Optimising in situ click chemistry: the screening and identification of biotin protein ligase inhibitors

*William Tieu, Tatiana P. Soares da Costa, Min Y. Yap, K. L. Keeling, Matthew C. J. Wilce, John C. Wallace, Grant W. Booker, Steven W. Polyak, Andrew D. Abell**

Supporting Information

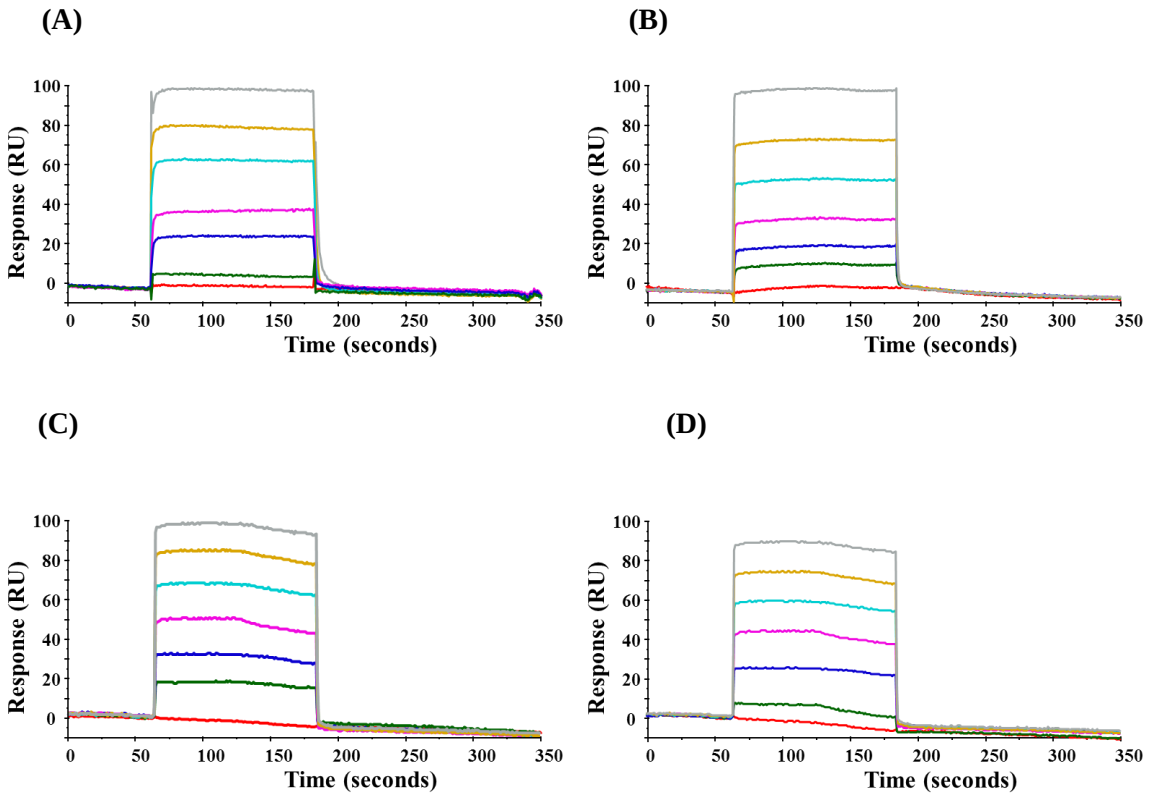
Figures and Tables

Methodology

Table S1. Kinetic characterisation of wild-type *Sa*BPL and *Sa*BPL-R122G.

	Biotin			***MgATP***		
	k_{cat} (s^{-1}) x 10^3	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1}) x 10^4	k_{cat} (s^{-1}) x 10^3	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1}) x 10^4
Wild-type *Sa*BPL	450±90	1.01±0.16	4500±190	360±70	0.18±0.03	20±0.2
*Sa*BPL-R122G	0.20±0.01	1.68±0.24	1.70±0.10	0.15±0.01	0.18±0.03	0.08±0.01

Figure S1. SPR analysis of biotin and AMP binding. Sensorgrams showing binding of various concentrations of biotin[a] to either wild-type *Sa*BPL (A) or *Sa*BPL-R122G (B). Sensorgrams showing binding of various concentrations of AMP[b] to either wild-type *Sa*BPL (C) or *Sa*BPL-R122G (D). AMP binding was measured after injection of saturating concentrations of biotin alkyne (100 µM) to pre-charge the BPLs. Enzymes were immobilised onto the same CM5 chip using standard amine coupling chemistry. One channel was left blank to detect non-specific binding to the sensorchip.



[a] Concentration of biotin: 500 µM (grey), 125 µM (yellow), 31.3 µM (cyan), 7.8 µM (magenta), 0.98 µM (blue), 0.24 µM (green), 0 µM (red).

[b] Concentration of AMP: 10 µM (grey), 5 µM (yellow), 2.5 µM (cyan), 1.25 µM (magenta), 0.6 µM (blue), 0.15 µM (green), 0 µM (red).

Figure S2. Overlaid CD spectra of wild-type *Sa*BPL (red) and *Sa*BPL-R122G (blue) demonstrating that the amino acid substitution at position 122 had no discernable effect upon the enzymes secondary structure.

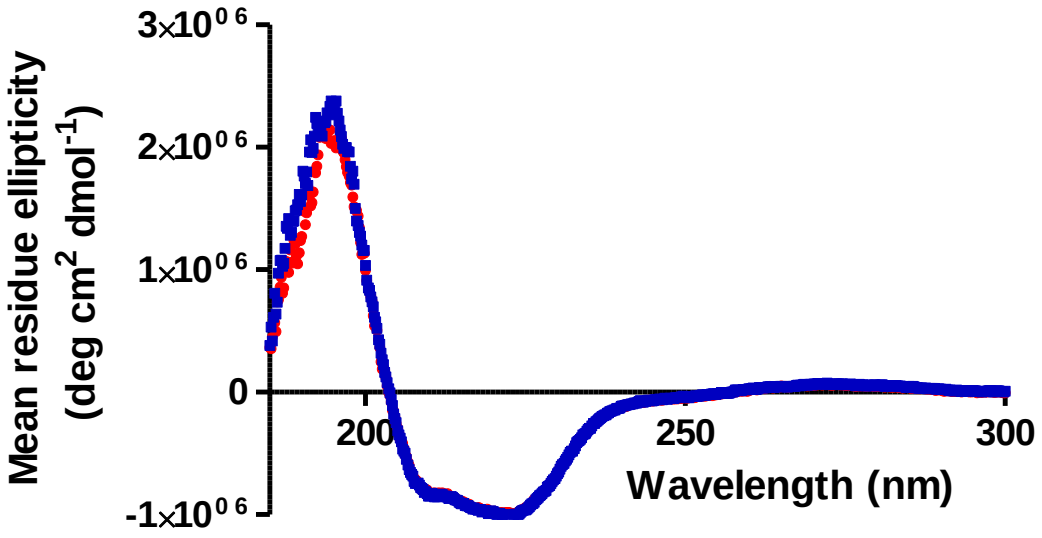


Figure S3. Streptavidin western blot analysis of soluble whole cell lysates. The lysates were separated by 12% SDS-PAGE and analysed by western blotting with streptavidin conjugated to Alexa488. Lane 1, Empty vector control; Lane 2, *Ec*BPL; Lane 3, *Ec*BPL-R118G; Lane 4, *Sa*BPL; Lane 5, *Sa*BPL-R122G. The arrows indicate the bands corresponding to biotin protein ligase (BPL) and the endogenous biotinylated protein in *E. coli*, the biotin carboxyl carrier protein (BCCP) subunit from acetyl CoA carboxylase. The streptavidin blot shows "promiscuous biotinylation" of non-target bacterial proteins caused by over-expression of the *Ec*BPL-R118G and *Sa*BPL-R122G muteins but not their wildtype equivalents. The promiscuous biotinylation phenotype is indicative of enhanced dissociation of biotinyl-5′-AMP from the enzyme where it can chemically modify non-target proteins.

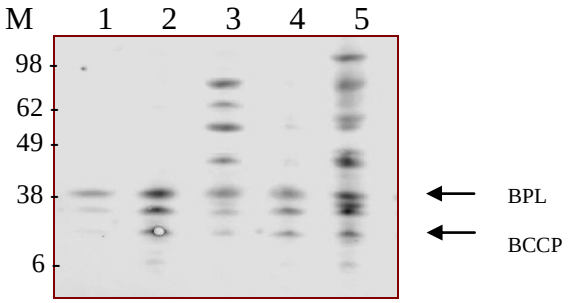


Figure S4 (aka s4) Analysis of products from *in situ* click reaction of alkyne **2** with azides **3–7** by HPLC (measured at 254 nm) and mass spectrometry. (A) In the presence of wild-type *Sa*BPL, (B) In the presence of *Sa*BPL-R122G showing triazole **8** at ~18 min, (C) In the absence of enzyme, (D) In the presence of bovine serum albumin instead of BPL, (E) In the presence of *Sa*BPL-R122G and biotinol-5´-AMP, which is evident at 16.7 min, (F) Mass spectrum of sample giving rise to (B) revealing triazole **8** (470.4 Da) and a trace of triazole **1** (570.3 Da).

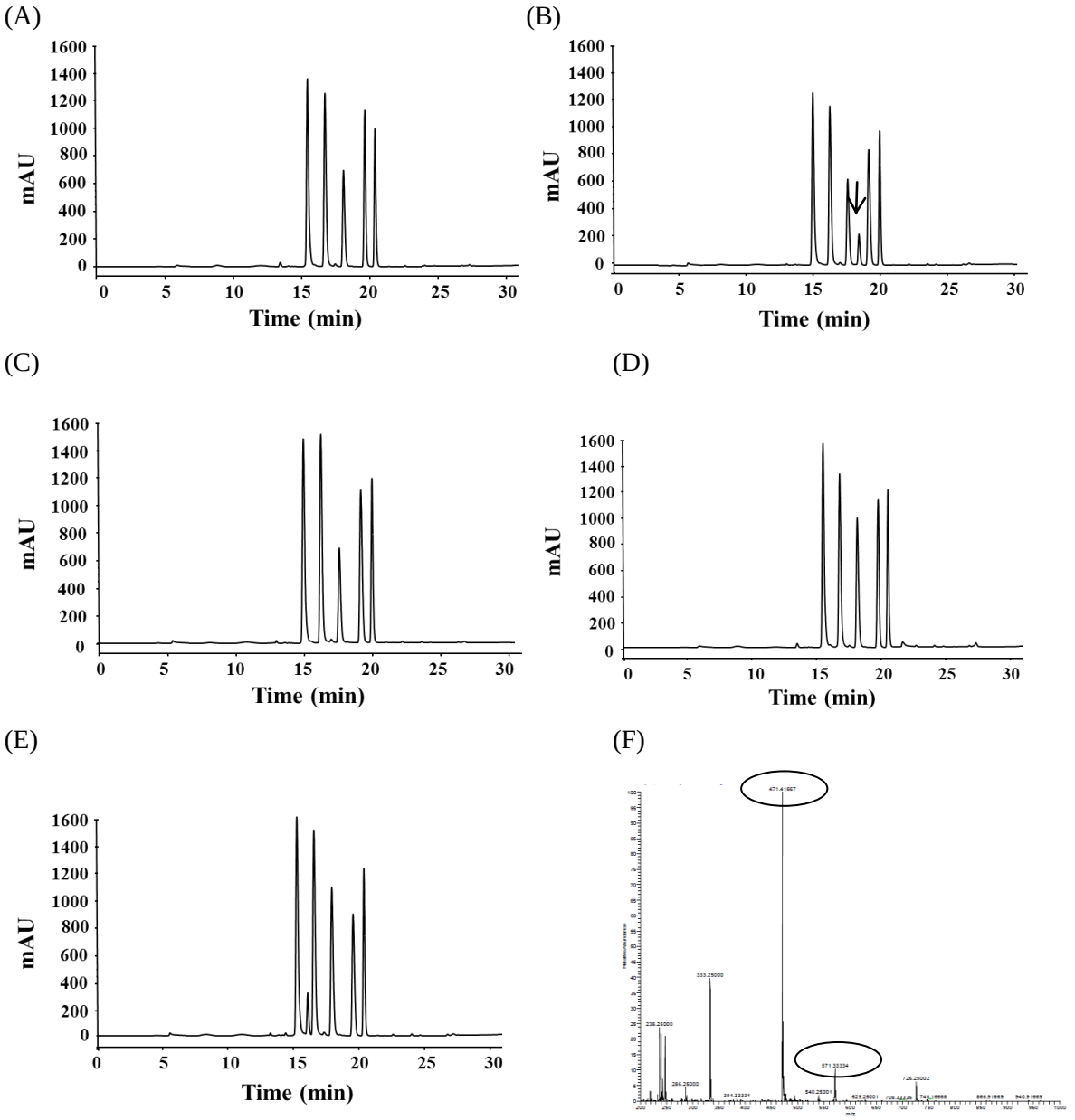
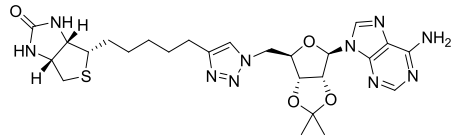
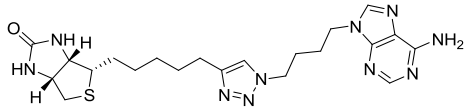
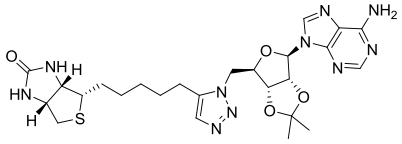
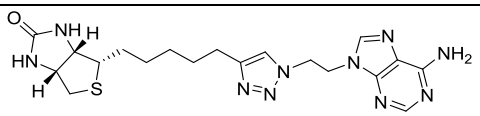
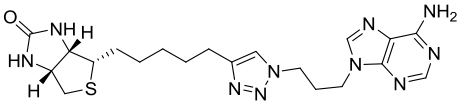
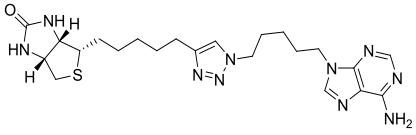
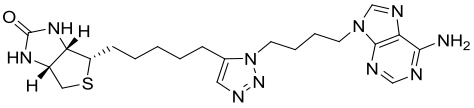
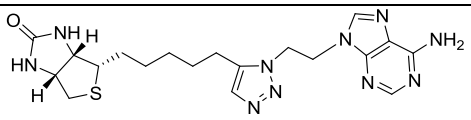
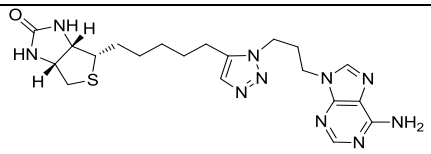
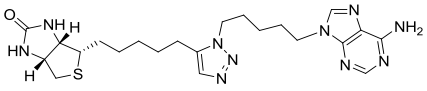
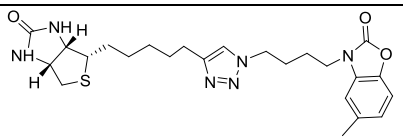


Table S2: HPLC retention times (R_t) and K_i values for synthetic **1**, **S1–S8**, **8**.

	Compound	HPLC R_t (min)[a]	*Sa*BPL K_i (μM)[b]
1		20.1	1.83 ± 0.33
8		18.5	0.66 ± 0.05
S1		20.1	>10
S2		17.8	>10
S3		18.0	>10
S4		19.0	>10
S5		18.2	>10
S6		17.5	>10
S7		18.4	>10
S8		19.0	>10
10		24.5	0.09 ± 0.01

[a] Compounds were detected on an RP-C18 column (4.6 mm x 250 mm), using eluents A (0.1% TFA in water) and B (0.08% TFA in 80% acetonitrile) at a flow rate of 0.5 mL/min. The gradient increased from 20% acetonitrile to 80% acetonitrile over 20 min followed by a 1 min 100% acetonitrile wash.
[b] Ki values were determined as described in Supplementary Methods.

Figure S5: HPLC analysis of *In situ* click formation of triazole **10** from biotin alkyne **2** and azide **9** a) in PBS buffer alone, b) in the presence of 2.5 μM wildtype *Sa*BPL, c) in the presence of 25 μM wildtype *Sa*BPL, d) in the presence of 2.5 μM mutant R122G *Sa*BPL and e) in the presence of 25 μM mutant R122G *Sa*BPL. The black arrows denote observed peaks that correspond to the retention time of an authentic sample of triazole **10**.

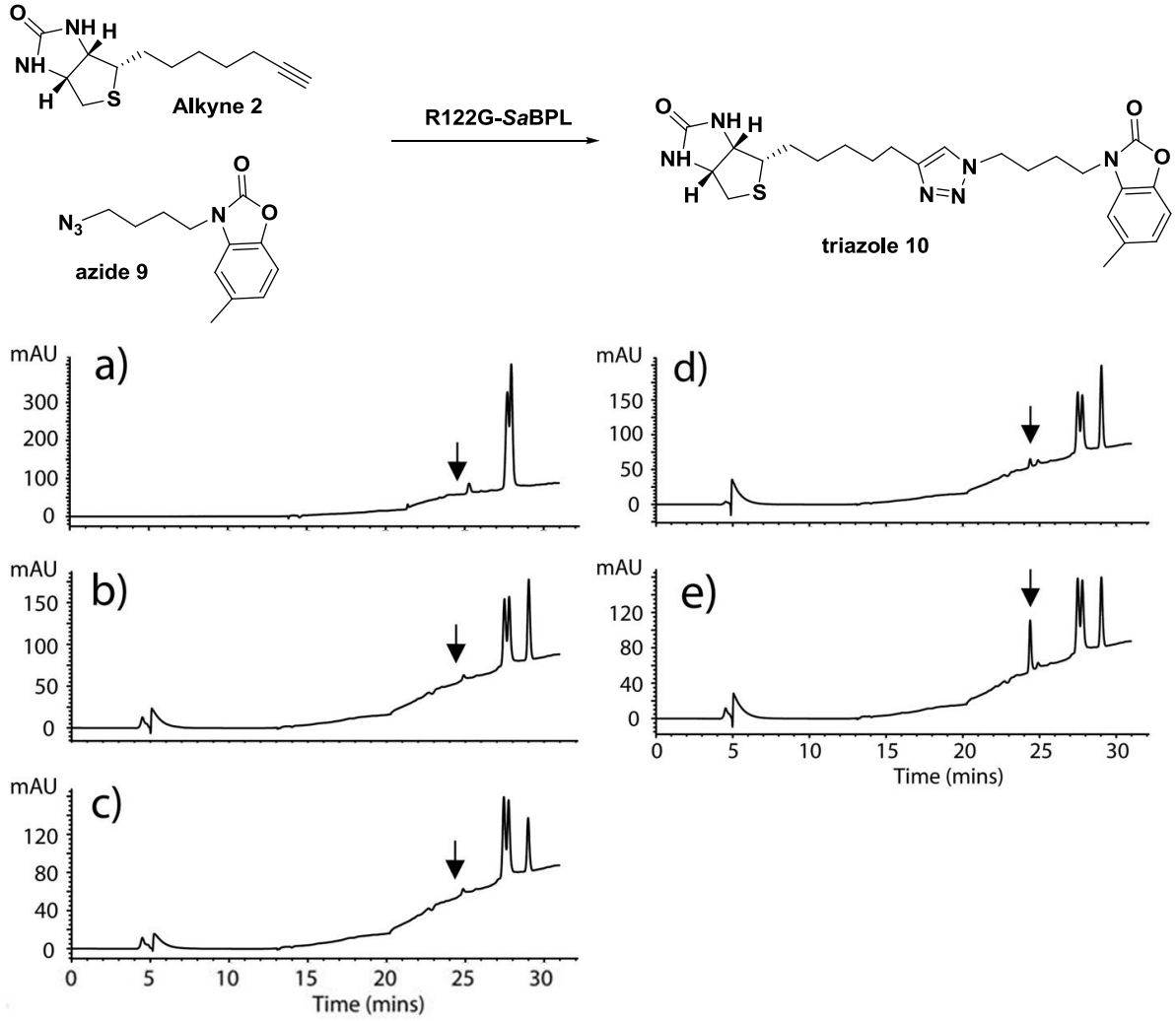


Table S3: X-ray data collection and refinement. Data collection and refinement statistics for holo *Sa*BPL structures with compound **8** bound.

	***Sa*BPL with 8**
Data collection[a]	
Space group	P 4(2) 2(1) 2
Cell dimensions	
a, *b*, *c* (Å)	94.6, 94.6, 130.7
α, β, γ (°)	90, 90, 90
Resolution (Å)	50.00 - 2.61 (2.68-2.61)
R_{sym} or R_{merge}	0.06 (0.40)
$I / \sigma I$	15.3 (3.54)
Completeness (%)	99.0 (94.3)
Redundancy	2.46 (2.47)
Refinement	
Resolution (Å)	50 - 2.60
No. reflections	125676
R_{work} / R_{free}	20.65/25.77
No. atoms	
Protein	2613
Ligand/ion	32
Water	147
B-factors	
Protein	52.4
Ligand/ion	37.9
Water	57.7
R.m.s. deviations	
Bond lengths (Å)	0.02
Bond angles (°)	1.9

[a] Values in parentheses refer to the highest resolution shell.

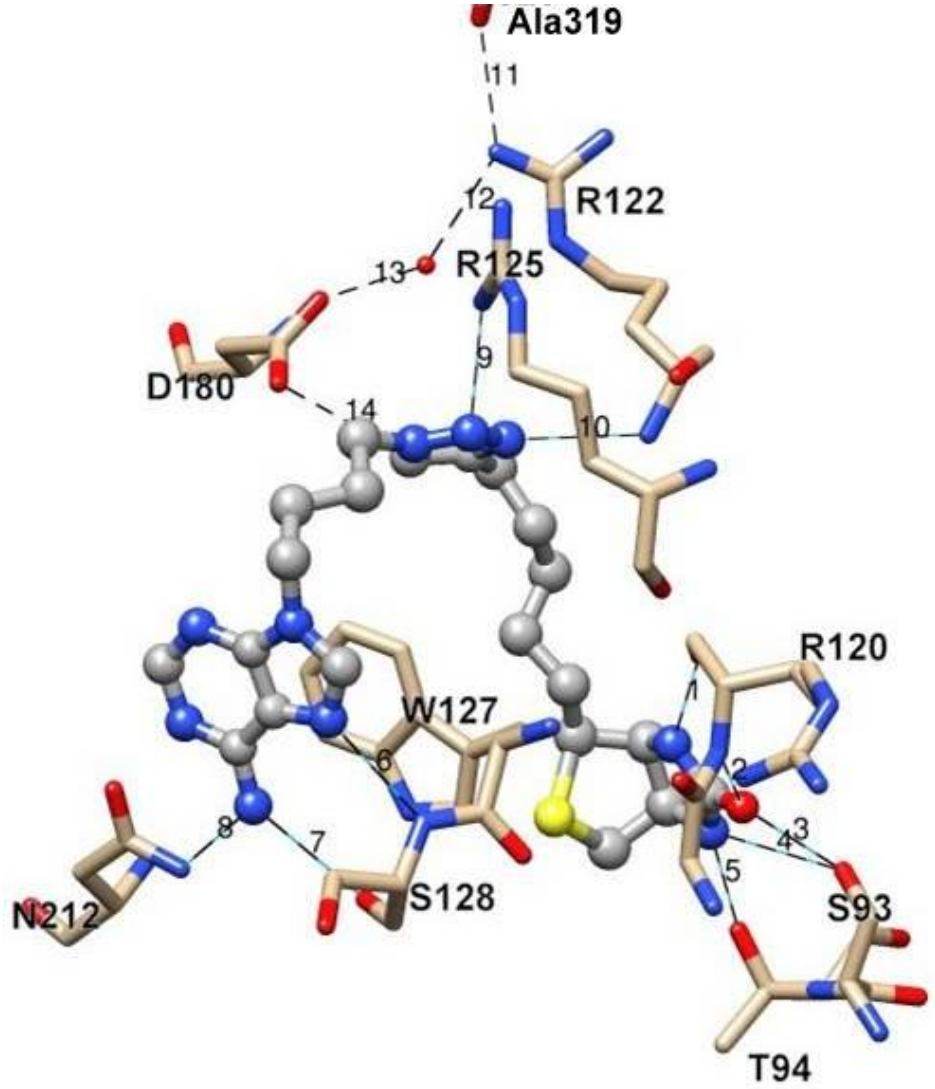
[b] $R_{merge} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$ where I is the intensity of individual reflections.

[c] $R_{pim} = \sum [1/(N-1)]^{1/2} \sum |I - \langle I \rangle| / \sum \langle I \rangle$ [1]

[d] $R_{factor} = \sum_h |F_o - F_c| / \sum_h |F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes for each reflection "h".

[e] R_{free} was calculated with 5% of the diffraction data selected randomly and excluded from refinement.

Figure S6: X-ray structure of triazole **8** bound to wild type *Sa*BPL with hydrogen bonding interactions between *Sa*BPL and triazole **8** indicated with black dashes.



Hydrogen bond ID	Distance (Å)	Hydrogen bond ID	Distance (Å)
1[a]	2.81	8[a]	2.94
2[a]	3.05	9[a]	2.92
3[a]	2.56	10[a]	3.00
4[a]	3.38	11[a]	2.89
5[a]	3.18	12[a]	2.94
6[a]	2.94	13[a]	2.74
7[a]	3.09	14[b]	3.43

[a] Hydrogen bonding interactions determined by UCSF Chimera 1.6.1 (build 35849) with distances determined between corresponding heteroatoms. [2]
[b] CH---O hydrogen bonding interactions similarly described by Brik and co-workers[3]

Protein Methods

In situ click experiment A: Preparation of triazole 1

The *in situ click* reaction of biotin alkyne **1** with azide **2** was performed at 37^{o}C for 48 hours in PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.46 mM KH_2PO_4, pH 7.4) using either wildtype or mutant *Sa*BPL (final enzyme concentration 2 µM). Control experiments were performed in parallel using either no enzyme or BSA in place of *Sa*BPL. Additionally, 500 µM of competitive inhibitor biotinol-5′-AMP was added to another control experiment with *Sa*BPL-R122G to address the specificity of the template guided reaction. The alkyne **2** and azide **3** were first solubilized in DMSO then added to all reactions to a final concentration of 500 µM each (final DMSO concentration 4% (v/v)). Aliquots of the reactions (50 µL) were diluted in 0.1% TFA (50 µL) and subjected to HPLC analysis. The analytes were fractionated using an analytical C18 reverse phase column (dimensions 4.6 mm x 250 mm), using eluents A (0.1% TFA in water) and B (0.08% TFA in 80% acetonitrile) at a flow rate of 0.5 mL/min. The gradient increased from 20% acetonitrile to 80% acetonitrile over 20 minutes followed by a 1 minute 100% acetonitrile wash. The system was set up for batch injections of 90 µL. Before each injection, the column was equilibrated with 100% Buffer A for 15 minutes. A duplicate sample was analysed using Orbitrap mass spectroscopy (Adelaide Proteomics Centre). To quantitate the yield of triazole product formed, the area under the peak on the UV trace was measured and compared against a standard curve established using known amounts of the synthesised triazole product.

In situ click experiment B: Library experiments

The *in situ* click reaction was performed essentially as described above with alkyne **2** (160 µM) reacted with a mixture azides **3–7** (final concentration of DMSO 4% (v/v)). HPLC of the reactions were performed as described above. LC-MS analysis was performed on an Agilent 1100 LC/MSD with a Bruker HCTultra PTM Discovery System (ES) eluting on a Agilent Poroshell 120 SB-C18 reverse-phase column (2.1 mm x 50 mm, 2.7 µm); 5 µL of sample was injected per run; flow rate at 0.2 mL/min; gradient elution (H_2O + 0.05% TFA)/(90% aqueous MeCN + 0.05% TFA) from 95:5 to 0:100 over 20 min followed by a wash of 5 min with 20:80 and a post-run time of 5 min using the starting solvent ratio. Detection was by electrospray

ionization with positive selected-ion monitoring tuned to the molecular masses of triazole **8** and **1** (M+H = 471 and M+H = 571 respectively). The cycloaddition products were identified by comparison of its retention time with those determined from analysis of the copper-catalyzed reaction and by its molecular weight.

In situ click experiment C: Preparation of Triazole 10

The *in situ click* reaction was performed at 37 ^{o}C for 48 hours in PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.46 mM KH_2PO_4, pH 7.4) using either wildtype or mutant *Sa*BPL (at final enzyme concentrations 2.5 μM). A second set of *in situ* click reactions were performed under the same conditions, except using 25 μM final concentrations of either wild type or mutant *SaBPL*. The alkyne **2** and azide **9** were first solubilized in DMSO then added to all reactions to a final concentration of 500 μM each (final DMSO concentration 4% (v/v)). Control experiments were performed in parallel and as detailed in '*in situ* click experiment 1'. Aliquots of the reactions (50 μL) were diluted in 0.1% TFA (50 μL) and subjected to HPLC analysis following the conditions described in '*in situ* click experiment 1'. The cycloaddition product **10** was identified by comparing its retention time with that of an authentic sample of **10**. Duplicate samples were analysed using Orbitrap mass spectroscopy (Adelaide Proteomics Centre). Detection of triazole **10** was confirmed with mutant *Sa*BPL at 25 μM concentration (found: 485.23 m/z, expected (M+H): 485.23), whilst triazole **10** was not detected by mass spectroscopy in samples containing wildtype enzyme and negative controls.

Protein expression and purification

DNA constructs required for recombinant expression of *Sa*BPL-H_6 in *E. coli* were described previously[4]. The R122G mutation was obtained using the QuikChange™ mutagenesis kit (Stratagene) with plasmid pGEM-(*Sa*BPL-H6)[4] and oligonucleotides B227 (5′-CGAAAGGTCGTGGT**GGT**TTTAATAGACATTGG-3′) and B228 (5′-CCAATGTCTATTAAA**ACC**ACCACGACCTTTCG-3′). The expression and purification of wildtype and the R122G *Sa*BPL-H_6 were performed using the same strategy as previously described by Pendini[4].

Surface Plasmon Resonance

SPR was performed using a Biacore™ T100 instrument (GE Healthcare). BPL was immobilised on a CM5 sensor chip by standard amine coupling chemistry. The carboxymethyl groups on the chip were activated by the addition of *N*-ethyl-*N´*-(3-diethylamino-propyl) carbodiimide and *N*-hydroxysuccinimide. BPL (120 µg/mL) in 10 mM sodium acetate buffer (pH 5.8) was coupled onto the surface, and 1 M ethanolamine hydrochloride was injected to block any unreacted sites. Typically, 6,500 resonance units of BPL were immobilised on the sensor chip. One channel was left blank which was subtracted from sample channel(s) to allow analysis methods to distinguish between actual and non-specific binding. Experiment was conducted at 25^{o} C at a flow rate of 30 µL/minute with a running buffer containing 10 mM HEPES pH 7.4, 150 mM NaCl. Zero concentration samples were used as blanks. The results for binding of biotin and AMP to both wild-type enzyme and *Sa*BPL-R122G showed fast on and off rates outside the range of kinetic quantification, so K_D values were estimated using steady-state affinity. A binding curve for AMP was obtained by injecting saturating concentrations of biotin alkyne **1** (100 µM) prior to injection of AMP. Binding of compounds to wild-type *Sa*BPL were fitted to a 1:1 ligand binding model. K_D values for binding of compounds to *Sa*BPL-R122G were estimated using steady-state affinity. Buffers for experiments conducted with non-water soluble compounds contained 4% (v/v) DMSO.

Circular Dichroism Analysis

Far–UV CD spectra were recorded at 20^{o} C using a Jasco CD J-185 spectrophotometer. Purified wild-type and mutant *Sa*BPL were dialysed overnight against the assay buffer (10 mM $NaPO_4$ pH 8.0) at 4^{o} C prior to analysis. Final concentrations of the proteins used were 0.25 mg/mL (6.5 µM). CD spectra were recorded from 185 nm to 300 nm at 0.2 nm increments, using a 1 mm path length quartz cell. The reported spectra are the average of five scans that were corrected for buffer blanks.

In vitro biotinylation assays

Quantitation of BPL catalysed ^{3}H-biotin incorporation into the biotin domain substrate was performed as previously described [5,6]. The IC_{50} value of each compound was determined from a dose-response curve by varying the concentration of the inhibitor under the same enzyme concentration. The data was

analysed with GraphPad Prism software using a non-linear fit of $\log_{10}$ (inhibitor) vs. normalized response. The K_i, the absolute inhibition constant for a compound, was determined using Eq1 [7]:

Eq1. $K_i = IC_{50} / (1 + [S] / K_m)$

where K_m is the affinity of the substrate for the enzyme ([biotin] =1.01 µM, Supplementary Table S2) and [S] is the substrate concentration ([biotin] =5 µM). The mode of inhibition was investigated by varying the concentrations of inhibitor alongside varying the concentrations of ^{3}H-biotin. The data was plotted as double reciprocal plots and assessed using Lineweaver-Burk analysis. Data reported here are the means of three independent assays (n = 3) ± standard error of the mean. Statistical analysis between two data sets was performed using a two-tailed unpaired t test using GraphPad Prism. See Supplementary Table S2 for enzyme inhibition results of **1**, **S1 – S8**, **8**).

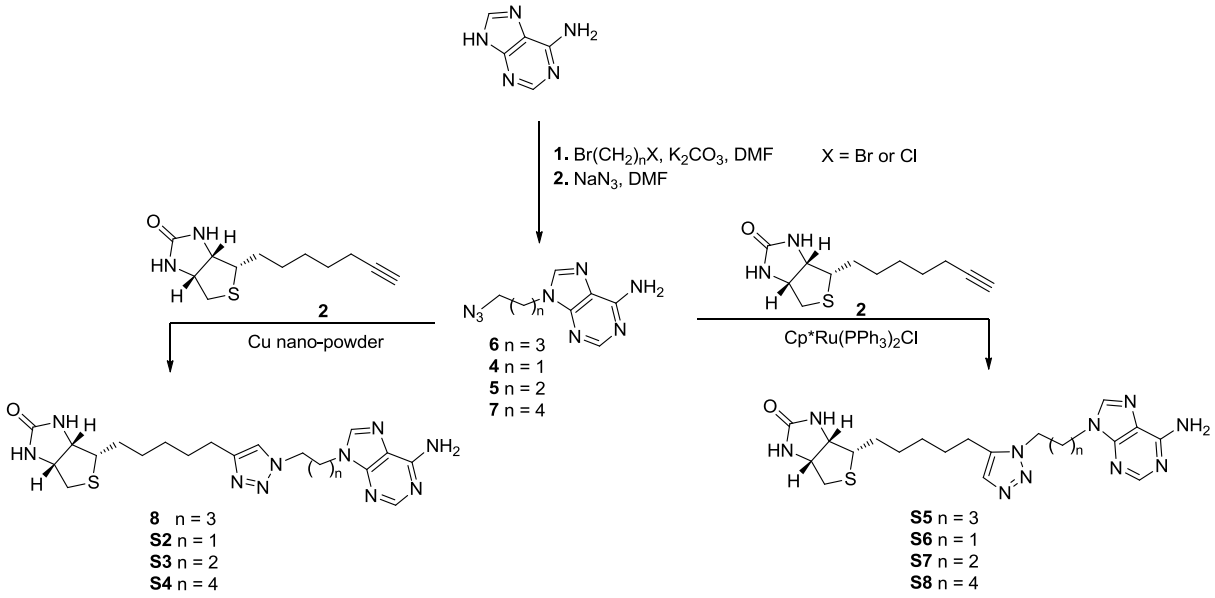
X-ray crystallography

Apo-*Sa*BPL was buffer exchanged into 50 mM Tris HCl pH 7.5, 50 mM NaCl, 1 mM DTT and 5% (v/v) glycerol, and concentrated to 5 mg/mL. Each compound was then added to BPL in a 10:1 molar ratio. The complex was crystallized using the hanging drop method at 4°C in 8 – 12% Peg 8000 in 0.1 M Tris pH 7.5 or 8.0, and 10% (v/v) glycerol as the reservoir. A single crystal was picked using a Hampton cryo-loop and streaked through cryoprotectant containing 25% (v/v) glycerol in the reservoir buffer prior to data collection. X-ray diffraction data was collected at the macromolecular crystallography beamline at the Australian Synchrotron using an ADSC Quantum 315r Detector. 90 images were collected for 1 second each at an oscillation angle of 1° for each frame. Data was integrated using XDS, and refined using the CCP4 suite of programs. PDB and cif files for the compounds were obtained using the PRODRG web interface. The models were built using cycles of manual modelling using COOT and refinement with REFMAC. The quality of the final models was evaluated using MOLPROBITY. Composite omit maps were inspected for each crystal structure and statistics for the data and refinement reported (Supplementary Table S3).

Synthetic chemistry methods

All reagents were from standard commercial sources and of reagent grade or as specified. Solvents were from standard commercial sources. Reactions were monitored by ascending TLC using precoated plates (silica gel 60 F_{254}, 250 µm, Merck, Darmstadt, Germany), spots were visualised under ultraviolet light at 254 nm and with either sulphuric acid-vanillin spray, potassium permanganate dip or Hanessian's stain. Column chromatography was performed with silica gel (40-63 µm 60 Å, Davisil, Grace, Germany). Melting points were recorded uncorrected on a Reichert Thermovar Kofler microscope. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Gemini 2000 (300 MHz) or a Varian Inova 600 MHz. Chemical shifts are given in ppm (δ) relative to the residue signals, which in the case of DMSO-d_6 were 2.50 ppm for ^{1}H and 39.50 ppm for ^{13}C, $CDCl_3$ were 7.26 ppm for ^{1}H and 77.23 ppm for ^{13}C. Structural assignment was confirmed with COSY, ROESY, HMQC and HMBC. High-resolution mass spectra (HRMS) were recorded on a Thermo Fisher Scientific LTQ orbitrap FT MS equipment (Δ < 2 ppm) and Brucker micrO TOF-Q. Compounds **1 – 3**[8], **S1**[8]**, 9**[8]**, 10**[8] and Cp*Ru(PPh$_3$)$_2$[9] catalyst were synthesised according to literature procedures.

Synthetic scheme for triazoles 8 and S2 – S8



General procedure for preparation of adenine azide 4 - 7

To a suspension of adenine (1 equiv.) and K_2CO_3 (2 equiv.) in DMF was added the appropriate di-halo-alkane (1.5 equiv) and the mixture stirred under nitrogen atmosphere at 50 ^oC for 8 h. The reaction mixture was cooled, diluted with DCM and washed with 0.5 M aqueous HCl, water and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to give crude alkylated adenine. See specific compounds for further details. A solution of alkylated adenine (1 equiv.), NaN_3 (1.5 equiv.) in DMF was stirred under a nitrogen atmosphere at 50 ^oC for 8h. The reaction was cooled, diluted with ethyl acetate and washed with water and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo* to give the crude azide. See specific compounds for further details.

9-(2-Azidoethyl)purin-6-amine (4)

Adenine (500 mg, 3.70 mmol) was alkylated with 1,2-dibromoethane (1370 mg, 7.41 mmol) in accordance to 'general procedure for azide' and the product purified by silica gel chromatography eluting with 10% methanol in dichloromethane to give the bromide as a white solid (640 mg, 72%). Data consistent with literature[10]. **FT-IR** (ATR) $\mathbf{\nu_{max}}$: 3280, 3101, 1667, 1596, 655 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d^6): δ 8.17 (1H, s), 8.15 (1H, s), 7.27 (2H, bs), 4.56 (2H, t, *J* = 6.0 Hz), 3.94 (2H, t, *J* = 6.0 Hz); **^{13}C NMR** (75 MHz; DMSO-d^6): δ 156.0, 152.5, 149.5, 140.9, 118.7, 44.64, 31.6.

The bromide (150 mg, 0.62 mmol) was treated according to 'general procedure for adenine azide' and the product was purified by silica gel chromatography eluting with 8% methanol in dichloromethane to give **4** as a white solid (111 mg, 86%). **FT-IR** (ATR) $\mathbf{\nu_{max}}$: 3268, 3093, 2097, 1674, 1599, 1574 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d_6): δ 8.16 (1H, s) 8.15 (1H, s), 7.25 (2H, bs), 4.34 (2H, t, *J* = 6.0 Hz), 3.81 (2H, t, *J* = 6.0 Hz); **^{13}C NMR** (75 MHz; 2% CD_3OD, $CDCl_3$): δ 156.0, 152.5, 149.6, 140.9, 118.7, 49.7, 42.4; **HRMS** calcd. for (M + H) $C_7H_9N_8$: requires 205.0950, found 205.0943.

9-(3-Azidopropyl)purin-6-amine (5)[11]

Adenine (250 mg, 1.85 mmol) was alkylated with 1,3-dibromopropane (737 mg, 3.70 mmol) in accordance with the 'general procedure for azide' and the product purified by silica gel chromatography eluting with 6% methanol in dichloromethane to

give the bromide as a white solid (178 mg, 38%). Characterisation data consistent with literature.[12] **FT-IR** (ATR)ν_{max}: 3288, 3108, 1661, 1598, 796 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d^6): δ 8.14 (2H, s) 7.22 (2H, bs), 4.26 (2H, t, *J* = 6.6 Hz), 3.50 (2H, t, *J* = 6.3 Hz), 2.37 (2H, m); **^{13}C NMR** (75 MHz; DMSO-d^6): δ 155.9, 152.4, 149.5, 140.8, 118.8, 41.7, 32.1, 31.3.

The bromide (161 mg, 0.63 mmol) was treated according to 'general procedure for adenine azide' and the product purified by silica gel chromatography eluting with 5% methanol in dichloromethane to give **5**[11] as a white solid (104 mg, 76%). **FT-IR** (ATR) ν_{max}: 3298, 3131, 2930, 2102, 1662, 1597 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d_6): δ 8.14 (1H, s), 8.13 (1H, s), 7.21 (2H, bs), 4.20 (2H, t, *J* = 7.2 Hz), 3.38 (2H, t, *J* = 7.2 Hz), 2.06 (2H, t, *J* = 7.2, 7.2 Hz); **^{13}C NMR** (75 MHz; 2% CD_3OD in $CDCl_3$): δ 155.9, 152.4, 149.6, 140.8, 118.7, 48.1, 40.5, 28.6.

9-(4-Azidobutyl)purin-6-amine (6)

Adenine (251 mg, 1.85 mmol) was alkylated with 1,4-dibromobutane (789 mg, 3.70 mmol) in accordance with the 'general procedure for azide' and the product purified by silica gel chromatography eluting with 6% methanol in dichloromethane to give the bromide as a white solid (427 mg, 86%). Characterisation data was consistent with literature report.[10] **FT-IR** (ATR) ν_{max}: 3202, 3108, 2868, 1668, 1605, 1228 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d_6): δ 8.15 (1H, s), 8.14 (1H, s), 7.21 (2H, bs), 4.18 (2H, t, *J* = 6.6 Hz), 3.55 (2H, t, *J* = 6.3 Hz), 1.88-1.97 (2H, m), 1.70-1.79 (2H, m); **^{13}C NMR** (75 MHz; 2% CD_3OD, $CDCl_3$): δ 155.9, 152.4, 149.6, 140.8, 118.7, 42.0, 34.3, 29.3, 28.1.

The bromide (150 mg, 0.56 mmol) was treated according to 'general procedure for adenine azide' and the product purified by silica gel chromatography eluting with 5% methanol in dichloromethane to give **6** as a white solid (121 mg, 92%). **^{1}H NMR** (300 MHz; 2% CD_3OD, $CDCl_3$): δ 8.30 (1H, s) 7.84 (1H, s), 4.25 (2H, t, *J* = 7.2 Hz), 3.36 (2H, t, *J* = 6.6 Hz), 1.95-2.04 (2H, m), 1.58-1.68 (2H, m); **^{13}C NMR** (75 MHz; DMSO-d_6): δ 155.9, 152.4, 149.9, 140.8, 118.6, 50.1, 42.4, 26.7, 25.5.

9-(5-Azidopentyl)purin-6-amine (7)

Adenine (250 mg, 1.85 mmol) was alkylated with 1-bromo-5-chloropentane (679 mg, 3.70 mmol) in accordance with 'general procedure for azide' and the

product purified by silica gel chromatography eluting with 5% methanol in dichloromethane to give the chloride as a white solid (380 mg, 89%). **FT-IR** (ATR) ν_{max}: 3215, 3109, 1608, 1570, 1300 cm^{-1}; **^{1}H NMR** (200 MHz; 5% CD_3OD, $CDCl_3$): δ 7.92 (1H, s), 7.49 (1H, s), 6.19 (2H, bs), 3.85 (2H, t, *J* = 7.2 Hz), 3.15 (2H, t, *J* = 6.4 Hz), 1.64-1.85 (2H, m), 1.32-1.45 (2H, m); **^{13}C NMR** (75 MHz; 2% CD_3OD in $CDCl_3$): δ 155.6, 152.8, 149.8, 140.5, 119.0, 44.6, 43.62, 31.9, 29.1, 24.1.

The chloride (110 mg, 0.46 mmol) was treated according to 'general procedure for adenine azide' and the product purified by silica gel chromatography eluting with 5% methanol in dichloromethane to give **7** as a white solid (95 mg, 84%). **FT-IR** (ATR) ν_{max}: 3211, 3108, 2942, 2089, 1667, 1607, 1572 cm^{-1}; **^{1}H NMR** (300 MHz; DMSO-d_6): δ 8.14 (1H, s), 8.13 (1H, s), 7.19 (2H, bs), 4.14 (2H, t, *J* = 7.2 Hz), 3.30(2H, t, *J* = 6.4 Hz), 1.76-1.88 (2H, m), 1.50-1.60 (2H, m), 1.21-1.32 (2H, m); **^{13}C NMR** (75 MHz; 2% CD_3OD in $CDCl_3$): δ 155.9, 152.9, 149.9, 140.2, 119.3, 51.1, 43.7, 29.6, 28.3, 23.8; **HRMS** calcd. for (M + H) $C_{10}H_{15}N_{8:}$ requires 247.1420, found 247.1420.

General CuAAC

A suspension of biotin alkyne **2** (1 equiv), azide (1 equiv) and copper nano-powder (2 mg) in acetonitrile or methanol was sonicated for 15 min. The reaction mixture was then stirred at ambient temperature for 4 h, concentrated *in vacuo* and the residue purified by silica gel-based chromatography. See specific compounds for further details.

(3a*S*,6a*R*)-4-[5-[1-[4-(6-Aminopurin-9-yl)butyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (8)

Biotin alkyne **2** (31 mg, 0.13 mmol) was reacted with azide **6** (36 mg, 0.14 mmol) and Cu nanopowder (2 mg, 0.03 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica eluting with 8% methanol in dichloromethane to give **8** as a crystalline white solid (46 mg, 73%). **MP:** 106 - 108 ^oC; **^{1}H NMR** (600 MHz; DMSO-d_6): δ 8.12 (1H, s), 8.11 (1H, s, Ar**H**), 7.86 (1H, s), 7.20 (2H, bs), 6.52 (1H, bs), 6.40 (1H, bs), 4.34-4.38 (3H, m), 4.34-4.38 (3H, m), 3.12-3.16 (m, 1H), 3.12-3.16 (1H, dd, *J* = 4.8, 12.0 Hz), 2.59-2.64 (3H, m), 1.76-1.80 (4H, m), 1.28-1.67 (8H, m); **^{13}C NMR** (150 MHz; DMSO-d_6): 162.8, 156.0, 152.4, 149.5, 146.8, 140.8, 121.8, 118.7, 61.2, 59.2, 55.6, 48.5, 42.2, 40.0, 28.8,

28.6, 28.5, 28.3, 27.0, 26.7, 25.1; **HRMS** calcd. for (M + H) $C_{21}H_{31}N_{10}OS$: requires 471.2403, found 471.2450.

(3a*S*,6a*R*)-4-[5-[1-[2-(6-Aminopurin-9-yl)ethyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S2)

Biotin alkyne **2** (15 mg, 0.063 mmol) was reacted with azide **4** (15 mg, 0.069) and Cu nanopowder (1 mg, 0.01 mmol) following 'general procedure CuAAC'. The residue was purified by silica gel chromatography on silica eluting with18% methanol in dichloromethane to give **S2** as a crystalline off white solid (12 mg, 41%). **MP:** 222 - 224 ^{o}C; **^{1}H NMR** (300 MHz; DMSO-d^{6}): δ 8.09 (1H, s), 7.75 (1H, s), 7.69 (1H, s), 7.29 (2H, bs), 6.69 (1H, bs), 6.40 (1H, bs), 4.79 (2H, t, *J* = 5.4 Hz), 4.60 (2H, t, *J* = 5.4 Hz), 4.29-4.33 (1H, m), 4.12-4.18 (1H, m), 3.06-3.12 (1H, m), 2.82 (1H, dd, *J* = 4.8, 12.3 Hz), 2.56-2.66 (3H, m), 1.14-1.67 (8H, m); **^{13}C NMR** (150 MHz; DMSO-d^{6}): δ 162.9, 155.9, 152.4, 149.4, 146.9, 140.4, 121.9, 118.4, 61.1, 59.2, 55.6, 48.5, 43.0, 39.9, 28.7, 28.4, 28.3, 28.2, 24.8; **HRMS** calcd. for (M + H) $C_{19}H_{27}N_{10}OS$: requires 443.2090, found 443.2094.

(3a*S*,6a*R*)-4-[5-[1-[3-(6-Amino-4,5-dihydropurin-9-yl)propyl]triazol-4-yl]pentyl]-1,3,3a,4,6, 6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S3)

Biotin alkyne **2** (15 mg, 0.063 mmol) was reacted with azide **5** (16 mg, 0.069 mmol) and Cu nanopowder (1 mg, 0.016 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica eluting with 12% methanol in dichloromethane to give **S3** as a crystalline white solid (19 mg, 65%). **MP:** 156 - 157 ^{o}C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ 8.13 (1H, s), 8.11 (1H, s), 7.23 (2H, bs), 6.47 (2H, bs), 6.37 (1H, bs), 4.35-4.34 (3H, m), 4.11-4.18 (3H, m), 3.06-3.13 (1H, m), 2.81 (1H, dd, *J* = 4.8, 13.2 Hz), 2.54 -2.61 (3H, m), 2.32-2.42 (2H, m), 1.33 – 1.59 (8H, m); **^{13}C NMR** (150 MHz; 5% CD_3OD, $CDCl_3$): δ 164.1, 155.7, 152.8, 149.5, 148.4, 140.9, 121.4, 119.1, 62.2, 60.1, 55.8, 46.9, 41.0, 40.4, 30.0, 29.0, 28.9, 28.8, 28.5, 25.2; **HRMS** calcd. for (M + H) $C_{20}H_{29}N_{10}OS$: requires 457.2247, found 457.2269.

(3a*S*,6a*R*)-4-[5-[1-[5-(6-Aminopurin-9-yl)pentyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S4)

Biotin alkyne **2** (21 mg, 0.088 mmol) was reacted with azide **7** (25 mg, 0.10 mmol) and Cu nanopowder (1 mg, 0.02 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica eluting with 8% methanol in dichloromethane to give **S4** as a crystalline white solid (33 mg, 74%). **MP**: 101 - 103 ^{o}C; **^{1}H NMR** (300 MHz, 1% CD_3OD $CDCl_3$): δ 8.27 (1H, s), 7.73 (1H, s), 7.20 (1H, s), 4.48-4.53 (1H, dd, *J* = 4.8, 7.8 Hz), 4.28-4.36 (3H, m), 4.13-4.19 (2H, m), 3.14-3.20 (1H, m), 2.94 (1H, dd, *J* = 4.8, 12.7 Hz), 2.72 (1H, d, *J* = 12.7), 2.66 (2H, t, *J* = 7.8 Hz), 1.26-1.96 (14H, m); **^{13}C NMR** (75 MHz; DMSO-d^{6}): δ 162.8, 155.9, 152.3, 149.5, 140.8, 137.1, 131.3, 118.6, 61.0, 59.2, 55.5, 46.5, 42.6, 29.0, 28.89, 28.61, 28.27, 28.20, 27.39, 23.01, 22.05; **HRMS** calcd. for (M+ H) $C_{22}H_{33}N_{10}O_4S$: requires 485.2560, found 485.2536.

General RuAAC

A solution of alkyne **2** (1 equiv.), azide (1 equiv.) and $Cp^{*}Ru(PPh_3)_2$ (0.1 equiv) in 1:1 DMF/THF was stirred at 80 ^{o}C for 4 h under a nitrogen atmosphere, concentrated *in vacuo* and the residue purified by silica gel-based chromatography. See specific compounds for further details.

(3a*S*,4*S*,6a*R*)-4-[5-[1-[4-(6-Aminopurin-9-yl)butyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S5)

Biotin alkyne **2** (41 mg, 0.17 mmol) was reacted with azide **6** (47 mg, 0.19 mmol) and $Cp^{*}Ru(PPh_3)_2Cl$ (27 mg, 0.03 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica eluting with 7% methanol in dichloromethane to give **S5** as a white solid (46 mg, 55%). **MP:** 214 – 216 ^{o}C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ 8.13 (1H, s), 8.12 (1H, s), 7.48 (1H, s), 7.24 (2H, bs), 6.56 (1H, bs), 6.40 (1H, bs), 4.25-4.34 (3H, m), 4.12-4.20 (3H, m), 3.07-3.13 (1H, m), 2.83 (1H, dd, *J* = 5.4, 12.9 Hz), 2.54-2.61 (3H, m), 1.23-1.85 (12H, m);**^{13}C NMR** (150 MHz; DMSO-d^{6}): 163.9, 155.6, 152.6, 149.4, 140.1, 137.1, 131.7, 118.8, 72.2, 62.0, 59.9, 46.7, 43.0, 40.3, 29.0, 28.7, 28.5, 27.8, 27.0, 26.7, 22.8; **HRMS** calcd. for (M + H) $C_{21}H_{31}N_{10}OS$: requires 471.2403, found 471.2440.

(3a*S*,4*S*,6a*R*)-4-[5-[3-[2-(6-Aminopurin-9-yl)ethyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S6)

Biotin alkyne **2** (15 mg, 0.045 mmol) was reacted with azide **4** (16 mg, 0.05 mg) and Cp*Ru(PPh_3)$_2$Cl (1 mg, 0.02 mmol) following 'general procedure RuAAC'. The crude material was purified by flash chromatography on silica eluting with 15% MeOH in DCM to give **S6** as a white solid, (13 mg, 41%). **^{1}H NMR (**300 MHz, DMSO-d_6): δ 8.17 (1H, s), 7.81 (1H, s), 7.48 (1H, s), 7.32 (2H, bs), 6.53 (1H, bs), 6.43(1H, bs), 4.77-4.81 (2H, m), 4.64-4.68 (2H, m), 4.34-4.39 (1H, m), 4.17-4.21 (1H, m), 3.14-3.22 (1H, m), 2.88 (1H, dd, *J* = 5.4, 12.6 Hz), 2.63 (1H, d, *J* = 12.6 Hz), 2.35 (2H, t, *J* = 7.8 Hz), 1.20-1.66 (8H, m); **^{13}C NMR (**150 MHz, DMSO-d_6): δ 162.7, 155.9, 152.2, 149.5, 137.9, 134.9, 131.4, 118.6, 61.1, 59.2, 55.5, 45.9, 42.9, 40.0, 28.5, 28.2, 28.1, 27.1, 21.7; **HRMS** calcd. for (M + H) $C_{19}H_{27}N_{10}OS$: requires 443.2090, found 443.2127.

(3a*S*,4*S*,6a*R*)-4-[5-[3-[3-(6-Aminopurin-9-yl)propyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S7)

Biotin alkyne **2** (24 mg, 0.11 mmol) was reacted with azide **5** (26 mg, 0.11 mmol) and Cp*Ru(PPh_3)$_2$Cl (16 mg, 0.02 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica eluting with 10% MeOH in DCM to give **S7** as a white solid, (25 mg, 52%). **MP:** 204 - 206 ^{o}C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ 8.13 (1H, s), 8.12 (1H, s), 7.49 (1H, s), 7.22 (2H, bs), 6.52 (1H, bs), 6.38 (1H, bs), 4.09-4.32 (6H, m), 3.05-3.11 (1H, m), 2.82 (1H, dd, *J* = 7.8, 12.6 Hz), 2.57 (1H, d, *J* = 12.6 Hz), 2.52 (2H, t, *J* = 7.2 Hz), 2.33-2.42 (2H, m), 1.22-1.61 (10H, m); **^{13}C NMR** (150 MHz; DMSO-d^6): δ 162.7, 155.9, 152.7, 149.5, 141.1, 137.2, 131.2, 118.8, 61.0, 59.2, 55.4, 44.2, 40.5, 29.4, 28.6, 28.2, 27.3, 22.0; **HRMS** calcd. for (M + H) $C_{20}H_{29}N_{10}OS$: requires 457.2247, found 457.2269.

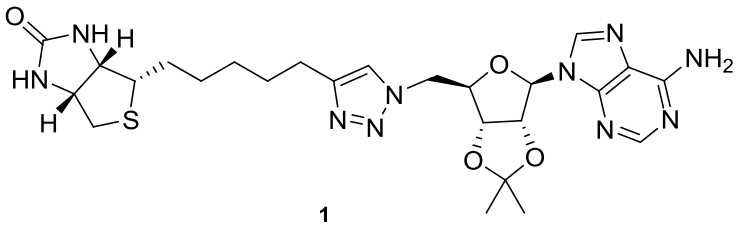
(3a*S*,6a*R*)-4-[5-[3-[5-(6-Aminopurin-9-yl)pentyl]triazol-4-yl]pentyl]-1,3,3a,4,6,6a-hexahydrothieno[3,4-*d*]imidazol-2-one (S8)

Biotin alkyne **2** (32 mg, 0.13 mmol) was reacted with azide **7** (39 mg, 0.15 mmol) and Cp*Ru(PPh_3)$_2$Cl (21 mg, 0.027 mmol) following 'general procedure CuAAC'. The crude material was purified by flash chromatography on silica using 5% methanol in dichloromethane to give **S8** as a white solid (32 mg, 48%). **MP:** 113 -

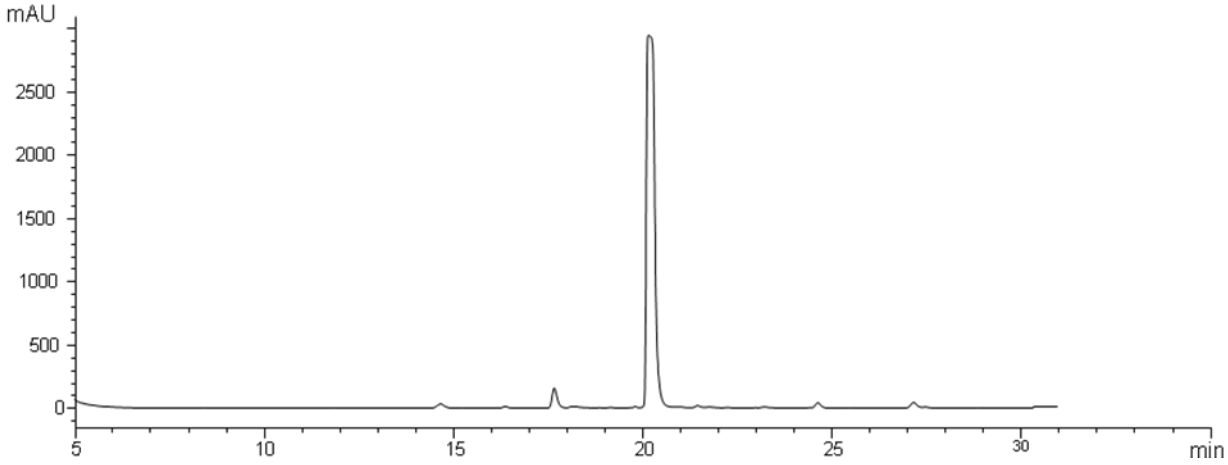
116 ^{o}C; **^{1}H NMR** (300 MHz; DMSO-d^{6}): δ 8.34 (0.5 H, s), 8.12 (1H, s), 8.11 (1H, s), 7.77 (1H, s), 7.20 (1.5H, bs), 6.49 (1H, bs), 6.38 (1H, bs), 4.24-4.32 (3H, m), 4.09-4.14 (3H, m), 3.06-3.17 (1H, m), 2.83 (1H, dd, *J* = 5.1, 12.3 Hz), 2.54-2.59 (3H, m,), 1.77-1.86 (4H, m), 1.13-1.57 (10H, m); **^{13}C NMR** (75 MHz; DMSO-d^{6}): δ 162.8, 156.2, 152.5, 149.5, 146.8, 140.8, 121.6, 118.9, 79.2, 62.4, 61.1, 55.6, 48.9, 44.5, 29.1, 28.8, 28.7, 28.6, 28.4, 28.2, 25.0, 22.9; **HRMS** calcd. for (M+ H) $C_{22}H_{33}N_{10}O_{4}S$: requires 485.2560, found 485.2559.

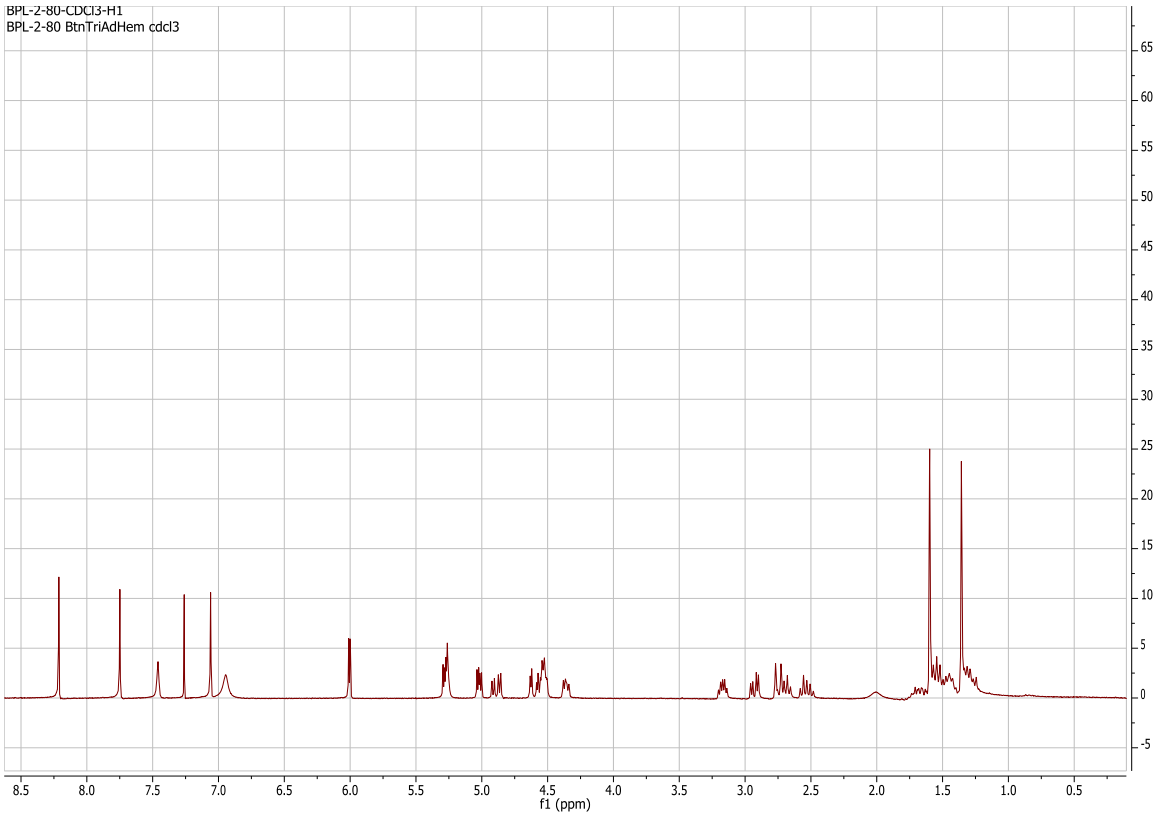
^{1}H NMR spectra and HPLC traces of compounds 1, 3-10 and S2 – S8

Compound 1

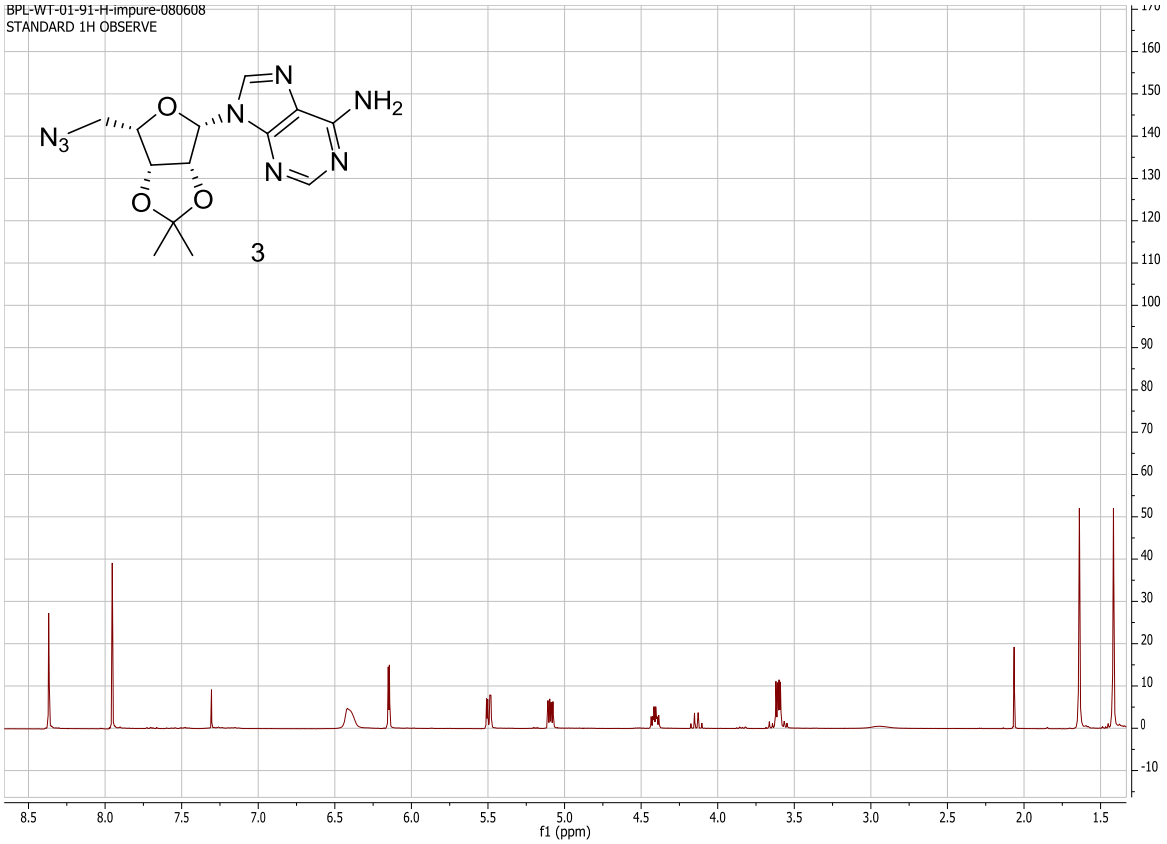


1

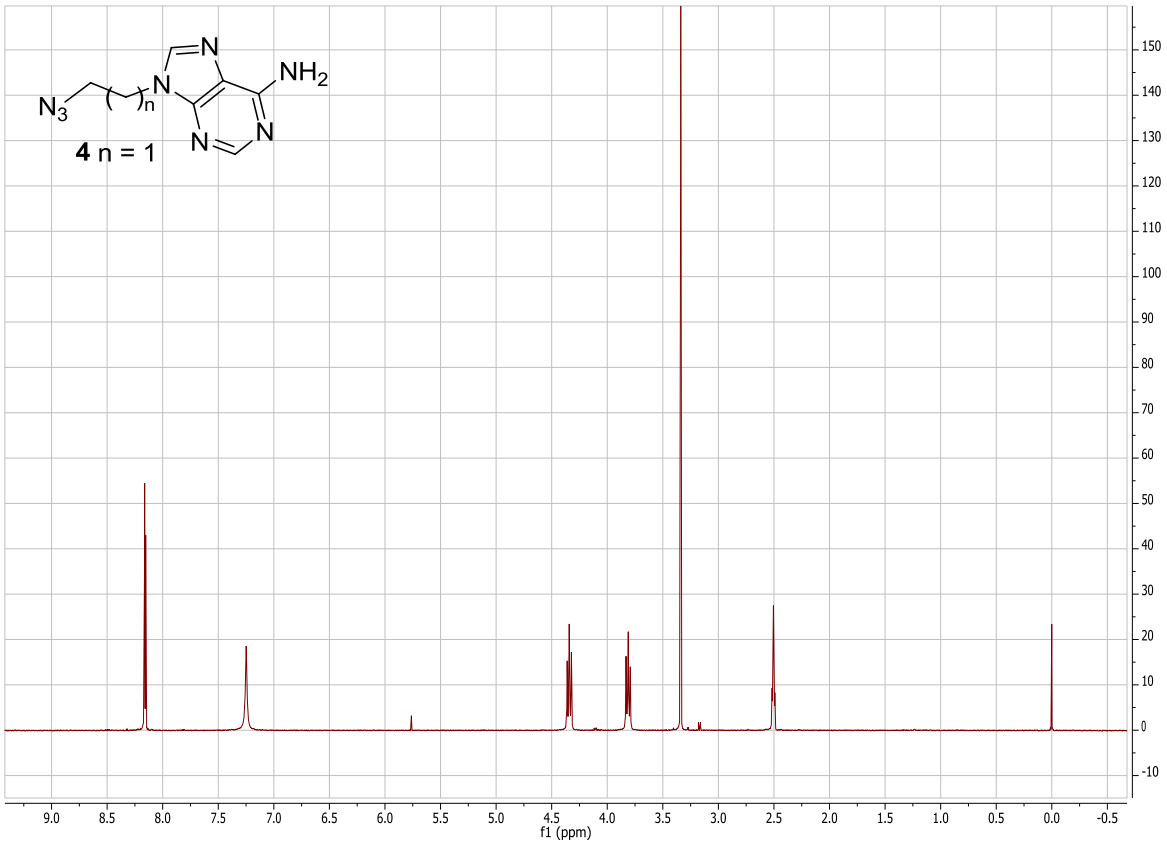




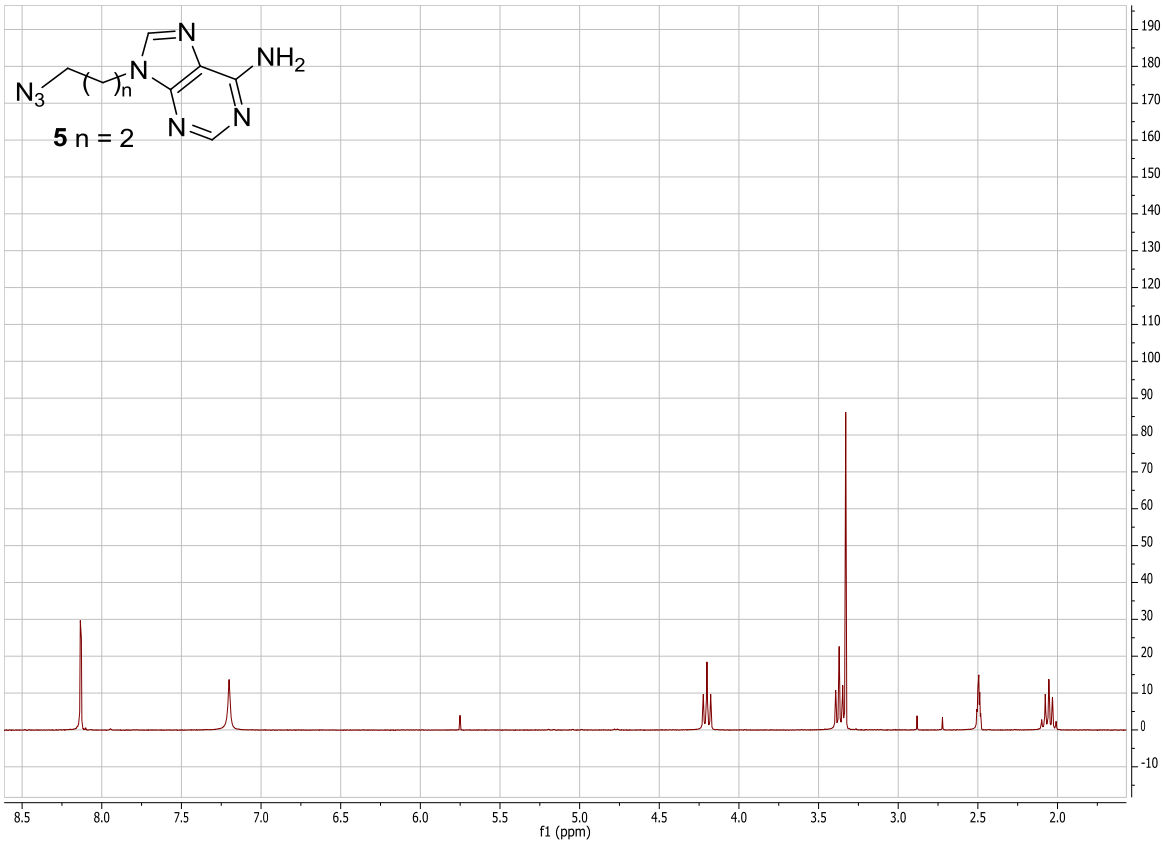
Compound 3



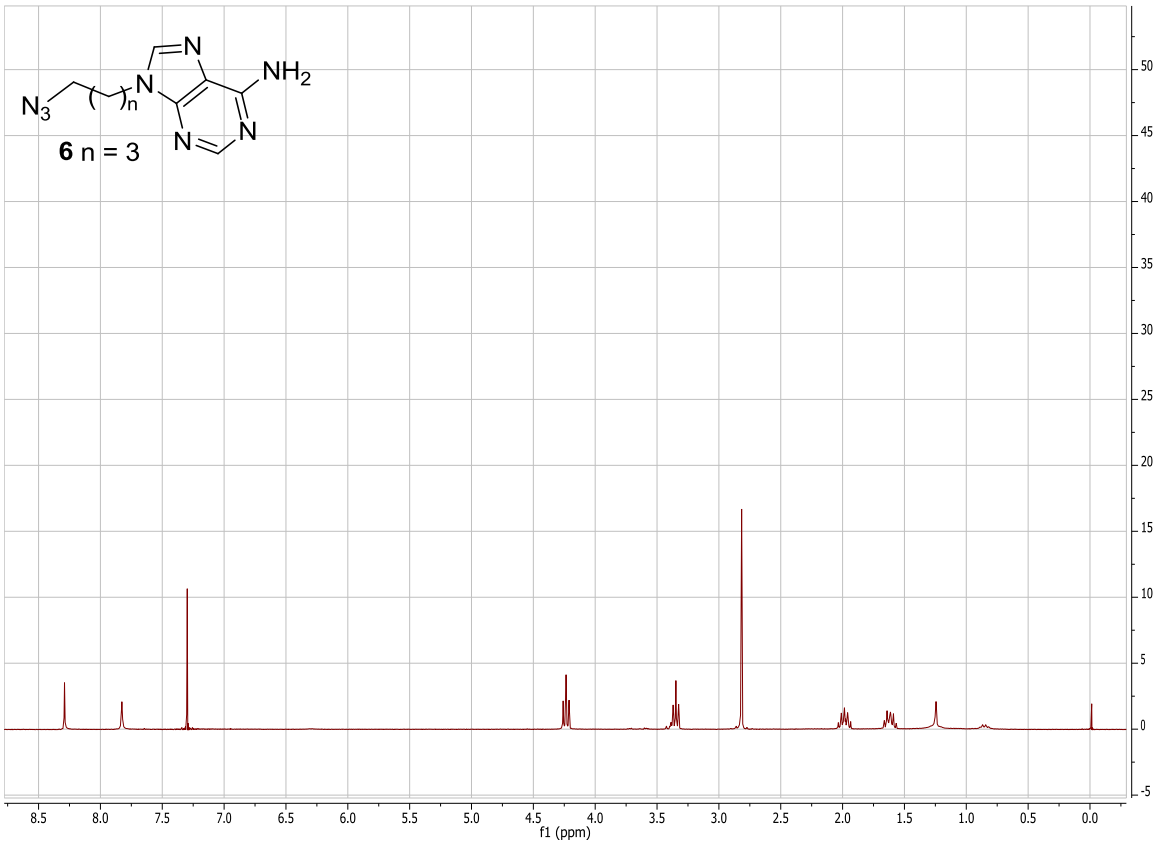
Compound 4



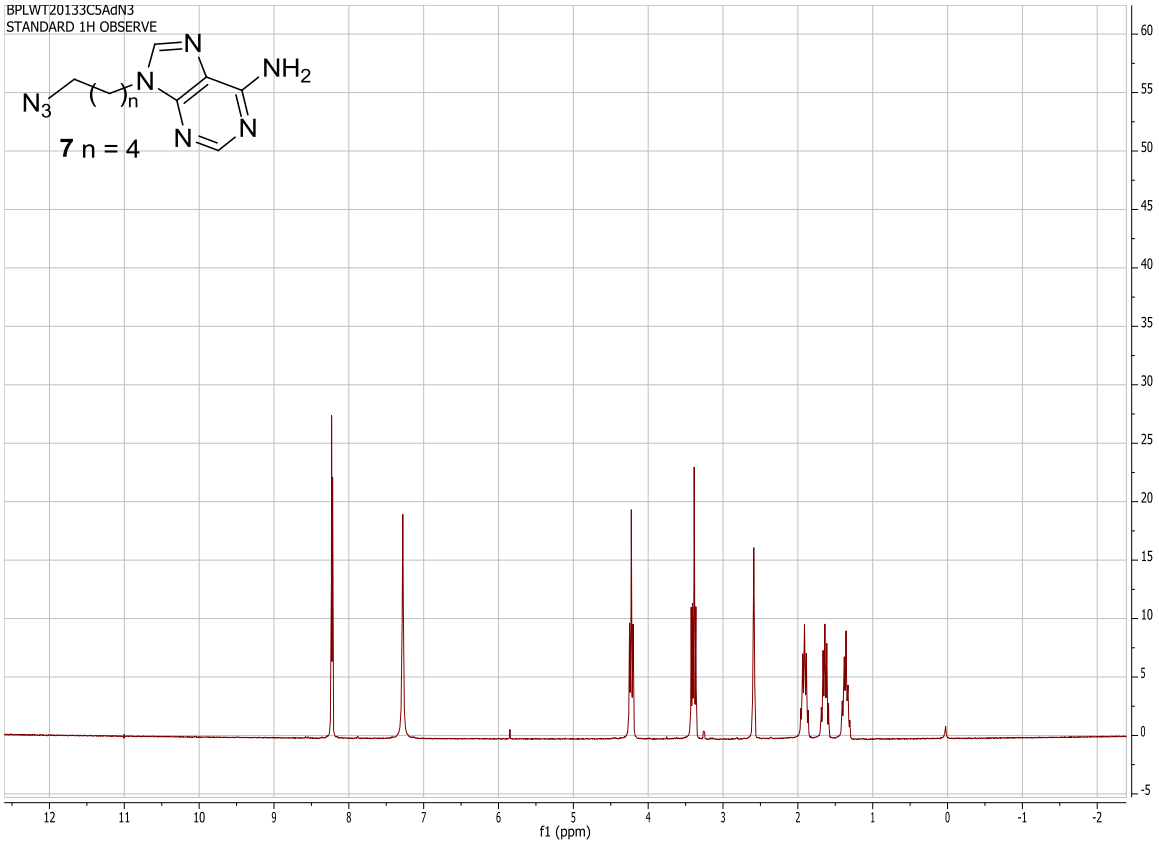
Compound 5



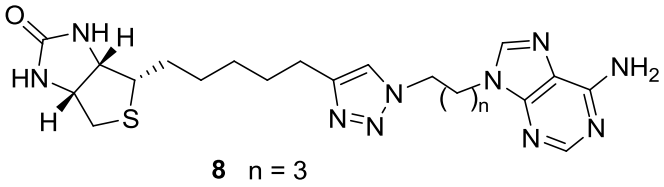
Compound 6



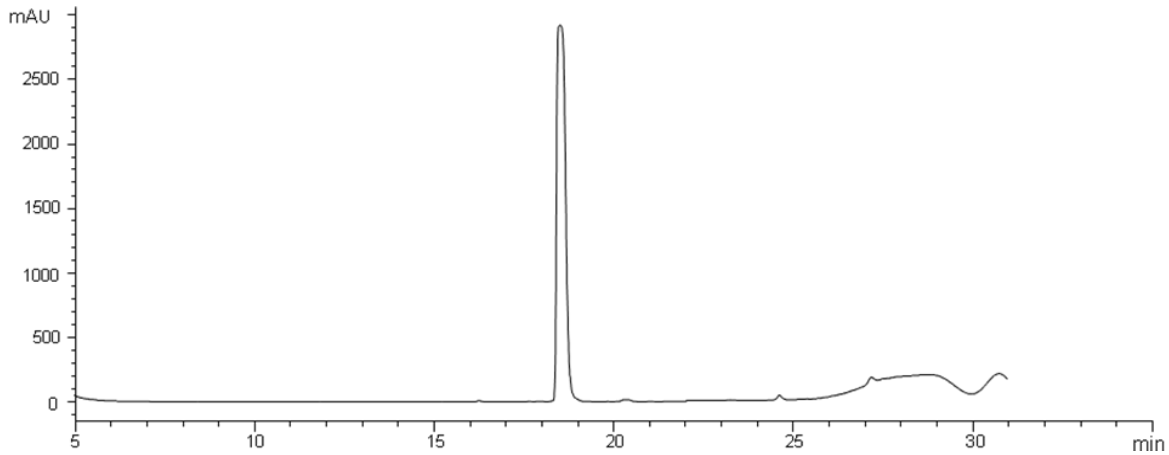
Compound 7

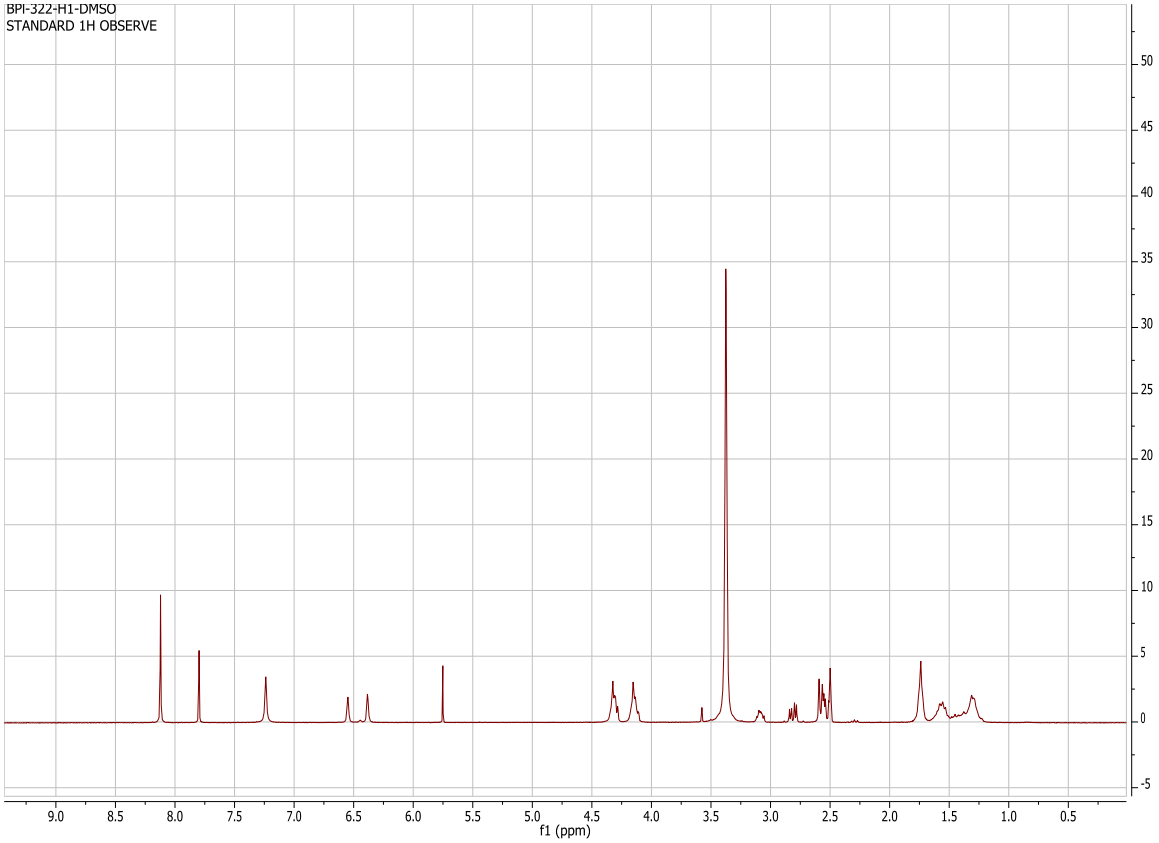


Compound 8

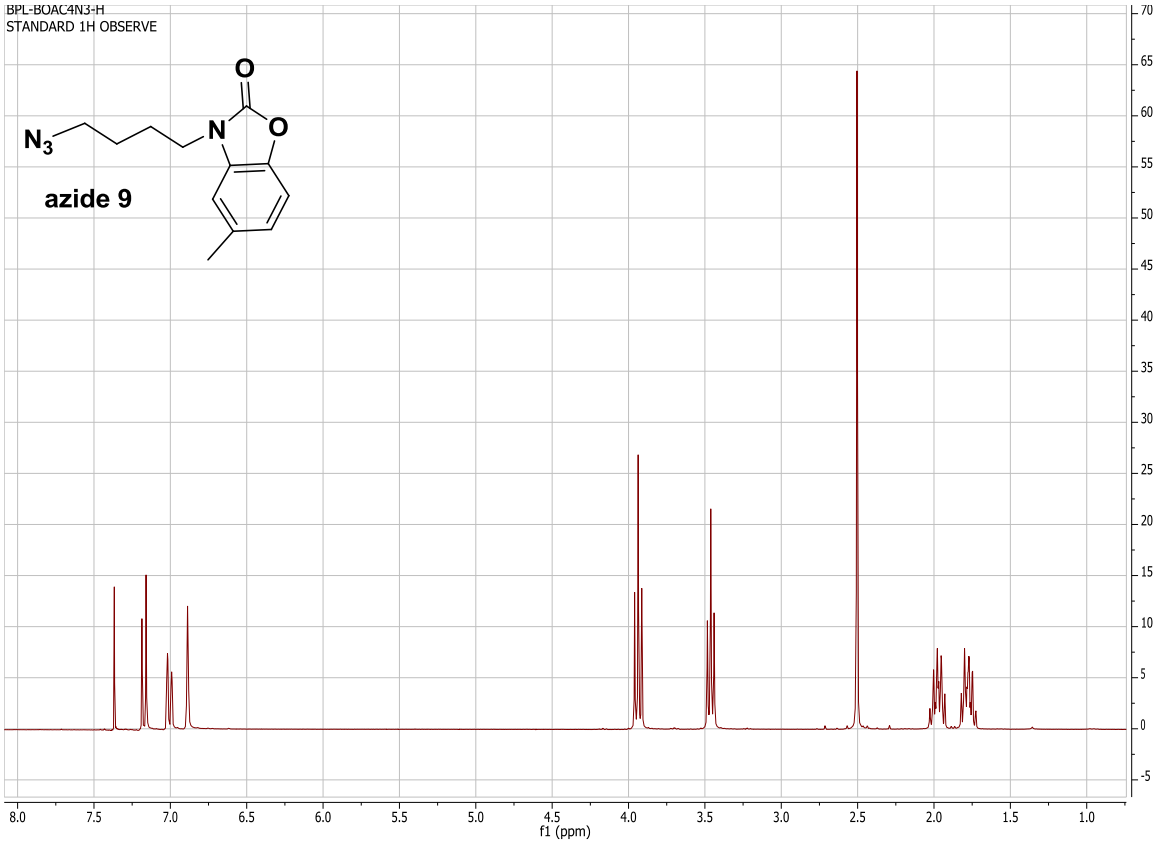


8 n = 3

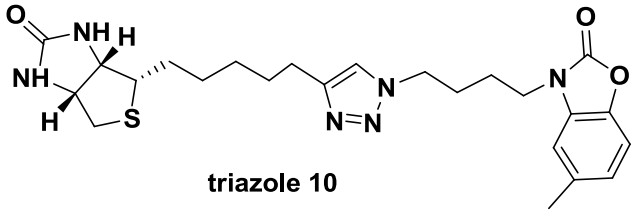



BPI-322-H1-DMSO
STANDARD 1H OBSERVE
f1 (ppm)

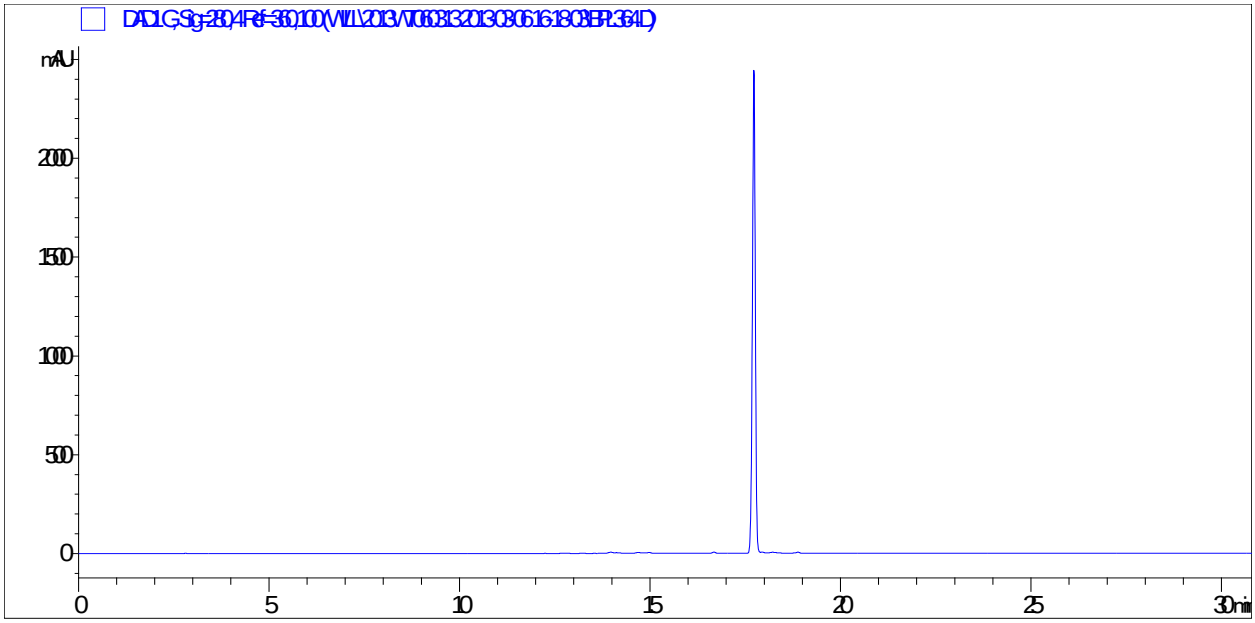
Compound 9

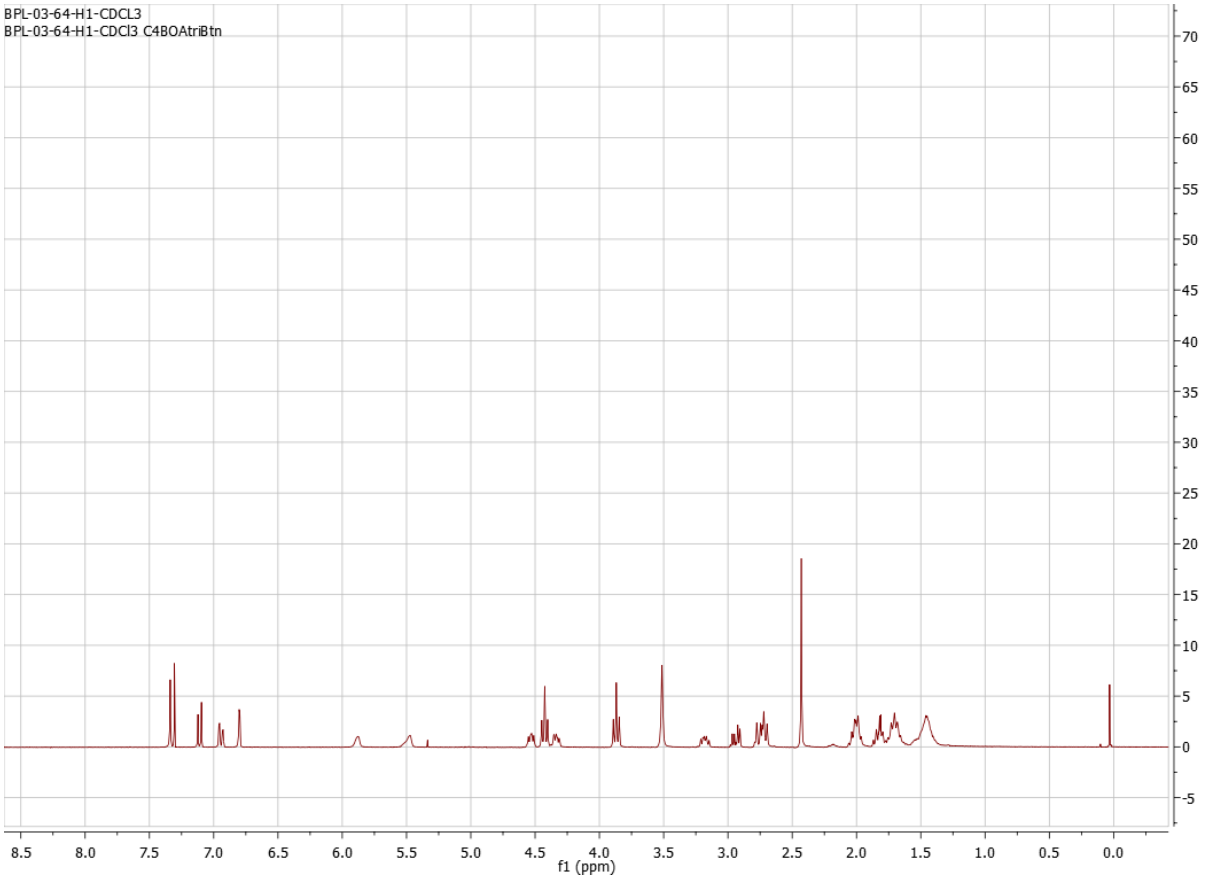

BPL-BOAC4N3-H
STANDARD 1H OBSERVE
N3
azide 9
f1 (ppm)

Compound 10

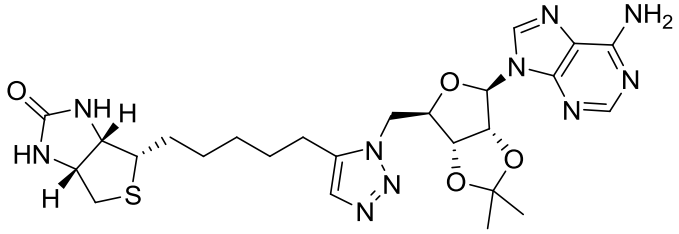


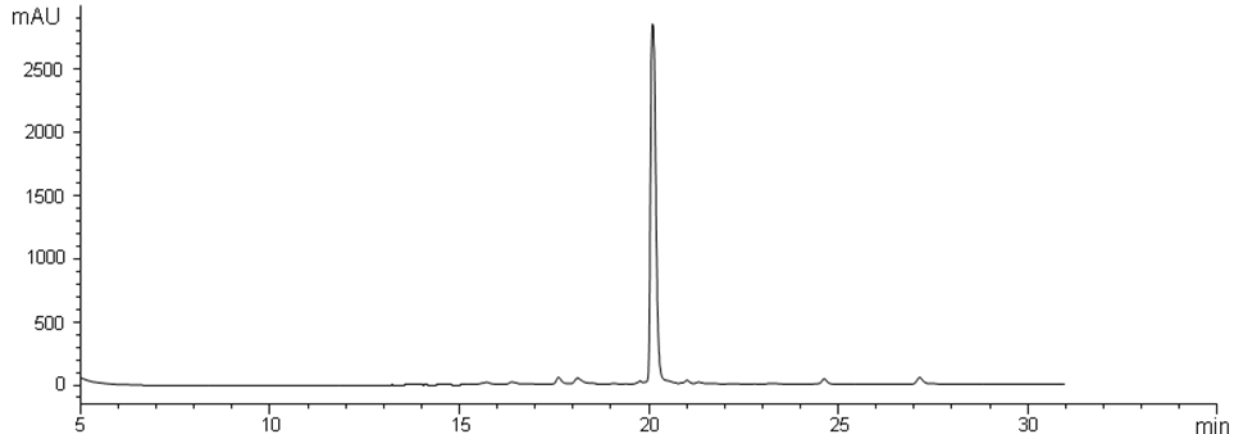
triazole 10

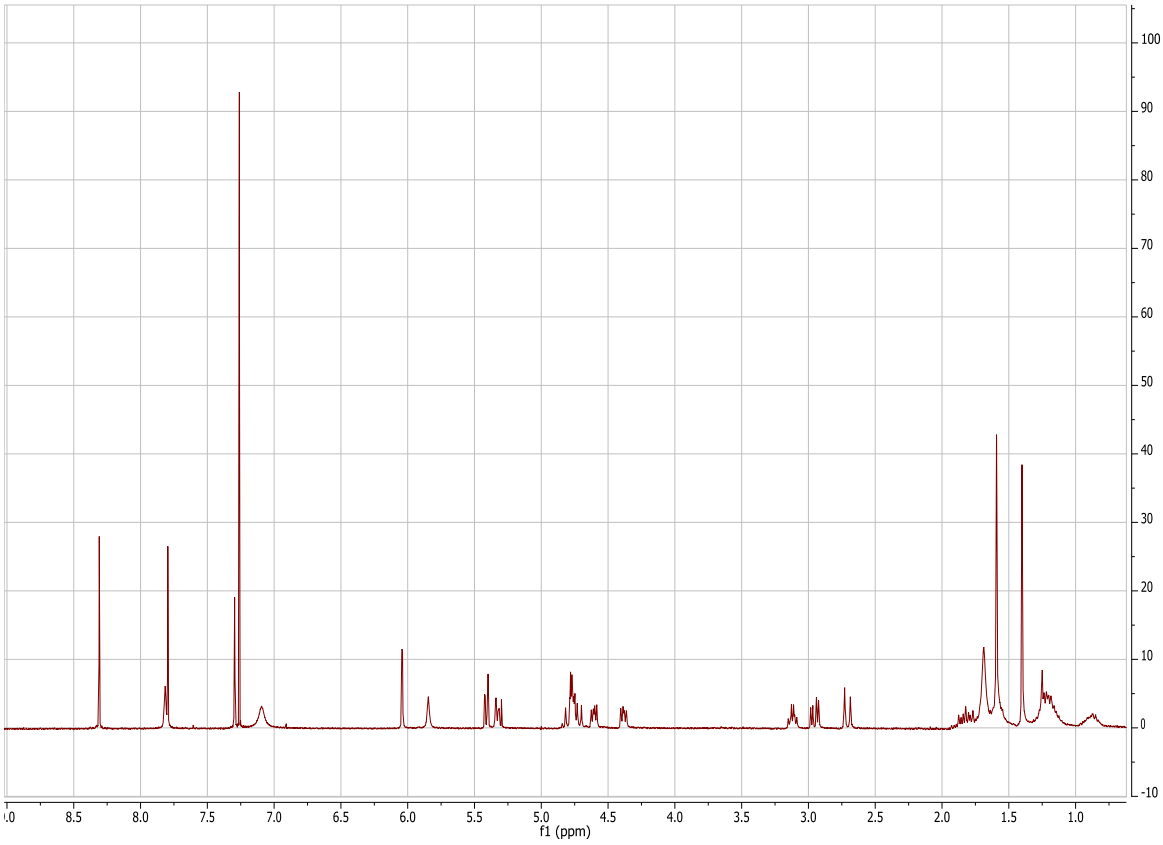




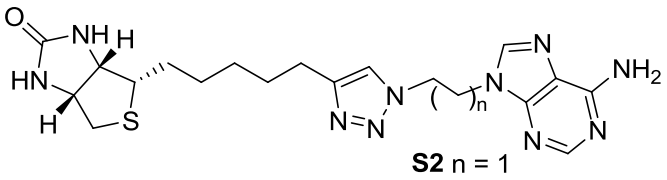
Compound S1



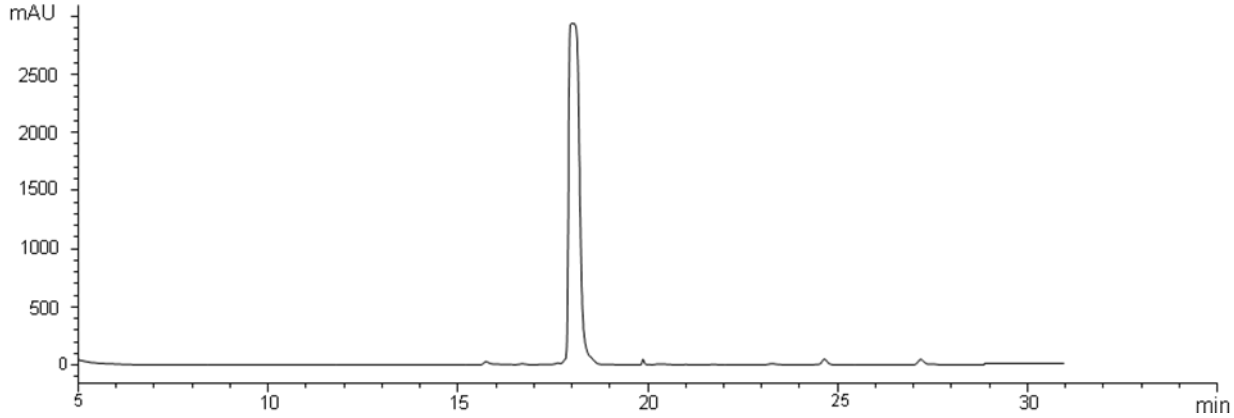


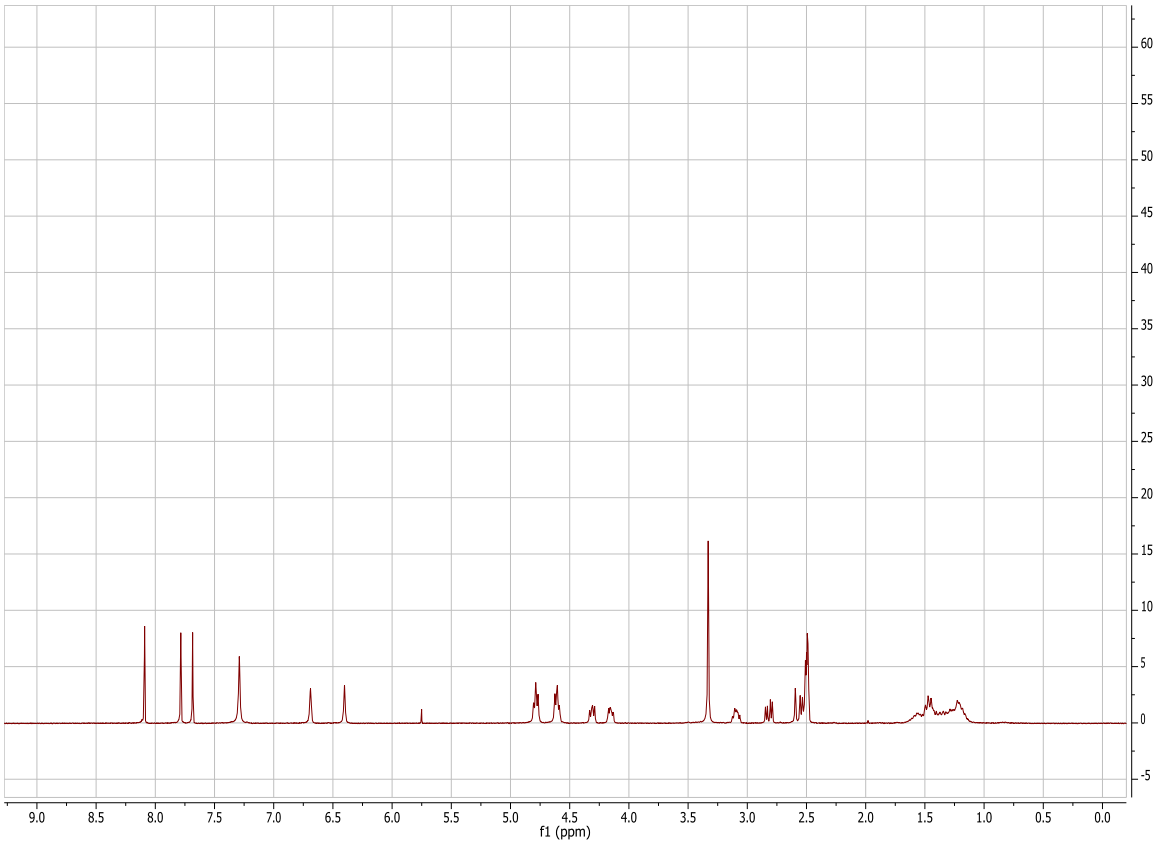


Compound S2

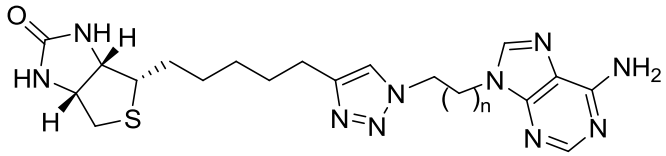


S2 n = 1

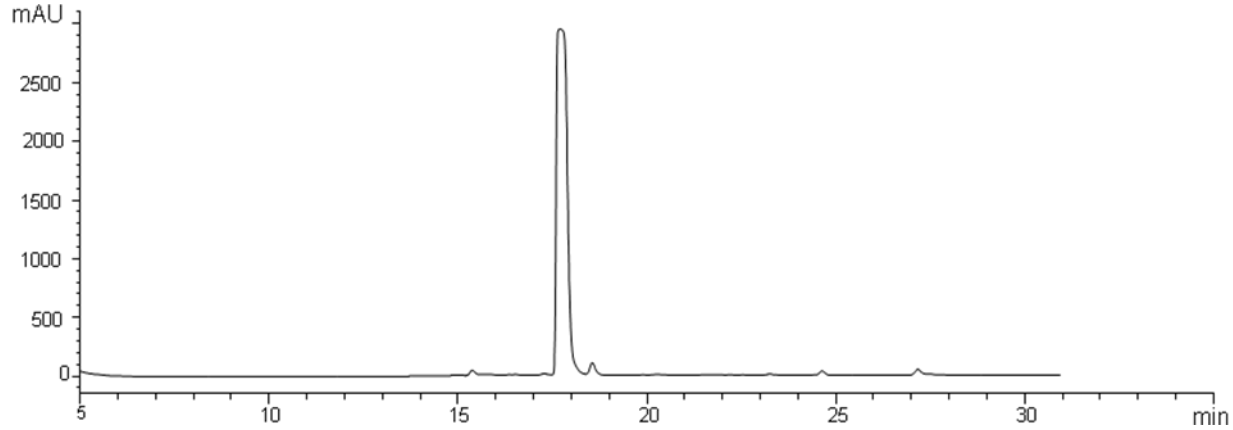


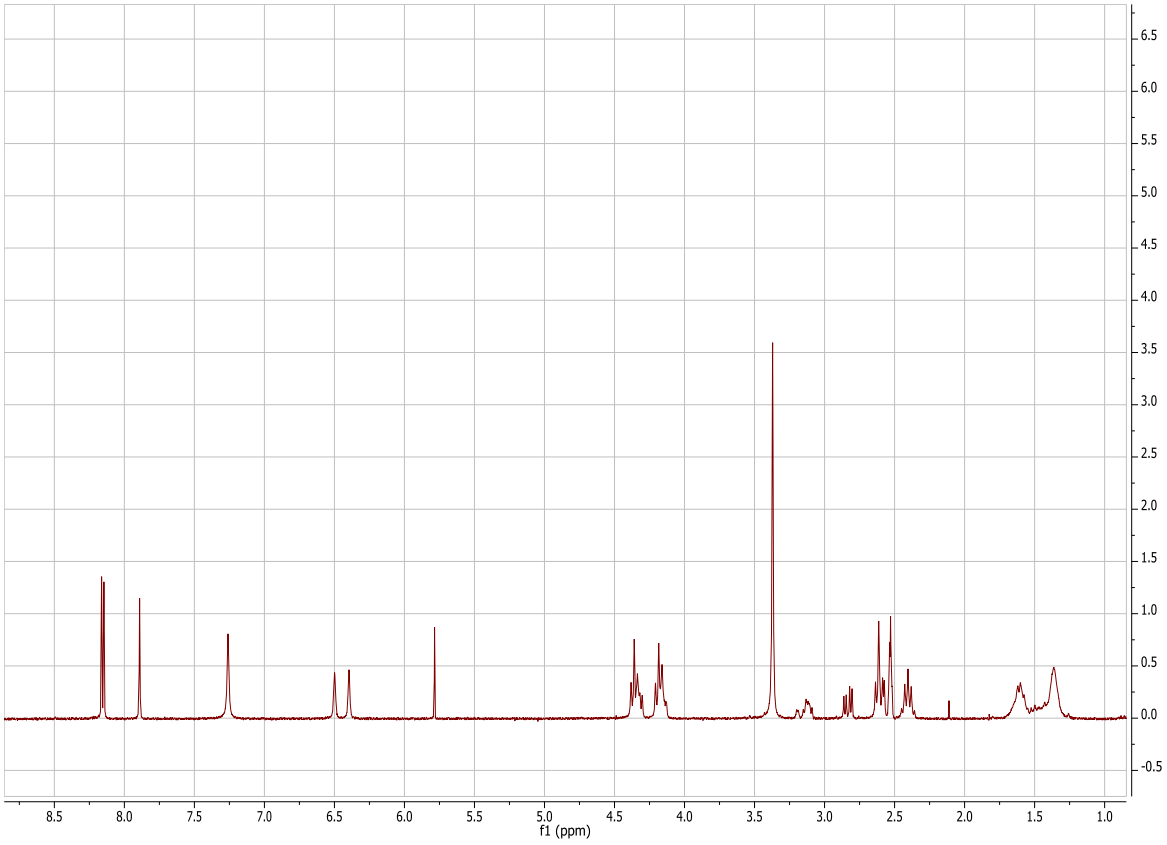


Compound S3

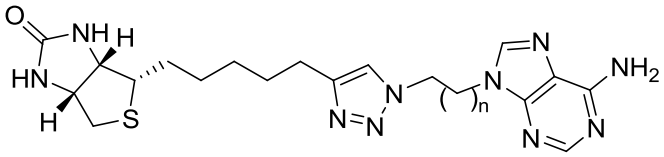


S3 n = 2

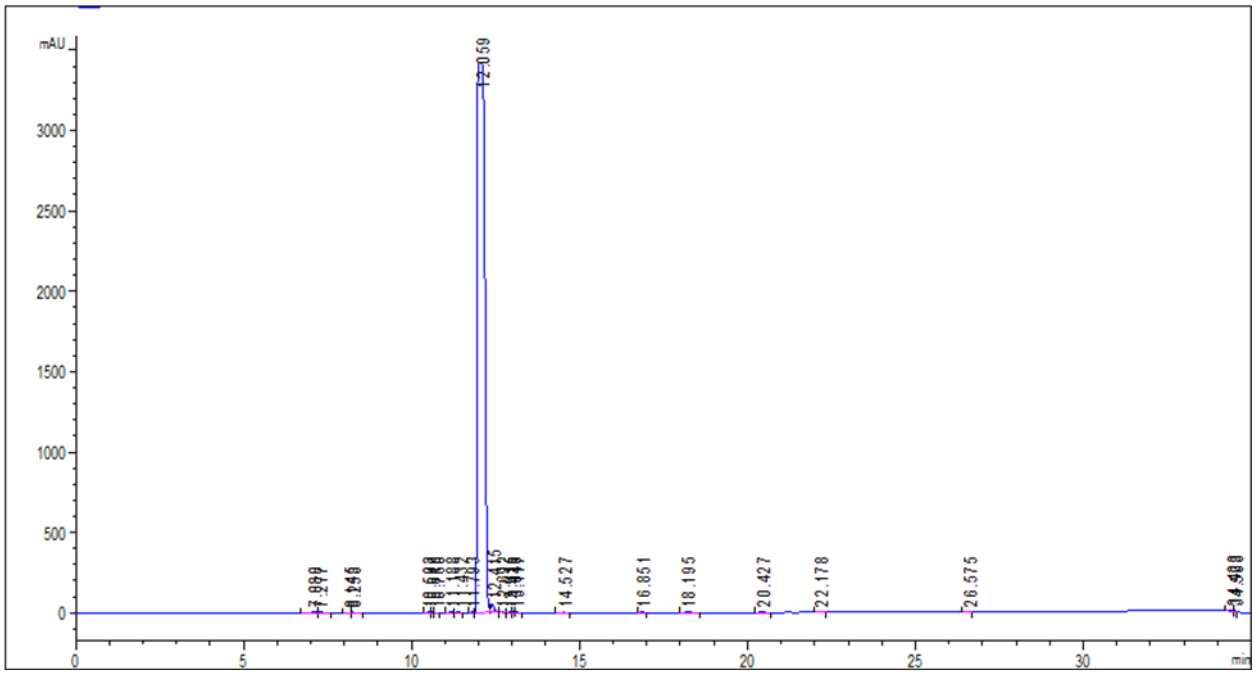


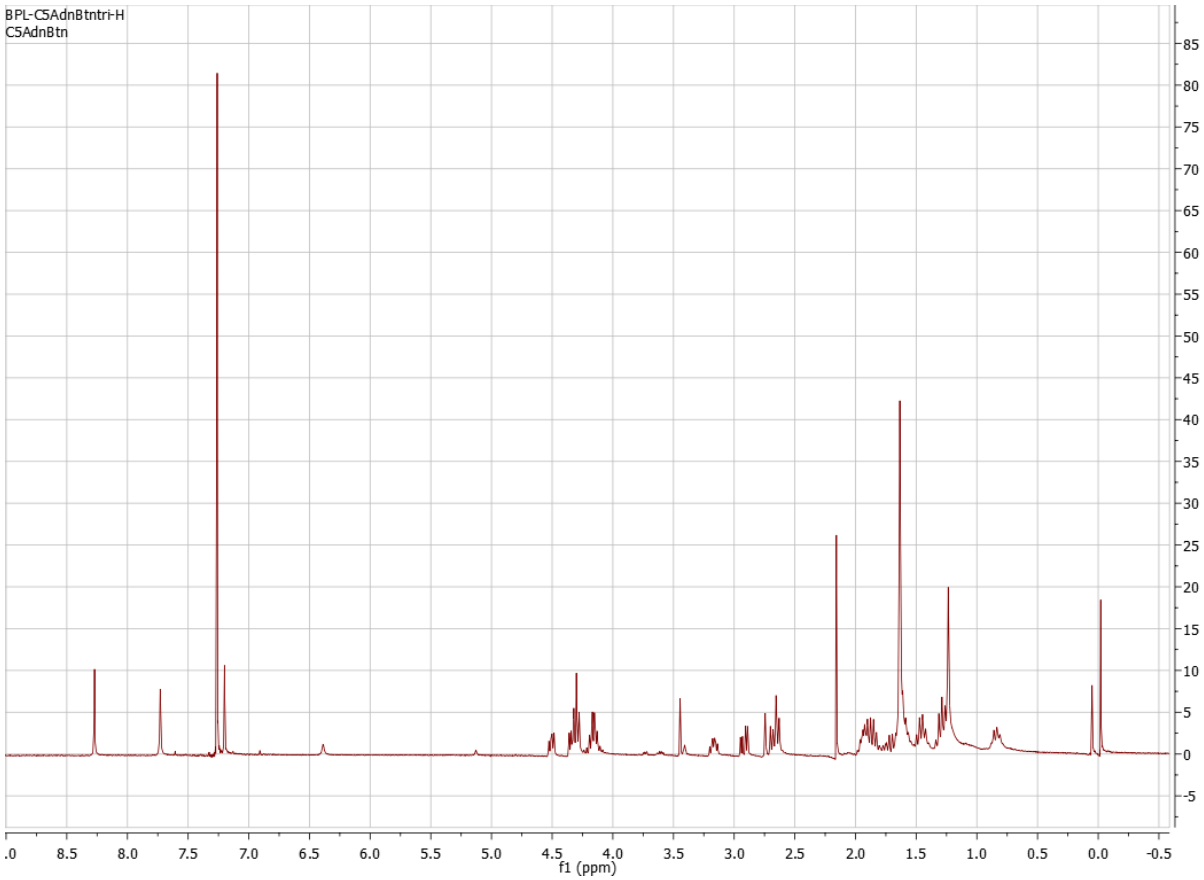


Compound S4

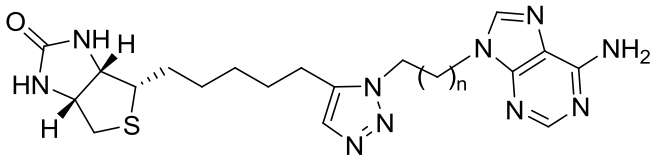


S4 n = 4

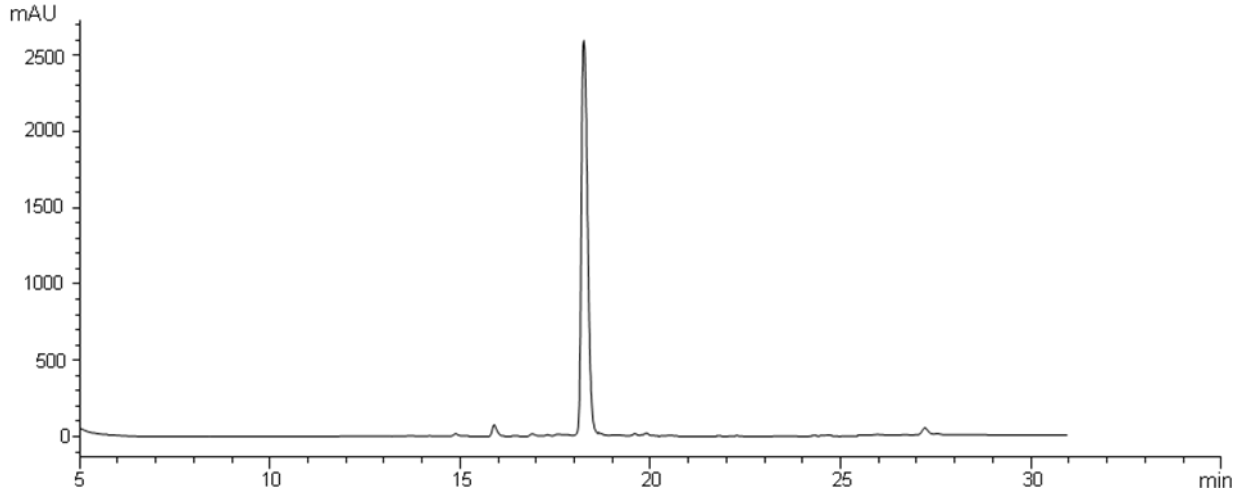


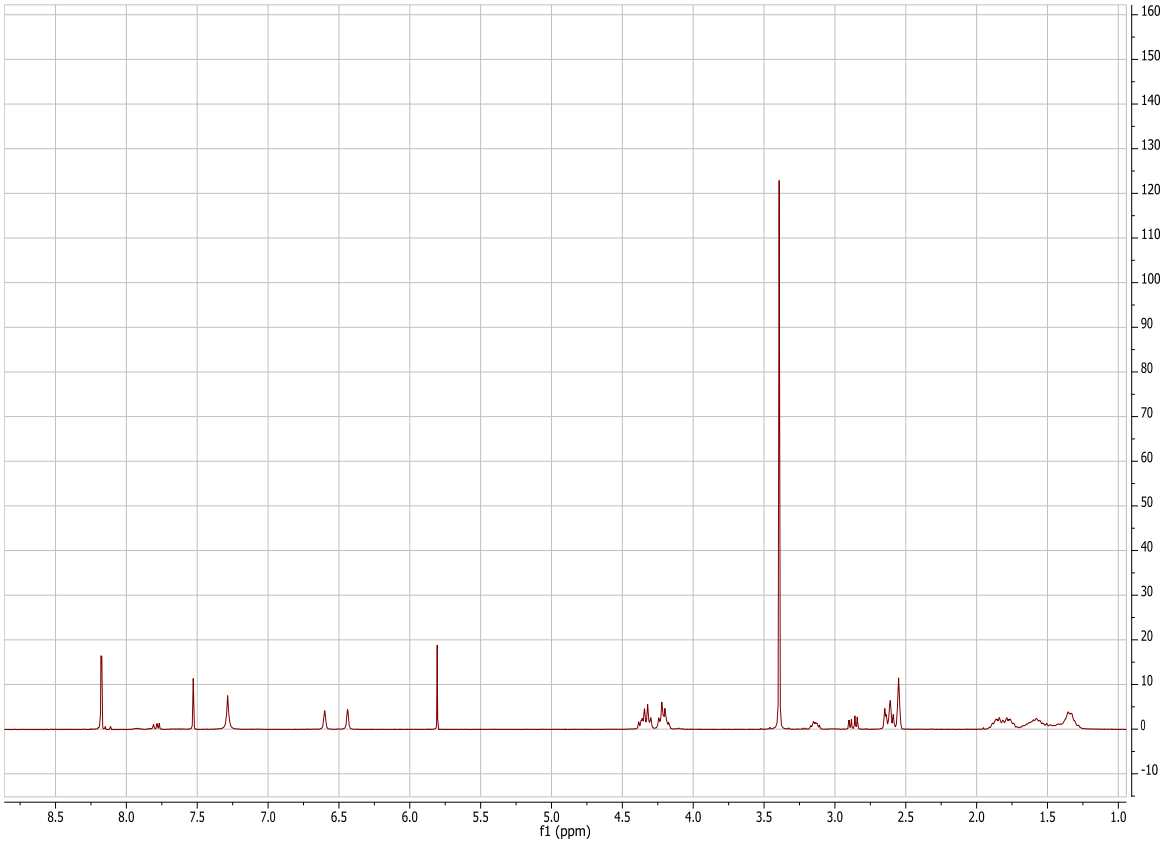


Compound S5

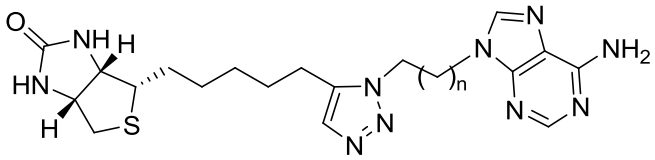


S5 n = 3

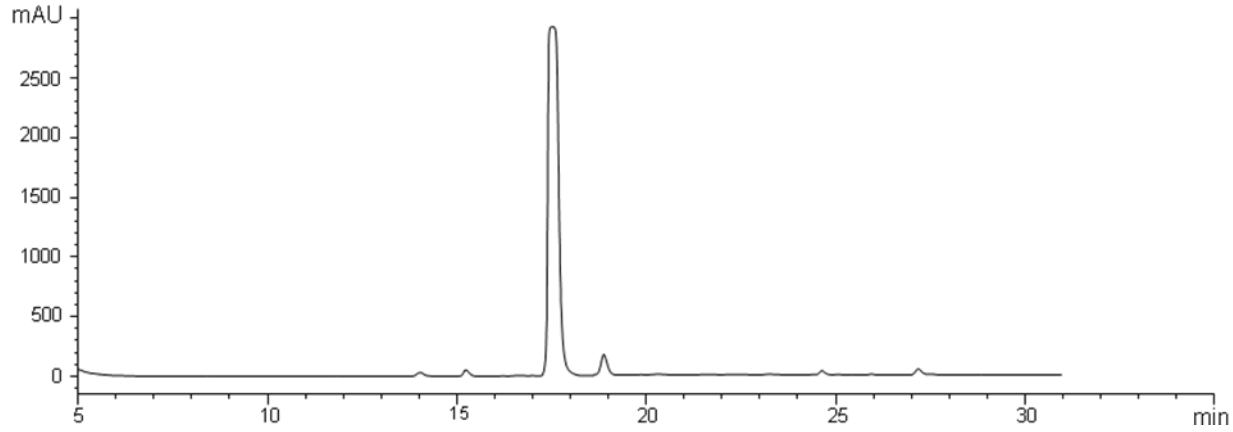


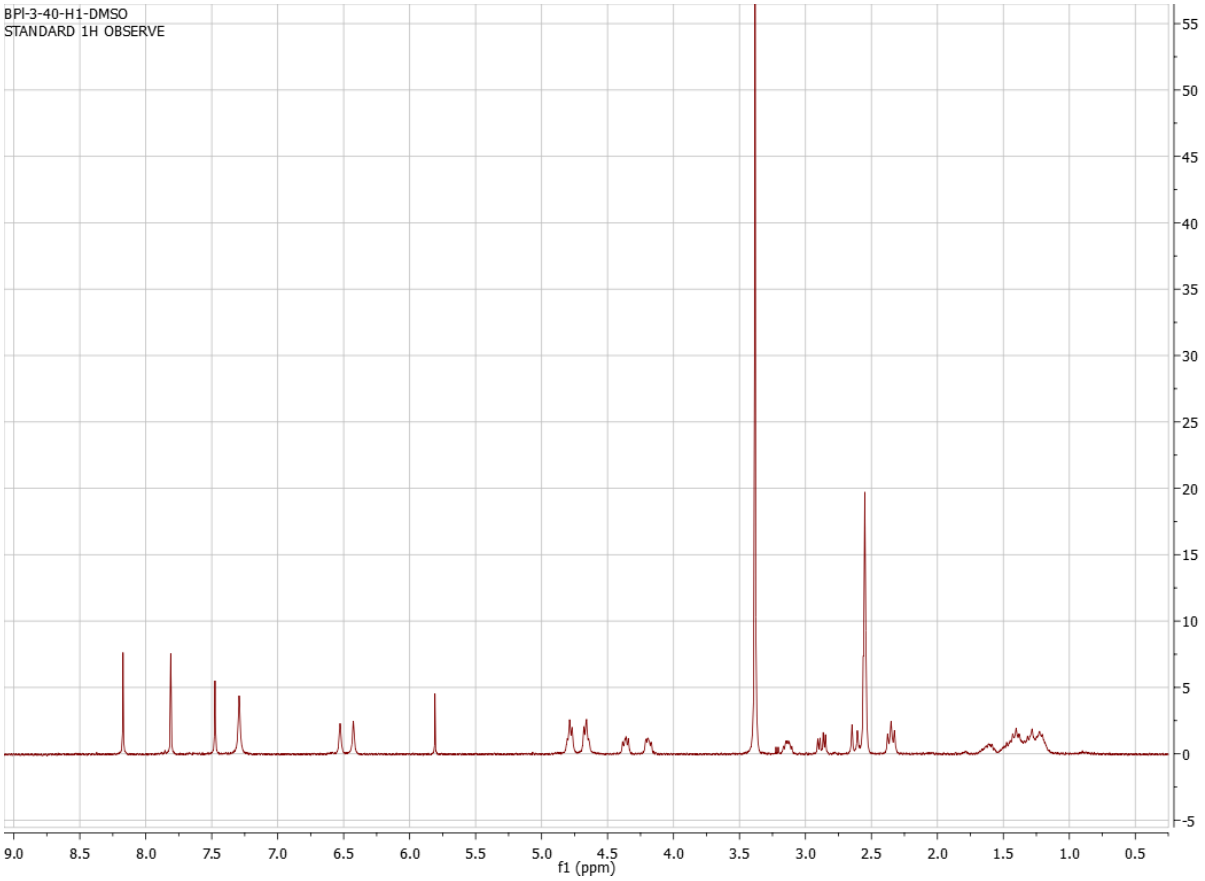


Compound S6

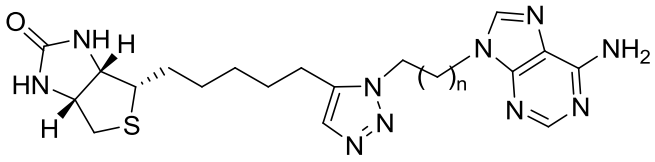


S6 n = 1

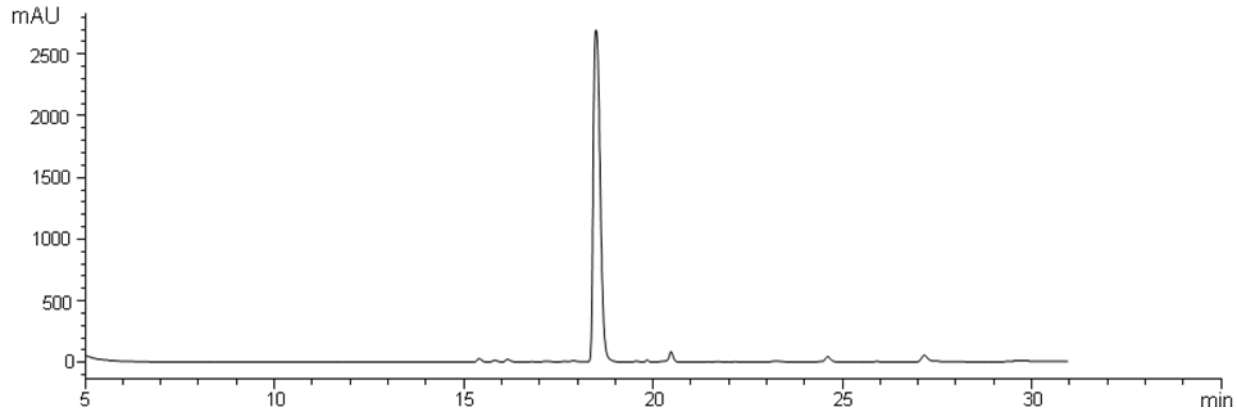


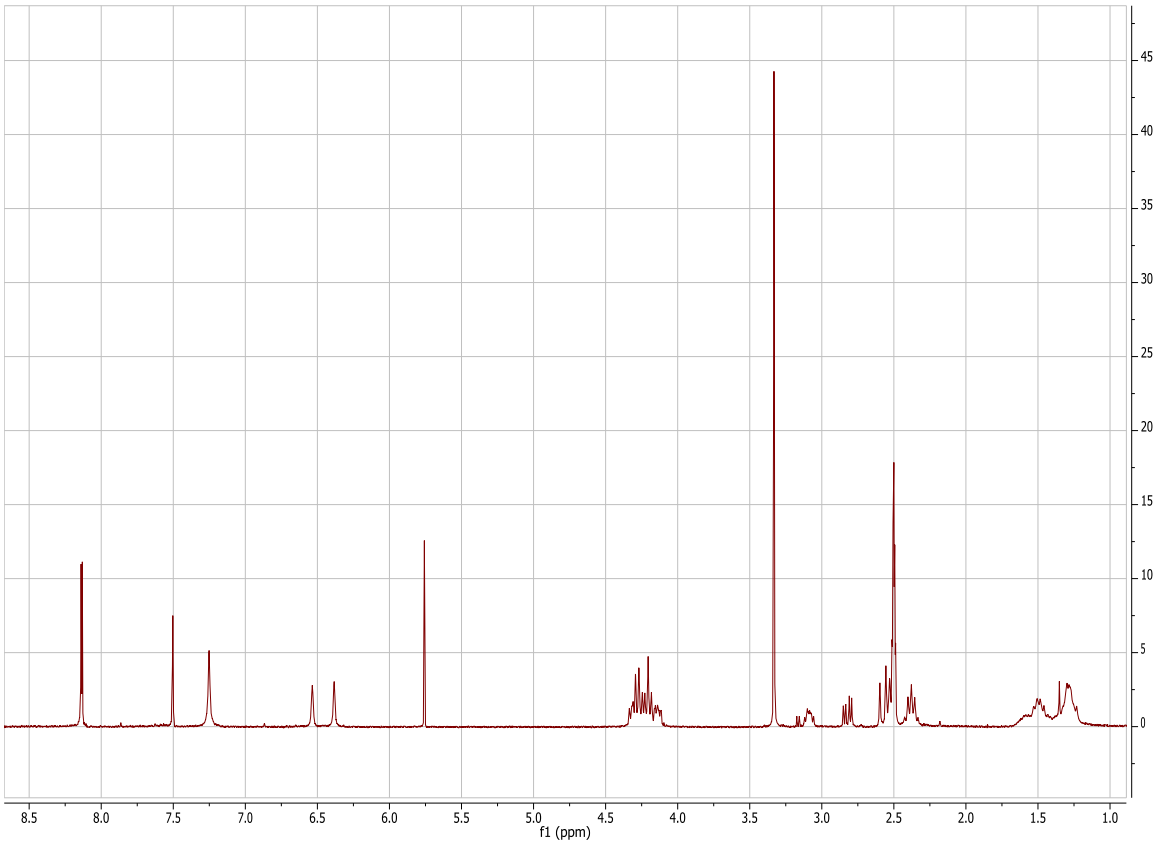


Compound S7

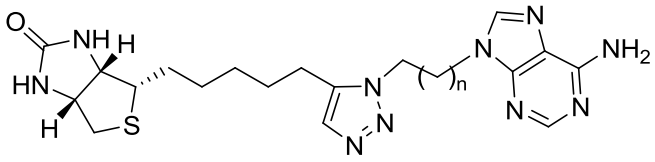


S7 n = 2

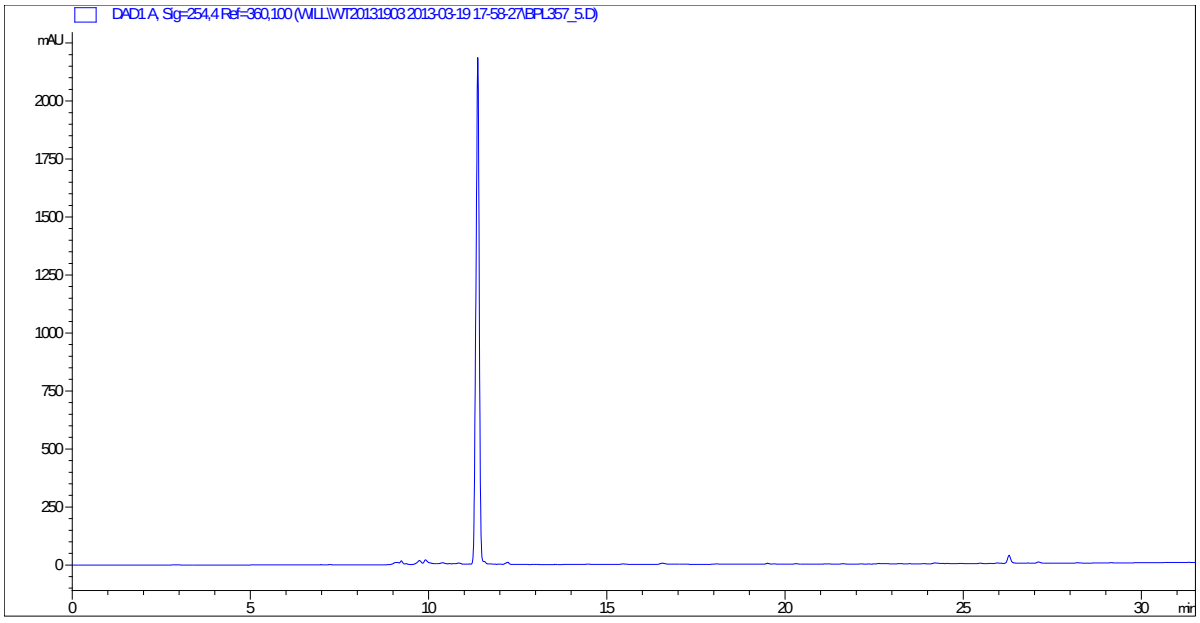


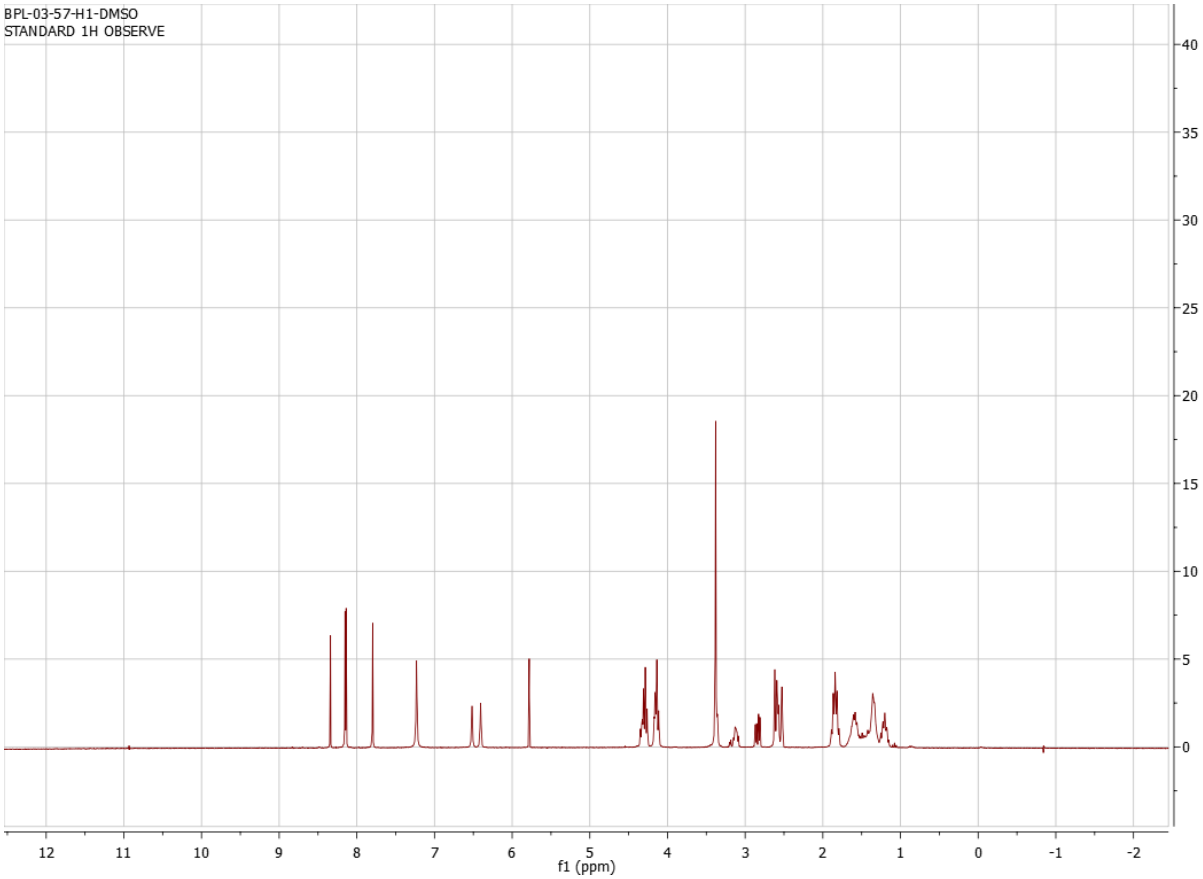


Compound S8



S8 n = 4





Supplementary References

(1) Weiss, M. S. *Journal of Applied Crystallography* **2001**, *34*, 130.

(2) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *Journal of Computational Chemistry* **2004**, *25*, 1605.

(3) Brik, A.; Alexandratos, J.; Lin, Y.-C.; Elder, J. H.; Olson, A. J.; Wlodawer, A.; Goodsell, D. S.; Wong, C.-H. *ChemBioChem* **2005**, *6*, 1167.

(4) Pendini, N. R.; Polyak, S. W.; Booker, G. W.; Wallace, J. C.; Wilce, M. C. J. *Acta Crystallographica Section F* **2008**, *64*, 520.

(5) Polyak, S. W.; Chapman-Smith, A.; Brautigan, P. J.; Wallace, J. C. *J Biol Chem* **1999**, *274*, 32847.

(6) Polyak, S. W.; Chapman-Smith, A.; Mulhern, T. D.; Cronan, J. E., Jr.; Wallace, J. C. *J Biol Chem* **2001**, *276*, 3037.

(7) Cheng, Y.; Prusoff, W. H. *Biochem Pharmacol* **1973**, *22*, 3099.

(8) Soares da Costa, T. P.; Tieu, W.; Yap, M. Y.; Pendini, N. R.; Polyak, S. W.; Sejer Pedersen, D.; Morona, R.; Turnidge, J. D.; Wallace, J. C.; Wilce, M. C. J.; Booker, G. W.; Abell, A. D. *Journal of Biological Chemistry* **2012**.

(9) Boren, B. C.; Narayan, S.; Rasmussen, L. K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V. V. *Journal of the American Chemical Society* **2008**, *130*, 8923.

(10) Nagahara, T.; Suemasu, T.; Aida, M.; Ishibashi, T.-a. *Langmuir* **2009**, *26*, 389.

(11) Zhang, Q.; Hua, G.; Bhattacharyya, P.; Slawin, Alexandra M. Z.; Woollins, J. D. *European Journal of Inorganic Chemistry* **2003**, *2003*, 2426.

(12) Itahara, T. *Journal of the Chemical Society, Perkin Transactions 2* **1998**, 1455.