A Chemical Probe Based on the PreQ1 Metabolite Enables Transcriptome-wide Mapping of Binding Sites

Sumirtha Balaratnam[1*], Curran Rhodes[1*], Desta Doro Bume[1], Colleen Connelly[1], Christopher C. Lai[1], James A. Kelley[1], Kamyar Yazdani[1], Philip J. Homan[2,3], Danny Incarnato[4], Tomoyuki Numata[5,6] and John S. Schneekloth Jr[1,†]

Supporting Information

Materials and methods:

General Information

Unless otherwise stated, all reagents and solvents were purchased from commercial vendors and used without a further purification. All ^{1}H, ^{13}C and ^{19}F NMR spectra were collected on a 500, 126 and 471 MHz NMR spectrometers in $(CD_3)_2SO$, respectively. The ^{1}H, ^{13}C, and ^{19}F NMR chemical shifts are given in parts per million (δ) with respect to an internal tetramethylsilane (TMS, δ = 0.00 ppm) standard and/or 3-chlorobenzotrifluoride (δ = –64.2 ppm relative to $CFCl_3$). NMR data are processed, annotated and reported in the following format: chemical shift (multiplicity (s = singlet, brs = broad singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constants (Hz), Integration) using MestReNova. The large-scale synthesis of 7-aminomethyl-7-deazagaunine (PreQ1) was carried out according to previously reported protocol.

Oligonucleotides

Deprotected and HPLC purified riboswitch aptamers or full-length riboswitch were purchased from Dharmacon with the following sequences and used without any further purifications.

Tt-PreQ1-RS: 5'-CUGGGUCGCAGUAACCCCAGUUAACAAAACAAG-3'
Cy5-Tt-PreQ1-RS: 5'-Cy5-CUGGGUCGCAGUAACCCCAGUUAACAAAACAAG-3'
Bs-PreQ1-RS: 5'-AGAGGUUCUAGCUACACCCUCUAUAAAAAACUAA-3'
Cy5-Bs-PreQ1-RS: 5'-Cy5-AGAGGUUCUAGCUACACCCUCUAUAAAAAACUAA-3'
Ss-PreQ1-RS: 5'-AGAGGUUCCUAGCUGAUACCCUCUAUAAAAAACUA-3'
Cy5-Ss-PreQ1-RS: 5'-Cy5-AGAGGUUCCUAGCUGAUACCCUCUAUAAAAAACUA-3'
SAM-II-WT-RS:5'-UCGCGCUGAUUUAACCGUAUUGCAAGCGCGUGAUAAAUGU
AGCUAAAAAGGG
miR21-hp: 5'-GGGUUGACUGUUGAAUCUCAUGGCAACCC-3'
Bs-PreQ1-RS full length (with aptamer domain and the expression platform): 5'-GCGGGAGAGGUUCUAGCUACACCCUCUAUAAAAAACUAGGACGAGCUGUAUCCU
UGGAUACGGCCUUUU-3'

Optimized conditions for photocrosslinking to PreQ1-riboswitch and click chemistry as analyzed by denaturing PAGE

UV photocrosslinking to PreQ1-riboswitch

A PreQ1 aptamer/full riboswitch (10 μM, 40 μL) in 1 mM $MgCl_2$ buffer was heated to 75 ºC for 5 minutes and allowed to cool to room temperature over an hour. At this point, a photoaffinity probe dissolved in DMSO (500 μM) was thoroughly mixed with the RNA and incubated for additional 30 minutes in dark at room temperature. Note that it is advised to keep the DMSO level at or below 5% in the reaction mixture. Then, the reaction mixture was irradiated with 365 nm light sources for 15 minutes using SPECTROLINKER XL-1000 UV Crosslinker. Consequently, reaction mixture was analyzed by MALDI-TOF and/or denaturing PAGE.

For the PAGE analysis, RNA samples were mixed with denaturing loading dye and followed by heating at 95°C for 5 minutes. Equal volumes of the samples were then subjected to 17% PAGE in the presence of 7 M urea. Then gel was stained in 1X SYBR® Gold Nucleic Acid Gel Stain (Invitrogen) for 15 minutes at room temperature. The gel images were obtained by scanning the gel on a AMERSHAM TYPHOON Biomolecular Imager. The Image J software was used for processing the images. The crosslinking efficiency was calculated by quantifying the higher molecular weight band intensity (cross-link adduct band) with Image J.

Click reaction on photocrosslinked RNA for labeling with fluorophores and biotin-azides

Biotin or fluorophore-azide probes were crosslinked to RNA according to the established protocol[1]. The RNA was resuspended in RNase and DNase free water and used for labeling experiments. Note that this crude solution contains a mixture of photocrosslinked and unmodified RNA (approximately 1:1.5 photocrosslinked to unmodified). To the above solution (40 μL of 66 μM), either Biotin-PEG_3-azide or Cy5-picolyl-azide (10 μL of 10 mM solution in DMSO), THPTA/$CuSO_4$ (25 μL of premixed solution of THPTA and $CuSO_4$ for 10 minutes in 2:1 ratio using 200 mM and 100 mM solutions in H_2O, respectively), and L-Sodium ascorbate (7.5 μL of 100 mM in H_2O) were added and the reaction mixture was stirred for 1 h. The resulting adducts were analyzed by MALDI-TOF and denaturing PAGE.

Single-round transcription termination assay

The transcription termination assays were carried out as previously reported[2].

X-Ray Crystallography Experiments and Data Collection

For X-ray crystallography, we used the PreQ1 riboswitch aptamer domain from *Thermoanaerobacter tengcongensis* (Tt). The wild-type aptamer and abasic mutant (ab13_14), in which the nucleobases at positions 13 and 14 were removed, were employed for crystallization. The crystals of Tt PreQ1 riboswitch aptamer domain in complex with **11** were prepared at 20 ˚C and cryoprotected by the same methods as

reported previously[2]. X-ray diffraction data were collected at the beamlines BL32XU and BL45XU of the SPring-8 (Hyogo, Japan) with the aid of an automatic data-collection system ZOO[3] and BL-1A of the Photon Factory (Ibaraki, Japan). Diffraction data were integrated and scaled with the programs KAMO[4] and XDS[5]. Data processing statistics are summarized in Table S1.

Table S1: Summary of data collection and refinement statistics

	ab13_14-**11**	Wild type-**11**
Data collection		
Space group	$P6_122$	$P6_322$
Cell dimensions		
a, b, c (Å)	52.7, 52.7, 177.0	113.2, 113.2, 59.5
Wavelength (Å)	1.0	1.0
Resolution (Å)	45.6-1.57 (1.66-1.57)	41.0-2.80 (2.97-2.80)
R_{merge}	0.124 (1.86)	0.177 (1.93)
$\langle I \rangle / \langle \sigma I \rangle$	17.2 (1.47)	9.43 (1.25)
Completeness (%)	99.9 (99.4)	98.0 (96.1)
Redundancy	20.5 (19.8)	8.2 (8.5)
Refinement		
Resolution (Å)	45.6-1.57	41.0-2.80
No. reflections	21,395	5,768
R_{work}[a]/R_{free}[b]	0.190/0.208	0.193/0.235
No. atoms		
RNA	694	686
Ligand	22	22
Ions	13	5
Water	86	2
B-factors (Å^2)		
RNA	40.2	87.1
Ligand	37.5	97.7
Ions	110.3	147.0
Water	46.3	61.1
R.m.s. deviations		
Bond lengths (Å)	0.004	0.006
Bond angles (°)	0.900	1.175
PDB ID	7E9E	7E9I

The values in parentheses are for the outermost shell.
[a]$R_{work} = \Sigma|F_o - F_c|/\Sigma F_o$ for reflections of working set.
[b]$R_{free} = \Sigma|F_o - F_c|/\Sigma F_o$ for reflections of test set (5.0% of total reflections).

Structure determination and refinement

The co-crystal structures were solved by the molecular replacement method with the program PHASER[6], using the previously determined Tt PreQ1 aptamer domain structure (PDB ID: 6E1W) as the search model. The solutions were subjected to simulated annealing, energy minimization, restrained isotropic B-factor, and TLS refinement with PHENIX[7]. The resulting electron density maps revealed the locations of the compound. The atomic models were then improved by iterative cycles of refinement with PHENIX[7] and manual rebuilding with COOT[8]. The current co-crystal structures of ab13_14 and the wild-type Tt PreQ1 aptamer domain in complex with **11** were refined to an R_{free} of 0.208 and 0.235 at 1.57 and 2.80 Å resolution, respectively.

PDB accession codes

Atomic coordinates and the structure factors of the co-crystal structures of ab13_14 and the wild-type Tt PreQ1 aptamer domain in complex with **11** have been deposited in the Protein Data Bank, under the accession codes 7E9E and 7E9I, respectively.

Orbitrap-LC/MS analysis to probe a crosslinked residue

A photocrosslinked RNA subjected to ethanol precipitation as specified above and resuspended in RNase and DNase free water (100 μL of 50 μM) was transferred to a PCR tube and treated with the NEB Nucleoside Digestion Mix® Reaction Buffer (10 μL, 10X) and the Nucleoside Digestion Mix (5 μL) and incubated at 37 ºC for 1 h. Then, there reaction mixture was analyzed by Orbitrap-LC/MS.

Photocrosslinking experiments in the presence of total cellular RNA

RNA Sample Preparation

MCF-7 cells were grown in culture media (DMEM + 10% FBS, 1% ABS, 5% glutamine) to 80% confluency in a T75 culture flask. The day of the experiment the cells were harvested by trypsinization, washed and resuspended in PBS. Total RNA was isolated from MCF-7 cells using the RNeasy mini kit (Qiagen, Cat #: 74104) with on column DNAse digestion. The concentration of eluted RNA was measured using a nanodrop and the RNA was split into 9 different tubes (5 μg RNA/condition) for compound treatment. 20 mM Compound stocks were prepared in DMSO. The total RNA was diluted up to 100 μL total volume with riboswitch buffer (50 mM Tris (pH 7.5), 100 mM KCl and 2 mM $MgCl_2$). 1 μM of the full length PreQ1 BS aptamer (70 nts, Dharmacon) was added to the appropriate tubes during the dilution with riboswitch buffer. The PreQ1 aptamer was refolded in riboswitch buffer before use by heating at 90°C for 5 minutes and allowing to cool to RT over the course of 1 hour.

Denaturing PAGE Gel Experiments

Compounds were added along with DMSO (to 5% total volume for all conditions), and the samples were incubated in the dark for 30 minutes to establish an equilibrium. After incubation, the solutions were moved to a clear plastic 96-well microtiter plate (VWR) for

UV crosslinking. The control sample (**Fig. 6**, lane 7), that would not be irradiated, was left in the dark in the original tube. Samples in the plate were irradiated for 15 minutes at 365 nm using a SPECTROLINKER XL-1000 UV Crosslinker. After UV irradiation the samples were moved into fresh Eppendorf tubes for click labeling with TAMRA azide. The click reaction was performed as described above. The RNA was then precipitated from solution by adding 10 μl of 3M sodium acetate buffer (PH = 5.5) and 3 volumes of ethanol to each sample. Tubes were stored at -80°C overnight to facilitate RNA precipitation. The pellet was washed the following morning with 70% ethanol and the samples were prepared in 2X TBE urea sample buffer (Alfa Aesar). Samples were boiled at 90°C for 5 minutes and loaded onto a 6% denaturing PAGE gel. The gel was run at 30 W for 1.5 hours to ensure proper migration of the PreQ1 aptamer. TAMRA fluorescence was assessed on an AMERSHAM TYPHOON Biomolecular Imager and the gel was subsequently stained with SYBR gold nucleic acid stain (Invitrogen) and re-imaged to assess total RNA in each lane. Images were processed using the ImageJ software package.

RNA Sequencing Experiments

For the sequencing experiments, the total RNA was prepared and treated (n = 4 for all conditions) as described above however, after UV crosslinking, the modified RNA was labeled with biotin-azide. After labeling the RNA was subjected to ethanol precipitation to wash away excess biotin-azide. The RNA pellets were dissolved in 50 μL of incubation buffer (50 mM Tris pH 7.5, 100 mM KCl, 0.01% Tween-20) and 50 μL of streptavidin linked magnetic beads (NEB) was added to each tube. After thirty minutes, the beads were pulled down on a magnetic tower and washed three times with 500 μL incubation buffer. Finally, 50 uL of elution buffer (95% formamide/H_2O, 10 mM EDTA) was added to each tube and the tubes were placed on a 95°C heat block for 4 minutes. The supernatant containing the eluted RNA was removed on the magnetic tower and the concentration was measured using a QuBit RNA HS Assay Kit. 200 ng of each sample was subsequently used to prepare sequencing libraries with the QiaSeq Stranded Total RNA Library Kit (Qiagen). Completed libraries were sent to Novogene for sequencing on a HiSeq 2500 (2 x 150 reads). Four replicates of each treatment condition were prepared and sequenced.

Sequencing Data Analysis

For the analysis of these pulldown experiments, a reference was built by using the longest isoform for every human RefSeq transcript, plus the sequence of the BS-PreQ1-RS full length aptamer (see above). Reads were aligned to this reference using Bowtie2 and the "--very-sensitive-local" preset[9]. Resulting BAM files were then processed to TSV count files, containing the number of reads mapping to each transcript. To identify transcripts significantly enriched in the PreQ1-dz-yne pulldown with respect to the control (adjusted p-value < 0.05), the DESeq2 package was used[10].

For the sequencing experiments without the aptamer spiked-in, the volcano plot generated by the DESeq2 package was used to identify potential hits that were

significantly enriched (FC > 1.5-fold, FDR < 0.05 when comparing **11** treated samples to **17** treated samples).

GEO accession codes

The raw fastq files along with processed alignment files and counts files from these sequencing experiments have been deposited on the NCBI's gene expression omnibus (GEO) database. These files can be accessed with the following GEO accession code: GSE168353.

Photocrosslinking experiments in the presence of cell lysates with biotinylated probe 11

The cell lysates were extracted from the MCF-7 cells with RIPA buffer (Santa Cruz). Then 1 μM of folded BS -PreQ aptamer (34 nucleotide length) was added into the 5μg of cell lysates and photocrosslinked to Bio-**11** as described above. Then equal volumes of the samples were subjected to 17% denaturing PAGE in the presence of 7 M urea. Meanwhile, pre-cut nylon membrane, whatman filter paper and sponge were soaked in 0.5 X TBE buffer. Once the gel was run, the transfer sandwich was prepared, and the sandwich was placed into the BioRad Trans-Blot Cell and filled with pre-chilled 0.5X TBE. The Trans-Blot Cell was run at 80V for 60 minutes in cold room with continuous stirring. Then the membrane was UV crosslinked at 1200 mJ for 3 minutes in UVP crosslinker. Then the membrane was transferred into the blocking buffer (ThermoFisher TM pierce starting blocking buffer at room temperature for 1 hour in the Shaker. Finally, the membrane was incubated with HRP-conjugated Streptavidin (ThermoFisher N100) in the 1: 2000 dilution at room temperature for 2 hours and the biotinylated probe conjugated band was visualized by Western Blotting Luminol Reagent (Cell signaling Technology #7003) in the ChemiDoc imager.

RT-qPCR

Samples for these experiments were prepared using the same protocol as described above for the RNA-sequencing experiments. However, in this case extra samples were prepared with the free PreQ1 ligand (25 μM and 100 μM concentrations) titrated into the solution containing compound 11 and the total RNA to allow for competition before UV crosslinking. RNA concentration was determined by using a QuBit RNA HS Assay Kit, then 200 ng of RNA was used for cDNA synthesis with qScript cDNA SuperMix (Quanta Biosciences) according to the manufacture's protocol. The 20 μl reactions were incubated in a thermocycler (Bio-rad) for 5 minutes at 22 °C, 30 minutes at 42 °C, 5 minutes at 85 °C, then held at 4 °C. Then 20 ng of cDNAs was subjected to qPCR using a Perfecta SYBR Green Super Mix (Quanta Biosciences) on an Eppendorf Mastercycler RealPlex2 in the presence of appropriate set of primers. The reactions were incubated at 95 °C for 10 minutes, followed by 40 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 20 seconds. Threshold cycles (C_T) is defined as the first cycle that showed a detectable increase in fluorescence due to the formation of PCR products and was used to determine the template amount in each sample. The relative fold change in expression was measured using the Livak method[11] and were normalized

to the relative level of RNAs of the control experiment (compound **17**). For example, to calculate the $\Delta\Delta(C_T)$ between each gene of interest and the average of the control samples: $\Delta(C_T) = C_T$ (target gene) – C_T(Gapdh); $\Delta\Delta(C_T) = \Delta C_T$ (treatment or knockdown) – ΔC_T (control); fold change = $2^{-\Delta\Delta (C_T)}$. The primers used in these experiments are shown in the Supplementary Information Table S2.

Table S2: The sequence information of the primers used in the qPCR.

Gene name		Sequence (5'-3')
RF00024 (ENSG00000277925.1)	Forward Primers	AAC CCT AAC TGA GAA GGG CG
	Reverse Primers	AGA ATG AAC GGT GGA AGG CG
HIST2H2BF (ENSG00000203814.6)	Forward Primers	TTC GCG CAA AAA TGC CG
	Reverse Primers	CTT CAG CAC CTT GTA CAC GTA A
HIST1H3F (ENSG00000277775.1)	Forward Primers	CAG CTC GTA AGT CCA CTG GC\
	Reverse Primers	CGA TTT CTG ATA GCG GCG GA
GAPDH (ENSG00000111640.14)	Forward Primers	AGG TCG GTG TGA ACG GAT TTG
	Reverse Primers	GGG GTC GTT GAT GGC AAC A

References:

1. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A. Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem Rev* (2021).

2. Connelly, C.M. et al. Synthetic ligands for PreQ1 riboswitches provide structural and mechanistic insights into targeting RNA tertiary structure. *Nat Commun* **10**, 1501 (2019).

3. Hirata, K. et al. ZOO: an automatic data-collection system for high-throughput structure analysis in protein microcrystallography. *Acta Crystallogr D* **75**, 138-150 (2019).

4. Yamashita, K., Hirata, K. & Yamamoto, M. KAMO: towards automated data processing for microcrystals. *Acta Crystallogr D* **74**, 441-449 (2018).

5. Kabsch, W. XDS. *Acta Crystallogr. D.* **66**, 125-132 (2010).

6. McCoy, A. J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).

7. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular graphics. *Acta Crystallogr D* **66**, 213-221 (2010).

8. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126-2132 (2004).

9. Langmead, B., Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).

10. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DeSeq2. *Genome Biology* **15**, 550 (2014).

11. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402-408 (2001).

Supporting Figures:

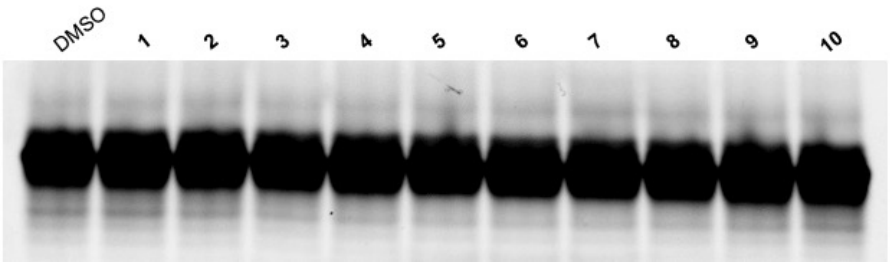


Figure S1: The PAGE analysis of electrophilic probe reactive towards to PreQ1 aptamer. No evidence for higher molecular weight covalent adducts was observed.

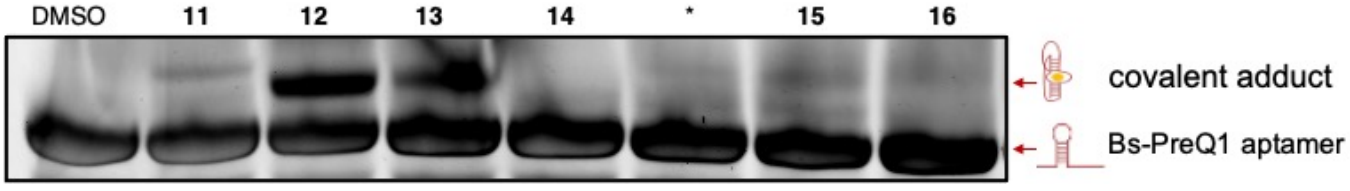


Figure S2: The PAGE analysis of photoaffinity probes reactivity towards to preQ1 aptamer. The unstabilized probes (**11**, **12**, and **13**) and stabilized probes (**15** and **16**) expect **14** showed formation the higher molecular weight covalent adduct.

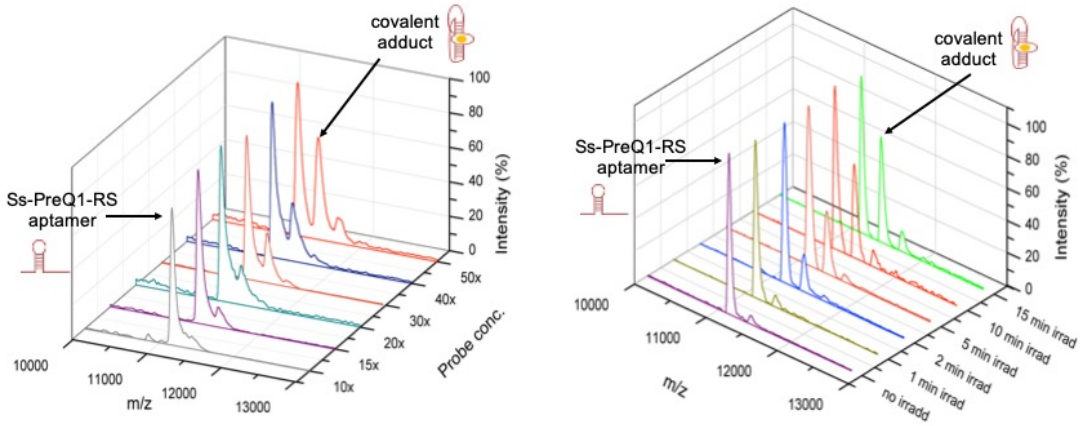


Figure S3: Orbitrap-LC/MS based characterization of dose and time dependent crosslinking efficiency of compound **11** towards PreQ1 aptamer.

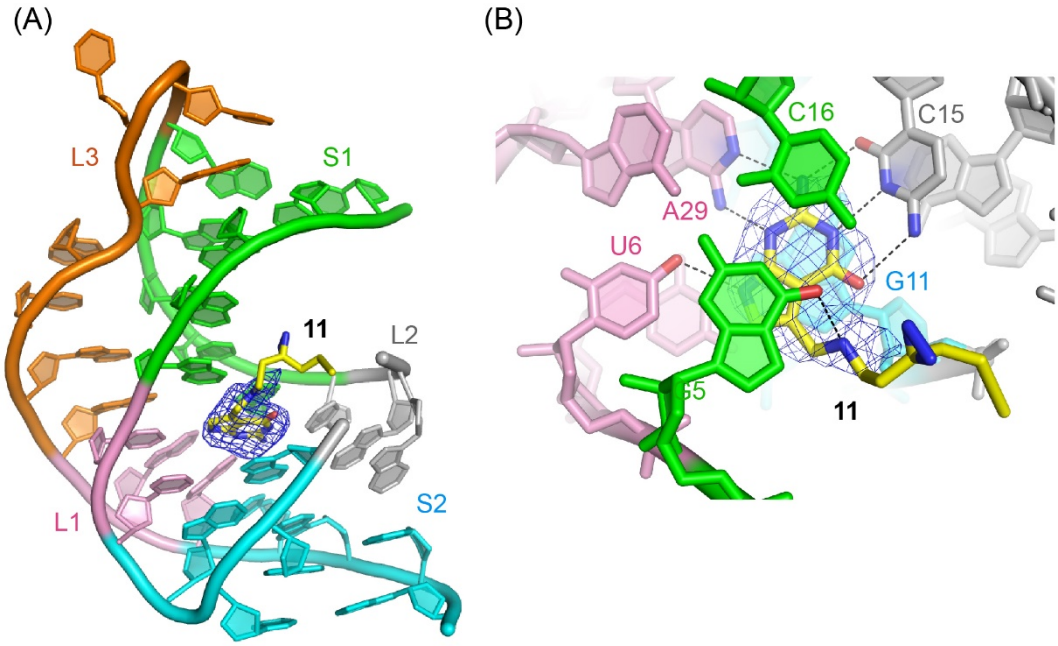


Figure S4: Crystal structures of the wild-type *Tt* PreQ1 riboswitch aptamer complexed with **11**. **(A)** Overall structure of the complex. **(B)** Close up view of the binding site of **11**. S1, S2, L1, L2, and L3 are colored green, cyan, pink, gray, and orange, respectively. The nucleotides between L3 and S2, which interact with L1, are also in pink. The $|F_o| - |F_c|$ electron density map for **11** is colored blue and contoured at 3.0 σ. The nucleotides interact with the compound are labeled, and hydrogen bonds are indicated as dotted lines.

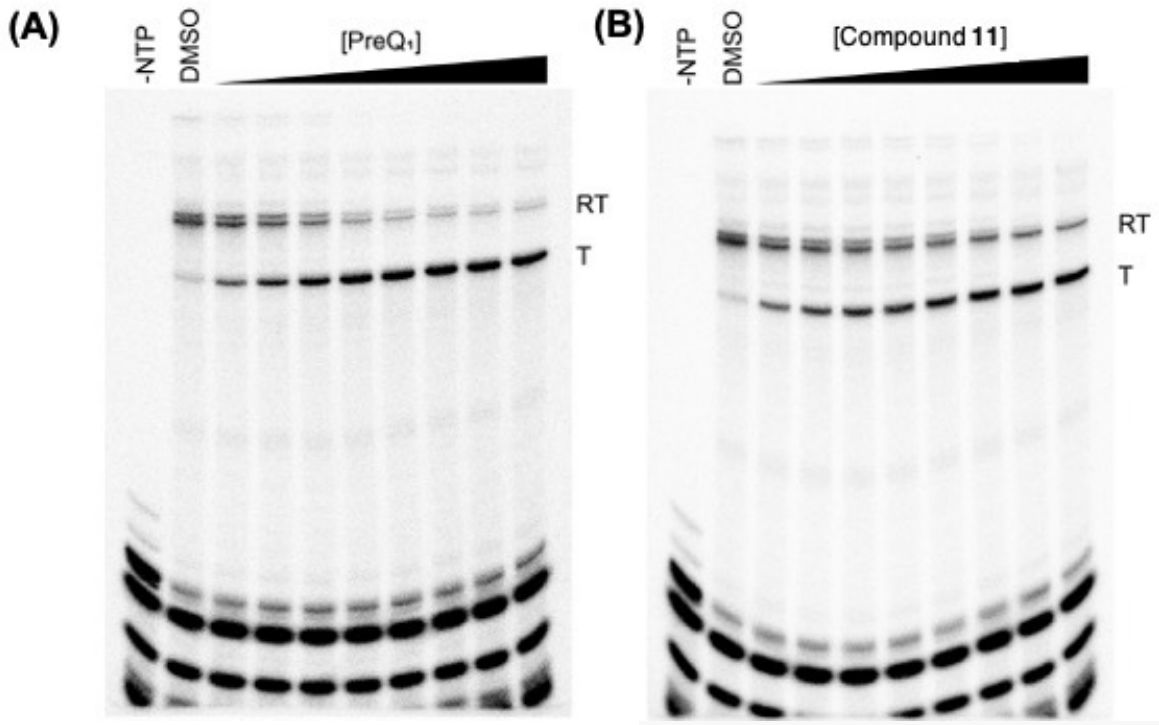


Figure S5: Representative PAGE gel images of the 32P-labeled RNA products of in vitro transcription of the *Ss* PreQ1-RS template in the presence of increasing concentrations of (A) PreQ1 and (B) **11**. Bands corresponding to the read-through transcription product (RT) and terminated transcription product (T) are indicated.

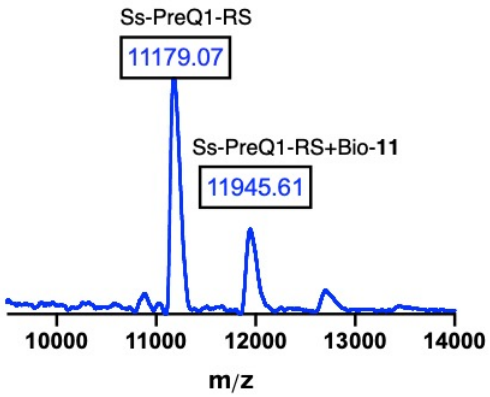


Figure S6: MALDI-TOF analysis of reactivity of Bio-**11** towards the PreQ1 aptamer. The higher molecular weight species confirmed the formation covalent adduct between the Bio-**11** and PreQ1 aptamer.

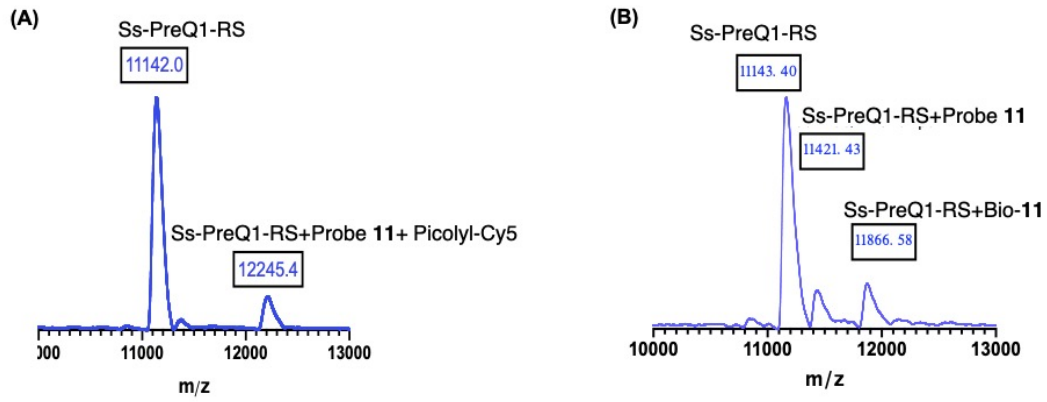


Figure S7: MALDI-TOF analysis of reactivity of **(A)** Cy5-picolyl azide probe **(B)** biotin-PEG-azide towards the PreQ1 aptamer. The higher molecular weight species confirmed the formation covalent adduct between the probe and PreQ1 aptamer.

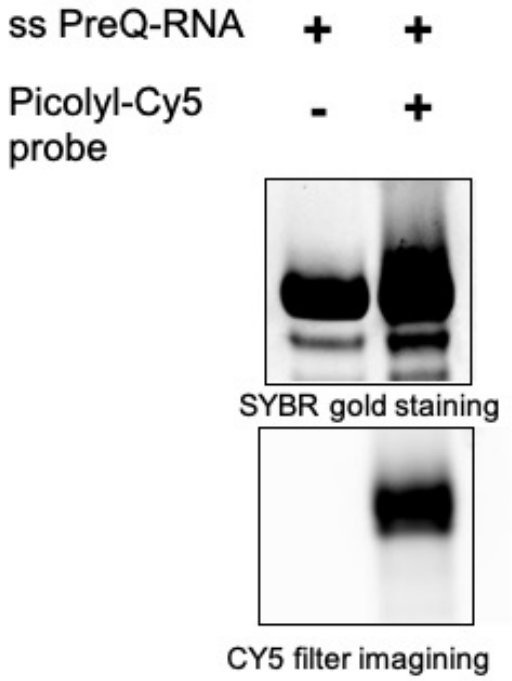


Figure S8: The PAGE characterization of CY5-picolyl probe cross-linked PreQ1 aptamer by SYBR gold staining (top) and CY5 filter imaging (bottom).

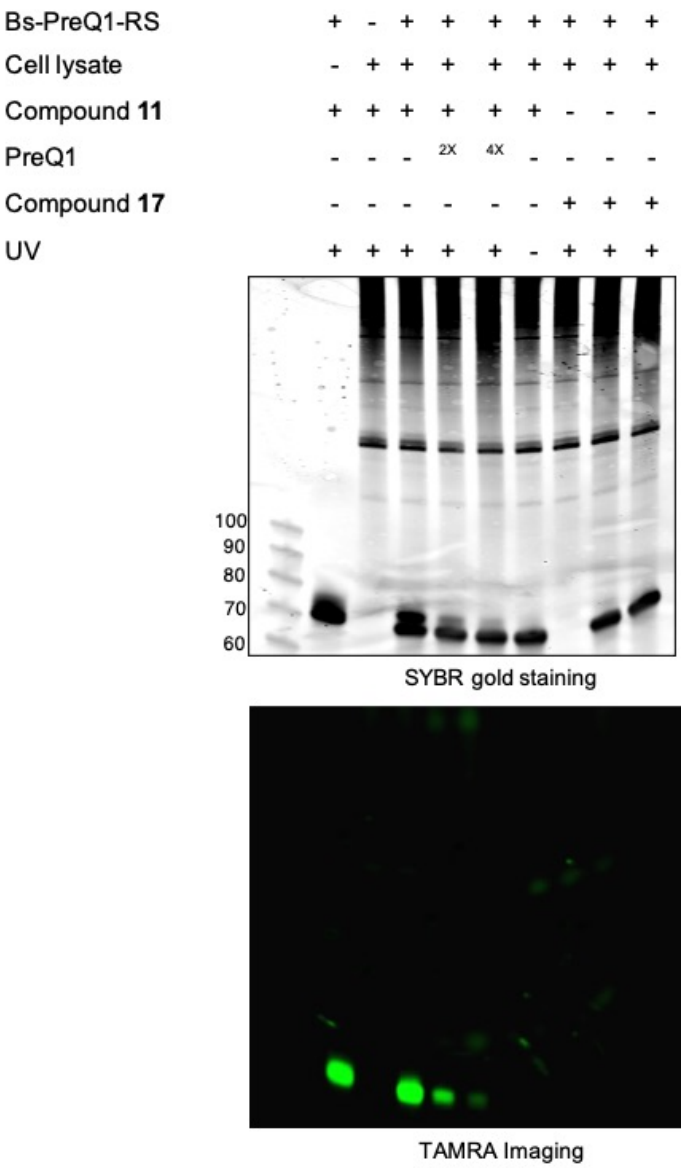
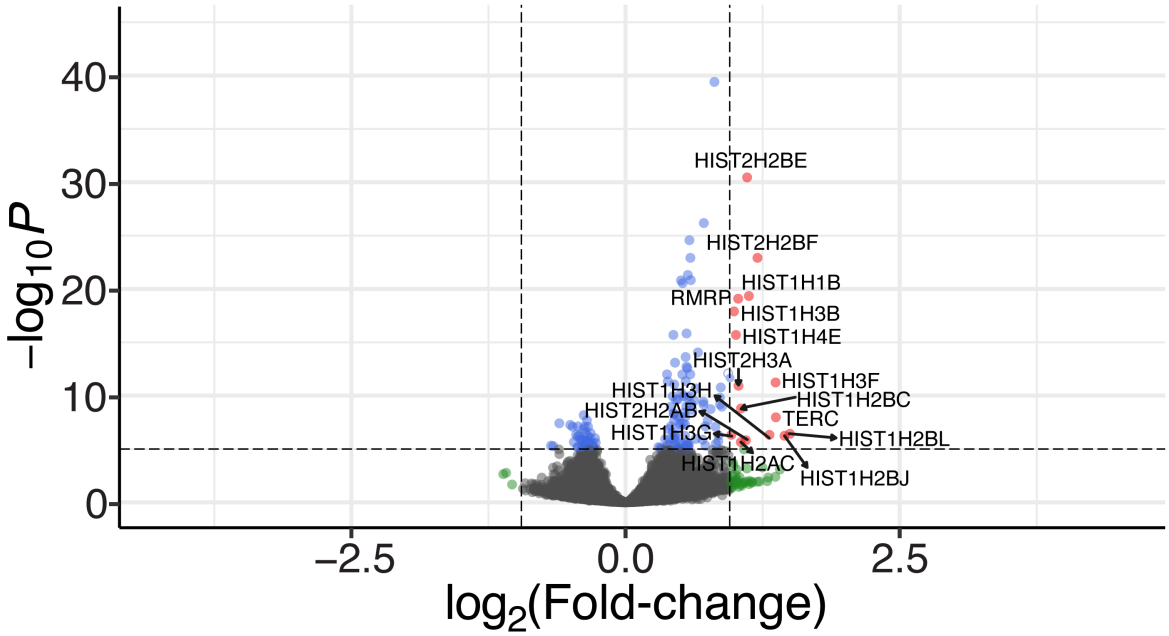


Figure S9: Competitive photocrosslinking experiments in MCF-7 total RNA imaged by SYBR gold stain (bottom) and TAMRA labeling (top). Compound **11** selectively label the PreQ1 aptamer in the presence of cellular RNA while no detectable crosslinked product was observed with compound **17**.



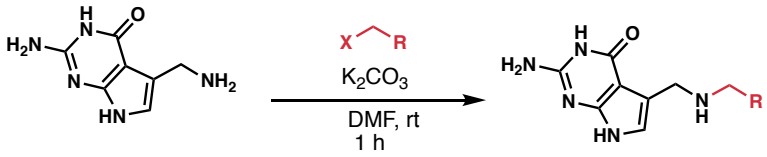
Gene Symbol	$\log_2$(Fold-change)	$-\log_{10}P$
HIST2H2BE	1.11	3.5e-31
HIST2H2BF	1.20	1.2e-23
HIST1H1B	1.12	4.6e-20
HIST1H3B	1.13	4.6e-20
RMRP	0.98	1.3e-18
HIST1H4E	1.01	2.0e-16
HIST2H3A	1.03	1.2e-11
HIST1H3F	1.37	5.9e-12
HIST1H2BC	1.05	1.6e-9
TERC	1.37	1.1e-8
HIST1H3H	1.31	4.7e-7
HIST1H3G	1.10	1.5e-6
HIST1H2BL	1.50	3.8e-7
HIST1H2BJ	1.45	5.8e-7
HIST1H2AC	1.05	2.3e-6
HIST2H2AB	1.40	8.4e-4

Figure S10: Volcano Plot showing differential expression analysis (DeSeq2) between **11** and **17** treated samples. The points in red are significantly enriched ($\log_2$(Fold-

change) > 0.95, -log$_{10}$P > 4.). These hits are labelled with the gene name on the plot. Corresponding values for enrichment and significance are provided in the accompanying table.

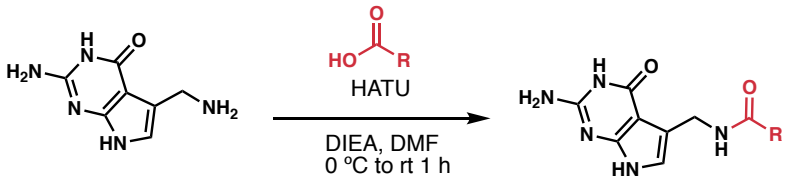
Figure S11: Synthesis and characterization of electrophoretic and photoaffinity probes

A) Substitution reaction between PreQ1 and alkyl halides



To a solution of PreQ1•2TFA (20 mg, 49 μmol, 1.0 equiv.) in DMF, K_2CO_3 (21 mg, 15 μmol, 3.1 equiv.) and alkyl halide (49 μmol, 1.0 equiv.) were added and the reaction mixture was stirred at room temperature for 1 h. Then, the reaction mixture was concentrated, and purified via column chromatography on a silica gel using gradient based MeOH:DCM (0-20 % MeOH) on CombiFlash® Rf system.

B) Coupling reaction between PreQ1 and carboxylic acid-based substrates

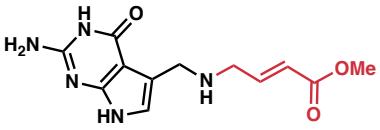


To a solution of a carboxylic acid substrate (37 μmol, 1.0 equiv.) in DMF (1 mL) cooled to 0 ºC, HATU (14 mg, 37 μmol, 1.0 equiv.) and diisopropylethylamine (20 μL, 118 μmol, 3.2 equiv.) were added and stirred for 10 minutes. Then, PreQ1•2TFA (15 mg, 37 μmol, 1.0 equiv.) was added, warmed the reaction mixture to room temperature and stirred for 1 hour. Then, the reaction mixture was concentrated and purified via column chromatography on a silica gel eluting with gradient based MeOH:DCM on CombiFlash® Rf system or preparative HPLC eluting with 10-90 % gradient of H_2O:MeCN containing 0.1% TFA.

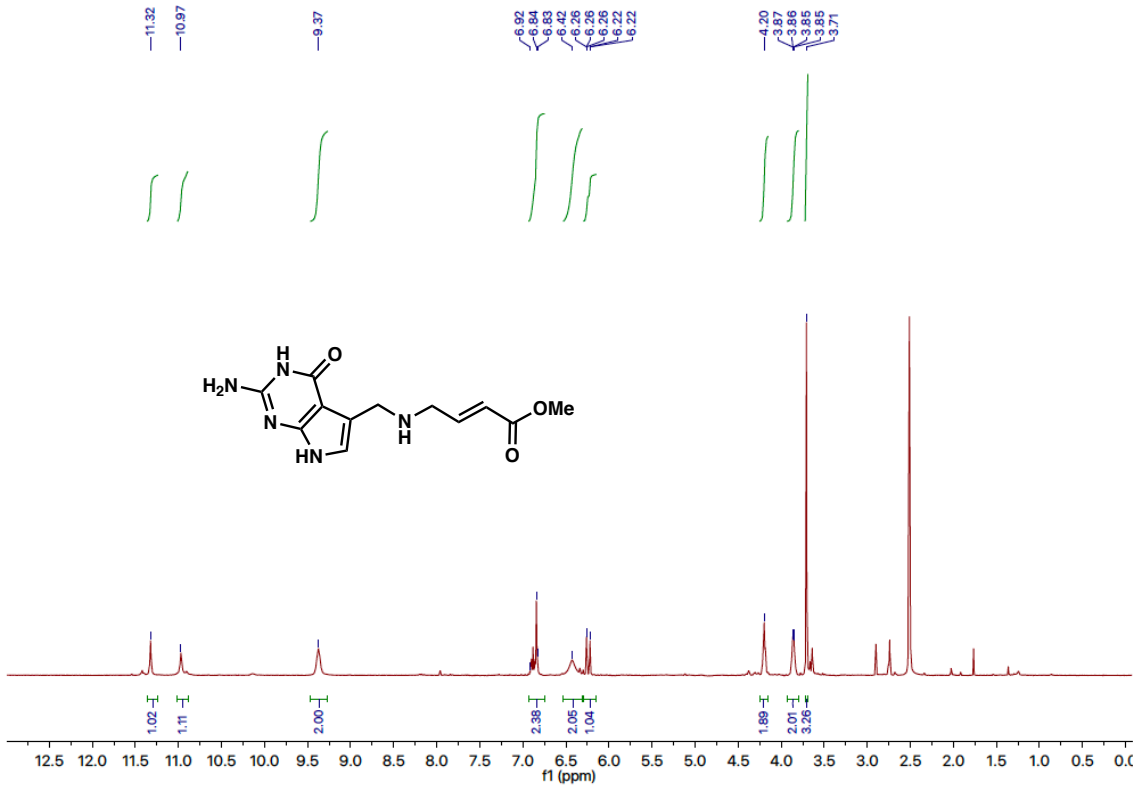
Mass Spectral Analysis

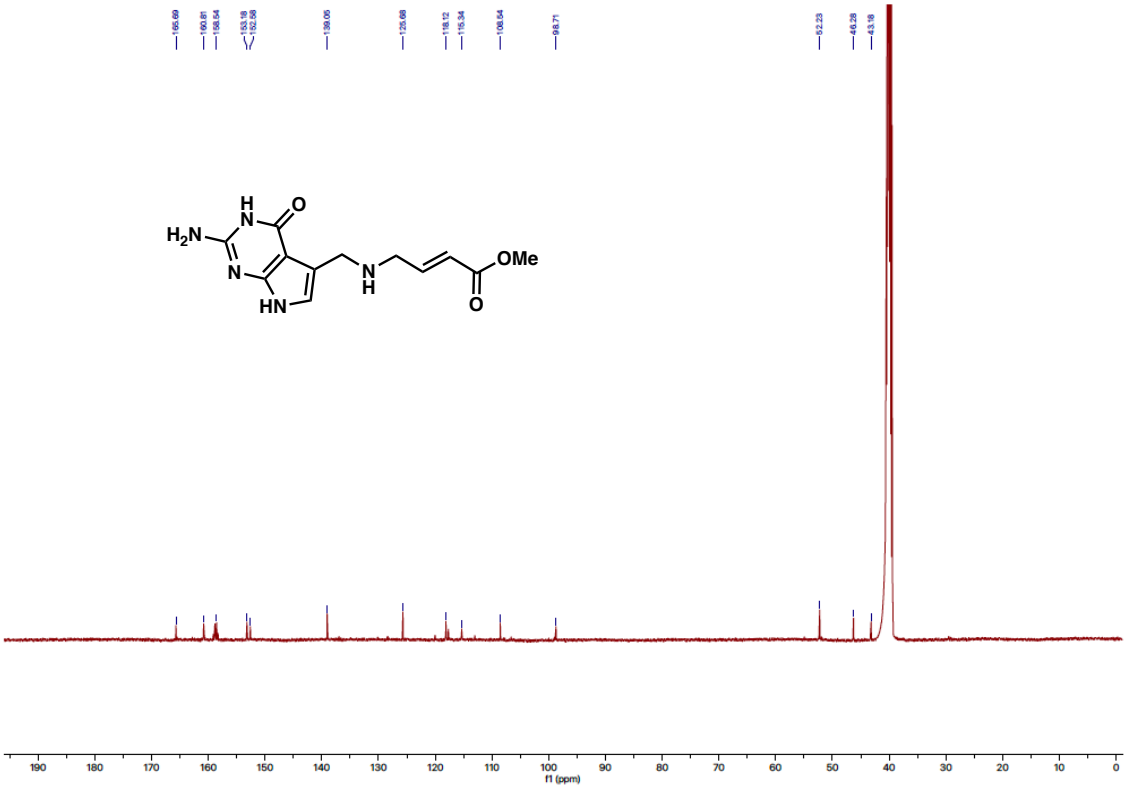
LC/MS analysis of accurate mass measurements was carried out on the Orbitrap LTQ-XL system. A 2.0-μL aliquot of an analytical solution (about 58 μg/ml in LC/MS-grade 1:1 CH_3CN/H_2O) was injected onto a 2.1 X 100 mm, 3.5-μm Zorbax SB-C18 narrow-bore HPLC column and eluted at flow rate of 250 μL/minutes with a 15-minutes combination of both isocratic and linear gradient elution using mobile phase varying from 2-90% CH_3CN/H_2O and containing 0.1% formic acid.

Compound 1 (Methyl (*E*)-4-(((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)amino)but-2-enoate).

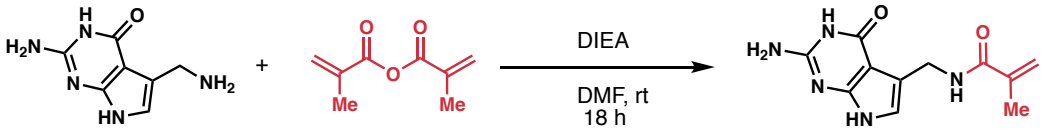


Prepared according to the general procedure **A** isolated as TFA salt (5.0 mg, 35% yield) . White solid. ^{1}H NMR (500 MHz, DMSO): δ 11.32 (s, 1H), 10.97 (s, 1H), 9.37 (brs, 2H), 6.92-6.83 (m, 1H), 6.84 (s, 1H), 6.42 (brs, 2H), 6.24 (d, *J* = 15.4 Hz, 1H), 4.20 (t, *J* = 5.2 Hz, 2H), 3.86 (m, 2H), 3.71 (s, 1H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 165.69, 160.81, 158.54 (q, *J* = 35.4 Hz), 153.18, 152.58, 139.05, 125.68, 118.12, 115.4 (q, *J* = 294.0 Hz), 108.54, 98.71, 52.23, 46.28, 43.18; HRMS m/z: calcd. for $C_{12}H_{16}N_5O_3$ $[M+H]^+$ 278.1248, found 278.1240.

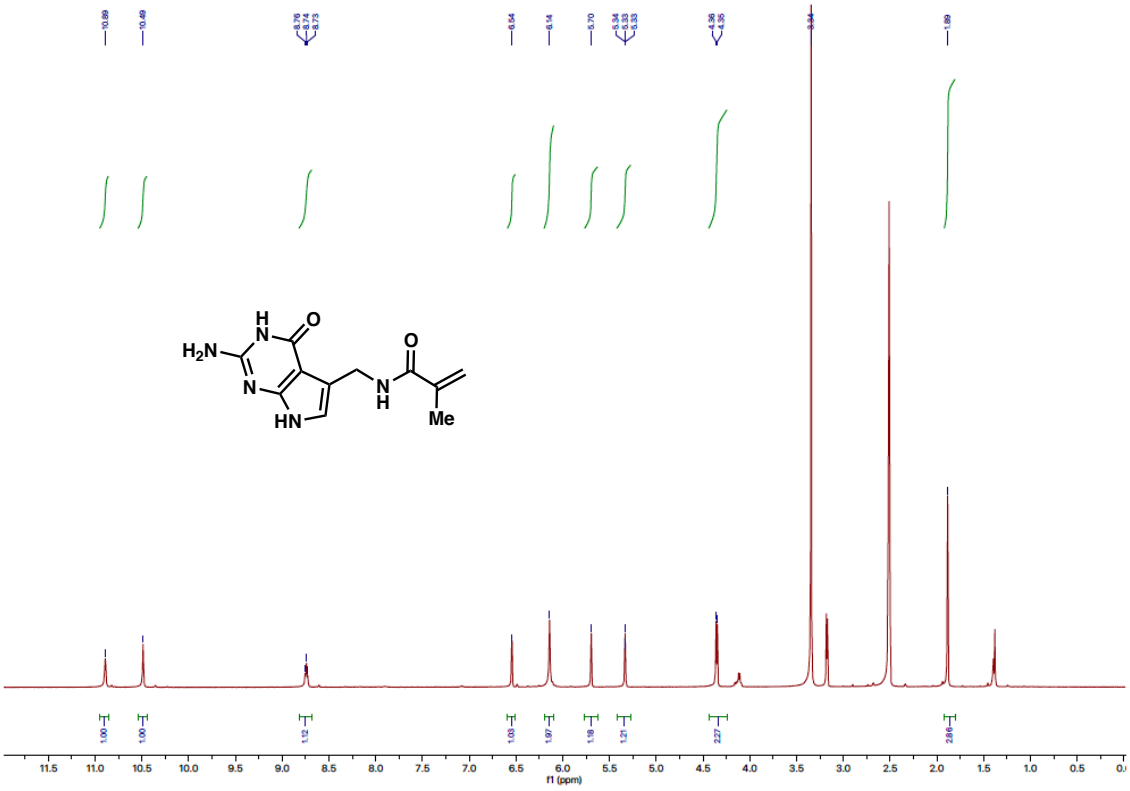


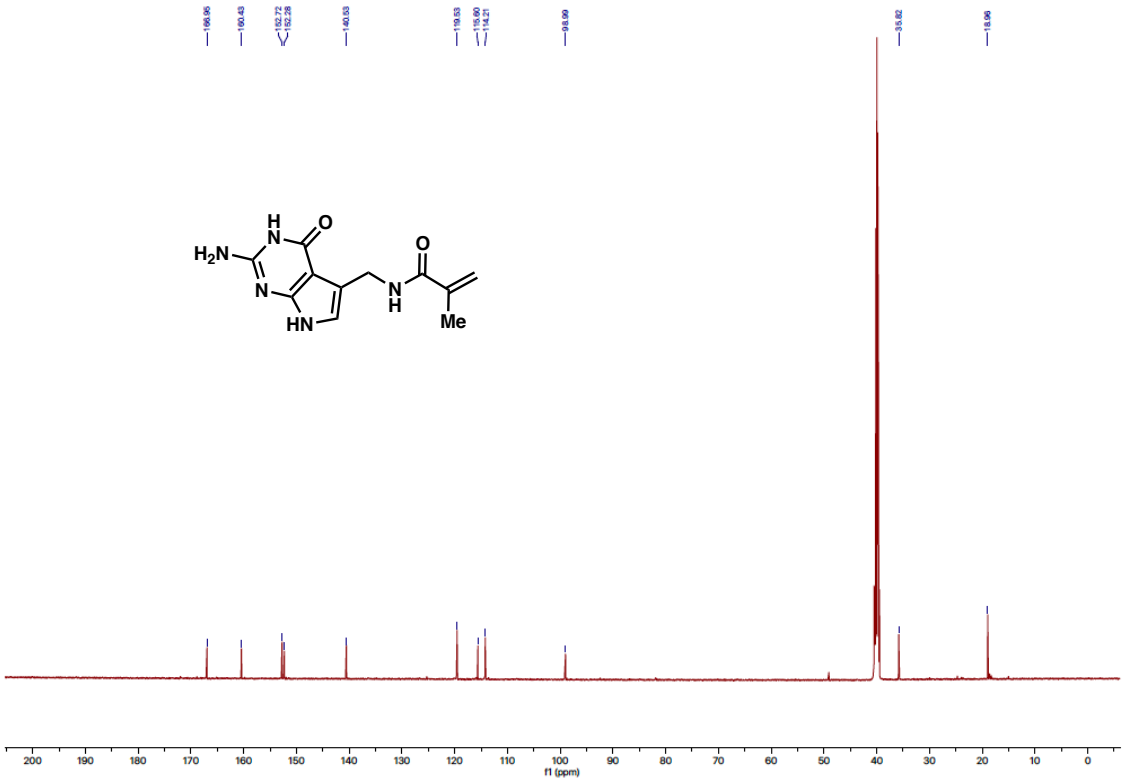


Compound 2 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)methacrylamide).

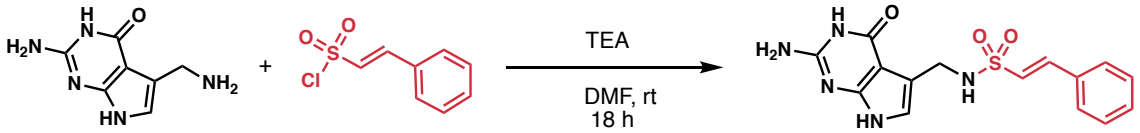


To a solution of PreQ1•2HCl (20 mg, 80 μmol, 1.0 equiv.) and DIEA (30 μL, 160 μmol, 2.0 equiv.) in DMF (1.0 mL), methacrylic anhydride (15 μL, 96 μmol, 1.2 equiv.) was added and the reaction mixture was stirred at room temperature for 18 h. Then, the reaction mixture was concentrated and via column chromatography on a silica gel using MeOH:DCM. White solid. ^{1}H NMR (500 MHz, DMSO): δ 10.89 (s, 1H), 10.49 (s, 1H), 8.74 (t, J = 5.4 Hz, 1H) , 6.54 (s, 1H), 6.14 (s, 2H), 5.70 (s, 1H), 5.33 (t, J = 1.6 Hz, 1H), 4.35 (d, J = 5.4 Hz, 2H),1.89 (s, 3H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 166.9, 160.4, 152.7, 152.3, 140.5, 119.5, 115.6, 114.2, 99.0, 35.8, 19.0; HRMS m/z: calcd. for $C_{11}H_{14}N_5O_2$ $[M+H]^+$ 248.1142, found 248.1137.

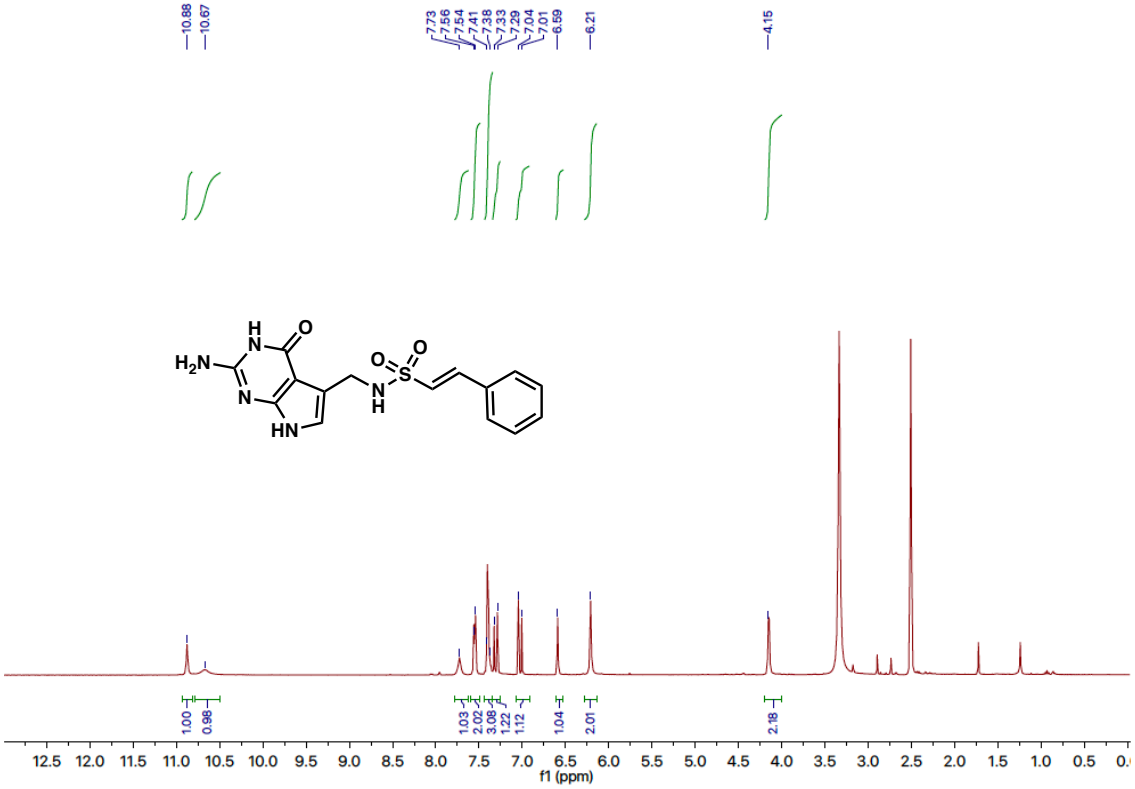


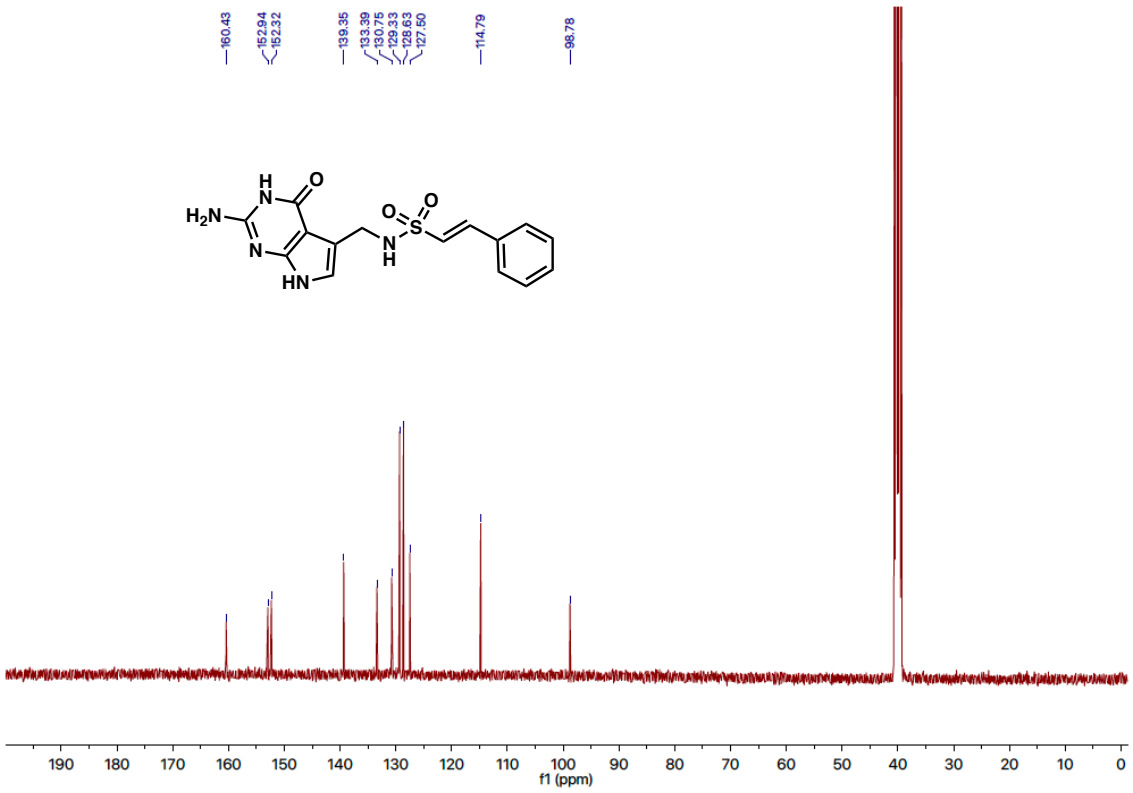


Compound 3 ((*E*)-*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-2-phenylethene-1-sulfonamide).

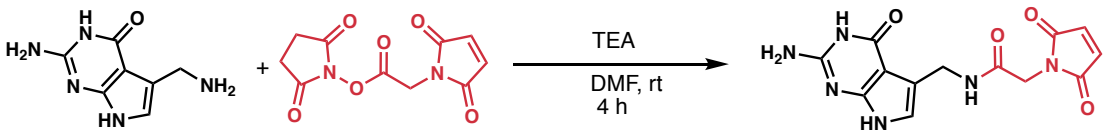


The reaction is run according the procedure for compound **6**. White solid. ^{1}H NMR (500 MHz, DMSO): δ 10.88 (s, 1H), 10.67 (brs, 1H), 7.73 (brs, 1H), 7.56-7.54 (m, 2H), 7.41-7.38 (m, 3H), 7.30 (d, J = 15.6 Hz, 1H) , 7.03 (d, J = 15.6 Hz, 1H) , 6.59 (s, 1H), 6.21 (s, 2H), 4.15 (d, J = 3.8 Hz, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 160.4, 152.9, 152.3, 139.4, 133.4, 130.8, 129.3, 128.6, 127.5, 114.8, 98.8; HRMS m/z: calcd. for $C_{15}H_{16}N_5O_3S$ $[M+H]^+$ 346.0968, found 346.0955.

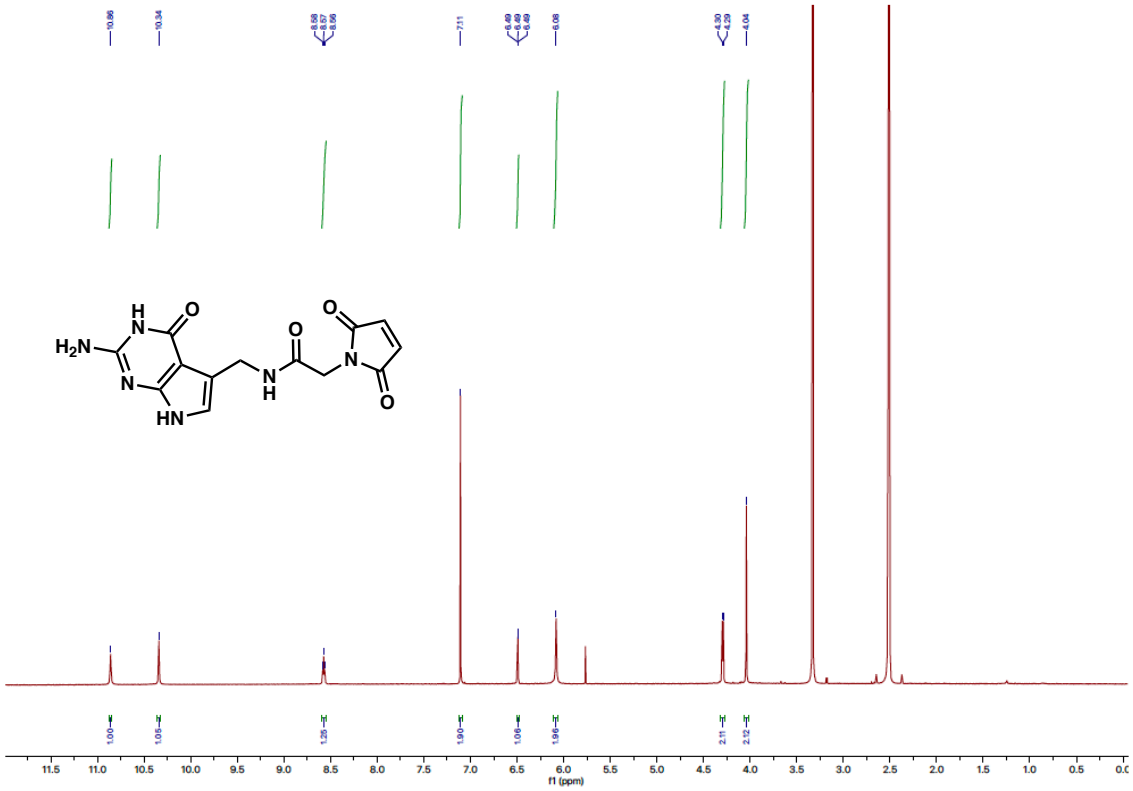


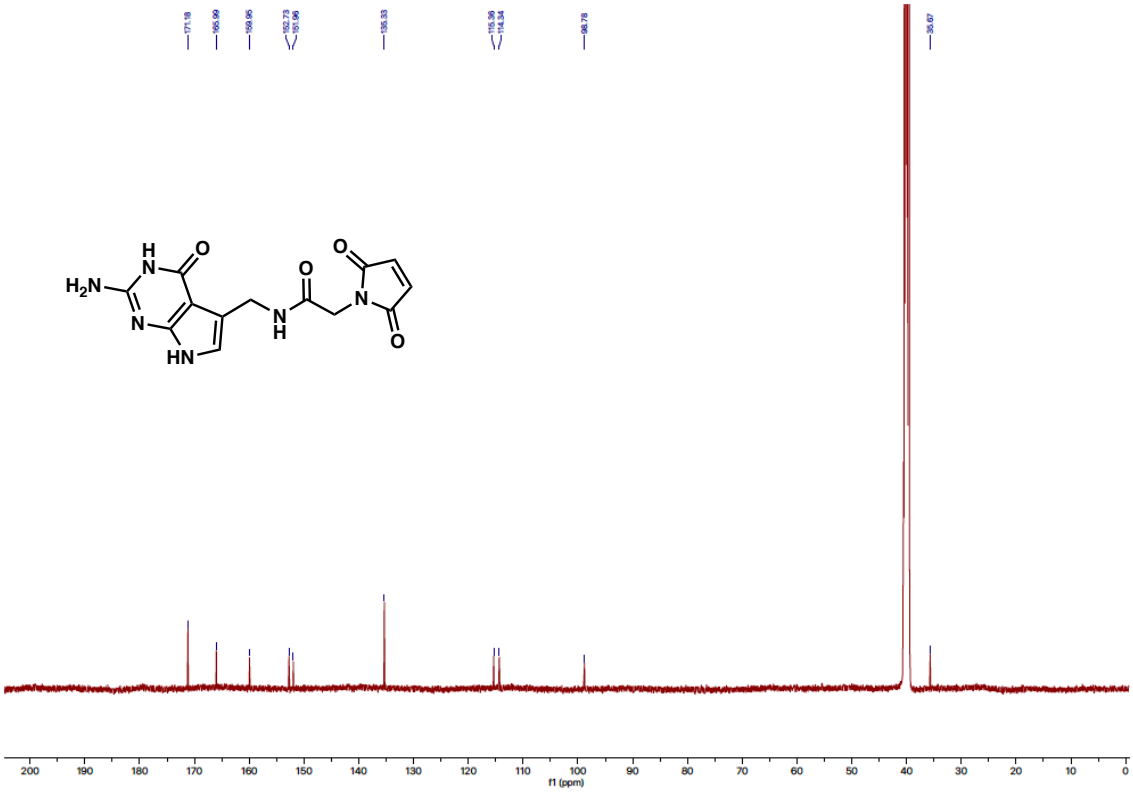


Compound 4 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-2-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)acetamide).

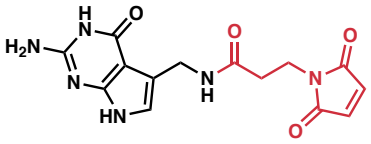


To a solution of PreQ1•2TFA (20 mg, 49 μmol, 1.0 equiv.) and triethylamine (15 μL, in DMF (1.0 mL) at room temperature, maleimidoacetic acid *N*-hydroxysuccinimide ester (12 mg, 49 μmol, 1.0 equiv.) was added and stirred for 4 h. Then, the reaction mixture was concentrated, and purified via column chromatography using a silica gel and eluting with gradient based MeOH:DCM to obtain the titled compound as a white solid. ^{1}H NMR (500 MHz, DMSO): δ 10.86 (s, 1H), 10.34 (s, 1H), 8.57 (t, J = 5.4 Hz, 1H), 7.11 (s, 2H), 6.49 (s, 1H), 6.08 (s, 2H), 4.30 (d, J = 5.4 Hz, 2H) , 4.04 (s, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 171.18, 165.99, 159.95, 152.73, 151.96, 135.33, 115.36, 114.34, 98.78, 35.67; HRMS m/z: calcd. for $C_{13}H_{13}N_6O_4$ $[M+H]^+$ 317.0993, found 317.0982.

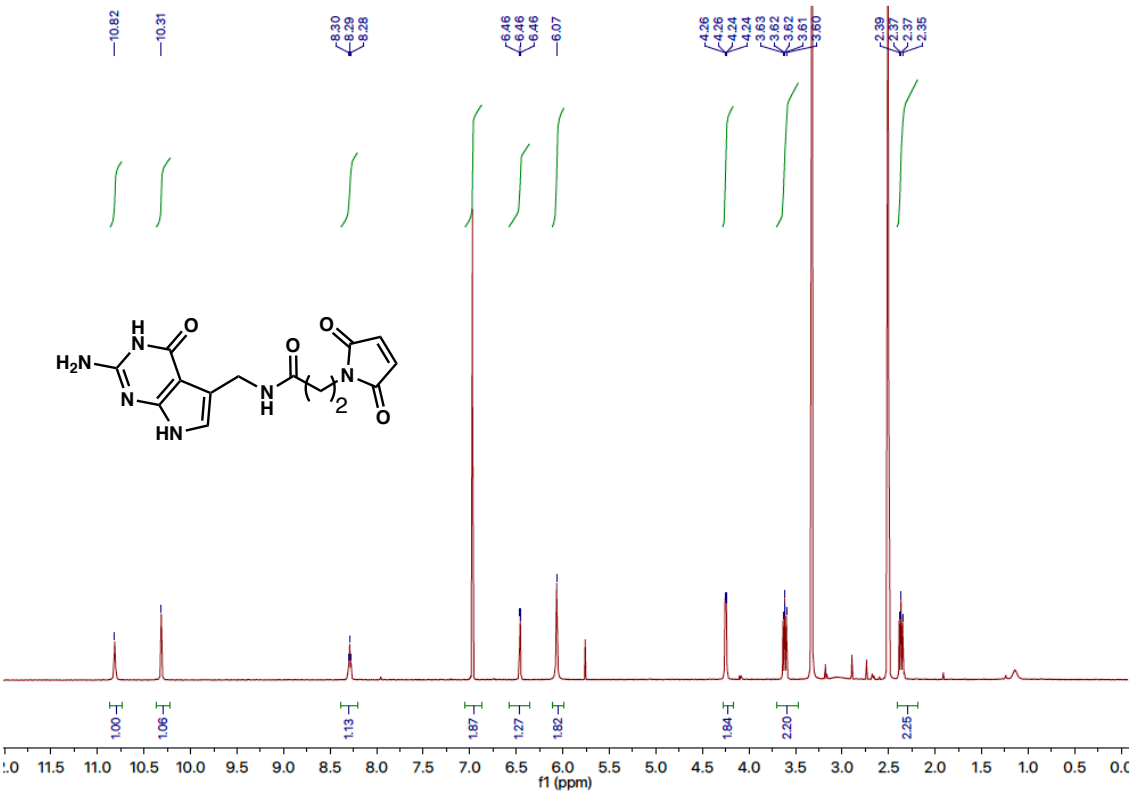

10.86
10.34
8.58
8.57
8.56
7.11
6.49
6.49
6.49
6.08
4.30
4.29
4.04
1.00
1.05
1.25
1.90
1.06
1.96
2.11
2.12
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)
H
N
O
H2N
N
HN
N
H
O
O
N
O

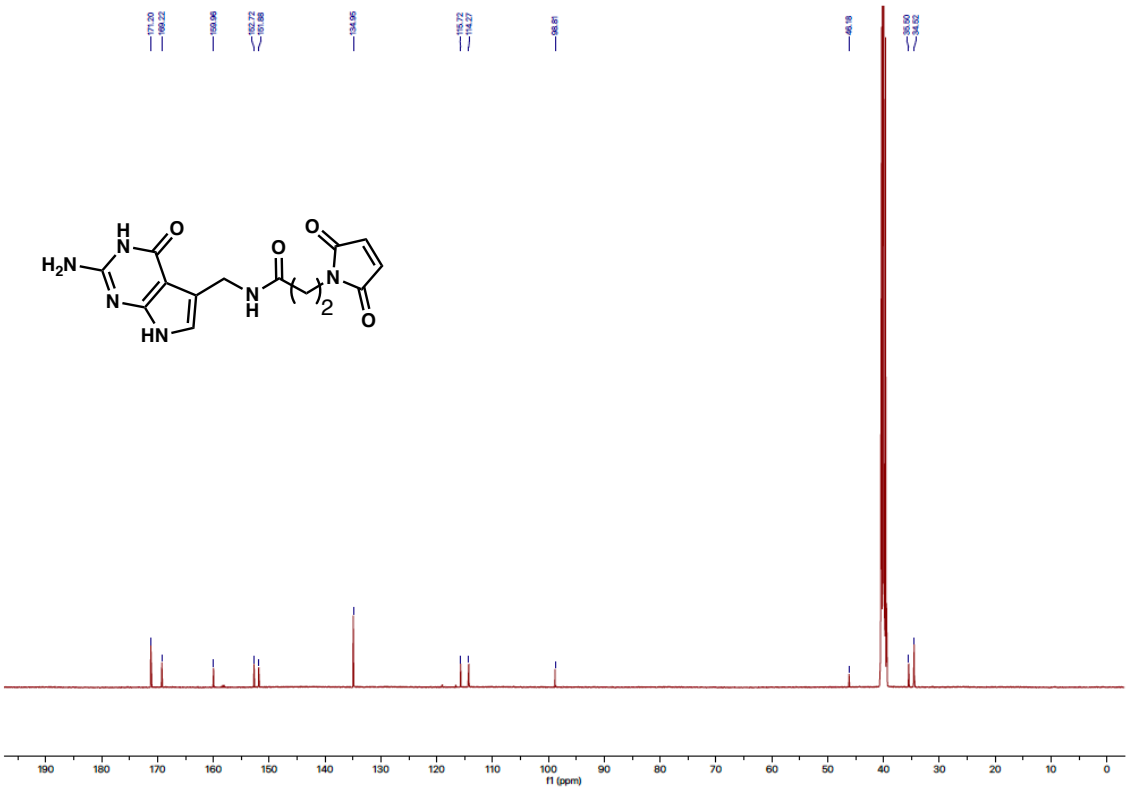

171.18
165.99
159.95
152.73
151.96
135.33
116.36
114.34
98.78
35.67
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)
H
N
O
H2N
N
HN
N
H
O
O
N
O

Compound 5 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-3-(2,5-dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)propenamide).

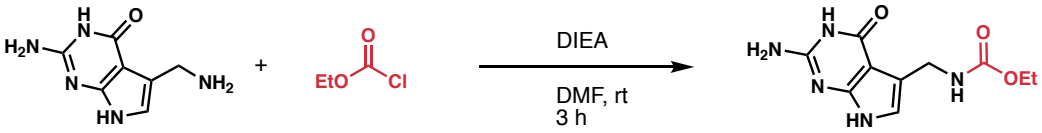


Prepared according to the procedure described for Compound **4**. White solid. ^{1}H NMR (500 MHz, DMSO): δ 10.82 (s, 1H), 10.31 (s, 1H), 8.29 (t, J = 5.4 Hz, 1H), 6.97 (s, 2H), 6.46 (brs, 1H), 6.07 (s, 2H), 4.24 (d, J = 5.1 Hz, 2H), 3.62 (m, 2H), 2.37 (m, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 171.2, 169.2, 160.0, 152.7, 151.9, 135.0, 115.7, 114.3, 98.8, 46.2, 35.5, 34.5; HRMS m/z: calcd. for $C_{14}H_{15}N_6O_4$ $[M+H]^+$ 331.1149, found 331.1138.

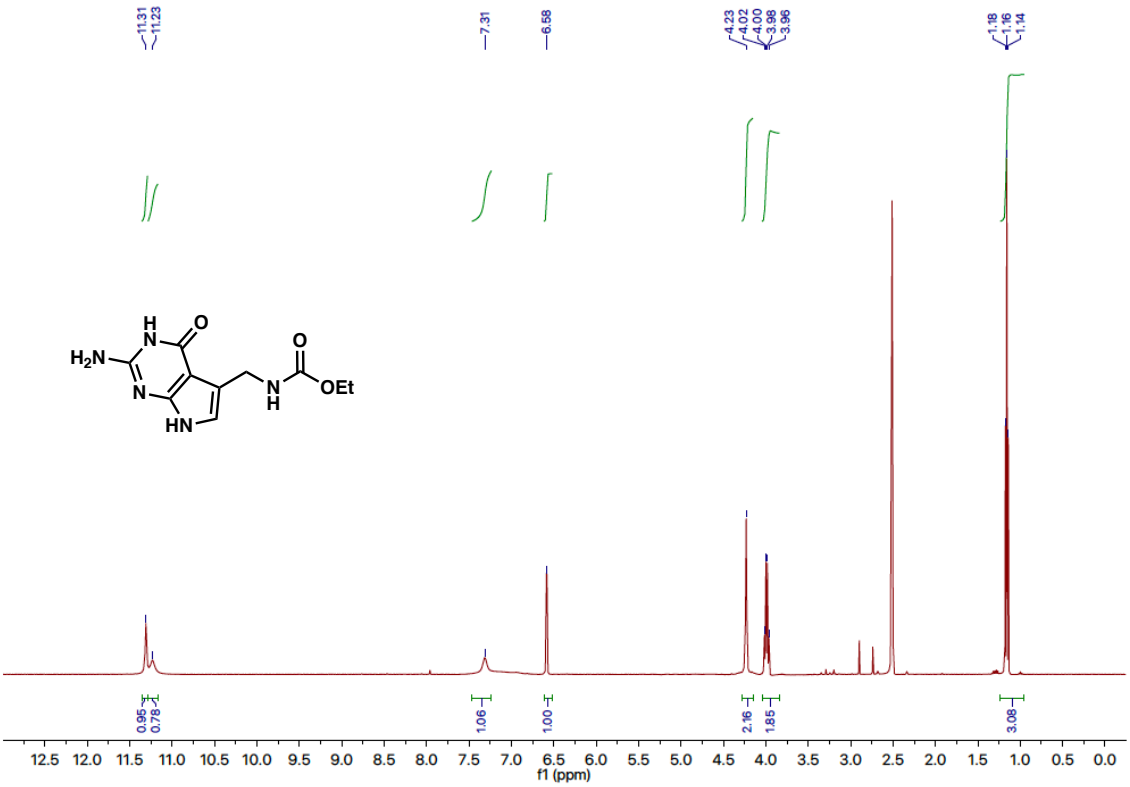


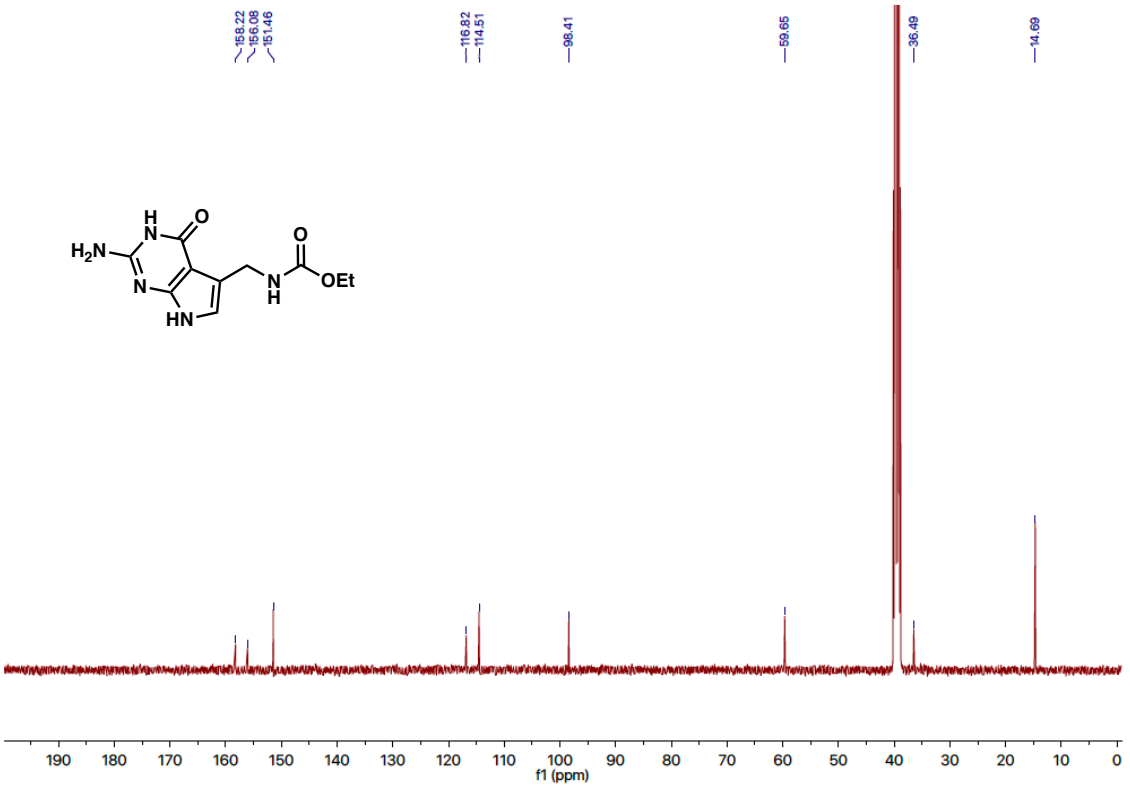


Compound 6 (Ethyl ((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)carbamate).

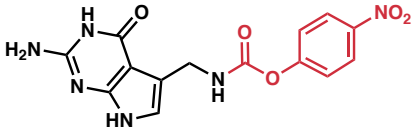


A solution of PreQ1•2HCl (20 mg, 80 μmol, 1.0 equiv.) in DMF (1.0 mL) and cooled to 0 °C was treated with diisopropylethylamine (27 μL, 160 μmol, 2.0 equiv.). Then, ethyl chloroformate (8 μL, 88 μmol, 1.1 equiv.) was added, the reaction mixture was allowed to warm to room temperature and stirred for 3 h. Consequently, the reaction mixture was concentrated, triturated with 20% MeOH/DCM, and the resulting precipitate was collected to obtain the titled compound as a white solid. ^{1}H NMR (500 MHz, DMSO): δ 11.31 (s, 1H), 11.23 (s, 1H), 7.31 (s, 1H), 6.58 (s, 1H), 4.23 (s, 2H), 3.99 (q, J = 7.1 Hz, 2H),1.16 (t, J = 7.1 Hz, 3H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 158.2, 156.1, 151.5, 116.8, 114.5, 98.4, 59.7, 36.5, 14.7; HRMS m/z: calcd. for $C_{10}H_{14}N_5O_3$ $[M+H]^+$ 252.1091, found 252.1084.

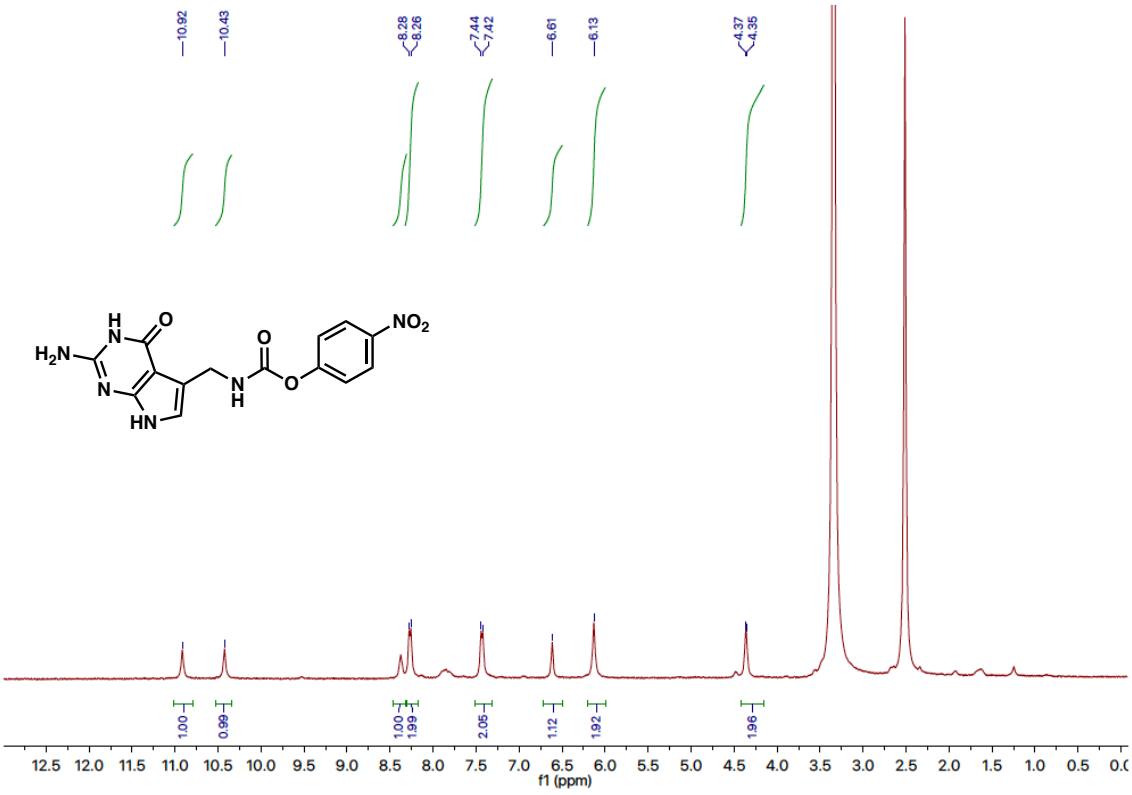

11.31
11.23
7.31
6.58
4.23
4.02
4.00
3.98
3.96
1.18
1.16
1.14
0.95
0.78
1.06
1.00
2.16
1.85
3.08
f1 (ppm)

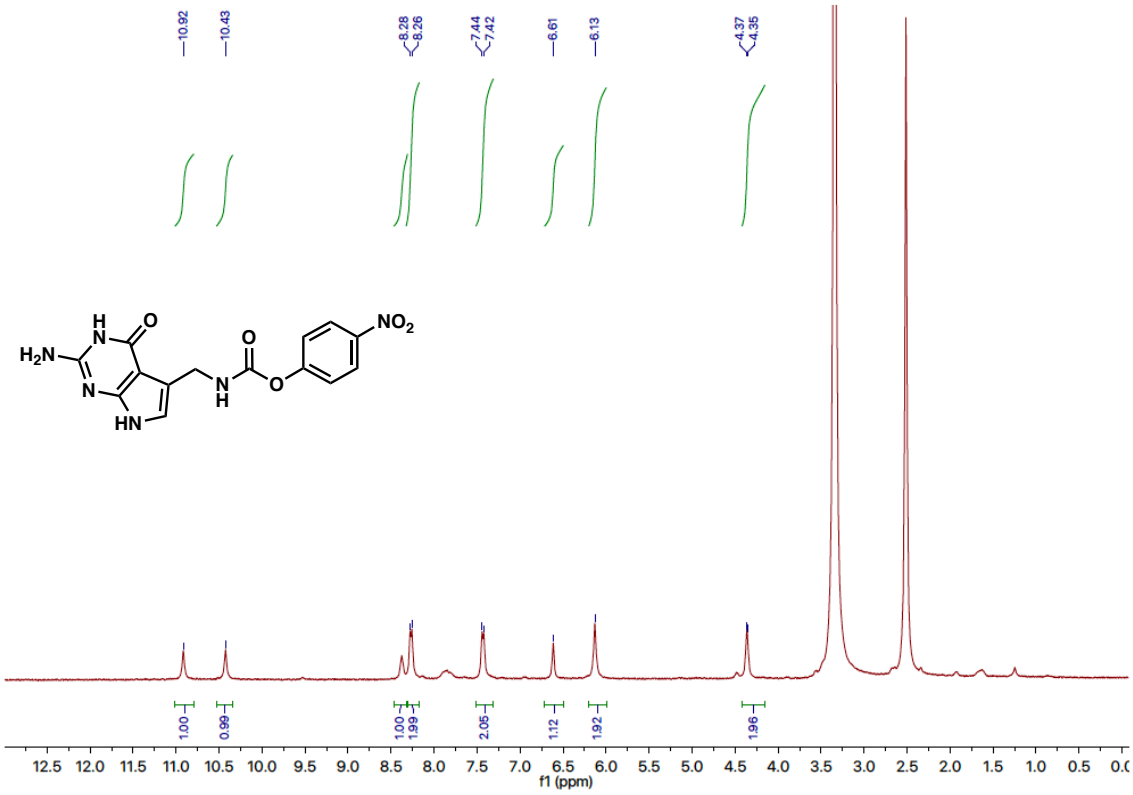

168.22
156.08
151.46
116.82
114.51
98.41
59.65
36.49
14.69
f1 (ppm)

Compound 7 (4-nitrophenyl ((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)carbamate).

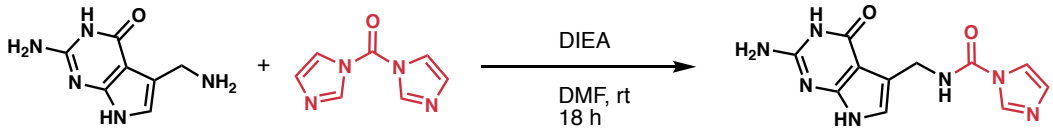


Prepared according the procedure for compound **9**. White solid. ^{1}H NMR (500 MHz, DMSO): δ 10.92 (s,1H), 10.43 (s, 1H), 8.37 (s, 1H), 8.27 (d, J = 8.2 Hz, 2H) , 7.43 (d, J = 8.2 Hz, 2H), 6.61 (s, 1H), 6.13 (s, 2H), 4.36 (brs, 2H); ^{13}C{^{1}H} NMR (125 MHz, DMSO): 160.10, 153.36, 152.81, 152.14, 144.58, 125.58, 125.58, 123.00, 115.12, 114.50, 98.80, 37.76; HRMS m/z: calcd. for $C_{14}H_{13}N_6O_5$ $[M+H]^+$ 345.0942, found 345.0933.

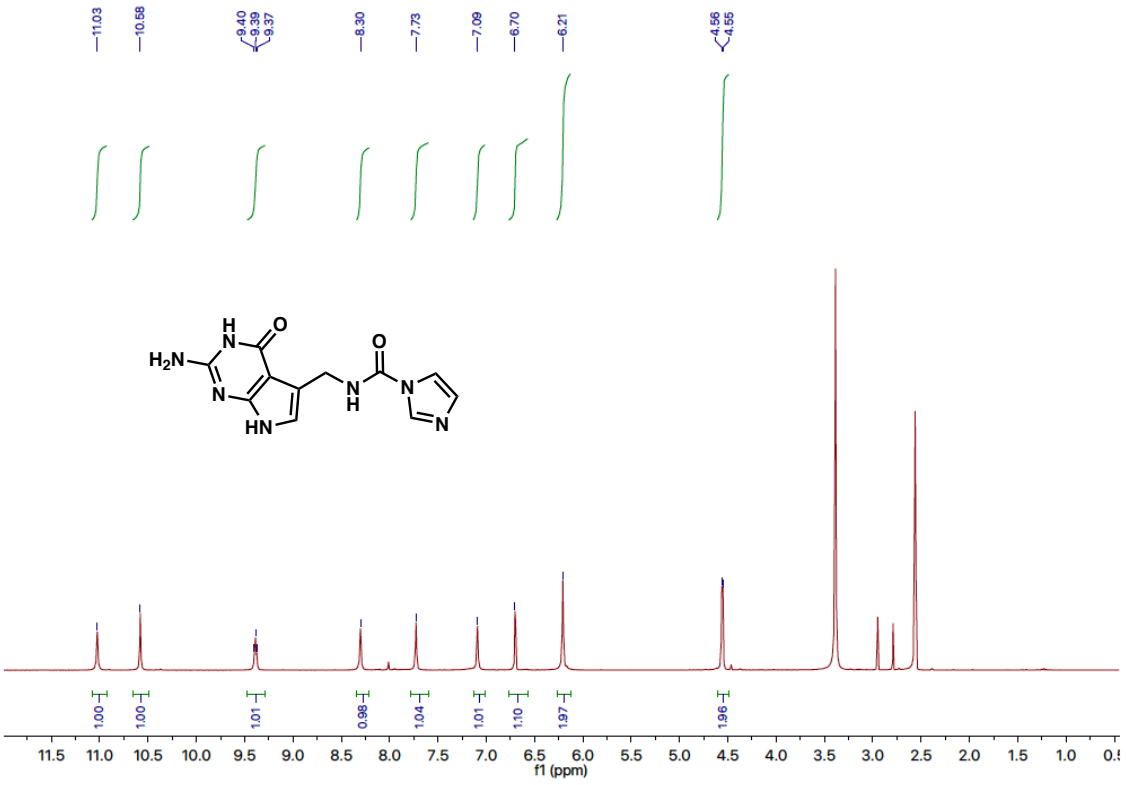


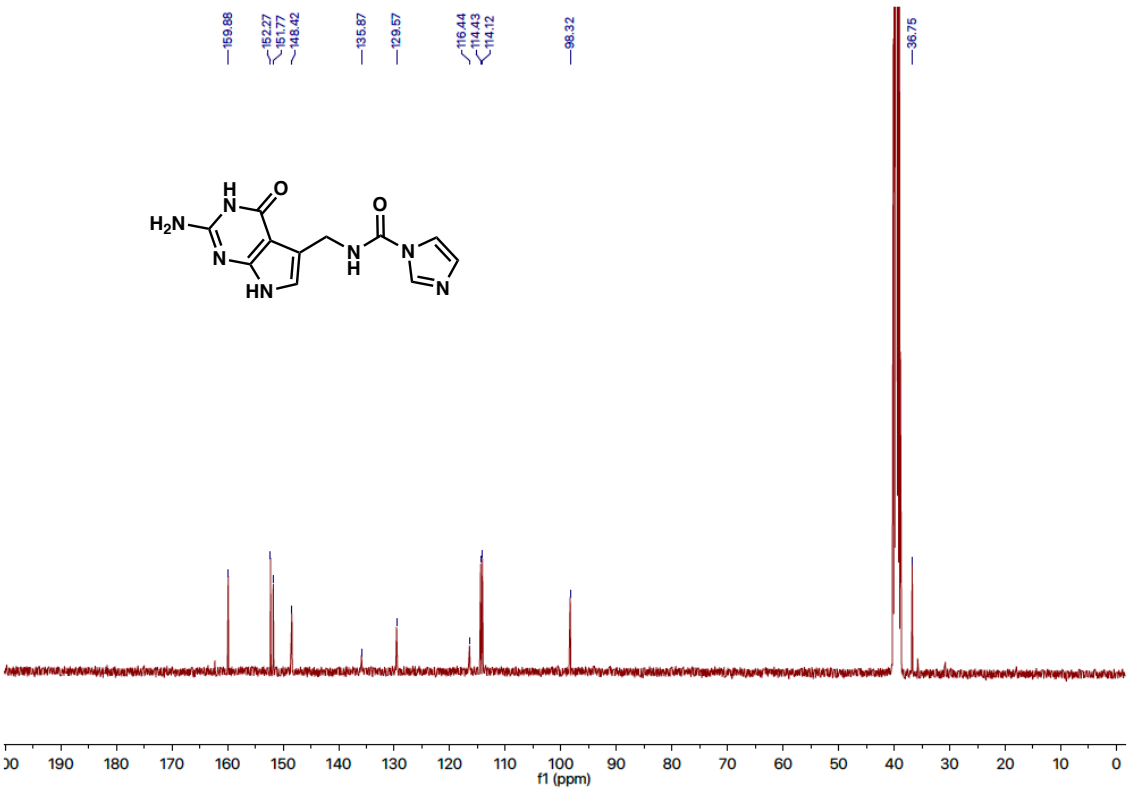


Compound 8 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-1*H*-imidazole-1-carboxamide).

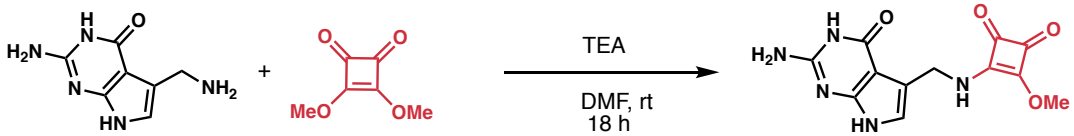


To a solution of PreQ1•2HCl (20 mg, 80 μmol, 1.0 equiv.) in DMF (1.0 mL), carbonyldiimidazole was added and the resulting reaction mixture was stirred at room temperature for 18 h. Then, the reaction mixture was concentrated, triturated with 20% MeOH/DCM, and the resulting precipitate was collected to obtain the titled compound as a white solid. ^{1}H NMR (500 MHz, DMSO): δ 11.03 (s, 1H), 10.58 (s, 1H) 9.39 (t, J = 5.3 Hz, 1H), 8.30 (s, 1H), 7.73 (s, 1H), 7.09 (s, 1H), 6.70 (s, 1H), 6.21 (s, 2H), 4.56 (d, J = 5.3 Hz, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 159.88, 152.27, 151.77, 148.42, 135.87, 129.57, 116.44, 114.43, 114.12, 98.32, 36.75; HRMS m/z: calcd. for $C_{11}H_{12}N_7O_2$ $[M+H]^+$ 274.1047, found 274.1040.

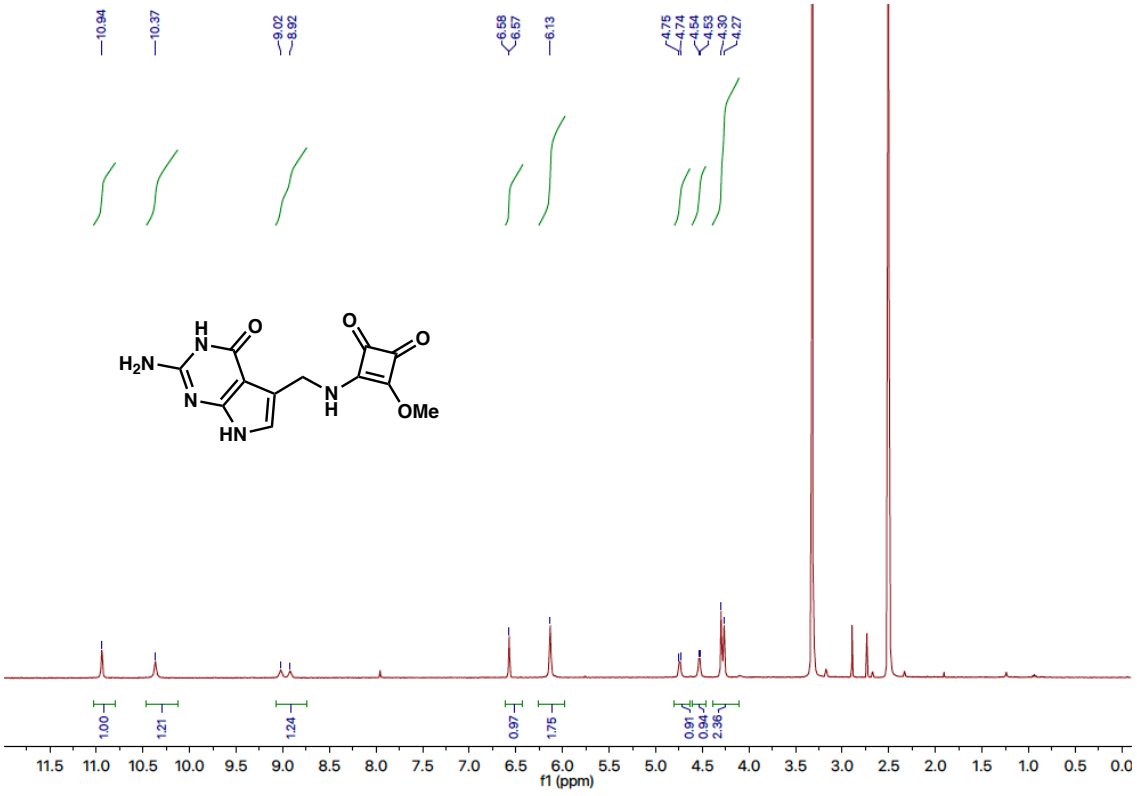

11.03
10.58
9.40
9.39
9.37
8.30
7.73
7.09
6.70
6.21
4.56
4.55
1.00
1.00
1.01
0.98
1.04
1.01
1.10
1.97
1.96
f1 (ppm)

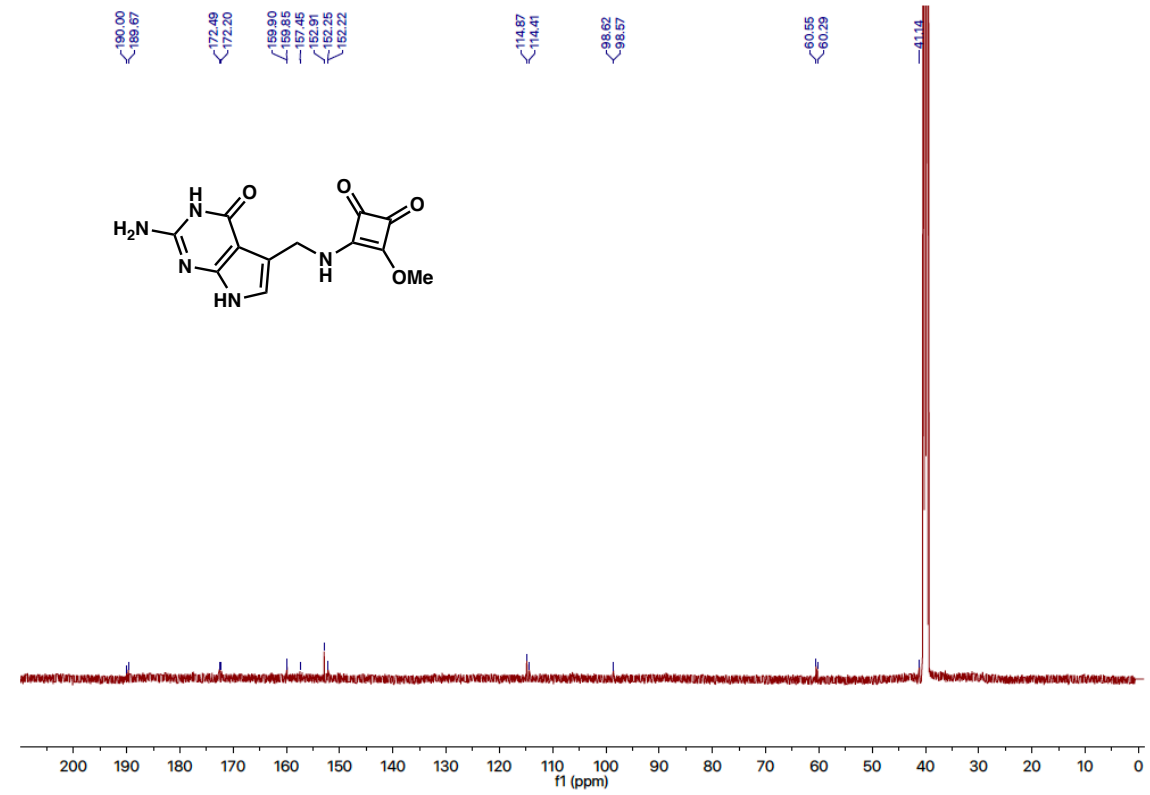

159.88
152.27
151.77
148.42
135.87
129.57
116.44
114.43
114.12
98.32
36.75
f1 (ppm)

Compound 9 (3-(((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)amino)-4-methoxycyclobut-3-ene-1,2-dione).

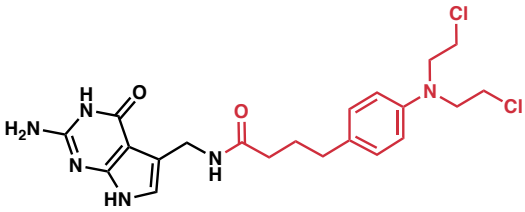


To a solution of PreQ1•2TFA in anhydrous DMF (600 μL) under N_2, dimethyl squarate (7.0 mg, 49 μmol, 1.0 equiv.) and triethylamine (15 μL, 103 μmol, 2.1 equiv.) dissolved in anhydrous DMF (400 μL) was added and stirred at room temperature for 18 h. Then, the reaction mixture was concentrated under reduced pressure, diluted with 1.0 mL of MeOH, and resulting precipitate was collected by filtration to obtain the titled compound as a white solid (57% yield). ^{1}H NMR (500 MHz, DMSO): δ 10.94 (s, 1H), 10.37 (s, 1H), 9.02-8.92 (s, 1H), 6.58 (d, J = 2.1 Hz, 1H), 6.13 (s, 2H), 4.75 (d, J = 5.6 Hz, 1H), 4.54 (d, J = 5.6 Hz, 1H), 4.30-4.27 (s, 3H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 190.0, 189.7, 172.5, 172.2, 159.9, 159.9, 157.5, 152.9, 152.5, 152.2, 114.9, 114.4, 98.6, 98.6, 60.6, 60.3, 41.1 (Note that the compound exist as a mixture of tautomers and as a result, there are more carbon peaks observed); HRMS m/z: calcd. for $C_{12}H_{12}N_5O_4$ $[M+H]^+$ 290.0884, found 290.0877.

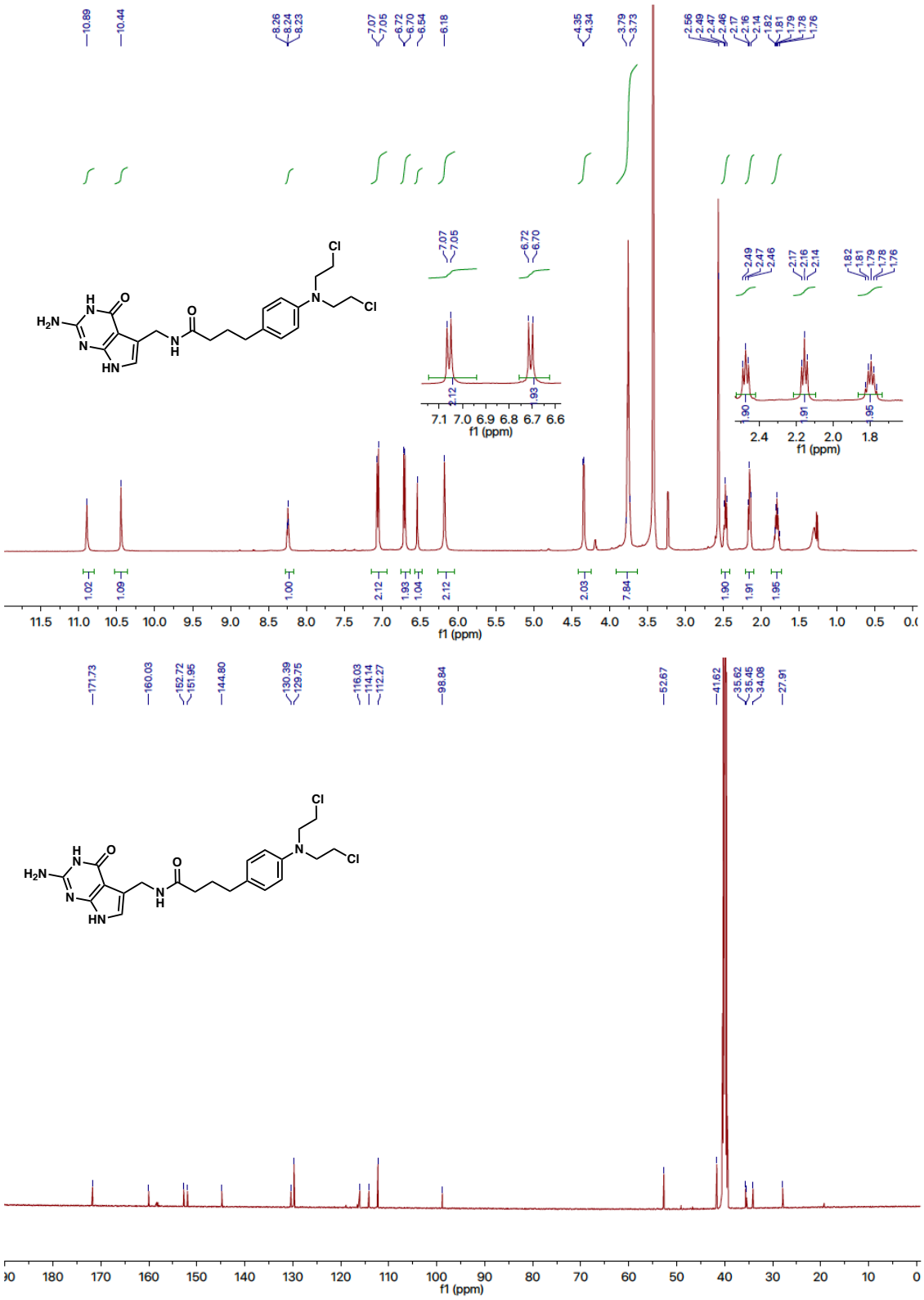




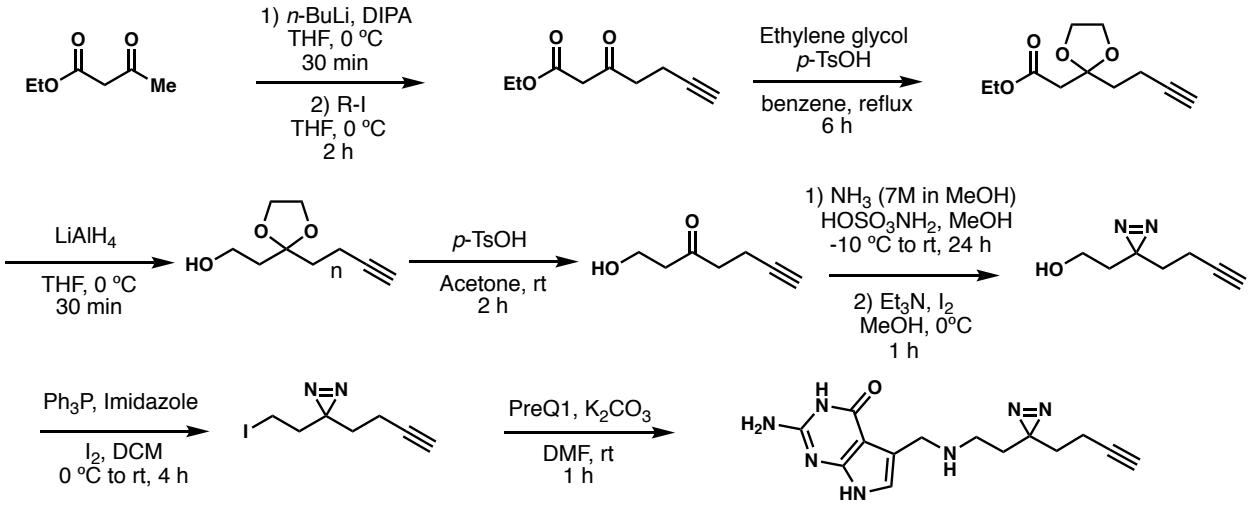
Compound 10 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-4-(4-(bis(2-chloroethyl)amino)phenyl)butanamide).



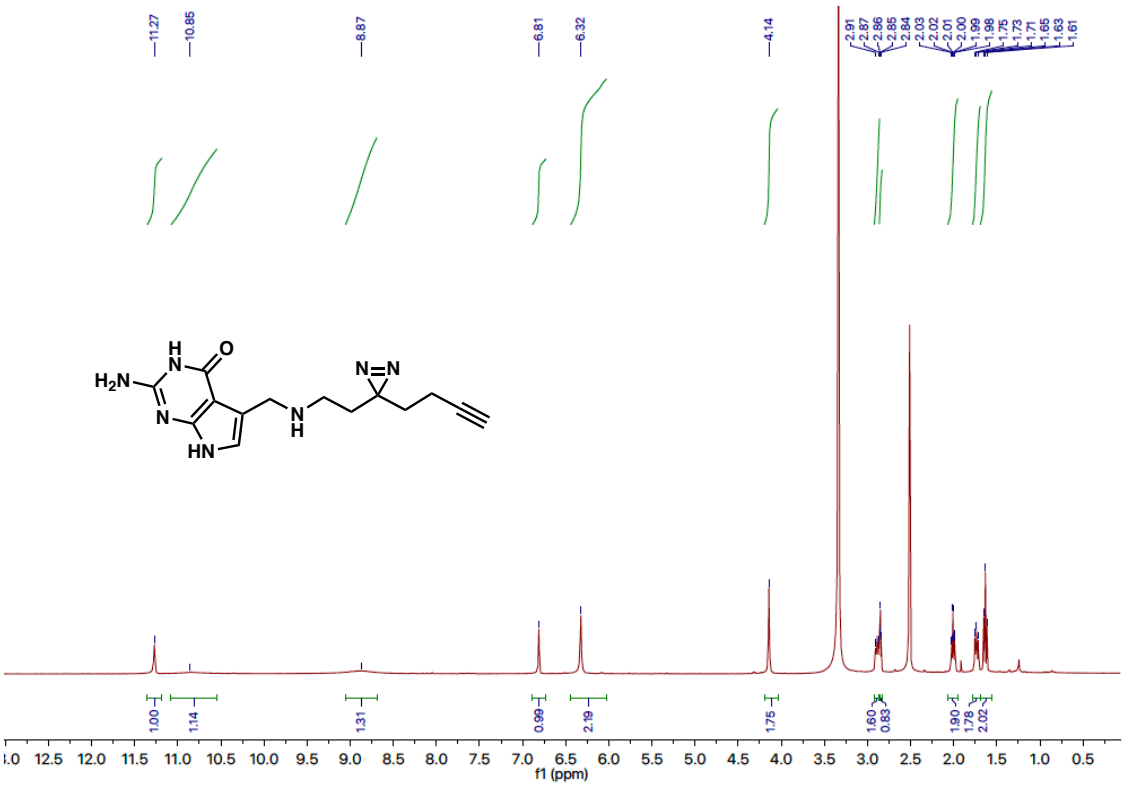
Prepared according to the general procedure **B**. ^{1}H NMR (500 MHz, DMSO): δ 10.89 (s, 1H), 10.44 (s, 1H), 8.24 (t, *J* = 5.4 Hz, 1H), 7.06 (d, *J* = 8.7 Hz, 2H) , 6.71 (d, *J* = 8.7 Hz, 2H) , 6.54 (s, 1H), 6.18 (s, 2H), 4.35 (d, *J* = 5.3 Hz, 2H), 3.79-3.73 (m, 8H), 2.47 (t, *J* = 7.5 Hz, 2H), 2.16 (t, *J* = 7.5 Hz, 2H),1.79 (qt, *J* = 7.5 Hz, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 71.7, 160.0, 152.7, 152.0, 144.8, 130.4, 129.8, 116.0, 114.1, 112.3, 98.8, 52.7, 41.6, 35.6, 35.5, 34.1, 27.9; HRMS m/z: calcd. for $C_{21}H_{27}Cl_2N_6O_2$ $[M+H]^+$ 465.1567, found 465.1555.

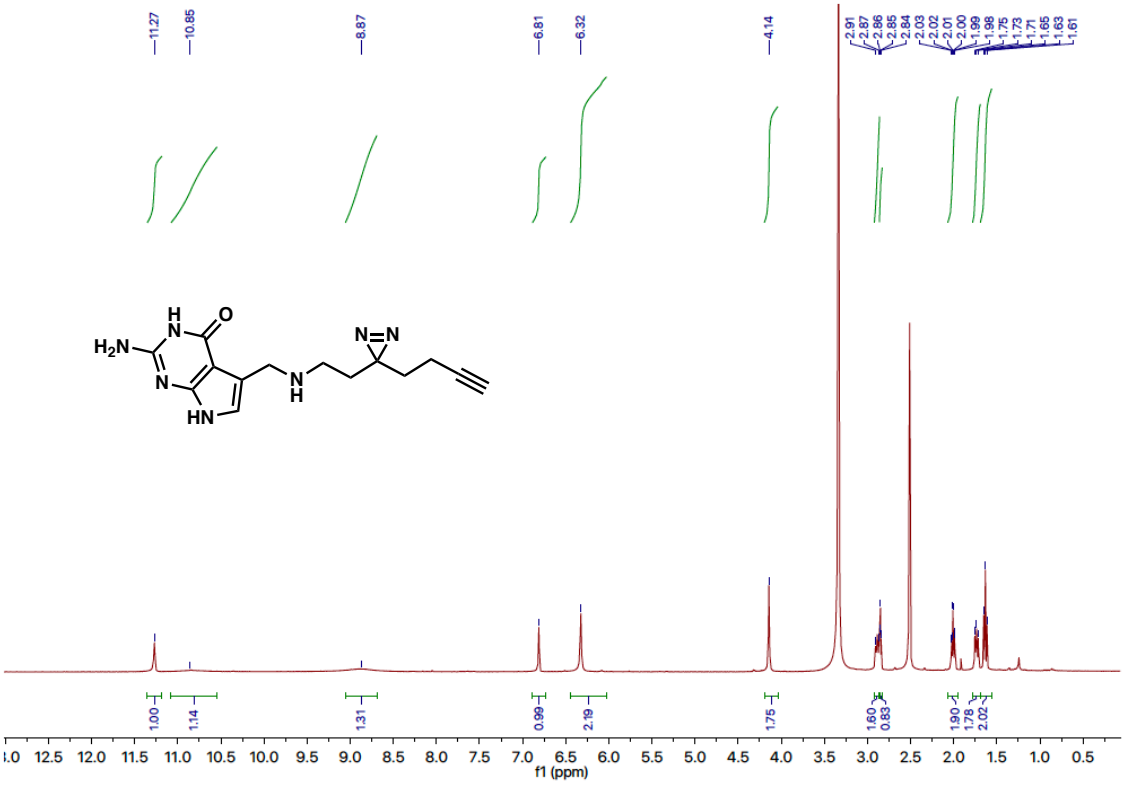


Compound 11 (2-amino-5-(((2-(3-(but-3-yn-1-yl)-3*H*-diazirin-3-yl)ethyl)amino)methyl)-3,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidin-4-one).

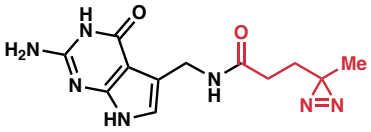


To a solution of PreQ1•2TFA (40 mg, 98 μmol, 1.0 equiv.) and K_2CO_3 (43 mg, 314 μmol, 3.2 equiv.) in DMF (1.0 mL), 3-(but-3-yn-1-yl)-3-(2-iodoethyl)-3*H*-diazirine (prepared according to the literature procedure (25 μL, 98 μmol, 1.0 equiv.) in MeCN (0.5 mL) was slowly added and stirred at room temperature for 1 h in dark. Then, the reaction mixture was concentrated and purified via column chromatography using a silica gel and eluting with gradient based MeOH:DCM to obtain the titled compound as a white solid. ^{1}H NMR (500 MHz, DMSO): δ 11.27 (s, 1H), 10.85 (brs, 1H), 8.87 (brs, 1H), 6.81 (s, 1H), 6.32 (s, 2H), 4.14 (s, 2H), 2.91-2.87 (m, 2H), 2.85 (t, J =2.7 Hz, 1H) , 2.00 (td, J = 7.3, 2.7 Hz, 2H), 1.75-171 (m, 2H), 1.63 (t, J = 7.3 Hz, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 160.8, 153.2, 152.7, 117.9, 108.7, 98.7, 83.4, 72.4, 43.4, 41.3, 31.2, 29.8, 26.9, 13.1; HRMS m/z: calcd. for $C_{14}H_{18}N_7O$ $[M+H]^+$ 300.1567, found 300.1558.

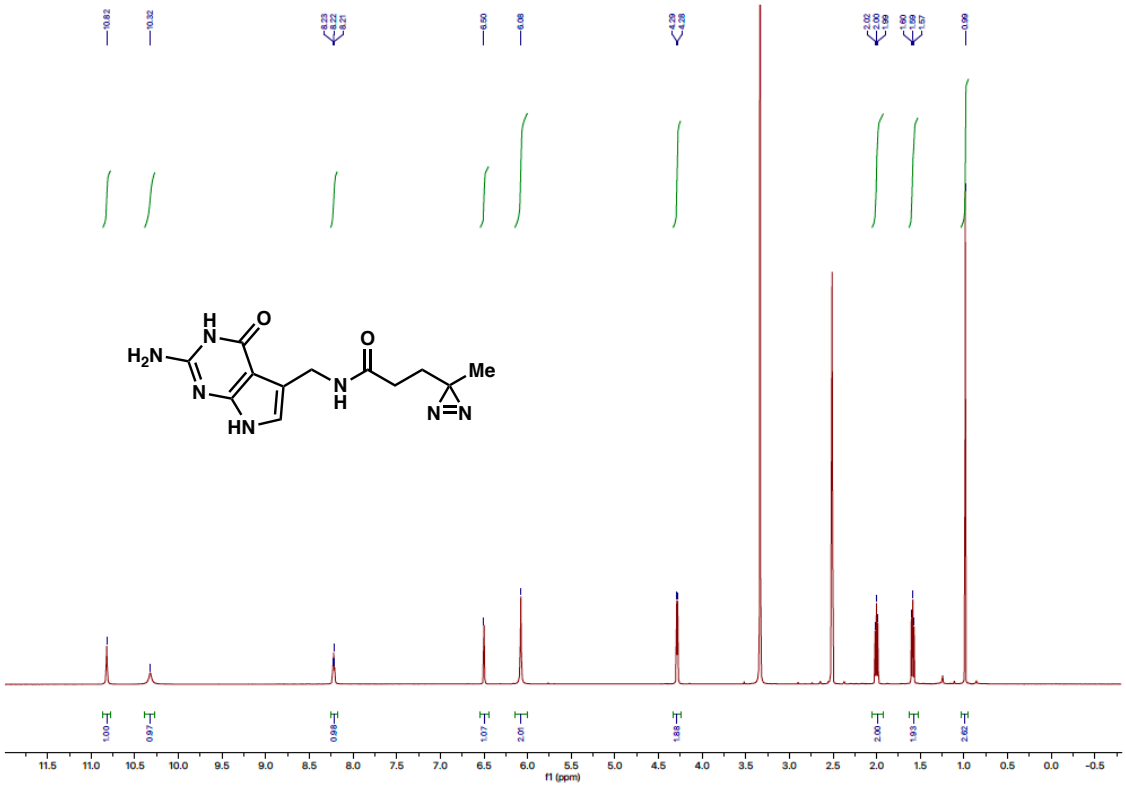

11.27
10.85
8.87
6.81
6.32
4.14
2.91
2.87
2.86
2.85
2.84
2.03
2.02
2.01
2.00
1.99
1.98
1.75
1.73
1.71
1.65
1.63
1.61
1.00
1.14
1.31
0.99
2.19
1.75
1.60
0.83
1.90
1.78
2.02
f1 (ppm)

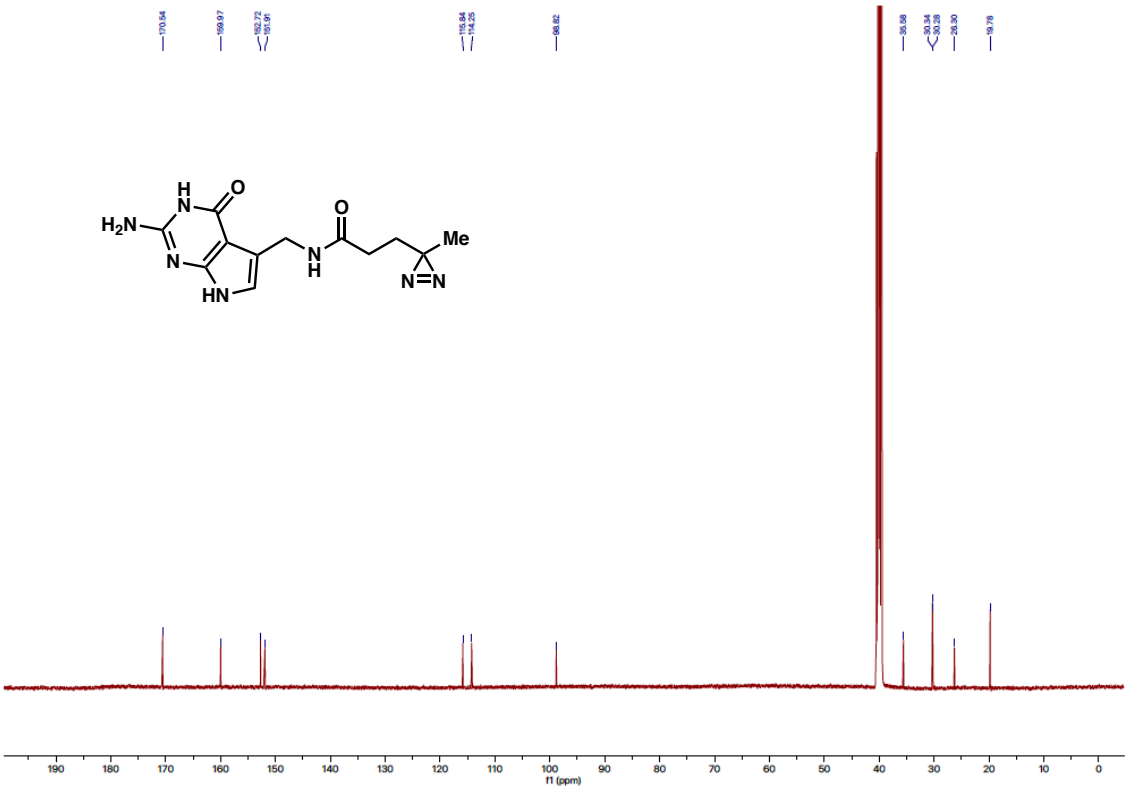

11.27
10.85
8.87
6.81
6.32
4.14
2.91
2.87
2.86
2.85
2.84
2.03
2.02
2.01
2.00
1.99
1.98
1.75
1.73
1.71
1.65
1.63
1.61
1.00
1.14
1.31
0.99
2.19
1.75
1.60
0.83
1.90
1.78
2.02
f1 (ppm)

Compound 12 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide).

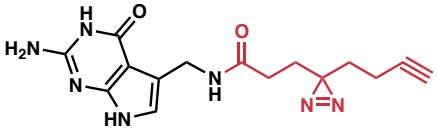


The 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid was prepared according the literature protocol. Then, compound 20 was prepared according to the general procedure **B**. White solid (12.0 mg, 85% yield). ^{1}H NMR (500 MHz, DMSO): δ 10.82, 10.32, 8.22 (t, *J* = 5.3 Hz, 1H) , 6.50 (s, 1H), 6.08 (s, 1H), 4.29 (d, *J* = 5.3 Hz, 2H) , 2.02-1.99 (m, 2H), 1.60-1.57 (m, 2H), 0.99 (s, 3H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 170.5, 160.0, 152.72, 151.91, 115.84, 114.25, 98.82, 35.58, 30.34, 30.28, 26.30, 19.78; HRMS m/z: calcd. for $C_{12}H_{16}N_7O_2$ $[M+H]^+$ 290.1360, found 290.1352.

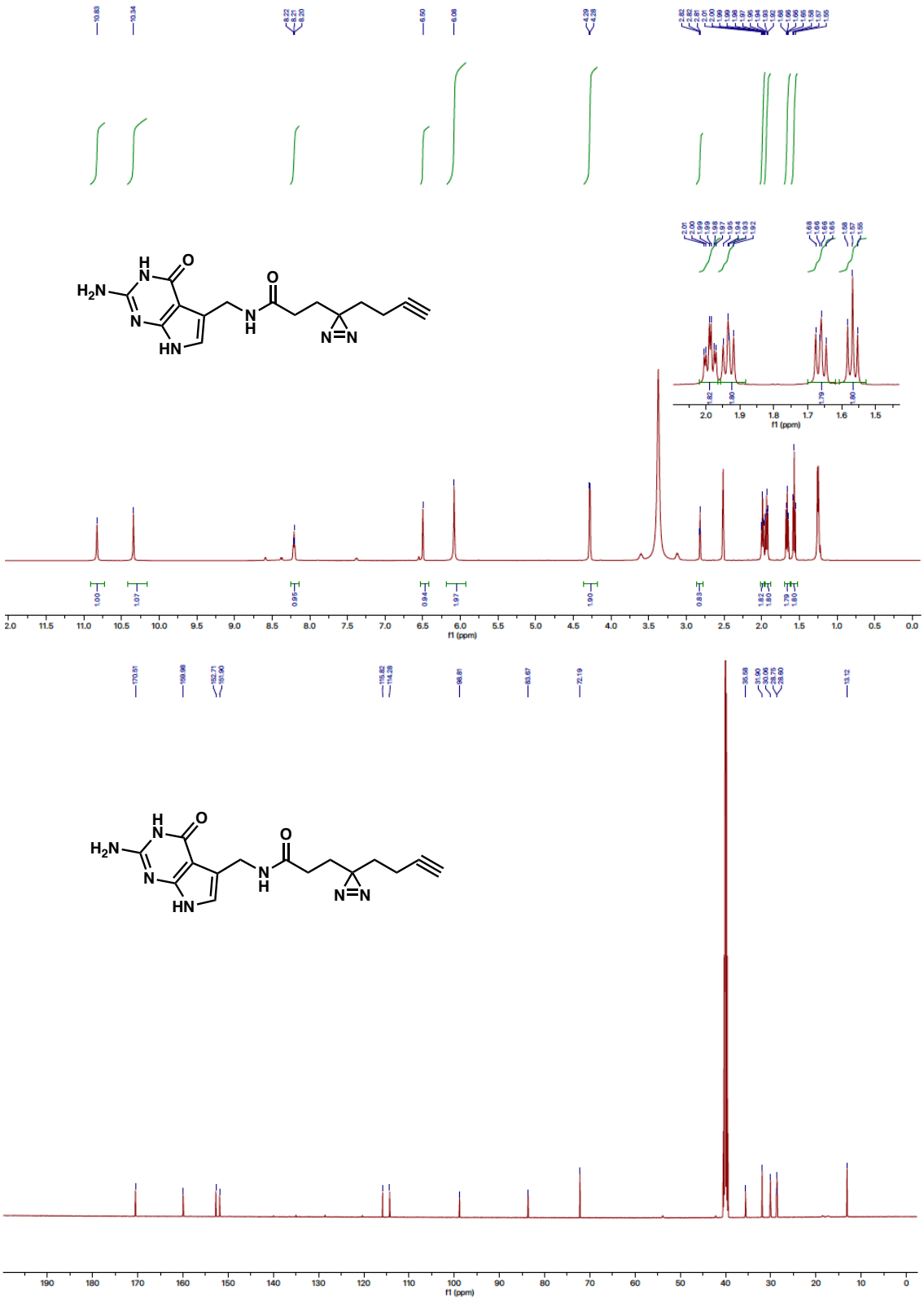




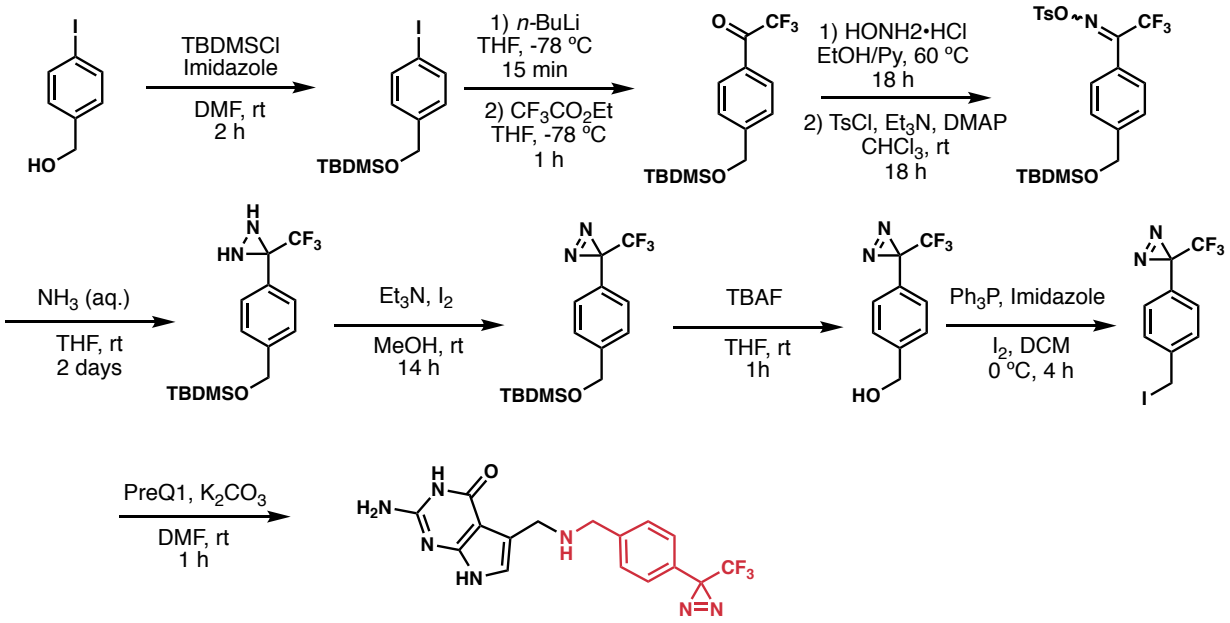
Compound 13 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-3-(3-(but-3-yn-1-yl)-3*H*-diazirin-3-yl)propenamide).



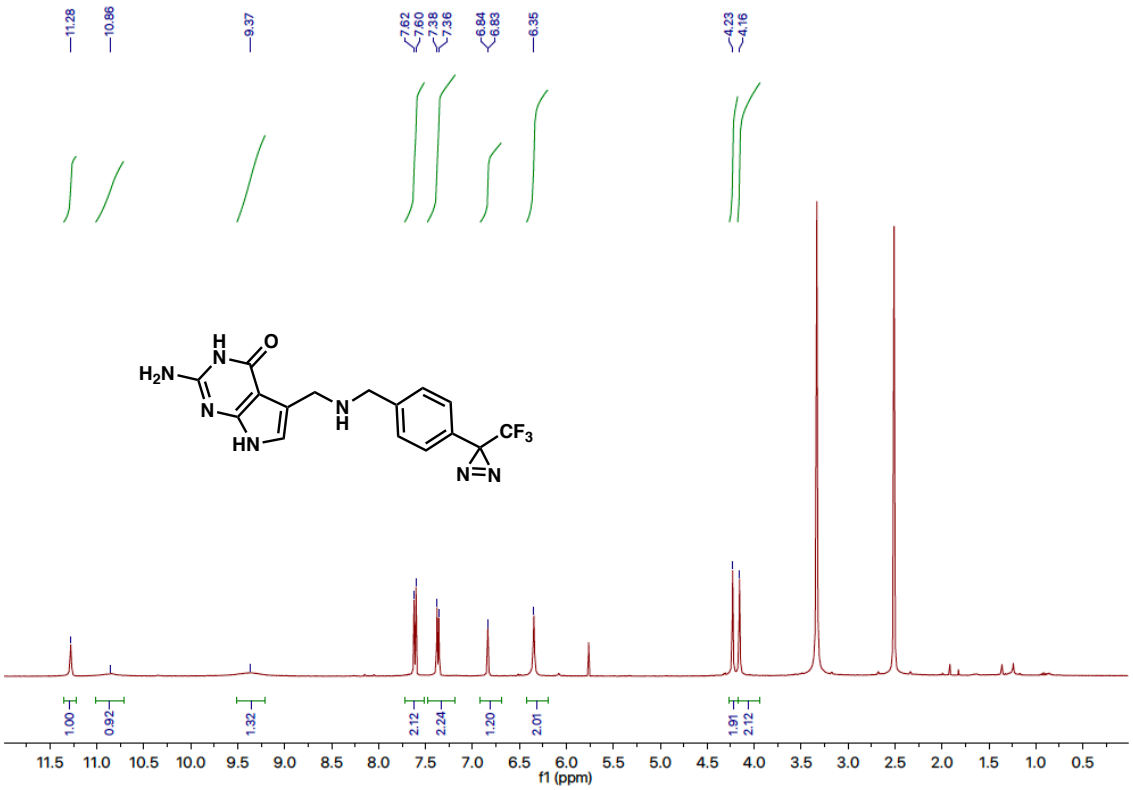
Prepared according to the general procedure **B** (Note that the reaction was performed in dark). White solid. ^{1}H NMR (500 MHz, DMSO): δ 10.83 (s, 1H), 10.34 (s, 1H), 8.21 (t, *J* = 5.4 Hz, 1H) , 6.50 (s, 1H), 6.08 (s, 2H), 4.28 (d, *J* = 5.4 Hz, 2H), 2.82 (t, *J* = 2.6 Hz, 1H), 2.00 (ddd, *J* = 7.5, 4.8, 2.6 Hz, 2H), 1.97-1.92 (m, 2H) 1.68-1.65 (m, 2H), 1.57 (t, *J* = 7.4 Hz, 2H). ^{13}C{^{1}H} NMR (126 MHz, DMSO): 170.5, 160.0, 152.7, 151.9, 115.8, 114.3, 98.8, 83.7, 72.2, 35.6, 31.9, 30.1, 28.8, 28.6, 13.1; HRMS m/z: calcd. for $C_{15}H_{18}N_7O_2$ $[M+H]^+$ 328.1516, found 328.1511.

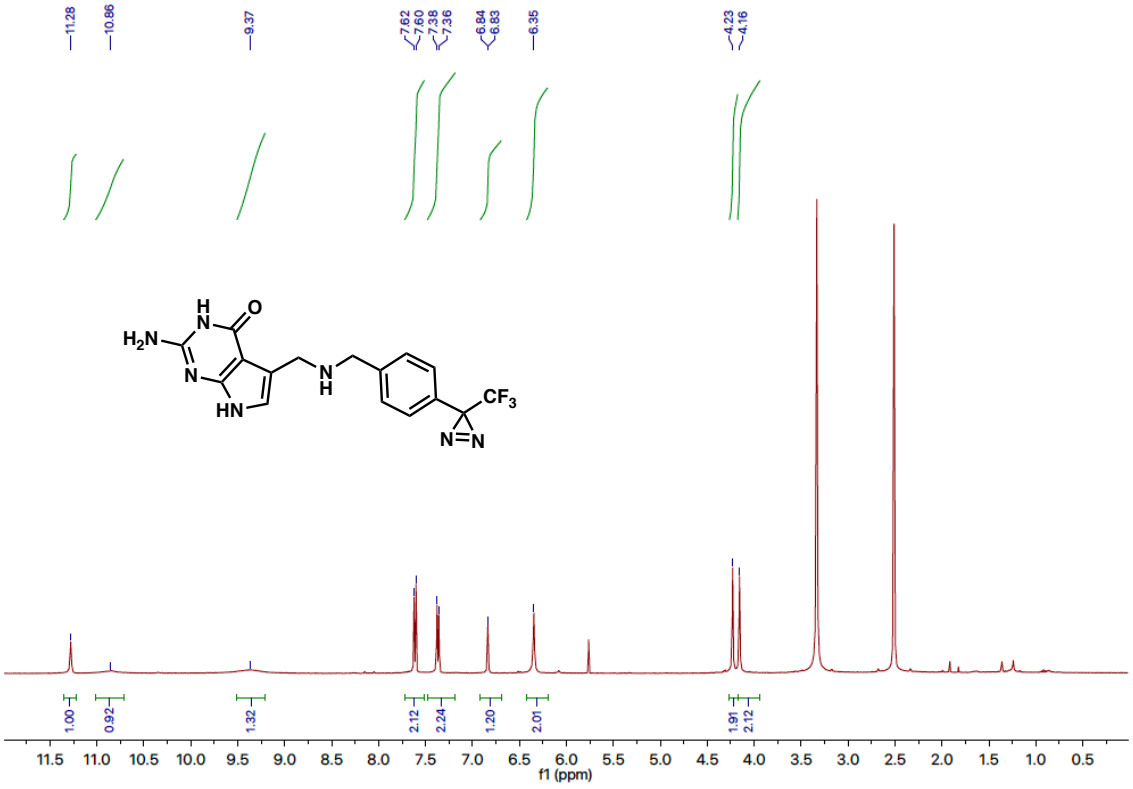


Compound 14 (2-amino-5-(((4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzyl)amino)methyl)-3,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidin-4-one).

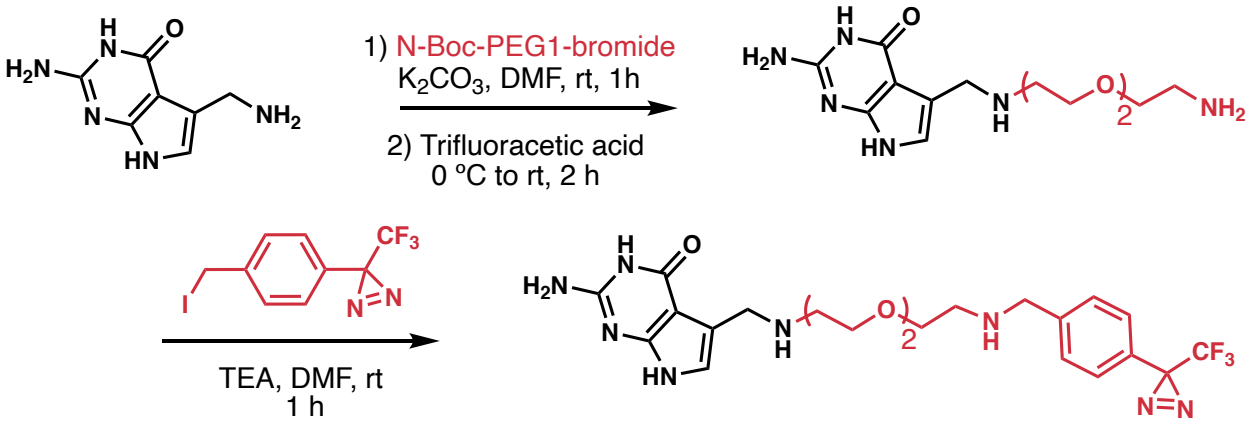


The 3-(4-(iodomethyl)phenyl)-3-(trifluoromethyl)-3*H*-diazirine was prepared according to the literature procedure. Then, compound **15** was prepared according the procedure for compound **14**. White solid. ^{1}H NMR (500 MHz, DMSO): δ 11.28 (s, 1H), 10.86 (s, 1H), 9.37 (s, 1H), 7.61 (d, J = 8.5 Hz, 2H) , 7.37 (d, J = 8.1 Hz, 2H) , 6.84 (d, J = 2.2 Hz, 1H) 6.35 (s, 2H), 4.23 (s, 2H), 4.16 (s, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): δ 160.9, 153.1, 152.8, 131.1, 128.4, 127.3, 123.4 (q, J = 276 Hz) , 117.7, 98.7, 49.3, 43.6, 40.9, 28.3 (q, J = 40.0 Hz); ^{19}F NMR (471 MHz, DMSO) δ -64.6; HRMS m/z: calcd. for $C_{16}H_{15}F_3N_7O$ $[M+H]^+$ 378.1285, found 378.1271.

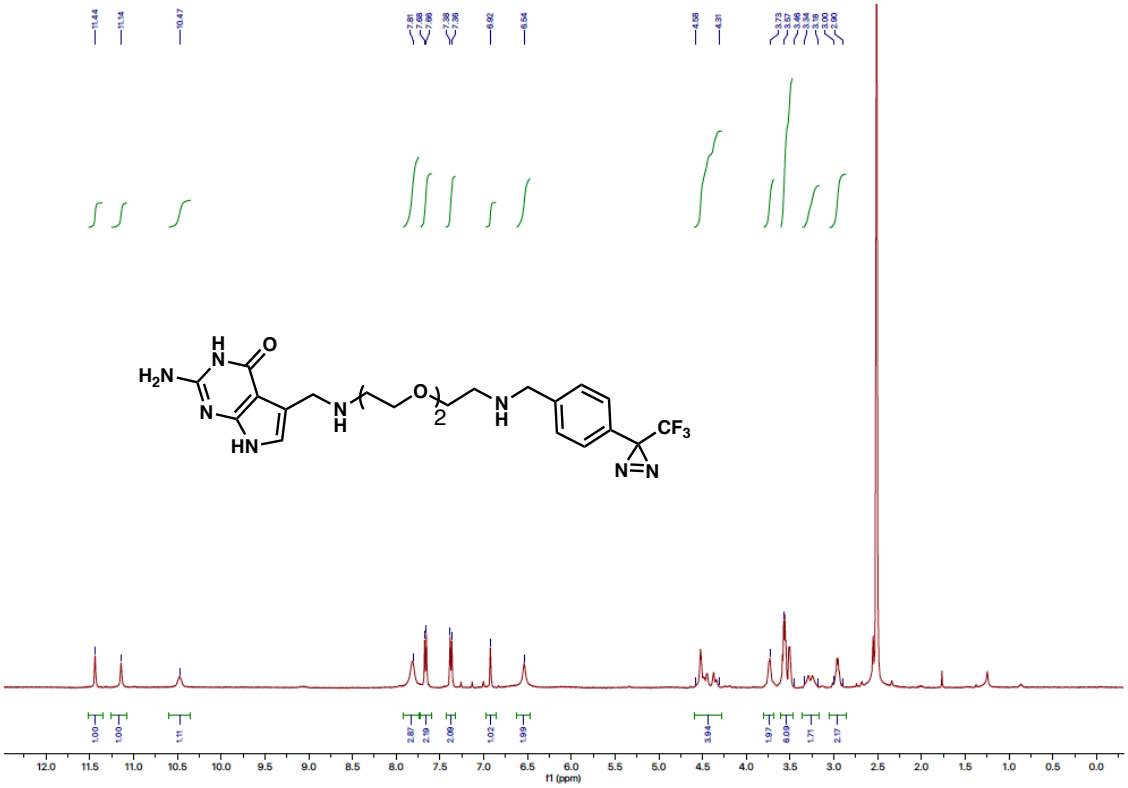

11.28
10.86
9.37
7.62
7.60
7.38
7.36
6.84
6.83
6.35
4.23
4.16
1.00
0.92
1.32
2.12
2.24
1.20
2.01
1.91
2.12
f1 (ppm)

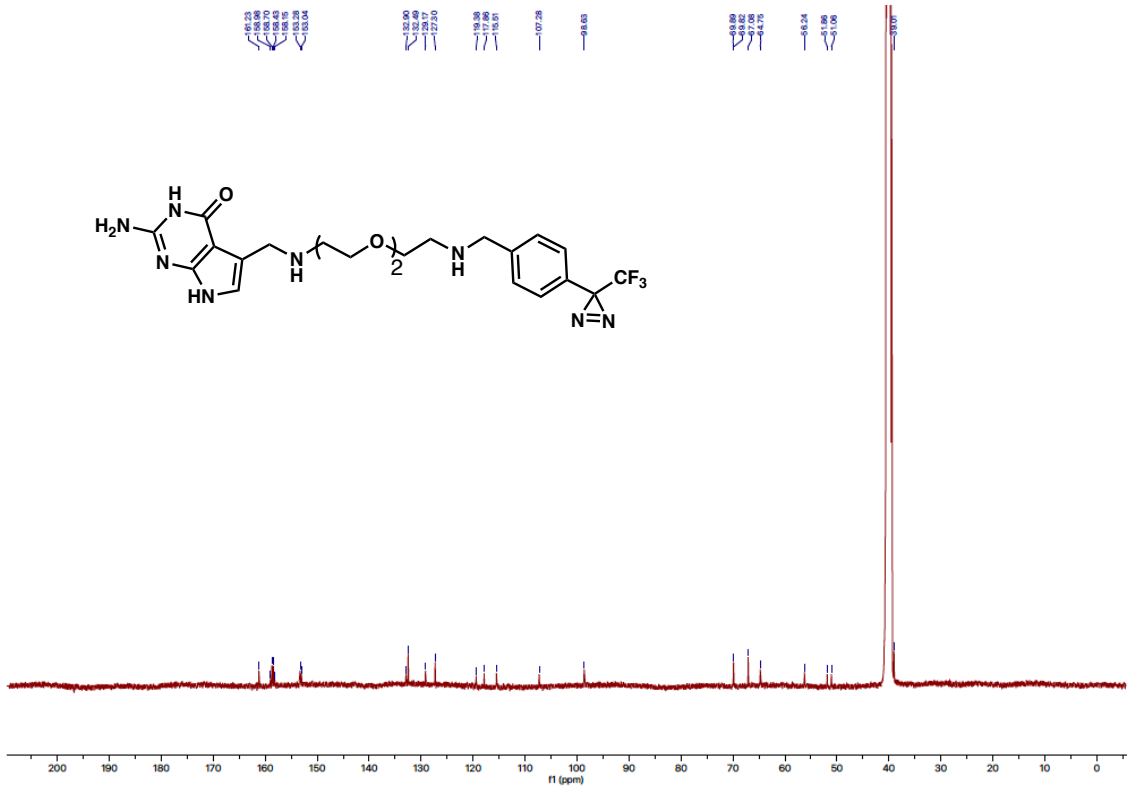

11.28
10.86
9.37
7.62
7.60
7.38
7.36
6.84
6.83
6.35
4.23
4.16
1.00
0.92
1.32
2.12
2.24
1.20
2.01
1.91
2.12
f1 (ppm)

Compound 15 (2-amino-5-(12-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8-dioxa-2,11-diazadodecyl)-3,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidin-4-one).

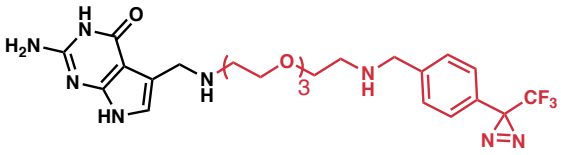


White solid. ^{1}H NMR (500 MHz, DMSO): δ 11.44 (s, 1H), 11.14 (s, 1H), 10.47 (s, 1H), 7.81 (brs, 3H), 7.64 (d, *J* = 8.1 Hz, 2H), 7.37 (d, *J* = 8.1 Hz, 2H), 6.92 (s, 1H), 6.54 (s, 2H), 4.58-4.31(m, 4H), 3.73 (m, 2H), 3.57-3.46 (m, 6H), 3.34-3.18 (m, 2H), 3.00-2.90 (m, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 161.2, 158.6 (q, *J* = 34.6 Hz) , 153.3, 153.0, 132.9, 132.5, 129.2, 127.3, 119.4, 116.7 (q, *J* = 293.6 Hz), 107.3, 98.6, 69.9, 69.8, 67.1, 64.8, 56.2, 51.9, 51.1, 39.0; HRMS m/z: calcd. for $C_{22}H_{28}F_3N_8O_3$ $[M+H]^+$ 509.2231; found 509.2224.

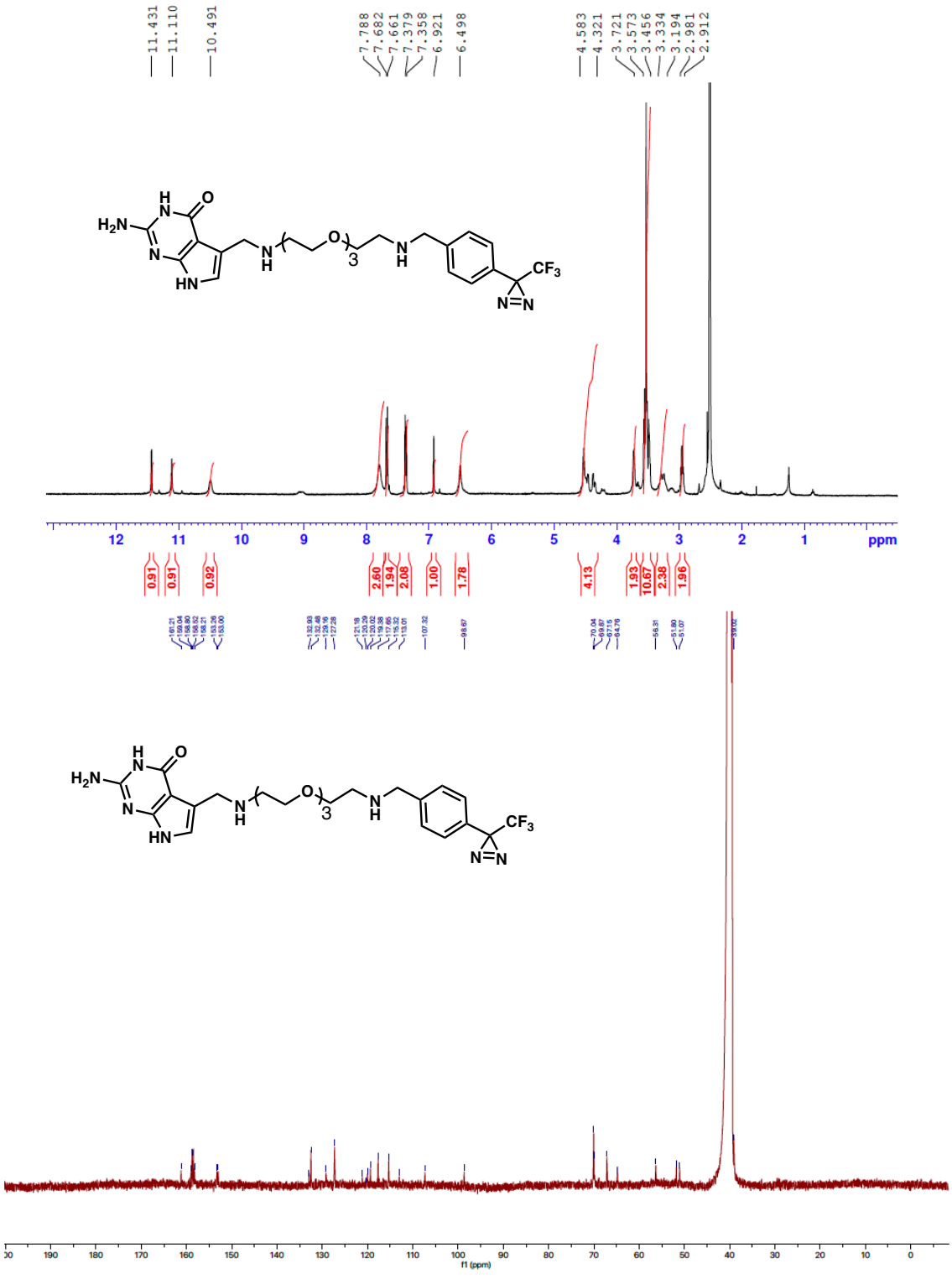




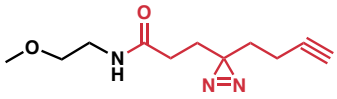
Compound 16 (2-amino-5-(15-(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)-5,8,11-trioxa-2,14-diazapentadecyl)-3,7-dihydro-4*H*-pyrrolo[2,3-*d*]pyrimidin-4-one).



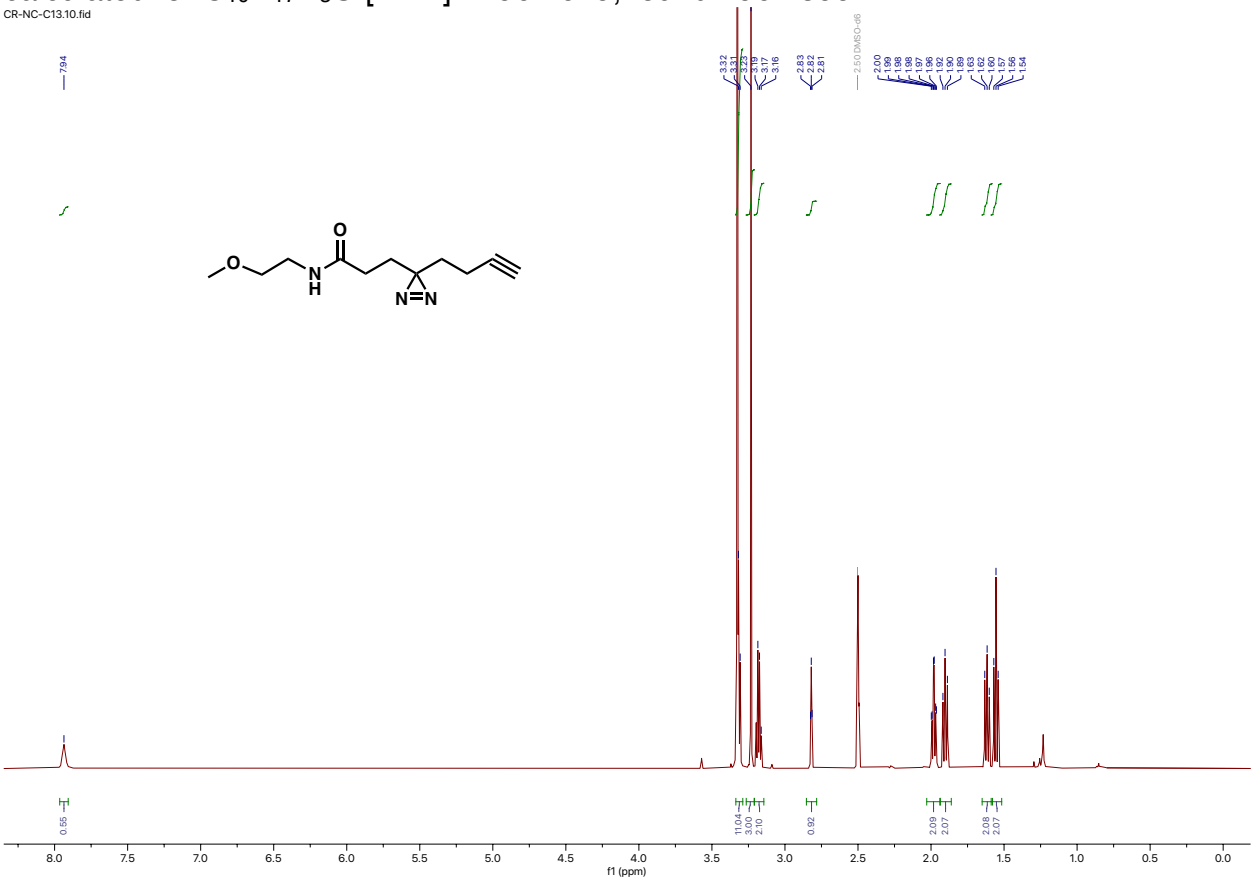
Prepared according to the procedure for compound **15**. White solid. ^{1}H NMR (500 MHz, DMSO): δ 11.43 (s, 1H), 11.11 (s, 1H), 10.48 (s, 1H) 7.78 (brs, 3H), 7.67 (d, *J* = 8.1 Hz, 2H), 7.37 (d, *J* = 8.1 Hz, 2H), 6.92 (s, 1H), 6.49 (s, 2H), 4.57-4.17 (m, 4H), 3.72 (m, 2H), 3.59-3.45 (m, 10H), 3.34-3.18 (m, 2H), 3.00-2.91 (m, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): 161.2, 158.6 (q, *J* = 36.0 Hz), 153.3, 153.0, 132.9, 132.5, 129.2, 127.3, 121.1, 119.4, 116.4 (q, *J* = 293.0 Hz), 107.3, 98.7, 70.0, 69.9, 67.2, 64.8, 56.3, 51.8, 51.1, 39.0; HRMS m/z: calcd. for $C_{24}H_{32}F_3N_8O_4$ $[M+H]^+$ 553.2493; found 553.2484.

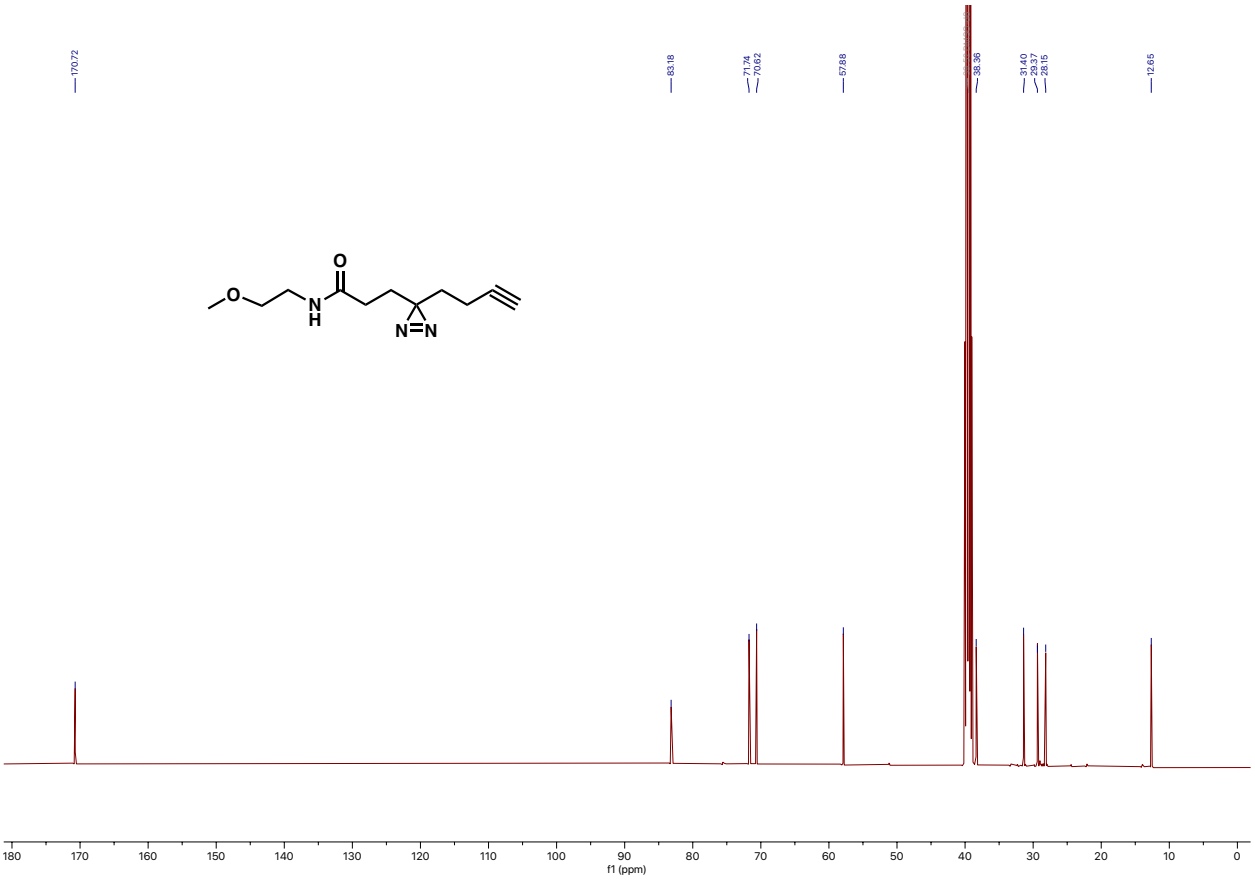


Compound 17 (*N*-((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)-3-(2-(2-((4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzyl)amino)ethoxy)ethoxy) propenamide).

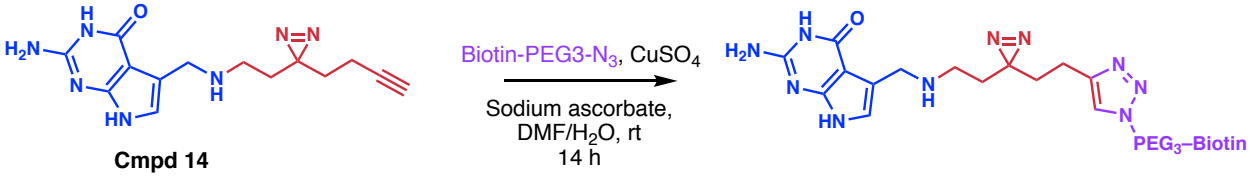


This compound was synthesized using the general procedure **B** for the coupling reaction. White solid. ^{1}H NMR (500 MHz, DMSO): δ 7.94 (s, 1H), 3.31 (d, J = 5.8 Hz, 2H), 3.23 (s, 3H), 3.21-3.14 (m, 2H), 2.82 (t, J = 2.7 Hz, 1H), 1.98 (td, J = 7.4, 2H), 1.94-1.86 (m, 2H), 1.62 (t, 2H), 1.56 (t, J = 7.4 Hz, 2H); ^{13}C{^{1}H} NMR (126 MHz, DMSO): δ 170.72, 83.18, 71.74, 70.62, 57.88, 38.36, 31.40, 29.37, 28.15, 12.65; HRMS m/z calculated for $C_{10}H_{17}N_3O$ [M+H]+ 196.2619, found 196.2599.



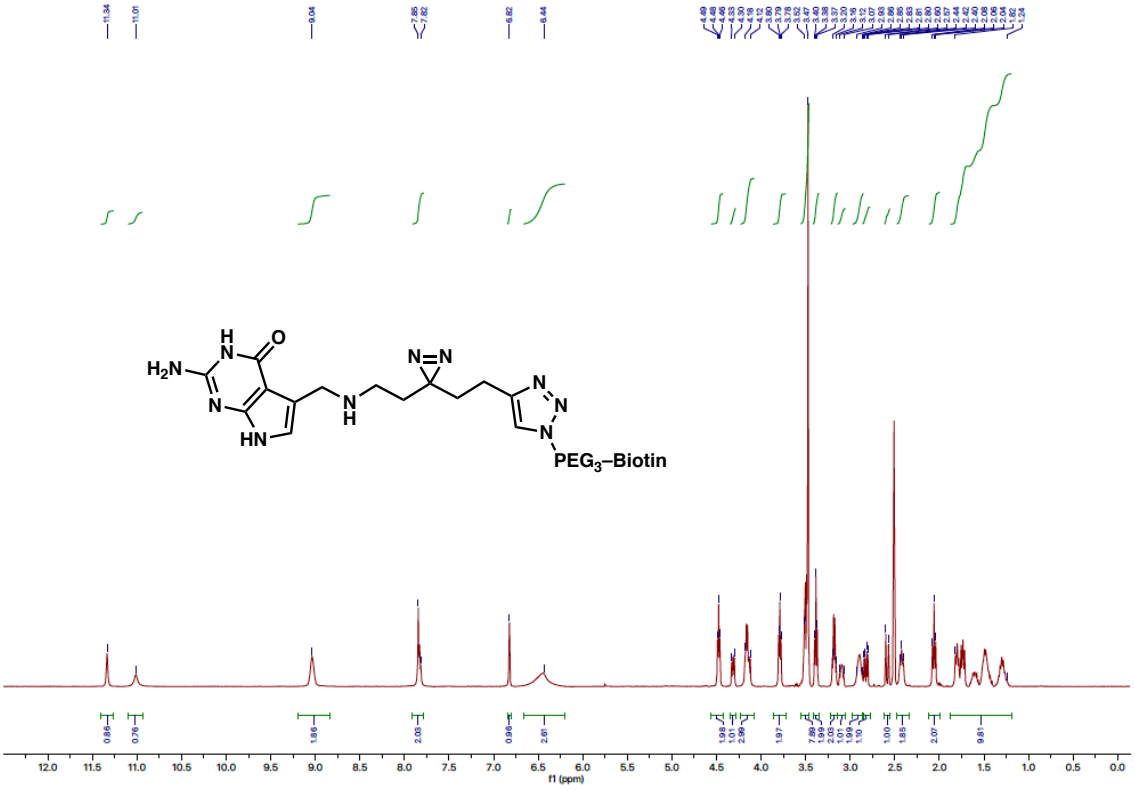


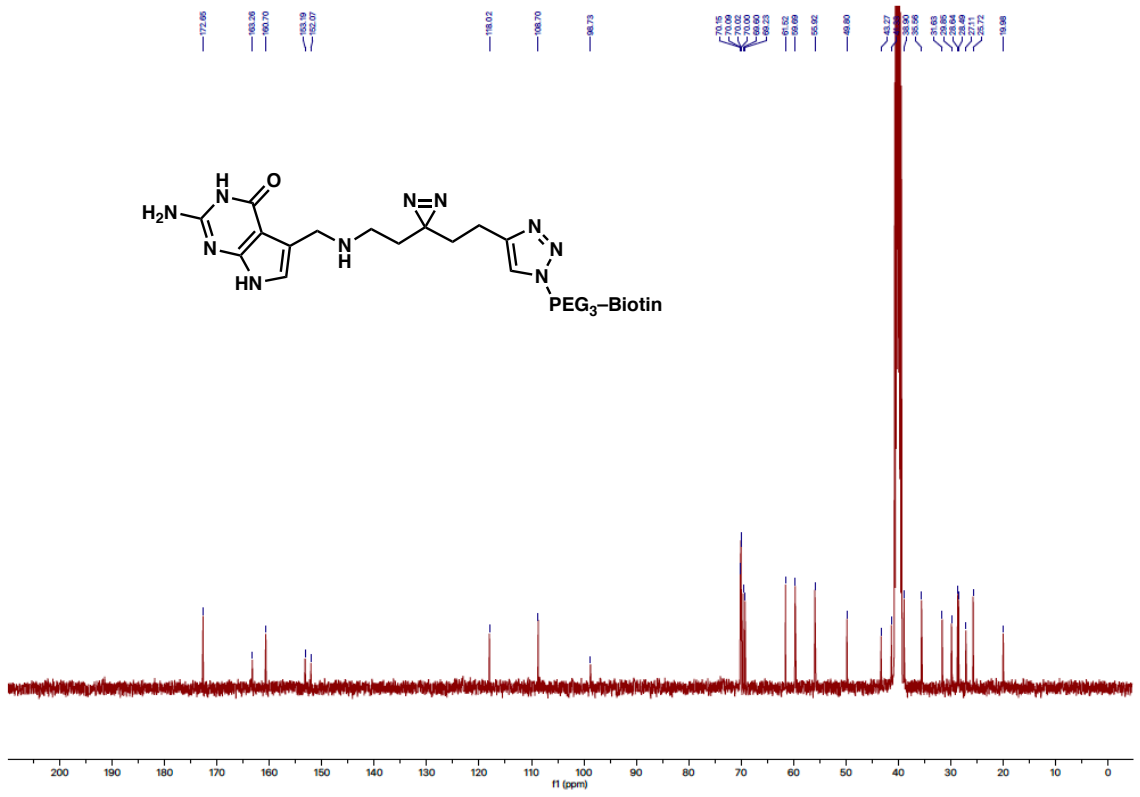
Compound Bio-11. (*N*-(2-(2-(2-(4-(2-(3-(2-(((2-amino-4-oxo-4,7-dihydro-3*H*-pyrrolo[2,3-*d*]pyrimidin-5-yl)methyl)amino)ethyl)-3*H*-diazirin-3-yl)ethyl)-1*H*-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethyl)-5-((3a*R*,4*R*,6a*S*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide).



To a solution of compound **14** (30.0 mg, 100 μmol, 1.0 equiv.) and Biotin-PEG-N_3 (45 mg, 100 μmol, 1.0 equiv.) in anhydrous DMF (2.0 mL), sodium ascorbate (100 μl of 0.2 M, 20 μmol, 0.2 equiv.) and $CuSO_4$ (100 μL of 0.2M, 20 μmol, 0.2 equiv.) and the reaction mixture was stirred at room temperature for 14 hours under N_2 atmosphere. Then, the reaction mixture was concentrated under reduced pressure and purified by preparative HPLC eluting with H_2O:MeCN containing 0.1% TFA to obtain the titled compound as white solid (40.0 mg, 54% yield). ^{1}H NMR (500 MHz, DMSO): δ 11.34 (s, 1H), 11.0 (s, 1H), 9.04 (s, 2H), 7.85-7.82 (m, 2H), 6.82 (s, 1H), 6.44 (brs, 3H), 4.48 (t, J = 5.2 Hz, 2H) , 4.32 (dd, J = 7.6, 4.8 Hz, 1H), 4.18-4.12 (m, 3H), 3.79 (t, J = 5.2 Hz, 2H), 3.52-3.47 (m, 8H), 3.38 (t, J = 6.0 Hz, 2H), 3.18 (q, J = 6.0 Hz, 2H), 3.12-3.07

(m, 1H), 2.93-2.86 (m, 2H), 2.82 (dd, J = 12.4, 5.0 Hz, 1H), 2.59 (d, J = 12.4 Hz, 1H), 2.44-2.40 (m 2H), 2.06 (t, J = 7.4 Hz, 2H), 1.82-1.24 (m, 10H); $^{13}C\{^{1}H\}$ NMR (126 MHz, DMSO): 172.7, 163.3, 160.7, 153.2, 152.1, 118.0, 108.7, 98.7, 70.2, 70.1, 70.02, 70.00, 69.6, 69.2, 61.5, 59.7, 55.9, 49.8, 43.3, 41.3, 38.9, 35.6, 31.6, 29.9, 28.6, 28.5, 27.1, 25.7, 20.0; HRMS m/z: calcd. for $C_{32}H_{50}N_{13}O_6S$ $[M+H]^+$ 744.3722, found 744.3716.




172.65
163.26
160.70
153.19
152.07
118.02
108.70
96.73
70.15
70.09
70.05
70.00
69.60
68.23
61.52
59.69
55.92
49.80
43.27
38.90
35.96
31.63
29.85
28.64
28.49
27.11
25.72
19.98
H
N
O
H2N
N
HN
N
H
N=N
N
N
N
PEG3–Biotin
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)