

## Supporting Information

for

### **An easily accessible, lower rim substituted calix[4]arene selectively binds *N,N*-dimethyllysine.**

Alok Shaurya, Graham A.E. Garnett, Melissa J. Starke, Mark C. Grasdahl, Charlotte C. Dewar, Anton Y. Kliuchynskyi, and Fraser Hof\*

---

## Table of Contents

<b>Experimental procedures</b> .....	<b>3-16</b>
<b>S1. General remarks</b> .....	<b>3-5</b>
S1.a. Chemicals.....	3
S1.b. Liquid Chromatography.....	3
S1.c. Solid Phase peptide synthesis.....	3
S1.d. Indicator Displacement Assay.....	3-4
S1.e. Isothermal calorimetry.....	5
<b>S2. Synthesis and characterization of compound 1</b> .....	<b>5-7</b>
<b>S3. Synthesis and characterization of control compound 2</b> .....	<b>8-9</b>
<b>S4. Characterization and purity of peptides</b> .....	<b>10-17</b>
<b>Result and Discussions</b> .....	<b>18-21</b>
<b>S5. Plots for Indicator displacement assay</b> .....	<b>17-18</b>
S15.1 Indicator Displacement Assays for <b>sCx4</b> .....	17
S15.2. Indicator Displacement Assays for compound <b>1</b> .....	18
<b>S6. Plots for Isothermal calorimetry</b> .....	<b>20-21</b>
S16.1 Isothermal calorimetry between <b>1</b> and H3K4me3.....	20
S16.2. Isothermal calorimetry between <b>1</b> and H3K4me2.....	21
<b>S7. NMR studies on host guest complex</b> .....	<b>22-23</b>
<b>References</b> .....	<b>23</b>

---

## Experimental Procedures

### S1. General remarks

**S1.a. Chemicals.** p-sulfonatocalix[4]arene (**sCx4**) was purchased from TCI America, (4-chlorosulfonyl)benzoic acid was purchased from Alfa Aesar and Tosyl chloride was purchased from ACP chemicals. All standard amino acids and solid phase peptide coupling reagents including Rink amide resin were purchased from Chem-Impex International as Fmoc-protected versions. Fmoc-protected versions of monomethyllysine, dimethyllysine and acetyl lysine were also purchased from Chem-Impex. Fmoc-protected trimethyllysine was bought from GL Biochem. All methylated arginines were purchased from AnaSpec Inc. All peptides used in IDA were synthesized in-house while the peptides used in ITC and NMR studies were purchased from GenScript.

**S1.b. Liquid chromatography.** All compounds were purified using RP-HPLC on Phenomenex Luna C18(2) column (4.6 mm x 250 mm) with 5  $\mu\text{m}$  particle and 100 Å pore size. The preparative RP-HPLC purifications were performed using a Shimadzu Prominence LC system using a gradient of acetonitrile in water as mobile phase. Both mobile phase eluents were spiked with 0.1% TFA. The compounds were detected using a DAD set to 280 nm.

Analytical LC-MS characterization was done on a Waters Acquity-H UPLC-MS, using a BEC C18 column with 1.7  $\mu\text{m}$  particle size. Like the preparative HPLC, a gradient of acetonitrile in water was used as mobile phase. Both mobile phase eluents were spiked with 0.1% TFA. The compounds were detected using a DAD set to 280 nm and an inline mass spectrometer.

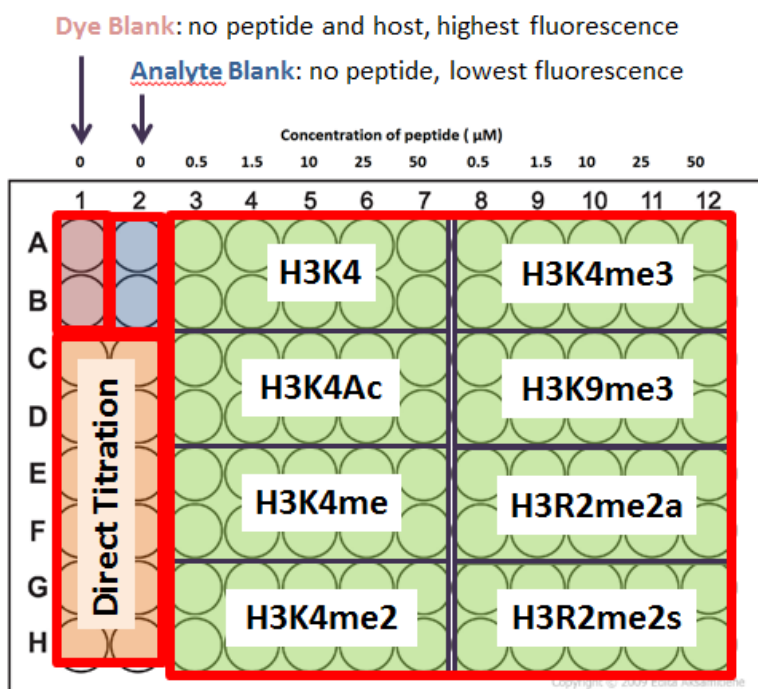
All enrichment studies were done using an AKTA Prime FPLC. The solvents used were A. 50 mM Sodium phosphate buffer pH 7.4 (binding buffer) and B. 50 mM Sodium phosphate + 2 M  $\text{NH}_4\text{Cl}$  (elution buffer). In a typical enrichment experiment, after injection the column was eluted with binding buffer for 30-minutes before switching to elution buffer, which was introduced slowly in 30-minute gradient. The flow rate was kept constant at 0.3 mL/min.

**S1.c. Solid Phase peptide synthesis.** All peptides used for IDA studies were made on an automated solid phase peptide synthesizer LIBERTY1 (CEM). The synthesis was done on Rink Amide resin (Chem Impex) using Fmoc chemistry and microwave assisted coupling methods programmed in the machine. After synthesis, the peptides were cleaved from the resin as C-terminal amide using a 95:2.5:2.5 mixture of TFA,  $\text{H}_2\text{O}$  and triisopropylsilane. The reaction continued for 4 hours under gentle bubbling of  $\text{N}_2$ . After that, the resin was filtered and washed with TFA. The combined TFA solution was concentrated on a rotavap. Cold diethyl ether was then used to precipitate crude peptides which were then air dried and then purified as described in SI1.2. The gradient used for purification started from 5% acetonitrile in water (with 0.1% TFA), increased to 20% acetonitrile in 15 minutes and then up to 90% in next 15 minutes. The peptides were characterized on Waters UPLC (as described in SI1.2). H3K4me3 peptide used in ITC and NMR studies was synthesized as above. H3K4me2 peptide used in ITC and NMR studies was bought from GenScript ([www.genscript.com](http://www.genscript.com)).

**S1.d. Indicator displacement assay (IDA).** IDA was performed in NUNC 96 black-well plated with optically clear bottoms. A total of 9 titrations were performed in a single plate which included the direct titration of calixarene into lucigenin, and 8 competitive titrations of individual peptides into the calixarene-lucigenin complex. The plate layout is shown in Figure SI1.4.1. Concentration of stock solutions were determined spectrophotometrically using  $A_{280}$  for peptides ( $\epsilon = 1280 \text{ M}^{-1}\text{cm}^{-1}$ ),  $A_{410}$  for lucigenin ( $\epsilon = 8900 \text{ M}^{-1}\text{cm}^{-1}$ ) and using qNMR for calixarenes with TSP as the internal integration standard.

Concentration of lucigenin was kept constant at 250 nM in all wells while the buffer concentration was kept constant at 10 mM sodium phosphate buffer at pH 7.4. To achieve consistency all reagents were added from one stock solution each with 10X concentration than required ( $[\text{lucigenin}]^{\text{stock}} = 2.5 \mu\text{M}$  and  $[\text{buffer}]^{\text{stock}} = 100 \text{ mM}$ ). 20  $\mu\text{L}$  of each solution was added to all wells. Direct titration wells had varying concentration of calixarene and no peptides. All competitive titration wells had a fixed concentration of calixarene (0.5  $\mu\text{M}$  working concentration added in from 5  $\mu\text{M}$  stock solution). Like calixarene and lucigenin, varying concentration of peptides were delivered from 10X concentrated stock solution (as shown in figure SI1.4.1). All the remaining volume was made up using  $\text{dH}_2\text{O}$  such that total volume in each well was 200  $\mu\text{L}$ .

Fluorescence of lucigenin was read across the plate using SpectraMax<sup>®</sup> M5/M5e microplate reader in top read mode. The excitation wavelength used was 369 nm while the emission was recorded at 485 nm. The plates were briefly centrifuged immediately prior to reading in a Beckman Coulter Allegra<sup>™</sup> X-12R centrifuge. The raw data was then fed into an in-house written python code which split the data into 9 different concentration response which were plotted as  $\text{dF}_{\text{obs}}$  vs [calixarene], for direct titration, and vs [peptide], for competitive titrations. The program used established numpy and scipy libraries to fit the direct titration data using 1:1 binding isotherm (*vide infra*) to obtain  $K_{\text{ind}}$ . This  $K_{\text{ind}}$  was then used further to fit the remaining 8 data sets with competitive binding isotherms (*vide infra*) and extract  $K_{\text{d}}$  for each peptide and calixarene.



**Figure S1.** Plate layout for Indicator Displacement Assay experiment

All IDA data were fit as explained below:

1. For all the equations shown below:

- $[H]^T$  is the total host concentration (host = calixarene,  $[H]^T = [H]^{free} + [H]^{bound}$ )
- $[I]^T$  is the total indicator concentration (indicator = lucigenin,  $[I]^T = [I]^{free} + [I]^{bound}$ )
- $[Pep]^T$  is the total peptide concentration ( $[Pep]^T = [Pep]^{free} + [Pep]^{bound}$ )
- $[HG]$  refers to host-guest complex. This would be calixarene-lucigenin complex in direct titration and calixarene-peptide complex in competitive titration.
- $[H]^{free}$  is written as  $[H]$  for simplicity.
- $dF_{obs}$  is change in fluorescence observed upon binding (negative for direct titration, positive for competitive titration). This is the raw data from the fluorometer.
- $\Delta F_{max}$  is the maximum change in fluorescence observed i.e. point where the curve plateaus. This is an iterable variable for all titration and was optimized.
- $K_{ind}$  is dissociation constant of host-indicator (calixarene-lucigenin) complex. This is an iterable variable in direct titration and was optimized. In all competitive titrations, this is kept constant at a value determined by the direct titration
- $K_d$  is dissociation constant of host-guest (calixarene-peptide) complex. This is an iterable variable in and was optimized

2. *Direct titration:* The general equation describing direct titration is a quadratic equation of the form<sup>[1]</sup>:  $[HG]^2 + b[HG] - c = 0$ , where  $b = K_{ind} + [H]^T + [I]^T$  and  $c = ([H]^T[I]^T)$ .  $[I]^T$  is kept constant while  $[H]^T$  is varied. The general solution to quadratic equation can be used to determine  $[HG]$  at any given  $[H]^T$ .  $[HG]$  is used to fit the data using the equation shown below.<sup>[1]</sup> Both  $\Delta F_{max}$  and  $K_{ind}$  are optimized.

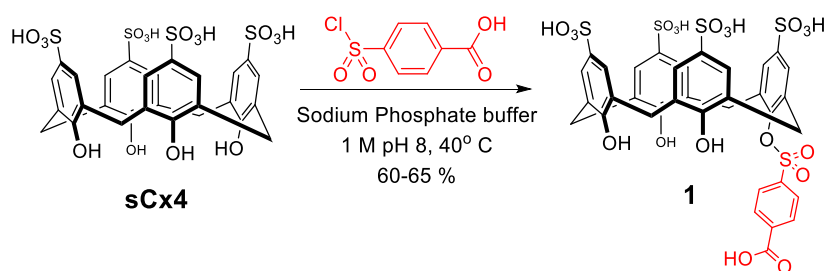
$$dF_{obs} = \frac{\Delta F_{max} [HG]}{[I]^T}$$

3. *Competitive titration:* The general equation describing competitive titration is a cubic equation of the form<sup>[2]</sup>:  $[H]^3 + a[H]^2 + b[H] + c = 0$ , where  $a = (K_d + K_{ind} + [H]^T + [I]^T + [Pep]^T)$ ,  $b = (K_d * K_{ind}) + (K_{ind} * ([Pep]^T - [H]^T)) + (K_d * ([I]^T - [H]^T))$  and  $c = -(K_d * K_{ind} * [H]^T)$ .  $[H]^T$  and  $[I]^T$  are kept constant while  $[Pep]^T$  is varied. An exact solution to this cubic equation has been published in literature<sup>[2]</sup> and can be used to determine  $[H]$  at any given  $[Pep]^T$ .  $[H]$  is used to fit the data using the equation shown below. Both  $\Delta F_{max}$  and  $K_d$  are optimized.

$$dF_{obs} = \Delta F_{max} \left( 1 - \frac{[H]}{K_{ind} + [H]} \right)$$

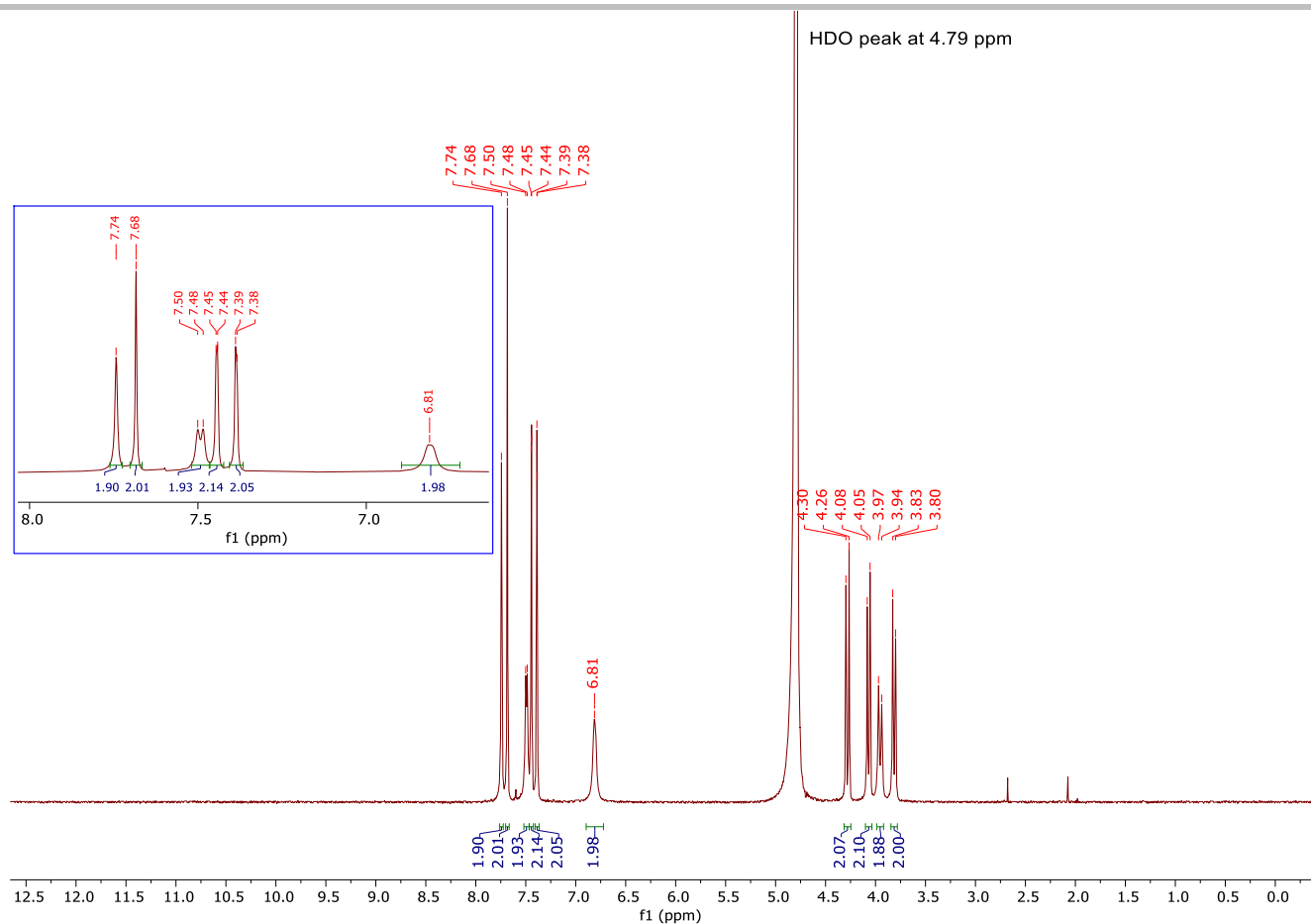
**S1.e. Isothermal calorimetry (ITC).** ITC was performed using Microcal VP-ITC (GE Healthcare). Titrations were performed in 10 mM sodium phosphate buffer at pH 7.5 at 303 K. Concentration of peptides were calculated using  $A_{280}$  while that of **1** was calculated using qNMR. All solutions were thoroughly degassed before experiment. In all cases, solution of **1** was loaded in the ITC cell while peptide solution was loaded into the syringe. The full titration involved 50 injections of 5  $\mu$ L (except first injection which was 2  $\mu$ L) at constant interval of 300 seconds. Binding curves were produced using supplied Origin software and fit using a 1-sites binding model. The first point of titration was discarded. Each titration was performed in duplicate and the final data reported is average of two runs. ITC data and curve fit are shown in Section SI5.

## S2. Synthesis and characterization of compound 1

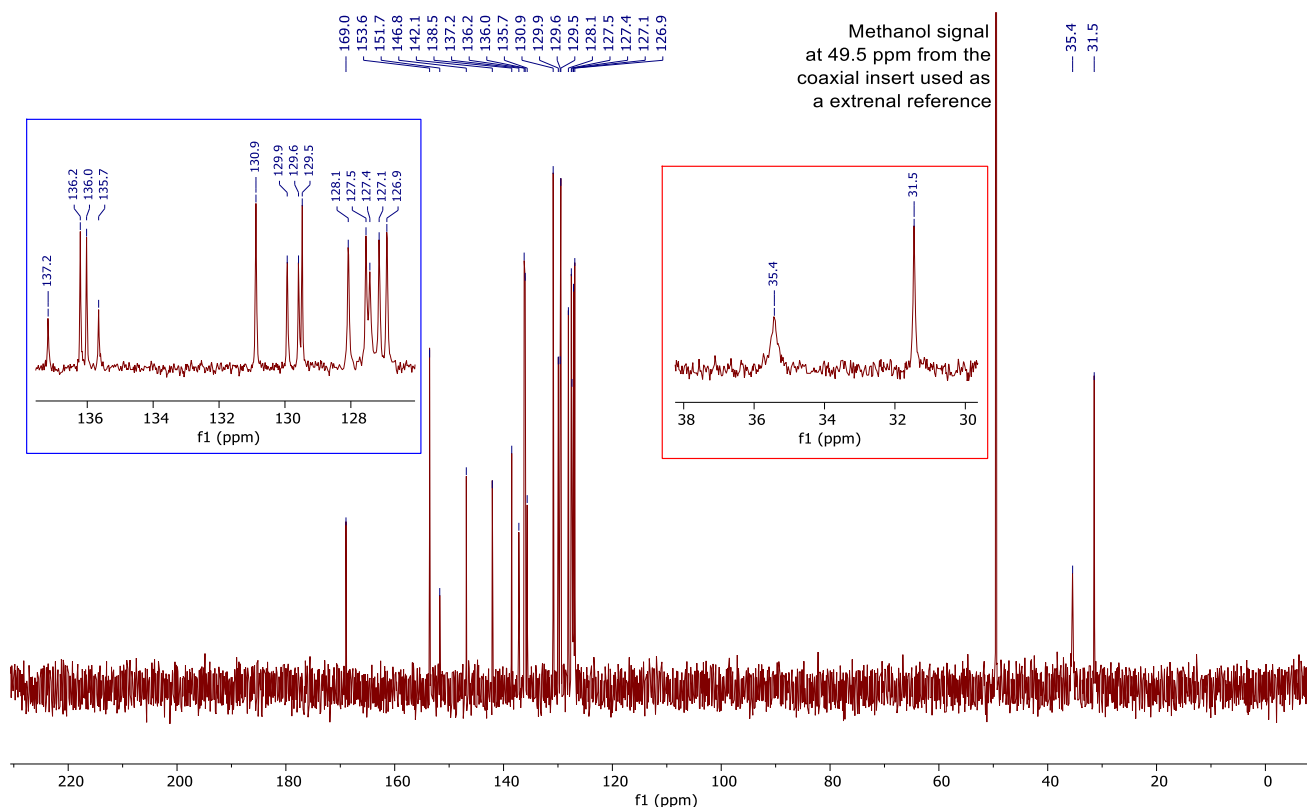


**Scheme S1** One step lower rim functionalization of commercially available calix[4]arene (**sCx4**) to compound **1**.

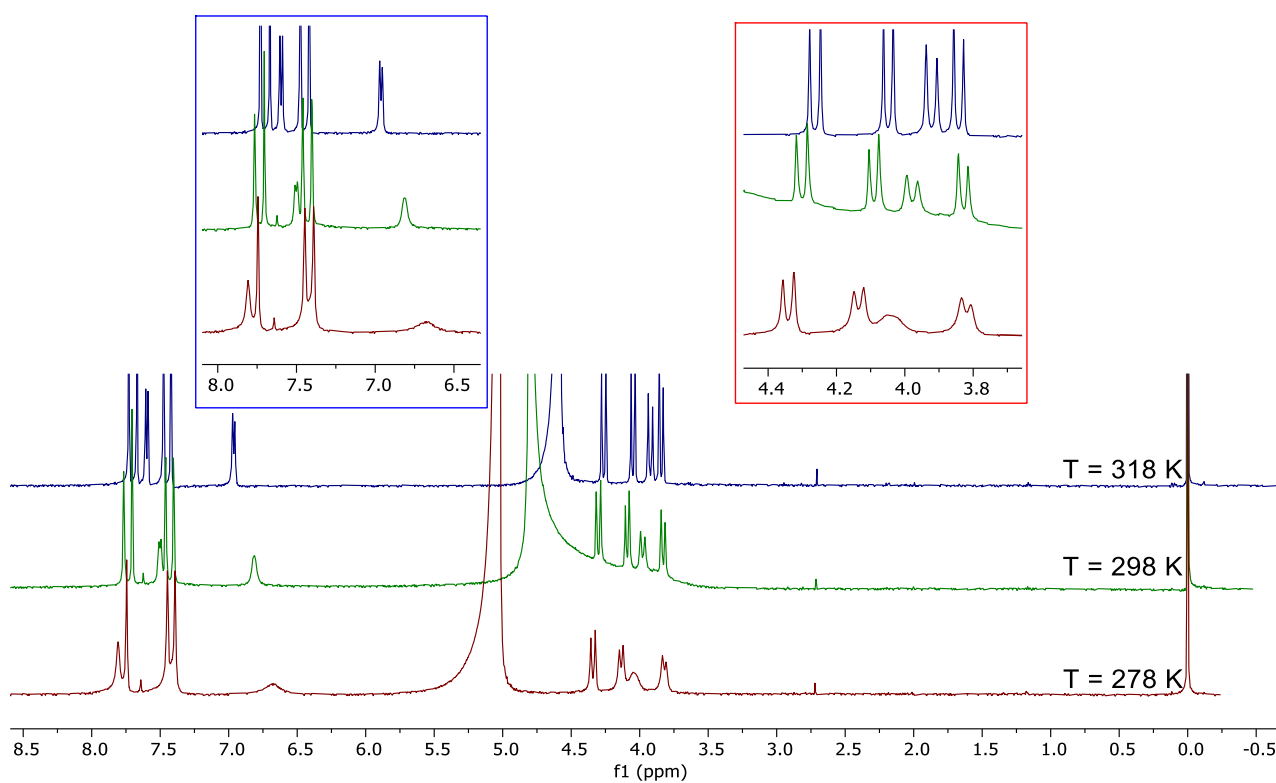
**sCx4** (200 mg, 0.27 mmol, 1 eq) was dissolved in sodium phosphate buffer (1 M, pH 8). To this solution, 36 mg of (4-chlorosulfonyl)benzoic acid (0.162 mmol, 0.6 eq) was added and the mixture was stirred for 1 hour at 40°C (This temperature was maintained to prevent the phosphate salts from crashing out). At this point, UPLC showed 25% reaction progress. After this, small portions of (4-chlorosulfonyl)benzoic acid (12-18 mg, 0.05-0.081 mmol, 0.2-0.3 eq) were added at 2-hour intervals and reaction monitored using UPLC-MS until <5% starting material was visible. During this time, the pH of the reaction mixture was regularly checked and maintained between 7.5 and 8.0. The total amount of (4-chlorosulfonyl)benzoic acid added was 114 mg (0.517 mmol, 1.9 eq) added in 6 portions. After completion, a few drops of saturated  $\text{FeCl}_3$  solution were added to precipitate most of the phosphate salt. The mixture was centrifuged, and the supernatant collected. The precipitate was washed twice with water. The combined supernatant and washes were then subjected to preparative RP-HPLC using a 30-minute gradient of acetonitrile in water (starting from 10% acetonitrile to 90% acetonitrile). The compound was obtained as a white powder with yields of 60-65% (3 iterations). **Note** – As reported by us previously<sup>[3]</sup>, p-sulfonatocalixarenes tend to aggregate in aqueous solution and as such show concentrated dependent chemical shift and broadening. To account for this, concentration of **1** in NMR experiments was calculated using qNMR and is reported along with characterization data.  **$^1\text{H}$  NMR (500 MHz  $\text{D}_2\text{O}$ , concentration = 7.8 mM):**  $\delta$  (ppm) = 3.81 (d, 2H,  $J = 14.3$  Hz, Ar- $\text{CH}_2$ -Ar), 3.95 (d, 2H,  $J = 16.1$  Hz, Ar- $\text{CH}_2$ -Ar), 4.07 (d, 2H,  $J = 14.3$  Hz, Ar- $\text{CH}_2$ -Ar), 4.28 (d, 2H,  $J = 16.1$  Hz, Ar- $\text{CH}_2$ -Ar), 6.81 (br, 2H, Ar-H), 7.39 (d, 2H,  $J = 1.6$  Hz, Ar-H), 7.44 (d, 2H,  $J = 1.7$  Hz, Ar-H), 7.49 (br d,  $J = 7.8$  Hz, 2H, Ar-H), 7.68 (s, 2H, Ar-H), 7.74 (s, 2H, Ar-H);  **$^{13}\text{C}$  NMR (75 MHz  $\text{D}_2\text{O}$ , referenced externally using 5% methanol in  $\text{D}_2\text{O}$ , concentration = 40.7 mM):**  $\delta$  (ppm) = 31.5, 35.4 (br), 126.9, 127.1, 127.4, 127.5, 128.1, 129.5, 129.6, 129.9, 130.9, 135.7, 136.0, 136.2, 137.2, 138.5, 142.1, 146.8, 151.7, 153.6, 169.0; **IR (ATR) ( $\text{cm}^{-1}$ )** = 3332 (br), 1707 (br m), 1593 (w), 1448 (w), 1376 (w), 1151 (m), 1111 (s), 1035 (s), 888 (m), 829 (w), 621 (s), 556 (s)  $\text{cm}^{-1}$ ; **HRMS (ESI):** Found 462.97945 (calculated for  $\text{C}_{35}\text{H}_{26}\text{O}_{20}\text{S}_5^{2-}$  [ $\text{M}-2\text{H}$ ] $^{2-}$ ) = 462.9816)



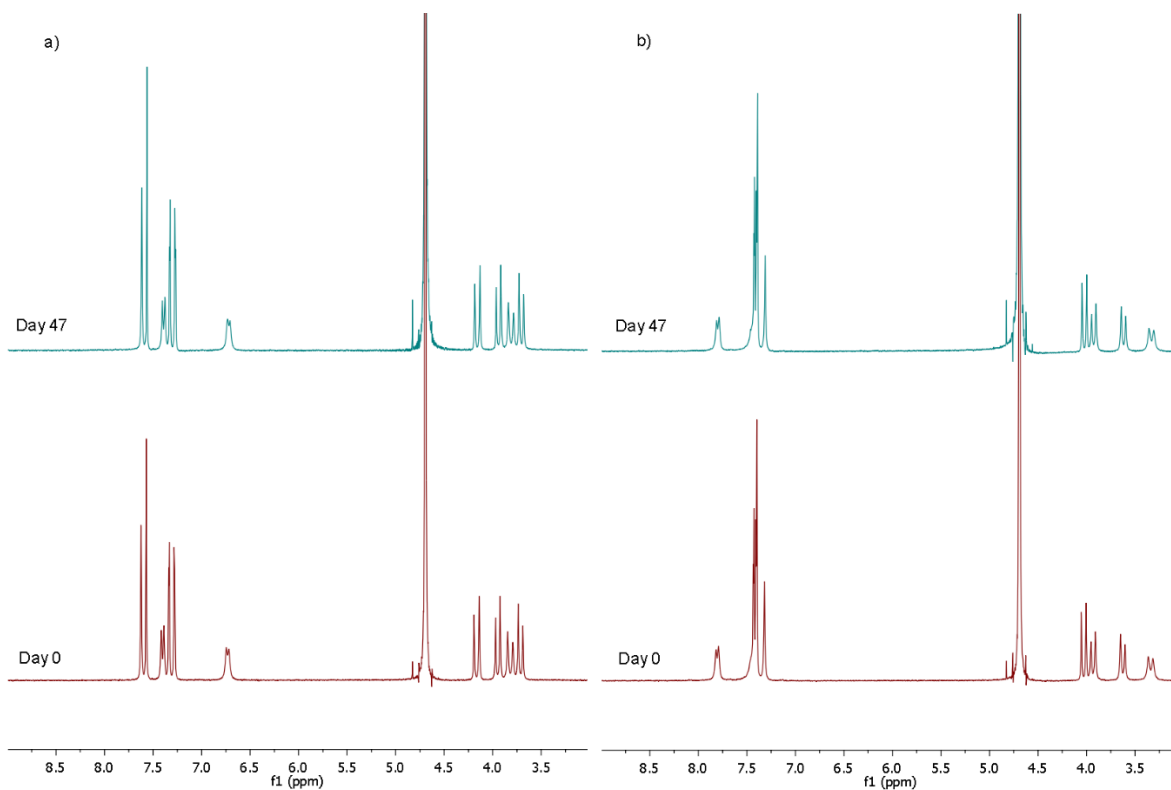
**Figure S2:**  $^1\text{H}$  NMR trace of **1** in  $\text{D}_2\text{O}$  collected at 500 MHz frequency. The sample concentration of **1** was 7.8 mM. Blue inset shows a blown up trace of aromatic region.



**Figure S3:**  $^{13}\text{C}$  NMR trace of **1** collected in  $\text{D}_2\text{O}$  at 75 MHz frequency. The sample concentration of **1** was 40.7 mM. Blue inset shows the blown up version of aromatic region (126-138 ppm) of this NMR while red inset shows the methylene region (30-38 ppm). Difference in lineshape of different methylenes is evident.

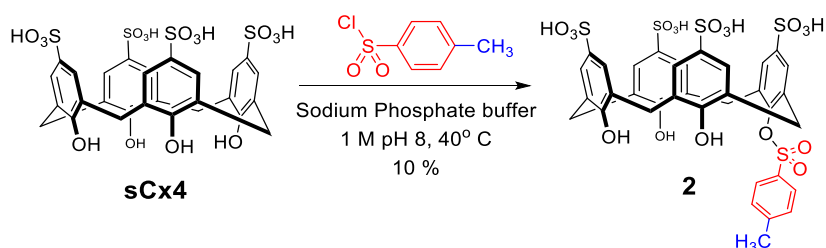


**Figure S4:**  $^1\text{H}$  NMR trace of **1** (3.2 mM) in  $\text{D}_2\text{O}$  at three different temperatures. The samples also contained 1.2 mM of Trimethylsilylpropionic acid (TSP) as an internal standard, the chemical shift of which was set to 0 ppm to account of temperature dependent shift of the HDO peak. Insets show blown up trace of aromatic (blue box) and aliphatic (red box) regions.



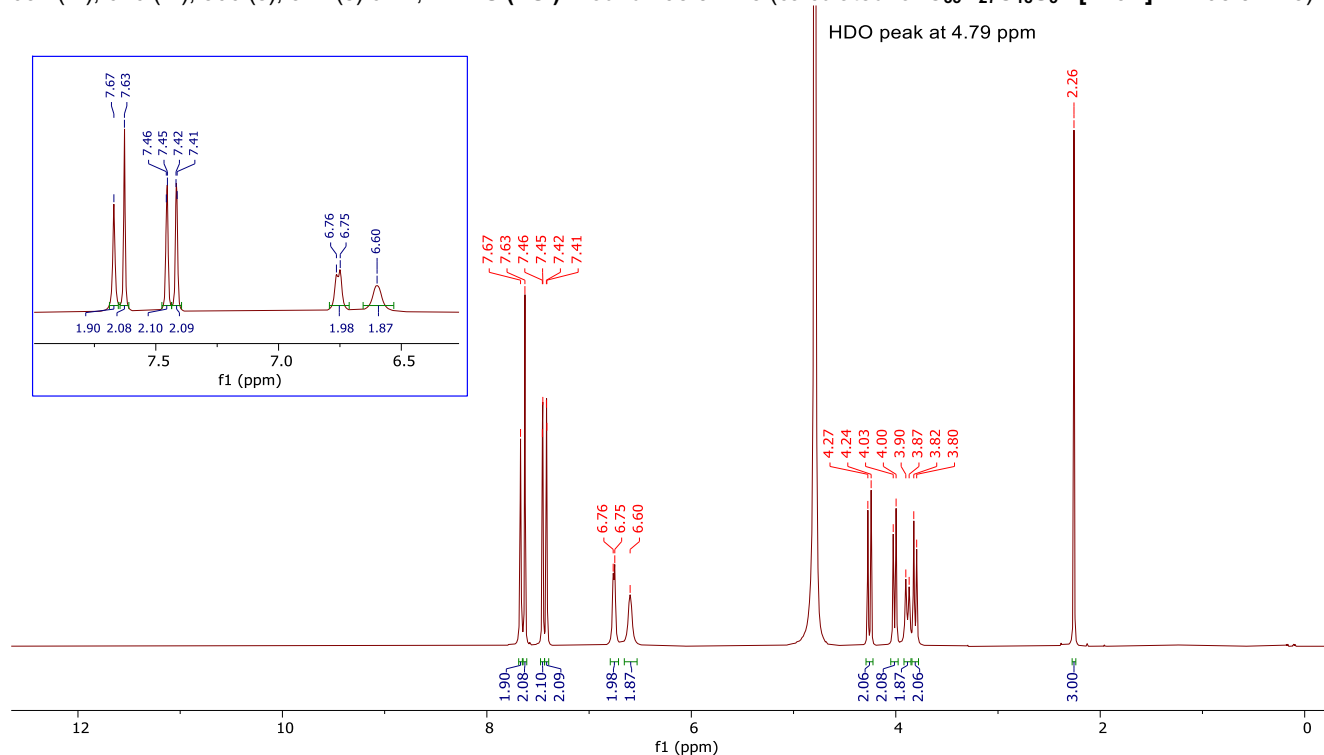
**Figure S5:** Exceptional stability of **1** in water as evident by  $^1\text{H}$  NMR spectra of a 10 mM sample taken in a)  $\text{D}_2\text{O}$  and b) 50 mM sodium phosphate buffer, pH 7.5 over the span of 7 weeks.

### S3. Synthesis and characterization of control compound 2



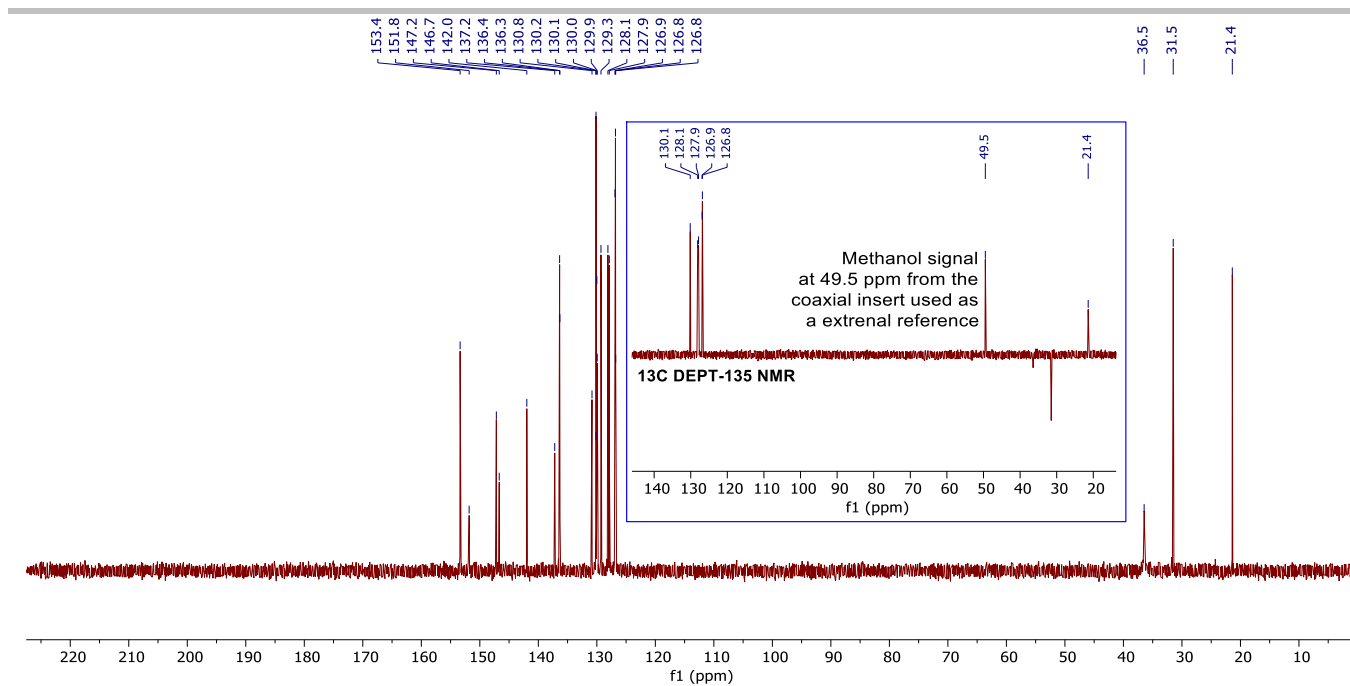
**Scheme S2.** One step lower rim functionalization of commercially available calix[4]arene (**sCx4**) to compound **2**.

Compound **2** was synthesized in the same manner as compound **1** on a 200 mg scale by only changing the sulfonyl chloride used. **Note:** The yield reported here is unoptimized (single iteration). Significant amount of disubstituted product (50 %) was found to have crashed out of the reaction mixture. **<sup>1</sup>H NMR (500 MHz D<sub>2</sub>O, concentration – 32.4 mM):**  $\delta$  (ppm) = 2.26 (s, 3H, Ar-CH<sub>3</sub>), 3.81 (d, 2H, J = 14.1 Hz, Ar-CH<sub>2</sub>-Ar), 3.88 (d, 2H, J = 16.1 Hz, Ar-CH<sub>2</sub>-Ar), 4.01 (d, 2H, J = 14.1 Hz, Ar-CH<sub>2</sub>-Ar), 4.26 (d, 2H, J = 16.1 Hz, Ar-CH<sub>2</sub>-Ar), 6.60 (br s, 2H, Ar-H), 6.76 (br d, 2H, Ar-H), 7.42 (s, 2H, Ar-H), 7.45 (s, 2H, Ar-H), 7.63 (s, 2H, Ar-H), 7.67 (s, 2H, Ar-H); **<sup>13</sup>C NMR (125 MHz D<sub>2</sub>O, referenced externally using 10% methanol in D<sub>2</sub>O, concentration – 32.4 mM):**  $\delta$  (ppm) = 21.4, 31.5, 36.5 (br), 126.8, 126.8 (2 overlapping signals, see figure SI3.3), 126.9, 127.9, 128.1, 129.3, 129.9, 130.0, 130.1, 130.2, 130.8, 136.3, 136.4, 136.2, 137.2, 142.0, 146.7, 147.2, 151.8, 153.4; **IR (ATR) (cm<sup>-1</sup>)** = 3371 (br), 1697 (br), 1595 (w), 1474 (w), 1377 (w), 1194 (m), 1156 (m), 1115 (s), 1039 (s), 886 (w), 788 (w), 657 (m), 626 (m), 566 (s), 547 (s) cm<sup>-1</sup>; **HRMS (ESI):** Found 298.32725 (calculated for C<sub>35</sub>H<sub>27</sub>O<sub>18</sub>S<sub>5</sub><sup>3-</sup> [M-3H]<sup>3-</sup> = 298.32725).

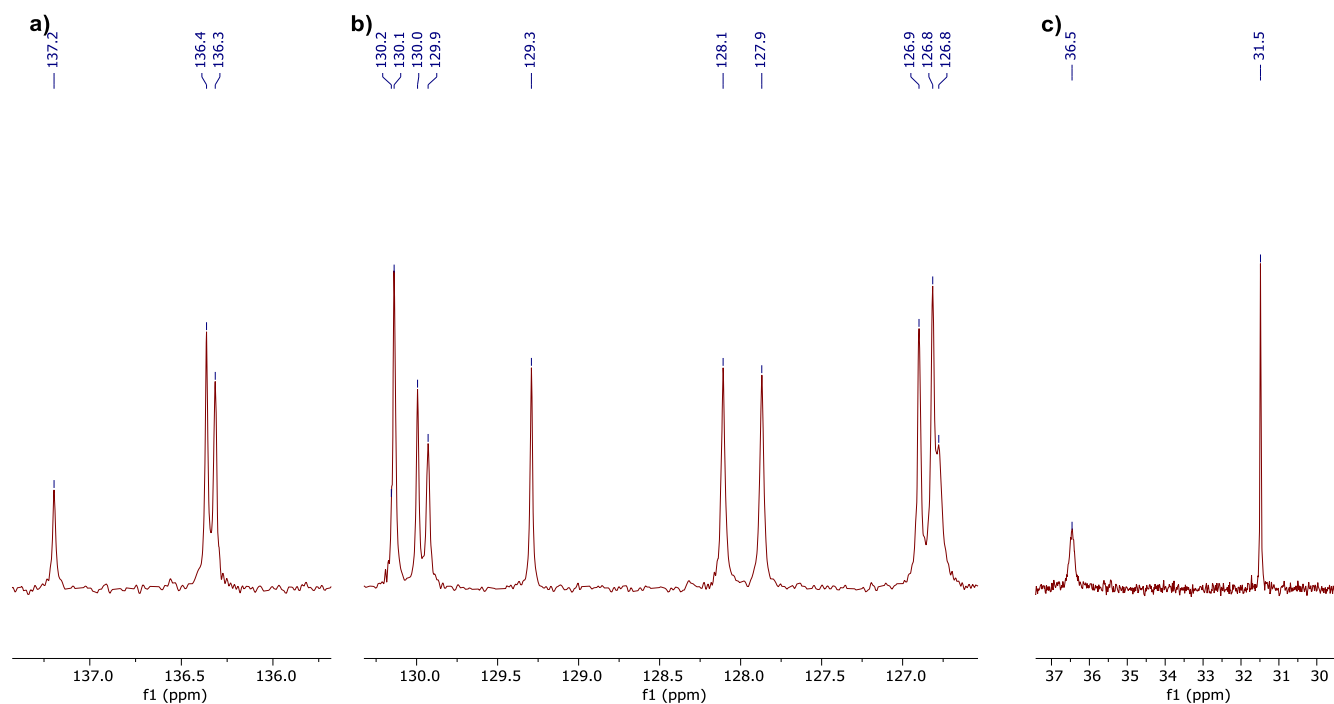


**Figure S6:** <sup>1</sup>H NMR trace of **2** in D<sub>2</sub>O collected at 500 MHz frequency. The sample concentration of **2** was 32.4 mM. Blue inset shows a blown up trace of aromatic region.



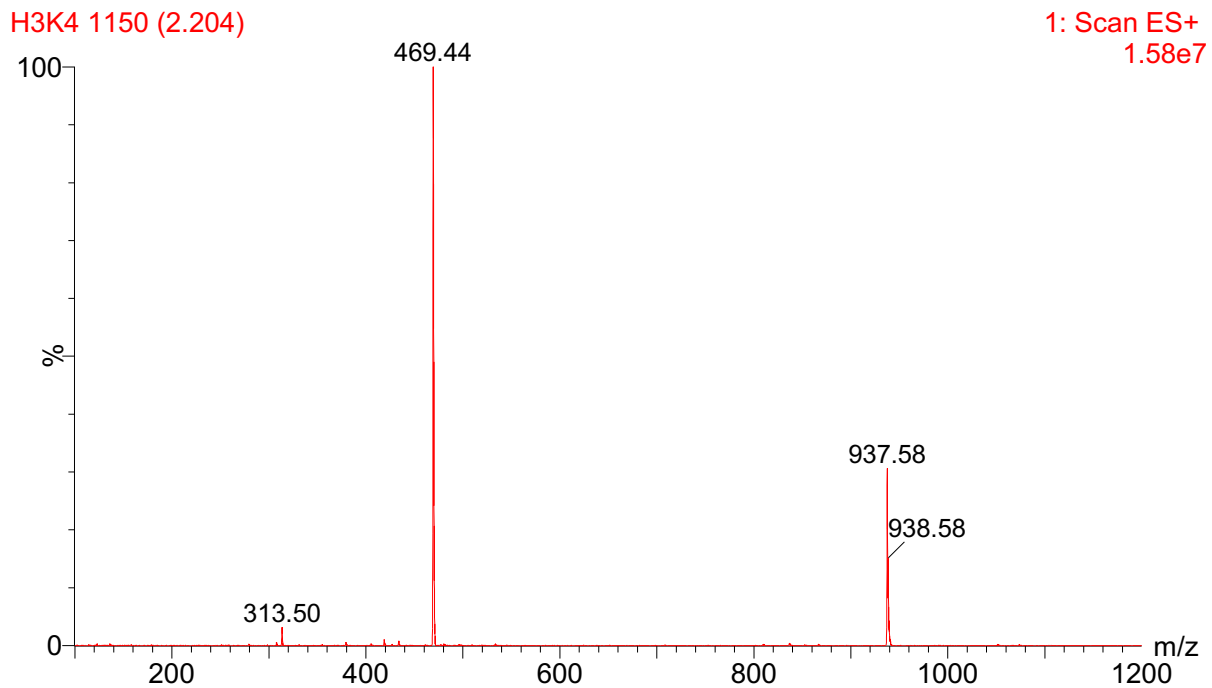
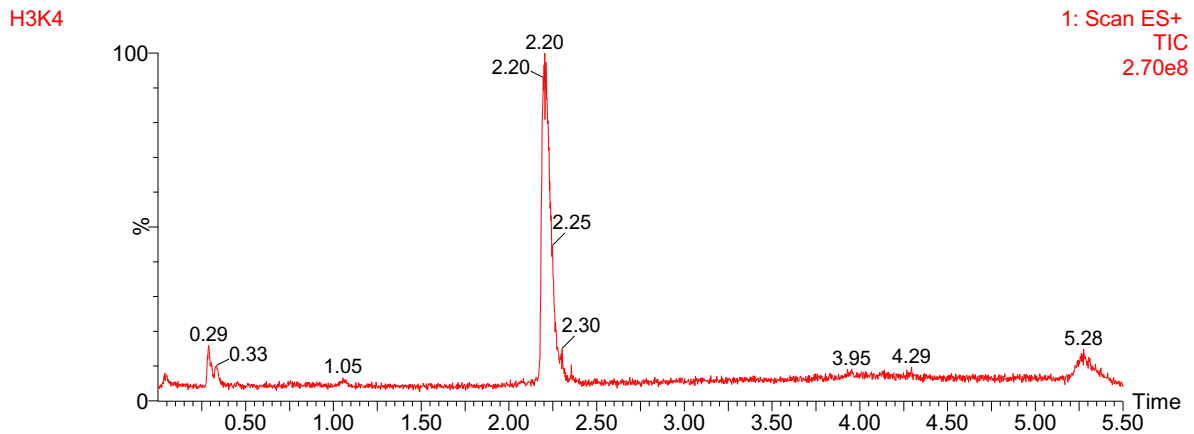
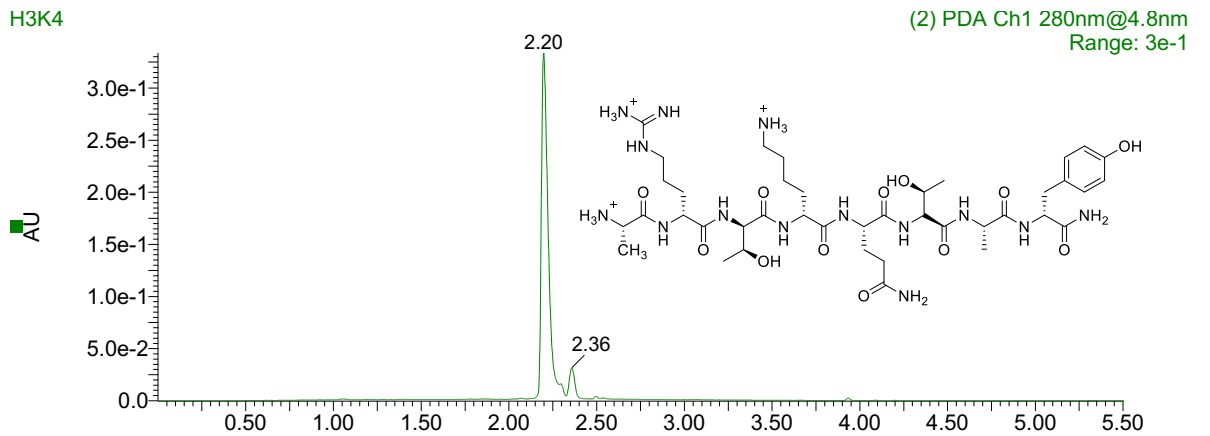


**Figure S7:**  $^{13}\text{C}$  NMR trace of **2** collected in  $\text{D}_2\text{O}$  at 125 MHz frequency. The sample concentration of **2** was 40.7 mM. Inset shows a 75 MHz  $^{13}\text{C}$  DEPT-135 NMR of the same sample with an external reference (10% MeOH in  $\text{D}_2\text{O}$ ) in a coaxial insert.  $^{13}\text{C}$  chemical shifts were referenced externally using the MeOH signal from DEPT spectra.



**Figure S8.** Blown up  $^{13}\text{C}$  NMR trace of **2** collected from a) 135 to 138 ppm, b) 126 to 130.5 ppm and c) 31 to 37 ppm (methylene region). The spectra was collected in  $\text{D}_2\text{O}$  at 125 MHz frequency. The sample concentration of **2** was 40.7 mM.

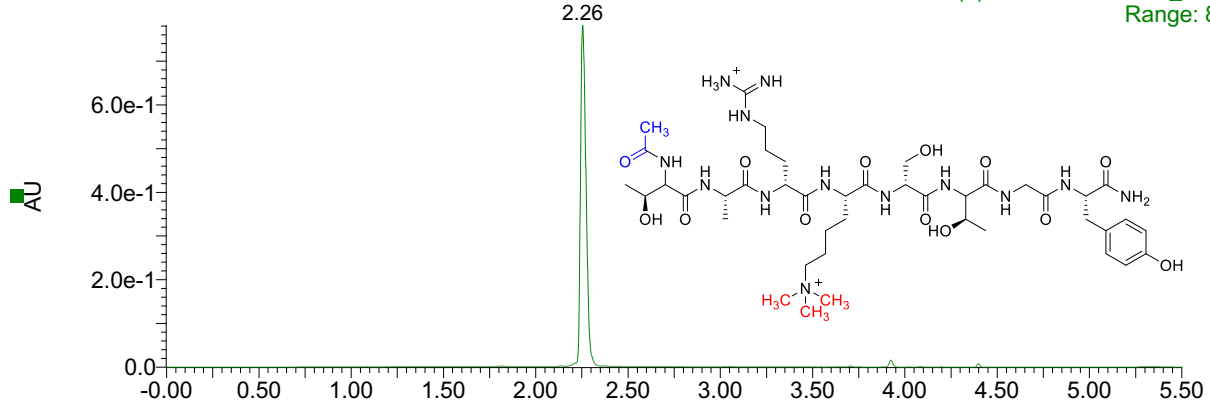
#### S4. Characterization and purity of peptides (UPLC-MS)



**Figure S9:** UPLC trace for peptide **H3K4**. Sequence (N to C terminus) – ARTKQTAY (N terminus = free amine, C terminus = amide)  
Prominent mass spectroscopy peaks – 937.58 ( $M+H^+$ ), 469.44 ( $M+2H^+$ ), 313.50 ( $M+3H^+$ )

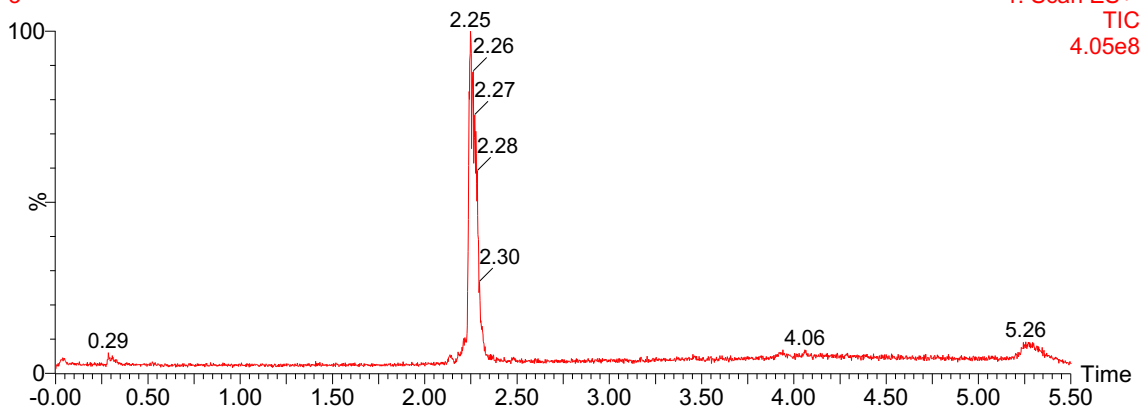
H3K9me3-3

(2) PDA Ch1 280nm@4.8nm  
Range: 8e-1



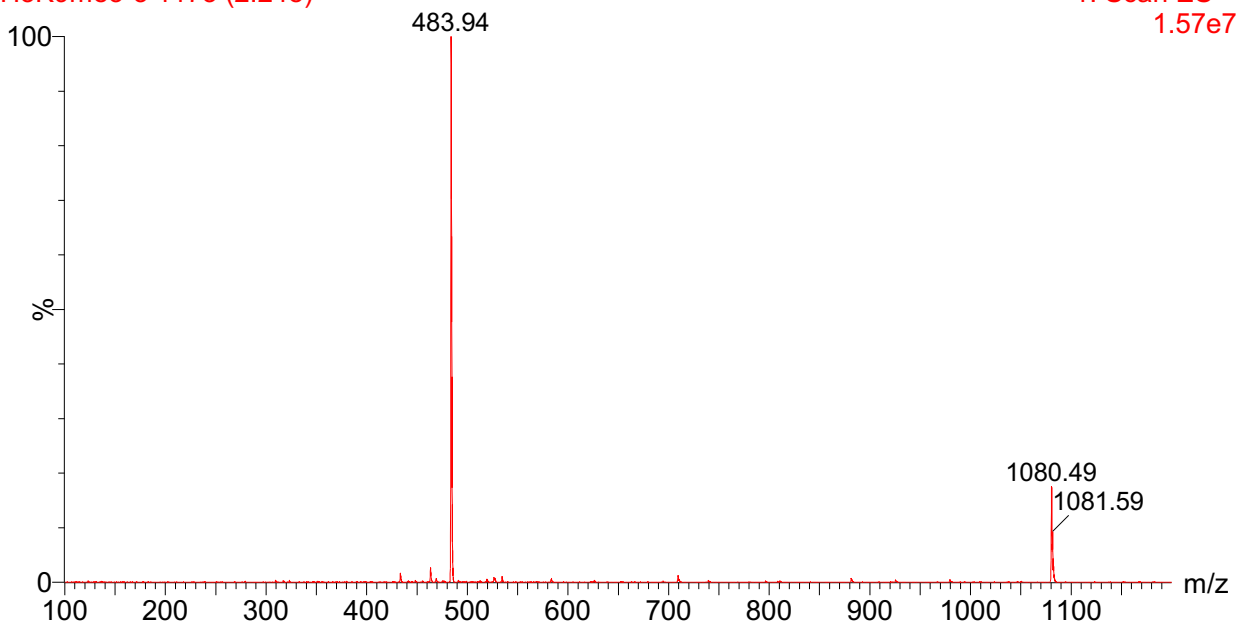
H3K9me3-3

1: Scan ES+  
TIC  
4.05e8



H3K9me3-3 1173 (2.248)

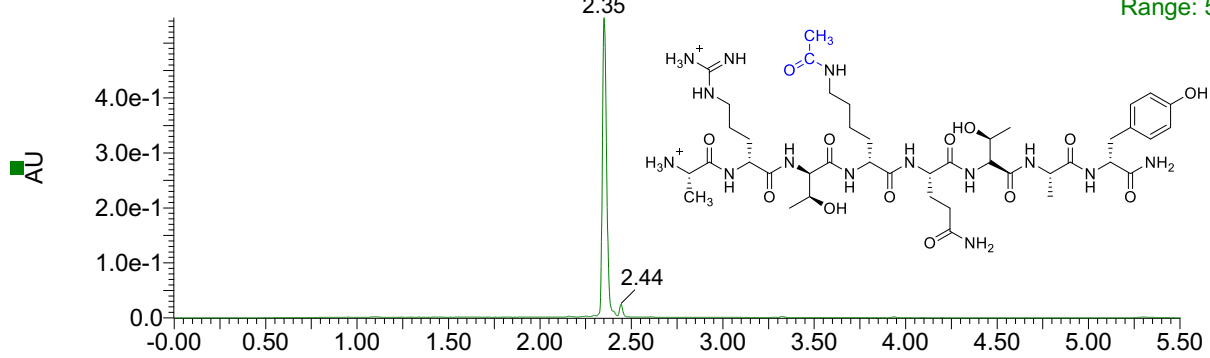
1: Scan ES+  
1.57e7



**Figure S10:** UPLC trace for peptide - **H3K9me3**. Sequence (N to C terminus) – TAR(Kme3)STGY (N terminus =acetyl, C terminus = amide)  
Prominent mass spectroscopy peaks – 1080.49 ( $M^+ + TFA$ ), 483.94 ( $M^+ + H^+$ )

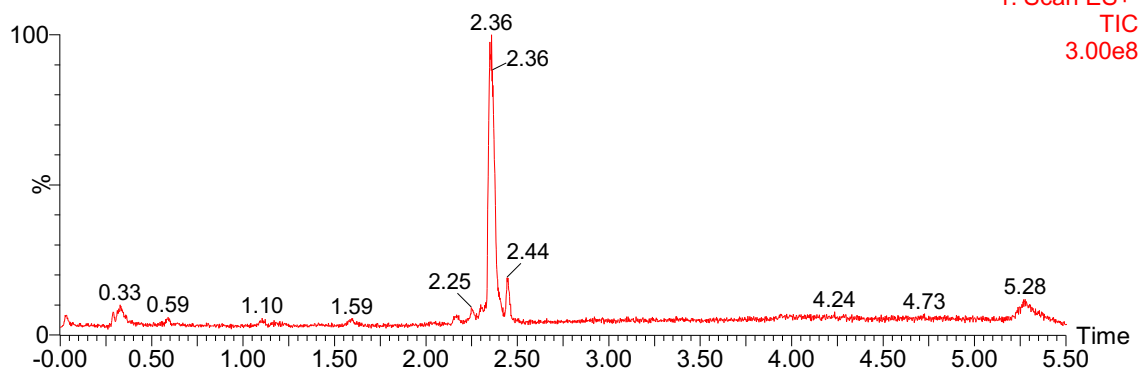
H3K4ac

(2) PDA Ch1 280nm@4.8nm  
Range: 5e-1



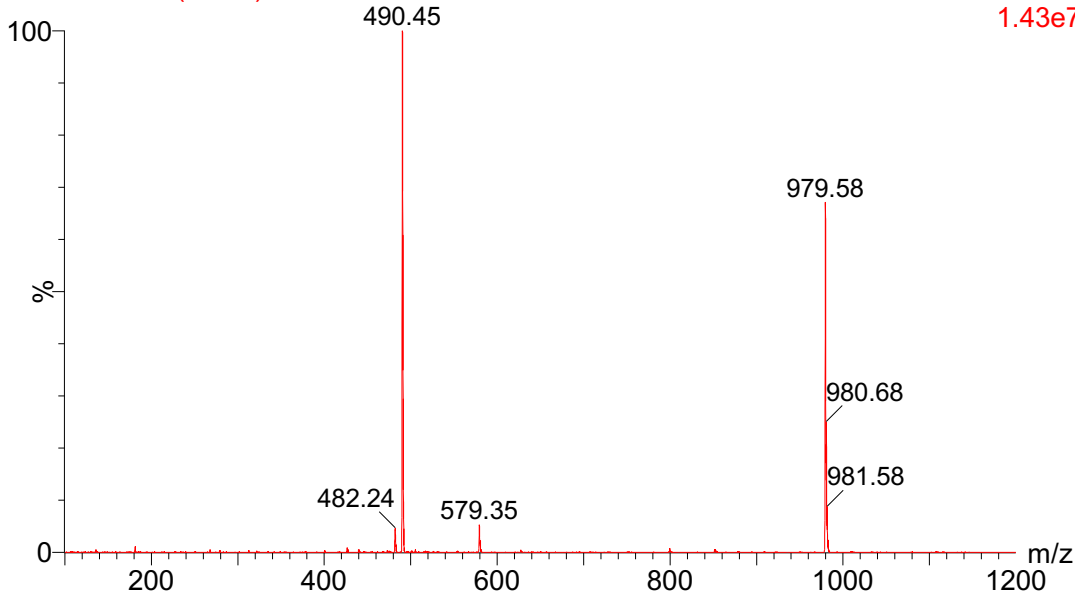
H3K4ac

1: Scan ES+  
TIC  
3.00e8



H3K4ac 1231 (2.359)

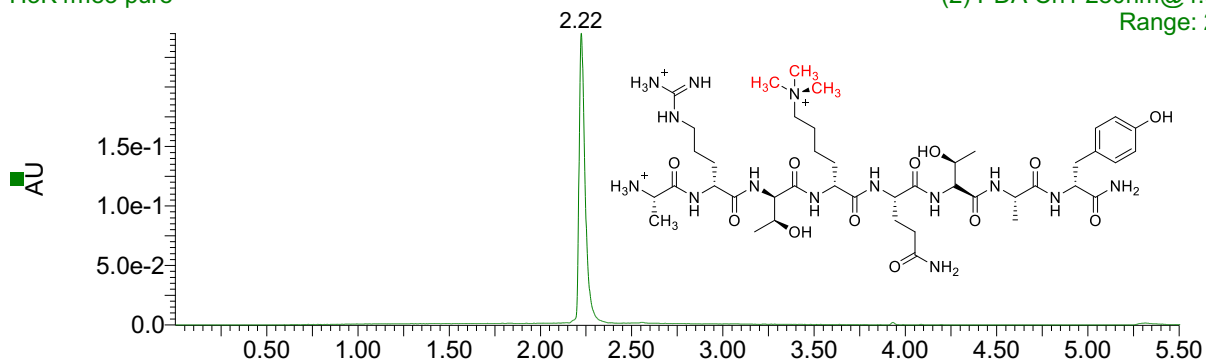
1: Scan ES+  
1.43e7



**Figure S11:** UPLC trace for peptide **H3K4ac**. Sequence (N to C terminus) – ART(Kac)QTAY (N terminus = free amine, C terminus = amide)  
Prominent mass spectroscopy peaks – 979.58 ( $M+H^+$ ), 490.45 ( $M+2H^+$ )

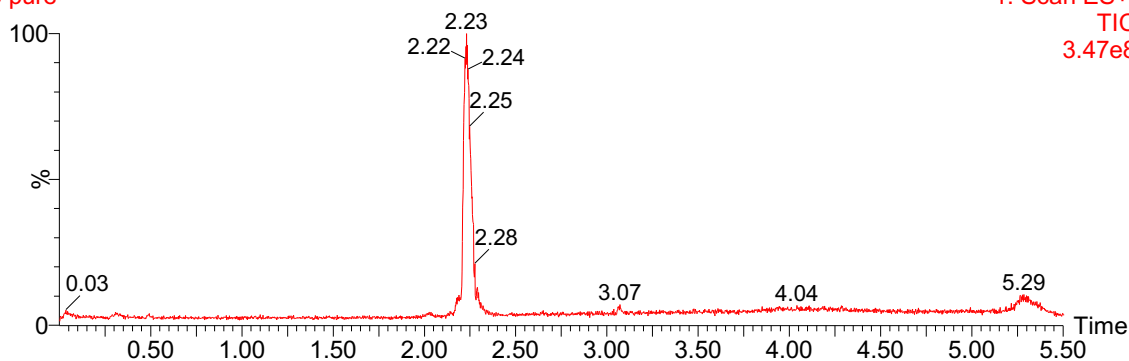
H3K4me3 pure

(2) PDA Ch1 280nm@4.8nm  
Range: 2e-1



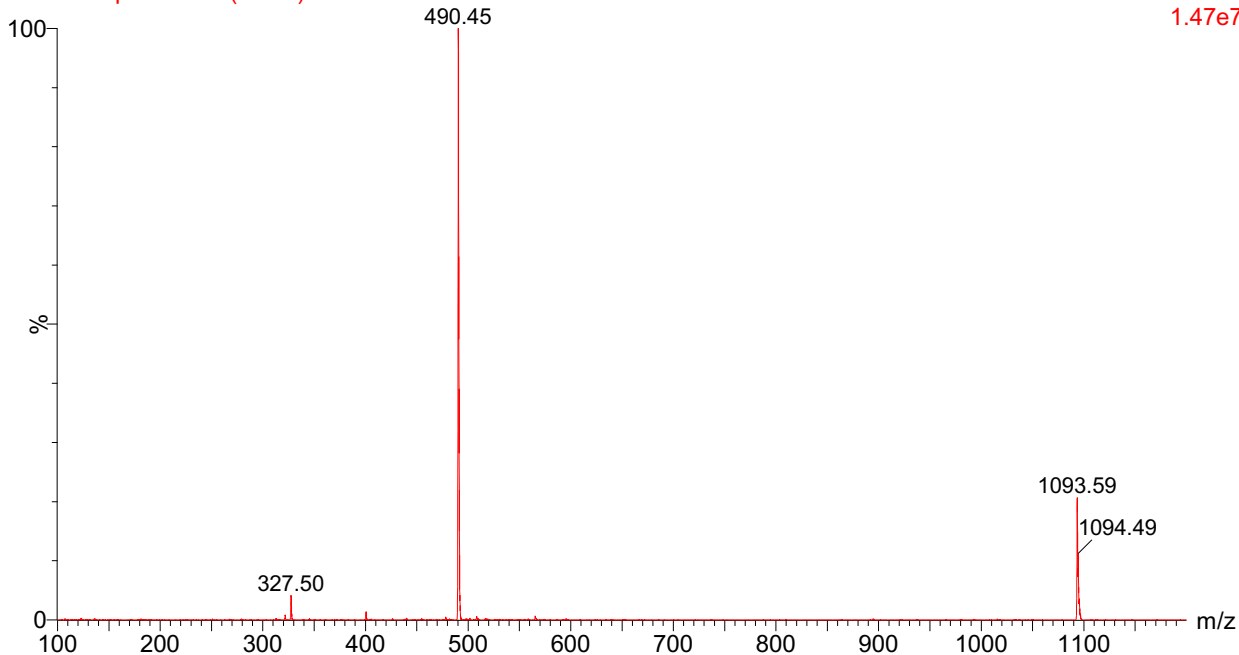
H3K4me3 pure

1: Scan ES+  
TIC  
3.47e8



H3K4me3 pure 1164 (2.231)

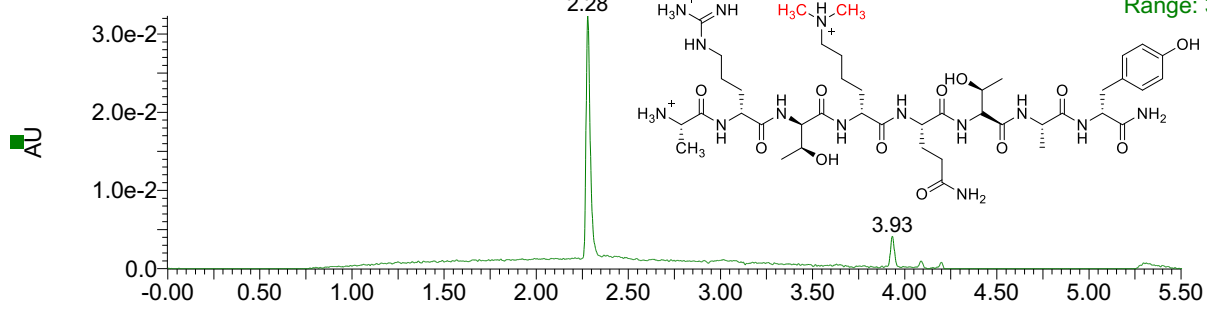
1: Scan ES+  
1.47e7



**Figure S12.** UPLC trace for peptide peptide - **H3K4me3**. Sequence (N to C terminus)– ART(Kme3)QTAY (N terminus = free amine, C terminus = amide). Prominent mass spectroscopy peaks – 1093.59 ( $M^+ + \text{TFA}$ ), 490.45 ( $M^+ + \text{H}^+$ ), 327.50 ( $M^+ + 2\text{H}^+$ )

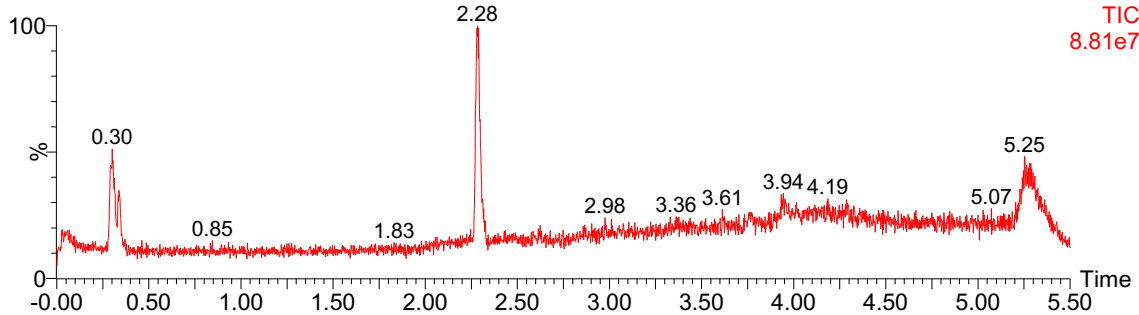
AS-2-114-H3K4me2-1

(2) PDA Ch1 280nm@4.8nm  
Range: 3e-2



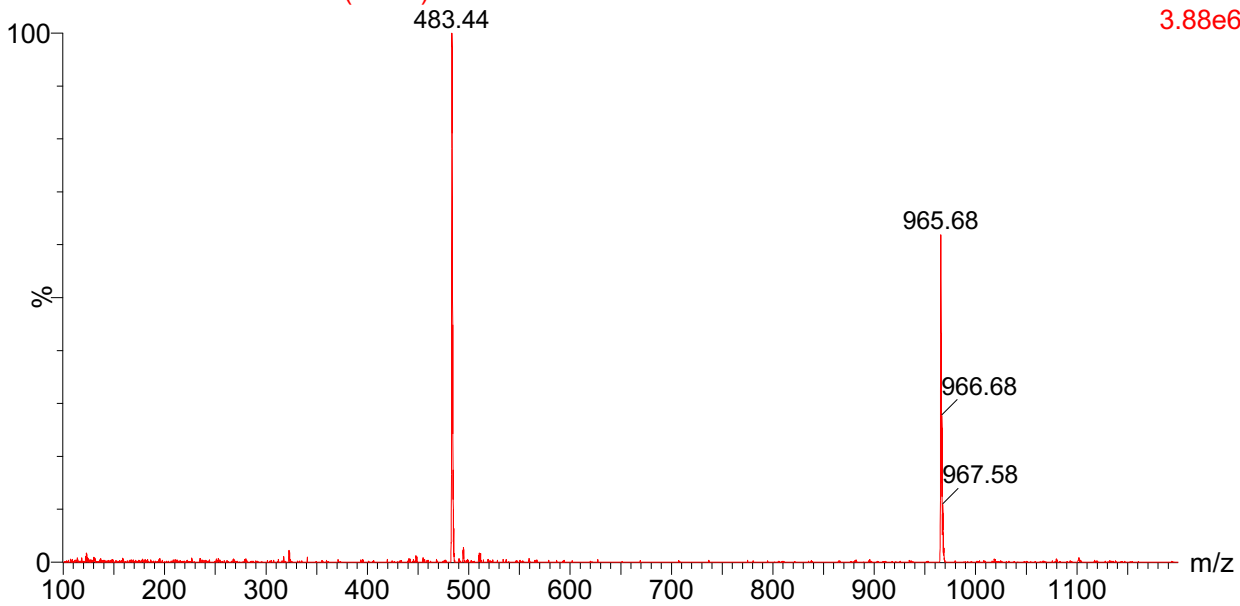
AS-2-114-H3K4me2-1

1: Scan ES+  
TIC  
8.81e7



AS-2-114-H3K4me2-1 1192 (2.285)

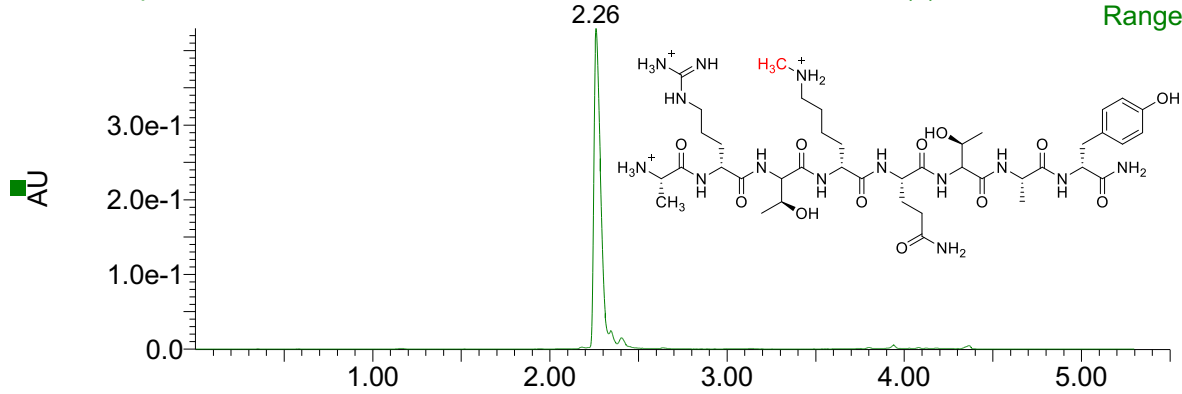
1: Scan ES+  
3.88e6



**Figure S13.** UPLC trace for peptide **H3K4me2** Sequence (N to C terminus) – ART(Kme2)QTAY (N terminus= free amine, C terminus= amide). Prominent mass spectroscopy peaks – 965.68 (M+H<sup>+</sup>), 483.44 (M+2H<sup>+</sup>)

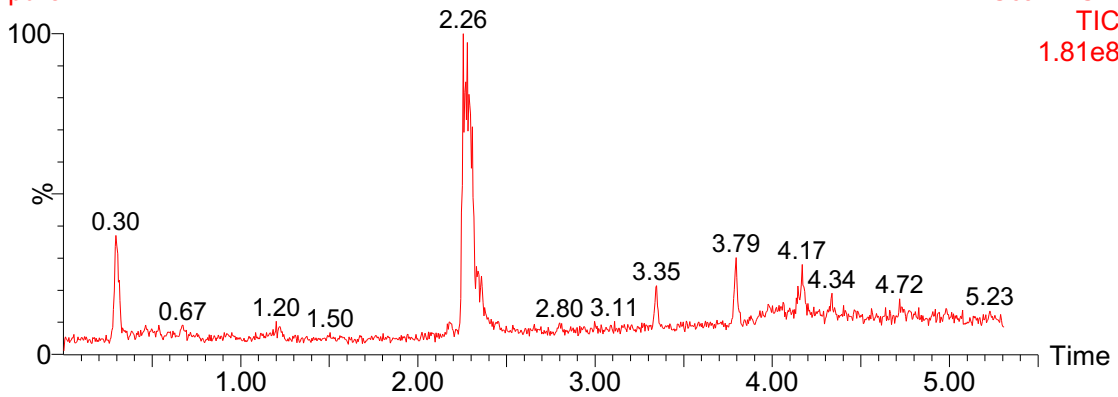
H3K4me1-pure

(2) PDA Ch1 280nm@4.8nm  
Range: 4e-1



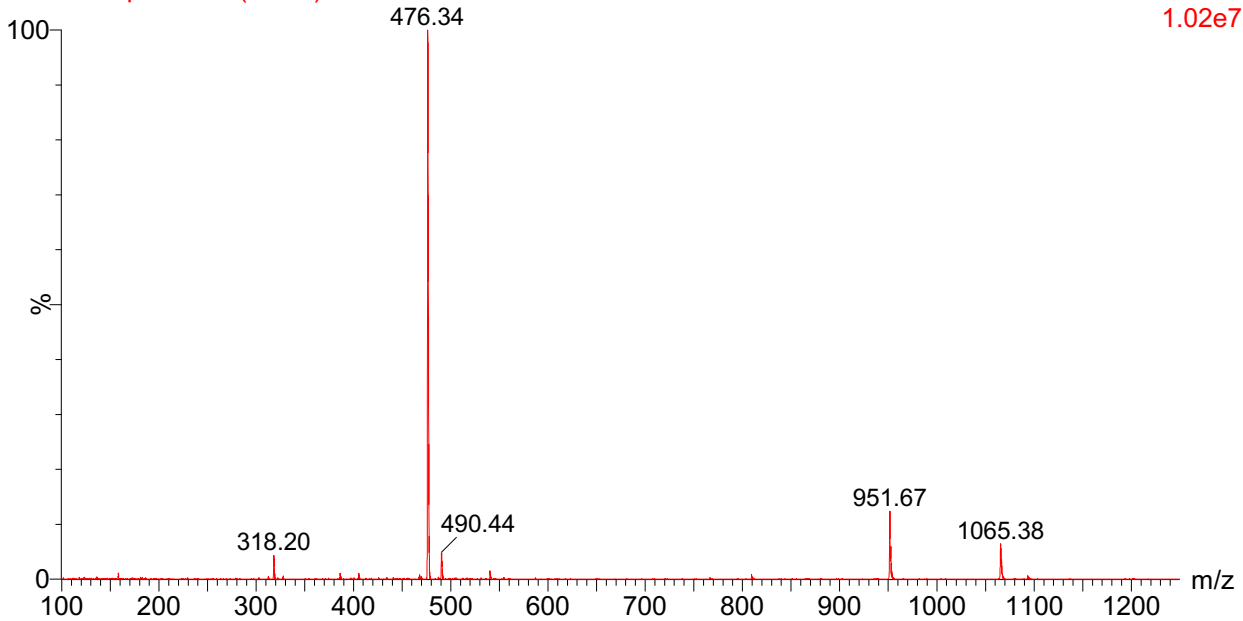
H3K4me1-pure

1: Scan ES+  
TIC  
1.81e8



H3K4me1-pure 484 (2.255)

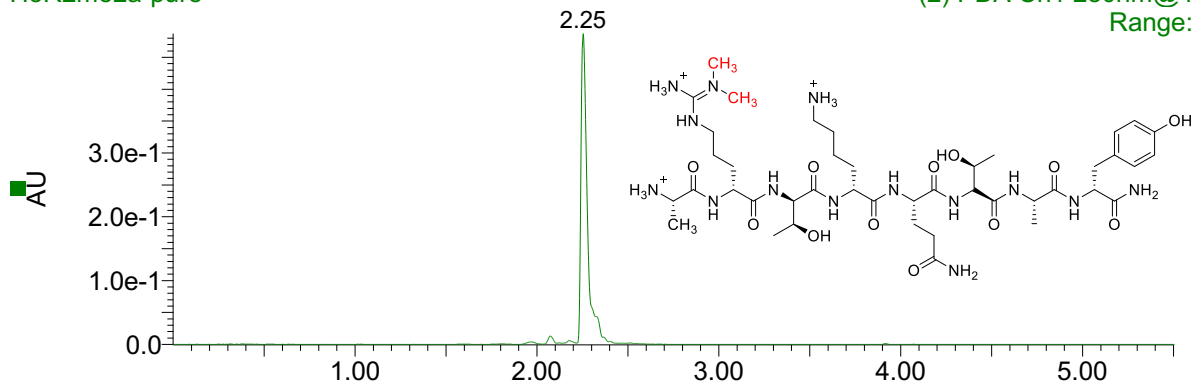
1: Scan ES+  
1.02e7



**Figure S14.** UPLC trace for peptide **H3K4me1**. Sequence (N to C terminus) – ART(Kme1)QTAY (N terminus= free amine, C terminus= amide). Prominent mass spectroscopy peaks – 951.67 (M+H<sup>+</sup>), 476.34 (M+2H<sup>+</sup>), 318.20 (M+3H<sup>+</sup>), 1065.38 (M+TFA+H<sup>+</sup>)

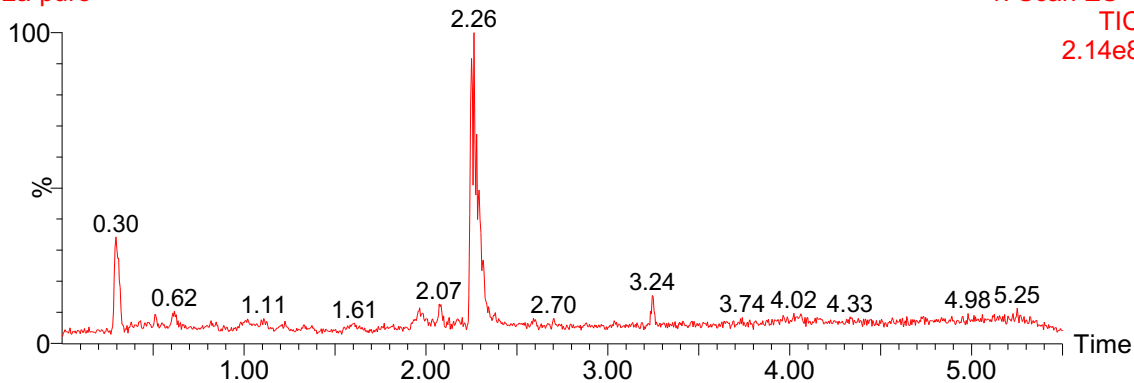
H3R2me2a-pure

(2) PDA Ch1 280nm@4.8nm  
Range: 5e-1



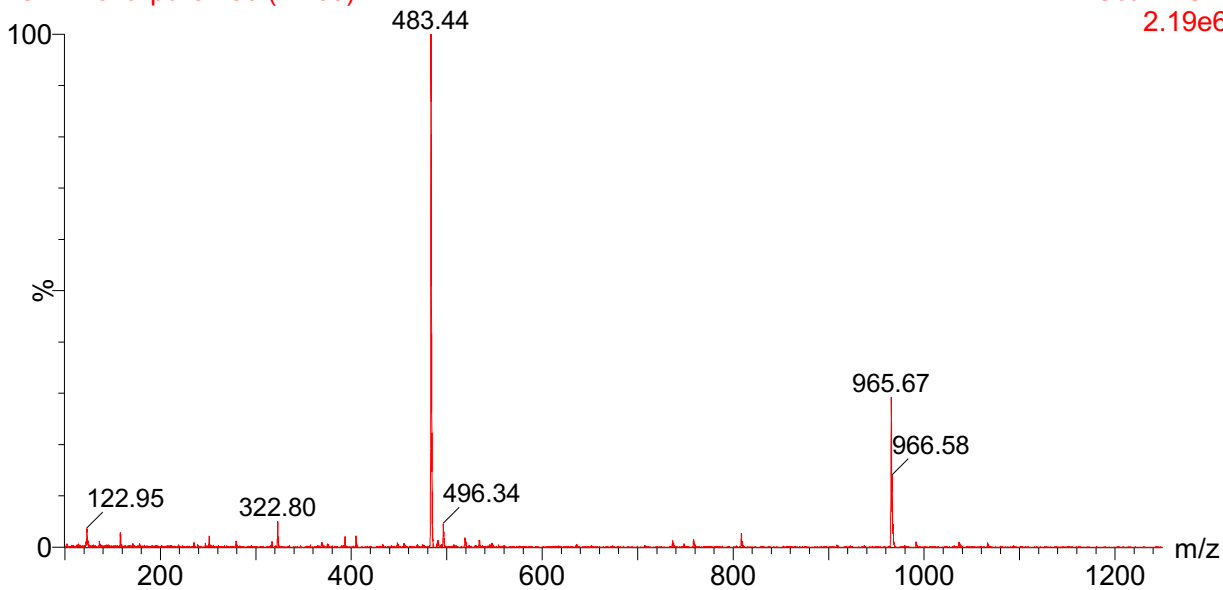
H3R2me2a-pure

1: Scan ES+  
TIC  
2.14e8



H3R2me2a-pure 486 (2.265)

1: Scan ES+  
2.19e6

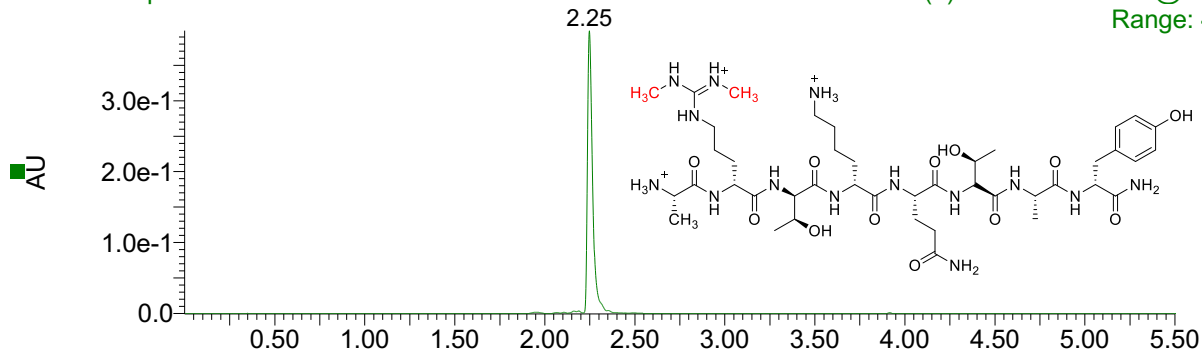


**Figure S15.** UPLC trace for peptide **H3R2me2a**; Sequence(N to C terminus)– A(Rme2a)TKQTAY (N terminus= free amine, C terminus= amide). Prominent mass spectroscopy peaks – 965.67 (M+H<sup>+</sup>), 483.44 (M+2H<sup>+</sup>)



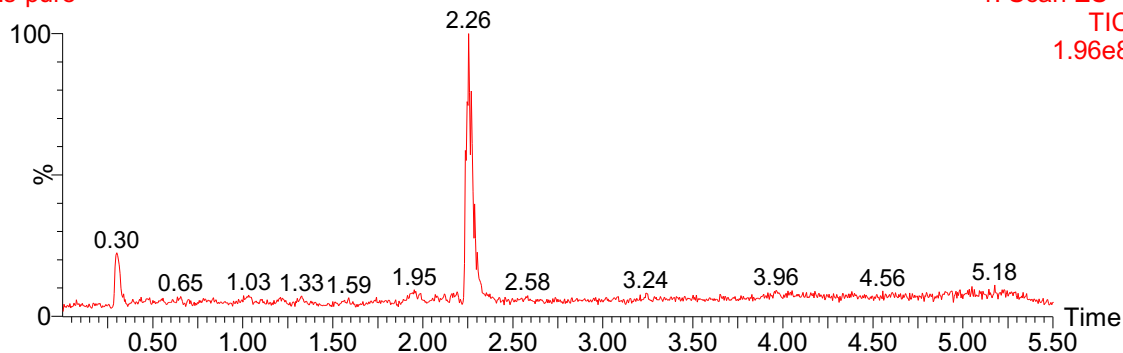
H3R2me2s-pure

(2) PDA Ch1 280nm@4.8nm  
Range: 4e-1



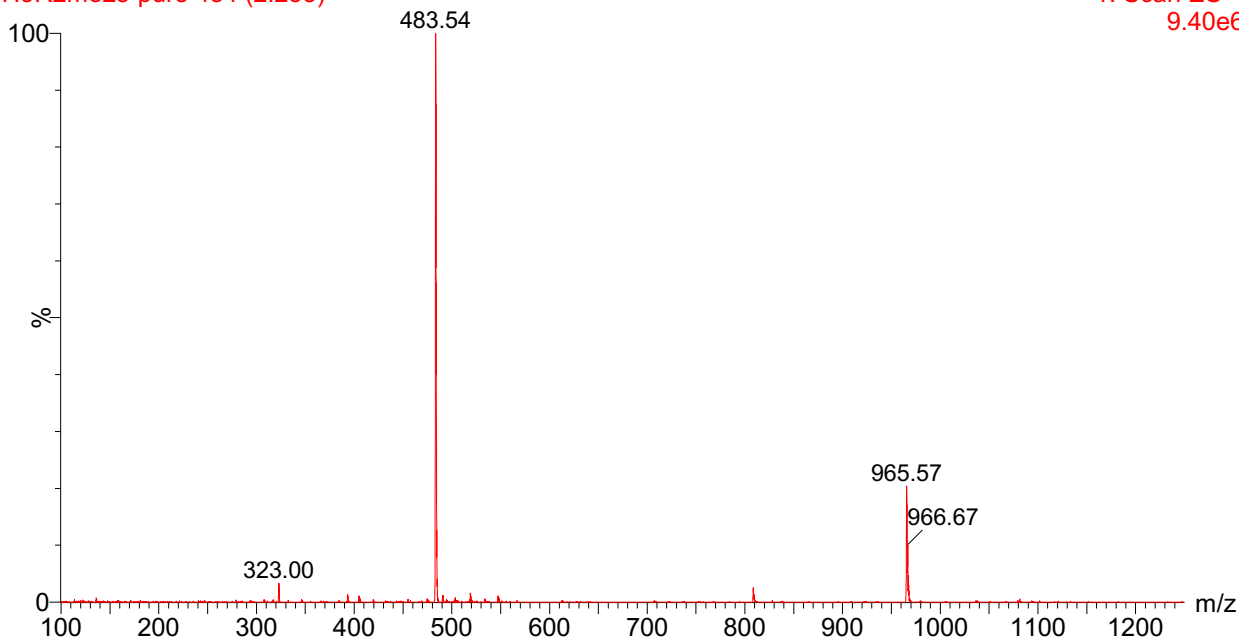
H3R2me2s-pure

1: Scan ES+  
TIC  
1.96e8



H3R2me2s-pure 484 (2.255)

1: Scan ES+  
9.40e6

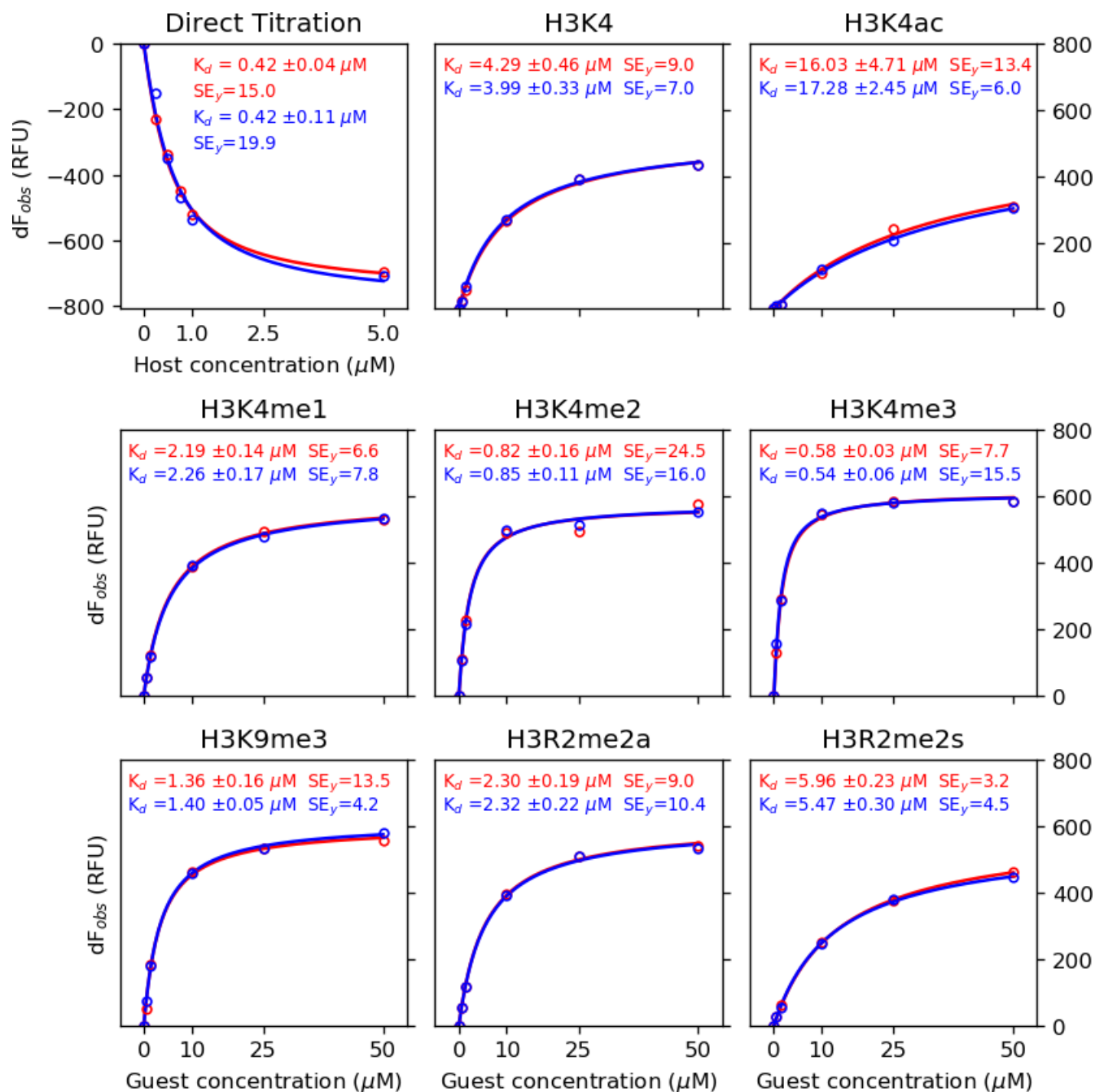


**Figure S16.** UPLC trace for peptide **H3R2me2s**; Sequence(N to C terminus)– A(Rme2a)TKQTAY (N terminus= free amine, C terminus= amide). Prominent mass spectroscopy peaks – 965.67 (M+H<sup>+</sup>), 483.44 (M+2H<sup>+</sup>)

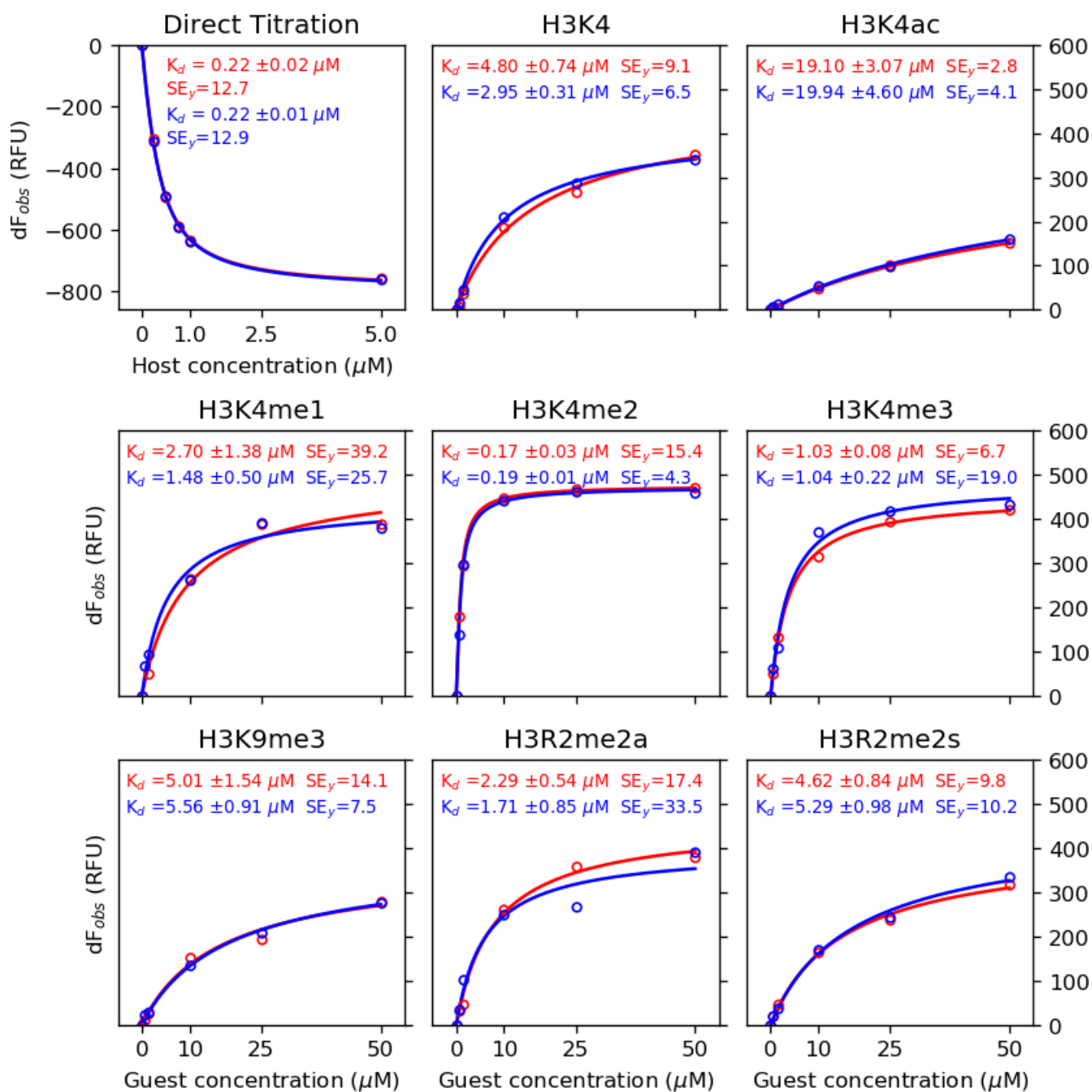
## Results and Discussion

### S5: Fitted curves from Indicator Displacement assay (IDA)

S5.a. Indicator Displacement Assays for **sCx4**. For detailed procedure, see S11.4.



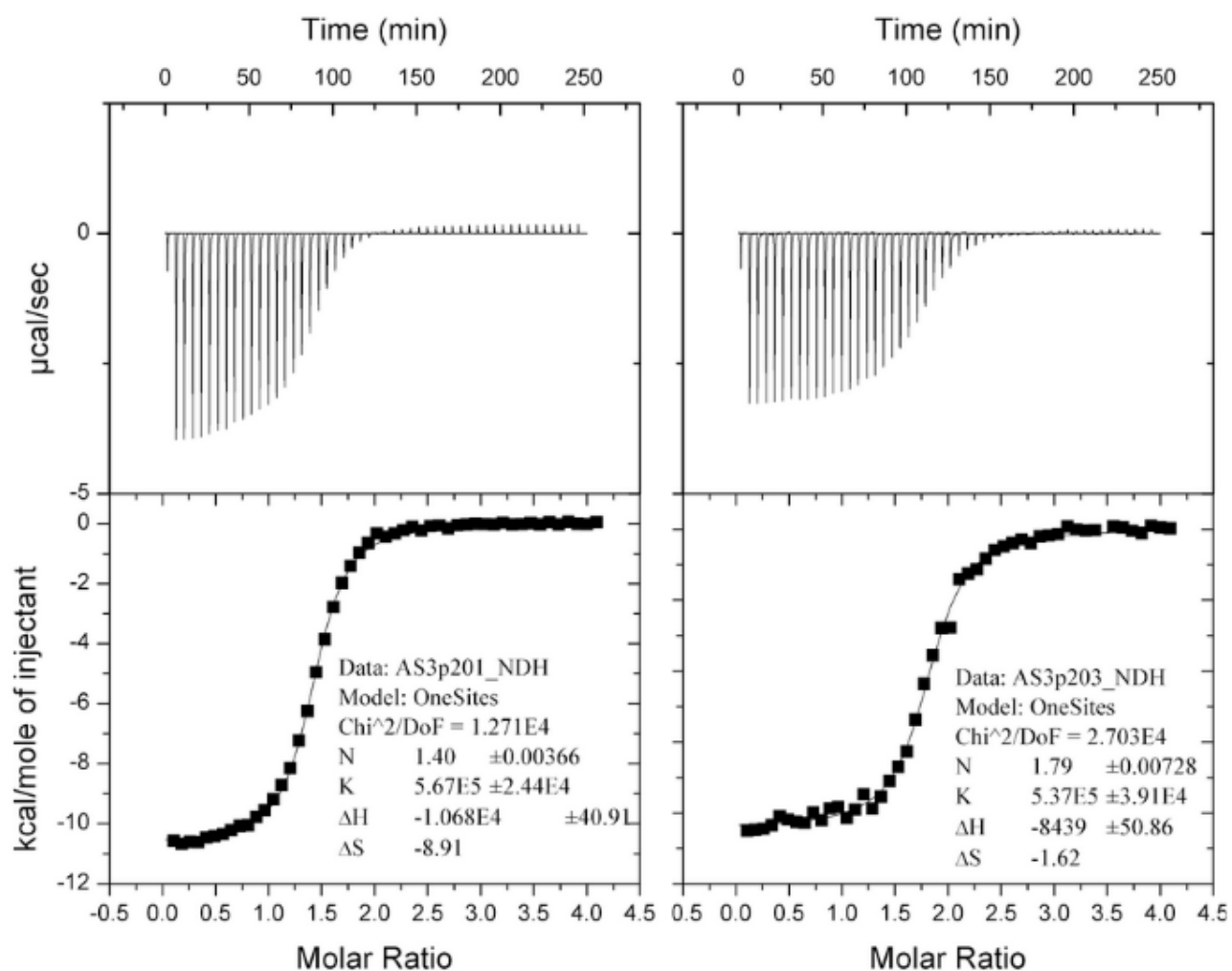
**Figure S17.** Plots for IDA of compound **sCx4** (host). The first plot (upper left) is the direct titration of lucigenin with **sCx4**. The next 8 plots are competitive titrations of **sCx4-Lucigenin** complex with 8 different peptides bearing different modified amino acids. The concentration of lucigenin was fixed at 250 nM for all titrations while the concentration of **sCx4** was fixed at 0.5  $\mu\text{M}$  in the competitive titrations. The experiment was done in a 96-well plate as described in S11.4. Duplicates are shown as red and blue data sets.



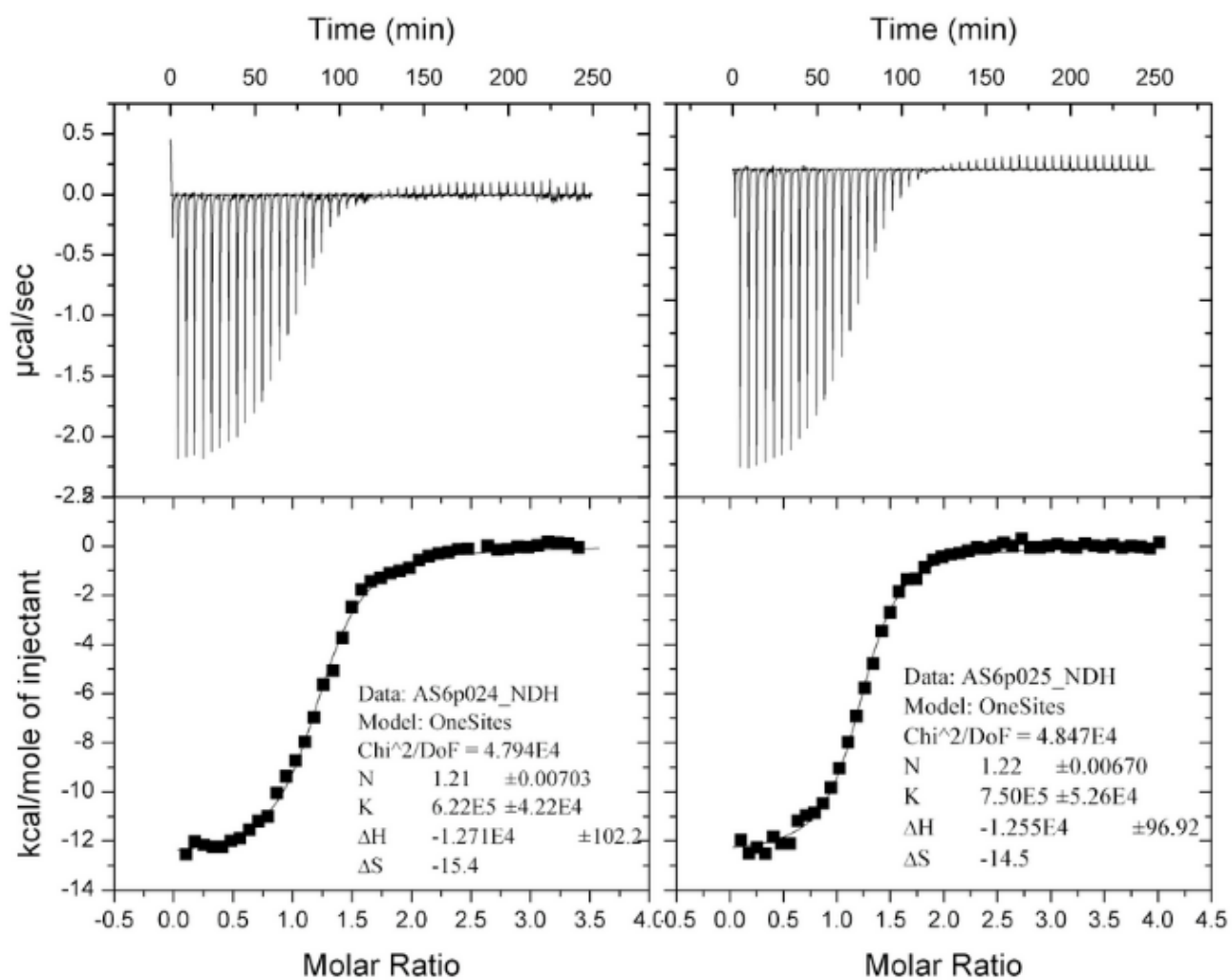
**Figure S18.** Plots for IDA of compound **1** (host). The first plot (upper left) is the direct titration of lucigenin with **1**. The next 8 plots are competitive titrations of **1-Lucigenin** complex with 8 different peptides bearing different modified amino acids. The concentration of lucigenin was fixed at 250 nM for all titrations while the concentration of calixarene **1** was fixed at 0.5  $\mu\text{M}$  in the competitive titrations. The experiment was done in a 96-well plate as described in SI1.4. Duplicates are shown as red and blue data sets.

**S6: Binding isotherms from Isothermal Titration Calorimetry (ITC)** – done as described in SI1.5

S6.a. ITC binding curves between compound **1** and **H3K4me3**: The full procedure is described in SI1.5

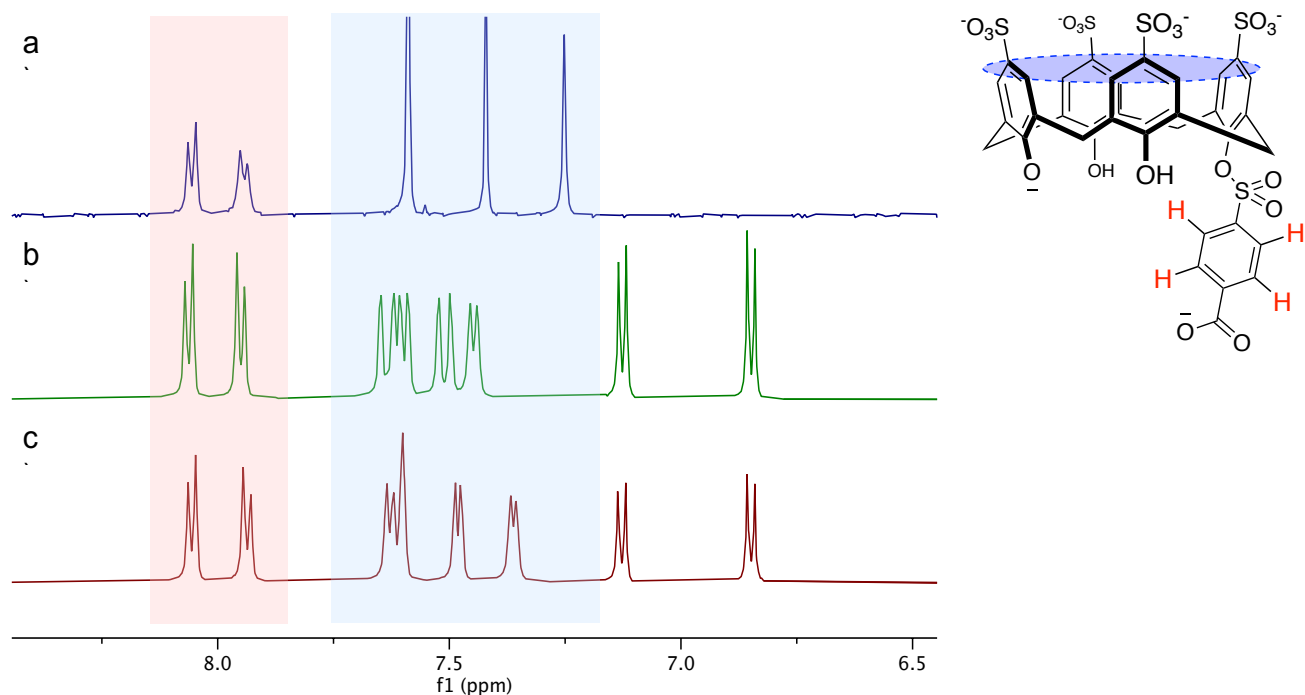


**Figure S19.** Replicate ITC curves for titration of **1** and **H3K4me3** peptide. The concentration of peptide was 2.1 mM (calculated using  $A_{280}$ ) while concentration of calixarene host **1** was kept at 100  $\mu\text{M}$  (determined using qNMR). ITC was run and data analysed as described in General Remarks section (SI1.5).

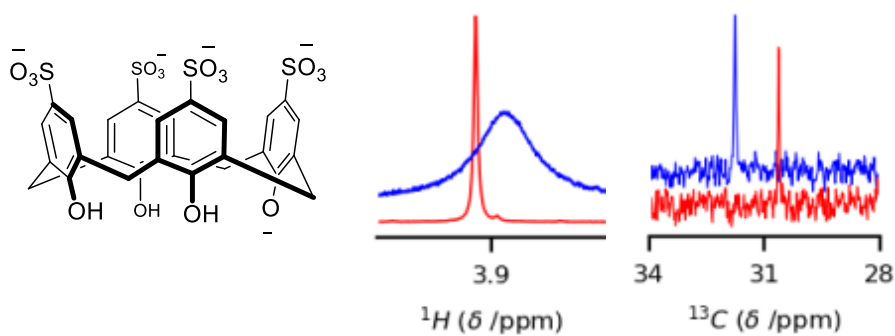


**Figure S20.** Replicate ITC curves for titration of **1** and **H3K4me2** peptide. The concentration of peptide was 1 mM (calculated using  $A_{280}$ ) while concentration of calixarene host **1** was kept at 50  $\mu$ M (determined using qNMR). ITC was run and data analyzed as described in General Remarks section (SI1.5).

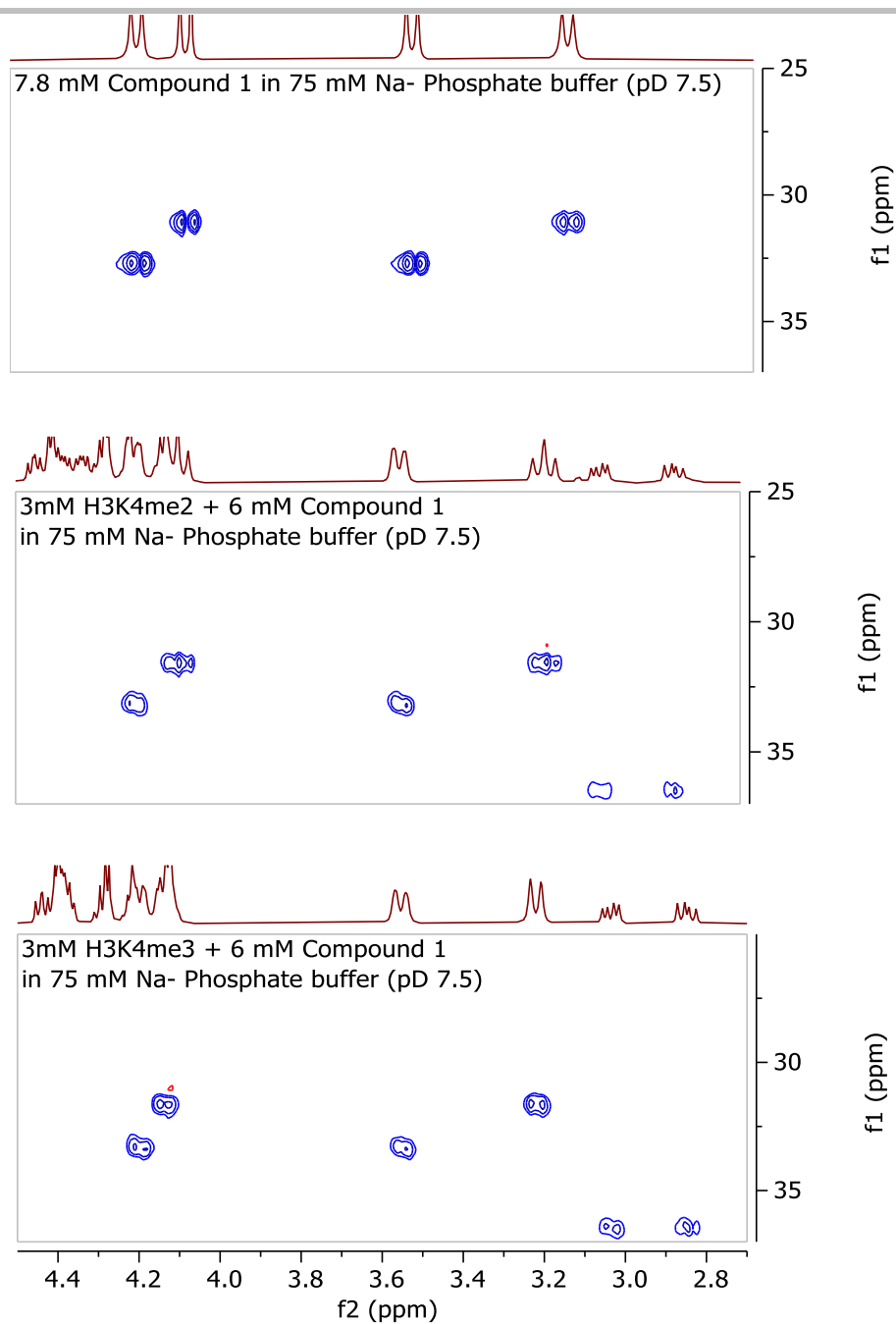
## S7. NMR studies on host guest complex



**Figure S21.** Comparison of aromatic peaks of compound **1** with and without the presence of a guest. a) 1.1 mM **1** in 10 mM sodium phosphate buffer pD 7.5, b) 1 mM **1** + 0.5 mM H3K4me3 peptide in 10 mM sodium phosphate buffer pD 7.5, c) 1 mM **1** + 0.5 mM H3K4me2 peptide in 10 mM sodium phosphate buffer pD 7.5. The structure of host **1** (in cone conformation) is shown on the right and the protons in the blue circle exhibit strong guest induced chemical shift perturbation (blue box in the NMR plot). The pendant benzoic acid's protons are marked in red (in both the structure and NMR plot) and show almost no movement upon complexation. The remaining two doublets in b) and c) are tyrosine resonances from the guest peptide.



**Figure S22.** Chemical shift and lineshape of  $^1\text{H}$  and  $^{13}\text{C}$  NMR trace of parent compound **sCx4** in both  $\text{D}_2\text{O}$  (pD < 2, red trace) and sodium phosphate buffer (pD 7.5, blue trace) support a  $\text{C}_{4v}$  symmetric cone conformation<sup>[4]</sup> with the expected intraannular ring inversion.



**Figure S23.** Comparison of  $^{13}\text{C}$  chemical shift, as judged from  $^1\text{H}$ - $^{13}\text{C}$  HSQC plots of compound **1** (top) in sodium phosphate buffer (75 mM, pD 7.5) and its complex with H3K4me2 (middle) and H3K4me3 (bottom) as guest (in the same buffer), suggest that the conformation of **1** doesn't change upon binding.<sup>[4]</sup>

## References

- [1] P. Thordarson, *Chem. Soc. Rev.*, 2011, **40**, 1305-1323. DOI 10.1039/c0cs00062k.
- [2] Z. X. Wang, *FEBS Lett.*, 1995, **360**, 111-114. DOI 10.1016/0014-5793(95)00062-E.
- [3] G. A. E. Garnett, K. D. Daze, J. A. Peña Díaz, N. Fagen, A. Shaurya, M. C. F. Ma, M. S. Collins, D. W. Johnson, L. N. Zakharov, F. Hof, *Chem. Commun.*, 2016, **52**, 2768-2771. DOI 10.1039/c5cc10527g.
- [4] O. A. Yesypenko, I. B. Vyacheslav, M. A. Klyachina, S. V. Shishkina, O. V. Shishkin, V. V. Pyrozhenko, I. F. Tsymbal, V. I. Kalchenko, *J. Incl. Phenom. Macrocycl. Chem.*, 2012, **74**, 265-275. DOI 10.1007/s10847-012-0109-9.