC system (Thermo Fisher, SJ). The samples were analyzed with a 120 min-HPLC gradient from 0% to 100% of buffer B (buffer A: 0.1% formic acid in water; buffer B: 0.1% formic acid in 20/80 water/acetonitrile) at 300 nL/min. Ionization was performed on a Q-Exactive mass spectrometer (Thermo Fisher, SJ) via an EASY-Spray ion source. Survey full-scan MS spectra (from m/z 300–1800) were acquired in the Orbitrap analyzer at a resolution of 70,000 at m/z 400 (target value of 1,000,000 ions, maximum injection time 20 ms). For CID acquisition, data dependent MS2 scan was acquired with a resolution of 35,000 at m/z 400 (target value of 500,000 ions, maximum injection time 1,000 ms). The activation time was set at 30 ms, the isolation width was 1.5 amu, the normalized activation energy was 35%, and the activation q was 0.25. Ions with unassigned charge state, and singly or highly (> 8) charged ions were rejected. Intensity threshold was set to 2.1×10^4 units. Peptide match was set to preferred and dynamic exclusion option was enabled (exclusion duration 40 s).

MS2 experiments by ETD were conducted on an Orbitrap Elite mass spectrometer (Thermo Scientific) equipped with a nano-ESI source. The mass spectrometer was programmed to record a full-scan ESI mass spectrum (m/z = 300-1700) followed by four data-dependent MS/MS scans (70-100 ms ETD activation; 2 Da isolation window). The normalized activation energy and the activation q were set same with CID method.

MS2 experiments by EThcD were conducted on an Orbitrap Fusion Lumos (Thermo Scientific) mass spectrometer equipped with a Nano-Spray ion source. Survey full-scan MS spectra (from m/z 350–1800) were acquired in the Orbitrap analyzer at a resolution of 60,000 (target value of 400,000 ions, maximum injection time 50 ms). EThcD was activated and calibrated charge dependent ETD parameters were used. Isolation window was set at 2 Da.

LC-MS/MS data analysis. CID mass spectrometric raw data files were processed with MASCOT against the Uniprot human database; Mass tolerance was set to be 20 ppm for precursor and 0.5 Da for product ion. Missed cleavages were no more than two for each peptide. O-GlcNAc-tag adduction to Serine (Ser) and Threonine (Thr), and Oxidation (Met), Acetyl (N terminus) were set as variable modifications; Carbamidomethyl (Cys) was chosen as a fixed modification. The identification data were filtered to a 1% false discovery rate (FDR) on both peptide and protein level using target-decoy strategy. MS2 spectra of tagged peptides were manually validated and annotated based on following criteria: (a) good coverage of y-and b-ion series, (b) extensive identification rate of intensive fragment ion peaks, (c) the probability of site localization was checked to be greater than 75%, and (d) Existence of a diagnostic peak corresponding to a fragment ion of photocleaved GalNAz moiety (349.2 Da or 427.2 Da for ¹H₇ coded, 356.2 Da or 434.3 Da for ²D₇ coded) or an identification of both light- and heavy-labeled peptides corresponding to the same O-GlcNAc-tag site. ETD raw data files were processed with MaxQuant software integrated with Andromeda search engine against the Uniprot human database. The parameters were set as CID data processing. For quantification, the intensity of labeled peptides was represented by the peak area of the extracted ion chromatogram of their monoisotopic peaks.

Quantification of O-GlcNAcylated peptides in sorafenib -sensitive and -resistant HepG2 cells. 4 mg cell lysates of each sample (-sensitive or -resistant) were chemoenzymatically labeled, conjugated with ¹H₇ coded and ²D₇ coded probes, respectively, and further treated for CID and ETD based MS/MS analysis according to the protocols shown above.

Reference:

- (1) Szychowski, J., Mahdavi, A., Hodas, J. J., Bagert, J. D., Ngo, J. T., Landgraf, P., Dieterich, D. C., Schuman, E. M., and Tirrell, D. A. (2010). Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J. Am. Chem. Soc.*, 132, 18351-18360.

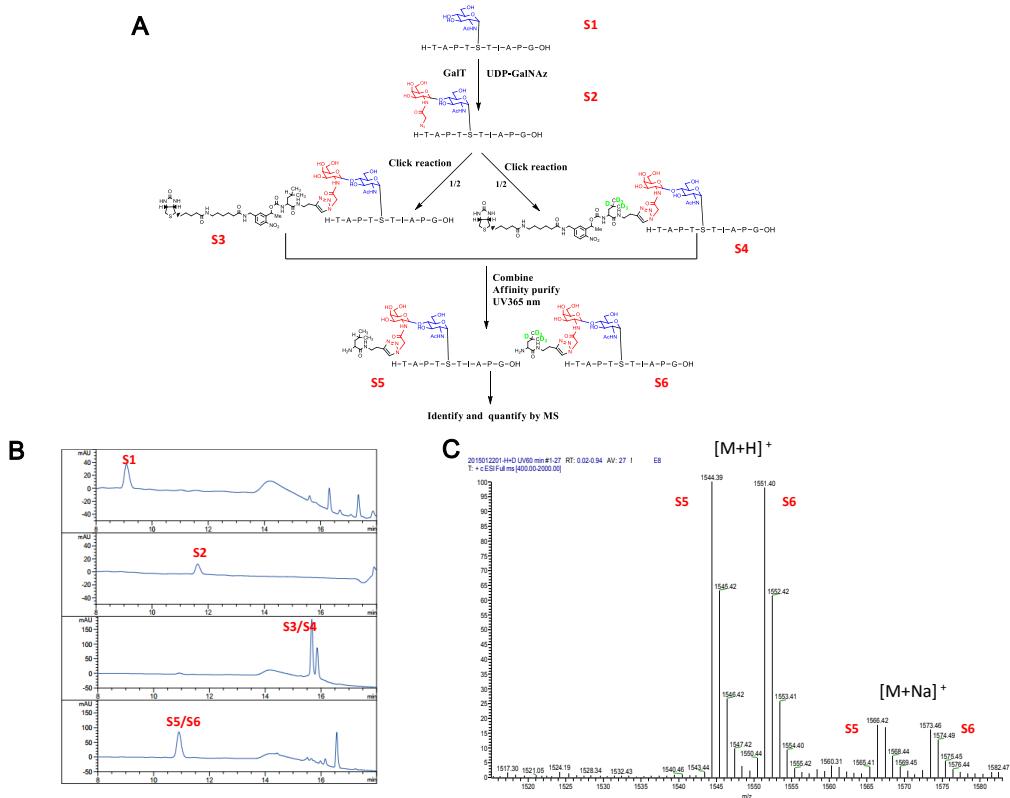


Figure S1. Labeling and quantification of a model O-GlcNAcylated peptide. (A) Glycopeptides were chemo-enzymatically tagged with GalNAz and the isotope-coded probes. Equal amounts of the isotopically labeled peptides were combined, affinity captured with streptavidin, and further subjected to UV photocleavage. (B) The reaction progress was monitored using liquid chromatography, and (C) relative quantification of O-GlcNAcylated peptide pairs was performed using an orbitrap mass spectrometry.

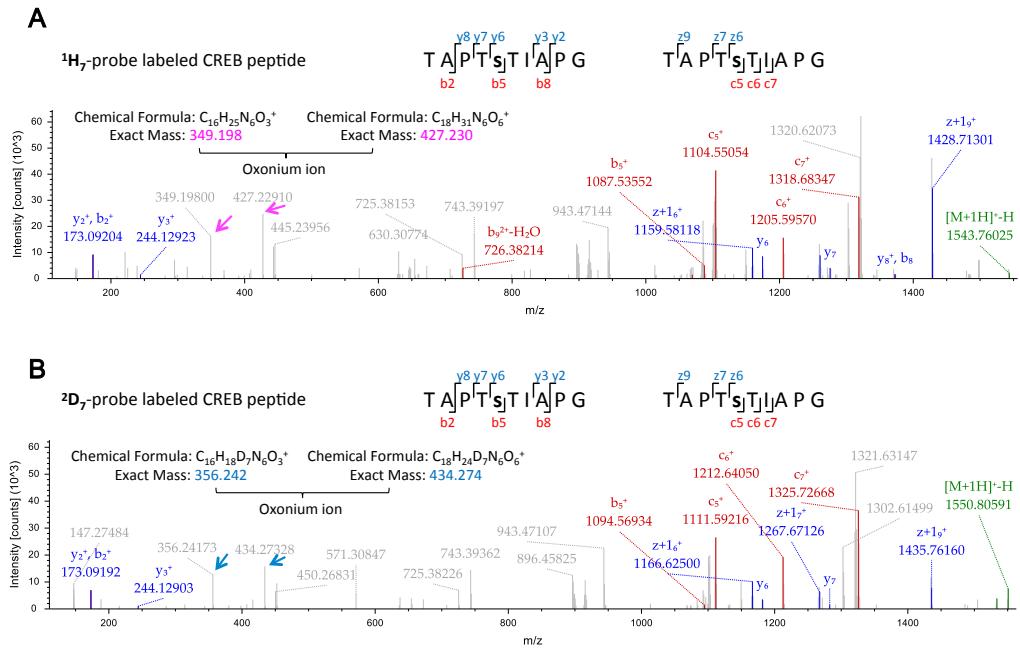


Figure S2. Fragment spectra of isotopically labeled CREB glycopeptides dissociated by EThcD. (A) $^1\text{H}_7$ coded and (B) $^2\text{D}_7$ coded probe exhibited the same site localization and similar fragment spectra, where S5 was the O-GlcNacylated site in both spectra. HCD dissociation results loss of tagged GlcNAc moiety and the corresponding characteristic oxonium ions. 349.2 Da, 427.2 Da for $^1\text{H}_7$ labeled (A), and 356.2 Da, 434.3 Da for $^2\text{D}_7$ labeled (B).

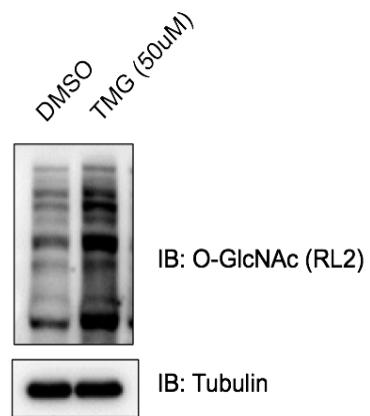


Figure S3. Global O-GlcNAc levels in 293T cells treated with or without OGA inhibitor TMG. Cells were treated with DMSO (control) or 50 μ M TMG for 12 h in the culture medium, and were collected and lysed by RIPA lysis buffer. Whole cell lysates were separated by SDS-PAGE and immunoblotted with a pan-anti-O-GlcNAc antibody (RL2).

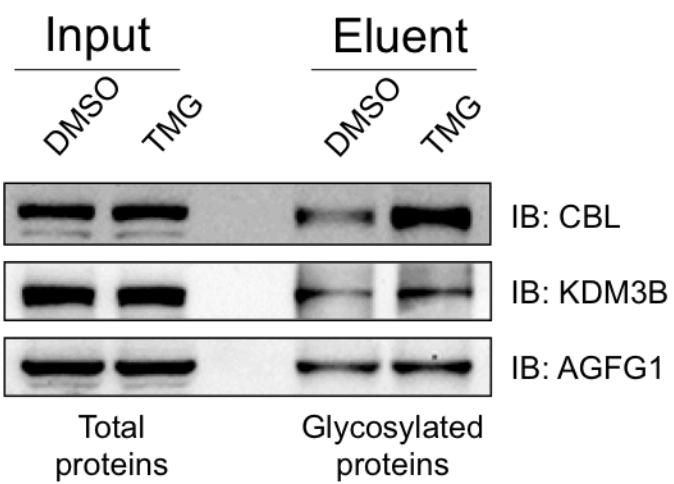


Figure S4. Immunoblotting analysis of O-GlcNAc levels in selected proteins treated with or without TMG. O-GlcNAcylated proteins from 293T cell lysates treated with or without TMG were chemoenzymatically tagged and conjugated with the probe **1**. Following affinity capturing and photocleavage, the eluents were immunoblotted with antibodies for specific proteins and the changes in O-GlcNAcylation were quantified based on the relative amounts of proteins released from the affinity beads.

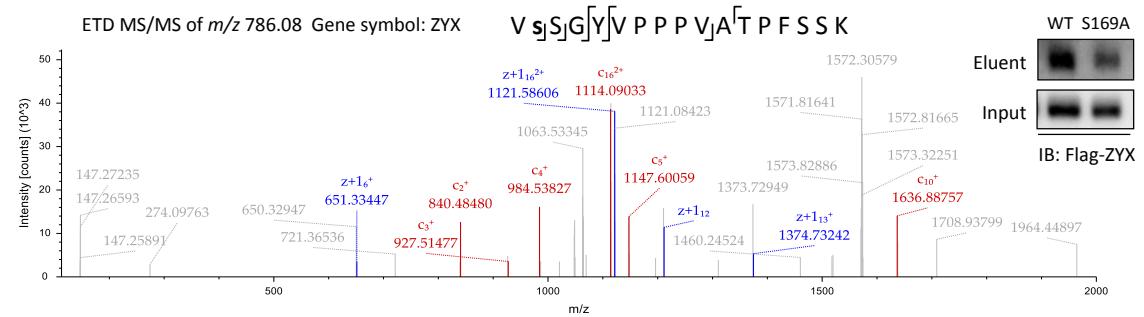
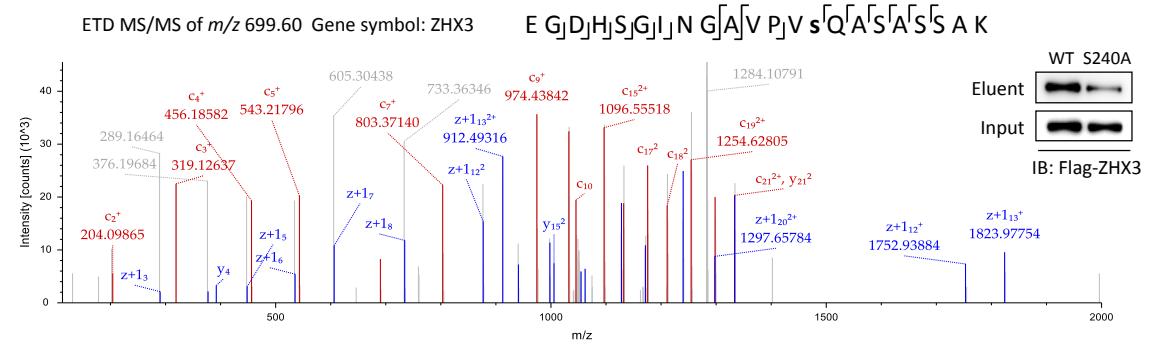
A**B**

Figure S5. O-GlcNAcylation site(s) mapping of selected proteins using ETD-based MS/MS, followed by site-directed mutagenesis and immunoblotting analysis.

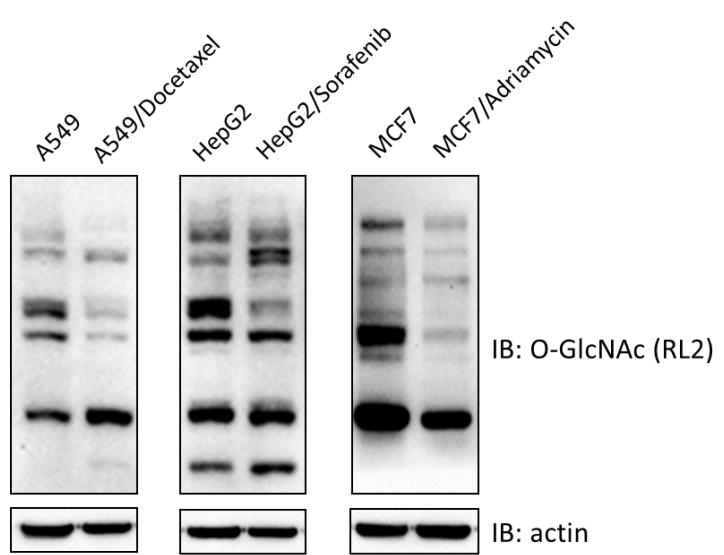


Figure S6. O-GlcNAc levels in drug -sensitive and -resistance cell lines.

Table S1. Quantified glycopeptides with CID and ETD from 293T cells cultured in the presence or absence of TMG.

- Sheet 1. Quantified glycopeptides with CID from 293T cells.
- Sheet 2. Quantified glycopeptides with ETD from 293T cells.
- Sheet 3. Glycopeptides taken more than 1.5-fold increase upon TMG treatment.
- Sheet 4. Glycopeptides taken more than 1.5-fold decrease upon TMG treatment.
- Sheet 5. Glycopeptides taken no measurable change upon TMG treatment.

Table S2. Identification of O-GlcNAcylation sites by ETD from control and TMG treated 293T cells.

- Sheet 1. Unambiguously identified O-GlcNAcylation sites from control 293T cells.
- Sheet 2. Unambiguously identified O-GlcNAcylation sites from TMG treated 293T cells.

Table S3. Quantified glycopeptides with CID and ETD from sorafenib-sensitive and -resistant HepG2 cells.

- Sheet 1. Quantified glycopeptides with CID from HepG2 cells.
- Sheet 2. Quantified glycopeptides with ETD from HepG2 cells.
- Sheet 3. Glycopeptides taken more than 1.2-fold increase in sorafenib-resistant HepG2 cells compared with its sensitive counterparts.
- Sheet 4. Glycopeptides taken more than 1.2-fold decrease in sorafenib-resistant HepG2 cells compared with its sensitive counterparts.

Table S4. Identification of O-GlcNAcylation sites by ETD from sorafenib-sensitive and -resistant HepG2 cells.