A Triarylphosphine-Trimethylpiperidine Reagent for One-step Derivatization and Enrichment of Protein Post-Translational Modifications and Identification by Mass Spectrometry

Bianbian Huo,†,‡ # Wanjun Zhang,‡ # Xinyuan Zhao,‡ Hangyan Dong,‡ Yongliang Yu,†,* Jianhua Wang,† Xiaohong Qian‡ and Weijie Qin‡,*

† Research Center for Analytical Sciences, College of Sciences, Northeastern University, Shenyang 110819, PR China
‡ State Key Laboratory of Proteomics, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Beijing Proteome Research Center, Beijing 102206, PR China

Corresponding Authors
*E-mail: aunp_dna@126.com (Weijie Qin).
*E-mail: yuyl@mail.neu.edu.cn (Yongliang Yu).

#These authors contribute equally to this work.
Materials and methods

Anti-TMT antibody resin was purchased from Thermo Fisher Scientific (Waltham, MA, USA). Bovine serum albumin (BSA), Thiamet G, Endoplasmic Reticulum and Golgi Isolation Kit were obtained from Sigma (St. Louis, MO, USA). Cytoplasmic and cytoskeletal/nuclear fractionation kit, sequencing grade porcine trypsin, PNGase F and galactose oxidase was received from Cell Signaling Technology (Boston, MA, USA), Promega (Madison, WI, USA) and New England Biolabs (Beverly, MA, USA), respectively. Plasma Membrane Protein Extraction Kit was bought from abcam (Cambridge, UK). The deionized water used in all of the experiments (resistance > 18 MΩ cm⁻¹) was prepared using a Millipore purification system (Billerica, MA, USA).

Synthesis and characterization of the triphenylphosphine functionalized trimethylpiperidine (TFT) and standard azide modified O-GlcNAc peptides

The synthesis route of TFT is shown in Figure S1, supporting information and the obtained TFT was characterized by NMR and mass spectrometry (Figure S2 a-c). Four synthetic azide-labeled O-GlcNAc (N₃-O-GlcNAc) peptides were obtained from Chinese peptide Company (Hangzhou, China). The N-acyl substituent of the GlcNAc moiety was modified with an azide group, which structurally mimics the metabolically labeled N₃-O-GlcNAc peptide obtained from cell lysates.
Synthesis and characterization of the triphenylphosphine functionalized trimethylpiperidine (TFT) and standard synthetic azide modified O-GlcNAc peptides

Figure S1. The synthesis route of triphenylphosphine functionalized trimethylpiperidine (TFT).
2,6-Dimethylpiperidine, bromoacetic acid, HOBt, DCC, EDCI, DMAP and N-Boc-ethylenediamine were all purchased from Sigma (St. Louis, MO, USA). 3-(diphenylphosphino)-4-(methoxycarbonyl) benzoic acid was obtained from Hangzhou Dayangchem (Hangzhou, China) and benzyl 3-aminopropanoate was bought from Jianchao Chemical (Shanghai, China).

A solution of 3.3 M NaOH (5 mL) was added to an ice-cooled solution of bromoacetic acid (2.0 g, 14.4 mmol) in water (8 mL) and the pH was adjusted to 14. Next, compound 1 (1 mL, 11.9 mmol) was added dropwise to the mixture and stirred in an ice bath for 1 h. The resulting solution was stirred at room temperature for 3 days. The product was extract with CH$_2$Cl$_2$ for three times. The organic phase was washed with brine and dried with anhydrous sodium sulfate. The solvent was removed under reduced pressure and the residue was purified by silica column chromatography (CH$_2$Cl$_2$/MeOH, 20/1, v/v) to give the compound 2 (1.3 g, 7.7 mmol, 65%).

Tionyl chloride (7.77 mL, 95.8 mmol) was added dropwise to an ice-cooled solution of compound 2 (3.0 g, 17.5 mmol) in dry CH$_2$Cl$_2$ (13 mL) and stirred for 2 h in ice bath. The solvent was removed under reduced pressure and then CH$_2$Cl$_2$ was added to the product and concentrated again to take away the residual tionyl chloride. The product was dissolved in 20 mL dry CH$_2$Cl$_2$ and then added dropwise with a solution of compound 3 and triethylamine (7.32 mL, 52.5 mmol) in 20 mL dry CH$_2$Cl$_2$. The resulting solution was stirred at room temperature overnight. The
product was extracted by CH₂Cl₂ for three times. The organic phase was washed with
brine and dried with anhydrous sodium sulfate. The solvent was removed under
reduced pressure and the residue was purified by silica column chromatography
(petroleum ether/ethyl acetate, 2/1, v/v) to give compound 4 (4.2 g, 12.6 mmol, 72%).

10% Pd/C (1.0 g) was added to a solution of compound 4 (7.0 g, 21.0 mmol) in
MeOH (250 mL). The resulting solution was hydrogenated under constant pressure at
room temperature for 4 h. The suspended solid was removed by filtration. The solvent
was removed under reduced pressure and the residue was purified by silica column
chromatography (CH₂Cl₂/MeOH, 15/1, v/v) to give compound 5 (4.8 g, 19.8 mmol,
95%).

HOBt (3.4 g, 25.2 mmol), DCC (10.4 g, 50.4 mmol) and triethylamine (10.7 mL,
75.6 mmol) was added to an ice-cooled solution of compound 5 (3.0 g, 12.6 mmol) in
DMF (30 mL) and stirred at 0 °C for 1 h. The resulting solution was added with a
solution of compound 6 (2.4 g, 15.1 mmol) in DMF (30 mL) and stirred at room
temperature overnight. The product was extracted with CH₂Cl₂ for three times. The
organic phase was washed with brine and dried with anhydrous sodium sulfate. The
solvent was removed under reduced pressure and the residue was purified by silica
column chromatography (petroleum ether/ethyl acetate, 2/1, v/v) to give the
compound 7 (4.1 g, 10.7 mmol, 85%).

Trifluoroacetic acid (12 mL) was added to an ice-cooled solution of compound 7
in CH₂Cl₂ (30 mL). The completion of the reaction was monitored using thin layer
cromatography (TLC) analysis. The product was washed twice with H₂O and dried
with anhydrous sodium sulfate. The solvent was removed under reduced pressure and the residue was purified by silica column chromatography (CH$_2$Cl$_2$/MeOH, 20/1, v/v) to give the compound $8$ (732 mg, 2.6 mmol, 90%).

EDCI (2.4 g, 12.4 mmol) and DMAP (200 mg, 1.1 mmol) was added to a solution of compound $9$ (3.0 g, 8.2 mmol) in DMF. After 30 min, a solution of compound $8$ (3.1 g, 10.7 mmol) in DMF was added and the resulting solution was stirred at room temperature overnight. The product was extract with CH$_2$Cl$_2$ for three times. The organic phase was washed with brine and dried with anhydrous sodium sulfate. The solvent was removed under reduced pressure and the residue was purified by silica column chromatography (petroleum ether/ethyl acetate, 1/2, v/v) to give the compound $10$ (2.6 g, 4.1 mmol, 50%).
Figure S2 (a). $^1$H NMR characterization of TFT.

8.07 (dd, $J = 8.0$, 3.5 Hz, 1H; C(O)NH), 7.95 (s, 1H; C(O)NH), 7.80 (d, $J = 7.9$ Hz, 1H; C(O)NH), 7.53 – 7.42 (m, 2H; Ar-H), 7.39 – 7.24 (m, 10H; Ar-H), 7.01 (s, 1H; Ar-H), 3.74 (s, 3H; C(O)OCH$_3$), 3.53 (dd, $J = 12.0$, 6.2 Hz, 2H; CH$_2$), 3.41 (d, $J = 16.7$ Hz, 4H; CH$_2$), 2.98 (s, 2H; CH$_2$), 2.54 – 2.27 (m, 4H; CH$_2$), 1.63 (d, $J = 12.5$ Hz, 1H; CH), 1.52 (d, $J = 12.3$ Hz, 2H; CH$_2$), 1.39 – 1.23 (m, 1H; CH), 1.22 – 1.07 (m, 2H; CH$_2$), 0.92 (d, $J = 5.9$ Hz, 6H; CH$_3$)
Figure S2 (b). $^{13}$C NMR characterization of TFT.

172.36 (C(O)), 166.90 (C(O)), 166.89 (C(O)), 166.79 (C(O)), 141.67 (Ar), 141.44 (Ar), 137.29 (Ar), 137.21 (Ar), 136.99 (Ar), 136.83 (Ar), 134.09 (Ar), 133.93 (Ar), 133.53 (Ar), 130.83 (Ar), 130.81 (Ar), 129.06 (Ar), 128.74 (Ar), 128.68 (Ar), 126.36 (Ar), 58.77 (NCH), 52.32 (NCH$_2$C(O)), 41.29 (CH$_2$), 39.68 (CH$_2$), 36.65 (CH$_2$), 35.09 (CH$_2$), 24.23 (CH$_3$), 21.39 (CH$_3$)
Figure S2 (c). Mass spectrometry characterization of TFT.

ESI-MS (m/z): calcd for C35H43N4O5P: 630.2971, found: 631.3027 [M+H]+.

The obtained compound 10 (TFT) was characterized by 1H NMR, 13C NMR and mass spectrometry.

LogP of TFT representing hydrophobicity of the reagent was calculated using the Advance Chemistry Development (ACD/Labs) Software V11.02.
Characterization of Compound 2

\[
\text{\textsuperscript{13}C NMR characterization of Compound 2} \\
\text{(126 MHz, DMSO-\textit{d}_6) } \delta \text{ 172.05, 59.04, 27.19, 22.44, 21.18, 17.12 ppm}
\]

\[
\text{\textsuperscript{1}H NMR characterization for Compound 2} \\
\text{(500 MHz, DMSO-\textit{d}_6) } \delta \text{ 3.19 (s, 2H), 1.90 (s, 2H), 1.67 – 1.34 (m, 6H), 1.11 (s, 3H), 1.10 (s, 3H) ppm.} \\
\text{ESI-MS (m/z): calcd for C_9H_{18}NO_2+: 172.1332 [M+H]^+, found: 172.1336 [M+H]^+}
\]
Characterization of Compound 4

\( ^{13} \text{C NMR characterization for Compound 4} \)

(126 MHz, CDCl\(_3\)) \( \delta \) 172.09, 135.64, 128.57, 128.27, 66.47, 58.69, 34.34, 24.17, 21.26 ppm.

\( ^{1} \text{H NMR characterization for Compound 4} \)

(500 MHz, CDCl\(_3\)) \( \delta \) 7.94 (s, 1H), 7.40 – 7.29 (m, 5H), 5.14 (s, 2H), 3.56 (q, \( J = 6.3 \) Hz, 2H), 3.03 (s, 2H), 2.59 (t, \( J = 6.1 \) Hz, 2H), 2.40 (s, 2H), 1.69 -1.53 (m, 3H), 1.38 – 1.11 (m, 3H), 0.94 (s, 3H), 0.93 (s, 3H) ppm.

ESI-MS (m/z): calcd for C\(_{19}\)H\(_{29}\)N\(_2\)O\(_3^+\): 333.2173 [M+H]\(^+\), found: 333.2179 [M+H]\(^+\)
**Characterization of Compound 5**

\[ \text{\textbf{13C NMR characterization for Compound 5}} \]

(126 MHz, CDCl\textsubscript{3}) \( \delta \) 176.36, 170.62, 58.92, 51.34, 35.15, 32.78, 23.75, 20.15 ppm.

**1H NMR characterization for Compound 5**

(500 MHz, CDCl\textsubscript{3}) \( \delta \) 8.12 (t, \( J = 5.7 \) Hz, 1H), 3.62 – 3.43 (m, 4H), 2.98 (s, 2H), 2.50 (t, \( J = 6.0 \) Hz, 2H), 1.80 – 1.33 (m, 6H), 1.14 (s, 3H), 1.12 (s, 3H) ppm.

ESI-MS (m/z): calcd for C\textsubscript{12}H\textsubscript{22}N\textsubscript{2}O\textsubscript{3}\textsuperscript{+}: 243.1703 [M+H]\textsuperscript{+}, found: 243.1704 [M+H]\textsuperscript{+}
Characterization for Compound 7

\[ \text{\textsuperscript{13}C NMR characterization for Compound 7} \]
(126 MHz, CDCl\textsubscript{3}) \( \delta \) 174.15, 171.52, 156.52, 79.43, 58.66, 40.34, 36.19, 34.94, 28.3, 24.14, 21.26 ppm.

\[ \text{\textsuperscript{1}H NMR characterization for Compound 7} \]
(500 MHz, CDCl\textsubscript{3}) \( \delta \) 7.99 (s, 1H), 6.89 (s, 1H), 5.28 (s, 1H), 3.55 (dd, \( J = 12.3, 6.3 \) Hz, 2H), 3.34 (dd, \( J = 10.8, 5.3 \) Hz, 2H), 3.25 (d, \( J = 4.7 \) Hz, 2H), 3.05 (s, 2H), 2.48 – 2.37 (m, 4H), 1.71 – 1.16 (m, 6H), 1.43 (s, 9H), 0.98 (s, 3H), 0.97 (s, 3H) ppm.

ESI-MS (m/z): calcd. for C\textsubscript{19}H\textsubscript{37}N\textsubscript{4}O\textsubscript{4}+: 385.2809 [M+H]\textsuperscript{+}, found: 358.2808 [M+H]\textsuperscript{+}
Characterization of Compound 8

$^{13}$C NMR characterization for Compound 8
(126 MHz, CD$_3$OD) $\delta$ 173.28, 165.74, 61.39, 41.92, 39.41, 36.75, 35.89, 34.76, 25.37, 22.08, 15.82 ppm.

$^1$H NMR characterization for Compound 8
(500 MHz, CD$_3$OD) $\delta$ 3.80 (s, 2H), 3.60 – 3.48 (m, 4H), 3.45 (t, $J = 5.9$ Hz, 2H), 3.07 (t, $J = 5.9$ Hz, 2H), 2.48 (t, $J = 6.7$ Hz, 2H), 2.01 – 1.44 (m, 6H), 1.24 (s, 3H), 1.22 (s, 3H) ppm.

ESI-MS (m/z): calcd for C$_{14}$H$_{29}$N$_4$O$_2$:285.2285[M+H]$^+$, found: 285.2285 [M+H]$^+$
Figure S2d. NMR and MS Characterization of compounds 2, 4, 5, 7 and 8.

**Cell Culture and protein digestion**

For metabolic based azide tagging of complex protein sample, HeLa cells were cultured in the DMEM containing 10% (v/v) fetal bovine serum and 1% (v/v) penicillin-streptomycin at 37 °C in a 5% CO\(_2\) incubator. The cells were cultured for 24 h after adding 50 μM Ac\(_4\)GlcNAz\(^1\) and 10 μM Thiamet G (final concentration). For quantitative O-GlcNAcylation analysis, the cells were treated with and without 10 Gy of ionizing radiation before harvesting the cells. Distinct cell fractions including cytoplasmic, nuclear/cytoskeletal, cell membrane, Endoplasmic Reticulum and Golgi apparatus were extracted from the obtained HeLa cells using corresponding cell fractionation kit. The cytoplasmic and nuclear/cytoskeletal fractions were combined and used for the subsequent O-GlcNAc proteins enrichment (referred to as nucleocytoplasmic protein). The proteins were digested using the FASP protocol\(^2\). Briefly, Nucleocytoplasmic proteins (2 μg/μL) were loaded onto 10 kDa ultrafiltration devices (Millipore, Danvers, MA, USA) and treated with 100 units of PNGase F in 25 mM of ammonium bicarbonate (pH=7.8) at 37 °C overnight and next with 10 units galactose oxidase & 15 units HRP in 25 mM sodium phosphate buffer (pH 7.0) containing 5% DMSO for 1 h at 35°C to remove the N-glycan and destroy the O-galactosamine modifications on the residual oligosaccharide modified N/O-glycoproteins from Endoplasmic Reticulum, Golgi apparatus and cell membrane, while sparing the O-GlcNAc modification on the nucleocytoplasmic proteins. After removal of the released N-glycans by ultrafiltration, DTT reduction and IAA alkylation was conducted. Finally, trypsin was added (with a 1:50 enzyme to protein
ratio) and further incubated overnight at 37 °C. The obtained peptides were quantified based on absorbance at 280 nm using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA), lyophilized under vacuum and stored at -80°C.

**Stable isotope dimethyl labeling, one step derivatization and enrichment of O-GlcNAc peptides by TFT**

For quantitative analysis of cellular O-GlcNAcylation under different conditions, peptides obtained from cells treated with and without ionizing radiation were labeled with isotopomeric dimethyl according to a reported protocol. In brief, the obtained peptides were dispersed in 84 μL of IP buffer (100 mM HEPES, 250 mM NaCl and 0.2 mM Na₂HPO₄, pH=7.4). Next, 8 μL of 4% CD₂O/CH₂O and 8 μL of 0.6 M NaBH₃CN was added to the ionizing radiation treated/non-treated samples for heavy/light isotopic labeling of the peptides, respectively. The dimethyl labeling reaction was carried out at room temperature for 2 hours with gentle agitation. The heavy/light labeled peptides were mixed at 1:1 ratio before enrichment. TFT reagent was first loaded to the anti-TMT antibody resin by mixing at ratio of 50 nmol TMT per mL of resin and incubated for 1 hour at RT with gentle agitation. Nest, 500 μg heavy/light labeled peptides in 100μl IP buffer was mixed 25 μl anti-TMT antibody resin and incubated for 4 hours at RT to allow Staudinger ligation between the azide tagged O-GlcNAc peptides and the triphenylphosphine of TFT reagent on the resin. After enrichment, the resin was washed with 3 × 300 μL of washing buffer containing 0.05% Tween-20 and PBS to remove excess reagents and non-specific binding.
peptides. The enriched O-GlcNAc peptides were eluted by 50 µL competitively Elution Buffer (Thermo-Fisher Scientific, USA). The obtained O-GlcNAc peptides were desalted using C_{18} Zip-Tip, freeze-dried and stored at -80°C until analysis.

**Sample recovery of the TFT-based enrichment and derivatization**

Sample recovery was evaluated using isotopic dimethyl labeled synthetic N\textsubscript{3}-O-GlcNAc peptides. 10ng “heavy” labeled N\textsubscript{3}-O-GlcNAc peptides were spiked in the tryptic digest of nucleocytoplasmic proteins from HeLa cell at 1:1:1:1:1000 ratio. The mixture was subjected to enrichment and derivatization by the TFT reagent. Next, the peptides released from the antibody resin were mixed with the same amount of the “light” labeled and TFT derivatized N\textsubscript{3}-O-GlcNAc peptides, which were used as the reference. The recovery of our method was determined by quantifying the ratio between the “light” and the “heave” peptides by MS.

**MALDI-TOF-MS, LC-MS\textsuperscript{2} analysis and data processing**

The mixture of tryptic digested BSA and the four synthetic azide-labeled O-GlcNAc (N\textsubscript{3}-O-GlcNAc) peptides with or without TFT derivatization and enrichment were redispersed in 5 µl of CHCA solution (5 mg/mL, 50% ACN, 0.1% TFA). 1 µl sample was spotted on the target plate and air-dried. MALDI-TOF-MS analysis was carried out using an ultrafleXtreme MALDI-TOF-TOF mass spectrometer (Bruker Daltonics Inc., USA). All the mass spectra (1000 laser shots for every spectrum) were acquired in positive reflection mode. LC-MS\textsuperscript{2} analysis was carried out using an EASY-nLC
1000 system coupled with an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, USA). The samples were separated using in-house made 12cm length reverse phase columns (150μm id) packed with Ultimate XB-C18 1.9μm resin (Welch materials). The constant flow rate was 600nl/min. For the MS scan, the scan range was from 300 to 1400 m/z with a resolution of 120,000. All MS² spectra were acquired in data-dependent acquisition mode. Higher-energy collision dissociation (HCD) with normalized collision energy of 35% was used for O-GlcNAc peptide identification. Electron-transfer and higher-energy collision dissociation (EThcD) with ETD reaction time 100 ms and HCD AS Collision energy = 35% was used for O-GlcNAc site assignment.

Data Processing

Raw files of tandem mass spectra obtained were processed against the uniprot-all-human database (20207-2015.7.21) using Maxquant. Trypsin was selected as the proteolytic enzyme allowing for a maximum of two missed cleavages. The mass tolerance of the precursor ion was set to 15 ppm and that of the fragment ions was set to 20 mmu. The false discovery rate was set ≤ 1% at spectra level and protein level. TFT-N₃-O-GlcNAc (mass=832.3560), N-terminal acylation and methionine oxidation were set as the variable modifications. Maximum of three common modifications and two rare modifications were allowed per peptide. Min. score and min. delta score for modified peptides was set ≥ 40 and ≥ 8 in Maxquant⁴. To screen the spectra for presence of the diagnostic fragmentation ions (DFIs) of TFT-N₃-O-GlcNAc, raw files
were converted to MGF files by ProteoWizard 3.0.7162 and next processed by spectrum-picking and optimization tools written using the Perl language. At least two DFI$s are required for confident identification of an O-GlcNAc peptide.

Quantitative comparison was conducted on the isotopic dimethyl labeled O-GlcNAc peptides. Quantitative O-GlcNAc peptides data from Maxquant were $\log_{10}$ transformed and missing values were imputed using random values generated from a normal distribution centered on the 1% quantile and the median SD of all identified peptides\textsuperscript{5}. T-test was applied to the O-GlcNAc peptides obtained from the ionizing radiation treated and non-treated cells to find the significantly changed ones. Two criteria were applied for differential O-GlcNAc peptides screening. (1) Quantitative ratios of the O-GlcNAc peptides from the ionizing radiation treated and non-treated cells $\geq 2$ or $\leq 0.5$ and (2) P value of t-test $\leq 0.01$. 
Figure S3. MALDI-TOF-MS characterization of the TFT tagging efficiency of the four synthetic \( N_3 \)-O-GlcNAc peptides. 1. Residual untagged \( N_3 \)-O-GlcNAc peptides. 2. TFT tagged \( N_3 \)-O-GlcNAc peptides.
<table>
<thead>
<tr>
<th>Peptides</th>
<th>Residual untagged peptides</th>
<th>TFT tagged peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S/N</td>
<td>Signal intensity</td>
</tr>
<tr>
<td>Peptide 1</td>
<td>N/A</td>
<td>18</td>
</tr>
<tr>
<td>Peptide 2</td>
<td>N/A</td>
<td>14</td>
</tr>
<tr>
<td>Peptide 3</td>
<td>N/A</td>
<td>4</td>
</tr>
<tr>
<td>Peptide 4</td>
<td>N/A</td>
<td>4</td>
</tr>
</tbody>
</table>

Table S1. Signal intensity and S/N of the TFT tagged N$_3$-O-GlcNAc peptides and the corresponding residual untagged N$_3$-O-GlcNAc peptides in MALDI-TOF-MS analysis.
Figure S4. Total ion current chromatography of the four N$_3$-O-GlcNAc peptides before (a) and after TFT tagging (b).
Peptide 1. LNPAVT[N$_3$-O-GlcNAc]CAGK,
Peptide 2. RQLFVT[N$_3$-O-GlcNAc]VVK
Peptide 3. AQPVQS[N$_3$-O-GlcNAc]KPQK
Peptide 4. AAAPAPVS[N$_3$-O-GlcNAc]EAVCR
Peptide 1: LNPAVT[N_3-O-GlcNAc]CAGK before (a) and after (b) TFT tagging.
Peptide 2: RQLFVT[N\textsubscript{3}-O-GlcNAc]VVK before (c) and after (d) TFT tagging.
Peptide 3: AQPVQS[N$_3$-O-GlcNAc]KPQK before (e) and after (f) TFT tagging.
Peptide 4: AAAPAPVS[N$_3$-O-GlcNAc]EAVCR before (g) and after (h) TFT tagging.

Figure S5. Charge states of the synthetic N$_3$-O-GlcNAc peptides before and after TFT tagging in ESI-MS. Relative Abundance in the y-axis of the figure refers to the intensity of the corresponding peaks relative to the strongest one in the figure.
HCD fragmentation spectra of Peptide 1 (LNPAVT[N\text{\textsubscript{3}}-O-GlcNAc]CAGK) before (a) and after (b) TFT tagging.
HCD fragmentation spectra of Peptide 2 (RQLFVT[N$_3$-O-GlcNAc]VVK) before (c) and (d) after TFT tagging.
HCD fragmentation spectra of Peptide 3 (AQPVQS[N3-O-GlcNAc]KPQK) before (e) and after (f) TFT tagging.
HCD fragmentation spectra of Peptide 4 (AAAPAPVS$[N_\gamma-O-GlcNAc]$EAVCR) before (g) and after (h) TFT tagging.
EThcD fragmentation spectra of LNPAVT[N3-O-GlcNAc]CAGK before (i) and after (j) TFT tagging.

(k)
EThcD fragmentation spectra of RQLFVT[N$_3$-O-GlcNAc]VVK before (k) and after (l) TFT tagging.
EThcD fragmentation spectra of AQPVQS[N\textsubscript{2}-O-GlcNAc]KPQK before (m) and after (n) TFT tagging.
EThcD fragmentation spectra of AAAPAPVS[N$_3$-O-GlcNAc]EAVCR before (o) and after (p) TFT tagging.

Figure S6. HCD and EThcD fragmentation spectra of the N$_3$-O-GlcNAc peptides before and after TFT tagging.
Table S2. Characteristic diagnostic fragmentation ions (DFI) of the TFT-tagged N$_3$-O-GlcNAc peptides found in the MS$^2$ spectra.

<table>
<thead>
<tr>
<th>Dignostic fragmentation ions (DFI)</th>
<th>Molecular formula</th>
<th>m/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFI-1</td>
<td>C34H41N4O5P$^{2+}$</td>
<td>308.14(Z=2)</td>
</tr>
<tr>
<td></td>
<td>C34H40N4O5P$^+$</td>
<td>615.27(Z=1)</td>
</tr>
<tr>
<td>DFI-2</td>
<td>C36H44N5O6P$^{2+}$</td>
<td>366.66(Z=2)</td>
</tr>
<tr>
<td></td>
<td>C36H43N5O6P$^+$</td>
<td>672.30(Z=1)</td>
</tr>
<tr>
<td>DFI-3</td>
<td>C42H54N6O10P$^+$</td>
<td>833.36(Z=1)</td>
</tr>
<tr>
<td>EThcD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFI-4</td>
<td>C33H42N4O4P$^+$</td>
<td>589.29(Z=1)</td>
</tr>
<tr>
<td>DFI-5</td>
<td>C34H41N4O5P$^+$</td>
<td>616.28(Z=1)</td>
</tr>
</tbody>
</table>
Table S3. Intensity and S/N in MALDI-TOF-MS analysis of the four N$_3$-O-GlcNAc peptides in mixture with tryptic digested BSA at a 1 : 100 (w/w) before (a) and after (b) derivatization and enrichment by TFT tagging.

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Before derivatization and enrichment</th>
<th>After derivatization and enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S/N</td>
<td>Signal intensity</td>
</tr>
<tr>
<td>Peptide 1</td>
<td>N/A</td>
<td>45</td>
</tr>
<tr>
<td>Peptide 2</td>
<td>N/A</td>
<td>41</td>
</tr>
<tr>
<td>Peptide 3</td>
<td>N/A</td>
<td>62</td>
</tr>
<tr>
<td>Peptide 4</td>
<td>N/A</td>
<td>49</td>
</tr>
<tr>
<td>Peptides</td>
<td>Sequence</td>
<td>Score</td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
<td>-------</td>
</tr>
<tr>
<td>Peptide 1</td>
<td>LNPAVTCAGK</td>
<td>58.595</td>
</tr>
<tr>
<td>Peptide 2</td>
<td>RQLFVTVVK</td>
<td>118.87</td>
</tr>
<tr>
<td>Peptide 3</td>
<td>AQPVQSKPQK</td>
<td>135.43</td>
</tr>
<tr>
<td>Peptide 4</td>
<td>AAAPAPVSEAVCR</td>
<td>87.647</td>
</tr>
</tbody>
</table>

Table S4. Identification results of the four N$_3$-O-GlcNAc peptides in the mixture with the tryptic digested peptides of the nucleocytoplasmic protein extracts from Hela cells (ratio=1:1000, w/w) obtained by the search engine Maxquant.
<table>
<thead>
<tr>
<th>Peptide</th>
<th>Recovery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>81.0</td>
</tr>
<tr>
<td>2</td>
<td>89.5</td>
</tr>
<tr>
<td>3</td>
<td>87.4</td>
</tr>
<tr>
<td>4</td>
<td>82.4</td>
</tr>
</tbody>
</table>

Table R5. Recovery of the four synthetic N-GlcNAc peptides using tryptic digested nucleocytoplasmic proteins of HeLa cell as the background (N-GlcNAc peptide 1/2/3/4: tryptic digested nucleocytoplasmic proteins = 1:1:1:1000, w/w).

Four “heavy” labeled synthetic N-GlcNAc peptides were spiked in tryptic digest of nucleocytoplasmic proteins from HeLa cell at 1:1:1:1000 (w/w) ratio. The mixture was subjected to enrichment and derivatization by the TFT reagent. Next, the peptides released from the antibody resin were mixed with the same amount of “light” labeled and TFT derivatized synthetic N-GlcNAc peptides, which were used as the reference. The recovery of our method was determined by quantifying the ratio between the “heave” and the “light” peptides using ESI-MS analysis.
Figure S7. Typical HCD fragmentation spectra (a, b, c) and EtHCD fragmentation spectra (d, e, f). og stands for O-GlcNAc.
Reference


