Glycoprotein mimics with tunable functionalization through global amino acid substitution and copper click chemistry

Brian M. Seifried[1,‡], Wenjing Qi[2,‡], Yun Jung Yang[1], Danielle J. Mai[1], Wendy B. Puryear[3], Jonathan A. Runstadler[3], Guosong Chen[2], and Bradley D. Olsen[1,4,]*

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

[2]The State Key Laboratory of Molecular Engineering of Polymers and Department of Macromolecular Science, Fudan University, Shanghai 200000, China

[3]Department of Infectious Disease and Global Health, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA 01536, United States

[4]Visiting Professor, Department of Macromolecular Science, Fudan University, Shanghai 200000, China

Table of Contents

S1. Materials used for Small Molecules

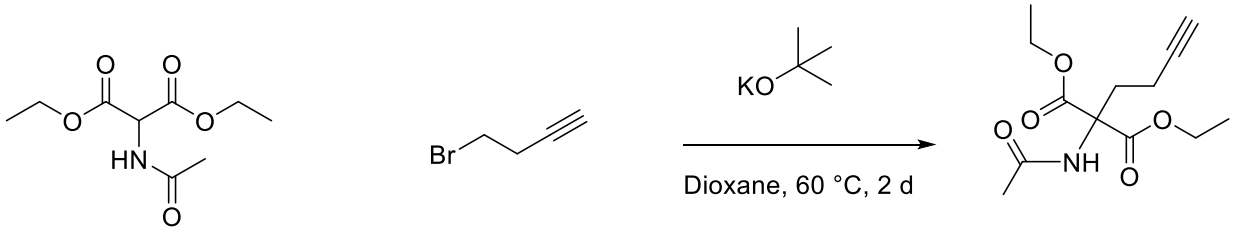
D-Mannose, D-Lactose, 6'-Sialyllactose sodium salt, and HBr (33%) in glacial acetic acid were purchased from J&K Chemical and used as received. Ion-exchange resin Dowex® 50WX2 supplied by Sigma-Aldrich Co. LLC. was washed with methanol three times before use. Sodium azide supplied by Xiya Reagent was used as received. All anhydrous solvents were distilled before use. All other chemical reagents were purchased from commercial sources (Sigma-Aldrich and VWR) and used as received unless otherwise noted.

S2. Instrumentation and Characterization of Small Molecules

NMR spectra were recorded on an AVANCE III HD 400 MHz, INOVA 500 MHz, or AVANCE 600 MHz spectrometer. The reactions were monitored and the R_f values were determined using analytical thin-layer chromatography (TLC). The TLC plates were visualized by immersion into 5% sulfuric acid solution in ethanol followed by heating using a heat gun. The residual undeuterated solvent peaks were used as references. The following abbreviations are used to denote the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. Coupling constants J are reported in Hertz (Hz).

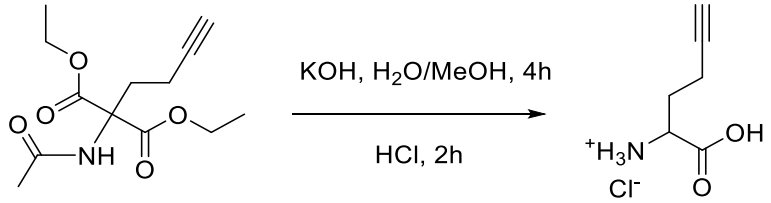
S3. Synthetic Procedures D,L-2-amino-5-hexynoic acid (HPG)

Synthesis of Diethyl α-acetamido-α-homopropargylmalonate



The procedure was adapted from the literature.(*1*) Diethylacetamidomalonate (24.69 g, 1.5 equiv, 113.6 mmol) was dissolved in dry dioxane (740 mL) under N_2 at room temperature. Potassium tert-butoxide (9.35 g, 1.0 equiv, 83.5 mmol) was added under vigorous stirring. The reaction mixture was heated to 60 °C and allowed to stir for 2 hours. *t*-BuOK(10 g, 1 equiv) was added via syringe, and the mixture was heated to reflux for 48 hours. The suspension was filtered, and the solids were washed with ether (3 × 50 mL). The solution was concentrated to yield a yellow oil via rotary evaporation. Silica column chromatography using 40% ethyl acetate in hexanes (Rf = 0.58) was performed to isolate pale yellow crystals. (5.5 g, 27% yield) ^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.84 (d, *J* = 11.1 Hz, 1H), 4.30 – 4.09 (m, 4H, H-1), 2.59 (q, *J* = 7.2, 6.2 Hz, 2H, H-2), 2.11 (td, *J* = 7.3, 2.7 Hz, 2H, H-3), 2.05 (m, 3H, H-4) 1.96 – 1.88 (m, 1H, H-5), 1.24 (td, *J* = 7.6, 7.1, 4.4 Hz, 6H). LRMS (ESI) m/z calculated for $C_{13}H_{19}NO_5$ $[M-H]^-$ 270.3, found 270.1. The ^{1}H NMR spectrum is presented in Figure S10.

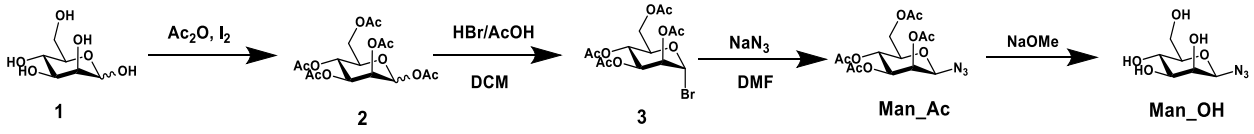
Synthesis of D,L-2-amino-5-hexynoic acid (**HPG**)



Protocols were adapted from the literature.(*2, 3*) Diethyl α-acetamido-α-homopropargylmalonate (7.46 g, 27.7 mmol) was dissolved in a mixture of MeOH/H_2O (5:1) (666 mL). KOH (3.10 g, 55.4 mmol) was added slowly. The mixture was refluxed for 4 hours. The solvents were evaporated to dryness, and the residue was redissolved in ethyl acetate. 2 N HCl was added to the mixture to adjust to pH = 1. The aqueous layer was extracted with ethyl acetate. The organic layers were combined and dried with sodium sulfate. The solvents were evaporated, and the crude compound was recrystallized in ethyl acetate. The crystals were dissolved in 2 N HCl (100 mL) and refluxed for 2 hours. The solvents were evaporated off, and the residue was recrystallized in acetone to yield the hydrochloride (1.7 g, 48%). ^{1}H NMR (500 MHz, Deuterium Oxide) δ 4.15 (dd, J = 7.2, 6.0 Hz, 1H, H-1), 2.53 – 2.38 (m, 3H, H-2), 2.22 (dtd, J = 13.0, 7.0, 5.9 Hz, 1H, H-3), 2.10 (dq, J = 14.2, 7.0 Hz, 1H, H-4). The ^{1}H NMR spectrum is presented in Figure S11.

S4. Synthetic Procedures for Sugar Azides

Azido monosaccharides, disaccharides, and trisaccharides with different protecting groups were prepared by reaction with sodium azide (**Man_OH**, **Man_Ac**, **Man_Bz**, **Lac_OH**, **Lac_Ac**, **Lac_Bz**, **SA_Ac**, **SA_OH**). Monosaccharide and disaccharide azides were prepared by a displacement of acetyl (Ac) or benzoyl (Bz) protected saccharide bromides. Mannose and lactose were first fully protected by Ac or Bz, followed by treatment with hydrobromic acid to produce the corresponding sugar bromides. Then the azidation was performed in DMF by the addition of sodium azide to acetylated or benzoylated saccharide bromides. The protecting groups facilitated compound isolation, while the sugar bromide intermediates increased the reaction stereoselectivity. Sugar azides with no protecting groups were prepared from sodium methoxide catalyzed acetyl removal reaction of azido saccharides, which unified the stereoselectivity and made all sugar azides into β isomers. Trisaccharide (SA) azides were prepared by modifying the methods presented in the literature.(*4, 5*) Acetyl 6'-Sialyllactose azide (**SA_Ac**) were synthesized in one-pot from DMC catalyzed 6'-sialyllactose sodium salt azidation and sequential acetylation in pyridine. The introduction of acetyl groups greatly benefited compounds isolation and the latter deprotection step unified the stereoselectivity of **SA_Ac** and **SA_OH**. Detailed synthetic procedures for each compound follow.



Scheme S1. Synthesis of Man_OH

Synthesis of Pentaacetyl-D-mannopyranose **(2)** .

The compound was synthesized as described in the literature.(*6*) D-mannose (10.00 g, 55.56 mmol) and iodine (0.22 g, 0.87 mmol) were added to acetic anhydride (100 mL, 1.06 mol). The reaction mixture was stirred at room temperature until the solid was totally dissolved. The reaction mixture was extracted with CH_2Cl_2 (3 × 100 mL) and washed with saturated $NaHCO_3$ solution and saturated NaCl solution. The organic phase was dried over anhydrous $MgSO_4$, filtered and evaporated under vacuum to yield 22 g (99%) of pentaacetyl-D-mannopyranose as a white solid. ^{1}H NMR (400 MHz $CDCl_3$) δ 6.09 – 6.05 (m, 1H , H-1), 5.35 – 5.32 (m, 2H , H-2 , H-4), 5.26 – 5.23 (m, 1H , H-3), 4.26 (dd, *J* = 12.3, 4.9 Hz, 1H , H-6a), 4.09 (dd, *J* = 12.3, 2.4 Hz, 1H , H-6b), 4.06 – 3.99 (m, 1H , H-5), 2.40 – 1.72 (m, 15H , $COCH_3$). The ^{1}H NMR spectrum is presented in Figure S12.

Synthesis of 2,3,4,6-Tetra-*O*-acetyl-α-D- mannopyranosyl bromide **(3)**

The compound was synthesized as described in the literature.(*7, 8*) Pentaacetyl-D- mannopyranose (8.00 g, 20.50 mmol) was dissolved in DCM (50 mL) and cooled to 0 °C. HBr (33%) in glacial acetic acid (22.00 mL, 120.75 mmol) was added slowly via a dropping funnel to the exclusion of moisture. The reaction mixture was warmed to room temperature over a period of 1 h under nitrogen atmosphere. After complete conversion of starting material as determined by thin-layer chromatography (TLC), the solution was diluted with DCM (50 mL) and poured into ice water. The organic layer was separated and washed with water (2 × 50 mL), sat. $NaHCO_3$ solution (50 mL), and brine (50 mL) prior to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure, and the crude product was purified by column chromatography using silica and petroleum ether/ethyl acetate gradient (*v*:*v* = 3:1→1:1). The product mannopyranosyl bromide was obtained as a colorless solid (7.59 g, 92%). The product was immediately used for the next reaction due to its reactivity.

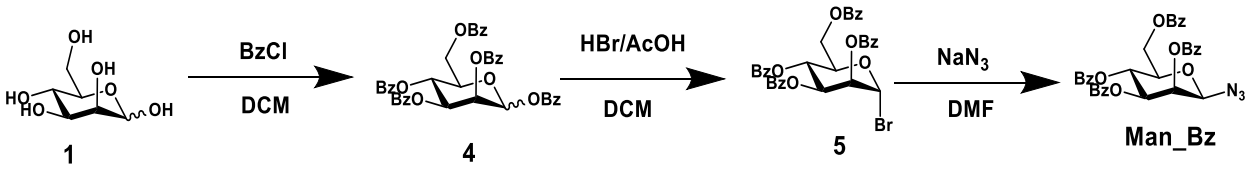
Synthesis of azido 2,3,4,6-Tetra-*O*-acetyl-β-D- mannopyranoside (**Man_Ac**)

The compound was synthesized according to the literature.(*9*) Mannopyranosyl bromide (1.96 g, 4.77 mmol) and NaN_3 (0.62 g, 9.54 mmol) were dissolved in dried DMF (30 mL) and stirred overnight at 25 °C. Solvent was removed under reduced pressure, and then DCM (50 mL) was added to the mixture. The organic layer was separated and washed with brine (50 mL) prior to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure, and the crude product was purified by column chromatography using silica and petroleum ether/ethyl acetate gradient (*v*:*v* = 3:1→1:1). The product **Man_Ac** was obtained as a colorless solid (1.44 g, 81%). ^{1}H NMR (400 MHz $CDCl_3$) δ 5.45 (dd, *J* = 3.3, 1.3 Hz, 1H, H-2), 5.27 (t, *J* = 10.0 Hz, 1H , H-4), 5.05 (dd, *J* = 10.1, 3.3 Hz, 1H, H-3), 4.74 (d, *J* = 1.3Hz, 1H , H-1), 4.29 (dd, *J* = 12.4, 5.6 Hz, 1H , H-6a), 4.21 (dd, *J* = 12.4, 2.5 Hz, 1H , H-6b), 3.77 (ddd, *J* = 10.0, 5.6, 2.6 Hz, 1H , H-5). 2.21 – 1.99 (m, 21H, $COCH_3$). 20.6, 20.7, 20.8, 20.8 (4C, s x 4, OCOCH3 x 4); 62.3 (1C, s, C6); 65.3 (1C, s, C4); 69.2 (1C, s,C3); 71.0 (1C, s, C2); 74.7 (1C, s, C5); 85.1 (1C, s, C1); 170.0, 170.0, 170.0, 169.6 (4C, s x 4, OCOCH3 x 4). The ^{1}H NMR spectrum is presented in Figure S13.

Synthesis of azido -β- mannopyranoside (**Man_OH**)

The compound was synthesized according to the literature.(*4*) **Man_Ac** (3.43 g, 5.08 mmol) was dissolved in methanol (30 mL), and sodium methoxide (3 M in methanol, 5 drops) was added. The reaction mixture was allowed to stir at room temperature for 1 h. Dowex 50WX2 H+ resin was added to adjust the pH to 7. The mixture was then filtered through Celite, and the solvent was evaporated. The product was obtained as a colorless solid (1.73 g, 90%). ^{1}H NMR (400 MHz, Deuterium Oxide): δ 4.86 (d, *J* = 1.1 Hz, 1H), 4.02 (dd, *J* = 3.1, 1.1 Hz, 1H), 3.95 (dd, *J* = 12.3,

2.3 Hz, 1H), 3.76 (dd, *J* = 12.4, 6.4 Hz, 1H), 3.66 (dd, *J* = 9.6, 3.1 Hz, 1H), 3.60 (t, *J* = 9.5 Hz, 1H), 3.48 (ddd, *J* = 9.4, 6.4, 2.2 Hz, 1H). The ^{1}H NMR spectrum is presented in Figure S14.



Scheme S2. Synthesis of Man_Bz

Synthesis of Pentabenzoyl-D-mannopyranose **(4)**

The compound was synthesized as described in the literature.(*10*) Benzoyl chloride (21.0 mL, 180 mmol) was added slowly to a suspension of mannose (5.14 g, 30.0 mol) , 4-dimethylaminopyridine (2.93 g, 24.0 mmol) and Et_3N (20.07 mL, 144 mmol) in 100 mL dichloromethane at 0 °C, then stirred at room temperature for 24 h under argon. Excess benzoyl chloride was quenched by adding MeOH with cooling. The organic layer was washed with water (3×30 mL), dried with anhydrous $MgSO_4$, and evaporated under vacuum. The crude residue was purified by flash chromatography using EtOAc/petroleum ether (1:5) as eluent to yield 15.8 g (90%) of pentabenzoyl-D-mannopyranose as yellow solid. ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.26 – 8.22 (m, 2H, Ph), 8.11 (ddt, *J* = 7.4, 1.3, 0.6 Hz, 4H , Ph), 8.03 – 7.96 (m, 2H , Ph), 7.93 – 7.85 (m, 2H , Ph), 7.78 – 7.69 (m, 1H , Ph), 7.68 – 7.63 (m, 1H , Ph), 7.63 – 7.57 (m, 3H , Ph), 7.56 – 7.52 (m, 1H , Ph), 7.51 – 7.36 (m, 7H , Ph), 7.36 – 7.29 (m, 2H , Ph), 6.65 (d, *J* = 2.0 Hz, 1H, H-1), 6.31 (t, *J* = 10.1 Hz, 1H , H-4), 6.09 (dd, *J* = 10.2, 3.3 Hz, 1H , H-3), 5.94 (dd, *J* = 3.3, 2.0 Hz, 1H , H-2), 4.72 (dd, *J* = 12.3, 2.5 Hz, 1H , H-6'), 4.60 (dt, *J* = 10.1, 3.1 Hz, 1H , H-5), 4.52 (dd, *J* = 12.3, 3.6 Hz, 1H , H-6). The ^{1}H NMR spectrum is presented in Figure S15.

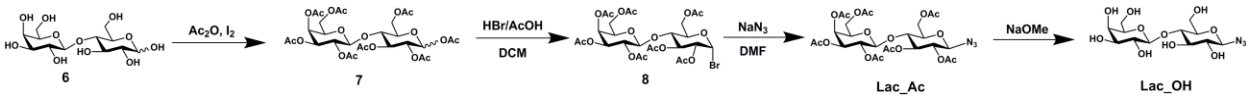
Synthesis of 2,3,4,6-Tetra-*O*-Benzoyl-α-D- mannopyranosyl bromide **(5)**

The compound was synthesized as described in the literature.(*7, 8*) Pentabenzoyl -D-mannopyranose (14.36 g, 20.50 mmol) was dissolved in DCM (50 mL) and cooled to 0 °C. HBr (33%) in glacial acetic acid (22.00 mL, 120.75 mmol) was added slowly via a dropping funnel with the exclusion of moisture. The reaction mixture was warmed to room temperature over a period of 1 h under nitrogen atmosphere. After detection of complete conversion of the starting material by TLC, the solution was diluted with DCM (50 mL) and poured into ice water. The organic layer was separated and washed with water (2 × 50 mL), sat. $NaHCO_3$ solution (50 mL), and brine (50 mL) prior to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure, and the crude product was purified by column chromatography using silica and petroleum ether/ethyl acetate gradient (*v*:*v* = 3:1→1:1). The product benzoyl mannopyranosyl bromide was obtained as a colorless solid (12.10 g, 90 %). The product was immediately used for the next reaction due to its reactivity.

Synthesis of azido 2,3,4,6-Tetra-*O*-acetyl-β-D- mannopyranoside (**Man_Bz**)

The compound was synthesized according to the literature.(*9*) Benzoyl mannopyranosyl bromide (3.14 g, 4.77 mmol) and NaN_3 (0.62 g, 9.54 mmol) were dissolved in dried DMF (30 mL) and stirred overnight at 25 °C. Solvent was removed under reduced pressure, and then DCM (50 mL) was added into the mixture. The organic layer was separated and washed with brine (50 mL) prior

to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure, and the crude product was purified by column chromatography using silica and petroleum ether/ethyl acetate gradient (*v*:*v* = 3:1→1:1). The product **Man_Bz** was obtained as a colorless solid (2.34 g, 80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.13-7.29 (m, 20H，Ph), 6.06 (t, *J* = 10.1 Hz, 1H , H-4), 5.95 (dd, *J* = 3.2, 1.2 Hz, 1H , H-2), 5.62 (dd, *J* = 10.2, 3.1 Hz, 1H , H-3), 5.08 (d, *J* = 1.2 Hz, 1H, H-1), 4.81 (dd, *J* = 12.3, 2.7 Hz, 1H , H-6'), 4.54 (dd, *J* = 12.3, 4.4 Hz, 1H , H-6), 4.24 (ddd, *J* = 9.9, 4.3, 2.7 Hz, 1H , H-5). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.1, 165.5, 165.2 (C=O), 133.6, 133.4, 133.2 (Ph), 130.0, 129.8, 128.6, 128.5, 128.4 (CH, Ph), 85.5 (C-1), 74.7 (C-5), 71.9 (C-3), 70.0 (C-2), 66.1 (C-4), 62.5 (C-6). The ^{1}H NMR spectrum is presented in Figure S16 and the ^{13}C NMR spectrum is presented in Figure S17.



Scheme S3. Synthesis of Lac_OH

Synthesis of azido 2,2′,3,3′,4′,6,6′-hepta-*O*-acetyl-β-lactoside (**Lac_Ac**)

Compound 7 was synthesized by the same method as compound 2. The masses of saccharides and the other reagents were kept the same. The product compound 7 was obtained as a colorless solid (15.8 g, 98%). Compound 8 was synthesized using the same method of compound 3. The moles of saccharides and the other reagents were kept the same. The product compound 8 was obtained as a colorless solid (13.1 g, 92%). **Lac_Ac** was synthesized with the same method as compound **Man_Ac**. The moles of saccharides and the other reagents were kept the same. The product **Lac_Ac** was obtained as a colorless solid (1.9 g, 60%).

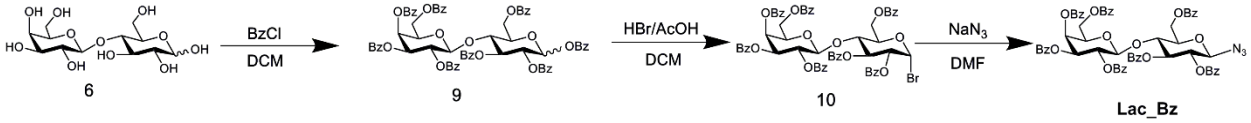
Octoacetyl-D-lactose **(7)** ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.26 (d, *J* = 3.7 Hz, 1H , H-1), 5.47 (dd, *J* = 10.3, 9.1 Hz, 1H , H-2), 5.37 (dd, *J* = 3.5, 1.2 Hz, 1H, H-1'), 5.13 (dd, *J* = 10.4, 7.8 Hz, 1H,H-2'), 4.99 (ddd, *J* = 17.0, 10.3, 3.6 Hz, 2H , H-3, H-3'), 4.56 – 4.39 (m, 2H, H-5, H-5'), 4.22 – 4.06 (m, 3H, H-6a, H-6b, H-6'a), 4.01 (ddd, *J* = 10.1, 4.3, 2.1 Hz, 1H , H-6'b), 3.89 (ddd, *J* = 7.5, 6.2, 1.2 Hz, 1H , H-4), 3.86 – 3.78 (m, 1H , H-4'), 2.39 – 1.88 (m, 24H). The ^{1}H NMR spectrum is presented in Figure S18.

Lac_Ac ^{1}H NMR NMR (400 MHz, $CDCl_3$) δ 5.35 (dd, *J* = 3.5, 1.2 Hz, 1H, H-4'), 5.21 (dd, *J* = 9.5, 9.0 Hz, 1H , H-3), 5.11 (dd, *J* = 10.4, 7.9 Hz, 1H , H-2'), 4.96 (dd, *J* = 10.4, 3.4 Hz, 1H , H-3'), 4.86 (dd, *J* = 9.5, 8.8 Hz, 1H , H-2), 4.63 (d, *J* = 8.8 Hz, 1H , H-1), 4.54 – 4.48 (m, 2H , H-1' , H-6a), 4.23 – 4.01 (m, 3H, H-6b, 2×H-6'), 3.93 – 3.75 (m, 2H, H-5', H-4), 3.70 (ddd, *J* = 10.0, 5.0, 2.1 Hz, 1H, H-5), 2.31 – 1.84 (m, 21H, COCH3). ^{13}C NMR (101 MHz, CDCl3) δ 170.3, 170.1, 170.0, 169.6, 169.5, 169.1, 101.1, 87.7, 75.7, 74.8, 72.5, 71.0, 70.9, 70.7, 69.0, 66.6, 61.7, 60.8, 20.8, 20.7, 20.6. The ^{1}H NMR spectrum is presented in Figure S19 and the ^{13}C NMR spectrum is presented in Figure S20.

Synthesis of azido β- Lactoside (**Lac_OH**)

Lac_OH was synthesized by the same method as compound **Man_OH**. The moles of saccharides and the other reagents were kept the same. The product **Lac_OH** was obtained as a colorless solid (1.71 g, 92%).

^{1}H NMR (400 MHz, D_2O) δ 4.46 (d, *J* = 7.8 Hz, 1H, H-1), 4.04-3.97 (m, 1H, H-1'), 3.93 (dd, *J* = 3.4, 0.7 Hz, 1H, H-6a), 3.87 – 3.76 (m, 3H, H-4, H-6b, H-6'), 3.75 – 3.73 (m, 1H, H-4'), 3.72-3.65 (m, 4H, H-3, H-3', H-5, H-5'), 3.55 (dd, *J* = 10.0, 7.7 Hz,1H, H-2), 3.35-3.29 (m, 1H, H-2'). The ^{1}H NMR spectrum is presented in Figure S21.



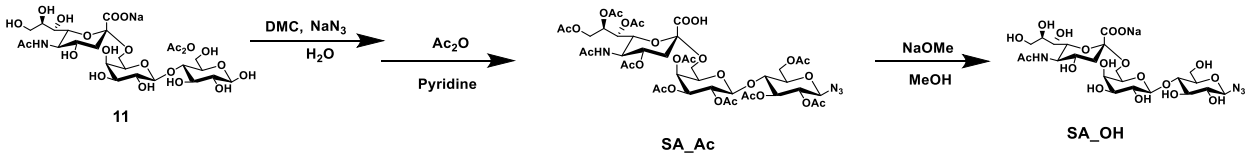
Scheme S4. Synthesis of Lac_Bz

Synthesis of azido 2,2′,3,3′,4′,6,6′-hepta-*O*-benzoyl-β-lactoside (**Lac_Bz**)

Compound 9 was synthesized by the same method as compound 4. The mass of saccharides and the other reagents were kept the same. The product compound 9 was obtained as a colorless solid (16.0 g, 91%). Compound 10 was synthesized using the same method of compound 5. The moles of saccharides and the other reagents were kept the same. The product compound 10 was obtained as a colorless solid (21.4 g, 92%). **Lac_Bz** was synthesized with the same method as compound **Man_Bz**. The product **Lac_Bz** was obtained as a colorless solid (4.8 g, 92%).

Octobenzoyl-D-lactose **(9)** ^{1}H NMR (400 MHz, $CDCl_3$): δ 6.13 (d, *J* = 8.1 Hz, 1H , H-1), 5.94 (dd, *J* = 9.6, 9.0 Hz, 1H , H-2), 5.83 – 5.69 (m, 3H , H-2', H-4', H-2), 5.37 (dd, *J* = 10.4, 3.4 Hz, 1H, H-3'), 4.88 (d, *J* = 8.0 Hz, 1H , H-1'), 4.63 – 4.47 (m, 2H, H-6), 4.45 – 4.33 (m, 1H , H-6'a), 4.12 (q, *J* = 7.2 Hz, 1H, H-6'b), 3.88 (t, *J* = 6.4 Hz, 1H , H-5), 3.84 – 3.65 (m, 2H , H-4, H-5'). The ^{1}H NMR spectrum is presented in Figure S22.

Lac_Bz ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04-7.16 (m, 35H， Ph), 5.88 – 5.80 (m, 1H), 5.80 – 5.69 (m, 2H), 5.48 – 5.38 (m, 2H), 4.90 (d, *J* = 7.9 Hz, 1H), 4.85 (d, *J* = 8.8 Hz, 1H), 4.63 (dd, *J* = 12.3, 1.9 Hz, 1H), 4.54 (dd, *J* = 12.3, 4.1 Hz, 1H), 4.29 (t, *J* = 9.5 Hz, 1H), 4.04 – 3.87 (m, 2H), 3.75 (dd, *J* = 6.7, 1.7 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 170.3, 170.1, 170.0, 169.6, 169.5, 169.1, 101.1, 87.7, 75.7, 74.8, 72.5, 71.0, 70.9, 70.7, 69.0, 66.6, 61.7, 60.8, 20.8, 20.7, 20.6. The ^{1}H NMR spectrum is presented in Figure S24 and the ^{13}C NMR spectrum is presented in Figure S25.



Scheme S5. Synthesis of SA_OH

Synthesis of Acetyl 6'-Sialyllactose azide (**SA_Ac**)

The compound was synthesized according to the literature.(*5*) 6'-Sialyllactose sodium salt (500.0 mg, 0.76 mmol), DMC (387 mg, 2.29 mmol), NaN_3 (496 mg, 7.63 mmol), and DIPEA (1.33 mL,

7.63 mmol) were mixed in 1.57 mL water and shaken at 0 °C.for 5 h and lyophilized. Then DCM (30 mL) was added, and the organic layer was separated and washed with brine (30 mL) prior to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure. Ac_2O (7.5 mL, 80 mmol) and pyridine (10 mL) were added to the residue, and the reaction mixture was stirred overnight at room temperature. The reaction was concentrated *in vaccuo*. DCM was added to the residue, which was then treated with 4 mL of 1 M HCl and washed with water prior to drying over anhydrous $MgSO_4$. After filtration, solvent was removed under reduced pressure, and the crude product was purified by column chromatography using silica and DCM/MeOH gradient (*v*:*v* = 20:1→10:1). The product **SA_Ac** was obtained as a colorless solid (400 mg, 48 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.74 (d, *J* = 9.7 Hz, 1H), 5.40 (d, *J* = 3.4 Hz, 1H), 5.35 (s, 1H), 5.19 (t, *J* = 9.5 Hz, 1H), 5.14 – 4.98 (m, 3H), 4.89 (t, *J* = 9.2 Hz, 1H), 4.63 (d, *J* = 8.7 Hz, 1H), 4.54 (dd, *J* = 22.7, 9.8 Hz, 2H), 4.32 (d, *J* = 12.0 Hz, 1H), 4.25 – 4.00 (m, 5H), 3.97 – 3.85 (m, 2H), 3.87 – 3.68 (m, 3H), 3.58 (dd, *J* = 10.8, 6.9 Hz, 2H), 2.28 – 1.86 (m, 33H). The ^{1}H NMR spectrum is presented in Figure S26.

Synthesis of Acetyl 6'-Sialyllactose azide (**SA_OH**)

The compound was synthesized according to the literature.(*4*) **SA_Ac** (100 mg, 0.09 mmol) was dissolved in methanol (10 mL), and sodium methoxide (3 M in methanol, 5 drops) was added. The reaction mixture was allowed to stir at room temperature for 2 h. Dowex 50WX2 H^+ resin was added to adjust the pH to 7. The mixture was then filtered through Celite, and the solvent was evaporated. The product was obtained as colorless solid (50 mg, 82%). ^{1}H NMR (400 MHz, D_2O): δ 4.67 (d, *J* = 8.8 Hz, 1H), 4.35 (dd, *J* = 7.9, 1H), 4.03 – 3.41 (m, 18H), 3.33 – 3.17 (m, 1H), 2.67 – 2.52 (m, 1H), 1.95 (s, *3*H), 1.72 (t, *J* = 12.3, 1H). The ^{1}H NMR spectrum is presented in Figure S27.

S5. Cloning Strategy, Peptide Sequences, and MALDI-TOF of ELPs

The gene sequence for the 50 pentapeptide methionine enriched sequence (**E50**) was purchased from Genscript in the pQE1 vector through cloning between SphI and HindIII. The highlighted restriction sites are in order SphI and HindIII. The sequence is provided below.

E50 Gene Sequence

E50 Amino Acid Sequence

MKHHHHHHQLHAGSVPGMGVPGMGVPGMGVPGMGVPGDGVPGMGVPGMGVPGMG
VPGMGVPGKGVPGMGVPGMGVPGMGVPGMGVPGDGVPGMGVPGMGVPGMGVPGM
GVPGKGVPGMGVPGMGVPGMGVPGMGVPGDGVPGMGVPGMGVPGMGVPGMGVPG
KGVPGMGVPGMGVPGMGVPGMGVPGDGVPGMGVPGMGVPGMGVPGMGVPGKGVP
GMGVPGMGVPGMGVPGMGVPGDGVPGMGVPGMGVPGMGVPGMGVPGKG

Table S1. MALDI-TOF Measurements of ELPs

	E50 [kg/mol]	**E50-MET** [kg/mol]
Predicted with 100% **HPG** Substitution	22.8	
Predicted with 100% **MET**		23.7
MALDI - TOF	23.1	24.0

S6. Additional Supporting Figures

a. MALDI-TOF MS

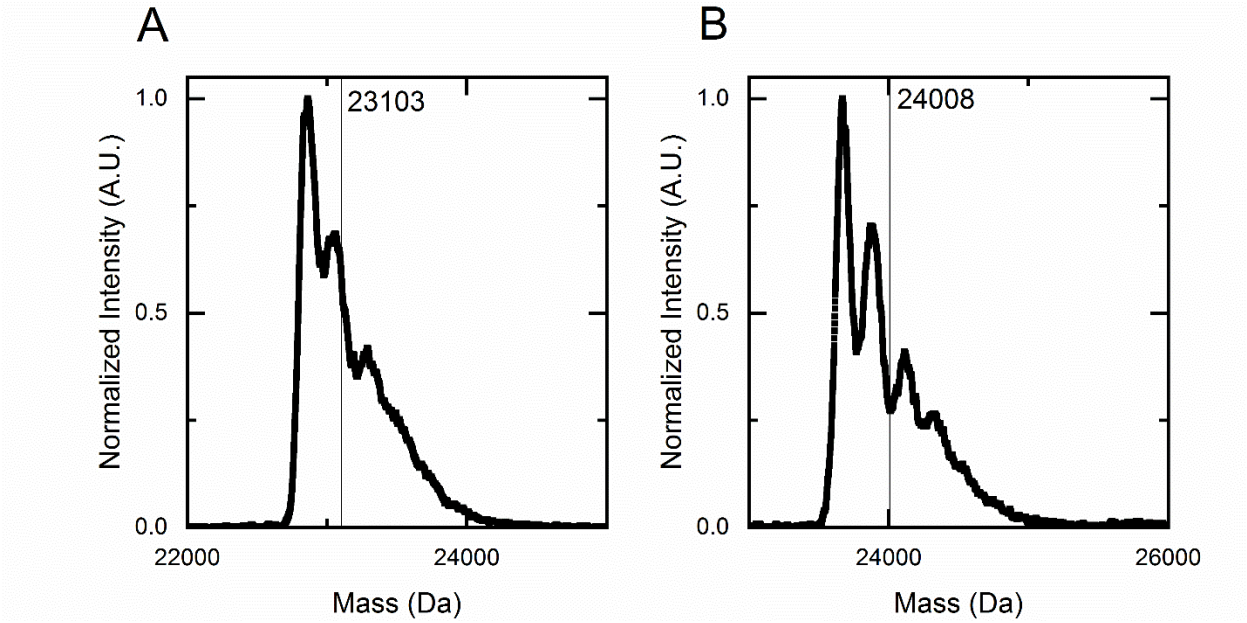


Figure S1. (A) MALDI-TOF MS of **E50** with **HPG** (positive linear mode) (B) MALDI-TOF MS of **E50-MET** (positive linear mode). The molecular weight denoted by the line is the weighted average.

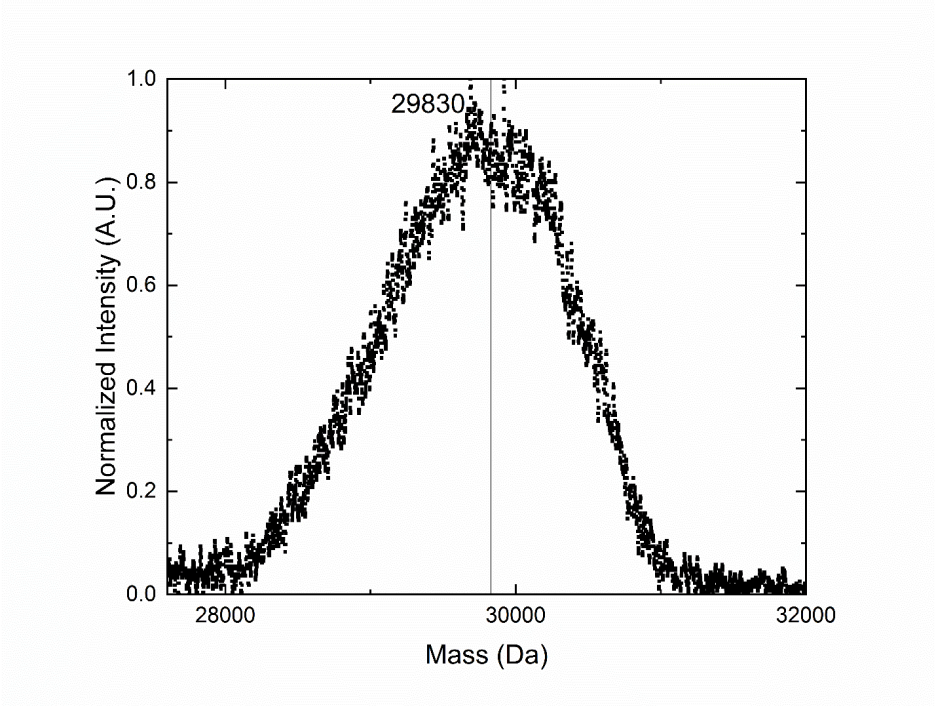


Figure S2. MALDI-TOF MS of **E50_Man_OH** (positive linear mode). The molecular weight denoted by the line is the weighted average.

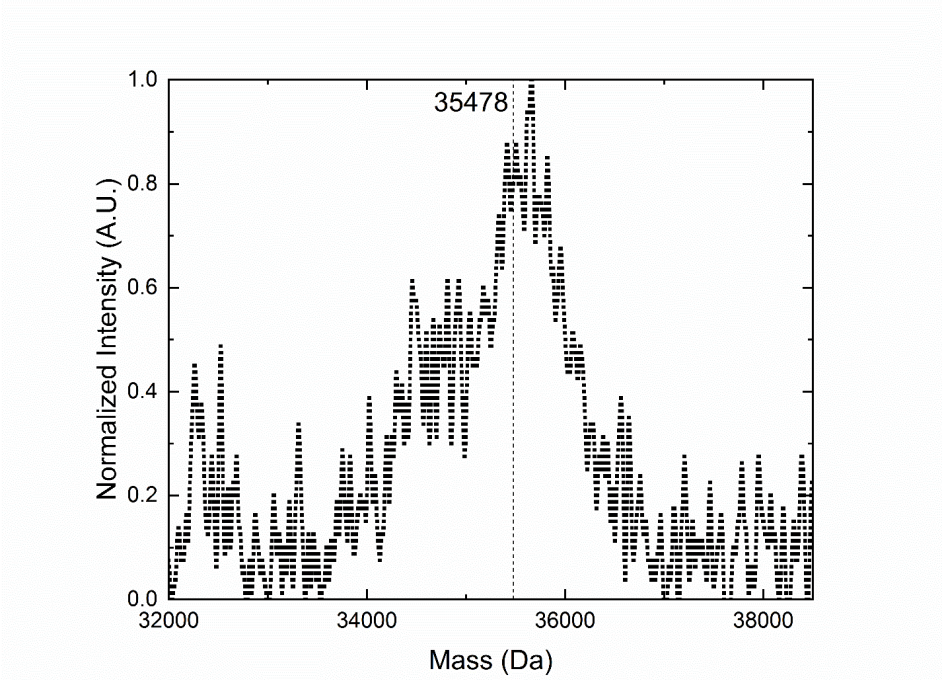


Figure S3. MALDI-TOF MS of **E50_Man_Ac**. Sample did not ionize well and exhibits poor signal to noise. (positive linear mode). The molecular weight denoted by the line is the weighted average.

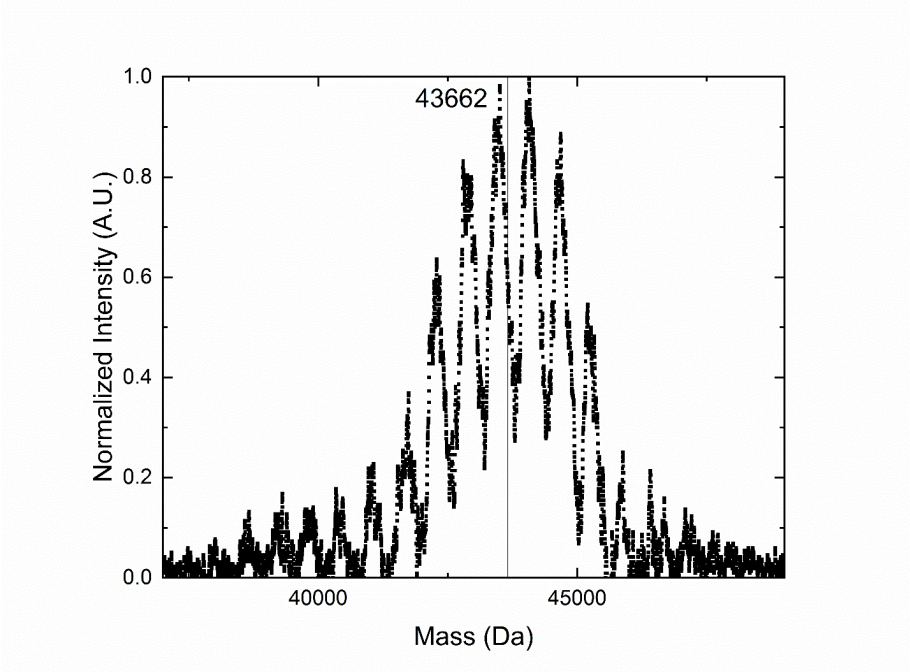


Figure S4. MALDI-TOF MS of **E50_Man_Bz** (positive linear mode). The molecular weight denoted by the line is the weighted average.

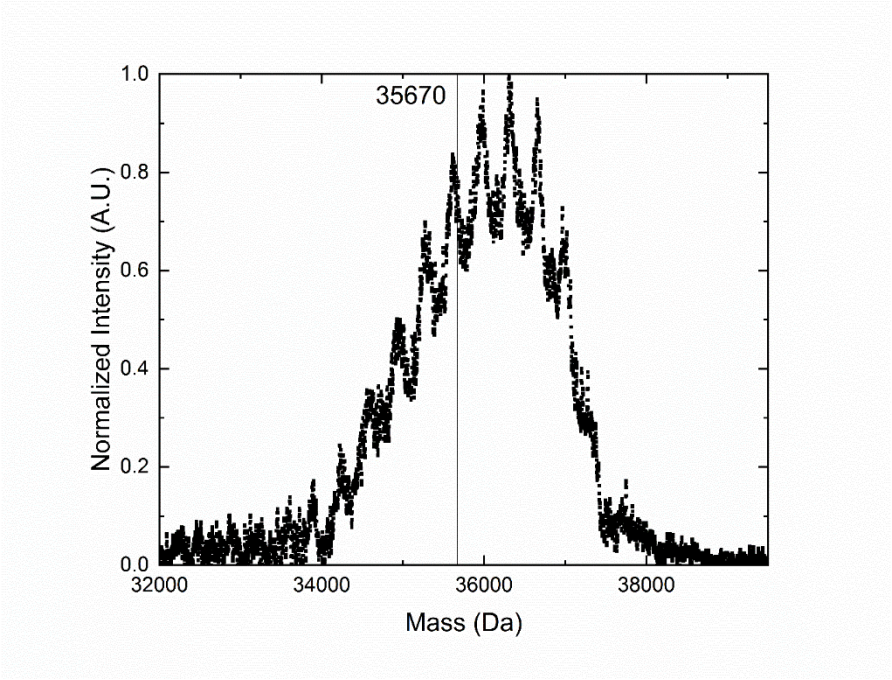


Figure S5. MALDI-TOF MS of **E50_Lac_OH** (positive linear mode). The molecular weight denoted by the line is the weighted average.

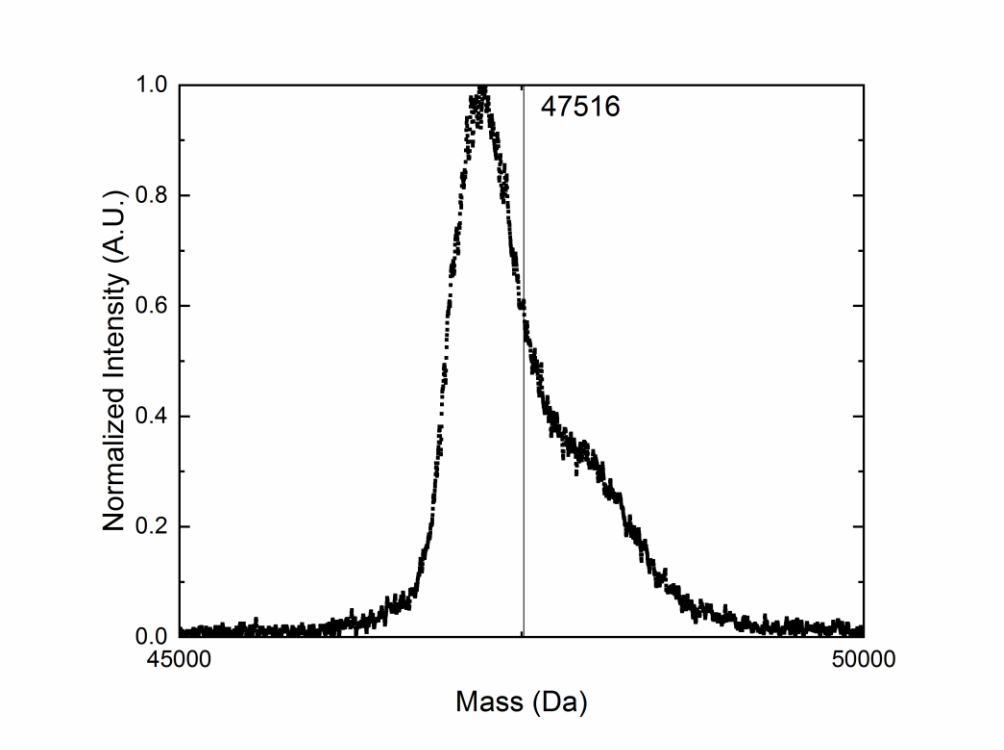


Figure S6. MALDI-TOF MS of **E50_Lac_Ac** (positive linear mode). The molecular weight denoted by the line is the weighted average.

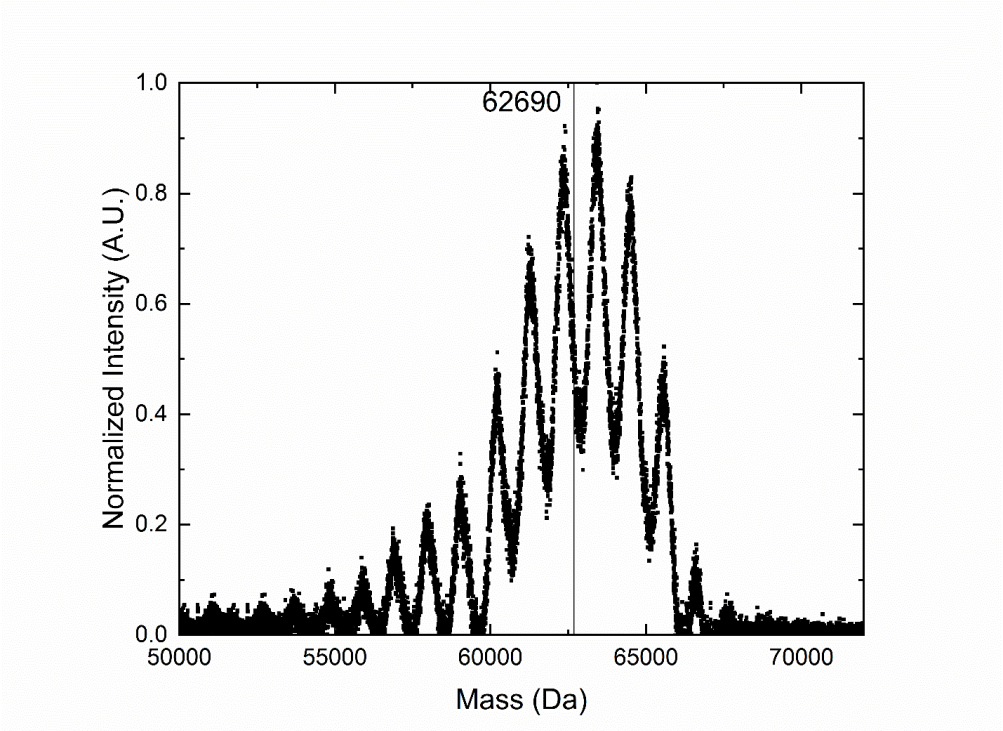


Figure S7. MALDI-TOF MS of **E50_Lac_Bz** (positive linear mode). The molecular weight denoted by the line is the weighted average.

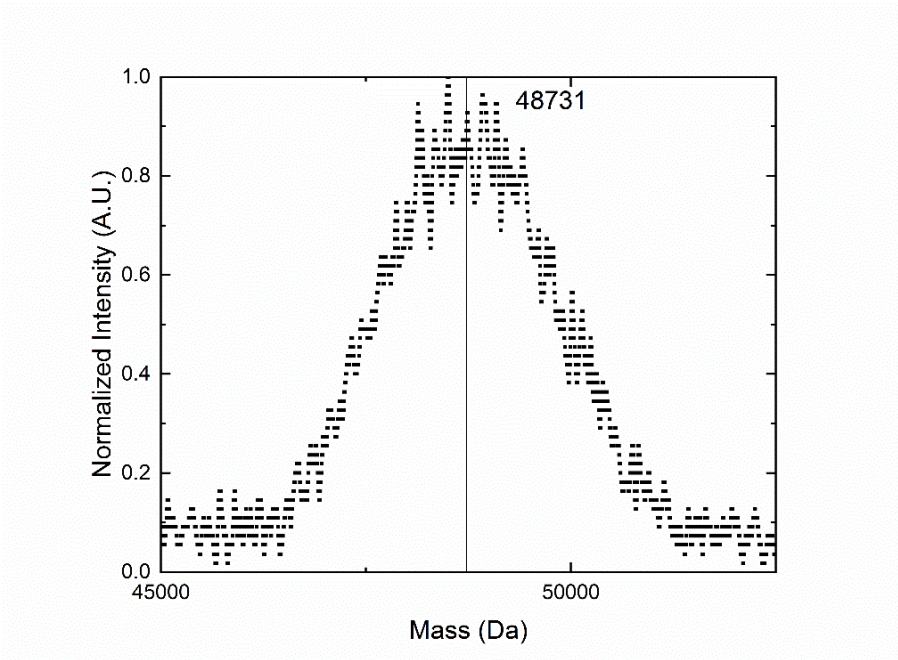


Figure S8. MALDI-TOF MS of **E50_SA_OH** (negative linear mode). The molecular weight denoted by the line is the weighted average.

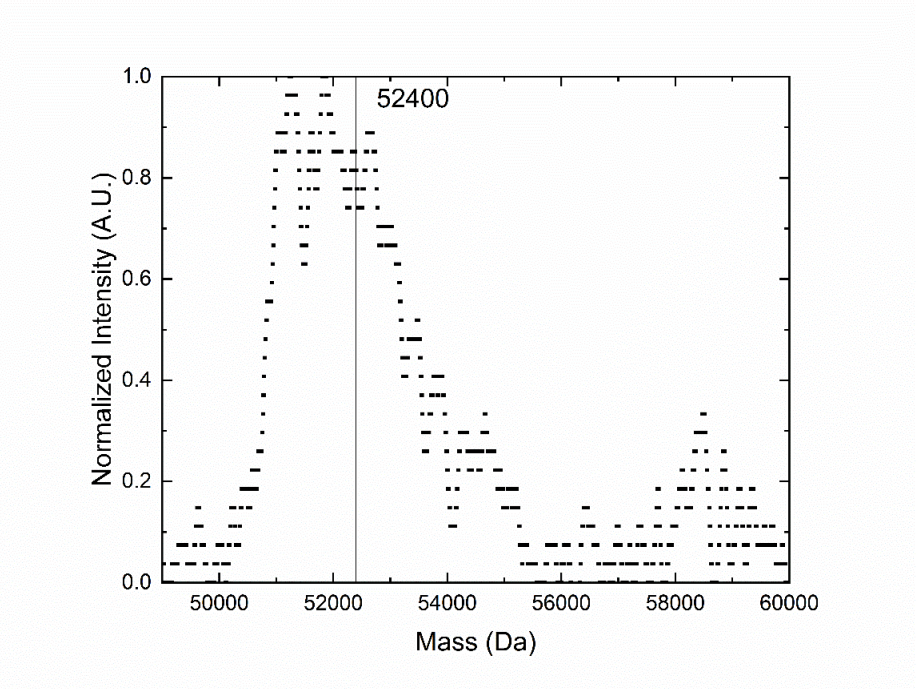


Figure S9. MALDI-TOF MS of **E50_SA_Ac**. Signals below 48000 Da were suppressed to improve signal to noise. Sample did not ionize well. (Positive linear mode). The molecular weight denoted by the line is the weighted average.

b. Small Molecule NMR spectra

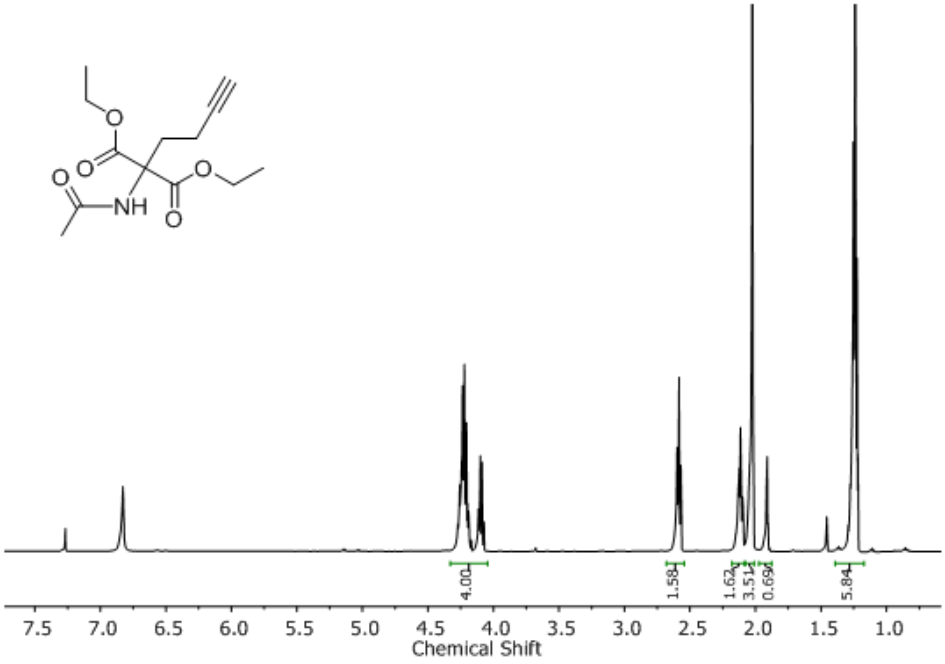


Figure S10. ^{1}H NMR spectrum of diethyl α-acetamido-α-homopropargylmalonate in $CDCl_3$ (500 MHz).

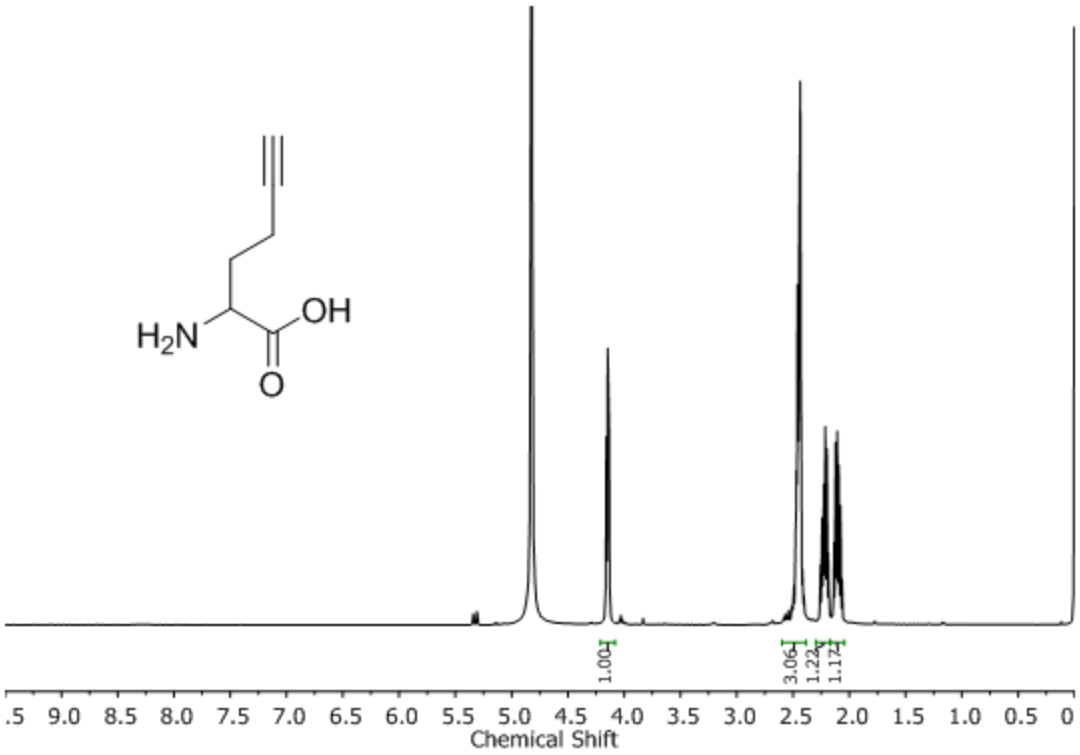


Figure S11. ^{1}H NMR spectrum of **HPG** in D_2O (500 MHz).

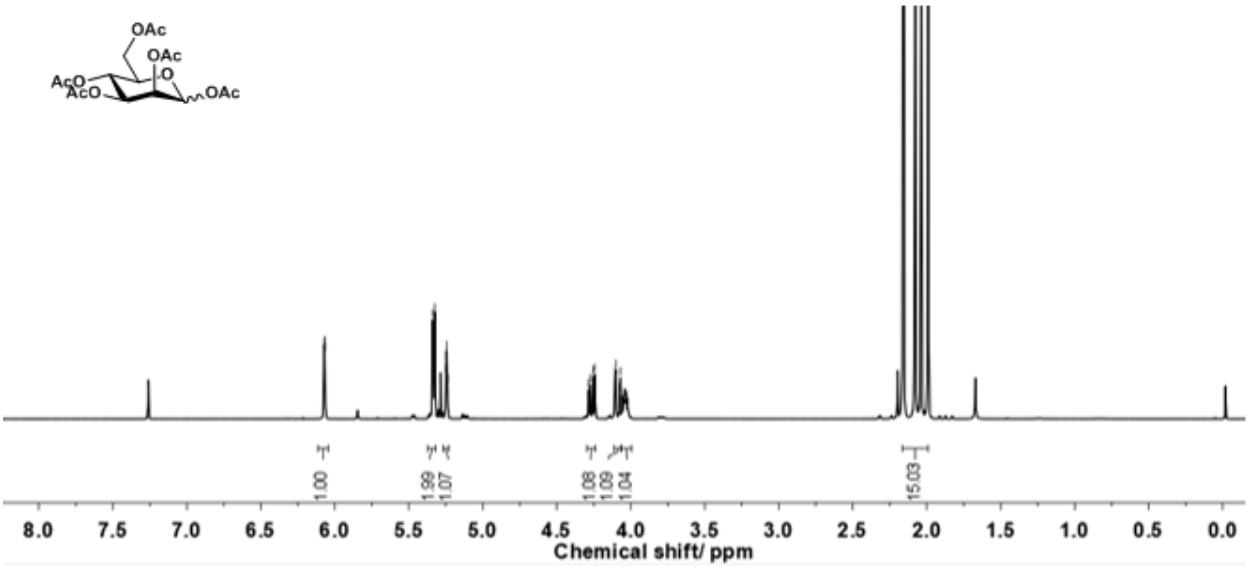


Figure S12. ^{1}H NMR spectrum of pentaacetyl-D-mannopyranose **(2)** in $CDCl_3$ (400 MHz).

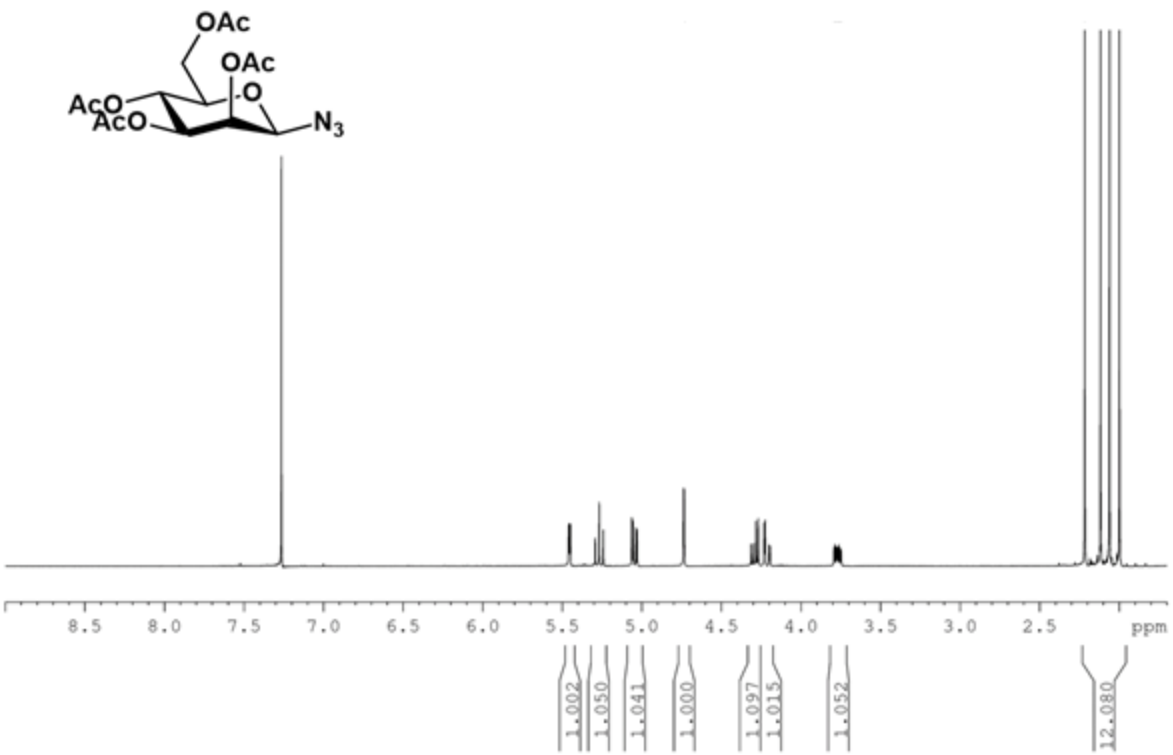


Figure S13. ^{1}H NMR spectrum of **Man_Ac** in $CDCl_3$ (400 MHz).

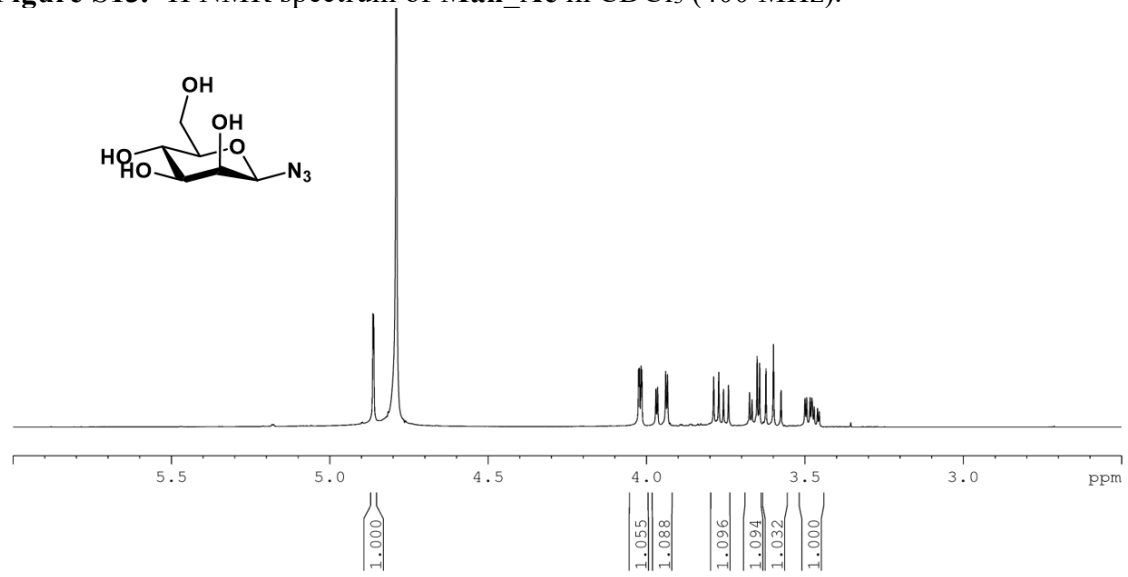


Figure S14. ^{1}H NMR spectrum of **Man_OH** in D_2O (400 MHz).

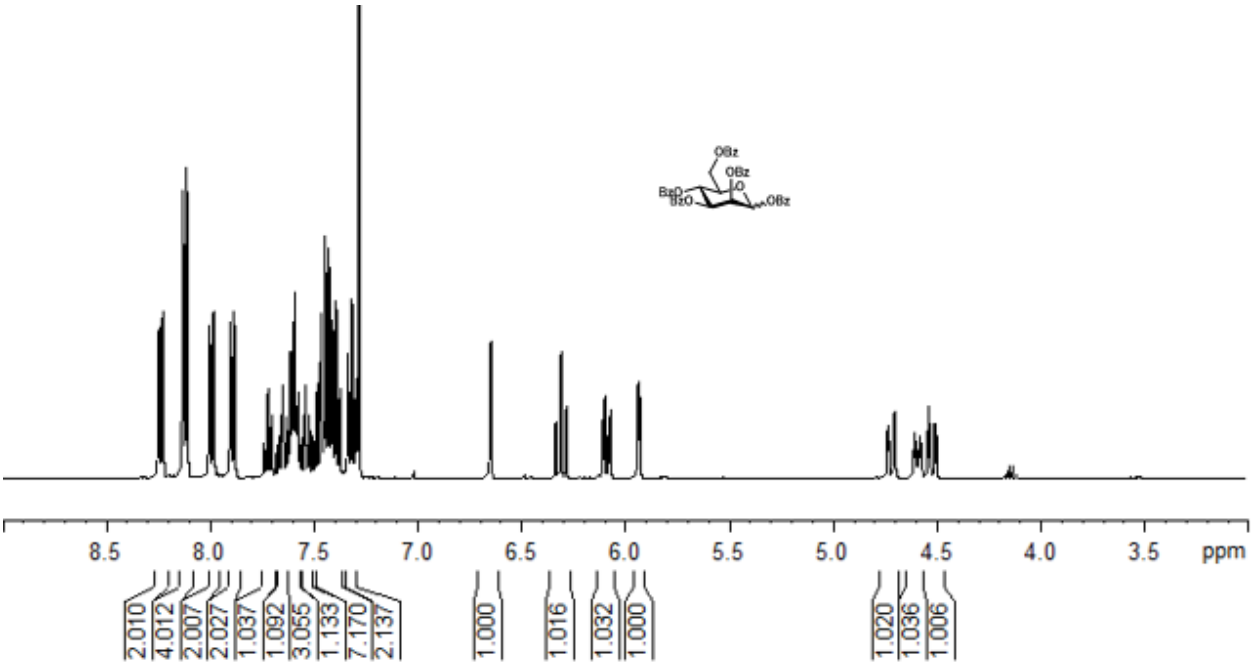


Figure S15. ^{1}H NMR spectrum of pentabenzoyl-D-mannopyranose **(4)** in $CDCl_3$ (400 MHz).

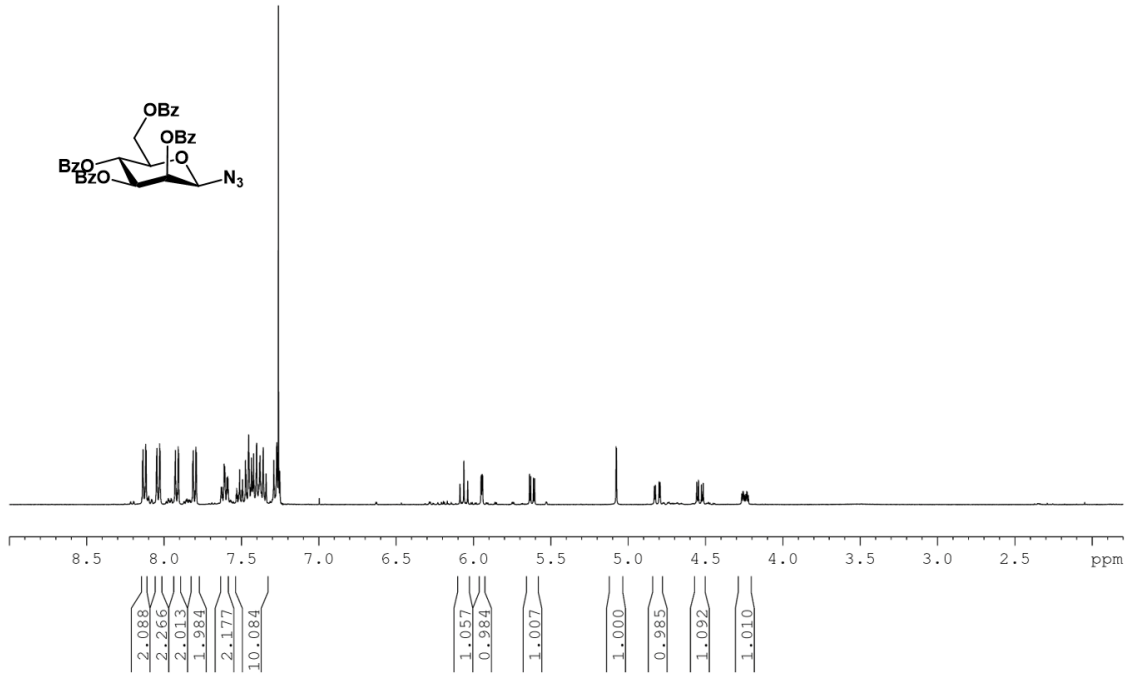


Figure S16. ^{1}H NMR spectrum of **Man_Bz** in $CDCl_3$ (400 MHz).

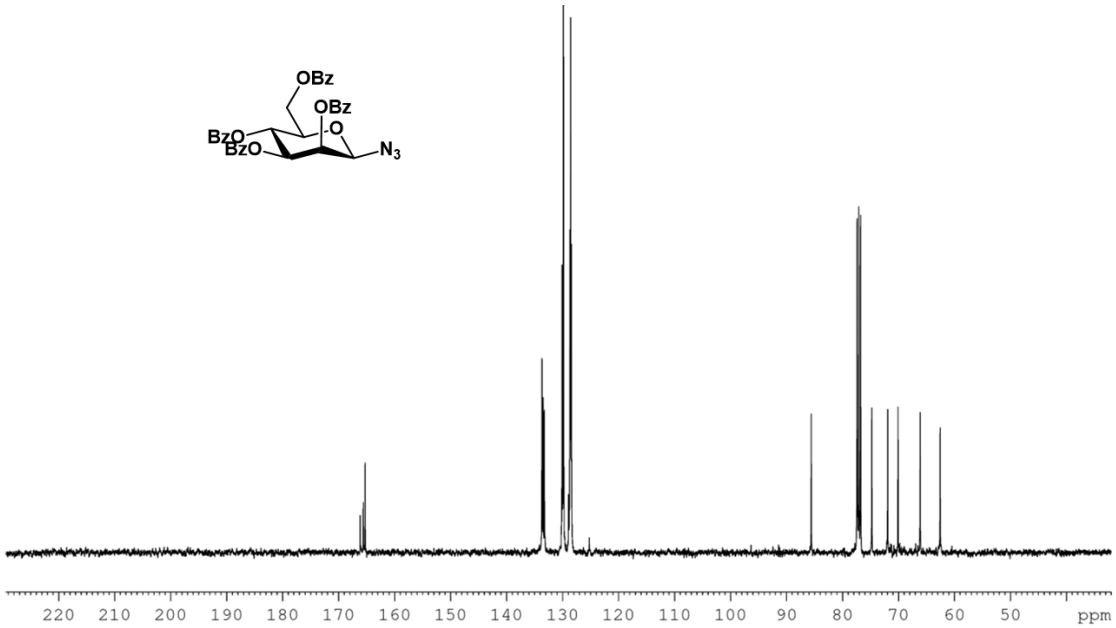


Figure S17. ^{13}C NMR spectrum of **Man_Bz** in $CDCl_3$ (101 MHz).

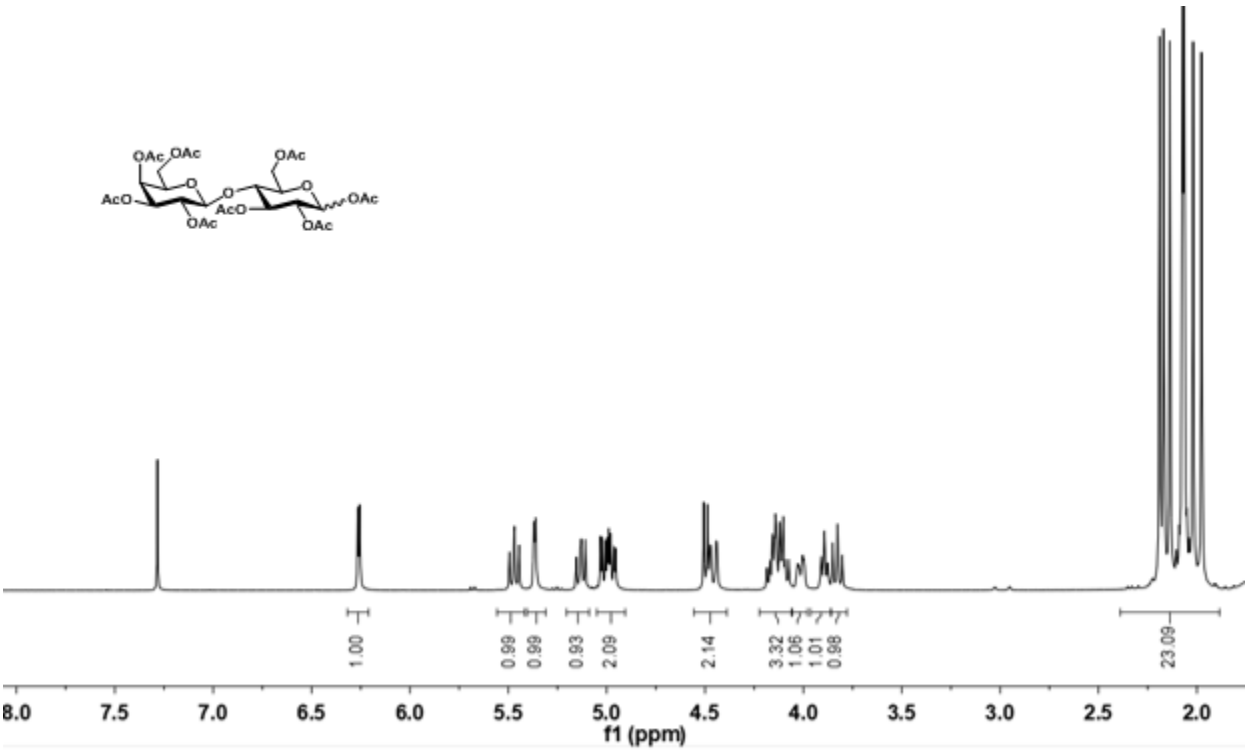


Figure S18. ^{1}H NMR spectrum of octoacetyl-D-lactose **(7)** in $CDCl_3$ (400 MHz).

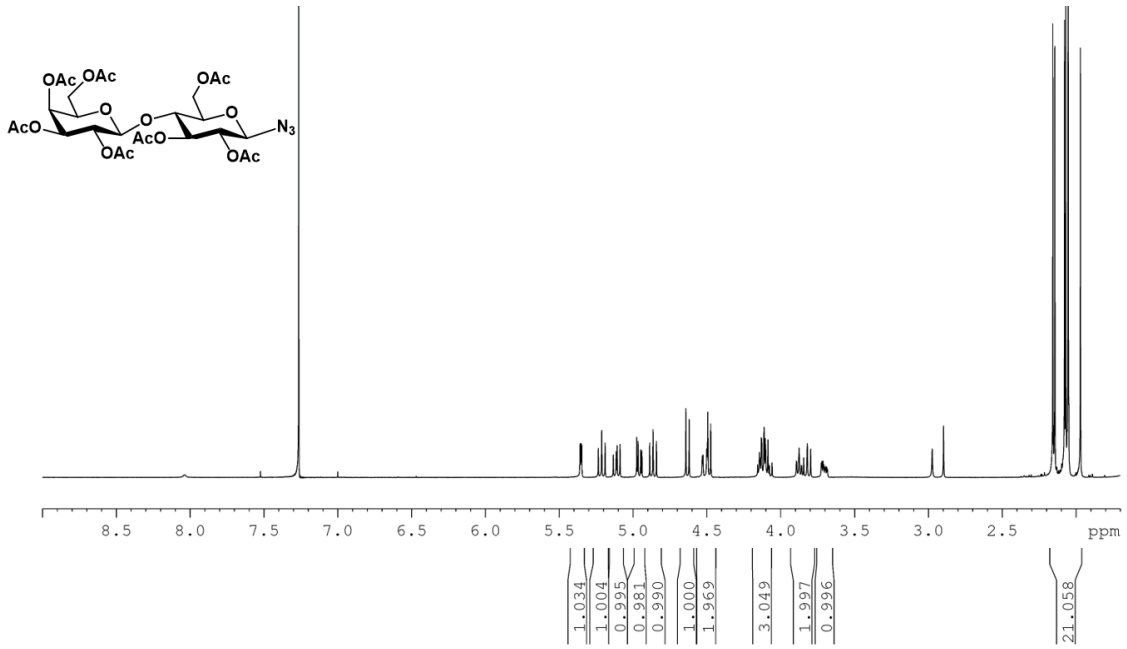


Figure S19. ^{1}H NMR spectrum of **Lac_Ac** in $CDCl_3$ (400 MHz).

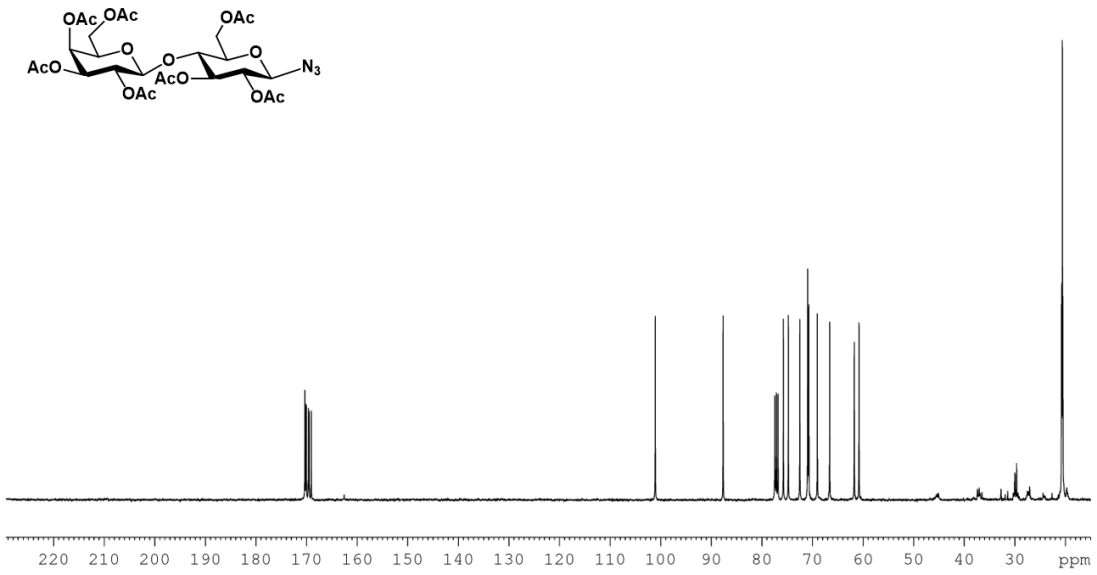


Figure S20. ^{13}C NMR spectrum of **Lac_Ac** in $CDCl_3$ (101 MHz).

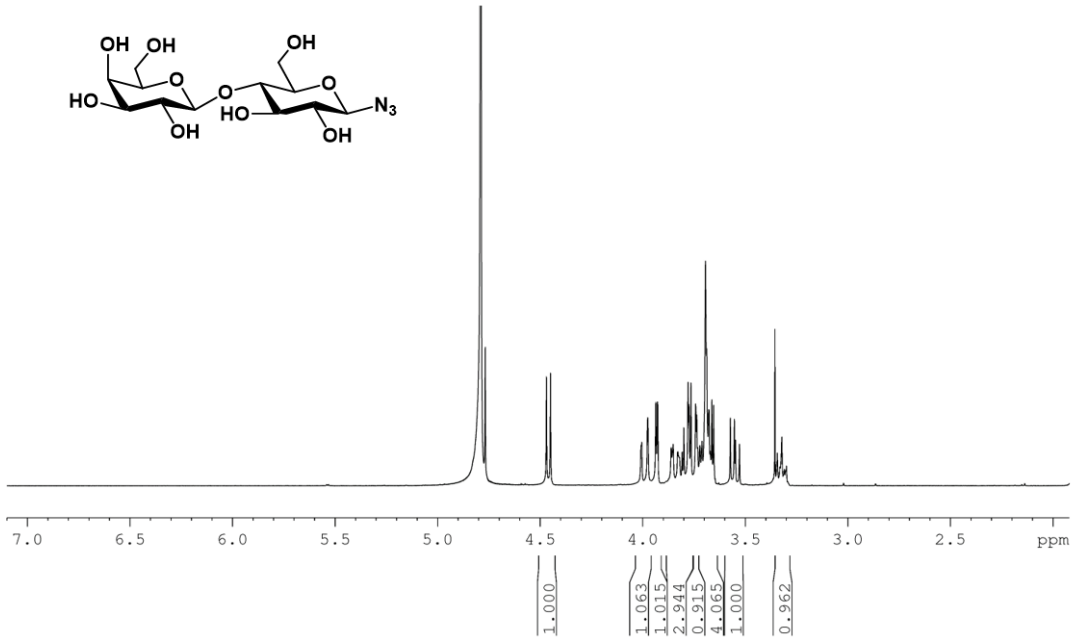


Figure S21. ^{1}H NMR spectrum of **Lac_OH** in D_2O (400 MHz).

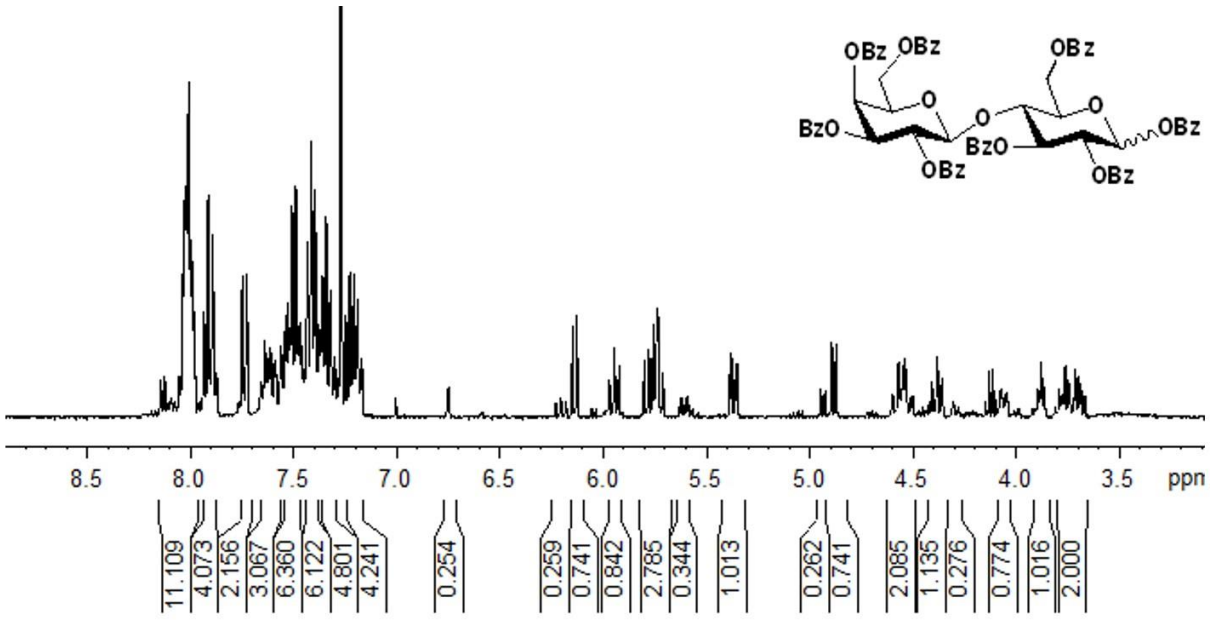


Figure S22. ^{1}H NMR spectrum of octobenzoyl-D-lactose **(9)** in $CDCl_3$ (400 MHz).

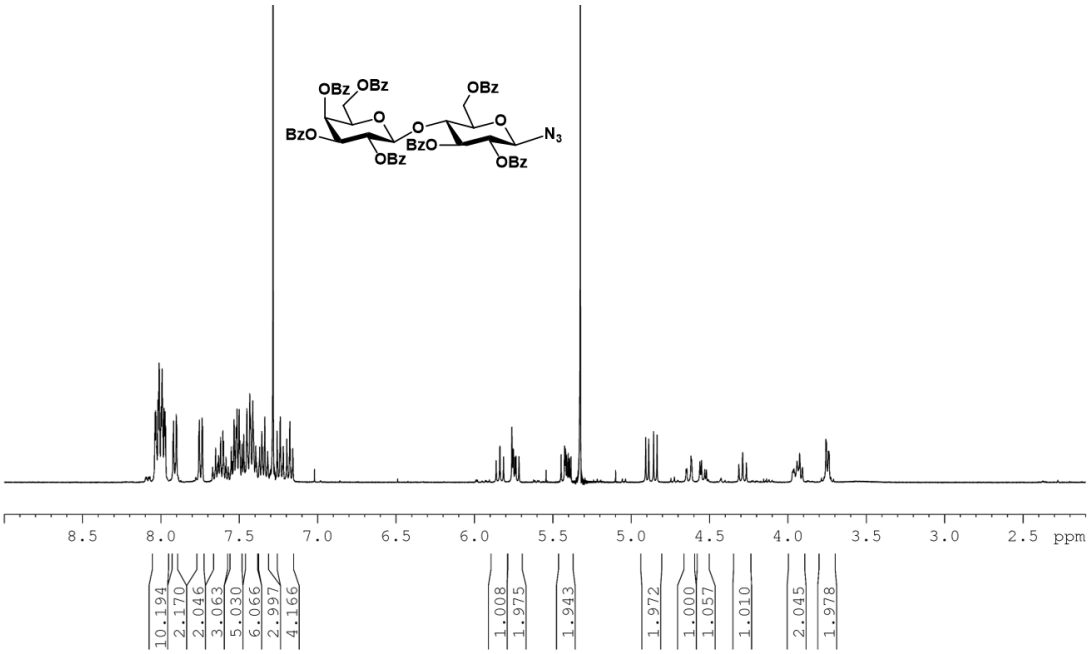


Figure S23. ^{1}H NMR spectrum of **Lac_Bz** in $CDCl_3$ (400 MHz).

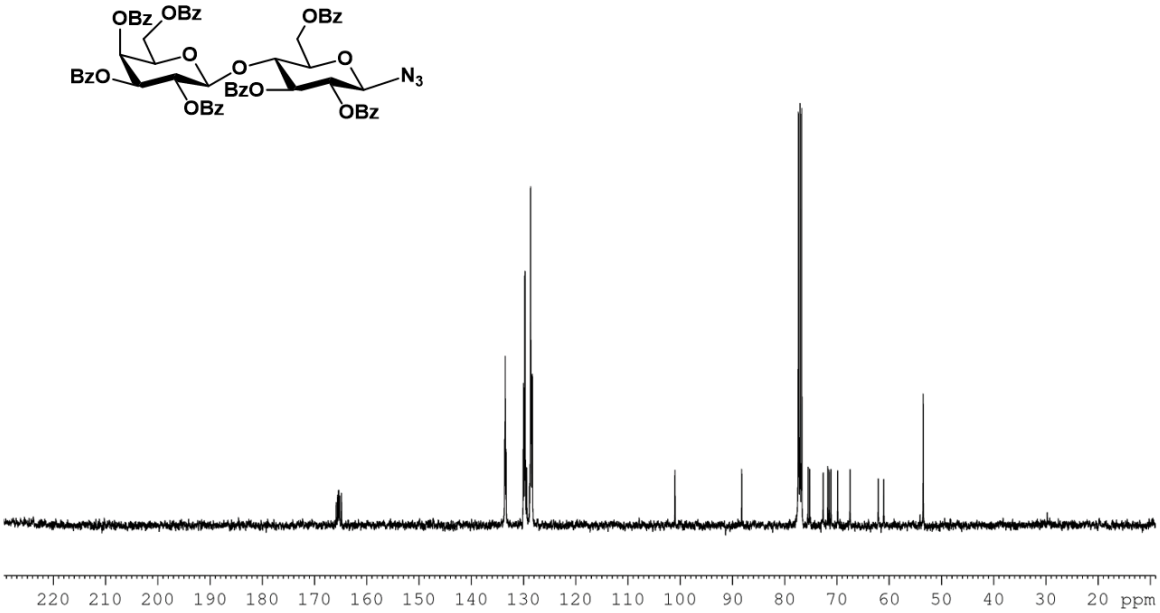


Figure S24. ^{13}C NMR spectrum of **Lac_Bz** in $CDCl_3$ (101 MHz).

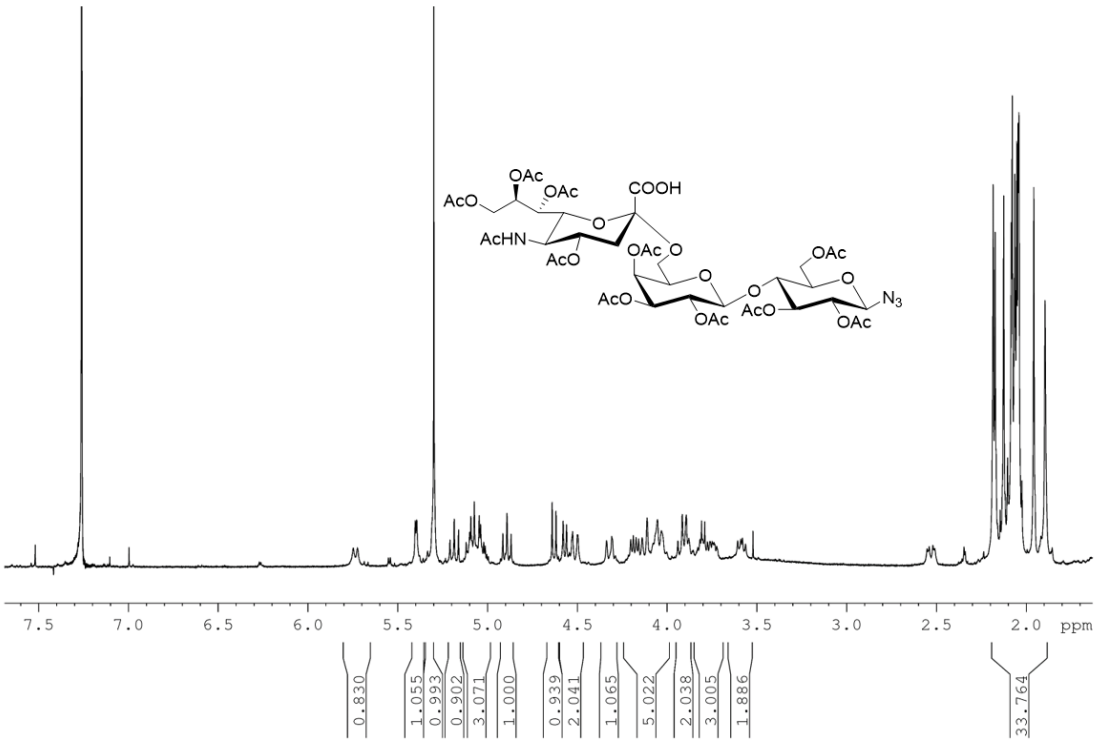


Figure S25. ^{1}H NMR spectrum of **SA_Ac** in $CDCl_3$ (400 MHz).

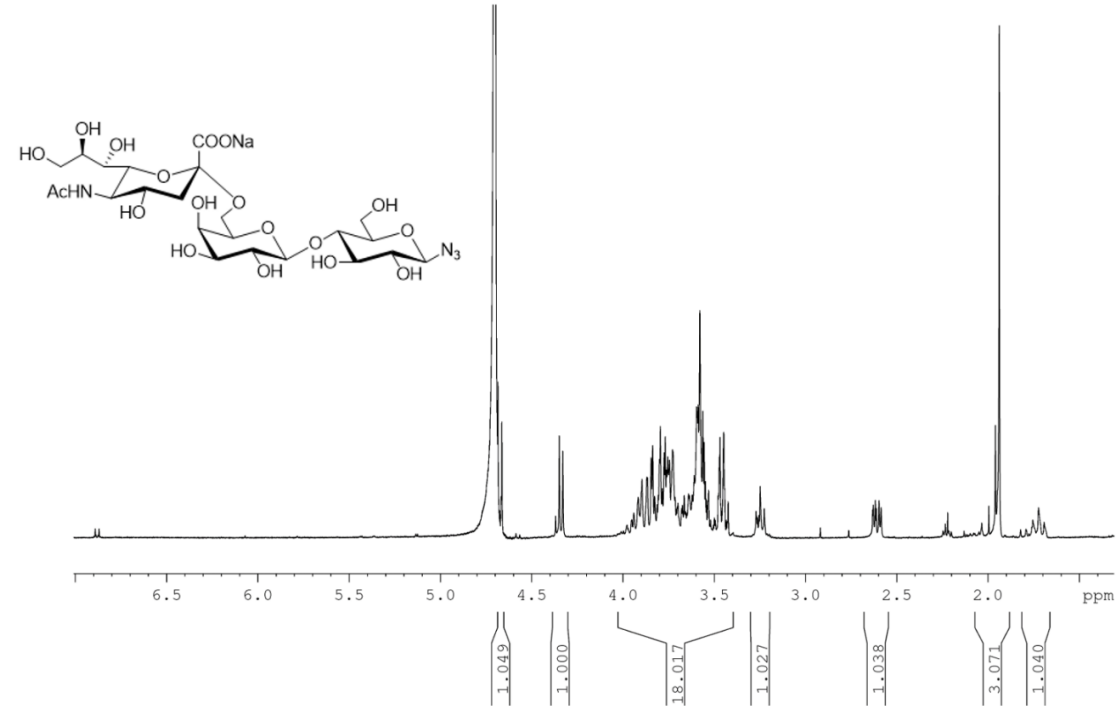


Figure S26. ^{1}H NMR spectrum of **SA_OH** in $CDCl_3$ (400 MHz).

c. Protein NMR spectra

Table S1. Conjugate ^{1}H NMR Reference Peaks

Conjugate	Valine (-CH_3) hydrogens referenced against	Solvent
E50_Man_OH	Triazole hydrogen (8.1 ppm) and hydrogens on carbon adjacent to triazole (6.1 ppm)	D_2O
E50_Man_Ac	Triazole hydrogen (8.1 ppm) and Mannose hydrogens (5.8 to 5.3 ppm)	d-DMF
E50_Man_Bz	Bz hydrogens (8.2 and 7.4 ppm)	d-DMF
E50_Lac_OH	Triazole hydrogen (8.1 ppm) and hydrogens on carbon adjacent to triazole (5.6 ppm)	D_2O
E50_Lac_Ac	Lactose hydrogens (6.5 and 5.0 ppm)	d-DMF
E50_Lac_Bz	Bz hydrogens (8.0 to 7.5 ppm)	d-DMF
E50_SA_OH	SA hydrogens (3.4, 3.1, and 1.2 ppm)	D_2O
E50_SA_Ac	Ac hydrogens (2.4 to 1.6 ppm)	d-DMF

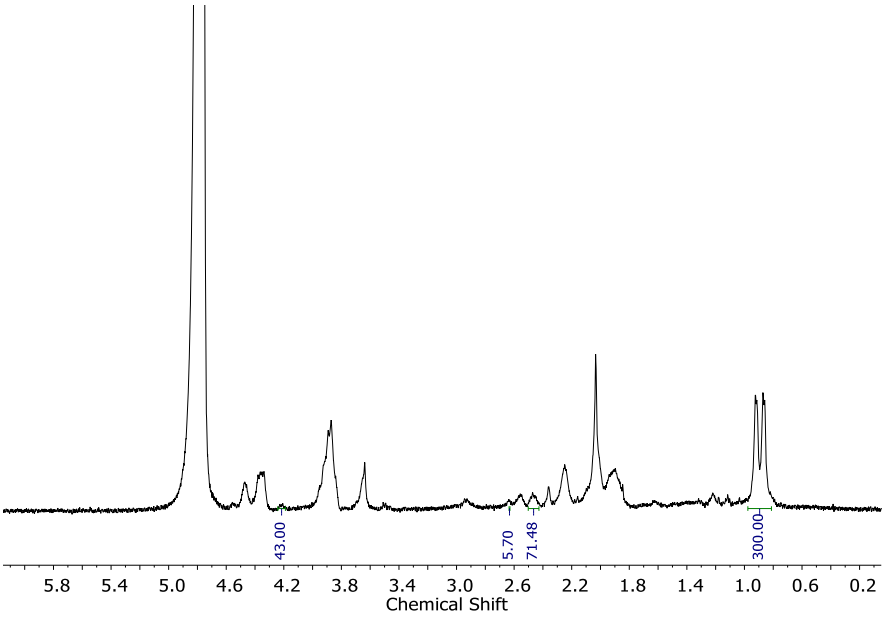


Figure S27. ^{1}H NMR of **E50** in D_2O (500 MHz). Peaks at 4.2 ppm and 2.5 ppm correspond to **HPG** protons. The peaks at 2.6 ppm correspond to unsubstituted methionine (-CH_2) hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

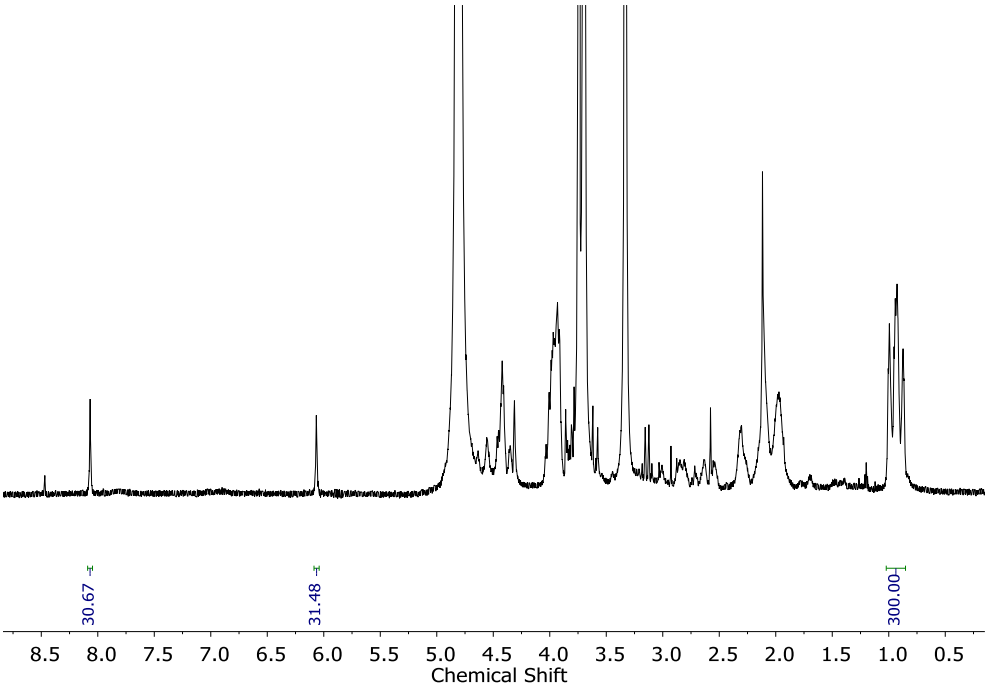


Figure S28. ^{1}H NMR of **E50_Man_OH** in D_2O (600 MHz). The peak at 8.1 ppm corresponds to triazole hydrogens. The peak at 6.1 ppm corresponds to the hydrogens on the carbon adjacent to the triazole. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

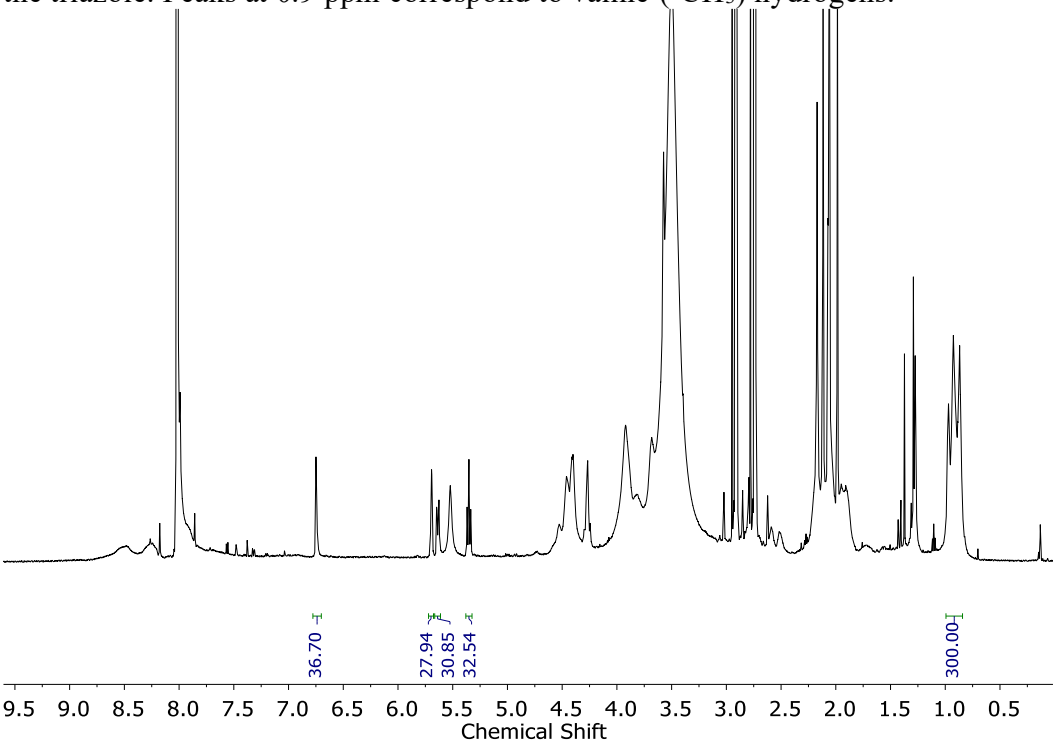


Figure S29. ^{1}H NMR of **E50_Man_Ac** in DMF (600 MHz). The peak at 6.7 ppm corresponds triazole hydrogens. Peaks from 5.8 ppm to 5.3 ppm correspond to Man hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

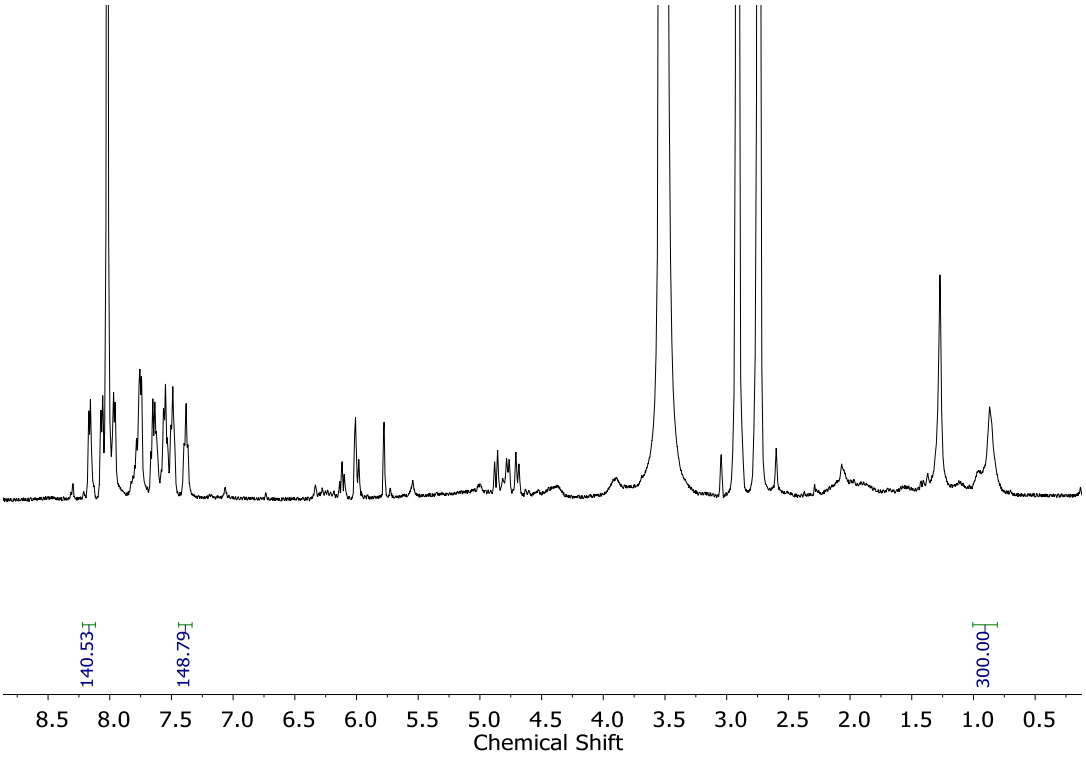


Figure S30. ^{1}H NMR of **E50_Man_Bz** in d-DMF (500 MHz). The peaks at 8.2 and 7.4 ppm corresponds to Bz hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

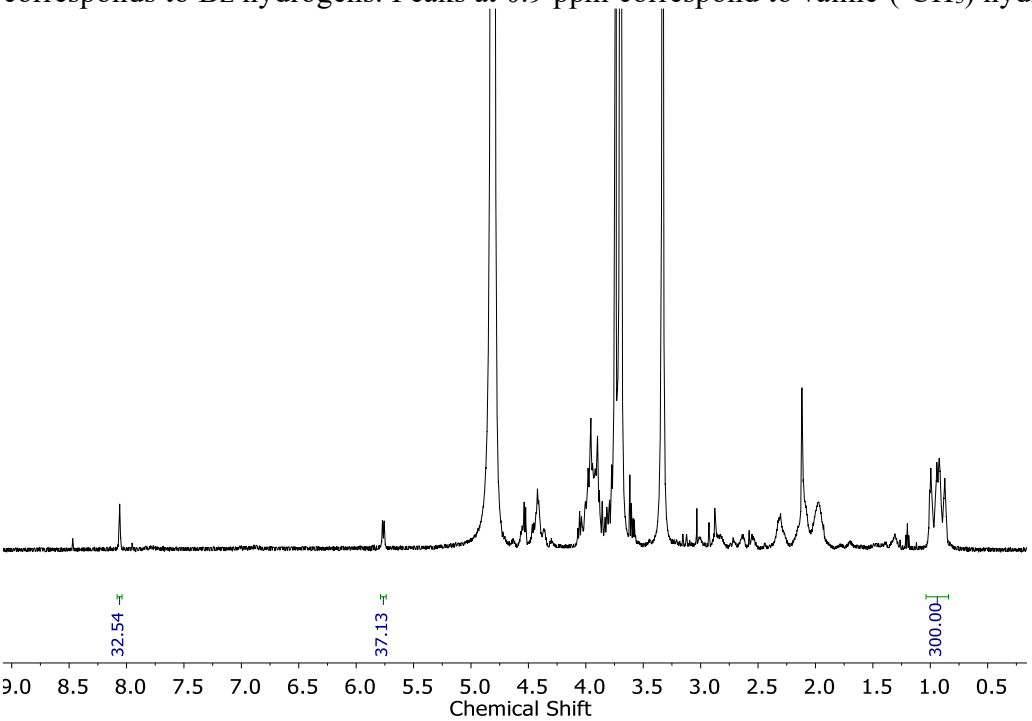


Figure S31. ^{1}H NMR of **E50_Lac_OH** in D_2O (600 MHz). The peak at 8.1 ppm corresponds to triazole hydrogens. The peak at 5.6 ppm corresponds to the hydrogens on the carbon adjacent to the triazole. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

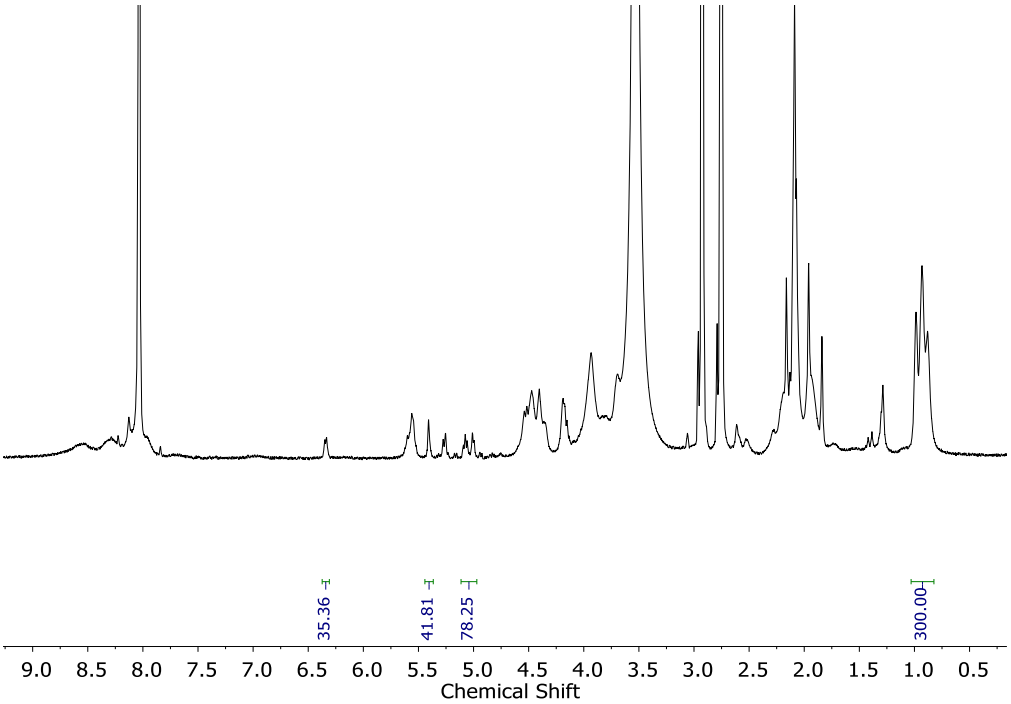


Figure S32. ^{1}H NMR of **E50_Lac_Ac** in d-DMF (500 MHz). The peaks from 6.5 to 5.0 ppm corresponds to lactose hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

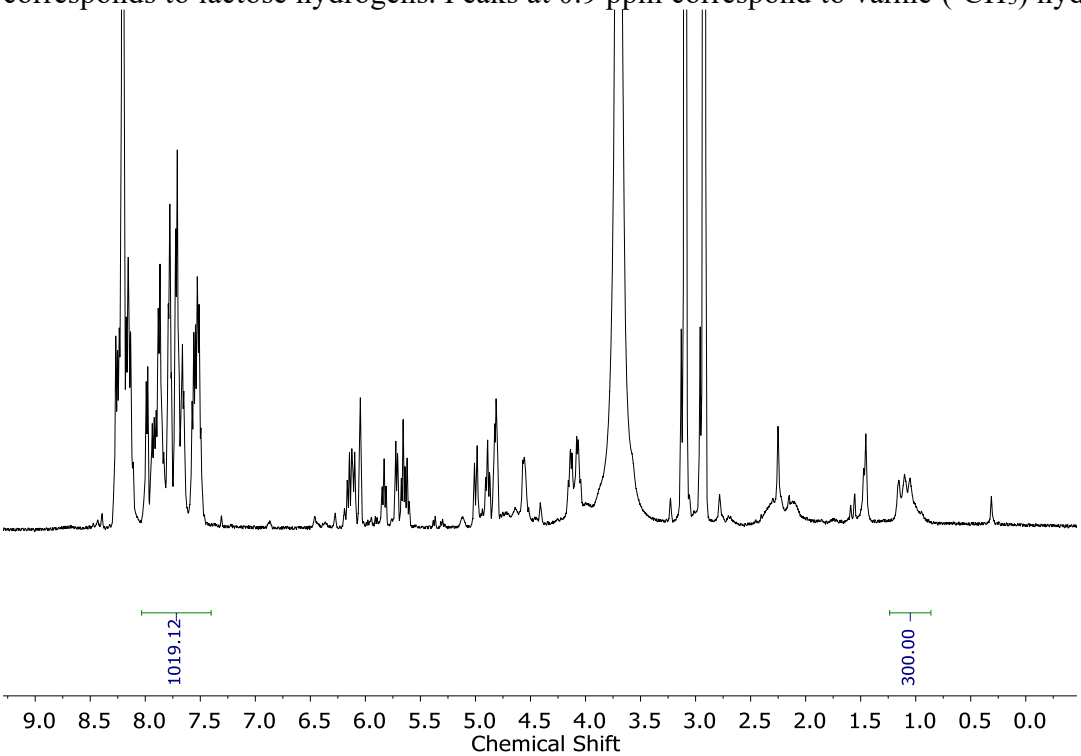


Figure S33. ^{1}H NMR of **E50_Lac_Bz** in d-DMF (500 MHz). The peaks from 8.0 to 7.5 ppm correspond to Bz hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

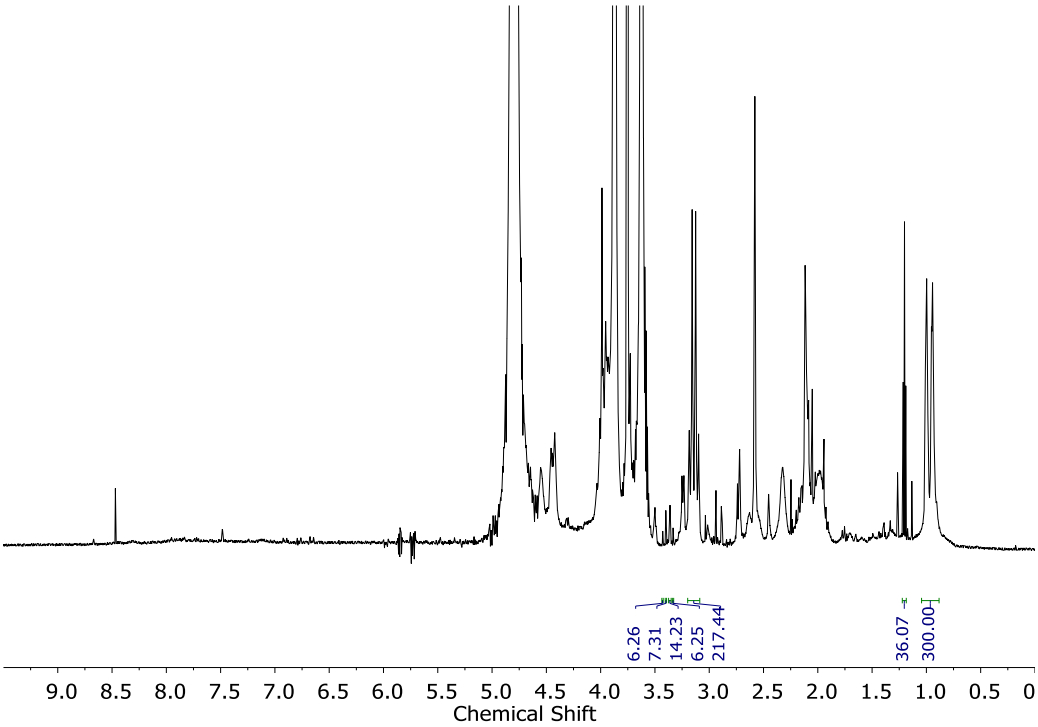


Figure S34. ^{1}H NMR of **E50_SA_OH** in d-D_2O (600 MHz). The peaks from 3.4 to 3.1 ppm and 1.2 ppm correspond to SA hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

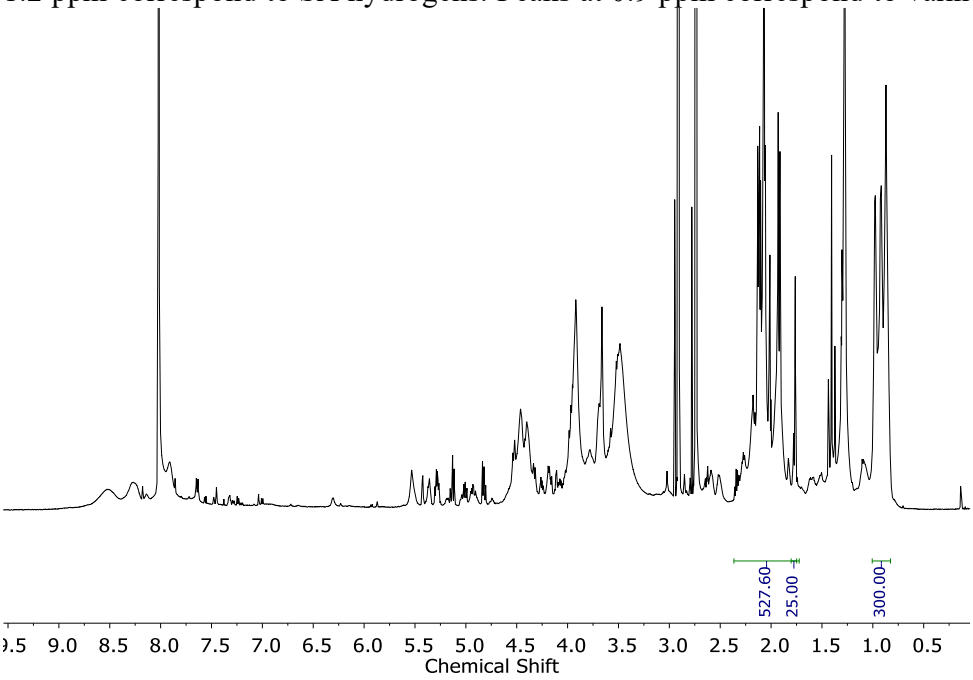


Figure S35. ^{1}H NMR of **E50_SA_Ac** in d-DMF (600 MHz). The peaks from 2.4 to 1.6 ppm correspond to Ac hydrogens. Peaks at 0.9 ppm correspond to valine (-CH_3) hydrogens.

d. Protein Gels

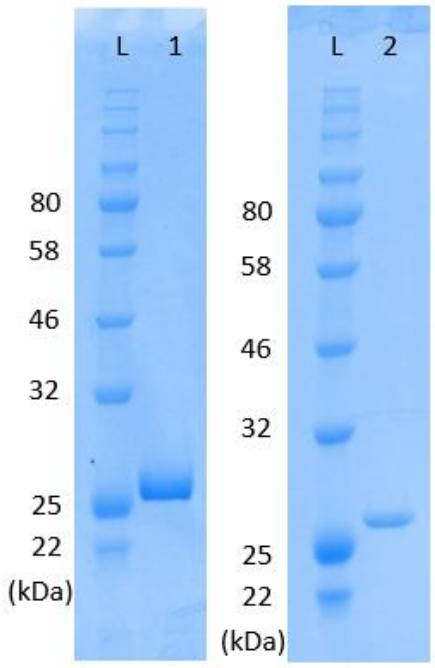


Figure S36. Stained Coomassie Blue SDS-PAGE gel of **E50** and **E50-MET**. L – Protein Ladder, 1– Purified **E50**, 2– Purified **E50-MET**. Molar masses are in kilodaltons.

e. Integrated ITC data

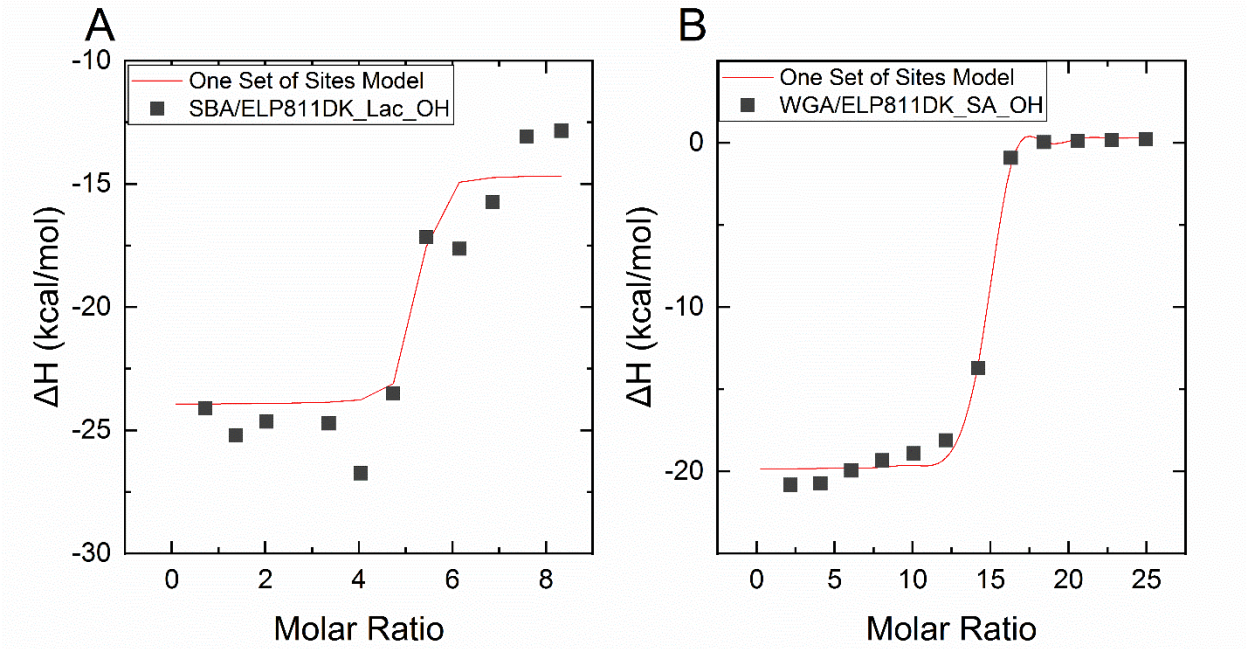


Figure S37. ITC results of integrated data for titration of (A) **E50_Lac_OH** (200 µM) to SBA (200 µM) in HEPES buffer at 25 ºC and (B) **E50_SA_OH** (10 µM) to WGA (100 µM) in HEPS buffer at 25 ºC.

f. Phase Diagrams

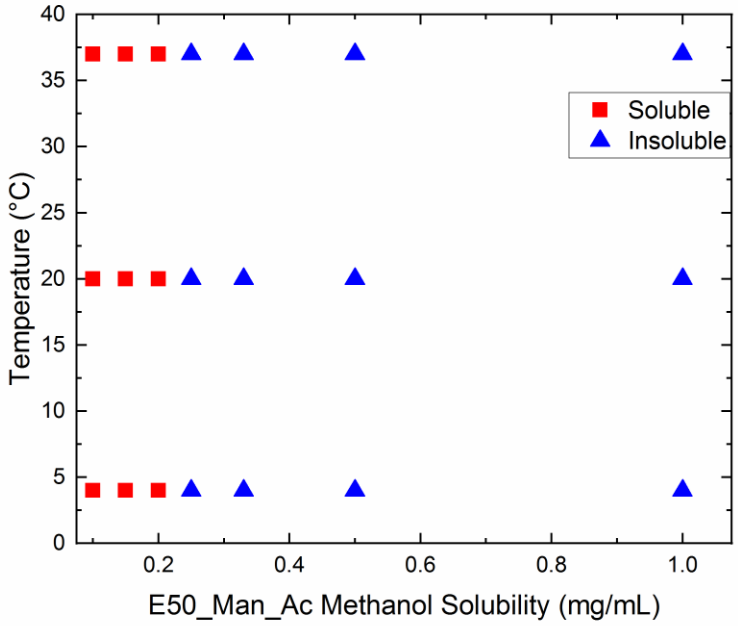


Figure S38. Solubility study for **E50_Man_Ac** in methanol.

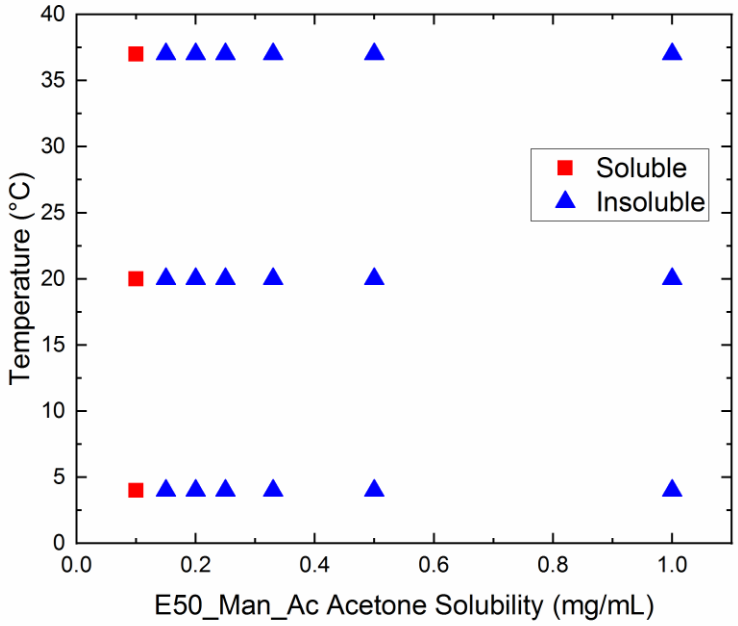


Figure S39. Solubility study for **E50_Man_Ac** in acetone.

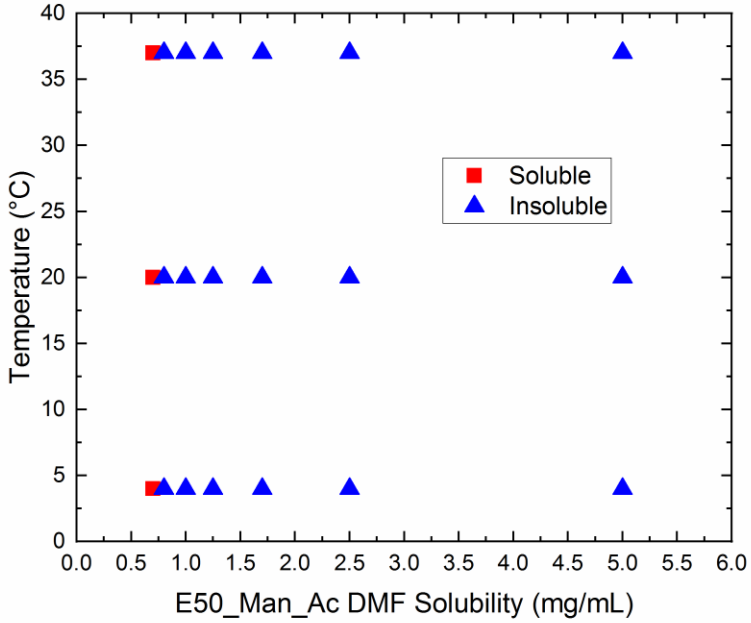


Figure S40. Solubility study for **E50_Man_Ac** in DMF.

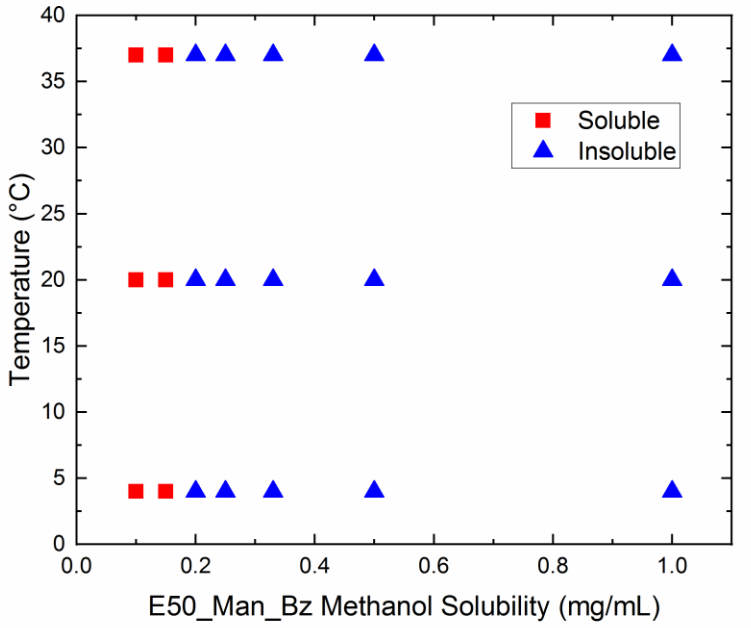


Figure S41. Solubility study for **E50_Man_Bz** in methanol.

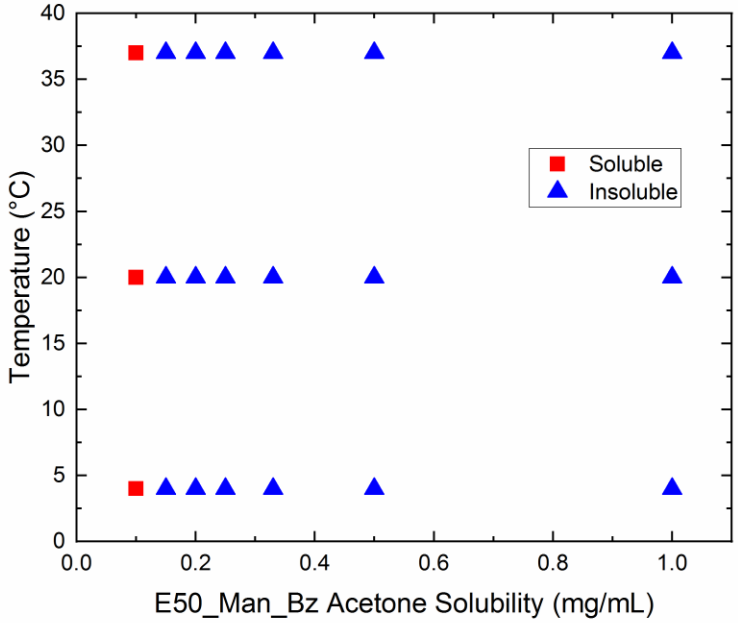


Figure S42. Solubility study for **E50_Man_Bz** in acetone.

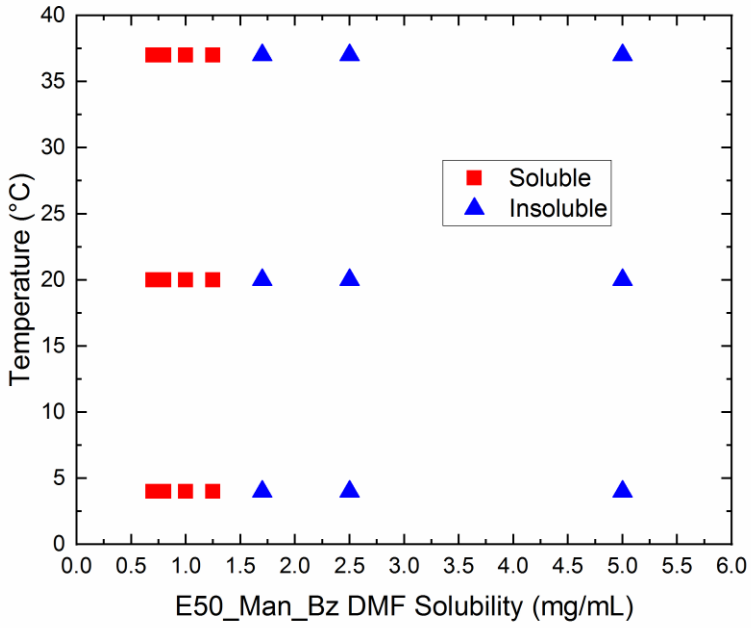


Figure S43. Solubility study for **E50_Man_Bz** in DMF.

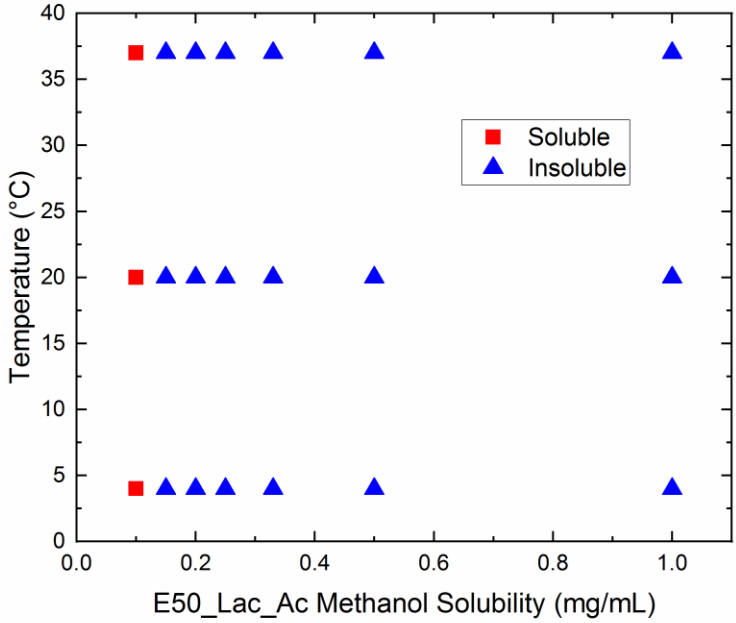


Figure S44. Solubility study for **E50_Lac_Ac** in methanol.

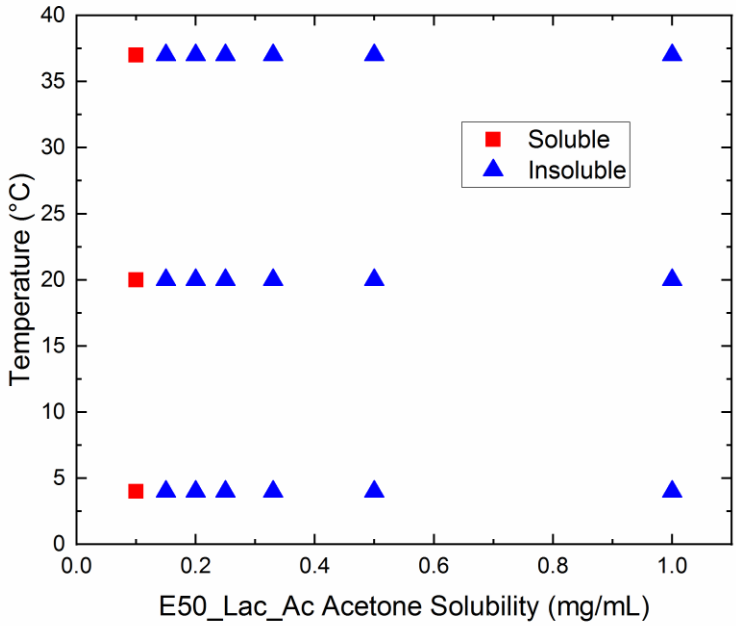


Figure S45. Solubility study for **E50_Lac_Ac** in acetone.

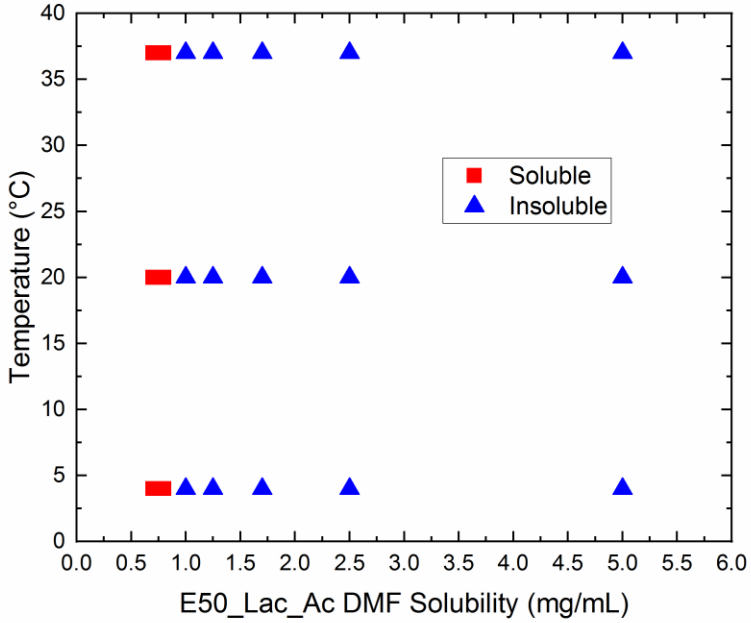


Figure S46. Solubility study for **E50_Lac_Ac** in DMF.

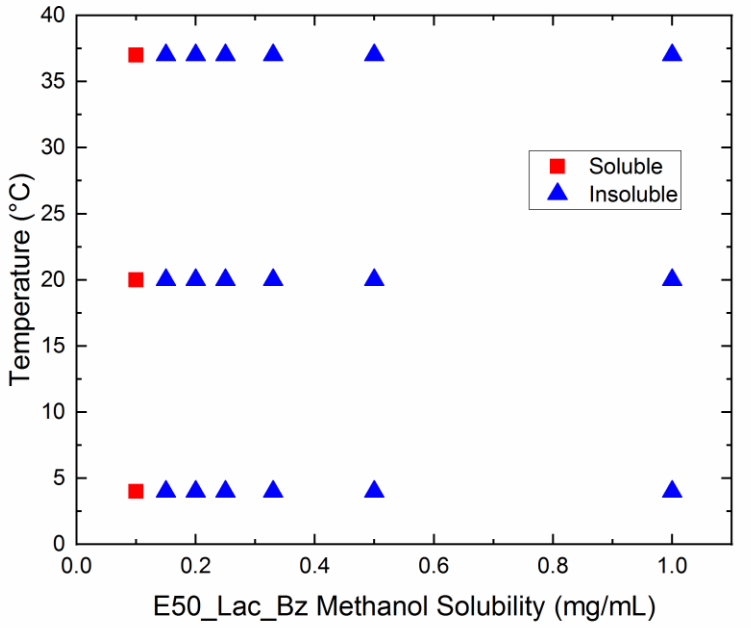


Figure S47. Solubility study for **E50_Lac_Bz** in methanol.

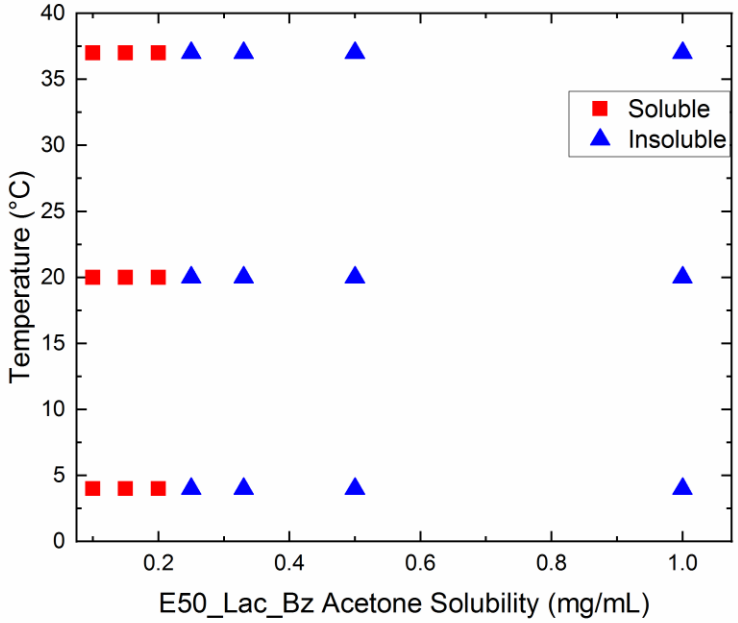


Figure S48. Solubility study for **E50_Lac_Bz** in acetone.

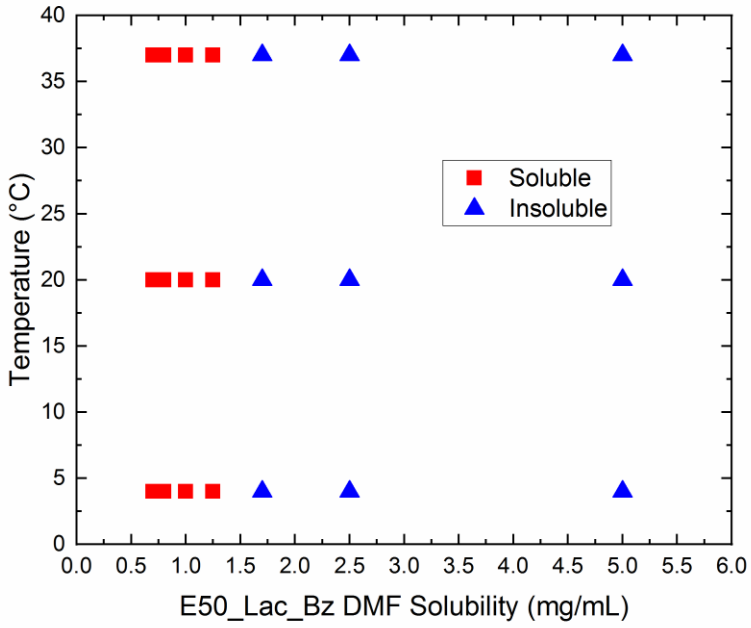


Figure S49. Solubility study for **E50_Lac_Bz** in DMF.

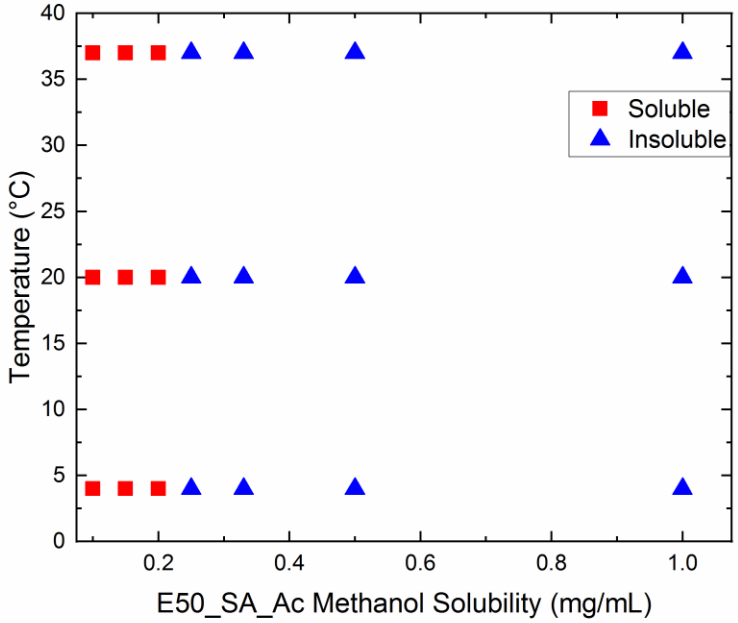


Figure S50. Solubility study for **E50_SA_Ac** in methanol.

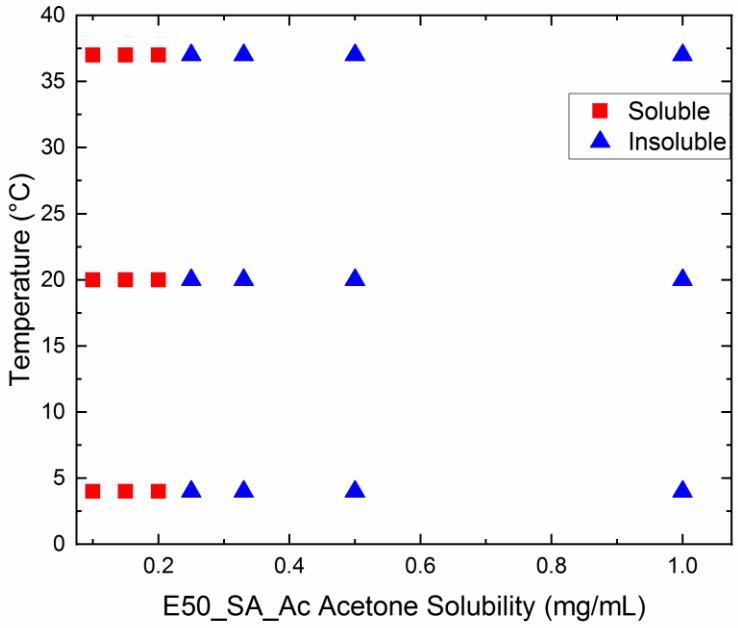


Figure S51. Solubility study for **E50_SA_Ac** in acetone.

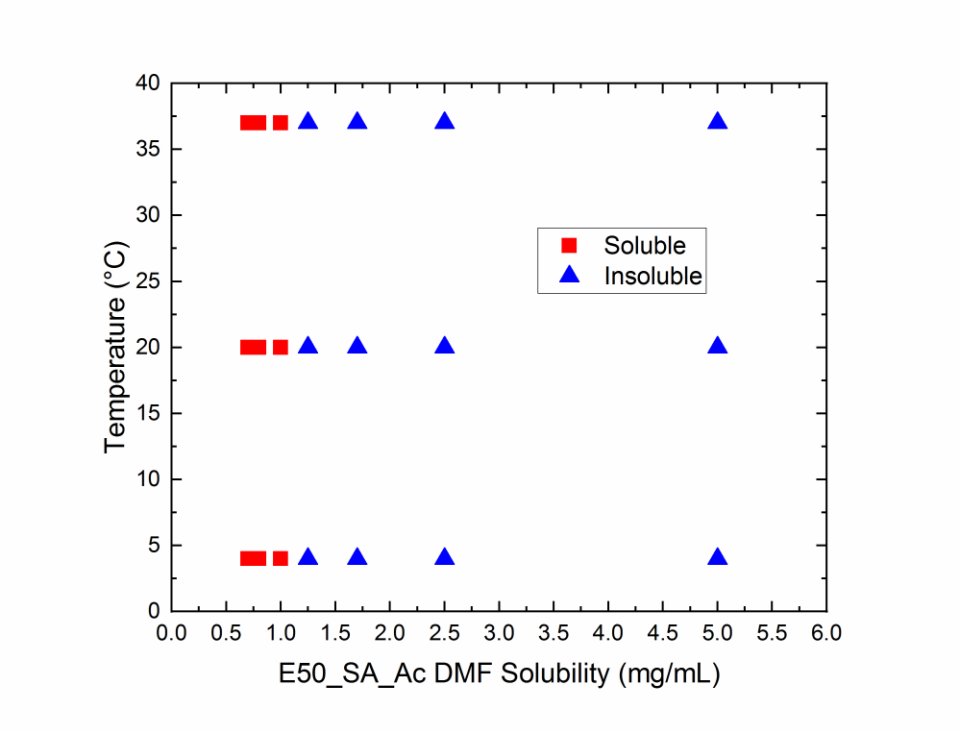


Figure S52. Solubility study for **E50_SA_Ac** in DMF.

S7. References

(1) Brea, R. J., López-Deber, M. P., Castedo, L., and Granja, J. R. (2006) Synthesis of ω-(Hetero)arylalkynylated α-Amino Acid by Sonogashira-Type Reactions in Aqueous Media. J. Org. Chem. 71, 7870-7873.

(2) van Hest, J. C. M., Kiick, K. L., and Tirrell, D. A. (2000) Efficient Incorporation of Unsaturated Methionine Analogues into Proteins in Vivo. J. Am. Chem. Soc. 122, 1282-1288.

(3) Medran, N. S., Villalba, M., Mata, E. G., and Testero, S. A. (2016) Gold-Catalyzed Cycloisomerization of Alkyne-Containing Amino Acids: Controlled Tuning of C–N vs. C–O Reactivity. Eur. J. Org. Chem. 2016, 3757-3764.

(4) Negishi, K., Mashiko, Y., Yamashita, E., Otsuka, A., and Hasegawa, T. (2011) Cellulose Chemistry Meets Click Chemistry: Syntheses and Properties of Cellulose-Based Glycoclusters with High Structural Homogeneity. Polymers 3, 489.

(5) Machida, T., Lang, K., Xue, L., Chin, J. W., and Winssinger, N. (2015) Site-Specific Glycoconjugation of Protein via Bioorthogonal Tetrazine Cycloaddition with a Genetically Encoded trans-Cyclooctene or Bicyclononyne. Bioconjugate Chem. 26, 802-806.

(6) Kartha, K. P. R., and Field, R. A. (1997) Iodine: A versatile reagent in carbohydrate chemistry IV. Per-O-acetylation, regioselective acylation and acetolysis. Tetrahedron 53, 11753-11766.

(7) Kopitzki, S., Jensen, K. J., and Thiem, J. (2010) Synthesis of Benzaldehyde-Functionalized Glycans: A Novel Approach Towards Glyco-SAMs as a Tool for Surface Plasmon Resonance Studies. Chem.: Eur. J 16, 7017-7029.

(8) Mereyala, H. B., and Gurrala, S. R. (1998) A highly diastereoselective, practical synthesis of allyl, propargyl 2,3,4,6-tetra-O-acetyl-β-d-gluco, β-d-galactopyranosides and allyl, propargyl heptaacetyl-β-d-lactosides. Carbohydr. Res. 307, 351-354.

(9) Benito-Alifonso, D., Tremel, S., Hou, B., Lockyear, H., Mantell, J., Fermin, D. J., Verkade, P., Berry, M., and Galan, M. C. (2014) Lactose as a “Trojan Horse” for Quantum Dot Cell Transport. Angew. Chem. 126, 829-833.
(10) Sandbhor, M. S., Soya, N., Albohy, A., Zheng, R. B., Cartmell, J., Bundle, D. R., Klassen, J. S., and Cairo, C. W. (2011) Substrate recognition of the membrane-associated sialidase NEU3 requires a hydrophobic aglycone. Biochemistry 50, 6753-62.