

**Cell Reports Physical Science, Volume 5**

**Supplemental information**

**Tuning pH-dependent cytotoxicity  
in cancer cells by peripheral fluorine  
substitution on pseudopeptidic cages**

**Lucía Tapia, Yolanda Pérez, Israel Carreira-Barral, Jordi Bujons, Michael Bolte, Carmen Bedia, Jordi Solà, Roberto Quesada, and Ignacio Alfonso**

## Supplemental items for:

# Tuning pH-dependent cytotoxicity in cancer cells by peripheral fluorine substitution on pseudopeptidic cages

Lucía Tapia,<sup>a</sup> Yolanda Pérez,<sup>b</sup> Israel Carreira-Barral,<sup>c</sup> Jordi Bujons,<sup>a</sup> Michael Bolte,<sup>d</sup> Carme Bedia,<sup>e</sup> Jordi Solà,<sup>a</sup> Roberto Quesada<sup>c</sup> and Ignacio Alfonso<sup>a,\*</sup>

a. Department of Biological Chemistry, Institute for Advanced Chemistry of Catalonia, Spanish Research Council, IQAC-CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain.

b. NMR Facility, Institute for Advanced Chemistry of Catalonia, Spanish Research Council, IQAC-CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain.

c. Departamento de Química, Facultad de Ciencias, Universidad de Burgos, 09001 Burgos, Spain.

d. Institut für Anorganische Chemie, J.-W.-Goethe-Universität, Max-von-Laue-Str.7 60438 Frankfurt/Main, Germany.

e. Institute of Environment Assessment and Water Research, Spanish Research Council, IDAEA-CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain.

## Table of Contents

Synthesis and characterization of cages	S-2
X-ray crystal analysis of <b>1h</b> ·4HCl	S-74
<sup>1</sup> H NMR titration of compounds <b>1h</b> , <b>1i</b> , <b>1m</b> and <b>1l</b> with TBACl	S-76
pK <sub>a</sub> determination by fluorescence spectroscopy	S-85
Transmembrane anion transport experiments in vesicles	S-87
NMR measurements	S-102
Cytotoxicity studies	S-114
Computational methods	S-122
Computational results	S-125

## Supplemental Experimental

### Procedures Synthesis and

#### characterization of cages

**Materials and methods:** Reagents and materials were purchased from commercial suppliers (Aldrich, Fluka or Merck) and were used without further purification. All the compounds prepared were fully characterized by the complete spectroscopic (NMR, ESI-MS) and analytical data. Preparative reverse phase purifications were performed on an Isolera Biotage instrument (KP-C18-HS, CH<sub>3</sub>CN and water with 0.1% TFA). Analytical RP-HPLC was performed with a Hewlett Packard Series 1100 (UV detector 1315A) modular system using a reverse-phase Kromasil 100 C8 (15 x 0.46 cm, 5 μm) column. CH<sub>3</sub>CN-H<sub>2</sub>O mixtures containing 0.1% TFA at 1 mL/min were used as mobile phase and monitoring wavelengths were set at 220, 254 and 280 nm.

**NMR spectroscopy (compounds characterization):** The NMR experiments were carried out at 25 °C on a VNMRS-400 NMR spectrometer (Agilent Technologies 400 MHz for <sup>1</sup>H and 100 MHz for <sup>13</sup>C) for characterization and a Bruker Avance-III 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance (<sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N) TCI cryoprobe (500 MHz for 1H and 125 MHz for 13C) for titrations. Chemical shifts are reported in ppm using tetrakis-trimethylsilylsilane as a reference. Data were processed with the software program MNova (Mestrelab Research).

**ESI mass spectrometry:** High resolution mass spectra (HRMS) were performed on Acquity UPLC System and a LCT Premier<sup>TM</sup> XE Benchtop orthogonal acceleration time-of-flight (oa-TOF) (Waters Corporation, Milford, MA) equipped with an electrospray ionization source. All sample solution (in the 1 x 10<sup>-4</sup> to 1 x 10<sup>-6</sup> M range) were prepared in methanol.

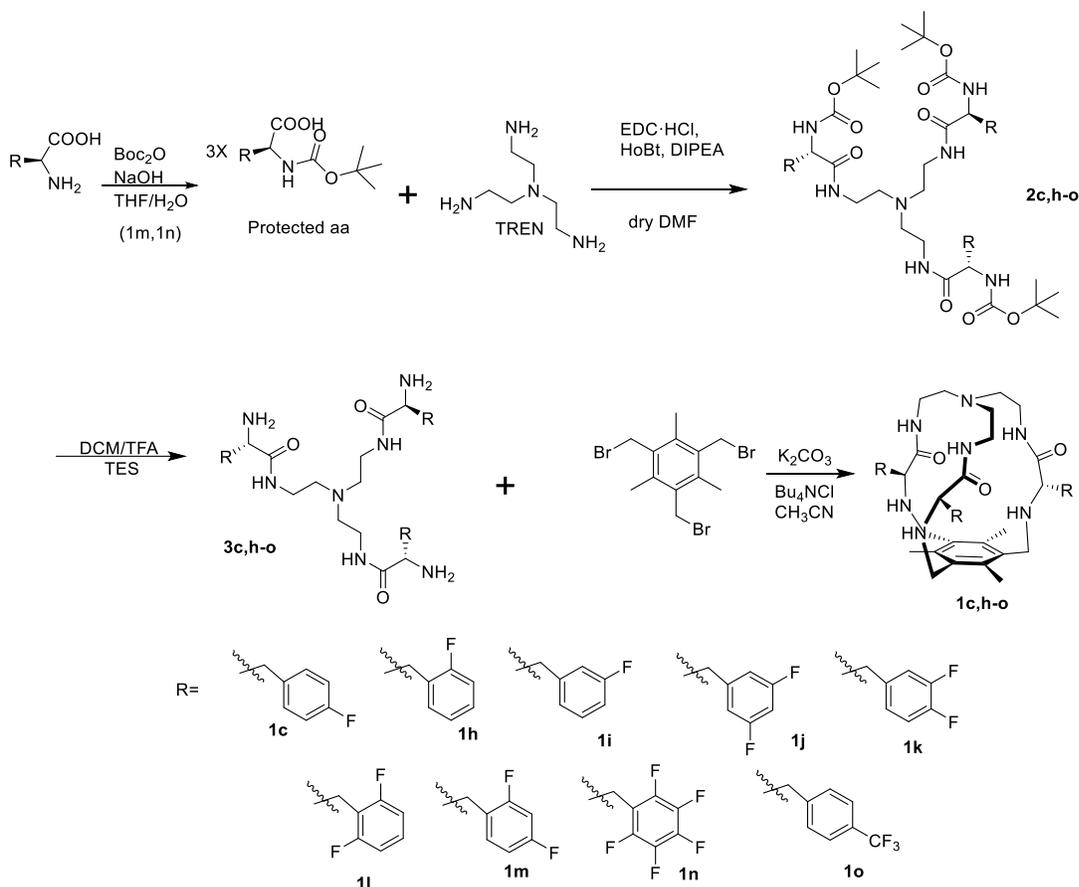
**X-RAY diffraction:** Data were collected on a STOE IPDS II two-circle diffractometer with a Genix Microfocus tube with mirror optics using MoK $\alpha$  radiation ( $\lambda = 0.71073 \text{ \AA}$ ). The data were scaled using the frame scaling procedure in the X-Area program system.<sup>1</sup> The structure was solved by direct methods using the program SHELXS<sup>2</sup> and refined against F<sup>2</sup> with full-matrix least-squares techniques using the program SHELXL<sup>3</sup>

**Fluorescence spectroscopy:** Fluorescence emission was recorded using a HITACHI F-7000 Fluorescence spectrophotometer equipped with stirrer. Two different types of cuvettes were used depending on the assay: standard 10 mm quartz glass cells from Hellma Analytics High Precision Cell made of Quartz SUPRASIL or disposables MAPM-F10-100 Labbox cuvettes. All measurements were performed at 25 °C unless specified.

**pH measurements** were made using a CRISON pHmeter 50 14 T.

## Synthesis of compounds 1h-o

The synthetic procedure followed is described in Scheme S1. Compounds **1m** and **1n** required an additional step (protection of the amine nitrogen of the amino acid) whereas for the rest of the compounds the synthesis started with the N-Boc protected amino acid. Compound **1c** was synthesized as previously reported.



**Scheme S1.** Synthetic scheme for the preparation of compounds **1c**, **1h-o**.

### Synthesis of **1h** (2F-Phe)

**2h**: Boc-Phe(2-F)-OH (419 mg, 1.48 mmol) was dissolved in dry DMF (2.5 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 0.344 mg, 1.79 mmol) and 1-Hydroxybenzotriazole hydrate (HOBT, 0.274 g, 1.79 mmol), *N,N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.38 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.44 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM : MeOH as eluent to give 330 mg of **2h** (0.351 mmol, 78 % yield).

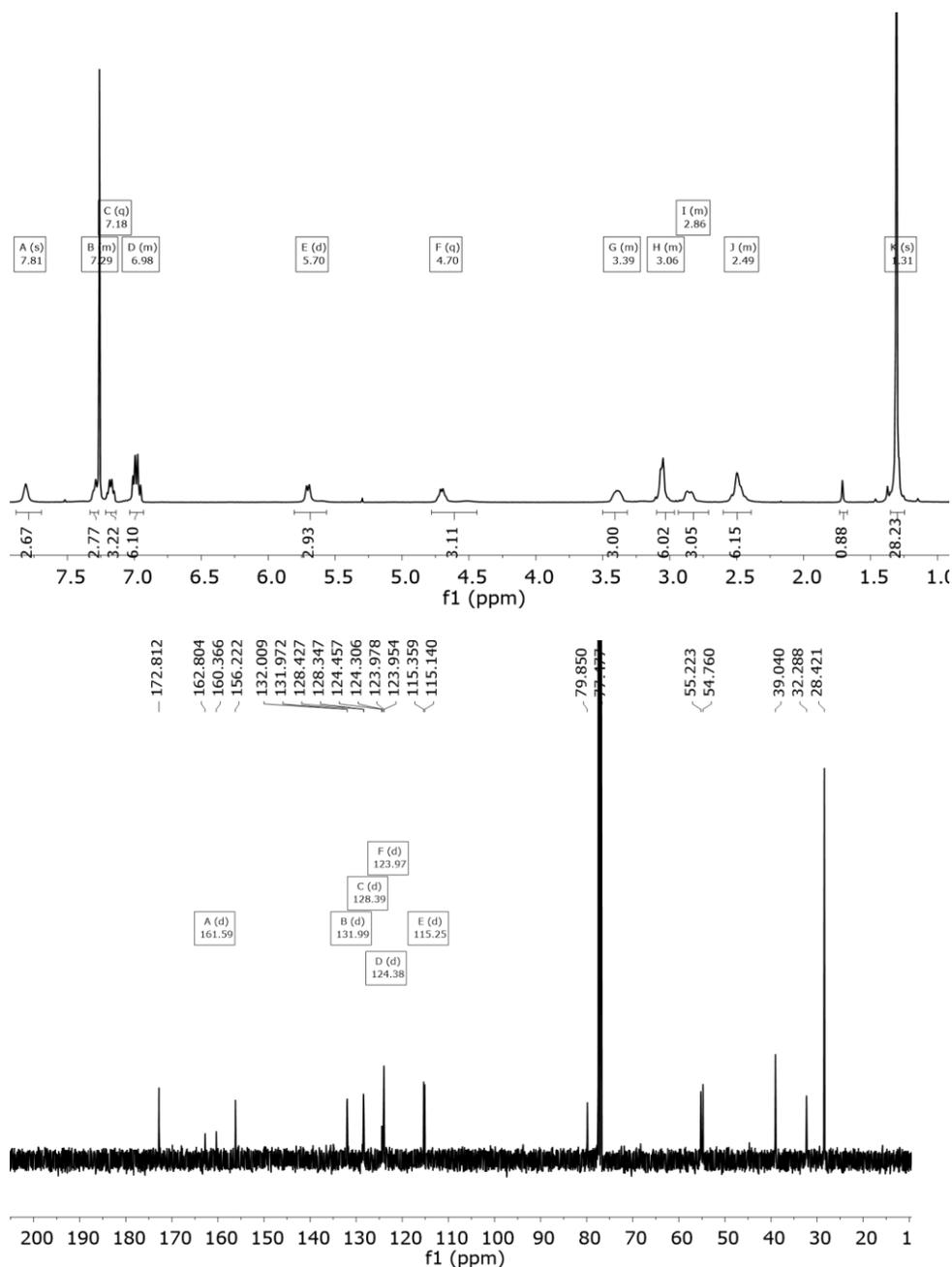
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.81 (s, 3H), 7.29 (m, 3H), 7.18 (m, 3H), 6.98 (m, 6H), 5.70 (d,  $J = 8.8$  Hz, 3H), 4.703 (m, 3H), 3.39 (m, 3H), 3.06 (m, 6H), 2.86 (m, 3H), 2.49 (m, 6H), 1.306 (s, 27H).

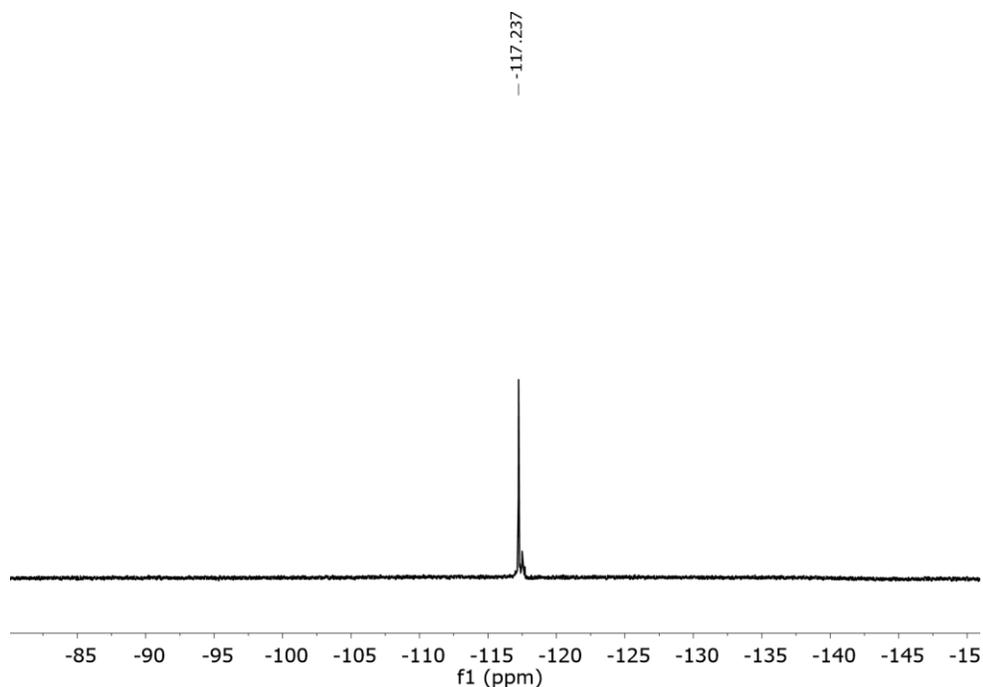
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.8, 161.6 (d,  $^1J_{\text{C-F}} = 245.4$  Hz), 156.2, 132.0 (d,  $^nJ_{\text{C-F}} = 3.7$  Hz), 128.4 (d,  $^nJ_{\text{C-F}} = 8.1$  Hz) 124.3 (d,  $^nJ_{\text{C-F}} = 15.2$  Hz), 124.0, 115.2 (d,  $^nJ_{\text{C-F}} = 22.1$  Hz), 79.9, 55.2, 54.8, 39.0, 32.3, 28.4.

$^{19}\text{F}$ - $\{^1\text{H}\}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -117.23.

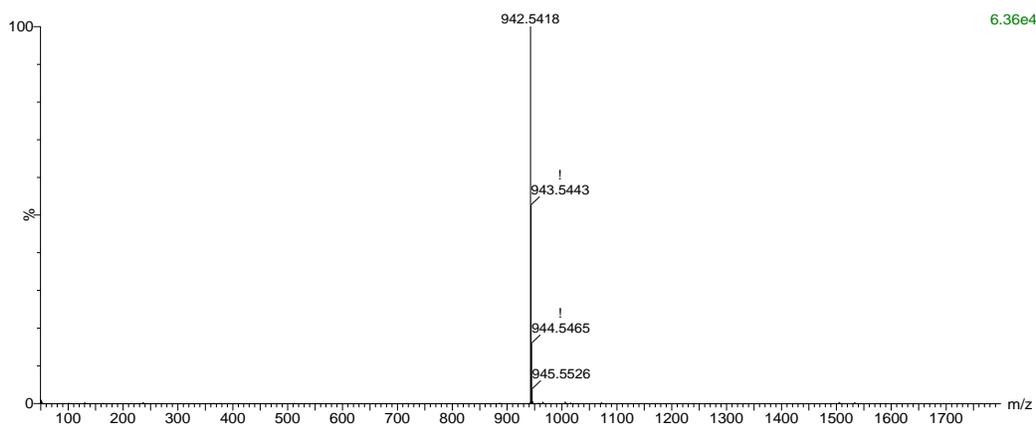
HRMS (ESI-TOF)  $m/z$  [ $2\mathbf{h} + \text{H}$ ] $^+$  Calc: 942.4947, found: 942.5418

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CD}_3\text{OD}$ ) spectra of  $2\mathbf{h}$ :





HRMS (ESI+) experimental spectrum of **2h**:



**3h**: **2h** (280 mg, 0.287 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.65 mL, 4.31 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3h**·4TFA as a white solid (279 mg, 0.250 mmol, 89 % yield).

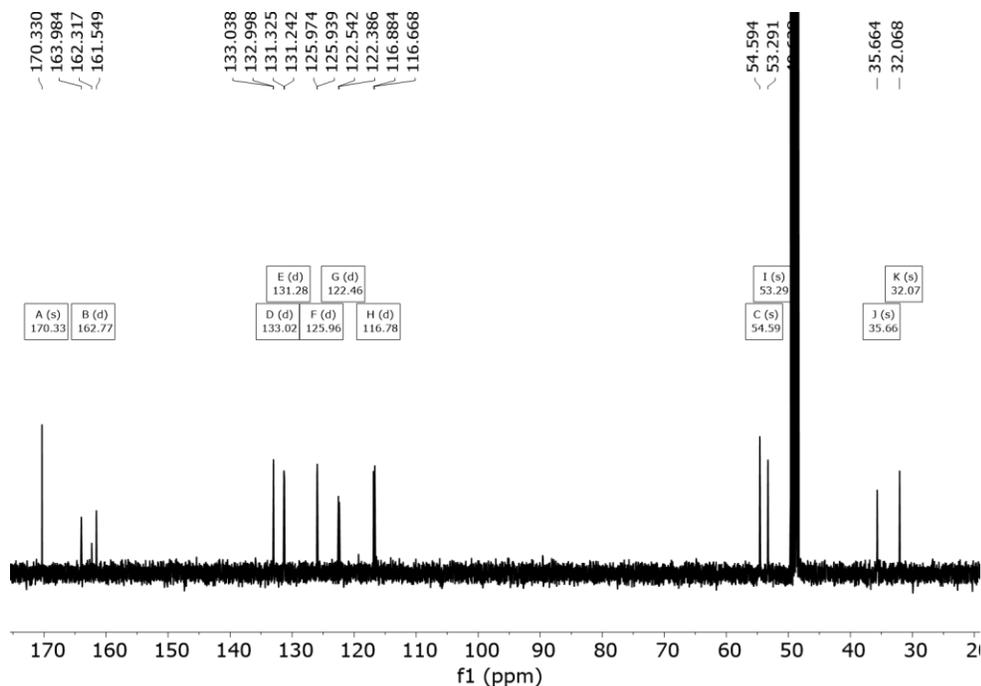
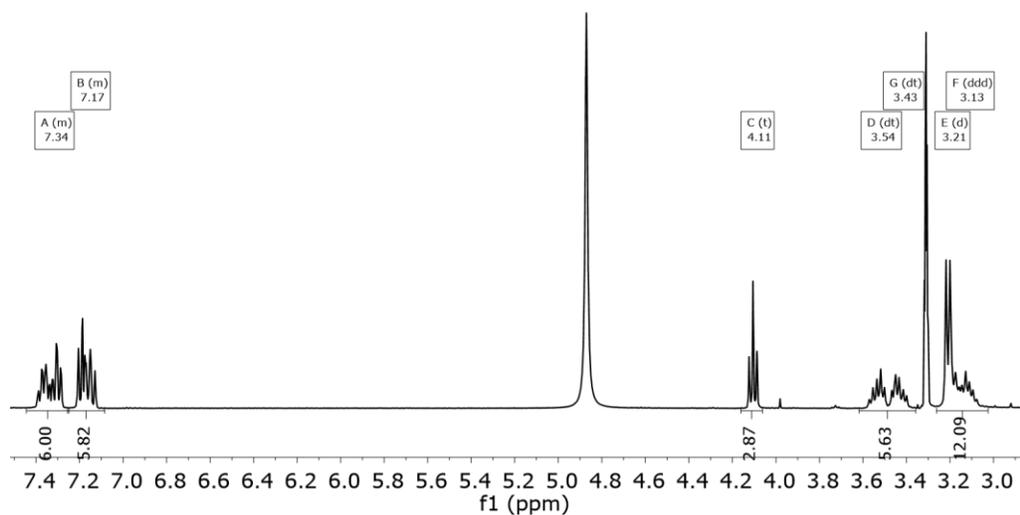
<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 7.34 (m, 6H), 7.17 (m, 6H), 4.11 (X subsystem form ABX,  $J_{AX} = 7.0$ ,  $J_{BX} = 6.5$  Hz, 3H), 3.54 (A subsystem form ABX,  $J_{AX} = 7$ ,  $J_{AB} = 14$  Hz, 3H), 3.43 (B subsystem form ABX,  $J_{BX} = 6.5$ ,  $J_{AB} = 14$  Hz, 3H), 3.21 (d,  $J = 7.3$  Hz, 6H), 3.13 (m, 6H).

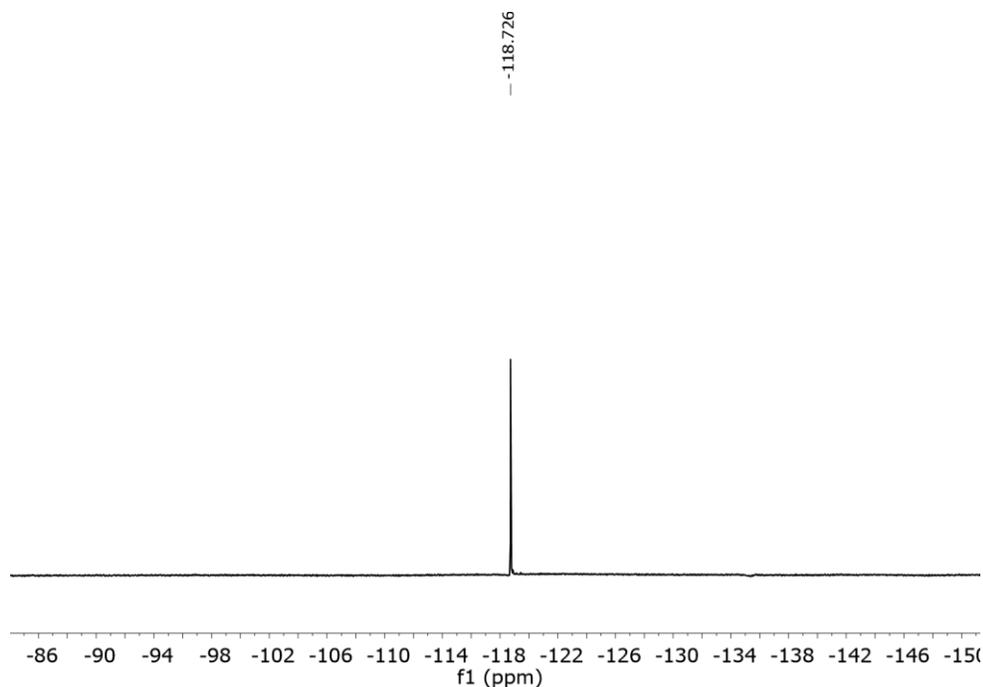
$^{13}\text{C}$  NMR (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  170.3, 162.8 (d,  $^1J_{\text{C-F}} = 245.0$  Hz), 133.0 (d,  $^nJ_{\text{C-F}} = 4.0$  Hz), 131.3 (d,  $^nJ_{\text{C-F}} = 8.3$  Hz), 125.9 (d,  $^nJ_{\text{C-F}} = 3.5$  Hz), 122.5 (d,  $^nJ_{\text{C-F}} = 15.7$  Hz), 116.8 (d,  $^nJ_{\text{C-F}} = 21.8$  Hz), 54.6, 53.2, 35.7, 32.1.

$^{19}\text{F}$ - $\{^1\text{H}\}$  NMR (376 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  -118.73.

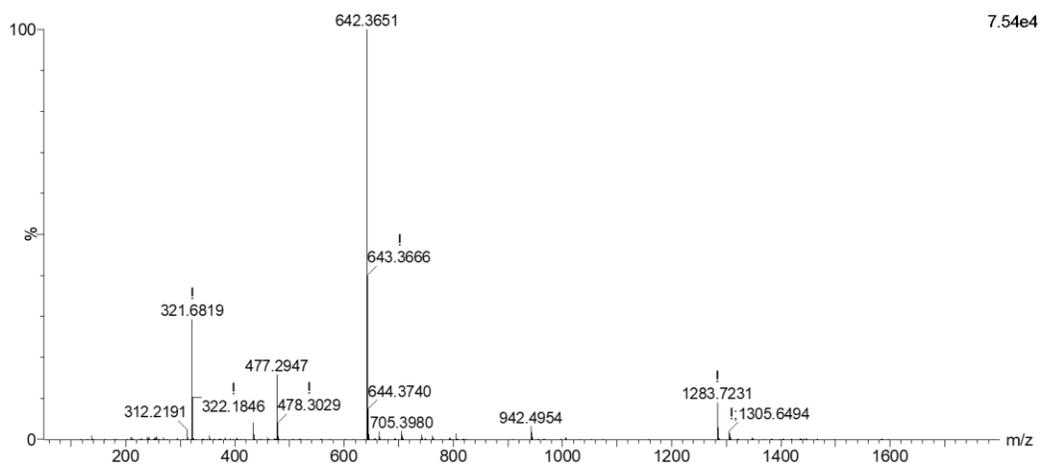
HRMS (ESI-TOF)  $m/z$  [ $\mathbf{3h} + \text{H}$ ] $^+$  Calc: 642.3374, found: 642.3651.

$^1\text{H}$ -NMR (400 MHz,  $\text{CD}_3\text{OD}$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CD}_3\text{OD}$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CD}_3\text{OD}$ ) spectra of  $\mathbf{3h}$ :

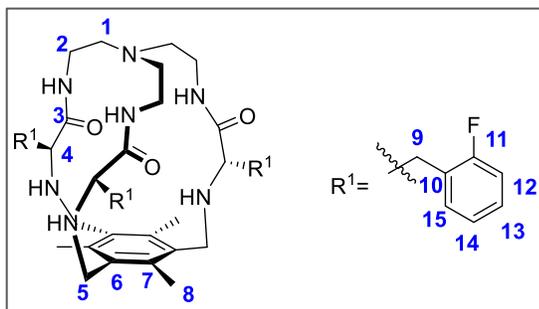




HRMS (ESI+) experimental spectrum of **3h**:



**1h (2F-Phe): 3h** (160 mg, 0.146 mmol) was dissolved in ACN (40 mL). Tetrabutylammonium chloride (20 mg, 0.073 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (58 mg, 0.146 mmol) and potassium carbonate (402 mg, 2.91 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 95 : 5 DCM : MeOH as eluent to give **1h** as a white solid (40 mg, 0.05 mmol, 35 % yield).



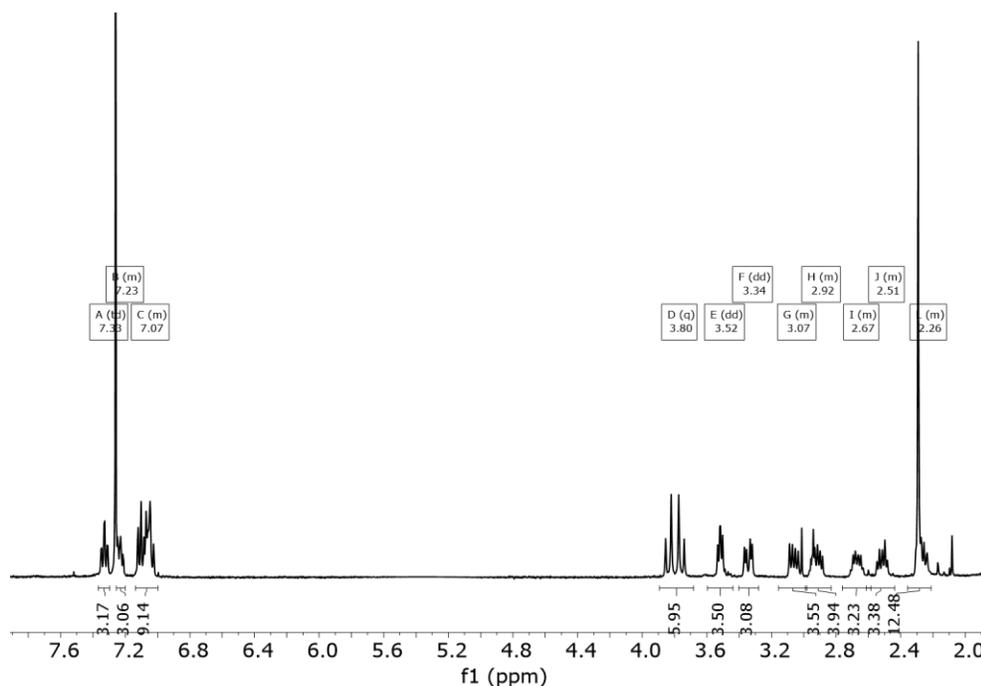
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.33 (td,  $J = 7.6, 1.8$  Hz, 3H (H12/13)), 7.22(m, 3H, (H12/13)), 7.07 (m, 6H, (H14,15)), 3.80 (AB<sub>q</sub>,  $\delta_A = 3.81, \delta_B = 3.74, J_{AB} = 13.6$  Hz, 6H, (H5)), 3.52 (X subsystem from ABX,  $J_{AX} = 5.1, J_{BX} = 7.1$  Hz, 3H (H4)), 3.34 (A subsystem from ABX,  $J_{AX} = 5.1, J_{AB} = 14.0$  Hz, 3H, (H9)), 3.06 (B subsystem from ABX,  $J_{BX} = 7.1, J_{AB} = 14.0$  Hz, 3H, (H9)), 2.92 (m, 3H, (H2)), 2.67 (m, 3H, (H2)), 2.53 (m, 3H, (H1)), 2.26 (s, 9H, (H8) overlapped with m, 3H, (H1)).

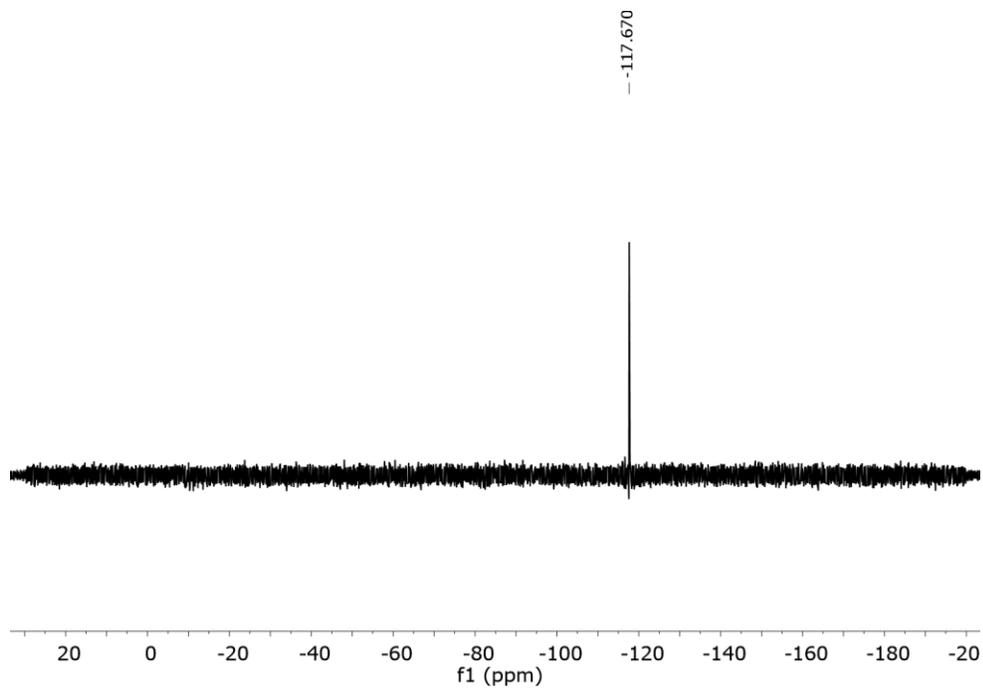
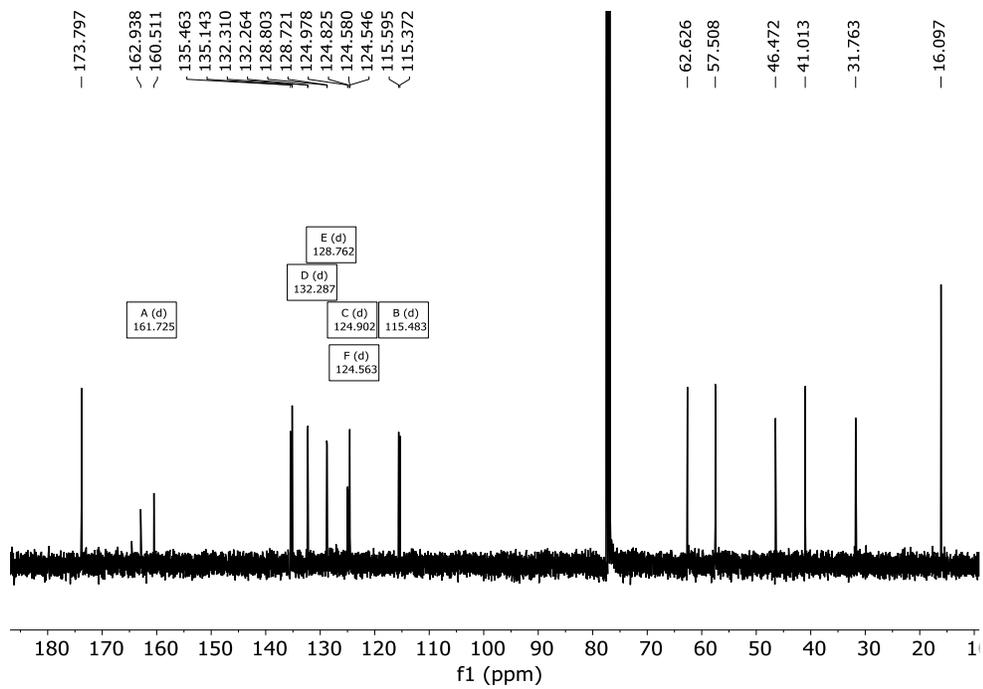
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.8 (C3), 161.7 (d,  $^1J_{\text{C-F}} = 244.2$  Hz, C11), 135.5 (C6/7), 135.1 (C6/7), 132.3 (d,  $^nJ_{\text{C-F}} = 4.7$  Hz, C13), 128.8 (d,  $^nJ_{\text{C-F}} = 8.2$  Hz, C14/15), 125.0 (d,  $^nJ_{\text{C-F}} = 15.4$  Hz, C10), 124.6 (d,  $^nJ_{\text{C-F}} = 3.5$  Hz, C14/15), 115.4 (d,  $^nJ_{\text{C-F}} = 22.4$  Hz, C12), 62.6 (C4), 57.5 (C1), 46.47 (C5), 41.0 (C2), 31.8 (C9), 16.1 (C8).

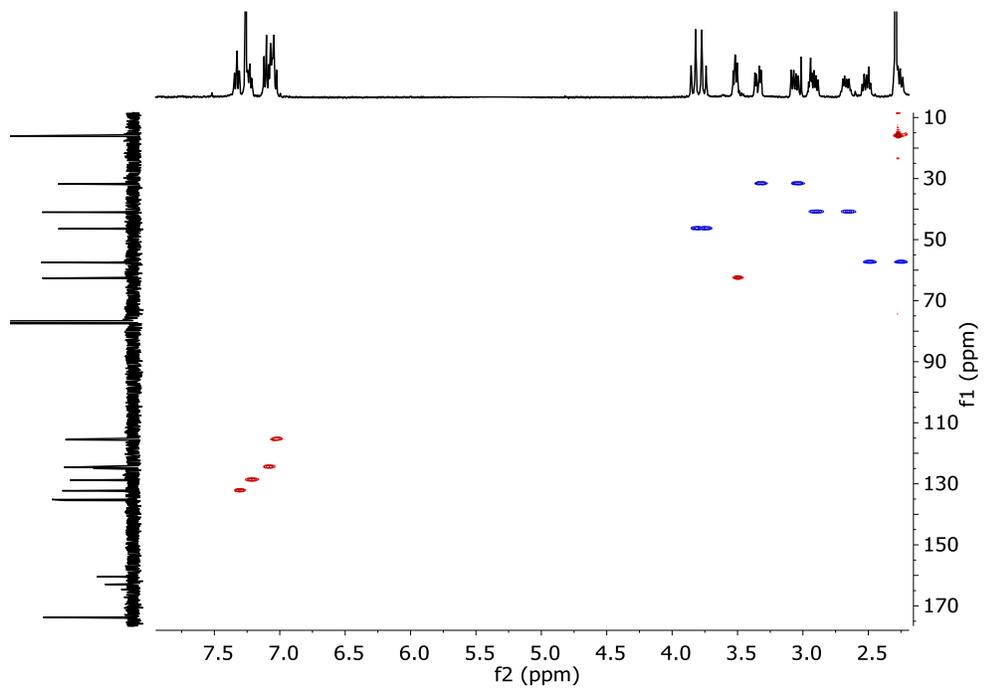
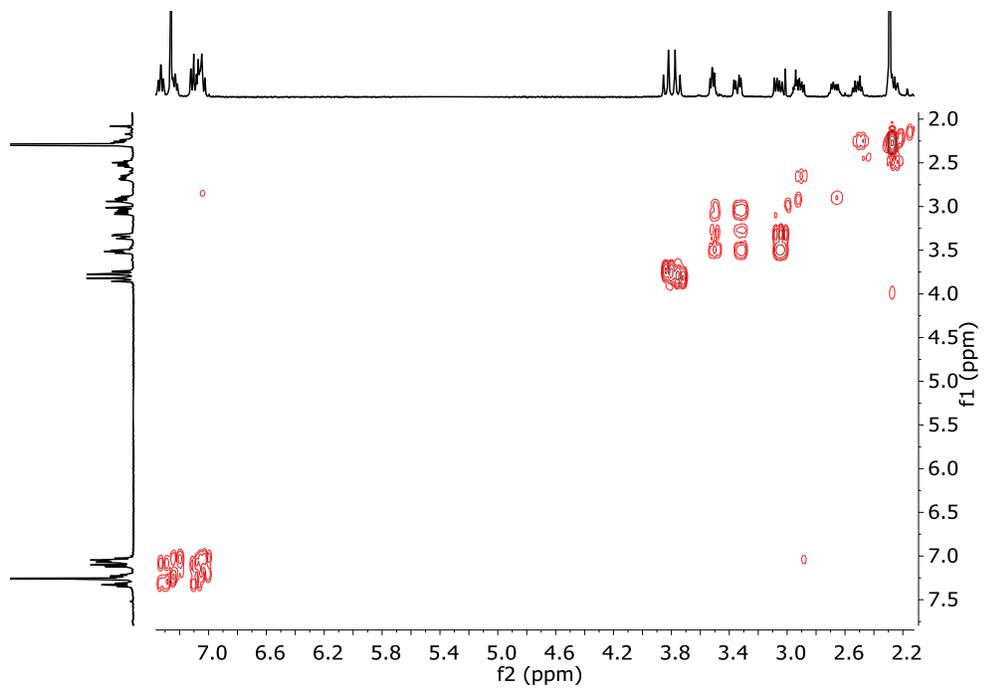
$^{19}\text{F}$ - $\{^1\text{H}\}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -117.67.

HRMS (ESI-TOF)  $m/z$  [ $\mathbf{1h} + \text{H}$ ] $^+$  Calc: 798.4313, found: 798.4215.

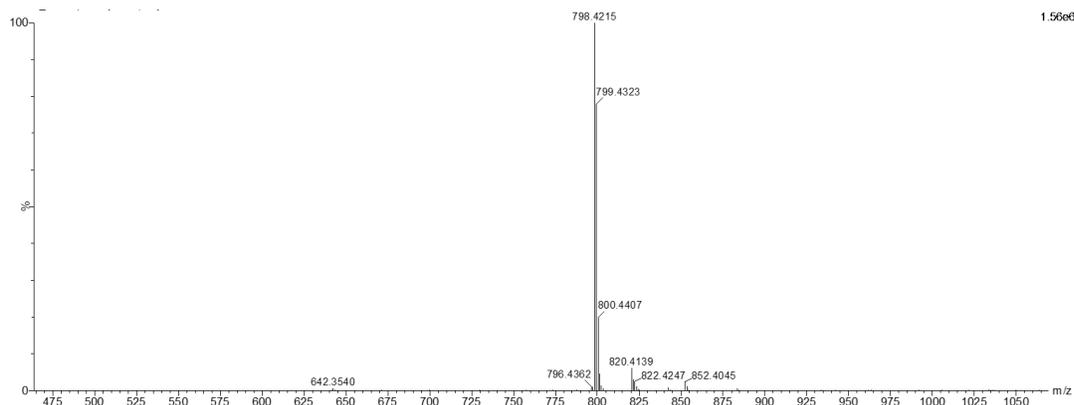
$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ), COSY and HSQC spectra of **1h**:







HRMS (ESI+) experimental spectrum of **1h**.



### Synthesis of **1i** (**3F-Phe**)

**2i**: Boc-Phe(3-F)-OH (500 mg, 1.76 mmol) was dissolved in dry DMF (3 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 384 mg, 2 mmol) and 1-hydroxybenzotriazole hydrate (HOBT, 306 mg, 2 mmol), *N,N*-diisopropylethylamine (DIPEA, 1 mL, 5.78 mmol) and tris(2-aminoethyl)amine (0.08 mL, 0.56 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was purified by flash chromatography using 95:5 DCM:MeOH to give 0.412 mg of **2i** (0.437 mmol, 90 % yield).

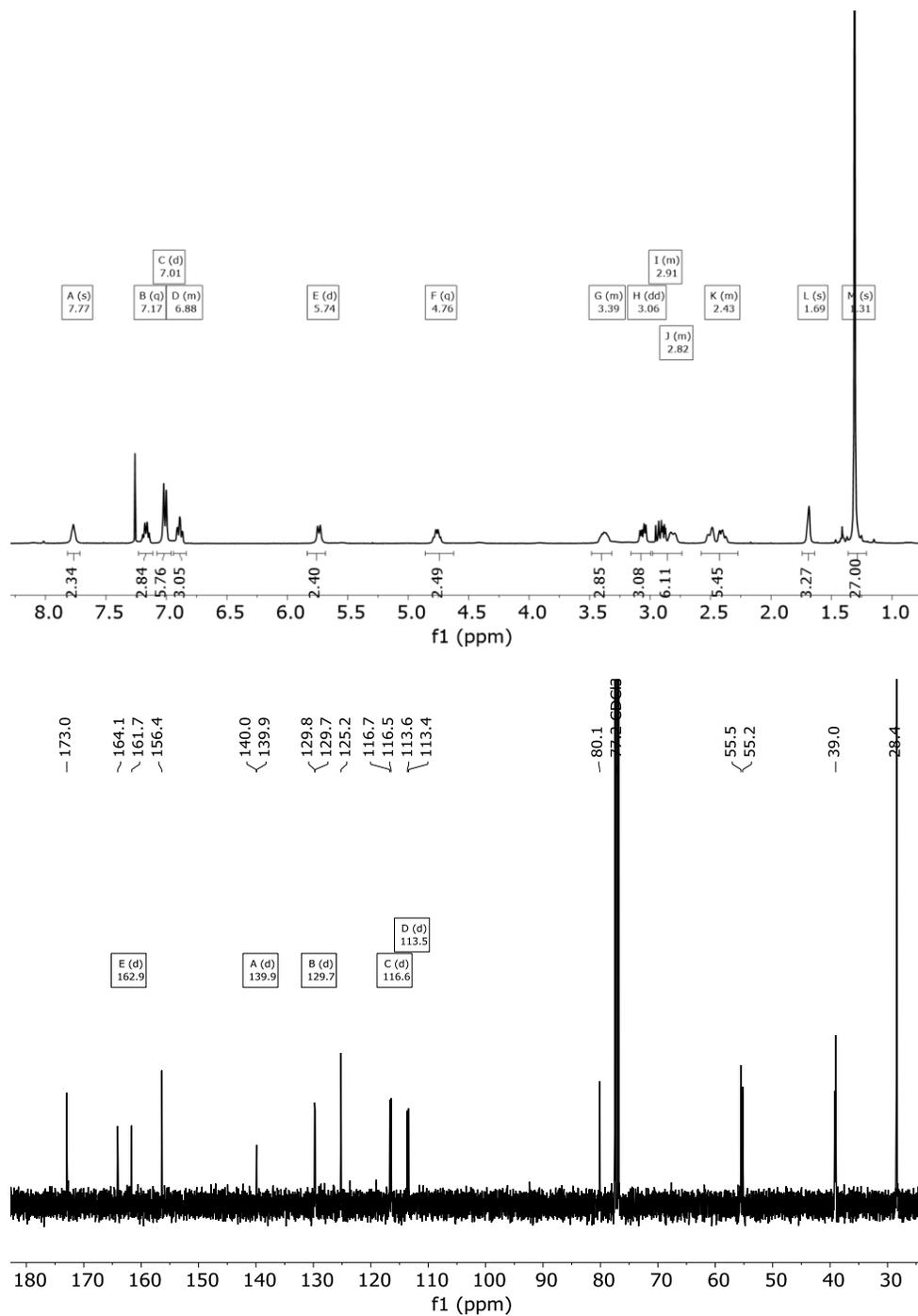
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.77 (s, 3H), 7.17 (q, *J* = 7.3 Hz, 3H), 7.01 (d, *J* = 8.1 Hz, 6H), 6.95 – 6.79 (m, 3H), 5.74 (d, *J* = 8.9 Hz, 3H), 4.76 (q, *J* = 8.2 Hz, 3H), 3.39 (m, 3H), 3.06 (dd, *J* = 13.6, 6.1 Hz, 3H), 2.91 (m, 3H), 2.82 (m, 3H), 2.43 (m, 6H), 1.69 (s, 3H), 1.31 (s, 27H).

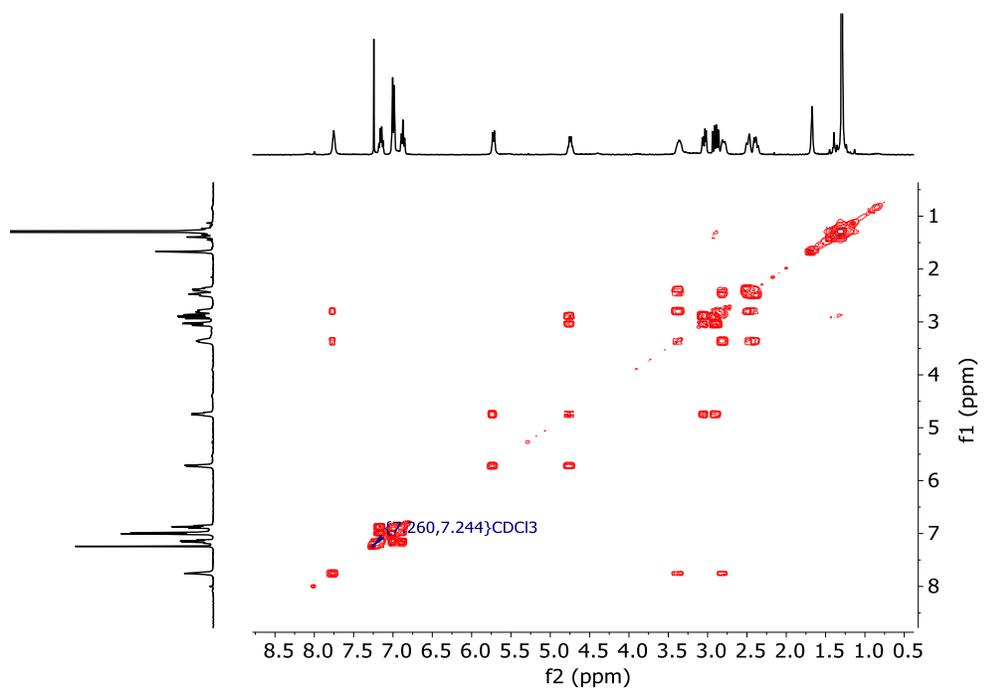
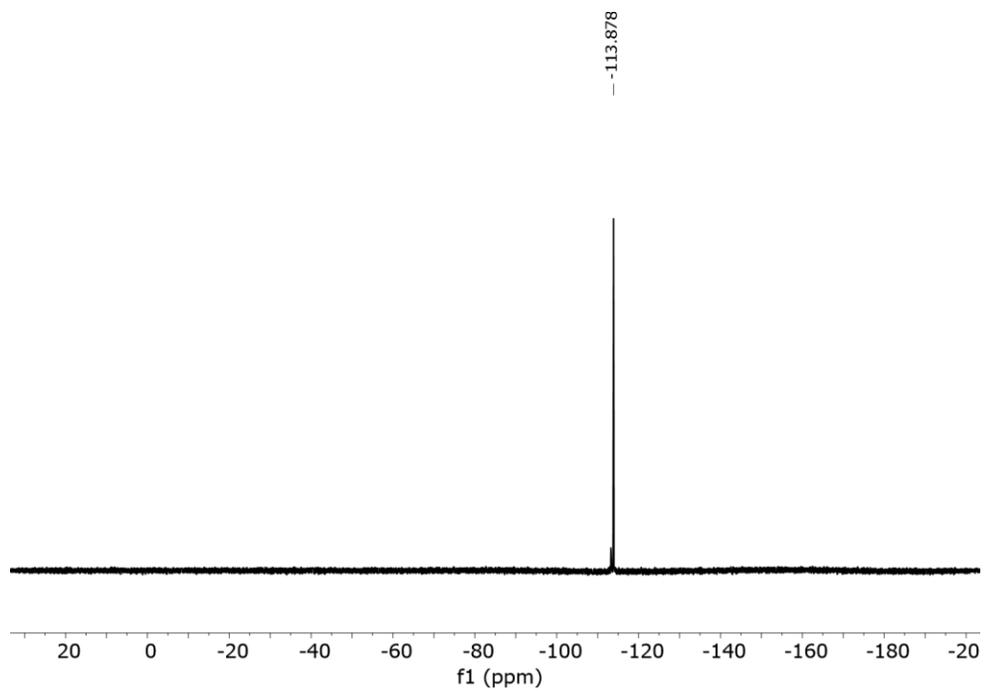
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 173.0, 162.9 (d, <sup>1</sup>*J*<sub>C-F</sub> = 245.1 Hz), 156.4, 139.9 (d, <sup>n</sup>*J*<sub>C-F</sub> = 7.4 Hz), 129.7 (d, <sup>n</sup>*J*<sub>C-F</sub> = 8.6 Hz), 125.2, 116.6 (d, <sup>n</sup>*J*<sub>C-F</sub> = 21.4 Hz), 113.5 (d, <sup>n</sup>*J*<sub>C-F</sub> = 20.7 Hz), 80.1, 55.5, 39.1, 39.0, 28.4.

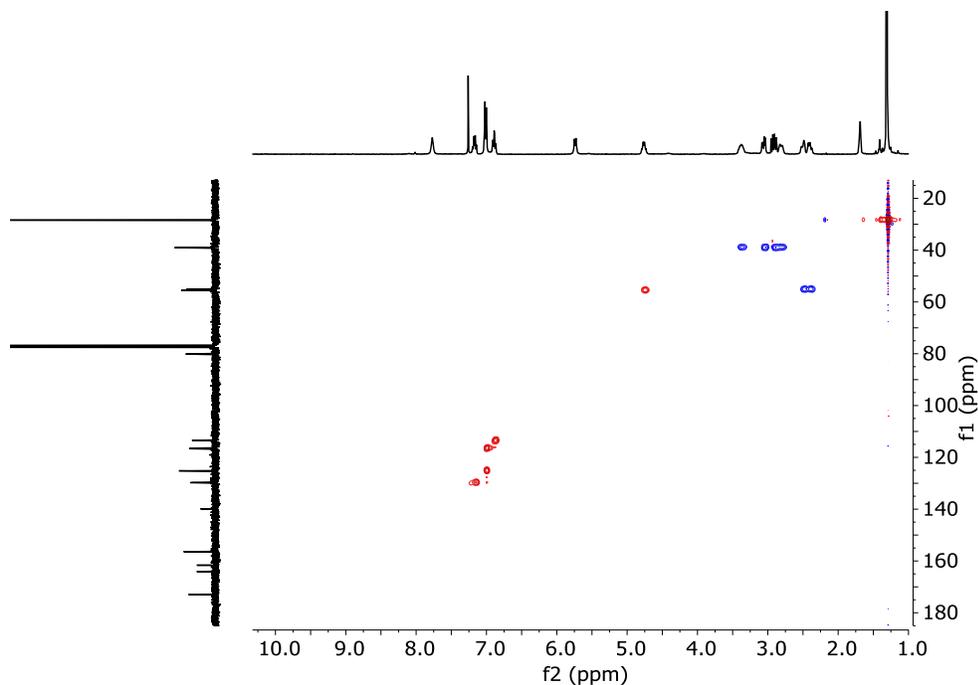
<sup>19</sup>F-<sup>{1</sup>H} NMR (376 MHz, CDCl<sub>3</sub>) δ -113.88.

HRMS (ESI-TOF) m/z [**2i** + H]<sup>+</sup> Calc: 942.4947, found: 942.5024.

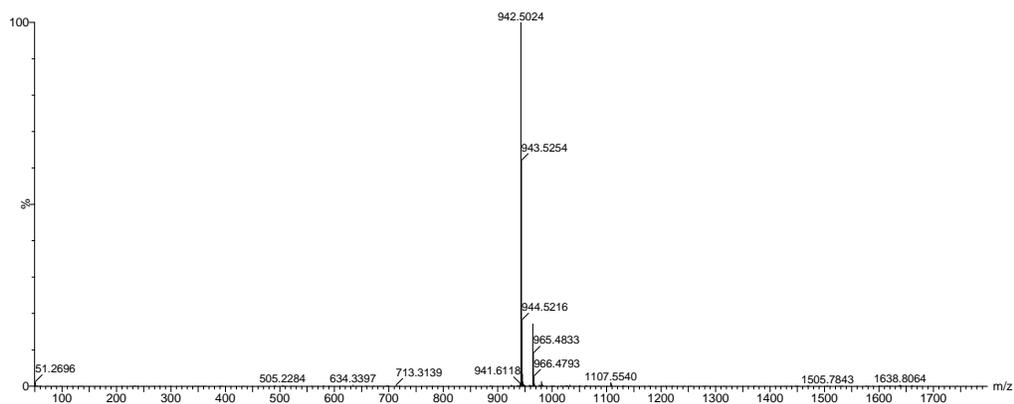
$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ ), COSY and HSQC spectra of **2i**:







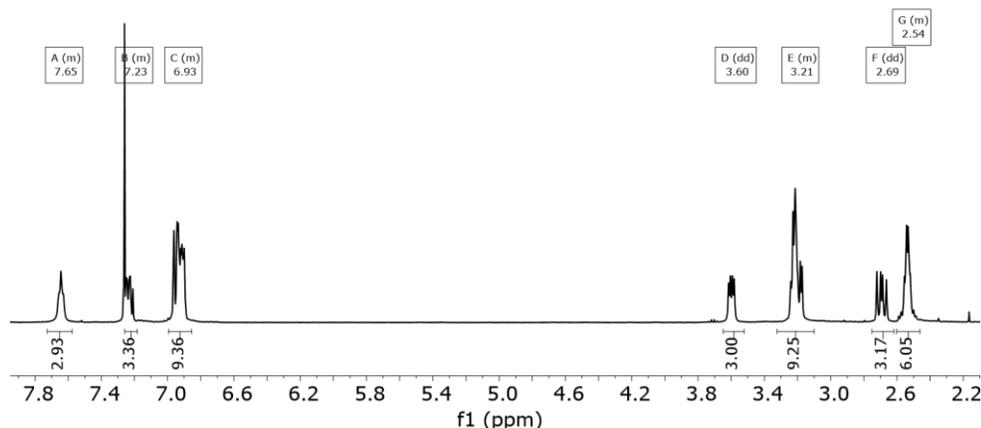
HRMS (ESI+) experimental spectrum of **2i**:



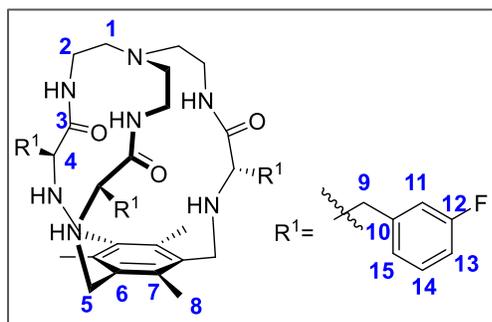
**3i**: **2i** (400 mg, 0.425 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.65 mL, 4.31 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3i**·4TFA as a white solid. The solid was dissolved in aqueous NaOH 0.1 M and extracted with DCM. Organic layers were dried affording **3i** as a white solid (255 mg, 0.397 mmol, 93% yield).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.65 (m, 3H), 7.23 (m, 3H), 6.93 (m, 9H), 3.60 (dd,  $J = 9.1, 4.3$  Hz, 3H), 3.21 (m, 9H), 2.69 (dd,  $J = 13.7, 9.1$  Hz, 3H), 2.54 (m, 6H).

<sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD) spectra of **3i**:



**1i (3F-Phe): 3i** (200 mg, 0.311 mmol) was dissolved in ACN (40 mL). Tetrabutylammonium chloride (30 mg, 0.1095 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (116 mg, 0.622 mmol) and potassium carbonate (500 mg, 3.62 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 97 : 3 DCM : MeOH as eluent to give **1i** as a white solid (90 mg, 0.112 mmol, 38 % yield).



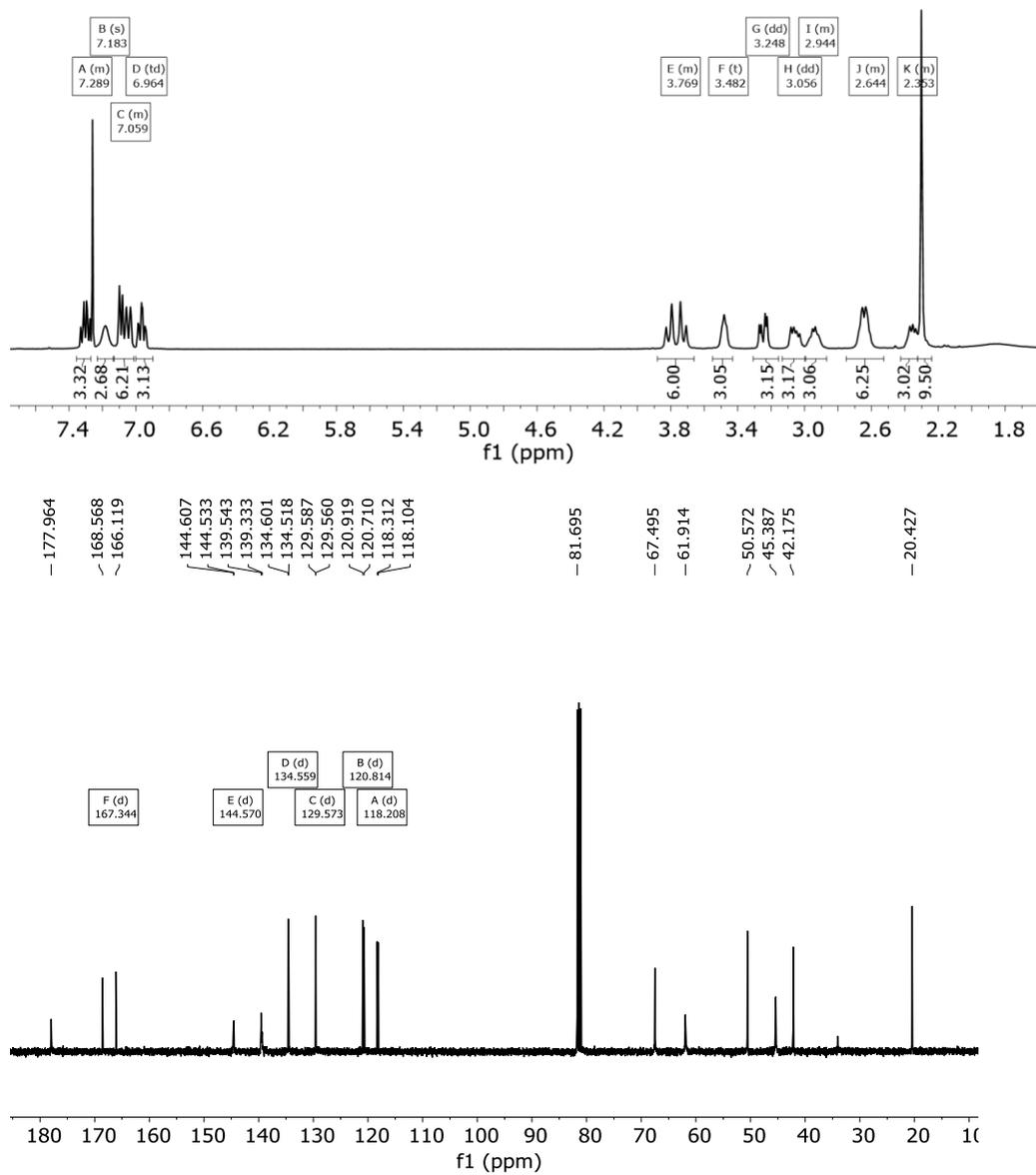
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.29 (m, 3H, **Har**), 7.18 (s, 3H (NH<sub>amide</sub>)), 7.06 (m, 6H, (**Har**)), 6.96 (m, 3H, (**Har**)), 3.77 ((AB<sub>q</sub>, δ<sub>A</sub> = 3.838, δ<sub>B</sub> = 3.757, J<sub>AB</sub> = 13.7 Hz 6H, (**H5**)), 3.48 (X subsystem from ABX, J<sub>AX</sub> = 5.1, J<sub>BX</sub> = 7.1 Hz, 3H (**H4**)), 3.25 (A subsystem from ABX, J<sub>AX</sub> = 5.1, J<sub>AB</sub> = 14.0 Hz, 3H, (**H9**)), 3.06 (B subsystem from ABX, J<sub>BX</sub> = 7.1, J<sub>AB</sub> = 14.0 Hz, 3H, (**H9**)), 2.94 (m, 3H, (**H2**)), 2.64 (m, 6H, (**H2+1**)), 2.35 (m, 3H, (**H1**)), 2.30 (s, 9H, (**H8**)).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 178.0 (**C3**), 167.3 (d, <sup>1</sup>J<sub>C-F</sub> = 246.4 Hz, **C10**), 144.6 (d, <sup>n</sup>J<sub>C-F</sub> = 7.4 Hz **C6/7/12**), 139.5 (**C6/7/12**), 139.3 (**C6/7/12**), 134.6 (d, <sup>n</sup>J<sub>C-F</sub> = 8.3 Hz, **C14/15/11/13**), 129.6 (d, <sup>n</sup>J<sub>C-F</sub> = 2.7 Hz, **C14/15/11/13**), 120.8 (d, <sup>n</sup>J<sub>C-F</sub> = 21.0 Hz, **C14/15/11/13**), 118.1 (d, <sup>n</sup>J<sub>C-F</sub> = 21.0 Hz, **C14/15/11/13**), 67.5 (**C4**), 61.9 (**C1**), 50.6 (**C5**), 45.4 (**C2**), 42.2 (**C9**), 20.4 (**C8**).

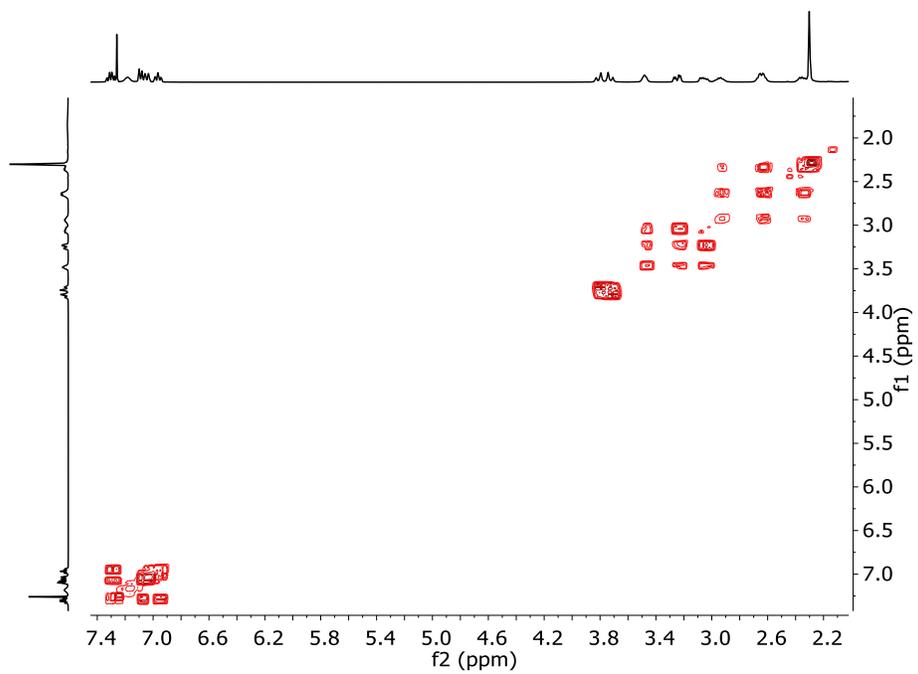
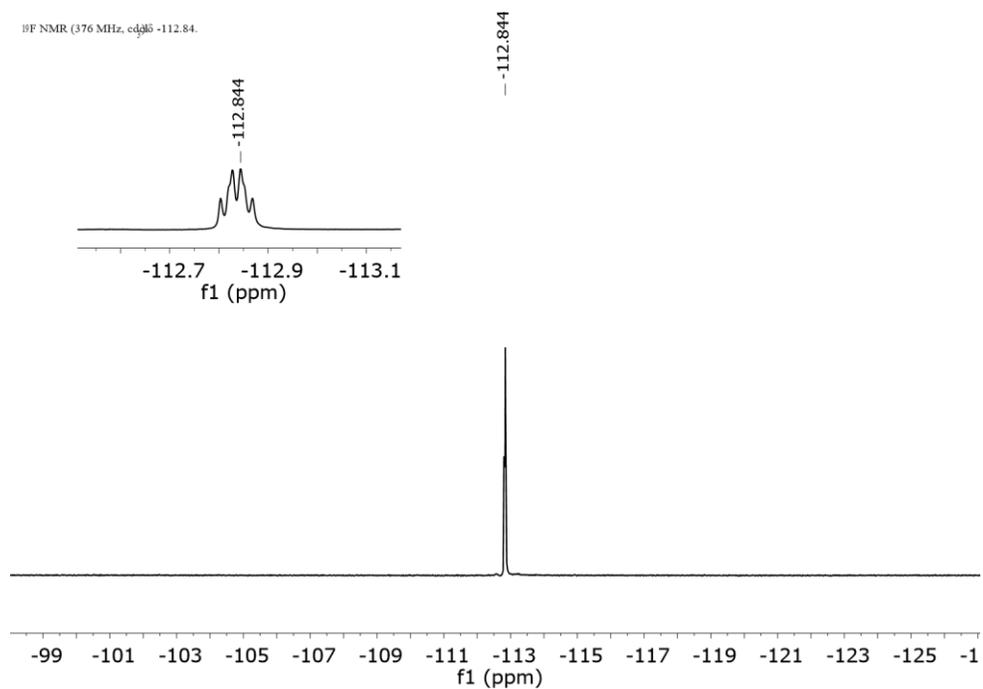
<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -112.8 (m, due to H-F coupling).

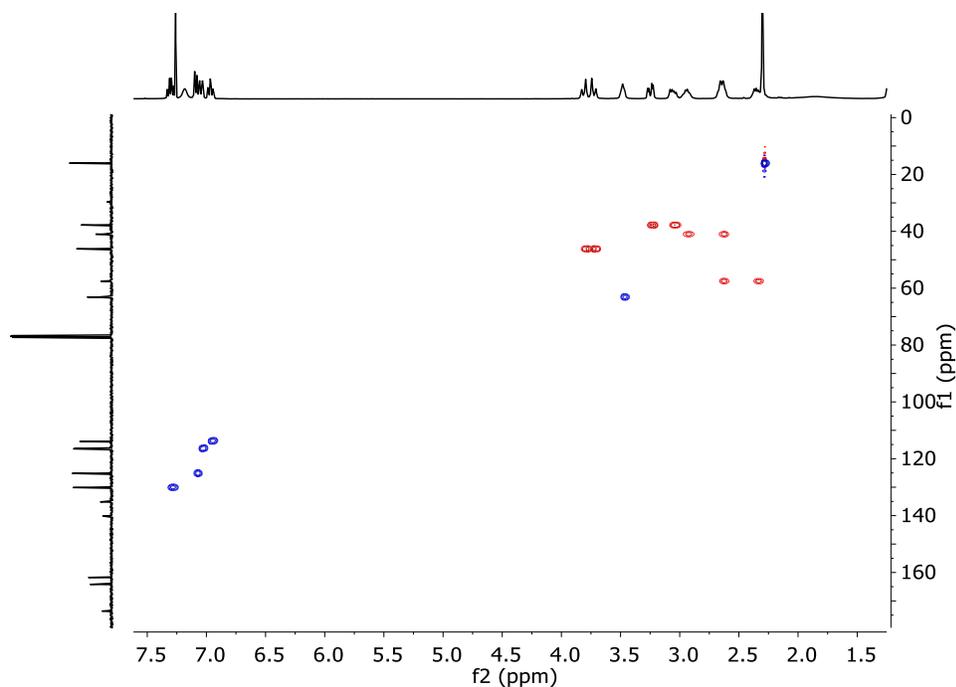
HRMS (ESI-TOF) m/z [**1i** + H]<sup>+</sup> Calc: 798.4313, found: 798.4662.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ ), COSY and HSQC spectra of **1i**:

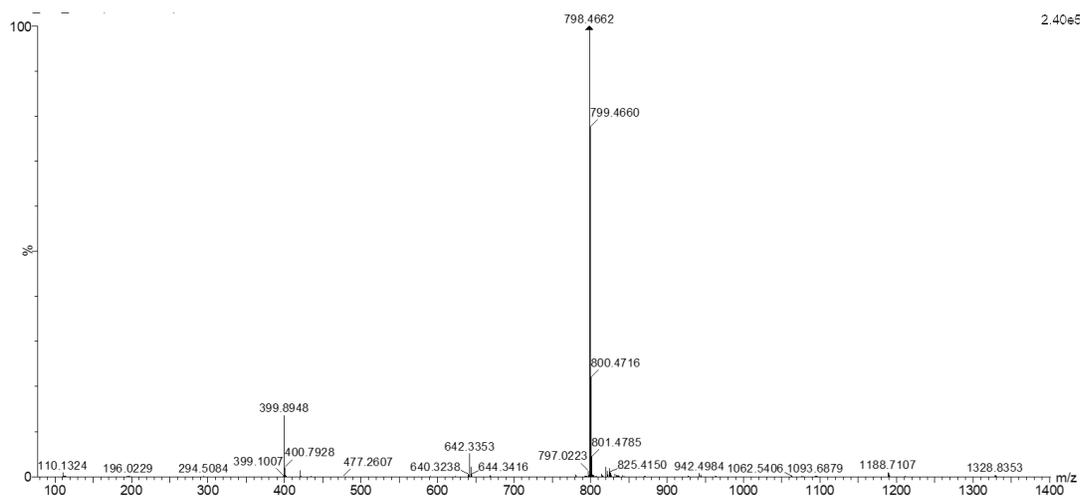


<sup>19</sup>F NMR (376 MHz, cd<sub>3</sub>CO) -112.84





HRMS (ESI+) experimental spectrum of **1h**:



### Synthesis of **1j** (3,5-diF-Phe)

**2j**: Boc-Phe(3,5-F)-OH (254 mg, 0.84 mmol) was dissolved in dry DMF (3 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 196 mg, 1.02 mmol) and 1-Hydroxybenzotriazole hydrate (HOBT, 157 mg, 1.024 mmol), *N,N*-diisopropylethylamine (DIPEA, 0.53 mL, 53.07 mmol) and tris(2-aminoethyl)amine (0.04 mL, 0.25 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was

purified by flash chromatography using 95 : 5 DCM : MeOH to give 0.220 mg of **2j** (0.256 mmol, 87 % yield).

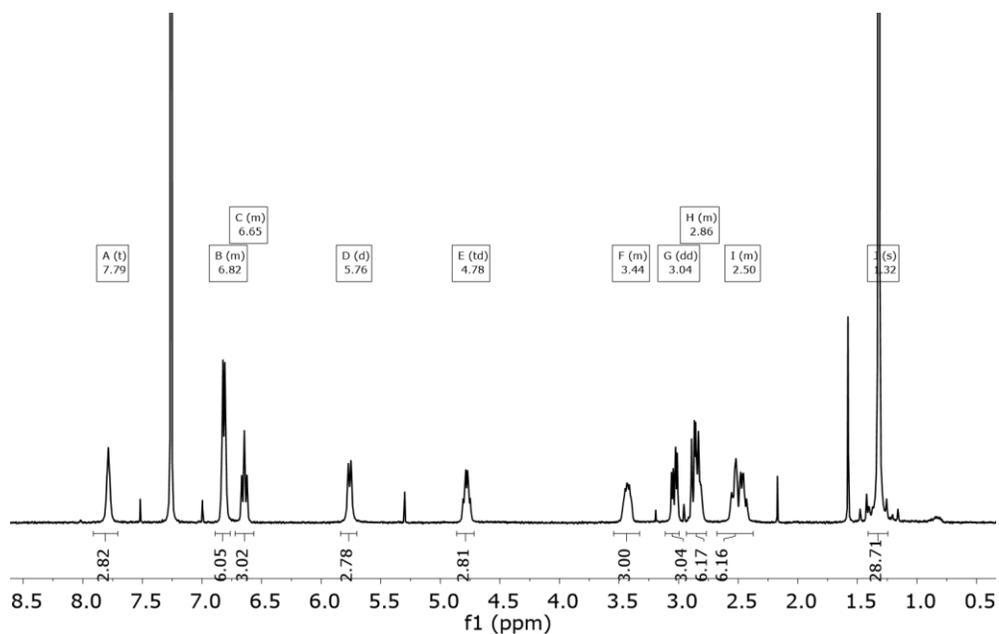
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.79 (t,  $J = 5.1$  Hz, 3H), 6.82 (m, 6H), 6.65 (m, 3H), 5.76 (d,  $J = 9.1$  Hz, 3H), 4.78 (td,  $J = 9.2, 5.8$  Hz, 3H), 3.44 (m, 3H), 3.04 (dd,  $J = 13.6, 5.9$  Hz, 3H), 2.86 (m, 6H), 2.55 (m, 6H), 1.32 (s, 27H).

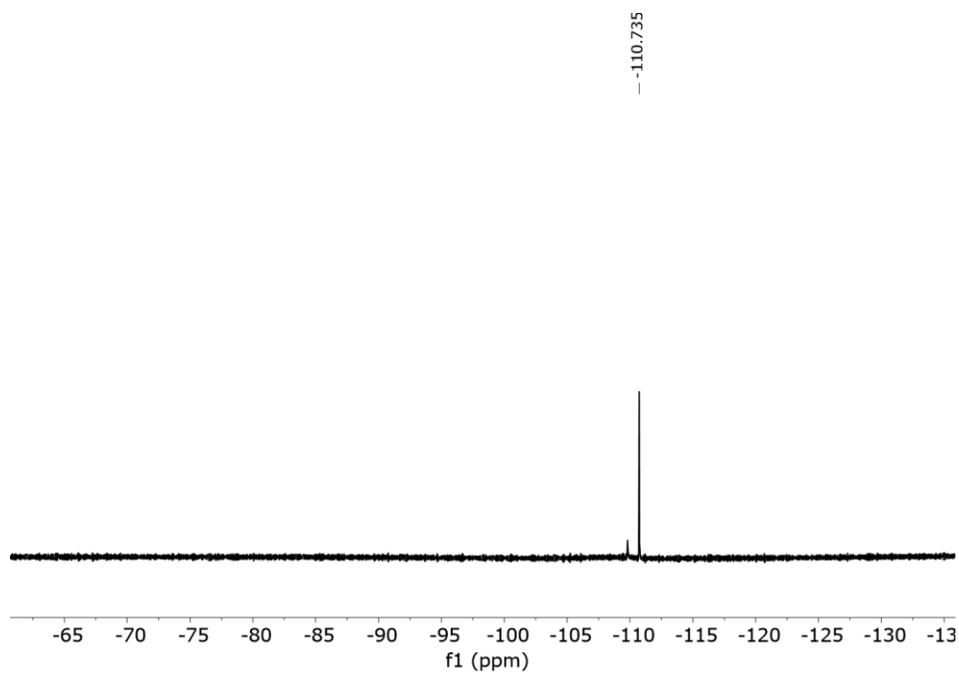
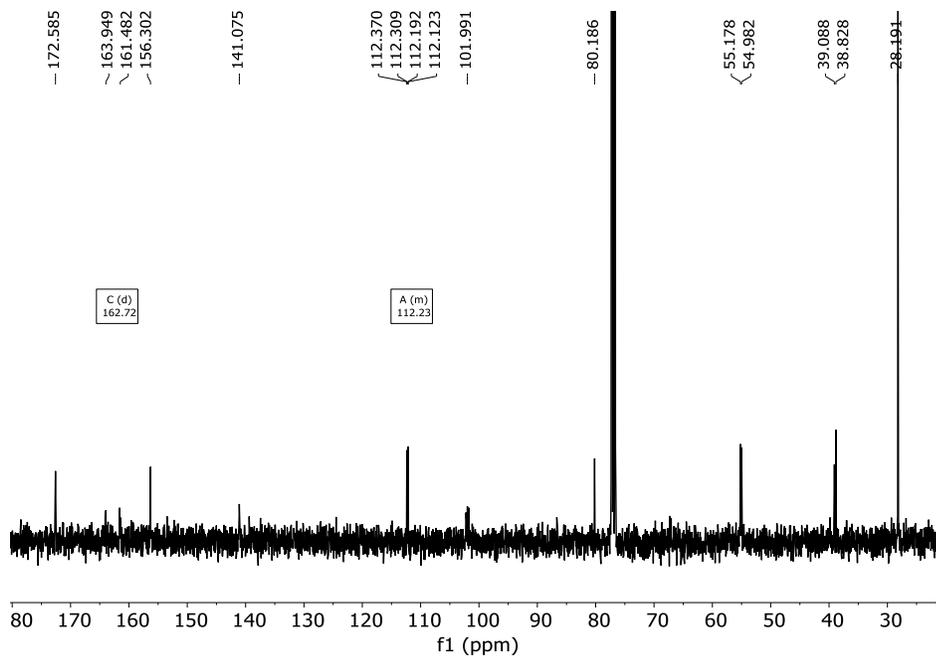
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.6, 162.7. (d,  $^nJ_{\text{C-F}} = 248$  Hz), 156.3, 141.1, 112.4 (m), 101.9 (m), 80.2, 55.2, 55.0, 39.1, 38.8, 28.2.

$^{19}\text{F}$   $\{-^1\text{H}\}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -110.74.

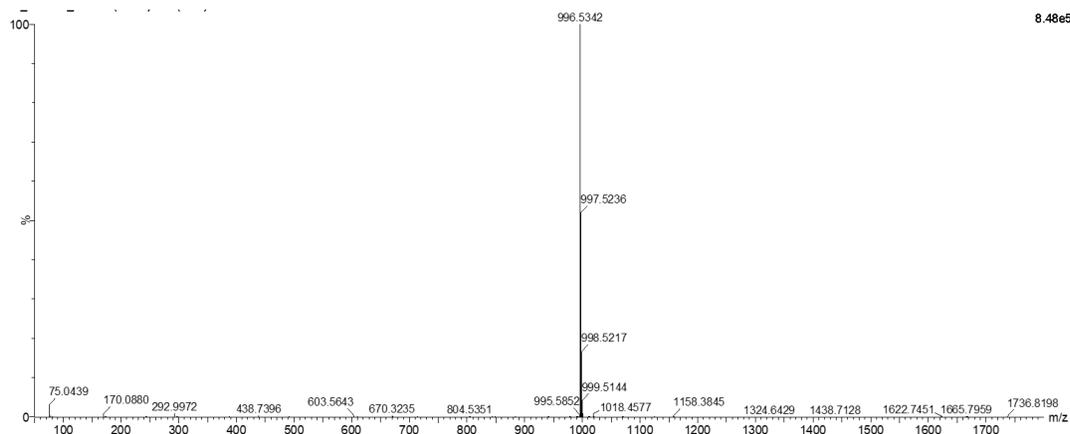
HRMS (ESI-TOF)  $m/z$  [**2j** + H] $^+$  Calc: 996.4664, found: 996.5342

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ) of **2j**:





HRMS (ESI+) experimental spectrum of **2j**:



**3j**: **2j** (220 mg, 0.225 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3j**·**4TFA** as a white solid (279 mg, 0.250 mmol, 91% yield).

TFA was removed by dissolving the solid in 1M NaOH and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>.

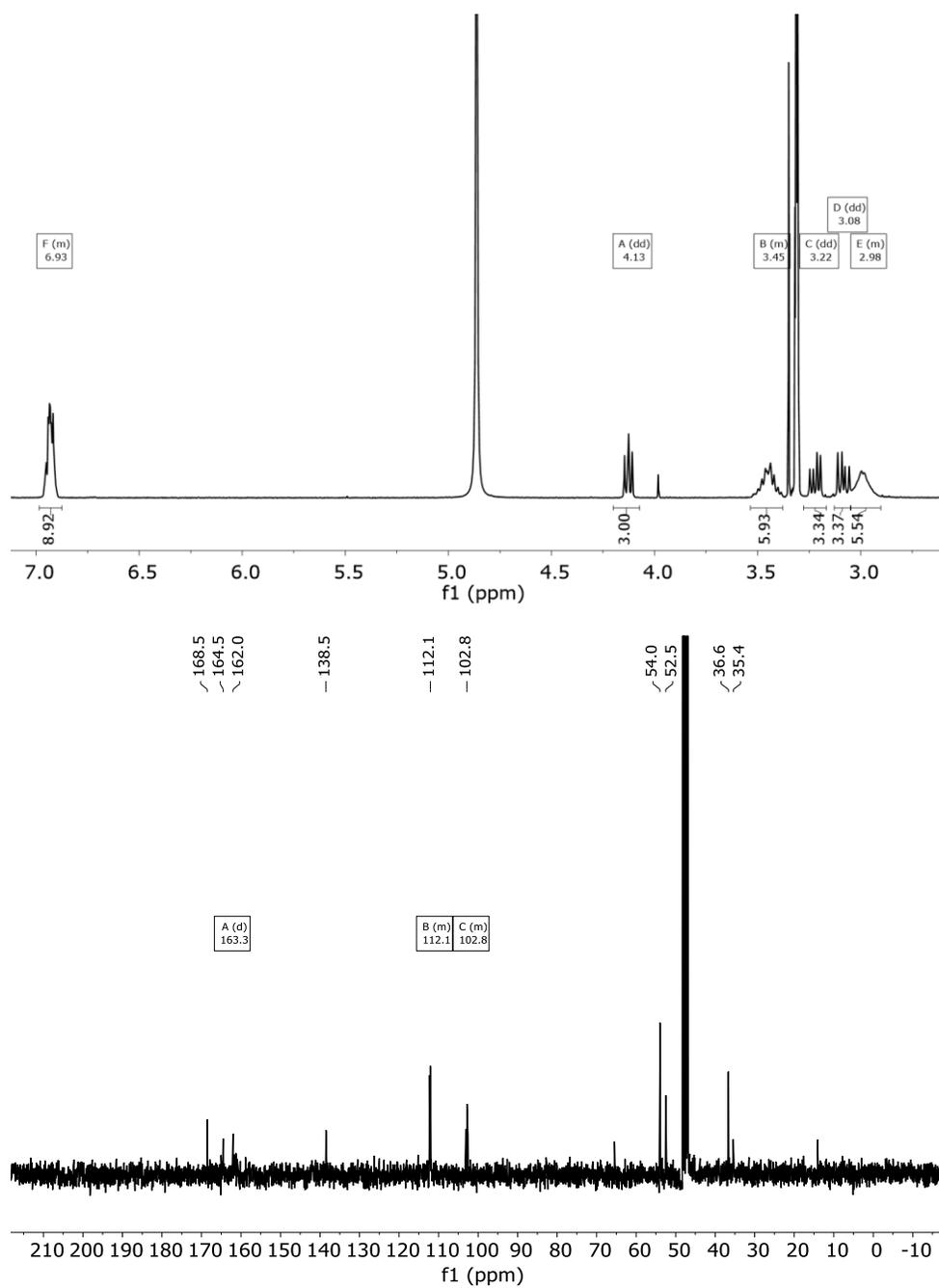
<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 6.93 (m, 9H), 4.13 (m, 3H), 3.45 (m, 6H), 3.22 (dd, *J* = 14.0, 6.7 Hz, 3H), 3.08 (dd, *J* = 14.0, 8.1 Hz, 3H), 2.98 (m, 6H).

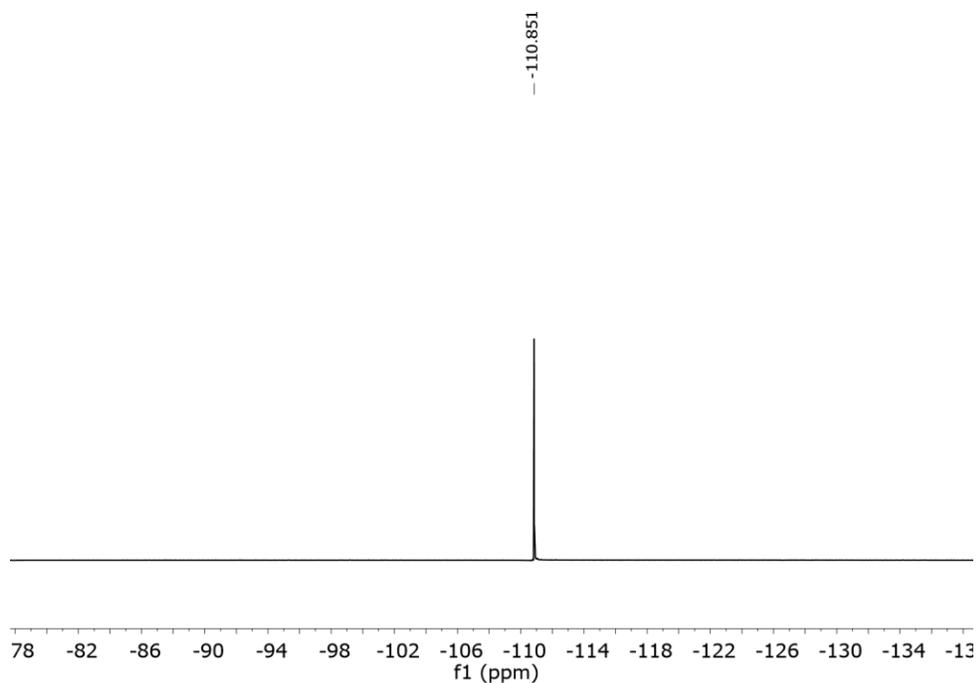
<sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 168.5, 163.3 (d, <sup>1</sup>*J*<sub>C-F</sub> = 248.3 Hz), 138.5 (m), 112.1 (m), 102.8 (m), 54.0, 52.5, 36.6, 35.4.

<sup>19</sup>F NMR—{<sup>1</sup>H}— (376 MHz, CD<sub>3</sub>OD) δ -110.74.

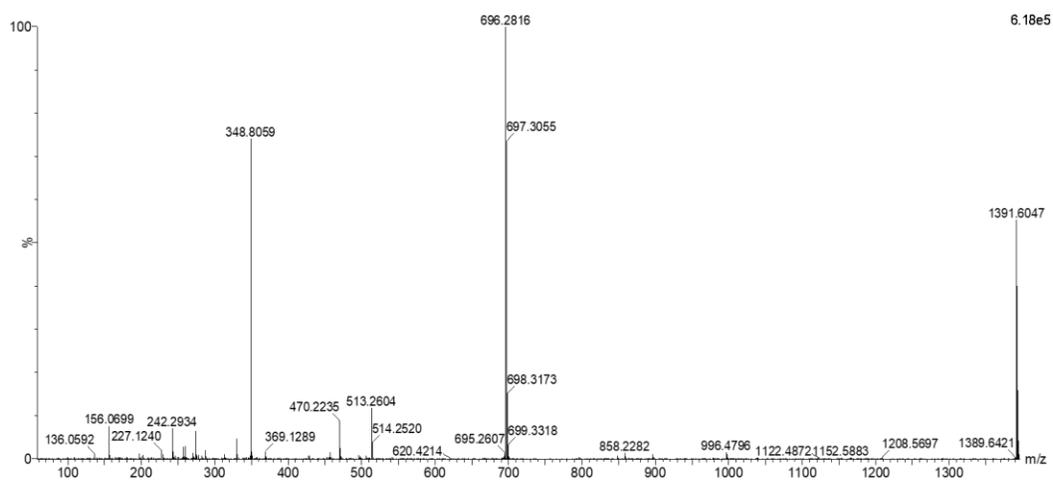
HRMS (ESI-TOF) m/z [**3j** + H]<sup>+</sup> Calc: 696.3091, found: 696.2816.

$^1\text{H-NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ ) and  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CD}_3\text{OD}$ ) of **3j**:

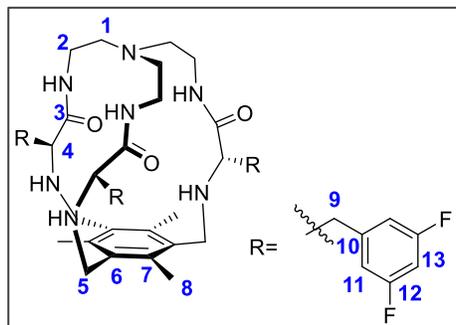




HRMS (ESI+) experimental spectrum of **3j**:



**1j (3,5-diF-Phe): 3j** (120 mg, 0.121 mmol) was dissolved in ACN (40 mL). Tetrabutylammonium chloride (24 mg, 0.087 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (68.8 mg, 0.171 mmol) and potassium carbonate (477 mg, 3.45 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 97 : 3 DCM : MeOH as eluent to give **1j** as a white solid (80 mg, 0.09 mmol, 55 % yield).



$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.18 (t,  $J = 5.4$  Hz, 3H,  $\text{NH}_{\text{amide}}$ ), 6.86 (m, 6H, **H11**), 6.73 ( $\text{tt}_{\text{ap}}$ ,  $J = 8.9, 2.4$  Hz, 3H, **H13**), 3.79 ( $\text{AB}_q$ ,  $\delta_{\text{A}} = 3.84$ ,  $\delta_{\text{B}} = 3.74$ ,  $J_{\text{AB}} = 13.4$  Hz, 6H, **H5**), 3.46 (X subsystem from ABX,  $J_{\text{AX}} = 7.3$ ,  $J_{\text{BX}} = 4.8$  Hz, 3H, **H4**), 3.22 (B subsystem from ABX,  $J_{\text{AB}} = 13.9$ ,  $J_{\text{BX}} = 4.8$  Hz, 3H, **H9**), 3.04 (dd, A subsystem from ABX,  $J_{\text{AX}} = 7.3$ ,  $J_{\text{AB}} = 13.9$  Hz, 3H, **H9**), 2.94 (m, 3H, **H2**), 2.65 (m, 6H, **H2+1**), 2.39 (m, 3H, **H1**), 2.32 (s, 9H, **H8**).

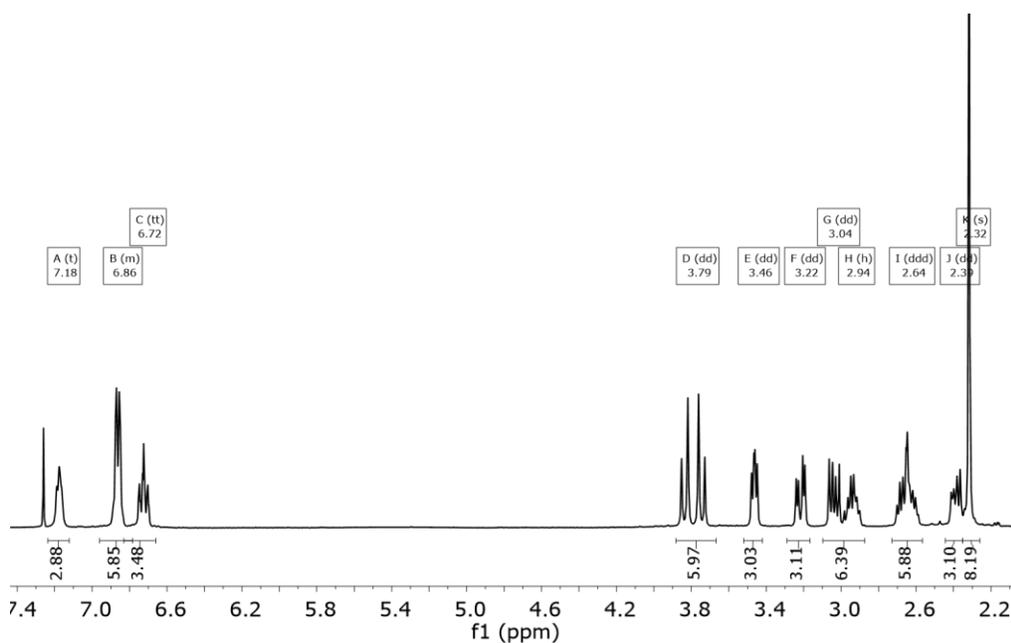
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.6 (**C3**), 163.3 (d,  $^1J_{\text{C-F}} = 249.0$  Hz,

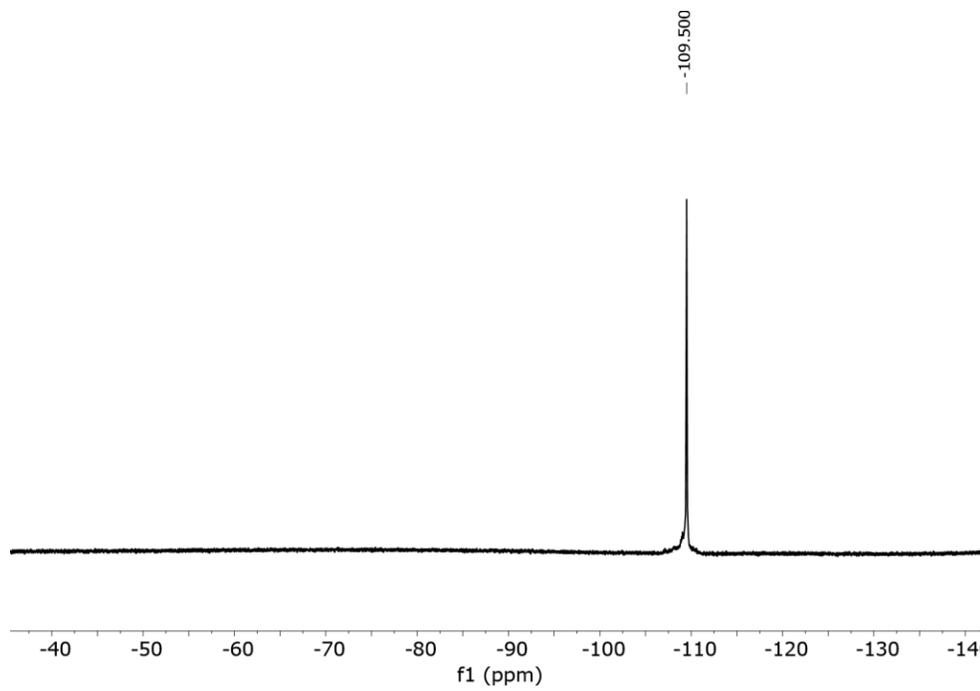
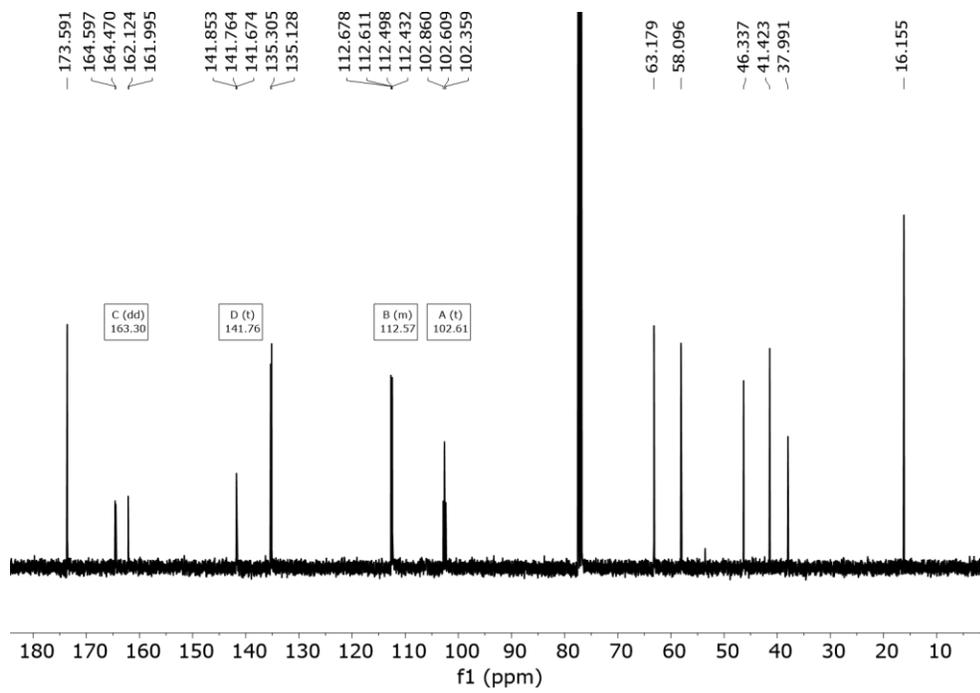
**C12**), 141.8 (t,  $^nJ_{\text{C-F}} = 9$  Hz, **C10**), 135.3 (**C6/7**), 135.1 (**C6/7**), 112.4 (m, **C11**), 102.6 (t,  $^nJ_{\text{C-F}} = 25.2$  Hz, **C13**), 63.2 (**C4**), 58.1 (**C1**), 46.3 (**C5**), 41.4 (**C2**), 38.0 (**C9**), 16.2 (**C8**).

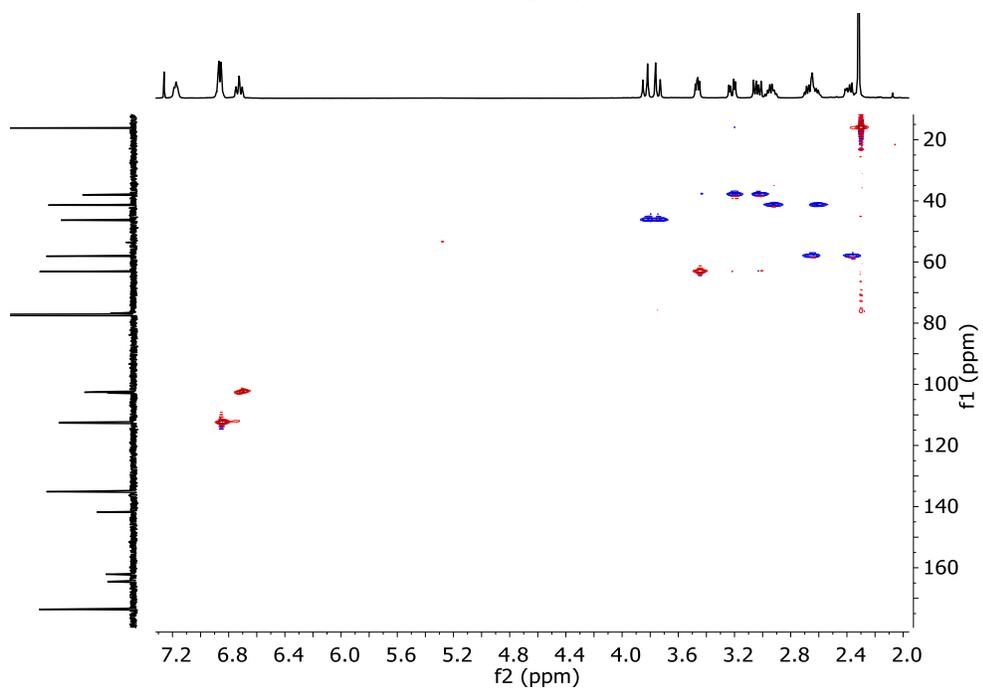
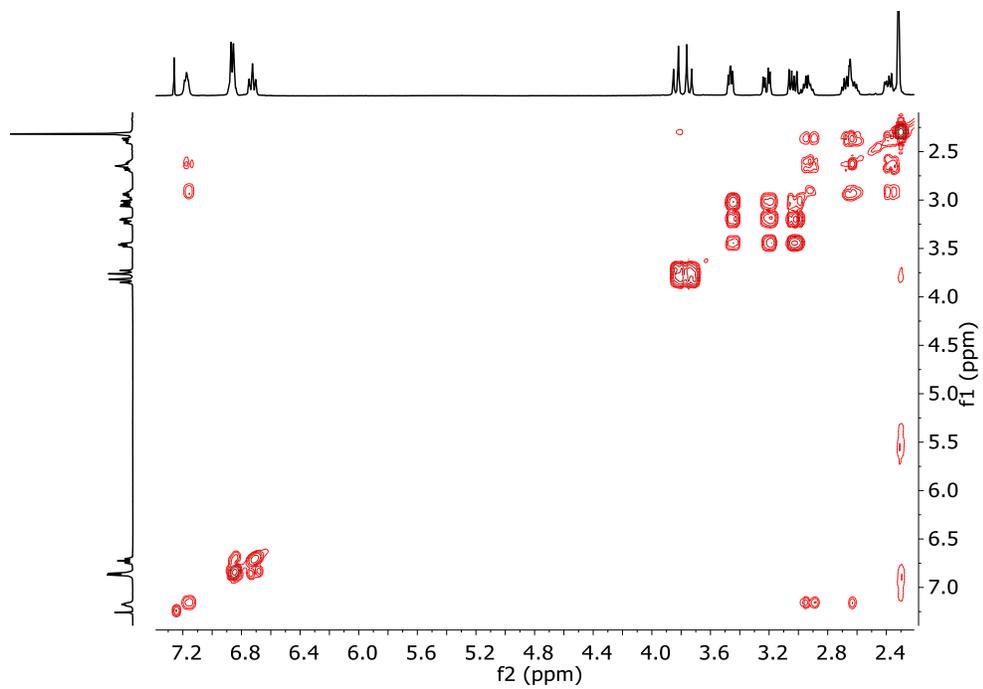
$^{19}\text{F}$  NMR- $\{^1\text{H}\}$ - (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -109.50.

HRMS (ESI-TOF)  $m/z$  [**1c** +  $\text{H}$ ] $^+$  Calc: 852.4030, found: 852.4062.

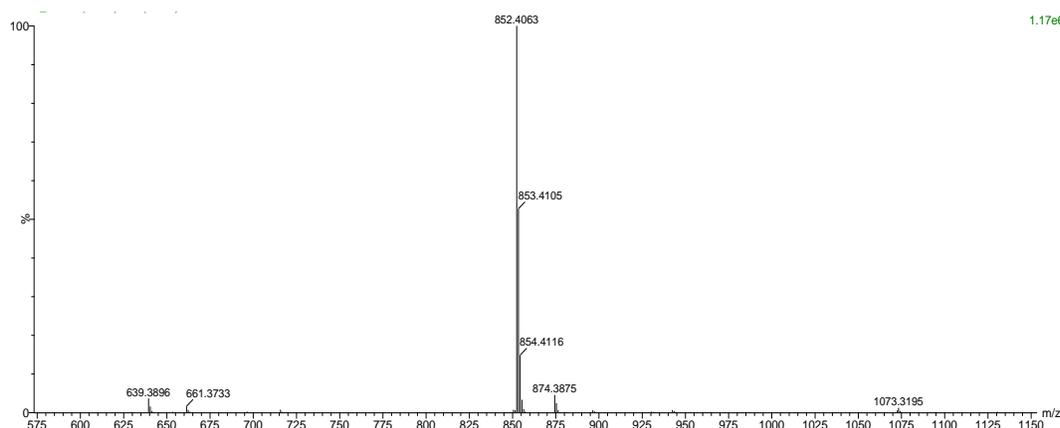
$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ), COSY, and HSQC spectra of **1j**:







HRMS (ESI+) experimental spectrum of **1j**:



### Synthesis of 1k (3,4-diF-Phe)

**2k**: Boc-Phe(3,4-F)-OH (445 mg, 1.48 mmol) was dissolved in dry DMF (4 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 343 mg, 1.793 mmol) and 1-hydroxybenzotriazole hydrate (HOBt, 274 mg, 1.79 mmol), *N,N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.37 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.448 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was purified by flash chromatography using 97 : 3 DCM : MeOH as eluent to give 0.320 mg of **2k** (0.321 mmol, 72% yield).

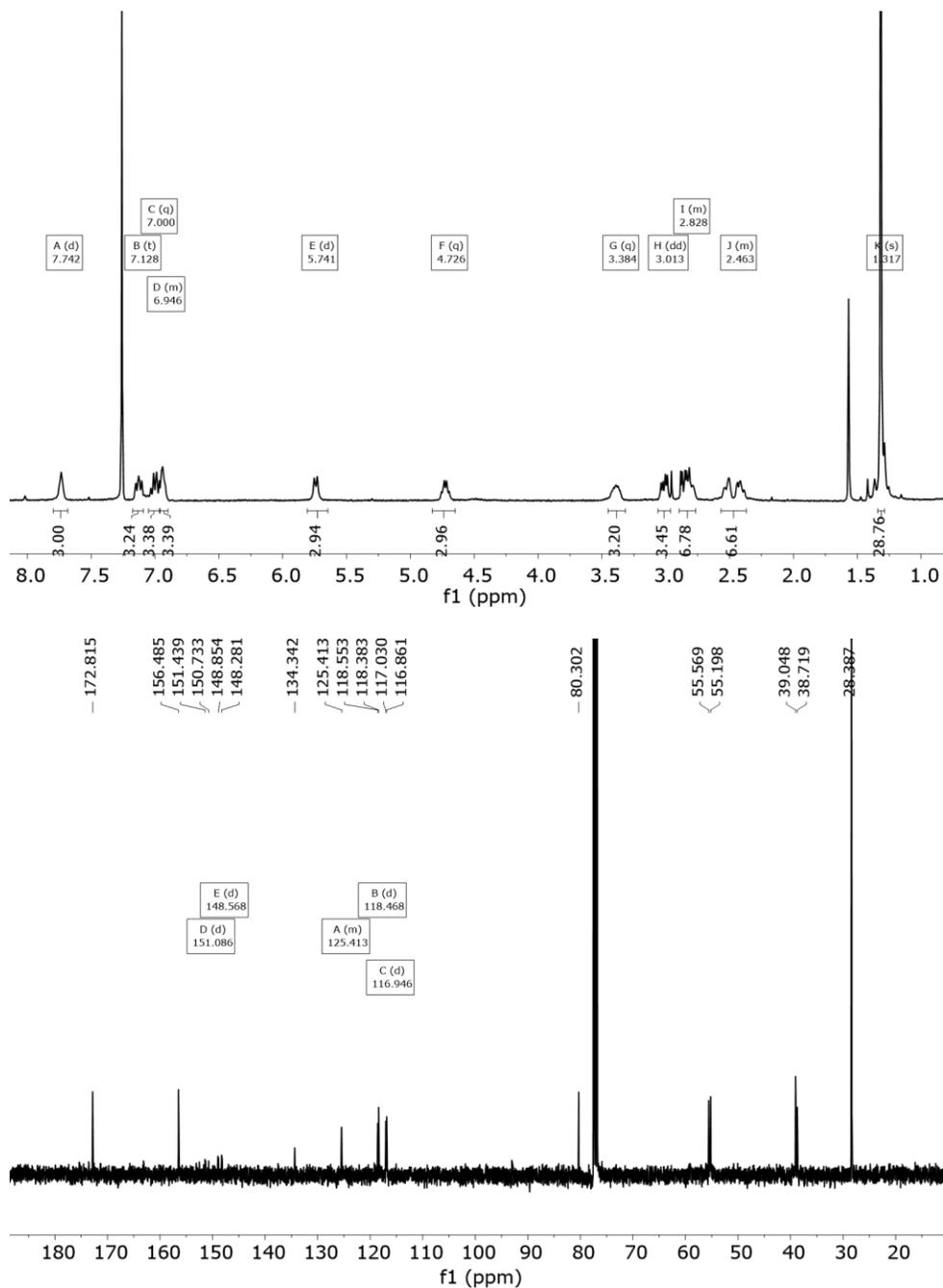
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.74 (d, *J* = 5.2 Hz, 3H), 7.13 (t, *J* = 9.6 Hz, 3H), 7.00 (q, *J* = 8.9 Hz, 3H), 6.94 (m, 3H), 5.74 (d, *J* = 9.1 Hz, 3H), 4.73 (q, *J* = 8.4 Hz, 3H), 3.38 (dd, *J* = 7.5, 6.6 Hz, 3H), 3.01 (dd, *J* = 13.7, 6.2 Hz, 3H), 2.83 (m, 6H), 2.46 (m, 6H), 1.32 (s, 27H).

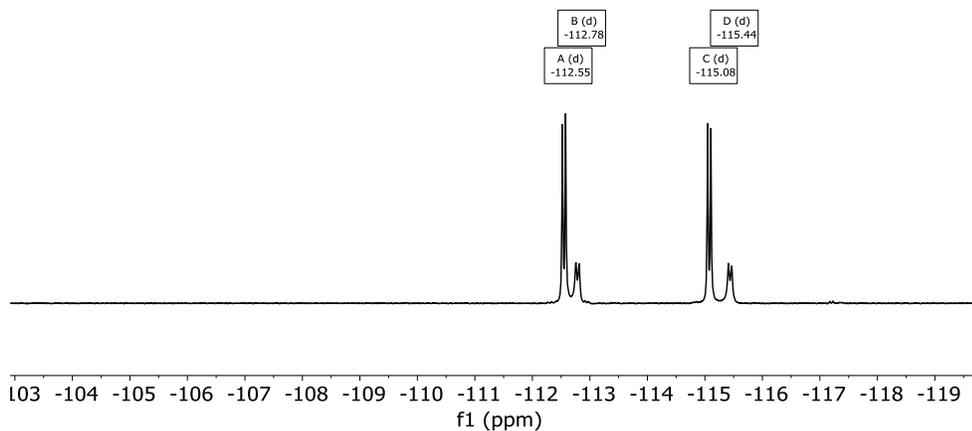
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.8, 156.5, 151-148 (2F, m), 125.5 (m), 118.4 (d, <sup>n</sup>*J*<sub>C-F</sub> = 17.1), 117.0 (d, <sup>n</sup>*J*<sub>C-F</sub> = 17.1), 116.9, 80.3, 55.6, 55.2, 39.1, 38.7, 28.4.

<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -112.55/-112.78 (d, *J* = 21.1 Hz), -115.08/-117.44 (d, *J* = 21.1 Hz). Each F gives two signals of different intensity due to the presence of rotamers.

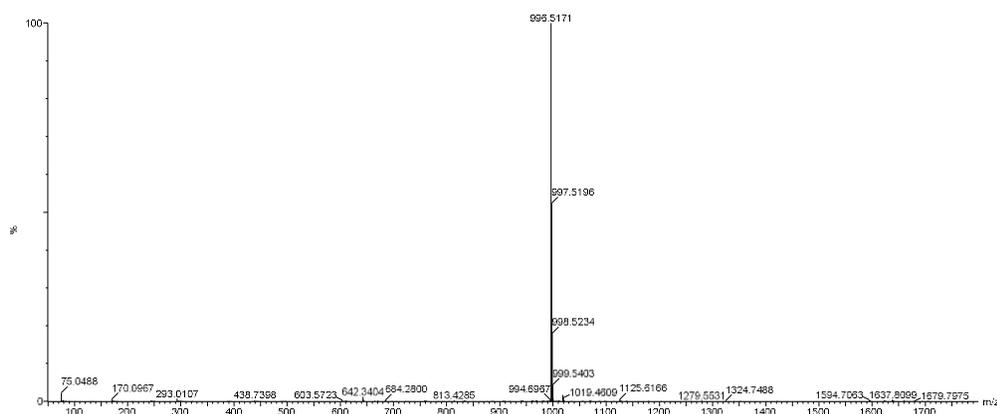
HRMS (ESI-TOF) *m/z* [**2k**+ H]<sup>+</sup> Calc: 996.4664; found: 996.5171.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ ) of **2k**:





HRMS (ESI+) experimental spectrum of **2k**:



**3k: 2k** (250 mg, 0.251 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3k-4TFA** as a white solid (360 mg, 0.35 mmol, 90% yield). TFA was removed by dissolving the solid in NaOH 1M and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>.

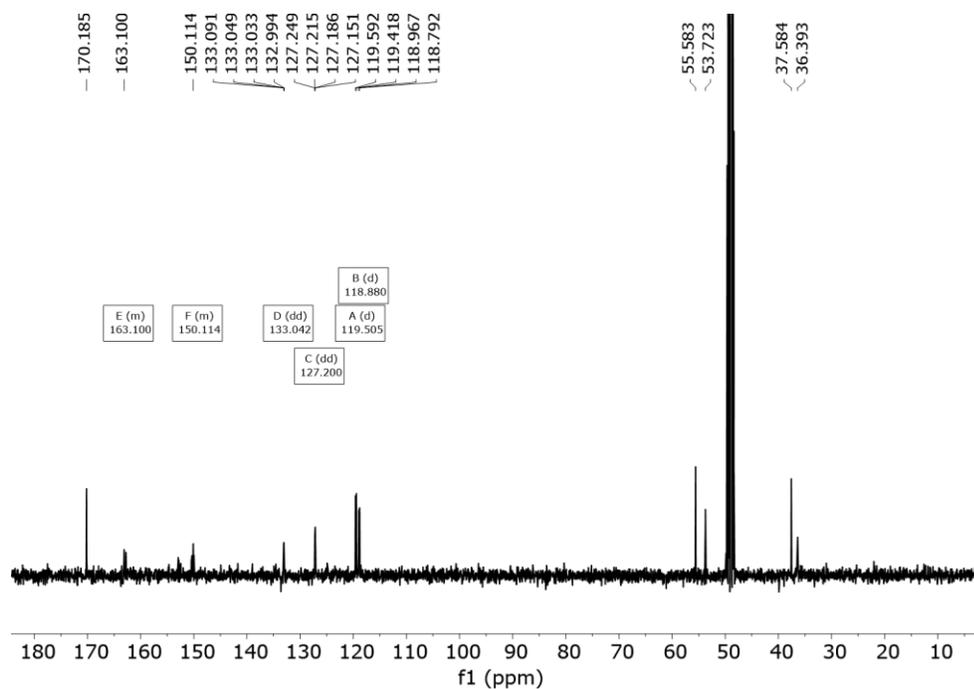
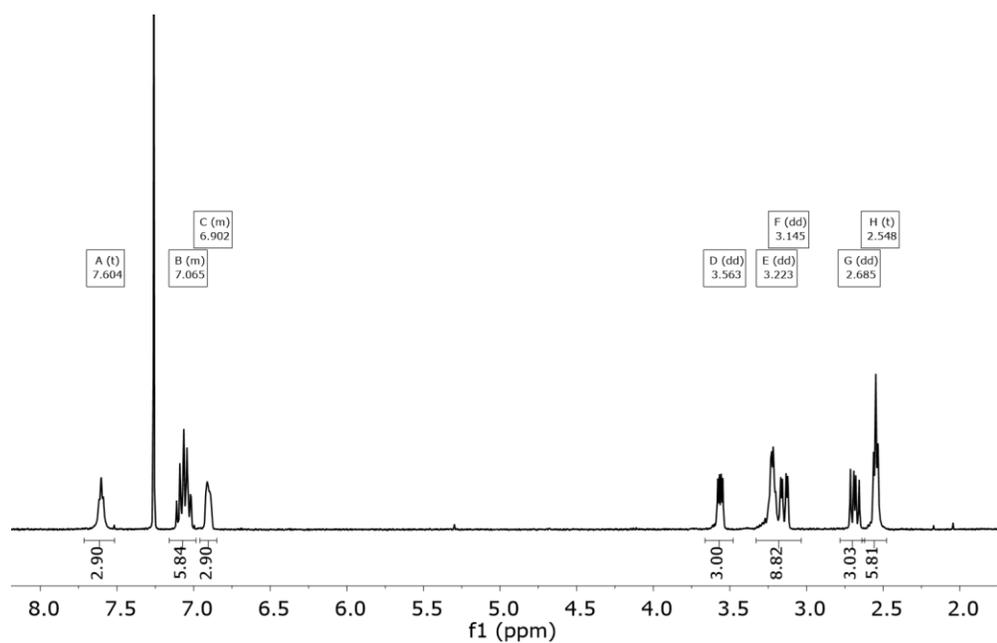
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.60 (m, 3H), 7.07 (m, 6H), 6.90 (m, 3H), 3.56 (X subsystem from ABX,  $J_{AX} = 8.9$ ,  $J_{BX} = 4.3$  Hz, 3H), 3.22 (A subsystem from ABX,  $J_{AX} = 8.9$ ,  $J_{AB} = 13.8$  Hz, 3H), 3.15 (B subsystem from ABX,  $J_{BX} = 4.3$ ,  $J_{AB} = 13.8$  Hz, 3H), 2.69 (dd<sub>ap</sub>,  $J = 13.8$ , 8.9 Hz, 3H), 2.55 (t<sub>ap</sub>,  $J = 5.7$  Hz, 6H).

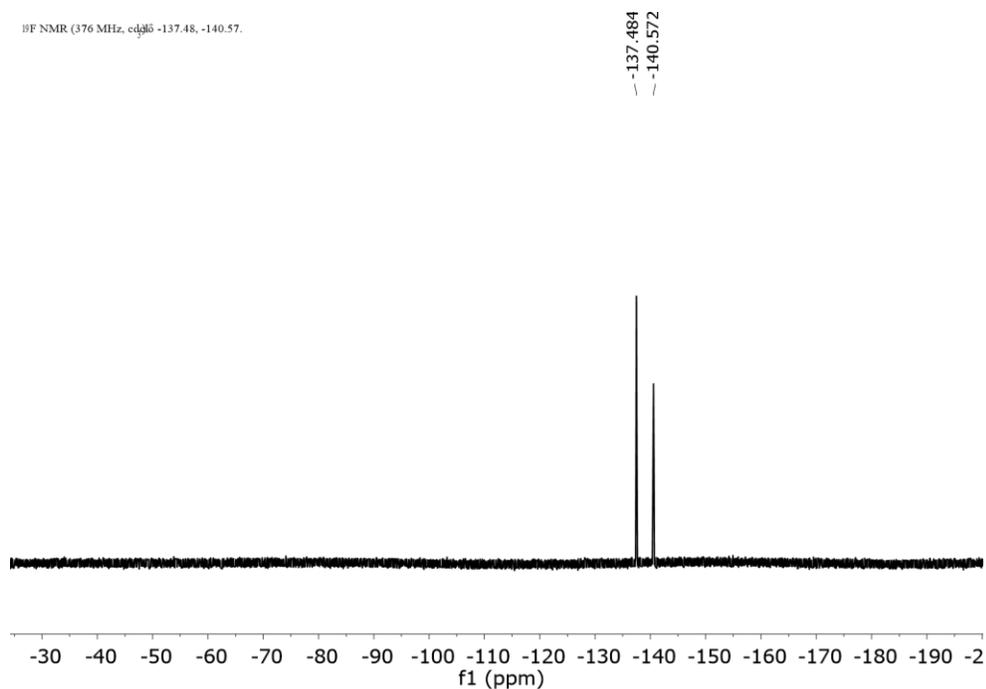
<sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 170.2, 163.1 (C-F, m), 150.1 (C-F, m), 133.1 (dd,  $^nJ_{C-F} = 5.7$ , 4.0 Hz), 127.3 (dd,  $^nJ_{C-F} = 6.4$ , 3.4 Hz), 119.6 (d,  $^nJ_{C-F} = 17.6$  Hz), 118.8 (d,  $^nJ_{C-F} = 17.6$  Hz), 55.6, 53.7, 37.6, 36.4.

<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -137.48, -140.57.

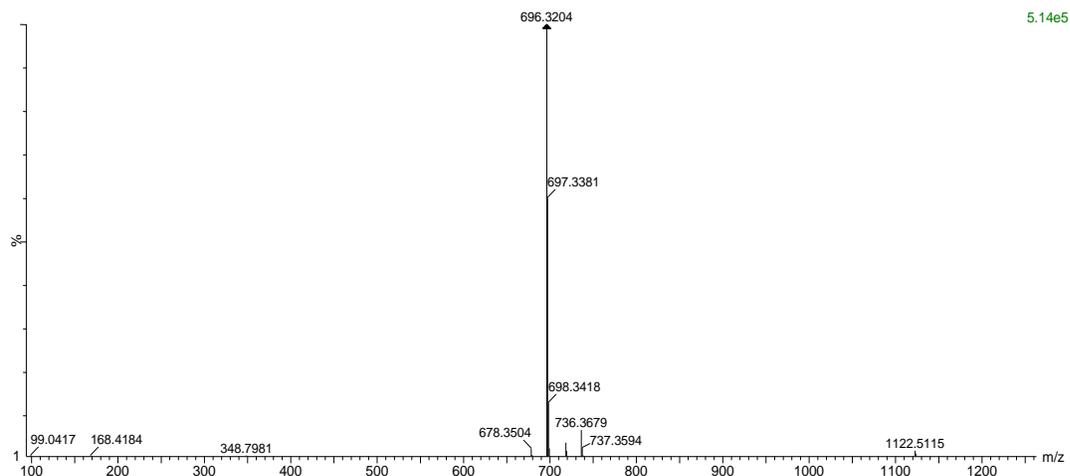
HRMS (ESI-TOF) m/z [**3k** + H]<sup>+</sup> Calc: 696.3091, found: 696.3204.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ ) and  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ ) of **3k**:

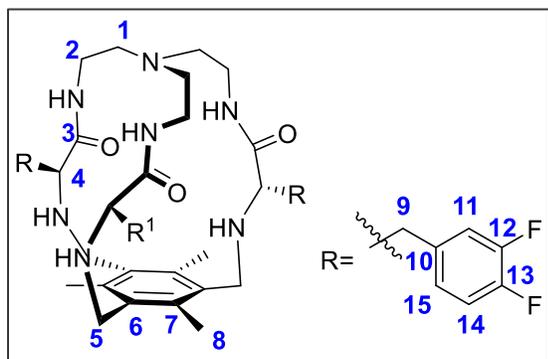




HRMS (ESI+) experimental spectrum of **3k**:



**1k (3,4-diF-Phe): 3k** (112 mg, 0.113 mmol) was dissolved in ACN (40 mL). Tetrabutylammonium chloride (22.4 mg, 0.080 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (64.3 mg, 0.160 mmol) and potassium carbonate (445 mg, 3.22 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 97 : 3 DCM : MeOH as eluent to give **1k** as a white solid (75 mg, 0.09 mmol, 55 % yield).



$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.15 (m, 9H, (H**11**,**14**,**15**), 7.03 (m, 3H, ( $\text{NH}_{\text{amide}}$ ), 3.78 ((AB<sub>q</sub>,  $\delta_{\text{A}}=3.81$ ,  $\delta_{\text{B}}=3.72$ ,  $J_{\text{AB}}=13.4$  Hz, 6H, H**5**)), 3.44 ((X subsystem from ABX,  $J_{\text{AX}}=4.7$ ,  $J_{\text{BX}}=7.2$  Hz, 3H (H**4**)), 3.19 (A subsystem from ABX,  $J_{\text{AX}}=4.7$ ,  $J_{\text{AB}}=14.0$  Hz, 3H, (H**9**)), 2.97 (B subsystem from ABX,  $J_{\text{BX}}=7.2$ ,  $J_{\text{AB}}=14.0$  Hz, 3H, (H**9**) + m, 3H (H**2**)), 2.65 (m, 6H, (H**2+1**)), 2.37 (m, 3H, (H**1**)), 2.32 (s, 9H, (H**8**)).

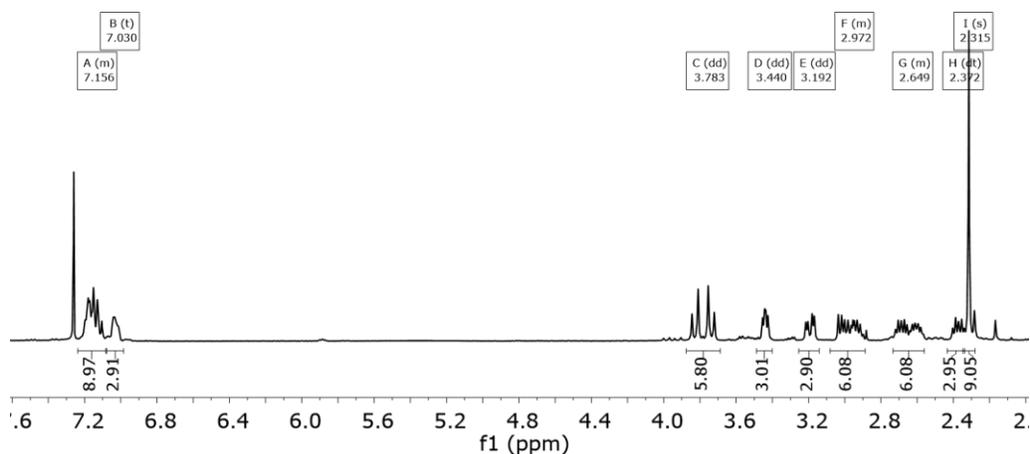
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.8 (C**3**), 160.4 (C-F, m, C**12**/**13**), 148.3 (C-F, m, C**12**/**13**), 135.3 (C**6**/**7**), 135.2

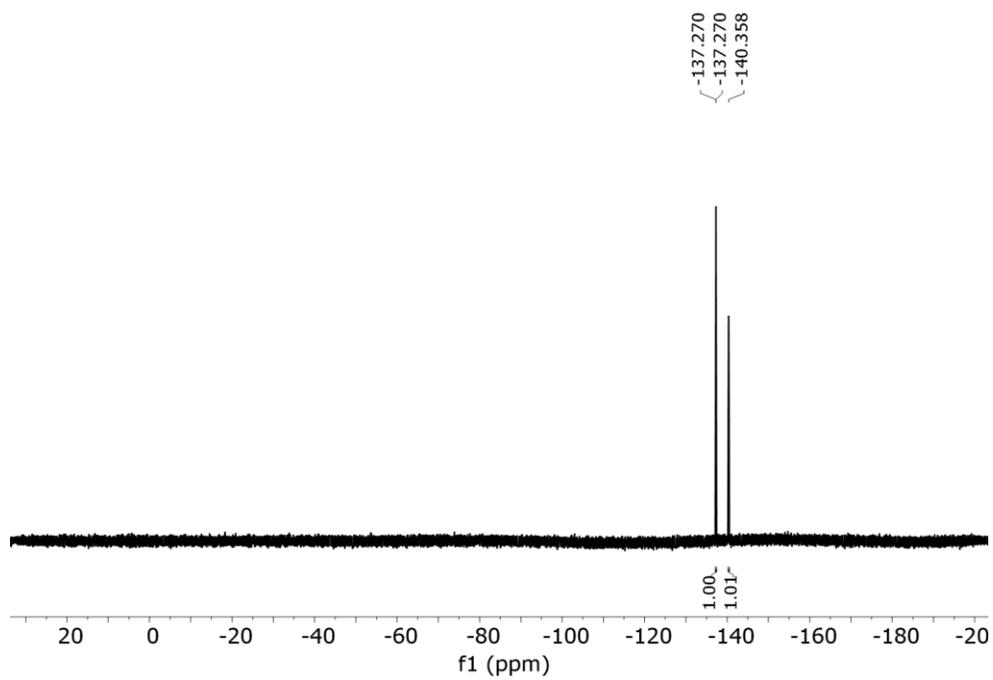
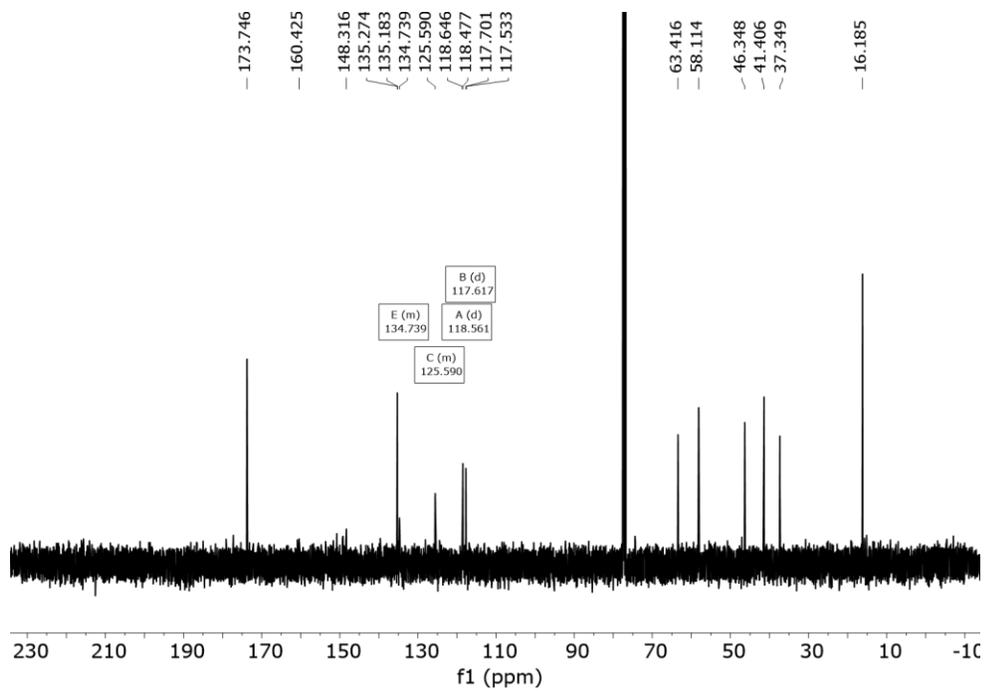
(C**6**/**7**), 134.8 (m, C**15**), 125.6 (m, C**10**), 118.6 (d,  $^nJ_{\text{C-F}}=17$  Hz, C**11**/**14**), 117.6 (d,  $^nJ_{\text{C-F}}=17$  Hz, C**11**/**14**), 63.4 (C**4**), 58.1 (C**1**), 46.4 (C**5**), 41.4 (C**2**), 37.4 (C**9**), 16.2 (C**8**).

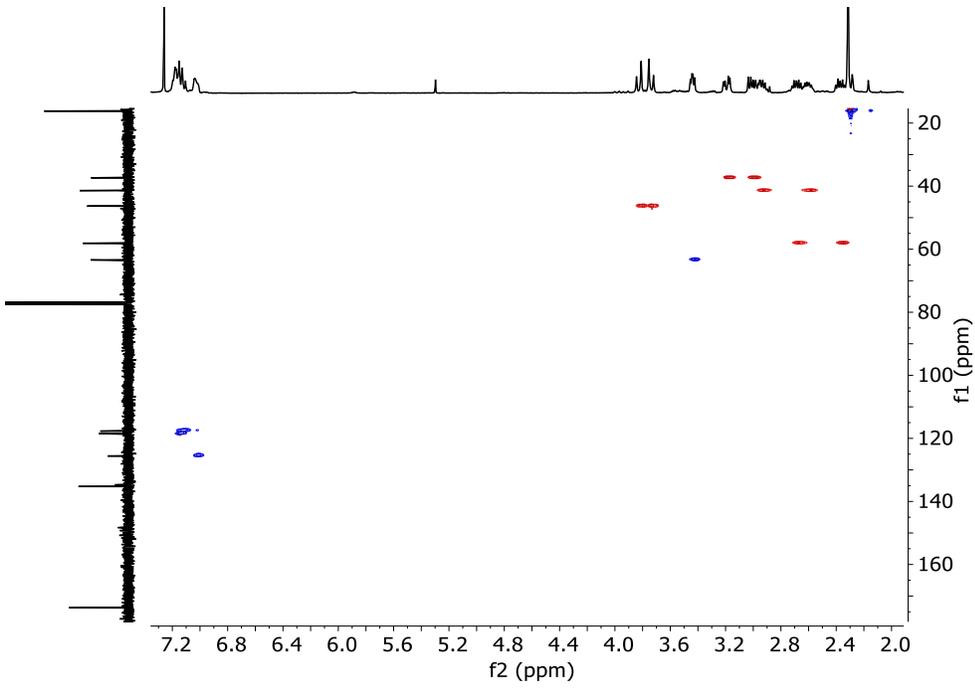
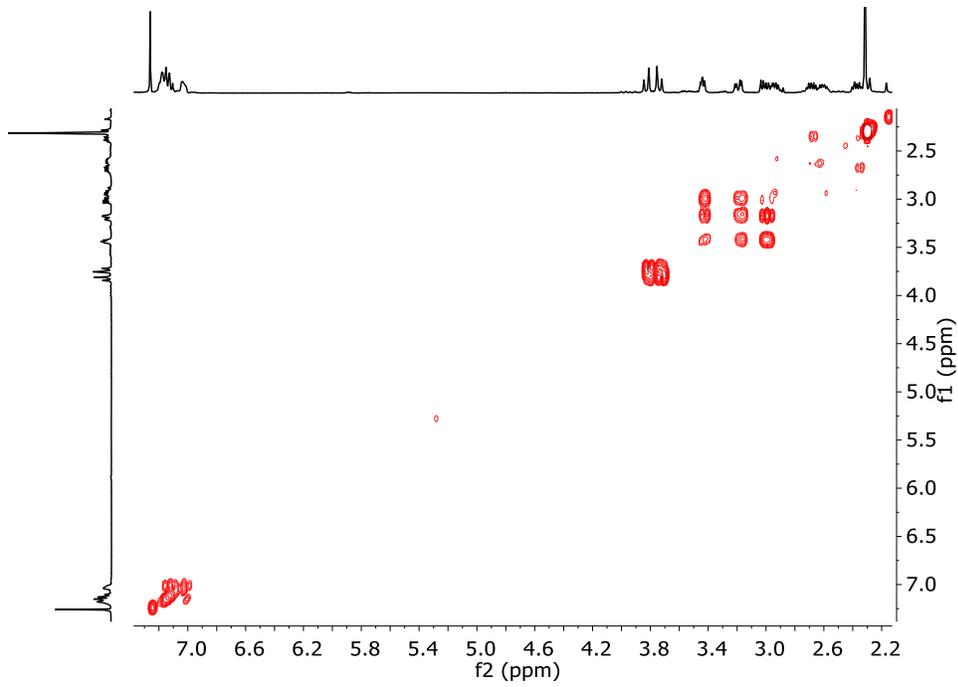
$^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -137.27, -140.36.

HRMS (ESI-TOF)  $m/z$  [**1k**+ 2H]<sup>+</sup>/2 Calc: 852.9469, found: 852.4105.

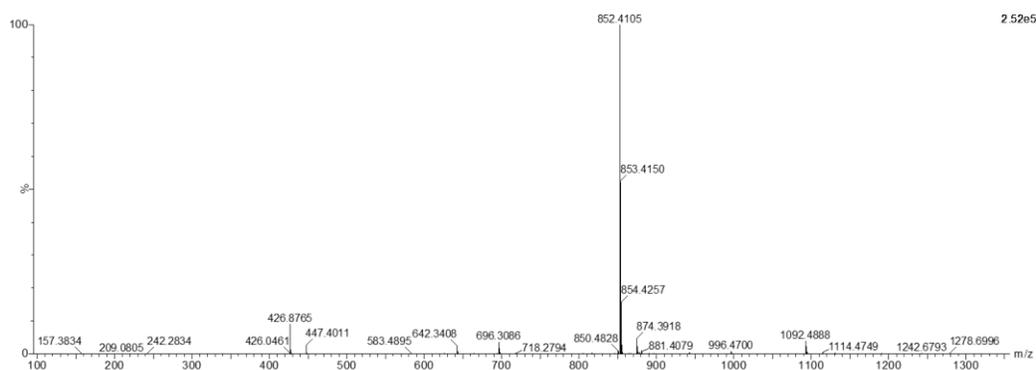
$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ), COSY and HSQC of **1k**:







HRMS (ESI+) experimental spectrum of **1k**:



### **Synthesis of 1l (2,6-diF-Phe)**

**Boc (2,6F-Phe)-OH:** To a solution of 2,6-difluoro-phenylalanine-HCl (0.5 g, 2.1 mmol) in THF/H<sub>2</sub>O 1:1 (8.5 mL) was added di-tert-butyl dicarbonate (0.5 g, 2.3 mmol) and sodium hydroxide (278 mg, 6.95 mmol). The reaction was stirred at room temperature for 16 hours. The THF was removed in vacuum and DCM was added to the reaction flask. 10% HCl was added drop wise to this solution while stirring vigorously until the precipitate ceased (pH  $\approx$  4). The reaction mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3x150mL). The organic layer was then dried over anhydrous sodium sulphate (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuum to afford **Boc (2,6F-Phe)-OH** as a white solid (605.8 mg, 2.01 mmol, 96 % yield).

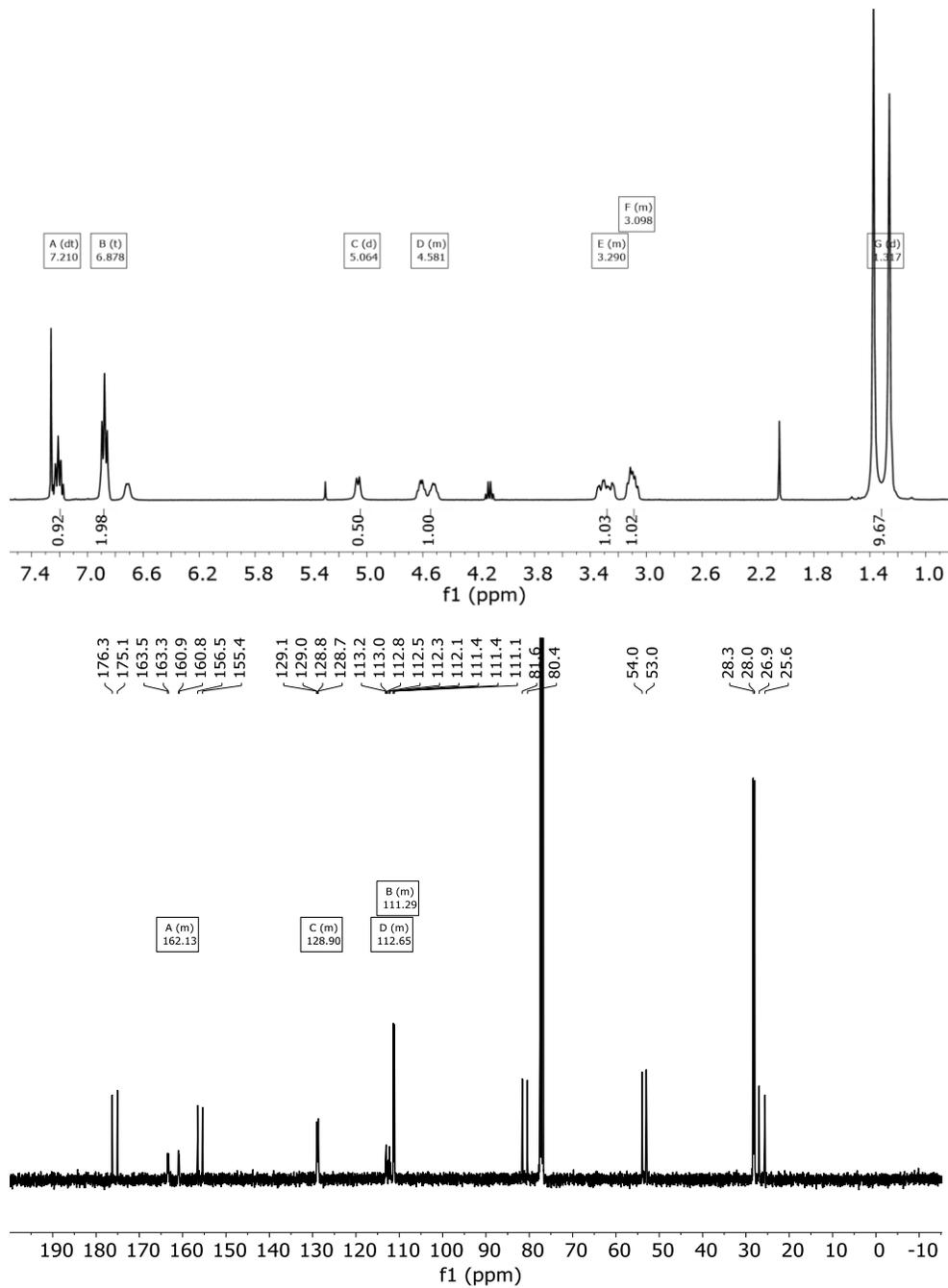
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.21 (m, 3H), 6.88 (t<sub>ap</sub>,  $J$  = 7.7 Hz, 3H), 5.06 (d,  $J$  = 8.5 Hz, 3H), 4.68 (m, 3H), 3.29 (m, 3H), 3.10 (m, 3H), 1.32 (m, 9H)

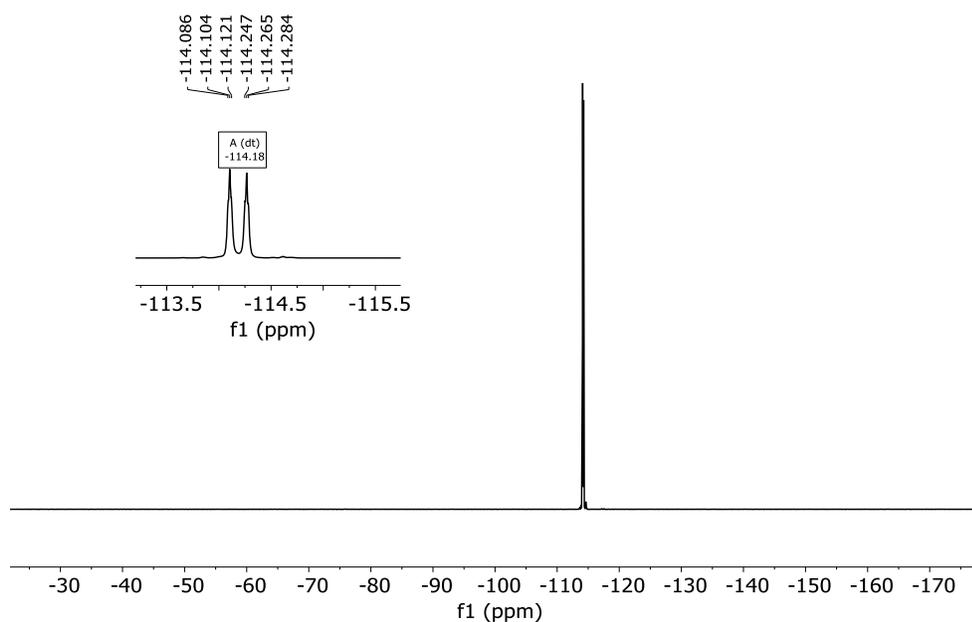
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) (most carbons give two signals due to the presence of rotamers in solution)  $\delta$  176.3/175.1, 163.4/160.9 (m, F-C<sub>Arr</sub> coupling), 156.5/155.4, 128.9 (m, F-C<sub>Arr</sub> coupling), 112.7 (m, F-C<sub>Arr</sub> coupling), 111.4/111.1 (m, F-C<sub>Arr</sub> coupling), 81.6/80.4, 54.0/53.0, 28.3/28.0, 29.9/25.6.

<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>)  $\delta$  -114.18 (dt,  $J$  = 61.2, 6.8 Hz). Spectra multiplicity due to H coupling.

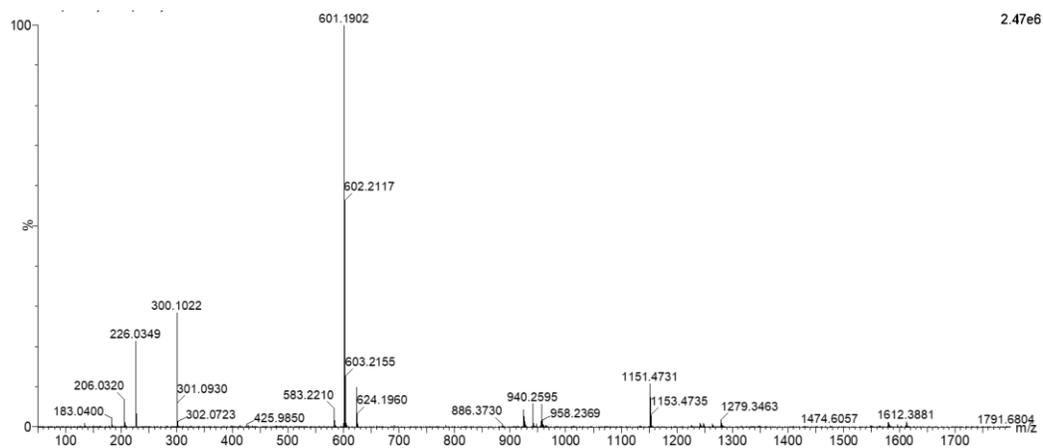
HRMS (ESI-TOF) m/z [Boc(2,6F-Phe)-OH - H]<sup>-</sup> Calc: 300.1053, found: 300.1022; also [2(Boc(2,6F-Phe)-OH) - H]<sup>-</sup> Calc: 601.2178, found: 601.1902.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F NMR}$  (376 MHz,  $\text{CDCl}_3$ ) spectra of **Boc (2,6F-Phe)-OH**:





HRMS (ESI-) experimental spectrum of **Boc (2,6F-Phe)-OH**:



**2I:** The protected amino acid (509 mg, 1.69 mmol) was dissolved in dry DMF (3 mL). In a separate flask, EDC·HCl (393 mg, 2.05 mmol) and HOBT (313 mg, 2.05 mmol) were dissolved in dry DMF and added to the amino acid solution and diisopropylethylamine (DIPEA, 1 mL, 6.15 mmol) was also added to the solution. The reaction mixture was stirred under nitrogen at room temperature for 10 minutes to allow the activation of the acid. Then, tris(2-aminoethyl)amine (0.08 mL, 0.512 mmol) was added. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5 % w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The

residue was purified by flash chromatography using 97 : 3 DCM : MeOH as eluent to give 0.350 mg of **21** (0.351 mmol, 69% yield).

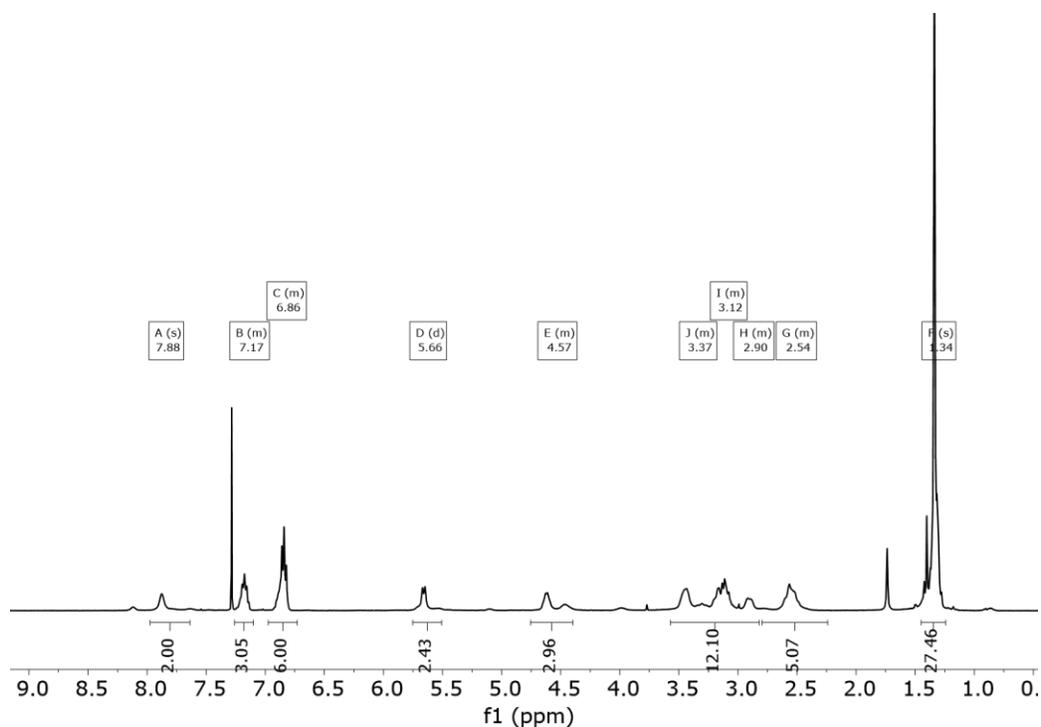
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.88 (s, 3H), 7.17 (m, 3H), 6.86 (m, 6H), 5.66 (d,  $J = 8.8$  Hz, 3H), 4.57 (m, 3H), 3.37 (m, 3H), 3.12 (m, 6H), 2.90 (m, 3H), 2.54 (m, 6H), 1.34 (s, 27H).

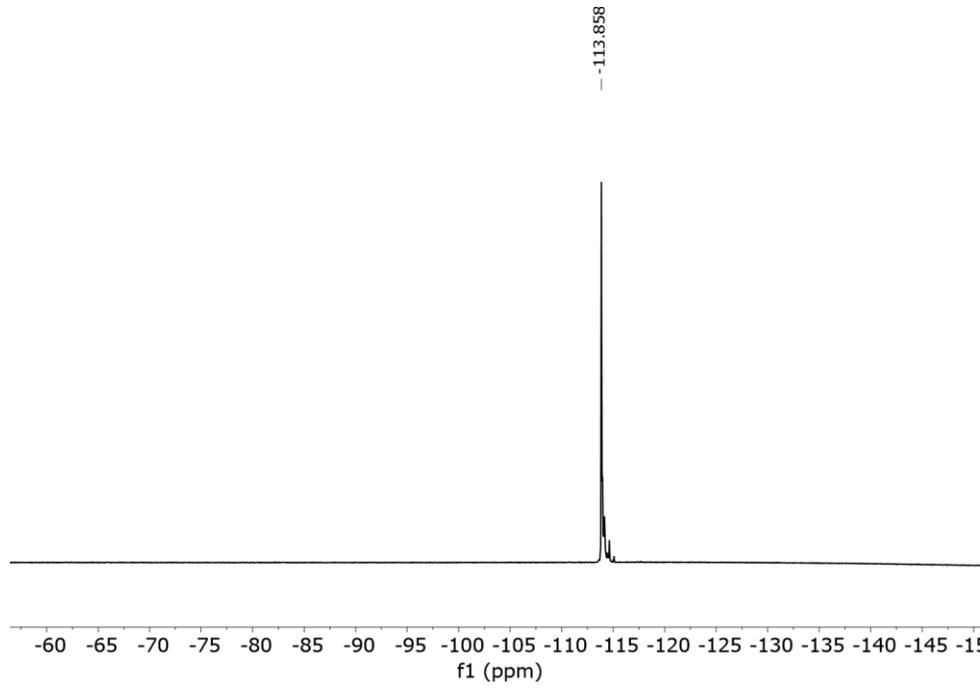
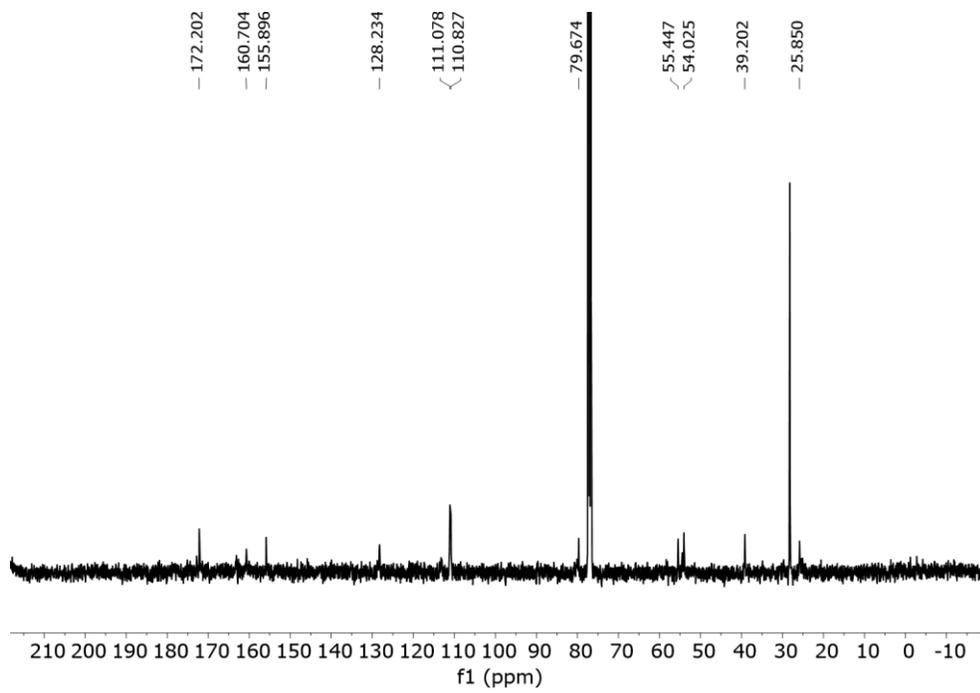
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.2, 161.7 (d,  $^1J_{\text{C-F}} = 250$  Hz), 155.9, 128.2, 111.1 (m), 110.8 (m), 79.7, 55.5, 54.0, 39.2, 25.9.

$^{19}\text{F}$ -{H} NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -113.86.

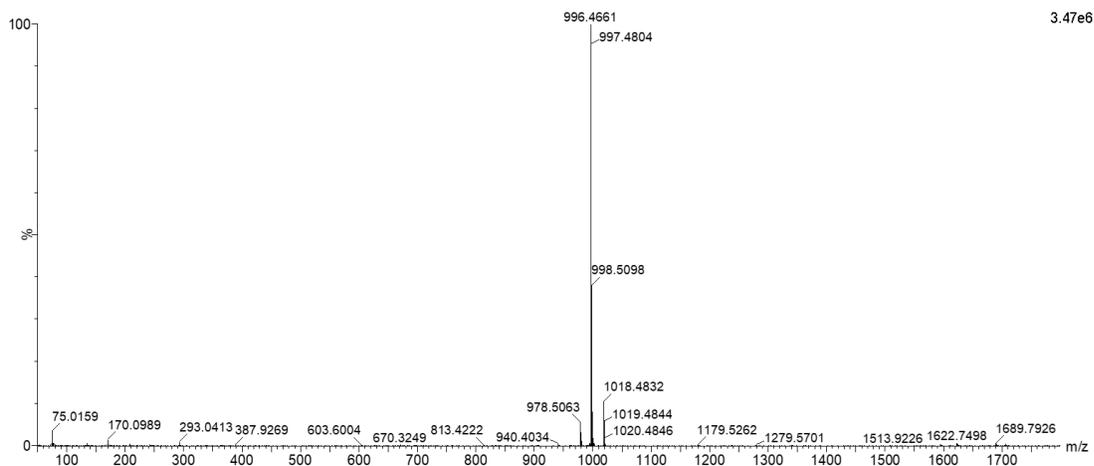
HRMS (ESI-TOF)  $m/z$  [**21** + H] $^+$  Calc: 996.4665 found: 996.4661.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ) spectra of **21**:





HRMS (ESI+) experimental spectrum of **2I**:



**3I**: **2I** (300 mg, 0.301 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3I**·4TFA as a white solid (279 mg, 0.250 mmol, 91% yield). TFA was removed by dissolving the solid in aqueous NaOH 1M and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>.

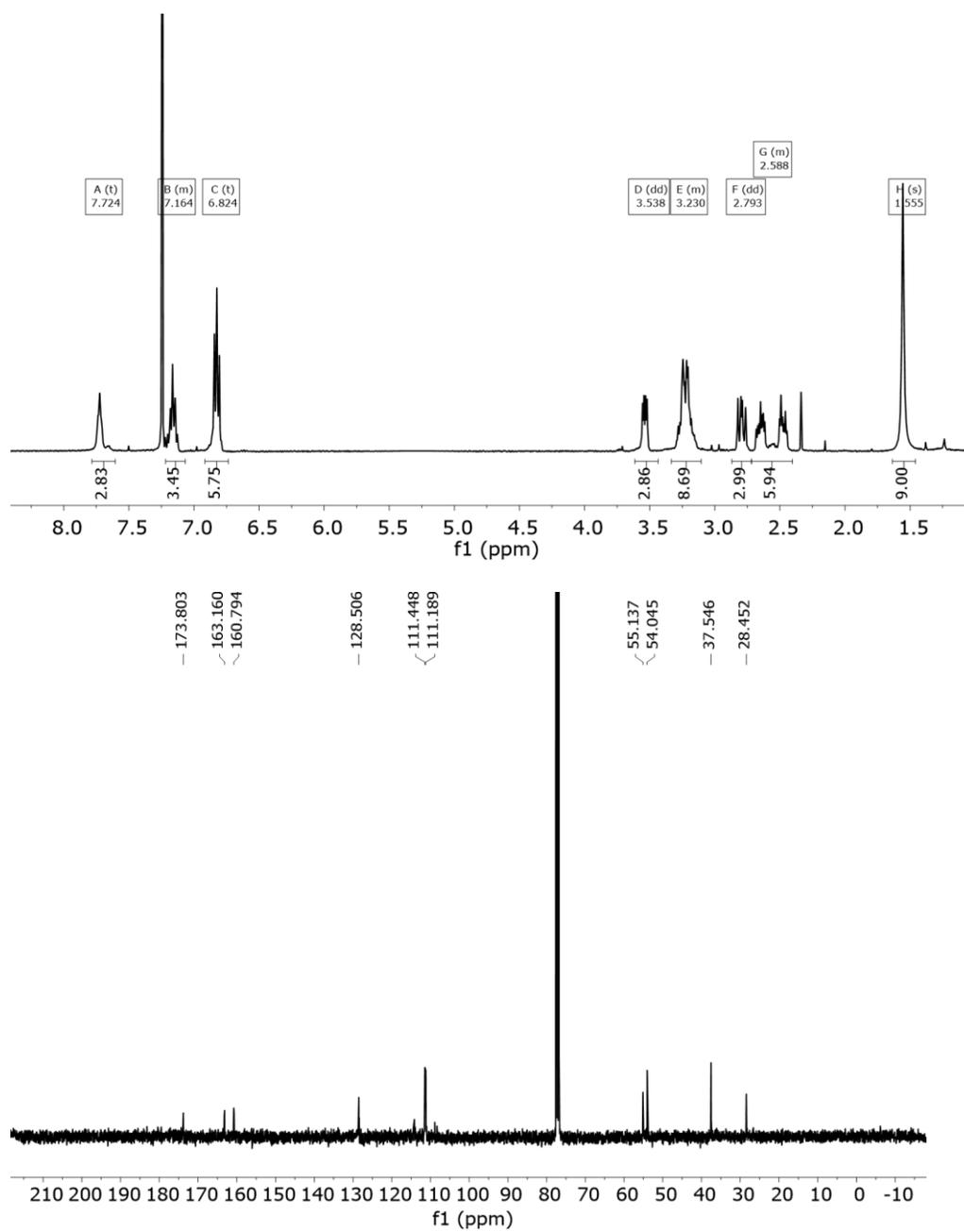
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.72 (t, *J* = 5.5 Hz, 3H), 7.16 (m, 3H), 6.82 (t, *J* = 7.7 Hz, 6H), 3.54 (X subsystem from ABX, *J*<sub>AX</sub>=4.5, *J*<sub>BX</sub>=9.6 Hz, 3H), 3.32 (m, 6H + A subsystem from ABX, *J*<sub>AX</sub>=4.5, *J*<sub>AB</sub>=13.9 Hz, 3H), 2.79 (B subsystem from ABX, *J*<sub>BX</sub>=9.5, *J*<sub>AB</sub>=13.9 Hz, 3H), 2.59 (m, 6H), 1.56 (s, 9H).

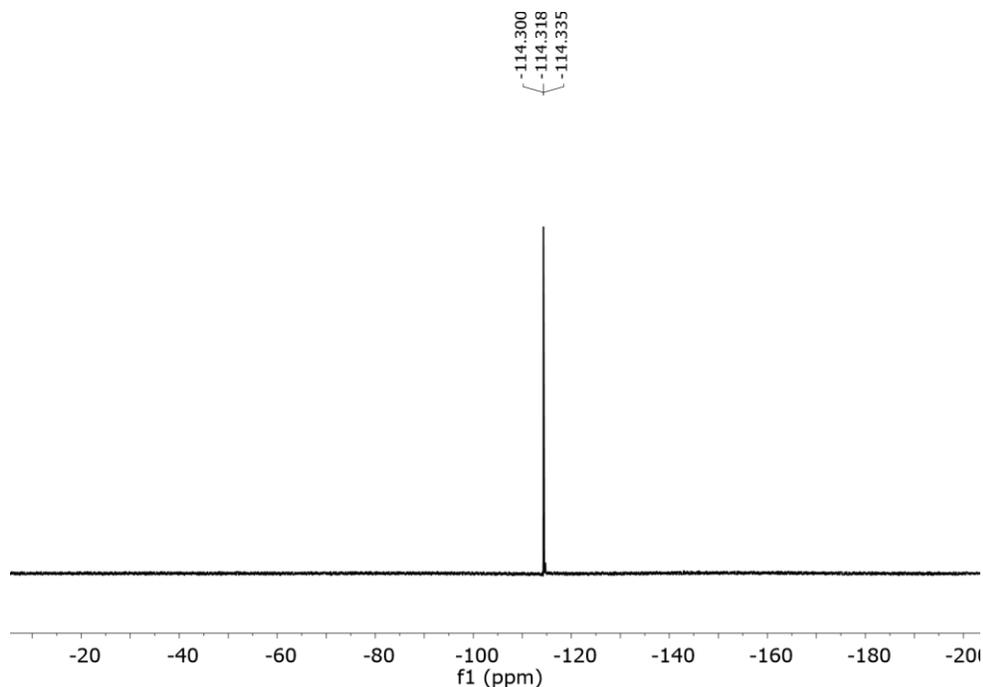
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 173.8, 161.6 (dd, <sup>1</sup>*J*<sub>C-F</sub> = 246.8, 8.8 Hz), 128.5 (t, <sup>n</sup>*J*<sub>C-F</sub> = 10.4, 8.8 Hz), 114.1, 111.4 (m), 111.2 (m), 108.9, 55.1, 54.1, 37.6, 28.5.

<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -114.32. Spectra multiplicity due to H coupling.

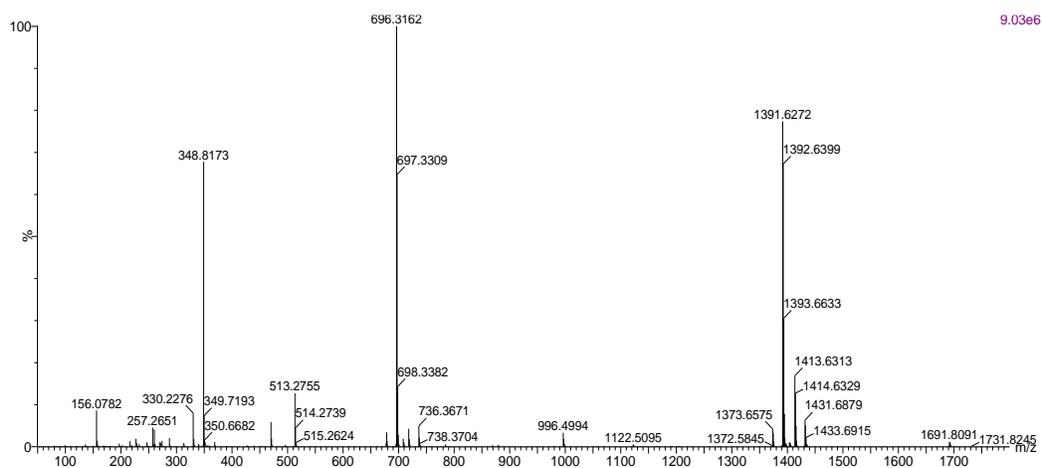
HRMS (ESI-TOF) m/z [**3I** + H]<sup>+</sup> Calc: 696.3091 found: 696.3162.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ) spectra of **31**:

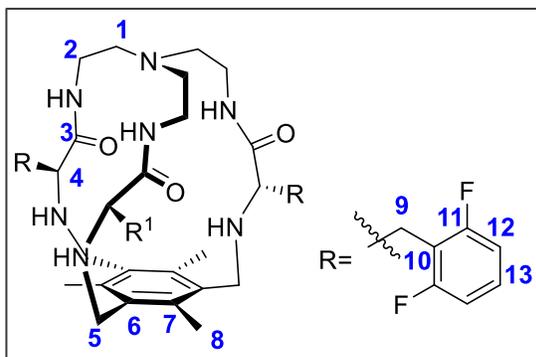




HRMS (ESI+) experimental spectrum of **3I**:



**1I (2,6-diF-Phe): 3I** (150 mg, 0.150 mmol) was dissolved in ACN (50 mL). Tetrabutylammonium chloride (30 mg, 0.097 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene ne (86 mg, 0.216 mmol) and potassium carbonate (597 mg, 4.32 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using DCM : MeOH 97 : 3 as eluent to give **1I** as a white solid (97 mg, 0.073 mmol, 53 % yield).



$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.22 (m, 3H, (H13)), 7.11 (m, 3H,  $\text{NH}_{\text{amide}}$ ), 6.91 (m, 6H, H12) 3.74 (( $\text{AB}_q$ ,  $\delta_A=3.80$ ,  $\delta_B=3.71$ ,  $J_{AB}=13.4$  Hz, 6H, H5)), 3.57 (X subsystem from ABX,  $J_{AX}=4.7$ ,  $J_{BX}=7.2$  Hz, 3H (H4)), 3.35 (A subsystem from ABX,  $J_{AX}=4.7$ ,  $J_{AB}=14.2$  Hz, 3H, (H9)), 3.03 (B subsystem from ABX,  $J_{BX}=7.2$ ,  $J_{AB}=14.2$  Hz, 3H, (H9) + m, 3H (H2)), 2.64 (m, 6H, (H2+1)), 2.38 (m, 3H, (H1)), 2.25 (s, 9H, (H8)).

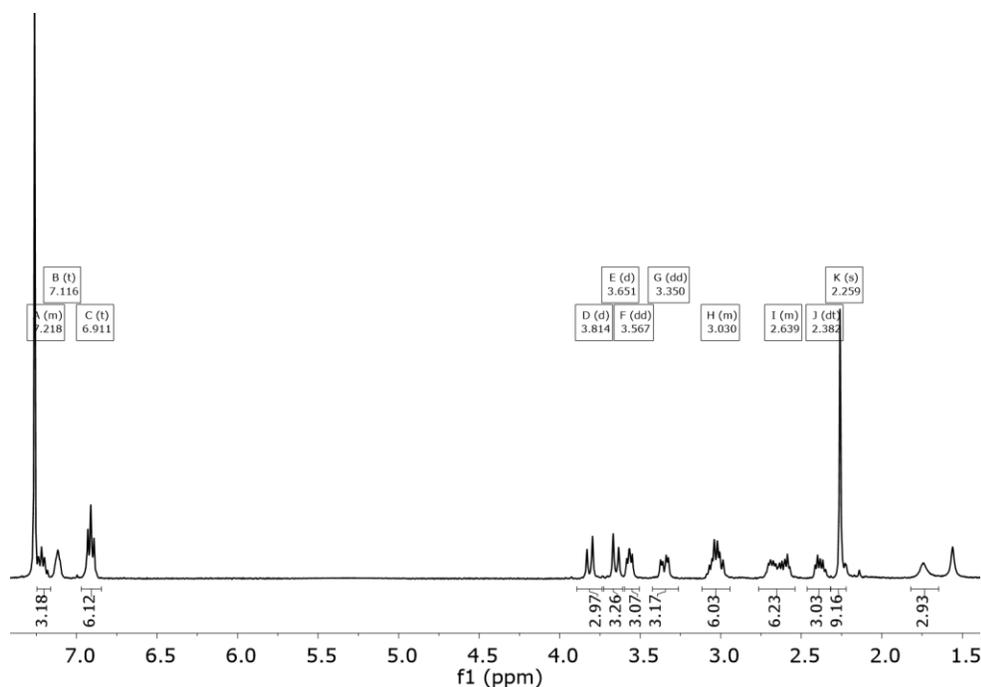
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  163.2 (C3), 160.7 (C11),

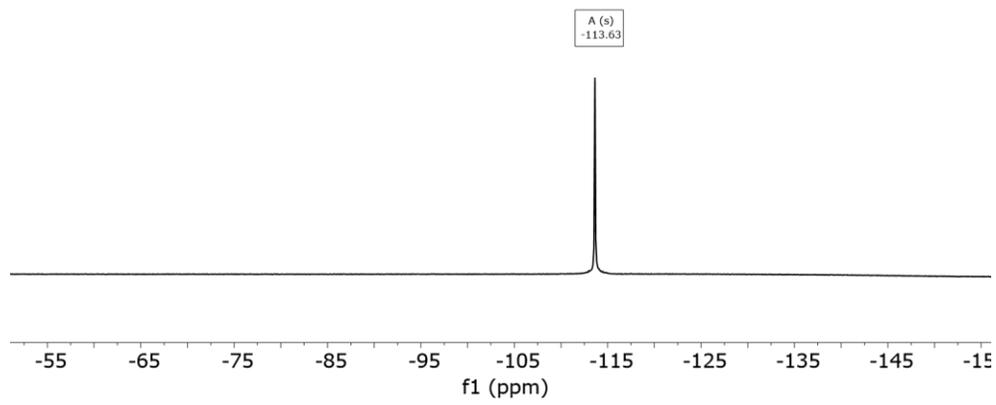
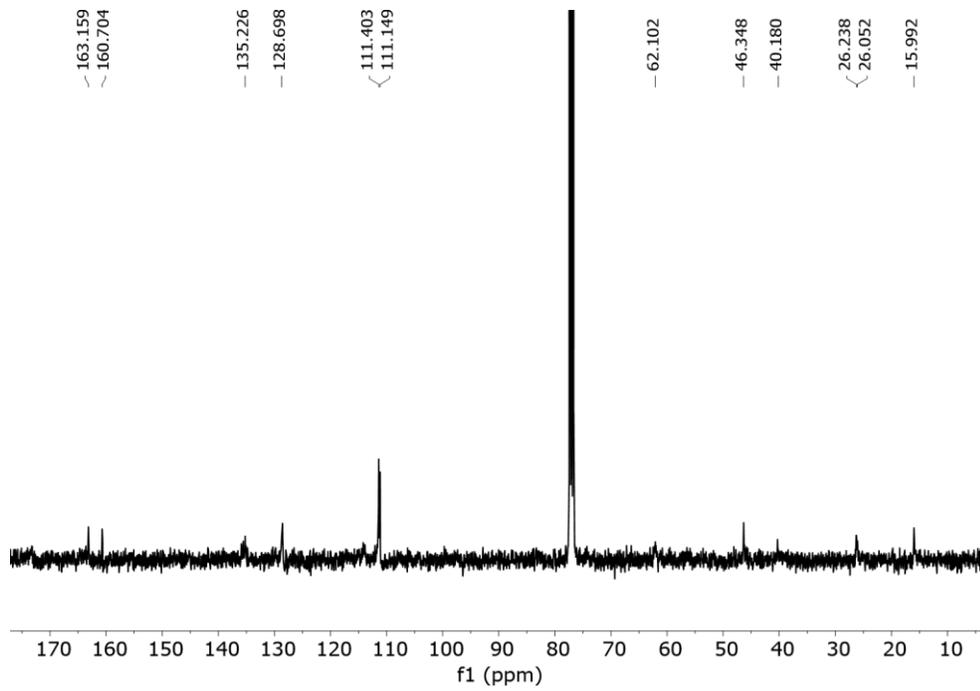
135.2 (C6+7), 128.7 (C13), 111.4 (C12), 111.2 (C10), 62.1 (C4+1) 46.4 (C5), 40.2 (C2), 26.2 (C9), 16.0 (C8)

$^{19}\text{F}$ - $\{^1\text{H}\}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -113.63

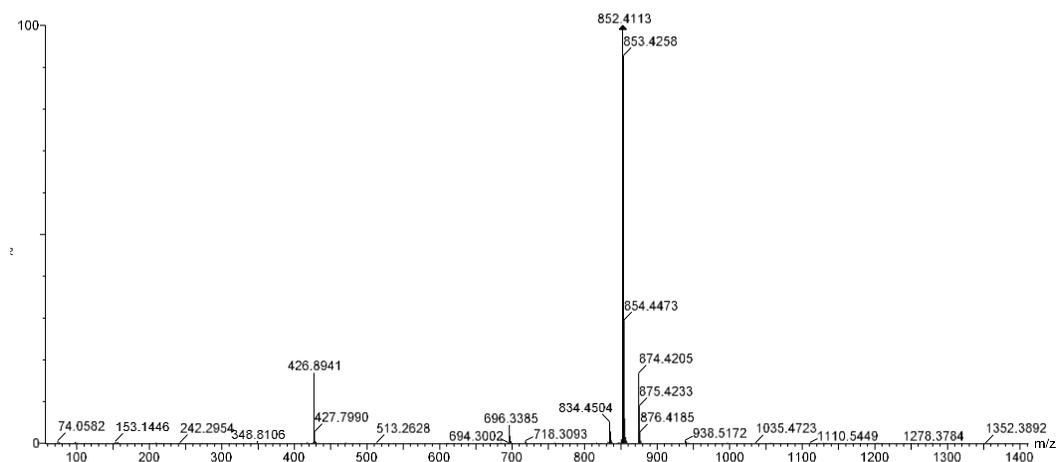
HRMS (ESI-TOF)  $m/z$  [ $\mathbf{11} - \text{H}$ ] $^+$  Calc: 852.4030; found: 852.4113.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ ) spectra of **11**:





HRMS (ESI+) experimental spectrum of **11**:



### **Synthesis of 1m (2,4-diF-Phe)**

**Boc-(2,4F)Phe-OH:** To a solution of 2,4-difluoro-phenylalanine (0.5 g, 2.1 mmol) in THF/H<sub>2</sub>O 1:1 (8.5 mL) was added di-tertButyl dicarbonate (0.5 g, 2.3 mmol) and sodium hydroxide (0.184 g, 4.62 mmol). The reaction was stirred at room temperature for 16 hours. The THF was removed in vacuum and DCM was added to the reaction flask. 10% HCl was added drop wise to this solution while stirring vigorously until the precipitate ceased (~pH = 4). The reaction mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3x150mL). The organic layer was then dried over anhydrous sodium sulphate (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuum to afford **Boc-(2,4F)Phe-OH** as a white solid (0.606 g, 2.03 mmol, 97% yield).

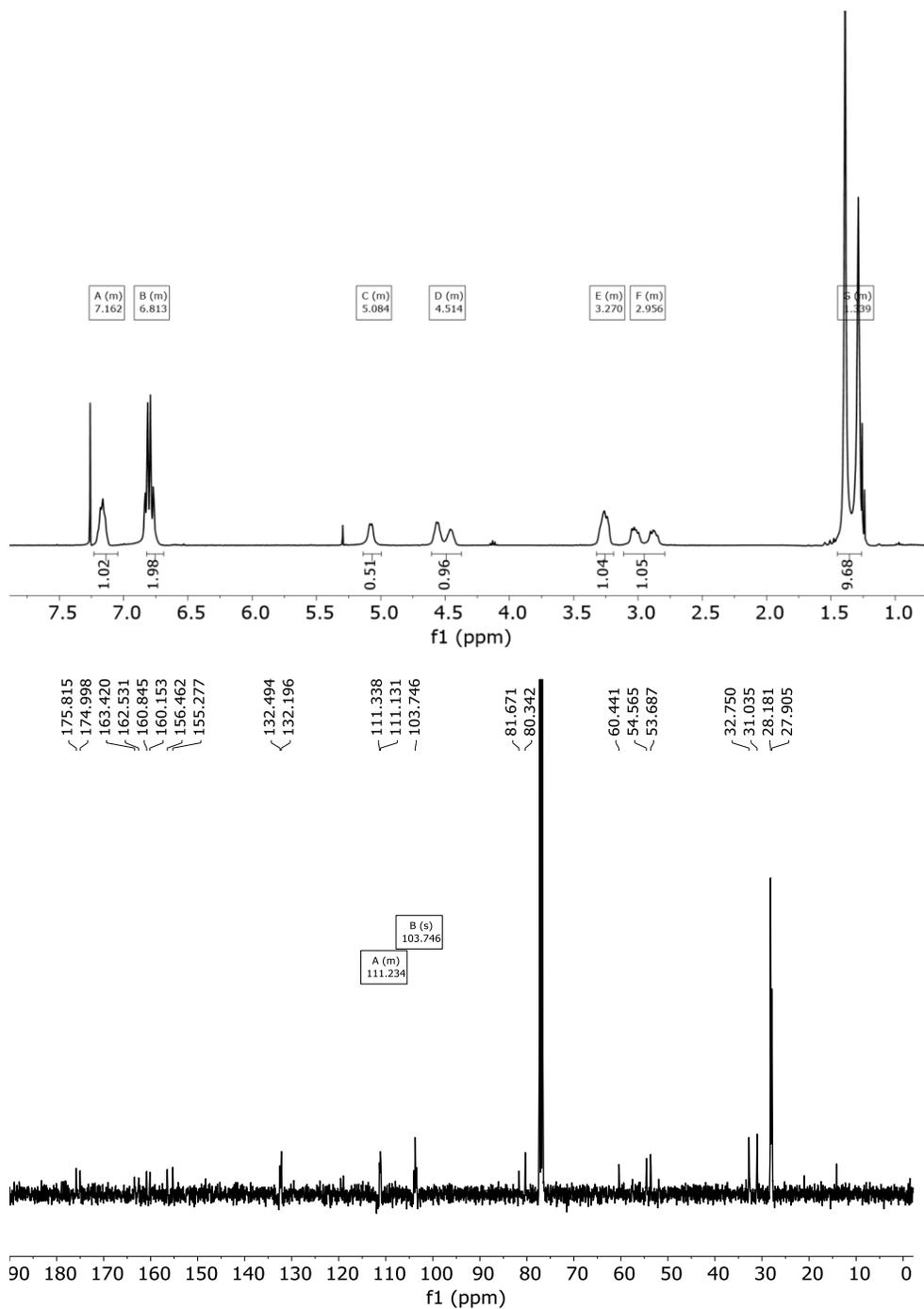
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.16 (m, 1H), 6.81 (m, 2H), 5.08 (m, 1H), 4.57-4.46 (m, 1H, rotamers), 3.27 (m, 1H), 3.03-2.88 (m, 1H, rotamers), 1.39-1.29 (m, 9H, rotamers).

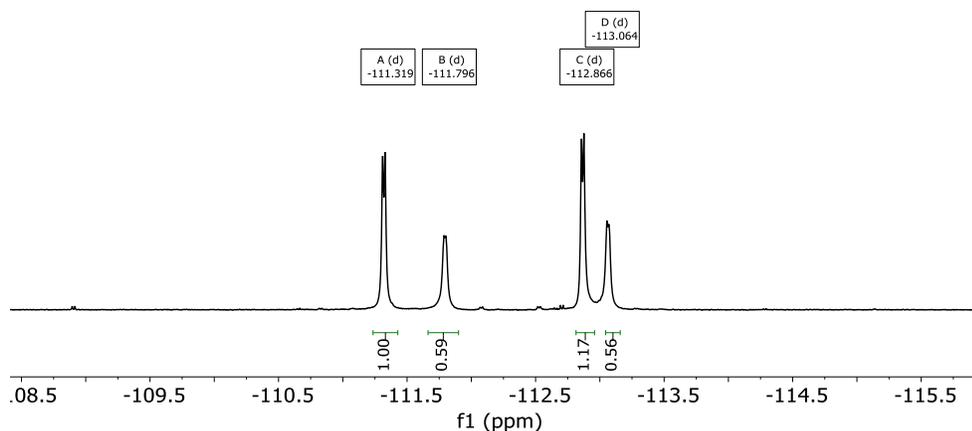
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 175.8/175.0, 163.4/162.5, 160.8/160.2, 156.5/155.3, 132.5/132.2, 111.3 (m, C-F coupling), 103.8 (m, C-F coupling), 81.7/80.3, 60.4, 54.6/53.7, 32.8/31.1, 28.2/27.9. Most of the C give two different signals due to the presence of rotamers.

<sup>19</sup>F -{H} NMR (376 MHz, CDCl<sub>3</sub>) δ -111.31/-112.87 (d, J = 7.4 Hz, 3F), -111.78/-113.06 (d, J = 7.1 Hz, 3F). The two F have two signals associated to rotamers.

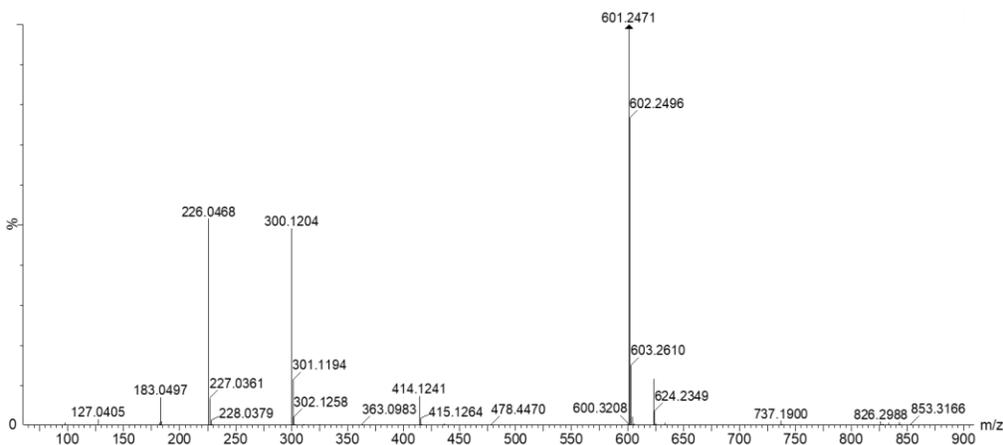
HRMS (ESI-TOF) m/z [Boc-(2,4F)Phe-OH - H]<sup>-</sup> Calc: 300.2824, found: 300.1204; also [2(Boc-(2,4F)Phe-OH - H)]<sup>-</sup> Calc: 601.2178, found: 601.2471.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CDCl}_3$ ) of **Boc-(2,4F)Phe-OH**:





HRMS (ESI+) experimental spectrum of **Boc-(2,4F)Phe-OH**:



**2m**: Boc-(2,4F)Phe-OH (0.50 g, 1.68 mmol) was dissolved in dry DMF (3 mL). In a separate flask, EDC·HCl (0.393 mg, 2.05 mmols) and HOBt (0.313 mg, 2.05 mmols) were dissolved in dry DMF and added to the amino-acid solution and diisopropylethylamine (DIPEA, 1 mL, 6.15 mmols) was also added to the solution. The reaction mixture was stirred under nitrogen at room temperature for 10 minutes to allow the activation of the acid. Then tris(2-aminoethyl)amine (0.08 mL, 0.512 mmols) was added. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was purified by flash chromatography using 97:3 DCM:MeOH as eluent to give 0.435 mg of **2m** (0.436 mmol, 84% yield).

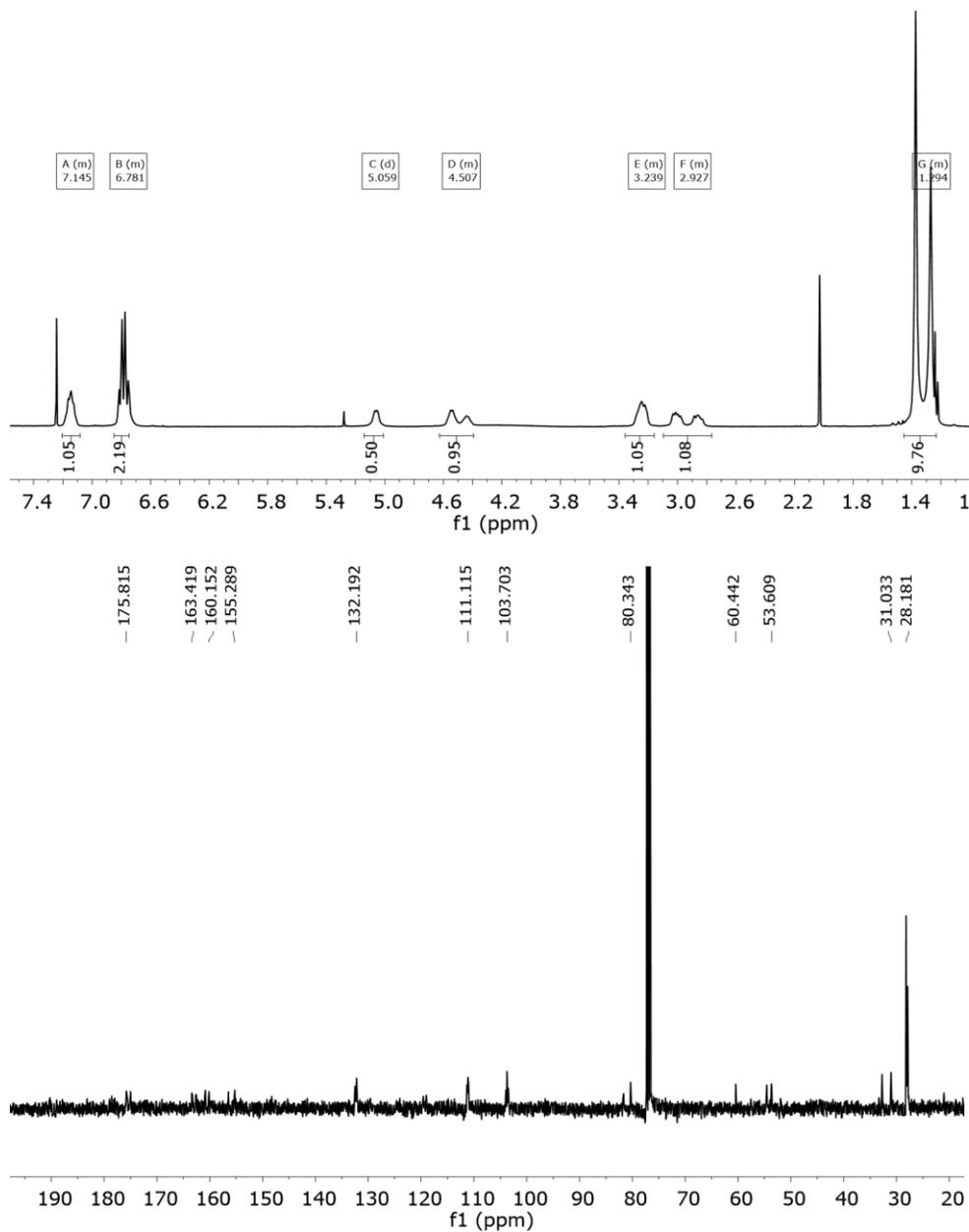
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.15 (m, 3H), 6.78 (m, 6H), 5.06 (d, *J* = 8.0 Hz, 3H), 4.51 (m, 3H), 3.24 (m, 3H), 2.93 (m, 3H), 1.35 (m, 27H (rotamers)).

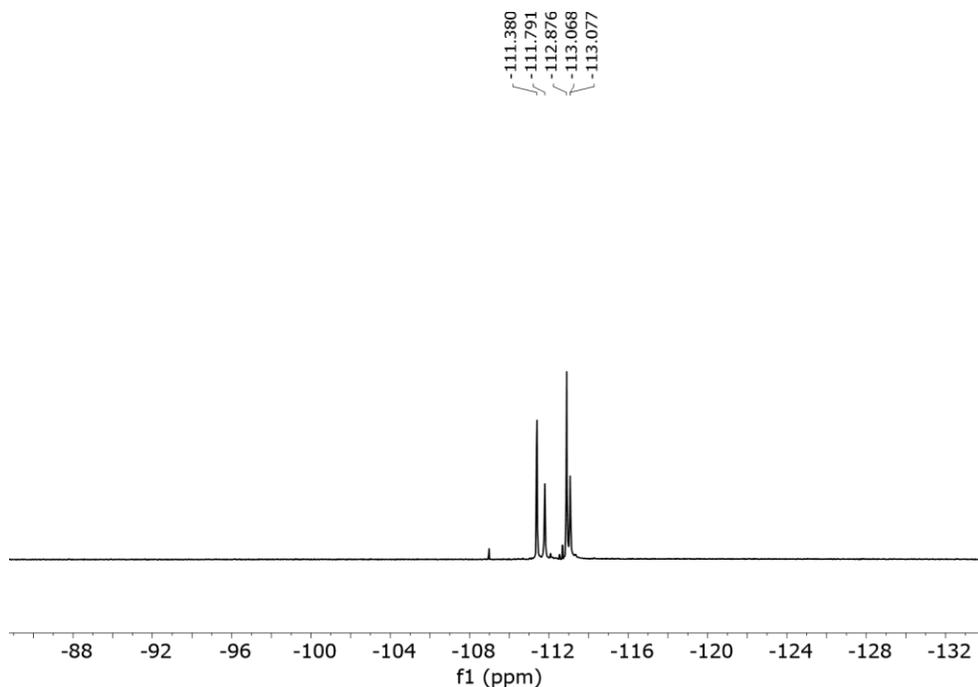
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  175.8, 163.4, 160.2-155.5 two d corresponding to C-F, 132.2 (m), 119.6, 111.1 (m), 103.7 (m), 80.3, 60.4, 53.7-53.6 (rotamers), 32.7-31.0 (rotamers), 28.2 (rotamers). Multiplicity due to C-F coupling and the presence of rotamers lead to a complex spectrum.

$^{19}\text{F}$  -{H} NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -111.38-111.97 (rotamers), -112.88-113.08 (rotamers).

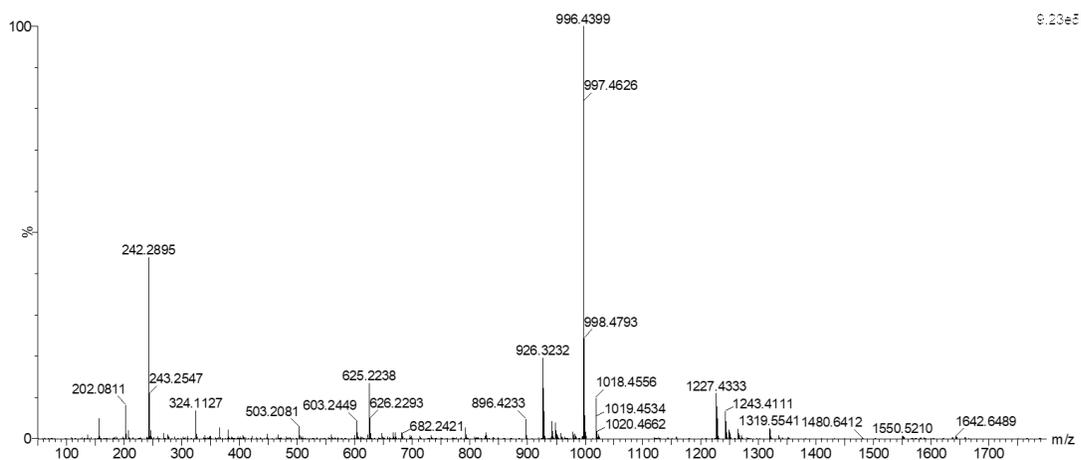
HRMS (ESI-TOF)  $m/z$  [ $2\mathbf{m} + \text{H}$ ] $^+$  Calc: 996.4664, found: 996.4399.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ ) spectra of  $2\mathbf{m}$ :





HRMS (ESI+) experimental spectrum of **2m**:



**3n**: **2m** (220 mg, 0.250 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3m**·4TFA as a white solid (279 mg, 0.245 mmol, 95% yield). TFA was removed by dissolving the solid in 1 M aqueous NaOH and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>.

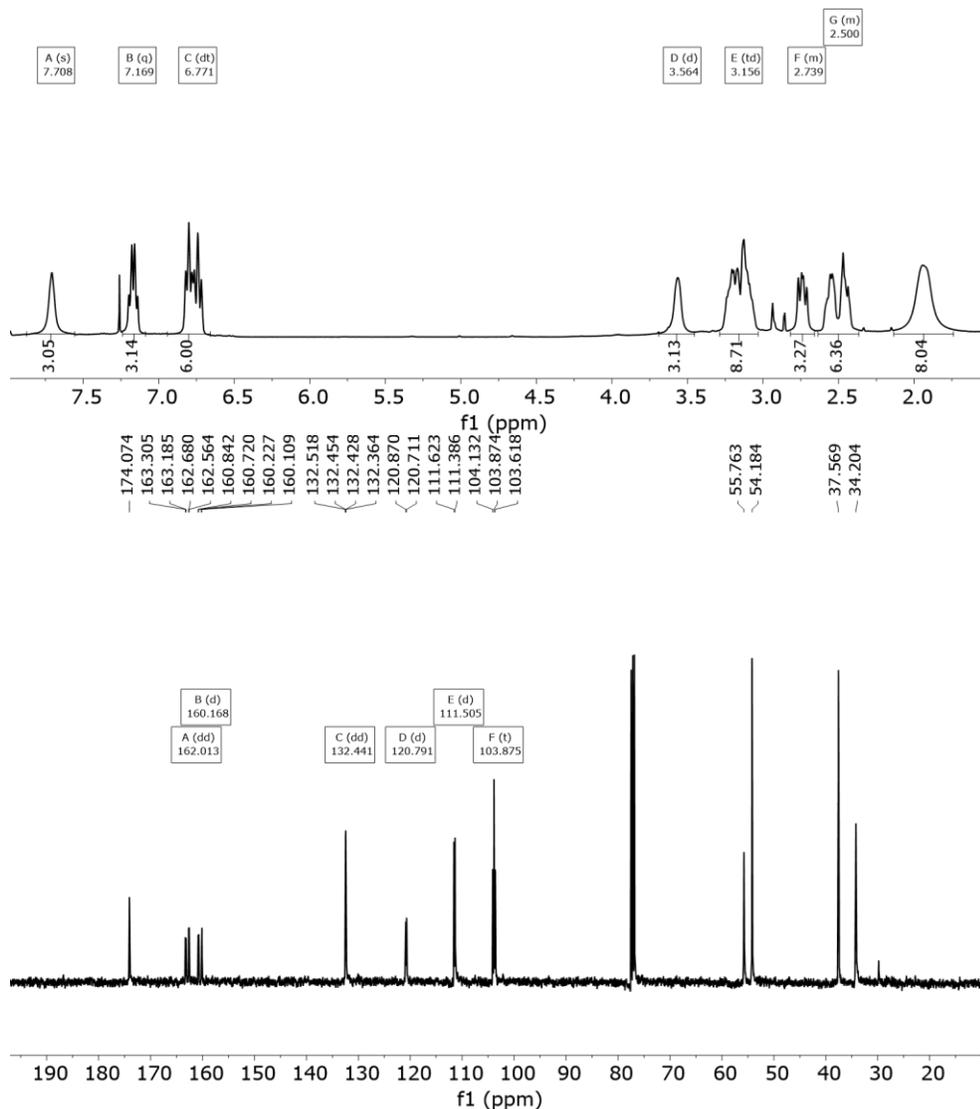
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.71 (s, 3H), 7.17 (q, *J* = 9.0 Hz, 3H), 6.77 (dt, *J* = 24.8, 9.0 Hz, 6H), 3.54 (m, 3H), 3.16 (m, 9H), 2.74 (m, 3H), 2.50 (m, 6H), 1.95 (m, 9H).

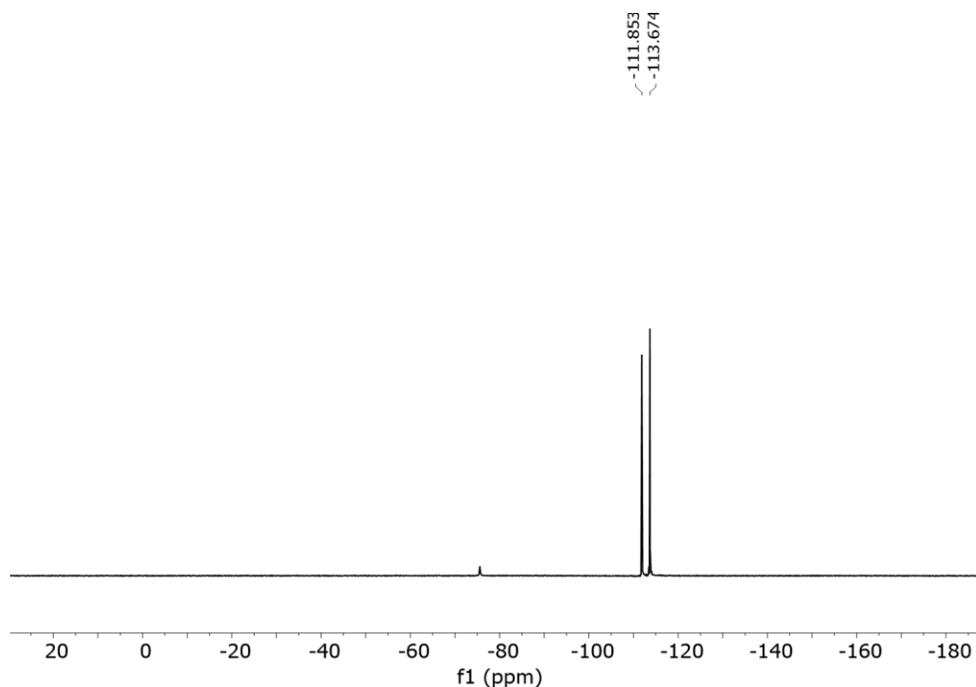
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.1, 162.01 (dd,  $^nJ_{\text{C-F}} = 247.9, 12.2$  Hz), 160.1 (dd,  $J_{\text{C-F}} = 247.9, 12.2$  Hz), 132.5 (dd,  $^nJ_{\text{C-F}} = 9.1, 6.5$  Hz), 120.9 (d,  $^nJ_{\text{C-F}} = 15.9$  Hz), 111.6 (d,  $^nJ_{\text{C-F}} = 23.8$  Hz), 103.9 (d,  $^nJ_{\text{C-F}} = 25.9$  Hz), 55.8, 54.8, 37.6, 34.2.

$^{19}\text{F}$  – {H} NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -113.67, -111.85

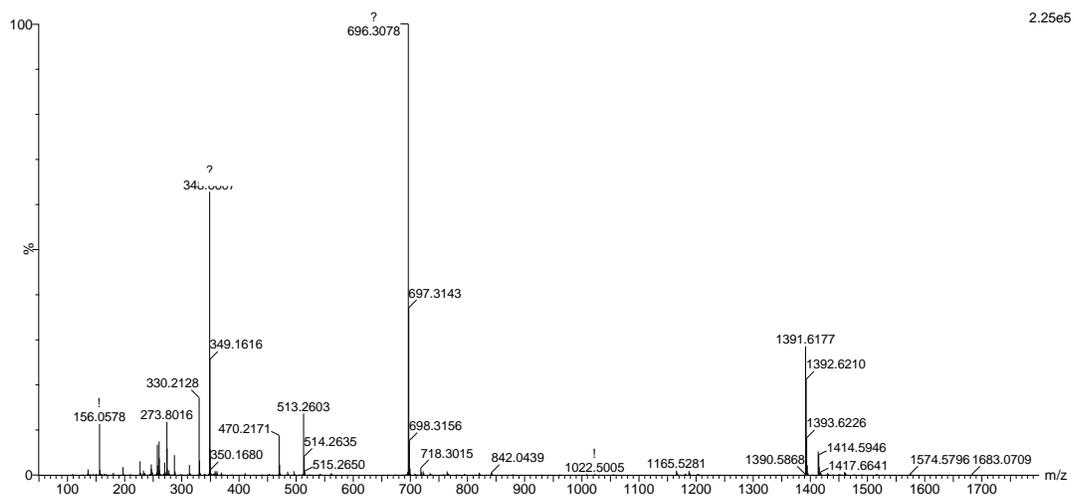
HRMS (ESI-TOF)  $m/z$  [ $\mathbf{3m} + \text{H}$ ] $^+$  Calc: 696.3091, found: 696.3078

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ ) spectra of  $\mathbf{3m}$ :

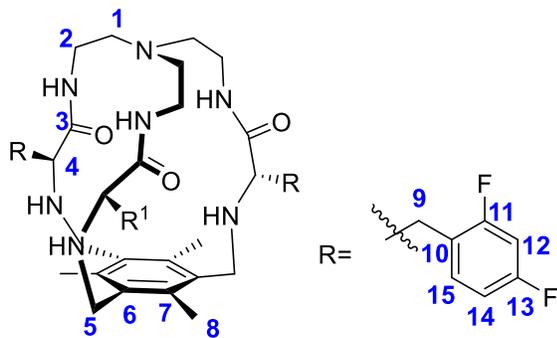




HRMS (ESI+) experimental spectrum of **3m**:



**1m (2,4-diF-Phe): 3m** (150 mg, 0.21 mmol) was dissolved in ACN (60 mL). Tetrabutylammonium chloride (30 mg, 0.10 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (86.1 mg, 0.216 mmol) and potassium carbonate (596 mg, 4.31 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography DCM : MeOH 97 : 3 as eluent to give **1m** as a white solid (80 mg, 0.094 mmol, 43 % yield).



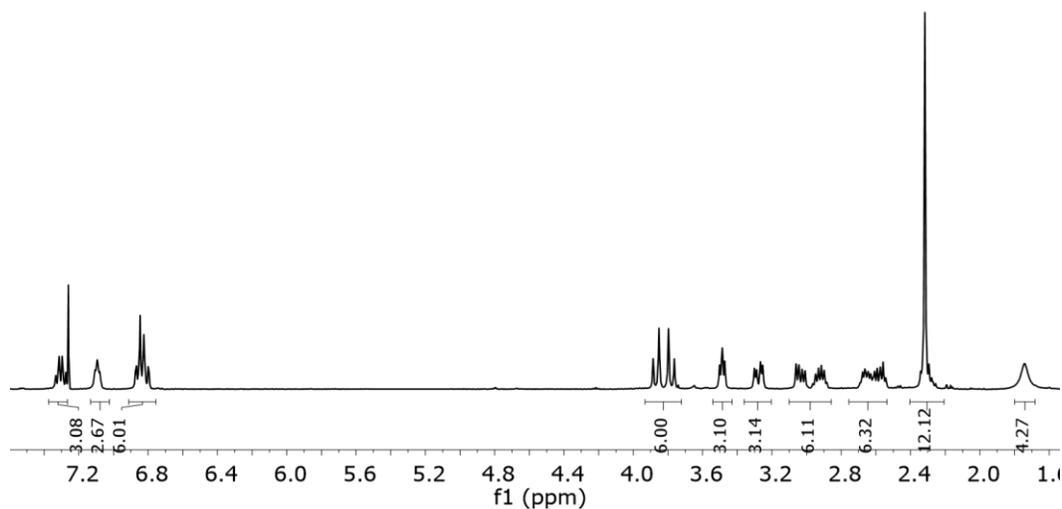
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.30 (m, 3H, H15), 7.1 (t,  $J = 5.8$  Hz, 3H,  $\text{NH}_{\text{amide}}$ ), 6.83 (m, 6H, H12+H14), 3.81 (AB<sub>q</sub>,  $\delta_{\text{A}} = 3.87$ ,  $\delta_{\text{B}} = 3.78$ ,  $J_{\text{AB}} = 13.6$  Hz, 6H, H5), 3.49 (X subsystem from ABX,  $J_{\text{AX}} = 5.2$ ,  $J_{\text{BX}} = 6.8$  Hz, 3H, H4), 3.28 (A subsystem from ABX,  $J_{\text{AX}} = 5.2$ ,  $J_{\text{AB}} = 14.1$  Hz, 3H, H9), 3.04 (B subsystem from ABX,  $J_{\text{AB}} = 14.1$ ,  $J_{\text{BX}} = 6.8$  Hz, 3H, H9), 2.93 (m, 3H, H2), 2.67 (m, 3H, H2), 2.56 (m, 3H, H1), 2.31 (s + m, 4H, H1 + H8), 1.74 (s, 3H,  $\text{NH}_{\text{amine}}$ ).

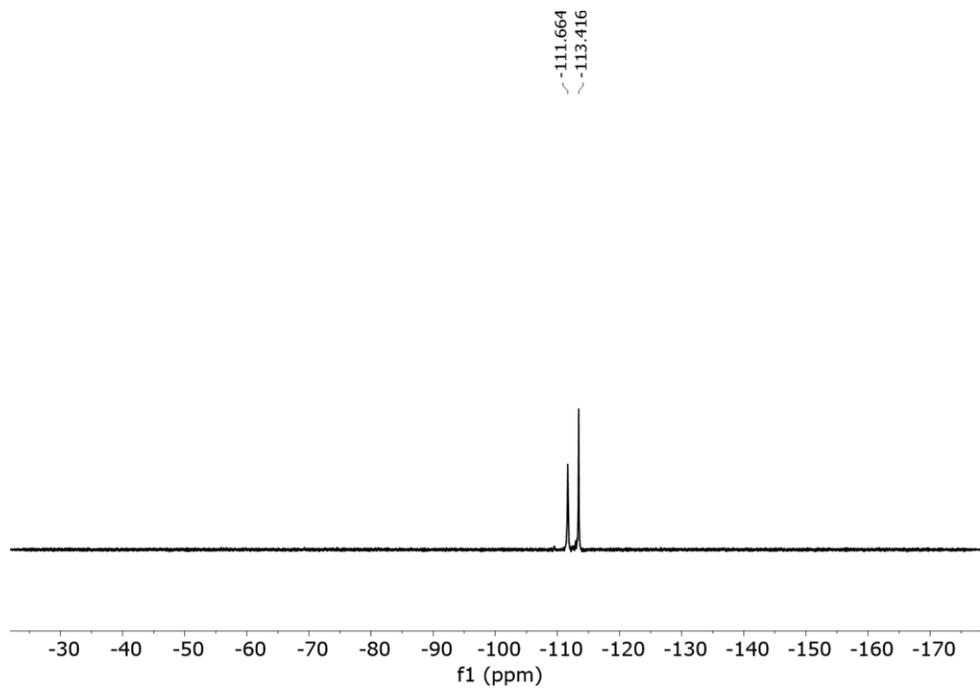
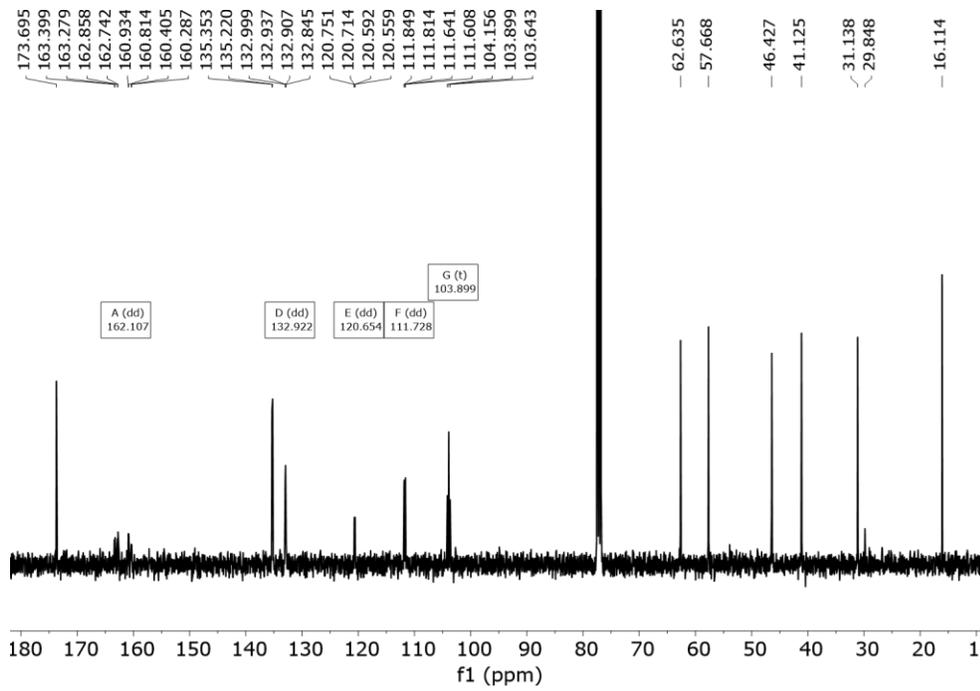
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.7 (C3), 162.1 (dd,  $^nJ_{\text{C-F}} = 247.0$ , 11.8 Hz) and 161.6 (dd,  $^nJ_{\text{C-F}} = 248.0$ , 12.1 Hz) (C11+C13), 135.3 and 135.2 (C6+C7), 132.9 (dd,  $^nJ_{\text{C-F}} = 9.6$ , 6.3 Hz, C15), 120.8 (dd,  $^nJ_{\text{C-F}} = 15.8$ , 3.5 Hz) C10), 111.6 (dd,  $^nJ_{\text{C-F}} = 20.9$ , 3.4 Hz, C14), 103.9 (t,  $^nJ_{\text{C-F}} = 247.0$ , 25.8 Hz, C12), 62.6 (C4), 57.7 (C1), 46.4 (C5), 41.1 (C2), 31.1 (C9), 16.1 (C8).

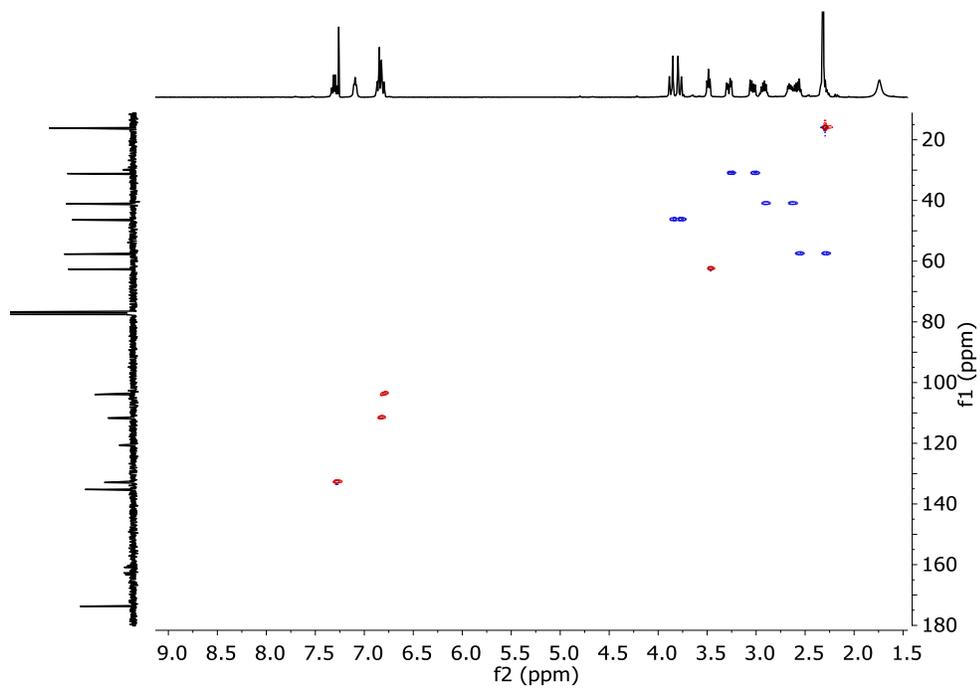
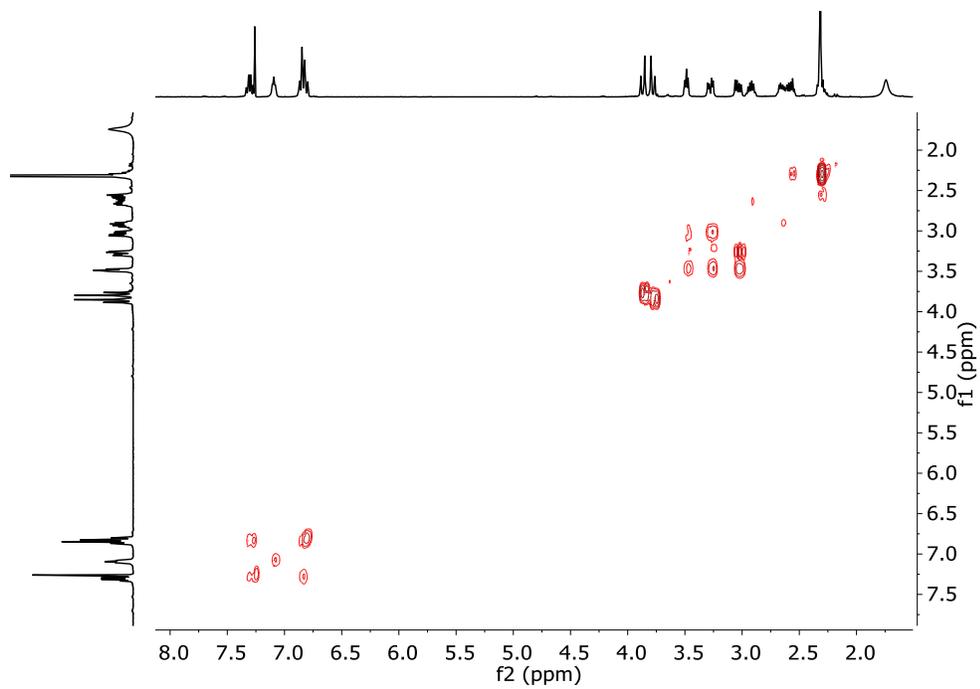
$^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -111.664, -113.416.

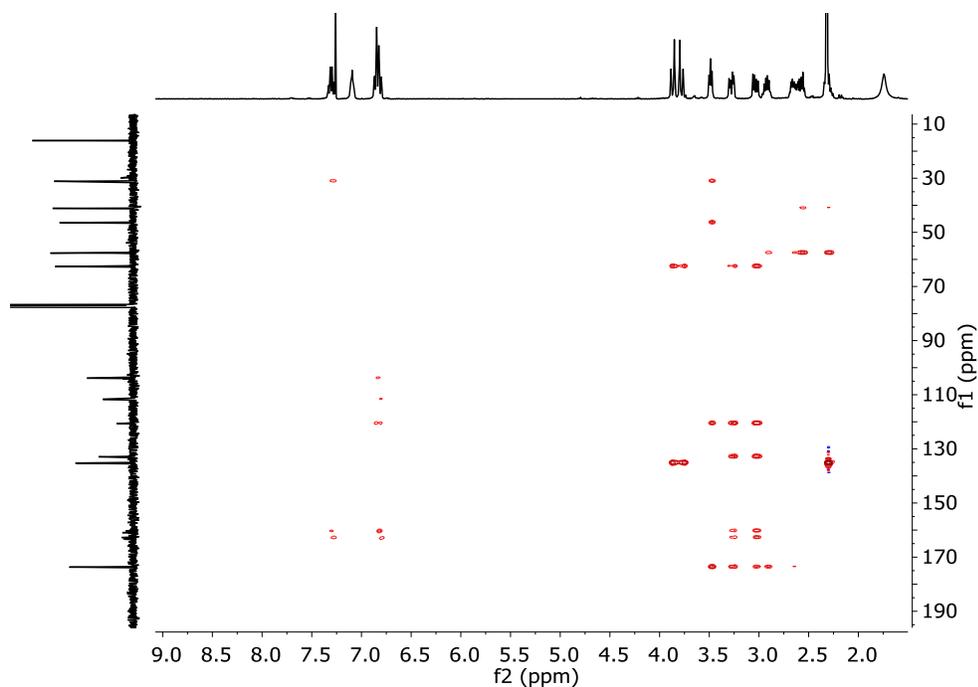
HRMS (ESI-TOF)  $m/z$  [ $\mathbf{1m} + \text{H}$ ]<sup>+</sup> Calc: 852.4030, found: 852.3644.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$ -NMR (376 MHz,  $\text{CDCl}_3$ ), COSY, HSQC and HMBC spectra of  $\mathbf{1m}$ :

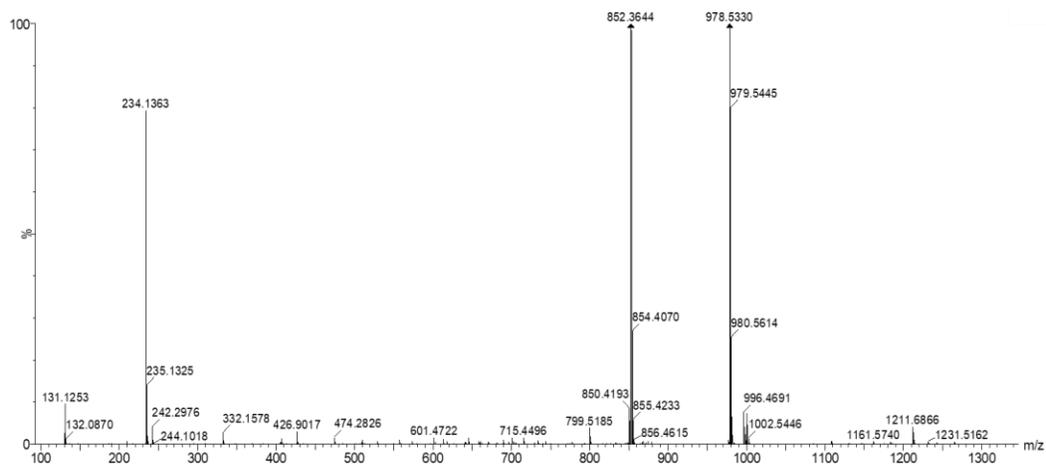








HRMS (ESI+) experimental spectrum of **1m**:



### Synthesis of **1n** (Penta-F-Phe)

**2n**: Boc-(Penta-F)Phe-OH (405 mg, 1.14 mmol) was dissolved in dry DMF (4 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 265 mg, 1.38 mmol) and 1-hydroxybenzotriazole hydrate (HOBt, 210 mg, 1.38 mmol), *N,N*-diisopropylethylamine (DIPEA, 0.72 mL, 4.15 mmol) and tris(2-aminoethyl)amine (0.051 mL, 0.32 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were washed with aqueous LiCl (5% w/w), dried over MgSO<sub>4</sub> and concentrated to

dryness. The residue was purified by flash chromatography using 98:2 DCM:MeOH as eluent to give 390 mg of **2n** (0.330 mmol, 92% yield).

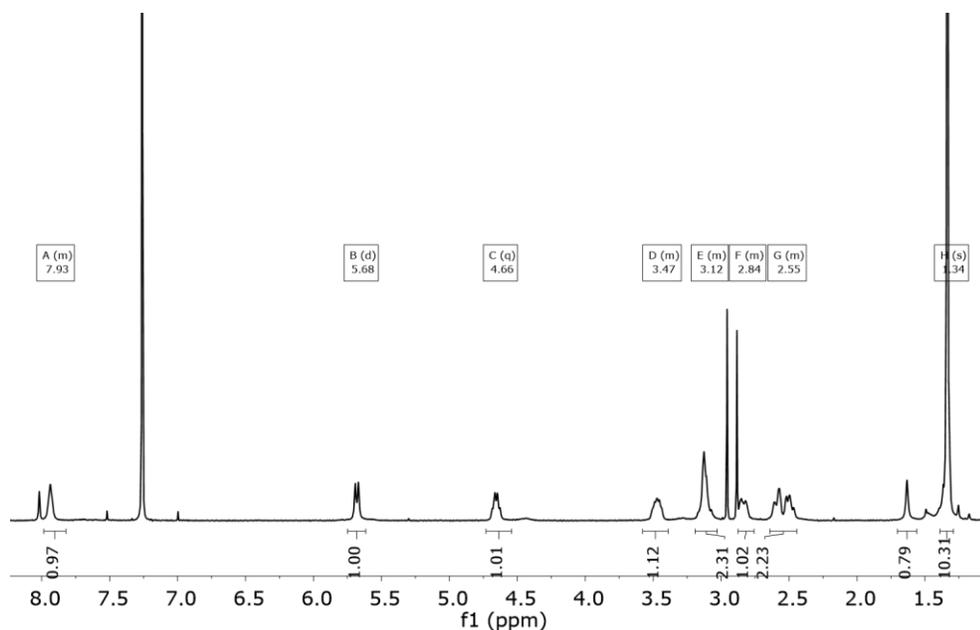
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.93 (m, 3H), 5.68 (d,  $J = 9.1$  Hz, 3H), 4.66 (m, 3H), 3.47 (m, 3H), 3.12 (m, 6H), 2.84 (m, 3H), 2.55 (m, 6H), 1.34 (s, 9H).

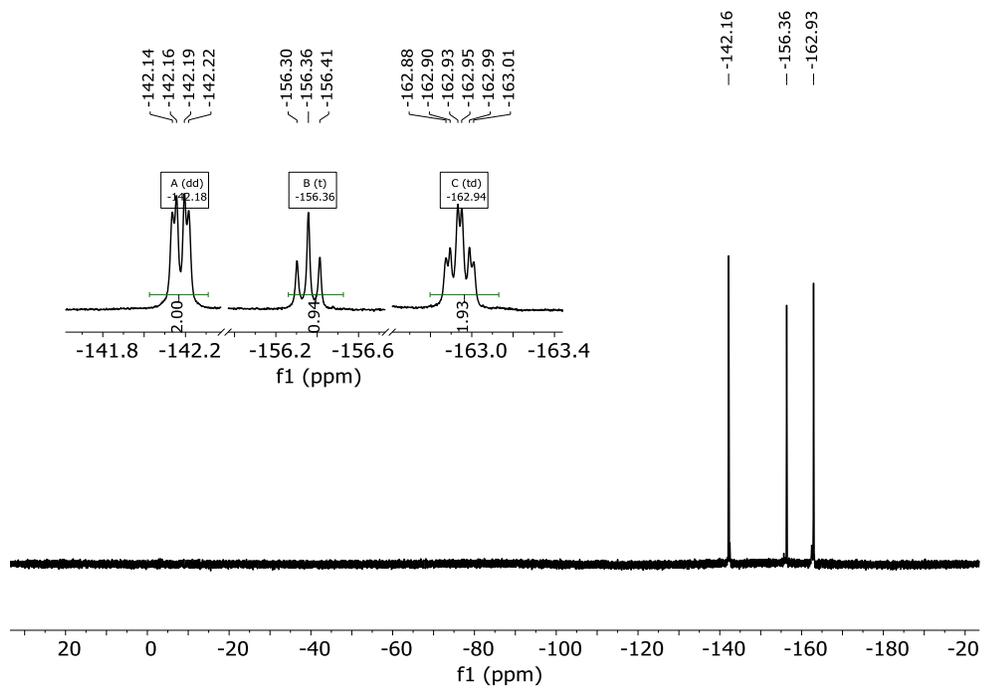
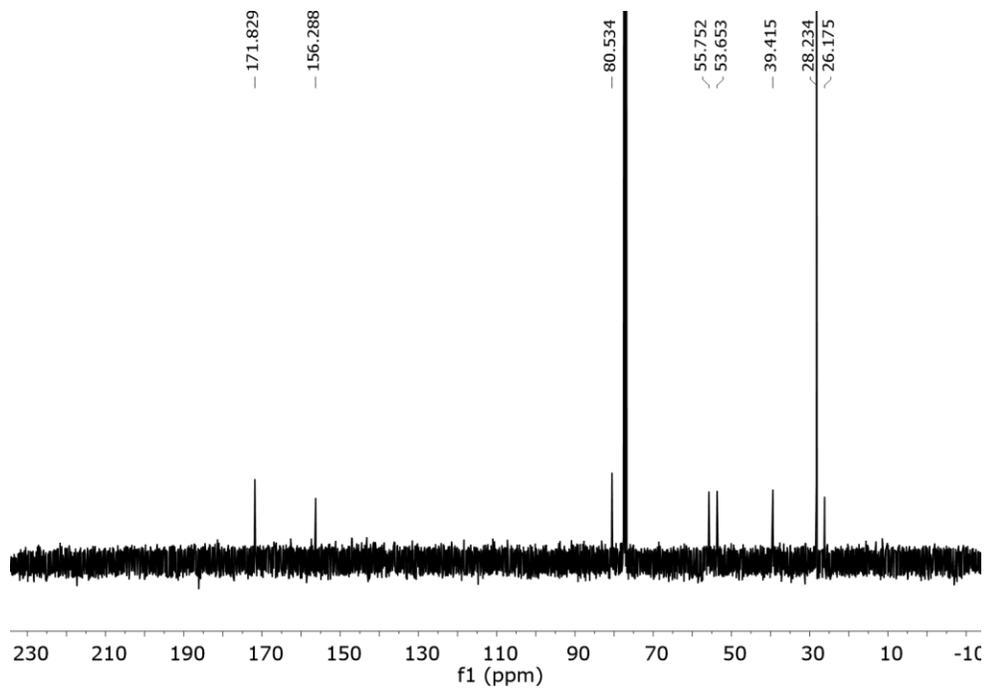
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.8, 156.3, 80.5, 55.8, 53.7, 39.4, 28.2, 26.2 ( $\text{C}_{\text{ar}}$  are not seen in this spectrum).

$^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -142.18 (dd,  $J = 22.0, 7.7$  Hz, 6F), -156.36 (t,  $J = 20.8$  Hz, 3F), -162.94 (td,  $J = 21.7, 7.6$  Hz, 6F)

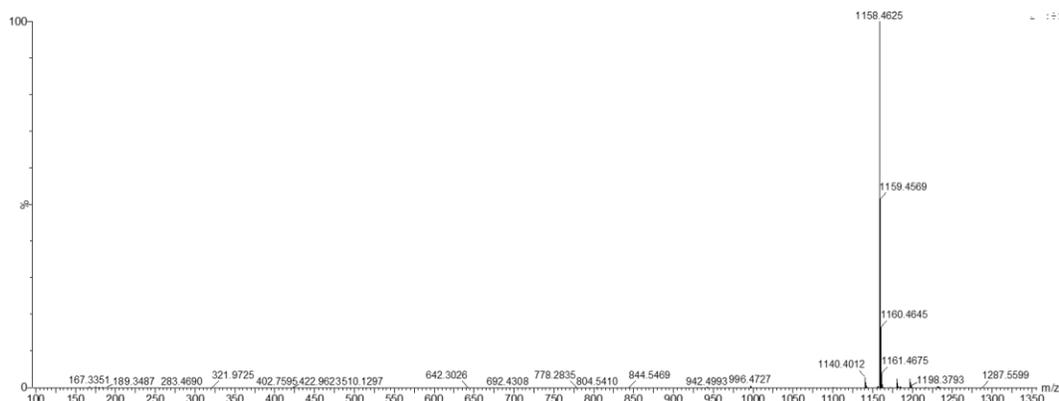
HRMS (ESI-TOF)  $m/z$  [**2n**+ H] $^+$  Calc: 1158.3816, found: 1158.4625.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ) of **2n**:





HRMS (ESI+) experimental spectrum of **2n**:



**3n**: **2n** (385 mg, 0.31 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.75 mL, 5.05 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3n**·4TFA as a white solid (395 mg, 0.28 mmol, 90% yield). TFA was removed by dissolving the solid in 1M aqueous NaOH and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>.

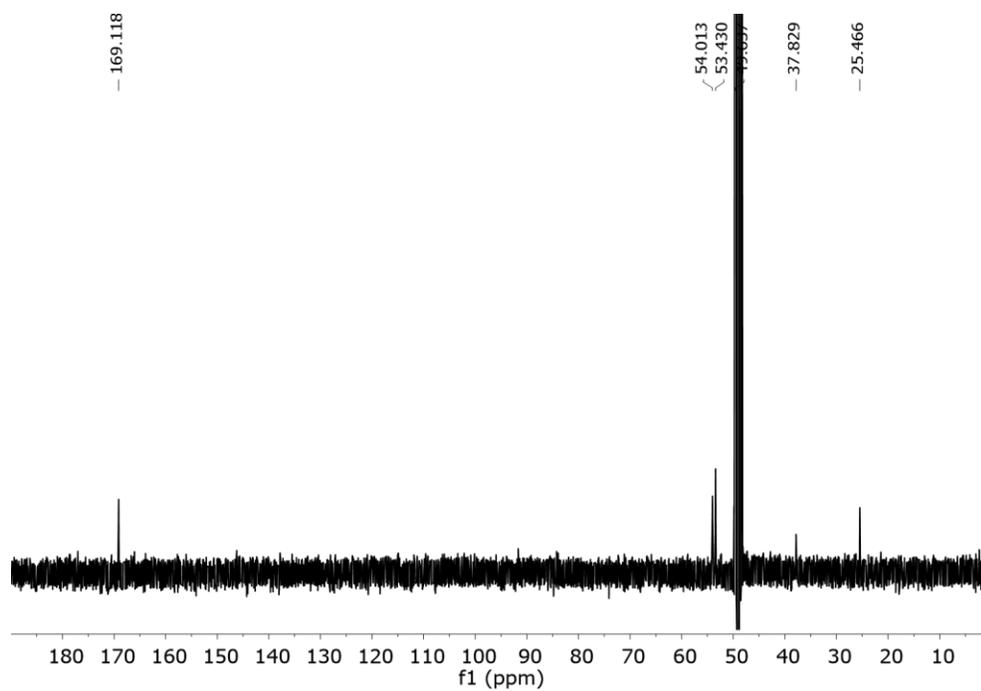
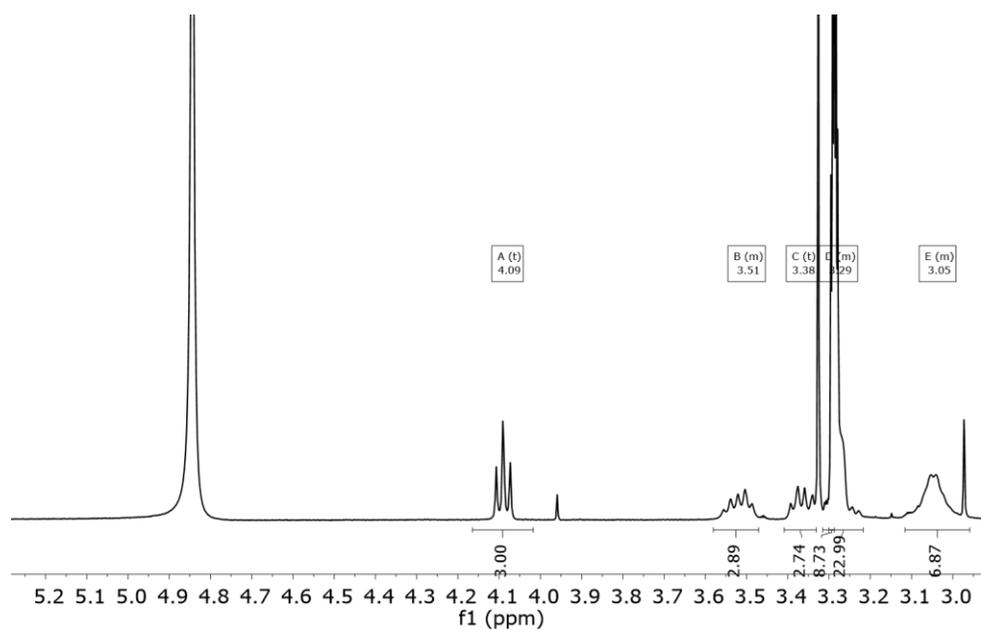
<sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ 4.09 (t, *J* = 6.8 Hz, 3H), 3.51 (m, 3H), 3.38 (t, *J* = 6.8 Hz, 3H), 3.26 (m, 6H), 3.05 (m, 6H).

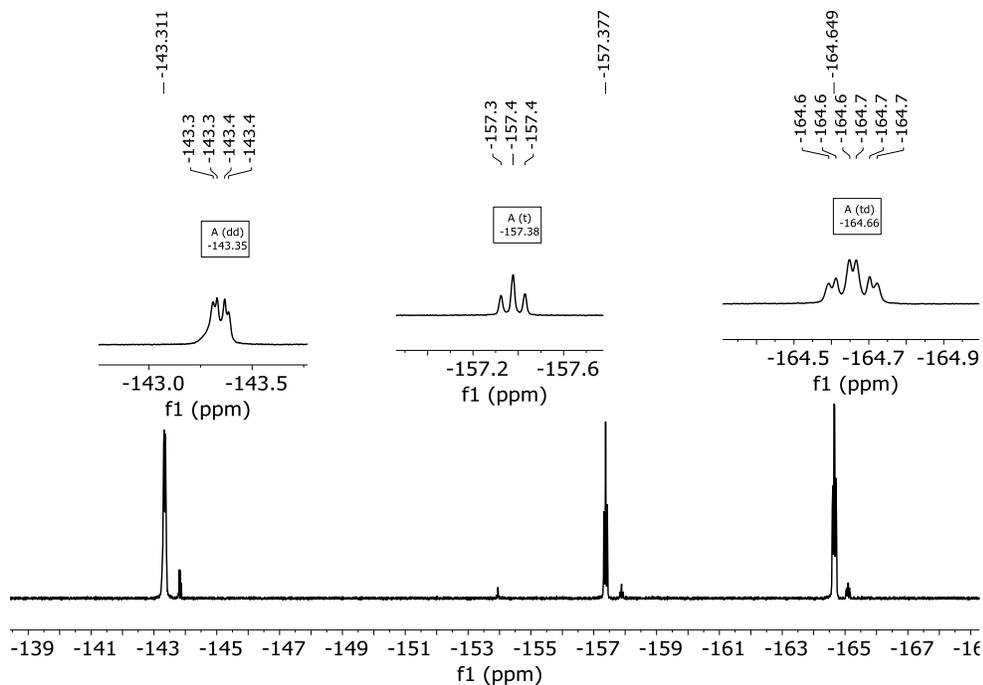
<sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 169.1, 54.0, 53.4, 49.6, 37.8, 25.5 (C<sub>ar</sub> are not seen in this spectrum).

<sup>19</sup>F –{<sup>1</sup>H}-NMR (376 MHz, CDCl<sub>3</sub>) δ -143.31 (dd, *J* = 20.9, 7.0 Hz, 6F), -157.38 (t, *J* = 19.9 Hz, 3F), -164.65 (td, *J* = 21.1, 7.4 Hz, 6F).

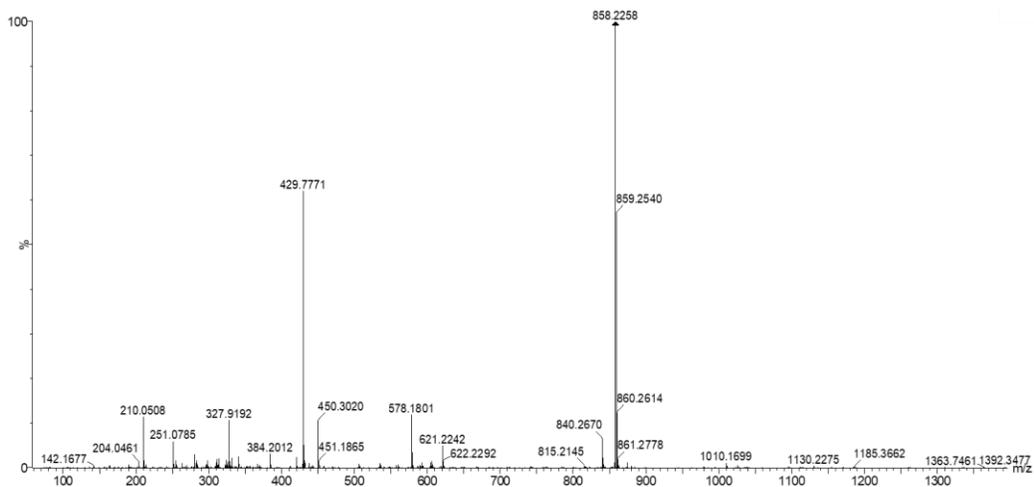
HRMS (ESI-TOF) m/z [**3n**+ H]<sup>+</sup> Calc: 858.2234; found: 858.2258.

$^1\text{H}$ -NMR (400 MHz,  $\text{CD}_3\text{OD}$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CD}_3\text{OD}$ ) and  $^{19}\text{F}$  NMR (376 MHz,  $\text{CD}_3\text{OD}$ ) of **3n**:

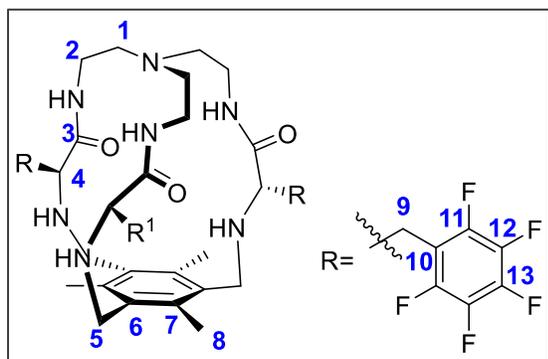




HRMS (ESI+) experimental spectrum of **3n**:



**1n (Penta-F-Phe): 3n** (205 mg, 0.156 mmol) was dissolved in ACN (60 mL). Tetrabutylammonium chloride (21 mg, 0.078 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (62.4 mg, 0.156 mmol) and potassium carbonate (432 mg, 3.12 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 97 : 3 DCM : MeOH as eluent to give **1n** as a white solid (40 mg, 0.039 mmol, 40% yield).



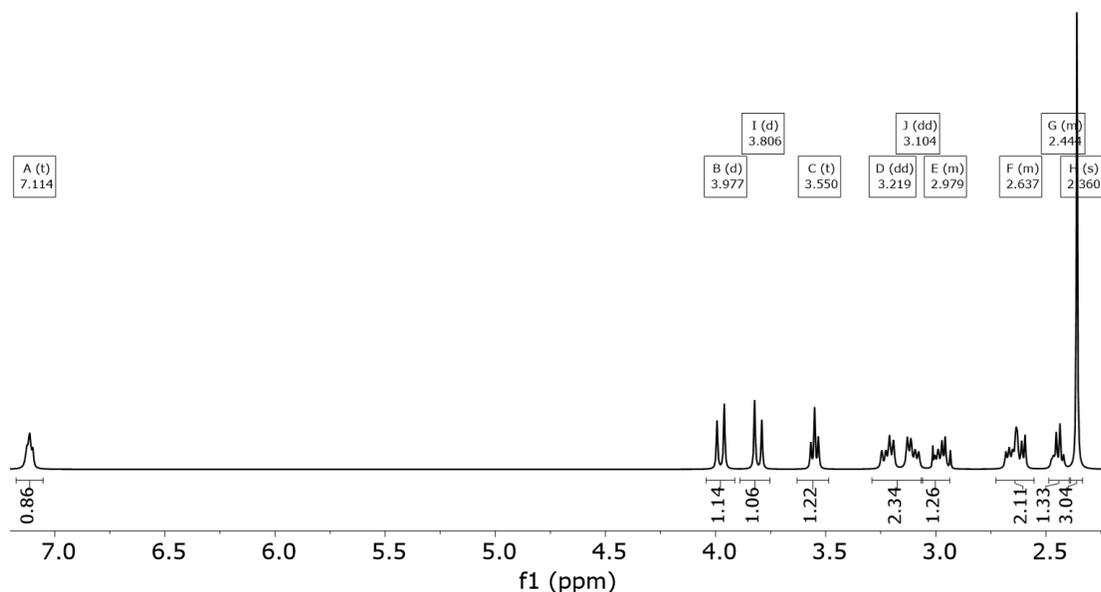
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.11 (m, 3H,  $\text{NH}_{\text{amide}}$ ), 3.89 (AB<sub>q</sub>,  $\delta_{\text{A}}=3.98$ ,  $\delta_{\text{B}}=3.81$ ,  $J_{\text{AB}} = 13.1$  Hz, 6H, H5), 3.55 (X subsystem from ABX,  $J_{\text{AX}}=7.3$ ,  $J_{\text{BX}}=6.5$  Hz, 3H, H4), 3.21 (A subsystem from ABX,  $J_{\text{AX}}=7.3$ ,  $J_{\text{AB}}=14$  Hz, 3H, H9), 3.10 (B subsystem from ABX,  $J_{\text{BX}}=6.5$ ,  $J_{\text{AB}}=14$  Hz, 3H, H9), 2.98 (m, 3H, H2), 2.64 (m, 6H, H2'+1), 2.44 (m, 3H, H1'), 2.36 (s, 9H, H8).

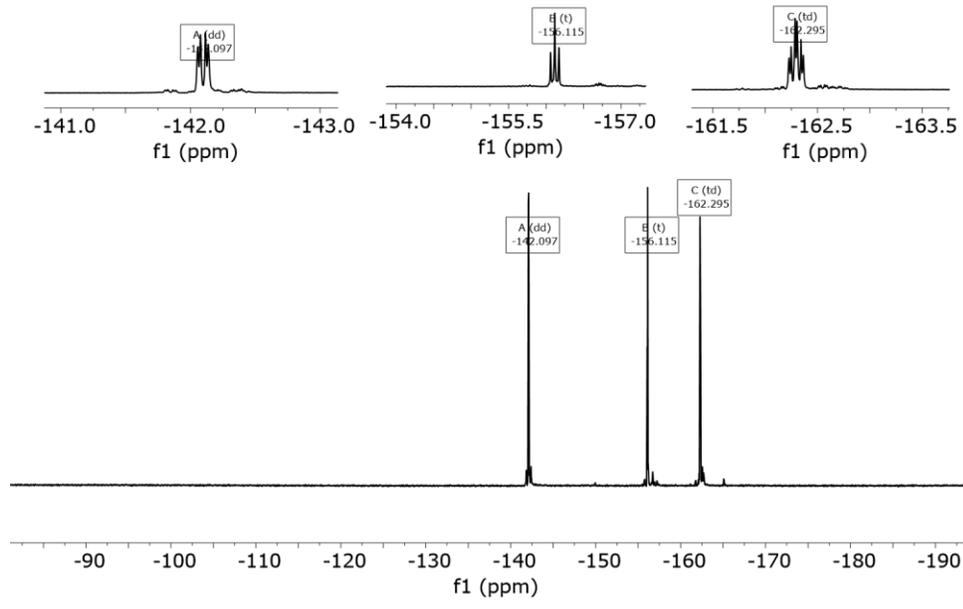
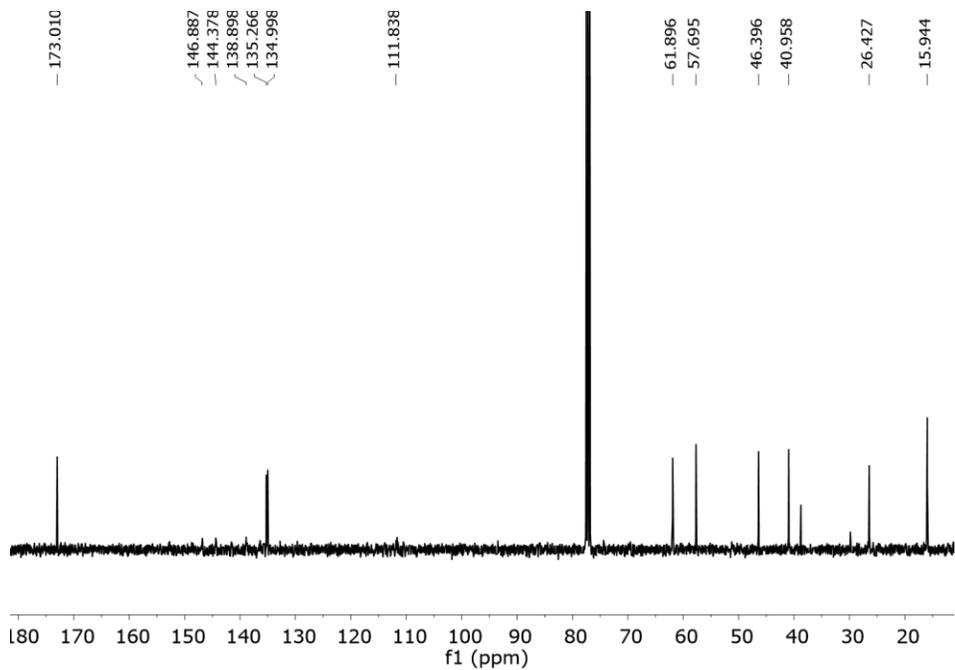
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.0 (C3), 146.9 (m, C11/12), 144.4 (m, C11/12), 138.9 (m, C13), 135.3 (C6), 135.0 (C7), 111.8 (m, C10), 61.9 (C4), 57.7 (C1), 46.4 (C5), 41.0 (C2), 26.4 (C9), 15.9 (C8).

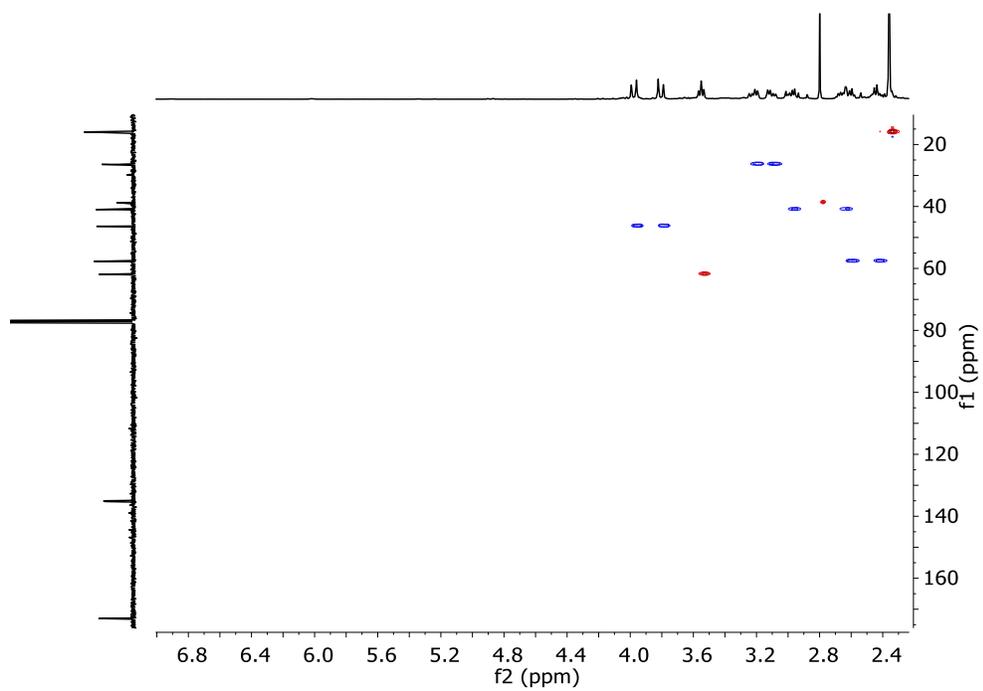
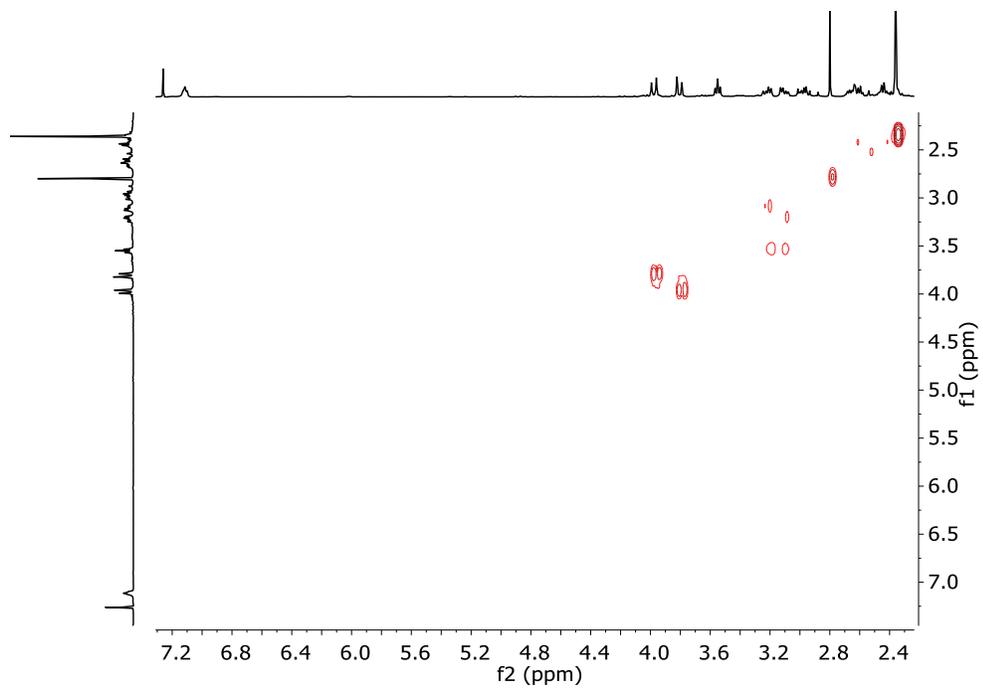
$^{19}\text{F}$  - $\{^1\text{H}\}$ -NMR (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -142.10 (dd,  $J = 22.7$ , 8.2 Hz), -156.12 (t,  $J = 20.9$  Hz), -162.29 (td,  $J = 21.9$ , 8.2 Hz).

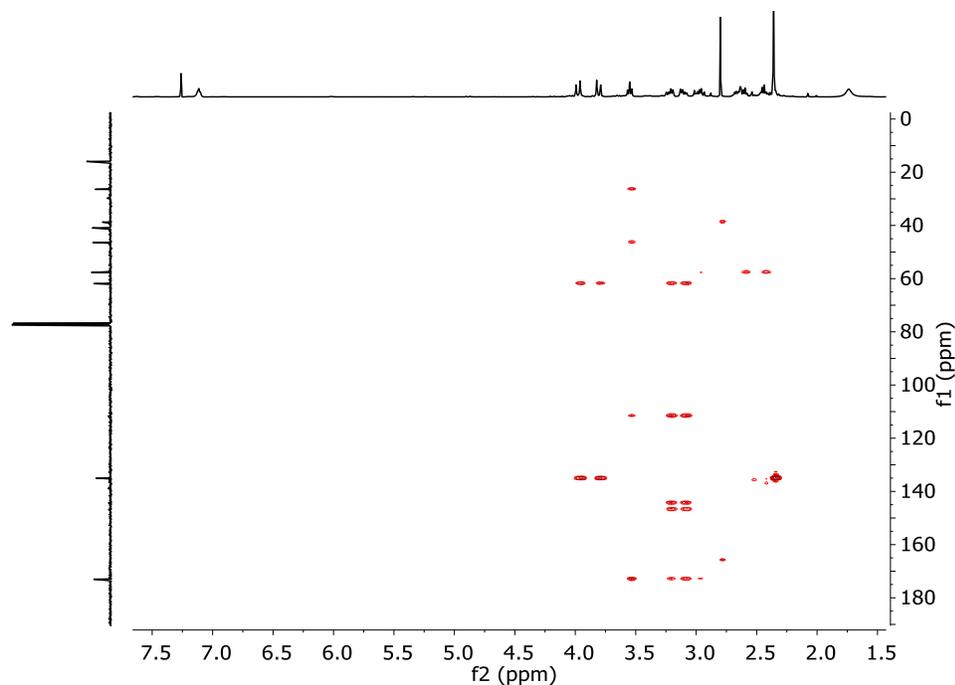
HRMS (ESI-TOF)  $m/z$  [ $3\text{n} + \text{H}$ ]<sup>+</sup> Calc: 1014.2008, found: 1014.3086.

$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$  NMR (376 MHz,  $\text{CDCl}_3$ ), COSY, HSQC and HMBC of **1n**:

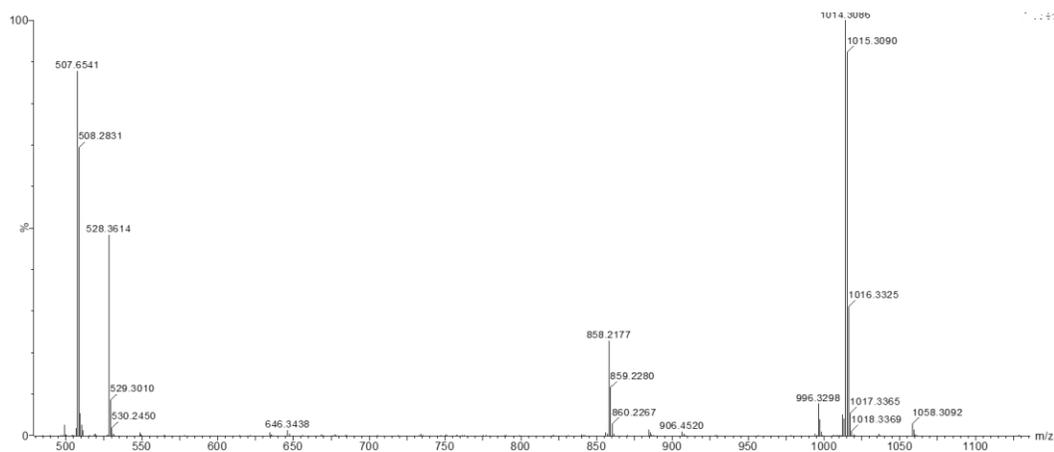








HRMS (ESI+) experimental spectrum of **1n**:



### Synthesis of **1o** (3-CF<sub>3</sub>-Phe)

**2o**: Boc-Phe(3-CF<sub>3</sub>)-OH (493 mg, 1.48 mmol) was dissolved in dry DMF (4 mL). *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl 343 mg, 1.79 mmol) and 1-hydroxybenzotriazole hydrate (HOBT, 274 mg, 1.793 mmol), *N,N*-diisopropylethylamine (DIPEA, 0.93 mL, 5.37 mmol) and tris(2-aminoethyl)amine (0.07 mL, 0.448 mmol) were added over the solution. The solution was stirred at room temperature for 16 hours, when no more conversion of the starting material was observed by TLC. The mixture was diluted with water and extracted with DCM (3 X 10 mL). Combined organic fractions were

washed with aqueous LiCl (5% w/w), dried over MgSO<sub>4</sub> and concentrated to dryness. The residue was purified by flash chromatography using 98:2 DCM:MeOH to give 216 mg of **2o** (0.197 mmol, 44% yield).

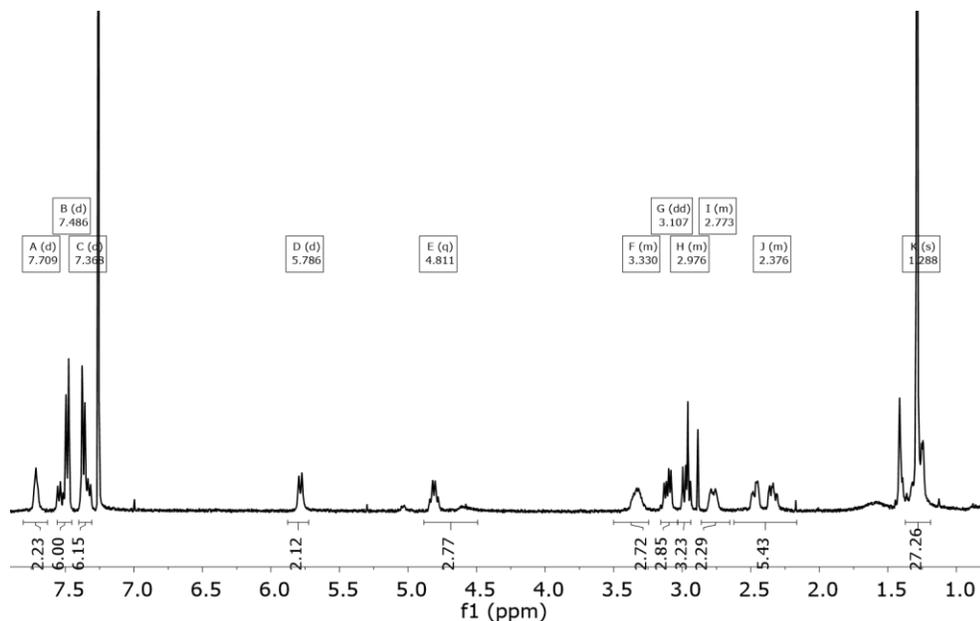
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.71 (d, *J* = 5.2 Hz, 3H), 7.49 (d, *J* = 7.9 Hz, 6H), 7.37 (d, *J* = 7.9 Hz, 6H), 5.79 (d, *J* = 9.1 Hz, 3H), 4.81 (q<sub>ap</sub>, *J* = 8.3 Hz, 3H), 3.33 (m, 3H), 3.11 (dd, *J* = 13.5, 6.6 Hz, 3H), 2.98 (m, 3H), 2.77 (m, 3H), 2.38 (m, 6H), 1.29 (s, 27H).

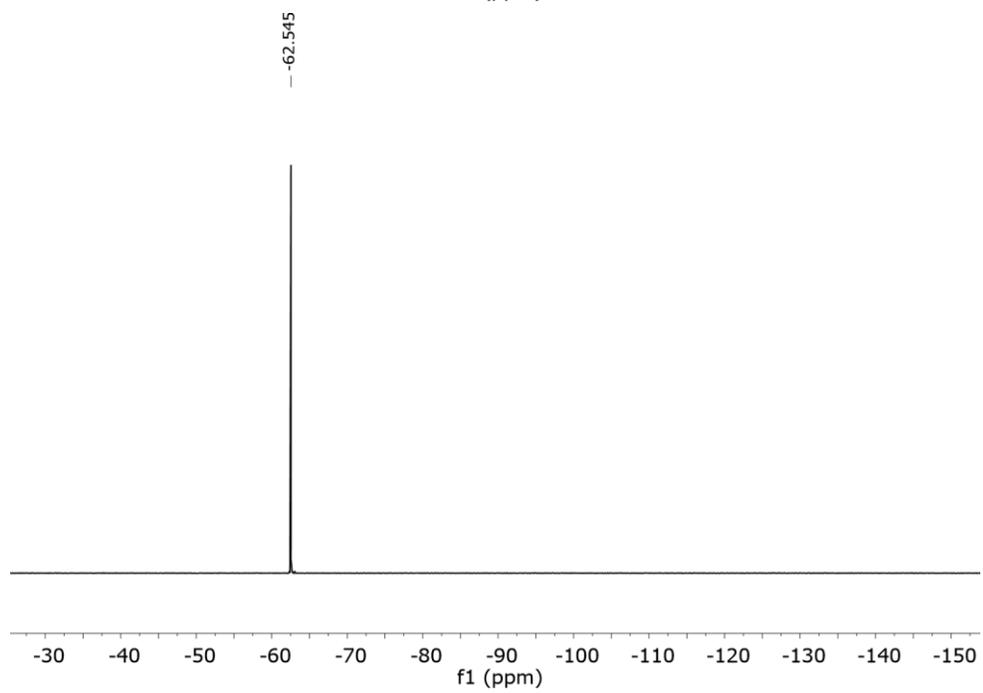
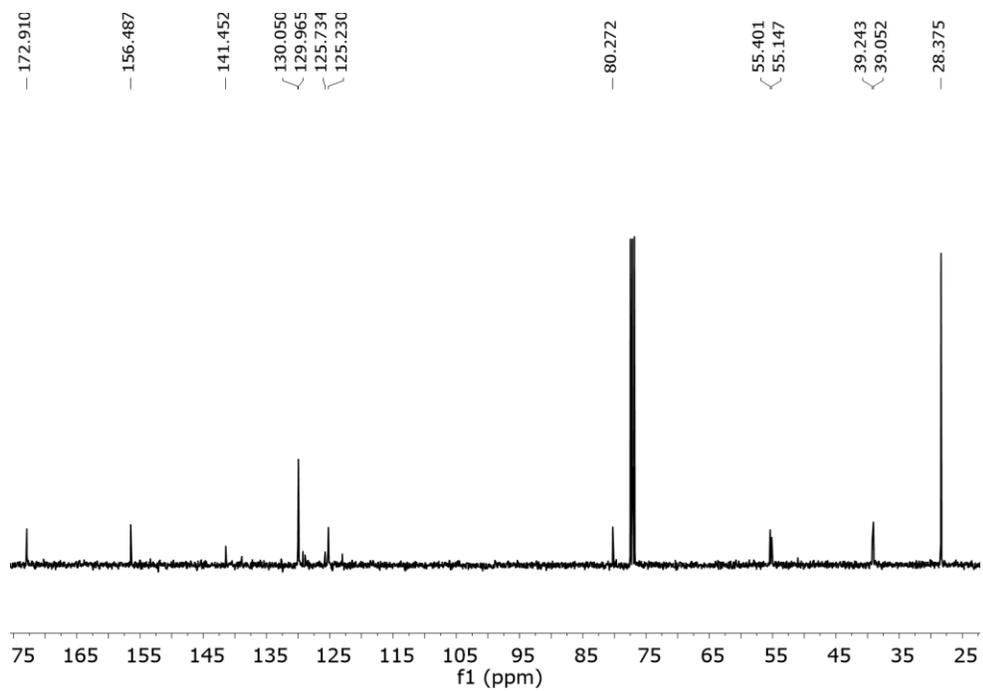
<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.9, 156.5, 141.5, 130.1, 130.0, 125.7, 125.2, 80.3, 55.4, 55.2, 39.2, 39.1, 28.4.

<sup>19</sup>F -{<sup>1</sup>H}- NMR (376 MHz, CD<sub>3</sub>OD) δ -62.55.

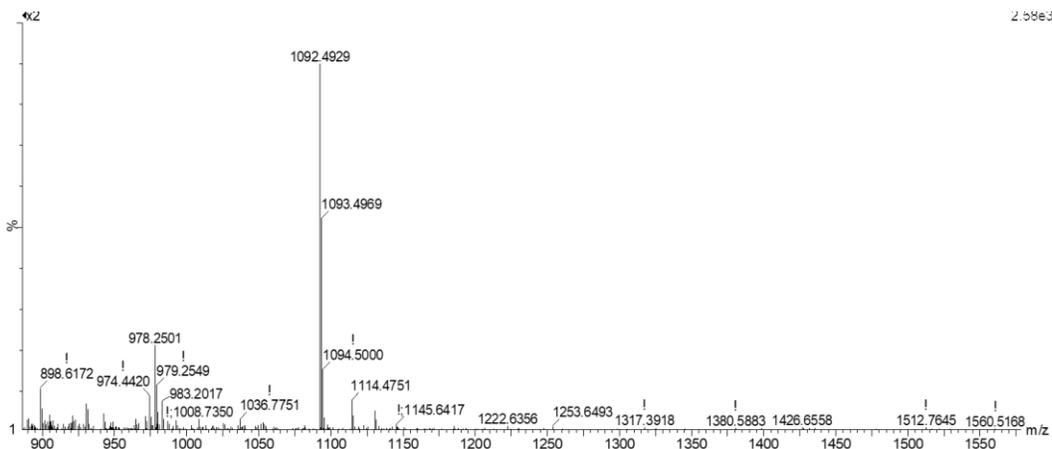
HRMS (ESI-TOF) *m/z* [**2o** + H]<sup>+</sup> Calc: 1092.4851, found: 1092.4929.

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>), <sup>13</sup>C-NMR (101 MHz, CDCl<sub>3</sub>) and <sup>19</sup>F NMR (376 MHz, CD<sub>3</sub>OD) spectra of **2o**:





HRMS (ESI+) experimental spectrum of **2o**:



**3o**: **2o** (200 mg, 0.183 mmol) was dissolved in DCM (1 mL) and triethylsilane (TES, 0.5 mL, 3.37 mmol) and TFA (1 mL) were added. The solution was stirred at room temperature for 3 hours. The solvents were then evaporated under an air current affording a yellow oil. The residue was washed several times with diethyl ether and dried affording **3o**-4TFA as a white solid. TFA was removed by dissolving the solid in 1M aqueous NaOH and extracting several times with CH<sub>2</sub>Cl<sub>2</sub>, dried and washed with hexane. Final product **3o** was isolated as a white solid (120 mg, 0.155 mmol, 85% yield).

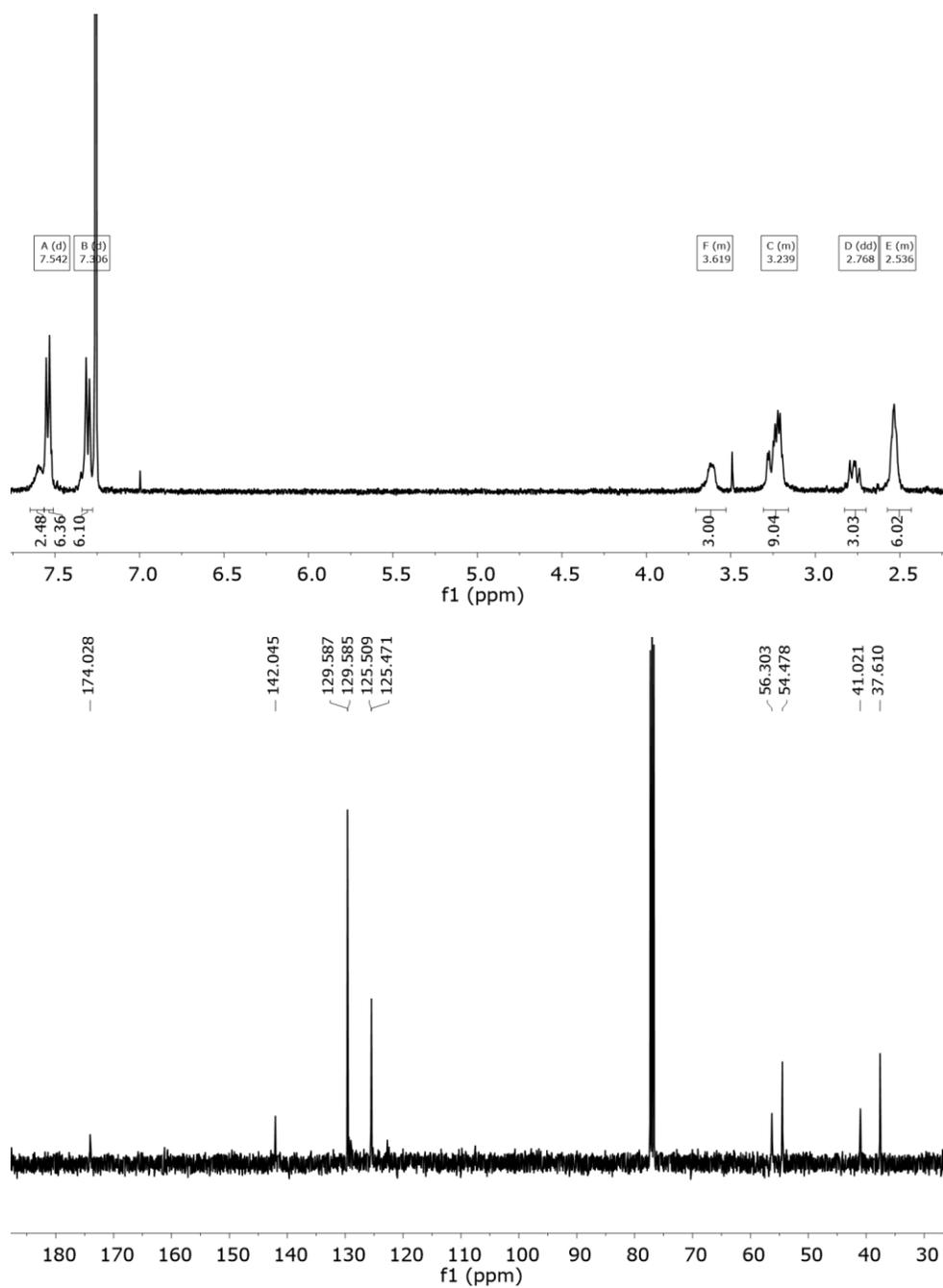
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.54 (d, *J* = 7.9 Hz, 6H), 7.31 (d, *J* = 8.0 Hz, 6H), 3.62 (m, 3H), 3.24 (m, 9H), 2.77 (dd, *J* = 13.6, 9.1 Hz, 3H), 2.54 (m, 6H).

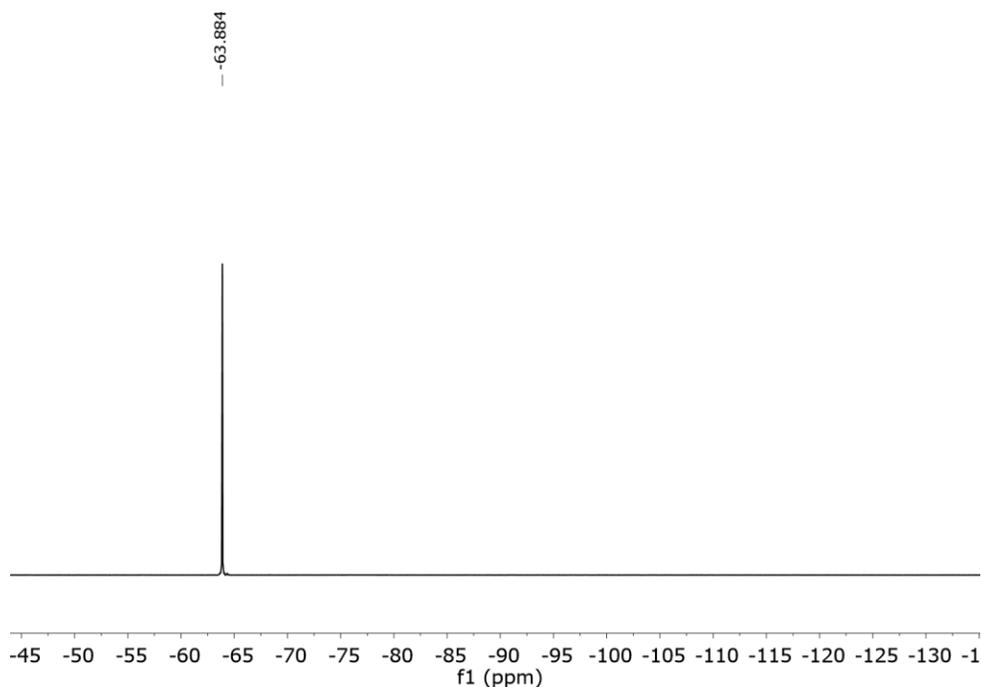
<sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ 174.0, 142.1, 129.6, 129.6, 125.5, 125.5, 56.3, 54.5, 41.0, 37.6.

<sup>19</sup>F NMR (376 MHz, CDCl<sub>3</sub>) δ -62.44.

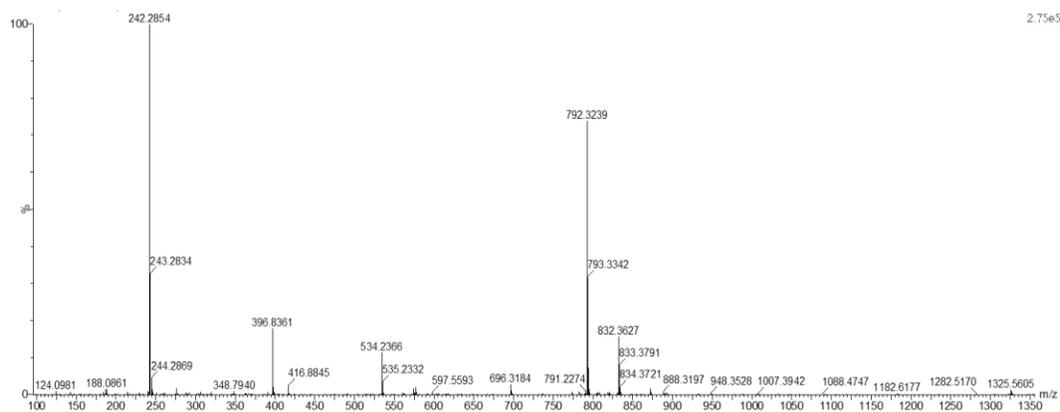
HRMS (ESI-TOF) m/z [**3o** + H]<sup>+</sup> Calc: 792.3278, found: 792.3239.

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C-NMR}$  (101 MHz,  $\text{CDCl}_3$ ) and  $^{19}\text{F-NMR}$  (376 MHz,  $\text{CD}_3\text{OD}$ ) spectra of **3o**:

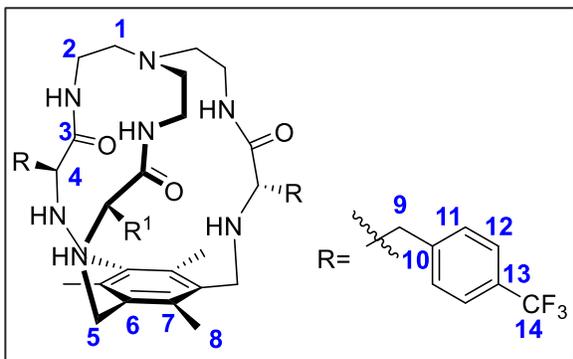




HRMS (ESI+) experimental spectrum of **3o**:



**1o (3-CF<sub>3</sub>-Phe): 3o** (120 mg, 0.151 mmol) was dissolved in acetonitrile (40 mL). Tetrabutylammonium chloride (21 mg, 0.076 mmol), 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (60.5 mg, 0.151 mmol) and potassium carbonate (419 mg, 3.03 mmol) were added over the solution. The reaction mixture was refluxed for 16 hours. After cooling down, the solution was filtered, solvent was evaporated and the resulting crude was purified by flash column chromatography using 98 : 2 DCM : MeOH as eluent to give **1o** as a white solid (70 mg, 0.0738 mmol, 50% yield).



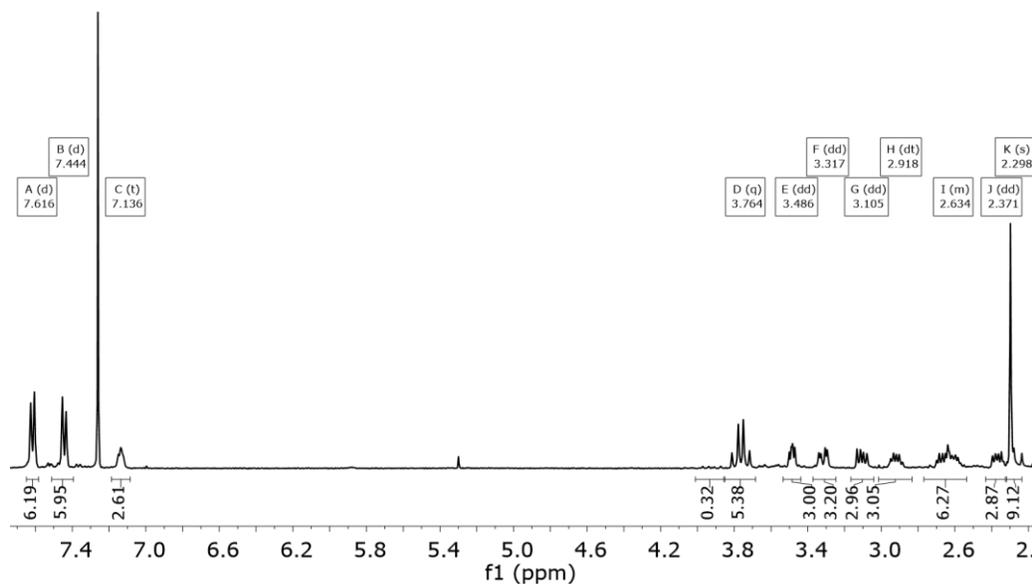
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.62 (d,  $J = 7.9$  Hz, 6H (H12)), 7.44 (d,  $J = 8.0$  Hz, 6H (H11)), 7.14 (t,  $J = 5.4$  Hz, 3H ( $\text{NH}_{\text{amide}}$ )), 3.76 (AB<sub>q</sub>,  $\delta_{\text{A}} = 3.79$ ,  $\delta_{\text{B}} = 3.73$ ,  $J_{\text{AB}} = 13.5$  Hz, 6H, (H5)), 3.49 (X subsystem from ABX,  $J_{\text{AX}} = 4.8$ ,  $J_{\text{BX}} = 7.4$  Hz, 3H (H4)), 3.32 (A subsystem from ABX,  $J_{\text{AX}} = 7.4$ ,  $J_{\text{AB}} = 13.8$  Hz, 3H, (H9)), 3.11 (B subsystem from ABX,  $J_{\text{BX}} = 7.4$ ,  $J_{\text{AB}} = 13.5$  Hz, 3H, (H9)), 2.92 (m, 3H, (H2)), 2.63 (m, 6H, (H2) overlapped with (H1)), 2.37 (m, 3H, (H1)), 2.30 (s, 9H (H8)).

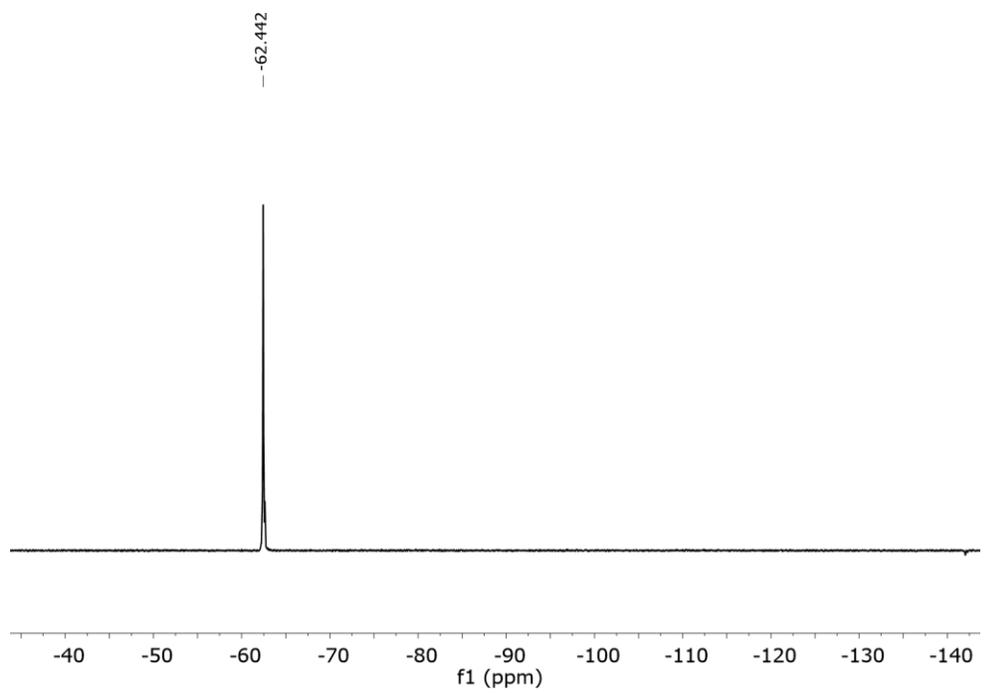
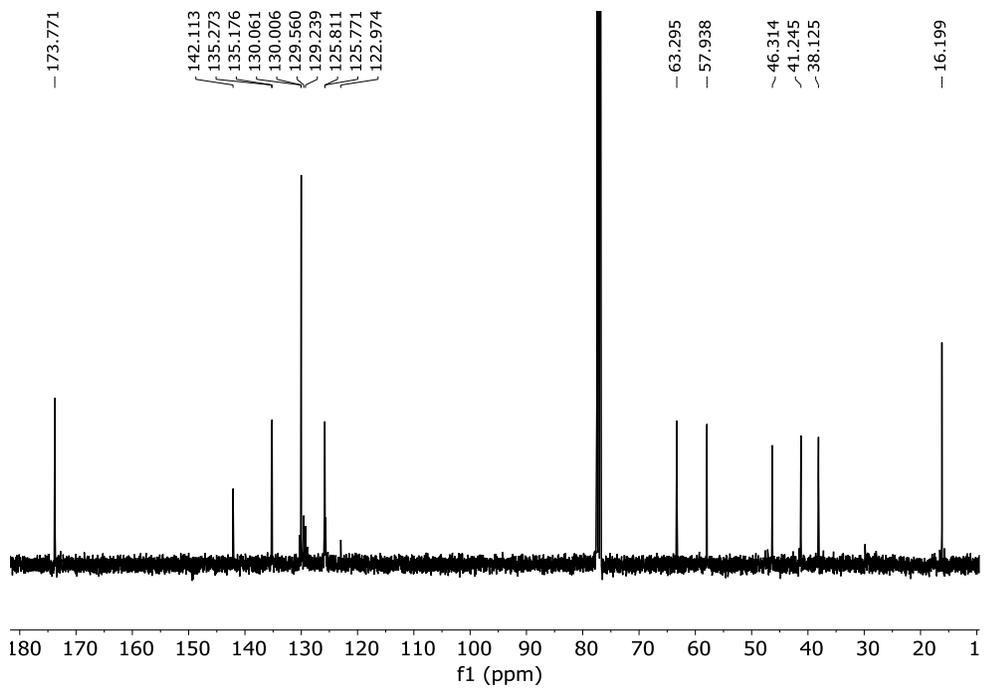
$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.8 (C3), 142.1 (C10), 135.3 (m, C-F coupling C14), 135.2 (C7), 130.1 (C11), 130.0 (C12), 129.2 (C4) 125.8 (m, C-F coupling 13), 125.7 (C6), 63.3 (C9), 57.9 (C1), 46.3 (C5), 41.2 (C2), 38.13 (C9), 16.2 (C8).

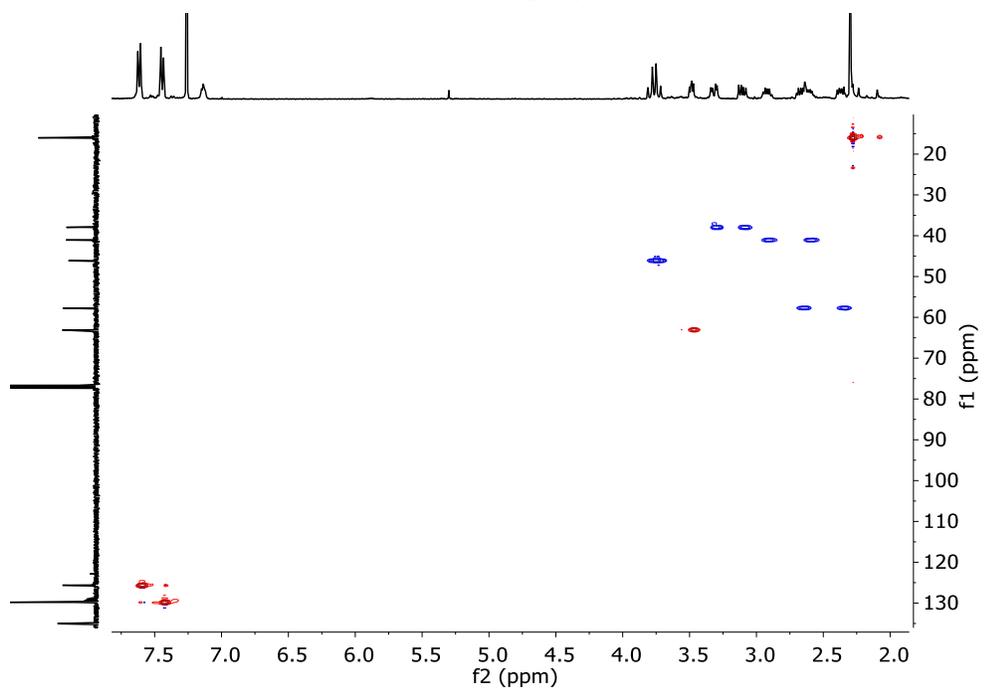
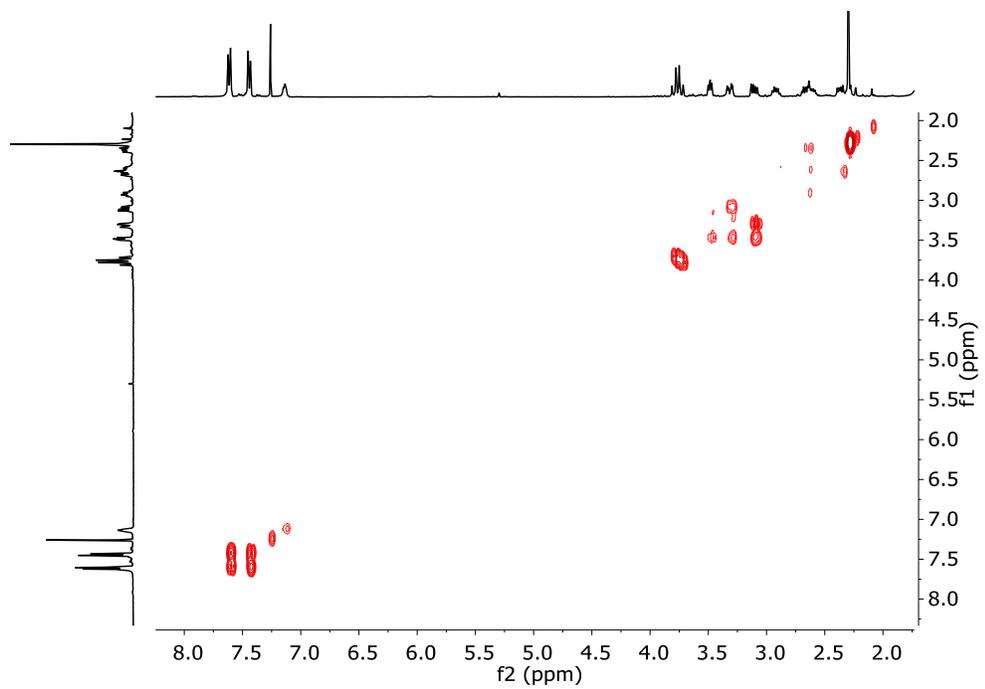
$^{19}\text{F}$  - $\{^1\text{H}\}$ -NMR (376 MHz,  $\text{CD}_3\text{OD}$ )  $\delta$  -62.44.

HRMS (ESI-TOF)  $m/z$  [ $\mathbf{1o} + \text{H}$ ]<sup>+</sup> Calc: 948.4217, found: 948.5418

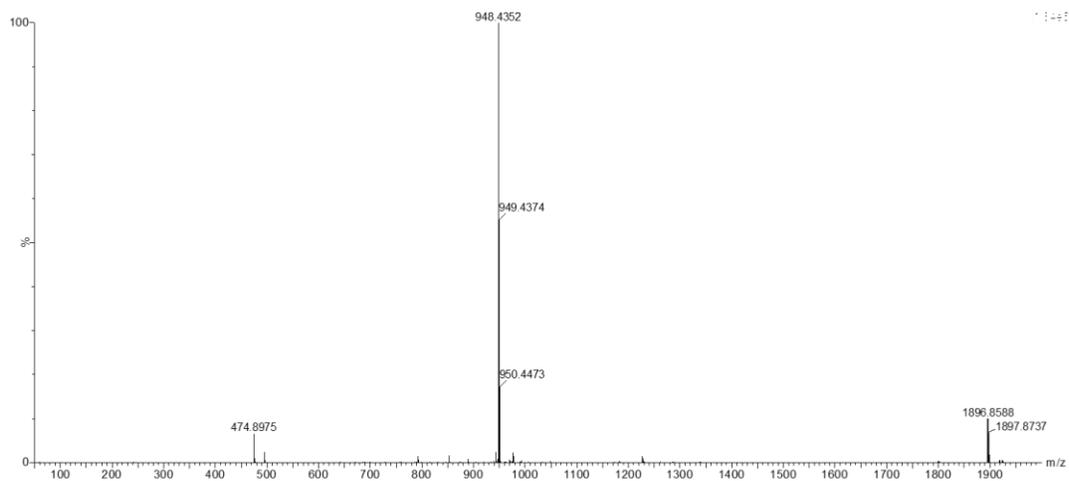
$^1\text{H}$ -NMR (400 MHz,  $\text{CDCl}_3$ ),  $^{13}\text{C}$ -NMR (101 MHz,  $\text{CDCl}_3$ ),  $^{19}\text{F}$  NMR (376 MHz,  $\text{CD}_3\text{OD}$ ), COSY and HSQC spectra of  $\mathbf{1o}$ :



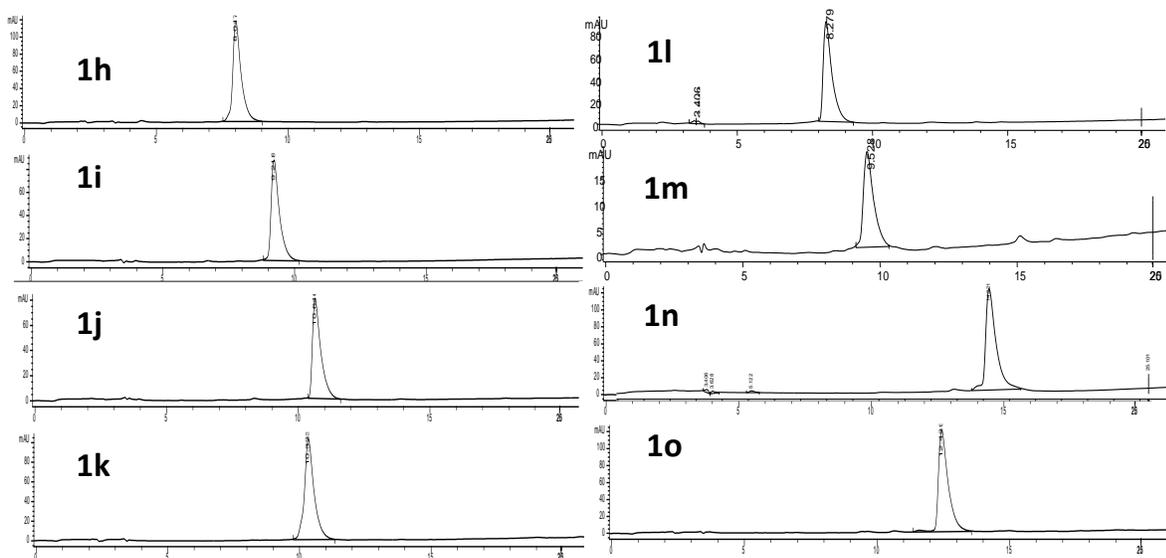




HRMS (ESI+) experimental spectrum of **1o**:



Reverse phase HPLC chromatograms of compounds **1h-o**:

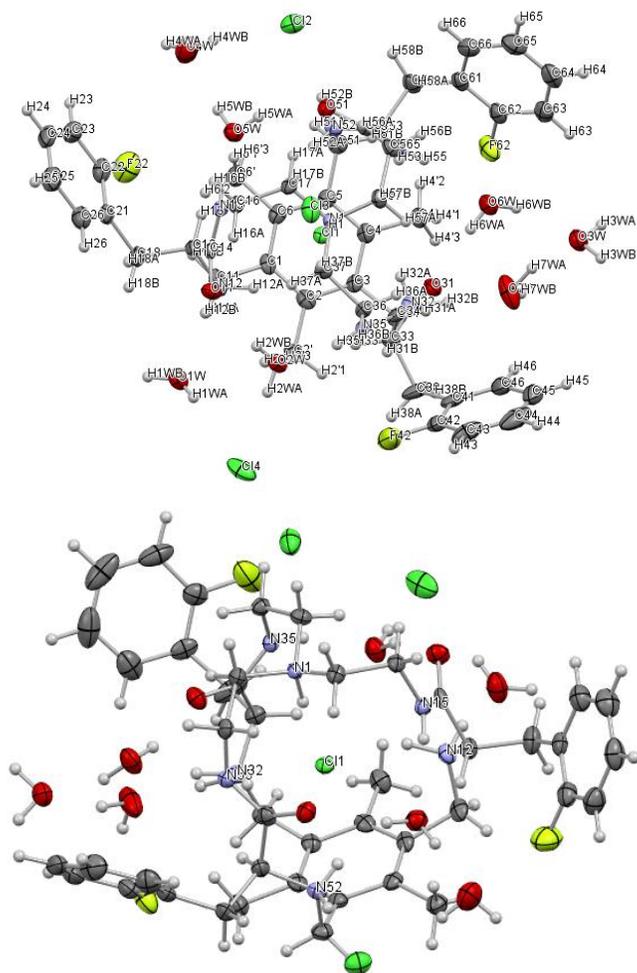


## X-ray crystal analysis of 1h·4HCl

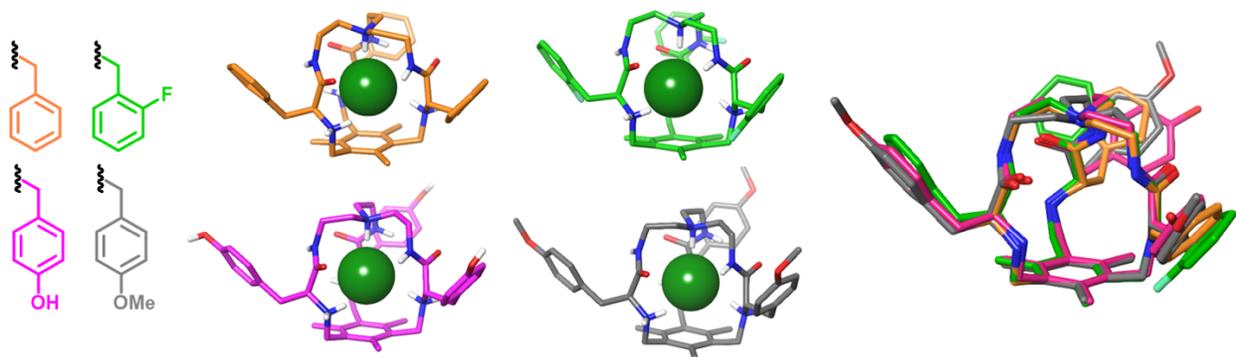
Crystals suitable for X-ray diffraction studies of compound **1h** were obtained by slowly evaporating a solution of **1h** fully protonated with TFA in acetonitrile with 5% H<sub>2</sub>O and an excess of TBACl.

**Table S1** Crystal data and structure refinement for **1h**

Identification code	n32	
Empirical formula	C <sub>45</sub> H <sub>72</sub> Cl <sub>4</sub> F <sub>3</sub> N <sub>7</sub> O <sub>10</sub>	
Formula weight	1069.89 g/mol	
Temperature	173(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	P 21 21 21	
Unit cell dimensions	a = 13.9421(4) Å	α = 90°.
	b = 17.0957(4) Å	β = 90°.
	c = 21.6955(7) Å	γ = 90°.
Volume	5171.1(3) Å <sup>3</sup>	
Z	4	
Density (calculated)	1.374 Mg/m <sup>3</sup>	
Absorption coefficient	0.301 mm <sup>-1</sup>	
F(000)	2264	
Crystal size	0.160 x 0.150 x 0.120 mm <sup>3</sup>	
Theta range for data collection	3.293 to 25.651°.	
Index ranges	-16<=h<=16, -18<=k<=20, -26<=l<=26	
Reflections collected	47047	
Independent reflections	9693 [R(int) = 0.0488]	
Completeness to theta = 10.000°	95.2 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.000 and 0.818	
Refinement method	Full-matrix least-squares on F <sup>2</sup>	
Data / restraints / parameters	9693 / 6 / 721	
Goodness-of-fit on F <sup>2</sup>	1.099	
Final R indices [I>2sigma(I)]	R1 = 0.0483, wR2 = 0.1050	
R indices (all data)	R1 = 0.0568, wR2 = 0.1088	
Absolute structure parameter	0.00(3)	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.649 and -0.236 e.Å <sup>-3</sup>	



Perspective view of **1h** with the atoms labels. Displacement parameters are drawn at the 50 % probability level. Perspective view of **1h** with the most important atoms labels.



**Figure S1.** A geometrical comparison and overlay of the crystal structures of **1h** (this work, green) with those of **1a** (orange),<sup>4,5</sup> **1b** (pink)<sup>6</sup> and **1d** (grey)<sup>6</sup> showed high structural similarity

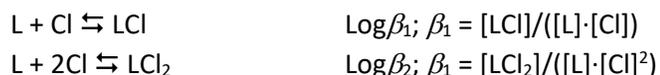
## **<sup>1</sup>H NMR titration of compounds 1h, 1i, 1m and 1l with TBACl**

The titrations were performed with the cage receptor as the fully protonated molecules, using trifluoroacetic acid (TFA). The corresponding tetra-TFA salts were prepared by dissolving each compound in methanol and adding an excess of trifluoroacetic acid, followed by the solvent evaporation and drying in vacuum.

Stock solutions of the cage were prepared by weighting the corresponding amount of the receptor and reaching a final concentration between 1 and 2 mM. The solvent used was 95:5 CD<sub>3</sub>CN:H<sub>2</sub>O, since this mixture generally allows a good solubility during the titration experiment and renders reasonably sharp and well-defined <sup>1</sup>H NMR spectra. Besides, under these conditions, the amide proton is detectable during the titration experiments. Additionally, a stock solution of the titrant containing 0.1 M TBACl was prepared by dissolving the salt in the stock solution of the cage. Thus, for each experiment, the solution of the titrant will be 0.1 M in TBACl and 0.001-0.002 M in the receptor therefore maintaining the concentration of the cage constant during the titration experiment. The stock solution of the cage was introduced in a NMR tube and the <sup>1</sup>H NMR spectrum (500 MHz, 303 K) was acquired, then small volumes of the stock solution of the titrant were added and the <sup>1</sup>H NMR spectrum recorded after each addition.

Different signals changed upon addition of chloride anion, and their variations were globally fitted using HypNMR 2008 version 4.0.71 software.

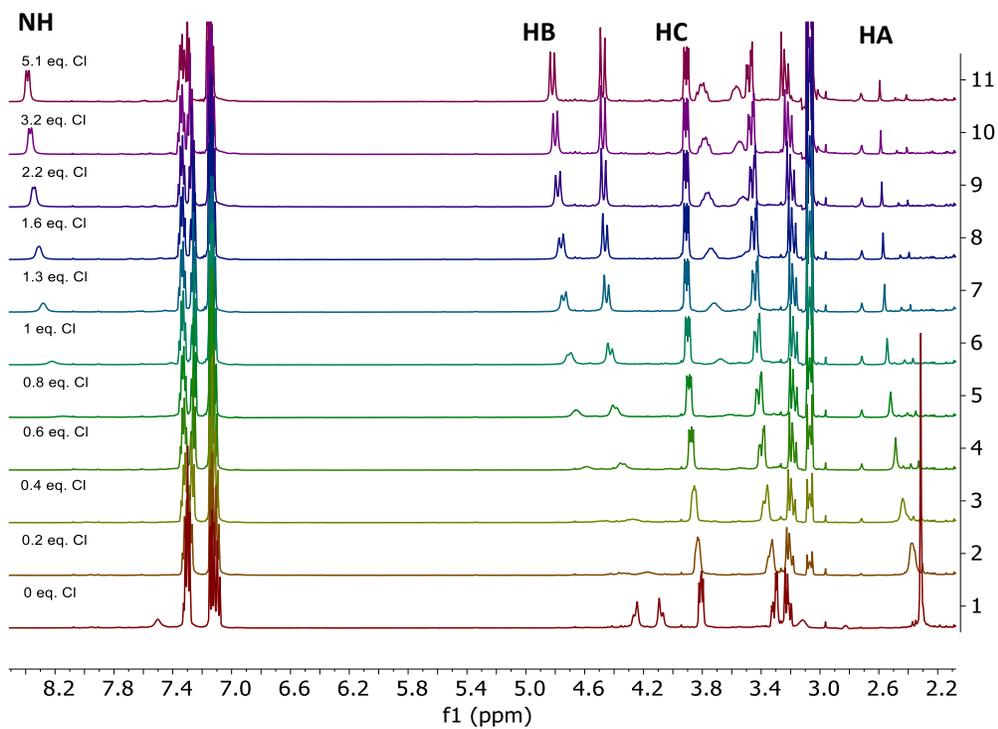
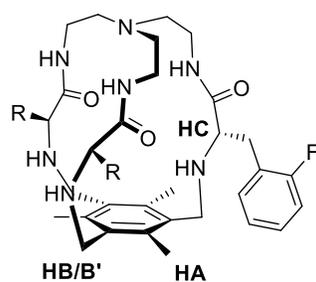
We tried different binding modes starting from the simplest 1:1 cage:chloride stoichiometry of the supramolecular complex, for which the fitting was unsatisfactory in some the cases. For these, the simplest binding mode that led to a satisfactory fitting of all the proton signals corresponded to the formation of both 1:1 and 1:2 cage:chloride complexes, with the stability constants ( $\beta_1$  and  $\beta_2$ ) defined as (L: tetraprotonated cage; Cl: chloride):



Because species with varied stoichiometry are formed in solution, for a suitable comparison of the systems, we calculated the BC50<sup>0</sup> parameter, using the BC50 calculator version 2.37.1 program.

Following, we show the stacked plot of the NMR spectra for the titration experiments, the corresponding data set introduced (experimental) and obtained (fit) during the fitting process, the output values (both  $\log \beta$  and BC50<sup>0</sup>) for the binding for every supramolecular complex, the plot of the experimental (symbols) and the fitted (lines) values of the chemical shifts, and the plot of the simulated species distribution obtained.

1h

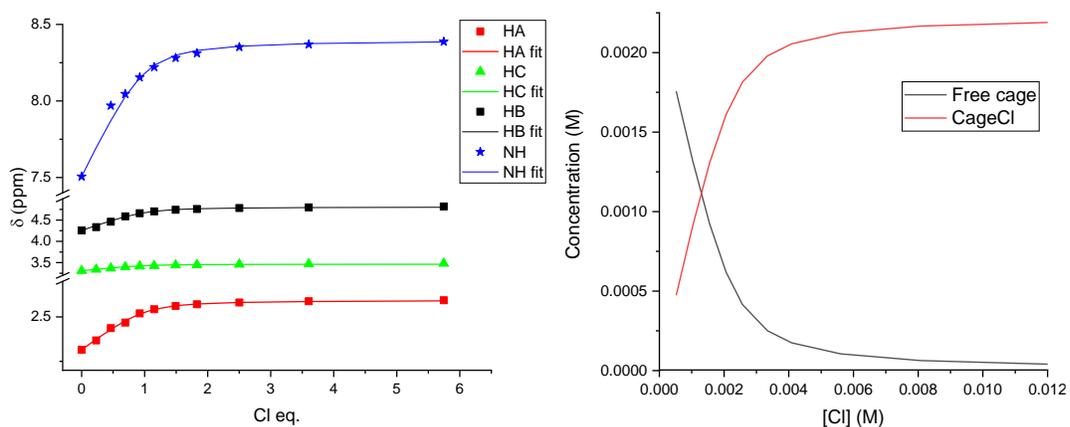


**Figure S2.** Stacked  $^1\text{H}$  NMR spectra for the titration of **1h**.

Data set of **1h** titration:

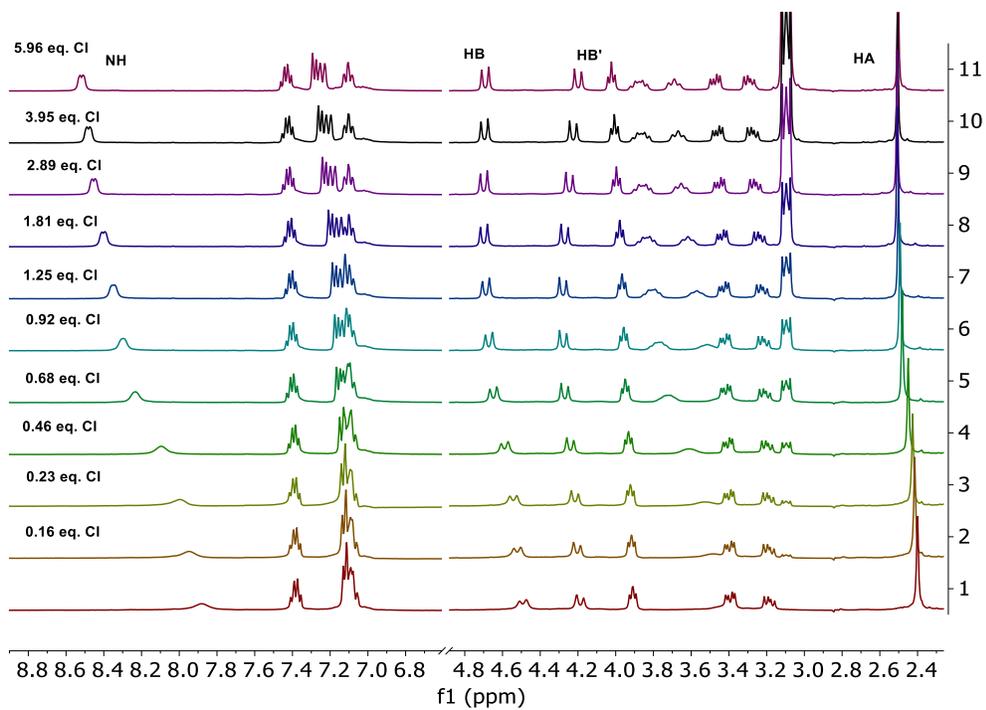
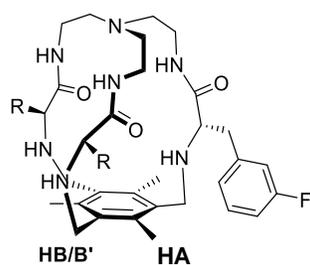
[Cl] (M)	NH	NH fit	HB	HB fit	HC	HC fit	HA	HA fit
0.0005	7.696	7.696	4.335	4.374	3.339	3.343	2.368	2.374
0.0010	7.970	7.871	4.461	4.482	3.368	3.373	2.437	2.429
0.0016	8.046	8.030	4.588	4.581	3.396	3.401	2.468	2.479
0.0021	8.154	8.152	4.661	4.656	3.415	3.422	2.520	2.517
0.0026	8.222	8.234	4.705	4.707	3.428	3.436	2.543	2.543
0.0033	8.281	8.299	4.742	4.748	3.441	3.447	2.562	2.563
0.0041	8.311	8.330	4.760	4.767	3.448	3.452	2.571	2.573
0.0056	8.352	8.358	4.782	4.784	3.458	3.457	2.581	2.581
0.0081	8.370	8.375	4.800	4.794	3.468	3.460	2.588	2.587

Results of the HypNMR fitting:  $\text{Log}\beta = 3.76 \pm 0.094$  (1 : 1),  $\text{BC}_{50} = 173 \pm 4 \mu\text{M}$



**Figure S3.** Plot of the experimental (symbols) and fitting (lines) data of **1h** titration (left). Species distribution as a function of the chloride concentration of **1h** titration (right).

1i

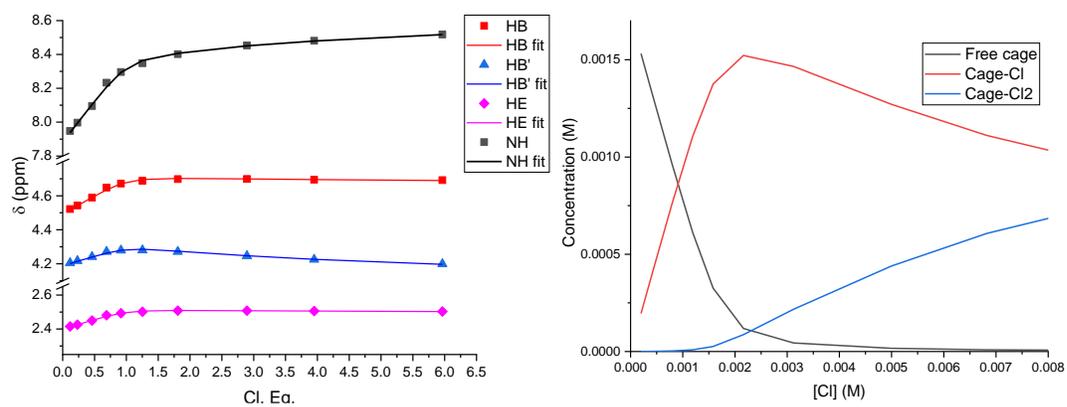


**Figure S4.** Stacked  $^1\text{H}$  NMR spectra for the titration of **1i**.

Data set of **1i** titration

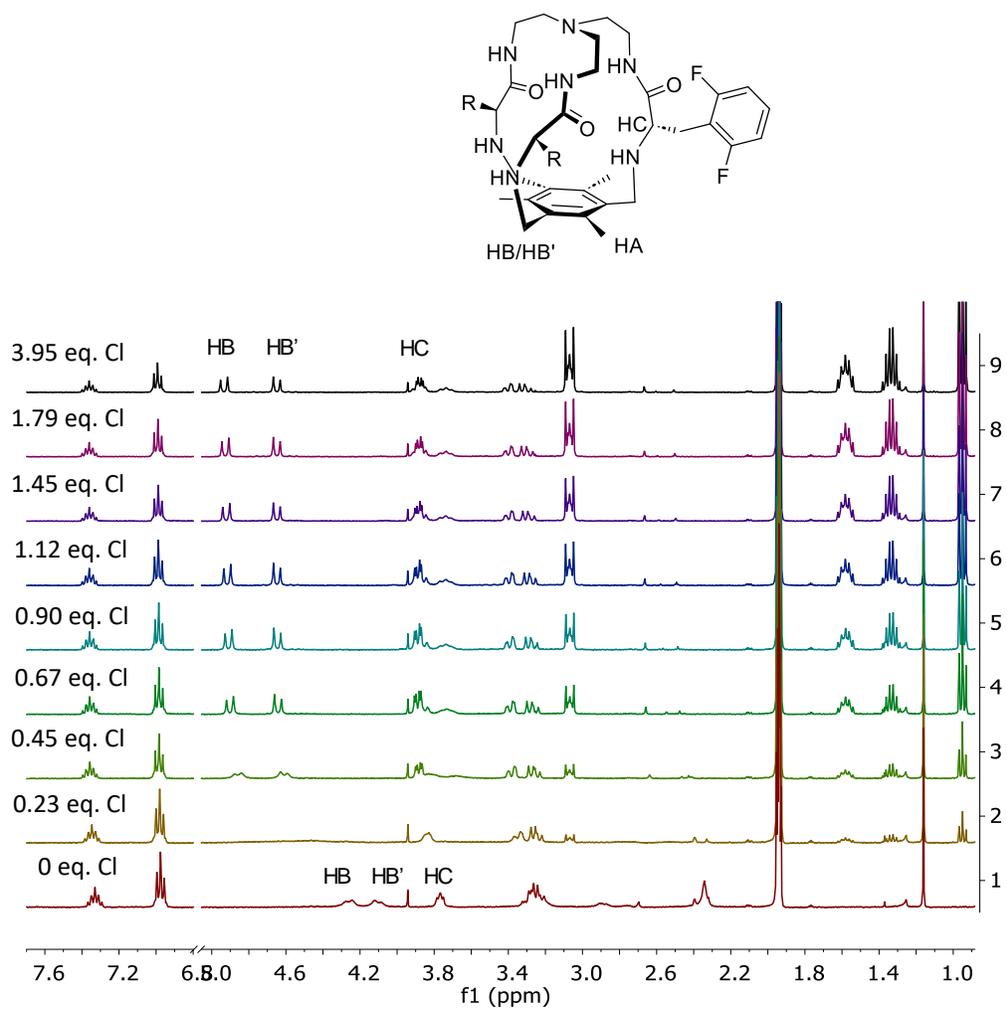
[Cl] M	NH	NH fit	HB	HB fit	HB'	HB' fit	HA	HA fit
0.0002	7.948	7.938	4.522	4.517	4.205	4.2012	2.415	2.413
0.0004	7.997	7.9943	4.543	4.5418	4.217	4.2142	2.426	2.4259
0.0008	8.095	8.1042	4.589	4.5899	4.241	4.2393	2.45	2.4509
0.0012	8.233	8.2076	4.648	4.6347	4.271	4.2622	2.481	2.4743
0.0016	8.296	8.2938	4.672	4.671	4.279	4.2794	2.493	2.4931
0.0022	8.348	8.3642	4.688	4.696	4.28	4.2855	2.502	2.5061
0.0031	8.402	8.407	4.698	4.7019	4.271	4.2737	2.508	2.5092
0.0050	8.453	8.4507	4.699	4.6992	4.246	4.2469	2.508	2.5077
0.0068	8.481	8.4796	4.695	4.6955	4.225	4.2258	2.506	2.5057

Results of the HypNMR fitting:  $\text{Log}\beta_1 = 4.4 \pm 0.1$  (1:1),  $\text{Log}\beta_2 = 6.5 \pm 0.2$  (1 : 2),  $\text{BC}_{50} = 36.8 \pm 8 \mu\text{M}$



**Figure S5.** Plot of the experimental (symbols) and fitting (lines) data of **1i** titration (left). Species distribution as a function of the chloride concentration of **1i** titration (right).

11

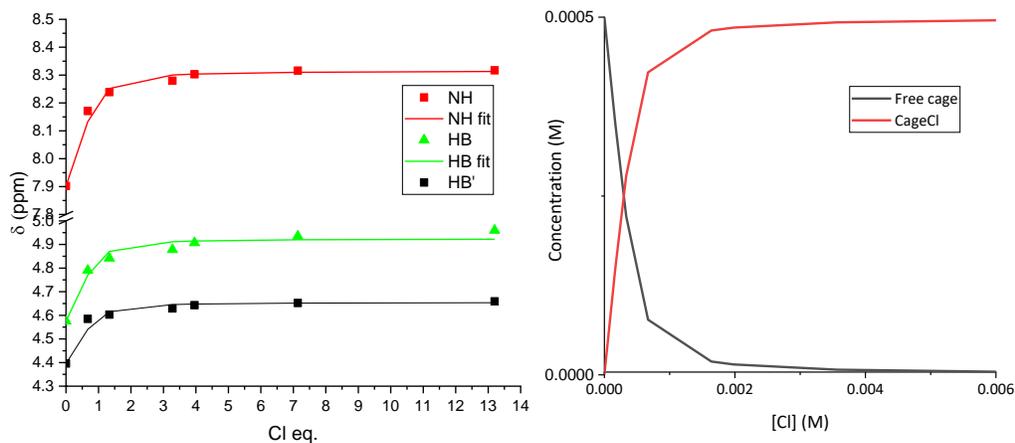


**Figure S6.** Stacked  $^1\text{H}$  NMR spectra for the titration of **11**.

Data set of **1I** titration:

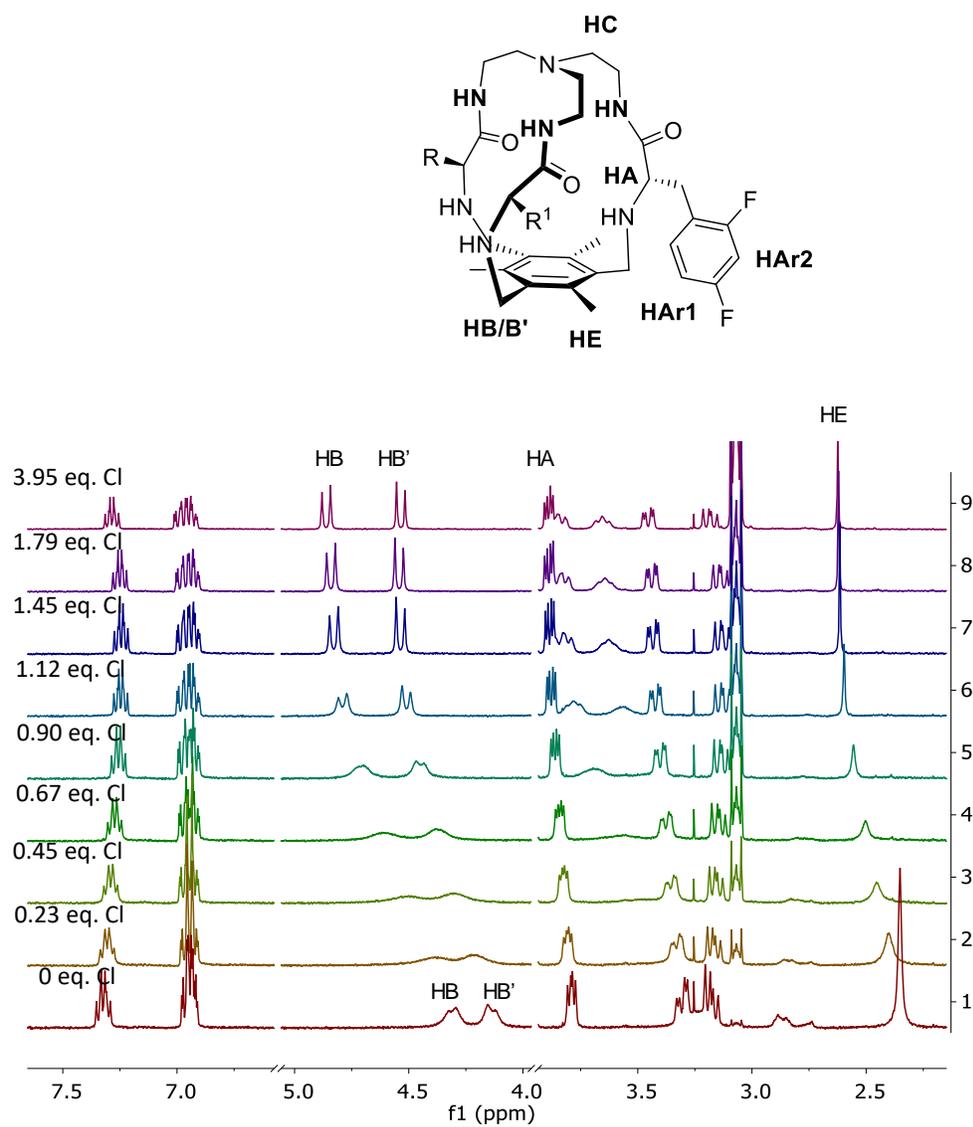
[Cl]	Cl eq.	NH	NH fit	HB	HB fit	HB'	HB' fit
0	0	7.902	7.902	4.575	4.575	4.396	4.396
0.000349	0.67	8.171	8.1326	4.79	4.7698	4.585	4.5404
0.000687	1.3378	8.239	8.252	4.842	4.8707	4.603	4.6153
0.00137	3.28	8.28	8.3004	4.878	4.9116	4.629	4.6456
0.001879	3.96	8.303	8.3038	4.907	4.9145	4.643	4.6477
0.00256	7.14	8.316	8.3099	4.935	4.9197	4.652	4.6515
0.00306	13.2	8.317	8.3129	4.96	4.9221	4.659	4.6534

Results of HypNMR fitting:  $\text{Log}\beta = 4.3 \pm 0.13$ ;  $\text{BC}_{50} = 50 \pm 15 \mu\text{M}$



**Figure S7.** Plot of the experimental (symbols) and fitting (lines) data of **1I** titration (left). Species distribution as a function of the chloride concentration of **1I** titration (right).

1m

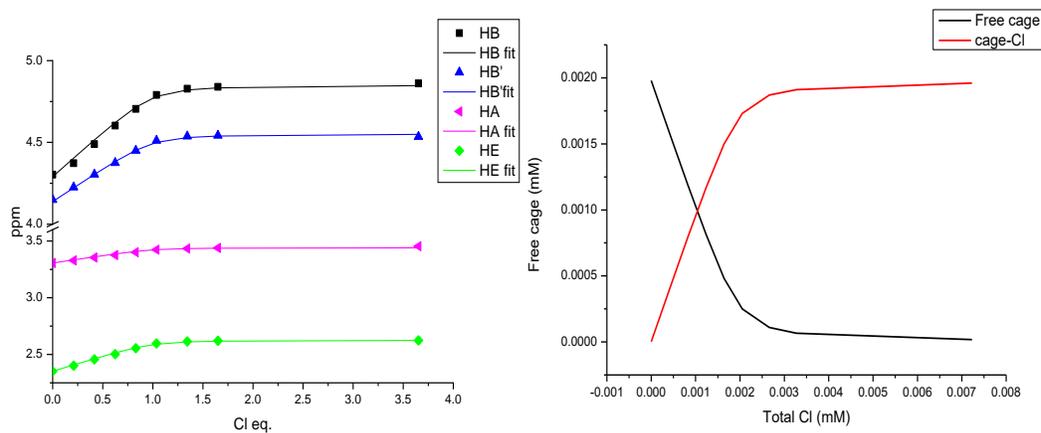


**Figure S8.** Stacked  $^1\text{H}$  NMR spectra for the titration of **1m**.

Data set of **1m** titration:

[Cl]	Cl eq.	HB	HB fit	HB'	HB'fit	HA	HA fit	HE	HE fit
0.000	0.000	4.303	4.2903	4.149	4.1382	3.306	3.3061	2.352	2.352
0.000	0.209	4.373	4.4039	4.226	4.222	3.329	3.3334	2.401	2.408
0.001	0.417	4.489	4.515	4.304	4.3041	3.354	3.3601	2.457	2.462
0.001	0.624	4.603	4.6215	4.375	4.3827	3.376	3.3857	2.503	2.514
0.001	0.830	4.705	4.7154	4.45	4.452	3.401	3.4083	2.555	2.560
0.002	1.036	4.79	4.7807	4.511	4.5002	3.423	3.424	2.597	2.591
0.002	1.344	4.828	4.8207	4.537	4.5297	3.433	3.4336	2.615	2.611
0.002	1.650	4.84	4.8332	4.542	4.539	3.439	3.4367	2.621	2.617
0.005	3.652	4.861	4.8471	4.535	4.5492	3.454	3.44	2.624	2.624

Results of HypNMR fitting:  $\text{Log}\beta = 4.3 \pm 0.13$ ;  $\text{BC}_{50} = 50 \pm 15 \mu\text{M}$



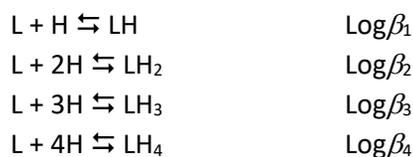
**Figure S9.** Plot of the experimental (symbols) and fitting (lines) data of **1m** titration (left). Species distribution as a function of the chloride concentration of **1m** titration (right).

## **pK<sub>a</sub> determination by fluorescence spectroscopy**

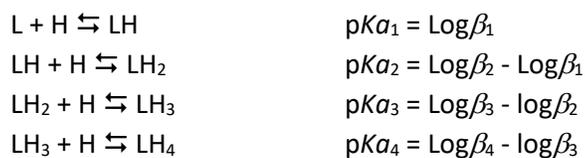
A 100-200  $\mu\text{M}$  solution of the corresponding cage in MeOH/H<sub>2</sub>O (40:60, 11 mL) containing 150 mM of KClO<sub>4</sub> (to maintain the ionic strength constant through the assay) was adjusted to different pH's (3-11) using small and controlled volumes of NaOH or HClO<sub>4</sub>. The fluorescence emission of the solution was measured at room temperature for each pH using  $\lambda_{\text{ex}} = 270 \text{ nm}$  (maximum absorbance of the phenylalanine) and  $\lambda_{\text{em}} = 290\text{-}420 \text{ nm}$ . The fluorescence emission spectra changes upon basification of the solution are associated to changes in the protonation of the cage.

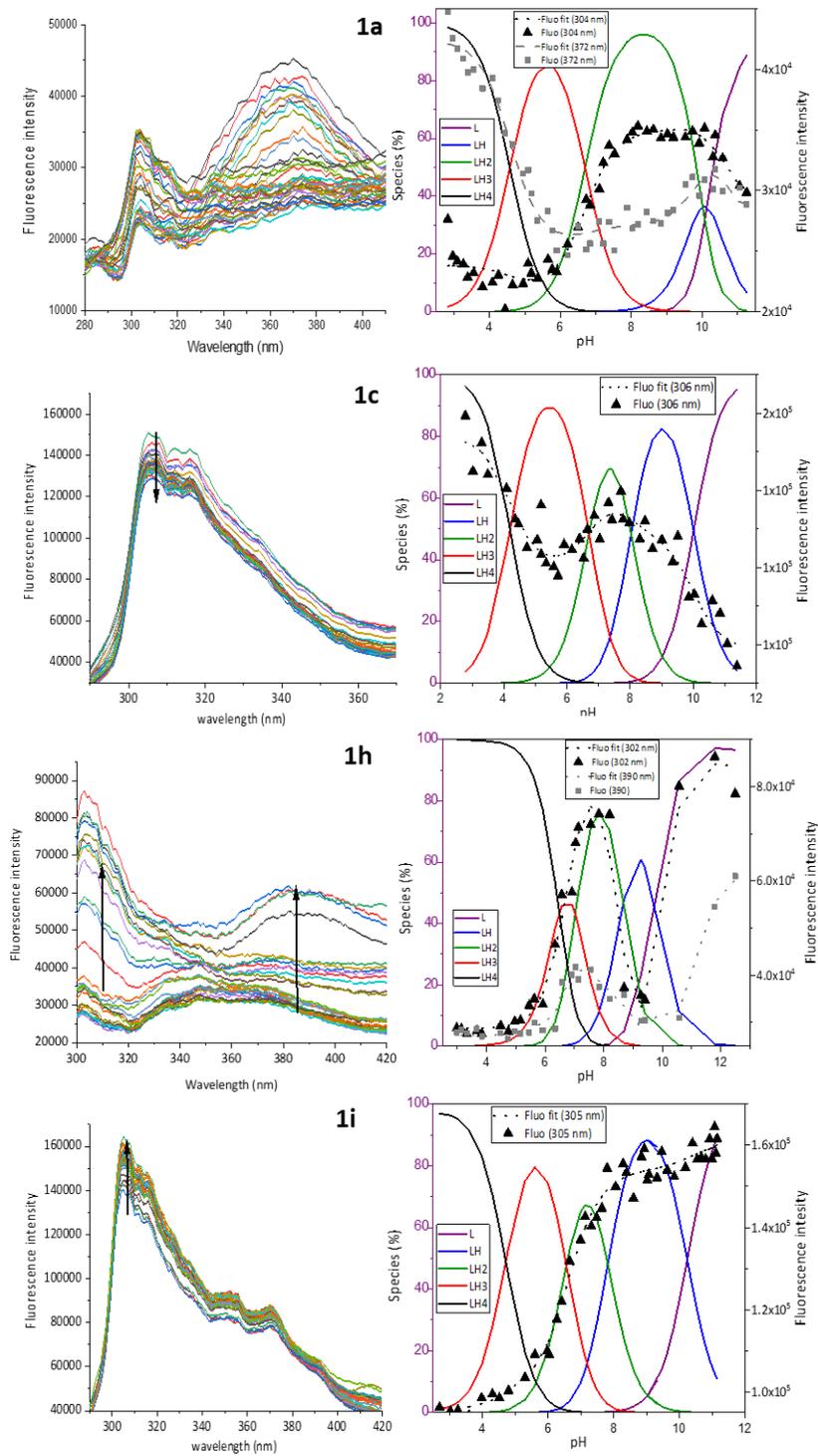
Titration data were fitted using HypSpec 2008 version 1.1.33 software to the protonation model for these systems (charges have been omitted for simplicity):

Global stability constants for each protonated species:



Acid-base properties (*pKa* values) for stepwise protonation processes:





**Figure S10.** Fluorescence spectra of **1a** (Phe), **1c** (4F-Phe), **1h** (2F-Phe) and **1i** (2F-Phe) at different pH values (left, arrows indicate increase of pH). Data fitting with HypSpec including selected wavelengths and the distribution of protonated species as a function of pH.

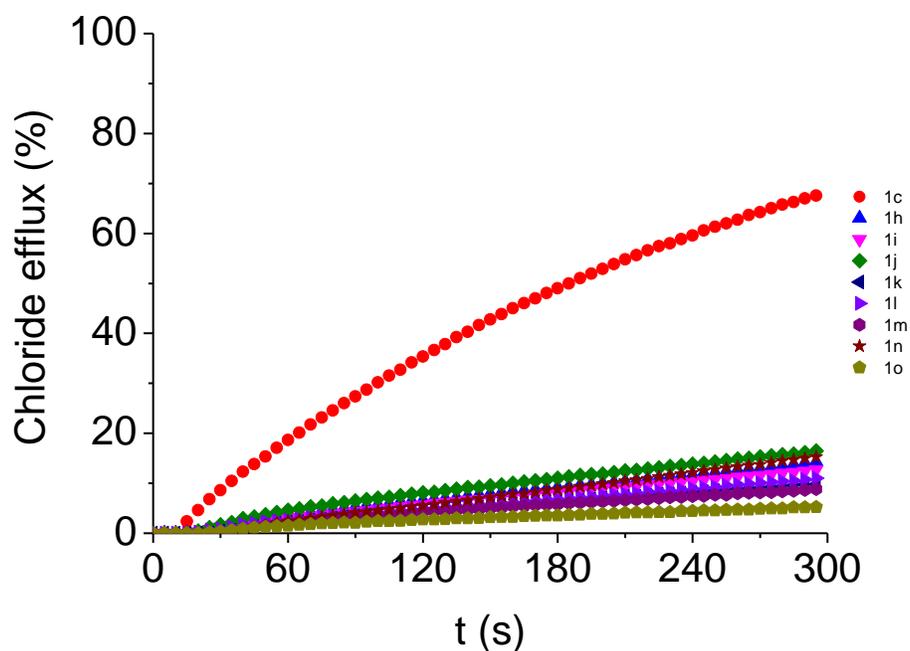
## **Transmembrane anion transport experiments in vesicles**

### **Preparation of phospholipid vesicles**

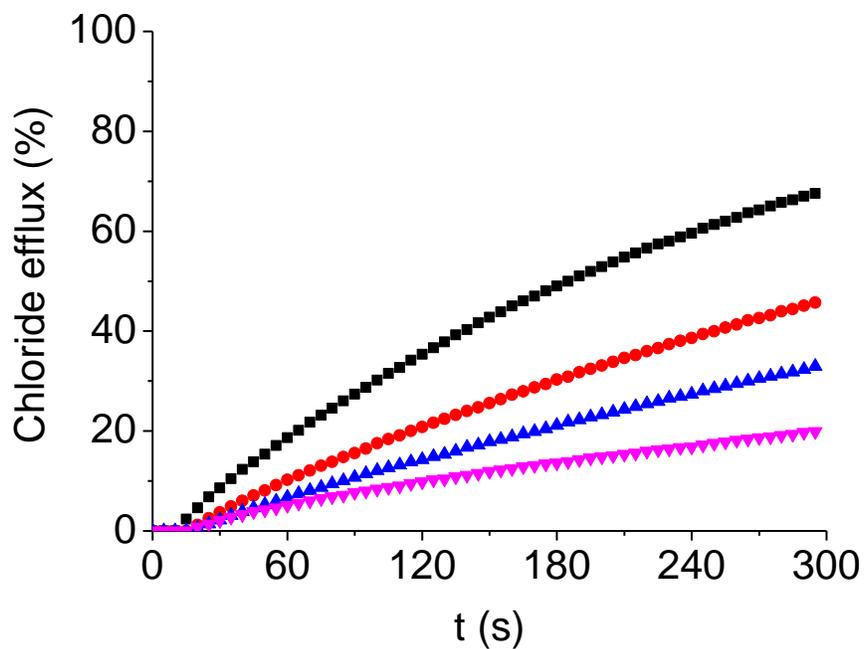
A chloroform solution of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocoline (POPC) (20 mg/mL) (Sigma Aldrich) (or a 7:3 POPC:cholesterol mixture) was evaporated to dryness employing a rotary evaporator and the resulting film was dried under high vacuum for eight hours. Different aqueous solutions were used to rehydrate the lipid film: (a) ISE assays: 489 mM NaCl, 5 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 500 mM, pH 7.2; (b) HPTS-based assays: 123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 1 mM HPTS, I.S. 150 mM, pH 6.5; (c) Carboxyfluorescein-based assays: 123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM CF, I.S. 150 mM, pH 7.2. The resulting suspension was vortexed and subjected to nine freeze-thaw cycles; subsequently, it was extruded twenty-nine times through a polycarbonate membrane (200 nm) employing a LiposoFast basic extruder (Avestin, Inc.). The resulting unilamellar vesicles were: (a) ISE assays: dialyzed against a NaNO<sub>3</sub> aqueous solution (489 mM NaNO<sub>3</sub>, 5 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 500 mM, pH 7.2) to remove the non-encapsulated chloride; (b) HPTS-based assays: subjected to size-exclusion chromatography, using Sephadex G-25 as the stationary phase and the inner solution without HPTS (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 150 mM, pH 6.5) as the mobile phase, to remove the non-encapsulated HPTS; (c) Carboxyfluorescein-based assays: subjected to size-exclusion chromatography, using Sephadex G-25 as the stationary phase and the inner solution without CF (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 150 mM, pH 7.2) as the mobile phase, to remove the non-encapsulated carboxyfluorescein. Vesicles were collected in a 10-mL volumetric flask, using either the external solution (ISE assays) or the inner one without the probe (HPTS- and carboxyfluorescein-based assays) to bring the suspension to the desired volume.

### **ISE experiments**

Unilamellar vesicles (average diameter: 200 nm) made of POPC and loaded with a sodium chloride aqueous solution (489 mM NaCl, 5 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 500 mM, pH 7.2) were suspended in a sodium nitrate aqueous solution (489 mM NaNO<sub>3</sub>, 5 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 500 mM, pH 7.2), the final POPC concentration in the vial being 0.5 mM and the final volume 5 mL. A certain volume of a solution of the corresponding compound in DMSO was added at  $t = 0$  s, and the chloride released was monitored using a chloride-selective electrode (HACH 9652C). At  $t = 300$  s a surfactant (Triton-X, 10% dispersion in water, 20  $\mu$ L) was added to lyse the vesicles and release all the encapsulated chloride. This value was regarded as 100% release and used to normalize the data.



**Figure S11.** Chloride efflux promoted by the studied compounds (25  $\mu\text{M}$  - 5% mol carrier to lipid concentration) in unilamellar POPC vesicles. Vesicles were loaded with a 489 mM NaCl aqueous solution buffered at pH 7.2 with 5 mM  $\text{NaH}_2\text{PO}_4$  and dispersed in a 489 mM  $\text{NaNO}_3$  aqueous solution buffered at pH 7.2 with 5 mM  $\text{NaH}_2\text{PO}_4$ . Each trace represents the average of at least three different trials performed with three batches of vesicles.

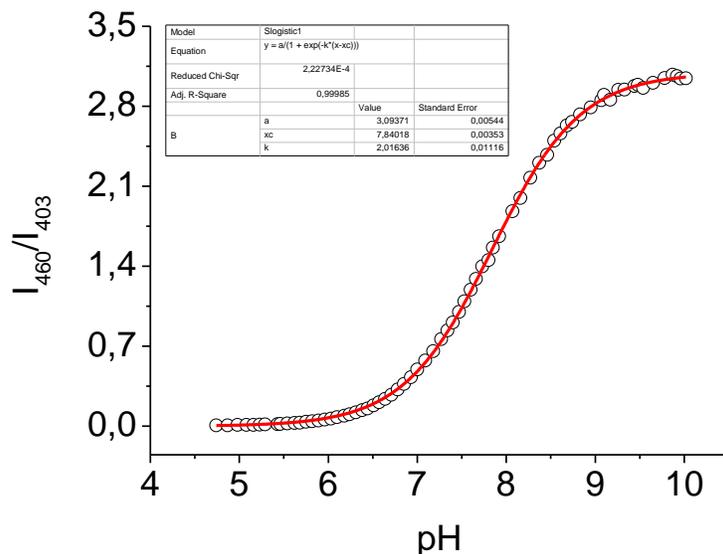


**Figure S12.** Chloride efflux promoted by compound **1c** (25  $\mu\text{M}$  - 5%, black trace; 15  $\mu\text{M}$  - 3%, red trace; 10  $\mu\text{M}$  - 2%, blue trace; 5  $\mu\text{M}$  - 1% mol carrier to lipid concentration, pink trace) in unilamellar POPC vesicles. Vesicles were loaded with a 489 mM NaCl aqueous solution buffered at pH 7.2 with 5 mM  $\text{NaH}_2\text{PO}_4$  and dispersed in a 489 mM  $\text{NaNO}_3$  aqueous solution buffered at pH 7.2 with 5 mM  $\text{NaH}_2\text{PO}_4$ . Each trace represents the average of at least three different trials performed with three batches of vesicles.

## Emission spectroscopy experiments

### HPTS-based assays

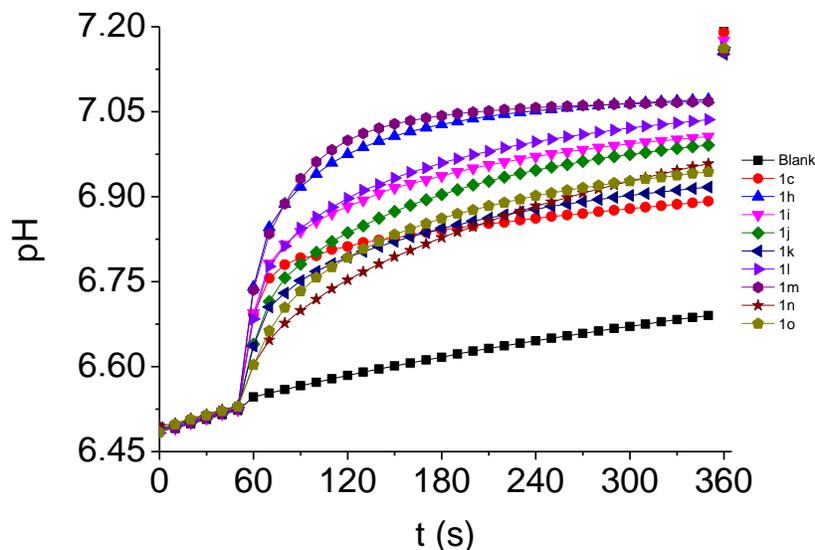
First of all, a calibration curve matching  $I_{460}/I_{403}$ , *i.e.*, the relationship between the emission intensities collected at 510 nm when exciting the sample at 460 nm and 403 nm (the excitation wavelengths of the dye's deprotonated and protonated forms, respectively) of an HPTS aqueous solution (15 nM), prepared with a  $\text{NaNO}_3/\text{NaCl}$  aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM), and the pH was built. In order to do it, aliquots of a sodium hydroxide aqueous solution (0.5 M), prepared with a  $\text{NaNO}_3/\text{NaCl}$  aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM), were successively added to the HPTS solution, and after each addition  $I_{460}/I_{403}$  and the pH value of the solution were recorded. Data were fitted to an S-logistic model, which provided an  $R^2 > 0.999$  (Figure S13).



**Figure S13.** Calibration curve obtained for HPTS (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 15 nM HPTS, I.S. 150 mM). Data were fitted to an S-logistic model (Origin Pro®).

7:3 POPC:cholesterol vesicles were loaded with a  $\text{NaNO}_3/\text{NaCl}$  aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 1 mM HPTS, I.S. 150 mM, pH 6.5) and treated according to the procedure described previously. The assays were conducted in 1-cm disposable cells, the final POPC concentration in the cuvette being 0.5 mM and the final volume 2.5 mL. Just before starting the measurements the required volume of the vesicles stock solution was suspended in the outer solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.5). At  $t = 60$  s an aliquot of a solution of the compound in DMSO (or the blank, DMSO, 12.5  $\mu\text{L}$ ) was added, and the ratio of emission intensities

recorded at 510 nm by excitation of the sample at 460 and 403 nm was recorded for five more minutes. At  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added, to lyse the vesicles and balance the pH.



**Figure S14.** Variation of pH upon addition of the studied compounds (50  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to 7:3 POPC:cholesterol vesicles (0.5 mM POPC). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 1 mM HPTS, I.S. 150 mM, pH 6.5), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.5) just before starting the measurements. At  $t = 60$  s the compound (or the blank, DMSO, 12.5  $\mu\text{L}$ ) was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each trace corresponds to the average of at least three different trials performed with three batches of vesicles. The pH value obtained after mixing the inner and outer solutions in the same proportions used during the assays was 7.23 (measured with a pH-meter).

The value of  $J_0$  (pH initial rate) for each compound was determined. To that end, the section of the curve corresponding to the 60 s passed by since the addition of the compound was selected and fitted to Eq. (1):

$$y = A - B \cdot e^{-k \cdot x} \quad \text{Eq. (1)}$$

where  $y$  represents the pH at a certain time;  $A$  the pH at  $t = 60$  s;  $B$  is  $\Delta\text{pH}$ , *i.e.*, the difference between the pH at  $t = 60$  s and the pH at  $t = 0$  s;  $k$  is the rate constant of the process; and  $x$  is the time. Deriving from Eq. (1) at  $t = 0$  s,  $J_0$  is obtained:

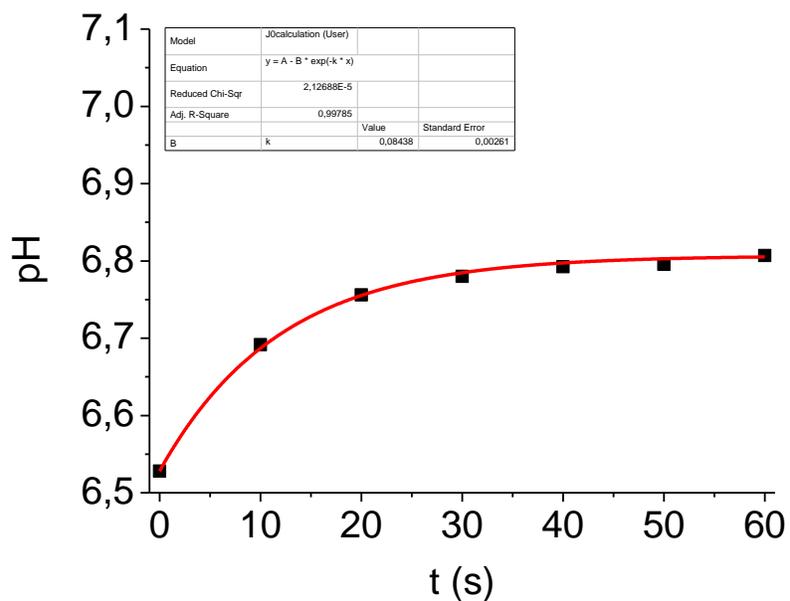
$$y = B \cdot k \quad \text{Eq. (2)}$$

where  $y$  represents  $J_0$ .

Therefore, fitting of the pH vs time data to Eq. (1) provides  $k$ , and with this value it is possible to determine the value of  $J_0$  for each compound (Eq. (2)). These values are presented in **Table S2**.

**Table S2.** Values of  $k$  and  $J_0$  determined for each compound.

Compound	pH at t = 0 s	pH at t = 60 s	$\Delta$ pH	$k$ (s <sup>-1</sup> )	$J_0$ (pH·s <sup>-1</sup> )
<b>1c</b>	6.52799	6.80693	0.27894	0.08438	0.02354
<b>1h</b>	6.52890	6.95975	0.43085	0.06437	0.02773
<b>1i</b>	6.52368	6.87065	0.34697	0.06571	0.02280
<b>1j</b>	6.52756	6.82075	0.29319	0.05111	0.01498
<b>1k</b>	6.52517	6.78145	0.25628	0.05742	0.01472
<b>1l</b>	6.52771	6.88148	0.35377	0.05867	0.02076
<b>1m</b>	6.52595	6.98267	0.45672	0.05719	0.02612
<b>1n</b>	6.53049	6.73723	0.20674	0.04371	0.00904
<b>1o</b>	6.52919	6.77550	0.24631	0.04275	0.01053



**Figure S15.** pH vs time curve and its fitting to Eq. (2) for compound **1c**.

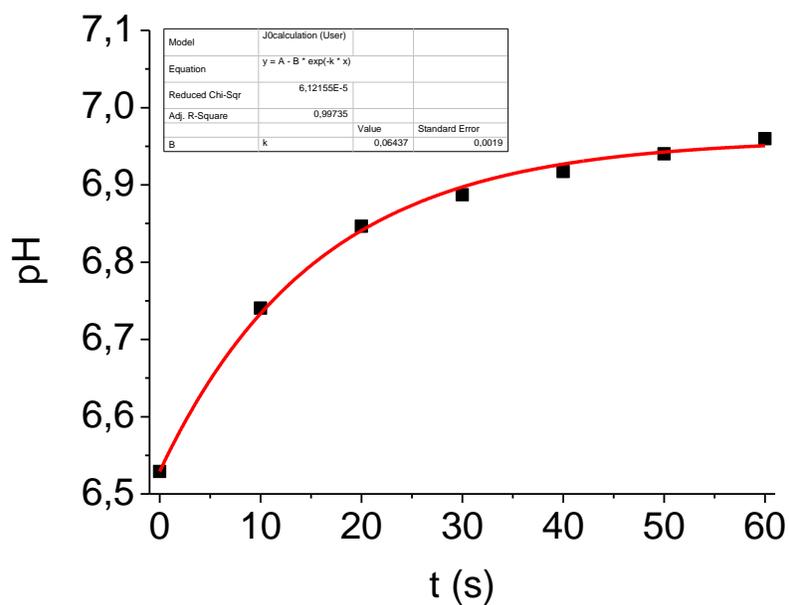


Figure S16. pH vs time curve and its fitting to Eq. (2) for compound **1h**.

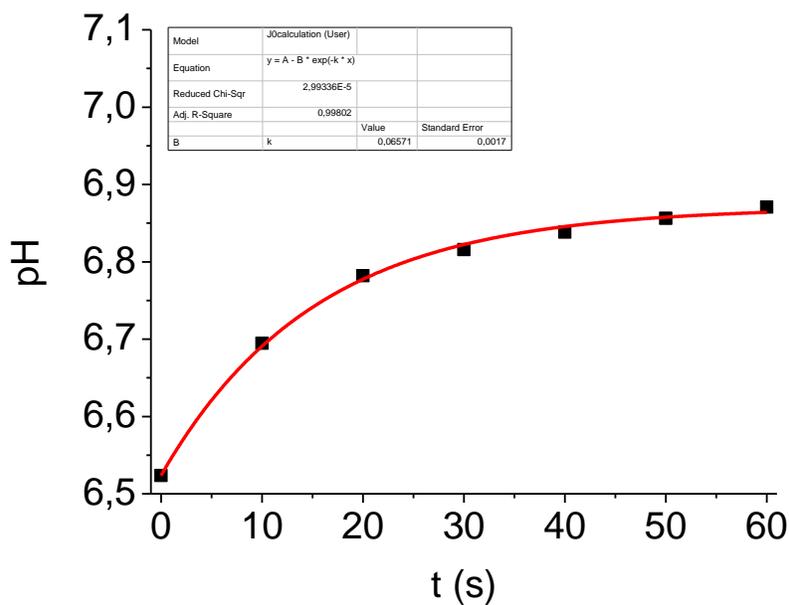


Figure S17. pH vs time curve and its fitting to Eq. (2) for compound **1i**.

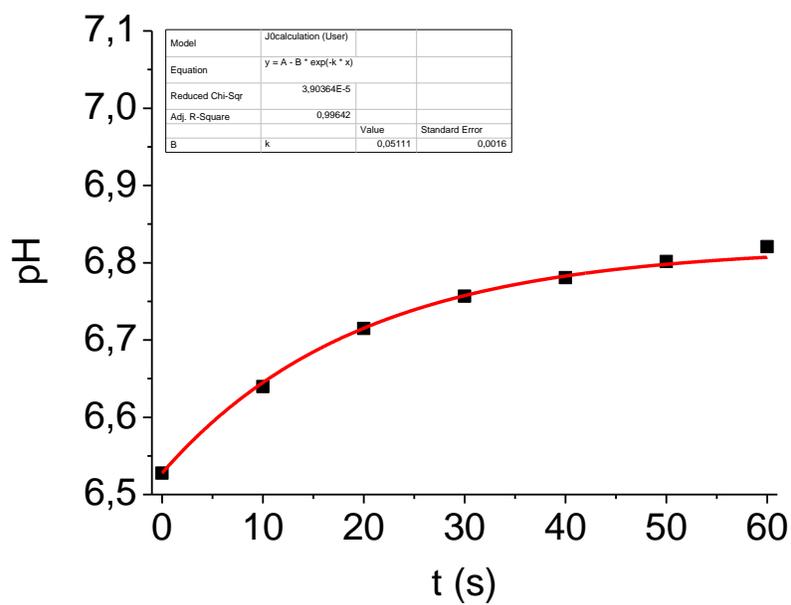


Figure S18. pH vs time curve and its fitting to Eq. (2) for compound **1j**.

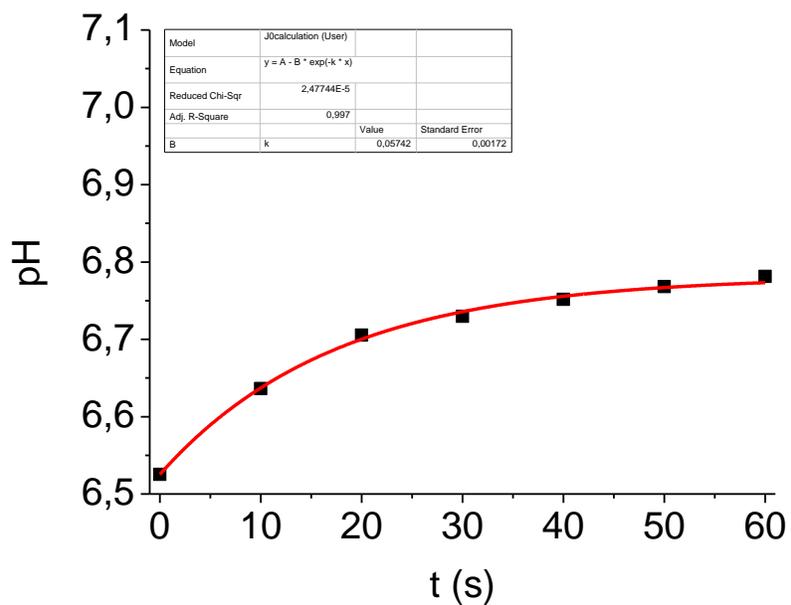


Figure S19. pH vs time curve and its fitting to Eq. (2) for compound **1k**.

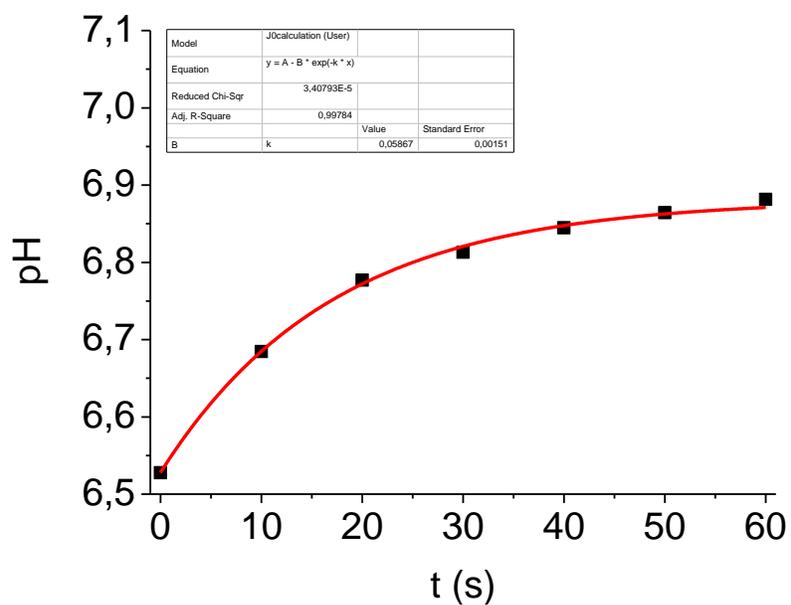


Figure S20. pH vs time curve and its fitting to Eq. (2) for compound **1l**.

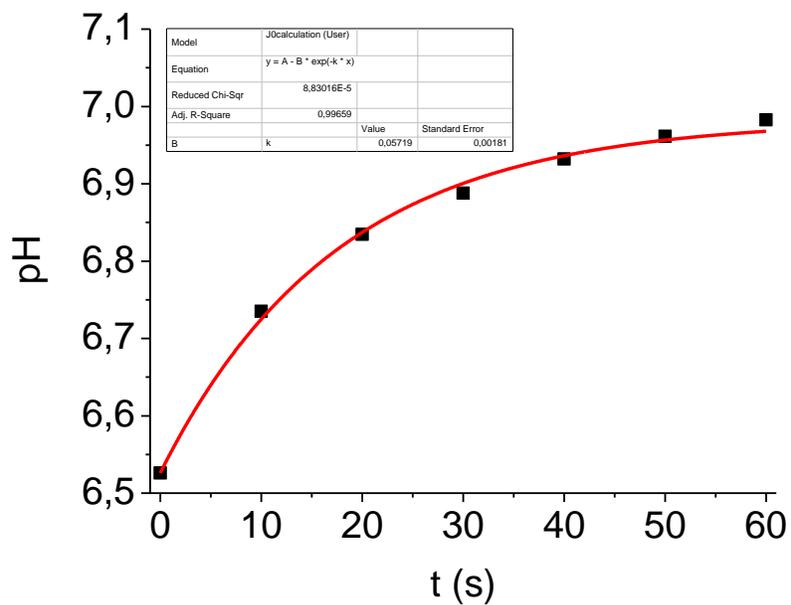


Figure S21. pH vs time curve and its fitting to Eq. (2) for compound **1m**.

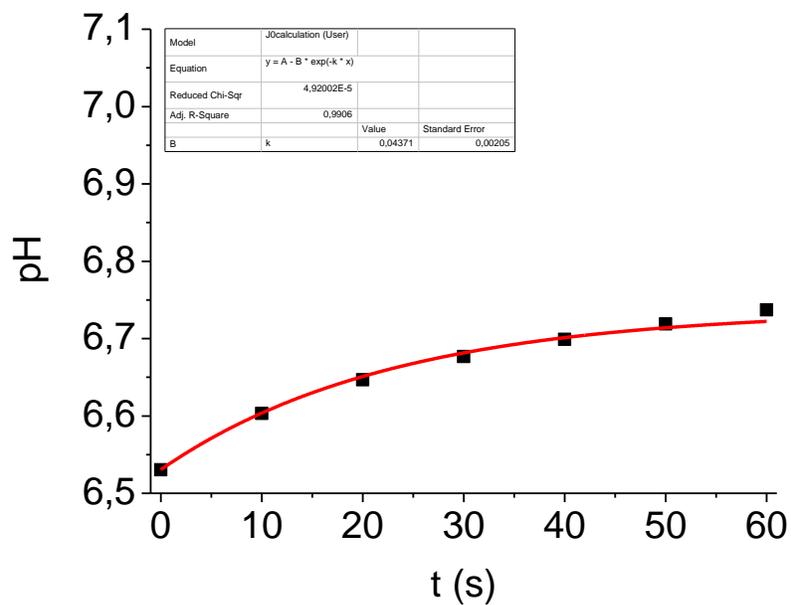


Figure S22. pH vs time curve and its fitting to Eq. (2) for compound **1n**.

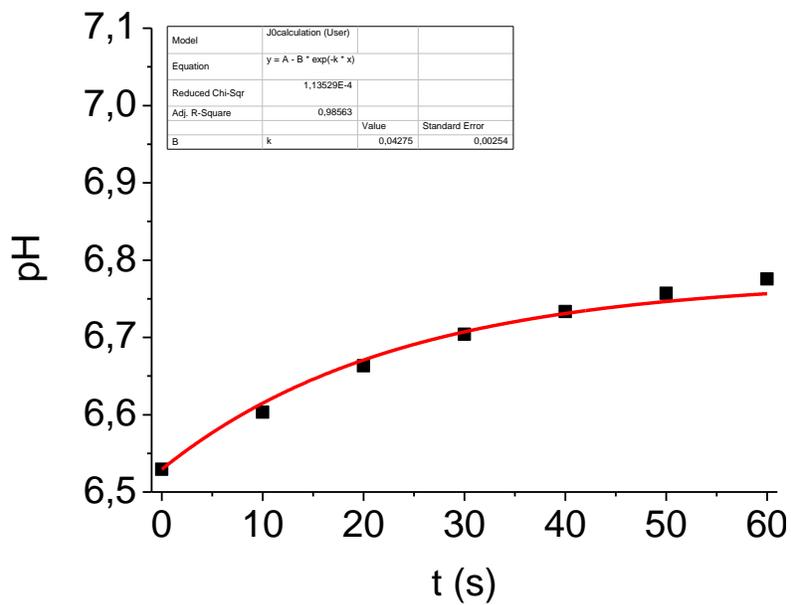
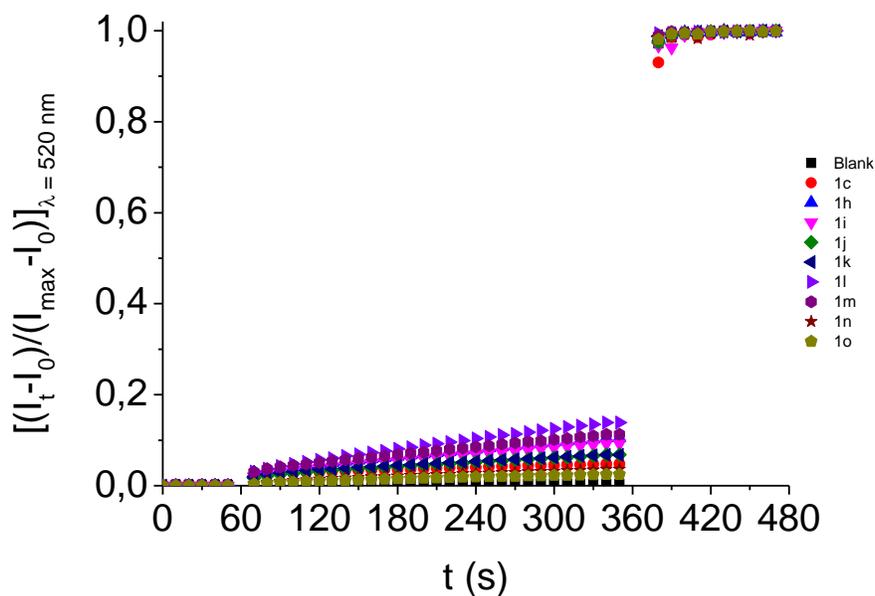


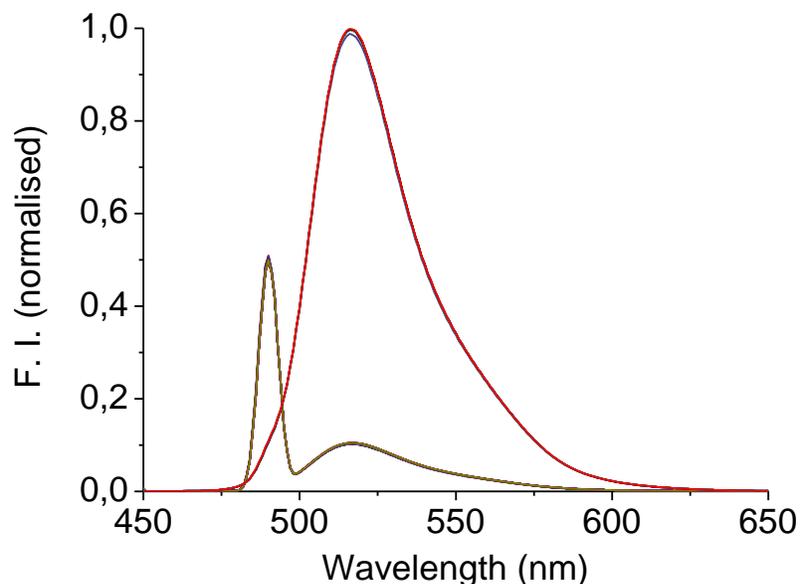
Figure S23. pH vs time curve and its fitting to Eq. (2) for compound **1o**.

### Carboxyfluorescein-based assays

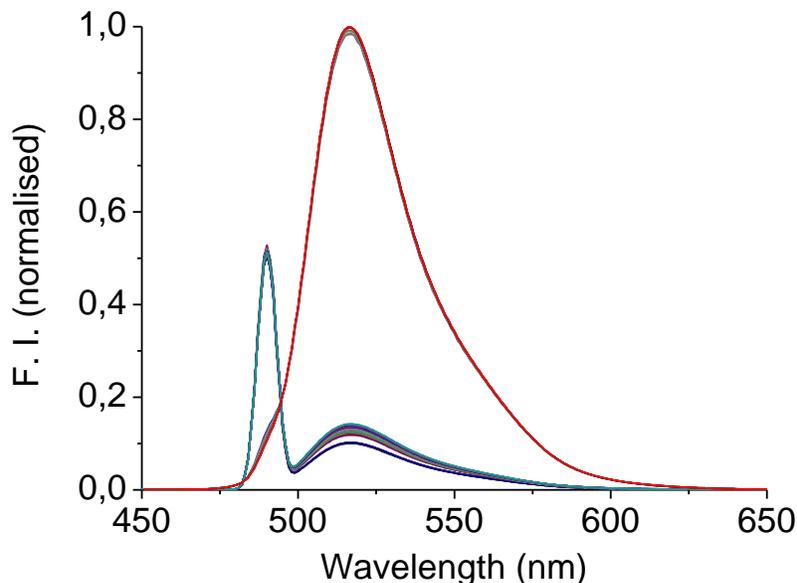
POPC vesicles were loaded with a  $\text{NaNO}_3/\text{NaCl}$  aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2) and treated according to the procedure described previously. The experiments were performed in 1-cm disposable cells, the final POPC concentration in the cuvette being 0.05 mM and the total volume 2.5 mL. Just before starting the measurements, the required volume of the vesicles stock solution was suspended in the outer solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added and the emission changes were recorded during five minutes. At  $t = 360$  s a pulse of a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added to lyse the vesicles and release all the entrapped carboxyfluorescein. This value was regarded as 100% release and used to normalize the data.



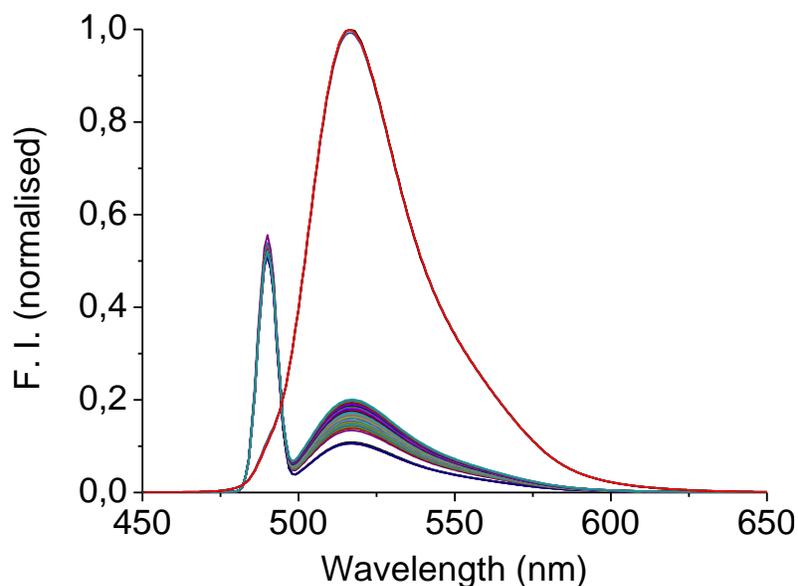
**Figure S24.** Carboxyfluorescein leakage observed upon addition of the studied compounds (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) or the blank (DMSO) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound (or the blank, DMSO, 1.25  $\mu\text{L}$ ) was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each trace corresponds to the average of six different trials performed with three batches of vesicles.



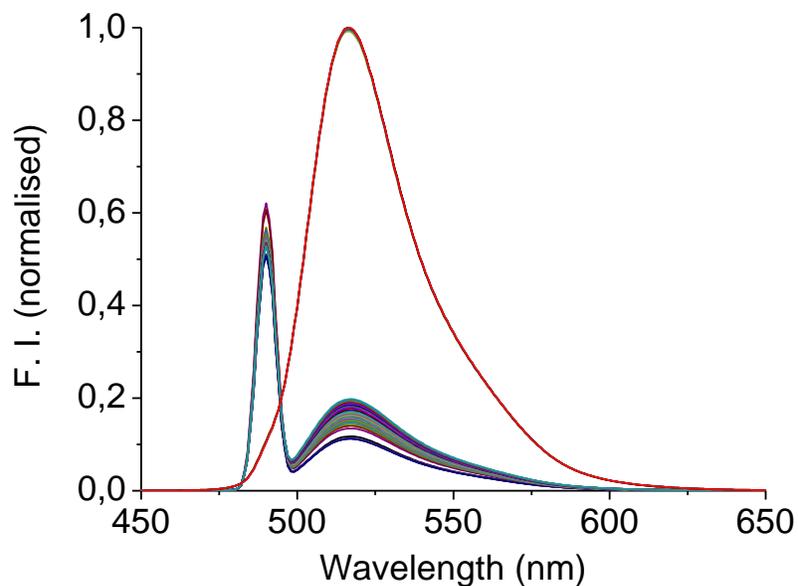
**Figure S25.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of **DMSO**, the blank, to POPC vesicles (0.05 mM). Vesicles, loaded with a NaNO<sub>3</sub> and NaCl buffered aqueous solution (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a NaNO<sub>3</sub> and NaCl buffered aqueous solution (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 150 mM, pH 7.2). At t = 60 s the blank (DMSO, 1.25 μL) was added, and at t = 360 s a detergent (Triton-X, 10% dispersion in water, 20 μL) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



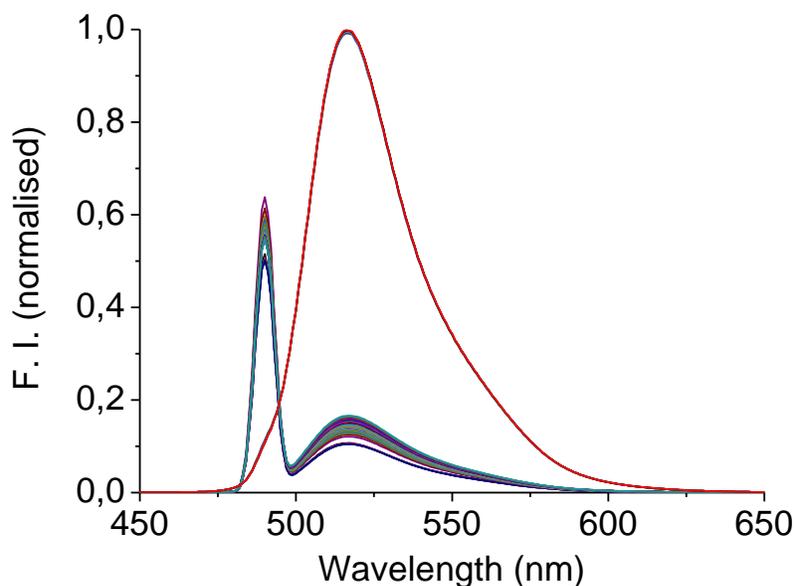
**Figure S26.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1c** (5 μM - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a NaNO<sub>3</sub> and NaCl buffered aqueous solution (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a NaNO<sub>3</sub> and NaCl buffered aqueous solution (123.9 mM NaNO<sub>3</sub>, 10 mM NaCl, 10 mM NaH<sub>2</sub>PO<sub>4</sub>, I.S. 150 mM, pH 7.2). At t = 60 s the compound was added, and at t = 360 s a detergent (Triton-X, 10% dispersion in water, 20 μL) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



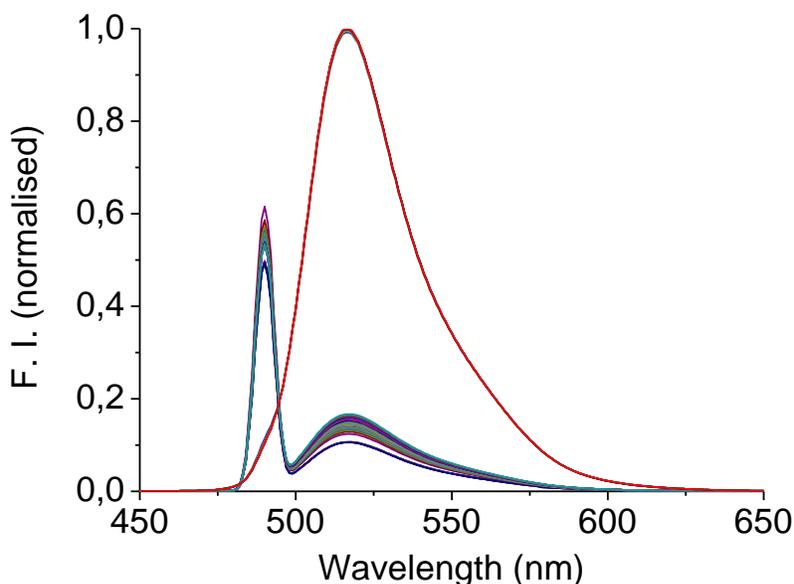
**Figure S27.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1h** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



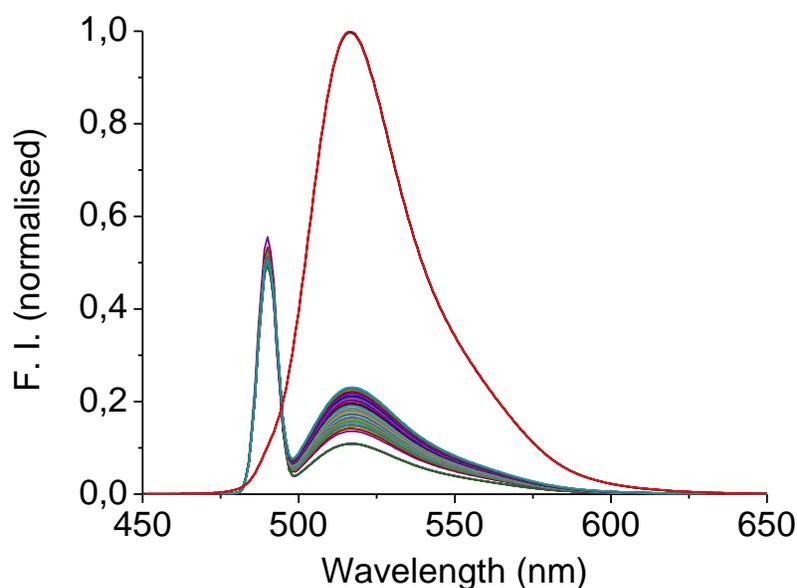
**Figure S28.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1i** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



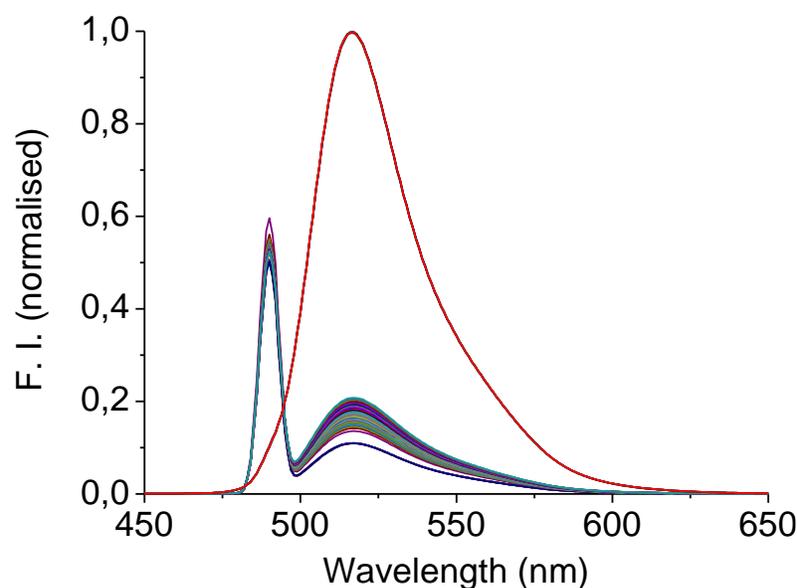
**Figure S29.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1j** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



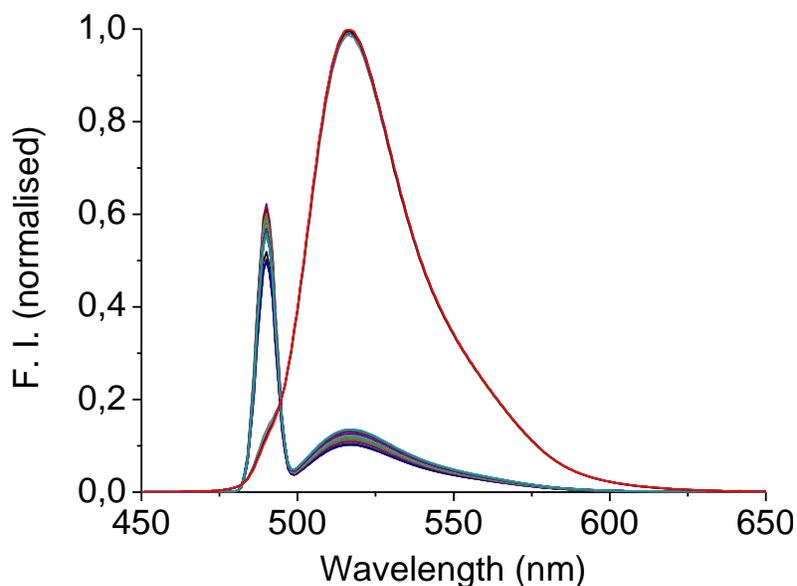
**Figure S30.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1k** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



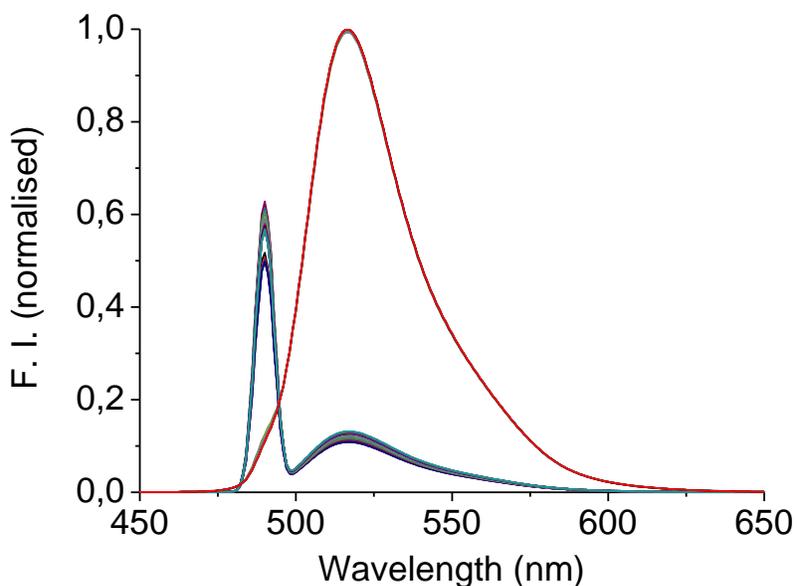
**Figure S31.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1l** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and NaCl buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM NaCl, 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and NaCl buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM NaCl, 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



**Figure S32.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1m** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and NaCl buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM NaCl, 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and NaCl buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM NaCl, 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



**Figure S33.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1n** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.



**Figure S34.** Carboxyfluorescein normalised fluorescence intensity recorded upon addition of compound **1o** (5  $\mu\text{M}$  - 10% mol carrier to lipid concentration) to POPC vesicles (0.05 mM). Vesicles, loaded with a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , 50 mM CF, I.S. 150 mM, pH 7.2), were suspended in a  $\text{NaNO}_3$  and  $\text{NaCl}$  buffered aqueous solution (123.9 mM  $\text{NaNO}_3$ , 10 mM  $\text{NaCl}$ , 10 mM  $\text{NaH}_2\text{PO}_4$ , I.S. 150 mM, pH 7.2). At  $t = 60$  s the compound was added, and at  $t = 360$  s a detergent (Triton-X, 10% dispersion in water, 20  $\mu\text{L}$ ) was added. Each spectrum corresponds to the average of six different trials performed with three batches of vesicles.

## **NMR measurements**

The NMR spectra of the cages in micelles were acquired using a Bruker Avance-III 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) TCI cryoprobe. We used 25 mM (monomer concentration) of DPC-d38 (dodecylphosphocholine) micelles, well above cmc (1.36 mM in PBS). For the incorporation of pseudopeptidic cages into DPC micelles, a homogeneous solution of detergent in  $\text{D}_2\text{O}$  was prepared and the cage was added from a weighted stock solution prepared in  $\text{DMSO-d}_6$ . After vigorous vortexing to ensure the incorporation of the cage to the micelles, the pH was adjusted. Due to the limited solubility of pseudopeptidic host in aqueous solution, the samples in DPC/ $\text{D}_2\text{O}$  solution have a final concentration of 0.4 mM cage with a  $\sim 1:1$  cage:DPC micelles ratio (using known aggregation number 55-60). The samples at these concentrations were stable and optically clear (without observable precipitation or phase separation) over all the NMR experiments. Regarding the pH adjustment, preliminary studies showed that the use of a buffer (in  $\text{D}_2\text{O}$ ) could interfere with anion recognition. Thus, all the cage/DPC samples were prepared in  $\text{D}_2\text{O}$  without any buffer and the pH (acidic or neutral) was adjusted using the minimum amount of  $\text{DCl}/\text{NaOD}$  (from a 0.2 M stock in  $\text{D}_2\text{O}$ ), minimizing the amount of chloride added at the end of the pH adjustment.

Two-dimensional (2D) spectra (NOESY, ROESY, COSY and HSQC) were acquired at 303 K. The water resonance was suppressed by “3-9-19” pulse sequence with gradients using flip-back pulse in NOESY experiments (pulse sequence `noesyfpgpphgw`), 3-9-19 suppression sequence in ROESY experiments (pulse sequence `roesygpph19`) and presaturation for COSY experiments (pulse sequence `cosygpprqf`).  $^1\text{H}$ - $^1\text{H}$  NOESY spectra were acquired with a mixing time of 150 ms (in DPC) and  $^1\text{H}$ - $^1\text{H}$  ROESY spectra with a mixing time of 75 ms. 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC experiments were acquired with Bruker pulse sequence `hsqc dietg p sisp`. 2D experiments were processed with the standard TOPSPIN program (Bruker Biospin, Karlsruhe, Germany). For PFG-NMR diffusion experiments, the field gradient strength (Gz) was calibrated by measuring the self-diffusion coefficient of residual  $\text{H}_2\text{O}$  in a 100 %  $\text{D}_2\text{O}$  sample at 298.0 K. A diffusion coefficient of  $1.91 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$  for the residual  $\text{H}_2\text{O}$  signal was then used for the back calculation of the field gradient strength. For  $\text{D}_2\text{O}$  samples, diffusion NMR experiments were acquired at 303 K using Shigemi tubes to minimize convection effects. The `stebppg1s19` pulse sequence with WATERGATE 3-91-9 for water suppression and one spoil gradient was used. The diffusion coefficient,  $D$ , was determined by fitting diffusion weighted intensities of selected peaks or integrals over a chosen range to the following equation:

$$I = I_0 \exp\{-\gamma^2 g^2 \delta^2 (\Delta - \delta/3) D\} \quad (\text{Eq. 1})$$

Where  $\gamma$  is the gyromagnetic ratio of proton and  $g$ ,  $\delta$  and  $\Delta$  are the amplitude, duration and separation of the single pair of gradient pulses, respectively. Due to the use of bipolar gradient elements,  $\Delta$  is replaced by  $(\Delta - \tau_1/2)$  in Eq. (1), with  $\tau_1$  being the time interval between the bipolar gradient pulses within the bipolar gradient encoding or decoding segment. Fittings to Eq. (1) were performed using Bruker Dynamics Center.

The pulsed-field-gradient signal attenuation was monoexponential for all measured resonances. In the case of DPC, this indicated that the exchange of lipid molecules between micelles and free solution is fast in the NMR time scale. Under these conditions, the observed diffusion coefficient is the weighted average of the free ( $D_{free}$ ) and micellar ( $D_{mic}$ ) values:

$$D_{obs} = f_{free}D_{free} + f_{mic}D_{mic} \text{ (Eq. 2)}$$

In which  $f_{free}$  and  $f_{mic}$  are the mole fraction of free and micellar DPC. The self-coefficient diffusion for DPC protons ( $(CH_2)_n$  and  $CH_3$ ) are in the range  $10\text{-}12 \times 10^{-11} \text{ m}^2/\text{s}$  (at 303K in  $D_2O$ ), similar to known values from DPC micelles in literature, a confirmation that in our experimental conditions we have micellar aggregates. The reason for DPC self-diffusion coefficients being higher than the values from absorbed cage is the presence of free monomeric DPC molecules (with concentration smaller than cmc) in fast exchange with micelles in solution. However, the observed cage resonances and their self-diffusion coefficients, smaller than DPC diffusion coefficients, mainly correspond to DPC-bound molecules and act as a probe for the cage incorporation to the micelles, so  $D_{mic} \text{ (DPC)} \sim D_{cage} \text{ (micelle-bound cages)}$ .

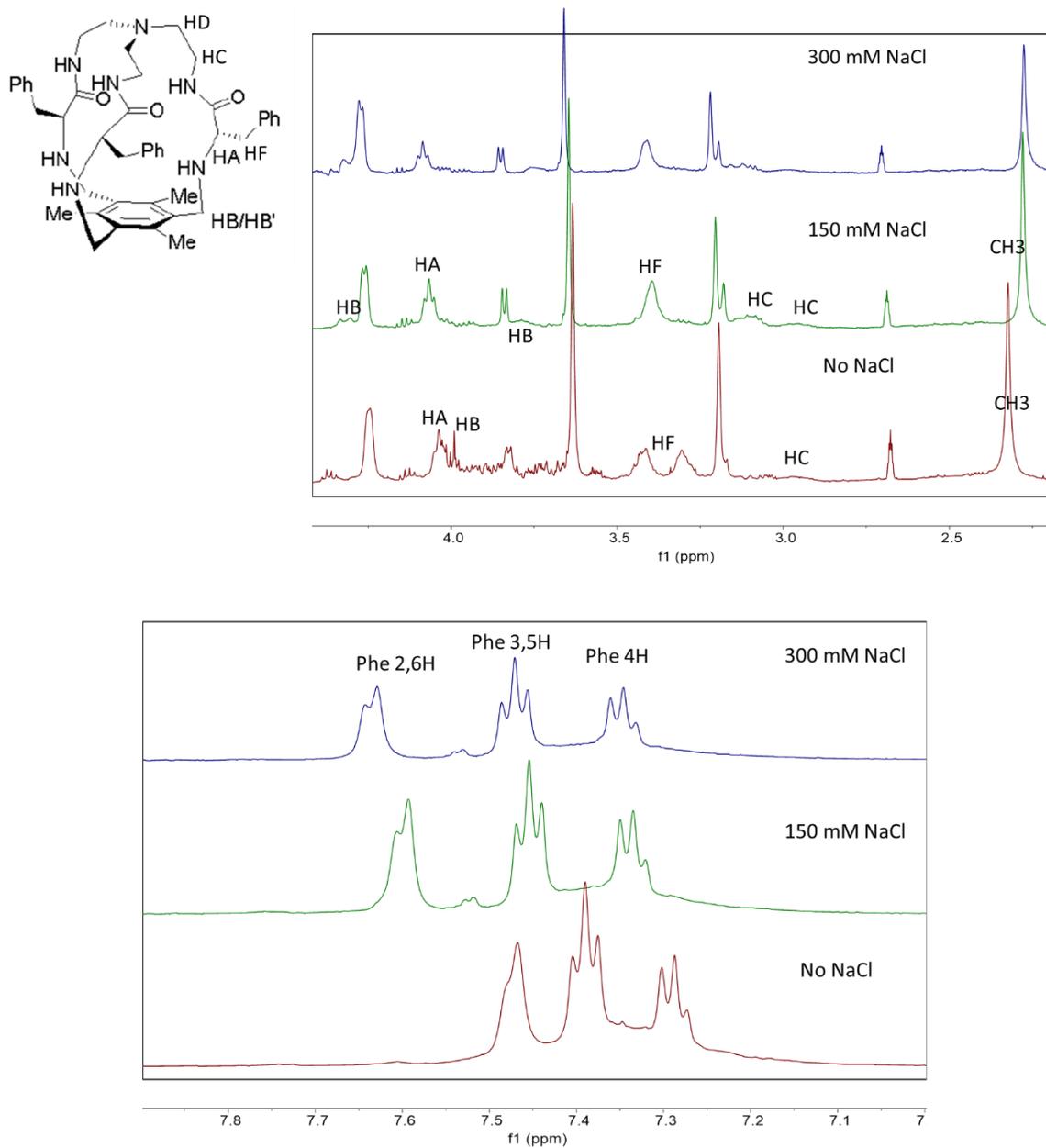
**Table S3.** Translational diffusion coefficient measured by NMR spectroscopy for cages **1a**, **1c** and **1h** in DPC-d38 micelles at neutral pH and without and with 150 mM NaCl. SD, standard deviation of the fitting to Eq. 1.

	<b>D (10<sup>-11</sup> m<sup>2</sup>s<sup>-1</sup>)</b>	<b>SD</b>	<b>D cage/ D CH<sub>2</sub> DPC</b>
<b>0.4 mM 1a cage:25 mM DPC (D<sub>2</sub>O, pH 7.0)</b>			
5e Aromatic region	9.085	1.013e-3	0.79
CH <sub>3</sub>	8.857	3.113e-4	0.77
DPC CH <sub>3</sub>	11.570	5.155e-3	
DPC CH <sub>2</sub>	11.440	4.465e-3	
<b>0.4 mM 1c cage:25 mM DPC (D<sub>2</sub>O, pH 7.1)</b>			
2,6H Aromatic region (free)	8.844	4.323e-4	0.70
3,5H Aromatic region (free)	8.914	3.134e-4	0.71
CH <sub>3</sub>	8.917	2.939e-4	0.71
DPC CH <sub>3</sub>	10.903	1.280e-3	
DPC CH <sub>2</sub>	12.590	1.531e-3	
<b>0.4 mM 1h cage:25 mM DPC (D<sub>2</sub>O, pH 7.1)</b>			
6H Aromatic region	9.060	1.017e-3	0.80
4H Aromatic region	9.010	5.860e-4	0.80
5H Aromatic region	9.070	6.180e-4	0.80
3H Aromatic region	8.990	7.882e-4	0.80
CH <sub>3</sub>	9.020	3.079e-4	0.80
DPC CH <sub>3</sub>	11.200	1.670e-4	
DPC CH <sub>2</sub>	11.300	6.803e-4	
<b>0.4 mM 1a cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 7.0)</b>			
5e Aromatic region	8.810	2.888e-2	0.82
CH <sub>3</sub>	8.640	3.285e-2	0.81
DPC CH <sub>3</sub>	11.360	9.161e-2	
DPC CH <sub>2</sub>	10.710	1.272e-2	
<b>0.4 mM 1c cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 7.1)</b>			
2,6H Aromatic region (free)	8.751	9.715e-3	0.76
3,5H Aromatic region (free)	8.808	7.750e-3	0.77
CH <sub>3</sub>	8.743	6.564e-3	0.76
DPC CH <sub>3</sub>	10.680	3.462e-3	
DPC CH <sub>2</sub>	11.500	3.900e-3	
<b>0.4 mM 1h cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 7.1)</b>			
6H Aromatic region	8.970	1.896e-2	0.81
4H Aromatic region	8.950	1.237e-2	0.81
5H Aromatic region	8.970	1.387e-2	0.81
3H Aromatic region	8.950	1.448e-2	0.81
CH <sub>3</sub>	8.928	5.661e-3	0.81
DPC CH <sub>3</sub>	11.200	3.306e-3	
DPC CH <sub>2</sub>	11.090	1.769e-2	

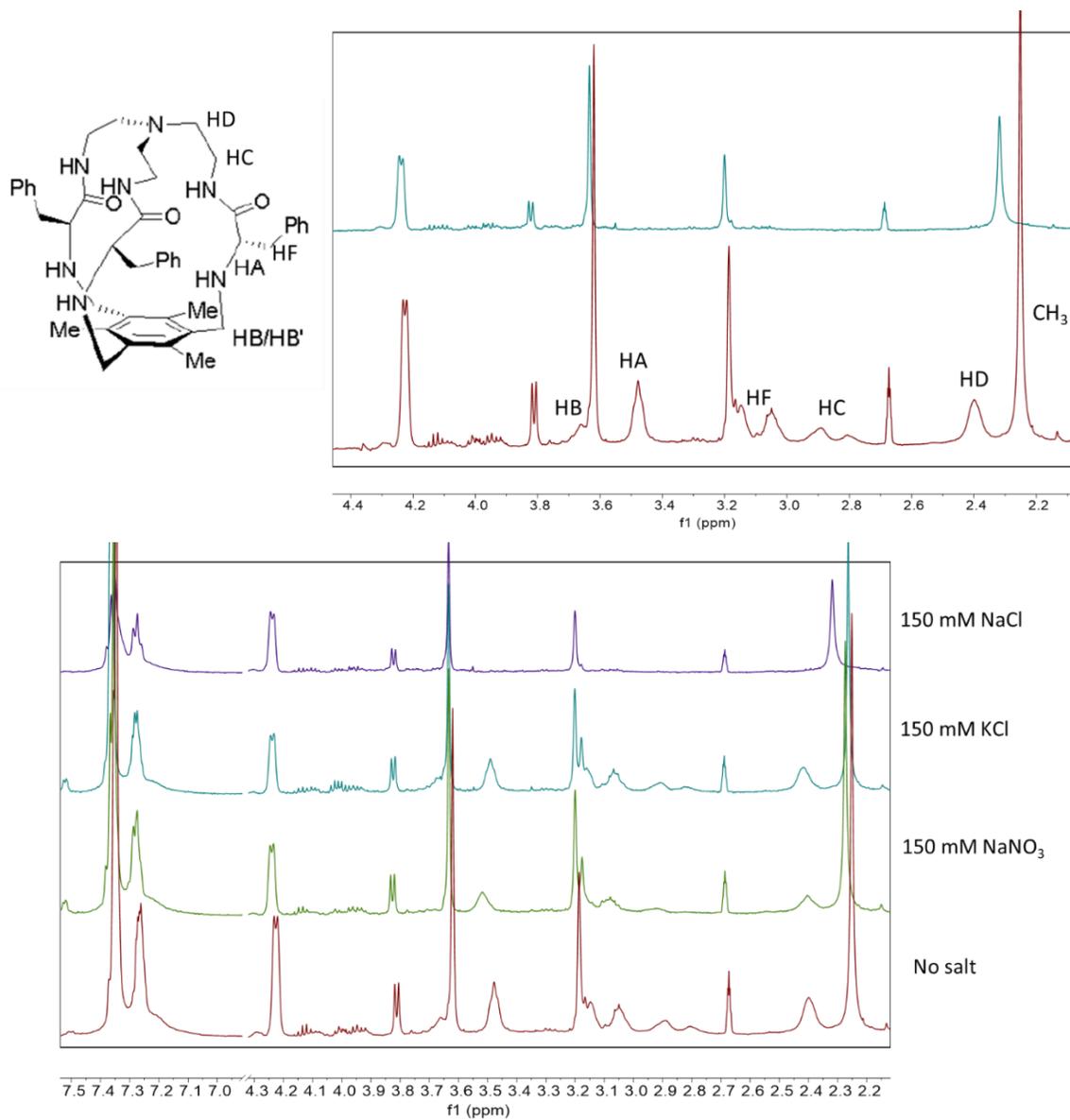
**Table S4.** Translational diffusion coefficient measured by NMR spectroscopy for cages **1a**, **1c** and **1h** samples in DPC-d38 micelles at acid pH and without and with 150 mM NaCl. SD, standard deviation of the fitting to Eq. 1.

	<b>D (<math>10^{-11}</math> <math>m^2s^{-1}</math>)</b>	<b>SD</b>	<b>D cage/ D CH<sub>2</sub> DPC</b>
<b>0.4 mM 1a cage:25 mM DPC (D<sub>2</sub>O, pH 2.3)</b>			
5e Aromatic region	10.300	2.044e-3	0.94
CH <sub>3</sub>	10.300	2.140e-3	0.94
DPC CH <sub>3</sub>	11.530	3.240e-3	
DPC CH <sub>2</sub>	10.990	4.722e-3	
<b>0.4 mM 1c cage:25 mM DPC (D<sub>2</sub>O, pH 2.6)</b>			
2,6H Aromatic region (free)	8.824	5.254e-4	0.79
3,5H Aromatic region (free)	8.856	4.820e-4	0.80
CH <sub>3</sub>	8.936	4.575e-4	0.80
DPC CH <sub>3</sub>	n.m.	n.m.	
	(broad)	(broad)	
DPC CH <sub>2</sub>	11.113	1.393e-2	
<b>0.4 mM 1h cage:25 mM DPC (D<sub>2</sub>O, pH 2.6)</b>			
6H Aromatic region	9.590	7.017e-3	0.86
4H Aromatic region	9.563	6.844e-3	0.85
5H Aromatic region	9.476	6.333e-3	0.85
3H Aromatic region	9.542	7.749e-3	0.85
CH <sub>3</sub>	9.624	3.240e-3	0.86
DPC CH <sub>3</sub>	10.070	8,319e-3	
DPC CH <sub>2</sub>	11.210	1.535e-2	
<b>0.4 mM 1a cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 2.3)</b>			
5e Aromatic region	11.240	1.572e-3	1.06
5e CH <sub>3</sub>	11.770	3.240e-3	1.11
DPC CH <sub>3</sub>	11.120	1.264e-3	
DPC CH <sub>2</sub>	10.600	3.235e-3	
<b>0.4 mM 1c cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 2.6)</b>			
2,6H Aromatic region (bound)	8.810	1.428e-3	0.79
3,5H Aromatic region (bound)	8.870	1.054e-3	0.79
CH <sub>3</sub>	8.867	1.075e-3	0.79
DPC CH <sub>3</sub>	n.m.	n.m.	
	(broad)	(broad)	
<b>DPC CH<sub>2</sub></b>	11.160	3.326e-4	
<b>0.4 mM 1h cage:25 mM DPC (D<sub>2</sub>O, 150 mM NaCl, pH 2.6)</b>			
6H Aromatic region	13.37	1.830e-2	1.25
4H Aromatic region	13.33	1.874e-2	1.25
5H Aromatic region	13.31	1.850e-2	1.25
3H Aromatic region	13.36	2.164e-2	1.25
CH <sub>3</sub>	13.74	1.016e-2	1.28
DPC CH <sub>3</sub>	10.41	1.519e-2	
DPC CH <sub>2</sub>	10.72	3.177e-3	

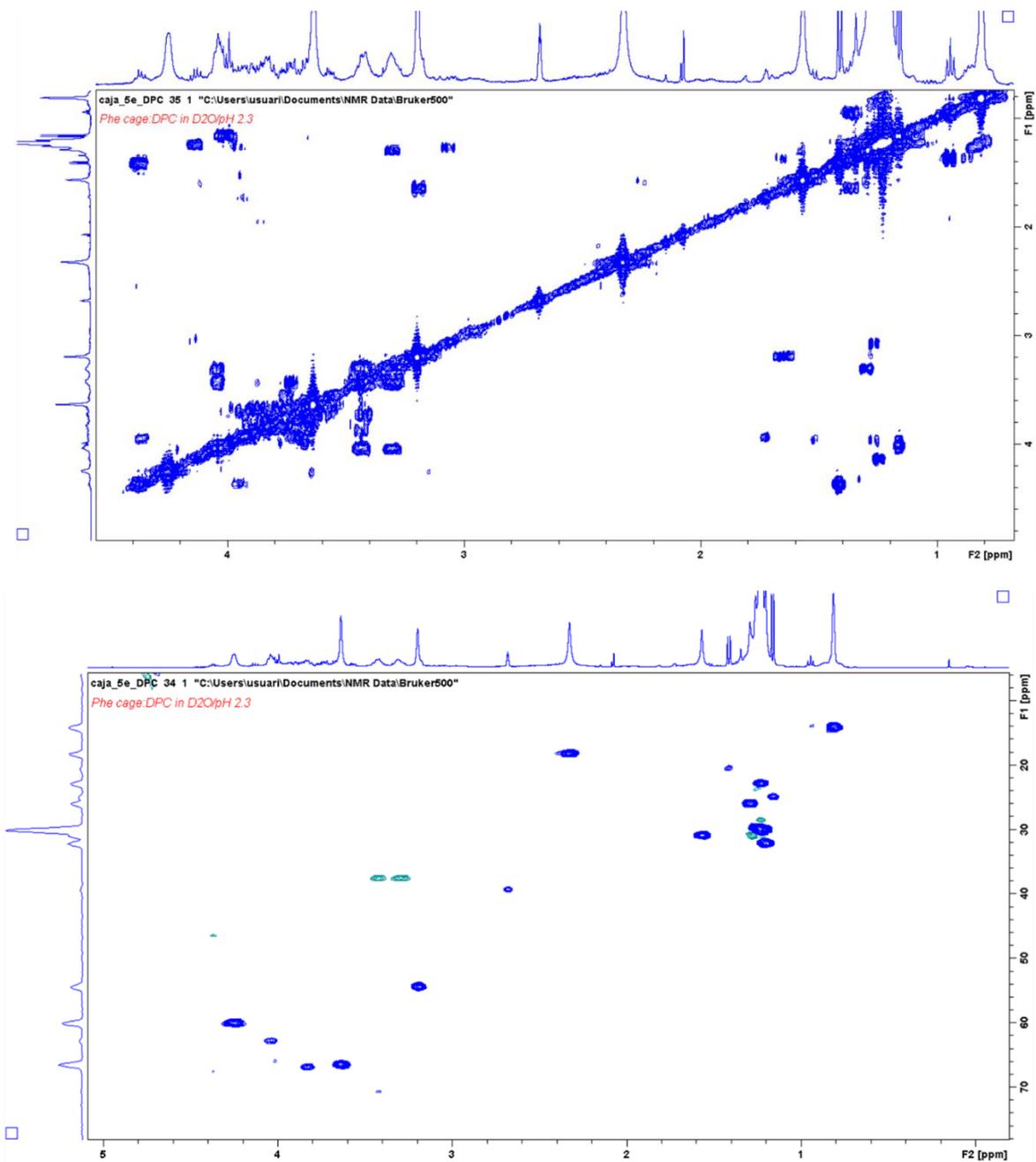
### NMR Spectra for 1a and 1h in DPC micelles



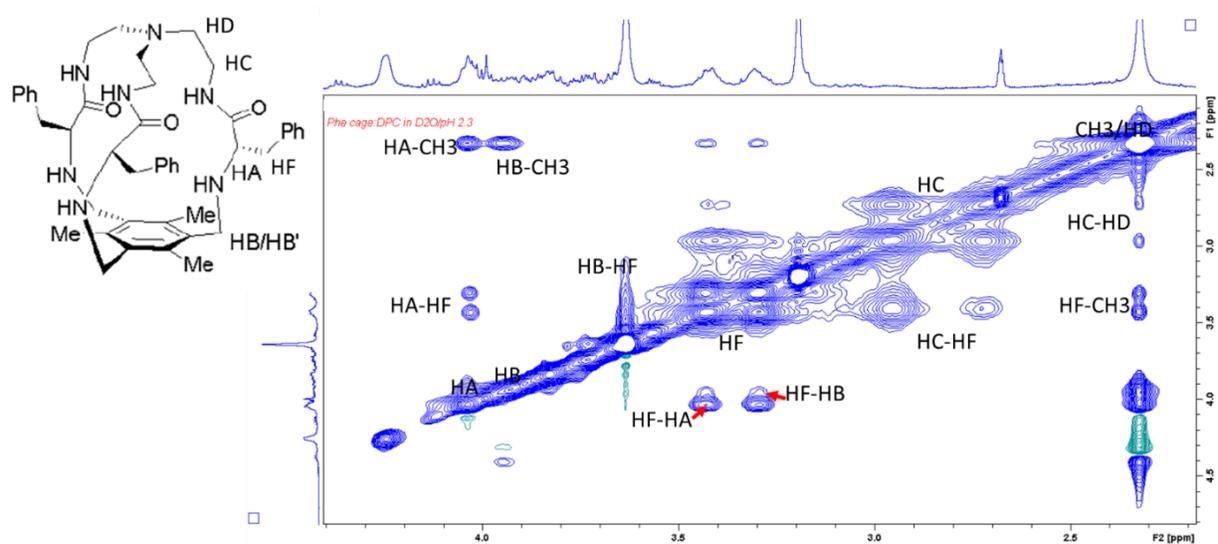
**Figure S35.** 1D <sup>1</sup>H NMR spectra of **1a** (Phe cage) in DPC-d38 micelles (D<sub>2</sub>O, pH 2.3) with 0, 150 and 300 mM NaCl. Peak assignments (top spectrum) were confirmed by 2D NMR experiments.



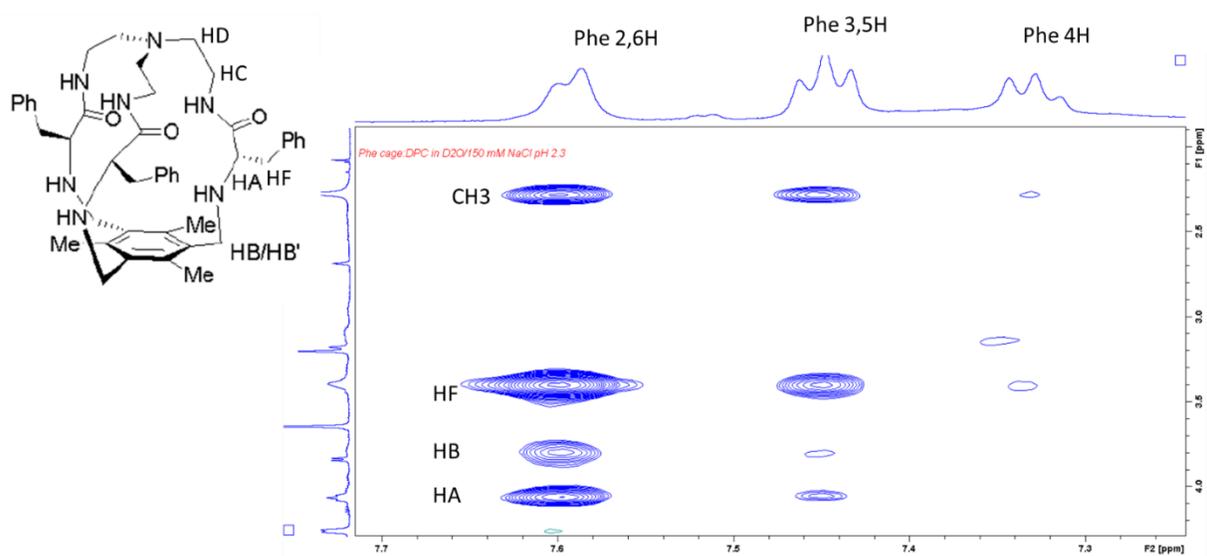
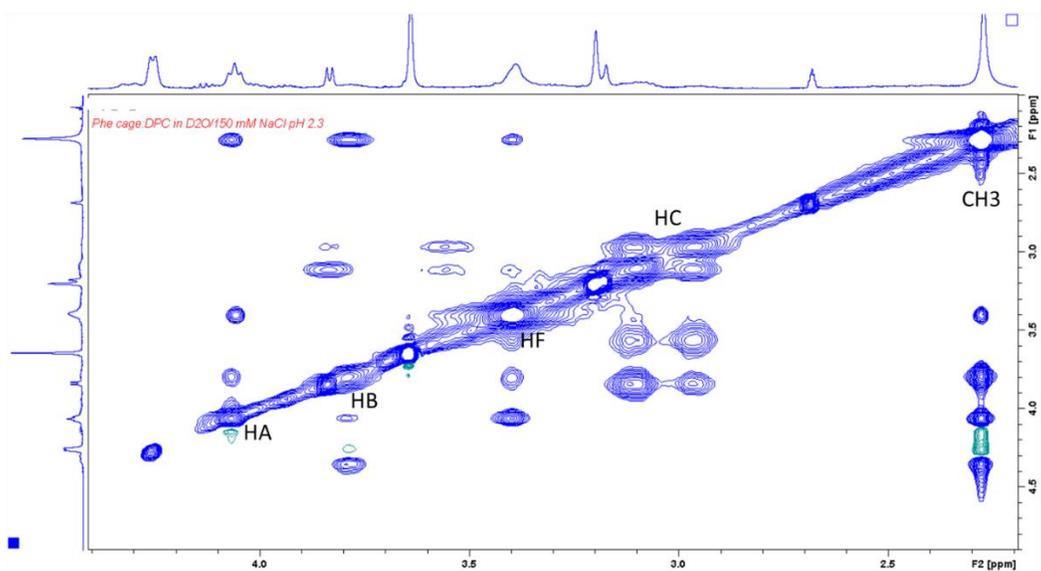
**Figure S36.** 1D  $^1\text{H}$  NMR spectra of **1a** (Phe cage) in DPC-d38 micelles (D<sub>2</sub>O, pH 7.0) with 0 and 150 mM NaNO<sub>3</sub>, KCl or NaCl. Peak assignments (top spectrum) were confirmed by 2D experiments.



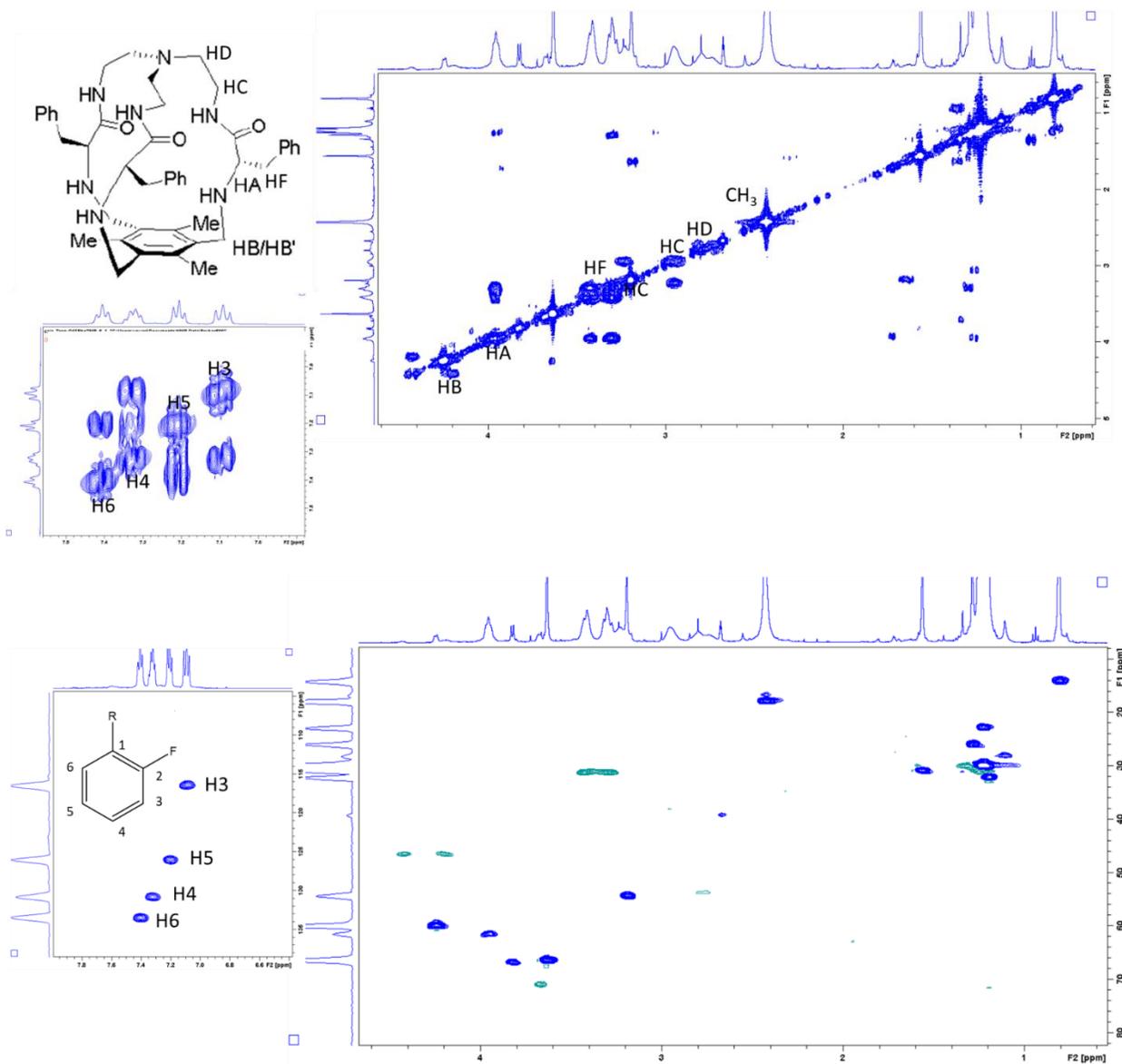
**Figure S37.** 2D <sup>1</sup>H-<sup>1</sup>H COSY and <sup>1</sup>H-<sup>13</sup>C HSQC NMR spectra of **1a** (Phe cage) in DPC-d38 micelles (D<sub>2</sub>O, pH 2.3).



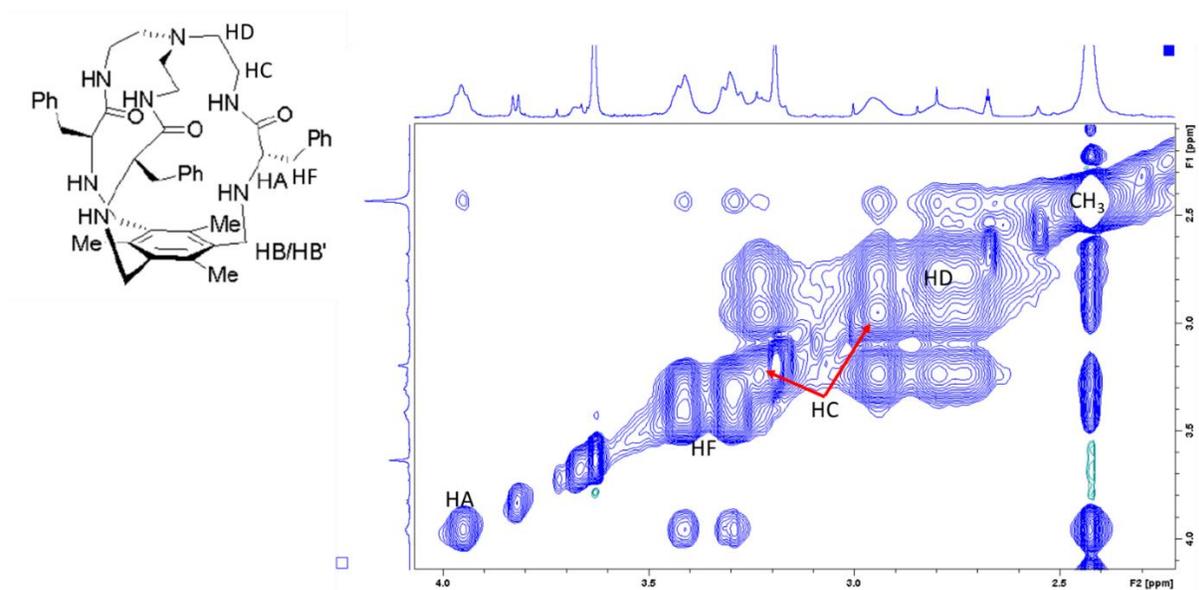
**Figure S38.** 2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectra of **1a** (Phe cage) in DPC-d38 micelles ( $\text{D}_2\text{O}$ , pH 2.3).



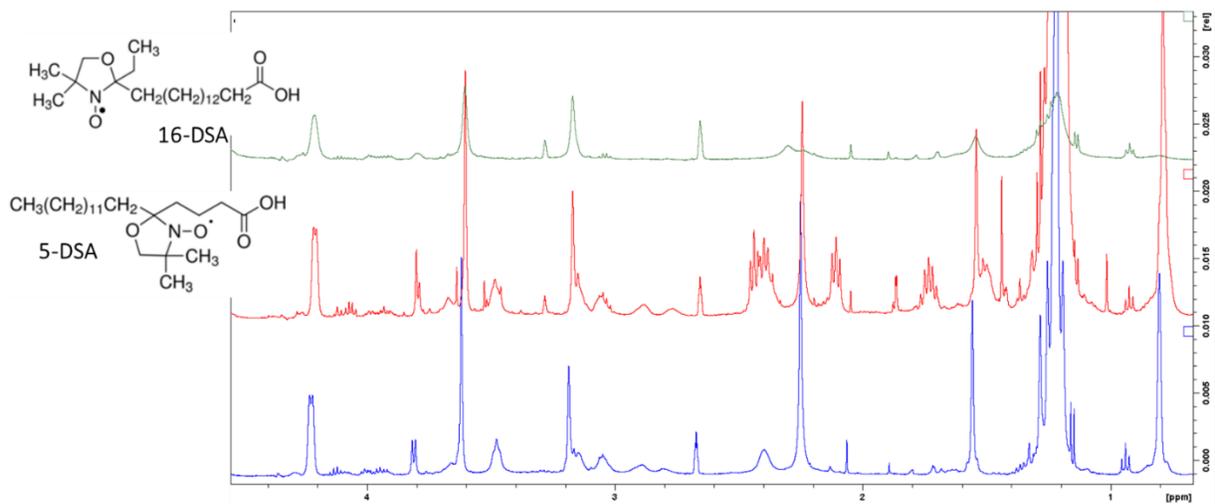
**Figure S39.** 2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectra of **1a** (Phe cage) in DPC-d38 micelles ( $\text{D}_2\text{O}$ , pH 2.3) with 150 mM NaCl.



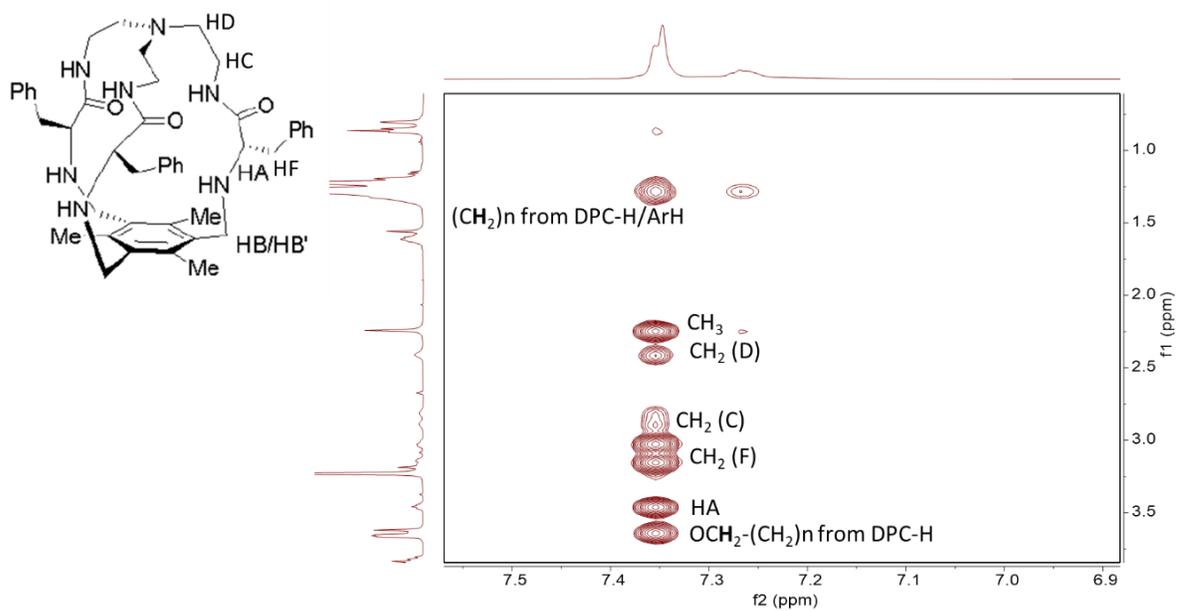
**Figure S40.** 2D  $^1\text{H}$ - $^1\text{H}$  COSY (up) and  $^1\text{H}$ - $^{13}\text{C}$  HSQC (down) NMR spectra of **1h** (2F-Phe cage) in DPC-d38 micelles (D<sub>2</sub>O, pH 2.3).



**Figure S41.** 2D  $^1\text{H}$ - $^1\text{H}$  NOESY NMR spectra of **1h** (2F-Phe cage) in DPC-d38 micelles ( $\text{D}_2\text{O}$ , pH 2.6).



**Figure S42.** 1D  $^1\text{H}$  NMR spectra of **1a** (Phe cage) in DPC-d38 micelles ( $\text{D}_2\text{O}$ , pH 7.0) doped with 5-DSA and 16-DSA doxyl stearic acid radicals. The resonances of **1a** are clearly broadened in presence of the 16-DSA radical, showing that the cage is mainly positioned within the core of the micelles.



**Figure S43.** 2D <sup>1</sup>H-<sup>1</sup>H NOESY NMR spectra of **1a** (Phe cage) in DPC-d38 micelles (D<sub>2</sub>O, pH 7.0) doped with 0.8 mM protonated DPC. The NOE cross-peaks confirm the positioning of the cages close to the aliphatic region of DPC.

## **Cytotoxicity studies**

### **Cell cultures**

Human lung adenocarcinoma cells, A549 were maintained in Dulbecco's modified eagle's medium (DMEM; Gibco Thermo Fisher Scientific, USA) with carbonate buffer. Cells were kept in the logarithmic growth phase by routine passage every 2-3 days using 0.025% trypsin-EDTA treatment.

pH 7.2 and 6.2 media were prepared replacing the carbonate by PIPES buffer (10 mM), and adjusting the pH with HCl or NaOH. pH 7.6 media (with carbonate buffer) is the same used for cell growth. pH 8.0 and 8.5 were prepared using 10 mM HEPES buffers adjusted to the desired pH.

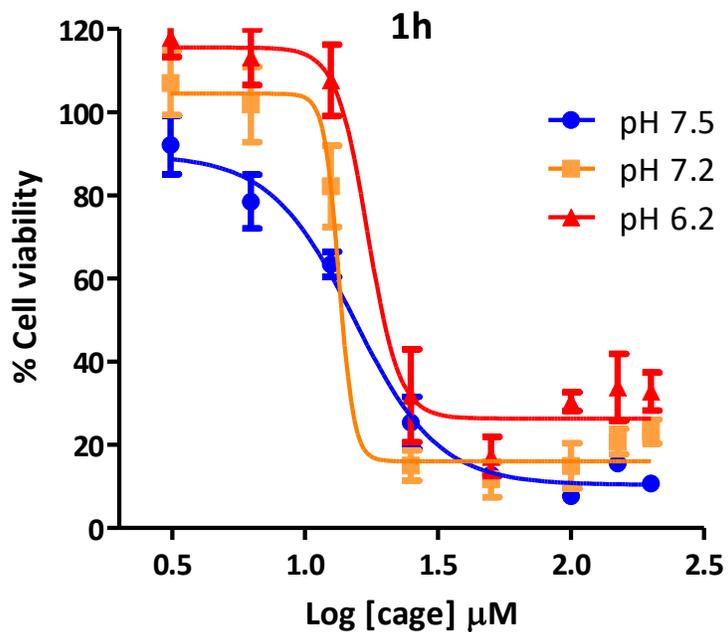
### **Cytotoxicity vs cancer cells**

The antiproliferative activity of compounds **1h-o** was evaluated using the MTT assay method. Cells were seeded 24 hours prior to treatment in 96-well plate (100  $\mu$ L of a suspension  $2.5 \cdot 10^5$  cell/mL). The following day, culture media from the wells was replaced by 100  $\mu$ L of fresh media at each pH containing the desired concentration of the compound to be tested. All the compounds were previously dissolved in DMSO at a concentration of 20 mM. The final concentration of DMSO used in the corresponding wells did not exceeded 1% (v/v). This concentration does not affect cell viability. Negative control cultures received the same concentration of solvent alone.

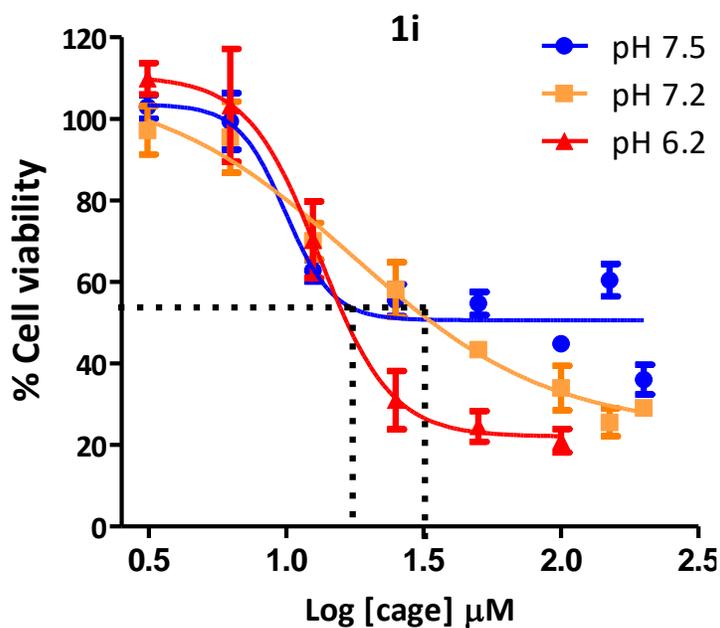
Cells were incubated for 24h in the presence of compounds **1h-o** at 37°C in a humidified atmosphere with 5% CO<sub>2</sub>. At the end of incubation, culture media was removed and 100  $\mu$ L of MTT solution (5 mg/mL diluted with plain culture media 1 : 5) was added to each well and incubated for 4 hours. Afterwards, MTT solution was discarded. The purple formazan crystal formed at the bottom of the wells was dissolved with 100  $\mu$ L of DMSO and stirred for 30 minutes at room temperature. The absorbance at 570 nm was read on a spectrophotometer plate reader. The proportion of surviving cells was calculated as (Absorbance of treated sample/Absorbance of control) x 100. Dose-response curves were constructed to obtain CC<sub>50</sub> values using Origin dose-response fitting function. All experimental data were derived from at least 3 independent experiments.

Control cells were grown in the same conditions (same cell culture media, buffer and pH) as treated cells to ensure that the difference in viability is only associated to the corresponding cages and not the buffer or the pH of the medium.

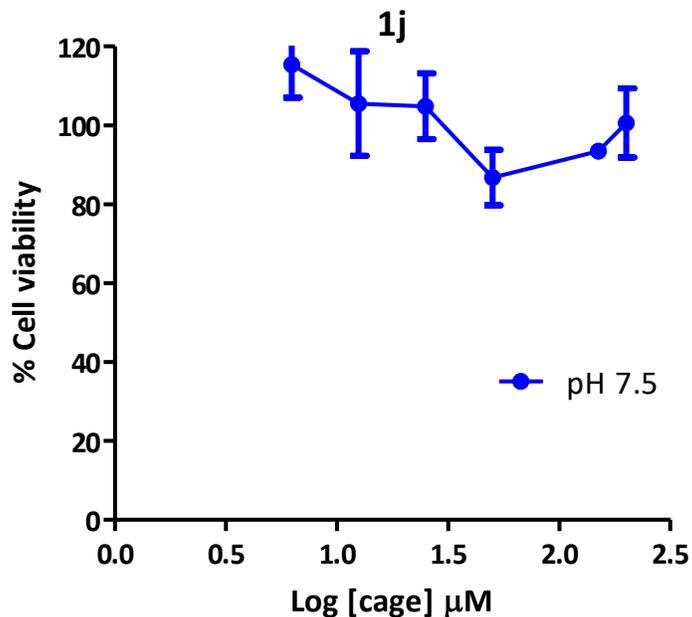
Statistical analysis: GraphPad Prism v5.0 software (GraphPad Software Inc., La Jolla, USA) was used for statistical analysis. For all experiments, the obtained results of the triplicates were represented as means with standard deviation (SD).



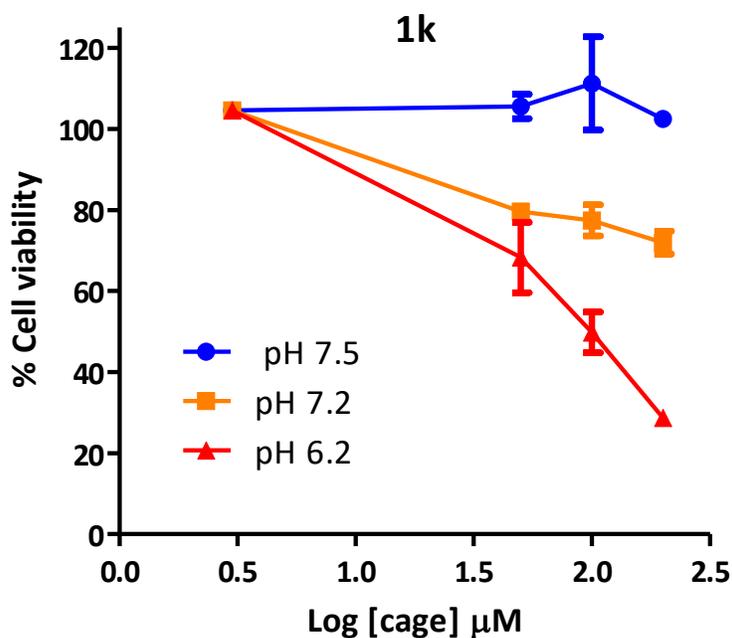
**Figure S44.** Cytotoxicity of compound **1h** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



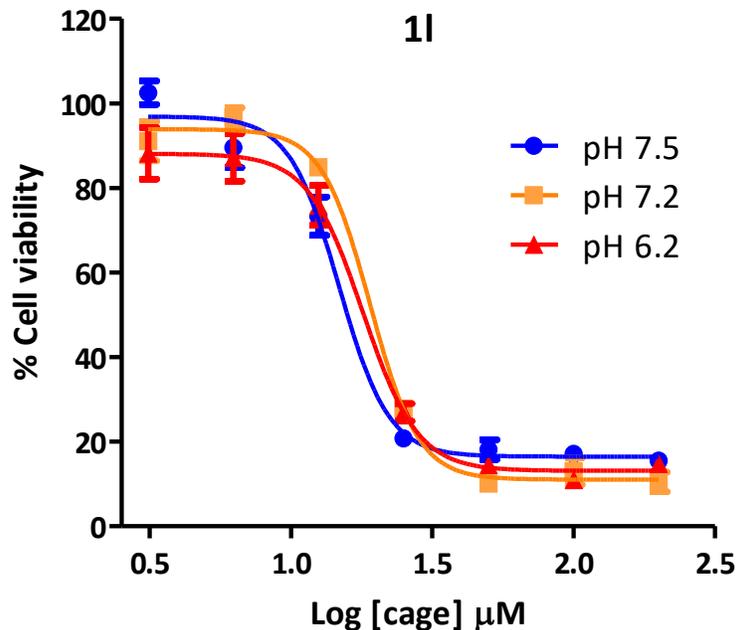
**Figure S45.** Cytotoxicity of compound **1i** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



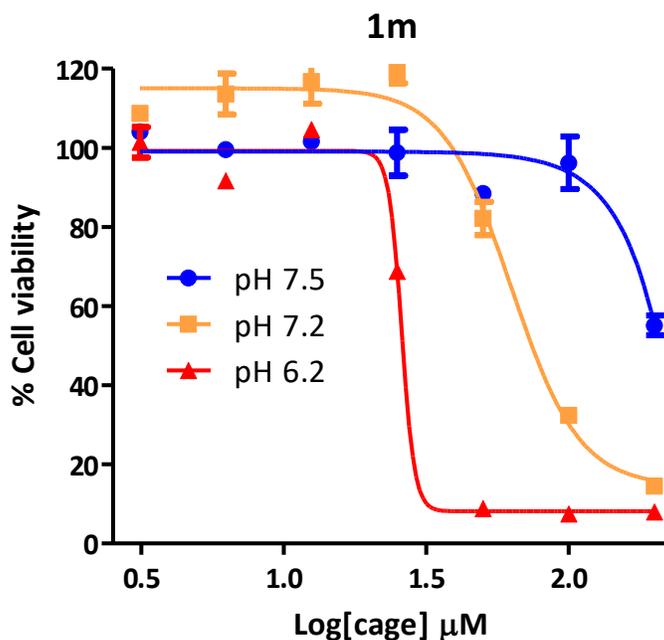
**Figure S46.** Cytotoxicity of compound **1j** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



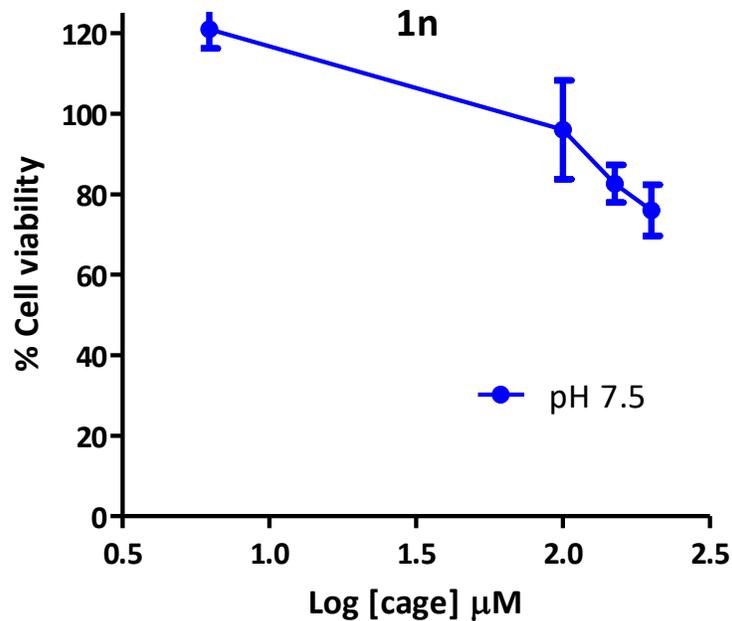
**Figure S47.** Cytotoxicity of compound **1k** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



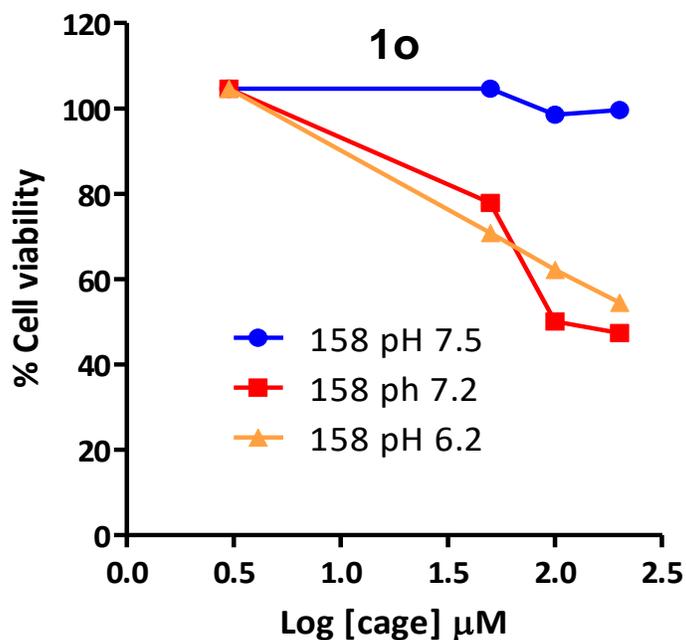
**Figure S48.** Cytotoxicity of compound **1l** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



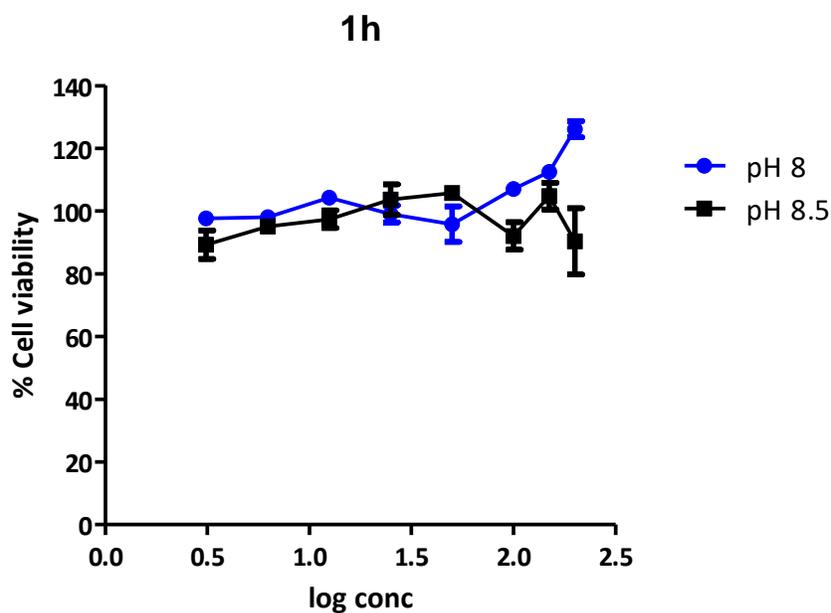
**Figure S49.** Cytotoxicity of compound **1m** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



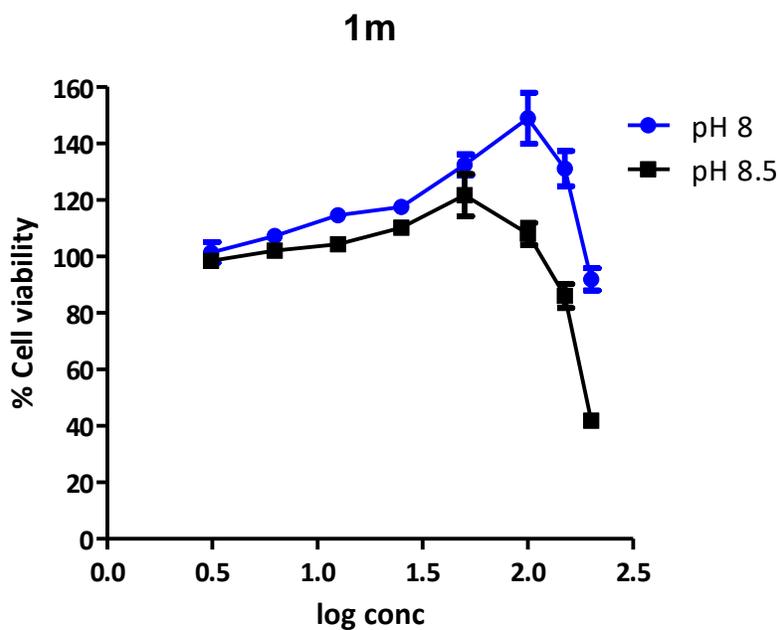
**Figure S50.** Cytotoxicity of compound **1n** in A549 cells at pH 7.5 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments. This compound shows very low aqueous solubility.



**Figure S51.** Cytotoxicity of compound **1o** in A549 cells at pH 7.5, 7.2 and 6.2 after 24 hours of incubation measured with MTT and dose-response calculated curves. Each trace corresponds to an average of at least three independent experiments.



**Figure S52.** Cytotoxicity of compound **1h** towards A459 cells at pH 8.0 and 8.5 after 24 hours of incubation measured with MTT. Each trace corresponds to an average of at least three independent experiments.

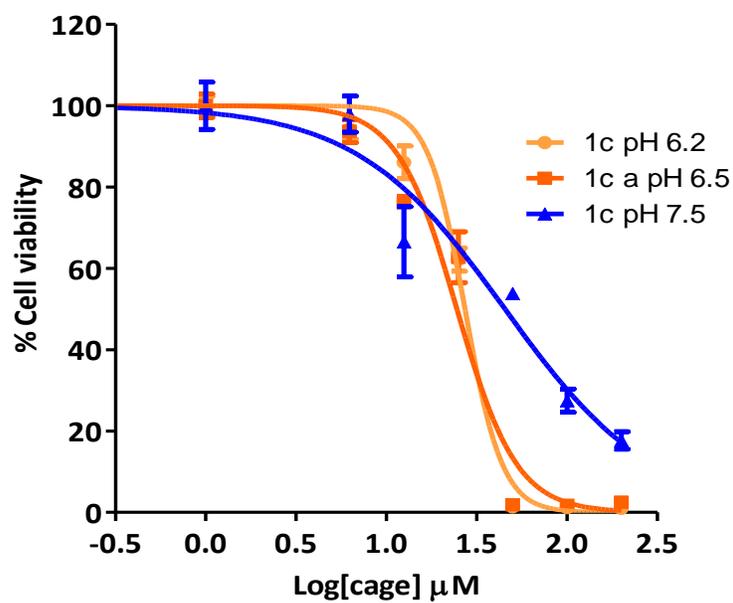


**Figure S53.** Cytotoxicity of compound **1m** towards A459 cells at pH 8.0 and 8.5 after 24 hours of incubation measured with MTT. Each trace corresponds to an average of at least three independent experiments.

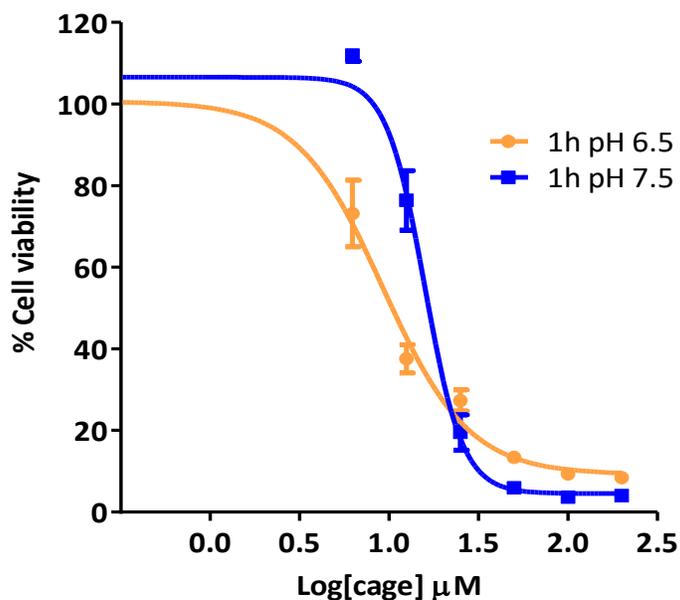
**3D models with Ca-alginate spheroids.** A549 cells incubated in a T-75 flask were suspended in DMEM and mixed in a 1 : 1 proportion with a Sodium Alginate 2.4 % solution in water. Cells suspension was prepared to get beads with 20000-40000 cells each. The mixture was taken up with a syringe coupled to a 24G needle. Then it was drop by drop poured into a 102 mM CaCl<sub>2</sub> solution with magnetic stirring forming the Ca-alginate spheroids. Finally, the alginate beads containing cells were washed with a 0.9% NaCl aqueous solution and transferred into DMEM.

A 96 well plate was used for the cytotoxicity assay. In each well a spheroid was suspended in 100 µL of DMEM buffer. After 24 hours, the buffer was replaced by the desired buffer (pH 6.2, 6.5 or 7.5) containing different concentrations of compounds **1c** or **1h**. Beads were incubated at 37°C with the cages for 48 hours. Then 20 µL of CellTiter-Blue were added to each well. The plate was incubated at 37°C for 3 hours and absorbance of each well was measured using a plate reader  $\lambda = 570$  nm.

The incubation time of the A549 in alginate spheroids with the cage compounds was 48 h since almost no cytotoxicity was detected at 24h, probably due to slower diffusion through the alginate gel. The cytotoxicity of compound **1h** to A549 in spheroids was equivalent to the cytotoxicity measured in a 2D culture (Figure S55). For compound **1c**, a higher cytotoxicity at neutral pH was also detected in spheroids compared to the 2D cell culture (Figure S54). The lower pH differences between the 2D and 3D models can be associated to the longer time exposures and to the possible interactions of the cage with alginate, as detected through fluorescence spectroscopy. Besides, buffering effects of the alginate gel micro-environment network could be playing a role at slightly acidic pH. Despite these slight differences, we have successfully recreated the cytotoxicity and the pH dependence of these two compounds from a simple 2D culture into a more complex three-dimensional system, which is more similar to a living tissue.



**Figure S54.** Cytotoxicity of compound **1c** at different pH in A549 cells present in alginate beads after 48 hours of incubation measured with resazurin and dose response calculated curves. Each trace corresponds to an average of at least three independent experiments.  $CC_{50}$  (pH 7.5) =  $45 \pm 10 \mu\text{M}$ ;  $CC_{50}$  (pH 6.5) =  $25 \pm 3 \mu\text{M}$ ;  $IC_{50}$  (pH 6.2) =  $27 \pm 4 \mu\text{M}$ .



**Figure S55.** Cytotoxicity of compound **1h** at different pH in A549 cells present in alginate beads after 48 hours of incubation measured with resazurin and dose response calculated curves. Each trace corresponds to an average of at least three independent experiments.  $CC_{50}$  (pH 7.5) =  $15 \pm 2 \mu\text{M}$ ;  $CC_{50}$  (pH 6.5) =  $9 \pm 2 \mu\text{M}$ .

## **Computational methods**

The package Schrödinger Suite 2022-4,<sup>7</sup> through its graphical interface Maestro,<sup>8</sup> was used to perform most modeling and visualization tasks. The program MacroModel<sup>9</sup> with the OPLS4 force field<sup>10</sup> and GB/SA water solvation conditions<sup>11</sup> was used for molecular mechanics energy minimization. The software Jaguar was used for DFT (Density Functional Theory) optimizations using the default B3LYP-D3 functional (B3LYP<sup>12,13</sup> plus Grimme dispersion corrections<sup>14</sup>) and the 6-31G\*\* basis set, with PBF implicit water solvation (the standard Poisson–Boltzmann continuum solvation model implemented in Jaguar<sup>15,16</sup>).

Starting from the crystal structure obtained for chloride-bound **1h**, compounds **1c**, **1h** and **1m** complexed with a single Cl<sup>-</sup> anion in the cage were modeled in different protonation states (ie. either with 1, 2, 3 or 4 protonated amino groups) and with 2 different conformations that differed on the orientation (clockwise or anticlockwise) of the Tren-axial arms of each cage. Since the axial nitrogen is in all cases the more basic one, it was always protonated. The Monte Carlo/Stochastic Dynamics (MD/SD) method implemented in MacroModel was used to sample the flexibility of the benzyl side chains of each of the 24 complexes thus generated. The following MD/SD settings were used: OPLS4 force-field with implicit water solvation, distance constraints were applied to keep the chloride anion in the cage, stochastic dynamics at 300 K with 1.5 fs timestep and with 10 and 1000 ps of equilibration and simulation time, 200 snapshots sampled at regular intervals and minimized. In this way, after removal of duplicate conformations using the Redundant Conformer Elimination module of Maestro, a number of distinct conformers were generated for each complex which were then submitted to further DFT optimization as described above. Removal of duplicates and of conformations with energies higher than 10 kcal·mol<sup>-1</sup> from the lowest energy minimum was then reapplied as above. The globularity ( $\Omega$ ) of each conformation, an index that measures how close to a spherical shape is the shape of a molecule, was calculated as the smallest eigenvalue divided by the largest eigenvalue of the covariance matrix of atomic coordinates ( $0 \leq \Omega \leq 1$ , being closer to 1 as the shape is closer to a sphere).

**Molecular Dynamics.** MD simulations of **1c** and **1h** in different states and interacting with a POPC model membrane were performed with AMBER 20<sup>17</sup> with resort to GPU acceleration.<sup>18–20</sup> The LIPID14 force-field<sup>21</sup> was used to parameterize the membrane POPC lipids, while GAFF2<sup>22</sup> parameters and atomic RESP charges<sup>23</sup> were used for the cage compounds. Cages **1c** and **1h** were simulated either free and neutral (no Cl<sup>-</sup> bound), or with a single bound Cl<sup>-</sup> and monoprotonated in the axial nitrogen (zero total charge). Simulations were performed in the presence of a POPC lipid bilayer, therefore two initial situations for the cage compounds were considered: either in the bulk water (W) or in the center of the POPC membrane (M).

Starting from the crystal structure obtained for chloride-bound **1h**, the structures of **1c** and **1h** either in neutral state or monoprotonated were built and optimized at HF/6-31G\* level with Gaussian 09 (Rev. E.01),<sup>24</sup> using the internal options `iop(6/33=2)`, `iop(6/42=6)` and `iop(6/50=1)` to write out the electrostatic potential (ESP) points and potentials, using 4 concentric layers per atom and 6 density points in each layer. The atomic charges of each molecule were then calculated by RESP fitting with Antechamber<sup>25</sup> and the simulation systems were built using the membrane bilayer builder of CHARMM-GUI.<sup>26–29</sup> To achieve the initial W or M cage placement situation, the CHARMM-GUI “Translate Molecule along Z axis” setting was

set to a value of 40 or 0 Å, respectively. In addition, the following settings were used: a hydration number of 50 (well above the hydration number for POPC lipid bilayers<sup>30</sup>), 64 POPC molecules per membrane leaflet, the replacement method for building the system and enough Na<sup>+</sup> and Cl<sup>-</sup> ions to achieve zero total charge and 0.15 M salt concentration. In this way, a PDB formatted file for each system was obtained which included, aside from the cage molecule, 128 POPC, 6400 TIP3P waters, 13 Na<sup>+</sup> and 13 or 14 Cl<sup>-</sup> ions (depending on the protonation state of the cage molecule), packed in an orthorhombic box. The systems corresponding to the complexes with one bound Cl<sup>-</sup> anion were achieved by manually modifying the coordinates of the last chloride ion in the PDB file to place it approximately in the center of the cage, with the assumption that during the equilibration step of the simulations the location of the anion would be adjusted to reach a stable position in the cage and that the bulk water molecules would occupy the cavity created by the movement of the chloride. The LEaP module from AMBER was then used to generate the topology files for the structures.

The systems thus prepared were minimized with 5000 steps of steepest-descent plus 5000 steps of conjugate-gradient algorithms, and then thermalized and equilibrated in three successive MD stages: (1) 25000 MD steps from 0 to 100 K, 2 fs per step, periodic boundary conditions (PBC) under the NVT ensemble, Langevin thermostat control and thermostat collision frequency of 1.0 ps<sup>-1</sup>, cutoff of 10 Å for short-range interactions and Particle Mesh Ewald (PME) method for long-range interactions, SHAKE on to constrain bonds to hydrogens, applying positional restraints to the solute and lipid molecules with a force constant of 10.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>; (2) 50000 steps from 100 to 303 K and 1 atm, PBC under the NPT ensemble with Berendsen barostat, anisotropic pressure scaling and pressure relaxation time of 2.0 ps, rest of conditions as before; (3) 2500000 steps at 300 K and 1 atm, no restraints for simulations of the free cages or distance restraints (5.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>) between the nitrogen atoms of the cage and the bound chloride atom for the chloride complexes, rest of conditions as before. Production NPT simulations (200 ns) were run using conditions as in step (3), saving frames and energy information every 10 ps (ie. 20000 frames per trajectory). Trajectory analysis was performed with CPPTRAJ<sup>31</sup> and visualization was performed with VMD-1.9.4a51<sup>32,33</sup> or PyMOL v. 2.5.4.<sup>34</sup>

**Steered Molecular Dynamics (SMD).** The thermalized systems prepared for the normal MD simulations (ie. those resulting from step (2) in the above MD protocol) were used as starting point to perform SMD simulations in the NPT ensemble. Thus, in order to study the interactions between the cages and the water and lipid environments, the free cages **1c** and **1h**, initially located in the water slab, were first equilibrated with 500000 steps MD (1 ns) applying distance restraints (5.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>) to keep their center of mass (COM) at Z-dimension = +40.0 Å relative to the COM of the phosphor atoms of the POPC molecules. Then, they were dragged across the lipid bilayer normal from Z = +40.0 to -40.0 Å at 0.5 Å ns<sup>-1</sup>, applying a force constant of 5.0 kcal mol<sup>-1</sup> Å<sup>-2</sup> to pull the COM of the cages. CPPTRAJ was used to determine the variability with simulation time of the number of hydrogen bonds between the cages and the molecules surrounding them, as well as the distances between the 4'-aromatic carbon atoms of the side arms of the cages and the tertiary nitrogen, as an indicator of the degree of folding of the side arms towards the cage core.

In order to generate starting configurations for Umbrella Sampling (US), SMD simulations were also performed starting with the free cages **1c** and **1h** located in the middle of the membrane (Z = 0.0), first equilibrating them as before and then pulling both in the positive or negative directions normal to the

plane of the model membrane, up to  $Z = +32.0$  or  $-32.0$  Å in 32 ns and applying a force constant of  $5.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ , as above. From these SMD simulations, 33 snapshots in each direction, with a 1 Å spacing between the cage-COMs of successive snapshots, were selected as starting configurations for independent US simulations that would allow the calculation of the Potential of Mean Force (PMF) profiles for the process of crossing the membrane. This approach was chosen because it has the advantage of showing less hysteresis for the calculation of PMFs than pulling from the water phase into the membrane.<sup>35</sup> The same SMD scheme was applied to a system containing a free chloride anion as solute in the center of the lipid bilayer, to determine the PMF of chloride crossing the lipid bilayer for comparison purposes.

Since AMBER allows to apply only a single restraint in SMD simulations, a slightly different approach had to be used to generate the initial US configurations for the chloride-bound complexes of protonated **1c** and **1h**. In this case, the cage-COM of each chloride complex was initially placed in the middle of the lipid bilayer ( $Z = 0.0$ ) and the system was submitted to a 1 ns MD equilibration step, applying distance restraints ( $5.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ ) to maintain the value of  $Z = 0.0$ . In addition, distance restraints were also applied between the chloride anion and the nitrogen atoms of the cage. Then, the last frame was used as input for a new equilibration step changing the location of the cage-COM to a value of  $Z = (\text{last } Z + 1.0)$ . This was repeated until  $Z = +32.0$  and the same was performed in the other direction (ie.  $Z = (\text{last } Z - 1.0)$  until  $Z = -32.0$ ). The last frame of each of these simulations was used as input configuration to generate independent US trajectories.

With these, two sets (positive and negative Z-direction) of US starting configurations were obtained for each of the five systems analyzed: the free chloride and the two cages (**1c** and **1h**) free or complexed to one chloride anion. Furthermore, the above protocols were repeated up to three times starting from different initial coordinates for each of the cage systems, in order to improve sampling.

**Umbrella Sampling (US) and Potential of Mean Force (PMF) calculation.** Umbrella sampling simulations with the 33 starting configurations for each set generated as described above were performed as follows. Systems were simulated for 30 ns (2 fs timestep) at 303 K and 1 atm, under the NPT ensemble, using distance restraints ( $5.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ ) to maintain the cage-COM at the corresponding Z-value (0, 1, 2, 3,... 32 or 0, -1, -2, -3,..., -32) and, for the simulations of the chloride complexes, also between the bound chloride anion and the nitrogen atoms of the cage. Frames were saved each 2 ps and the first 20 ns of each simulation were discarded as equilibration period. PMF profiles were obtained from the last 10 ns of simulation using the Weighted Histogram Analysis Method,<sup>36,37</sup> as implemented by A. Grossfield in the software WHAM.<sup>38</sup>

## Computational results

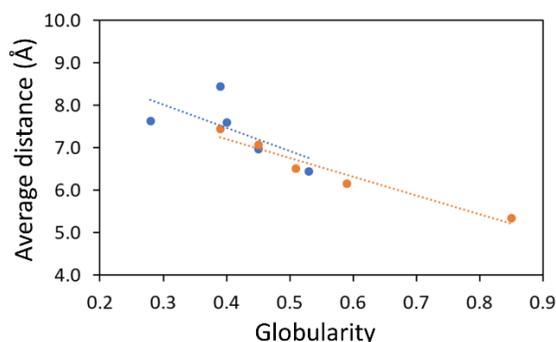
### DFT calculations.

A conformational search was conducted on compounds **1c** and **1h**, with varying protonation states (ranging from 1 to 4 protonated amino groups) and complexed with a single chloride anion, to identify potential conformations in solution. Following DFT-optimization and elimination of redundant conformations, a relatively large number of conformers were determined for each complex (Table S5).

**Table S5.** Number of DFT-optimized conformers generated for the Cl<sup>-</sup> complexes of compounds **1c** and **1h** with varying number of protonated amino groups and within 10 kcal·mol<sup>-1</sup> of the lowest energy minimum determined in each case.

	H <sup>+</sup>	2H <sup>+</sup>	3H <sup>+</sup>	4H <sup>+</sup>
<b>1c</b>	13	44	33	40
<b>1h</b>	32	71	21	105

Analysis of their DFT-optimized conformations revealed that these compounds can form several intramolecular interactions that stabilize the chloride complex (Tables S6-S7). Specifically, compound **1h**, featuring a 2'-fluorine substituent in the aromatic ring of the three arms of the cages, can establish intramolecular F··H-bonds with adjacent NH groups from amine and amide moieties. In contrast, compound **1c**, which only has 4'-fluorine substituents, lacks this capability. Consequently, low-energy conformations characterized by increased symmetry and reduced extension are observed for **1h** compared to **1c**, as evidenced by the globularity parameters determined for the conformations of each compound (Tables S6-S7). This globularity parameter correlates with shorter distances between the aromatic side rings and the axial N-atom, as illustrated in the graph presented in Figure S56 for the monoprotonated chloride complexes of **1c** and **1h**.

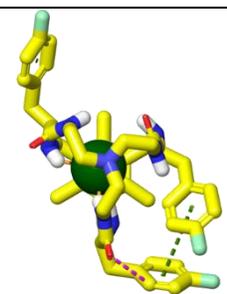
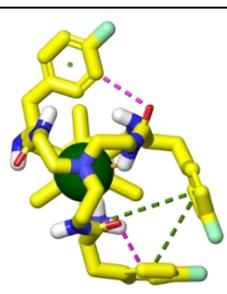
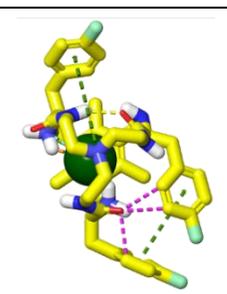
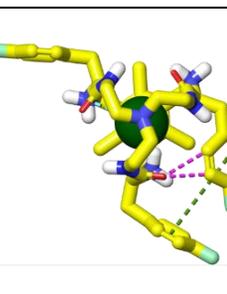
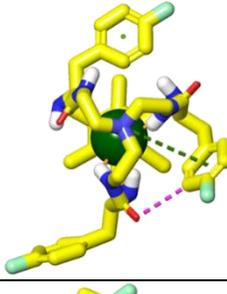
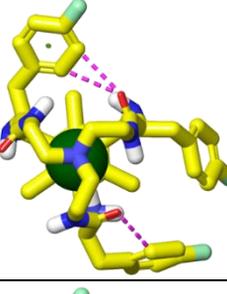
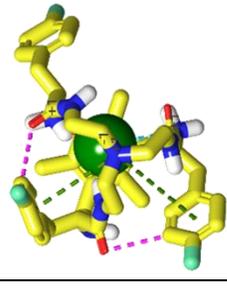
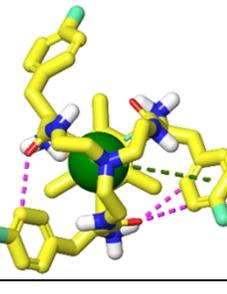
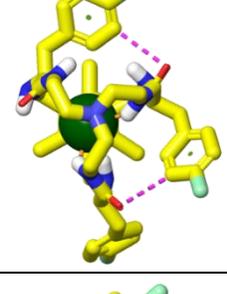
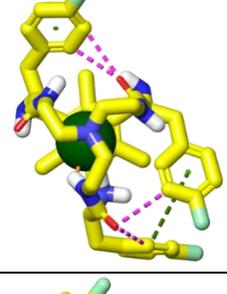
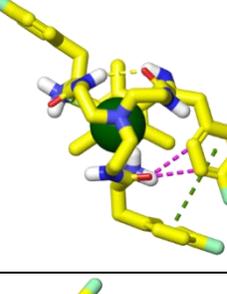
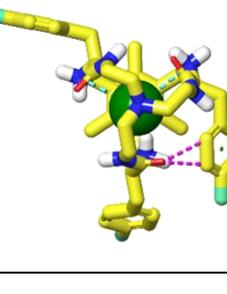
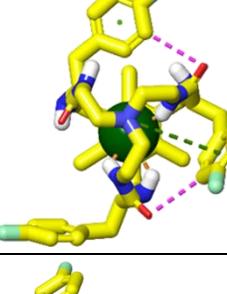
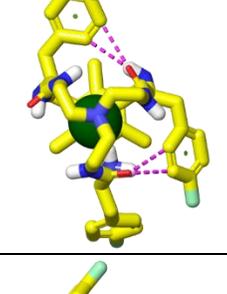
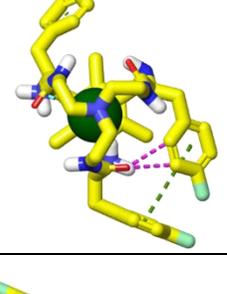
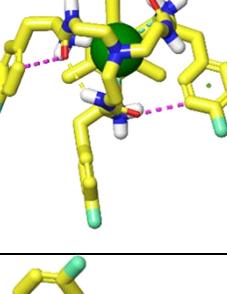
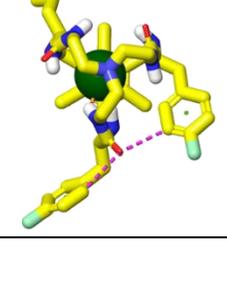
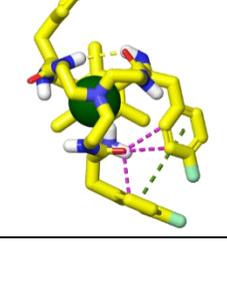
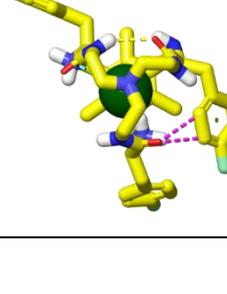
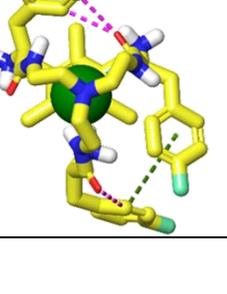


**Figure S56.** Average distance between the axial N-atom and the 4'-carbon atoms of the three side aromatic rings of the monoprotonated chloride complexes of **1c** (blue) and **1h** (orange) in the low energy conformations shown in Tables S6 and S7, represented against their corresponding globularity values.

Therefore, while it cannot be ruled out that a more exhaustive conformational search might uncover additional low-energy conformers, these results imply that conformers featuring such intramolecular

interactions could be relatively prevalent in solution. Moreover, it could be reasoned that they might experience even greater stabilization in hydrophobic environments such as lipid membranes. These conformers, particularly those with three F··H-bonds, exhibit a more compact structure compared to those where F··H interactions are absent. It can be hypothesized that these conformers, at least in part, contribute to the distinctive properties (e.g., chromatographic mobility and chloride transport) that set these compounds apart from those with a fluorine atom in different positions of the aromatic ring or with other types of substituents, and which are unable to adopt such compact conformations.

**Table S6.** Lowest energy DFT-optimized conformations (B3LYP-D3/6-31G\*\* level with PBF implicit water solvation) for the Cl<sup>-</sup> complexes of **1c** in different protonation states and properties: <sup>a</sup> relative solution phase energy (kcal mol<sup>-1</sup>), <sup>b</sup> globularity, <sup>c</sup> O··H-bonds (yellow dashed lines), <sup>d</sup> F··H-bonds (yellow), <sup>e</sup> Cl··H-bonds (orange), <sup>f</sup> aromatic H-bonds (magenta), <sup>g</sup> salt bridges (cyan) and <sup>h</sup> π-cation/π-π interactions (green).

Total charge:	0	+1	+2	+3			
	0.0 <sup>a</sup> 0.28 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 4 <sup>e</sup> 1 <sup>f</sup> 1 <sup>g</sup> 1 <sup>h</sup>		0.0 <sup>a</sup> 0.45 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		0.0 <sup>a</sup> 0.36 <sup>b</sup> 1 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 3 <sup>f</sup> 3 <sup>g</sup> 2 <sup>h</sup>		0.0 <sup>a</sup> 0.37 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 4 <sup>g</sup> 1 <sup>h</sup>
	2.25 <sup>a</sup> 0.45 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 4 <sup>e</sup> 1 <sup>f</sup> 1 <sup>g</sup> 1 <sup>h</sup>		0.04 <sup>a</sup> 0.39 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 3 <sup>f</sup> 2 <sup>g</sup> 0 <sup>h</sup>		2.25 <sup>a</sup> 0.45 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 3 <sup>g</sup> 2 <sup>h</sup>		1.18 0.56 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 3 <sup>f</sup> 4 <sup>g</sup> 1 <sup>h</sup>
	2.32 <sup>a</sup> 0.40 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 4 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 0 <sup>h</sup>		0.54 <sup>a</sup> 0.34 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 4 <sup>f</sup> 2 <sup>g</sup> 1 <sup>h</sup>		2.27 <sup>a</sup> 0.31 <sup>b</sup> 1 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		2.33 <sup>a</sup> 0.41 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 4 <sup>g</sup> 0 <sup>h</sup>
	3.11 <sup>a</sup> 0.53 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 4 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 1 <sup>h</sup>		0.75 <sup>a</sup> 0.40 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 4 <sup>f</sup> 2 <sup>g</sup> 0 <sup>h</sup>		2.27 <sup>a</sup> 0.37 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 1 <sup>e</sup> 2 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		2.49 <sup>a</sup> 0.53 <sup>b</sup> 1 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 4 <sup>g</sup> 0 <sup>h</sup>
	3.94 <sup>a</sup> 0.39 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 0 <sup>h</sup>		0.80 <sup>a</sup> 0.36 <sup>b</sup> 1 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 3 <sup>f</sup> 2 <sup>g</sup> 1 <sup>h</sup>		2.83 <sup>a</sup> 0.35 <sup>b</sup> 1 <sup>c</sup> 0 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 3 <sup>g</sup> 0 <sup>h</sup>		2.89 <sup>a</sup> 0.37 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 3 <sup>e</sup> 3 <sup>f</sup> 4 <sup>g</sup> 1 <sup>h</sup>

**Table S7.** Lowest energy DFT-optimized conformations (B3LYP-D3/6-31G\*\* level with PBF implicit water solvation) for the Cl<sup>-</sup> complexes of **1h** in different protonation states and properties: <sup>a</sup> relative solution phase energy (kcal mol<sup>-1</sup>), <sup>b</sup> globularity, <sup>c</sup> O··H-bonds (yellow dashed lines), <sup>d</sup> F··H-bonds (yellow), <sup>e</sup> Cl··H-bonds (orange), <sup>f</sup> aromatic H-bonds (magenta), <sup>g</sup> salt bridges (cyan) and <sup>h</sup> π-cation/π-π interactions (green).

Total charge:	0	+1	+2	+3			
	0.0 <sup>a</sup> 0.51 <sup>b</sup> 0 <sup>c</sup> 1 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 2 <sup>h</sup>		0.0 <sup>a</sup> 0.59 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		0.0 <sup>a</sup> 0.37 <sup>b</sup> 2 <sup>c</sup> 2 <sup>d</sup> 3 <sup>e</sup> 0 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		0.0 <sup>a</sup> 0.43 <sup>b</sup> 2 <sup>c</sup> 3 <sup>d</sup> 2 <sup>e</sup> 0 <sup>f</sup> 4 <sup>g</sup> 2 <sup>h</sup>
	0.63 <sup>a</sup> 0.85 <sup>b</sup> 0 <sup>c</sup> 3 <sup>d</sup> 4 <sup>e</sup> 3 <sup>f</sup> 1 <sup>g</sup> 3 <sup>h</sup>		2.04 <sup>a</sup> 0.56 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 2 <sup>e</sup> 2 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		6.15 <sup>a</sup> 0.47 <sup>b</sup> 0 <sup>c</sup> 3 <sup>d</sup> 4 <sup>e</sup> 1 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		0.08 <sup>a</sup> 0.64 <sup>b</sup> 0 <sup>c</sup> 3 <sup>d</sup> 1 <sup>e</sup> 3 <sup>f</sup> 4 <sup>g</sup> 3 <sup>h</sup>
	0.82 <sup>a</sup> 0.59 <sup>b</sup> 0 <sup>c</sup> 3 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 2 <sup>h</sup>		2.13 <sup>a</sup> 0.44 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		6.59 <sup>a</sup> 0.40 <sup>b</sup> 1 <sup>c</sup> 2 <sup>d</sup> 2 <sup>e</sup> 1 <sup>f</sup> 3 <sup>g</sup> 3 <sup>h</sup>		0.88 <sup>a</sup> 0.67 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 2 <sup>e</sup> 4 <sup>f</sup> 4 <sup>g</sup> 2 <sup>h</sup>
	3.13 <sup>a</sup> 0.45 <sup>b</sup> 0 <sup>c</sup> 0 <sup>d</sup> 5 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 1 <sup>h</sup>		2.71 <sup>a</sup> 0.37 <sup>b</sup> 2 <sup>c</sup> 1 <sup>d</sup> 3 <sup>e</sup> 0 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		7.61 <sup>a</sup> 0.49 <sup>b</sup> 1 <sup>c</sup> 2 <sup>d</sup> 2 <sup>e</sup> 1 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		0.96 <sup>a</sup> 0.34 <sup>b</sup> 0 <sup>c</sup> 1 <sup>d</sup> 2 <sup>e</sup> 1 <sup>f</sup> 4 <sup>g</sup> 2 <sup>h</sup>
	3.61 <sup>a</sup> 0.39 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 4 <sup>e</sup> 2 <sup>f</sup> 1 <sup>g</sup> 1 <sup>h</sup>		2.99 <sup>a</sup> 0.58 <sup>b</sup> 0 <sup>c</sup> 1 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 2 <sup>g</sup> 2 <sup>h</sup>		8.11 <sup>a</sup> 0.43 <sup>b</sup> 0 <sup>c</sup> 2 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 3 <sup>g</sup> 1 <sup>h</sup>		0.99 <sup>a</sup> 0.62 <sup>b</sup> 0 <sup>c</sup> 3 <sup>d</sup> 3 <sup>e</sup> 2 <sup>f</sup> 4 <sup>g</sup> 3 <sup>h</sup>

## Molecular Dynamics Simulations.

Molecular dynamics simulations were conducted to investigate the interaction and translocation of the pseudopeptidic cages across a model POPC lipid bilayer. To examine their passive diffusion, cages **1c** and **1h** in their neutral state (i.e., with all amine groups unprotonated) were simulated under two distinct initial conditions: within the aqueous phase (W) and at the center of the POPC model membrane (M). Figures S57-top and S58 illustrate that diffusion of both compounds from the bulk water into the POPC membrane occurs during the initial phase of the simulation. Subsequently, they stabilize within a region where their center of mass (COM) is approximately 5 to 10 Å below the membrane surface. Remarkably, a similar location was reached in simulations where the cages were initially placed at the center of the membrane, indicating independence from the starting placement of the cages. Comparable findings were observed when simulations were conducted with cages **1c** and **1h** monoprotonated at the tertiary nitrogen and complexed with a single chloride anion (Figures S57-middle and S59). In these simulations, restraints were necessary to prevent dissociation of the anion from the complex. Notably, without restraints, rapid dissociation of the chloride occurred when the complexes were initially positioned in the bulk water (not shown). Similarly, when the complexes were placed at the middle of the lipid bilayer, both unrestrained complexes persisted for the initial 15-30 ns of the simulations. However, upon reaching the water-lipid interface, the chloride was released, and the cages remained within the same region as before (Figures S57-bottom and S60). ***These results collectively suggest that the most stable location for these cages, whether free or complexed with a chloride anion, lies within a region below the surface of the membrane near the water-lipid interface.*** This positioning facilitates their interaction with the polar and non-polar moieties of POPC, as well as with the water molecules and ions at the interface, thereby fulfilling the requirements for their role as anion transporters.

Regarding the orientation adopted by the cages relative to the membrane, it can be approximated by assessing the relative positions of the axial N-atom and the centroid of the axial aromatic ring in relation to the lipid surface. Figure S61 illustrates, for each of the aforementioned simulations, the temporal variation of the distance between the axial nitrogen or the centroid of the axial aromatic ring and the lipid surface ( $d(N)$  and  $d(CAr)$ , respectively; positive values indicate the location of the N or the centroid in the water phase, while negative values indicate their location in the lipid phase). Considering a vector between the centroid and the axial N-atom, if  $d(N) > d(CAr)$ , that is, if  $D = d(N) - d(CAr)$  is positive, it signifies that the vector is oriented towards the positive direction of the Z-axis (perpendicular to the plane of the lipid surface). Conversely, if  $D < 0$ , it indicates orientation towards the negative direction, and if  $D = 0$ , it suggests that the vector is parallel to the lipid surface. Figure 6 also presents the plots of this  $D$ -value over time, revealing that when the cages are in the water phase (i.e., initial phase of the **1c(W)**, **1h(W)**, **1c/Cl(W)**, and **1h/Cl(W)** simulations),  $D$  fluctuates rapidly between positive and negative values,

indicating a rapid tumbling motion with no clear orientation preference. In contrast, when the cages are in the lipid phase (i.e., later phase of the **1c(W)**, **1h(W)**, **1c/Cl<sup>-</sup>(W)**, and **1h/Cl<sup>-</sup>(W)** simulations, and all the **M** simulations), such fluctuations are much less frequent or virtually absent, consistent with the higher viscosity of the environment which restricts the mobility of the cages. Although there is no distinct orientation tendency in the simulations of the neutral cages, the simulations of the chloride complexes reveal that when the cages are within the lipid, they tend to orient with the protonated nitrogen towards the membrane surface (i.e.,  $D > 0$ ). This tendency could be attributed to interactions between the positive nitrogen and the negatively charged carboxylates present in the polar moiety of the POPC molecules.

Moreover, steered molecular dynamics (SMD) simulations were employed to characterize the process of fully crossing the membrane. To accomplish this, **1c** and **1h**, initially located in the bulk water, were pulled through the water phase, across the membrane, and then again through the water phase at the other side (Figure 62A). During this translation, as previously observed, both cages exhibited a rapid tumbling motion when in the water phase and a significantly slower tumbling when at the lipid-water interface or inside the lipid phase (Figure 62B). Analysis of hydrogen bonding interactions (Figure 62C) revealed that both cages formed some intramolecular H-bonds, more in the case of compound **1h** compared to **1c**, and only a few interactions with the POPC molecules. However, both cages engaged in a larger number of interactions with water molecules, even when inside the membrane. This is attributed to several water molecules accompanying the diffusing cages during the SMD simulations, although the number of H-bonds with water substantially decreased as the cages approached the bilayer core. This observation aligns with findings reported by Marques et al. for their simulations of squaramide anion transporters.<sup>39</sup> To examine the shape of the cages, the distance between the aromatic 4'-carbon of the three aromatic side rings and the axial N-atom was determined. The graphs in Figure 62D indicated that during these simulations, both compounds predominantly adopted conformations where these distances were around 5-6 Å, with occasional transitions to conformations where they increased to 8-10 Å. As mentioned previously, shorter distances can be associated with conformations featuring a more folded disposition for the side ring moieties and higher globularity, while longer distances can be associated with conformations displaying a more extended disposition of the side rings and lower globularity. Consequently, these graphs suggest that under the conditions of these SMD simulations, both compounds, particularly **1c**, predominantly exhibited a folded conformation with intermittent transitions to more extended ones. However, this result should be interpreted cautiously because the force applied to pull the molecules during the SMD simulations could influence their conformation, potentially favoring the more folded ones which would pose a lower resistance during the pulling process.

#### **Umbrella sampling (US) simulations and Potential of Mean Force (PMF) calculations.**

To explore the energetic barriers involved in the process of membrane crossing, the PMF profiles for **1c** and **1h**, both in neutral state or monoprotinated and complexed with one chloride ion, were calculated

by combining US simulations with the Weighted Histogram Analysis Method (WHAM).<sup>36–38</sup> Additionally, a PMF profile for the crossing of a free chloride anion was computed for comparison purposes.

Prior to these, SMD simulations were conducted for each system to generate the starting positions along the membrane's normal. Two approaches were considered for these SMD simulations: either pulling the cages from the water phase across the membrane to the opposite side, as detailed in the previous section, or pulling the cages from the middle of the membrane to the water phase on both sides. Previous research has shown that the latter option can lead to faster PMF convergence in certain cases.<sup>35</sup> Considering that pulling from the water phase into the lipid may carry associated water molecules, potentially influencing the PMF results, we opted for this second option, i.e. to pull the cages from inside the lipid towards the water phase on both sides. Furthermore, taking into account the slow tumbling within the lipid phase, we anticipated that the direction of pulling would yield different results depending on the initial orientation of the cages. Therefore, to better represent the actual process, we conducted up to three series of US simulations in each direction, ranging from  $Z=0$  to  $+32$  or  $-32$  Å in 1 Å increments, for each cage, either free or complexed. At each  $Z$ -value,  $Z$ -restrained simulations were carried out for 30 ns, with only the final 10 ns utilized for the PMF calculations (i.e. discarding the initial 20 ns as equilibration period). Figure S63 illustrates the results of the six PMF profiles obtained for each case, demonstrating acceptable reproducibility among the three profiles in each direction. Since the observed transport process in reality would involve the average of many crossing events of numerous cages through the membrane in either direction, we considered that a more accurate depiction of the real PMF profile could be obtained by averaging the six individual profiles for each case and assuming that the real energy profiles should be symmetric. Figure 5B of the main text displays the results of this averaging, representing the cumulative sampling of almost 2  $\mu$ s for each cage (33 simulations, 10 ns per simulation across 6 series of simulations). Additionally, it also presents the PMF profile for the crossing of a free chloride anion.

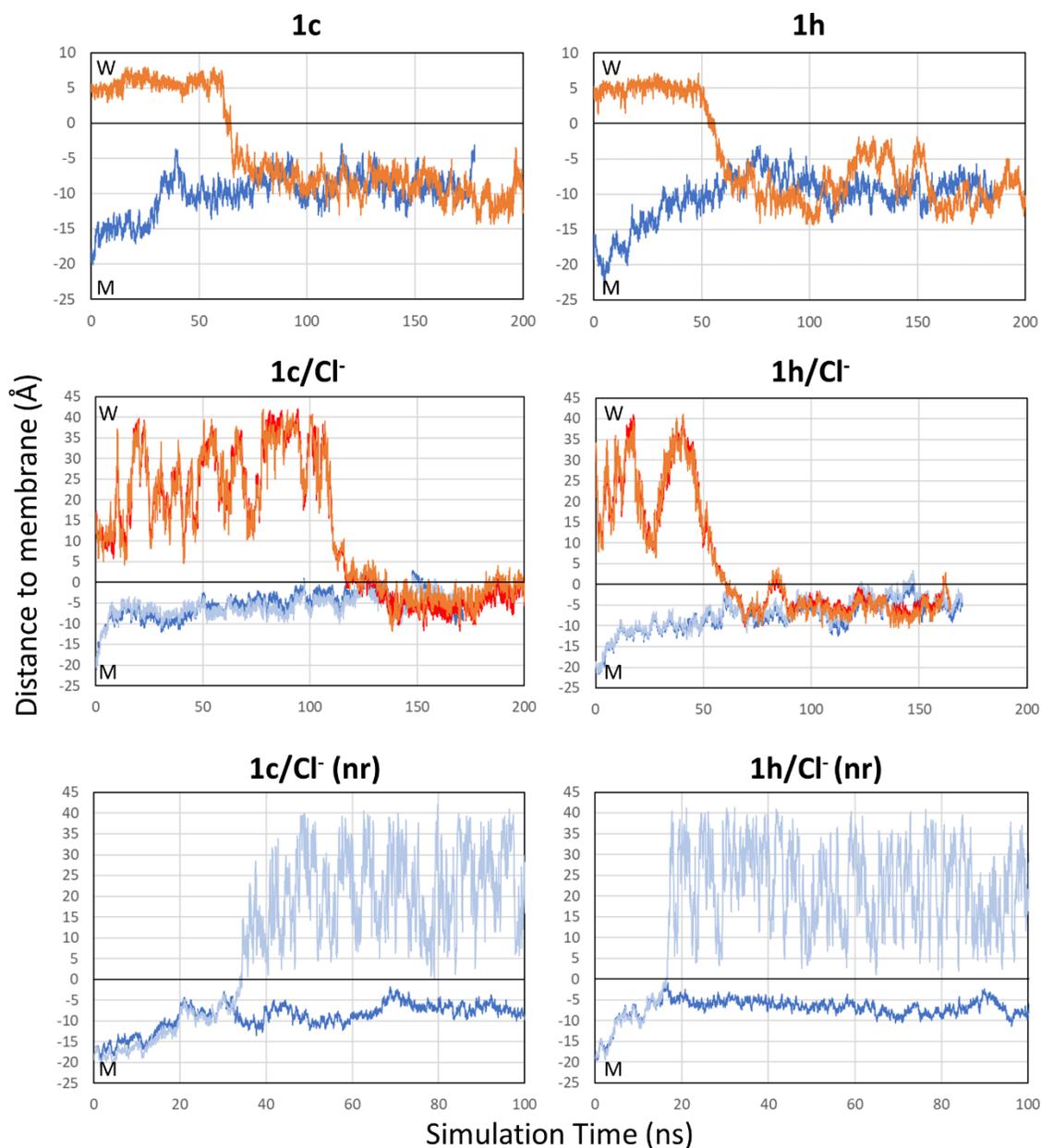
Therefore, according to the PMF-profiles obtained, the crossing of a single chloride anion through the model POPC membrane exhibits an energetic barrier of about 21 kcal mol<sup>-1</sup> at  $Z = 0$  (i.e. at the middle of the membrane). This energy maximum is relative to an arbitrary value of 0 when the chloride anion is in the bulk water, and it reasonably aligns with previously reported data ( $\sim 18$  kcal mol<sup>-1</sup>).<sup>39</sup> As anticipated for anion carriers, this barrier disappears for the chloride complexes of monoprotonated **1c** and **1h**, indicating no net energy expense in transitioning from one side of the membrane to the other. Surprisingly, the PMF-profiles for both cage-chloride complexes are nearly identical, which would suggest that the chloride-transport capacities of both cages are largely comparable. Hence, their PMF-profiles exhibit energetic minima of approximately -10 kcal mol<sup>-1</sup> at  $Z = \pm 12.5$  Å, separated by a maximum around -2 kcal mol<sup>-1</sup> at  $Z = 0$ . Considering that the membrane extends approximately from  $Z = +20$  to  $-20$  Å, these energetic minima correspond to the location where the chloride-bound cages were previously observed to stabilize in the simulations depicted in Figures S57 and S59. In contrast, when looking at the PMF-profiles of neutral and free **1c** and **1h** some remarkable differences emerge both among themselves and in comparison to their

complexed forms. First, despite both **1c** and **1h** exhibit energy minima at  $Z = \pm 10 \text{ \AA}$ , these energy basins are approximately  $5 \text{ kcal mol}^{-1}$  deeper for the former ( $-16$  vs  $-11 \text{ kcal mol}^{-1}$ ), with both minima being separated by energy maxima at approximately  $4 \text{ kcal mol}^{-1}$  above ( $-12$  and  $-7 \text{ kcal mol}^{-1}$ ) in both cages. Again, the energy minima observed for **1c** and **1h** align with the location where both cages were observed to stabilize in the simulations depicted in Figures S57 and S58. This stable location near the water/lipid interface is appropriate for the interaction of the cages with water and ion molecules close to the lipid surface, facilitating their role as ion binders. The lower energy values for free **1c** compared to free **1h** suggest that the former is more stabilized in the lipid environment, a finding consistent with the experimentally observed differences in lipophilicity (Table 1 of main text). Moreover, while the PMF profiles of the chloride-bound cages display minimal energetic differences, the disparities observed between free **1c** and **1h** suggest that the former faces a higher energy barrier to overcome in reaching the water phase ( $16$  vs  $11 \text{ kcal mol}^{-1}$ ), or even the membrane surface at  $Z = \pm 20 \text{ \AA}$  (approximately  $8$  vs  $6 \text{ kcal mol}^{-1}$  for **1c** and **1h**, respectively), from its preferred location within the membrane.

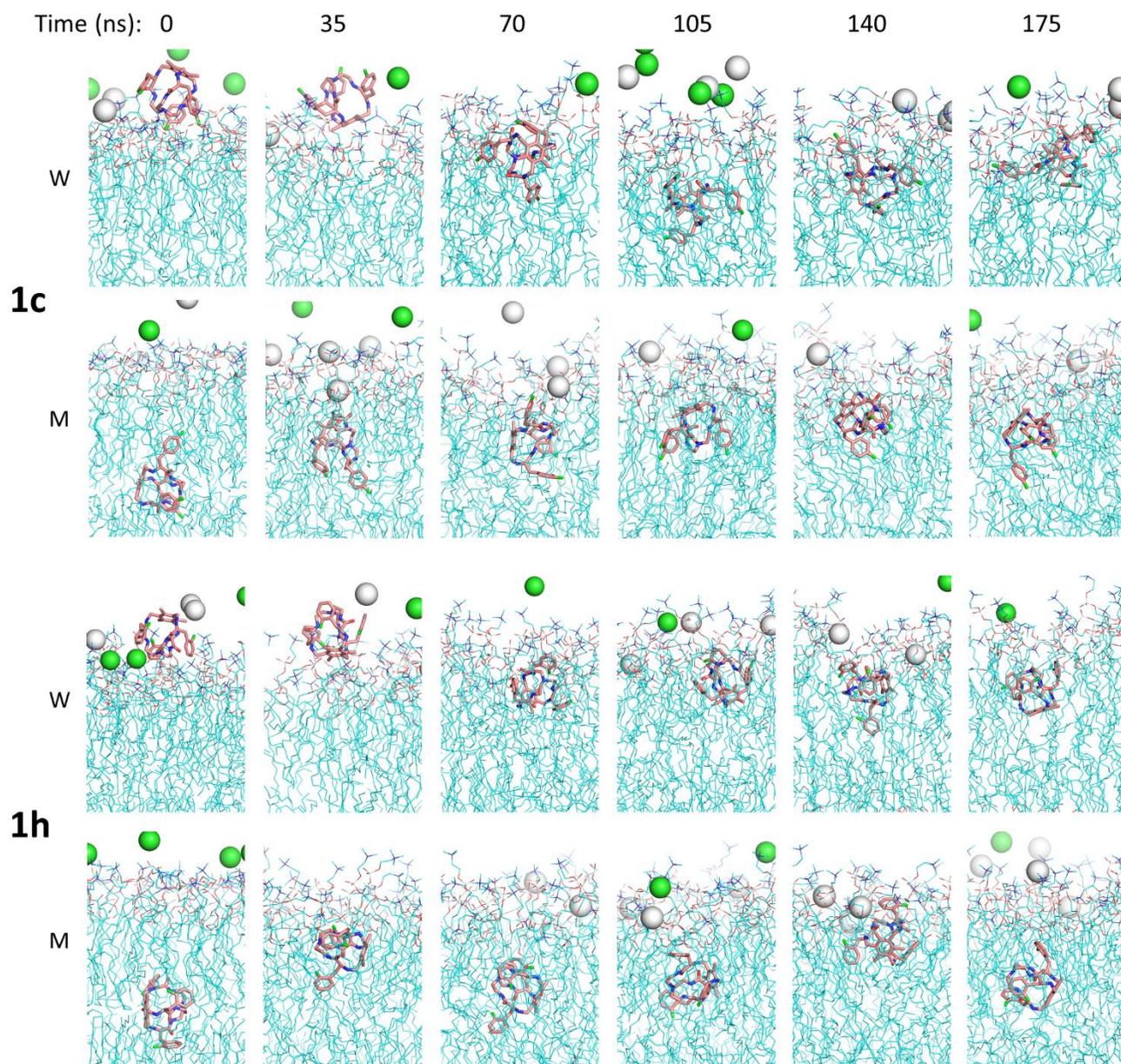
Formally, chloride transport can be conceptualized as a two-step cycle: first, the translation of the chloride-bound protonated cage from one side of the membrane to the opposite side, where it releases the anion along with a proton, to keep charge neutrality at both sides; second, the return of the neutral cage back to the initial side of the membrane to undergo reprotonation and load a new anion, thus reinitiating the cycle. Therefore, the above results suggest that despite the first step of the cycle is energetically similar for both **1c** and **1h**, the second step may be slightly more kinetically favored for **1h** relative to **1c**. Remarkably, this tendency is in line with the observed transport rates for each cage (Table 1 of main text).

To elucidate the structural basis for these differences, we examined the geometry of the cages during the US simulations and calculated the distances between the 4'-carbon atom on each aromatic side-ring and the axial nitrogen. This served as a way to determine whether the side arms of the cages adopt folded or extended conformations during the simulations. Figure S64 presents the frequency diagrams of these distances. Observations reveal that the most frequent distances for cage **1c** fall within the ranges of  $7.00$ - $7.25$ ,  $5.00$ - $5.25$ , and  $8.25$ - $8.50 \text{ \AA}$ , whereas for **1h**, these distances are concentrated within the range of  $5.00$ - $5.25 \text{ \AA}$  for the three aromatic rings. This suggests that while the side arms of cage **1c** predominantly adopt a semi-extended conformation, with one folded arm and the other two more or less extended, those of **1h** predominantly assume conformations where they are mostly folded and close to the axial nitrogen. This finding correlates well with the observed NOEs for **1h** in lipid micelles (Figure 4, main text). Regarding the chloride complexes of both monoprotonated cages, Figure S64 indicates that their distances are considerably larger, suggesting predominant conformations where the side arms are mostly in an extended disposition. These structural data could help explain the observed differences in the PMF profiles. It could be speculated that the predominantly folded conformations for **1h** would expose less surface area of the aromatic side chains to the environment, thereby reducing the number of potential

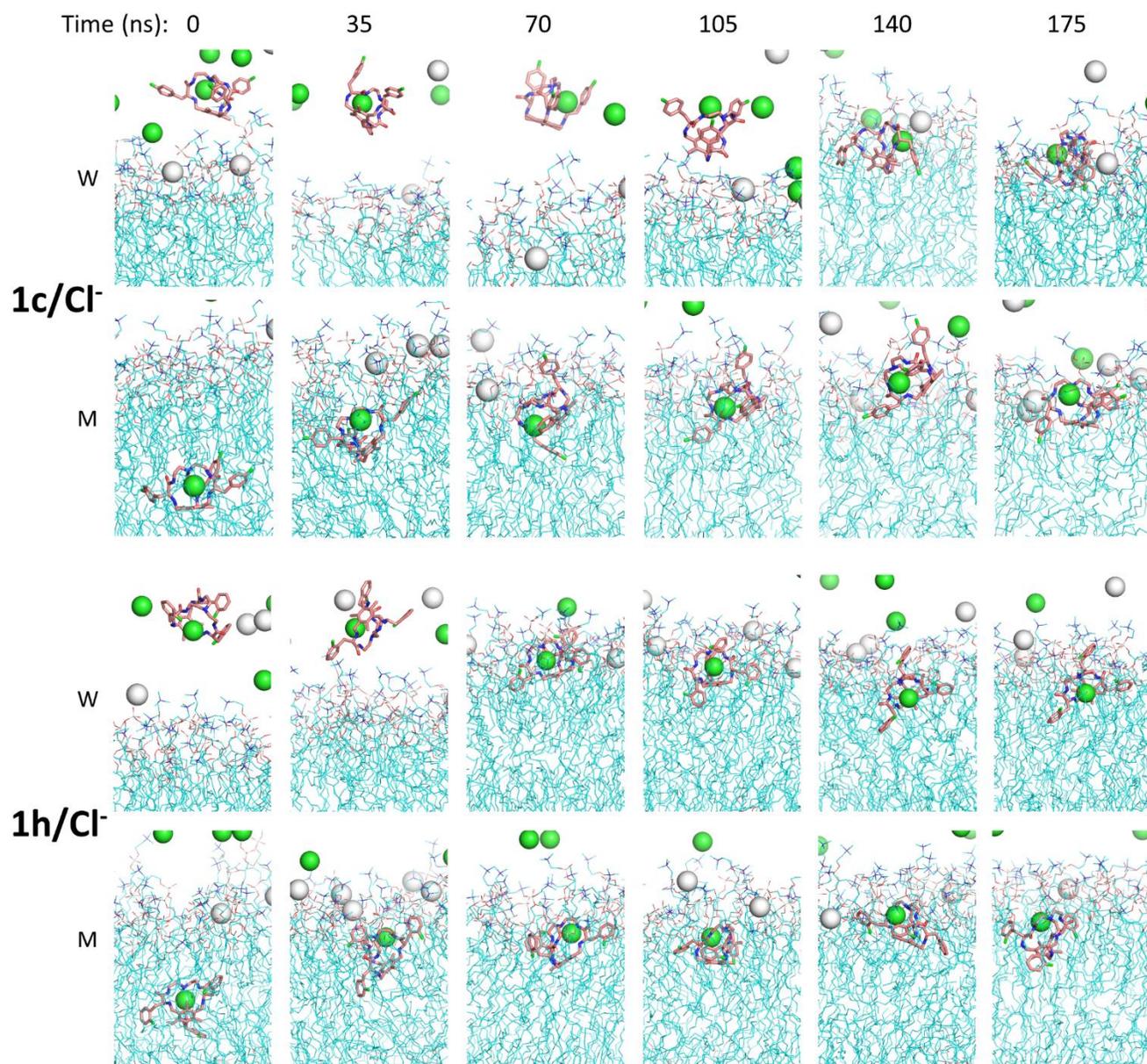
hydrophobic interactions with the membrane. Conversely, the more unfolded conformations predominant in **1c** would allow for the establishment of more hydrophobic interactions with the membrane, thereby preferentially stabilizing it relative to **1h**. Moreover, when both cages are monoprotonated and bound to chloride, the electrostatic repulsion between the negatively charged chloride and the electronic cloud of the aromatic side-rings would favor more extended conformations for both. However, the potential to establish more hydrophobic interactions between the aromatic rings and the lipids could be compensated by the lower stability of a charged complex in a hydrophobic environment, as observed.



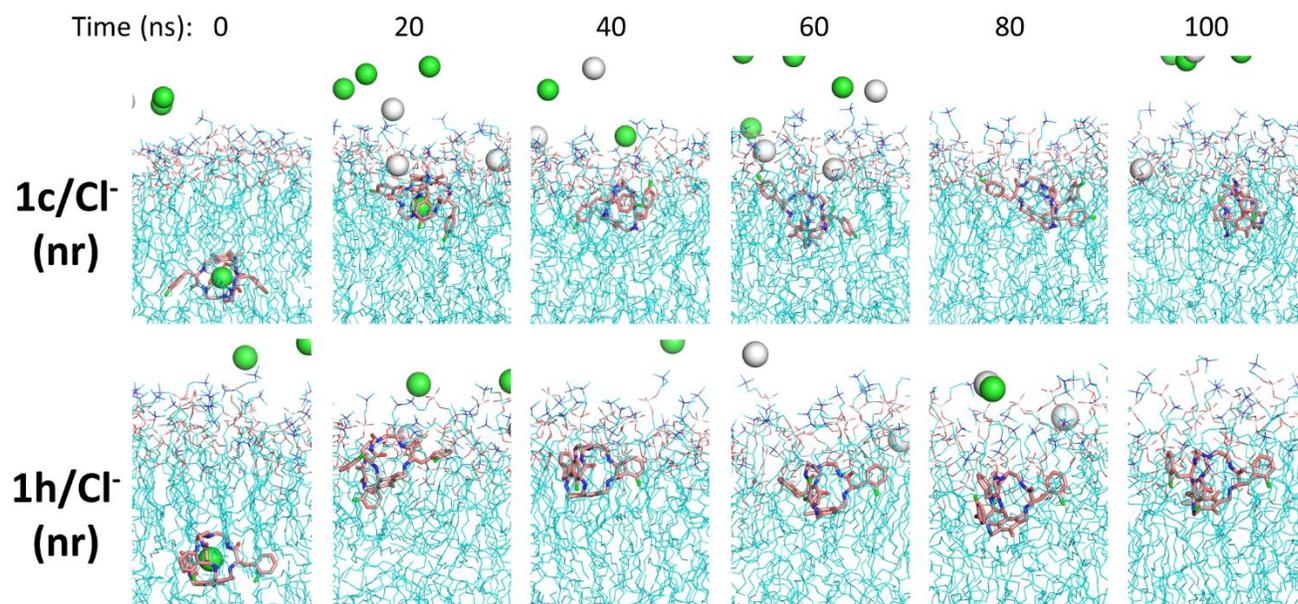
**Figure S57.** Diffusion across the water/lipid interface in simulations of **1c** and **1h** in different states: free (**1c** and **1h** simulations), monoprotated and chloride bound with distance restrictions between the chloride and the nitrogen atoms of the cage (**1c/Cl<sup>-</sup>** and **1h/Cl<sup>-</sup>**), and protonated and chloride bound without restrictions (**1c/Cl<sup>-</sup> (nr)** and **1h/Cl<sup>-</sup> (nr)**). The initial location of the cage was either in the water phase (W) or in the middle of the membrane (M). Graphs show the distance between the cage center of mass (blue or orange) or the chloride anion (light blue or red) and the membrane surface (represented by a black line at ordinate 0). Good superposition of cages and chloride traces indicates that the anion remains bound to the cage.



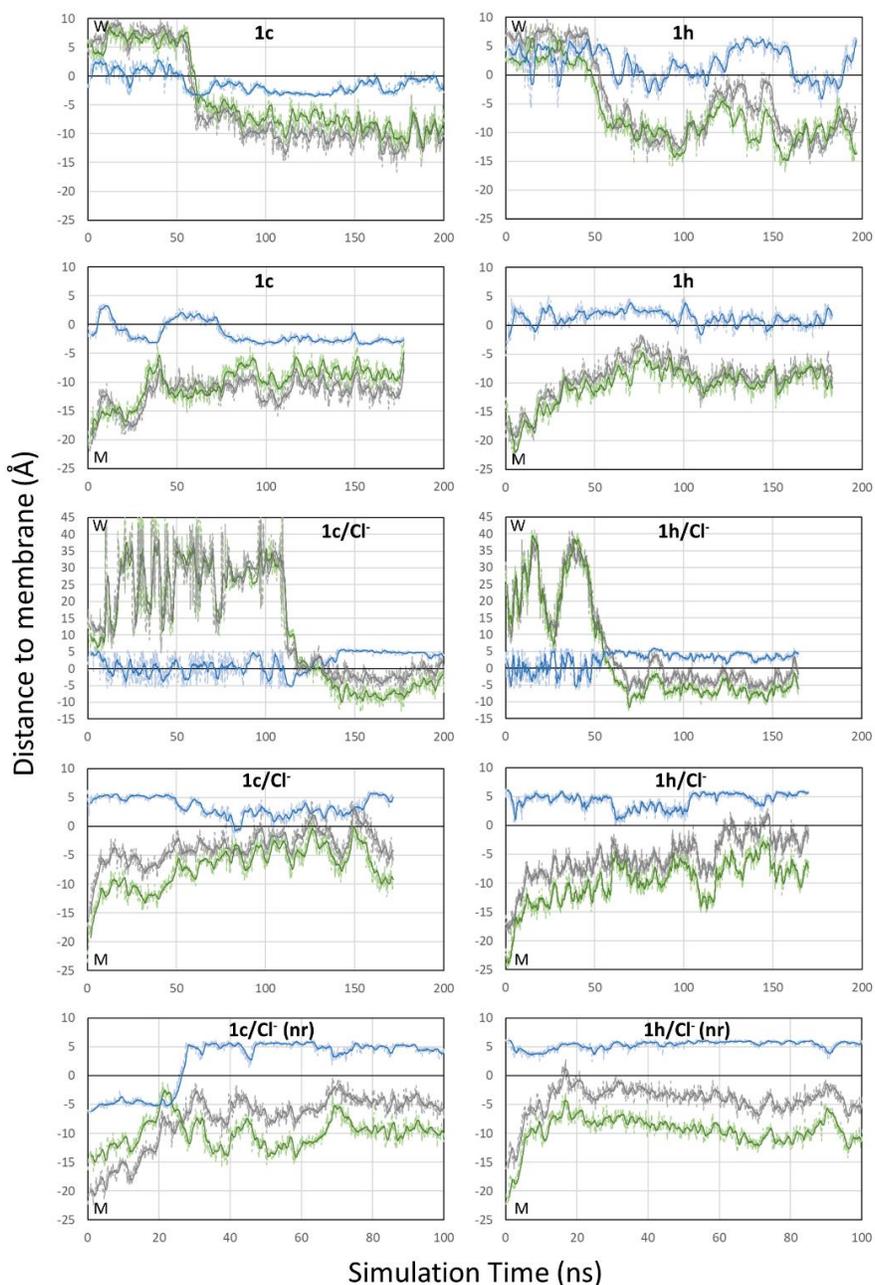
**Figure S58.** Snapshots of the simulations of unprotonated **1c** and **1h** that illustrate their free diffusion through the lipid membrane with time. W (water) and M (membrane) indicate the initial location of the cage. Water molecules and hydrogen atoms are omitted for clarity.



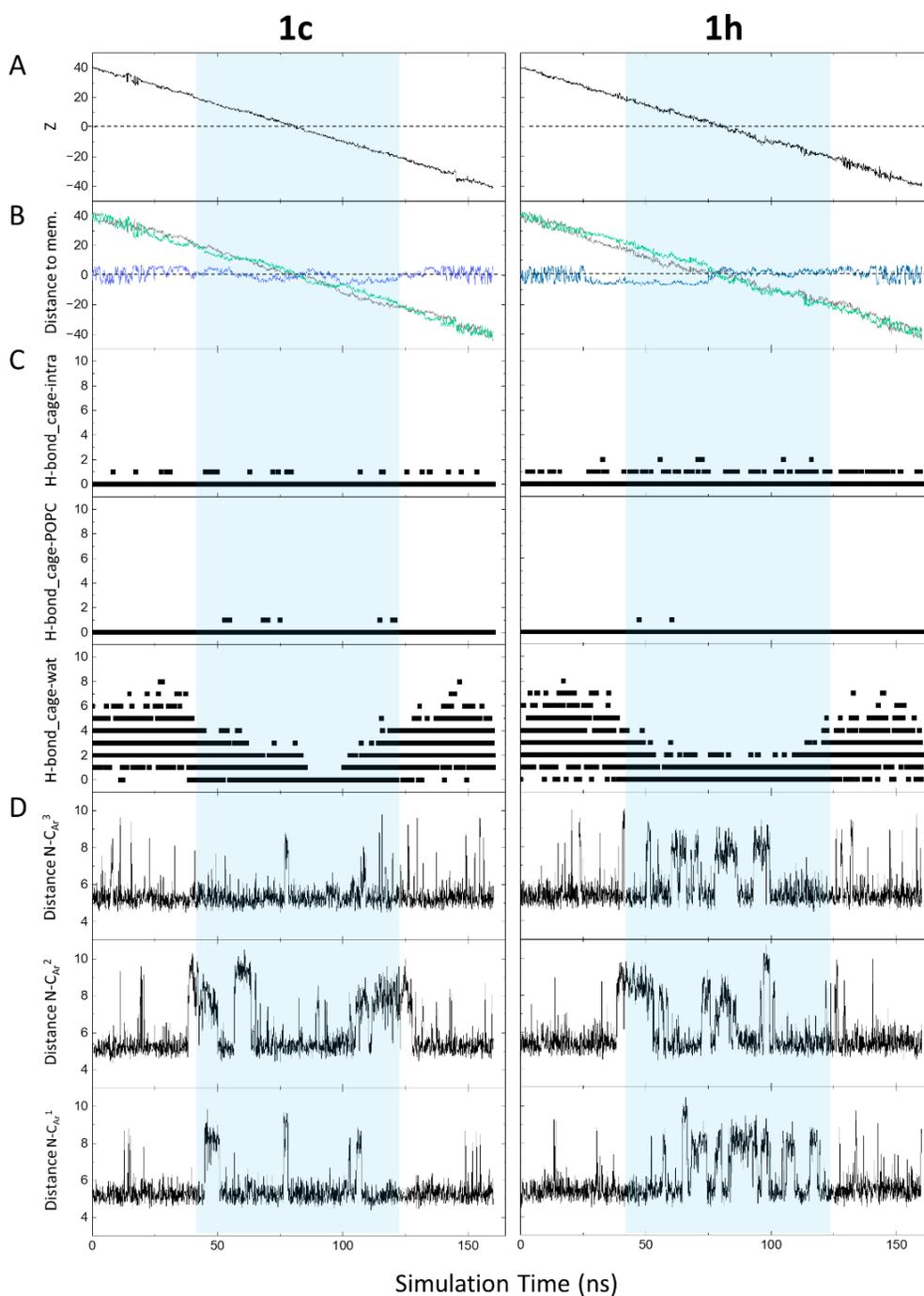
**Figure S59.** Snapshots of the simulations of **1c** and **1h**, protonated and complexed with one chloride anion, which illustrate their diffusion through the lipid membrane with time. W (water) and M (membrane) indicate the initial location of the cage. In these simulations distance restraints were applied to avoid dissociation of the bound chloride.



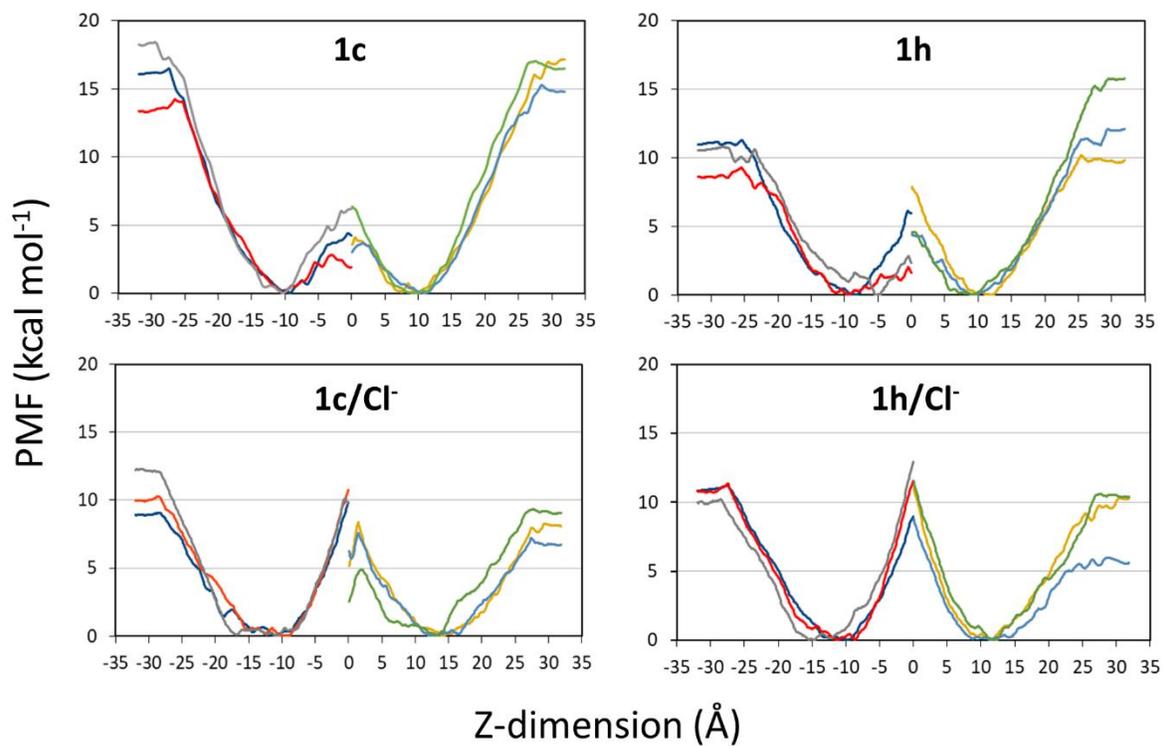
**Figure S60.** Snapshots of the simulations of **1c** and **1h** protonated and complexed to chloride, which illustrate their diffusion through the lipid membrane with time. Complexes were initially located in the middle of the membrane and no restraints (nr) were applied during the simulation.



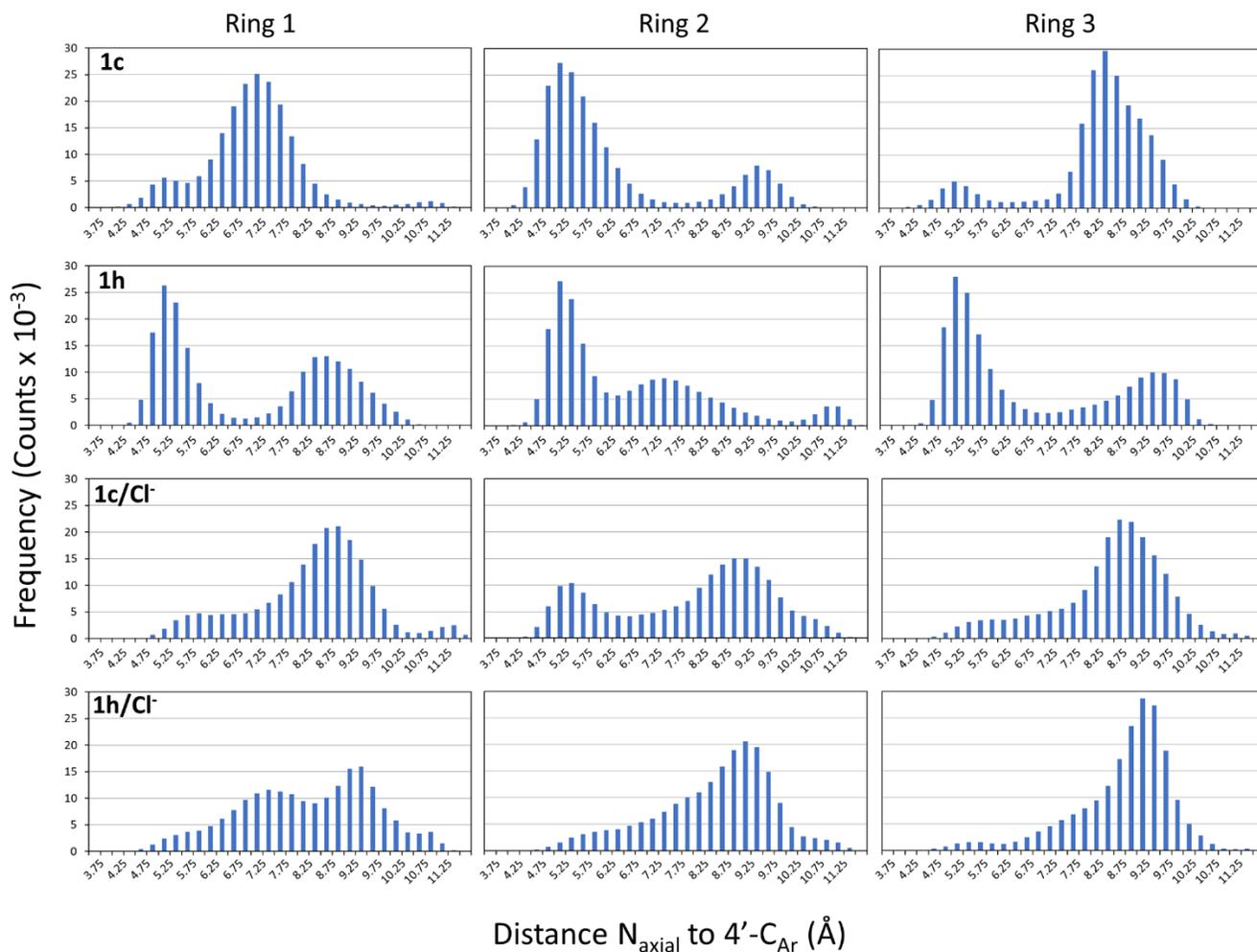
**Figure S61.** Orientation of cages **1c** and **1h** relative to the membrane surface during the simulations (see Figure S57 legend for conditions of each simulation). Graphs show the distance between the membrane surface and the axial N-atom ( $d(N)$ , gray) or the centroid of the axial aromatic ring C-atoms ( $d(C_{Ar})$ , green), as well as the difference between both distances ( $D = d(N) - d(C_{Ar})$ , blue). Discontinuous and continuous lines correspond to raw and smoothed data, respectively. Considering a vector between the centroid of the axial aromatic ring C-atoms and the axial N-atom, a value of  $D = 0$  indicates that this vector is parallel to the surface of the membrane, a  $D > 0$  value indicates that the vector is oriented towards the positive direction of the Z-axis (which is perpendicular to the plane of the membrane surface) and a  $D < 0$  value indicates that it is oriented towards the negative direction of the Z-axis.



**Figure S62.** Results from the steered molecular dynamics simulations of unprotonated **1c** and **1h** crossing the membrane from  $Z = 40$  to  $-40$  Å in 160 ns. (A) Distance of the cage center of mass relative to the membrane central plane ( $Z$  coordinate). (B) Distance between the membrane central plane, represented by a dashed line at  $Z=0$ , and the axial N-atom (gray) or the centroid of the axial aromatic ring (green), and the difference between both (blue). (C) Number of cage hydrogen bonds: intramolecular, with POPC and with water molecules. (D) Graphs of distances between the axial nitrogen and the three aromatic carbons in 4'-position for each cage. White and blue background regions indicate when the cages are in the aqueous or membrane phase, respectively.



**Figure S63.** Raw results from the PMF calculations for **1c** and **1h**, and the chloride complexes of both monoprotonated cages. Each graph represents the results of three US series with Z-values between 0 and 32 Å and three with Z-values between 0 and -32 Å.



**Figure S64.** Frequency distribution of the distances between the axial nitrogen and the aromatic carbon in 4'-position of each of the three side rings of each cage. Data correspond to the US simulations of **1c** and **1h** and their corresponding monoprotonated forms complexed with chloride (**1c/Cl<sup>-</sup>** and **1h/Cl<sup>-</sup>**), and each graph represents the sum of about  $2 \times 10^5$  snapshots and almost 2  $\mu$ s of simulation time (33 simulations from 0 to 32 Å relative to membrane central plane, 1001 snapshots per simulation, 10 ns of simulation time per simulation, 6 repeats).

## References

1. Sheldrick, G.M. (2008). A short history of SHELX. *Acta Crystallogr. Sect. A* *64*, 112–122. <https://doi.org/10.1107/S0108767307043930>.
2. Spek, A.L. (2009). Structure validation in chemical crystallography. *Acta Crystallogr. Sect. D* *65*, 148–155. <https://doi.org/10.1107/S090744490804362X>.
3. Sheldrick, G.M. (2015). Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C* *71*, 3–8. <https://doi.org/10.1107/S2053229614024218>.
4. Martí, I., Bolte, M., Burguete, M.I., Vicent, C., Alfonso, I., and Luis, S. V. (2014). Tight and selective caging of chloride ions by a pseudopeptidic host. *Chem. - A Eur. J.* *20*, 7458–7464. <https://doi.org/10.1002/chem.201303604>.
5. Martí, I., Rubio, J., Bolte, M., Burguete, M.I., Vicent, C., Quesada, R., Alfonso, I., and Luis, S. V. (2012). Tuning chloride binding, encapsulation, and transport by peripheral substitution of pseudopeptidic tripodal small cages. *Chem. - A Eur. J.* *18*, 16728–16741. <https://doi.org/10.1002/chem.201202182>.
6. Tapia, L., Pérez, Y., Bolte, M., Casas, J., Solà, J., Quesada, R., and Alfonso, I. (2019). pH-Dependent Chloride Transport by Pseudopeptidic Cages for the Selective Killing of Cancer Cells in Acidic Microenvironments. *Angew. Chemie - Int. Ed.* *58*, 12465–12468. <https://doi.org/10.1002/anie.201905965>.
7. Schrödinger (2022). Release 2022-4 at Schrödinger, LLC.
8. Schrödinger Release 2022-4 (2022). Maestro, version 13.4 at Schrödinger, LLC.
9. Schrödinger Release 2022-4 (2022). MacroModel, version 13.8 at Schrödinger, LLC.
10. Lu, C., Wu, C., Ghoreishi, D., Chen, W., Wang, L., Damm, W., Ross, G.A., Dahlgren, M.K., Russell, E., Von Bargen, C.D., et al. (2021). OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J. Chem. Theory Comput.* *17*, 4291–4300. <https://doi.org/10.1021/acs.jctc.1c00302>.
11. Still, W.C., Tempczyk, A., Hawley, R.C., and Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* *112*, 6127–6129. <https://doi.org/10.1021/ja00172a038>.
12. Becke, A.D. (1993). A new mixing of Hartree-Fock and local-density-functional theories. *J. Chem. Phys.* *98*, 1372–1377. <https://doi.org/10.1063/1.464304>.
13. Lee, C., Yang, W., and Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B Condens. Matter* *37*, 785–789. <https://doi.org/10.1103/PhysRevB.37.785>.
14. Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* *132*, 154104–154119. <https://doi.org/10.1063/1.3382344>.
15. Schrödinger (2022). Release 2022-4: Jaguar at Schrödinger, LLC.
16. Bochevarov, A.D., Harder, E., Hughes, T.F., Greenwood, J.R., Braden, D.A., Philipp, D.M., Rinaldo,

- D., Halls, M.D., Zhang, J., and Friesner, R.A. (2013). Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem.* *113*, 2110–2142. <https://doi.org/10.1002/QUA.24481>.
17. Case, D.A., Belfon, K., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, T.E.I.I.I., Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Giambasu, G., et al. (2020). AMBER 2020 at University of California.
  18. Goetz, A.W., Williamson, M.J., Xu, D., Poole, D., Grand, S.L., and Walker, R.C. (2012). Routine microsecond molecular dynamics simulations with AMBER - Part I: Generalized Born. *J. Chem. Theory Comput.* *8*, 1542–1555. <https://doi.org/10.1021/ct200909j>.
  19. Salomon-Ferrer, R., Goetz, A.W., Poole, D., Grand, S.L., and Walker, R.C. (2013). Routine microsecond molecular dynamics simulations with AMBER - Part 2: Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* *9*, 3878–3888. <https://doi.org/10.1021/ct400314y>.
  20. Grand, S.L., Goetz, A.W., and Walker, R.C. (2013). SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.* *184*, 374–380. <https://doi.org/10.1016/j.cpc.2012.09.022>.
  21. Dickson, C.J., Madej, B.D., Skjervik, Å.A., Betz, R.M., Teigen, K., Gould, I.R., and Walker, R.C. (2014). Lipid14: The Amber Lipid Force Field. <https://doi.org/10.1021/ct4010307>.
  22. He, X., Man, V.H., Yang, W., Lee, T.S., and Wang, J. (2020). A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* *153*. <https://doi.org/10.1063/5.0019056>.
  23. Bayly, C.I., Cieplak, P., Cornell, W.D., and Kollman, P.A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* *97*, 10269–10280. <https://doi.org/10.1021/J100142A004>.
  24. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., et al. Gaussian 09, Revision E.01; Gaussian, Inc.; Wallingford CT, 2009.
  25. Wang, J., Wang, W., Kollman, P.A., and Case, D.A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* *25*, 247–260. <https://doi.org/10.1016/j.jmglm.2005.12.005>.
  26. CHARMM-GUI. <https://www.charmm-gui.org/>.
  27. Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* *29*, 1859–1865. <https://doi.org/10.1002/JCC.20945>.
  28. Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V., Venable, R.M., et al. (2014). CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* *35*, 1997–2004. <https://doi.org/10.1002/JCC.23702>.
  29. Jo, S., Lim, J.B., Klauda, J.B., and Im, W. (2009). CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* *97*, 50–58. <https://doi.org/10.1016/J.BPJ.2009.04.013>.
  30. Kučerka, N., Tristram-Nagle, S., and Nagle, J.F. (2006). Structure of fully hydrated fluid phase lipid

- bilayers with monounsaturated chains. *J. Membr. Biol.* **208**, 193–202. <https://doi.org/10.1007/s00232-005-7006-8>.
31. Roe, D.R., and Cheatham 3rd, T.E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **9**, 3084–3095. <https://doi.org/10.1021/ct400341p>.
  32. VMD <http://www.ks.uiuc.edu/Research/vmd/>.
  33. Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD - Visual Molecular Dynamics. *J. Molec. Graph.* **14**, 33–38.
  34. Schrödinger (2023). The PyMOL Molecular Graphics System, Version 2.5 at Schrödinger, LLC.
  35. Filipe, H.A.L., Moreno, M.J., Róg, T., Vattulainen, I., and Loura, L.M.S. (2014). How to tackle the issues in free energy simulations of long amphiphiles interacting with lipid membranes: Convergence and local membrane deformations. *J. Phys. Chem. B* **118**, 3572–3581. <https://doi.org/10.1021/jp501622d>.
  36. Kumar, S., Rosenberg, J.M., Bouzida, D., Swendsen, R.H., and Kollman, P.A. (1992). THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021. <https://doi.org/10.1002/jcc.540130812>.
  37. Roux, B. (1995). The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91**, 275–282. [https://doi.org/10.1016/0010-4655\(95\)00053-l](https://doi.org/10.1016/0010-4655(95)00053-l).
  38. Grossfield, A. WHAM: the weighted histogram analysis method at Grossfield Lab, University of Rochester, New York, 2017. [http://membrane.urmc.rochester.edu/?page\\_id=126](http://membrane.urmc.rochester.edu/?page_id=126).
  39. Marques, I., Costa, P.M.R., M, Q.M., Busschaert, N., Howe, E.N.W., Clarke, H.J., Haynes, C.J.E., Kirby, I.L., Rodilla, A.M., Perez-Tomas, R., et al. (2018). Full elucidation of the transmembrane anion transport mechanism of squaramides using in silico investigations. *Phys Chem Chem Phys* **20**, 20796–20811. <https://doi.org/10.1039/c8cp02576b>.