Supplementary Information

Synthesis and biological activity of a CXCR4-targeting bis(cyclam) lipid

Anna D. Peters,[a,b,†] Catriona McCallion,[a,b,c,†] Andrew Booth,[a,b] Julie A. Adams,[c] Karen Rees-Unwin,[d] Alain Pluen,[e] John Burthem,[c,d,*] Simon J. Webb[a,b,*]

a. School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom

b. Manchester Institute of Biotechnology, University of Manchester, 131 Princess St, Manchester M1 7DN, United Kingdom

c. Department of Haematology, Manchester Royal Infirmary, Manchester University NHS Foundation Trust, Manchester M13 9WL, United Kingdom

d. Division of Cancer Sciences, University of Manchester, Manchester, United Kingdom

e. Division of Pharmacy and Optometry, School of Health Sciences, University of Manchester, M13 9PL, United Kingdom

† Contributed equally to the work

Table of Contents

S.1. NMR spectra of new compounds and known compounds without reported spectral data 5

S.1.1. ^{1}H and ^{13}C NMR spectra of compound **9** 5

Figure 1: ^{1}H NMR spectrum of **9** (400 MHz, D_2O, 298 K) 5

Figure 2: ^{13}C NMR spectrum of **9** (100 MHz, D_2O, 298 K) 5

S.1.2. ^{1}H and ^{13}C NMR spectra of compound **3b** 6

Figure 3: ^{1}H NMR spectrum of **3b** (400 MHz, $CDCl_3$, 298 K) 6

Figure 4: ^{13}C NMR spectrum of **3b** (100 MHz, $CDCl_3$, 298 K) 6

S.1.3. ^{1}H and ^{13}C NMR spectra of compound **5a** 7

Figure 5: ^{1}H NMR spectrum of **5a** (400 MHz, $CDCl_3$, 298 K) 7

Figure 6: ^{13}C NMR spectrum of **5a** (100 MHz, $CDCl_3$, 298 K) 7

S.1.4. ^{1}H and ^{13}C NMR spectra of compound **5b** 8

Figure 7: ^{1}H NMR spectrum of **5b** (400 MHz, $CDCl_3$, 298 K) 8

Figure 8: ^{13}C NMR spectrum of **5b** (100 MHz, $CDCl_3$, 298 K) 8

S.1.5. ^{1}H and ^{13}C NMR spectra of compound **7** 9

Figure 9: ^{1}H NMR spectrum of **7**.$(AcOH)_n$ (500 MHz, D_2O, 298 K) 9

Figure 10: ^{1}H NMR spectrum of **7**.$(AcOH)_3$ (500 MHz, CD_3OD, 348 K) 9

Figure 11: ^{13}C NMR spectrum of **7**.$(AcOH)_3$ (500 MHz, D_2O, 298 K) 10

S.1.6. ^{1}H and ^{13}C NMR spectra of compound **1** 11

Figure 12: ^{1}H NMR spectrum of **1**.$(AcOH)_n$ (500 MHz, D_2O, 373 K) 11

Figure 13: ^{13}C NMR spectrum of **1**.$(AcOH)_n$ (125 MHz, D_2O, 298 K) 11

S.1.7. ^{1}H and ^{13}C NMR spectra of compound **2** 12

Figure 14: ^{1}H NMR spectrum of **2** (500 MHz, $CDCl_3$, 298 K) 12

Figure 15: ^{13}C NMR spectrum of **2** (125 MHz, $CDCl_3$, 298 K) 12

S.1.8. HMBC and HSQC spectra of compound **2** 13

Figure 16: ^{13}C NMR spectrum of **2**, with the aromatic region inset 13

Figure 17: 2D HMBC spectrum of **2**. Both the 1 and 4 (ArC-O) aromatic carbon resonances and the 2 and 5 (ArC-C) aromatic carbon resonances are observed (boxed) 14

Figure 18: Detailed image showing the aromatic resonances associated with the 2 and 5 carbons (124.06 and 124.98 ppm, circled in blue) and the 1 and 4 carbons (151.13 and 152.28 ppm, circled in red). 14

Figure 19: 2D HSQC NMR spectrum of **2**, with the peaks associated with the 3 and 6 (ArC-H) aromatic carbon resonances (115.32 and 116.48 ppm) indicated by the green box. 15

Figure 20: HSQC NMR peaks associated with the 3 and 6 (ArC-H) aromatic carbon resonances of **2** (circled in blue on structure) 15

S.2. NMR and VT ^{1}H NMR spectroscopy data for compound 6 16

Figure 21: ^{1}H NMR spectrum of **6** (400 MHz, CD_3CN, 345 K) 16

Figure 22: ^{1}H VT NMR spectra of **6**; VT steps: 235, 253, 273, 293, 313, 343 & 345 K (500 MHz, CD_3CN); full spectra 17

Figure 23: ^{1}H VT NMR spectra of **6**; VT steps: 235, 243, 253, 263, 273, 283, 293, 303, 313, 323, 343, 345 K (500 MHz, CD_3CN). 17

S.3. Spectroscopic and HPLC data for crude product 5a before Finkelstein reaction 18

Figure 24: HRMS of crude product **5a** before Finkelstein reaction. 18

Table 1: Summary of compounds observed in HRMS spectrum of crude **5a** 18

Figure 25: HPLC of crude product **5a** before Finkelstein reaction 19

Figure 26: ^{1}H NMR spectrum of crude product **5a** before Finkelstein reaction (400 MHz, $CDCl_3$, 298 K) 19

S.4. Analytical HPLC traces for compounds 5a, 7 and 1 20

Figure 27: HPLC trace of **5a** 20

Figure 28: HPLC trace of **7** 20

Figure 29: HPLC trace of **1** 21

S.5. Dynamic light scattering (DLS) histograms for 2 22

Figure 30: A representative histogram set showing the r_h distribution with respect to a) intensity b) mass c) number for a suspension of 2 illustrating the bimodal distribution of hydrodynamic radii. 22

S.6. AFM images of dropcast suspensions of 2 23

Figure 31: An AFM image of 2 drop-cast onto freshly cleaved mica, where b) depicts the central region of a), imaged at higher magnification. A large number of small self-assembled particles are observed, alongside larger loosely assembled aggregates. 23

Figure 32: Particle analysis performed on image Figure 31a: the histogram has been clipped at 36 nm for clarity. 24

S.7. Antibody staining of CXCR4 25

Figure 33: CXCR4 staining on live primary CLL cells after 24 h culturing, where the blue curve shows CXCR4 expression without stimulation or blocking, and the pink shows CXCR4 staining after incubation with 10 µM 1. 25

S.8. Western Blot: B-cell receptor control and full protein stain 26

Figure 34: Western Blots showing phosphorylated and total ERK, following stimulation with a) biotinylated anti-IgM antibody plus avidin and b) CXCL12. 26

Figure 35: Full protein stain of the SDS-PAGE gels used to separate the proteins for both experiments; a) B-cell receptor stimulation, b) CXCR4 stimulation. Stain was performed by

treating the gels with water/acetic acid/methanol followed by staining with Coommassie Brilliant Blue R-250 overnight at room temperature..... 27

S.1. NMR spectra of new compounds and known compounds without reported spectral data

S.1.1. ^{1}H and ^{13}C NMR spectra of compound 9

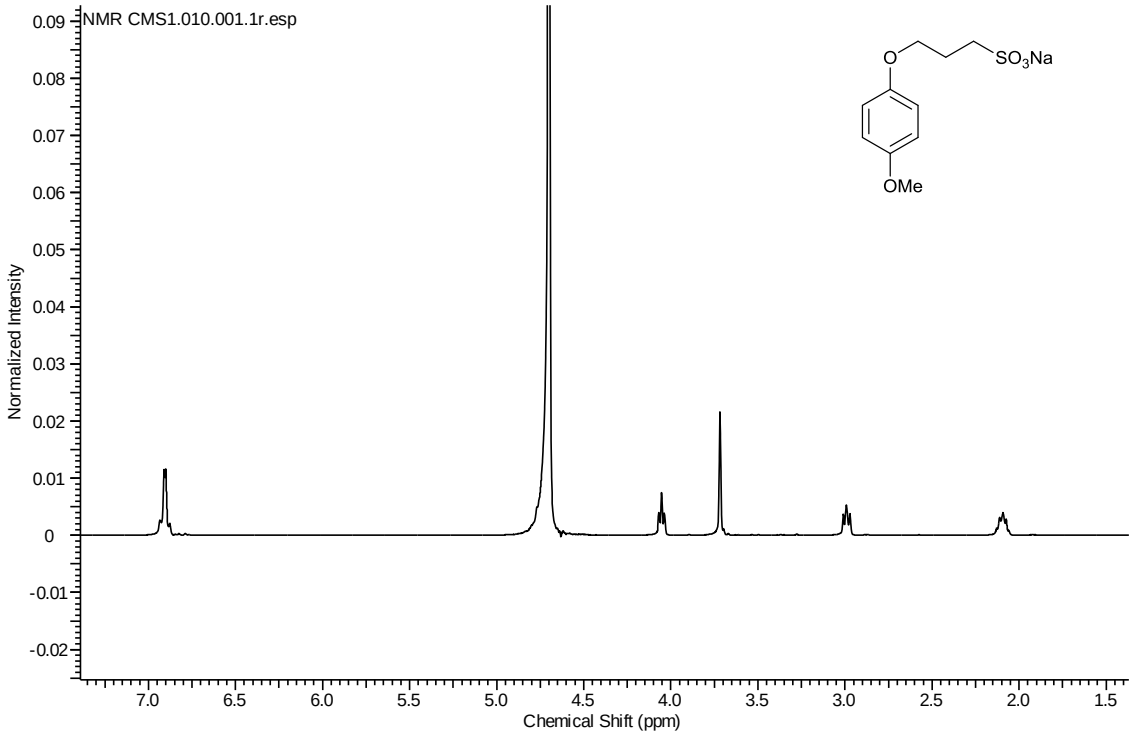

Figure 1: ^{1}H NMR spectrum of **9** (400 MHz, D_2O, 298 K)

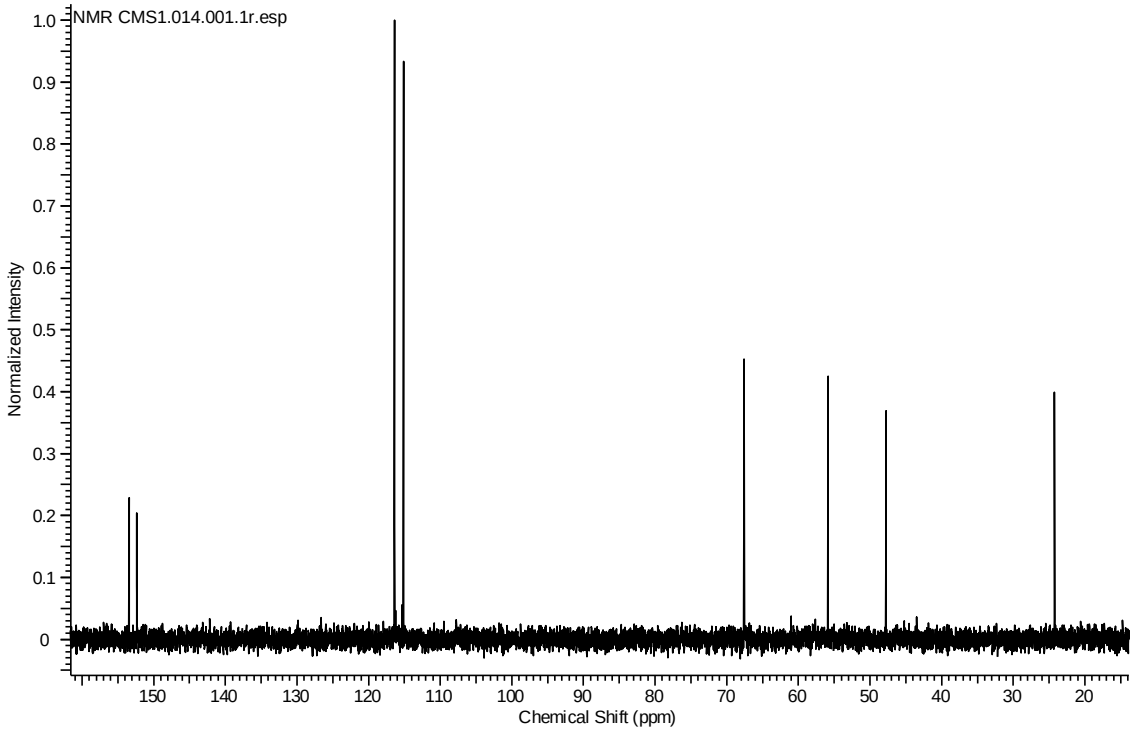

Figure 2: ^{13}C NMR spectrum of **9** (100 MHz, D_2O, 298 K)

S.1.2. ^{1}H and ^{13}C NMR spectra of compound 3b

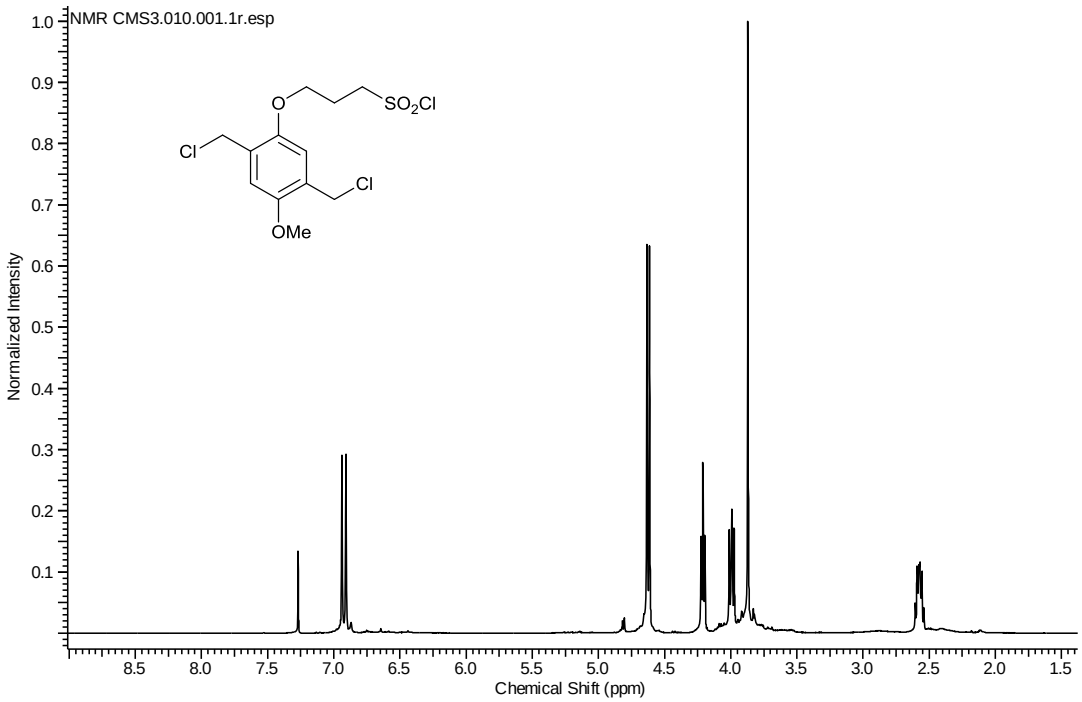

Figure 3: ^{1}H NMR spectrum of **3b** (400 MHz, $CDCl_3$, 298 K)

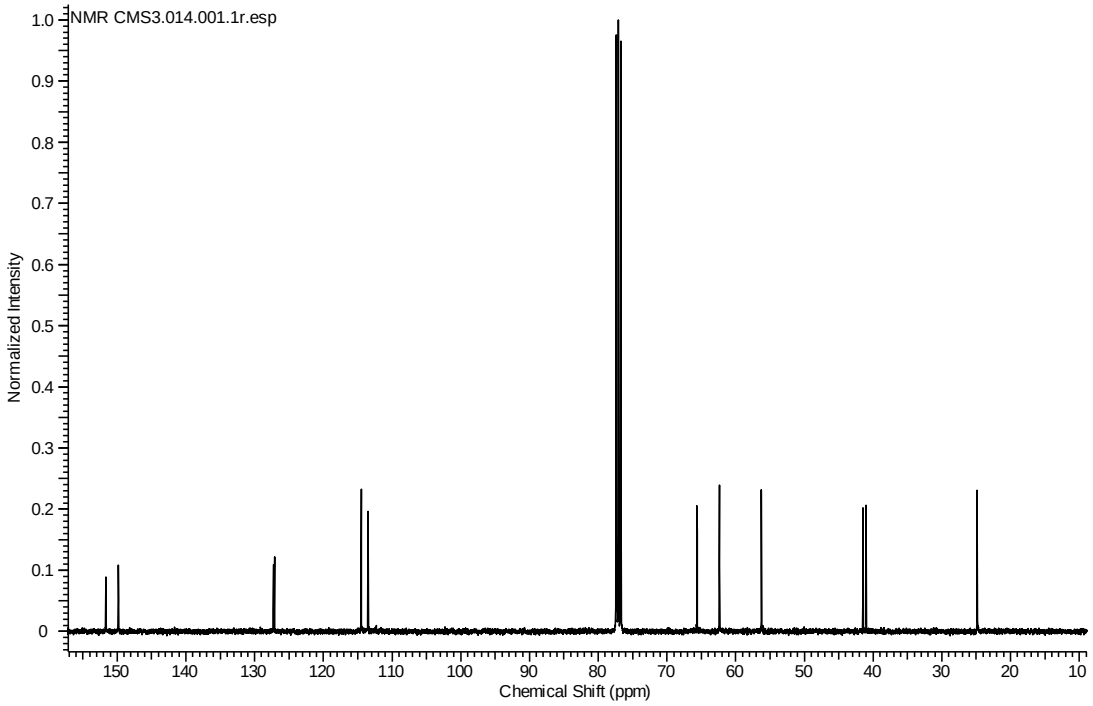

Figure 4: ^{13}C NMR spectrum of **3b** (100 MHz, $CDCl_3$, 298 K)

S.1.3. ^{1}H and ^{13}C NMR spectra of compound 5a

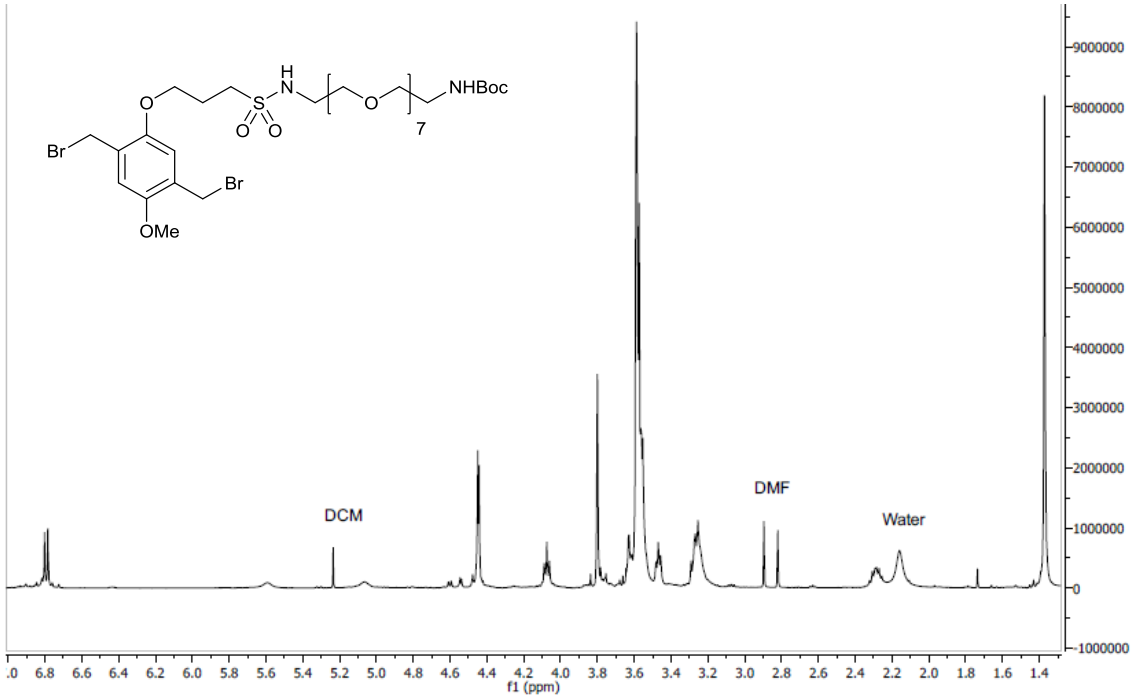

Figure 5: ^{1}H NMR spectrum of **5a** (400 MHz, $CDCl_3$, 298 K)

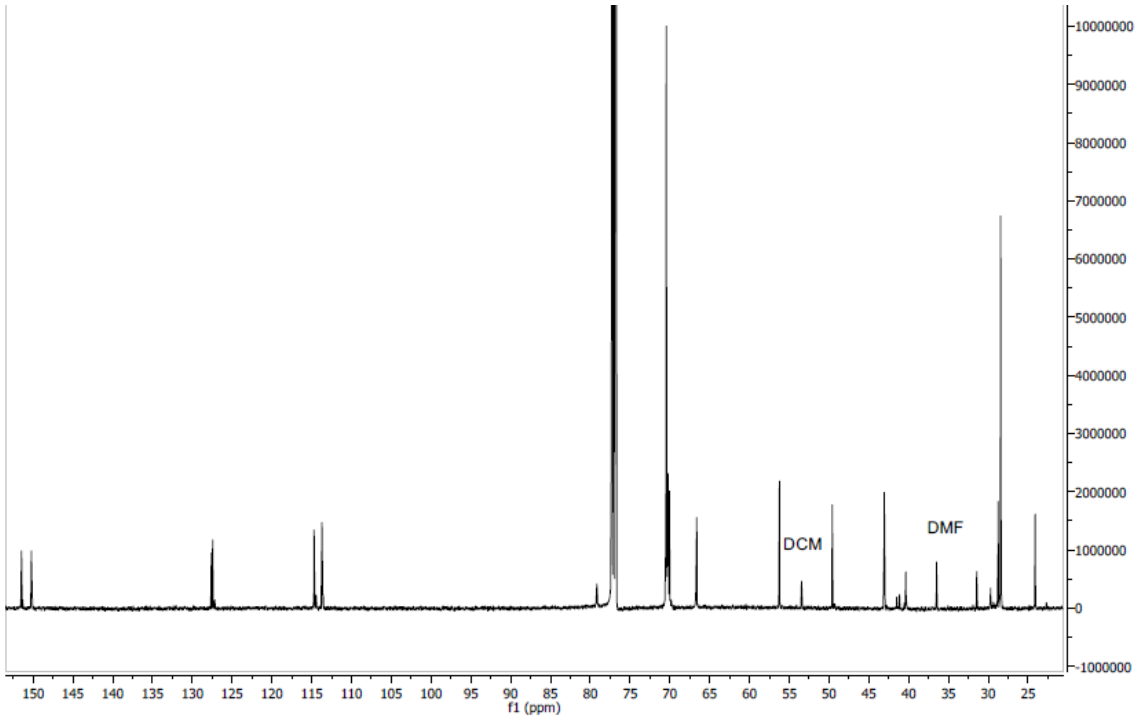

Figure 6: ^{13}C NMR spectrum of **5a** (100 MHz, $CDCl_3$, 298 K)

S.1.4. ^{1}H and ^{13}C NMR spectra of compound 5b

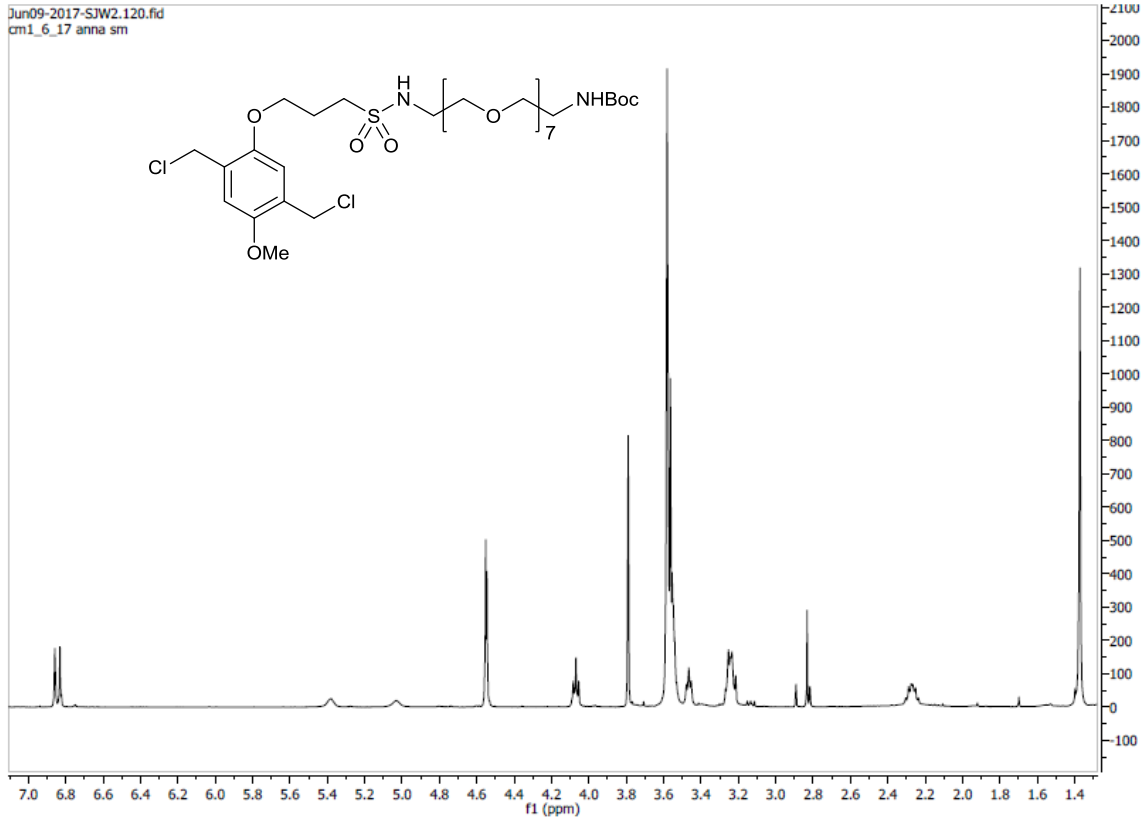

Figure 7: ^{1}H NMR spectrum of **5b** (400 MHz, $CDCl_3$, 298 K)

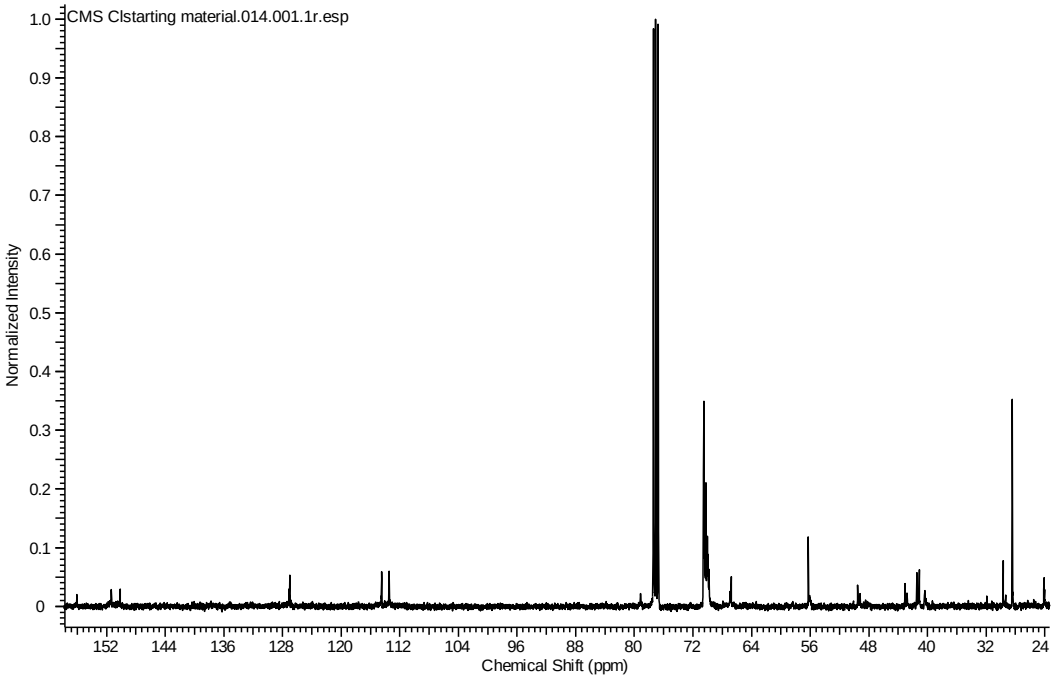

Figure 8: ^{13}C NMR spectrum of **5b** (100 MHz, $CDCl_3$, 298 K)

S.1.5. ^{1}H and ^{13}C NMR spectra of compound 7

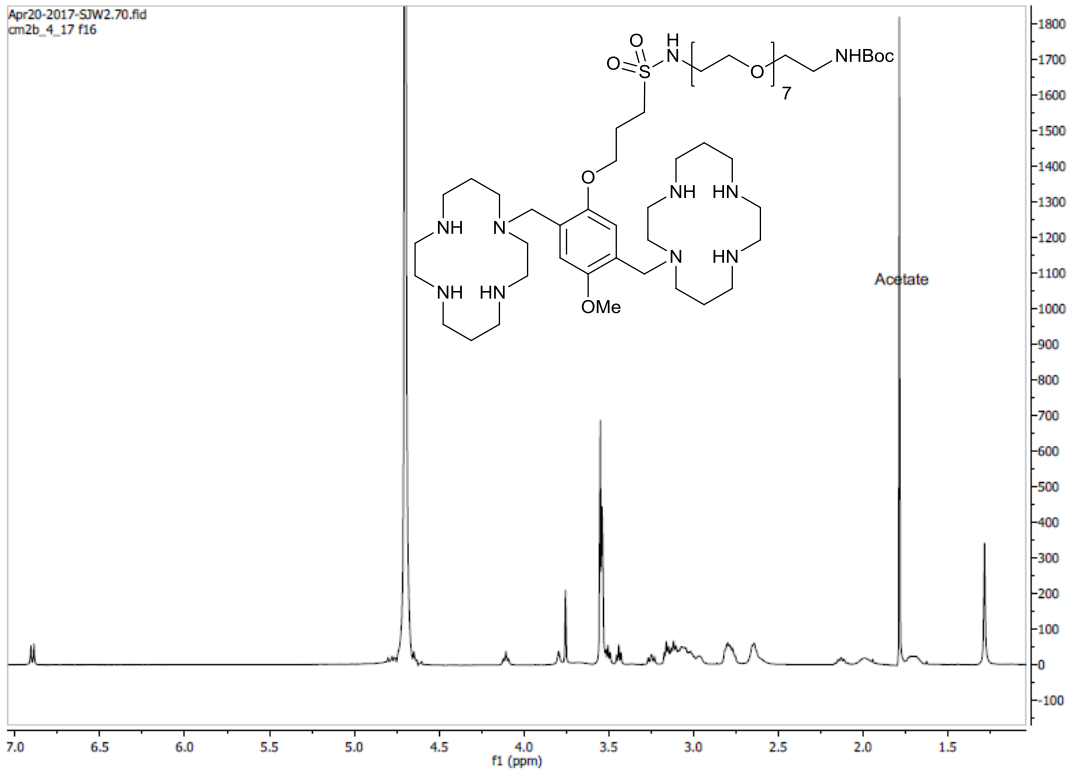

Figure 9: ^{1}H NMR spectrum of **7.**$(AcOH)_n$ (500 MHz, D_2O, 298 K)

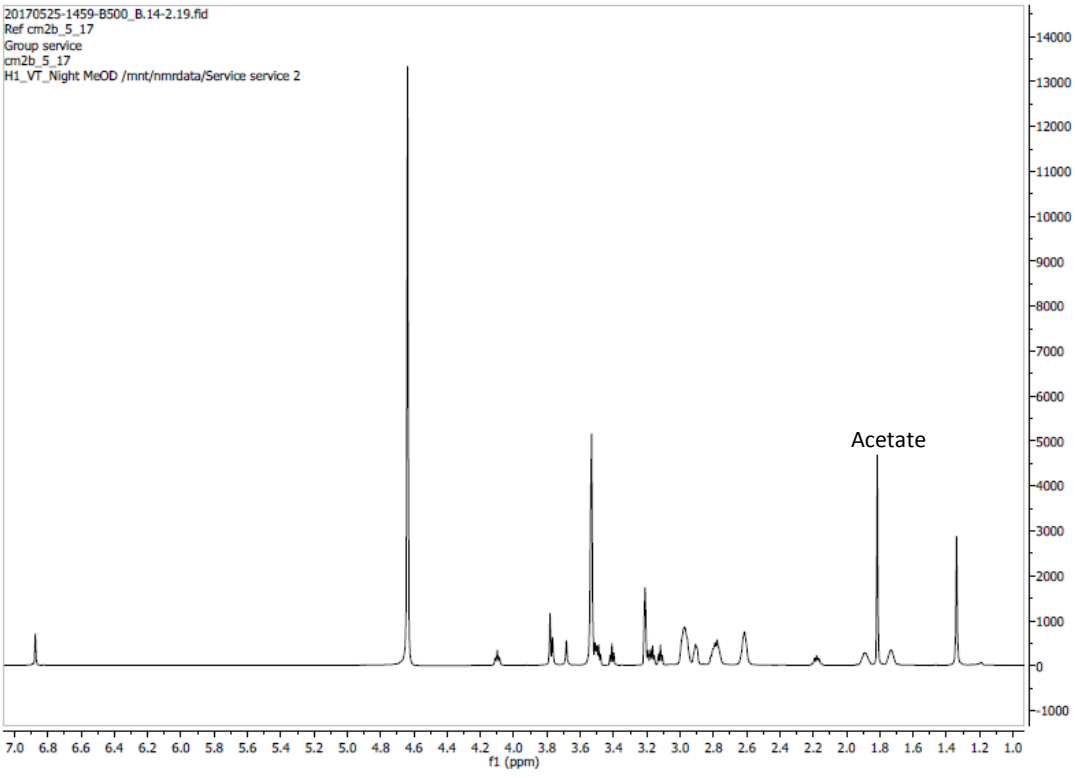

Figure 10: ^{1}H NMR spectrum of **7.**$(AcOH)_3$ (500 MHz, CD_3OD, 348 K)

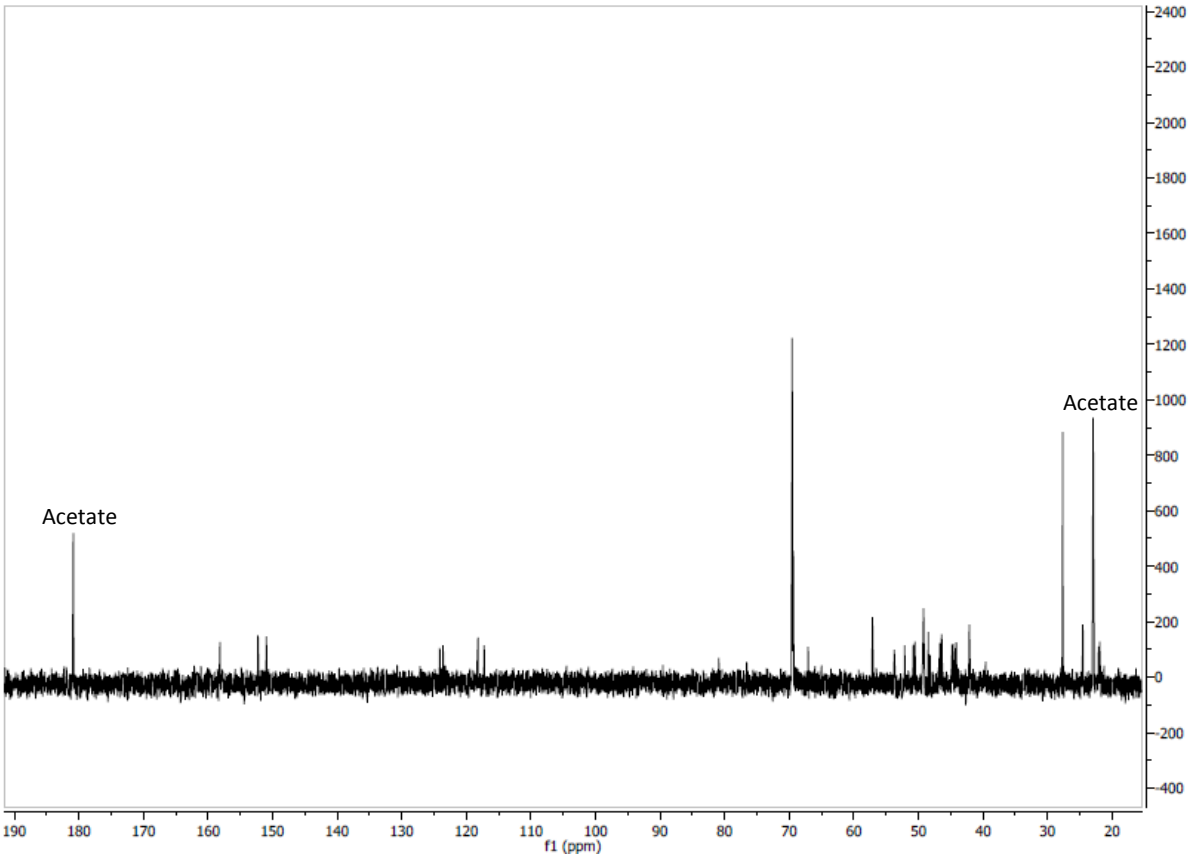

Figure 11: ^{13}C NMR spectrum of **7.**$(AcOH)_3$ (500 MHz, D_2O, 298 K)

S.1.6. ^{1}H and ^{13}C NMR spectra of compound 1

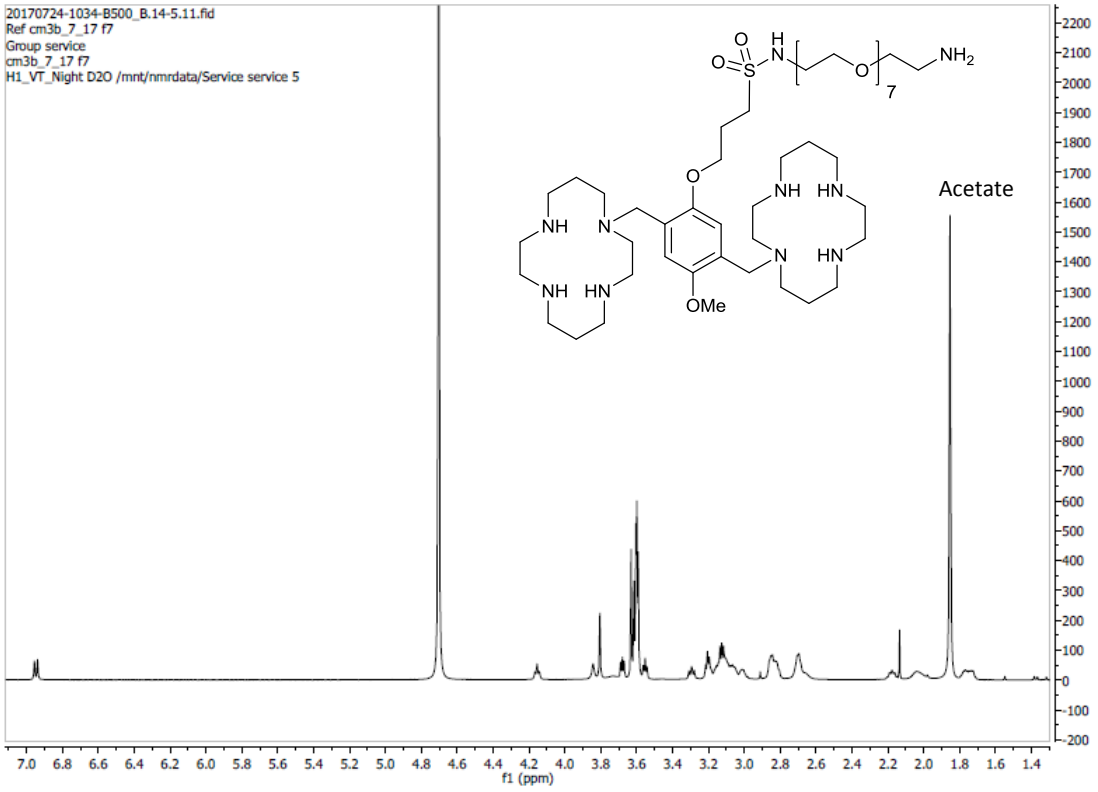

Figure 12: ^{1}H NMR spectrum of **1**.$(AcOH)_n$ (500 MHz, D_2O, 373 K)

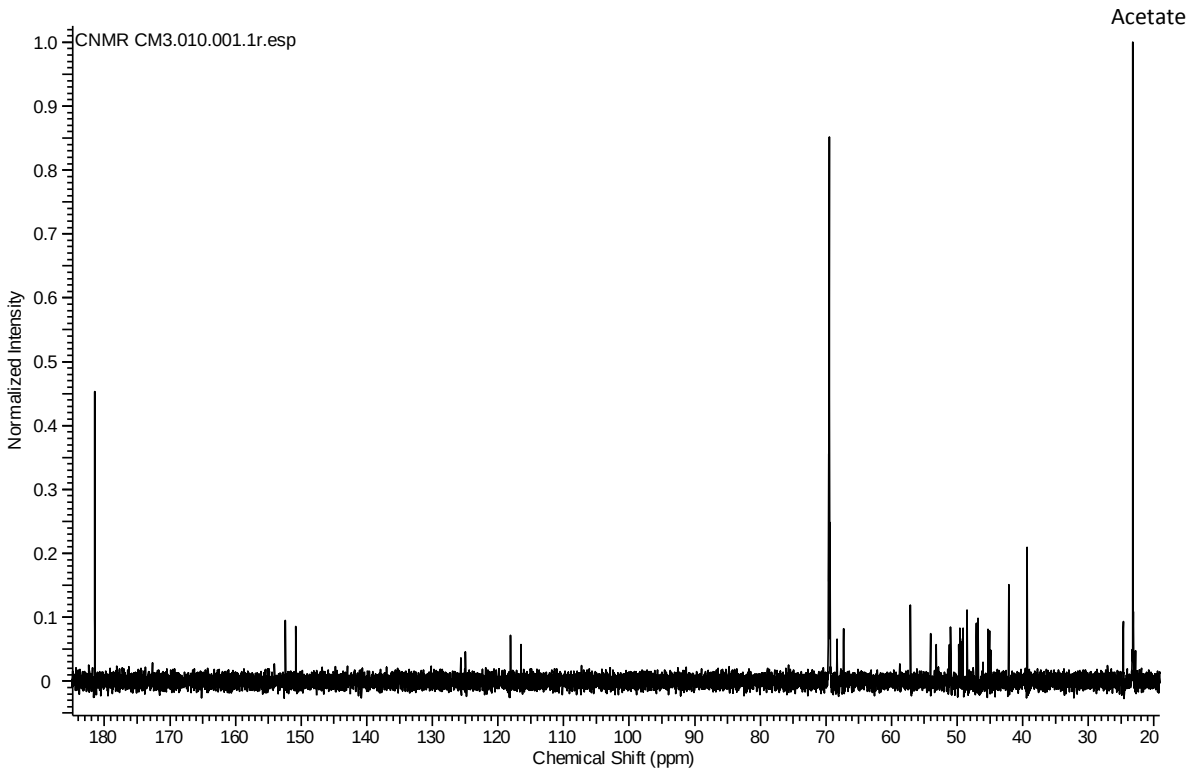

Figure 13: ^{13}C NMR spectrum of **1**.$(AcOH)_n$ (125 MHz, D_2O, 298 K)

S.1.7. ^{1}H and ^{13}C NMR spectra of compound 2

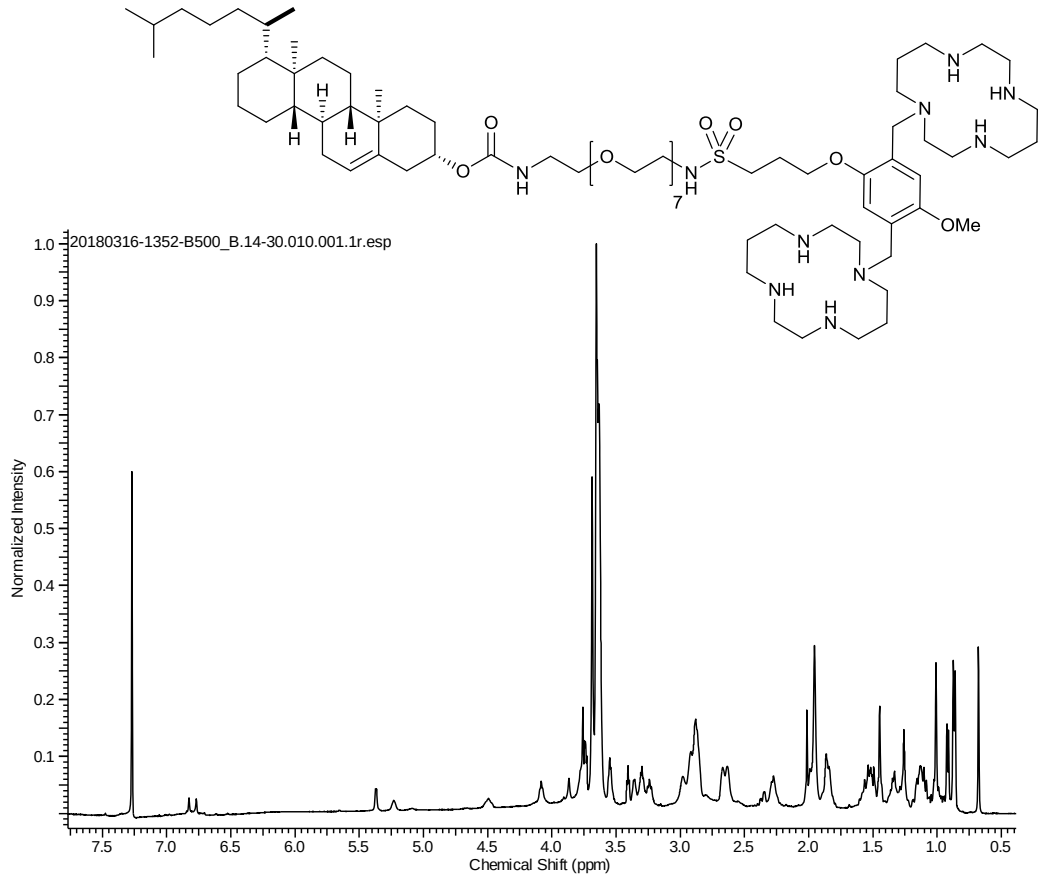

Figure 14: ^{1}H NMR spectrum of **2** (500 MHz, $CDCl_3$, 298 K)

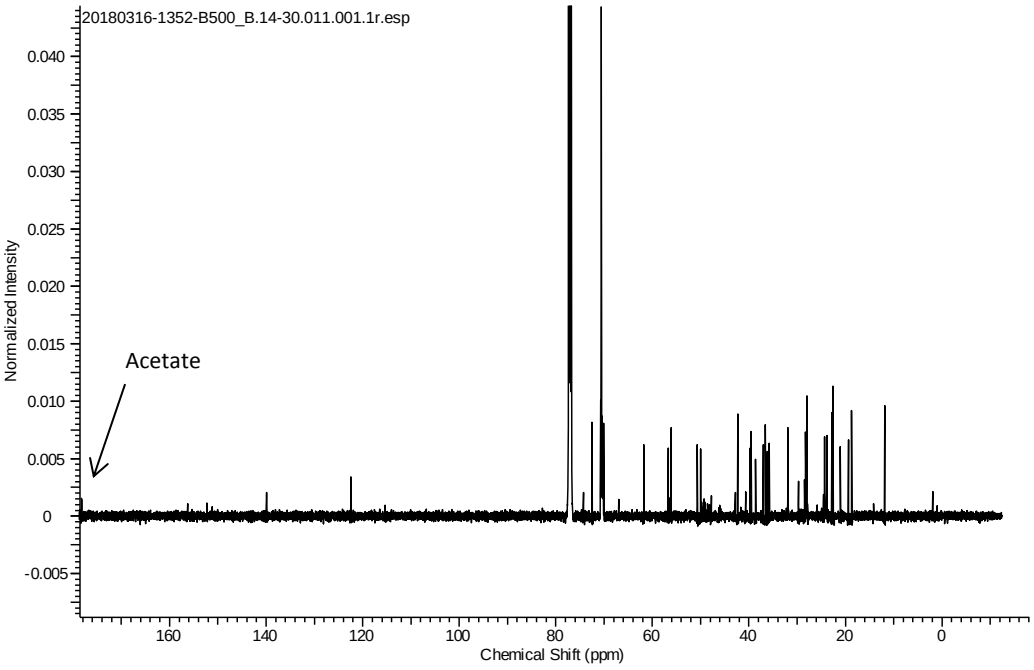

Figure 15: ^{13}C NMR spectrum of **2** (125 MHz, $CDCl_3$, 298 K)

S.1.8. HMBC and HSQC spectra of compound 2

In the bis(cyclam) precursors **7** and **1**, the resonances for the aromatic carbons at positions 2 and 5 on the ring (those aromatic carbons attached to the bis(cyclam) arms) are much weaker than the other aromatic resonances, including those of the quaternary centres attached to oxygen. In the ^{13}C NMR for **2** they were not observed even after long spectral acquisition times (although the 3 and 6 (ArC-H) carbon resonances are found at 115.32 and 116.48 ppm, while the 1 and 4 (ArC-O) carbon resonances are at 151.13 and 152.28 ppm).

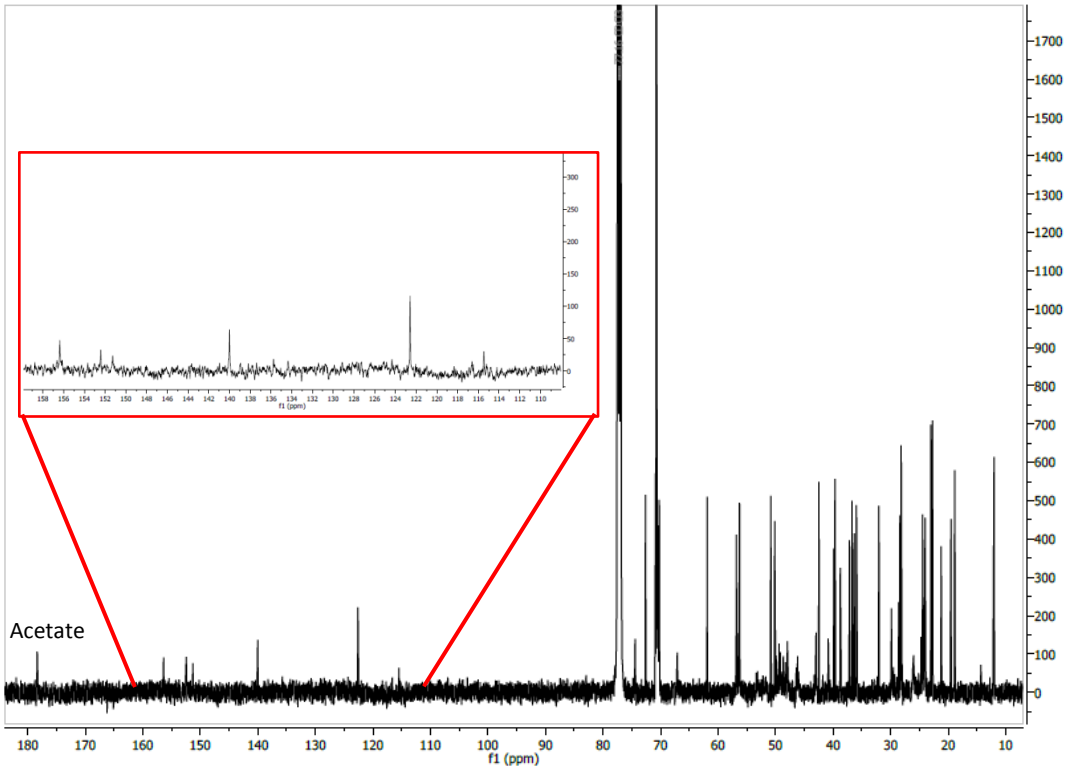

Figure 16: ^{13}C NMR spectrum of **2**, with the aromatic region inset.

To locate the weak carbon resonances, 2D ^{13}C-^{1}H spectra were acquired. The HSQC and HMBC spectra both confirmed the assignments of the aromatic peaks in the ^{13}C spectrum, and in the HMBC spectrum the tertiary carbon peaks were observed. Weak cross-peaks in the HMBC were also observed correlating to coupling between the tertiary carbons and the CH_2 on the cyclam arms, further confirming their position.

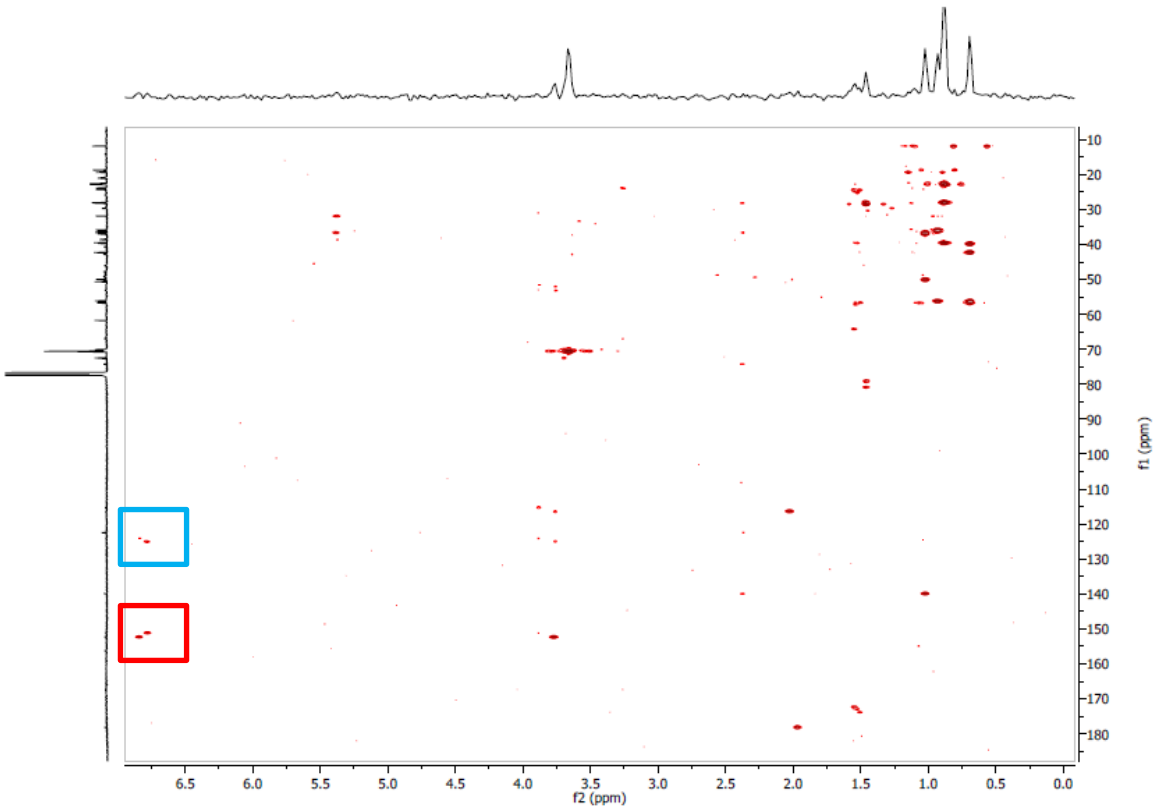

Figure 17: 2D HMBC spectrum of **2.** Both the 1 and 4 (ArC-O) aromatic carbon resonances and the 2 and 5 (ArC-C) aromatic carbon resonances are observed (boxed).

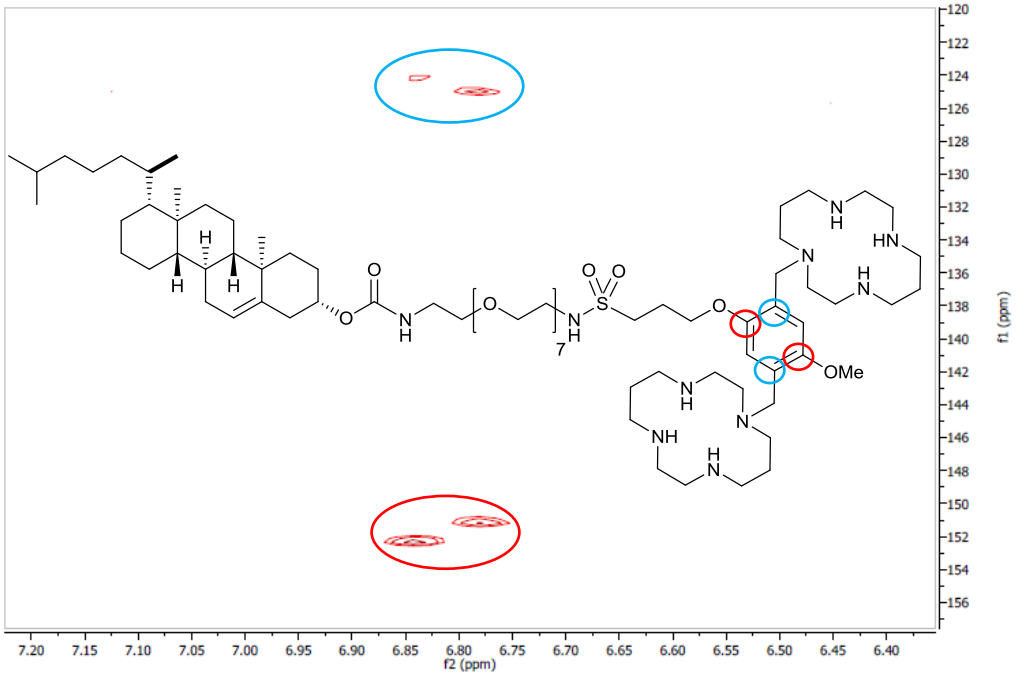

Figure 18: Detailed image showing the aromatic resonances associated with the 2 and 5 carbons (124.06 and 124.98 ppm, circled in blue) and the 1 and 4 carbons (151.13 and 152.28 ppm, circled in red).

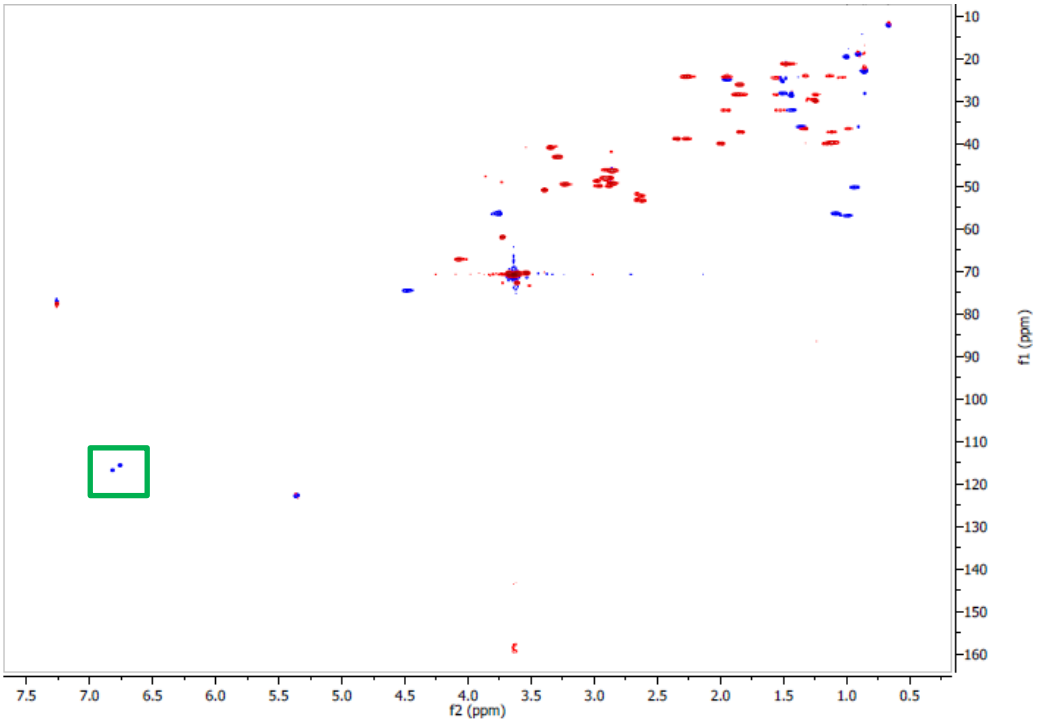

Figure 19: 2D HSQC NMR spectrum of **2**, with the peaks associated with the 3 and 6 (ArC-H) aromatic carbon resonances (115.32 and 116.48 ppm) indicated by the green box.

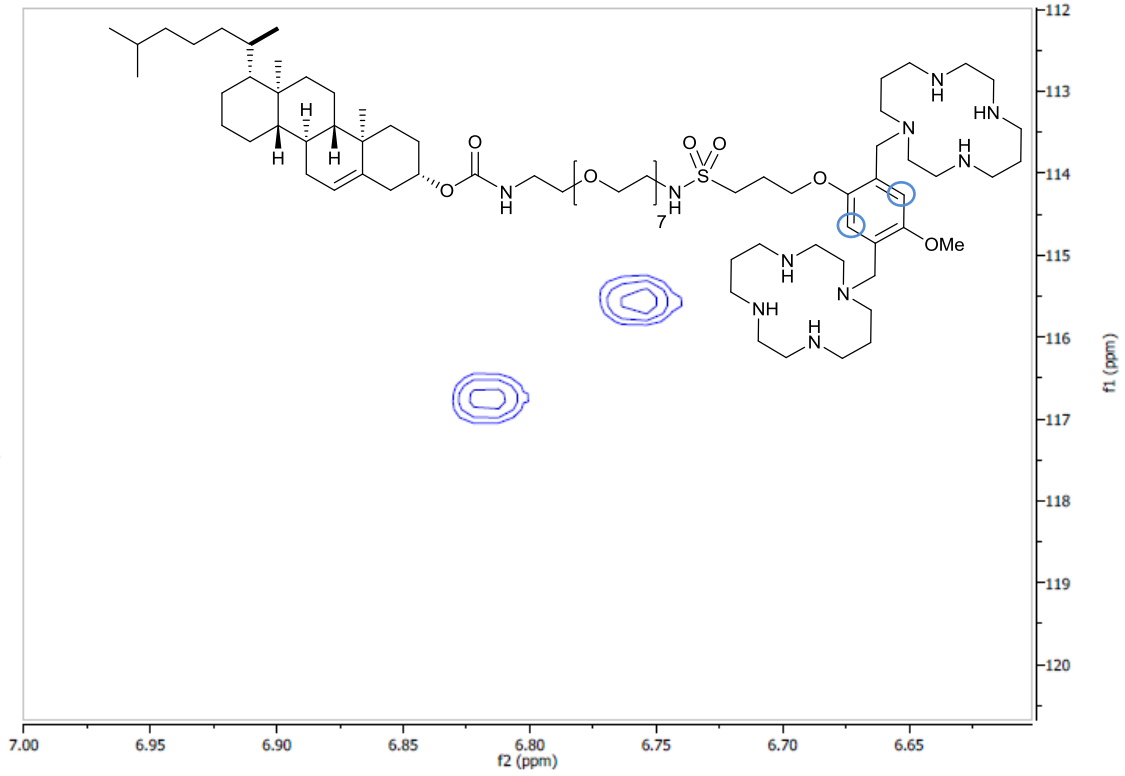

Figure 20: HSQC NMR peaks associated with the 3 and 6 (ArC-H) aromatic carbon resonances of **2** (circled in blue on structure)

S.2. NMR and VT ^{1}H NMR spectroscopy data for compound 6

Variable temperature (VT) NMR was performed to elucidate the causes for the peak broadness in the ^{1}H NMR spectrum of **6** and to facilitate peak assignment. With increasing temperature, the peaks across the spectrum become sharper, including in the aromatic and PEG regions. In particular, the peaks assigned to the cyclam region (3.0-3.3 ppm and 2.5-2.1 ppm) became much sharper, with multiplet structures far more easily discernible. The principal reason for the peak broadness of B is suggested to be due to rotamers of the Boc groups on the cyclam groups.

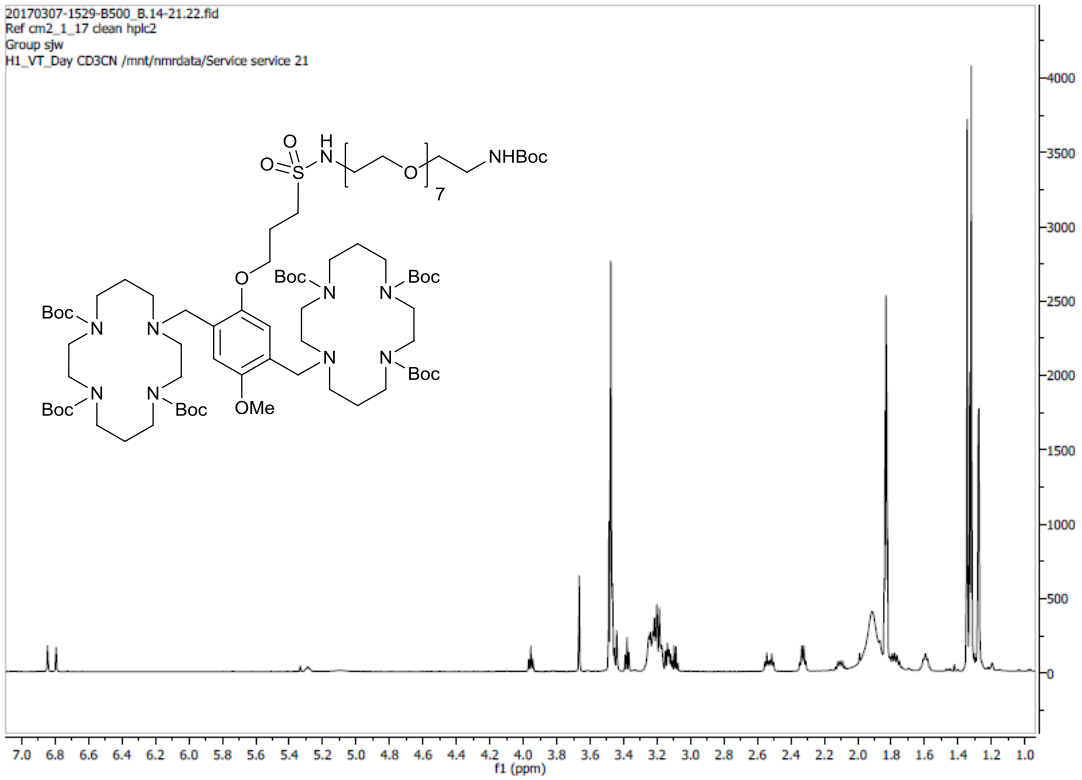

Figure 21: ^{1}H NMR spectrum of **6** (400 MHz, CD_3CN, 345 K)

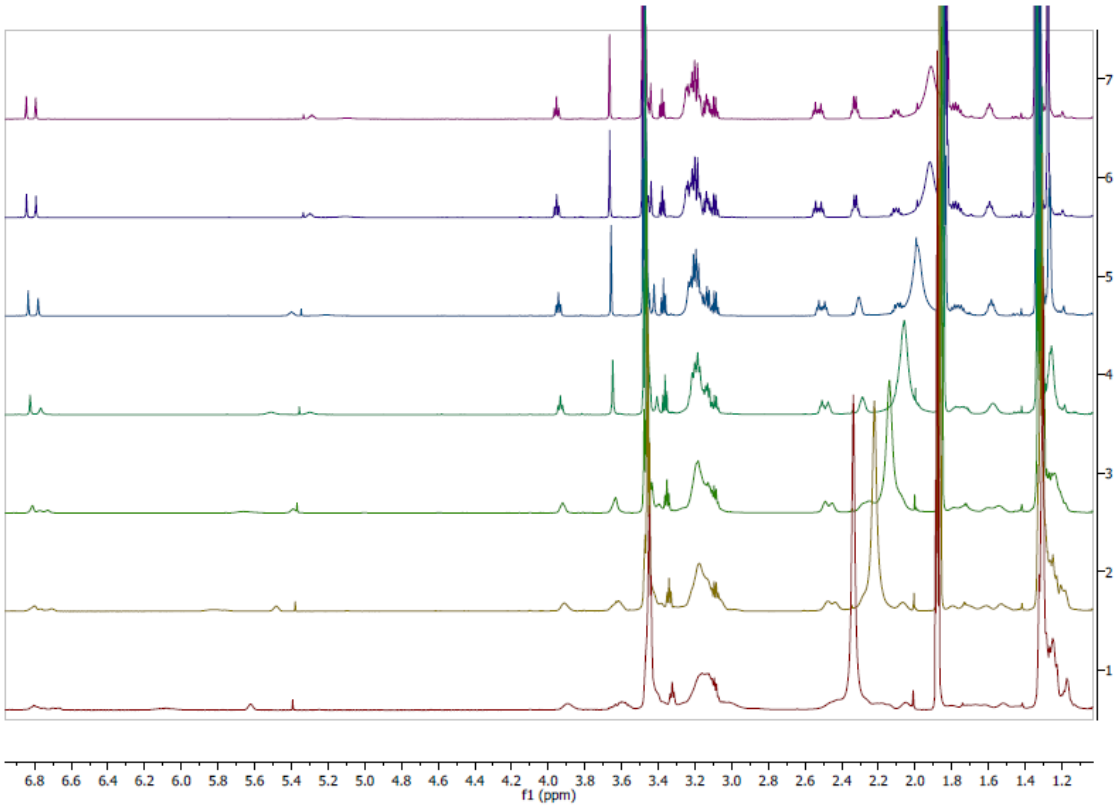

Figure 22: ^{1}H VT NMR spectra of **6**; VT steps: 235, 253, 273, 293, 313, 343 & 345 K (500 MHz, CD_3CN); full spectra

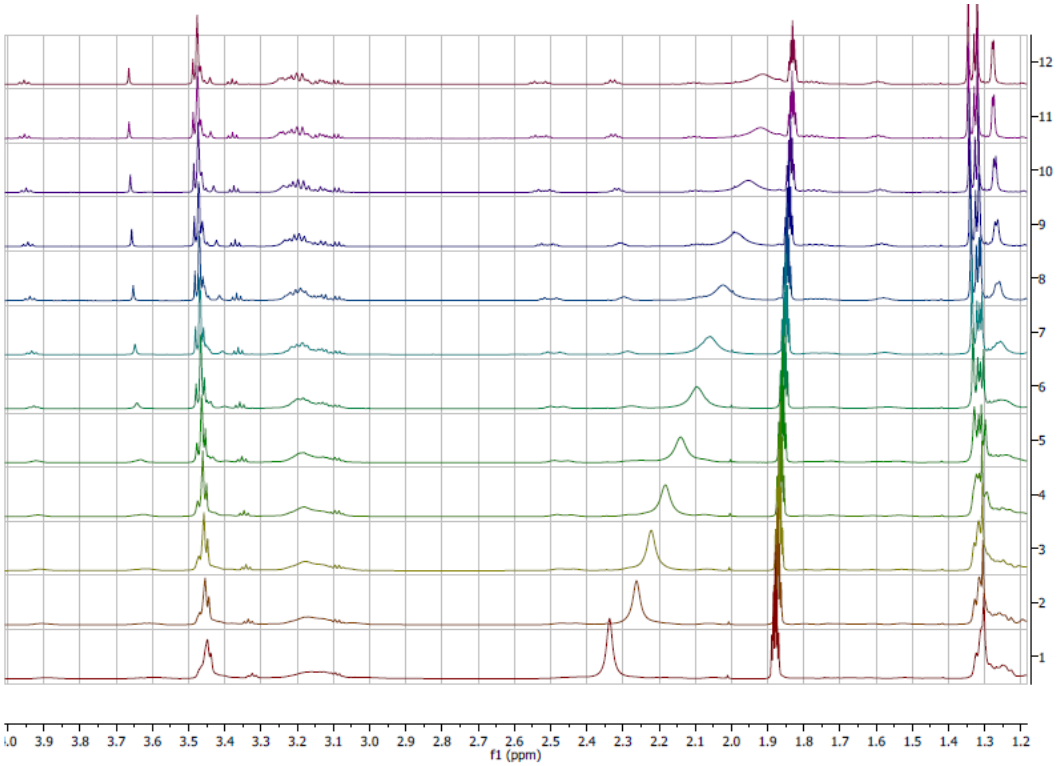

Figure 23: ^{1}H VT NMR spectra of **6**; VT steps: 235, 243, 253, 263, 273, 283, 293, 303, 313, 323, 343, 345 K (500 MHz, CD_3CN).

S.3. Spectroscopic and HPLC data for crude product 5a before Finkelstein reaction

After conjugation of **4** to **3a**, the mass spectrometry (Figure 24), HPLC (Figure 25) and NMR data (Figure 26) suggested a mixture of products had formed, specifically that the bromide had been partially displaced by chloride (generated during reaction of the sulfonyl chloride). In the mass spectrum, a range of species are observed (Figure 24, summarised in Table 1). In the NMR spectrum, the presence of a mixture was indicated by a multiplet in the aromatic region (6.95-6.84 ppm) and the CH_2X protons give rise to two pairs of singlets (4.403, 4.408 ppm and 4.501, 4.496 ppm) in a 1:1 ratio (in **5a** the CH_2Br proton resonances appear at 4.42 and 4.45 ppm). These data suggest that pre-Finkelstein, the **5a** reaction mixture has a *ca*. 1:1 ratio of CH_2Br: CH_2Cl.

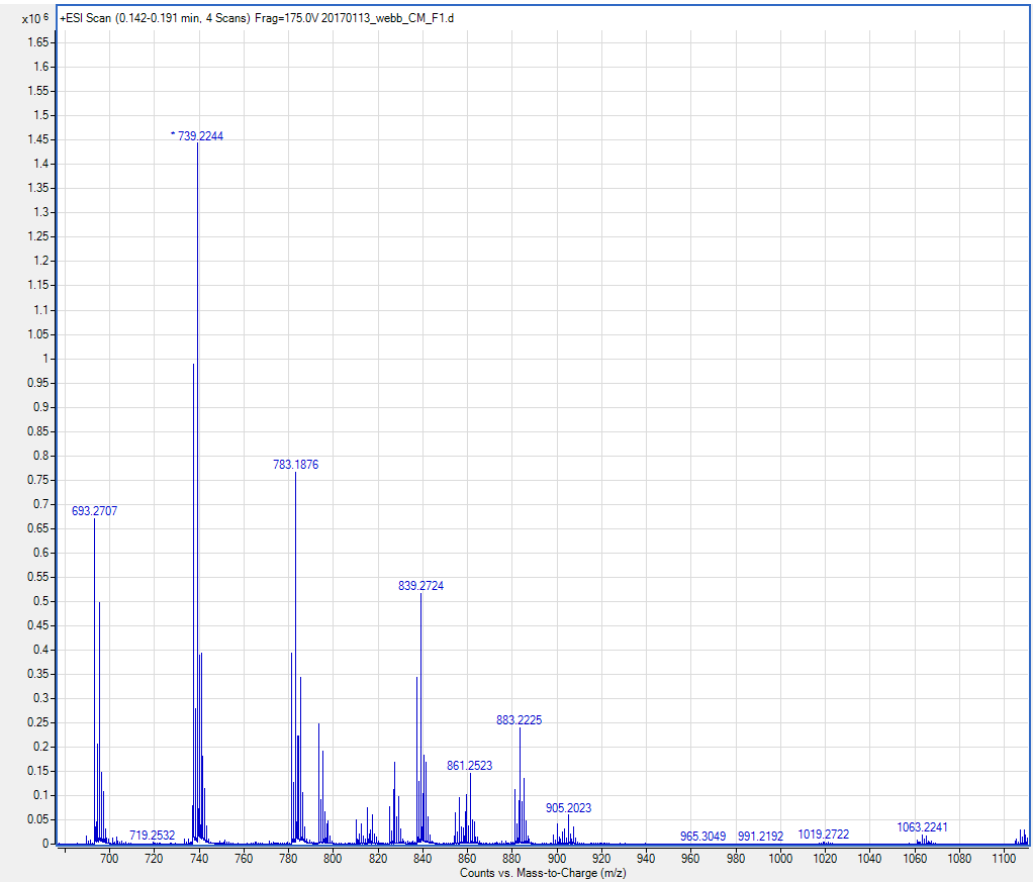

Figure 24: HRMS of crude product **5a** before Finkelstein reaction.

Table 1: Summary of compounds observed in HRMS spectrum of crude **5a**

m/z	Compound
693	[**5a – 2Br + 2Cl – Boc** + H]$^+$
739	[**5a – Br + Cl – Boc** + H]$^+$
783	[**5a – Boc** + H]$^+$
839	[**5a – Br + Cl** + H]$^+$
861	[**5a – Br + Cl** + Na]$^+$
883	[**5a** + H]$^+$
905	[**5a** + Na]$^+$

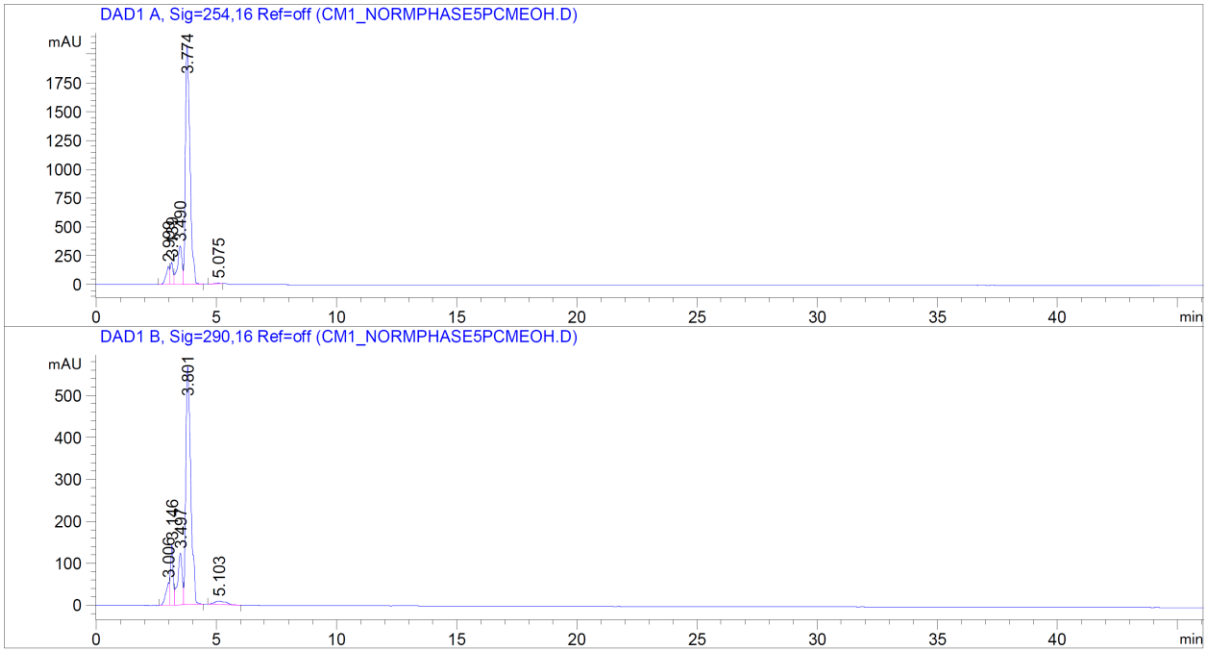

Figure 25: HPLC of crude product **5a** before Finkelstein reaction

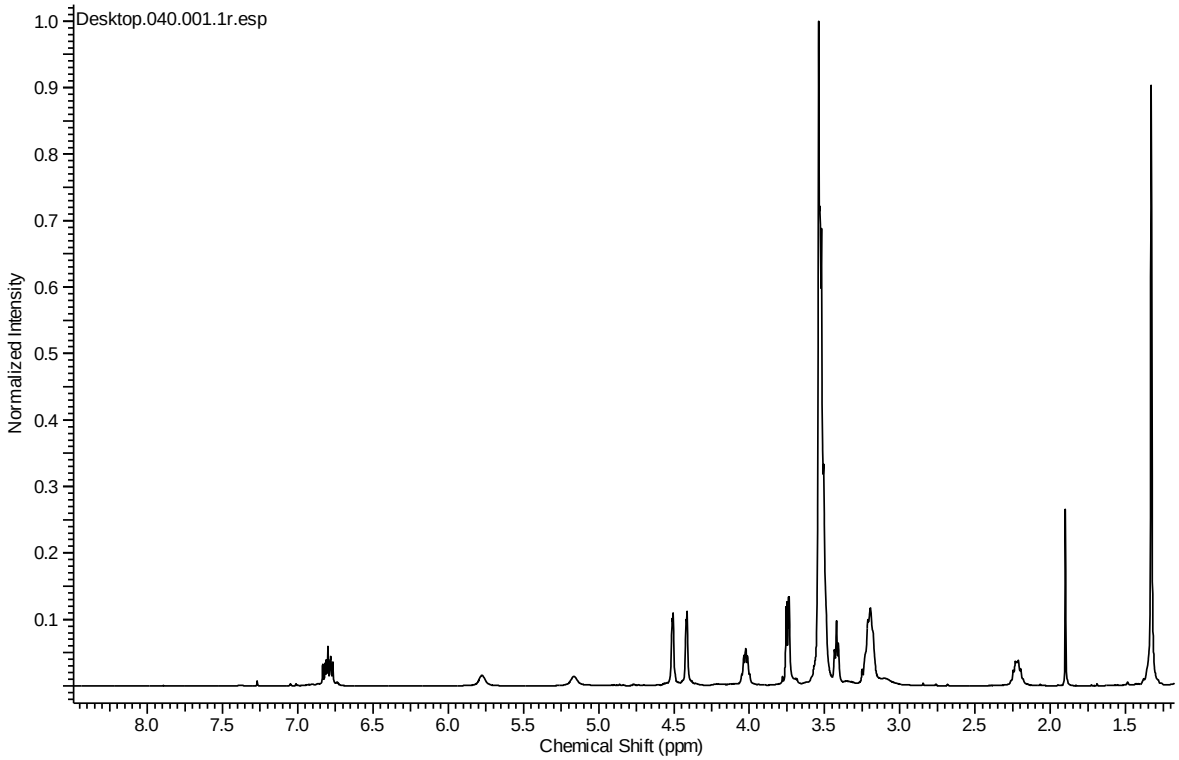

Figure 26: ^{1}H NMR spectrum of crude product **5a** before Finkelstein reaction (400 MHz, $CDCl_3$, 298 K)

S.4. Analytical HPLC traces for compounds 5a, 7 and 1

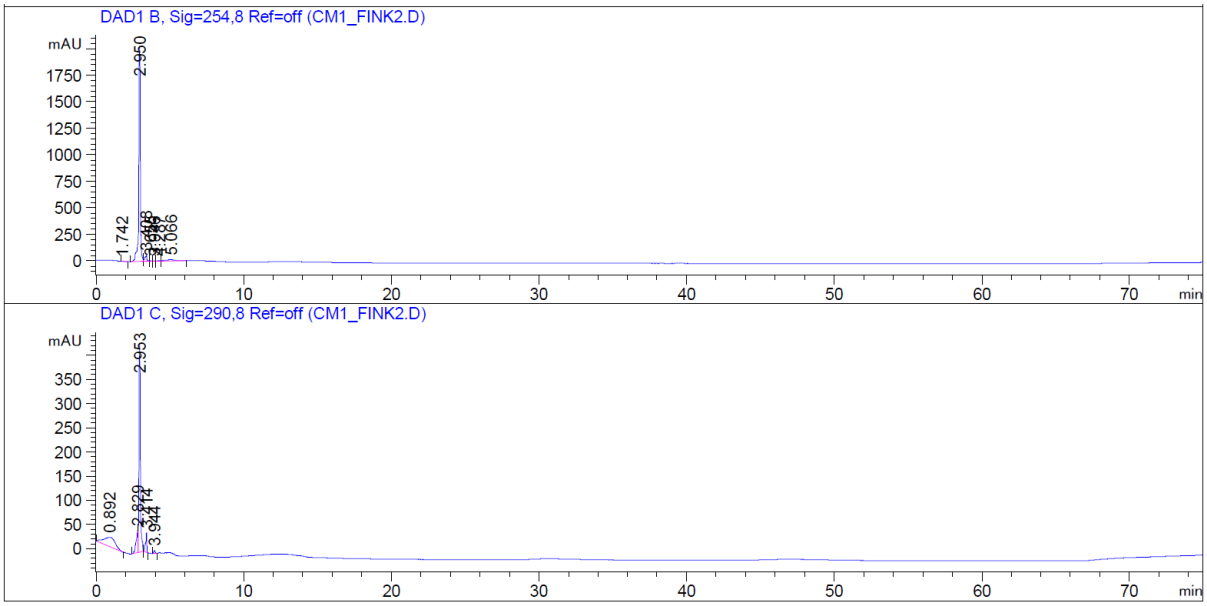

Figure 27: HPLC trace of **5a**

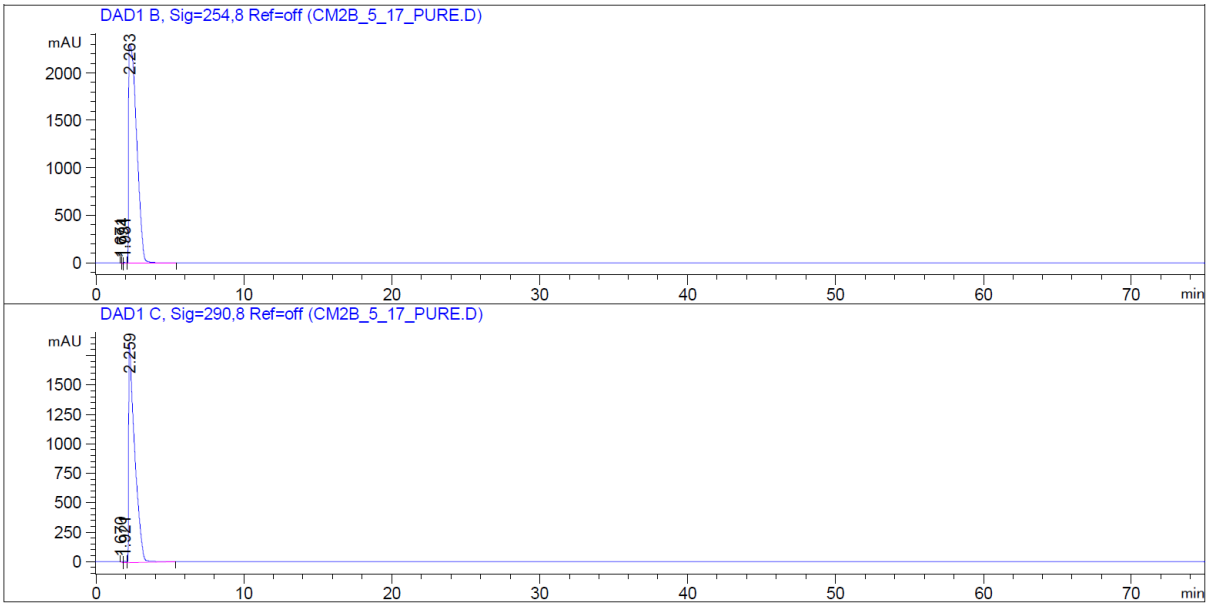

Figure 28: HPLC trace of **7**

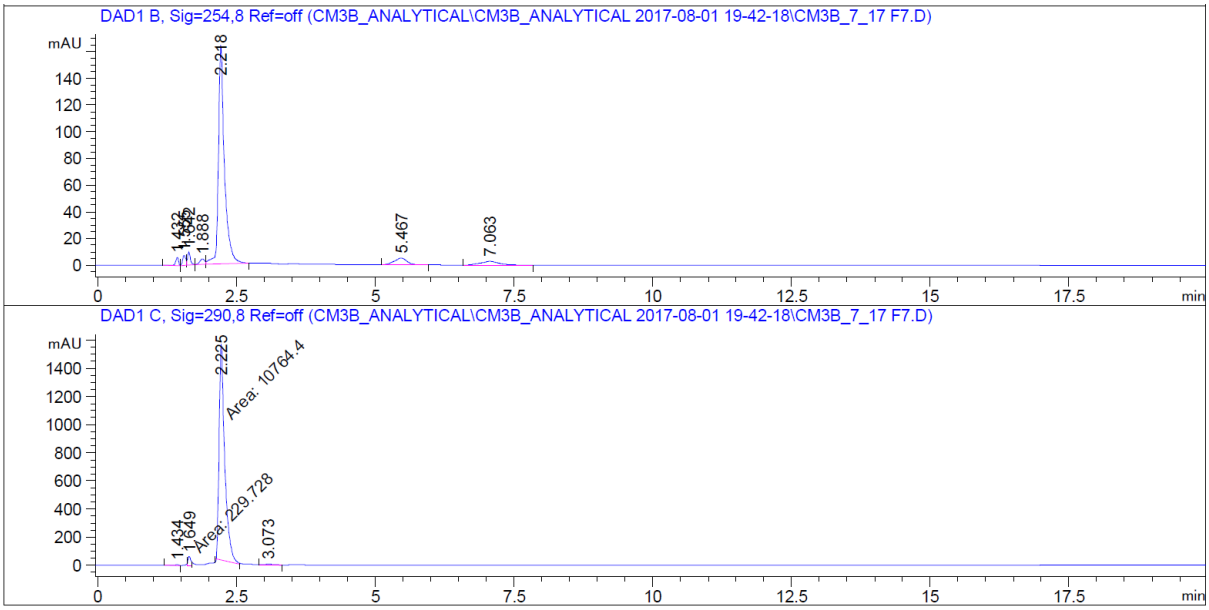

Figure 29: HPLC trace of **1**

S.5. Dynamic light scattering (DLS) histograms for 2

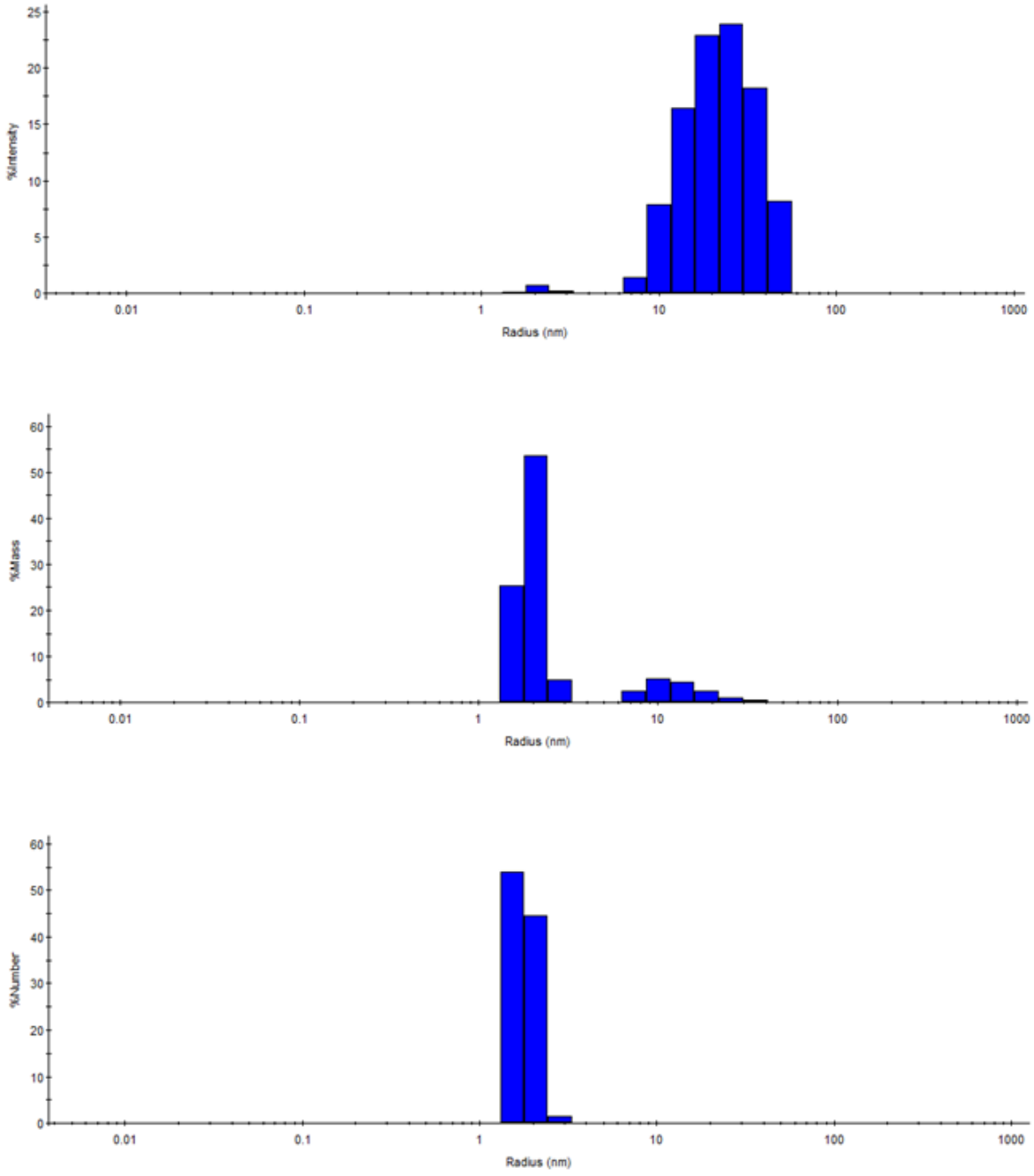

Figure 30: A representative histogram set showing the r_h distribution with respect to a) intensity b) mass c) number for a suspension of **2** illustrating the bimodal distribution of hydrodynamic radii.

S.6. AFM images of dropcast suspensions of 2

AFM was used to investigate the structure of the objects formed by the self-assembly of **2** (Figure 31a). A high concentration of **2** in water (100 µM) was used, well above the critical aggregation concentration. The suspensions were drop-cast and imaged in air, using the ScanAsyst mode. The AFM tip used was a ScanAsyst-Air, with tip radius 2 nm. These studies revealed a large number of small self-assembled particles, with radii typically in the range 6-40 nm, as well as some larger aggregates (e.g. 160 nm radius). The height of these particles was typically only a few nanometers.

Particle analysis was performed on the image shown in Figure 31a using ImageJ thresholding and particle analysis tools – radii were extracted from calculated areas, assuming circularity of the measured objects. Within the resolution of the AFM (> 3 nm), the radii of the observed particles fall within the nanoscale region, similar to the radii measured using DLS (Figure 31b). The geometric mean of the entire population is 8 nm, which lies between the two major populations observed in DLS. However, it should be noted that both the effects of drying and interactions with the mica substrate surface may have changed particle size and shape in comparison to the particles observed in aqueous suspension (for example, the particles are flattened in the AFM).

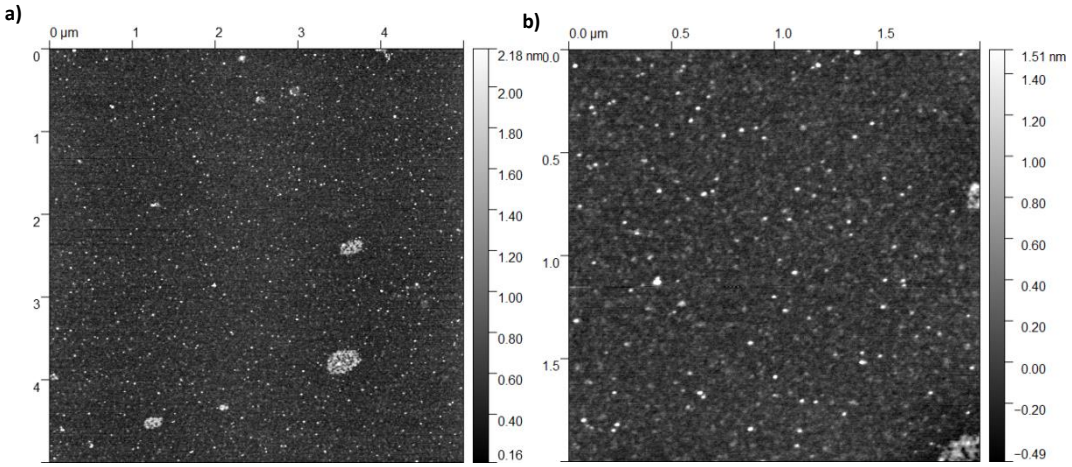

Figure 31: An AFM image of **2** drop-cast onto freshly cleaved mica, where b) depicts the central region of a), imaged at higher magnification. A large number of small self-assembled particles are observed, alongside larger loosely assembled aggregates.

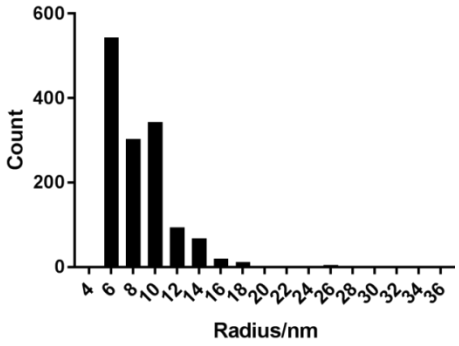

Figure 32: Particle analysis performed on image Figure 31a: the histogram has been clipped at 36 nm for clarity.

S.7. Antibody staining of CXCR4

Chronic lymphocytic leukaemia cells are known to show high levels of CXCR4 expression, however different primary cell samples can differ in their expression levels. Therefore the CLL cells that were to be used to analyse the affinity of **1** and **2** for CXCR4 were assessed for their CXCR4 expression levels prior to subsequent experiments. As shown in Figure 33, these CLL cells have very high levels of CXCR4 present, as expected for this disease type.

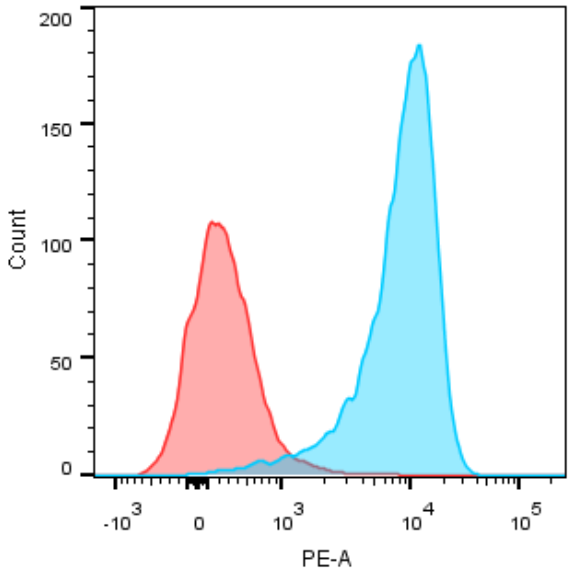

Figure 33: CXCR4 staining on live primary CLL cells after 24 h culturing, where the blue curve shows CXCR4 expression without stimulation or blocking, and the pink shows CXCR4 staining after incubation with 10 μM **1**.

S.8. Western Blot: B-cell receptor control and full protein stain

In order to further investigate CXCR4 antagonism by **2** in the presence of CXCL12, Western Blots were performed. Monitoring p-ERK levels can provide evidence that downstream signalling associated with CXCL12 activation has been disrupted. To ensure that the effect of **2** was specific to CXCR4, and not a result of non-specific signalling inhibition, the effect of both **2** and AMD3100 was assessed during the stimulation of the B-cell receptor.

The details for the CXCR4 Western Blot assay are described in the manuscript text. For the B-cell receptor Western Blot assay, a similar procedure was followed; cells were cultured overnight at a density of 8×10^6 cells/mL in a 24 well plate. The B-cell receptor was stimulated by anti-IgM binding and biotin-avidin crosslinking as previously described in the literature (Wang *et al.*, *J. Biol. Chem.,* 2012*,* **287**, 11037-11048). An aliquot of biotin-goat anti-human IgM solution (final conc. 10 µg/mL, Bio-Rad Laboratories, Inc., UK) was added to 6×10^6 cells for 30 minutes on ice, followed by an aliquot of avidin solution (final conc. 25 µg/mL, Sigma Aldrich, UK) for 30 minutes at 37 °C. The cells were then placed back on ice, pelleted and washed three times with ice-cold PBS. Cell lysis and Western Blot were performed as described in the Experimental Section of the manuscript.

A comparison of both sets of blots shows that while both **2** and AMD3100 reduce CXCL12-induced signalling to similar levels as nil stimulation (Figure 34b), there is no effect on signalling caused by activation of the B-cell receptor (Figure 34a). In addition, total levels of ERK were unaffected by the addition of **2** or AMD3100 in both experiments. The assays show that both **2** and AMD3100 acted as pure antagonists to CXCR4 with no non-specific knockdown of signalling following stimulation of the B-cell receptor.

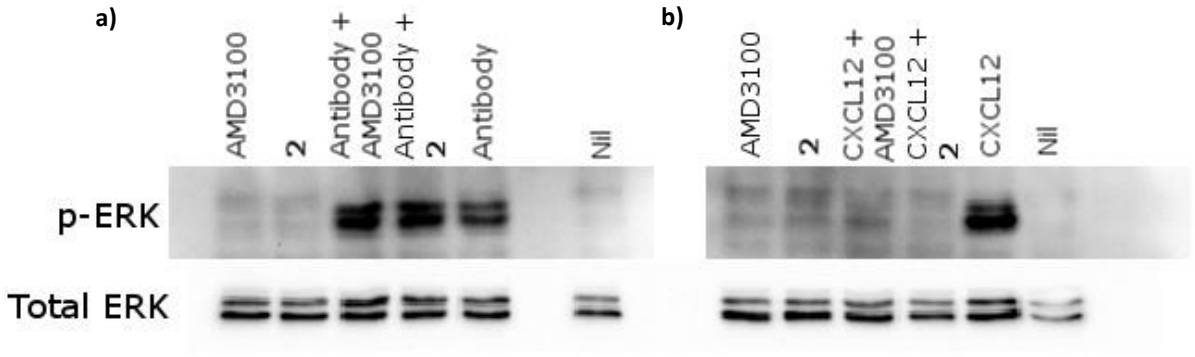

Figure 34: Western Blots showing phosphorylated and total ERK, following stimulation with a) biotinylated anti-IgM antibody plus avidin and b) CXCL12.

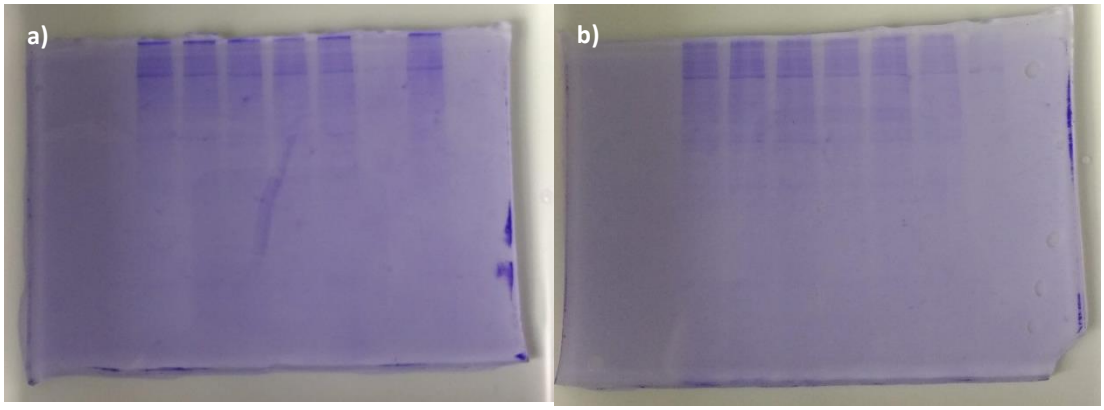

Figure 35: Full protein stain of the SDS-PAGE gels used to separate the proteins for both experiments; a) B-cell receptor stimulation, b) CXCR4 stimulation. Stain was performed by treating the gels with water/acetic acid/methanol followed by staining with Coommassie Brilliant Blue R-250 overnight at room temperature.