Sequence Decoding of 1D to 2D Self-Assembling Cyclic Peptides

Sandra Díaz[+], Ignacio Insua[+], Ghibom Bhak and Javier Montenegro*

Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Orgánica. Universidade de Santiago de Compostela. Santiago de Compostela, 15782, Spain.

[+] These authors contributed equally to this work.

E-mail: javier.montenegro@usc.es

Supporting Information

1. Abbreviations

ACN: acetonitrile, AFM: atomic force microscopy, All: allyl, CP: cyclic peptide, DCM: dichloromethane, DIEA: *N,N*-diisopropylethylamine, DMF: *N,N*-dimethylformamide, ESI: electrospray ionisation, Fmoc: 9-fluorenylmethoxycarbonyl, HOBt: hydroxybenzotriazole, HR-MS: high resolution mass spectrometry, HPLC-MS: high performance liquid chromatography-mass spectrometry, HBTU: *N,N,N',N'*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate, NMR: nuclear magnetic resonance spectroscopy, OAc: acetate, Ph: phenyl, PyAOP: (7-azabenzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate, SPPS: solid phase peptide synthesis STEM: scanning transmission electron microscopy, TFA: trifluoroacetic acid, ThT: thioflavin T, TIS: triisopropylsilane.

2. Materials and methods

Chemical reagents were acquired from Acros Organics, Aldrich, Fisher Scientific, Iris Biotech and Novabiochem. Dichloromethane was dried under reflux over calcium hydride. D_2O was purchased from EMD Millipore. Glass slides for fluorescence microscopy were obtained from Ibidi (Cat# 80827). TEM grids (Cu carbon type-B, 300 mesh) and PELCO® mica discs for AFM were acquired from Ted Pella. ^{1}H NMR spectra were recorded on a Varian 300 MHz spectrometer. Chemical shifts are reported in ppm (δ) referenced to D_2O's residual signal (δ = 4.79). HPLC-MS analyses were carried out on an Agilent 1260 Infinity II fitted with an Agilent SB-C18 column and connected to a 6120 Quadrupole MS detector. HR-MS was acquired on a Bruker MicroTOF. Epifluorescence micrographs were taken with a Nikon Eclipse T*i* (60x immersion objective, Excitation=475/35 nm; Emission=530/43 nm). STEM images were acquired on a FESEM Ultra plus (Zeiss) operating at 20 kV from unstained samples. AFM analysis was carried out on a NX-10 microscope in non-contact mode and using ACTA 10M cantilevers (Park Systems).

3. Cyclic peptide synthesis and purification

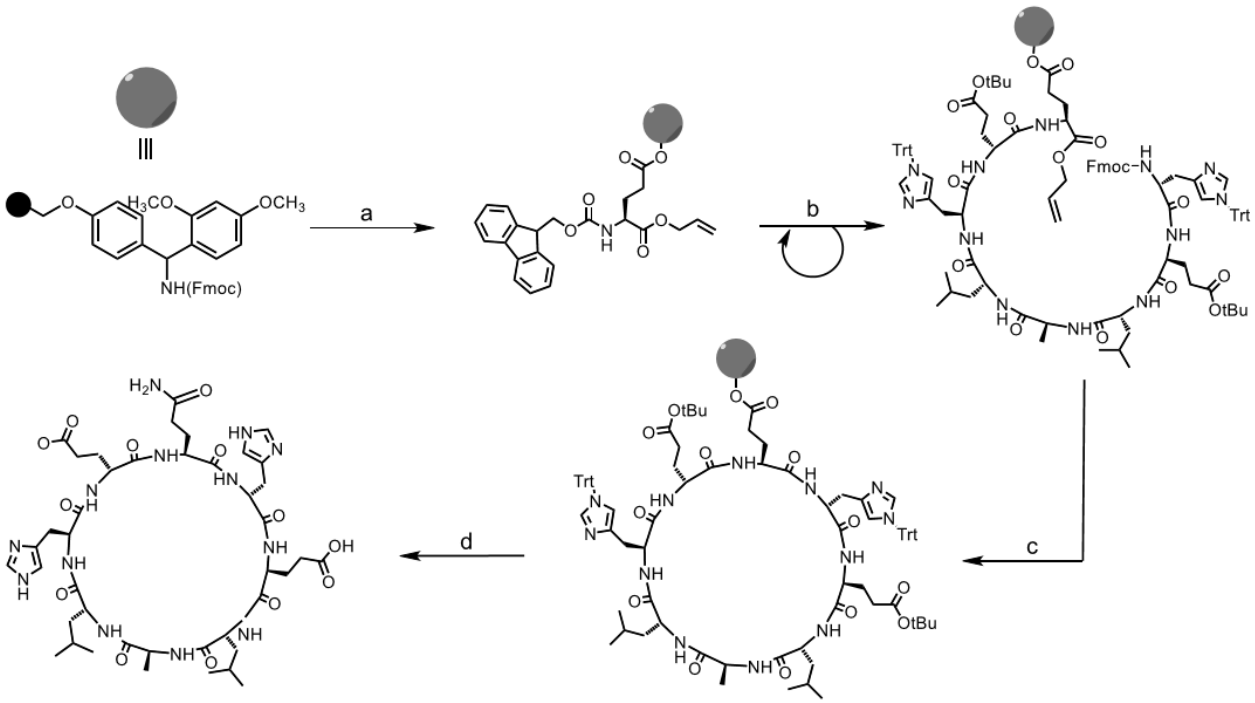
Cyclic peptides were prepared manually in solid phase. Rink Amide resin was swollen in DMF for 30 min and then treated with piperidine 20% v/v in DMF (3 mL) to remove the Fmoc protecting group. The resin was washed with DMF (3 x 3 mL) and a solution of Fmoc-L-Glu(OH)-OAll (163.6 mg, 0.1 mmol), HBTU (132.5 mg, 0.35 mmol) and DIEA (95 µL, 0.6 mmol) in DMF (3 mL) was added over the resin and shaken for 1 h. Then the mixture was filtered and washed with DMF (3 x 3mL) and DCM (3 x 3 mL). We followed cycles of Fmoc removal (2 mL of piperidine 20% v/v in DMF for 15 min) and amino acid coupling (0.4 mmol of amino acid, 0.35 mmol HBTU and 0.6 mmol DIEA in DMF for 30 min) with their respective washes (DMF 3 x 3 mL).

For OAll removal, the resin was washed with DMF and DCM (3 x 3 mL) and then reacted with a solution of PPh_3 (39.3 mg, 0.15 mmol), *N*-methylmorpholine (110 µL, 0.01 mmol) and phenylsilane (123 µL, 1.0 mmol) and $Pd(OAc)_2$ (6.7 mg, 0.03 mmol) in dry DCM (4 mL) overnight (16-20 h). The resin was then washed with DCM (3 x 3 mL) and DIEA in DMF (2% v/v, 3 x 3 mL), and soaked in a solution of sodium diethyldithiocarbamate (0.5% w/v in DMF, 2 x 3 mL) for 30 min to remove all traces of Pd. Finally, the resin was stirred with piperidine/DMF (20% v/v, 3 mL) for 30 min for *N*-terminal Fmoc removal.

Cyclisation was carried out by reacting a solution of PyAOP (208.6 mg, 0.4 mmol) and DIEA (95.1 mL, 0.6 mmol) in DMF (3 mL) for 2 h. After washing with DMF (3 x 3 mL) and DCM (3 x 3 mL), the cyclisation was repeated twice more in the same conditions.

The peptide was cleaved from the resin by addition of a freshly prepared TFA cocktail (4 mL, TFA:DCM:H_2O:TIS, 0.9:0.05:0.025:0.025), this mixture was shaken for 2 h and then filtered. The resin was washed with TFA (0.5 mL) twice and concentrated under nitrogen. The concentrated reaction crude was precipitated dropwise into 40 mL of cold diethyl ether under stirring. The resulting suspension was centrifuged and the pellet was then dissolved in a 1:0.25 mixture of MilliQ H_2O:ACN and purified by semipreparative HPLC using a Phenomenex Luna C18 100Å column [gradient of A (H_2O+0.1% v/v TFA) : B (ACN+0.1% v/v TFA)→95:5 (0 min) to 25:75 (30 min)]. Peptide fractions were concentrated *in vacuo* to remove ACN and TFA and the remaining solution was freeze dried.

A white powder was obtained for all cyclic peptides: **CP_A** (35 mg, 36 %); **CP_L** (40 mg, 40 %); **CP_F** (41 mg, 39 %); **CP_{2F}** (44 mg, 38 %); **CP_{3F}** (39 mg, 35 %); **CP_{oE}** (41 mg, 38 %); **CP_{2E2H}** (38 mg, 35 %); **CP_K** (22 mg, 21 %); **CP_{4E}** (24 mg, 22 %).



Scheme S1. a) Rink Amide resin functionalisation: i. Piperidine 20% v/v in DMF, 30 min; ii. Fmoc-L-Glu-OAll, HBTU, DIEA, DMF, 1 h. **b)** Peptide elongation: i. Piperidine 20% v/v in DMF, 15 min; ii. Fmoc-amino acid, HBTU, DIEA, DMF, 30 min. (Repeat 7 cycles). **c)** Cyclisation: i. OAllyl removal: $Pd(OAc)_2$, PPh_3, phenylsilane, 4-methylmorpholine, DCM, overnight (16-20 h); ii. Fmoc removal: piperidine 20% v/v in DMF, 30 min; iii. Peptide cyclisation: PyAOP, DIEA, DMF, 2 h (repeat twice). **d)** Peptide cleavage: TFA, DCM, H_2O, TIS, 2 h.

4. Cyclic peptide characterisation

4.1. CP_A

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 9.5 min; m/z = 958.5 ($[M+H]^+$), 479.8 ($[M+2H]^{2+}$).

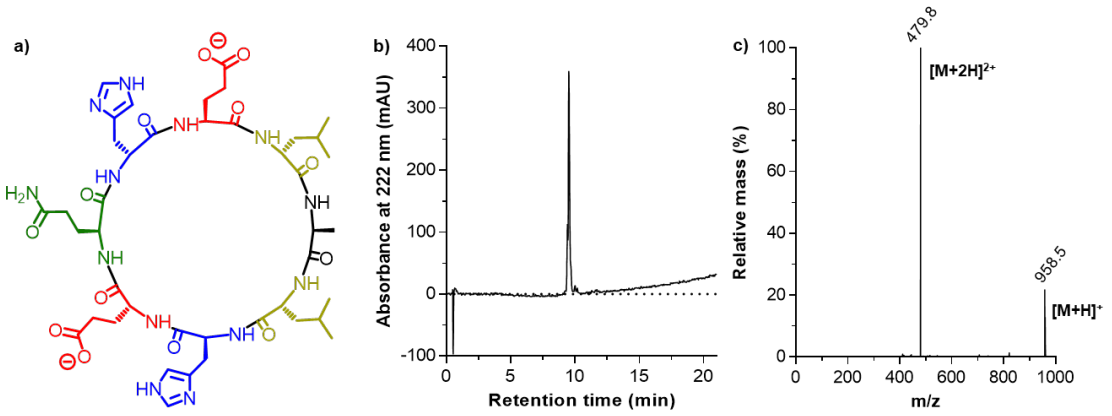


Figure S1. a) Structure of **CP_A**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV) m/z = 958.4741 (calculated for $[M+H]^+$); 958.4744 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.78-0.95 (m, 17H, Leu *i*-Bu x2, Ala-CH_3), 1.31-1.38 (m, 4H, Leu-CH_2), 1.49-1.69 (m, 4H, Glu-CH_2- x2), 1.82-2.13 (m, 4H, Gln x2), 2.27-2.34 (m, 4H, Glu-CH_2- x2), 3.01-3.43 (m, 4H, His-CH_2- x2), 4.17-4.47 (m, 8H, H_α), 7.22-7.34 (m, 2H, His-CH= x2), 8.66 (s, 2H, His-CH= x2) ppm.

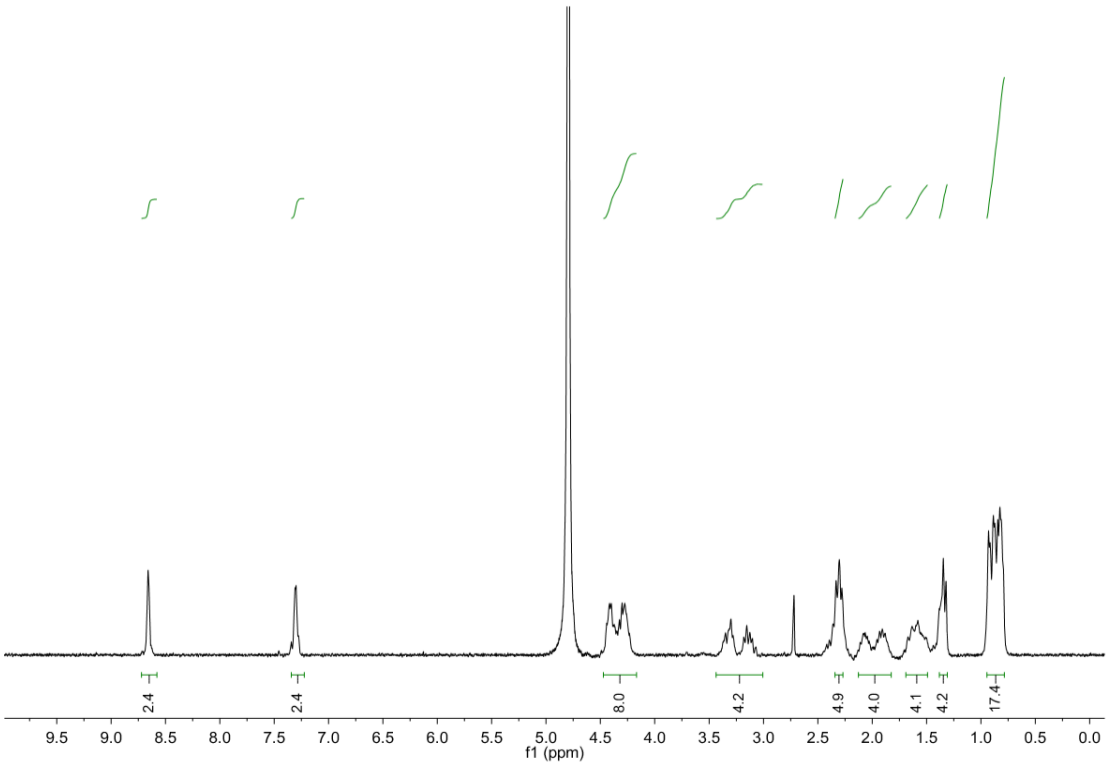


Figure S2. ^{1}H NMR (300 MHz) of **CP_A** in D_2O.

4.2. CP_L

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t =11.1 min; m/z = 1000.5 ($[M+H]^+$), 500.8 ($[M+2H]^{2+}$).

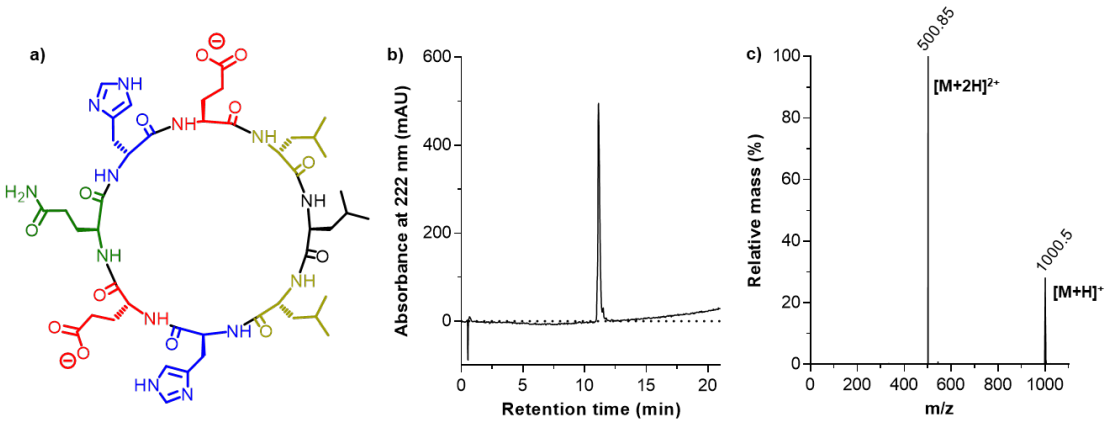


Figure S3. a) Structure of **CP_L**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1000.5211 (calculated for $[M+H]^+$); 1000.5206 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.78-0.95 (m, 27H, Leu x3), 1.49-2.37 (m, 12H, Glu x2, Gln), 3.01-3.44 (m, 4H, His-CH_2- x2), 4.18-4.76 (m, 8H, H_α), 7.28-7.31 (m, 2H, His-CH= x2), 8.66 (s, 2H, His-CH= x2) ppm.

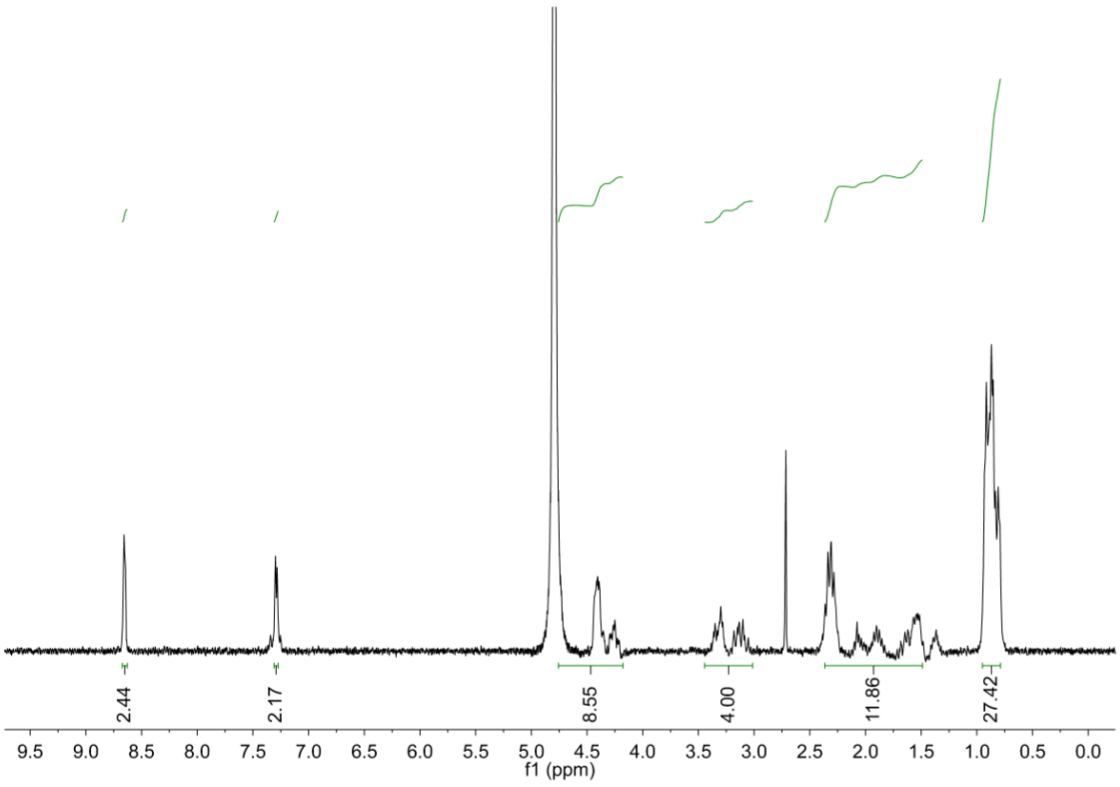


Figure S4. ^{1}H NMR (300 MHz) of **CP_L** in D_2O.

4.3. CP_F

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 11.4 min; m/z = 1034.5 ($[M+H]^+$), 517.8 ($[M+2H]^{2+}$).

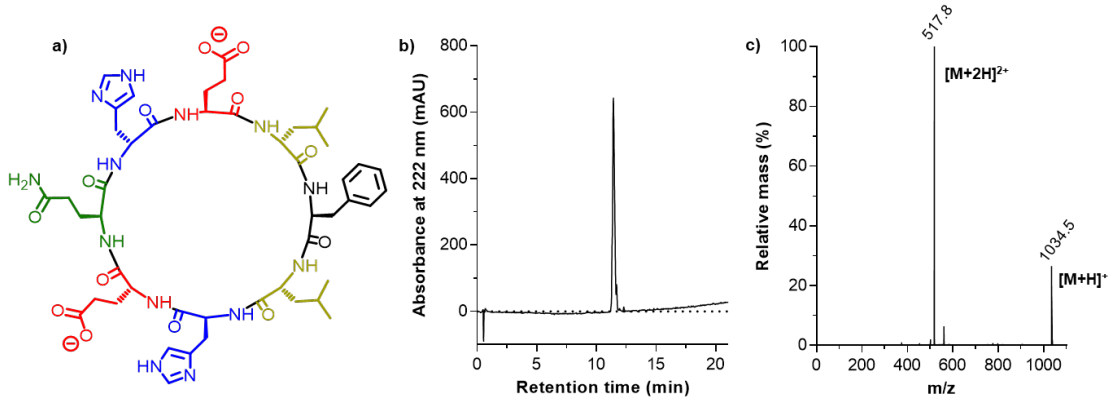


Figure S5. a) Structure of **CP_F**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1034.5054 (calculated for $[M+H]^+$); 1034.5048 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.67-0.90 (m, 16H, Leu *i*-Bu x2, Leu-CH_2- x2), 1.22-1.38 (m, 2H, Leu-CH- x2), 1.69-2.41 (m, 12H, Glu x2, Gln), 2.88-3.25 (m, 4H, His-CH_2- x2), 3.25-3.42 (m, 2H, Phe-CH_2-), 4.01-4.68 (m, 8H, H_α), 7.19-7.35 (m, 7H, His-CH= x2, Phe-CH= x2), 8.63 (s, 2H, His-CH= x2) ppm.

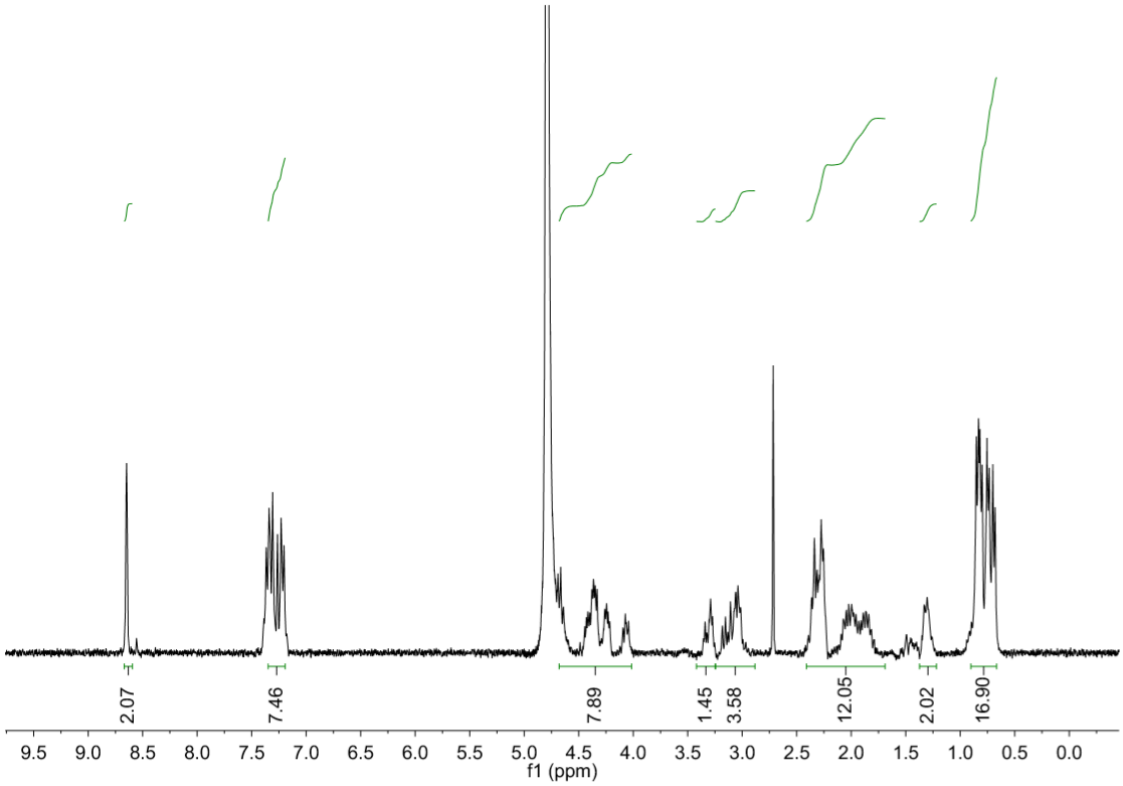


Figure S6. ^{1}H NMR (300 MHz) of **CP_F** in D_2O.

4.4. CP_{2F}

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 11.9 min; m/z = 1141.5 ($[M+H]^+$), 571.3 ($[M+2H]^{2+}$).

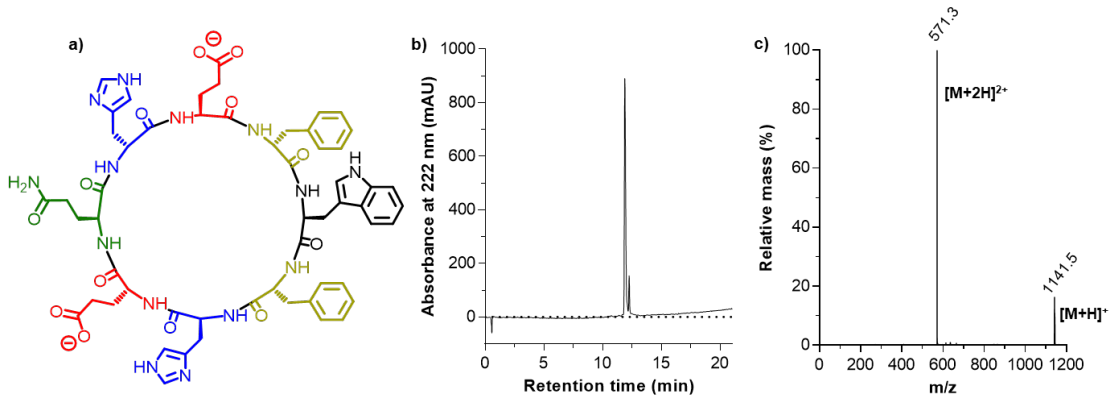


Figure S7. a) Structure of **CP_{2F}**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1141.4850 (calculated for $[M+H]^+$); 1141.4845 (found):

^{1}H-NMR (300 MHz, D_2O) δ: 1.77-2.02 (m, 8H, Gln, Glu-CH_2- x2), 2.13-2.22 (m, 4H, Glu-CH_2- x2), 2.76-3.26 (m, 10H, Phe-CH_2- x2, Trp-CH_2-, His-CH_2- x2), 4.11-4.70 (m, 8H, H_α), 7.02-7.45 (m, 17H, Phe-CH= x15, His-CH= x2), 8.53-8.65 (m, 2H, His-CH= x2) ppm.

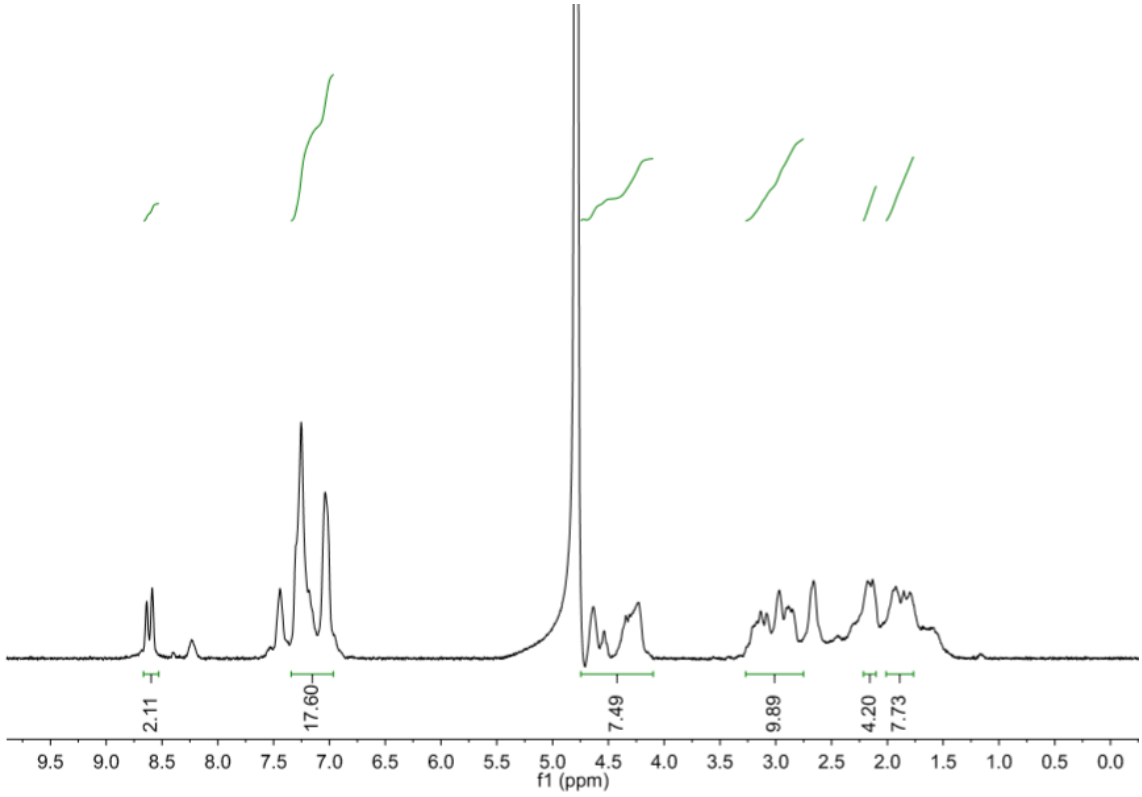


Figure S8. ^{1}H NMR (300 MHz) of **CP_{2F}** in D_2O.

4.5. CP_{3F}

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 12.2 min; m/z = 1102.5 ($[M+H]^+$), 551.8 ($[M+2H]^{2+}$).

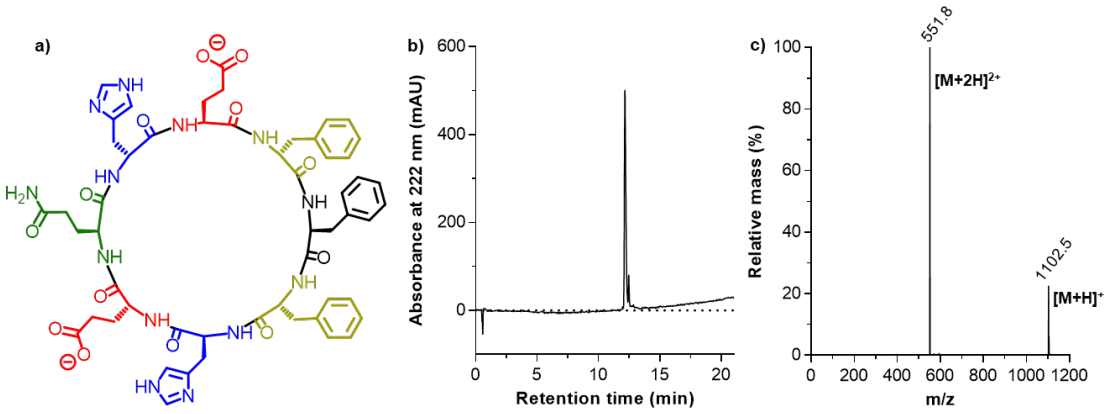


Figure S9. a) Structure of **CP_{3F}**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1102.4741 (calculated for $[M+H]^+$), 1102.4737 (found).

^{1}H NMR (300 MHz, D_2O) δ: 1.36-2.15 (m, 12 H, Glu, Gln), 2.62-3.25 (m, 10 H, His-CH_2- x2, Phe-CH_2- x2, Trp-CH_2-), 4.19-4.80 (m, 8H, H_α), 6.98-7.41 (m, 17H, Phe x3, His x2), 8.62 (s, 2H, His x2) ppm.

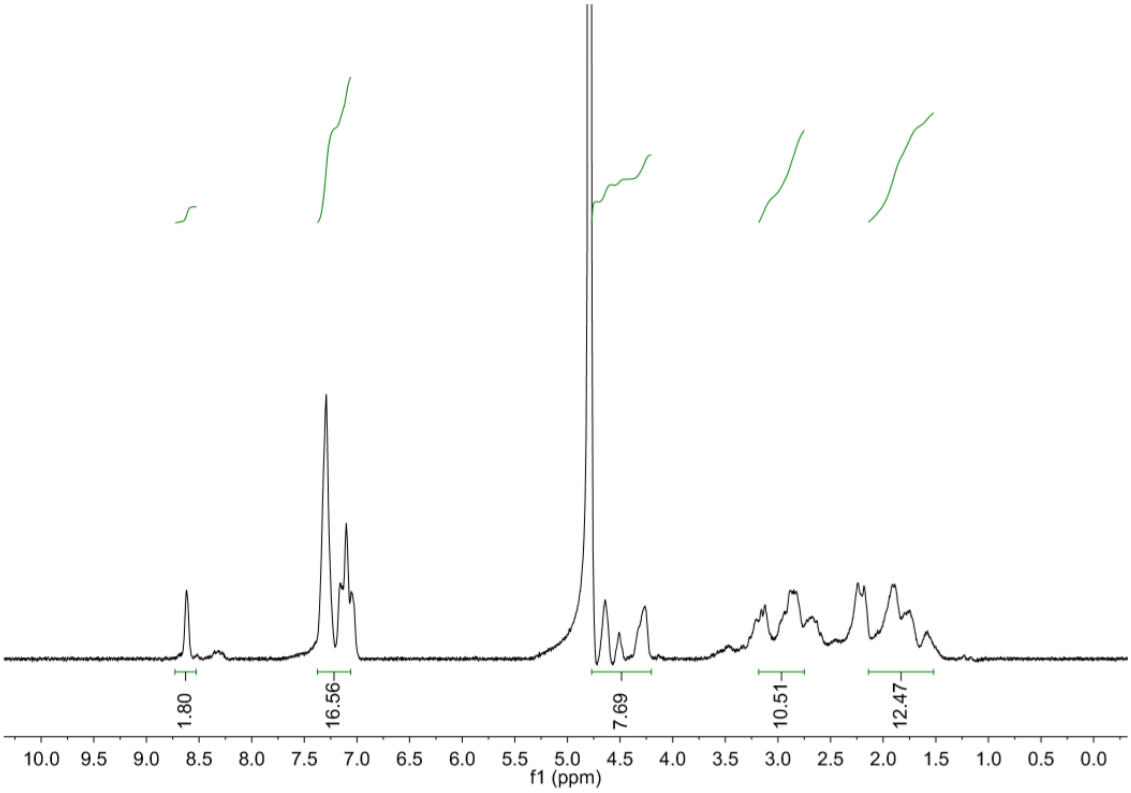


Figure S10. ^{1}H NMR (300 MHz) of **CP_{3F}** in D_2O.

4.6. CP_{oE}

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 10.9 min; m/z = 1073.5 ($[M+H]^+$), 537.3 ($[M+2H]^{2+}$).

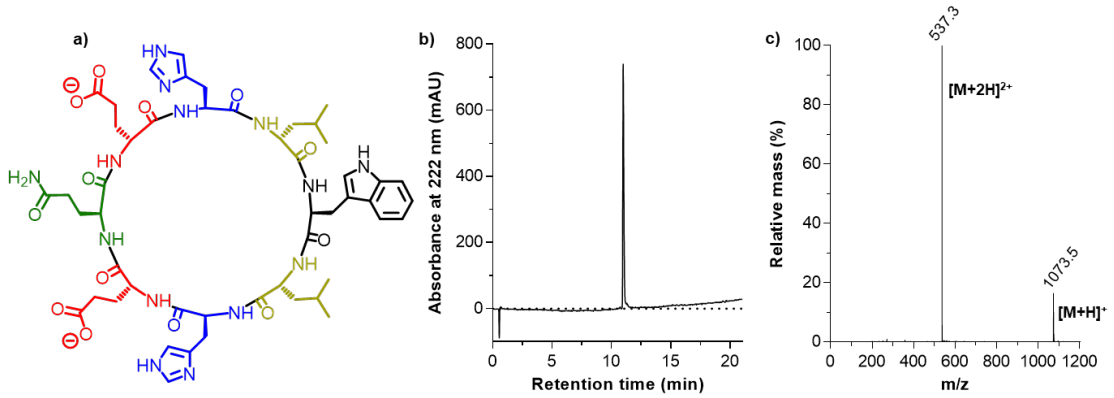


Figure S11. a) Structure of **CP_{oE}**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1073.5163 (calculated for $[M+H]^+$); 1073.5172 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.59-0.84 (m, 14H, Leu *i*-Bu x2, Leu-CH- x2), 1.15-1.35 (m, 4H, Leu-CH_2- x2), 1.85-2.38 (m, 12H, Glu x2, Gln), 2.97-3.43 (m, 6H, His-CH_2- x2, Trp-CH_2-), 3.80-4.70 (m, 8H, H_α), 7.11-7.18 (m, 2H, His-CH= x2), 7.18-7.37 (m, 3H, Trp-CH= x2, Trp-CH=), 7.40-7.69 (m, 2H, Trp-CH= x2), 8.67 (s, 2H, His-CH= x2) ppm.

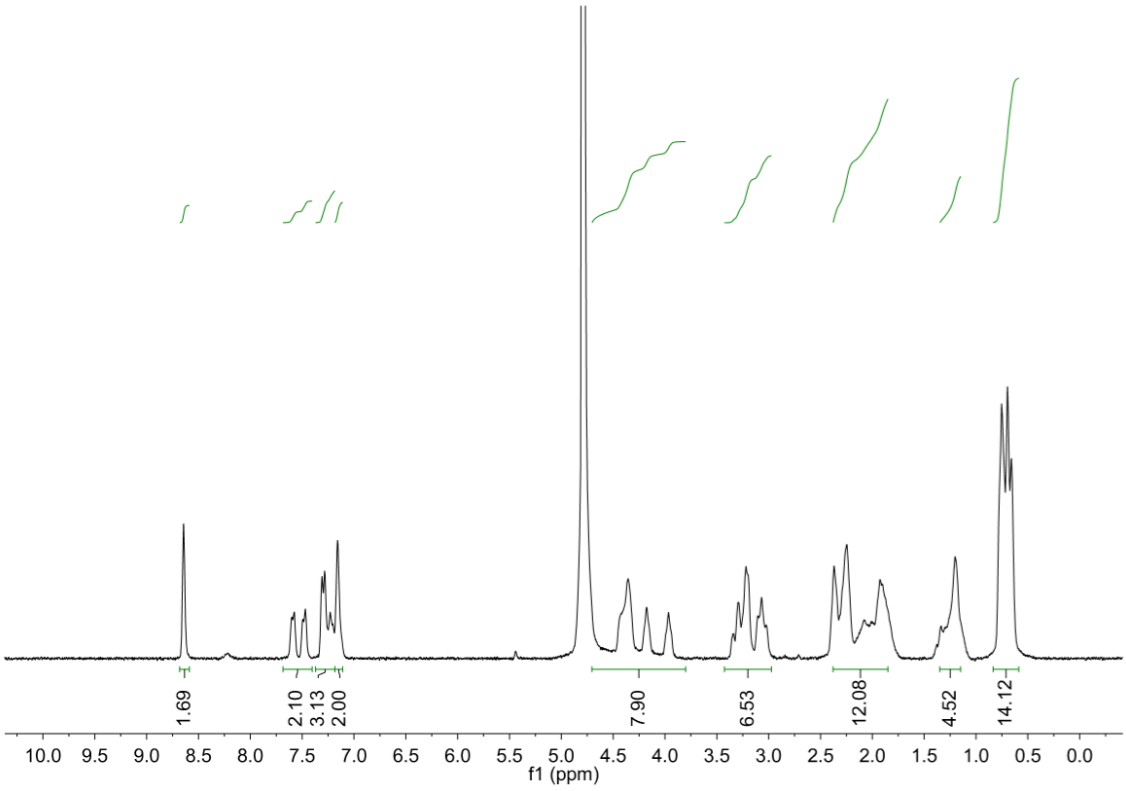


Figure S12. ^{1}H NMR (300 MHz) of **CP_{oE}** in D_2O.

4.7. CP_{2E2H}

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 11.0 min; m/z = 1073.5 ($[M+H]^+$), 537.4 ($[M+2H]^{2+}$).

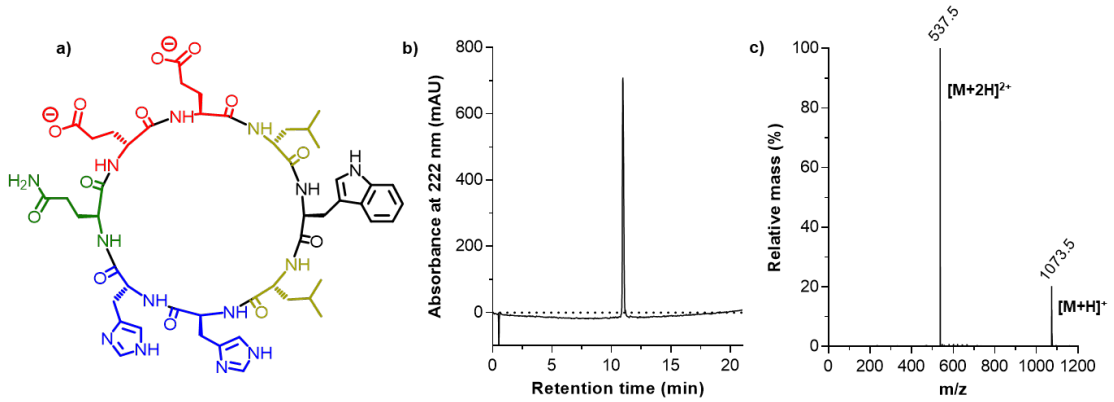


Figure S13. a) Structure of **CP_{2E2H}**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1073.5163 (calculated for $[M+H]^+$); 1073.5155 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.49-0.82 (m, 14H, Leu *i*-Bu x2, Leu-CH- x2), 1.18-1.37 (m, 4H, Leu-CH_2- x2), 1.86-2.38 (m, 12H, Glu x2, Gln), 2.87-3.38 (m, 6H, His-CH_2- x2, Trp-CH_2-), 3.83-4.69 (m, 8H, H_α), 6.95-7.37 (m, 2H, His-CH= x2, 3H, Trp-CH= x3), 7.48 (d, *J* = 7.5 Hz, 1H, Trp-CH=), 7.61 (d, *J* = 7.5 Hz, 1H, Trp-CH=), 8.62 (s, 2H, His-CH= x2) ppm.

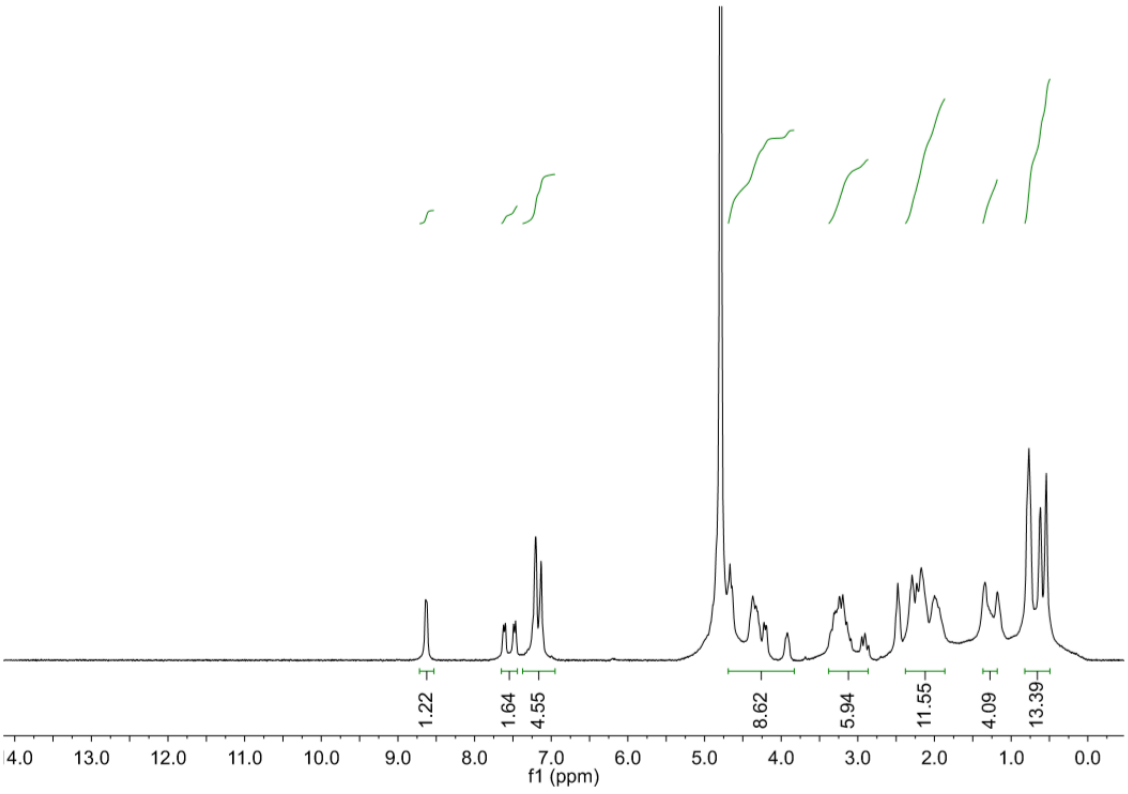


Figure S14. ^{1}H NMR (300 MHz) of **CP_{2E2H}** in D_2O.

4.8. CP_K

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; Gradient (A:B): 100:0 (0 min) → 25:75 (21 min); R_t = 11.2 min; m/z = 1055.6 ($[M+H]^+$), 528.4 ($[M+2H]^{2+}$).

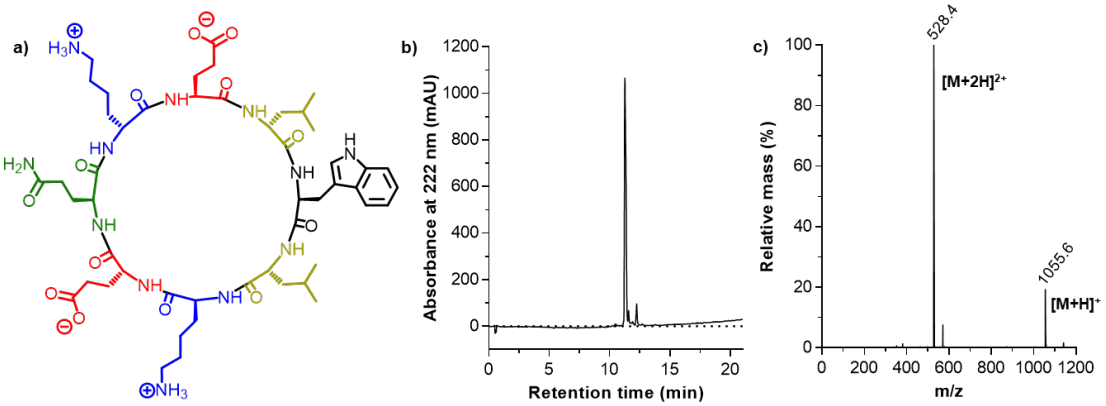


Figure S15. a) Structure of **CP_K**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1055.5884 (calculated for $[M+H]^+$); 1055.5877 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.41-0.91 (m, 12H, Leu *i*-Bu x2), 1.15-1.28 (m, 2H, Leu-CH- x2), 1.32-1.46 (m, 4H, Leu-CH_2- x2), 1.47-2.38 (m, 24H, Glu x2, Lys-CH_2- x6, Gln), 2.88-3.02 (m, 4H, Lys-CH_2- x2), 3.19-3.35 (m, 2H, Trp-CH_2-), 4.10-4.66 (m, 8H, H_α), 7.04-7.29 (m, 3H, Trp-CH= x3), 7.47-7.70 (m, 2H, Trp-CH= x2) ppm.

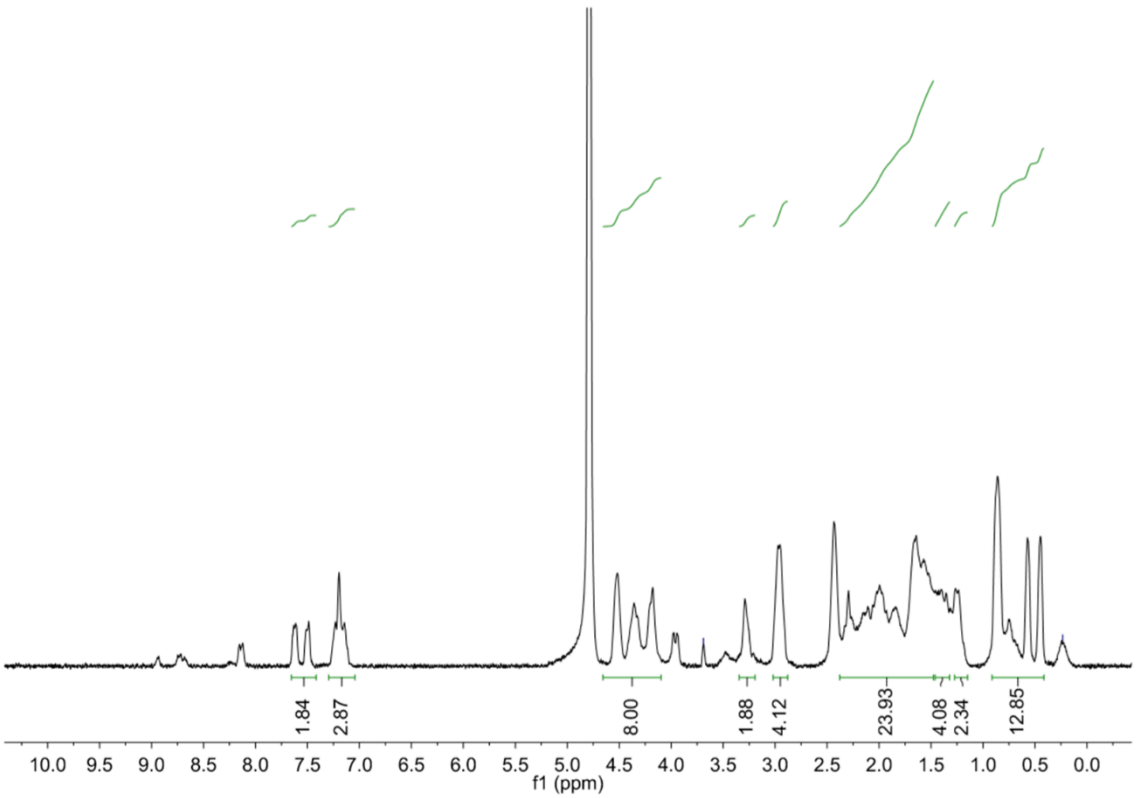


Figure S16. ^{1}H NMR (300 MHz) of **CP_K** in D_2O.

4.9. CP_{4E}

UHPLC-MS (C18-ESI, +eV) A = H_2O + 0.1% TFA; B = ACN + 0.1% TFA; 100:0 (0 min) → 25:75 (21 min); R_t = 12.2 min; m/z = 1057.5 ($[M+H]^+$), 529.4 ($[M+2H]^{2+}$).

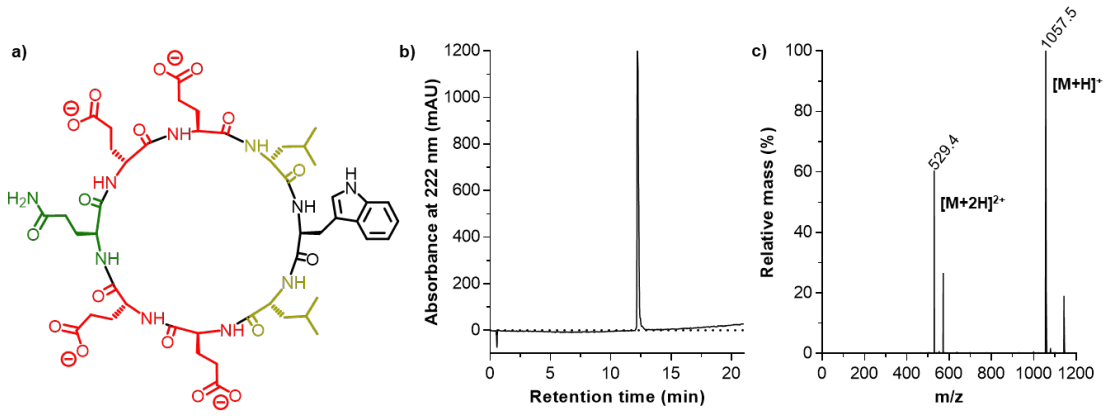


Figure S17. a) Structure of **CP_{4E}**; b) Chromatogram after purification; c) MS spectrum of the peak.

HR-MS (ESI, +eV): m/z = 1057.4837 (calculated for $[M+H]^+$); 1057.4836 (found).

^{1}H-NMR (300 MHz, D_2O) δ: 0.65-0.81 (m, 12H, Leu *i*-Bu x2), 1.20-1.29 (m, 2H, Leu-CH- x2), 1.34-1.46 (m, 4H, Leu-CH_2- x2), 1.87-2.43 (m, 20H, Glu x4, Gln), 3.12-3.36 (m, 2H, Trp-CH_2-), 4.01-4.43 (m, 8H, H_α), 7.03-7.34 (m, 3H, Trp-CH= x3), 7.47-7.69 (m, 2H, Trp-CH= x2) ppm.

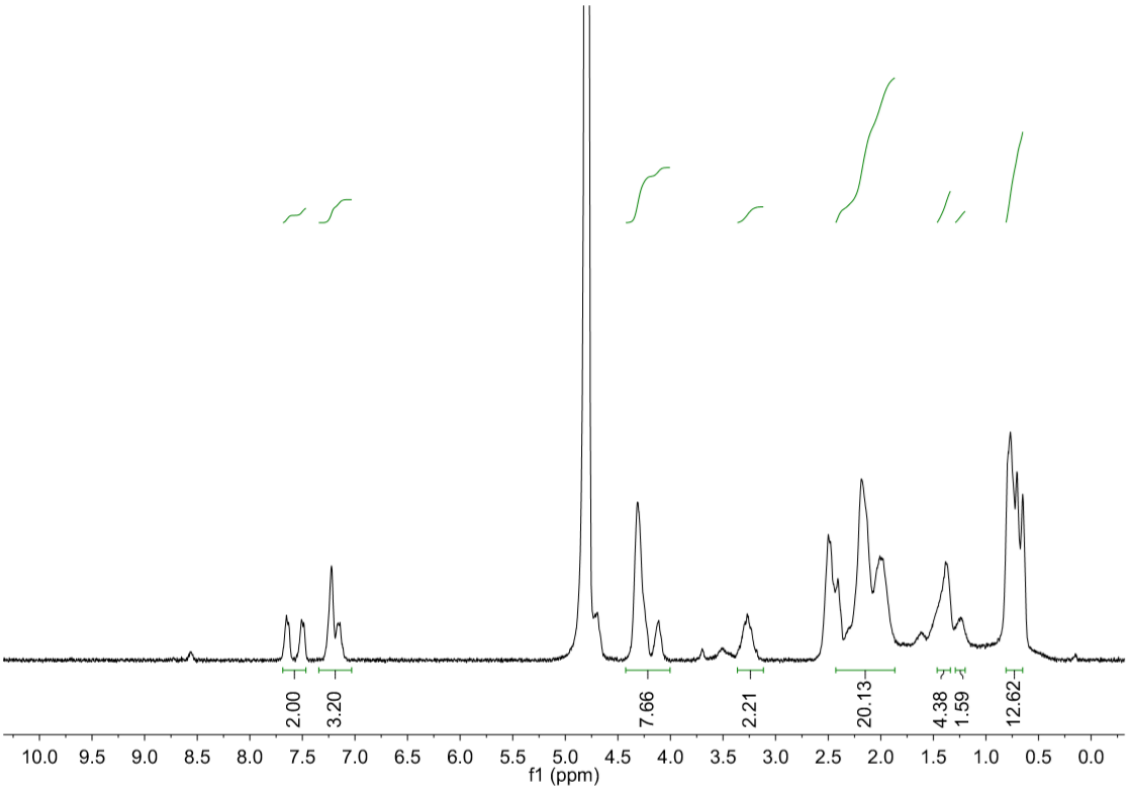


Figure S18. ^{1}H NMR (300 MHz) of **CP_{4E}** in D_2O.

5. Cyclic peptide self-assembly *versus* pH and NaCl

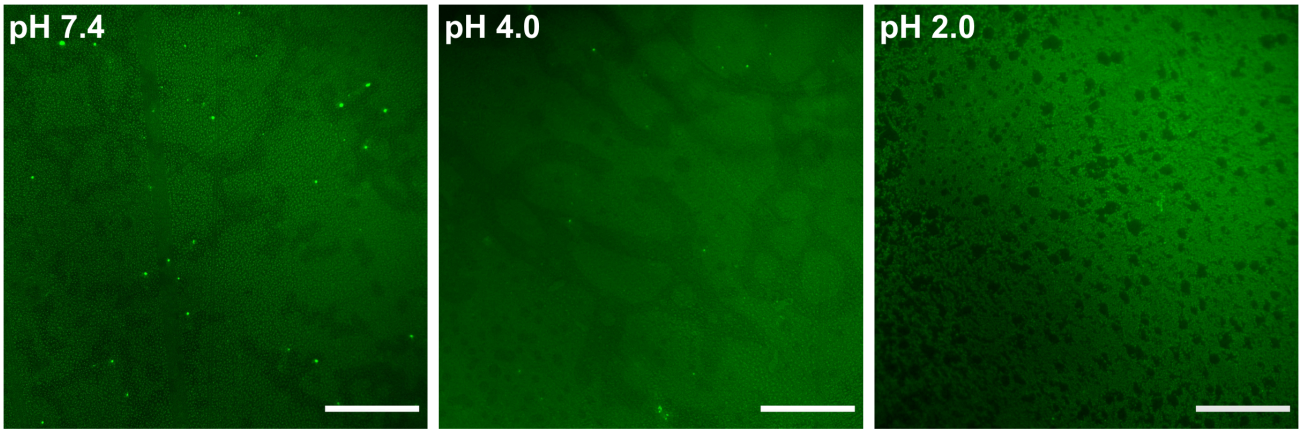


Figure S19. Epifluorescence micrographs of **CP_{2E2H}** in 20 mM sodium phosphate buffer adjusted at pH 7.4, 4.0 or 2.0 (see insets) in presence of thioflavin-T (10 μM). Scale bars = 50 μm.

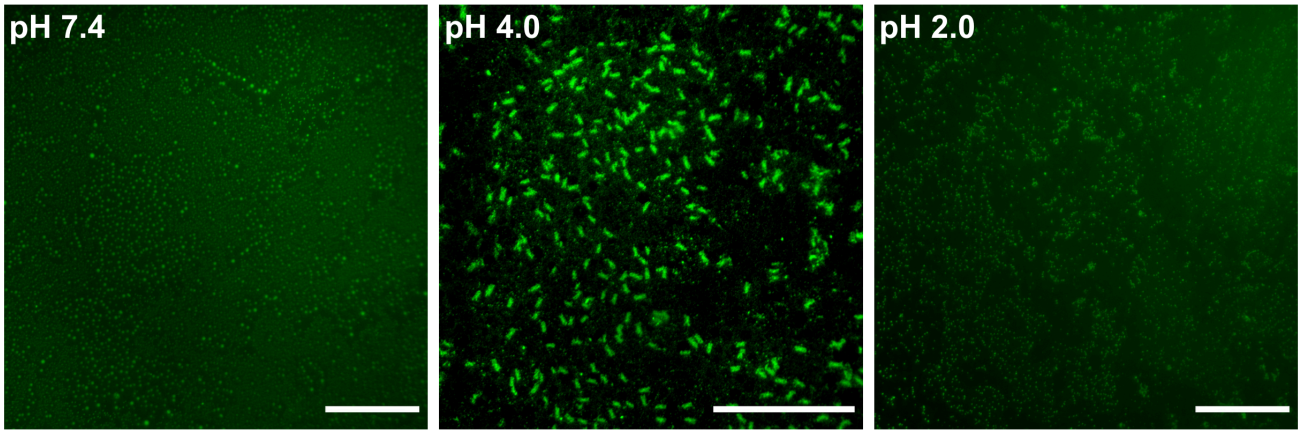


Figure S20. Epifluorescence micrographs of **CP_{4E}** in 20 mM sodium phosphate buffer adjusted at pH 7.4, 4.0 or 2.0 (see insets) in presence of thioflavin-T (10 μM). Scale bars = 50 μm.

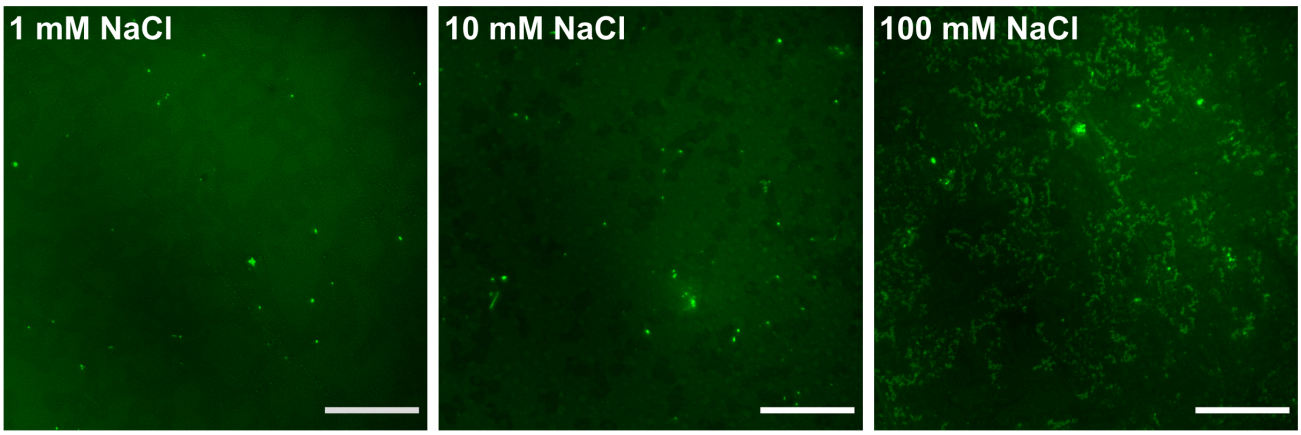


Figure S21. Epifluorescence micrographs of $\mathbf{CP_{2E2H}}$ in 20 mM sodium phosphate buffer at pH 7.4 with increasing concentrations of NaCl (see insets) in presence of thioflavin-T (10 μM). Scale bars = 50 μm.

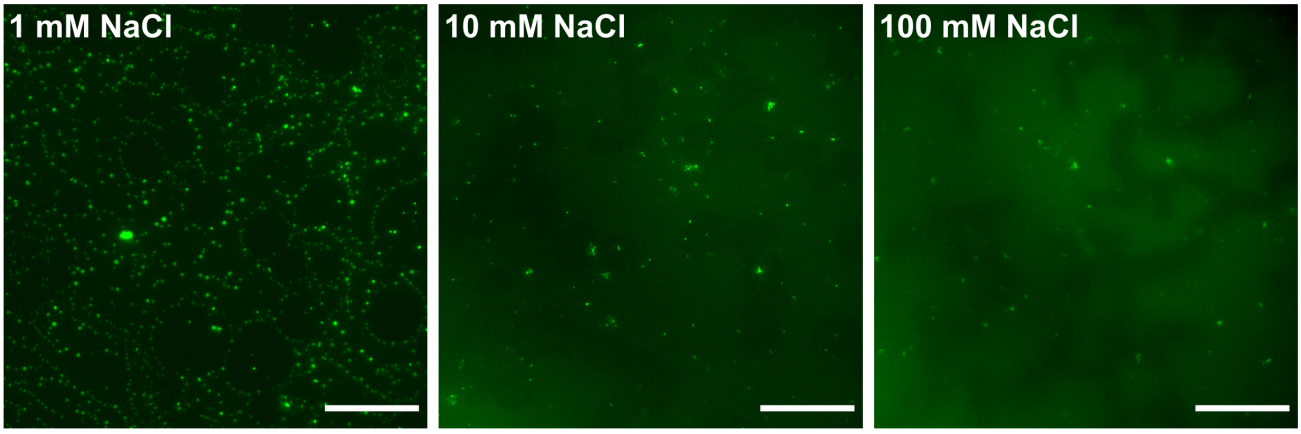


Figure S22. Epifluorescence micrographs of $\mathbf{CP_{4E}}$ in 20 mM sodium phosphate buffer at pH 7.4 with increasing concentrations of NaCl (see insets) in presence of thioflavin-T (10 μM). Scale bars = 50 μm.