

Chemically Tunable Mucin Chimeras Assembled on Living CellsJessica R. Kramer^a, Bibiana Onoa^b, Carlos Bustamante^{b,c,d,e,f,g}, and Carolyn R. Bertozzi^{a,g}^aDepartment of Chemistry, Stanford University, Stanford, CA 94305; ^bInstitute for Quantitative Biosciences, University of California, Berkeley, CA 94720;^cDepartment of Chemistry, University of California, Berkeley, CA 94720; ^dDepartment of Molecular and Cell Biology, University of California, Berkeley,CA 94720; ^eDepartment of Physics, University of California, Berkeley, CA 94720; ^fKavli Energy Nanosciences Institute at Berkeley, University of California,Berkeley, CA 94720; ^gHoward Hughes Medical Institute, University of California, Berkeley, CA 94720; and ^hHoward Hughes Medical Institute, Stanford University, Stanford, CA 94305

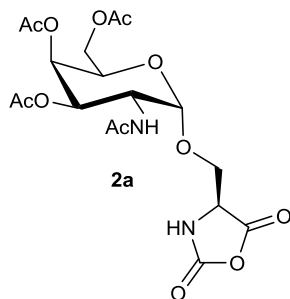
Table of Contents:	Page
I. General Materials and Methods	2
II. Experimental Procedures	3
α -GalNAc-Ser <i>N</i> -carboxyanhydride, 2a	3
β -GlcNAc-Ser <i>N</i> -carboxyanhydride, 2b	3
L-Alanine <i>N</i> -carboxyanhydride, S1	4
γ - <i>tert</i> -Butyl-L-glutamate <i>N</i> -carboxyanhydride, S2	4
<i>N</i> ϵ -Trifluoroacetyl-L-lysine <i>N</i> -carboxyanhydride, S3	5
Synthesis of tetrazine isothiocyanate, S6	5
<i>N</i> α -Alloc-L-leucine 3-azido-propyl amide, S9	6
Preparation of azido amido-amidate nickelacycle, initiator 3	6
General procedure for polymerization of NCAs	7
Kinetic study of the polymerization rate of α -GalNAc Ser and TFA-Lys NCAs	7
General glycopolyptide deprotection procedure	8
General procedure for statistical copolymerizations	8
Copolyptide deprotection	8
General procedure for glycopolyptide end-group modification	9
Circular dichroism spectroscopy	9
TEM sample preparation and imaging	9
Atomic force microscopy imaging and persistence length calculations	10
Mammalian cell culture, transfection, and conjugation of synthetic mucin	10
Cell viability assay	11
III. Supporting Figures and Tables	
Table S1. Polymerization data for β -GlcNAc Ser NCA	11
Table S2. Polymerization data for statistical copolymerization of <i>t</i> Bu-L-Glu and α -GalNAc Ser NCAs	12
Table S3. Polymerization data for statistical copolymerization of TFA-L-Lys and α -GalNAc-Ser NCAs	12
Table S4. Polymerization data for statistical copolymerization of L-Ala and α -GalNAc-Ser NCAs	13
Figure S1. Kinetic data for polymerization of α -GalNAc-Ser and TFA-Lys NCAs with initiator 3	13

Figure S2. ATR-IR of (α -GalNAc-Ser) ₃₀₀	14
Figure S3. CD of (α -GalNAc-Ser) _n , and (α -GalNAc-Ser) ₃₀₀ in PBS	14
Figure S4. CD of poly(α -GalNAc-Ser- <i>stat</i> -L-Glu) _n and CD of poly(α -GalNAc-Ser- <i>stat</i> -L-Ala) _n	15
Figure S5. Transmission electron microscopy of (α -GalNAc-Ser) ₃₀₀	15
Figure S6. Example persistence length plot	16
Figure S7. AFM, example measurements of poly(α -GalNAc-Ser)	16
Figure S8. Cell viability assay	17
IV. Spectral Data	18
V. References	28

I. General Materials and Methods

Unless stated otherwise, reactions were conducted in oven-dried glassware under an atmosphere of nitrogen using anhydrous solvents. THF and DCM were purified by first purging with dry nitrogen, followed by passage through columns of activated alumina. Anhydrous DMF was purchased and stored over 4Å molecular sieves. Deionized water (18 MΩ-cm) was obtained by passing in-house deionized water through a Millipore Milli-Q Biocel A10 purification unit. All commercially obtained reagents were used as received without further purification unless otherwise stated. Reaction temperatures were controlled using an IKA temperature modulator, and unless stated otherwise, reactions were performed at room temperature (RT, approximately 20 °C). Flash chromatography was performed using Silicycle SiliaFlash P60 silica gel. Analytical thin layer chromatography was performed using glass-backed Analtech Uniplate silica gel plates containing a fluorescent indicator, and visualized using a combination of UV, anisaldehyde, and phosphomolybdic acid staining. NMR spectra were obtained on Bruker AVQ-400, AVB-400, DRX-500, or AV-600 spectrometers at ambient temperature at the UC Berkeley College of Chemistry NMR Facility. NMR spectra are reported relative to deuterated solvent signals. Data for ¹H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz) and integration. Splitting patterns are designated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet and br, broad. ¹³C NMR spectra were recorded on Bruker Spectrometers (at 125 MHz). Data for ¹³C NMR spectra are reported in terms of chemical shift. Low resolution mass spectrometry was performed using an Agilent 6120 Single Quad mass spectrometer. High resolution mass spectrometry was performed at the UC Berkeley Mass Spectrometry Laboratory. All attenuated total reflectance (ATR) Fourier Transform infrared (FTIR) samples were recorded on a Bruker Alpha spectrometer, and are reported in terms of frequency of absorption (cm⁻¹). Size exclusion chromatography (SEC) was performed using a Viscotek TDA 302 SEC fitted with a Shodex SB-803 HQ column running 0.10 M LiBr in DMF as the eluent at 60 °C with differential refractive index detection. All GPC/LS samples were prepared at concentrations of 5 mg/mL.

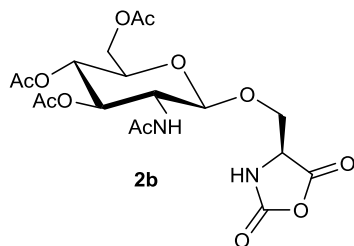
II. Experimental Procedures



O*-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- α -D-galactopyranosyl)-L-serine *N*-carboxyanhydride, α -GalNAc-Ser NCA **2a*

O-(2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- α -D-galactopyranosyl)-L-serine **1** was prepared according to literature procedures.^{1,2} **1** (0.869 g, 2.00 mmol, 1.0 eq) was dissolved in anhydrous THF (45 mL) under N₂. α -Pinene was added (1.09 g, 8.00 mmol, 4.0 eq) followed by phosgene (0.395 g, 4.00 mmol, 2.0 eq, 15 wt % solution in toluene). After 24 hours, the reaction was condensed under high vacuum and the condensate in the traps was quenched with ammonium hydroxide. **Caution! Phosgene is extremely hazardous and all manipulations must be performed in a well-ventilated chemical fume hood with proper personal protection and necessary precautions taken to avoid exposure.** The crude NCA was purified by anhydrous silica chromatography³ on a gradient from 5% dry THF in dry DCM to 35% dry THF in dry DCM. Fractions were visualized with a combination of UV light, and phosphomolybic acid or potassium permanganate dip stains. Fractions containing NCA were combined and condensed to give pure α -GalNAc Ser NCA **2a** (0.682 g, 74% yield) as a white solid.

¹H NMR (500 MHz, CDCl₃, 25 °C): δ 7.59 (s, 1H), 6.53 (s, 1H), 5.41 (s, 1H), 5.12-5.03 (m, 2H), 4.55 (s, 2H), 4.23-4.12 (m, 4H), 3.86 (s, 1H), 2.16 (s, 3H), 2.08-1.98 (m, 9H); ¹³C NMR (125 MHz, CDCl₃, 25 °C): δ 171.25, 171.13, 170.45, 170.45, 167.78, 152.81, 98.30, 67.90, 67.82, 66.99, 62.23, 53.45, 47.48, 21.13, 21.01, 20.81, 20.78. FTIR (THF): cm⁻¹ 3359, 2956, 2855, 1856, 1791, 1719, 1655, 1441, 1367, 1343, 1236, 1184, 118, 1055, 1033, 981, 920, 848, 7567; HRMS-ESI negative ion mode (*m/z*) [M - H]⁻ Calcd for C₁₈H₂₃N₂O₁₂, *m/z*, 459.1251; found 459.1255.

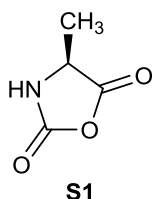


O*-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- β -D-glucopyranosyl)-L-serine *N*-carboxyanhydride, β -GlcNAc-Ser NCA, **2b*

O-(2-Acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- β -D-glucopyranosyl)-L-serine was prepared according to literature procedures.⁴ **1** (0.174 g, 0.400 mmol, 1.0 eq) was dissolved in anhydrous THF (15 mL) under N₂. α -Pinene was added (0.218 g, 1.60 mmol, 4.0 eq) followed by phosgene

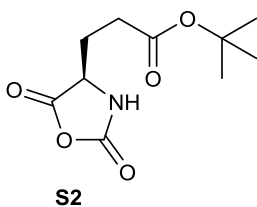
(0.0790 g, 0.0800 mmol, 2.0 eq, 15 wt % solution in toluene). After 48 hours, the reaction was condensed under high vacuum and the condensate in the traps was quenched with ammonium hydroxide. **Caution! Phosgene is extremely hazardous and all manipulations must be performed in a well-ventilated chemical fume hood with proper personal protection and necessary precautions taken to avoid exposure.** The crude NCA was purified by anhydrous silica chromatography³ on a gradient from 10% dry THF in dry DCM to 100% dry THF. Fractions were visualized with a combination of UV light, phosphomolybic acid, and potassium permanganate dip stains. Fractions containing NCA were combined and condensed to give pure β -GlcNAc-Ser NCA **2b** (0.0847 g, 46% yield) as a white solid.

¹H NMR (600 MHz, DMF-d₇, 25 °C): δ 7.59 (s, 1H), 6.53 (s, 1H), 5.41 (s, 1H), 5.12-5.03 (m, 2H), 4.55 (s, 2H), 4.23-4.12 (m, 4H), 3.86 (s, 1H), 2.16 (s, 3H), 2.08-1.98 (m, 9H); ¹³C NMR (125 MHz, DMF-d₇, 25 °C): δ 171.40, 171.24, 170.82, 170.80, 170.56, 169.80, 101.52, 74.60, 72.68, 69.88, 63.03, 54.77, 53.94, 23.51, 21.02, 20.96, 20.96. FTIR (THF): cm⁻¹ 3014, 2996, 2974, 2956, 2921, 2852, 1859, 1786, 1742, 1668, 1524, 1458, 1422, 1400, 1371, 1255, 1224, 1138, 1046, 1032, 943, 923, 798; HRMS-ESI negative ion mode (*m/z*) [M - H]⁻ Calcd for C₁₈H₂₃N₂O₁₂, *m/z*, 459.1251; found 459.1249.



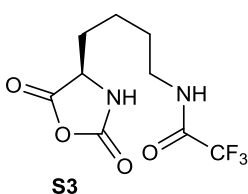
L-Alanine N-carboxyanhydride, S1

L-Alanine (1.87 g, 21.0 mmol, 1.0 eq) was dissolved in anhydrous THF (60 mL) under N₂. Triphosgene (3.11 g, 10.5 mmol, 0.50 eq) was added and the reaction was then heated to 50 °C. After 4 hours at 50 °C, the reaction has cleared. After a total of 5 hours, the reaction was condensed under high vacuum and the condensate in the traps was quenched with ammonium hydroxide. **Caution! Triphosgene generates phosgene, which is extremely hazardous. All manipulations must be performed in a well-ventilated chemical fume hood with proper personal protection and necessary precautions taken to avoid exposure.** The crude NCA was purified by anhydrous silica chromatography³ on a gradient from 5% dry THF in dry DCM to 20% dry THF in dry DCM. Fractions were visualized with a combination of UV light, and phosphomolybic acid or potassium permanganate dip stains. Fractions containing NCA were combined and condensed to give L-Ala NCA **S1** (2.22, 90% yield) as a white solid. Spectral data was in agreement with the literature.⁵



γ -tert-Butyl-L-glutamate N-carboxyanhydride, S2

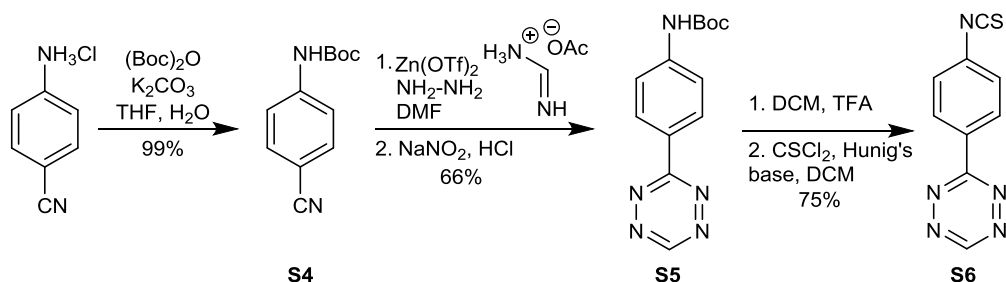
γ -*tert*-Butyl-L-glutamate (0.397 g, 1.95 mmol, 1.0 eq) was dissolved in anhydrous THF (12 mL) under N₂. α -Pinene was added (1.33 g, 1.55 mL, 9.77 mmol, 5.0 eq) was added followed by phosgene (0.386 g, 3.91 mmol, 2.0 eq, 15 wt % solution in toluene). The reaction was then heated to 45 °C for 4 hours under N₂. The reaction was condensed under high vacuum and the condensate in the traps was quenched with ammonium hydroxide. **Caution! Phosgene is extremely hazardous and all manipulations must be performed in a well-ventilated chemical fume hood with proper personal protection and necessary precautions taken to avoid exposure.** The crude NCA was purified by anhydrous silica chromatography³ on a gradient from dry DCM to 10% dry THF in dry DCM. Fractions were visualized with a combination of UV light, and phosphomolybic acid or potassium permanganate dip stains. Fractions containing NCA were combined and condensed to give *t*Bu-L-Glu NCA **S2** (0.411, 92% yield) as an off-white solid. Spectral data was in agreement with the literature.⁶



***N*ε-Trifluoroacetyl-L-Lysine *N*-carboxyanhydride, S3**

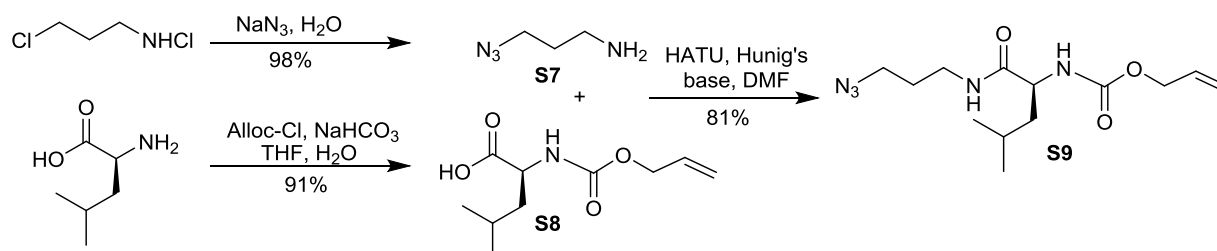
*N*ε-Trifluoroacetyl-L-lysine NCA (TFA-L-Lys) was prepared as previously described and spectral data was in agreement with the literature.⁷

Synthesis of Tetrazine Isothiocyanate S6:



4-Aminomethyl benzonitrile (0.415 g, 2.46 mmol, 1.0 eq) and K₂CO₃ (0.748 g, 5.40 mmol, 2.2 eq) were dissolved in 2:1 THF: H₂O (6 mL) then cooled in an ice bath. Boc anhydride (0.644 g, 2.95 mmol, 1.2 equiv) was added. The reaction was stirred for 16 hours as the ice bath expired. The reaction mixture was poured into water (60 mL) and extracted with DCM (3 x 30 mL). The combined DCM extracts were washed with brine, dried with MgSO₄ and condensed to 0.569 g of **S4** as a white solid which was used directly in the preparation of **S5**. **S5** was prepared according to a literature procedure.⁸ To remove the Boc group, **S5** (11.0 mg, 0.0380 mmol) was dissolved in dry DCM (1.5 mL) and cooled in an ice bath. 4M HCl in dioxane (0.75 mL) was added. The reaction was stirred for 15 min, then the ice bath was removed and the reaction stirred at room temperature for 30 min. The reaction mixture was condensed and the compound was used directly in the next reaction. The amine hydrochloride of **S5** (9.3 mg, 0.0420 mmol, 1.0 eq) was suspended in dry DCM (2 mL) and cooled in an ice bath. Thiophosgene (19.1 mg, 12.7 μ L, 0.210 mmol, 5.0 eq) was added and the reaction stirred for 16 hours as the ice bath expired. The

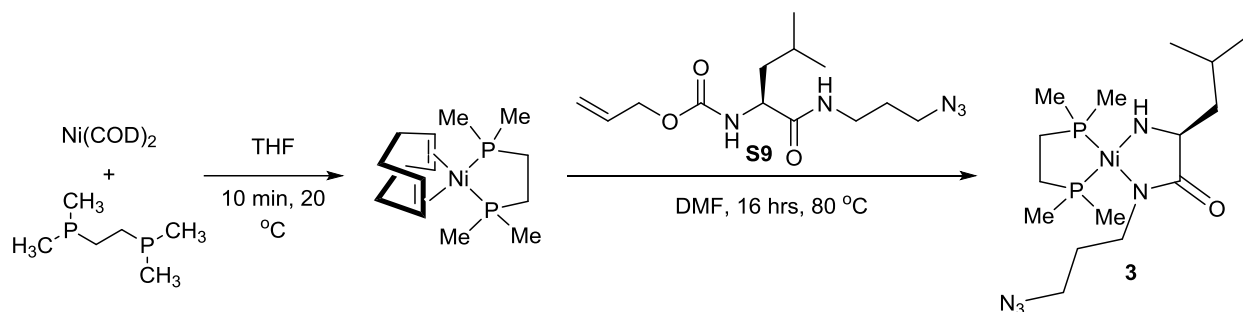
reaction was condensed under high vacuum. The crude residue was taken up in DCM and purified by flash chromatography on silica in 3:1 DCM:hexanes. The bright pink tetrazine isothiocyanate **S6** was condensed to 6.6 mg (75%). Spectral data was in agreement with the literature.⁹



*N*_α-Alloc-L-leucine 3-azido-propyl amide, **S9**

Azidopropylamine¹⁰ **S7** and alloc-leucine¹¹ **S8** were prepared according to literature routes. **S8** (2.39 g, 11.1 mmol, 1.0 eq) was dissolved in anhydrous DMF (5 mL) under N₂ and cooled in an ice bath. Hunig's base (1.43 g, 11.1 mmol, 1.0 eq) was added followed by HATU (4.23 g, 11.1 mmol, 1.0 eq). After 5 minutes, **S7** (1.37 g, 13.3 mmol, 1.2 eq) was added as a solution in anhydrous DMF (2 mL). The reaction was stirred for 16 hours as the ice bath was allowed to expire. The reaction was poured into EtOAc (250 mL), washed with 3 portions of water (3 x 100 mL), then brine. The organic phase was dried with magnesium sulfate, filtered, and condensed. The crude yellow oil was purified by flash chromatography on silica to give **S9** (2.68 g, 81% yield) as a white solid.

¹H NMR (500 MHz, CDCl₃, 25 °C): δ 7.36 (s, 1H), 6.30 (s, 1H), 5.95-5.85 (m, 1H), 5.30 (d, *J* = 21.5, 1H), 5.22 (d, *J* = 13.1, 1H), 5.15 (d, *J* = 9.9, 1H), 4.57 (d, *J* = 6.4, 2H), 4.14-4.08 (m, 1H), 3.37-3.32 (m, 4H), 1.82-1.75 (m, 2H), 7.10-1.63 (m, 2H), 1.56-1.47 (m, 1H), 0.94 (dd, *J* = 6.8, 5.2 6H); ¹³C NMR (125 MHz, CDCl₃, 25 °C): δ 173.30, 156.69, 132.99, 118.15, 66.19, 54.01, 49.42, 41.84, 37.28, 29.04, 25.11, 23.26, 22.42. FTIR (THF): cm⁻¹ 3302, 3085, 2966, 2932, 287, 2096, 1703, 1656, 1534, 1452, 1368, 1243, 1171, 1122, 1045, 339.92, 932.39, 778.62, 649.64; HRMS-ESI (*m/z*) [M + H]⁺ Calcd for C₁₃H₂₄N₅O₃, *m/z*, 298.1879; found 298.1877.



Preparation of azido amido-amidate nickelacycle initiator **3**:

Inside a dinitrogen filled glovebox, Ni(COD)₂ (16.0 mg, 0.0582 mmol, 1.0 eq) was suspended in minimal THF (400 μL) in a pressure vessel. 1,2-Bis(dimethylphosphino)ethane (dmpe) (19.4 μL, 0.116 mmol, 2.0 eq.) was added to the bright yellow Ni(COD)₂ solution followed by stirring for 10 minutes. The solution turned orange with the formation of dmpeNi(COD). *N*_α-Alloc-L-leucine 3-azido-propyl amide **S3** (17.3 mg, 0.0582 mmol, 1.0 eq) was added as a solution in DMF (800 μL). The pressure vessel was sealed, removed from the glovebox, and placed in an oil bath at 80

°C. After 16 hours at 80 °C, the vessel was cooled to room temperature and returned to the glovebox. The orange colored amido-amidate nickelacycle solution was used directly in NCA polymerization reactions.

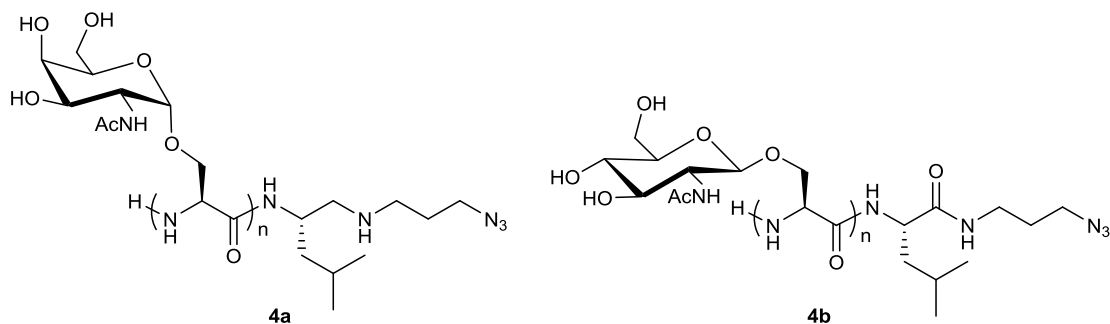
General Procedure for Polymerization of glycosylated Ser NCAs:

All polymerization reactions were performed in a dinitrogen filled glove box. To a solution of α -GalNAc Ser NCA **2a** or β -GlcNAc Ser NCA **2b** in DMF (50 mg/mL) was rapidly added a solution of initiator **3** via syringe. The reaction was stirred at 20 °C and polymerization progress was monitored by ATR-IR. Polymerizations were generally complete within 2 hours. Aliquots were removed for GPC analysis immediately upon polymerization completion. Reactions were dried under reduced pressure, diluted with minimal THF, then precipitated into acidic water (pH = 3, 1 mM HCl) to remove the nickel. Polymers were collected by centrifugation and dried under reduced pressure to yield peracetylated glycopolypeptides as white solids (87–99% yield). Polymer absolute molecular weights were determined by ¹H NMR by integrations calibrated to the azido-Leu end group for low molecular weight polypeptides, or to polypeptides endcapped with PEG-isocyanate according to literature protocols.⁷ Briefly, low polydispersity amino-terminal PEG (MW=1000, purchased from Nanocs) was converted to PEG-isocyanate (PEG-NCO) via treatment with phosgene. Peracetylated polypeptides were treated with 3.0 eq of PEG-NCO per polypeptide chain in THF for 16 hours at RT. Polypeptides were deprotected and dialyzed as described in section ‘General Glycopolypeptide Deprotection Procedure’. Polypeptide integrals were compared to the ethylene glycol peak at δ 3.66 ppm, correlating to 88 H. See the spectral data section for example spectra.

Kinetic study of the polymerization rate of α -GalNAc Ser NCA:

A kinetic study was performed to check that the rates of glycosylated and non-glycosylated NCA polymerizations with initiator **3** are comparable, and to ensure that random copolymers are indeed statistical in their distribution of residues rather than tapered or block-like. ATR-FTIR was used to monitor conversion of NCA to polypeptide, based on comparison of the initial and reaction concentrations of NCA anhydride carbonyl peaks at ca. 1790 cm⁻¹ and 1850 cm⁻¹, and appearance of polypeptide peaks at ca. 1650 cm⁻¹ and 1540 cm⁻¹.¹² The glycosyl acetate peak at ca. 1750 cm⁻¹ was used as an internal standard since this concentration does not change over the course of the polymerization. α -GalNAc-Ser NCA polymerization was compared to TFA-Lys NCA polymerization. Polymerization of α -GalNAc-Ser and TFA-Lys NCAs was initiated at an M:I ratio of 200:1 as described in the section ‘General Procedure for Polymerization of glycosylated Ser NCAs’. Two μ L aliquots of the polymerization reactions were analyzed by ATR-FTIR at various timepoints after addition of initiator **3**. Studies were performed in duplicate and standard deviations were less than 0.05.

The polymerization rate constants ($k_{p,obs}$) were calculated by plots of natural logarithm of monomer concentration versus time, and linear fitting of the data using the following equation: $[-\ln(M_t/M_i)=k_{p,obs}t]$ where M_t = concentration of monomer at time (t), M_i = initial concentration of monomer, $k_{p,obs}$ = rate constant, and t = time. Reaction half-lives were calculated from the equation $t_{1/2} = (\ln 2)/k_{p,obs}$.



General Glycopolypeptide Deprotection Procedure:

To a solution of peracetylated glycopeptide in methanol (10 mg/mL) was added hydrazine monohydrate (4.0 eq per OAc group) and the reaction was stirred for 6 hours, with formation of a white precipitate. Water was added to make the solution 1:1 methanol:water and the reaction was allowed to proceed for a total of 48 hours at room temperature. Reactions were quenched by addition of drops of acetone and then condensed under reduced pressure. The solid glycopolypeptides were dissolved in water and transferred to dialysis cassettes (3500 Da molecular weight cutoff) and dialyzed against 4L of deionized water for 3 days, with water changes twice per day. This procedure removes all small molecule and residual nickel impurities. Dialyzed polymers were lyophilized to dryness to give poly(*O*-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- α -D-galactopyranosyl)-L-serine), (α -GalNAc-Ser) **4a** or poly(*O*-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl- β -D-glucopyranosyl)-L-serine), poly(β -GlcNAc-Ser) **4b** as white fluffy solids. (~90% yield after dialysis). PEG-encapped glycopolypeptides were deprotected using the exact same conditions.

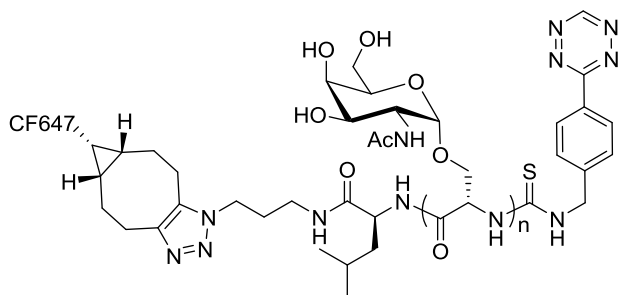
General Procedure for Statistical Copolymerizations:

All polymerization reactions were performed in a dinitrogen filled glove box. To a solution of α -GalNAc-Ser NCA **2a** in DMF (50 mg/mL) was added L-Ala NCA **S1**, *t*Bu-L-Glu NCA **S2**, or TFA-L-Lys NCA **S3** as a solution in DMF (50 mg/mL) in the molar ratios described in Tables S1, S2, and S3. Initiator **3** was rapidly added as a solution via syringe. The reaction was stirred at 20 °C and polymerization progress was monitored by ATR-IR. Polymerizations were generally complete within 2 hours. Aliquots were removed for GPC analysis immediately upon polymerization completion. Reactions were dried under reduced pressure, diluted with minimal THF, then precipitated into acidic water (pH = 3, 1 mM HCl). Polymers were collected by centrifugation and dried under reduced pressure to yield peracetyled statistical glycopolypeptides.

Copolyptide Deprotection:

Acetate groups were removed with hydrazine monohydrate in MeOH as described in the general glycopolypeptide deprotection procedure. After 48 hours with hydrazine monohydrate in 1:1 water:MeOH, copolypeptide solutions were quenched by addition of drops of acetone and then condensed under reduced pressure. L-Ala/ α -GalNAc-Ser copolypeptides were taken up in water and directly dialyzed to give poly(α -GalNAc-Ser-*stat*-L-Ala) as fluffy white solids (92-99% yield). *t*Bu-L-Glu/ α -GalNAc-Ser copolypeptides were redissolved in TFA (10 mg/mL) as stirred at room temperature for 4 hours. The TFA was evaporated and the copolypeptides were taken up

in water and directly dialyzed to give poly(α -GalNAc-Ser-*stat*-L-Glu) as fluffy white solids (92-99% yield). TFA-L-Lys/ α -GalNAc-Ser copolypeptides were redissolved in 1:1 MeOH:water and K_2CO_3 was added (3 eq per TFA group). The reaction was stirred for 8 hours at 50 °C. After cooling to room temperature the solution was acidified to ~ pH 4 with 1M HCl. The methanol was evaporated under reduced pressure and the copolypeptide solution was dialyzed to give poly(α -GalNAc-Ser-*stat*-L-Lys) as fluffy white solids (92-99% yield). All copolypeptides were found to be fully water soluble with the exception of 1:3 α -GalNAc Ser:L-Ala. Amino acid ratios were calculated by 1H NMR integration, see spectral data section. All copolypeptide NMR spectra were obtained in D_2O with the exception of 1:3 α -GalNAc Ser:L-Ala, which was obtained in d-TFA.



General Procedure for Glycopeptide End-Group Modification, CF647-poly(α -GalNAc-Ser)-Tz 5:

To a solution of glycopeptide in Milli-Q water (10 mg/mL) was added CF647-BCN (2.0 eq per N_3 group, photostable Cy5 equivalent purchased from Biotium). The reaction was covered in foil and incubated for 24 hours at room temperature. The reaction mixture was transferred to a 3500 MWCO dialysis cassette and dialyzed against 4L of Milli-Q water for 24 hours at RT covered with foil, the water was changed 3 times. The dialysate was removed from the cassette and lyophilized to dryness. The lyophilized powder was taken up in PBS buffer, pH 8 (10 mg/mL). Tetrazine isothiocyanate **S6** (15.0 eq per NH_2 group) was added and the reaction incubated for 16 hours at 4 °C. The reaction was extracted with 3 portions of diethyl ether to remove excess tetrazine, then transferred to a 3500 MWCO dialysis cassette and dialyzed against 4L of Milli-Q water for 24 hours at 4 °C in the dark, with 2 water changes. The dialysate was lyophilized to dryness to yield CF647-poly(α -GalNAc-Ser)-Tz **5** as a blue solid.

Circular Dichroism Spectroscopy

Circular dichroism spectra were recorded on a JASCO CD spectrophotometer running in conventional scanning mode with samples in a quartz cuvette of 0.1 cm path length and prepared at 0.5 mg/mL using Milli-Q water. All spectra were recorded as an average of 3 scans. The spectra are reported in units of molar ellipticity $[\theta]$, ($deg \cdot cm^2 \cdot dmol^{-1}$). The formula used for calculating molar ellipticity, $[\theta]$, was $[\theta] = (\theta \times 100 \times M_w) / (c \times l)$ where θ is the experimental ellipticity in millidegrees, M_w is the average molecular weight of a residue in g/mol, c is the peptide concentration in mg/mL; and l is the cuvette pathlength in cm.

TEM Sample Preparation and Imaging:

Solutions of (α -GalNAc Ser)₃₀₀ in Milli-Q water (20 μ L of 10 μ M) were deposited on plasma-treated carbon-coated hydrophilic grids for 2 minutes then rinsed with Milli-Q water. Samples were then stained with uranyl acetate solution for 1 min. Grids were imaged using a FEI Tecnai 12 Transmission electron microscope.

Atomic Force Microscopy (AFM) Imaging and Persistence Length Calculations

Sample preparation and imaging:

Solutions of glycopeptide in Milli-Q water (10 μ L of 10-20 nM) were deposited on freshly cleaved bare mica for 1 minute, rinsed with Milli-Q water, then gently dried under a stream of N₂ perpendicular to the mica surface.

AFM measurements were taken with a MultiMode NanoScope V atomic force microscope (Bruker Co.) equipped with a vertical engagement scanner E. The samples were imaged in tapping mode using silicon cantilevers (Nanosensors), excited at their resonance frequency (280-350 kHz) with free amplitudes (A_0) of 2-10 nm. The image amplitude (set point A_s) and A_0 ratio (A_s/A_0) was kept at ~ 0.8 in a repulsive tip-sample interaction regime, and phase oscillations no greater than ± 5 degrees. The surface was rastered following the fast scan axis (x) at rates of 1.5- 2 Hz, the retrace line was captured to reconstruct the AFM micrographs. All samples were measured at room temperature in air, at a relative humidity of 30%.

Image processing and persistence length calculation:

Raw AFM images were flattened and leveled using Gwyddion 2.3.¹³ Molecular end-to-end distances (R) and contour lengths (L) were measured from filaments selected by height threshold segmentation (package features) from the particle and pore analysis module included in SPIP™. Only molecules that were equilibrated on the mica surface, fully contained within images, and greater than 25 nm in length were measured. The persistence length (P) (i.e. how far a polymeric chain persists in a given direction) was calculated using the worm-like-chain (WLC) model. It has been demonstrated that the WLC model satisfactorily relates 2D biopolymer conformations equilibrated onto mica surfaces, with their actual 3D conformations.^{14,15} Thus, to estimate the polypeptide persistence length, we used the equation which relates P to the mean-square R as follows:

$$\text{Equation 1: } \langle R^2 \rangle_{eq} = 4PL \left(1 - \frac{2P}{L} \left(1 - e^{-\frac{L}{2P}} \right) \right).$$

Mammalian cell culture, transfection, and conjugation of synthetic mucin.

HEK 293T cells (purchased from ATCC) were seeded and grown on polylysine coated 8-well glass culture dishes (LabTek) at 37 °C in a CO₂ incubator. 24 Hours before transfection, cells were plated in DMEM with 1% FBS containing no antibiotics. At $\sim 80\%$ confluence, cells were simultaneously transfected using Lipofectamine 2000 (Life Technologies) with two plasmids, *p4CMVE-U6-PylT* and *pMmPylRS-EGFR(128TAG)-GFP*. Three hours before addition of plasmid DNA, media was exchanged for DMEM with 0.1% FBS, no antibiotics, and containing 1 mM of amino acid **6** or **7**, or no added amino acid as indicated. According to the manufacturer protocols, transfection was conducted using 0.48 μ g total plasmid DNA (0.24 μ g each plasmid) and 1.0 μ L Lipofectamine 2000 per well in a total of 60 μ L Optimem. Plasmid DNA was combined in Optimem and added to a solution of Lipofectamine in Optimem, the solution was gently mixed, and incubated for 10 minutes before addition to cells. Sixteen hours after transfection, media was replaced with DMEM with 1% FBS and cells were grown for another 24

hours. Cells were washed twice to remove any remaining amino acid **6** or **7**. Media was then replaced with DMEM with 0.1% FBS and 5 μM CF647-(α -GalNAc-Ser)₃₀₀-Tz, **5**. After 4 hours at 37 °C in a CO₂ incubator cells, media was replaced with DMEM with 1% FBS and containing 2 μM Hoechst 33342 (purchased from Biotium). Cells were incubated for 15 min to allow nuclear staining. Cells were then washed 2x for 5 minutes per wash in DMEM with 1% FBS. Cells were then imaged in dye free DMEM with 1% FBS on a Zeiss AxioVert 200M inverted microscope using a PlanNeofluor 40x/0.75 objective. A 175W xenon lamp housed in a Sutter DG4 illuminator linked to the microscope by an optical fiber assured shuttering and illumination. Images were acquired using the Cy5 filter for labeling with CF647-glycopeptide, the FITC filter for production of EGFR-GFP, and the DAPI filter for labeling with Hoescht. Images were acquired and processed using SlideBook 5.0, and are shown as a single z plane. Images acquired using the Cy5 filter are false-colored red, the FITC filter are false-colored green, and the DAPI filter are false-colored blue. All fluorescence images were exposed for 1 second.

Cell Viability Assay.

HEK 293T cell viability under various conditions was determined using a fluorescence based viability assay kit based on ethidium homodimer III (EthIII, purchased from Biotium). In a 12-well plate, HEK 293T cells were transfected with *p4CMVE-U6-PylT* and *pMmPylRS-EGFR(128TAG)-GFP* in the presence of **6** as previously described. Mucin chimeras were prepared using 5 μM (α -GalNAc Ser)₃₀₀-Tz (not labelled with CF647 since this would interfere with the EthIII fluorescence) as previously described. To determine toxicity of the synthetic mucin with untreated cells, cells grown under the same conditions but not transfected were treated with 5 μM (α -GalNAc Ser)₃₀₀-Tz for 4 hours. As a positive control for cell death, HEK 293T cells grown under the same conditions were treated with 0.3 mM Triton X-100 for 5 min immediately before EthIII treatment.¹⁶ Untreated HEK 293T cells grown under the same conditions were used as a negative control. After the various treatments, the viability assay was performed according to the manufacturer's protocols. Briefly, cells were treated with 4 μM EthD-III in serum free DMEM for 30 minutes at room temperature. Cells were then trypsinized, pelleted, and resuspended in PBS with 0.1% FBS. EthIII fluorescence inside dead cells (excitation/emission 530 nm/635 nm) was quantified by flow cytometry. Data shown is from 3 biological replicates measured in duplicate.

III. Supporting Figures and Tables

Table S1. Polymerization data for β -GalNAc-Ser NCAs at various monomer to initiator ratios using amido-amidate nickelacycle initiator **3** in DMF at 20 °C.

[M]/[I]	M_n ^[a]	DP ^[b]	PDI ^[c]	yield (%) ^[d]
50	29,148	70	1.19	94
100	57,880	139	1.14	98
200	120,756	290	1.09	99

[a] Molecular weight after polymerization as determined by ¹H NMR and GPC. [b] DP = number average degree of polymerization. [c] Polydispersity index (M_w/M_n) after polymerization as determined by GPC. [d] Total isolated yield of glycopolypeptide.

Table S2. Polymerization data for statistical copolymerization of *t*Bu-L-Glu and α -GalNAc-Ser NCAs at a monomer to initiator ratio ([M]/[I]) of [100]/[1] using amido-amidate nickelacycle initiator **3** in DMF at 20 °C.

[M]/[I]	α -GalNAc-Ser: <i>t</i> Bu-L-Glu	M_n ^[c]	DP ^[d]	PDI ^[e]	yield (%) ^[f]
100	95:5 ^[a] , N/A ^[b]	47,769	118	1.14	92
100	75:25 ^[a] , 75:25 ^[b]	43,748	122	1.09	93
100	60:40 ^[a] , 63:37 ^[b]	38,546	119	1.07	95
100	25:75 ^[a] , 27:73 ^[b]	30,619	126	1.11	91

[a] Feed ratio of two monomers. [b] Amino acid ratio by ¹H NMR, peaks were too small to integrate for the 95:5 sample. [c] Molecular weight after polymerization as determined by ¹H NMR and GPC. [d] DP = number average degree of polymerization. [e] Polydispersity index (M_w/M_n) after polymerization as determined by GPC. [f] Total isolated yield of glycopolypeptide.

Table S3. Polymerization data for statistical copolymerization of TFA-L-Lys and α -GalNAc-Ser NCAs at a monomer to initiator ratio ([M]/[I]) of [100]/[1] using amido-amidate nickelacycle initiator **3** in DMF at 20 °C.

[M]/[I]	α -GalNAc-Ser:TFA-L-Lys	M_n ^[c]	DP ^[d]	PDI ^[e]	yield (%) ^[f]
100	95:5 ^[a] , N/A ^[b]	50,439	124	1.12	94
100	75:25 ^[a] , 72:28 ^[b]	43,463	118	1.10	96
100	60:40 ^[a] , 63:37 ^[b]	37,685	111	1.09	99
100	25:75 ^[a] , 30:70 ^[b]	31,851	117	1.07	97
300	75:25 ^[a] , 73:27 ^[b]	68,161	317	1.16	98
300	50:50 ^[a] , 52:48 ^[b]	53,717	325	1.13	98

[a] Feed ratio of two monomers. [b] Amino acid ratio by ¹H NMR, peaks were too small to integrate for the 95:5 sample. [c] Molecular weight after polymerization as determined by ¹H NMR and GPC. [d] DP = number average degree of polymerization. [e] Polydispersity index (M_w/M_n) after polymerization as determined by GPC. [f] Total isolated yield of glycopolypeptide.

Table S4. Polymerization data for statistical copolymerization of L-Ala and α -GalNAc-Ser NCAs at a monomer to initiator ratio ([M]/[I]) of [100]/[1] using amido-amidate nickelacycle initiator **3** in DMF at 20 °C.

[M]/[I]	α -GalNAc-Ser: L-Ala	$M_n^{[c]}$	DP ^[d]	PDI ^[e]	yield (%) ^[f]
100	95:5 ^[a] , 96.5:3.5 ^[b]	48,293	121	1.13	91
100	75:25 ^[a] , 79:21 ^[b]	40,596	123	1.18	96
100	60:40 ^[a] , 64:36 ^[b]	32,834	118	1.19	99
100	25:75 ^[a] , 22:78 ^[b]	19,989	127	1.24	99

[a] Feed ratio of two monomers. [b] Amino acid ratio by ¹H NMR. [c] Molecular weight after polymerization as determined by ¹H NMR and GPC. [d] DP = number average degree of polymerization. [e] Polydispersity index (M_w/M_n) after polymerization as determined by GPC. [f] Total isolated yield of glycopolypeptide.

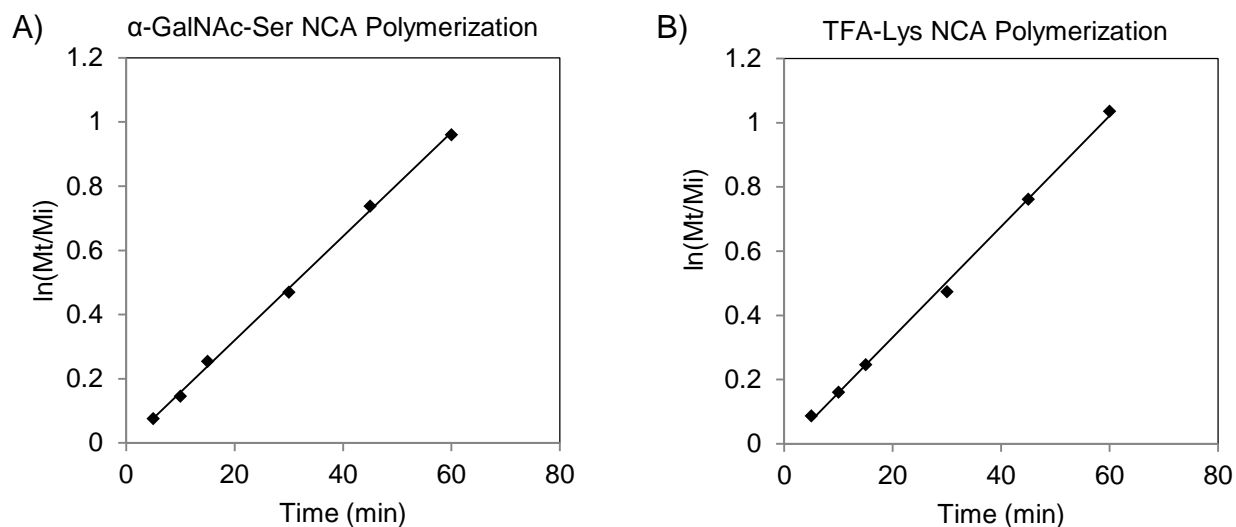


Figure S1. Kinetic Plot of NCA Polymerization with Initiator 3

A) α -GalNAc-Ser NCA, $k_{p,obs} = 0.0162 \text{ min}^{-1}$, $t_{1/2} = 42.7 \text{ min}$; and B) TFA-Lys NCA $k_{p,obs} = 0.0173 \text{ min}^{-1}$, $t_{1/2} = 40.1 \text{ min}$. Initial NCA concentration was 0.109 M, M:I = 200:1.

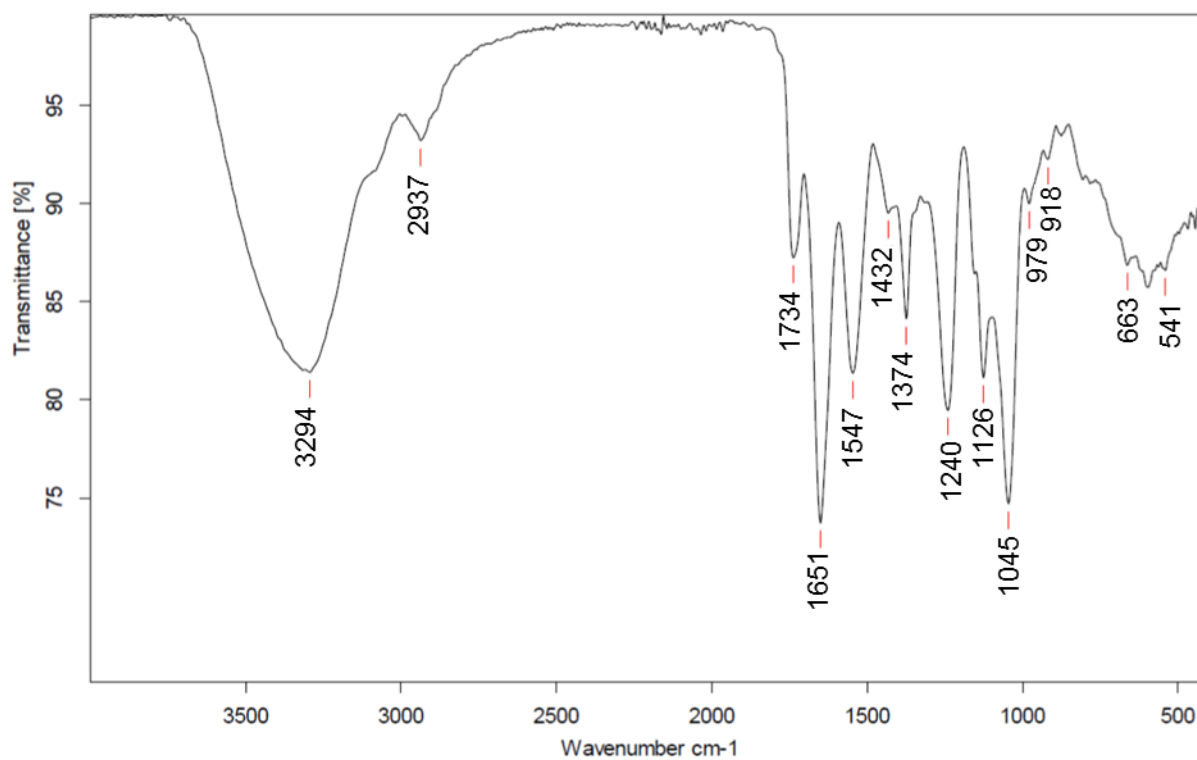


Figure S2. ATR-IR of $(\alpha\text{-GalNAc-Ser})_{300}$.

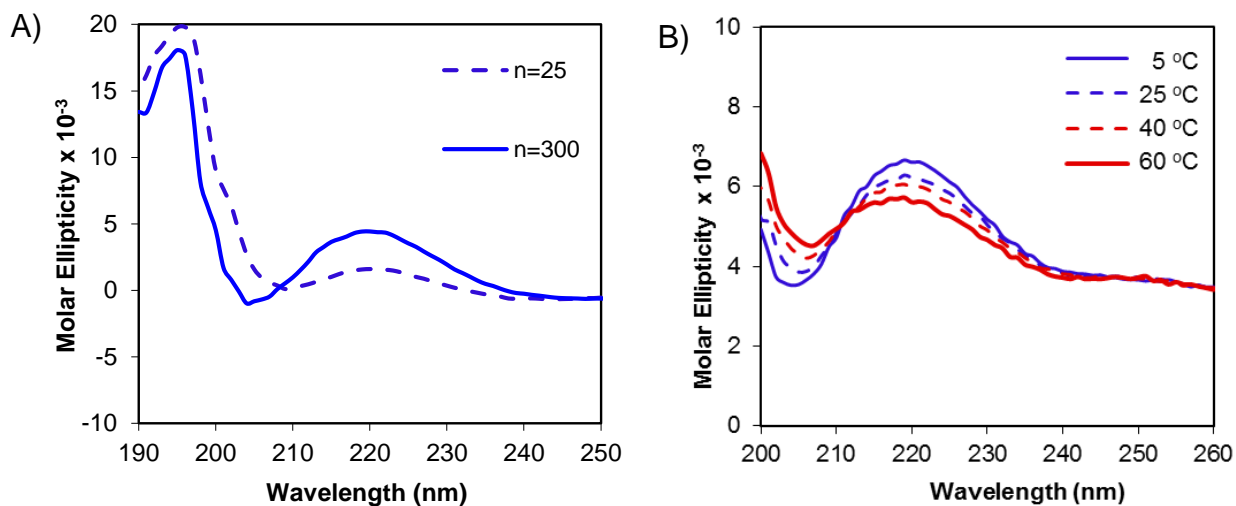


Figure S3. A) CD of $(\alpha\text{-GalNAc-Ser})_n$, 0.5 mg/mL in Milli-Q water; B) CD of $(\alpha\text{-GalNAc-Ser})_{300}$, 0.5 mg/mL in PBS at varied temperatures.

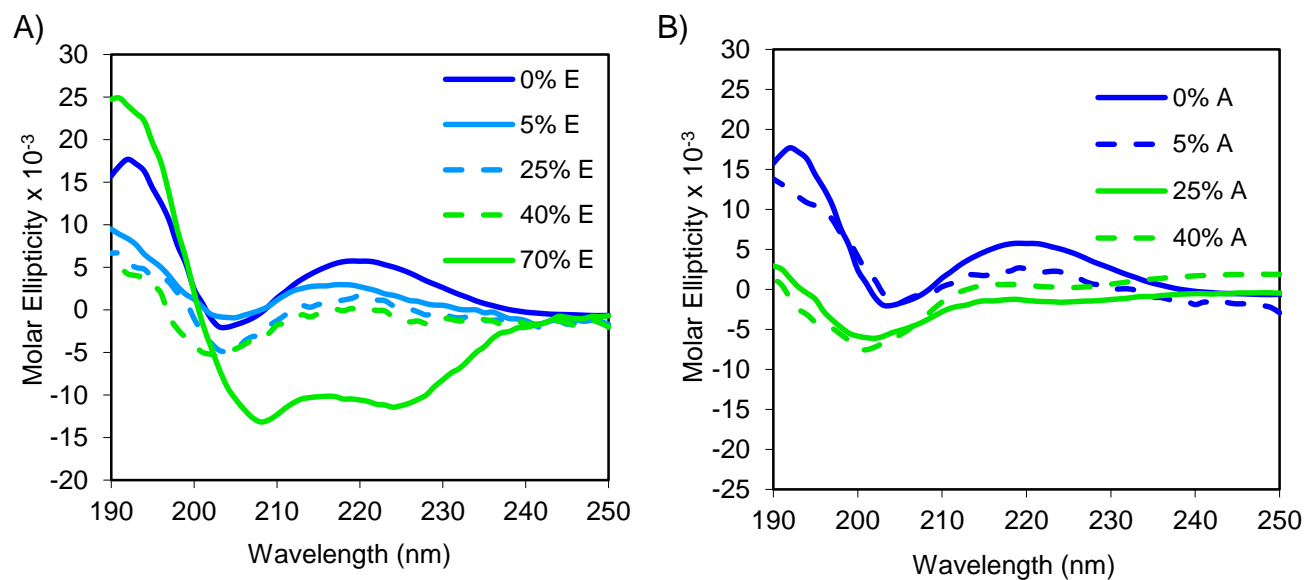


Figure S4. A) CD of poly(α -GalNAc-Ser-*stat*-L-Glu)_n, copolypeptide data are listed in Table S2; B) CD of poly(α -GalNAc-Ser-*stat*-L-Ala)_n, copolypeptide data are listed in Table S4. Data is not shown for the 75% Ala 25% α -GalNAc-Ser copolypeptide since it was not soluble in water. 0% E and 0% A are poly(α -GalNAc-Ser)₁₂₅, PDI = 1.13. All spectra are 0.5 mg/mL in Milli-Q water.

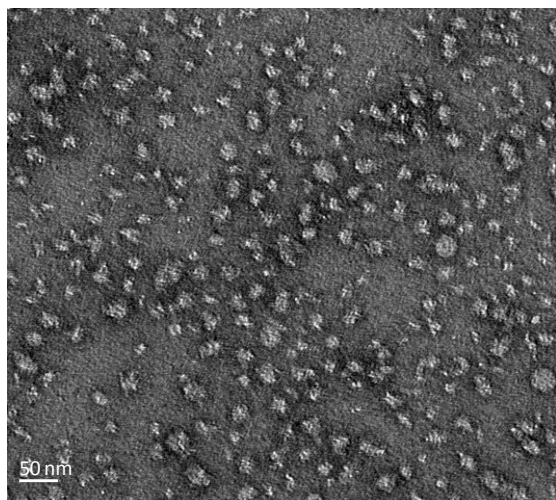


Figure S5. Transmission electron microscopy of 10 μ M (α -GalNAc-Ser)₃₀₀, stained with uranyl acetate.

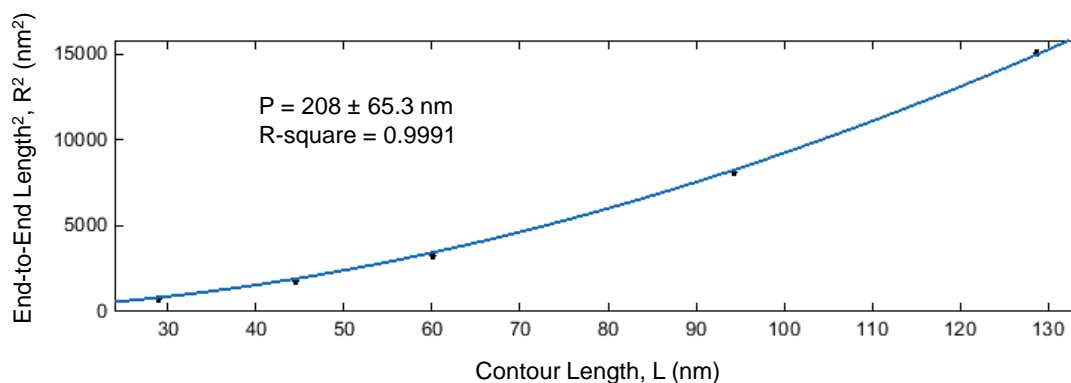


Figure S6. Example plot of the mean squared end-to-end distance versus contour length for poly(α -GalNAc-Ser) and corresponding fit to equation 1.

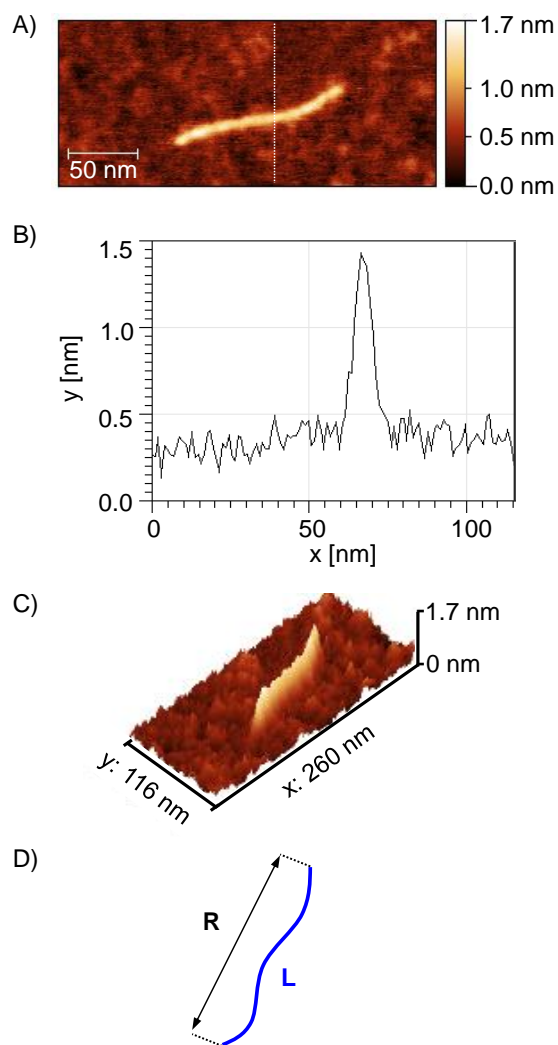


Figure S7. AFM image of poly(α -GalNAc-Ser)₃₀₀, A) Top down 2-D view of poly(α -GalNAc-Ser)₃₀₀, white line is cross-section site represented in panel B; B) Representative height profile analysis, the measured width of 13 nm for the molecule correlates well with the convolution of

the AFM tip; C) 3D perspective of image shown in panel A; D) cartoon representation of measurement of R and L used to calculate persistence length.

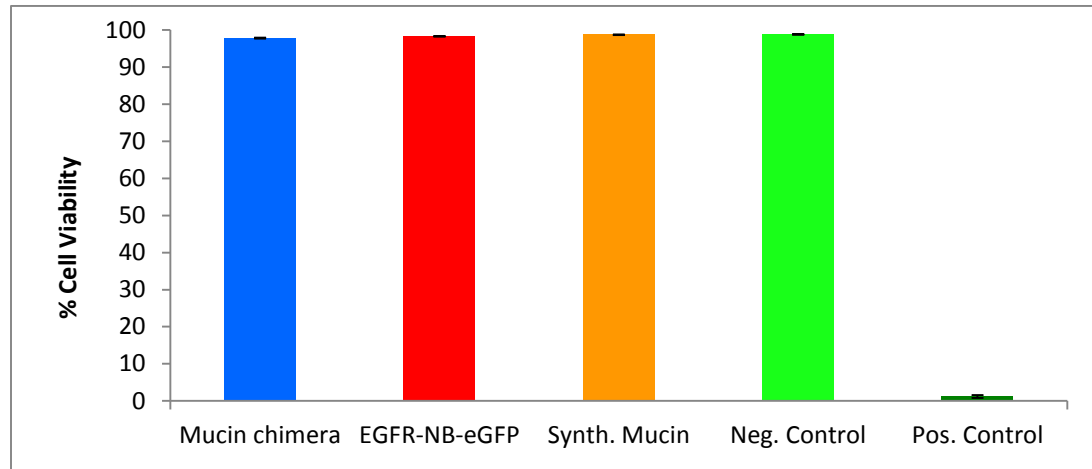
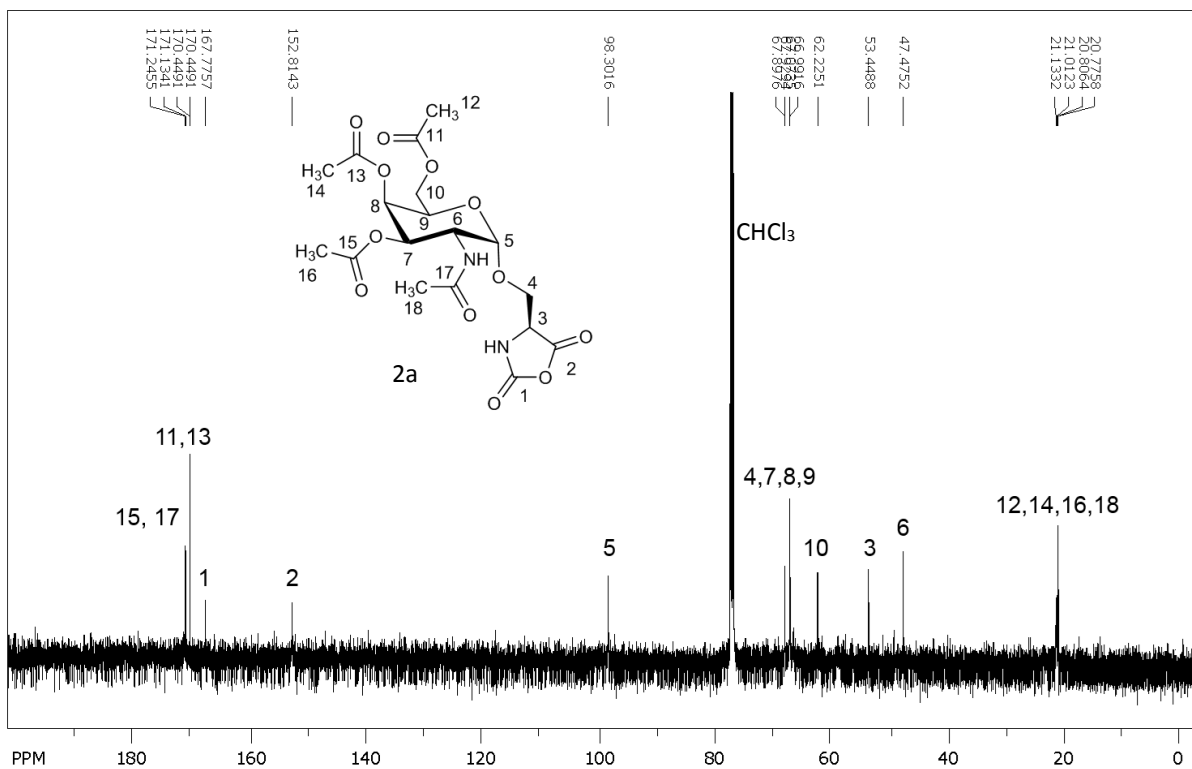
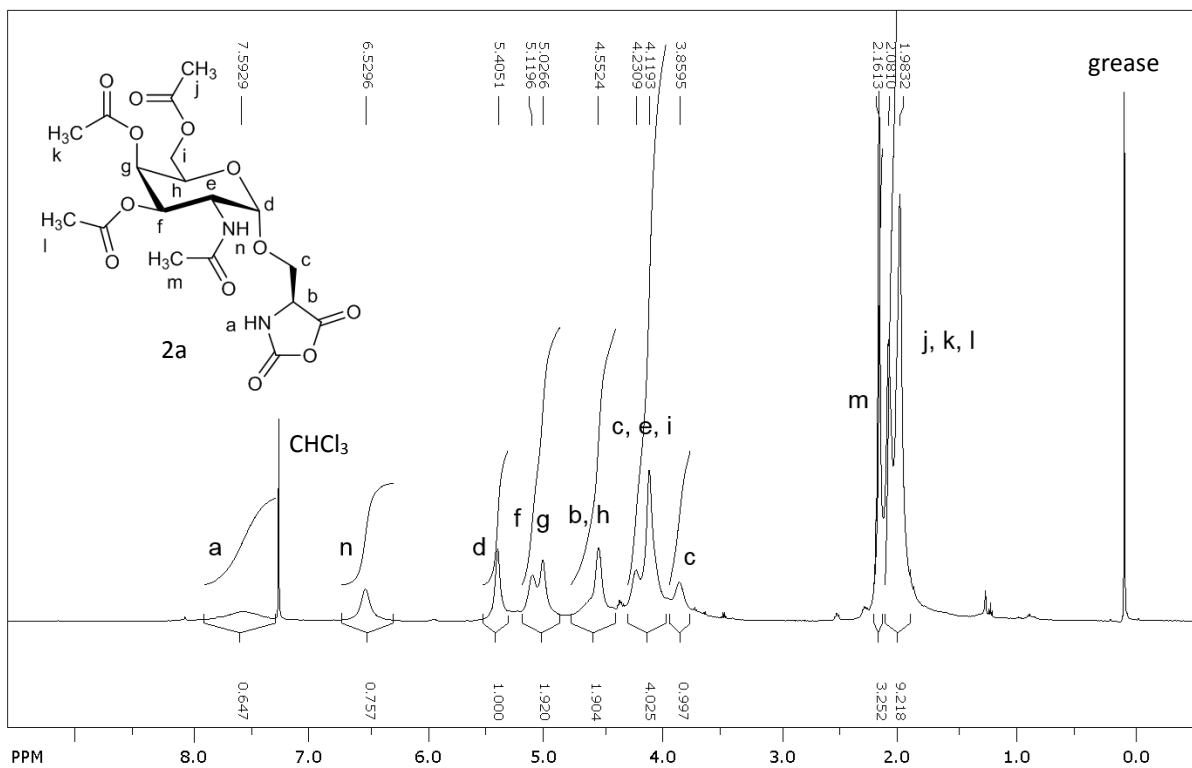
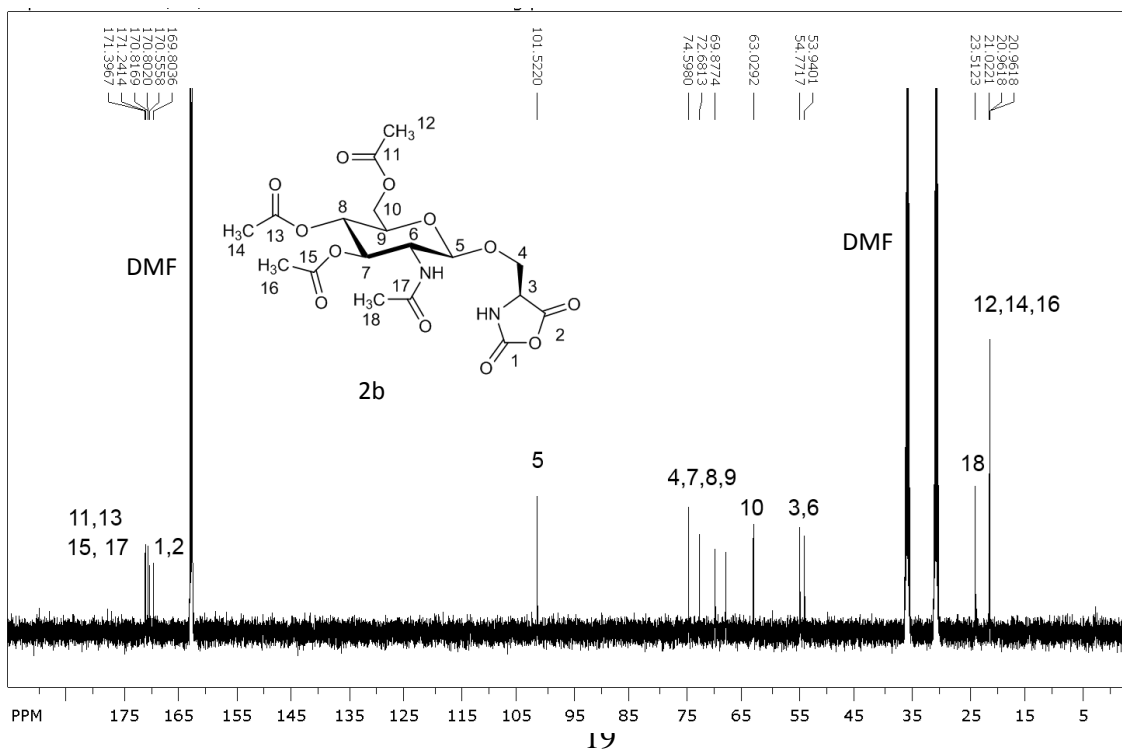
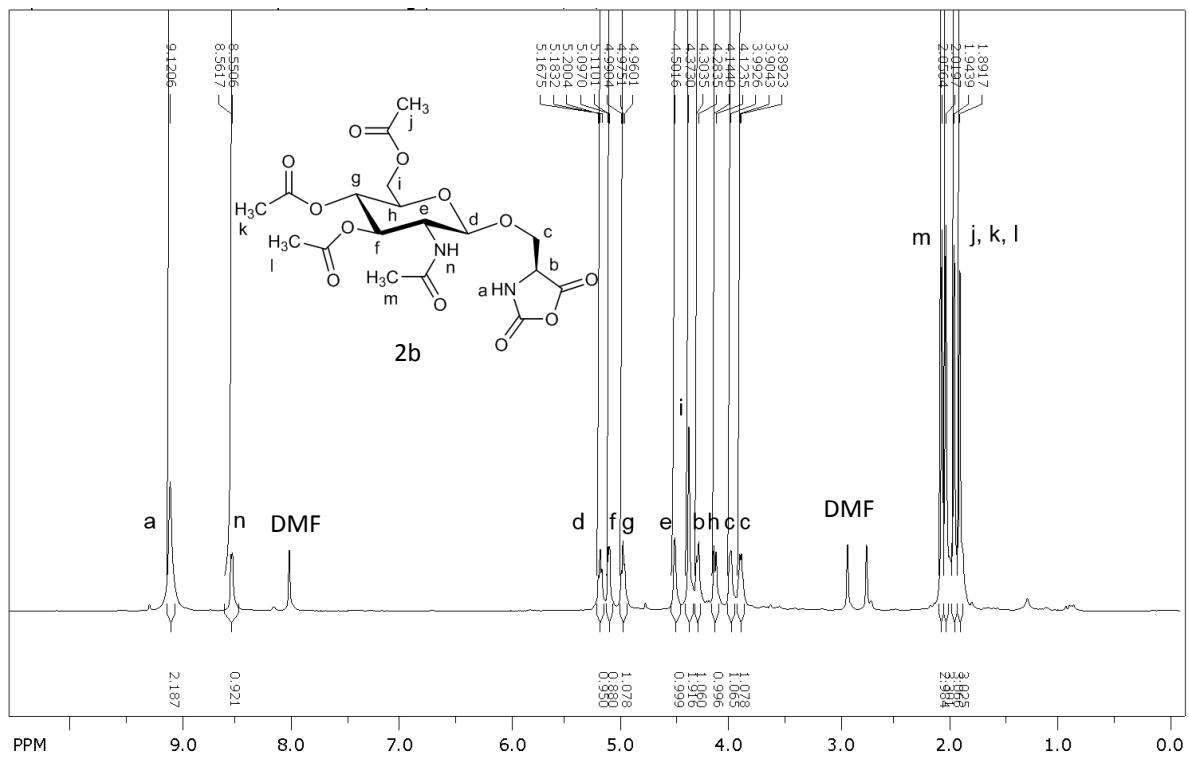


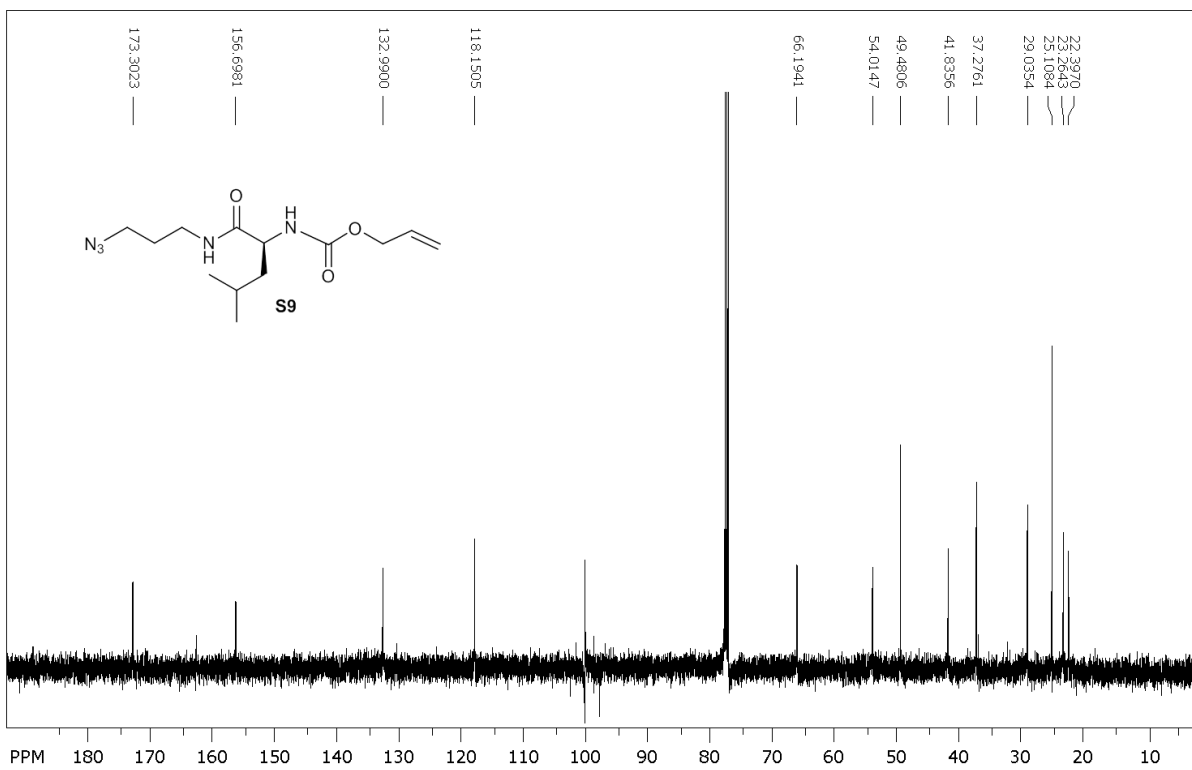
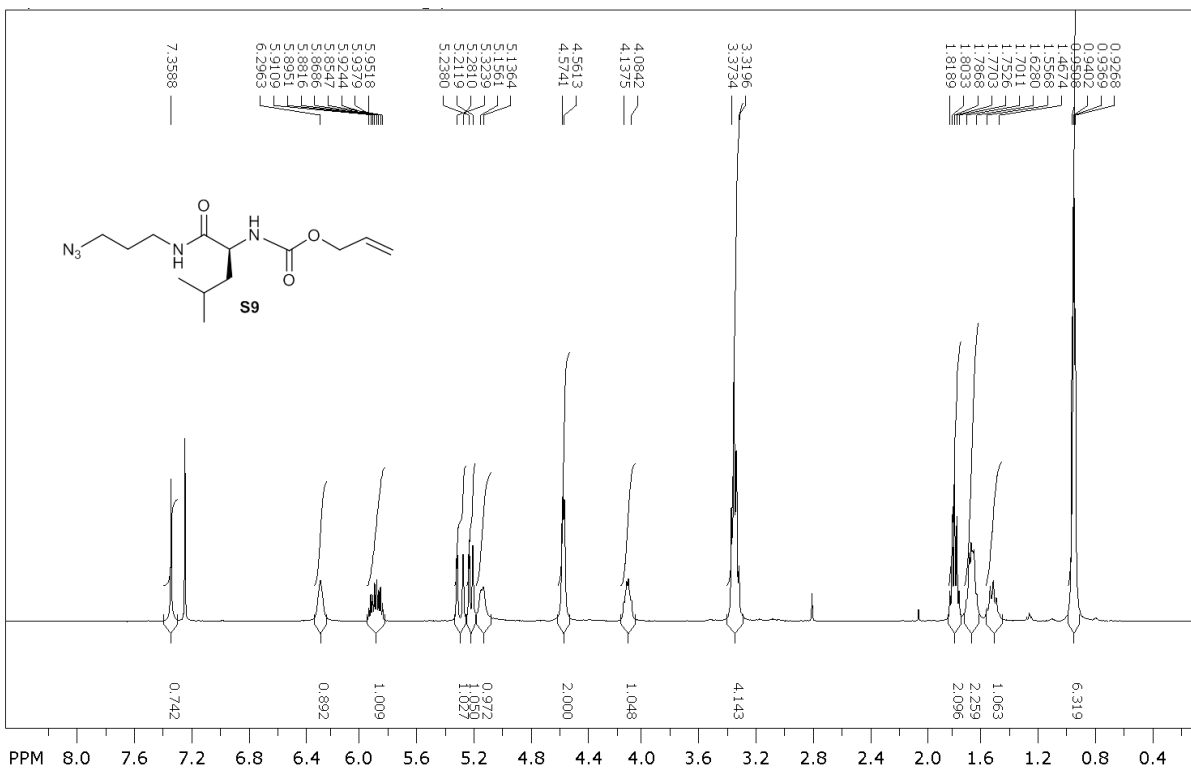
Figure S8. Ethidium Homodimer III Cell Viability Assay Under Various Conditions

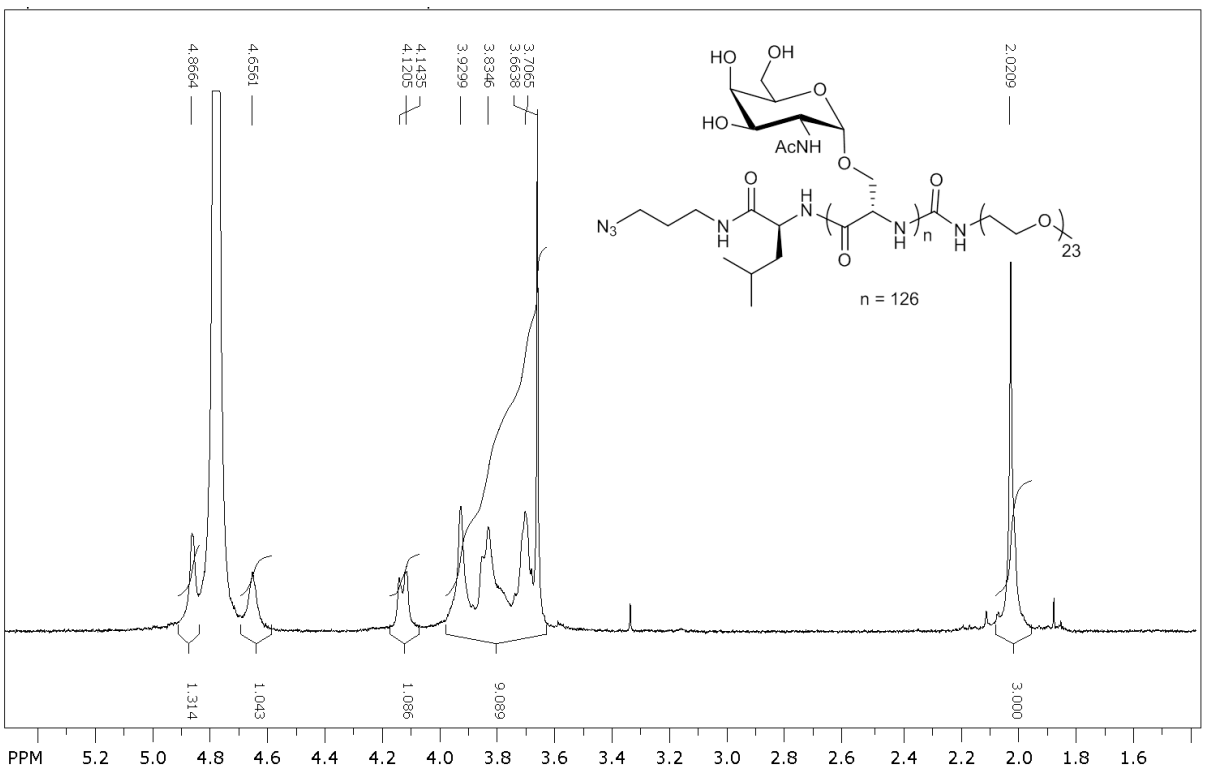
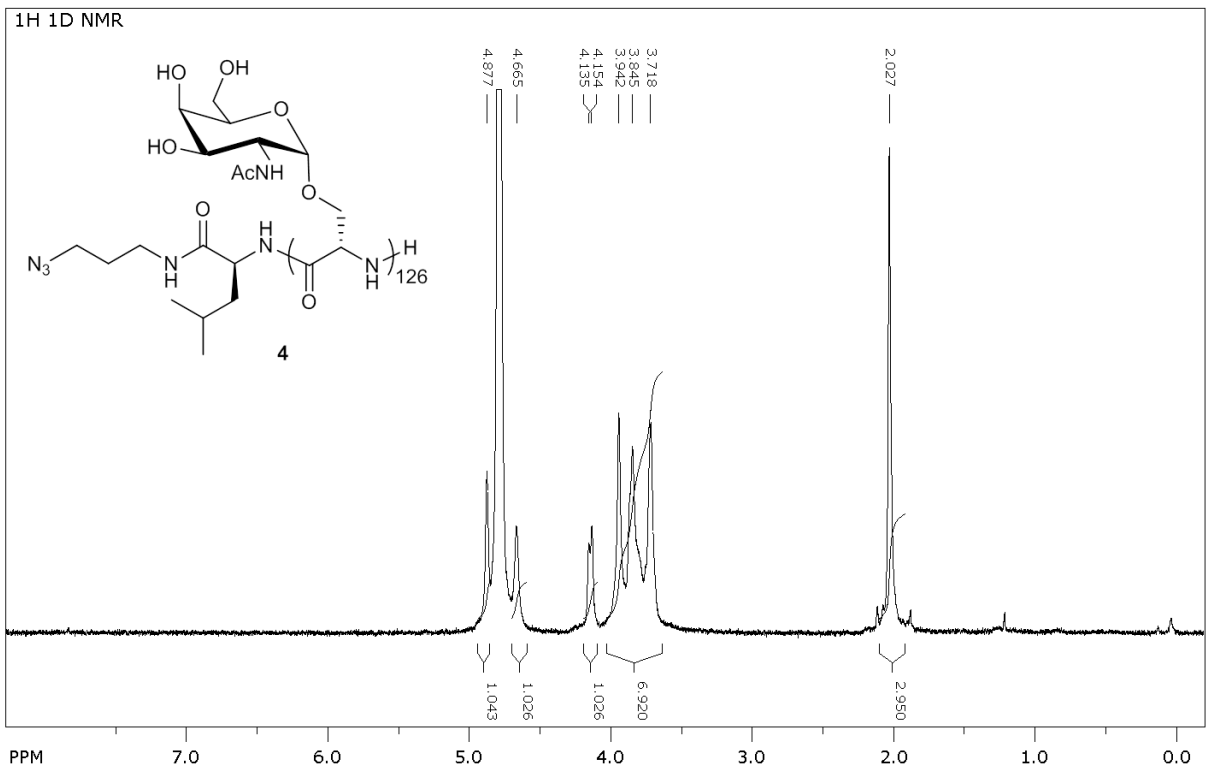
Mucin chimera is cells transfected to express EGFR-NB-eGFP, EGFR-NB-eGFP is cells transfected with *p4CMVE-U6-PylT* and *pMmPylRS-EGFR(128TAG)-eGFP* in the presence of **6**. Synth. Mucin is untreated cells incubated with 5 μ M $(\alpha$ -GalNAc-Ser)₃₀₀-Tz, Neg. Control is untreated cells, Pos. Control is cells treated with 0.3 mM Triton X-100.

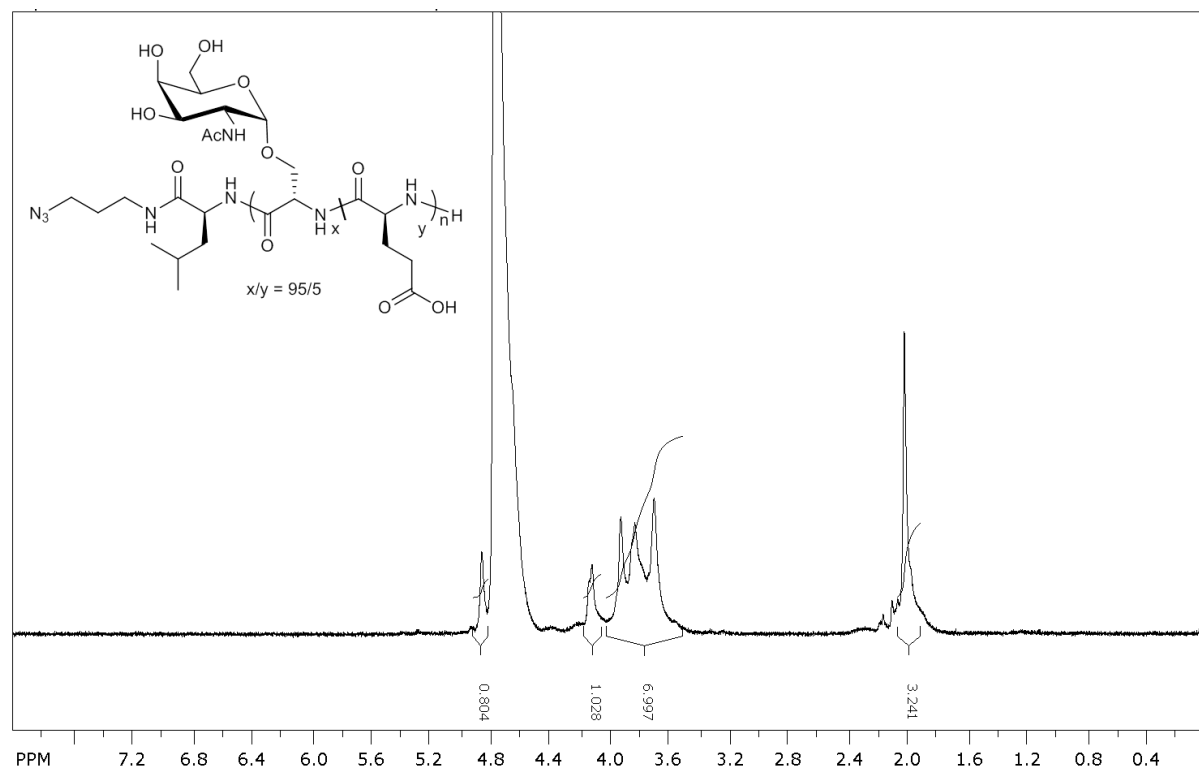
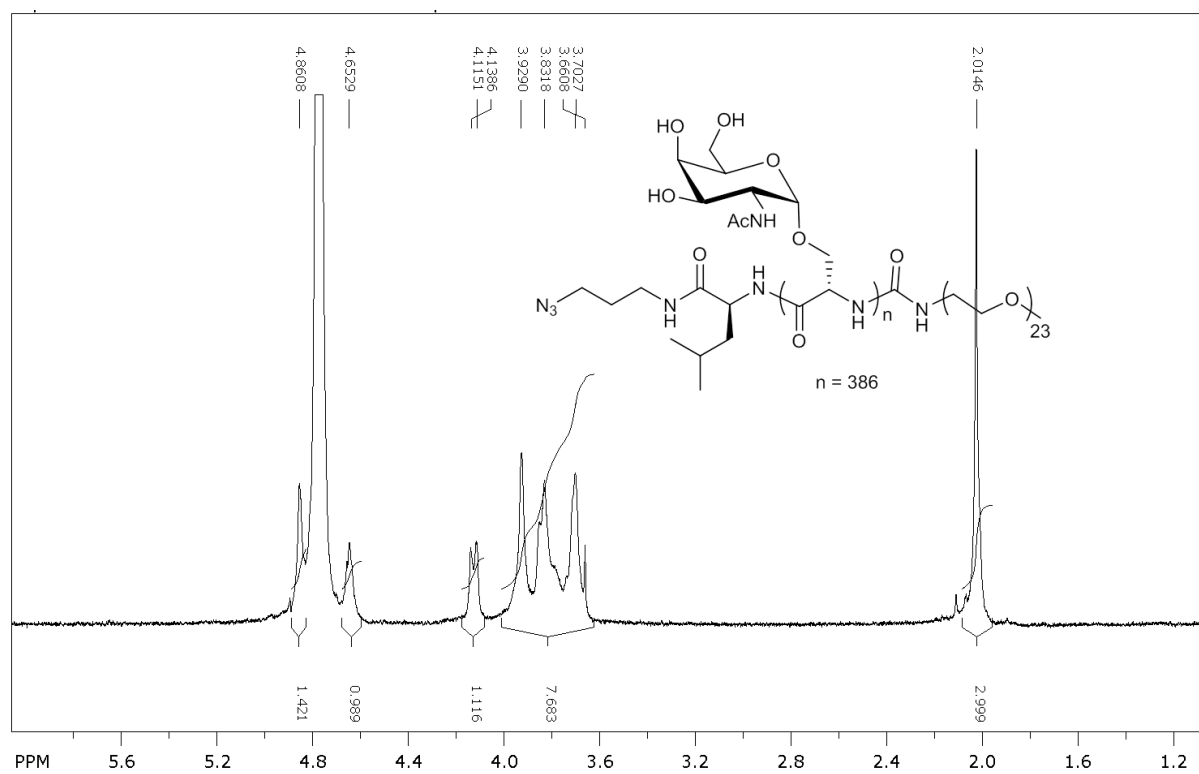
IV. Spectral Data

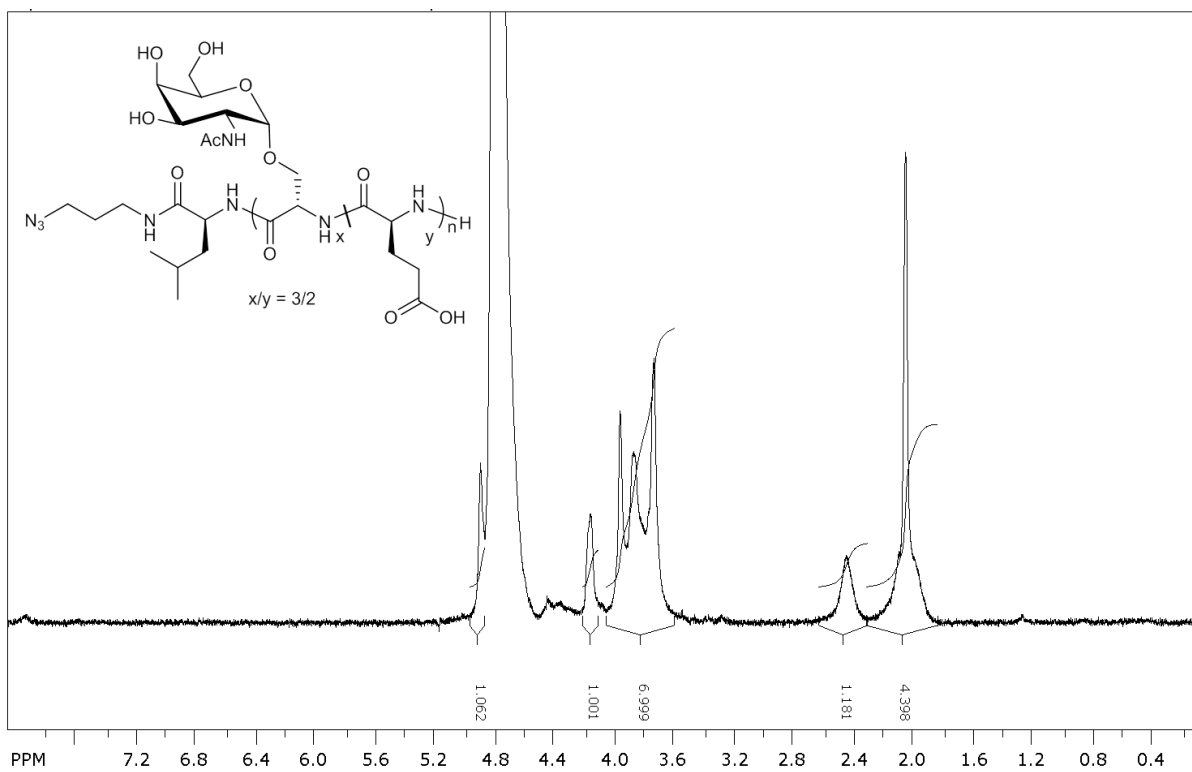
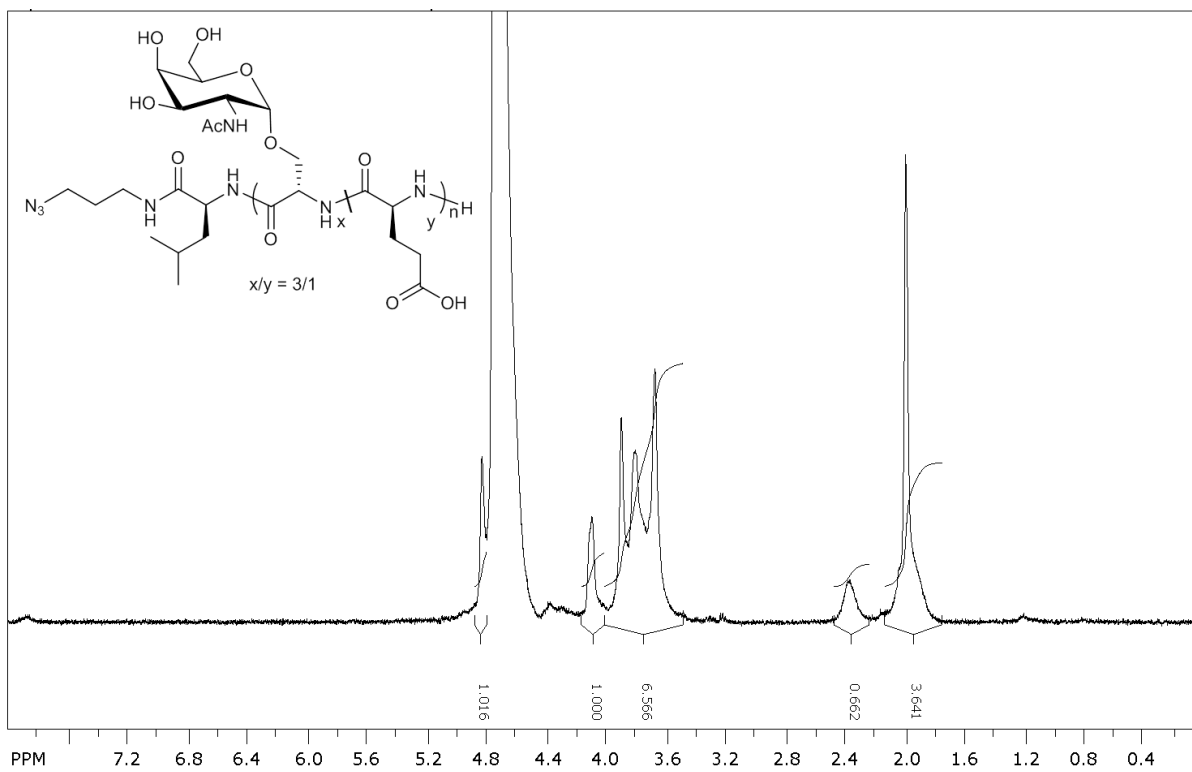


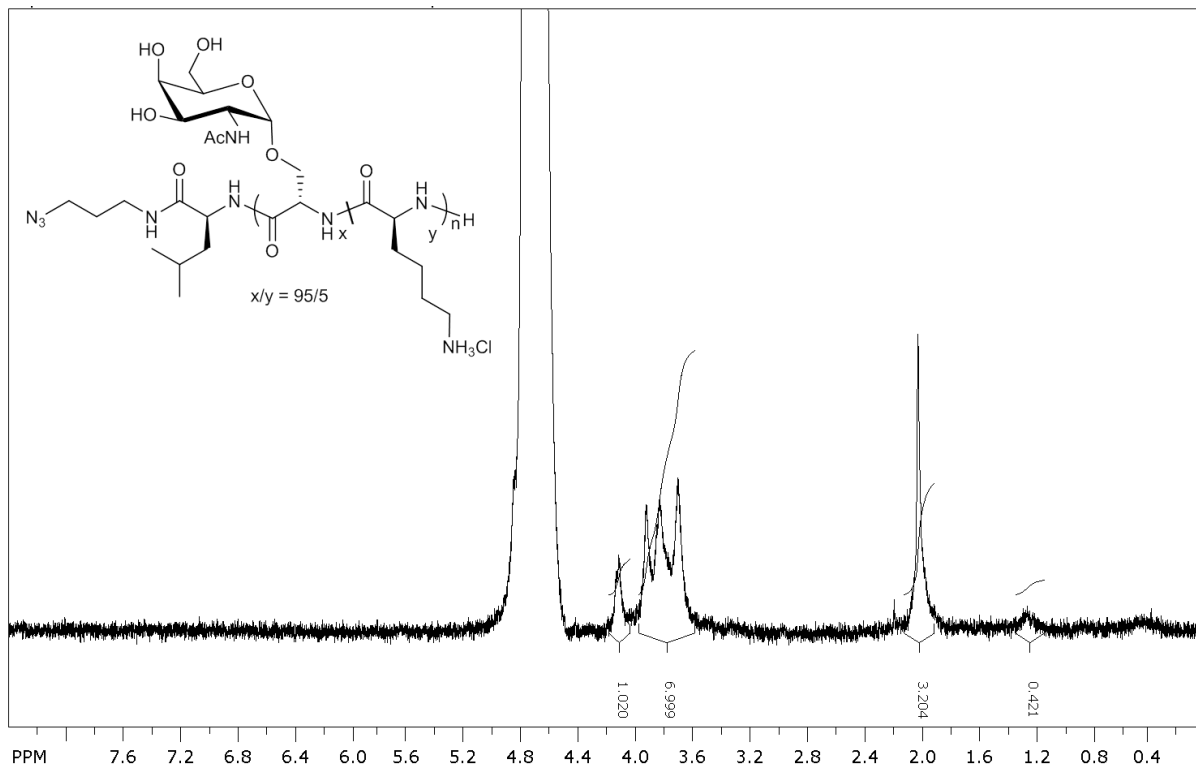
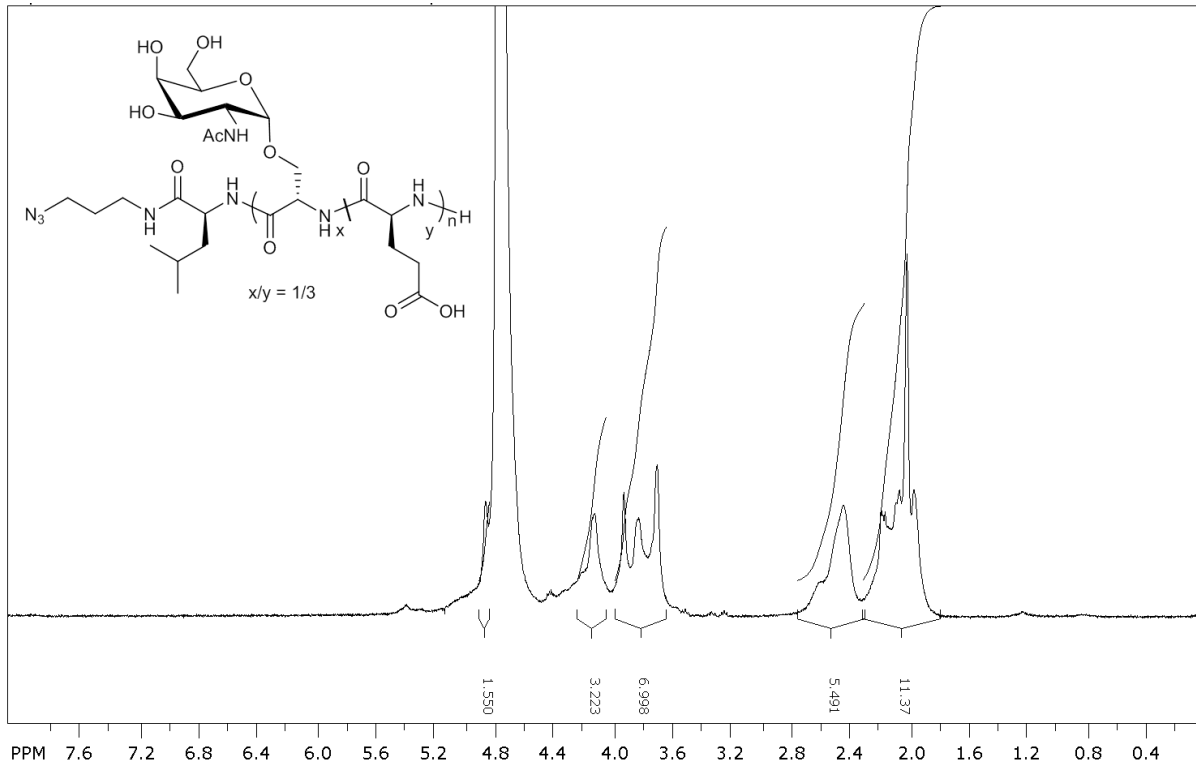


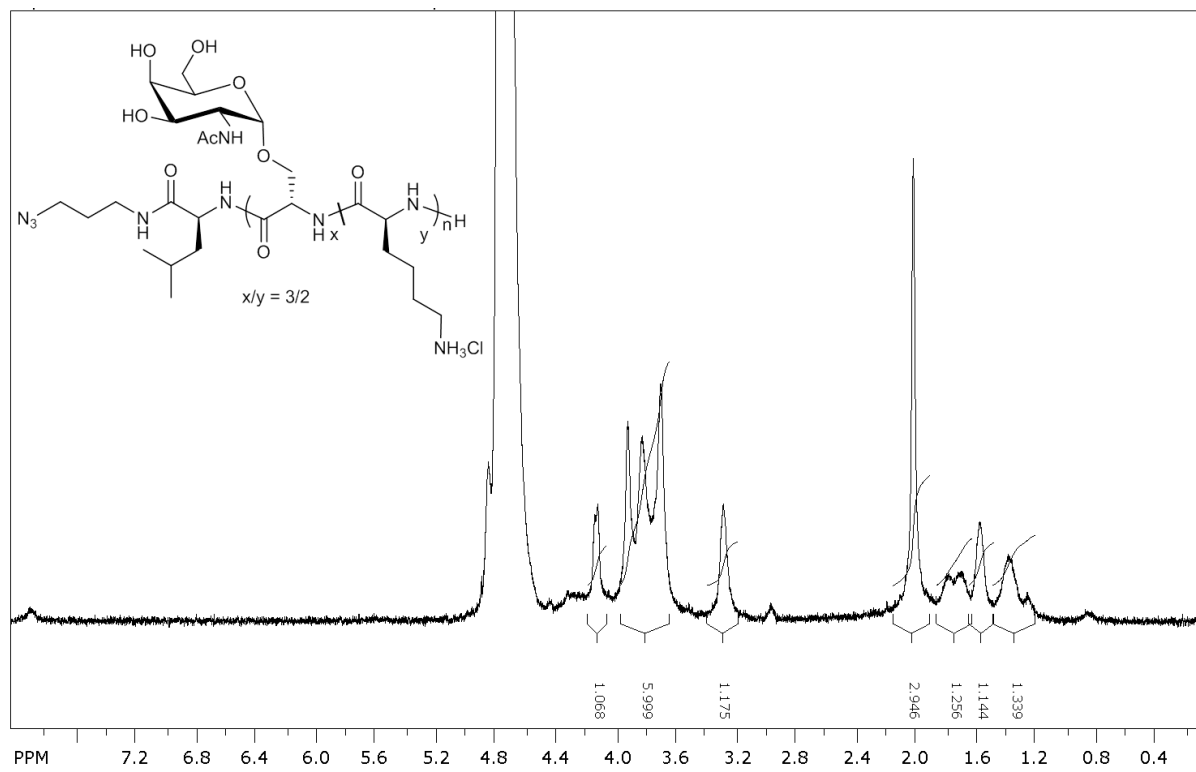
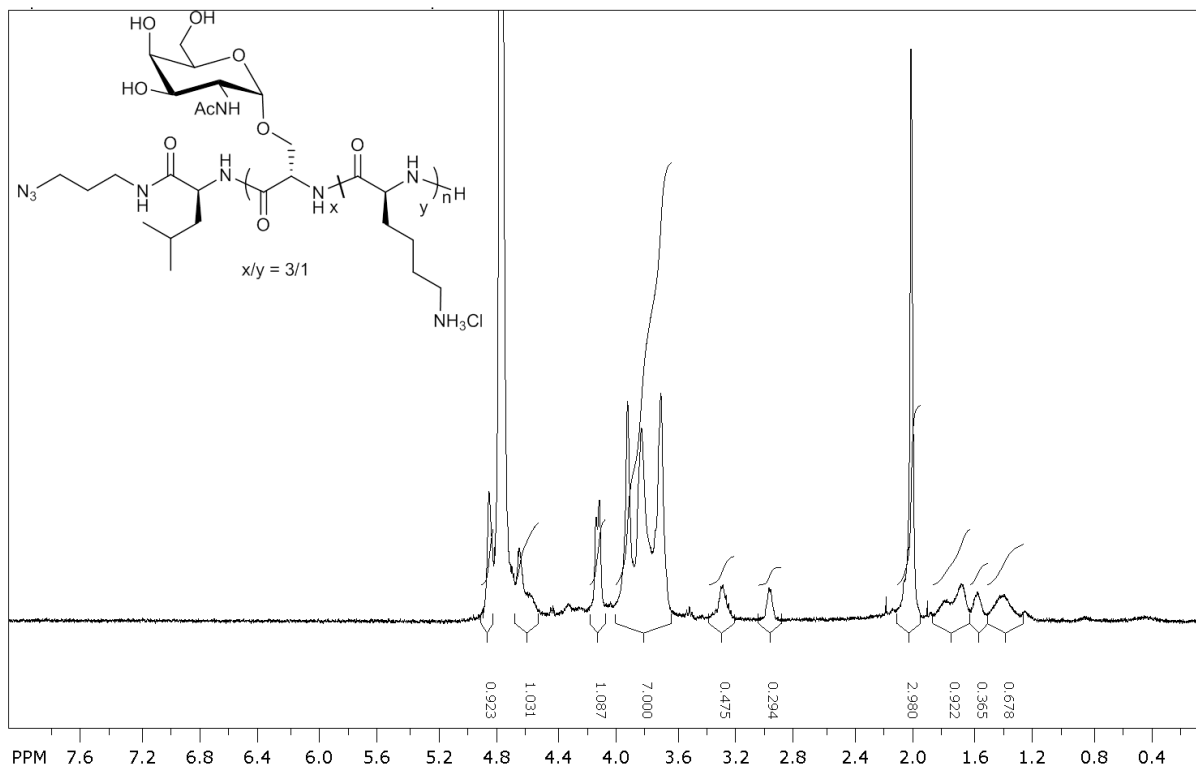


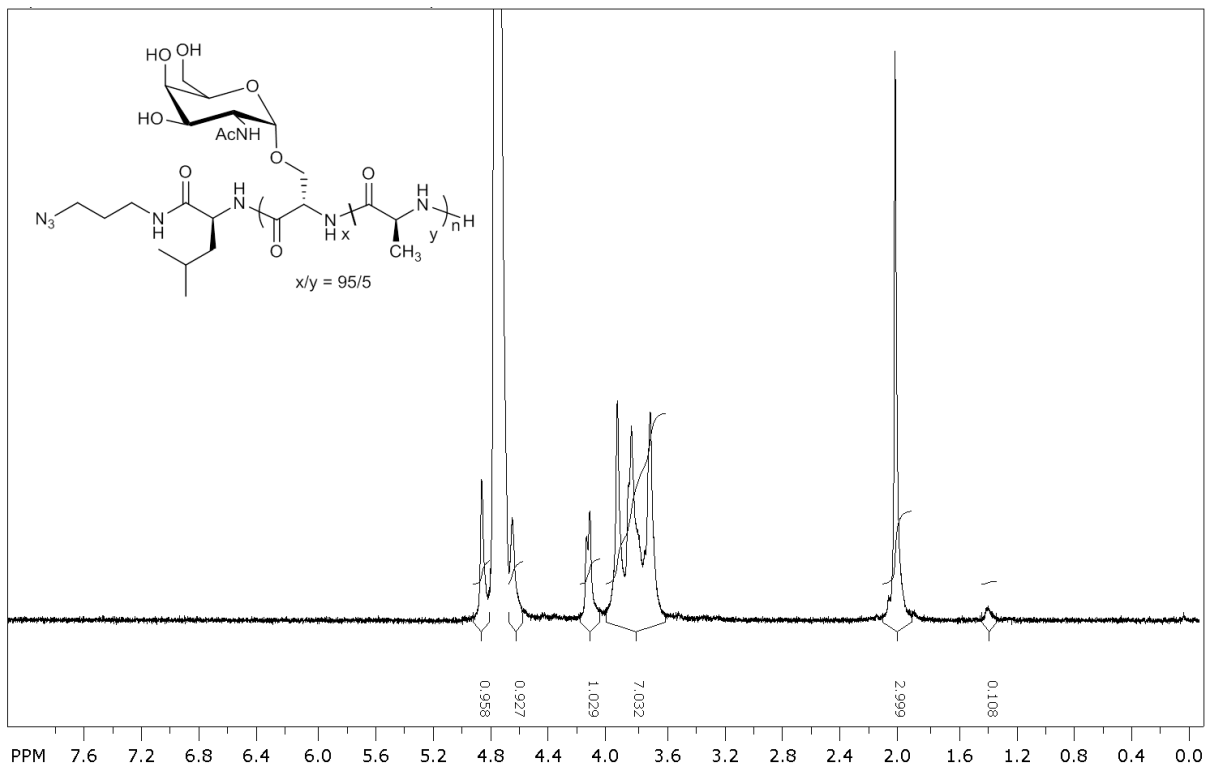
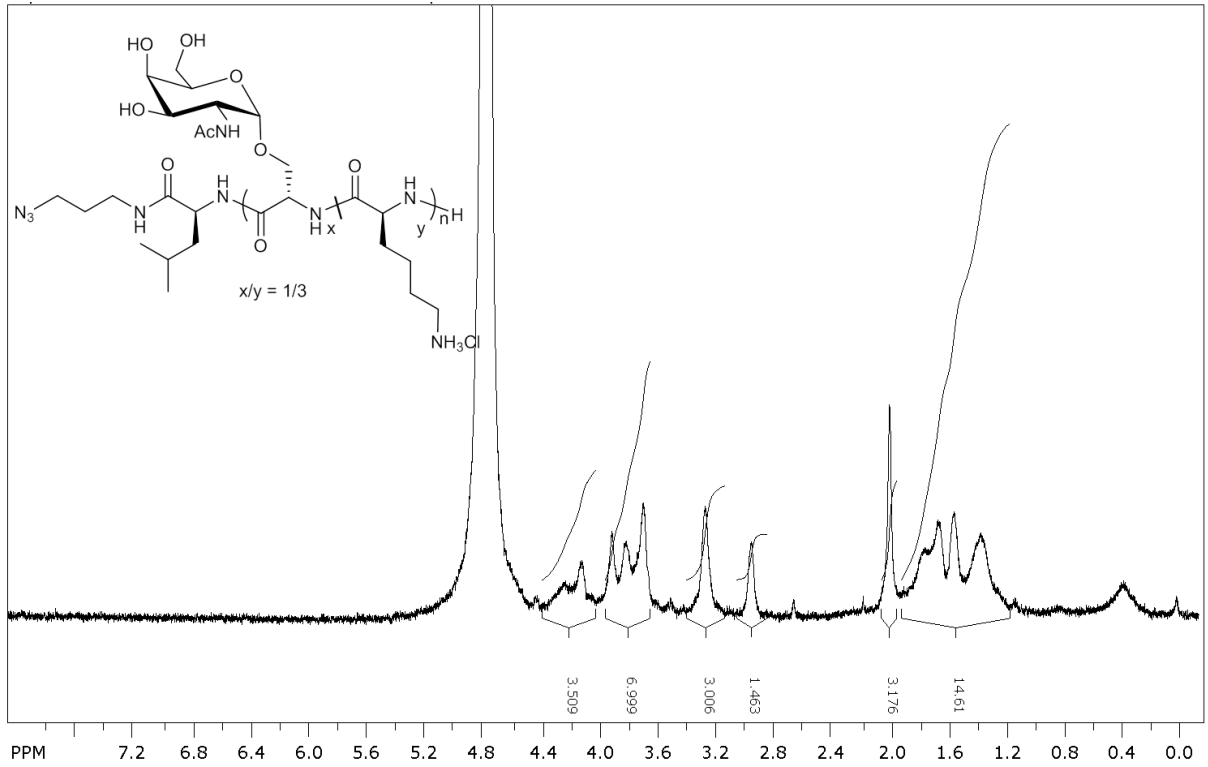


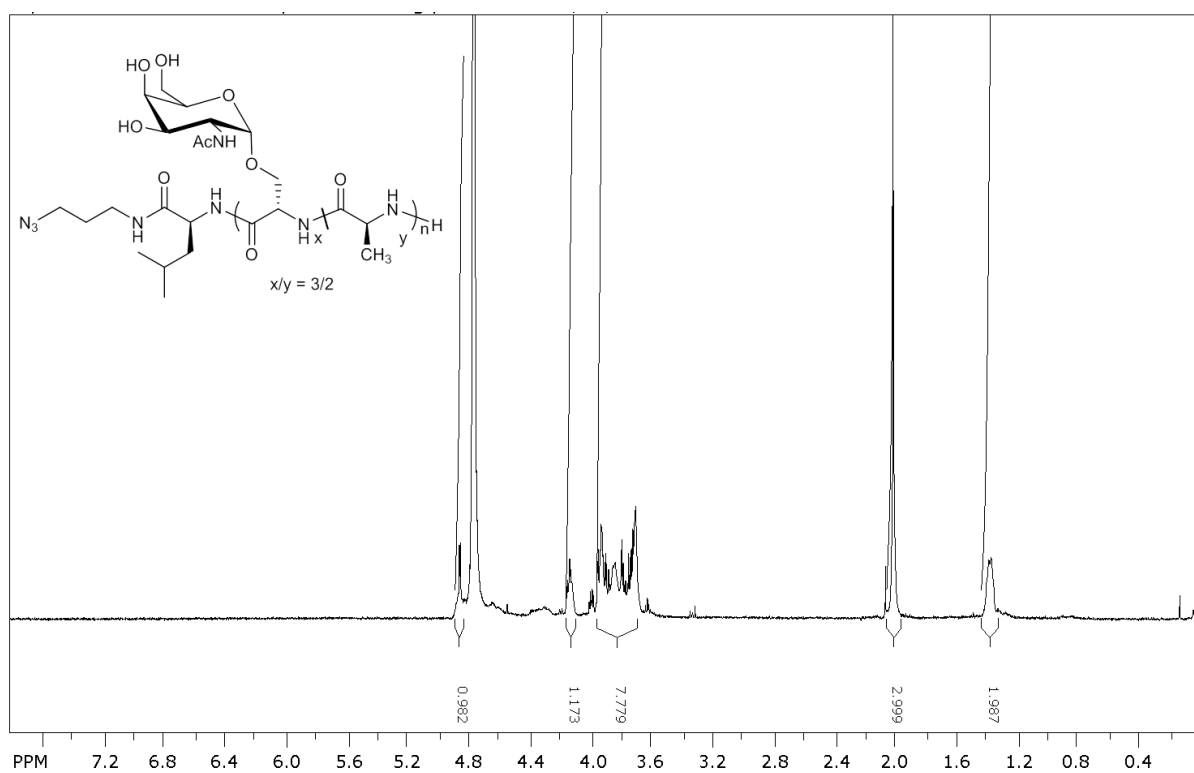
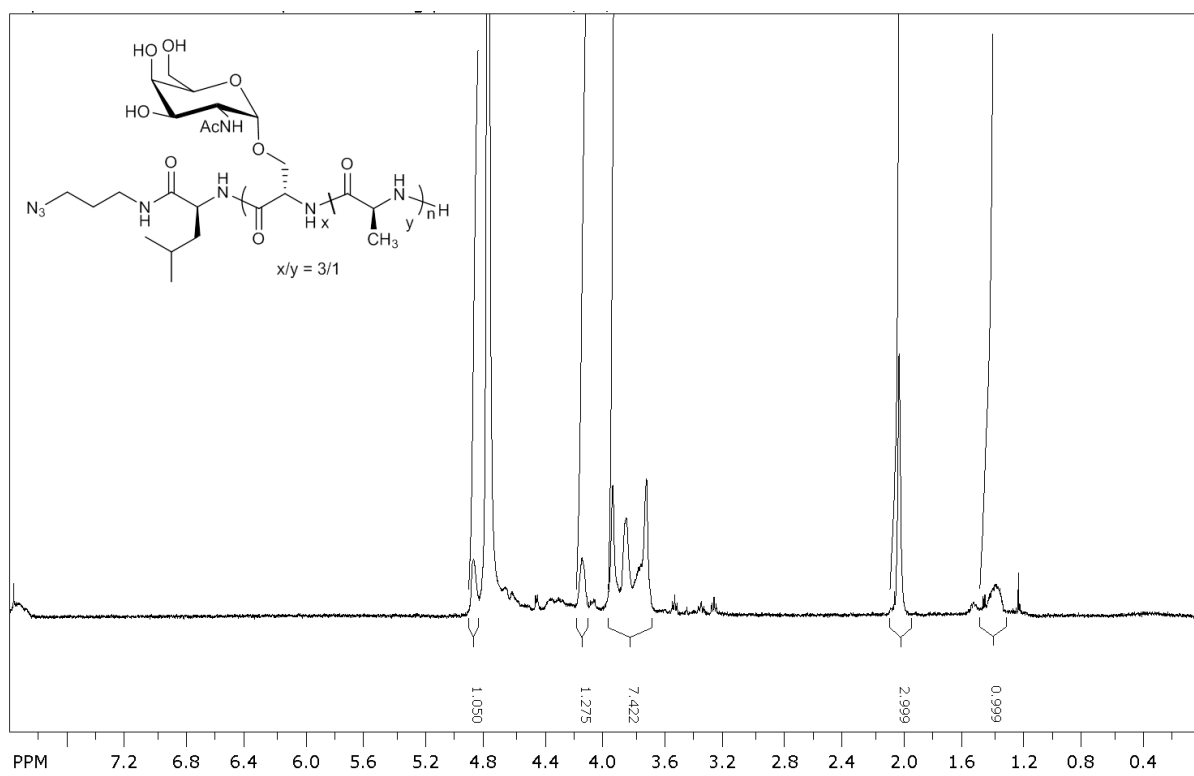


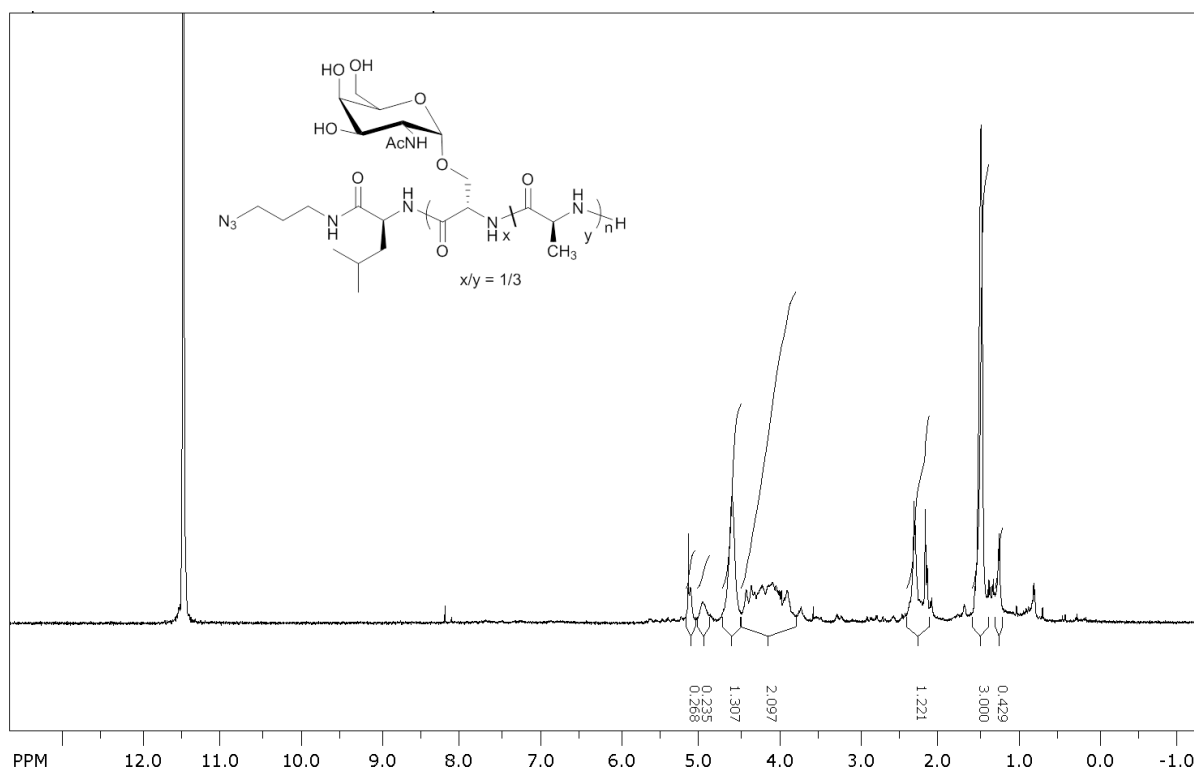












V. References

1. Winans KA, King DS, Rao VR, Bertozzi CR (1999) A chemically synthesized version of the insect antibacterial glycopeptide, dipterucin, disrupts bacterial membrane integrity. *Biochemistry* 38(36):11700–11710.
2. Martins-Teixeira MB, Campo VL, Biondo M, Sesti-Costa R, Carneiro ZA, Silva JS, Carvalho I (2013) Selective glycosylation affords mucin-related GalNAc amino acids and diketopiperazines active on *Trypanosoma cruzi*. *Bioorgan Med Chem* 21(7):1978–1987.
3. Kramer J, Deming TJ (2010) General method for purification of α -amino acid-*N*-carboxyanhydrides using flash chromatography. *Biomacromolecules* 11(12):3668–3672.
4. Campo VL, Carvalho I, Allman S, Davis BG, Field RA (2007) Chemical and chemoenzymatic synthesis of glycosyl-amino acids and glycopeptides related to *Trypanosoma cruzi* mucins. *Org Biomol Chem* 16(5):2645–2657.
5. Nagai A, Sato D, Ishikawa J, Ochiai B, Kudo H, Endo T (2004) A facile synthesis of *N*-carboxyanhydrides and poly(α -amino acid) using di-*tert*-butyltricarboxylate. *Macromolecules* 37(7):2332–2334.
6. Nguyen L-T T, Vorenkamp EJ, Daumont CJM, ten Brinke G, Schouten AJ (2010) Double-brush Langmuir–Blodgett monolayers of α -helical diblock copolypeptides. *Polymer* 51(5):1042–1055.
7. Kramer JR Deming TJ (2012) Preparation of multifunctional and multireactive polypeptides via methionine alkylation. *Biomacromolecules* 13(6):1719–1723. For TFA-Lys NCA spectral data see Guillermain C, Gallot B (2002) Homopolymers and copolymers of $N\epsilon$ -4-

-
- phenylbenzamido-L-lysine and N ϵ -trifluoroacetyl-L-lysine: Synthesis and liquid-crystalline properties. *Macromol Chem Phys* 203(10-11):1346–1356.
8. Yang J, Karver MR, Li W, Sahu S, Devaraj NK (2012) Metal-Catalyzed One-Pot Synthesis of Tetrazines Directly from aliphatic nitriles and hydrazine. *Angew Chem Int Ed* 51(21):5222 – 5225.
9. Knight JC, Richter S, Wuest M, Way JD, Wuest F (2013) Synthesis and evaluation of an 18F-labelled norbornene derivative for copper-free click chemistry reactions. *Org Biomol Chem* 11:3817–3825.
10. Hannant J, Hedley JH, Pate J, Walli A, Al-Said SAF, Galindo MA, Connolly BA, Horrocks BR, Houlton A, Pike AR (2010) Modification of DNA-templated conductive polymer nanowires via click chemistry. *Chem Commun* 46:5870–5872.
11. Brzezinska K, Deming TJ (2001) Synthesis of ABA triblock copolymers via acyclic diene metathesis polymerization and living polymerization of alpha-amino acid-N-carboxyanhydrides. *Macromolecules* 34(13):4348–4354.
12. Habraken, GJM, Wilsens KHRM, Koninga CE, Heise A (2011) Optimization of N-carboxyanhydride (NCA) polymerization by variation of reaction temperature and pressure. *Polym Chem* 2:1322–1330. Zou J, Fan J, He X, Zhang S, Wang H, Wooley KL (2013) A Facile glovebox-free strategy to significantly accelerate the syntheses of well-defined polypeptides by N-carboxyanhydride (NCA) ring-opening polymerizations. *Macromolecules* 46: 4223–4226.
13. Necas D, Klapetek P (2012) Gwyddion: an open-source software for SPM data analysis. *Central European Journal of Physics* 10:181-188.
14. Mantelli S, Muller P, Harlepp S, Maaloum M (2011) Conformational analysis and estimation of persistence length of DNA using atomic force microscopy in solution. *Soft Matter* 7:3412-3416.
15. Rivetti C, Guthold M, Bustamante C (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J Mol Biol* 264(5):919–932.
16. Koley D, Bard AJ (2010) Triton X-100 concentration effects on membrane permeability of a single HeLa cell by scanning electrochemical microscopy (SECM). *Proc Natl Acad Sci* 107(39):16783–16787.