Electronic Supplementary Material (ESI) for Chemical Communications

SupplementaryInformation

Enhancethebindingaffinityoffluorophore-aptamerpairsfor RNAtaggingwithneomycinconjugation

Jongho Jeon[a‡], Kyung Hyun Lee[a‡] and Jianghong Rao*[a]

[a] *Molecular Imaging Program at Stanford (MIPS), Departments of Radiology and Chemistry, Stanford University , Stanford, California 94305-5484 USA.; E-mail: jrao@stanford.edu*

Fax: + 1 650 736 7925; Tel: + 1 650 736 8563

‡ These authors contributed equally to this work.

Contents

Fluorescence titration and equation for K_d and F_{max} values

All RNAs (Structure and sequences are shown in Fig. S4) were prepared as previously reported (J. Lee, K. H. Lee, J. Jeon, A. Dragulescu-Andrasi, F. Xiao and J. Rao, *ACS Chem. Biol.*, 2010, **5**, 1065–1074). Fluorescence titration was carried out in PBS buffer (pH7.4) containing 1 mM $MgCl_2$ using Horiba FluoroMax-3 fluorometer. Excitation and emission wavelength were 555 nm and 610 nm, respectively. Excitation and emission slit were 5 nm each.

$$(F_i) = \left(\frac{[ligand]F_{\max}}{K_d + [ligand]} \right) \quad \text{(Equation 1)}$$

$$(F_i)_{Total} = (F_{i1})_{site1} + (F_{i2})_{site2} \quad \text{(Equation 2)}$$

$$(F_i)_{Total} = \left(\frac{[ligand]F_{\max 1}}{K_{d1} + [ligand]} \right)_{site1} + \left(\frac{[ligand]F_{\max 2}}{K_{d2} + [ligand]} \right)_{site2} \quad \text{(Equation 3)}$$

For the calculation of dissociation constants (K_d), one-site and two-site binding models were employed for non-linear fitting using Graphpad Prism Software. One-site binding models (Equation 1) have been previously employed for fluorescence measurements (A. Shoji, M. Kuwahara, H. Ozaki and H. Sawai, *J. Am. Chem. Soc.*, 2007, **129**, 1456-1464.). F_i is the relative fluorescence intensity at a given ligand concentration ([*ligand*], RNA aptamer at μM) and the total possible fluorescence intensity at saturation of a given binding site is F_{max}. However, for models containing two different binding sites with different associated F_{max} and K_d values, the total relative fluorescence intensity due to binding becomes the sum of the F_i of each specific type of binding site (Equation 2). Therefore, the total fractional binding obtainable through fluorescence titration (Equation 3) was obtained by substitution of Equation 1 into Equation 2, allowing for the evaluation of K_d using a two-site binding model.

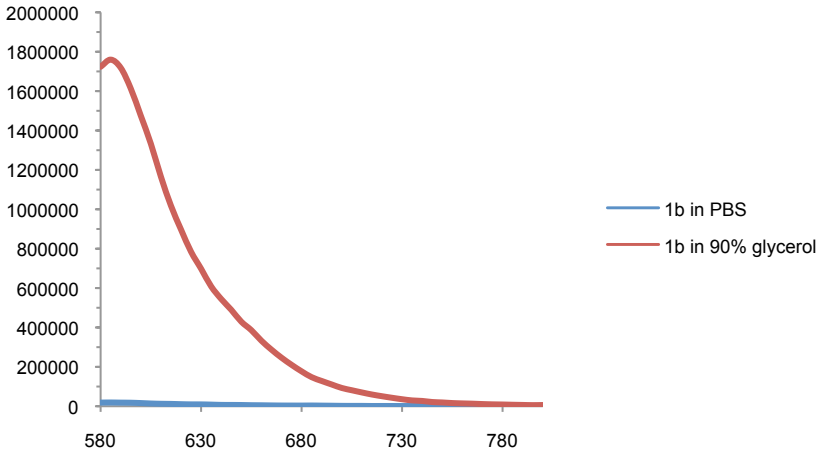


Figure S1. Fluorescence emission spectra of the compound 1b (1 μM) in PBS pH = 7.4 (blue) and 90% glycerol (red). In 90% glycerol aqueous solution, **1b** showed ca. 90-fold enhanced fluorescence intensity. The compound **1a** also displayed similar fluorescent enhancement (ca. 85-fold) in 90% glycerol solution.

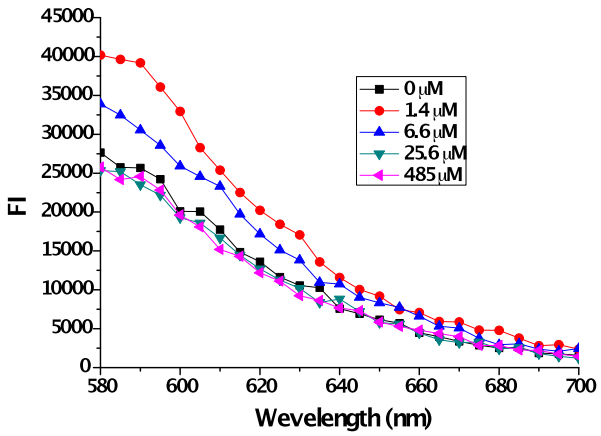


Figure S2. Fluorescence intensity change of ASR 4 in the presence of neomycin. Fluorescence titration was carried out in PBS buffer (pH7.4) containing 1 mM $MgCl_2$ using Horiba FluoroMax-3 fluorometer. Increasing concentrations of neomycin were added in the presence of 1 μM of ASR **4**. Excitation and emission wavelength were 555 nm and 610 nm, respectively. Excitation and emission slit were 5 nm each.

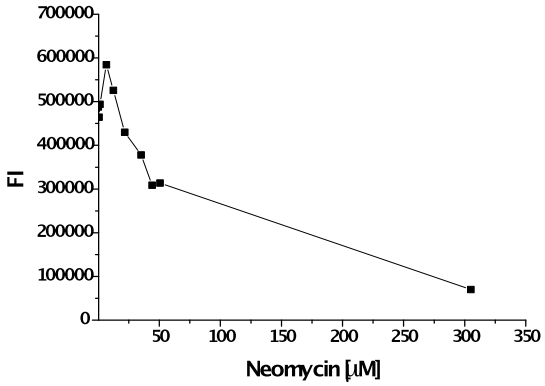


Figure S3. Competition assay with neomycin for ASR 4 bound to Apt10L. Fluorescence titration was carried out in PBS buffer (pH7.4) containing 1 mM $MgCl_2$ using Horiba FluoroMax-3 fluorometer. Increasing concentrations of neomycin were added in the mixture of ASR **4** and Apt10L RNA (1 μM of ASR **4** and 10 μM of Apt10L RNA). Excitation and emission wavelength were 555 nm and 610 nm, respectively. Excitation and emission slit were 5 nm each.

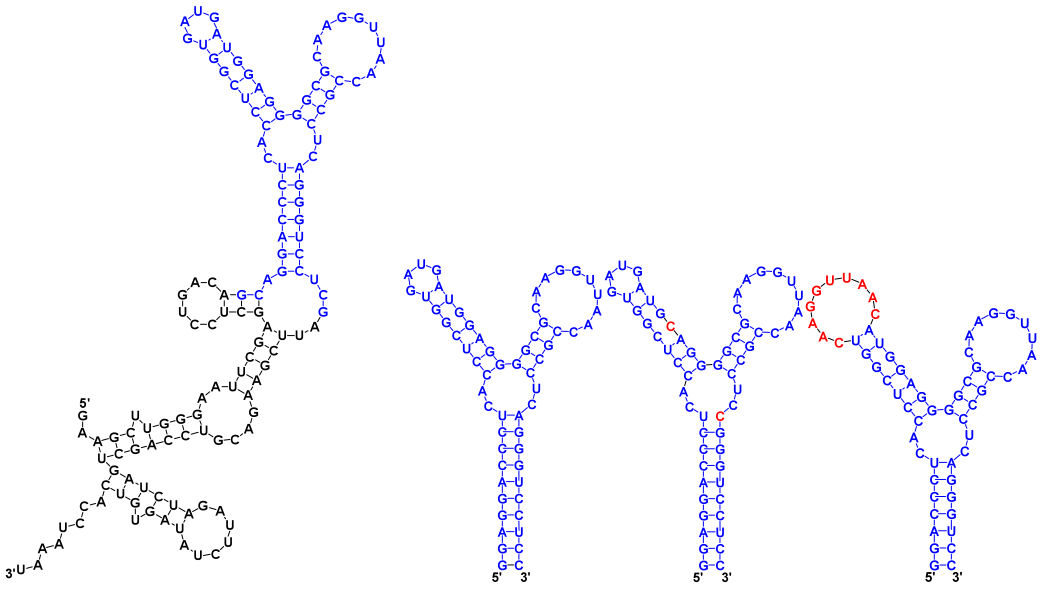
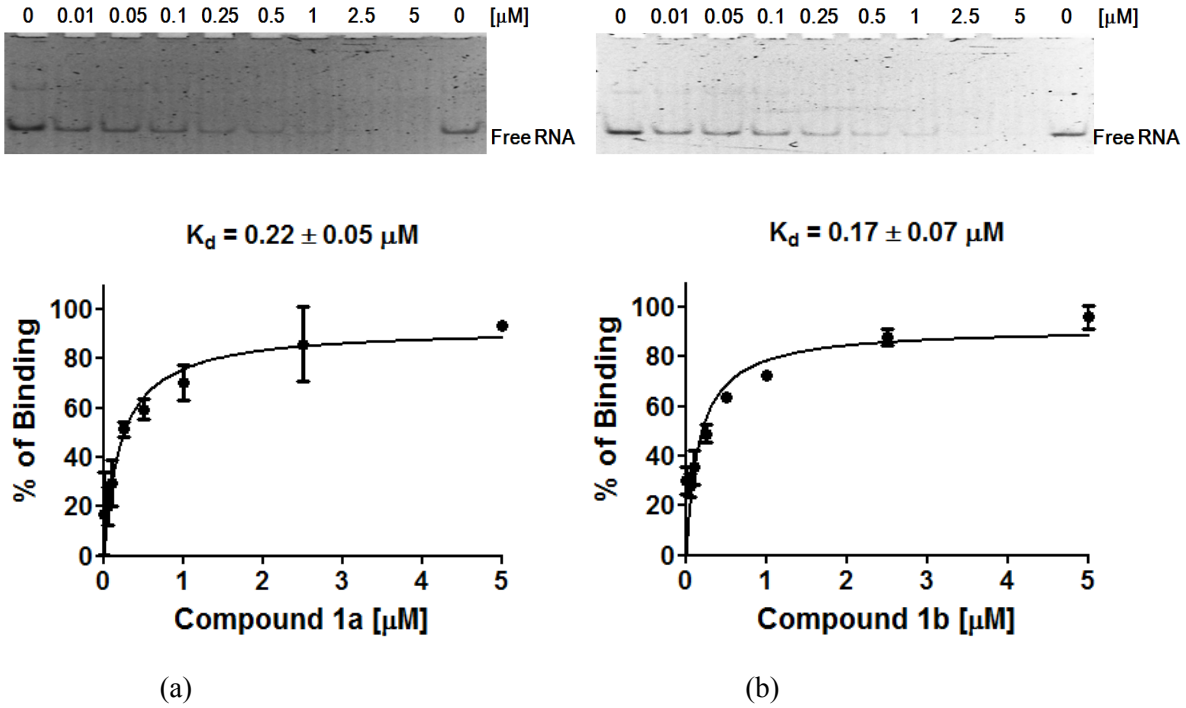
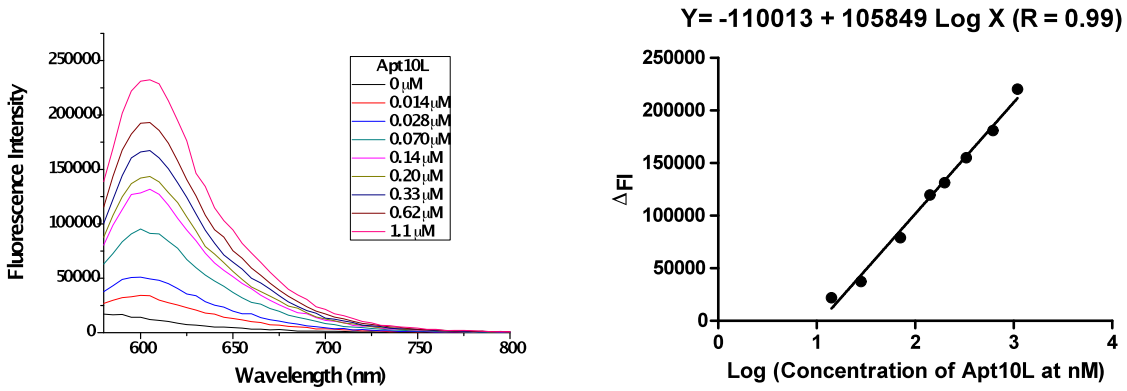


Figure S4. Predicted secondary structures and sequences of Apt10 variants and mutants. Structures of Apt10L **(a)**, Apt10M **(b)**, Apt10M1 **(c)**, and Apt10M-Lm3 **(d)** are shown as previously described (J. Lee, K. H. Lee, J. Jeon, A. Dragulescu-Andrasi, F. Xiao and J. Rao, *ACS Chem. Biol.*, 2010, **5**, 1065–1074). Secondary structures were predicted by Michael Zuker's M-fold program. Representative structures are shown and usually additional 1~3 structures are predicted with different energy level. Red sequences stand for the substituted sequences. Blue sequences represent the region of randomized 60 nucleotides.

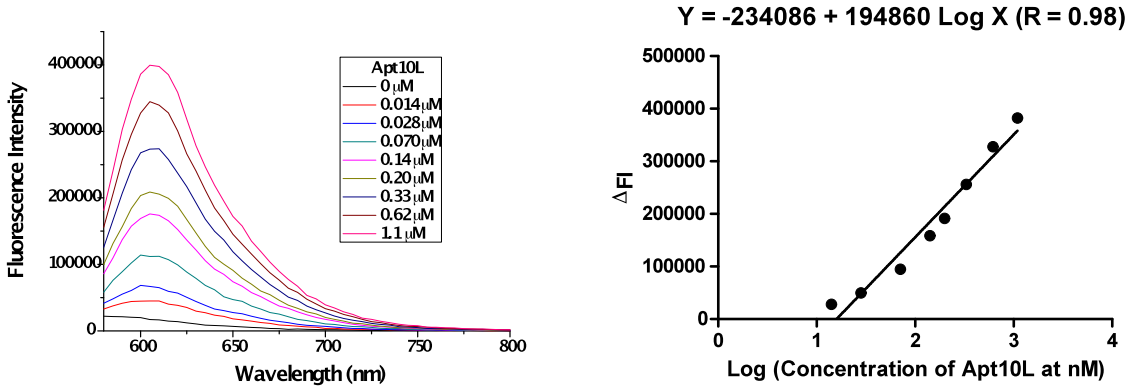


(a) (b)

Figure S5. Gel mobility shift assay for K_d calculation. All binding reaction samples were performed on a 10 μL scale with final concentrations of Apt10M RNA at 50 nM and various concentrations of compound **1a** or **1b** in binding buffer (0.01 M PBS with 1 mM $MgCl_2$ at pH 7.4). Apt10M RNA was heated to 65 °C for 5 min and slowly cooled to room temperature over 1 h. The RNA solution and compound **1a** or **1b** were mixed in binding buffer to make final concentration of 50 nM of Apt10M and indicated concentrations of **1a** and **1b** and incubated for 15 min at room temperature. To the resulting sample solutions was added the loading buffer (2.5 μL, 0.14 nM bromophenol blue, 0.19 nM xylene cyanol FF, 30% glycerol in 0.5x Tris-Borate and EDTA (TBE) buffer). Each sample (12.5 μL) was loaded on the 10% native polyacrylamide gel (1 mM $MgCl_2$, 1% glycerol) and run at 120V at 4 °C for 2 hr. The gel was stained with SYBR® Green EMSA nucleic acid gel stain solution (Invitrogen) using manufacturer's protocol and individual bands were quantified using ImageJ software. Neomycin conjugates are highly positively charged molecules so bound fractions could not migrate well on the gel. Therefore, the bound fraction was calculated from amount of remaining free RNA aptamer, and the K_d value was calculated by non-linear fitting using Graphpad Prism Software, as described in Page S-3. Shown are representative gel images and graphs of K_d calculation for **1a (a)** and **1b (b)**. Calculated K_d value for **1a** is 0.22 ± 0.05 μM and **1b** is 0.17 ± 0.07 μM.

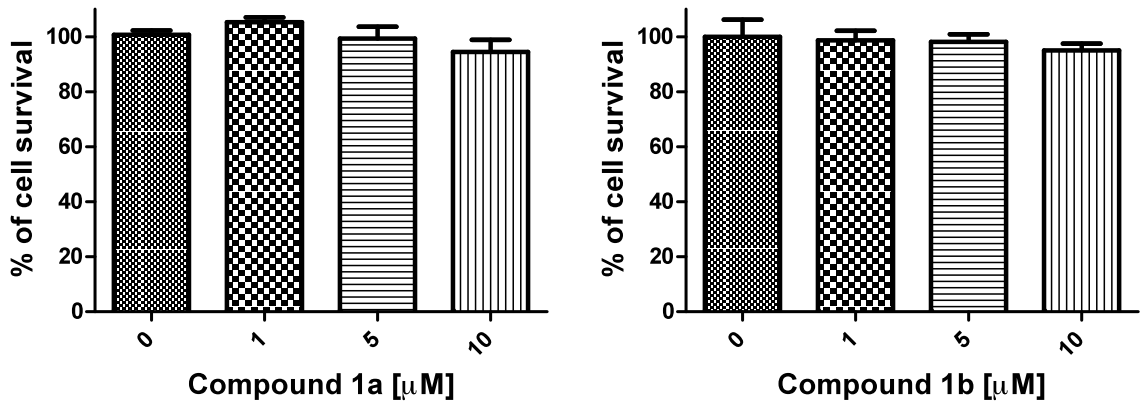


(a)



(b)

Figure S6. Dynamic linear range determination for 1a and 1b. The linear relationship between the fluorescence intensity change (ΔFI, Y) of **1a** or **1b** (1 μM in 0.01M PBS with 1 mM $MgCl_2$) and the logApt10L concentration (X) were determined by linear fitting. The regression equation was calculated with goodness of fit (R = 0.99) for **1a (a)** and (R = 0.98) for **1b (b)**. According to the calculated equation, **1a** and **1b** are linear starting from 11 nM and 16 nM of Apt10L in solution, respectively.



(a) (b)

Figure S7. Cell viability assay in HeLa cells. 5 x 10^3 of HeLa cells were grown in each well of 96-well plate in DMEM media containing 10% FBS for 24 hr at 37 °C in a humidified, 5% CO_2 atmosphere. 1, 5 and 10 μM of compound **1a** or **1b** were added in each well and then incubated for 20 hr. 20 μL per well of CellTiter96® AQ_{ueous} One Solution reagent (Promega) was added and then incubated for 2 hr. The absorbance was recorded at 490 nm using TECAN 96-well plate reader. Background absorbance of each compound was subtracted using samples which contain compound alone in media at each concentration. As a result, cell viability was not significantly affected by **(a) 1a** and **(b) 1b** up to 10 μM (~100-fold higher concentration than K_d of each compound).

Job plot to determine the binding stoichiometry between conjugate 1a (or 1b) and Apt10L or M RNA.

Fluorescence titration was carried out in PBS buffer (pH 7.4) containing 1 mM $MgCl_2$ using Horiba FluoroMax-3 fluorometer. Various concentrations of dye and RNA were prepared such as 500 nM/ 0 nM, 450 nM/ 50 nM, 400 nM/ 100 nM, 350 nM/ 150 nM, 300 nM/ 200 nM. 250 nM/ 250 nM, 200 nM/ 300 nM, 150 nM/ 350 nM, 100 nM/ 400 nM, 50 nM/ 450 nM and 0 nM/ 500 nM. Excitation and emission wavelength were 555 nm and 610 nm, respectively. Excitation and emission slit were 5 nm each.

Synthetic procedure of compounds and spectroscopic data

General procedure. Chemicals were purchased from Sigma-Aldrich, TCI, or Acros (unless noted otherwise), and were used without further purification. ^{1}H nuclear magnetic resonance (NMR) spectra were recorded on a Varian DPX 400 (400 MHz), with chemical shifts (δ) reported in ppm relative to the solvent peaks (CD_3OD: 3.31 ppm, DMSO-d_6: 2.50 ppm) and coupling constants reported in Hz. Mass spectra were recorded on a Brucker Daltonics Ultraflex MALDI TOF/TOF. UV spectra were measured on an Agilent 8453 UV/Vis spectrometer. Reverse phase analytical HPLC (RP-HPLC) was run on a Phenomenex Luna 5u C18(2) 100A column using Dionex separation module and products were detected at 214 nm, 254 nm, 325 nm and 550 nm using Dionex Ultramate 3000 Photodiode Array detector. For HPLC systems, solvents used were acetonitrile (0.1% TFA) and water (0.1%TFA). The standard gradient that was used for the analysis and purification was MeCN (0.1% TFA) /H_2O (0.1% TFA) 20/80 to 80/20. A flow rate of 12 mL/min for preparative HPLC was used.

Compound **3**. To a solution of compound **2** (392 mg, 0.264 mmol) in DMF (10 mL) was added sodium azide (20 equiv., 344 mg, 5.29 mmol) and the resulting reaction mixture was heated to 80 ^{o}C for 7 h. After the reaction was completed, the crude product was partitioned between water (100 mL) and ethyl acetate (100 mL). The aqueous layer was separated and extracted with ethyl acetate (100 mL × 2). The combined organic layer was dried with $MgSO_4$ and then evaporated. The crude product was purified with SiO_2 column chromatography (DCM: MeOH = 100 : 5) to give *N*-Boc protected azido neomycin (86 %, 281 mg, 0.227 mmol). ^{1}H-NMR (400 MHz, CD_3OD): δ 5.46 (br s, 1H), 5.16 (s, 1H), 4.92 (s, 1H), 4.35 (t, 1H, *J* = 5.2 Hz), 4.29 (br s, 1H), 4.03 (m, 1H), 3.90 (m, 2H), 3.72-3.79 (m, 3H), 3.45-3.65 (m, 8H), 3.22-3.16 (m, 6H), 1.92-2.00 (m, 1H), 1.64 (m, 1H), 1.46 (s, 54H); ^{13}C-NMR (100 MHz, CD_3OD): δ 157.8, 157.6, 157.3, 157.1, 157.0, 156.8, 110.2, 99.1, 97.6, 86.1, 79.7, 79.5, 79.5, 79.3, 79.2, 78.6, 77.8, 74.3, 74.1, 72.2, 71.8, 71.7, 70.5, 67.9, 55.6, 52.3, 51.3, 51.1, 50.2, 41.4, 40.8, 39.3, 34.8, 27.8, 27.7, 27.6; MALDI-Tof: $[M + Na]^+$ calcd. for $C_{53}H_{93}N_9NaO_{24}^+$ 1262.623, found: 1262.208.

Boc groups of *N*-Boc protected azido neomycin (250 mg, 0.202 mmol) was deprotected in TFA/DCM/Triisopropylsilane (800 μL/400 μL/40 μL) for 2 h. After the solvent was evaporated under reduced pressure, the crude product was dissolved in water (30 mL). The aqueous phase was washed with DCM (30 mL × 2) and Et_2O (30 mL × 2). And the aqueous phase was evaporated under reduced pressure and further dried using lyophilizer to give the compound **3** (quantitative, 267 mg, 0.202 mmol) as white solid. ^{1}H-NMR (400 MHz, CD_3OD): δ 6.02 (d, 1H, *J* = 3.6), 5.40 (d, 1H, *J* = 2.8), 5.31 (d, 1H, *J* = 1.6), 4.46 (t, 1H, *J* = 5.2), 4.38 (dd, 1H, *J* = 4.4, 2.8), 4.29 (m, 2H), 4.16 (m, 2H), 4.05 (dd, 1H, *J* = 10.0, 8.8), 3.99 (td, 1H, *J* = 8.8, 3.2), 3.91-3.82 (m, 2H), 3.70 (m, 1H), 3.62-3.54 (m, 2H), 3.52-3.13 (m, 10H), 2.53-2.45 (m ,1H), 2.05 (q, 1H, *J* = 12.4); ^{13}C-NMR (100 MHz, CD_3OD):

δ 162.6, 162.2, 161.9, 161.5, 121.3, 118.4, 115.5, 112.6, 111.2, 95.6, 95.1, 86.1, 80.1, 76.4, 75.4, 73.8, 72.8, 71.9, 70.8, 70.6, 68.3, 68.1, 68.0, 54.0, 52.4, 51.7, 50.2, 49.0, 40.7, 40.4, 39.2, 28.3; MALDI-Tof: $[M + H]^+$ calcd. for $C_{23}H_{46}N_9O_{12}^+$ 640.327, found 640.196.

Compound **4** was prepared by following the procedures in our previous paper (Ref. 5 in the main text).

Compound **5a**. To a solution of the compound **4** (49 mg, 93 μmol) in DMF (2 mL) was added HBTU (53 mg, 0.139 mmol), DIPEA (0.185 mmol, 32 μL) and propargylamine (12 μL, 0.185 mmol). After being stirred for 2 h at room temperature, the reaction mixture was partitioned between 0.1 M aq. HCl (100 mL) and dichloromethane (100 mL). The aqueous layer was separated and extracted with dichloromethane (100 mL × 2). The combined organic layer was dried with $MgSO_4$ and then evaporated. The crude product was purified with SiO_2 column chromatography (DCM: MeOH from 100/5 to 100/10) to give the compound **5** (67 %, 35 mg). ^{1}H-NMR (400 MHz, DMSO-d_6): δ 7.96 (d, 1H, J = 6.4), 7.64-7.47 (m, 4H), 7.46-7.31 (m, 3H), 7.17 (d, 1H, J = 6.4), 7.10-6.96 (m, 3H), 6.92 (br s, 1H), 6.83 (br s, 2H), 4.34 (s, 2H), 3.88 (s, 2H), 3.53 (s, 3H), 3.23 (s, 3H), 3.11 (s, 1H); ^{13}C-NMR (100 MHz, DMSO-d_6): δ 167.6, 162.3, 158.2, 157.8, 157.6, 157.2, 147.4, 145.3, 133.4, 131.2, 130.3, 129.8, 129.4, 128.7, 128.3, 127.3, 115.5, 115.4, 115.3, 115.0, 97.7, 96.9, 81.4, 74.1, 55.1, 41.9, 41.0, 28.7; MALDI-Tof: $[M + H]^+$ calcd. for $C_{32}H_{27}N_3O_5S^+$ 565.167, found 565.131.

Compound **5b**. To a solution of the compound **4** (40 mg, 76 μmol) in DMF (2 mL) was added HBTU (53 mg, 0.139 mmol), DIPEA (0.303 mmol, 53 μL) and 3,6,9,12-tetraoxapentadec-14-yn-1-amine hydrochloride (30 mg, 0.114 mmol). After being stirred for 2 h at room temperature, the reaction mixture was partitioned between 0.1 M aq. HCl (100 mL) and dichloromethane (100 mL). The aqueous layer was separated and extracted with dichloromethane (100 mL × 2). The combined organic layer was dried with $MgSO_4$ and then evaporated. The crude product was purified with SiO_2 column chromatography (DCM: MeOH = 100 : 7 to 100 : 10) to give the compound **5b** (80 %, 45 mg). ^{1}H-NMR (400 MHz, CD_3OD): δ 8.21 (dd, 1H, J = 8.0, 1.6), 7.73 (td, 1H, J = 8.0, 1.6), 7.66 (td, 1H, J = 8.0, 1.6), 7.56 (t, 2H, J = 8.0), 7.45 (t, 1H, J = 8.0), 7.36 (d, 2H, J = 8.0), 7.26 (t, 2H, J = 9.6), 7.17 (d, 1H, J = 14.8), 7.02 (dd, 1H, J = 9.6, 2.4), 6.93 (m, 2H), 6.85 (dd, 1H, J = 9.6, 2.4), 4.38 (d, 1H, J = 18.0), 4.32 (d, 1H, J = 18.0), 4.16 (d, 2H, J = 2.4), 3.68-3.58 (m, 14H), 3.56 (t, 2H, J = 5.2), 3.41 (t, 2H, J = 5.2), 3.34 (s, 3H), 2.85 (t, 1H, J = 2.4); ^{13}C-NMR (100 MHz, CD_3OD): δ 168.8, 160.8, 158.2, 157.7, 145.2, 144.7, 132.7, 132.5, 130.5, 130.1, 130.0, 129.8, 128.2, 128.1, 126.7, 115.6, 115.4, 115.3, 114.4, 97.4, 96.6, 79.4, 74.9, 70.4, 70.3, 70.2, 70.1, 69.2, 68.9, 57.9, 55.4, 40.6, 39.9, 39.4; MALDI-Tof: $[M + H]^+$ calcd. for $C_{40}H_{44}N_3O_9S^+$ 742.280, found 742.349.

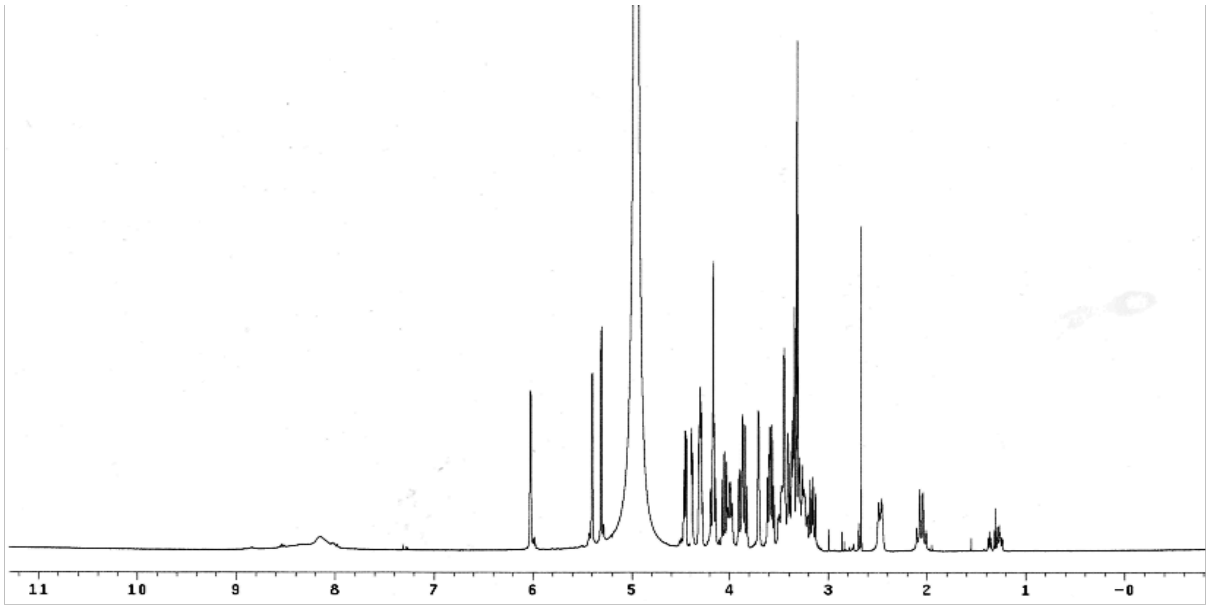
Compound **1a**. To a solution of the compound **5a** (26 mg, 45 μmol) in DMSO (2 mL) was added the

compound **3** (89 mg, 68 μmol) in DMSO (1 mL), followed by the addition of 200 μL of $CuSO_4$ solution (0.2 M in H_2O) and 120 μL of sodium ascorbate solution (1.0 M in H_2O). After being stirred for 2 h at room temperature, the crude product was purified with preparative HPLC (C-18, 20% isocratic acetonitrile (0.1% TFA) in water (0.1% TFA) for 3 minutes then 20-80% acetonitrile (0.1% TFA) in water (0.1% TFA) over 25 minutes, flow rate = 12 mL/min, R_t of the product = 13.5 min) to give the compound **1a** (52 mg, 61%). ^{1}H-NMR (400 MHz, CD_3OD): δ 8.10 (d, 1H, *J* = 7.6), 7.98 (br s, 1H), 7.84 (s, 1H), 7.66-7.54 (m, 2H), 7.45 (t, 1H, *J* = 7.6), 7.35 (t, 1H, *J* = 7.6), 7.24 (d, 2H, *J* = 7.6), 7.19 (d, 1H, *J* = 7.6), 7.10 (t, 1H, *J* = 9.2), 7.04 (d, 1H, *J* = 9.2), 6.94-6.79 (m, 3H), 6.75 (d, 1H, *J* = 9.2), 5.80 (br s, 1H), 5.26 (dd, 1H, *J* = 8.0, 3.2), 5.19 (s, 1H, *J* = 4.4), 4.74 – 3.90 (m, 16H), 3.78-3.56 (m, 4H), 3.49 (s, 3H), 3.45-3.05 (m, 9H), 2.34 (m, 1H), 1.90 (m, 1H); MALDI-Tof: $[M]^+$ calcd. for $C_{55}H_{72}N_{12}O_{17}S^+$ 1204.486, found 1204.781; $[M + Na]^+$ $C_{55}H_{72}N_{12}NaO_{17}S^+$ 1227.476, found 1226.781.

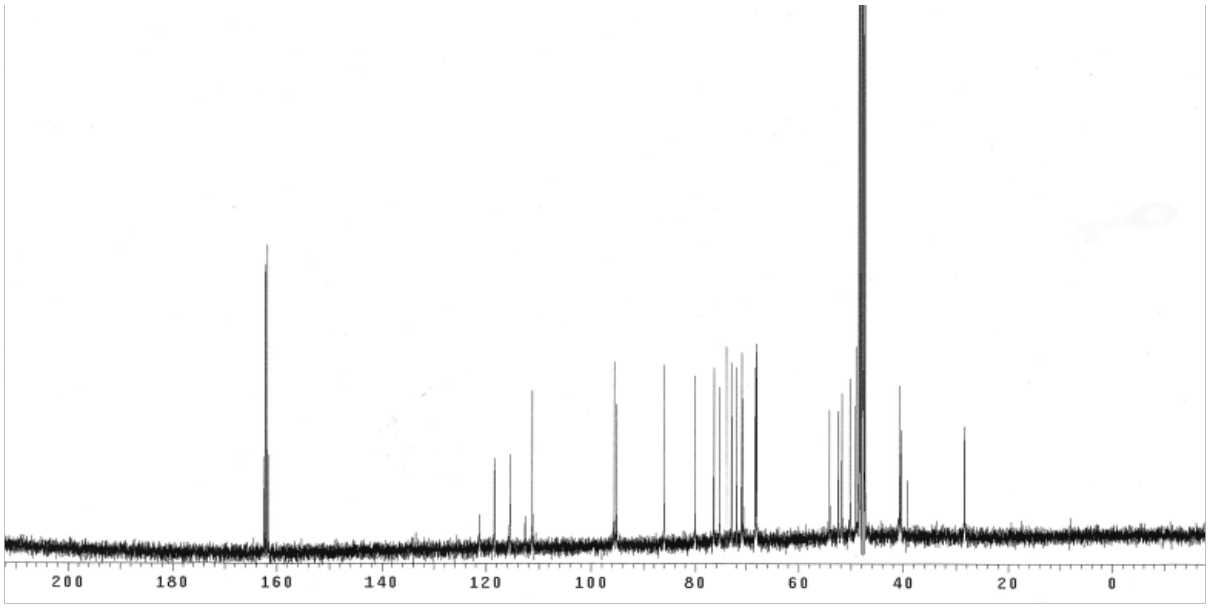
Compound **1b**. To a solution of the compound **5b** (12 mg, 16.2 μmol) in DMSO (1 mL) was added the compound **3** (32 mg, 24.3 μmol) in DMSO (1 mL), followed by the addition of 80 μL of $CuSO_4$ solution (0.2 M in H_2O) and 50 μL of sodium ascorbate solution (1.0 M in H_2O). After being stirred for 2 h at room temperature, the crude product was purified with preparative HPLC (C-18, 20% isocratic acetonitrile (0.1% TFA) in water (0.1% TFA) for 3 minutes then 20-80% acetonitrile (0.1% TFA) in water (0.1% TFA) over 25 minutes, flow rate = 12 mL/min, R_t of the product = 13.5 min) to give the compound **1b** (24 mg, 72%). ^{1}H-NMR (400 MHz, D_2O): δ 7.93 (m, 2H), 7.58 (t, 1H, *J* = 7.6), 7.43 (t, 1H, *J* = 7.6), 7.30-7.18 (m, 3H), 6.99 (d, 2H, *J* = 7.6), 6.95 (d, 1H, *J* = 9.6), 6.88 (d, 1H, *J* = 9.2), 6.80-6.68 (m, 2H), 6.57 (m, 1H), 6.43 (m, 2H), 5.84 (d, 1H, *J* = 4.0), 5.23 (d, 1H, *J* = 4.0), 5.12 (s, 1H), 4.64-4.32 (m, 4H), 4.27-4.09 (m, 4H), 4.04 (t, 1H, *J* = 2.8), 3.94-3.56 (m, 6H), 3.55-3.31 (m, 20H), 3.29-3.12 (m, 10H), 3.06 (s, 3H), 2.31 (m, 1H), 1.71 (q, 1H, *J* = 8.8); MALDI-Tof: $[M + H]^+$ calcd. for $C_{63}H_{89}N_{12}O_{21}S^+$ 1381.599, found 1381.626; $[M + Na]^+$ $C_{63}H_{88}N_{12}NaO_{21}S^+$ 1403.581, found 1404.181.

^{1}H and ^{13}C spectra

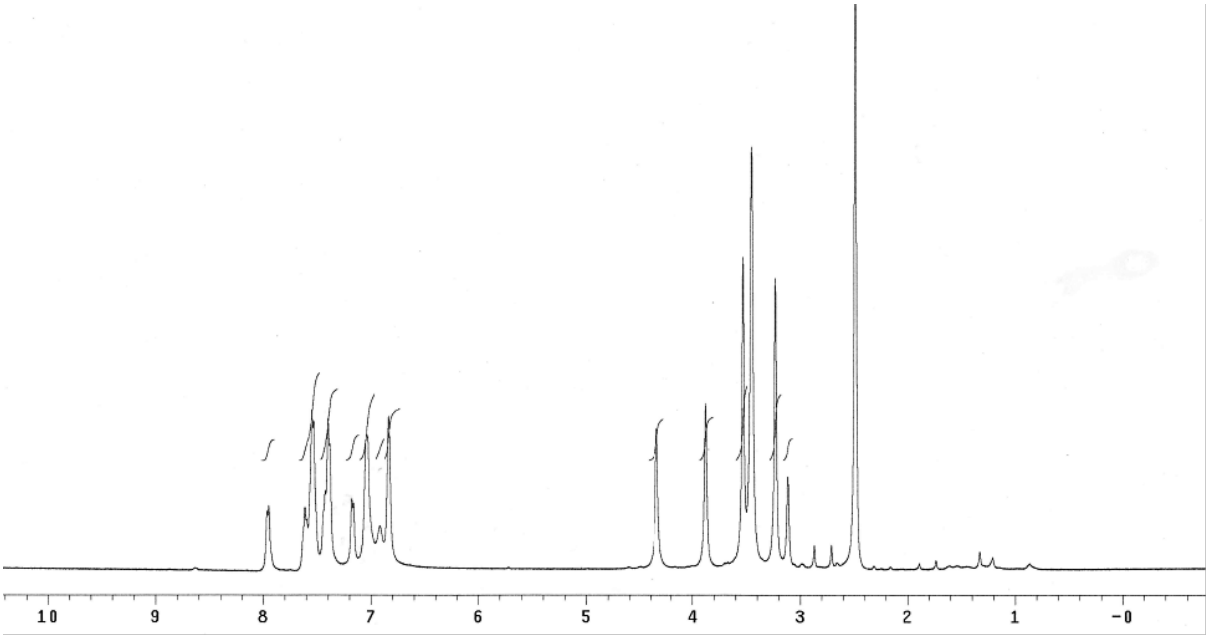
Compound **3** (^{1}H-NMR)



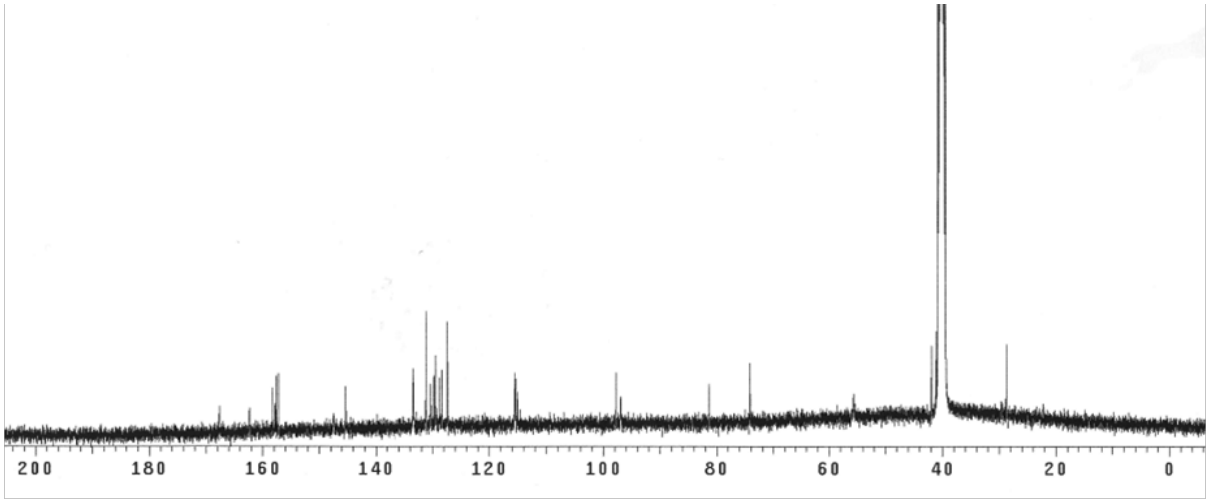
Compound **3** (^{13}C-NMR)



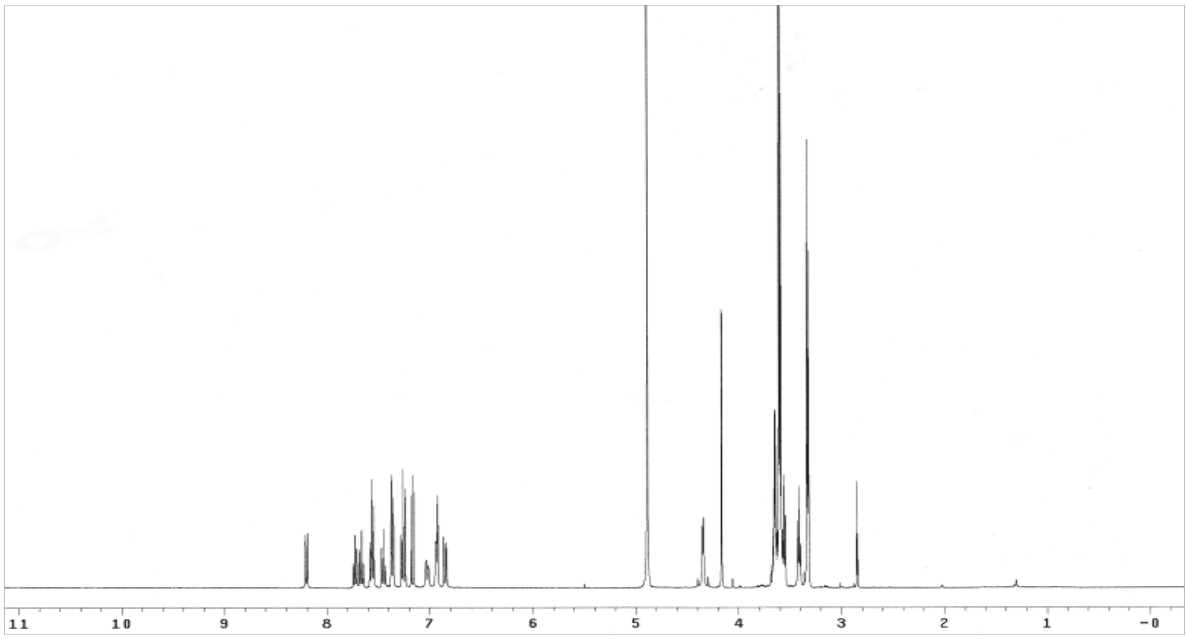
Compound **5a** (^{1}H-NMR)



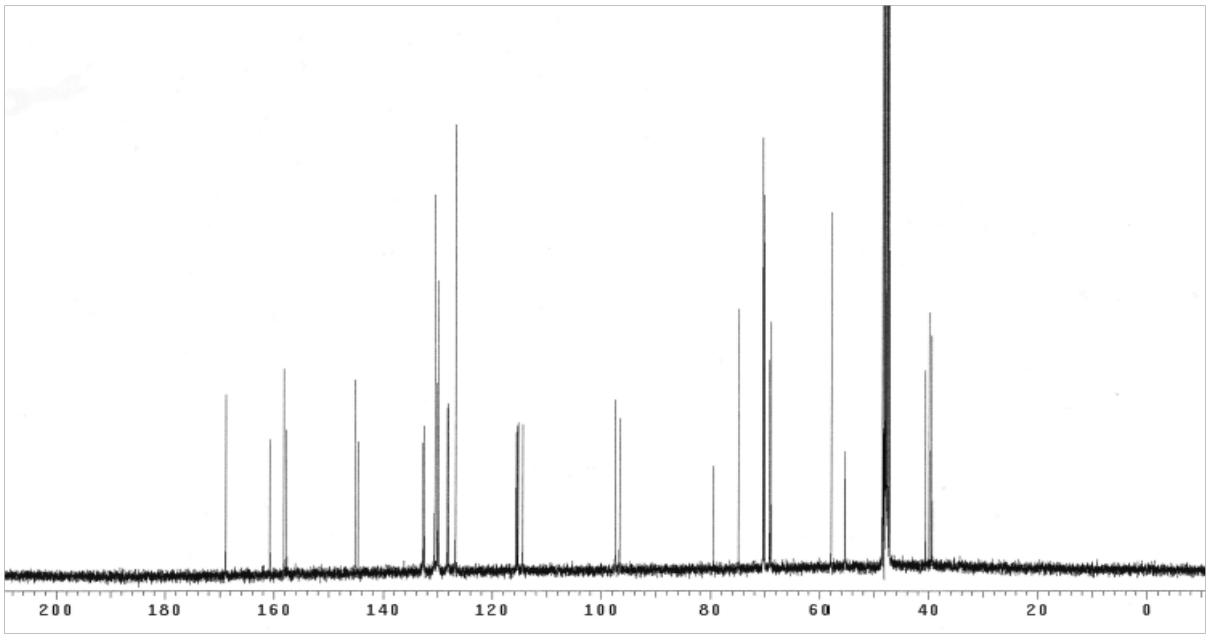
Compound **5a** (^{13}C-NMR)



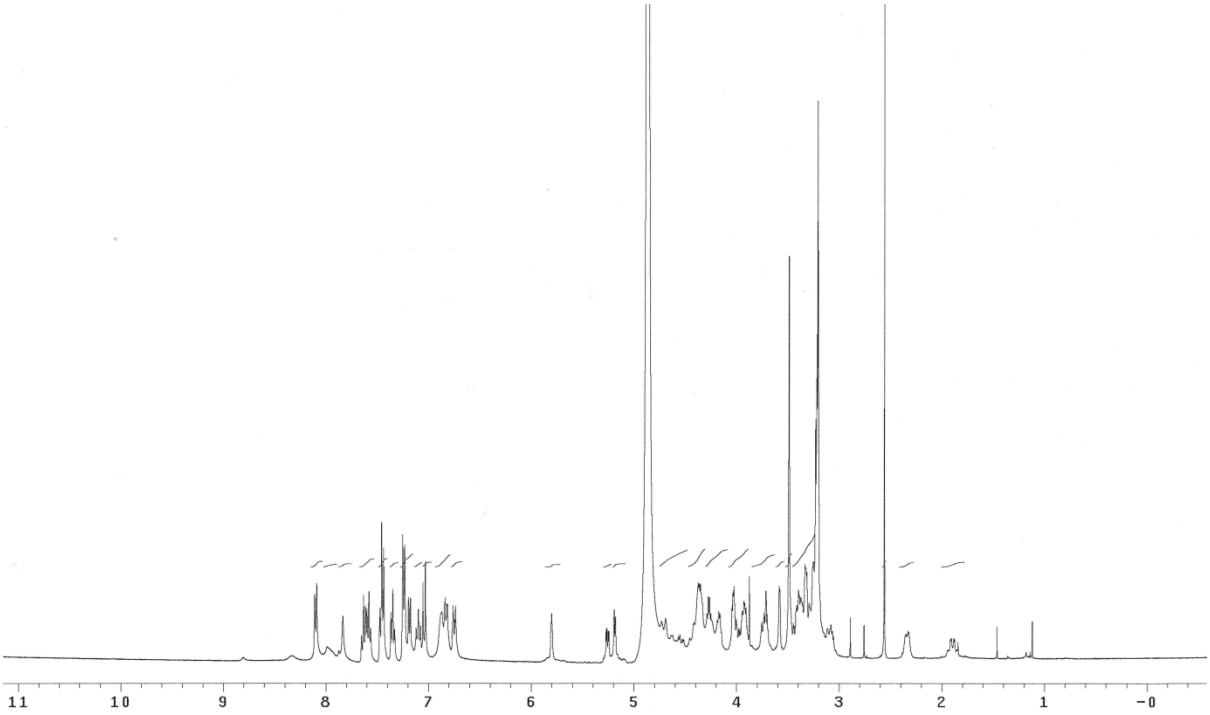
Compound **5b** (^{1}H-NMR)



Compound **5b** (^{13}C-NMR)



Compound **1a** (^{1}H-NMR)



Compound **1b** (^{1}H-NMR)