Squaratides: A tunable platform for anion binding in peptide-inspired macrocycles

Farhad Ali. Mohammed,[a,b] Hua Tong,[a] Chris S. Hawes,[c] Shayon Bhattacharya,[b,d] Damien Thompson, [b,d] Katrina A. Jolliffe[e] and Robert 4B. P. Elmes*[a,b,f]

[a.] *Department of Chemistry, Maynooth University, National University of Ireland, Maynooth, Co. Kildare, Ireland.*

Tel: +353 1 7084615; E-mail: robert.elmes@mu.ie

[b.] *Synthesis and Solid State Pharmaceutical Centre (SSPC), Ireland.*

[c.] *School of Chemical and Physical Sciences, Keele University, Keele ST5 5BG, U.K.*

[d.] *Department of Physics, Bernal Institute, University of Limerick, V94 T9PX Limerick, Ireland*

[e.] *ARC Centre of Excellence for Innovations in Peptide and Protein Science, School of Chemistry, University of Sydney, 2006, NSW, Australia*

[f.] *Kathleen Lonsdale Institute for Human Health Research, Maynooth University, National University of Ireland, Co. Kildare, W23 F2H6 Maynooth, Ireland*

Supporting Information

1. Methods

Solvents and Reagents

All chemicals purchased were obtained from commercial suppliers and used without further purification. Anhydrous DMF was purchased from Sigma Aldrich.

Chromatography

Flash chromatography was performed with Merck Silica Gel 60. TLC was performed on Merck Silica Gel F254 plates and visualized under UV light (λ = 254 nm). Compounds were lyophilized on a Labconco Freezone 1 Dry system.

Spectroscopy

NMR spectra were recorded on a Bruker Ascend 500 NMR spectrometer, operated at 500 MHz for ^{1}H NMR and 126 MHz for ^{13}C NMR analysis at 298 K. The residual solvent peak was used as an internal standard for DMSO-d_6 and TMS for $CDCl_3$. Chemical shifts (δ) were reported in ppm. NMR spectra were processed, and stack plots were produced using MestRe Nova 6.0.2 and 14.3 software. Proton and carbon signals were assigned with the aid of 2D NMR experiments (COSY, HSQC and HMBC). Multiplicity was given as s = singlet, bs = broad singlet, d = doublet, dd = doublet of doublets, t = triplet, q/quar = quartet, p/quin = pentet/quintet, m = multiplet, dm = doublet of multiplets as appropriate, and J values are given in Hz. In some cases, ^{13}C resonances could not be observed due to the limited solubility of some macrocycles in the NMR solvent, resulting in reduced signal-to-noise despite extended acquisition times. In these cases, compound identity and purity were confirmed by ^{1}H NMR and LCMS analysis.

Infrared spectra were obtained via ATR as a solid on a zinc selenide crystal in the region of 4000 – 400 cm^{-1} on a Perkin Elmer Spectrum 100 FT-IR spectrophotometer.

LC-MS was performed on an Agilent Technologies 1200 Series instrument consisting of a G1322A quaternary pump and a G1314B UV detector (254 nm) coupled to an Advion Expression L Compact Mass Spectrometer (ESI) operating in positive mode. High-resolution

ESI spectra were recorded on an Agilent 6545 QTof instrument coupled with 6200 series TOF/6500 series software of the B9044.0 version.

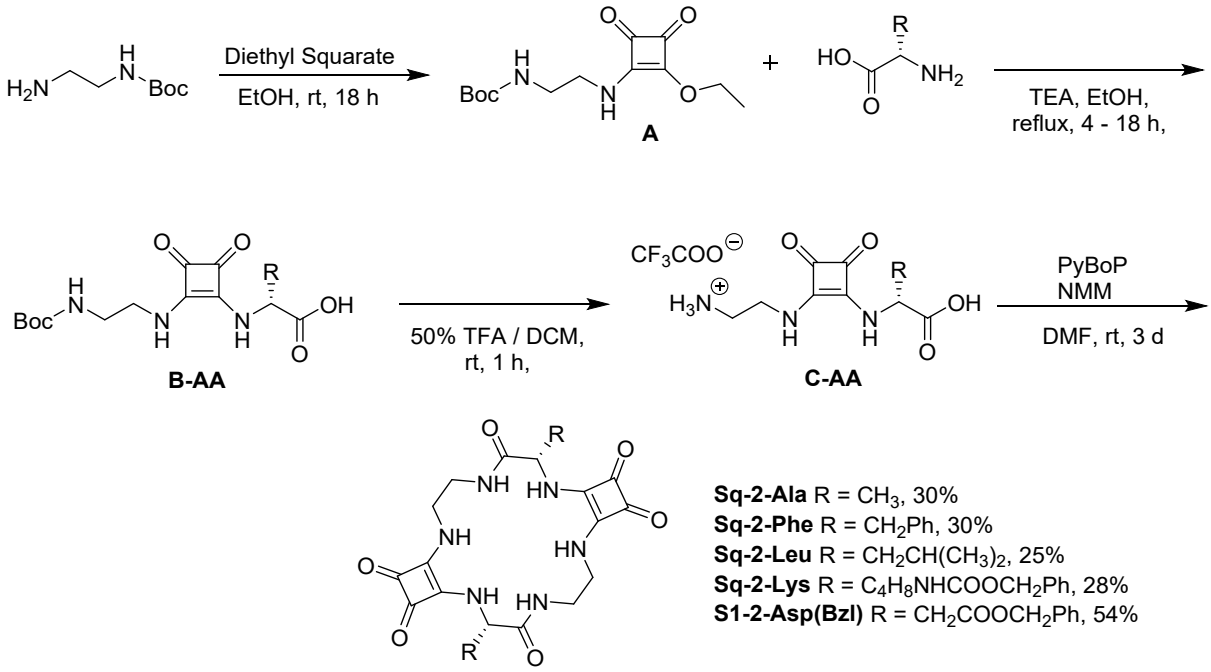
^{1}H NMR Anion Titration

All tetrabutylammonium halide salts (TBAX) and the receptors were lyophilised before use, and halide salts were stored under vacuum in a desiccator. Solutions of the TBA salts were made up in DMSO-d6, which was dried over 3Å molecular sieves before use, to a concentration of 250 mM. An aliquot of stock solution of receptor in DMSO-d_6 was diluted to 1 mL (2.5 mM). 600 µL of this solution was added to an NMR and the ^{1}H NMR spectrum was recorded. Subsequent additions of aliquots of TBAX solutions were added to the NMR tube and shaken vigorously to ensure homogenisation. This process was repeated up to 20 equivalents of halide was reached. The ^{1}H NMR spectra were analysed and processed, and stack plots were generated using MestReNova 6.0.2 software. The changes in chemical shifts of the squaramide NH protons when observed, were plotted against the amount of anion added to generate binding isotherms. Non-linear global curve fitting of the isotherms using the

BindFit program were carried out for three binding modes for each titration (H:G 1:1, 1:2 and 2:1). The most appropriate binding model was used to obtain the association constant for that replicate.

Experimental Procedures

Overall Synthetic Scheme towards Cyclic Squaratide Dimers



Scheme S1. General synthetic scheme towards Cyclic Squaratide dimers

Synthesis of Sq-2-AA dimers

General method A:

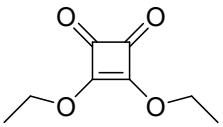
To a stirring solution of **A** (1 eq., 1.4 mmol, 0.4 g) in EtOH (24 mL), TEA (6 eq., 8.4 mmol, 1.18 mL) the relevant L – amino acid (2 eq., 2.8 mmol, 0.251 g) was added. The resulting mixture was stirred under reflux for 24 h. The reaction mixture was concentrated in vacuo, suspended in DCM (20 mL) and filtered. the solvent was removed in vacuo to afford crude product, which was purified via column chromatography (SiO_2), and triturated in diethyl ether to yield the title compound.

General method B:

The corresponding AA-Sq-EDA(Boc) monomers **(B-AA))** underwent boc deprotection using a 50:50 TFA: DCM solution, allowing the mixture to stir for 1 hour before being removed in vacuo, redissolved in MeOH and precipitated out of diethyl ether. The deprotected monomers (1 eq., 0.8 mmol, 0.274 g) was dissolved in DMF (9 mL) and NMM (6 eq., 4.82 mmol, 0.53 mL) was added and the resulting solution was stirred for 10min. To this solution PyBOP (3 eq., 2.41 mmol, 1.25 g) was added and allowed to stir for 48 hrs. The DMF was removed in vacuo and the crude mixture was triturated in Ethyl Acetate. The mixture was added to a falcon tube and centrifuged under 4000 RPM for 10 min (2

x Ethyl Acetate and 1x Diethyl Ether). The supernatant was dispelled, and the precipitate was dissolved in 50:50 MeCN/H_2O and lyophilized to afford the Cyclic Squaratides (Sq-2-AA)).

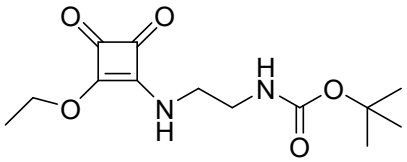
3,4-diethoxyl-cyclobut-3-ene-1,2-dione (DeSq)



To a suspension of squaric acid (1 eq., 87.7 mmol, 10 g) in EtOH (125 mL) was added triethyl orthoformate (5 eq., 438 mmol, 66.2 mL) and the resulting solution was heated at reflux for 3 days. The solvent was removed in vacuo and the orange/yellow oily residue was purified by flash column chromatography eluting with DCM to yield a clear yellow oil in a 90% yield. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 4.65 (q, *J* = 7.1 Hz, 4H), 1.37 (t, *J* = 7.1 Hz, 6H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 189.1, 183.7, 70.2, 54.9, 39.5, 15.3.

Tert-butyl (2-((2-ethoxy-3,4-dioxocyclobut-1-en-1- yl)amino)ethyl)carbamate

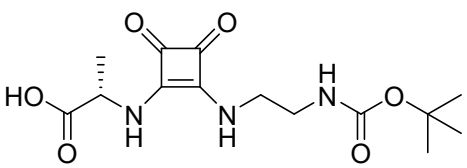
(A)



To a stirring solution of 3,4-diethoxyl-cyclobut-3-ene-1,2-dione (1.5 eq. 14 mmol, 2.39 g) in EtOH (50 mL), a solution of *N*-Boc-ethylenediamine (1 eq. 9.4 mmol, 1.5 g) in EtOH (70 mL) was added dropwise and the resulting solution was stirred under RT overnight. The solvent was removed in vacuo to yield a yellow mixture, which was then purified by column chromatography (SiO_2), using a 0 - 5% EtOH:DCM gradient as eluent to afford the product (**A**) as a yellow solid (2.2 g, 82 %). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.74 – 8.53 (t, 1H), 6.89 (d, *J* = 5.7 Hz, 1H), 4.63 (q, *J* = 6.5 Hz, 2H), 3.49 – 3.43 (m, 1H), 3.30 (d, *J* = 7.8 Hz, 1H), 3.08 (q, J = 6.1 Hz, 2H),(1.35 (s, 12H). **^{13}C NMR** (125 MHz, DMSO-d_6) δ 189.7, 182.7, 176.9, 173.2, 156.1, 78.3, 69.2, 44.6, 44.2, 40.8, 28.7, 16.2.

(2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut-1-en-1-yl)-L-alanine

(B-Ala)

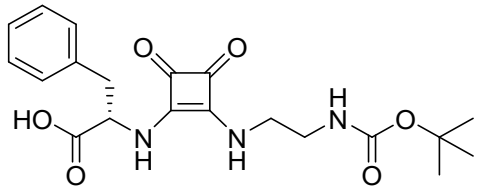


B-Ala was synthesised as per **General method A** and was purified by flash chromatography in 0 - 30 % EtOH: DCM as eluent to afford the title compound as a white powder in a 64 % yield. **^{1}H NMR** (500

MHz, DMSO-d_6) δ 13.08 (s, 1H), 7.74 (s, 1H), 7.54 (s, 1H), 6.93 (t, *J* = 5.1 Hz, 1H), 4.63 (s, 1H), 3.52 (s, 2H), 3.09 (q, *J* = 5.9 Hz, 2H), 1.41 (d, *J* = 7.2 Hz, 3H), 1.36 (s, 9H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.2, 182.8, 174.1, 168.6, 167.4, 156.2, 78.3, 51.7, 43.6, 41.6, 28.7, 20.4.

(2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut-1-en-1- yl)-L-phenylalanine

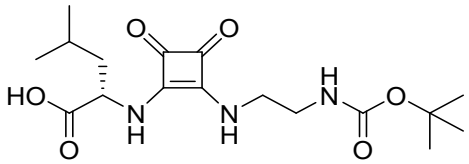
(B-Phe)



(B-Phe) was synthesised as per **General method A** and was purified by flash chromatography in 0 - 30 % EtOH: DCM as eluent to afford the title compound as a Beige powder in a 45 % yield. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.20 (s, 1H), 7.66 (s, 1H), 7.57 (s, 1H), 7.29 (t, *J* = 7.2 Hz, 2H), 7.22 (t, *J* = 7.2 Hz, 1H), 7.18 (d, *J* = 7.2 Hz, 2H), 6.92 (s, 1H), 4.92 (s, 1H), 3.50 (s, 2H), 3.16 (d, *J* = 9.6 Hz, 1H), 3.07 (q, *J* = 4.1 Hz, 2H), 3.02 (dd, *J* = 9.6 Hz, 1H), 1.36 (s, 9H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.3, 182.7, 172.7, 168.6, 167.3, 156.2, 136.7, 129.8, 128.8, 127.2, 78.3, 57.2, 43.5, 41.6, 39.7, 28.7.

(2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut-1-en-1- yl)-L-leucine

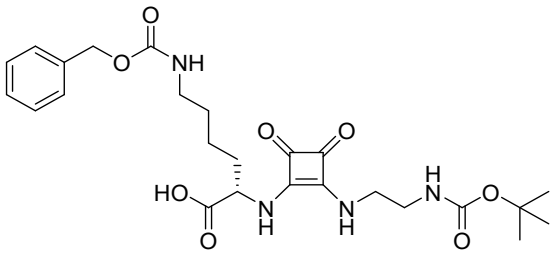
(B-Leu)



(B-Leu) was synthesised as per **General method A** and was purified by flash chromatography in 0 - 30 % EtOH: DCM as eluent to afford the title compound as a white powder in a 76 % yield. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.05 (s, 1H), 7.67 (s, 1H), 7.45 (s, 1H), 6.94 (s, 1H), 4.67 (s, 1H), 3.53 (s, 2H), 3.12 – 3.07 (q, 2H), 1.64 (d, *J* = 4.5 Hz, 1H), 1.64 – 1.57 (m, 2H), 1.36 (s, 9H), 0.90 (dd, *J* = 6.3 Hz, 6H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.2, 182.2, 174.0, 168.2, 158.4, 54.6, 43.9, 42.4, 41.6, 28.7, 24.7, 23.3.

N6 -((benzyloxy)carbonyl)-N2 -(2-((2-((tertbutoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut-1-en-1-yl)-L-lysine

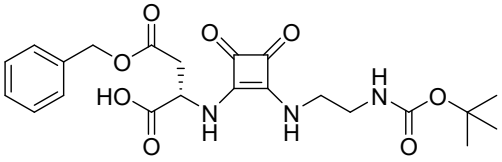
(B-(Lys(Cbz))



(B-(Lys(Cbz)) was synthesised as per **General method A** and was purified by flash chromatography in 0 - 30 % EtOH: DCM as eluent to afford the title compound as a yellow powder in a 78 % yield. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.94 (s, 1H), 7.72 (s, 1H), 7.53 (s, 1H), 7.36 (s, 2H), 7.34 (s, 1H), 7.33 (s, 2H), 7.24 (s, 1H), 6.92 (s, 1H), 4.99 (s, 2H), 4.63 (s, 1H), 3.52 (q, 2H), 3.10 (q, 2H), 2.97 (m, 2H), 1.75 (dm , 2H), 1.42 (m, 2H), 1.36 (s, 9H), 1.34 (m, 2H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 182.9, 182.4, 173.2, 168.2, 167.2, 156.2, 155.9, 137.4, 128.5, 127.8, 127.4, 78.0, 65.3, 55.6, 43.3, 41.2, 40.2, 33.2, 29.1, 28.3, 15.3.

(S)-4-(benzyloxy)-2-((2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4- dioxocyclobut-1-en-1-yl)amino)-4-oxobutanoic acid

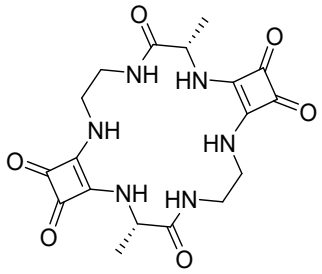
(B-Asp(Bzl))



(B-Asp(Bzl)) was synthesised as per **General method A** and was purified by flash chromatography in 0 - 30 % EtOH: DCM as eluent to afford the title compound as a yellow powder in an 80 % yield. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.87 (s, 1H), 8.05 (s, 1H), 7.33 (d, *J* = 7.2 Hz, 2H), 7.32 (m, 2H), 7.31 – 7.27 (m, 1H), 7.25 (s, 1H), 5.02 (s, 2H), 4.59 (s, 1H), 3.52 (m, 2H), 2.93 (q, *J* = 7.1 Hz, 3H), 2.83 (dd, *J* = 6.6 Hz, 2H), 1.35 (s, 9H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.1, 182.7, 172.9, 170.8, 168.5, 167.8, 156.2, 136.7, 128.8, 128.3, 128.2, 78.1, 65.8, 55.0, 43.6, 41.4, 40.0, 28.7.

(3S,14S)-3,14-dimethyl-2,5,8,13,16,19-hexaazatricyclo[18.2.0.09,12]docosa1(20),9(12)-diene-4,10,11,15,21,22-hexaone

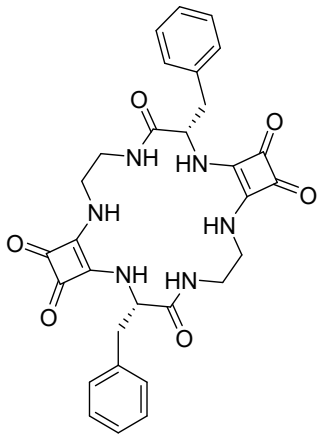
(Sq-2-Ala)



Sq-2-Ala was synthesised as per **General method B** to afford the title compound as a beige powder in a 34 % yield. **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 14 min. **HRMS** (ESI+): m/z Calc. for $C_{18}H_{22}N_6O_6$ ([M+H]$^+$): 419.1608, found 419.1681 (1.68 ppm). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.06 (s, 2H), 7.54 (s, 2H), 7.17 (s, 2H), 4.54 (s, 2H), 3.78 (d, 3H), 2.88 (s, 2H), 1.24 (dd, *J* = 5.4 Hz, 6H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.0, 182.1, 172.4, 168.8, 168.3, 52.4, 44.0, 39.0, 20.6.

(3S,14S)-3,14-dibenzyl-2,5,8,13,16,19-hexaazatricyclo[18.2.0.09,12]docosa1(20),9(12)-diene-4,10,11,15,21,22-hexaone

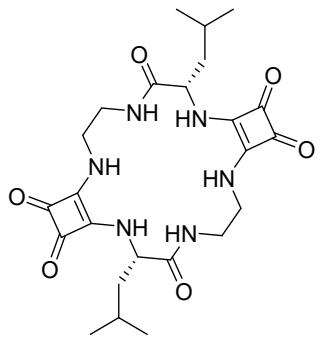
(Sq-2-Phe)



Sq-2-Phe was synthesised as per **General method B** to afford the title compound as a beige powder in a 46 % yield. **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 27 min. **HRMS** (ESI+): m/z Calc. for $C_{30}H_{30}N_6O_6$ ([M+H]$^+$): 571.2235, found 571.2309 (1.38 ppm). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.10 (s, 2H), 7.53 (d, J = 7.1 Hz, 1H), 7.23 (t, *J* = 7.2 Hz, 10H), 7.05 (s, 1H), 4.72 (s, 2H), 3.76 (d, *J* = 7.3 Hz, 2H), 2.97 (s, 2H), 2.83 (s, 2H), 2.74 (s, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 182.0, 181.2, 171.9, 168.5, 167.7, 137.5, 129.7, 128.6, 126.9, 58.0, 44.1, 38.9, 38.9.

(3S,14S)-3,14-diisobutyl-2,5,8,13,16,19- hexaazatricyclo[18.2.0.09,12]docosa-1(20),9(12)-diene-4,10,11,15,21,22- hexaone

(Sq-2-Leu)

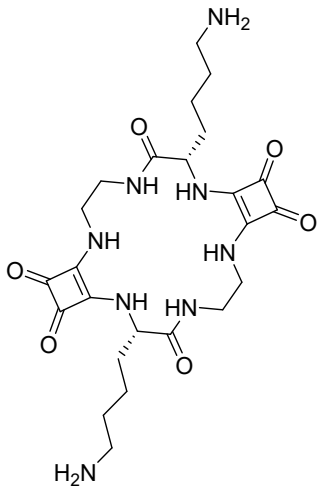


Sq-2-Leu was synthesised as per **General method B** to afford the title compound as a beige powder in a 26 % yield. **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 25 min. **HRMS** (ESI+): m/z Calc. for $C_{24}H_{34}N_6O_6$ ([M+H]$^+$): 503.2548, found 503.2621 (1.61 ppm). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.13 (d, *J*

= 7.5 Hz, 1H), 7.46 (d, *J* = 7.2 Hz, 1H), 7.08 (d, *J* = 7.0 Hz, 1H), 4.52 (q, *J* = 7.9 Hz, 2H), 3.87 (q, *J* = 10.4 Hz, 1H), 3.75 (q, J = 9.1 Hz, 1H), 2.82 (d, *J* = 13.1 Hz, 1H), 1.60 – 1.47 (m, 3H), 1.42 (s, 3H), 0.86 (dd, J = 6.1 Hz, 12H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ δ 55.4, 44.2, 42.5, 24.6, 23.5, 22.1.

(3S,14S)-3,14-bis(4-aminobutyl)-2,5,8,13,16,19-hexaazatricyclo[18.2.0.0^{9,12}]docosa-1(20),9(12)-diene-4,10,11,15,21,22-hexaone

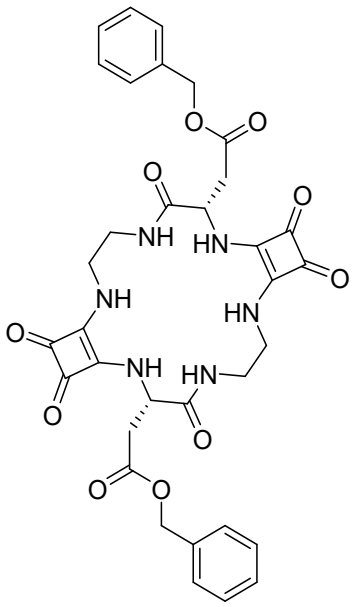
(Sq-2-Lys)



Sq-2-Lys was synthesised as per **General method B** to afford the title compound as a beige powder in a 34 % yield. The resulting benzyl protected groups were deprotected in a HBr 33 % wt AcOH solution (3 mL) which was stirred for 20 min before being filtered and triturated in diethyl ether. The mixture was added to a falcon tube and centrifuged under 4000 RPM for 5 min (3x Diethyl Ether). The supernatant was dispelled, and the precipitate was dissolved in 50:50 MeCN/H_2O and lyophilized to yield the title compound as a beige powder in a 89 % yield. min). **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 9 min. **HRMS** (ESI+) m/z: Calc. for $C_{24}H_{36}N_8O_6$ ($[M+H]^+$): 533.2756, Found: 533.2828 (-0.33 ppm). **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 28 min. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.08 (s, 2H), 7.71 (s, 6H), 7.61 (s, 2H), 7.28 (s, 2H), 4.46 (s, 2H), 3.82 (d, *J* = 40.5 Hz, 4H), 2.89 (s, 2H), 2.75 (s, 5H), 1.66 (s, 2H), 1.53 (s, 7H), 1.29 (s, 5H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 182.9, 181.8, 171.8, 168.2, 166.7, 165.5, 158.7, 158.5, 148.9, 56.5, 49.5, 44.0, 40.8, 40.4, 40.4, 40.3, 40.2, 40.1, 39.9, 39.8, 39.1, 33.4, 27.1, 27.1, 22.4.

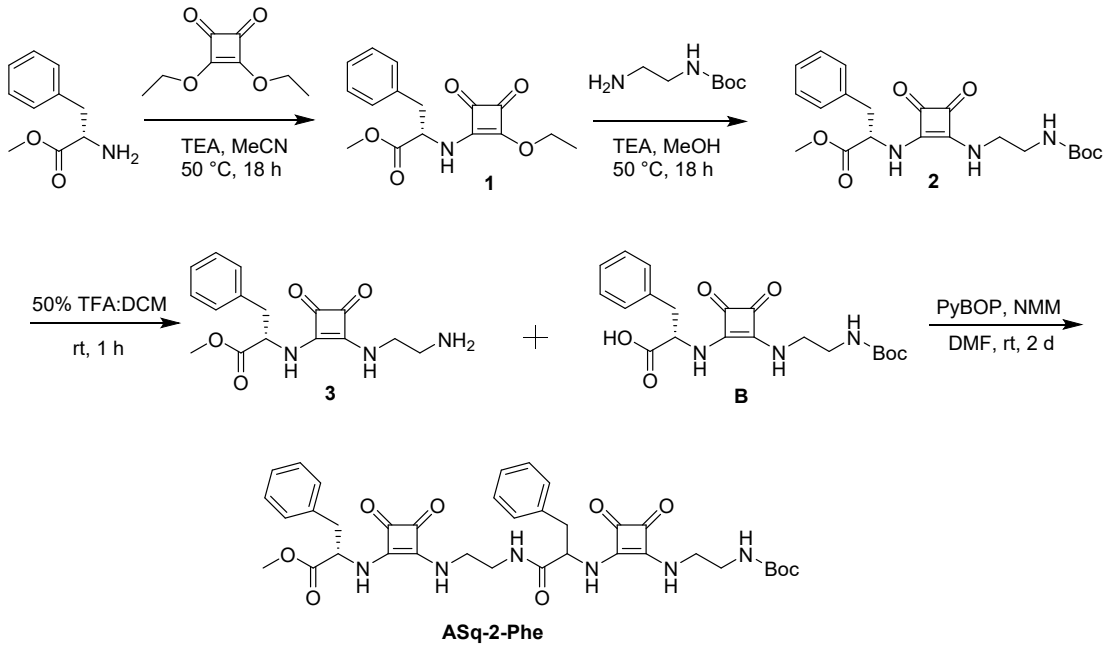
Dibenzyl 2,2'-((3S,14S)-4,10,11,15,21,22-hexaoxo-2,5,8,13,16,19-hexaazatricyclo [18.2.0.09,12]docosa-1(20),9(12)-diene-3,14-diyl)diacetate

(Sq-2-Asp(Bzl))



Sq-2-Asp(Bzl) was synthesised as per **General method B** to afford the title compound as a beige powder in a 34 % yield. **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 28 min. **HRMS** (ESI+): m/z Calc. for $C_{34}H_{34}N_6O_6$ ($[M+H]^+$): 687.2338 , found 687.2411 (0.22 ppm). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 7.83 (s, 2H), 7.57 (s, 2H), 7.34 (m, 11H), 7.10 (s, 1H), 5.07 (d, 4H), 4.91 (m, 2H), 3.58 (m, 4H), 3.23 (m, 2H), 2.99 (m, 2H), 2.78 (dd, *J* = 8.2 Hz, 2H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.4, 181.0, 172.0, 170.6, 170.2, 167.4, 167.1, 136.4, 128.8, 128.5, 128.3, 66.3, 53.8, 43.0, 40.6, 37.7.

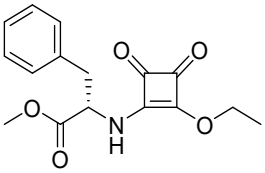
Overall Synthetic Scheme towards Acyclic Squaratide Dimer ASq-2-Phe



Scheme S2. General synthetic scheme towards ASq-2-Phe.

Methyl (2-ethoxy-3,4-dioxocyclobut-1-en-1-yl)-L-phenylalaninate

(1-Phe)

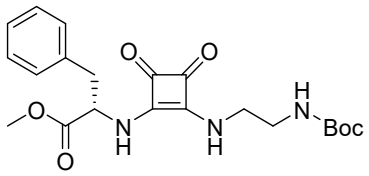


L-Phenylalanine methyl ester HCl (1 eq., 2.45 mmol, 0. 528 g) was dissolved in MeCN (50 mL) with TEA (1.2 eq., 2.94 mmol, 0.41 mL) and the solution was added dropwise into a stirring solution of diethyl squarate (1.2 eq., 2.94 mmol, 0.5 g) in MeCN (10 mL).183 The resulting solution was stirred under 50 °C overnight. The volatiles were removed in vacuo to yield a yellow mixture. The mixture was partitioned between 5 mL of water and 20 mL of AcOEt. The water phase was washed by AcOEt of 20 mL by 3 times in total. The organic phase was combined and dried by MgSO4 and filtered and the solvent was removed in vacuo. The residue was purified by flash chromatography eluting with 1) DCM; 2) 3 % EtOH: DCM to afford the product **(1-Phe)** as an orange oil (0.5129 g, 69.0 %). The product was dissolved in MeOH (4 mL) for further use. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 9.09 (dd, *J* = 8.1 Hz, 1H), 7.29 (t, *J* = 7.2 Hz, 2H), 7.24 (m, 1H), 7.23 (d, J = 7.3 Hz, 2H), 4.91 (s, 0.5H), 4.62 (q, J = 6.5 Hz, 1H), 4.52 (dd, *J* = 11.9, 7.0 Hz, 1H), 4.48 (s, 1H), 3.70 (d, *J* = 4.2 Hz, 3H), 3.24 (dd, *J* = 4.5 Hz, 1H), 2.94 (q,

J = 13.4 Hz, 1H), 1.31 (dt, *J* = 6.9 Hz, 3H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 189.3, 188.8, 183.2, 183.0, 177.7, 177.4, 173.2, 172.7, 171.4, 171.0, 137.1, 129.7, 128.8, 127.2, 69.5, 58.7, 57.8, 53.0, 37.8, 16.0, 16.0.

Methyl (2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut1-en-1-yl)-L-phenylalaninate

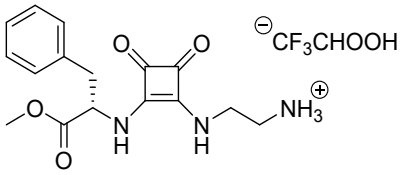
(2-Phe)



To a stirring solution containing 1 mL MeOH solution of Methyl (2-ethoxy-3,4-dioxocyclobut-1-en-1-yl)-L-phenylalaninate (1 eq, 0.421 mmol, 0.1278 g) in MeOH (20 mL), Boc-EDA (1.5eq., 0.632 mmol, 0.101 g) dissolved in MeOH (30 mL) and TEA (0.2 eq., 0.084 mmol, 0.012 mL) were added, and the solution was stirred and refluxed under 50 °C overnight. The volatiles were removed in vacuo to yield an orange-yellow oil. The residue was purified by flash chromatography eluting with 4 % EtOH: DCM to afford **2-Phe** as a light-yellow solid (0.1392 g, 79.1 %). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 7.74 (s, 1H), 7.53 (s, 1H), 7.29 (t, *J* = 7.3 Hz, 2H), 7.23 (t, *J* = 7.3 Hz, 1H), 7.16 (d, *J* = 6.0 Hz, 2H), 6.92 (s, 1H), 5.02 (s, 1H), 3.68 (s, 2H), 3.53 (s, 2H), 3.15 (bd, 1H), 3.07 (s, 2H), 3.07 – 3.00 (m, 1H), 1.36 (s, 9H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.5, 182.7, 171.8, 168.6, 166.4, 136.4, 129.7, 128.9, 127.3, 78.3, 57.2, 52.8, 43.6, 41.6, 39.7, 28.7.

(S)-2-((2-((1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-3,4- dioxocyclobut-1-en-1-yl)amino)ethan-1-aminium 2,2,2-trifluoroacetate

(3-Phe)

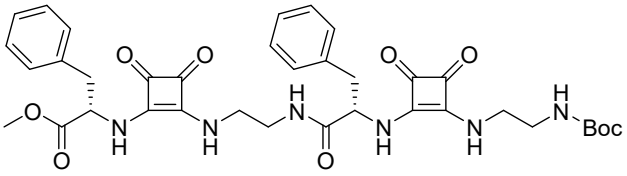


To a flask containing Methyl (2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut1-en-1-yl)-L-phenylalaninate (0.306 mmol ,0.1276 g), DCM: TFA 1:1 solution (4 mL) was added. The resulting solution was sealed by a stopper softly and stirred for 1h. The volatiles were removed in vacuo to yield a yellow clear oil. Nitrogen gas was blown into the flask to remove as much TFA as possible. Fully converted. **^{1}H NMR** (500 MHz, DMSO-d_6) δ 7.95 (s, 1H), 7.89 (s, 3H), 7.73 (s, 1H), 7.30 (t, *J* = 7.2 Hz, 2H), 7.23 (t, *J* = 7.2 Hz, 1H), 7.19 (d, *J* = 7.2

Hz, 2H), 5.03 (s, 1H), 3.69 (s, 3H), 3.68 – 3.63 (m, 2H), 3.18 (dd, *J* = 5.4 Hz, 1H), 3.04 (dd, *J* = 13.8, 8.0 Hz, 2H), 3.00 (d, J = 5.0 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.5, 182.9, 171.8, 136.6, 129.7, 128.9, 127.3, 57.4, 52.9, 41.4, 39.4, 27.5.

Methyl (2-((2-((S)-2-((2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4- dioxocyclobut-1-en-1-yl)amino)-3-phenylpropanamido)ethyl)amino)-3,4- dioxocyclobut-1-en-1-yl)-L-phenylalaninate

(ASq-2-Phe)



To a (S)-2-((2-((1-methoxy-1-oxo-3-phenylpropan-2-yl)amino)-3,4- dioxocyclobut-1-en-1-yl)amino)ethan-1-aminium 2,2,2-trifluoroacetate DMF solution (1.5 mL) (1 eq., 0.263 mmol, 0.1134 g), NMM (6 eq., 1.577 mmol, 0.173 mL) was added and stirred for 10 min. (2-((2-((tert-butoxycarbonyl)amino)ethyl)amino)-3,4-dioxocyclobut-1-en-1- yl)-L-phenylalanine (1 eq., 0.263 mmol, 0.106 g) was dissolved in DMF (2.5 mL) so the final concentration of 3.20 to react was to be close to 6.75 x 10-2 mmol/mL. The later solution was added into the former solution and PyBoP (3 eq., 0.789 mmol, 0.410 g) was added and the resulting solution was stirred under room temperature for 3 days. The solvent was removed in vacuo to yield a brown oil. 10 mL AcOEt was added to triturate the oil into a mixture of beige precipitate and yellow liquid. The mixture was sonicated and transferred into a centrifuge tube and centrifuged under 4000 RPM for 5 min. The liquid was poured out of the tube and a new 10 mL of AcOEt was added. The tube was sealed and rocked to fully mix the liquid and solid. This process was repeated for 3 times or until the mixture became slightly cloudy. The precipitate was dried in fume hood and then frozen in 5 mL H_2O and lyophilized in 2 days to yield a beige solid (0.4298 g, 23.3 %). **LC-MS** (0-100%) gradient of A in B, 55 min) t_R = 31 min. **HRMS (ESI)**: m/z Calc. for $C_{36}H_{42}N_6O_9$ 725.2905 [M+Na]+ , found 725.2922 (1.94 ppm). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.52 (s, 1H), 7.75 (ds, 2H), 7.67 – 7.56 (ds, 2H), 7.31 – 7.26 (m, 2H), 7.24 (m, 2H), 7.21 (m, 1H), 7.19 (m, 1H), 7.18 (m, 2H), 7.16 (m, 2H), 6.90 (s, 1H), 5.02 (s, 1H), 4.86 (s, 1H), 3.67 (s, 3H), 3.52 (m, 2H), 3.50 (m, 2H), 3.25 (s, 2H), 3.14 (s, 1H), 3.07 – 3.03 (m, 4H), 2.85 (m, 1H), 1.36 (s, 9H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 183.5, 183.1, 182.7, 182.5, 171.7, 171.0, 168.5, 167.3, 156.2, 136.4, 129.9, 129.8, 128.9, 128.6, 127.3, 128.0, 78.3, 57.8, 57.3, 52.8, 43.4, 43.0, 41.6, 41.1, 40.3, 39.6, 28.7.

2. Spectroscopic Characterisation Data

Characterisation and Spectroscopic data of cyclic Sq-2-AA dimers

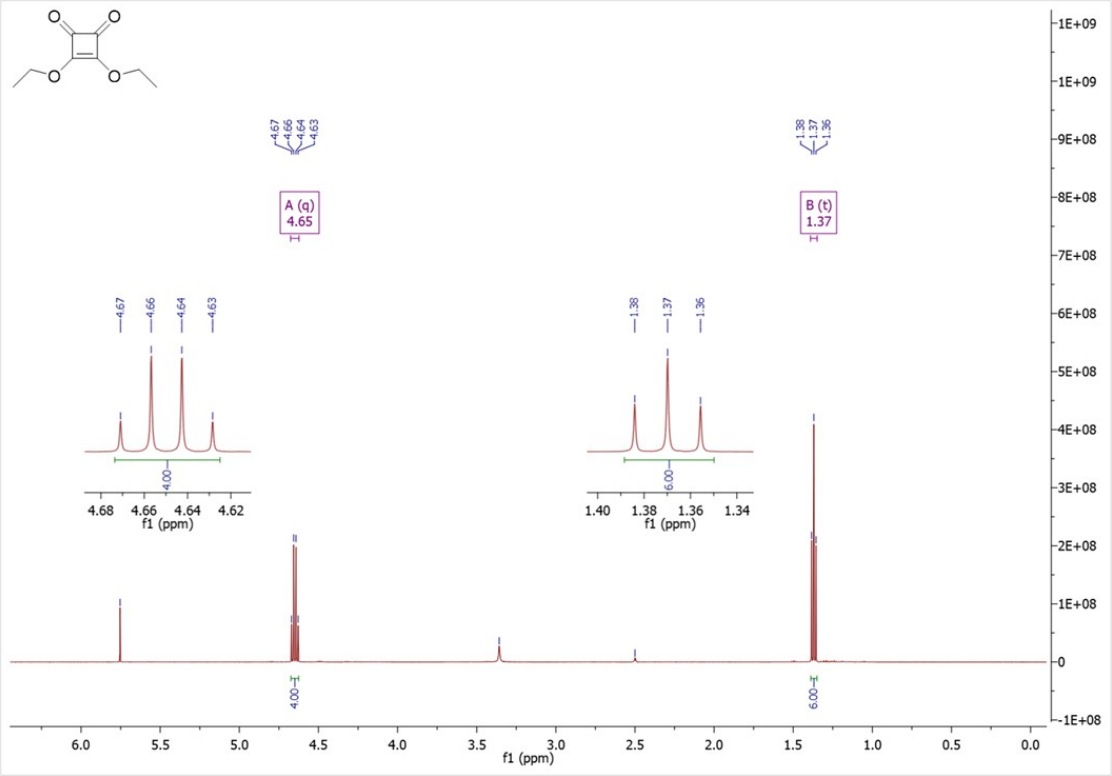


Figure S1: ^{1}H NMR spectrum of **DeSq** in DMSO-d_6.

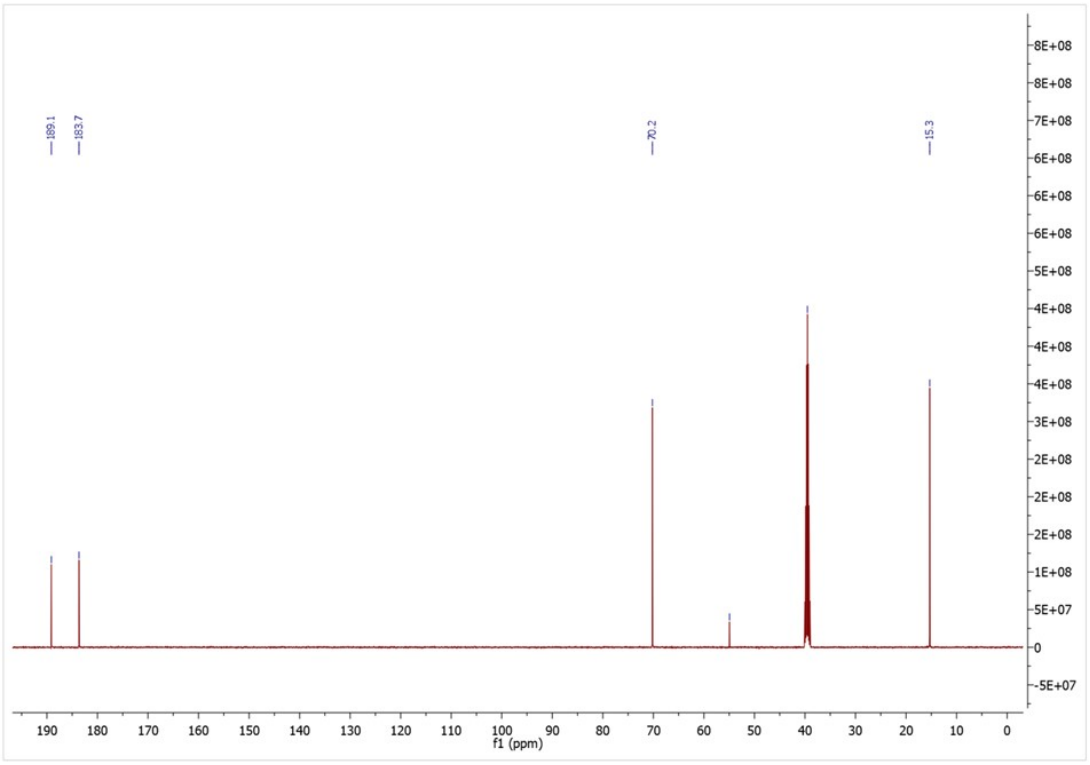


Figure S2: ^{13}C NMR spectrum of **DeSq** in DMSO-d_6.

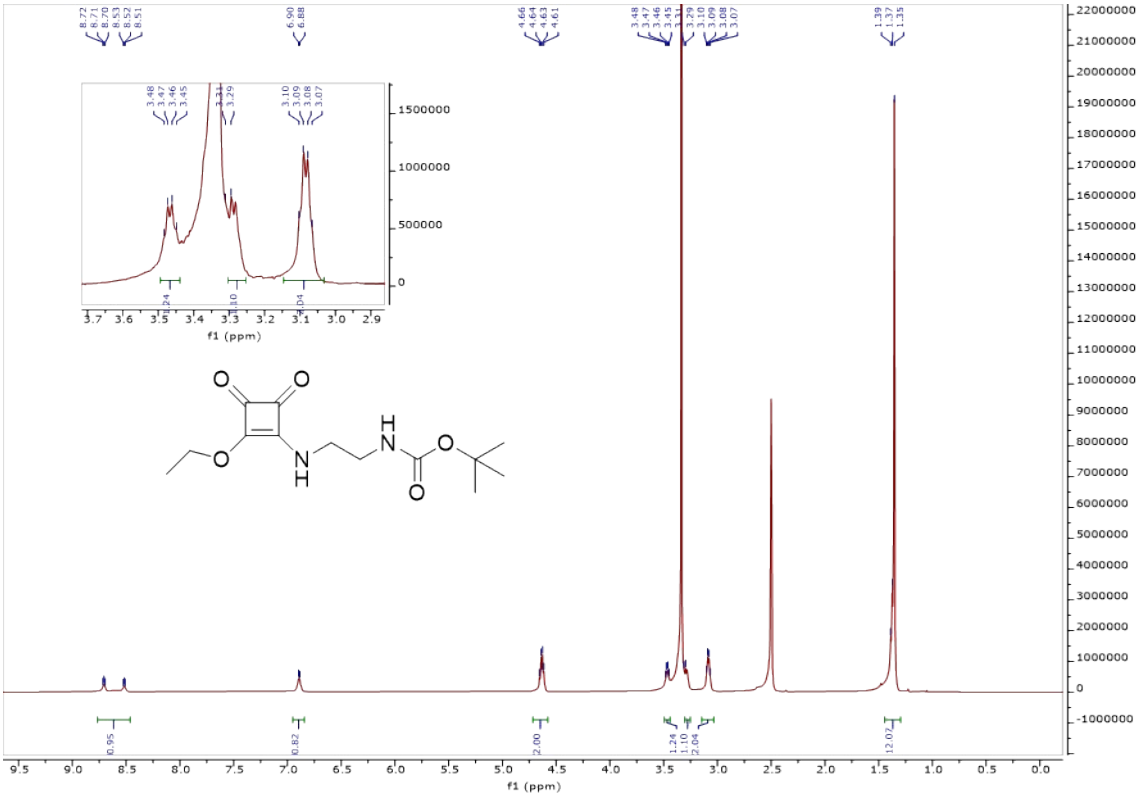


Figure AS3: ^{1}H NMR spectrum of **A** in DMSO-d_6.

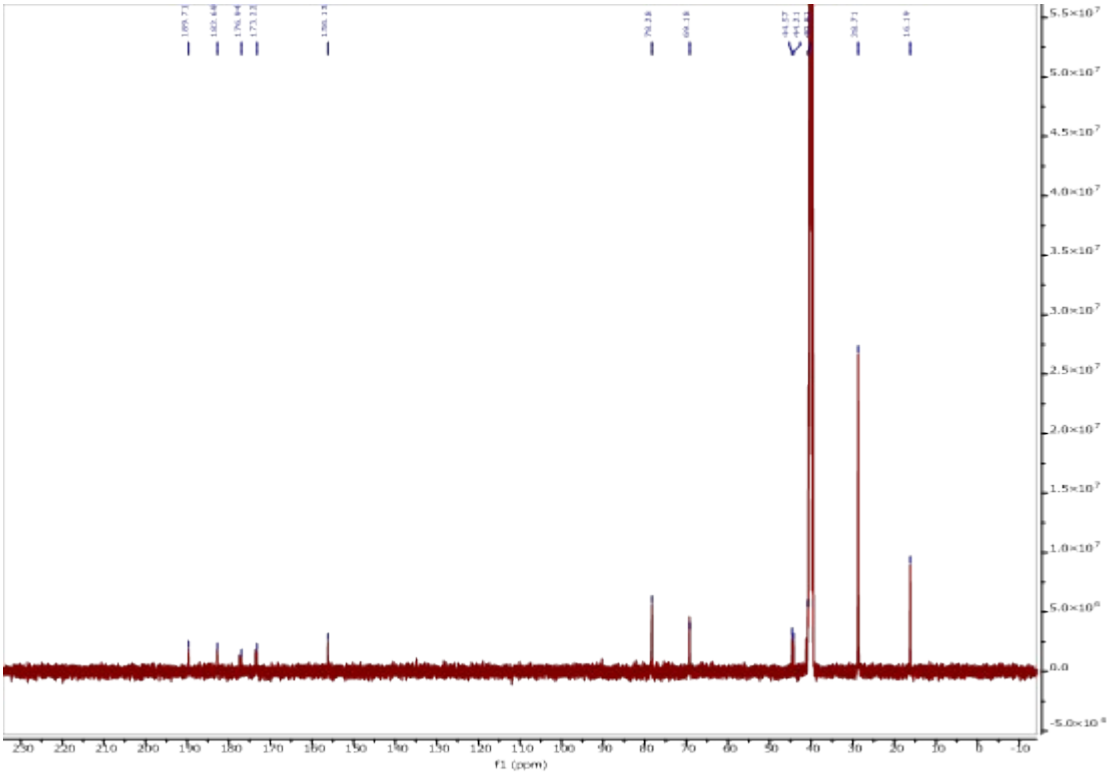


Figure S4: ^{13}C NMR spectrum of **A** in DMSO-d_6.

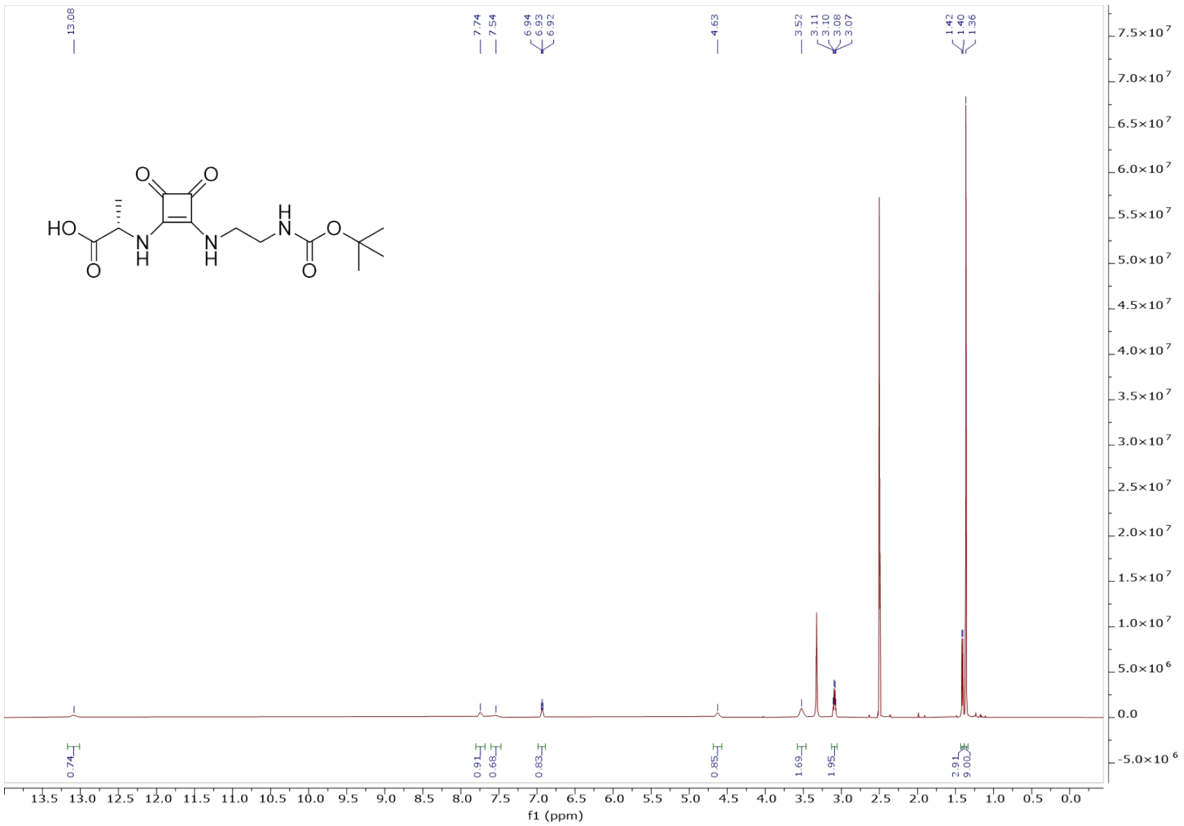


Figure S5: ^{1}H NMR spectrum of **B-Ala** in DMSO-d_6.

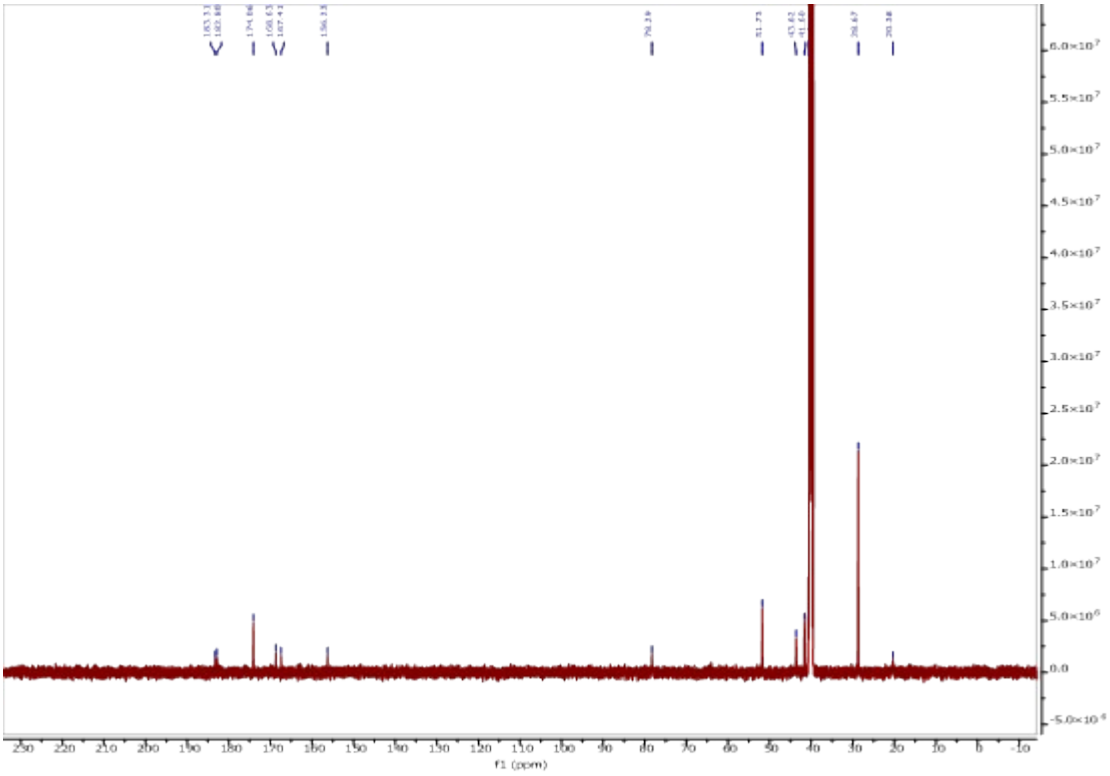


Figure S6: ^{13}C NMR spectrum of **B-Ala** in DMSO-d_6.

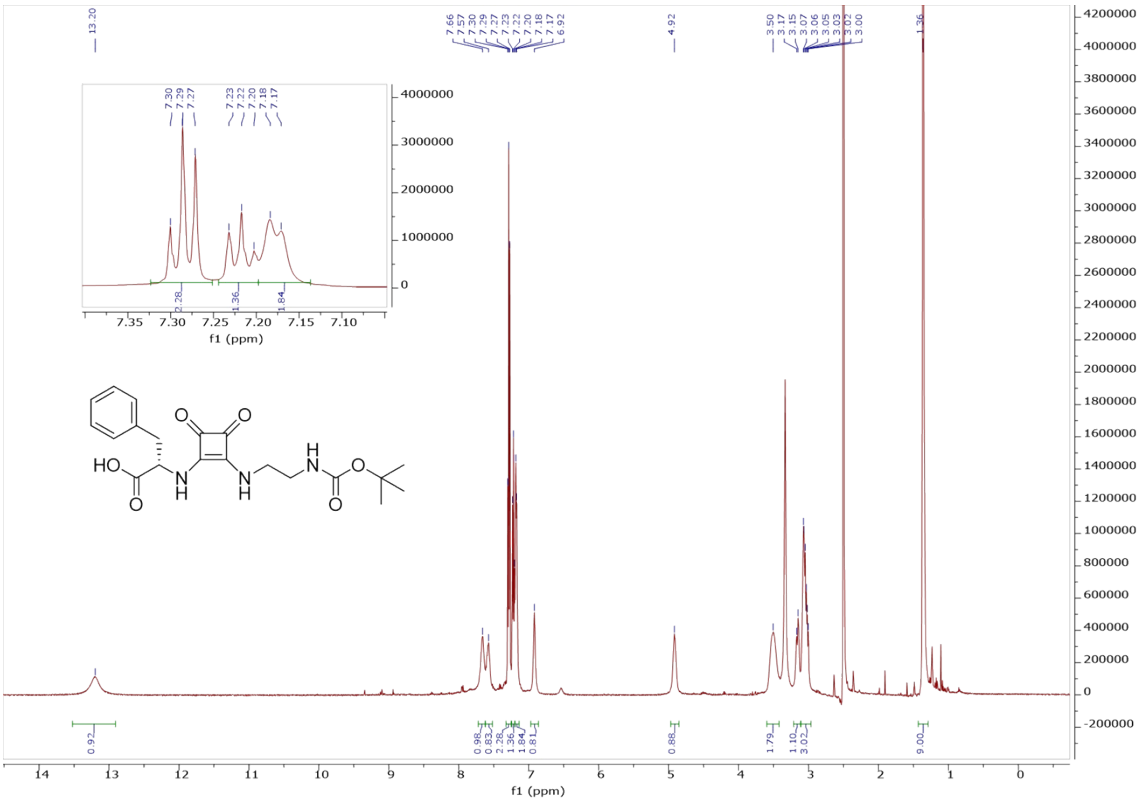


Figure S7: ^{1}H NMR spectrum of **B-Phe** in DMSO-d_6.

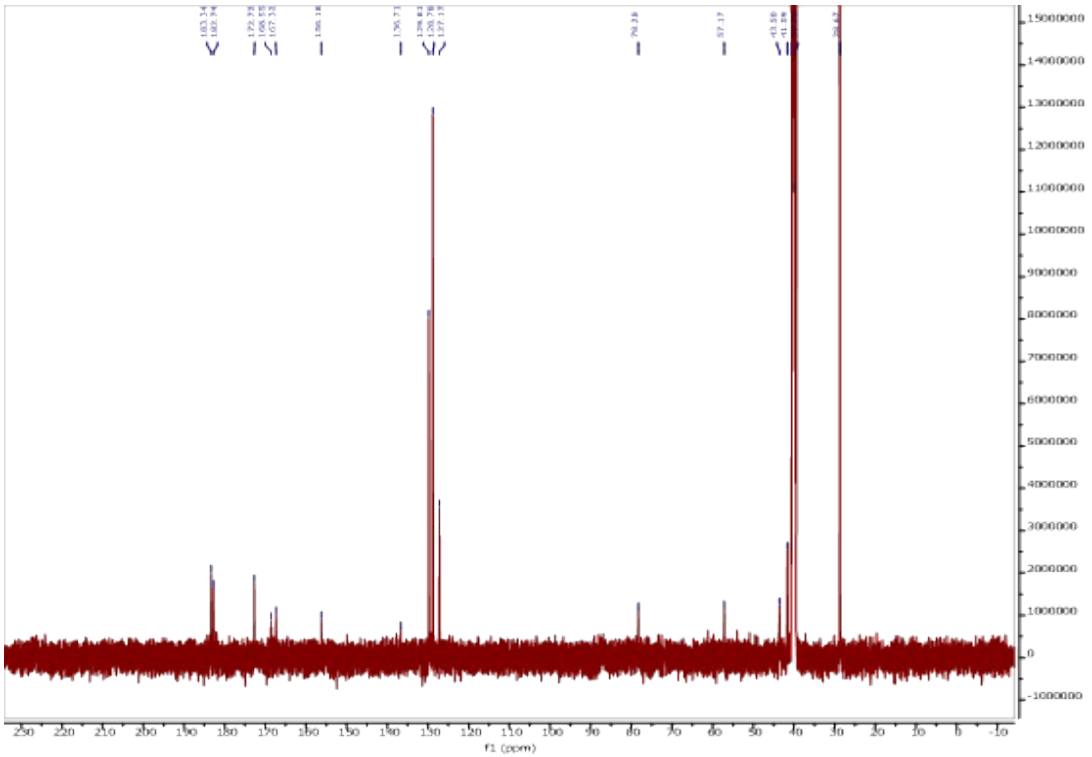


Figure S8: ^{13}C NMR spectrum of **B-Phe** in DMSO-d_6.

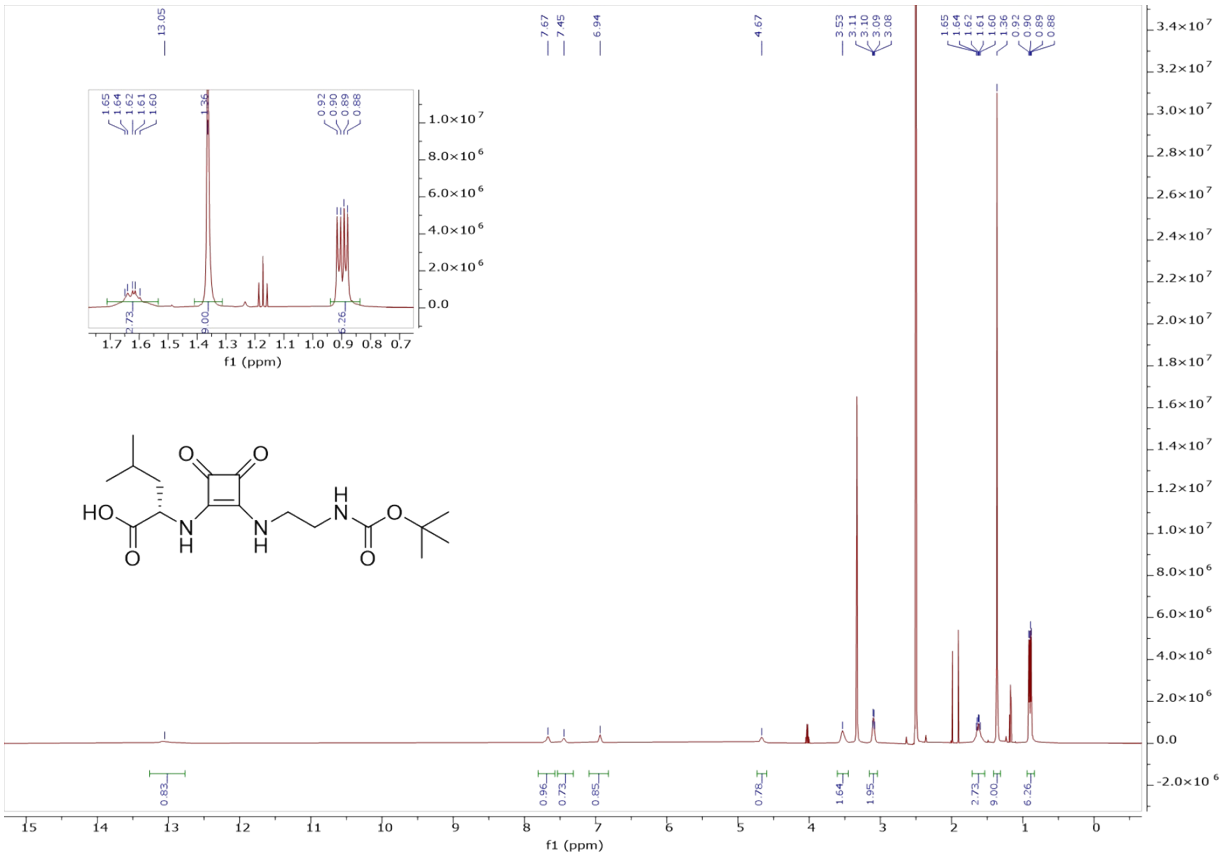


Figure S9: ^{1}H NMR spectrum of **B-Leu** in DMSO-d_6.

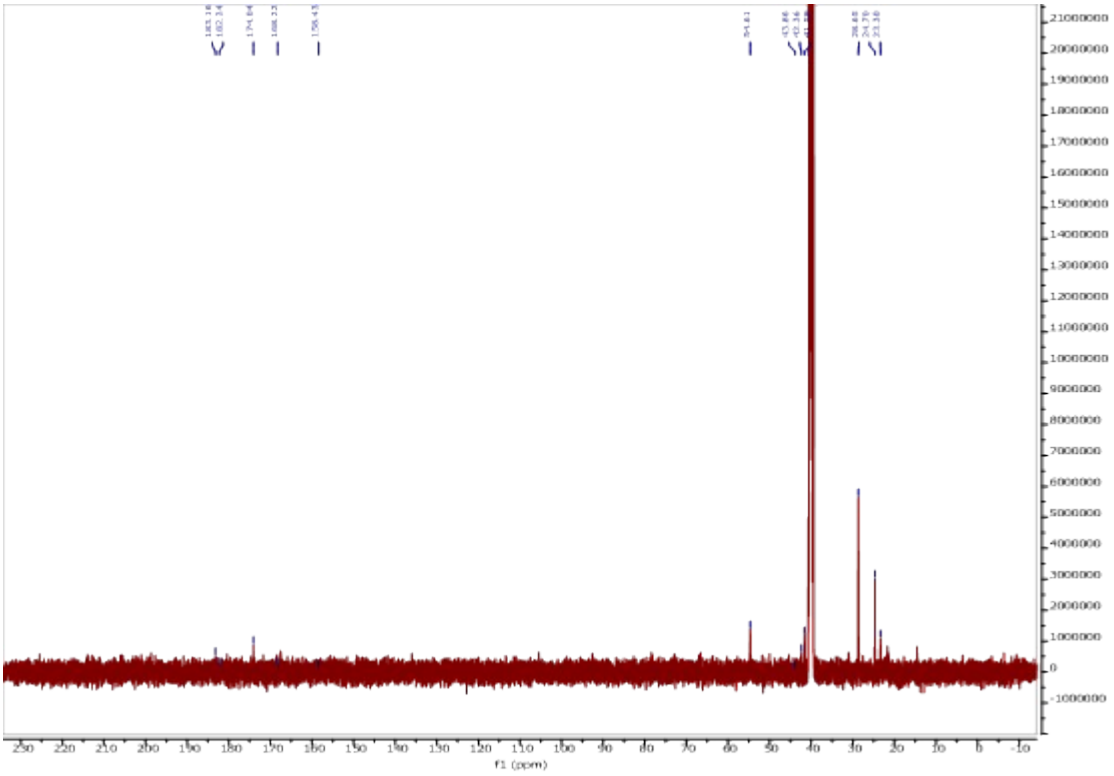


Figure S10: ^{13}C NMR spectrum of **B-Leu** in DMSO-d_6.

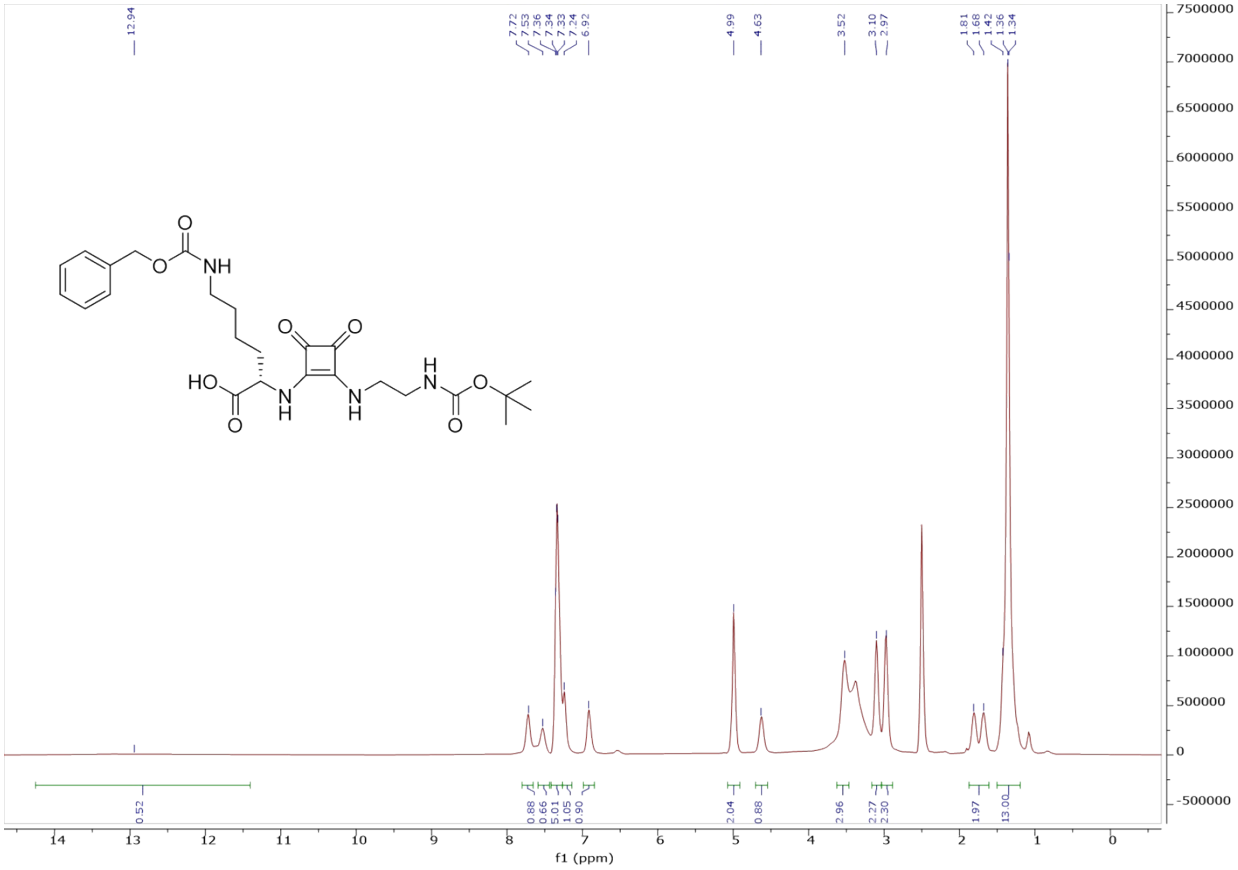


Figure S11: ^{1}H NMR spectrum of **B-Lys(Cbz)** in DMSO-d_6.

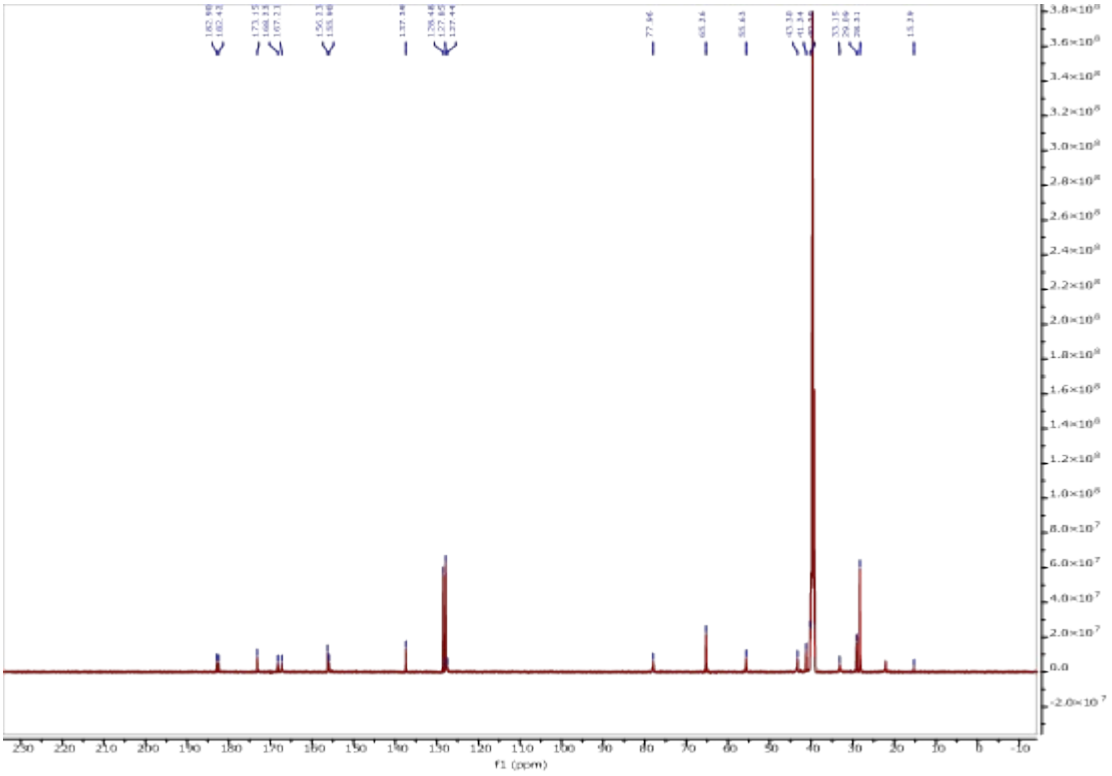


Figure S12: ^{13}C NMR spectrum of **B-Lys(Cbz)** in DMSO-d_6.

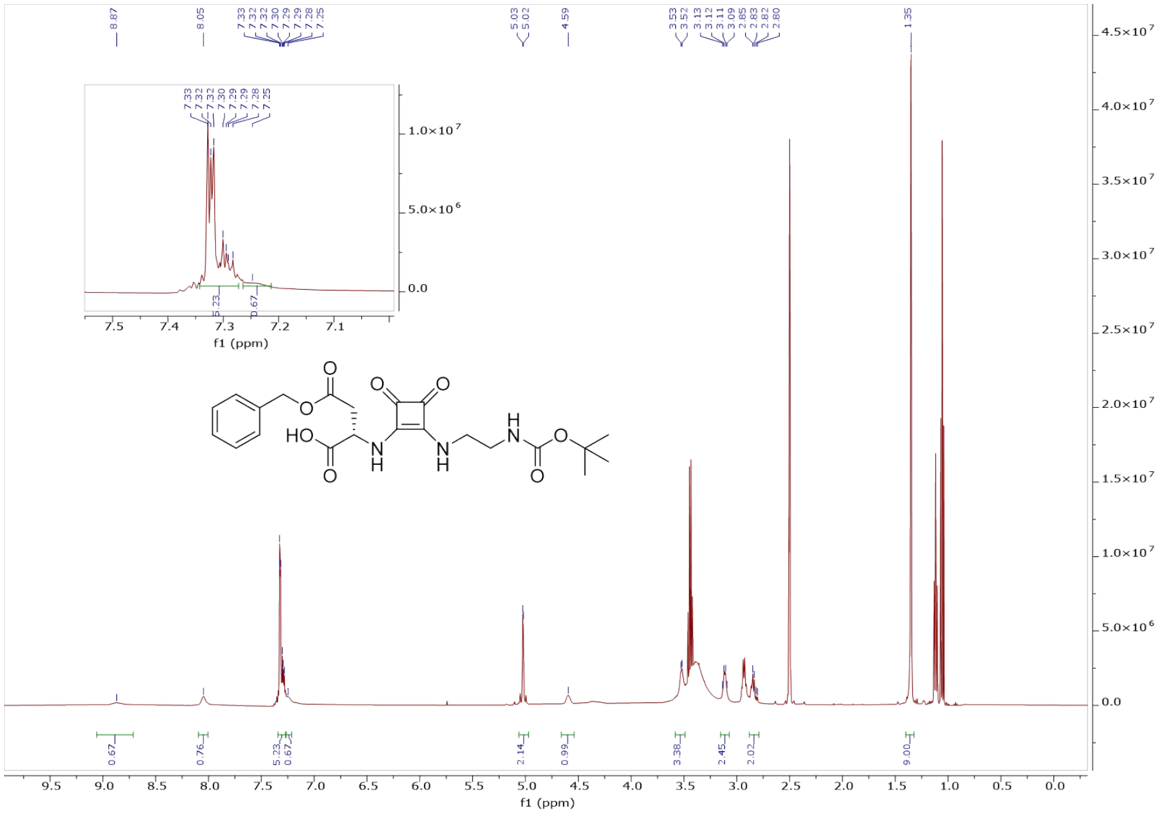


Figure S13: ^{1}H NMR spectrum of **B-Asp(Bzl)** in DMSO-d_6.

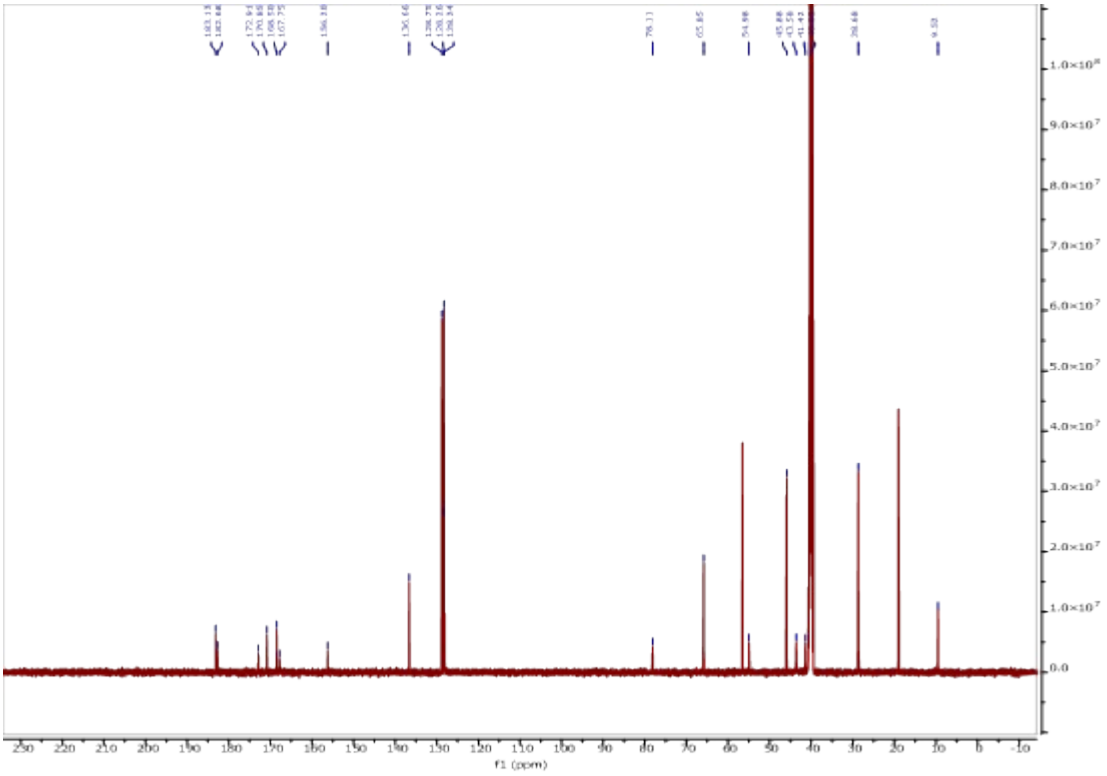


Figure S13: ^{13}C NMR spectrum of **B-Asp(Bzl)** in DMSO-d_6.

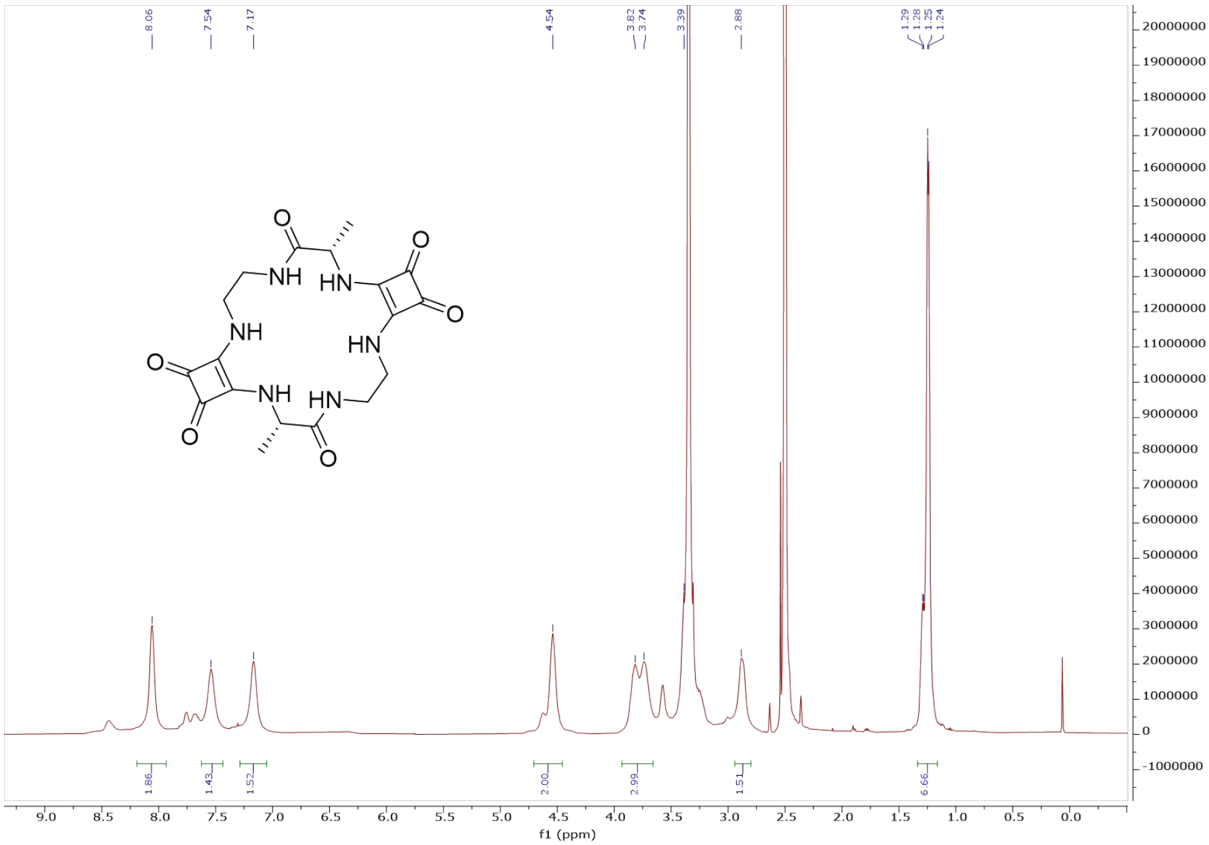


Figure S14: ^{1}H NMR spectrum of **Sq-2-Ala** in DMSO-d_6.

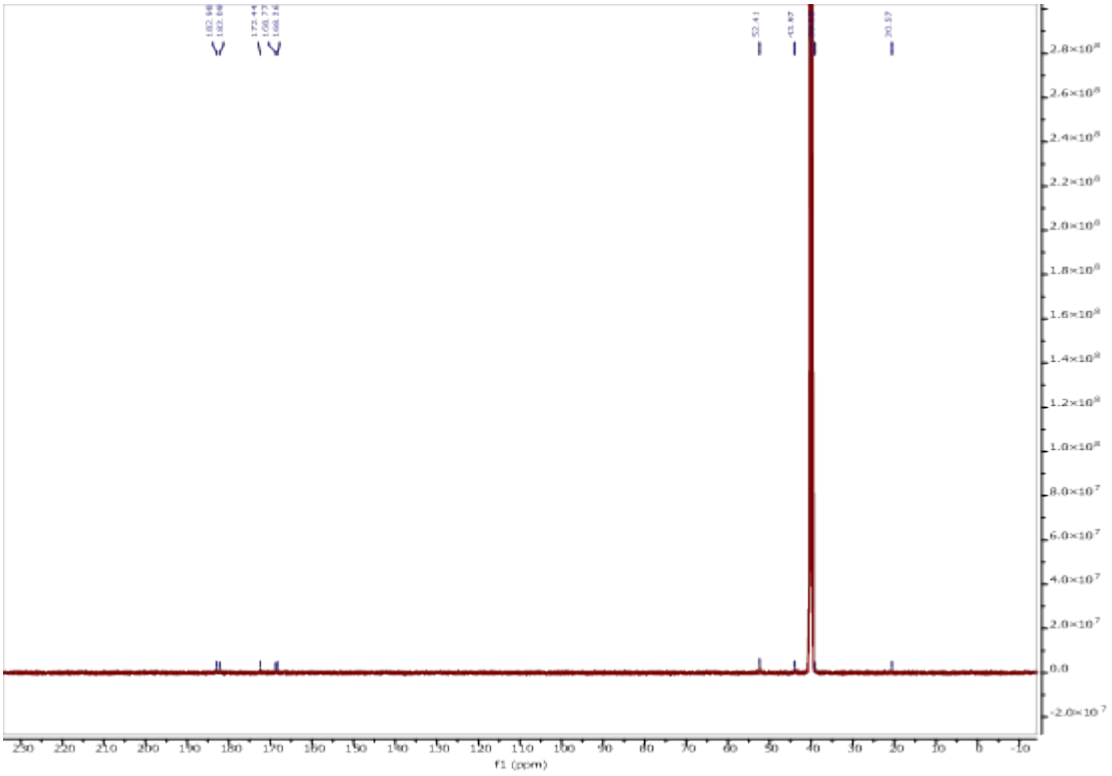


Figure S15: ^{13}C NMR spectrum of **Sq-2-Ala** in DMSO-d_6.

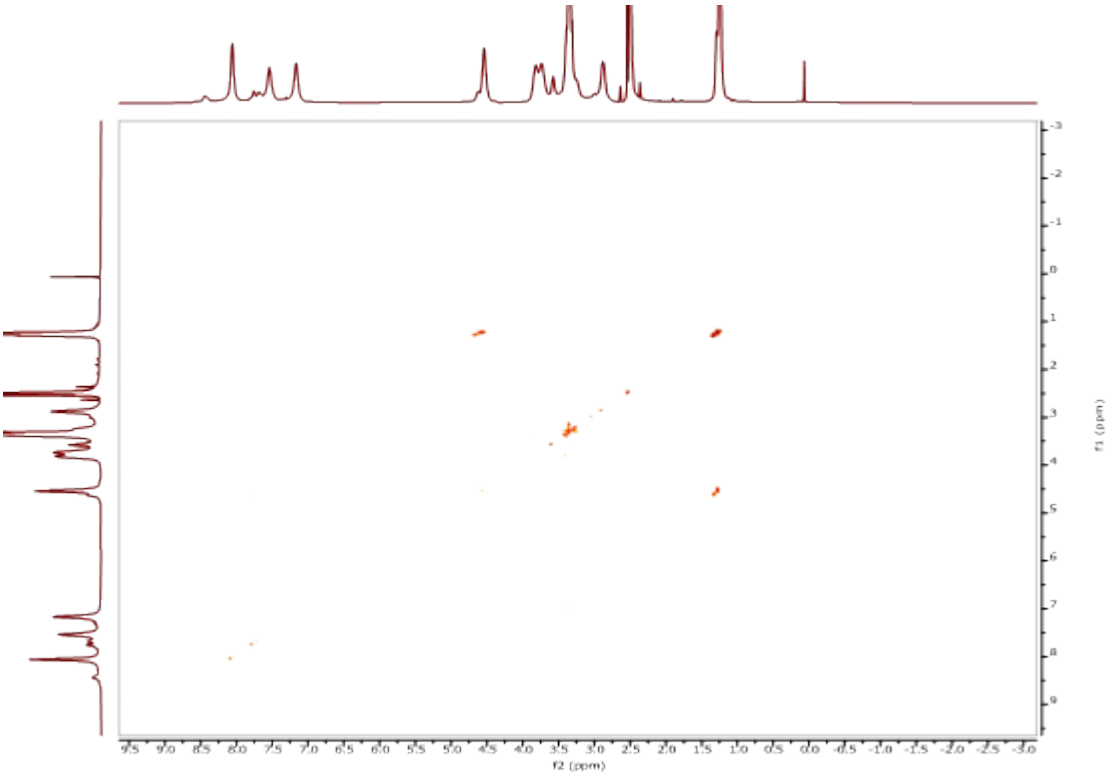


Figure S16: COSY spectrum of **Sq-2-Ala** in DMSO-d_6.

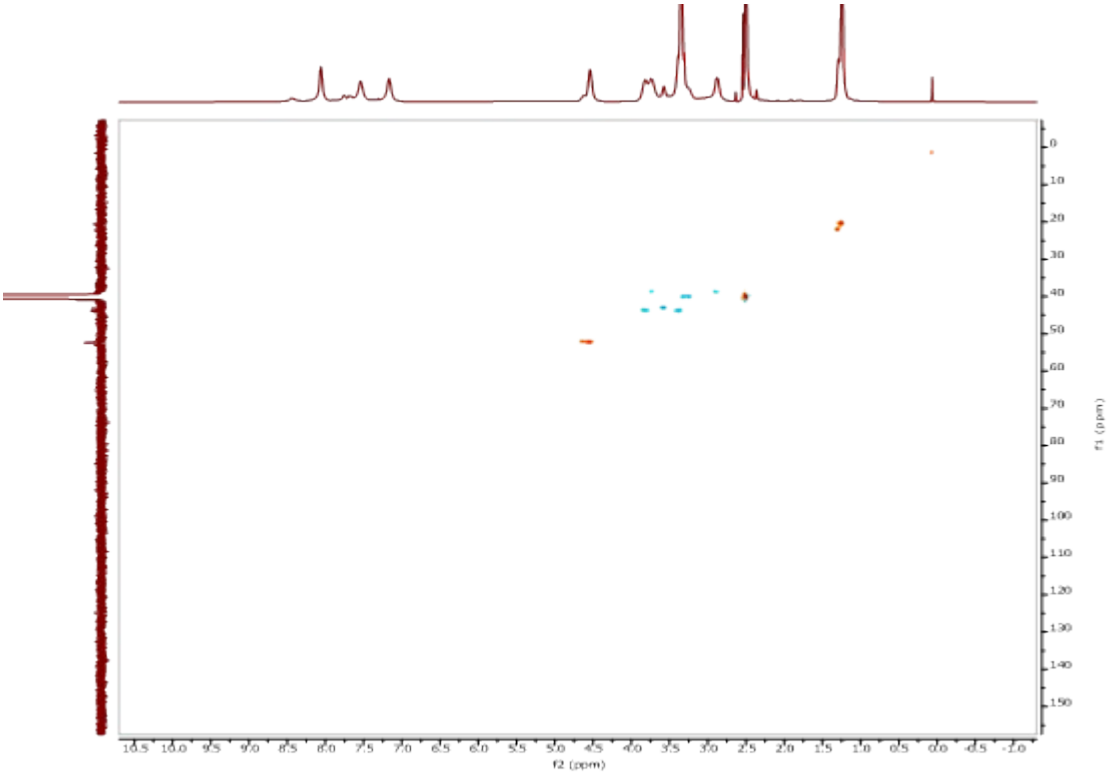


Figure S17: HSQC spectrum of **Sq-2-Ala** in DMSO-d_6.

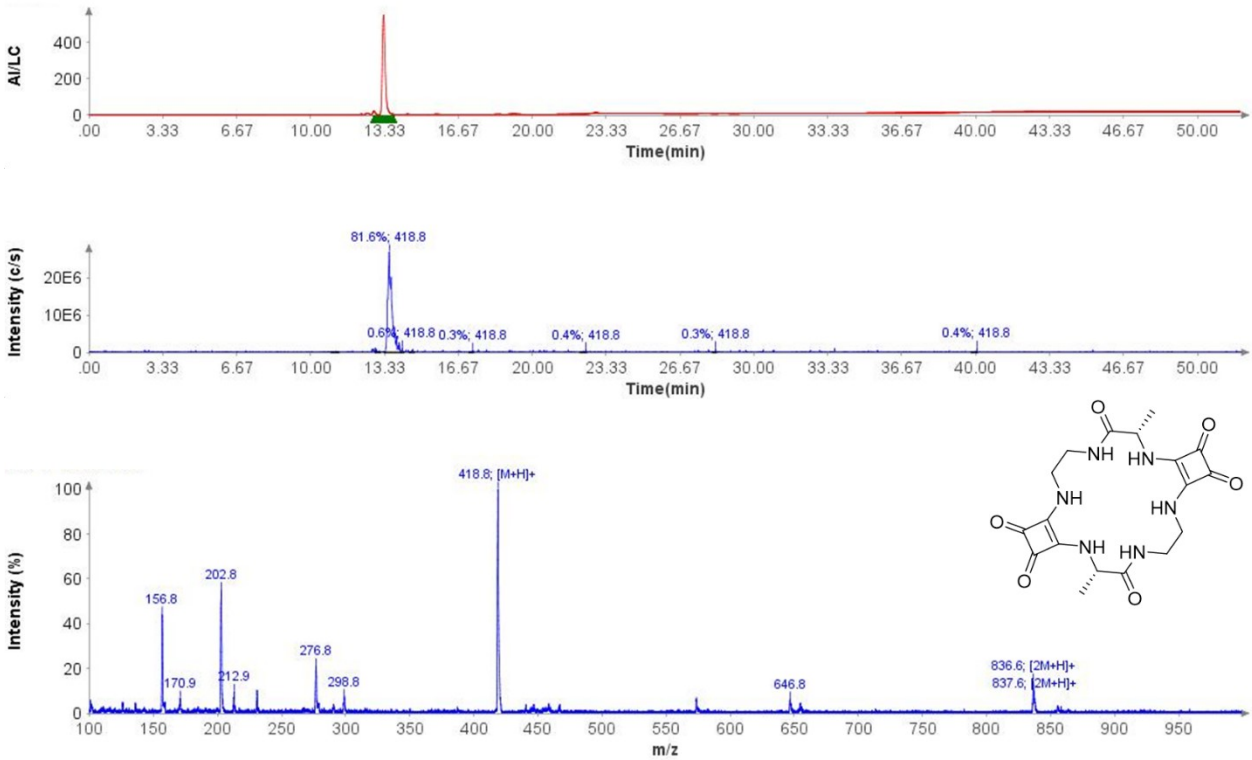


Figure S18: LCMS of **Sq-2-Ala**

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C18 H22 N6 O6	0.69	441.1500	418.1608	418.1601	1.68	97.96

Figure: Extracted ion chromatogram (EIC) of compound.

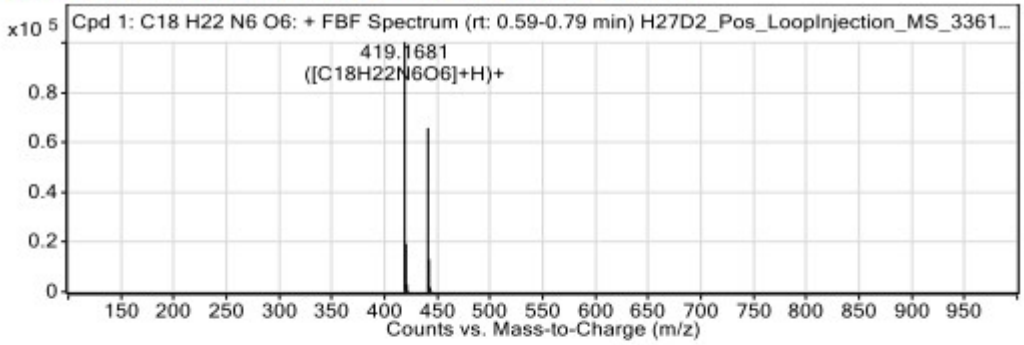


Figure: Full range view of Compound spectra and potential adducts.

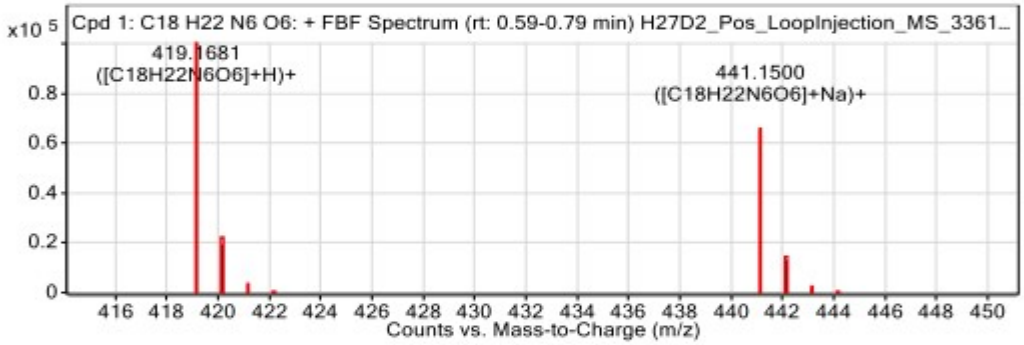


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
419.1681	1	100720.4	C18H22N6O6	(M+H)+
420.1706	1	19354.7	C18H22N6O6	(M+H)+
421.1730	1	3027.0	C18H22N6O6	(M+H)+
422.1768	1	449.2	C18H22N6O6	(M+H)+
441.1500	1	65932.4	C18H22N6O6	(M+Na)+
442.1529	1	13102.1	C18H22N6O6	(M+Na)+
443.1548	1	1856.9	C18H22N6O6	(M+Na)+
444.1550	1	334.9	C18H22N6O6	(M+Na)+

Figure S19: HRMS data of **Sq-2-Ala.**

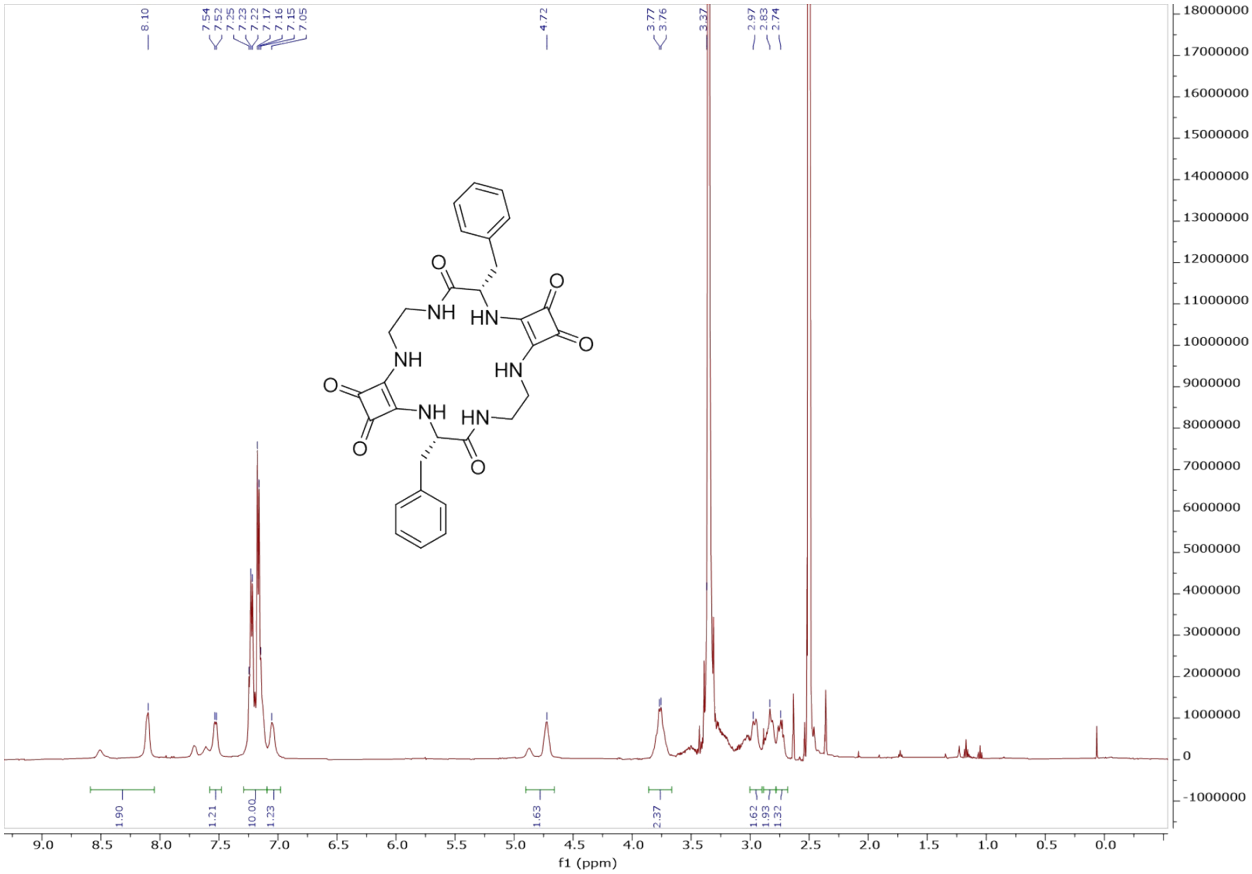


Figure S20: ^{1}H NMR spectrum of **Sq-2-Phe** in DMSO-d_6.

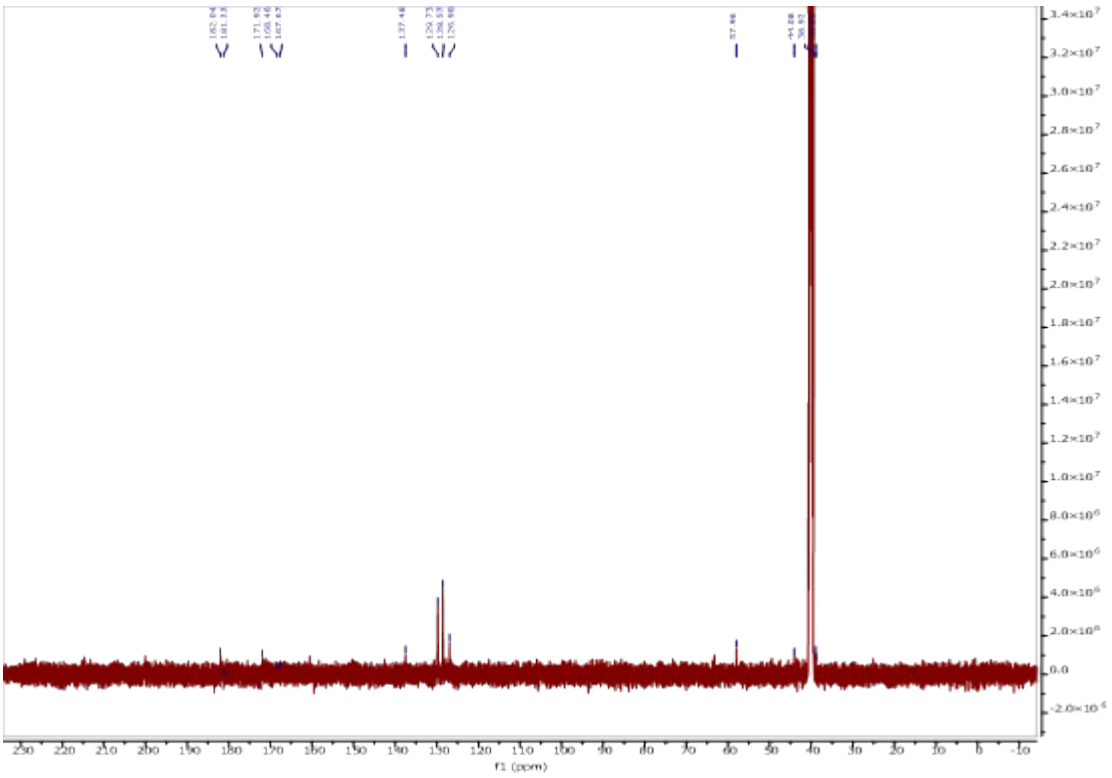


Figure S21: ^{13}C NMR spectrum of **Sq-2-Phe** in DMSO-d_6.

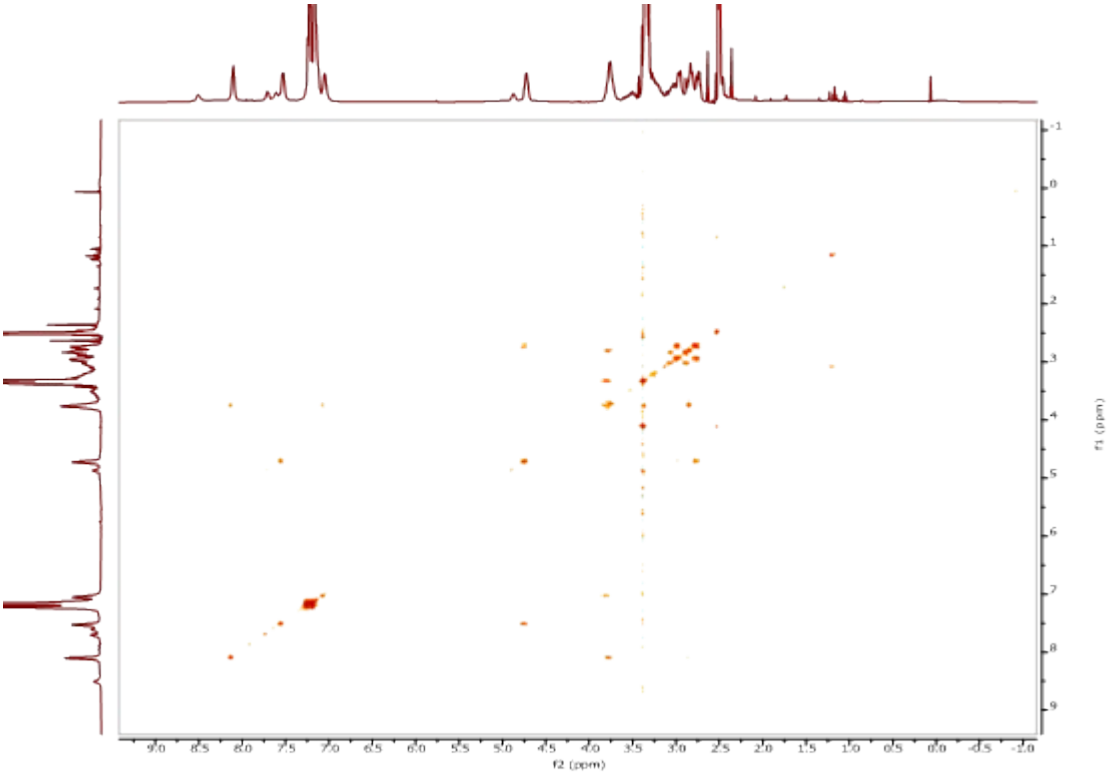


Figure AS22: COSY spectrum of **Sq-2-Phe** in DMSO-d_6.

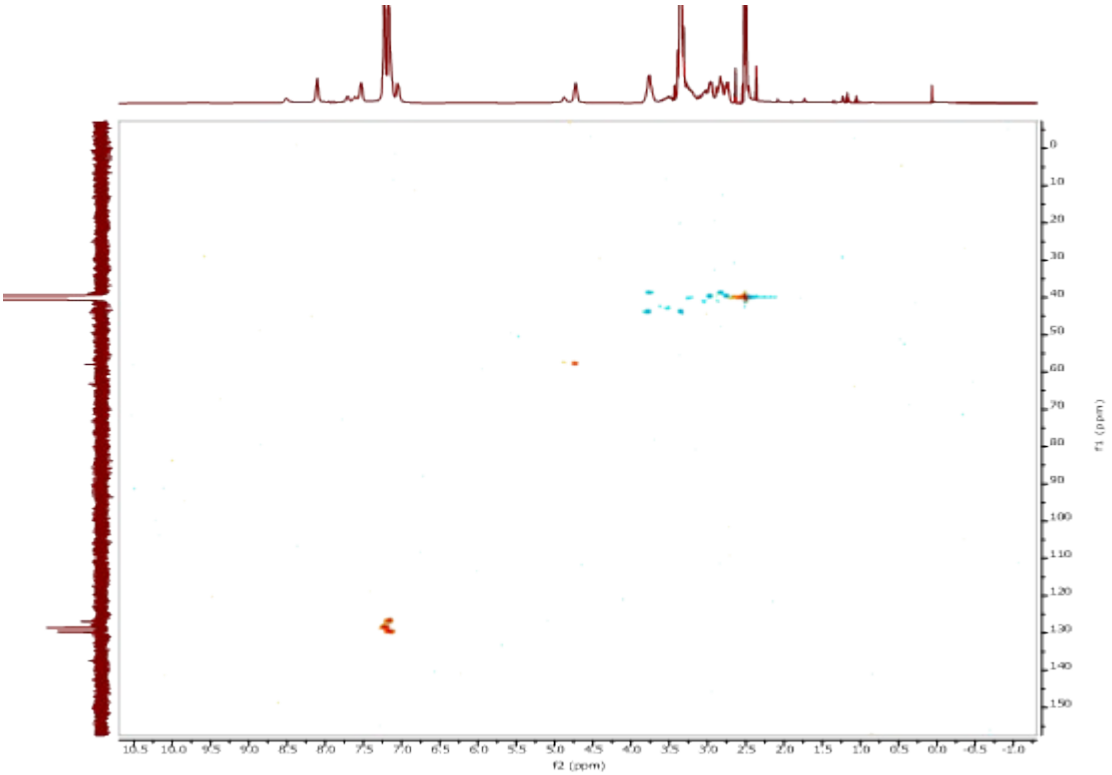


Figure S23: HSQC spectrum of **Sq-2-Phe** in DMSO-d_6.

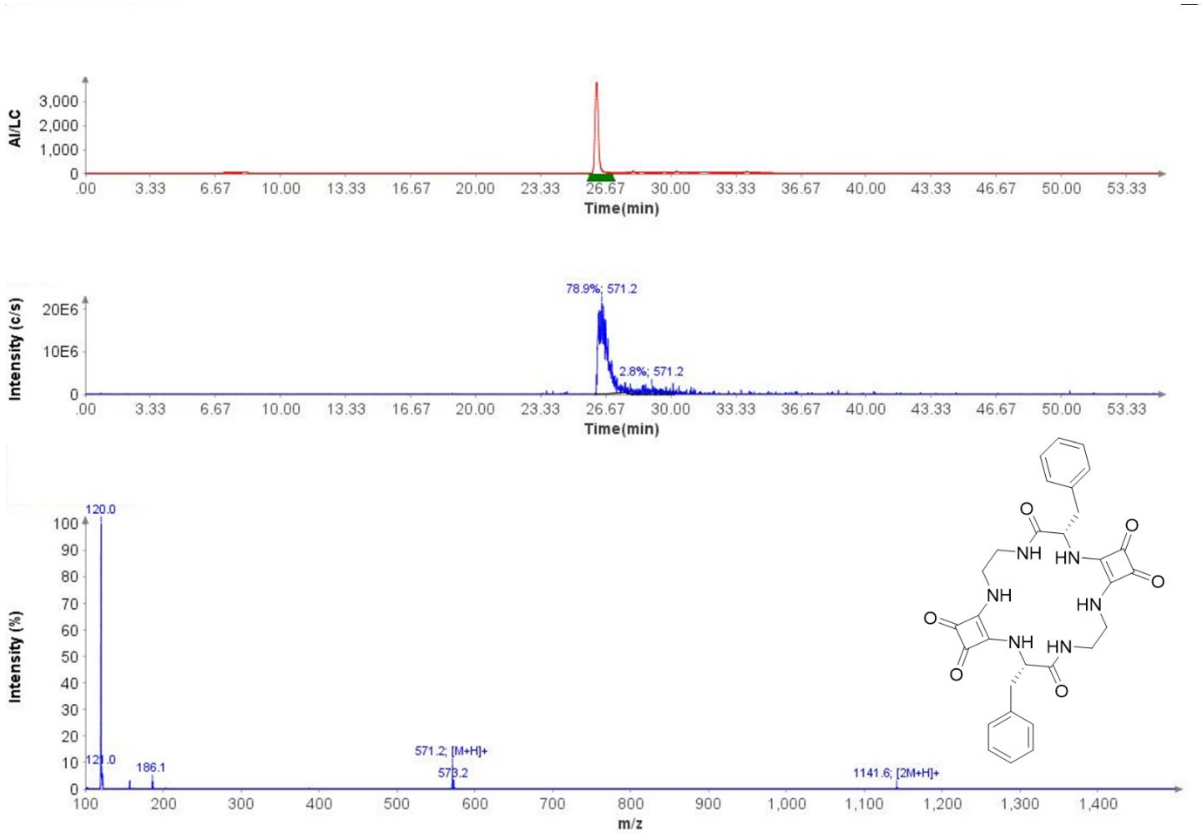


Figure S24: LCMS data of **Sq-2-Phe**.

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C30 H30 N6 O6	0.71	571.2309	570.2235	570.2227	1.38	98.59

Figure: Extracted ion chromatogram (EIC) of compound.

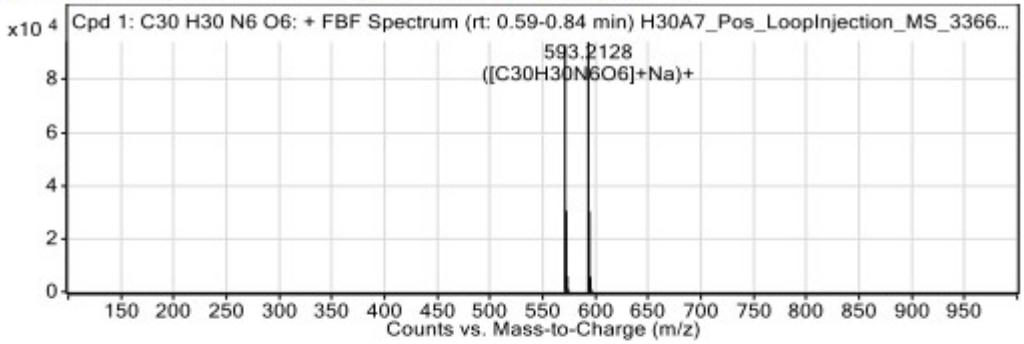


Figure: Full range view of Compound spectra and potential adducts.

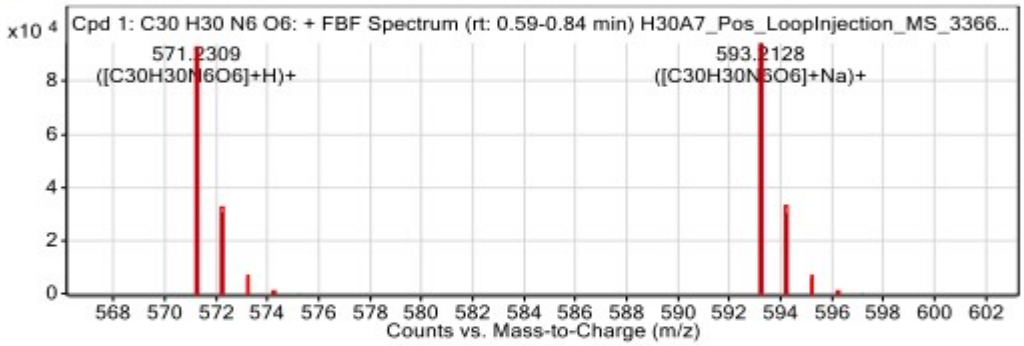


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
571.2309	1	93039.2	C30H30N6O6	(M+H)+
572.2337	1	30574.3	C30H30N6O6	(M+H)+
573.2362	1	5949.9	C30H30N6O6	(M+H)+
574.2380	1	944.1	C30H30N6O6	(M+H)+
575.2424	1	154.9	C30H30N6O6	(M+H)+
593.2128	1	94281.0	C30H30N6O6	(M+Na)+
594.2154	1	30331.3	C30H30N6O6	(M+Na)+
595.2179	1	5687.7	C30H30N6O6	(M+Na)+
596.2209	1	880.1	C30H30N6O6	(M+Na)+
597.1953	1	149.3	C30H30N6O6	(M+Na)+

Figure S25: HRMS data of **Sq-2-Phe.**

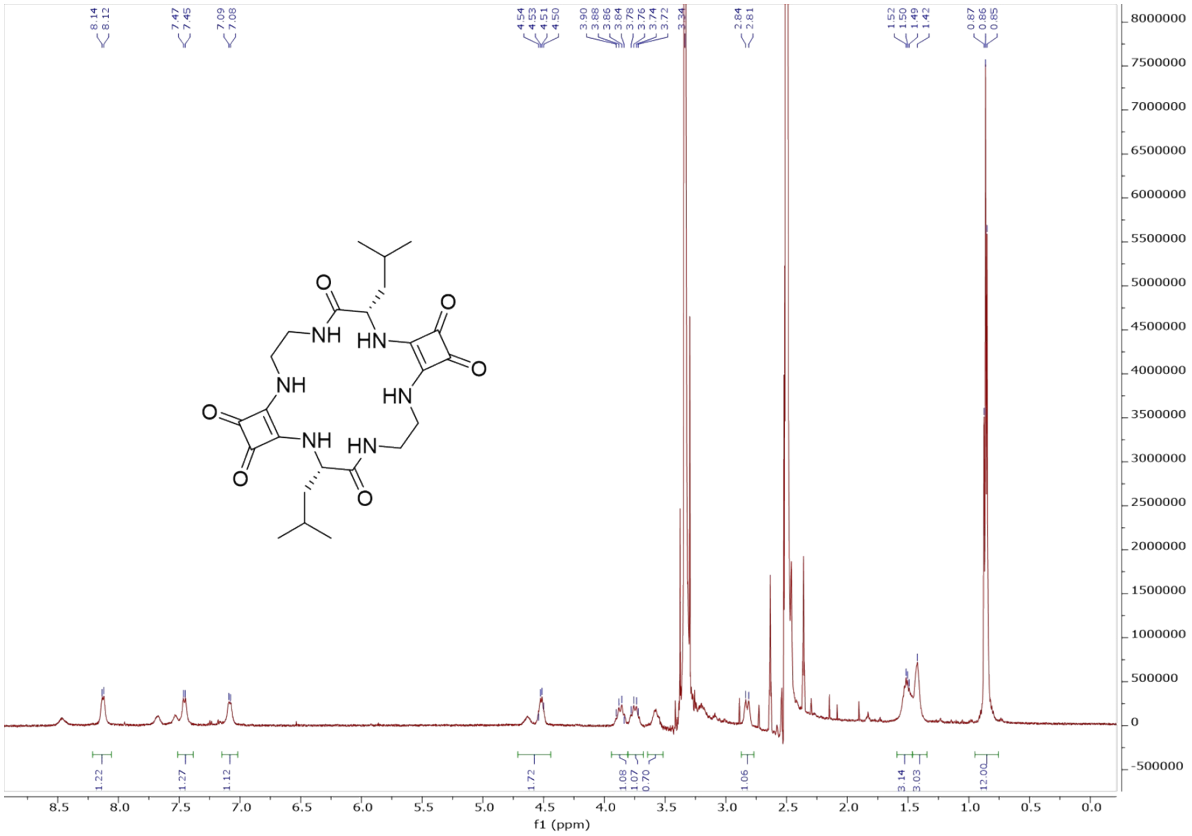


Figure S26: ^{1}H NMR spectrum of **Sq-2-Leu** in DMSO-d_6.

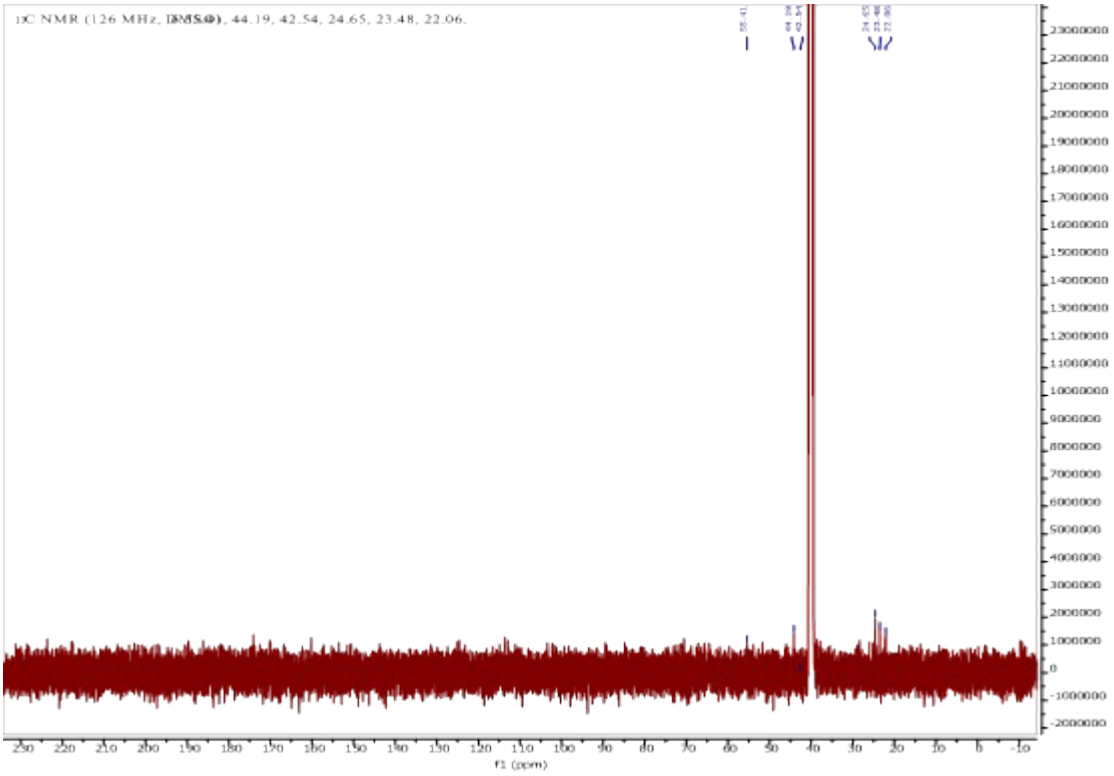


Figure S27: ^{13}C NMR spectrum of **Sq-2-Leu** in DMSO-d_6.

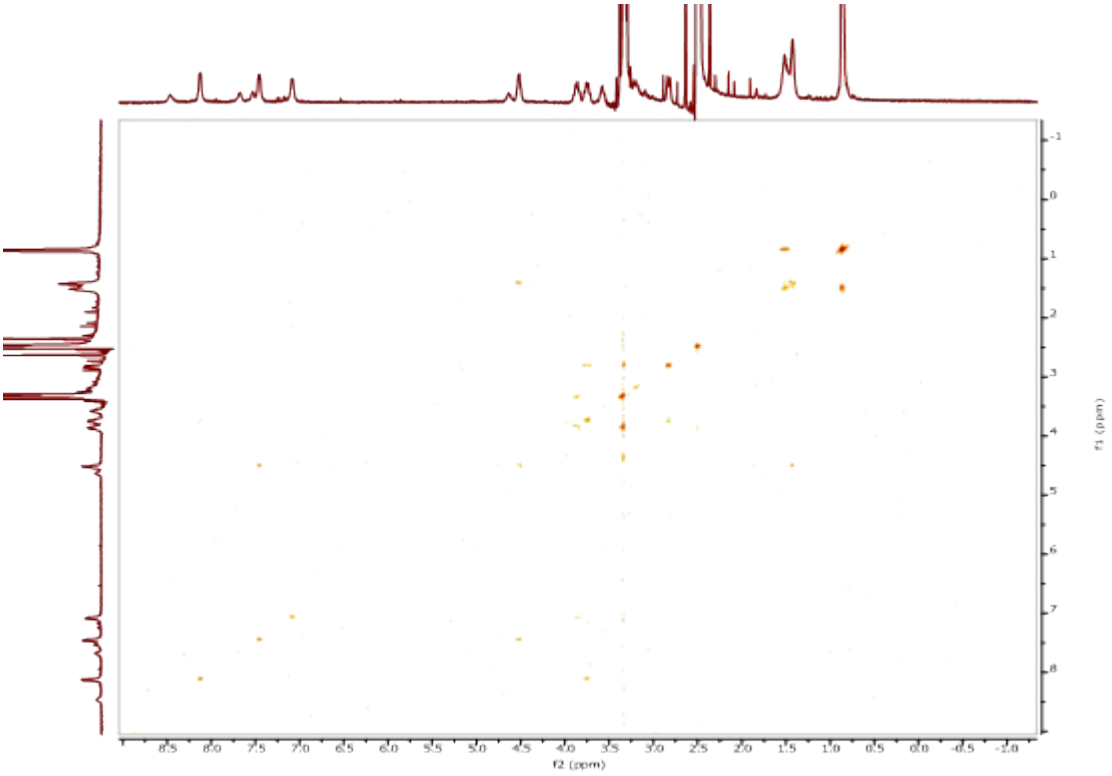


Figure S28: COSY spectrum of **Sq-2-Leu** in DMSO-d_6.

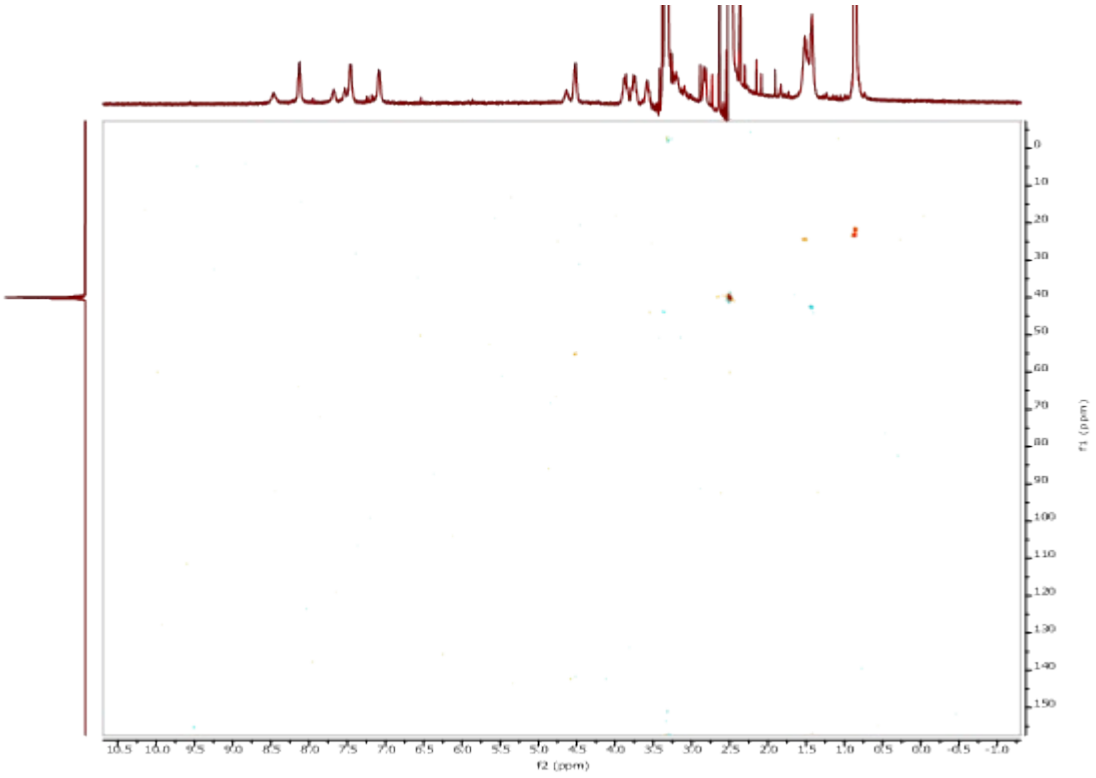


Figure S29: HSQC spectrum of **Sq-2-Leu** in DMSO-d_6.

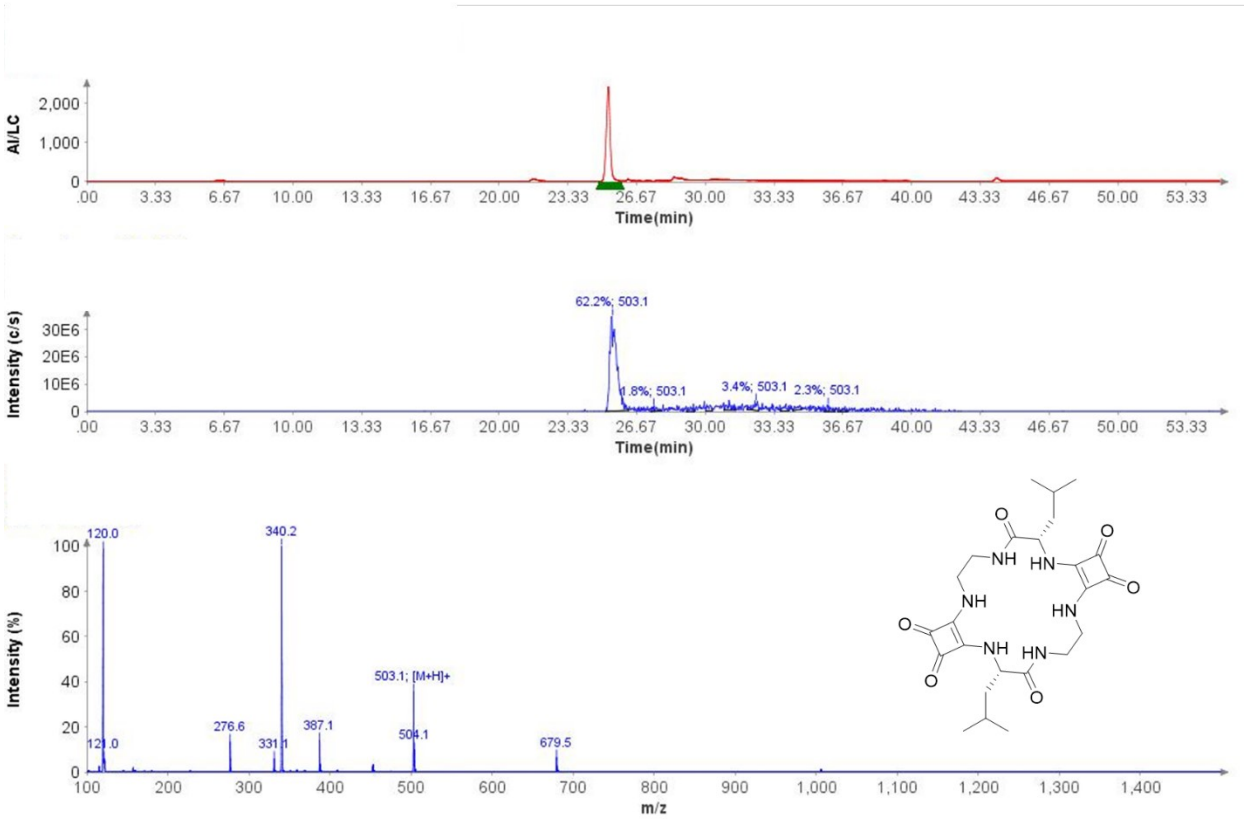


Figure S30: LCMS data of **Sq-2-Leu.**

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C24 H34 N6 O6	0.69	503.2621	502.2548	502.2540	1.61	98.86

Figure: Extracted ion chromatogram (EIC) of compound.

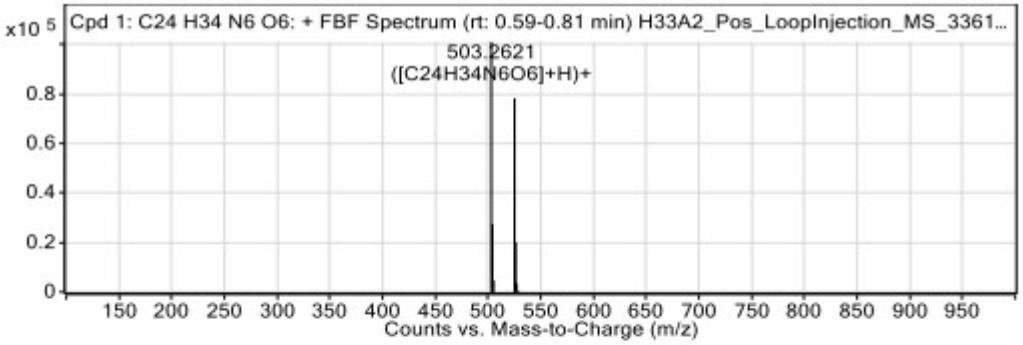


Figure: Full range view of Compound spectra and potential adducts.

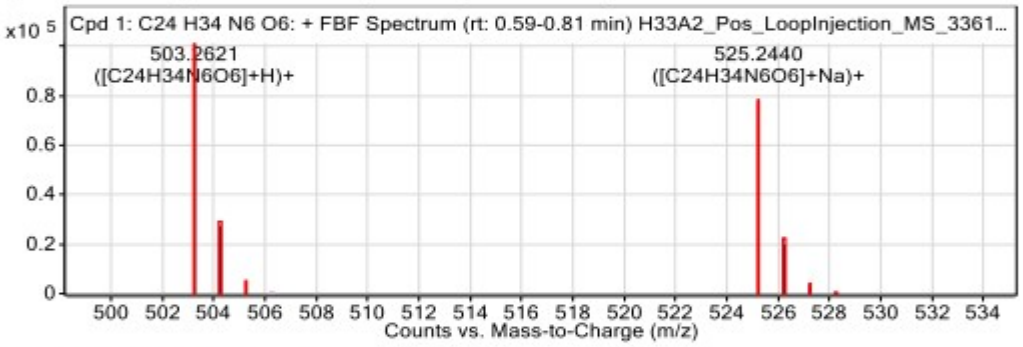


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
503.2621	1	101174.8	C24H34N6O6	(M+H)+
504.2649	1	27291.0	C24H34N6O6	(M+H)+
505.2682	1	4574.2	C24H34N6O6	(M+H)+
525.2440	1	78263.4	C24H34N6O6	(M+Na)+
526.2465	1	19998.0	C24H34N6O6	(M+Na)+
527.2492	1	3481.3	C24H34N6O6	(M+Na)+
528.2534	1	525.3	C24H34N6O6	(M+Na)+

Figure S31: HRMS data of **Sq-2-Leu.**

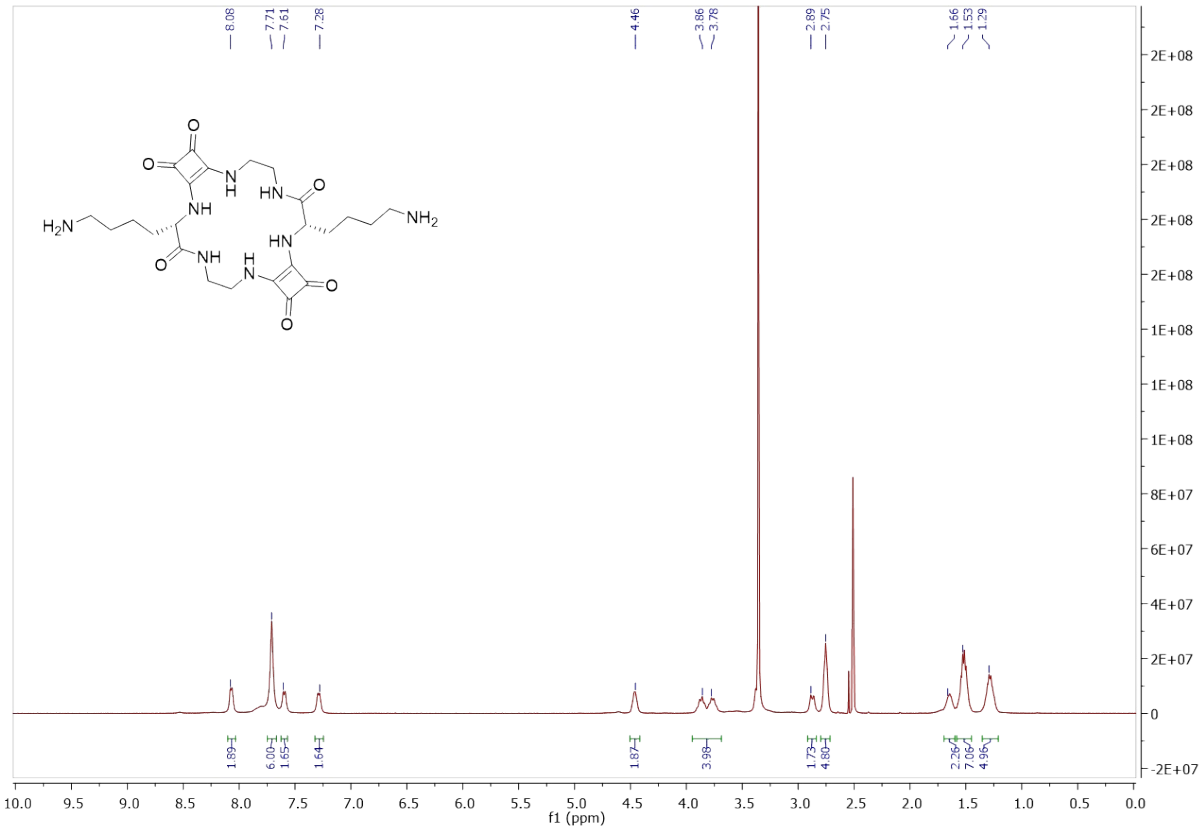


Figure S32: ^{1}H NMR spectrum of **Sq-2-Lys** in DMSO-d_6.

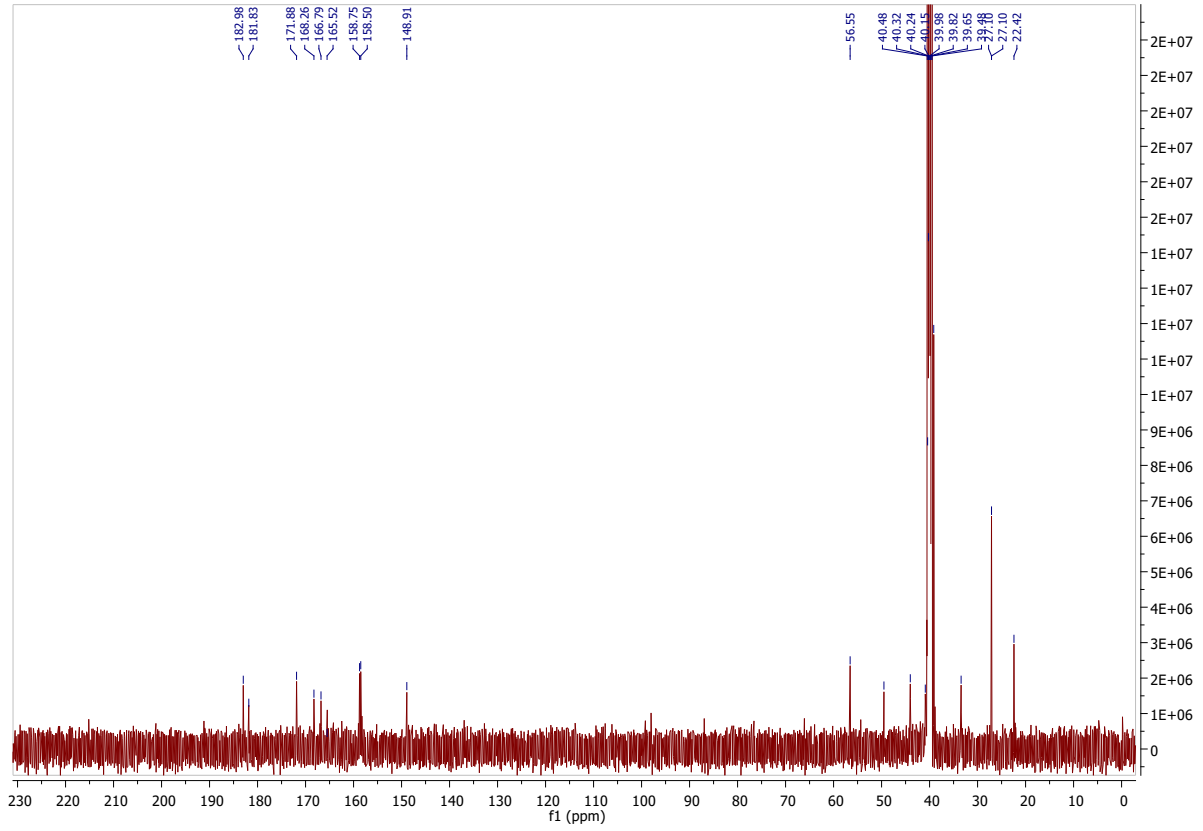


Figure S33: ^{13}C NMR spectrum of **Sq-2-Lys** in DMSO-d_6.

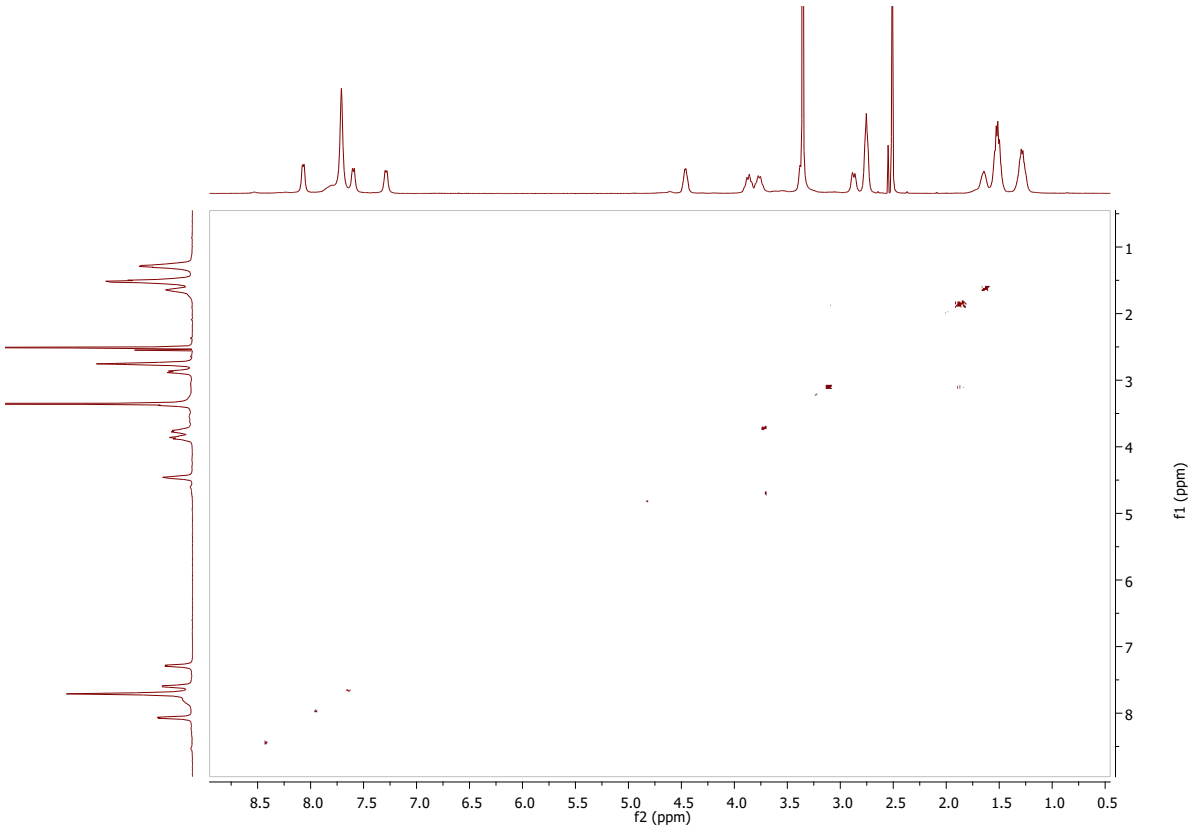


Figure S35: COSY spectrum of **Sq-2-Lys** in DMSO-d_6.

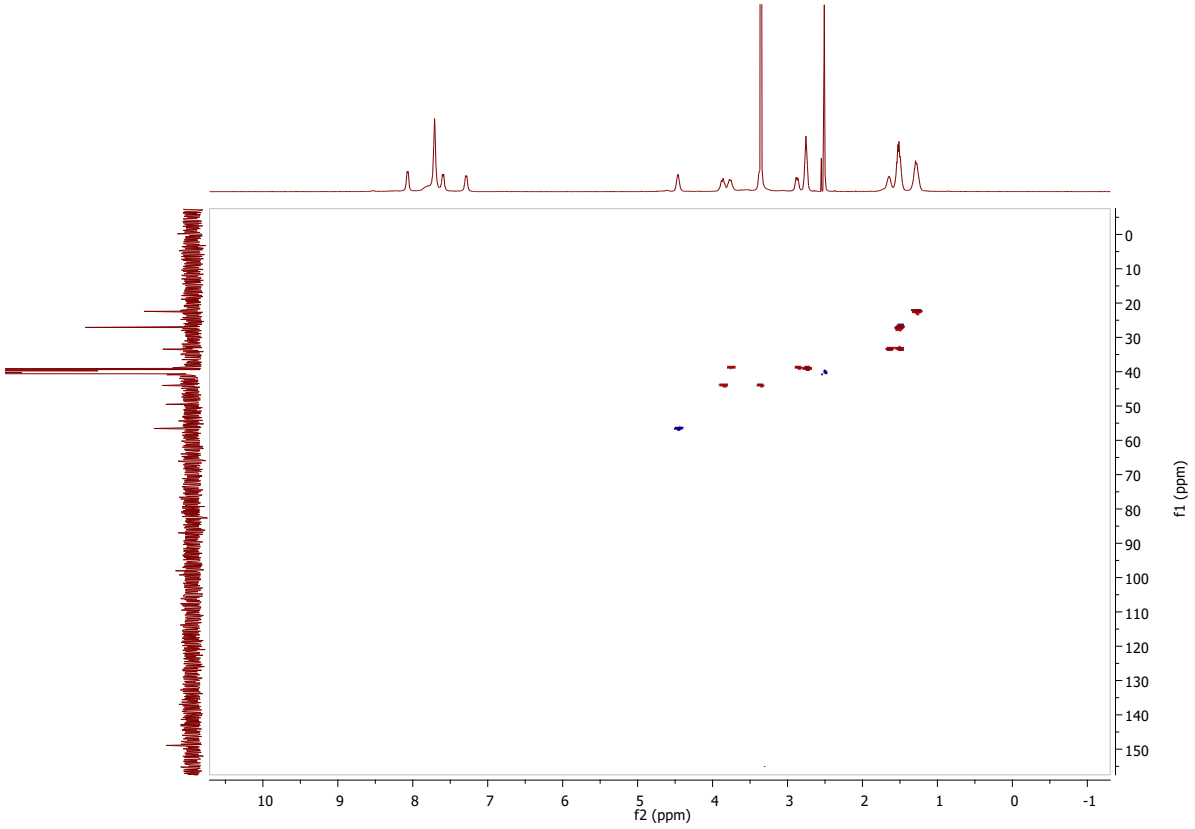


Figure S36: HSQC spectrum of **Sq-2-Lys** in DMSO-d_6.

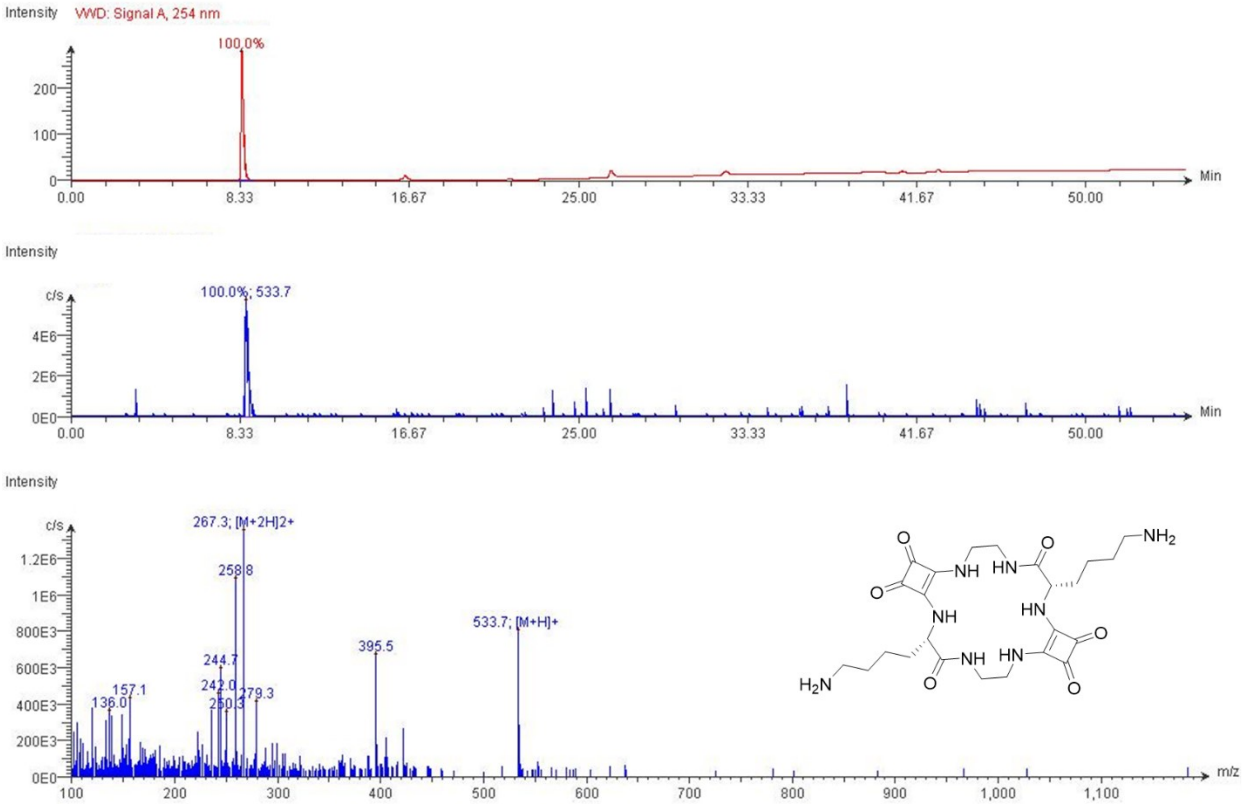


Figure S37: LCMS data of **Sq-2-Lys.**

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C24 H36 N8 O6	0.75	533.2828	532.2756	532.2758	-0.33	98.94

Mass errors of between -5.00 and 5.00 ppm with isotope match scores above 60% are considered confirmation of molecular formulae

Figure: Extracted ion chromatogram (EIC) of compound.

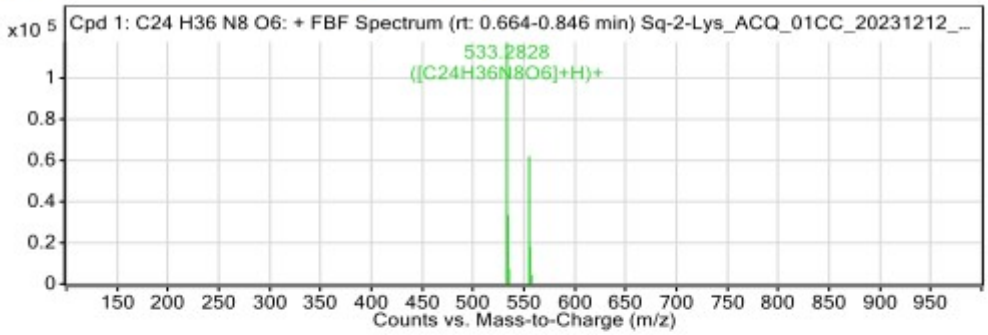


Figure: Full range view of Compound spectra and potential adducts.

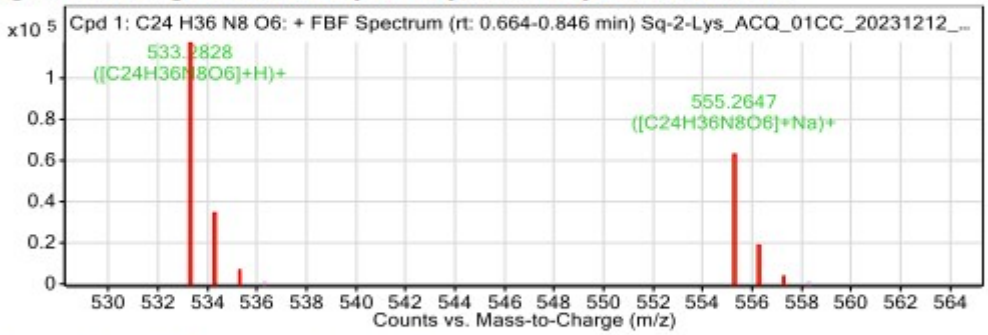


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
533.2828	1	117360.2	C24H36N8O6	(M+H)+
534.2856	1	33513.8	C24H36N8O6	(M+H)+
535.2884	1	7453.1	C24H36N8O6	(M+H)+
555.2647	1	61934.6	C24H36N8O6	(M+Na)+
556.2677	1	18164.1	C24H36N8O6	(M+Na)+
557.2744	1	4436.2	C24H36N8O6	(M+Na)+

Figure S38: HRMS data of **Sq-2-Lys.**

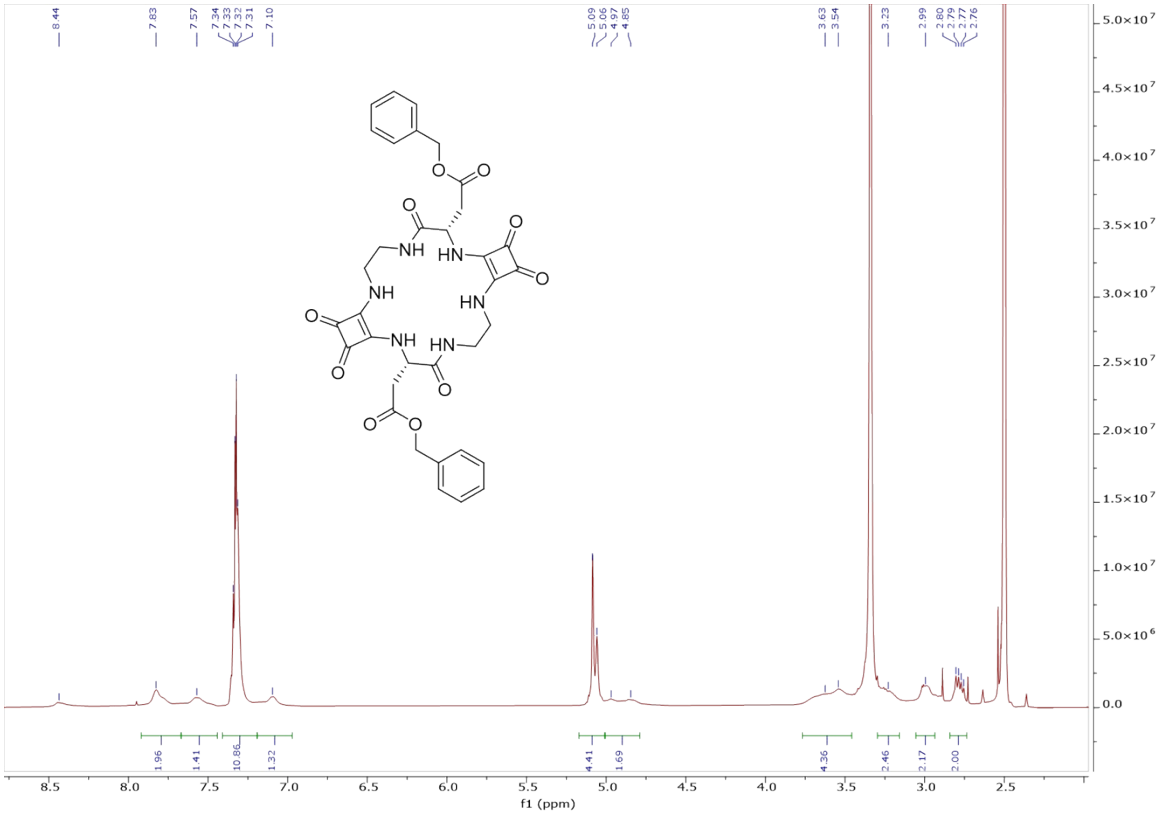


Figure S39: ^{1}H NMR spectrum of **Sq-2-Asp(Bzl)** in DMSO-d_6.

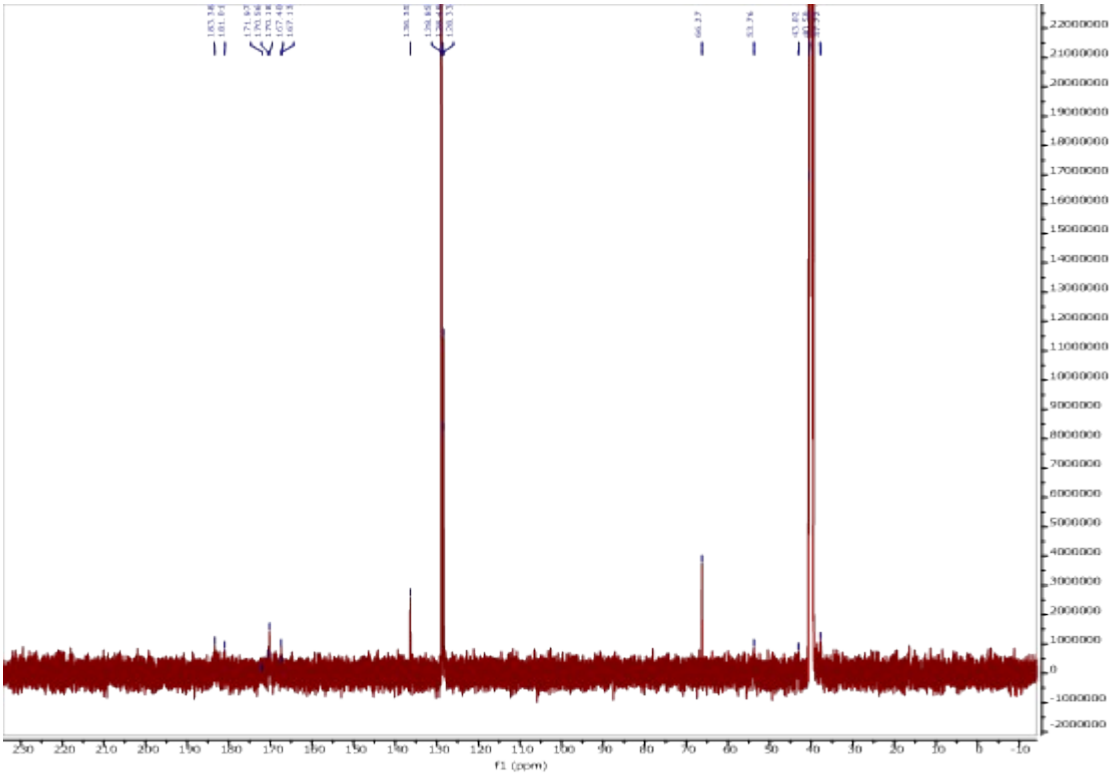


Figure S40: ^{13}C NMR spectrum of **Sq-2-Asp(Bzl)** in DMSO-d_6.

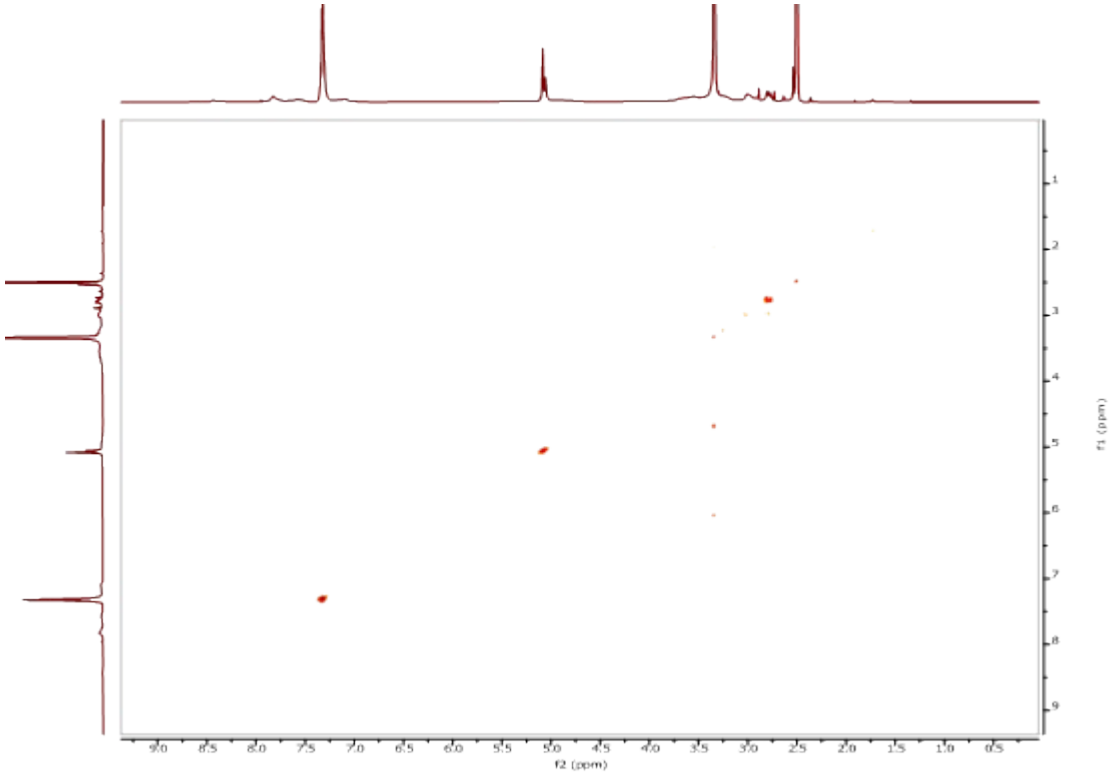


Figure S41: COSY spectrum of **Sq-2-Asp(Bzl)** in DMSO-d_6.

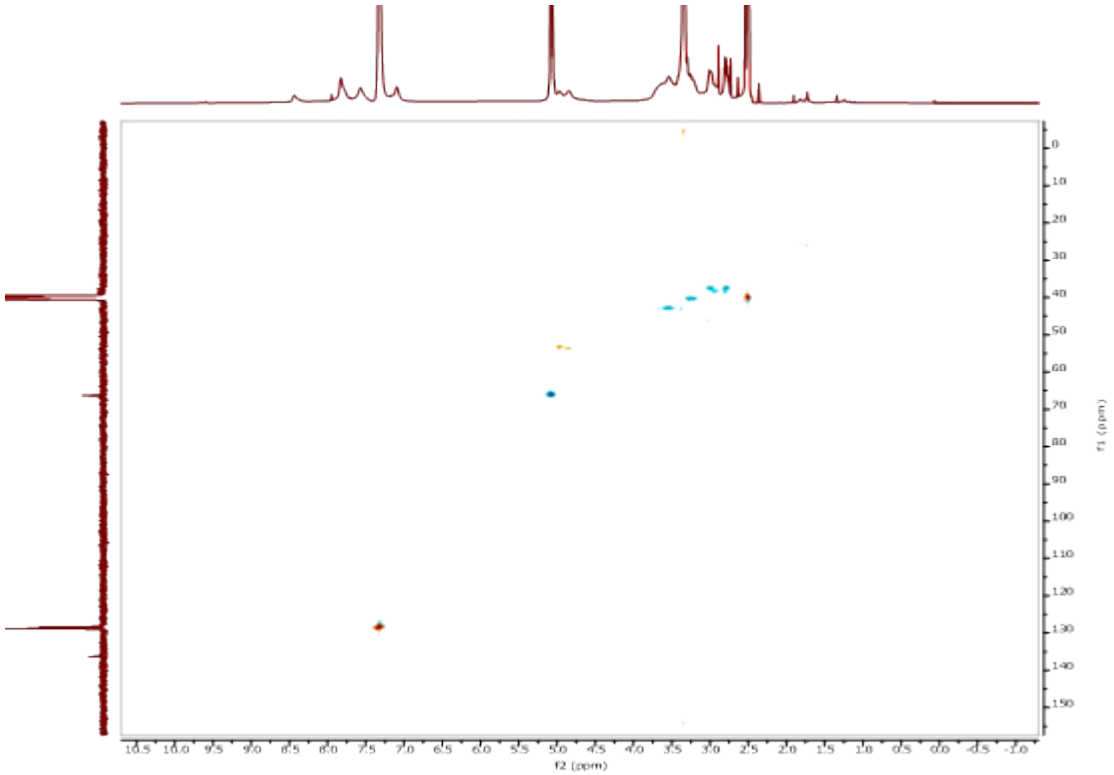


Figure S42: HSQC spectrum of **Sq-2-Asp(Bzl)** in DMSO-d_6.

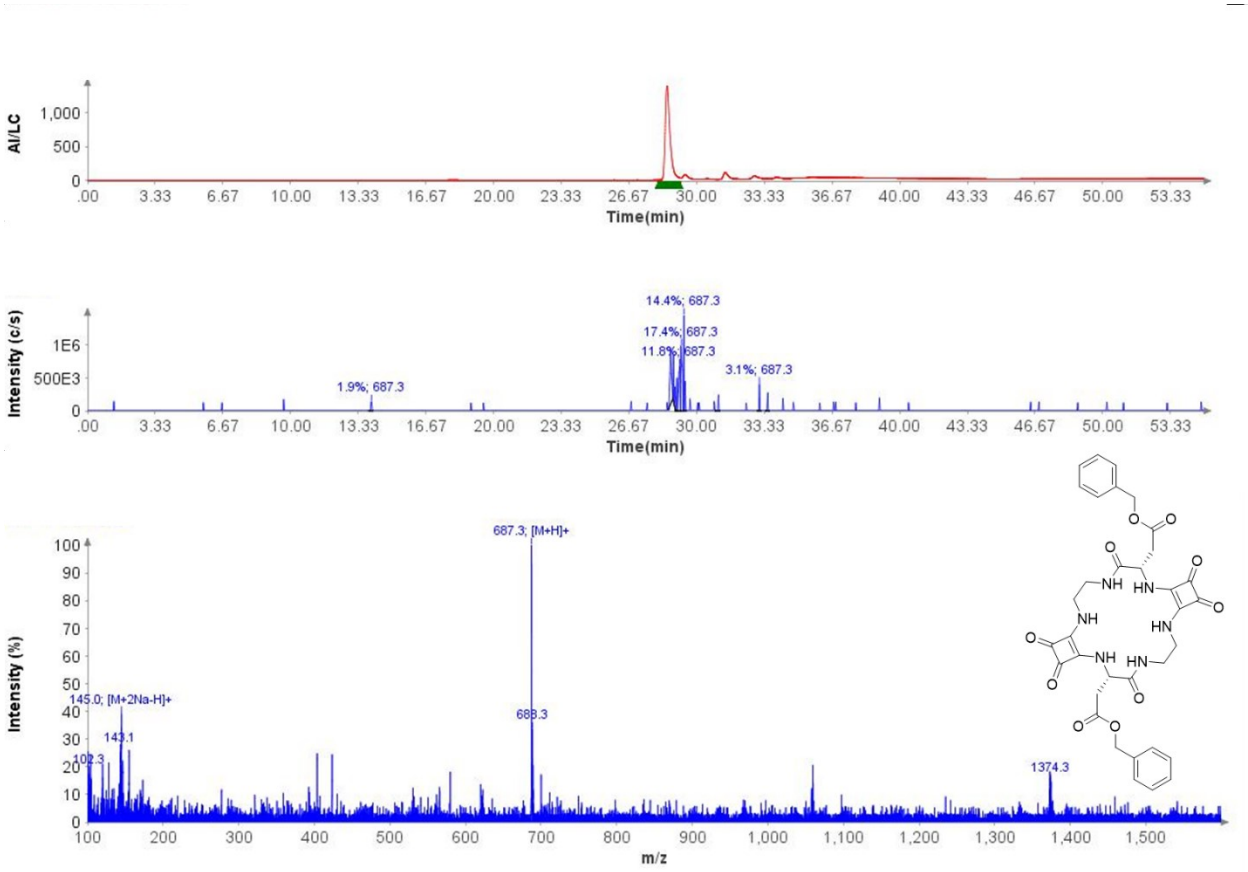


Figure S43: LCMS data of **Sq-2-Asp(Bzl).**

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C34 H34 N6 O10	0.73	709.2229	686.2338	686.2336	0.22	99.81

Figure: Extracted ion chromatogram (EIC) of compound.

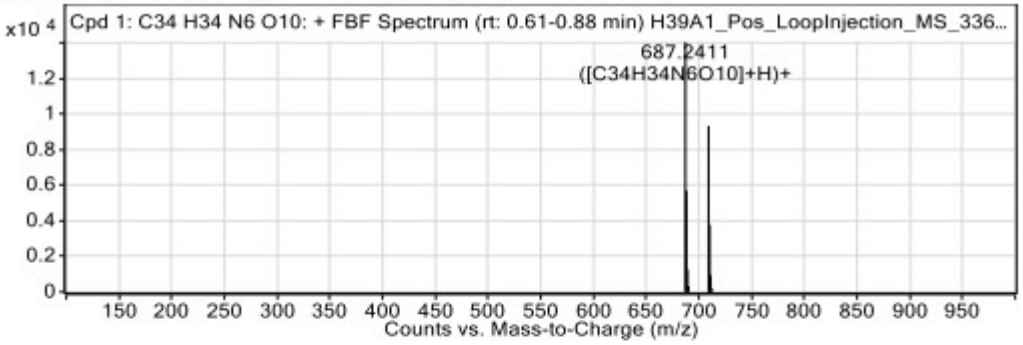


Figure: Full range view of Compound spectra and potential adducts.

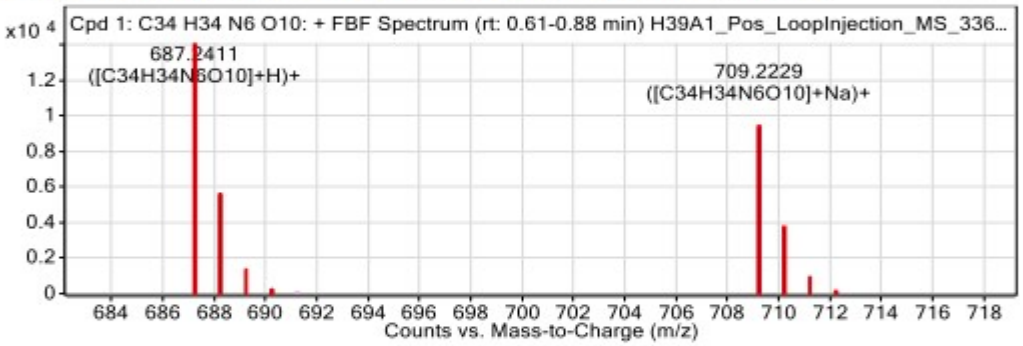


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
687.2411	1	14083.9	C34H34N6O10	(M+H)+
688.2443	1	5686.3	C34H34N6O10	(M+H)+
689.2466	1	1272.8	C34H34N6O10	(M+H)+
690.2499	1	317.7	C34H34N6O10	(M+H)+
709.2229	1	9337.3	C34H34N6O10	(M+Na)+
710.2259	1	3771.1	C34H34N6O10	(M+Na)+
711.2284	1	937.2	C34H34N6O10	(M+Na)+
712.2314	1	198.5	C34H34N6O10	(M+Na)+

Figure S44: HRMS data of **Sq-2-Asp(Bzl).**

Characterisation and Spectroscopic data of Acyclic Squaratide (ASq-2-Phe)

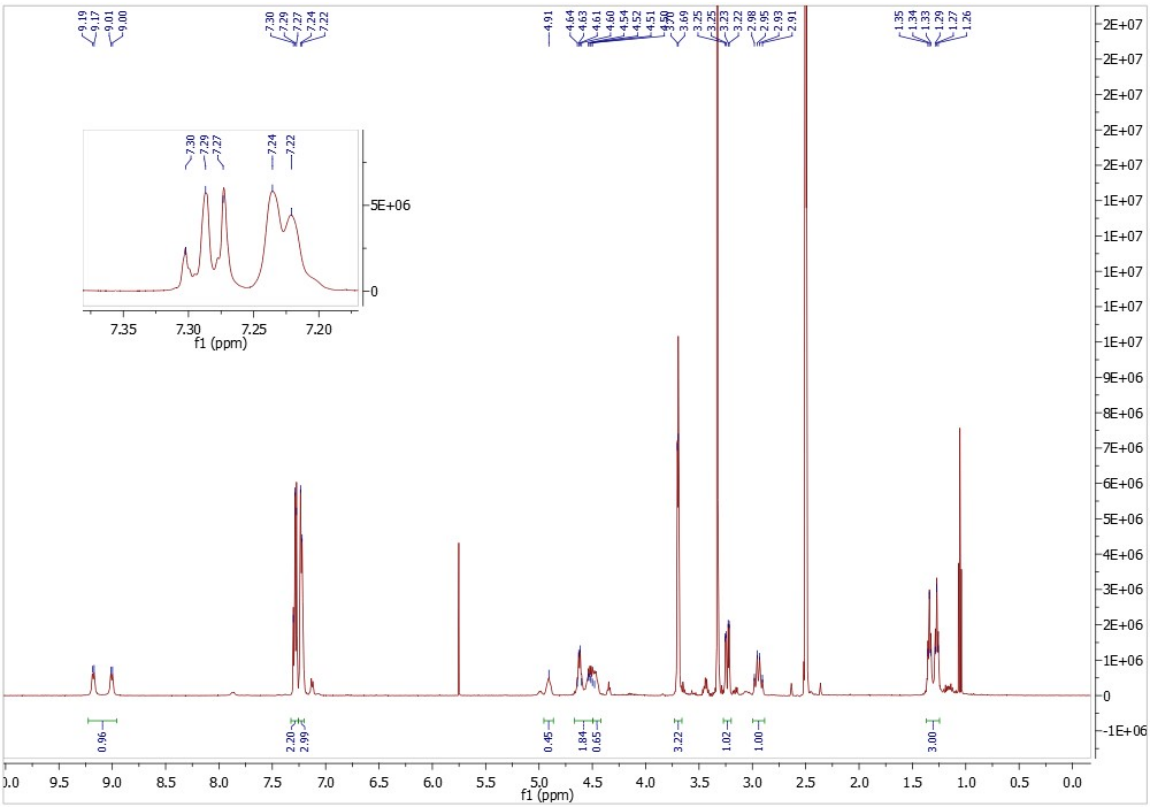


Figure S45: ^{1}H NMR spectrum of **1** in DMSO-d_6.

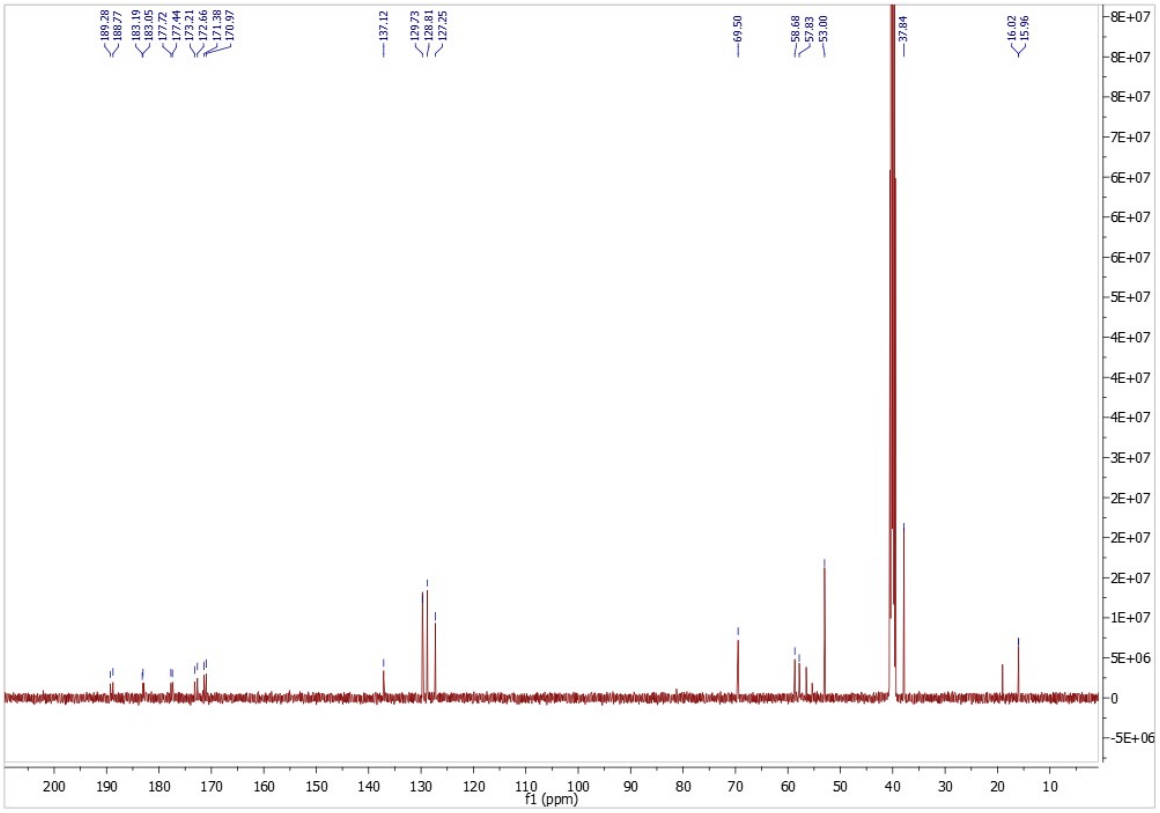


Figure S46: ^{13}C NMR spectrum of **1** in DMSO-d_6.

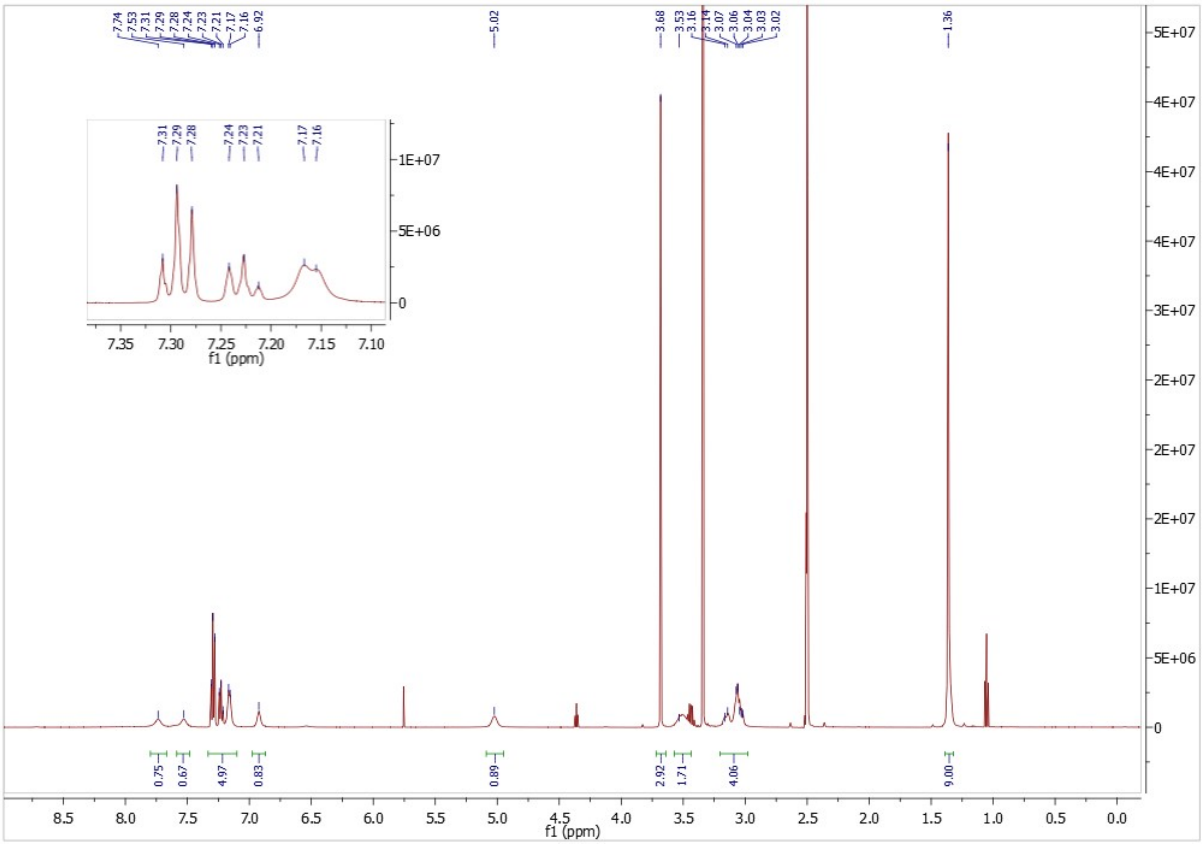


Figure S47: ^{1}H NMR spectrum of **2** in DMSO-d_6.

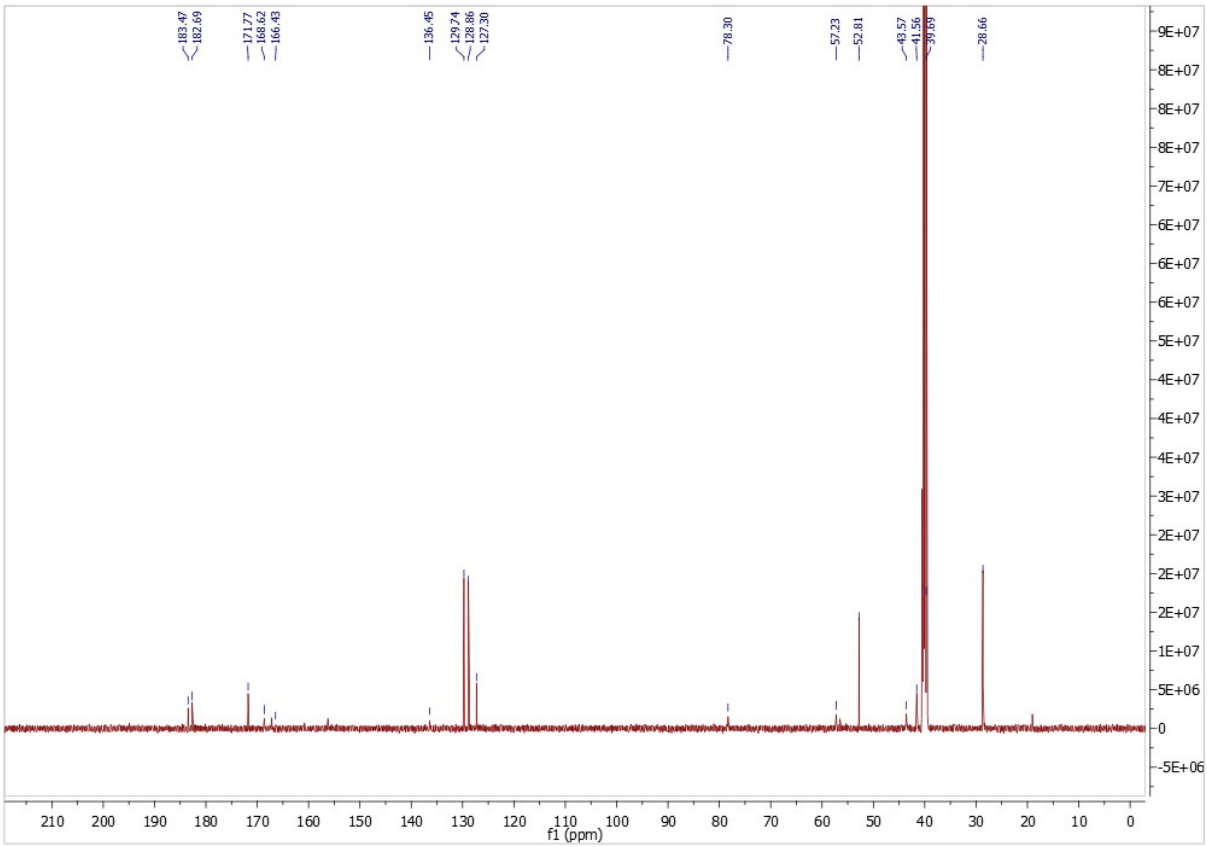


Figure S48: ^{13}C NMR spectrum of **2** in DMSO-d_6.

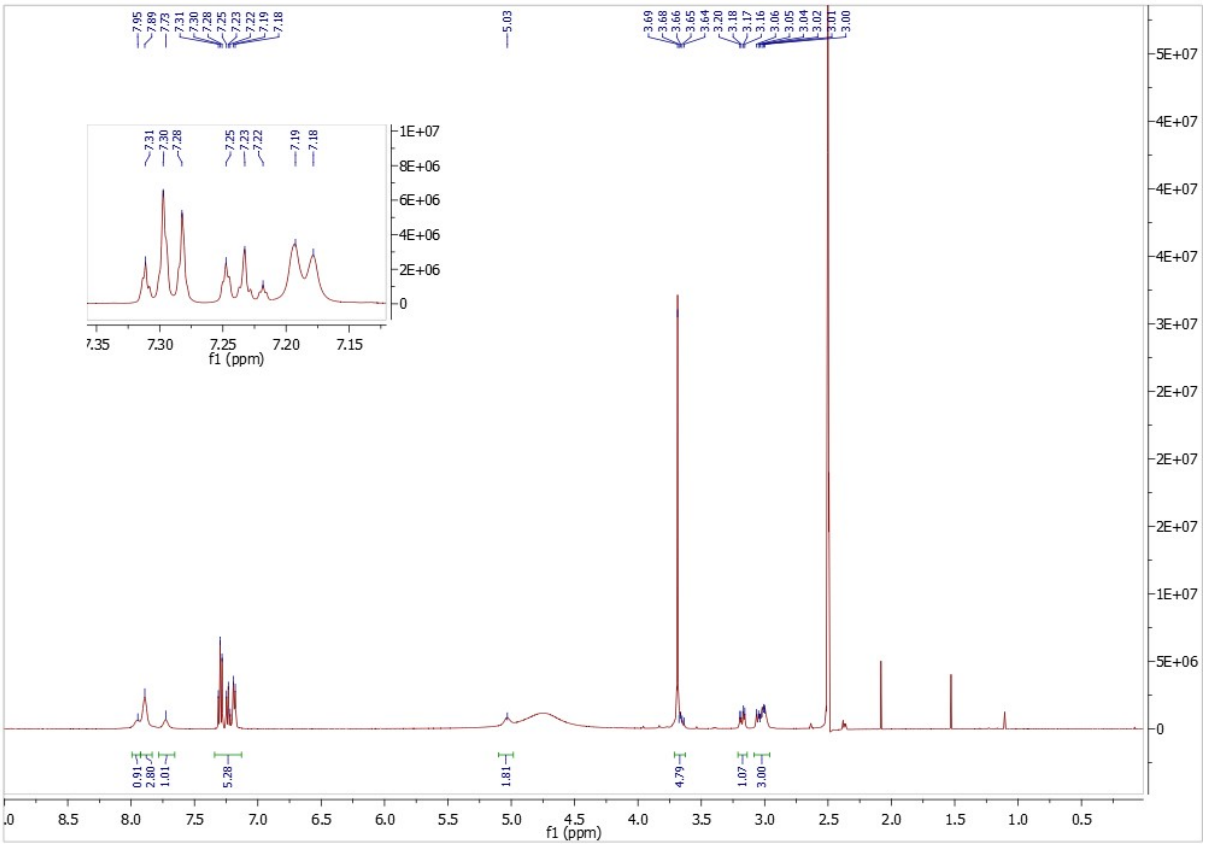


Figure S49: ^{1}H NMR spectrum of **3** in DMSO-d_6.

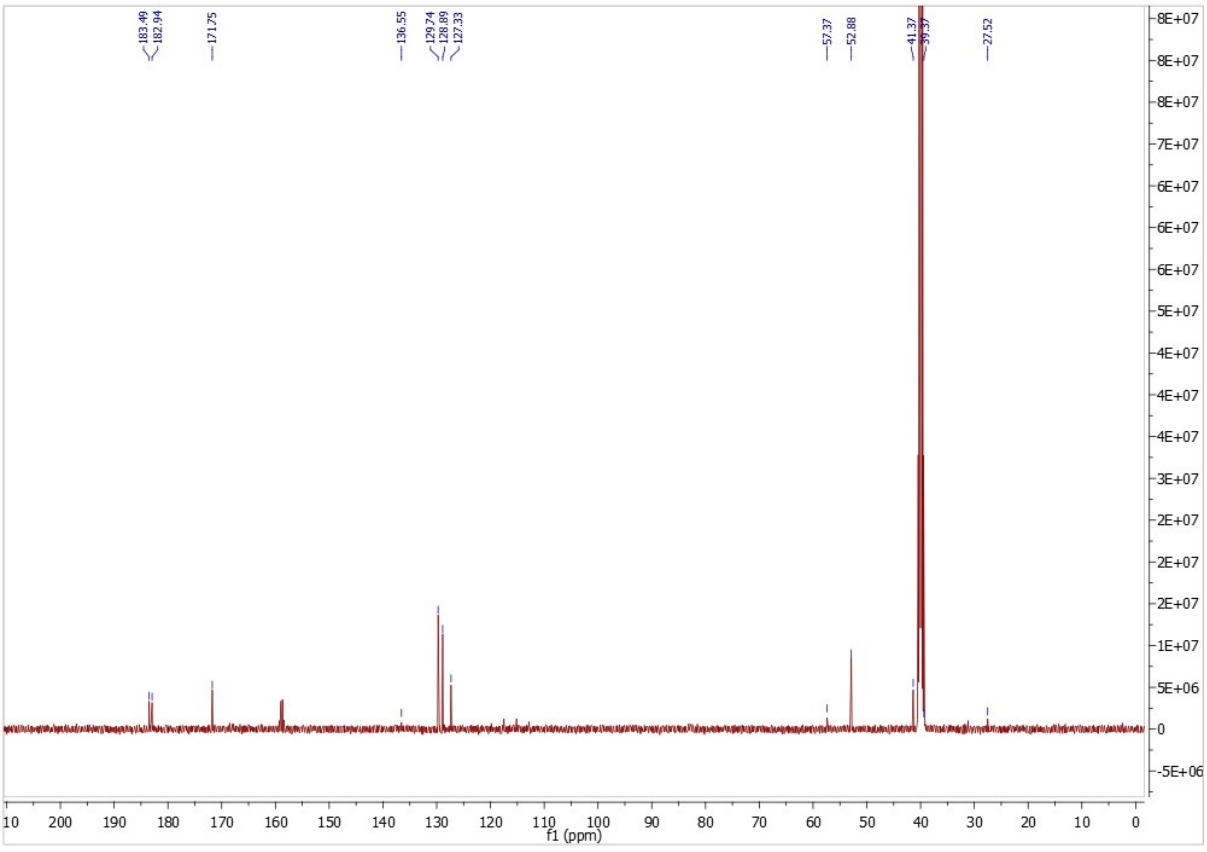


Figure S50: ^{13}C NMR spectrum of **3** in DMSO-d_6.

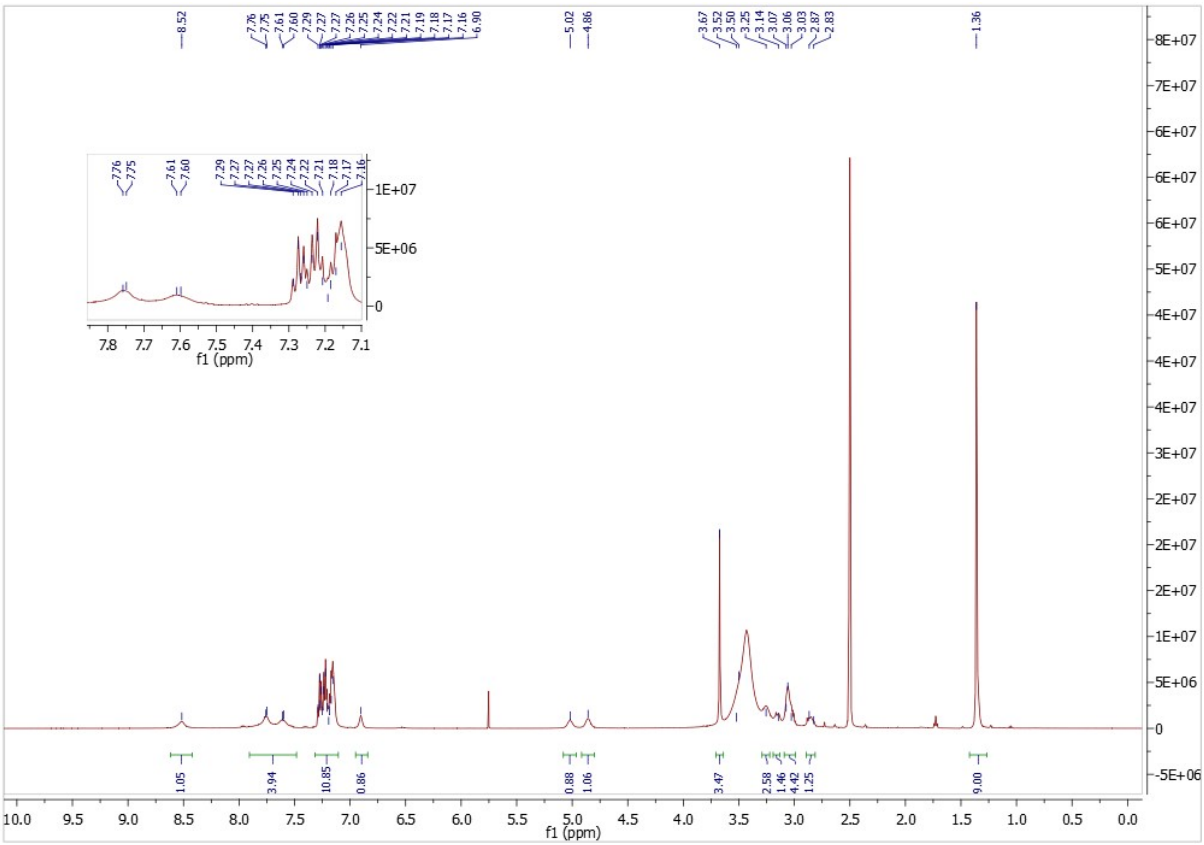


Figure S51: ^{1}H NMR spectrum of **ASq-2-Phe** in DMSO-d_6.

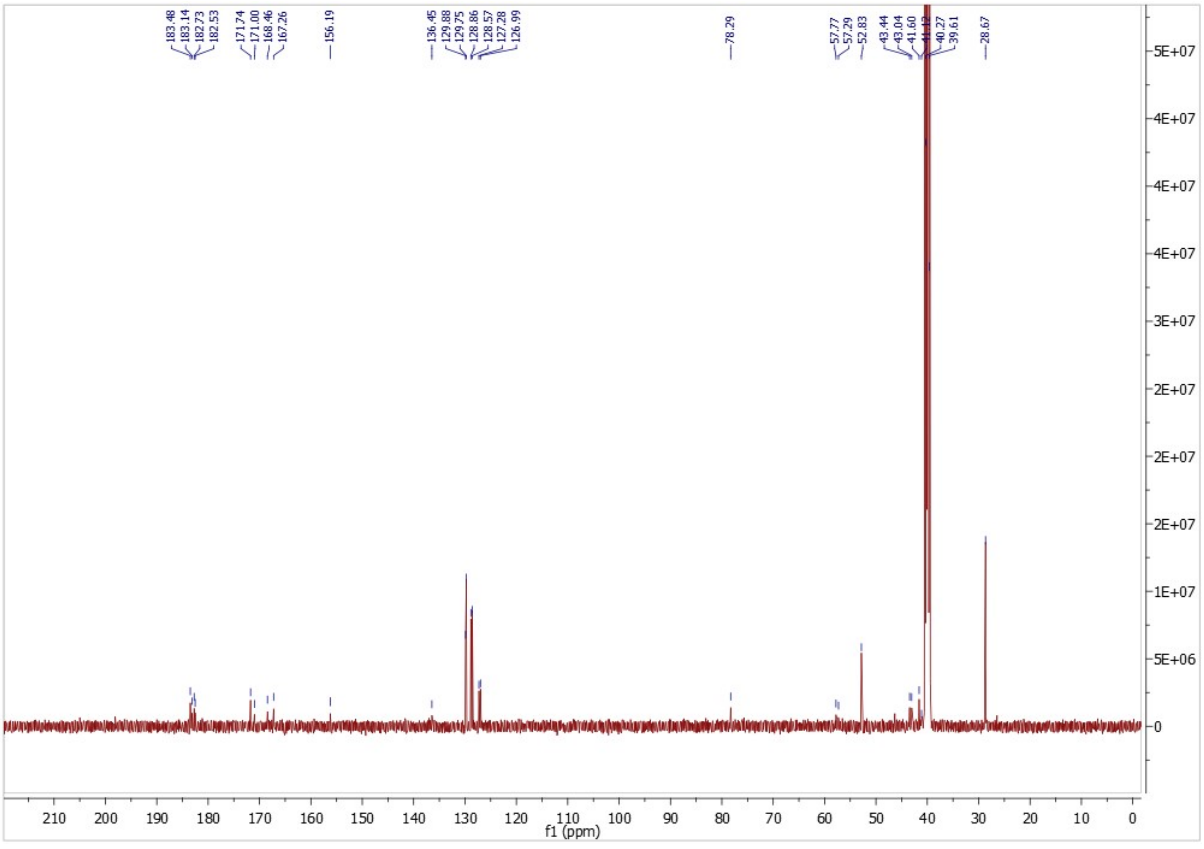


Figure S52: ^{13}C NMR spectrum of **ASq-2-Phe** in DMSO-d_6.

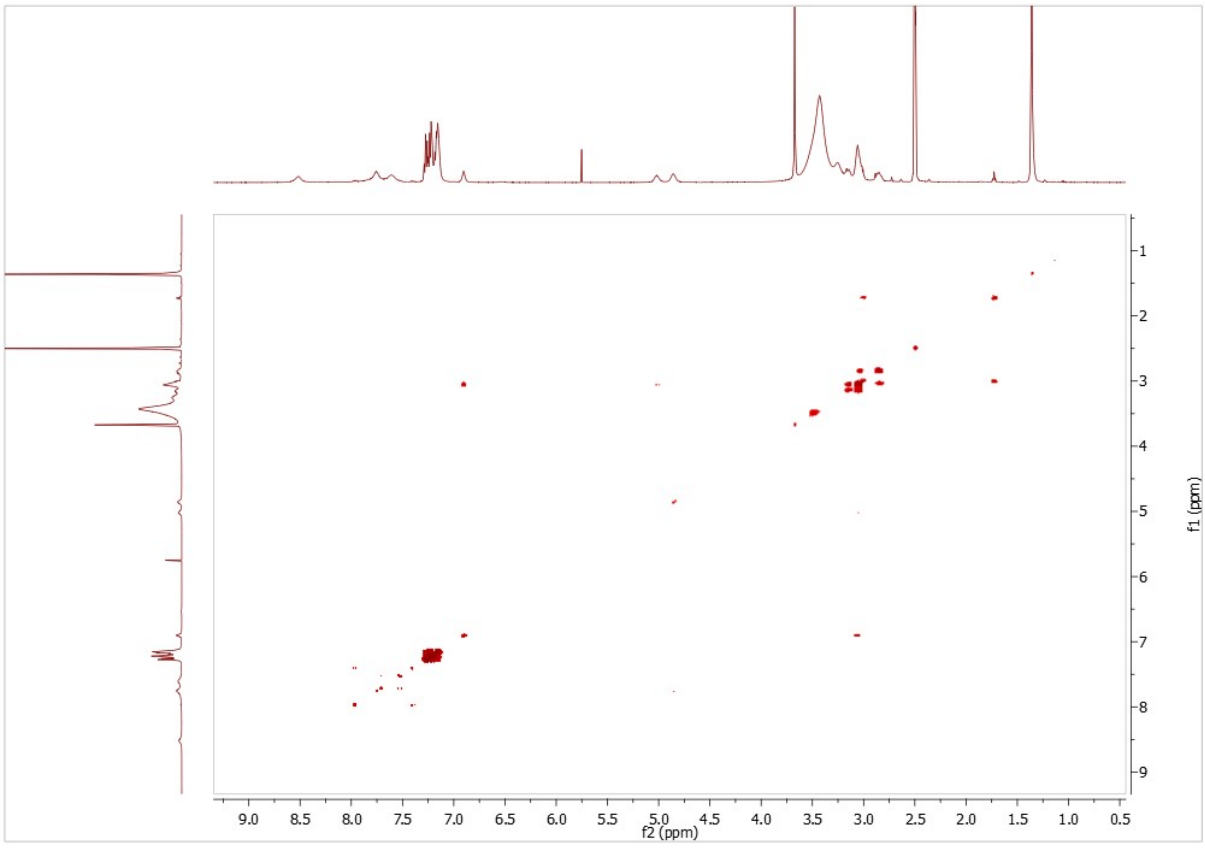


Figure S53: COSY spectrum of **ASq-2-Phe** in DMSO-d_6.

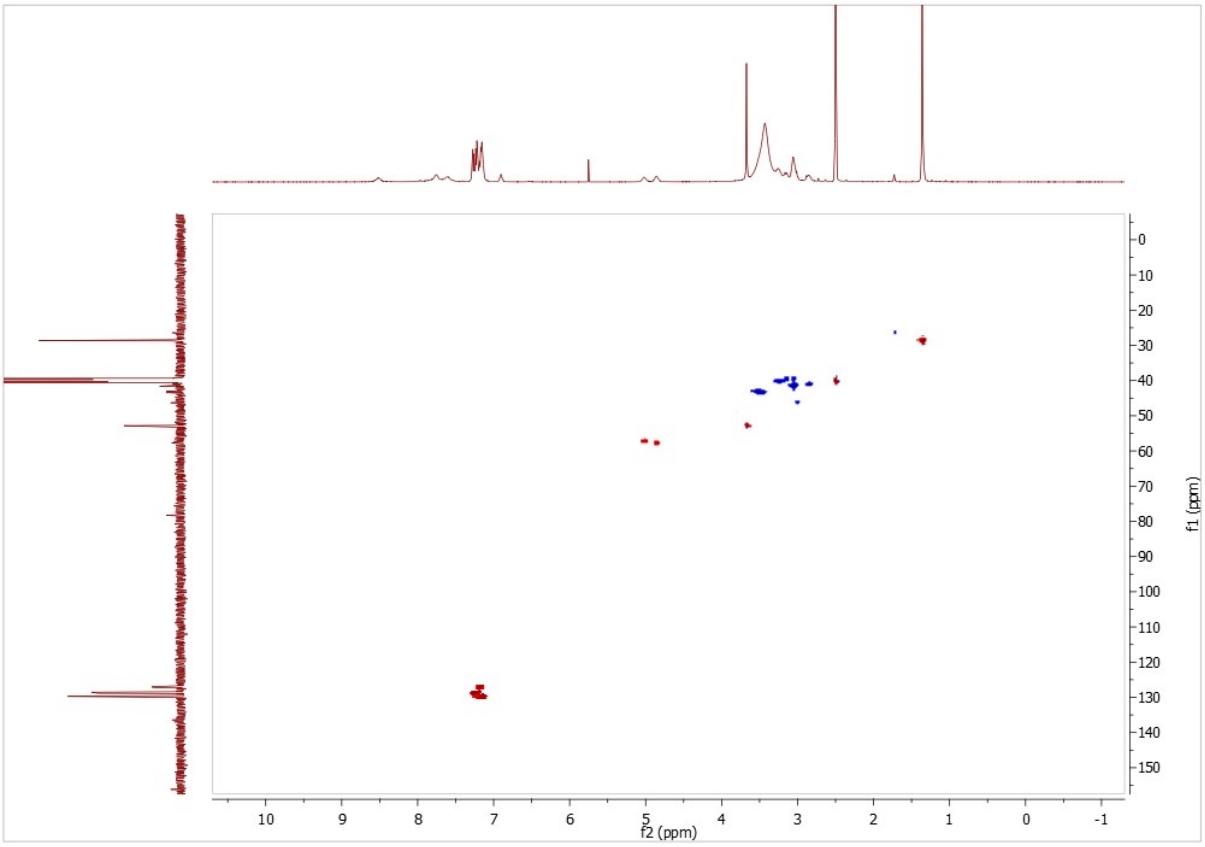


Figure S54: HSQC spectrum of **ASq-2-Phe** in DMSO-d_6.

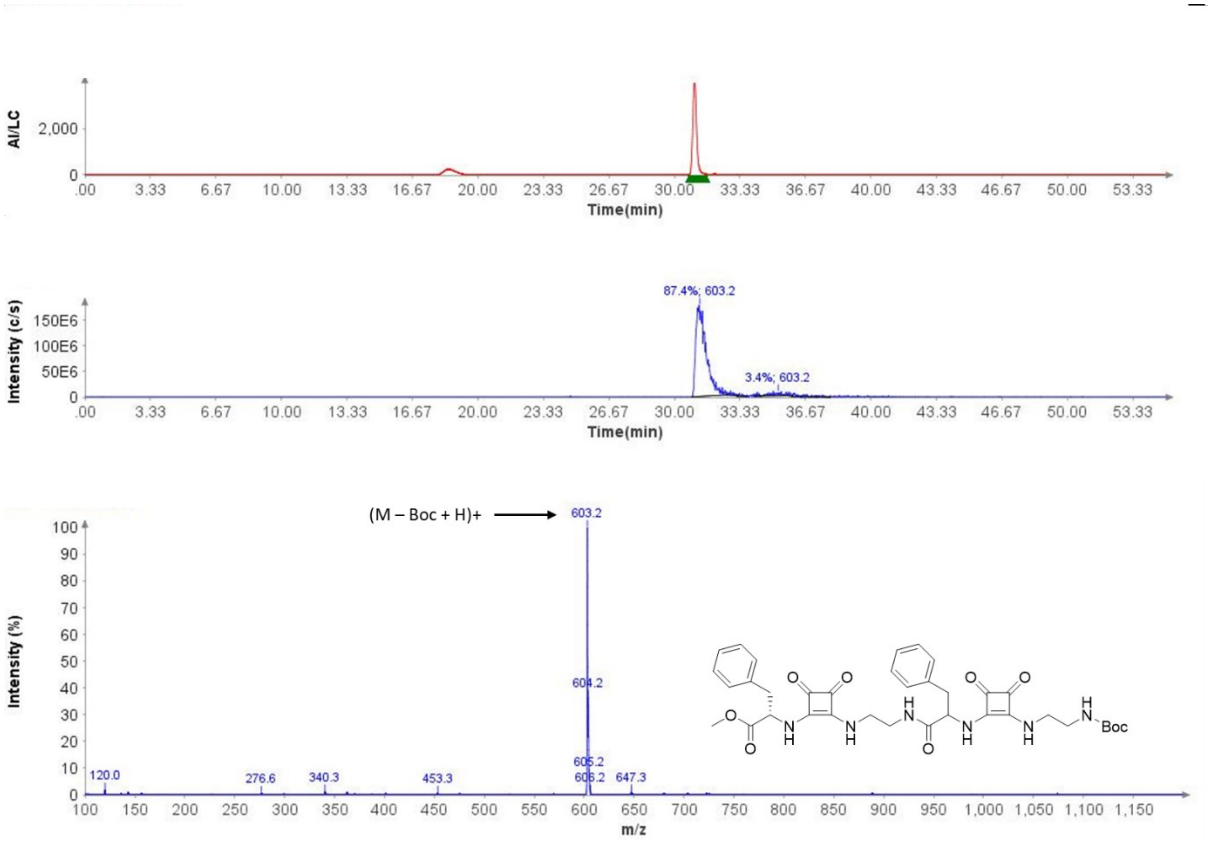


Figure S55: LCMS data of **ASq-2-Phe**.

Compound Table

Compound Label	RT (min)	Observed mass (m/z)	Neutral observed mass (Da)	Theoretical mass (Da)	Mass error (ppm)	Isotope match score (%)
Cpd 1: C36 H42 N6 O9	0.71	725.2922	702.3027	702.3013	1.94	98.05

Figure: Extracted ion chromatogram (EIC) of compound.

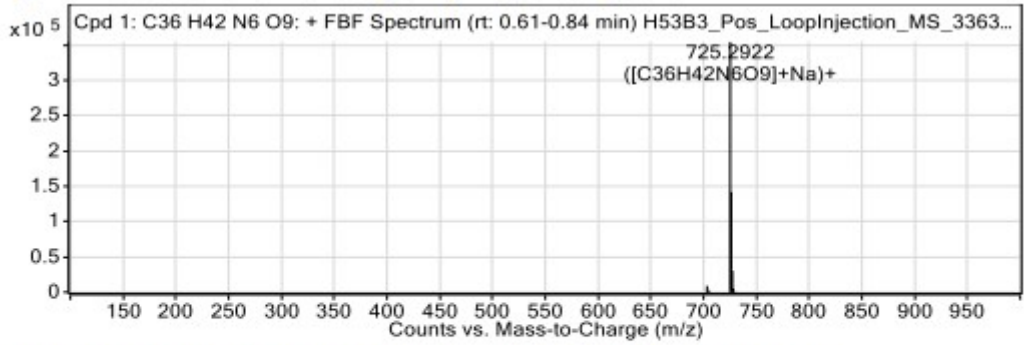


Figure: Full range view of Compound spectra and potential adducts.

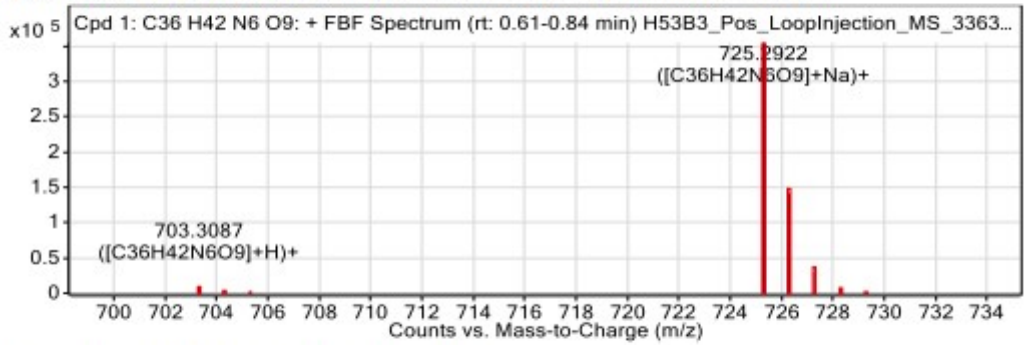


Figure: Zoomed Compound spectra view
(red boxes indicating expected theoretical isotope spacing and abundance)

Compound isotope peak List

m/z	z	Abund	Formula	Ion
703.3087	1	8984.4	C36H42N6O9	(M+H)+
704.3123	1	3564.0	C36H42N6O9	(M+H)+
705.3158	1	1012.7	C36H42N6O9	(M+H)+
706.3205	1	220.1	C36H42N6O9	(M+H)+
725.2922	1	355158.7	C36H42N6O9	(M+Na)+
726.2948	1	141965.0	C36H42N6O9	(M+Na)+
727.2968	1	30350.3	C36H42N6O9	(M+Na)+
728.2992	1	5129.1	C36H42N6O9	(M+Na)+
729.2873	1	911.7	C36H42N6O9	(M+Na)+

Figure S56: HRMS data of **ASq-2-Phe**

3. Temperature Dependent NMR

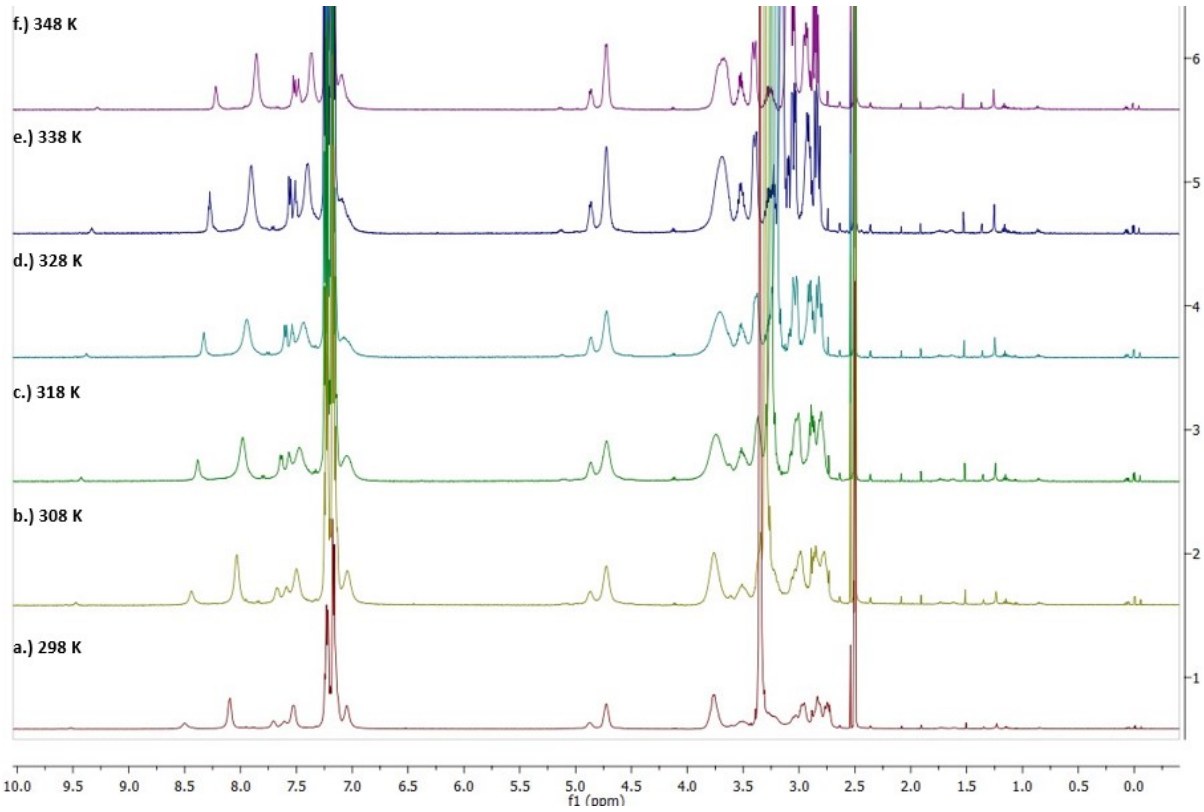


Figure S57: The VT – NMR of **Sq-2-Phe** from 298 K (a.) to 348 K (f.).

4. ^{1}H NMR Anion Binding Screen

Cyclic dimers

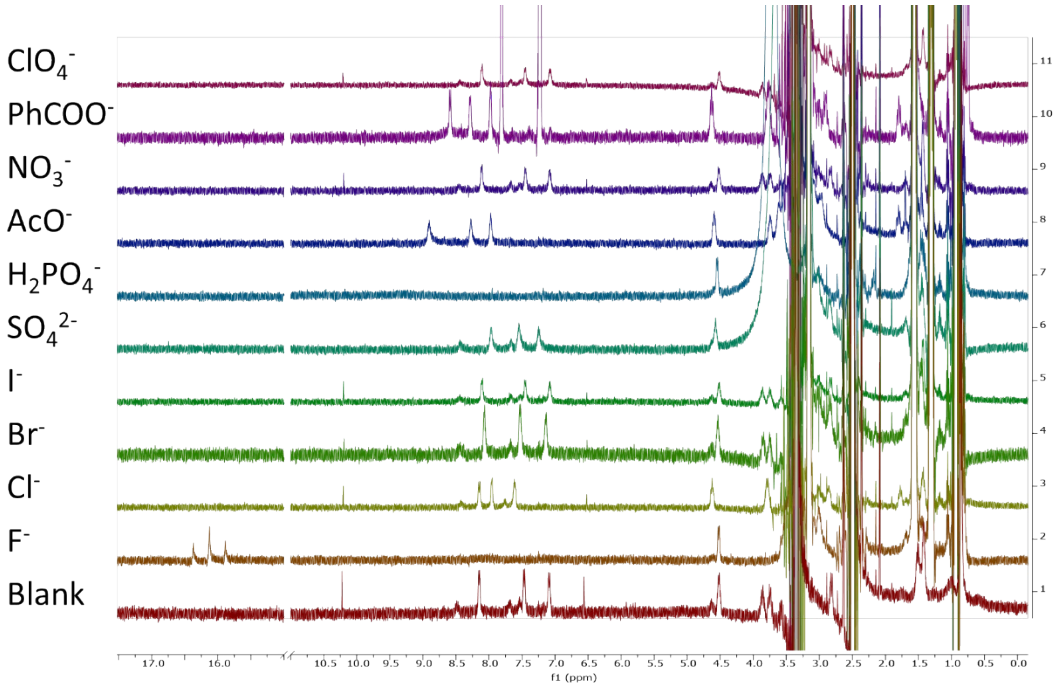


Figure S58: ^{1}H NMR stack plot of **Sq-2-Leu** (2.5 x 10^{-3} M) titrated with different TBA salts of anions (10 eq.).

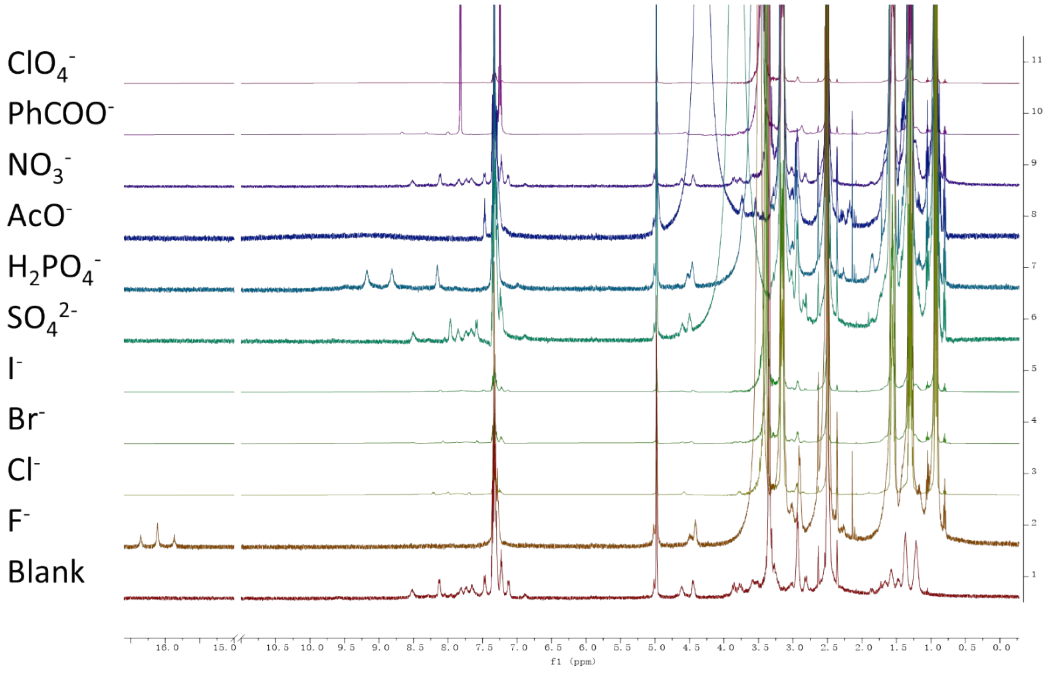


Figure S59: ^{1}H NMR stack plot of **Sq-2-Lys(Cbz)** (2.5 x 10^{-3} M) titrated with different TBA salts of anions (10 eq.).

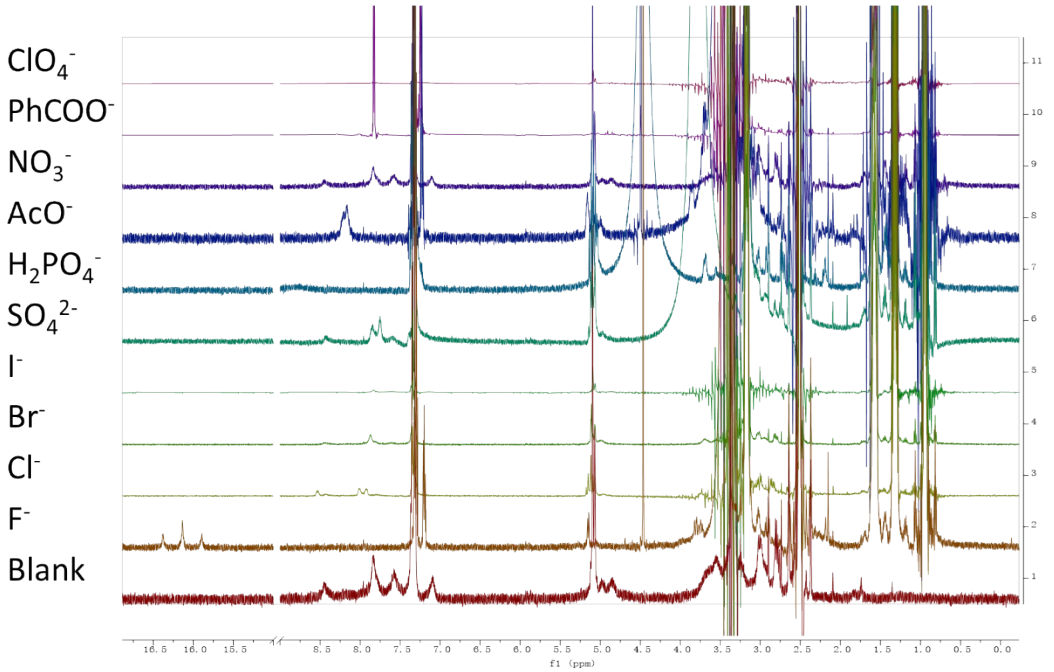


Figure S60: ^{1}H NMR stack plot of **Sq-2-Asp(Bzl)** (2.5 x 10^{-3} M) titrated with different TBA salts of anions (10 eq.).

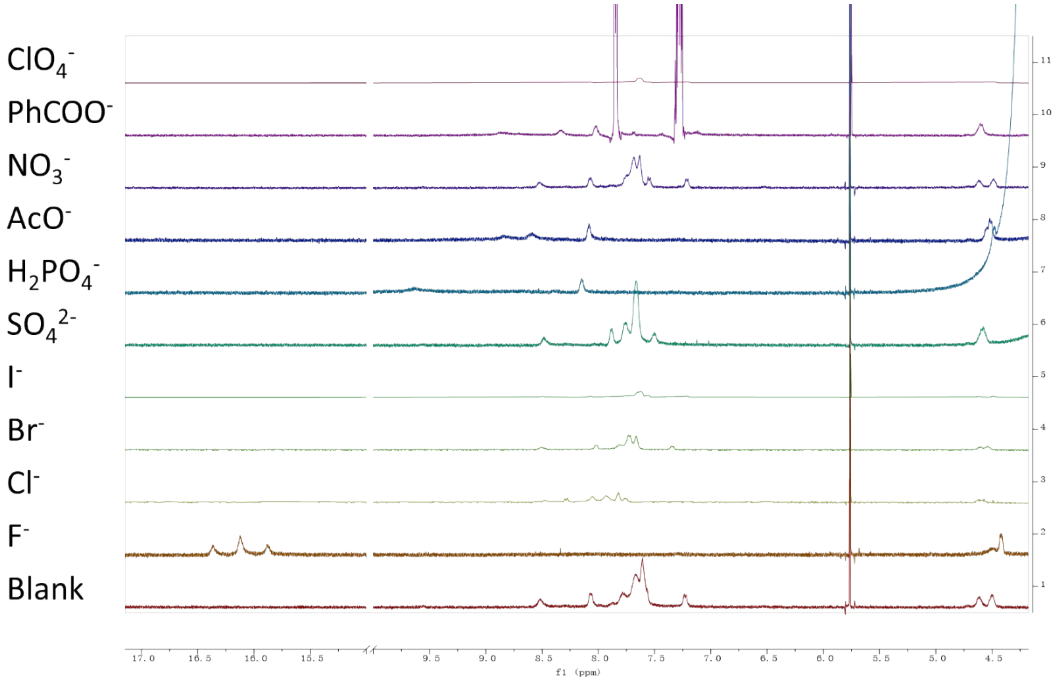


Figure S61: ^{1}H NMR stack plot of **Sq-2-Lys** (2.5 x 10^{-3} M) titrated with different TBA salts of anions (10 eq.).

Acyclic Dimers

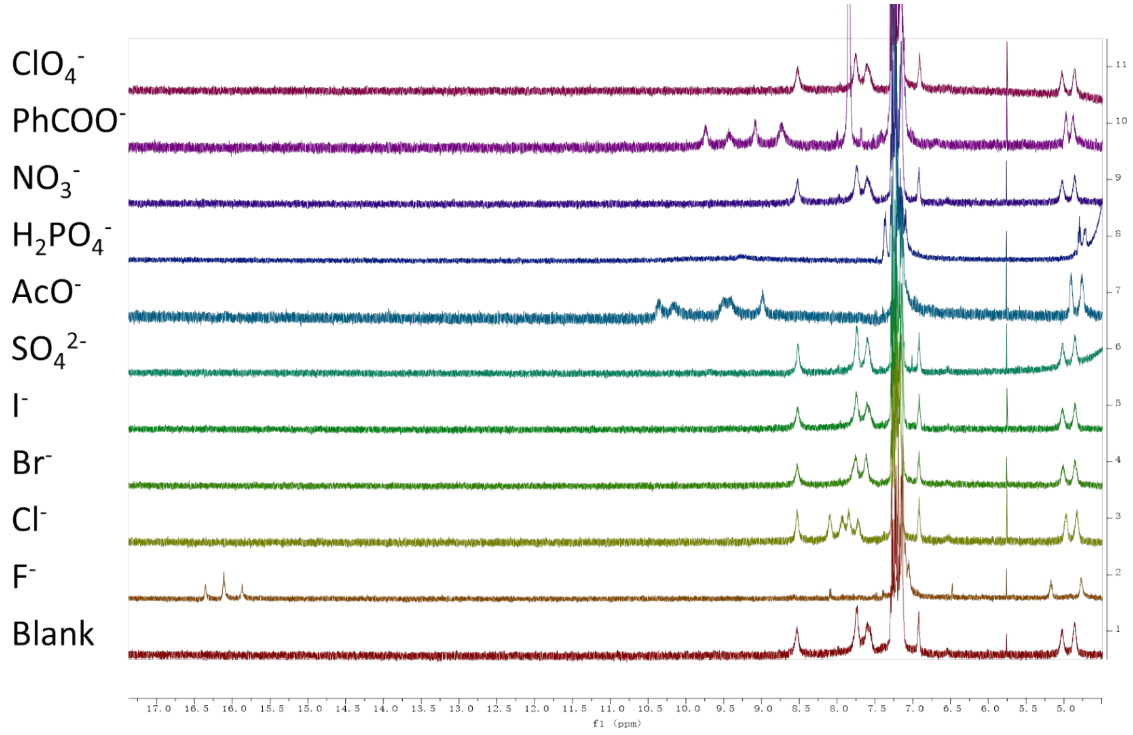


Figure S62: ^{1}H NMR stack plot of **ASq-2-Phe** (2.5 x 10^{-3} M) titrated with different TBA salts of anions (10 eq.).

5. ^{1}H NMR Anion Binding and Fitting Data

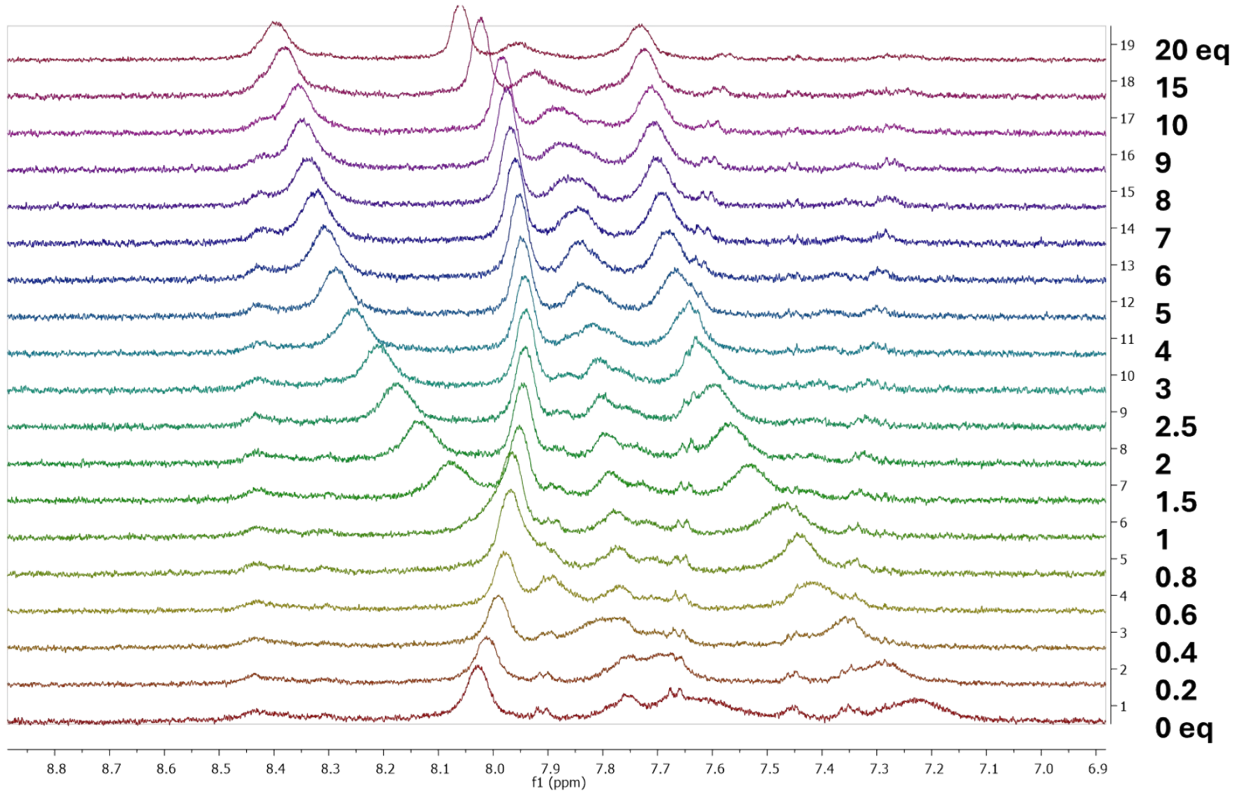


Figure S63: ^{1}H NMR titration of compound **Sq-2-Ala** with TBACl in DMSO-d_6

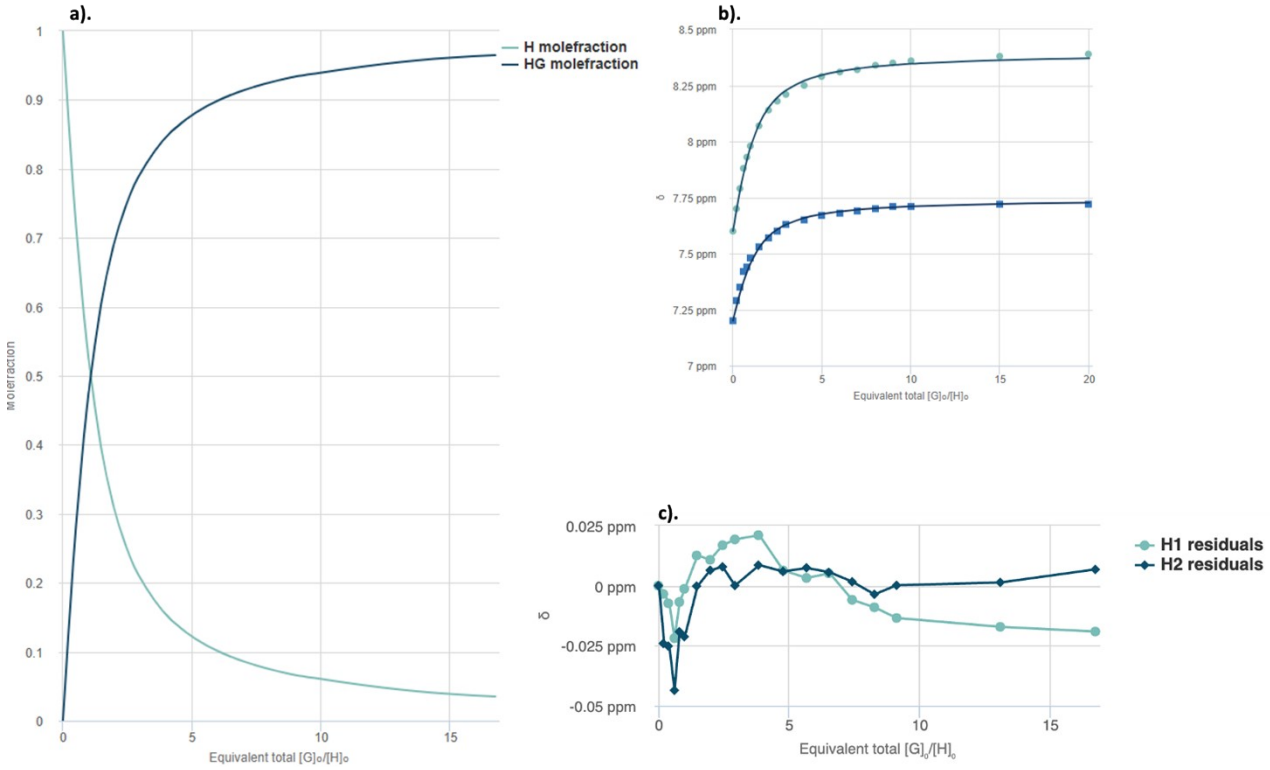


Figure S64: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Ala. b).** Fitting binding isotherms of **Sq-2-Ala** with TBACl in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 1:1 binding model (K_a = 692.55 M^{-1}). **c).** Residuals plot of **Sq-2-Ala**.

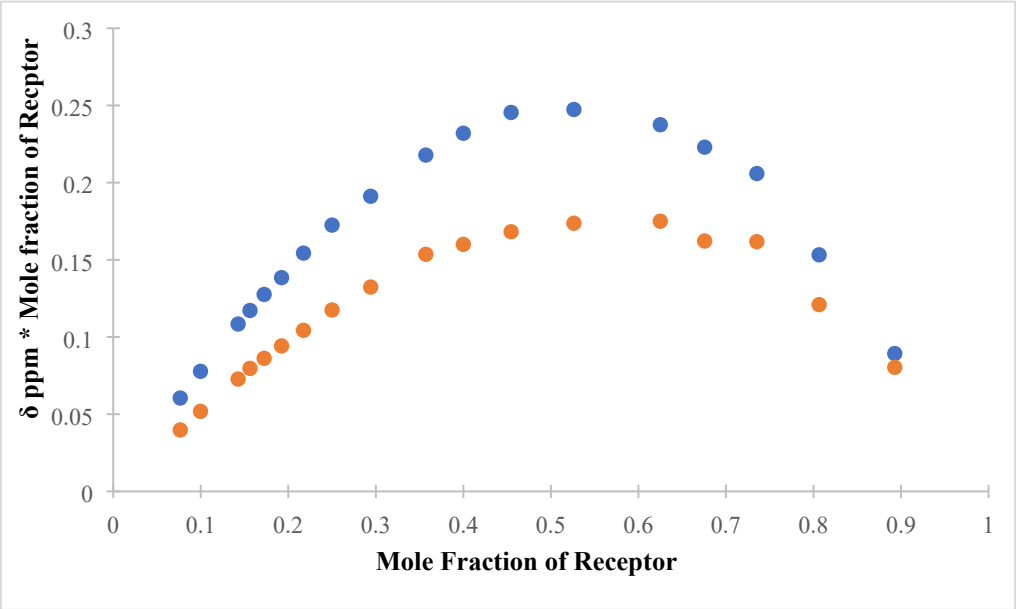


Figure S65: Jobs plot of **Sq-2-Ala** with TBACl in DMSO-d_6

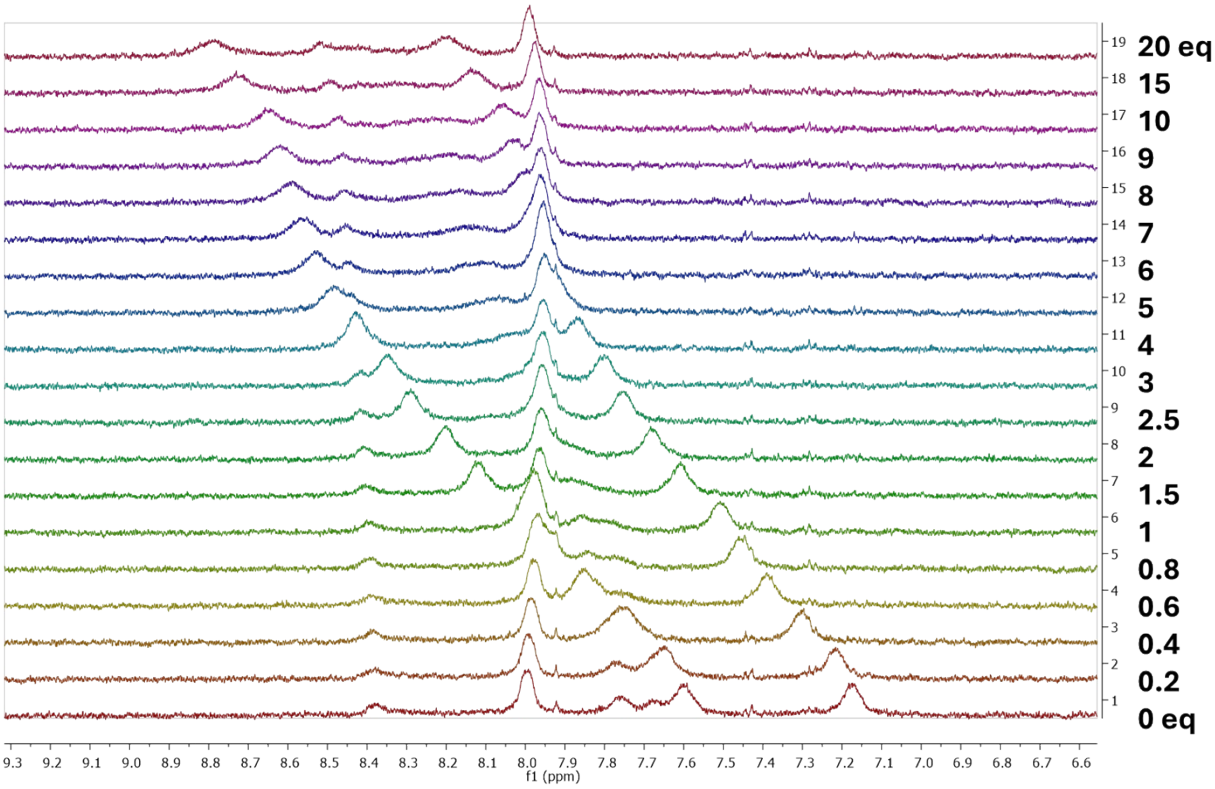


Figure S66: ^{1}H NMR titration of compound **Sq-2-Ala** with TBAAcO in DMSO-d_6

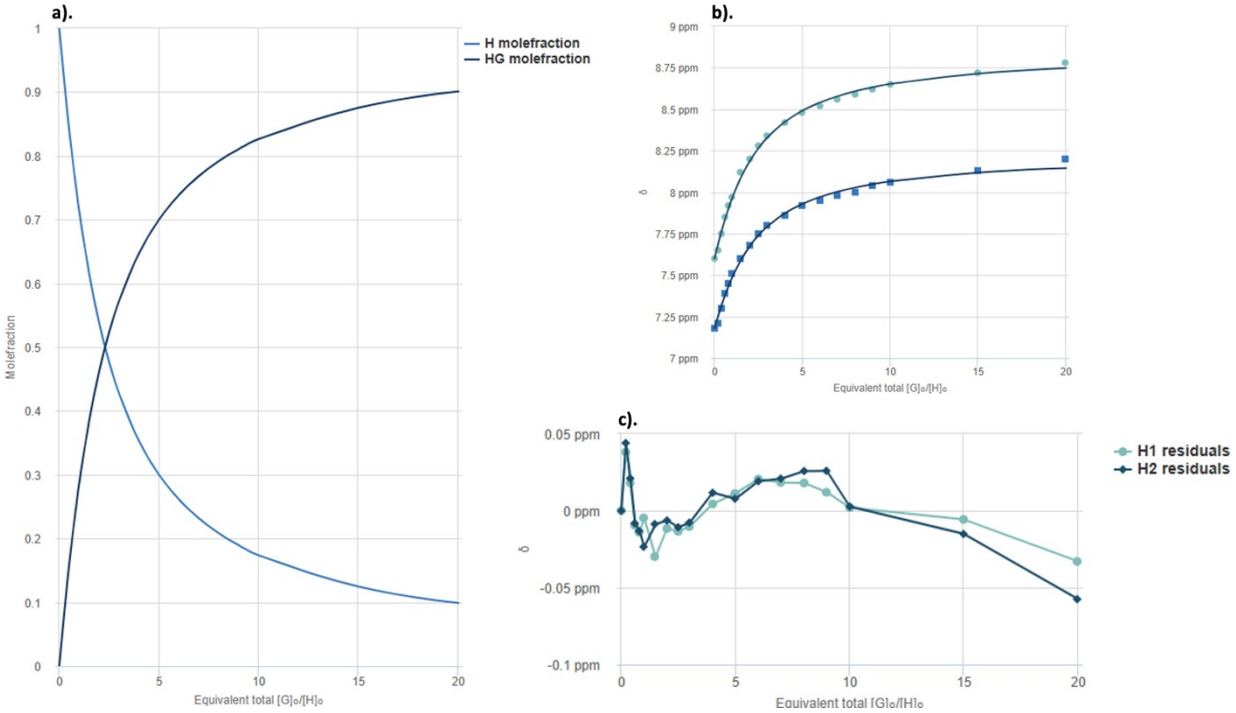


Figure S67: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Ala. b).** Fitting binding isotherms of **Sq-2-Ala** with TBAAcO in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 1:1 binding model (K_a = 227.24 M^{-1}). **c).** Residuals plot of **Sq-2-Ala**.

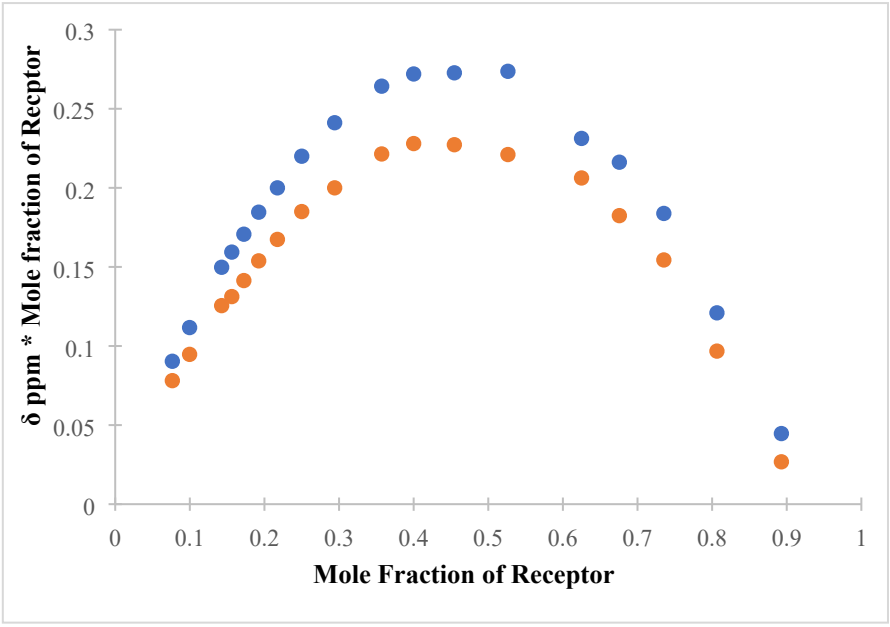


Figure S68: Jobs plot of **Sq-2-Ala)** with TBAAcO in DMSO-d_6

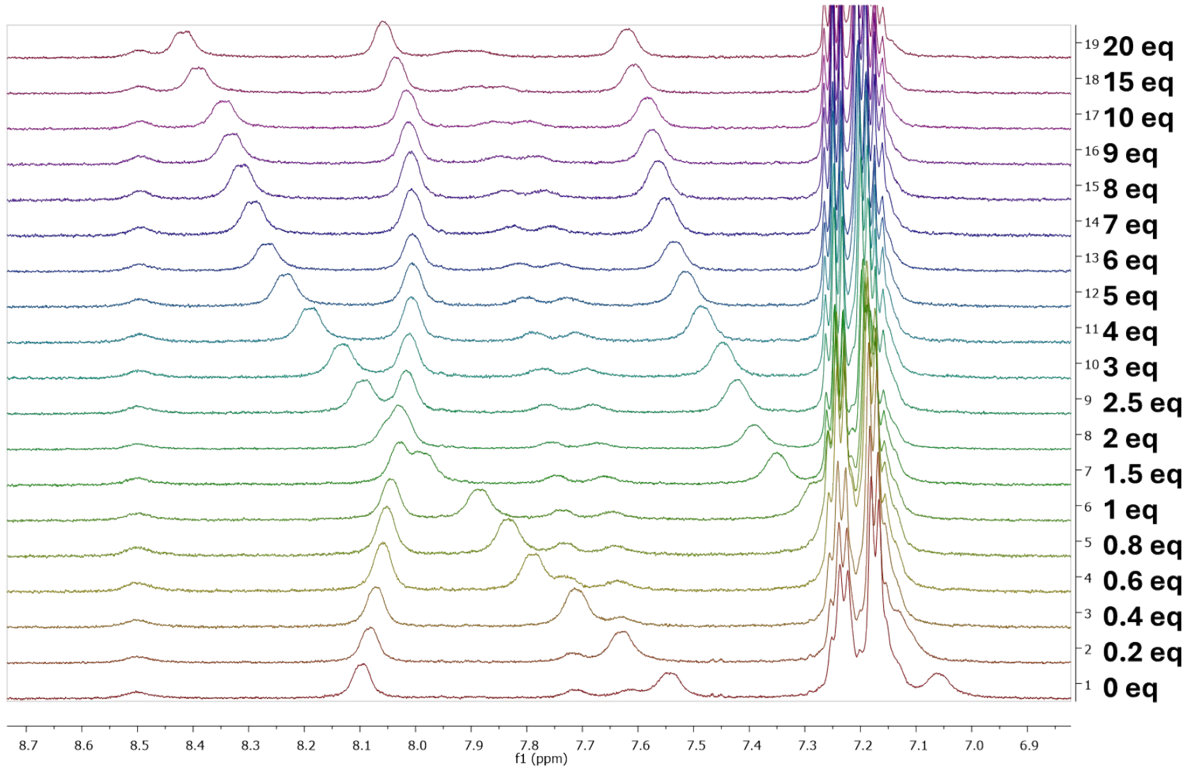


Figure S69: ^{1}H NMR titration of compound **Sq-2-Phe** with TBACl in DMSO-d_6

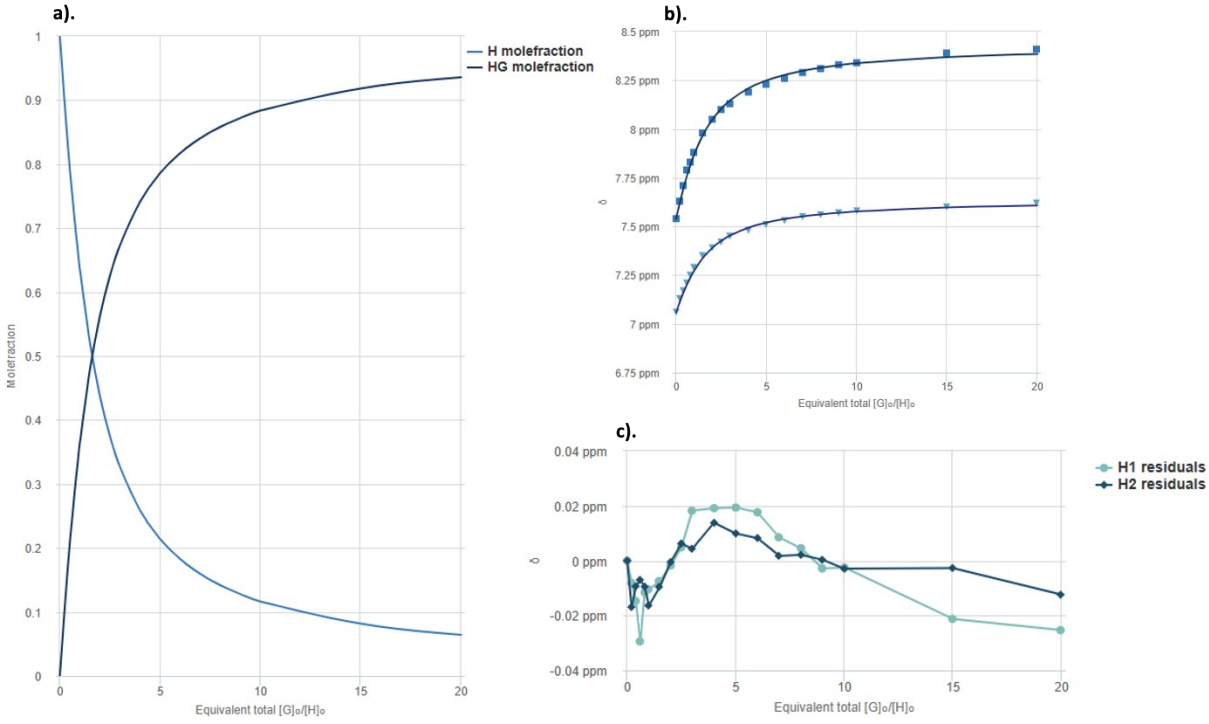
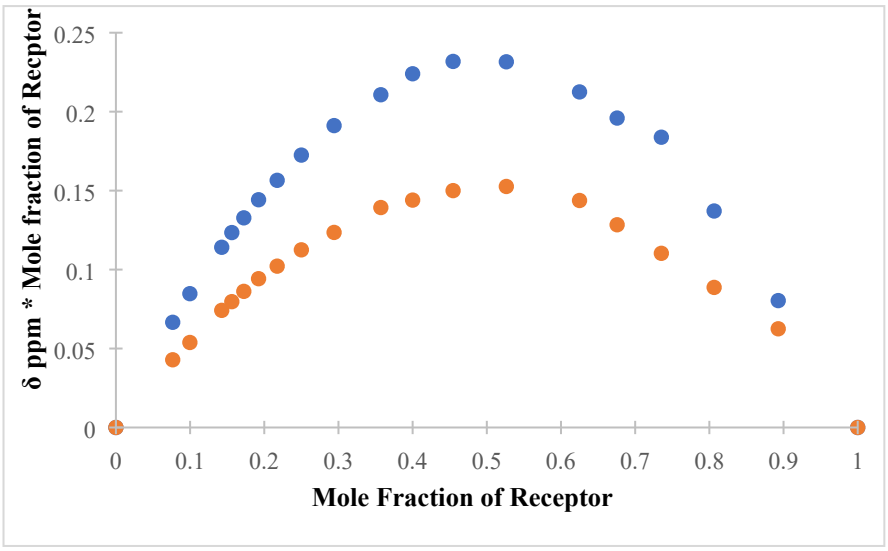


Figure S70: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Phe b).** Fitting binding isotherms of **Sq-2-Phe** with TBACl in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 1:1 binding model (K_a = 366 M^{-1}). **c).** Residuals plot of **Sq-2-Phe**



FigureS71: Jobs plot of **Sq-2-Phe** with TBACl in DMSO-d_6

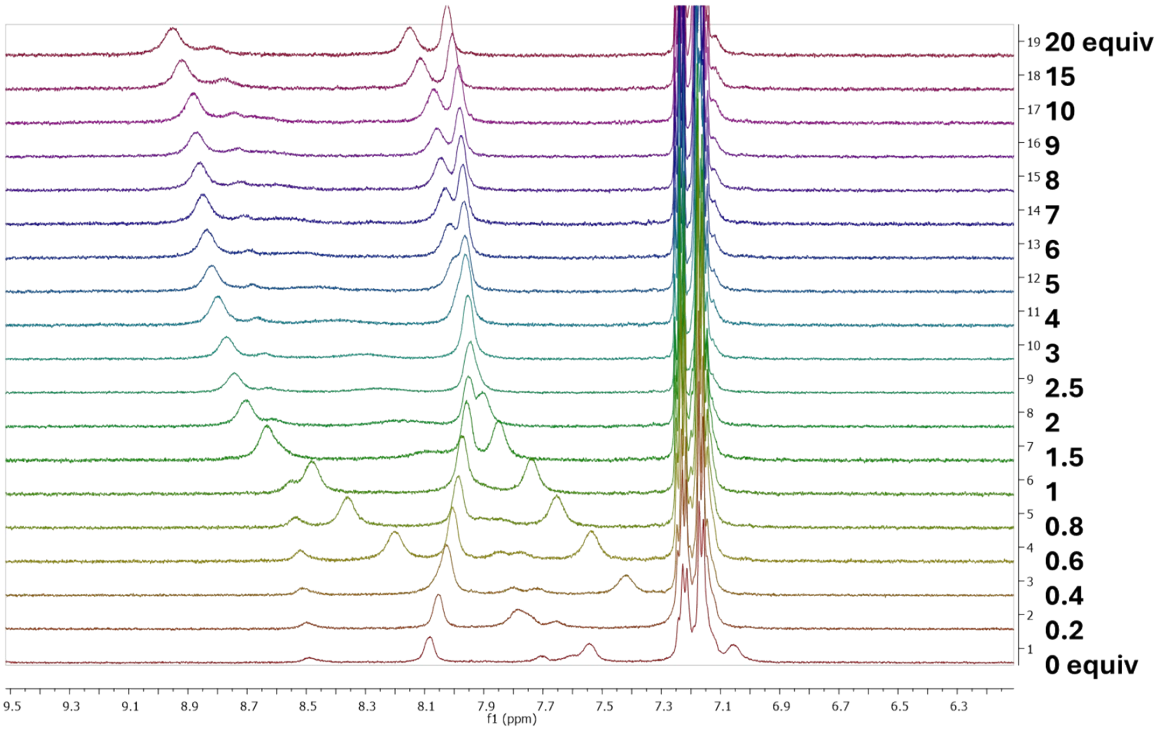


Figure S72: ^{1}H NMR titration of compound **Sq-2-Phe** with TBAAcO in DMSO-d_6

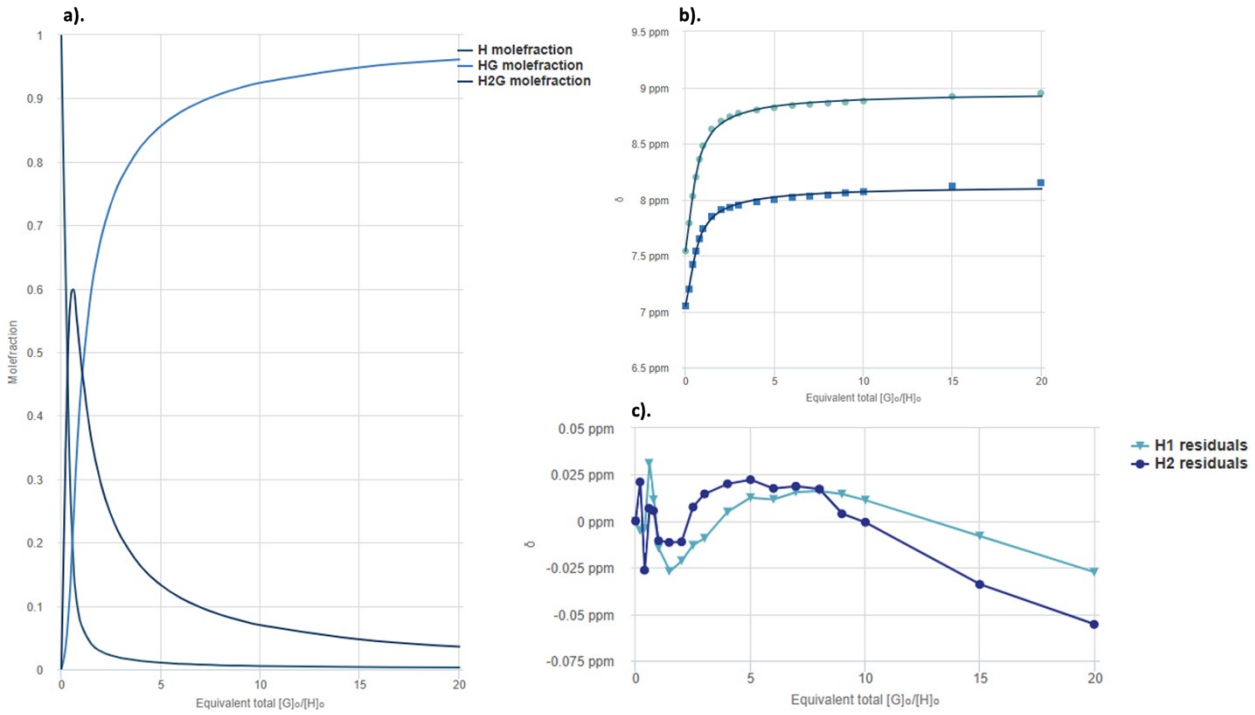


Figure S73: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Phe. b).** Fitting binding isotherms of **Sq-2-Phe** with TBAAcO in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 2:1 binding model (K_{21} = 3 x 10^3 M^{-1}). **c).** Residuals plot of **Sq-2-Phe.**

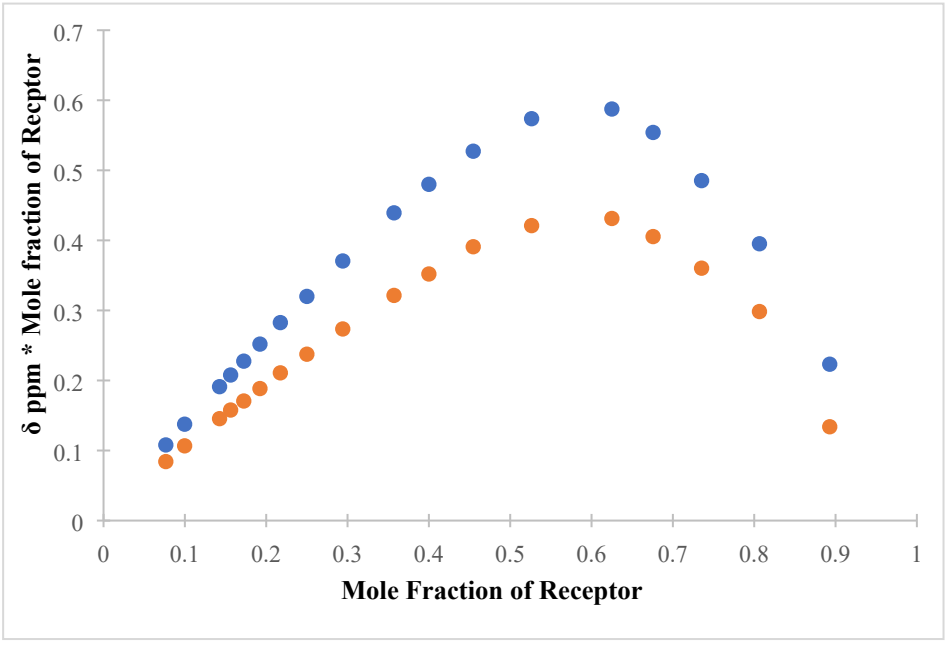


Figure S74: Jobs plot of **Sq-2-Phe** with TBAAcO in DMSO-d_6

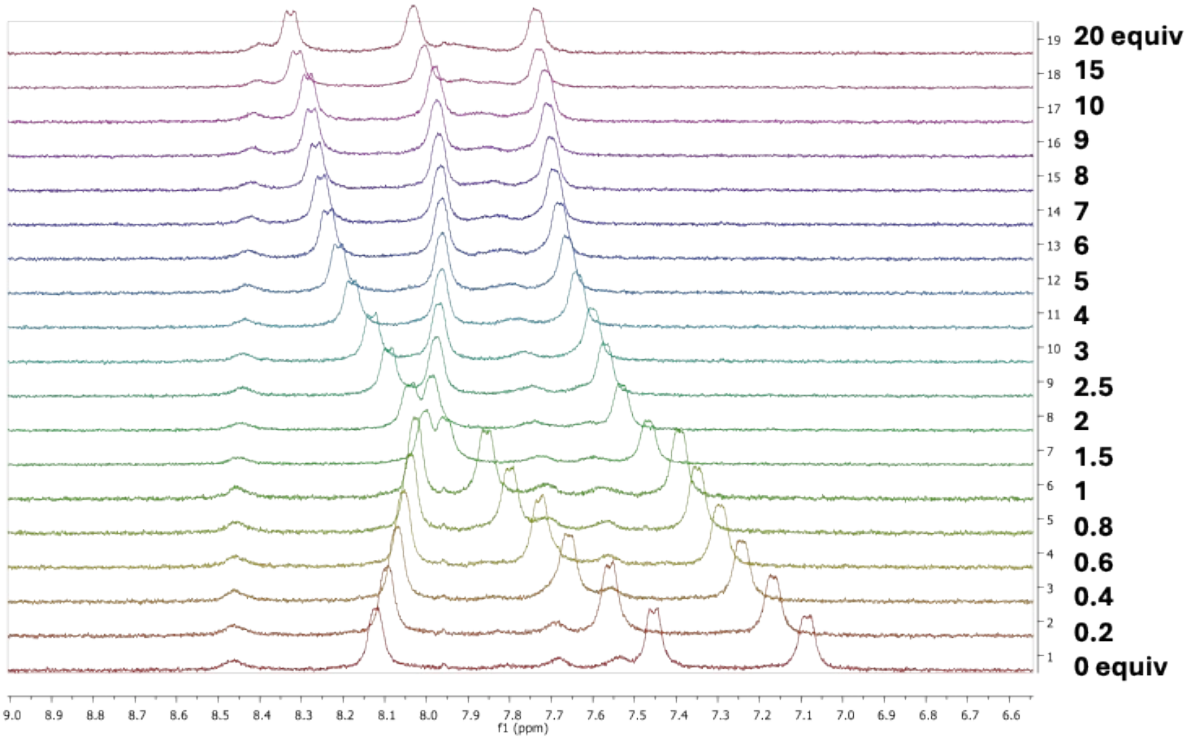


Figure S75: ^{1}H NMR titration of compound **Sq-2-Leu** with TBACl in DMSO-d_6

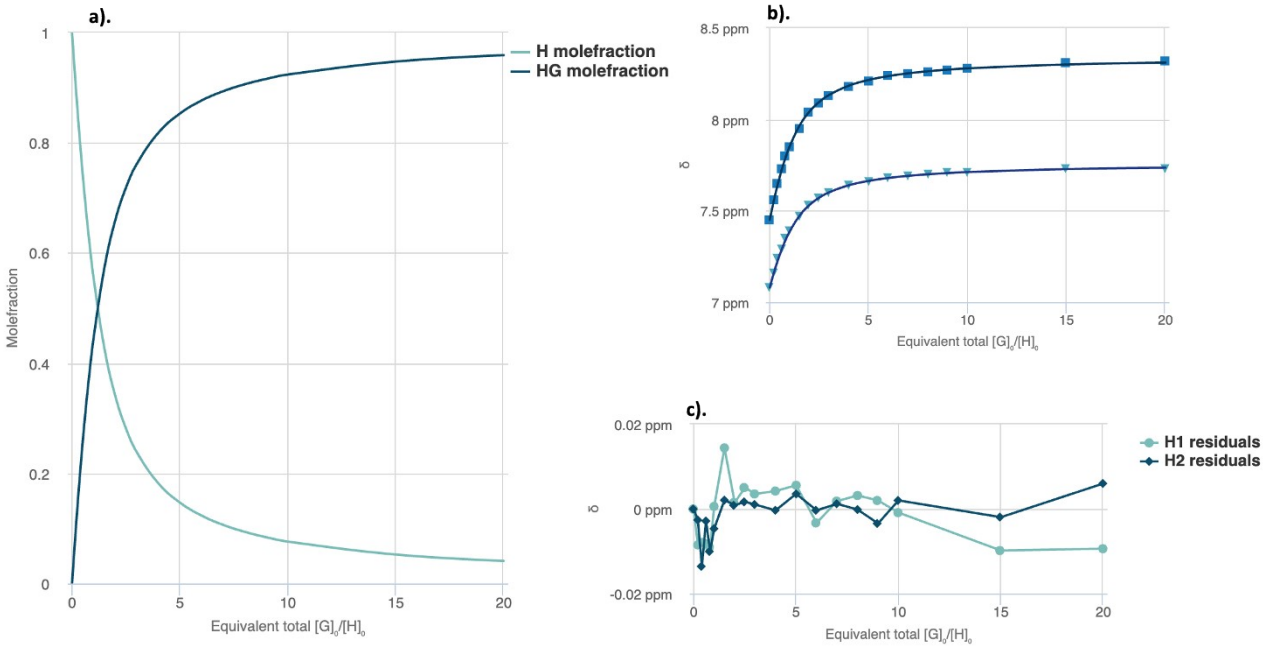


Figure S76: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Leu. b).** Fitting binding isotherms of **Sq-2-Leu** with TBACl in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 1:1 binding model (K_a = 588 M^{-1}). **c).** Residuals plot of **Sq-2-Leu**.

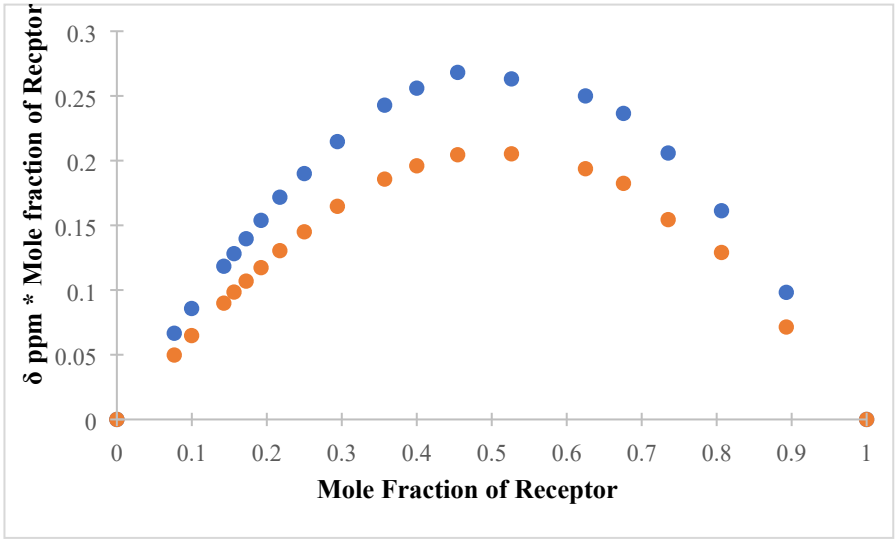


Figure S77: Jobs plot of **Sq-2-Leu** with TBACl in DMSO-d_6

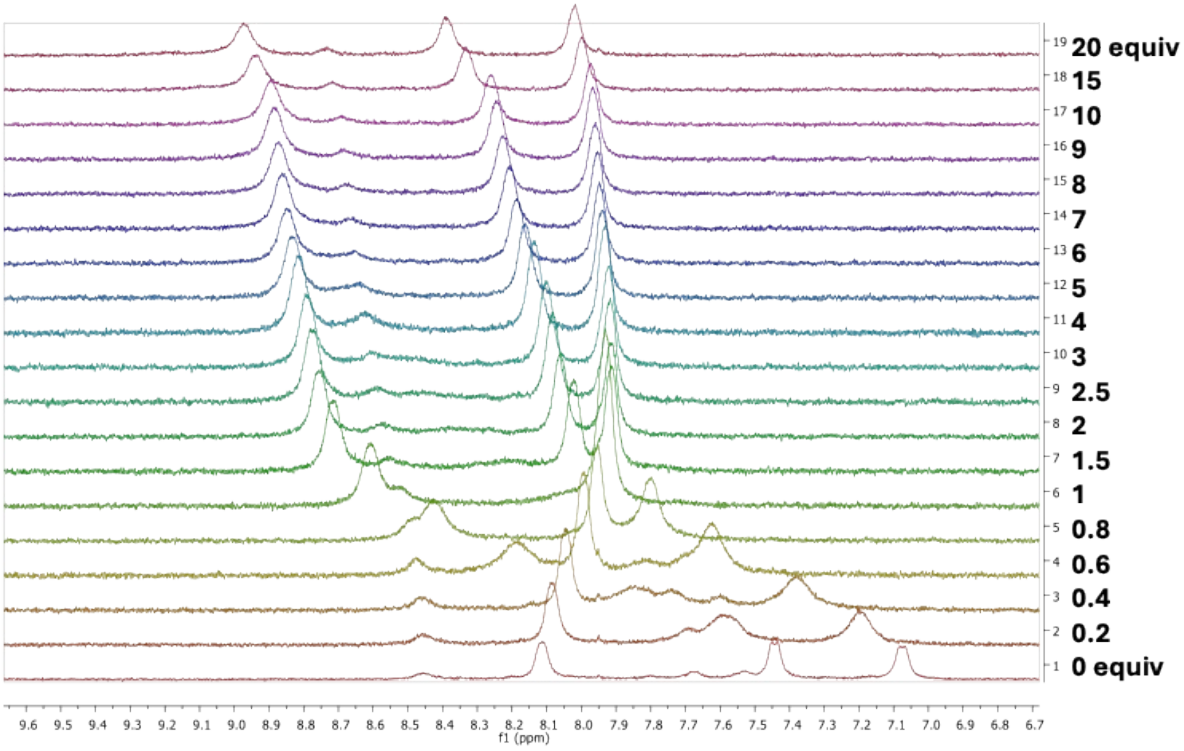


Figure S78: ^{1}H NMR titration of compound **Sq-2-Leu** with TBAAcO in DMSO-d_6

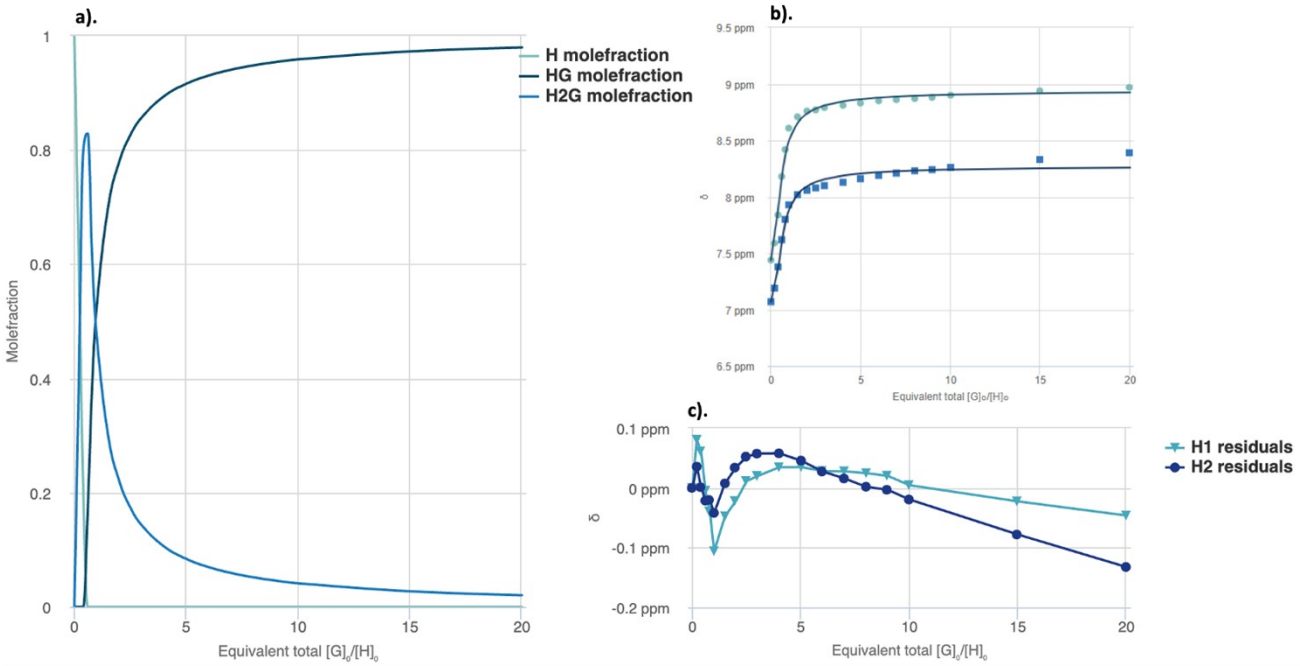


Figure S79: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Leu. b).** Fitting binding isotherms of compound **Sq-2-Leu** with TBAAcO in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 2:1 binding model (K_{21} = 10^4 M^{-1}). **c).** Residuals plot of **Sq-2-Leu**.

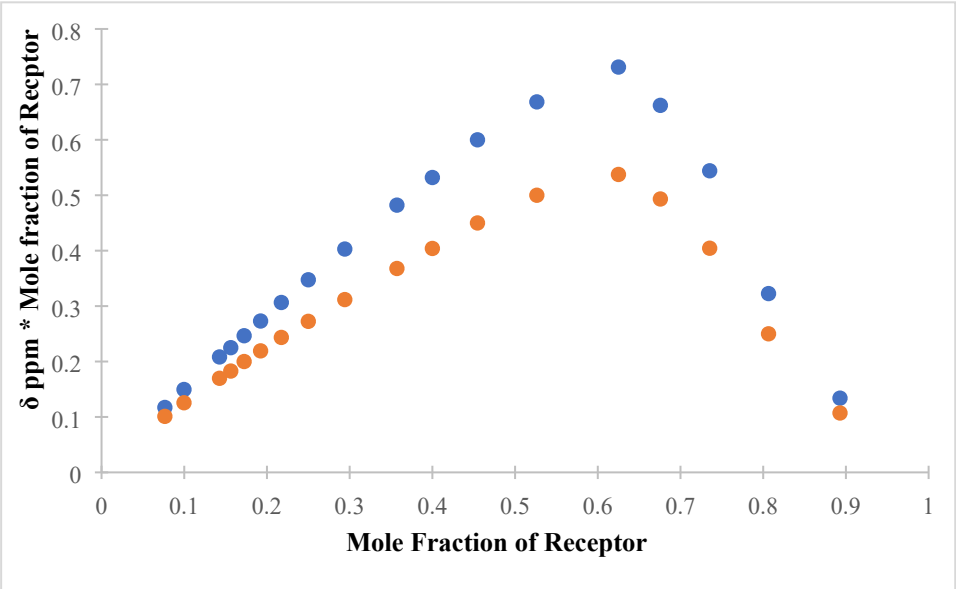


Figure S80: Jobs plot of **Sq-2-Leu** with TBAAcO in DMSO-d_6

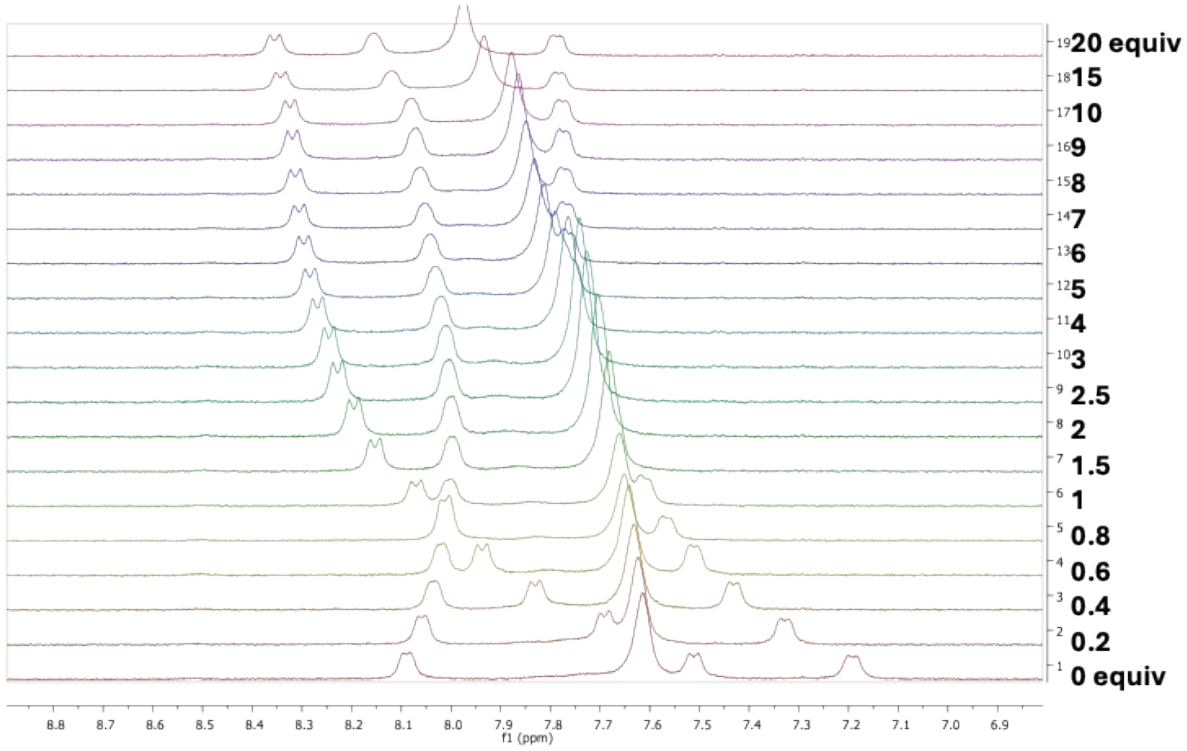


Figure S81: ^{1}H NMR titration of compound **Sq-2-Lys** with TBACl in DMSO-d_6

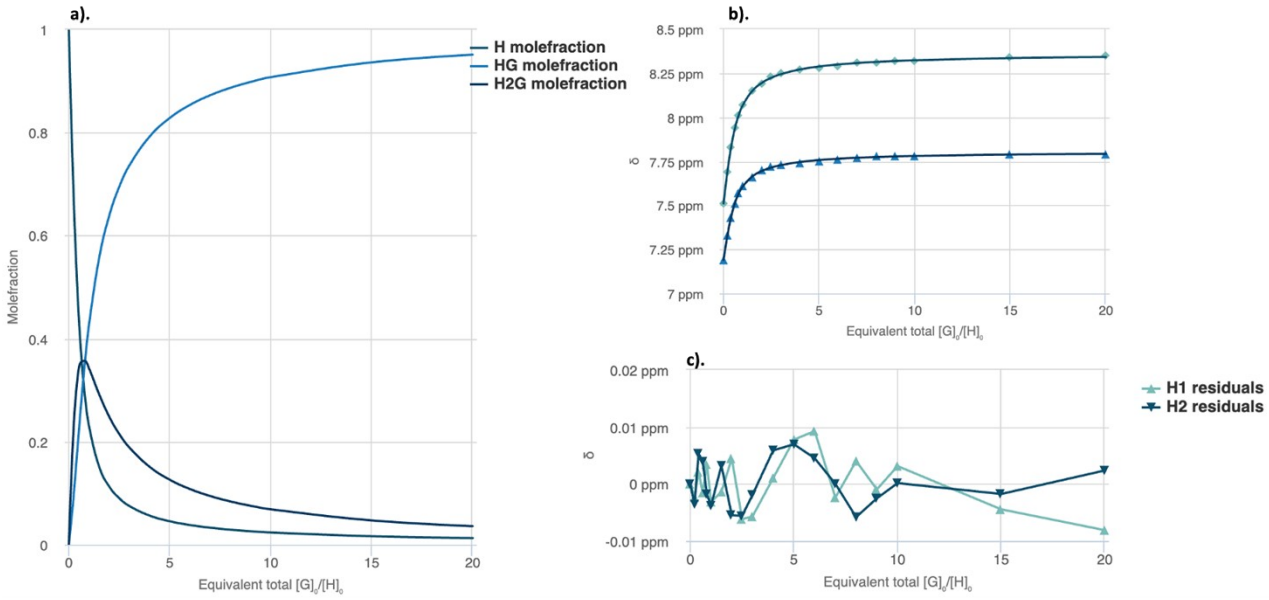


Figure S82: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Lys. b).** Fitting binding isotherms of **Sq-2-Lys** with TBACl in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 2:1 binding model (K_a = 703 M^{-1}). **c).** Residuals plot of **Sq-2-Lys**.

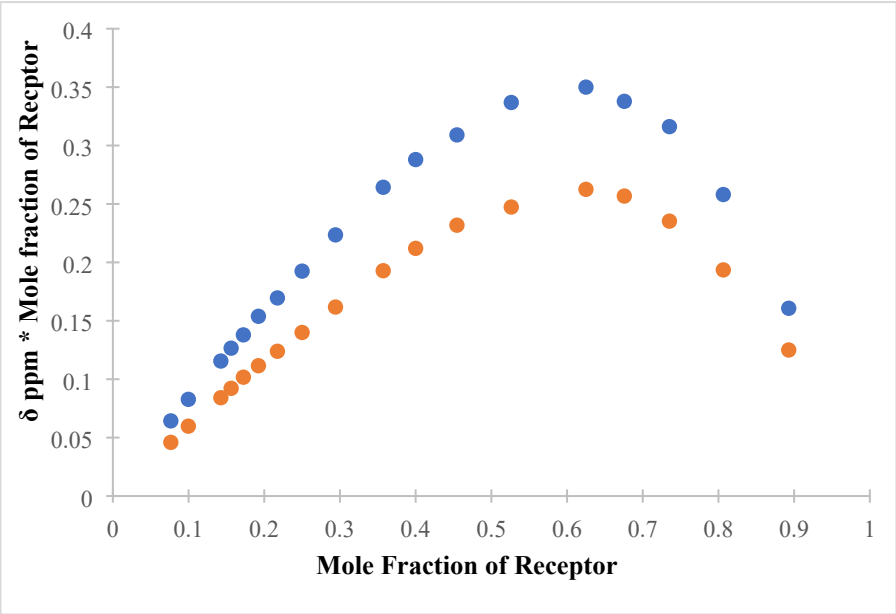


Figure S83: Jobs plot of **Sq-2-Lys** with TBACl in DMSO-d_6

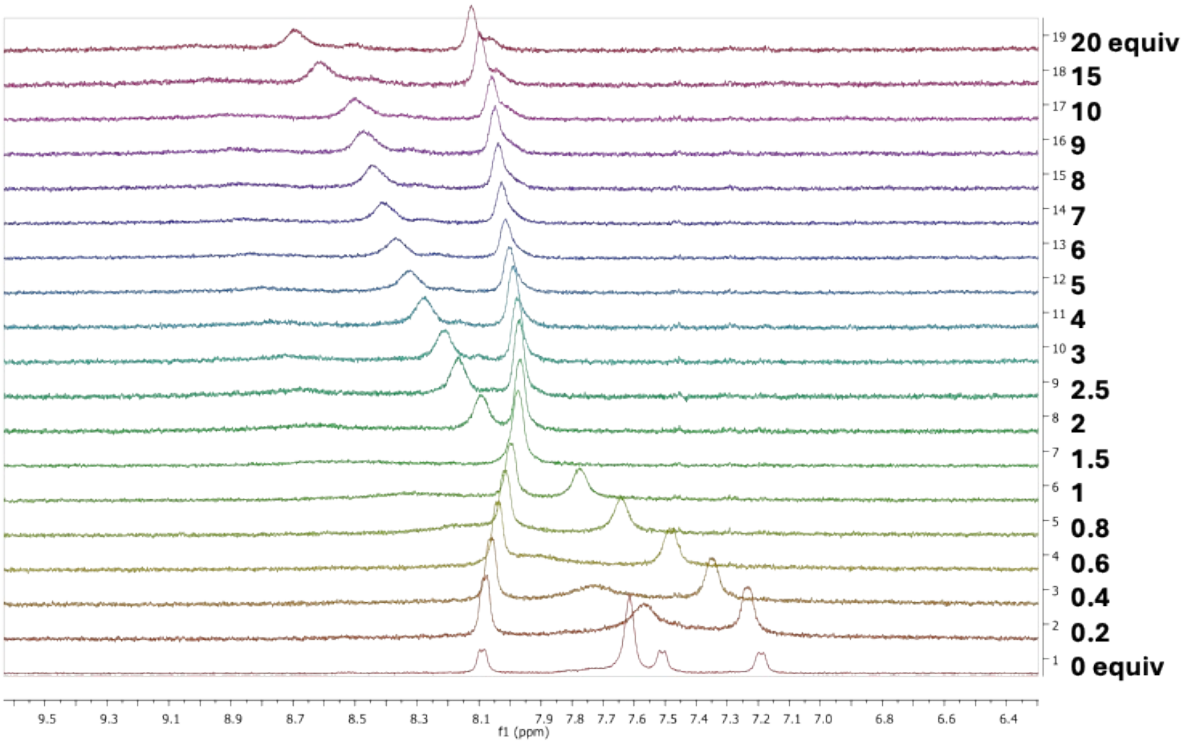


Figure S84: ^{1}H NMR titration of compound **Sq-2-Lys** with TBAAcO in DMSO-d_6

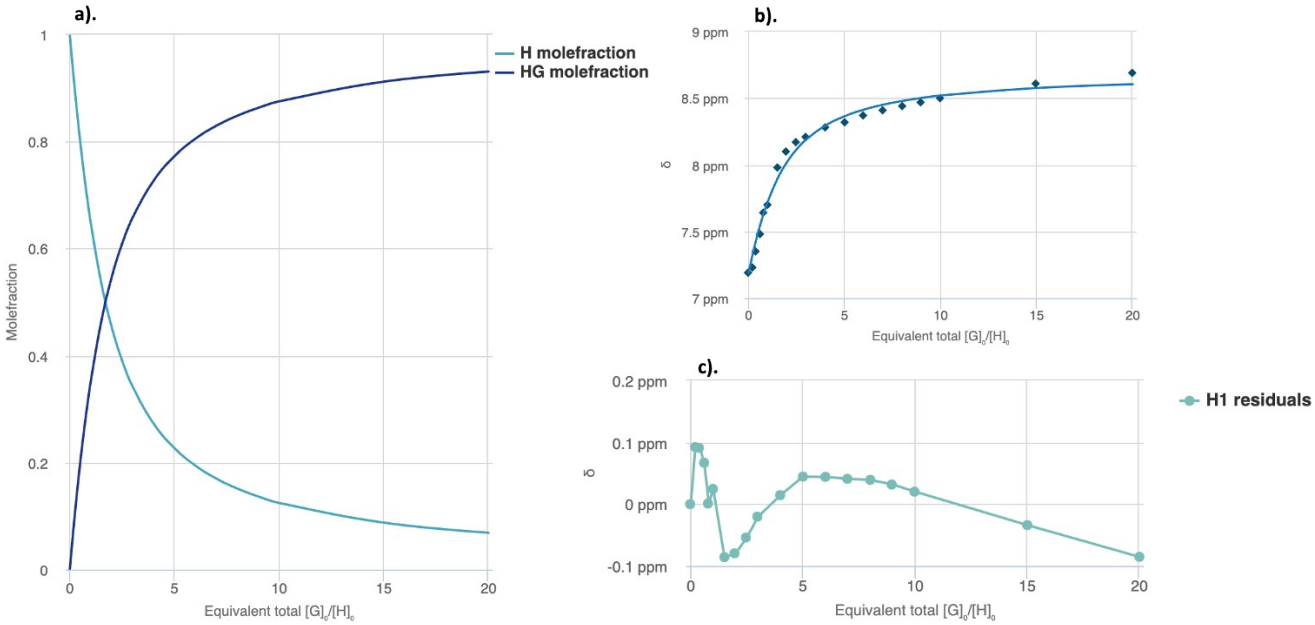


Figure S85: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Lys. b).** Fitting binding isotherms of **Sq-2-Lys** with TBAAcO in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 1:1 binding model (K_a = 333 M^{-1}). **c).** Residuals plot of **Sq-2-Lys**.

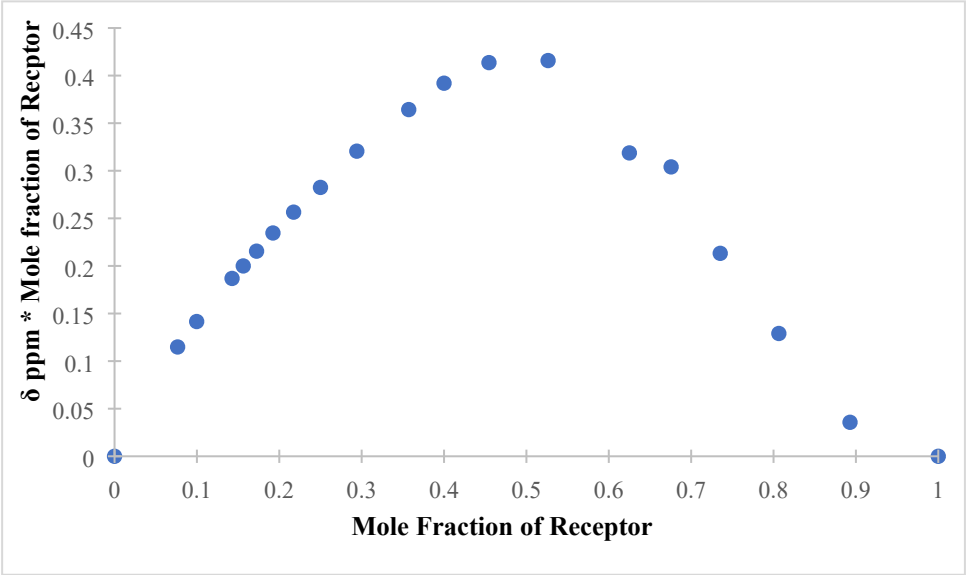


Figure S86: Jobs plot of **Sq-2-Lys** with TBAAcO in DMSO-d_6

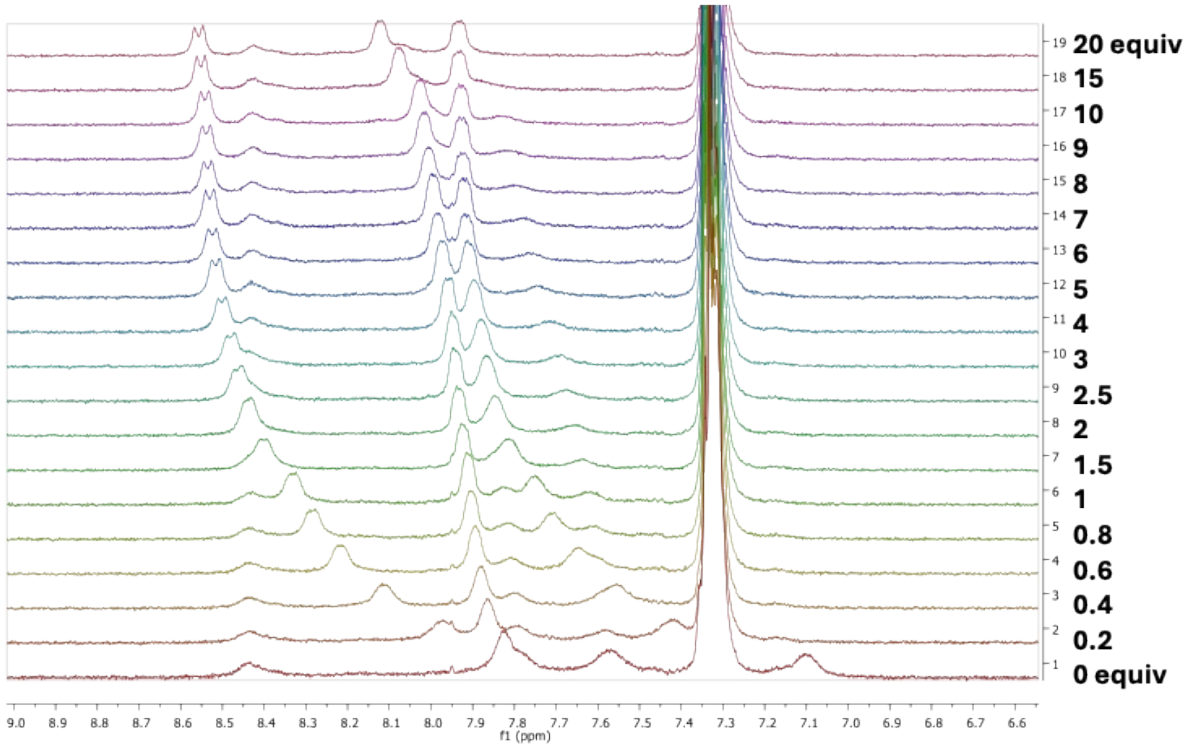


Figure S87: ^{1}H NMR titration of compound **Sq-2-Asp(Bzl)** with TBACl in DMSO-d_6

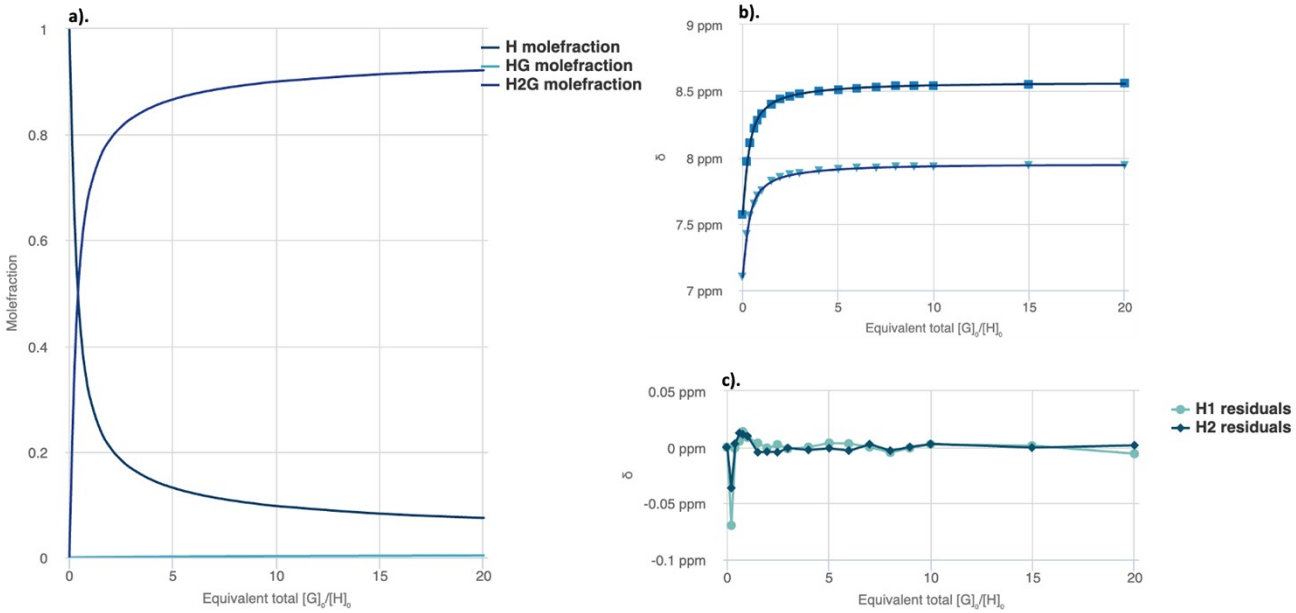


Figure S88: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Asp(Bzl). b).** Fitting binding isotherms of **Sq-2-Asp(Bzl)** with TBACl in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 2:1 binding model (K_{21} = 10^4 M^{-1}). **c).** Residuals plot of **Sq-2-Asp(Bzl)**.

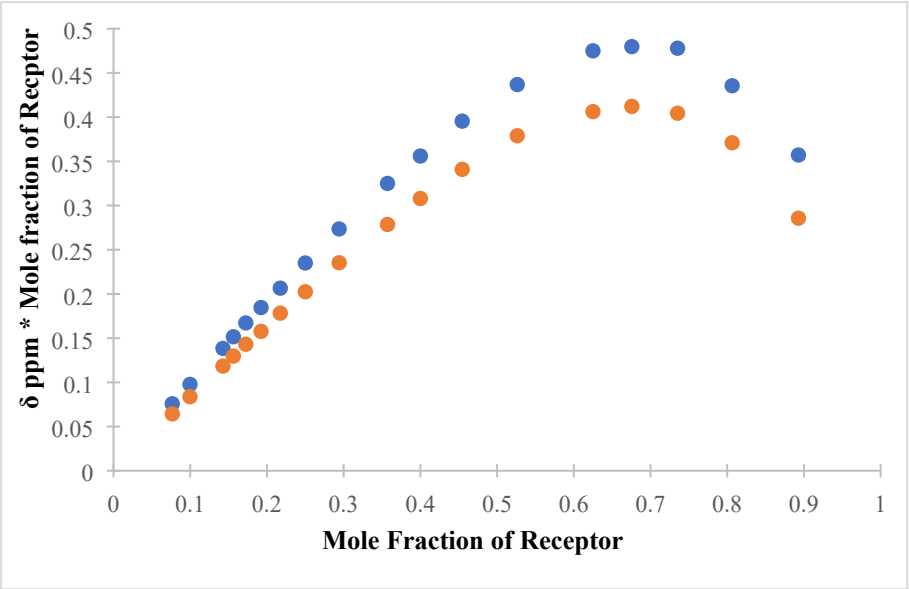


Figure S89: Jobs plot of **Sq-2-Asp(Bzl)** with TBACl in DMSO-d_6

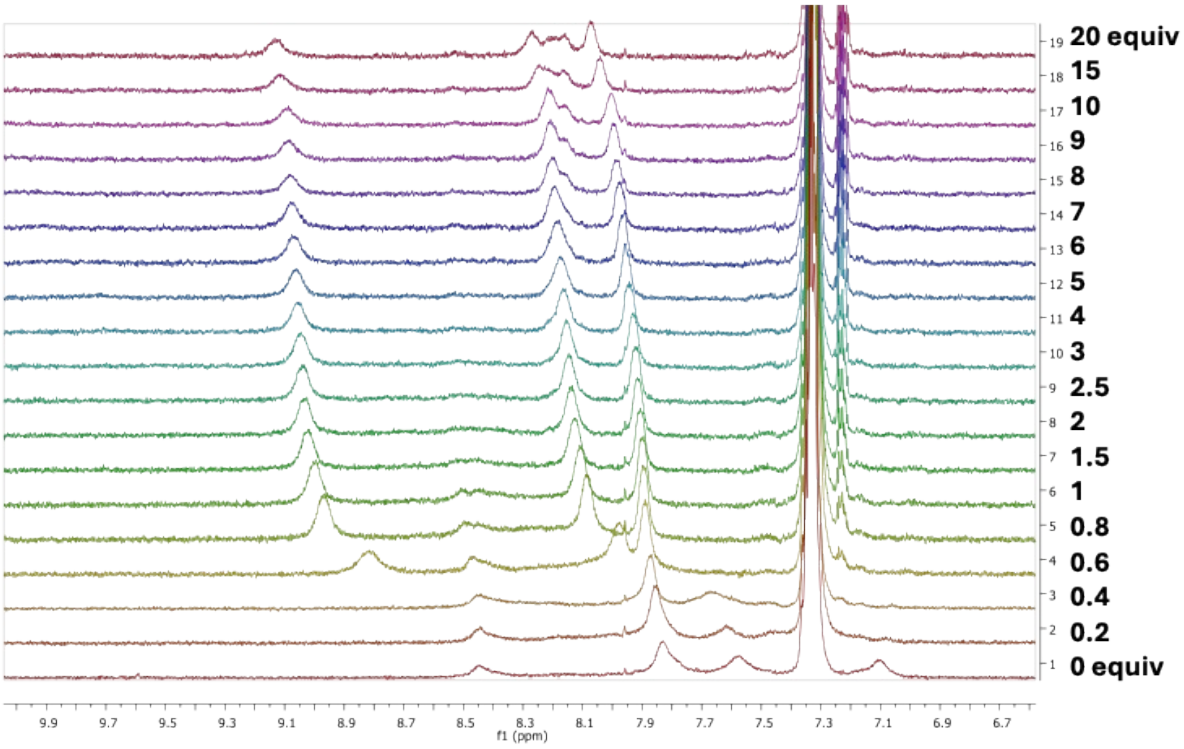


Figure S90: ^{1}H NMR titration of compound **Sq-2-Asp(Bzl)** with TBAAcO in DMSO-d_6

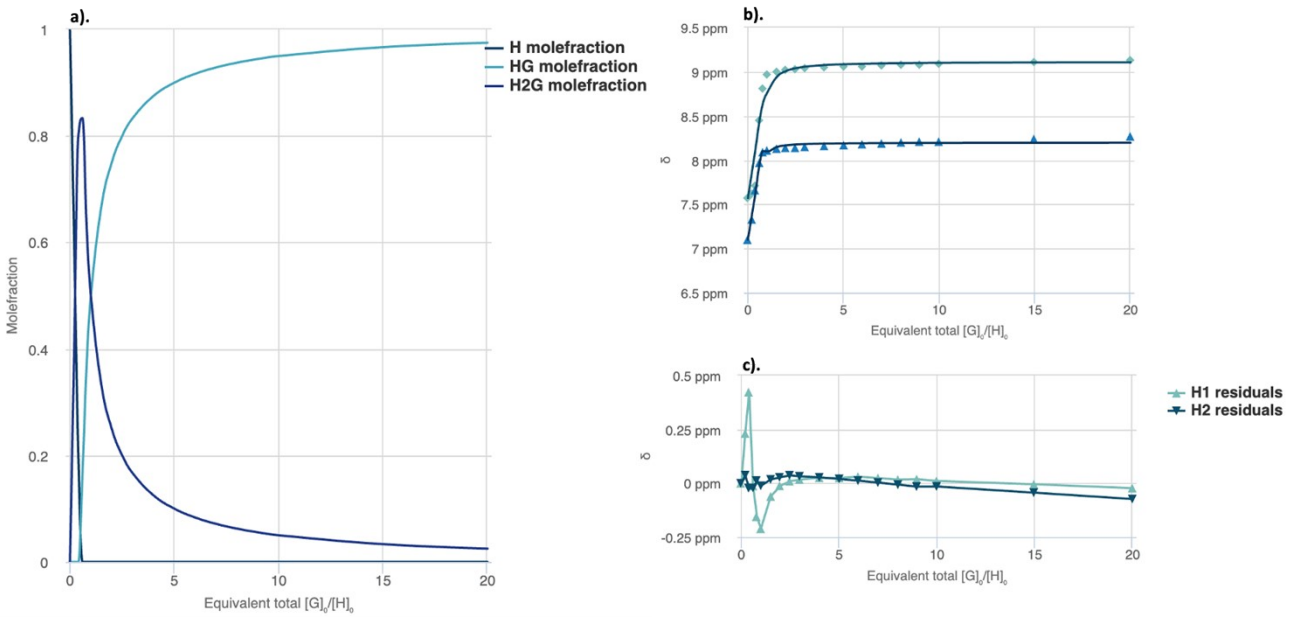


Figure S91: a). Mole fraction plot of Host vs Host:guest fraction with increasing guest concentration for **Sq-2-Asp(Bzl). b).** Fitting binding isotherms of **Sq-2-Asp(Bzl)** with TBAAcO in DMSO-d_6 at 298 K, showing the changes in chemical shifts for the squaramide NH protons fitted to the 2:1 binding model (K_{21} = 10^4 M^{-1}). **c).** Residuals plot of **Sq-2-Asp(Bzl)**.

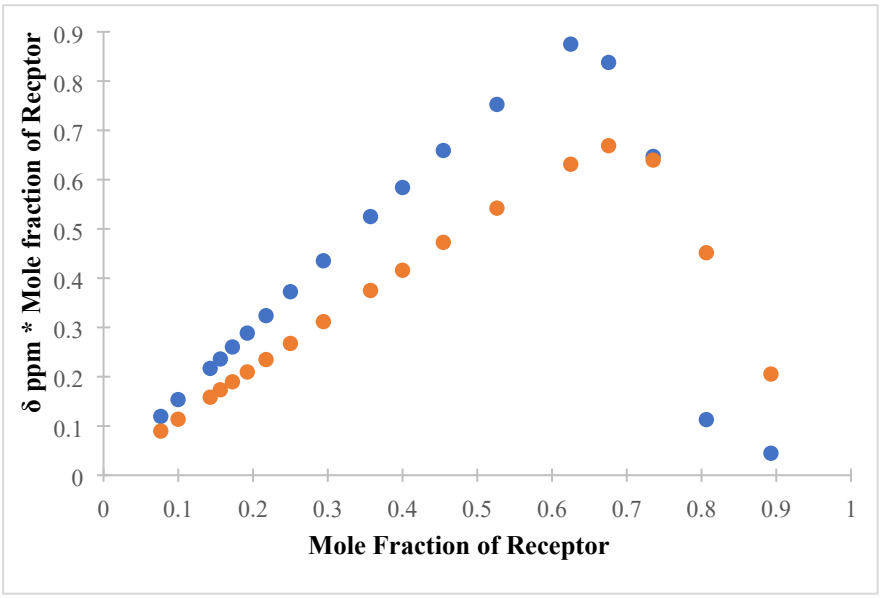


Figure S92: Jobs plot of **2 Sq-2-Asp(Bzl)** with TBAAcO in DMSO-d_6.

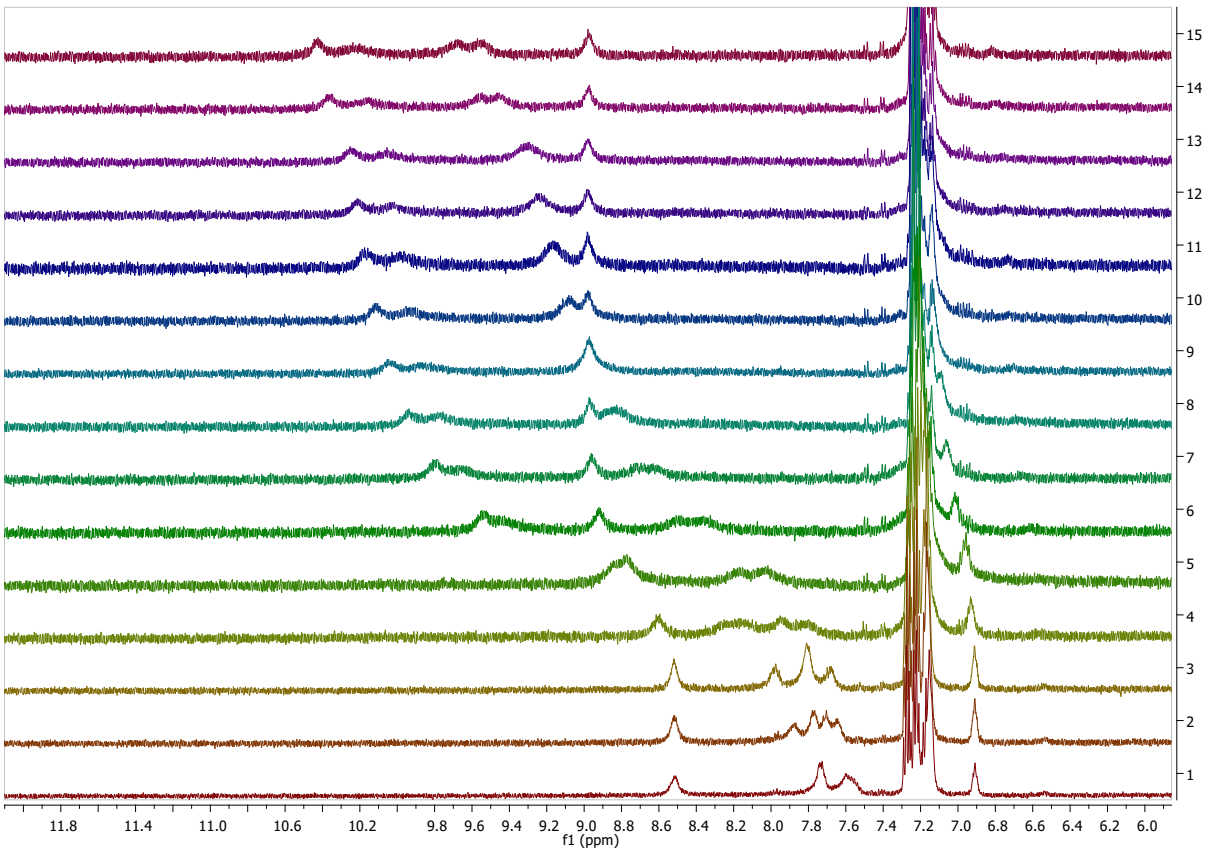


Figure S93: ^{1}H NMR titration of compound **ASq-2-Phe** with TBAAcO in DMSO-d_6

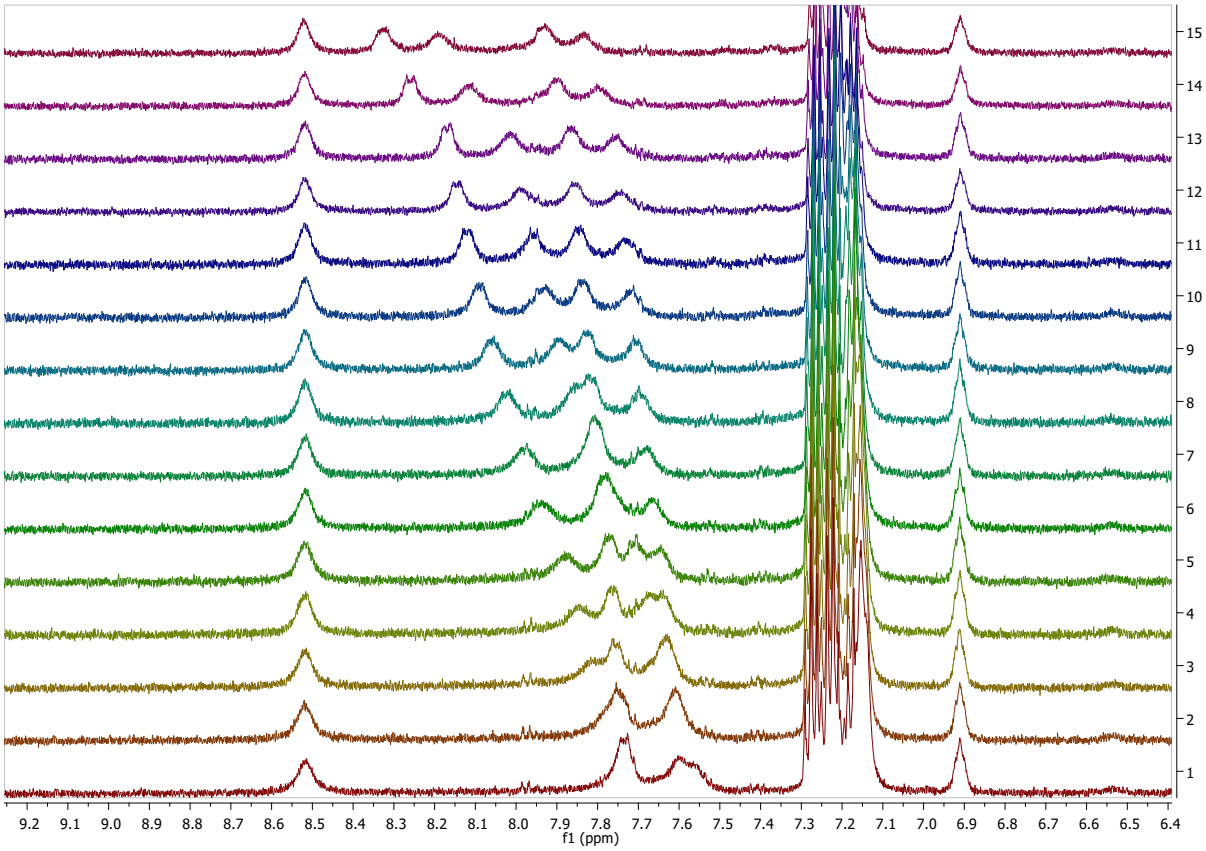


Figure S94: ^{1}H NMR titration of compound **ASq-2-Phe** with TBACl in DMSO-d_6

6. ^{1}H NMR Concentration Dependent Study

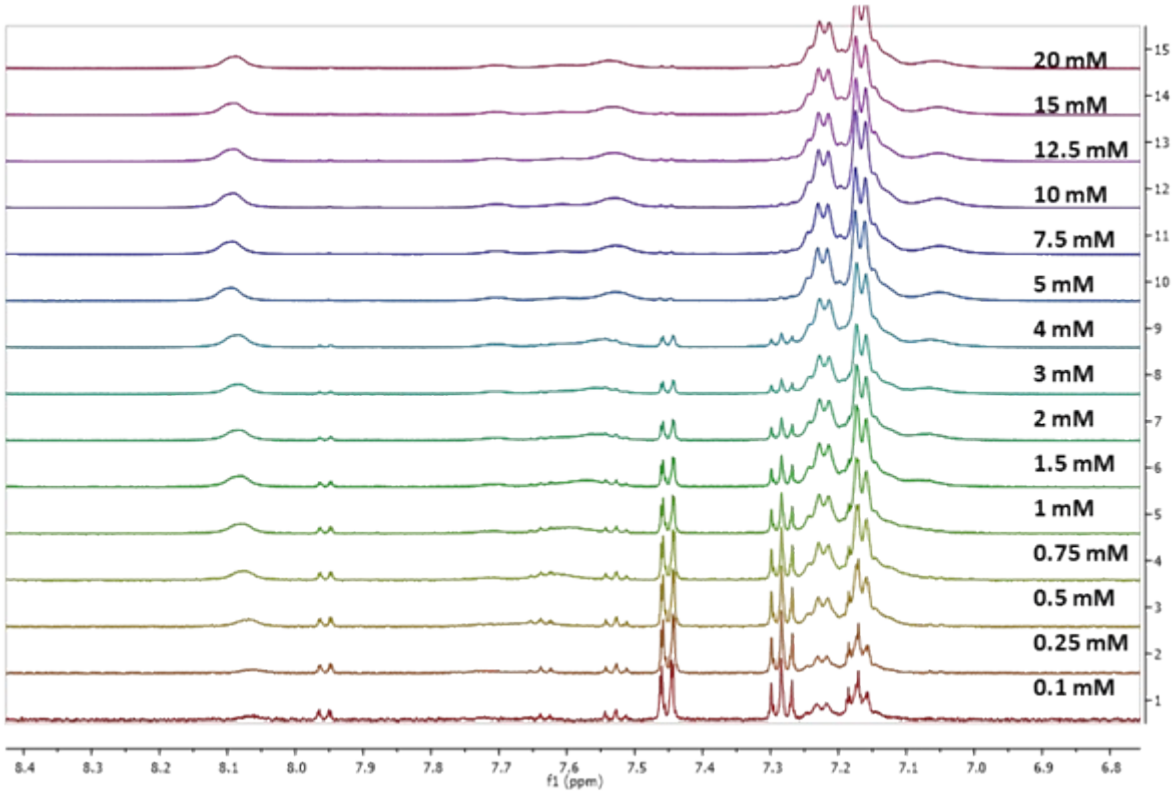


Figure S95: Concentration dependent ^{1}H NMR spectra of **Sq-2-Phe**

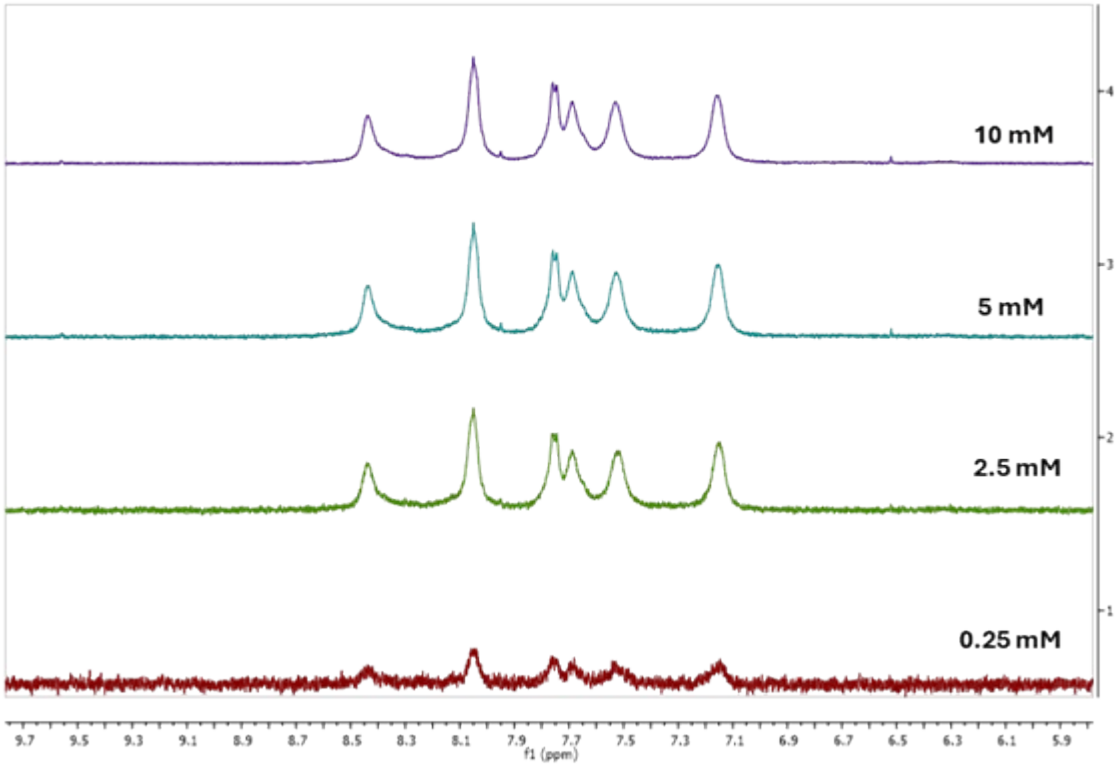


Figure S96: Concentration dependent ^{1}H NMR spectra of **Sq-2-Ala**

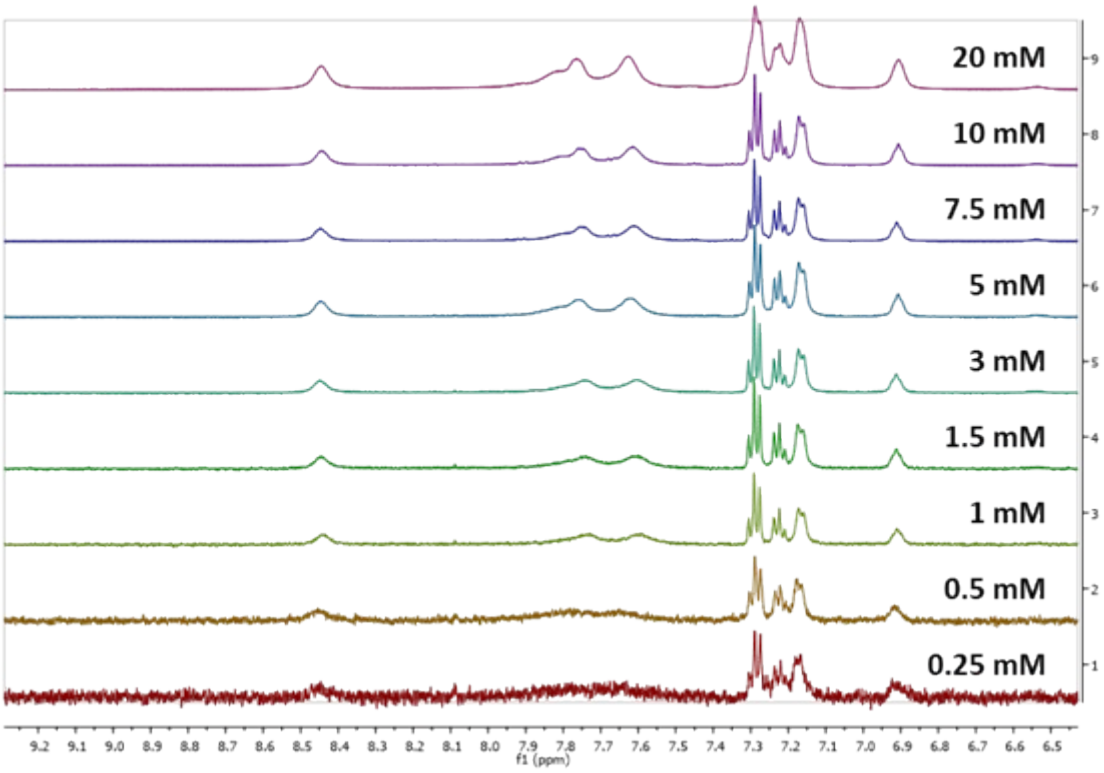


Figure S97: Concentration dependent ^{1}H NMR spectra of **ASq-2-Phe**

7. Circular Dichroism Spectroscopy

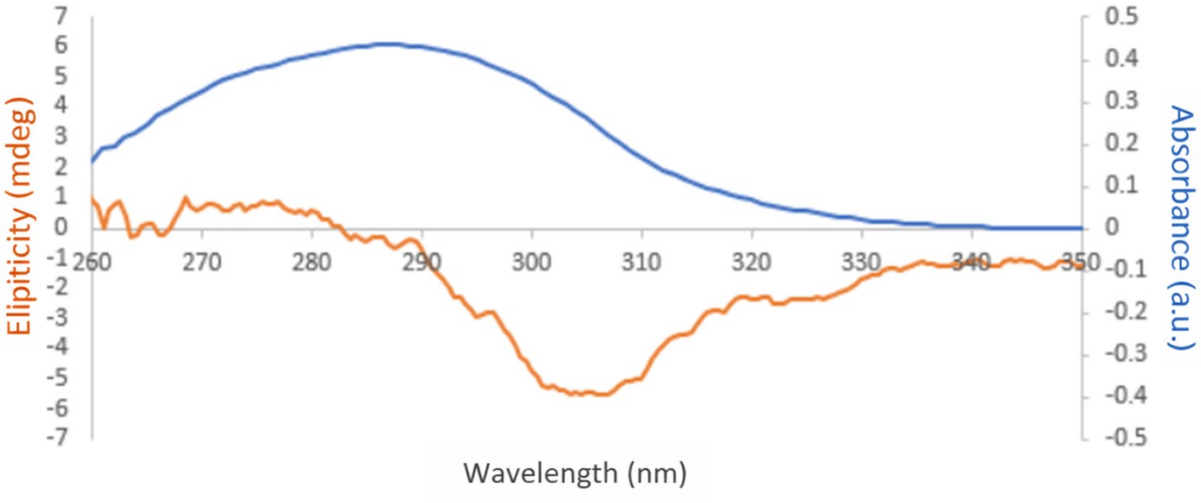


Figure S98: Absorption (Blue) and circular dichroism (orange) spectra of **Sq-2-Phe**.

8. X-ray Crystallography

Crystal and refinement parameters are given in Table S1. All data were collected on a Bruker D8 Quest ECO diffractometer using graphite-monochromated Mo Kα radiation (λ = 0.71073 Å) and a Photon II-C14 CPAD detector. Crystals were mounted on Mitegen micromounts in NVH immersion oil, and all collections were carried out at 150 K using an Oxford cryostream. Data collections were carried out using ϕ and ω scans, with collections and data reductions carried out in the Bruker APEX-3 suite of programs.[1] Multi-scan absorption corrections were applied for all datasets using SADABS.[2] The data were solved with the intrinsic phasing routine in SHELXT,[3] and all data were refined on F^2 with full-matrix least squares procedures in SHELXL,[4] operating within the OLEX-2 GUI.[5] All non-hydrogen atoms were refined with anisotropic displacement parameters. Hydrogen atoms were placed in riding positions and refined with isotropic displacement parameters equal to 1.2 or 1.5 times the isotropic equivalent of their carrier atom. Absolute configuration was determined based on the anomalous dispersion contribution from the sulfur atoms in the DMSO solvates and matches the expected configuration on the basis of the synthesis. CCDC 2287336.

Table S1: Crystal data and structure refinement for **Sq-2-Ala**.

Identification code	SQ-2-Ala
Empirical formula	$C_{22}H_{34}N_6O_8S_2$
Formula weight	574.67
Temperature/K	150.0
Crystal system	monoclinic
Space group	$P2_1$
a/Å	8.8611(3)
b/Å	10.5198(4)
c/Å	15.0198(5)
α/°	90
β/°	103.0770(10)
γ/°	90
Volume/Å^3	1363.79(8)
Z	2
ρ_{calc}g/cm^3	1.399
μ/mm^{-1}	0.252
F(000)	608.0
Crystal size/mm^3	0.36 × 0.29 × 0.16
Radiation	MoKα (λ = 0.71073)
2Θ range for data collection/°	6.106 to 61.062
Index ranges	-12 ≤ h ≤ 12, -15 ≤ k ≤ 15, -21 ≤ l ≤ 21
Reflections collected	38125
Independent reflections	8257 [R_{int} = 0.0228, R_{sigma} = 0.0182]
Data/restraints/parameters	8257/1/349
Goodness-of-fit on F^2	1.047
Final R indexes [I>=2σ (I)]	R_1 = 0.0246, wR_2 = 0.0658
Final R indexes [all data]	R_1 = 0.0255, wR_2 = 0.0664
Largest diff. peak/hole / e Å^{-3}	0.29/-0.29
Flack parameter	0.002(9)
CCDC No.	2287336

Table S2: Hydrogen bonding parameters for **Sq-2-Ala**

D	H	A	d(D-H)/Å	d(H-A)/Å	d(D-A)/Å	D-H-A/°
N1	H1	O7	0.88	1.87	2.7330(16)	165.3
N2	H2	O7	0.88	2.25	3.0364(17)	148.7
N3	H3	O8[1]	0.88	2.17	3.0172(18)	160.7
N4	H4	O2[2]	0.88	2.08	2.9180(15)	157.9
N5	H5	O2[2]	0.88	2.04	2.8818(16)	159.3
N6	H6	O8[3]	0.88	1.95	2.7737(16)	155.9

[1]1-*x*,1/2+*y*,-*z*; [2]-1+*x*,+*y*,+*z*; [3]1-*x*,-1/2+*y*,-*z*

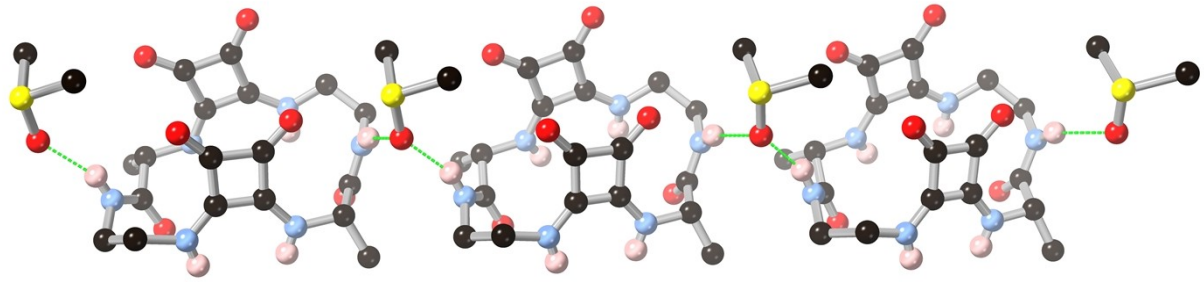


Figure S99: The hydrogen bonding mode of the amide groups within the **Sq-2-Ala** macrocycle forming a complementary 1-dimensional chain through the DMSO solvate, perpendicular to that formed by the squaramide···squaramide contacts. Selected hydrogen atoms and DMSO molecules are omitted for clarity.

9. DFT models of anion binding to Sq-2-Ala receptor

Density functional theory (DFT) calculations were carried out using the Gaussian 16 [6] package. Geometry optimisations and frequency calculations were performed at the B3LYP/cc-pvtz level of theory. Dispersion interactions were included using the Grimme D2 dispersion correction.[7] All calculations were conducted in the gas phase in order to isolate intrinsic host–guest interaction trends across the monovalent anion series. The structures of the free receptor (Sq-2-Ala), the isolated anions (Cl^-, Br^-, I^-, and NO_3^-), and their corresponding host–guest complexes were fully optimised without symmetry constraints. Binding energies were calculated according to Equation S1:

$$\Delta E_{bind} = E_{anion:Sq-2-Ala} - (E_{Sq-2-Ala} + E_{anion}) \quad \text{(Eq. S1)}$$

where $E_{anion:Sq-2-Ala}$ is the total electronic energy of the anion–receptor complex, and $E_{Sq-2-Ala}$ and E_{anion} are the total energies of the isolated receptor and anion, respectively.

To account explicitly for basis set superposition error (BSSE) [8], binding energies were corrected using the counterpoise (CP) method applied to the optimised geometries. All reported binding energies therefore correspond to BSSE-corrected gas-phase interaction energies (ΔE_{bind}).

Zero-point energy and thermal corrections were evaluated and found to be small (< 0.05 eV) and do not affect the relative binding trends discussed. While absolute binding energies are expected to be overestimated in the gas phase, the BSSE-corrected results provide a reliable basis for comparing relative anion selectivity and hydrogen-bonding strength of the Sq-2-Ala host.

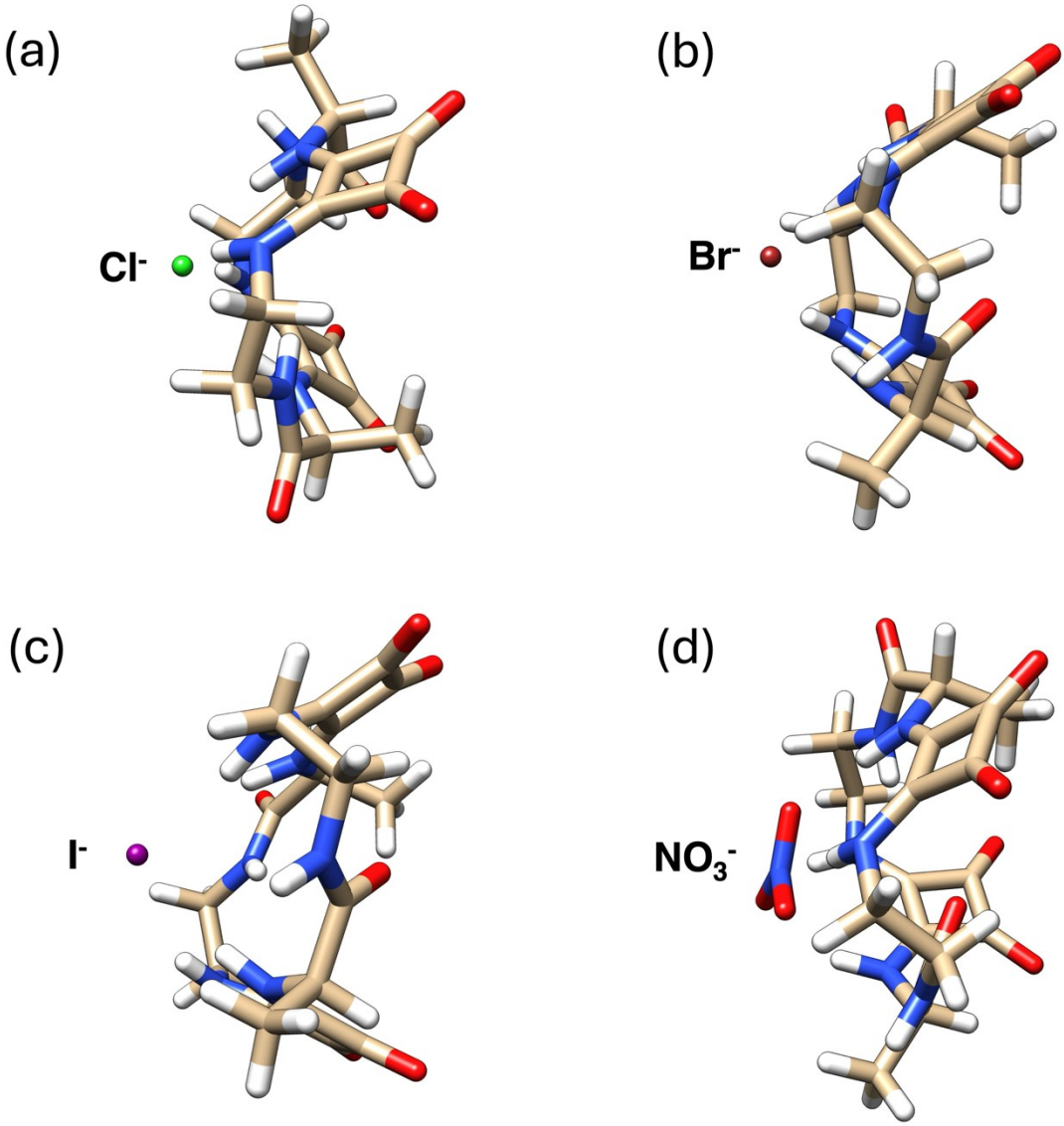


Figure S100: Side-view representations of the DFT-optimised Sq-2-Ala–anion complexes for (a) Cl^-, (b) Br^-, (c) I^-, and (d) NO_3^-. The views highlight the relative distance of each anion from the squaramide hydrogen-bond donor plane, illustrating significant and size-dependent binding of halides and weaker binding of nitrate.

References

1. *Bruker APEX-3*, Bruker-AXS Inc., Madison, WI, 2016.

2. *SADABS*, Bruker-AXS Inc., Madison, WI, 2016.

3. G. M. Sheldrick, *Acta Crystallogr., Sect. A: Found. Adv.*, 2015, **71**, 3–8.

4. G. M. Sheldrick, *Acta Crystallogr., Sect. C: Struct. Chem.*, 2015, **71**, 3–8.

5. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.

6. Gaussian 16 Rev. C.01 v. C.01 (Wallingford, CT, 2016).

7. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. *J Chem Phys, 2010* **132**, 154104.

8. Richard, R. M., Bakr, B. W. & Sherrill, C. D. *J Chem Theory Comput 2018* **14**, 2386–2400.