Accepted Manuscript

Synthesis and evaluation of *N*6-substituted azide- and alkyne-bearing *N*-mustard analogs of *S*-adenosyl-L-methionine

Mohamed Ramadan , Natalie K. Bremner-Hay , Steig A. Carlson , Lindsay R. Comstock

PII: S0040-4020(14)00735-2

DOI: 10.1016/j.tet.2014.05.055

Reference: TET 25604

To appear in: *Tetrahedron*

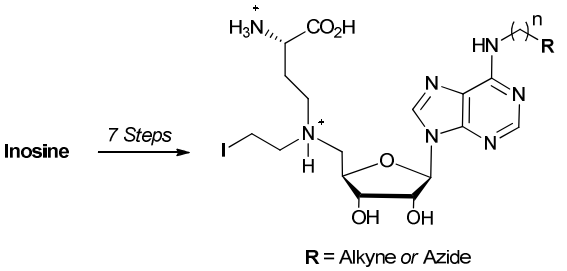
Received Date: 2 April 2014

Revised Date: 14 May 2014

Accepted Date: 15 May 2014

Please cite this article as: Ramadan M, Bremner-Hay NK, Carlson SA, Comstock LR, Synthesis and evaluation of *N*6-substituted azide- and alkyne-bearing *N*-mustard analogs of *S*-adenosyl-L-methionine, *Tetrahedron* (2014), doi: 10.1016/j.tet.2014.05.055.

Graphical Abstract



Synthesis and evaluation of *N*6-substituted azide- and alkyne-bearing *N*-mustard analogs of *S*-adenosyl-L-methionine

Mohamed Ramadan, Natalie K. Bremner-Hay, Steig A. Carlson and Lindsay R. Comstock*

Department of Chemistry, Wake Forest University, Winston-Salem, NC 27106, USA

*Corresponding author Tel.: 1-336-758-5514; e-mail address: comstolr@wfu.edu (L. Comstock).

Keywords:

S-adenosyl-L-methionine
DNA Methyltransferase
SAM analog
N-mustard

Abstract

The synthesis of a family of *N*-mustard analogs of *S*-adenosyl-L-methionine (SAM) containing azides and alkynes at the *N*6-position of the adenosine base has been accomplished from commercially available inosine. Further biochemical analysis of these analogs indicates successful modification of pUC19 plasmid DNA in an enzyme-dependent fashion with DNA methyltransferases M.TaqI and M.HhaI.

1. Introduction

Methylation of proteins, nucleic acids, polysaccharides, small molecules, and lipids plays a dynamic role in cellular processes.[1, 2] More importantly, DNA and protein methylation have been shown to play a critical role in the epigenetic control of gene activity. [3] To advance our understanding of how methylation events contribute to and alter cellular function, small molecules which mimic the native methyl donor, *S*-adenosyl-L-methionine (SAM), will hold tremendous promise by which to dissect and understand biological methylation. Our interests in

developing these tools are reflected in the synthesis and biochemical study of SAM-based analogs containing reactive functionalities capable of undergoing chemoselective ligations.[4, 5]

The design of SAM analogs containing these functionalities has evolved from work carried out by Weinhold *et al.* to generate simple analogs containing aziridines[6-8] to work by S. Rajski to incorporate azides capable of undergoing copper-catalyzed azide-alkyne cycloaddition (CuAAC or Click) chemistry or the Staudinger ligation.[9, 10] Pivotal to the advancement of such analogs as useful biochemical tools was the incorporation of the amino acid recognition moiety found in SAM (**1**, **Figure 1**). Incorporating the amino acid has proven to be a more effective SAM mimic compared to the aziridine-functionalized SAM analogs, as demonstrated by its efficient enzymatic incorporation onto DNA, small molecule, and peptide substrates.[11-13] More recently, work carried out in our laboratory has brought together the requisite amino acid moiety for improved enzyme-cofactor interactions, along with the reactive azides and alkynes at the *C*8-position, to generate *N*-mustard SAM analogs **2-4** (**Figure 1**).[14, 15] The utility of **2-4** has been demonstrated with a small panel of DNA methyltransferases to undergo subsequent ligation chemistries to generate fluorescently-tagged DNA products.

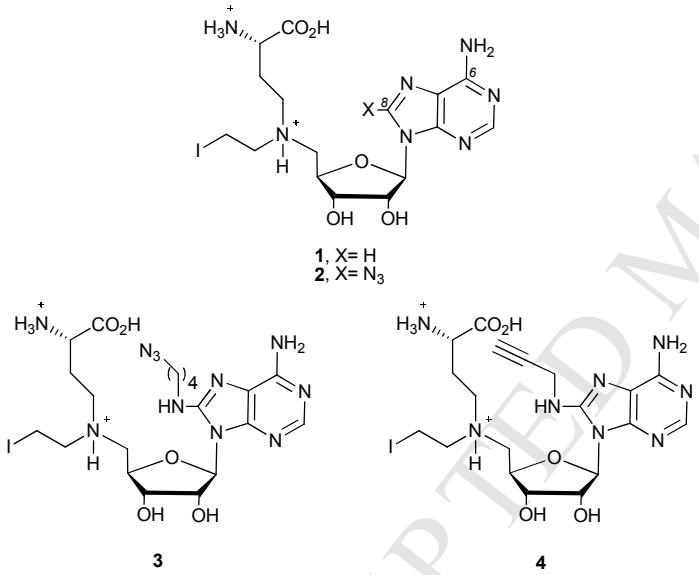


Fig. 1 *C*8-Substituted *N*-mustard Analogs of SAM.

While the *C*8-substituted analogs **2-4** have proven useful with a small panel of DNA methyltransferases, the accessibility of this position may be limited in a number of other methyltransferases, including protein methyltransferases.[16, 17] Analysis of known crystal structures indicates the 6-position of the adenosine base to be accessible and supports expansion of our small library of SAM analogs (**Figure 2**) to contain alkyne (**5a-c**) and azide (**6a-b**) handles at the *N*6 position of the adenine base. We anticipate that this family of SAM analogs may be accommodated by classes of methyltransferases (MTase) that do not tolerate *C*8-substitution and will be important biochemical tools for future studies in identifying both sites and substrates of biological methylation.

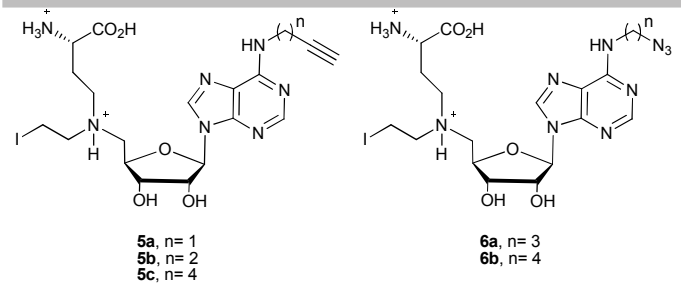


Fig. 2 *N6*-Substituted Azide- and Alkyne *N*-mustard Analogs of SAM.

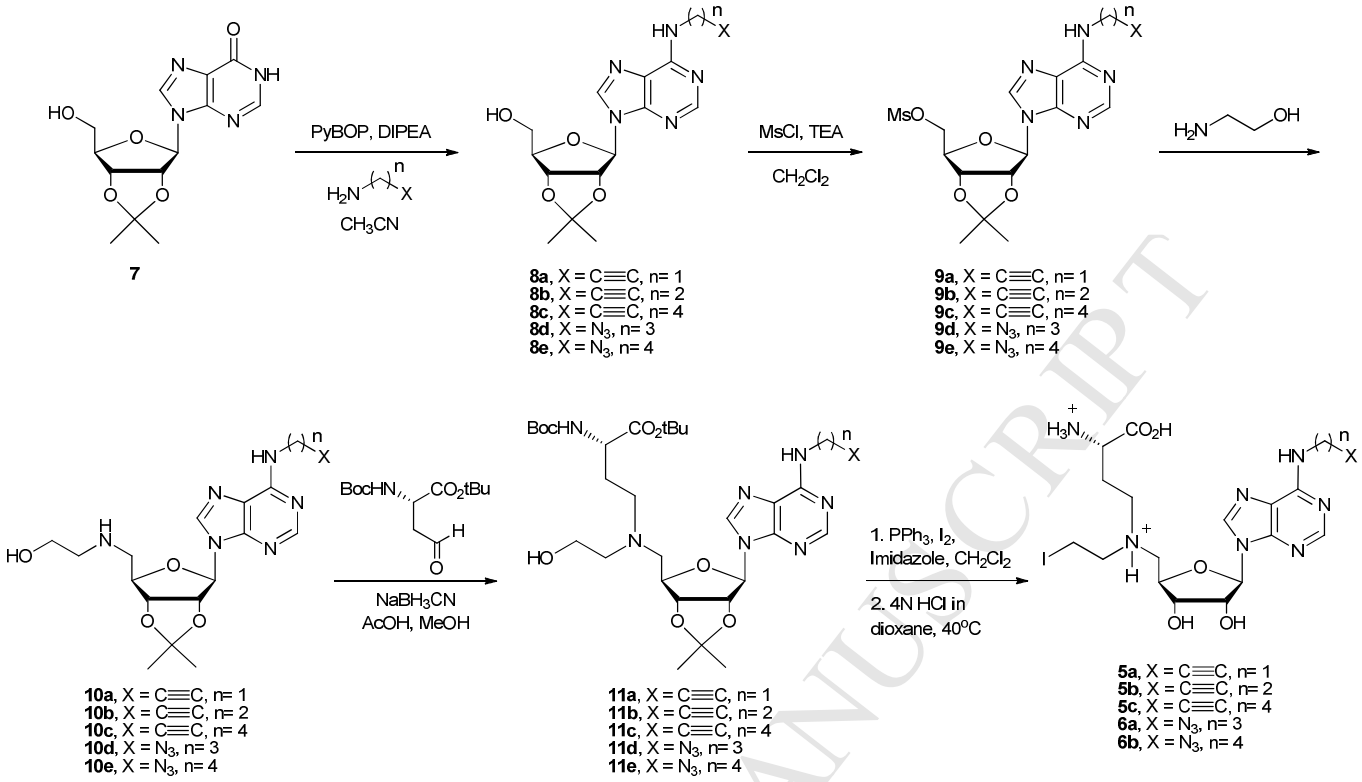
2. Results and Discussion

In the early stages of developing a synthetic pathway to obtain **5a**-**c** and **6a**-**b**, we sought to simplify and shorten the synthetic pathway to generate *N*-mustard analogs of SAM from previously reported methods.[14, 15] Although incorporation of the amino acid functionality would again be incorporated via reductive amination using previously described procedures, it was anticipated that production of the requisite ethanolamine intermediate on the ribose sugar could be easily improved upon. Ultimately, a procedure to aminate the 6-position of the adenine base to incorporate the desired alkyne and azide linkers was pivotal to the completion of analogs **5a**-**c** and **6a**-**b**.

Investigation of previously reported transformations that generate similar alkyl linkages at the 6-position of the adenine base indicated two potential starting points. The first strategy either began with or carried out a multi-step synthesis to generate 6-chloropurine riboside derivatives, followed by displacement of the chloride with primary amines, to yield the desired amination product.[18, 19] Alternatively, direct incorporation of primary amines to the 6-position through a one-step coupling with inosine using BOP derivatives has been shown to occur in high yields.[20, 21] Ultimately, we chose to proceed with a PyBOP coupling to inosine, as a significant reduction in the number of transformations to incorporate the alkyne and azide functionalities was required. While it was envisioned that this coupling to the inosine base could occur at multiple stages in the synthetic pathway, it was carried out early on to generate intermediates that were more non-polar and easier to manipulate synthetically.

Beginning with isopropylidene inosine[22], a small panel of alkyne- and azide-functionalized amines were coupled to the 6-position with PyBOP in the presence of DIPEA to generate *N*6-substituted adenosines **8a**-**e** in high yields (**Scheme 1**). Incorporation of the desired ethanolamine functionality was carried out in a two-step process via activation of the 5′ ribose alcohol with methanesulfonyl chloride to obtain mesylates **9a**-**e**, which were immediately displaced with ethanolamine to afford alcohols **10a**-**e** in moderate yields.[23] Subsequent incorporation of the amino acid functionality followed previously reported procedures via reductive amination with *tert*-Butyl (S)-2-[*N*-(tert-butoxycarbonyl)amino]-4-oxobutanoate[24] in the presence of $NaBH_3CN$ and acetic acid to provide the fully protected amino alcohols **11a**-**e**. To complete the synthesis, incorporation of the requisite iodine was directly followed by global deprotection with 4N HCl in dioxane,[14, 15] to provide the desired alkyne- and azide -substituted *N*-mustard SAM analogs, **5a**-**c** and **6a**-**b**. It is important to note that iodination of the alcohol requires the reagents

to be added in the proper order to prevent reduction of azides **11d** and **11e** to a primary amine by the triphenylphosphine.



Scheme 1. Synthesis of *N*6-substituted SAM analogs **5a-c** and **6a-b**.

Having synthesized the small panel of *N*6-functionalized *N*-mustard analogs of SAM, we next set out to investigate their ability to be enzymatically transferred to DNA in a site-specific manner. These experiments employed pUC19 plasmid DNA as the substrate, as it is considered a good model for a DNA MTase due to its large size and its closer resemblance to genomic DNA. The generality of the SAM analogs amongst various DNA MTases was assessed using a small panel of prokaryotic enzymes, including M.TaqI[25] and M.HhaI[26], which methylate adenine *N*6 and cytosine *C*5, respectively.

Enzymatic transfer of **5a-c** and **6a-b** to R.EcoRI-linearized plasmid DNA was carried out using the well-established restriction/ protection assay,[10] which exploits the resistance of DNA modified at the site corresponding to the MTase of interest to strand cleavage by its' restriction endonuclease. Experiments employing M.TaqI are shown in **Figure 3A** and **3B** and indicate successful transfer of **5a-c** and **6a-b** to pUC19. Examination of **Figure 3A** for the alkyne analogs **5a-c** demonstrate an increase in the amount of DNA alkylation with a gradient ranging from 0.1 to 10 μM (lanes 3–5, 7–9 and 11–13) for the three analogs, relative to the DNA controls in lanes 1 and 14, as reflected by the increased protection from R.TaqI digestion. At the highest concentration of analog in lanes 5, 9,

and 13, a significant portion of the plasmid DNA is full length and smaller restriction fragments almost completely disappear, indicating nearly quantitative transfer. In the absence of M.TaqI, protection of the DNA was not

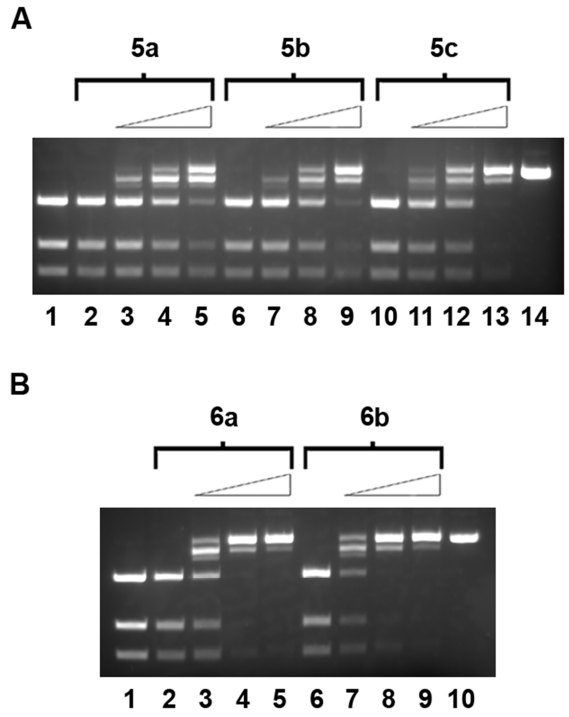


Figure 3. DNA alkylation reactions with M.TaqI. Restriction/protection assay was carried out with R.EcoRI-linearized pUC19 for analogs **5a-c** and **6a-b**. The extent of DNA modification was analyzed on a 2 % agarose gel containing ethidium bromide. **A**. Lane components: 1) DNA, R.TaqI; 2) DNA, **5a** (10 μM), R.TaqI; 3) DNA, **5a** (100 nM), M.TaqI, R.TaqI; 4) DNA, **5a** (1 μM), M.TaqI, R.TaqI; 5) DNA, **5a** (10 μM), M.TaqI, R.TaqI; 6-9) same as 2-5, but with **5b**; 10-13) same as 2-5, but with **5c**; 14) DNA. **B**. Lane components: 1) DNA, R.TaqI; 2) DNA, **6a** (10 μM), R.TaqI; 3) DNA, **6a** (100 nM), M.TaqI, R.TaqI; 4) DNA, **6a** (1 μM), M.TaqI, R.TaqI; 5) DNA, **6a** (10 μM), M.TaqI, R.TaqI; 6-9) same as 2-5, but with **6b**; 10) DNA.

observed (lanes 2, 6 and 10) and indicates that non- specific DNA alkylation by any of the analogs is not sufficient to protect TaqI-mediated plasmid scission. Similar trends were observed for the azide analogs, as shown in **Figure 3B**. Nearly quantitative transfer was observed, as an increase in DNA alkylation (lanes 3–5 and 7–9) with a gradient of **6a** and **6b**, relative to the controls in lanes 1 and 10. Again, no protection of the DNA was detected (lanes 2 and 6) in the absence of M.TaqI. The results demonstrated here agrees with the previous report of M.TaqI successfully transferring a Cy3 functionalized aziridine SAM analog linked to the 6-position of the adenine base with plasmid DNA.[27]

Having demonstrated successful transfer of the SAM analogs to plasmid DNA with M.TaqI, the generality of **5a-c** and **6a-b** to serve as cofactors for another DNA MTase was explored. Specifically, M.HhaI-mediated transfer of **5a-c** and **6a-b** to pUC19 was carried out. Reaction analysis using the restriction/protection assay demonstrated

successful incorporation of **5a-c** and **6a-b** to plasmid DNA (see **Figure S1**) and supports the ability of M.HhaI to accommodate substitution at the *N*6 position. These results reveal for the first time the ability of M.HhaI to bind and catalyze the transfer of SAM analogs substituted at the 6-position, indicating its flexibility in accommodating substitution at multiple positions on the adenine base.

3. Conclusions

In conclusion, we have successfully synthesized a small library of *N*6 functionalized alkyne and azide *N*-mustard analogs (**5a-c** and **6a-b**) of SAM to expand our chemical toolbox to study biological methylation using native methyltransferases. The synthetic route taken here significantly improved on those previously reported for **2**-**4**, as the number of required transformations is reduced to 7 steps beginning from inosine (compared to 12 and 11 steps for **3** and **4**, respectively). Due to the efficiency and simplicity of the synthetic transformations described here, it is feasible that such steps may be used in future efforts to not only shorten and improve the syntheses of **2**-**4**, but to also generate additional SAM analogs bearing reactive functionalities.

The biological utility of **5a-c** and **6a-b** was further demonstrated by their near quantitative MTase-dependent modification of plasmid DNA. Plasmid DNA experiments indicated that both M.TaqI and M.HhaI were able to tolerate substitution at the *N*6 position of the adenine base. Based on these findings, future experiments to investigate additional MTases, including the mammalian DNA MTase DNMT1 and protein MTase PRMT1, to accommodate and catalyze the transfer of functionalized *N*-mustard analogs must be carried out. Such experiments will be essential to validate the SAM analogs described here as biochemical probes of DNA and protein methylation for future cell-based studies.

4. Experimental

4.1 General Methods

All reagents and solvents were purchased from commercial sources and used without additional purification. Anhydrous solvents were obtained from a Meyer solvent system. Reactions were performed at room temperature under argon, unless otherwise indicated. ^{1}H NMR spectra were recorded at either 300 or 500 MHz; ^{13}C NMR spectra were recorded at either 75 or 125 MHz, with the solvent peak used as the internal standard. 2D NMR spectra were recorded at 500 MHz. Chemical shifts are reported in δ (ppm). Analytical and semi-preparative HPLC procedures are described in Supporting Information.

For the purposes of biological experiments, solutions of **5a-c** and **6a-b** were made up using H_2SO_4 (2.5 mM), and the concentrations of these samples were determined by UV/Vis spectroscopy. By using the absorption values obtained at 260 nm and a molar extinction constant of 9710 $M^{-1}cm^{-1}$ (for **5a**), 10970 $M^{-1}cm^{-1}$ (for **5b**), 11270 M^{-}

$^{1}cm^{-1}$ (for **5c**), 12370 $M^{-1}cm^{-1}$ (for **6a**), and 11530 $M^{-1}cm^{-1}$ (for **6b**), stock solution concentrations were calculated using Beer's law. Stock solutions were stored at -80 °C and used within 2 weeks.

pUC19, all enzymes, and their corresponding buffers were obtained from New England Biolabs. All agarose gels were prepared with a high-melt agarose in 1X TAE. Ethidium bromide images were obtained with a ChemiDoc MP system (Bio-Rad).

4.2 General Procedure for PyBOP Coupling

N,*N*-Diisopropylethylamine (1.04 mL, 6.0 mmol) was added to a solution of 2′,3′-*O*-Isopropylidene inosine (**7**)[22] (0.925 g, 3.0 mmol) and PyBOP (2.341 g, 4.5 mmol) in 6.0 mL HPLC grade acetonitrile and stirred until the solution turned clear (15 min). The required alkyne or azide amine (6.0 mmol) was added drop wise and the reaction was stirred overnight. The reaction was diluted with Et_2O and washed twice each with 1M acetic acid, saturated $NaHCO_3$, and brine. The organic layer was dried over Na_2SO_4 and concentrated. The crude residue was purified by column chromatography (silica gel, 4:1 to 1:2 Hexane:EtOAc).

4.2.1 *N*6-(Prop-2"-ynyl)-2',3'-bis-(*O*-isopropylidene) adenosine (**8a**). The reaction was carried out with propargylamine to obtain **8a** (0.90 g, 87%) as a white foam. MP 134-137°C. ^{1}H NMR (CD_3OD) δ 8.30 (s, 2H), 6.17 (d, *J* = 3.5 Hz, 1H), 5.29 (dd, *J* = 6.1, 3.5 Hz, 1H), 5.05 (dd, *J* = 6.1, 2.4 Hz, 1H), 4.43-4.37 (m, 3H), 3.81 (dd, *J* = 12.1, 3.5 Hz, 1H), 3.73 (dd, *J* = 12.1, 3.9 Hz, 1H), 2.62 (t, *J* = 2.5 Hz, 1H), 1.63 (s, 3H), 1.39 (s, 3H); ^{13}C NMR (CD_3OD) δ 155.6, 153.7, 149.6, 141.6, 121.3, 115.3, 92.9, 88.1, 85.3, 83.0, 81.1, 72.2, 63.6, 30.9, 27.6, 25.5. HRMS-ESI: calcd for $C_{16}H_{19}N_5O_4$ ($M + Na^+$) 368.1329, obsd 368.1330.

4.2.2 *N*6-(But-3"-ynyl)-2',3'-bis-(*O*-isopropylidene) adenosine (**8b**). The reaction was carried out with 3-butyn-1-amine to obtain **8b** (0.90 g, 84%) as a white foam. MP 50-52°C. ^{1}H NMR (CD_3OD) δ 8.27 (s, 1H), 8.25 (s, 1H), 6.14 (d, *J* = 3.6, 1H), 5.27 (dd, *J* = 6.1, 3.6 Hz, 1H), 5.04 (dd, *J* = 6.1, 2.3 Hz, 1H), 4.40-4.36 (m, 1H), 3.82-3.68 (m, 4H), 2.57 (td, *J* = 7.0, 2.7 Hz, 2H), 2.32 (t, *J* = 2.6 Hz, 1H), 1.62 (s, 3H), 1.38 (s, 3H); ^{13}C NMR (CD_3OD) δ 156.0, 153.7, 149.3, 141.3, 121.1, 115.2, 92.9, 88.0, 85.2, 82.9, 82.3, 71.0, 63.6, 40.7, 27.6, 25.5, 20.0. HRMS-ESI: calcd for $C_{17}H_{21}N_5O_4$ ($M + Na^+$) 382.1486, obsd 382.1487.

4.2.3 *N*6-(Hex-5"-ynyl)-2',3'-bis-(*O*-isopropylidene) adenosine (**8c**). The reaction was carried out with 5-hexyn-1-amine to obtain **8c** (1.05 g, 90%) as a white foam. MP 78-80°C. ^{1}H NMR (CD_3OD) δ 8.26 (s, 1H), 8.22 (s, 1H), 6.14 (d, *J* = 3.6 Hz, 1H), 5.27 (dd, *J* = 6.1, 3.6 Hz, 1H), 5.04 (dd, *J* = 6.1, 2.3 Hz, 1H), 4.39-4.36 (m, 1H), 3.80 (dd, *J* = 12.2, 3.4 Hz, 1H), 3.71 (dd, *J* = 12.2, 3.9 Hz, 1H), 3.62 (bs, 2H), 2.27-2.19 (m, 3H), 1.83-1.76 (m, 2H), 1.68-1.63 (m, 2H), 1.61 (s, 3H), 1.38 (s, 3H); ^{13}C NMR (CD_3OD) δ 156.3, 153.8, 149.1, 141.1, 121.0, 115.2, 92.9, 88.0, 85.2, 84.7, 83.0, 69.8, 63.6, 41.1, 29.7, 27.6, 27.0, 25.6, 18.8. HRMS-ESI: calcd for $C_{19}H_{25}N_5O_4$ ($M + Na^+$) 410.1799, obsd 410.1800.

4.2.4 *N*6-(3"-azidopropyl)-2',3'-bis-(*O*-isopropylidene) adenosine (**8d**). The reaction was carried out with 3-azidopropan-1-amine[16] to obtain **8d** (1.02 g, 87%) as a light yellow oil. ^{1}H NMR (CD_3OD) δ 8.26 (s, 1H), 8.24 (s, 1H), 6.13 (d, *J* = 3.6 Hz, 1H), 5.26 (dd, *J* = 6.1, 3.6 Hz, 1H), 5.03 (dd, *J* = 6.1, 2.3 Hz, 1H), 4.39-4.35 (m, 1H), 3.80 (dd, *J* = 12.2, 3.4 Hz, 1H), 3.74-3.68 (m, 3H), 3.44 (t, *J* = 6.7 Hz, 2H), 1.94 (p, *J* = 6.7 Hz, 2H), 1.61 (s, 3H), 1.37 (s, 3H); ^{13}C NMR (CD_3OD) δ 156.4, 153.8, 141.2, 121.1, 115.3, 92.9, 88.0, 85.3, 83.0, 63.6, 50.1, 29.9, 27.6, 25.6; additional carbons identified by HMBC, ($CDCl_3$) δ 147.2, 37.9. HRMS-ESI: calcd for $C_{16}H_{22}N_8O_4$ ($M + Na^+$) 413.1656, obsd 413.1657.

4.2.5 *N*6-(4"-azidobutyl)-2',3'-bis-(O-isopropylidene) adenosine (**8e**). The reaction was carried out with 4-azidobutan-1-amine[16] to obtain **8e** (1.10 g, 91%) as a yellow oil. ^{1}H NMR ($CDCl_3$) δ 8.32 (s, 1H), 7.79 (s, 1H), 6.77 (bs, 1H), 6.15 (bt, *J* = 5.5 Hz, 1H), 5.85 (d, *J* = 4.9 Hz, 1H), 5.23-5.19 (m, 1H), 5.12 (dd, *J* = 6.0, 1.2 Hz, 1H), 4.55-4.53 (m, 1H), 3.98 (dd, *J* = 12.8, 1.6 Hz, 1H), 3.81-3.76 (m, 1H), 3.69 (bs, 2H), 3.34 (t, *J* = 6.3 Hz, 2H), 1.80-1.68 (m, 4H), 1.64 (s, 3H), 1.38 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.4, 152.8, 147.5, 139.6, 121.3, 114.1, 94.5, 86.2, 83.1, 81.8, 63.5, 51.2, 40.1, 27.8, 27.1, 26.4, 25.4. HRMS-ESI: calcd for $C_{17}H_{24}N_8O_4$ ($M + Na^+$) 427.1813, obsd 427.1814.

4.3 General Procedure for Mesylation

To a solution of the *N*6 substituted adenosine (1.5 mmol) in 7.5 mL CH_2Cl_2 was added triethylamine (0.50 mL, 3.6 mmol) and methanesulfonyl chloride (140 μL, 1.8 mmol). The reaction was stirred for 1 h and then quenched with saturated NH_4Cl. The layers were separated and the organic layer was washed twice with brine. The combined aqueous layers were extracted twice with CH_2Cl_2, dried over Na_2SO_4 and concentrated. Due to the inherent reactivity of the mesylate, melting point and HRMS was not performed (LRMS is provided). The crude residue was purified by column chromatography (silica gel, 50:1 CH_2Cl_2:MeOH).

4.3.1 *N*6-(Prop-2"-ynyl)-5'-methanesulfonate-2',3'-bis-(*O*-isopropylidene) adenosine (**9a**). The reaction was carried out with **8a** to give mesylate **9a** (0.552 g, 87 %) as a white foam. ^{1}H NMR ($CDCl_3$) δ 8.42 (s, 1H), 7.89 (s, 1H), 6.31 (t, *J* = 5.6 Hz, 1H), 6.12 (d, *J* = 2.0 Hz, 1H), 5.47 (dd, *J* = 6.4, 2.1 Hz, 1H), 5.16 (dd, *J* = 6.3, 3.0 Hz, 1H), 4.54-4.38 (m, 5H), 2.89 (s, 3H), 2.27 (t, *J* = 2.6 Hz, 1H), 1.61 (s, 3H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 154.3, 153.3, 148.7, 139.9, 120.7, 114.9, 91.0, 84.9, 84.1, 81.5, 80.2, 71.7, 68.6, 37.6, 30.5, 27.2, 25.4. LRMS-ESI: calcd for $C_{17}H_{21}N_5O_6S$ ($M + H^+$) 424.1, obsd 424.2.

4.3.2 *N*6-(But-3"-ynyl)-5'-methanesulfonate-2',3'-bis-(*O*-isopropylidene) adenosine (**9b**). The reaction was carried out with **8b** to give mesylate **9b** (0.551 g, 84 %) as a white foam. ^{1}H NMR ($CDCl_3$) δ 8.37 (s, 1H), 7.86 (s, 1H), 6.20-6.16 (m, 1H), 6.11 (d, *J* = 2.1 Hz, 1H), 5.48 (dd, *J* = 6.3, 2.1 Hz, 1H), 5.17 (dd, *J* = 6.4, 3.0 Hz, 1H), 4.53-4.41

(m, 3H), 3.85 (bs, 2H), 2.91 (s, 3H), 2.60 (td, *J* = 6.5, 2.6 Hz, 2H), 2.04 (t, *J* = 2.6 Hz, 1H), 1.62 (s, 3H), 1.40 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 154.8, 153.3, 148.3, 139.5, 120.6, 114.7, 90.9, 84.8, 84.1, 81.6, 81.5, 70.4, 68.6, 39.3, 37.5, 27.1, 25.3, 19.6. LRMS-ESI: calcd for $C_{18}H_{23}N_5O_6S$ ($M + H^+$) 438.1, obsd 438.2.

4.3.3 *N*6-(Hex-5''-ynyl)-5'-methanesulfonate-2',3'-bis-(*O*- isopropylidene) adenosine (**9c**). The reaction was carried out with **8c** to give mesylate **9c** (0.621 g, 89 %) as a light yellow foam. ^{1}H NMR ($CDCl_3$) δ 8.36 (s, 1H), 7.83 (s, 1H), 6.11 (d, *J* = 2.0 Hz, 1H), 5.82 (bt, *J* = 6.0 Hz, 1H), 5.48 (dd, *J* = 6.3, 2.0 Hz, 1H), 5.18-5.15 (m, 1H), 4.55-4.39 (m, 3H), 3.69 (bs, 2H), 2.90 (s, 3H), 2.27 (td, *J* = 7.0, 2.7 Hz, 2H), 1.97 (t, *J* = 2.7 Hz, 1H), 1.87-1.78 (m, 2H), 1.71-1.63 (m, 2H), 1.62 (s, 3H), 1.40 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 154.9, 153.1, 147.9, 139.1, 120.1, 114.5, 90.6, 84.6, 84.0, 83.9, 81.2, 68.7, 68.6, 39.9, 37.2, 28.6, 26.9, 25.5, 25.1, 17.9. LRMS-ESI: calcd for $C_{20}H_{27}N_5O_6S$ ($M + H^+$) 466.2, obsd 466.2.

4.3.4 *N*6-(3''-azidopropyl)-5'-methanesulfonate-2',3'-bis-(*O*-isopropylidene) adenosine (**9d**). The reaction was carried out with **8d** to give mesylate **9d** (0.667 g, 95 %) as a light yellow foam. ^{1}H NMR ($CDCl_3$) δ 8.37 (s, 1H), 7.86 (s, 1H), 6.12-6.09 (m, 2H), 5.48 (dd, *J* = 6.3, 2.1 Hz, 1H), 5.17 (dd, *J* = 6.2, 2.9 Hz, 1H), 4.55-4.40 (m, 3H), 3.78 (bs, 2H), 3.45 (t, *J* = 6.6 Hz, 2H), 2.91 (s, 3H), 1.98 (p, *J* = 6.6 Hz, 2H), 1.62 (s, 3H), 1.40 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.1, 153.4, 148.3, 139.4, 120.5, 114.8, 91.0, 84.9, 84.1, 81.5, 68.7, 49.2, 38.1, 37.6, 29.1, 27.2, 25.4. LRMS-ESI: calcd for $C_{17}H_{24}N_8O_6S$ ($M + H^+$) 469.2, obsd 469.2.

4.3.5 *N*6-(4''-azidobutyl)-5'-methanesulfonate-2',3'-bis-(*O*-isopropylidene) adenosine (**9e**). The reaction was carried out with **8e** to give mesylate **9e** (0.607 g, 84 %) as a light yellow foam. ^{1}H NMR ($CDCl_3$) δ 8.29 (s, 1H), 7.77 (s, 1H), 6.04 (d, *J* = 2.0 Hz, 1H), 5.80 (bt, *J* = 6.0 Hz, 1H), 5.41 (dd, *J* = 6.4, 2.1 Hz, 1H), 5.10 (dd, *J* = 6.5, 2.9 Hz, 1H), 4.48-4.32 (m, 3H), 3.64 (bs, 2H), 3.28 (t, *J* = 6.4 Hz, 2H), 2.84 (s, 3H), 1.77-1.59 (m, 4H), 1.55 (s, 3H), 1.33 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.1, 153.4, 148.1, 139.3, 120.5, 114.8, 91.0, 84.9, 84.1, 81.5, 68.7, 51.2, 40.1, 37.6, 27.2, 27.1, 26.3, 25.4. LRMS-ESI: calcd for $C_{18}H_{26}N_8O_6S$ ($M + H^+$) 483.2, obsd 483.2.

4.4 General Procedure for Ethanolamine Displacement

The *N*6 substituted mesylate (1.0 mmol) was brought up in 3 mL of 2-aminoethanol (50 mmol) and stirred for 3 days. The volatiles were evaporated off under high vacuum with gentle heating using a distillation apparatus. Upon addition of brine to the resulting material, the product was extracted with EtOAc, washed with brine, dried over Na_2SO_4 and concentrated. The crude residue was purified by column chromatography (silica gel, 9:1 to 7:3 CH_2Cl_2:MeOH).

4.4.1 *N*6-(Prop-2''-ynyl)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**10a**). The reaction was carried out with **9a** to obtain **10a** as a white foam (0.271 g, 70 %). MP 74-76°C. ^{1}H NMR ($CDCl_3$) δ 8.40 (s,

1H), 7.89 (s, 1H), 6.57 (bs, 1H), 5.99 (d, *J* = 3.1 Hz, 1H), 5.47 (dd, *J* = 6.4, 3.0 Hz, 1H), 5.03 (dd, *J* = 6.4, 3.3 Hz, 1H), 4.45 (bs, 2H), 4.38-4.33 (m, 1H), 3.63 (t, *J* = 5.1 Hz, 2H), 2.95 (dd, *J* = 12.6, 4.6 Hz, 1H), 2.90 (dd, *J* = 12.6, 6.2 Hz, 1H), 2.83-2.68 (m, 2H), 2.50 (bs, 2H), 2.26 (t, *J* = 2.5 Hz, 1H), 1.61 (s, 3H), 1.38 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 154.2, 153.3, 148.9, 139.9, 120.8, 114.7, 91.1, 85.8, 83.6, 82.4, 80.2, 71.6, 61.0, 51.3, 51.1, 30.6, 27.4, 25.5. HRMS-ESI: calcd for $C_{18}H_{24}N_6O_4$ ($M + Na^+$) 411.1751, obsd 411.1753.

4.4.2 *N*6-(But-3''-ynyl)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**10b**). The reaction was carried out with **9b** to obtain **10b** as a white gum (0.273 g, 68 %). ^{1}H NMR ($CDCl_3$) δ 8.35 (s, 1H), 7.87 (s, 1H), 6.38 (bt, *J* = 6.6 Hz, 1H), 6.00 (d, *J* = 3.0 Hz, 1H), 5.44 (dd, *J* = 6.5, 3.0 Hz, 1H), 5.04 (dd, *J* = 6.4, 3.4 Hz, 1H), 4.42-4.37 (m, 1H), 3.81 (bs, 2H), 3.65 (t, *J* = 5.1 Hz, 2H), 3.47 (bs, 2H), 3.09-2.98 (m, 2H), 2.90-2.75 (m, 2H), 2.58 (td, *J* = 6.6, 2.7 Hz, 2H), 2.05 (t, *J* = 2.6 Hz, 1H), 1.61 (s, 3H), 1.38 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 154.8, 153.3, 139.6, 120.5, 114.9, 91.0, 85.0, 83.6, 82.3, 81.7, 70.4, 60.3, 51.0, 50.7, 27.4, 25.5, 19.7; additional carbons identified by HMBC, ($CDCl_3$) δ 148.5, 39.3. HRMS-ESI: calcd for $C_{19}H_{26}N_6O_4$ ($M + Na^+$) 425.1908, obsd 425.1911.

4.4.3 *N*6-(Hex-5''-ynyl)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**10c**). The reaction was carried out with **9c** to obtain **10c** as a white gum (0.245 g, 57%). ^{1}H NMR ($CDCl_3$) δ 8.35 (s, 1H), 7.82 (s, 1H), 5.98 (d, *J* = 3.1 Hz, 1H), 5.85 (bs, 1H), 5.49 (dd, *J* = 6.5, 3.1 Hz, 1H), 5.06 (dd, *J* = 6.5, 3.2 Hz, 1H), 4.39-4.34 (m, 1H), 3.74-3.66 (m, 2H), 3.63 (t, *J* = 5.2 Hz, 2H), 2.97 (dd, *J* = 12.7, 4.6 Hz, 1H), 2.92 (dd, *J* = 12.6, 5.9 Hz, 1H), 2.86-2.70 (m, 2H), 2.42 (bs, 2H), 2.26 (td, *J* = 7.0, 2.6 Hz, 2H), 1.96 (t, *J* = 2.6 Hz, 1H), 1.87-1.77 (m, 2H), 1.71-1.63 (m, 2H), 1.61 (s, 3H), 1.38 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.2, 153.4, 139.4, 120.7, 114.7, 91.2, 85.6, 84.1, 83.5, 82.4, 68.9, 60.9, 51.3, 51.0, 28.9, 27.5, 25.8, 25.6, 18.3; additional carbons identified by HMBC, ($CDCl_3$) δ 148.2, 40.0. HRMS-ESI: calcd for $C_{21}H_{30}N_6O_4$ ($M + Na^+$) 453.2221, obsd 453.2223.

4.4.4 *N*6-(3''-azidopropyl)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**10d**). The reaction was carried out with **9d** to obtain **10d** as a yellow oil (0.260 g, 60 %). ^{1}H NMR ($CDCl_3$) δ 8.36 (s, 1H), 7.85 (s, 1H), 6.13 (bt, *J* = 5.9 Hz, 1H), 5.99 (d, *J* = 3.1 Hz, 1H), 5.49 (dd, *J* = 6.4, 3.1 Hz, 1H), 5.05 (dd, *J* = 6.4, 3.2 Hz, 1H), 4.40-4.36 (m, 1H), 3.79-3.72 (m, 2H), 3.64 (t, *J* = 5.1 Hz, 2H), 3.45 (t, *J* = 6.6 Hz, 2H), 3.00-2.88 (m, 2H), 2.86-2.70 (m, 2H), 2.52 (bs, 2H), 1.97 (p, *J* = 6.7 Hz, 2H), 1.62 (s, 3H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.1, 153.4, 139.6, 120.7, 114.7, 91.2, 85.6, 83.5, 82.4, 61.0, 51.3, 51.0, 49.2, 29.1, 27.4, 25.5; additional carbons identified by HMBC, ($CDCl_3$) δ 148.3, 38.0. HRMS-ESI: calcd for $C_{18}H_{27}N_9O_4$ ($M + Na^+$) 456.2078, obsd 456.2080.

4.4.5 *N*6-(4''-azidobutyl)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**10e**). The reaction was carried out with **9e** to obtain **10e** as a yellow oil (0.290 g, 65 %). ^{1}H NMR ($CDCl_3$) δ 8.36 (s, 1H), 7.84 (s, 1H),

5.98 (d, *J* = 3.2 Hz, 1H), 5.92 (bs, 1H), 5.49 (dd, *J* = 6.5, 3.2 Hz, 1H), 5.06 (dd, *J* = 6.5, 3.3 Hz, 1H), 4.40-4.35 (m, 1H), 3.70 (bs, 2H), 3.63 (t, *J* = 5.2 Hz, 2H), 3.35 (t, *J* = 6.3 Hz, 2H), 2.99-2.90 (m, 2H), 2.88-2.70 (m, 2H), 2.36 (bs, 2H), 1.82-1.65 (m, 4H), 1.62 (s, 3H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 155.2, 153.4, 139.5, 120.7, 114.7, 91.2, 85.7, 83.5, 82.4, 61.0, 51.3, 51.3, 51.1, 27.5, 27.2, 26.4, 25.6; additional carbons identified by HMBC, ($CDCl_3$) δ 148.3, 39.9. HRMS-ESI: calcd for $C_{19}H_{29}N_9O_4$ (M + Na^+) 470.2235, obsd 470.2236.

4.5 General Procedure for Reductive Amination

$NaBH_3CN$ (0.048 g, 0.771 mmol) and glacial acetic acid (29.4 μL, 0.514 mmol) were added to a solution of the *N*6-substituted *N*-hydroxyethyl adenosine (0.566 mmol) and *tert*-butyl (*S*)-2-[*N*-(*tert*-butoxycarbonyl)amino]-4-oxobutanoate[24] (0.1405 g, 0.514 mmol) in 2 mL dry MeOH. The reaction mixture was stirred overnight. After diluting the reaction with EtOAc and saturated $NaHCO_3$, the organic layer was washed with saturated $NaHCO_3$, dried over Na_2SO_4 and concentrated. The crude residue was purified by column chromatography (silica gel, 8:4:2:1 Pet Ether/EtOAc/CH_2Cl_2/MeOH).

4.5.1 *N*6-(Prop-2''-ynyl)-5'-(*N*α-Boc-diaminobutyric acid *O*-*tert*-butyl ester)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**11a**). The reaction was carried out with **10a** to yield **11a** (0.250 g, 68 %) as a white foam. MP 64-68°C. ^{1}H NMR ($CDCl_3$) δ 8.42 (s, 1H), 7.88 (s, 1H), 6.04 (d, *J* = 2.1 Hz, 1H), 6.01 (bs, 1H), 5.45 (dd, *J* = 6.5, 2.0 Hz, 1H), 5.39 (d, *J* = 8.2 Hz 1H), 5.06 (dd, *J* = 6.5, 4.0 Hz, 1H), 4.51-4.45 (bm, 2H), 4.31 (td, *J* = 6.8, 4.0 Hz, 1H), 4.24-4.18 (bm, 1H), 3.58-3.51 (m, 1H), 3.46-3.39 (m, 1H), 2.84 (dd, *J* = 13.5, 6.8 Hz, 1H), 2.71-2.47 (m, 5H), 2.27 (t, *J* = 2.5 Hz, 1H), 1.99-1.91 (m, 1H), 1.74-1.64 (m, 2H), 1.61 (s, 3H), 1.45 (s, 18H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 171.9, 155.6, 154.3, 153.2, 148.9, 140.1, 120.8, 114.8, 90.6, 85.5, 84.1, 83.5, 82.0, 80.2, 79.9, 71.6, 59.5, 57.4, 56.5, 52.8, 51.3, 30.6, 30.0, 28.5, 28.1, 27.3, 25.6. HRMS-ESI: calcd for $C_{31}H_{47}N_7O_8$ (M + Na^+) 668.3378, obsd 668.3376.

4.5.2 *N*6-(But-3''-ynyl)-5'-(*N*α-Boc-diaminobutyric acid *O*-*tert*-butyl ester)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**11b**). The reaction was carried out with **10b** to yield **11b** (0.283 g, 76 %) as a white foam. MP 60-62°C. ^{1}H NMR ($CDCl_3$) δ 8.35 (s, 1H), 7.85 (s, 1H), 6.15 (bs, 1H), 6.02 (d, *J* = 2.1 Hz, 1H), 5.44-5.41 (m 2H), 5.05 (dd, *J* = 6.5, 4.0 Hz, 1H), 4.29 (td, *J* = 6.8, 4.0 Hz, 1H), 4.24-4.20 (bm, 1H), 3.82 (bs, 2H), 3.55-3.52 (m, 1H), 3.44-3.41 (m, 1H), 2.85 (dd, *J* = 13.4, 6.8 Hz, 1H), 2.69-2.49 (m, 8H), 2.03 (t, *J* = 2.6 Hz, 1H), 1.99-1.94 (m, 1H), 1.71-1.66 (m, 1H), 1.60 (s, 3H), 1.44 (s, 9H), 1.43 (s, 9H), 1.37 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 172.0, 155.7, 154.9, 153.3, 139.8, 114.8, 90.5, 85.5, 84.1, 83.5, 82.1, 81.6, 79.9, 70.4, 59.4, 57.4, 56.3, 52.7, 51.0, 30.0, 28.5, 28.1, 27.3, 25.6, 19.8; additional carbons identified by HMBC, ($CDCl_3$) δ 148.2, 120.5, 39.2. HRMS-ESI: calcd for $C_{32}H_{49}N_7O_8$ (M + Na^+) 682.3535, obsd 682.3535.

4.5.3 *N*6-(Hex-5''-ynyl)-5'-(*N*α-Boc diaminobutyric acid *O*-*tert*-butyl ester)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**11c**). The reaction was carried out with **10c** to yield **11c** (0.263 g, 68 %) as a white foam. ^{1}H NMR ($CDCl_3$) δ 8.34 (s, 1H), 7.82 (s, 1H), 6.02 (d, *J* = 2.0 Hz, 1H), 5.84 (bs, 1H), 5.44-5.41 (m, 2H), 5.05 (dd, *J* = 6.5, 4.0 Hz, 1H), 4.29 (td, *J* = 6.8, 4.0 Hz, 1H), 4.24-4.20 (bm, 1H), 3.67 (bs, 2H), 3.56-3.52 (m, 1H), 3.43-3.40 (m, 1H), 2.87-2.83 (m, 1H), 2.69-2.49 (m, 5H), 2.25 (td, *J* = 7.0, 2.7 Hz, 2H), 2.00-1.94 (m, 2H), 1.84-1.78 (m, 2H), 1.71-1.62 (m, 3H), 1.60 (s, 3H), 1.44 (s, 9H), 1.43 (s, 9H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 172.0, 155.7, 155.1, 153.4, 139.5, 120.6, 114.7, 90.5, 85.4, 83.5, 82.1, 79.9, 68.9, 59.4, 57.4, 56.3, 52.7, 51.0, 29.9, 28.9, 28.5, 28.1, 27.3, 25.8, 25.6, 18.3; additional carbons identified by HMBC, ($CDCl_3$) δ 148.0, 83.90, 83.87, 40.0. HRMS-ESI: calcd for $C_{34}H_{53}N_7O_8$ ($M + Na^+$) 710.3848, obsd 710.3855.

4.5.4 *N*6-(3''-azidopropyl)-5'-(*N*α-Boc diaminobutyric acid *O*-*tert*-butyl ester)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**11d**). The reaction was carried out with **10d** to yield **11d** (0.278 g, 71 %) as a white foam. MP 52-54°C. ^{1}H NMR ($CDCl_3$) δ 8.37 (s, 1H), 7.85 (s, 1H), 6.03 (d, *J* = 2.1 Hz, 1H), 5.95 (bs, 1H), 5.46-5.41 (m, 2H), 5.05 (dd, *J* = 6.6, 4.0 Hz, 1H), 4.33-4.28 (m, 1H), 4.25-4.20 (bm, 1H), 3.81-3.73 (bm, 2H), 3.59-3.52 (m, 1H), 3.48-3.41 (m, 3H), 3.01 (bs, 1H), 2.89-2.82 (m, 1H), 2.71-2.47 (m, 5H), 2.02-1.92 (m, 3H), 1.76-1.66 (m, 1H), 1.61 (s, 3H), 1.45 (s, 18H), 1.39 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 172.0, 155.7, 155.1, 153.3, 139.7, 120.7, 114.8, 90.6, 85.7, 84.1, 83.5, 82.1, 79.9, 59.4, 57.4, 56.3, 52.7, 51.1, 49.3, 30.0, 29.2, 28.5, 28.1, 27.3, 25.6; additional carbons identified by HMBC, ($CDCl_3$) δ 148.2, 38.0. HRMS-ESI: calcd for $C_{31}H_{50}N_{10}O_8$ ($M + Na^+$) 713.3705, obsd 713.3709.

4.5.5 *N*6-(4''-azidobutyl)-5'-(*N*α-Boc diaminobutyric acid *O*-*tert*-butyl ester)-5'-*N*-hydroxyethyl-5'-deoxy-2',3'-bis-(*O*-isopropylidene) adenosine (**11e**). The reaction was carried out with **10e** to yield **11e** (0.259 g, 65 %) as a white foam. MP 50-52 °C. ^{1}H NMR ($CDCl_3$) δ 8.34 (s, 1H), 7.83 (s, 1H), 6.02 (bd, *J* = 2.0 Hz, 1H), 5.91 (bs, 1H), 5.44-5.41 (bm, 2H), 5.05 (dd, *J* = 6.5, 4.0 Hz, 1H), 4.29 (td, *J* = 6.8, 4.0 Hz, 1H), 4.24-4.20 (bm, 1H), 3.68 (bs, 2H), 3.55-3.51 (m, 1H), 3.43-3.40 (m, 1H), 3.34 (t, *J* = 6.7 Hz, 2H), 2.84 (dd, *J* = 13.6, 7.0 Hz, 1H), 2.69-2.48 (m, 6H), 1.99-1.93 (m, 1H), 1.80-1.67 (m, 5H), 1.60 (s, 3H), 1.44 (s, 9H), 1.43 (s, 9H), 1.37 (s, 3H); ^{13}C NMR ($CDCl_3$) δ 172.0, 155.7, 155.1, 153.3, 139.6, 120.6, 114.7, 90.5, 85.5, 84.1, 83.5, 82.1, 79.9, 59.5, 57.4, 56.3, 52.7, 51.3, 51.1, 29.9, 28.5, 28.1, 27.3, 27.2, 26.4, 25.6; additional carbons identified by HMBC, ($CDCl_3$) δ 148.1, 39.9. HRMS-ESI: calcd for $C_{32}H_{52}N_{10}O_8$ ($M + Na^+$) 727.3862, obsd 727.3867.

4.6 General Procedure for Iodination and Global Deprotection

Iodination of the alcohol requires the reagents to be added in the proper order to prevent reduction of azides **11d** and **11e** by triphenylphosphine. Thus, the following order of addition was followed for incorporating the iodine for all analogs described. I_2 (0.0247 g, 0.097 mmol) was added to triphenylphosphine (0.0247 g, 0.094 mmol) and imidazole (0.0064 g, 0.094 mmol) in 250 μL CH_2Cl_2 at 0°C. Once triphenylphosphine was completely consumed

(monitored by TLC), the amino acid functionalized *N*-hydroxyethyl adenosine (0.0619 mmol) in CH_2Cl_2 (150 μL) was then added and stirred for 1 h. The reaction was diluted with ice-chilled CH_2Cl_2 and H_2O and the organic layer was washed with three times with H_2O. After evaporating *in vacuo*, 4N HCl-dioxane (370 μL) was added to the iodinated adenosine in CH_2Cl_2 (1.8 mL) and the mixture was heated to 40°C and stirred for 2 h. Ice-chilled H_2O was added and the aqueous layer was extracted with CH_2Cl_2 several times prior to lyophilization to afford a crude light yellow solid. Compound purification was carried out using HPLC (see Supplementary Information for details). NMR analysis of the resulting analogs required the addition of TFA to the deuterated solvent for product stability. Residual signals in the ^{13}C spectrum resulted and appeared in the range of 162.0-161.0 and 121.0-113.0 ppm. Several carbon signals were broadened and indistinguishable from noise by ^{13}C and required characterization by HMBC. Additional carbon signals from the alkyl linkers were doubled and listed as the major and minor signals where observed.

4.6.1 *N*6-(Prop-2''-ynyl)-5'-(diaminobutyric acid)-5'-*N*-iodoethyl-5'-deoxyadenosine ammonium hydrochloride (**5a**). The reaction was carried out with **11a** and the crude product was purified by HPLC (product eluted at 8.23 min; see Figure S2 in SI) to afford **5a** as a white solid (0.0098 g, 28 %). MP 131 °C (dec). ^1H NMR (1 % TFA in CD_3OD) δ 8.47 (s, 1H), 8.41 (bs, 1H), 6.12 (d, *J* = 4.2 Hz, 1H), 4.77 (t, *J* = 4.3 Hz, 1H), 4.50-4.42 (m, 4H), 4.10 (dd, *J* = 7.6, 5.6 Hz, 1H), 3.81-3.75 (m, 1H), 3.68-3.61 (m, 3H), 3.58-3.41 (m, 4H), 2.78 (t, *J* = 2.5 Hz, 1H), 2.45-2.37 (m, 1H), 2.31-2.25 (m, 1H); ^{13}C NMR (1 % TFA in CD_3OD) δ 170.8, 143.5, 121.5, 92.1, 80.2, 79.2, 74.4, 73.9, 73.5, 57.4, 56.1, 51.8, 51.4, 31.6, 26.1, -8.0; additional carbons identified by HMBC, (1 % TFA in CD_3OD) δ 152.6, 149.8, 148.9. HRMS-ESI: calcd for $C_{19}H_{27}IN_7O_5$ (M^+) 560.1113, obsd 560.1114.

4.6.2 *N*6-(But-3''-ynyl)-5'-(diaminobutyric acid)-5'-*N*-iodoethyl-5'-deoxyadenosine ammonium hydrochloride (**5b**). The reaction was carried out with **11b** and the crude product was purified by HPLC (product eluted at 8.33 min; see Figure S3 in SI) to afford **5b** as a white solid (0.010 g, 28 %). MP 136 °C (dec). ^1H NMR (1 % TFA in CD_3OD, 40ºC) δ 8.46 (s, 2H), 6.14 (d, *J* = 4.0 Hz, 1H), 4.77 (t, *J* = 4.5 Hz, 1H), 4.50-4.45 (m, 2H), 4.12 (dd, *J* = 7.5, 5.5 Hz, 1H), 3.82-3.74 (m, 2H), 3.68-3.63 (m, 3H), 3.58-3.43 (m, 5H), 2.70-2.67 (m, 2H), 2.43-2.40 (m, 1H), 2.35 (t, *J* = 2.6 Hz, 1H), 2.33-2.27 (m, 1H); ^{13}C NMR (1 % TFA in CD_3OD) δ 170.9, 143.5, 121.3, 92.0, 81.3, 80.3, 74.4, 73.5, 71.8, 57.4, 56.1, 51.9, 51.6, 41.7, 26.1, 19.6, -7.8; carbons identified to be doubled, (1 % TFA in CD_3OD) δ 41.7, 19.6 (major) and 44.0, 20.4 (minor); additional carbons identified by HMBC, (1 % TFA in CD_3OD) δ 151.4, 148.4, 147.5. HRMS-ESI: calcd for $C_{20}H_{29}IN_7O_5$ (M^+) 574.1269, obsd 574.1271.

4.6.3 *N*6-(Hex-5''-ynyl)-5'-(diaminobutyric acid)-5'-*N*-iodoethyl-5'-deoxyadenosine ammonium hydrochloride (**5c**). The reaction was carried out with **11c** and the crude product was purified by HPLC (product eluted at 8.93 min; see Figure S4 in SI) to afford **5c** as a white solid (0.013 g, 35 %). MP 142 °C (dec). ^1H NMR (1 % TFA in CD_3OD) δ 8.44 (s, 2H), 6.12 (d, *J* = 4.1 Hz, 1H), 4.74 (t, *J* = 4.5 Hz, 1H), 4.48-4.43 (m, 2H), 4.11 (dd, *J* = 7.4, 5.8 Hz, 1H),

3.76-3.60 (m, 6H), 3.53-3.42 (m, 4H), 2.49-2.38 (m, 1H), 2.32-2.23 (m, 3H), 2.19 (bs, 1H), 1.93-1.87 (m, 2H). 1.70-1.64 (m, 2H); ^{13}C NMR (1 % TFA in CD_3OD) δ 170.9, 143.8, 92.1, 84.3, 80.2, 74.5, 73.5, 70.1, 57.4, 56.1, 51.8, 51.5, 26.7, 26.1, 18.7, -8.0; carbons identified to be doubled, (1 % TFA in CD_3OD) δ 42.8, 28.6 (major) and 45.0, 29.8 (minor); due to peak doubling in the range of 151.0-146.0 and 121.5-120.0, the following carbons were identified by HMBC, (1 % TFA in CD_3OD) δ 151.5, 147.5, 146.5, 121.3. HRMS-ESI: calcd for $C_{22}H_{33}IN_7O_5$ (M^+) 602.1582, obsd 602.1584.

4.6.4 *N*6-(3"-azidopropyl)-5'-(diaminobutyric acid)-5'-*N*-iodoethyl-5'-deoxyadenosine ammonium hydrochloride (**6a**). The reaction was carried out with **11d** and the crude product was purified by HPLC (product eluted at 8.62 min see Figure S5 in SI) to afford **6a** as a white solid (0.0094 g, 25 %). MP 125 °C (dec). ^{1}H NMR (1 % TFA in CD_3OD, 40ºC) δ 8.42 (bs, 2H), 6.11 (bd, *J* = 4.1 Hz, 1H), 4.75 (t, *J* = 4.5, 1H), 4.47-4.42 (m, 2H), 4.09 (dd, *J* = 7.4, 5.5 Hz, 1H), 3.83-3.64 (m, 3H), 3.62-3.55 (m, 3H), 3.53-3.40 (m, 6H), 2.44-2.35 (m, 1H), 2.29-2.22 (m, 1H), 2.02 (p, *J* = 6.8 Hz, 2H); ^{13}C NMR (1 % TFA in CD_3OD) 171.0, 143.7, 121.4, 92.0, 80.3, 74.5, 73.5, 57.4, 56.1, 52.0, 51.7, 40.4, 26.2, -7.7; carbons identified to be doubled, (1 % TFA in CD_3OD) δ 42.6, 29.0 (major) and 43.0, 30.0 (minor); additional carbons identified by HMBC, (1 % TFA in CD_3OD) δ 151.5, 147.9, 147.3. HRMS-ESI: calcd for $C_{19}H_{30}IN_{10}O_5$ (M^+) 605.1440, obsd 605.1441.

4.6.5 *N*6-(4"-azidobutyl)-5'-(diaminobutyric acid)-5'-*N*-iodoethyl-5'-deoxyadenosine ammonium hydrochloride (**6b**). The reaction was carried out with **11e** and the crude product was purified by HPLC (product eluted at 8.95 min; see Figure S6 in SI) to afford **6b** as a white solid (0.0096 g, 25 %). MP 120 °C (dec). ^{1}H NMR (1 % TFA in CD_3OD) δ 8.44 (s, 1H), 8.41 (s, 1H), 6.11 (bs, 1H), 4.76-4.74 (m, 1H), 4.47-4.42 (m, 2H), 4.09 (dd, *J* = 7.5, 5.5 Hz, 1H), 3.78-3.58 (m, 6H), 3.52-3.38 (m, 6H), 2.43-2.36 (m, 1H), 2.28-2.21 (m, 1H), 1.87-1.81 (m, 2H), 1.76-1.70 (m, 2H) ; ^{13}C NMR (1 % TFA in CD_3OD) δ 171.0, 143.9, 92.1, 80.2, 74.5, 73.5, 57.4, 56.0, 52.0, 51.6, 44.9, 42.7, 26.9, 26.1, -7.9; carbon identified to be doubled, (1 % TFA in CD_3OD) δ 27.1 (major) and 28.0 (minor); due to peak doubling in the range of 151.5-146.0 and 121.5-120.0, the following carbons were identified by HMBC, (1 % TFA in CD_3OD) δ 151.2, 148.2, 147.1, 121.2. HRMS-ESI: calcd for $C_{20}H_{32}IN_{10}O_5$ (M^+) 619.1596, obsd 619.1597.

4.7 Linearization of pUC19

Plasmid DNA was linearized prior to conducting MTase-dependent labeling. The following procedure was carried out as described by the manufacturer: pUC19 was linearized with R.EcoRI (2 U per μg DNA) in R.EcoRI buffer (final DNA concentration of 0.1 μg/μL) at 37°C for 1 h and followed heat inactivation at 65°C for 20 min. The linearized pUC19 was used without subsequent purification.

4.8 Restriction/protection assay with M.TaqI

DNA reaction mixtures for each SAM analog were prepared as previously described.[10, 15] Briefly, DNA reaction mixtures were prepared by the addition of appropriate analog stock solutions in NEBuffer 4 (final reaction volume was 20μL). The final DNA concentration was 0.025 μg/μL and M.TaqI was 54 U per μg DNA; the final concentrations of the SAM analogs ranged from 100 nM to 10 μM for the specific reaction sequence. All reactions were incubated at 65°C for 4 h. The extent of DNA modification by the SAM analogs was analyzed by the addition of R.TaqI (4 U per μg DNA in 10 μL NEBuffer 4), followed by incubation at 65°C for 1 h. Proteinase K (0.02 U in 5 μL H_2O) was added to each reaction, which were then incubated for 1 h at 37°C. Agarose gel loading dye was added to each reaction and the extent of DNA modification was visualized by electrophoresis on a 2 % agarose gel containing ethidium bromide

Acknowledgements

Financial support of this research from Wake Forest University and the National Science Foundation (CHE – 1307425) are gratefully acknowledged.

Supplementary data

Supplementary data of NMR spectra (^{1}H and ^{13}C NMR) for all compounds described, HPLC methods and chromatograms of **5a-c** and **6a-b**, and M.HhaI reaction procedure and agarose gel image are available. Supplementary data related to this article can be found at ...

References and notes

1. Cheng, X.; Roberts, R. J., *Nucleic Acids Res.* **2001**, *29*, 3784-3795.
2. Cheng, X.; Blumenthal, R. M., *Biochemistry* **2010**, *49*, 2999-3008.
3. Ballestar, E., *Adv Exp Med Biol* **2011**, *711*, 1-11.
4. Kohn, M.; Breinbauer, R., *Angew. Chem. Int. Ed.* **2004**, *43*, 3106-3116.
5. Best, M. D., *Biochemistry* **2009**, *48*, 6571-6584.
6. Pignot, M.; Siethoff, C.; Linscheid, M.; Weinhold, E., *Angew. Chem. Int. Ed.* **1998**, *37*, 2888-2891.
7. Pljevaljcic, G.; Pignot, M.; Weinhold, E., *J. Am. Chem. Soc.* **2003**, *125*, 3486-3492.
8. Pljevaljcic, G.; Schmidt, F.; Weinhold, E., *ChemBioChem* **2004**, *5*, 265-269.
9. Comstock, L. R.; Rajski, S. R., *J. Org. Chem.* **2004**, *69*, 1425-1428.
10. Comstock, L. R.; Rajski, S. R., *Nucleic Acids Res.* **2005**, *33*, 1644-1652.
11. Weller, R. L.; Rajski, S. R., *ChemBioChem* **2006**, *7*, 243-245.
12. Zhang, C.; Weller, R. L.; Thorson, J. S.; Rajski, S. R., *J Am Chem Soc* **2006**, *128*, 2760-2761.
13. Osborne, T.; Roska, R. L.; Rajski, S. R.; Thompson, P. R., *J Am Chem Soc* **2008**, *130*, 4574-4575.
14. Mai, V.; Comstock, L. R., *J Org Chem* **2011**, *76*, 10319-10324.
15. Du, Y.; Hendrick, C. E.; Frye, K. S.; Comstock, L. R., *Chembiochem* **2012**, *13*, 2225-2232.
16. Xiao, B.; Jing, C.; Wilson, J. R.; Walker, P. A.; Vasisht, N.; Kelly, G.; Howell, S.; Taylor, I. A.; Blackburn, G. M.; Gamblin, S. J., *Nature* **2003**, *421*, 652-656.
17. Zhang, X.; Yang, Z.; Khan, S. I.; Horton, J. R.; Tamaru, H.; Selker, E. U.; Cheng, X., *Mol Cell* **2003**, *12*, 177-185.
18. Ashton, T. D.; Scammells, P. J., *Biorg Med Chem Lett* **2006**, *16*, 4564-4566.
19. Ashton, T. D.; Baker, S. P.; Hutchinson, S. A.; Scammells, P. J., *Bioorg Med Chem* **2008**, *16*, 1861-73.
20. Wan, Z. K.; Binnun, E.; Wilson, D. P.; Lee, J., *Org Lett* **2005**, *7*, 5877-80.
21. Hong, W.; Dowden, J., *Chinese Chemical Letters* **2011**, *22*, 1439-1442.
22. Hampton, A., *J Am Chem Soc* **1961**, *83*, 3640–3645.

23. Townsend, A. P.; Roth, S.; Williams, H. E.; Stylianou, E.; Thomas, N. R., *Org Lett* **2009**, *11*, 2976-9.
24. Fukuyama, T.; Lin, S.-C.; Li, L., *J. Am. Chem. Soc.* **1990**, *112*, 7050-7051.
25. Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E., *Nature Structural Biology* **2001**, *8*, 121-125.
26. Sankpal, U. T.; Rao, D. N., *Critical Rev. in Biochem. and Mol. Bio.* **2002**, *37*, 167-197.
27. Schmidt, F. H.; Huben, M.; Gider, B.; Renault, F.; Teulade-Fichou, M. P.; Weinhold, E., *Bioorg. Med. Chem.* **2008**, *16*, 40-48.

Synthesis and evaluation of *N*6-substituted azide- and alkyne-bearing *N*-mustard analogs of *S*-adenosyl-L-methionine

*Mohamed Ramadan, Natalie K. Bremner-Hay, Steig A. Carlson and Lindsay R. Comstock**

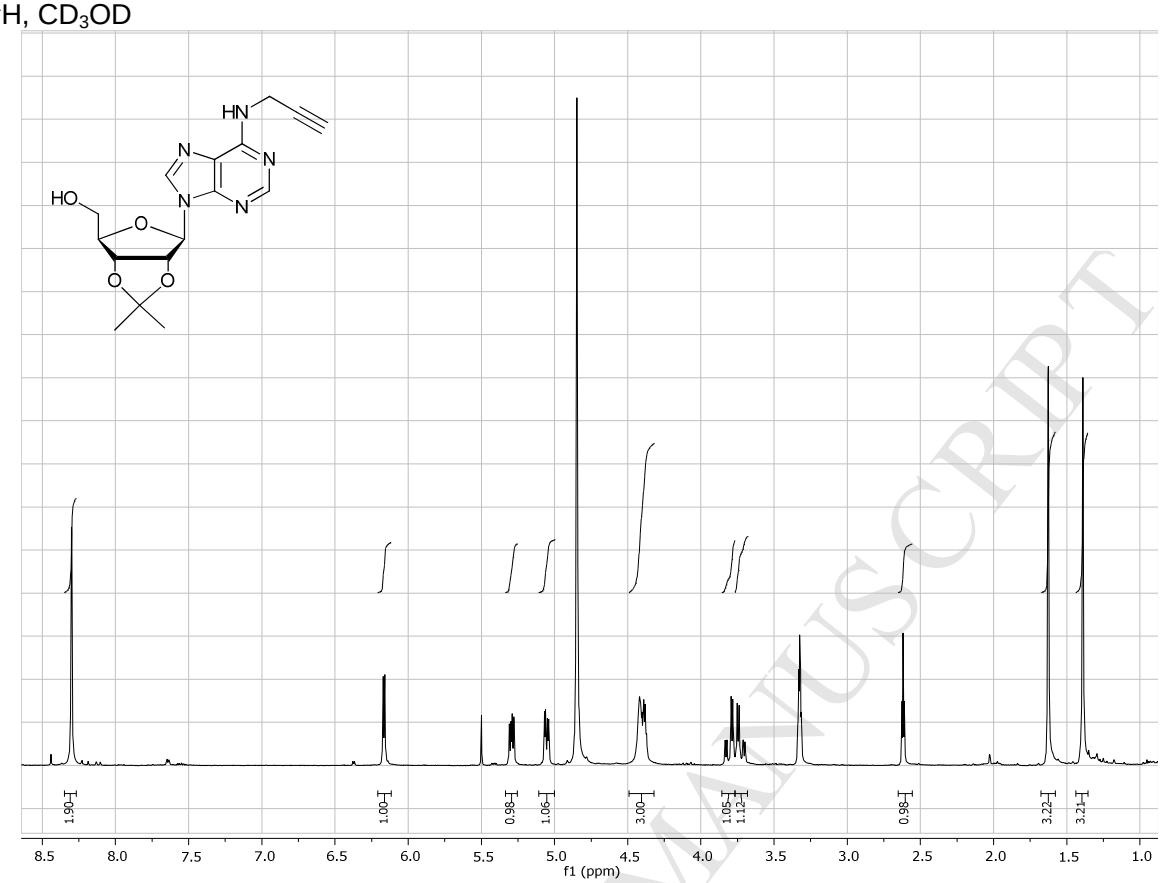
Department of Chemistry, Wake Forest University, Winston-Salem, North Carolina, 27106, *USA*

E-mail: comstolr@wfu.edu

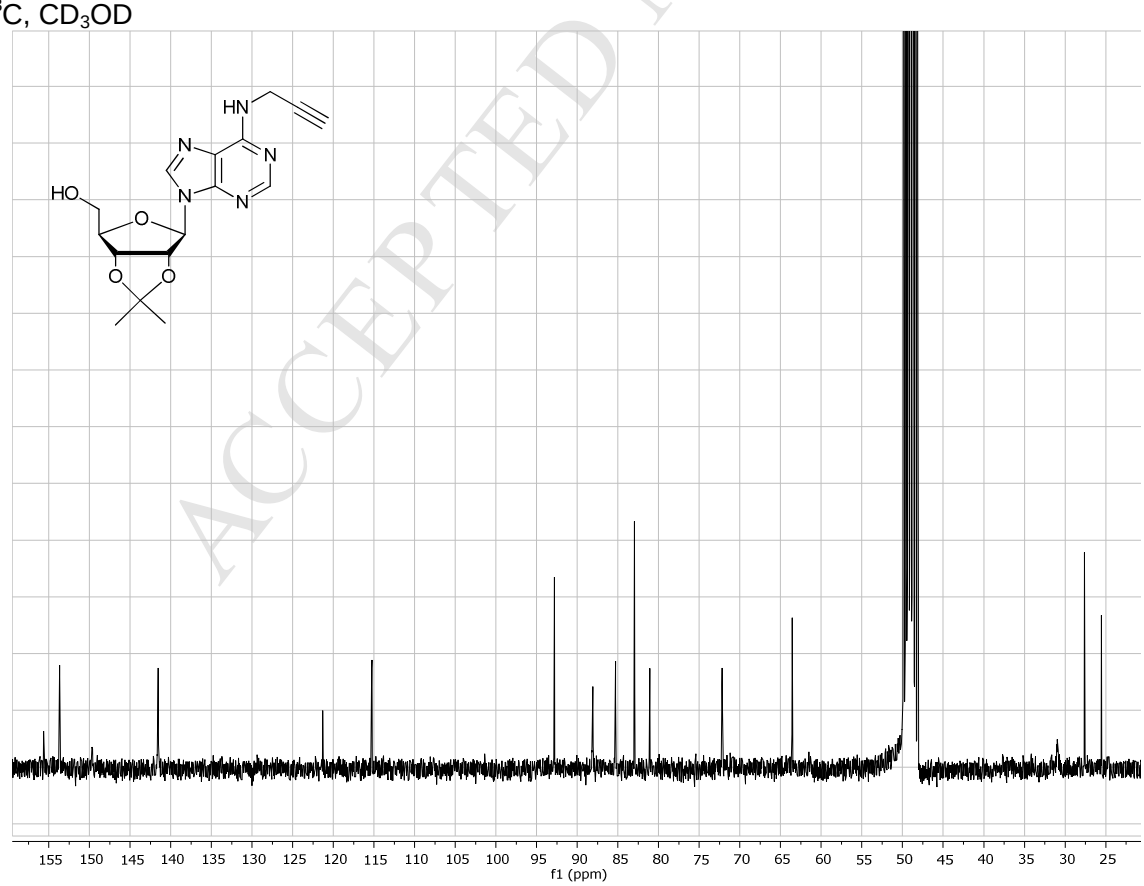
SUPPLEMENTARY DATA

CONTENTS

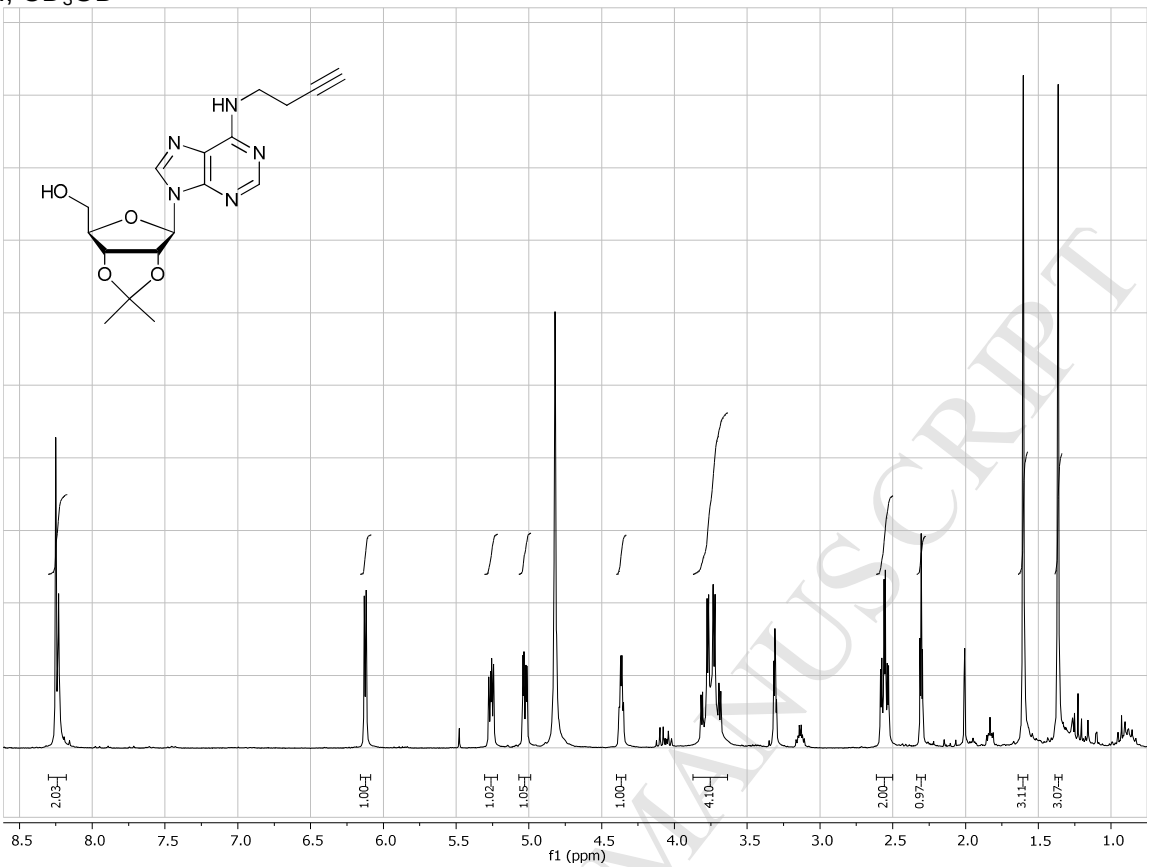
8a – ^{1}H, CD_3OD



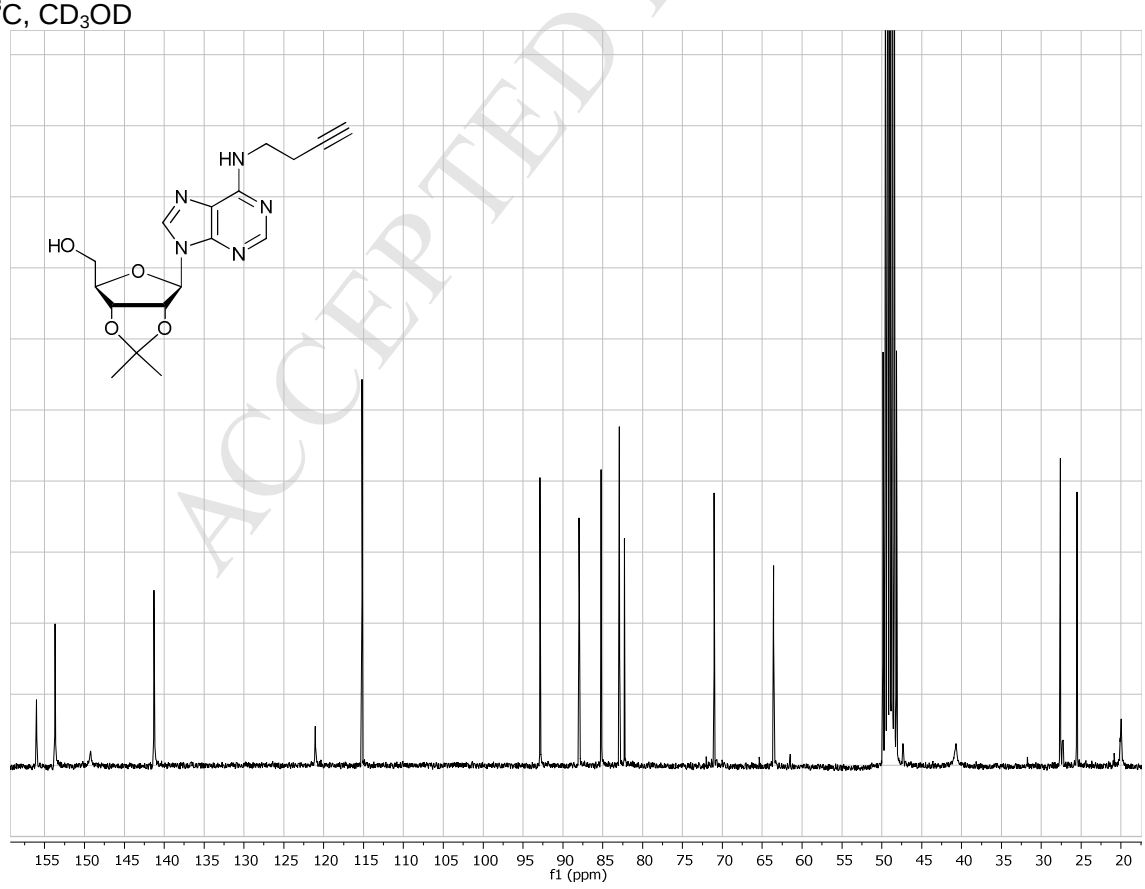
8a – ^{13}C, CD_3OD



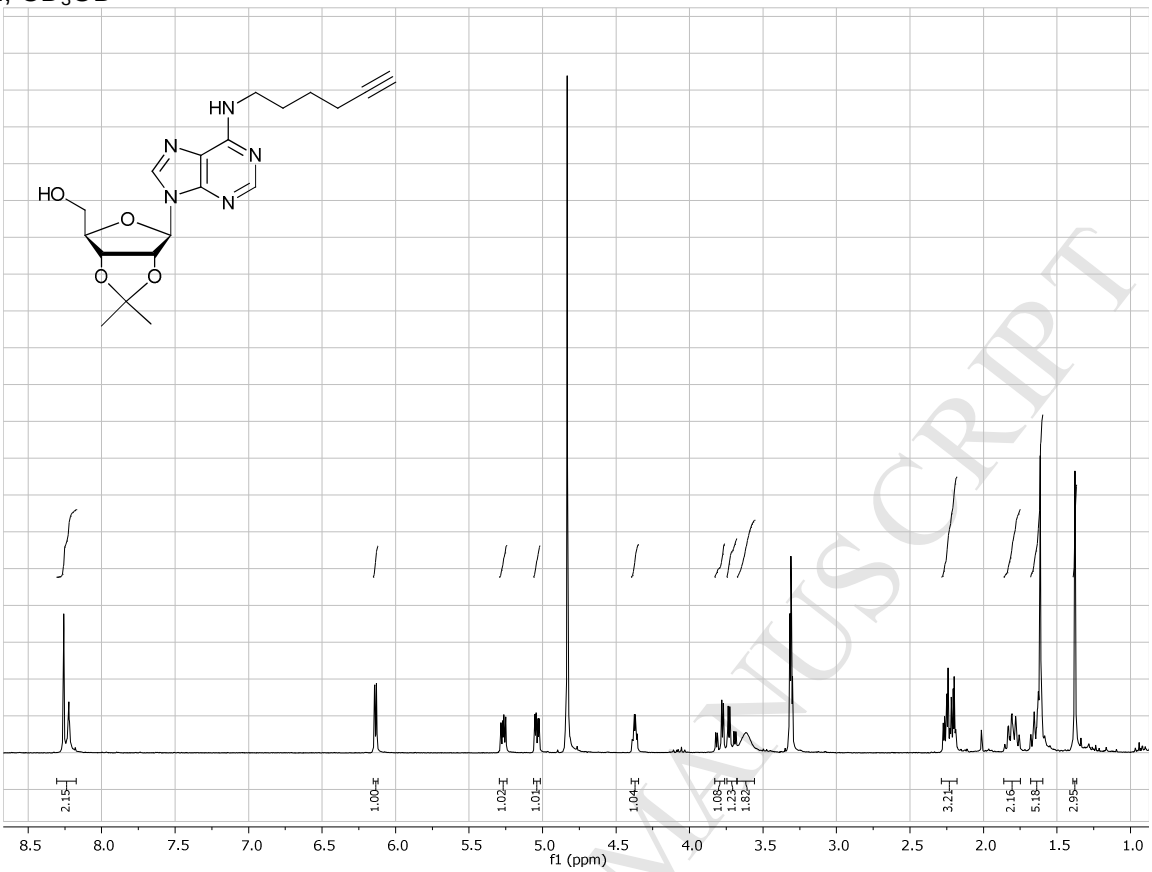
8b – ^{1}H, CD_3OD



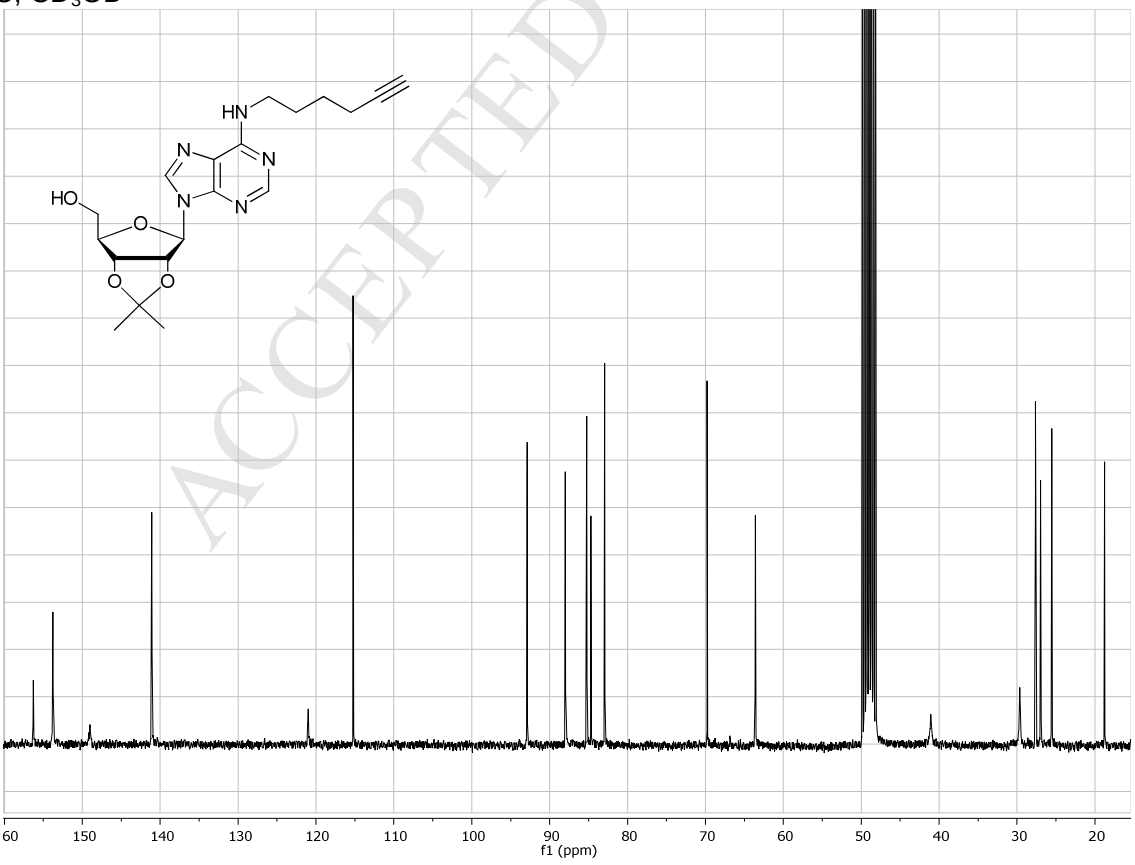
8b – ^{13}C, CD_3OD



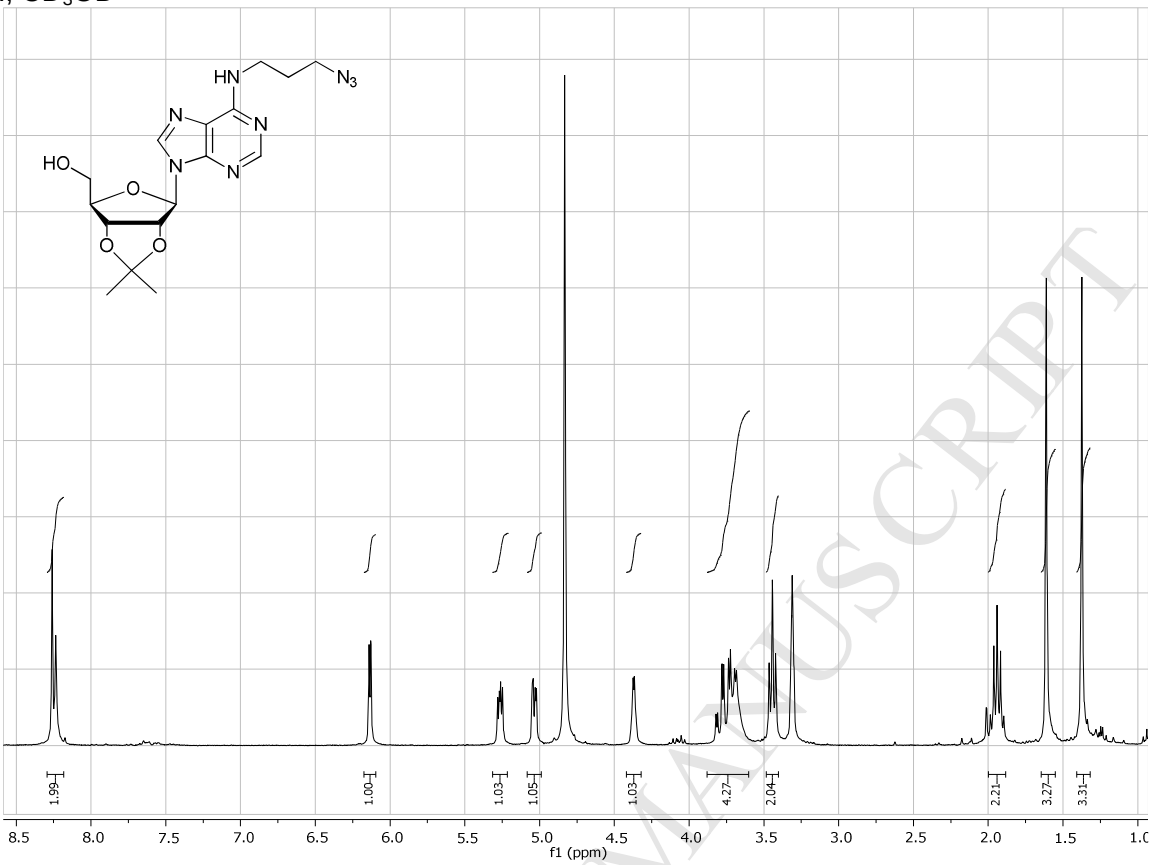
8c – ^{1}H, CD_3OD



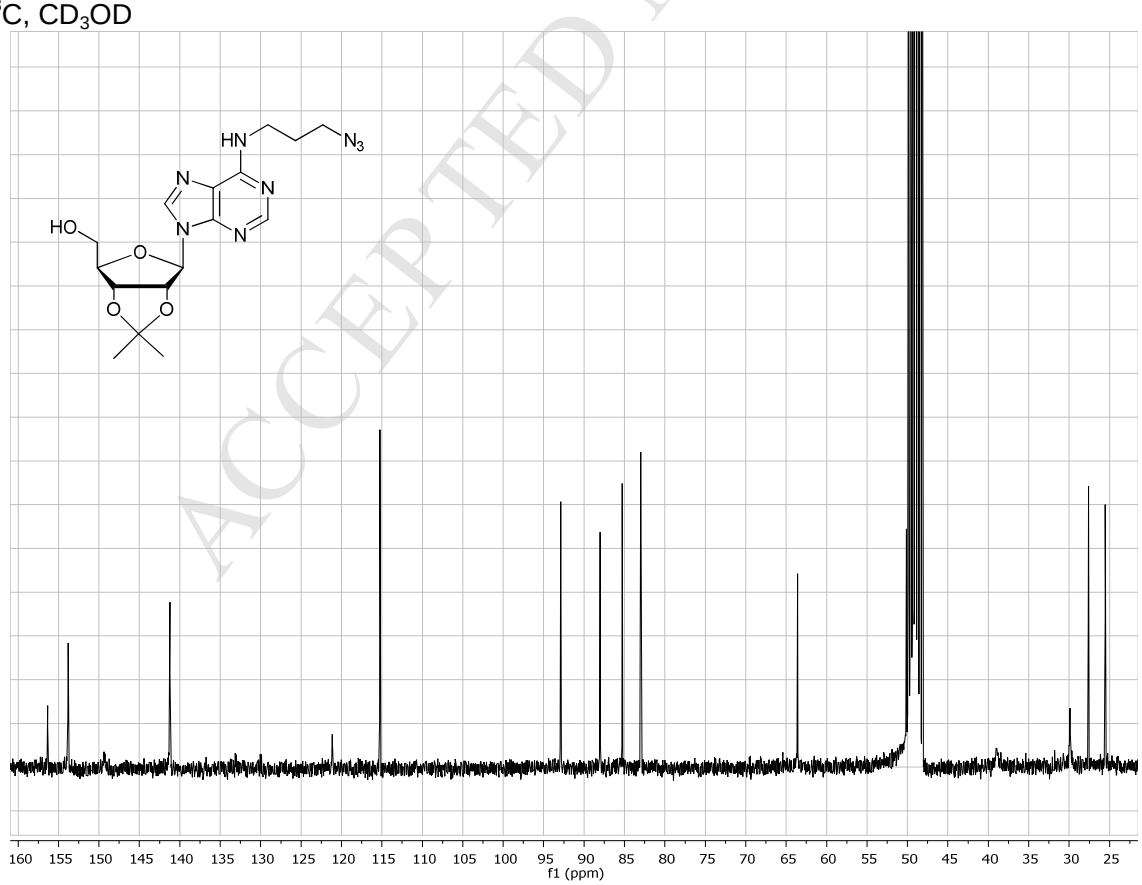
8c – ^{13}C, CD_3OD



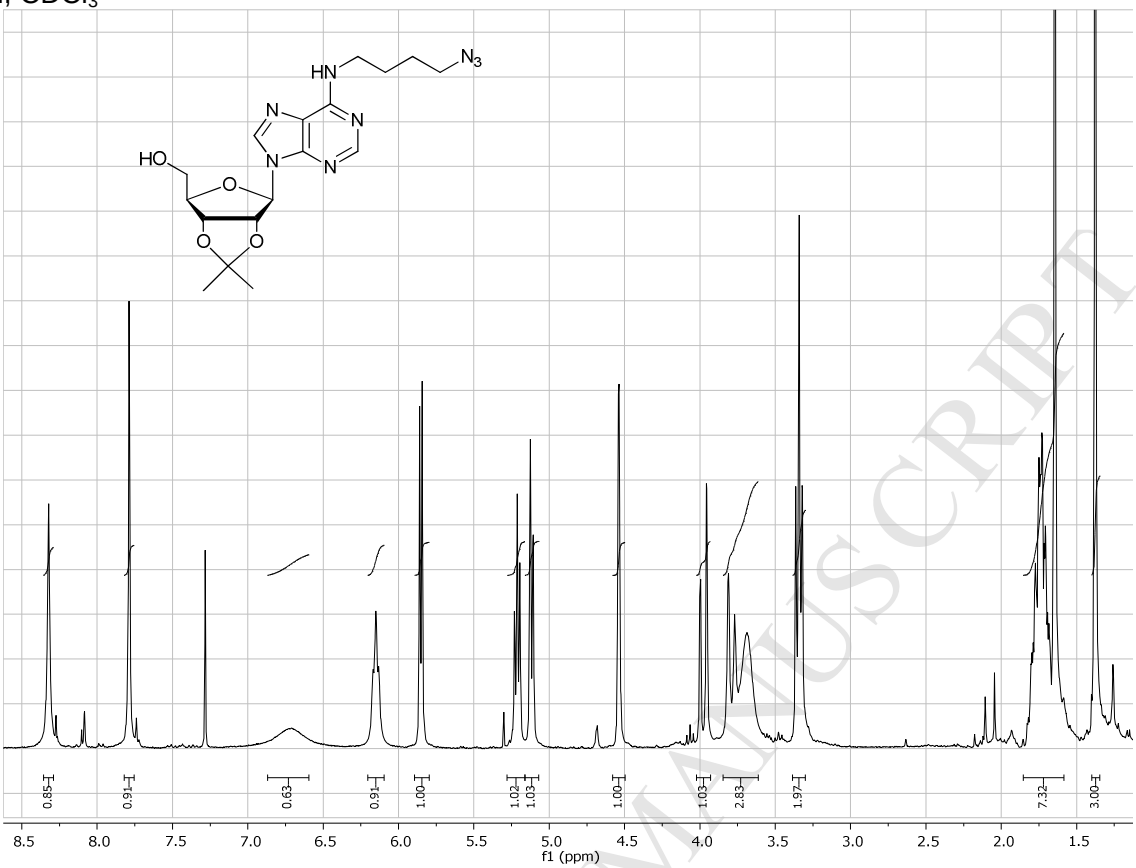
8d – ^{1}H, CD_3OD



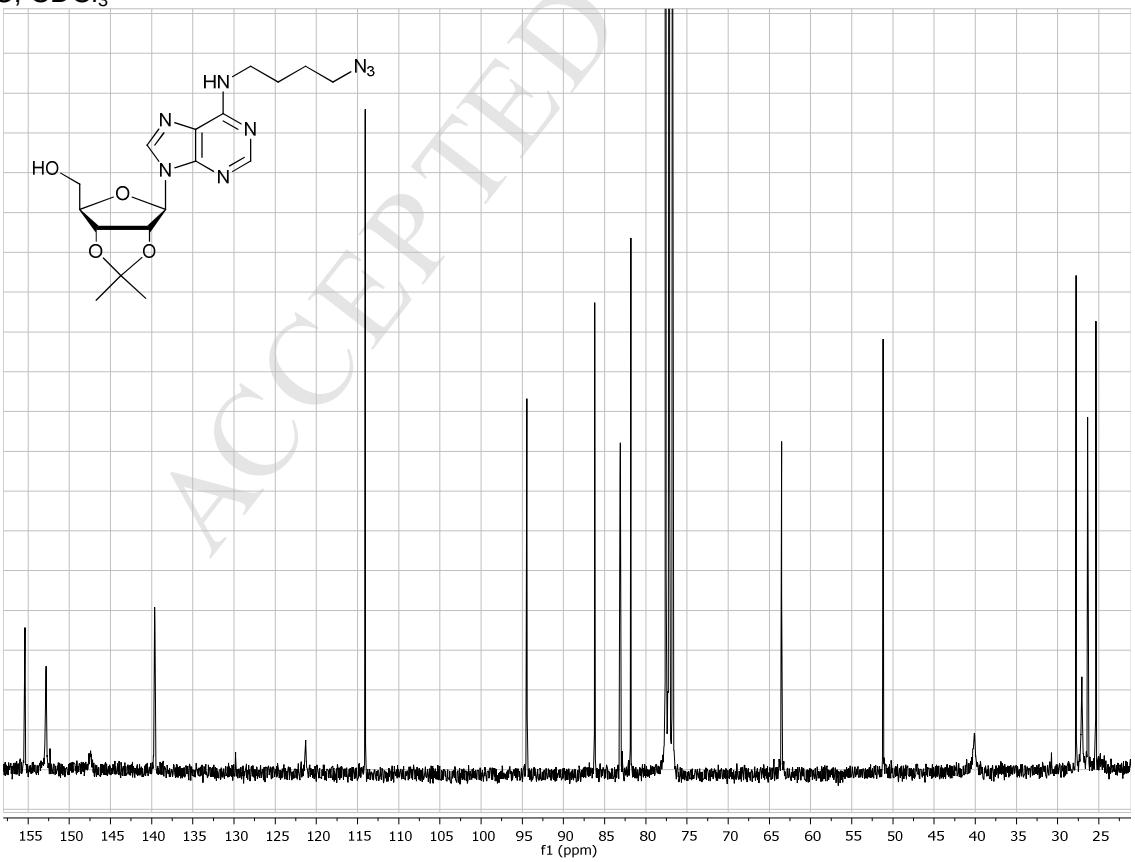
8d – ^{13}C, CD_3OD



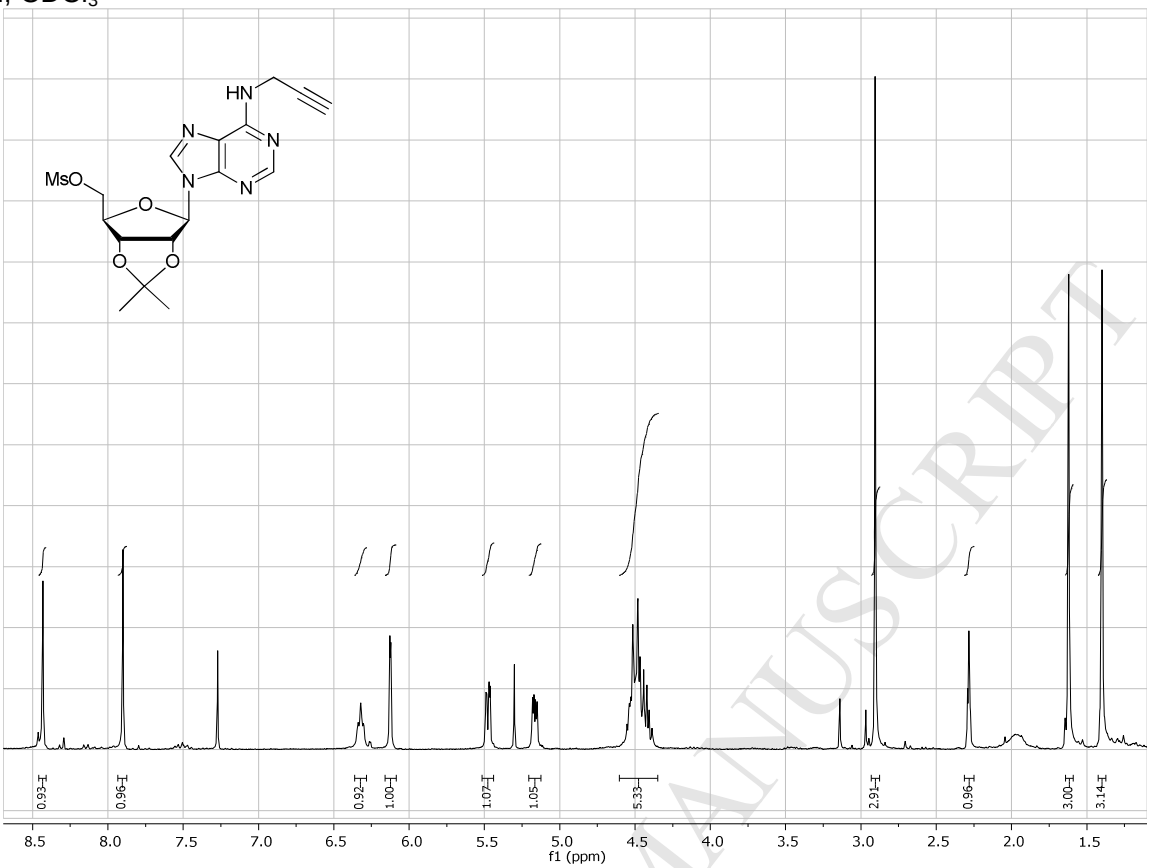
8e – ^{1}H, $CDCl_3$



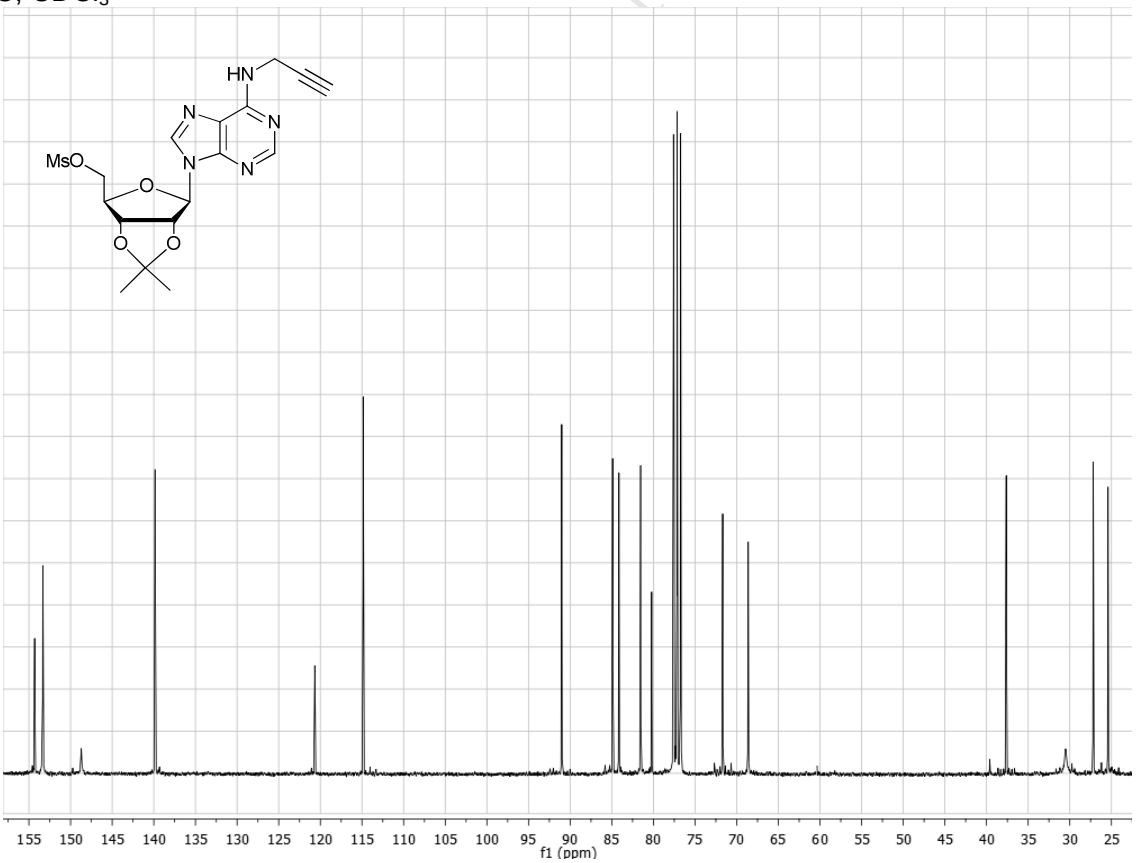
8e – ^{13}C, $CDCl_3$



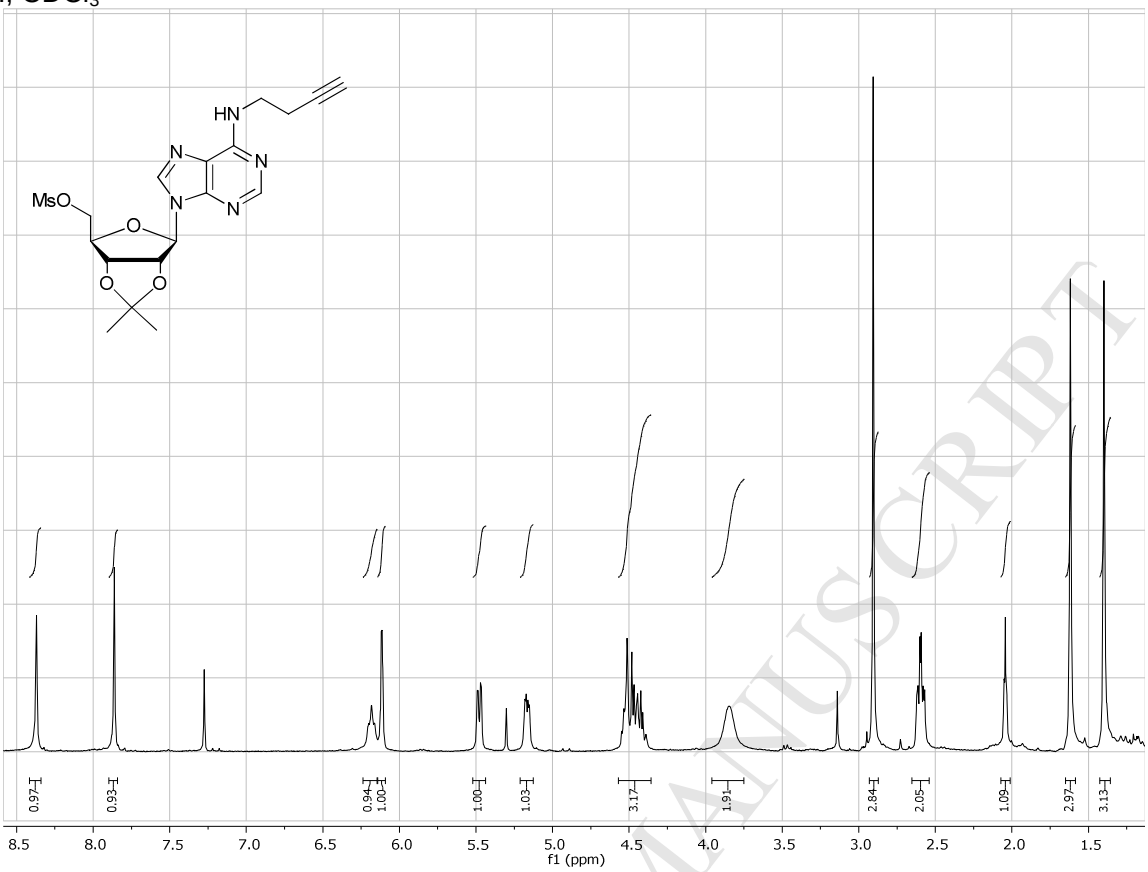
9a – ^{1}H, $CDCl_3$



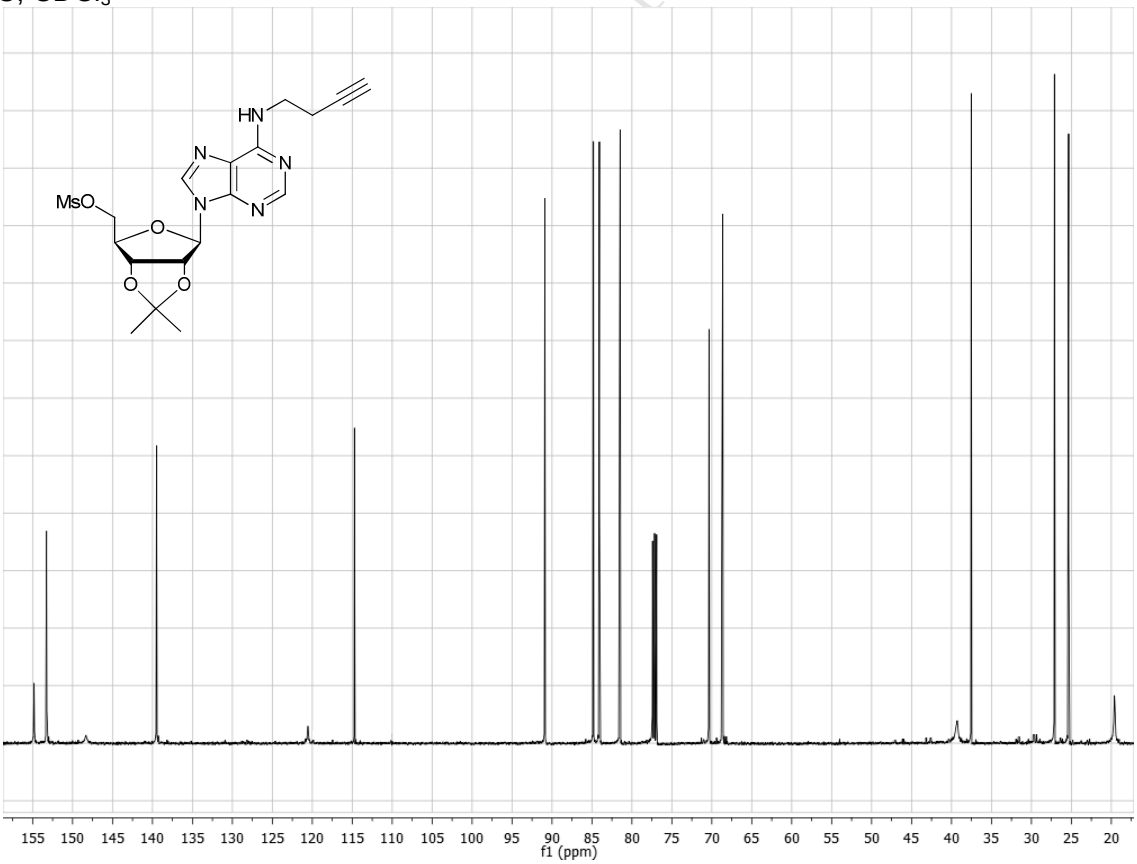
9a – ^{13}C, $CDCl_3$



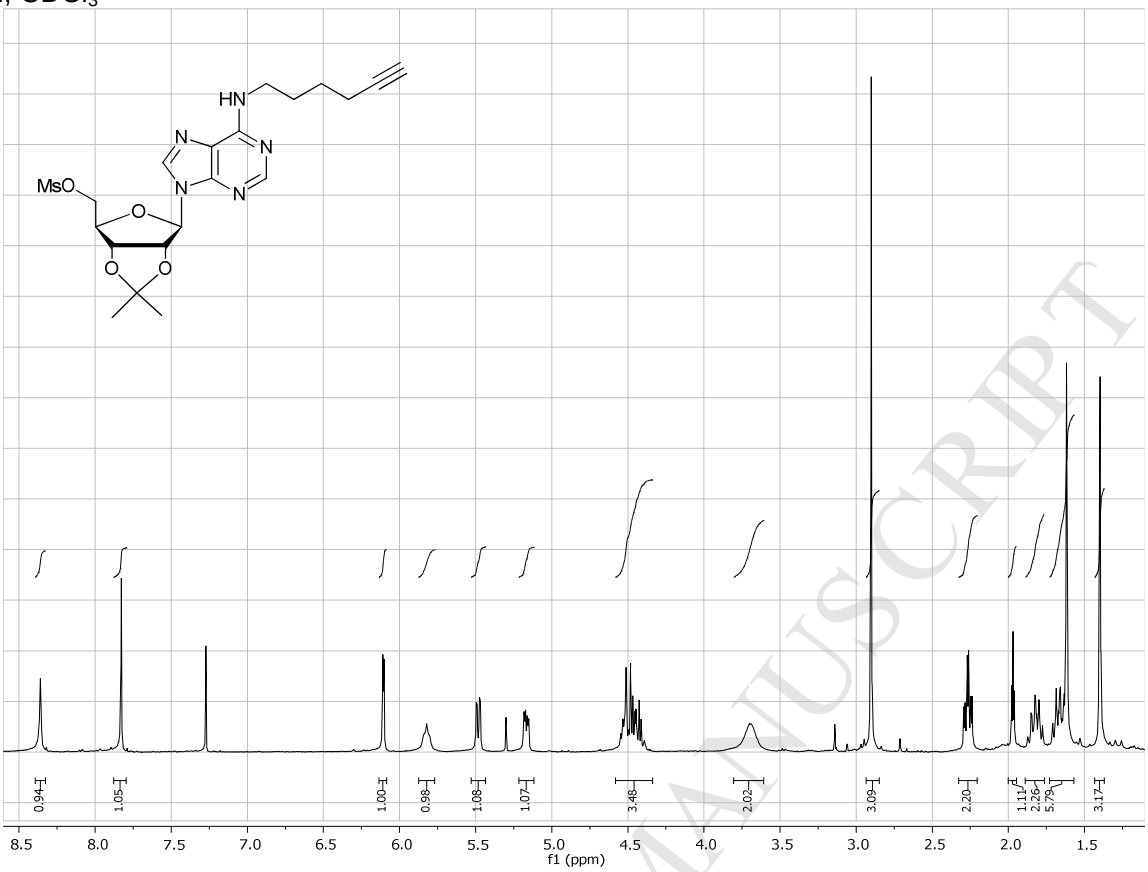
9b – ^{1}H, $CDCl_3$



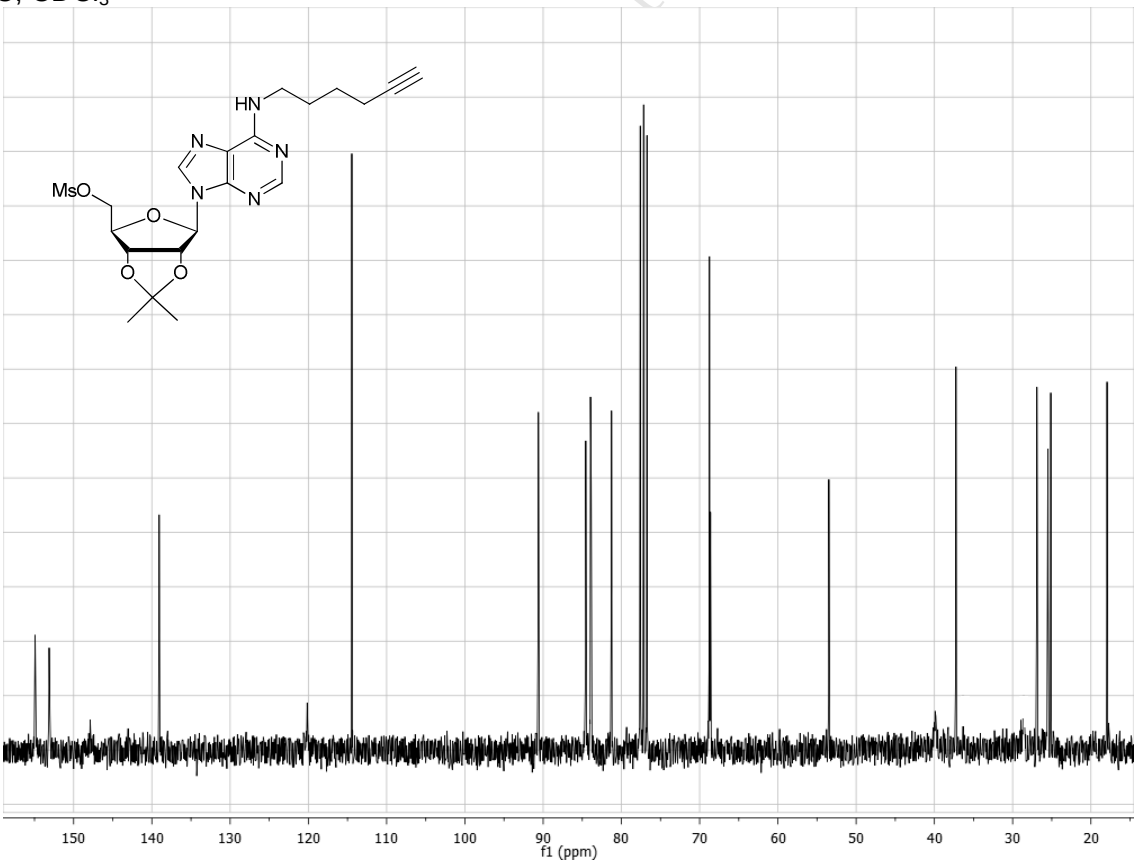
9b – ^{13}C, $CDCl_3$



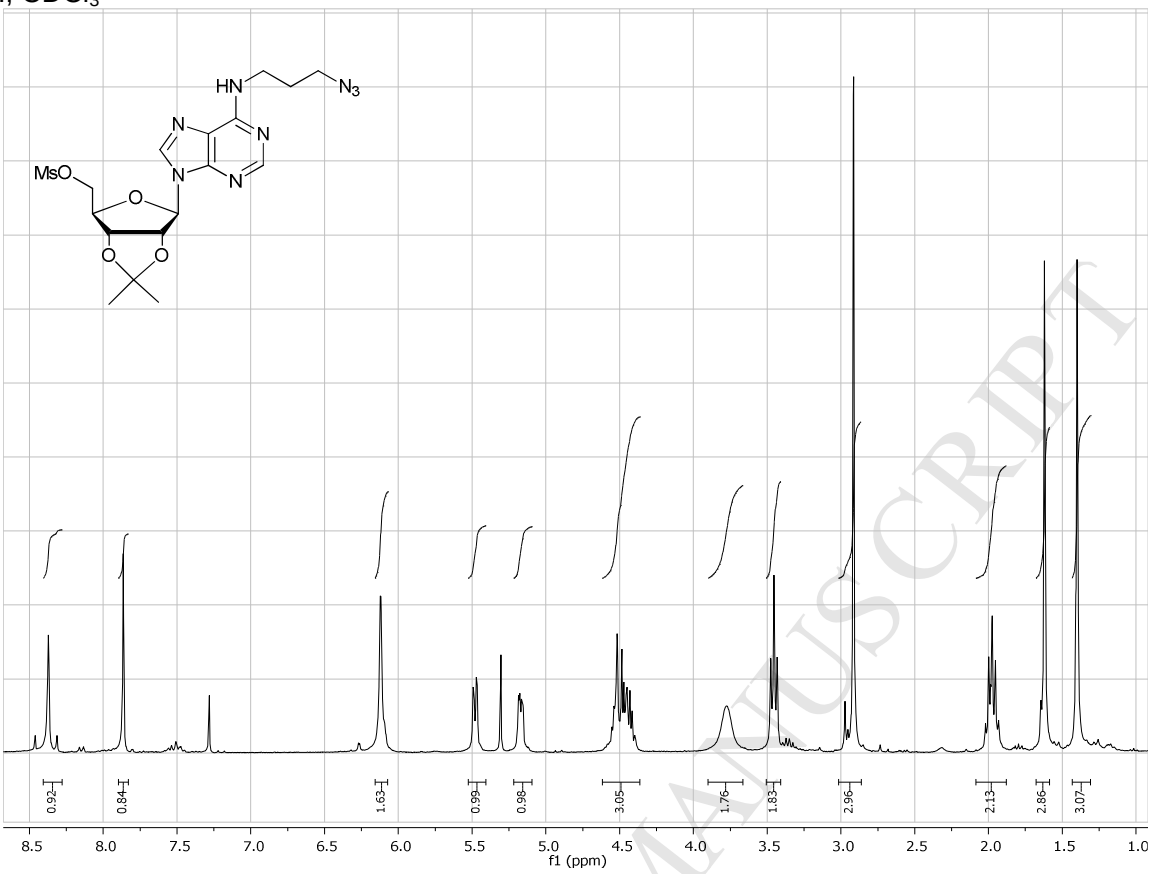
9c – ^{1}H, $CDCl_3$



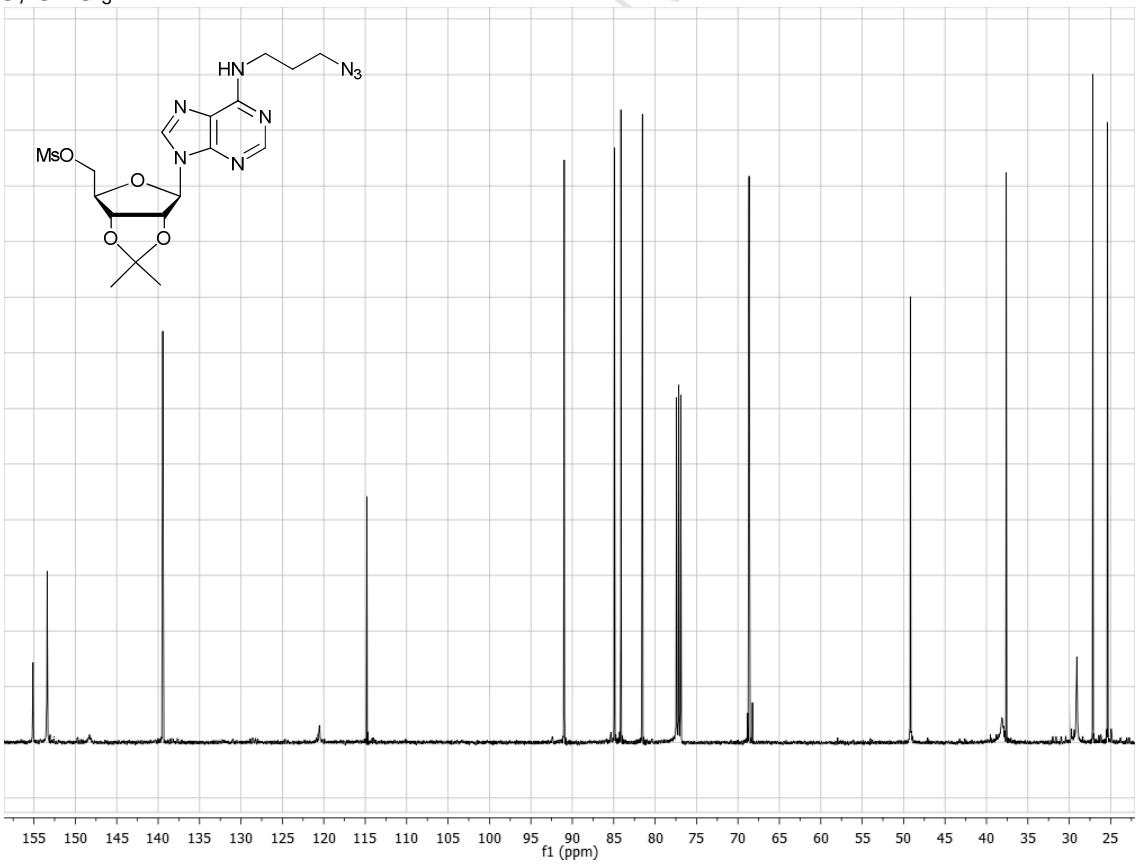
9c – ^{13}C, $CDCl_3$



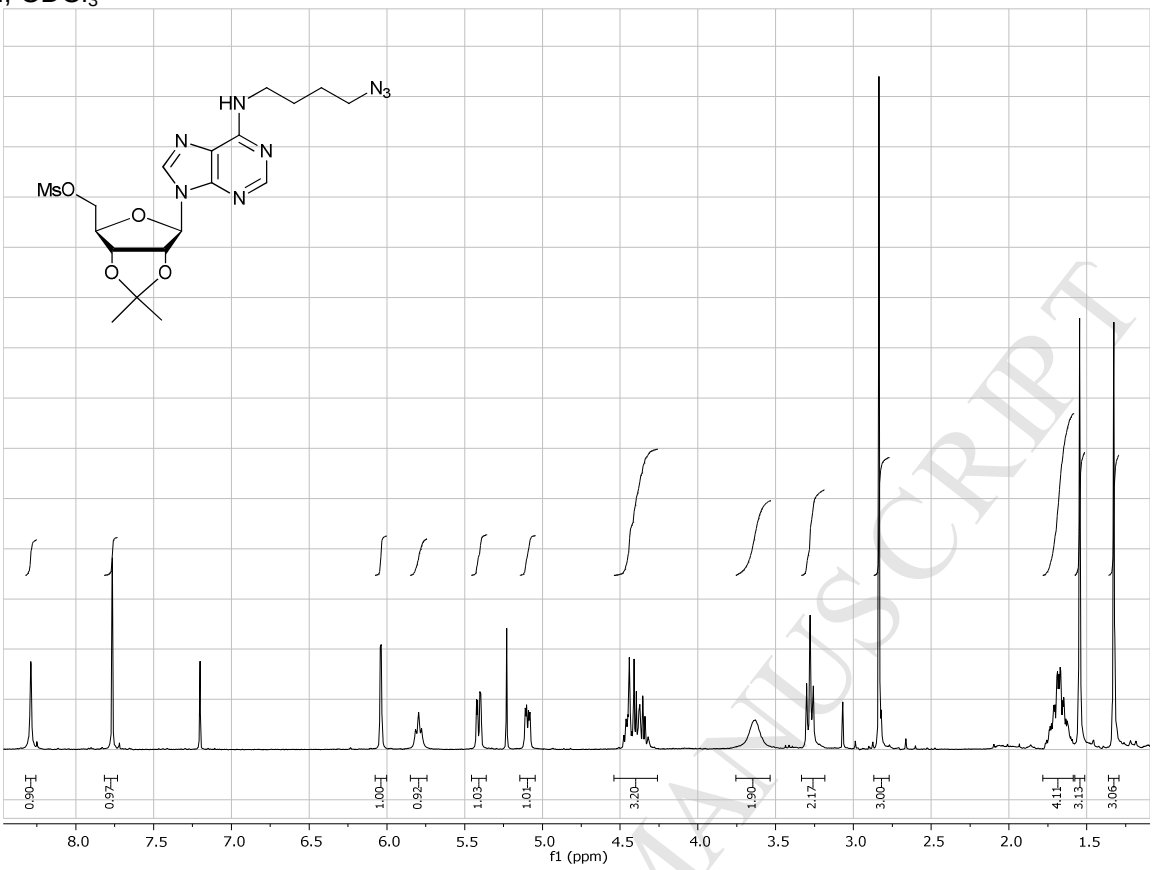
9d – ^{1}H, $CDCl_3$



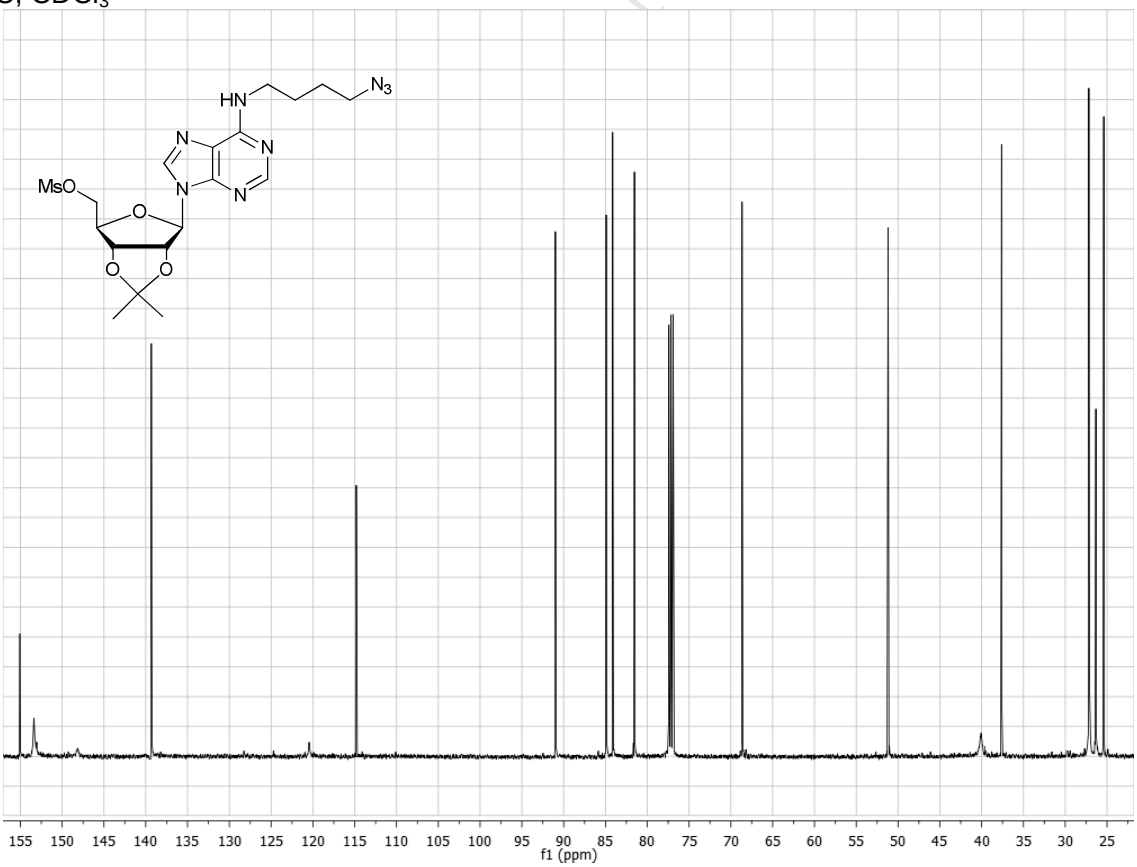
9d – ^{13}C, $CDCl_3$



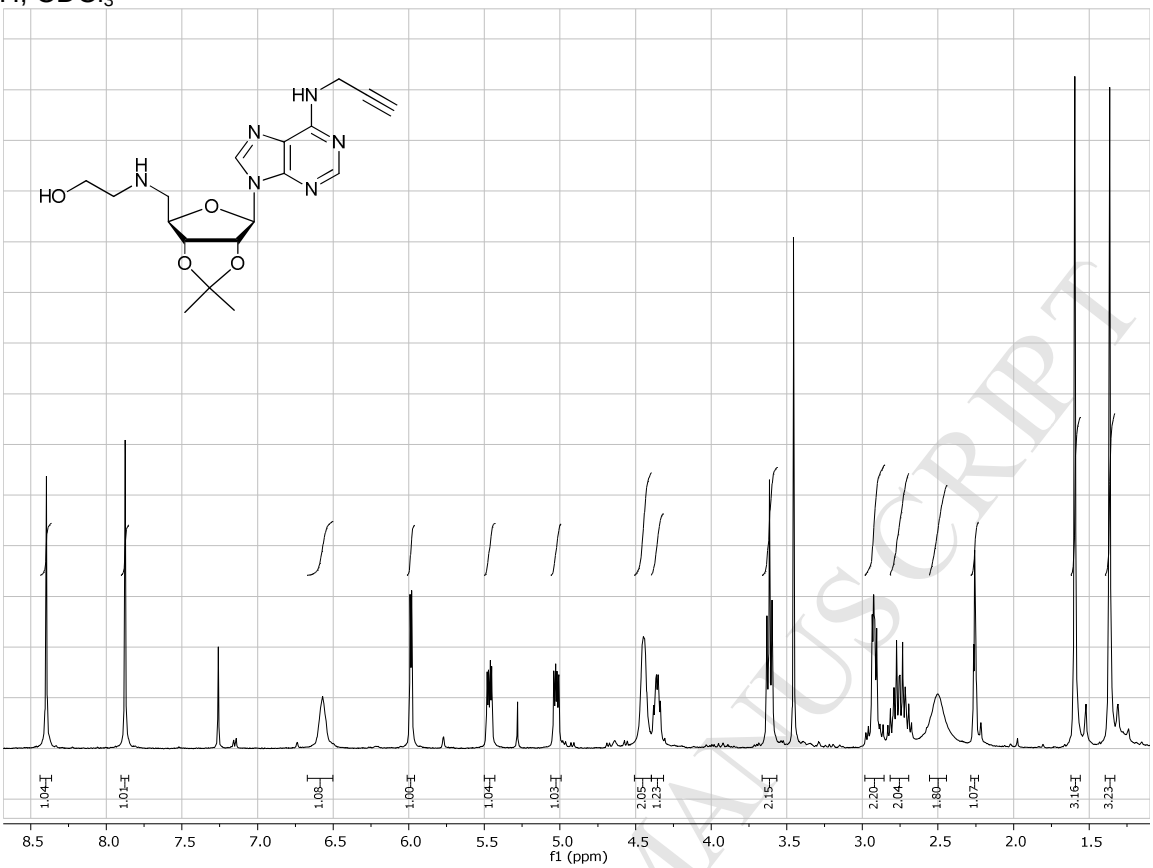
9e – ^{1}H, $CDCl_3$



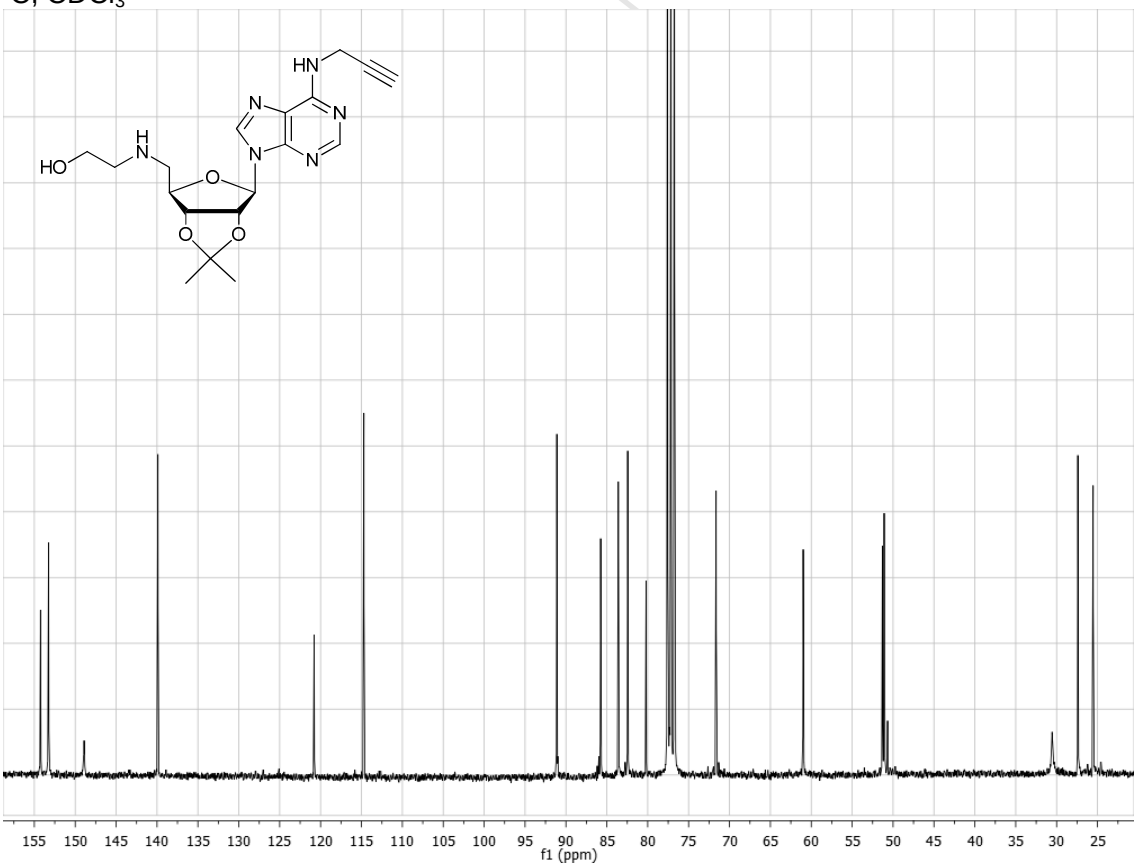
9e – ^{13}C, $CDCl_3$



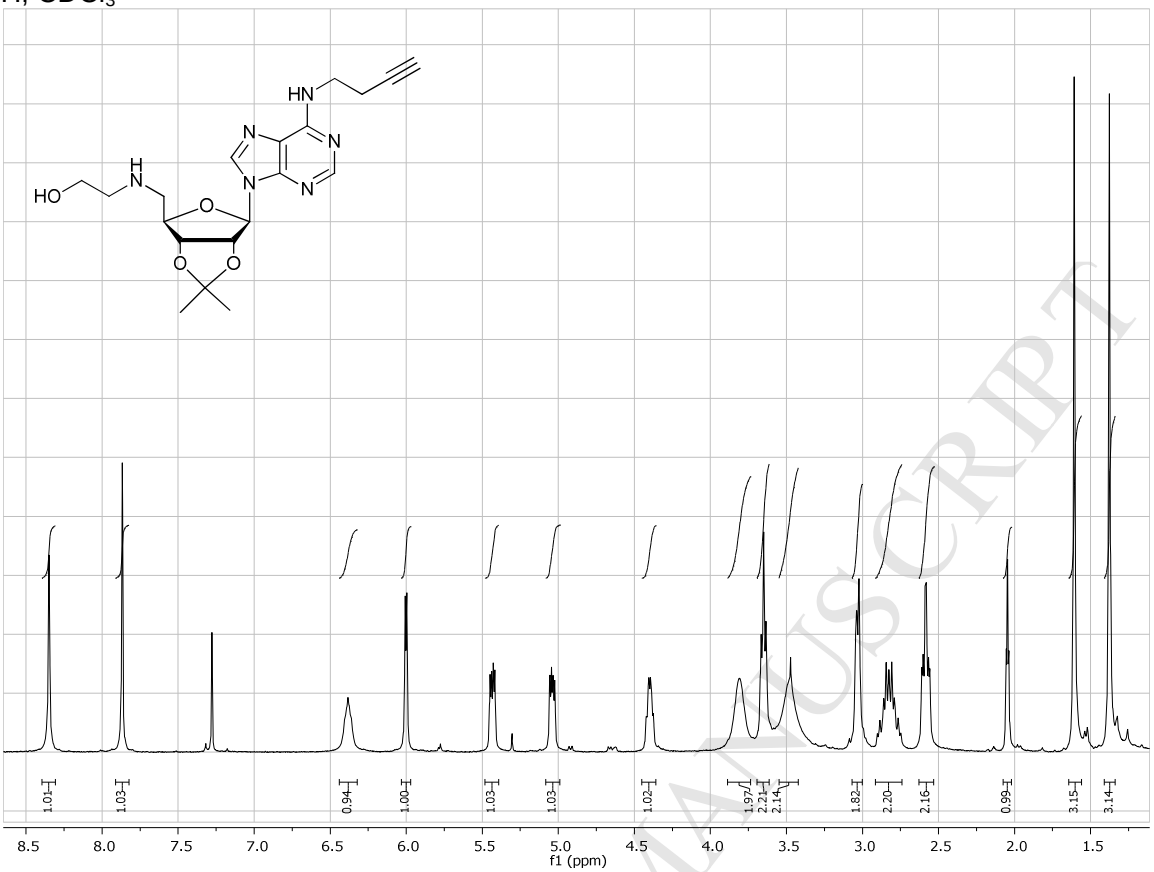
10a – ^{1}H, $CDCl_3$



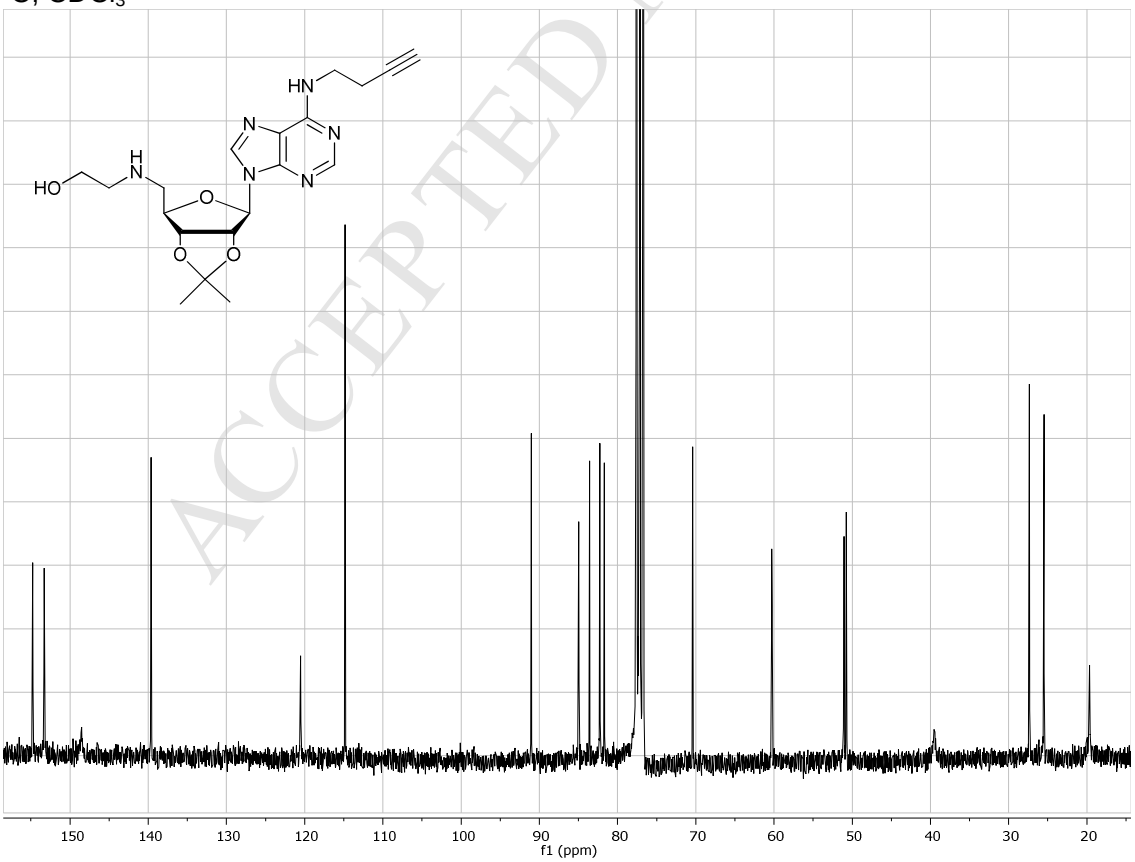
10a – ^{13}C, $CDCl_3$



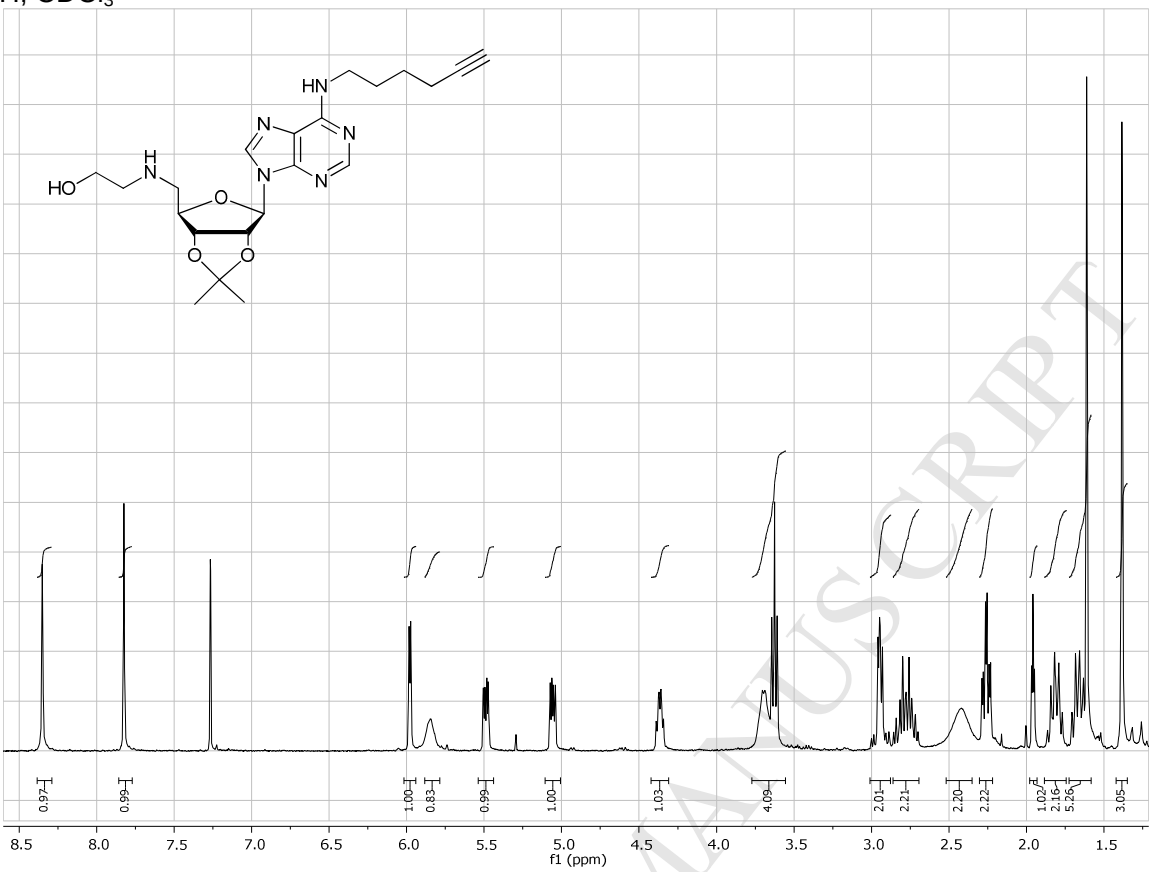
10b – ^{1}H, $CDCl_3$



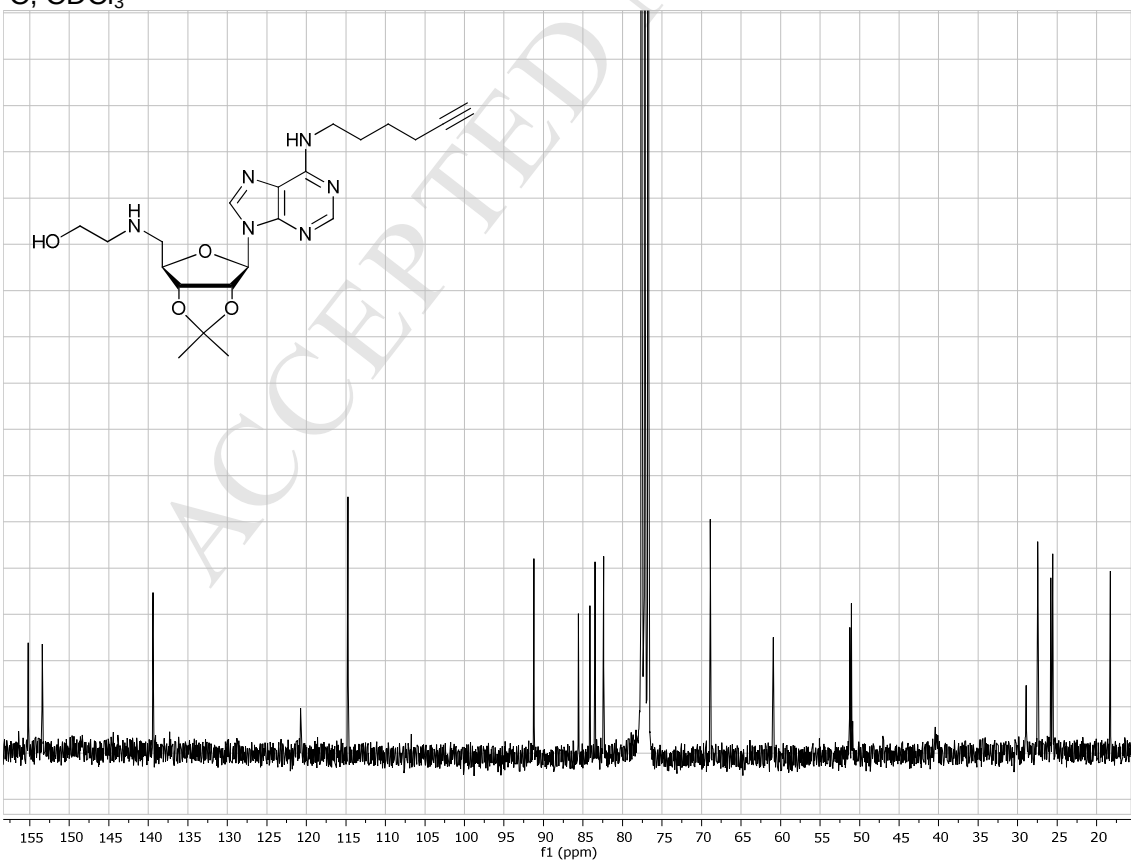
10b – ^{13}C, $CDCl_3$



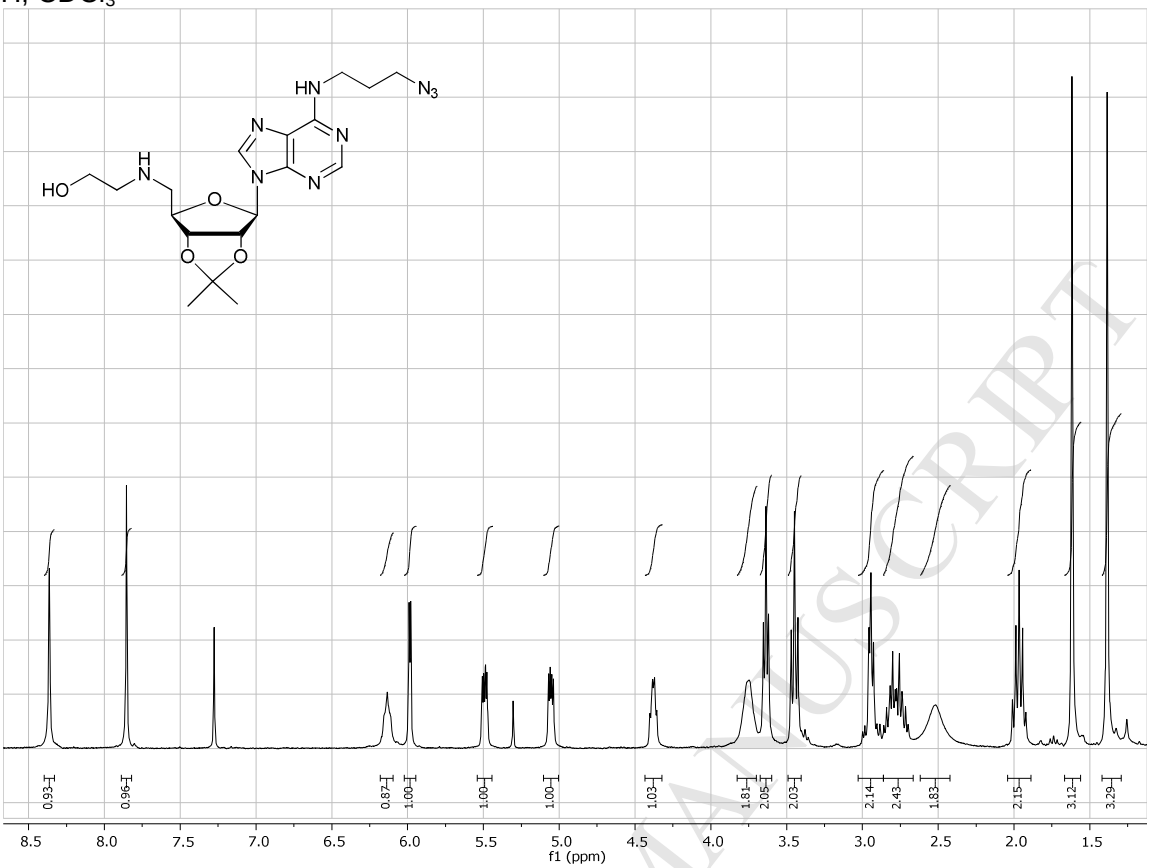
10c – ^{1}H, $CDCl_3$



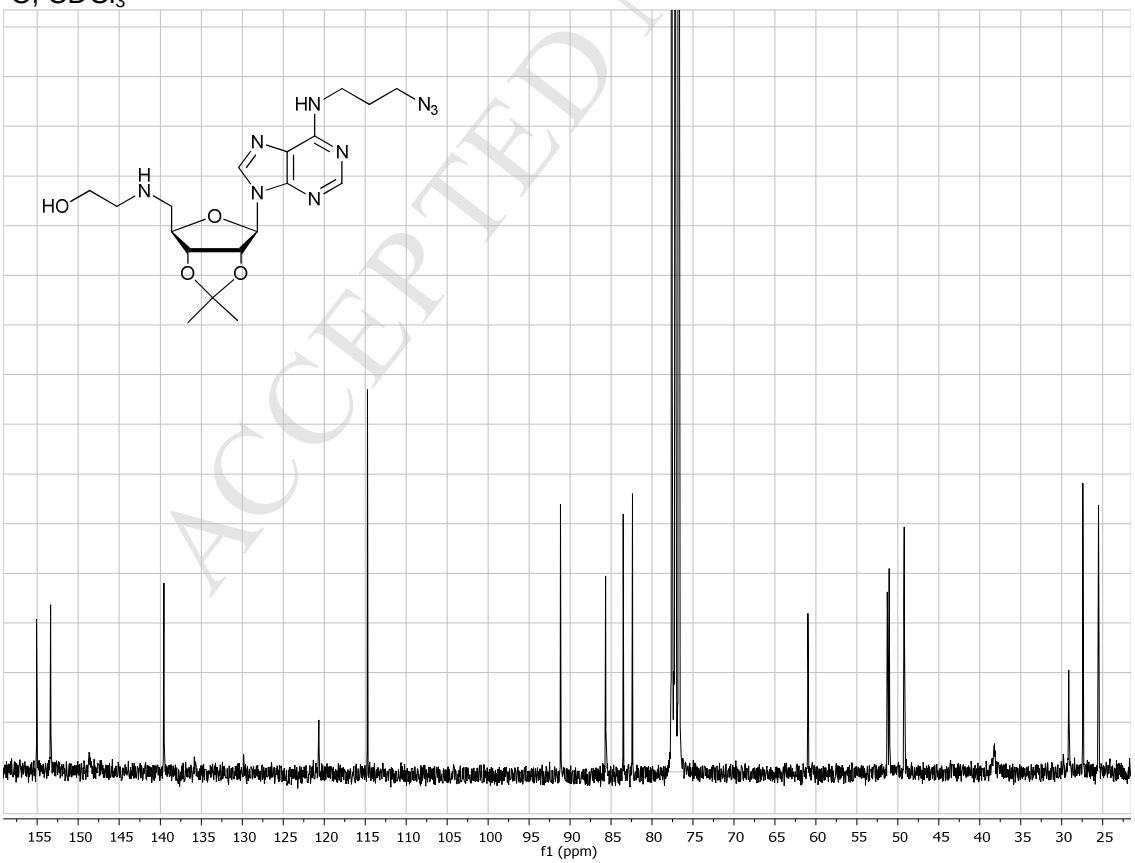
10c – ^{13}C, $CDCl_3$



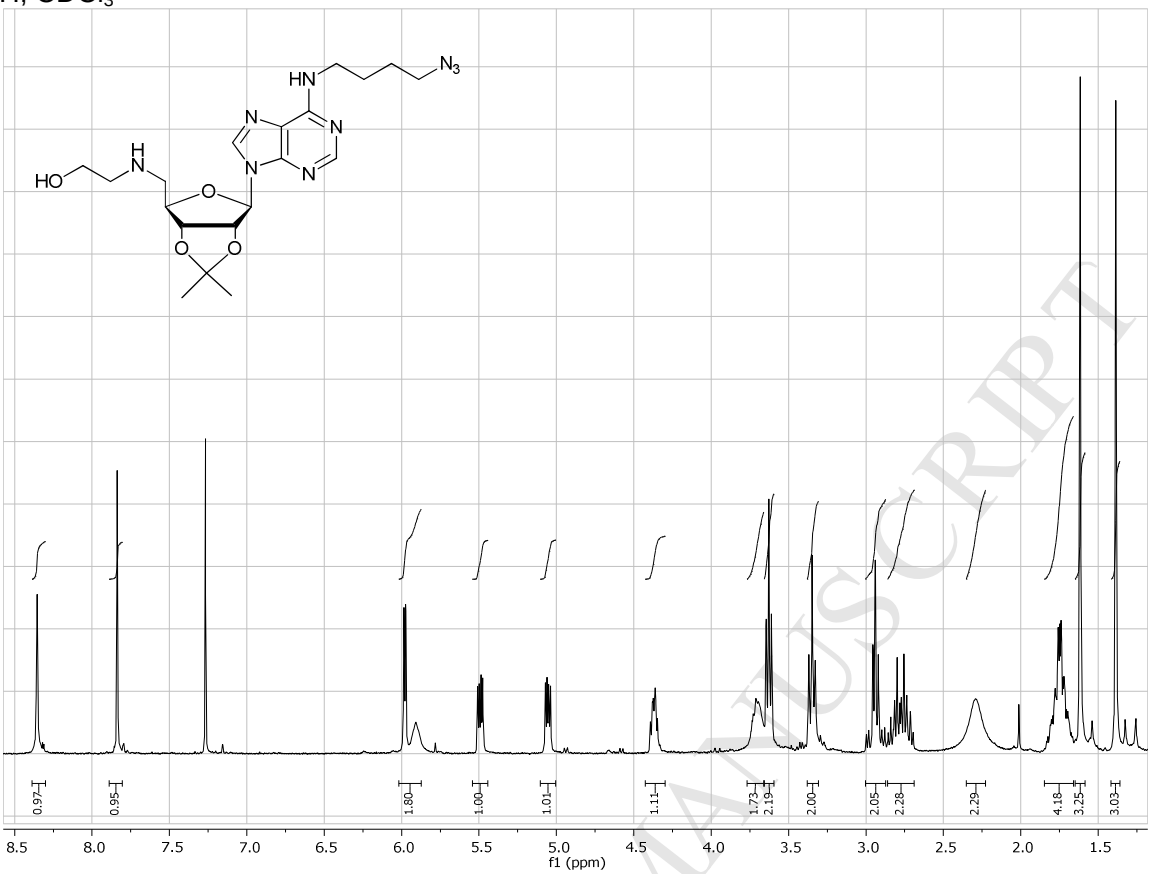
10d – ^{1}H, $CDCl_3$



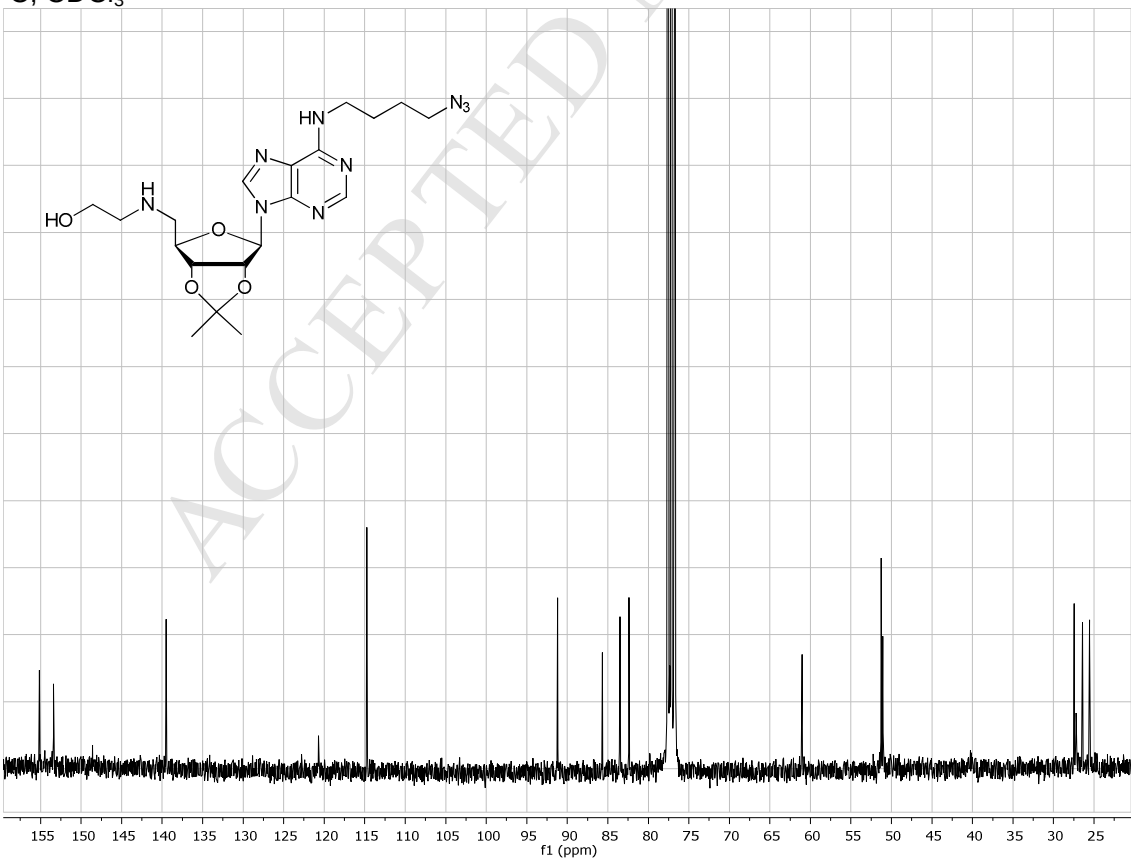
10d – ^{13}C, $CDCl_3$



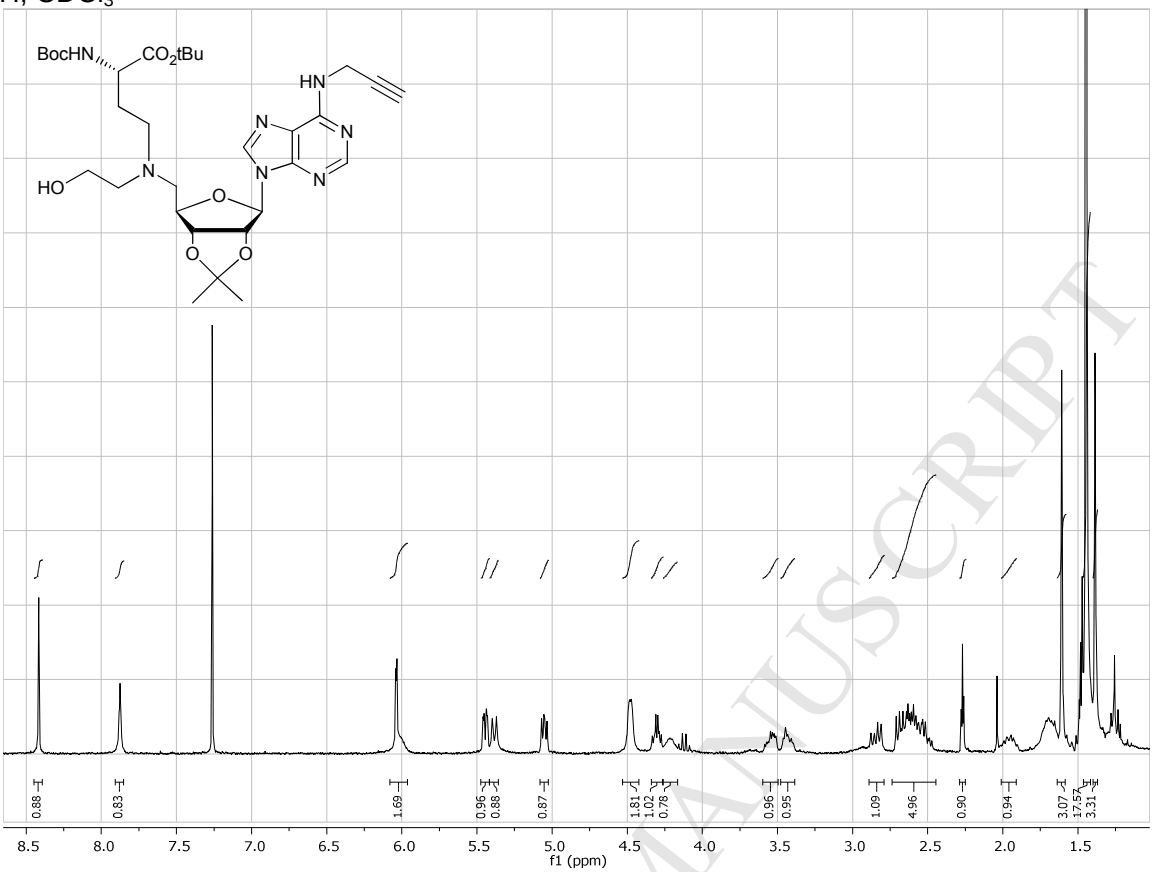
10e – ^{1}H, $CDCl_3$



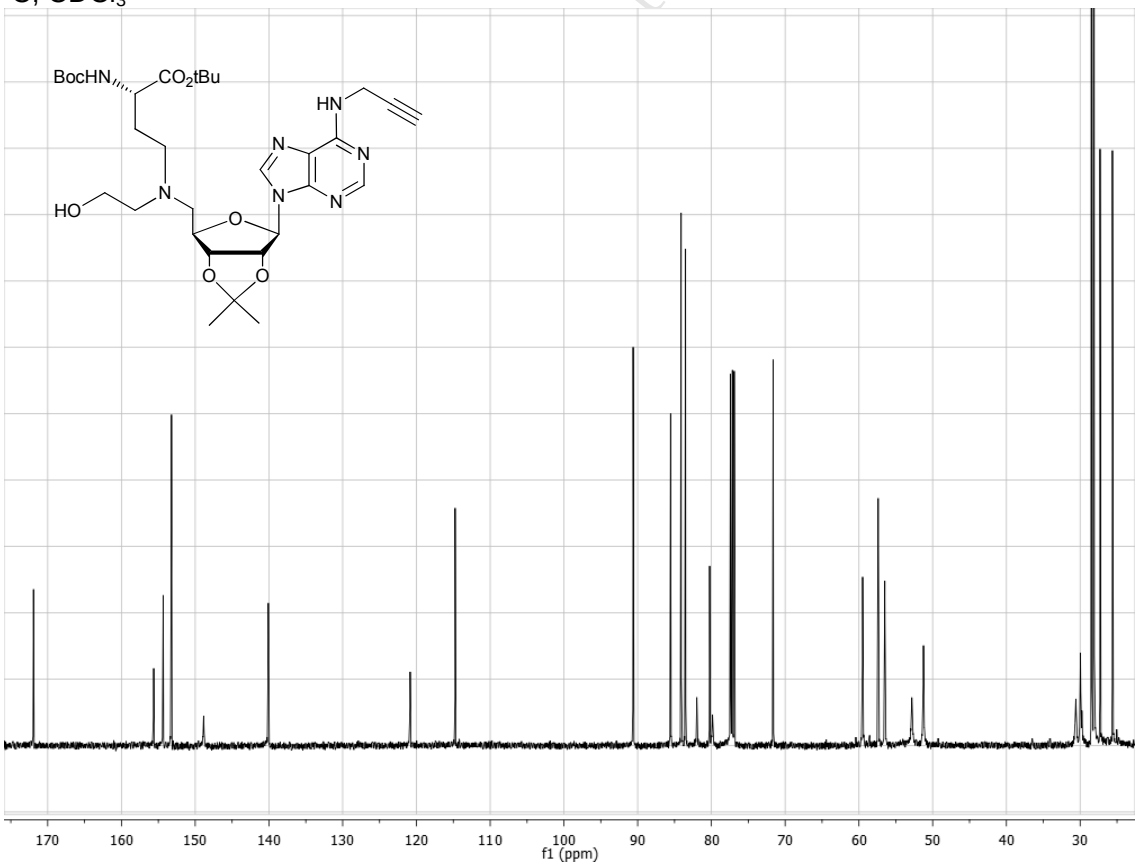
10e – ^{13}C, $CDCl_3$



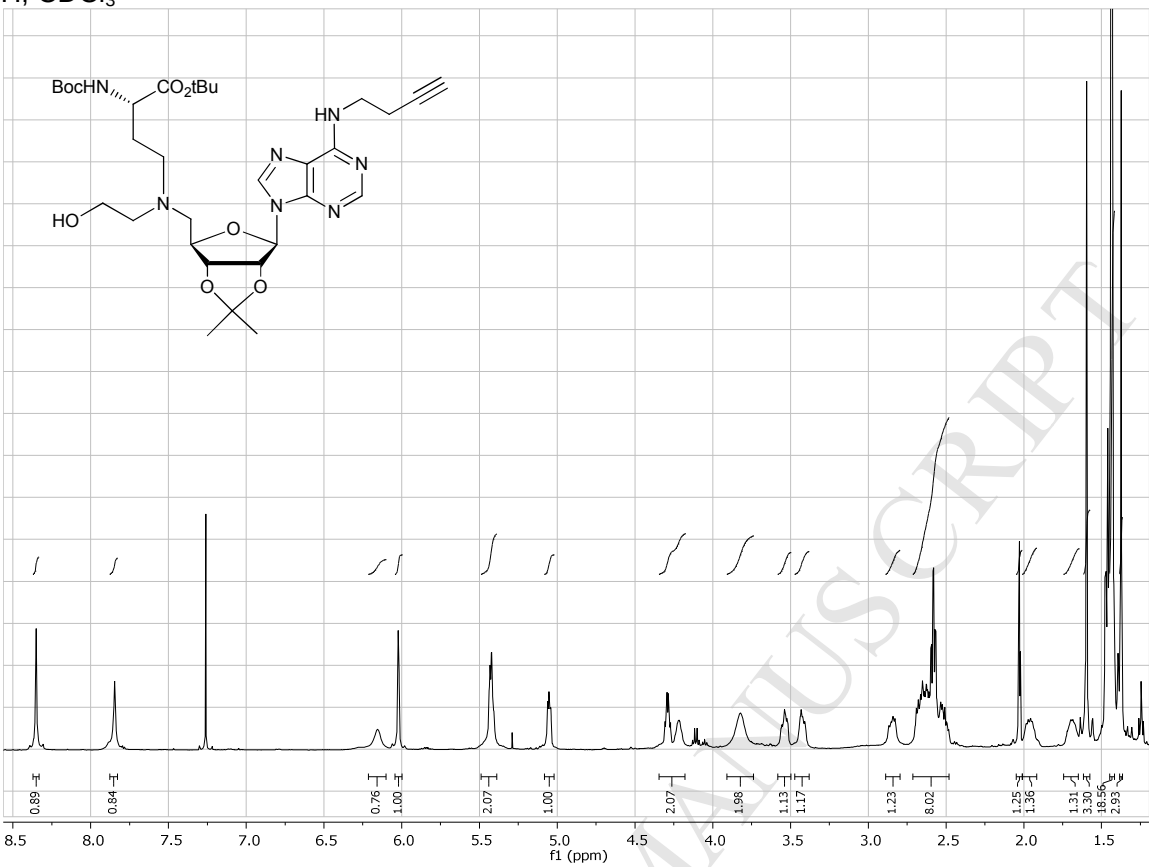
11a – ^{1}H, $CDCl_3$



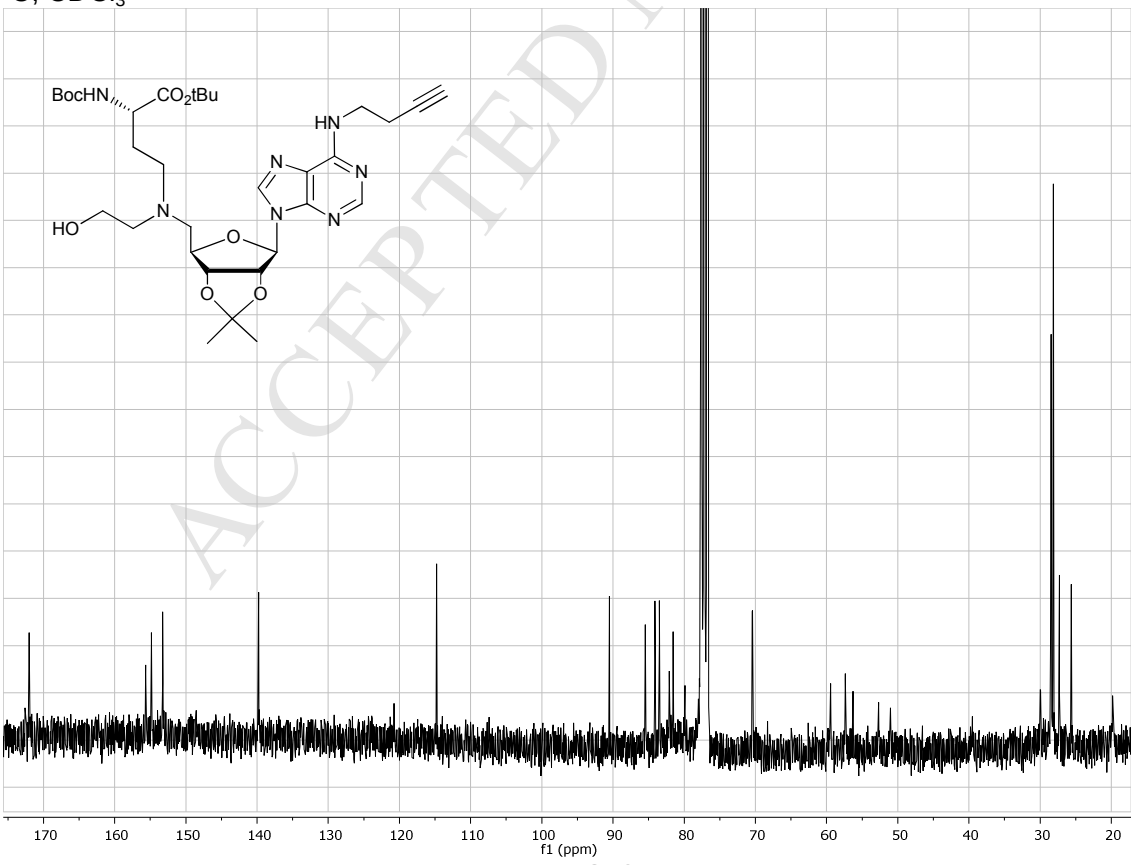
11a – ^{13}C, $CDCl_3$



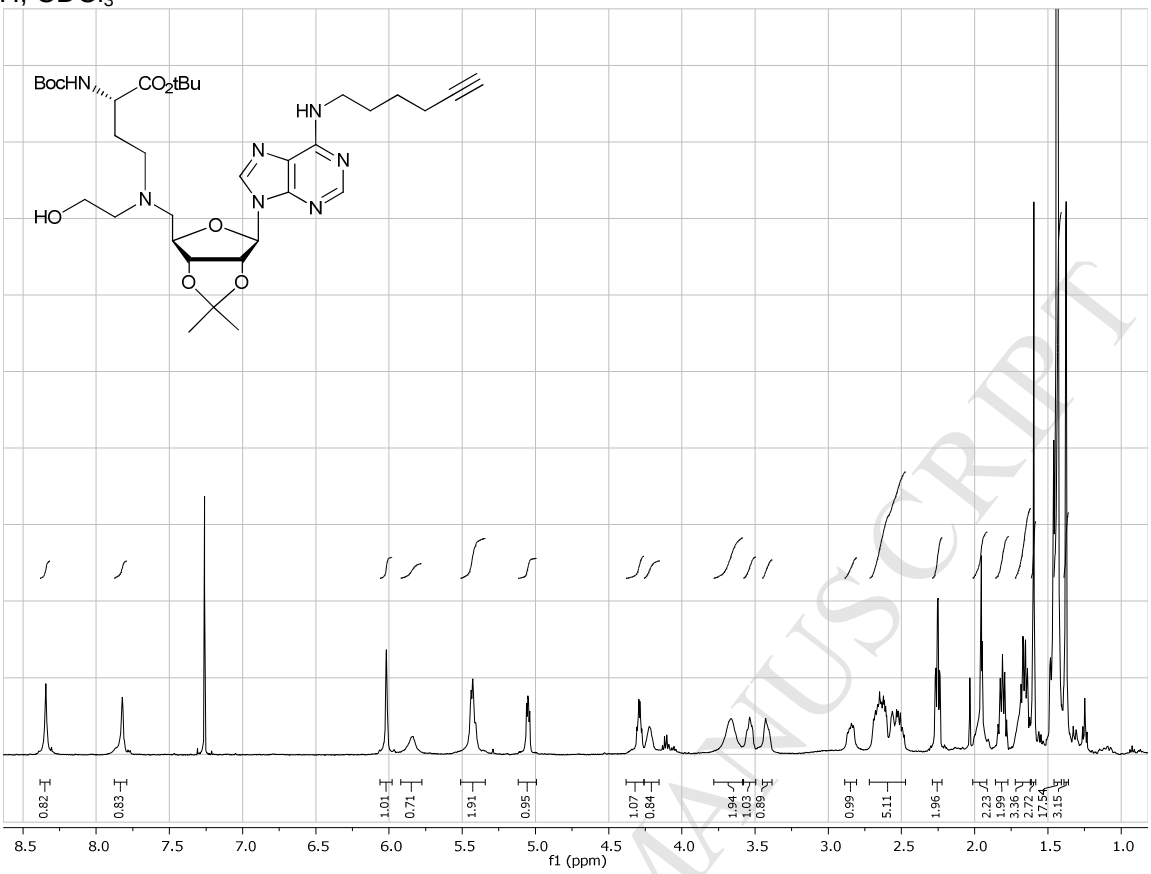
11b – ^{1}H, $CDCl_3$



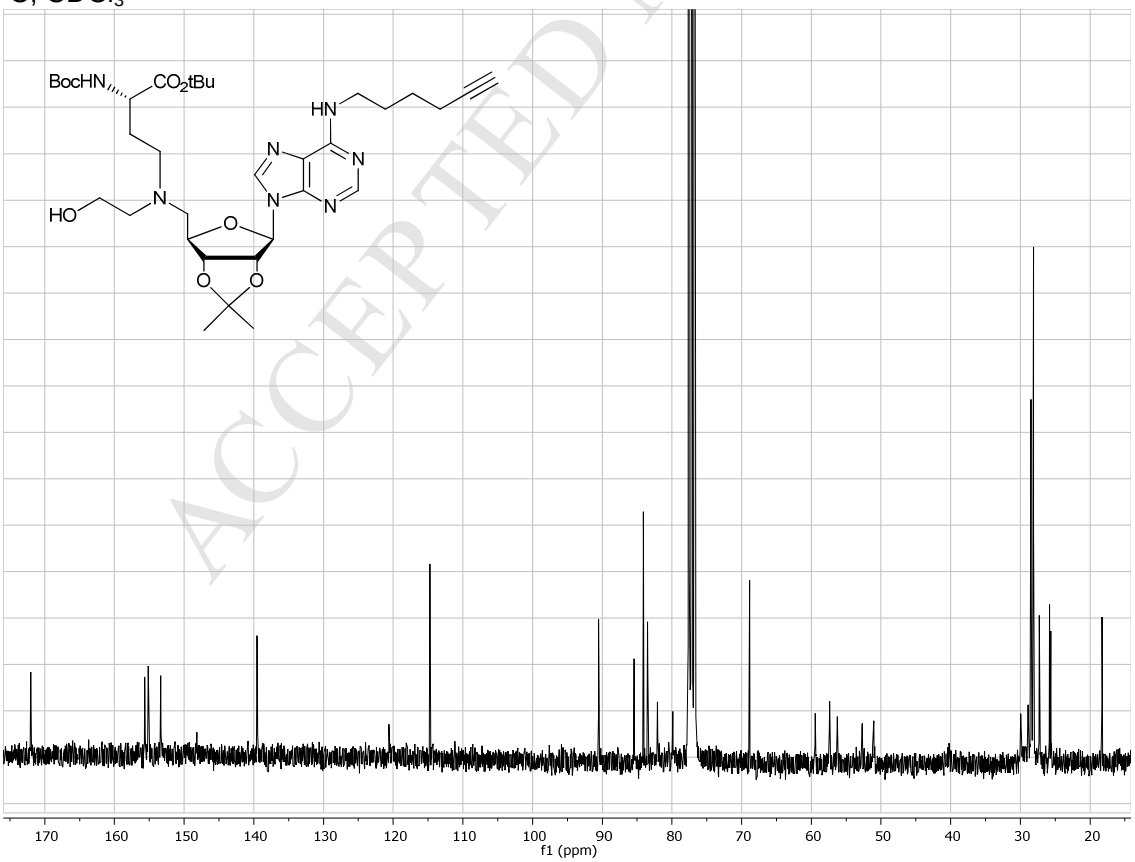
11b – ^{13}C, $CDCl_3$



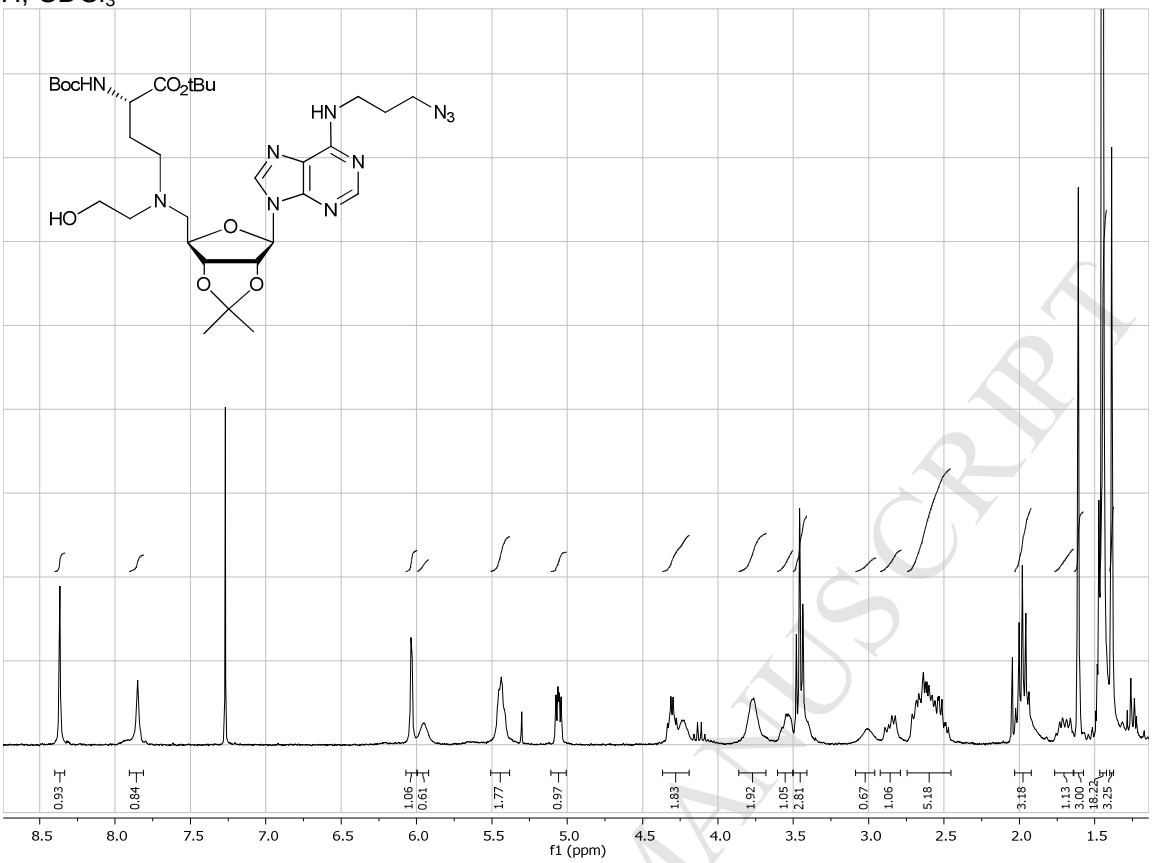
11c – ^{1}H, $CDCl_3$



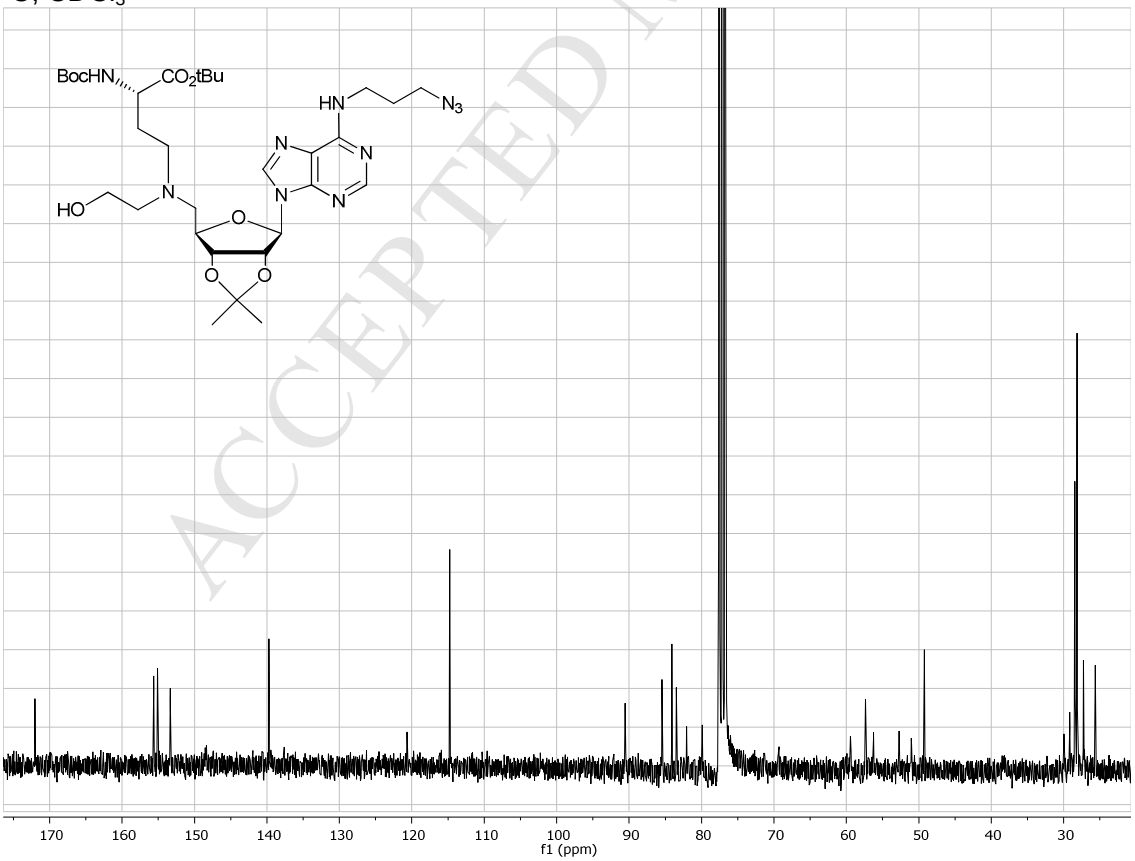
11c – ^{13}C, $CDCl_3$



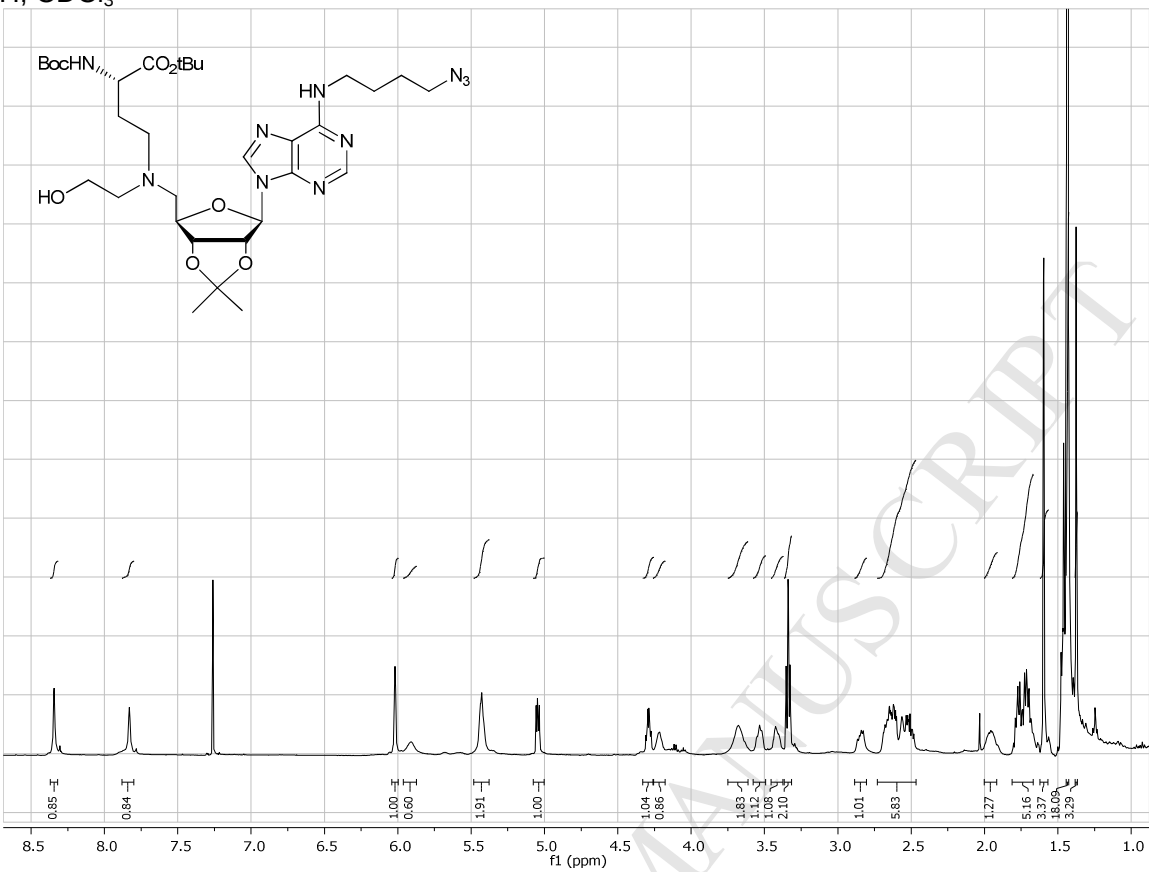
11d – ^{1}H, $CDCl_3$



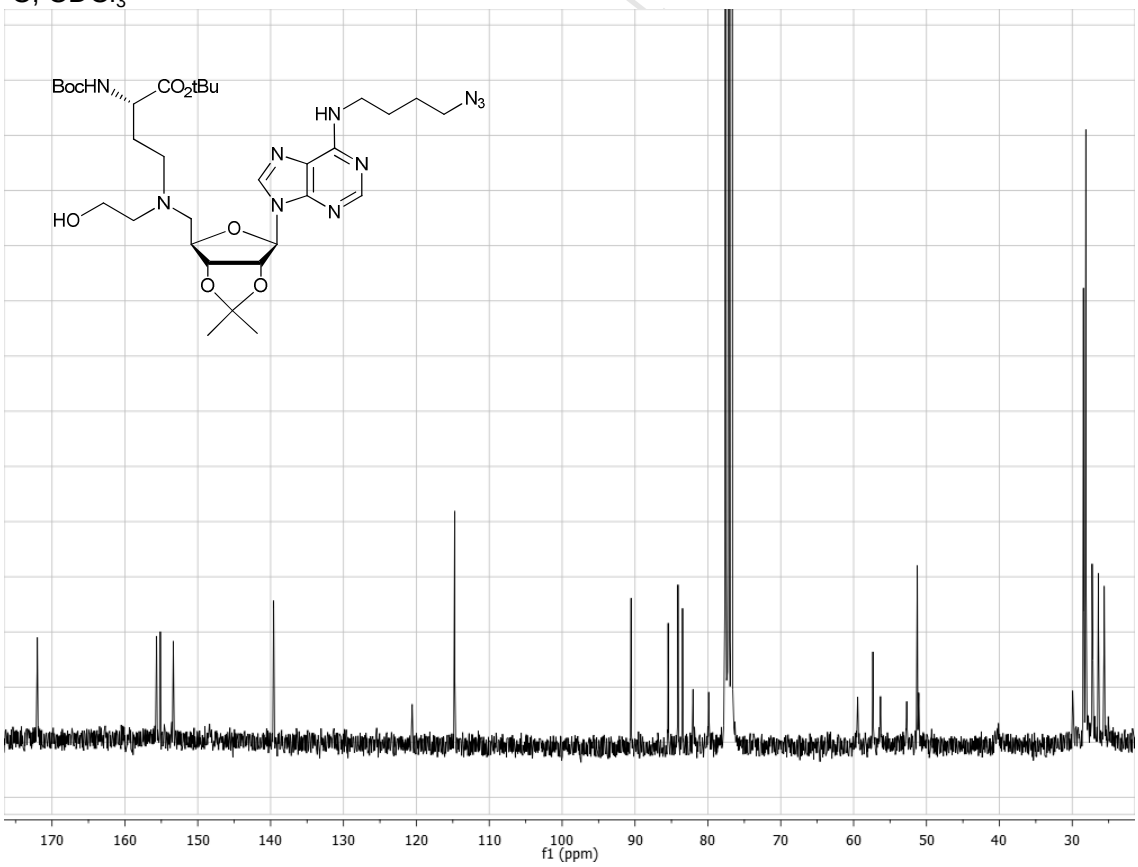
11d – ^{13}C, $CDCl_3$



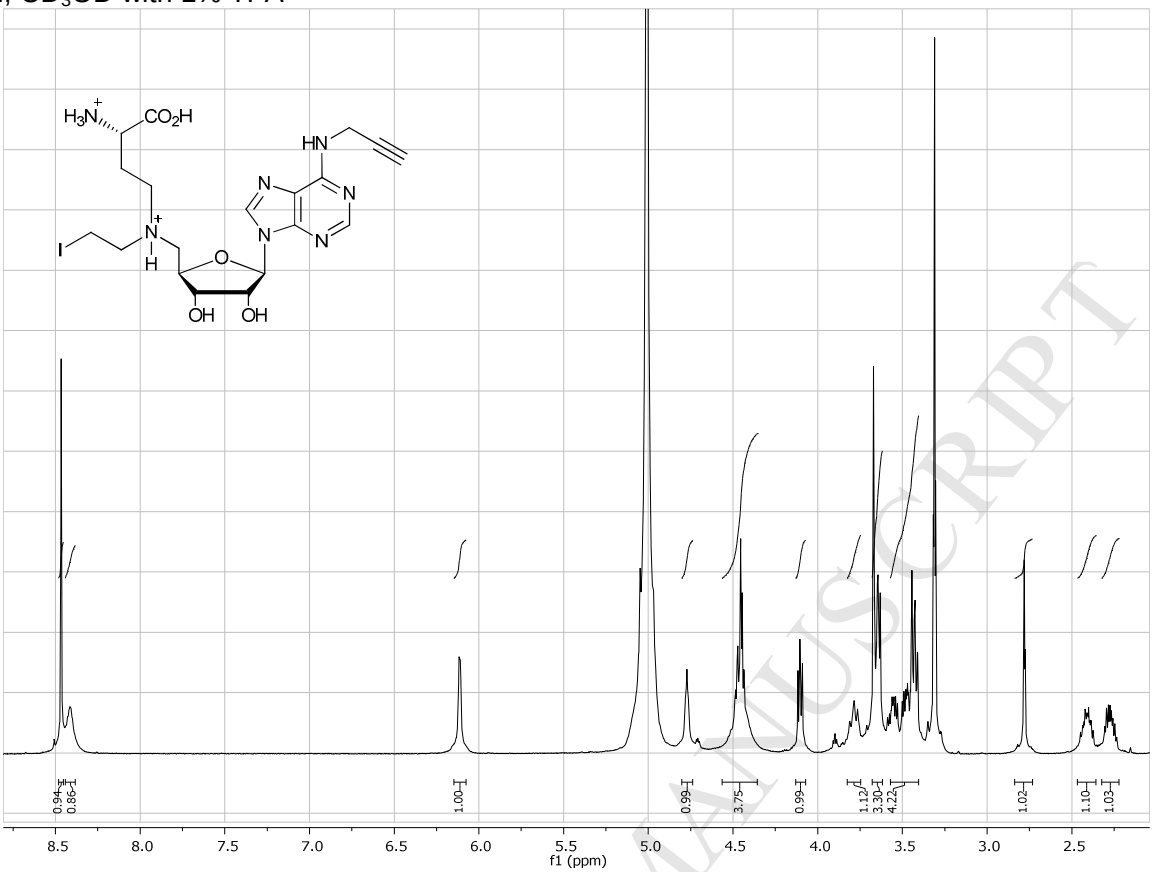
11e – ^{1}H, $CDCl_3$



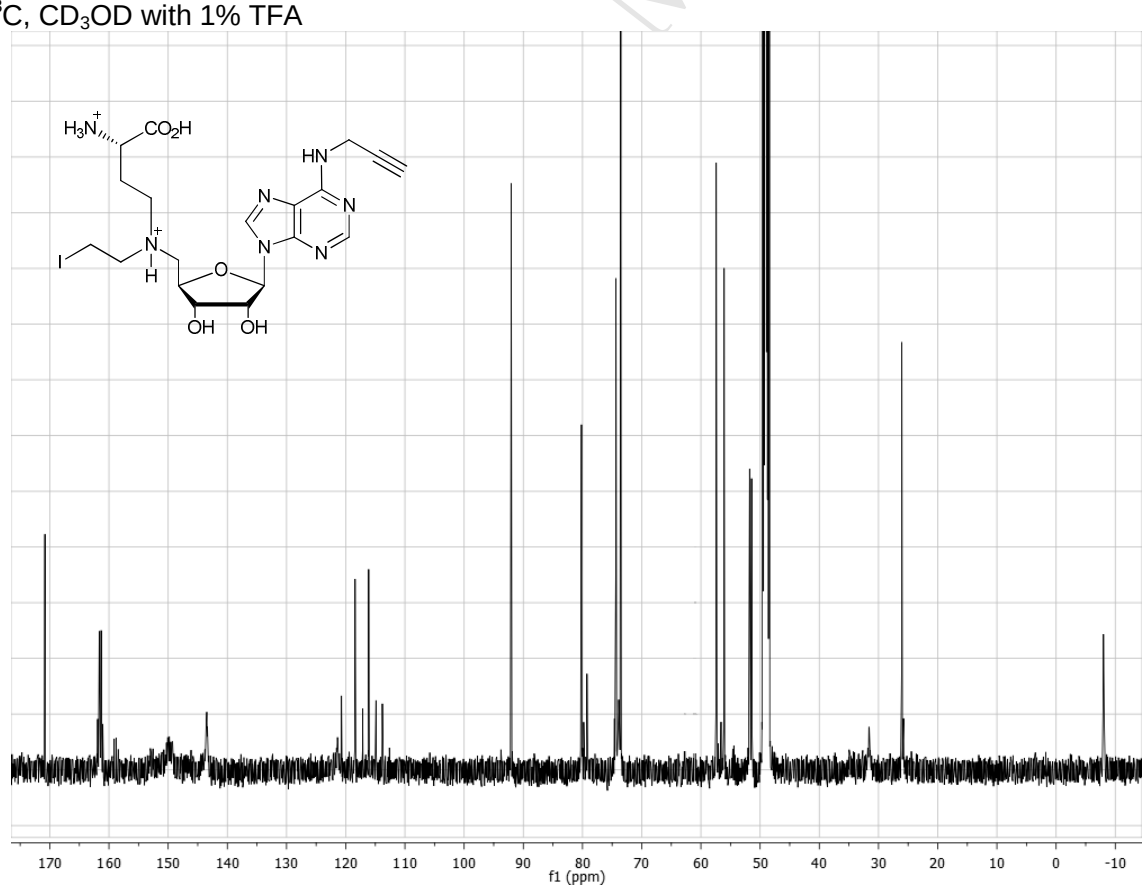
11e – ^{13}C, $CDCl_3$



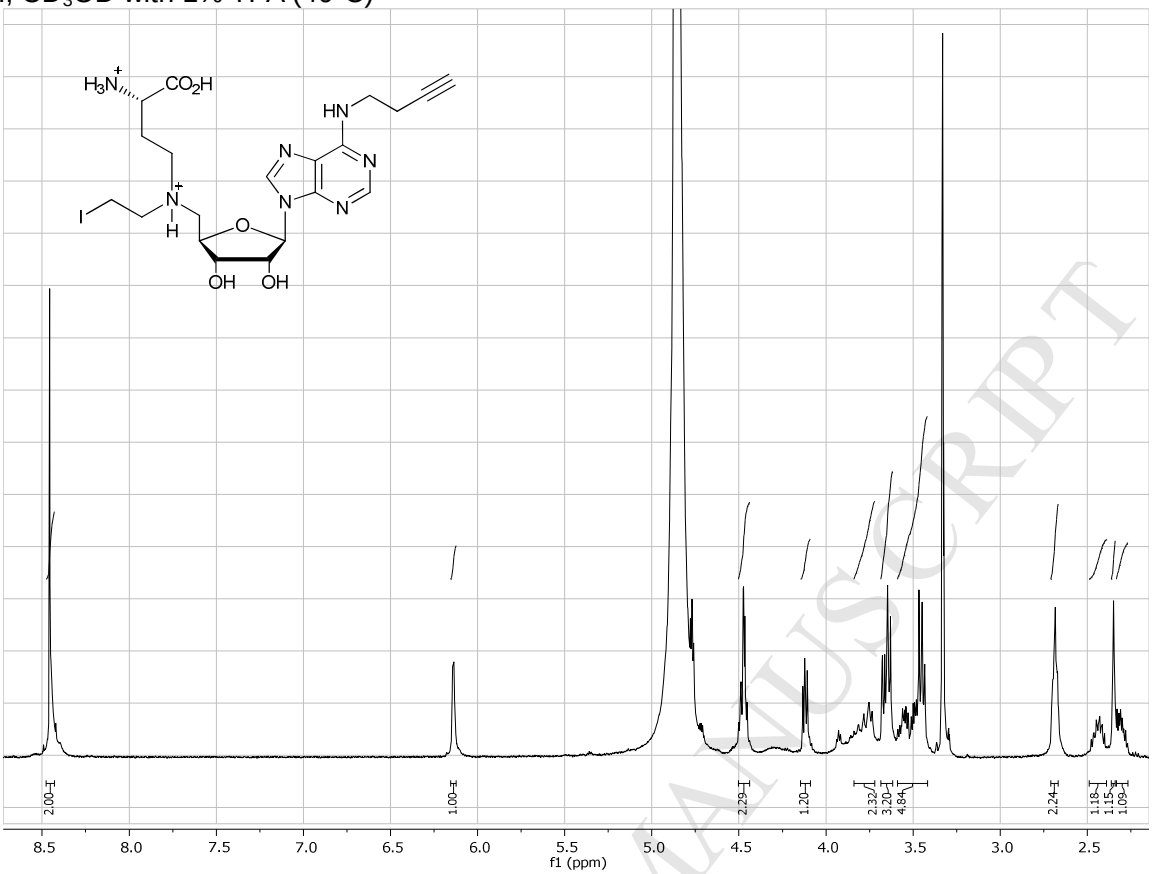
5a – ^{1}H, CD_3OD with 1% TFA



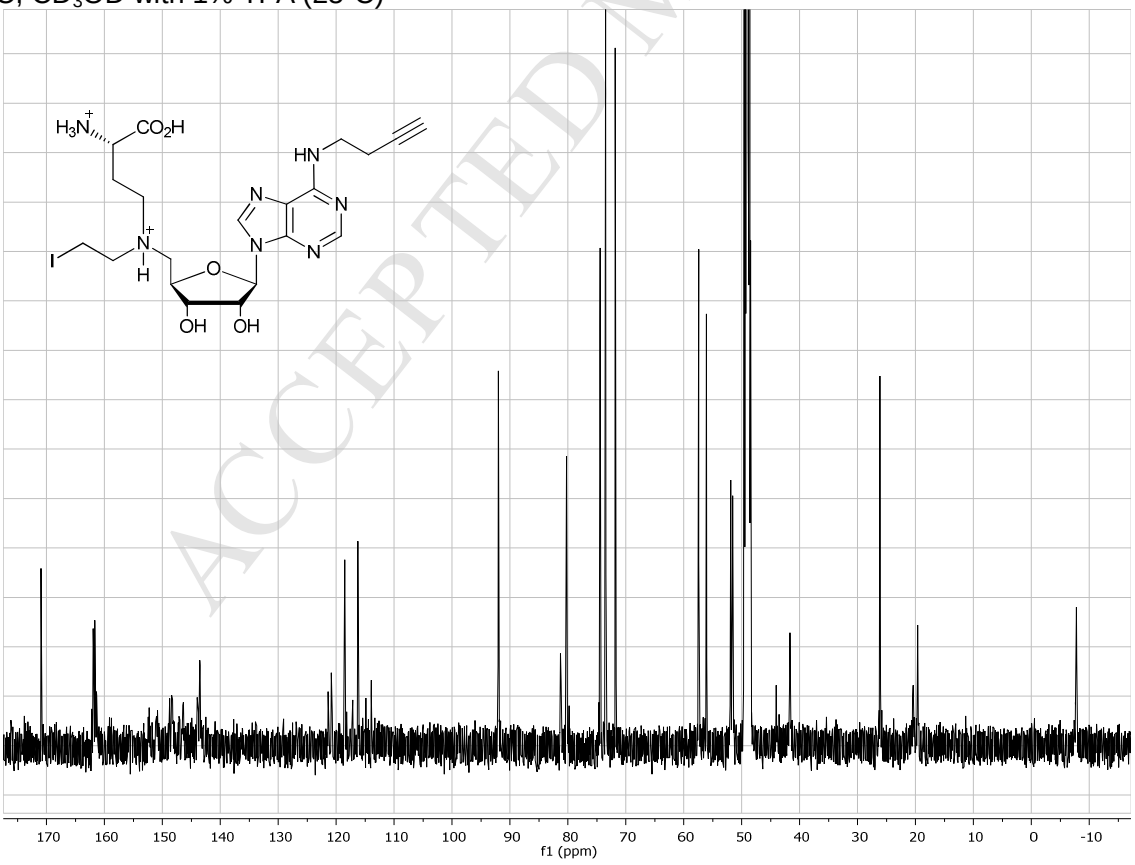
5a – ^{13}C, CD_3OD with 1% TFA



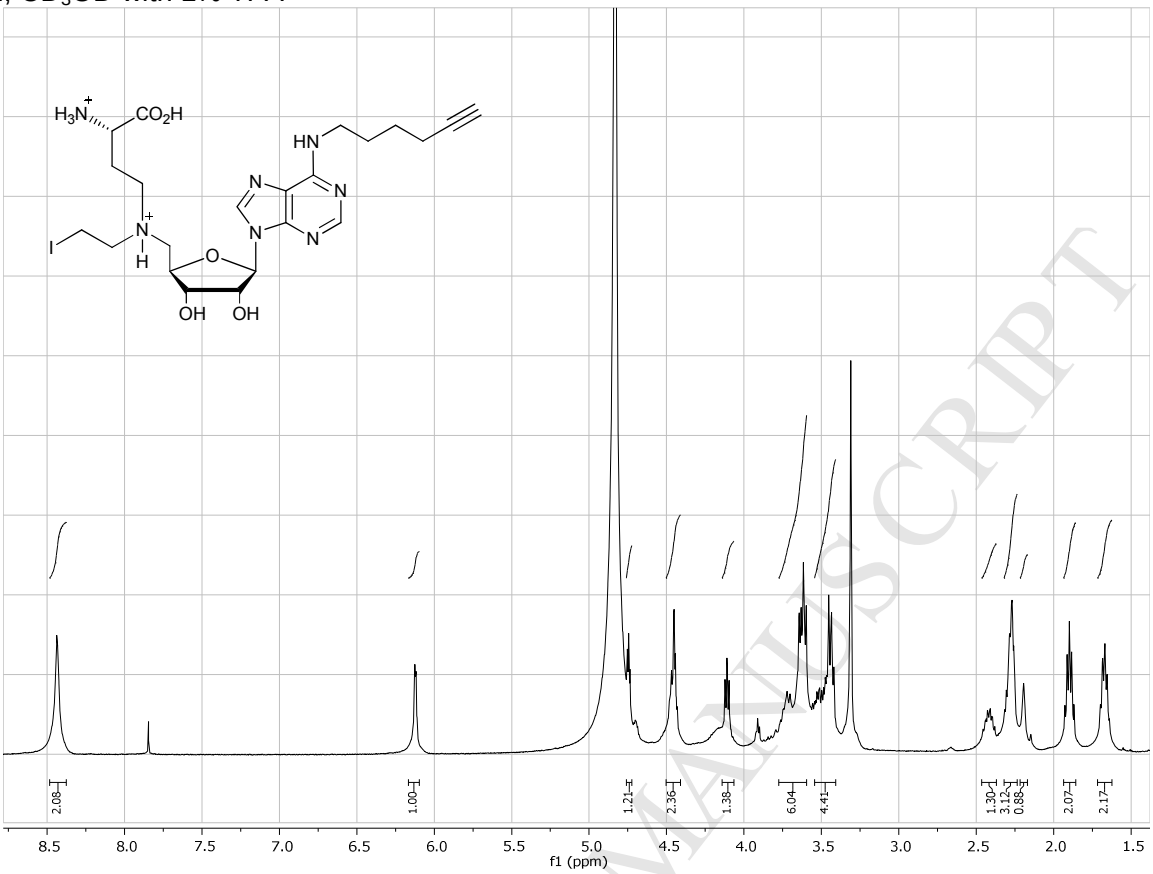
5b – ^{1}H, CD_3OD with 1% TFA (40ºC)



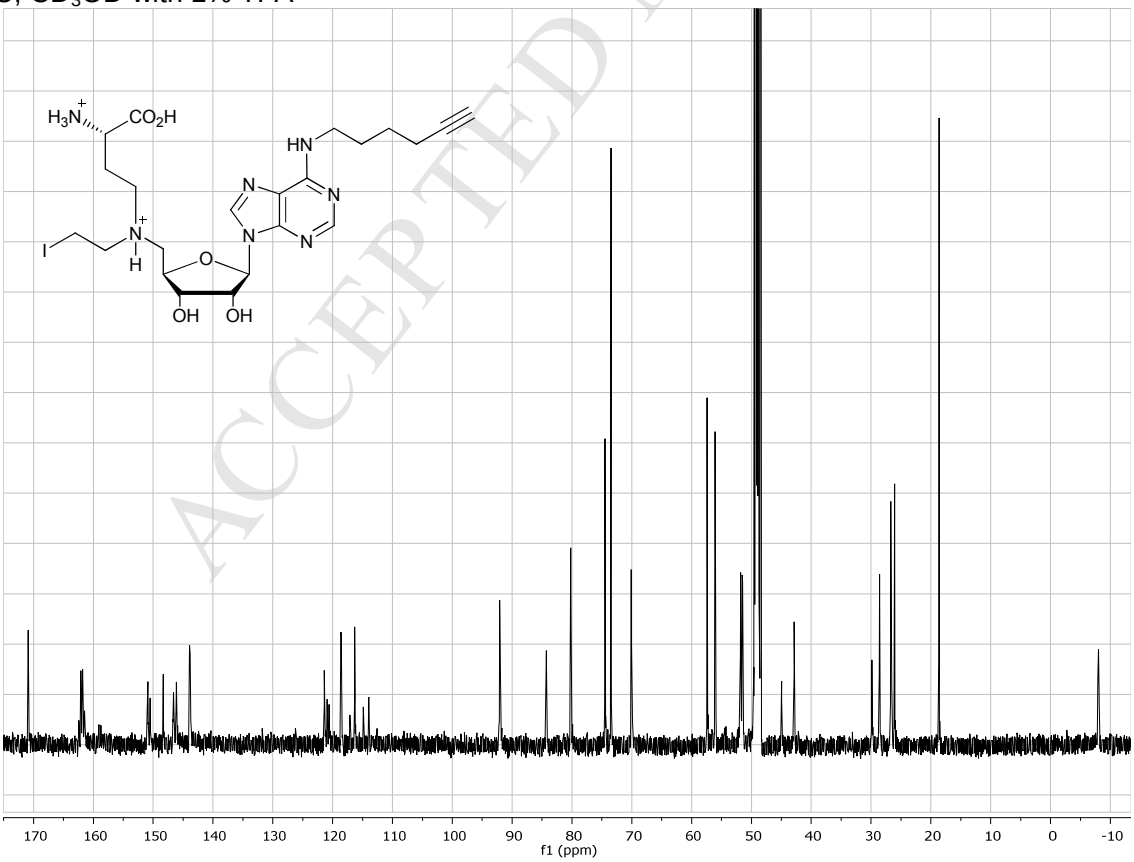
5b – ^{13}C, CD_3OD with 1% TFA (25ºC)



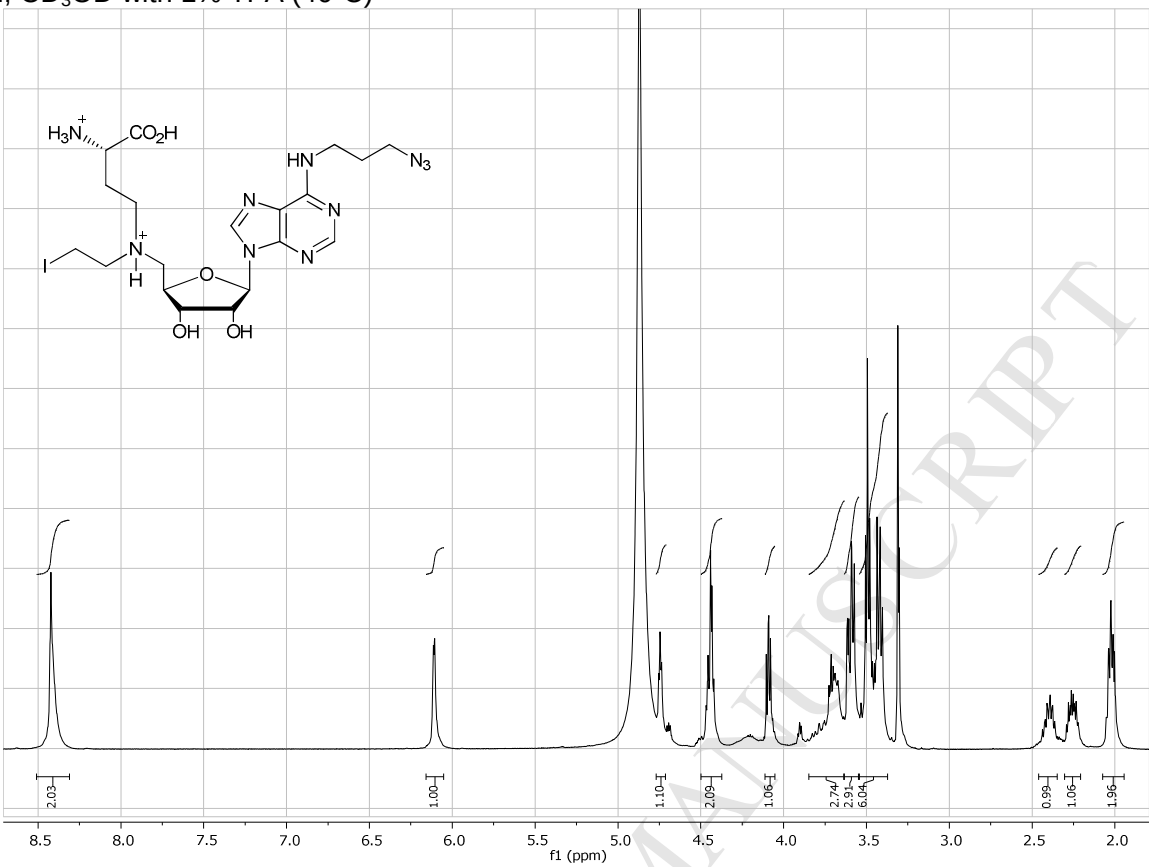
5c – ^{1}H, CD_3OD with 1% TFA



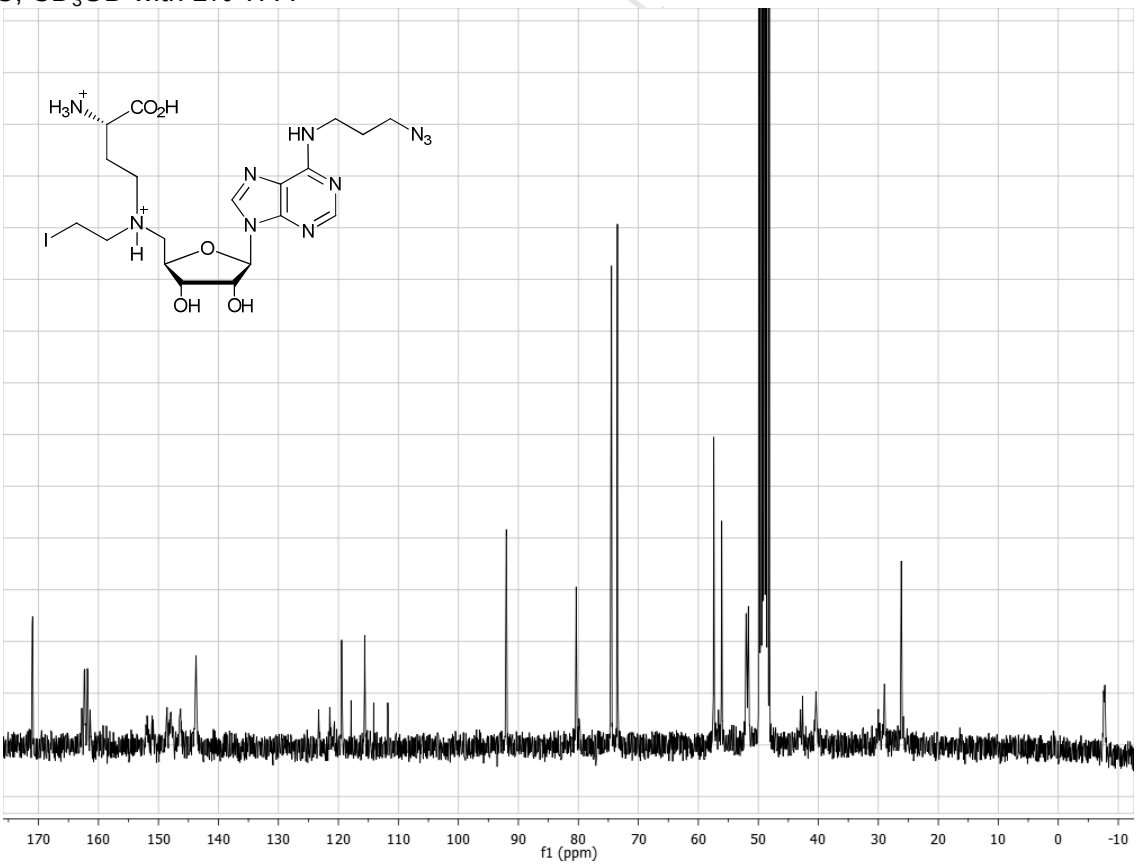
5c – ^{13}C, CD_3OD with 1% TFA



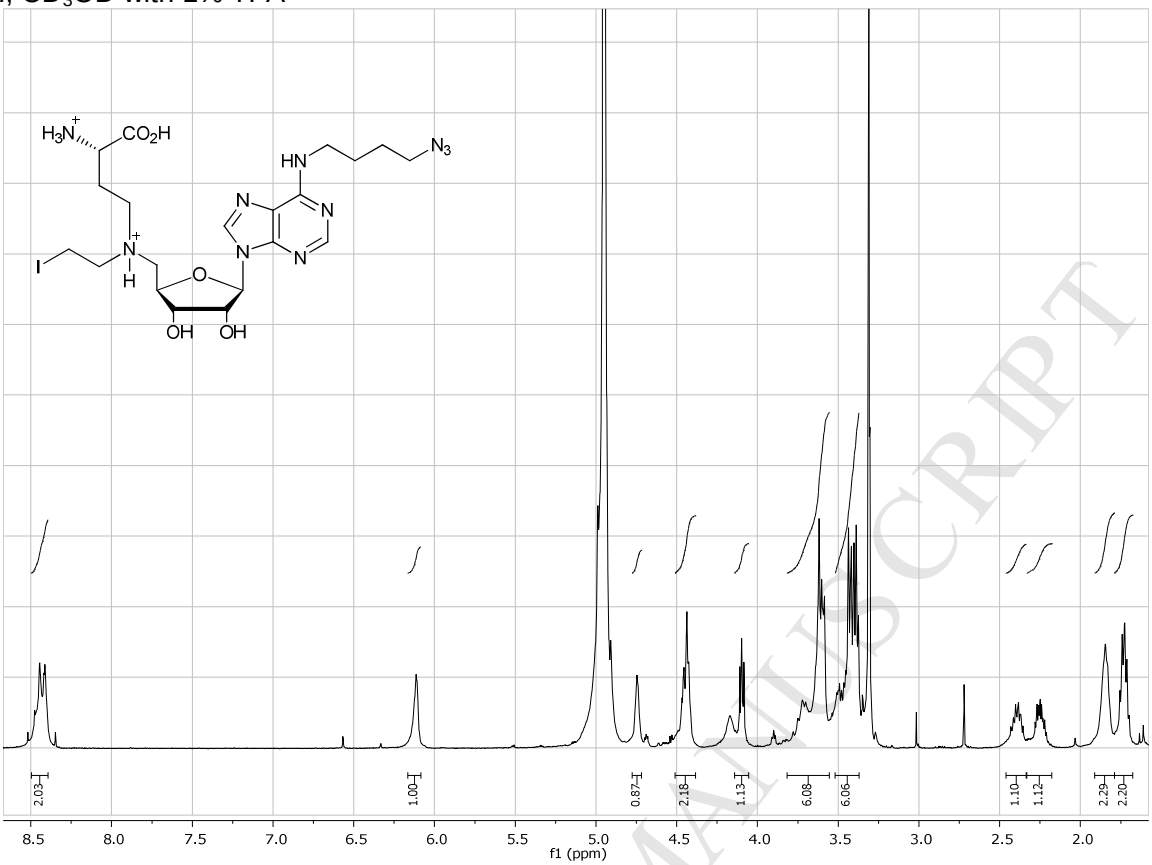
6a – ^{1}H, CD_3OD with 1% TFA (40ºC)



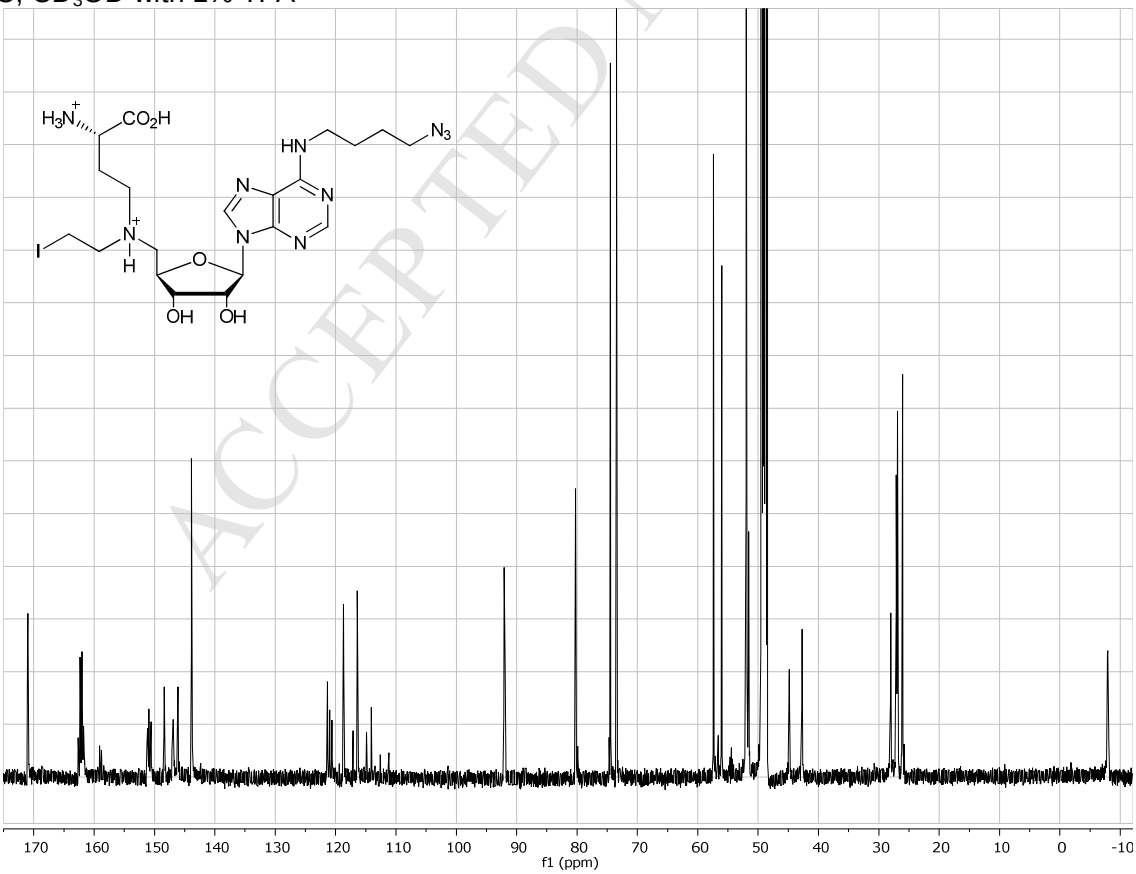
6a – ^{13}C, CD_3OD with 1% TFA



6b – ^{1}H, CD_3OD with 1% TFA



6b – ^{13}C, CD_3OD with 1% TFA



Restriction/protection assay with M.HhaI: DNA reaction mixtures for each SAM analog were prepared by the addition of appropriate stock solutions in M.HhaI buffer (total volume of 10 μL). R.EcoRI-linearized pUC19 (final concentration of 0.025 μg/μL) was treated with M.HhaI (50 U) and SAM analogs (final concentration at 10 μM and 50 μM for the specific reaction sequence). All reactions were incubated at 37 ℃ for 4 h, followed by 65 ℃ for 15 min. The extent of DNA modification was analyzed by the addition of R.HhaI (3 U in an additional 10 μL NEBuffer 4), followed by incubation at 37℃ for 1 h. Proteinase K (0.02 U in 5 μL H_2O) was added to each reaction and incubated an additional hour at 37℃. Agarose gel loading dye was added to each reaction and the extent of alkylation was visualized by electrophoresis on a 1 % agarose gel containing ethidium bromide. It is interesting to note that reaction analysis at lower concentrations of **5c** and **6b** does not generate clearly defined bands since pUC19 contains 17 recognition sites (compared to 4 for M.TaqI). Consequently, the unprotected DNA is digested into a complex mixture of products and appears as a "streak" following electrophoresis.

Supplementary Figure 1

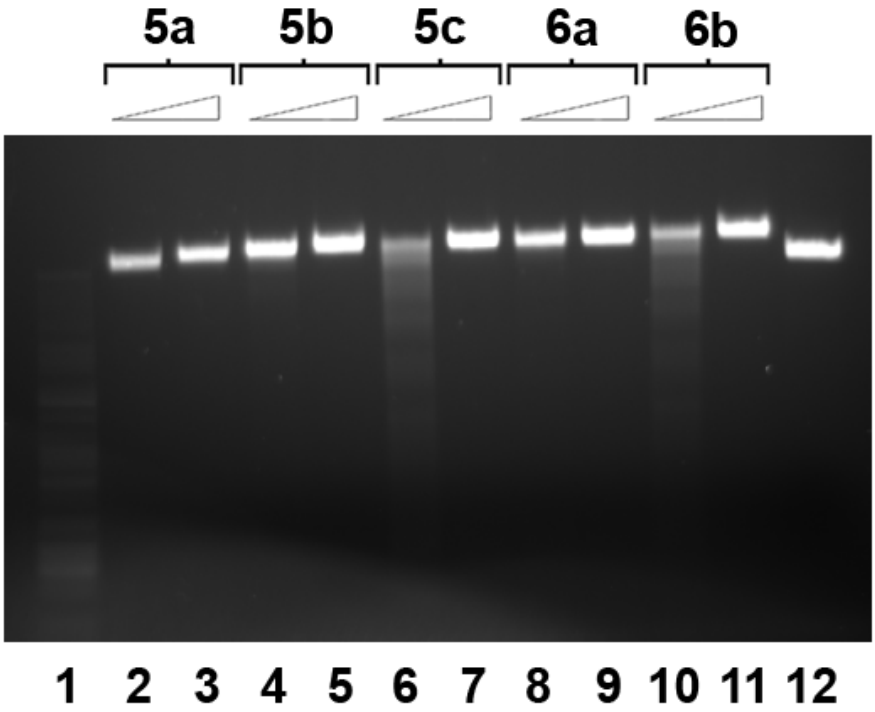


Figure S6. DNA alkylation reaction of R.EcoRI-linearized pUC19 with **5a**-**c** and **6a**-**b** using M.HhaI. Lane 1: DNA, R.HhaI; Lane 2: DNA, 10 μM **5a**, M.HhaI, R.HhaI; Lane 3: DNA, 50 μM **5a**, M.HhaI, R.HhaI; Lane 4: DNA, 10 μM **5b**, M.HhaI, R.HhaI; Lane 5: DNA, 50 μM **5b**, M.HhaI, R.HhaI; Lane 6: DNA, 10 μM **5c**, M.HhaI, R.HhaI; Lane 7: DNA, 50 μM **5c**, M.HhaI, R.HhaI; Lane 8: DNA, 10 μM **6a**, M.HhaI, R.HhaI; Lane 9: DNA, 50 μM **6a**, M.HhaI, R.HhaI; Lane 10: DNA, 10 μM **6b**, M.HhaI, R.HhaI; Lane 11: DNA, 50 μM **6b**, M.HhaI, R.HhaI; Lane 12: DNA.

General HPLC Methods: Analytical and semi-preparative HPLC was performed with the UV/VIS detector set at 254 nm. An Alltima^{TM} C18 analytical column (100 Å, 5 μm, 250 x 4.6 mm) and Alltima^{TM} C18 semi-preparative column (100 Å, 5 μm, 250 x 10.0 mm) were used for HPLC analysis and isolation of **5a**-**c** and **6a**-**b**. All mobile phases were filtered through a 0.22 μm membrane filter prior to use. Compound separation utilized a gradient system comprising of 0.1 % TFA in H_2O (solvent A) and HPLC grade ACN (solvent B) using a flow rate of 1 mL/min (analytical scale) or 4 mL/min (semi-preparative scale). The analytical gradient was run isocratically with 10 % B for 1 min followed by a linear gradient of 10 - 80 % B over a 9 min period. The gradient was then increased to 100 % B over the next 2 min and run isocratically for an additional 3 min. The preparative gradient was run isocratically with 10 % B for 1 min followed by a linear gradient of 10 - 60 % B over a 19 min period. The gradient was then increased to 100 % B over the next 2 min and run isocratically for an additional 2 min.

Supplemental Figure 2

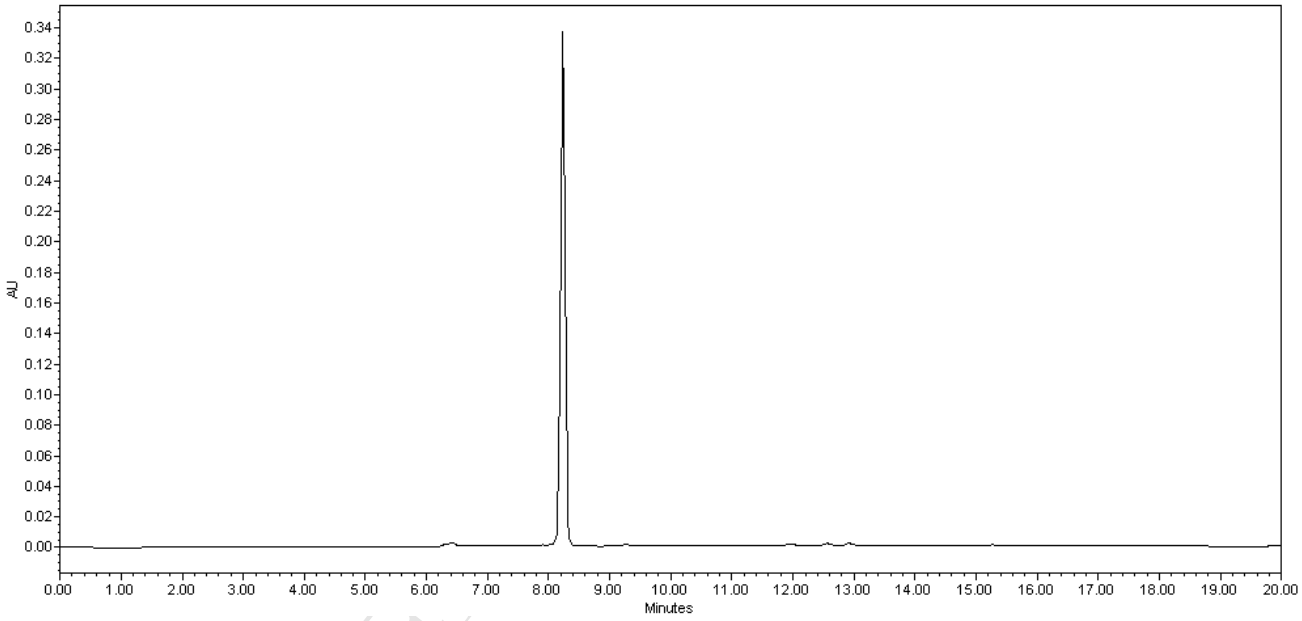


Figure S2. Analytical HPLC of **5a** after purification. **5a** eluted at 8.23 min.

Supplemental Figure 3

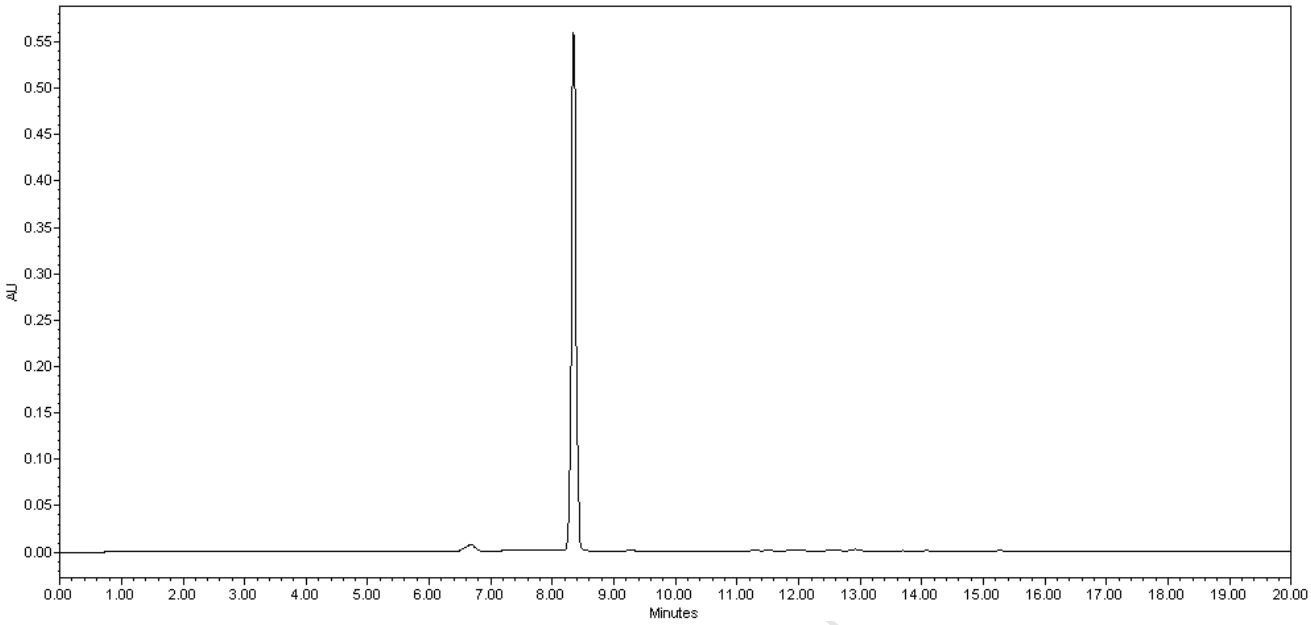


Figure S3. Analytical HPLC of **5b** after purification. **5b** eluted at 8.33 min.

Supplemental Figure 4

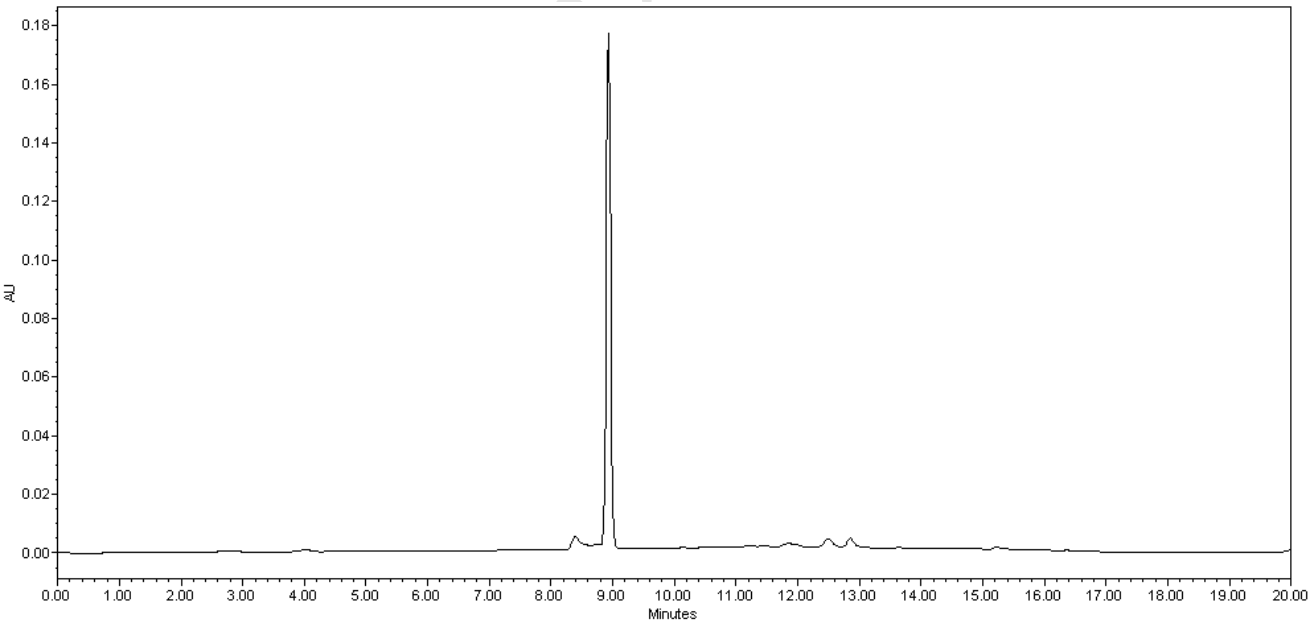


Figure S4. Analytical HPLC of **5c** after purification. **5c** eluted at 8.93 min.

Supplemental Figure 5

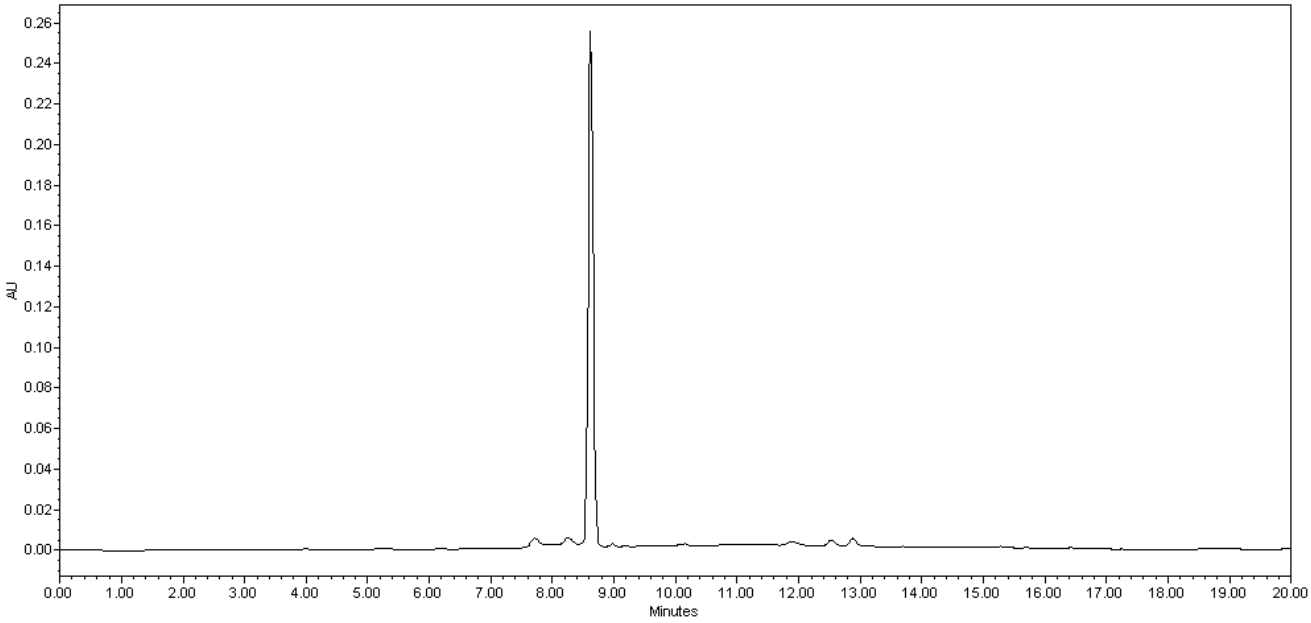


Figure S5. Analytical HPLC of **6a** after purification. **6a** eluted at 8.62 min.

Supplemental Figure 6

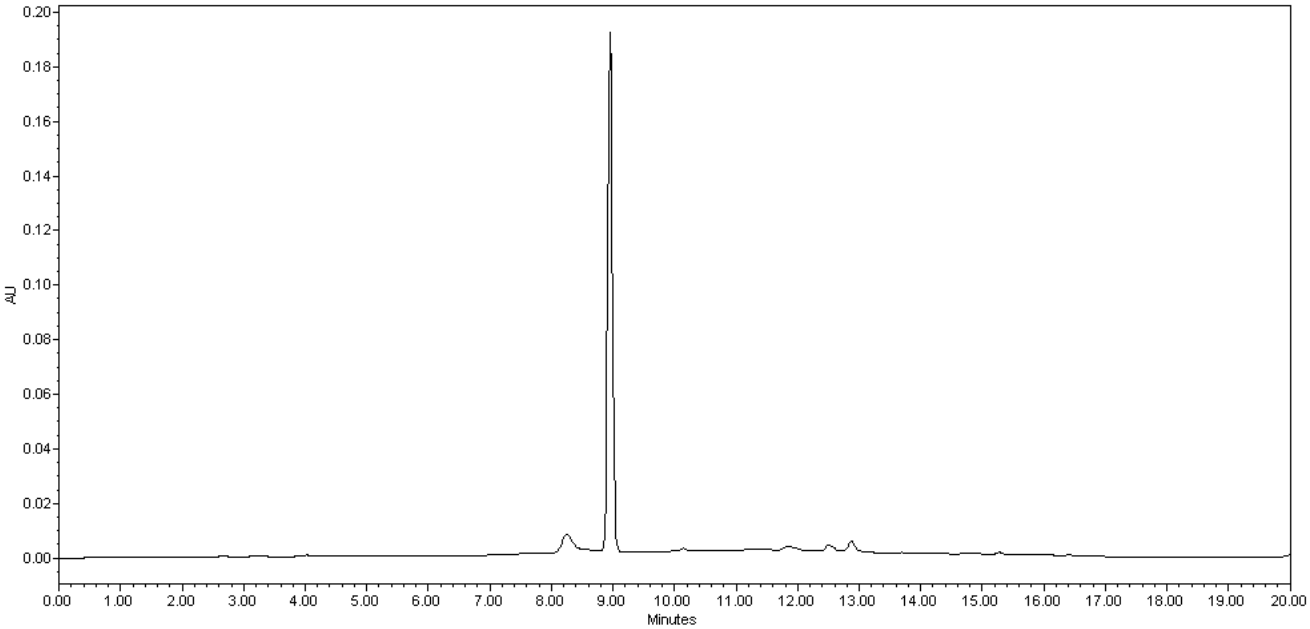


Figure S6. Analytical HPLC of **6b** after purification. **6b** eluted at 8.95 min.