Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Tuning the properties of a cyclic RGD-containing tetrapeptide through backbone fluorination

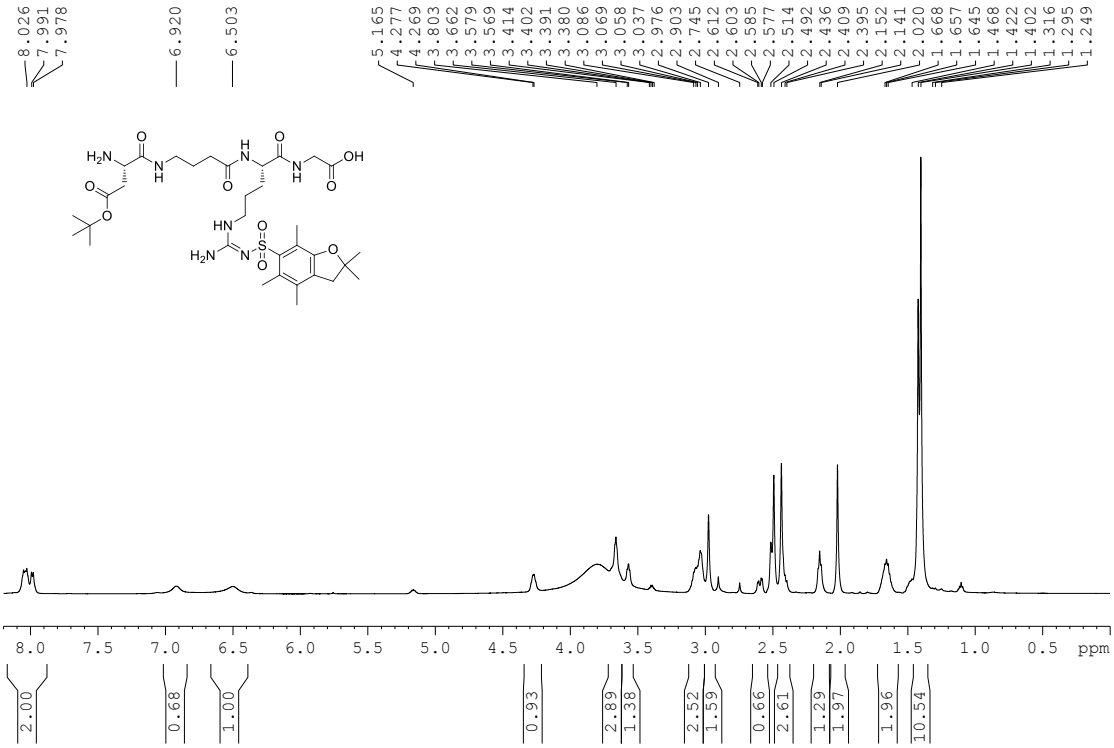
*Catherine Au, Christina Gonzalez, Yun Cheuk Leung, Flora Mansour, Johny Trinh, Zhiyong Wang, Xiang-Guo Hu, Renate Griffith, Eddy Pasquier and Luke Hunter**

SUPPORTING INFORMATION

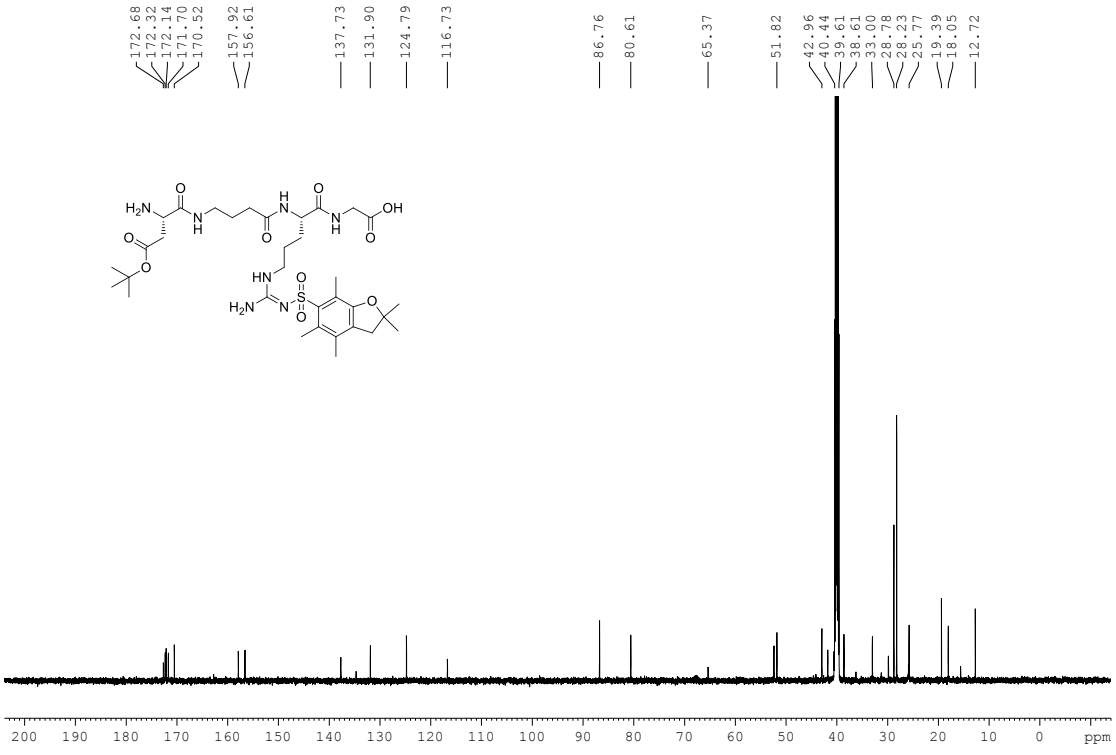
Contents:

1. NMR spectra of novel compounds

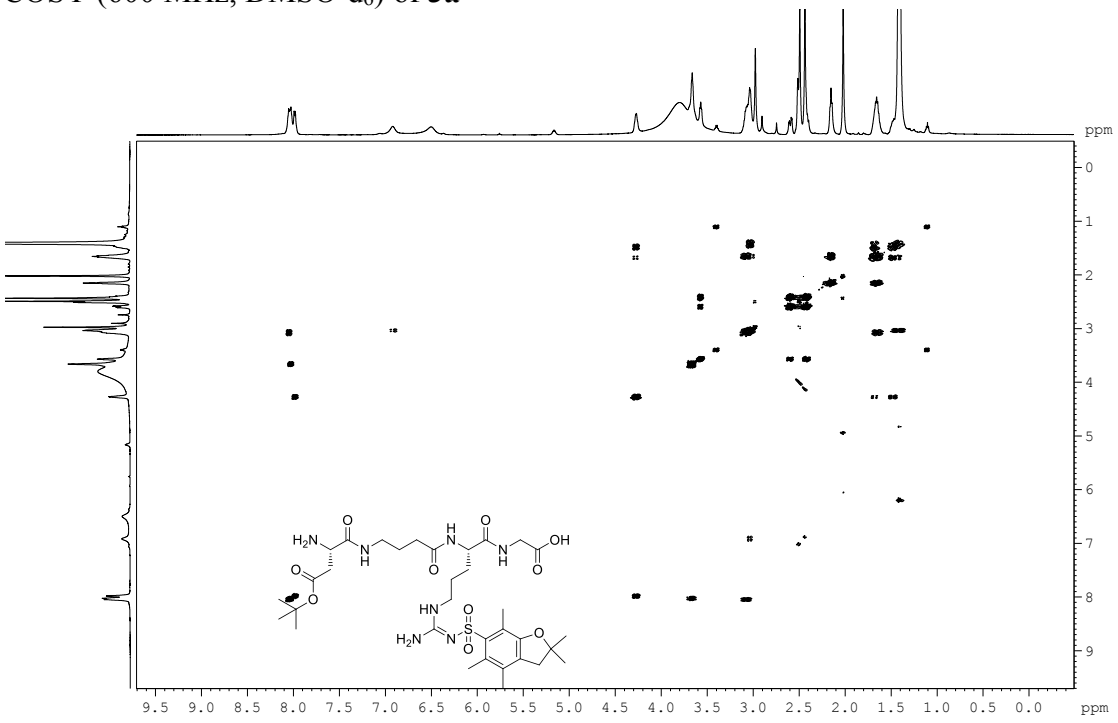
^{1}H NMR (600 MHz, DMSO-d$_6$) of **3a**



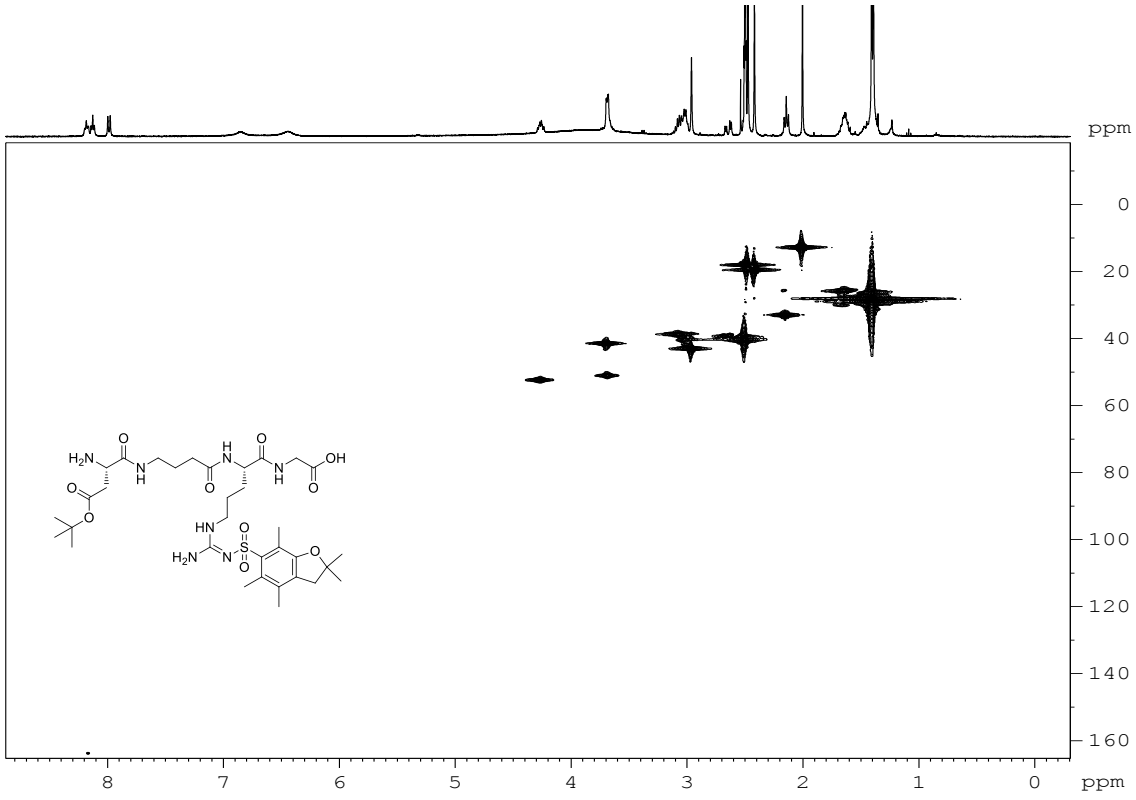
^{13}C{^{1}H} NMR (125 MHz, DMSO-d$_6$) of **3a**



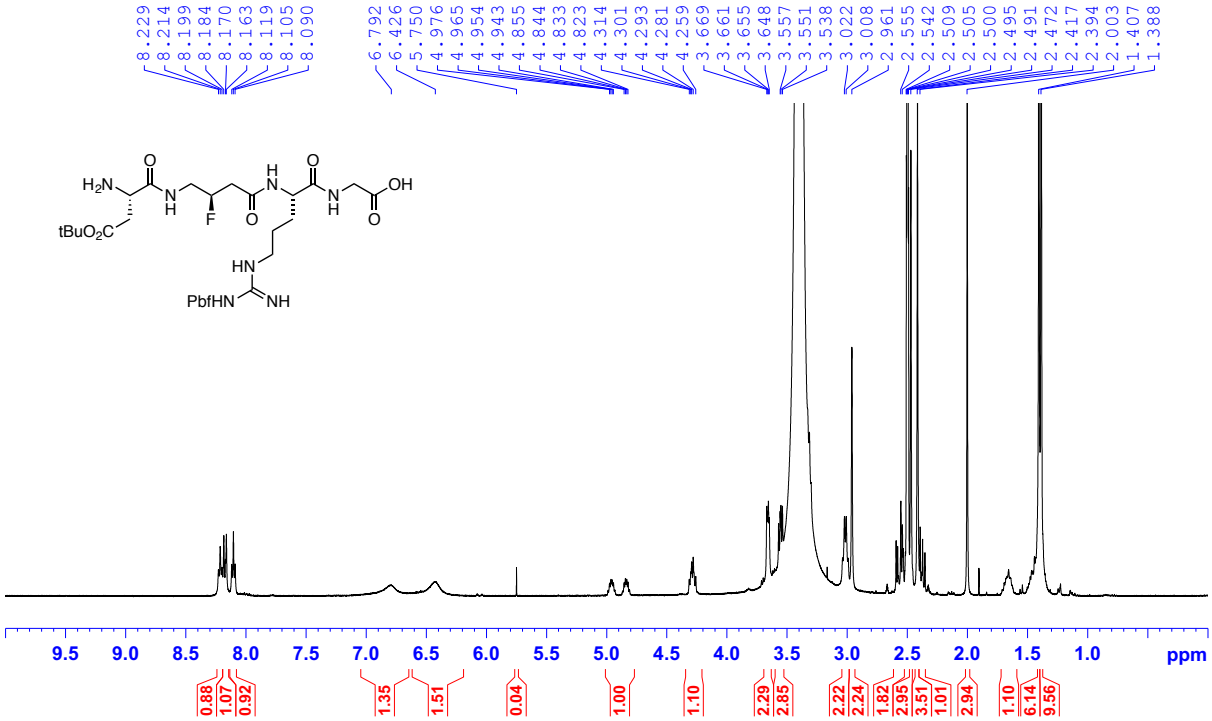
^{1}H-^{1}H COSY (600 MHz, DMSO-d_6) of **3a**



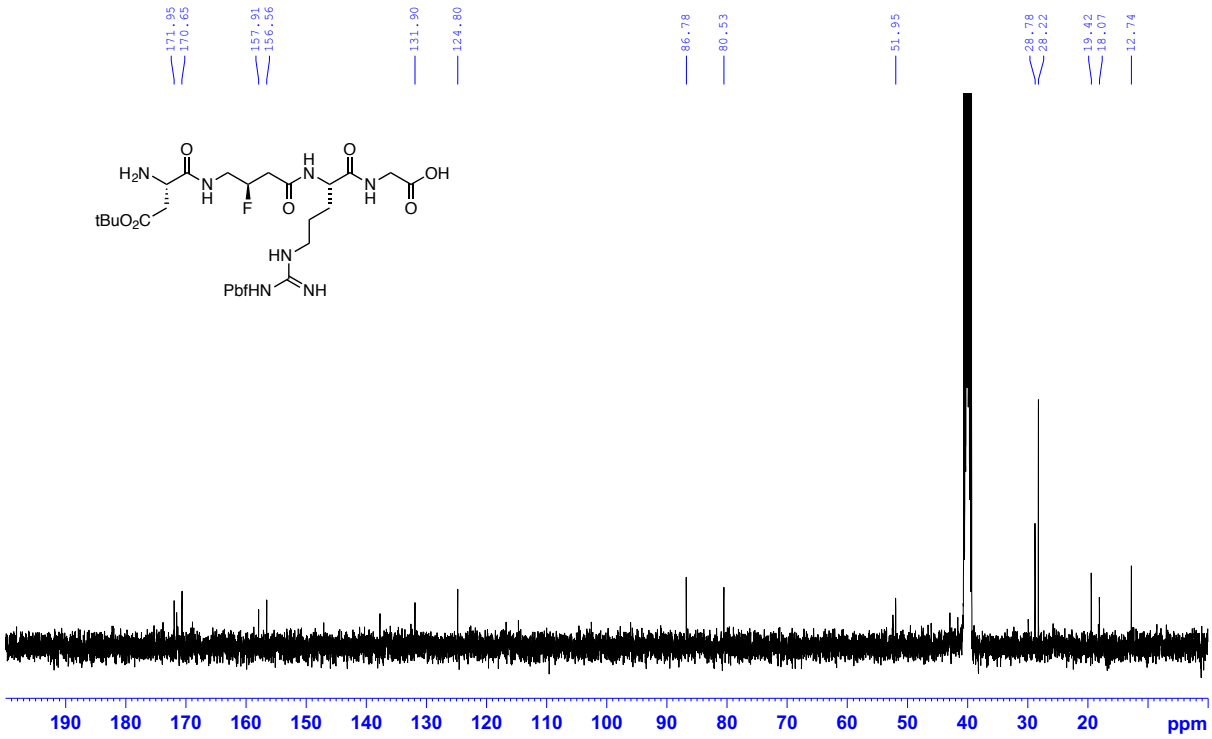
^{1}H-^{13}C HSQC (300 MHz, DMSO-d_6) of **3a**



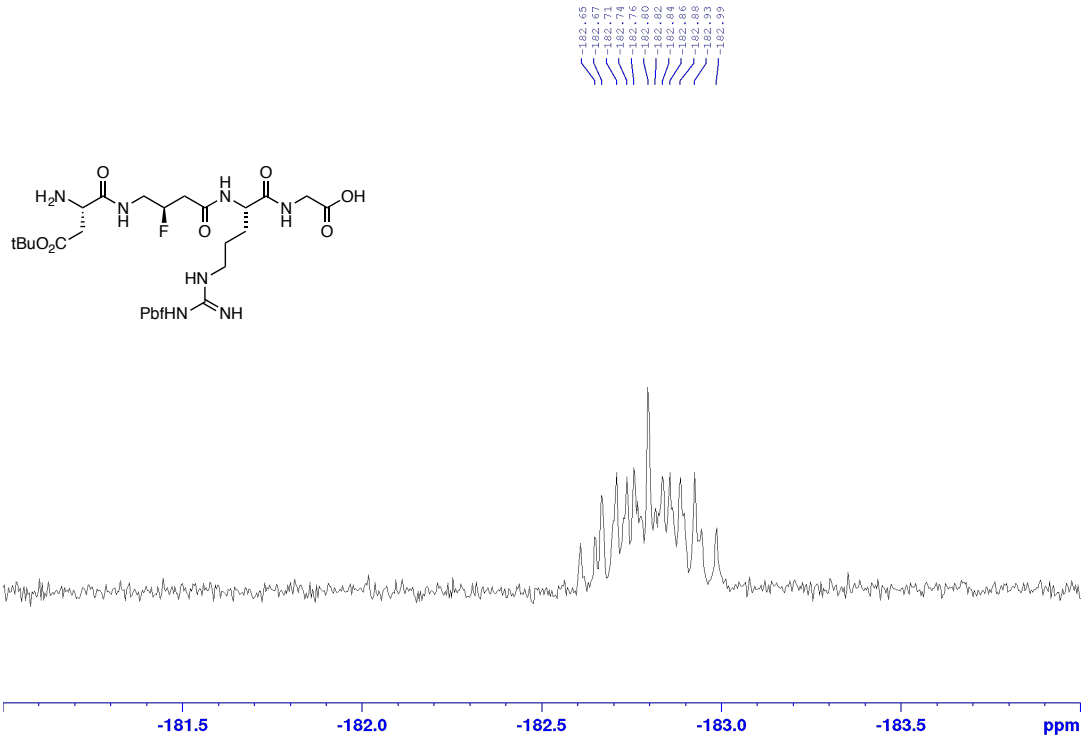
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3b**



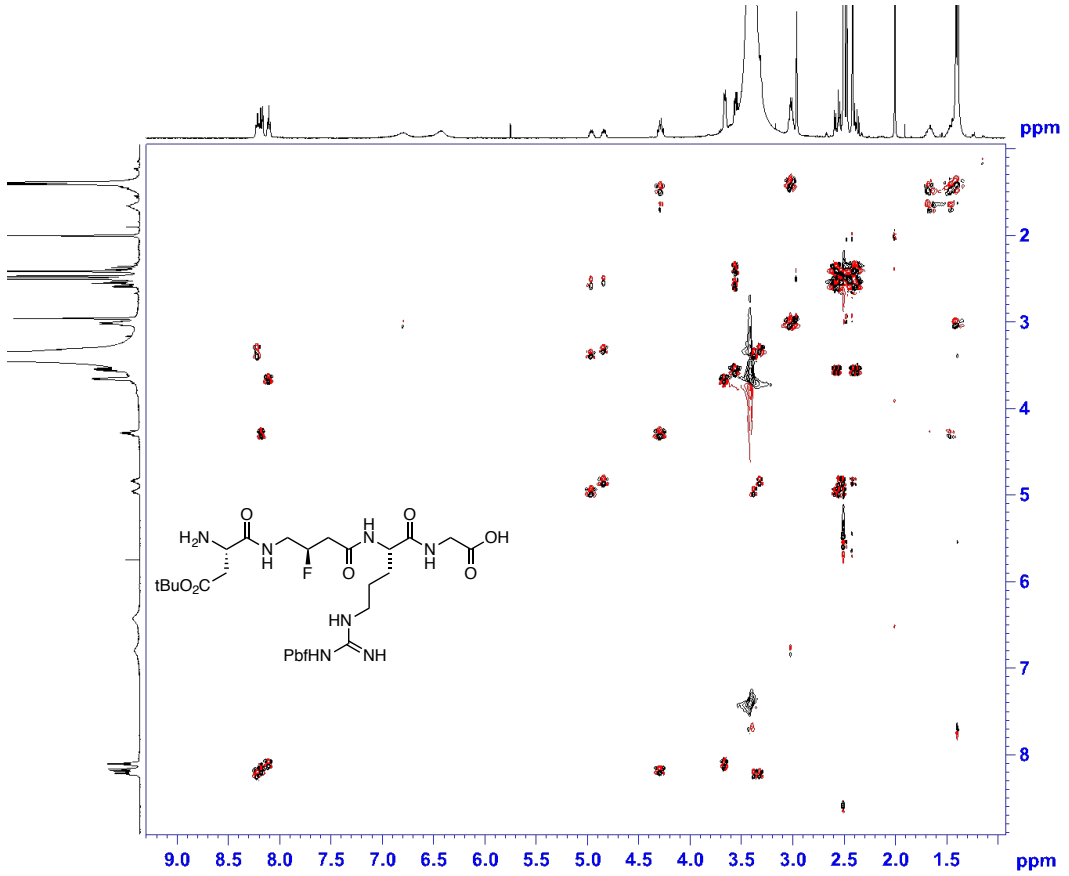
^{13}C{^{1}H} NMR (100MHz, DMSO-d$_6$) of **3b**



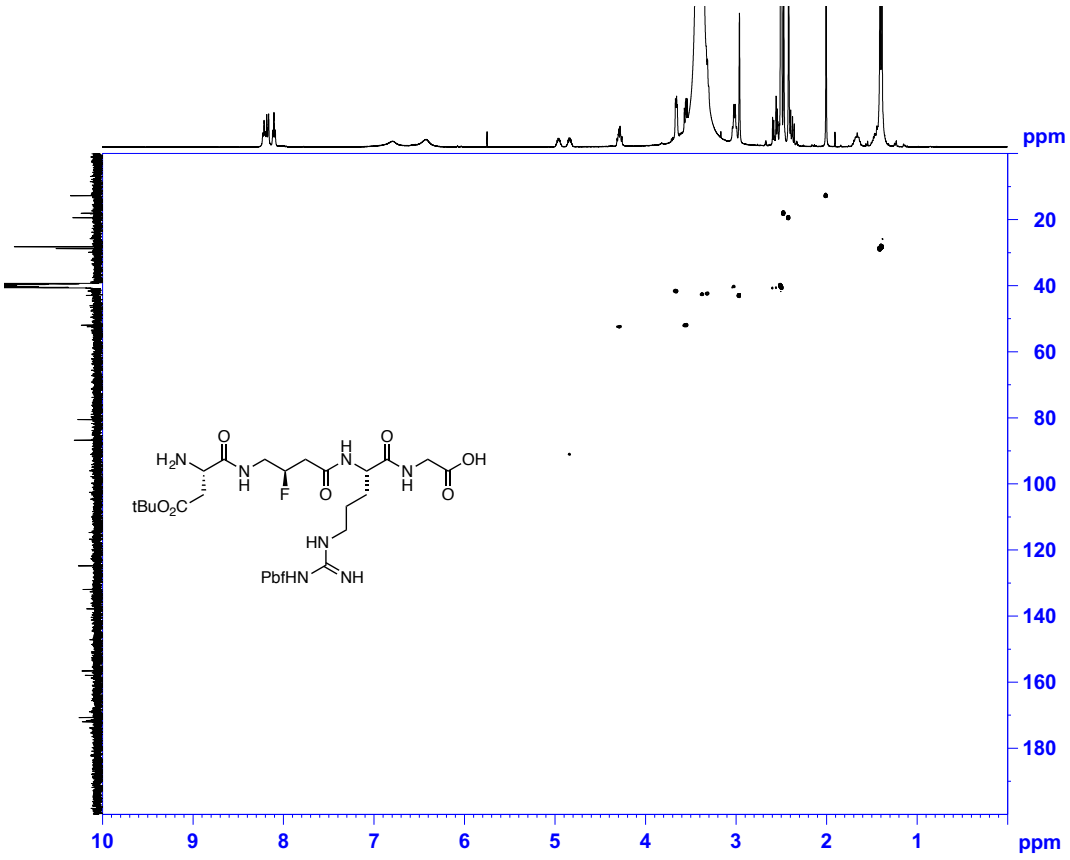
^{19}F NMR (376 MHz, DMSO-d$_6$) of **3b**



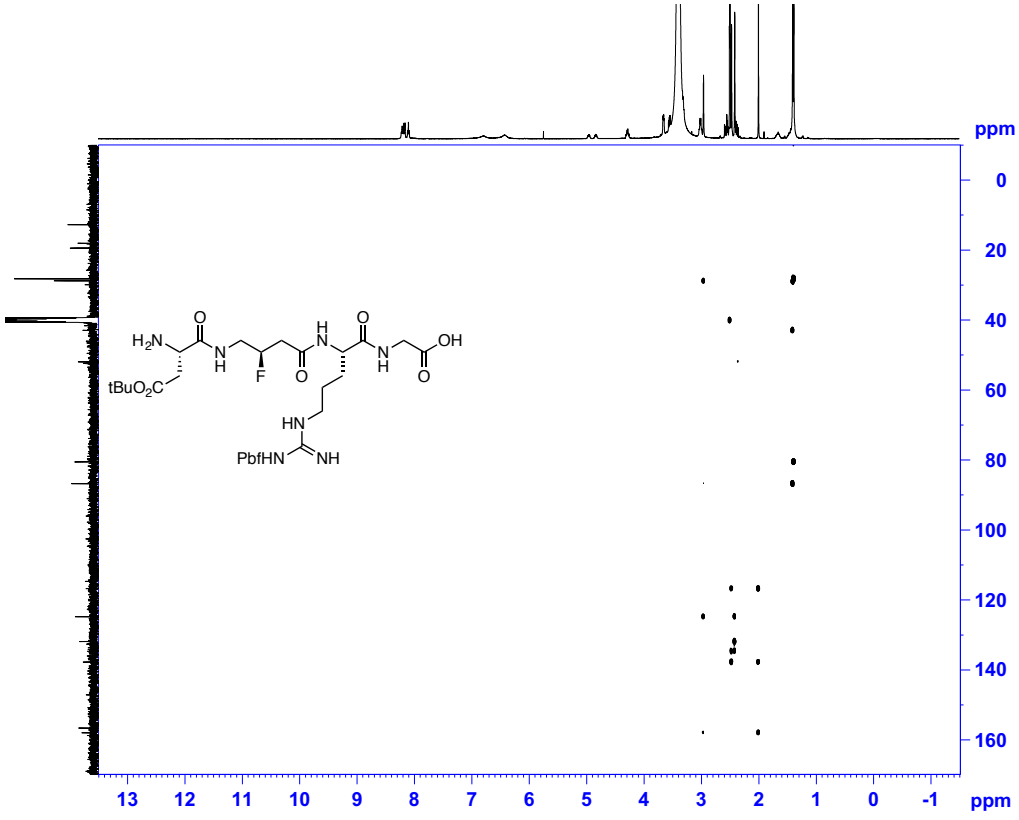
^{1}H-^{1}H COSY (400 MHz, DMSO-d$_6$) of **3b**



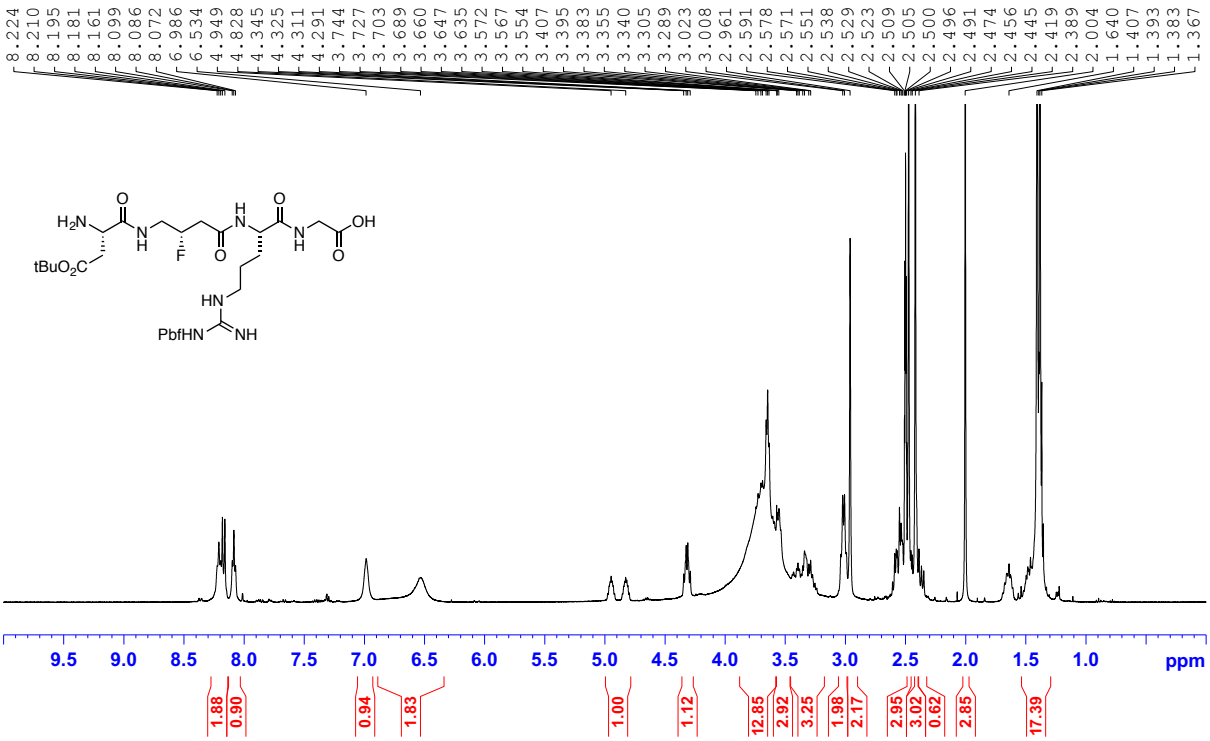
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **3b**



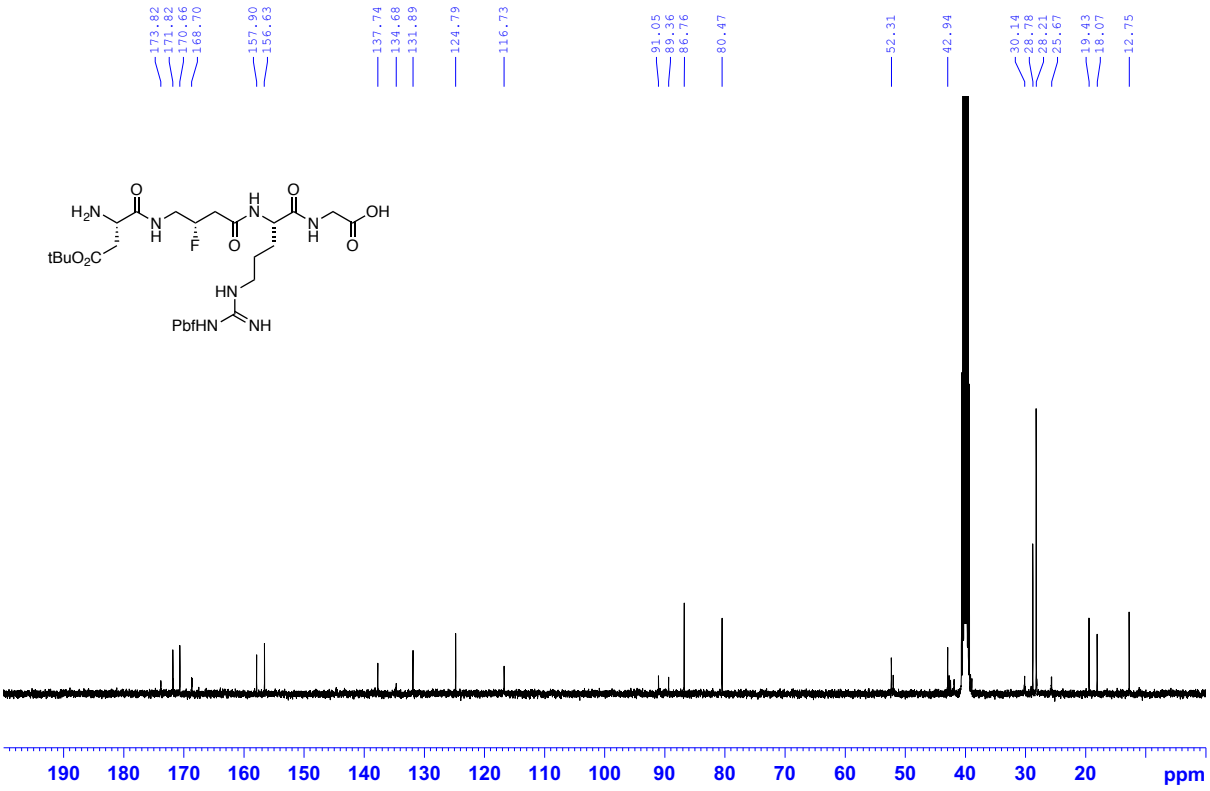
^{1}H-^{13}C HMBC (400 MHz, DMSO) of **3b**



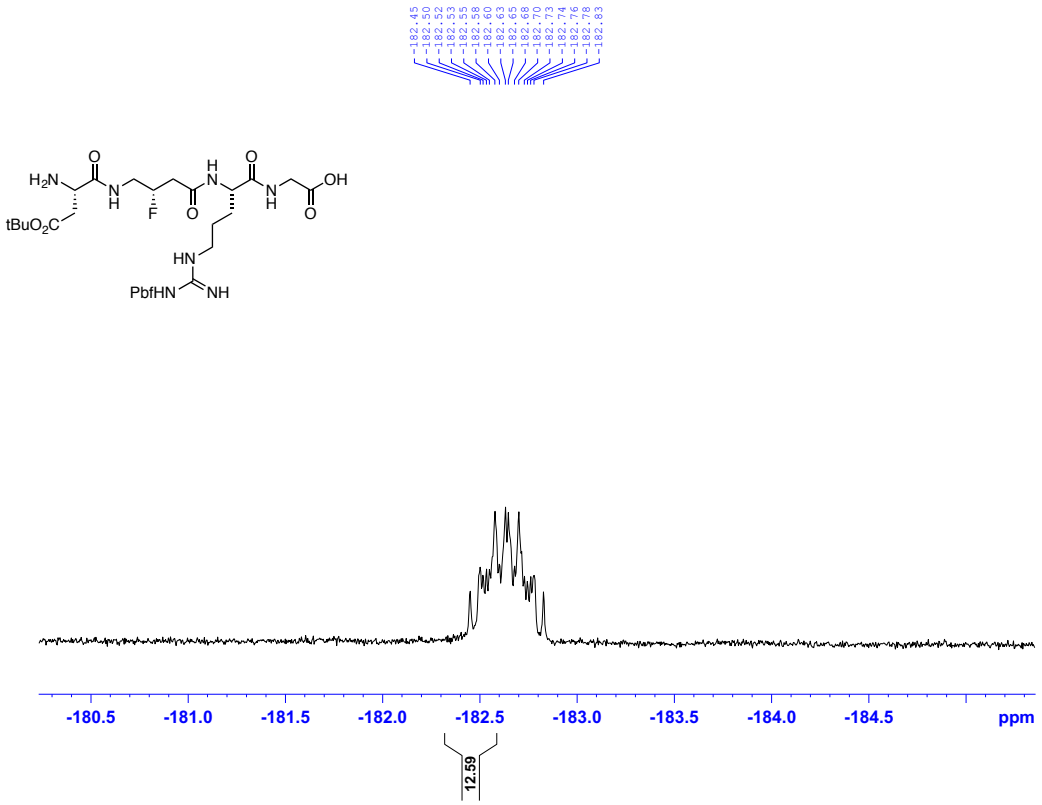
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3c**



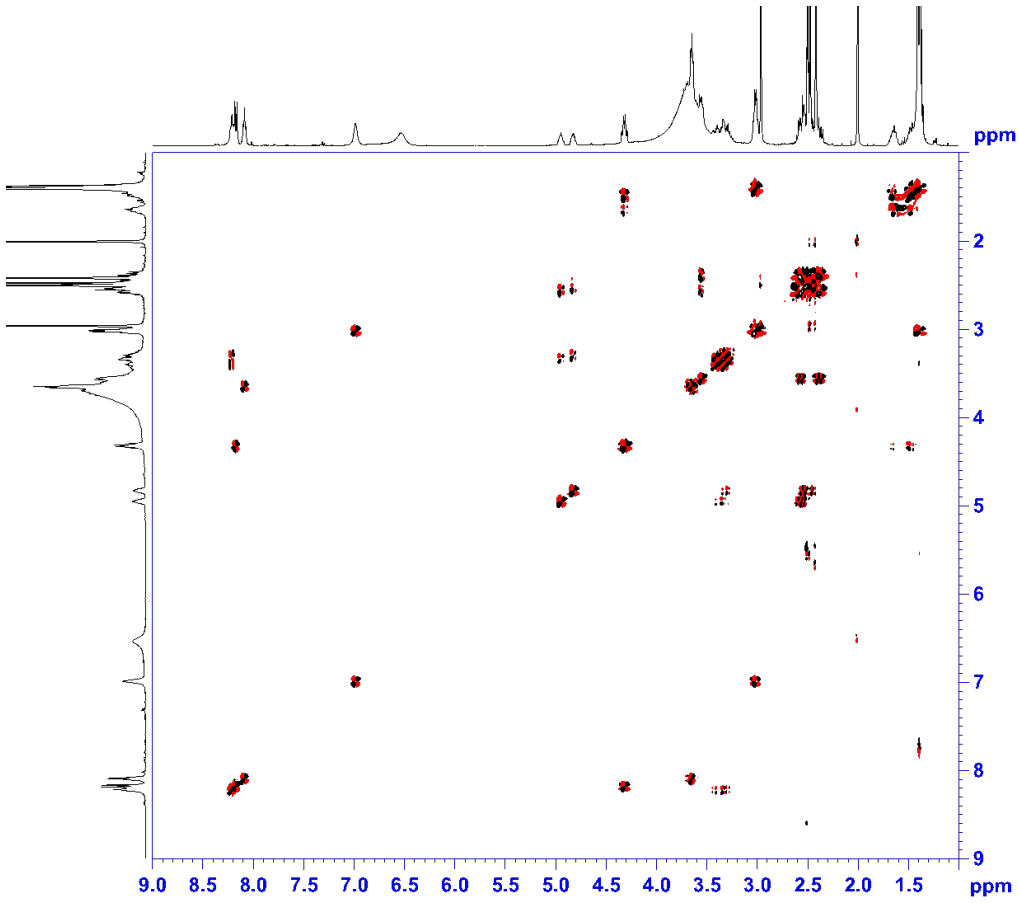
^{13}C{^{1}H} NMR (100 MHz, DMSO-d$_6$) of **3c**



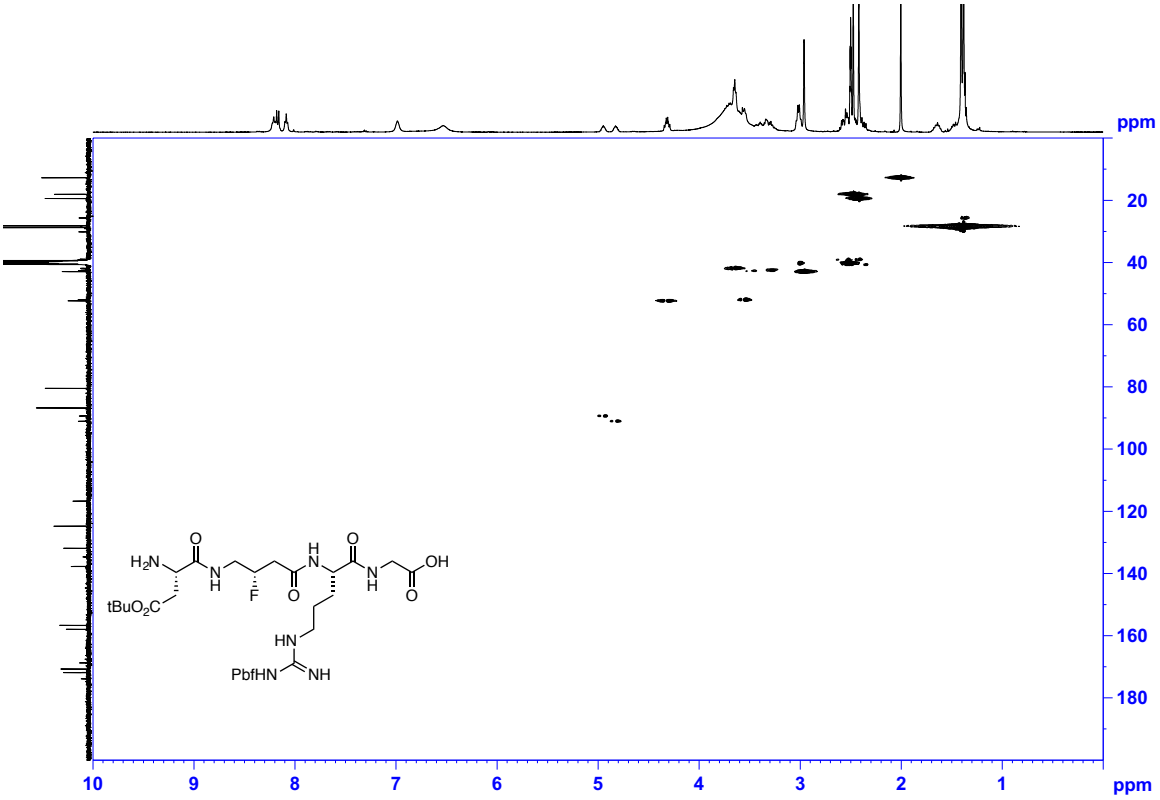
^{19}F NMR (376 MHz, DMSO-d_6) of **3c**



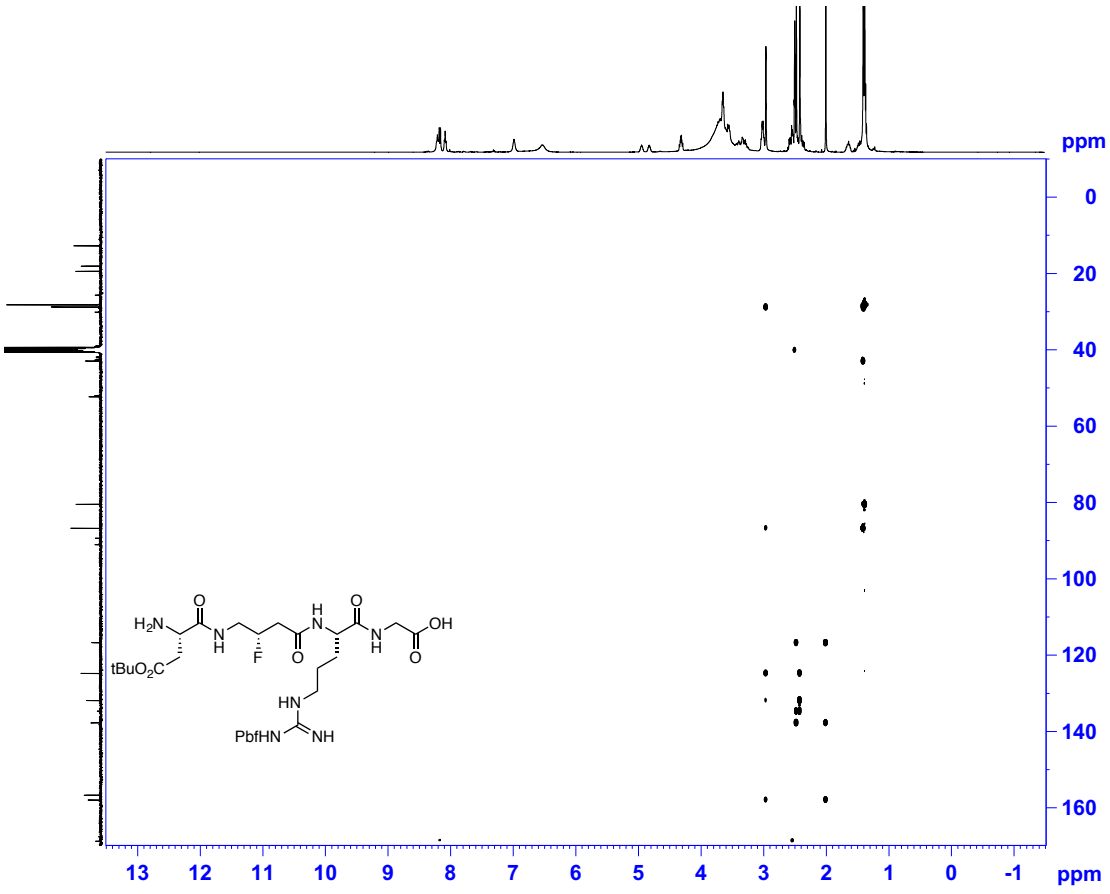
^{1}H-^{1}H COSY (400MHz, DMSO) of **3c**



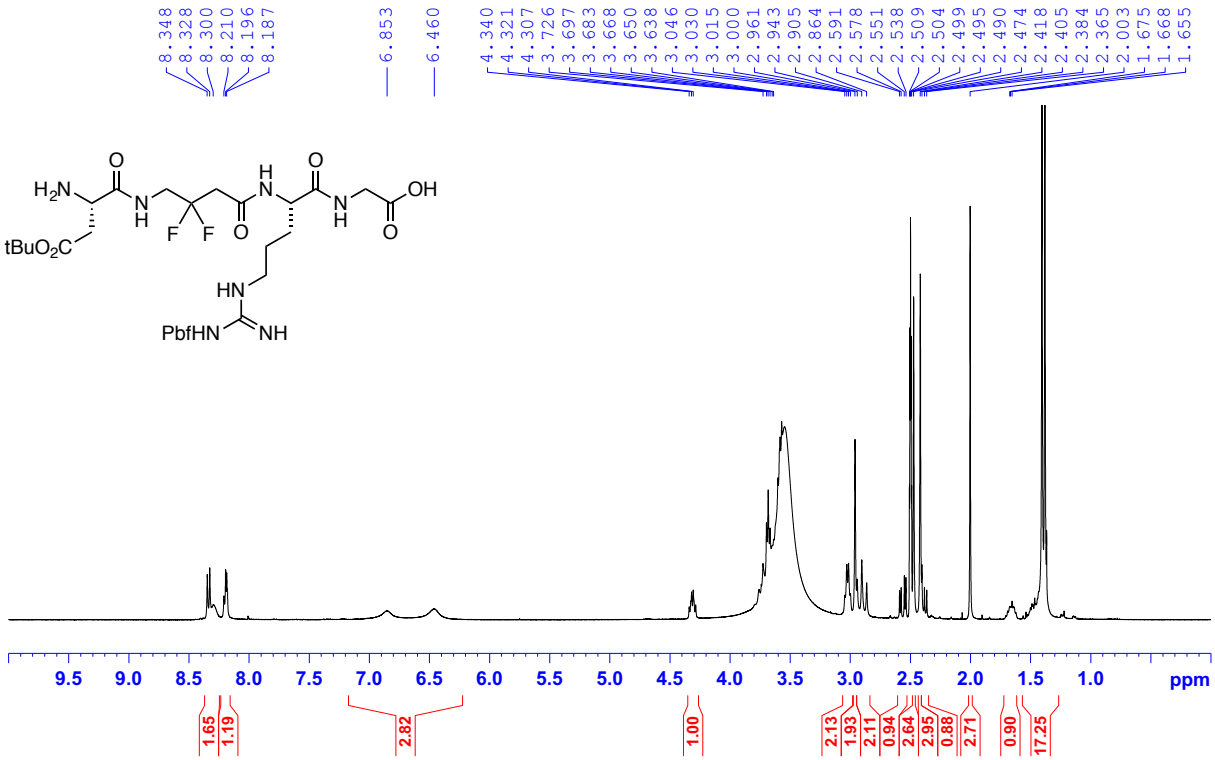
^{1}H-^{13}C HSQC (400 MHz, DMSO-d$_6$) of **3c**



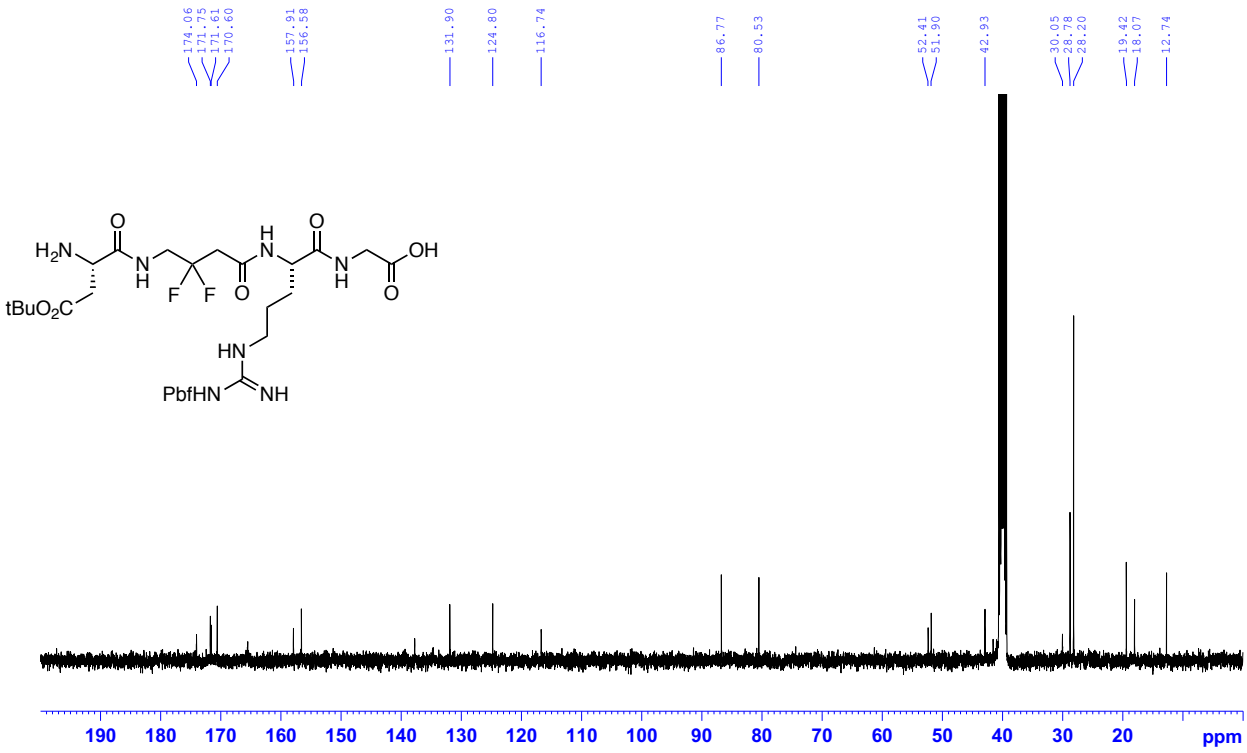
^{1}H-^{13}C HMBC (400 MHz, DMSO-d$_6$) of **3c**



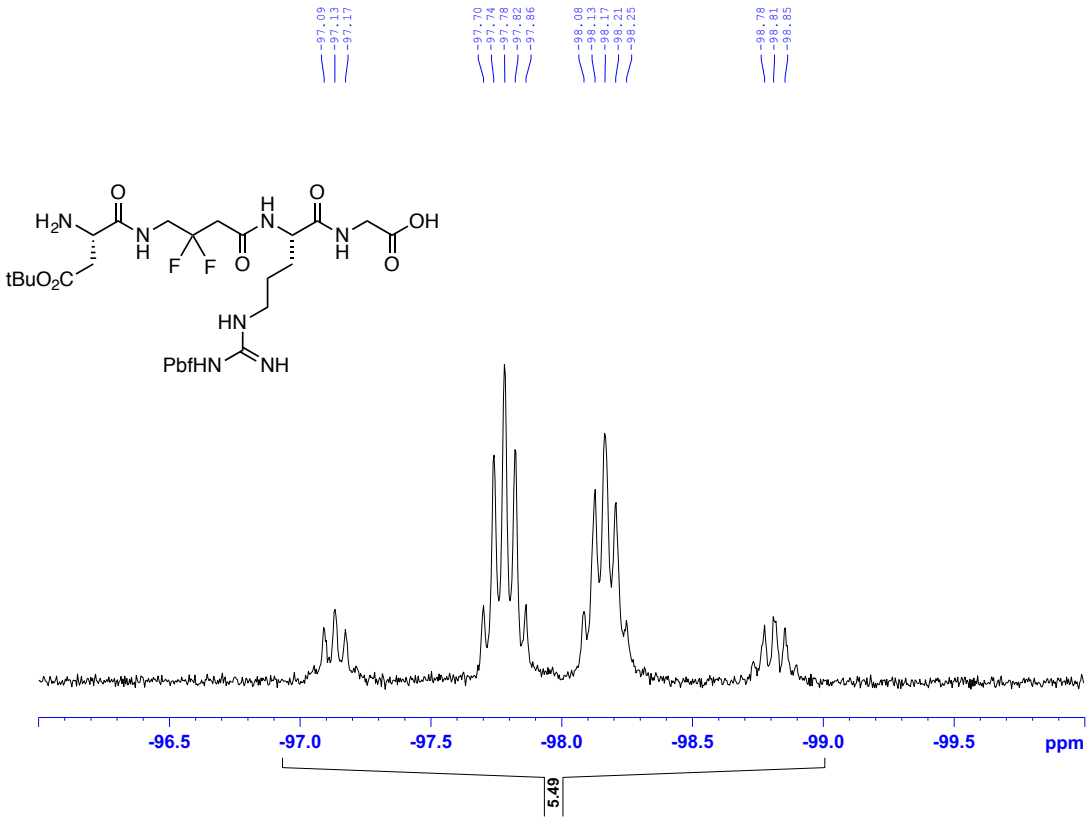
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3d**



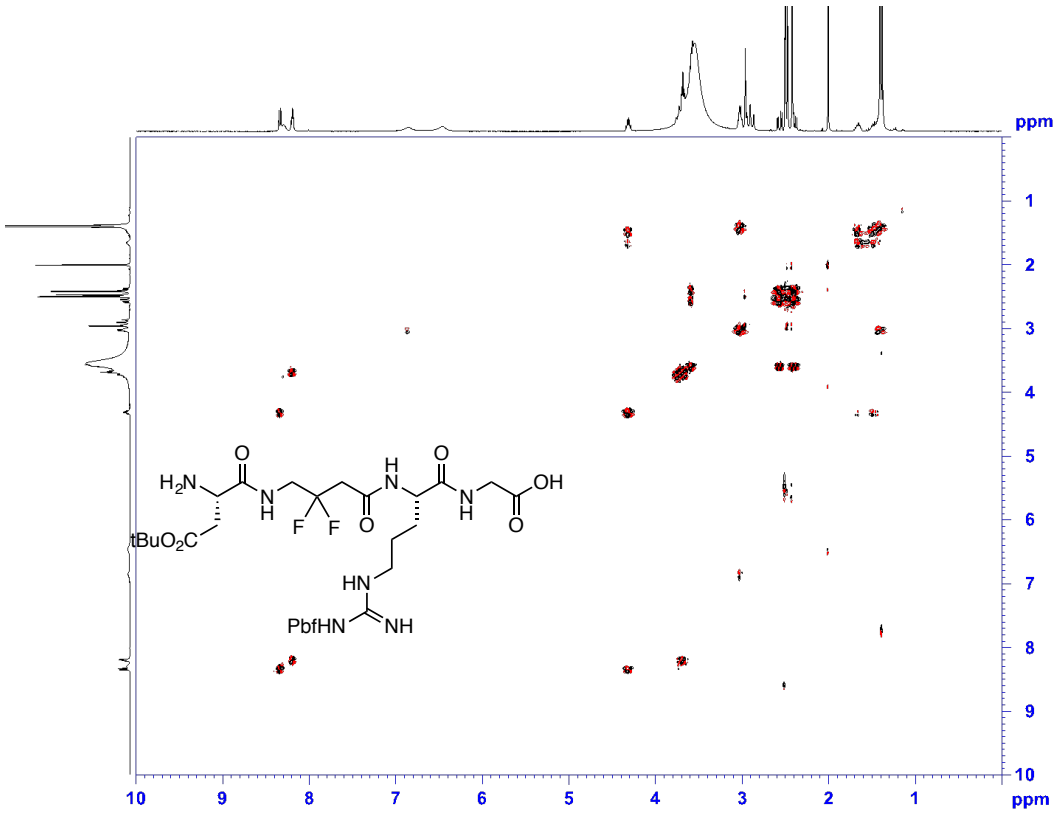
^{13}C{^{1}H} NMR (100 MHz, DMSO-d$_6$) of **3d**



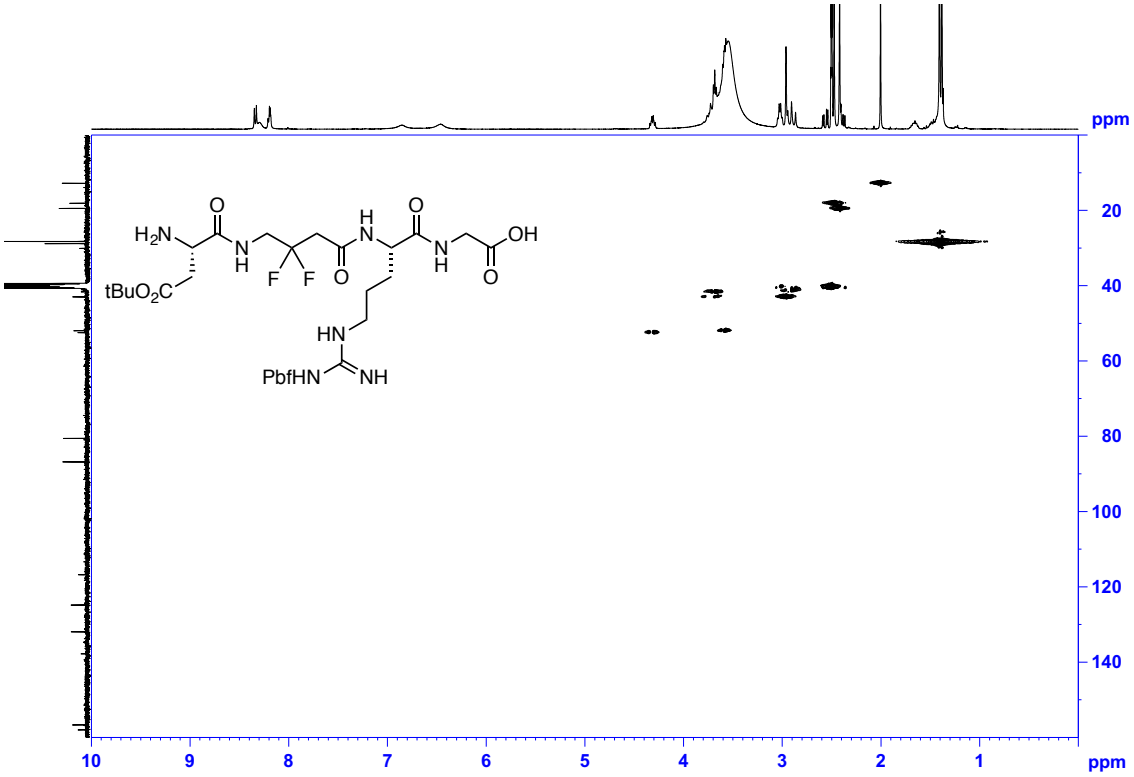
^{19}F NMR (376 MHz, DMSO-d$_6$) of **3d**



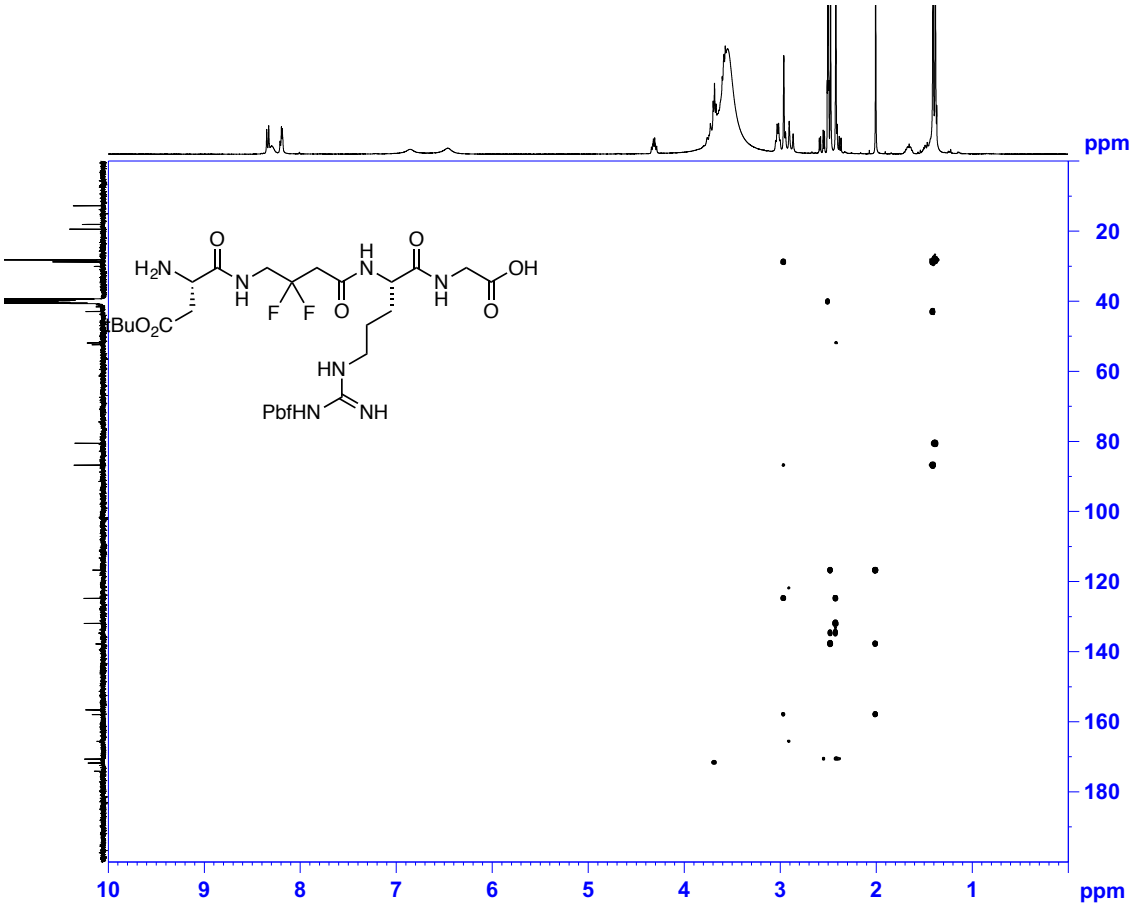
^{1}H-^{1}H COSY (400 MHz, DMSO-d$_6$) of **3d**



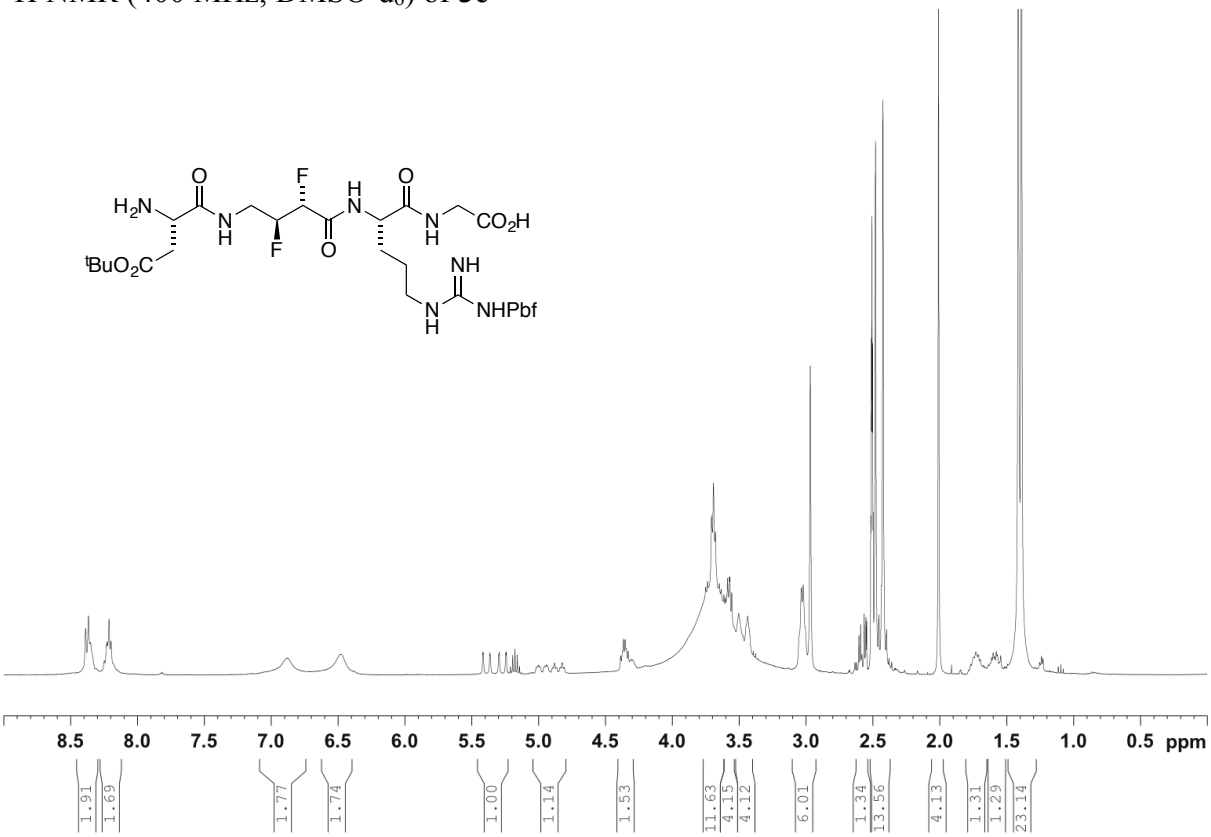
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **3d**



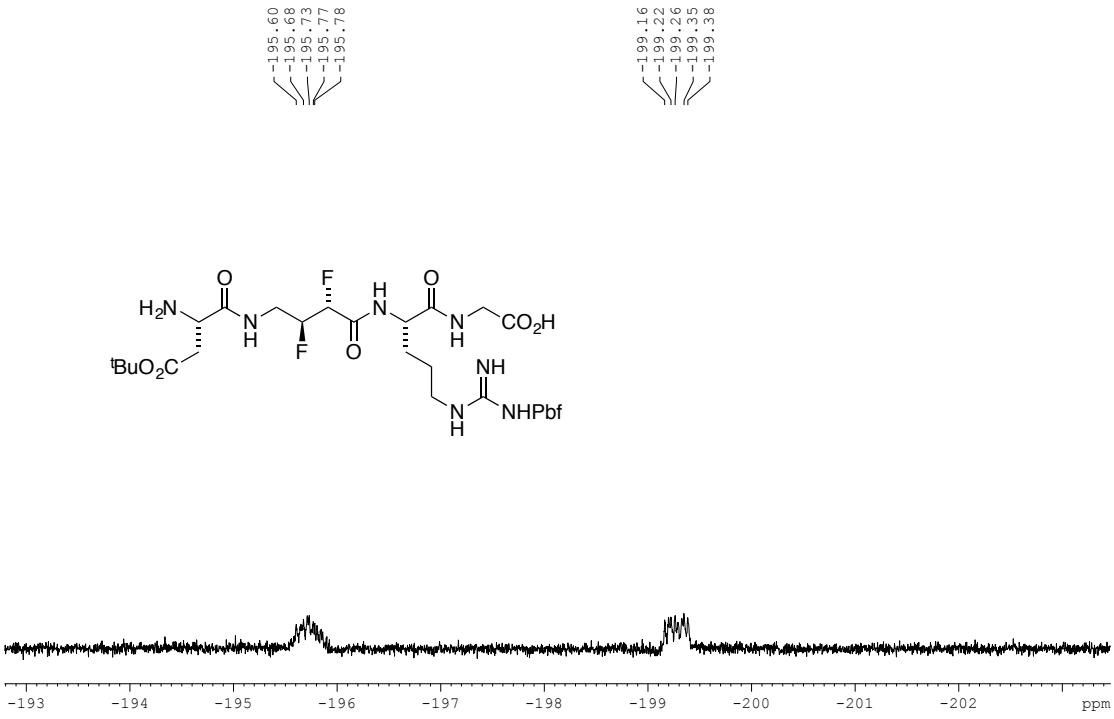
^{1}H-^{13}C HMBC (400 MHz, DMSO-d_6) of **3d**



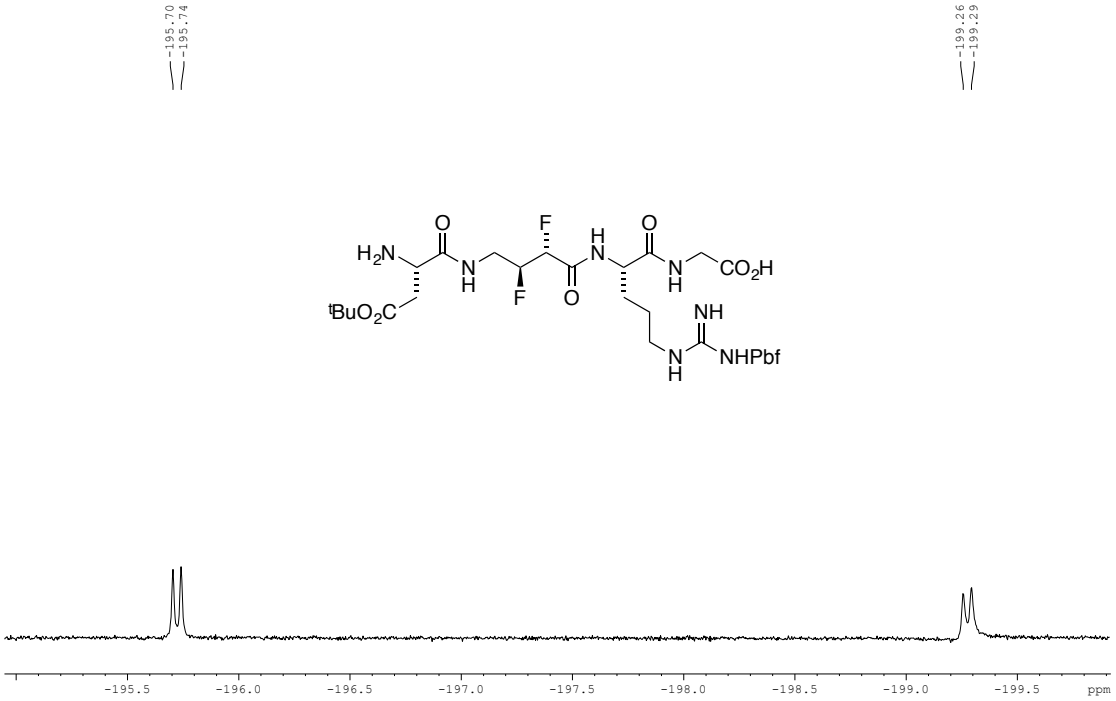
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3e**



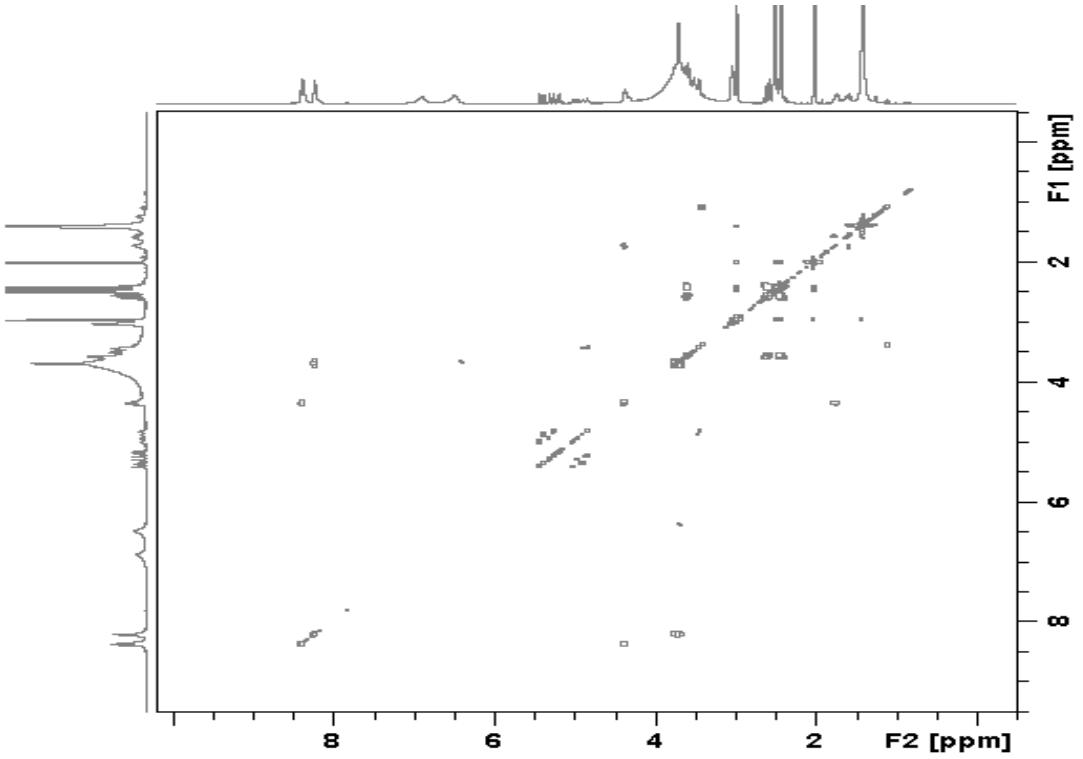
^{19}F NMR (376 MHz, DMSO-d$_6$) of **3e**



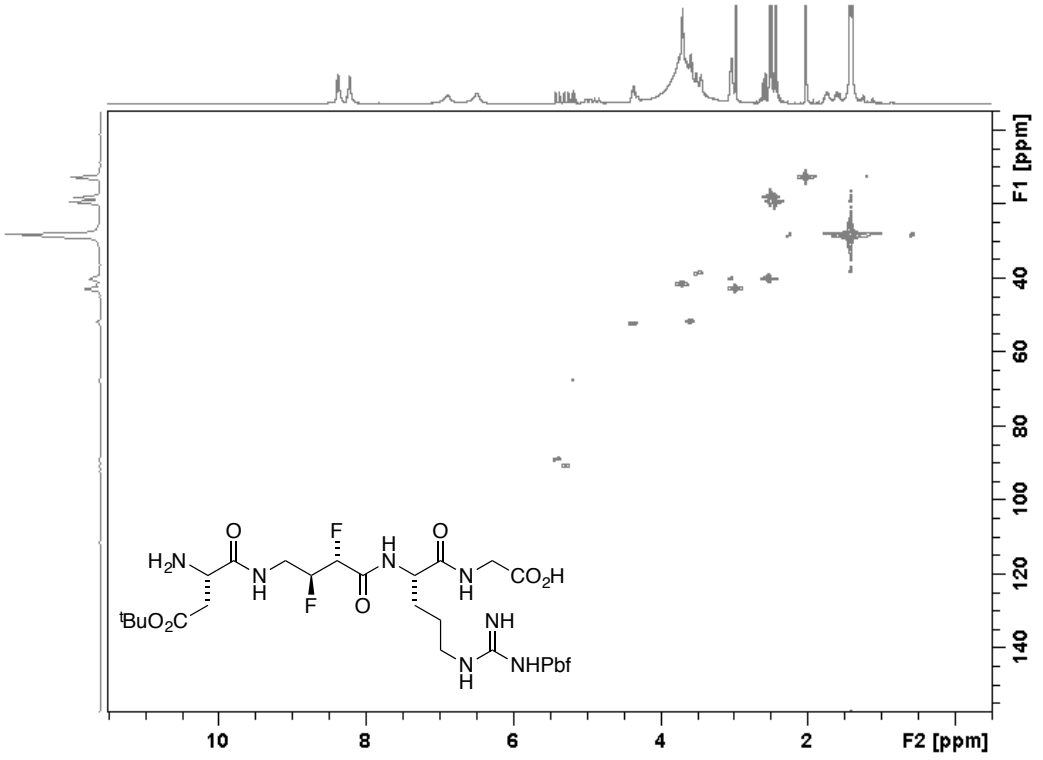
$^{19}F\{^{1}H\}$ NMR (376 MHz, DMSO-d_6) of **3e**



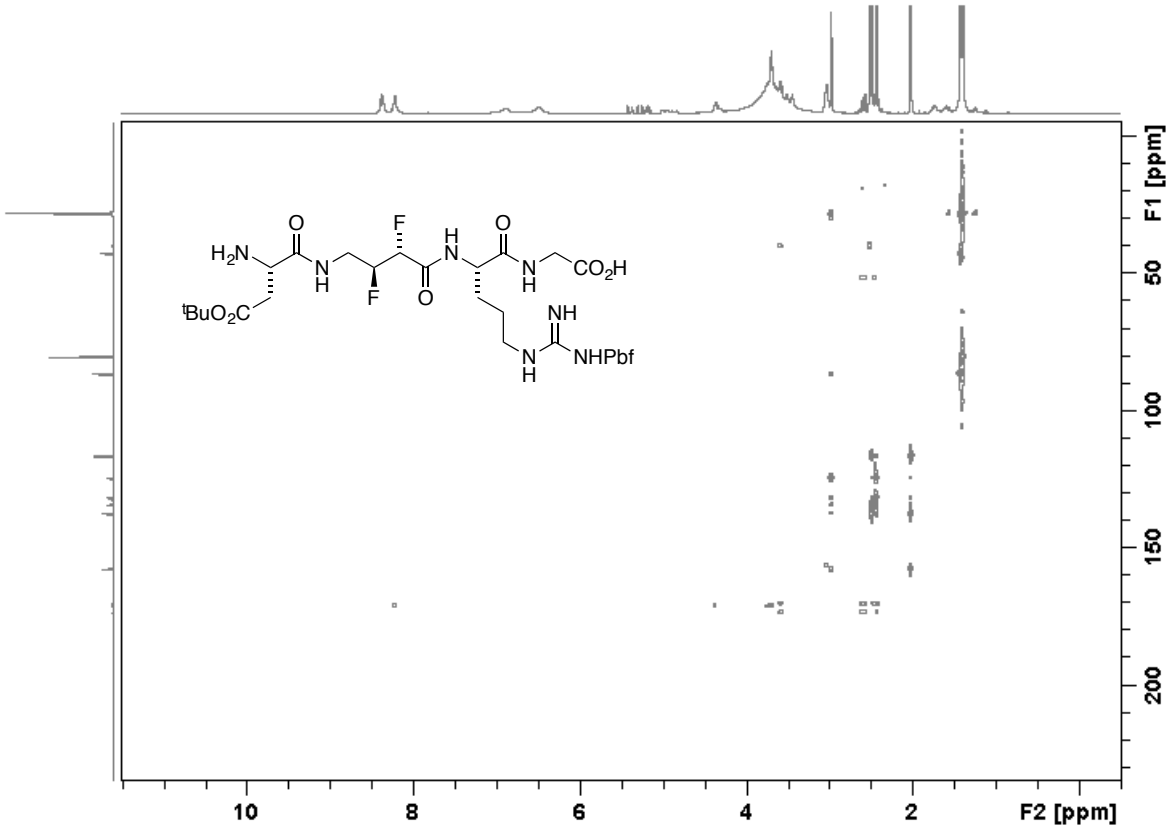
^{1}H-^{1}H COSY (400 MHz, DMSO-d_6) of **3e**



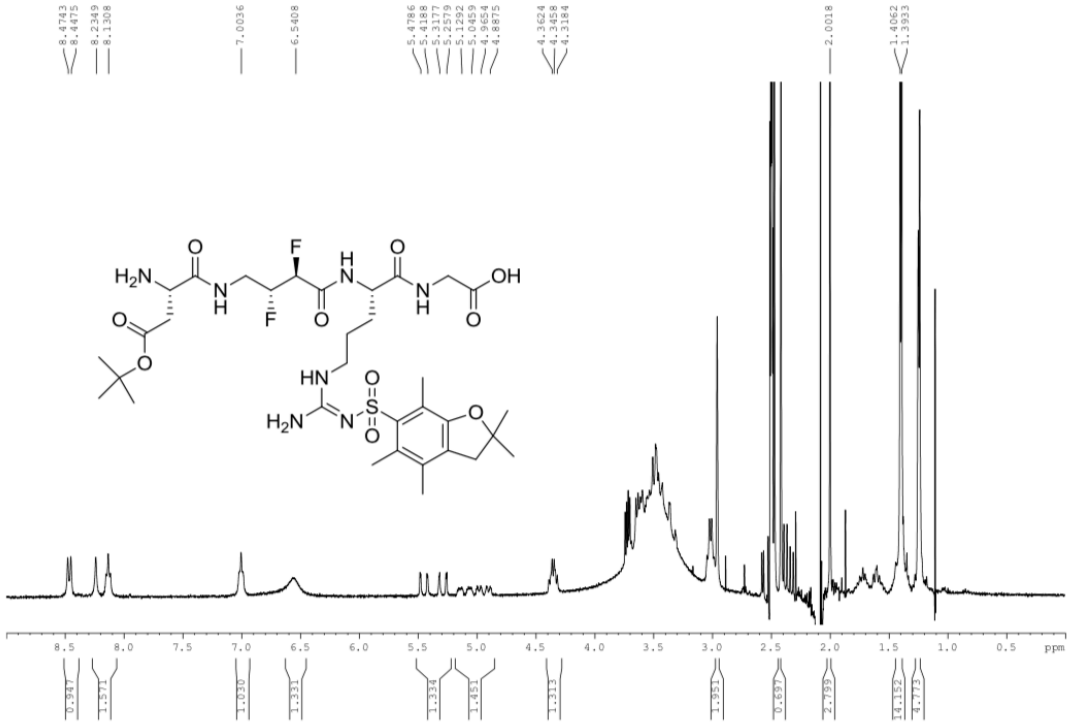
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **3e**



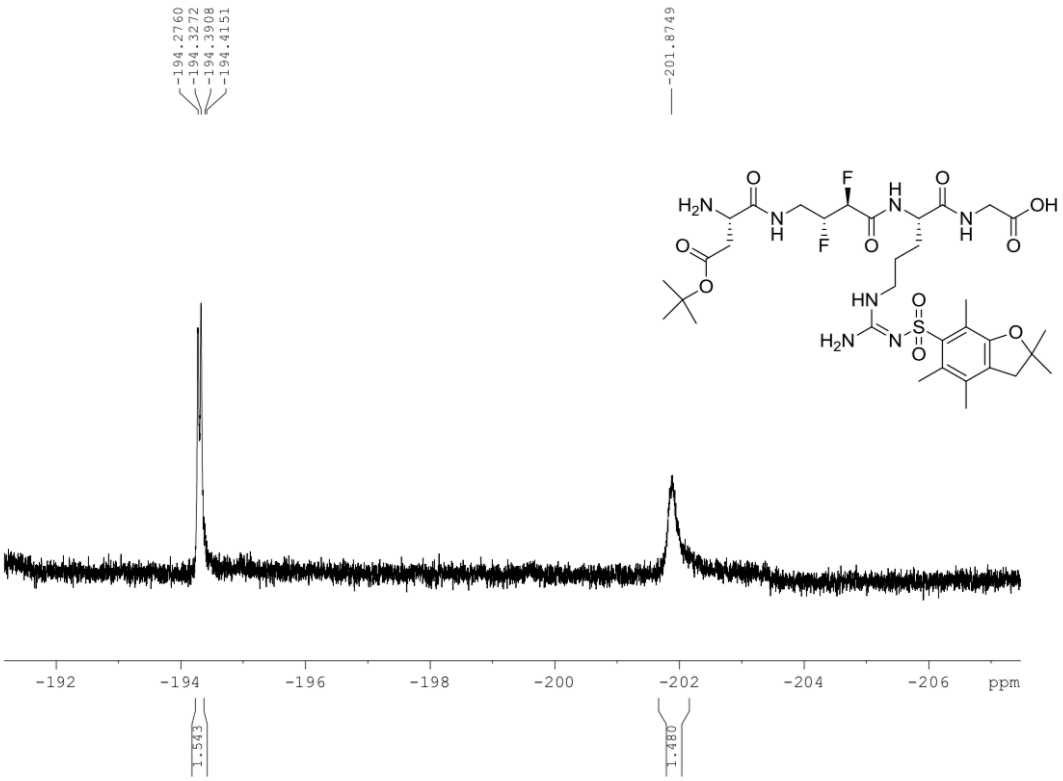
^{1}H-^{13}C HMBC (400 MHz, DMSO-d_6) of **3e**



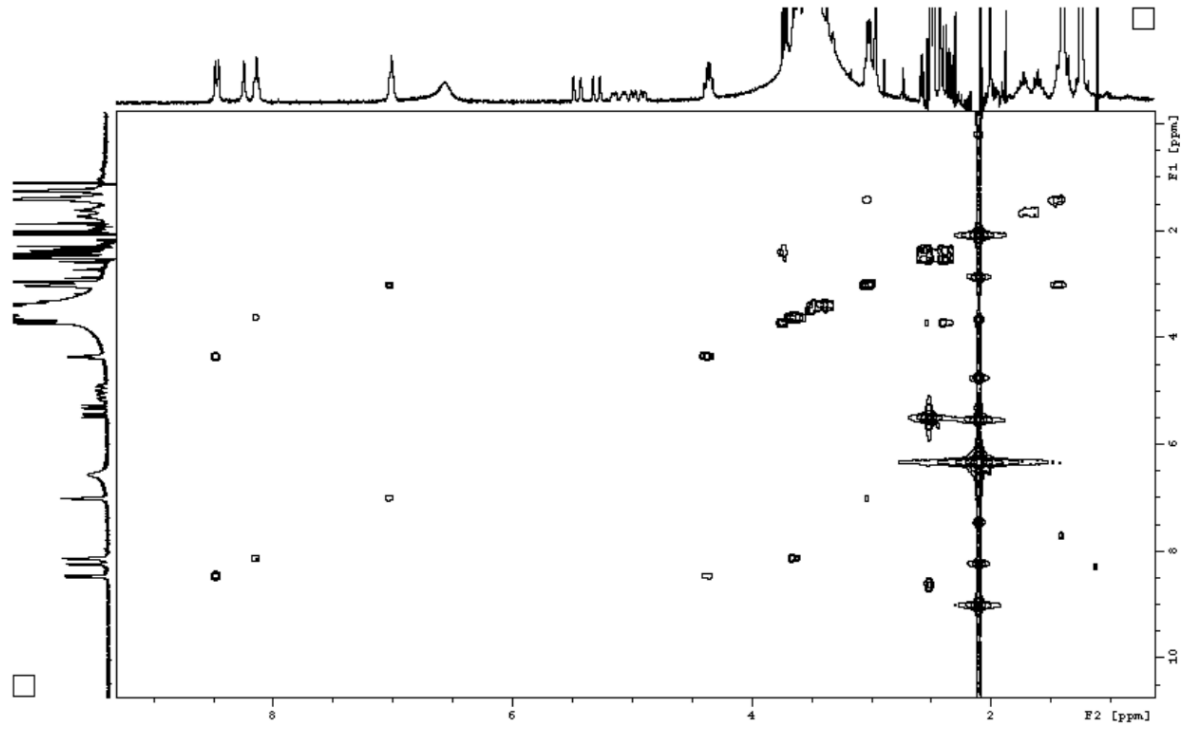
^{1}H NMR (300 MHz, DMSO-d$_6$) of **3f**



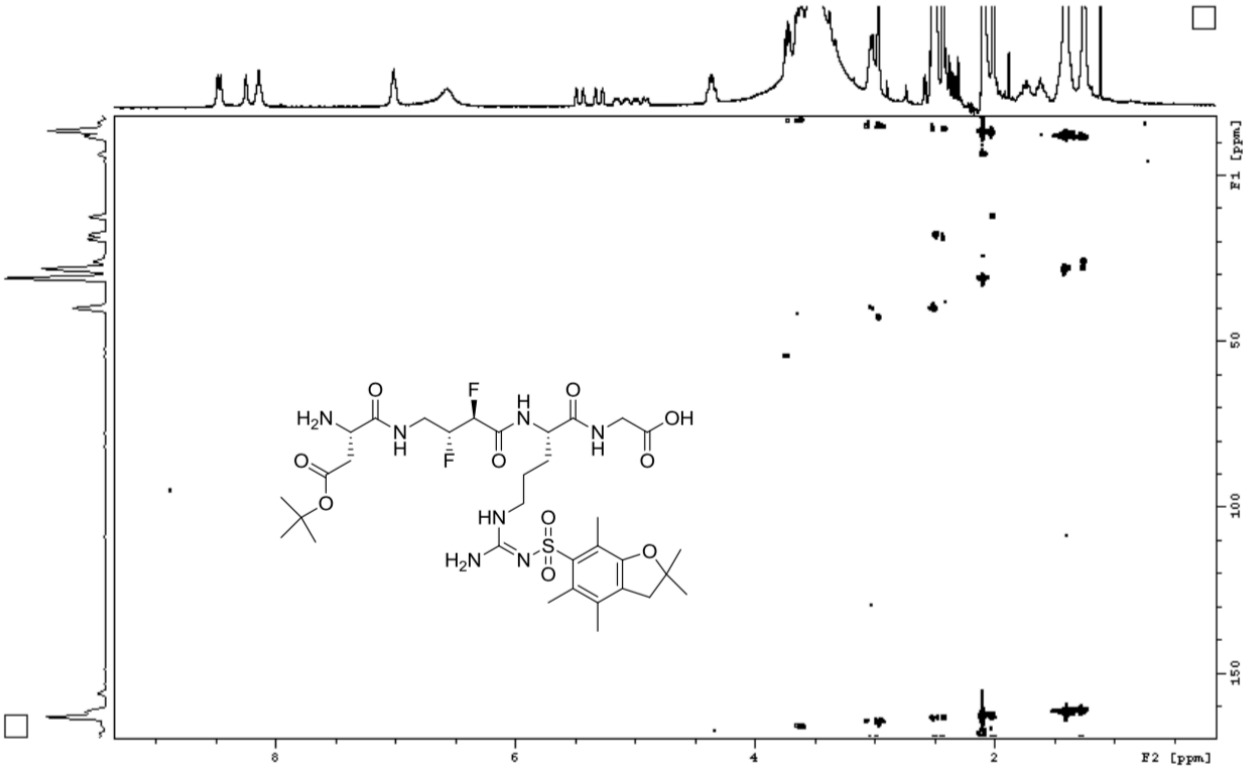
^{19}F{^{1}H} NMR (282 MHz, DMSO-d$_6$) of **3f**



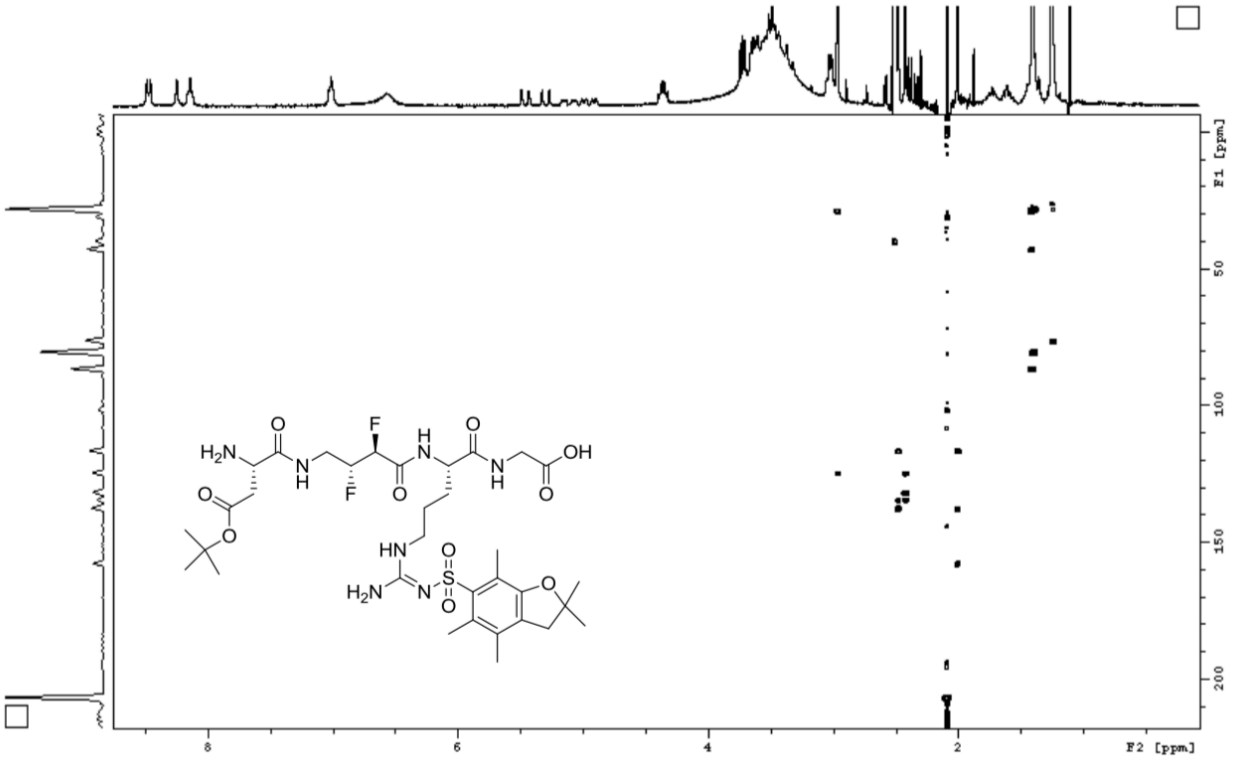
^{1}H-^{1}H COSY (300 MHz, DMSO-d_6) of **3f**



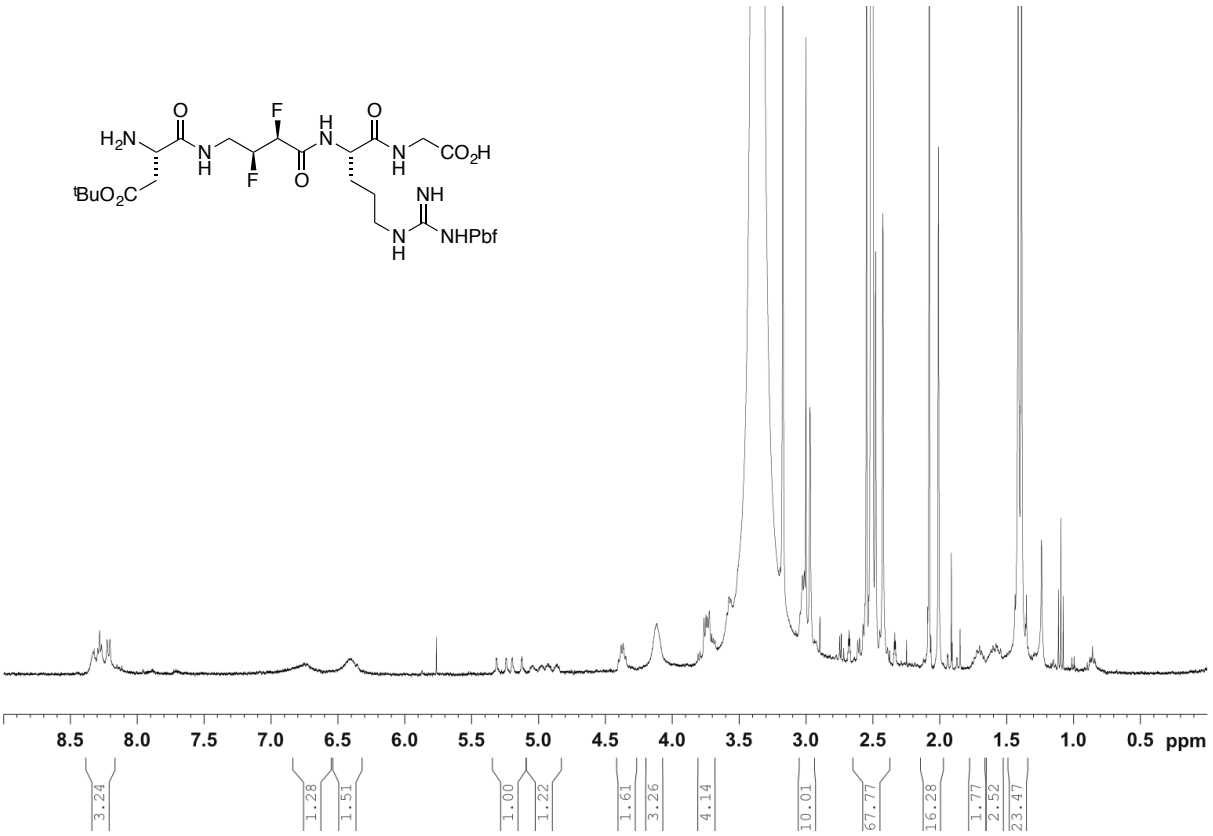
^{1}H-^{13}C HSQC (300 MHz, DMSO-d_6) of **3f**



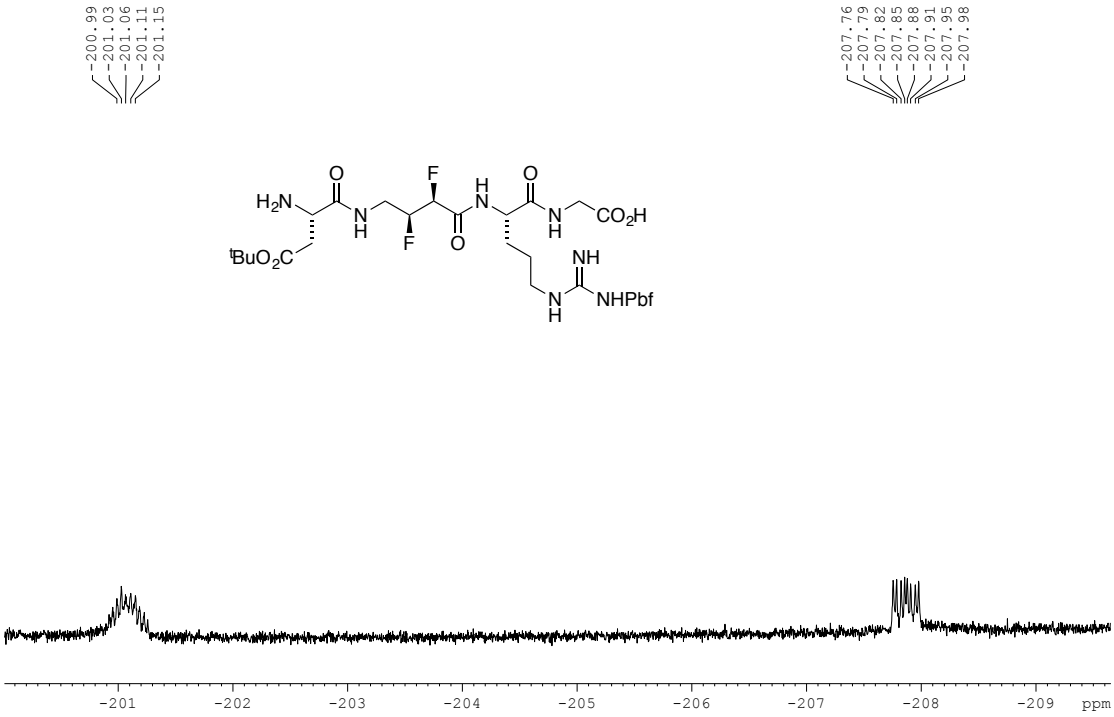
^{1}H-^{13}C HMBC (300 MHz, DMSO-d$_6$) of **3f**



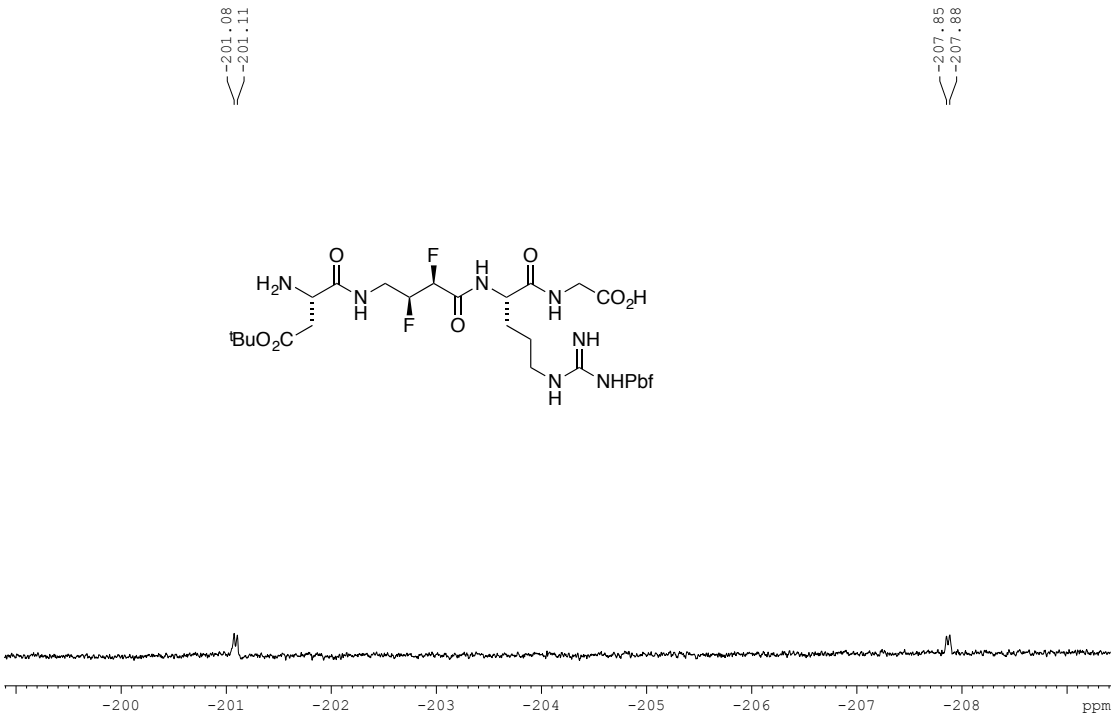
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3g**



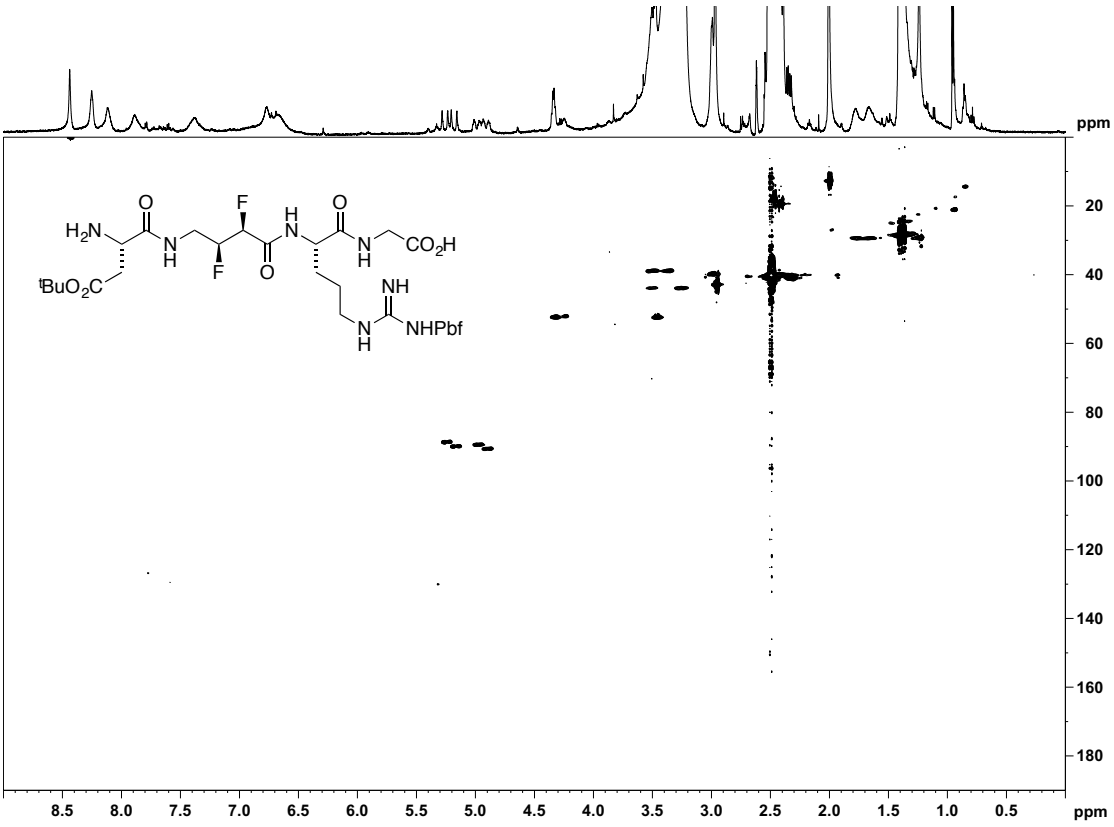
^{19}F NMR (376 MHz, DMSO-d_6) of **3g**



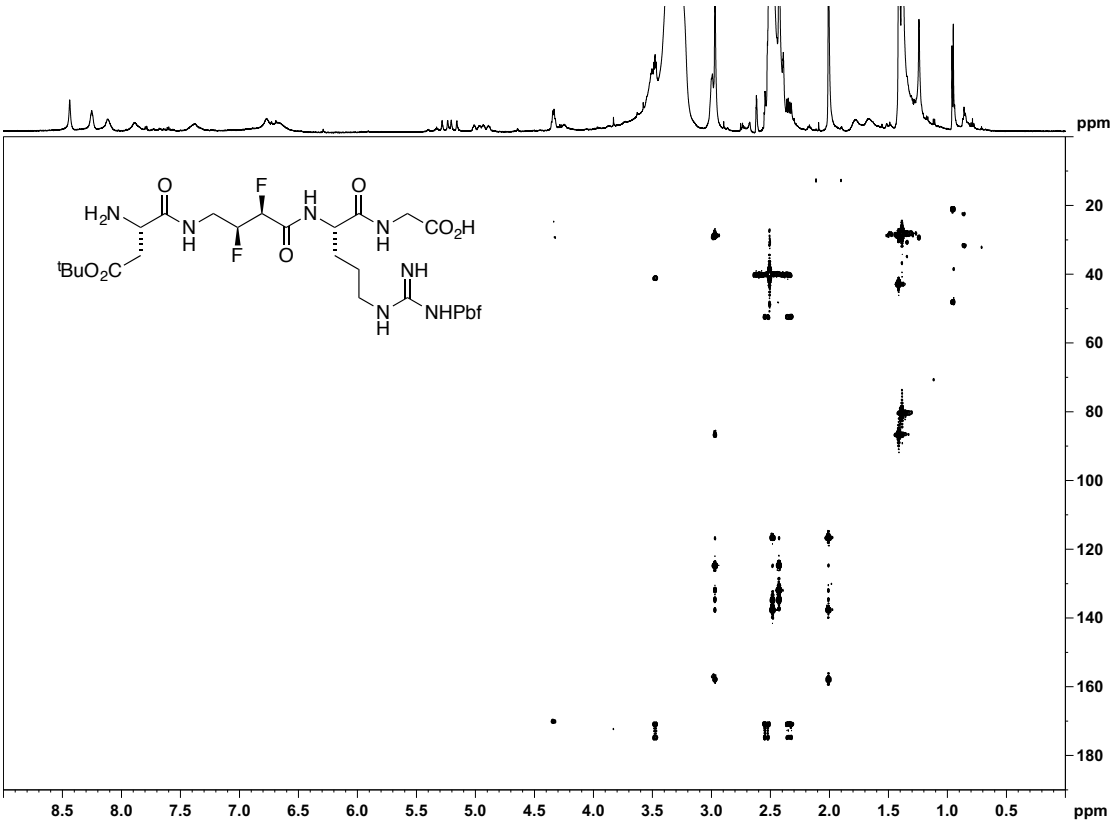
$^{19}F\{^{1}H\}$ NMR (376 MHz, DMSO-d_6) of **3g**



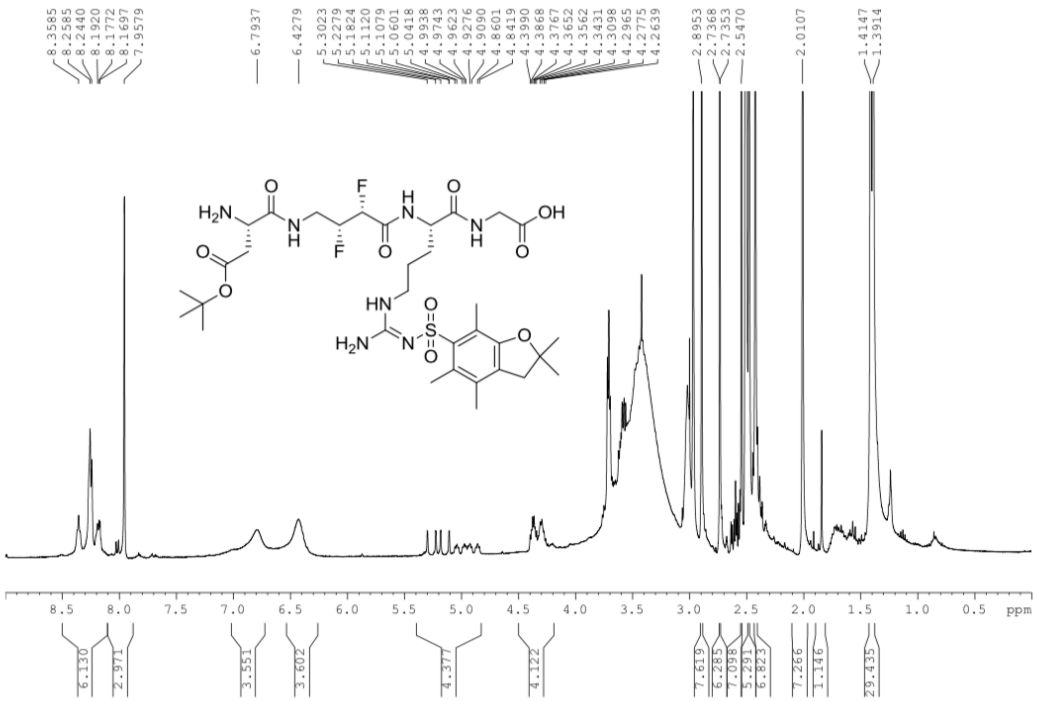
^{1}H-^{13}C HSQC (600 MHz, DMSO-d_6) of **3g**



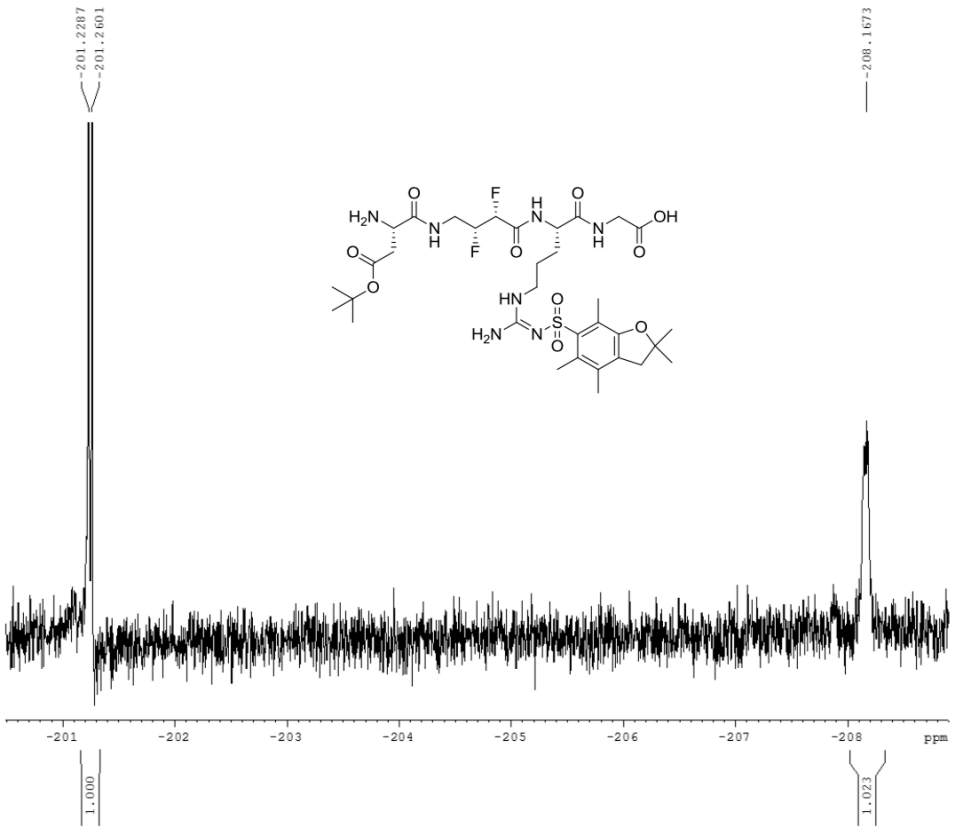
^{1}H-^{13}C HMBC (600 MHz, DMSO-d_6) of **3g**



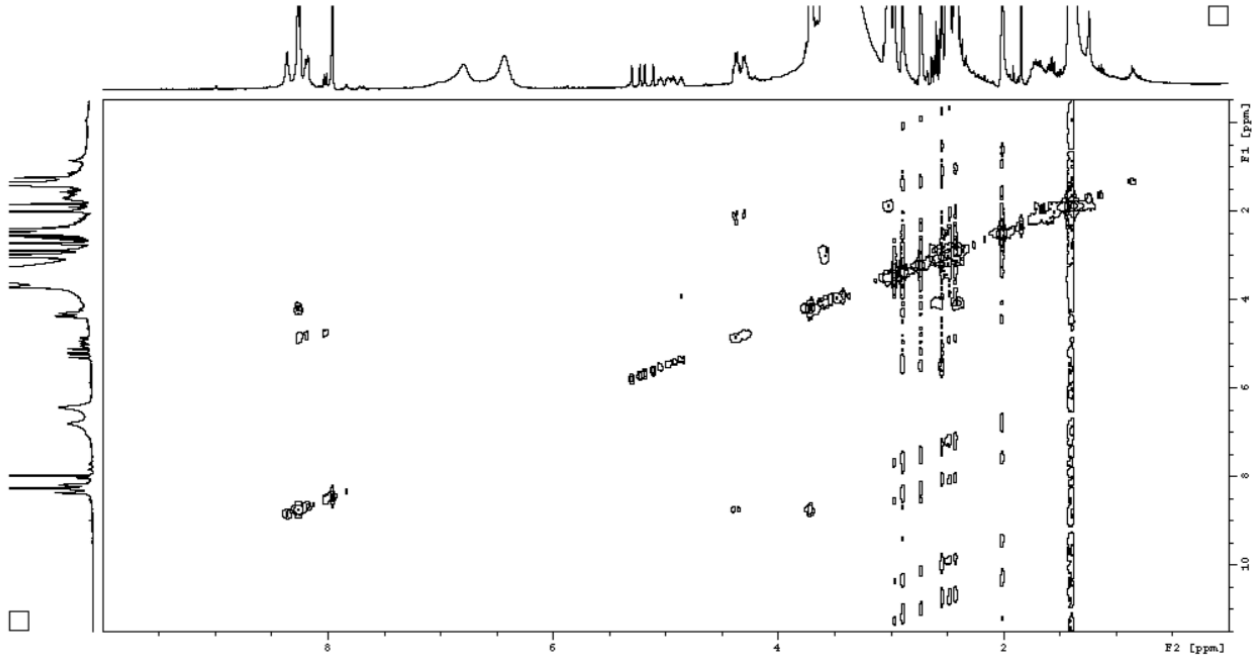
^{1}H NMR (400 MHz, DMSO-d$_6$) of **3h**



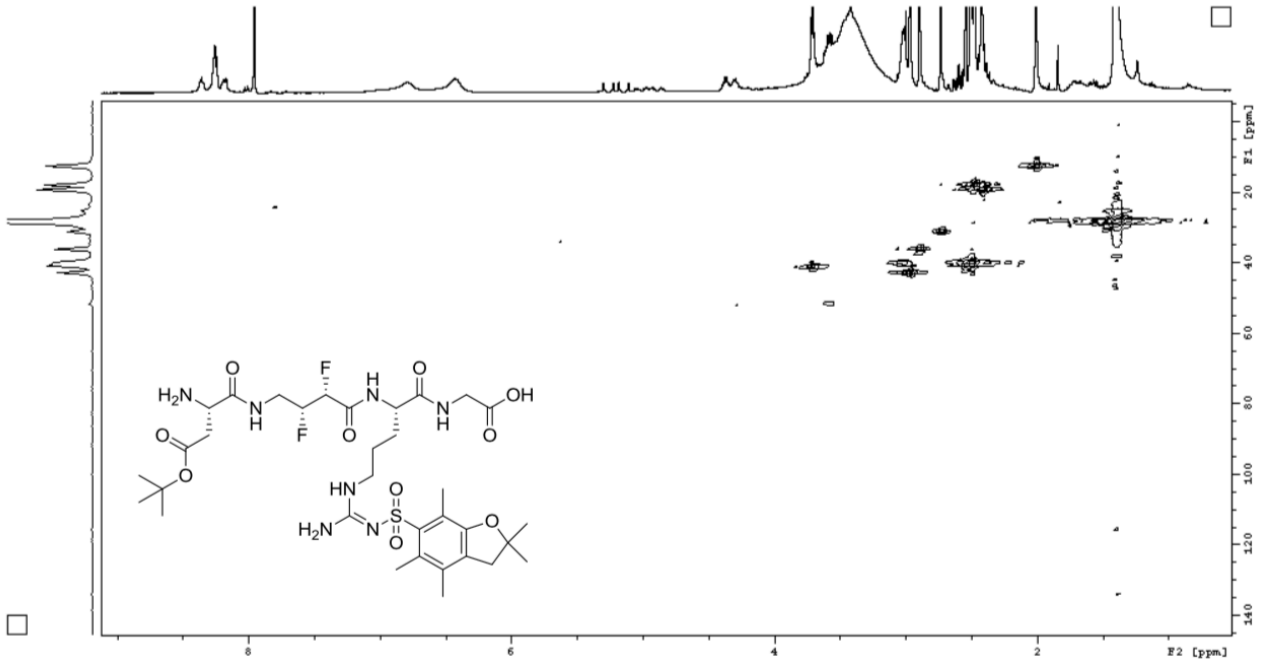
^{19}F{^{1}H} NMR (376 MHz, DMSO-d$_6$) of **3h**



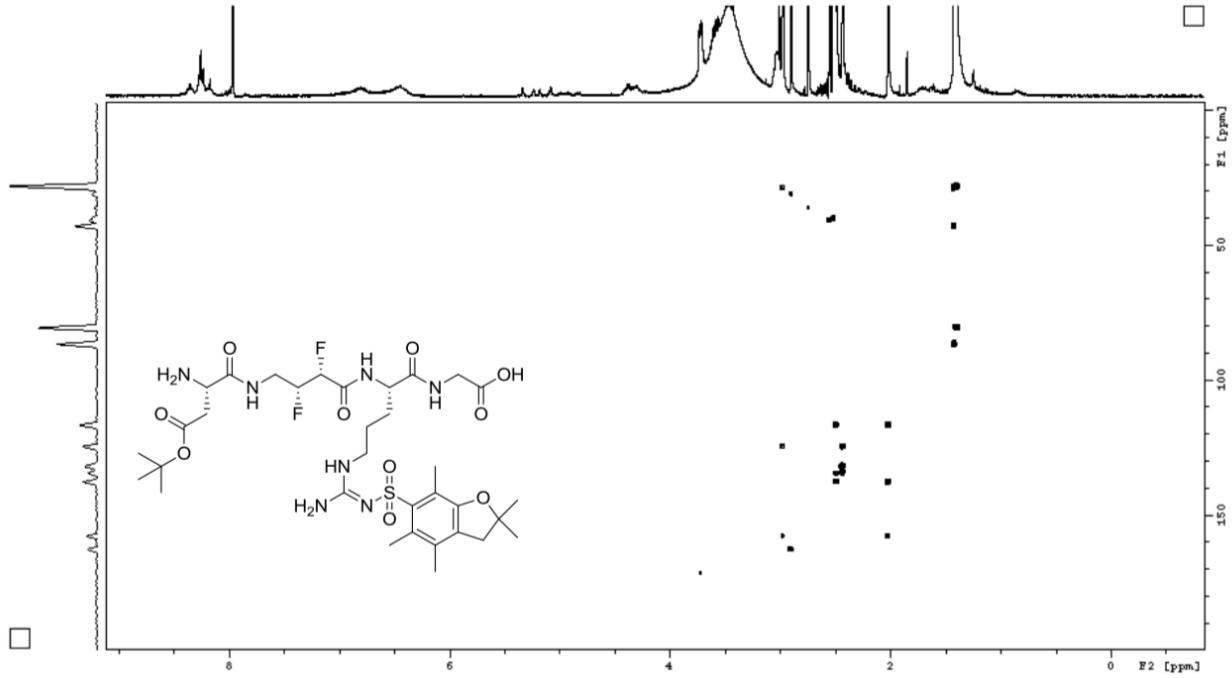
^{1}H-^{1}H COSY (400 MHz, DMSO-d_6) of **3h**



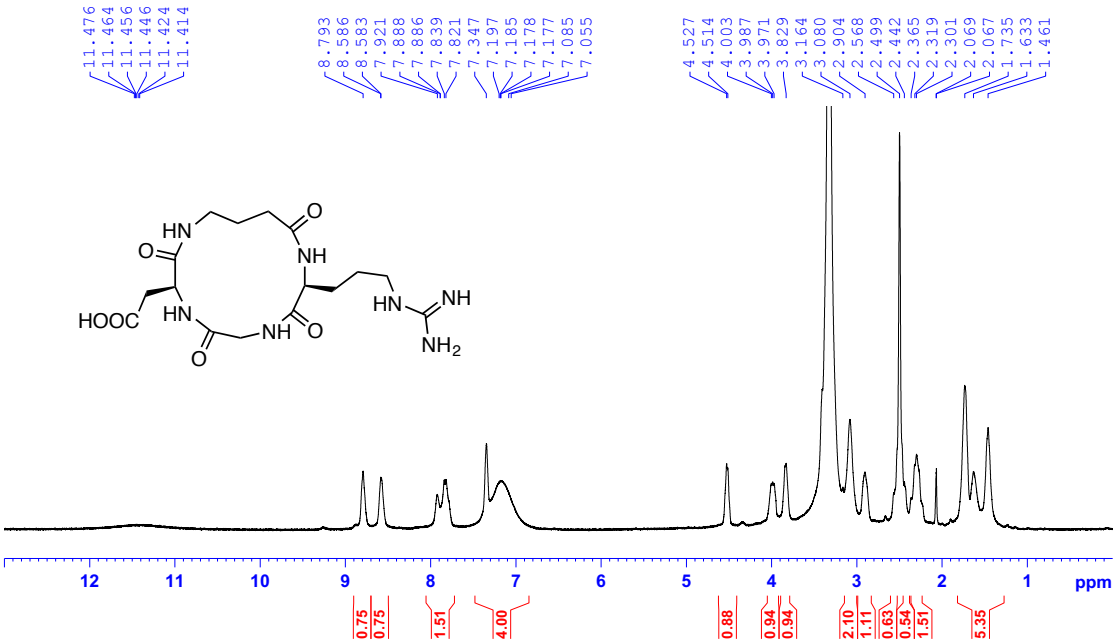
^{1}H-^{13}C HSQC (300 MHz, DMSO-d_6) of **3h**



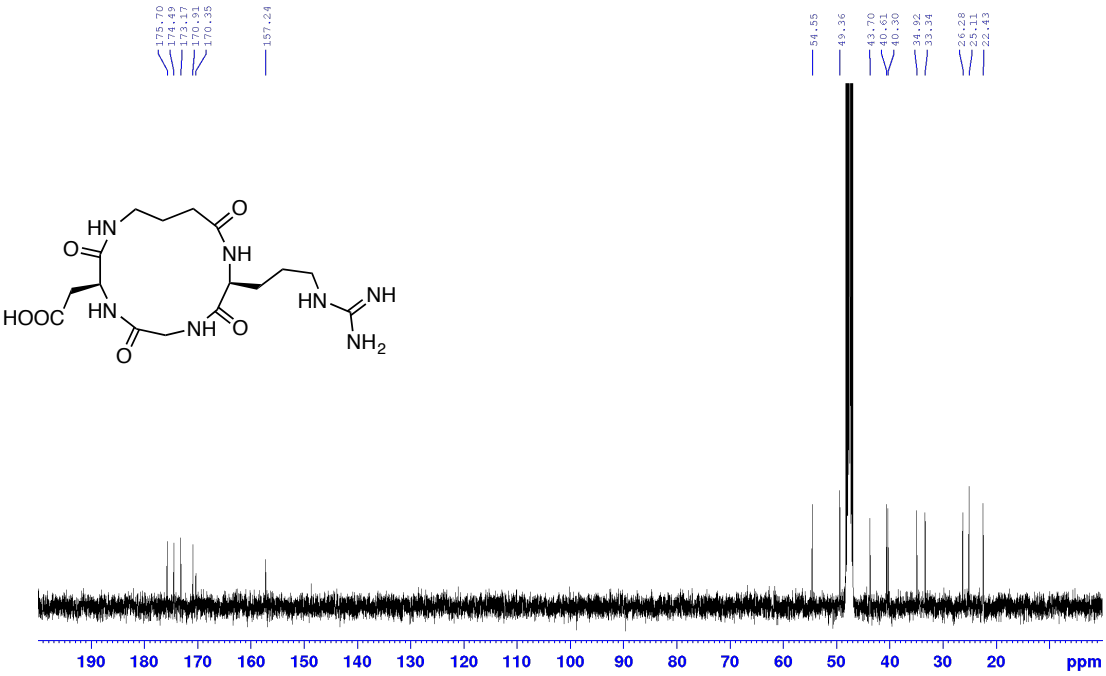
^{1}H-^{13}C HMBC (300 MHz, DMSO-d_6) of **3h**



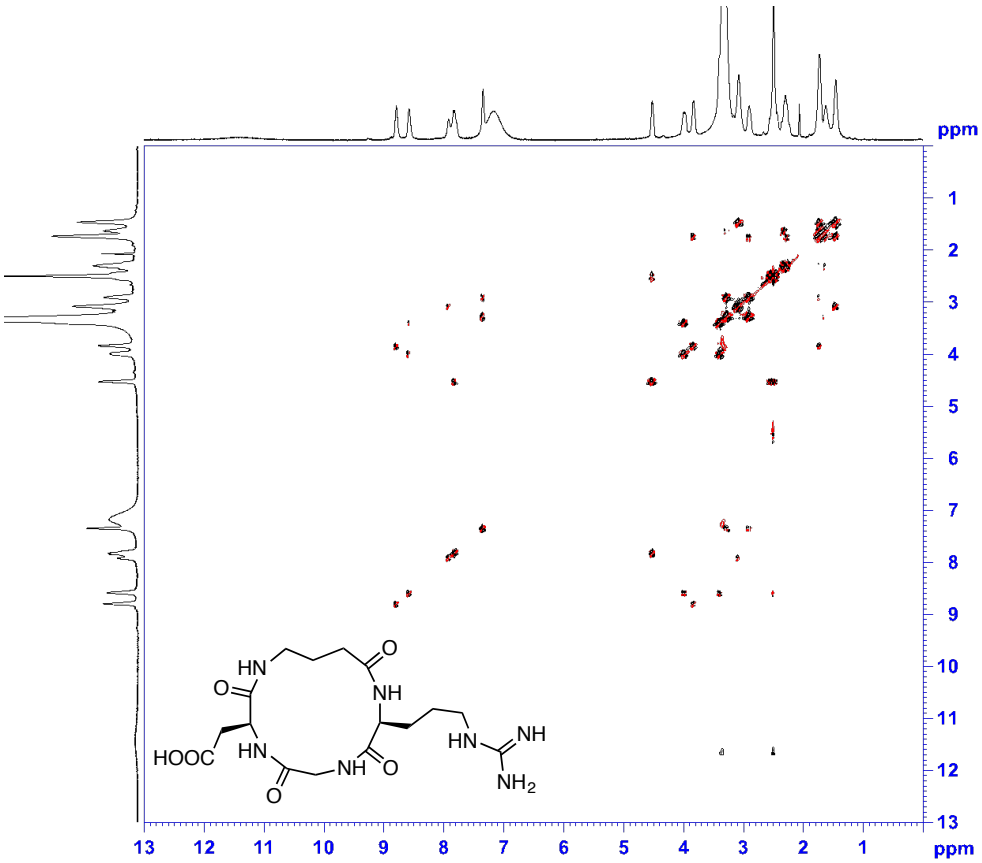
^{1}H NMR (400 MHz, DMSO-d_6) of **4a**



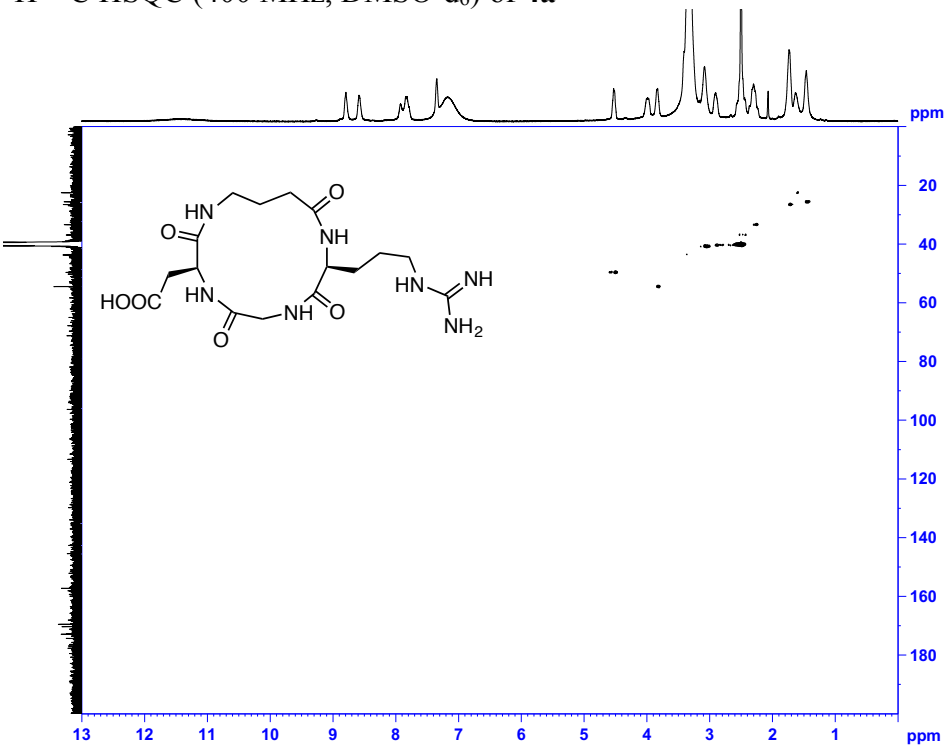
$^{13}C\{^{1}H\}$ NMR (100 MHz, DMSO-d_6) of **4a**



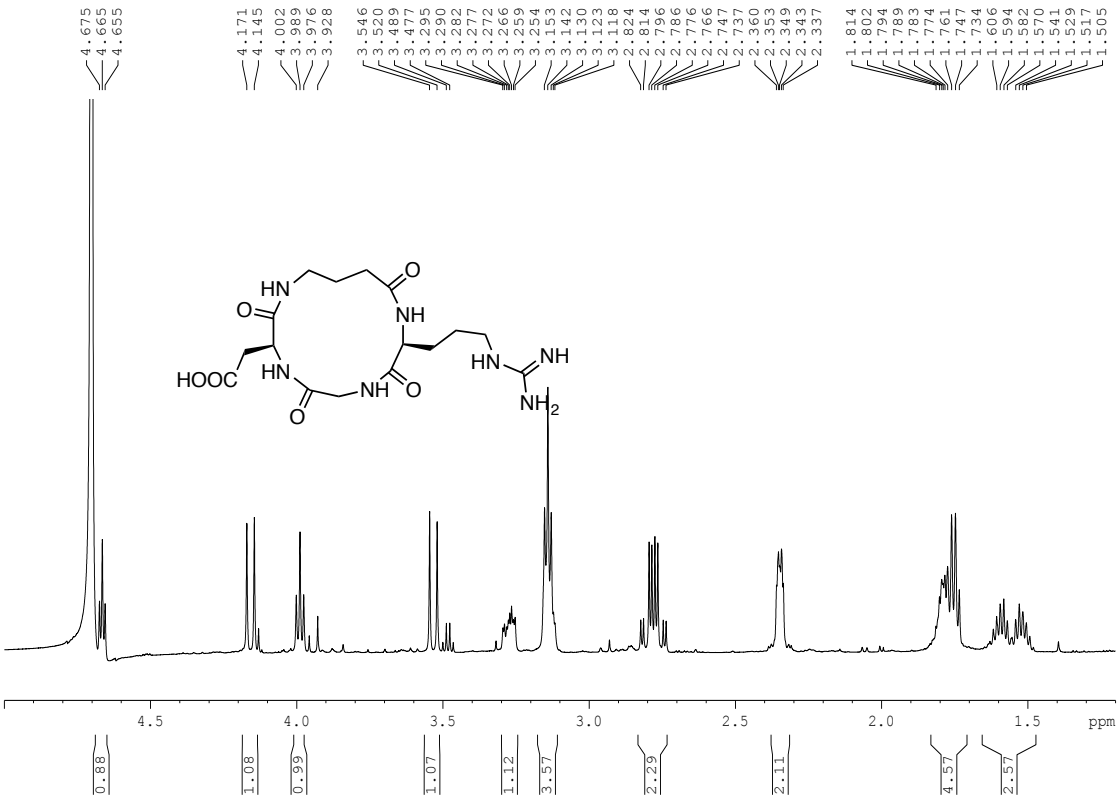
^{1}H-^{1}H COSY (400 MHz, DMSO-d_6) of **4a**



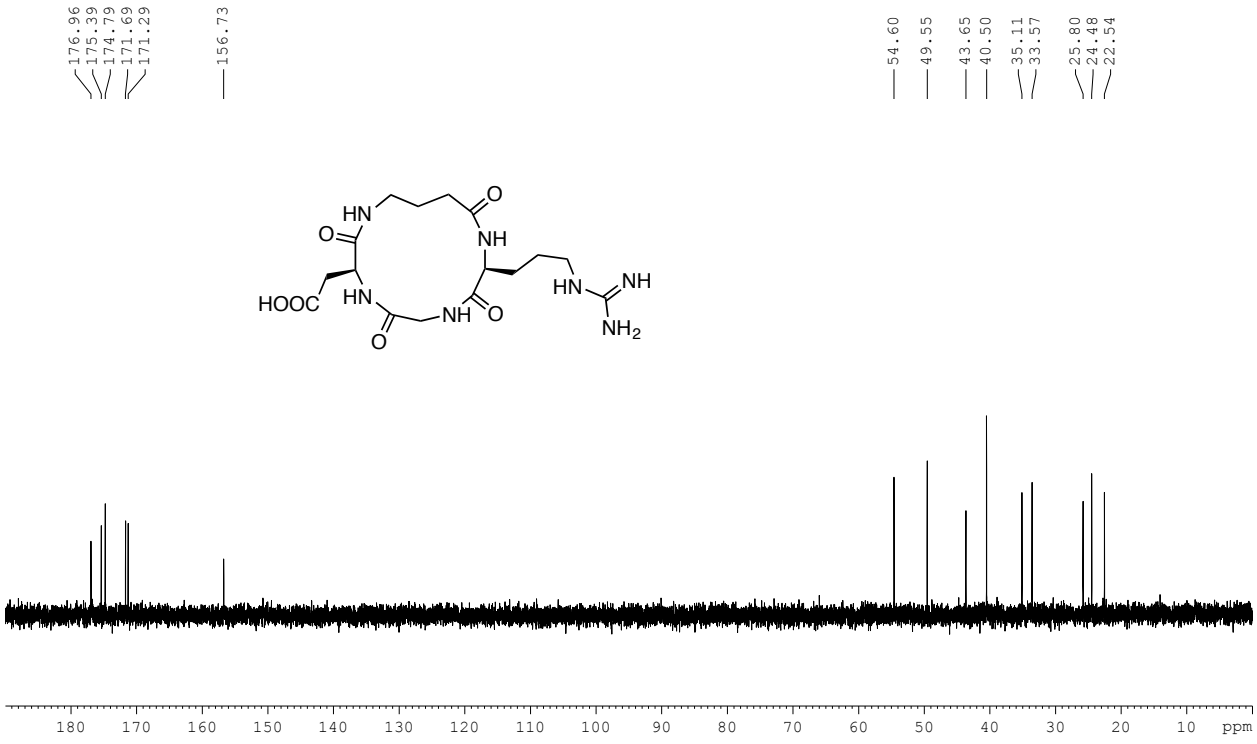
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **4a**



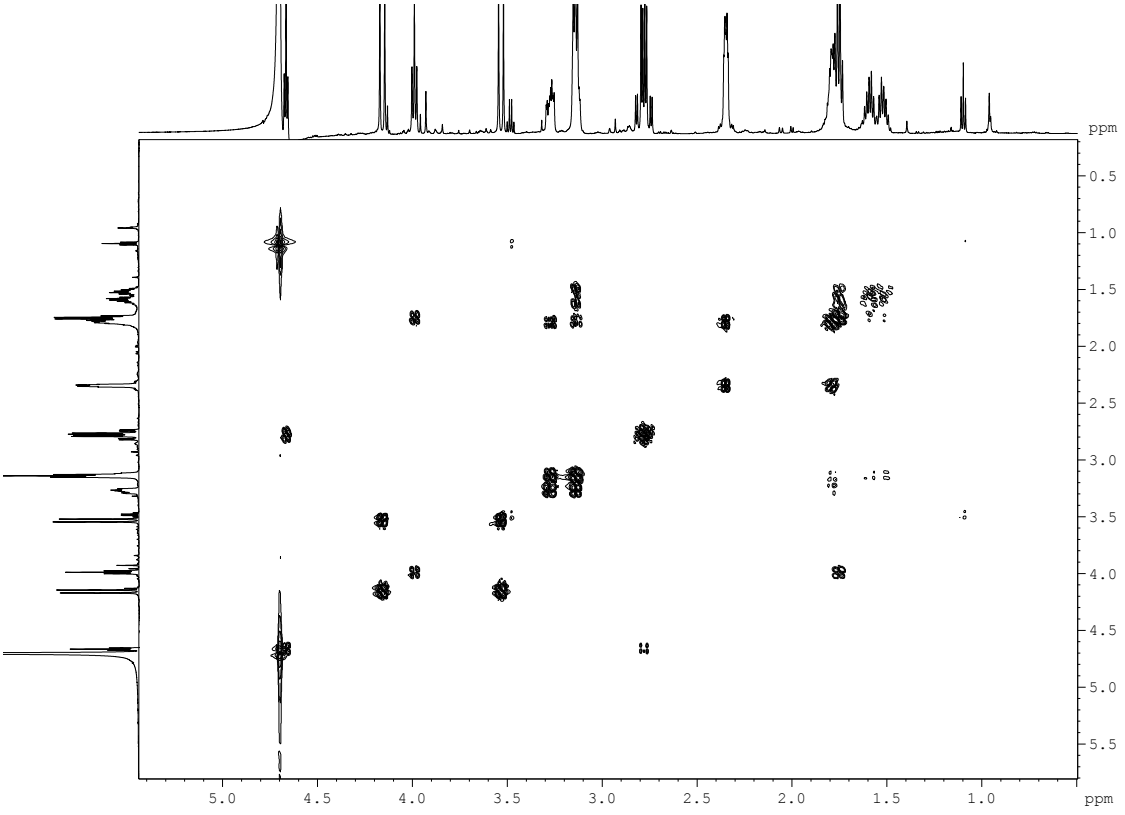
^{1}H NMR (600 MHz, D_2O) of **4a**



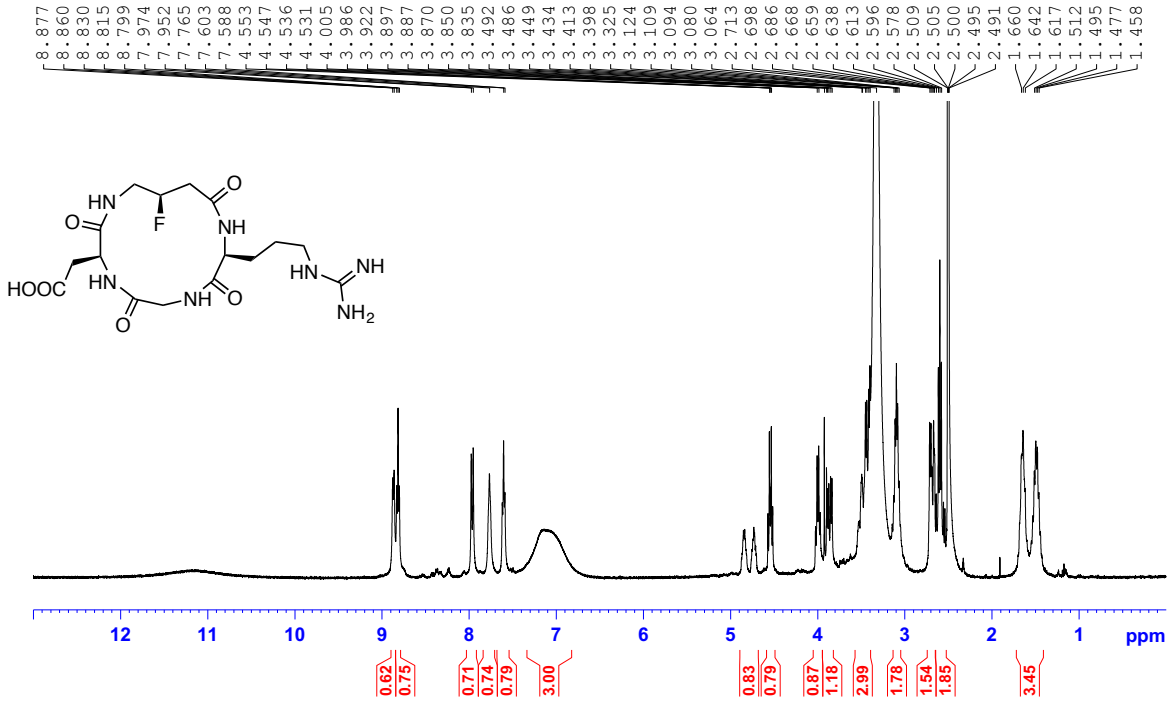
$^{13}C\{^{1}H\}$ NMR (600 MHz, D_2O) of **4a**



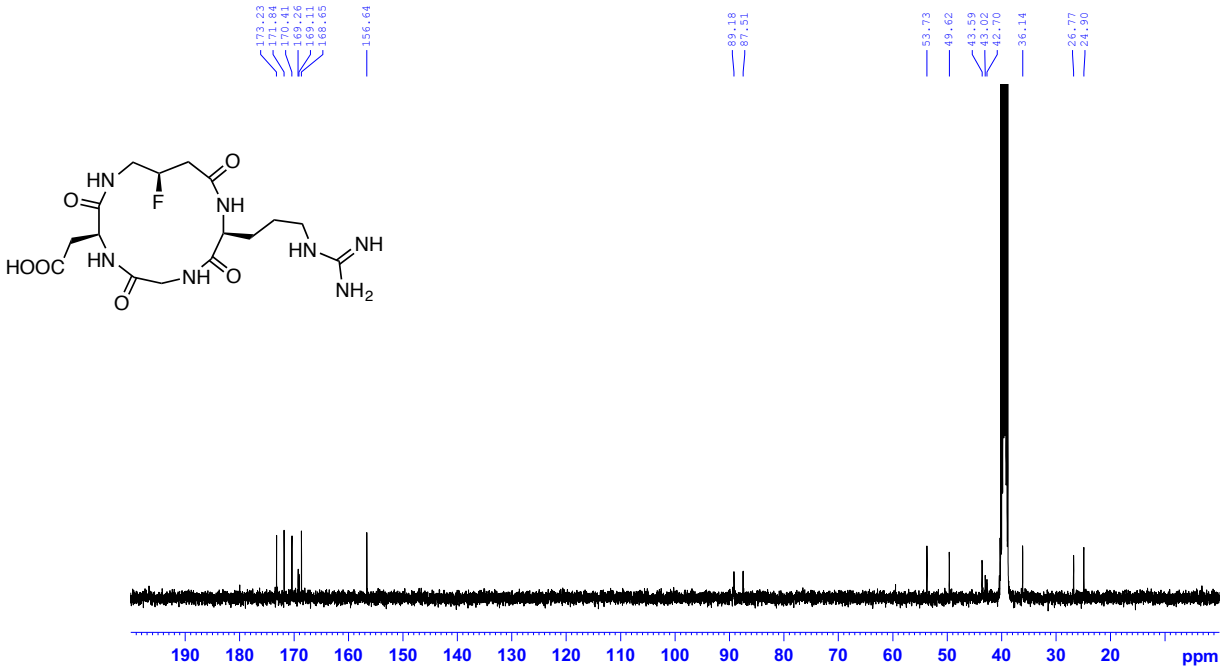
^{1}H-^{1}H COSY (600 MHz, D_2O) of **4a**



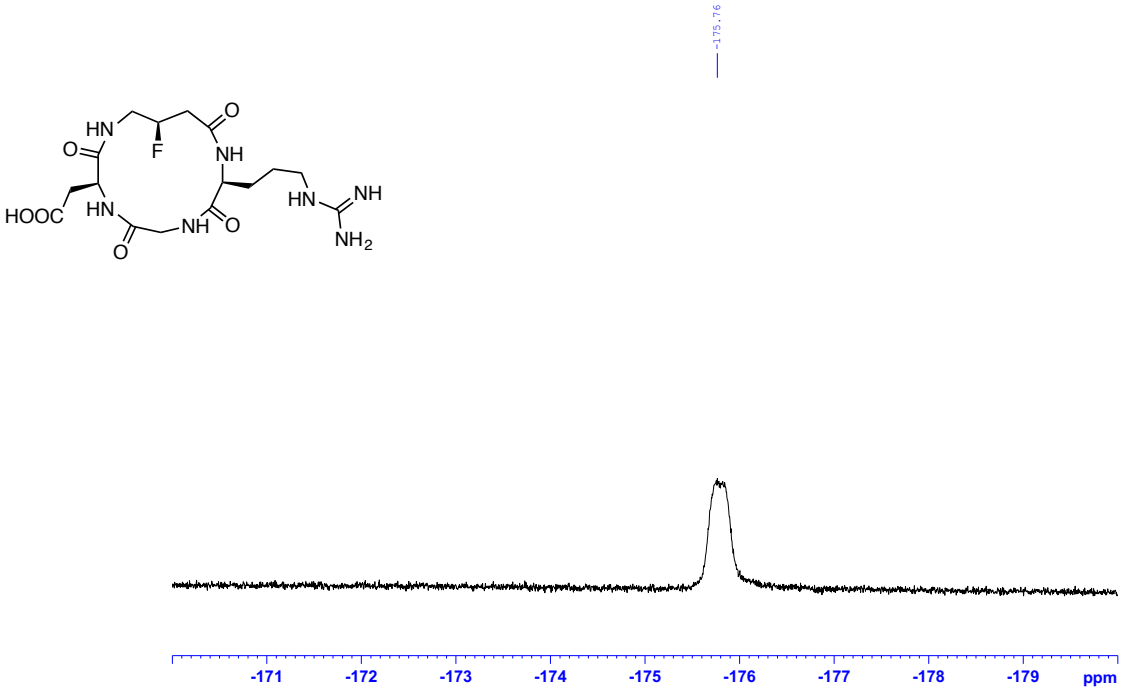
^{1}H NMR (400 MHz, DMSO-d$_6$) of **4b**



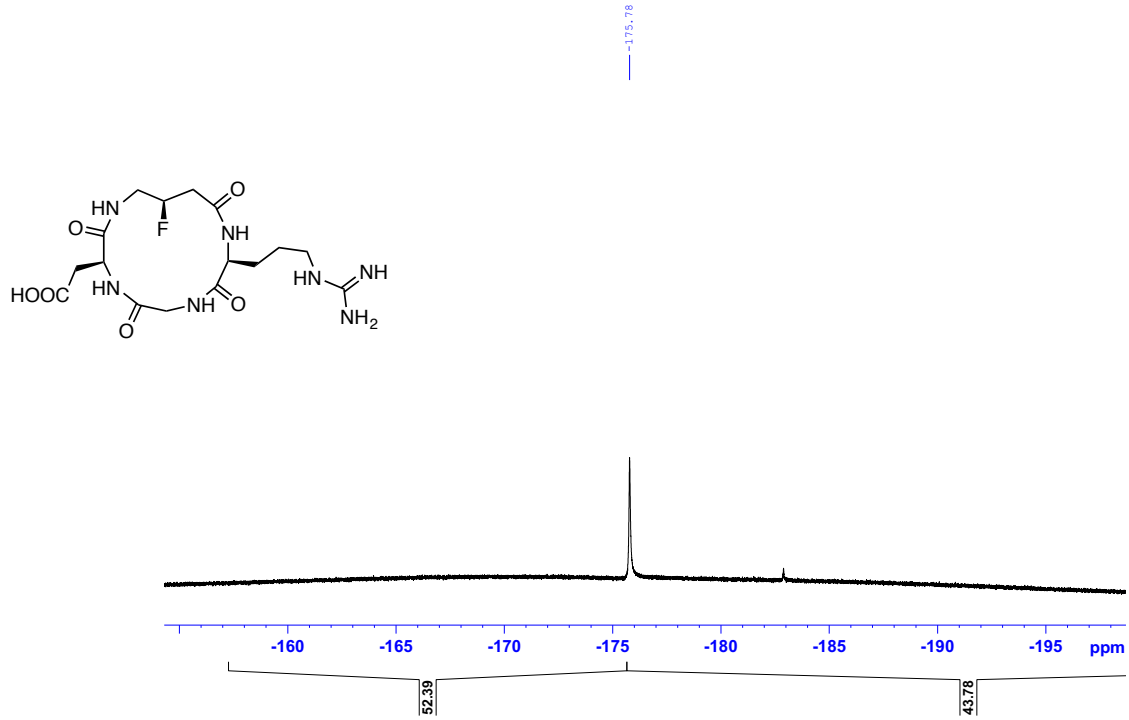
^{13}C{^{1}H} NMR (100 MHz, DMSO-d$_6$) of **4b**



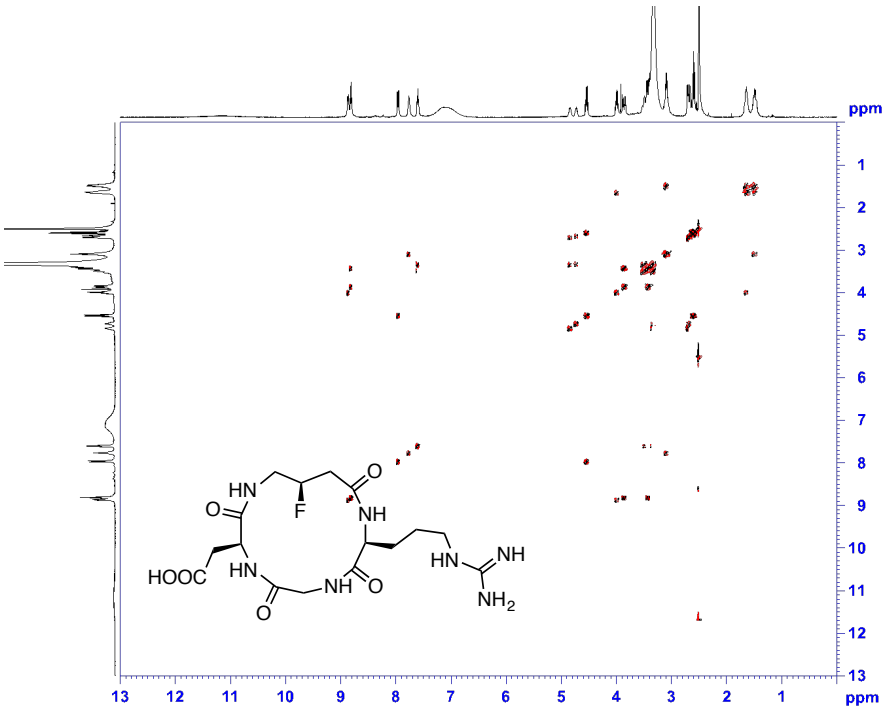
^{19}F NMR (376 MHz, DMSO-d_6) of **4b**



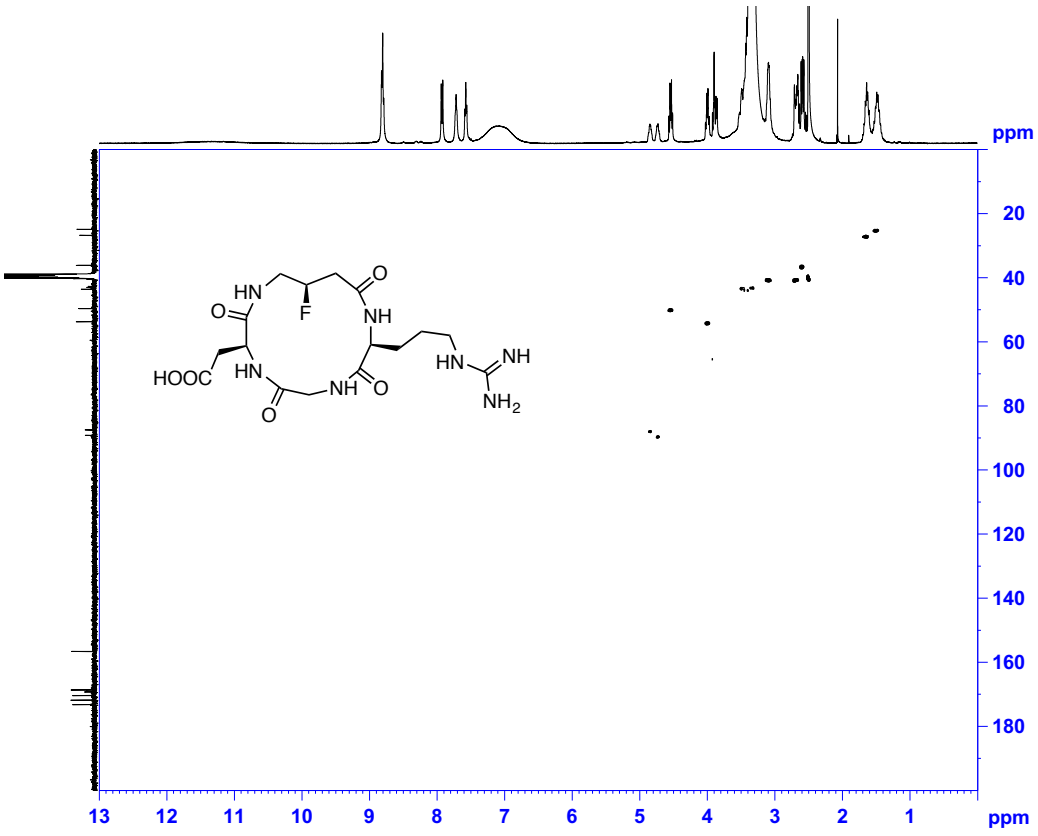
^{19}F{^{1}H} NMR (376 MHz, DMSO-d_6) of **4b**



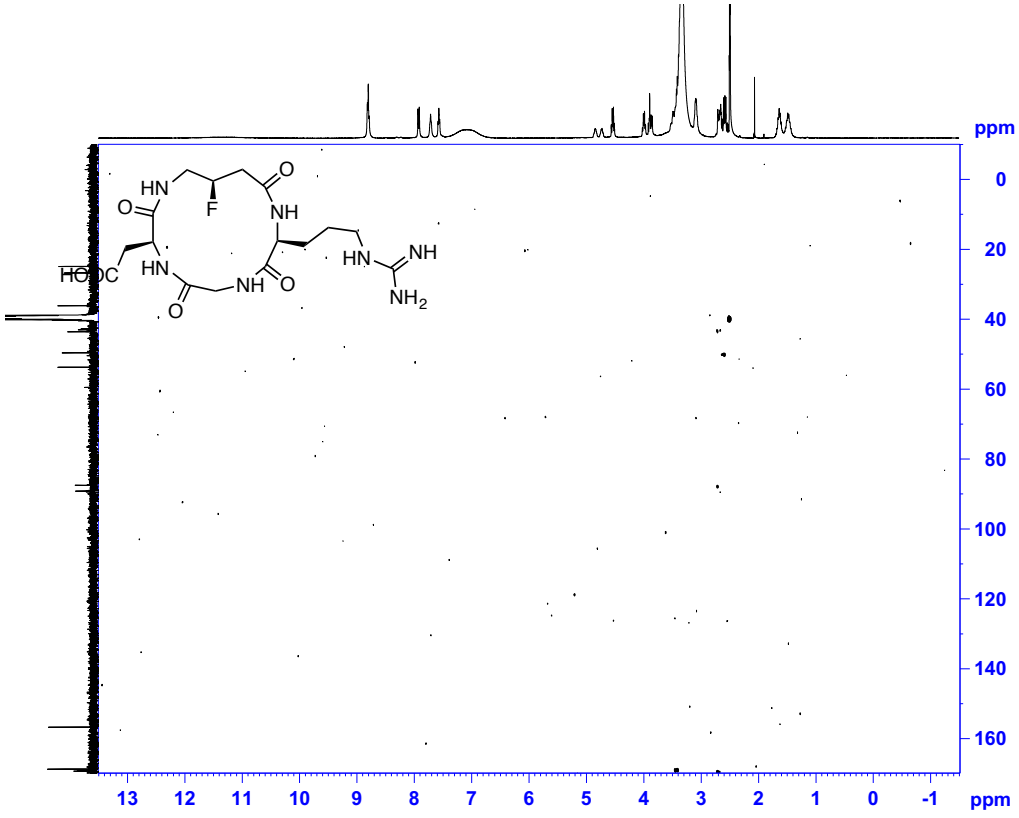
^{1}H-^{1}H COSY (400 MHz, DMSO-d_6) of **4b**



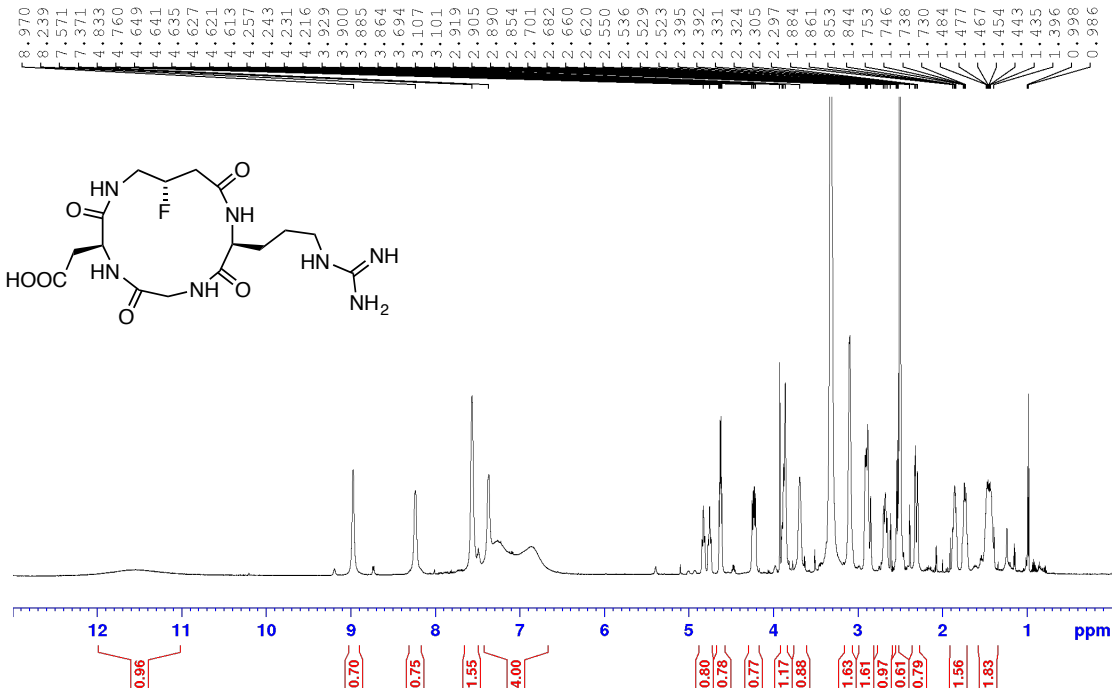
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **4b**



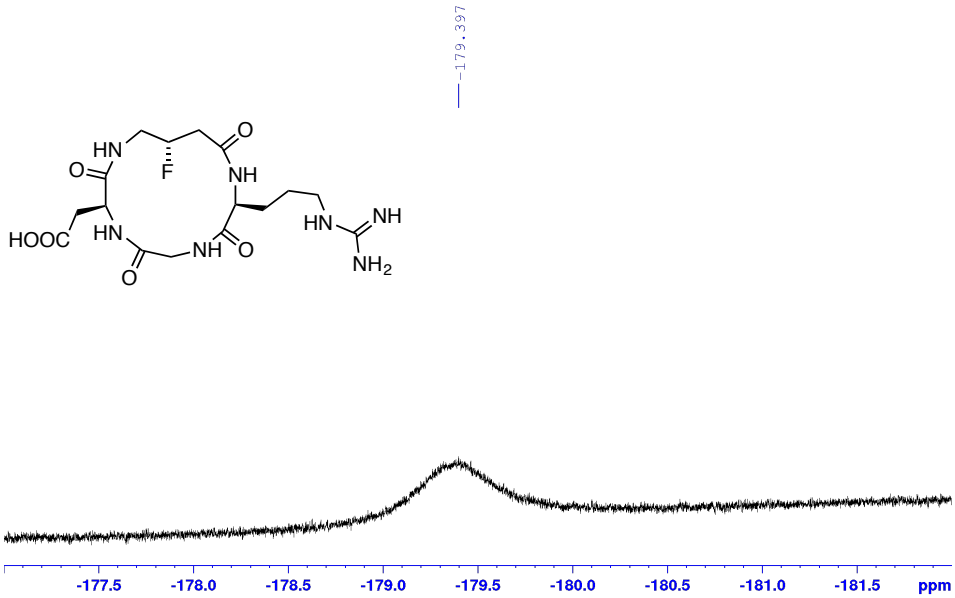
^{1}H-^{13}C HMBC (400MHz, DMSO-d_6) of **4b**



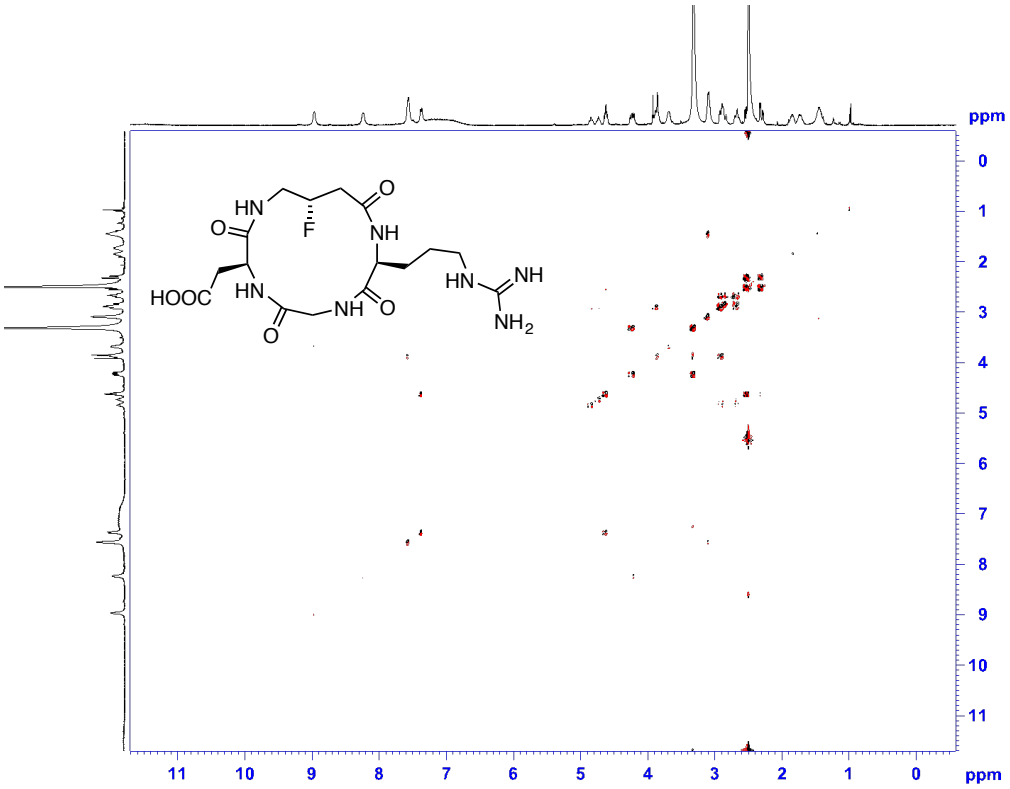
^{1}H NMR (400 MHz, DMSO-d_6) of **4c**



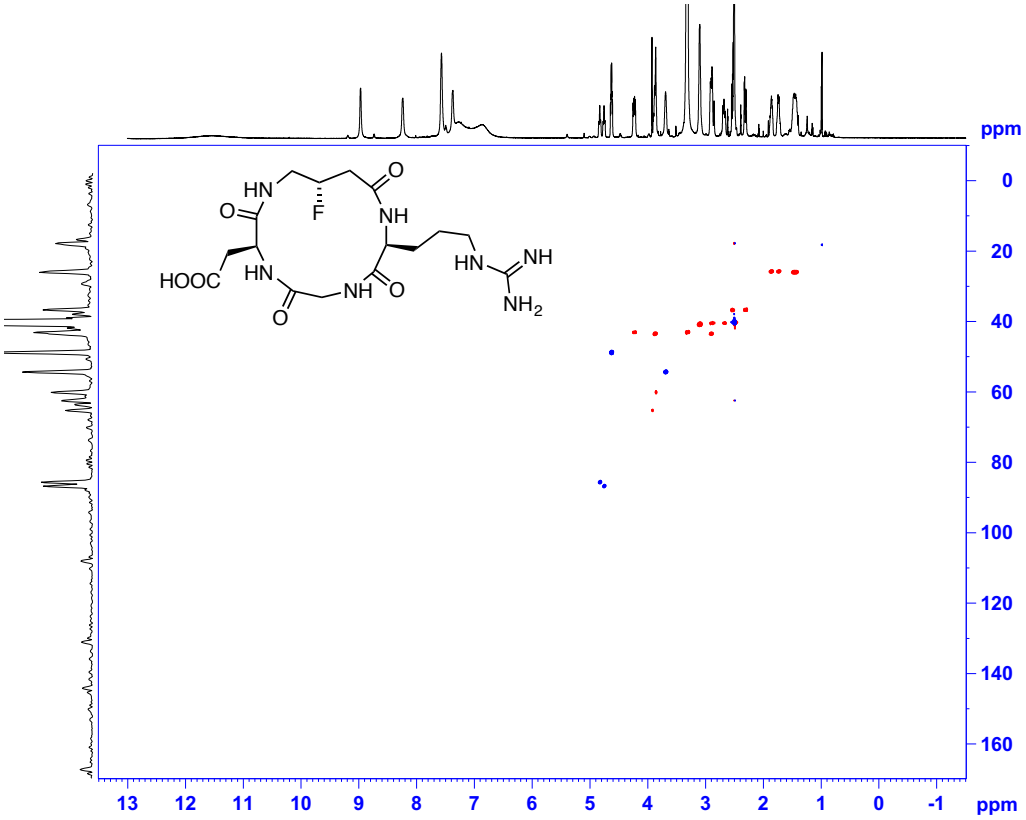
^{19}F NMR (376 MHz, DMSO-d$_6$) of **4c**



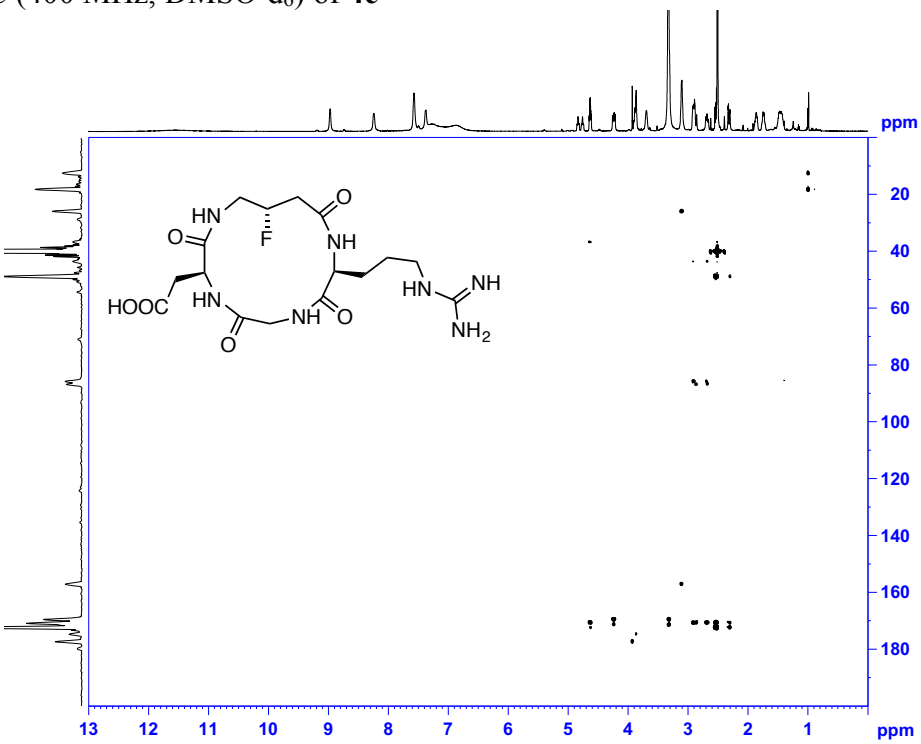
^{1}H-^{1}H COSY (400 MHz, DMSO-d$_6$) of **4c**



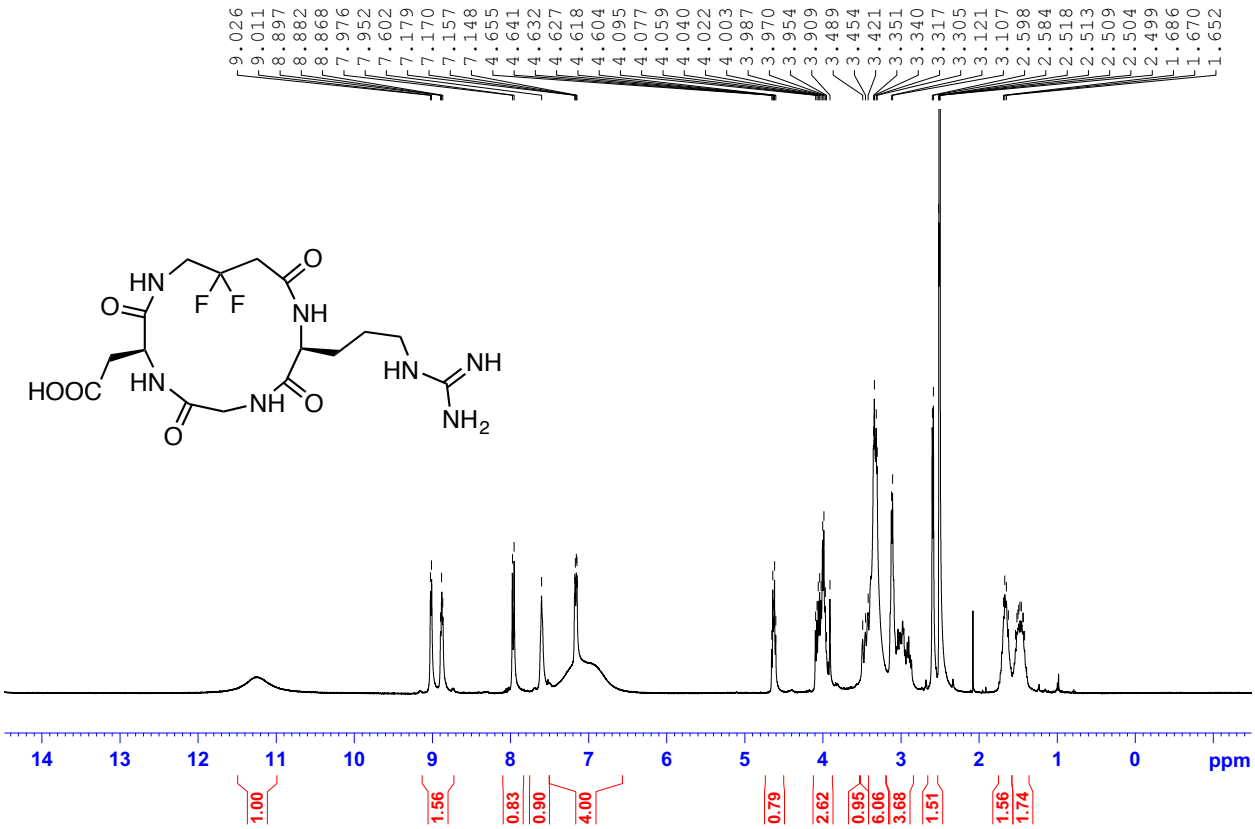
^{1}H-^{13}C HSQC (400 MHz, DMSO-d_6) of **4c**



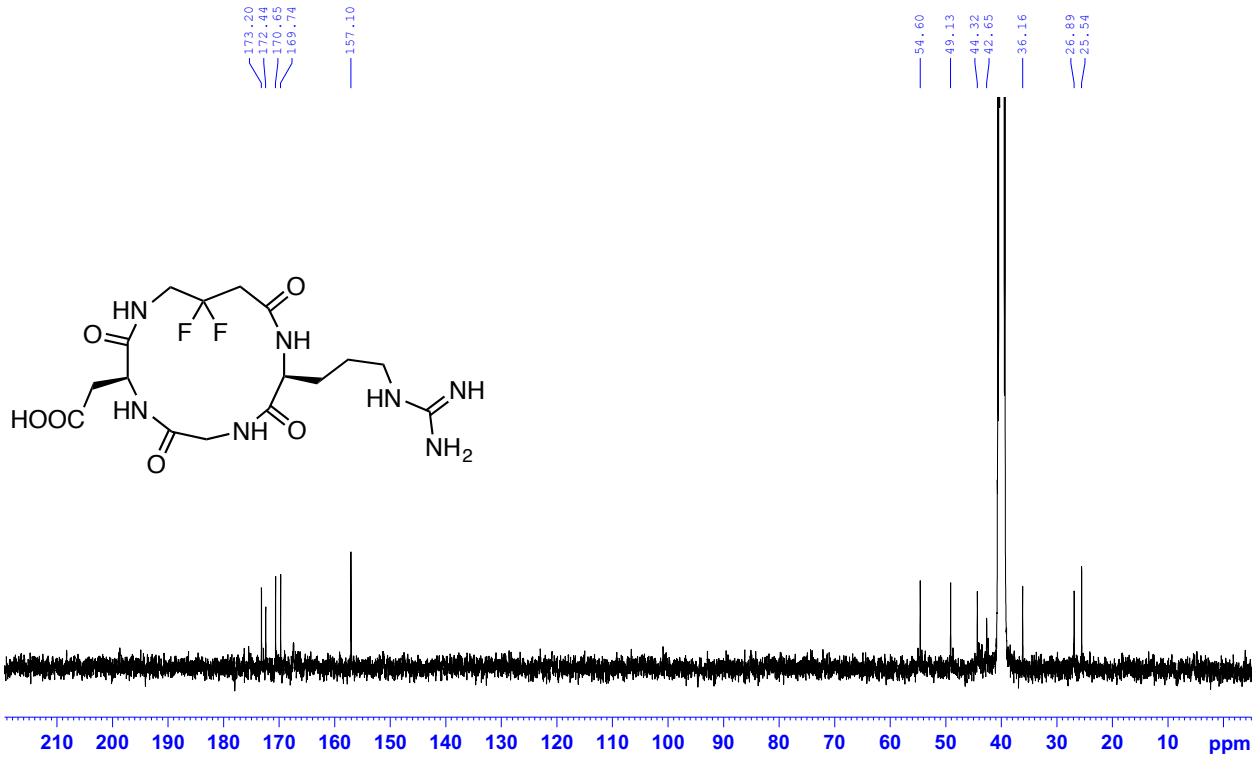
^{1}H-^{13}C HMBC (400 MHz, DMSO-d_6) of **4c**



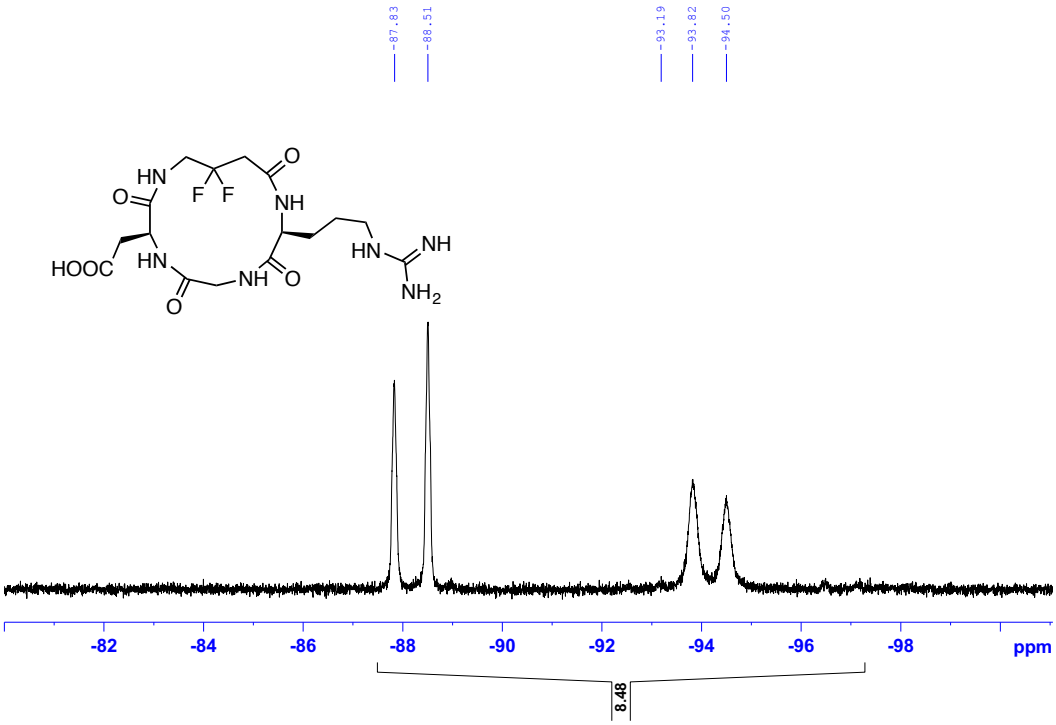
^{1}H NMR (400 MHz, DMSO-d$_6$) of **4d**



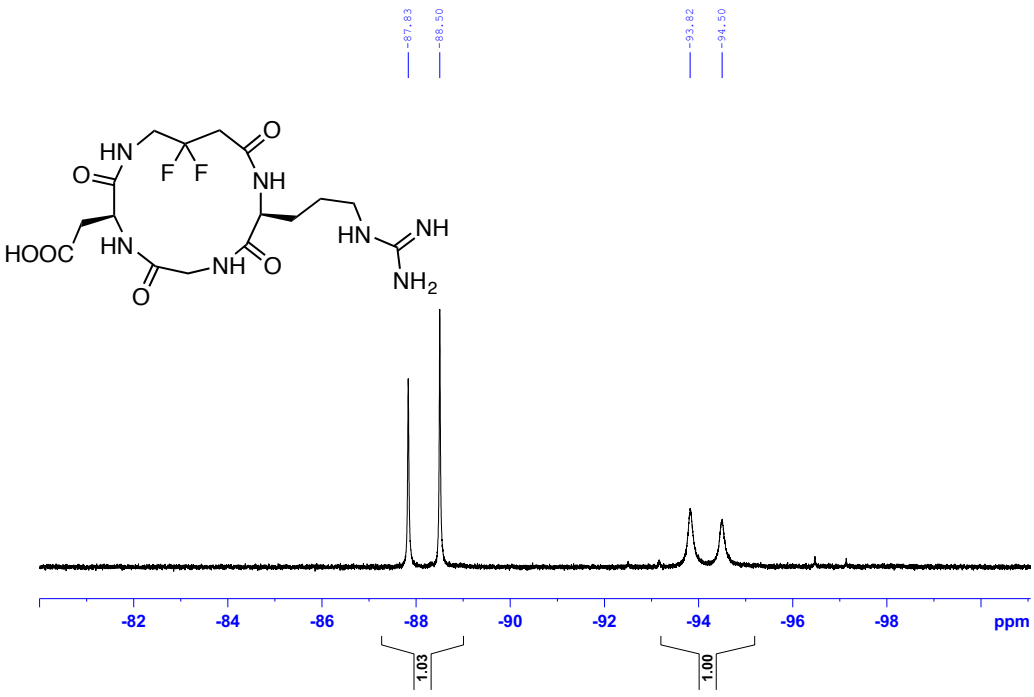
^{13}C{^{1}H} NMR (100 MHz, DMSO-d$_6$) of **4d**



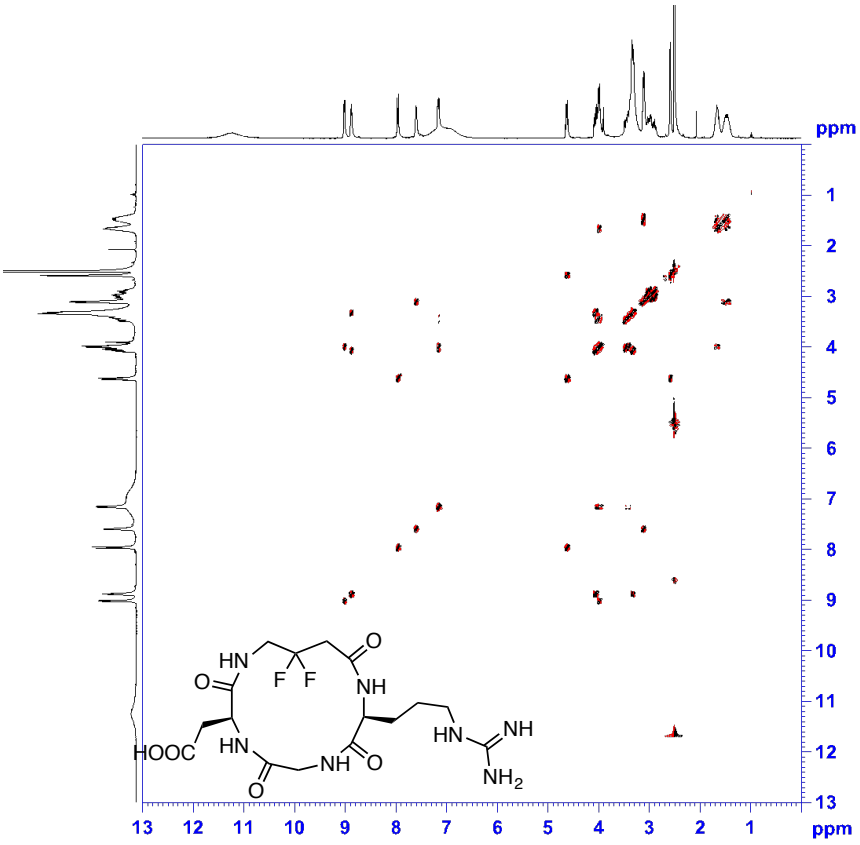
^{19}F NMR (376 MHz, DMSO-d$_6$) of **4d**



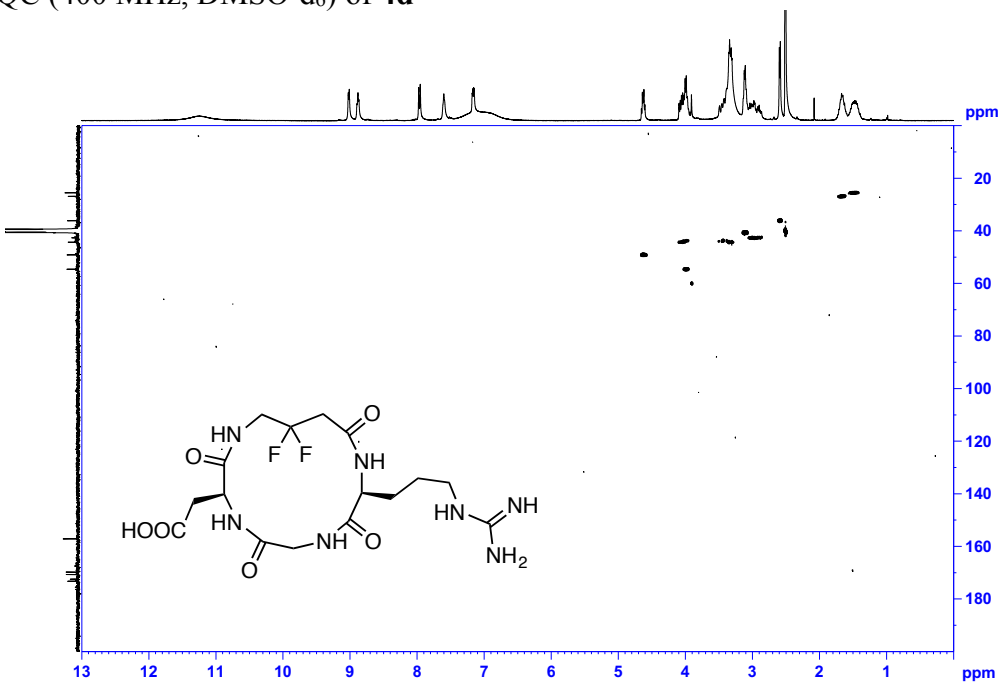
^{19}F{^{1}H} NMR (376 MHz, DMSO-d$_6$) of **4d**



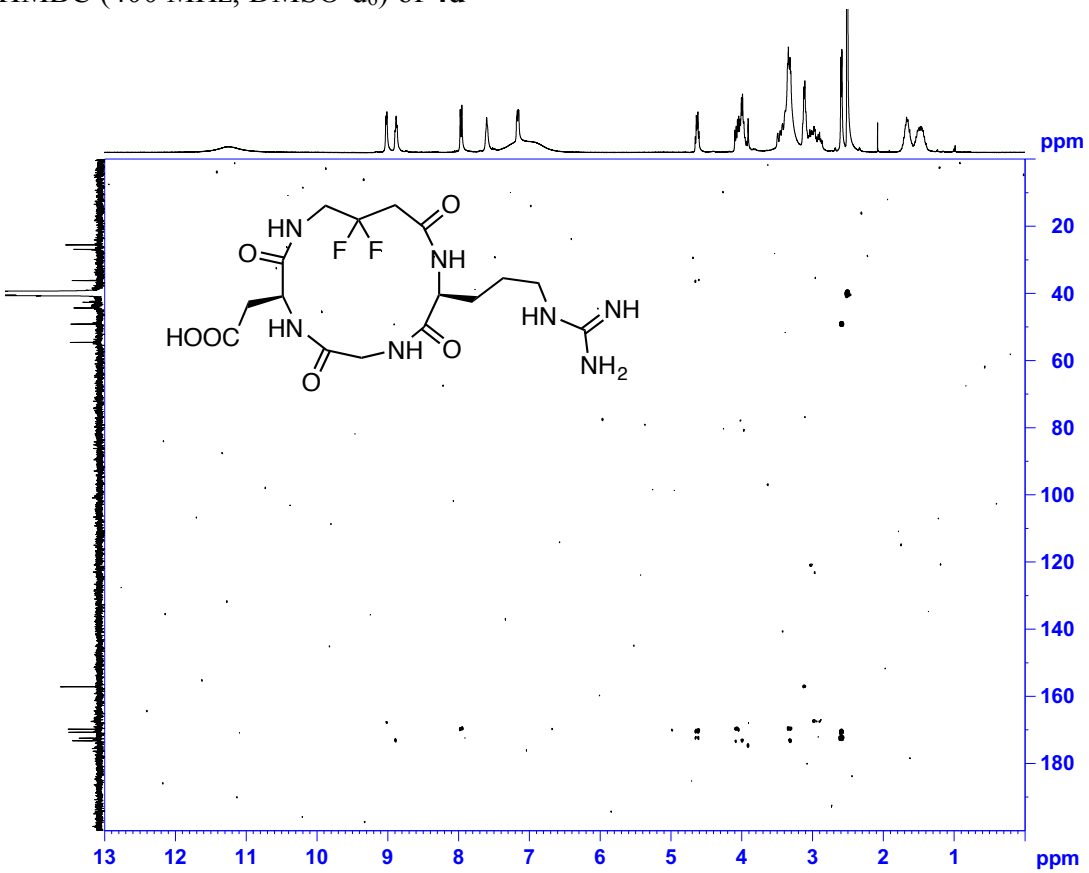
^{1}H-^{1}H COSY (400MHz, DMSO-d$_6$) of **4d**



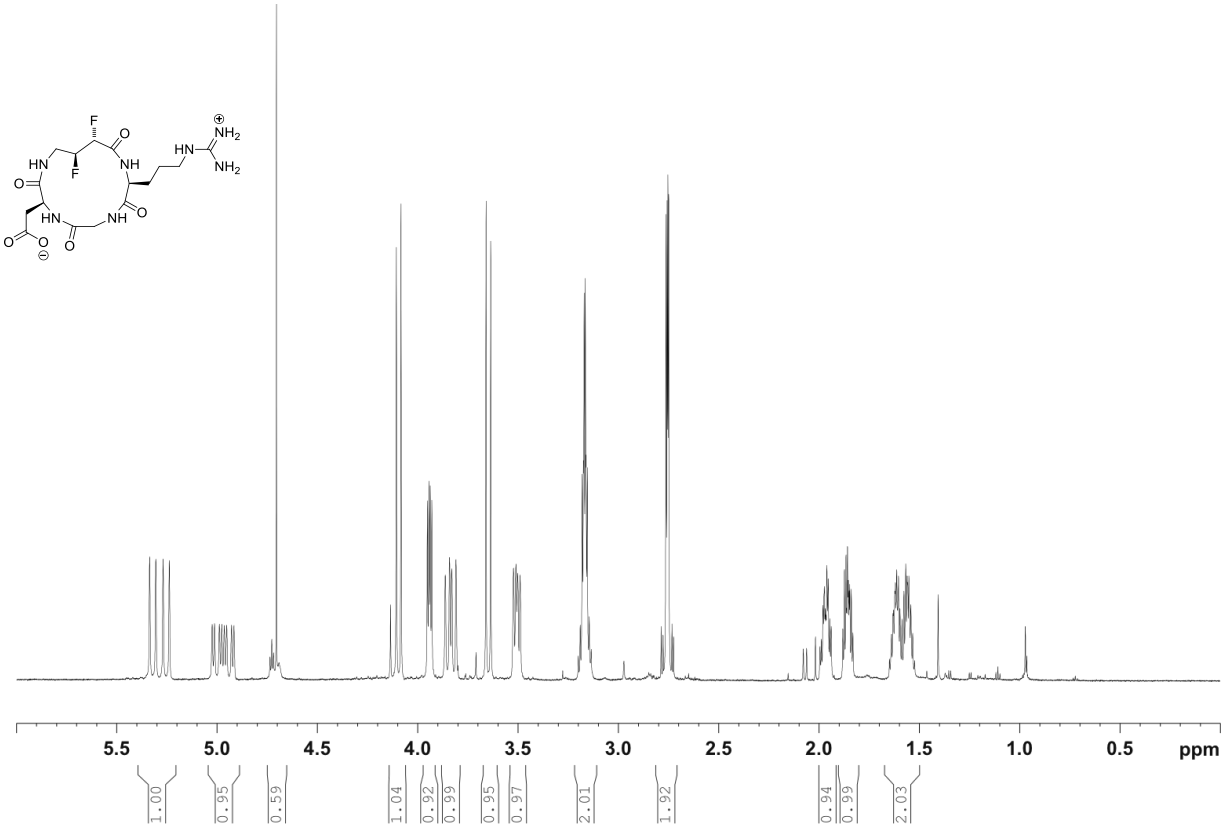
^{1}H-^{13}C HSQC (400 MHz, DMSO-d$_6$) of **4d**



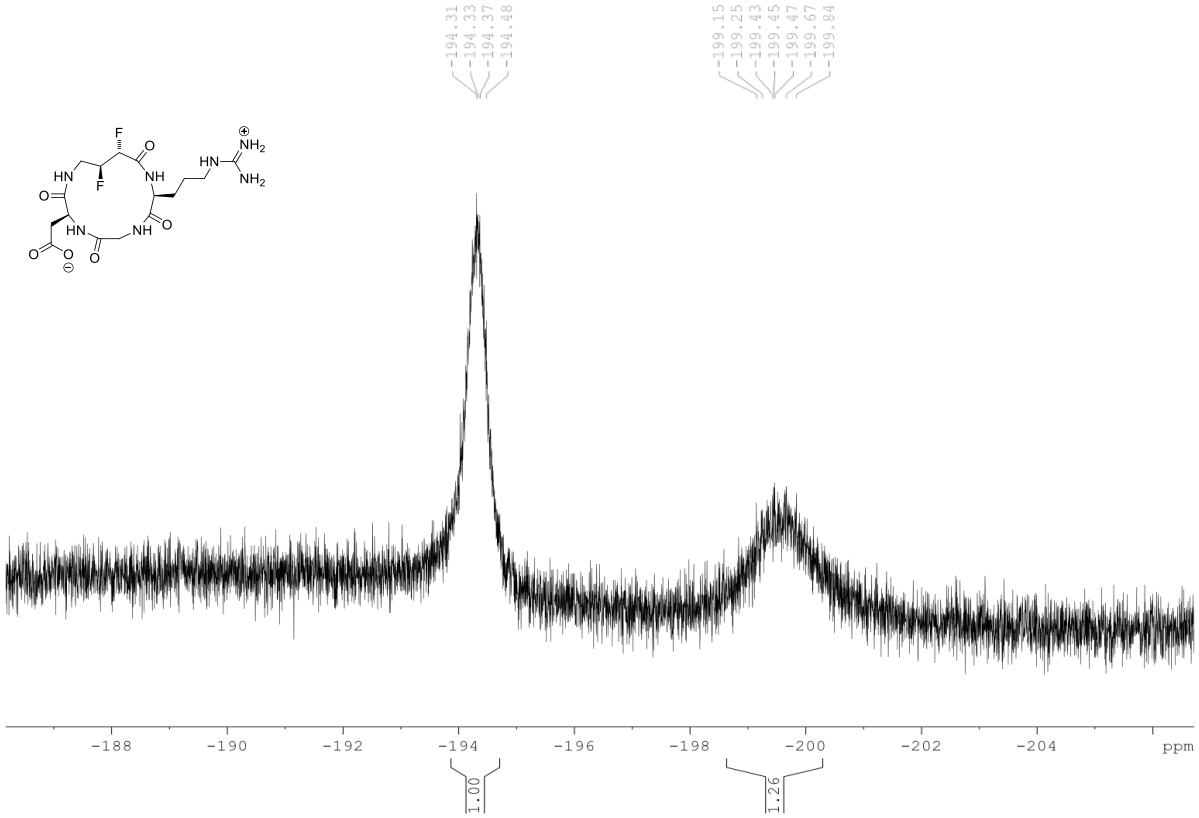
^{1}H-^{13}C HMBC (400 MHz, DMSO-d_6) of **4d**



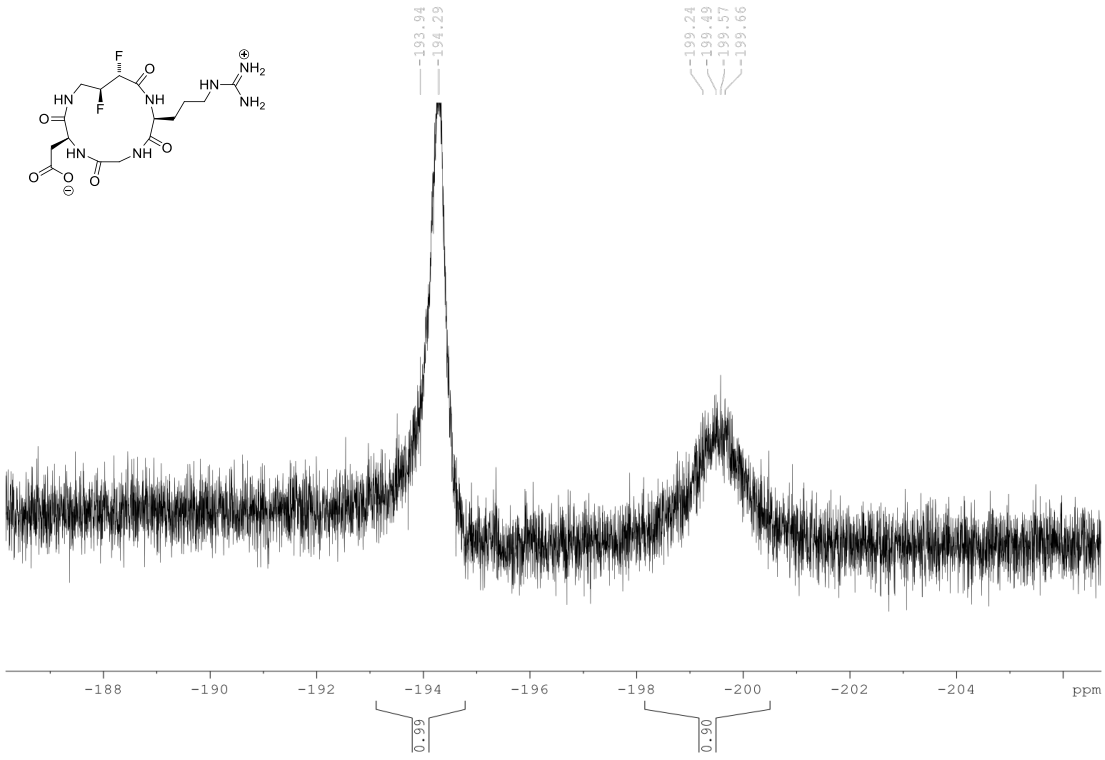
^{1}H NMR (700 MHz, D_2O) of **4e**



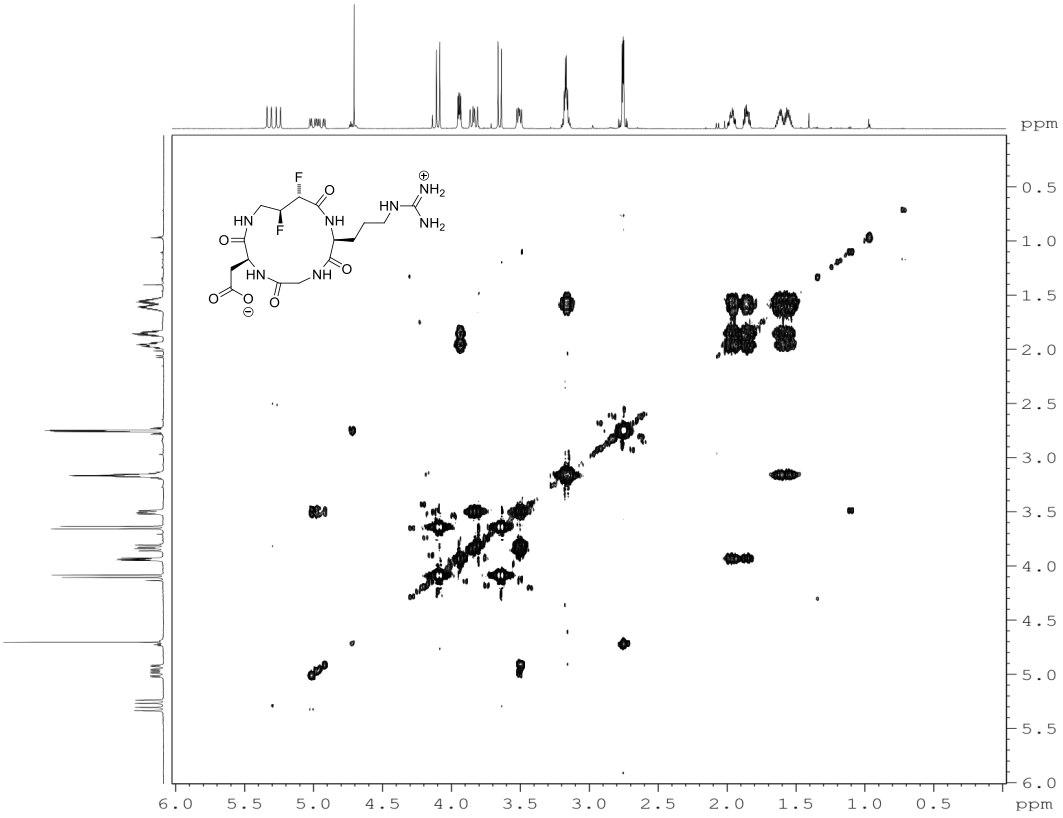
^{19}F NMR (376 MHz, D_2O) of **4e**



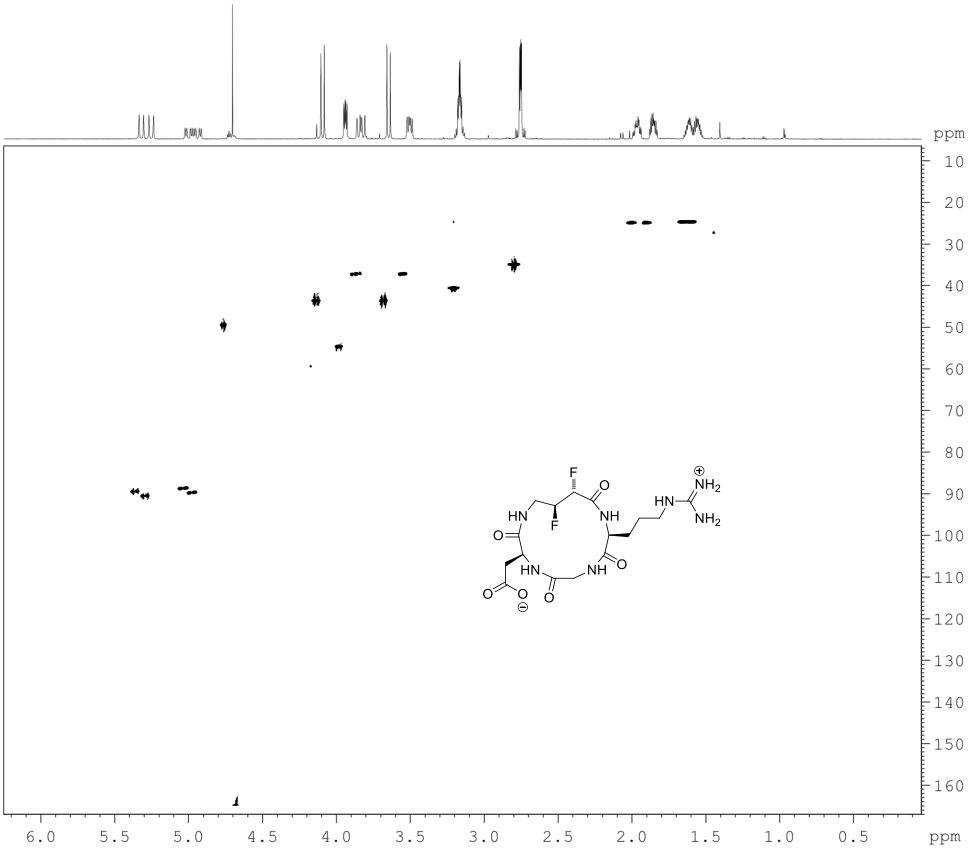
$^{19}F\{^1H\}$ NMR (376 MHz, D_2O) of **4e**



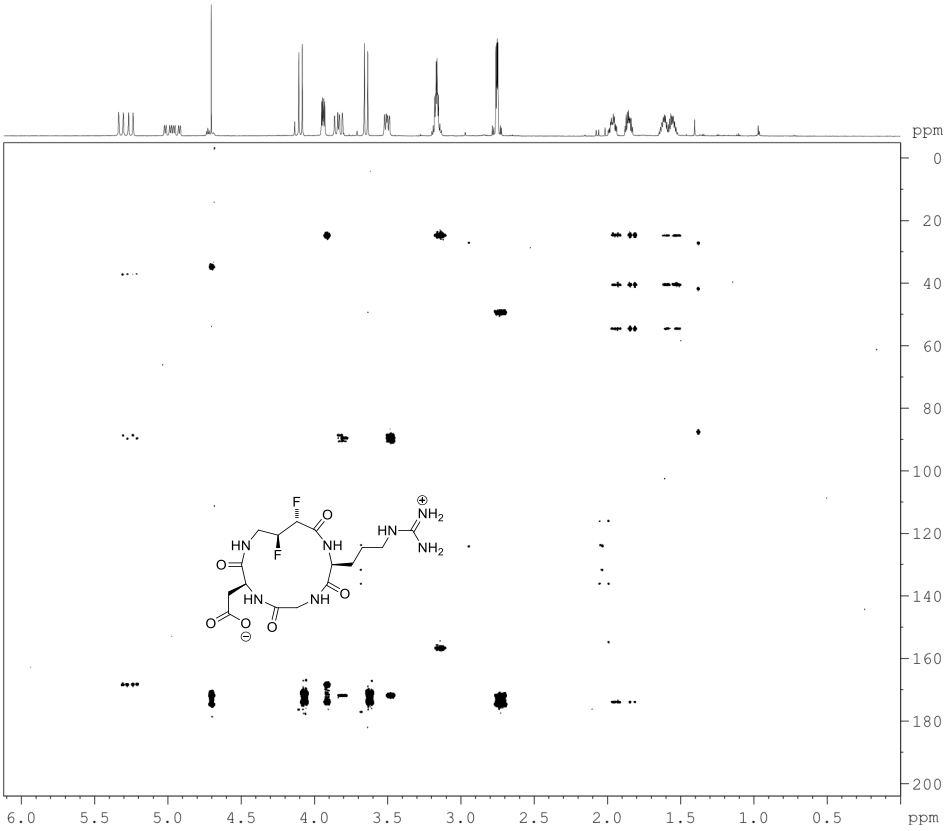
^{1}H-^{1}H COSY (700 MHz, D_2O) of **4e**



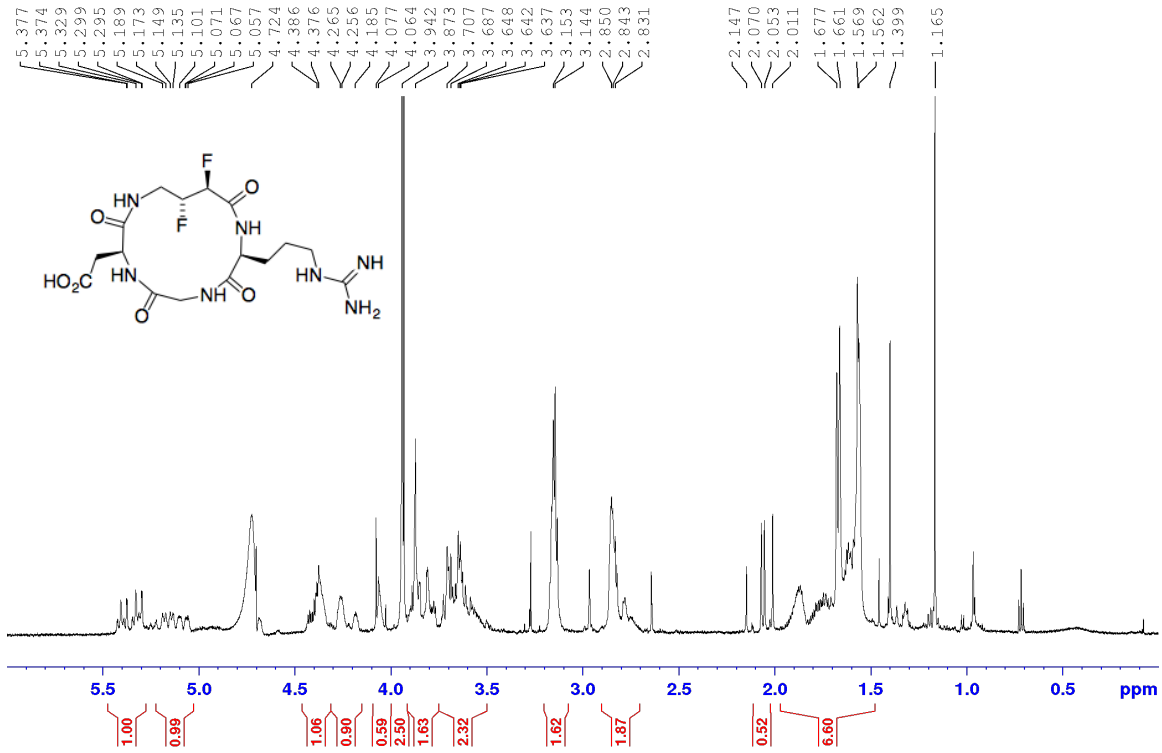
^{1}H-^{13}C HSQC (700 MHz, D_2O) of **4e**



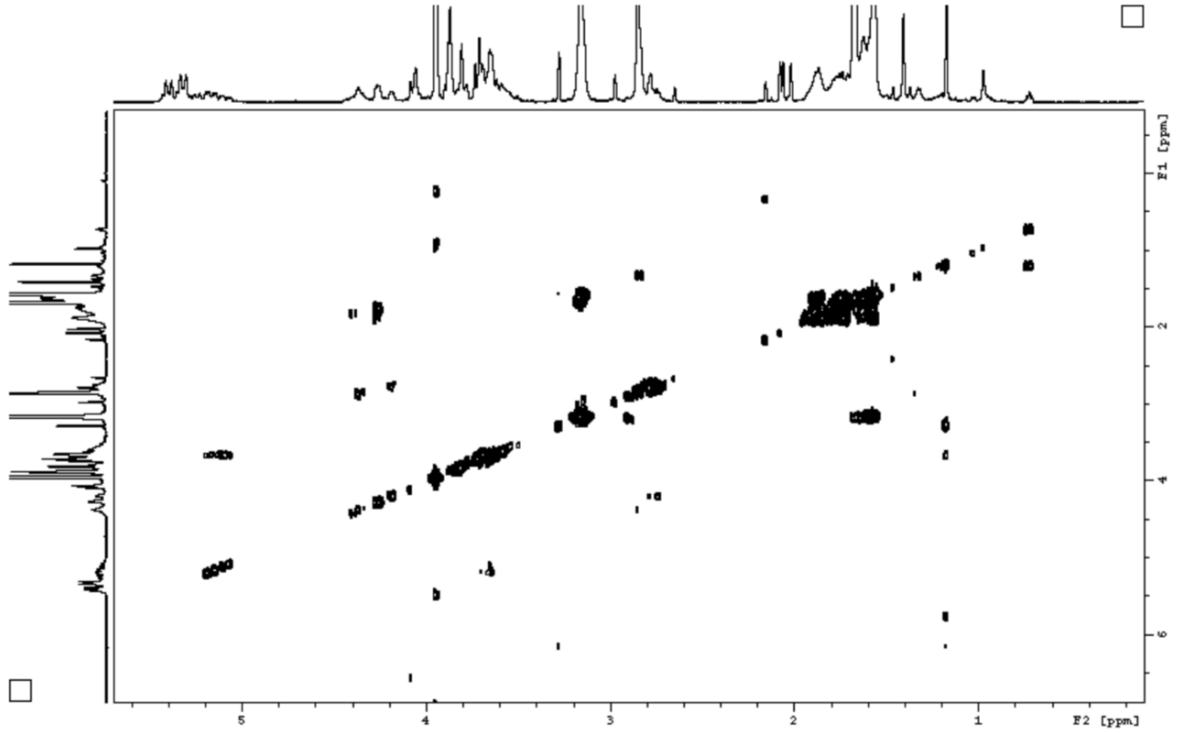
^{1}H-^{13}C HMBC (700 MHz, D_2O) of **4e**



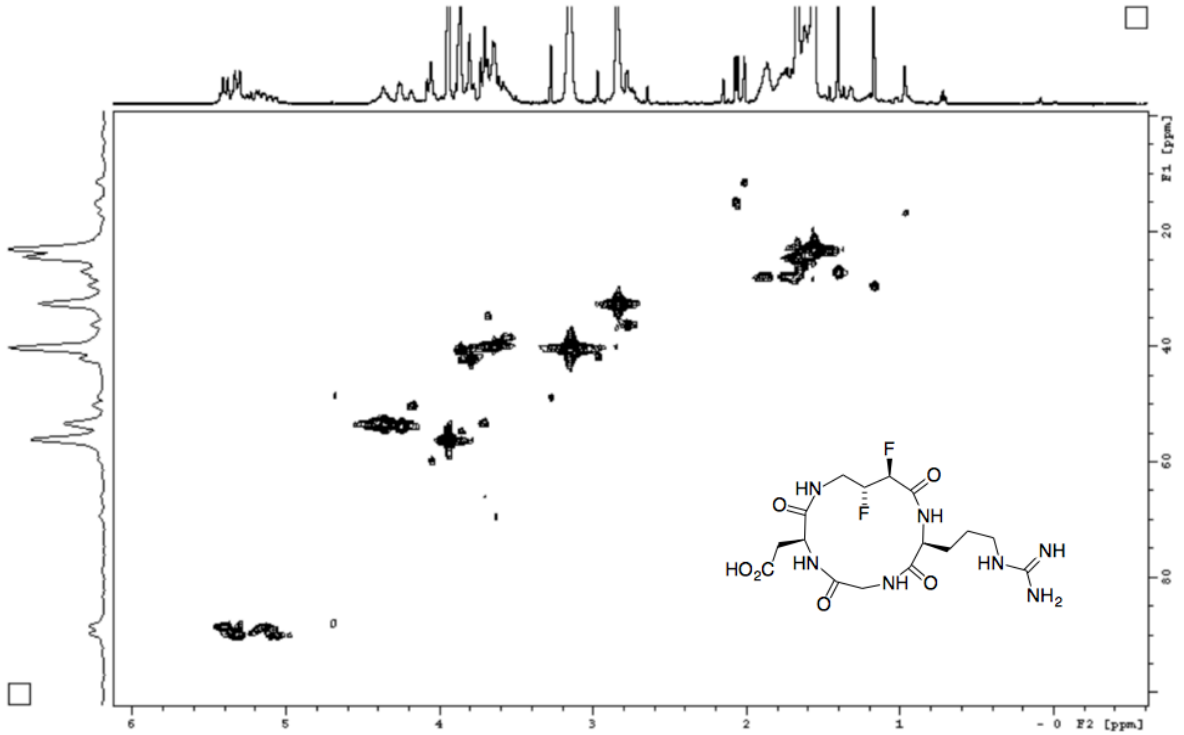
^{1}H NMR (700 MHz, D_2O) of **4f**



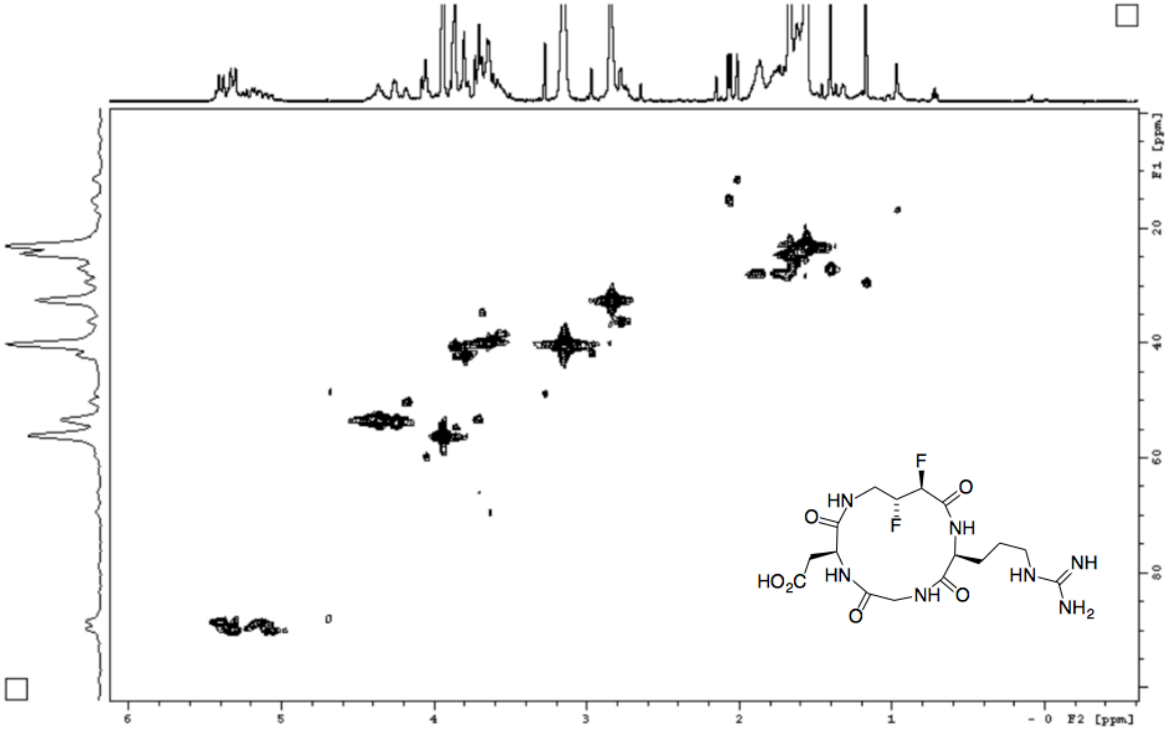
^{1}H-^{1}H COSY (600 MHz, D_2O) of **4f**



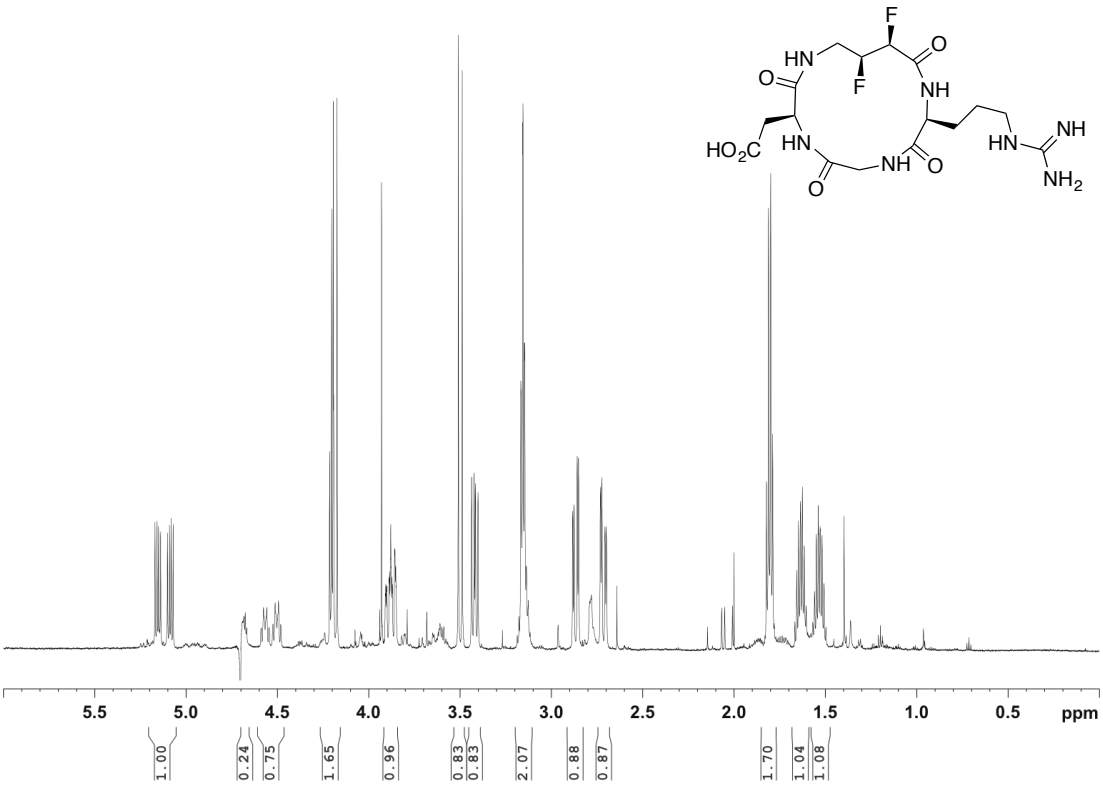
^{1}H-^{13}C HSQC (600 MHz, D_2O) of **4f**



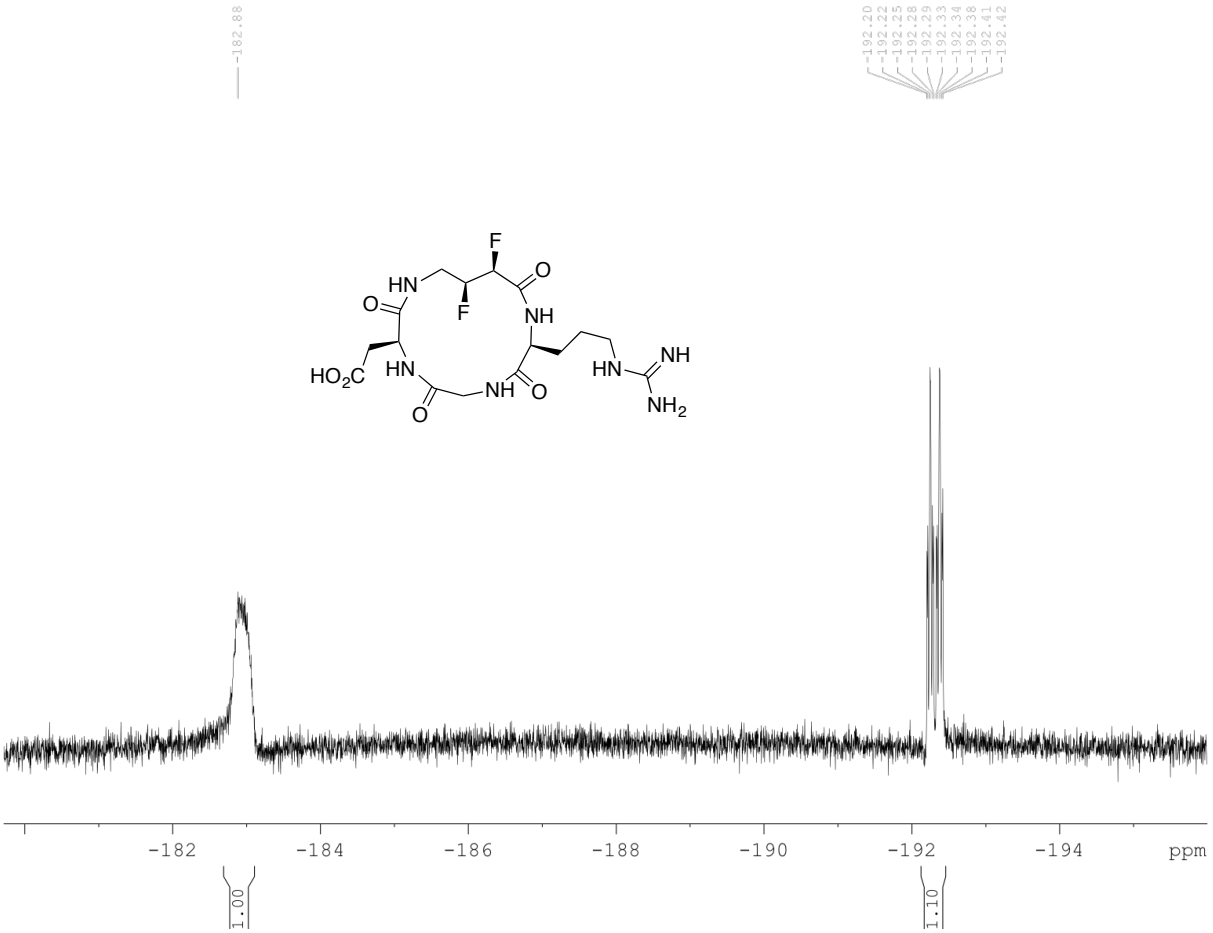
^{1}H-^{13}C HMBC (600 MHz, D_2O) of **4f**



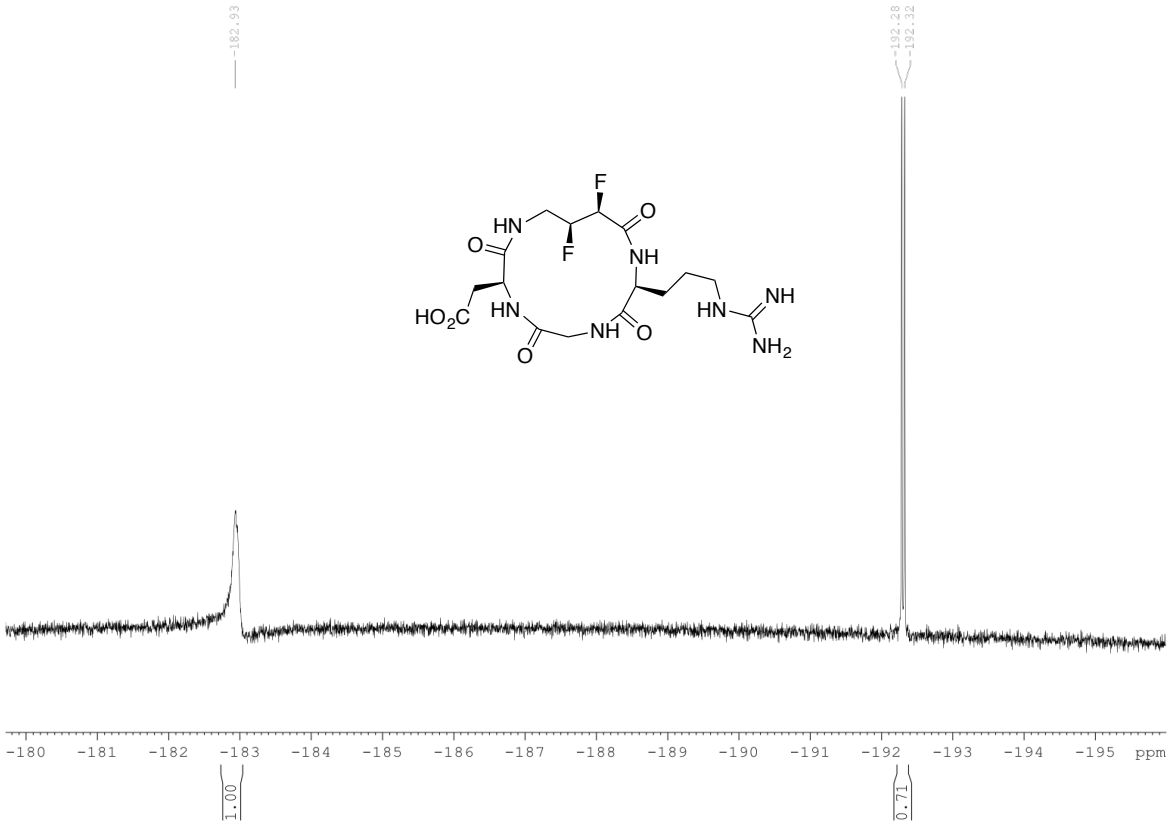
^{1}H NMR (700 MHz, D_2O) of **4g**



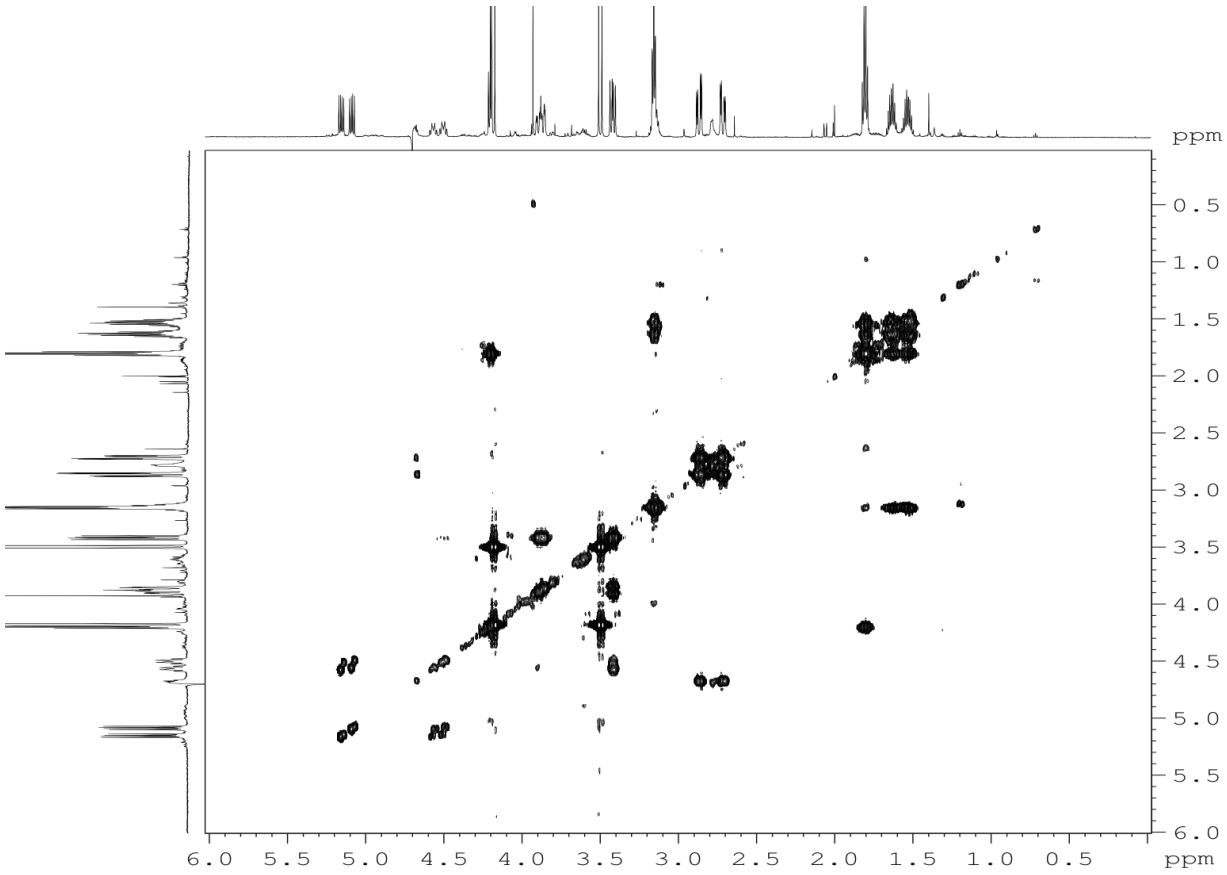
^{19}F NMR (376 MHz, D_2O of **4g**



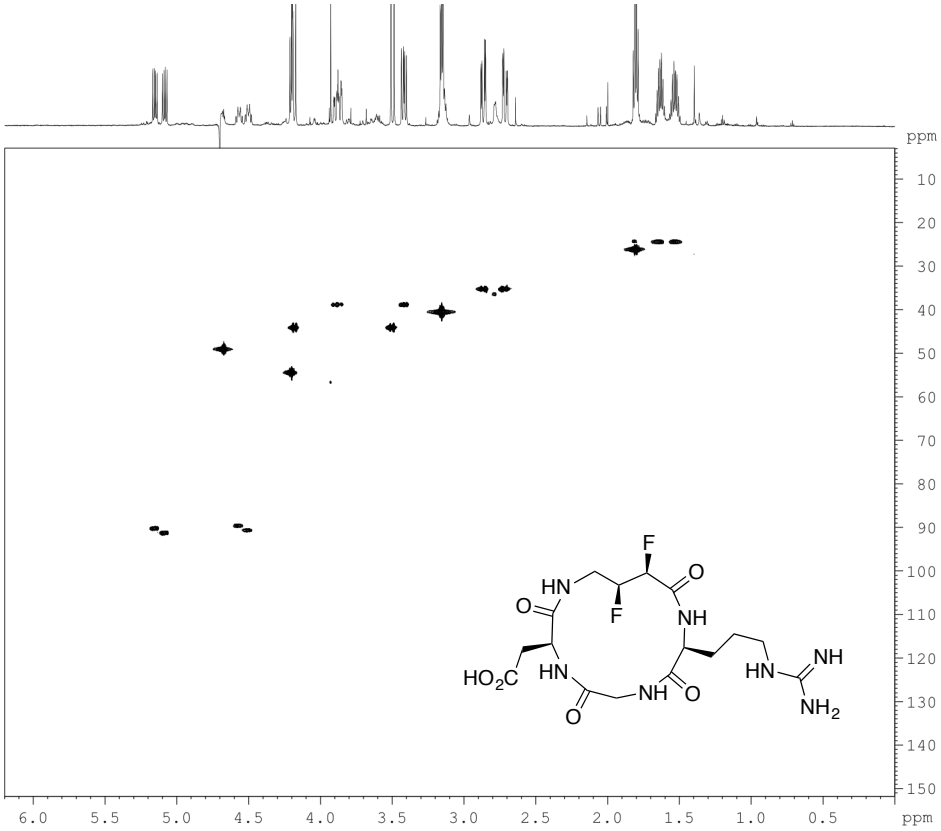
$^{19}F\{^{1}H\}$ NMR (376 MHz, D_2O) of **4g**



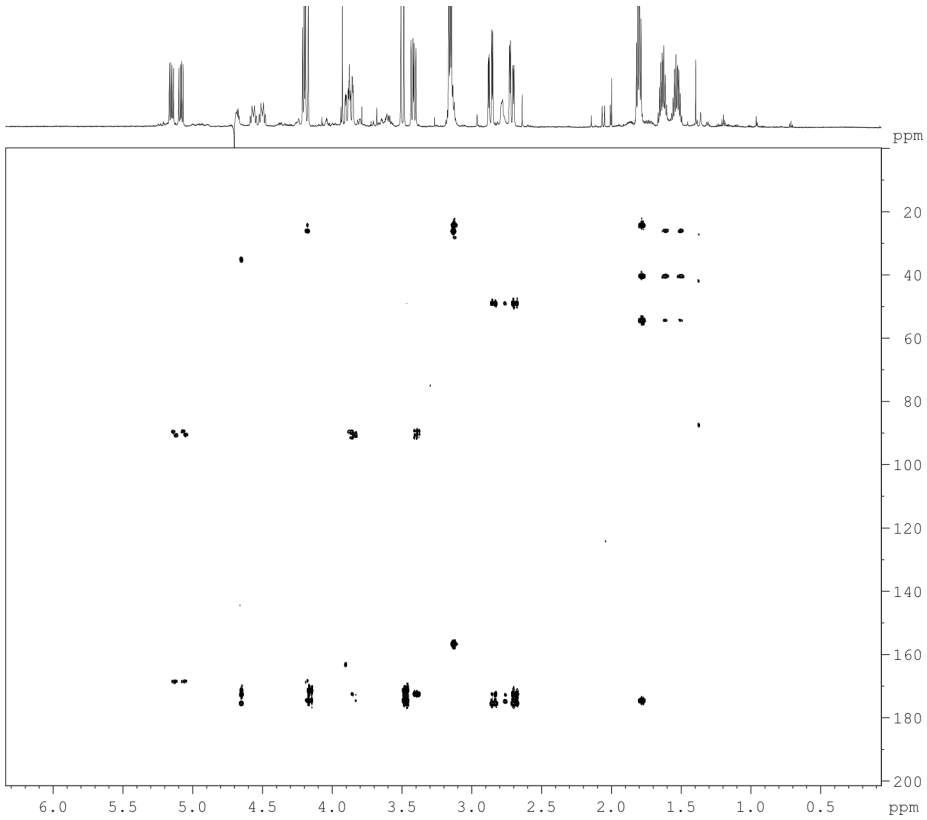
^{1}H-^{1}H COSY (700 MHz, D_2O) of **4g**



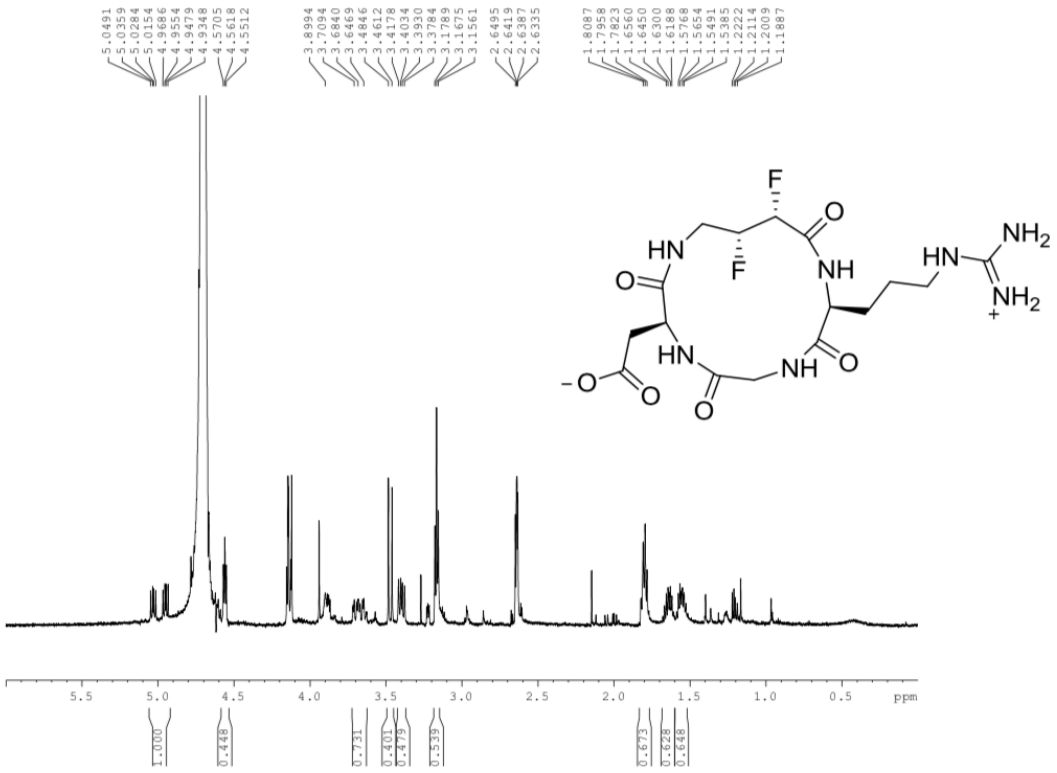
^{1}H-^{13}C HSQC (700 MHz, D_2O) of **4g**



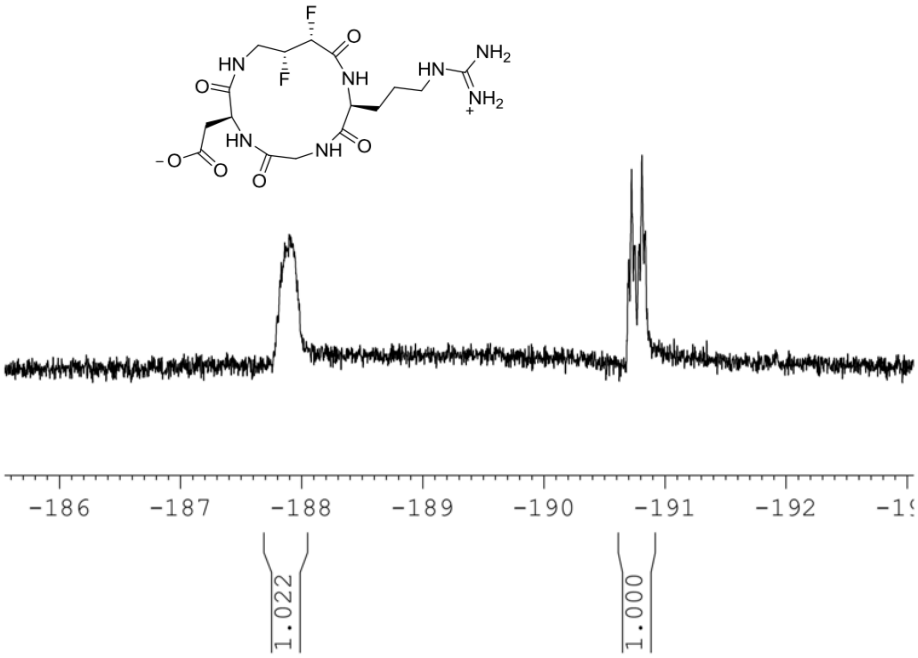
^{1}H-^{13}C HMBC (700 MHz, D_2O) of **4g**



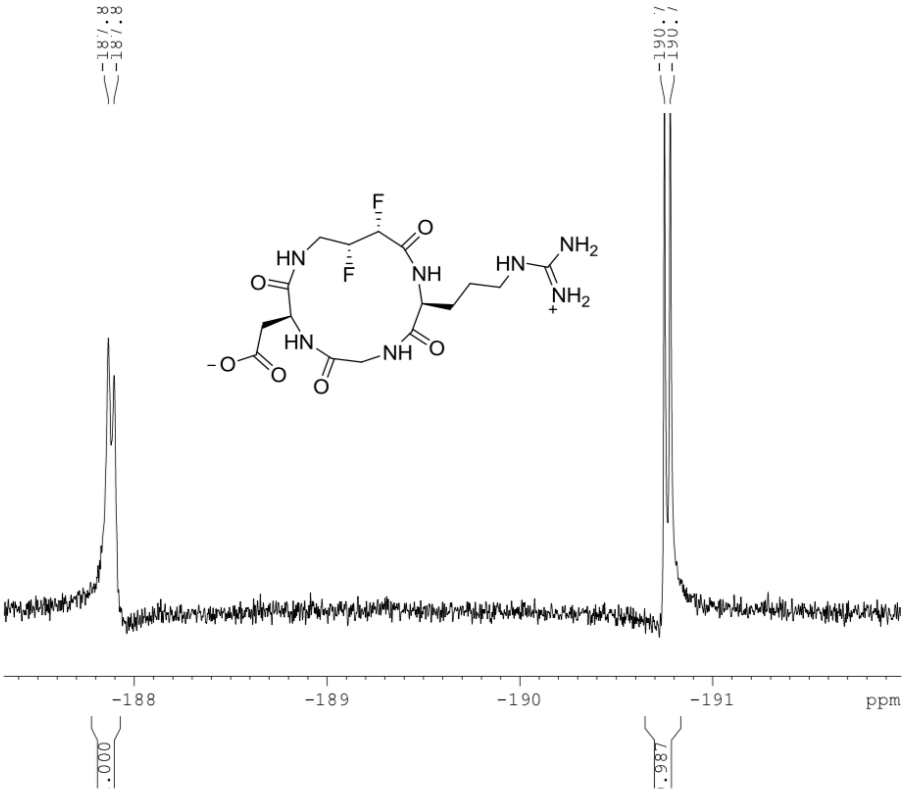
^{1}H NMR (600 MHz, D_2O) of **4h**



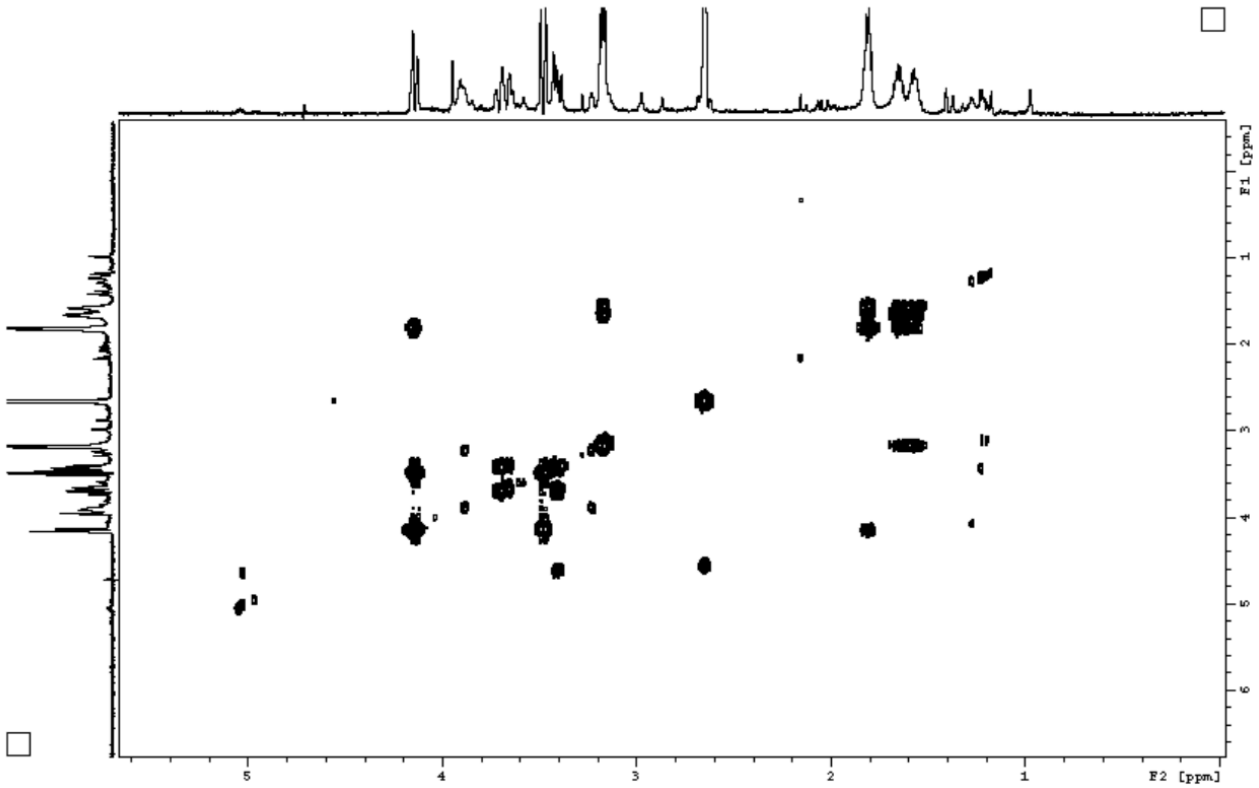
^{19}F NMR (564 MHz, D_2O) of **4h**



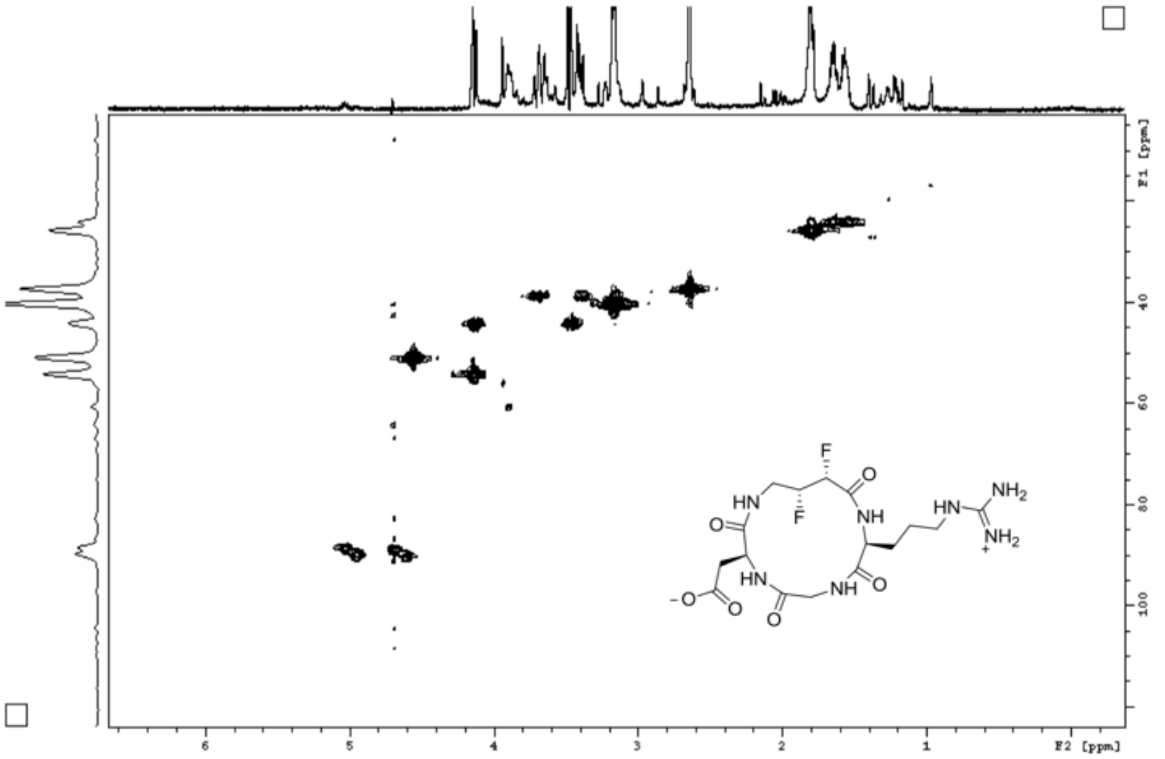
$^{19}F\{^{1}H\}$ NMR (564 MHz, D_2O) of **4h**



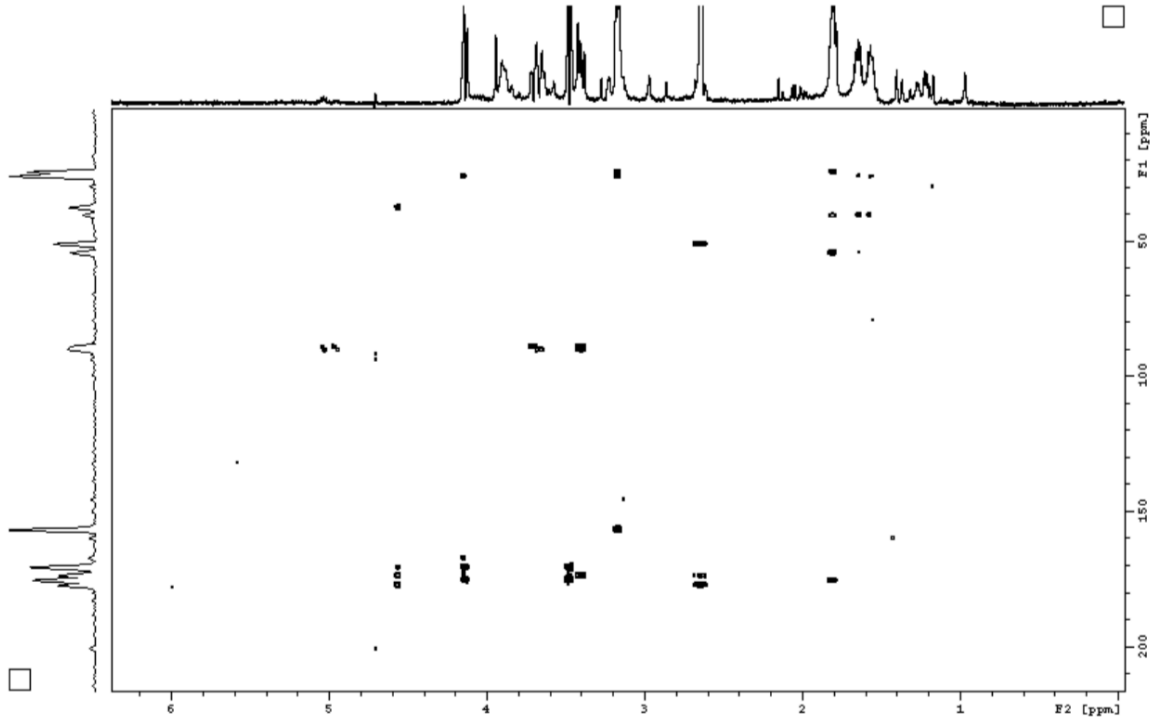
^{1}H-^{1}H COSY (600 MHz, D_2O) of **4h**



^{1}H-^{13}C HSQC (600 MHz, D_2O) of **4h**



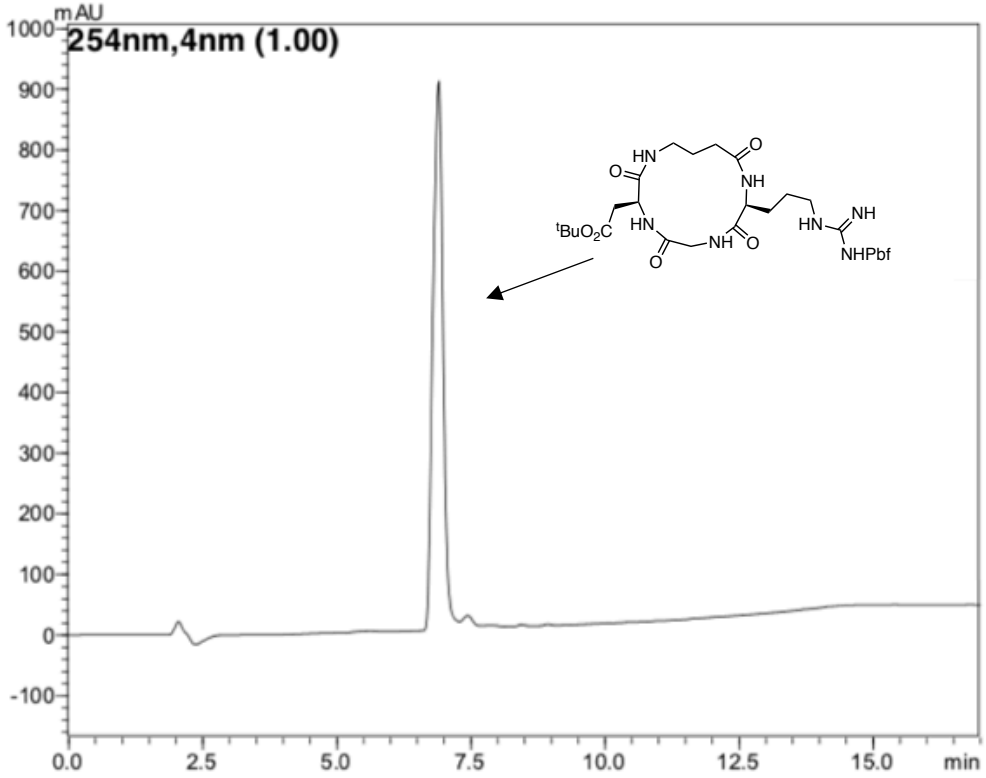
^{1}H-^{13}C HMBC (600 MHz, D_2O) of **4h**



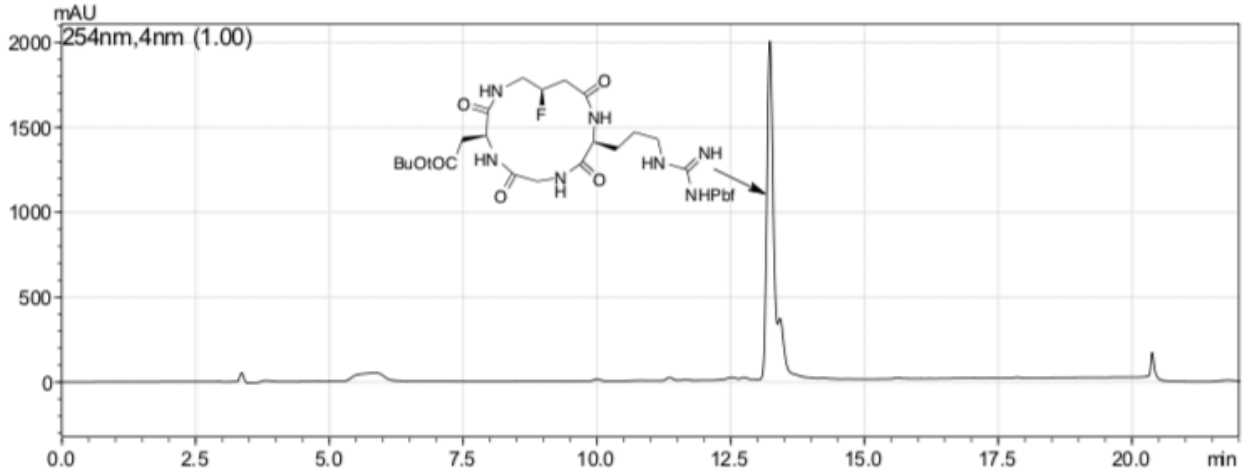
2. HPLC traces of cyclic peptides

HPLC analysis of the final peptides (**4a**–**h**) was not very informative, because (i) these compounds lack a chromophore, and (ii) they are highly polar, typically eluting close to the solvent front. Therefore, HPLC traces of the penultimate protected cyclic peptides (**5a**–**h**) are reproduced here instead. It is assumed that the final deprotection reactions (**5**→**4**) were quantitative. HPLC conditions are described in the main manuscript.

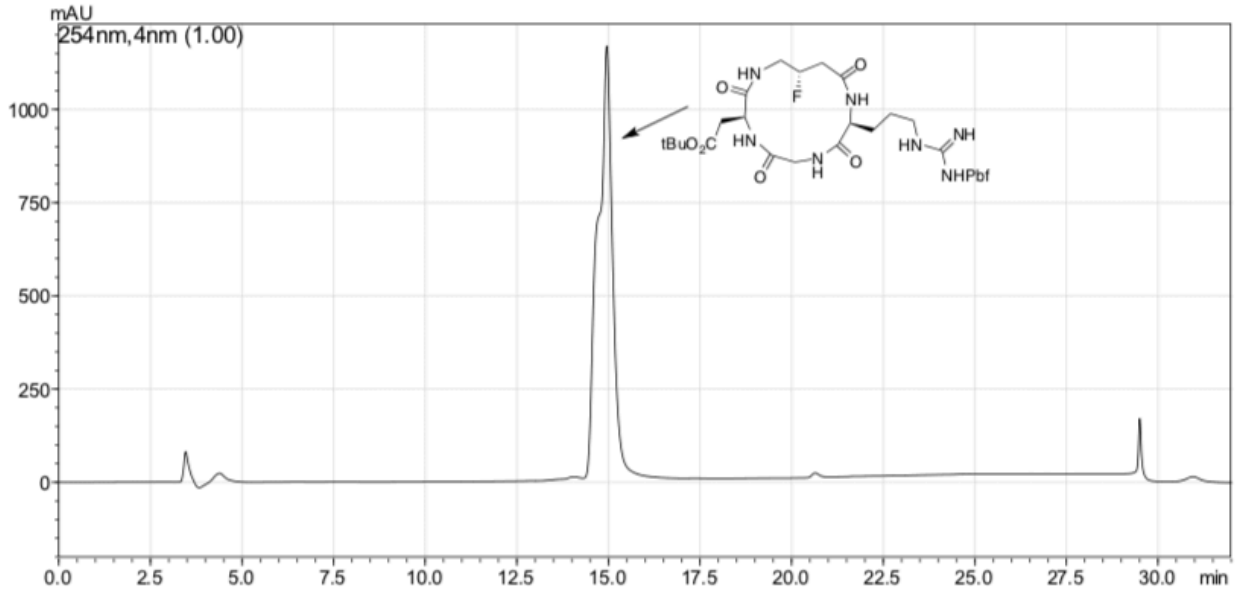
Cyclic peptide 5a



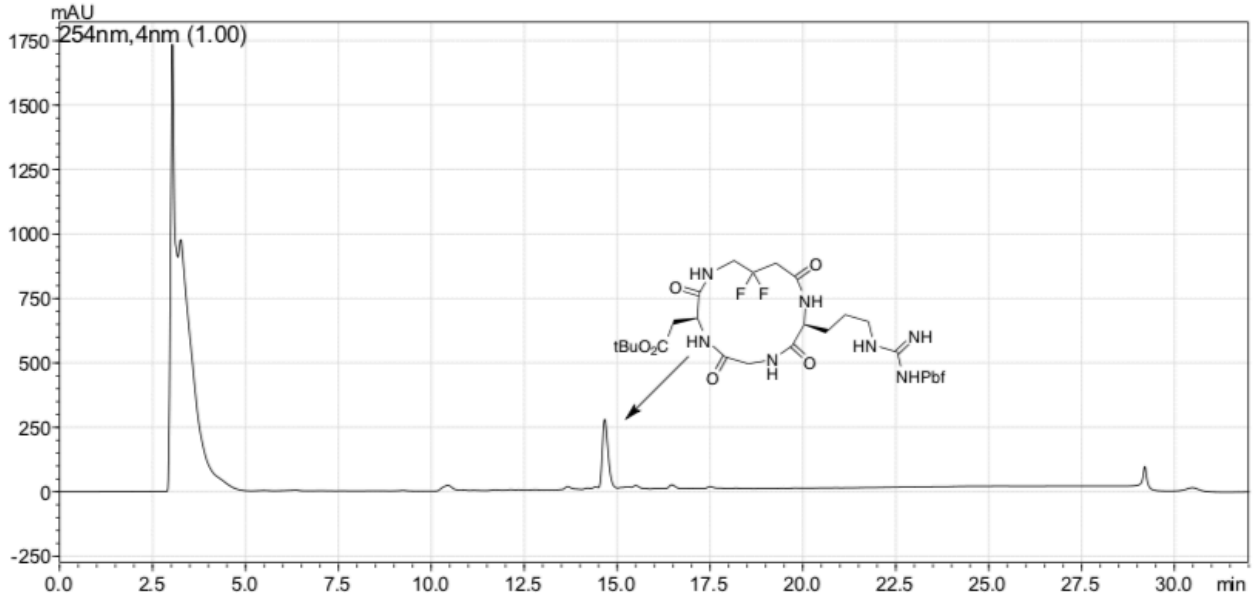
Cyclic peptide 5b



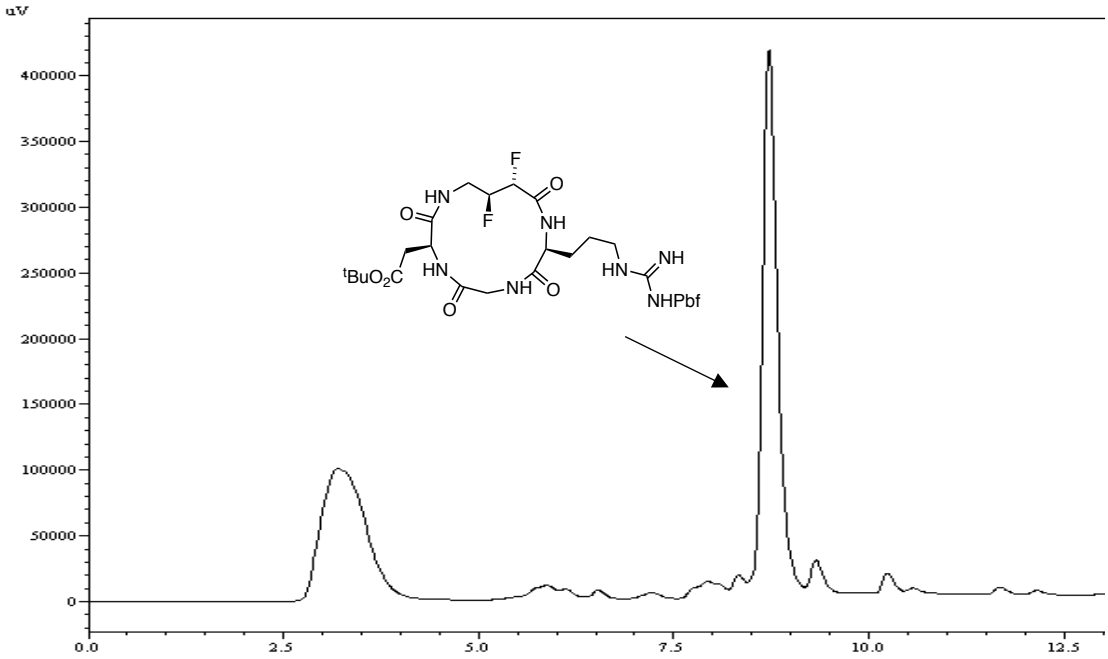
Cyclic peptide 5c



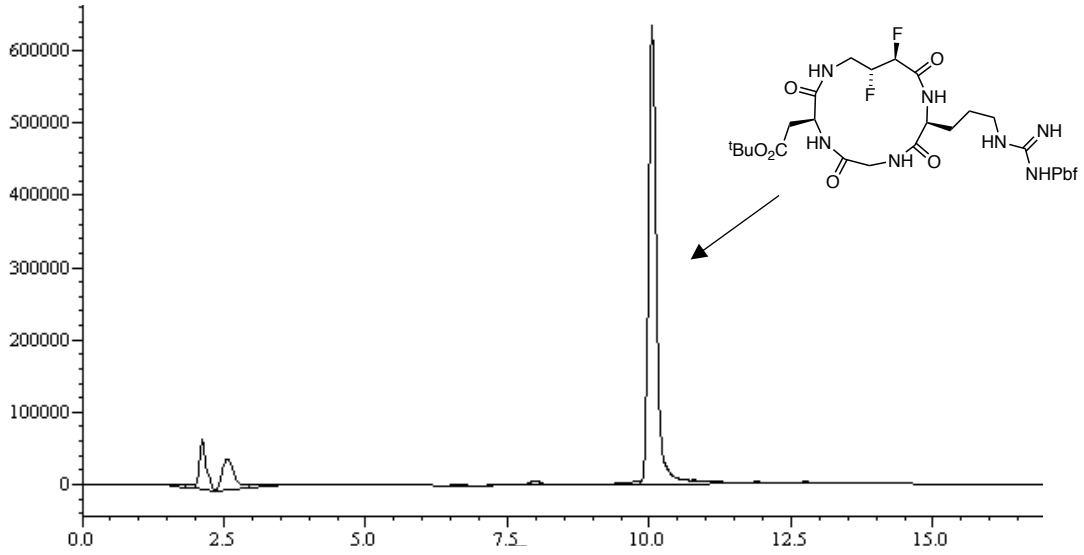
Cyclic peptide 5d



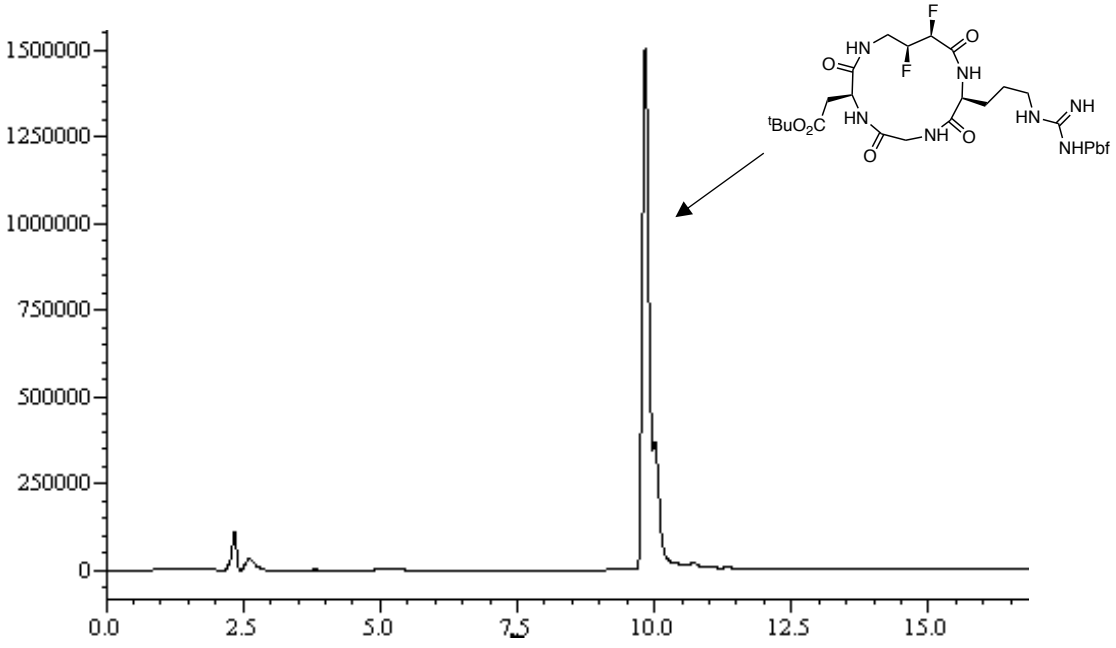
Cyclic peptide 5e



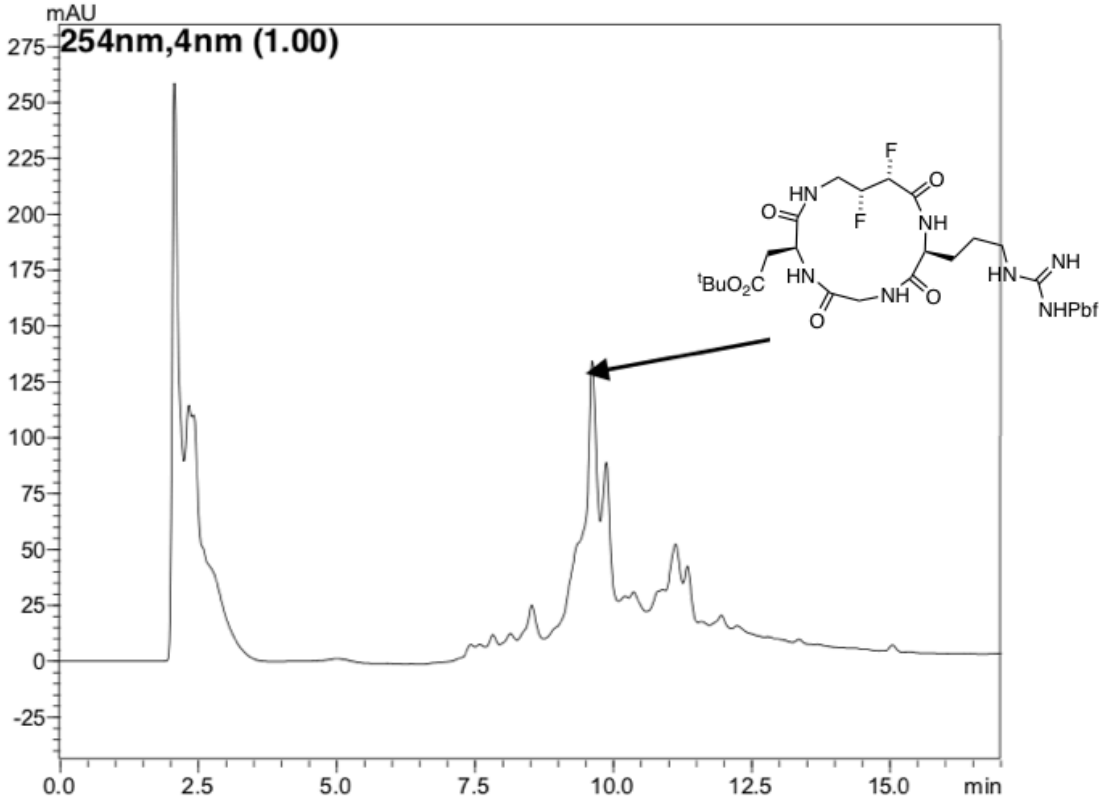
Cyclic peptide 5f



Cyclic peptide 5g

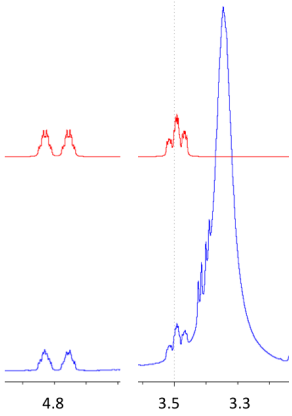
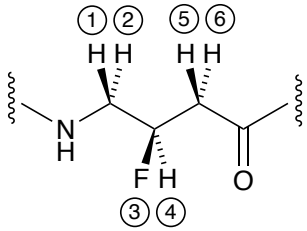
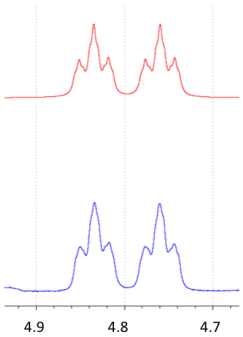
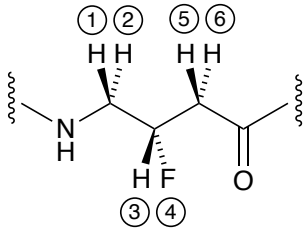


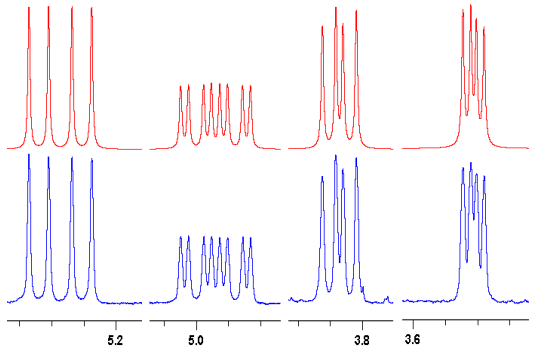
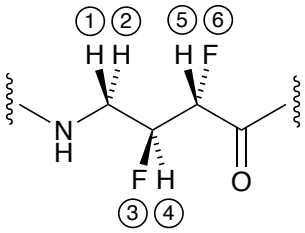
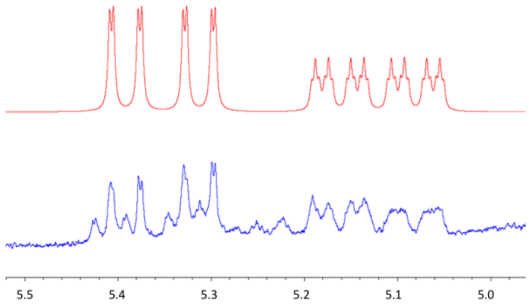
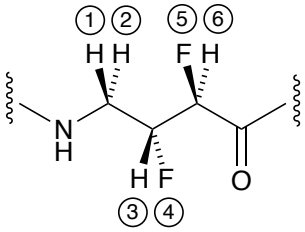
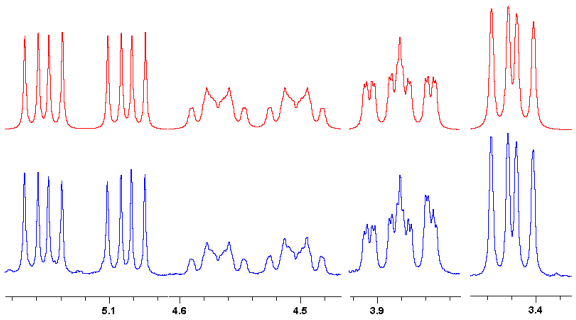
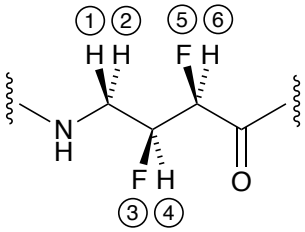
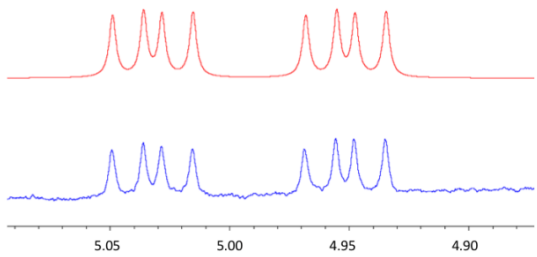
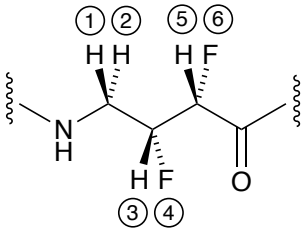
Cyclic peptide 5h
[plus substantial impurities as discussed in main manuscript]



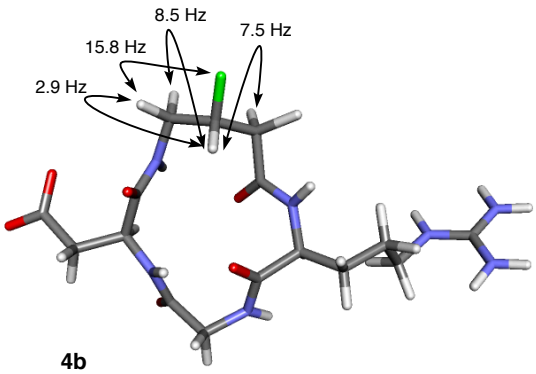
3. Conformational analysis

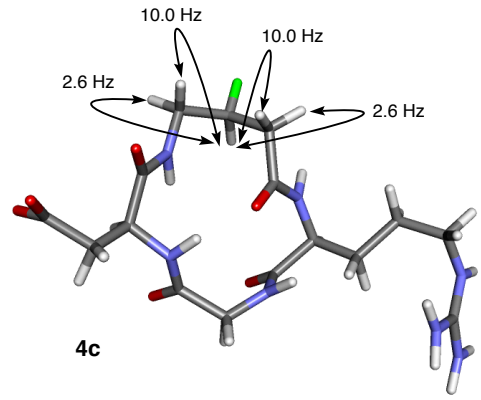
NMR *J*-based analyses were performed in order to glean information on certain dihedral angles within the GABA residues of **4b**–**h**. Signals from the ^{1}H NMR spectra of **4b**–**h**, corresponding to the GABA residue, were simulated using the DAISY module of the Bruker Topspin software. This allowed the associated *J*-values to be elucidated. These *J*-values were interpreted in a qualitative sense: "large" *J*-values (*i.e.* $^{3}J_{HH} > 6$ Hz, $^{3}J_{HF} > 25$ Hz) were taken to indicate a predominantly *anti* alignment, and "small" *J*-values (*i.e.* $^{3}J_{HH} < 3$ Hz, $^{3}J_{HF} < 15$ Hz) were taken to indicate a predominantly *gauche* alignment. Some signals could not be accurately simulated, due to signal overlap and/or multiplet complexity. The ^{1}H NMR spectra of **4b**–**d** were recorded in D_2O, whereas the ^{1}H NMR spectra of **4a**,**e**–**h** were recorded in both D_2O and DMSO-d_6. For the latter group, it was found that DMSO-d_6 generally gave better signal dispersion / resolution. It was also observed, where direct comparisons could be made, that the *J*-values changed little from one solvent to the other.

Cyclic peptide	NMR solvent	Selected ^{1}H NMR signals from GABA residue (simulated = red; experimental = blue)	Elucidated *J*-values
4b	D_2O		 H^1–H^4 = 2.9 Hz H^1–F^3 = 15.8 Hz H^2–H^3 = 8.5 Hz H^4–H^5 = 5.0 Hz H^4–H^6 = 7.5 Hz
4c	D_2O		 H^1–H^3 = 10.0 Hz H^2–H^3 = 2.6 Hz H^3–H^5 = 10.0 Hz H^3–H^6 = 2.6 Hz

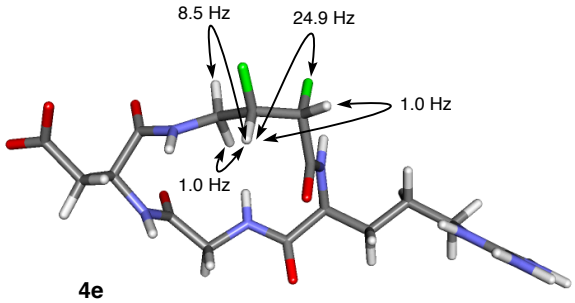
4d	D_2O	It was impossible to accurately simulate these signals due to complexity / overlap.	n/a
4e	DMSO-d_6		 H^1–H^4 = 8.5 Hz H^2–H^4 = 1.0 Hz F^3–H^5 = 21.4 Hz H^4–H^5 = 1.0 Hz H^4–F^6 = 24.9 Hz
4f	DMSO-d_6		 H^3–F^5 = 22.8 Hz H^3–H^6 = 2.5 Hz F^4–H^6 = 18.3 Hz
4g	DMSO-d_6		 H^1–H^4 = 9.2 Hz H^2–H^4 = 1.7 Hz F^3–H^6 = 13.9 Hz H^4–F^5 = 13.0 Hz H^4–H^6 = 7.8 Hz
4h	DMSO-d_6		 H^3–H^5 = 7.8 Hz F^4–H^5 = 12.4 Hz

Molecular dynamics simulations and energy minimisations were performed using the CHARMm force field [F. A. Momany and R. Rone, *J. Comput. Chem.*, 1992, **13**, 888–900] in Discovery Studio (version 4.0, Biovia). The initial molecule geometries were arbitrarily set using a sketching tool. Dihedral restraints were placed on the fluorinated segment of each molecule, where appropriate, to bias that segment towards the conformation indicated by *J*-based analysis. The structures were minimised using the Smart minimiser algorithm in Discovery Studio using the Generalised Born with Molecular Volume implicit solvent model with a dielectric constant of 50 to mimic DMSO. The RMS gradient was set to 0.1 kcal mol^{-1} $Å^{-1}$ and minimisation was run for 10,000 iterations or until convergence. Molecular dynamics simulations were performed using the same implicit solvent model described above. Molecules were equilibrated for 10 ps before being subjected to a 1,000 ps production run with a timestep of 1 fs. Equilibration, heating and production steps were conducted at 500 K. The default NPT (constant pressure and temperature) ensembles were used. A series of 10–20 "snapshots" were taken at regular time intervals during each run, and all of these snapshots were minimised using the procedure described above (*i.e.* without any dihedral restraints). The resulting final structures were inspected, and any structures that had dihedral angles within the fluorinated segment that were inconsistent with the NMR data (see Table above) were discarded. The remaining structures were graphically overlaid by tethering at the α-carbons of the arginine, glycine and aspartate residues. The most populous cluster of structures was assumed to represent the preferred shape for each of **4a**–**h**. This subset of the final structures is shown in Figure 3 of the main manuscript. A representative structure of each of **4a**–**h** is shown below, with diagnostic *J*-values superimposed.

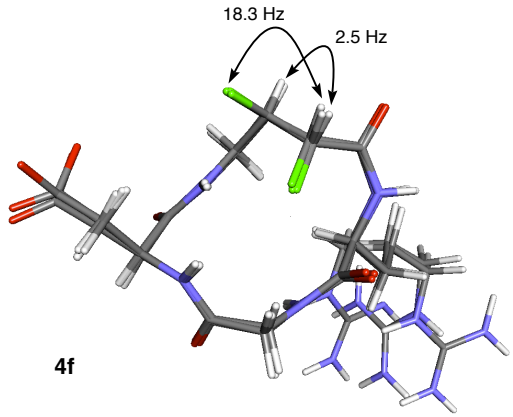



10.0 Hz
10.0 Hz
2.6 Hz
2.6 Hz

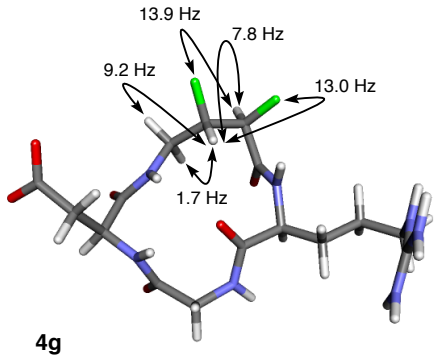
4c


8.5 Hz
24.9 Hz
1.0 Hz
1.0 Hz

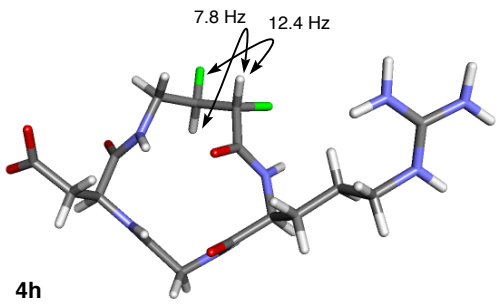
4e


18.3 Hz
2.5 Hz

4f


13.9 Hz
7.8 Hz
9.2 Hz
13.0 Hz
1.7 Hz

4g


7.8 Hz
12.4 Hz

4h