Supporting Information

Site-selective and Rewritable labeling of DNA through Enzymatic, Reversible and Click Chemistries

Andrew A. Wilkinson, Elodie Jagu, Krystian Ubych, Steven Coulthard, Ashleigh E. Rushton, Jack Kennefick, Qiang Su, Robert K. Neely,* Paco Fernandez-Trillo.*

School of Chemistry, University of Birmingham, Edgbaston, Birmingham, UK, B15 2TT

r.k.neely@bham.ac.uk, f.fernandez-trillo@bham.ac.uk

Table of Contents

1. Safety Statement

No unexpected safety hazards were encountered. Azides can be explosive upon contact. This hazard depends on the molecular mass of the azide and it is recommended that the number of N does not exceed the number of C, and that the overall number of C and O is at least three times bigger than the number of N.[1] Enzymes can cause skin and eye irritation, and can be sensitizers. When handling enzymes, care should be taken to avoid inhalation of aerosols and contact with skin and eyes.[2]

2. General Materials and Methods

Reagents were purchased from Sigma-Aldrich except 6-heptyn-1-ol, which was purchased from Fluorochem. *N*-Boc-hydroxylamine was purchased from Alfa Aesar. *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-4-formylbenzamide (**5**) was purchased from BroadPharm®. TAMRA-dibenzylcyclooctyne conjugate **6** was purchased from Click Chemistry Tools LLC. CutSmart® buffer, pUC19 plasmid, proteinase K, HaeII and TaqαI (R.TaqI) were purchased from New England BioLabs (NEB). The DNA sequence for 14 bp oligo was 3'-GCCGCTCGATGCCG-5' 5'-CGGCGAGCTACGGC-3', and was purchased from Integrated DNA technologies®. Flash chromatography was performed using Geduran Si 60 from Merck and TLC (F254) analysis was performed using 60 Å silica gel from VWR International. TLC plates were visualized by staining with potassium permanganate stain or UV-absorption. NMR data was acquired on a Bruker Avance III operating at 300 or 400 MHz. MS spectra were obtained on a Xevo® G2-XS-ToF (Waters) and Synapt-G2-S from electrospray ionization (ESI) and time-of-flight (TOF) measurement in negative or positive ion mode. IR spectra were acquired on Agilent Technologies Cary 600 Series FTIR Spectrometer. UV-Vis absorbance measurements were performed on Shimadzu BioSpec-nano.

2.1. Structure of Dyes

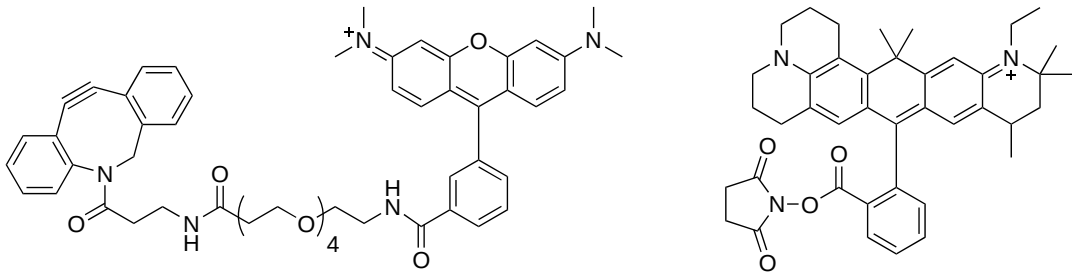


Figure S1 Structure of the dyes used in this work.

2.2. LC Analysis, Concentration and Yield Determination

Analytical RP-HPLC for **3a,b** and **4a,b** was performed on Shimadzu LC-20 Prominence equipped with ACE 5 C18 (250 x 4.6 mm, flow rate 1 mL/min). Elution with 10 mM ammonium acetate pH 5.5/MeCN gradient: 3-30% MeCN over 20 minutes, 30-97% over 15 minutes hold 97% for 10 minutes. Preparative RP-HPLC was performed on Agilent Technologies 1260 Infinity equipped with ACE 5 C18 (250 x 21.2 mm, 100 Å, flow rate 10 mL/min). Elution with 10 mM ammonium acetate pH 5.5/MeCN gradient: 3-30% MeCN over 30 minutes, 30-97% over 20 minutes, hold 97% 5 minutes.

Analytical RP-HPLC for oligonucleotides was performed on Agilent Technologies 1260 Infinity equipped with Phenomenex Gemini® C18 (150 x 4.6 mm, 5 µm, 100 Å, flow rate 1 mL/min) at 60 °C. Elution with a 0.1 M triethyl amine acetate buffer, pH 7.0/MeCN: gradient A: 5-18% MeCN over 25 min, to 100% 5

[1] Bräse, S., Gil, C., Knepper, K. and Zimmermann, V. *Angew. Chem., Int. Ed.* **2005**, *19*, 5188-5240.

[2] https://amfep.org/_library/_files/amfep-guide-on-safe-handling-of-enzymes-updated-in-2013.pdf. Accessed 10/01/2019

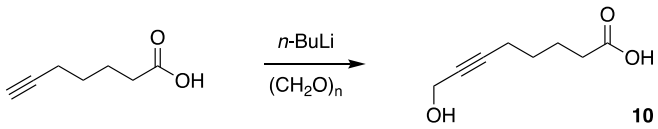
min, hold at 100% 10 min, lower to 5% for 5 min; gradient B: 5-31% MeCN over 50 min, to 100% 10 min, hold at 100% 5 min, lower to 5% for 10 min. Gradient A was used for unlabeled and alkylated oligonucleotides and gradient B was used for all remaining samples. The UV-detection was carried out at 260 nm.

Analytical RP-UPLC-MS for oligonucleotides was performed on Waters Ltd. Xevo-G2-XS Tof equipped with ACQUITY UPLC® Oligonucleotide BEH C18 (50 x 2.1 mm, 1.7 µm, 130 Å, flow rate 0.2 mL/min) at 60 °C. Elution with a 0.5 % triethyl amine water/MeCN gradient: 0-100% MeCN over 20 min, hold at 100% 5 min, lower to 0% for 5 min. The UV-detection was carried with a diode array and the trace at 260 nm represented. MS detection was done on negative mode.

The concentration and yields for **3** and **4** were determined by UV absorbance measurements at 260 nm, performed in 0.1 % acetic acid. For the calculation, molar extinction coefficient was used: ε_{260} = 15,400 $dm^3\,mol^{-1}\,cm^{-1}$.

3. Synthesis of Cofactors

3.1. 8-Hydroxyoct-6-ynoic acid (10)



A solution of 6-heptynoic acid (2 g, 15.87 mmol) was made in dry THF (42 mL) under argon, to this, HMPA (34.9 mmol, 6.13 mL) was added and the solution was cooled to -78 °C. To this, *n*-BuLi (1.6 M in hexanes, 34.9 mmol, 21.8 mL) was added dropwise whilst maintaining the temperature below -60 °C. The solution was then warmed to -40 °C and stirred for 1 h. After 1 h, paraformaldehyde (1.47 g, 47.6 mmol) was added via powder funnel under an argon flow. The reaction mixture was then warmed to 45 °C for 4 h. After reaction, the mixture was quenched with 1 M HCl to pH 4-5 and extracted with EtOAc. Organic layer was then dried over Na_2SO_4 and the EtOAc was removed by rotary evaporation giving the crude product. Purification was completed using flash column chromatography (silica gel, Hex:EtOAc, 6:4): Yield = 68%; Rf = 0.27 (Hex:EtOAc, 6:4); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 12.03 (s, 1H), 5.03 (s, 1H), 4.02 (d, *J* = 2.6 Hz, 2H), 2.29 – 2.14 (m, 4H), 1.63 – 1.50 (m, 2H), 1.50 – 1.39 (m, 2H); ^{13}C-NMR (101 MHz, DMSO) δ 174.3, 83.8, 80.5, 49.1, 33.1, 27.6, 23.7, 23.7, 17.7; IR (cm^{-1}) *ν* max: 3003, 1700, 1411, 1218, 1133, 1002; HRMS-ESI(-): calcd for $C_8H_{11}O_3$: 155.0708, found: 155.0713 [M-H].

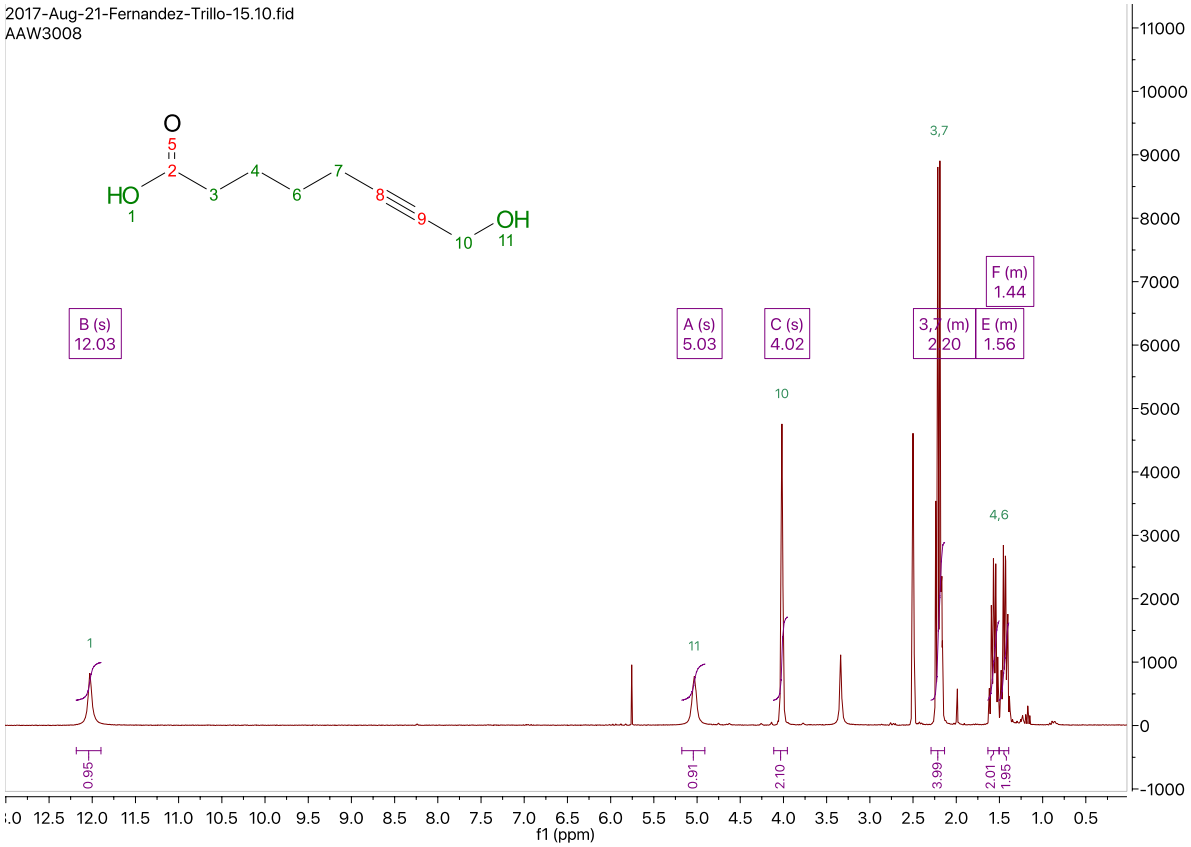


Figure S2 ^{1}H-NMR spectrum of 8-hydroxyoct-6-ynoic acid (**10**)

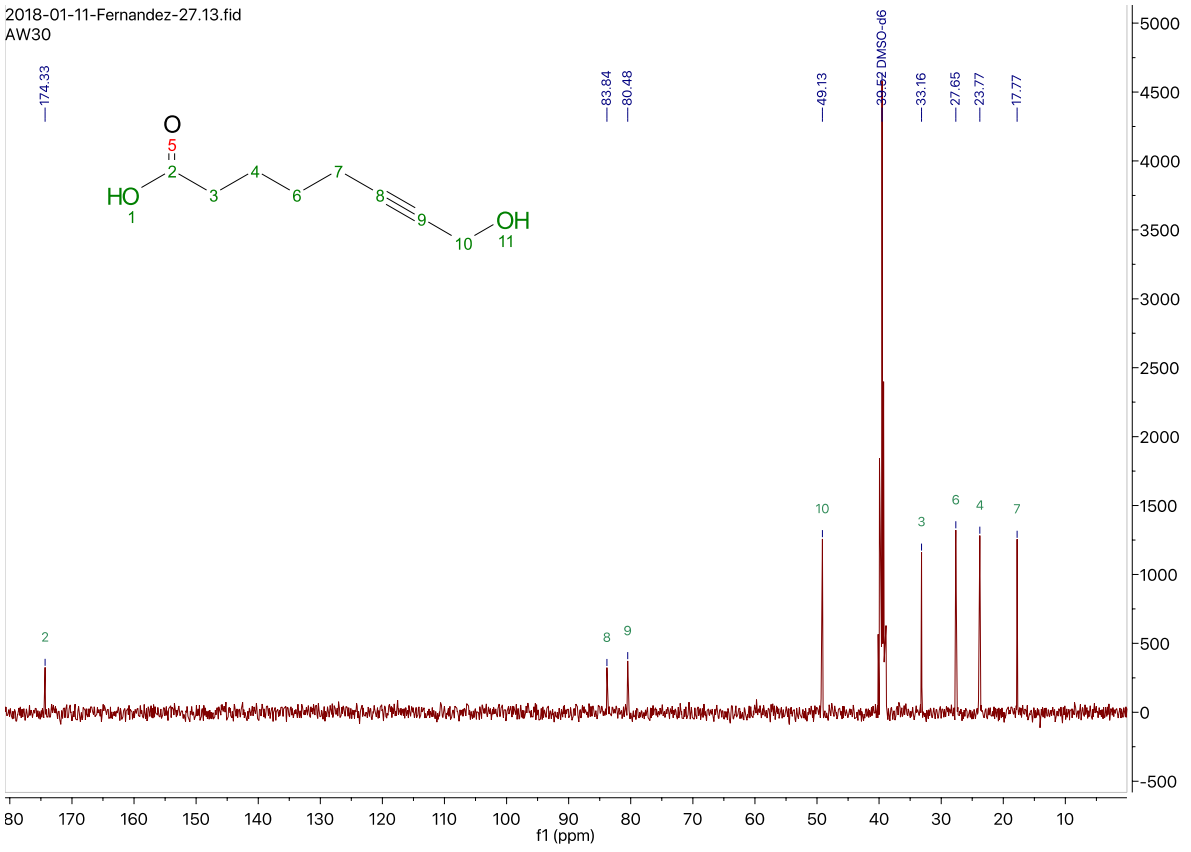


Figure S3 ^{13}C-NMR spectrum of 8-hydroxyoct-6-ynoic acid (**10**)

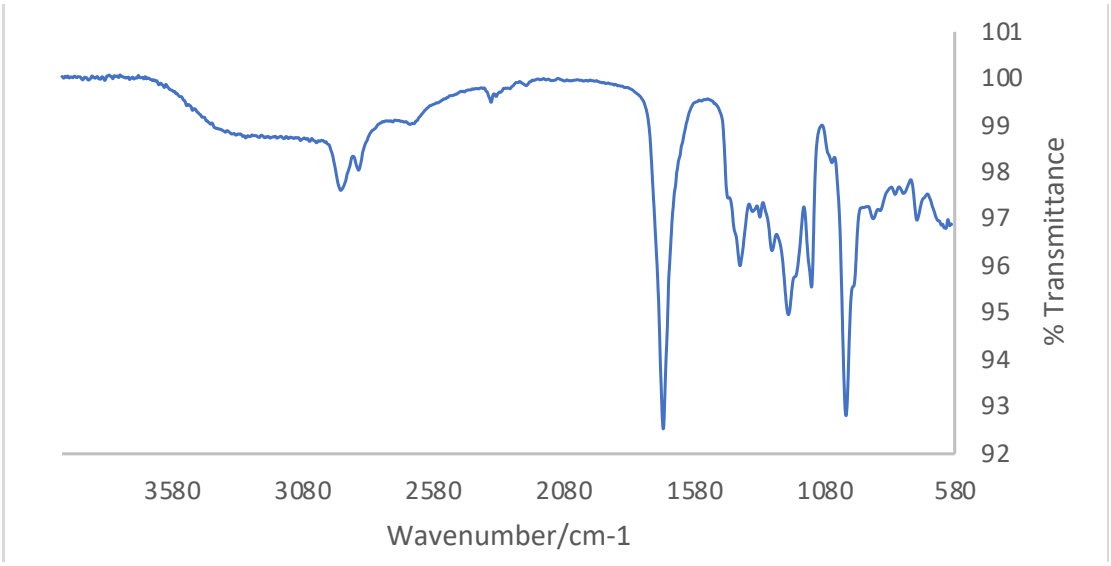
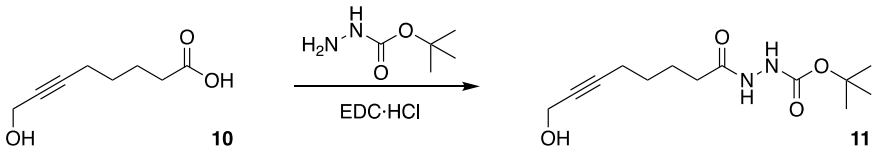


Figure S4 Infrared spectrum of 8-hydroxyoct-6-ynoic acid (**10**)

3.2. *tert*-Butyl-2-(8-hydroxyoct-6-ynoyl)hydrazine-1-carboxylate (11)



8-hydroxyoct-6-ynoic acid (**10**) (1.35 g, 8.65 mmol) and *tert*-butylcarbazate (1.4 g, 10.38 mmol) were dissolved in 2:1 THF:H_2O (13.5:6.75 mL). To this, EDC·HCl (1.87 g, 9.52 mmol) was added slowly over 15 minutes. The mixture was left to stir for 3 h and then extracted with EtOAc. The organic layer was washed with 0.1 M HCl, water and brine and the organic layer was collected, dried over anhydrous sodium sulfate and the solvent was removed under reduced pressure yielding the product as a white solid: Yield = 63%; Rf = 0.14 (Hex:EtOAc 1:1) ^{1}H-NMR (400 MHz, DMSO-d_6) δ 9.47 (s, 1H), 8.66 (s, 1H), 5.04 (t, *J* = 5.9 Hz, 1H), 4.02 (dt, *J* = 5.9, 2.2 Hz, 2H), 2.19 (tt, *J* = 7.1, 2.2 Hz, 2H), 2.06 (t, *J* = 7.2 Hz, 2H), 1.58 (p, *J* = 7.3 Hz, 2H), 1.50 – 1.32 (m, 12H); ^{13}C-NMR (101 MHz, DMSO) δ 172.0, 155.6, 84.3, 80.9, 79.4, 49.6, 33.0, 28.5, 28.1, 24.7, 18.2; IR (cm^{-1}) *ν* max: 3300, 1668, 1370, 1245, 1157; HRMS-ESI(+): calcd for $C_{13}H_{22}N_2O_4Na$: 293.1477, found: 293.1477 [M+Na].

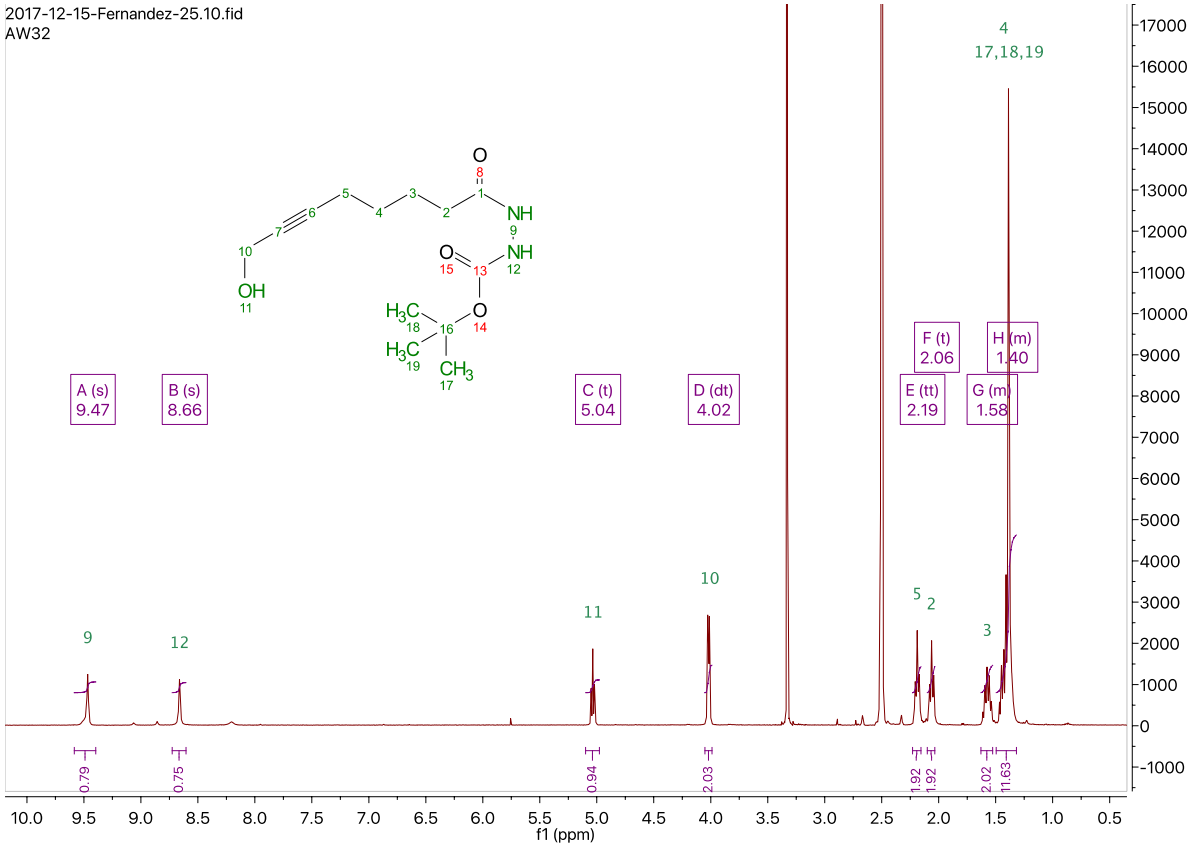


Figure S5 ^{1}H-NMR spectrum of *tert*-butyl-2-(8-hydroxyoct-6-ynoyl)hydrazine-1-carboxylate (**11**)

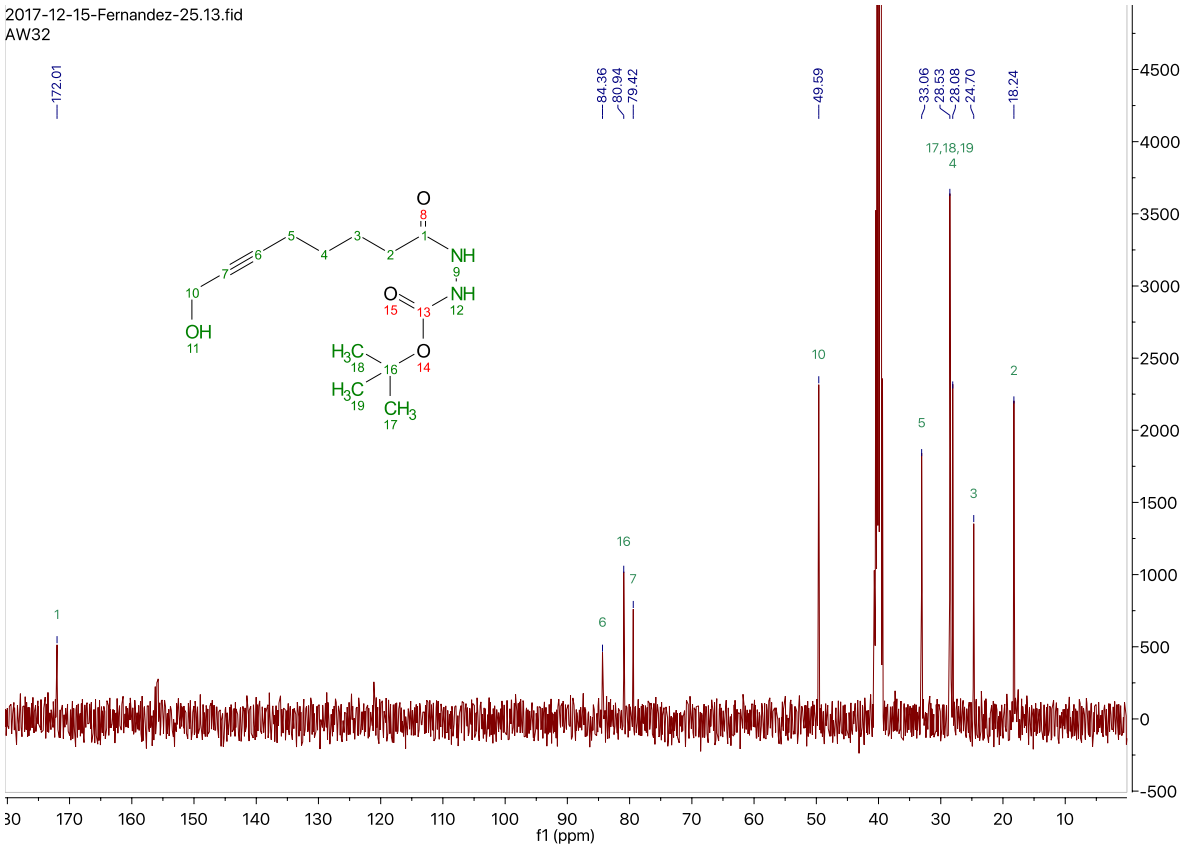


Figure S6 ^{13}C-NMR spectrum of *tert*-butyl-2-(8-hydroxyoct-6-ynoyl)hydrazine-1-carboxylate (**11**)

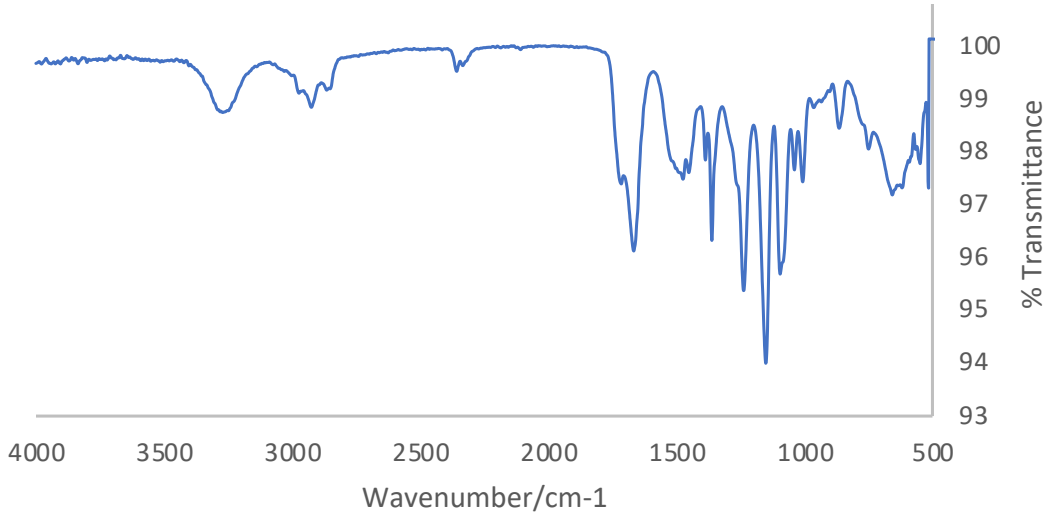
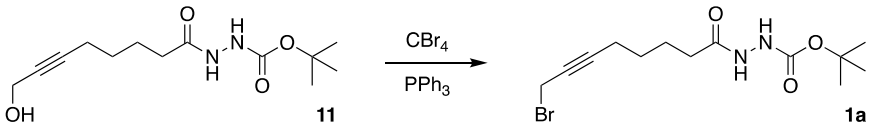


Figure S7 Infrared spectrum of *tert*-butyl-2-(8-hydroxyoct-6-ynoyl)hydrazine-1-carboxylate (**11**)

3.3. *tert*-Butyl-2-(8-bromooct-6-ynoyl)hydrazine-1-carboxylate (1a)



A solution of *tert*-butyl-2-(8-hydroxyoct-6-ynoyl)hydrazine-1-carboxylate (**11**) (300 mg, 1.11 mmol) was made in dry DCM (3.33 mL) and cooled on ice. Triphenylphosphine (437 mg, 1.67 mmol) was added and left to dissolve, once dissolved tetrabromomethane (552 mg, 1.67 mmol) was added slowly. The reaction was then brought to room temperature and left to stir for 1 h. After reaction, the solvent was removed under reduced pressure and the crude mixture was purified by flash column chromatography (Hex:EtOAc, 7:3): Yield = 55%; Rf = 0.15 (Hex:EtOAc 7:3); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 9.48 (s, 1H), 8.67 (s, 1H), 4.21 (t, J = 2.3 Hz, 2H), 2.27 (tt, J = 6.9, 3.4 Hz, 2H), 2.06 (t, J = 7.4 Hz, 2H), 1.65 – 1.31 (m, 13H); ^{13}C-NMR (101 MHz, DMSO) δ 171.4, 155.2, 87.7, 78.9, 76.3, 54.9, 39.5, 32.5, 28.0, 27.3, 24.1, 17.9, 17.2; HRMS-ESI(+): calcd for $C_{13}H_{21}N_2O_3NaBr$: 355.0633, found: 355.0634 [M+Na].

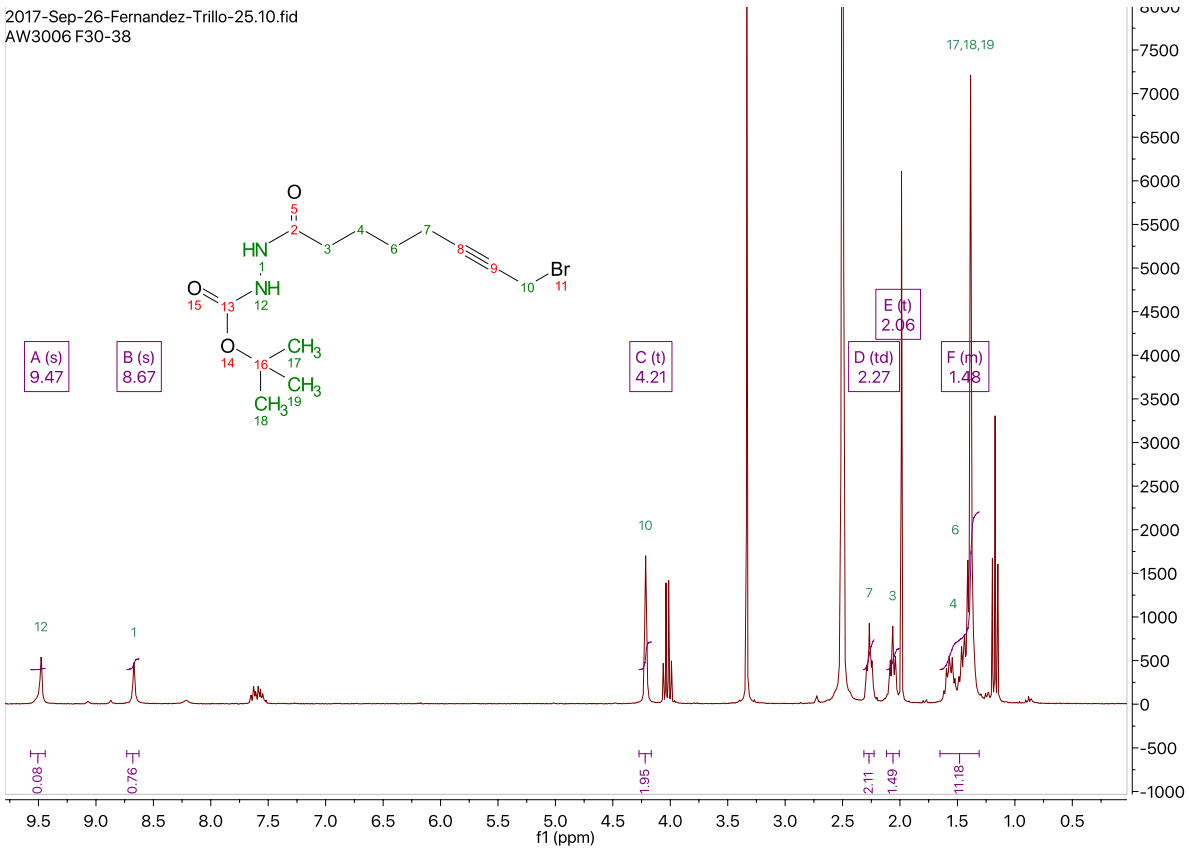


Figure S8 ^{1}H-NMR spectrum of *tert*-butyl-2-(8-bromooct-6-ynoyl)hydrazine-1-carboxylate (**1a**)

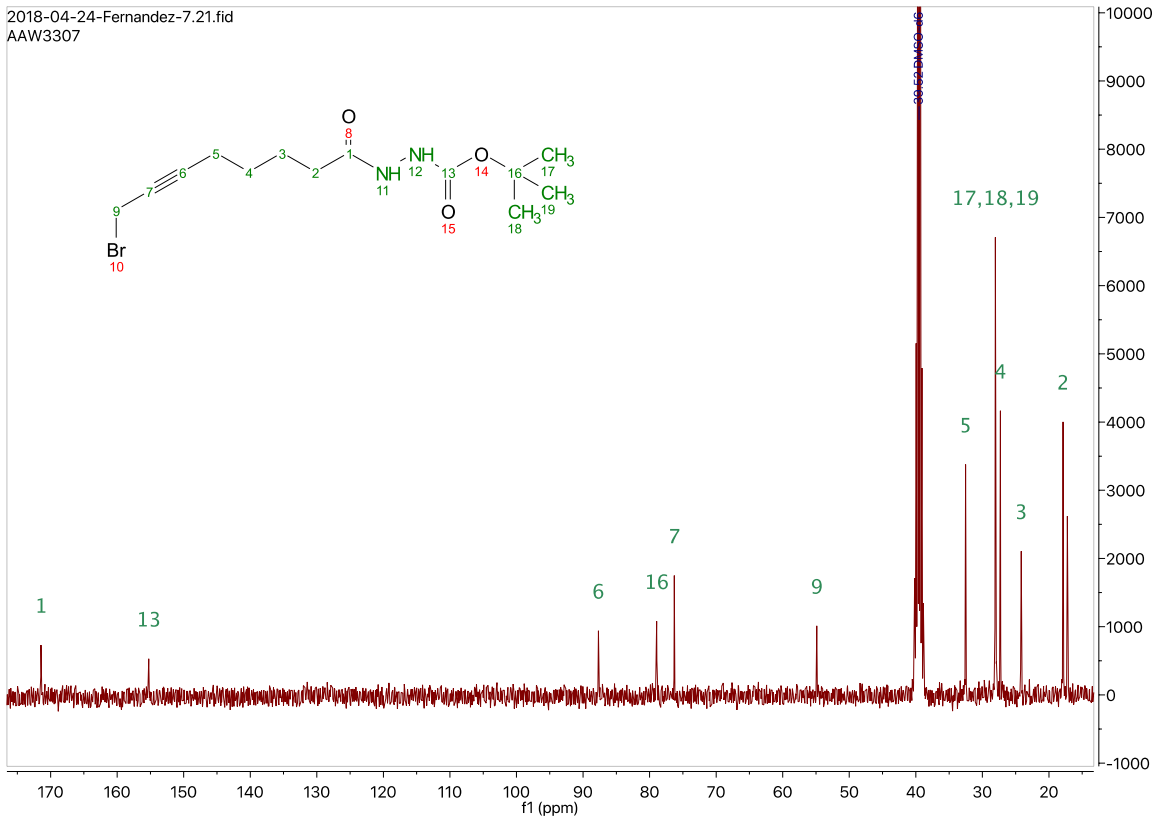
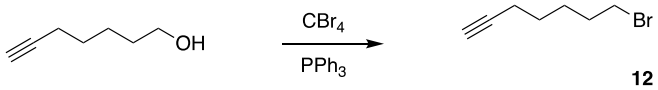


Figure S9 ^{13}C-NMR spectrum of *tert*-butyl-2-(8-bromooct-6-ynoyl)hydrazine-1-carboxylate (**1a**)

3.4. 7-Bromo-hept-1-yne (12)



12

A solution of 6-heptyn-1-ol (5 g, 44.6 mmol) was made in dry DCM (60 mL) and cooled on ice. To this, triphenylphosphine (17.6 g, 67 mmol) was added, upon complete dissolution, tetrabromomethane (22.2 g, 67 mmol) was added slowly. The reaction mixture was brought to room temperature and stirred for 1 h. After completion, the solvent was removed under reduced pressure. Hexane was added to the crude forming a white suspension. The hexane fraction was filtered, collected and then the solvent was removed. An oily residue remained which was purified by flash column chromatography with hexane: Yield = 91%, Rf = 0.45 (hexane); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 3.53 (t, J = 6.7 Hz, 2H), 2.75 (t, J = 2.7 Hz, 1H), 2.23 – 2.10 (m, 2H), 1.89 – 1.74 (m, 2H), 1.50 – 1.43 (m, 4H); ^{13}C-NMR (101 MHz, DMSO) δ 84.3, 71.3, 35.0, 31.7, 27.0, 27.7, 17.6; IR (cm^{-1}) ν max: 3290, 2918, 1437, 1120, 635, 540.

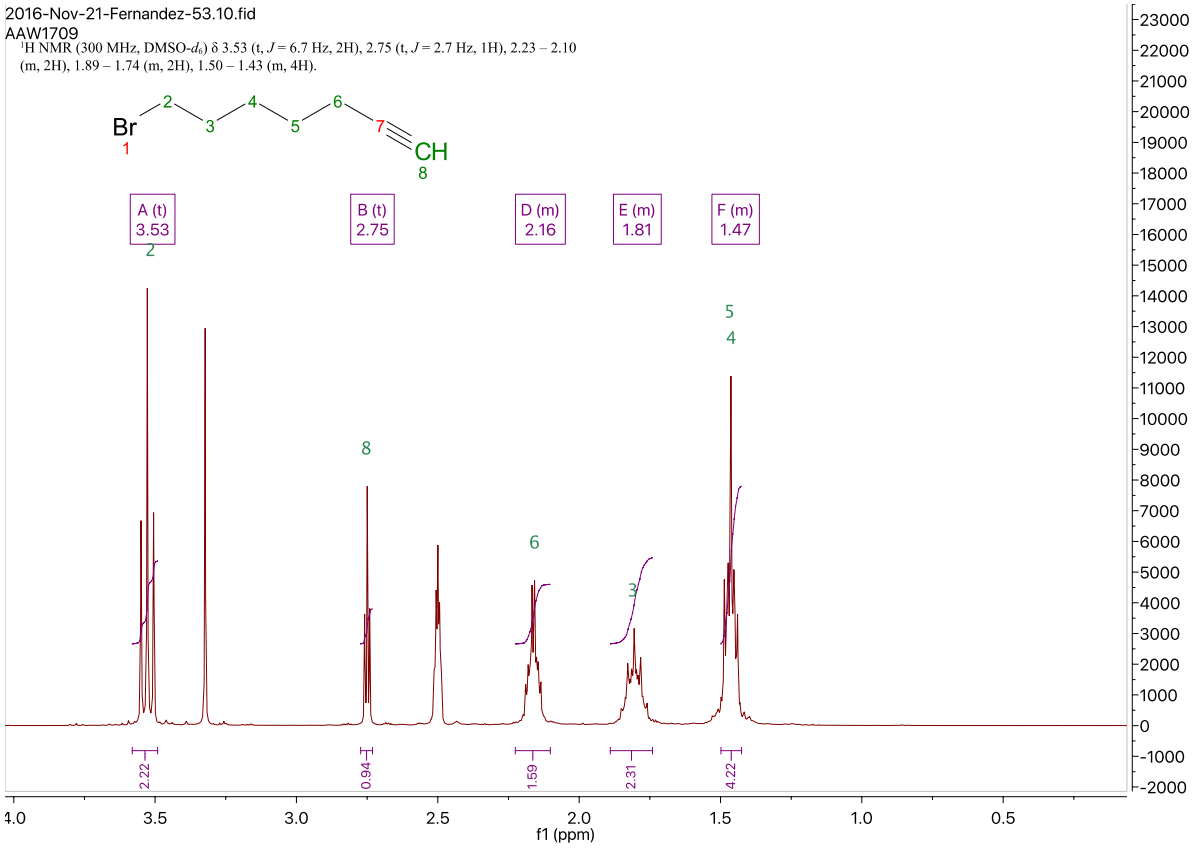


Figure S10 ^{1}H-NMR spectrum of 7-bromo-hept-1-yne (**12**)

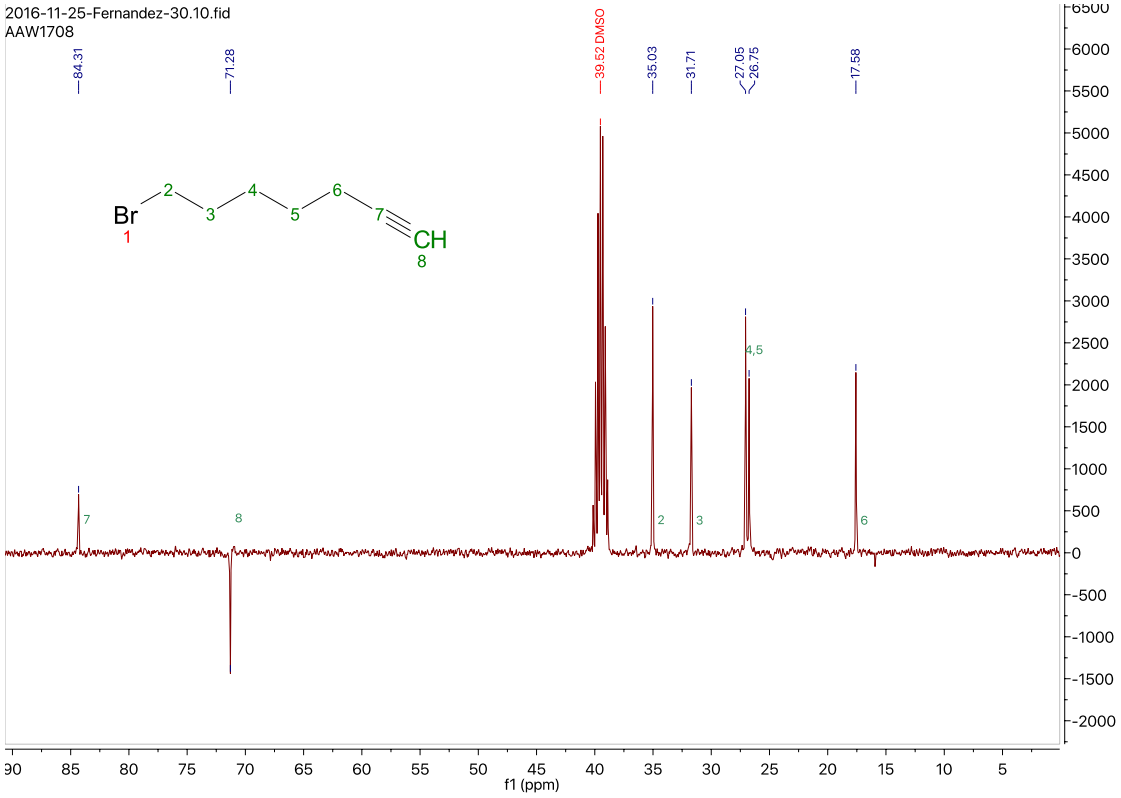


Figure S11 ^{13}C-NMR spectrum of 7-bromo-hept-1-yne (**12**)

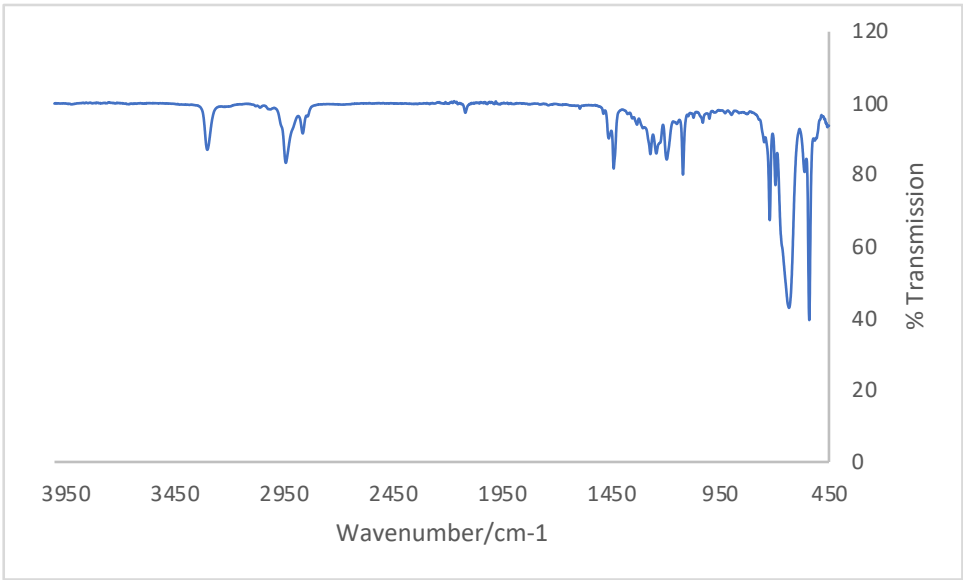
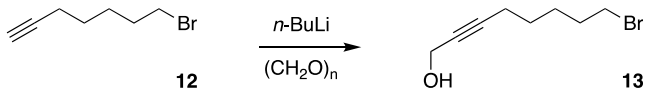


Figure S12 Infrared spectrum of 7-bromo-hept-1-yne (**12**)

3.5. 8-Bromooct-2-yn-1-ol (13)



A solution of 7-bromohept-1-yne (**12**) (20.56 mmol, 3600 mg) was made in dry THF (12.3 mL) and cooled to -78 °C under Argon. To this, a solution of *n*-BuLi in hexanes (1.6 M, 13 mL) was added dropwise, whilst maintaining the temperature below -60 °C. The reaction mixture was then warmed to 0 °C in an ice bath at which point paraformaldehyde (1718 mg, 55.5 mmol) was added under a flow of argon and stirred for 30 minutes. The mixture was then warmed to room temperature and left to stir, while the temperature was maintained below 30 °C until the exothermic reaction had stopped. The mixture was

then heated to 45 °C for 2 h. Once complete the reaction was extracted with ether and sat. NH_4Cl. The organic layer was collected, dried over anhydrous sodium sulfate and the solvents were removed under reduced pressure to yield the crude product as an oil. Purification was completed by flash column chromatography (Hex:EtOAc 9:1). The product was then collected as a colorless oil: Yield = 55%; Rf = 0.15 (Hex:EtOAc 9:1); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 5.04 (t, J = 5.7 Hz, 1H), 4.03 (dt, J = 5.5, 2.1 Hz, 2H), 3.54 (t, J = 6.7 Hz, 2H), 2.20 (m, 2H), 1.88 – 1.75 (m, 2H), 1.52 – 1.40 (m, 4H); ^{13}C-NMR (101 MHz, $CDCl_3$) δ 86.1, 78.8, 77.2, 51.5, 33.7, 32.4, 27.8, 27.5, 18.7.

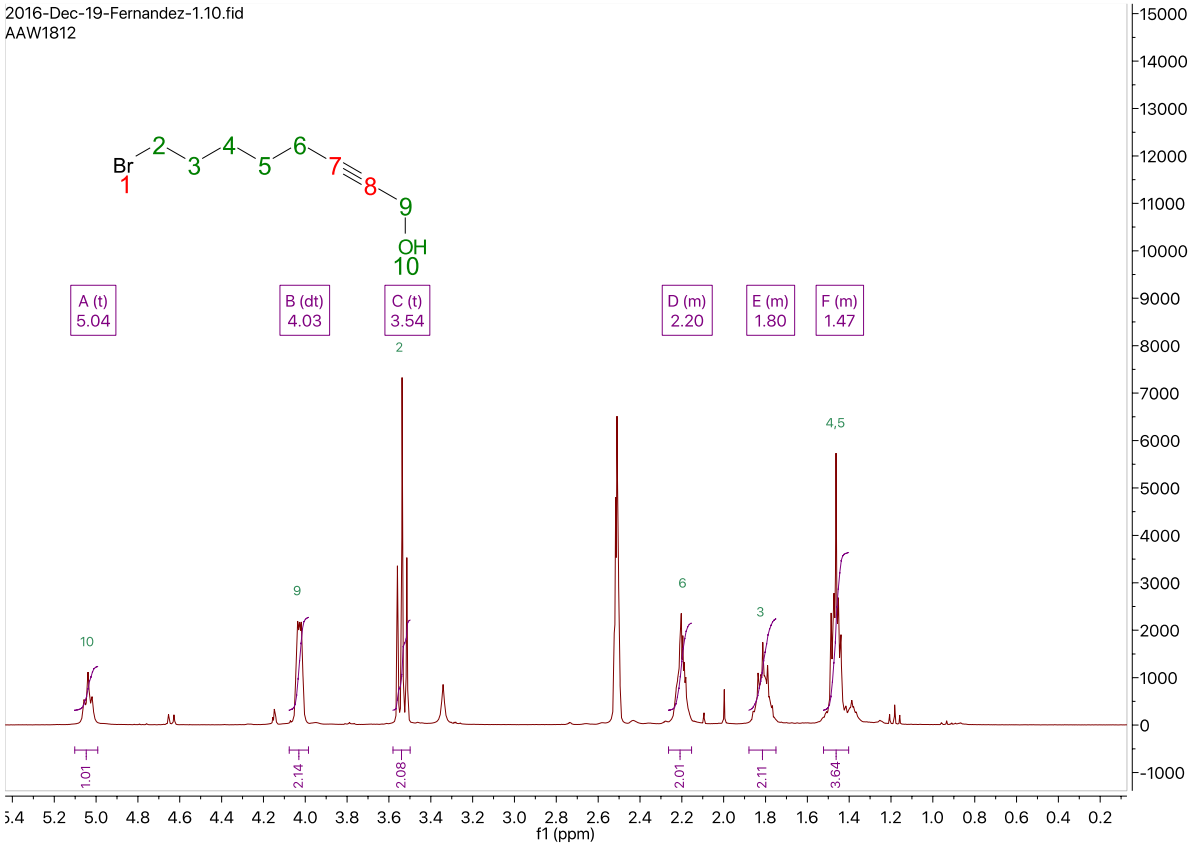


Figure S13 ^{1}H-NMR spectrum of 8-bromooct-2-yn-1-ol (**13**)

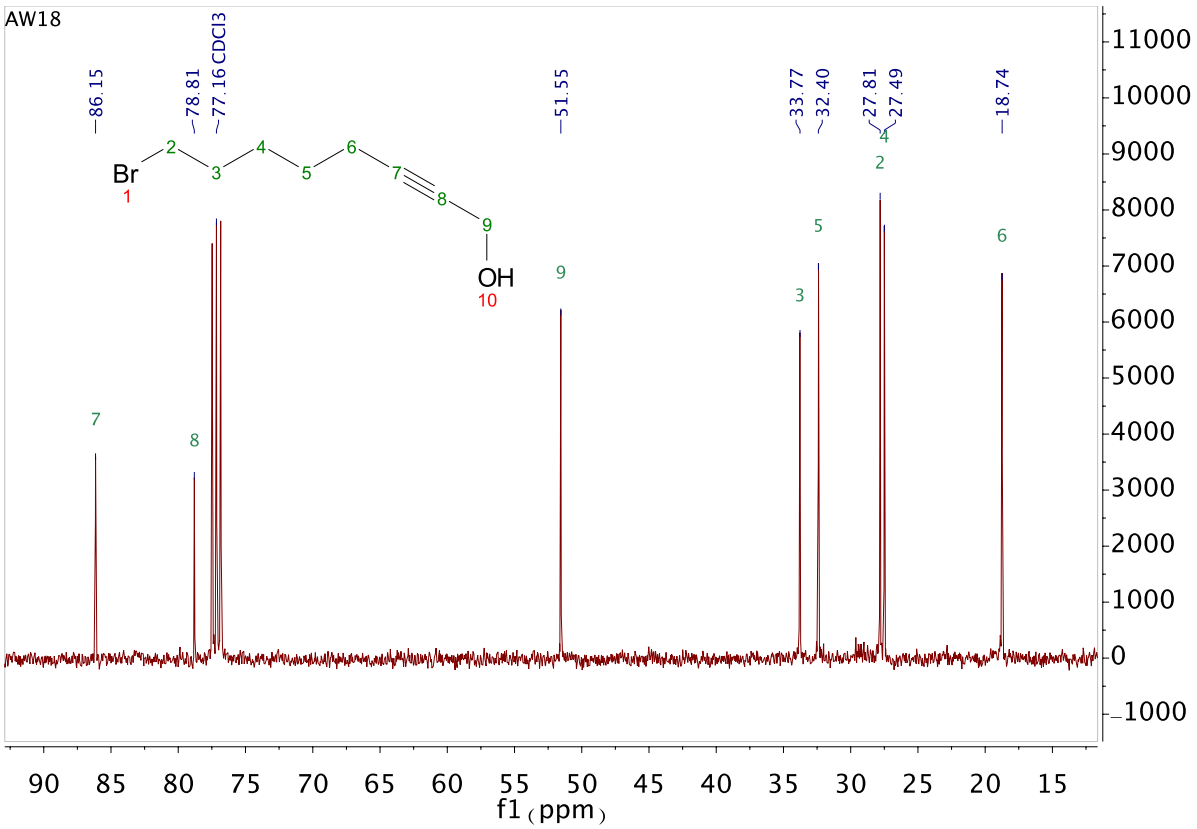
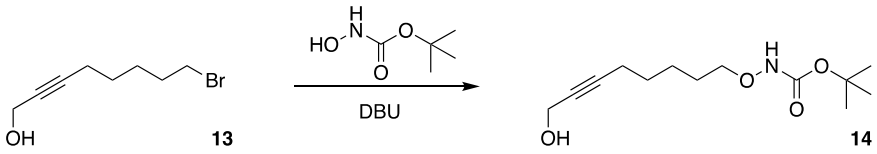


Figure S14 ^{1}H-NMR spectrum of 8-bromooct-2-yn-1-ol (**13**)

3.6. *tert*-Butyl-((8-hydroxyoct-6-yn-1-yl)oxy)carbamate (14)



To a solution of *N*-Boc-hydroxylamine (890 mg, 6.55 mmol) in DMF (4.3 mL), 8-bromooct-2-yn-1-ol (**13**) (1200 mg, 5.85 mmol) and 1,8-Diazabicyclo[5.4.0]undec-7-ene (1000 mg, 6.55 mmol) were added. The solution was stirred at 50 °C for 20 h. Once complete, the reaction was extracted with DCM and 15% citric acid solution. The organic phases were dried over anhydrous sodium sulfate, collected and the solvent was removed under reduced pressure. A colorless oil was collected as the crude product. This was further purified by flash column chromatography (Hex:EtOAc 8:2). The product was collected as a colorless oil: Yield = 73 %; Rf = 0.27 (Hex:EtOAc 8:2); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 9.91 (s, 1H), 5.03 (t, *J* = 5.9 Hz, 1H), 4.02 (dt, *J* = 5.9, 2.2 Hz, 2H), 3.66 (t, *J* = 6.2 Hz, 2H), 2.17 (tt, *J* = 6.7, 1.7 Hz, 2H), 1.40 (m, 15H); ^{13}C-NMR (101 MHz, DMSO) δ 156.1, 84.0, 80.4, 79.4, 75.1, 49.1, 28.1, 28.0, 27.1, 24.8, 17.9; MS: m/z [M+H] = 258.2.

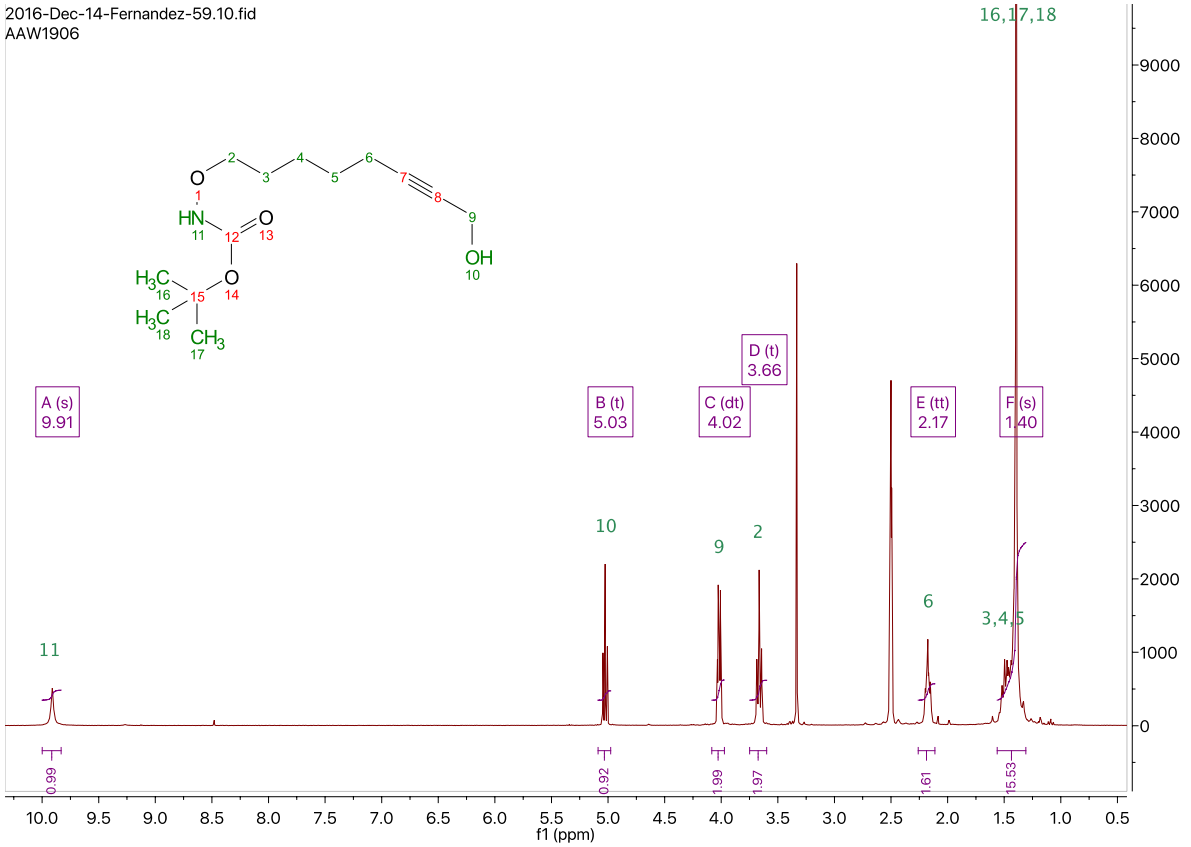


Figure S15 ^{1}H-NMR spectrum of *tert*-butyl-((8-hydroxyoct-6-yn-1-yl)oxy)carbamate (**14**)

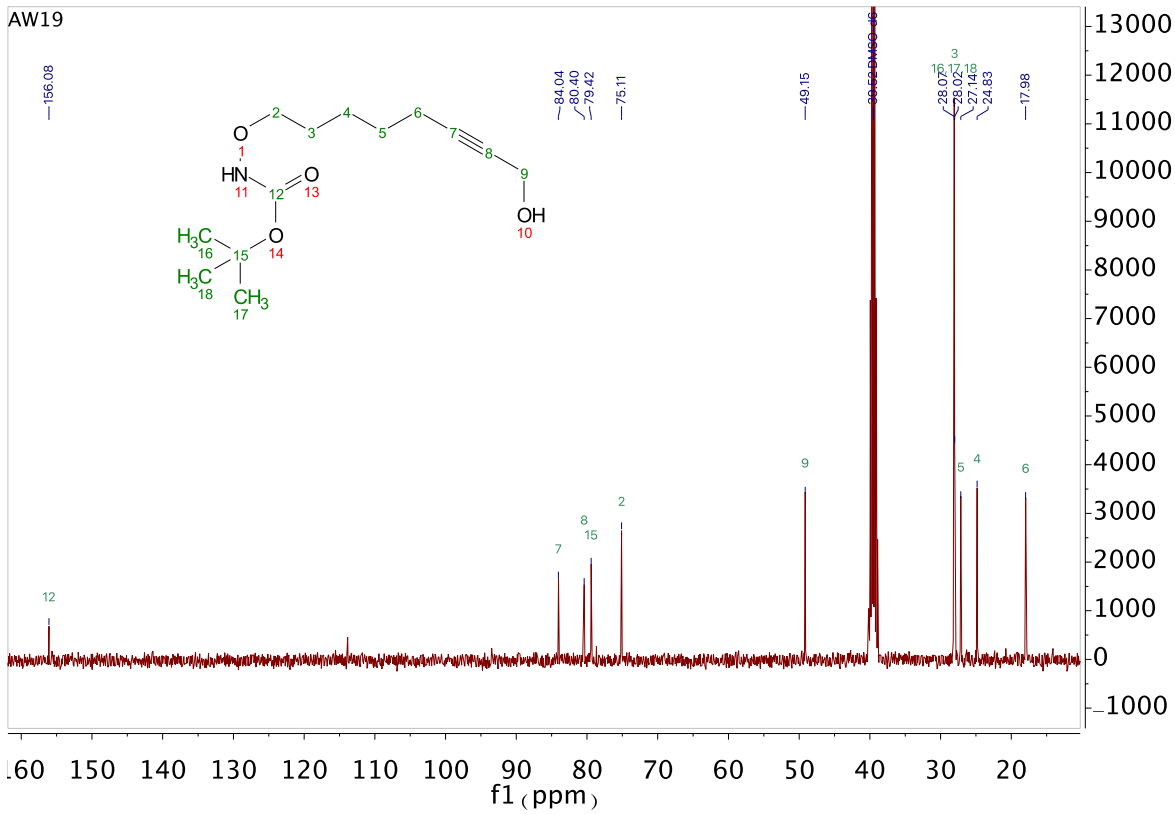
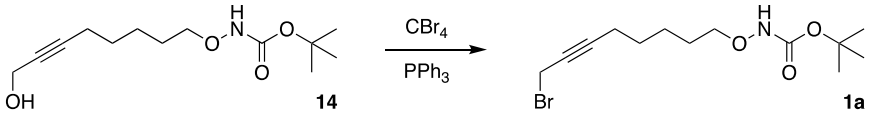


Figure S16 ^{13}C-NMR spectrum of *tert*-butyl-((8-hydroxyoct-6-yn-1-yl)oxy)carbamate (**14**)

3.7. *tert*-Butyl-((8-bromooct-6-yn-1-yl)oxy)carbamate (1b)



A solution of *tert*-butyl((8-hydroxyoct-6-yn-1-yl)oxy)carbamate (**14**) (1 g, 3.89 mmol) was made in dry DCM (5.2 mL) and cooled on ice. To this, triphenylphosphine (1.53 g, 67 mmol) was added. Upon complete dissolution tetrabromomethane (1.94 g, 67 mmol) was added slowly. The reaction mixture was brought to room temperature and allowed to stir for 1 h . After completion, the solvent was removed under reduced pressure. Purification was completed using flash column chromatography (Hex:EtOAc 8:2): Yield = 67%; Rf 0.52 (Hex:EtOAc, 8:2); ^{1}H-NMR (300 MHz, DMSO-d_6) δ 9.90 (s, 1H), 4.21 (t, J = 2.4 Hz, 2H), 3.66 (t, J = 6.2 Hz, 2H), 2.25 (tt, J = 6.9, 2.4 Hz, 2H), 1.40 (m, 15H); ^{13}C-NMR (101 MHz, DMSO) δ 156.0, 87.8, 79.3, 76.2, 75.0, 39.5, 28.0, 27.6, 27.0, 24.7, 18.0, 17.2; IR (cm^{-1}) ν max: 1712, 607; MS: m/z [M+Na] = 342.35/344.35, [M-^{t}BuOH] = 246.38/248.38.

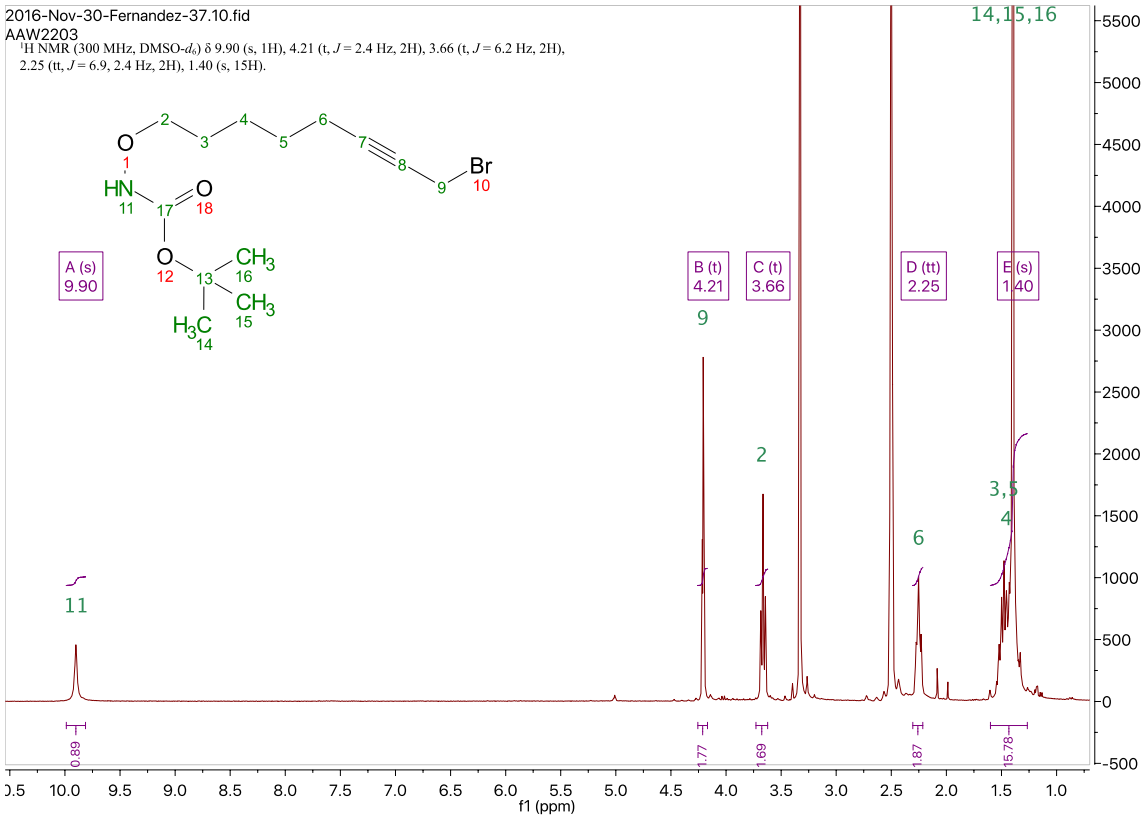


Figure S17 ^{1}H-NMR spectrum of *tert*-butyl-((8-bromooct-6-yn-1-yl)oxy)carbamate (**1b**)

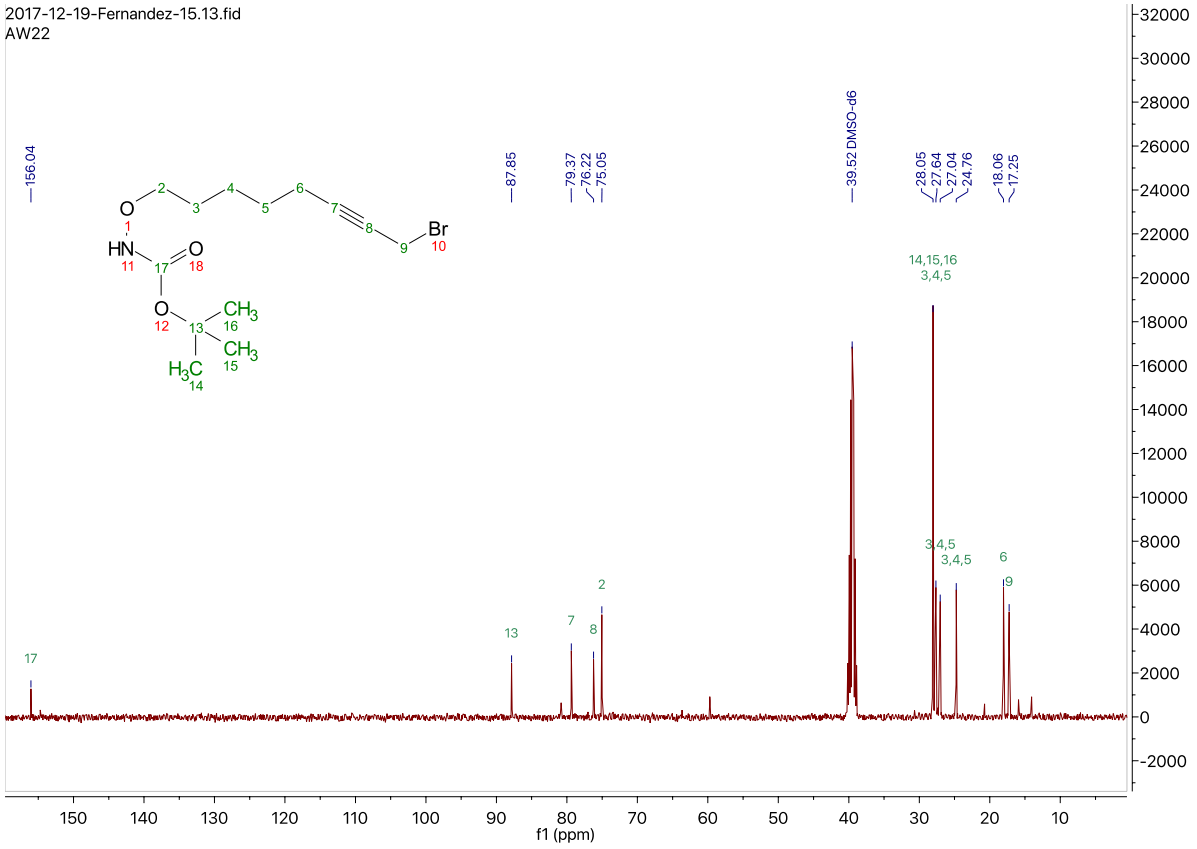


Figure S18 ^{13}C-NMR spectrum of *tert*-butyl-((8-bromooct-6-yn-1-yl)oxy)carbamate (**1b**)

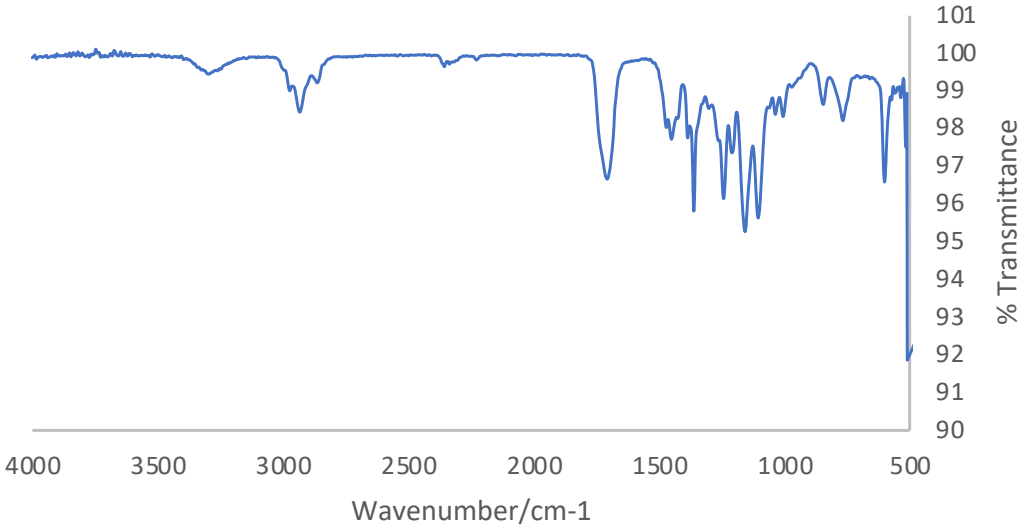


Figure S19 Infrared spectrum of *tert*-butyl-((8-bromooct-6-yn-1-yl)oxy)carbamate (**1b**)

3.8. Synthesis of 2a and 2b: General Procedure for the Coupling of Bromides 1 with AdoHcy

A solution of S-adenosyl-L-homocysteine (**AdoHcy**) (15 mg, 0.04 mmol) was made in a 1:1 mixture of formic and acetic acid (300 µL). Linker **1** (1.2 mmol, 30 equiv.) was then added dropwise, on ice. The reaction mixture was warmed to 35 °C and left to stir overnight. After overnight stirring, the reaction mixture was extracted with diethyl ether and the aqueous layer was collected and dried by lyophilization: MS: m/z [M+H] = 638, HRMS-ESI(+): calcd for $C_{23}H_{34}N_7O_6S$: 536.2291, found: 536.2302 [M+H-Homocysteine] (**2a**), [M+H] = 624 (**2b**). Compounds **2** were immediately used in the next step without further purification.

3.9. Synthesis of Cofactors 3a and 3b: Cofactor Deprotection

The crude product (**2a** or **2b**) was dissolved in TFA (400 µL) and left to stir for 2 h at room temperature. After reaction, the acid was removed under a flow of argon. The crude reaction mixture was then dissolved in water (2 mL) and purified by Prep-HPLC. Retention times: **3a** isomer I = 17.51 min, isomer II = 18.73 min, **3b** isomer I = 25.47 min, isomer II = 28.24 min: MS: m/z [M+H] = 538, HRMS-ESI(+): calcd for $C_{18}H_{26}N_7O_4S$: 436.1767, found: 436.1768 [M+H-Homocysteine] (**3a**), [M+H] = 524 (**3b**). Compounds **3** were immediately used in the next step without further characterization.

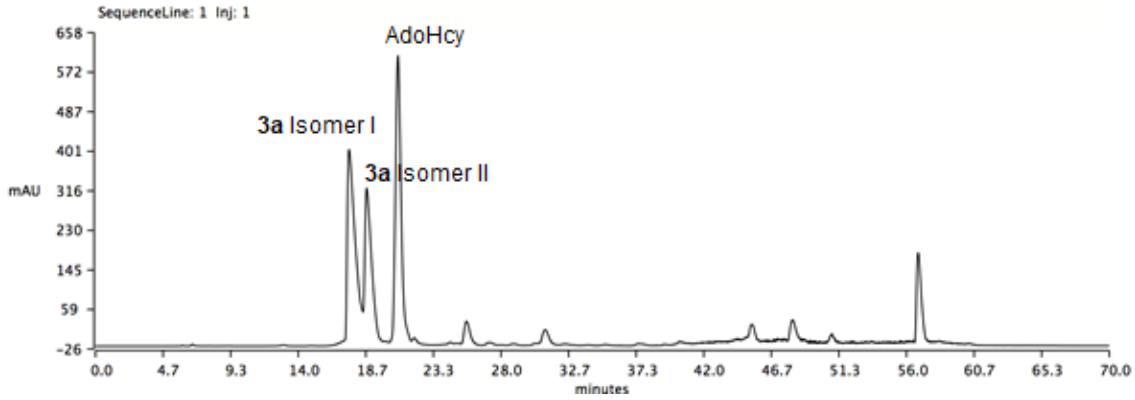


Figure S20 Preparative-HPLC chromatogram of cofactor **3a**. Conditions for analytical HPLC: 3-30% over 30 minutes, 30-97% over 20 minutes, hold 97% 5 minutes. 20 mM ammonium acetate pH 5.5. Flow rate 10 mL/min

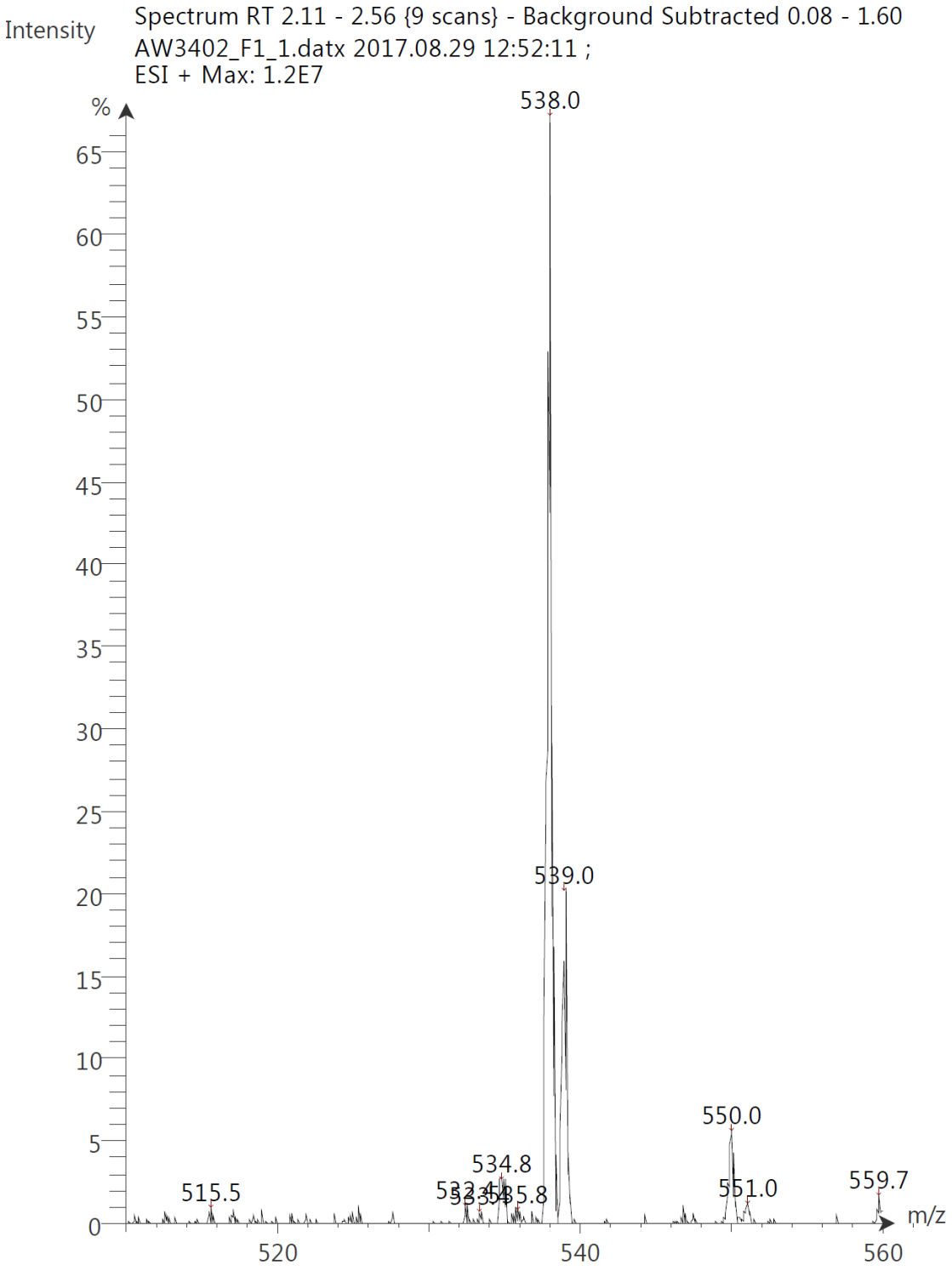


Figure S21 Mass spectrum of the peak at 17 min from Figure S20. 538.0: M + H^+, 550.0: M + Na^+.

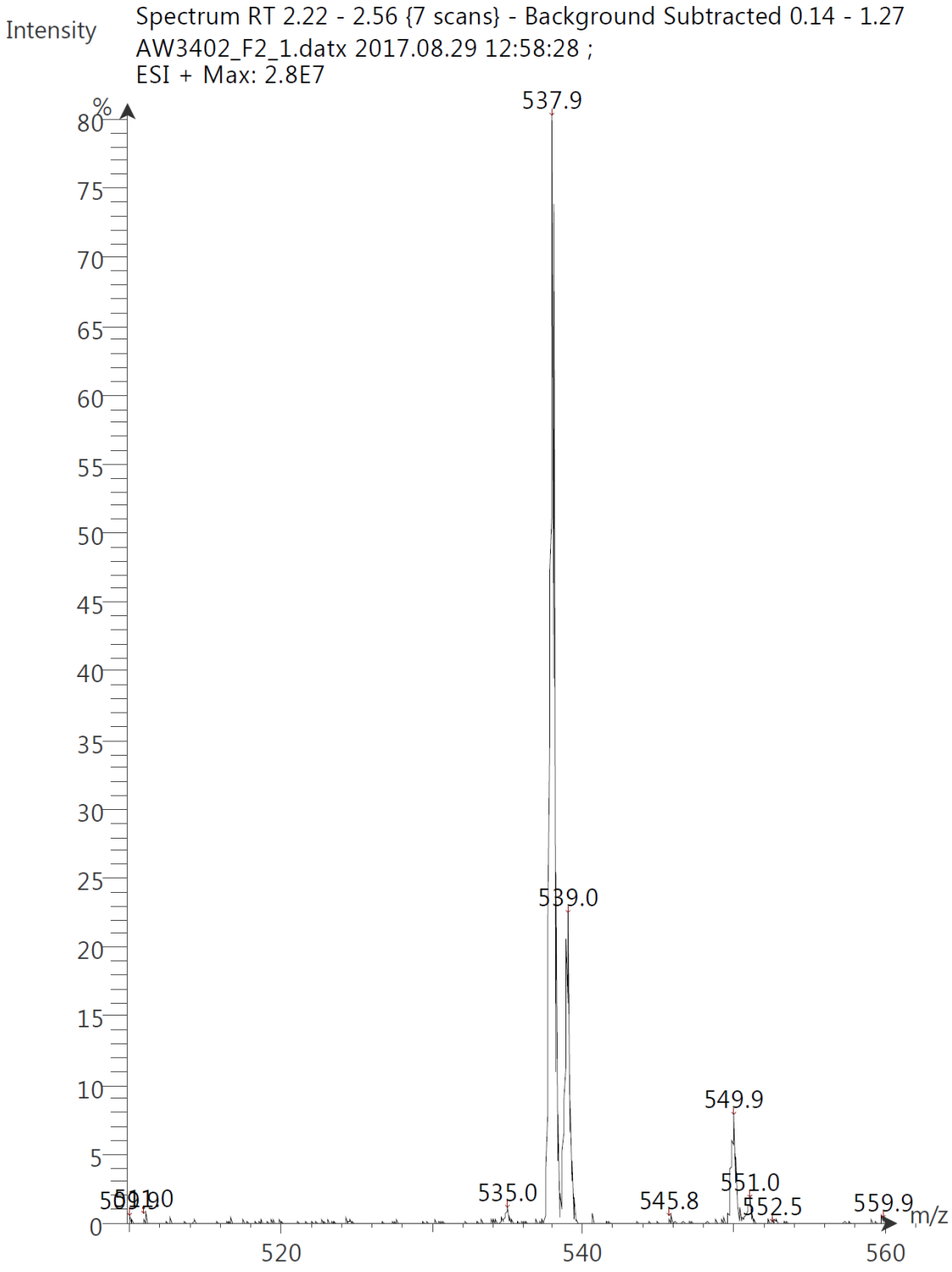


Figure S22 Mass spectrum of the peak at 18.5 min from Figure S20. 537.9: M + H^+, 549.9: M + Na^+.

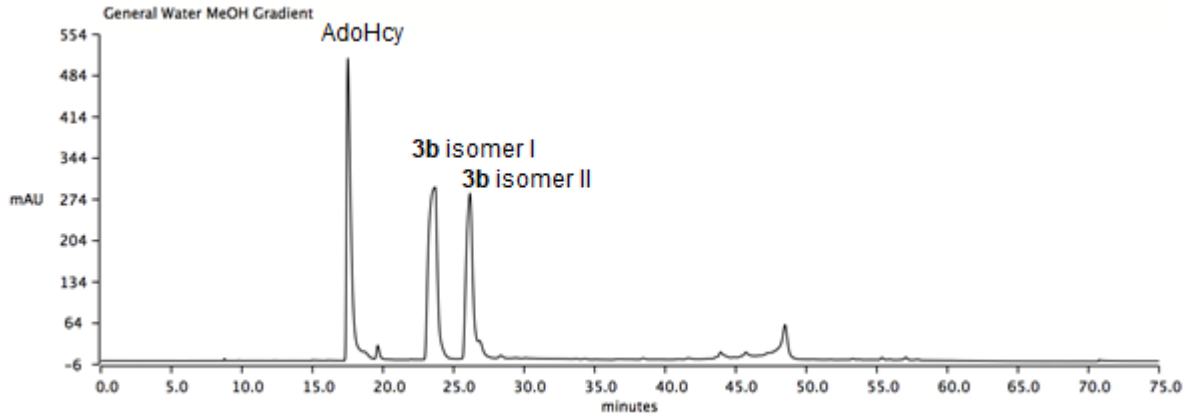


Figure S23 Preparative-HPLC chromatogram of cofactor **3b**. Conditions for analytical HPLC: 3-30% over 30 minutes, 30-97% over 20 minutes, hold 97% 5 minutes. 20 mM ammonium acetate pH 5.5. Flow rate 10 mL/min.

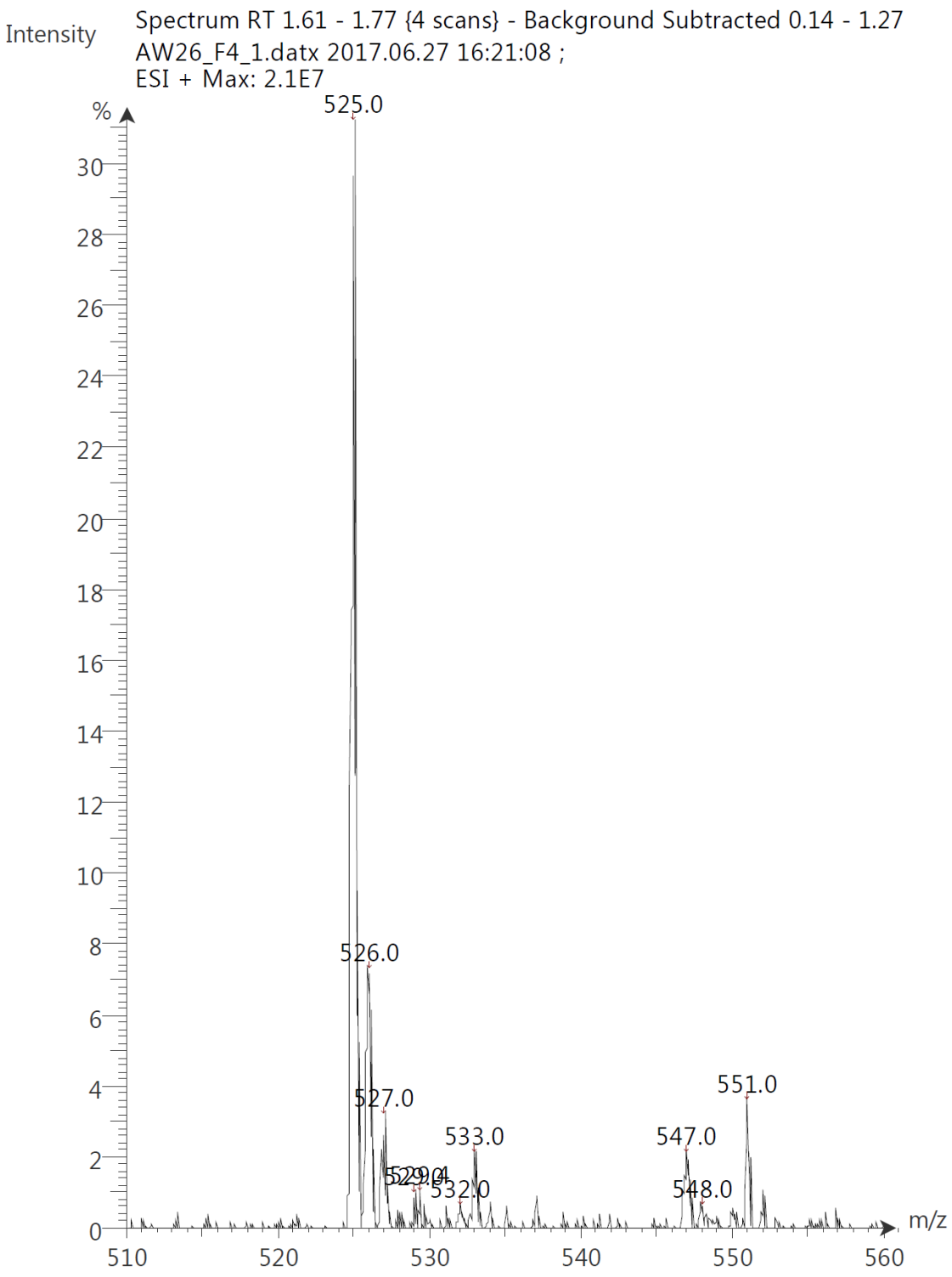


Figure S24 Mass spectrum of the peak at 24.5 min from Figure S23. 525.0: M + H^+.

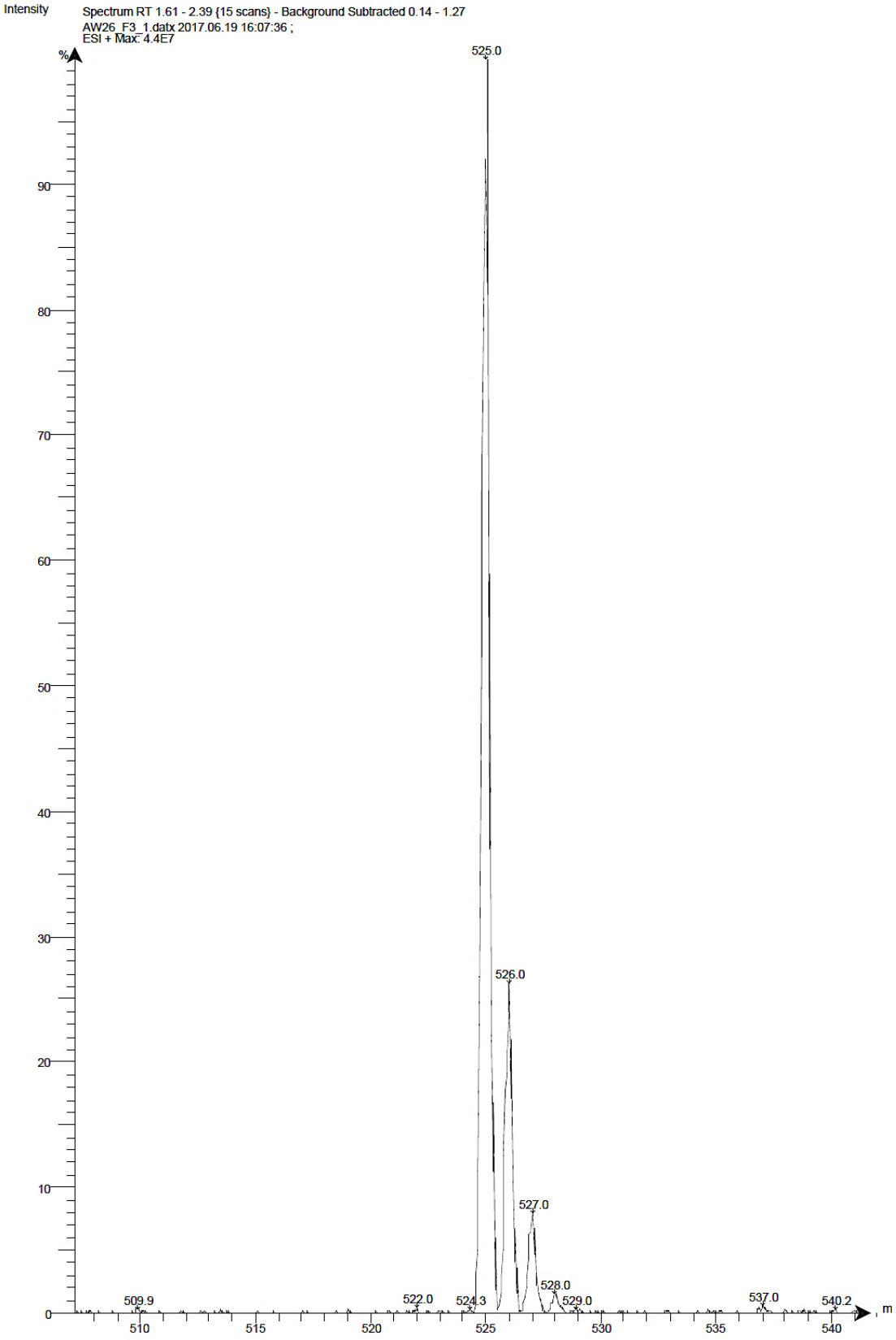


Figure S25 Mass spectrum of the peak at 26 min from Figure S23. 525.0: M + H^+.

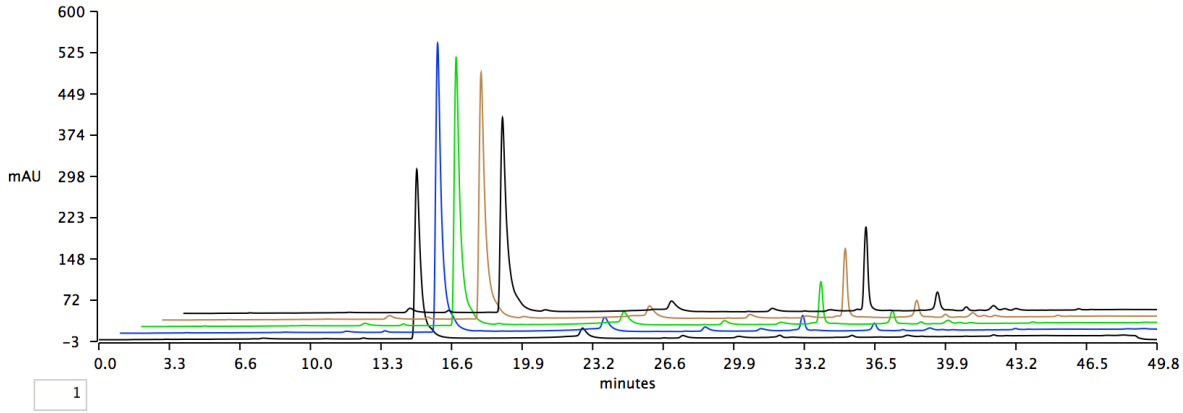


Figure S26 Consecutive analytical HPLC chromatograms of cofactor **3a** after incubation at 37 °C (Isomer I) following isolation via preparative HPLC. Conditions for analytical HPLC: 3-30% over 20 minutes, 30-97% over 15 minutes, hold 97% for 10 minutes. 10 mM ammonium acetate pH 5.5. Flow rate 1 mL/min. 14 min = **3a** (Isomer I), 33 min = **3a** (Isomer I) – Homoserine lactone.

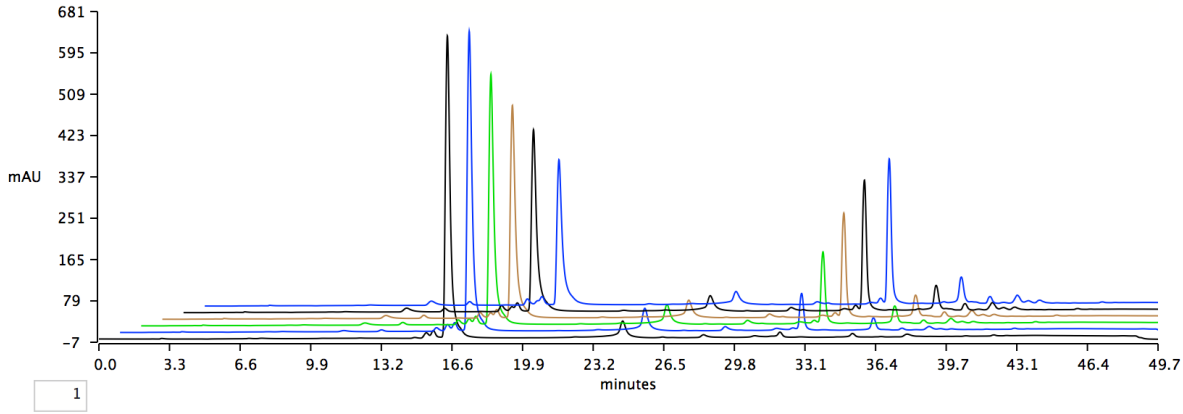


Figure S27 Consecutive analytical HPLC chromatograms of **AdoMet** derivative **3a** after incubation at 37 °C (Isomer II) following isolation via preparative HPLC. Conditions for analytical HPLC: 3-30% over 20 minutes, 30-97% over 15 minutes, hold 97% for 10 minutes. 10 mM ammonium acetate pH 5.5. Flow rate 1 mL/min. 16.5 min = **3a** (Isomer I), 33 min = **3a** (Isomer I) – Homoserine lactone.

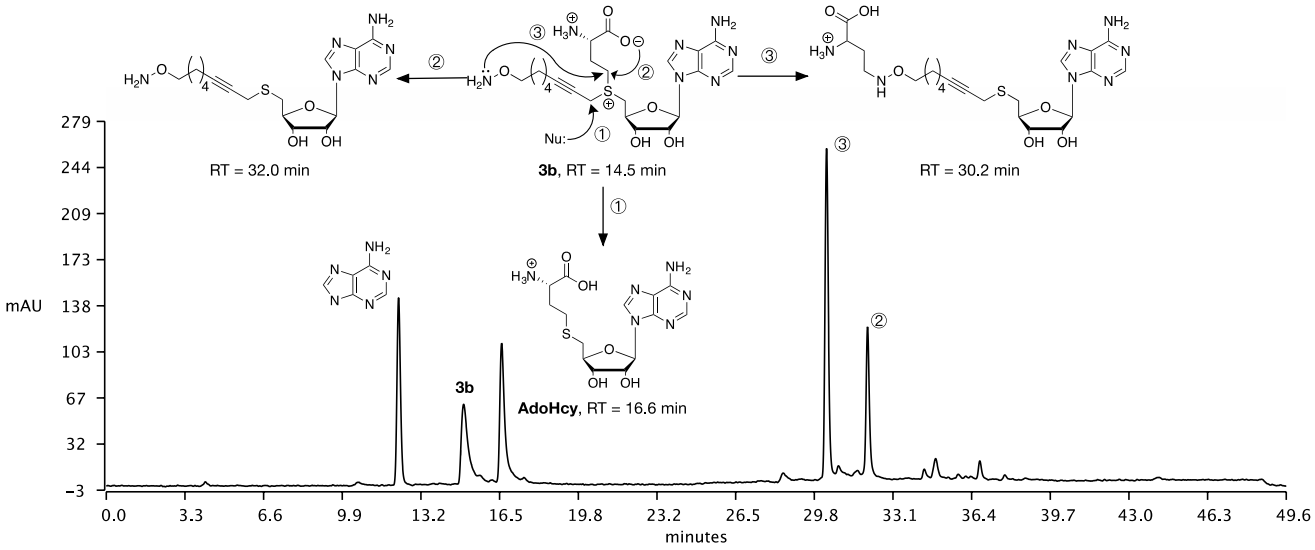


Figure S28 Analytical HPLC chromatogram of **AdoMet** derivative **3b** following isolation via preparative HPLC and freeze-drying, and proposed degradation pathways. Conditions for analytical HPLC: 3-30% over 20 minutes, 30-97% over 15 minutes, hold 97% for 10 minutes. 10 mM ammonium acetate pH 5.5. Flow rate 1 mL/min.

3.10. Synthesis of Cofactors 4a and 4b: Aldehyde Coupling

Crude deprotected cofactors **3** were mixed, directly after collection from preparative HPLC, with aldehyde **5** (1.2 equiv.) and the mixture rolled for 30 min at room temperature. The reaction mixture was then dried by lyophilization,[3] and then resuspended in 100 µL 0.1% Acetic Acid and stored at -20 °C without further purification. MS: m/z [M+H] = 869 (**4a**), [M+H] = 856 (**4b**); HRMS-ESI(+): calcd for C38H54N11O10S: 856.3776, found: 856.3780 [M+H]+ (**4a**). Calcd for C38H53N12O10S: 869.3728, found: 869.3729 [M+H]+ (**4b**). HPLC retention times: 33.9 min (**4a**), 31.9 (**4b** iso 1), 32.0 (**4b** iso 2).

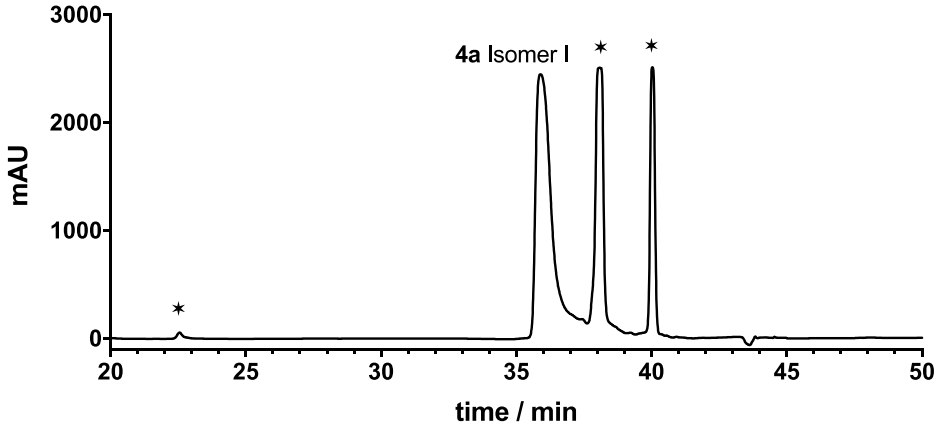


Figure S29 Analytical HPLC chromatogram of cofactor **4a** (Isomer I) after freeze-drying. Conditions for analytical HPLC: 0.1 mL/min 3-100% over 60 min. 0.1% formic acid water methanol gradient. Excess aldehyde 5 at 40 min. * correspond to peaks from aldehyde **5**.

[3] Acetone should be avoided during lyophilization, as exchange with the acyl hydrazone/oxime moiety can be observed.

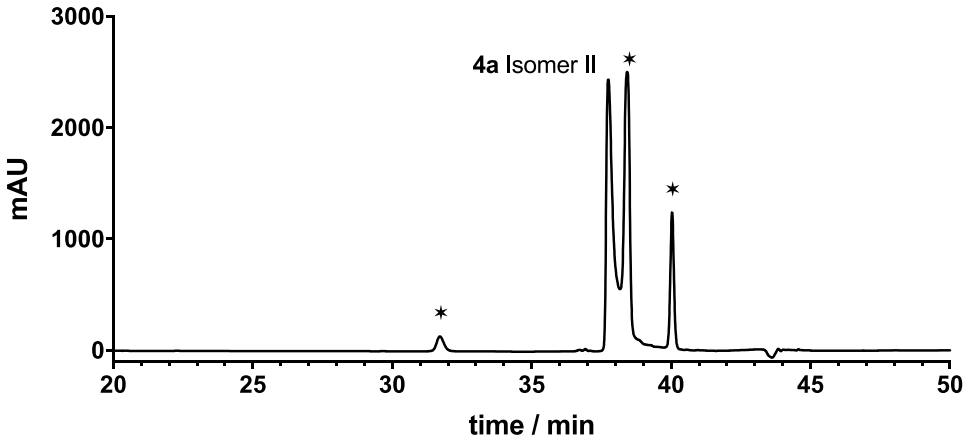


Figure S30 Analytical HPLC chromatogram of cofactor **4a** (Isomer II) after freeze-drying. Conditions for analytical HPLC: 0.1 mL/min 3-100% over 60 min. 0.1% formic acid water methanol gradient. Excess aldehyde **5** at 40 min. * correspond to peaks from aldehyde **5**.

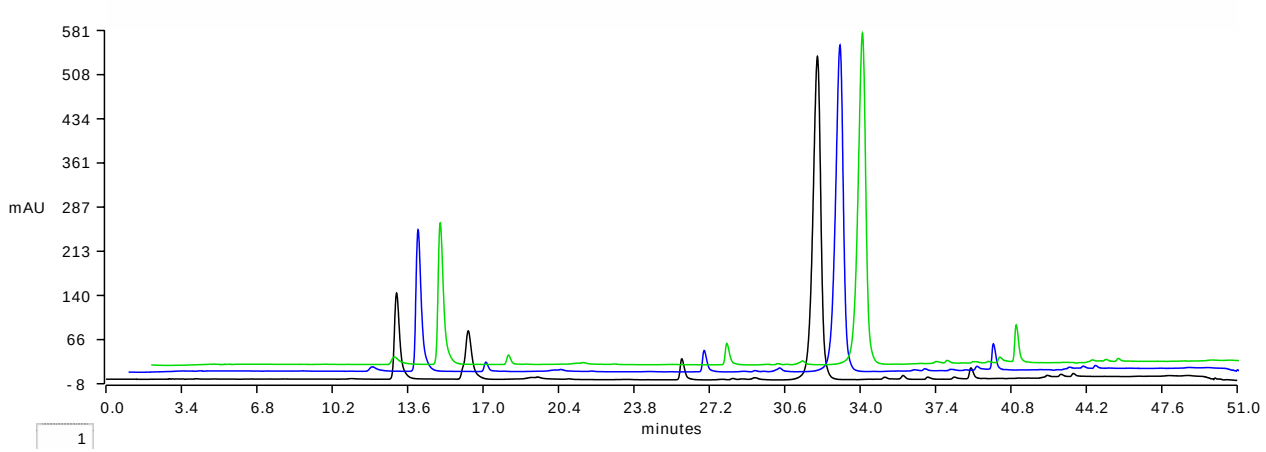


Figure S31 Consecutive analytical HPLC chromatograms of **AdoMet** derivative **4b** after incubation at 50 °C (Isomer II). Conditions for analytical HPLC: 3-30% over 20 minutes, 30-97% over 15 minutes, hold 97% for 10 minutes. 10 mM ammonium acetate pH 5.5. Flow rate 1 mL/min. 13 min = Adenine, 31 min = **4b** (Isomer II), 38 min = **3a** (Isomer I) – Homoserine lactone.

4. MTase Expression and Purification

4.1. M.TaqI

Escherichia coli T7 Express Competent (High Efficiency) (C2566) cells were transformed with pET28a::M.TaqI and allowed to recover in 495 µL SOC, 250 rpm and 37°C, for 30 minutes to 1 h. 150 µL of recovered cells were spread onto a sterile LB-agar (Melford) plate supplemented with kanamycin to a final concentration of 50 µg/mL and grown overnight at 37 °C. A colony was picked and used to inoculate 10 mL LB media (Melford) supplemented with kanamycin final concentration of 50 µg/mL and left shaking overnight at 37 ^{0}C 250 rpm. 5 mL of overnight culture was used to inoculate 2x 400 mL of LB media (Melford), supplemented with kanamycin to a final concentration of 50 µg/mL and left shaking 180 rpm at 37 °C until an OD600 reading of 0.4-0.6 was reached. IPTG was added to a final concentration 0.5 mM and left shaking at 20 ^{0}C, 180 rpm for 16 h. Cells were harvested by centrifugation (20 minutes at 3993 xg, 4 °C). The supernatant was discarded, and the pelleted cells were resuspended in 20 mL of 1% PBS supplemented with EDTA-free Protease Inhibitor Cocktail (Sigma) (pH 7.4). Resuspended cells

were then disrupted using an emulsiflex at a pressure of 12K psi and passed through the funnel repeatedly for a total of 5x. The resulting cell lysate was then centrifuged at 13500 rpm, 4 °C for 14 minutes. The supernatant was collected, and further EDTA-free Protease Inhibitor Cocktail tablets added. The supernatant was then passed through a 0.22 μM Sartorius syringe filter. 1 mL of Ni-NTA agarose beads were centrifuged at 500 xg, for 5 minutes. The supernatant was removed, 1 mL of 1% PBS was added and the beads were washed by inverting several times followed by centrifuged at 500 xg for 5 minutes and removal of the PBS (repeated 5x). A final 1 mL of 1% PBS was added to the beads to create a 50% slurry, 50 μL of which was mixed per 1 mL of cell lysate and left rotating end over end for 1 h at 4 °C. Biorad EconoColumns were washed with 1x PBS and the sample was added to the column. Next, 25 mL of wash buffer (PBS, 1 M NaCl, 20 mM imidazole) followed by 5 mL elution buffer (PBS, 200 mM NaCl, 250 mM imidazole) and left for 10 minutes, then eluted. An Amicon Ultra 0.5 mL 10 KDa kit was used to exchange the buffer following the manufacturer's instructions. Storage buffer was made up of PBS, 5 mM EDTA and 5 mM β-mercaptoethanol. Protein presence was checked using SDS-PAGE and stored in 50% glycerol at -20 °C.

4.2. M.MpeI

E. coli T7 Express Competent (High Efficiency) (C2566) cells were transformed with pET28a::mutated mpeIM (M.Mpe1) and allowed to recover in 200 μL of LB (Melford) at 37 °C, 200 rpm for 1 h. 50 μL of recovered cells were spread onto a sterile LB-agar (Melford) plate supplemented with kanamycin to a final concentration of 50 μg/mL and grown overnight at 37 °C. A colony was selected and used to inoculate 25 mL of sterile LB (Melford) supplemented with kanamycin (50 μg/mL), grown overnight at 37 °C, 200 rpm. From this, 20 mL was added to a 5 L flask containing 2 L of LB (Melford) supplemented with kanamycin (50 μg/mL) and grown at 37 °C, 180 rpm until an OD600 reading of 0.4-0.8 was reached. The temperature was then reduced to 28 °C and the culture was supplemented with IPTG to a final concentration of 0.5 mM and incubated for 10 h at 180 rpm. The culture was then split into 2x 1 L centrifuge containers and harvested by centrifugation (20 minutes at 3993 x g, 4 °C). Cell pellets were washed in 50 mL of pre-chilled phosphate buffered saline (PBS) and re-harvested (20 minutes at 3993 x g, 4 °C). Cell pellets were then resuspended in 20 mM Tris-HCl, 40 mM imidazole, 0.5 mM TCEP, EDTA-free Protease Inhibitor Cocktail (Sigma) (pH 7.4) and 1 mg/mL of DNAase I and cells were lysed using an EmulsiFlex-C3 (Avastin). Cell lysate was cleared via ultracentrifugation (30 min at 48,384 x g, 4 °C) and the supernatant collected and filtered (0.45 μM filter). A 5 mL HisTrap HP column (GE healthcare) was equilibrated with 25x column volumes (CV)) of equilibration buffer (20 mM Tris-HCl, 40 mM imidazole, 0.5 mM TCEP, EDTA-free Protease Inhibitor Cocktail (Sigma)). The supernatant was added to the column and allowed to circulate overnight at 4 °C. A HisTrap HP column was then loaded onto an AKTA and the column was washed with 30x CV of equilibration buffer followed by 30x CV of high salt wash buffer (20 mM Tris-HCl, 40 mM imidazole, 0.5 mM TCEP, 500 mM NaCl and EDTA-free Protease Inhibitor Cocktail (Sigma)). M.MpeI was then eluted into 20 mM Tris-HCl, 0.5 mM TCEP, 500 mM NaCl, EDTA-free Protease Inhibitor Cocktail (Sigma) using a 0-500 mM gradient of imidazole, collected into 2 mL fractions. Fractions containing the most protein were pooled, and buffer exchanged into 10 mM Tris-HCl, 0.5 mM TCEP, 1 mM EDTA (pH 7.5) and 50% glycerol. A final protein concentration of 1.7 mg/mL was achieved.

5. Gel Electrophoresis

5.1. General procedure

On ice, a master mix was created by mixing 79.5 μL molecular grade water, 10 μL (10x NEB CutSmart buffer) and 3 μL pUC19 (1000 ng/μL). 2x 9.25 μL was taken for tubes 4, 8 (restriction controls) and 0.25 μL of each cofactor isomer and 0.5 μL water was added. To the remaining master mix 4 μL M.TaqI (0.3 mg/mL) was added and mixed. The master mix was then split into 2x 19.5 μL (1,5) and 4x 10 μL and labeled (2,3,6,7). A 2x serial dilution was made by adding 0.5 μL **4a** to the final concentration 300 μM to tubes 1/5 and mixed. 10 μL was then taken from tubes 1/5 and added to 2/6 and continued until tubes 6/7 discarding the final 10 μL. Additional controls for this experiment were also set up (Table 1). All samples were incubated at 50 °C for 1 h before adding 0.5 μL restriction enzyme (R.TaqI) to all tubes

except 9. Samples were again incubated for 1 h at 50 °C. 0.5 µL proteinase K was added to all tubes and incubated at 50 °C for 1 h before being run on 1% agarose gel to analyze. A 2-Log DNA Ladder (0.1–10.0 kb) (NEB) was used in all the gels. Relevant bands have been labeled.

Table 1: Controls set up for all protection assays (Lanes are representative for Figure S33)

	AdoMet control (9/10)	No cofactor (11)	No MTase (12)
10x NEB CutSmart	2 µL	1 µL	1 µL
pUC19 (1000 ng/µL)	1 µL	0.5 µL	0.5 µL
AdoMet (3.5 mM)	0.5 µL	-	-
M.TaqI	0.125 µL	0.125 µL	n/a
water	16.5 µL	8.5 µL	8.5 µL

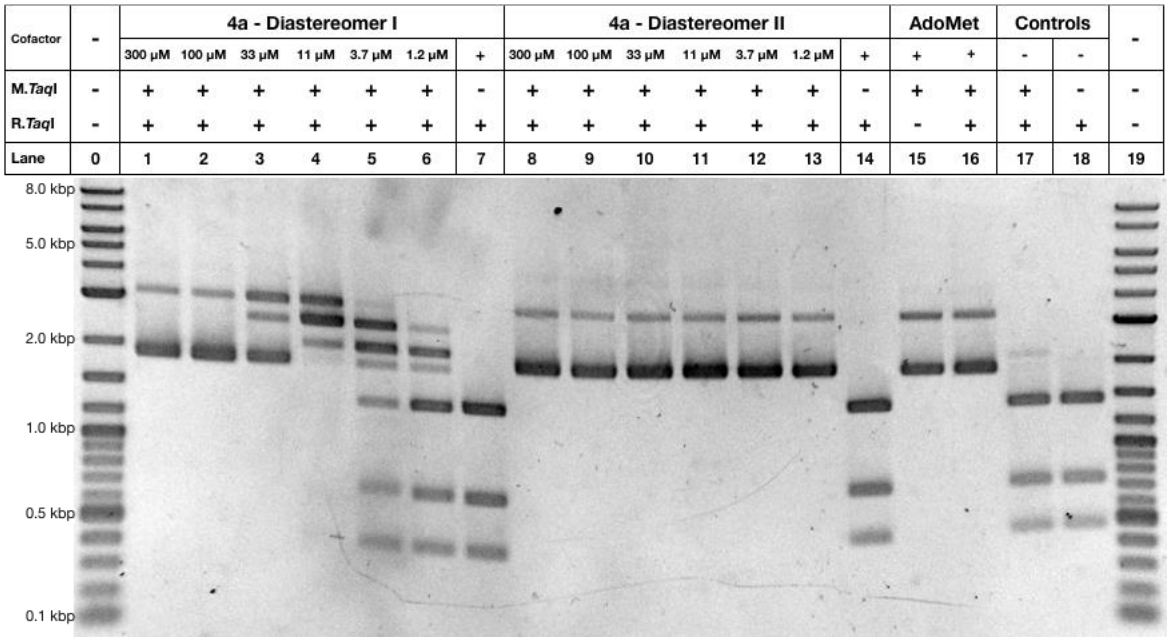


Figure S32 Gel electrophoresis of pUC19 following enzymatic treatment with M.TaqI and/or R.TaqI in the presence and absence of **AdoMet** or **AdoMet** derivative **4a.**

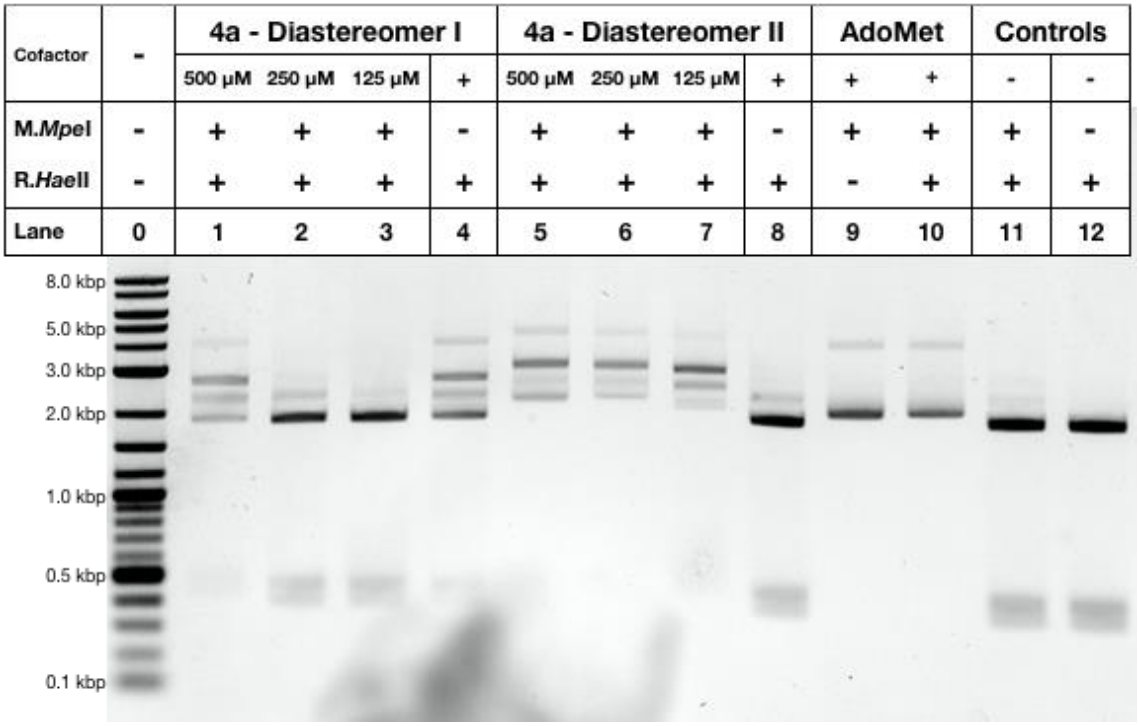


Figure S33 Gel electrophoresis of pUC19 following enzymatic treatment with M.MpeI and/or R.HaeII in the presence and absence of **AdoMet** or **AdoMet** derivative **4a**.

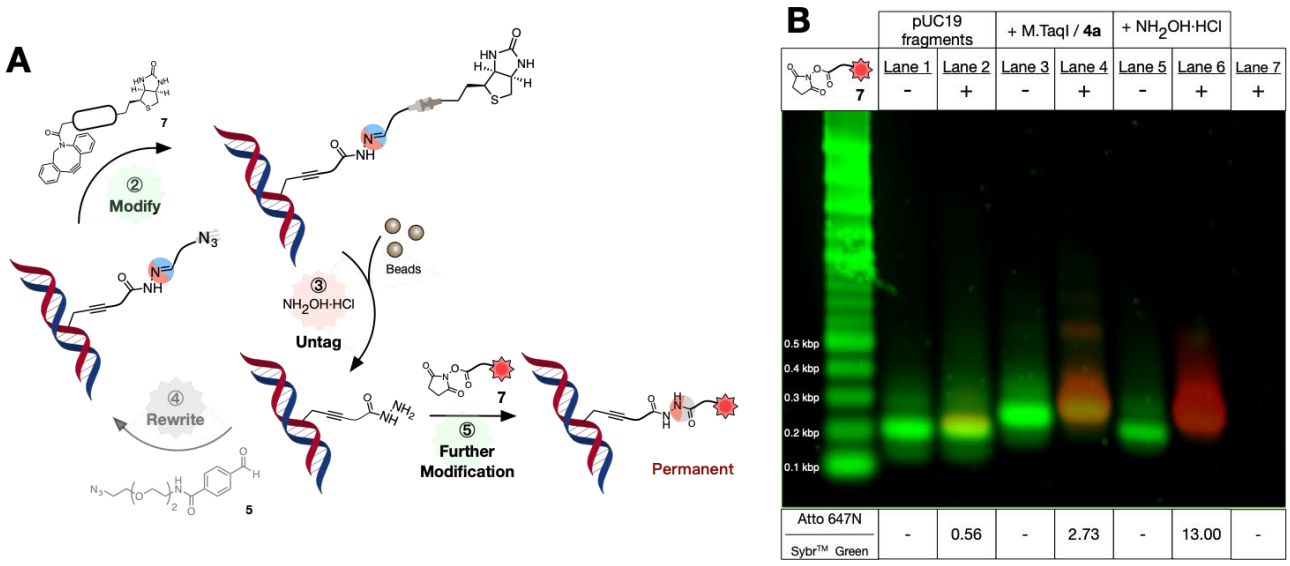


Figure S34: (A) Schematic representation of the further functionalization with NHS-activated ATTO 647N (**7**). (B) Gel electrophoresis of pUC19 fragments. DNA was either unlabeled (lanes 1 and 2), incubated with M.TaqI and **4a** (lanes 3 and 4), or incubated with M.TaqI and **4a** followed by incubation with $H_2NOH{\cdot}HCl$ in 10 mM ammonium acetate pH 4 (lanes 5 and 6). The NHS ester of ATTO 647N (**7**) is shown as a control (lane 7). Conditions: DNA concentration; 7 ng/µL, untag buffer; 10 mM ammonium acetate, pH 6.8, 1 M NaCl, 0.01% SDS. DNA stained with SYBR™ Green. Gel was visualized using a ChemiDoc MP Imager (Bio-Rad) (SYBR™ Green: excitation, trans-UV; emission filter, 590/110 nm; ATTO 647N (**7**): excitation, epi-red illumination; emission filter: 700/50 nm). SYBR™ Green channel was colored green and ATTO 647N (**7**) was colored red for visualization. Full chemical structures of the fluorescent dyes are available in Figure S1.

6. Oligonucleotide LC

6.1. Enzymatic Labeling

For each sample a solution of oligo (120 µL, 10 µM), buffer (40 µL, 10x NEB CutSmart buffer), M.TaqI (45 µL), water (189 µL) and cofactor (6 µL, 20 mM) was made. Samples were incubated at 50 °C for 1.5 h. After incubation, proteinase K (2.5 µL) was added and the samples were incubated at 50 °C for a further 1 h. The samples were then purified using the Qiagen Qiaquick nucleotide clean up kit and eluted into 50 µL water and their concentration was measured by Shimadzu BioSpec-nano. Samples not to be un-tagged were taken and stored in the fridge until LC analysis.

6.2. Untagging – Competitive Exchange of the Schiff-Base

To the labeled DNA, a solution of $H_2NOH{\cdot}HCl$ in water (10 µL, 10 equiv. per site) was added. The pH of the solution was then adjusted using 100 mM ammonium acetate buffer (pH 4.0, 7 µL). The samples were then incubated at 50 °C for 1.5 h and then stored in the fridge until analysis.

Table 2: Mass spec data collected of the fractions collected during HPLC for each oligonucleotide strand.

	Strand 1 5'-CGGCG<u>A</u>GCTACGGC-3'		**Strand 2 3'-GCCGCTCG<u>A</u>TGCCG-5'**	
	Calculated	Observed	Calculated	Observed
Unlabeled	4289.8	N/A	4240.8	N/A
Methylated	4304.8	4304.28	4255.8	4255.26
Acyl Hydrazone	4773.8	4773.36	4724.8	4724.76
Oxime	4760.8	4760.32	4711.8	4711.32
Hydrazide	4441.8	4441.8	4392.8	4392

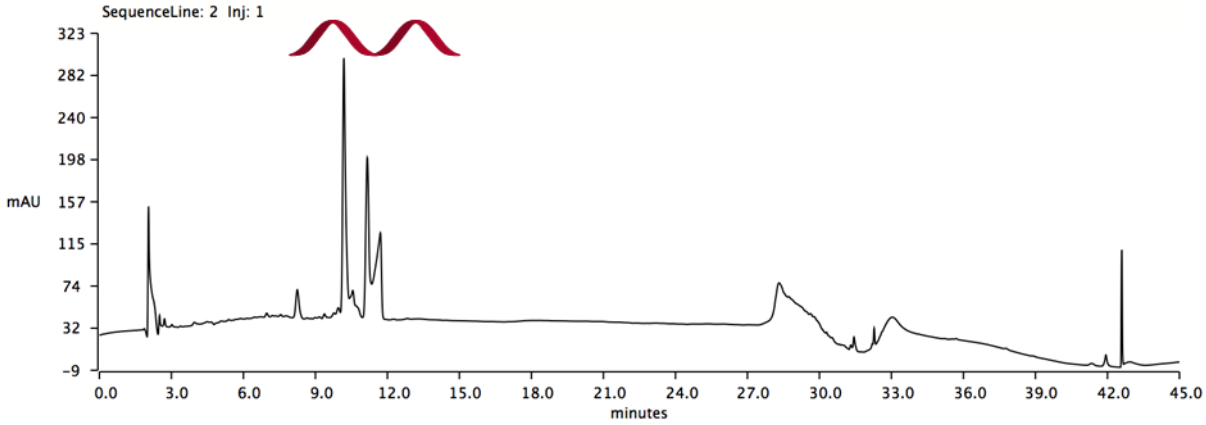


Figure S35 Analytical-HPLC chromatogram of 14 bp oligo DNA. Conditions: 0.1 M TEAA buffer against acetonitrile. 5-18% MeCN over 25 minutes, 18-100% over 5 minutes, hold 100% 10 minutes, drop to 5% and hold 5 minutes, 60 °C.

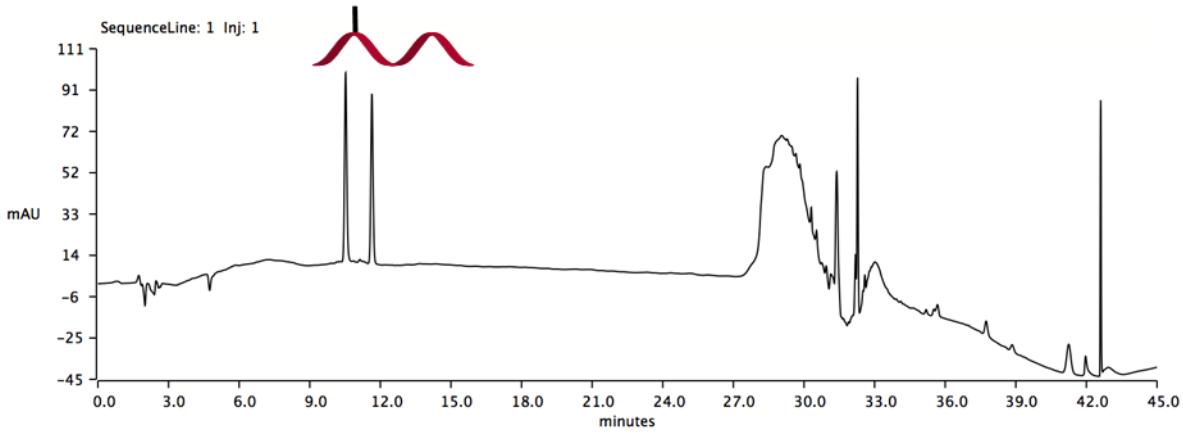


Figure S36 Analytical-HPLC chromatogram of 14 bp oligo DNA incubated in the presence of M.TaqI and **AdoMet**. Conditions: 0.1 M TEAA buffer against acetonitrile. 5-18% MeCN over 25 minutes, 18-100% over 5 minutes, hold 100% 10 minutes, drop to 5% and hold 5 minutes, 60 °C.

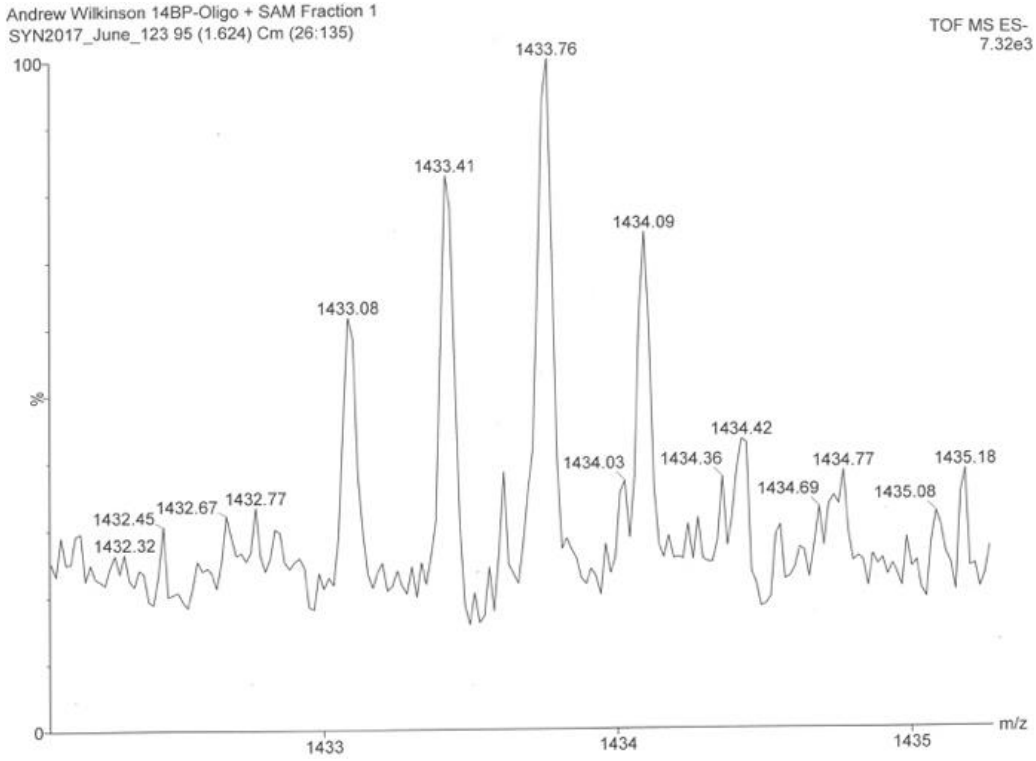


Figure S37. Mass spectrum of the peak at 10.53 min from Figure S36 (methylated 14 bp oligo strand 1, M – 3H)

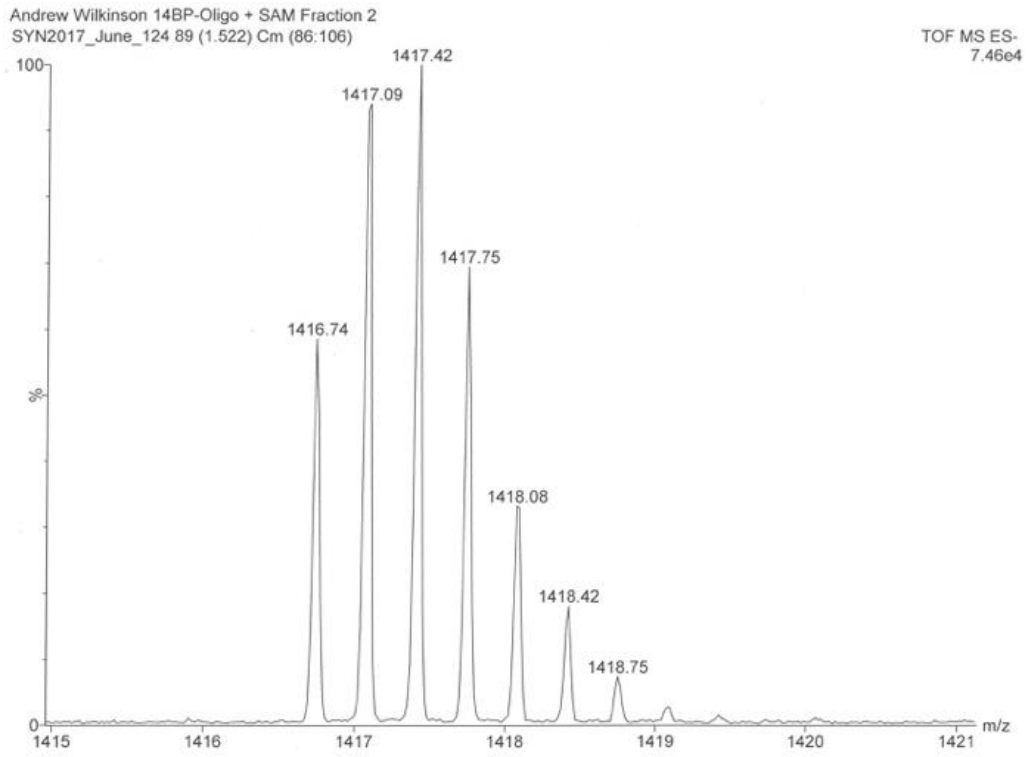


Figure S38. Mass spectrum of the peak at 11.64 min from Figure S36 (methylated 14 bp oligo strand 2, M – 3H)

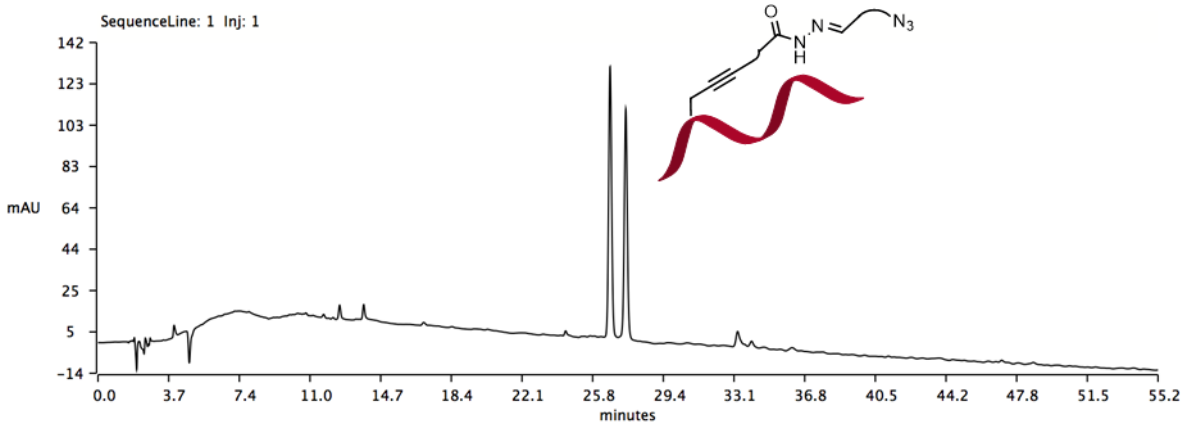


Figure S39 Analytical-HPLC chromatogram of 14 bp oligo DNA incubated in the presence of M.TaqI and **4a**. Conditions: 0.1 M TEAA 5-36% over 50 minutes, 60 °C.

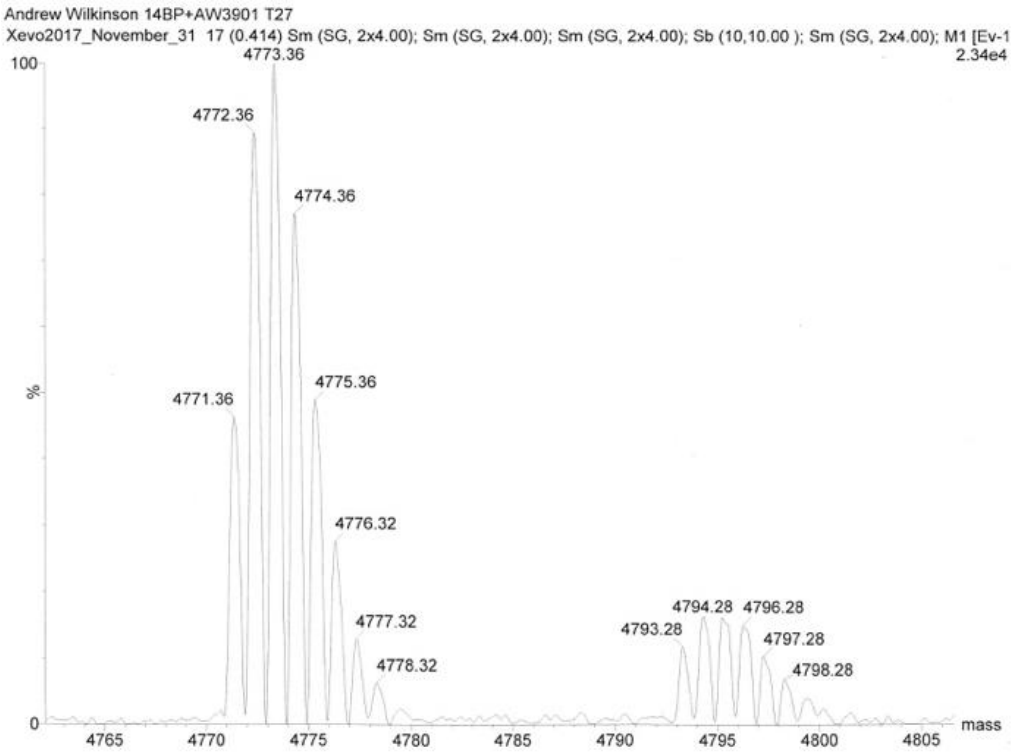


Figure S40. Mass spectrum of the peak at 27 min from Figure S39 (14 bp oligo strand 1 incubated with M.TaqI and **4a**. 4773.36: M – H; 4795.28: M + Na – H).

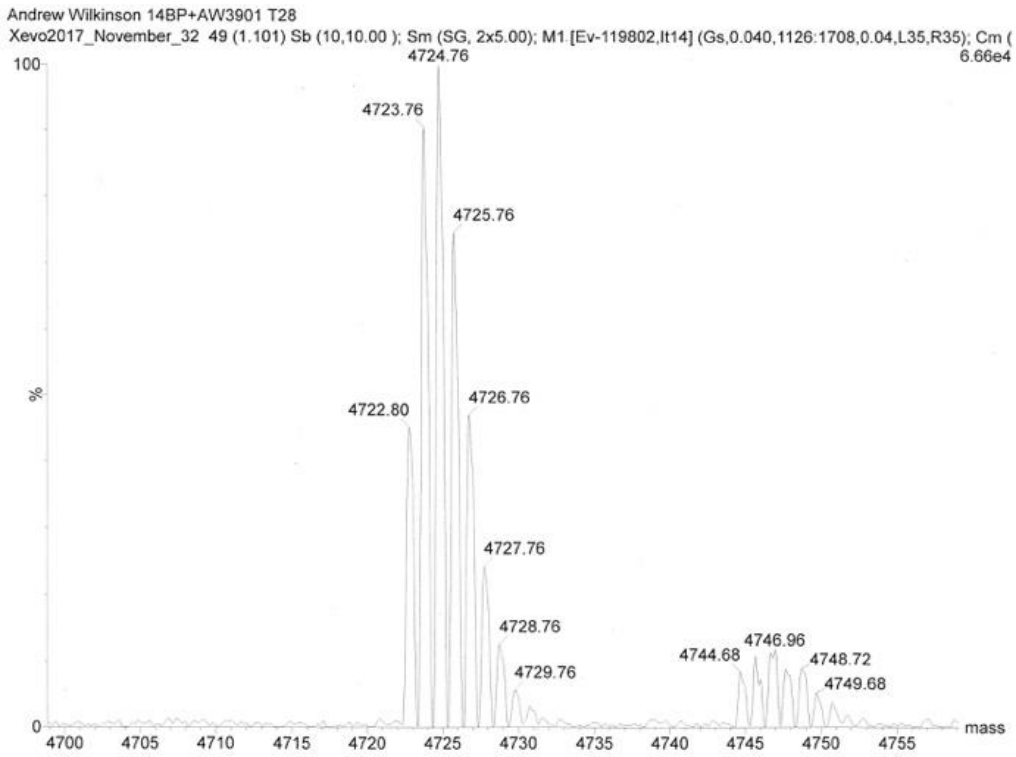


Figure S41. Mass spectrum of the peak at 28 min from Figure S39 (14 bp oligo strand 2 incubated with M.TaqI and **4a**. 4724.76: M – H; 4746.96: M + Na – H).

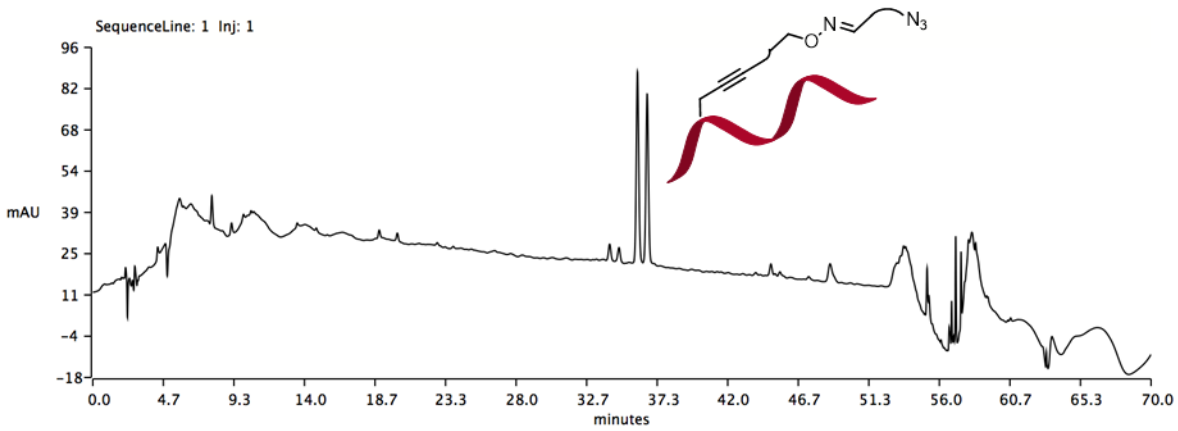


Figure S42 Analytical-HPLC chromatogram of 14 bp oligo DNA incubated in the presence of M.TaqI and **4b**. Conditions: 0.1 M TEAA buffer against acetonitrile. 5-18% MeCN over 25 minutes, 18-100% over 5 minutes, hold 100% 10 minutes, drop to 5% and hold 5 minutes, 60 °C.

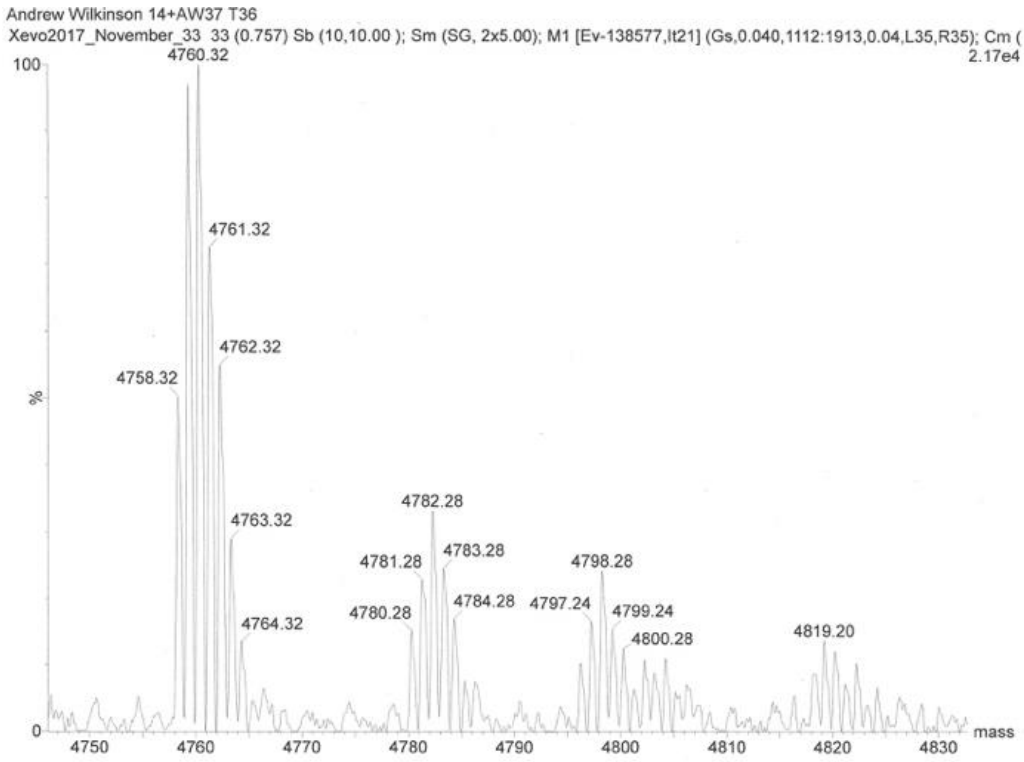


Figure S43. Mass spectrum of the peak at 36 min from Figure S42 (14 bp oligo strand 1 incubated with M.TaqI and **4b**. 4760.32: M – H; 4795.28: M + Na – H; 4795.28: M + K – H; 4819.20: M + Na + K – 2H).

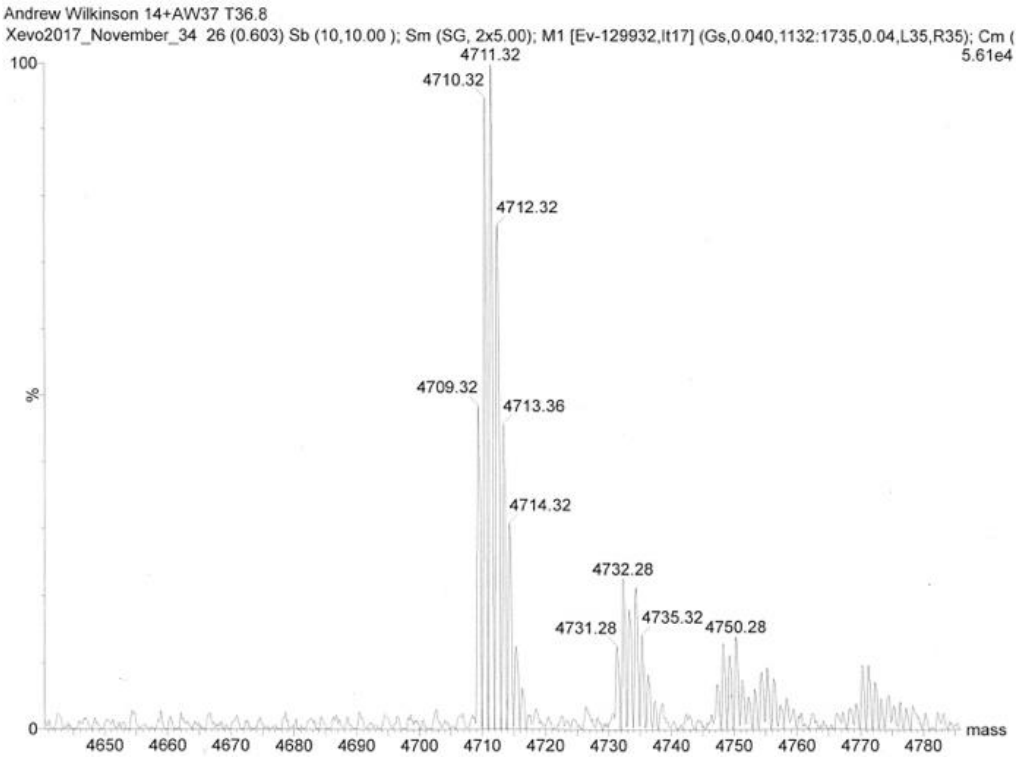


Figure S44. Mass spectrum of the peak at 36.8 min from Figure S42 (14 bp oligo strand 2 incubated with M.TaqI and **4b.** 4710.32: M – H; 4732.28: M + Na – H; 4750.28: M + K – H).

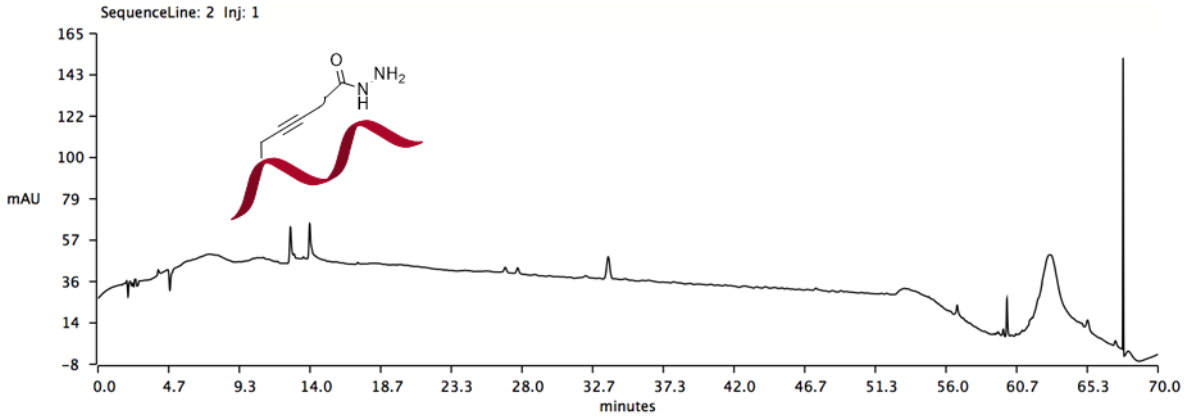


Figure S45 Analytical-HPLC chromatogram of 14 bp oligo DNA incubated in the presence of M.TaqI and **4a**, followed by treatment with $H_2NOH{\cdot}HCl$ (10 equiv.) in a pH 4 10 mM ammonium acetate buffer. Conditions: 0.1 M TEAA buffer against acetonitrile. 5-18% MeCN over 25 minutes, 18-100% over 5 minutes, hold 100% 10 minutes, drop to 5% and hold 5 minutes, 60 °C.

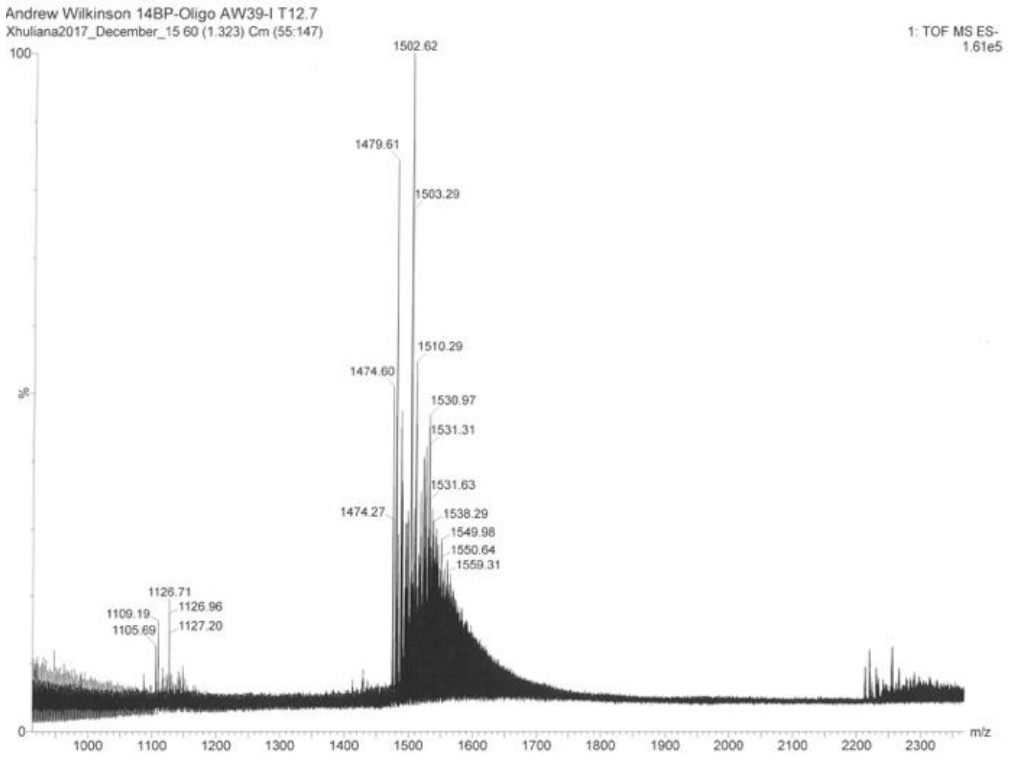


Figure S46. Mass spectrum of the peak at 12.7 min from Figure S45 (14 bp oligo strand 1 incubated with M.TaqI and **4b**, followed by treatment with $H_2NOH{\cdot}HCl$. 1502.96: M – 3H).

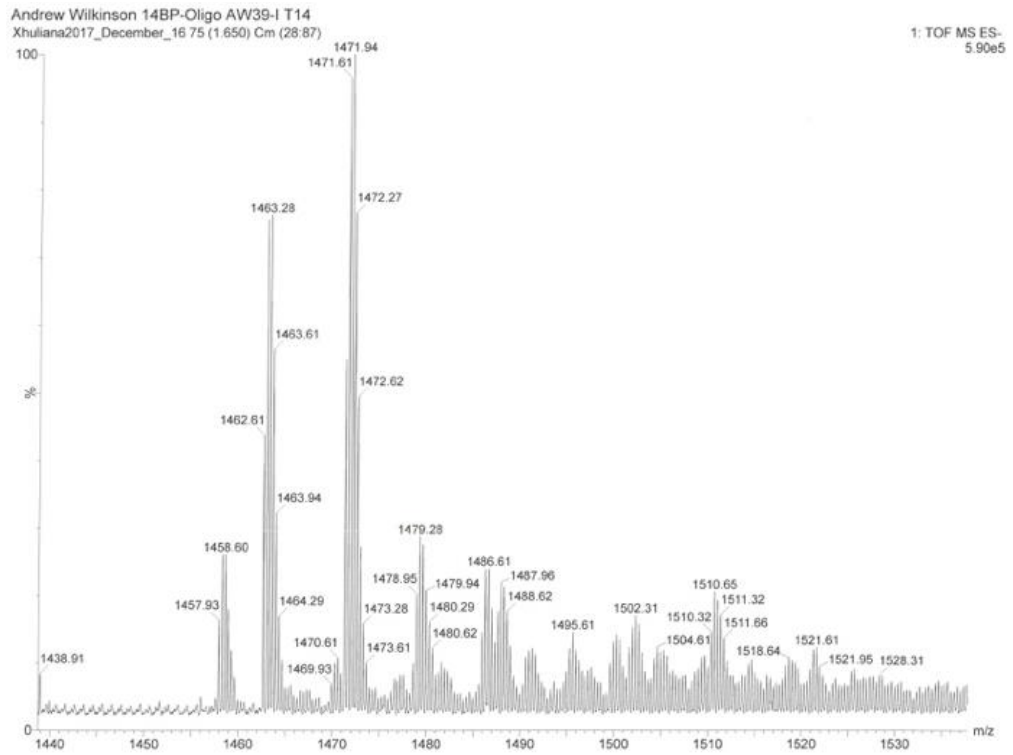


Figure S47. Mass spectrum of the peak at 14 min from Figure S45 (14 bp oligo strand 2 incubated with M.TaqI and **4b**, followed by treatment with $H_2NOH{\cdot}HCl$. 1463.28: M – 3H; 1471.94: M + Na – 4H).

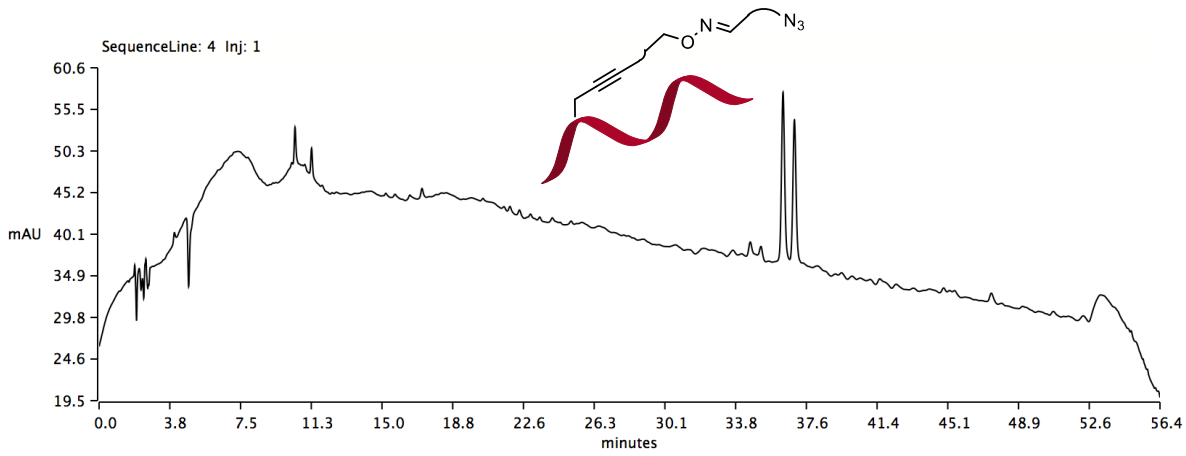


Figure S48 Analytical-HPLC chromatogram of 14 bp oligo DNA incubated in the presence of M.TaqI and **4b**, followed by treatment with $H_2NOH{\cdot}HCl$ (10 equiv.) in a pH 4 10 mM ammonium acetate buffer. Conditions: 0.1 M TEAA buffer against acetonitrile. 5-18% MeCN over 25 minutes, 18-100% over 5 minutes, hold 100% 10 minutes, drop to 5% and hold 5 minutes, 60 °C.

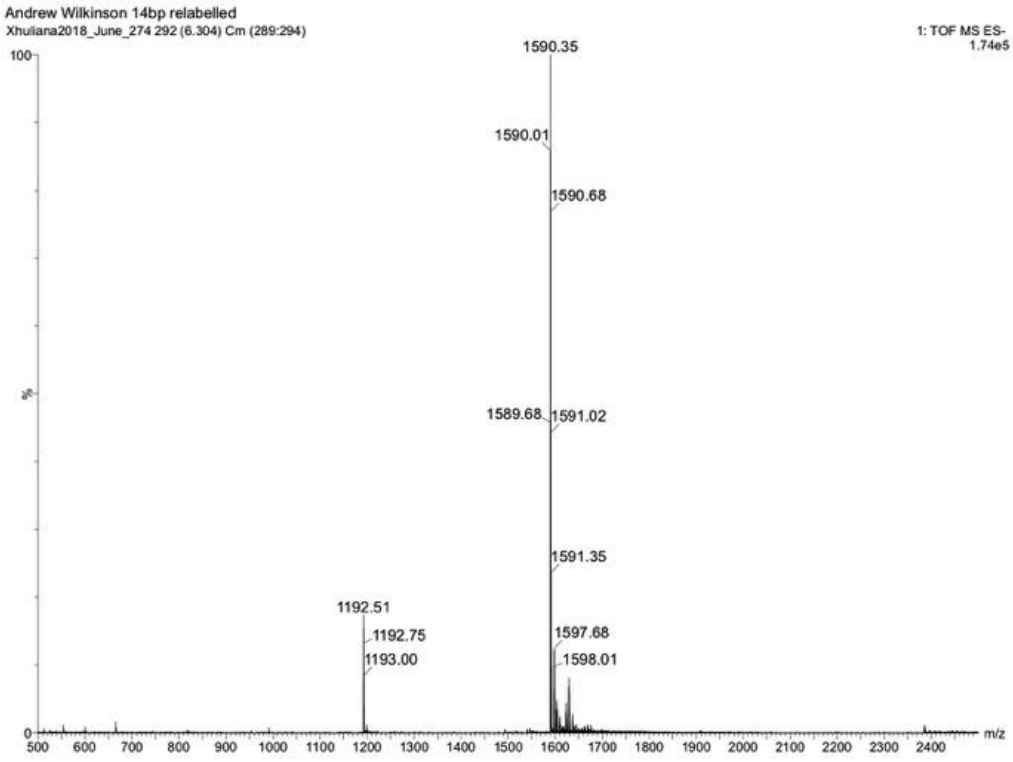


Figure S49 MS spectra of the 14 bp oligo DNA, strand 1, after re-witting with aldehyde **5** (1192.51: M – 4H; 1590.35: M – 3H; 1598.01: M + Na – 4H).

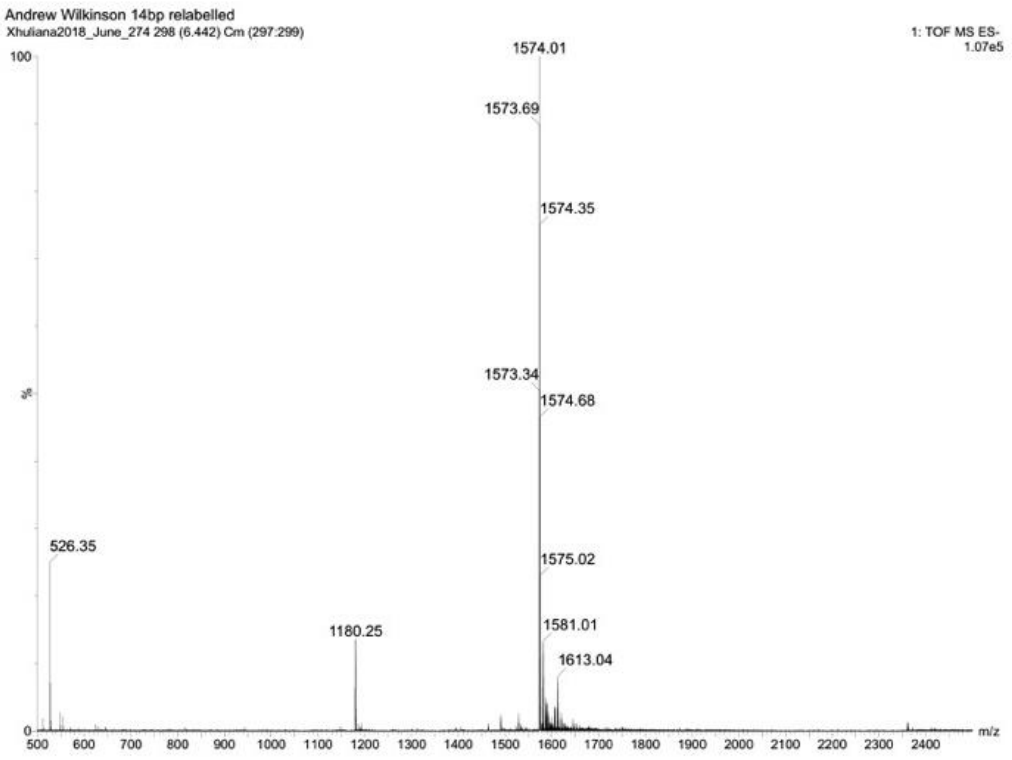


Figure S50 MS spectra of the 14 bp oligo DNA, strand 2, after rewritting with aldehyde **5** (1180.25: M – 4H; 1574.01: M – 3H).

7. PCR Product: Amplification, Labeling, Capture and Release

7.1. PCR Amplification

The PCR product was prepared by amplification of a section of the plasmid pUC19 using the corresponding primers (Forward primer: 5'-GCC AGG AAC CGT AAA AAG-3' and Reverse primer: 5'-AGA AAG GCG GAC AGG TAT-3'). The DNA fragment amplified was 203 bp with 31 CpG sites. For amplification, a mastermix of buffer (200 µL, NEB 2x high-fidelity mastermix), forward primer (4 µL, 100 µM), reverse primer (4 µL, 100 µM), pUC19 (2 µL, 10 ng/µL) and water (190 µL) was prepared. The mastermix was then aliquoted into 50 µL portions. Amplification was completed using the program: 98 °C for 45 s, 40 cycles 98 °C for 15 s, 62 °C for 45 s, and 72 °C for 60 s. After amplification, the DNA was purified on silica column (Sigma-Aldrich PCR clean up kit) and eluted into water.

7.2. PCR Product Labeling

Once collected, the DNA fragment was labeled using the CpG specific DNA methyltransferase M.MpeI. For labeling, samples were made up of: DNA (66 µL, 40 ng/µL), buffer (10 µL, 10x NEB CutSmart buffer), cofactor (3 µL, 15 mM), M.MpeI (5 µL), and water (16 µL). Samples were incubated at 37 °C for 1 h after which proteinase K (2 µL) was added and incubated at 50 °C for 1 h. Finally, a solution of diazo-DBCO-Biotin (2 µL, 50 mM, Jena Bioscience) was added and incubated for a further 1 h at 37 °C. After labeling, the DNA was purified using a Qiagen purification kit and the DNA was eluted into a high salt Tris buffer (10 mM Tris, 1 M NaCl, pH 8.5, Tris A) ready for capture.

7.3. DNA Capture

After labeling, the DNA was captured onto streptavidin-coated magnetic beads (Dynabeads myone C1). For capture, 10 µL of the bead stock was washed 2x with 70 µL Tris A and then all supernatant was removed. The bead pellet was then resuspended in the DNA solution previously eluted (500 ng, 70 µL, Tris A). This solution was shaken at room temperature for 15 minutes. After capture, the bead mixture was placed on a magnetic rack and the DNA concentration of the eluent was measured. The eluent was then placed into a fresh bead pellet to capture the remaining DNA. This process was repeated 3x until the majority of DNA was captured.

After capture, the bead pellets were resuspended in water and washed 2x. The beads were then finally placed within the relevant buffer ready for release.

7.4. General DNA Release

To release captured DNA (500 ng, 200 bp) the bead pellets were suspended in ammonium acetate buffer (63 µL, 11.1 mM, 1 M NaCl, 0.01% SDS, var. pH) and to this, a solution of $H_2NOH{\cdot}HCl$ (7 µL, var. equiv.) was added. The bead solution was then shaken at 50 °C for 1 h . After shaking, the solution was placed on a magnet, the supernatant was removed and the DNA concentration was measured. This process was repeated until all DNA had been released.

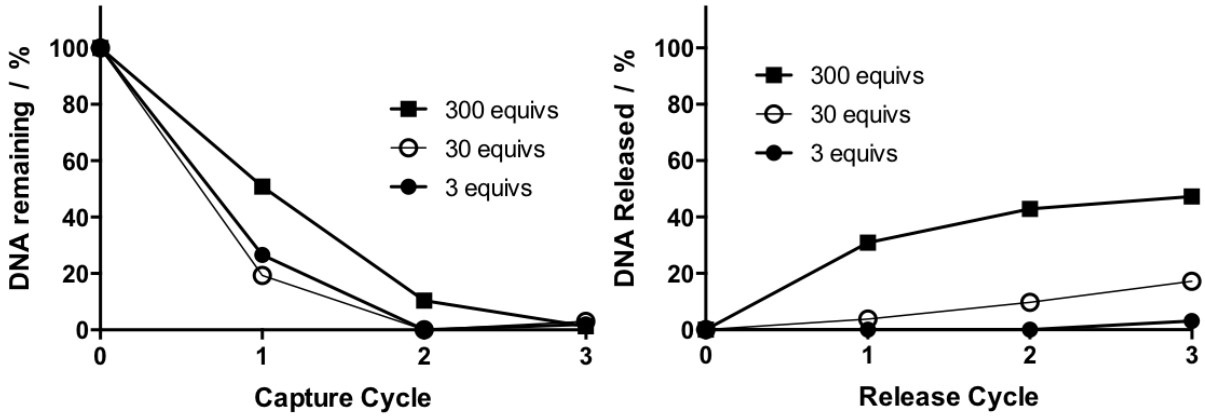


Figure S51 Left: Percentage of DNA remaining in solution following capture of DNA. Traces have been labeled with the corresponding release experiment. Right: Percentage of DNA released from the magnetic beads following treatment 3 (●), 30 (○) or 300 (□) equiv. of $H_2NOH{\cdot}HCl$ in acetate buffer at pH 4.9. In all cases, DNA (2500 ng, 19.9 pmol) was incubated with M.MpeI and **4a** and then purified prior to capture. Capture was then performed in all cases using approx. 1 µg DNA (8 pmol, 22 ng/µL). n = 1. The amount of DNA was quantified using a Qubit fluorometer and normalized to the starting amount of DNA (Left) or the amount of DNA captured (Right).

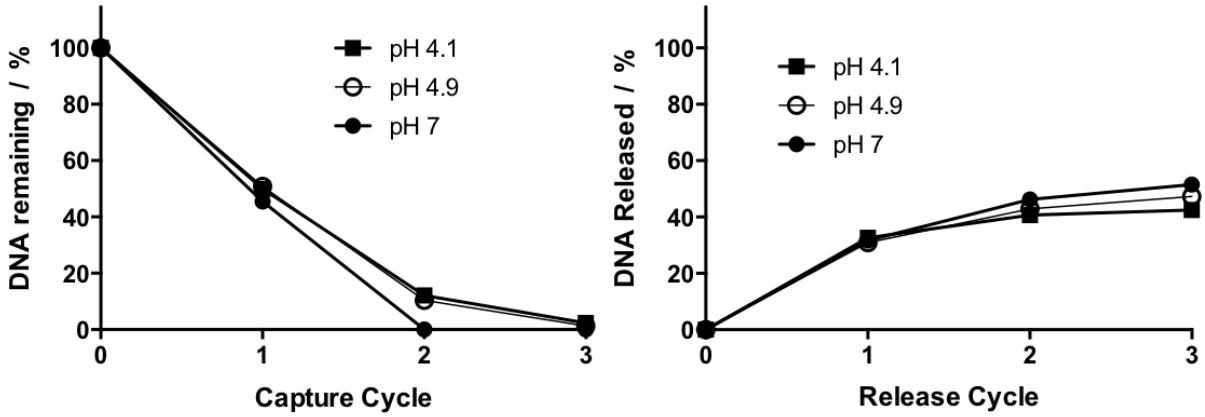


Figure S52 Left: Percentage of DNA remaining in solution following capture for these experiments. Traces have been labeled with the corresponding release experiment. Right: Percentage DNA released from the magnetic beads following treatment with $H_2NOH{\cdot}HCl$ (300 equiv.) in acetate buffer at pH 4.1 (□), 4.9 (●) or 7.4 (○). In all cases, DNA (2500 ng, 19.9 pmol) was incubated with M.MpeI and **4a** and then purified prior to capture. Capture was then performed in all cases using approx. 1 µg DNA (8 pmol, 22 ng/µL). n = 1. The amount of DNA was quantified using a Qubit fluorometer and normalized to the starting amount of DNA (Left) or the amount of DNA captured (Right).

7.5. DNA Quantification Using Qubit Fluorometer

DNA quantification was performed using the Qubit® 3.0 Fluorometer. A broad-range (BR) assays kit was used according to the instructions provided by the supplier.

7.6. Real Time PCR Amplification of Purified DNA

DNA following capture and release (**Figure 6**) was prepared by PCR amplification of a known DNA fragment (332 bp) from the Lambda bacteriophage genome (NEB). Real time PCR was performed using the SsoAdvanced™ Universal SYBR™ Green Supermix (Bio-Rad) with appropriate primers (forward primer: 5'- GTG GTG AAA GGG CAG AGC A -3' and reverse primer: 5'- AGG GCG AGA TGC TCA ATG C -3'). The amplification program was set as: 98 °C for 45 s, 40 cycles 98 °C for 15 s, 68 °C for 45 s, and 72 °C for 60 s. The captured/released DNA was diluted 500 times, prior to running the qPCR experiment.

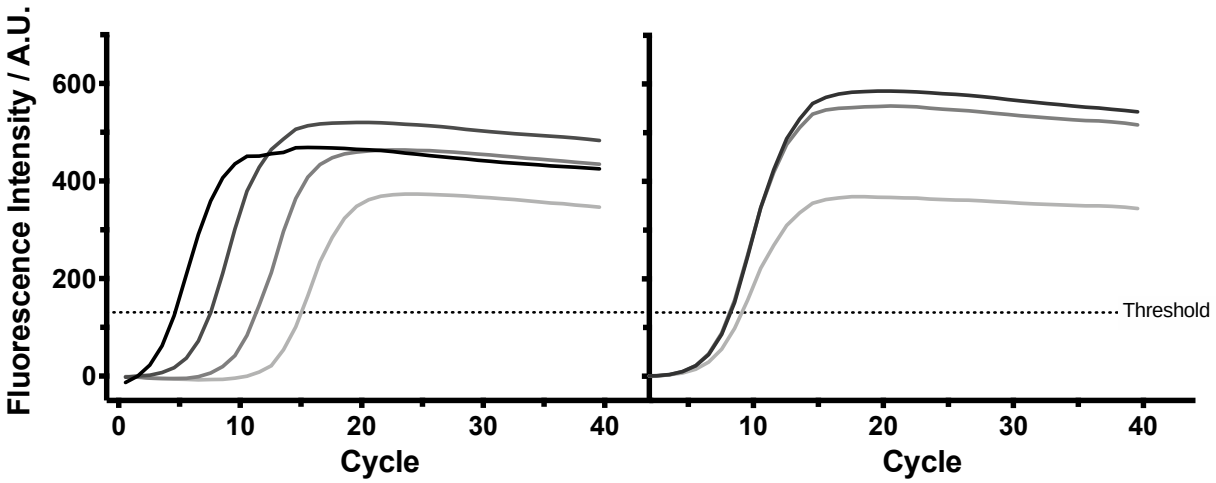


Figure S53 Real-time PCR of pUC19 fragments following capture and release experiment (**Figure 6**): (Left) A standard curve is generated using samples with known concentrations, from left to right, at 1000 pg/µL, 100 pg/µL, 10 pg/µL and 1 pg/µL. (Right) Amplification curves for the captured/released DNA, where concentrations in the PCR reaction were calculated to be 83 pg/µL, 79 pg/µL and 45 pg/µL. These correspond to released DNA concentrations of 41.5 ng/µL, 39.5 ng/µL and 22.5 ng/µL respectively.