Development of Trifunctional Probes for Glycoproteomic Analysis

Charng-Sheng Tsai,[a] Po-Yu Liu,[a,b] Hsin-Yung Yen,[a,b] Tsui-Ling Hsu,[a] Chi-Huey Wong*[a,c]

[a]Genomics Research Center, Academia Sinica, 128 Academia Road Section 2, Nankang, Taipei 115, Taiwan.

[b]Institute of Biochemical Science, National Taiwan University, 1 Roosevelt Road Section 4, Taipei 106, Taiwan

[c]Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA.

chwong@gate.sinica.edu.tw

-Supporting Information-

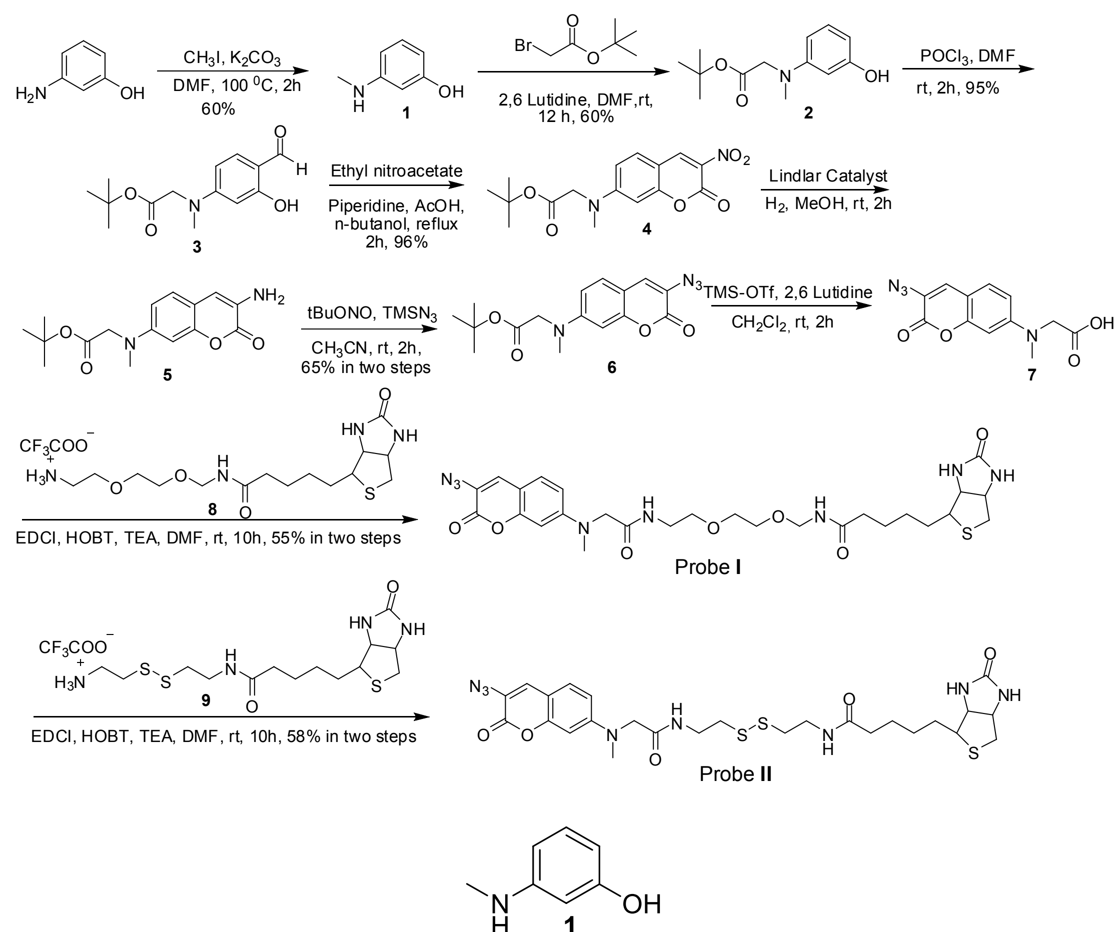
Synthetic procedures and spectroscopic data for probe I and probe II, the experimental procedures of CuAAC, Labeling ManNAcyne-tagged glycoproteomes with trifunctional probes, Affinity enrichment of probe II-labeled sialylated glycoproteins, Silver staining, 2D protein gel electrophoresis and fluorescence imaging as well as the table of identified proteins from 2D gel.

General Reagents and Methods:

Unless otherwise noted, all compounds and reagents were purchased from Acros or Sigma-Aldrich. All chemicals were purchased as reagent grade and used without further purification. Reactions were monitored with analytical thin layer chromatography (TLC) in EM silica gel 60 F254 plates and visualized under UV (254 nm) and/or staining with acidic ceric ammonium molybdate or ninhydrin. Flash column chromatography was performed on silica gel 60 Geduran (35-75 μm, EM Science). 1H NMR spectra were recorded on a Bruker DRX-600 (600MHz) spectrometer at 20 °C. Chemical shifts (J ppm) were assigned according to the CDCl3 (δ = 7.24 ppm). ^{13}C NMR spectra were obtained using Attached Proton Test (APT) on a Bruker DRX-600 (150MHz) spectrometer and were reported in J ppm scale using the signal of CDCl3 (δ= 77.00 ppm) for calibration. Fluorescence spectra were obtained on a Hitachi F-4500 spectrometer. Protease inhibitors were purchased from Roche Applied Sciences, NeutrAvidin agarose resins were from Thermo Scientific. Amicon Ultra-4 P10 and Immobilon PVDF membrane were from Millipore. NuPAGE® Bis-Tris Mini gels (4-12%), PBS and cell culture media and reagents were from Invitrogen. Rehydration solution and IPG buffer were from GE health-care.

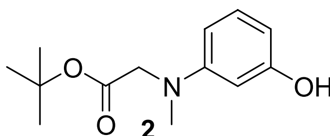
Mineral oil and pH 3-10 IPG strips were from ProteomIQ. Protein concentration was measured by either BCA protein assay (Thermo Scientific) or Bradford assay (Bio-rad). Flamingo staining reagent was purchased from Bio-rad. Chemiluminescence on protein blots was visualized and quantified using FUJI LAS3000 imaging system (Fujifilm). Fluorescence in protein gel was detected using Alpha Innotech FluoroChem SP with 419 nm excitation and an emission filter 490±25nm (*Omega*).

Synthetic Scheme of Probe I and Probe II

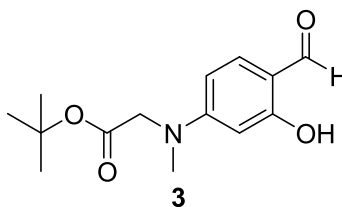


A mixture of 3-aminophenol (4.0 g, 37 mmol) and methyl iodide (5.1 g, 37 mmol) was heated in DMF (20 mL) in the presence of potassium carbonate (5.0 g, 37 mmol) at 100℃ for 2 h. The solvent was removed *in vacuo* and the resulting residue was purified using flash chromatography (Hex/EtOAc = 4/1). Compound **1** was isolated as a white solid (4.3 g, 16.5 mmol, 60% yield). TLC (Hex/EtOAc = 2/1) $R_f = 0.31$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ 7.01 (t, J = 8.0 Hz, 1 H), 6.22-6.16 (m, 2 H), 6.08 (t, J = 2.3 Hz, 1 H), 2.75 (s, 3 H); ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 156.7, 150.7, 130.1, 105.8, 104.8, 99.8, 30.8; IR (neat) : 3332, 1614, 1557, 1498, 11188, 1161, 833, 765 cm^{-1};

FAB-MS : m/z 123.1 (M + 1); HR-FAB MS calcd for C_7H_9NO $(M + 1)^+$: 123.0681; found : 123.0684.

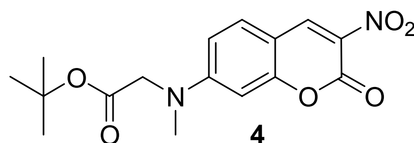


A solution of **1** (3.4 g, 27.6 mmol), 2,6-lutidine (15 mL), *tert*-butyl bromoacetate (3.93 mL, 30.93 mmol), and DMF (60 mL) was stirred at ambient temperature for 20 h. The solvent was removed *in vacuo* and the resulting residue was purified using flash chromatography (Hex/EtOAc = 2/1). Compound **2** was isolated as a white solid (4.3 g, 16.5 mmol, 60% yield). mp = 98-100 °C; TLC (Hex/EtOAc = 2/1) $R_f = 0.30$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ 7.04 (t, J = 8.1 Hz, 1 H), 6.22 (dd, J = 2.3, 8.1 Hz, 1 H), 6.18 (dd, J = 2.3, 8.1 Hz, 1 H), 6.14 (t, J = 2.3 Hz, 1 H), 5.29 (s,1H), 3.91 (s, 2H), 2.99 (s, 3 H), 1.40 (s, 9 H); ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 170.5, 156.7, 150.5,130.0, 104.9, 104.2, 99.5, 81.7, 55.38, 39.6, 28.0; IR (neat) : 3395, 1732, 1615, 1581, 1504, 1368, 1238, 1154, 1119 cm^{-1}; FAB-MS : m/z 237.1 (M + 1); HR-FAB MS calcd for $C_{13}H_{19}NO_3$ $(M + 1)^+$: 237.1365; found : 237.1369.

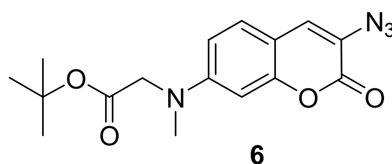


A solution of DMF (3.3 mL, 45.3 mmol) and phosphorous oxychloride (1.6 mL, 18.1 mmol) were stirred at ambient temperature for 2 h. Compound **2** (4.3 g, 16.5 mmol) in DMF (10 ml) was added via cannula. The reaction was stirred at ambient temperature for 10 h. The mixture was slowly poured over ice, and the aqueous layer was extracted with CH_2Cl_2 (3 x 50 mL). The organic layer was dried over $MgSO_4$ and the solvent was removed *in vacuo*. The resulting residue was purified by flash chromatography (Hex/EtOAc = 3/1) and compound **3** was isolated as a white solid (4.2 g, 15.6 mmol, 95% yield). mp = 78-80 °C; TLC (Hex/EtOAc = 2/1) $R_f = 0.45$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ 11.49 (s, 1 H), 9.49 (s, 1 H), 7.25 (d, J = 8.8 Hz, 1 H), 6.20 (dd, J = 2.3, 8.8 Hz, 1 H), 6.03 (d, J = 2.3 Hz, 1 H), 3.96 (s, 2 H), 3.06 (s, 3 H), 1.40 (s , 9 H).; ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 192.7, 168.5, 163.9, 155.5, 135.2,

112.3, 104.6, 97.7, 82.2, 54.7, 39.7, 27.8; IR (neat) : 1740, 1635, 1524, 1341, 1227, 1150, 1111 cm^{-1}; FAB-MS : m/z 266.1 (M + 1); HR-FAB MS calcd for $C_{14}H_{20}NO_4$ $(M + 1)^+$: 266.1392; found : 266.1397.

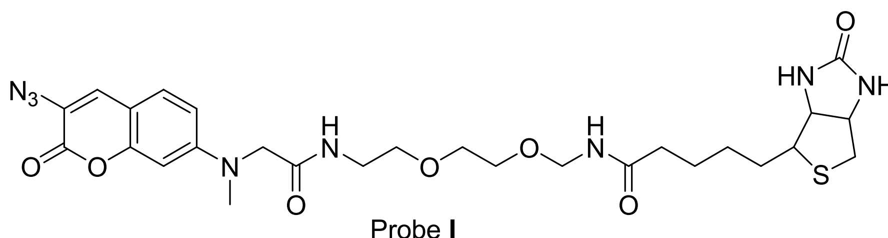


A mixture containing n-butanol (40 mL), compound **3** (3.5 g, 13.2 mmol), ethyl nitroacetate (1.6 mL, 13.2 mmol), molecular sieves 4 Å (300 mg), piperidine (0.3 mL) and acetic acid (0.6 mL) was refluxed for a period of 2 h. Upon cooling to room temperature, a bright yellow solid formed, which was collected and dissolved in DMF (20 mL) at 80 °C. It was filtered again to remove the molecular sieves. The filtrate, upon addition to 100 ml of ice-cold water, yielded compound **4** as a bright yellow solid (4.3 g, 12.7 mmol, 96 %). mp = 164-166 °C; TLC (Hex/EtOAc = 2/1) $R_f = 0.33$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ 8.68(s, 1 H). 7.45 (d, J = 9.0 Hz, 1 H), 6.67 (dd, J = 2.1, 9.0 Hz, 1 H), 6.43 (d, J = 2.0 Hz, 1 H), 4.07 (s, 2 H), 3.18 (s, 3 H), 1.43 (s, 9 H).; ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 167.59, 158.1, 155.8, 153.1, 143.5, 132.3, 128.1, 111.2, 106.9, 97.6, 83.1, 54.9, 40.2, 28.0; IR (neat) : 1737, 1730, 1620, 1588, 1502, 1320, 1275, 1220, 1151, 1112 cm^{-1}; FAB-MS : m/z 335.1 (M + 1); HR-FAB MS calcd for $C_{16}H_{19}N_2O_6$ $(M + 1)^+$: 335.1243; found : 335.1245.



To a solution of compound **4** (1.2 g, 3.60 mmol) in 50 mL of MeOH was added a spatula of Lindlar catalyst. The system was evacuated and filled with H_2 three times. It was kept under H_2 atmosphere with a balloon and stirred at rt for 2 h. The Pd/C catalyst was filtered off through Celite 535, and the filtrate concentrated. The amino intermediate was directly for next step without any purification. The amino intermediate was dissolved in CH_3CN (10 mL) and cooled to 0℃ in an ice bath.To this stirred mixture was added t-BuONO (742.0 mg, 7.20 mmol, 873 μL) followed by $TMSN_3$ (830.0 mg, 7.20 mmol, 943.0 μL) dropwise. The resulting solution was stirred

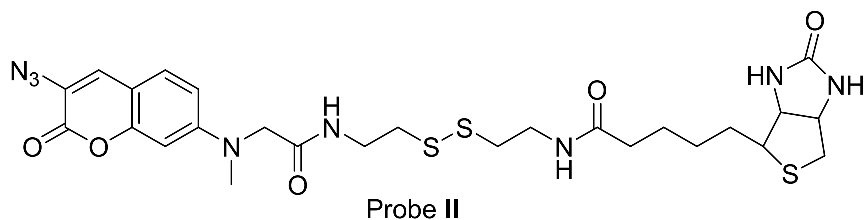
at room temperature for 2 h. The reaction mixture was concentrated under vacuum and the crude proguct was purified by flash chromatography (Hex/EtOAc = 3/1) and compound **6** was isolated as a yellow solid (772.0 mg, 2.34 mmol, 65% yield). mp = decompose; TLC (Hex/EtOAc = 2/1) $R_f = 0.45$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ 7.19 (d, J = 8.8 Hz, 1 H), 7.08 (s, 1 H), 6.56 (dd, J = 2.5, 8.8 Hz, 1 H), 6.47 (d, J= 2.5 Hz, 1 H), 3.98 (s , 2 H), 3.09 (s, 3 H), 1.41 (s, 9 H).; ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 168.9, 158.2, 153.5, 151.1, 128.0, 127.2, 120.8, 109.8, 109.3, 98.3, 82.3, 54.9, 39.8, 28.0; IR (neat) : 2121, 1732, 1621, 1522, 1392, 1368, 1330, 1224, 1153, 1112. cm^{-1}; FAB-MS : m/z 331.1 (M + 1); HR-FAB MS calcd for $C_{16}H_{19}N_4O_4$ $(M + 1)^+$: 331.1406; found : 331.1403.



Probe **I**

To a solution of compound **6** (140.0 mg, 0.42 mmol) 2,6-lutidine (494 μl, 4.2 mmol) and TMS-OTf (385 μl, 2.1 mmol) in 4 mL of CH_2Cl_2 were stirred at ambient temperature for 2 h. The solvent was removed *in vacuo* and the resulting residue was purified using flash chromatography (CH_2Cl_2/MeOH, 9:1). The acid intermediate was isolated as a brown solid. To a solution of the acid intermediate and the TFA salt of dioxaoctanediamine derivative of biotin **8** (207.0 mg, 0.42 mmol) in 4 mL of DMF was added EDCI (98.0 mg, 0.51 mmol), HOBT (13.5 mmol, 0.1 mmol), and TEA (1 mL). The reaction mixture was stirred at rt for 10 h. The solvent was removed *in vacuo*. The resulting residue was purified by flash chromatography (CH_2Cl_2/MeOH, 95:5) and the biotinylated product probe **1** was isolated as brown oil (142 mg, 0.23 mmol, 55% yield) TLC (CH_2Cl_2/MeOH = 8:2) $R_f = 0.40$. ^{1}H NMR : ($CDCl_3$, 600 MHz) δ δ 7.21(t, J = 5.8 Hz, 1 H), 7.18 (d, J = 8.8 Hz, 1 H), 7.06 (s, 1 H), 6.81 (t, J = 5.6 Hz, 1 H), 6.62 (s, 1 H), 6.58 (dd, J = 2.4, 8.8 Hz, 1 H), 6.49 (d, J = 2.4 Hz, 1 H), 5.82 (s, 1 H), 4.42 (m, 1 H), 4.22 (m, 1 H), 3.98 (s, 2 H), 3.56-3.28 (m, 12 H), 3.08 (s, 3 H), 2.82 (dd, J = 4.8, 12.8 Hz, 1 H), 2.67 (d, J = 12.8 Hz, 2 H), 1.70-1.50 (m,4 H), 1.42-1.29 (m, 2 H).; ^{13}C-NMR ($CDCl_3$, 150 MHz) δ 173.5, 169.3, 164.2, 158.0, 153.3, 151.2, 128.1, 127.1, 120.8, 110.2, 109.6, 98.6, 70.0, 69.8, 69.7, 69.6, 61.7, 60.1, 56.8, 55.6, 40.4, 39.9, 39.1, 39.0, 35.8, 28.1, 28.0, 25.5; IR (neat) : 3356, 2126, 1697, 1621, 1457, 1257, 1148, 1118. cm^{-1}; HR-ESI MS calcd for $C_{28}H_{38}N_8O_7SNa$ $(M + Na)^+$:

653.2482; found : 653.2475.



Probe **II**

A solution of compound **6** (150.0 mg, 0.45 mmol) 2,6-lutidine (530 μl, 4.5 mmol) and TMS-OTf (412 μl, 2.3 mmol) in 4 mL of CH_2Cl_2 was stirred at ambient temperature for 2 h. The solvent was removed *in vacuo* and the resulting residue was purified using flash chromatography (CH_2Cl_2/MeOH, 9:1). The acid intermediate was isolated as a brown solid. To a solution of the acid intermediate and the TFA salt of cystamine derivative of biotin **9** (221.0 mg, 0.45 mmol) in 4 mL of DMF was added EDCI (105.0 mg, 0.55 mmol), HOBT (13.5 mmol, 0.1 mmol) and TEA (1 mL). The reaction mixture was stirred at rt for 12 h. The solvent was removed *in vacuo*. The resulting residue was purified by flash chromatography (CH_2Cl_2/MeOH, 95:5) and probe **II** was isolated as brown oil (153 mg, 0.24 mmol, 58% yield) TLC (CH_2Cl_2/MeOH = 9:1) R_f = 0.35. ^{1}H NMR : (*d*-DMSO, 600 MHz) δ 8.20 (t, J = 5.5 Hz, 1 H), 7.96 (t, J = 5.3 Hz, 1 H), 7.51 (s, 1 H), 7.40 (d, J = 8.8 Hz, 1 H), 6.67 (d, J = 8.8 Hz, 1 H), 6.57 (s, 1 H), 6.43 (s, 1 H), 6.37 (s, 1 H), 4.29 (t, J = 6.6 Hz, 1 H), 4.12 (m, 2 H), 4.04 (s, 2 H), 3.30 (m, 4 H), 3.16 (s, 3 H), 3.06 (m, 4 H), 2.77 (m, 5 H), 2.55 (d, J = 12.8 Hz, 1 H), 2.05 (t, J = 7.2 Hz, 2 H), 1.65-1.40 (m, 4 H), 1.38-1.30 (m, 2 H).; ^{13}C-NMR (*d*-DMSO, 150 MHz) δ 172.3, 171.4, 169.0, 162.8, 157.7, 153.3, 151.5, 128.5, 128.4, 118.9, 110.1, 108.5, 97.6, 61.1, 59.2, 55.5, 55.1, 48.7, 40.1, 38.0, 37.9, 37.3, 37.2, 35.2, 28.2, 28.1, 25.3; IR (neat) : 212, 1710, 1625, 1449, 1250, 1141, 1112. cm^{-1}; HR-ESI MS calcd for $C_{26}H_{34}N_8O_5S_3Na$ $(M + 1)^+$: 657.1712; found : 657.1716.

The experimental procedures of CuAAC

Typically the CuAAC reaction was performed in a volume of 200 μL with 1 mM probe, 1 mM alkynyl group, 16 mM $CuSO_4$, 32 mM tris(triazolyl)amine and 16 mM sodium ascorbate in DMSO/H_2O (1:1). The reactions could be monitored by long wavelength UV irradiation (Figure 1).

Labeling ManNAcyne-tagged glycoproteomes with trifunctional probes

Prostate cancer PC3 cells (2x10^6/12 ml; from ATCC) were seeded in T75 cell culture flasks and cultured in RPMI 1640 supplemented with 10% FCS and 25 μM

ManNAcyne or control ManNAc, at 37°C for 2 days. Cells were scraped, spun down and lysed with 0.5 mL lysis buffer (1% NP-40, 150 mM NaCl, Roche protease inhibitor, and 100 mM sodium phosphate pH 7.5). After removing insolubles in cell lysates, the extracts were measured for protein concentration. Cell lysates extracted from ManNAcyne- or control ManNAc-treated PC3 cells (0.5 mL, 1-2 mg/mL) were reacted with 1 mM probe I or probe II in the presence of 100 μM tris(triazolyl)amine, 1 mM $CuSO_4$, and 2 mM sodium ascorbate at room temperature for 1 h. Free probes and excess reagents in the click mixtures were removed by Amicon Ultra-4 (10 KD cut-off) centrifugal filters. The centrifuged protein samples were then diluted to 4 ml 2D rehydration solution (2% CHAPS, 3M urea, 1M thiourea in PBS) and proceeded to another round of centrifugation. The procedure was repeated twice for buffer exchange. The concentrated protein samples were further diluted in 4 ml 2D rehydration solution.

Affinity enrichment of probe II-labeled sialylated glycoproteins

Protein samples in 4 ml 2D rehydration solution (2%CHAPS, 3M urea, 1M thiourea in PBS) was incubated with NeutrAvidin agarose beads (800 ⌠l for 1.5 mg protein samples) at 4 ºC for 3 h. Following incubation, the beads were spun down and washed once with 2D rehydration solution (0.2%CHAPS, 0.3 M urea, 0.1 M thiourea in PBS) and three times with PBS buffer. After washing, the proteins that bound to NeutrAvidin agarose beads were eluted by incubating in 500 ⌠l of 50 mM TCEP at 37ºC for 1h. Eluted proteins were resolved by SDS-PAGE or 2D protein gel electrophoresis and analyzed for coumarin fluorescence. The protein gels were then subjected to Coomassie blue stain, silver stain or flamingo stain for total proteins visualization.

Silver staining

Protein gels were fixed in an aqueous solution of 50% methanol/12% acetic acid/0.05% formalin for overnight and then washed three times with 20% ethanol (20 min incubation/time). The gels were then sensitized in 0.02 $Na_2S_2O_3$ for 2 min and then washed three times in H_2O (5 min incubation/time). The gels were stained in 0.2% $AgNO_3$/0.076% formalin for 20 min, followed by washing twice with H_2O (1 min incubation/time). The signals were then developed in 6% Na_2CO_3/0.05% formalin/0.0004% $Na_2S_2O_3$ until protein bands appeared, and the reaction was stopped by placing the gels in 50% methanol/12% acetic acid for 5 min. The gels were then preserved in 1% acetic acid before the images were taken.

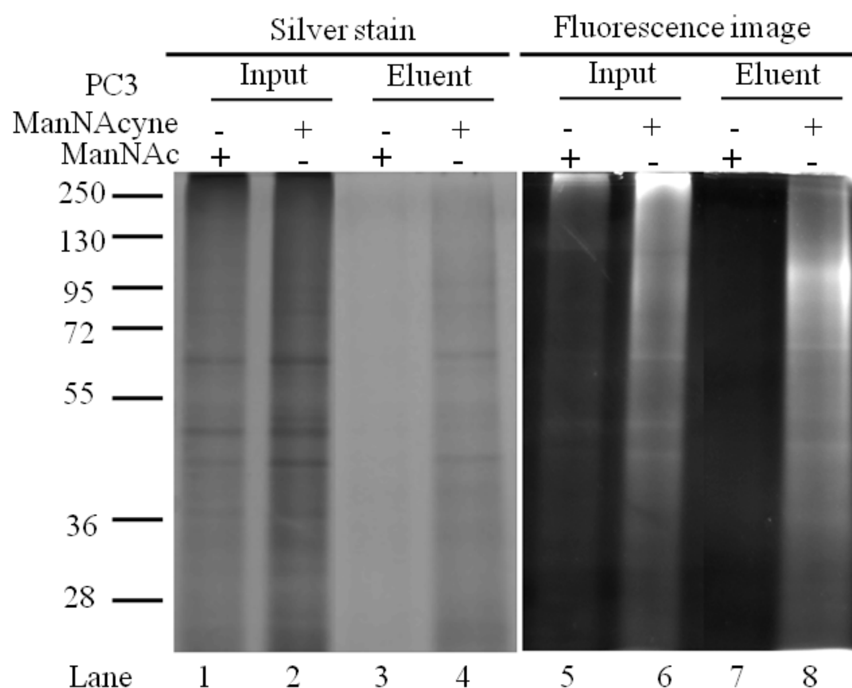


Figure S1. Fluorescence gel imaging of enriched sialylated proteins labeled by probe **II**. Proteins that captured from 100 μg of ManNAcyne- or control ManNAc-treated cell extracts were eluted and examined for fluorescence of coumarin (Lanes 7-8; Ex: 419 nm and Em: 490 nm) or by protein silver staining (Lanes 3-4). Total protein extracts (10 μg/lane) were loaded for comparison (Input; lanes 1-2 by silver stain, and lanes 5-6 by detecting coumarin)

2D protein gel electrophoresis and fluorescence imaging

Precipitated eluents were resuspended in 7 M urea 2 M thiourea 4% CHAPS 2% IPG buffer (GE health), DTT 40 mM and 1% bromophenol blue. The samples were centrifuged at 10,000 rpm for 1 min, the supernatant was loaded onto pH 3-10 strips, and the sample was rehydrated overnight. The strips has gradient for 1hr from 100V to 500V, and 1hr from 500V to 1000V, and holding at 10000V for reaching 45000Vhr, after which they were incubated in 1x SDS equilibration buffer plus DTT 10mg/ml for 5min, and then reacted with IAA 25 mg/ml 5min. The strip reacted with IAA must be packaged by aluminum foil to avoid the light. Then, the two-dimension is by SDS-PAGE. The strips were resolved on a SDS10% gel for 150V for 4.5 hours. The fluorescence signal on gel was imaged using *Alpha Innotech FluoroChem SP* with UV lamp 419 nm excitation and a filter 490±25 nm(Omega).

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2010

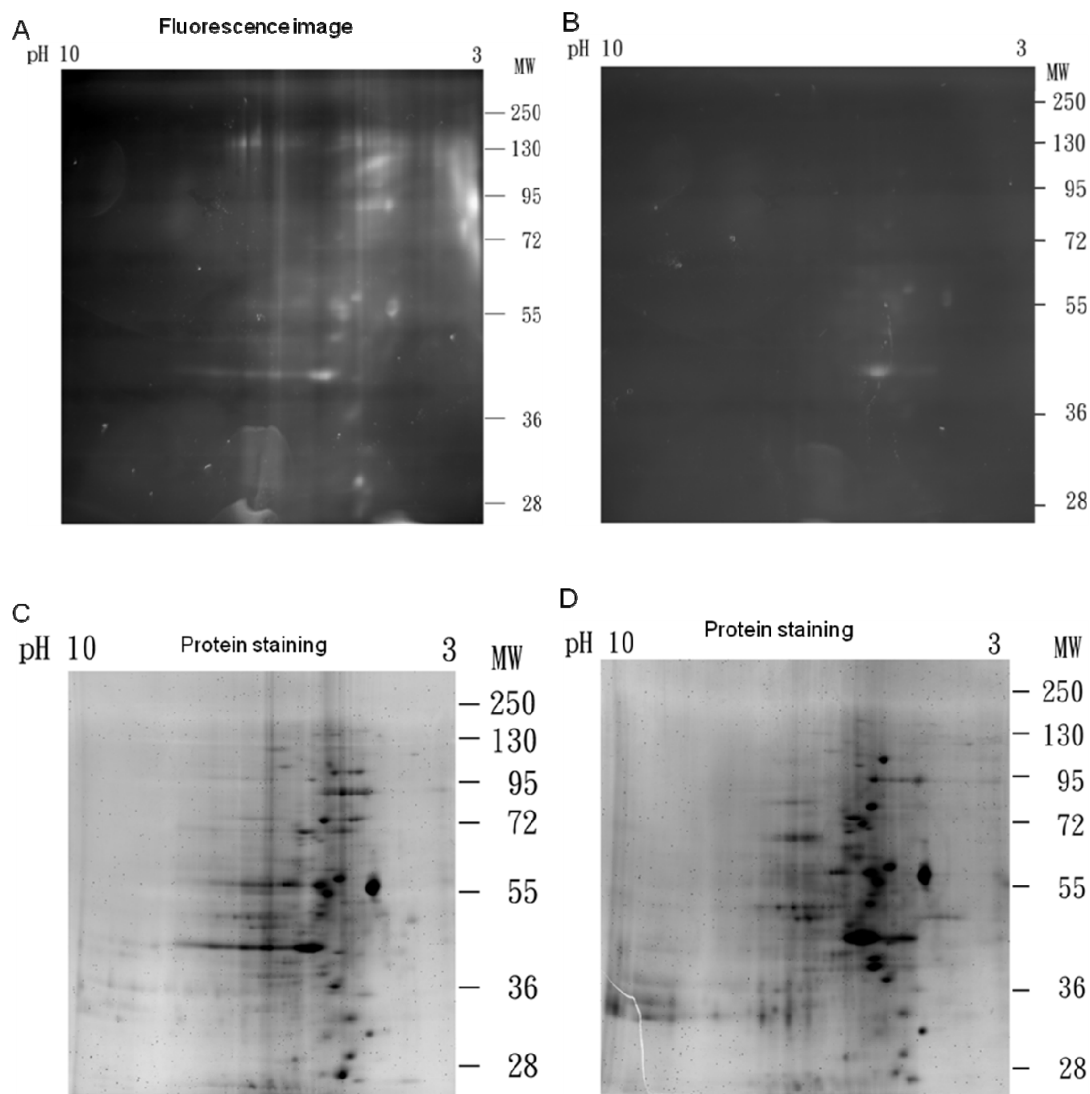


Figure S2. Probe **I** clicked with PC3 cell lysate was resolved in 2D electrophresis (A) fluorescence image : PC3 cell lysate treated with ManNAcyne then clicked with probe **I** (B) fluorescence image : PC3 cell lysate clicked with probe **I** (C) protein staining: PC3 cell lysate treated with ManNAcyne then clicked with probe **I** (D) protein staining: PC3 cell lysate clicked with probe **I**

Spot no.	Protein name	IPI number	Protein score	Coverage rate (%)	*M*r (kDa)/p*I*	
1	Calereticulin	IPI00020599	204	21	48.10/ 4.29	LKEEEEDKK DKGLQTSQDAR FYALSASFEPFSNK IDNSQVESGSLEDDWDFLPPKK LFPNSLDQTDMHGDSEYNIMFGPDICGPGTK
2	Protein disulfide-isomerase	IPI00010796	398	21	57.08/ 4.76	LKAEGSEIR SNFAEALAAHK FFRNGDTASPK KSNFAEALAAHK YQLDKDGVVLFK LGETYKDHENIVIAK VDATEESDLAQQYGVR ILFIFIDSDHTDNQR HNQLPLVIEFTEQTAPK
3	Glucosidase 2 subunit beta	IPI00026154	72	4	59.39/ 4.33	SLKDMEESIR TVKEEAEKPER
4	Isoform 1 of Nucleophosmin	IPI00549248	377	31	32.56/ 4.64	TVSLGAGAKDELHIVEAEAMNYEGSPIK DSKPSSTPR GPSSVEDIKAK TPKGPSSVEDIK SIRDTPAKNAQK SNQNGKDSKPSSTPR DELHIVEAEAMNYEGSPIK
5	Tropomyosin alpha-3 chain	IPI00218319	576	50	29.02/ 4.75	TIDDLEDKLK CREMDEQIR KLVIIEGDLER TEERAELAESR EQAEAEVASLNRR LEEAEKAADESER AREQAEAEVASLNR KIQVLQQQADDAEER RIQLVEEELDRAQER ALKDEEKMELQEIQLK IQVLQQQADDAEERAER KIQVLQQQADDAEERAER LATALQKLEEAEKAADESER

Table S1. Proteins identified from 2D gel.