Electronic Supplementary Information

Simple toroids to multi-torus structures from self-assembling peptides

Souvik Dutta,[a] and V. Haridas[a,b]*

[a]Department of Chemistry, Indian Institute of Technology Delhi, Hauz Khas, New Delhi-110016, India.

[b]Department of Chemistry, Indian Institute of Technology Palakkad, Palakkad, Kerala-678623, India.

*V. Haridas. Email: haridasv@chemistry.iitd.ac.in, haridasv@iitpkd.ac.in.

1. General Information

1.1 Materials

All solvents employed were distilled prior to utilization. All amino acids were of L-configuration. Reactions were monitored by thin layer chromatography (TLC) using TLC plates purchased from Merck, India. Silica gel with a mesh size of 100-200 was employed for purification in column chromatography. NMR was recorded using Brucker-DPX-400 and 500 MHz spectrometers. Tetramethylsilane (TMS) was used as the internal standard. Coupling constants are recorded in Hz, and the ^{1}H NMR data are denoted as s (singlet), d (doublet), br (broad), t (triplet) and m (multiplet), dd (double of doublet). IR spectra (KBr) pellets were recorded on Nicolet, Protégé 460 spectrometer. Bruker MicrO-TOF-QII model High-resolution mass spectrometer (HRMS) was used for characterization. Agilent Cary-60 dual beam spectrophotometer was used for recording UV-visible spectra measurements. FluoroMax-4 spectrofluorometer (HORIBA JOBIN YVON Scientific), was employed to record fluorescence spectra. A slit width of 2 nm excitation was used for data collection. Samples were taken 3 mL quartz cuvette. Melting points were recorded in Fisher-Scientific melting point apparatus. All spectroscopic experiments were done at room temperature. The confocal imaging was done in LEICA STELLARIS Stimulated Emission Depletion Microscope.

1.2 Methods

1.2.1 Scanning Electron Microscopy (SEM)

Solutions of compounds **1** and **3** were prepared by dissolving 1 mg of each compound in 1 mL of methanol. About 5 µL of the solution was drop-casted onto the coverslip, attached to a stub using carbon tape, and left to dry at room temperature. Further, it was coated with gold (~10 nm) and analyzed by ZEISS EVO 50 Scanning Electron Microscope. The images were captured at room temperature and were processed using Image J software.

1.2.2 Transmission Electron Microscopy (TEM)

Samples for TEM were prepared by dissolving the compound in methanol. About 5 µL aliquot of the sample solution was drop-casted on a 200 mesh copper grid and allowed to dry at room temperature. Samples were observed using a TECHNAI G2 (20S-TWIN) electron microscope. Images were captured at room temperature and were processed using Image J software.

1.2.3 Atomic Force Microscopy (AFM)

Solutions of compounds were prepared by dissolving 1 mg of each compound in 1 mL of methanol. About 50 µL of the solution was drop-casted onto a glass slide and left to dry at room temperature. Further, it was analyzed by BRUKER Dimension Icon XR Atomic Force Microscope. The images captured at room temperature were processed using Image J software.

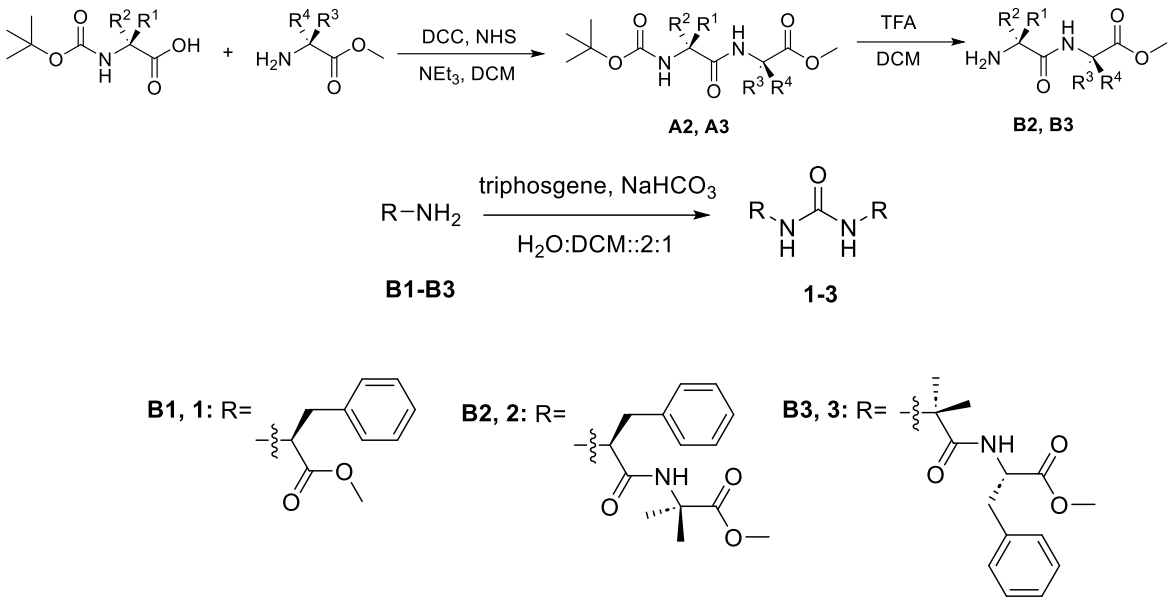
1.2.4 Confocal Microscopy

Solutions of compounds were prepared by dissolving them in methanol. About 10 µL of the solution was drop-casted onto a glass slide and left to dry at room temperature.

Further, it was covered with a cover slip and analyzed by LEICA STELLARIS Stimulated Emission Depletion Microscope. The images were captured at room temperature and were processed using Image J software.

1.2.5 NMR titration study

To 10 mM host solution in $CDCl_3$, small amounts of DMSO-d_6 were added and the NMR spectra were recorded.



Scheme S1. Synthesis of Urea-based peptides **1-3**

1.3 General Synthetic Procedure

1.3.1. General Synthetic Procedure of dipeptides A2 and A3

N-hydroxy succinimide (NHS) (276 mg, 2.40 mmol, 1.2 equiv.) and dicyclohexylcarbodiimide (DCC) (495 mg, 2.40 mmol, 1.2 equiv.) were added to an ice-cooled solution of Boc-protected amino acid (2.00 mmol, 1.0 equiv.) in dry dichloromethane (DCM) and stirred. After 5 minutes, amino acid methyl ester hydrochloride (2.40 mmol, 1.2 equiv.) in ~50 mL dry DCM and NEt_3 (0.33 ml, 2.40 mmol, 1.2 equiv.) were added and stirred for 24 h. The reaction mixture was then filtered, and the clear filtrate obtained was washed sequentially with 0.2 N H_2SO_4, aq. $NaHCO_3$ solution and finally with water. The organic part was dried over anhydrous Na_2SO_4, evaporated, and purified by silica gel column chromatography using ethyl acetate and hexane as eluents to afford pure dipeptides **A2** and **A3**.

1.3.2. Synthetic Procedure of Peptide 1

A solution of phenylalanine methyl ester hydrochloride (431.4 mg, 2.00 mmol, 1.0 equiv.) in dry DCM was prepared and was divided into two equal parts. One part (1.00 mmol) was added dropwise to a stirred solution of triphosgene (0.33 mmol) in two phase solution of dry DCM (20 mL) and saturated solution of $NaHCO_3$ (40 mL). After 5 min of stirring, the other part of phenyl alanine methyl ester hydrochloride (1.00 mmol) solution was added further into the reaction mixture and stirred for additional 4

h. The organic phase was then washed with water, dried with anhydrous Na_2SO_4, and evaporated to give crude product, which was purified by column chromatography.

1.3.3. General Synthetic Procedure of Peptides 2 and 3

To an ice-cooled solution of dipeptide (550 mg, 1.51 mmol) in ~20 mL dry DCM was added TFA (8 mL) and kept stirred for 3 h. The reaction was monitored by TLC and after completion, the reaction mixture was evaporated under a high vacuum with a KOH trap to afford N-deprotected derivative.

An ice-cold solution of N-deprotected derivative (398.9 mg, 1.51 mmol, 1.0 equiv.) in dry DCM was prepared and was divided into two equal parts. One part (0.76 mmol) was added dropwise to a stirred solution of triphosgene (0.25 mmol) in two phase solution of dry DCM (20 mL) and saturated solution of $NaHCO_3$ (40 mL). After 5 min of stirring, the other part of N-deprotected derivative (0.76 mmol) solution was added into the reaction mixture and stirred for additional 4 h. The organic phase was then washed with water, dried with anhydrous Na_2SO_4, and evaporated to give crude product, which was purified by column chromatography.

1.3.4 Analytical data of dipeptides A2, A3 and Urea-based peptides 1-3.

1.3.4.1 Dipeptide A2

Yield: 652 mg, 89%

MP: 106-108 °C.

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.42 (s, 9H), 1.43 (s, 3H), 1.45 (s, 3H), 3.00 (dd, J = 14.0, 7.5 Hz, 1H), 3.08 (dd, J = 13.0, 5.5 Hz, 1H), 3.71(s, 3H), 4.30 (m, 1H), 5.14 (br s, 1H), 6.36 (s, 1H), 7.19 – 7.32 (m, 5H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 24.5, 24.7, 28.3, 38.5, 52.6, 56.4, 80.1, 126.9, 128.7, 129.5, 136.8, 155.4, 170.3, 174.6.

IR (KBr): 3302, 2983, 1745, 1660, 1539, 1451, 1389, 1279, 1156, 1023 cm^{-1}.

HRMS: Calcd. for $C_{19}H_{28}N_2O_5Na$ m/z = 387.1896, found m/z = 387.1897.

1.3.4.2 Dipeptide A3

Yield: 659 mg, 90%

MP: 142-144 °C.

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.42 (s, 12H), 1.45 (s, 3H), 3.12 (m, 2H), 3.70 (s, 3H), 4.85 (m, 1H), 6.86 (s, 1H), 7.11 – 7.29 (m, 6H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 25.4, 28.3, 38.1, 52.2, 53.3, 56.7, 77.2, 127.0, 128.5, 129.3, 136.0, 154.5, 172.0, 174.2.

IR (KBr): 3329, 2982, 1754, 1688, 1658, 1522, 1453, 1367, 1284, 1167, 1087 cm^{-1}.

HRMS: Calcd. for $C_{19}H_{28}N_2O_5Na$ m/z = 387.1896, found m/z = 387.1888.

1.3.4.5 Peptide 1

Yield: 342 mg, 89%

MP: 152-154 °C

$[\alpha]^{26}_D$: +22° (c 0.5 mg/mL, CH_3OH)

^{1}H NMR (500 MHz, $CDCl_3$) δ 3.06 (m, 4H), 3.69 (s, 6H), 4.76 (dt, J = 8.0 Hz, 5.5 Hz, 2H), 4.94 (d, J = 8 Hz, 2H), 7.09 (m, 4H), 7.24 (m, 6H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 38.6, 52.3, 54.0, 127.0, 128.6, 129.4, 136.1, 156.2, 173.1.

IR (KBr): 3354, 3029, 2950, 1742, 1715, 1634, 1574, 1499, 1442, 1360, 1302, 1241, 1209, 1178, 1115, 1080, 1028 cm^{-1}.

HRMS: Calcd. for $C_{21}H_{24}N_2O_5Na$ m/z = 407.1583, found m/z = 407.1580.

1.3.4.6 Peptide 2

Yield: 363 mg, 87%

MP: 146-148 °C

$[\alpha]^{26}_D$: -8° (c 0.5 mg/mL, CH_3OH)

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.40 (m, 12H), 2.97 (m, 2H), 3.07 (m, 2H), 3.66 (s, 6H), 4.47 (m, 2H), 5.47 (d, J = 6.5 Hz, 2H), 6.47 (s, 2H), 7.18-7.26 (m, 10H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 24.7, 24.8, 38.7, 52.6, 55.4, 56.3, 126.9, 128.6, 129.6, 137.1, 157.0, 171.0, 174.7.

IR (KBr): 3320, 2948, 1743, 1629, 1548, 1455, 1387, 1284, 1230, 1193, 1152 cm^{-1}.

HRMS: Calcd. for $C_{29}H_{39}N_4O_7$ m/z = 555.2819, found m/z = 555.2814; for $C_{29}H_{38}N_4O_7Na$ m/z = 577.2638, found m/z = 577.2632, for $C_{29}H_{38}N_4O_7K$ m/z = 593.2378, found m/z = 593.2367.

1.3.4.7 Peptide 3

Yield: 380 mg, 91%

MP: 160-162 °C

$[\alpha]^{26}_D$: +2° (c 0.5 mg/mL, CH_3OH)

^{1}H NMR (500 MHz, $CDCl_3$) δ 1.37 (s, 6H), 1.44 (s, 6H), 3.05 (dd, J = 14 Hz, 7 Hz, 2H), 3.14 (dd, J = 14 Hz, 6 Hz, 2H), 3.68 (s, 6H), 4.79 (m, 2H), 5.06 (s, 2H), 7.13-7.28 (m, 12H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 25.2, 26.3, 37.8, 52.3, 53.6, 56.9, 127.0, 128.5, 129.4, 136.3, 156.5, 172.4, 175.2.

IR (KBr): 3420, 1739, 1656, 1536, 1215 cm^{-1}.

HRMS: Calcd. for $C_{29}H_{38}N_4O_7Na$ m/z = 577.2638, found m/z = 577.2632.

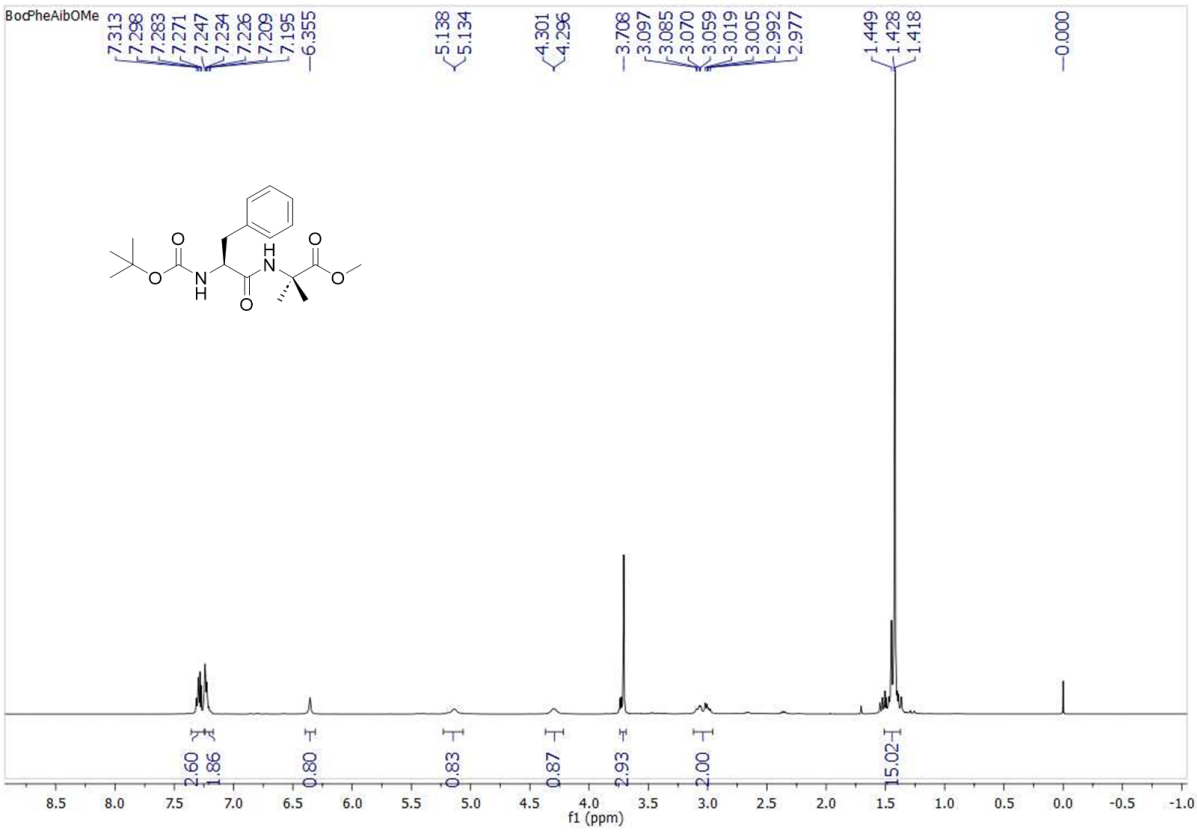


Figure S1. ^{1}H NMR (500 MHz, $CDCl_3$) of **A2**

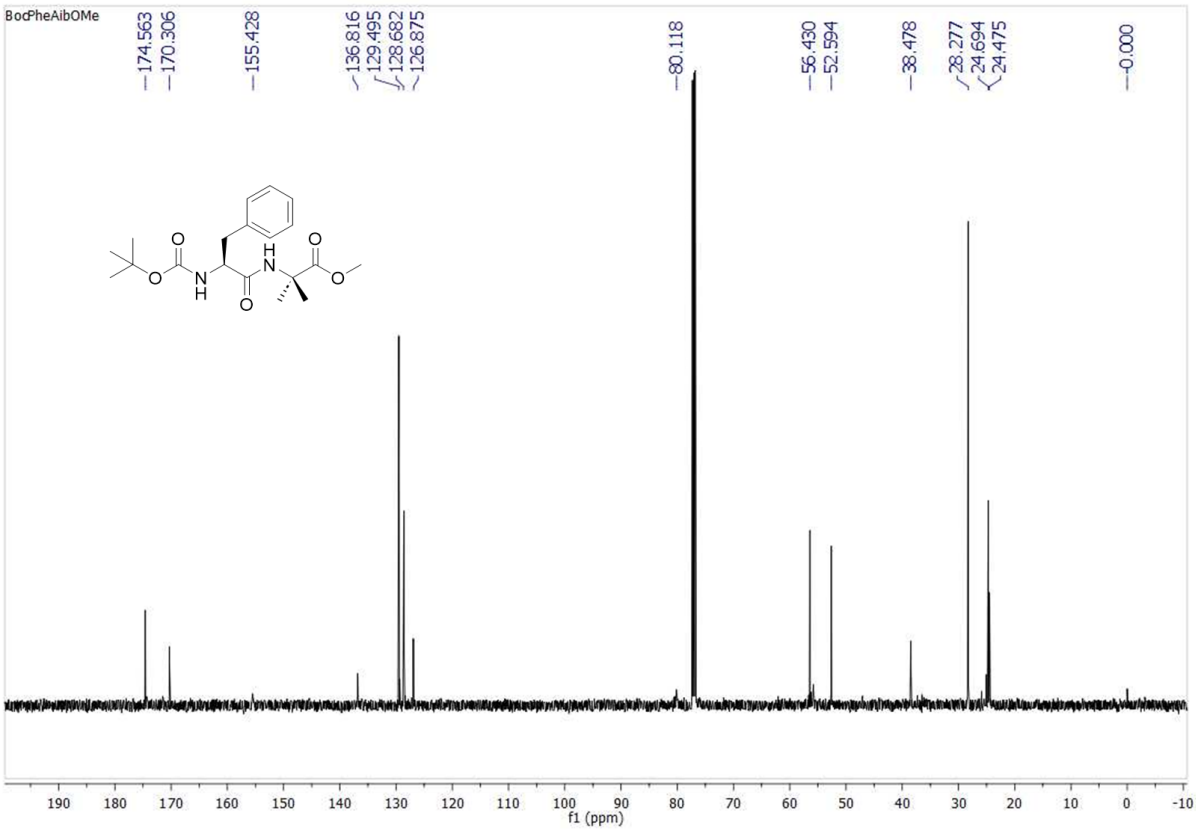


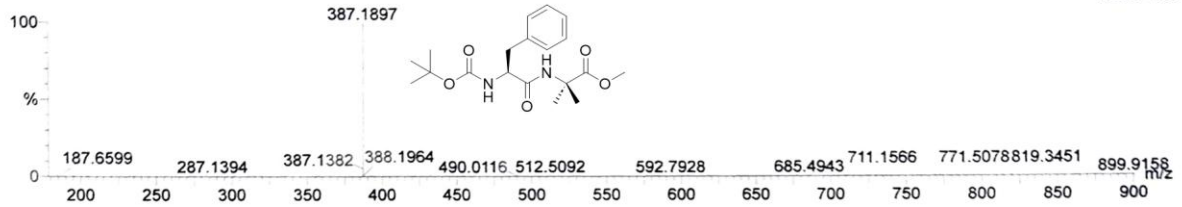
Figure S2. ^{13}C NMR (125 MHz, $CDCl_3$) of **A2**

Elemental Composition Report **Page 1**

Single Mass Analysis
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

Monoisotopic Mass, Even Electron Ions
376 formula(e) evaluated with 2 results within limits (up to 10 closest results for each mass)
Elements Used:
C: 0-45 H: 0-50 N: 0-4 O: 0-8 Na: 0-1

XEVO -G2XSQTOF#TFC2176 Capillary V 3, Cone V 40 , Desolvation Gas 800 17-Jan-2024
POSITIVE ION MODE ESI
VH 5
17012024_13 25 (0.519) 1: TOF MS ES+
1.29e+005



Minimum:				-1.5				
Maximum:		5.0	5.0	50.0				
Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	Norm	Conf(%)	Formula
387.1897	387.1896	0.1	0.3	6.5	43.2	0.444	64.16	C19 H28 N2 O5 Na
	387.1880	1.7	4.4	5.5	43.8	1.026	35.84	C16 H27 N4 O7

Figure S3. ESI-Mass spectrum of **A2**

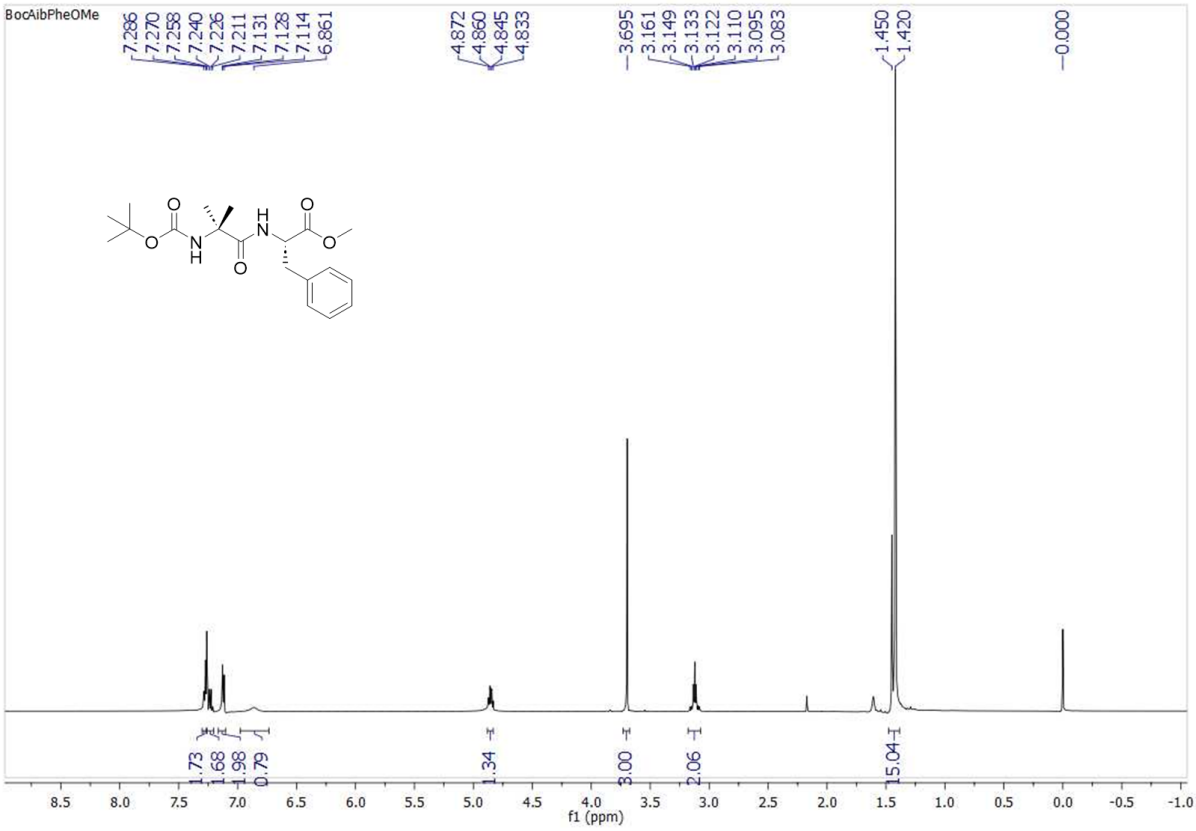


Figure S4. ^{1}H NMR (500 MHz, $CDCl_3$) of **A3**

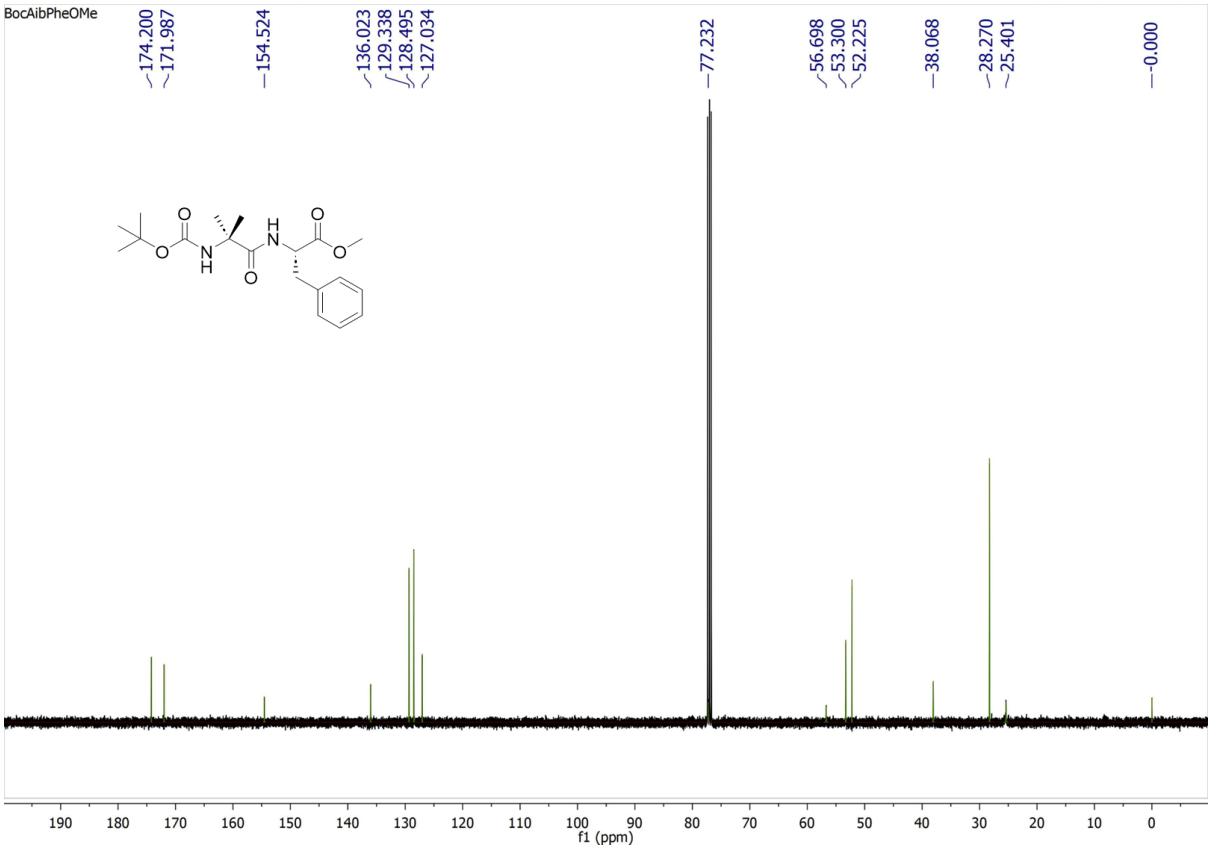


Figure S5. ^{13}C NMR (100 MHz, $CDCl_3$) of **A3**

Elemental Composition Report **Page 1**

Single Mass Analysis
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

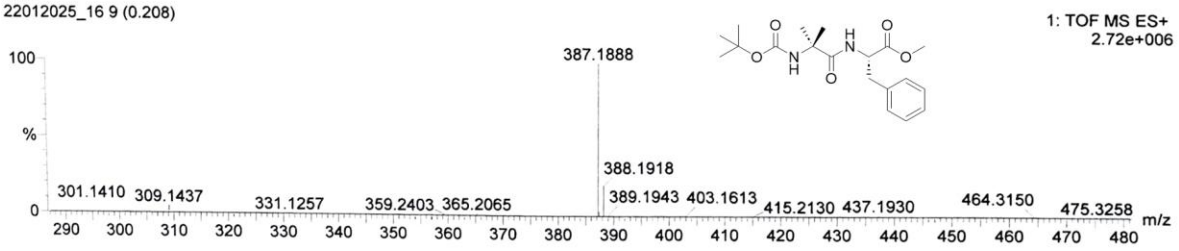
Monoisotopic Mass, Even Electron Ions
144 formula(e) evaluated with 1 results within limits (up to 10 closest results for each mass)
Elements Used:
C: 0-30 H: 0-35 N: 0-3 O: 0-5 Na: 0-1

XEVO -G2XSQTOF#TFC2176
POSITIVE ION MODE
VH 1
22012025_16 9 (0.208)

Capillary V 3, Cone V 40 , Desolvation Gas 800
ESI

22-Jan-2025

1: TOF MS ES+
2.72e+006



Minimum:				-1.5				
Maximum:		5.0	5.0	50.0				
Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	Norm	Conf(%)	Formula
387.1888	387.1896	-0.8	-2.1	6.5	432.5	n/a	n/a	C19 H28 N2 O5 Na

Figure S6. ESI-Mass spectrum of **A3**

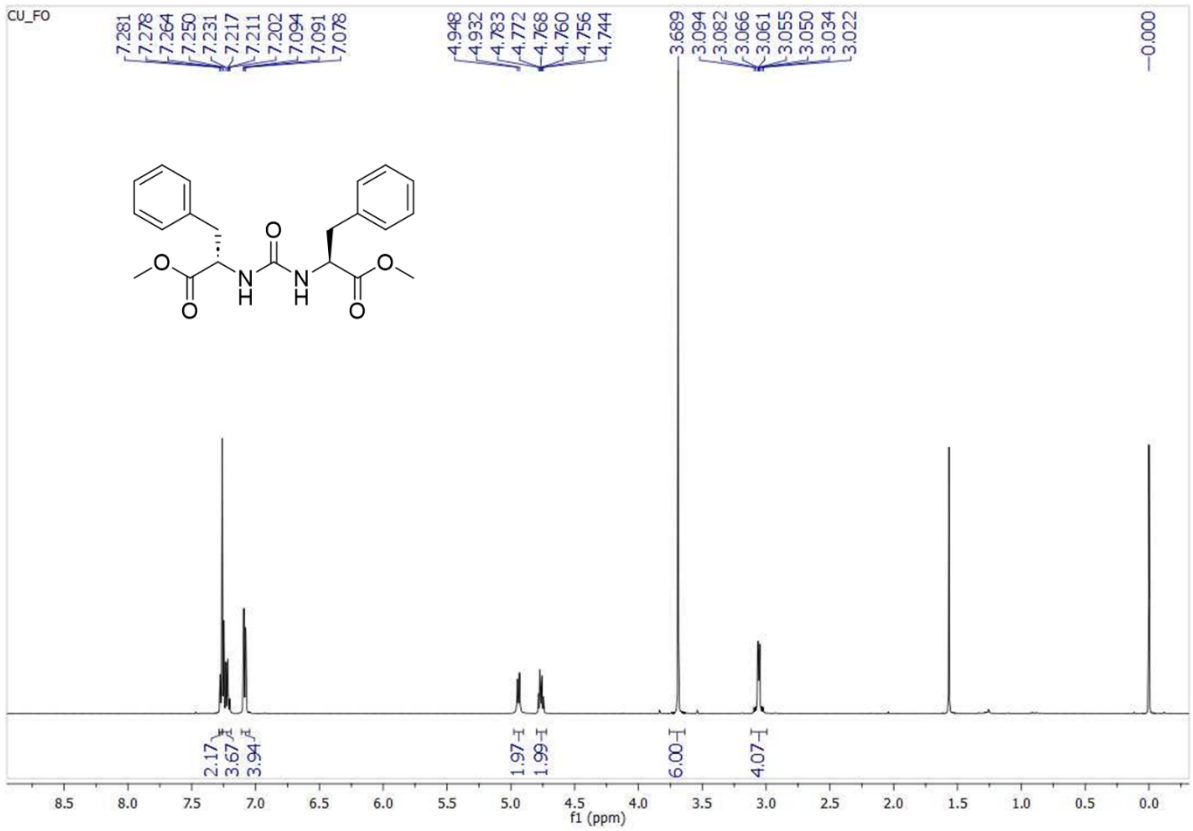


Figure S7. ^{1}H NMR (500 MHz, $CDCl_3$) of Peptide **1**

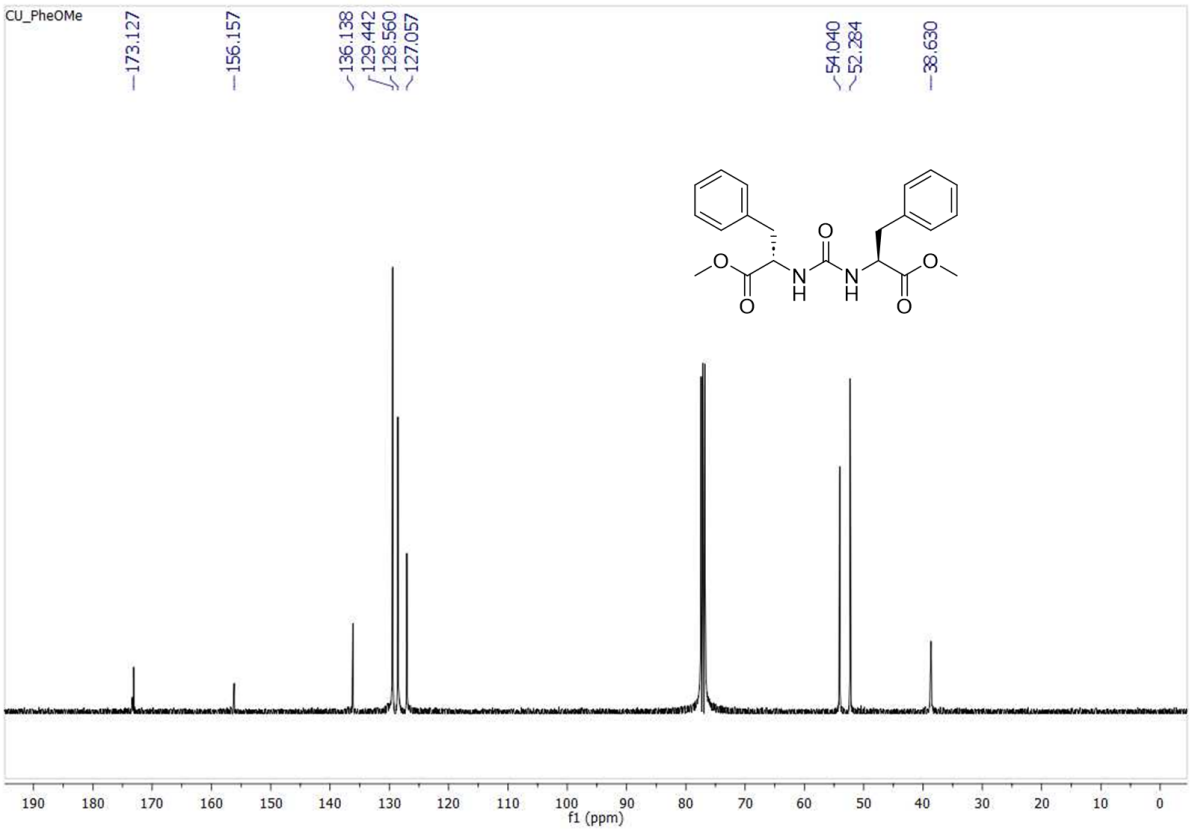


Figure S8. ^{13}C NMR (100 MHz, $CDCl_3$) of Peptide **1**

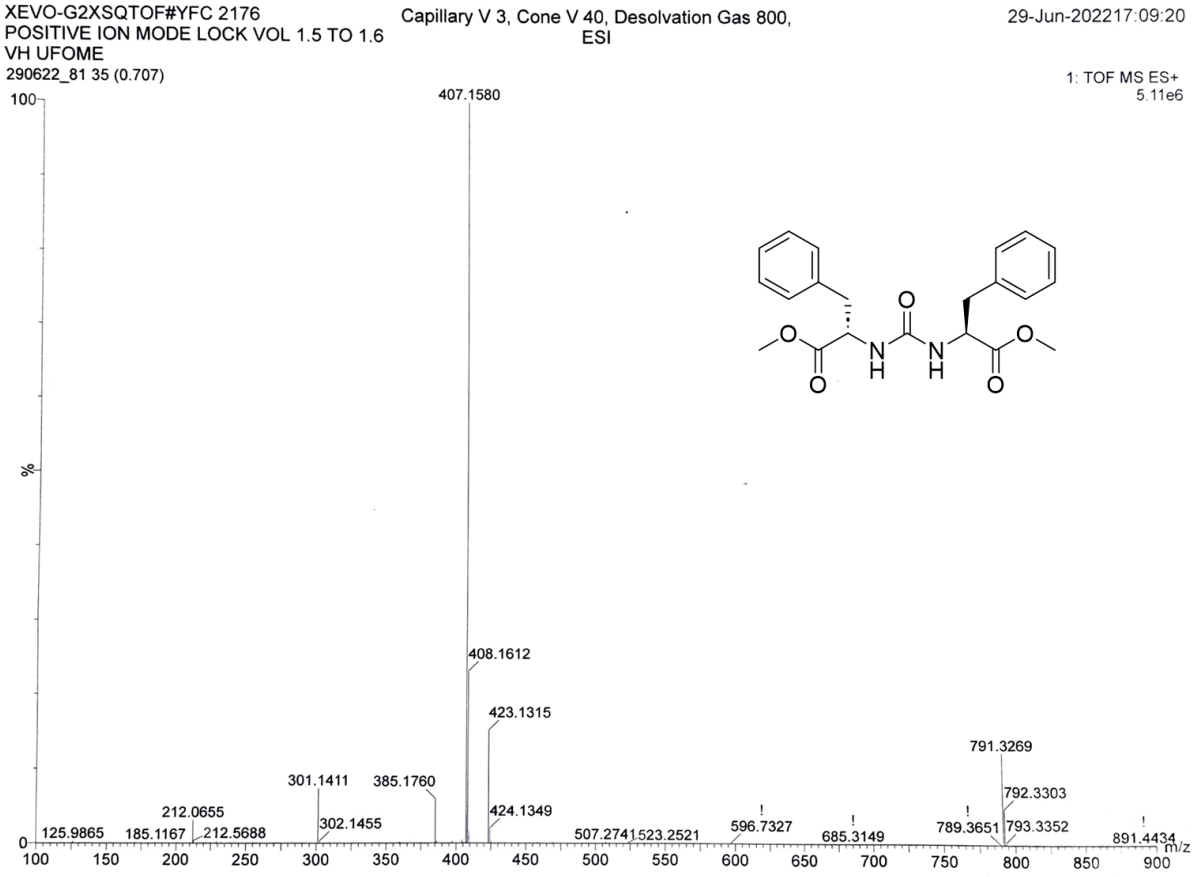


Figure S9. ESI-Mass spectrum of **1**

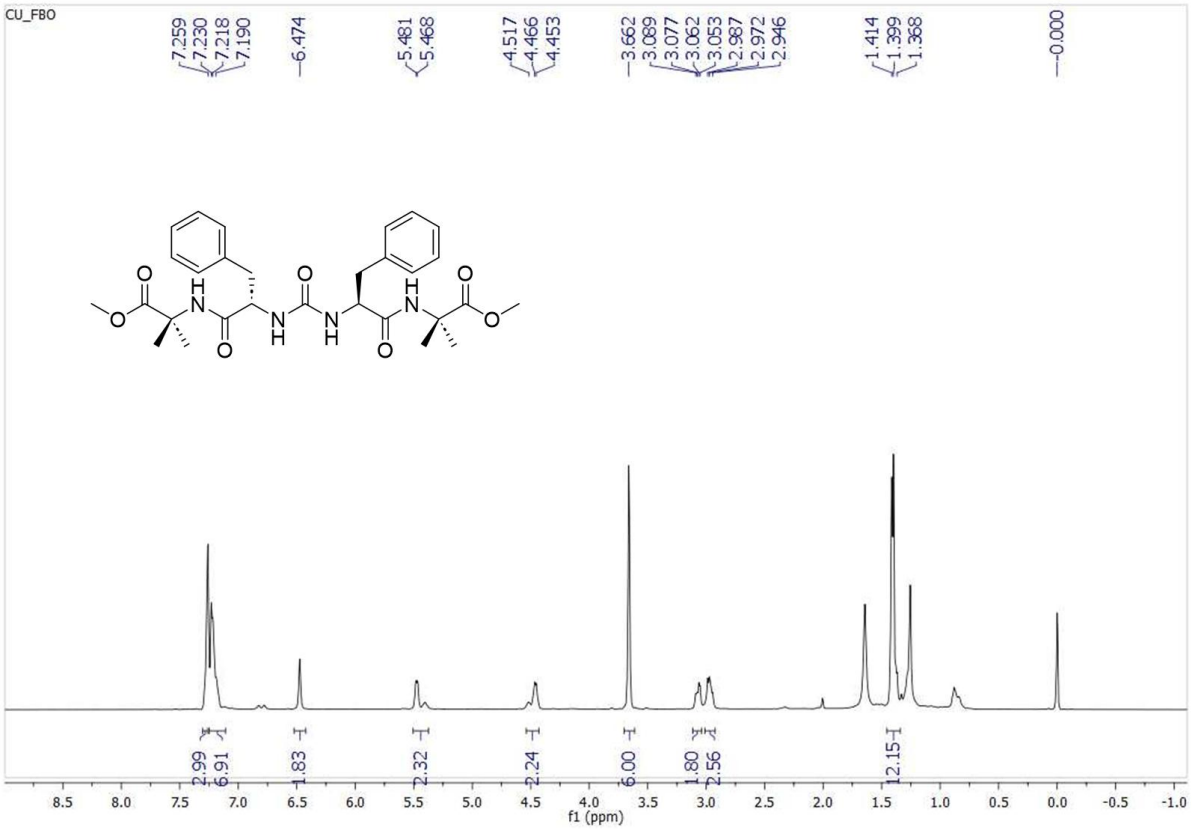


Figure S10. ^{1}H NMR (500, $CDCl_3$) of Peptide **2**

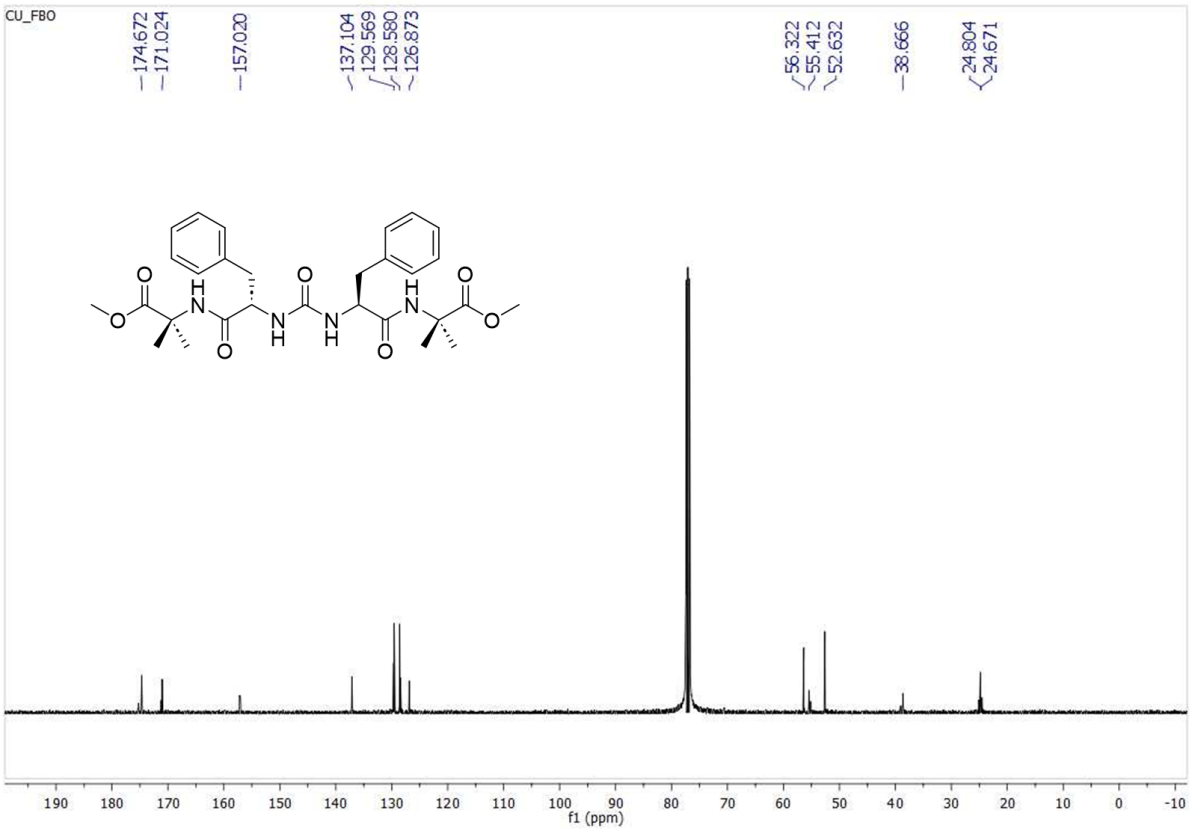


Figure S11. ^{1}H NMR (100 MHz, $CDCl_3$) of Peptide **2**

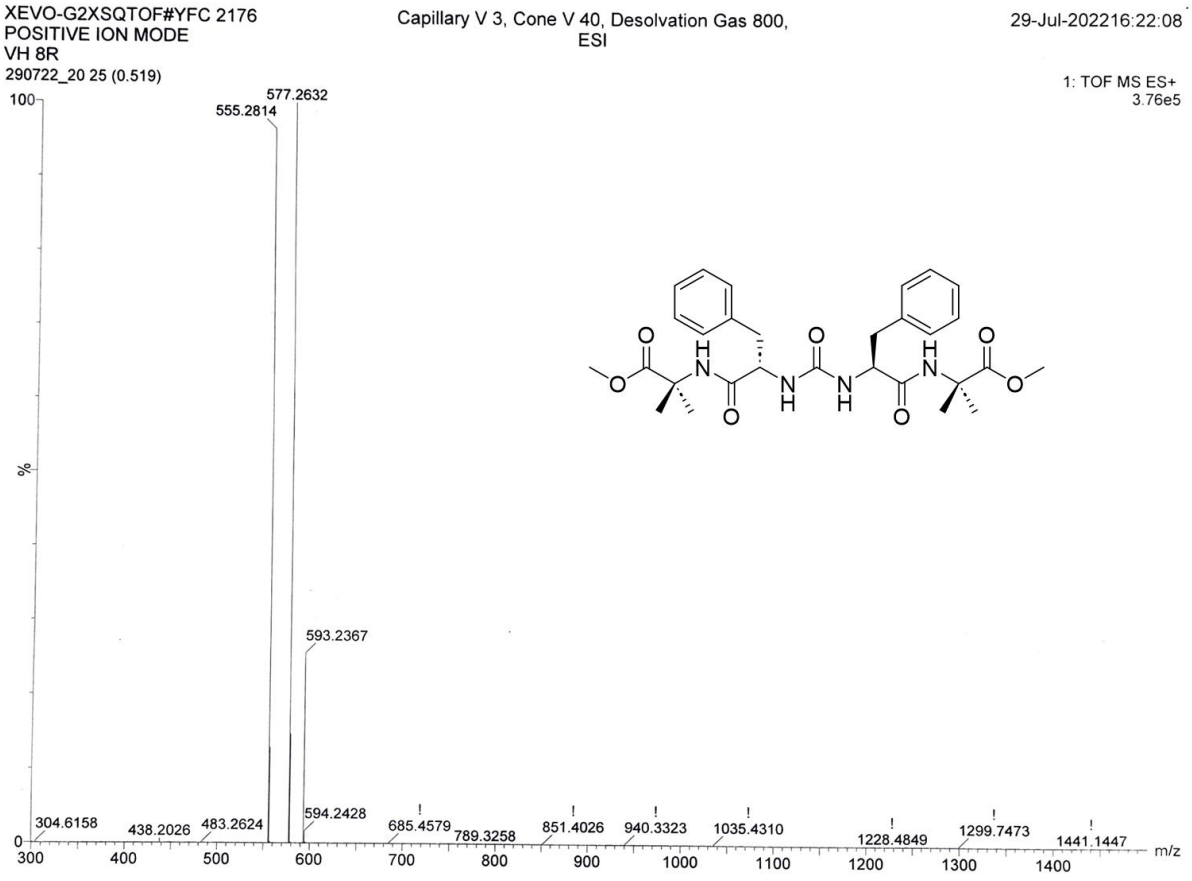


Figure S12. ESI-Mass spectrum of **2**

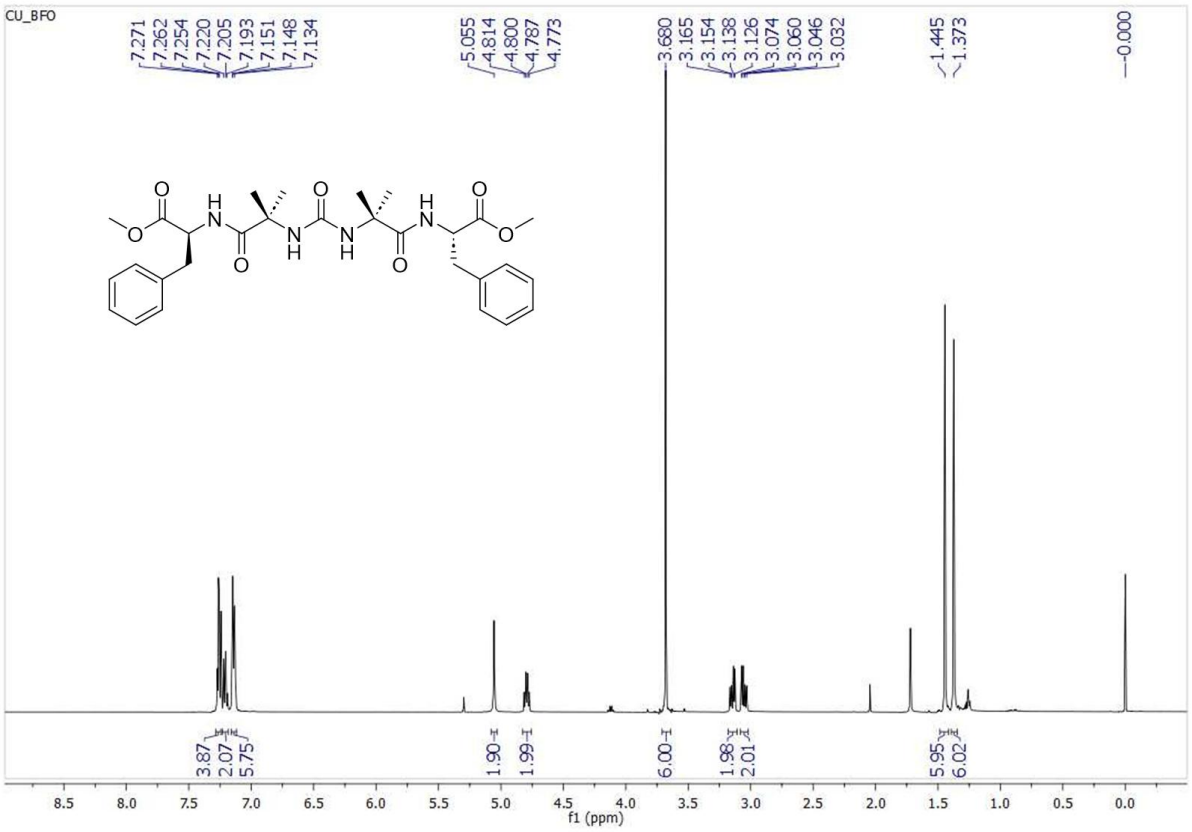


Figure S13. ^{1}H NMR (500, $CDCl_3$) of Peptide **3**

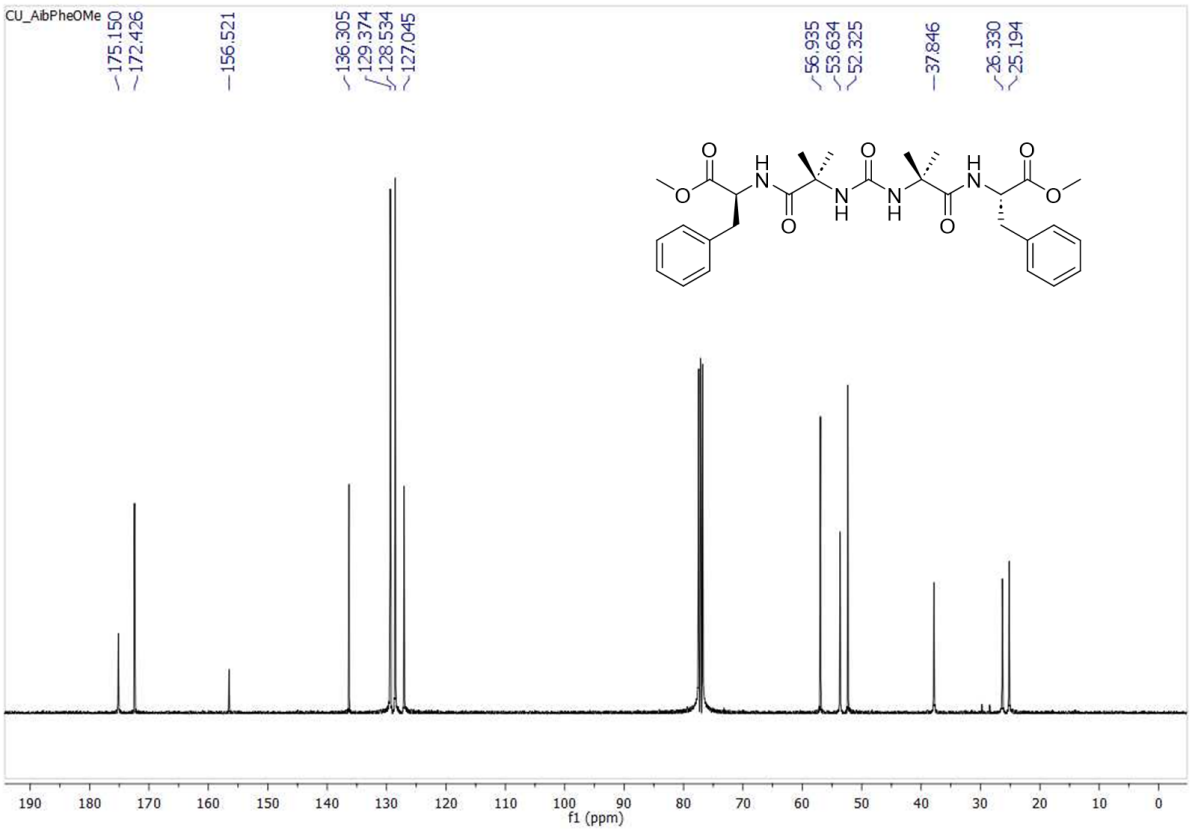


Figure S14. ^{1}H NMR (100 MHz, $CDCl_3$) of Peptide **3**

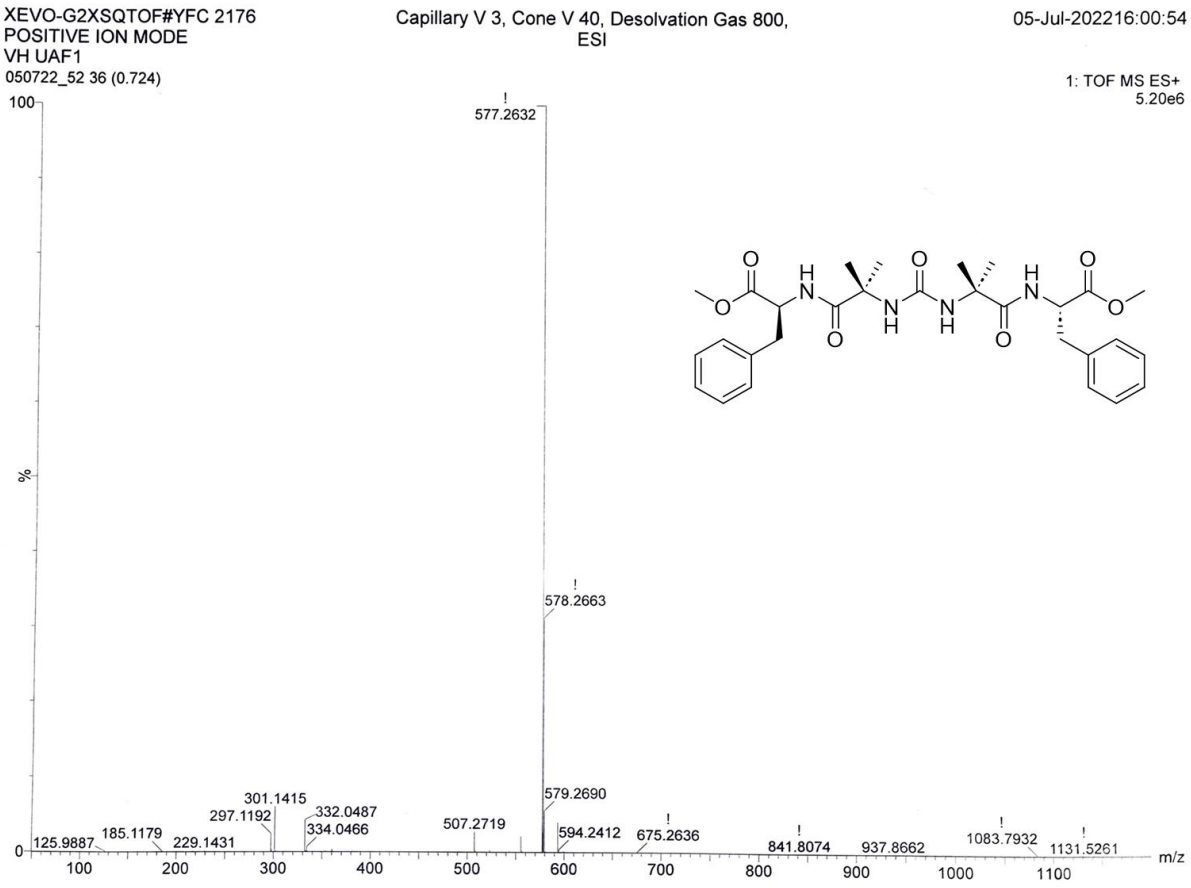


Figure S15. ESI-Mass spectrum of **3**

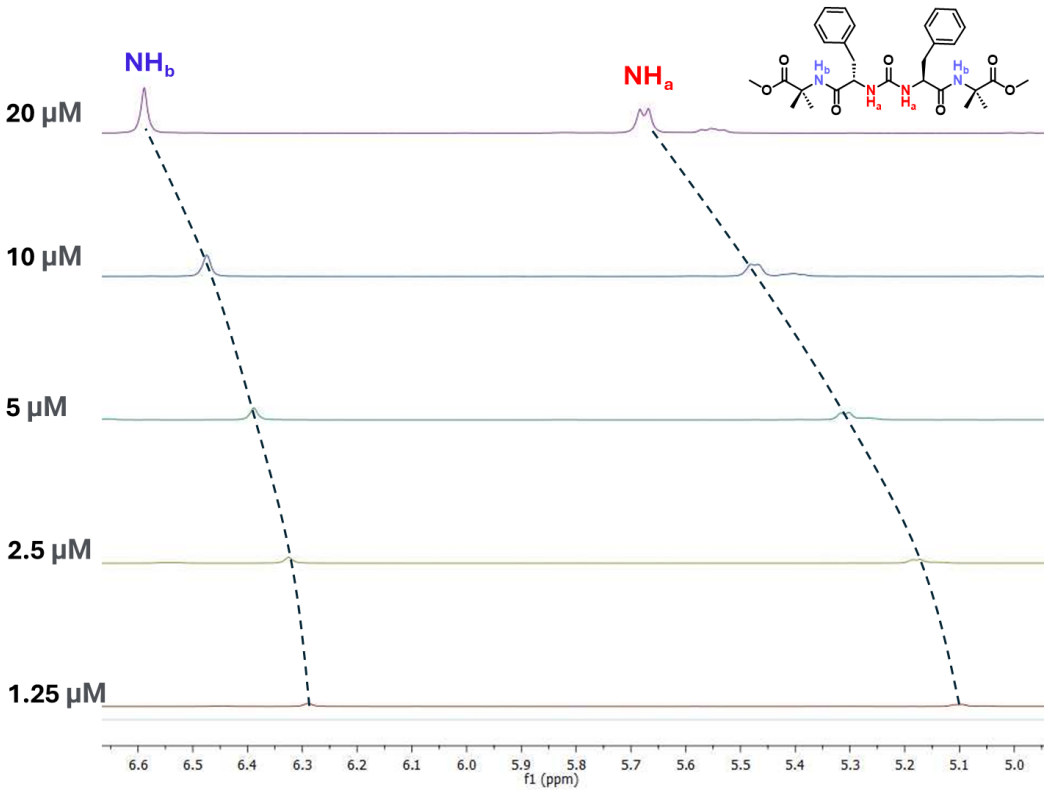


Figure S16. Partial ^{1}H NMR (500 MHz, $CDCl_3$) spectra of **2** at different concentrations.

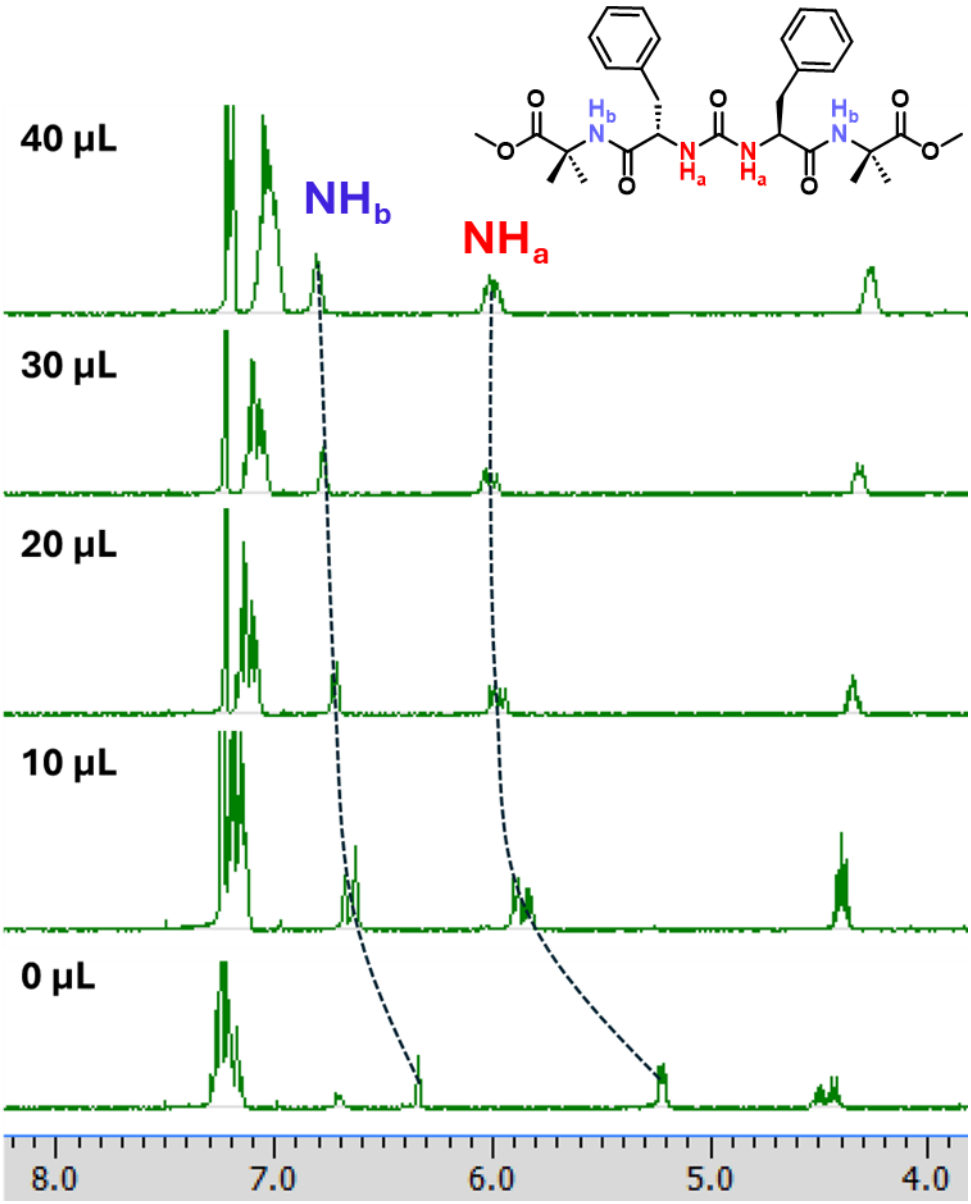


Figure S17. Partial ^{1}H NMR (400 MHz, $CDCl_3$) spectra of **2** upon the addition of various amounts of DMSO-d_6, indicating shifts of NHs.

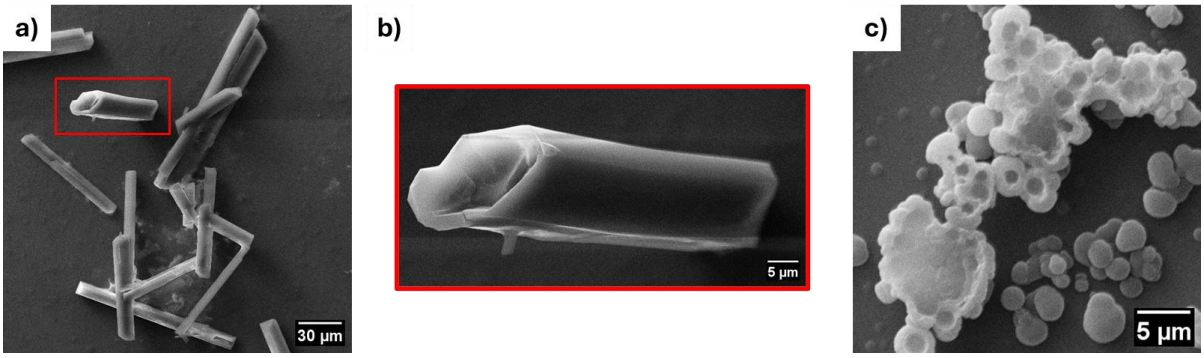


Figure S18. a) SEM image of peptide **3**, b) zoomed-in image of a single microtube showing an open end of the tube, c) An SEM image of peptide **2**, illustrating the fusion of the inner cavities of several toroids, culminating in the creation of a larger, shared cavity.

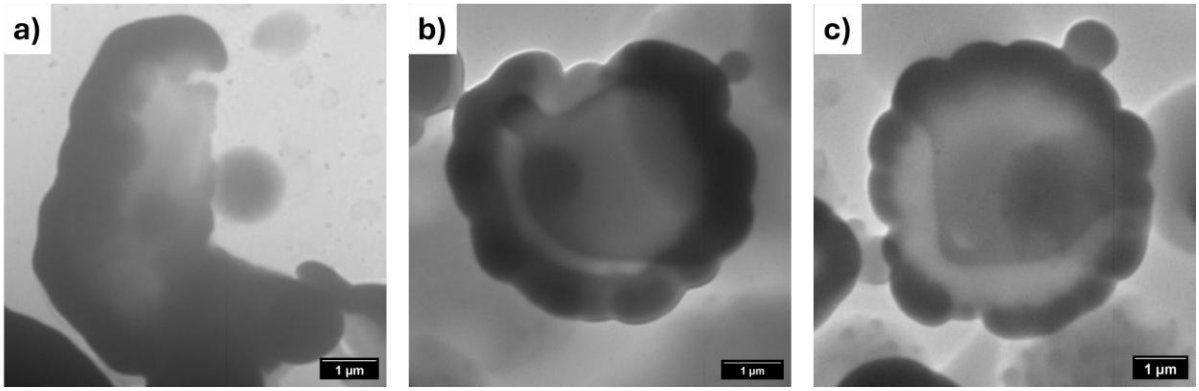


Figure S19. TEM images of peptide **2** showing intermediate stages of toroid formation.

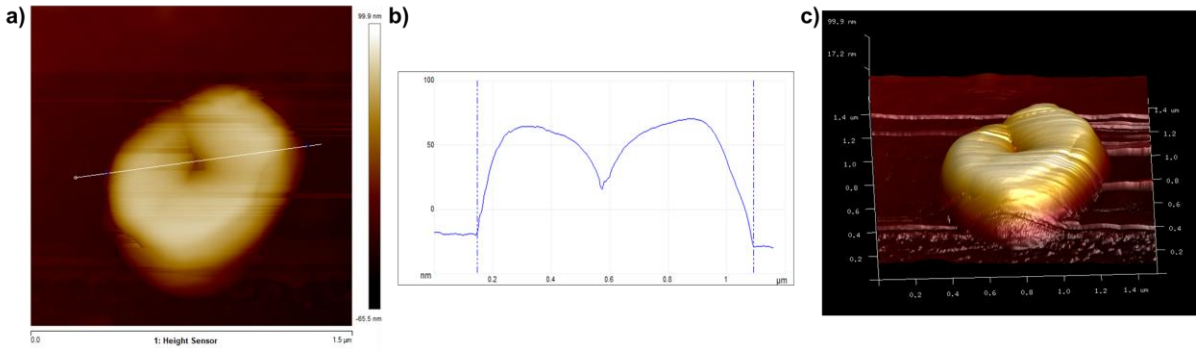


Figure S20. a) AFM image of toroid formed from self-assembly of peptide **2**; b) Height profile diagram of the toroid in Figure a; c) Three-dimensional AFM image of the toroid.

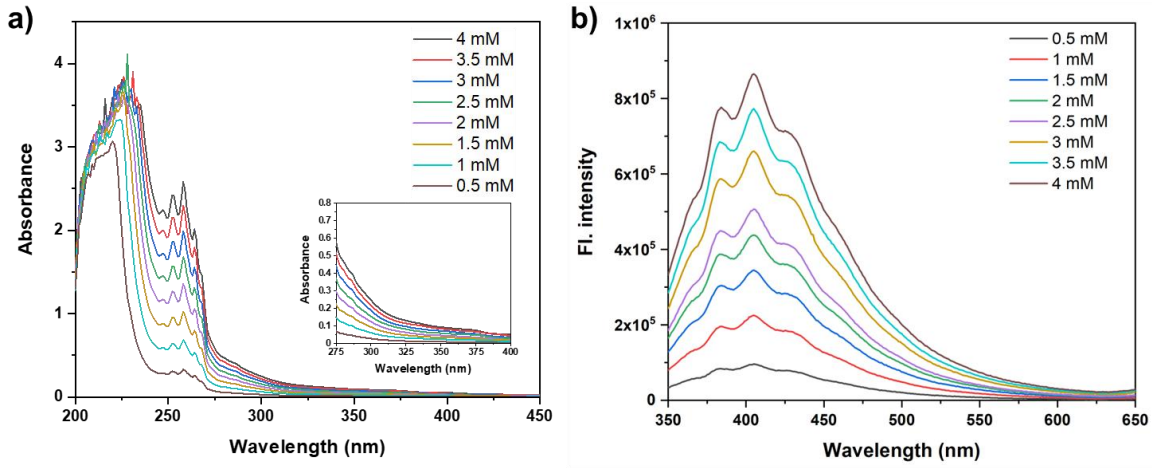


Figure S21. a) UV-visible absorption spectra of **2** at different concentrations. The inset shows partial absorption spectra indicating absorption at higher wavelengths, b) Fluorescence emission spectra for different concentrations of **2** upon excitation at 335 nm.

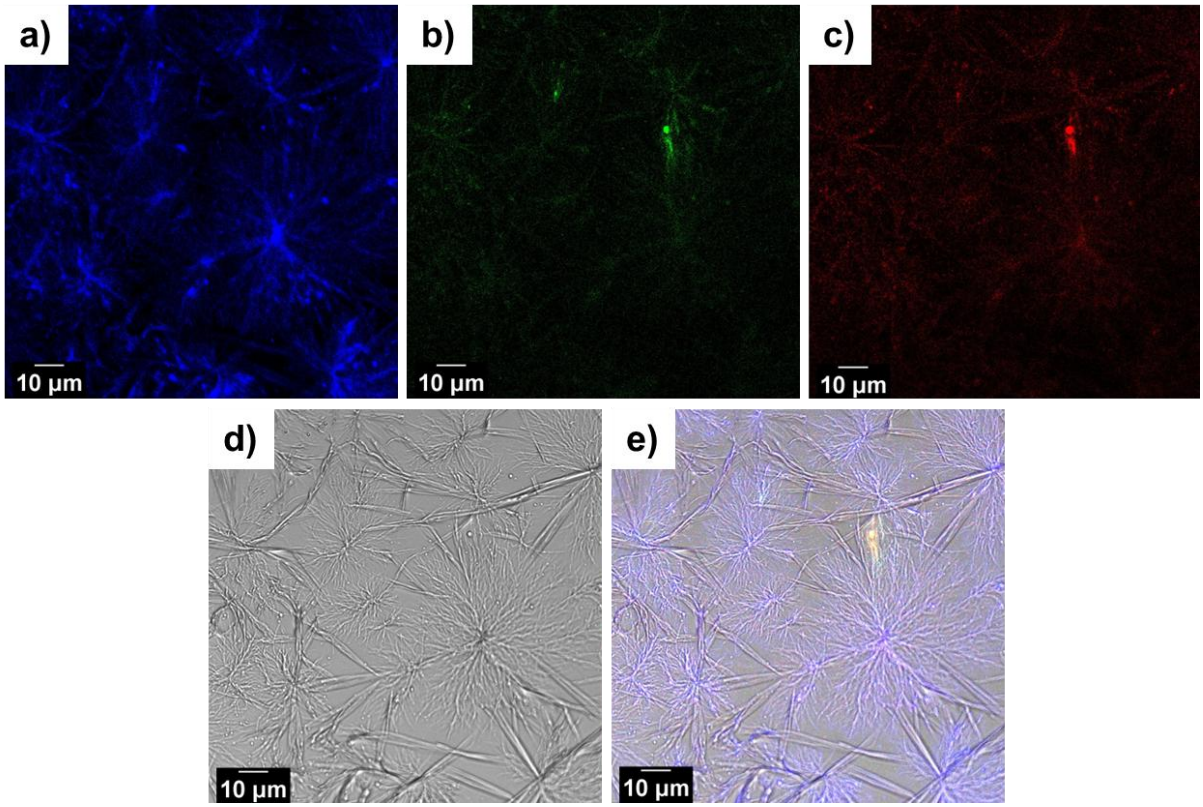


Figure S22. Confocal images of compound **1** upon excitation at a) 405 nm, b) 488 nm, and c) 560 nm; d) optical image of compound **1**; e) overlayed image of a-d.