Supporting Information

Development of a cyclic, cell penetrating peptide compatible with in vitro selection strategies

Nicolas A. Abrigo[a,b], Kara K. Dods[a,b], Chelsea A. Makovsky [a,b], Sandeep Lohan [a,b], Koushambi Mitra[a,b], Kaylee M. Newcomb[a,b], Anthony V. Le[a,b] and Matthew C. T. Hartman[*a,b]

**E-mail: mchartman@vcu.edu (Matthew C. T. Hartman)*

[a] *Chemistry, Virginia Commonwealth University, 1001 W Main Street, Richmond, 23284, VA, United States*

[b] *Massey Cancer Center, Virginia Commonwealth University, 401 College St, Richmond, 23219, Virginia, United States. Tel.: 8046284095*

Table of Contents:

Reagents **S3**

Methods **S3**

References **S5**

Scheme S1 ..... **S5**

Figure S1 ..... **S6**

Figure S2 ..... **S7**

Figure S3 ..... **S8**

Figure S4 ..... **S9**

Figure S5 ..... **S10**

Figure S6 ..... **S11**

Figure S7 ..... **S12**

Figure S8 ..... **S13**

Figure S9 ..... **S14**

Figure S10 ..... **S15**

Figure S11 ..... **S16**

Figure S12 ..... **S17**

Figure S13 ..... **S18**

Figure S14 ..... **S19**

Figure S15 ..... **S20**

Figure S16 ..... **S21**

Table S1 ..... **S22**

NMR Spectra **S23**

Reagents

Cyclization reagents. The bis(bromomethyl) xylene linkers (ortho, meta, and para) were purchased from Sigma Aldrich. The bis(bromomethyl) naphthalene linkers (1,4) (1,8) (2,3) and (2,6) were purchased from Sigma Aldrich. Protected amino acids were purchased from CEM.

DTT (Fisher Scientific), TCEP (Alfa Aesar), Ammonium Carbonate (Sigma Aldrich), Ammonium Bicarbonate (Sigma Aldrich), Acetonitrile (Fisher Scientific), DMF (VWR), DCM (Fisher Scientific), MeOH (Fisher Scientific), Piperidine (Sigma Aldrich), Diisopropylethylamine (VWR), TFA (Sigma Aldrich), Triisopropylsilane (Sigma Aldrich), DODT (Sigma Aldrich)

Methods

Bis-bromomethyl naphthalene synthesis for linkers d and f [1]**.** The dimethylnaphthalene compound (125 mg, 0.8 mmol, Sigma Aldrich) and N-bromosuccinimide (303 mg, 1.7 mmol, Sigma Aldrich) were added to a two-neck round bottom flask connected to a reflux condenser. Carbon tetrachloride (4 mL) and benzoyl peroxide (8.5 mg, 0.035 mmol, Sigma Aldrich) were added and the reaction was heated to reflux with stirring. The reaction was monitored by TLC using 2.5% ethyl acetate in hexanes. The Rf of the product was lower than the starting material. After completion of the reaction (1.5-2 h), the solvent was evaporated, and the mixture purified by flash chromatography leading to the title compounds as white solids.

Linker d. 2,7-bis(bromomethyl)naphthalene. Flash chromatography using 2.5% EtOAc/hexanes led to a white powder. A small sample was re-purified using 100:1 hexanes:ether to give an analytically pure sample. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 7.82 (d, 2H, *J* = 8.5 Hz), 7.81 (s, 2H), 7.52 (dd, 2H, *J* = 8.5 Hz, 1.6 Hz), 4.66 (s, 4H). ^{13}C NMR ($CDCl_3$, 150 MHz): 135.85, 132.98, 132.70, 128.59, 127.88, 127.50, 33.69. This data matched that in the literature[2].

Linker f. 1,5-bis(bromomethyl)naphthalene. Flash chromatography using 2.5% EtOAc/hexanes led to a white powder. ^{1}H NMR ($CDCl_3$, 600 MHz): δ 8.19 (d, 2H, *J* = 8.4 Hz), 7.60 (d, 2H, *J* = 6.9 Hz), 7.55 (dd, 2H, *J* = 6.9 Hz, 8.4 Hz), 4.96 (s, 4H). ^{13}C NMR ($CDCl_3$, 150 MHz): 134.13, 131.72, 128.13, 126.05, 125.41, 125.40, 31.62. This data matched that in the literature[2].

Peptide purification. The peptides were cleaved from the solid support using a cleavage cocktail containing 92.5% trifluoroacetic acid, 2.5% water, 2.5% 2,2'-(ethylenedioxy)-diethanethiol, 2.5% triisopropylsilane and incubated at room temperature for 2 h. The cleavage time was increased for peptides containing multiple arginine residues (+15 min for every additional arginine after the third arginine residue) but did not exceed 4 h total cleavage time to avoid degradation of the peptide. Small scale cleavages were performed to examine reaction efficiency for each step (1 mg resin : 100 µL cleavage cocktail). Large scale cleavages were performed for purification purposes (100 mg resin : 1 mL cleavage cocktail). Both were precipitated with at least 10 equivalents of cold diethyl ether and the peptide was pelleted using a centrifuge at max speed.

Peptides were purified by reverse-phase liquid chromatography using a two solvent system on a Shimadzu HPLC with a Restek Viva C18 5 µm, 250 x 4.6 mm column at a rate of 1 mL/min. The two solvents were water with 0.1% TFA (buffer A) and acetonitrile with 0.1% TFA (buffer B). Generally, peptides were dissolved in a small volume (20-50 µL for small scale and 200-300 µL for large scale) of

dimethylsulfoxide or MeCN/water prior to injection. The carboxyfluorescein-labeled peptides were collected using the absorbance wavelength of the fluorophore (443 nm). Carboxynaphthofluorescein-labeled peptides were collected using the absorbance wavelength of the fluorophore (606 nm). Collected fractions of purified peaks were frozen at -80 °C and lyophilized prior to use in cellular assays.

An Applied Biosystems Voyager DE-Pro instrument was used for matrix-assisted laser desorption/ionization analysis. All mass spectra were acquired using positive reflectron mode. CHCA matrix was freshly made each day at a concentration of 10 mg/mL in acetonitrile:water (1:1) with 0.1% TFA. The matrix was applied to the MALDI plate first and the sample was added on top. Volumes used for spotting matrix and sample depended on the well size for the MALDI plate but were generally 0.5 or 1 µL. Sinapinic Acid was used for peptides >2500 Da. Sinapinic Acid was freshly prepared at a concentration of 10 mg/mL in acetonitrile:water (1:1) with 0.1% TFA.

Peptide concentration determination. Peptides were dissolved in 20 µL DMSO at a concentration of 250 µM to 1 mM. 2 µL of the DMSO/peptide solution was added to 8 µL of 1M-Tris (pH 8.0) to make a 20% DMSO solution. Peptide concentration was measured by absorbance using a Nanodrop spectrophotometer. Carboxyfluorescein labeled peptides were measured at 495 nm and carboxynaphthofluorescein labeled peptides were measured at 602 nm. A total of three readings were obtained and the concentration was determined using Beer's Law. The extinction coefficient for carboxyfluorescein is ε = 75,000 M^{-1} cm^{-1} and for carboxynaphthofluorescein is ε = 44,000 M-1 cm^{-1}. The concentrations of peptides were validated by comparing peak areas of HPLC injected peptides.

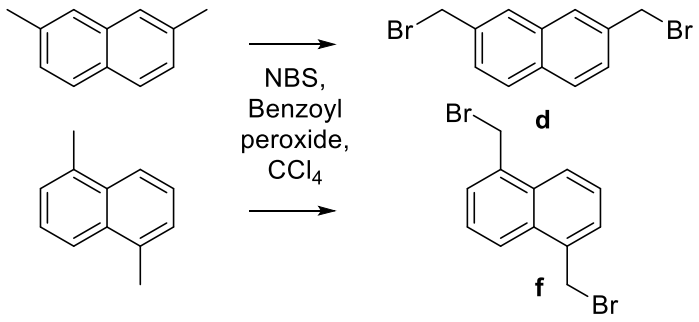
Statistics. Data were analyzed using SigmaPlot 14. Statistical tests for each figure are reported. For 1-way ANOVA, groups were stratified by peptide (Fig. 4, 6B, 8B). For 2-way ANOVA, stratification was first based on peptide (Fig. 5B, 6C), then by presence of serum. Tests were set to alpha 0.05 and specific p-values were calculated for each comparison. Critical statistical comparisons are displayed in figures, but not all comparisons could be visualized. For additional statistical information, please contact the authors.

References

1 R. G. Harvey, J. Pataki, C. Cortez, P. Di Raddo and C. Yang, A new general synthesis of polycyclic aromatic compounds based on enamine chemistry, *J. Org. Chem.*, 1991, **56**, 1210–1217.

2 T. Higashi, K. Uemura, K. Inami and M. Mochizuki, Unique behavior of 2,6-bis(bromomethyl)naphthalene as a highly active organic DNA crosslinking molecule, *Bioorganic Med. Chem.*, 2009, **17**, 3568–3571.

Scheme S1



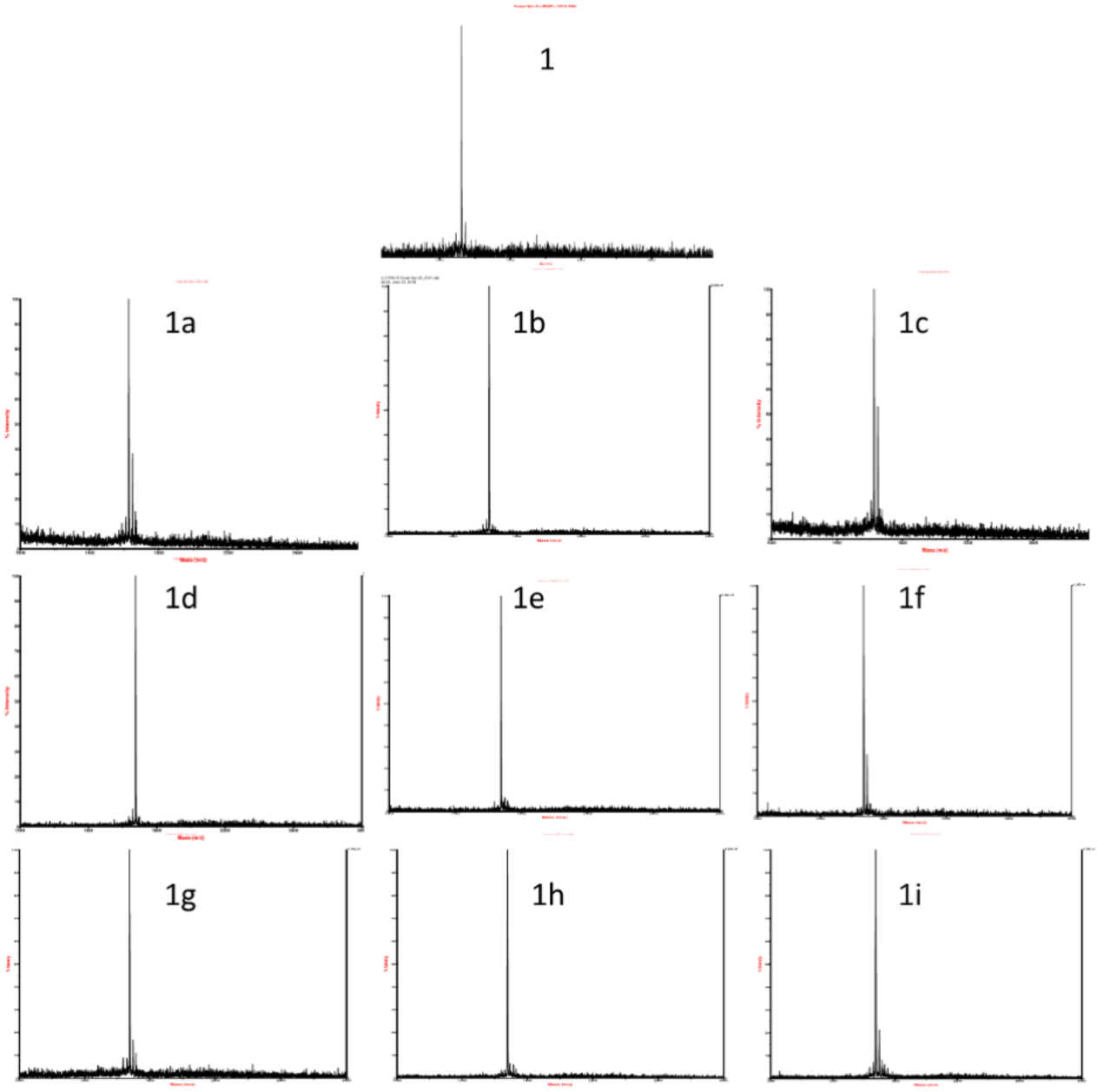


Figure S1. MALDI-MS Spectra for peptides 1 and 1a-1i

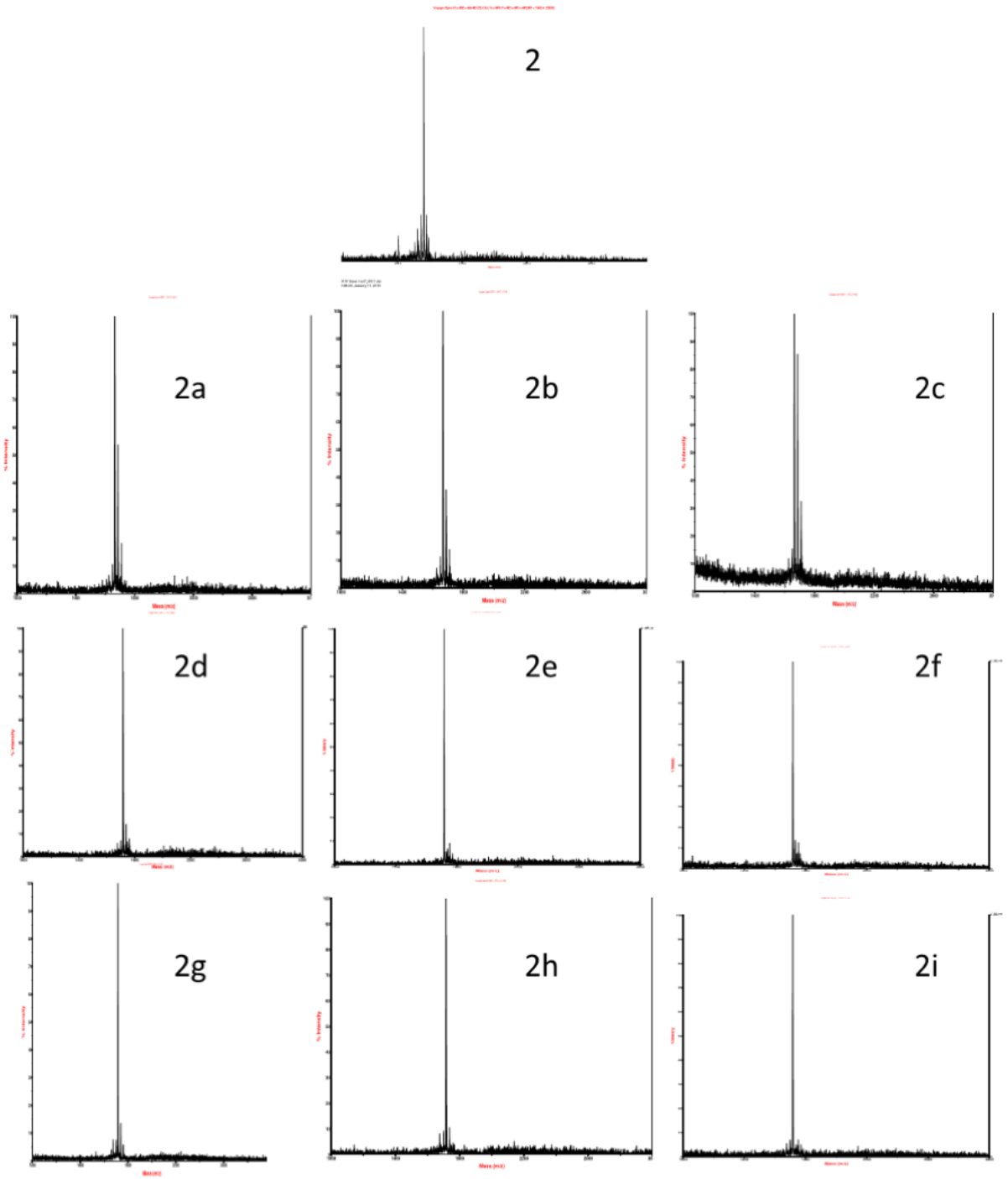


Figure S2. MALDI-MS Spectra for peptides 2 and 2a-2i

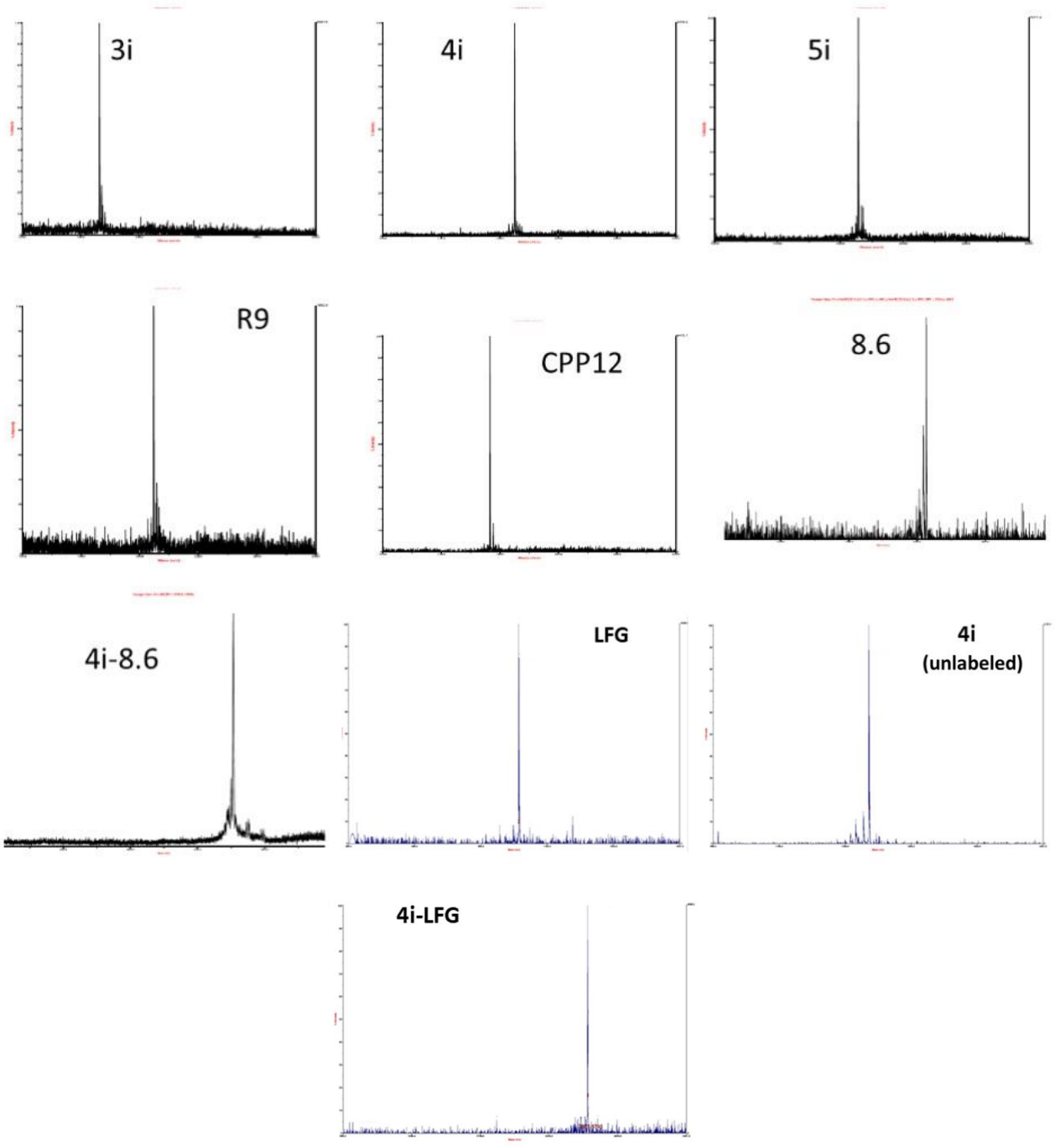


Figure S3. MALDI-MS Spectra for peptides 3i, 4i, 5i, CPP12, 8.6, 4i-8.6, LFG, 4i (unlabeled), and 4i-LFG

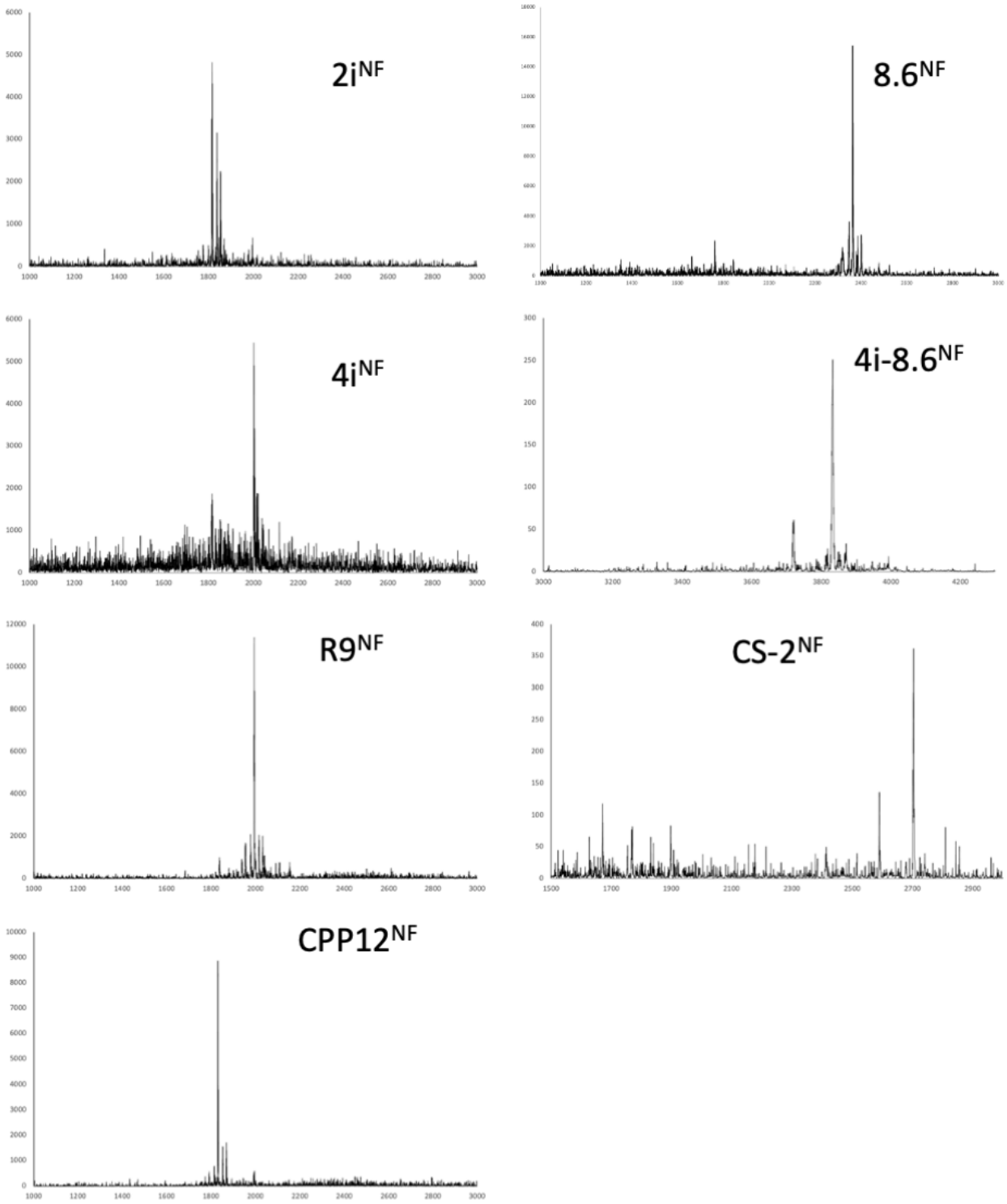


Figure S4. MALDI-MS Spectra for NF-labeled peptides. Masses higher than the peptide indicate sodium (+23) and potassium (+39) adducts.

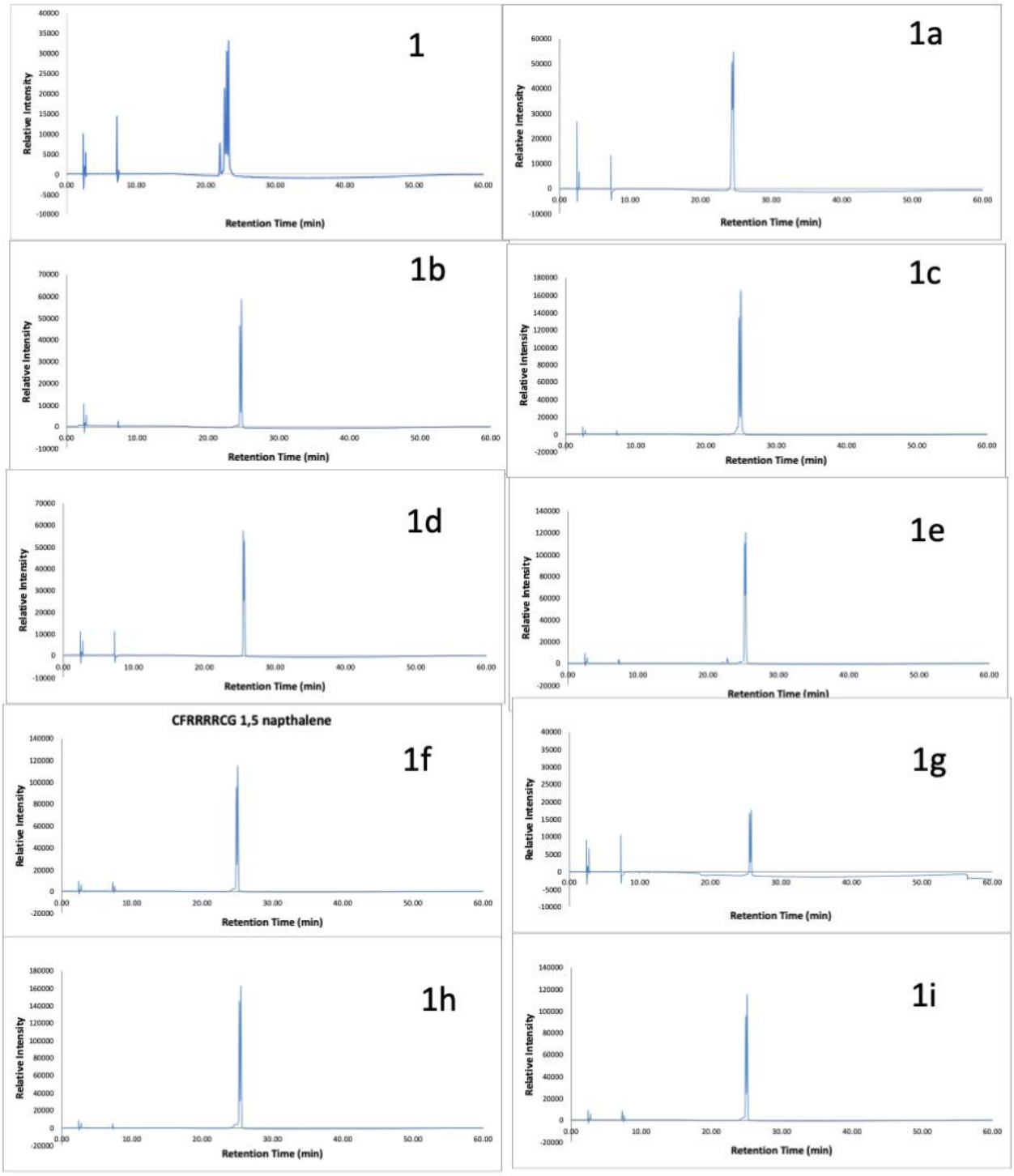


Figure S5. HPLC traces for peptides 1 and 1a-1i.

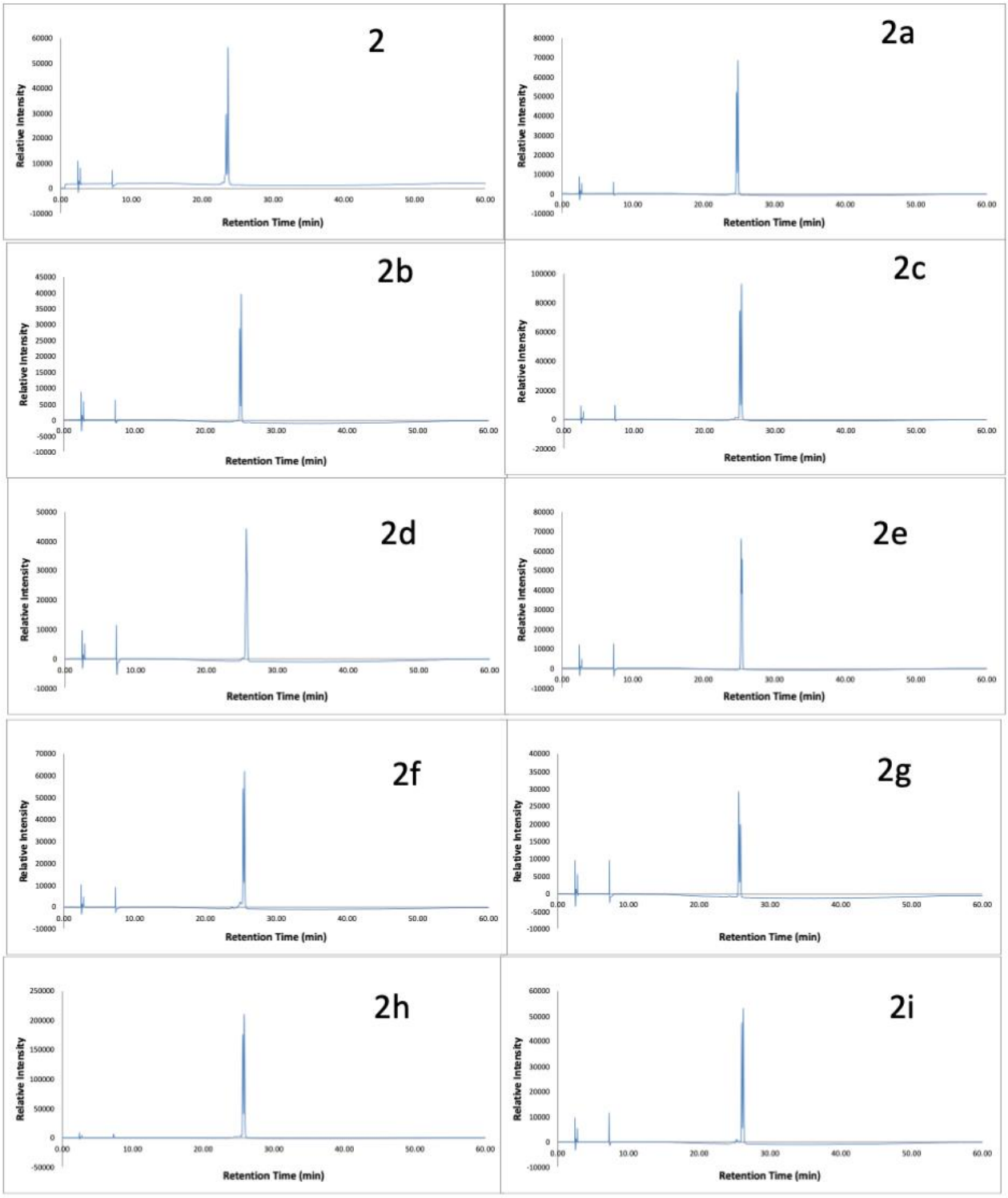


Figure S6. HPLC traces for peptides 2 and 2a-i.

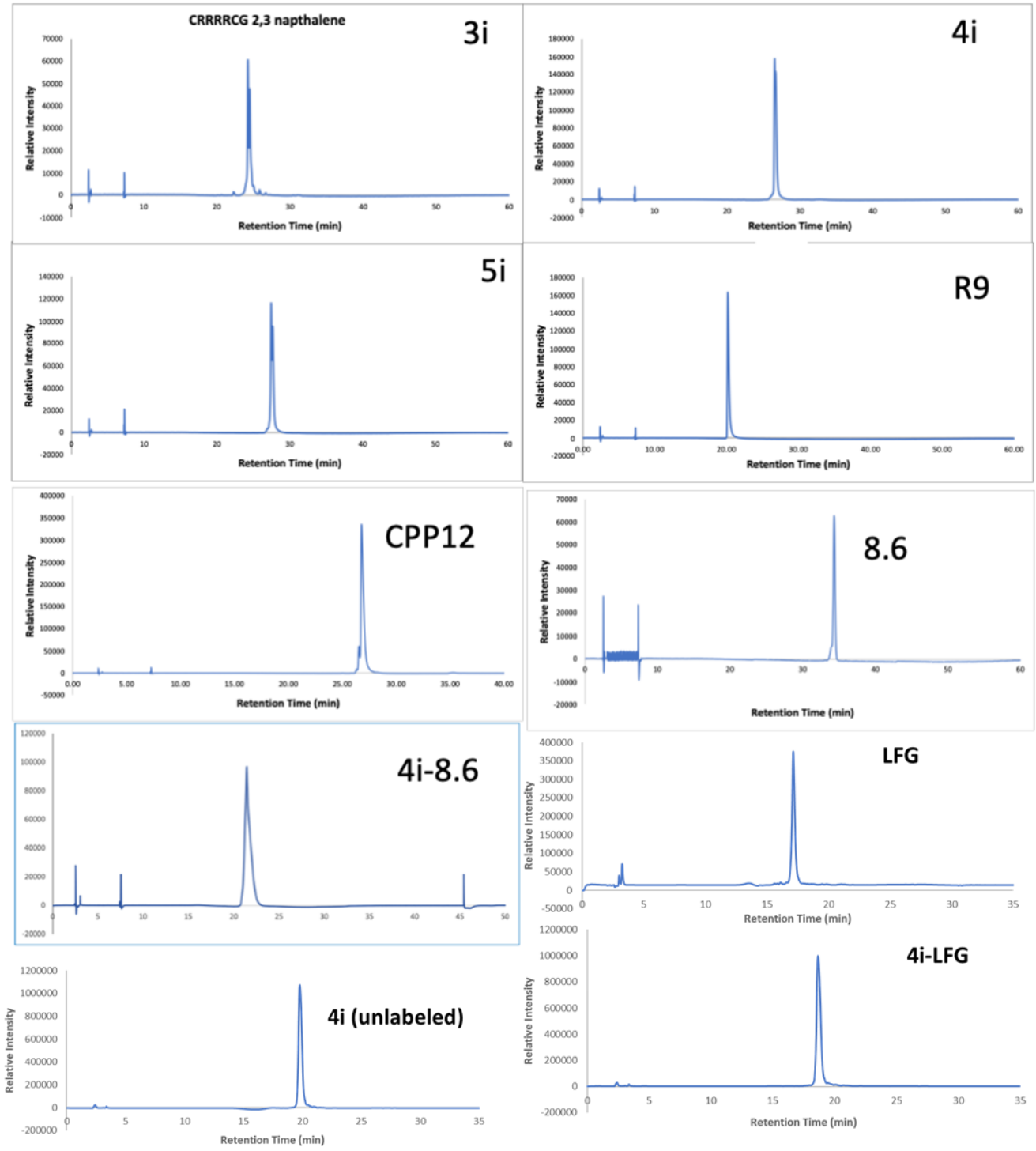


Figure S7. HPLC traces for peptides 3i, 4i, 5i, CPP12, 8.6, 4i-8.6, LFG, 4i (unlabeled), and 4i-LFG

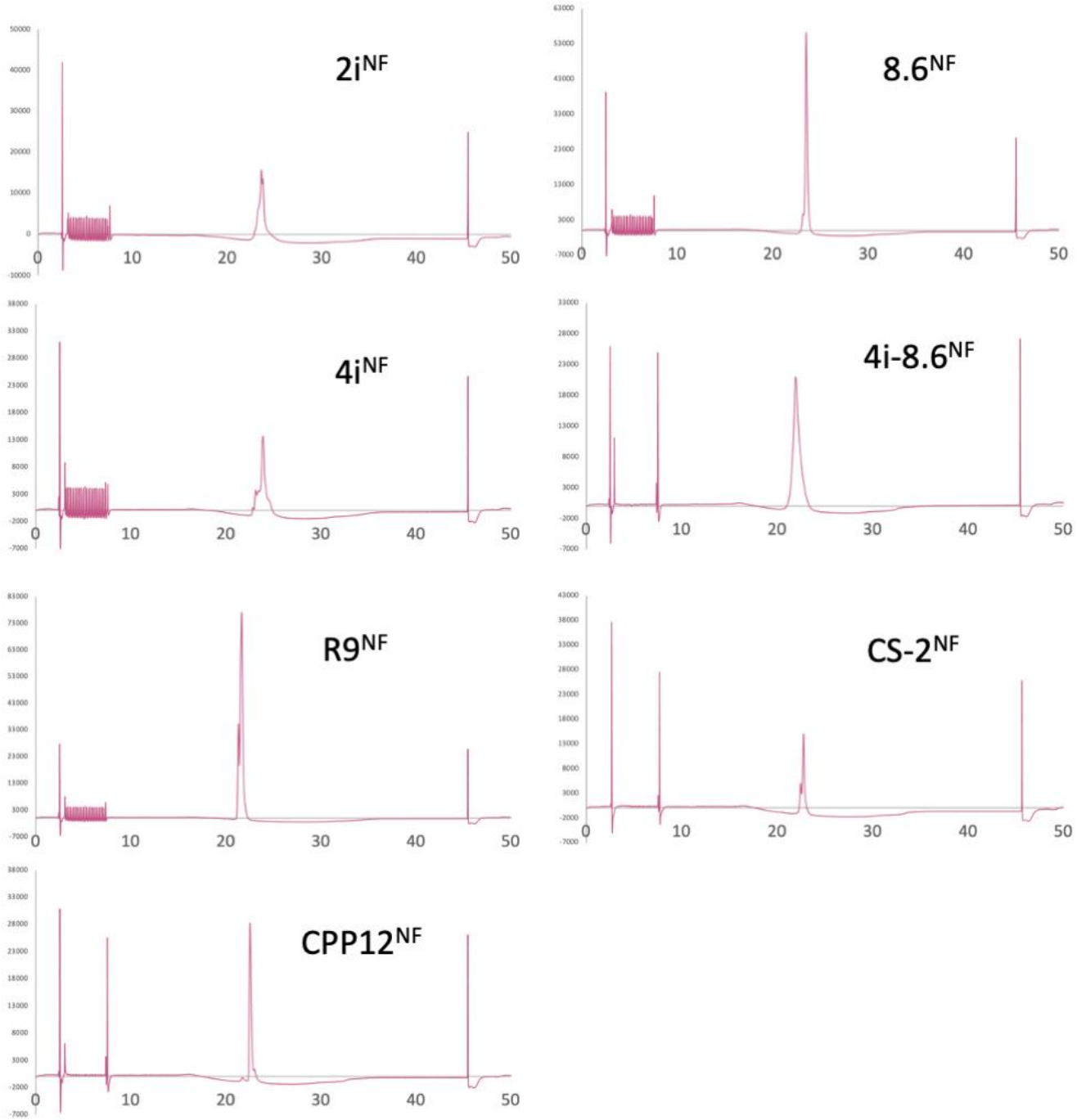


Figure S8. HPLC traces for NF-labeled peptides

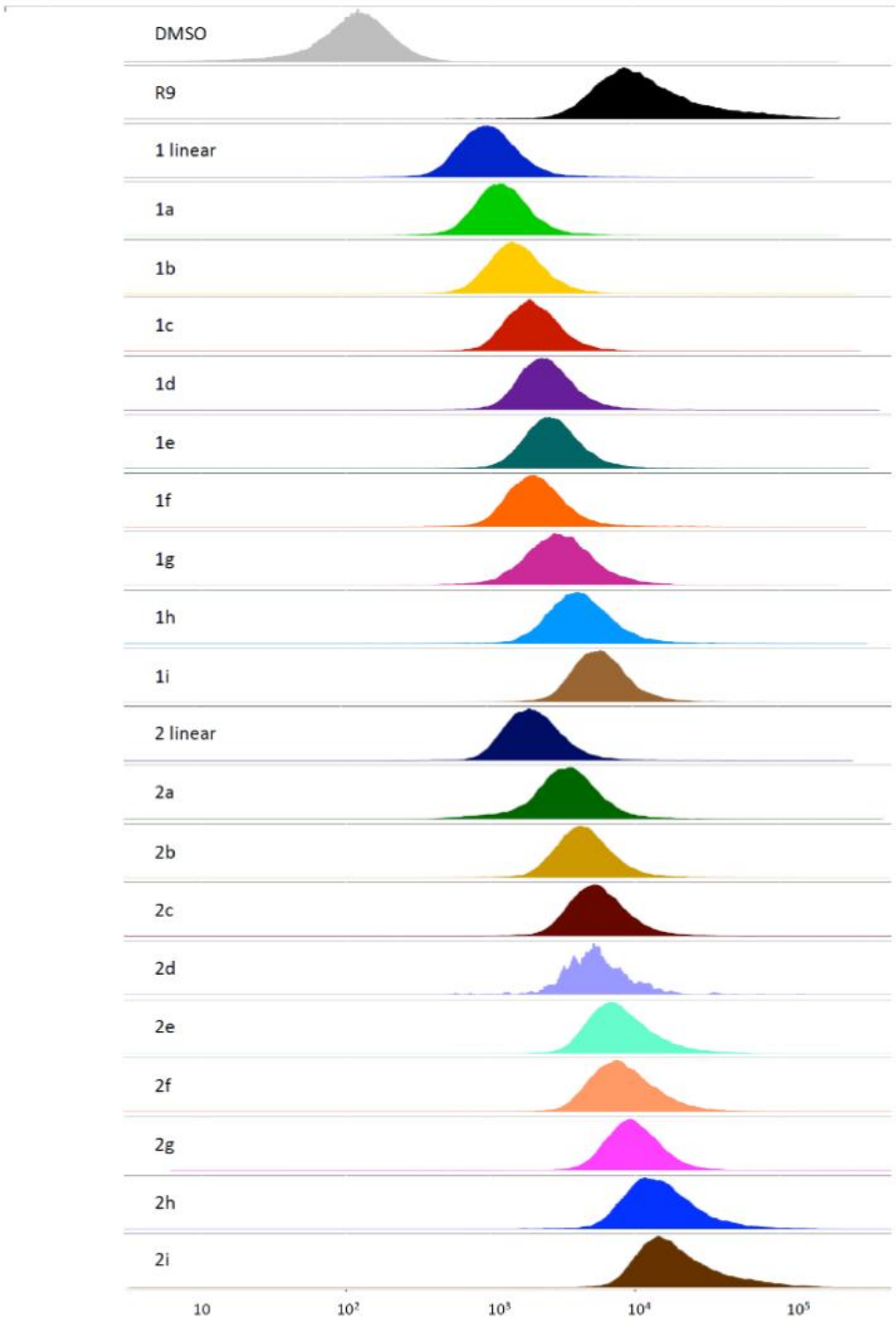


Figure S9. Flow cytometry histogram plots MDA-MB-231 cells incubated with various peptides. Each peptide was incubated at a concentration of 5 μM for 2 h. Each peptide was tested in triplicate and the plots show an average of all three experiments. Fluorescence values were adjusted based on internal peptide fluorescence quenching. The y-axes were adjusted within each experiment to enhance visualization.

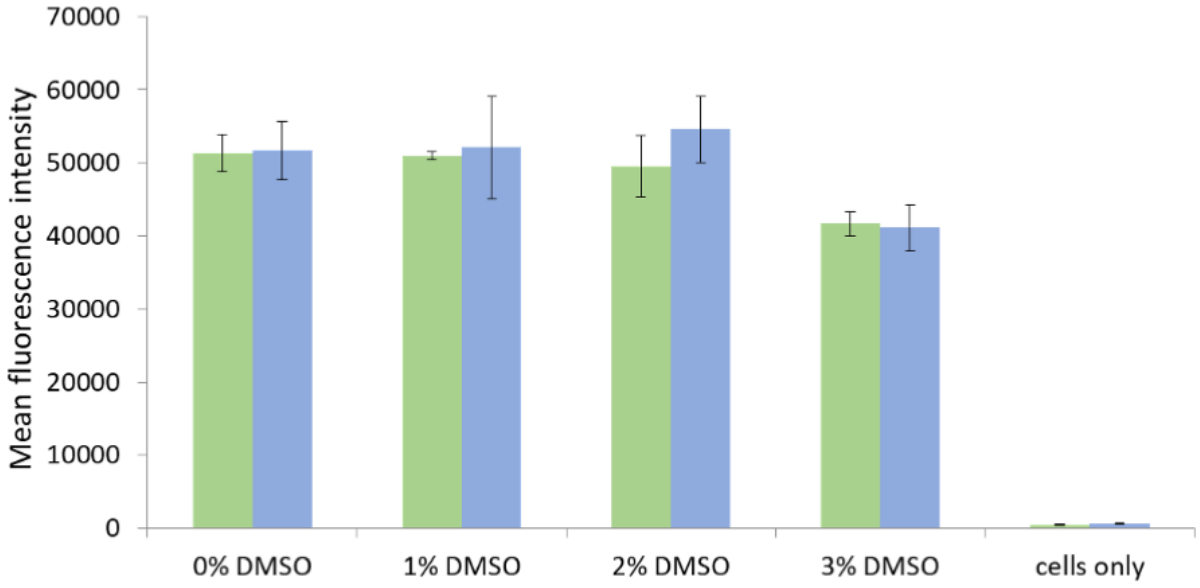


Figure S10. Effect of DMSO and extracellular Trypan Blue addition on mean fluorescence intensity. Green bars = standard treatment. Blue bars = trypan blue quench. MD-MB-231 cells were incubated with 5 µM of peptide **R9** under the standard conditions in the presence of increasing concentrations of DMSO. Cells were treated with 0.25% trypsin for 5 min at 37 °C to lift them from the plate and digest any extracellular bound peptide. For the trypan blue experiments, cells were treated with a 0.05% trypan blue solution prior to injection into the BD FACSCanto II Analyzer.

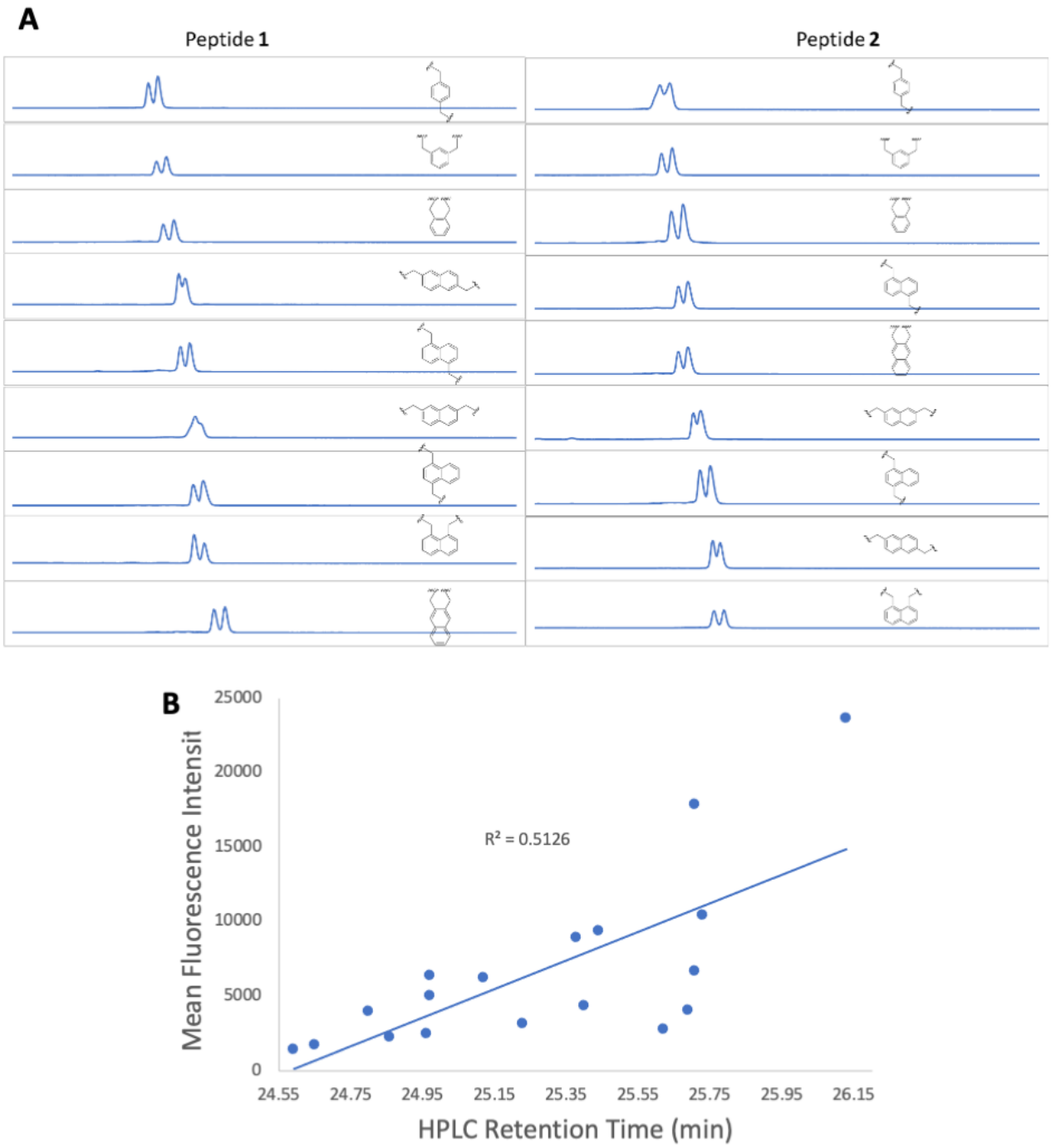


Figure S11. Comparison of hydrophobicity and cellular uptake. (A) HPLC traces of each cyclized version of peptide **1** and **2** in the time region of 22-32 minutes. (B) Plot of retention time vs. mean fluorescence for peptides displayed in A, showing the correlation between hydrophobicity and cellular uptake. Retention time was taken as the average between the two HPLC peaks.

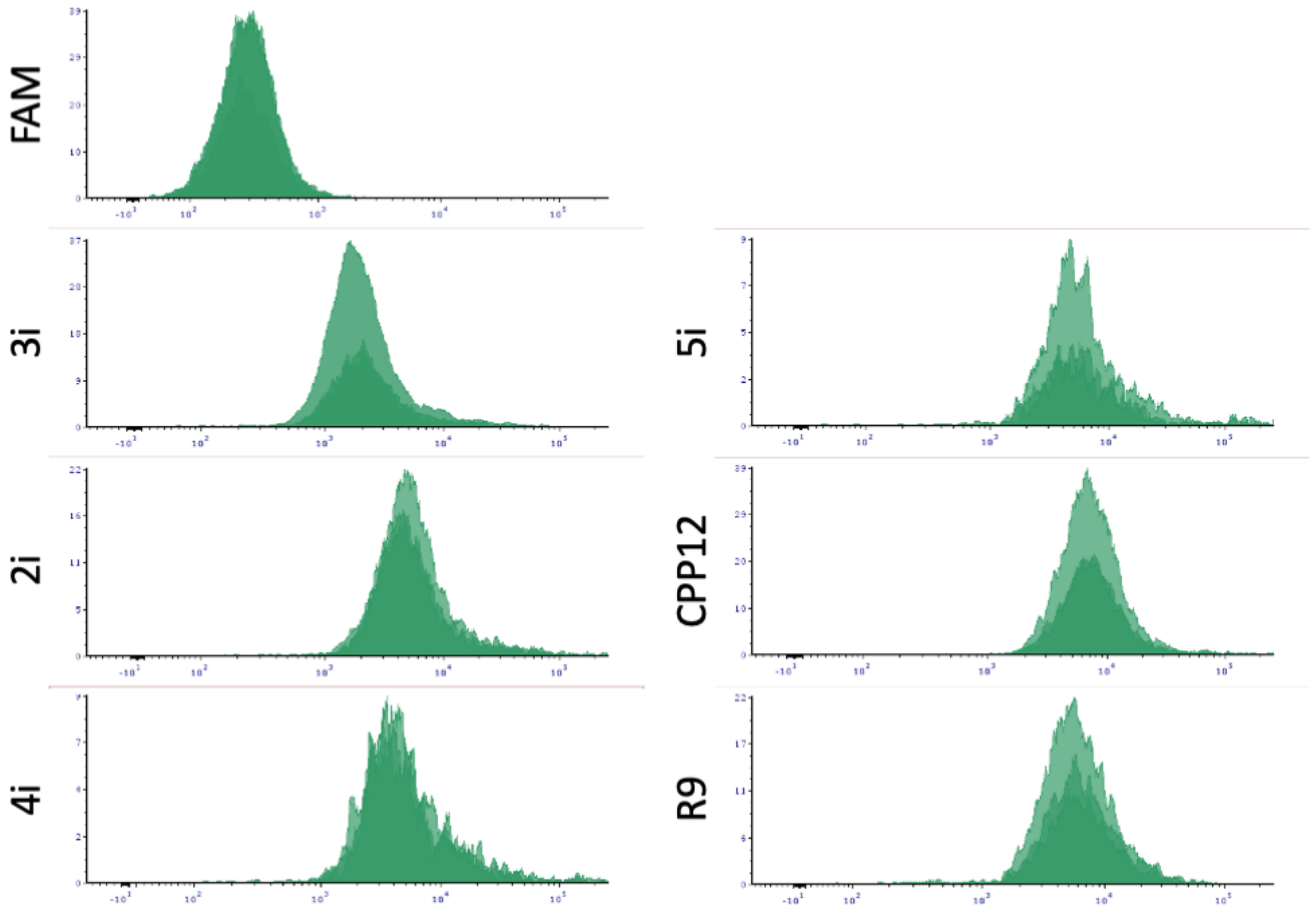


Figure S12. Flow cytometry histogram plots MDA-MB-231 cells incubated with peptides 2i-5i against positive controls. Each peptide was incubated at a concentration of 5 µM for 2 h. The y-axes were adjusted within each experiment to enhance visualization. The y-axes are labeled with the peptide used. All histograms are representative of the 488-530/30A laser-bandpass filter. Histograms contain overlays from each replicate of the peptide.

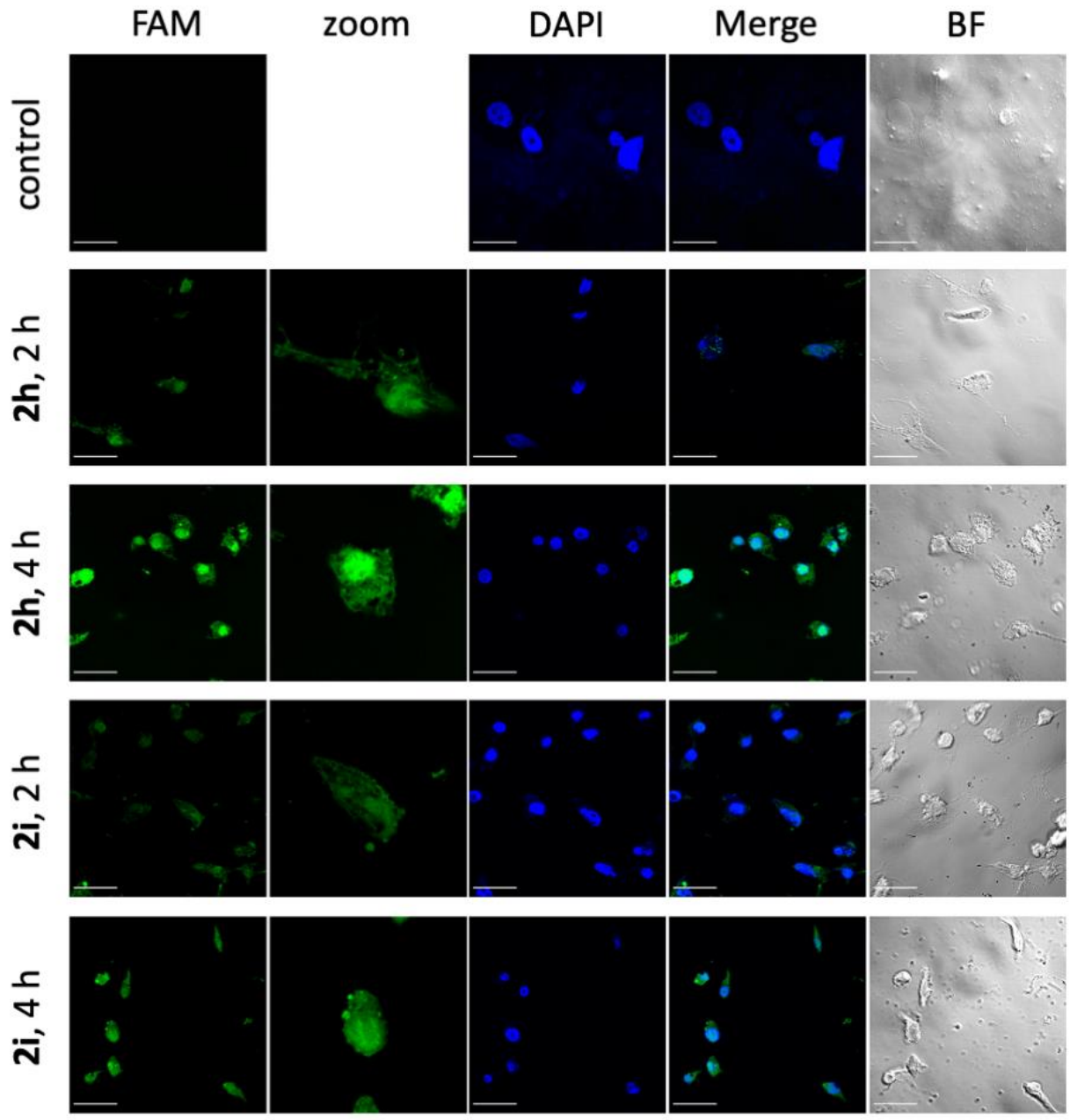


Figure S13. Analysis of cellular uptake by various peptides by confocal microscopy. Each peptide was incubated for with MDA-MB-231 cells at a concentration of 5 μM for the time indicated. Scale bar = 30 μm.

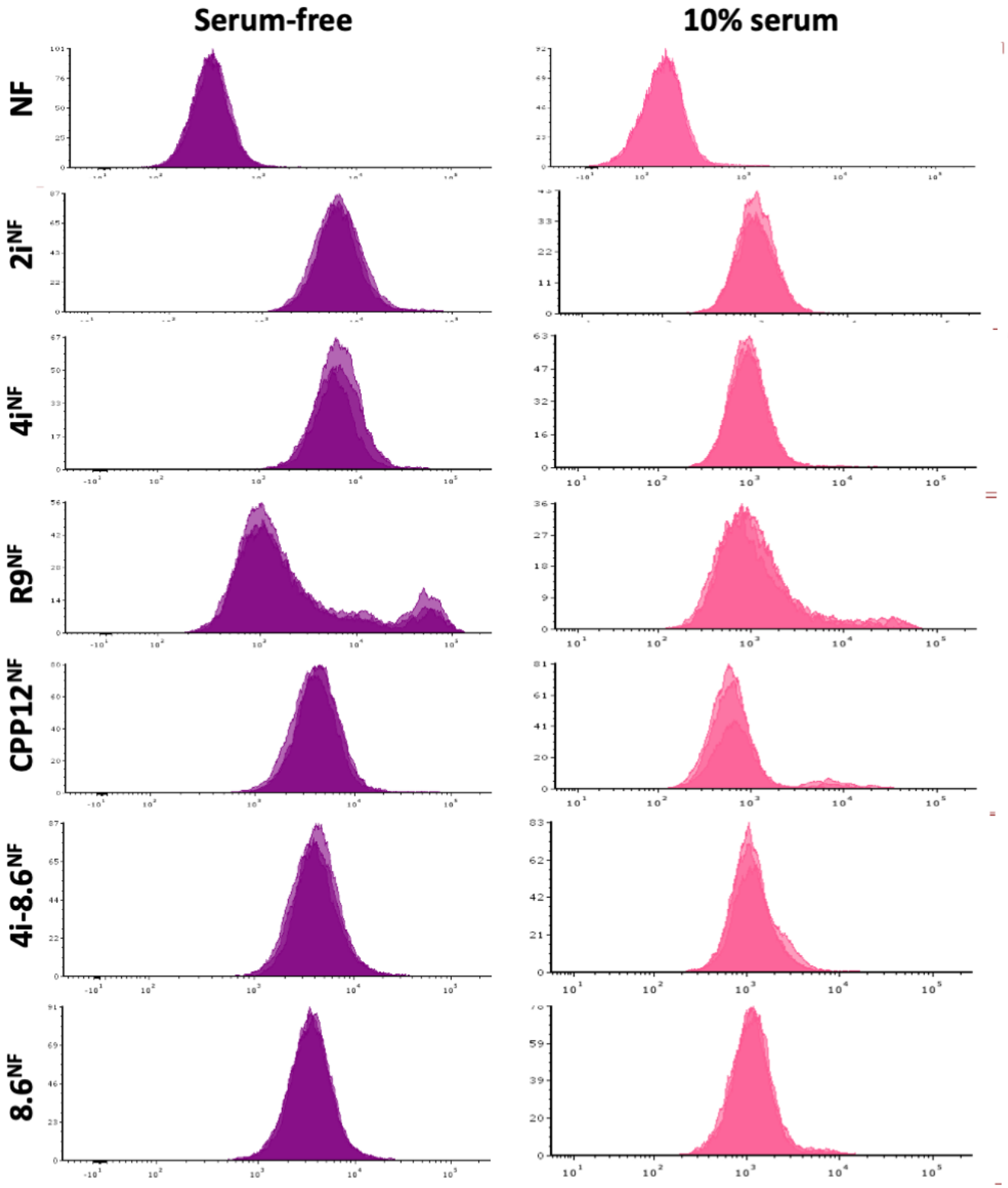


Figure S14. Flow cytometry histogram plots MDA-MB-231 cells incubated with naphthofluorescein-labeled peptides. Each peptide was incubated at a concentration of 5 µM for 2 h. Histograms on the left correspond with serum-free analysis. Those on the right were incubated under 10% FBS. The y-axes were adjusted within each experiment to enhance visualization. The y-axes are labeled with the peptide used. NF signifies naphthofluorescein alone. All histograms are representative of the 633-660/20A laser-bandpass filter. Histograms contain overlays from each replicate of the peptide.

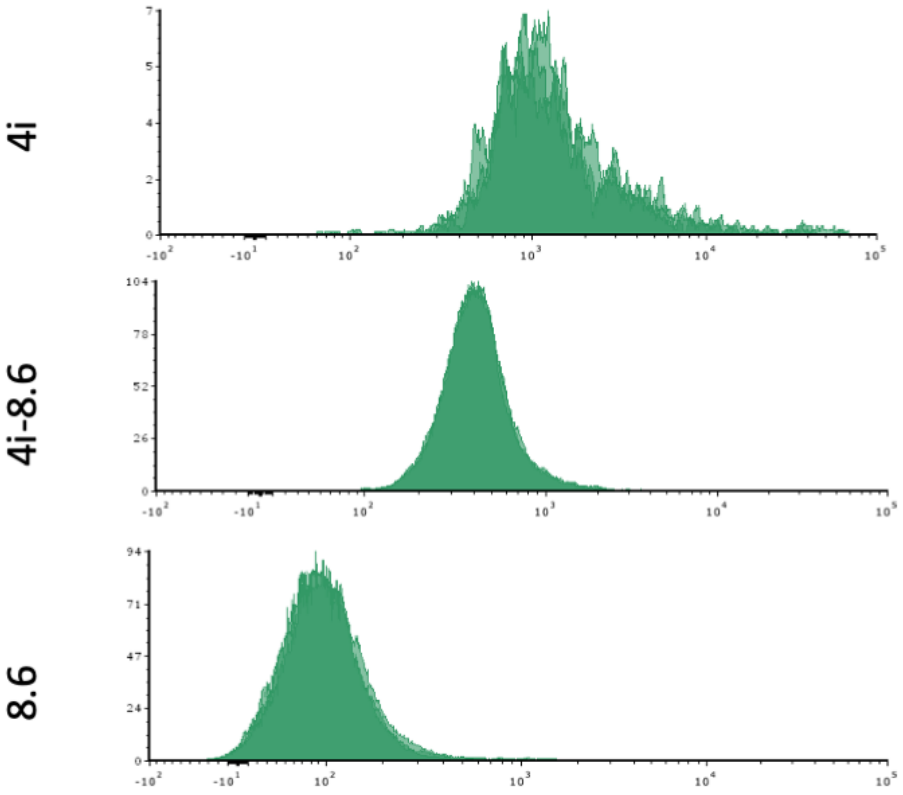


Figure S15. Flow cytometry histogram plots MDA-MB-231 cells incubated with peptides 4i, 4i-8.6, and 8.6. Each peptide was incubated at a concentration of 5 μM for 2 h. The y-axes were adjusted within each experiment to enhance visualization. The y-axes are labeled with the peptide used. All histograms are representative of the 488-530/30A laser-bandpass filter. Histograms contain overlays from each replicate of the peptide.

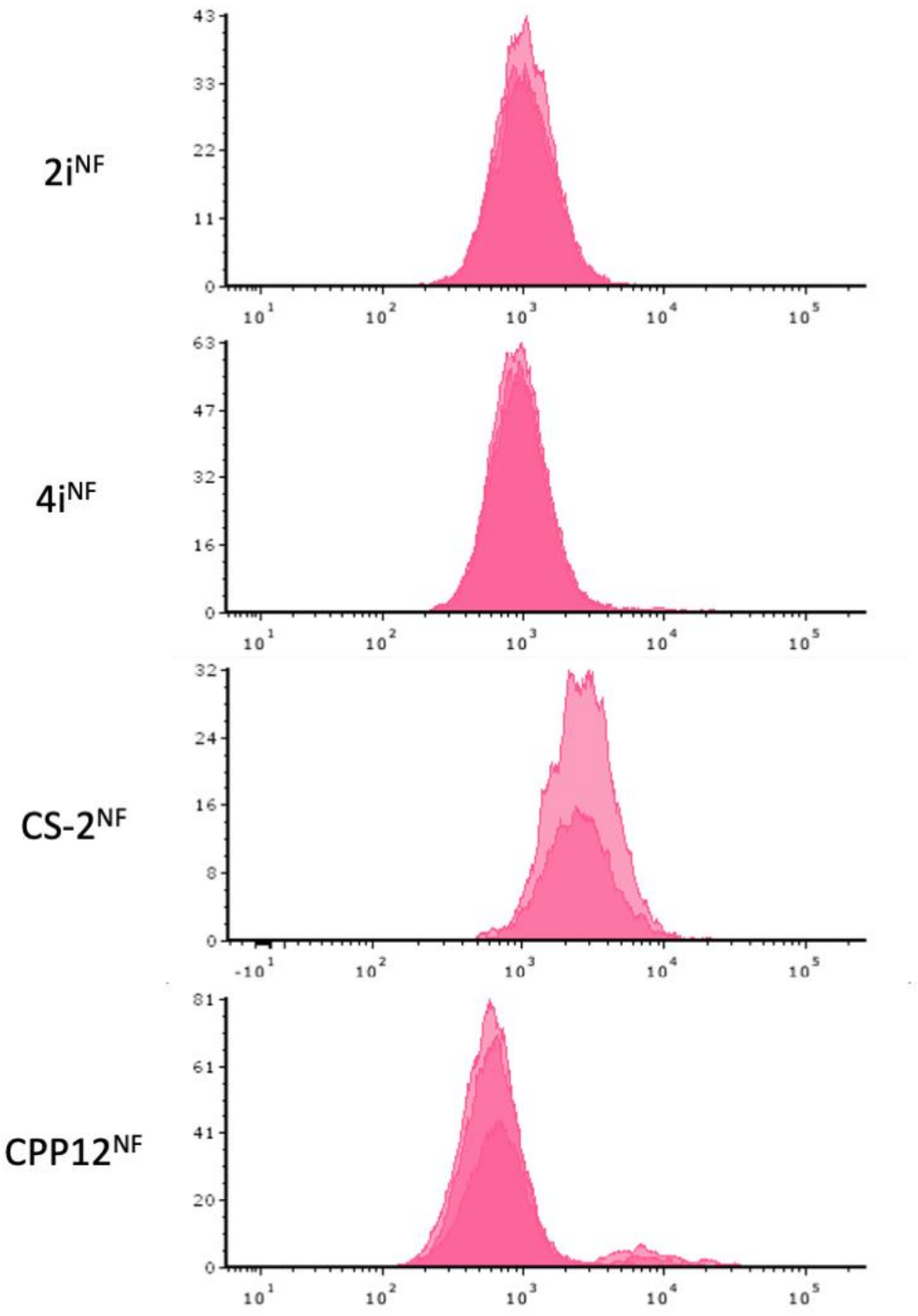


Figure S16. Flow cytometry histogram plots MDA-MB-231 cells incubated with naphthofluorescein-labeled peptides. Each peptide was incubated at a concentration of 5 µM for 2 h. The y-axes were adjusted within each experiment to enhance visualization. The y-axes are labeled with the peptide used. **2i** and **4i** are repeated from Figure S12 for comparison. All histograms are representative of the 633-660/20A laser-bandpass filter. Histograms contain overlays from each replicate of the peptide.

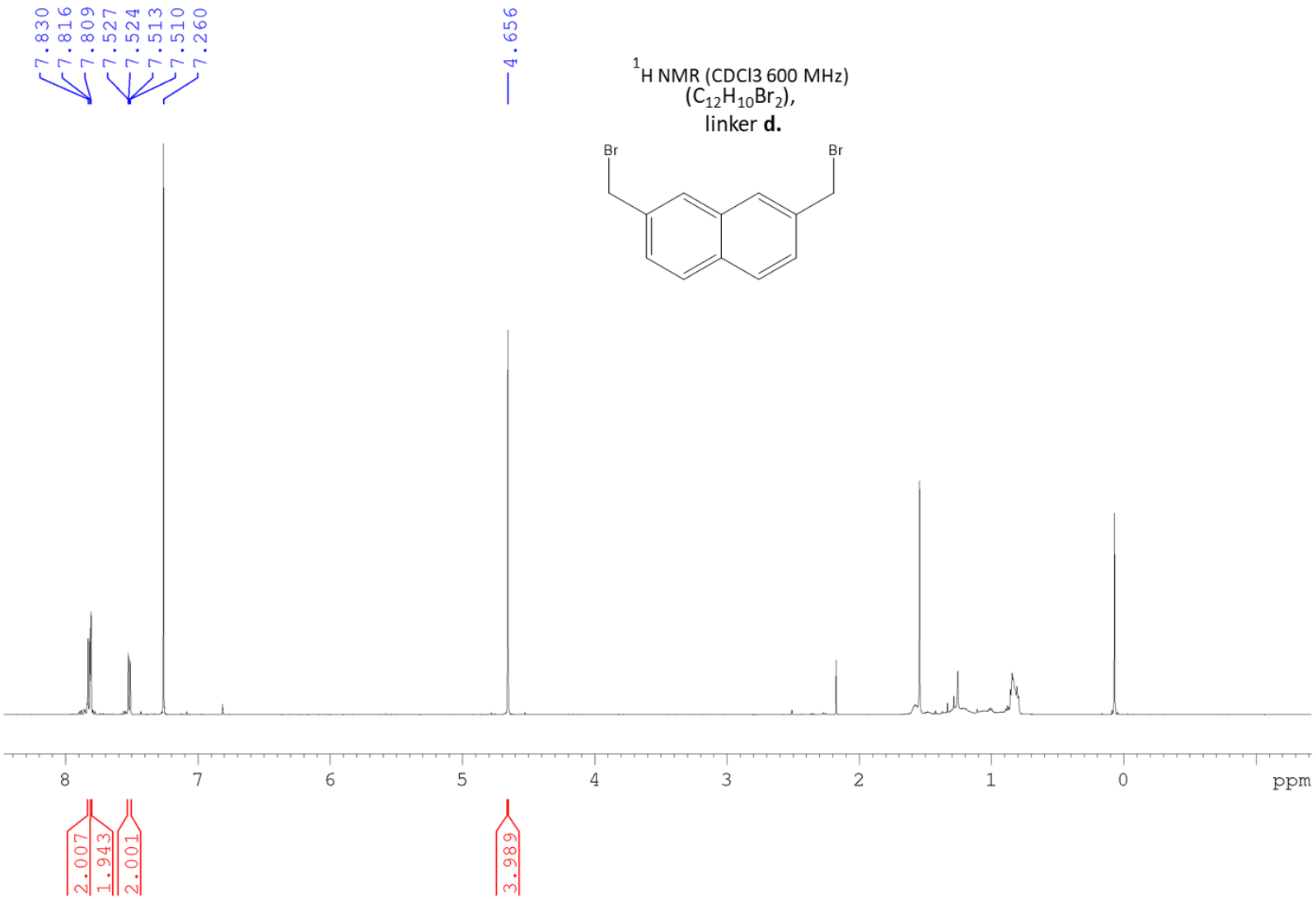
Table S1

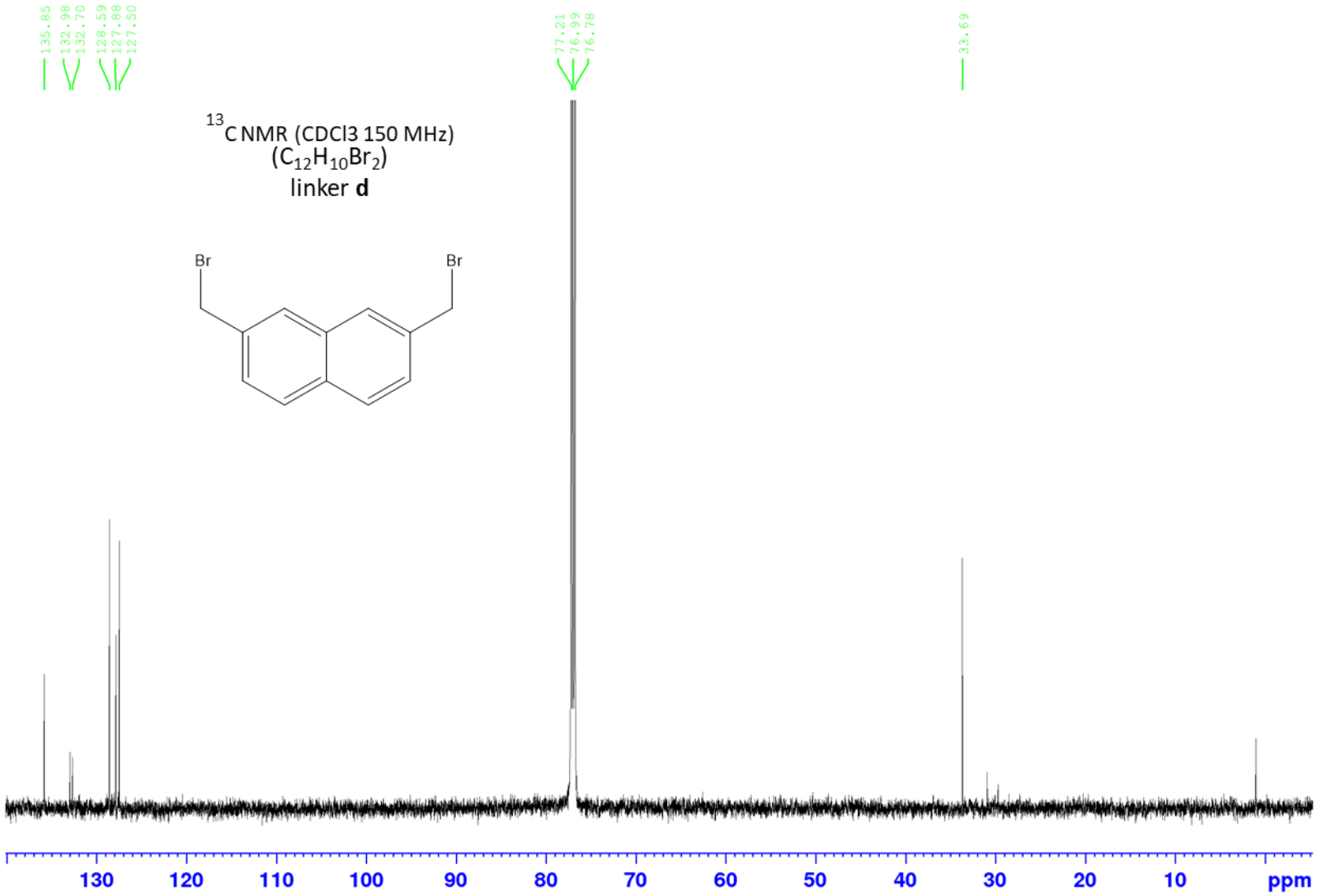
Peptide	Expected $[M+H]^+$	Observed $[M+H]^+$
1a	1625.73	1625.87
1b	1625.73	1625.29
1c	1625.73	1625.89
1d	1675.74	1675.51
1e	1675.74	1675.45
1f	1675.74	1675.45
1g	1675.74	1675.28
1h	1675.74	1675.43
1i	1675.74	1675.29
1	1523.68	1522.30
3i	1528.67	1527.96
4i	1900.83	1900.12
5i	1911.84	1911.17
R9	1894.08	1893.54
CPP12	1730.83	1730.48
LFG	1013.63	1013.75
4i-LFG	2425.29	2425.86

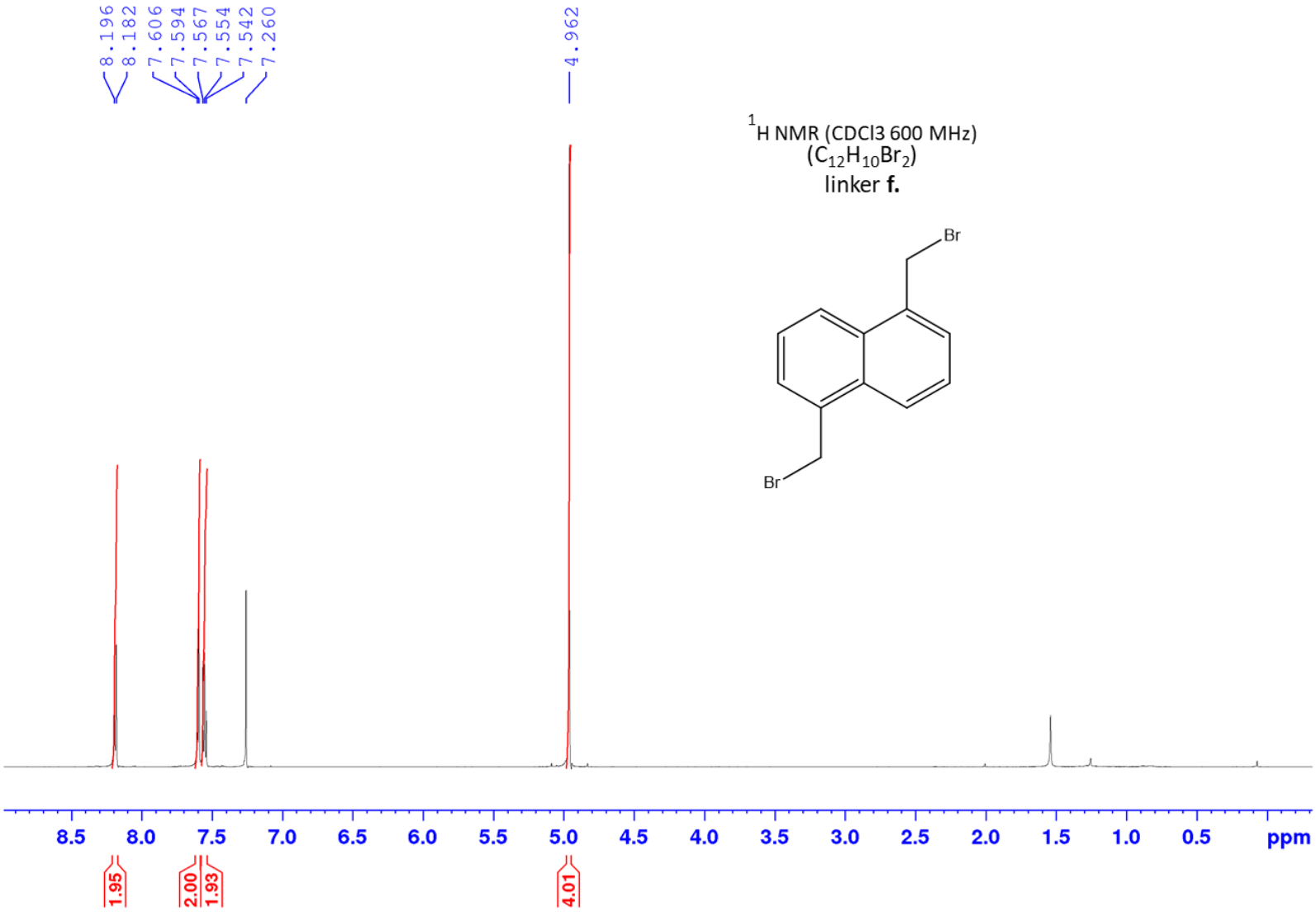
Peptide	Expected $[M+H]^+$	Observed $[M+H]^+$
2a	1664.74	1664.79
2b	1664.74	1664.77
2c	1664.74	1664.72
2d	1714.75	1714.47
2e	1714.75	1714.31
2f	1714.75	1714.49
2g	1714.75	1714.35
2h	1714.75	1714.30
2i	1714.75	1714.07
2	1562.69	1561.40
4i-8.6 [a]	3764.39	3765.16
8.6	2252.94	2253.48
2h [NF]	1814.78	1815.73
2i [NF]	1814.78	1814.65
R9 [NF a]	1884.14	1884.04
4i (unlabeled)	1471.70	1471.53

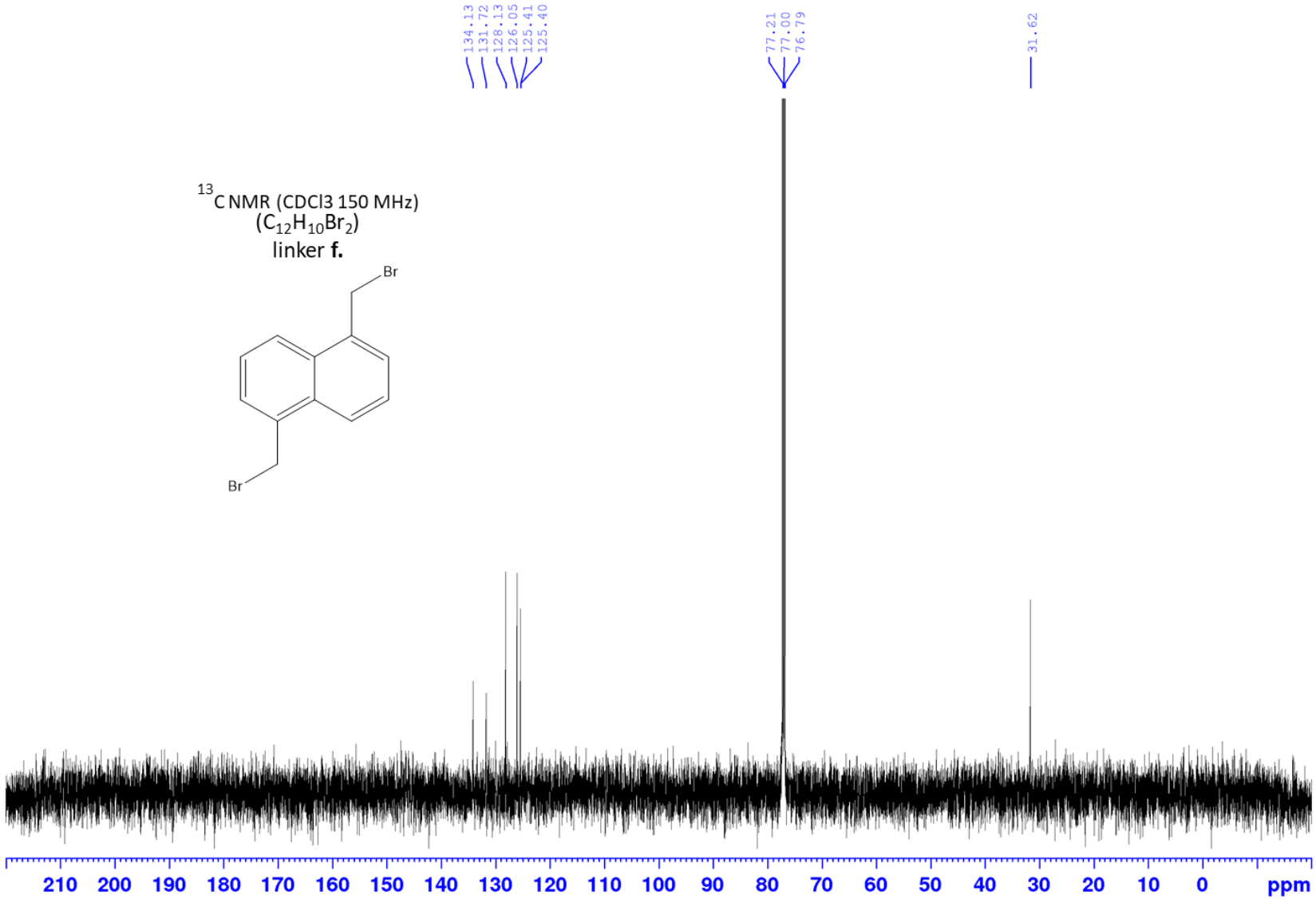
[a]Calculated $[M+H]^+$ is an average mass as the MALDI-MS spectra did not show isotopic resolution

NMR Spectra


7.830
7.816
7.809
7.527
7.524
7.513
7.510
7.260
4.656
^{1}H NMR (CDCl3 600 MHz)
($C_{12}H_{10}Br_2$),
linker **d.**
Br
Br
8
7
6
5
4
3
2
1
0
ppm
2.007
1.943
2.001
3.989


135.85
132.98
132.70
128.59
127.88
127.50
77.21
76.99
76.78
33.69
13C NMR (CDCl3 150 MHz)
(C12H10Br2)
linker d
Br
Br
130
120
110
100
90
80
70
60
50
40
30
20
10
ppm


8.196
8.182
7.606
7.594
7.567
7.554
7.542
7.260
4.962
^{1}H NMR (CDCl3 600 MHz)
($C_{12}H_{10}Br_2$)
linker **f.**
Br
Br
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm
1.95
2.00
1.93
4.01


134.13
131.72
128.13
126.05
125.41
125.40
77.21
77.00
76.79
31.62
13C NMR (CDCl3 150 MHz)
(C12H10Br2)
linker f.
Br
Br
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 ppm