Illuminating HIV-1 gp120–Ligand Recognition through Computationally-Driven Optimization of Antibody-Recruiting Molecules

Christopher G. Parker, Markus K. Dahlgren, Ran N. Tao, Don T. Li, Eugene F. Douglass, Takuji Shoda, Navneet Jawanda, Krasimir A. Spasov, Sangil Lee, Nannan Zhou, Robert A. Domaoal, Richard E. Sutton, Karen S. Anderson,William L. Jorgensen, Mark Krystal, David A. Spiegel

Table of Contents

Abbreviations

AcOH = acetic acid
Boc = *tert*-butoxycarbonyl
BSA = bovine serum albumin
DCM = dichloromethane
DEPBT = 3-(Diethoxyphosphorylloxy)-1,2,3-benzotriazin-4-(3H)-one
DIPEA = *N*,*N*-diisopropylethylamine
DMF = *N*,*N*-dimethylformamide
DMSO = dimethylsulfoxide
DNP = 2,4-dinitrophenyl
DPBS = Dulbecco's phosphate-buffered saline
EDC = 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EtOAc = ethyl acetate
Fmoc = 9-fluorenylmethyloxycarbonyl
HI-FBS = heat inactivated fetal bovine serum
HOBt = hydroxybenzotriazole
MeCN = acetonitrile
MeOH = methanol
MW = microwave reactor
NaAsc = sodium ascorbate
pyr = pyridine
*t*BuOH = t-butanol
TEA = triethylamine
TFA = trifluoroacetic acid
THF = tetrahydrofurane

Chemistry

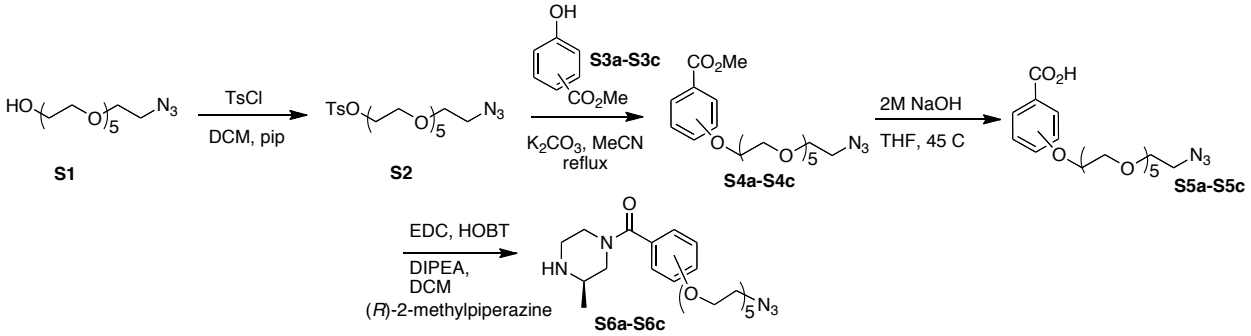
Materials and General Information

Purchased starting materials were used as received unless otherwise noted. All moisture sensitive reactions were performed in an inert, dry atmosphere of nitrogen in flame dried glassware. Reagent grade solvents were used for extractions and flash chromatography. Reaction progress was checked by analytical thin-layer chromatography (TLC, Merck silica gel 60 F-254 plates). The plates were monitored either with UV illumination, or by charring with anisaldehyde (2.5% *p*-anisaldehyde, 1% AcOH, 3.5% H_2SO_4(conc.) in 95% EtOH) or ninhydrin (0.3% ninhydrin (w/v), 97:3 EtOH-AcOH) stains. Flash column chromatography was performed using silica gel (230-400 mesh). The solvent compositions reported for all chromatographic separations are on a volume/volume (v/v) basis.

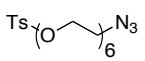
Instrumentation

^{1}H-NMR spectra were recorded at either 400 or 500 MHz and are reported in parts per million (ppm) on the δ scale relative to solvent as an internal standard ($CDCl_3$ = 7.26 ppm, DMSO-d_6 = 2.49 ppm, CD_3OD = 3.30). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br = broad, m = multiplet), coupling constants (Hz), and integration. ^{13}C-NMR spectra were recorded at either 100 or 125 MHz and are reported in parts per million (ppm) on the δ scale relative to solvent as an internal standard ($CDCl_3$ = 77.0 ppm, DMSO-d_6 = 39.5 ppm, CD_3OD = 49.0 ppm). Analytical ultra high-performance liquid chromatography/mass spectrometry (UPLC/MS) was performed on a Waters UPLC/MS instrument equipped with a reversephase C_{18} column (1.7 μm particle size, 2.1 Å~ 50 mm), dual atmospheric pressure chemical ionization (API)/electrospray (ESI) mass spectrometry detector, and photodiode array detector. Samples were eluted with a linear gradient of 20% acetonitrile−water containing 0.1% formic acid→100% acetonitrile containing 0.1% formic acid over 3 min, followed by 100% acetonitrile containing 0.1% formic acid for 1 min, at a flow rate of 800 μL/min. High-resolution liquid chromatography-mass spectrometry (HR-LC/MS) was performed on a Waters UPLC/HRMS instrument equipped with a dual API/ESI high-resolution mass spectrometer, and a photodiode array detector. High Pressure Liquid Chromatography (HPLC) using a Dynamax Rainin Solvent Delivery System equipped with a Varian Prostar Detector (Galaxie Chromatography Data System version 1.8.505.5), and absorbance measurements were made at 214 and 254 nm simultaneously. A Waters Xterra Prep MS C18 7.8x150mm column was used for semi-preparative purifications using a water:acetonitrile (A:B) gradient containing 0.1% TFA at 5.0 mL/min, as specified below for individual compounds. Flash column chromatography (unless otherwise noted) was performed using silica gel (230-400 mesh) using Teledyne Isco CombiFlash Rf 200 equipped with a UV detector and fraction collector. Attenuated total reflectance Fourier transform infrared spectra (ATR-FTIR) were obtained using a Thermo Electron Corporation Nicolet 6700 FTIR spectrometer. Data are represented as follows: frequency of absorption (cm–1), intensity of absorption (s = strong, m = medium, w = weak, br = broad).

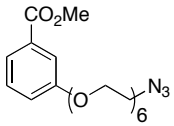
Synthetic Procedures



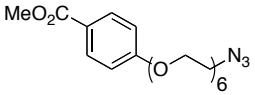
Scheme S1. General synthesis of (R)-methylpiperazine-PEG6-azides from **S1**.



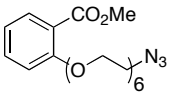
(**S2**) To a solution of 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol (**S1**, 6.45 g, 21 mmol) in CH_2Cl_2 (35 mL), added a solution of 4-toluenesulfonyl chloride (4.56 g, 25.2 mmol, 1.2 equiv) in pyridine (9 mL) dropwise via addition funnel. Let homogeneous mixture stir at room temperature until TLC (20:1 CH_2Cl_2/CH_3OH) indicated reaction completion (14 hrs). Reaction mixture was diluted with additional CH_2Cl_2, the organic layer was washed twice with aq. HCl (2M, 30 mL), dried over anhydrous $MgSO_4$, filtered and all solvents were evaporated yielding **S2** as a colorless oil (8.67 g, 18.8 mmol, 90%). Crude material was used without further purification. **IR (ATR-FTIR)** 2867 (m), 2100 (s), 1353 (s), 1291 (m), 1175 (s), 1095 (s), 1017 (w), 920 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 7.76 (d, J = 8.3, 2H), 7.31 (d, J = 8.0, 2H), 4.12 (m, 2H), 3.68 – 3.56 (m, 20H), 3.35 (m 2H), 2.41 (s, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 145.21, 133.39, 130.23, 128.37, 71.11, 71.06, 71.03, 71.01, 70.99, 70.94, 70.92, 70.89, 70.40, 69.66, 69.06, 51.07, 22.03. **HRMS (ESI+)** calc'd for $C_{19}H_{32}N_3O_8S$ $[M+H]^+$ *m/z* 462.1905. Found 462.2037.



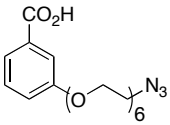
(**S4a**) To a flame-dried flask containing a solution of methyl 3-hydroxybenzoate (**S3a**, 190 mg, 1.25 mmol, 1.16 equiv) in anhydrous MeCN (10 mL), added K_2CO_3 (173 mg, 1.25 mmol, 1.16 equiv) and S2 (500 mg, 1.08 mmol). Resulting mixture was heated to reflux under an atmosphere of N_2 until TLC (5:1 Hexanes/EtOAc) indicated reaction completion (14 hrs). Reaction was quenched with saturated NH_4Cl (30 mL) and extracted with DCM (3 x 50 mL). The combined organic layers were dried over anhyd. $MgSO_4$, filtered, and all solvents were evaporated. Crude **S4a** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 25g pre-packed dry loading column. Run using 10% EtOAc:Hexanes to 50% EtOAc:Hexanes gradient over 40 column volumes, followed by EtOAc flush) to yield **S4a** as a clear viscous oil (331 mg, 70%). **IR (ATR-FTIR)** 2868 (m), 2101 (s), 1719 (s), 1586 (w), 1444 (m), 1277 (s), 1230 (w), 1097 (s), 940 (w) cm-1. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.64 – 7.57 (m, 1H), 7.54 (dd, J = 1.5, 2.5, 1H), 7.31 (t, J = 8.0, 1H), 7.10 (ddd, J = 1.0, 2.7, 8.3, 1H), 4.19 – 4.10 (m,s 2H), 3.88 (s, 3H), 3.86 – 3.82 (m, 2H), 3.73 – 3.57 (m, 20H), 3.40 – 3.33 (m, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 166.90, 158.76, 131.38, 129.40, 122.15, 120.08, 114.73, 70.86, 70.68, 70.64, 70.59, 70.03, 69.64, 67.63, 52.17, 50.67. **HRMS (ESI+)** calc'd for [M+H]+ $C_{20}H_{32}N_3O_8$ m/z 442.2184. Found 442.2308; for [M + Na]+, 464.2003, found 464.2113.



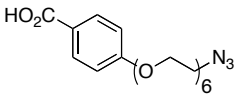
(S4b) Prepared **S4b** in the same manner as compound **S4a** starting from methyl 4-hydroxybenzoate (**S3b**, 190mg, 1.25 mmol) and **S2** (500 mg, 1.08 mmol). Crude **S4b** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with pre-packed dry loading column. Run using 0% EtOAc in hexanes to 50% EtOAc in hexanes gradient over 30 column volumes to yield **S4b** as a clear sticky residue in 73% yield (350 mg, 0.79 mmol). **IR (ATR-FTIR)** 2869 (m), 2104 (s), 1715 (s), 1606 (s), 1511 (w), 1283 (w), 1255 (s), 1170 (w), 1107 (s) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 7.97 (d, J = 9.0, 2H), 6.92 (d, J = 9.0, 2H), 4.24 – 4.09 (m, 2H), 3.87 (m, 5H), 3.75 – 3.70 (m, 2H), 3.70 – 3.61 (m, 16H), 3.37 (t, J = 5.0, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 166.74, 162.52, 131.48, 122.68, 114.15, 70.84, 70.64, 70.63, 70.58, 70.54, 69.97, 69.49, 67.54, 51.77, 50.63. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{20}H_{32}N_3O_8$ *m/z* 442.2184. Found 442.2298; for $[M + Na]^+$, 464.2003, found 464.2114.



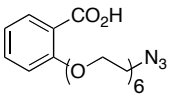
(S4c) Prepared **S4c** in the same manner as compound **S4c** starting from methyl 2-hydroxybenzoate (**S3c**, 113mg, 0.743 mmol) and **S2** (300 mg, 0.65 mmol). Crude **S4c** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 5% MeOH in DCM gradient over 45 column volumes to yield **S4c** as a clear sticky residue in 77% yield (221 mg, 0.50 mmol). **IR (ATR-FTIR)** 2869 (m), 2102 (s), 1725 (s), 1601 (w), 1450 (w), 1303 (m), 1248 (s), 1083 (s), 943 (w) cm^{-1}. **^{1H} NMR (400 MHz, $CDCl_3$)** δ 7.76 (dd, J = 1.8, 7.9, 1H), 7.42 (ddd, J = 1.8, 7.5, 8.4, 1H), 7.02 – 6.92 (m, 2H), 4.21 – 4.16 (m, 2H), 3.91 – 3.86 (m, 2H), 3.85 (s, 3H), 3.77 – 3.71 (m, 2H), 3.69 – 3.58 (m, 18H), 3.37 (dd, J = 4.2, 9.3, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 166.78, 161.01, 158.36, 133.42, 131.64, 120.73, 120.59, 113.81, 71.03, 70.69, 70.65, 70.60, 70.05, 69.60, 68.96, 68.87, 63.06, 51.95, 50.70. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{20}H_{32}N_3O_8$ *m/z* 442.2184. Found 442.2281; for $[M + Na]^+$, 464.2003, found 464.2118.



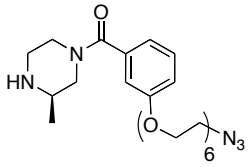
(**S5a**) A solution of **S4a** (318 mg, 0.72 mmol) in THF (12 mL) and aq. NaOH (2M, 5 mL) was heated to 45 ^oC for 20 hrs when TLC indicated reaction completion (20:1 CH_2Cl_2/CH_3OH). The solution was acidified to a pH of 1 using 6M aq. HCl and then extracted with CH_2Cl_2 (3 x 30 mL). The combined organic layers were dried over anhyd. $MgSO_4$, filtered, and all solvents were evaporated, resulting in **17** as a clear viscous oil (280 mg, 90%), which was used without further purification. **IR (ATR-FTIR)** 2882 (br), 2872 (m), 2104 (s), 1714 (s), 1447 (m), 1372 (w), 1259 (s), 1109 (s) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 7.67 (d, J = 7.7, 1H), 7.59 (s, 1H), 7.33 (t, J = 8.0, 1H), 7.13 (dd, J = 1.9, 8.2, 1H), 4.21 – 4.12 (m, 2H), 3.91 – 3.82 (m, 2H), 3.74 – 3.69 (m, 2H), 3.69 – 3.58 (m, 16H), 3.40 – 3.29 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 171.36, 159.23, 131.15, 129.89, 123.17, 121.25, 115.63, 71.27, 71.09, 71.07, 71.04, 71.01, 70.98, 70.97, 70.43, 70.42, 70.07, 68.10, 51.07. **HRMS (ESI+)** calc'd for $C_{19}H_{30}N_3O_8$ $[M+H]^+$ *m/z* 428.2027. Found 428.2130; for $[M + Na]^+$, calc'd 450.1847, found 450.1943.



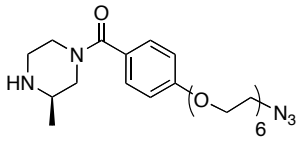
(**S5b**) Prepared by the same method as described for **S5a**, starting from **S4b** (320 mg, 0.72 mmol), resulting in **S5b** (258 mg, 0.60 mmol, 83%) as a clear viscous oil which was used without further purification. **IR (ATR-FTIR)** 2890 (br), 2873 (m), 2110 (s), 1713 (s), 1606 (m), 1512 (w), 1283 (w), 1255 (s), 1170 (w), 1107 (s) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 8.03 (d, J = 9.0, 2H), 6.94 (d, J = 9.0, 2H), 4.29 – 4.10 (m, 2H), 3.92 – 3.83 (m, 2H), 3.78 – 3.59 (m, 18H), 3.45 – 3.29 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 171.40, 163.54, 132.56, 122.34, 114.67, 71.21, 71.00, 71.00, 70.97, 70.91, 70.37, 69.86, 68.00, 51.01. **HRMS (ESI+)** calc'd for $C_{19}H_{30}N_3O_8$ $[M+H]^+$ *m/z* 428.2027. Found 428.2140; for $[M + Na]^+$, calc'd 450.1847, found 450.1945.



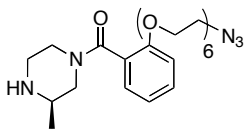
(S5c) Prepared by the same method as described for **S5a**, starting from **S4c** (220 mg, 0.50 mmol), resulting in **S5c** (199 mg, 0.466 mmol, 93%) as a clear viscous oil which was used without further purification. **IR (ATR-FTIR)** 2893 (br) 2870 (m), 2101 (s), 1727 (s), 1602 (m), 1454 (m), 1243 (s), 1097 cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 8.12 (dd, J = 1.8, 7.8, 1H), 7.52 (ddd, J = 1.8, 7.4, 8.4, 1H), 7.15 – 7.08 (m, 1H), 7.02 (d, J = 7.9, 1H), 4.42 – 4.30 (m, 2H), 3.90 (m, 2H), 3.72 – 3.57 (m, 18H), 3.37 – 3.33 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 165.96, 157.82, 135.16, 133.99, 122.86, 118.99, 113.83, 71.19, 71.09, 71.06, 71.03, 71.02, 70.98, 70.95, 70.92, 70.41, 70.39, 69.57, 69.05, 51.07. **HRMS (ESI+)** calc'd for $C_{19}H_{30}N_3O_8$ $[M+H]^+$ *m/z* 428.2027. Found 428.2062; for $[M + Na]^+$, calc'd 450.1847, found 450.1864.



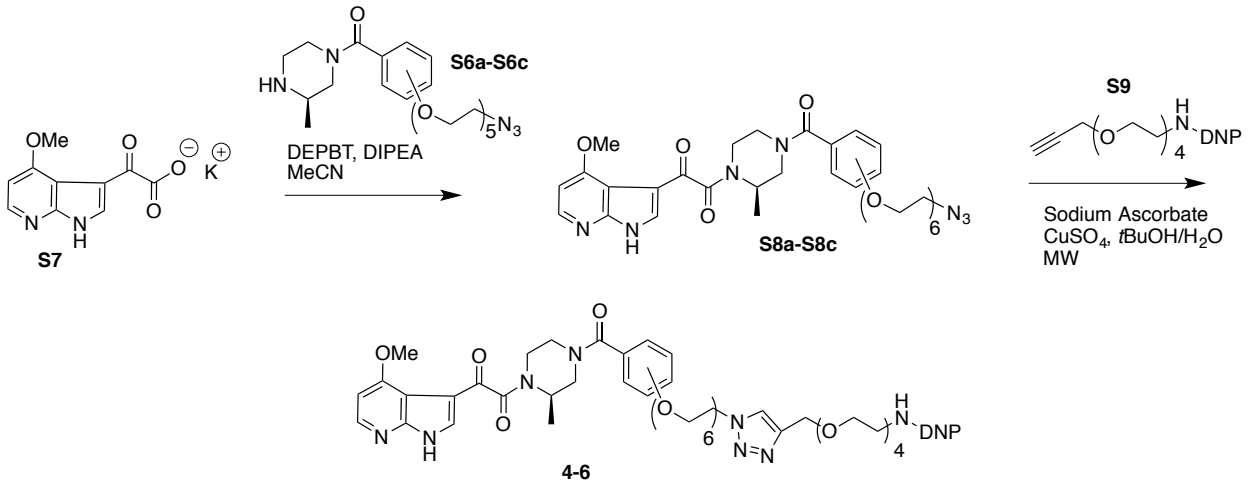
(**S6a**) To **S5a** (32 mg, 0.075 mmol) in CH_3CN (1 mL), added EDC-HCl (28 mg, 0.15 mmol, 2 equiv), HOBT (23 mg, 0.15mmol, 2 equiv) followed by (*R*)-2-methylpiperazine (7.5 mg , 0.075 mmol, 1 equiv) in CH_3CN (1 mL) dropwise. Let reaction stir under an atmosphere of nitrogen until TLC (10:1 CH_2Cl_2/MeOH) indicated reaction completion after 1.5 hr. All volatiles removed by rotary evaporation and residue was resuspended in CH_2Cl_2 (10mL) and washed with sat. $NaHCO_3$ (2 x 10 mL), followed by brine (2 x 10 mL). The organic layer was dried over anhyd. $MgSO_4$, filtered and all solvents were evaporated. Crude **S6a** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 4g pre-packed dry loading column. Run using 100% CH_2Cl_2 to 10% MeOH:CH_2Cl_2 gradient over 30 column volumes) to yield **S6a** as a colorless sticky residue (32 mg, 0.063 mmol 84%). (**IR (ATR-FTIR)** 3432 (br), 2933 (m), 2101 (s), 1649 (s), 1496 (w), 1387 (s), 1254 (m), 1096 (s) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.29 (t, J = 7.8, 1H), 6.95 (m, 3H), 4.57 (br s, 1H), 4.20 – 4.06 (m, 2H), 3.91 – 3.83 (m, 2H), 3.79 – 3.47 (m, 19H), 3.38 (m, 2H), 2.78 (m, 6H), 1.05 (m, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 170.64, 159.55, 137.96, 130.29, 119.94, 116.73, 113.77, 71.53, 71.39, 71.37, 71.32, 71.28, 70.72, 70.34, 68.27, 51.37. **HRMS (ESI+)** calc'd for $C_{24}H_{40}N_5O_7$ $[M+H]^+$ *m/z* 510.2922. Found 510.3221; for $[M + Na]^+$, calc'd 532.2742, found 532.2897.



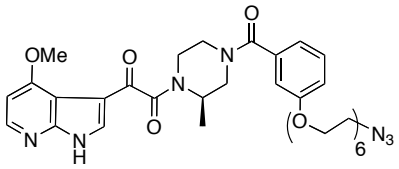
(**S6b**) Prepared by the same method as described for **S6a**, starting from **S5b** (100 mg, 0.23 mmol). Crude **S6b** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 10% MeOH in DCM gradient over 30 column volumes to yield **S6b** as a clear sticky residue (31 mg, 45%). **IR (ATR-FTIR)** 3494 (br), 2869 (m), 2102 (s), 1606 (s), 1435 (m), 1248 (m), 1104 (s), 945 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 7.33 (d, J = 6.8, 2H), 6.90 (d, J = 6.9, 2H), 4.51 (br s, 1H), 4.12 (m, 2H), 3.84 (m, 2H), 3.78 – 3.53 (m, 20H), 3.36 (m, 2H), 2.87 (m, 5H), 1.03 (br s, 3H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 170.21, 159.90, 129.06, 128.20, 114.36, 70.86, 70.70, 70.68, 70.63, 70.59, 70.04, 69.60, 67.51, 50.67, 19.47. **HRMS (ESI+)** calc'd for $C_{24}H_{40}N_5O_7$ $[M+H]^+$ *m/z* 510.2922. Found 510.3228; for $[M + Na]^+$, calc'd 532.2742, found 532.2911.



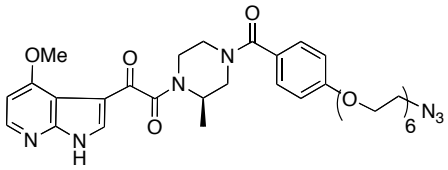
(**S6c**) Prepared by the same method as described for **S6a**, starting from **S5c** (125 mg, 0.29 mmol). Crude **S6c** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 10% MeOH in DCM gradient over 30 column volumes to yield **S6c** as a clear sticky residue (32 mg, 22%) **IR (ATR-FTIR)** 3457 (br), 2870 (m), 2103 (s), 1600 (s), 1447 (m), 1247 (m), 1103 (s), 943 (m) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.35 – 7.28 (m, 1H), 7.21 (m, 1H), 6.98 (m, 1H), 6.93 – 6.85 (m, 1H), 4.71 – 4.47 (m, 1H), 4.25 – 4.02 (m, 2H), 3.89 – 3.72 (m, 2H), 3.71 – 3.58 (m, 18H), 3.42 – 2.34 (m, 9H), 1.07 (apparent ddd, J = 6.3, 35.0, 44.8, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 167.79, 167.48, 154.39, 154.33, 154.31, 130.37, 130.35, 130.31, 128.27, 128.25, 127.91, 127.76, 126.15, 126.12, 121.45, 121.29, 112.22, 112.19, 112.02, 70.82, 70.79, 70.77, 70.68, 70.63, 70.61, 70.55, 70.53, 70.38, 70.35, 70.32, 69.96, 69.74, 69.70, 69.53, 68.04, 67.95, 67.90, 53.61, 53.40, 51.06, 50.96, 50.83, 50.65, 50.53, 48.21, 47.78, 46.74, 46.69, 46.68, 45.59, 45.48, 45.10, 41.79, 41.28, 19.23, 19.05, 18.75, 18.49. Note: conformational isomerism observed during NMR experiments, resulting in apparent signal multiplicity. **HRMS (ESI+)** calc'd for $C_{24}H_{40}N_5O_7$ $[M+H]^+$ *m/z* 510.2922. Found 510.3275; for $[M + Na]^+$, calc'd 532.2742, found 532.2910.



Scheme S2. General synthesis of BMS-3778806 ARM-Hs with linker attachment on the phenyl ring, starting from the known potassium salt **S7**.

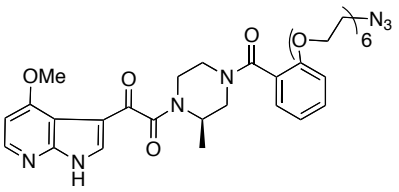


(**S8a**) Potassium (4-methoxy-7-azaindol-3-yl)-oxoacetate (**S7**)[1] (21 mg, 0.084 mmol), **S6a** (43 mg, 0.084 mmol), 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(*3H*)-one (DEPBT) (26.5 mg, 0.089 mmol) and Hunig's base (11.6 mg, 0.090 mmol) were combined in DMF (1 ml) and the mixture stirred at room temperature under nitrogen for overnight. The DMF was removed in *vacuo*. Crude **S8a** was purified by column chromatography (0% to 20% MeOH in DCM) to deliver 25 mg (42%) of **S8a** as a clear colorless oil. **IR (ATR-FTIR)** 3453 (br), 3114 (w), 2872 (m), 2112 (s), 1634 (s), 1583 (m), 1439 (m), 1300 (m), 1100 (m), 946 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 13.13 (s, 1H), 8.29 (d, *J* = 5.6 Hz, 1H), 8.09 (d, *J* = 17.0 Hz, 1H), 7.31 (s, 1H), 7.02 – 6.87 (m, 3H), 6.74 (d, *J* = 5.7 Hz, 1H), 5.07 – 4.21 (m, 2H), 4.18 – 4.08 (m, 2H), 4.08 – 3.99 (m, 4H), 3.85 (s, 3H), 3.76 – 3.55 (m, 18H), 3.54 – 3.32 (m, 3H), 3.27 – 3.07 (m, 2H), 1.50 – 1.08 (m, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.22, 166.96, 166.93, 159.38, 150.75, 145.98, 136.81, 136.77, 136.75, 135.44, 135.28, 135.26, 130.25, 119.73, 119.68, 116.85, 116.80, 114.56, 113.71, 108.25, 108.21, 101.36, 71.24, 71.18, 71.09, 71.06, 71.02, 70.97, 70.93, 70.42, 70.03, 68.04, 56.61, 56.56, 51.09, 45.27, 16.53, 15.59, 15.57. **HRMS (ESI+)** calc'd for [M+H]+ $C_{34}H_{46}N_7O_{10}$ *m/z* 712.3301. Found 712.3453

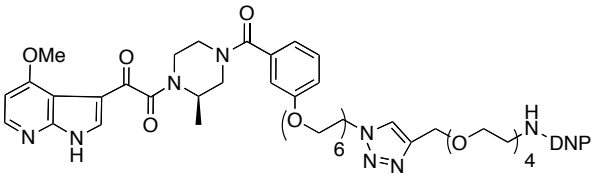


(**S8b**) Prepared by the same method as described for **S8a**, starting from **S7** (25 mg, 0.097 mmol) and **S6b** (51 mg, 0.097 mmol). Crude **S8b** was was purified by flash chromatography (CombiFlash Automated Chromatographer, 4g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 20% MeOH in DCM gradient over 100 column volumes to yield **S8b** as a clear sticky residue (31 mg, 45%). **IR (ATR-FTIR)** 3376 (br), 3105 (w), 2889 (m), 2100 (m), 1635 (s), 1514 (m), 1436 (m), 1300 (m), 1098 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 12.85 (s, 1H), 8.29 (d, *J* = 5.7 Hz, 1H), 8.08 (d, *J* = 16.8 Hz, 1H), 7.37 (d, *J* = 8.7 Hz, 2H), 6.93 (d, *J* = 8.5 Hz, 2H), 6.74 (d, *J* = 5.8 Hz, 1H), 5.01 – 4.22 (m, 2H), 4.20 – 4.09 (m, 4H), 4.05 (d, *J* = 4.7 Hz, 3H), 3.91 – 3.80 (m, 2H), 3.76 – 3.57 (m, 18H), 3.55 – 3.31 (m, 3H), 3.30 – 2.92 (m, 2H), 1.41 – 1.15 (m, 3H). **^{13}C NMR (125 MHz, CDCl3)** δ 184.96, 184.83, 166.61, 161.60, 161.45, 160.24, 151.34, 151.31, 146.53, 146.44, 134.77, 134.57, 129.11, 127.24, 127.20, 114.59, 114.53, 114.06, 100.91, 70.84, 70.67, 70.64, 70.59, 70.55, 70.00, 69.54, 67.53, 56.00, 55.95, 50.65, 49.87, 44.88, 16.09, 15.15. **HRMS (ESI+)** calc'd for [M+H]+ $C_{34}H_{46}N_7O_{10}$ *m/z* 712.3301. Found 712.3230; for [M + Na]+, calc'd 734.3120, found 734.3057.

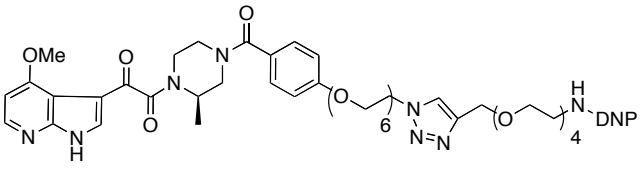
(**S8c**) Prepared by the same method as described for **S8a**, starting from **S7** (10 mg, 0.039 mmol) and **S6c** (20 mg, 0.039 mmol). Crude **S8c** was purified by flash chromatography (CombiFlash Automated Chromatographer, 4g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 20% MeOH in DCM



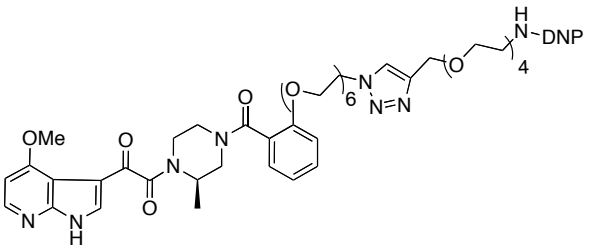
gradient over 100 column volumes to yield S8b as a clear sticky residue (15 mg, 54%) **IR (ATR-FTIR)** 3401 (br), 3093 (w), 2866 (m), 2087 (m), 1632 (s), 1439 (m), 1300 (m), 1099 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 11.79 (br s, 1H), 8.35 – 8.19 (m, 1H), 8.01 (m, 1H), 7.41 – 7.17 (m, 2H), 7.10 – 6.63 (m, 3H), 5.10 – 4.49 (m, 2H), 4.47 – 2.81 (m, 32H), 1.52 – 1.13 (m, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 184.88, 184.79, 168.33, 166.82, 161.48, 161.39, 154.33, 154.17, 151.41, 146.74, 146.72, 146.71, 146.69, 136.01, 135.76, 130.85, 130.80, 130.54, 130.47, 128.62, 128.00, 125.59, 121.39, 113.83, 113.62, 111.78, 111.50, 111.46, 111.30, 70.63, 70.54, 70.49, 70.47, 70.42, 70.40, 70.36, 70.20, 70.15, 70.06, 69.90, 69.78, 69.64, 69.52, 69.50, 69.42, 67.76, 67.63, 67.53, 55.98, 55.86, 55.84, 50.60, 50.57, 50.02, 47.25, 46.85, 46.21, 45.73, 44.68, 41.51, 36.23, 35.57, 16.36, 15.19.



(6) To a solution of **S8a** (12 mg, 0.017 mmol) dissolved in *t*-BuOH (0.5 ml) and water (0.5 ml) in a microwave reaction vessel, was added **S9**[2] (6.8 mg, 0.017 mmol). The mixture was stirred for 5 min before aqueous $CuSO_4$ (0.1 M, 8.5 µl) and aqueous sodium ascorbate (0.1 M, 17 µl) were added. The reaction vessel was then capped and heated in a microwave reactor for 25 min at 125 °C. Solvents were evaporated. The product was purified using HPLC (20%-60% B over 45 min) to give pure **6** (15.5 mg, 82%). **IR (ATR-FTIR)** 3375 (br), 3119 (w), 2874 (m), 1608 (s), 1515 (m), 1465 (m), 1249 (m), 1016 (m), 953 (w) cm^{-1}. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 11.39 (s, 2H), 9.10 (d, *J* = 2.6 Hz, 1H), 8.79 (s, 1H), 8.28 (d, *J* = 6.5 Hz, 1H), 8.24 (dd, *J* = 9.5, 2.6 Hz, 1H), 8.13 (d, *J* = 13.3 Hz, 1H), 7.80 (s, 1H), 7.32 (s, 1H), 7.04 – 6.86 (m, 5H), 5.05 – 4.35 (m, 7H), 4.33 – 4.03 (m, 4H), 3.92 – 3.35 (m, 41H), 3.32 – 2.87 (m, 2H), 1.59 – 1.08 (m, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 190.78, 183.77, 165.58, 161.87, 161.81 159.1, 148.6, 144.1, 137.9, 136.2, 130.4, 130.1, 124.4, 116.8, 115.3, 114.3, 113.5, 101.2, 70.8, 70.6 70.5, 70.4, 70.4, 70.4, 69.8, 69.7, 69.3, 68.8, 67.74, 64.0, 57.8, 50.8, 43.3, 16.36, 15.19. **HRMS (ESI+)** calcd for $C_{51}H_{69}N_{10}O_{18}^+$ $[M+H]^+$ *m/z* 1109.4786, found 1109.4825.

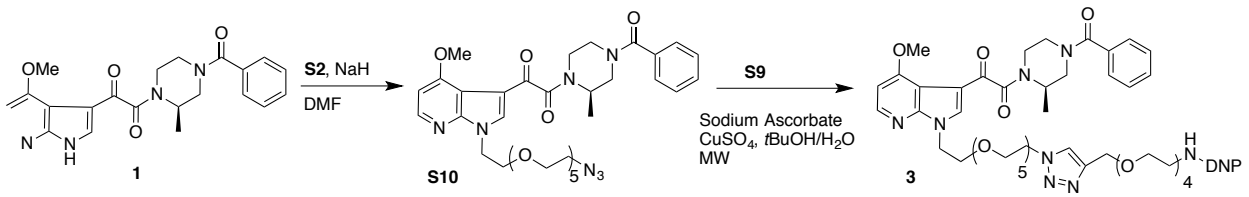


(**5**) Prepared by the same method as described for **6**, starting from **S8b** (7 mg, 0.01 mmol). Crude **S4** was purified using HPLC (20%-60% B over 45 min) to give pure **5** (8 mg, 72%). **IR (ATR-FTIR)** 3355 (br), 3107 (w), 2874 (m), 1622 (s), 1559 (m), 1431 (m), 1336 (m), 1133 (s), 951 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 14.89 (br s, 1H), 9.12 (d, J = 2.7, 1H), 8.79 (br s, 1H), 8.37 – 8.18 (m, 2H), 8.11 (d, J = 11.8, 1H), 7.77 (s, 1H), 7.38 (d, J = 8.7, 2H), 7.05-6.85 (m, 4H), 4.67 (s, 2H), 4.55-4.40 (m, 3H), 4.27-2.82 (m, 42H), 1.29 (br s, 3H). **^{13}C NMR (125 MHz, DMSO-*d_6*)** δ 184.98, 184.89, 169.94, 169.91, 166.38, 166.18, 160.27, 160.18, 159.43, 151.16, 148.29, 146.79, 146.72, 143.69, 135.64, 135.55, 134.80, 129.79, 129.57, 129.01, 127.42, 124.14, 123.48, 115.58, 114.11, 112.51, 112.49, 106.22, 100.78, 69.83, 69.70, 69.66, 69.60, 69.52, 69.44, 68.86, 68.74, 68.61, 68.17, 67.18, 63.43, 55.61, 55.57, 49.30, 49.21, 42.57, 15.48, 14.82. **HRMS (ESI+)** calcd for $C_{51}H_{69}N_{10}O_{18}^+$ $[M+H]^+$ *m/z* 1109.4786, found 1109.4869.

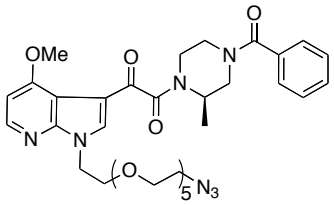


(4) Prepared by the same method as described for **6**, starting from **S8c** (15 mg, 0.02 mmol). Crude **S3** was purified using HPLC (20%-60% B over 45 min) to give pure **4** (13 mg, 59%). **^{1}H NMR (500 MHz, $CDCl_3$)** δ 11.39 (br signal, 1H), 9.28 – 9.03 (m, 1H), 8.87 – 8.60 (m, 1H), 8.44 – 7.65 (m, 4H), 7.44 – 7.29 (m, 2H), 7.11 – 6.61 (m, 4H), 5.10 – 2.88 (m, 52H), 1.54 – 1.04 (m, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.60, 169.27, 168.88, 167.21, 167.05, 154.81, 154.61, 148.84, 145.05, 136.42, 130.94, 130.62, 129.08, 128.34, 126.18, 125.47, 124.73, 124.49, 122.10, 114.66, 71.19, 71.11, 70.94, 70.89, 70.84, 70.80, 70.77, 70.74, 70.65, 70.61, 70.56, 70.42, 70.28, 70.11, 69.99, 69.84, 69.77, 68.99, 64.81, 56.40, 56.27, 50.61, 50.55, 50.51, 50.45, 47.63, 47.23, 46.59, 43.66, 36.59, 30.06, 26.15, 16.76, 15.69. **IR (ATR-FTIR)** 3452 (br), 3109

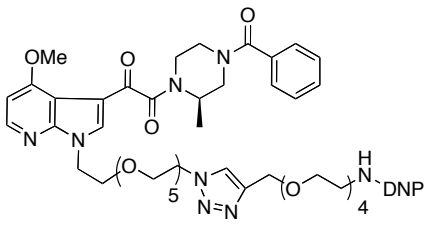
(w), 2887 (m), 1630 (s), 1512 (m), 1437 (m), 1336 (m), 1289 (m), 1122 (s) cm^{-1}. **HRMS (ESI+)** calcd for $C_{51}H_{69}N_{10}O_{18}^{+}$ $[M+H]^{+}$ *m/z* 1109.4786, found 1109.4751.



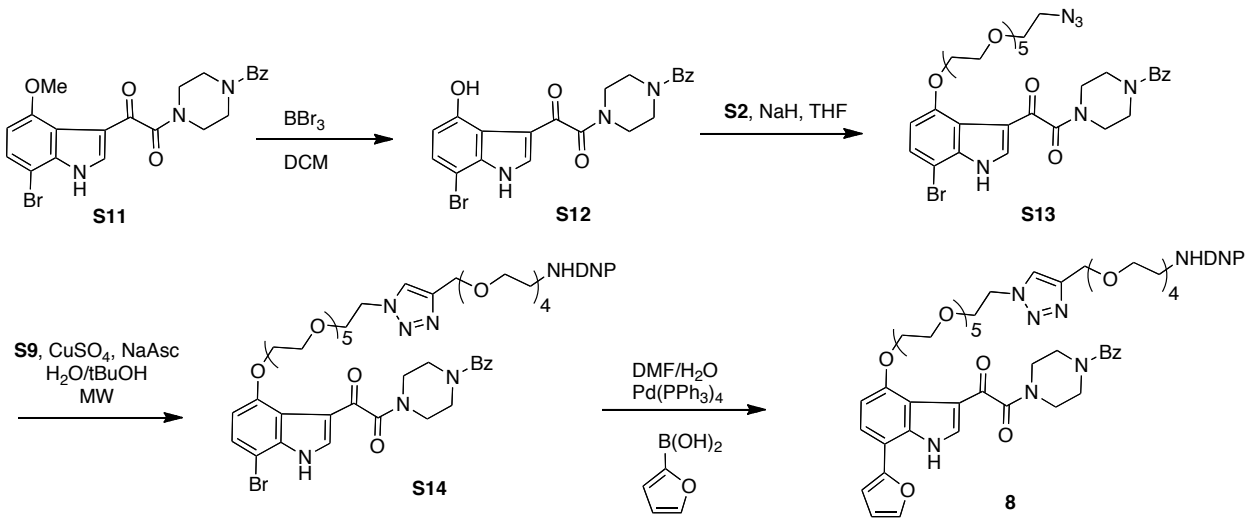
Scheme S3. General synthesis of BMS-378806 derived ARM-H (**3**) with linker attachment at the N7 position, starting from BMS-378806 (**1**)



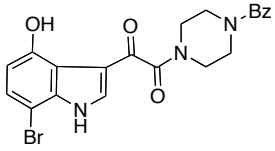
(**S10**) To a solution of BMS-378806[1] (**1**, 10.0 mg, 0.025 mmol) in anhydrous DMF (300 uL), added NaH (95%, 1.0 mg, 0.042 mmol, 1.6 equiv) and let stir under an atmosphere of nitrogen for 1 hr. when **S2** (13.7 mg, 0.30 mmol, 1.2 equiv) in anhydrous DMF (200 uL) was added via syringe. The subsequent suspension was stirred under nitrogen for 20 hr. when UPLC/MS analysis showed consumption of the limiting reagent and formation of product. Reaction was quenched with sat. NH_4Cl (8 mL) and extracted with DCM (3 x 10 mL). The combined organic layers were dried over anhyd. Na_2SO_4, filtered and dried via rotary evaporation. Crude **S10** was purified by HPLC (10-66% B over 49 min, Rt = 24.83 min), like fractions were combined and volatiles were removed via rotary evaporation, resulting in 2.1 mg (0.003 mmol, 12%) of **S10** as a clear residue. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 8.31 (s, 1H), 8.13 (d, J = 20.2, 1H), 7.46 (br signal, 5H), 6.76 (br signal, 1H), 5.03 – 2.90 (m, 34H), 1.29 (br signal, 3H). **HRMS (ESI+)** calcd for $C_{34}H_{46}N_7O_9^{+}$ $[M+H]^{+}$ *m/z* 696.3352 found 696.3368.



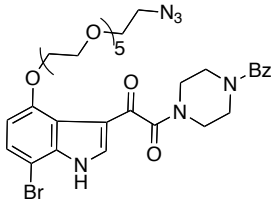
(**3**) To a microwave vial containing **S10** (2.1 mg, 0.003 mmol) in water (250 uL) and *t*BuOH (250 uL), added **S9** (1.5 mg, 0.004 mmol, 1.25 equiv), 0.1 M CuSO4 (5 uL) and 0.1 M sodium ascorbate (10 uL). Capped vial heated to 125 C for 20 min when UPLC/MS indicated formation of product and consumption of **S10**. Volatiles removed and crude **6** purified by HPLC (0-66% B, 49 min). Like fractions were combined and volatiles removed, yielding pure **3** as a yellow sticky residue (1.25 mg, 0.001 mmol, 38%). **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.13 (d, J = 2.6, 1H), 8.90 – 8.67 (m, 1H), 8.36 – 8.29 (m, 1H), 8.25 (dd, J = 2.7, 9.5, 1H), 8.10 (d, J = 19.1, 1H), 7.73 (s, 1H), 7.42 (br signal, 5H), 6.96 (d, J = 9.7, 1H), 6.79 – 6.68 (m, 1H), 4.66 (br s, 2H), 4.51 (br s, 4H), 4.17 – 2.58 (m, 46H), 1.45 – 1.19 (m, 3H). **HRMS (ESI+)** calcd for $C_{51}H_{69}N_{10}O_{17}^{+}$ $[M+H]^{+}$ *m/z* 1093.4837, found 1093.4933.



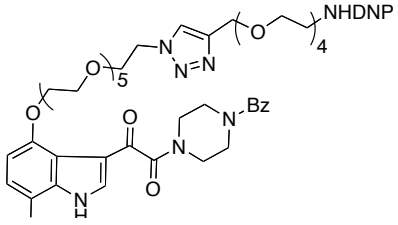
Scheme S4. Synthesis of BMS-furan derived ARM-H with linker attachment at the C4 position, starting from known indole **S11**.



(**S12**) To a flame-dried flask, added **S11** (750 mg, 1.6 mmol)[3] followed by 35 mL anhydrous CH_2Cl_2. Resulting mixture was cooled to -78 ^oC with dry ice/acetone bath and BBr_3 (1.0M in CH_2Cl_2, 13.5 mL, 13.5 mmol, 8.4 equiv) was carefully added via syringe under an atmosphere of N_2. The resulting purple mixture was allowed to warm to RT over a period of 2 hr and then heated to reflux until indicated no remaining starting material (72 hrs). The mixture was allowed to cool to RT and carefully quenched with aq. NaOH (0.5M, 50 mL) and extracted with EtoOAc (3 x 200 mL). The combined organic layers were then washed with saturated $NaHCO_3$ and brine, dried over anhydrous $MgSO_4$, filtered and all solvents were evaporated, yielding crude S**11** as a red solid. Crude **S12** purified by flash chromatography (3:1 hexanes/acetone → 1:1 → 100% acetone). Compound can be further purified by washing several times with a 1:1 mixture of DCM/hexanes, resulting in pure S**12** as an off yellow powder (218 mg, 0.48 mmol, 30%). **IR (ATR-FTIR)** 3301 (br), 2868 (w), 1627 (s), 1510 (w), 1410 (m), 1286 (m), 1101 (m), 1008 (w) cm^{-1}. **^{1}H NMR (500 MHz, DMSO)** δ 13.05 (s, 1H), 10.75 (s, 1H), 8.42 (br s, 1H), 7.55 (s, 5H), 7.48 (d, J = 8.2, 1H), 6.69 (d, J = 7.7, 1H), 4.00 – 3.51 (m, 8H), 3.45 (s, 3H). **^{13}C NMR (125 MHz, DMSO)** δ 188.13, 169.21, 164.45, 150.65, 139.77, 136.99, 135.42, 129.66, 128.56, 128.39, 126.94, 115.83, 114.74, 109.21, 93.98. **HRMS (ESI+)** calcd for $C_{21}H_{19}BrN_3O_4^+$ $[M+H]^+$ *m/z* 456.0553 and 458.0538, found 456.0560 and 458.0541.

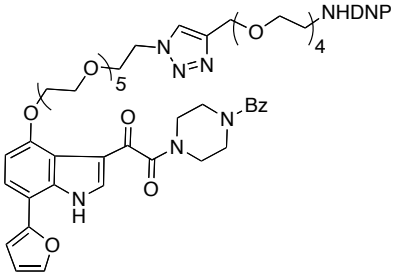


(**S13**) To a flame-dried flask containing a solution of NaH (>95%, 1.1 mg, 0.046 mmol) in anhydrous THF (1.0 mL), added **S2** (10 mg, 22.0 μmol) n THF (1.0 mL) followed by slow addition of **9** in THF (1.0 mL). The resulting green mixture was allowed to stir at RT under an atmosphere of N_2. After 12, 24 and 36 hr, carefully added addition NaH (2 mg each addition). After 5 days, TLC (1:1 hexanes/acetone) indicated reaction completion. Reaction was quenched with H_2O (5 mL) and extracted with CH_2Cl_2 (3 x 10 mL). The combined organic layers were washed with brine (30 mL), dried over anhydrous over anhydrous $MgSO_4$, filtered and all solvents were evaporated. Crude **S13** was purified by prepatory thin layer chromatography (AnalTech Uniplate 1000 μm; eluting with 1:1 hexanes/acetone), resulting in **S13** as a clear residue (6.0 mg, 0.008 mmol, 40%). **IR (ATR-FTIR)** 3095 (w), 2870 (m), 2106 (s), 1631 (s), 1513 (m), 1431 (m), 1251 (m), 1098 (m), 843 (w) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 8.01 (s, 1H), 7.37 (d, J = 23.6, 5H), 7.25 (d, J = 8.3, 1H), 6.56 (d, J = 8.5, 1H), 4.22 (s, 2H), 3.92 – 3.23 (m, 30H)**. ^{13}C NMR (125 MHz, $CDCl_3$)** δ 171.09, 167.21, 153.13, 135.46, 130.58, 129.09, 127.69, 127.52, 116.48, 107.33, 70.92, 70.58, 70.41, 70.36, 70.32, 70.25, 70.03, 69.23, 50.94. **HRMS (ESI+)** calcd for $C_{33}H_{42}BrN_6O_9^+$ $[M+H]^+$ *m/z* 745.2197 and 747.2176. Found 745.2057 and 747.2036.

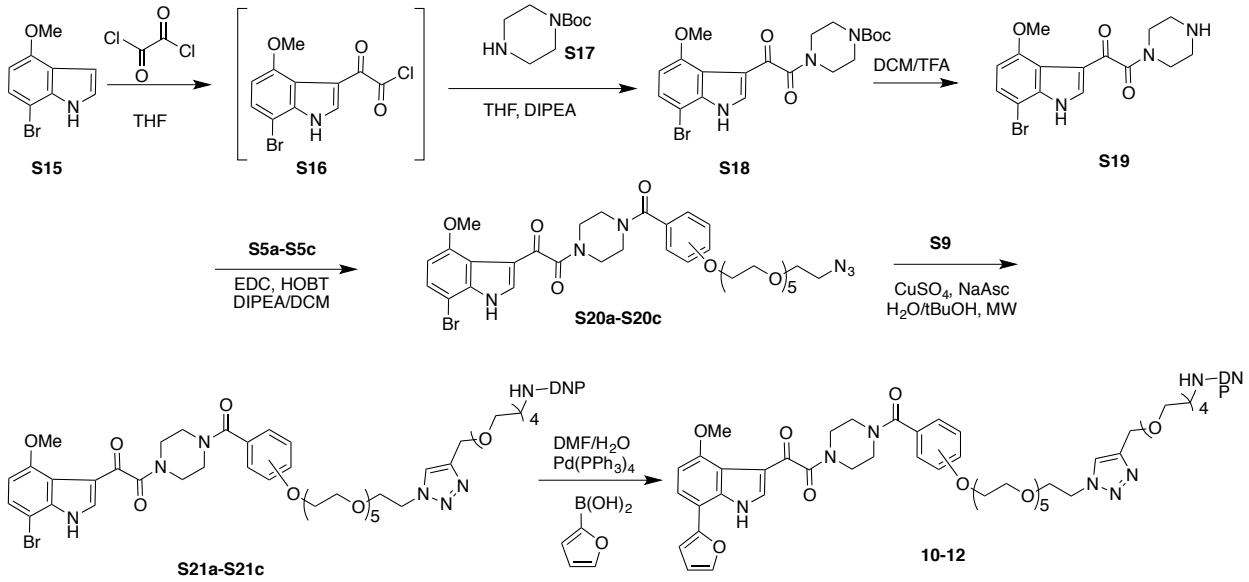


(**S14**) To a solution of **S13** (3.5 mg, 4.7 μmol) in *t*BuOH (260 μL) and H_2O

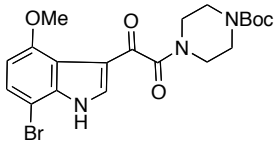
(250 μL) in a microwave vial, added **S9** (3.6 mg, 9 μmol, 2 equiv), 0.1 M $CuSO_4$ (3 μL) and 0.1M sodium ascorbate (6 μL). Heated to 125 °C in microwave reactor for 30 min when UPLC/MS indicated formation of product and consumption of **S13**. Volatiles were removed and crude **S13** was purified by flash chromatography (CombiFlash Automated Chromatographer, 4g column, dryloaded with pre-packed dry loading column. Run using 0% MeOH in DCM to 10% MeOH in DCM gradient over 80 column volumes to yield **S14** as a clear sticky residue (4 mg, 3.5 μmol 74%). **IR (ATR-FTIR)** 3356 (w), 2869 (m), 1621 (s), 1507 (m), 1428 (m), 1302 (m), 1116 (m), 832 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.73 (s, 1H), 9.09 (d, J = 2.7, 1H), 8.74 (s, 1H), 8.23 (d, J = 9.5, 1H), 7.98 (s, 1H), 7.61 (s, 1H), 7.42 (s, 5H), 7.30 (d, J = 8.5, 1H), 6.91 (d, J = 9.6, 1H), 6.62 (d, J = 8.5, 1H), 4.62 (s, 2H), 4.46 (t, J = 5.1, 2H), 4.26 (m, 2H), 3.91-3.54 (m, 44H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.27, 170.62, 166.82, 152.97, 148.37, 136.72, 135.99, 135.12, 134.92, 130.41, 130.17, 130.11, 128.65, 127.09, 127.04, 124.21, 116.44, 116.14, 114.12, 106.53, 96.63, 70.79, 70.65, 70.62, 70.55, 70.49, 70.41, 69.64, 69.49, 69.30, 68.85, 68.53, 50.25, 43.20. **HRMS (ESI+)** calcd for $C_{50}H_{65}BrN_9O_{17}^+$ $[M+H]^+$ *m/z* 1142.3682 and 1144.3661. Found 1142.3701 and 1144.3665.



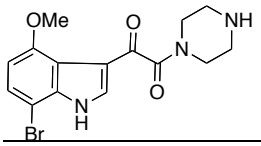
(**8**) To a solution of **S14** (3 mg, 2.63 μmol) in 300 μL dimethylformamide (DMF) in a microwave vial, added $NaHCO_3$ (0.287 mg, 3.4 μmol, 1.3 equiv; in 185 μL H_2O at 0.0185M) and 2-furanylbornic acid (0.4 mg, 3.4 μmol, 1.3 equiv). Removed O_2 from solution by bubbling with N_2 for at least 10 min. Carefully added $Pd(PPh_3)_4$ (0.15 mg, 0.13 μmol, 5 mol%), capped vial and heated in a microwave reactor for 12 min at 150 °C. Evaporated all solvents and purified crude residue by HPLC (0-60% B gradient, over 60 min). Yielded **8** as a yellow solid (2.1mg, 1.8 μmol 68%). In addition to ^{1}H NMR, purity affirmed by **UPLC/MS** t_R 1.57 min, $[M+H]^+$ *m/z* 1130.9. **IR (ATR-FTIR)** 3365 (w), 2916 (m), 2848 (m), 1621 (s), 1512 (m), 1426 (m), 1300 (m), 1134 (m), 833 (w) cm^{-1} **^{1}H NMR (500 MHz, $CDCl_3$)** δ 10.51 – 10.07 (m, 1H), 9.08 (d, J = 2.6, 1H), 8.82 – 8.67 (m, 1H), 8.21 (dd, J = 2.3, 9.6, 1H), 8.04 (s, 1H), 7.68 – 7.60 (br s, 1H), 7.58 – 7.53 (br s, 1H), 7.49 – 7.37 (m, 5H), 6.90 (d, J = 9.5, 1H), 6.77 (d, J = 8.2, 1H), 6.69 (d, J = 3.0, 1H), 6.55 (dd, J = 1.5, 3.2, 1H), 4.66 – 4.57 (m, 2H), 4.49 – 4.38 (m, 2H), 4.35 – 4.26 (m, 2H), 3.97-3.24 (m, 44H). **HRMS (ESI+)** calcd for $C_{54}H_{68}N_9O_{18}^+$ $[M+H]^+$ *m/z* 1130.4677. Found 1130.4617.



Scheme S5. General synthesis BMS-furan derived ARM-Hs with linker attachment on the phenyl ring.



(**S18**) To a flame-dried flask containing **S15**[1] (50.0 mg, 0.233 mmol) in anhyd. THF (600 μL), added oxalyl chloride (97 μL, 1.11 mmol, 5 equiv) and let stir under an atmosphere of N_2 until TLC (5:1 hexanes/EtOAc) indicated consumption of starting material (5-12 hr, depending on scale). All volatiles were removed by rotoevaporation and resulting green residue (**S16**) was immediately suspended in anhyd. THF (1 mL), followed by the addition of N-Boc piperazine[2] (**S17**, 52 mg, 0.28 mmol, 1.2 equiv) and DIPEA (78 μL, 2 equiv). Resulting mixture was stirred under an atmosphere of N_2 at RT for 12 hr. and then at reflux for 30 min (if needed) when TLC (5:1 hexanes/EtOAc) indicated reaction completion. Reaction was allowed to cool to RT, poured into H_2O (10 mL) and extracted with EtOAc (3 x 10 mL). The combined organic layers were dried over anhyd. $MgSO_4$, filtered, and all solvents were evaporated. Crude **S** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 4g pre-packed dry loading column. Run using 100% Hexanes to 50% EtOAc:Hexanes gradient over 30 column volumes, followed by EtOAc flush) to yield **S18** as a light brown powder (78 mg, 75%). **IR (ATR-FTIR)** 3205 (br), 3027 (w), 2976 (w), 1693 (s), 1645 (m), 1419 (m), 1287 (w), 1166 (s), 1104 (w), 1001 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.44 (s, 1H), 7.94 (d, J = 3.1, 1H), 7.28 (d, J = 8.5, 1H), 6.56 (d, J = 8.5, 1H), 3.90 (s, 3H), 3.71 (m, 2H), 3.60 – 3.51 (m, 2H), 3.46 (m, 4H), 1.47 (s, 9H). **: (ESI+)** *m/z* $[M+H]^+$ 466; [M + Na]+ 488; Rt = 1.34. **^{13}C NMR (100 MHz, CDCl3)** δ 186.41, 167.17, 154.69, 153.55, 136.34, 133.69, 126.84, 117.11, 115.90, 104.95, 96.65, 80.65, 56.04, 45.96, 41.22, 28.49. **HRMS (ESI+)** calc'd for $C_{20}H_{27}BrN_3O_5^+$ $[M+H]^+$ *m/z* 467.1050 and 469.1035. Found 467.1173 and 469.1058.



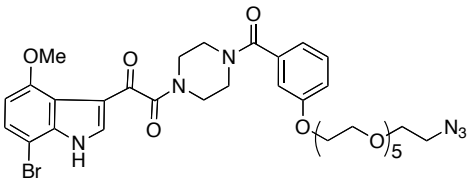
(**S19**) To a solution of **S18** (75 mg, 0.161 mmol) in CH_2Cl_2 (1.5 mL), added trifluoroacetic acid (TFA) (0.5 mL), resulting in an immediate color change from clear to yellow. The resulting solution was stirred at RT for 30 min when TLC (20:1 CH_2Cl_2/CH_3OH) indicated complete starting material consumption. All volatiles were

1 . US PATENT: US20030069245

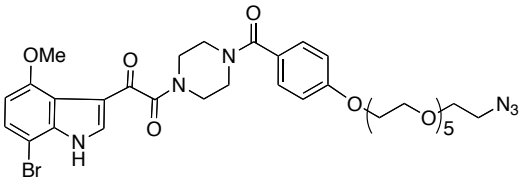
2. Faust, A.; Waschkau, B.; Waldeck, J.; Holtke, C.; Breyholtz, H.; Wagner, S.; Kopka, K.; Heindel, W.; Schafer, M.; Bremer, C. *Bioconjug. Chem.* **2008**, *19*, 1001-1008.

Faust, A.; Waschkau, B.; Waldeck, J.; Holtke, C.; Breyholtz, H.; Wagner, S.; Kopka, K.; Heindel, W.; Schafer, M.; Bremer, C. *Bioconjug. Chem.* **2008**, *19*, 1001-1008.

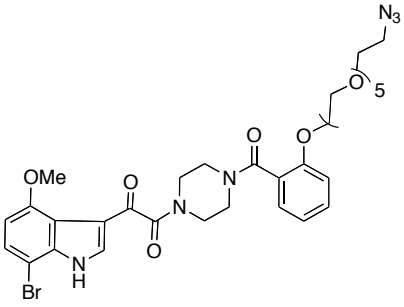
removed by rotoevaporation and residue was redissolved in CH_2Cl_2 (15 mL) and NaOH (2M, until a pH of 11 is achieved) and then extracted with CH_2Cl_2 (3 x 15 mL). The combined organic layers were dried over anhyd. $MgSO_4$, filtered, and all solvents were evaporated, resulting in **S19** as an off-white solid (51 mg, 86%). **IR (ATR-FTIR)** 3334 (br), 2934 (m) 2871 (w), 1640 (m), 1509 (m), 1437 (m), 1285 (m), 1171 (m), 1093 (m) cm^{-1}**.** **^{1}H NMR (500 MHz, $CDCl_3$)** δ 7.75 (s, 1H), 7.05 (d, J = 8.4, 1H), 6.35 (d, J = 8.5, 1H), 4.29 (s, 1H), 3.63 (s, 3H), 3.52 – 3.42 (m, 2H), 3.26 – 3.15 (m, 2H), 3.08 – 3.00 (m, 1H), 2.78 – 2.68 (m, 2H), 2.67 – 2.56 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 190.40, 171.60, 157.34, 141.04, 139.66, 139.63, 130.79, 119.67, 119.25, 108.64, 100.80, 59.59, 52.81, 52.64, 52.47, 52.30, 52.13, 51.96, 51.79, 50.13, 48.82, 48.48, 45.20. **HRMS (ESI+)** calc'd for $C_{15}H_{17}BrN_3O_3^+$ $[M+H]^+$ *m/z* 366.0448 and 368.0433. Found 366.0564 and 368.0650.



(**S20a**) To a flame-dried flask containing a solution of **S19** (20 mg, 0.055 mmol) in CH_2Cl_2 (2 mL), added **S5a** (25.8 mg, 0.06 mmol, 1.1 equiv), EDC-HCl (11.5 mg, 0.06 mmol, 1.1 equiv), HOBT (9.2 mg, 0.06 mmol, 1.1 equiv) and DIPEA (30 uL, 0.16mmol, 3 equiv). Resulting mixture was stirred at RT under an atmosphere of N_2 at RT for 8 hr when TLC (9:1 CH_2Cl_2/CH_3OH) indicated reaction completion. Mixture was diluted with CH_2Cl_2 (10 mL) and washed with sat. $NaHCO_3$ (15 mL), sat. NH_4Cl (15 mL) and brine (15 mL). The combined organic layers were dried over anhyd. $MgSO_4$, filtered, and all solvents were evaporated, resulting in crude **S20a** as a sticky solid. Crude **S20a** was purified by flash chromatography (CombiFlash Automated Chromatographer, 4g column, dryloaded with 4g pre-packed dry loading column. Run using 100% CH_2Cl_2 to 10% CH_3OH in CH_2Cl_2 gradient over 40 column volumes) to yield **S20a** as a clear sticky solid (38 mg, 90%). **IR (ATR-FTIR)** 3558 (br), 3111 (w), 2869 (m), 2107 (s), 1634 (s), 1579 (m), 1438 (m), 1286 (m), 1127 (s), 1093 (s), 969 (w) cm^{-1}**.** **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.66 (s, 1H), 7.96 (d, J = 3.1, 1H), 7.34 (br peak, 1H), 7.29 (d, J = 8.4, 1H), 6.96 (br peak, J = 7.5, 3H), 6.57 (d, J = 8.5, 1H), 4.12 (br peak, 2H), 3.91 (s, 3H), 3.84 (s, 4H), 3.74 – 3.39 (m, 24H), 3.36 (t, J = 5.0, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.43, 170.83, 167.67, 159.35, 153.86, 136.96, 136.68, 134.98, 130.25, 127.23, 119.67, 116.84, 116.65, 116.07, 113.72, 105.24, 97.15, 71.20, 71.06, 71.04, 70.99, 70.95, 70.40, 70.01, 68.00, 56.54, 51.06. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{34}H_{45}BrN_6O_{10}^+$ *m/z* 775.2302 and 777.2282. Found 775.2197 and 777.2182.

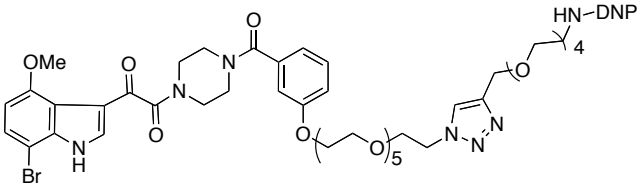


(S20b) Prepared **S20b** in the same manner as compound **S20a** from **S19** and **S5b** in 88% yield. **IR (ATR-FTIR)** 3502 (br), 3150 (w), 2865 (m), 2160 (s), 1630 (s), 1512 (m), 1431 (m), 1287 (m), 1248 (s), 1115 (s), 1005 (w) cm^{-1}**.** **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.72 (s, 1H), 7.95 (s, 1H), 7.38 (d, J = 8.6, 2H), 7.28 (d, J = 8.4, 1H), 6.91 (d, J = 8.6, 2H), 6.56 (d, J = 8.6, 1H), 4.18 – 4.10 (m, 2H), 3.90 (s, 3H), 3.87 – 3.81 (m, 2H), 3.80 – 3.41 (m, 26H), 3.40 – 3.32 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.05, 170.62, 167.17, 160.31, 153.47, 148.37, 144.75, 136.54, 135.94, 134.55, 130.38, 130.14, 129.20, 127.15, 126.77, 124.15, 123.78, 116.29, 115.66, 114.51, 114.15, 104.75, 96.70, 70.80, 70.65, 70.60, 70.57, 70.55, 70.51, 70.48, 70.46, 70.42, 69.60, 69.52, 69.37, 68.55, 67.55, 64.46, 56.09, 50.17**.** **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{34}H_{45}BrN_6O_{10}^+$ *m/z* 775.2297 and 777.2282. Found 775.2408 and 777.2392.

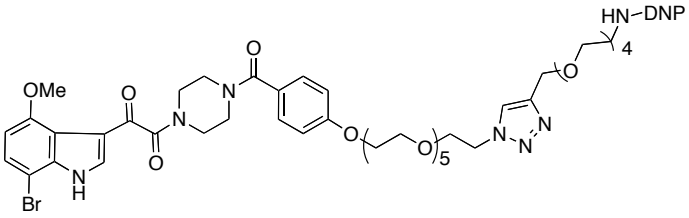


(S20c) Prepared **S20c** in the same manner as compound **S20a** from **S19** and **S20c** in 81% yield. **IR (ATR-FTIR)** 3488 (br), 3115 (w), 2873 (m), 2109 (s), 1629 (s), 1513 (m), 1435 (m), 1355 (w), 1186 (s), 1110 (s), 1007 (w) cm^{-1}**.** **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.11 (d, J = 2.6, 1H), 8.78 (s, 1H), 8.23 (dd, J = 2.4, 9.5, 1H), 7.68 (s, 1H), 7.41-7.14 (m, 1H), 7.09-6.72 (m, 2H), 6.41 (s, 1H), 4.64 (s, 2H), 4.41 (s, 2H), 4.29-3.98 (m, 2H), 3.92 – 3.05 (m, 27H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.42, 185.35, 167.98, 167.92, 167.01, 166.98, 154.38, 154.18, 153.96, 153.95, 136.87, 136.85, 136.02, 136.01, 135.86, 130.74, 130.67, 128.21, 128.11, 127.09, 127.04, 125.54, 125.46, 121.49, 121.42, 116.24, 116.00, 115.88, 115.83, 111.89, 111.57, 104.95, 104.90, 100.01, 96.50, 96.44, 70.70, 70.67, 70.63, 70.59, 70.56, 70.52, 70.50, 70.48, 70.41, 70.31, 70.22, 70.18, 70.11, 69.99, 69.95, 69.91, 69.66, 69.55, 67.74, 67.52, 56.20, 56.15, 50.65, 50.58, 50.56, 47.02, 46.49, 46.11,

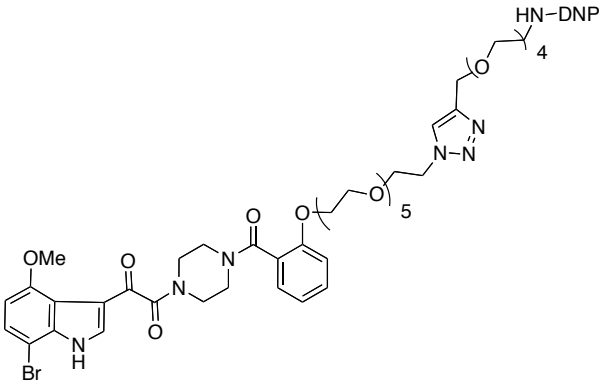
45.86, 41.93, 41.48, 41.33, 41.23. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{34}H_{45}BrN_6O_{10}^+$ *m/z* 775.2297 and 777.2282. Found 775.2434and 777.2401.



(**S21a**) To a solution of **S20a** (45 mg, 0.058 mmol) dissolved in a mixture of *t*BuOH/water (1.5 mL/1.4 mL), added **S9** (40 mg, 0.10 mmol, 1.7 equiv) within a microwave vial, added 0.1 M aq. $CuSO_4$ (30 uL) and 0.1M aq. sodium ascrobate (60 uL). The resulting mixture was heated to 125 ^oC fo 25 min when TLC (9:1 DCM/MeOH) indicated reaction completion. Volatiles were removed and crude **S21a** purified by flash column chromatography (100% CH_2Cl_2 → 20:1 CH_2Cl_2/CH_3OH → 10:1 CH_2Cl_2/CH_3OH), resulting in **S21a** as a sticky yellow solid (61 mg, 0.052 90%). **IR (ATR-FTIR)** 3360 (br), 3120 (w), 2872 (m), 1619 (s), 1523 (w), 1334 (m), 1304 (w), 1092 (s), 920 (w), 832 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.37 (s, 1H), 9.04 (d, J = 2.6, 1H), 8.73 (br peak, 1H), 8.18 (dd, J = 2.6, 9.5, 1H), 7.96 (d, J = 1.7, 2H), 7.69 (s, 1H), 7.28 (m, 1H), 7.24 (d, J = 8.5, 2H), 6.92 (m, 3H), 6.89 (d, J = 9.6, 1H), 6.52 (d, J = 8.5, 1H), 4.62 (s, 2H), 4.47 (t, J = 5.0, 2H), 4.09 (br peak, 4H), 3.87 (s, 3H), 3.85 – 3.71 (m, 14H), 3.71 – 3.29 (m, 28H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.98, 170.34, 167.19, 158.88, 153.44, 148.35, 144.70, 136.58, 136.30, 135.90, 134.82, 130.32, 130.12, 129.81, 126.75, 124.13, 123.84, 119.24, 116.35, 116.12, 115.66, 114.16, 113.23, 104.72, 96.72, 70.76, 70.63, 70.58, 70.54, 70.52, 70.50, 70.46, 70.45, 70.41, 69.58, 69.56, 69.35, 68.52, 67.56, 64.45, 56.08, 50.16, 43.19. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{51}H_{68}BrN_9O_{18}^+$ *m/z* 1172.3782 and 1174.3767 . Found 1172.4041 and 1174.4006.

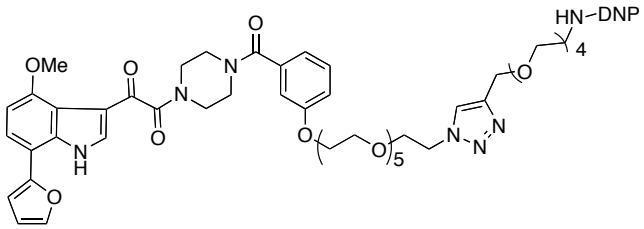


(**S21b**) Prepared in a similar manner as with **S21a** from **S20b** (15 mg, 0.019 mmol). Crude **S21b** purified by flash column chromatography (100% CH_2Cl_2 → 20:1 CH_2Cl_2/CH_3OH → 10:1 CH_2Cl_2/CH_3OH), resulting in **S21b** as a sticky yellow solid (19 mg, 0.016 mmol) 85% yield). **IR (ATR-FTIR)** 3508 (br), 3109 (w), 2869 (m), 1618 (s), 1512 (m), 1427 (m), 1246 (m), 1092 (s), 951 (w), 844 (w) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 10.11 (s, 1H), 9.07 (d, J = 2.7, 1H), 8.75 (s, 1H), 8.21 (dd, J = 2.6, 9.5, 1H), 7.99 (d, J = 3.1, 1H), 7.70 (s, 1H), 7.37 (d, J = 8.6, 2H), 7.27 (d, J = 10.0, 2H), 6.92 (d, J = 10, 2H), 6.91 (d, J = 10, 2H) 6.55 (d, J = 8.5, 1H), 4.64 (s, 2H), 4.49 (t, J = 5.1, 2H), 4.17 – 4.10 (m, 2H), 3.89 (s, 3H), 3.86 – 3.43 (m, 44H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.05, 170.62, 167.17, 160.31, 153.47, 148.37, 144.75, 136.54, 135.94, 134.55, 130.38, 130.14, 129.20, 127.15, 126.77, 124.15, 123.78, 116.29, 115.66, 114.51, 114.15, 104.75, 96.70, 70.80, 70.65, 70.60, 70.57, 70.55, 70.51, 70.48, 70.46, 70.42, 69.60, 69.52, 69.37, 68.55, 67.55, 64.46, 56.09, 50.17, 43.21. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{51}H_{68}BrN_9O_{18}^+$ *m/z* 1172.3782 and 1174.3767 . Found 1172.4026 and 1174.4011.

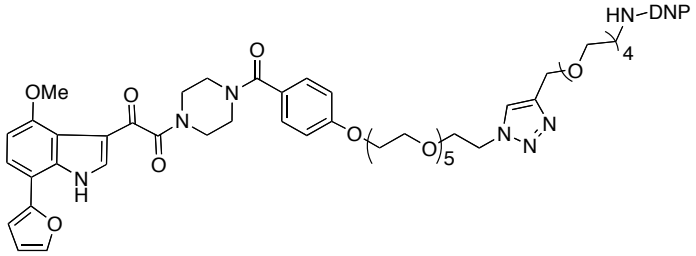


(S21c) Prepared in a similar manner as with **S21a** from **S20c** (20 mg, 0.026 mmol). Crude **S21c** purified by flash column chromatography (100% CH_2Cl_2 → 20:1 CH_2Cl_2/CH_3OH → 10:1 CH_2Cl_2/CH_3OH), resulting in **S21c** as a sticky yellow solid (28 mg, 0.24 mmol, 94%) yield**. IR (ATR-FTIR)** 3417 (br), 2984 (w) 2876 (m), 1649 (s), 1632 (s), 1502 (w), 1388 (m), 1254 (w), 1098 (m), 865 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.45 (two singlets, J = 85.3, 1H), 9.11 (d, J = 2.7, 1H), 8.84 – 8.70 (m, 1H), 8.24 (apparent dd, J = 2.5, 9.4, 1H), 8.03 (apparent d, J = 3.3, 1H), 7.66 (apparent d, J = 4.6, 1H), 7.41 – 7.27 (m, 4H), 7.07 – 6.79 (m, 4H), 6.59 (apparent dd, J = 8.5, 13.8, 1H), 4.64 (s, 2H), 4.46 – 4.36 (m, 2H),

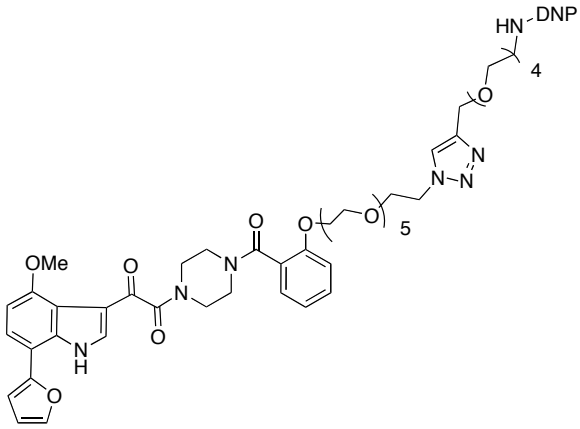
4.26 – 3.98 (m, 4H), 3.93 (apparent d, J = 16.0, 3H), 3.87 – 3.16 (m, 42H). (Note: For *ortho*-analogs, observed broadening/doubling of signals due to rotational isomerism). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 185.48, 168.02, 167.10, 154.37, 154.00, 148.50, 136.97, 136.29, 136.12, 135.96, 130.79, 130.39, 130.13, 128.29, 127.04, 125.47, 124.33, 121.43, 115.75, 114.06, 111.93, 111.38, 104.78, 96.52, 77.37, 77.06, 76.74, 70.72, 70.62, 70.52, 70.36, 70.23, 70.04, 69.62, 69.30, 68.59, 67.73, 67.57, 64.48, 56.18, 50.13, 47.11, 46.22, 43.27, 41.93, 41.32. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{51}H_{68}BrN_9O_{18}^+$ *m/z* 1172.3782 and 1174.3767 . Found 1172.4007 and 1174.3992.



(**12**) To a solution of **S21a** (20 mg, 0.017 mmol) in *t*BuOH/water (1.87 mL/1.12 mL) in a microwave vial, added 2-furanyl boronic acid (2.7 mg, 0.024 mmol, 1.4 equiv) and $NaHCO_3$ (2.0 mg, 0.024 mmol, 1.4 equiv). Bubbled nitrogen gas through (to remove oxygen) for 10 min. prior to adding $Pd(PPh_3)_4$ (0.98 mg, 0.85 μmol, 5 mol%) and heated subsequent mixture in microwave reactor for 12 min at 150 °C. Volatiles were removed and crude **11** was purified by HPLC (0-60% B gradient, 46 min run time), resulting in **12** as a sticky yellow solid (12.7 mg, 0.011 mmol, 65%). **IR (ATR-FTIR)** 3357 (w), 3115 (w), 2872 (m), 1620 (s), 1519 (m), 1425 (m), 1284 (m), 1104 (s), 799 (w) cm^{-1}. **^{1}H NMR (400 MHz, CDCl3)** δ 10.28 – 10.07 (m, 1H), 9.09 (d, J = 2.7, 1H), 8.84 – 8.66 (m, 1H), 8.21 (d, J = 9.5, 1H), 8.08 (d, J = 3.2, 1H), 7.72 (s, 1H), 7.55 (s, 1H), 7.45 (d, J = 8.3, 1H), 7.37 – 7.28 (m, 1H), 7.08 – 6.82 (m, 5H), 6.72 (d, J = 8.4, 1H), 6.67 (d, J = 3.4, 1H), 6.54 (dd, J = 1.8, 3.4, 1H), 4.64 (s, 2H), 4.50 (br signal, 2H), 4.12 (br signal, 2H), 3.95 (s, 3H), 3.89 – 3.38 (m, 44H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 170.51, 167.50, 158.89, 153.67, 152.94, 148.38, 141.23, 136.30, 135.97, 134.41, 134.01, 130.41, 130.15, 129.82, 124.18, 121.11, 119.29, 116.43, 115.43, 115.18, 114.11, 113.29, 111.81, 109.42, 104.51, 104.02, 70.65, 70.60, 70.46, 70.44, 70.37, 70.31, 70.23, 69.54, 69.31, 68.58, 67.57, 64.21, 55.97, 50.24, 43.11. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{55}H_{70}N_9O_{19}^+$ *m/z* 1160.4782. Found 1160.4890

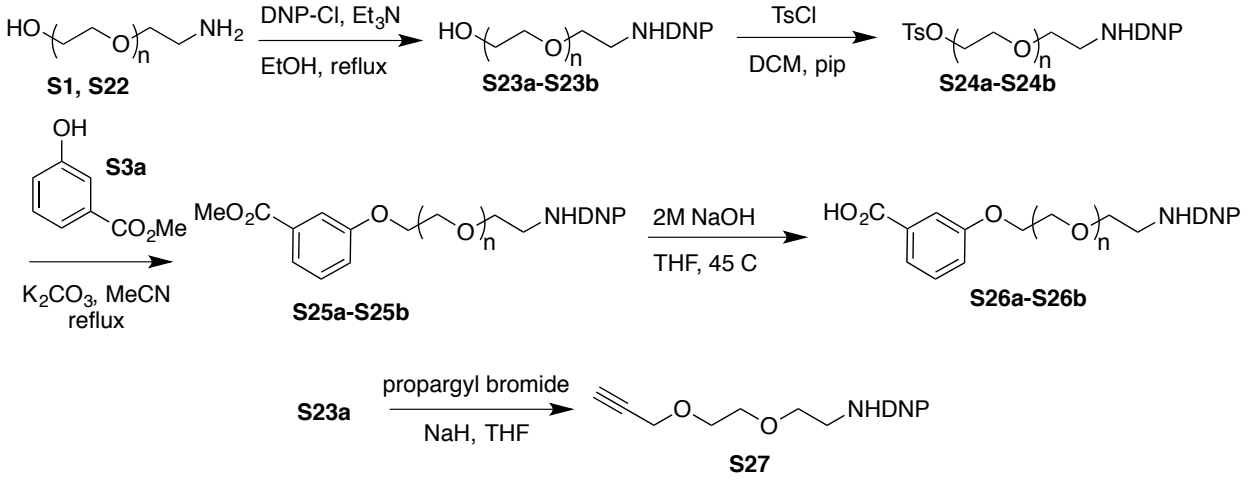


(**11**) Prepared **10** in the same manner as **12** using **S21b** (20 mg, 0.017 mmol). Crude **11** was purified by HPLC (0-60% B gradient, 46 min run time), resulting in **10** as a sticky yellow solid (9.9 mg, 50%). **IR (ATR-FTIR)** 3356 (w), 3112 (w), 2867 (m), 1622 (s), 1520 (m), 1427 (m), 1285 (m), 1103 (s), 1007 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.05 (s, 1H), 9.04 (d, J = 2.7, 1H), 8.71 (br peak, 1H), 8.17 (dd, J = 2.5, 9.6, 1H), 8.01 (d, J = 3.2, 1H), 7.65 (s, 1H), 7.48 (s, 1H), 7.38 (d, J = 8.3, 1H), 7.32 (m, 2H), 6.88 – 6.81 (m, 3H), 6.65 (d, J = 8.4, 1H), 6.60 (d, J = 3.4, 1H), 6.48 (dd, J = 1.8, 3.4, 1H), 4.59 (s, 2H), 4.44 (t, J = 5.1, 2H), 4.07 (m, 2H), 3.89 (s, 3H), 3.82 – 3.45 (m, 44H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 186.18, 170.49, 167.36, 160.28, 153.55, 153.01, 148.40, 144.67, 141.22, 136.06, 134.23, 134.01, 130.21, 129.34, 127.15, 124.26, 123.89, 121.02, 115.55, 115.26, 114.44, 114.06, 111.81, 109.39, 104.51, 103.99, 70.85, 70.72, 70.66, 70.59, 70.54, 69.63, 69.56, 69.40, 68.57, 67.55, 64.43, 56.06, 50.26, 43.22. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{55}H_{70}N_9O_{19}^+$ *m/z* 1160.4782. Found 1160.4719.

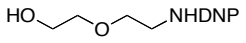


(**10**) Prepared **10** in the same manner as **12** using **S21c** (20 mg, 0.017 mmol). Crude **10** was purified by HPLC (0-60% B gradient, 60 min run time), resulting in **9** as a sticky yellow solid (10.5mg, 53%). **IR (ATR-FTIR)** 3352 (w), 3226 (w), 2869 (m), 1619 (s), 1512 (m), 1424 (m), 1283 (m), 1103 (s), 832 (w) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 10.48 (two singlets, J = 50.2, 1H), 9.10 (d, J = 2.6, 1H), 8.77 (s, 1H), 8.22 (dd, J = 2.5, 9.5, 1H), 8.08 (m, 1H), 7.73 (apparent d, J = 6.8, 1H), 7.54 (apparent dd, J = 1.2, 8.4, 1H), 7.46 (apparent dd, J = 8.3, 10.3, 1H), 7.41 – 7.28 (m, 2H), 7.02 (two triplets, J = 7.5, 25.4, 1H), 6.96 – 6.84 (m,

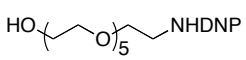
2H), 6.73 (apparent dd, J = 8.4, 12.8, 1H), 6.68 (apparent t, J = 3.8, 1H), 6.54 (apparent ddd, J = 1.8, 3.4, 5.3, 1H), 4.65 (s, 2H), 4.49 – 4.40 (m, 2H), 4.29 – 3.99 (m, 4H), 3.97 (d, J = 11.5, 3H), 3.88 – 3.20 (m, 42H) (Note: For *ortho*-analogs, observed broadening/doubling of signals due to rotational isomerism). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.30, 186.24, 168.41, 168.37, 167.67, 154.67, 154.35, 153.33, 148.82, 141.64, 136.42, 135.85, 131.21, 131.13, 130.86, 130.59, 124.64, 124.36, 121.92, 121.88, 121.73, 115.71, 115.65, 115.62, 114.54, 112.46, 112.26, 112.24, 112.22, 109.78, 105.00, 104.97, 104.43, 104.39, 71.13, 71.07, 70.99, 70.92, 70.91, 70.88, 70.84, 70.78, 70.66, 70.05, 69.96, 69.74, 68.99, 68.18, 64.82, 56.43, 56.40, 47.39, 46.90, 46.62, 46.31, 43.67, 42.28, 41.86, 41.64. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{55}H_{70}N_9O_{19}^+$ *m/z* 1160.4782. Found 1160.4688.



Scheme S6. General synthesis of benzoic acid substituted linkers.

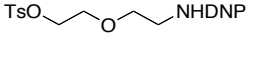


(**S23a**) 1-chloro-2,4-dinitrobenzene (3.857g, 19.05 mmol, 1 equiv) was added to a solution of 2-(2-aminoethoxy)ethanol (**S22,** 2.0 g, 19.05 mmol) in EtOH (75 mL), followed by TEA (3.86 g, 7.4 mL, 38.1 mmol, 2 equiv). Resulting yellow homogenous mixture was heated to reflux (90 ^oC) for 2 hrs at which time TLC (1:1 hexanes/EtOAc) indicated reaction completion. Solution allowed to cool and volatiles were removed via rotary evaporation. Resulting yellow solid was partitioned between DCM (50 mL) and water (100 mL), then extracted with DCM (3 X 50 mL). The resulting organic extracts were combined, dried over anhyd. Na_2SO_4, filtered and dried by rotary evaporation. Purified crude **S23a** by flash column chromatography (100% hexanes → 5:1 hexanes/EtOAc → 1:1 hexanes EtOAc → 100% EtOAc) and combined like fractions, yielding pure **S23a** as a yellow solid (4.98 g, 18.3 mmol, 96%). **IR (ATR-FTIR)** 3357 (w), 3110 (br), 2873 (m), 1617 (s), 1585 (m), 1423 (w), 1331 (m), 1121 (m), 1057(m), 919 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.11 (d, J = 2.7, 1H), 8.82 (s, 1H), 8.26 (ddd, J = 0.6, 2.7, 9.5, 1H), 6.95 (d, J = 9.5, 1H), 3.90 – 3.55 (m, 8H), 2.01 (br s, 1H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 148.78, 136.56, 130.96, 130.73, 124.70, 114.38, 72.98, 68.90, 62.28, 43.57. **HRMS (ESI+)** calcd for $C_{10}H_{14}N_3O_6^+$ $[M+H]^+$ *m/z* 272.0877. Found 272.0869.

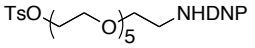


(**S23b**) Prepared **S23b** in the same manner as **S23a** starting from 17-amino-3,6,9,12,15-pentaoxaheptadecan-1-ol (**S1**, 200 mg, 0.71 mmol). Crude **S23b** was purified by flash column chromatography (5:1 hexanes/EtOAc → 1:1 hexanes EtOAc → 100% EtOAc) and combined like fractions, yielding pure **S23b** as a yellow oil (210 mg, 0.47 mmol, 66%). **IR (ATR-FTIR)** 3360 (w), 3109 (br), 2870 (m), 1618 (s), 1586 (m), 1424 (w), 1332 (m), 1090 (s), 918 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.11 (d, J = 2.7, 1H), 8.78 (s, 1H), 8.24 (dd, J = 2.6, 9.5, 1H), 6.95 (d, J = 9.6, 1H), 3.82 (t, J = 5.3, 2H), 3.74 – 3.53 (m, 22H), 2.57 (br s, 1H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 148.85, 136.46, 130.92, 130.63, 124.67, 114.52, 72.87,

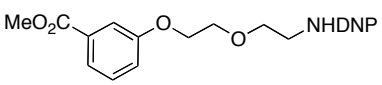
71.17, 71.07, 71.03, 71.00, 70.97, 70.92, 70.74, 69.00, 62.13, 43.68. **HRMS (ESI+)** calcd for $C_{18}H_{30}N_3O_{10}$+ $[M+H]^+$ *m/z* 448.1926 Found 448.1970.



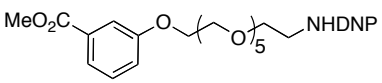
(**S24a**) To a solution of **S23a** (1.0 g, 3.68 mmol) in DCM (150 mL), added tosyl chloride (2.0 g, 11.03 mmol, 3 equiv), TEA (4.15 mL) and DMAP (200 mg, 1.63 mmol, 0.45 equiv). Resulting solution was stirred for 20 hrs when TLC (1:1 hexanes/EtOAc) indicated reaction completion, subsequently, sat. NH_4Cl (200 mL) was added and the organic layer was extracted with DCM (3 X 150 mL). Combined organic layers were dried over anhyd. Na_2SO_4, filtered and dried by rotary evaporation. Purified crude **S24a** by flash column chromatography (5:1 hexanes/EtOAc → 1:1 hexanes EtOAc) and combined like fractions, yielding pure **S24a** as a yellow solid (1.4 g, 3.3 mmol, 92%). **IR (ATR-FTIR)** 3360 (w), 3111 (w), 2926 (w), 1619 (s), 1588 (m), 1425 (m), 1334 (m), 1175 (m), 1095(w), 919 (m) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.09 (d, J = 2.4, 1H), 8.71 (s, 1H), 8.23 (dd, J = 2.5, 9.5, 1H), 7.77 (d, J = 8.2, 2H), 7.32 (d, J = 8.1, 2H), 6.95 (d, J = 9.5, 1H), 4.27 – 4.06 (m, 2H), 3.76 (dt, J = 4.8, 9.1, 4H), 3.57 (dd, J = 5.2, 10.4, 2H), 2.42 (s, 3H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 148.83, 145.51, 136.55, 133.18, 130.91, 130.66, 130.31, 128.31, 124.58, 114.60, 69.42, 69.36, 69.28, 43.50, 22.04. **HRMS (ESI+)** calcd for $C_{17}H_{20}N_3O_8S^+$ $[M+H]^+$ *m/z* 426.0966. Found 426.0936.



(**S24b**) Prepared **S24b** in the same manner as compound **S24a** from **S23b** (200 mg, 0.447 mmol). Crude **S24b** was purified by flash column chromatography (5:1 hexanes/EtOAc → 1:1 hexanes EtOAc → 100% EtOAc) and combined like fractions, yielding pure **S24b** as a yellow oil (180 mg, 0.30 mmol, 67%). **IR (ATR-FTIR)** 3360 (w), 2871 (w), 1619 (s), 1587 (m), 1425 (w), 1333 (m), 1175 (m), 1094 (m), 917 (m) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.21 – 9.13 (m, 1H), 8.83 (s, 1H), 8.29 (dd, J = 2.7, 9.5, 1H), 7.81 (d, J = 8.1, 2H), 7.37 (d, J = 8.5, 2H), 6.98 (d, J = 9.5, 1H), 4.17 (dt, J = 6.8, 13.7, 2H), 3.86 (t, J = 5.1, 2H), 3.78 – 3.55 (m, 20H), 2.47 (s, 3H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 148.47, 144.85, 136.06, 133.00, 130.52, 130.24, 129.85, 127.96, 124.26, 114.18, 70.78, 70.76, 70.71, 70.64, 70.63, 70.61, 70.57, 70.52, 69.29, 68.69, 68.61, 43.29, 21.64. **HRMS (ESI+)** calcd for $C_{25}H_{36}N_{3O12}S^+$ $[M+H]^+$ *m/z* 602.2014. Found 602.2043.

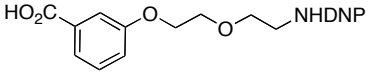


(**S25a**) To a solution of **S24a** (100 mg, 0.24 mmol) in MeCN (3 mL), added K_2CO_3 (38.3 mg, 0.28 mmol, 1.16 equiv) and 3-hydroxybenzoate (**S3a**, 43 mg, 0.28 mmol, 1.16 equiv) and heated to reflux (75°C) for 20 hr when TLC indicated reaction completion (1:1 hexanes/EtOAc). Reaction was quenched with sat. NH_4Cl (20 mL) and then extracted with DCM (3 X 25 mL). The combined organic layers were dried over anhyd. Na_2SO_4, filtered and dried by rotary evaporation. Purified crude **S25a** by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 12g pre-packed dry loading column. Run using 100% hexanes to 100% EtOAc in hexanes gradient over 40 column volumes) and combined like fractions, yielding pure **S25a** as a yellow solid (90 mg, 0.22 mmol, 93%). **IR (ATR-FTIR)** 3362 (w), 3110 (w), 2955 (w), 1720 (m), 1621 (s), 1588 (m), 1445 (w), 1336 (s), 1289 (m), 1100 (w), 922 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.11 (d, J = 2.7, 1H), 8.92 – 8.74 (br s, 1H), 8.24 (dd, J = 2.7, 9.5, 1H), 7.64 (d, J = 7.7, 1H), 7.54 (dd, J = 1.5, 2.5, 1H), 7.33 (t, J = 8.0, 1H), 7.10 (ddd, J = 0.7, 2.6, 8.2, 1H), 6.95 (d, J = 9.5, 1H), 4.24 – 4.17 (m, 2H), 4.00 – 3.77 (m, 7H), 3.69 – 3.50 (m, 2H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 167.23, 158.92, 148.81, 136.58, 131.89, 131.00, 130.63, 129.91, 124.69, 122.80, 120.34, 115.11, 114.39, 70.28, 69.39, 68.02, 52.62, 43.69. **HRMS (ESI+)** calcd for $C_{18}H_{20}N_3O_8^+$ $[M+H]^+$ *m/z* 406.1245. Found 406.1244.

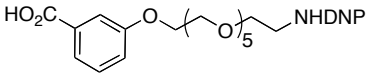


(**S25b**) Prepared **S25b** in the same manner as compound **S25a** from **S24b** (50 mg, 0.085 mmol). Crude **S25b** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 12g pre-packed dry loading column. Run using 100% DCM to 20% MeOH in DCM gradient over 30 column volumes) and combined like fractions, yielding pure **S25a** as a yellow solid (30 mg, 0.07 mmol, 84%). **IR (ATR-FTIR)** 3359 (w), 2871 (w), 1720 (m), 1621 (s), 1587 (m), 1445 (w), 1336 (s), 1289 (m), 1101 (m) cm^{-1}. **^{1}H NMR (400 MHz,**

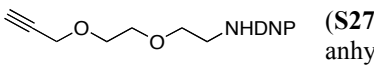
$CDCl_3$) δ 9.11 (dd, J = 1.5, 2.6, 1H), 8.78 (s, 1H), 8.24 (dd, J = 2.6, 9.5, 1H), 7.60 (d, J = 7.7, 1H), 7.53 (s, 1H), 7.31 (t, J = 7.9, 1H), 7.10 (dd, J = 2.6, 8.2, 1H), 6.93 (d, J = 9.5, 1H), 4.19 – 4.11 (m, 2H), 3.89 (s, 3H), 3.87 – 3.83 (m, 2H), 3.81 (t, J = 5.2, 2H), 3.75 – 3.54 (m, 18H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 167.29, 159.15, 148.82, 136.46, 131.80, 130.91, 130.62, 129.79, 124.67, 122.55, 120.43, 115.16, 114.48, 71.27, 71.18, 71.10, 71.05, 71.03, 71.00, 70.98, 70.04, 69.00, 68.06, 52.57, 43.68. **HRMS (ESI+)** calcd for $C_{26}H_{36}N_3O_{12}^+$ $[M+H]^+$ *m/z* 582.2294, found 582.2359.



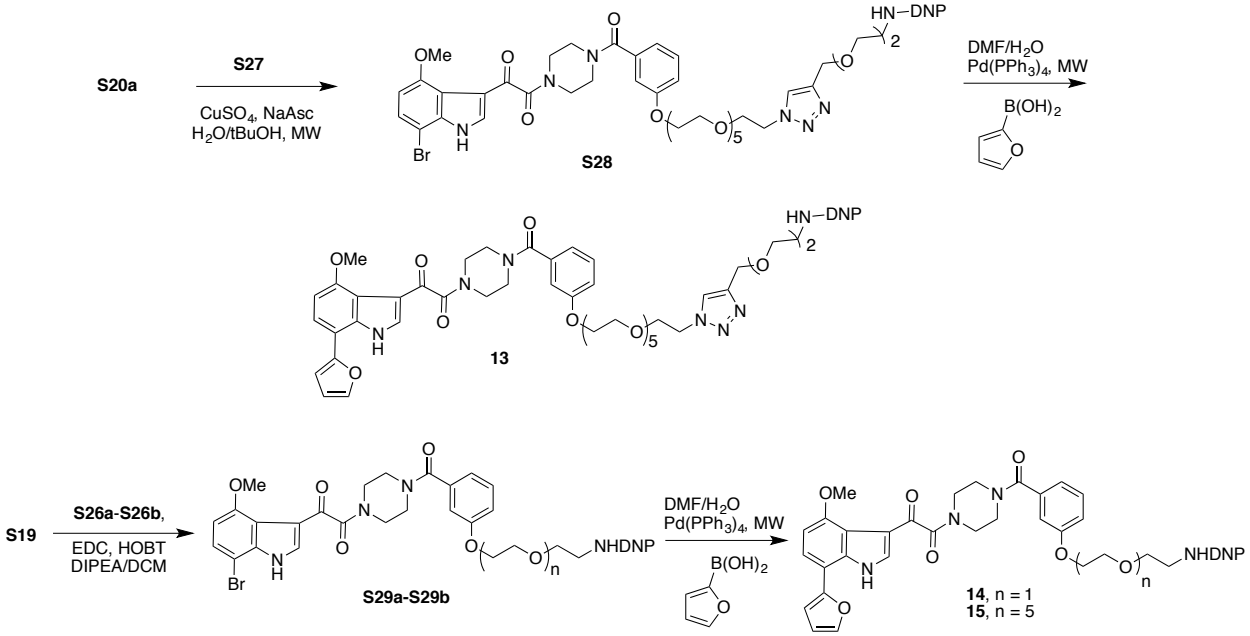
(**S26a**) Prepared **S26a** in the same manner as compound **S5a** from **S25a** (100 mg, 0.247 mmol). Crude **S26a** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 12g pre-packed dry loading column. Run using 100% DCM to 20% MeOH in DCM gradient over 30 column volumes) and combined like fractions, yielding pure **S26a** as a yellow solid (65 mg, 0.16 mmol, 67%). **IR (ATR-FTIR)** 3356 (w), 3111 (w), 2873 (m), 1717 (m), 1621 (s), 1587 (m), 1447 (w), 1336 (s), 1305 (m), 1103 (m), 920 (w) cm^{-1}. **^{1}H NMR (500 MHz, DMSO)** δ 13.14 – 12.63 (br s, 1H), 8.89 (br s, 1H), 8.81 (d, J = 2.7, 1H), 8.22 (dd, J = 2.8, 9.6, 1H), 7.50 (d, J = 7.9, 1H), 7.40 – 7.34 (m, 2H), 7.29 (d, J = 9.7, 1H), 7.14 (dd, J = 2.2, 7.8, 1H), 4.20 – 4.13 (m, 2H), 3.86 – 3.81 (m, 2H), 3.79 (t, J = 5.4, 2H), 3.73 – 3.66 (m, 2H). **^{13}C NMR (125 MHz, DMSO)** δ 167.38, 158.74, 148.77, 135.23, 132.56, 130.15, 130.00, 129.97, 123.88, 121.97, 119.55, 116.03, 115.99, 114.86, 79.60, 79.34, 79.07, 69.18, 68.82, 67.69, 43.03, 42.92. **HRMS (ESI+)** calcd for $C_{17}H_{18}N_3O_8^+$ $[M+H]^+$ *m/z* 392.1088. Found 392.1064.



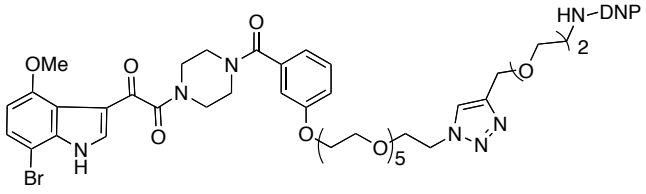
(**S26b**) Prepared **S26b** in the same manner as compound **S5a** from **S25a** (100 mg, 0.172 mmol). Crude **S26b** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 12g pre-packed dry loading column. Run using 100% DCM to 20% MeOH in DCM gradient over 30 column volumes) and combined like fractions, yielding pure **S26a** as a yellow oil (50 mg, 0.088 mmol, 52%) **IR (ATR-FTIR)** 3357 (w), 2919 (m), 1717 (m), 1621 (s), 1587 (m), 1447 (w), 1336 (s), 1305 (m), 1104 (m), 921 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.11 (d, J = 2.5, 1H), 8.78 (s, 1H), 8.23 (dd, J = 2.4, 9.5, 1H), 7.66 (d, J = 7.5, 1H), 7.60 (s, 1H), 7.34 (t, J = 7.9, 1H), 7.14 (dd, J = 1.9, 8.2, 1H), 6.92 (d, J = 9.5, 1H), 4.18 (br s, 2H), 3.87 (br s, 2H), 3.81 (d, J = 5.1, 2H), 3.76 – 3.52 (m, 18H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 159.23, 148.81, 136.47, 130.91, 130.63, 129.89, 124.69, 123.15, 121.23, 115.72, 114.47, 71.28, 71.14, 71.09, 71.03, 71.00, 70.96, 70.94, 70.13, 68.99, 68.15, 43.64. **HRMS (ESI+)** calcd for $C_{25}H_{34}N_3O_{12}^{tdl+}$ $[M+H]^+$ *m/z* 568.2137. Found 568.2162.



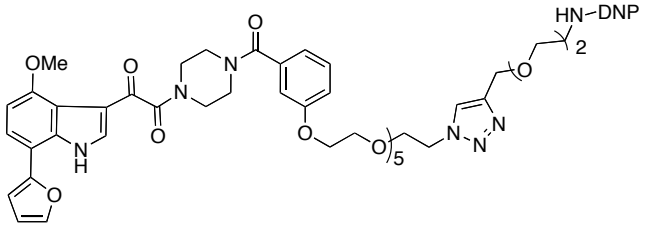
(**S27**) To a flame-dried flask containing **S23a** (500 mg, 1.84 mmol) dissolved in anhyd. THF (10 mL) chilled with an ice-bath, carefully added NaH (95%, 68 mg, 2.8 mmol, 1.5 equiv) followed by propargyl bromide (80% in toluene, 305 µL, 1.1 equiv). Resulting mixture was allowed to warm to room temperature and stirred for an additional 18 hr, when TLC (20:1 DCM/MeOH) indicated reaction completion. Reaction was quenched carefully with sat. NH_4Cl (20 mL) and extracted with DCM (3 X 20 mL). Organic layers were combined, dried over Na_2SO_4, filtered and dried via rotary evaporation. Crude **S19** was purified by flash column chromatography (5:1 hexanes/EtOAc → 1:1 hexanes/EtOAc) and like fractions were combined, yielding pure **S27** as a yellow solid (443 mg, 1.43 mmol, 78%) **IR (ATR-FTIR)** 3359 (w), 3288 (w), 2872 (m), 1617 (s), 1585 (m), 1423 (m), 1331 (m), 1086 (m), 918 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.13 (d, J = 2.7, 1H), 8.79 (s, 1H), 8.26 (ddd, J = 0.6, 2.7, 9.5, 1H), 6.96 (d, J = 9.5, 1H), 4.20 (d, J = 2.4, 2H), 3.88 – 3.78 (m, 2H), 3.73 (d, J = 3.6, 4H), 3.61 (dd, J = 5.2, 10.5, 2H), 2.43 (t, J = 2.4, 1H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 148.86, 136.52, 130.96, 130.65, 124.69, 114.51, 79.82, 75.13, 70.95, 69.55, 69.12, 58.91, 43.67. **HRMS (ESI+)** calcd for $C_{13}H_{16}N_3O_6^+$ $[M+H]^+$ *m/z* 310.1034. Found 310.1031.



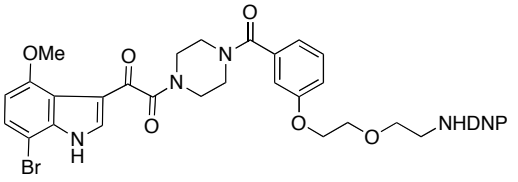
Scheme S7. General synthesis of *meta*-linker substituted ARM-Hs of varying linker lengths.



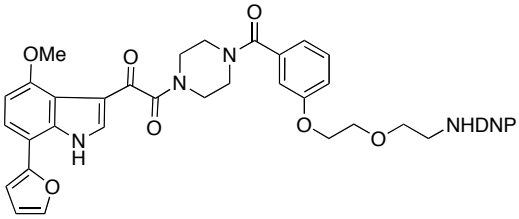
(**S28**) Prepared in a similar manner as with **S21a** from **S20a** (11 mg, 0.014 mmol) and **S27** (7.5 mg, 0.024 mmol, 1.7 equiv). Crude **S28** was purified by flash column chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 4 g pre-packed dry loading column. Run using 100% DCM to 10% MeOH in DCM gradient over 65 column volumes) yielding pure **S28** as a yellow solid (12 mg, 0.011 mmol, 94%). **IR (ATR-FTIR)** 3358 (w), 3115 (w), 2869 (m), 1620 (s), 1587 (m), 1433 (w), 1334 (m), 1285 (m), 1092 (m), 969 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$) δ** 10.08 – 9.76 (m, 1H), 9.07 (d, J = 2.6, 1H), 8.84 – 8.71 (m, 1H), 8.20 (dd, J = 2.6, 9.5, 1H), 8.00 (s, 1H), 7.90 – 7.65 (m, 1H), 7.30 (s, 2H), 7.04 – 6.82 (m, 4H), 6.56 (d, J = 8.5, 1H), 4.68 (br s, 2H), 4.53 (br s, 2H), 4.11 (br s, 2H), 3.91 (s, 3H), 3.90-3.35 (m, 36H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 169.76, 166.53, 158.30, 152.97, 147.74, 135.97, 135.76, 135.41, 133.88, 129.79, 129.61, 129.25, 126.29, 123.58, 118.70, 115.85, 115.78, 115.05, 113.49, 112.66, 104.26, 96.06, 70.21, 69.98, 69.94, 69.89, 69.85, 69.05, 69.01, 68.74, 67.89, 67.01, 55.53, 42.59. **HRMS (ESI+)** calcd for $C_{47}H_{59}BrN_9O_{16}^+$ $[M+H]^+$ *m/z* 1084.3263 and 1086.3243. Found 1084.3269 and 1086.3331.



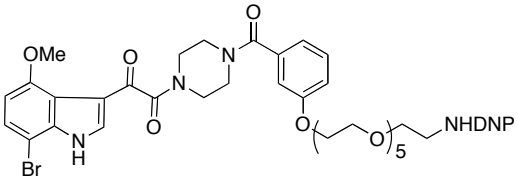
(**16**) Prepared **16** in the same manner as **12** using **S28** (7 mg, 0.007 mmol). Crude **13** was purified by HPLC (30-60% B gradient, 60 min run time), resulting in **16** as a sticky yellow solid (10.5mg, 53%). **IR (ATR-FTIR)** 3357 (w), 3110 (w), 2869 (m), 1621 (s), 1586 (m), 1437 (w), 1335 (m), 1284 (m), 1106 (m), 967 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.11 (br s, 1H), 9.09 (d, J = 2.5, 1H), 8.78 (br s, 1H), 8.22 (dd, J = 2.5, 9.4, 1H), 8.09 (s, 1H), 7.84 (s, 1H), 7.56 (s, 1H), 7.46 (d, J = 8.2, 1H), 7.32 (br s, 1H), 7.05 – 6.85 (m, 4H), 6.73 (d, J = 8.4, 1H), 6.68 (d, J = 3.4, 1H), 6.55 (dd, J = 1.6, 3.3, 1H), 4.71 (br s, 2H), 4.55 (br s, 2H), 4.13 (br s, 2H), 3.96 (s, 3H), 3.92 – 3.29 (m, 36H). **HRMS (ESI+)** calcd for $C_{51}H_{62}N_9O_{17}^+$ $[M+H]^+$ *m/z* 1072.4258. Found 1072.4297.



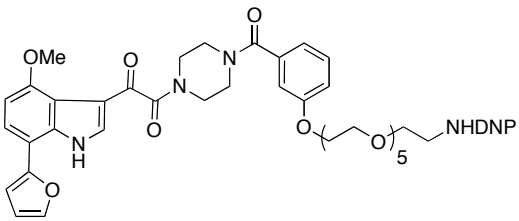
(**S29a**) Prepared in a similar manner as **S20a** from **S19** (20 mg, 0.055 mmol) and **S26a** (24 mg, 0.06 mmol, 1.1 equiv). Crude **S29a** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 4 g pre-packed dry loading column. Run using 100% DCM to 20% MeOH in DCM gradient over 65 column volumes) and combined like fractions, yielding pure **S29a** as a yellow oil (30 mg, 0.04 mmol, 74%). **IR (ATR-FTIR)** 2928 (w), 2857 (w), 1659 (s), 1503 (m), 1438 (w), 1384 (m), 1255 (w), 1089 (m), 865 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.49 (s, 1H), 9.06 (s, 1H), 8.90 – 8.71 (m, 1H), 8.21 (dd, J = 2.5, 9.5, 1H), 7.95 (d, J = 3.1, 1H), 7.30 (m, 2H), 7.11 – 6.79 (m, 4H), 6.57 (d, J = 8.5, 1H), 4.17 (br s, 2H), 4.01 – 3.30 (m, 17H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 186.02, 170.34, 167.12, 158.79, 153.57, 148.38, 136.50, 136.46, 136.14, 134.07, 130.52, 130.24, 129.89, 126.94, 124.23, 119.54, 116.64, 116.34, 115.63, 114.06, 113.52, 104.95, 96.67, 69.78, 68.83, 67.55, 56.16, 43.23. **HRMS (ESI+)** calcd for $C_{32}H_{32}BrN_6O_{10}^+$ $[M+H]^+$ *m/z* 739.1363 and 741.1343. Found 739.1431 and 741.1406.



(**18**) Prepared **18** in the same manner as **12** using **S29a** (12 mg, 0.003 mmol). Crude **14** was purified by was purified by HPLC (30-60% B gradient, 46 min run time) resulting in **18** as a sticky yellow solid (55%). **IR (ATR-FTIR)** 3359 (w), 3111 (w), 2929 (w), 1620 (s), 1587 (m), 1425 (w), 1335 (m), 1283 (m), 1128 (m), 968 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.05 (br s, 1H), 9.09 (s, 1H), 8.82 (br s, 1H), 8.24 (apparent d, J = 9.3, 1H), 8.11 (d, J = 3.1, 1H), 7.57 (s, 1H), 7.47 (d, J = 8.2, 1H), 7.33 (br s, 1H), 6.96 (dd, J = 8.9, 19.0, 4H), 6.74 (d, J = 8.3, 1H), 6.68 (d, J = 3.2, 1H), 6.56 (dd, J = 1.8, 3.4, 1H), 4.17 (br s, 2H), 4.05 – 3.28 (m, 17H). **HRMS (ESI+)** calcd for $C_{36}H_{35}N_6O_{11}^+$ $[M+H]^+$ *m/z* 727.2358. Found 727.2409.

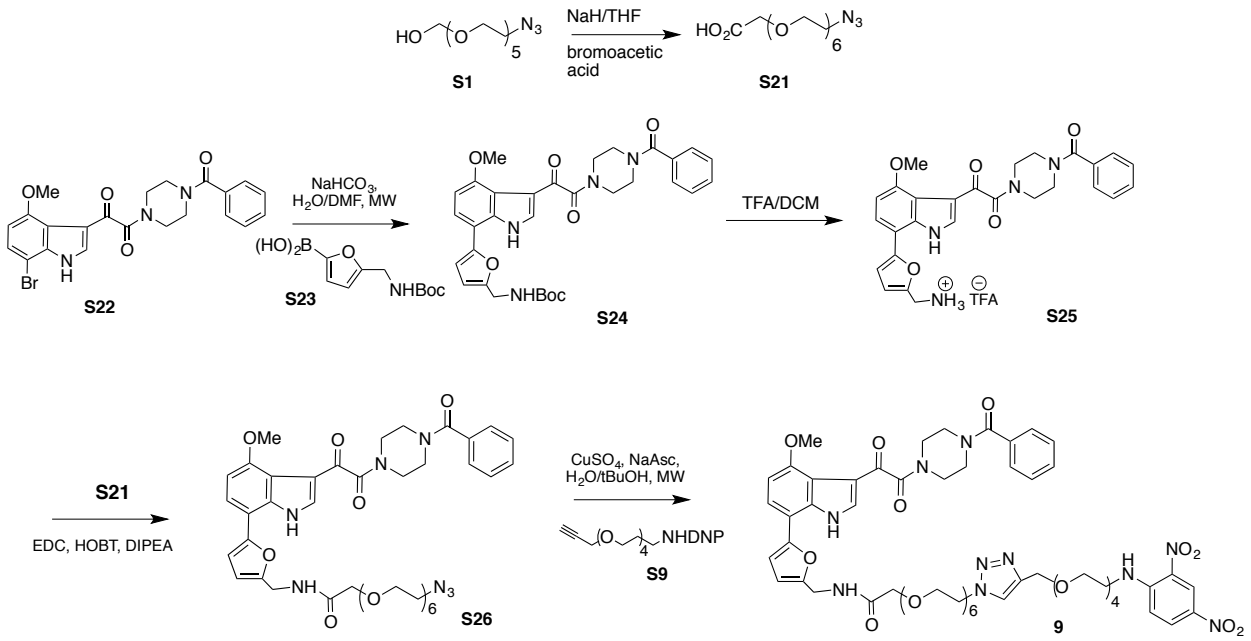


(**S29b**) Prepared in a similar manner as **S20a** from **S19** (20 mg, 0.055 mmol) and **S26b** (34 mg, 0.06 mmol, 1.1 equiv). Crude **S29b** was purified by flash chromatography (CombiFlash Automated Chromatographer, 12g column, dryloaded with 4 g pre-packed dry loading column. Run using 100% DCM to 20% MeOH in DCM gradient over 65 column volumes) and combined like fractions, yielding pure **S29b** as a yellow oil (30 mg, 0.04 mmol, 74%). **IR (ATR-FTIR)** 3356 (w), 3111 (w), 2870 (w), 1621 (s), 1586 (m), 1435 (w), 1335 (m), 1285 (m), 1093 (m), 970 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 9.64 (s, 1H), 9.08 (d, J = 2.6, 1H), 8.75 (s, 1H), 8.21 (d, J = 9.5, 1H), 7.98 (d, J = 3.1, 1H), 7.29 (d, J = 8.4, 2H), 7.00 – 6.87 (m, 4H), 6.56 (d, J = 8.5, 1H), 4.11 (s, 2H), 3.90 (s, 3H), 3.86 – 3.42 (m, 26H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 184.96, 169.38, 166.12, 157.91, 152.53, 147.36, 135.47, 135.45, 135.30, 135.00, 134.98, 133.20, 129.41, 129.18, 128.84, 125.90, 123.20, 118.29, 115.50, 115.48, 115.42, 115.38, 114.59, 113.08, 112.29, 103.87, 95.63, 69.78, 69.71, 69.64, 69.60, 69.54, 69.51, 68.61, 67.55, 66.62, 55.12, 42.21. **HRMS (ESI+)** calcd for $C_{40}H_{48}BrN_6O_{14}^+$ $[M+H]^+$ *m/z* 915.2412 and 917.2391. Found 915.2469 and 917.2467.

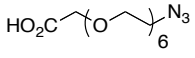


(**17**) Prepared **17** in the same manner as **12** using **S29a** (30 mg, 0.032 mmol). Crude **17** was purified by HPLC (30-60% B gradient, 46 min run time), resulting in **17** as a sticky yellow solid (16.7 mg, 0.018 mmol, 58%). **IR (ATR-FTIR)** 3358 (w), 3112 (w), 2914 (m), 1621 (s), 1588 (m), 1426 (m), 1335 (m), 1284 (m), 1093 (m), 953 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 10.06 (s, 1H), 9.10 (d, J = 2.5, 1H), 8.75 (s, 1H), 8.22 (dd, J =

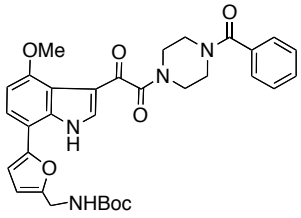
2.6, 9.5, 1H), 8.08 (d, J = 3.0, 1H), 7.56 (d, J = 1.4, 1H), 7.46 (d, J = 8.3, 1H), 7.31 (s, 1H), 7.05 – 6.84 (m, 4H), 6.73 (d, J = 8.4, 1H), 6.67 (d, J = 3.4, 1H), 6.55 (dd, J = 1.8, 3.3, 1H), 4.12 (s, 2H), 3.96 (s, 3H), 3.87 – 3.20 (m, 30H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.08, 170.47, 167.34, 158.83, 153.66, 153.02, 148.31, 141.21, 141.18, 136.19, 136.04, 134.21, 134.01, 130.40, 130.22, 130.17, 130.15, 129.81, 124.19, 121.11, 119.35, 116.52, 115.50, 115.15, 114.05, 113.37, 111.84, 109.38, 104.50, 104.07, 70.48, 70.40, 70.24, 70.21, 70.05, 69.99, 69.49, 68.44, 67.57, 55.98, 43.01**. HRMS (ESI+)** calc'd for $C_{44}H_{50}N_6O_{15}$ $[M+H]^+$ *m/z* 903.3413, found 903.3472



Scheme S8. Synthesis of BMS-furan derived ARM-H (**9**) with linker attached on furan

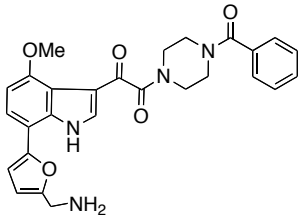


(**S21**) Azido polyethylene glycol **S1** (0.60 g, 2 mmol, 1 equiv.) was dissolved in dry THF (10 mL) and cooled at 0°C, then sodium hydride (0.15 g, 6.3 mmol, 3.1 equiv.) was added in portions followed by bromoacetic acid (0.35 g, 2.5 mmol, 1.25 equiv.). The suspension was stirred at room temperature under nitrogen overnight. Water (1 mL) was added carefully and then stirred for 5 min. The reaction mixture was concentrated *in vacuo*. DCM was added and organic layer was washed with 2N HCl and brine. The organic layer was dried over Na_2SO_4 and all solvents were evaporated. Pure **S21** (0.72 g, 98%) was obtained as an oil. **IR (ATR-FTIR)** 2892 (br), 2870 (m), 2101 (s), 1739 (m), 1285 (m), 1092 (s), 943 (m), 850 (m) cm^{-1}. **^{1}H NMR (500 MHz, $CDCl_3$)** δ 4.06 (s, 2H), 3.68 – 3.62 (m, 2H), 3.62 – 3.49 (m, 20H), 3.30(t, J = 5.0, 1H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 172.37, 70.84, 70.50, 70.48, 70.43, 70.42, 70.38, 70.31, 70.23, 69.86, 68.57, 50.54. **HRMS (ESI+)** calc'd for $C_{14}H_{29}N_3O_8^+$ $[M+H]^+$ *m/z* 366.1871. Found 366.1988; for $[M + Na]^+$, calc'd 388.1690, found 388.1795

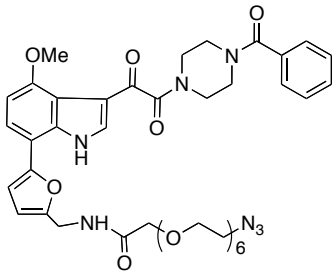


(**S24**) To a microwave vial (2.0 -5.0 mL) containing **S22** (52 mg, 0.11 mmol) in DMF/water (3.0/1.8mL), added 5-((BOC-Amino)methyl)furan-2-boronic acid (**S23**, Combi-Blocks LLC, San Diego CA; 37 mg, 0.154 mmol, 1.4 equiv) and $NaHCO_3$ (12.8 mg, 0.154 mmol, 1.4 equiv). Oxygen was removed from the solvent by bubbling nitrogen gas in solution for 10 min, to which $Pd(PPh_3)_4$ (6.3 mg, 0.0055 mmol, 5 mol %) was added. The subsequent heterogenous solution was capped and heated in microwave reactor for 12 min at 150 °C when UPLC/MS analysis showed reaction completion. The volatile solvents were removed by rotary evaporation and

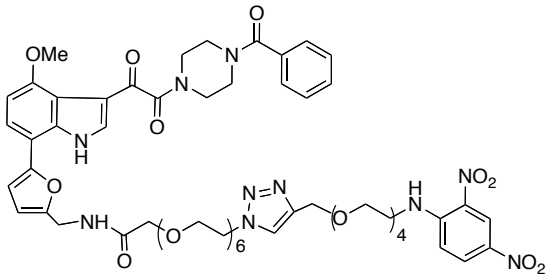
crude material was purified by flash chromatography (flash chromatography (CombiFlash Automated Chromatographer, 12g colum; gradient elution ranging from 0% MeOH:DCM to 15% MeOH:DCM was performed over 30 column volumes) to yield pure **S22** as a yellow solid (52 mg, 0.088 mmol, 80%). **IR (ATR-FTIR)** 3350 (br), 2977 (w) 2912 (w), 1703 (m), 1632 (s), 1512 (m), 1283 (m), 1168 (m), 1008 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 11.26 (s, 1H), 8.06 (s, 1H), 7.45 – 7.32 (m, 6H), 6.65 (d, J = 8.4, 1H), 6.49 (d, J = 3.3, 1H), 6.21 (d, J = 3.3, 1H), 5.43 (s, 1H), 4.26 (d, J = 6.3, 2H), 3.90 (s, 3H), 3.86 – 3.24 (m, 8H), 1.42 (s, 9H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 186.81, 171.10, 168.10, 157.62, 153.92, 152.77, 151.30, 135.66, 135.51, 134.65, 132.51, 132.43, 132.42, 132.40, 130.54, 129.06, 128.99, 128.89, 127.48, 120.46, 115.89, 115.32, 110.02, 109.03, 104.26, 80.41, 56.35, 37.81, 28.89. **HRMS (ESI+)** calc'd for $C_{32}H_{35}N_4O_7$ $[M+H]^+$ *m/z* 587.2500. Found 587.2418



(**S25**) To **S24** (51 mg, 0.86 mmol) in DCM (800 μL), added trifluoroacetic acid (250 μL), resulting in a color change from yellow to dark brown. Solution was stirred at room temperature, open to air for 1 hr when TLC (20:1 DCM/MeOH) showed reaction completion. Volatiles were removed by rotary evaporation, co-evaporating several times with chloroform, resulting in **S23** as a yellow solid which was used purified by running through a short silica plug (eluted with 10% MeOH in DCM). **IR (ATR-FTIR)** 3507 (br), 3024 (w) 2870 (w), 1636 (s), 1512 (m), 1430 (m), 1284 (m), 1006 (w) cm^{-1}. **^{1}H NMR (400 MHz, 10% MeOH-d_4 in $CDCl_3$)** δ 8.04 (s, 1H), 7.34 (m, 6H), 6.64 (d, J = 8.4, 1H), 6.50 (d, J = 3.3, 1H), 6.30 (d, J = 3.3, 1H), 3.93 (s, 2H), 3.86 (s, 3H), 3.80-3.31 (m, 8H). **^{13}C NMR (125 MHz, 10% MeOH-d_4 in $CDCl_3$)** δ 186.55, 171.48, 168.33, 162.98, 154.19, 154.00, 150.89, 136.85, 135.27, 130.72, 129.15, 127.45, 121.83, 115.89, 114.87, 111.00, 110.03, 105.43, 104.37, 56.37, 56.33, 53.89, 38.17.. **HRMS (ESI+)** calc'd for $C_{27}H_{27}N_4O_5^+$ $[M+H]^+$ *m/z* 487.1976 Found 487.2079; for $C_{27}H_{24}N_3O_5^{\bullet}$ [M• + •NH_2], calc'd 470.1716, found 470.1765

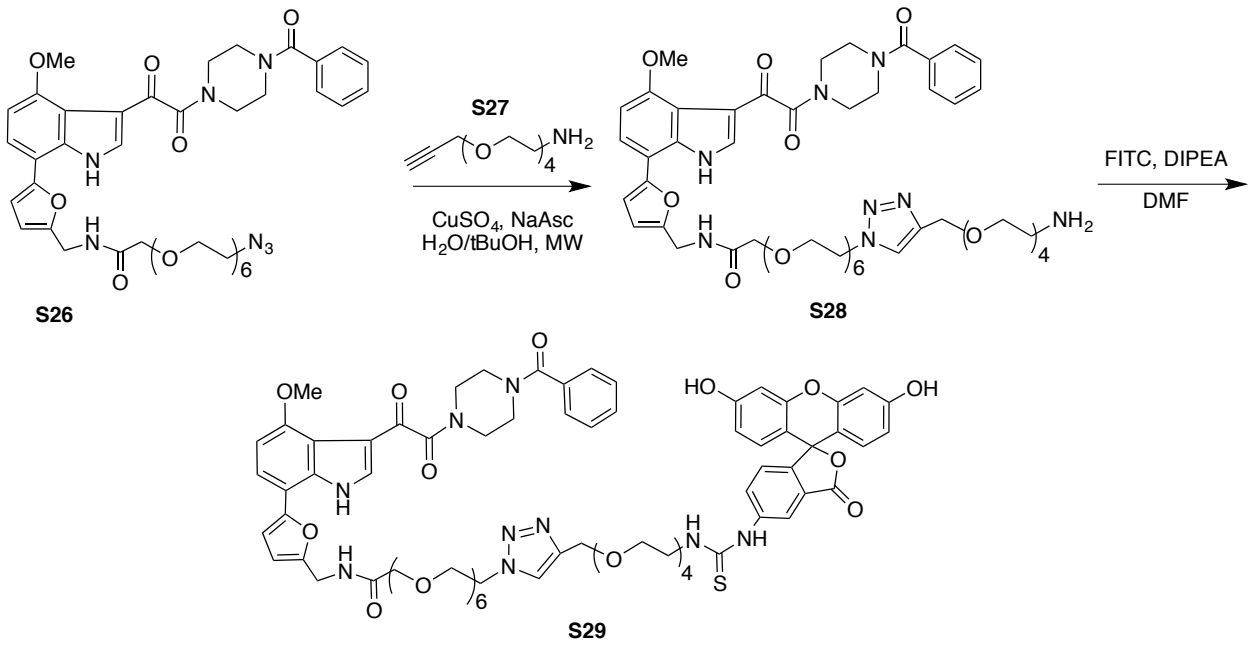


(**S26**) To **S25** (350 mg, 0.959 mmol 1.3 equiv)) dissolved in DCM (20 mL), added diisopropylethylamine (250 μL), EDC HCl (300 mg, 1.56 mmol), HOBt (250 mg, 1.63 mmol) followed by **S21** (350 mg, 0.719 mmol). After 5 hrs, TLC (20:1 DCM:MeOH) showed reaction completion. Reaction mixture was diluted with DCM (10 mL) and washed with saturated $NaHCO_3$ (3 x 30 mL), 2 M HCl (1 x 30 mL) and brine (1 x 30 mL) and subsequently dried over anhydrous $MgSO_4$. Solution was filtered and volatile solvents were removed by rotary evaporation. Crude yellow-orange product was purified by flash chromatography (flash chromatography (CombiFlash Automated Chromatographer, 24g colum; gradient elution ranging from 0% MeOH:DCM to 10% MeOH:DCM was performed over 30 column volumes) to yield pure **S26** as a light yellow sticky solid (360 mg, 0.088 mmol, 60%). **IR (ATR-FTIR)** 3409 (br), 2958 (w), 2872 (m), 2108 (s), 1634 (s), 1513 (m), 1421 (m), 1284 (m), 1105 (s), 987 (w) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 11.32 (s, 1H), 8.12 (d, J = 3.3, 1H), 8.07 (s, 1H), 7.38 (d, J = 8.1, 6H), 6.67 (d, J = 8.4, 1H), 6.51 (d, J = 3.2, 1H), 6.29 (d, J = 3.2, 1H), 4.52 (d, J = 6.2, 2H), 4.01 (s, 2H), 3.92 (s, 3H), 3.80-3.42 (m, 28H), 3.32 – 3.25 (m, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 186.09, 171.79, 170.68, 167.64, 153.69, 152.30, 150.30, 136.06, 135.15, 134.25, 130.13, 128.68, 127.10, 120.16, 115.62, 114.81, 109.58, 108.87, 104.03, 103.85, 71.05, 70.58, 70.55, 70.50, 70.41, 70.37, 70.02, 69.97, 56.02, 50.61, 35.97. **HRMS (ESI+)** calc'd for $C_{41}H_{52}N_7O_{12}^+$ $[M+H]^+$ *m/z* 834.3668. Found 834.3729; for $[M + Na]^+$, calc'd 856.3488, found 856.3502.

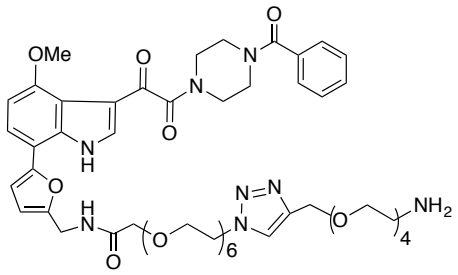


(**9**) To a microwave vial containing **S26** (200 mg, 0.240 mmol) in tBuOH/water (2 mL/2 mL), added alkyne **S9** (100 mg, 0.261 mmol, 1.1 equiv), followed by 120 μL of 0.1M $CuSO_4$ and 240 μL of 0.1M sodium ascorbate. The reaction mixture was capped and heated to 130 °C for 30 minutes in microwave reactor when LC/MS showed reaction completion. Volatile solvents were removed by rotary evaporation and crude product was purified by HPLC (30-50% B, 36 min). Like fractions were combined and volatile solvents were removed by rotary evaporation, resulting in **9** as a yellow sticky solid (281 mg, 0.228 mmol, 95%). **IR (ATR-FTIR)** 3357 (br), 3112 (w) 2872 (m), 1620 (s), 1512 (m), 1428 (m), 1335 (m), 1283 (m), 1105 (s) cm^{-1}. **^{1}H NMR (400 MHz, $CDCl_3$)** δ 11.21 (s, 1H), 9.09 (d, J = 2.5, 1H), 8.77 (s, 1H), 8.18 (m, 3H), 7.74 (s, 1H), 7.52 – 7.34 (m, 6H),

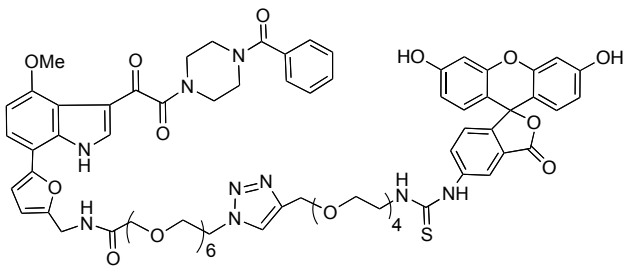
6.92 (d, J = 9.5, 1H), 6.69 (d, J = 8.4, 1H), 6.54 (d, J = 3.2, 1H), 6.32 (d, J = 3.1, 1H), 4.65 (s, 2H), 4.58 (d, J = 6.0, 2H), 4.45 (t, J = 4.9, 2H), 4.08 (s, 2H), 3.94 (s, 3H), 3.83 – 3.28 (m, 46H). **^{13}C NMR (125 MHz, $CDCl_3$)** δ 185.72, 172.32, 171.17, 167.58, 153.43, 152.07, 149.92, 148.25, 135.90, 135.73, 134.26, 133.90, 130.27, 130.16, 129.96, 128.59, 126.91, 124.32, 123.91, 120.28, 115.28, 114.53, 114.06, 109.34, 109.09, 104.28, 103.81, 80.25, 77.25, 77.00, 76.75, 70.76, 70.39, 70.30, 70.21, 70.11, 69.95, 69.85, 69.64, 69.59, 68.84, 68.38, 63.44, 55.80, 50.52, 42.97, 35.99. **HRMS (ESI+)** calc'd for $C_{58}H_{75}N_{10}O_{20}^+$ $[M+H]^+$ *m/z* 1231.5154. Found 1231.5178; for $[M + Na]^+$, calc'd 1253.4973, found 1253.5045.



Scheme S9. Synthesis of ARM-H based FP tracer (**S29**)



(**S28**) To a solution of **S26** (52 mg, 0.062 mmol) and alkyne **S27**[2] (19 mg, 0.082 mmol, 1.3 equiv) in *t*BuOH/water (900 uL/ 800 uL) in a microwave vial, added 0.1 M aq. CuSO4 (31 µL) and 0.1M aq. sodium ascorbate (62 µL). Reaction was subsequently heated in microwave reactor for 30 min at 130 °C when UPLC/MS indicated reaction completion. Volatiles were removed and crude **S28** was purified by flash chromatography (CombiFlash Automated Chromatographer, 4g colum; gradient elution 100% DCM for 5 column volumes, 0 → 50% MeOH in DCM for 40 CVs, the hold at 50% for 10 CVs, then 50-70% MeOH in DCM for remaining 20 CVs) to yield pure **S28** as a clear sticky solid (15.1 mg, 0.014 mmol, 23%). **^{1}H NMR (400 MHz, $CDCl_3$)** δ 11.65 (brs, 1H), 8.55 (brs, 1H), 8.19 (s, 1H), 7.83 (s, 1H), 7.42 (m, 6H), 6.69 (d, J = 8.4, 1H), 6.55 (d, J = 3.3, 1H), 6.33 (d, J = 3.3, 1H), 4.63 (s, 2H), 4.57 (d, J = 6.3, 2H), 4.52 – 4.43 (m, 2H), 4.24 (d, J = 2.4, 2H), 4.04 (s, 2H), 3.94 (s, 3H), 3.91 – 3.49 (m, 42H), 3.03 (s, 2H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 186.09, 171.79, 170.68, 167.64, 153.69, 152.30, 150.30, 136.06, 135.15, 134.25, 130.13, 128.68, 127.10, 120.16, 115.62, 114.81, 109.58, 108.87, 104.03, 103.85, 71.05, 70.58, 70.55, 70.50, 70.41, 70.37, 70.02, 69.97, 56.02, 50.61, 35.97. **HRMS (ESI+)** calc'd for $[M+H]^+C_{52}H_{73}N_8O_{16}^+$ *m/z* 1065.5139, found 1065.5206.



(**S29**) To a dried vial containing **S28** (70 mg, 0.066 mmol) in anhydrous DMF, added fluorescein isothiocyanate (43 mg, 0.111 mmol, 1.7 equiv) followed by DIPEA (600 μL), resulting in a color change from fluorescent yellow to fluorescent orange-yellow. Resulting mixture was protected from light and stirred under an atmosphere of nitrogen for 7 hrs when UPLC/MS indicated reaction completion. Volatiles were removed and crude **S29** was purified by HPLC (35-45% B over 46 min, t_R = 23.02 min). Like fractions were combined and volatiles were removed, resulting in **16** fluorescent yellow residue (36 mg, 0.025 mmol, 38%). **^{1}H NMR (400 MHz, 95%CDCl$_3$/5%MeOH-d$_4$)** δ 8.12 – 8.05 (m, 2H), 7.88 (dd, J = 1.9, 8.3, 1H), 7.65 (s, 1H), 7.33 (m, 6H), 7.00 (d, J = 8.3, 1H), 6.65 (m, 5H), 6.55 – 6.45 (m, 3H), 6.26 (d, J = 3.2, 1H), 4.51 (s, 2H), 4.45 (s, 2H), 4.36 (m, 2H), 3.96 (s, 2H), 3.86 (s, 3H), 3.83 – 3.36 (m, 44H), 3.32 (s, 2H). **^{13}C NMR (125 MHz, 95%CDCl$_3$/5%MeOH-d$_4$)** δ 185.86, 171.32, 170.95, 167.51, 153.04, 151.66, 149.83, 143.85, 141.08, 135.72, 134.13, 133.53, 129.94, 128.29, 126.44, 123.70, 120.39, 116.26, 114.84, 113.93, 109.27, 108.85, 104.34, 103.61, 102.16, 70.24, 69.76, 69.73, 69.69, 69.66, 69.63, 69.47, 69.01, 68.68, 63.54, 55.29, 49.71, 49.03, 43.71, 35.32. **HRMS (ESI+)** calc'd for $[M+H]^+$ $C_{73}H_{84}N_9O_{21}S$ *m/z* 1454.5497. Found 1454.5425.

Computations

Protein Preparation

The structure of gp120 was downloaded (PDB code 1G9N[4]) and imported into Maestro.[5] Chains C, H, and L were manually deleted. The N-terminus of Leu86 of chain G was acetyl capped. Hydrogen atoms were added and disulfide bridges were created between the residue pairs (Cys119-Cys205), (Cys218-Cys247), (Cys228-Cys239), (Cys296-Cys331), (Cys378-Cys445), and (Cys385-Cys418), using the protein preparation module.[6] Tautomer forms of the histidines were automatically assigned and the structure was energy minimized. The tautomers of histidines in the protein structure were assigned as follows: His105 – HSE, His249 – HSE, His330 – HSE, His216 – HSD, His374 – HSD. HSE is the neutral form of histidine that has the hydrogen on the epsilon nitrogen. HSD is the neutral form of histidine that has the hydrogen on the delta nitrogen.

Docking and Induced Fit Docking

All structures were docked and scored using Glide 5.5 in standard precision (SP)[7] and extra precision (XP) modes.[8] For rigid protein docking a receptor grid was prepared with the centroid located between residues Trp112, Val255, Glu370, Phe382, Trp427, and Met475. The grid box size was set to automatic. Induced fit docking was performed using the induced fit protocol in Maestro.[9] The centroid of the grid for the induced fit docking was located between residues Trp112, Val255, Thr257, Phe376 Phe382, Trp427, and Tyr435. In order to investigate the robustness of this method and the outcome of our results, we repeated these docking simulations using two additional gp120 starting structures (PDB codes 3JWD and 3TGQ). The previously prepared 3D structure of indol-furan was converted to smiles code and then converted back to 3D, using Ligprep, to avoid any conformational bias. As shown in Figure S4, we observed two different binding orientations, consistent with what was observed using starting structure PDB 1G9N.

Molecular Dynamics Simulations (PDB 1G9N)

For each ligand, BOSS was used to assign atom types, OPLS force field parameters,[10] and calculate CM1A charges.[11] The BOSS program scales CM1A charges by a factor of 1.14 for molecules of neutral charge. A program was written to automate the generation of formatted input files for the ligand. Topology for the carbohydrate groups

were created using standard OPLS-AA atom types for carbohydrates.[12] Parameters and CM1A charges for the carbohydrates were obtained using BOSS.

The protein complex was imported into VMD 1.8.7.[13] The protein-ligand complex was embedded in a triclinic box of approximately 16600 TIP3 water molecules,[14] extending 14 Å from all sides of the complex. The net charge of the system was neutralized by addition of one chlorine ion to the solvent box. The total number of atoms was approximately 54900 atoms.

Molecular dynamics simulations were performed using NAMD 2.8 for Linux_x86_64-CUDA.[15,16] Geometric mean was enabled to combine Lennard-Jones sigma parameters for different atom types and 1-4 interactions were scaled by 0.5. Van der Waals and short range electrostatic interactions were cut off at 10 Å. Switching with a switch distance of 8 Å was enabled to apply a smoothing function to both the electrostatics and van der Waals forces. Short range non-bonded interactions were calculated each time-step. Each cycle in the simulations consisted of 10 time-steps. A pair list distance of 14 Å was enabled, which was updated every cycle. Long range electrostatics were computed using the particle mesh Ewald method using a grid spacing of 1 Å.[17] The full electrostatics were calculated every two time-steps. Each time-step was 1 fs. Each system was initially subjected to 5000 steps of conjugate gradient minimization. Simulations were performed using Langevin dynamics applied to all heavy atoms with a damping coefficient of 0.5 ps^{-1} .[18] For pressure control at 1 atm, the Nosé-Hoover Langevin piston was used with a Langevin piston period of 100 fs and a Langevin piston decay of 50 fs..[19,20] An initial 100 ps equilibration was performed, during which time the temperature was brought up to 310 K. Harmonic constraints (10 kcal/mol/$Å^2$) were applied to all heavy solute atoms during the 100 ps simulation. The harmonic constraints were then removed and the simulations were continued, removing initial center of mass movement, using otherwise unchanged settings. Coordinates were saved every 10 ps and analyzed using the program VMD.

BOSS to NAMD file conversion program

The program reads an out-file and a z-matrix, generated by the program BOSS, in addition to the force field files oplsaa.par and oplsaa.sb (which are available through distributions of BOSS). It will output formatted parameter and topology files in the CHARMM format that readily can be used with the program NAMD. The program was written in C++ and can be compiled in a Unix or Linux environment. The program can be obtained by sending a request to either of the following two email addresses:

Markus Dahlgren:

markusan2000@yahoo.se

Peter Dahlgren:

peter.t.dahlgren@gmail.com

Quantum Mechanics Calculations

Torsional energy profile for the rotation around the indole C7-furan C2 bond was calculated at the MP2(fc)/6-311G(d,p) level of theory using Gaussian09.[21]

RMSD Analysis

The RMSD values of all α-carbon atoms within 5Å of the ligand in the final frame of the 214 ns MD simulation relative to same atoms of the crystal structure of gp120 (1G9N) were calculated using the program USCF Chimera.[22]

Data was fitted and graphed using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com).

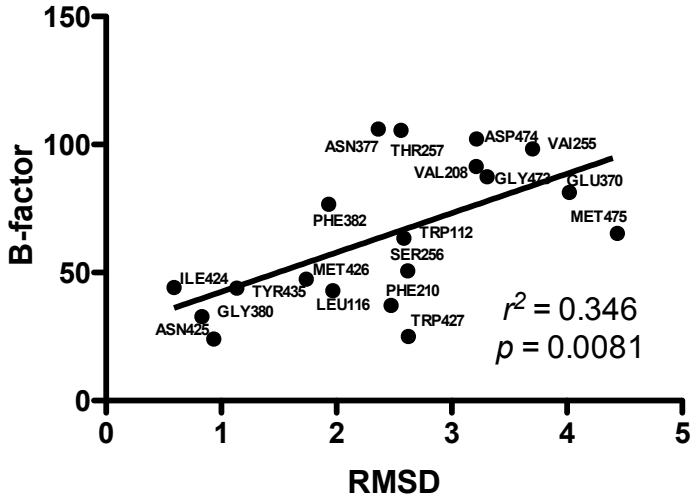
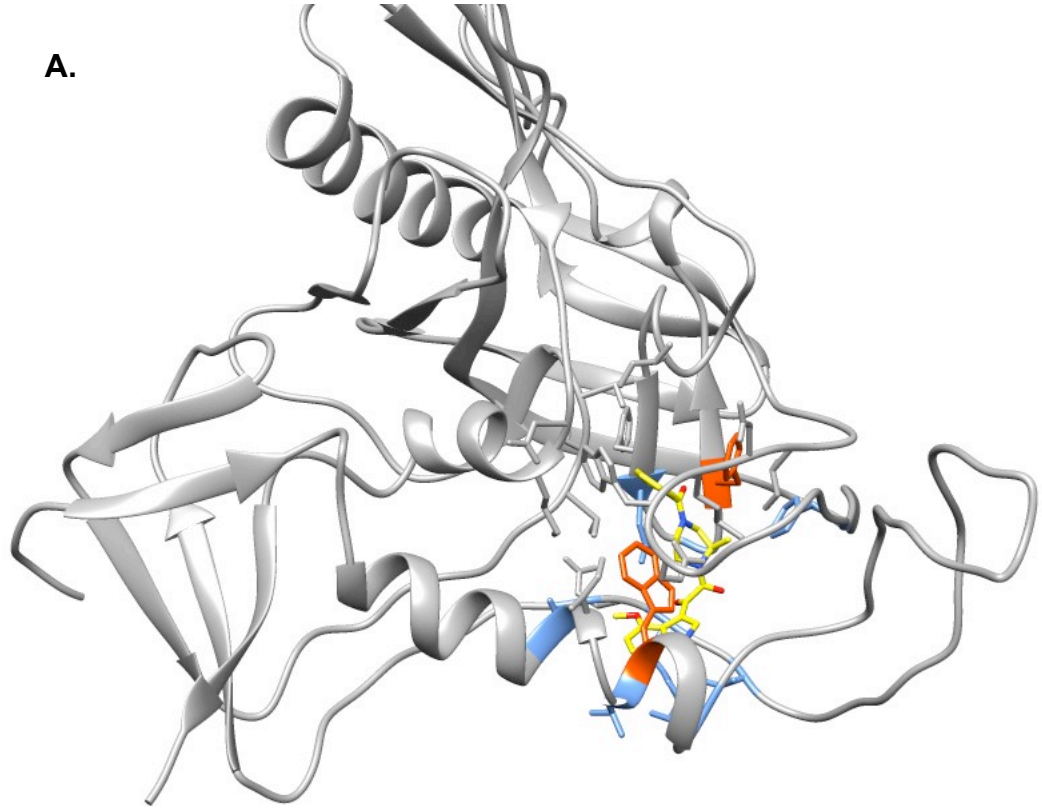
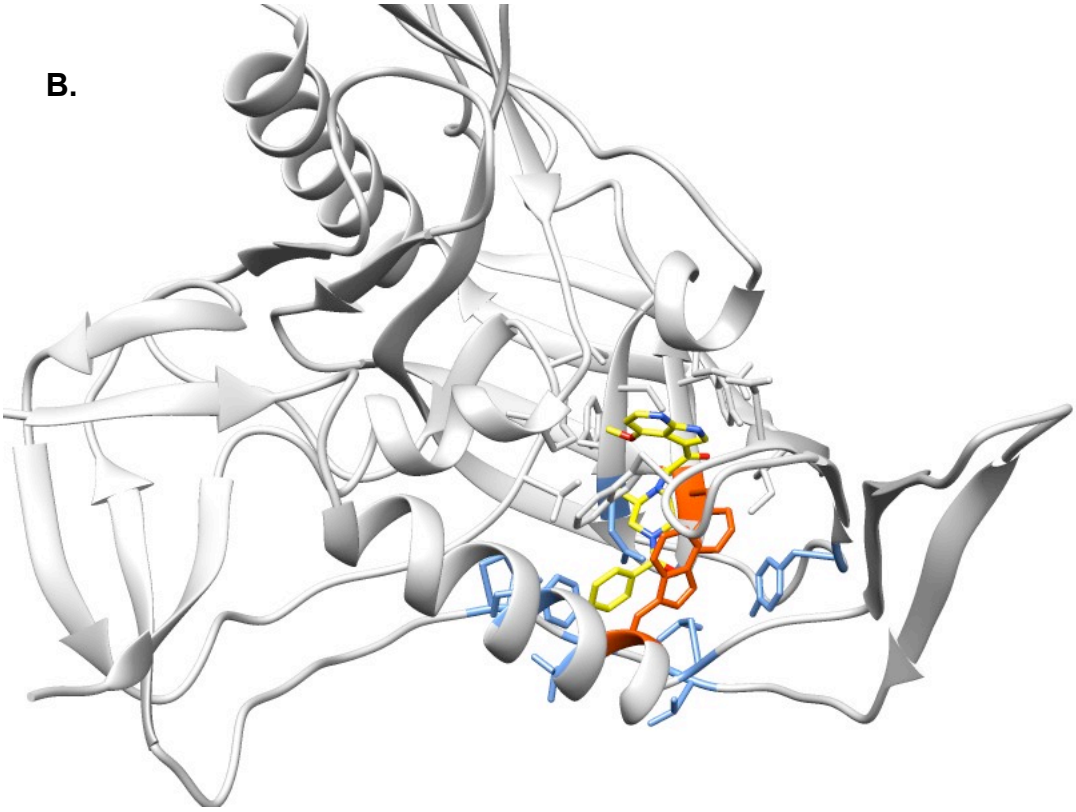


Figure S1. Comparison of residue B-factor and RMSD values.

A.



B.



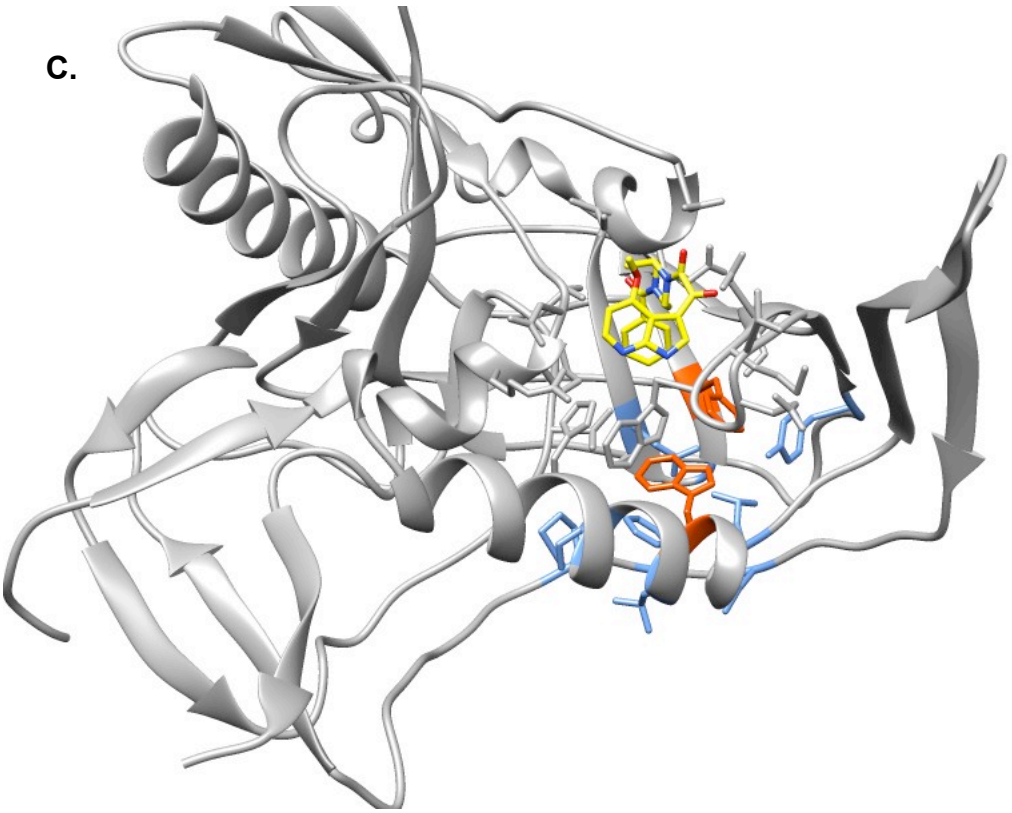
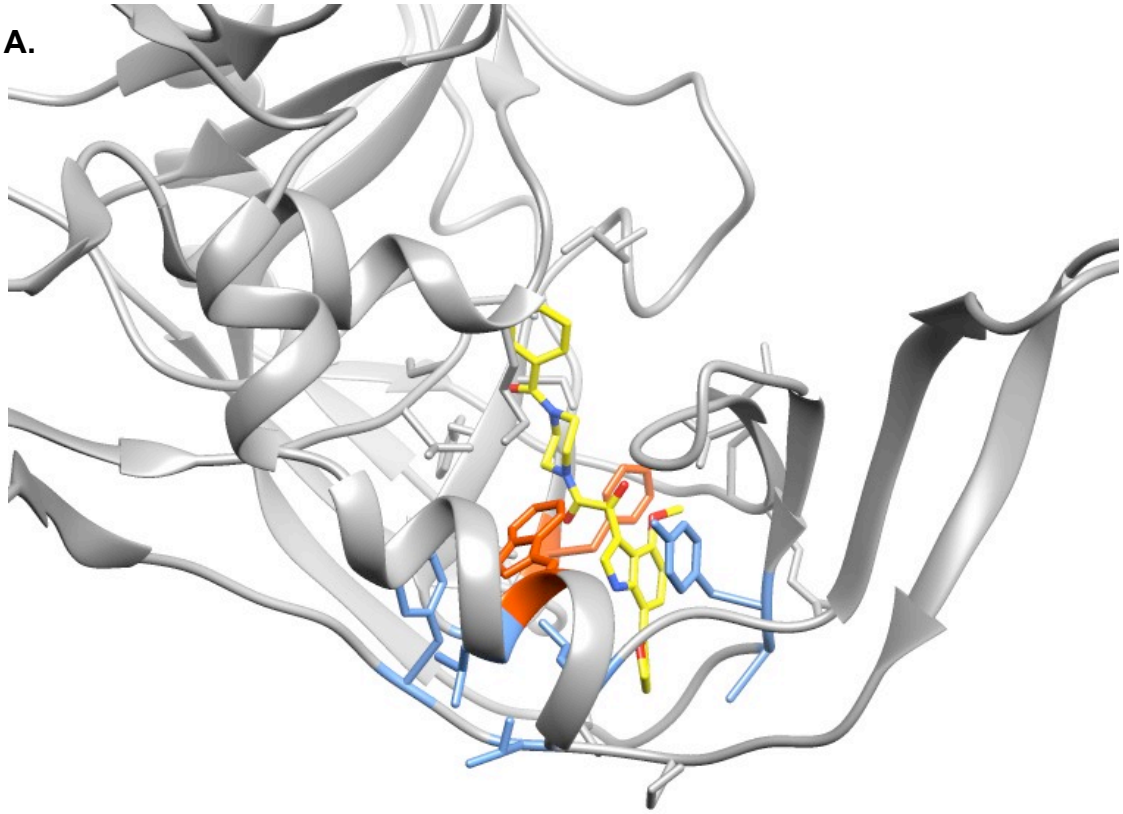
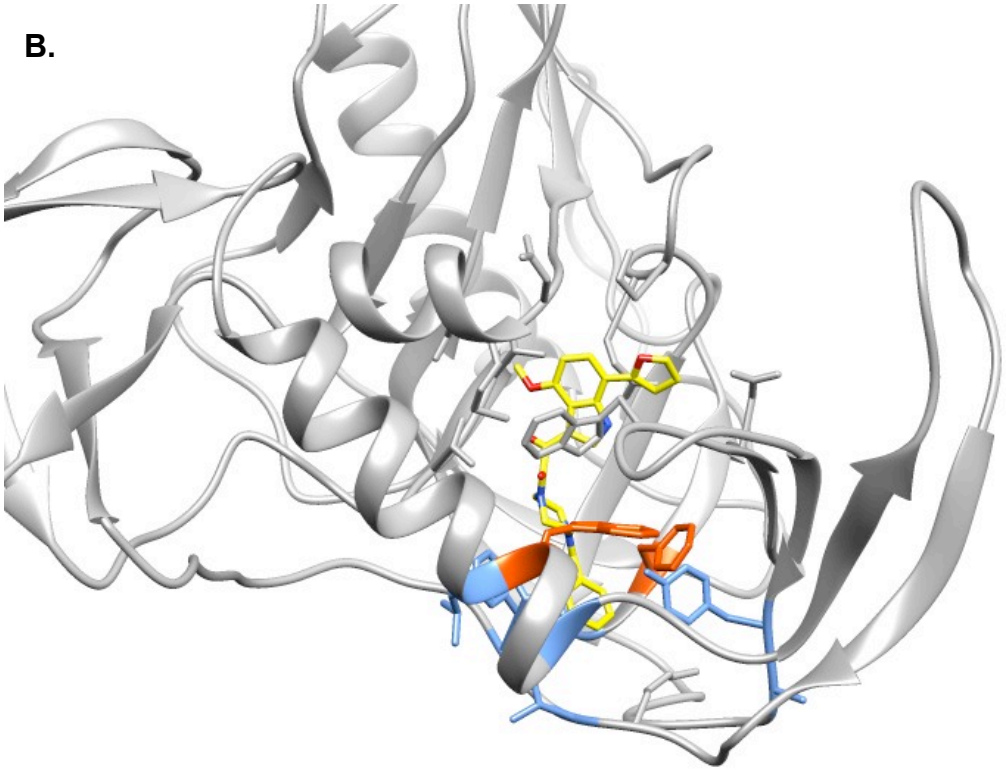


Figure S2. Full structure from Figure 2 in main text. Results of computational studies exploring the interaction between BMS-378806 and gp120. Analysis of rigid-body computational data revealed presence of two "gate-keeper" residues (W112 and F382, orange), which blocked a small accessory hydrophobic pocket (blue). BMS-378806 was docked using "induced-fit" protocol, wherein residues W112 and F382 were mutated to alanine and then re-incorporated after docking. The resulting poses suggested that BMS-378806 could access an accessory hydrophobic pocket in two low-energy orientations (A,B), and a higher-energy exposed orientation (C).

A.



B.



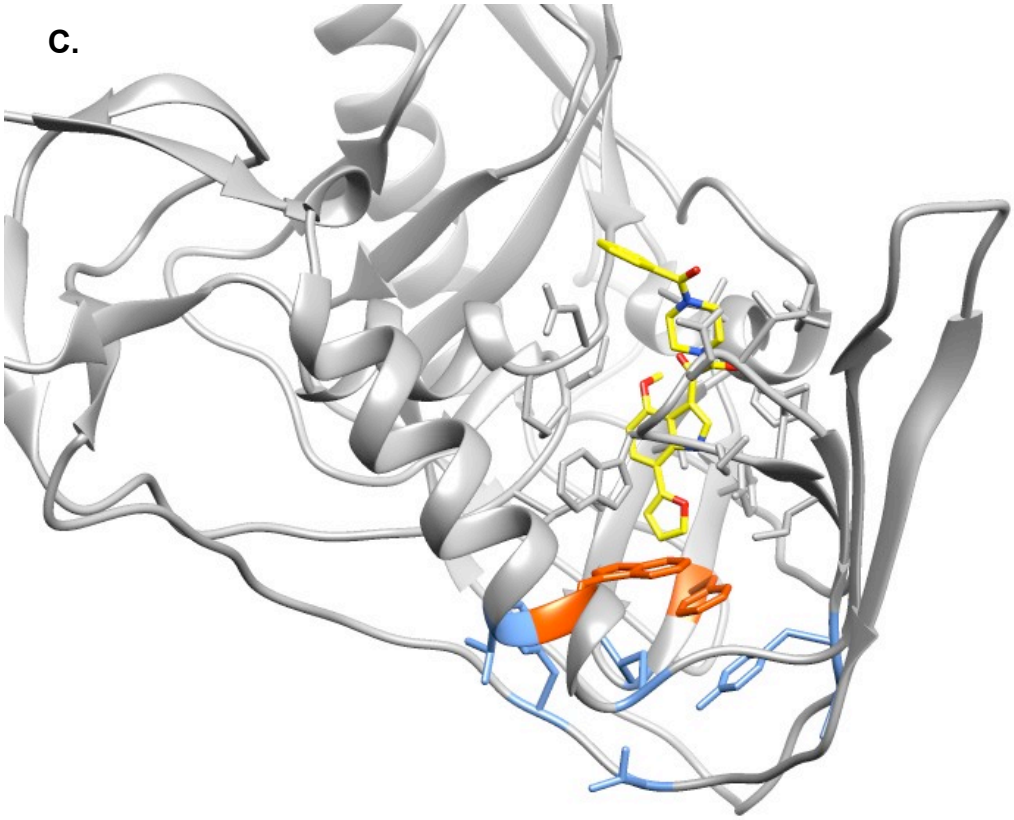


Figure S3. Full structure from Figure 3 in main text. Results of computational studies of the complex between **7** and gp120 (PDB 1G9N). Induced-fit docking of the compound in which the "gate-keeper" residues (Trp112 and Phe382, orange) were "trimmed" (mutated to alanine), resulted in two low-energy ligand orientations (panels A and B, respectively), and one high-energy orientation in which the ligand is mostly exposed (Orientation 3, panel C). Subsequent MD simulations indicate that in Orientations 1 and 2, compound **7** accesses an accessory hydrophobic pocket (blue). MD simulation for Orientation 1 was performed for 25 ns with an equilibrium structure observed after 11 ns. For Orientation 2, MD simulation was performed for 214 ns, with equilibrium reached after 178 ns that was stable for the remainder of the simulation.

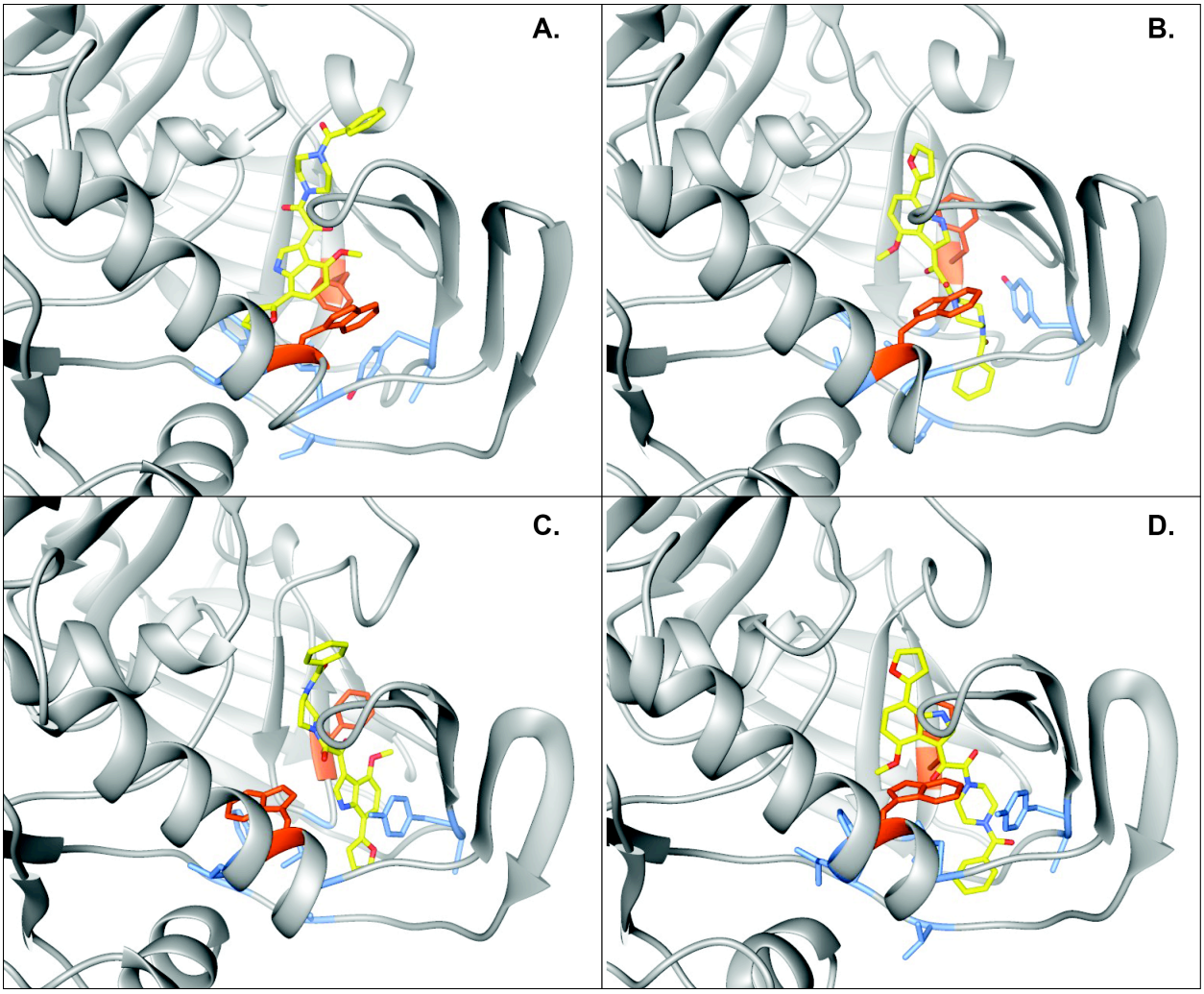


Figure S4. Induced-fit docking procedure of 7 (yellow) applied to two additional gp120 starting structures (gray): PDB 3JWD (panels A and B) and 3TGS (panels C and D). Induced-fit docking of the compound in which the "gate-keeper" residues (Trp112 and Phe382, orange) which blocked a hydrophobic pocket (blue) were "trimmed" (mutated to alanine), resulted in two low-energy ligand orientations (panels A/C and B/D respectively) in which the phenyl ring is solvent exposed (A and C) or the furan is solvent exposed (B and D).

Biology

General Information

All reagents and proteins used are commercially available and used as received unless otherwise noted. Unless otherwise noted, all micro-plate based assays were quantitated using a BioTek Synergy 3 Microplate reader and data was fitted and graphed using GraphPad Prism version 5.00 for Windows (GraphPad Software, San Diego California USA, www.graphpad.com).

Buffer A = 50mM TrisHCl, 100mM NaCl, 1% BSA, pH 7.5
Buffer B = 20mM TrisHCl, 500mM NaCl, 0.05% Tween 20, pH 7.5
Buffer C = 50mM carbonate/bicarbonate buffer, pH 9.6 (Aldrich)
FP Buffer = DPBS (Gibco) + 0.05% Tween 20

Anti-DNP IgG1 Production

The U7.6 hybridoma cell lines was kindly donated by the Eshhar Lab at the Weizmann Institute and culture and purification conditions were adapted from publications from the Eshhar Lab[23,24] and manufacturers instructions for CELLine1000 bioreactors (VWR International #37003-008) and Protein G Spin Column Kit (Thermo Scientific #89961).

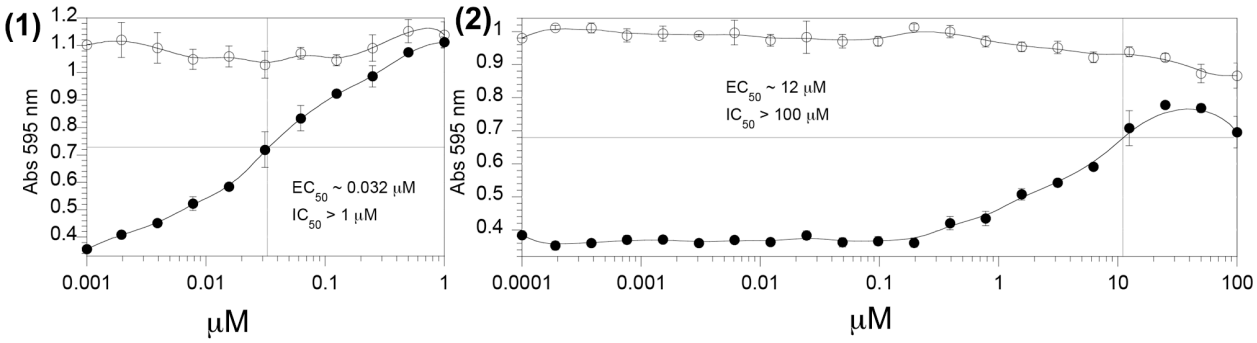
Briefly, cells were thawed and initially cultured in 75cm2 T-flasks using High Glucose Dulbecco's Modified Eagle Medium with pyruvate and L-glutamine (DMEM) (Invitrogen #11995-065**)** with 15% Horse Serum(HS) (Invitrogen #16050-114) and 1% Penacillin/Streptamycin (P/S) (Invitrogen #10378-016). Cells were passaged every 3 days until of sufficient viability (>90%) to transition to CELLine1000 Bioreactors (VWR International #37003-008). The cell compartment of the CELLine1000 Bioreactors were inoculated with 50-100million cells by pelleting and resuspending 3-4 confluent T-flasks in cell compartment media (DMEM + 15% IgG Depleted Fetal Bovine Serum (Invitrogen #16250-078) + 1% P/S and otherwise following manufacturer's instructions. The nutrient compartment of the CELLine1000 Bioreactor was filled with one liter cell "compartment media" (DMEM + 5% HS + 1% P/S) and stored at 37 degrees in a 5% CO2 atmosphere for 5 days until harvest.

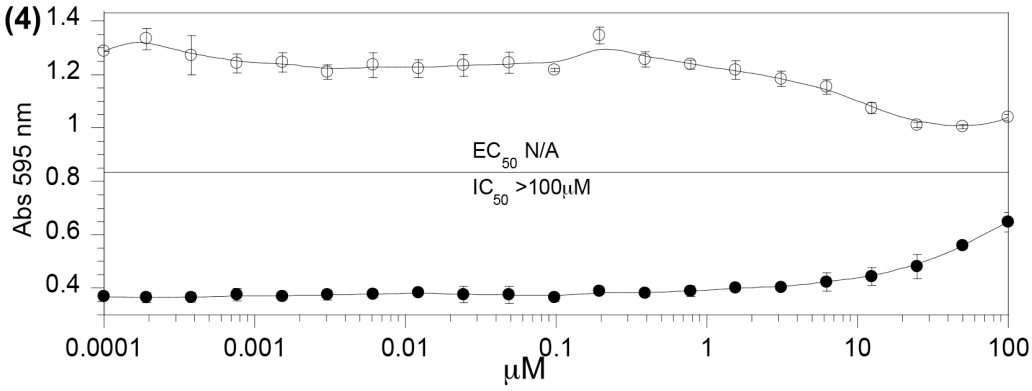
Harvests were collected after approximately 5 days, when cells had reached 20-30 million cells/mL, by pelleting (0.3 rcf for 10 min) the cell-compartment suspension and collecting the supernatant. For long-term cultures the cell-compartment must be washed with 3x15mL of wash media (DMEM + 1%P/S) after every harvest but before re-inoculation.

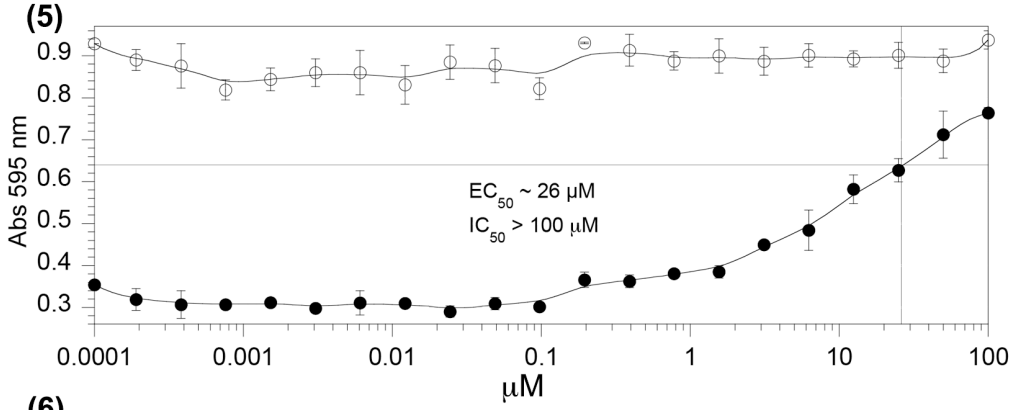
The supernatant (containing the anti-dinitrophenol (αDNP) IgG1) was then clarified by pelleting cellular debris(3 rcf for 10 min) and filtering through a 0.22 micron filter at which point it was stored at -80 degrees. The IgG was purified by 5mL Protein G Spin Columns (Thermo Scientific #89961) according to manufacters instructures. >95% purity was confirmed by reducing SDS-PAGE and activity (0.5nM (including avidity)) and Isotype (IgG1) were confirmed using sandwich ELISA format with DNP-modified Bovine Serum Albumin and a anti-mouse IgG1 secondary antibody (Novus Biologicals #NB7511).

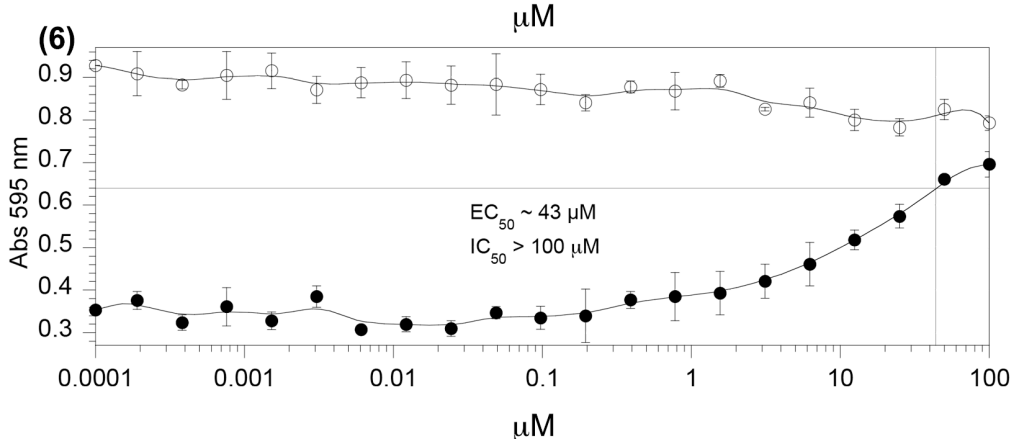
MT-2 Cell Assay

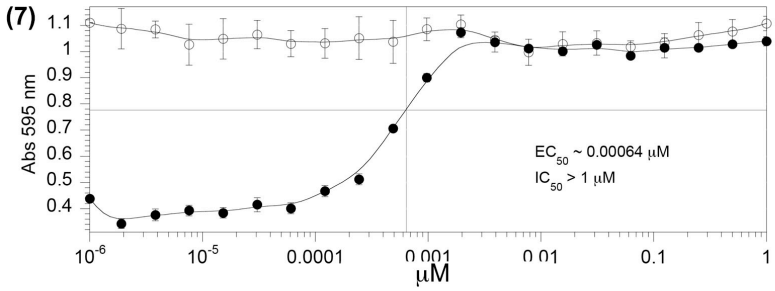
This procedure was adapted from our previously reported protocol.[2] Antiviral activity and cellular toxicity were determined using the MTT colorimetric method. MT-2 cell at a concentration of 1 x 10^5 cells per millilitre were infected with wild-type HIV IIIB at a multiplicity of infection (MOI) of 0.1. Infected and mock-infected cells were incubated in growth medium (RPMI 1640, 10% HI-FBS, kanamycin) for 5 days with varying concentrations of each compound being tested in triplicate in a 96-well plate. MTT (thiazolyl blue tetrazolium bromide), a cell-permeable tetrazolium dye was then added to each well. Active mitochondria reduce the yellow tetrazolium salt to a blue formazan precipitate. After 5 h, stop solution (86% isopropanol, 4% NP-40, 10% H_2O, and 0.3% concentrated HCl) was added to lyse the cells and stop the reaction. The plates were gently shaken gently overnight on a horizontal rotator, and quantitated the following morning. Cell viability was measured spectrophotometrically by quantifying the amount of purple precipitate via determining the absorbance at 595 nm using a Multiskan Plus from Labsystems (Helsinki, Finland) microplate reader. The average of these triplicate samples was then plotted versus inhibitor concentration to generate dose–response curves (± SD). The 50% effective concentration (EC_{50}) and 50% cytotoxic concentration (CC_{50}) of the compounds were defined as the concentrations required to inhibit viral replication and to reduce the number of viable cells by 50%, respectively. Positive controls were done during each set of experiments using d4T and the appropriate parent NNRTI (HI-236 or TMC-derivative). Data were quantitated using KaleidaGraph (Synergy Software). Assay was repeated in triplicate at least two times for each molecule examined. See Figure S5 for representative concentration curves for each molecule.

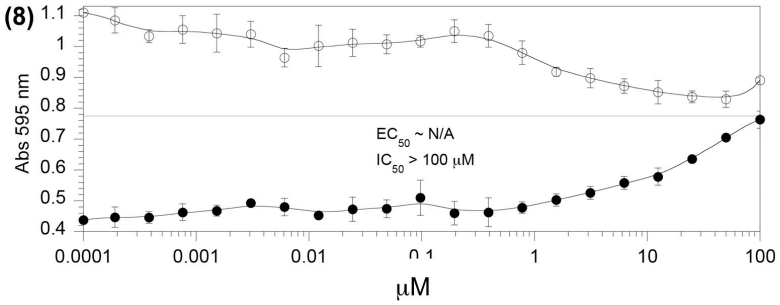

(1)
Abs 595 nm
EC50 ~ 0.032 μM
IC50 > 1 μM
μM
(2)
Abs 595 nm
EC50 ~ 12 μM
IC50 > 100 μM
μM

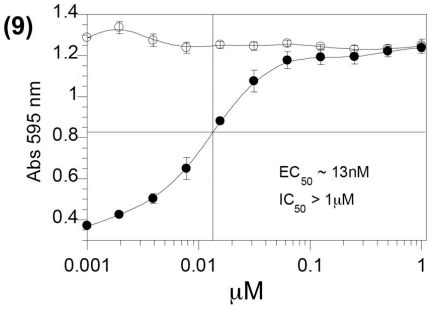

(4)
Abs 595 nm
EC50 N/A
IC50 >100μM
μM

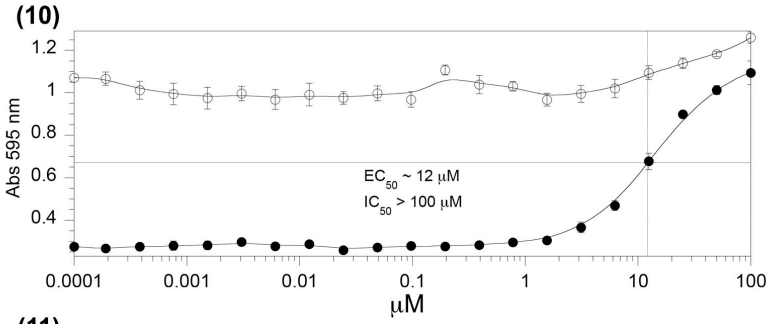

(5)
Abs 595 nm
EC50 ~ 26 μM
IC50 > 100 μM
μM

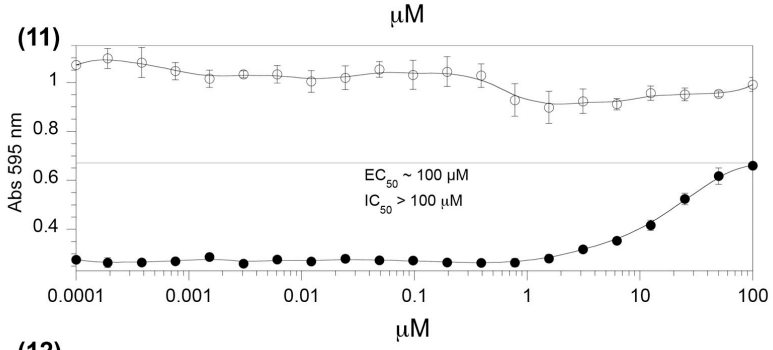

(6)
Abs 595 nm
EC50 ~ 43 μM
IC50 > 100 μM
μM


(7)
Abs 595 nm
EC50 ~ 0.00064 μM
IC50 > 1 μM
μM


(8)
Abs 595 nm
EC50 ~ N/A
IC50 > 100 μM
μM


(9)
Abs 595 nm
EC50 ~ 13nM
IC50 > 1μM
μM


(10)
Abs 595 nm
EC50 ~ 12 μM
IC50 > 100 μM
μM


(11)
Abs 595 nm
EC50 ~ 100 μM
IC50 > 100 μM
μM

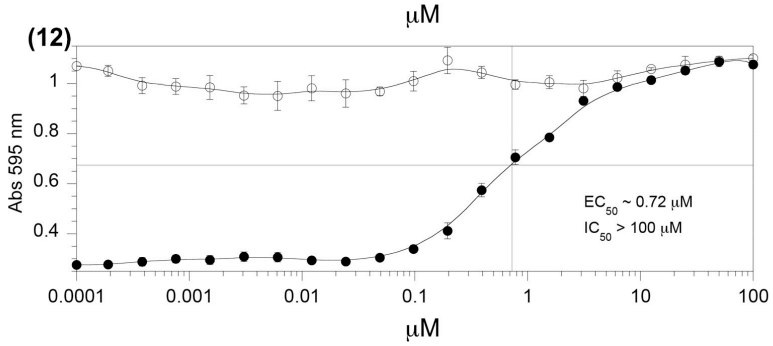

(12)
Abs 595 nm
EC50 ~ 0.72 μM
IC50 > 100 μM
μM

Figure S5. HIV-1 viral replication assay. Survival of MT-2 cells is measured in the presence of increasing concentrations of ARM-H alone (white circles), or ARM-H plus live HIV-1 virus (black circles). Data represented as means (±SD) from triplicate experiments.

Viral Pseudotype Inhibition Assay

Inhibition of HIV-1 pseudotyped vector supernatants by ARM-H compounds was tested by standard neutralization assay (or pseudotype inhibition assay), by measuring reduction in luciferase reporter gene expression as described previously.[25-27]

For laboratory clones

Production of Viral Pseudotypes for Pseudotype Inhibition Assay

Various HIV-1 pseudotype particles were prepared by calcium phosphate co-transfection of plasmids encoding different HIV-1 envelopes (env) (e.g., ADA, JRFL) and VSV-G, the latter as non-specific control glycoprotein, along with an env-deficient HIV-1 backbone vector (pHIV-eYFP-IRES-blasti) in 1:1 ratio of DNA amounts. Supernatants were harvested 48 h after transfection, filtered through 0.45 μm membrane, and stored until use in aliquots at -80° C.

Pseudotype Inhibition Assay

100 μL of pseudotyped viruses were incubated with 3-fold serial dilutions of ARM-H test compounds in duplicate for 1h at room temperature or at 37°C under sterile conditions. After 1h, either compounds mixed with viral supernatants were added to TZM-bl cells (~40% confluence) seeded overnight in 96-well assay format, or freshly trypsinized cells (10,000 cells/well in 100 μL growth medium) were added to virus-compound mixes, for single round virus infection assays. Several control wells were prepared for each assay in duplicates, including background controls with cells only, compounds mixed with culture medium only (control for intrinsic fluorescence/signal if any), and virus plus cells without any compound ("0" compound concentration or maximum signal point on data charts).

Media was replenished after 24 h and luciferase reporter gene activity was measured as relative light units (RLUs) 48 h after infection. No difference in neutralization was observed between the two assay formats, and the latter format was used in all of the assays reported here. Luminescence measurements were performed using Biotek SynergyMx plate reader. The 50% inhibitory concentration (IC_{50}) was defined as the compound concentration where the luminescence was reduced to half maximal value. Data were plotted using GraphPad Prism software (v.5/6), which also calculated the IC_{50} values for each fitted curve. For the luciferase assay, wells were washed with PBS twice, 100 μL triton/glycylglycine lysis buffer was added to each well and incubated at 4°C for 10 min. Luciferase assay buffer with ATP was then added, followed by luciferin substrate. Within seconds, emitted light, proportional to the amount of luciferase enzyme present in lysed TZM-bl cells, was measured as relative light units (RLU). Assay was repeated in triplicate at least two times for each molecule examined.

For clinical clones

Envelope clones

Functional gp160 envelope clones were obtained from baseline samples from subjects enrolled in proof of concept studies for BMS attachment inhibitors (AIs) BMS-488043 or BMS-663068 (a prodrug of BMS-626529 in clinical development).[28,29] A total of 12 functional subtype B clinical envelope clones from 9 subjects were used in the cell-cell fusion susceptibility studies. Envelopes were chosen to encompass a wide range of susceptibilities to AIs, from highly susceptible (<1 nM EC_{50}) to greatly reduced sensitivity (> 6 mM EC_{50}).

Cell-cell fusion assay

A cell-cell fusion assay was used to evaluate the sensitivity of HIV-1 envelopes to BMS-626529 and **9**. Briefly, 100 ng of each envelope clone present in the pCMV-HA expression vector (Clontech Labs, Mountain View, CA) was mixed with 1 μg of the pTet-off plasmid (Clontech) and transfected with lipofectamine 2000 (Life Technologies,

Grand Island, NY) into 1.2 x 10^6 HeLa cells to create effector cells. One day after transfection, the effector cells were mixed with target HeLa cells expressing CD4, CCR5 and CXCR4 and containing an integrated copy of pTRE2Hyg-Luc (Clontech), a tetracycline inducible luciferase reporter gene. The effector and target cells were mixed at a ratio of 1:1 and seeded into a 96 well plate in the presence of 5-fold serial dilutions of BMS-626529 or XX. All clinical envelopes were examined in the same experiment in quadruplicate, along with the LAI envelope clone.[30] After 18 hours of incubation, the extent of cell fusion was measured through luciferase expression using the Steady-Glo luciferase assay (Promega, Madison, WI). Susceptibility was determined based upon differences from luciferase expression observed in an untreated sample and expressed as EC_{50} or Fold-Change, the latter normalized by the LAI EC_{50}. The experiment was repeated three times and values were averaged.

Results

	EC_{50} (nM)		Fold Change (vs LAI)		FC Ratio
Clone ID	**9**	BMS-626529	**9**	BMS-626529	**9**/'529
27N4-20	7.4±2.6	0.2±0.08	0.22	0.20	1.1
36N3-5	5.5±2.1	0.4±0.20	0.16	0.40	0.40
P15-2-18	62±38	1.2±0.5	1.8	1.2	1.5
P23-1-13	110±50	2.4±0.5	3.2	2.4	1.3
P22-2-45	73±29	2.5±1.1	2.1	2.5	0.84
66N2-8	160±40	8.6±2.1	4.7	8.6	0.55
66N2-16	350±120	22±12	10	22	0.45
12N1-13	550±230	56±12	16	56	0.29
21N4-169	4300±1800	120±40	130	120	1.1
21N4-170	11000±4000	860±120	320	860	0.38
16N4-17	>40000	6200±1700	>1200	6200	
16N4-25	>40000	6400±1900	>1200	6400	
LAI	34±22	1.0±0.3	1	1	

Table S1. Inhibitory activity of ARM-H 9 and BMS-626529 against HIV-1 clinical envelope clones.

CD4 Inhibition ELISA

This procedure was adapted from our previously reported protocol.[2] 96 well plates (Nunc; Immuno) were coated overnight (12 hr) at 4 ^oC with soluble recombinant HIV-1 $gp120_{JRFL}$ (Immune Technology; Yonkers, NY) at 1 µg/ml in 0.05M Buffer C. Plates were washed with DPBS (Gibco, 1 x 100µL) and then blocked with Buffer A for 1 hr at room temperature. After washing with Buffer B (3 x 100µL) , varying concentrations of the inhibitor (including a "no molecule" control) were added simultaneously with recombinant human T-cell CD4 (ImmunoDiagnostics, Inc; Woburn, MA) in Buffer A in triplicate so that the final concentration/well of CD4 is 0.1 µg/mL and plates were incubated for 1 hr at room temperature. Plates were washed with Buffer B (3x100µL) and then incubated with mouse OKT4 anti-CD4 IgG antibody (Biolegend; San Diego, CA) at 0.36 µg/ml in Buffer A at RT for 1 hr. Following washes with Buffer B, plates were incubated with horse radish peroxidase (HRP)-conjugated goat anti-mouse antibody (1:2500; Biolegend; San Diego, CA). Following washes with Buffer B (3 x 100 µL), bound antibody was detected with 3,3,5,5-tetramethylbenzidine (TMB, Pierce Protein Research Products), the chromogenic substrate for HPR, and absorbance was read at 450 nm after stopping reaction with 2N H_2SO_4 (100 µL). The mean (±SD) of these triplicate samples was then plotted versus inhibitor concentration and a non-linear fit curve was generated using GraphPad Prism. The 50% inhibitory concentration (IC_{50}) was defined as the concentration of inhibitor to reduce the amount of bound CD4 to sgp120 by 50% of the maximum bound. Inhibitory assay was performed in triplicate at least two times for each molecule. See Figure S6 for representative concentration curves.

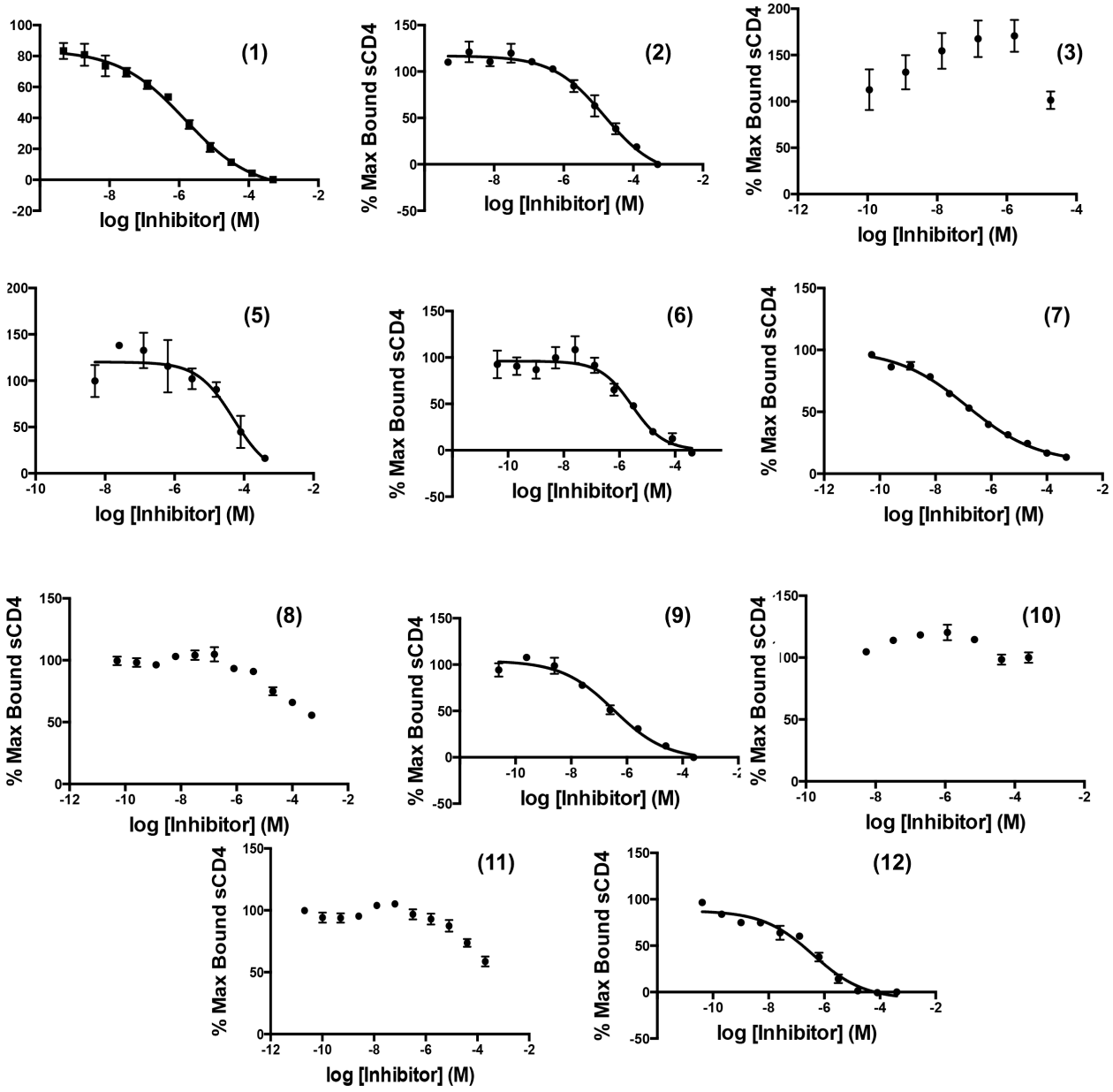


Figure S6. Competition ELISA monitoring the binding of sCD4 to immobilized gp120. Data represented as means (±SD) from triplicate experiments.

Anti-DNP IgG Recruiting ELISA

Nunc-Immuno 96-well plates were coated with soluble gp120 and blocked as described above. After the Buffer B wash, varying concentrations of ARM-H's (100 µL) were added (including a "no molecule" control) to the plate in triplicate and incubated for 1 hr at RT. After washing (3x100 µL) the plate with Buffer B, wells were incubated with mouse monoclonal anti-dinitrophenyl (anti-DNP) IgG1 antibodies (see above for production) at 5 µg/ml in Buffer A at room temperature for 1 hr. Following washes with Buffer B, plates were incubated with horse radish peroxidase (HRP)-conjugated goat anti-mouse antibody (1:2500; Biolegend; San Diego, CA). Following washes with Buffer B (3 x 100 µL), bound antibody was detected with 3,3,5,5-tetramethylbenzidine (TMB; Pierce Protein Research Products), and the absorbance was read at 450 nm after stopping reaction with 2N H_2SO_4 (100 µL). The mean (±SD) of these triplicate samples (with background absorbance subtracted) was then plotted versus molecule concentration and a non-linear fit curve was generated using GraphPad Prism. The 50% effective concentration

(EC_{50}) was defined as the concentration of ARM-H to bind 50% of the maximum bound anti-DNP in the ternary complex with sgp120. Anti-DNP IgG recruiting ELISA was performed in triplicate at least two times for each molecule. See Figure S7 for representative concentration curves.

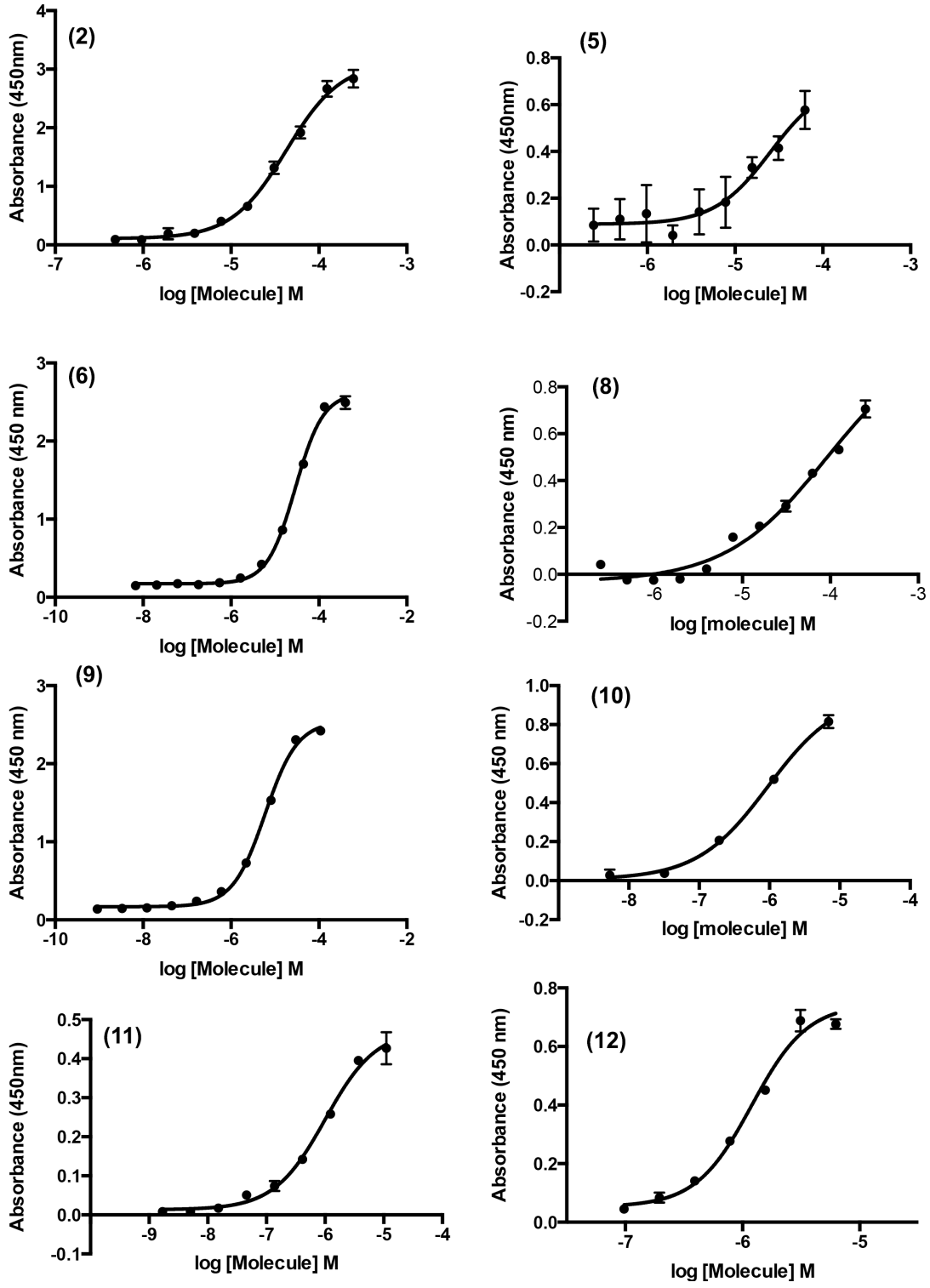
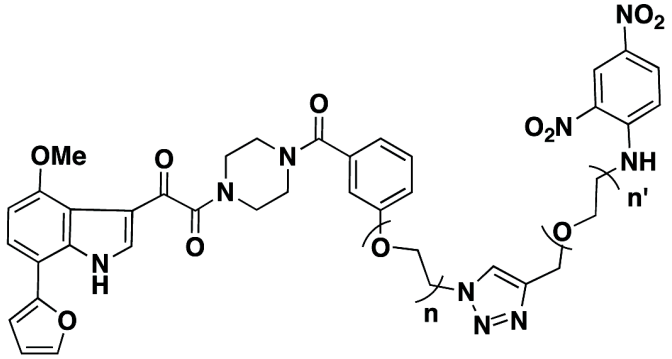


Figure S7. ELISA showing ARM-H concentration dependent increase in absorbance when mouse anti-DNP antibodies were allowed to bind to complex of ARM-H and gp120, as described above. Data represented as means (±SD) from triplicate experiments.

It is noteworthy that improvements in the gp120 binding constants (confirmed via fluorescence polarization and functional assays) do not appear to improve the EC50/potency of ARM-H's in this assay (Tables 1 and 2). While this may seen non-intuitive, it reflects that fact that "antibody recruiting" is dependent on the formation of a 3-component complex (gp120 – ARMH – antibody) as opposed to more common 2-component complexes (e.g. Receptor – Ligand) where EC_{50} generally correlates with the K_d. In the case of 3-component complexes there are two binding interactions to consider, of which only one or both may contribute to the EC_{50} as we have described in a recent publication.[31] Combining that mathematical model with the conditions of this assay we expect the K_d of the ARM-H/Antibody interaction to correlate with the "antibody recruiting" EC_{50} and thus be unaffected by changes in the gp120/ARM-H K_d.

Effect of Linker Length on CD4-inhibition and Antibody Recruiting

It has been reported that the addition of polyethylene glycol (PEG) can impair protein interactions by limiting the k_{on} rates of ligands to proteins.[32] To examine the potential effects of PEG length, we iteratively decreased the linker length of compounds based upon the ARM-H **12** scaffold (SI Scheme S6 and S7). Decreasing the linker length and removal of the triazole did not exhibit a significant effect on the inhibitory activity of the ARM-H's (Table S2). However, at lengths of two PEG units (**18**), we observed almost a complete loss of antibody binding capability. We hypothesize that the shorter linker lengths induce steric repulsion between the two proteins, but longer lengths do not significantly affect gp120 binding.



Cmpd	n	n'	EC_{50} Ab Recruiting (μM)	IC_{50} CD4 Inhibition (μM)
12	6	4	0.973	0.330
16	6	2	2.50	0.363
17	6	N.T.	13.9	0.220
18	2	N.T.	195	0.130

Table S2. Examination of the effect of ARM-H linker length on antibody recruiting and CD4 inhibition. Values represent means from triplicate experiments. Note: N.T. = no triazole.

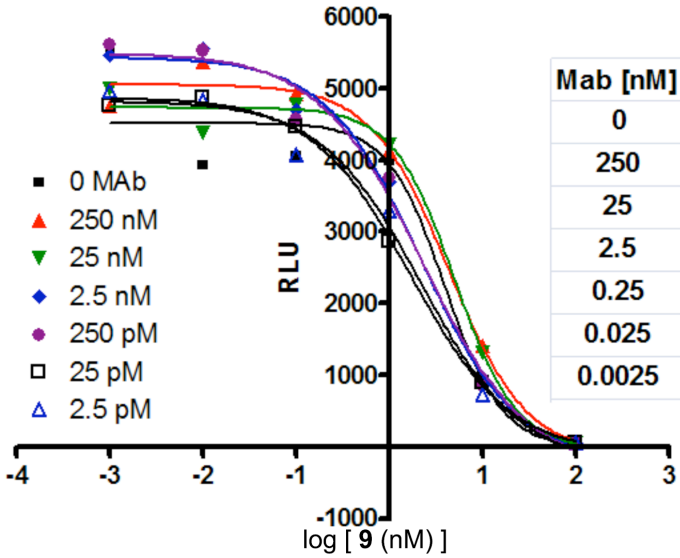
Viral and CD4 inhibition studies in the presence of anti-DNP antibodies

HIV has evolved several strategies to evade multivalent antibody binding, which are believed to have caused dramatic reductions in the neutralizing capacity of anti-HIV antibodies. First, Env trimers are expressed at very low levels on virus-infected cells, which results in a large average inter-spike distance (~30 nm).[33] Because the antigen-binding (Fab) regions of antibodies are separated by 10-15 nm,[34] they are unable to crosslink multiple Env

complexes such as to benefit from multivalency effects.[35] Second, due to their inherent geometrical constraints, these proteins are unable to crosslink gp120 monomers within a single Env trimer.[35,36] Indeed, protein chimeras with increased length and flexibility have been shown to possess dramatic enhancements in neutralizing potency due to the ability to crosslink gp120 monomers within a trimer.[34,37,38]

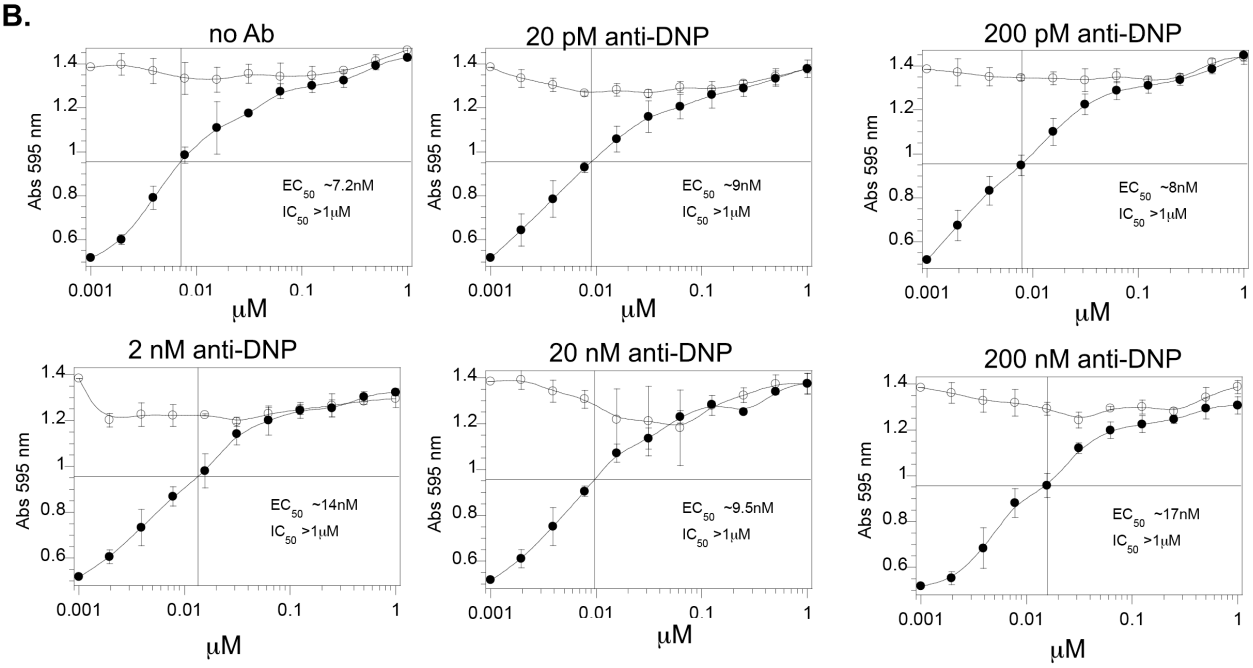
Provided this, we have examined the effect of inclusion of mouse monoclonal anti-DNP antibodies both in the viral pseudotype inhibition assays as well as for the MT-2 assays. These experiments were repeated in the same manner as described above (See Figure S8 below). Interestingly, in the CD4-inhibtion ELISA format, we indeed see significant enhancement of ARM-H-mediated CD4 inhibition in the presence of anti-DNP antibody. We hypothesize that immobilized gp120 monomers are within close enough proximity to enable anti-DNP antibody mediated gp120 binding enhancement. However, in viral and cell based assays, in which the Env is expressed as a trimer on the surfaces of HIV and infected cells, we observe no antibody-mediated enhancement. These observations suggest that the linker of our ARM-Hs are not long enough to effectively 'crosslink' the Fab's of the DNP antibodies in order to gain a similar avidity enhancement. Crystallographic studies from Kwong and colleagues[34] have demonstrated that the separation between the binding sites of IgG (10-15 nm) is similar to that between the Cα domains of multiple CD4's bound to the gp120 trimer (13 nm).[34,36] Based on this analysis, we hypothesize that ARM-H analogs possessing a 2-4 nm flexible linker should be able to link both binding sites of an IgG antibody to two monomers of the Env trimer.[34,36] Previous studies have measured the effective solution length of PEG linked molecules, and[39,40] suggest that a 2-4 nm distance translates to approximately 10-30 PEG units.[105,118] The current ARM-H structures examined have ~10 PEG units and therefore may be too short to effictively crosslink the Fab regions of IgG.. Such observations are indicative of non-cooperative ternary complex pre-equilibria (See Douglass, E. F. *et al. J. Am. Chem Soc.* **2013**, *135*, 6092)

A.

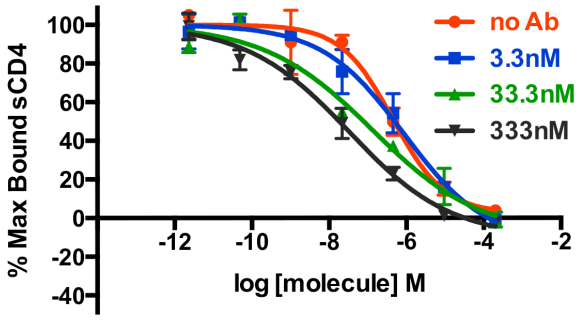


Mab [nM]	EC_{50} [nM]
0	3.923
250	4.337
25	5.003
2.5	2.208
0.25	2.264
0.025	1.707
0.0025	2.182

B.



C.



[Anti-DNP antibody]	no Ab	3.3nM	33.3nM	333nM
IC50	4.467e-007	8.790e-007	1.366e-007	3.239e-008

Figure S8. Effects of anti-DNP antibodies on CD4 and HIV-1 inhibition. (A) In HIV-1 (JRFL) pseudotype inhibition assay (B) MT-2 (HIV-1 IIIB) viral inihibition assay, ARM-H (**9**) co-incubated with increasing concentrations of mouse monoclonal anti-DNP IgG antibodies. (C) In CD4-gp120 (JRFL) inhibition ELISA, ARM-H (**9**) co-incubated with increasing amounts anti-DNP IgG antibodies. All assay performed as described above, with the only difference being the inclusion of anti-DNP antibodies. Data represented as means (±SD) from triplicate experiments.

Fluorescence Polarization – gp120 titrations to determine tracer binding affinity

Prepared a 10 μM stock solution of fluorescent tracer **S29** in FP buffer and from this stock, made a 50 nM solution and, with a multichannel pipette equipped with Finttip pipette tips, added stock (20 μL) to black, 384-well plate (Corning) in triplicate, except for the first column of the dilution series. To the first column, added a solution of 50 nM **17** and 1 μM stock of soluble HIV-1 $gp120_{JRFL}$ (Immune Technology; Yonkers, NY). From this column, using a multichannel pipette, made serial dilutions (1:5 v/v) into the wells containing tracer stock or just FP buffer (negative control). Protected plate from light and incubated at room temperature, monitoring anisotropy and polarization (Ex = 485±20 nm, Em = 528±20 nm) until observe equilibrium (no change in signal, 60 min). Data from titrations analyzed by a least-squares non-linear fit, generated using GraphPad Prism. This assay was performed three times in triplicate. Variation in fluorescence anisotropy occurs upon association of fluorescent tracer with gp120, and thus under equilibrium conditions, allows for the direct calculation of binding dissociation constants.[41] The observed large change in polarization affords a high signal-to-range, and hence sensitivity, providing a broad concentration range over which competition experiments can be conducted.[42] From the average of three replicate experiments performed in triplicate, a soluble HIV-1 $gp120_{JRFL}$:**S29** K_d^t of 36.3±6.7 nM (mean ± SE) was observed.

Competitive Fluorescence Polarization

Competition experiments were performed with fluorescent tracer **S29**, soluble gp120 and unlabelled competing ligands in order to determine each ligand's IC_{50}, which were then used to calculate subsequent K_d of the inhibitor (K_d^I) based on known total [gp120] (R_o), [bound tracer] (L_b), tracer K_d (K_d^t) and total [tracer] (L_o) using *equation 1*.[41,42] (Note: L_b and K_d^t determined from gp120 titration with tracer **S29** as described above).

$$K_d^I = \frac{L_b IC_{50} K_d^t}{R_o L_o + L_o(R_o - L_o + L_b - K_d^t)}$$

Equation 1

To wells of a black 384-well plate (Corning), added a standardized stock solution of soluble gp120 with tracer **17** (typically 100nM/50nM or 200nM/50nM gp120:tracer) with multichannel pipette in FP buffer. To the first column of this aliquoted solution, added the highest desired concentration of competing ligand in FP buffer containing gp120 and tracer. From this solution, made serial dilutions (in quadruplicate) of competing ligand with multichannel pipette (1:5 v/v). Protected plate from light and incubated at room temperature, monitoring anisotropy and polarization (Ex = 485±20 nm, Em = 528±20 nm) until observe equilibrium (no change in signal, typically 30 min). Data from titrations analyzed by a least-squares non-linear fit, generated using GraphPad Prism in order to determine competing ligand IC_{50} and subsequently K_d^I.

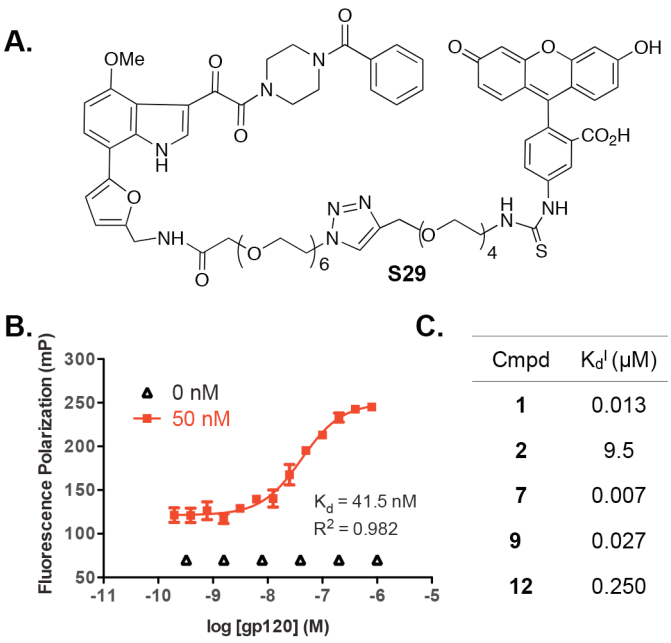


Figure S9. Fluorescence polarization (FP) was used to determine the binding affinities (KdI) of compounds to soluble gp120. (A) FITC conjugated to the gp120-binding terminus afforded ARM-H-FITC (**9**), which possessed a gp120 binding affinity (B) of 42 nM. (C) The binding affinities of unlabeled analogs were determined in subsequent competition experiments using **S29** as a fluorescent tracer reporter.

Complement Dependent Cytotoxicity (CDC)

Cell Culture

JRFL-TOE (Tet-off/TRE regulated expression of HIV-1 JRFL, HEK-293T, '+ gp120')[43] growth medium: high glucose DMEM + 10% FBS + 200 ug/mL hygromycin B + 50 ug/mL G418
HEK 293T/17 (ATCC # CRL-11268, '- gp120') growth medium: high glucose DMEM + 10% FBS + 1% P/S

Procedure

Assay was conducted in a similar fashion as previously reported.[2] Briefly, JRFL-TOE (+gp120) or 293T (-gp120) cells (100,000 cells/mL) were incubated with varying concentrations of ARM-H analogs or corresponding azide precursor in presence/absence of 10% v/v rabbit complement serum (Sigma) and 50 ug/mL rat anti-DNP Ab (Invitrogen) in corresponding growth media (100 uL total volume) at 37 °C for 3 hours. Cell viability was determined using the luciferase-based CellTiter-Glo Luminescent Cell Viability Assay (Promega). Max killing luminescent signal was determined 5% H_2O_2. Complement mediated cell death and cytotoxicity was calculated as: [1-((sample-max killing)/(untreated-max killing))] X 100 and plotted using GraphPad Prism. Assay repeated three times in triplicate and representative data plotted as mean (±SE) from triplicate experiment. Raw data was subjected to Dixon Q-test analysis at the 90% confidence interval and statistical outliers were removed accordingly.

References

(1) Wang, T.; Zhang, Z.; Wallace, O. B.; Deshpande, M.; Fang, H.; Yang, Z.; Zadjura, L. M.; Tweedie, D. L.; Huang, S.; Zhao, F.; Ranadive, S.; Robinson, B. S.; Gong, Y. F.; Ricarrdi, K.; Spicer, T. P.; Deminie, C.; Rose, R.; Wang, H. G.; Blair, W. S.; Shi, P. Y.; Lin, P. F.; Colonno, R. J.; Meanwell, N. A. *J Med Chem* **2003**, *46*, 4236.

(2) Parker, C. G.; Domaoal, R. A.; Anderson, K. S.; Spiegel, D. A. *J Am Chem Soc* **2009**, *131*, 16392.

(3) Falkler, W. A., Jr.; Lai, R.; Vincent, J. W.; Dober, L.; Spiegel, C.; Hayduk, S. *J Periodontol* **1982**, *53*, 762.

(4) Kwong, P. D.; Wyatt, R.; Majeed, S.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. *Structure* **2000**, *8*, 1329.

(5) ; 9.2 ed.; Schrodinger, LLC: New York, NY, 2011.

(6) ; Schrödinger, LLC: New York, NY, 2009.

(7) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J Med Chem* **2004**, *47*, 1739.

(8) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J Med Chem* **2006**, *49*, 6177.

(9) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J Med Chem* **2006**, *49*, 534.

(10) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J Am Chem Soc* **1996**, *118*, 11225.

(11) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J Comput Aided Mol Des* **1995**, *9*, 87.

(12) Damm, W.; Frontera, A.; TiradoRives, J.; Jorgensen, W. L. *J Comput Chem* **1997**, *18*, 1955.

(13) Humphrey, W.; Dalke, A.; Schulten, K. *J Mol Graph Model* **1996**, *14*, 33.

(14) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926.

(15) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J Comput Chem* **2005**, *26*, 1781.

(16) Jorgensen, W. L.; Tirado-Rives, J. *Proc Natl Acad Sci U S A* **2005**, *102*, 6665.

(17) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993**, *98*, 10089.

(18) Brunger, A.; Brooks, C. L.; Karplus, M. *Chem Phys Lett* **1984**, *105*, 495.

(19) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *Journal of Chemical Physics* **1994**, *101*, 4177.

(20) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *Journal of Chemical Physics* **1995**, *103*, 4613.

(21) Lepine, S.; Allegood, J. C.; Park, M.; Dent, P.; Milstien, S.; Spiegel, S. *Cell Death Differ* **2011**, *18*, 350.

(22) Irfan, F. B.; Irfan, B. B.; Spiegel, D. A. *J Surg Res* **2012**, *176*, 84.

(23) Eshhar, Z.; Ofarim, M.; Waks, T. *J Immunol* **1980**, *124*, 775.

(24) Tawfik, D. S.; Chap, R.; Eshhar, Z.; Green, B. S. *Protein Eng* **1994**, *7*, 431.

(25) Zhou, T.; Georgiev, I.; Wu, X.; Yang, Z. Y.; Dai, K.; Finzi, A.; Kwon, Y. D.; Scheid, J. F.; Shi, W.; Xu, L.; Yang, Y.; Zhu, J.; Nussenzweig, M. C.; Sodroski, J.; Shapiro, L.; Nabel, G. J.; Mascola, J. R.; Kwong, P. D. *Science* **2010**, *329*, 811.

(26) Maceyka, M.; Harikumar, K. B.; Milstien, S.; Spiegel, S. *Trends Cell Biol* **2012**, *22*, 50.

(27) Studer, E.; Zhou, X.; Zhao, R.; Wang, Y.; Takabe, K.; Nagahashi, M.; Pandak, W. M.; Dent, P.; Spiegel, S.; Shi, R.; Xu, W.; Liu, X.; Bohdan, P.; Zhang, L.; Zhou, H.; Hylemon, P. B. *Hepatology* **2012**, *55*, 267.

(28) Hanna, G. J.; Lalezari, J.; Hellinger, J. A.; Wohl, D. A.; Nettles, R.; Persson, A.; Krystal, M.; Lin, P.; Colonno, R.; Grasela, D. M. *Antimicrob Agents Chemother* **2011**, *55*, 722.

(29) Nettles, R. E.; Schurmann, D.; Zhu, L.; Stonier, M.; Huang, S. P.; Chang, I.; Chien, C.; Krystal, M.; Wind-Rotolo, M.; Ray, N.; Hanna, G. J.; Bertz, R.; Grasela, D. *The Journal of infectious diseases* **2012**, *206*, 1002.

(30) Zhou, N.; Nowicka-Sans, B.; McAuliffe, B.; Ray, N.; Eggers, B.; Fang, H.; Fan, L.; Healy, M.; Langley, D. R.; Hwang, C.; Lataillade, M.; Hanna, G. J.; Krystal, M. *The Journal of antimicrobial chemotherapy* **2013**.

(31) Douglass, E. F.; Miller, C. J.; Sparer, G.; Shapiro, H.; Spiegel, D. A. *J Am Chem Soc* **2013**, *135*, 6092.

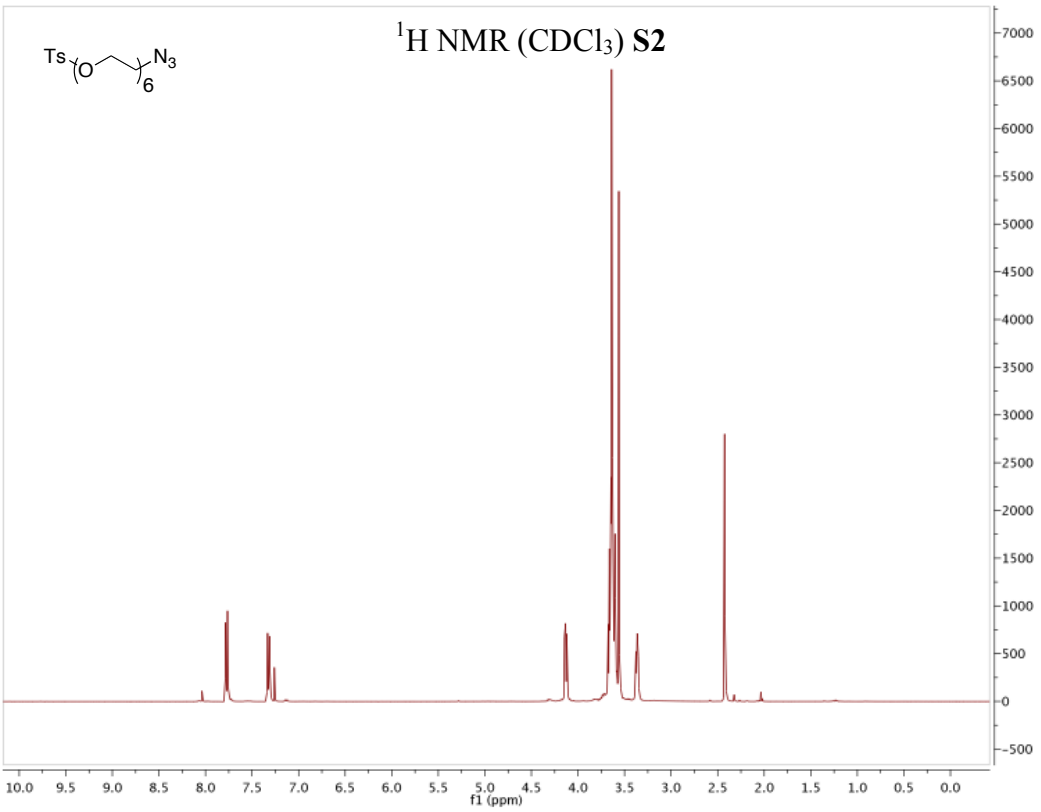
(32) Kubetzko, S.; Sarkar, C. A.; Pluckthun, A. *Mol Pharmacol* **2005**, *68*, 1439.

(33) Zhu, P.; Liu, J.; Bess, J.; Chertova, E.; Lifson, J. D.; Grisé, H.; Ofek, G. A.; Taylor, K. A.; Roux, K. H. *Nature* **2006**, *441*, 847.

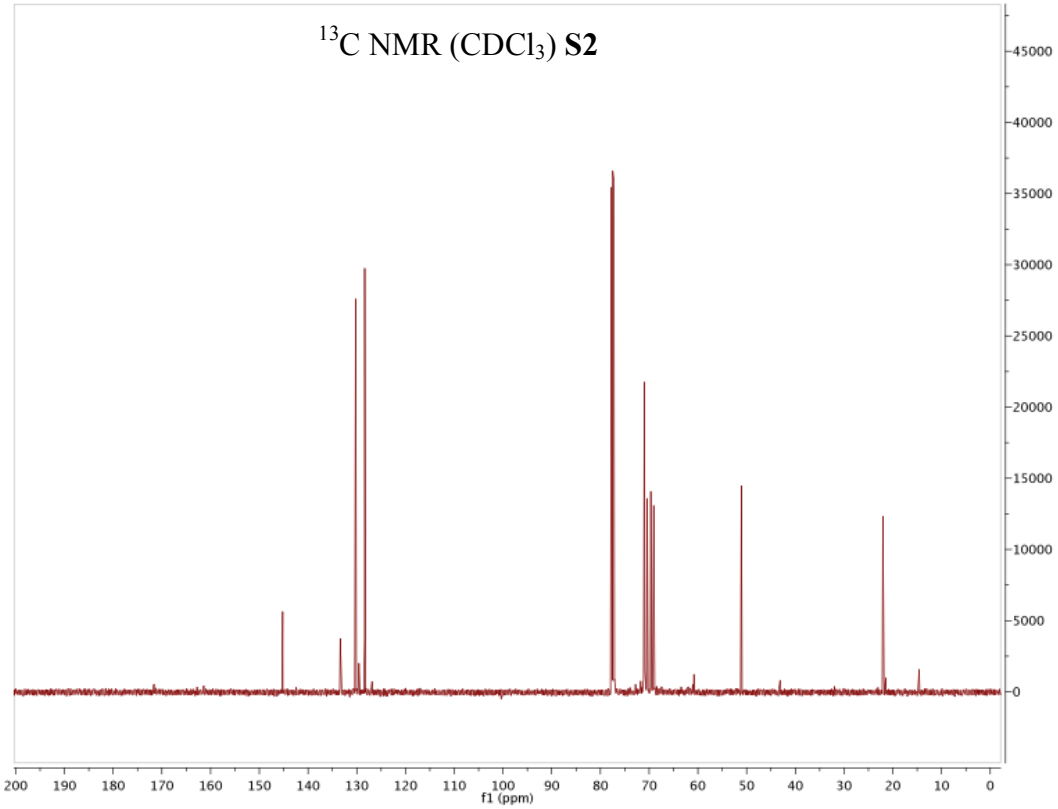
(34) Kwong, P. D.; Wyatt, R.; Sattentau, Q. J.; Sodroski, J.; Hendrickson, W. A. *Journal of virology* **2000**, *74*, 1961.
(35) Klein, J. S.; Bjorkman, P. J. *PLoS Pathog* **2010**, *6*, e1000908.
(36) Klein, J. S.; Gnanapragasam, P. N. P.; Galimidi, R. P.; Foglesong, C. P.; West, A. P.; Bjorkman, P. J. *Proceedings of the National Academy of Sciences* **2009**, *106*, 7385.
(37) Bennett, A.; Liu, J.; Van Ryk, D.; Bliss, D.; Arthos, J.; Henderson, R. M.; Subramaniam, S. *Journal of Biological Chemistry* **2007**, *282*, 27754.
(38) Li, H.; Song, H.; Heredia, A.; Le, N.; Redfield, R.; Lewis, G. K.; Wang, L. X. *Bioconjugate chemistry* **2004**, *15*, 783.
(39) Baird, E. J.; Holowka, D.; Coates, G. W.; Baird, B. *Biochemistry* **2003**, *42*, 12739.
(40) Jeppesen, C.; Wong, J. Y.; Kuhl, T. L.; Israelachvili, J. N.; Mullah, N.; Zalipsky, S.; Marques, C. M. *Science* **2001**, *293*, 465.
(41) Kloiber, K.; Weiskirchen, R.; Krautler, B.; Bister, K.; Konrat, R. *J Mol Biol* **1999**, *292*, 893.
(42) Huang, X. *J Biomol Screen* **2003**, *8*, 34.
(43) Lin, P. F.; Blair, W.; Wang, T.; Spicer, T.; Guo, Q.; Zhou, N. N.; Gong, Y. F.; Wang, H. G. H.; Rose, R.; Yamanaka, G.; Robinson, B.; Li, C. B.; Fridell, R.; Deminie, C.; Demers, G.; Yang, Z.; Zadjura, L.; Meanwell, N.; Colonno, R. *Proc Natl Acad Sci U S A* **2003**, *100*, 11013.

Spectra of Compounds

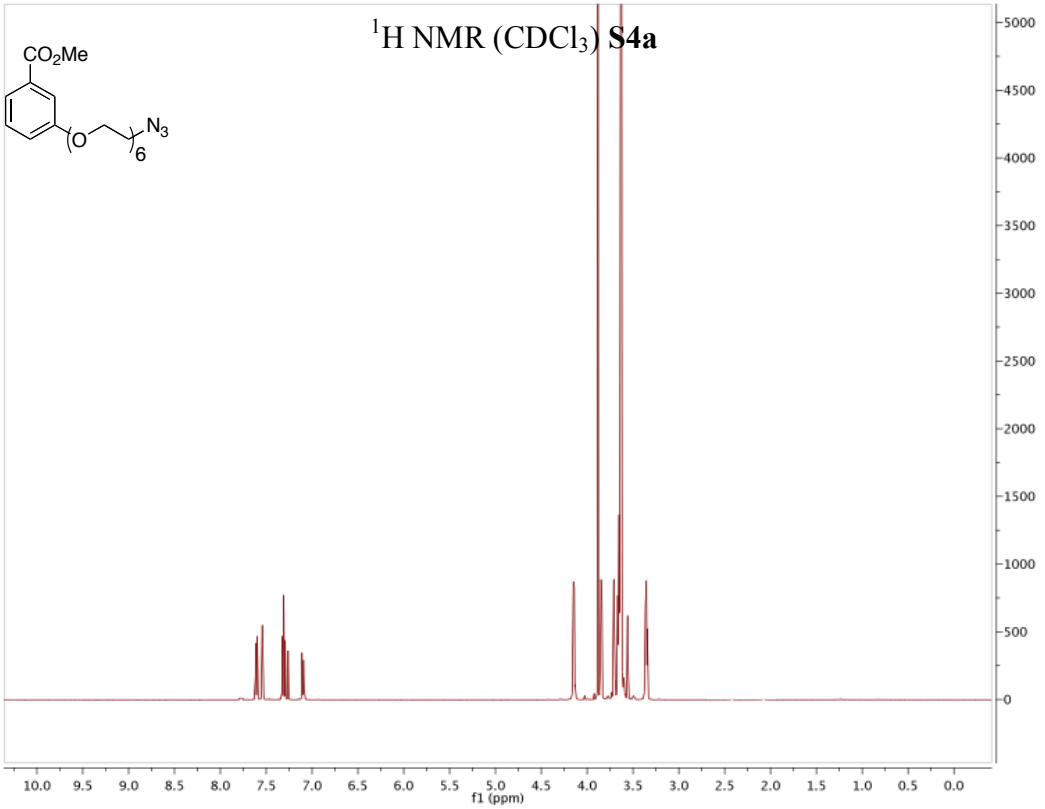
^{1}H NMR ($CDCl_3$) **S2**



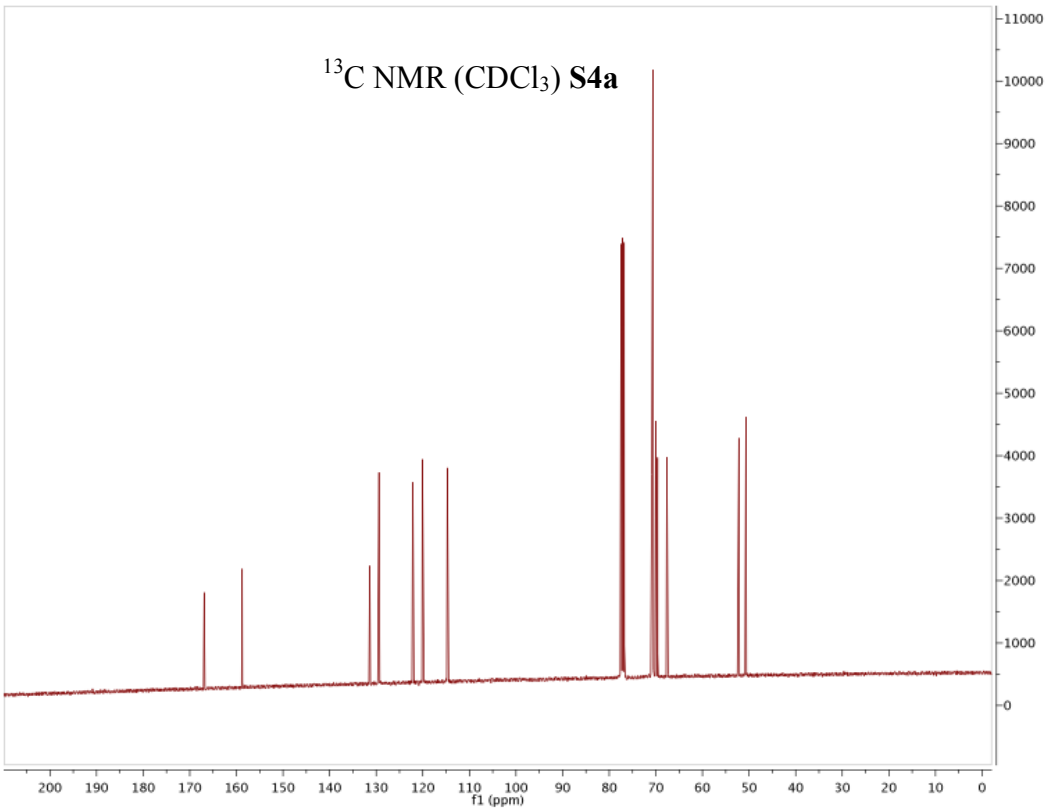
^{13}C NMR ($CDCl_3$) **S2**

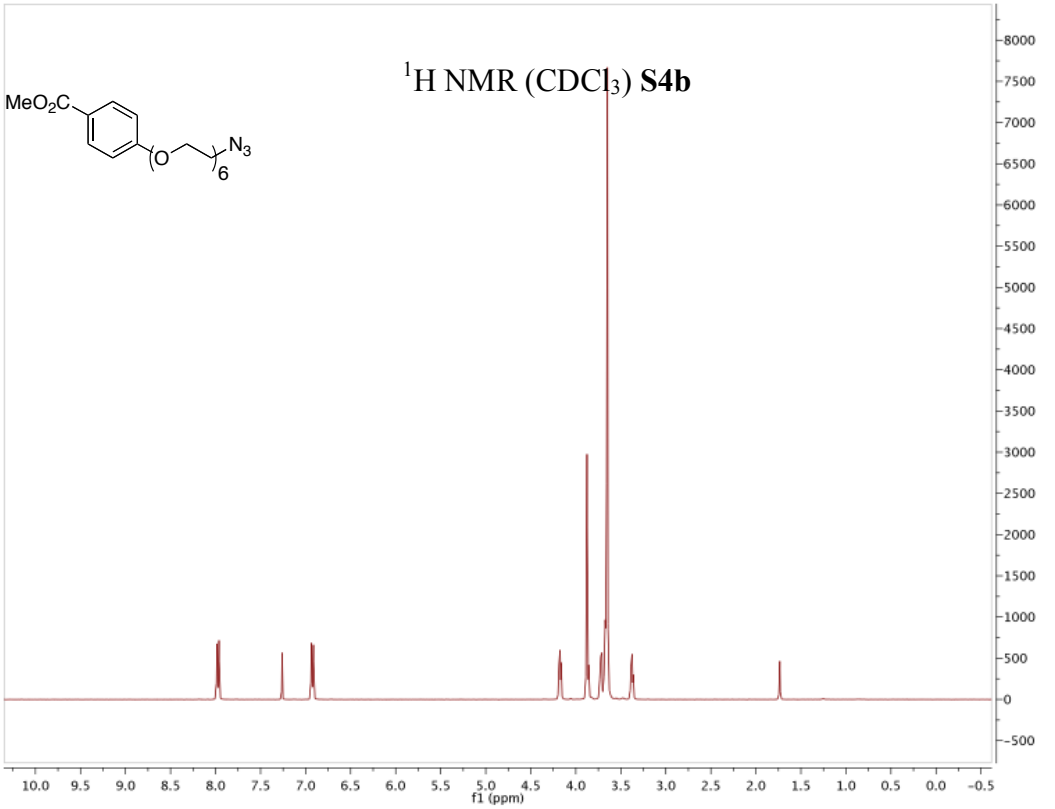


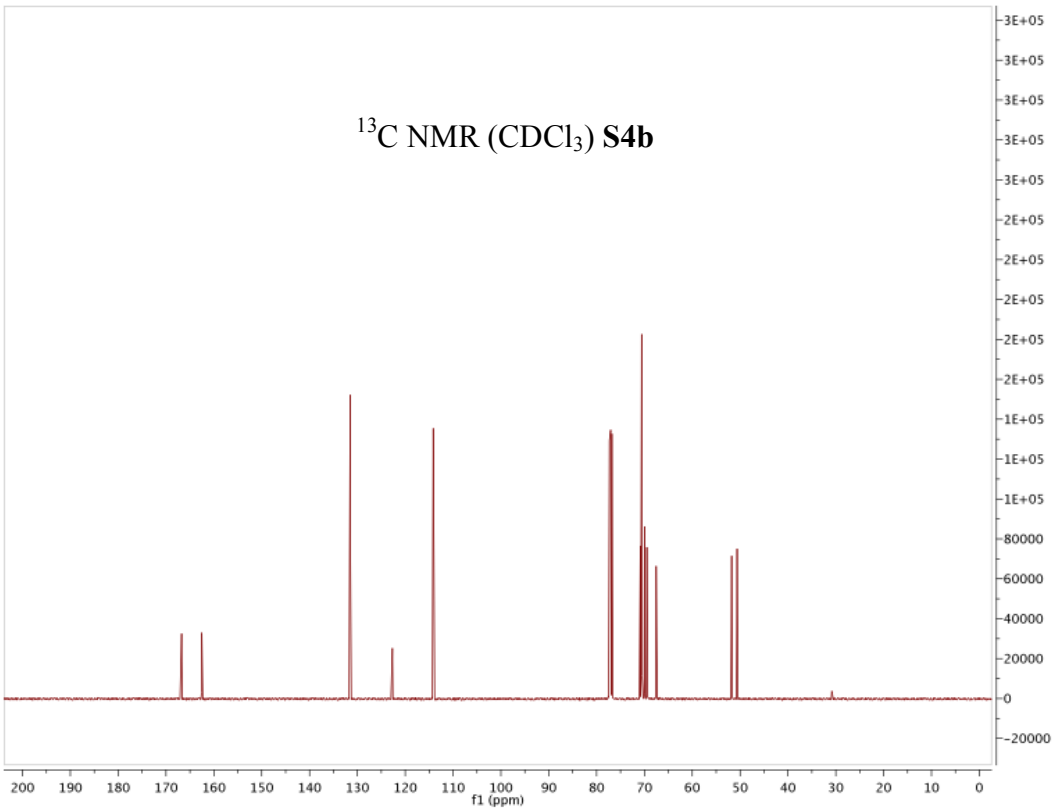
^{1}H NMR ($CDCl_3$) **S4a**



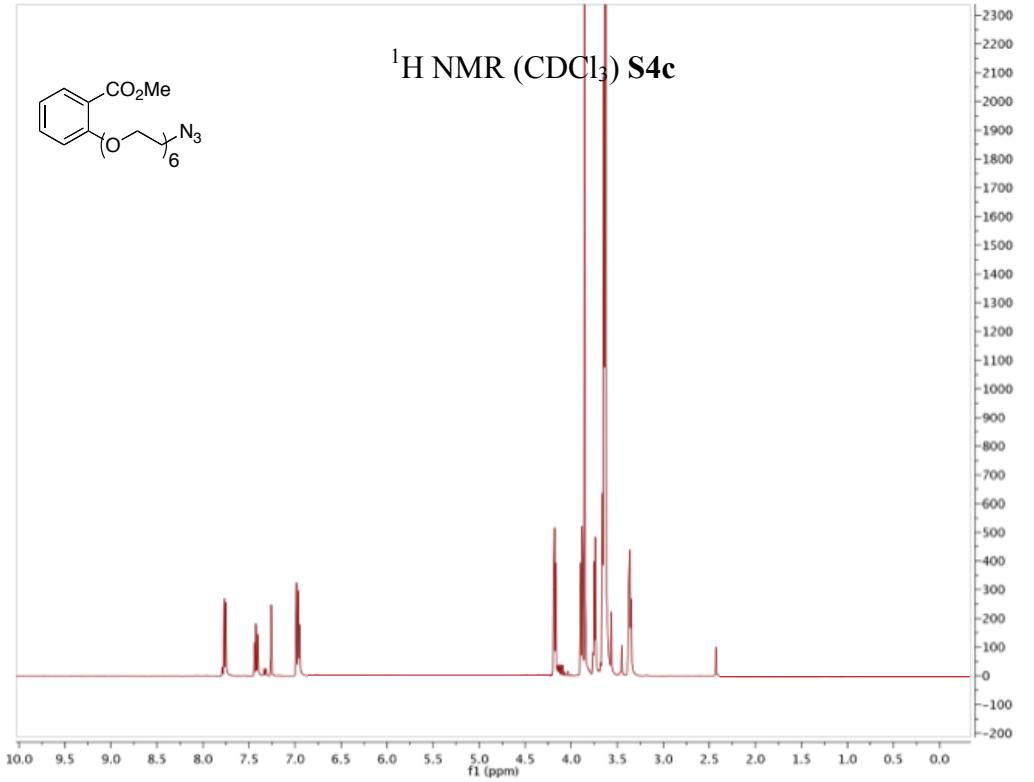
^{13}C NMR ($CDCl_3$) **S4a**



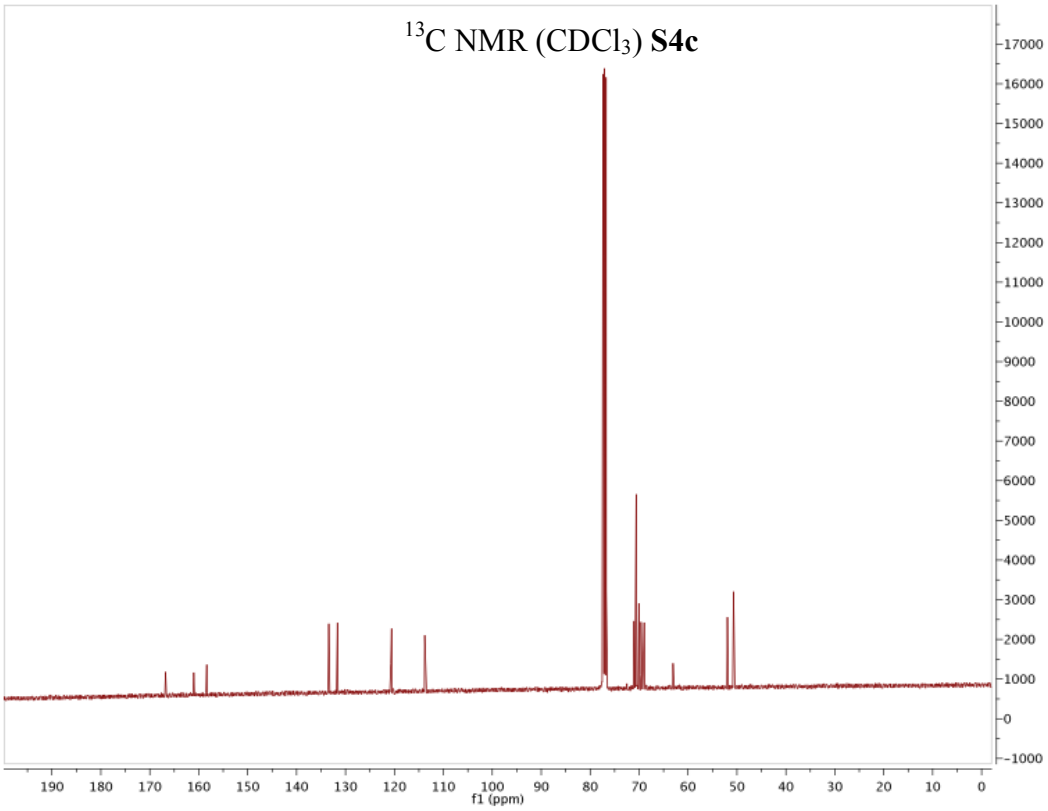

MeO2C
O
6
N3
1H NMR (CDCl3) S4b
f1 (ppm)


13C NMR (CDCl3) S4b
f1 (ppm)

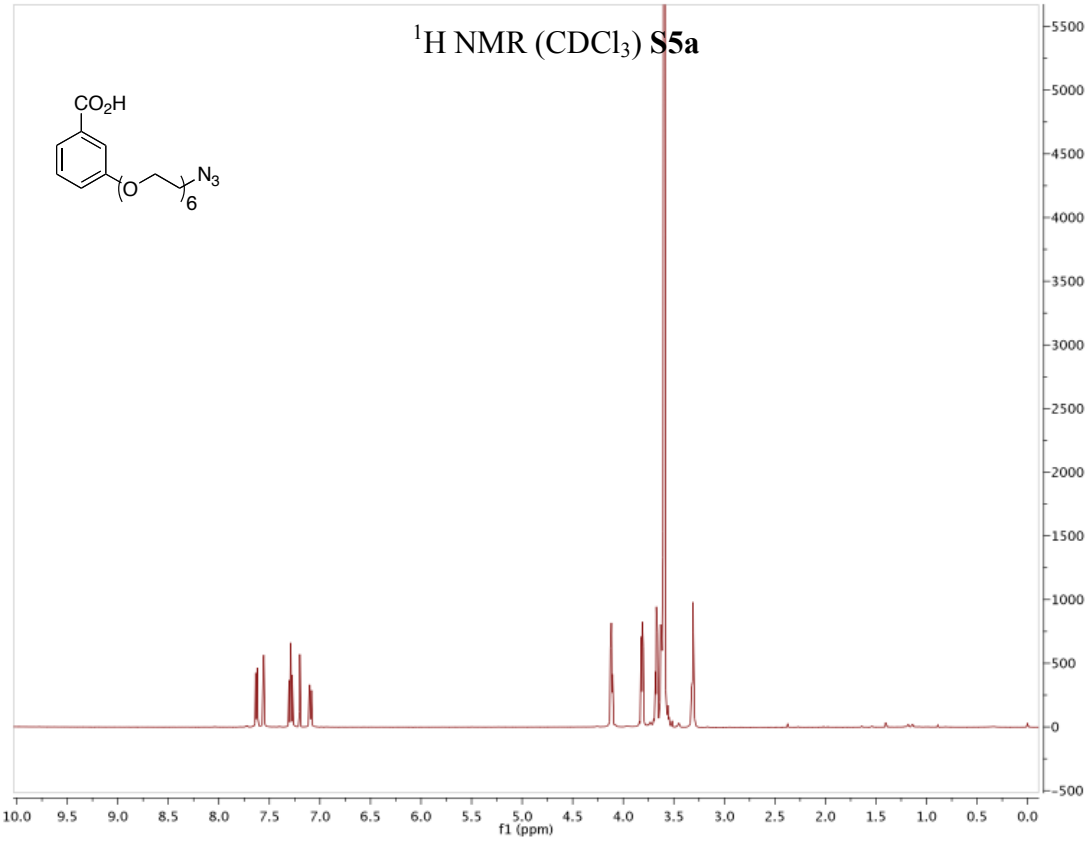
^{1}H NMR ($CDCl_3$) **S4c**



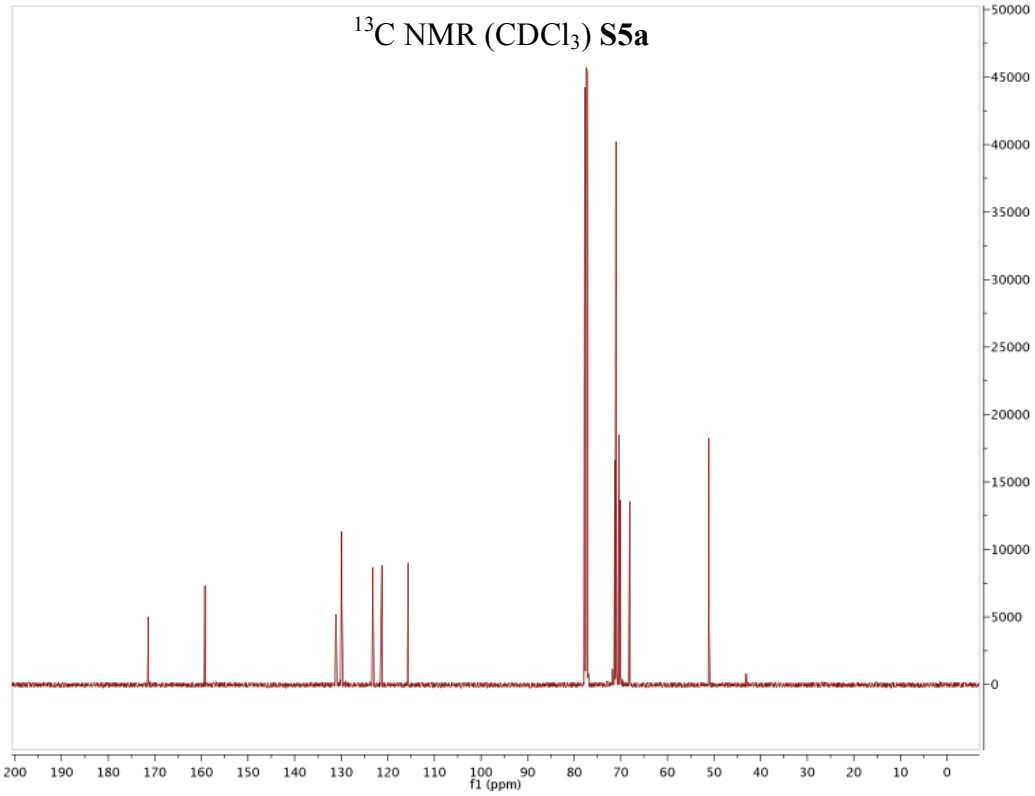
^{13}C NMR ($CDCl_3$) **S4c**



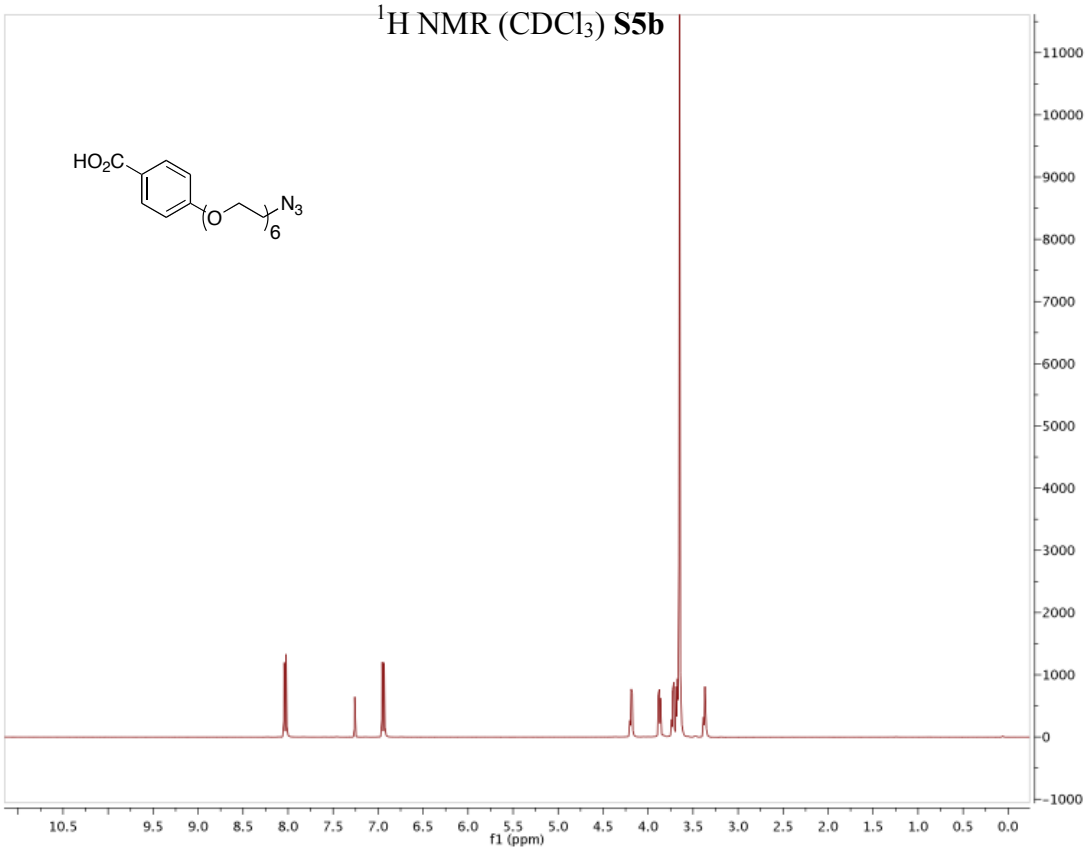
^{1}H NMR ($CDCl_3$) **S5a**



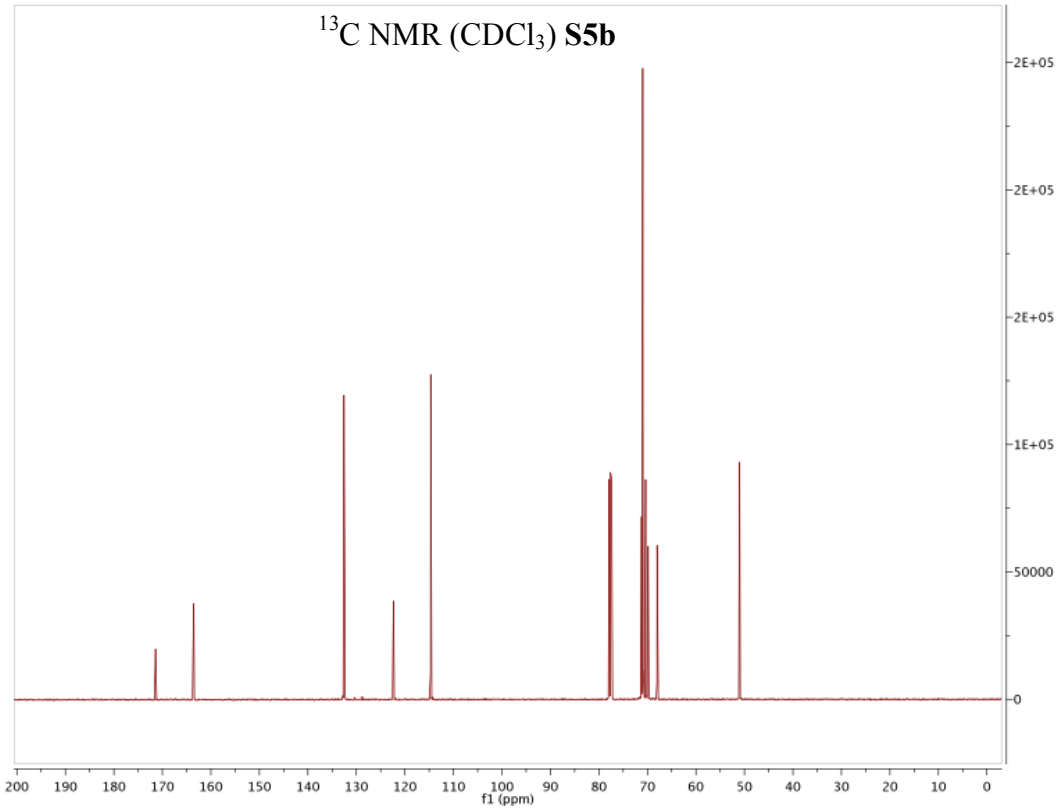
^{13}C NMR ($CDCl_3$) **S5a**



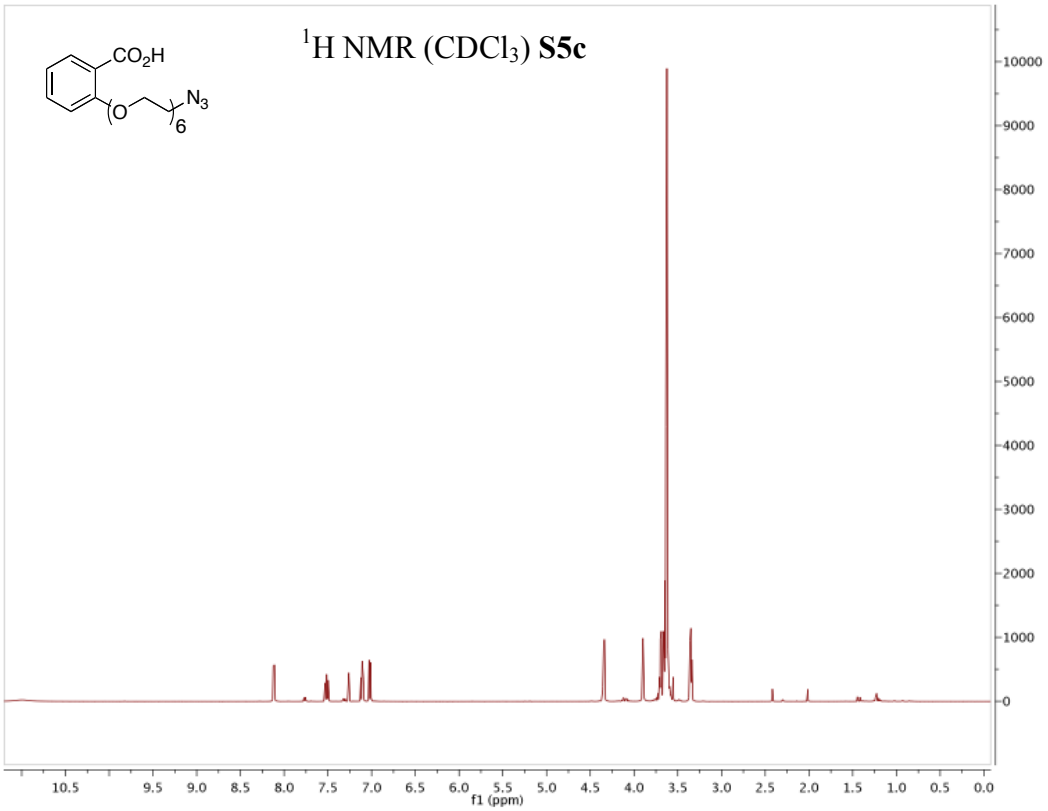
^{1}H NMR ($CDCl_3$) **S5b**



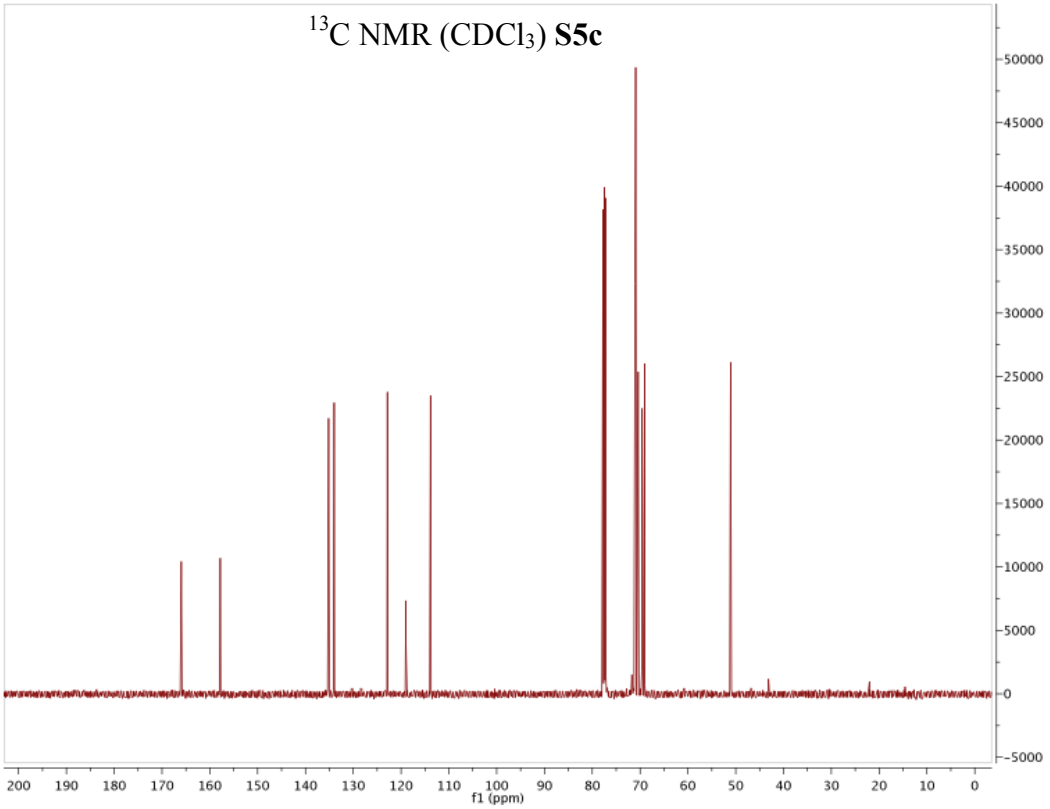
^{13}C NMR ($CDCl_3$) **S5b**



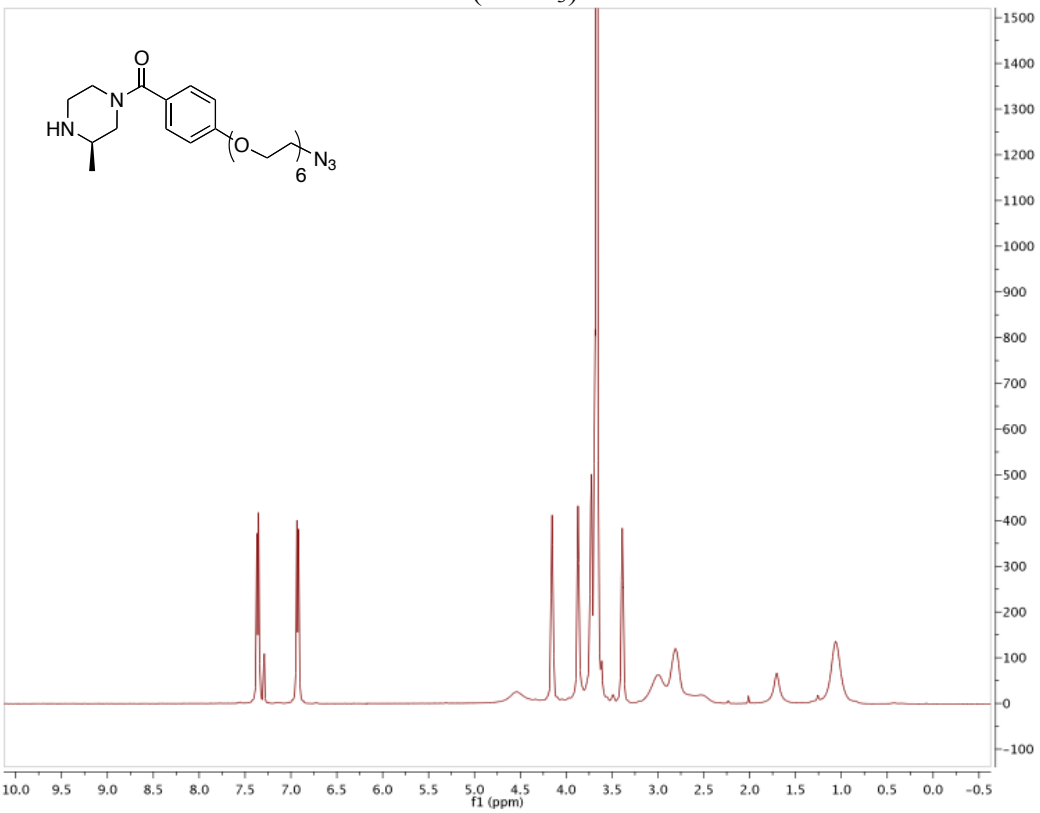
^{1}H NMR ($CDCl_3$) **S5c**



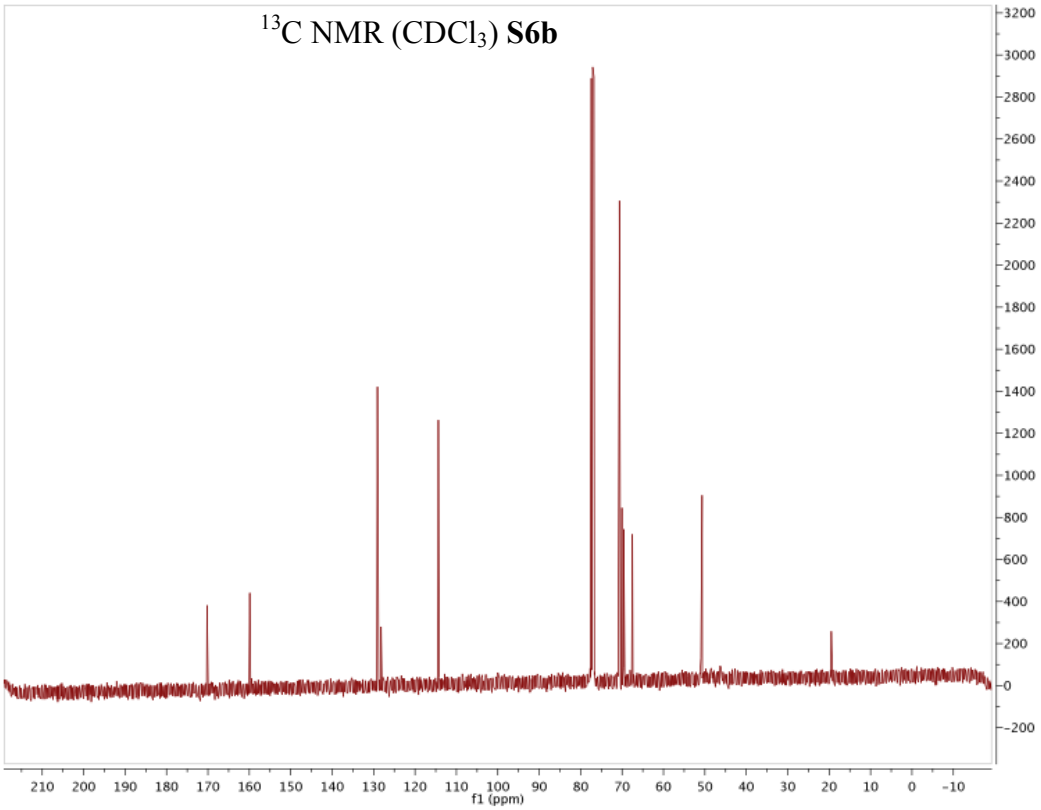
^{13}C NMR ($CDCl_3$) **S5c**



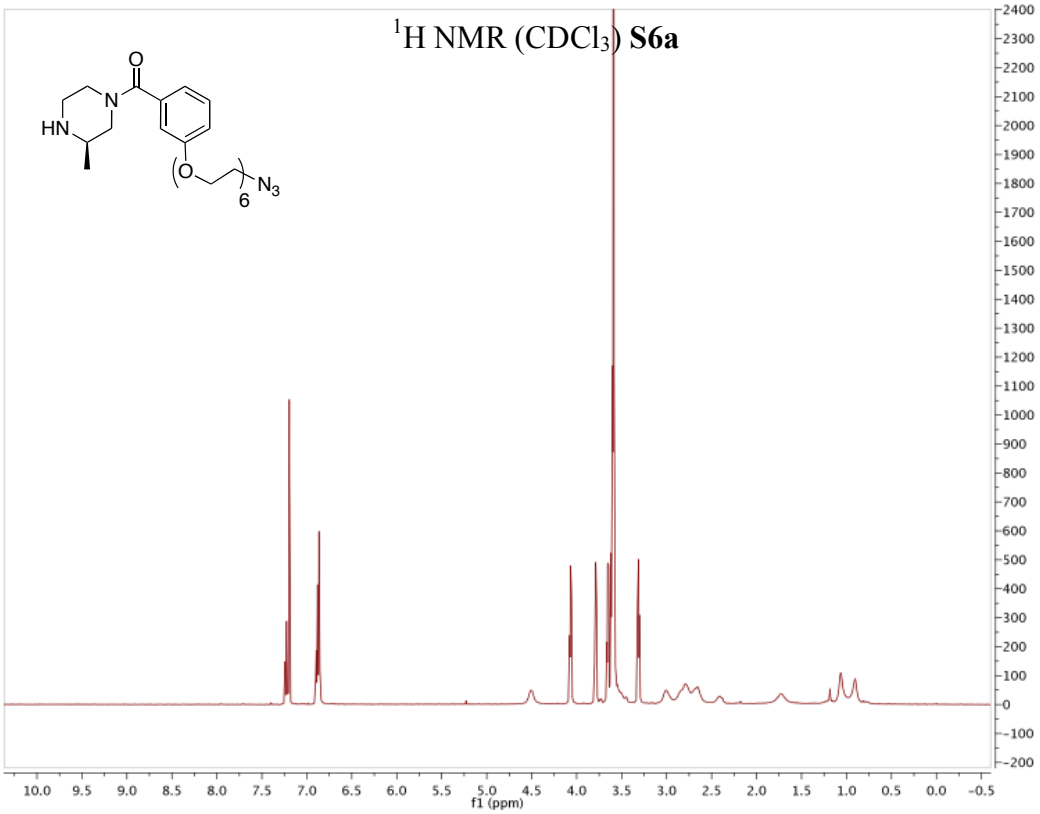
^{1}H NMR ($CDCl_3$) **S6b**



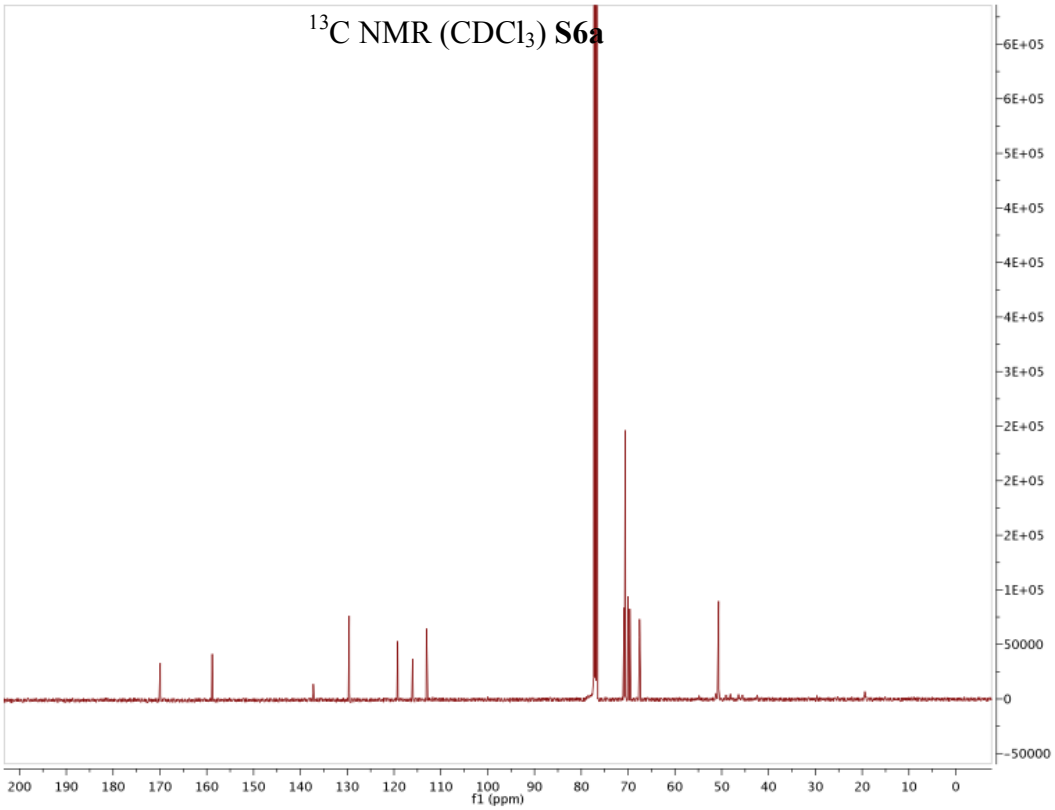
^{13}C NMR ($CDCl_3$) **S6b**

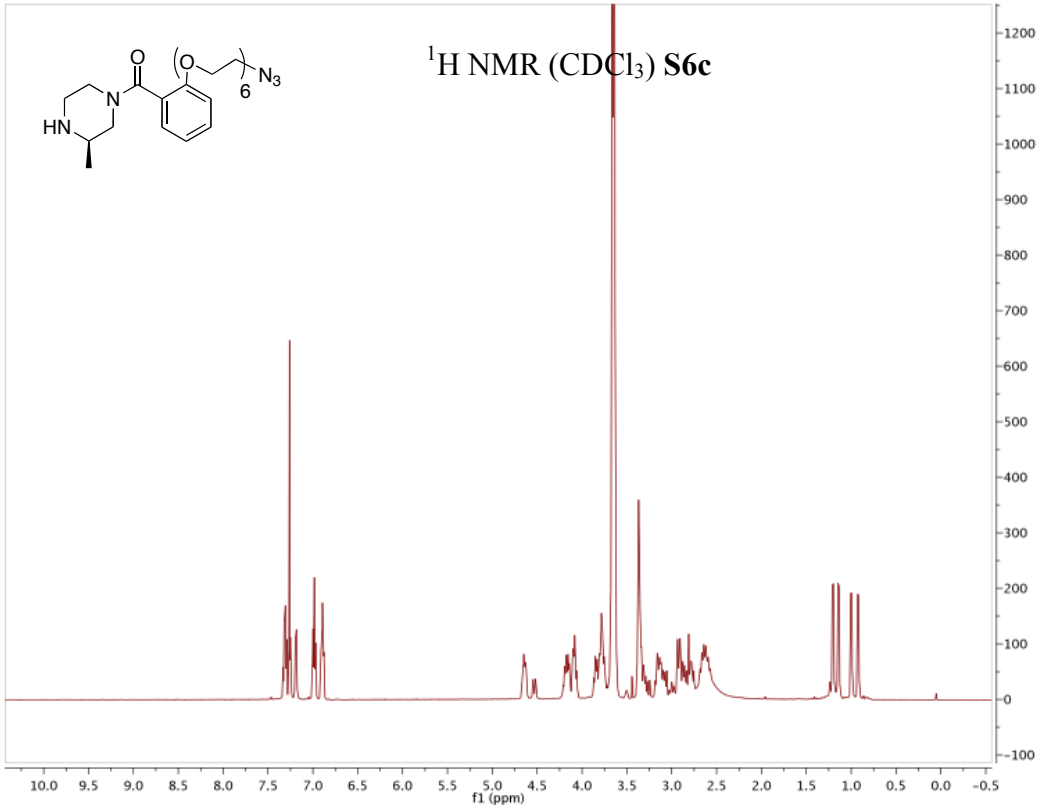


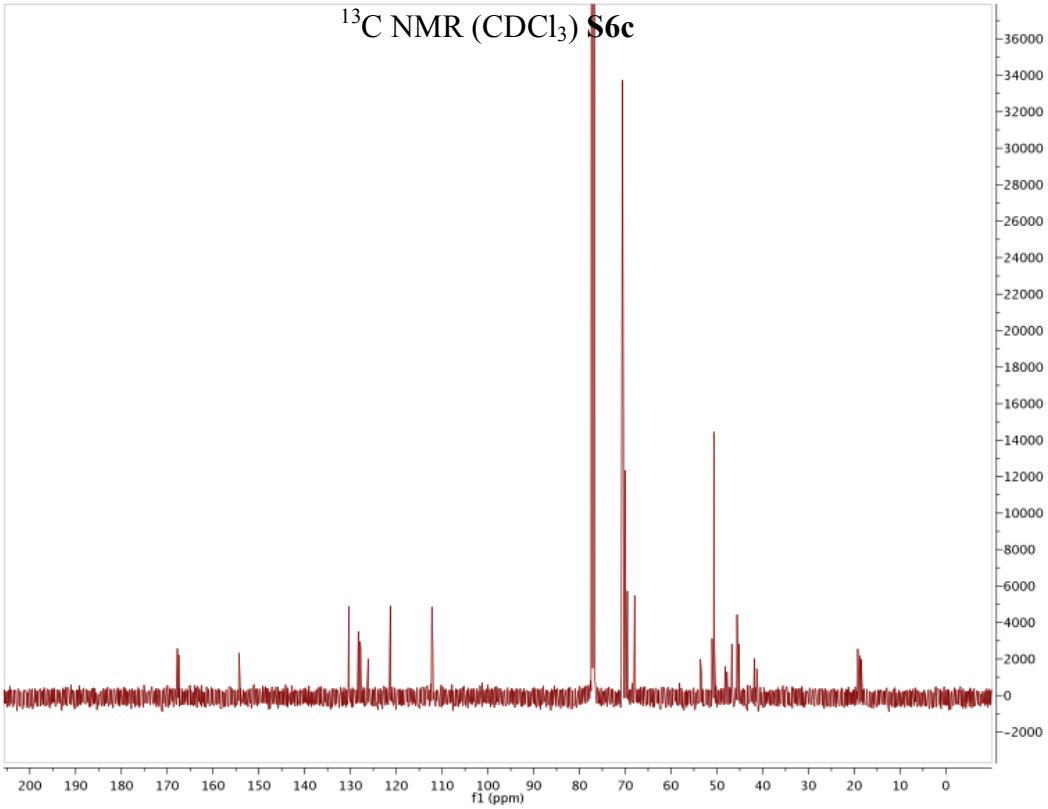
^{1}H NMR ($CDCl_3$) **S6a**



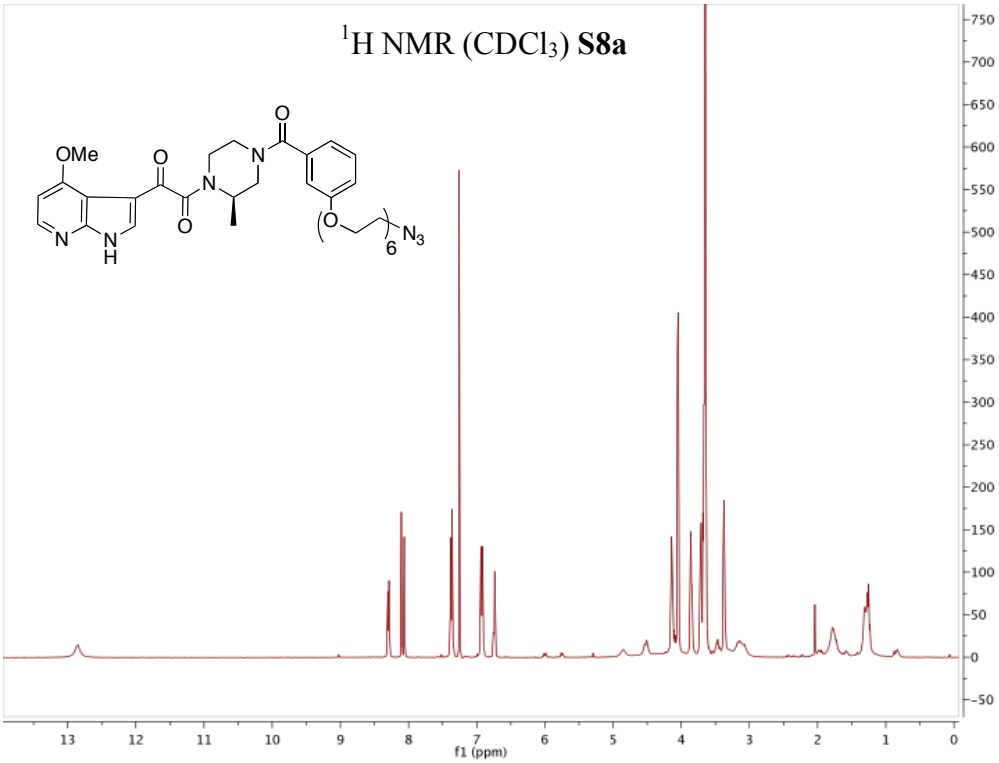
^{13}C NMR ($CDCl_3$) **S6a**



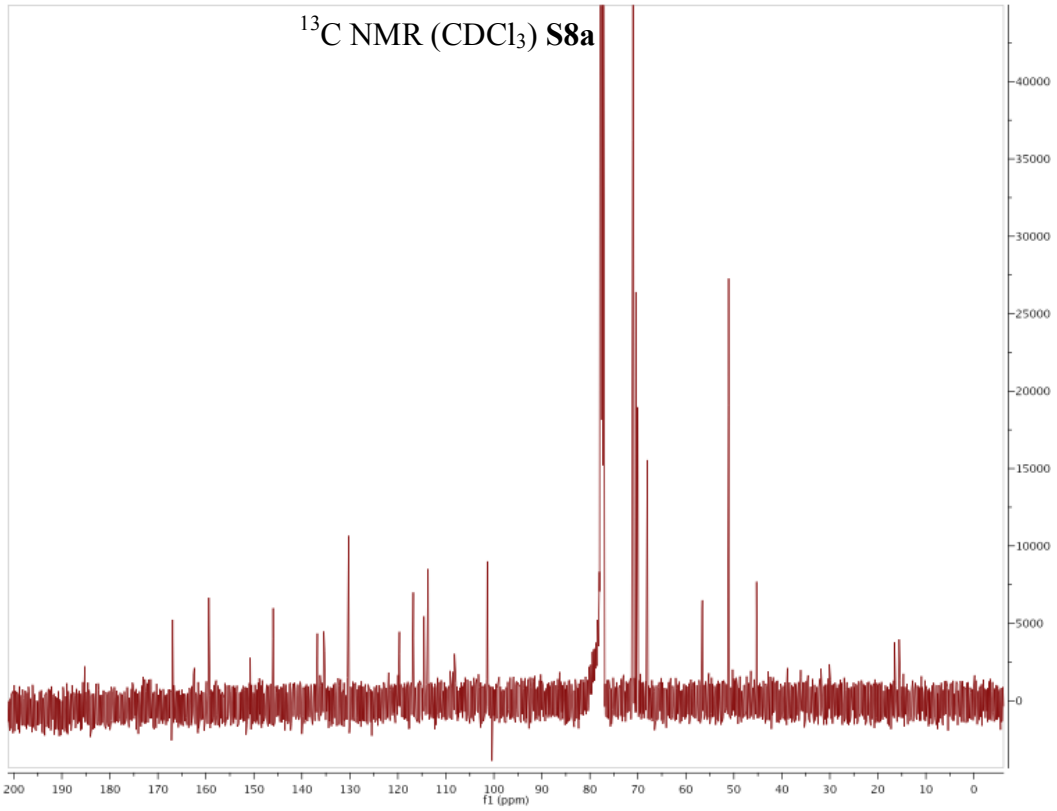

^{1}H NMR ($CDCl_3$) **S6c**
O
O
N
HN
6
N_3
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
-100
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)


^{13}C NMR ($CDCl_3$) **S6c**
36000
34000
32000
30000
28000
26000
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
-2000
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

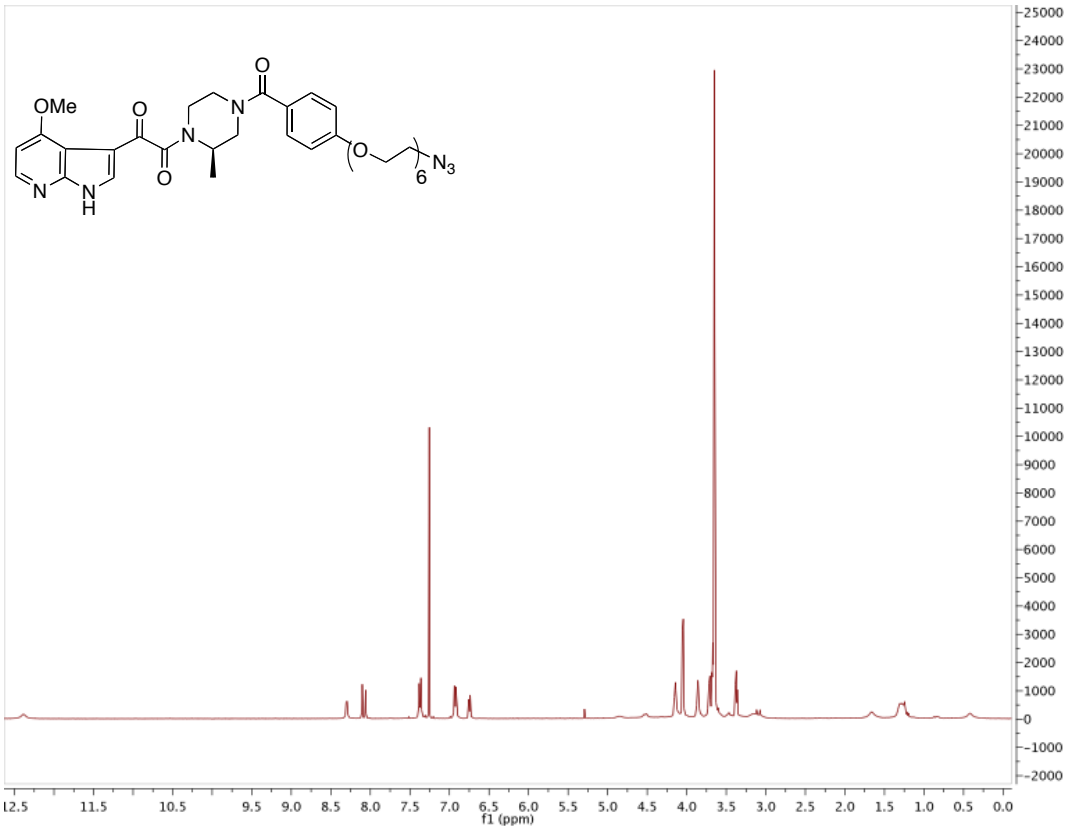
^{1}H NMR ($CDCl_3$) **S8a**



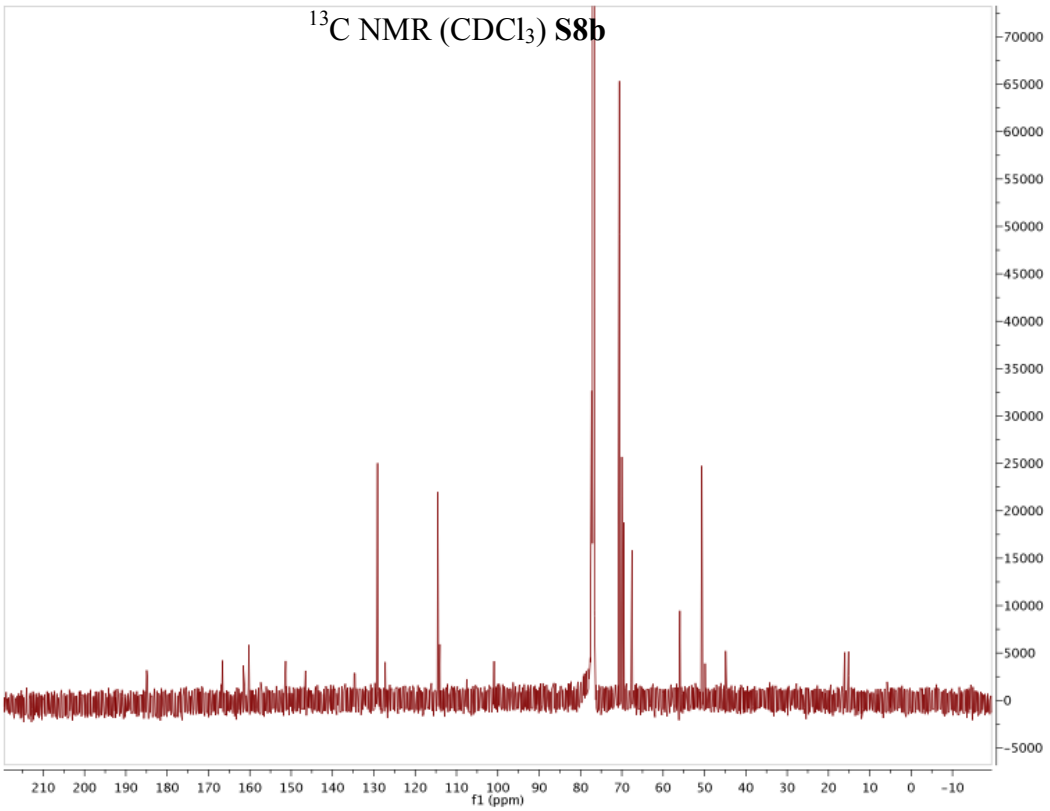
^{13}C NMR ($CDCl_3$) **S8a**



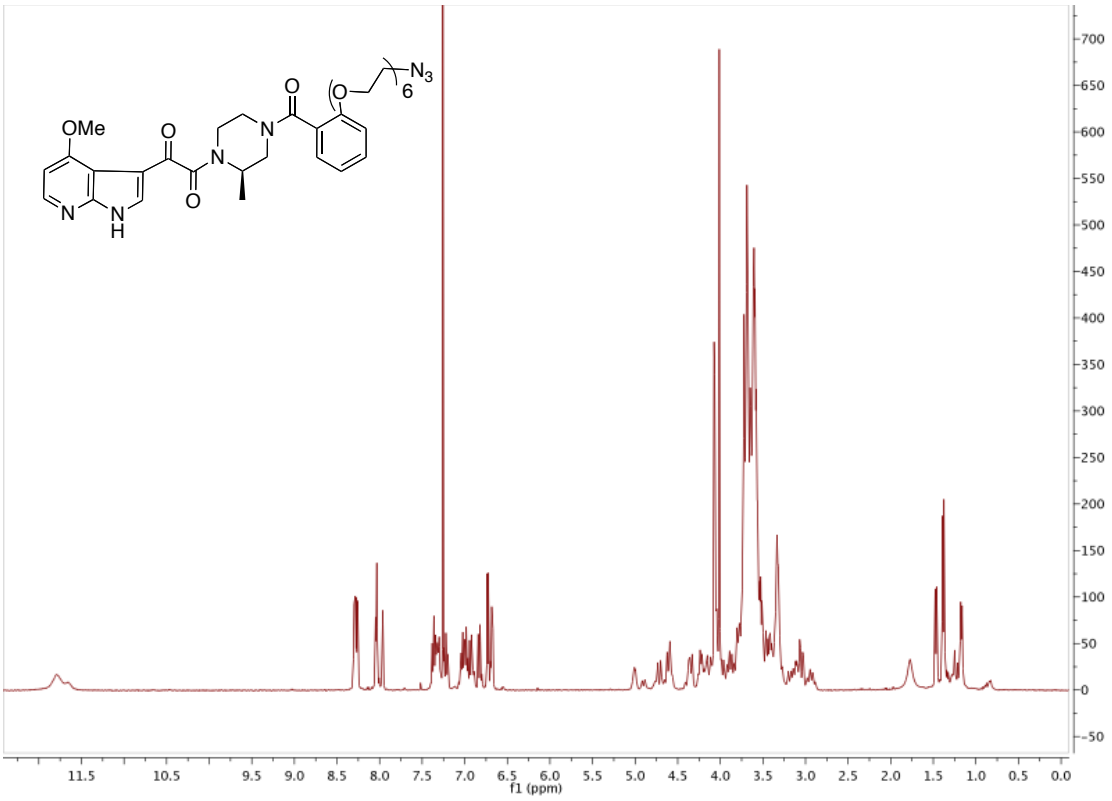
^{1}H NMR ($CDCl_3$) **S8b**



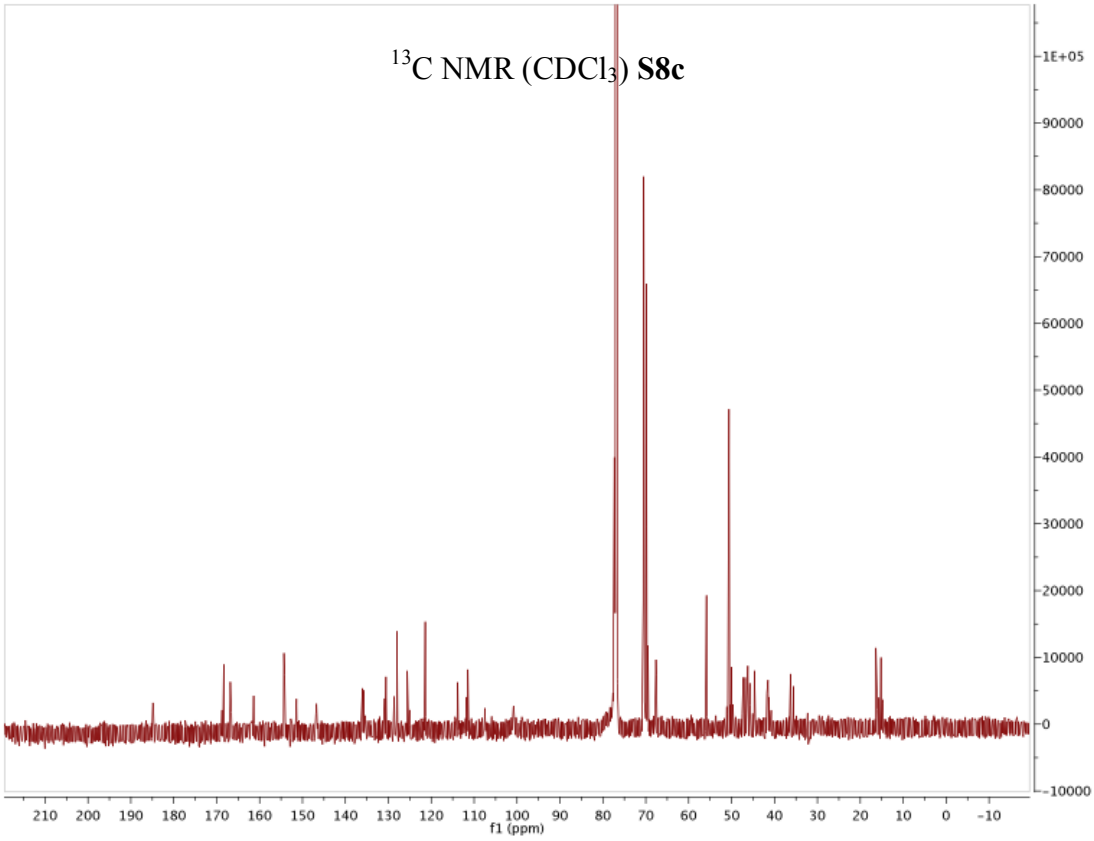
^{13}C NMR ($CDCl_3$) **S8b**



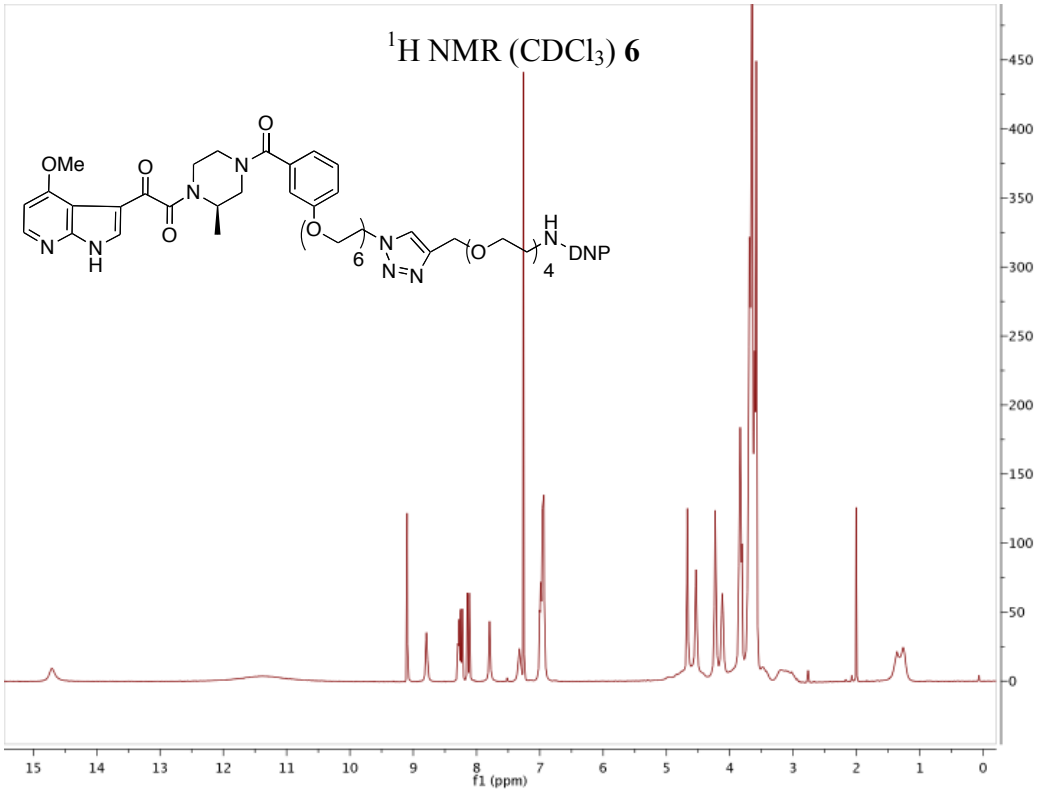
^{1}H NMR ($CDCl_3$) **S8c**



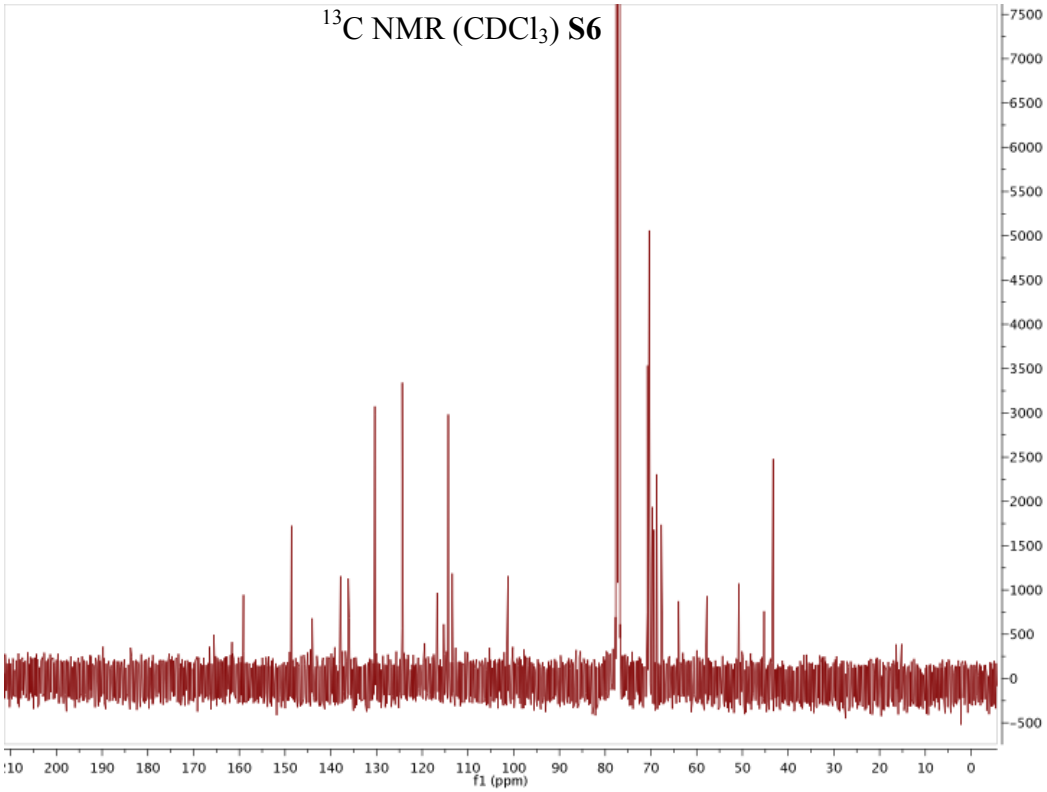
^{13}C NMR ($CDCl_3$) **S8c**



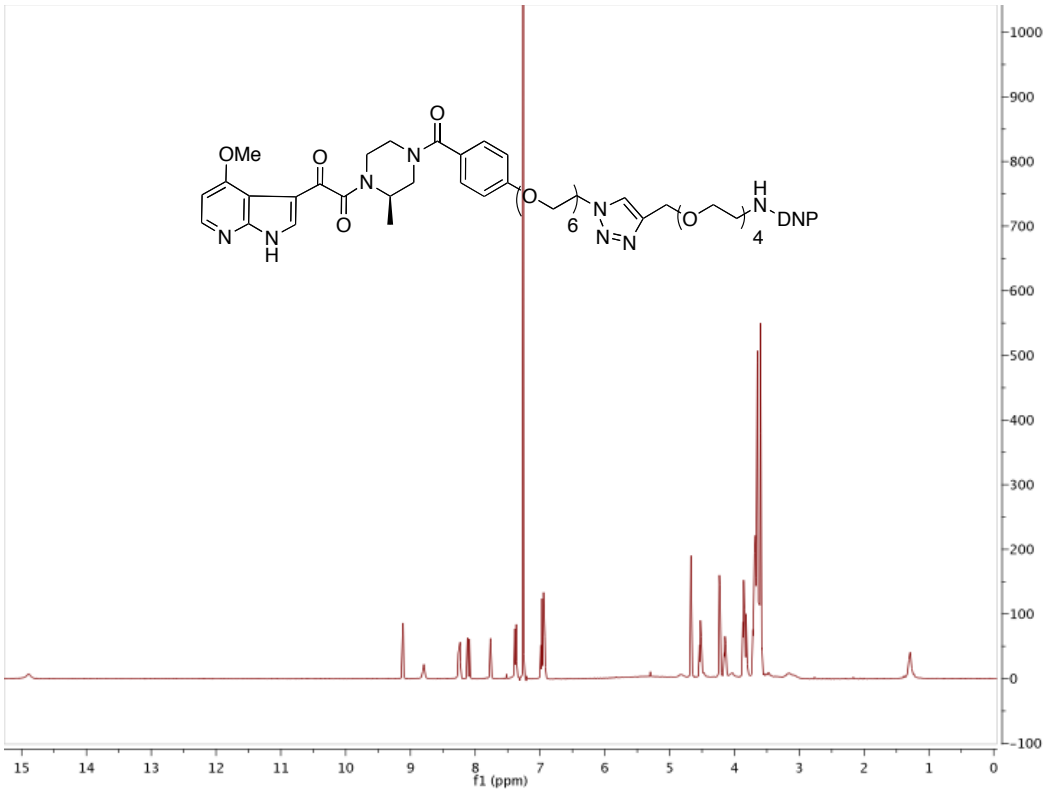
^{1}H NMR ($CDCl_3$) **6**



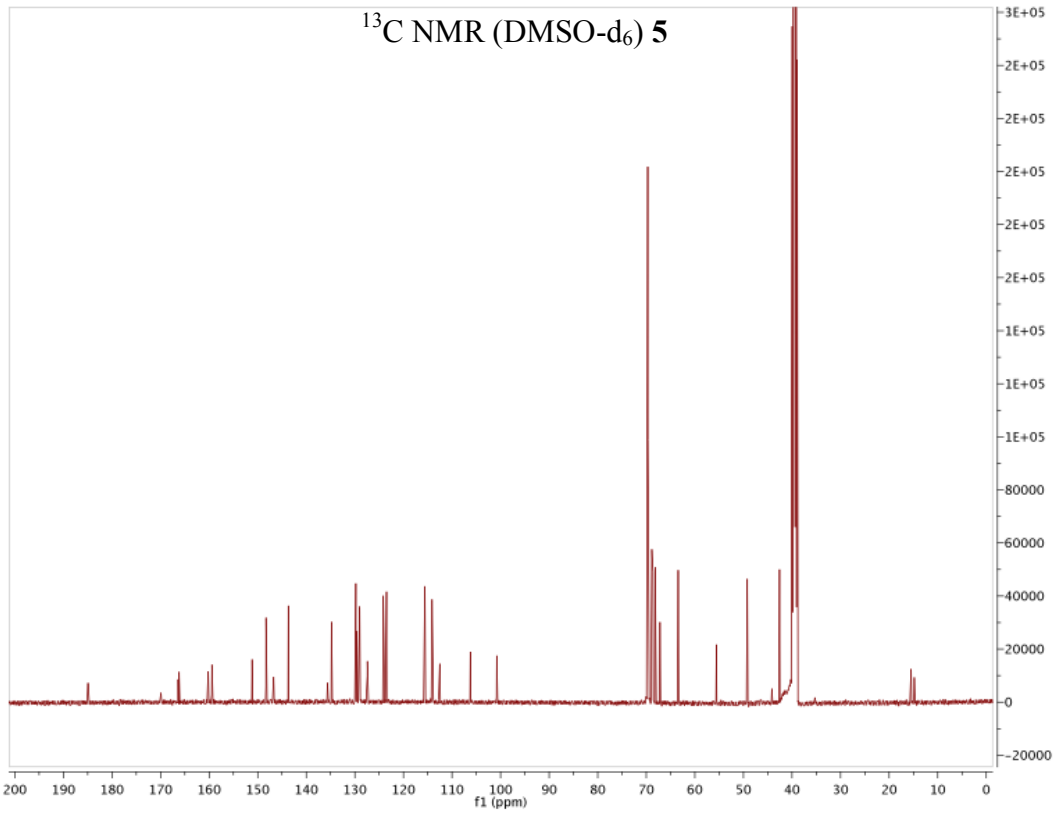
^{13}C NMR ($CDCl_3$) **S6**



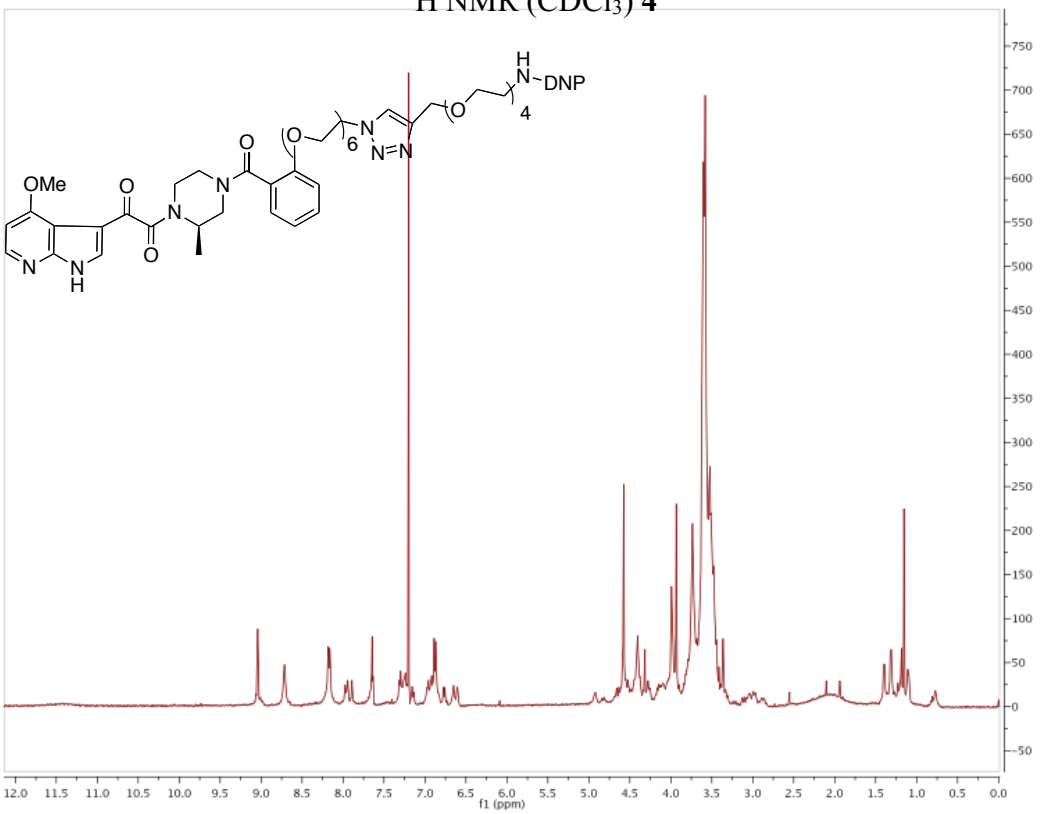
^{1}H NMR ($CDCl_3$) **5**



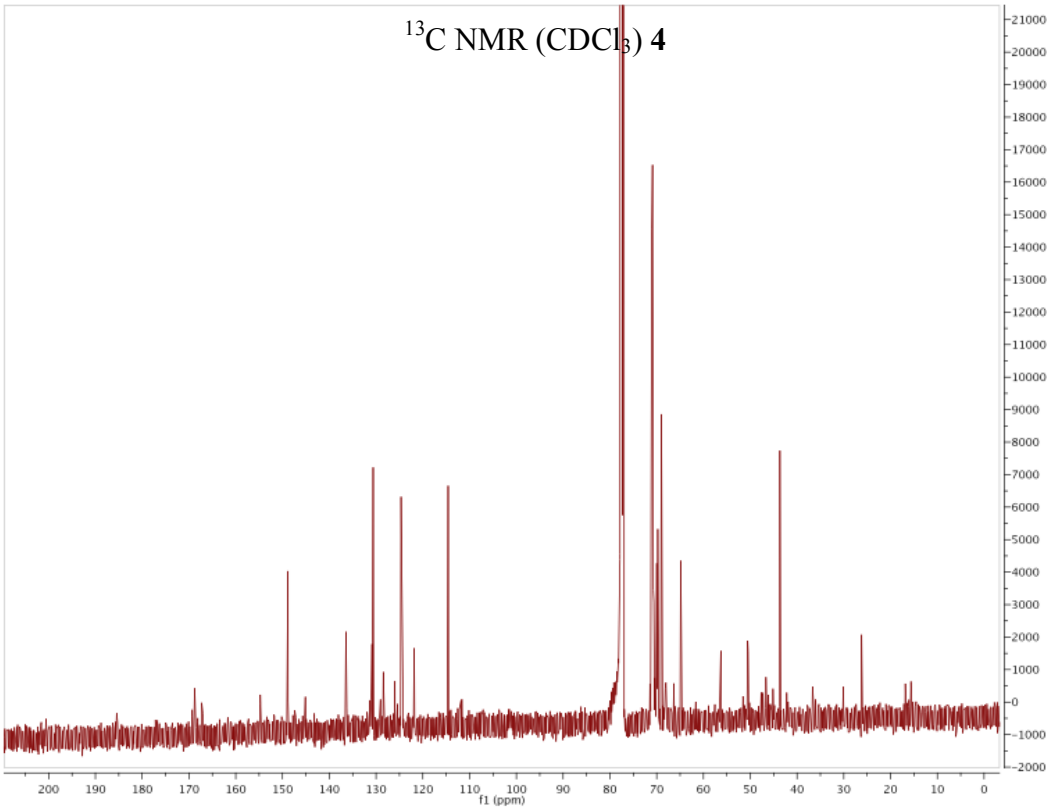
^{13}C NMR ($DMSO\text{-}d_6$) **5**



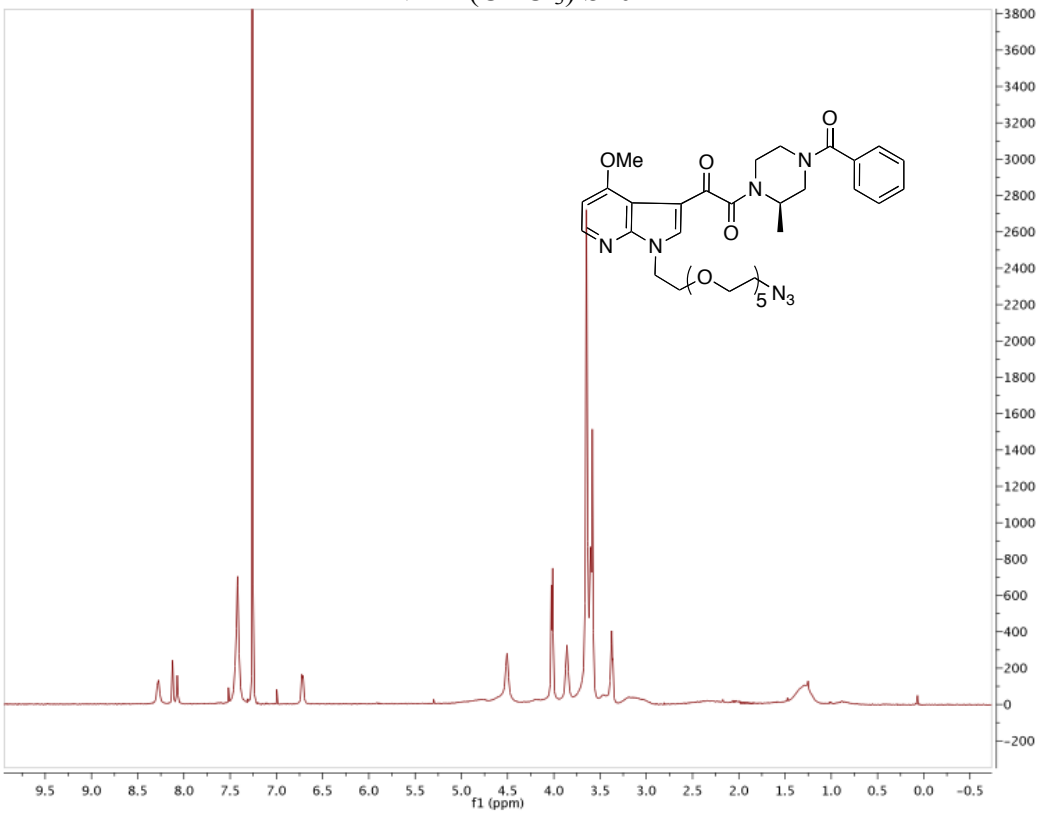
^{1}H NMR ($CDCl_3$) **4**



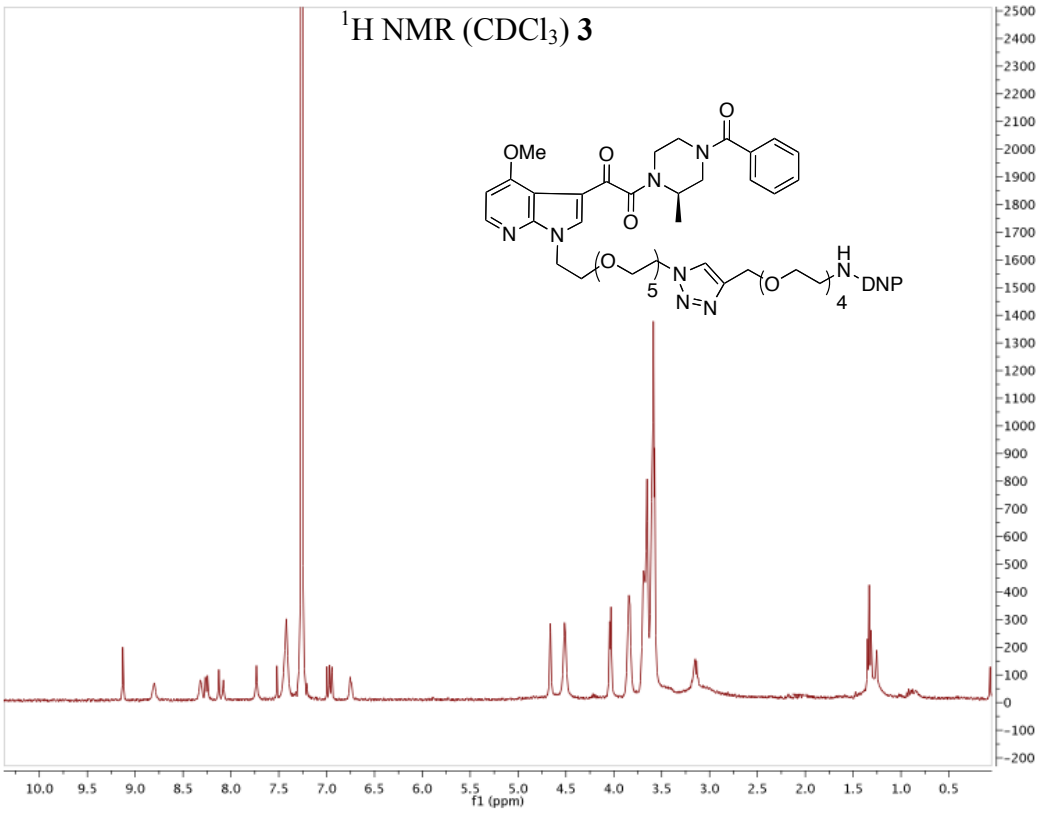
^{13}C NMR ($CDCl_3$) **4**



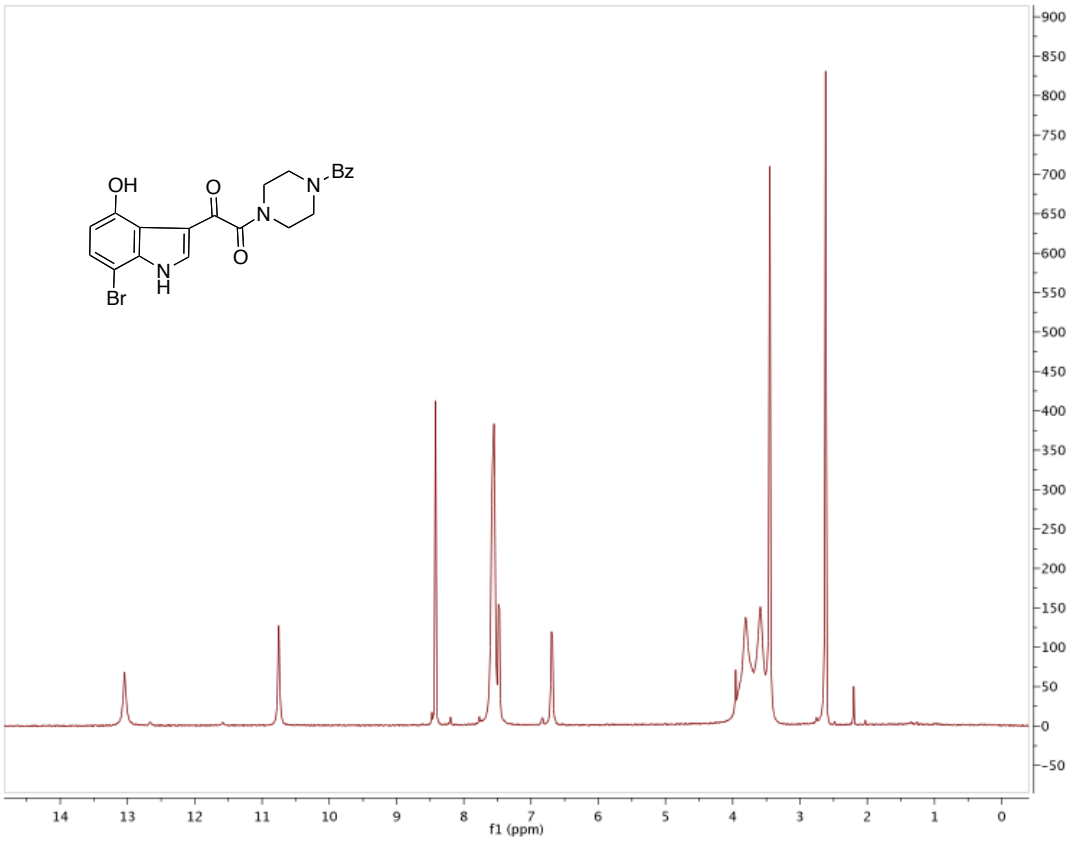
^{1}H NMR ($CDCl_3$) **S10**



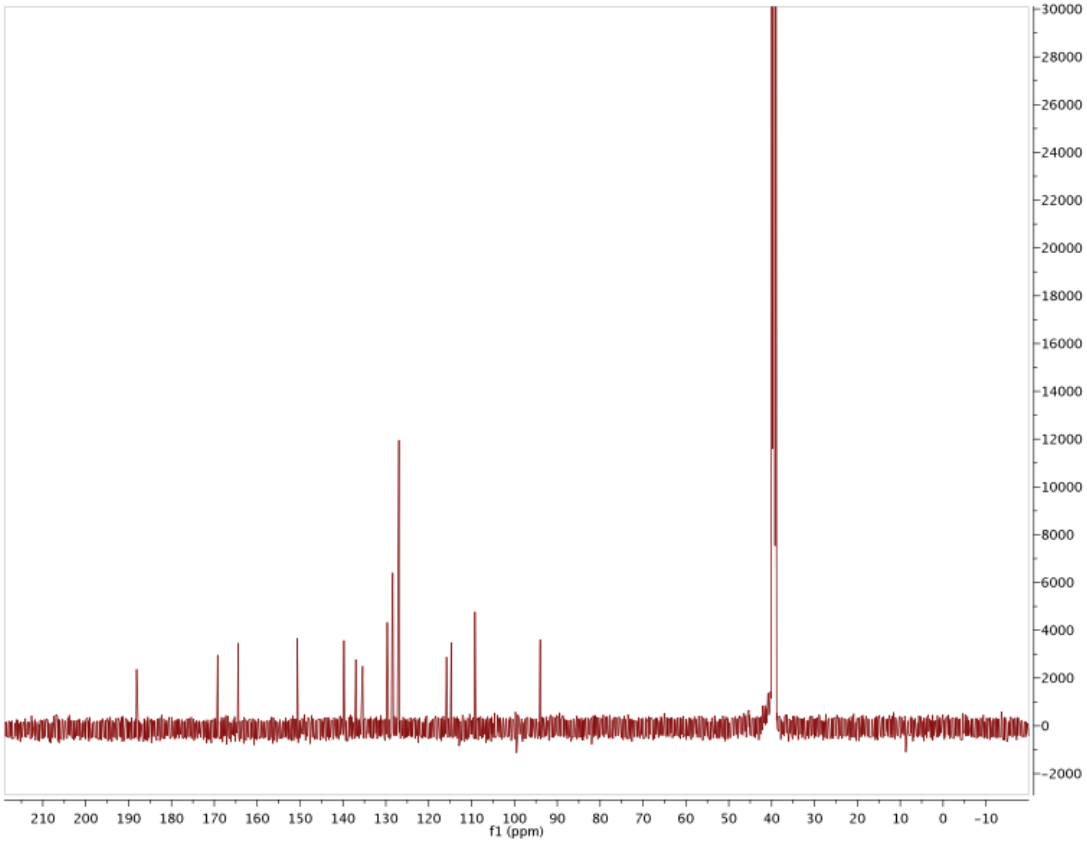
^{1}H NMR ($CDCl_3$) **3**



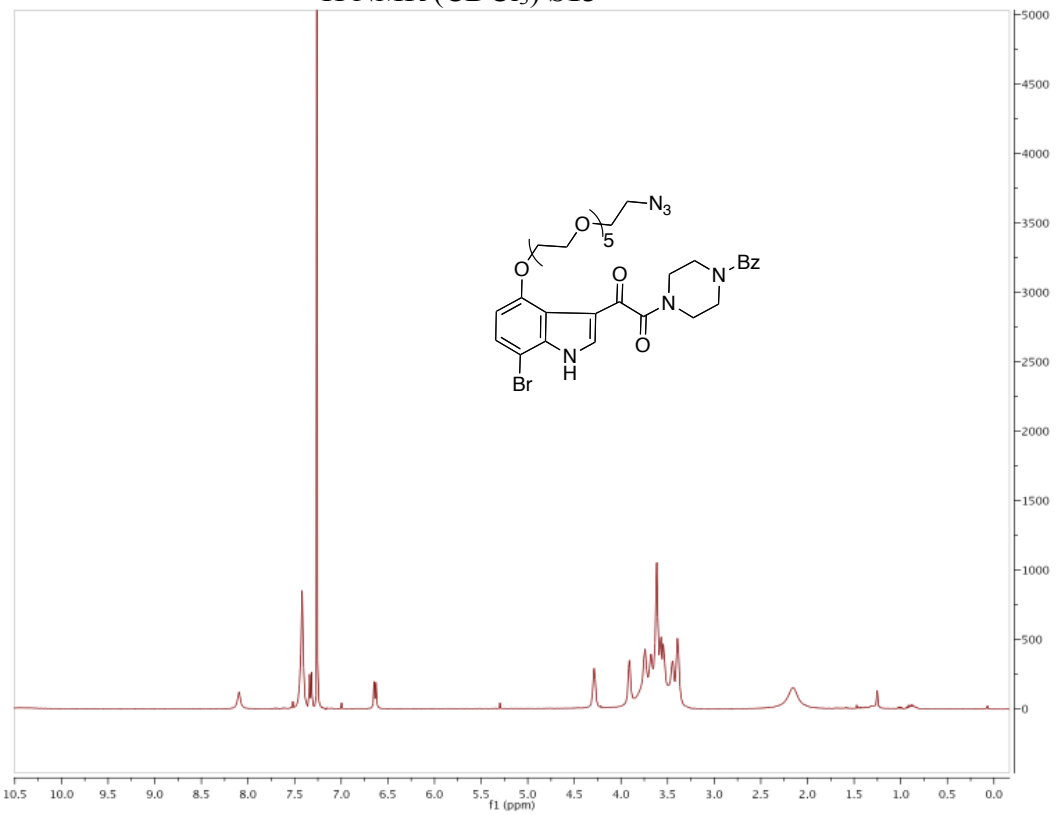
^{1}H NMR (DMSO-d$_6$) **S12**



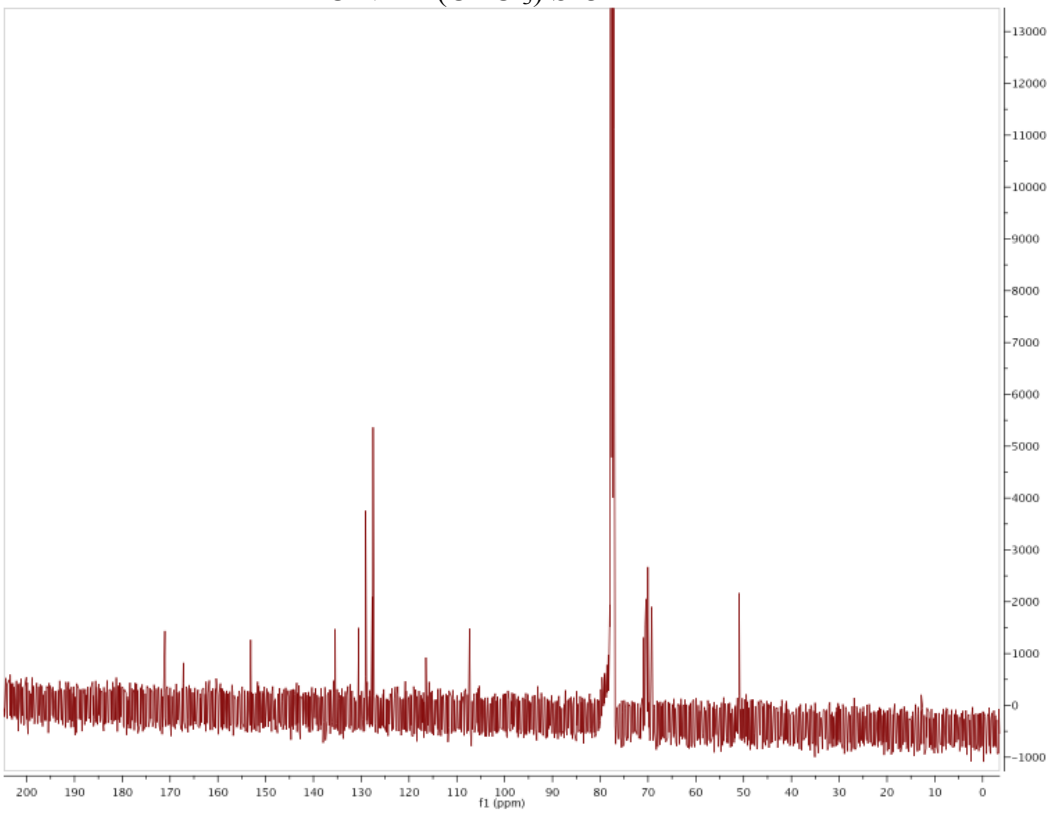
^{13}C NMR (DMSO-d$_6$) **S12**



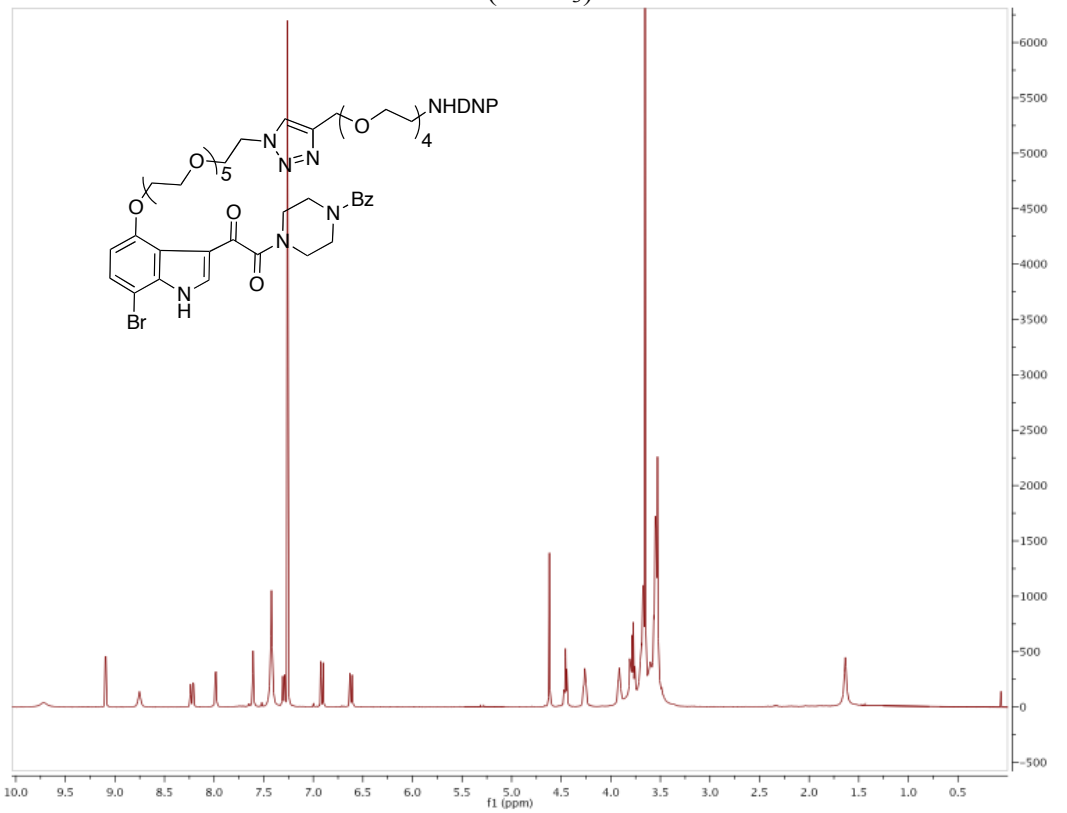
^{1}H NMR ($CDCl_3$) **S13**



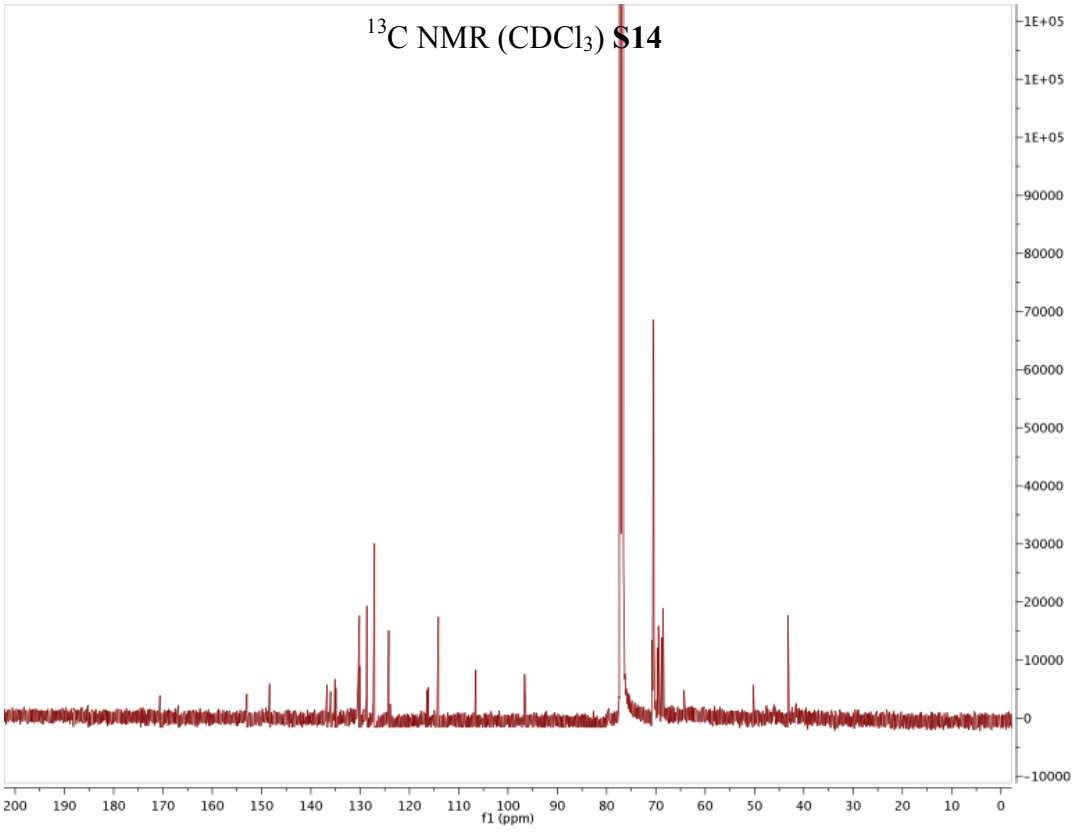
^{13}C NMR ($CDCl_3$) **S13**



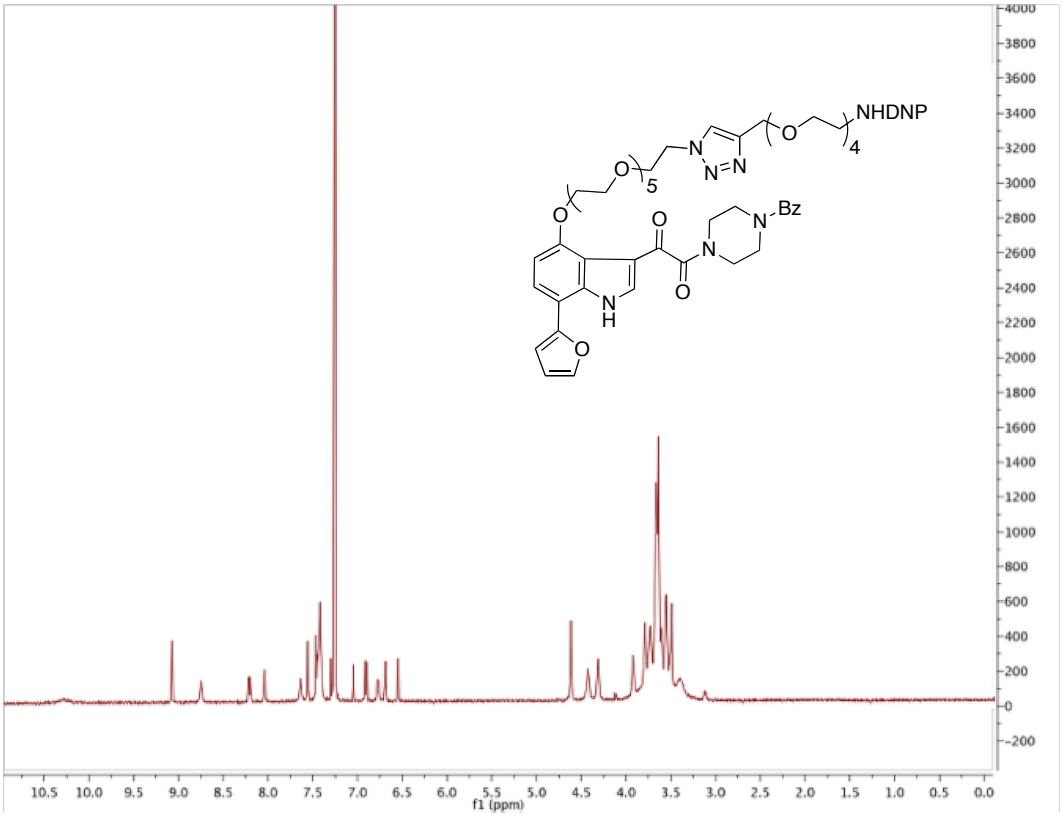
^{1}H NMR ($CDCl_3$) **S14**



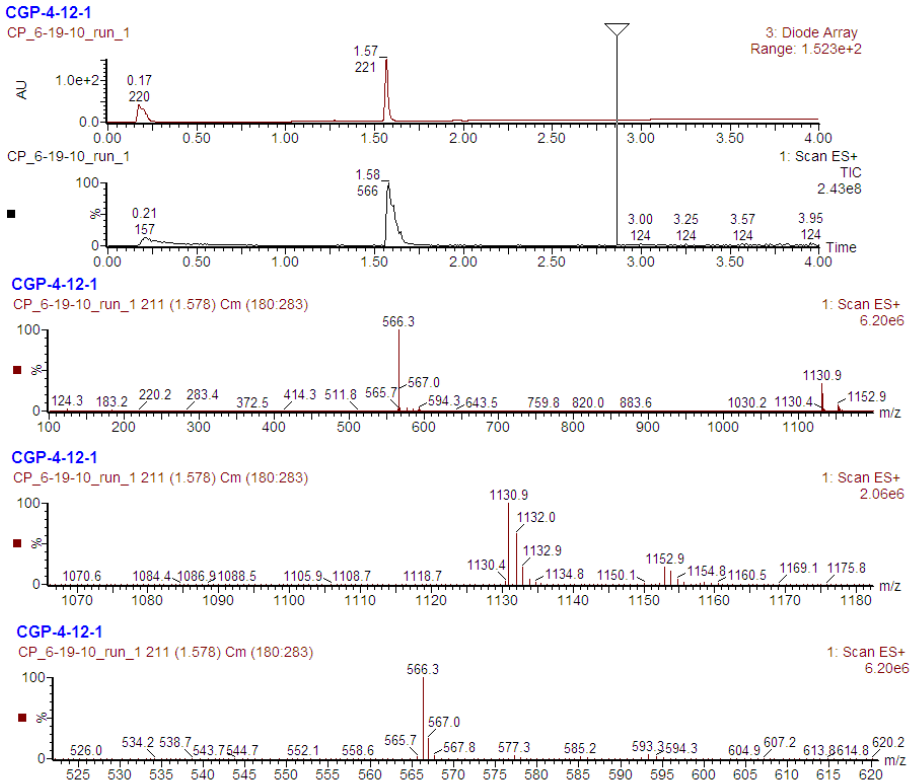
^{13}C NMR ($CDCl_3$) **S14**



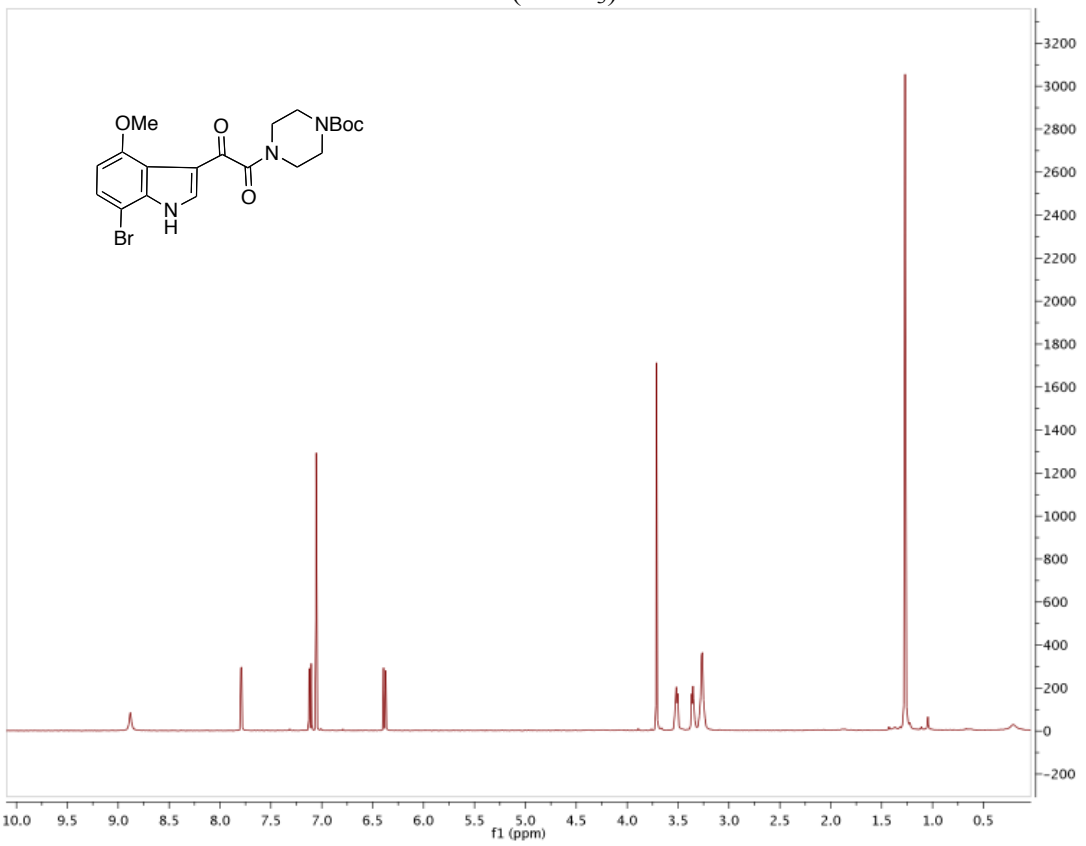
^{1}H NMR ($CDCl_3$) **8**



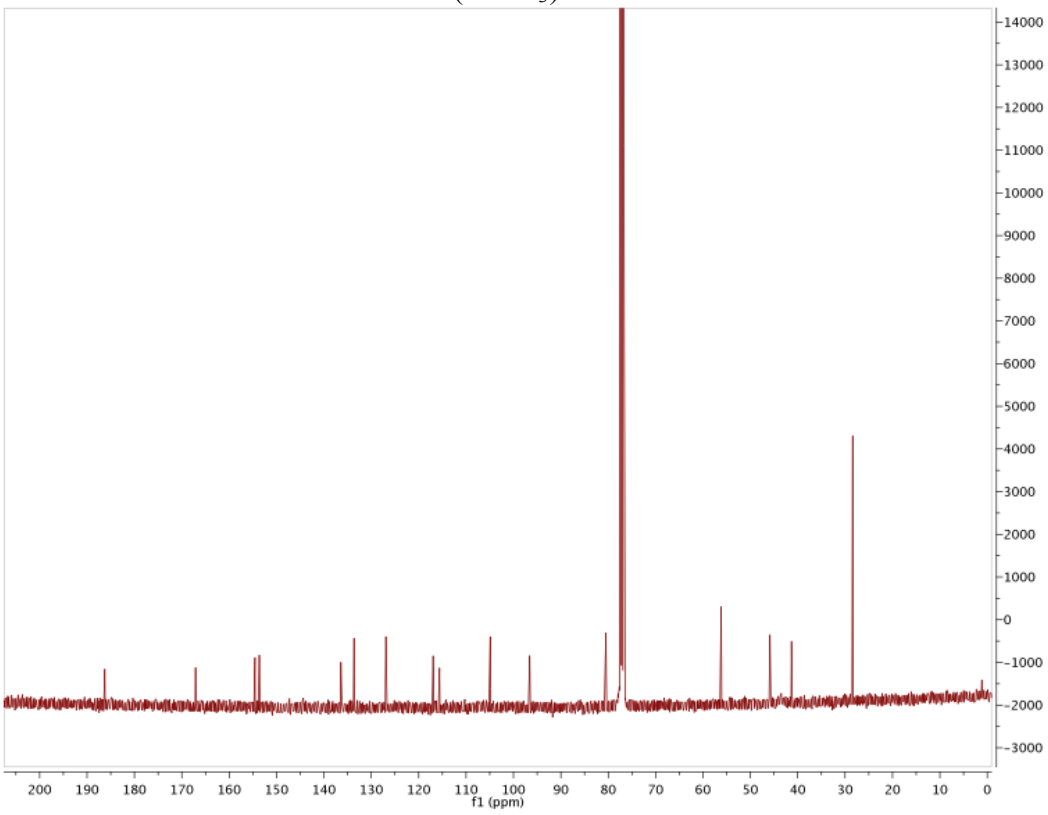
UPLC/MS Trace (**8**)



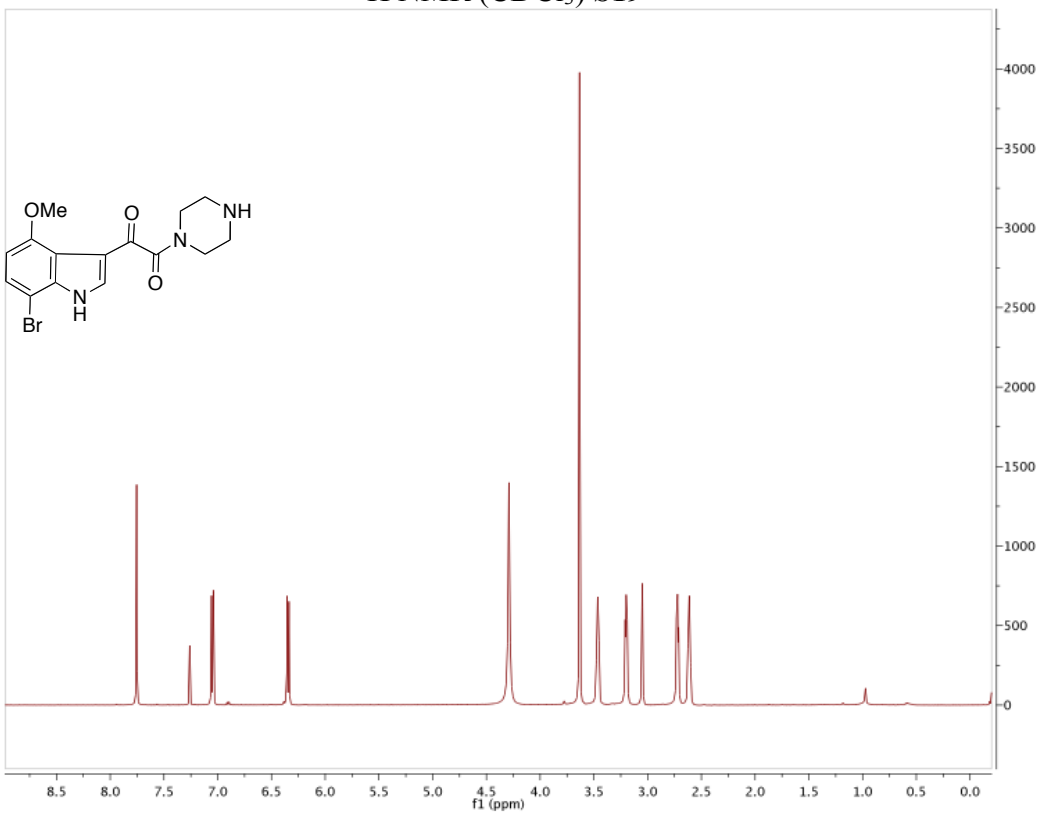
^{1}H NMR ($CDCl_3$) **S18**



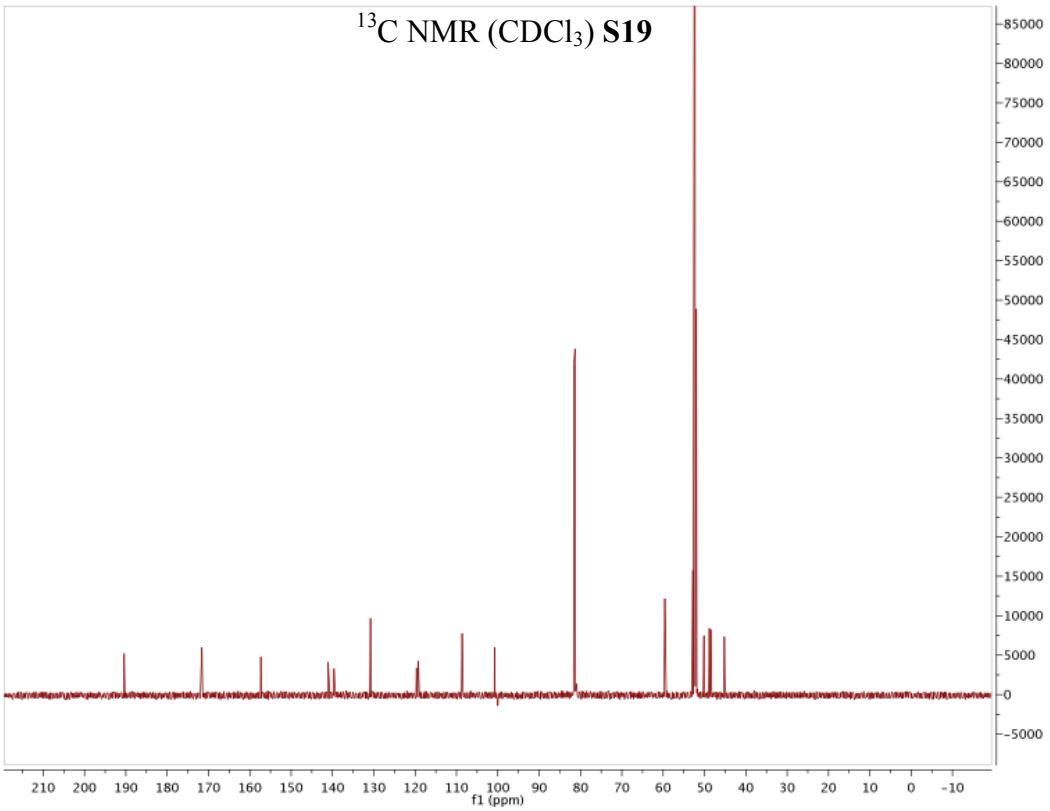
^{13}C NMR ($CDCl_3$) **S18**



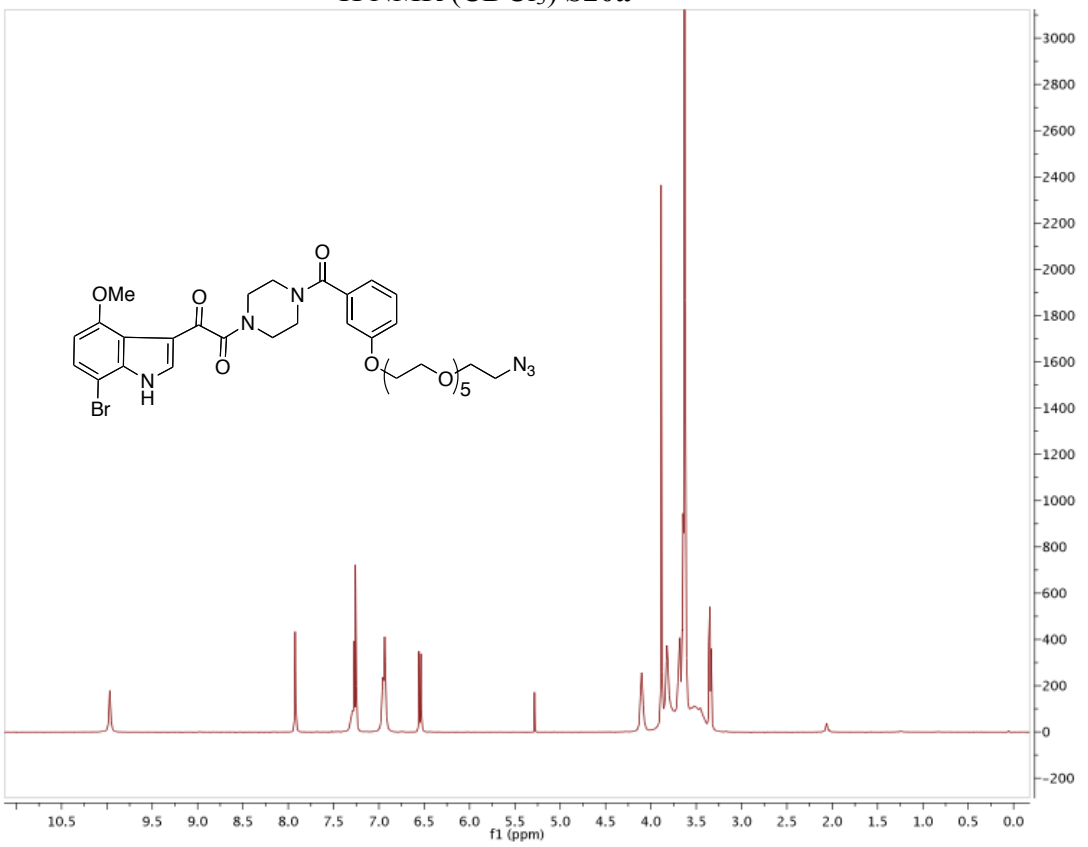
^{1}H NMR ($CDCl_3$) **S19**



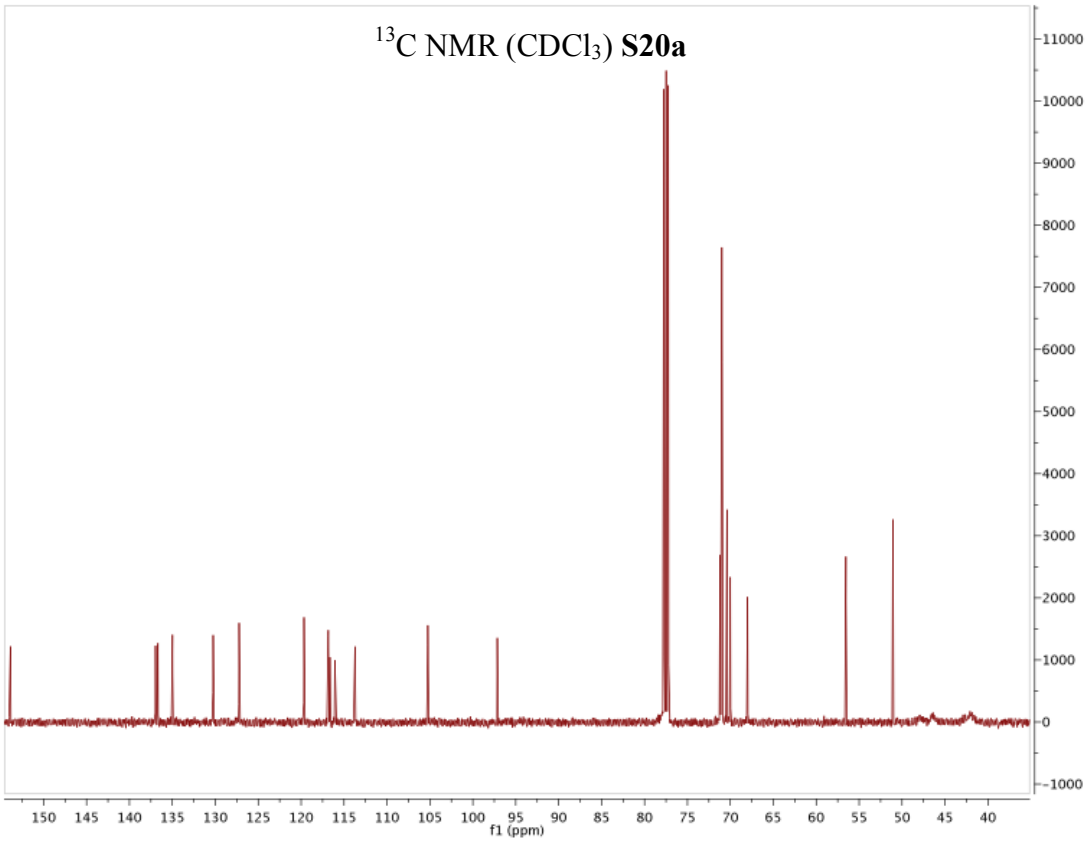
^{13}C NMR ($CDCl_3$) **S19**



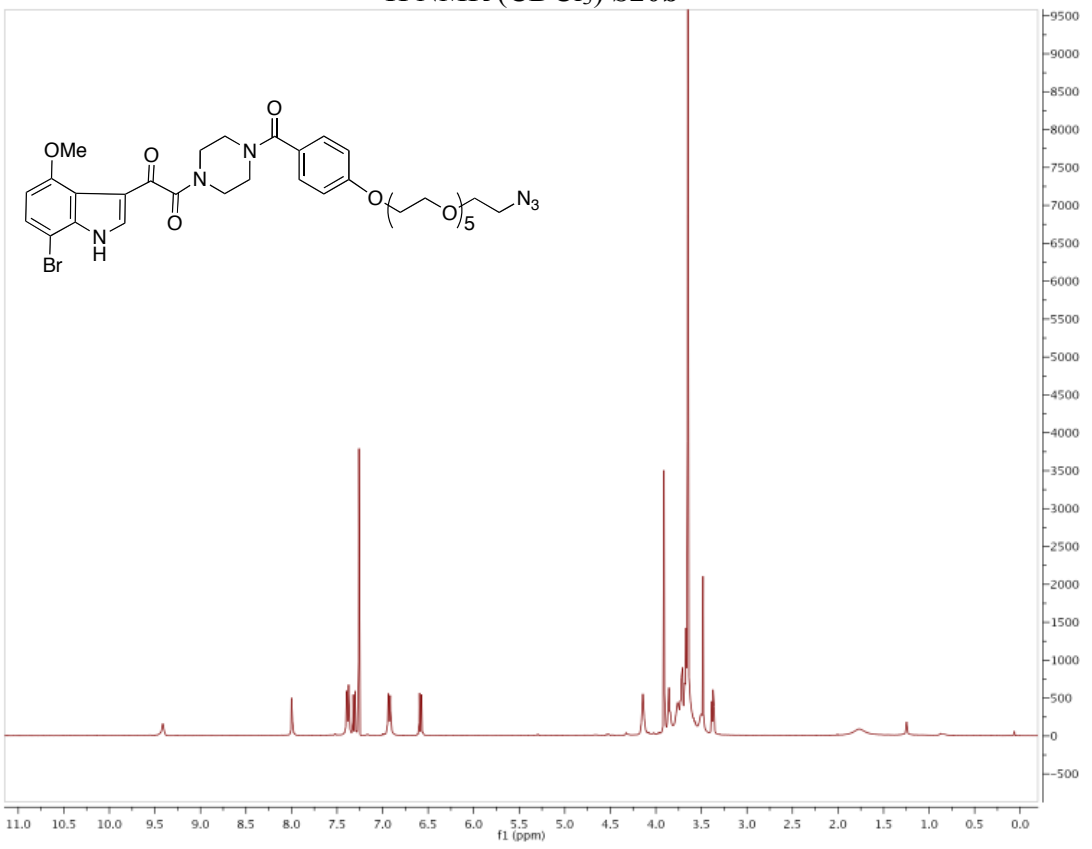
^{1}H NMR ($CDCl_3$) **S20a**



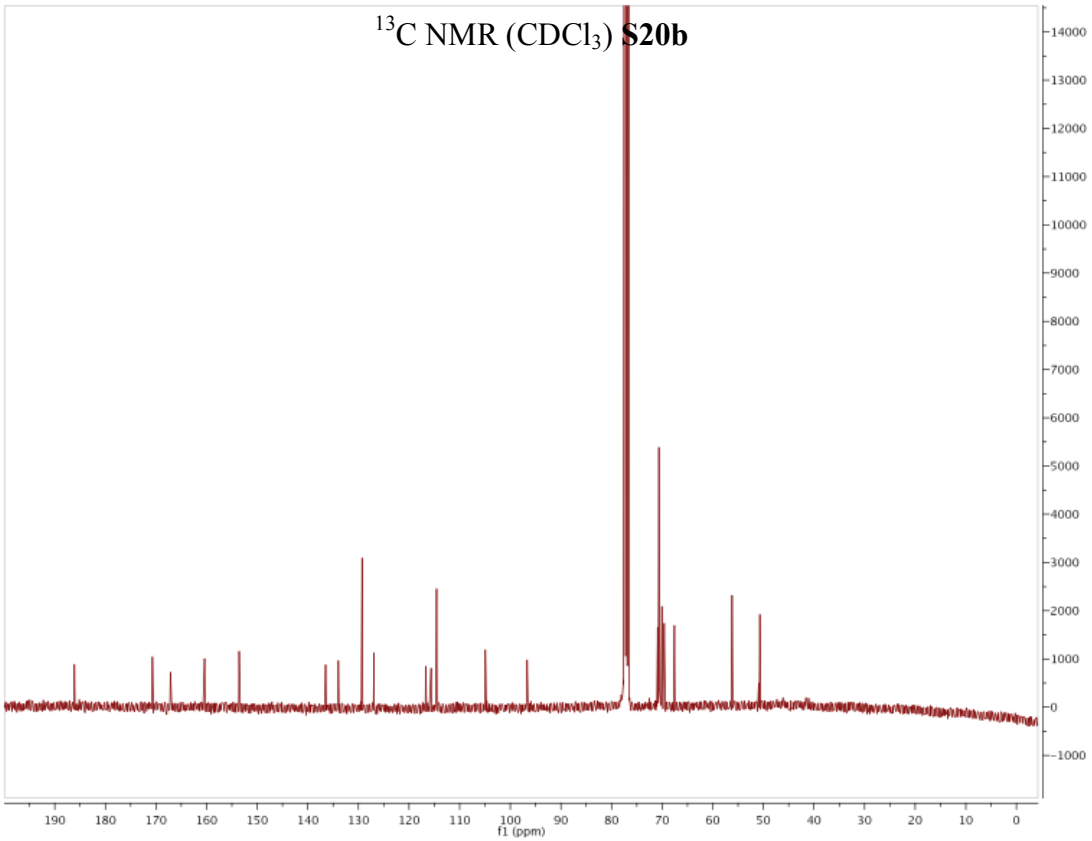
^{13}C NMR ($CDCl_3$) **S20a**



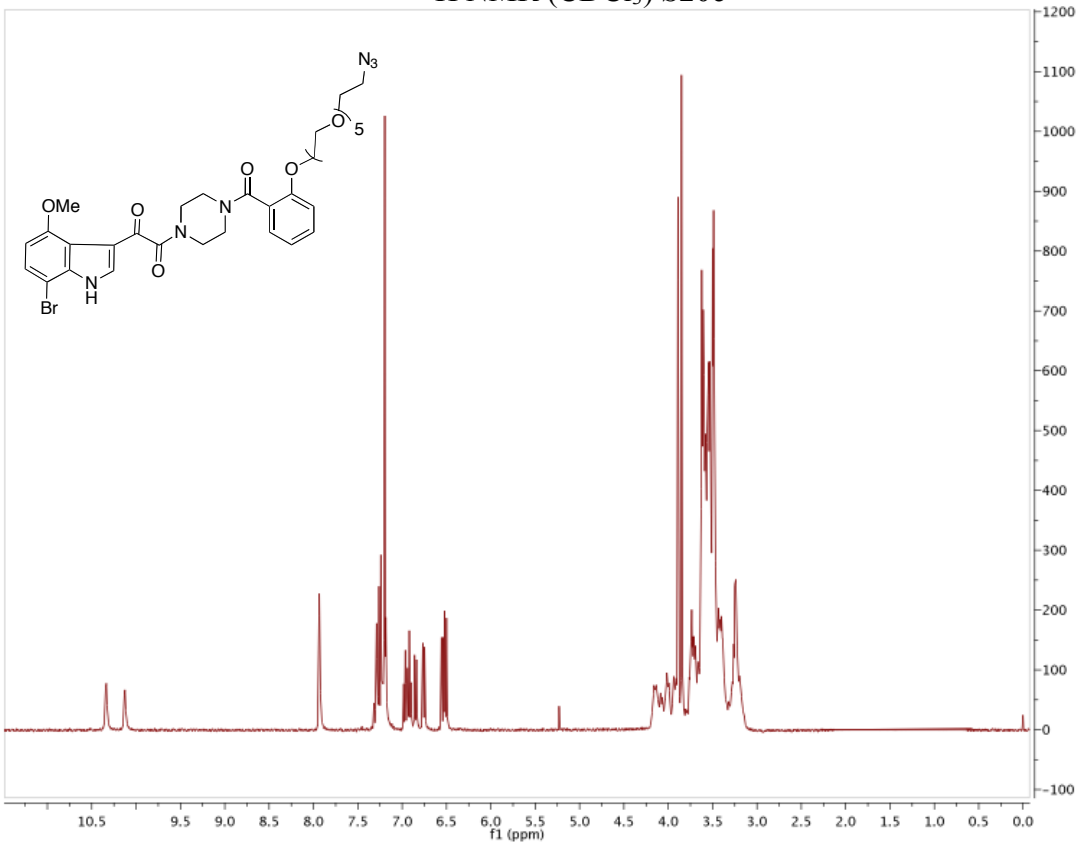
^{1}H NMR ($CDCl_3$) **S20b**



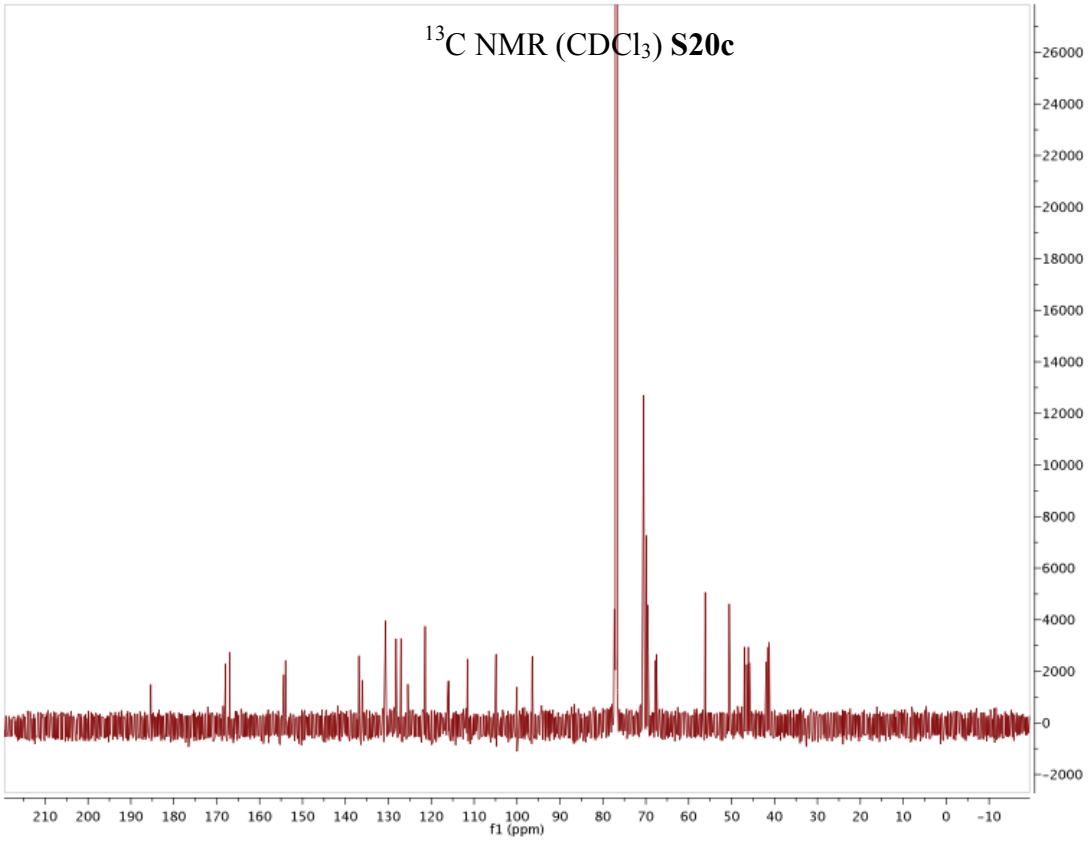
^{13}C NMR ($CDCl_3$) **S20b**



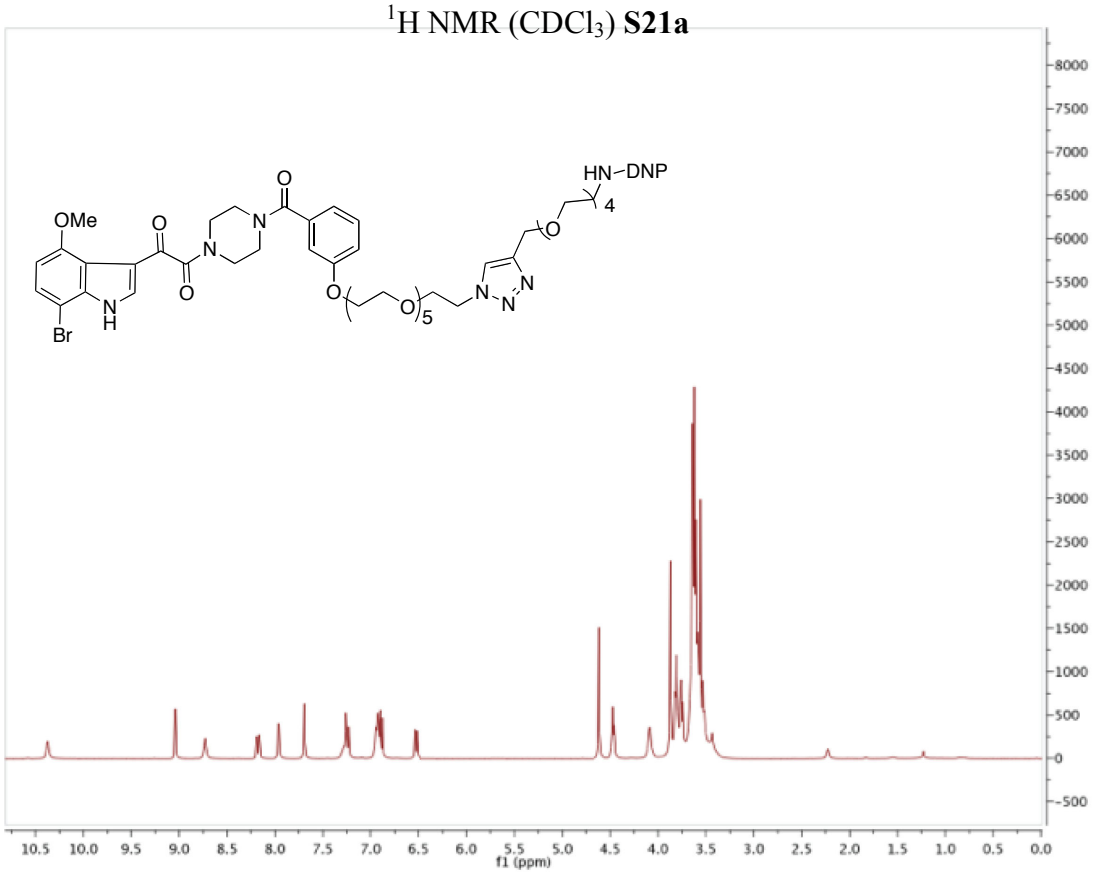
^{1}H NMR ($CDCl_3$) **S20c**



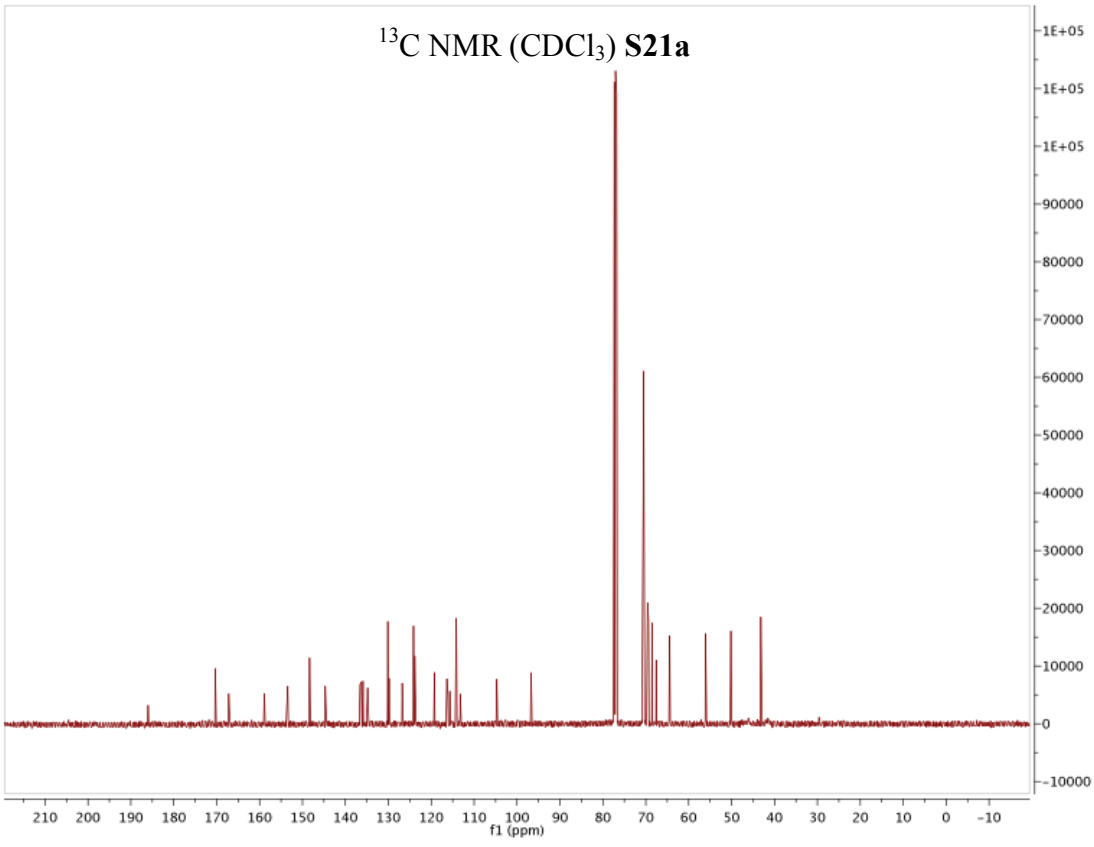
^{13}C NMR ($CDCl_3$) **S20c**



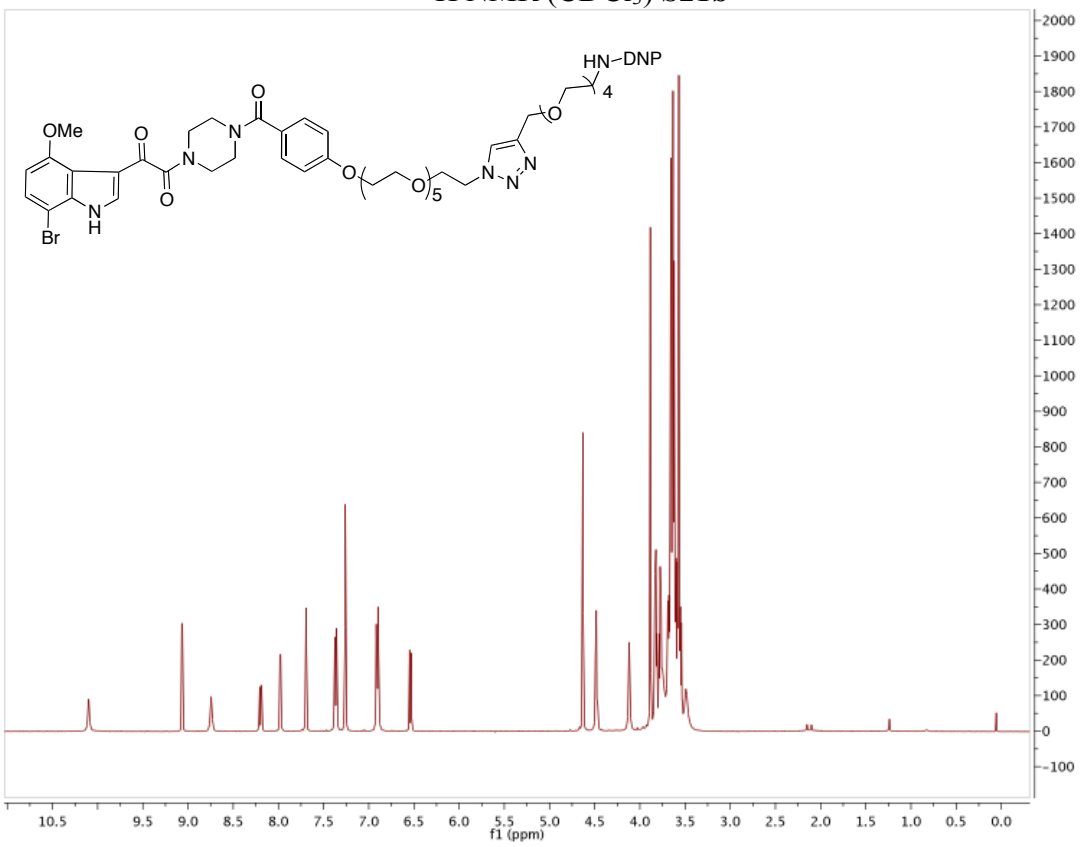
^{1}H NMR ($CDCl_3$) **S21a**



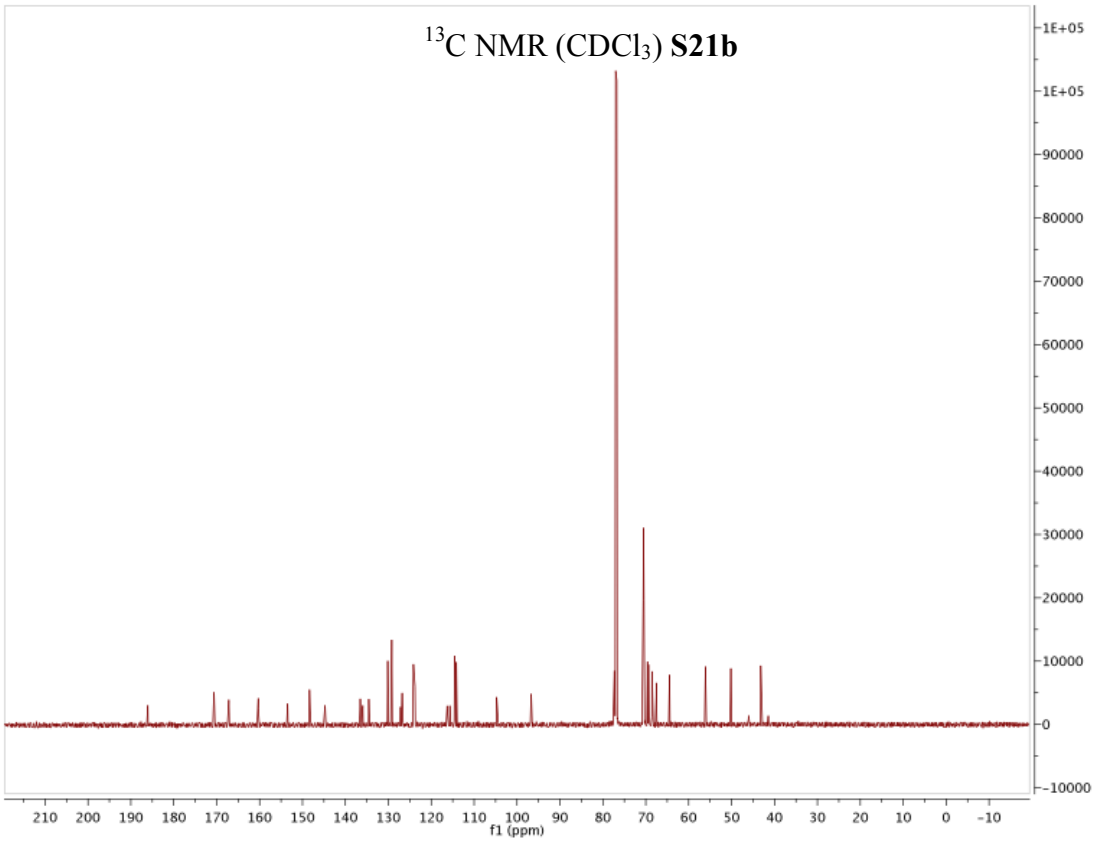
^{13}C NMR ($CDCl_3$) **S21a**



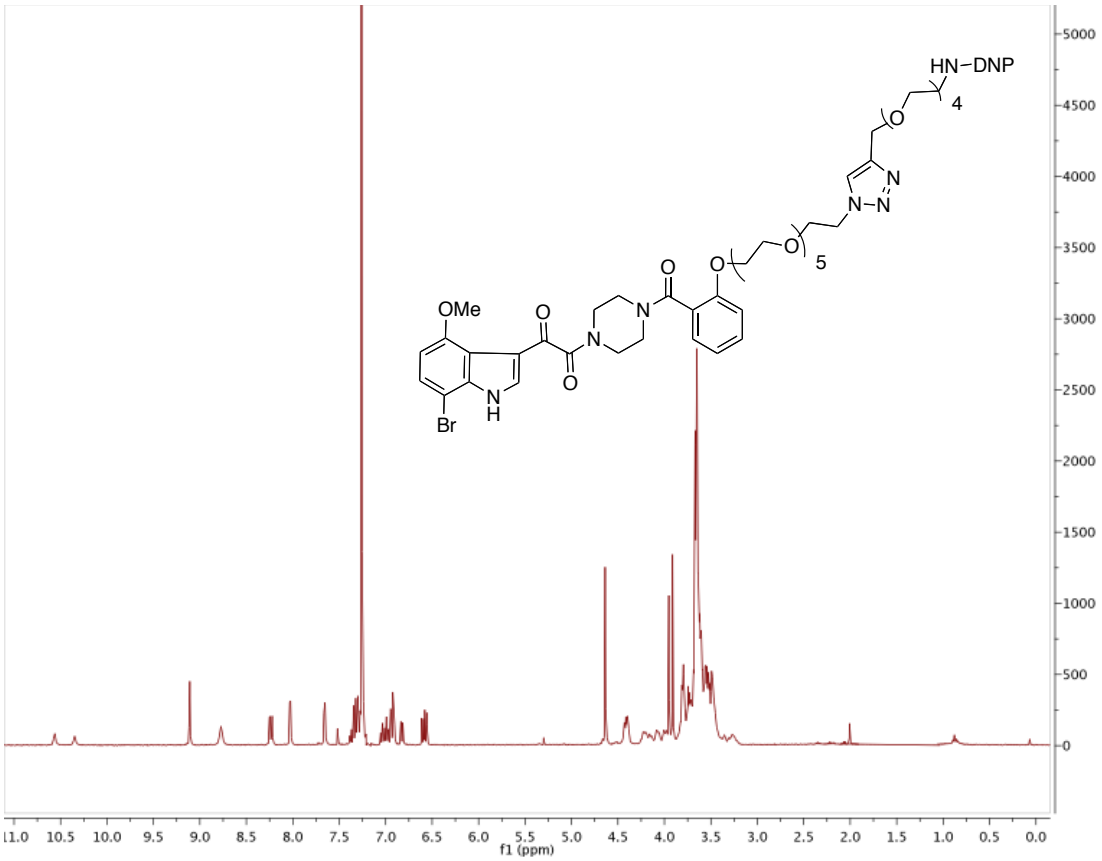
^{1}H NMR ($CDCl_3$) **S21b**



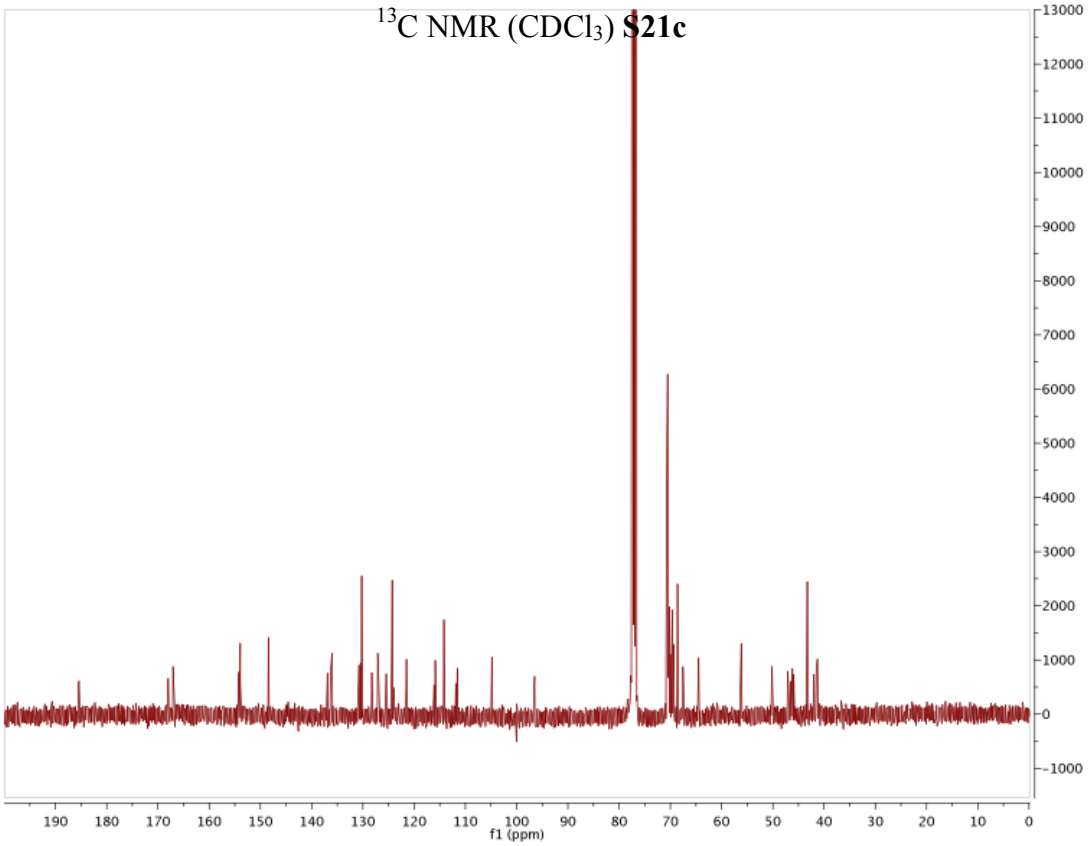
^{13}C NMR ($CDCl_3$) **S21b**



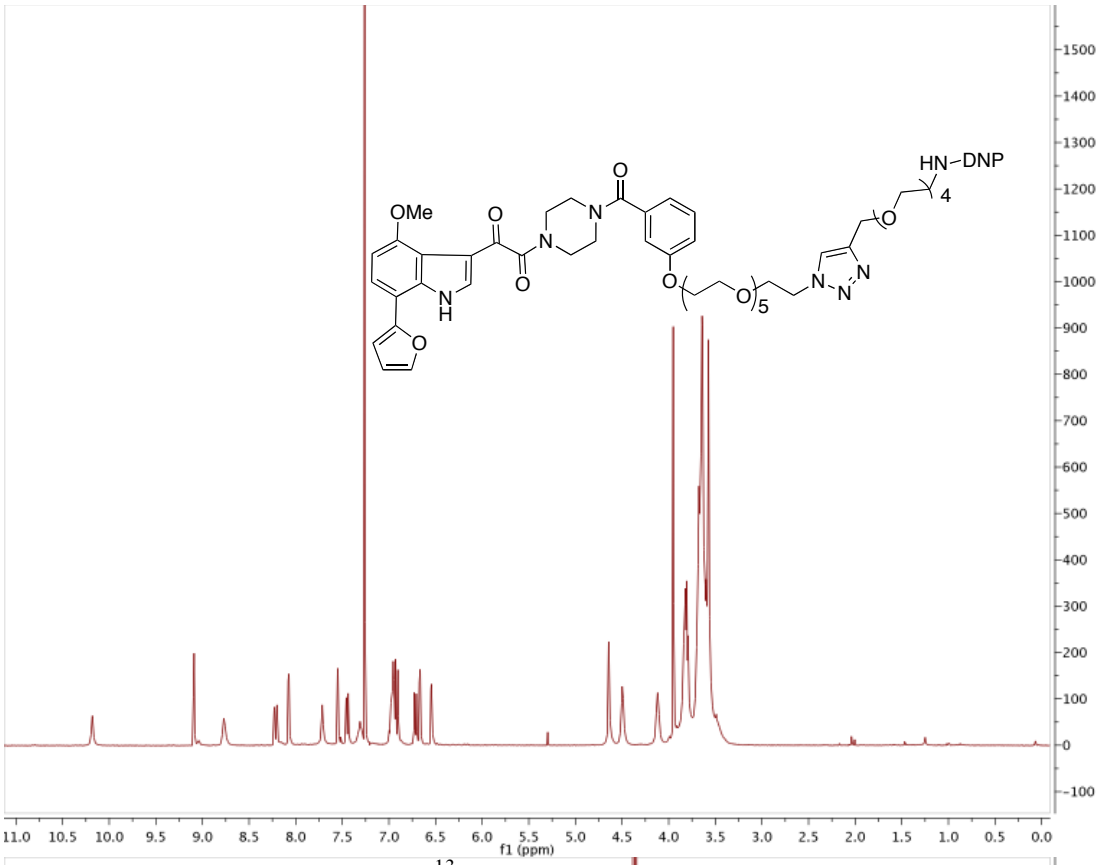
^{1}H NMR ($CDCl_3$) **S21c**



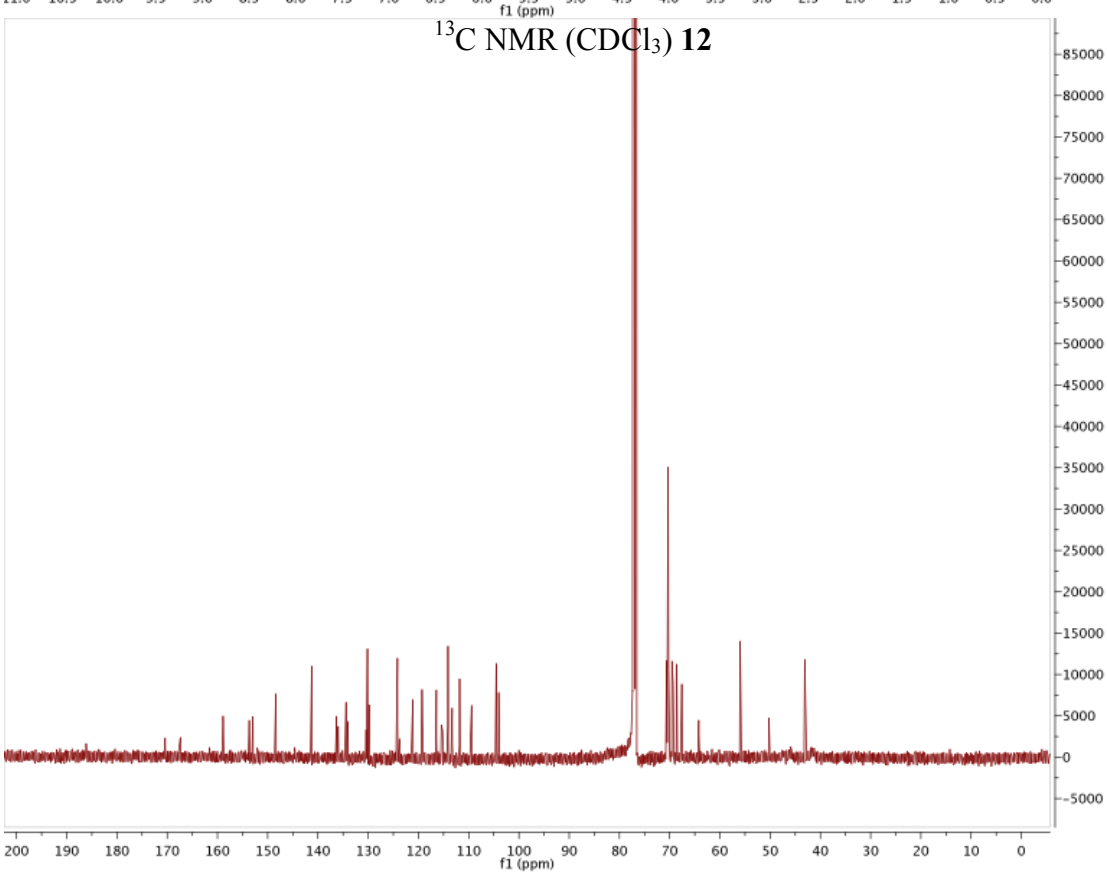
^{13}C NMR ($CDCl_3$) **S21c**



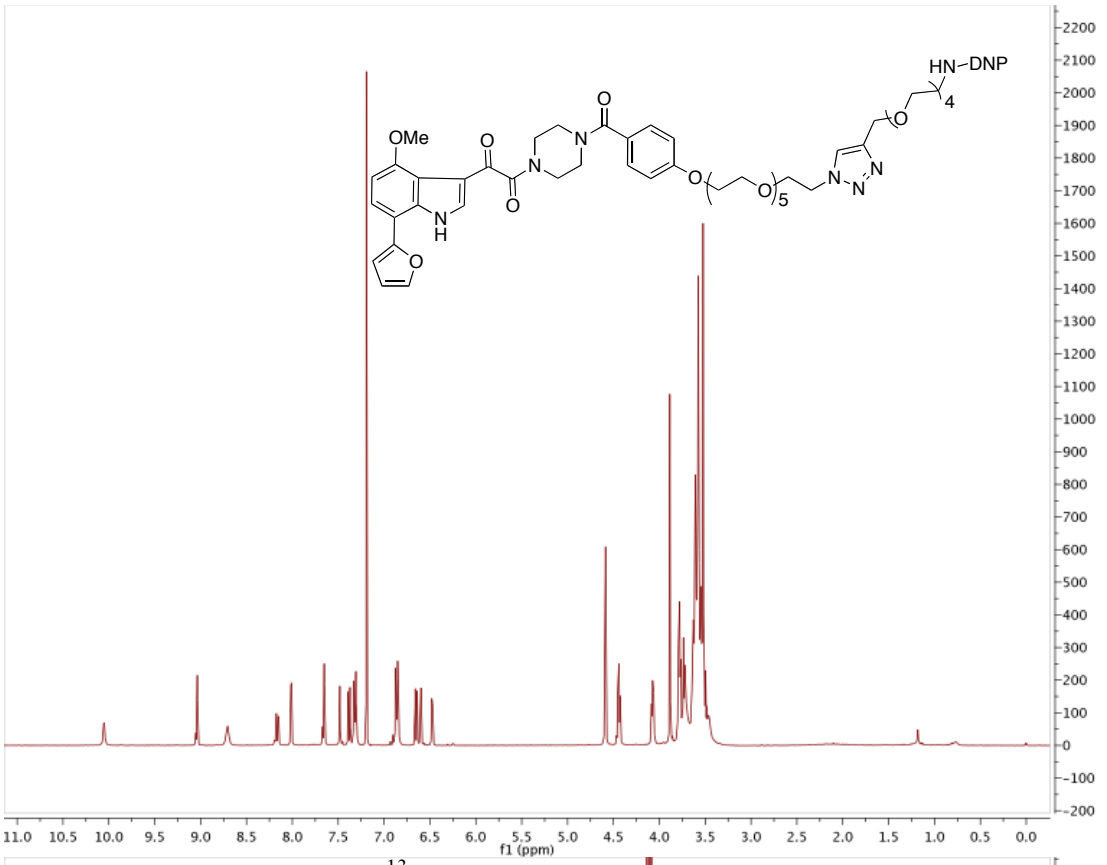
^{1}H NMR ($CDCl_3$) **12**



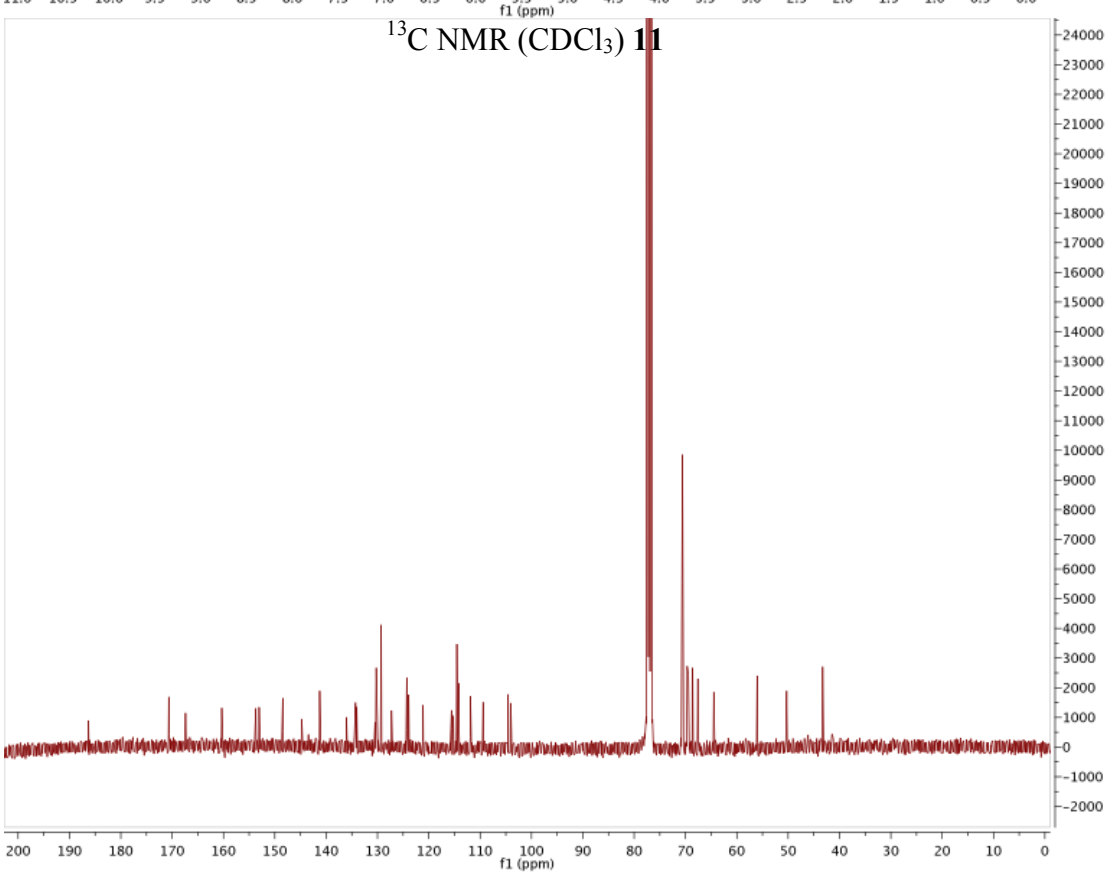
^{13}C NMR ($CDCl_3$) **12**



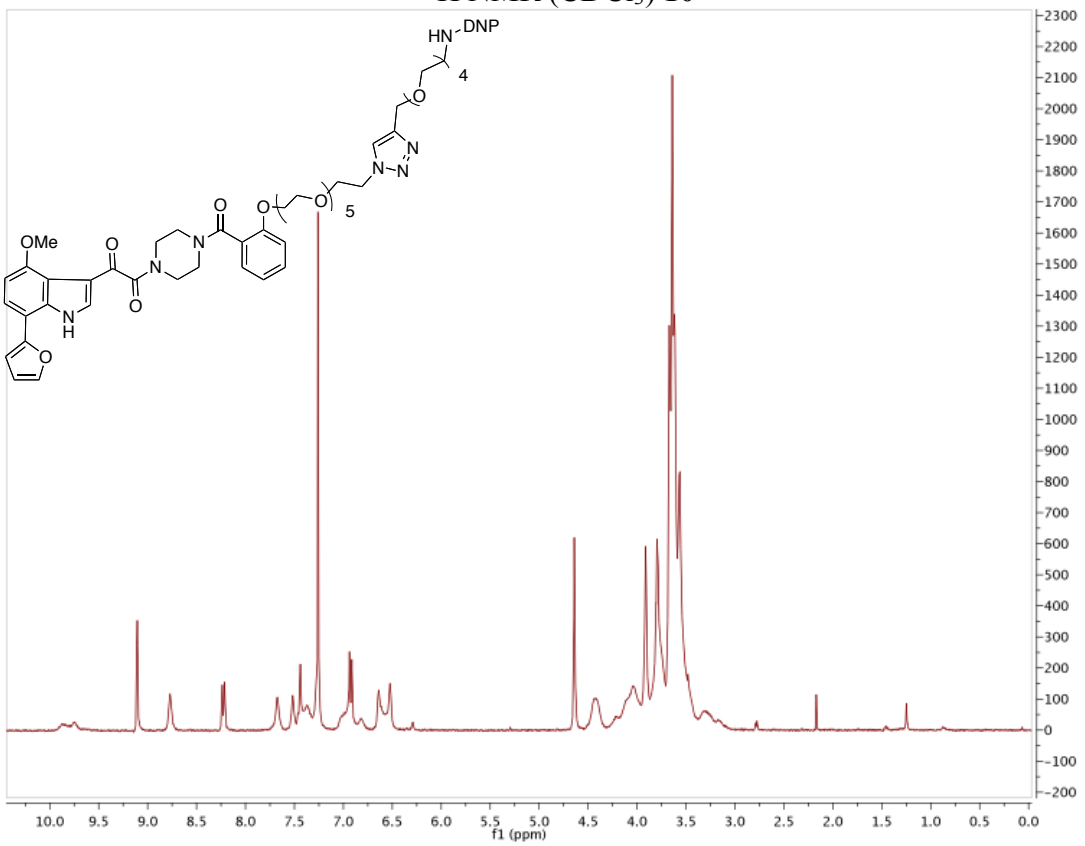
^{1}H NMR ($CDCl_3$) **11**



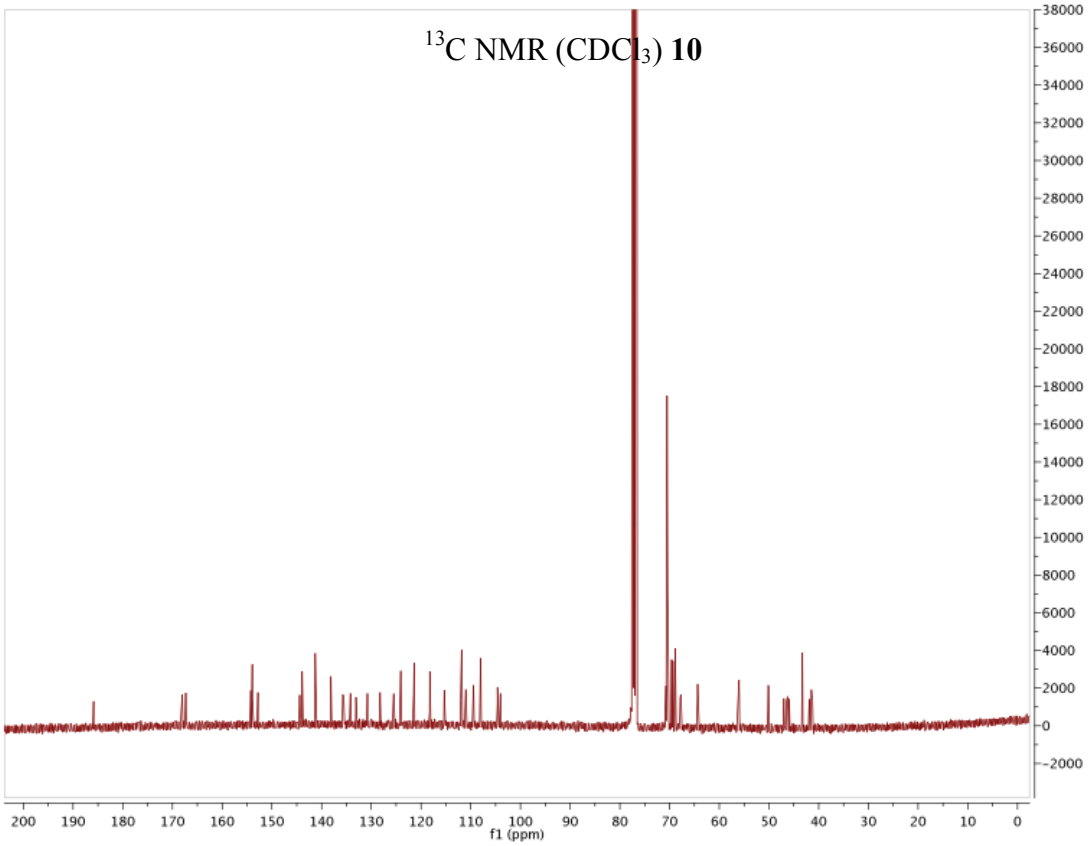
^{13}C NMR ($CDCl_3$) **11**



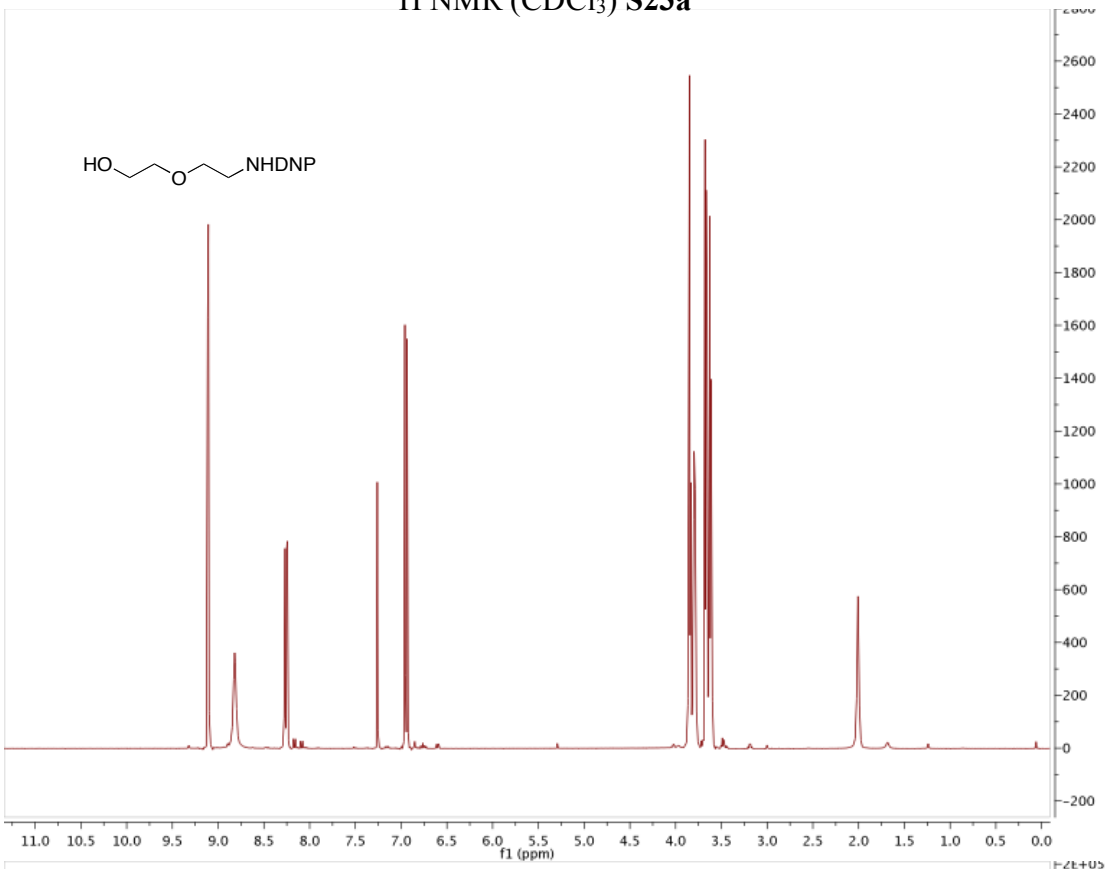
^{1}H NMR ($CDCl_3$) **10**



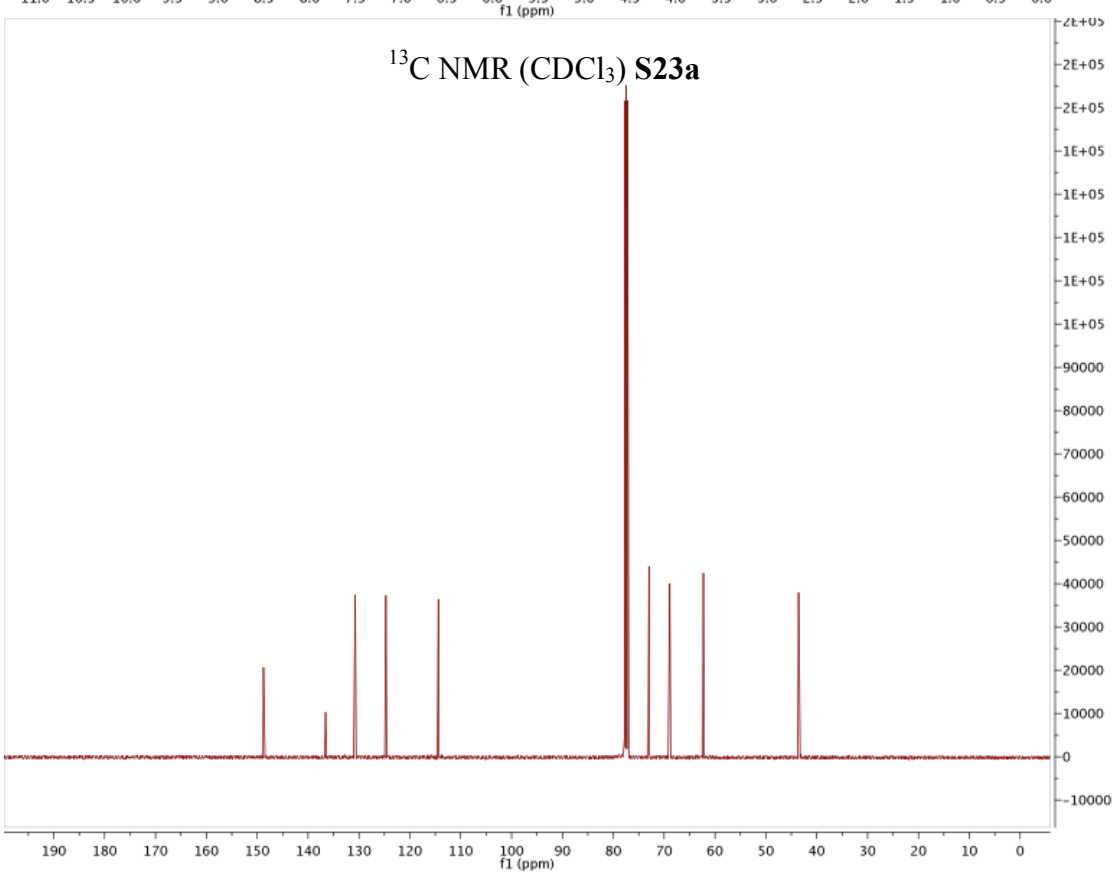
^{13}C NMR ($CDCl_3$) **10**



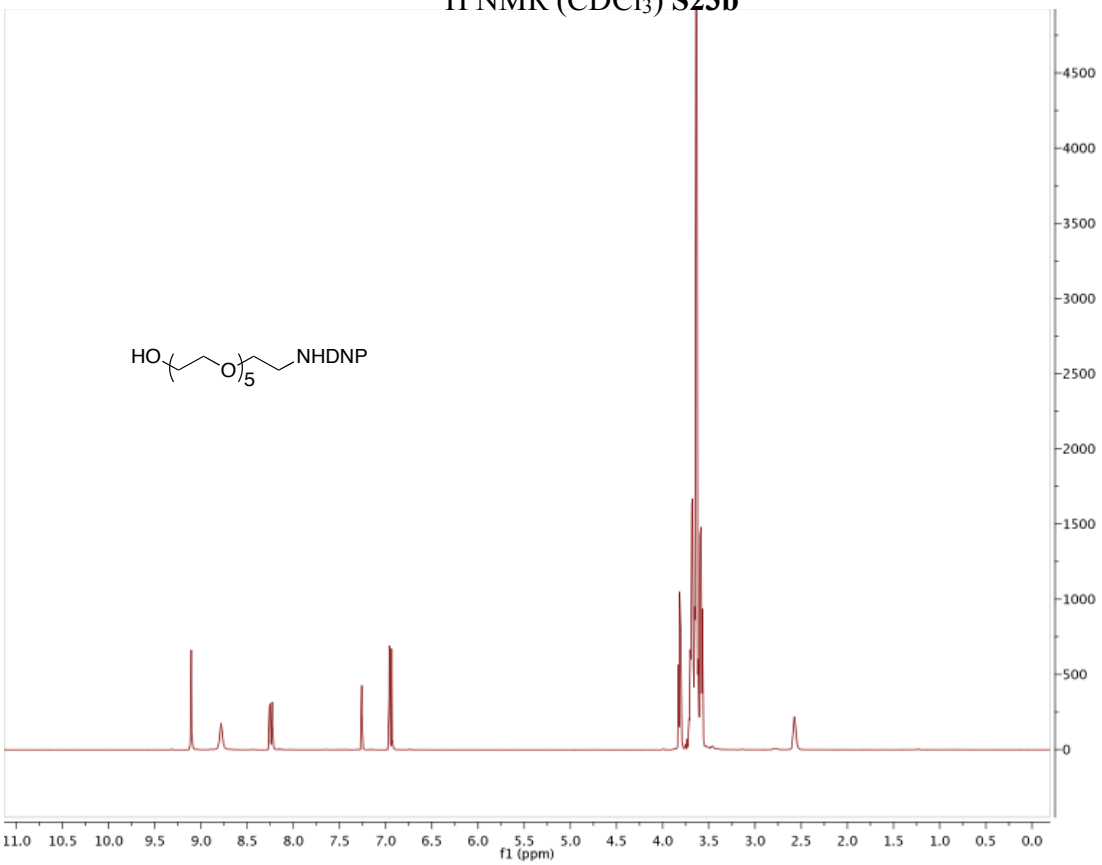
^{1}H NMR ($CDCl_3$) **S23a**



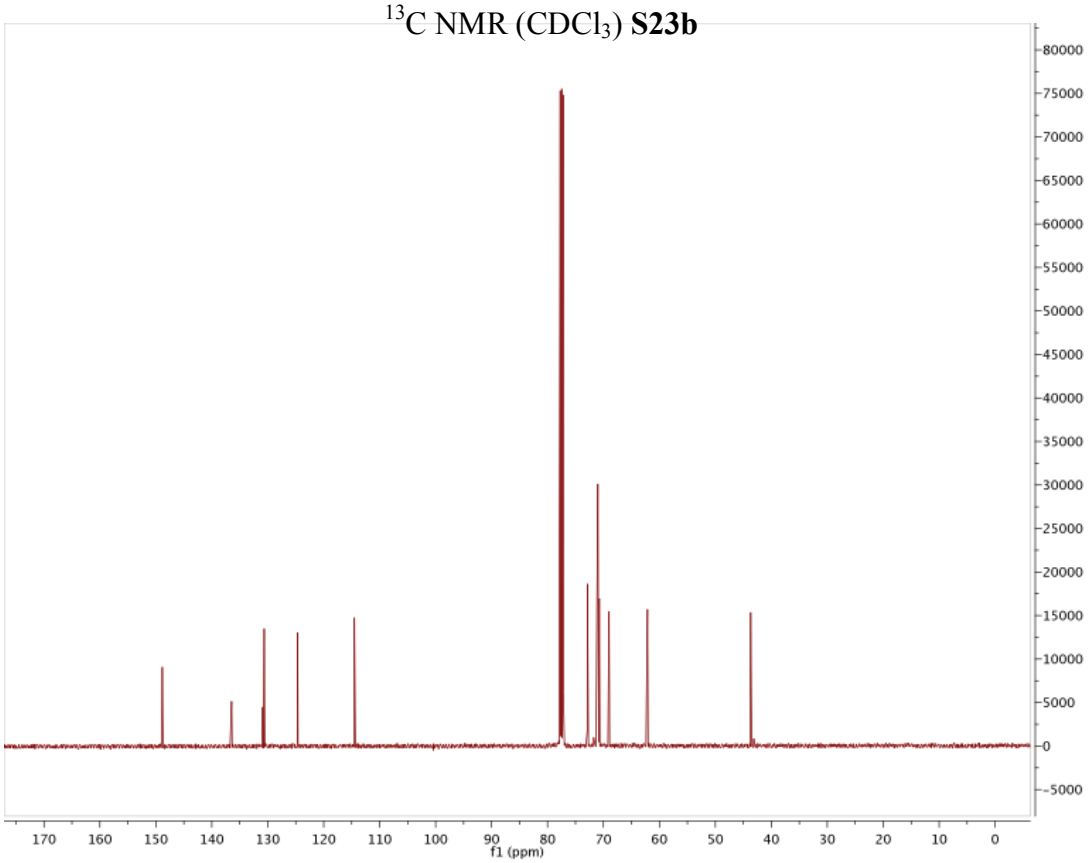
^{13}C NMR ($CDCl_3$) **S23a**



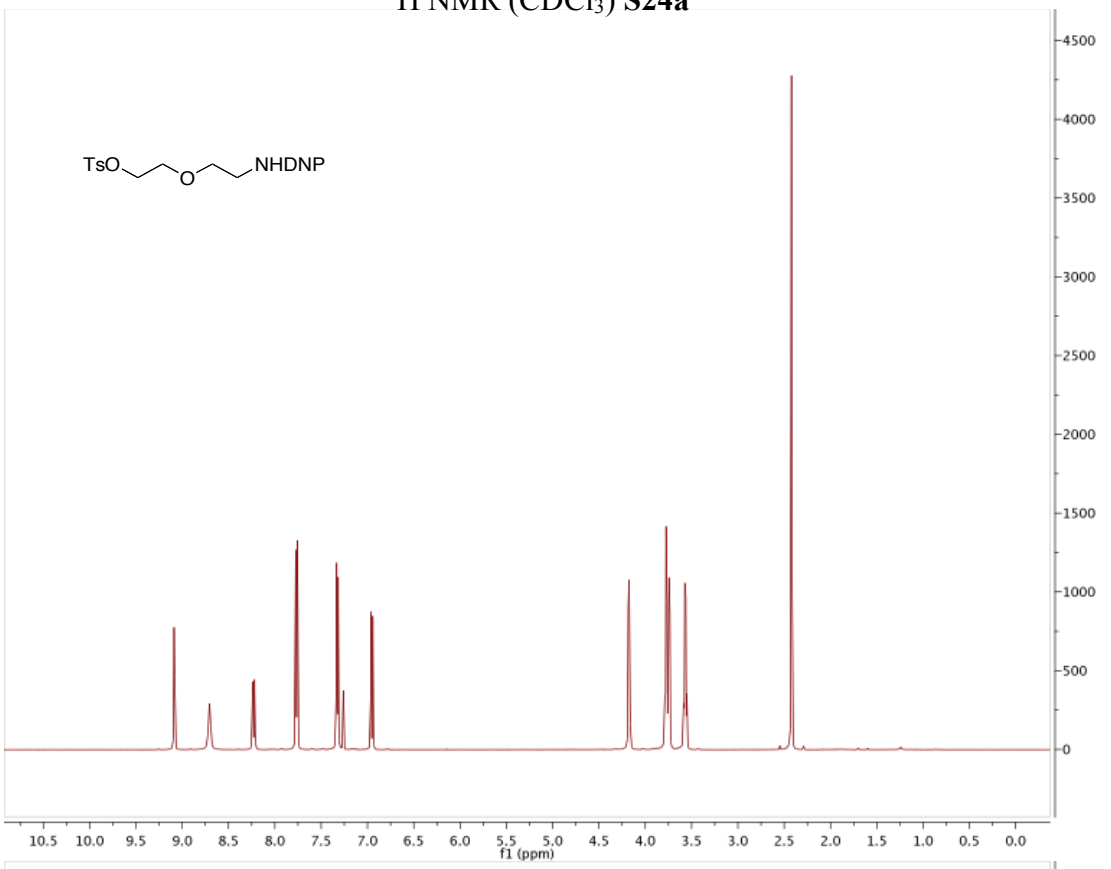
^{1}H NMR ($CDCl_3$) **S23b**



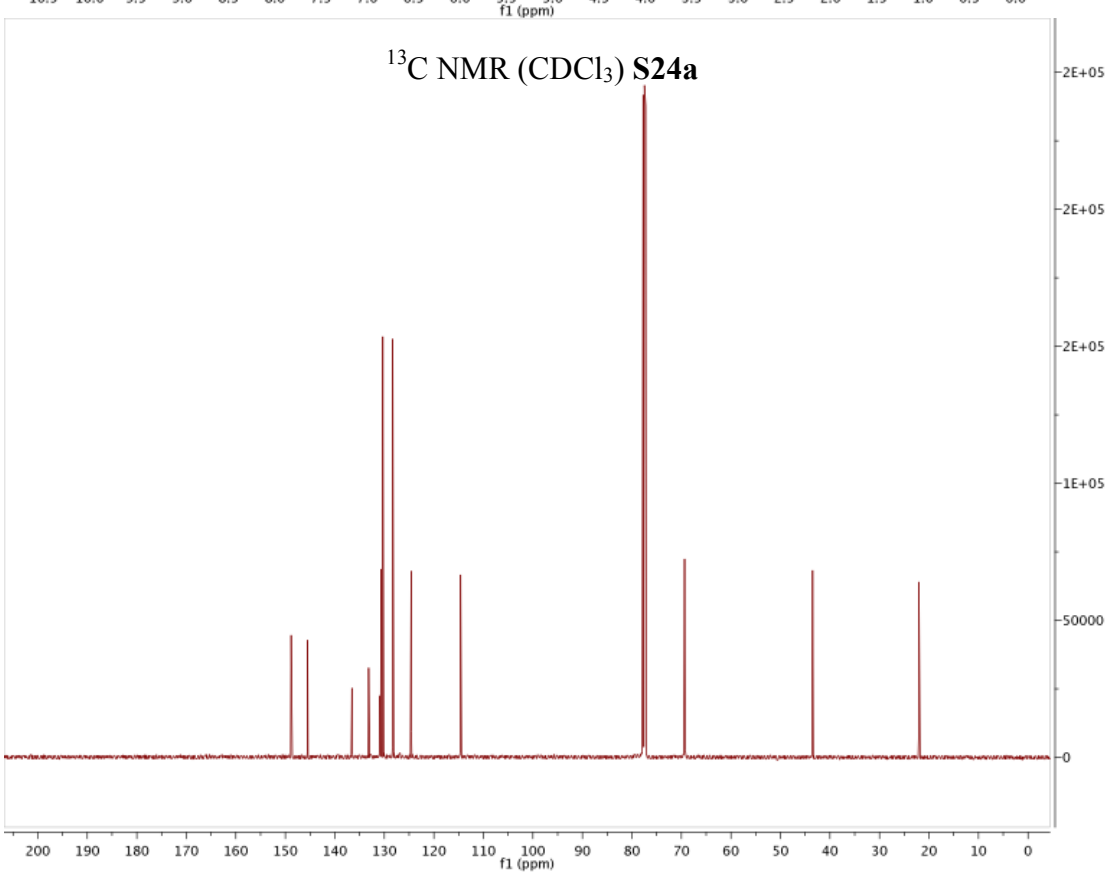
^{13}C NMR ($CDCl_3$) **S23b**



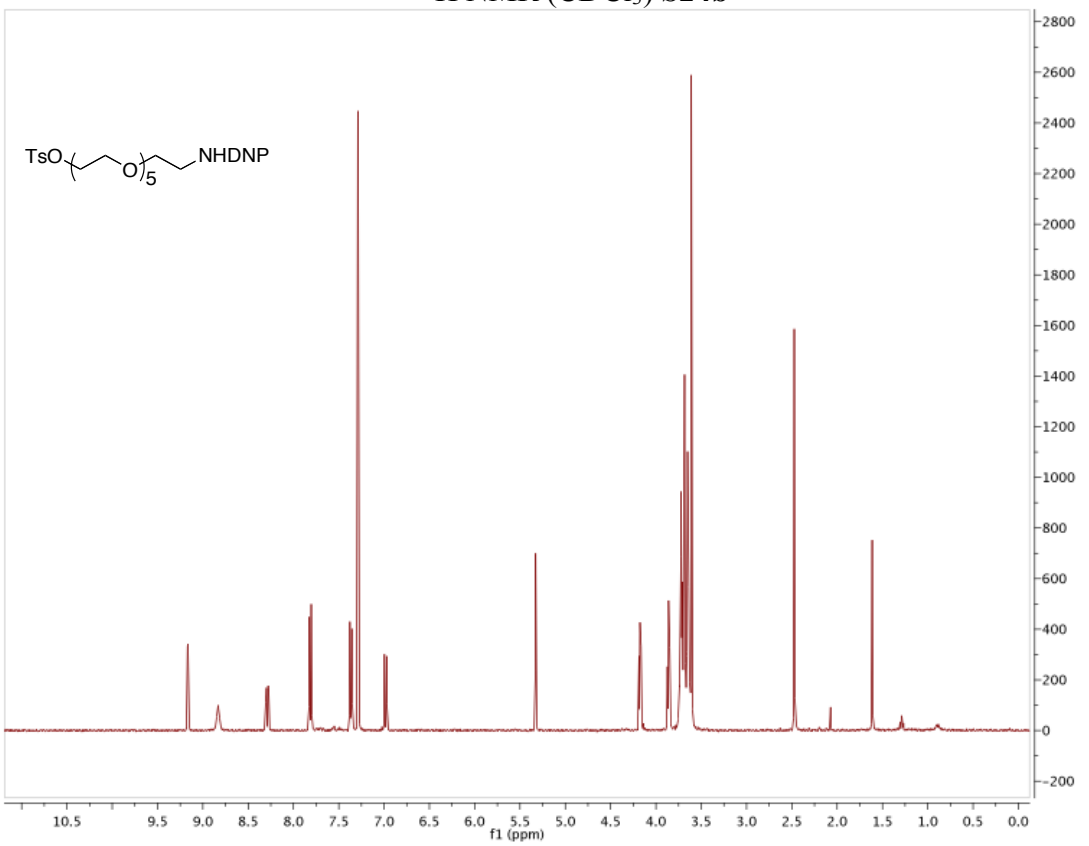
^{1}H NMR ($CDCl_3$) **S24a**



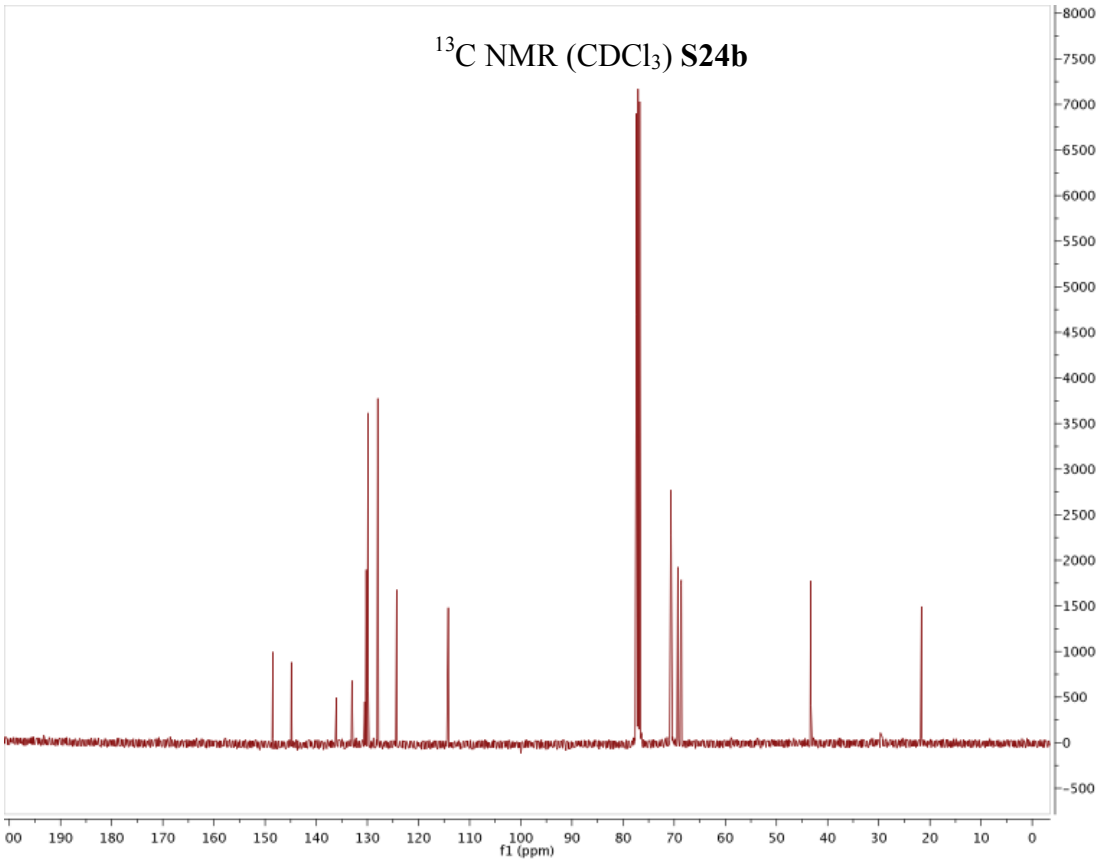
^{13}C NMR ($CDCl_3$) **S24a**



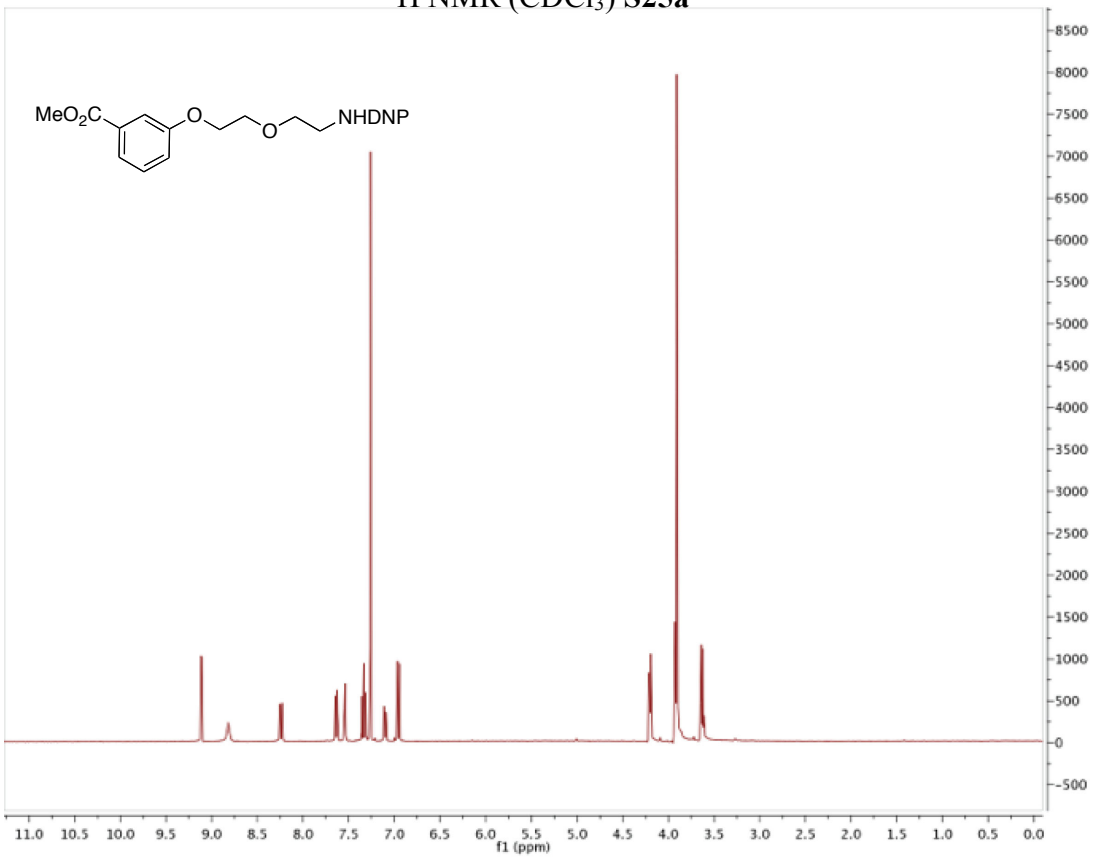
^{1}H NMR ($CDCl_3$) **S24b**



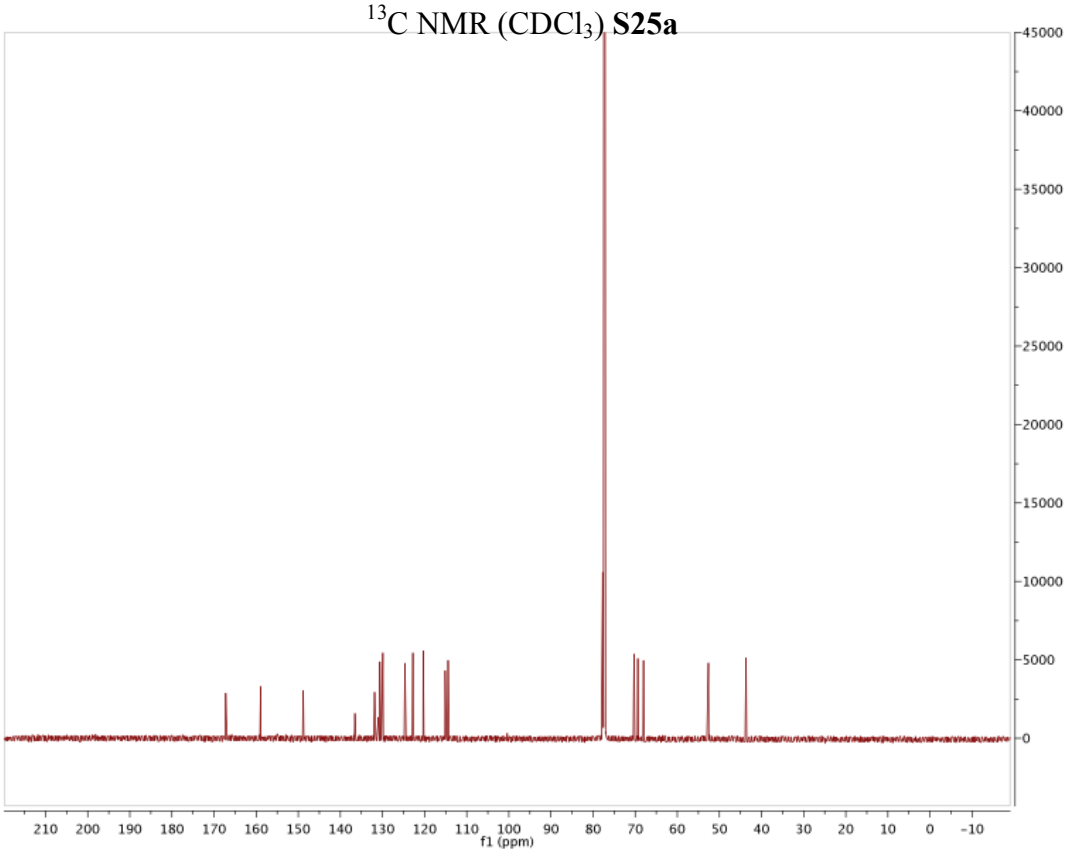
^{13}C NMR ($CDCl_3$) **S24b**



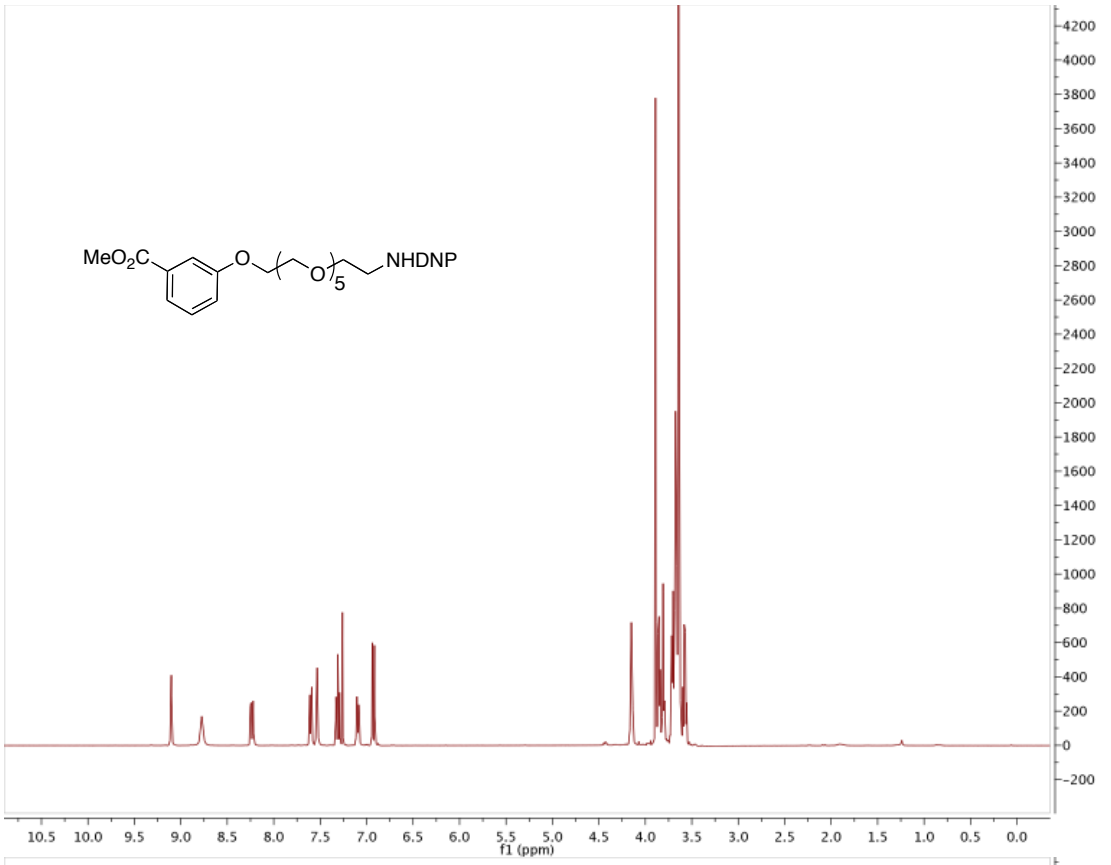
^{1}H NMR ($CDCl_3$) **S25a**



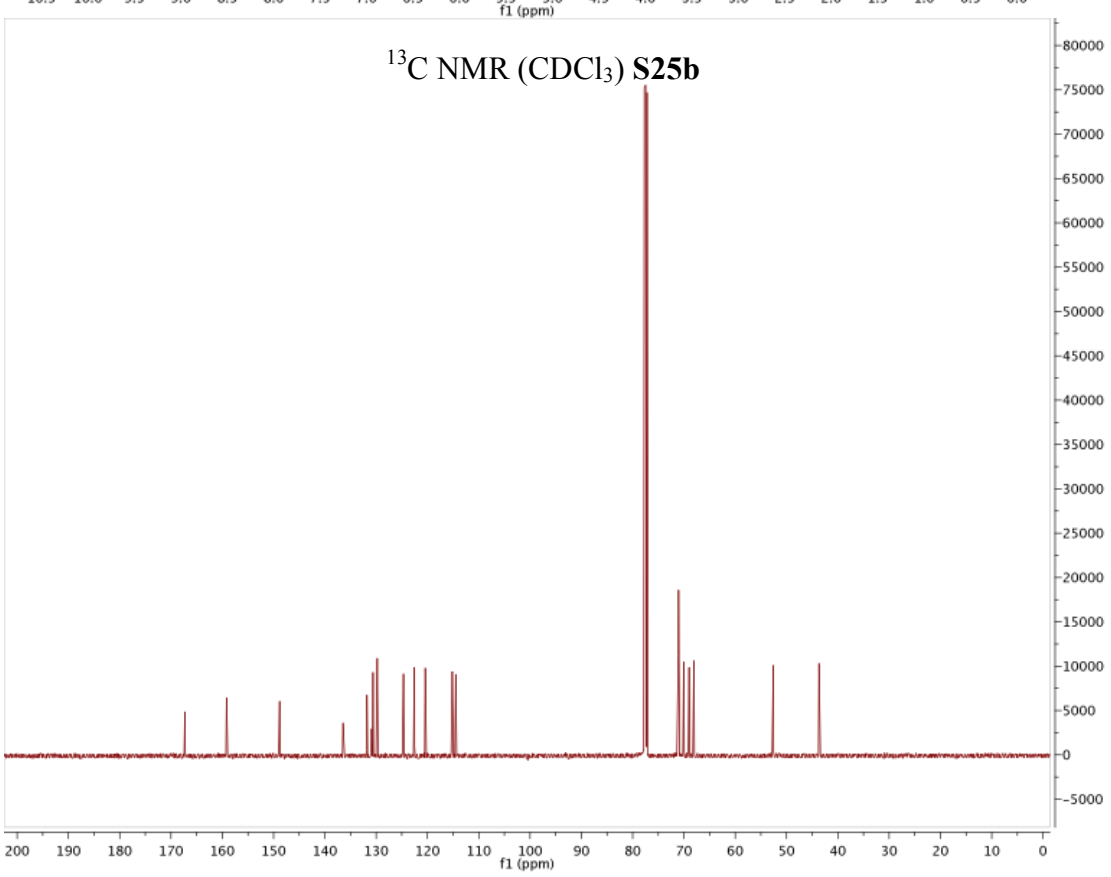
^{13}C NMR ($CDCl_3$) **S25a**



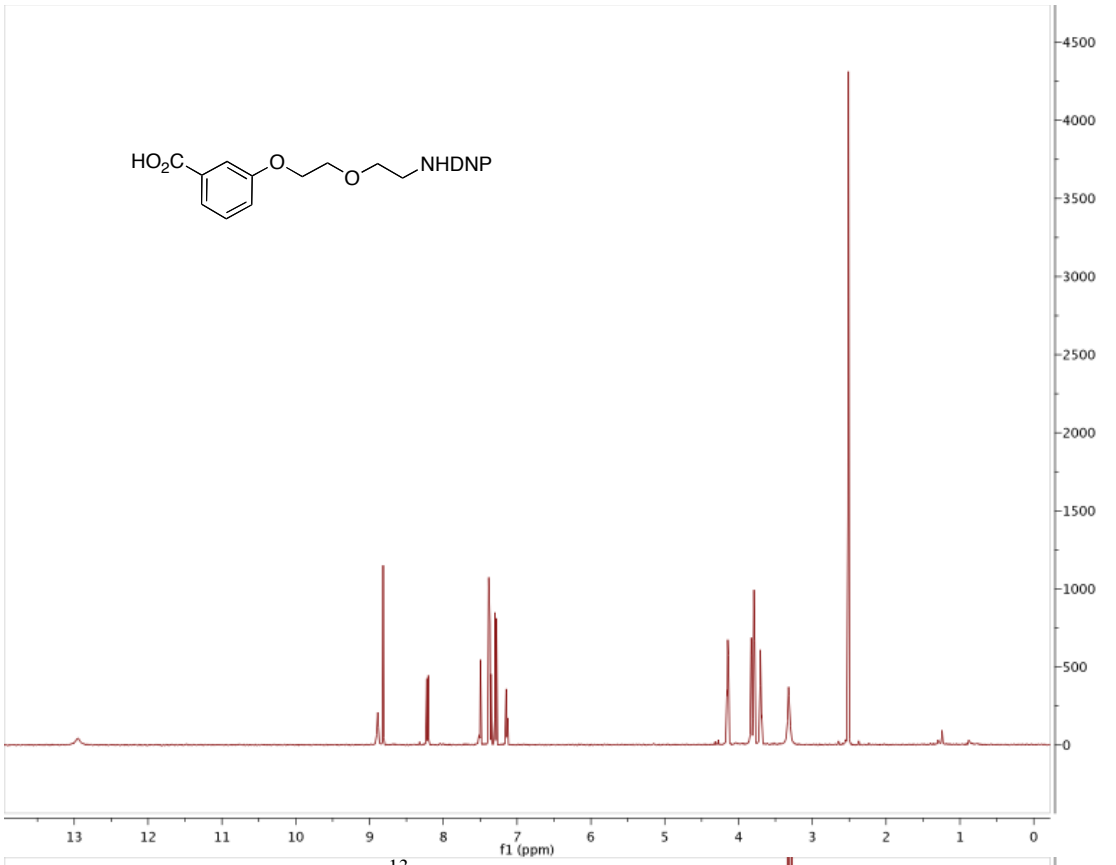
^{1}H NMR ($CDCl_3$) **S25b**



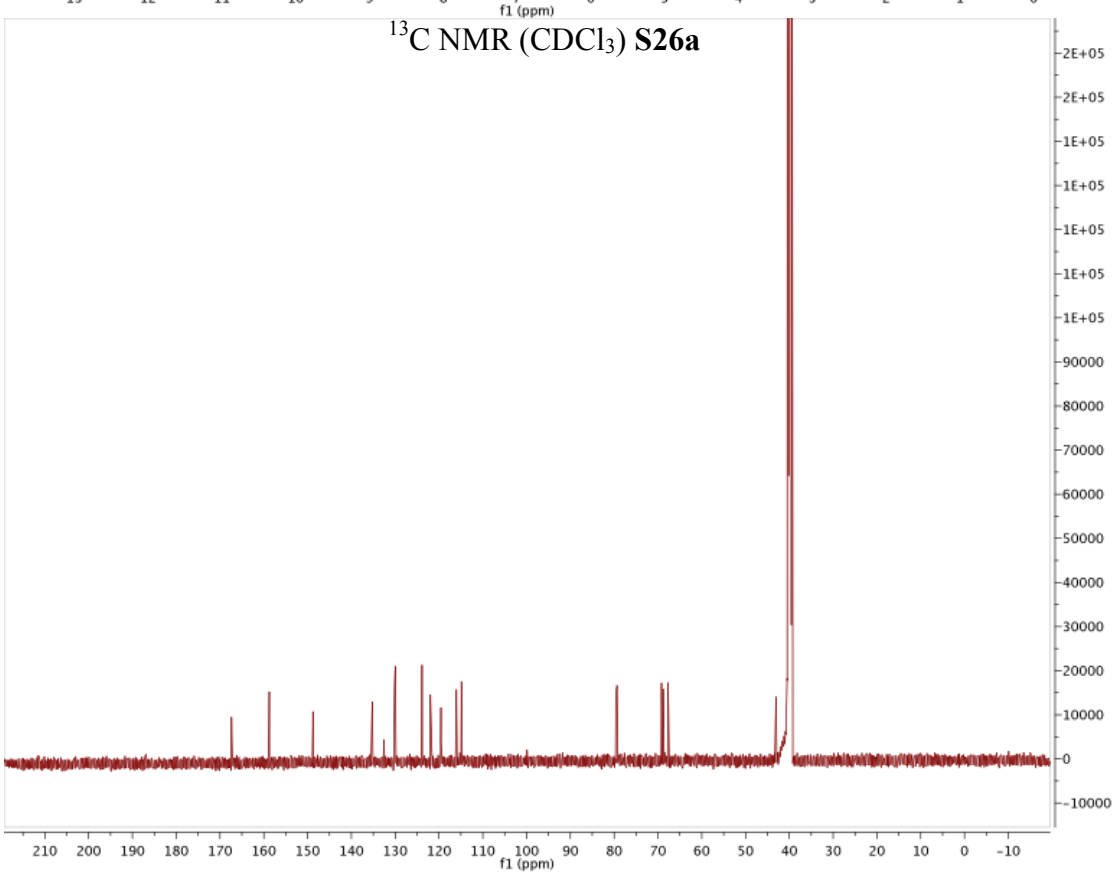
^{13}C NMR ($CDCl_3$) **S25b**



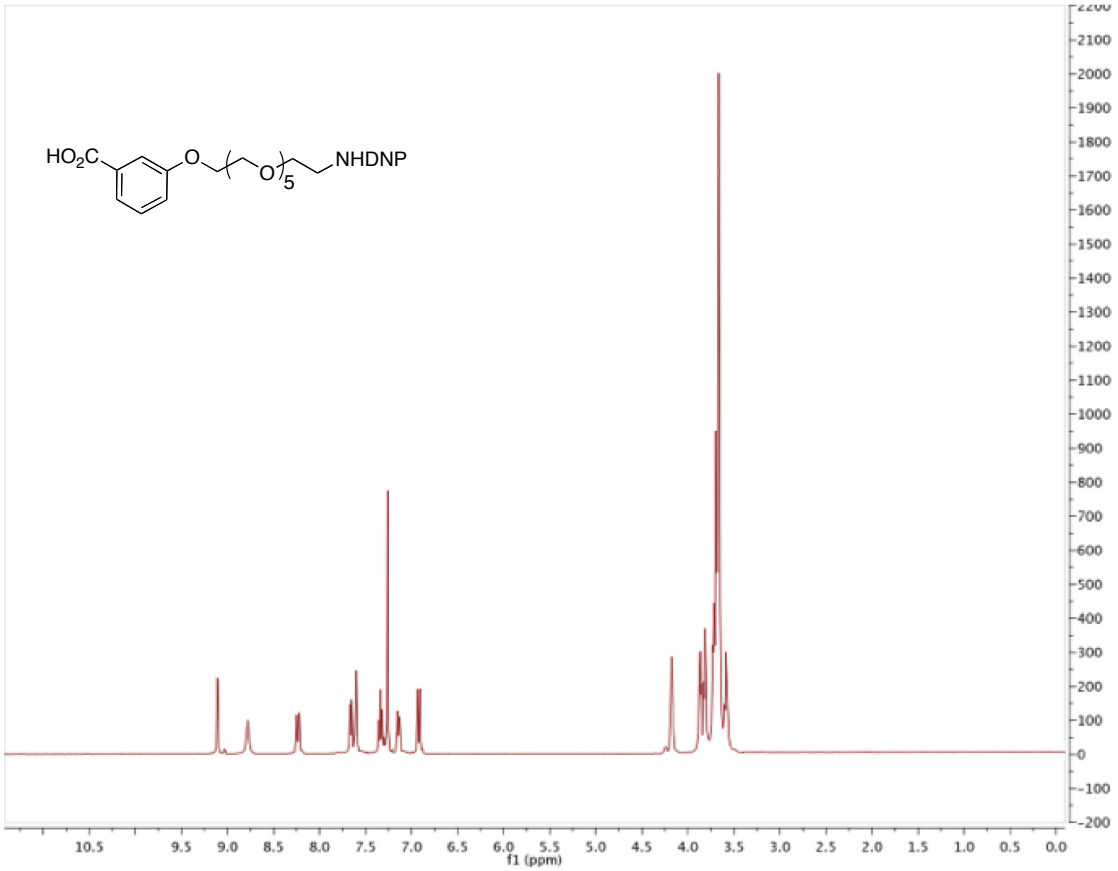
^{1}H NMR ($CDCl_3$) **S26a**



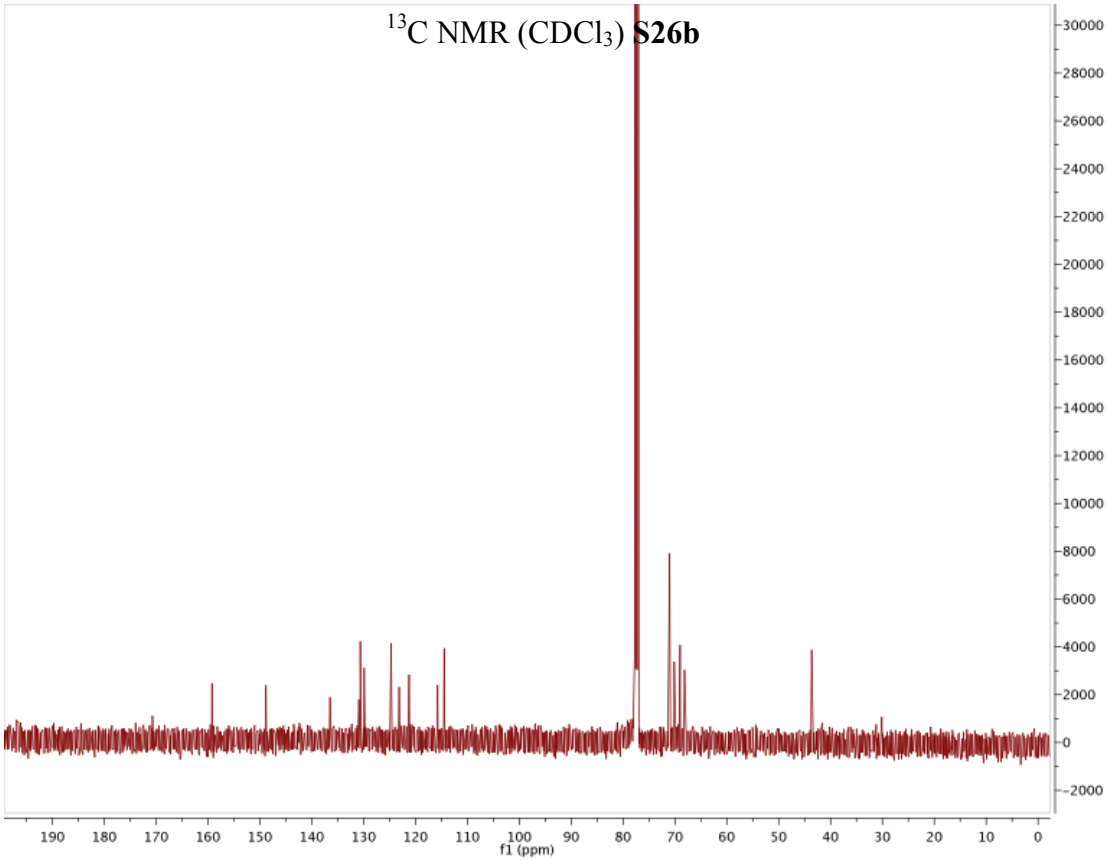
^{13}C NMR ($CDCl_3$) **S26a**



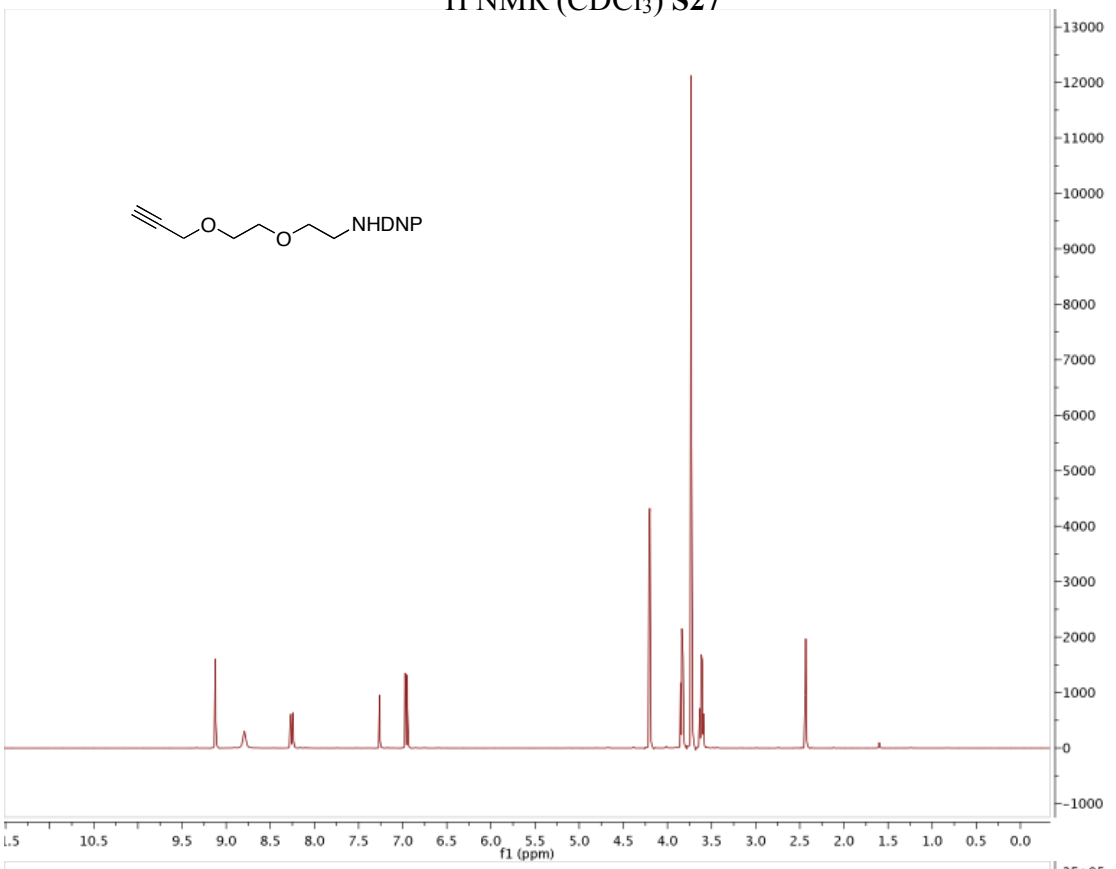
^{1}H NMR ($CDCl_3$) **S26b**



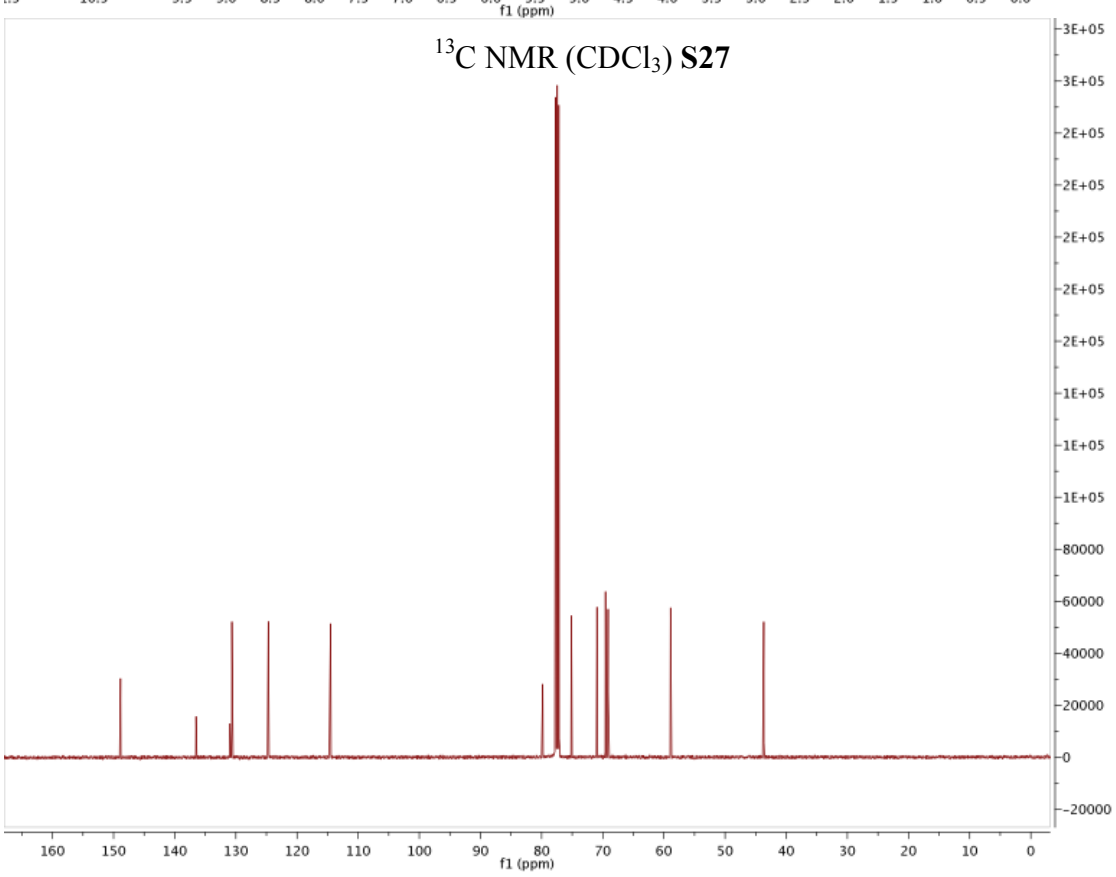
^{13}C NMR ($CDCl_3$) **S26b**



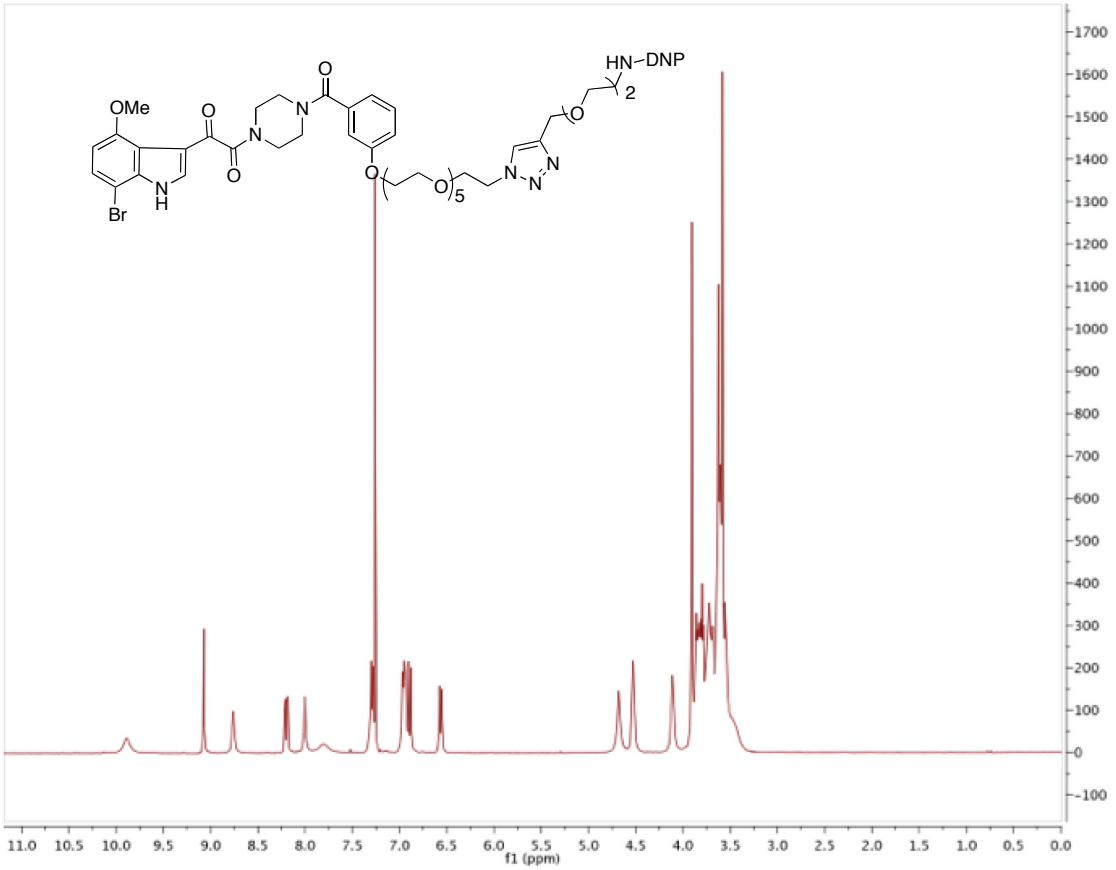
^{1}H NMR ($CDCl_3$) **S27**



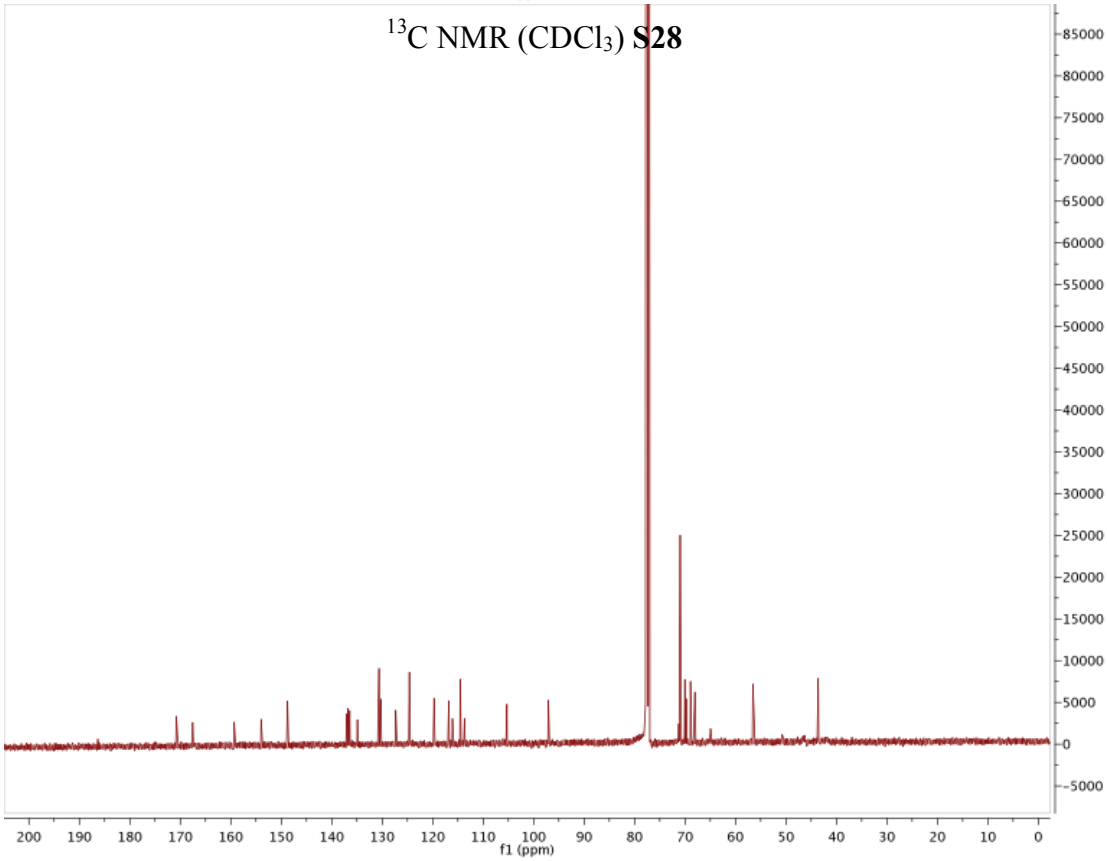
^{13}C NMR ($CDCl_3$) **S27**



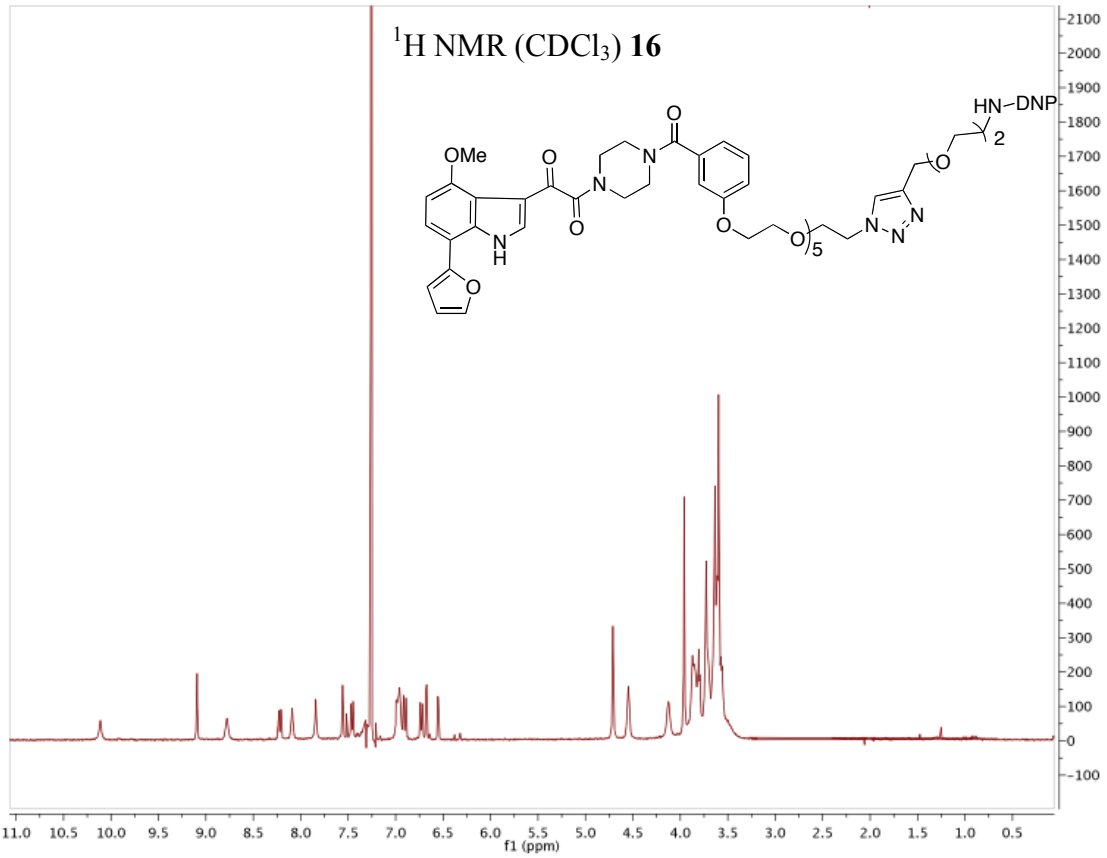
^{1}H NMR ($CDCl_3$) **S28**



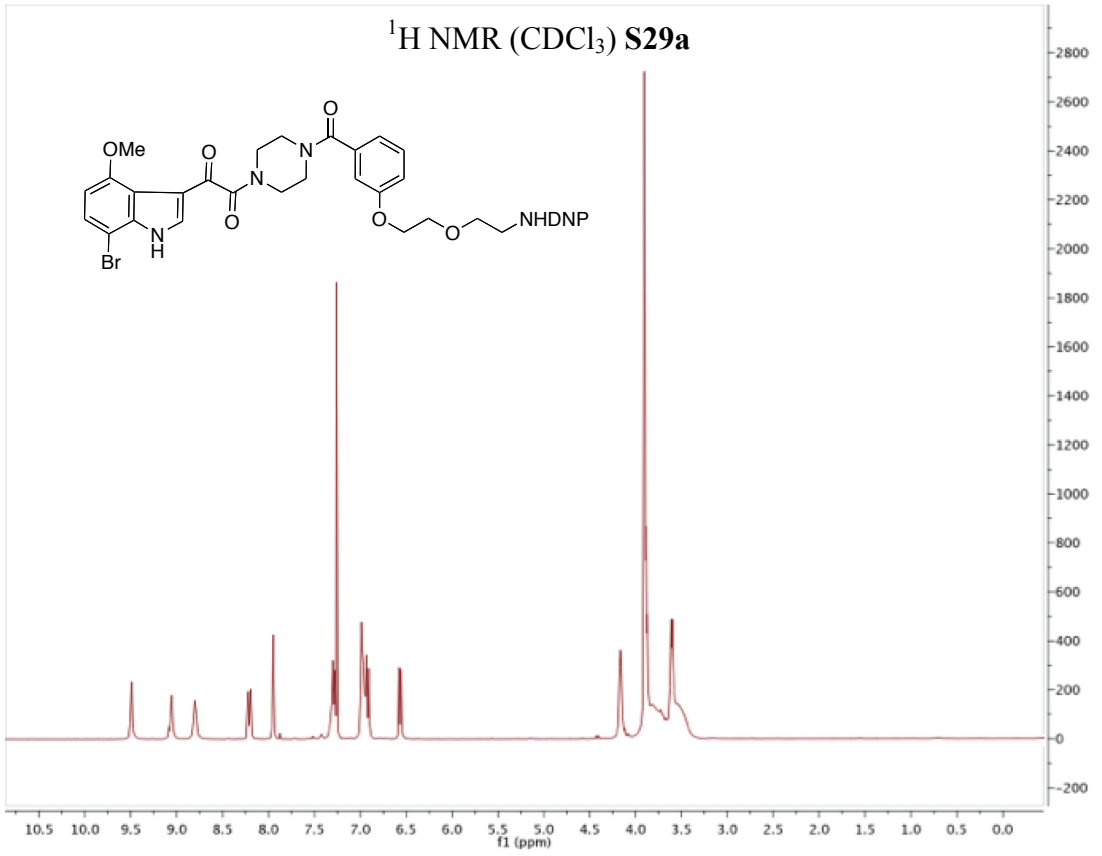
^{13}C NMR ($CDCl_3$) **S28**



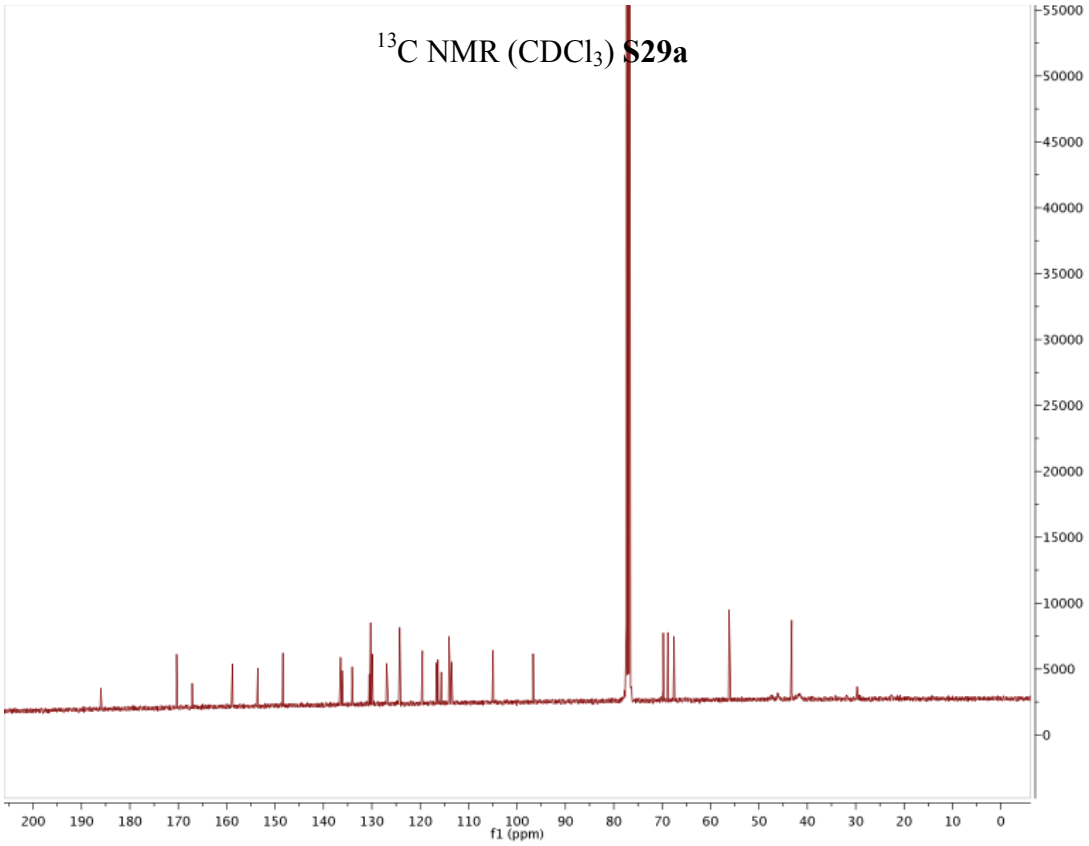
^{1}H NMR ($CDCl_3$) **16**



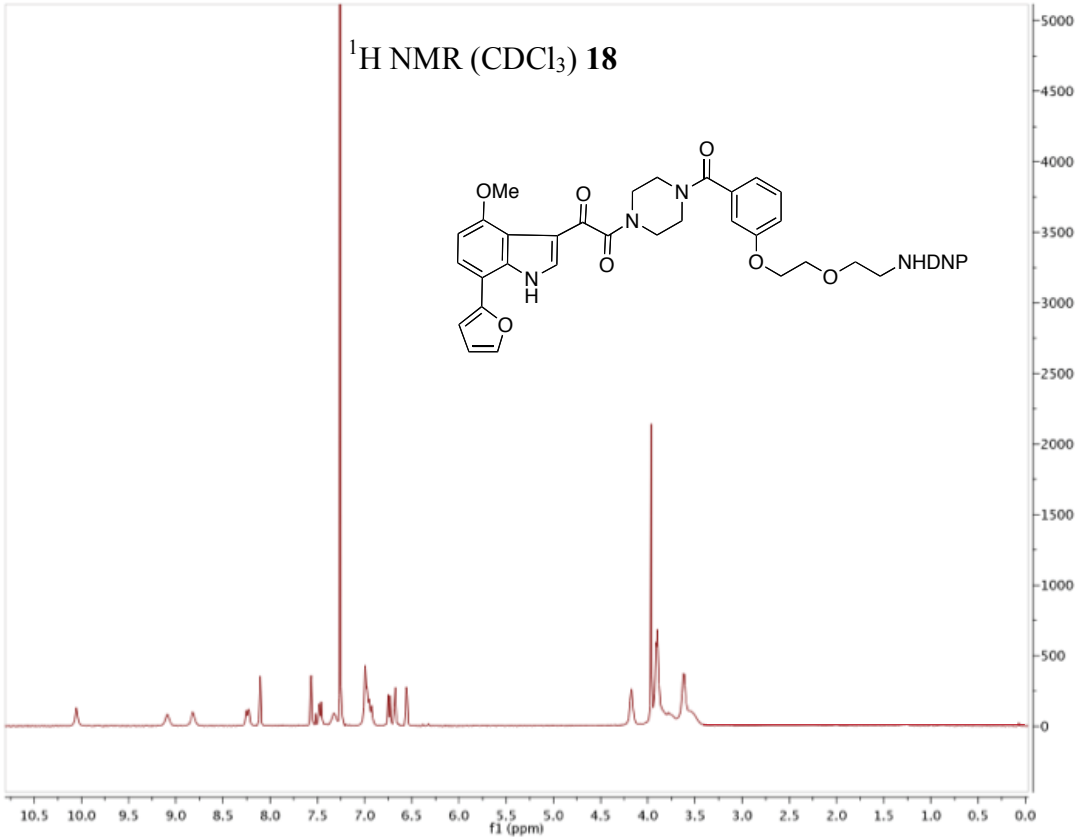
^{1}H NMR ($CDCl_3$) **S29a**



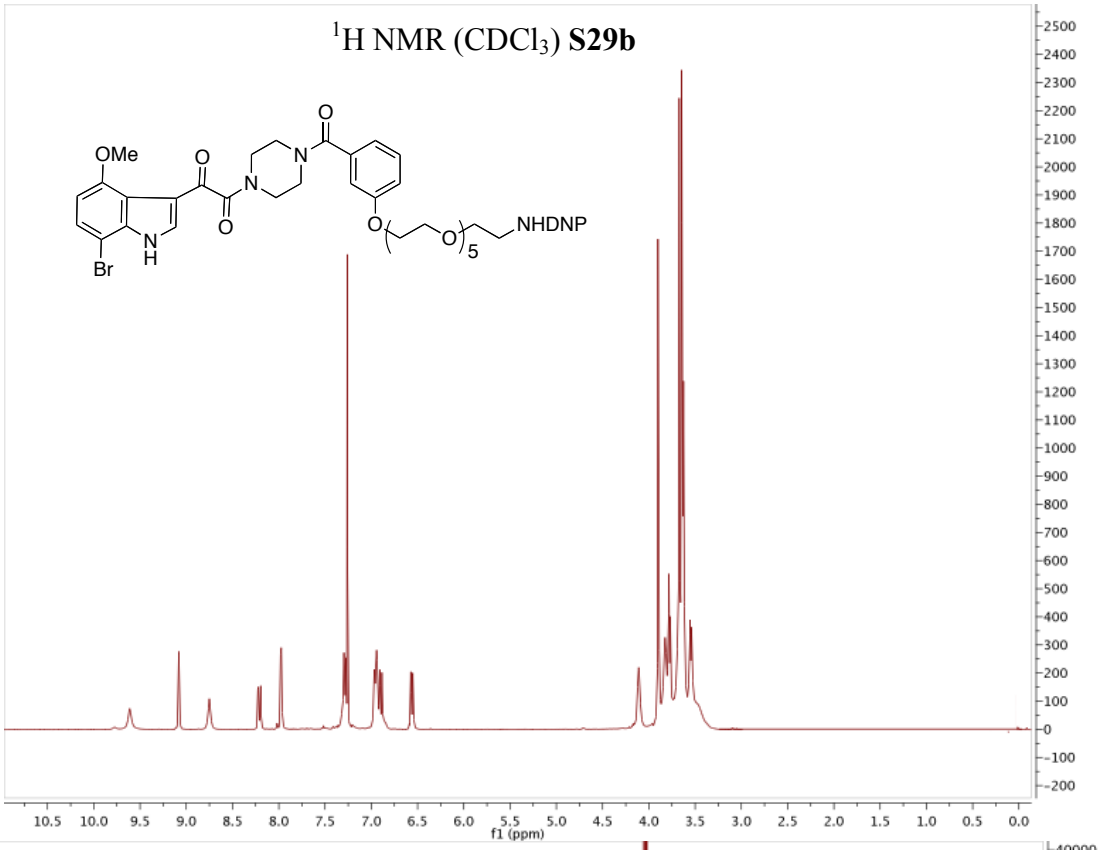
^{13}C NMR ($CDCl_3$) **S29a**



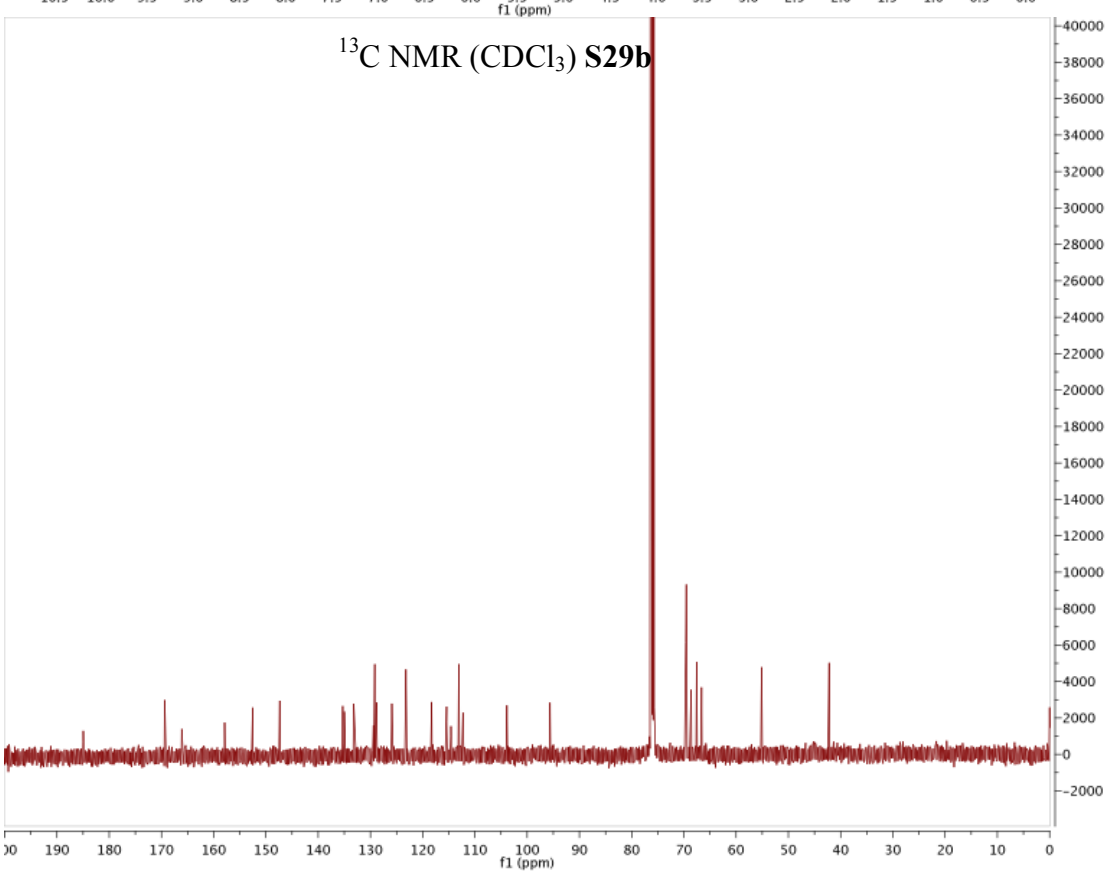
^{1}H NMR ($CDCl_3$) **18**



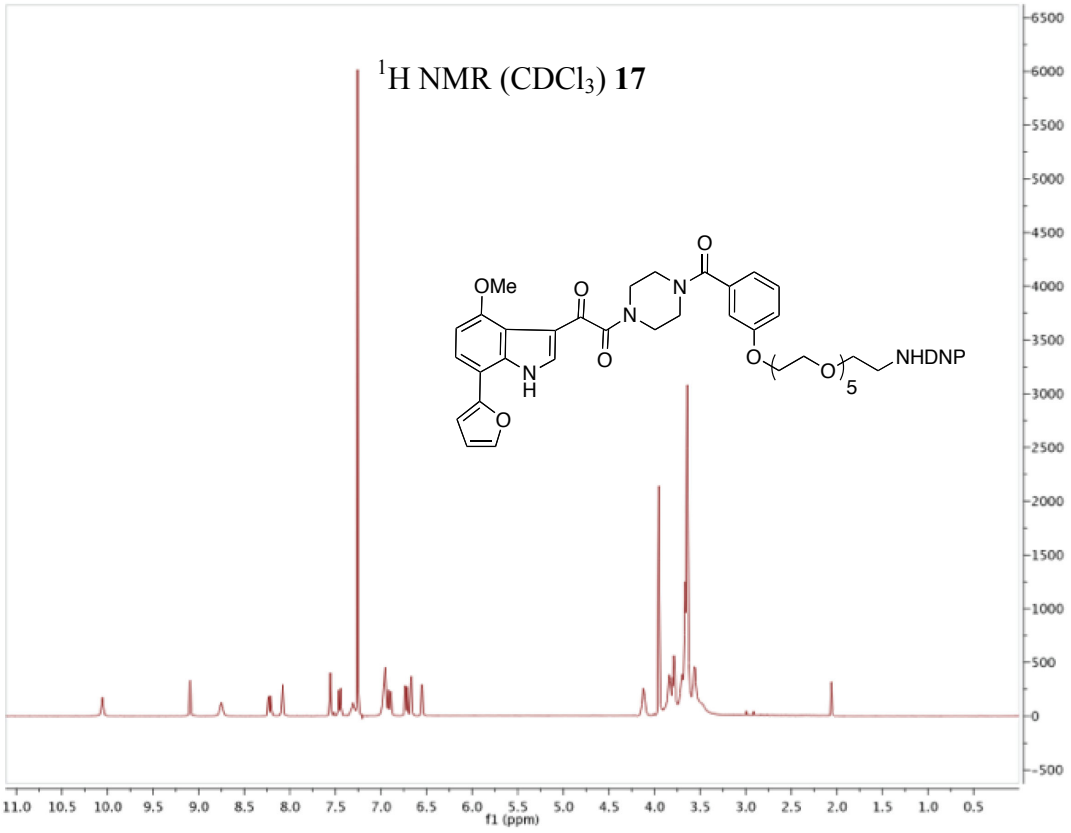
^{1}H NMR ($CDCl_3$) **S29b**



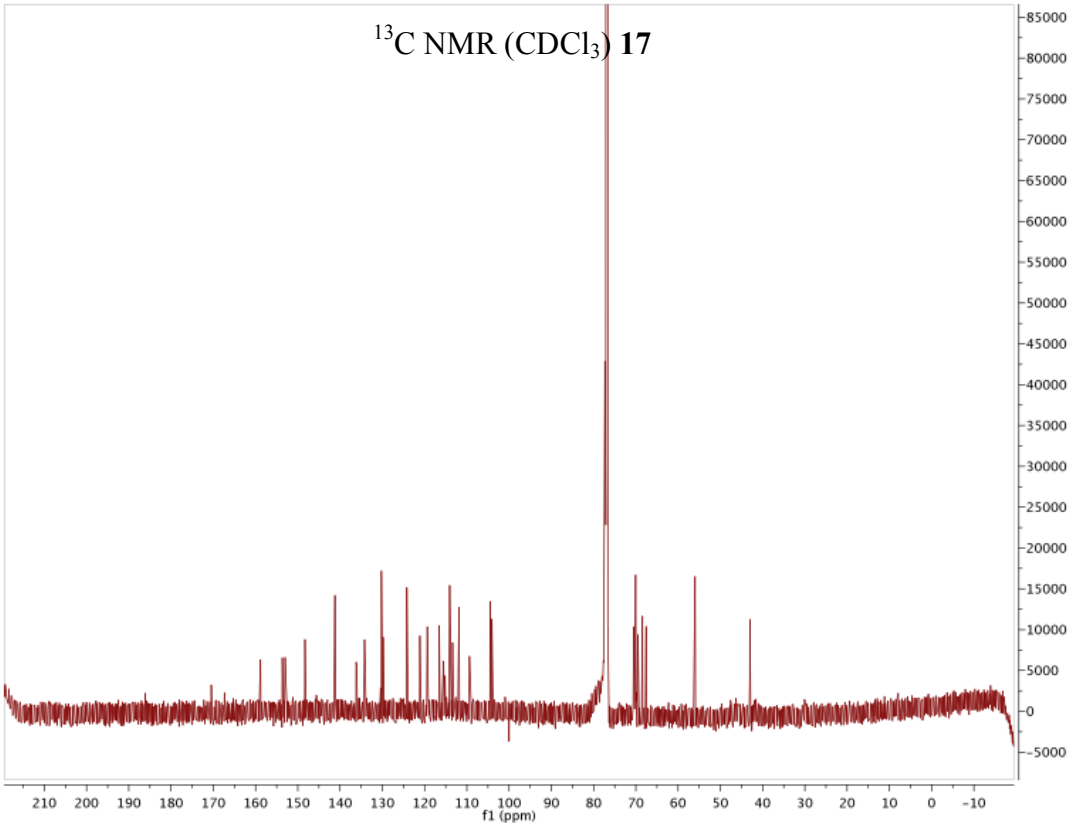
^{13}C NMR ($CDCl_3$) **S29b**



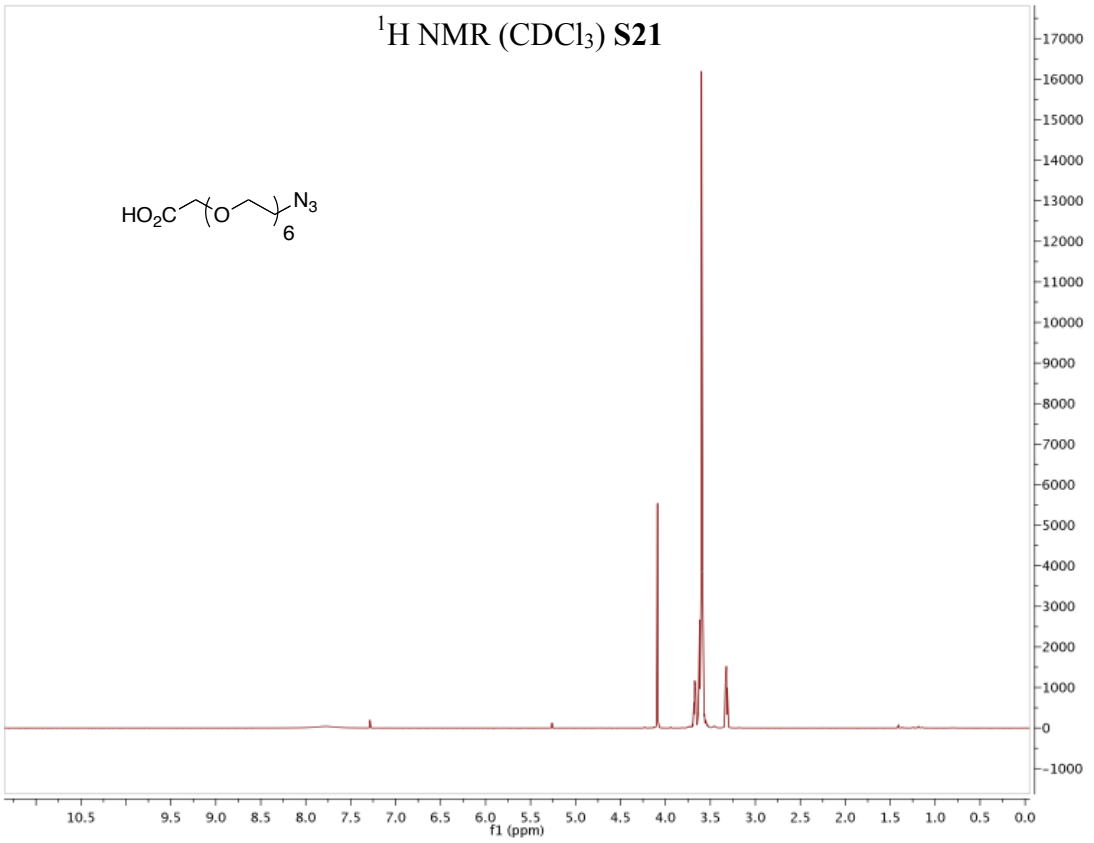
^{1}H NMR ($CDCl_3$) **17**



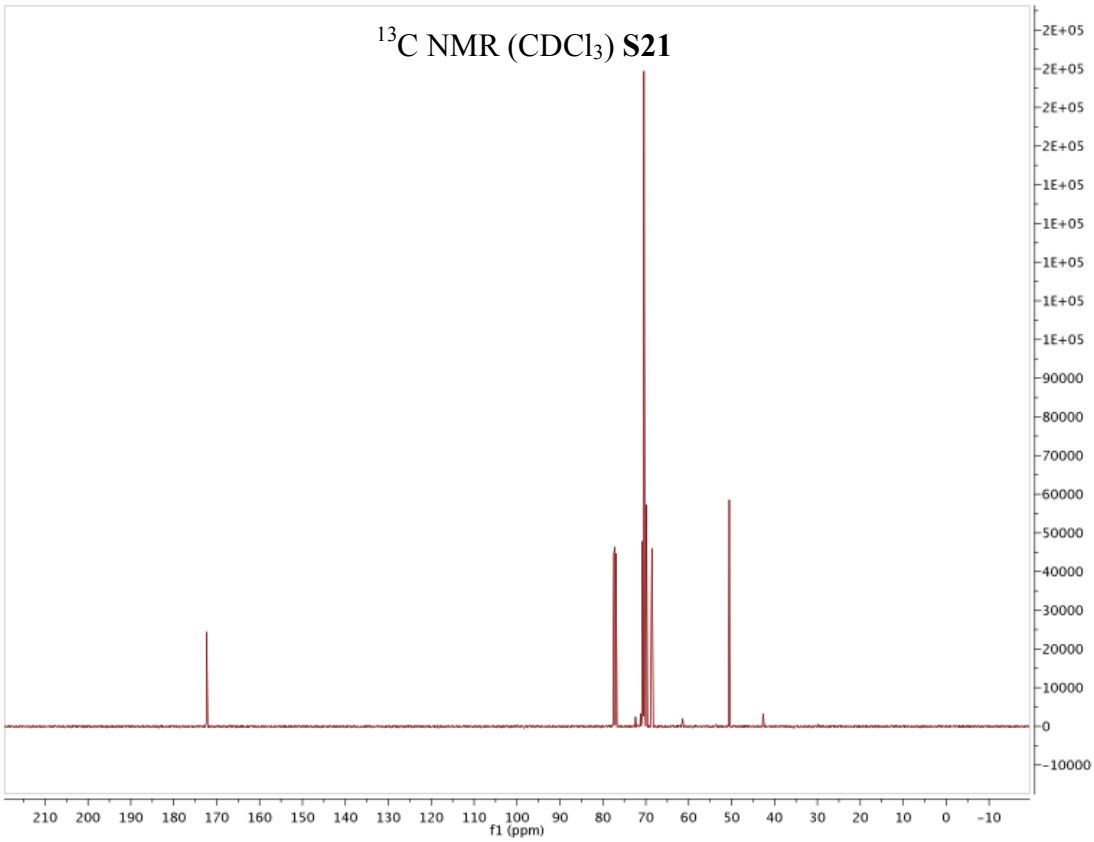
^{13}C NMR ($CDCl_3$) **17**



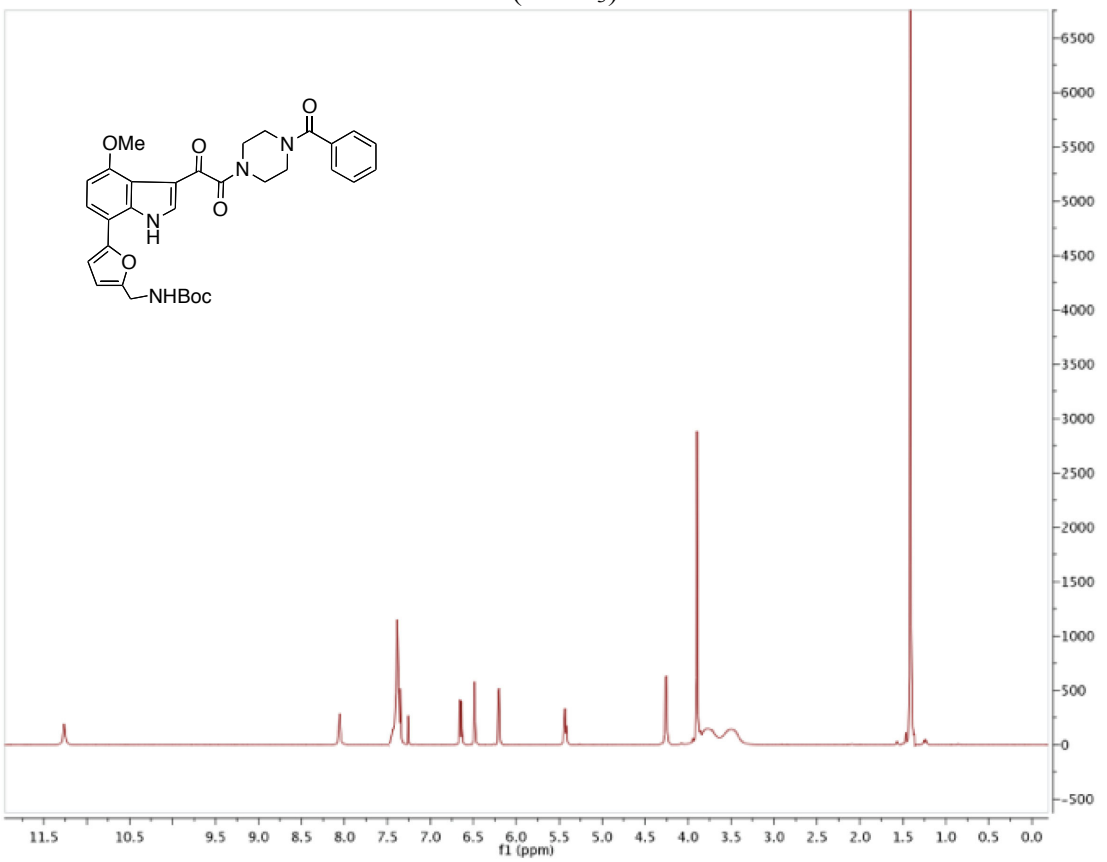
^{1}H NMR ($CDCl_3$) **S21**



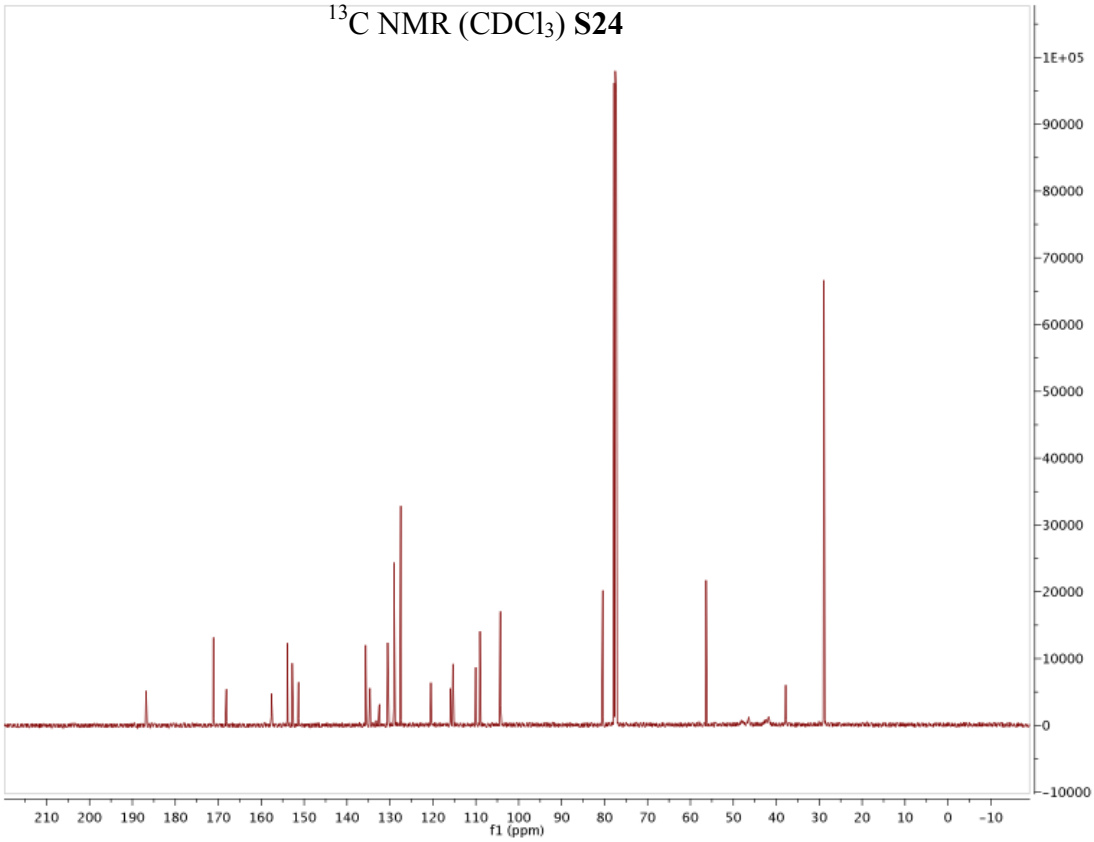
^{13}C NMR ($CDCl_3$) **S21**



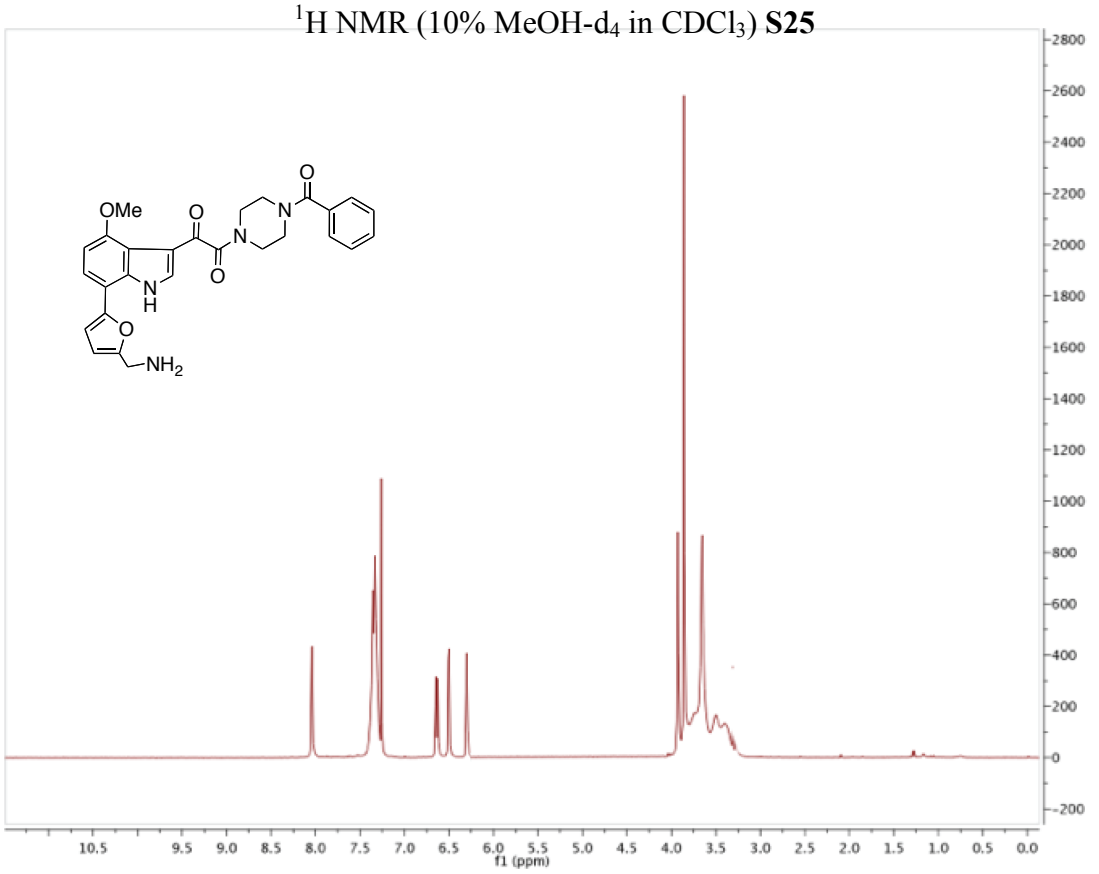
^{1}H NMR ($CDCl_3$) **S24**



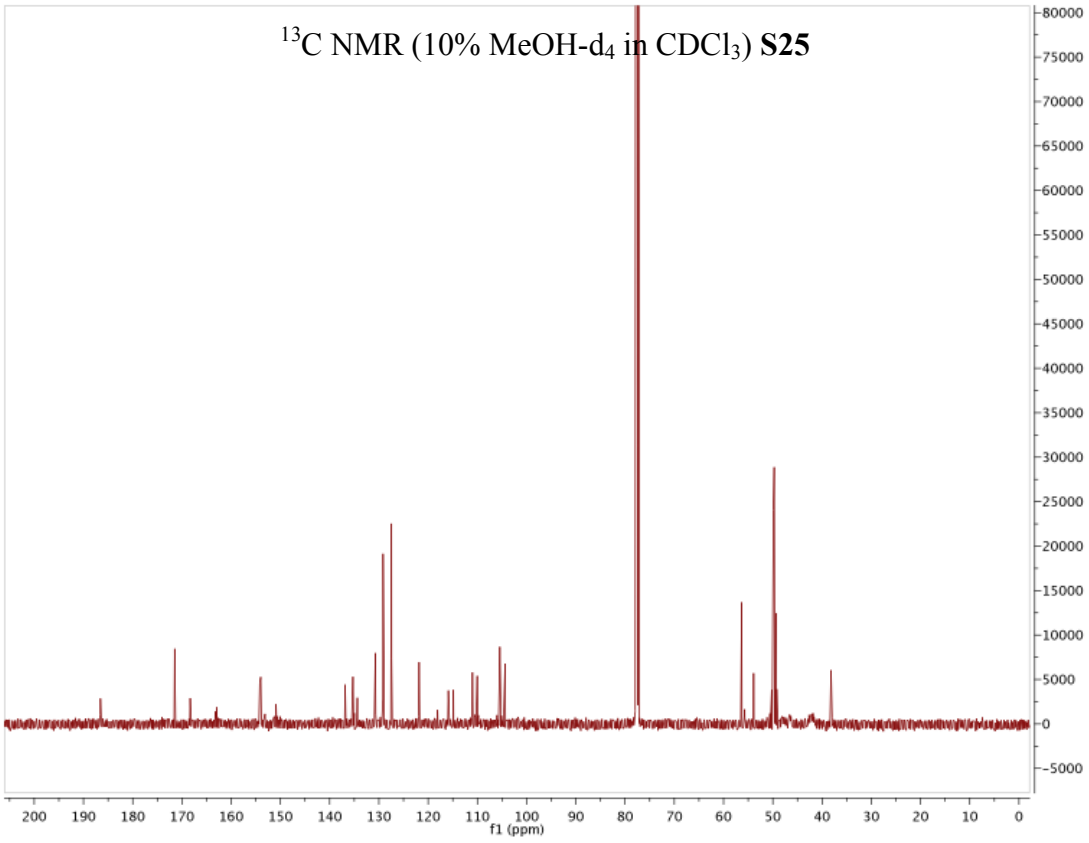
^{13}C NMR ($CDCl_3$) **S24**



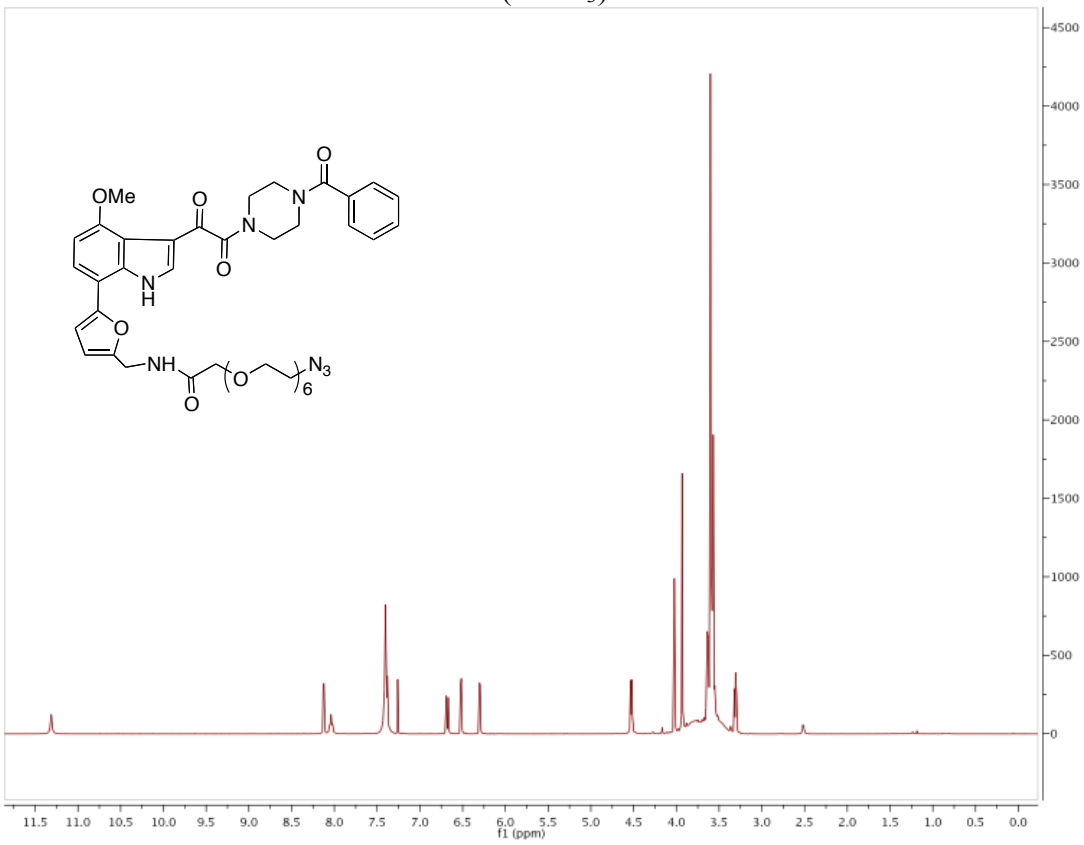
^{1}H NMR (10% MeOH-d_4 in $CDCl_3$) **S25**



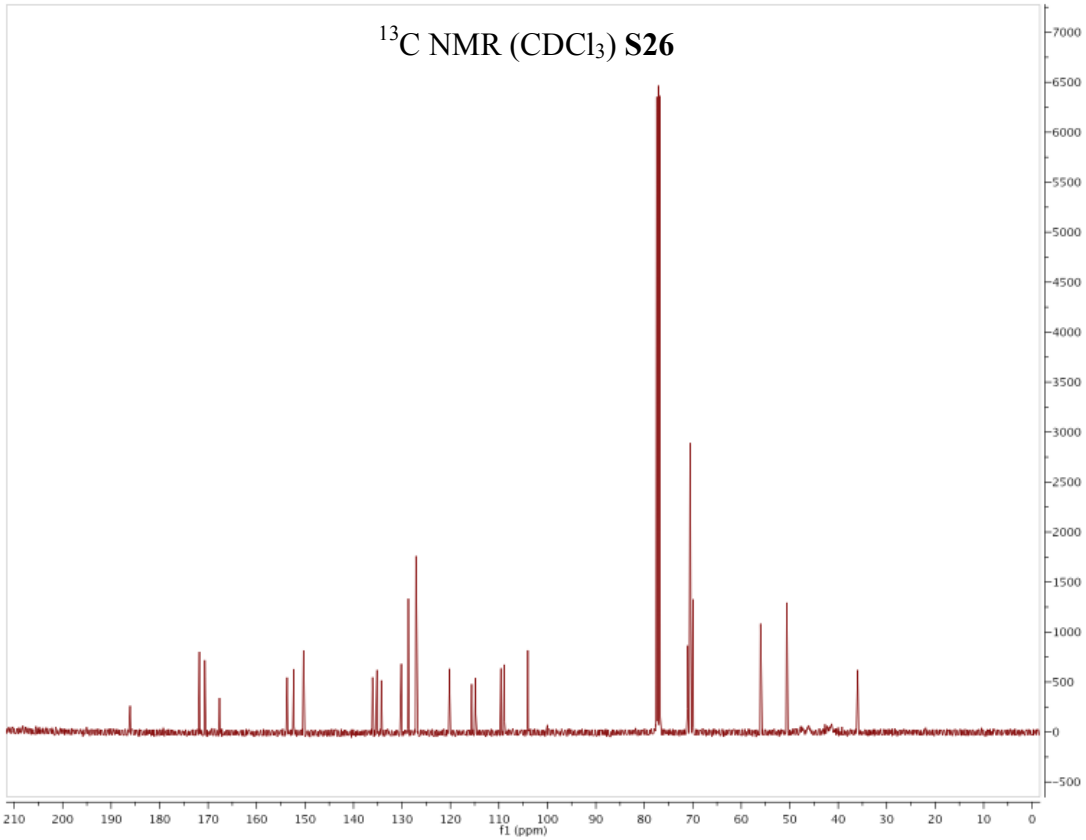
^{13}C NMR (10% MeOH-d_4 in $CDCl_3$) **S25**



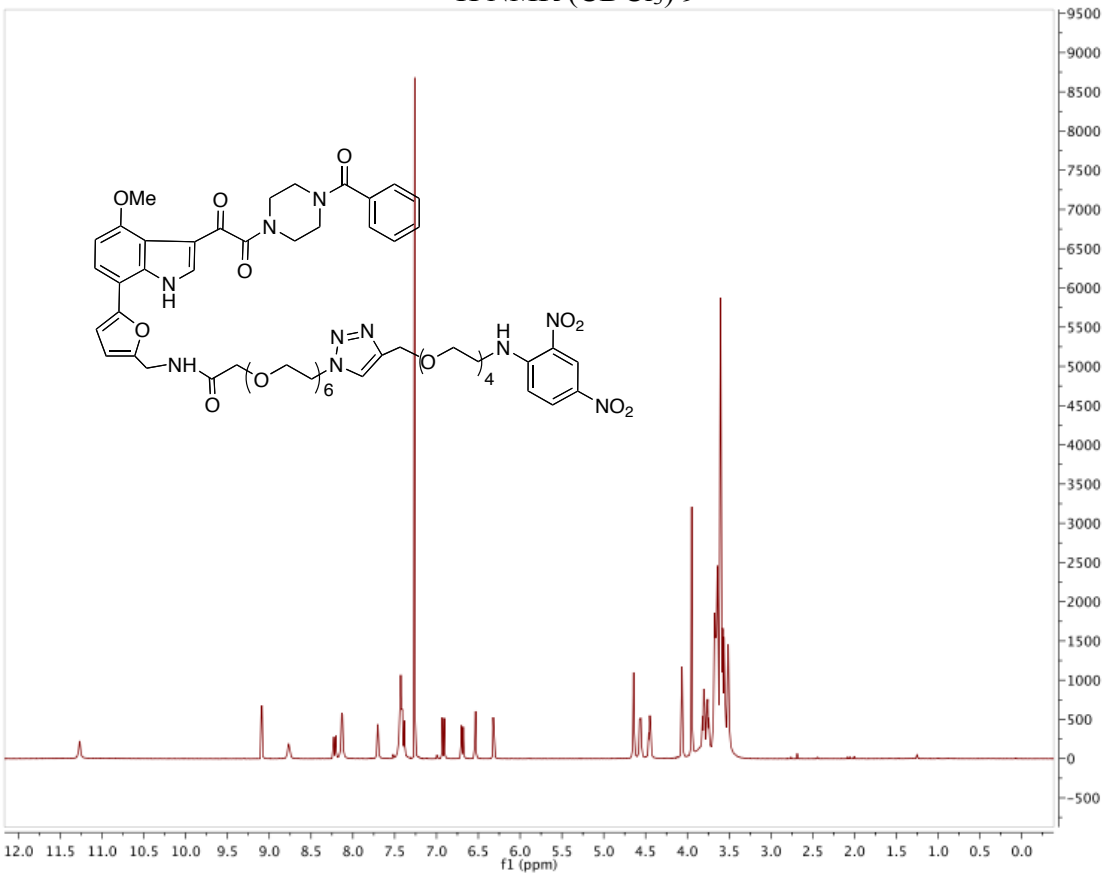
^{1}H NMR ($CDCl_3$) **S26**



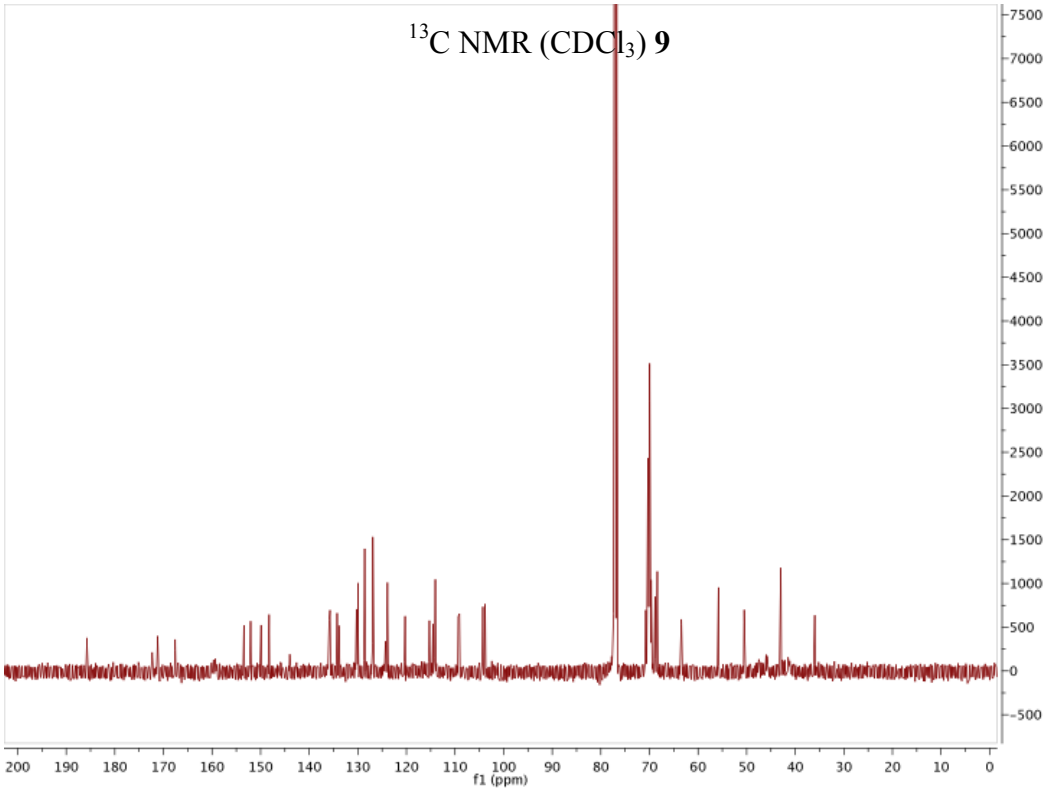
^{13}C NMR ($CDCl_3$) **S26**



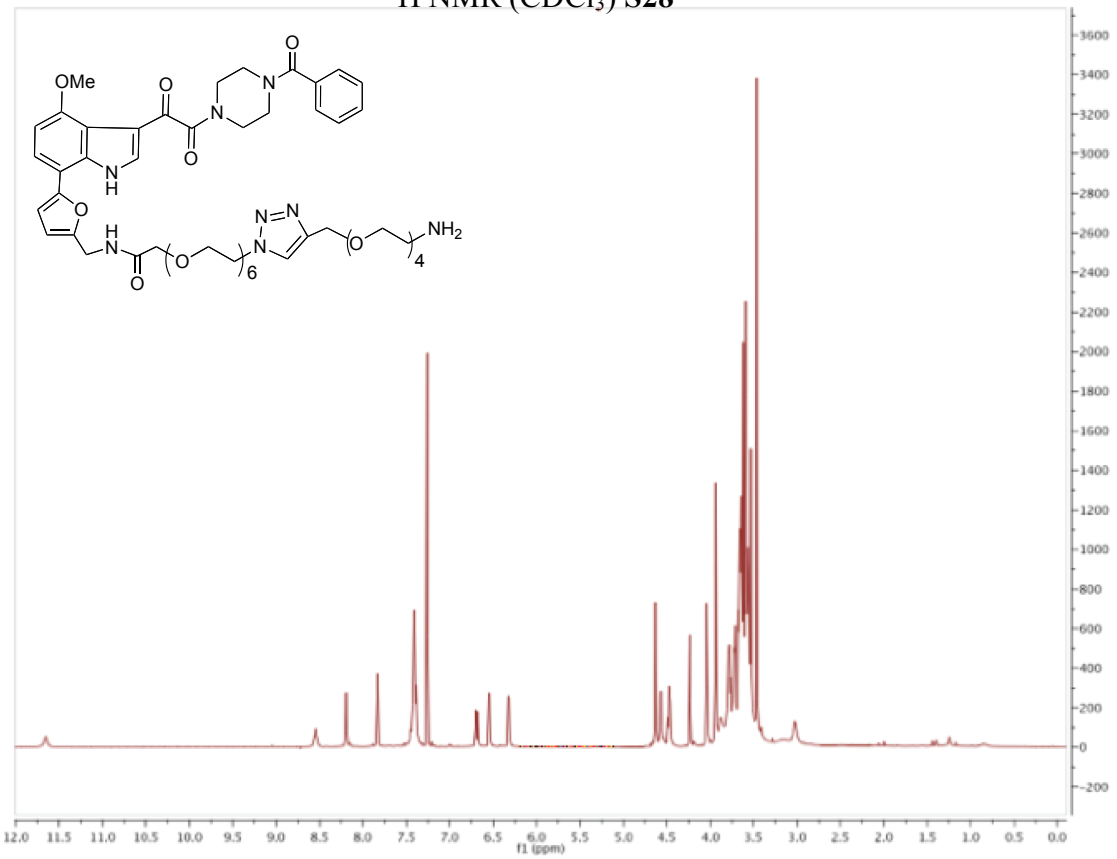
^{1}H NMR ($CDCl_3$) **9**



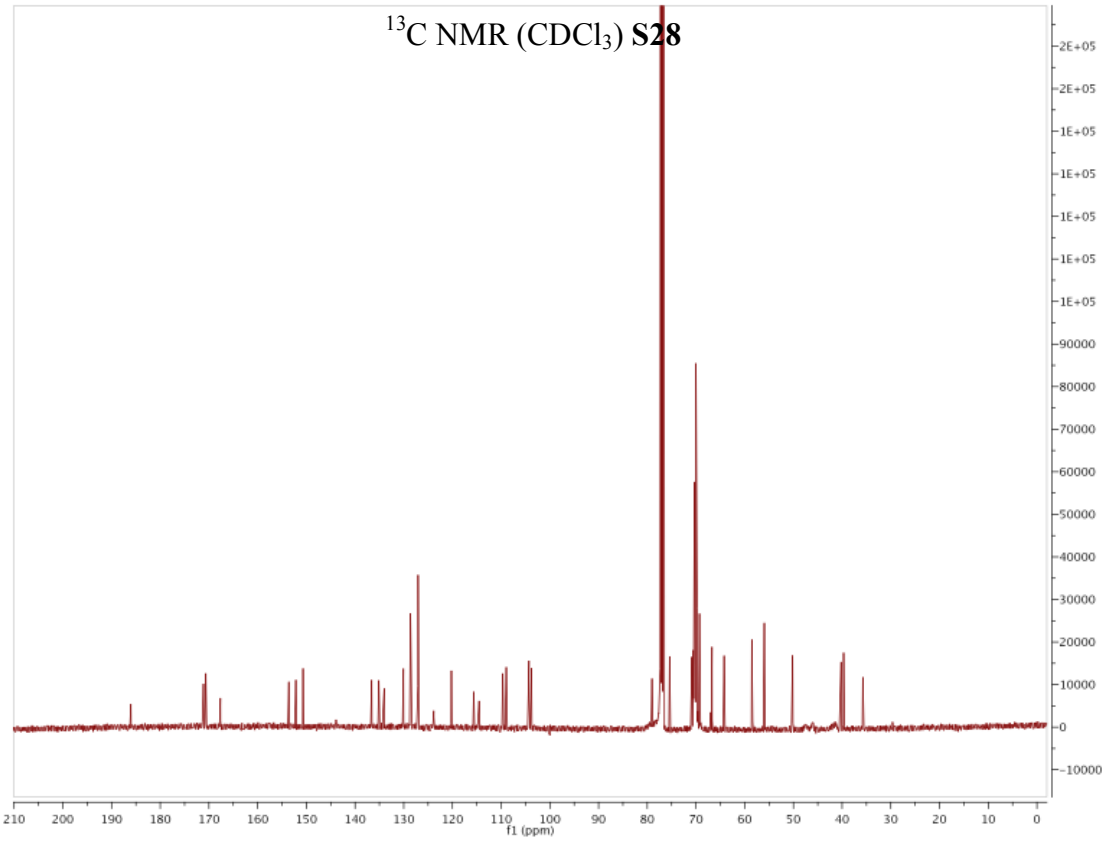
^{13}C NMR ($CDCl_3$) **9**



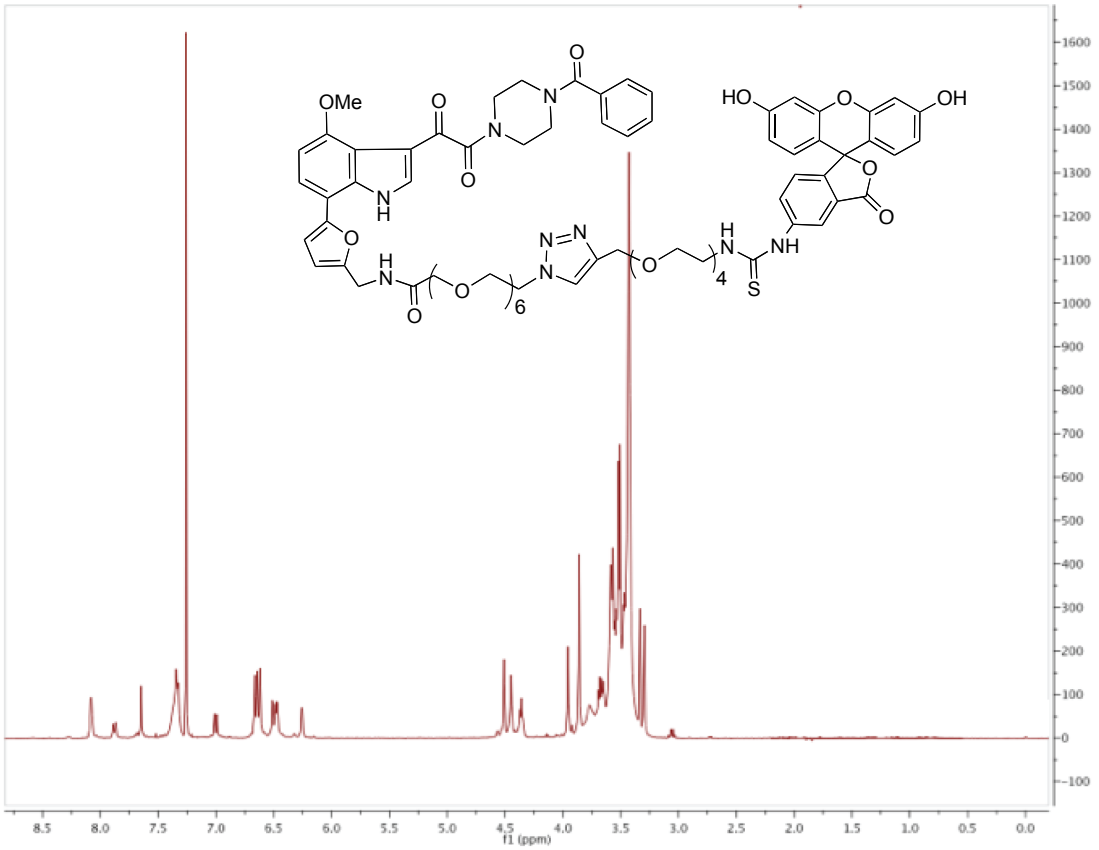
^{1}H NMR ($CDCl_3$) **S28**



^{13}C NMR ($CDCl_3$) **S28**



^{1}H NMR (5% MeOH-d_4 in $CDCl_3$) **S29**



^{13}C NMR (5% MeOH-d_4 in $CDCl_3$) **S29**