

Supporting Information for

**Synthesis of Oxime-Linked Mucin Mimics containing the Tumor-Related T_N
and ST_N Antigens**

Lisa A. Marcaurelle, Youngsook Shin, Scarlett Goon and Carolyn R. Bertozzi

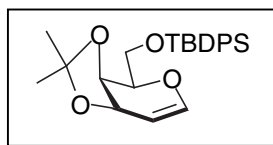
*Center for New Directions in Organic Synthesis, Departments of Chemistry and Molecular and
Cell Biology, and Howard Hughes Medical Institute,
University of California, Berkeley, 94720*

General Methods. N^α-Fmoc-amino acids, MBHA resin and Fmoc-Glu(tBu) resin were purchased from Novabiochem. Fmoc-dehydroleucine was from BACHEM. All other chemical reagents were obtained from commercial suppliers and used without further purification. The following solvents were distilled under a nitrogen atmosphere prior to use: THF was dried and deoxygenated over Na and benzophenone; CH₂Cl₂ and CH₃CN were dried over CaH₂; methanol was dried over Mg and I₂. Unless otherwise noted, all air and moisture sensitive reactions were performed under an argon or nitrogen atmosphere. Analytical thin-layer chromatography (TLC) was conducted on Analtech Uniplat silica gel plates with detection by ceric ammonium molybdate (CAM) and UV light. For flash chromatography, 60 Å silical gel (Bodman) was employed. High-pressure liquid chromatography (HPLC) was performed on a Rainin Dynamax SD-200 HPLC system using Microsorb and Dynamax C₁₈ reverse-phase columns or amino

propyl silica gel columns (analytical: 4.6 x 250 mm, 1 mL/min; preparative: 25 x 250 mm, 20 mL/min) and UV detection was performed with a Rainin Dynamax UV-1 detector.

^1H NMR spectra were obtained at 500 MHz with Bruker DRX 500 spectrometer. Chemical shifts are reported in parts per million (δ) relative to CHCl_3 (7.26 ppm) for spectra run in CDCl_3 . ^{13}C NMR spectra were obtained at 125 MHz on a Bruker DRX 500 spectrometer. High-resolution FAB mass spectra were recorded at the Mass Spectrometry Facility at the University of California at Berkeley. Electrospray ionization mass spectrometry (ESI-MS) was performed on a Hewlett-Packard 1100 mass spectrometer.

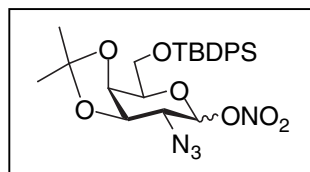
3,4-*O*-Isopropylidene 6-*O*-tert-butyldiphenylsilyl-D-galactal (**14a**).



To a solution of 13.1 g (34.0 mmol) of 6-*O*-TBDPS-D-galactal (**15**) dissolved in 150 mL of DMF was added 120 mL (1.0 mol) of dimethoxypropane and a catalytic amount of PPTS (0.2 mmol). The solution was warmed to 50 °C and stirred for 1 h. The reaction mixture was quenched with anhydrous Na_2CO_3 , filtered and concentrated *in vacuo*. The crude product was purified by silica gel chromatography (15:1 hexanes/EtOAc) to give 13.7 g (95%) **14 a** as a clear yellow syrup: ^1H NMR (500 MHz, CDCl_3) δ 7.69-7.73 (m, 4 H), 7.38-7.44 (m, 6 H), 6.35 (d, 1 H, J = 6.3 Hz), 4.78 (ddd, 1 H, J = 6.2, 2.7, 1.6 Hz), 4.67 (dd, 1 H, J = 6.1, 2.8 Hz), 4.48 (app d, 1 H, J = 6.1 Hz), 4.05 (app t, 1 H, J = 6.9 Hz), 3.93-4.00 (m, 2 H), 1.46 (s, 3 H), 1.38 (s, 3 H), 1.08 (s, 9 H); ^{13}C NMR (125 MHz, CDCl_3) δ 144.81, 165.82, 135.76, 133.59, 133.49, 129.94, 127.92, 127.87, 110.42, 102.81, 75.14, 72.05, 68.75, 63.11, 28.38, 27.06, 26.99, 19.47; HRMS (FAB) calcd for $\text{C}_{25}\text{H}_{32}\text{LiO}_4\text{Si}$ ($\text{M}+\text{Li}$) $^+$ 431.2230, found 431.2225.

2-Azido-2-deoxy-3,4-O-isopropylidene-6-O-tert-butyldiphenylsilyl-D-galactopyranosyl

nitrate (mixture of anomers) (**14b**). To a solution of **14a** (11.7 g, 26.7 mmol) dissolved in 140

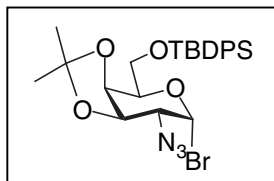


mL of CH₃CN at -15 °C under N₂ atmosphere, was added 82.7 g (45.4 mmol) of ceric ammonium nitrate (CAN) and 2.69 g (41.4 mmol) of sodium azide. After vigorous overhead stirring for 4 h, the mixture was

diluted with 150 mL of cold Et₂O and transferred to a separatory funnel. The organic phase was washed with cold H₂O (3 x 75 mL) and dried over Na₂SO₄. Concentration of the solution *in vacuo*, followed by purification by silica gel chromatography (15:1 hexanes/EtOAc) gave 10.3 g (71%) of **14b** as a mixture of anomers (2:1 α/β): ¹H NMR (500 MHz, CDCl₃) δ 7.66-7.70 (m, 8 H), 7.38-7.46 (m, 12 H), 6.17 (d, 1 H, *J* = 4.0 Hz), 5.46 (d, 1 H, *J* = 9.0 Hz), 4.36-4.37 (m, 2 H), 4.29 (dd, 1 H, *J* = 5.3, 2.0 Hz), 4.24 (app dt, 1 H, *J* = 6.6, 1.8 Hz), 4.14 (dd, 1 H, *J* = 7.3, 5.3 Hz), 4.03 (app dt, 1 H, *J* = 5.3, 2.0 Hz), 3.90-3.95 (m, 3 H), 3.85 (dd, 1 H, *J* = 10.1, 6.3 Hz), 3.78-3.80 (m, 1 H), 3.53 (dd, 1 H, *J* = 8.9, 7.4 Hz), 1.55 (s, 3 H), 1.53 (s, 3 H), 1.37 (s, 3 H), 1.05 (app s, 18 H); ¹³C NMR (125 MHz, CDCl₃) δ 135.79, 135.76, 133.28, 133.20, 130.05, 133.01, 127.97, 127.95, 127.88, 111.15, 110.53, 92.20, 97.13, 77.73, 74.53, 73.53, 72.28, 72.12, 71.11, 62.40, 61.79, 59.62, 28.32, 28.22, 26.88, 26.19, 19.36, 19.33; HRMS (FAB) calcd for C₂₅H₃₂LiO₇Si (M+Li)⁺ 535.2200, found 535.2192.

2-Azido-2-deoxy-3,4-O-isopropylidene-6-O-tert-butyldiphenylsilyl- α -D-galacto-pyranosyl

bromide (14). To a solution of **14b** (9.00g, 17.0 mmol) in CH₃CN (85 mL) was added LiBr

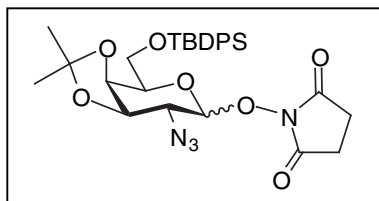


(7.47 g, 86.0 mmol) at 0 °C. After stirring for 5 h at rt the solution was diluted with CHCl₃ and washed with H₂O (3 x). The organic layer was dried over Na₂SO₄ and concentrated to give 8.9 g (96%) of **14** as a clear

yellow syrup: ¹H NMR (500 MHz, CDCl₃) δ 7.69 (ddd, 4 H, *J* = 16.6, 7.9, 1.4 Hz), 7.38-7.46 (m, 6 H), 6.25 (d, 1 H, *J* = 3.5 Hz), 4.40-4.44 (m, 2 H), 4.38 (dd, 1 H, *J* = 5.0, 2.6 Hz), 4.01 (dd, 10.0, 7.4 Hz), 3.89 (dd, 1 H, *J* = 10.0, 6.2 Hz), 3.63 (dd, 1 H, *J* = 5.0, 2.6 Hz), 1.53 (s, 3 H), 1.39 (s, 3 H), 1.07 (s, 9 H); ¹³C NMR (125 MHz, CDCl₃) δ 135.86, 135.77, 133.27, 133.25, 130.05, 130.01, 127.97, 127.88, 110.27, 88.60, 74.71, 73.43, 72.15, 62.66, 62.03, 28.56, 26.93, 26.43, 19.41; HRMS (FAB) calcd for C₂₅H₃₂LiN₃O₄SiBr (M+Li)⁺ 552.1506, found 552.1516.

2-Azido-2-deoxy-3,4-O-isopropylidene-6-O-tert-butyldiphenylsilyl- α -D-galacto-pyranosyl

NHS (16). To a solution of bromide **14** (14.4 g, 26.3 mmol) in dry CH₂Cl₂ (400 mL) was added



4 Å MS (30 g), *N*-hydroxysuccinimide (12.1 g, 105 mmol) and AgClO₄ (13.6 g, 66.0 mmol). After stirring at rt for 2 days in the dark the mixture was diluted with CH₂Cl₂ and filtered over Celite.

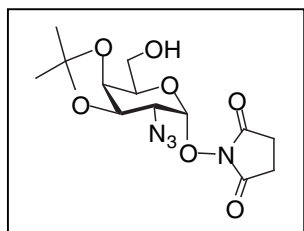
The filtrate was washed with sat'd NaHCO₃ (1 x) and brine (3 x), dried over Na₂SO₄, filtered and concentrated. Purification of the resulting residue by flash chromatography (20–30% EtOAc in hexanes) afforded 9.9 g (65%) of **16** as a mixture of anomers (3:1 α/β).

(**16 α**): ¹H NMR (500 MHz, CDCl₃) δ 7.72 (m, 4 H), 7.42 (m, 6 H), 5.48 (d, 1 H, *J* = 3.5 Hz), 5.08 (app t, 1 H, *J* = 4.5 Hz), 4.50 (m, 2 H), 3.90 (dd, 1 H, *J* = 10.0, 7.5 Hz), 3.84 (dd, 1 H,

$J = 9.5, 6.5$ Hz), 3.64 (dd, 1 H, $J = 7.5, 3.5$ Hz), 2.69 (app d, 4 H, $J = 1.5$ Hz), 1.52 (s, 3 H), 1.38 (s, 3 H), 1.06 (s, 9 H); ^{13}C NMR (125 MHz, CDCl_3) δ 170.46, 135.58, 129.63, 127.66, 127.58, 109.89, 101.67, 72.79, 72.34, 69.85, 60.30, 60.25, 28.30, 26.64, 26.06, 25.39, 19.20; HRMS (ESI-pos) calcd for $\text{C}_{29}\text{H}_{36}\text{NaN}_4\text{O}_7\text{Si}$ ($\text{M}+\text{Na}$) $^+$ 603.2251, found 603.2245.

(**16 β**): ^1H NMR (500 MHz, CDCl_3) δ 7.65 (m, 4 H), 7.41 (m, 6 H), 4.91 (d, 1 H, $J = 8.5$ Hz), 4.21 (dd, 1 H, $J = 5.5, 2.0$ Hz), 4.03 (dd, 1 H, $J = 7.5, 5.0$ Hz), 3.92 (m, 3 H), 3.66 (app t, 1 H, $J = 8.5$ Hz), 2.66 (m, 4 H), 1.55 (s, 3 H), 1.33 (s, 3 H), 1.03 (s, 9 H); ^{13}C NMR (125 MHz, CDCl_3) δ 170.15, 135.53, 135.48, 133.07, 132.84, 129.83, 129.77, 127.74, 127.70, 110.79, 103.86, 74.50, 72.20, 63.33, 62.26, 28.06, 26.62, 26.00, 25.32, 19.13; HRMS (ESI-pos) calcd for $\text{C}_{29}\text{H}_{36}\text{NaN}_4\text{O}_7\text{Si}$ ($\text{M}+\text{Na}$) $^+$ 603.2251, found 603.2244.

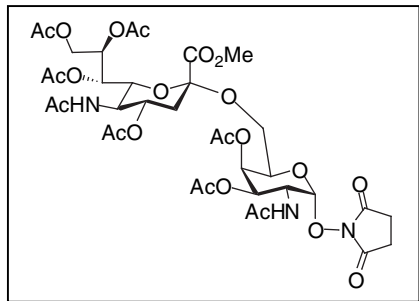
2-Azido-2-deoxy-3,4-O-isopropylidene- α -D-galactopyranosyl NHS (**12**).



To 3.6 g (6.2 mmol) of **16** (mixture of anomers) in THF (20 mL) was added AcOH (710 μL , 12.4 mmol) followed by TBAF (12.4 mL, 1.0 M in THF). After stirring at rt for 6 h the solution was concentrated. Purification of the resulting residue by flash chromatography (50

80% EtOAc in hexanes) afforded 916 mg (45%) of **12**: (^1H NMR (500 MHz, CDCl_3) δ 5.44 (d, 1 H, $J = 4.0$ Hz), 4.82 (m, 1 H), 4.46 (dd, 1 H, $J = 7.5, 6$ Hz), 4.33 (dd, 1 H, $J = 6.0, 2.5$ Hz), 3.83 (m, 3 H), 2.93 (m, 1 H), 2.75 (s, 4 H), 1.48 (s, 3 H), 1.35 (s, 3 H); ^{13}C NMR (125 MHz, CDCl_3) δ 171.07, 110.58, 102.02, 73.39, 72.96, 70.42, 62.39, 58.93, 52.64, 27.65, 25.80, 25.63, 25.43, 20.03, 13.56; HRMS (FAB) calcd for $\text{C}_{13}\text{H}_{19}\text{NN}_4\text{O}_7$ ($\text{M}+\text{H}$) $^+$ 343.1254, found 343.1261.

Disaccharide 18. To a mixture of **12** (978 mg, 2.86 mmol), sialyl phosphite **13** (2.68 g, 3.65

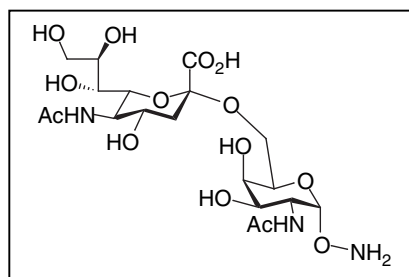


mmol) and 4 Å MS (5.0 g) in anhydrous THF (70 mL) at -35°C was added TMSOTf (140 μL , 0.77 mmol). After stirring for 1 h the reaction was quenched with sat'd NaHCO_3 and warmed to rt. The mixture was filtered through Celite, diluted with EtOAc washed with NaHCO_3 (1 x) and H_2O (3 x), dried

(Na_2SO_4), filtered and concentrated. Purification of the crude residue by flash chromatography (40 – 50 % acetone in hexanes) afforded disaccharide **17** (1.01 g, 44%) as a mixture of anomers (3:1 α/β). To **17** (2.43 g, 2.98 mmol) in MeOH (60 mL) was added *p*-TsOH (57 mg, 0.30 mmol). After stirring at rt overnight the reaction was neutralized by the addition of solid NaHCO_3 , filtered and concentrated. Following purification by flash chromatography (50 – 60% acetone in hexanes), the product (1.37 g, 1.76 mmol) was treated with Ac_2O and pyridine in the presence of DMAP (40 mg). After stirring at rt overnight the mixture was co-evaporated with toluene (3 x), and purified by flash chromatography (30 – 50% acetone in hexanes) to afford the acetylated product (814 mg, 54%). The resulting compound was (804 mg, 0.94 mmol) was dissolved in 1:1 MeOH/ CH_2Cl_2 and stirred under an atmosphere of H_2 in the presence of Pd/C (800 mg) and Ac_2O (176 μL , 1.87 mmol). After stirring for 2 h the mixture was filtered over Celite, rinsed with MeOH and concentrated. The crude product was purified by reversed phase-HPLC (20 – 50% CH_3CN in H_2O over 45 min, 20 mL/min, 220 nm detection) to afford disaccharide **18** (328 mg, 49%) as a white solid: ^1H NMR (500 MHz, CDCl_3) δ 7.12 (d, 1 H, $J=9.0$ Hz), 5.41 (m, 2 H), 5.30 (m, 1 H), 5.25 (dd, 1 H, $J=9.0, 2.0$ Hz), 5.13 (m, 2 H), 4.80 (m, 1 H), 4.72 (m, 1 H), 4.39 (app t, 1 H, $J=6.0$ Hz), 4.20 (dd, 1 H, $J=12.0, 2.5$ Hz), 4.01 (m, 3 H), 3.85 (dd, 1 H, $J=10.5, 7.5$ Hz), 3.75 (s, 3 H), 3.65 (s, 3 H), 3.20 (dd, 1 H, $J=10.0, 4.0$ Hz), 3.05

(m, 2 H), 2.44 (s, 4 H), 2.15 (s, 3 H), 2.11 (s, 3 H), 2.05 (s, 3 H), 1.98 (s, 9 H), 1.96 (s, 3 H), 1.83 (app s, 4 H); ^{13}C NMR (125 MHz, CDCl_3) δ 172.53, 171.01, 170.78, 170.66, 170.50, 170.25, 170.21, 170.12, 169.52, 167.65, 107.60, 98.37, 72.42, 70.82, 68.74, 67.79, 67.40, 63.64, 62.40, 52.89, 51.86, 49.06, 46.90, 37.45, 28.43, 23.06, 20.92, 20.76, 20.65, 20.61, 20.58; ESI-MS (pos) found $(\text{M} - \text{NHS} + \text{H})^+$ 761.

Aminooxy-ST_N 11. To compound **18** in MeOH (27 mL) was added 0.5 M NaOMe (3 mL in



MeOH). The reaction mixture was stirred at rt for 24 h, neutralized with Dowex® 50W-H ion—exchange resin, filtered and concentrated *in vacuo*. The resulting product was dissolved in 3:1 MeOH/H₂O (7 mL) and stirred in the presence

of LiOH (45 mg, 1.9 mmol) at 4 °C overnight. The mixture was neutralized with Dowex® 50W-H ion—exchange resin, filtered and concentrated *in vacuo*. The crude product was dissolved in 10% aqueous N₂H₄·H₂O (17 mL) and stirred at rt overnight. The solution was concentrated *in vacuo* and the resulting product was purified by HPLC on aminopropyl silica gel (90–30% CH₃CN in KH₂PO₄ buffer (pH 5.5) over 45 min, 20 mL/min, 220 nm detection). The product was desalted by reversed phase-HPLC using water as the eluant (20mL/min, 220 nm detection) to afford aminooxy-ST_N **11** (138 mg, 70%) as a white solid: ^1H NMR (500 MHz, D₂O) δ 4.79 (d, 1 H, J = 4.0 Hz), 4.05 (dd, 1 H, J = 11.5, 4.0 Hz), 3.93 (dd, 1 H, J = 7.5, 4.5 Hz), 3.87 (d, 1 H, J = 3.0 Hz), 3.74 (m, 6 H), 3.50 (m, 6 H), 3.53 (dd, 1 H, J = 8.5, 1.0 Hz), 2.60 (dd, 1 H, J = 12.5, 5.0 Hz), 1.90 (s, 3 H), 1.89 (s, 3 H), 1.55 (app t, 1 H, J = 12.5 Hz); ^{13}C NMR (125 MHz, D₂O) δ 180.94, 174.97, 174.54, 173.44, 173.36, 100.48, 100.30, 72.57, 71.10, 69.30, 68.28, 68.13, 67.39,

63.68, 62.57, 51.76, 49.09, 40.13, 32.68, 21.96, 21.86; ESI-MS (pos) found (M + H)⁺ 528, (M + Na)⁺ 551.

Synthesis of GlyCAM fragments 5 and 6. Peptides **5** and **6** were synthesized on MBHA resin (0.1 mmol) and Fmoc-Glu(tBu) Wang resin (0.1 mmol), respectively, using N^α-Fmoc-protected amino acids and DCC-mediated HOBt ester activation in NMP (Perkin-Elmer ABI 431A synthesizer, user-devised cycles). Peptide cleavage/deprotection was accomplished by treatment with 90% aq. TFA for 5 h. The peptides were precipitated with Et₂O, dissolved in 50% aqueous CH₃CN and lyophilized. The crude peptides were purified by preparative RP-HPLC with a gradient of 10-60% CH₃CN in water (0.1% TFA) over 50 min and their identity confirmed by ESI-MS (neg): Peptide **5** calcd. 1356.1, found 1355.7; Peptide **6** calcd, 2021.8, found 2021.7.

Synthesis of mucin-mimics 19 and 20. To 200 μL of peptides **5** and **6** (50 mM ketone) was added 50 μL of 1 M NaOAc buffer (pH 5.5) and 250 μL of aminooxy GalNAc (**7**) (100 mM). The reaction mixture was incubated at 37 °C for 24 h and the oxime-linked products (**19** and **20**) were isolated by RP-HPLC with a gradient of 0-30% CH₃CN in water (0.1% TFA) over 20 min. ESI-MS (neg): Glycopeptide **19** calcd. 2665.3, found 2665.6; Glycopeptide **20** calcd. 3331.3, found 3330.2.

Synthesis of mucin-mimics 21 and 22. To 100 μL of peptides **5** and **6** (50 mM ketone) was added 25 μL of 1 M NaOAc buffer (pH 5.5) and 125 μL of aminooxy ST_N (**11**) (100 mM). The reaction mixture was incubated at 37 °C for 24 h and the oxime-linked products (**21** and **22**)

were isolated by RP-HPLC with a gradient of 0-30% CH₃CN in water (0.1% TFA) over 20 min.

ESI-MS (neg): Glycopeptide **21** calcd. 4412.9, found 4413.2; Glycopeptide **22** calcd. 5078.9, found 5080.0.