

Supplementary Information

Design of a Fluorescent Ligand Targeting the S-adenosylmethionine

Binding Site of the Histone Methyltransferase MLL1

Yepeng Luan^{1†}, *Levi Blazer*^{2‡}, *Hao Hu*¹, *Taraneh Hajian*², *Jing Zhang*¹, *Hong Wu*², *Scott Houliston*² and *Cheryl H. Arrowsmith*^{2,3}, *Masoud Vedadi*^{2,4*}, *Yujun George Zheng*^{1*}

¹Department of Pharmaceutical and Biomedical Sciences, University of Georgia, Athens, Georgia 30602, United States.

²Structural Genomics Consortium, University of Toronto, Toronto, Ontario, M5G 1L7, Canada

³Princess Margaret Cancer Centre and Department of Medical Biophysics, University of Toronto, Toronto, ON, M5G 2M9.

⁴Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, M5S 1A8.

Content:

Detailed protein purification procedures	Page 2
Differential Static Light Scattering	Page 5
Saturation-Transfer Difference (STD) NMR	Page 6
Supplementary Figure 1	Page 7
Supplementary Figure 2	Page 8
Supplementary Figure 3	Page 9
MS and NMR spectra for compounds	Page 10

Detailed protein purification procedures:

MLL1 SET Domain: MLL SET domain (residues 3785-3969) in pET28GST-LIC was expressed in *E.coli* BL21 (DE3) V2RpRARE in TB medium (Sigma, Cat# T-9179) in the presence of 100 µg/mL of Kanamycin. Cells were grown at 37°C using the LEX system (HarbingerBiotech) in 2 L bottles (VWR, 89000-242) to an OD600 of 0.8. The temperature was reduced to 15 °C and recombinant protein expression was induced by the addition of isopropyl-1-thio-D-galactopyranoside (IPTG, Bioshop, Cat# IPT001.100) to final concentration of 1 mM. Recombinant protein expression was allowed to continue overnight at 15°C. Cells were harvested by centrifugation (Beckman Coulter, Aventi, Roto JLA8.10) and cell pellets were frozen in liquid nitrogen and stored at -80°C. Cell paste was thawed and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 2 mM β-mercaptoethanol, 5% glycerol) with protease inhibitor (1mM phenylmethyl sulfonyl fluoride, PMSF). The cells were lysed by sonication at frequency of 8.5 for 5 min. The cleared lysate was loaded onto a GST Bind column (Novagen, Cat# 70541) and the column was washed with 50 column volumes of 50 mM Tris-HCl, pH 7.5, containing 500 mM NaCl and 5% glycerol. MLL1 SET domain was liberated by incubating the column with 25 units of thrombin (Sigma, Cat# T6884) overnight at 4°C. The MLL protein was collected further purified over a Superdex200 (26x60, GE Healthcare, Cat# 17-1043-02) size exclusion column equilibrated in 20 mM Tris-HCL, pH 8.0 buffer containing 500 mM NaCl. Fractions containing MLL were pooled, 3 mM β-mercaptoethanol added and concentrated by using Amicon Ultra centrifugal filter.

MLL1 Complexes and complex components: DNA fragments encoding two human components of MLL1 complex - RbBP5 (1–538) and WDR5 (1–334) were amplified by PCR and sub-cloned into the pFastBac-Dual vector (Invitrogen). DNA fragments encoding a truncated form of human MLL1 (3745-3969) were amplified by PCR and sub-cloned into pFBOH-MHL (a derivative of the pFBOH-LIC vector). The resulting plasmids were transformed into DH10Bac™ Competent E. coli (Invitrogen) and recombinant viral DNA bacmids were purified. Recombinant baculovirus generation was performed in Sf9 cells. Sf9 cells grown in HyQ® SFX insect serum-free medium (ThermoScientific) were co-infected with 10 mL of each P3 viral stock per 0.8 L of suspension cell culture and incubated at 27°C for 72 hours. Harvested cells were re-suspended in 20 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl, 5 mM imidazole and 5% glycerol, 1X Complete EDTA-free protease inhibitor cocktail tablet (Roche). The cells were lysed chemically by rotating for 30 minutes in the presence of 0.6% NP40, 50 U/mL Benzonase Nuclease (Sigma), and 2 mM 2-Mercaptoethanol followed by sonication at frequency of 8 with a 50% duty cycle for 4 min (Sonicator 3000, Misoni). Crude extract was clarified by high-speed centrifugation (60 min at 36,000 ×g at 4°C). Recombinant protein complex was purified by loading the cleared lysate on a Talon metal affinity resin (Clontech) followed by washing the column with 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5% glycerol, 15 mM imidazole. The recombinant protein complex was eluted from the cobalt-based affinity column in the same buffer supplemented with 250 mM imidazole. Recovered protein was concentrated and loaded onto a Superdex-200 size exclusion column (GE Healthcare) equilibrated with 20 mM Bis-Tris propane pH 7.0, 250 mM NaCl. Trimeric MLL1 complex was recovered and purity was confirmed on SDS-PAGE and LC-MS.

ASH2L: A cDNA construct for full-length human ASH2L (1-628) was generously provided to us by Dr. Jean-Francois Couture (University of Ottawa) and was cloned into the pHIS2 vector. Protein was overexpressed as polyhistidine-tagged fusion protein in *E. coli* Rosetta cells (Novagen). Harvested cells were re-suspended in sodium phosphate buffer (50 mM, pH 7.0), containing 500 mM NaCl, 5 mM 2-Mercaptoethanol, and 1X Complete EDTA-free protease inhibitor cocktail tablet (Roche). The cells were lysed by sonication at frequency of 8.5 with a 50% duty cycle for 4 min (Sonicator 3000, Misoni). The crude extract was clarified by high-speed centrifugation (60 min at 36,000 ×g at 4°C). Recombinant protein was purified by loading the cleared lysate on a Talon metal affinity resin (Clontech) followed by washing 10 column volume with sodium phosphate buffer (50 mM, pH 7.0), 500 mM NaCl and 5 mM 2-Mercaptoethanol. Protein was eluted with 250 mM Imidazole in the same buffer. Protein was dialyzed against 50 mM sodium phosphate buffer pH 7.0, 300 mM NaCl, and 5 mM 2-mercaptoethanol in the presence of TEV protease and purified protein was recovered from the flow through of a Ni-NTA column. Purified protein was dialyzed against 50 mM sodium phosphate pH 7.0, 150 mM NaCl and 5 mM 2-Mercaptoethanol and concentrated for storage. Purity was confirmed on SDS-PAGE and LC-MS.

DPY-30: A DNA fragment encoding human DPY30 (residues 1-99) was amplified by PCR and sub-cloned into the pET28-MHL vector. Recombinant protein was overexpressed in *E. coli* BL21 (DE3) -V2R- pRARE2 during an overnight incubation at 15 °C in the presence of 1 mM isopropyl-1-thio-D-galactopyranoside (IPTG). Harvested cells were re-suspended in 20 mM Tris-HCl, pH 7.5, containing 500 mM NaCl, 5 mM imidazole and 5% glycerol, 1X Complete EDTA-free protease inhibitor cocktail tablet (Roche). Cells were lysed chemically by rotating 30

min with 0.5% CHAPS and 22.5 U/mL Benzonase nuclease (Sigma) followed by sonication at frequency of 8.5 with a 50% duty cycle for 4 min (Sonicator 3000, Misoni). The crude extract was clarified by high-speed centrifugation (60 min at 36,000 ×g at 4°C). Lysate loaded onto a DE52 ion-exchange resin (Whatman) and passed through a Hispur™ Ni-NTA resin (Thermo Scientific) column. The column was washed and eluted by running 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5% glycerol, containing 30 mM and 250 mM imidazole, respectively. Purified protein was dialyzed against 50 mM sodium phosphate pH 7.0, 150 mM NaCl and 5 mM 2-Mercaptoethanol and concentrated for storage. Purity was confirmed on SDS-PAGE and LC-MS.

To form tetrameric complex, components were mixed in equimolar ratios and incubated for at least 30 minutes on ice. Pentameric complex was prepared by mixing the trimeric complex with equimolar ratios of ASH2L and DPY30 for 2 hours against sodium phosphate buffer (50 mM, pH 7.0), 300 mM NaCl and 5 mM 2-Mercaptoethanol. The complex was then purified by size exclusion chromatography on a Superdex-200 16/60 column.

Differential Static Light Scattering (DSLS). DSLS experiments were performed according to a previously published protocol¹. MLL1 SET domain was diluted to 0.26 g/L in 20 mM bis-tris propane pH 8.0 in the presence of SAH, FL-NAH, or an appropriate vehicle control (DMSO). Forty microliters of these solutions were added to a clear-bottom black 384-well plate (Nunc, Cat. #242764) and samples were covered with 40 µl of mineral oil (Sigma, Cat.# MKBQ1755V). Aggregation assays were performed over a temperature gradient of 25 to 85°C at a rate of 1°C/minute in a StarGazer™ Instrument (Harbinger Biotech, Toronto, ON). T_{agg} values were

calculated using Boltzmann regression analysis in the Bioactive (v. 2.1.10) software and analyzed using GraphPad Prism (v. 6.03).

Saturation-Transfer Difference (STD) NMR. STD NMR² spectra were collected at 25°C using 256 transients on a Bruker 800MHz spectrometer. Protein saturation was for 2 s using an E-Burp2 pulse train centered at -0.5 ppm, and a bandwidth of 300 Hz. Off-resonance saturation pulses were centered at 30 ppm. H₂O suppression was achieved using excitation sculpting³. An exponential line broadening (LB=0.5 Hz) window function was applied for processing.

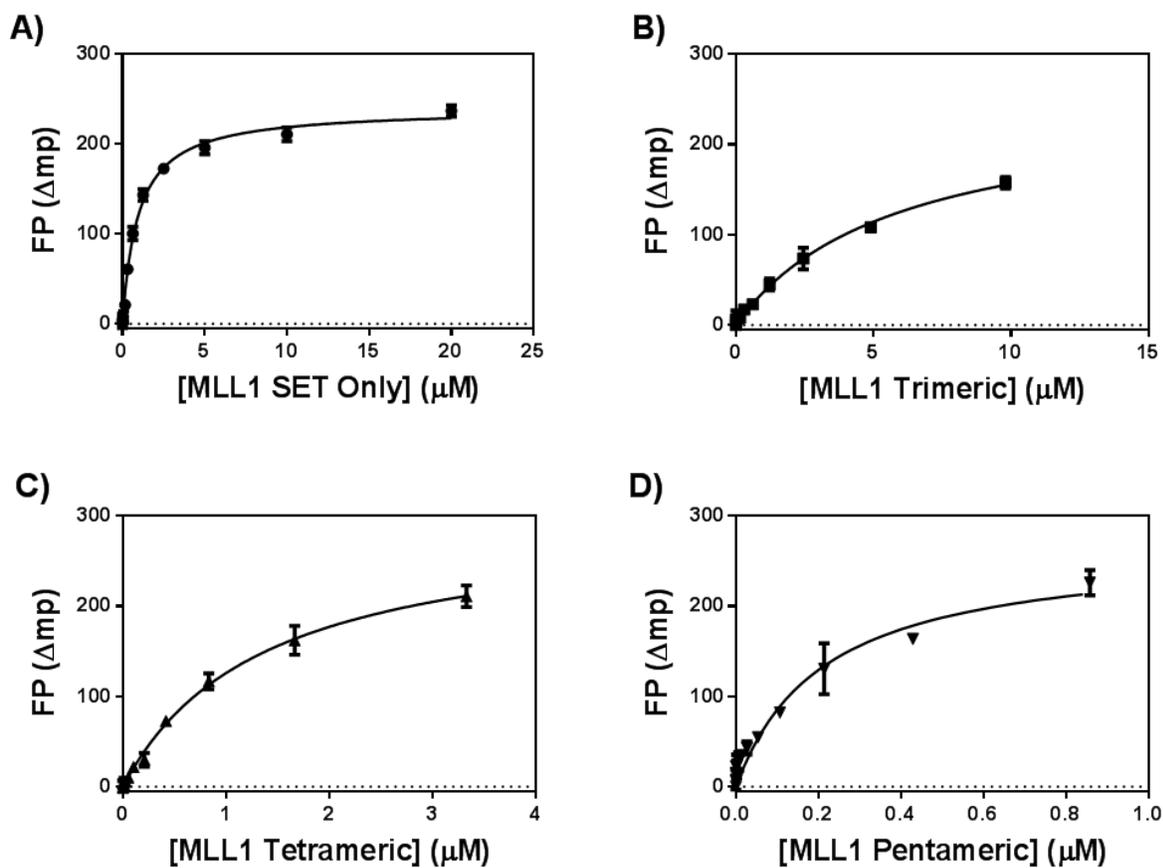


Figure S1. Confirmation of FL-NAH-MLL1 interaction by fluorescence polarization. Saturation binding of FL-NAH to A) MLL1 SET domain, B) Trimeric, C) Tetrameric, and D) Pentameric MLL1 complexes. The data presented in panel A are also presented in **Figure 2**, but are provided here again for comparison. Binding data calculated from these experiments are presented in **Table 1**. Data are presented as the mean \pm SD from at least three independent experiments.

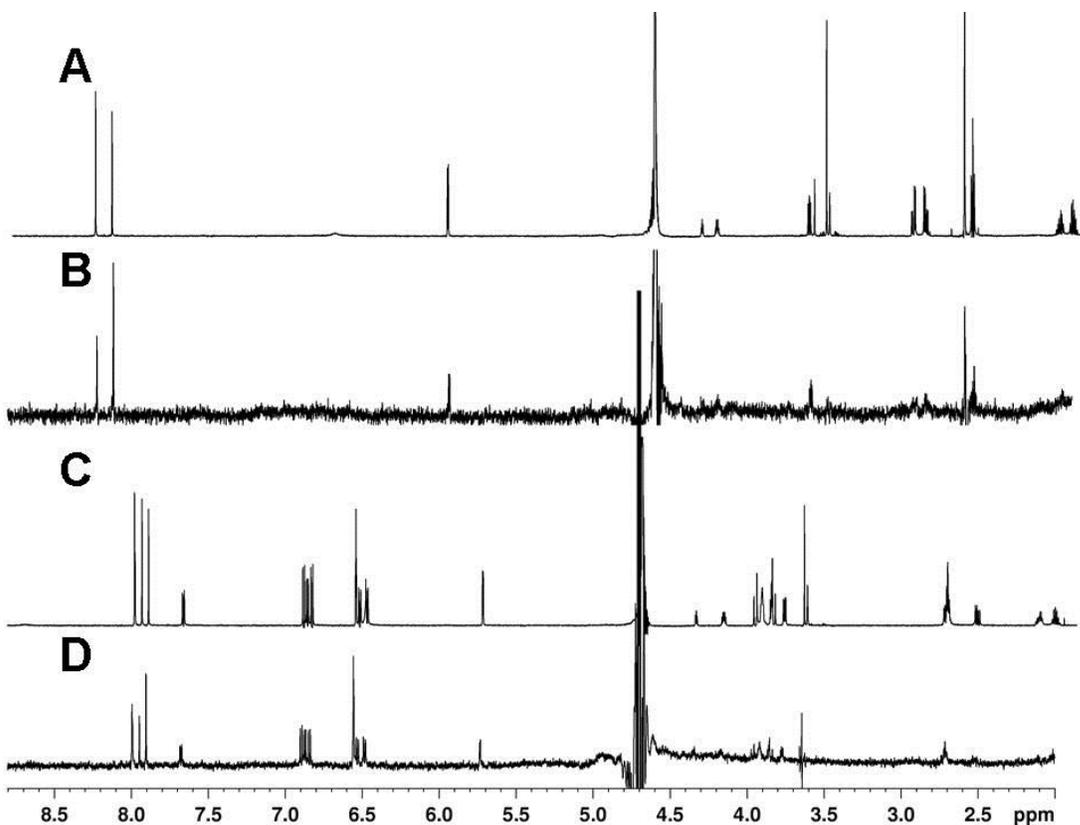


Figure S2. Confirmation of FL-NAH and SAH binding to MLL1 SET domain by STD NMR. Reference (A) and corresponding STD spectrum (B) of SAH (1.0 mM) in the presence of MLL1 SET domain (40 μM). Reference (C) and corresponding STD spectrum (D) of FL-NAH (1.0 mM) in the presence of MLL1-SET domain (40 μM). In both cases, STD signals are observed most strongly for downfield aromatic resonances.

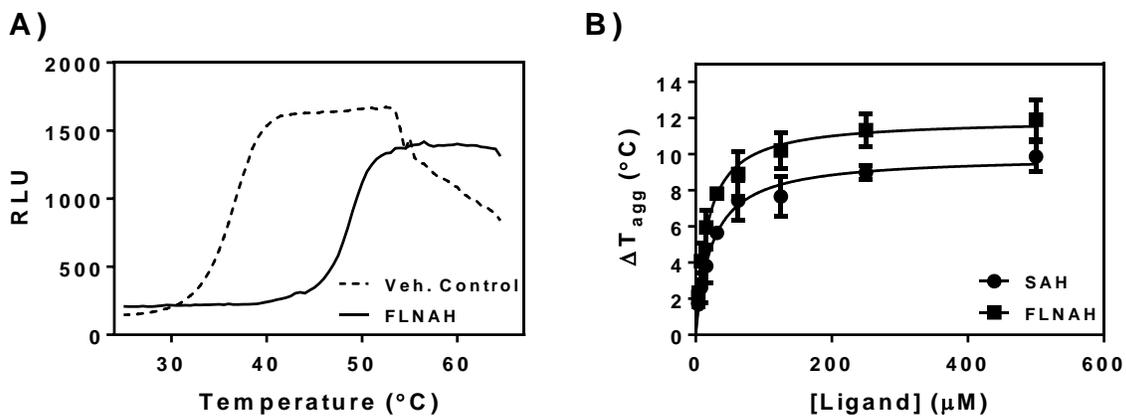
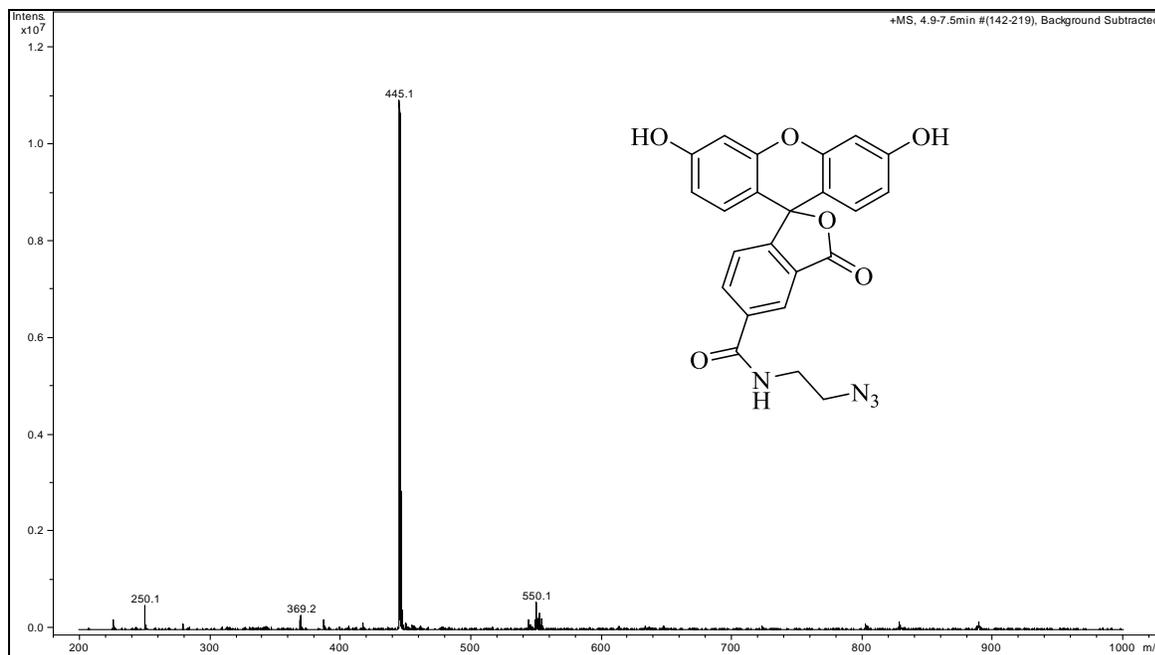


Figure S3. SAH and FL-NAH bind to the isolated SET domain of MLL1. Presence of FL-NAH stabilized the SET domain of MLL1 in a concentration-dependent manner similar to SAH as monitored by Differential Static Light Scattering (DSLS). A) Raw data trace from a single experiment with 500 μM FL-NAH or a vehicle control. Representative data shown. B) Concentration-response data of SAH and FL-NAH stabilization of the isolated SET domain of MLL1. At 500 μM , ΔT_{agg} values of 10 ± 0.8 °C and 12 ± 1 °C were obtained for SAH and FL-NAH, respectively. Data are presented as the mean \pm SD from three independent experiments.

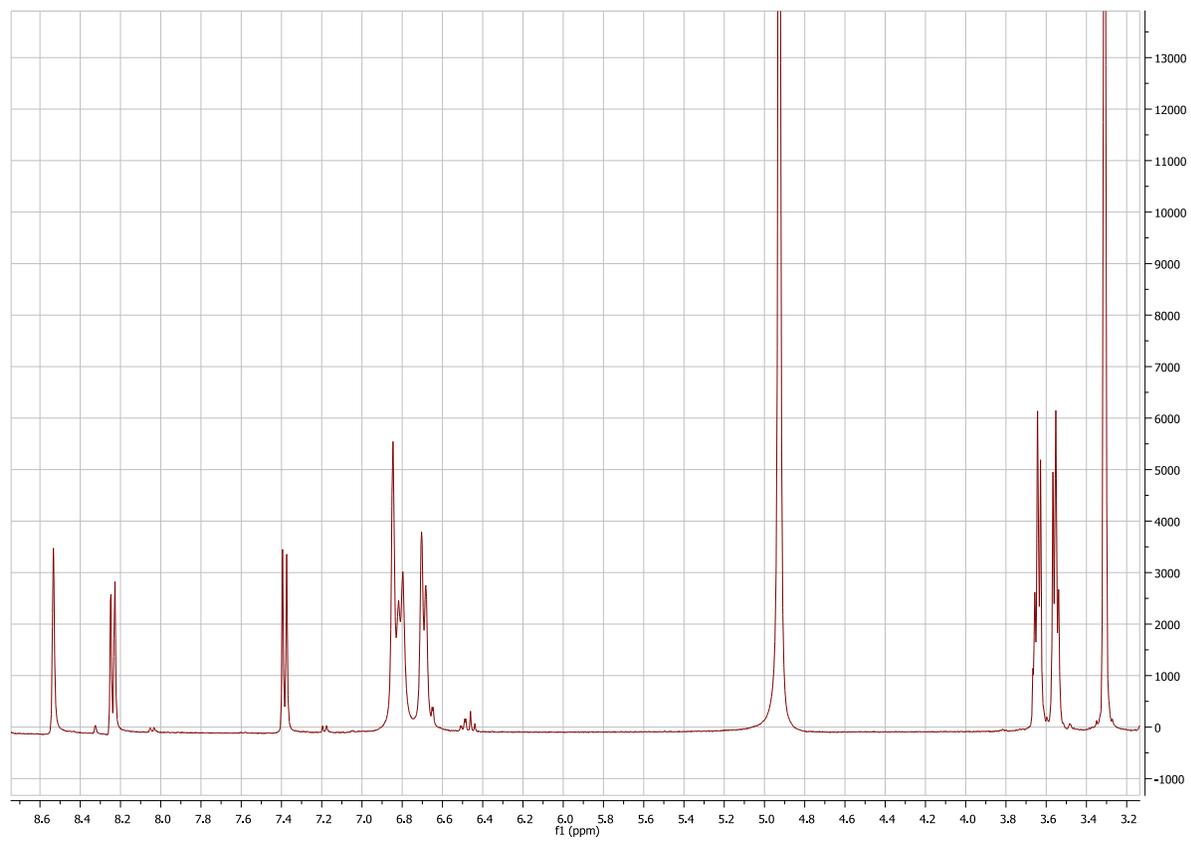
MS and NMR spectra for compounds:

Compound 1

ESI-MS

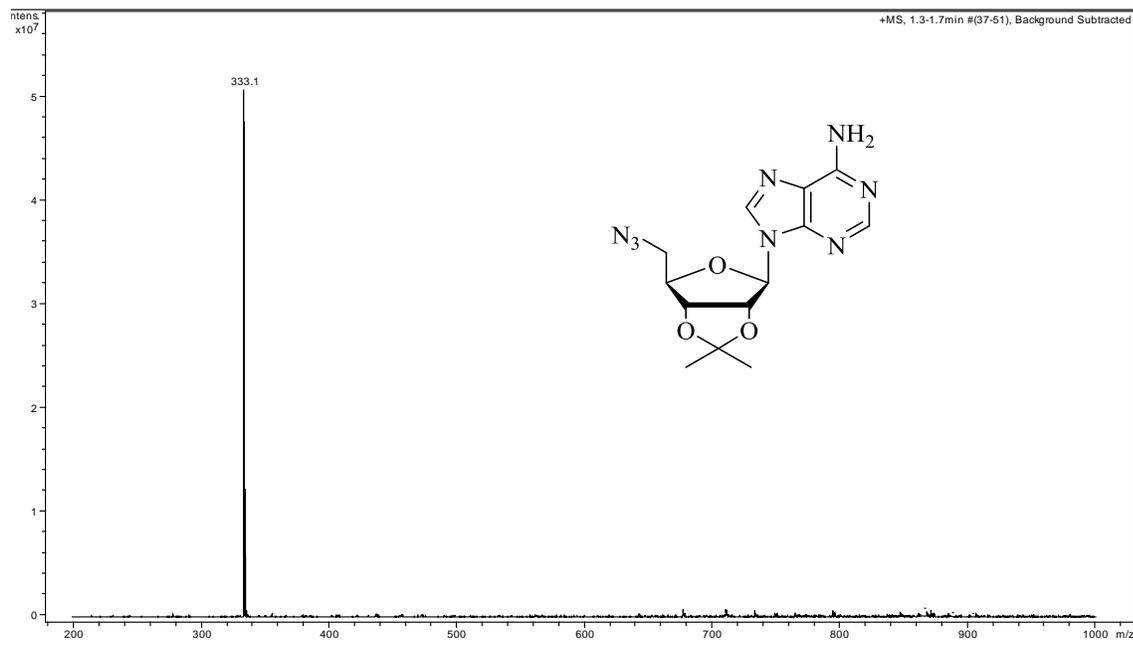


$^1\text{H-NMR}$

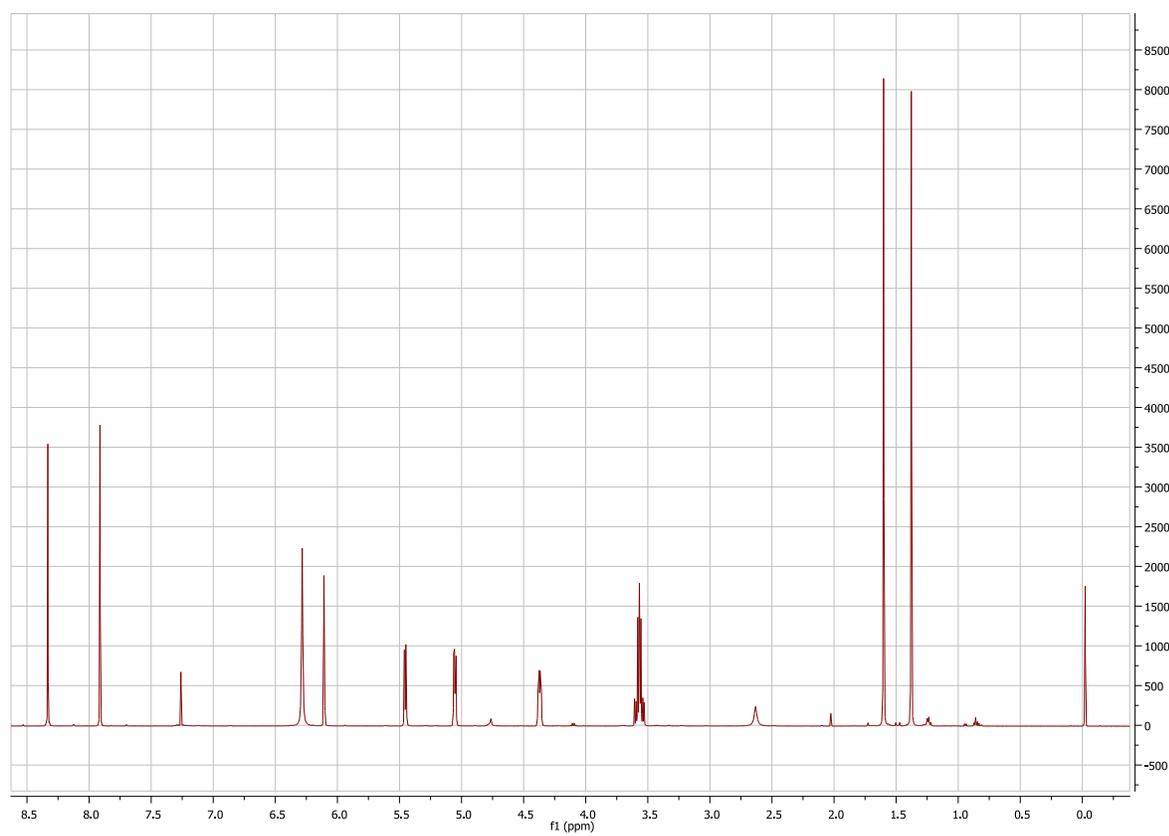


Compound 2

ESI-MS

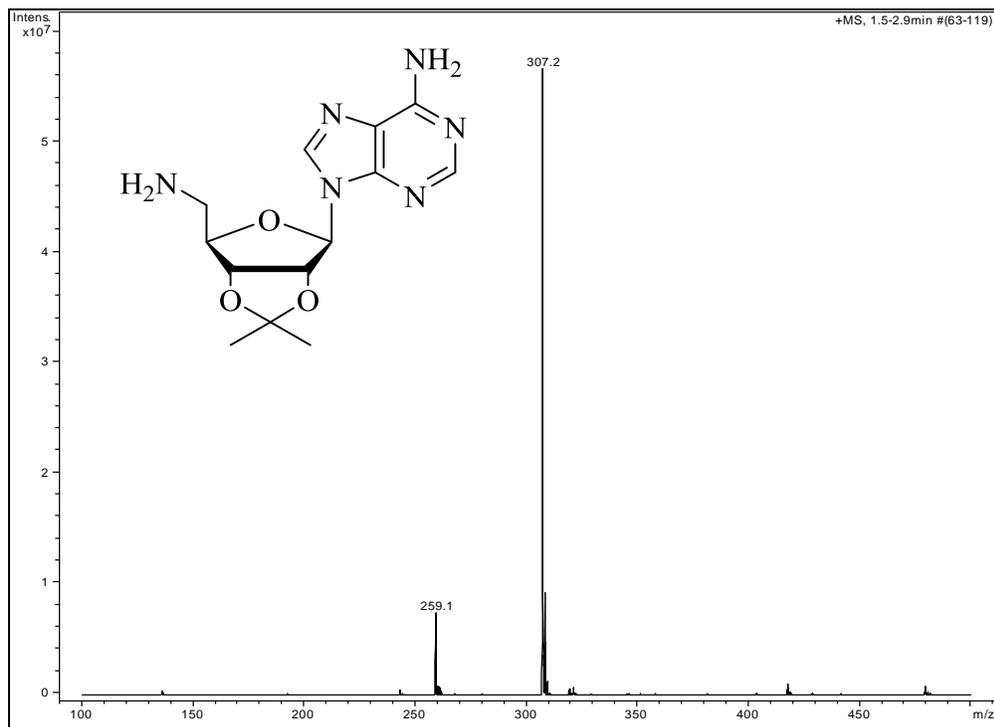


¹H-NMR

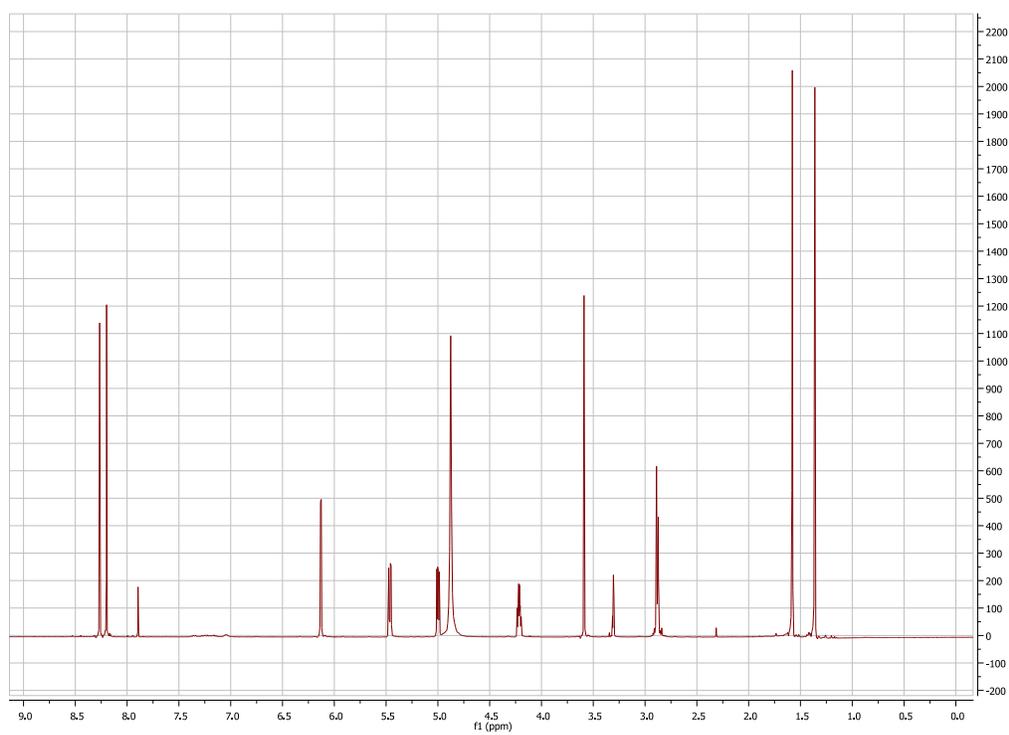


Compound 3

ESI-MS

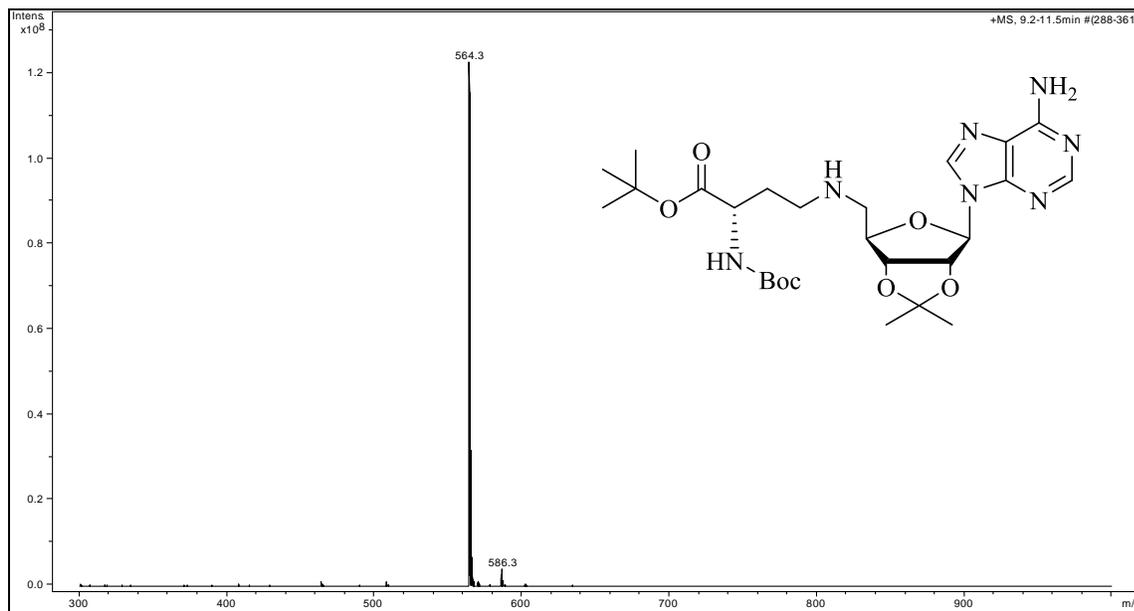


¹H-NMR

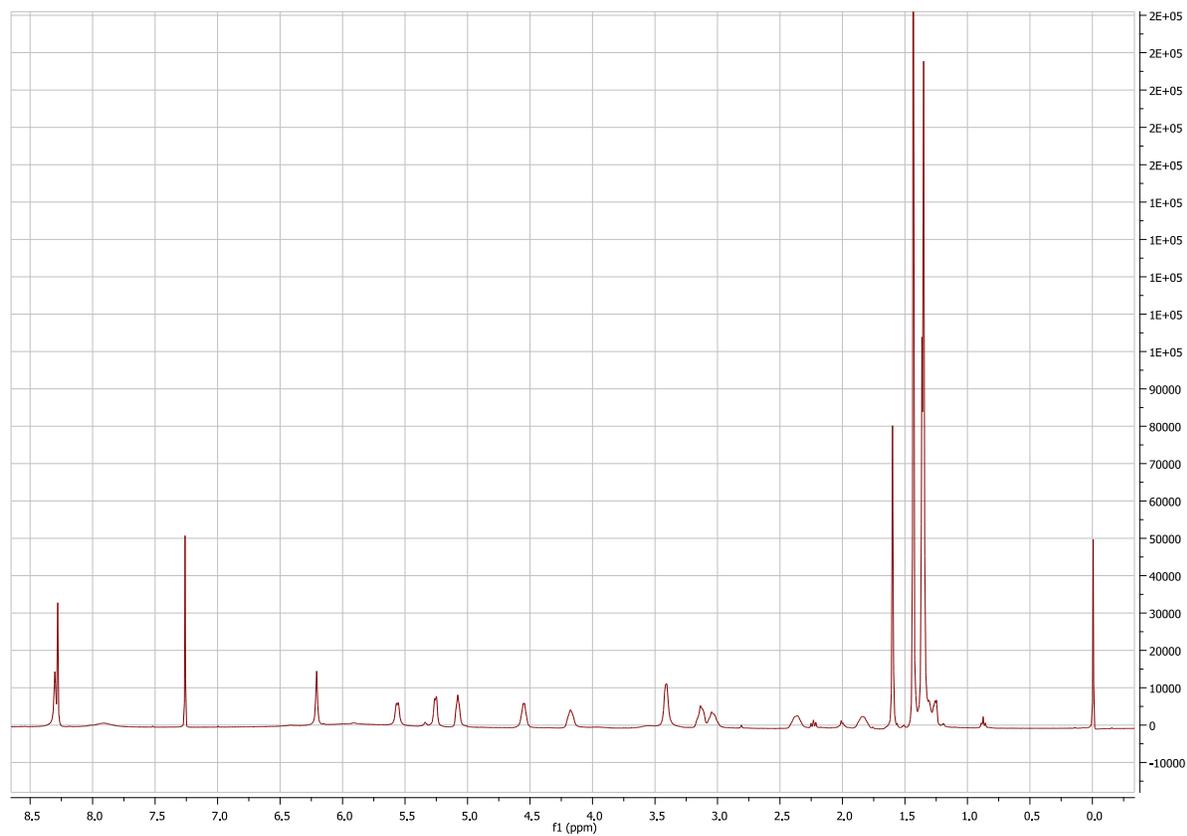


Compound 4

ESI-MS

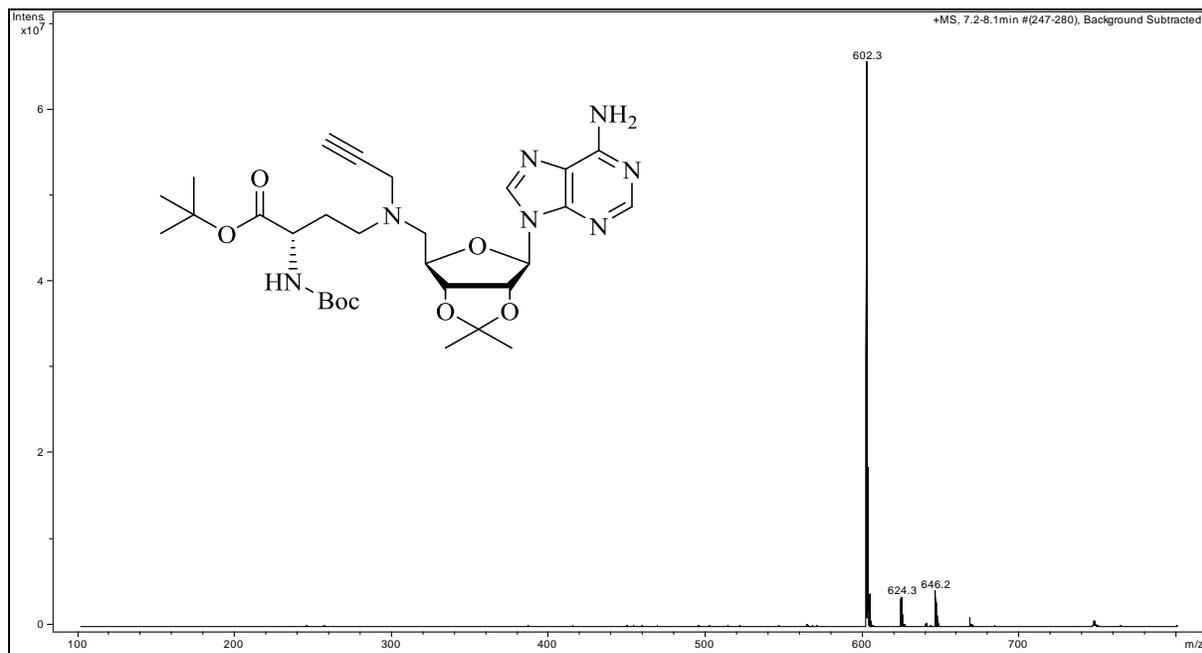


¹H-NMR

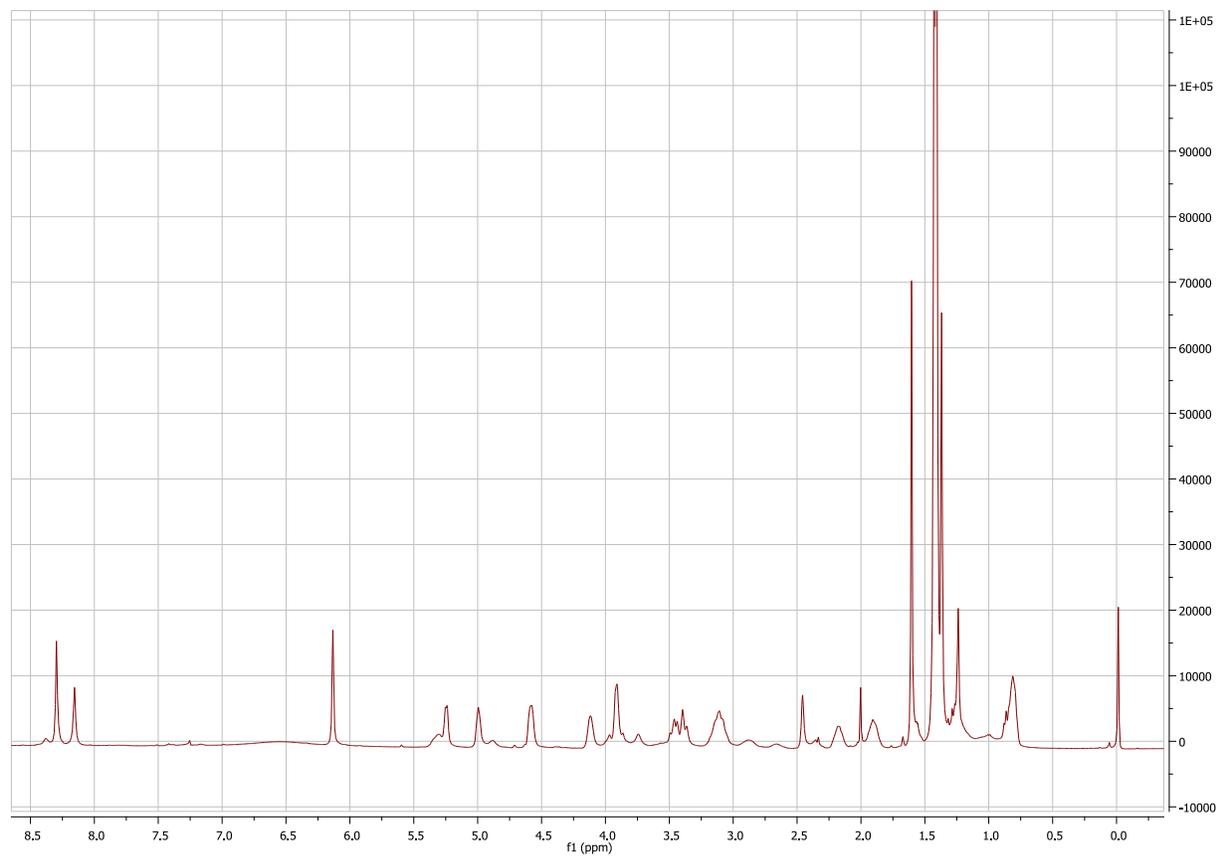


Compound 5

ESI-MS



¹H-NMR



Compound 6 (FL-NAH)

Analytical HPLC condition:

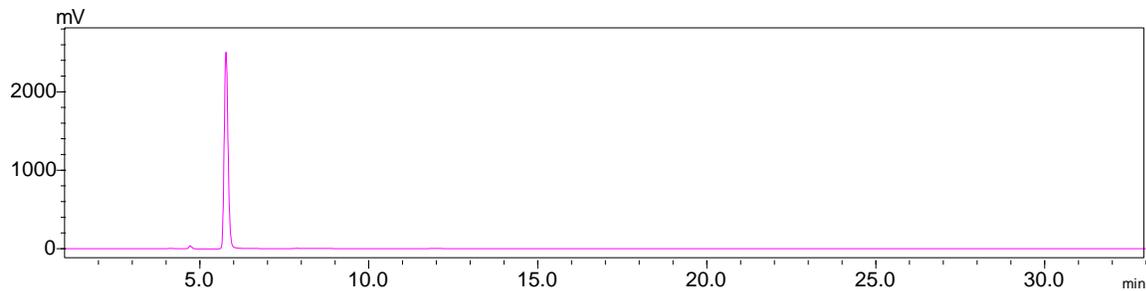
Column: Waters Nova-Park C18, 60 Å, 4 µm 150 × 3.9 mm;

Flow rate: 1 mL/min;

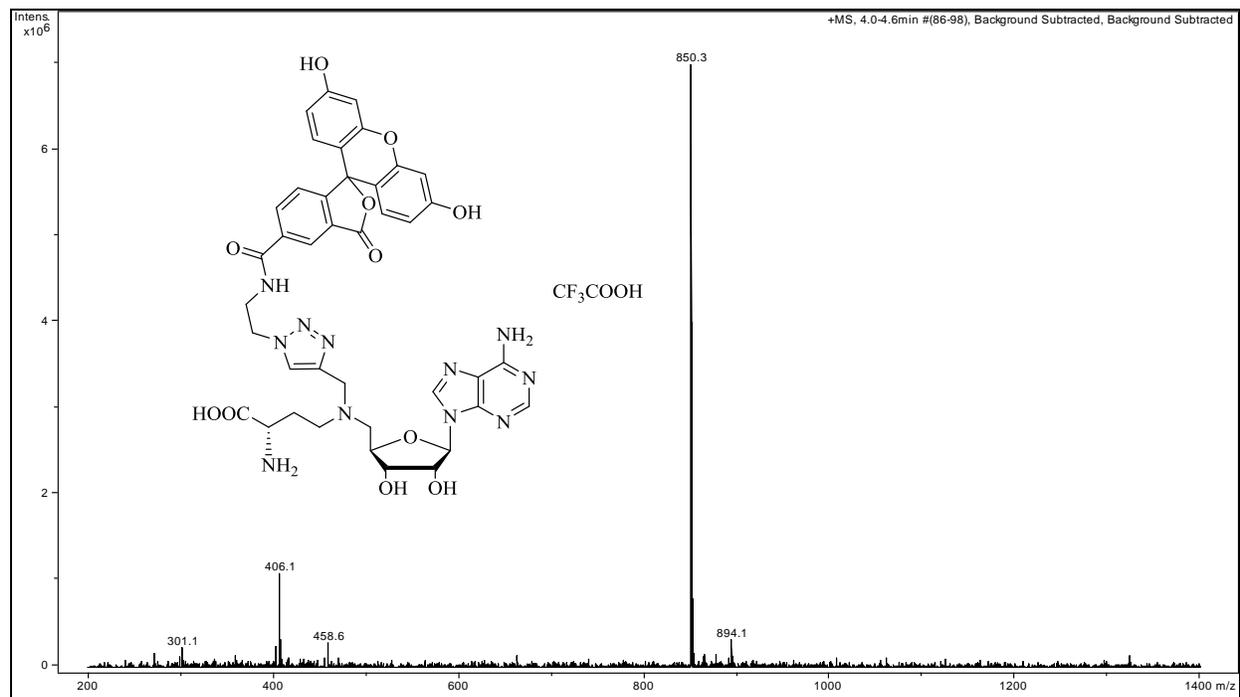
Detection wavelength: 214nm;

Eluent: (A) water (HPLC grade, +0.1% TFA), (B) acetonitrile (HPLC grade, +0.1% TFA);

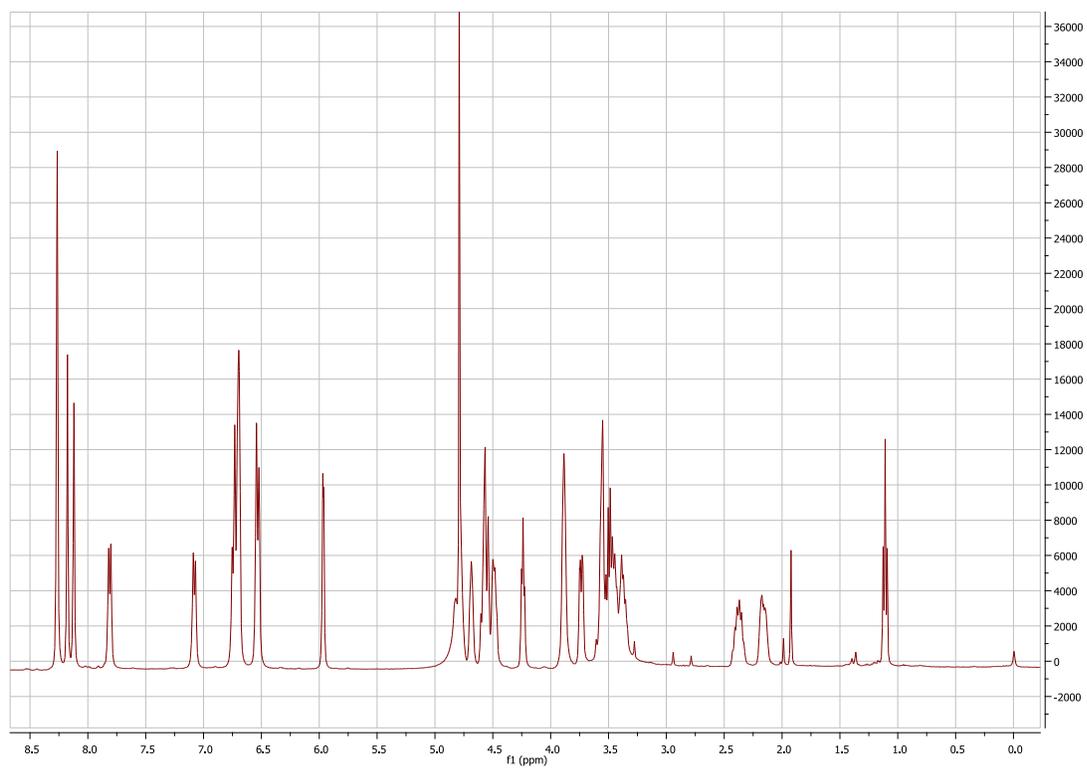
Gradient: B%=30% for 33 min.



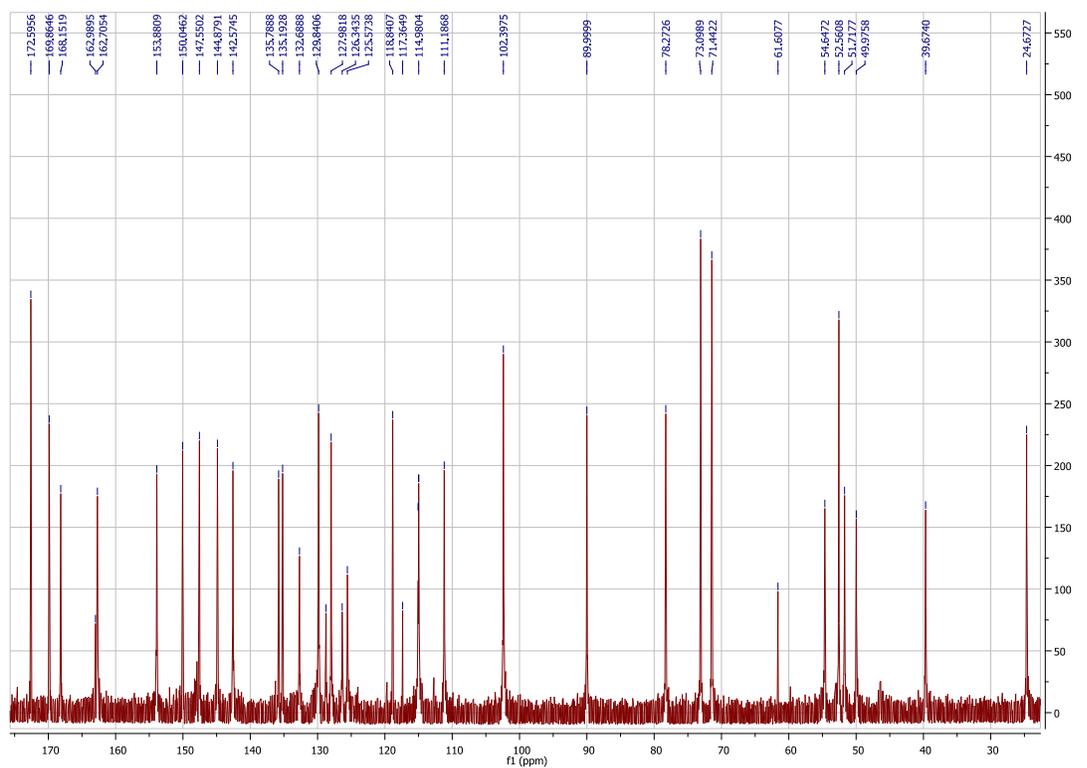
ESI-MS



¹H-NMR



¹³C-NMR



References:

1. M. Vedadi, F. H. Niesen, A. Allali-Hassani, O. Y. Fedorov, P. J. Finerty, Jr., G. A. Wasney, R. Yeung, C. Arrowsmith, L. J. Ball, H. Berglund, R. Hui, B. D. Marsden, P. Nordlund, M. Sundstrom, J. Weigelt and A. M. Edwards, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**, 15835-15840.
2. M. Mayer and B. Meyer, *Angew. Chem. Int. Ed.*, 1999, **38**, 1784-1788.
3. T. L. Hwang and A. J. Shaka, *J. Magn. Reson.* , 1995, **112**, 275-279.