Electronic Supplementary Material (ESI) for Chemical Communications

Design Supramolecular Amino Acids to Template Peptide Folding

Davoud Mozhdehi and Zhibin Guan

Department of Chemistry, University of California, Irvine, CA 92697

Supporting information

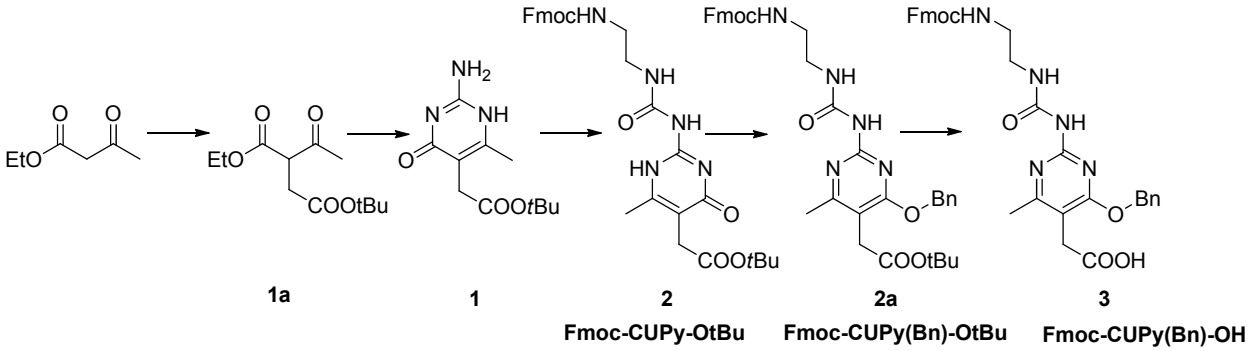
Table of Contents

General information

Wang resin was purchased from Aapptec and mono-Fmoc ethylene diamine hydrochloride was purchased from Novabiochem. Commercial solvents and reagents were used without further purification, unless otherwise stated. Reactions were carried out under an atmosphere of nitrogen. Column chromatography was performed on 40–63 μm silica gel (EMD Science) using flash chromatography. Solvents were removed by rotary evaporation. Residual solvents were removed under vacuum (< 0.01 mmHg). Precipitated products were dried under vacuum (< 0.01 mmHg) or by air suction through a filter funnel. High resolution mass spectra were obtained by electrospray ionization (ESI) on a Waters Micromass LCT Premier (instrument variation σ < 5 ppm). NMR spectra were recorded using 500 MHz Bruker AVANCETM, 600 MHz Bruker AVANCETM spectrometers. 125 MHz ^{13}C NMR spectra were obtained using a cryoprobe. ^{1}H spectra NMR were referenced with TMS (δ = 0.00 ppm), $CDCl_3$ (δ = 7.26 ppm), CD_2Cl_2 (δ = 5.32 ppm) and DMSO-d_6 (δ = 2.50 ppm). ^{13}C NMR spectra were referenced either with TMS (δ = 0.00 ppm), $CDCl_3$ (δ = 7.26 ppm) and DMSO-d_6 (δ = 77.2 ppm). Chemical shifts are reported in parts per million (ppm) on the δ scale. Analytical reverse phase HPLC (RP-HPLC) was performed on a Beckman gold system (Agilent Zorbax® Eclipse-XDB-C18 80 Å; 150 x 4.6 mm; particle size 5 μm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA), and preparative RP-HPLC was carried out on a Waters 600 HPLC machine (Agilent prep-C18 column, 250 x 21.2 mm; particle size 10 μm; solvent A: H_2O/0.1% TFA, solvent B= CH_3CN/0.1% TFA).

Synthesis of supramolecular amino acids

Synthesis of Fmoc-CUPy(Bn)-OH



Scheme S1: The synthetic scheme for Fmoc-CUPy(Bn)-OH (**3**).

Synthesis of 1a

To a round bottom flask containing ethyl acetoacetate (10 g, 76 mmol) and *tert*-butyl bromoacetate (12.6 g, 63.4 mmol) in 210 mL of DMF, was added potassium carbonate (11.39 g, 82 mmol). The reaction mixture was stirred at room temperature overnight. The mixture was then partitioned between 250 mL of EtOAc and 250 mL of water. The layers were then separated, and the aqueous layer was extracted with 2 × 200 mL of EtOAc. The combined organic layers were then dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford a yellow oil. The crude oil was dissolved in chloroform and purified by flash chromatography (10:90 EtOAc:hexane). The fractions containing pure product were combined and evaporated to afford compound **1a** as yellow oil (11.1g, 72%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K) δ 1.2 1.3(m, 3H), δ 1.4 (s, 9H), δ 2.3 (d, 3H,), δ 2.8 (dd, 2H), δ 3.0 (dd, 2H), δ 4.2 (m, 2H). MS(ESI) *m*/*z* calcd for $C_{12}H_{20}O_5$ $(M + Na)^+$ 267.13, found 267.1.

Synthesis of 1

To the solution of **1a** (10 g, 41 mmol) in 100 mL of ethanol was added guanidine hydrocarbonate (5.9 g, 33 mmol). The mixture was heated to reflux for 12 hours, cooled to the

room temperature. The solvent was evaporated *in vacuo* to afford orange crude oil. The residue was diluted with chloroform and purified by flash chromatography (12:88 MeOH:$CHCl_3$). The fractions containing pure product were combined and concentrated to afford compound **1** as yellow solid (4.8 g, 49%). ^{1}H NMR (500 MHz, DMSO-d_6, 298 K) δ 1.4 (s, 9H), δ 2.0 (s, 3H), δ 3.2 (s, 2H), δ 6.3 (s, 2H). MS(ESI) *m*/*z* calcd for $C_{11}H_{17}N_3O_3$ $(M + H)^+$ 240.13, found 240.1.

Synthesis of 2 (Fmoc-CUPy-O^{t}Bu)

To a suspension of **1** (2.98 g, 12.4 mmol) in 60 mL of chloroform was added 1,1'-carbonyldiimidazole (4.44 g, 27.4 mmol). The reaction mixture was heated at reflux for 2 h under nitrogen gas flow until clear solution was obtained. To the reaction mixture was added 100 mL of chloroform, and the organic layer was extracted with 3 × 50 mL of brine. The aqueous layers were then combined and extracted with 2 × 50 mL of chloroform. The combined chloroform layers were dried (Na_2SO_4) and filtered. Solvent was then evaporated *in vacuo* which resulted in a light yellow powder and was used without further purification.[1] To the CDI activated imidazolide (3.7 g, 11 mmol) and mono-Fmoc ethylenediamine hydrochloride, NovaBiochem, (4.60 g, 14.4 mmol) in 40 mL of chloroform was added 4.83 mL, (27.7 mmol), of N,N-diisopropylethylamine ,DIPEA. The reaction mixture was stirred for 6 hours under the flow of nitrogen gas. The solvent was then evaporated under reduced pressure. The residue was diluted with chloroform and purified by flash chromatography (6:94 MeOH:$CHCl_3$) to afford compound **2** as off white solid (5.39 g, 89%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K) δ 1.4 (s, 9H), δ 2.2 (s, 3H), δ 3.33 (m, 2H), δ 3.4 (m, 2H and m, 2H), δ 4.21 (t, H, *J* = 8), δ 4.35 (d, 2H, *J* = 7), δ 6.5 (s, 2H) ,δ 7.25 7.45 (m, 5H), δ = 7.6 (d, 2H, *J* = 7.5), δ 7.75 (d, 2H, *J* = 7.5), δ 10.2 (s, H), δ 11.8 (s, H), δ 12.9 (s, H). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K) δ 17, 28, 30, 32, 40.3, 40.5, 47, 66, 81, 112, 119, 125, 127, 127.6, 141.3, 144, 146, 156.5, 157, 170. HRMS (ESI) *m*/*z* calcd for $C_{29}H_{33}N_5O_6$ $(M + Na)^+$ 570.2328, found 570.2314.

Synthesis of 2a Fmoc-CUPy(Bn)-O^{t}Bu

To a round bottom flask containing **2** (5.4 g, 9.9 mmol) and benzyl bromide (2.58 g, 14.8 mmol) in 100 mL of DMF wad added potassium carbonate (2.04 g, 14.8 mmol). The reaction was then stirred at room temperature overnight. 110 mL of water was added to the reaction mixture and the aqueous layer was extracted with 2 × 100 mL chloroform. The organic layer was then washed with 100 mL of brine, dried ($MgSO_4$) and filtered. Finally the solvent was evaporated to afford a yellow solid. The crude solid was dissolved in chloroform and purified by flash column chromatography (50:50 EtOAc:hexane). The fractions containing pure product were combined and evaporated to afford **2a** as off-white solid (4.2 g, 67%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K) δ 1.4 (s, 9H), δ 2.32 (s, 3H), δ 3.32 (m, 2H), δ 3.46 (s, 2H), δ 3.50 (m, 2H), δ 4.21 (t, H, *J* = 8), δ 4.35 (d, 2H, *J* = 7), δ 5.32 (s, 2H) , δ 7.04 (s, 1H), δ 7.25 7.45 (m, 10H), δ = 7.6 (d, 2H, *J* = 7.5), δ 7.75 (d, 2H, *J* = 7.5), δ 9.4 (s, H). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K) δ 22, 28, 32, 39.5, 42, 47, 67, 68, 81,107, 119, 125, 127.1, 127.7,128.1,128.5, 136,141.3, 144, 146, 155.5,156, 157, 168, 169, 170. HRMS (ESI) *m*/*z* calcd for $C_{36}H_{39}N_5O_6$ $(M + Na)^+$ 660.2798, found 660.2791.

Synthesis of 3 (Fmoc-CUPy(Bn)-OH)

2a, (3.5 g, 5.5 mmol) was dissolved in 6 mL of 30% Trifluoroacetic acid, TFA, in DCM (% v/v). The clear solution was stirred until the deprotection was complete, judged by TLC or mass spectroscopy. The reaction mixture was then concentrated *in vacuo* and the product was precipitated by slow addition of reaction mixture to the 25 mL of cold Et_2O. The product was collected by centrifugation at 3000 rpm for 10 min and decantation of organic layer. The solid was dispersed in 20 mL of Et_2O and centrifuged at 3000 rpm for 10 min. After decantation of organic layer, **3** was obtained as off-white solid (2.9 g, 90%). ^{1}H NMR (500 MHz, DMSO-d_6, 298 K) δ 2.29 (s, 3H), δ 3.16 (m, 2H), δ 3.31 (m, 2H), δ 3.50 (s, 2H), δ 4.21 (t, H, *J* = 7), δ 4.4 (d, 2H, *J* = 7), δ 5.38 (s, 2H) , δ 7.25 7.45 (m, 10H), δ = 7.6 (d, 2H, *J* = 7.5), δ 7.8 (d, 2H, *J* = 7.5), δ

9.15 (s, H), δ 9.5 (s, H), δ 13.2 (s, H). ^{13}C NMR (125 MHz, DMSO-d_6, 298 K) δ 22, 31, 41, 47, 66, 68, 107, 120, 125, 127.5, 128.1, 128.2, 128.5, 128.8, 137, 141, 144, 155, 156, 157, 168, 172. HRMS (ESI) *m*/*z* calcd for $C_{32}H_{31}N_5O_6$ $(M + H)^+$ 582.2352 , found 582.2341.

Fmoc-CUPy(Bn)-OH HPLC trace

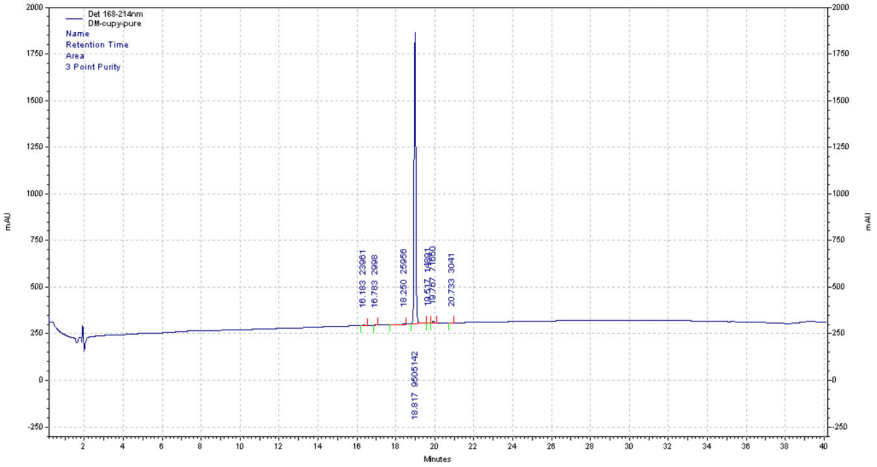
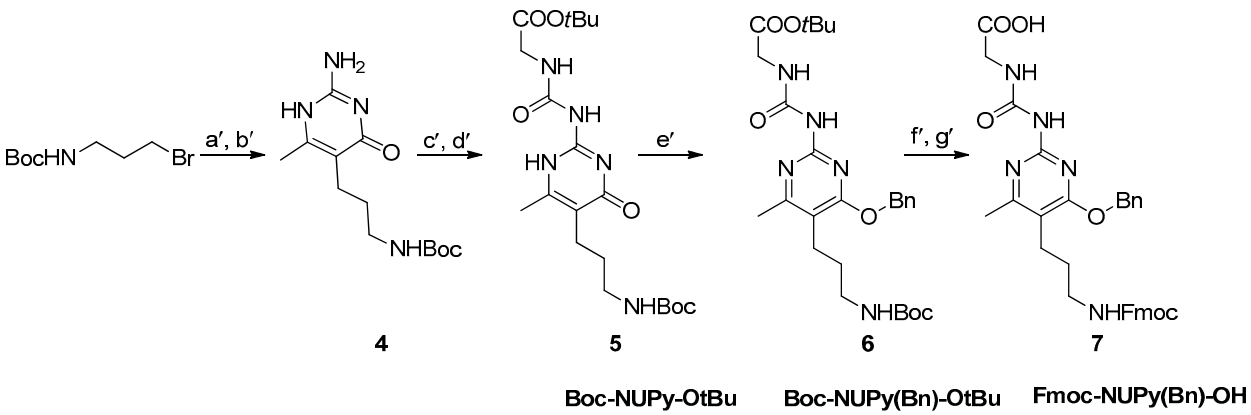


Figure **S1.** The HPLC trace of purified Fmoc-CUPy(Bn)-OH (**3**).

Synthesis of Fmoc-NUPy(Bn)-OH



Scheme S2. The synthetic scheme for Fmoc-NUPy(Bn)-OH (**7**).

Synthesis of 4

To a round bottom flask containing ethyl acetoacetate (4.89 g, 37.6 mmol) and tert-butyl (3-bromopropyl)carbamate (8.14 g, 34.2 mmol) in 115 mL of DMF, was added potassium carbonate (8.03 g, 58.1 mmol). The reaction mixture was stirred at room temperature overnight. The mixture was then partitioned between 150 mL of EtOAc and 250 mL of water. The layers were then separated, and the aqueous layer was extracted with 2 × 100 mL of EtOAc. The combined organic layers were then dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford a yellow oil. The crude oil was then dissolved in 170 mL of ethanol and guanidine hydrocarbonate (6.99 g, 37.6 mmol) was added to this mixture. The mixture was heated to reflux for 12 hours and subsequently cooled to room temperature. The solvent was evaporated *in vacuo* to afford orange crude oil. The residue was diluted with chloroform and purified by flash chromatography (12:88 MeOH:$CHCl_3$). The fractions containing pure product were combined and concentrated to afford compound **4** as a yellow solid (5.79 g, 60%). ^{1}H NMR (500 MHz, DMSO-d_6, 298 K) δ 1.36 (s, 9H), δ 1.43 (m, 2H), δ 2.92 (s, 3H), δ 2.23 (m, 2H), δ 2.86 (m, 2H), δ 6.24 (s, 2H), δ 6.74 (s,1H). MS (ESI) *m*/*z* calcd for $C_{13}H_{22}N_4O_3$ $(M+H)^+$: 283.34,found 283.1.

Synthesis of 5 (Boc-NUPy-O^{t}Bu)

To a suspension of **4** (5.70 g, 20.19 mmol) in 100 mL of chloroform was added 1,1'-carbonyldiimidazole (8.33 g, 50.5 mmol). The reaction mixture was heated at reflux for 2 h under nitrogen gas flow until clear solution was obtained. To the reaction mixture was added 200 mL of chloroform, and the organic layer was extracted with 2 $\times$ 200 mL of brine. The aqueous layers were then combined and extracted with 200 mL of chloroform. The combined chloroform layers were dried (Na_2SO_4) and filtered. Solvent was then evaporated *in vacuo* which resulted in a light yellow powder that was used without further purification.[1] To the CDI activated imidazolide solution (6.00 g, 15.94 mmol) and glycine *tert*-butyl ester hydrochloride (4.01 g, 23.91 mmol) was added 8.33 mL of DIPEA (47.8 mmol). The reaction mixture was stirred for six hours under the flow of nitrogen gas. The solvent was then evaporated under reduced pressure. The residue was diluted with chloroform and purified by flash chromatography (8:92 MeOH:$CHCl_3$) to afford compound **5** as an off white solid (4.97 g, 71%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K) δ 1.4 (s, 9H and s, 9H), δ 1.65 (m, 2H), δ 2.26 (s, 3H), δ 2.50 (t, 2H, J = 7), δ 3.06 (m, 2H), δ 3.89 (d, 2H, J = 5), δ 5.5 (s, H), δ 10.64 (s, H), δ 12.04 (s, H), δ = 12.75 (s, H). MS (ESI) *m*/*z* calcd $C_{20}H_{33}N_5O_6$ $(M+H)^+$: 440.51, found 440.1.

Synthesis of 6 (Boc-NUPy(Bn)-O^{t}Bu)

To a round bottom flask containing **5** (3.5 g, 7.96 mmol) and benzyl bromide (2.04 g, 11.9 mmol) in 80 mL of DMF wad added potassium carbonate (1.65 g, 11.95 mmol). The reaction was then stirred at room temperature overnight. 200 mL of water was added to the reaction mixture and the aqueous layer was extracted with 2 $\times$ 150 mL chloroform. The organic layer was then washed with 100 mL of brine, dried ($MgSO_4$) and filtered. Finally the solvent was evaporated to afford a yellow solid. The crude solid was dissolved in chloroform and purified by flash column chromatography (50:50 EtOAc: hexane). The fractions containing pure product were combined and evaporated to afford compound **6** as off-white solid (2.18 g, 52%). ^{1}H NMR

(500 MHz, $CDCl_3$, 298 K) δ 1.4 (s, 9H), δ 1.5 (s, 9H), δ 1.5 (m, 2H), δ 2.4 (s, 3H), δ 2.57 (t, 2H, *J* = 8), δ 3.11 (m, 2H), δ 4.05 (d, 2H, *J* = 5), δ 5.4 (s, 2H), δ 7.08 (s, H), δ 7.32 7.41 (m, 5H), δ = 9.7 (s, H).^{13}C NMR (125 MHz, $CDCl_3$, 298 K) δ 22, 28, 32, 39.5, 42, 47, 67, 68, 81,107, 119, 125, 127.1, 127.7,128.1,128.5, 136,141.3, 144, 146, 155.5,156, 157, 168, 169, 170. HRMS (ESI) *m*/*z* calcd for $C_{27}H_{39}N_5O_6$ $(M + Na)^+$ 552.2798, found 552.2803.

Synthesis of 7 (Fmoc-NUPy(Bn)-OH)

Compound **6** (2.0 g, 3.78 mmol) was dissolved in 5 mL of 30% TFA in DCM (% v/v). The clear solution was stirred until the deprotection was complete, judged by TLC or mass spectroscopy. The reaction mixture was then concentrated *in vacuo* and the product was precipitated by slow addition of reaction mixture to the 25 mL of cold Et_2O. The product was collected by centrifugation at 3000 rpm for 10 min. The organic layer was decanted, and the solid was dispersed in 80 mL of 1:1 mixture of acetonitrile and water. To the resulting suspension was added K_2CO_3 (0.8 g, 5.79 mmol). Fmoc-OSu (1.3 g, 3.85 mmol) was then slowly added to the mixture under vigorous stirring, after that the reaction mixture was stirred for 12 hours. pH of the solution was adjusted to 1 by slow addition of concentrated HCl. The precipitate was then extracted from aqueous layer by 3 × 100mL of EtOAc. The organic layer was then washed with 2 × 100 mL of cold water and 100 mL of brine. The solvent was evaporated and the purity of product was analyzed using NMR. Further purification was accomplished using preparative HPLC where needed to afford compound **7** as white solid (1.35 g, 60%). ^{1}H NMR (500 MHz, DMSO-d_6, 298 K) δ 1.6 (m, 2H), δ 2.36 (s, 3H), δ 2.52 (solvent peak and m, 2H), δ 3.0 (m, 2H), δ 3.9 (d, 2H, *J* = 4.5), δ 4.1 (t, 2H, *J* = 4.5), δ 4.3 (d, 2H, *J* = 6.5), δ 5.3 (s, 2H), δ 7.25 7.45 (m, 10H), δ = 7.7 (d, 2H, *J* = 7.5), δ 7.9 (d, 2H, *J* = 7.5), δ 9.4 (s, H), δ 9.6 (s, H), δ 12.8 (s, H). ^{13}C NMR (125 MHz, DMSO-d_6, 298 K) δ 21, 22, 29, 43, 47, 65, 68, 112, 121, 126, 127, 128, 128.3, 128.4, 128.8, 136, 141, 144, 154, 155, 156, 165, 167, 172. HRMS (ESI) *m*/*z* calcd for $C_{33}H_{33}N_5O_6$ $(M + H)^+$ 596.2509, found 596.2513.

Fmoc-NUPy(Bn)-OH HPLC trace

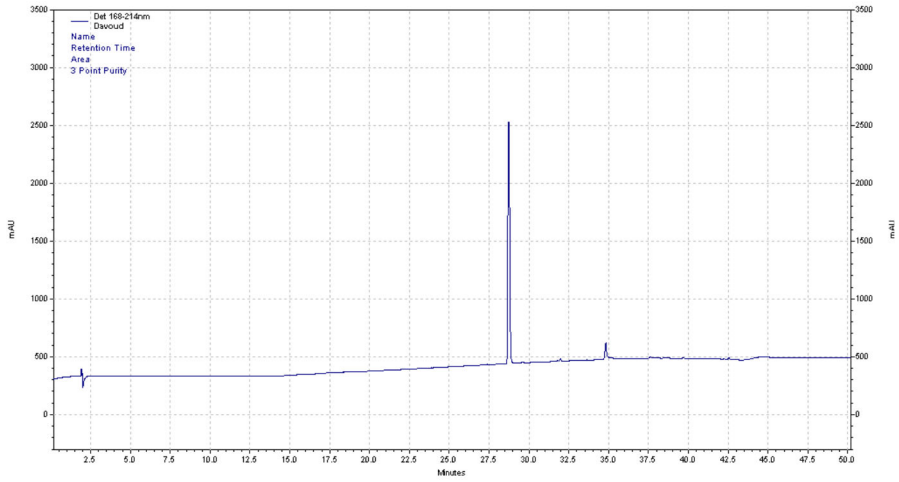


Figure S2. The HPLC trace of the purified Fmoc-NUPy(Bn)-OH (**7**).

Molecular modeling procedure

Maestro Molecular modeling package with MacroModel engine was utilized to aid the design process. A model of each peptide was created and subjected to energy minimization using AMBER* forcefield with $CHCl_3$, GB/SA implied solvent parameter after constraining the distance between hydrogen bonding edge of UPy amino acids to 2±0.1 Å. Minimized structures were then visually inspected for the orientation of UPy supramolecular amino acids.

The models that passed previous test, were further optimized by 5000 steps Monte Carlo conformational search and minimized using AMBER* force field and $CHCl_3$ implied solvent parameter. Constrained parameters were chosen similar to energy minimization calculations. The 30 lowest energy structures were then visually inspected for unwanted torsions or strains.

Peptide synthesis

Typical peptide synthesis procedure

Weighted solid support resin, 90 mg Wang resin 0.58mmol/g, was added to a Bio-Rad Poly-Prep chromatography column and was soaked with DMF (2 × ca. 5 mL, 10 min each) and drained under nitrogen pressure. Standard amino acids (3 equiv) were activated with HCTU (3 equiv) and DIPEA (6 equiv) in ca. 1 mL solution of DMF and after 30 s were added to the resin. The resin was gently agitated for 20 min and was subjected to second coupling. Fmoc-CUPy(Bn)-OH or Fmoc-NUPy(Bn)-OH was coupled to the resin using 2 equiv with HCTU (2 equiv) and DIPEA (4 equiv) in a ca. 1-mL DMF. The solution was added to the resin and the resin was gently agitated for 4–12 h. Fmoc deprotection was carried out by soaking the resin with a solution of 20% piperidine–DMF (2 × ca. 5 mL for 10 min) and gently agitating the resin. After deprotection of the Fmoc group, the resin was soaked with DMF (5 × ca. 3 mL for 1 min) and CH_2Cl_2 (5 × ca. 3 mL for 1 min). Couplings to the amino groups were monitored by the 2,4,6-trinitrobenzene sulphonic acid to determine coupling completeness. The peptide was cleaved from resin using 5 mL of TFA/water/triisopropylsilane (9/0.5/0.5, v/v/v) and mixed for ca. 3 h. The cleavage mixture was then concentrated *in vacuo* and precipitated by addition to cold ether followed by centrifugation and decantation of ether layer. Purification was accomplished using preparative RP-HPLC (water–CH_3CN buffers with 0.1 % TFA) after esterification.

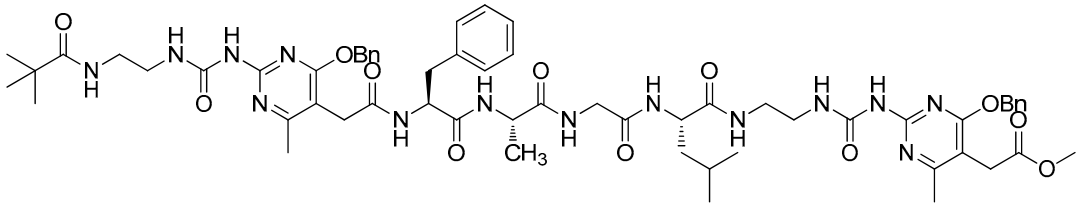
Typical Esterification Procedure

The crude white precipitate from the peptide synthesis was then dissolved in small amount of DCM and was slowly added to cooled solution of 1M acetyl chloride in anhydrous methanol and was stirred overnight.[2] The solution was then concentrated *in vacuo* and the residues was dissolved in water and acetonitrile and purified by RP-HPLC. Typical peptide yields were between 16-20%.

Typical Hydrogenolysis Procedure

A 17x60 mm scintillation vial, with plastic septum, a gas inlet adapter fitted with a hydrogen balloon, and a magnetic stirring bar was charged with 0.6-1 mL of methanol, and protected peptide (5.0 mg in 250 µL DCM). 5% $Pd/CaCO_3$ (2 mg) was added to the reaction mixture under a nitrogen flow. The vial filled with hydrogen and then maintained under the hydrogen atmosphere for 6 h. The reaction mixture was then diluted by addition of 4 mL of acetonitrile and 1 mL of water. The Pd catalyst was removed by centrifugation and the reaction mixture was purified using RP-HPLC.

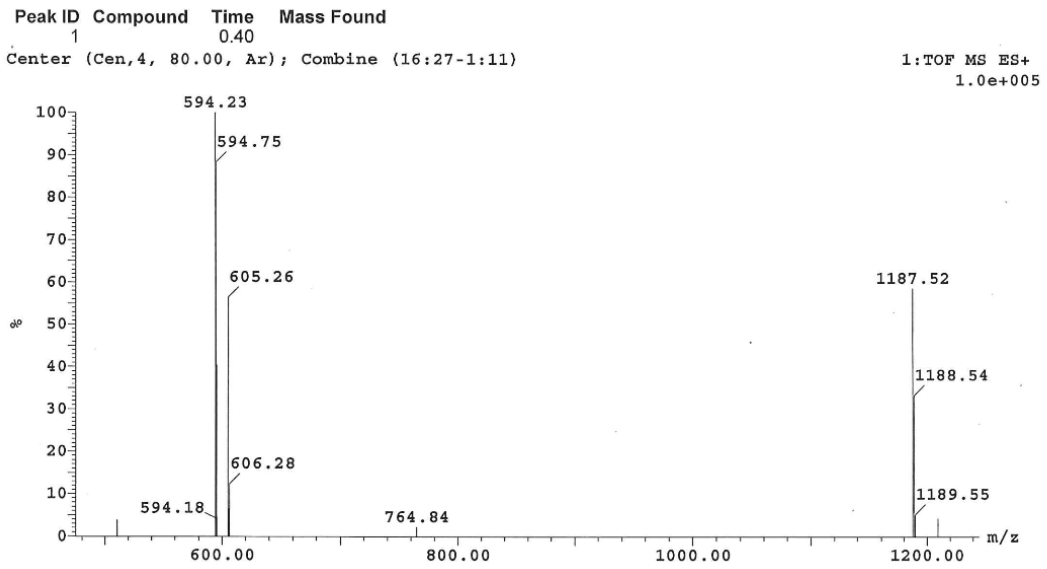
Synthesis of Peptide I



Chemical Formula: $C_{60}H_{78}N_{14}O_{12}$
Exact Mass: 1186.59

Peptide	Sequence
I	Piv-CUPy(Bn)-FAGL-CUPy(Bn)-OMe

MS (ESI), m/z calcd for $[(C_{60}H_{78}N_{14}O_{12} + H)/ 1]^+$ = 1187.59; found 1187.52; MS (ESI), m/z calcd for $[(C_{60}H_{78}N_{14}O_{12} + 2H)/ 2]^+$ = 594.29; found 594.23.



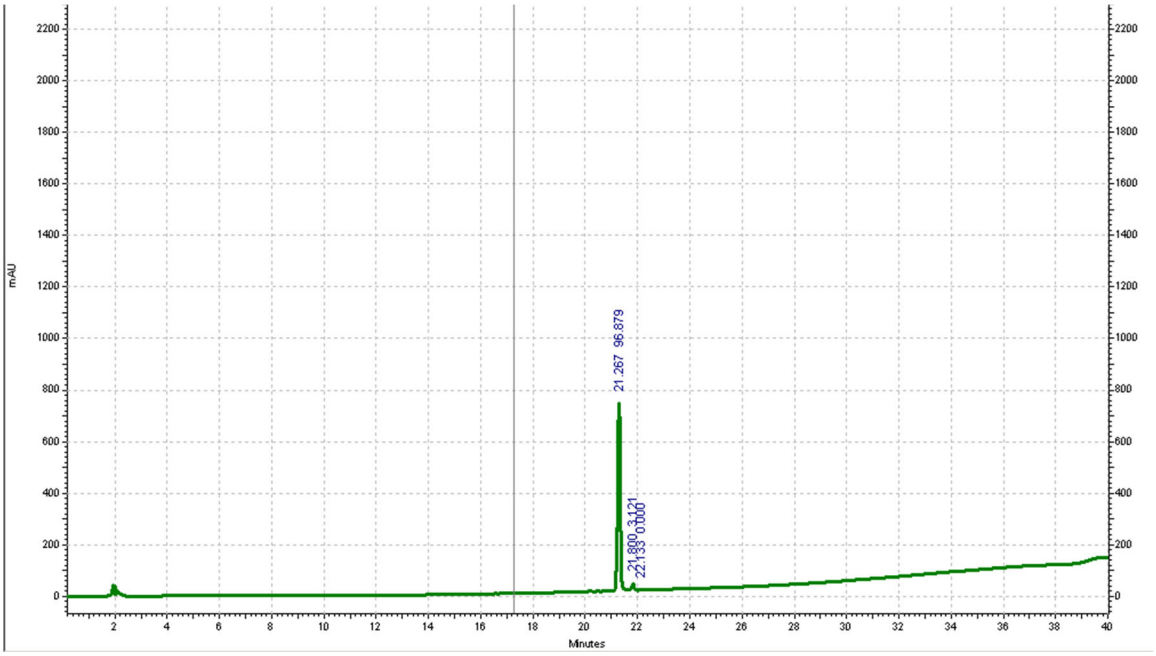
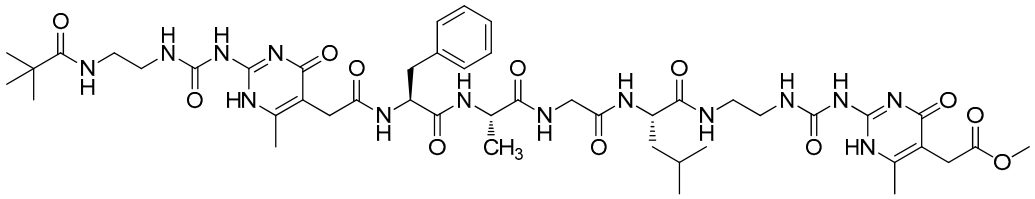


Figure S3. ESI-MS and analytical HPLC trace for Peptide **I**. (Agilent Zorbax® Eclipse-XDB-C18 80 Å; 150 x 4.6 mm; particle size 5 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA, 0-100% CH3CN over 35 minutes)

Synthesis of Peptide II



Chemical Formula: $C_{46}H_{66}N_{14}O_{12}$
Exact Mass: 1006.50

Peptide	Sequence
II	Piv-CUPy(Bn)-FAGL-CUPy(Bn)-OMe

MS (ESI), m/z calcd for $[(C_{46}H_{66}N_{14}O_{12} + H)/ 1]^+$ = 1007.50; found 1007.45; MS (ESI), m/z calcd for $[(C_{46}H_{66}N_{14}O_{12} + 2H)/ 2]^+$ = 504.25; found 504.23.

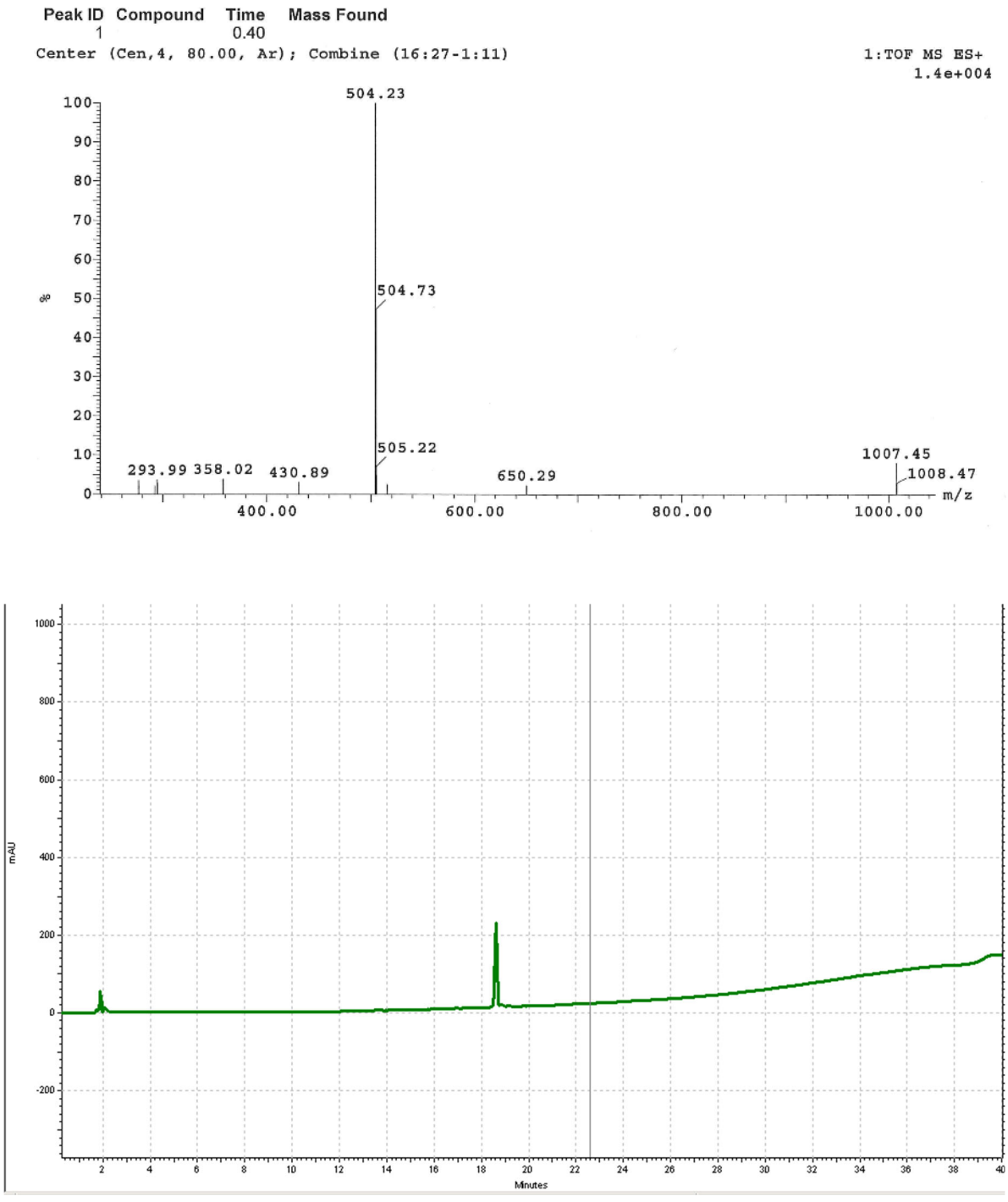
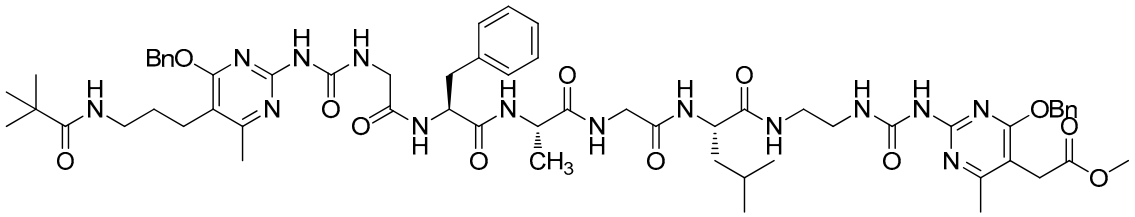


Figure S4. ESI-MS and analytical HPLC trace for Peptide II. (Agilent Zorbax® Eclipse-XDB-C18 80 Å; 150 x 4.6 mm; particle size 5 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA**,** 0-100% CH3CN over 35 minutes)

Synthesis of Peptide III



Chemical Formula: $C_{61}H_{80}N_{14}O_{12}$
Exact Mass: 1200.61

Peptide	Sequence
III	Piv-NUPy(Bn)-FAGL-CUPy(Bn)-OMe

MS (ESI), m/z calcd for $[(C_{61}H_{80}N_{14}O_{12} + H)/ 1]^+$ = 1201.61; found 1201.45; MS (ESI), m/z calcd for $[(C_{61}H_{80}N_{14}O_{12} + 2H)/ 2]^+$ = 601.30; found 601.24.

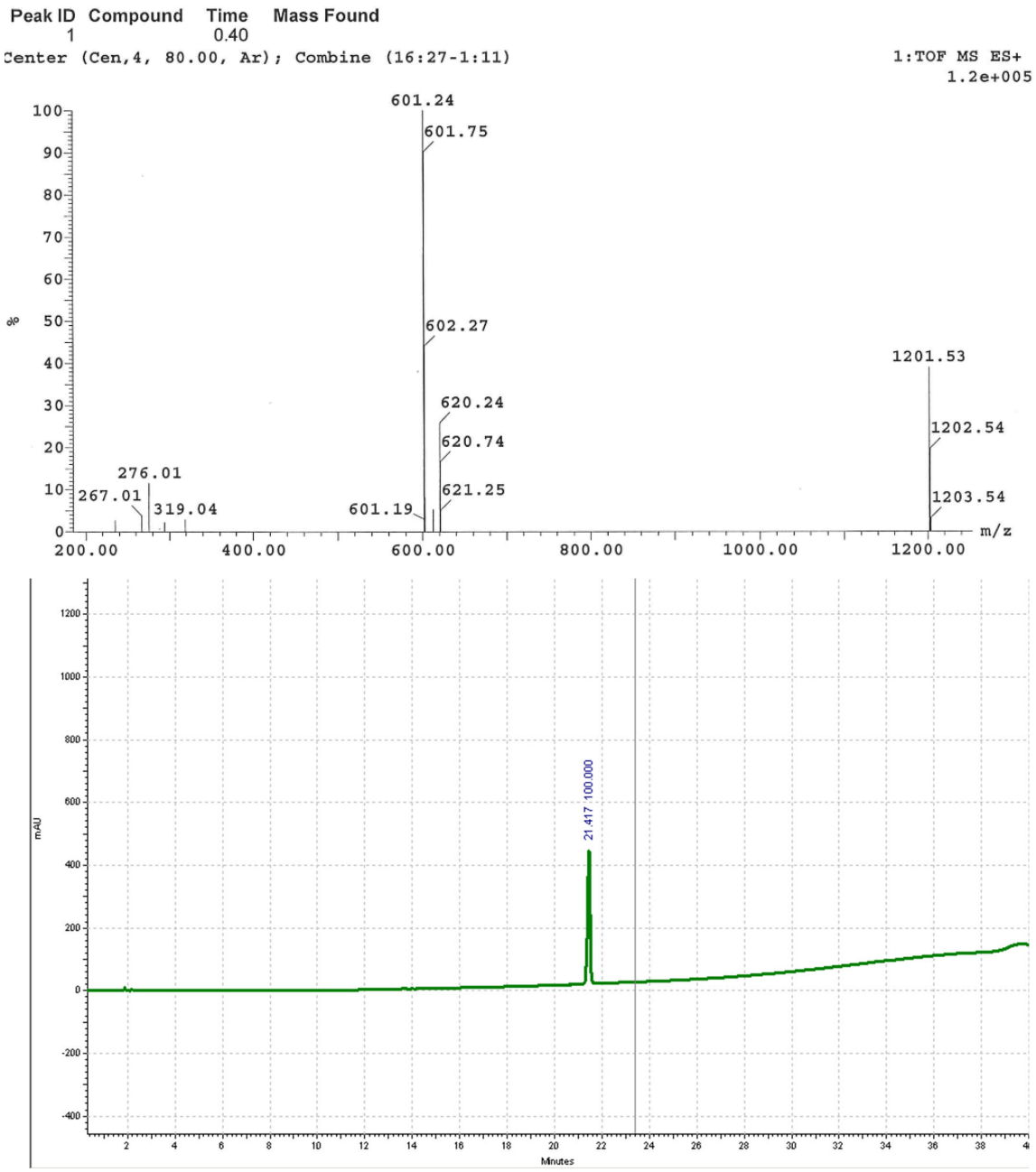
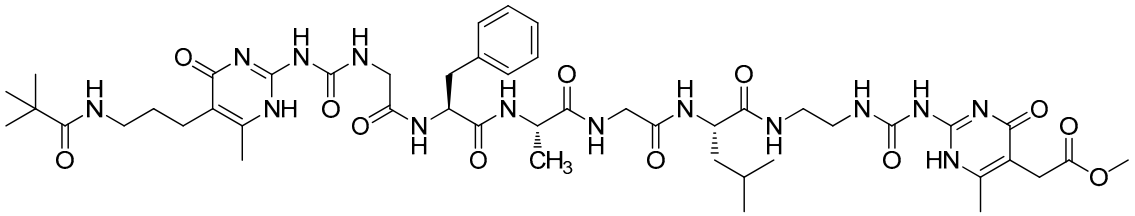


Figure S5. ESI-MS and analytical HPLC trace for Peptide **III**. (Agilent Zorbax® Eclipse-XDB-C18 80 Å; 150 x 4.6 mm; particle size 5 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA, 0-100% CH3CN over 35 minutes)

Synthesis of Peptide IV



Chemical Formula: $C_{47}H_{68}N_{14}O_{12}$
Exact Mass: 1020.51

Peptide	Sequence
IV	Piv-NUPy-FAGL-CUPy-OMe

MS (ESI), m/z calcd for $[(C_{61}H_{80}N_{14}O_{12} + 2H)/ 2]^+$ = 511.25; found 511.22.

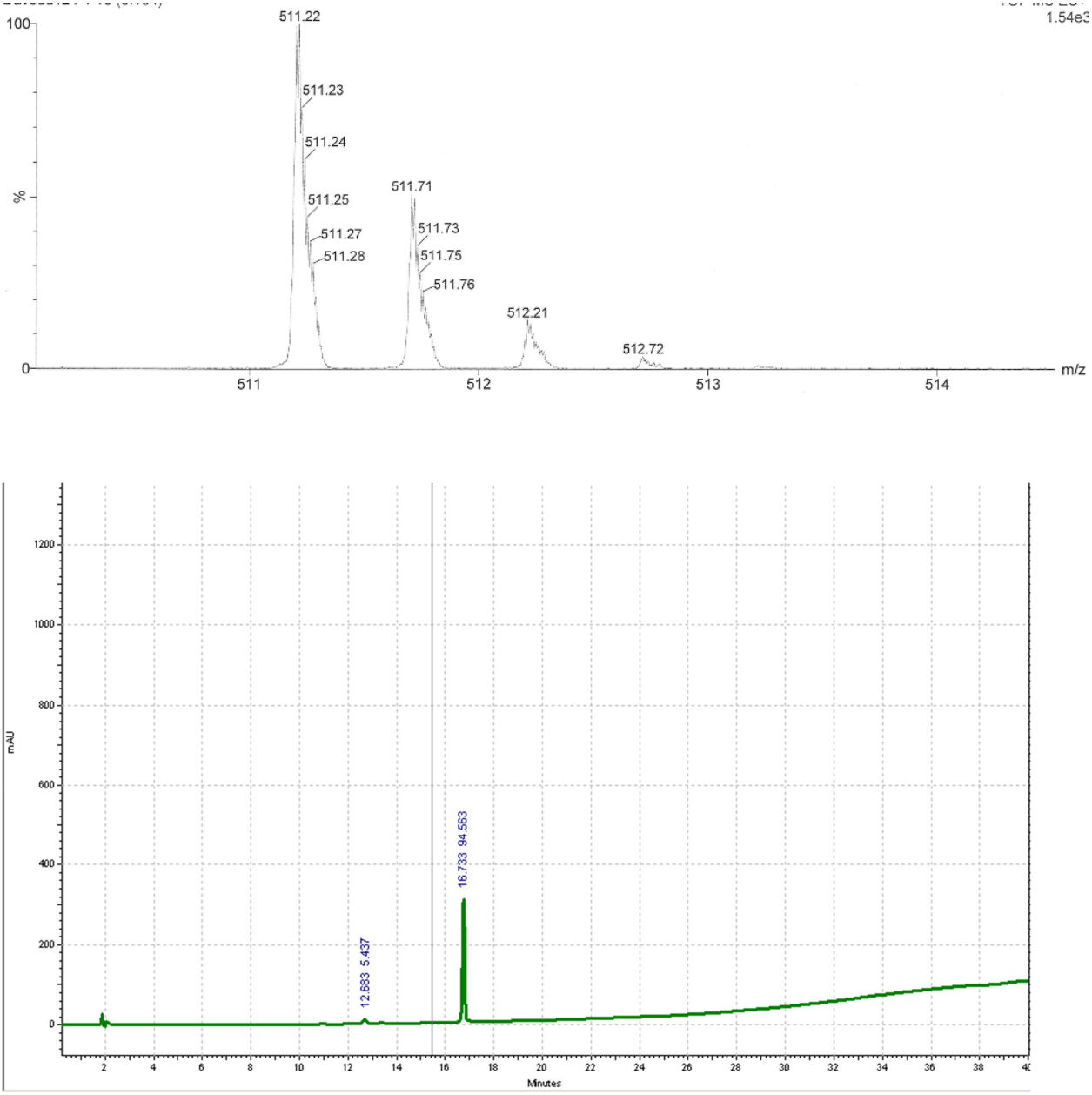
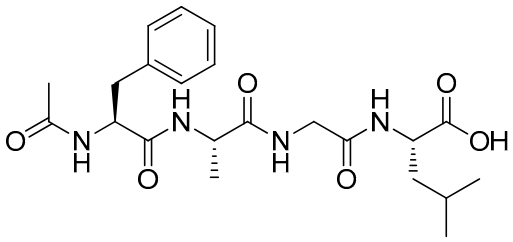


Figure S6. ESI-MS and analytical HPLC trace for Peptide **IV**. (Agilent Zorbax® Eclipse-XDB-C18 80 Å; 150 x 4.6 mm; particle size 5 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA, 0-100% CH3CN over 35 minutes)

Synthesis of Ac-FAGL-OH



Chemical Formula: $C_{22}H_{32}N_4O_6$
Exact Mass: 448.23

MS (ESI), m/z calcd for $[(C_{22}H_{32}N_4O_6 + H)]^+$ = 449.23; found 449.33.

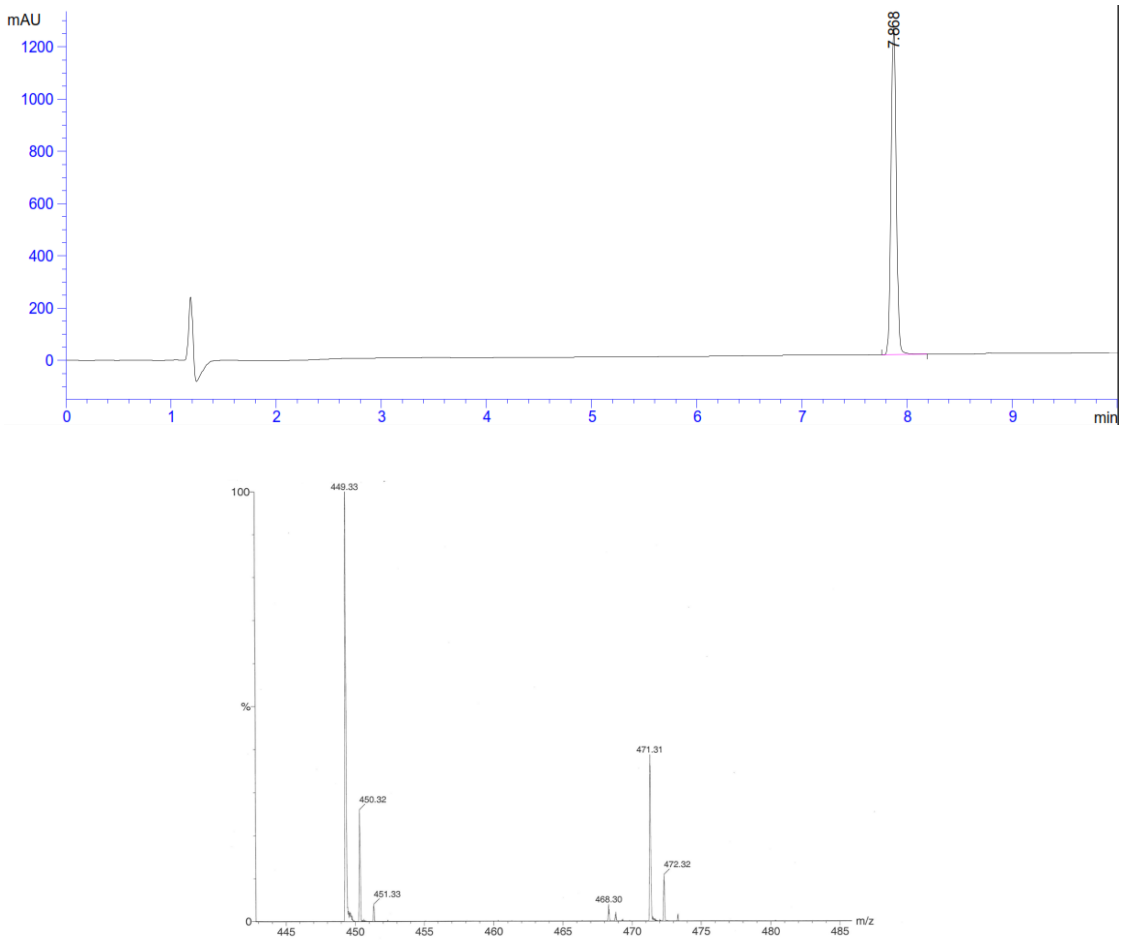
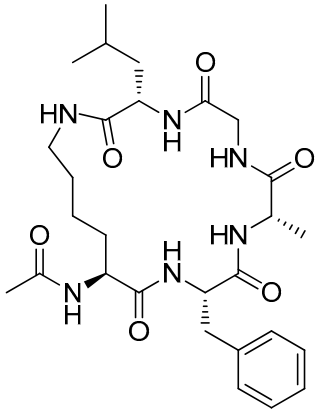


Figure S7. ESI-MS and analytical HPLC trace for Ac-FAGL-OH. (Kinetex™ LC Column 80 Å; 100 x 4.6 mm; particle size 2.6 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA**,** 0-100% CH_3CN over 20 minutes)

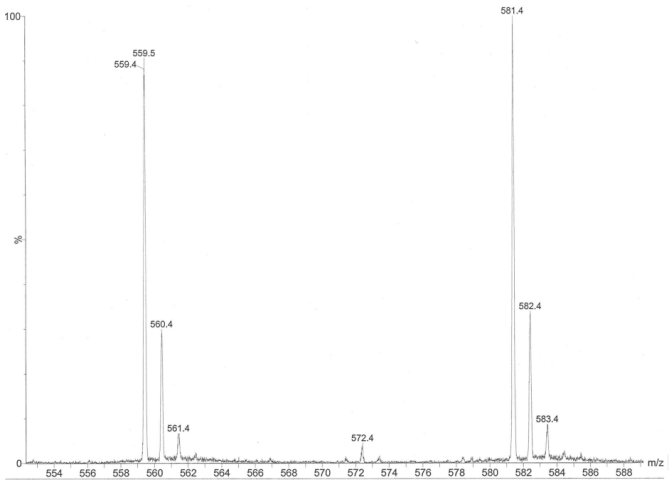
Synthesis of c-(KFAGL)

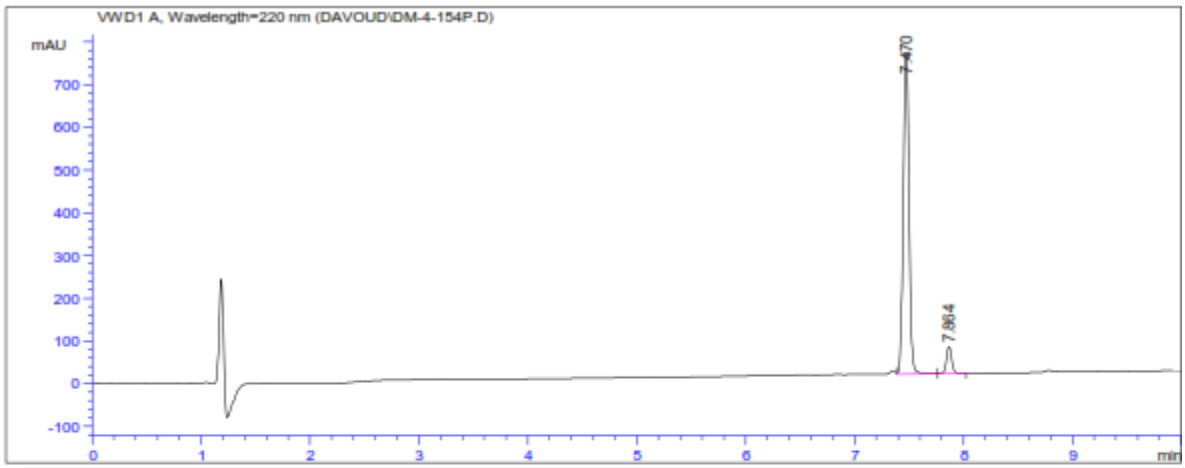
Crude Ac-LysFAGL-OH was synthesized according to the general peptide synthesis procedure on the wang resin. The crude peptide mixture was dissolved in 200mL DMF (1mM) and 5eq of HCTU and DIPEA was added to the mixture. The reaction was stirred for 24 hours after which the DMF was completely removed. The residue was dissolved in CH_3CN/water (1:1) and purified by HPLC.



Chemical Formula: $C_{28}H_{42}N_6O_6$
Exact Mass: 558.32

MS (ESI), m/z calcd for $[(C_{28}H_{42}N_6O_6 + H]^+$ = 559.32; found 559.40.






```
================================================================================
                        Area Percent Report
================================================================================

Sorted By             :      Signal
Multiplier:                   :      1.0000
Dilution:                     :      1.0000
Sample Amount:                :     20.00000  [ng/ul]    (not used in calc.)
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=220 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]  mAU   *s    [mAU   ]      %
----|-------|----|-------|----------|----------|--------|
   1   7.470 VV    0.0580 2781.83252  747.85638  93.0145
   2   7.864 VB    0.0541  208.91887   61.61217   6.9855
```

Figure S8. ESI-MS and analytical HPLC trace for c-(KFAGL). (Kinetex™ LC Column 80 Å; 100 x 4.6 mm; particle size 2.6 µm; solvent A: H_2O/0.1% TFA, solvent B: CH_3CN/0.1% TFA**,** 0-100% CH_3CN over 20 minutes)

600 MHz ^{1}H NMR spectrum of peptide **I** in 20% DMSO-d_6 in CD_2Cl_2

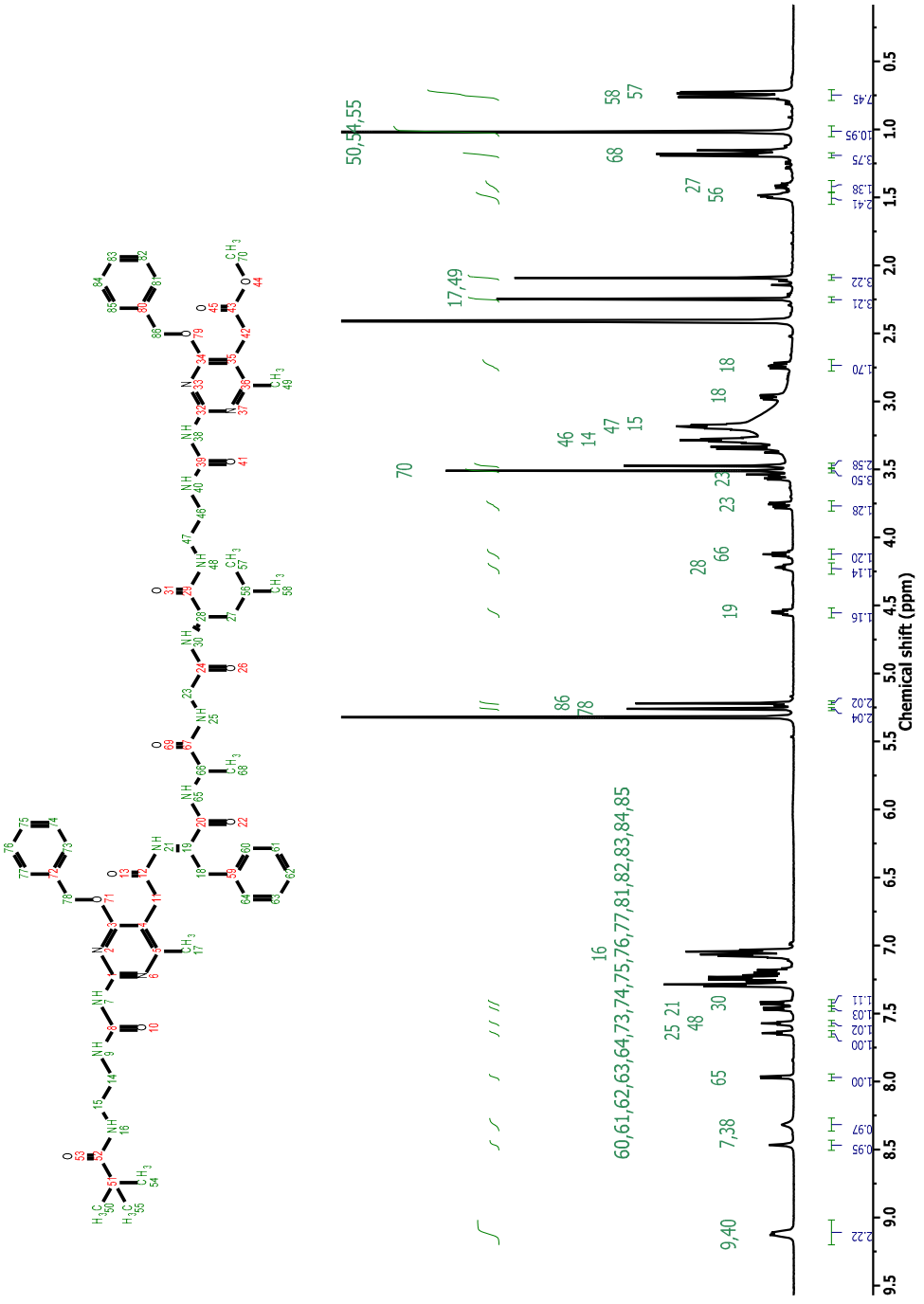


Figure S9. 600 MHz ^{1}H NMR spectrum of peptide **I** in 20% DMSO-d_6 in CD_2Cl_2.

600 MHz COSY spectrum of peptide I in 20% DMSO-d_6 in CD_2Cl_2

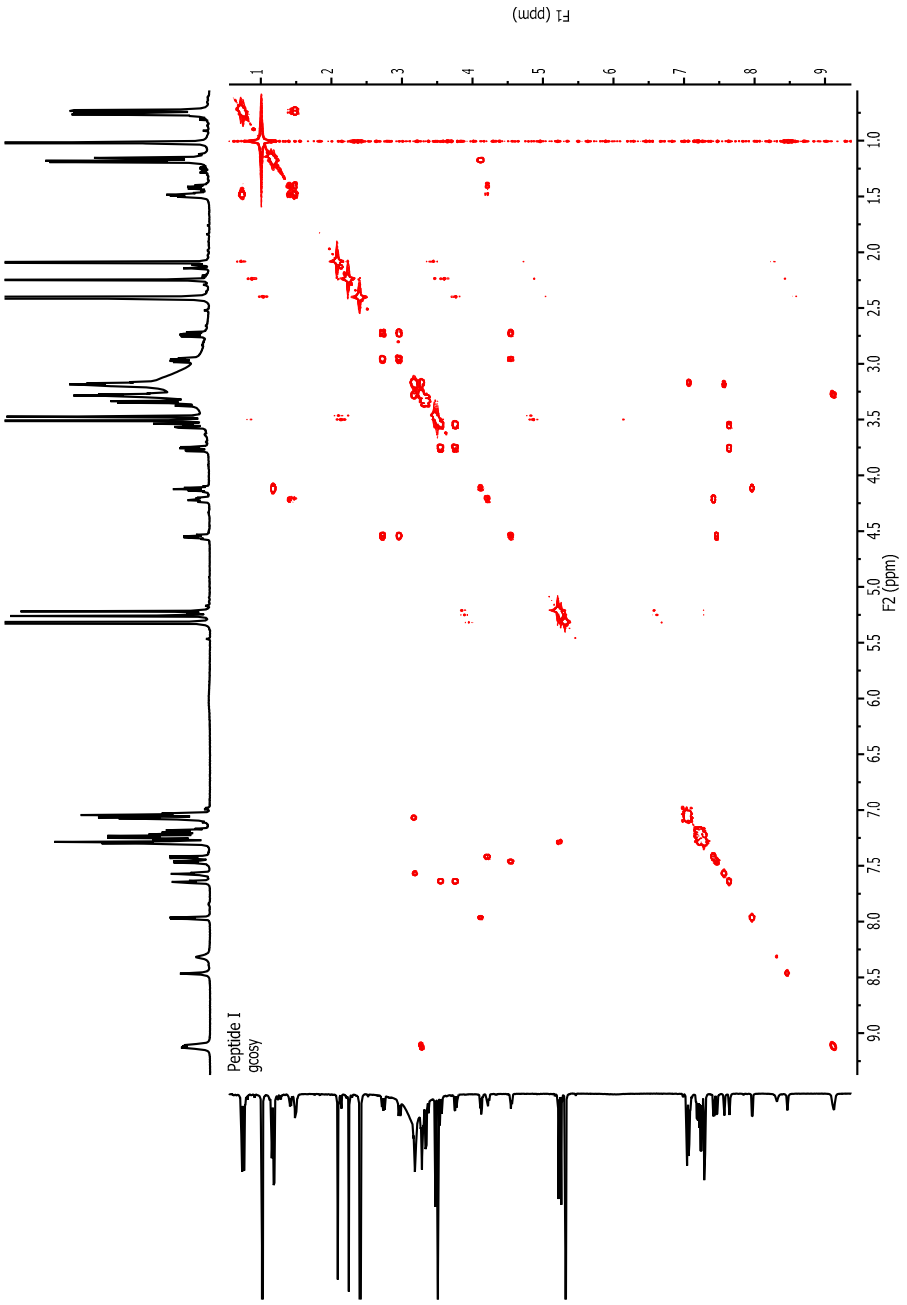


Figure S10. 600 MHz COSY spectrum of peptide **I** in 20% DMSO-d_6 in CD_2Cl_2.

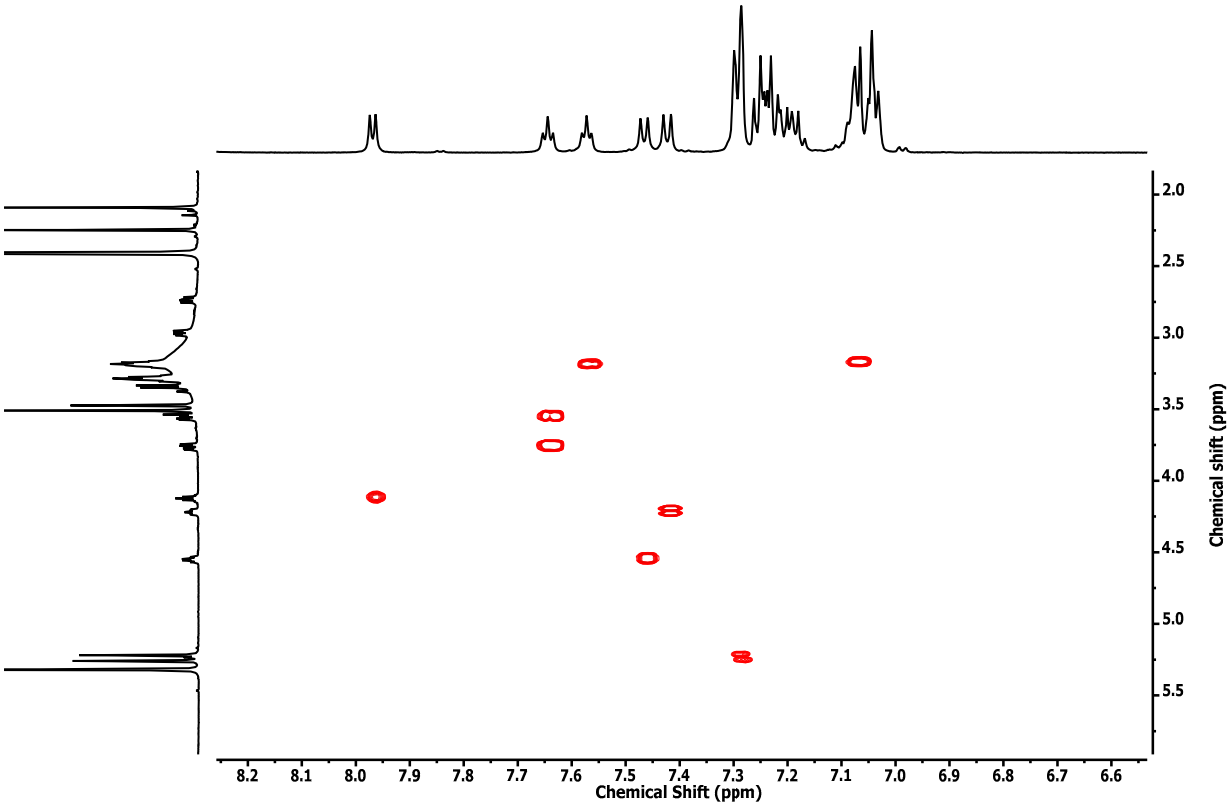


Figure S11. Cross-section of 600 MHz COSY spectrum of peptide **I** in 20% DMSO-d_6 in CD_2Cl_2 that is used to assign amide peaks.

500 MHz ^{1}H NMR spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2

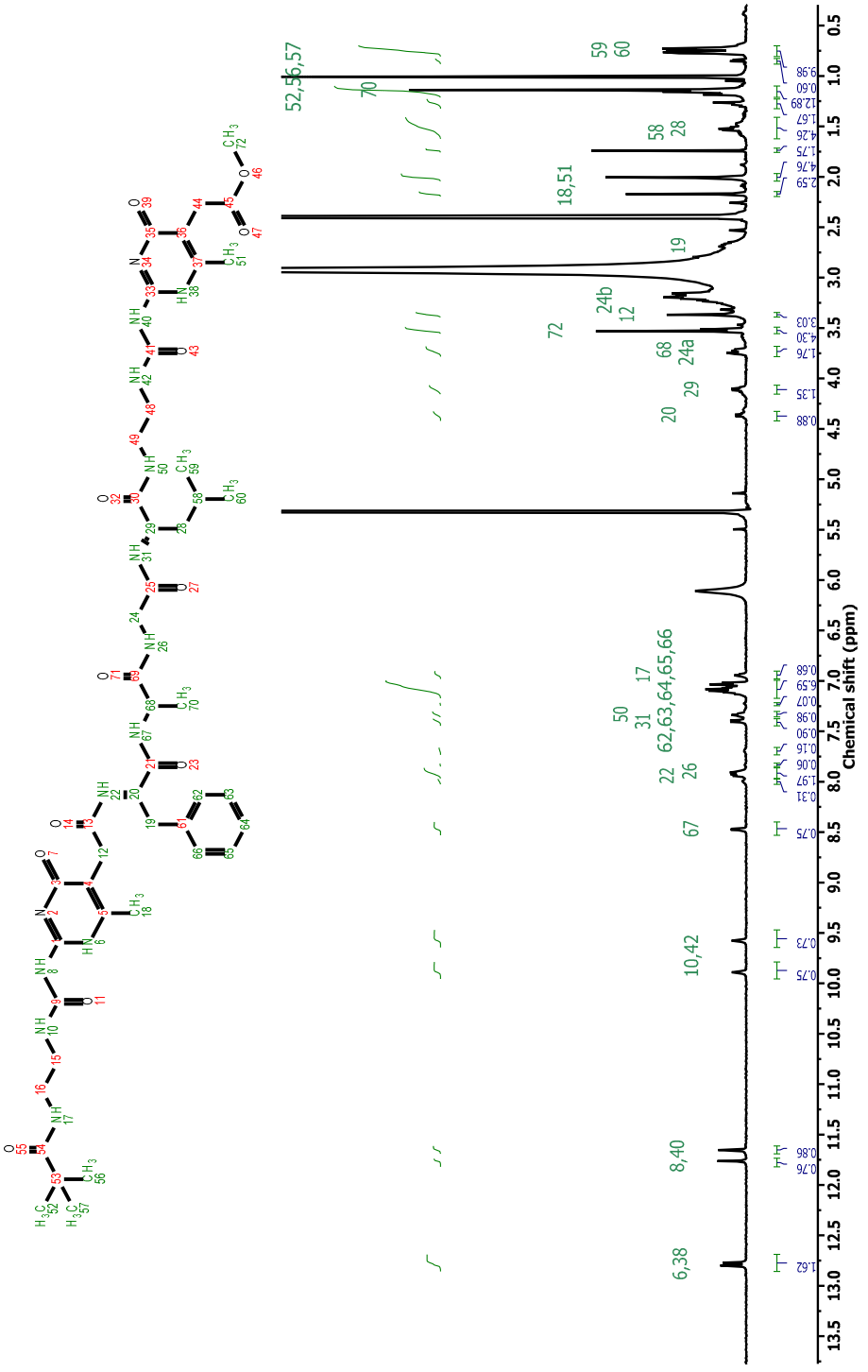


Figure S12. 500 MHz ^{1}H NMR spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz COSY spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2

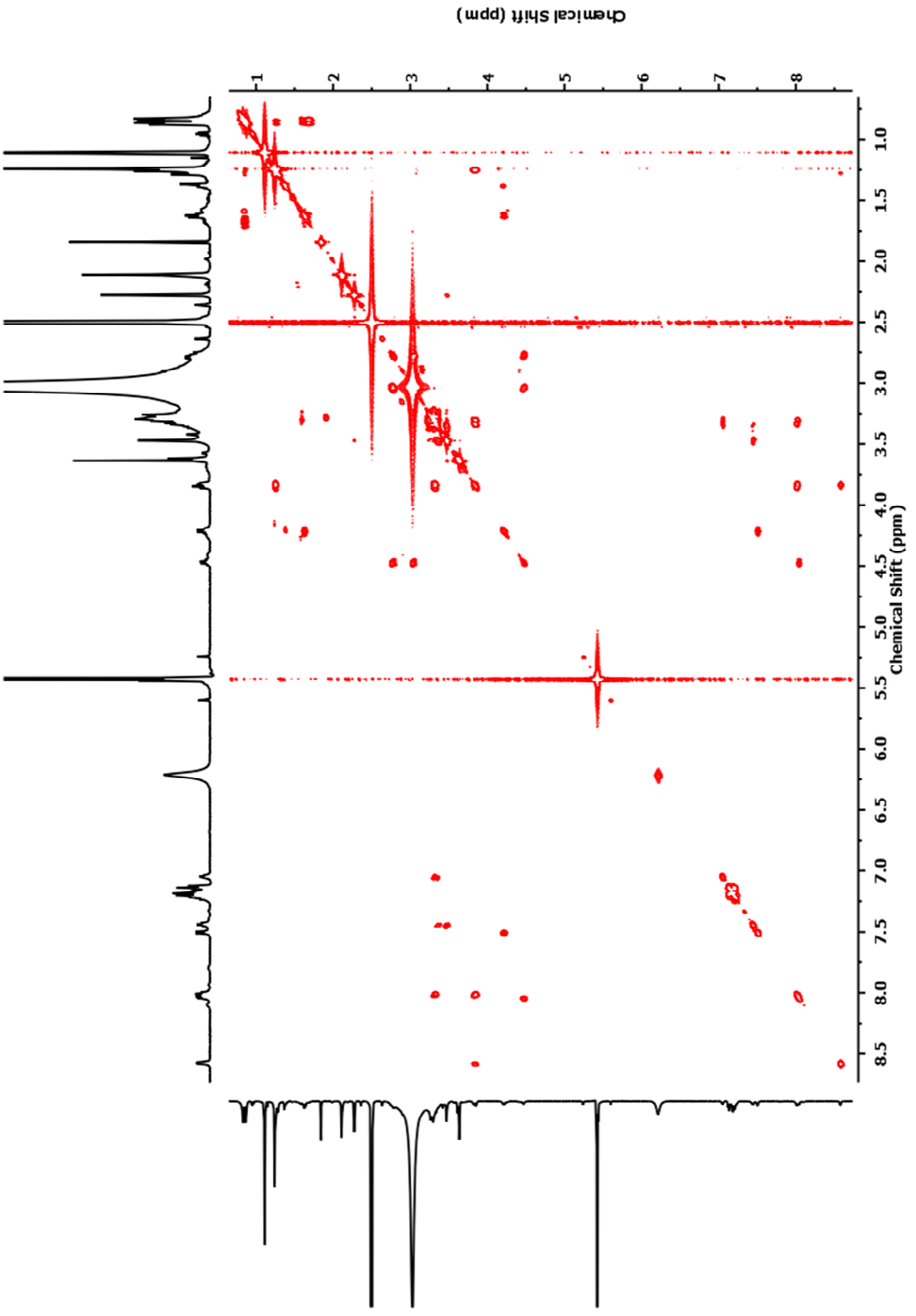


Figure S13. 500 MHz COSY spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz TOCSY spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2

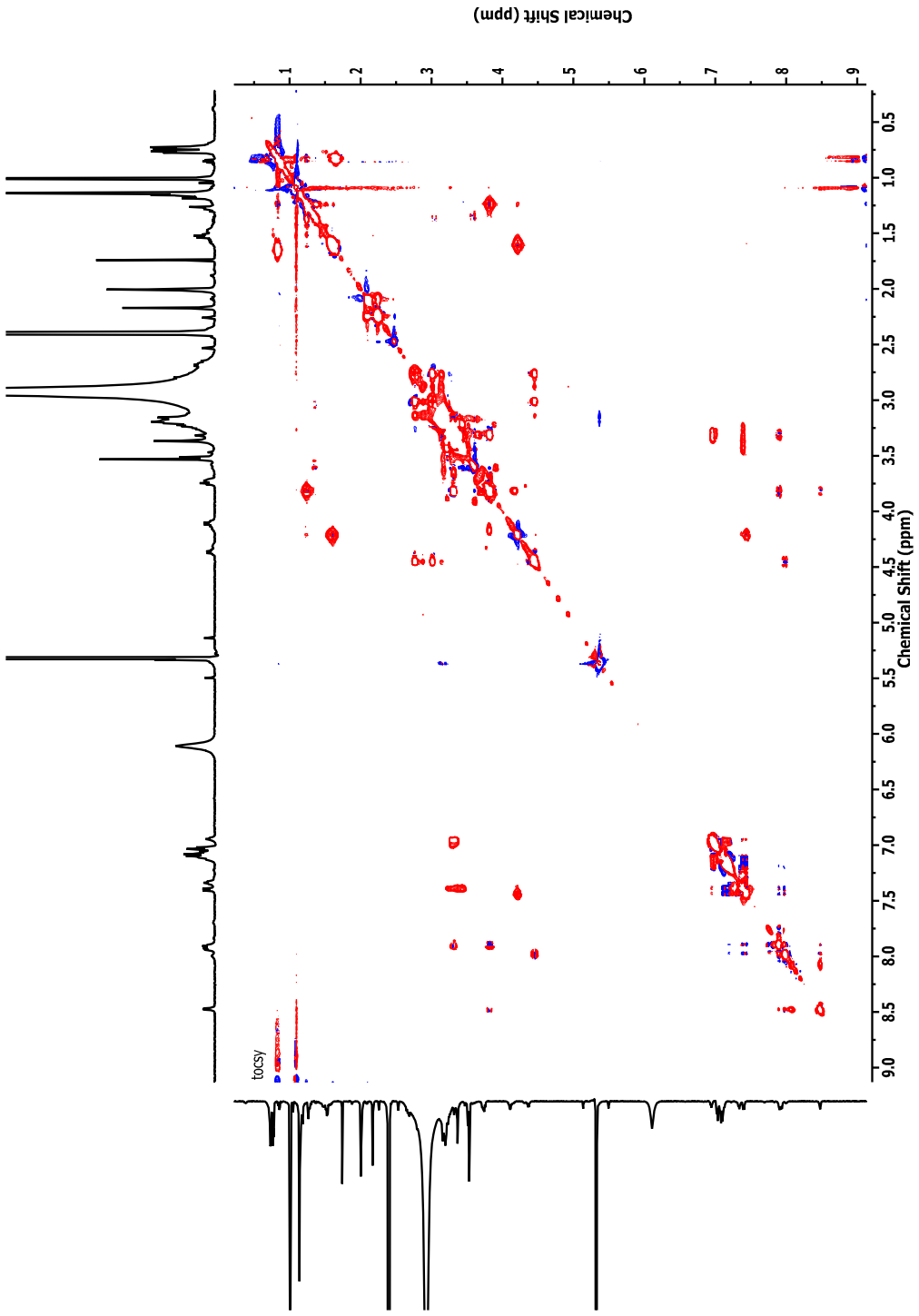


Figure S14. 600 MHz TOCSY spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2.

600 MHz ^{1}H NMR spectrum of peptide III in 20% DMSO-d_6 in CD_2Cl_2

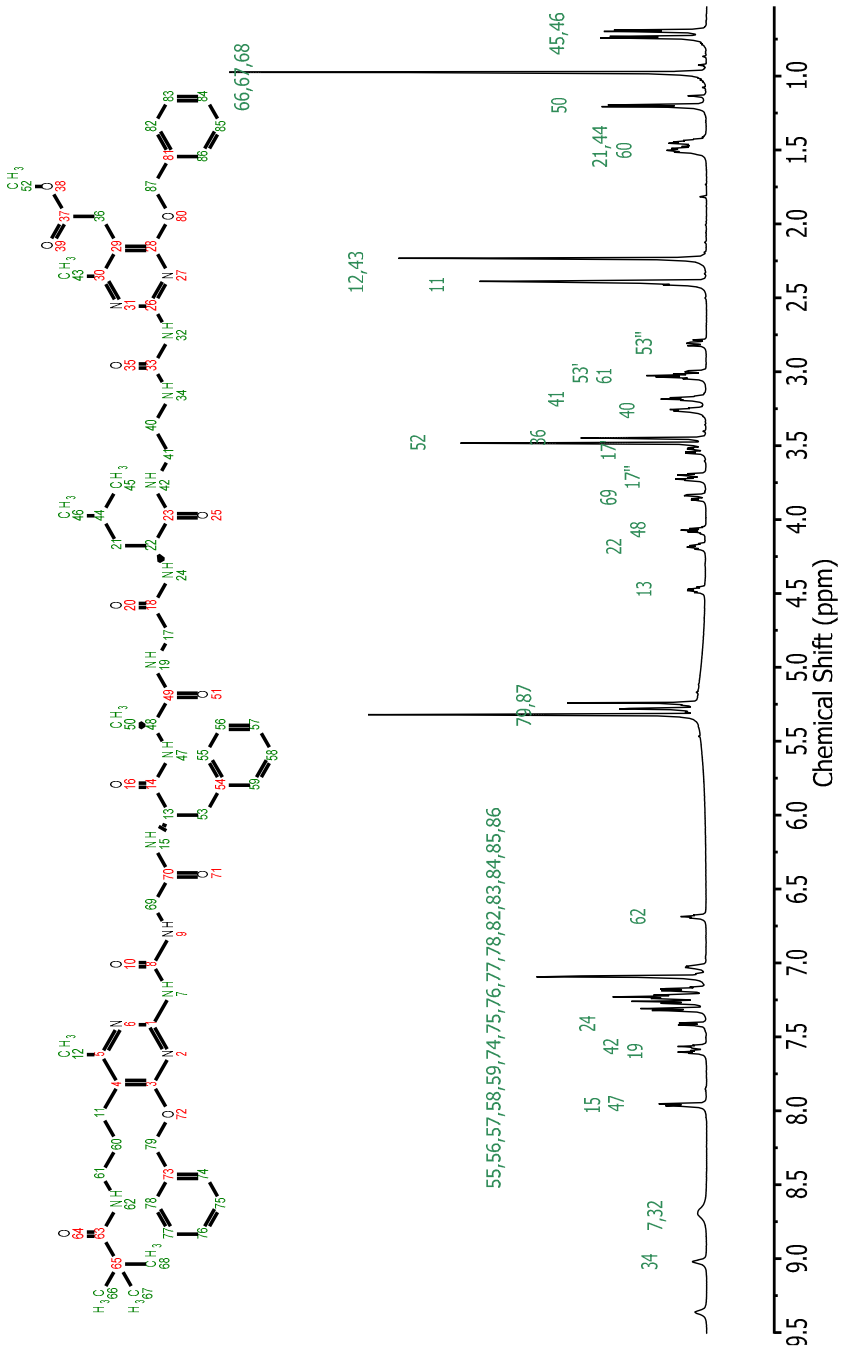


Figure S15. 600 MHz ^{1}H NMR spectrum of peptide III in 20% DMSO-d_6 in CD_2Cl_2.

600 MHz COSY spectrum of peptide III in 20% DMSO-d_6 in CD_2Cl_2

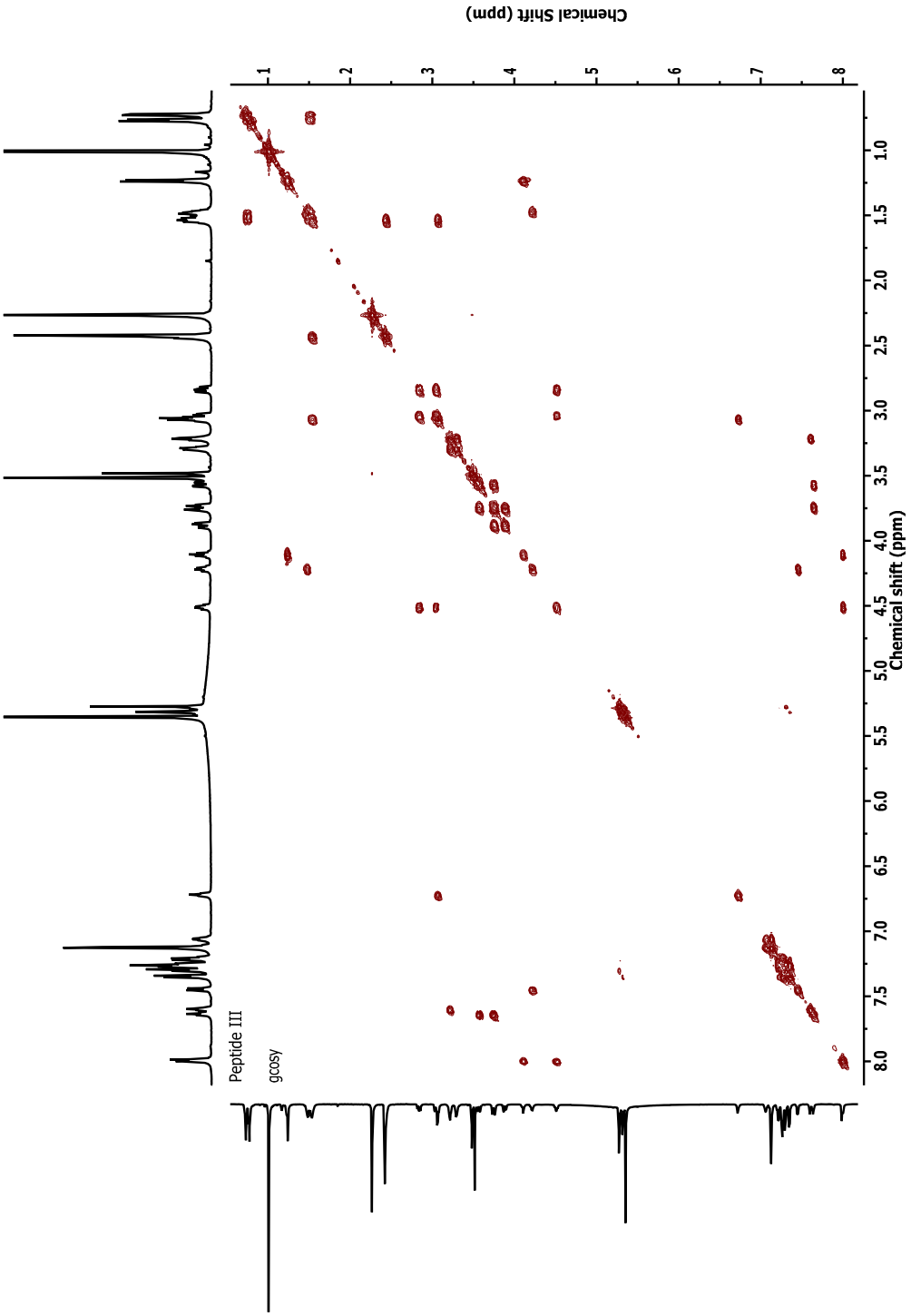


Figure S16. 600 MHz COSY spectrum of peptide **III** in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz ^{1}H NMR spectrum of peptide IV in 20% DMSO-d_6 in CD_2Cl_2

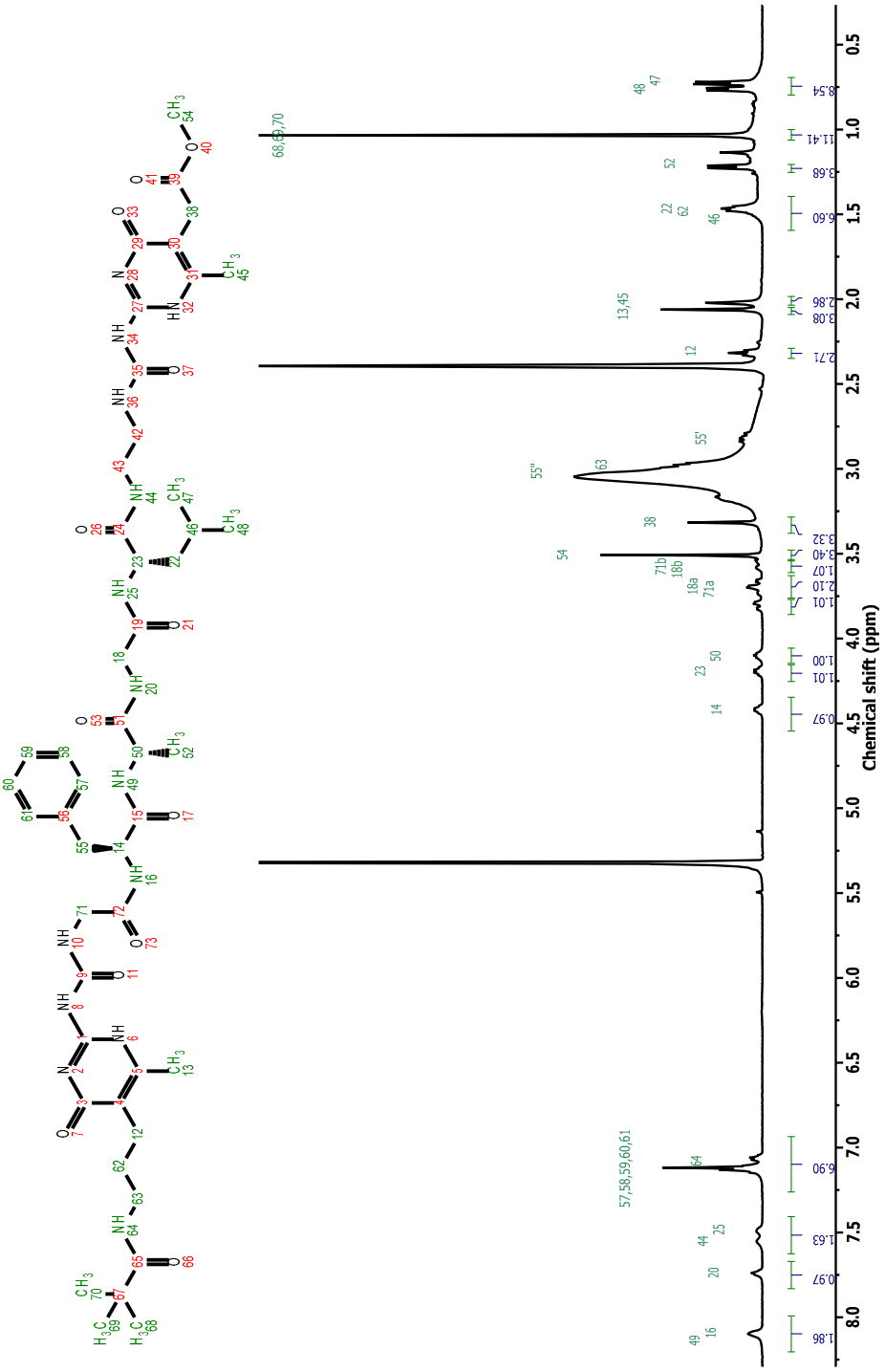


Figure S17. 500 MHz ^{1}H NMR spectrum of peptide **IV** in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz COSY spectrum of peptide IV in 20% DMSO-d_6 in CD_2Cl_2

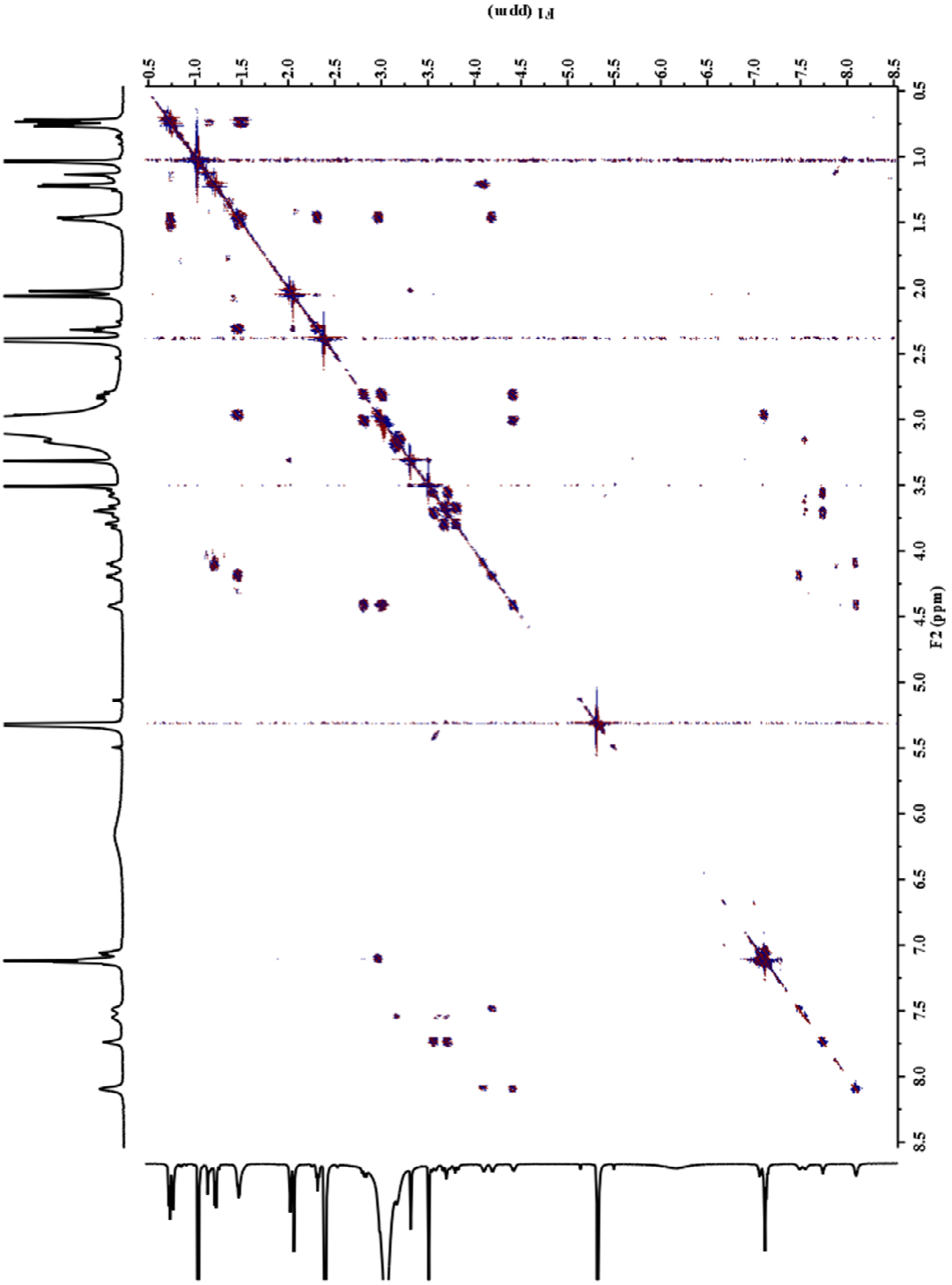


Figure S18. 500 MHz COSY spectrum of peptide **IV** in 20% DMSO-d_6 in CD_2Cl_2.

Comparison of peptide I and II ^{1}H NMR spectra in 20% DMSO-d_6 in CD_2Cl_2 (folding conformer)

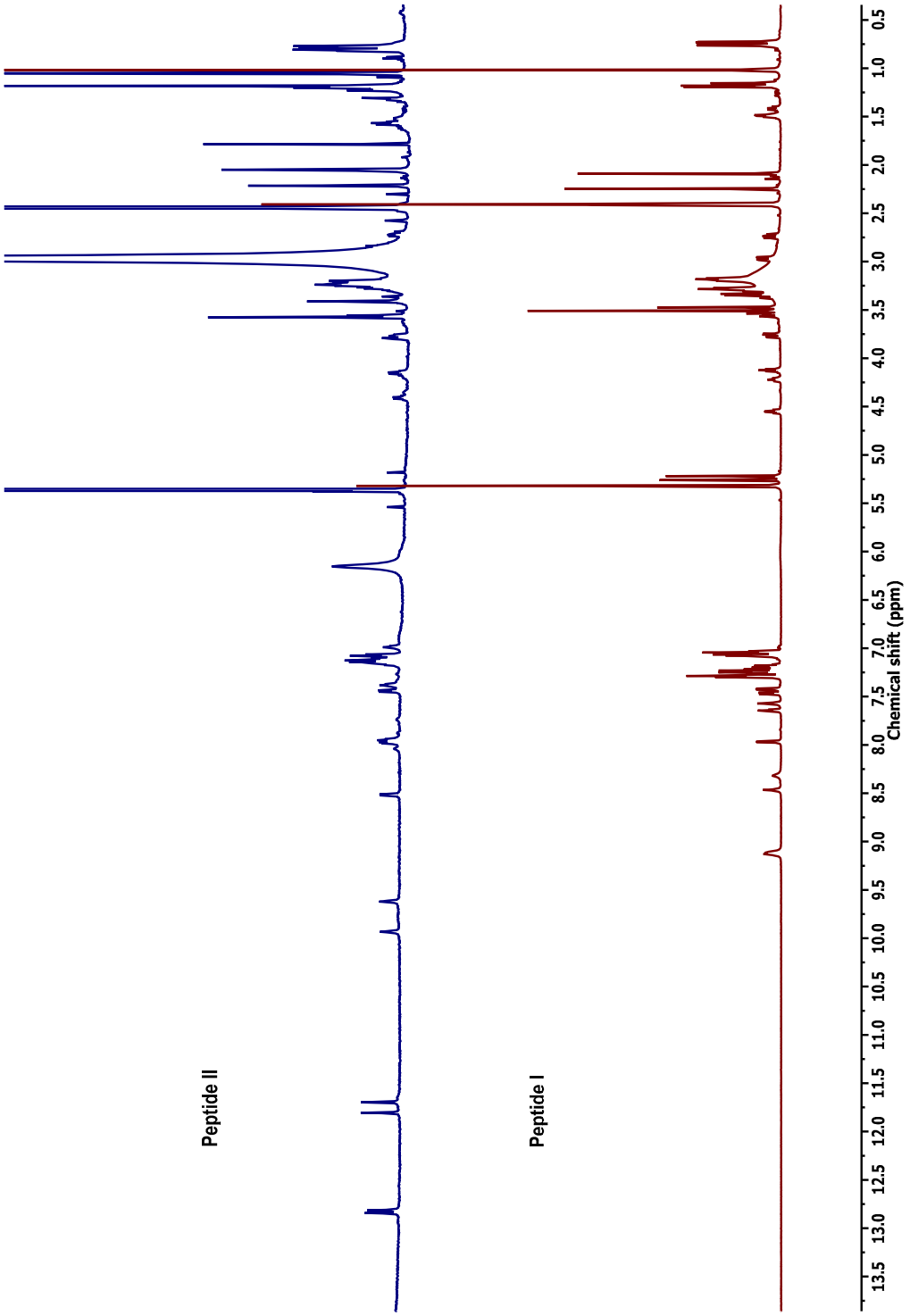


Figure S19. Comparison of ^{1}H NMR spectrum of peptide **I** and **II** in 20% DMSO-d_6 in CD_2Cl_2 (folding conformer).

Comparison of peptide III and IV ^{1}H NMR spectra in 20% DMSO-d_6 in CD_2Cl_2 (nonfolding conformer)

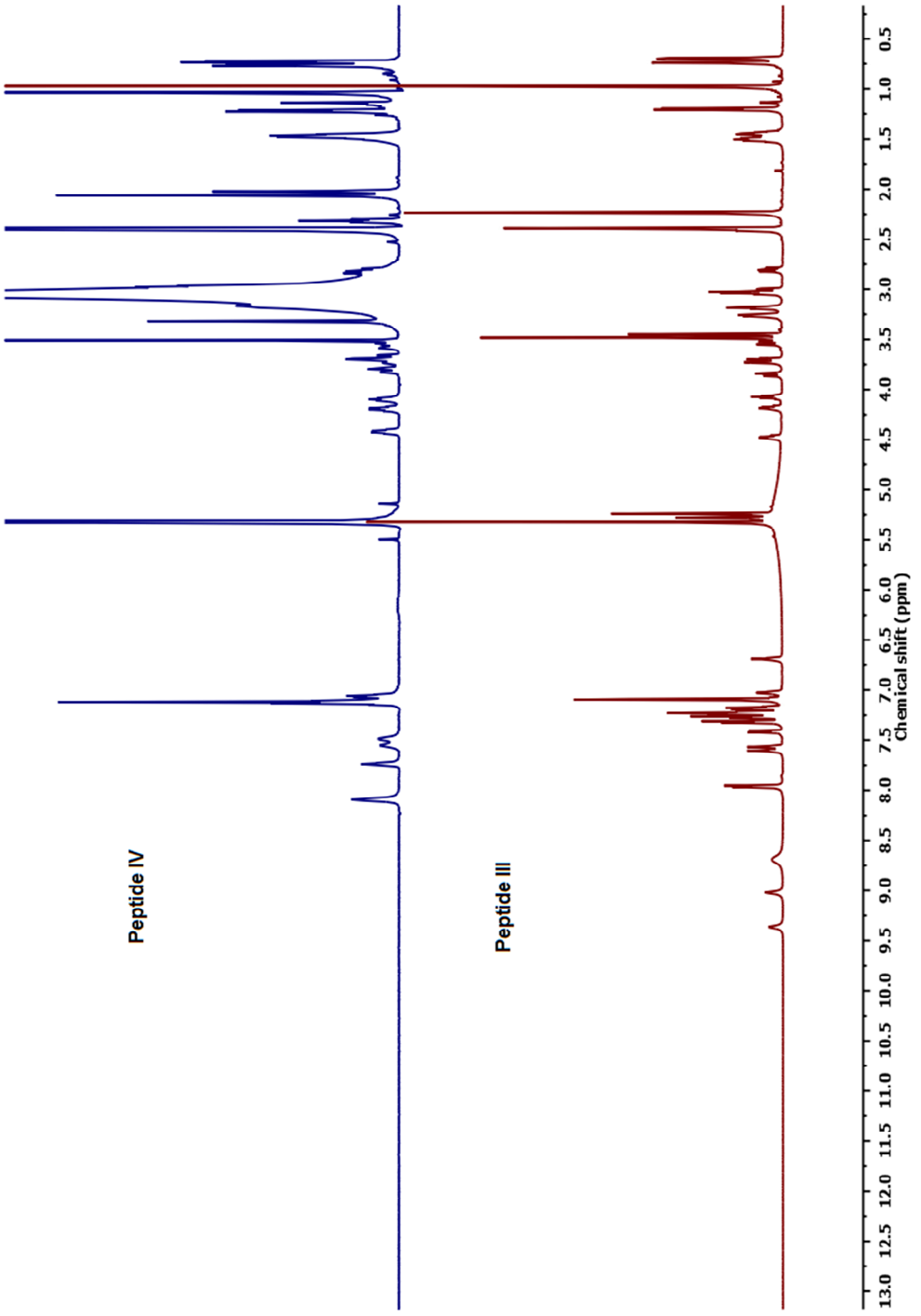
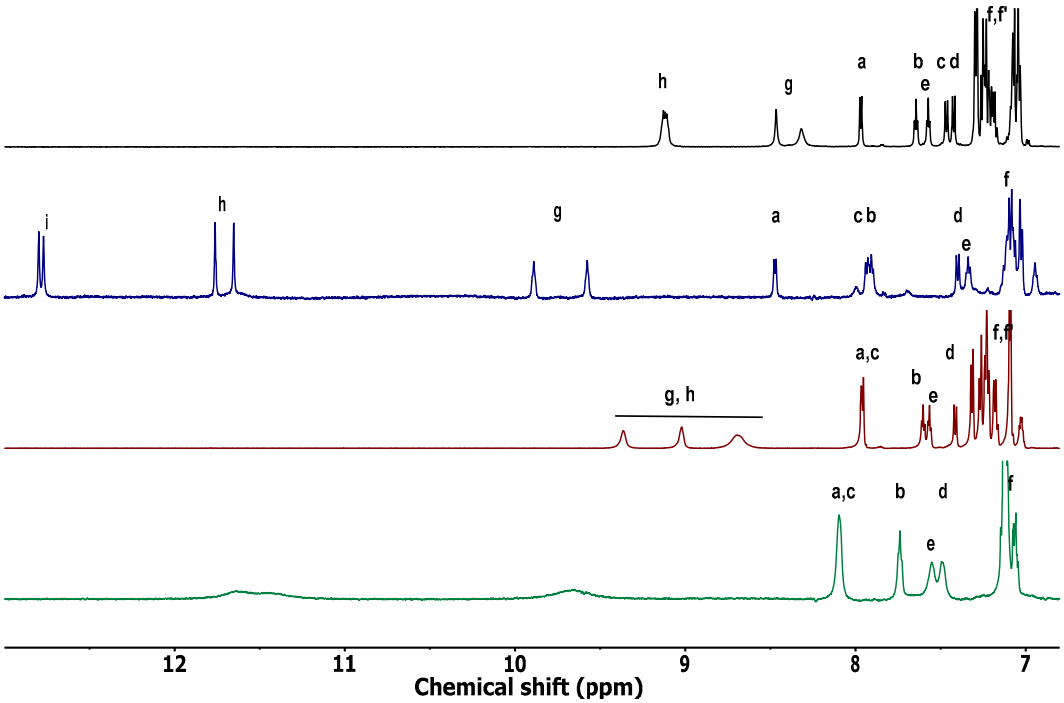


Figure S20. Comparison of ^{1}H NMR spectrum of peptide **III** and **IV** in 20% DMSO-d_6 in CD_2Cl_2 (nonfolding conformer).

Comparison of ^{1}H NMR spectra of peptide I-IV (full assignment)



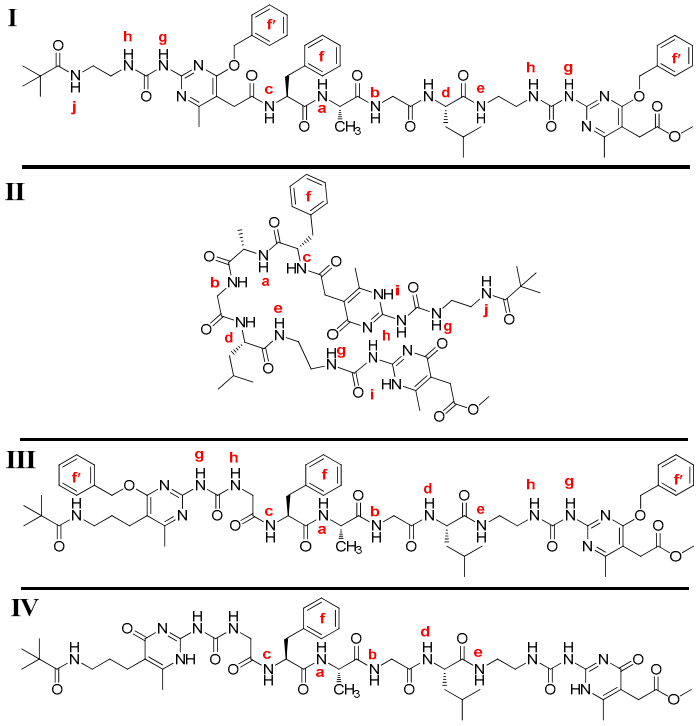


Figure S21. Comparison of ^{1}H NMR spectra of peptide **I-IV.** For discussion, refer to Figure 2 and its explanation.

500 MHz ^{1}H NMR spectrum of Ac-FAGL-OH in 20% DMSO-d_6 in CD_2Cl_2

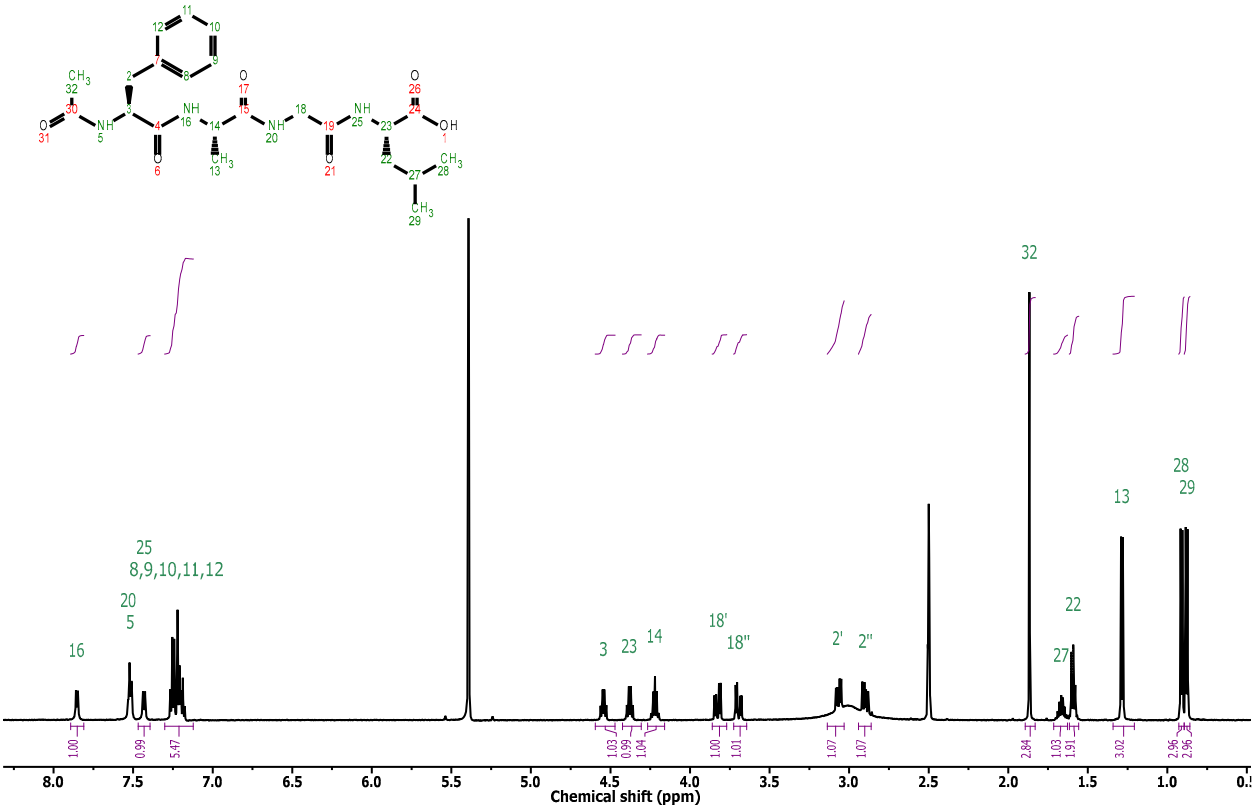


Figure S22. 500 MHz ^{1}H NMR spectrum of Ac-FAGL-OH in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz COSY spectrum of Ac-FAGL-OH in 20% DMSO-d_6 in CD_2Cl_2

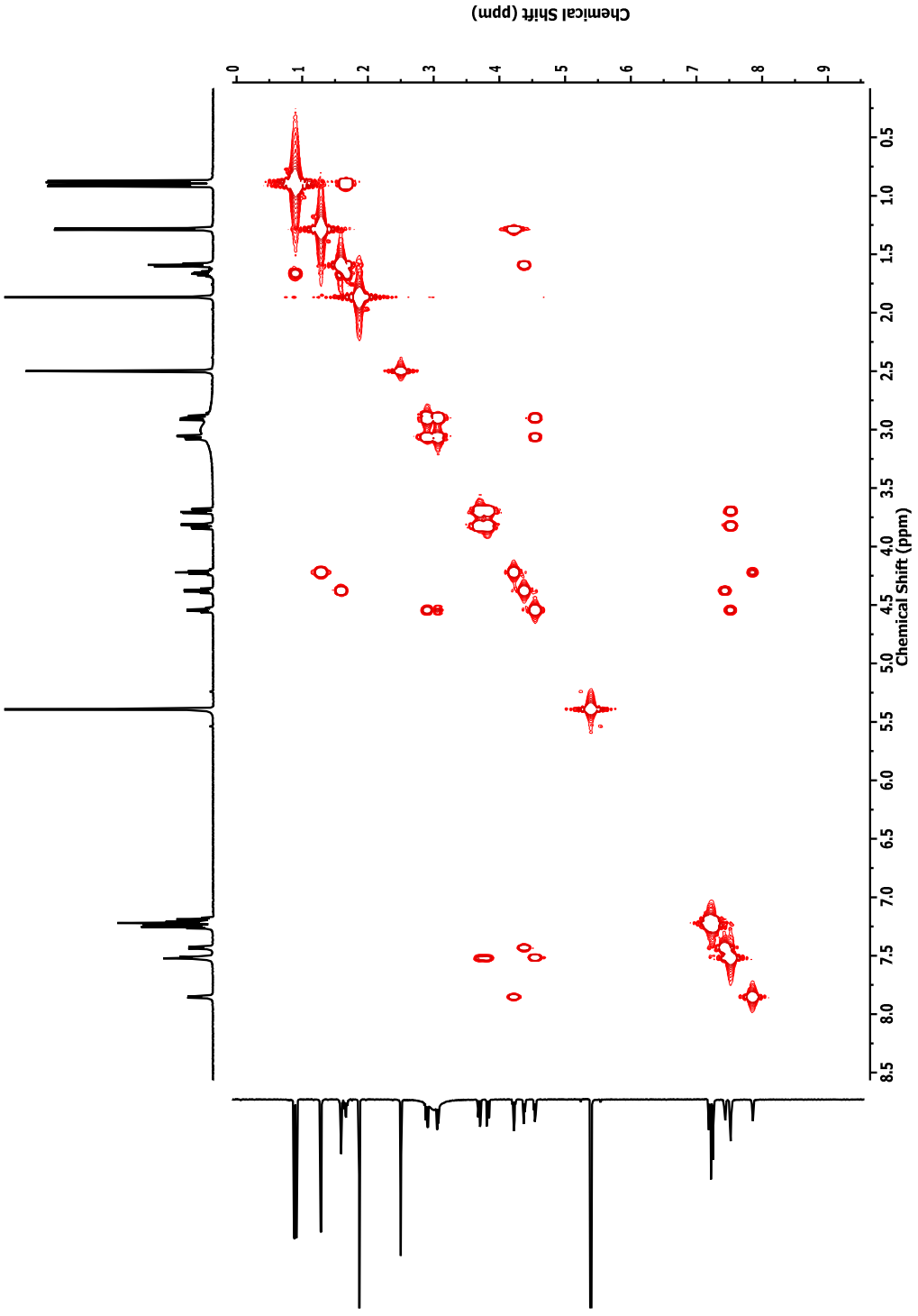


Figure S23. 500 MHz COSY spectrum of Ac-FAGL-OH in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz ^{1}H NMR spectrum of c-(KFAGL) in 20% DMSO-d_6 in CD_2Cl_2

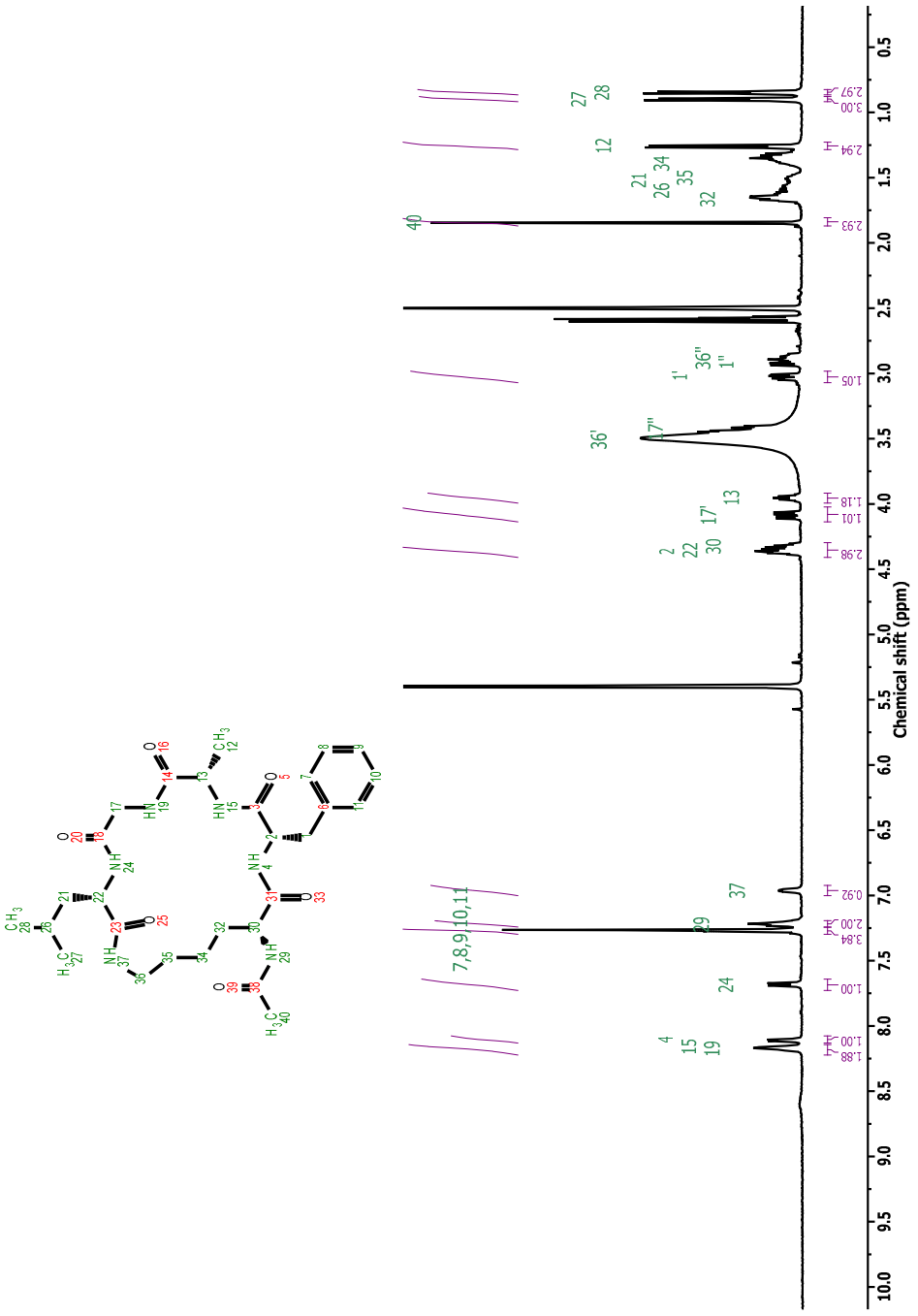


Figure S24. 500 MHz ^{1}H NMR c-(KFAGL) in 20% DMSO-d_6 in CD_2Cl_2.

500 MHz COSY spectrum of c-(KFAGL) in 20% DMSO-d_6 in CD_2Cl_2

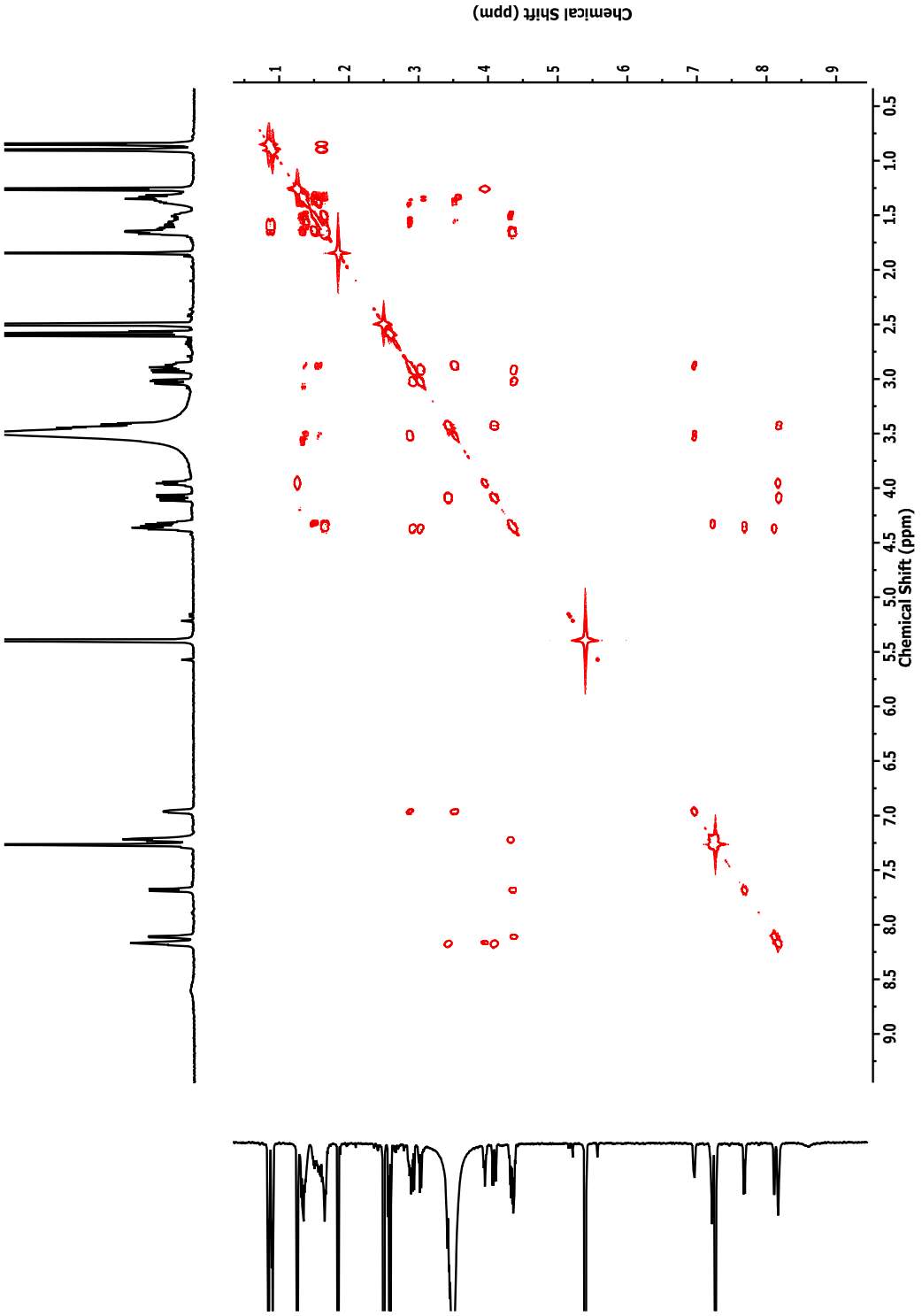


Figure S25. 500 MHz COSY spectrum of c-(KFAGL) in 20% DMSO-d_6 in CD_2Cl_2.

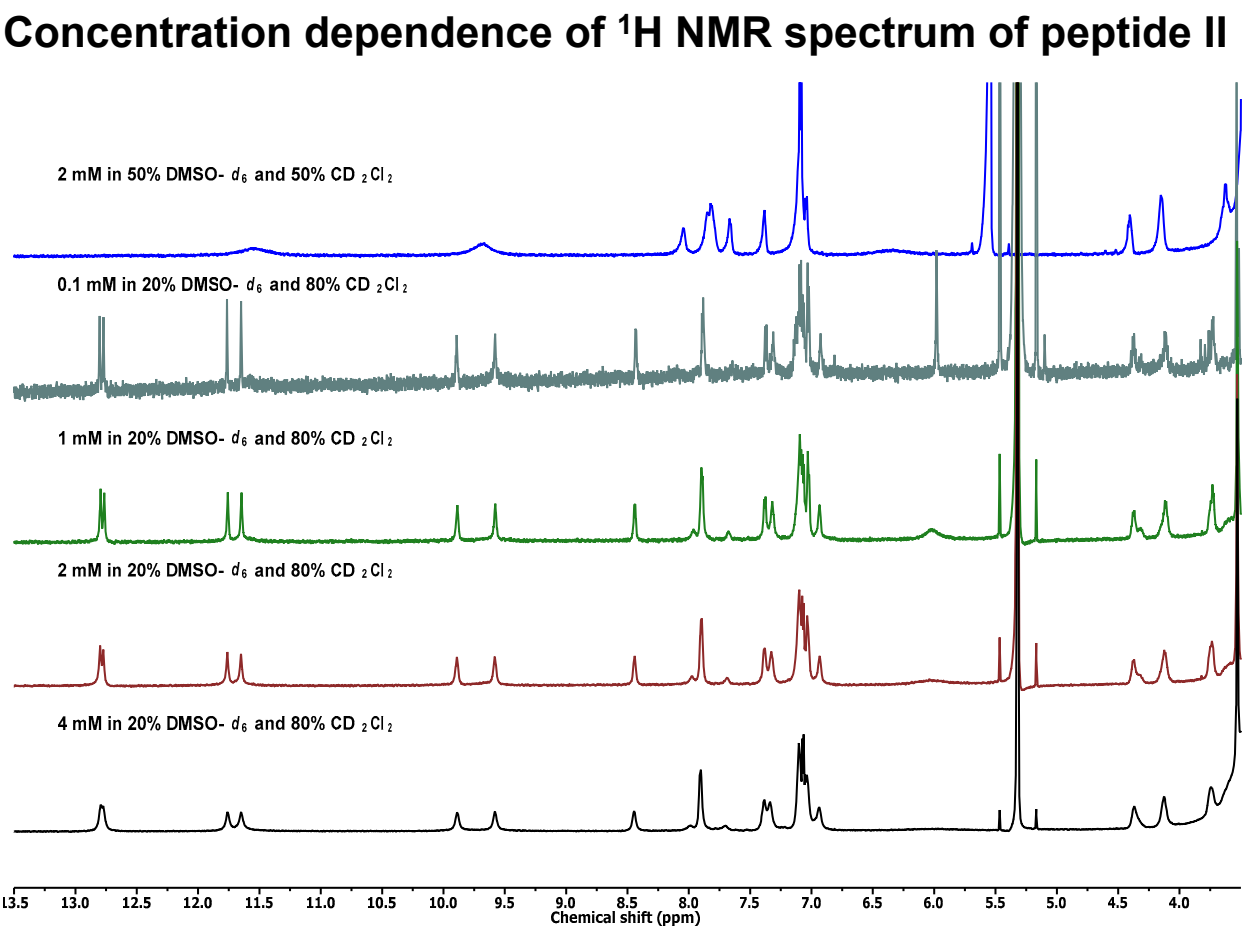


Figure S26. Concentration dependence of ^{1}H NMR spectrum of peptide **II** and the effect of DMSO on the peptide structure and spectrum. No change is observed over the range of 0.1-4 mM, however, addition of more DMSO perturbs the hydrogen bonding of the UPy units and denatures the peptide.

Plot of ∆δHα values for c-(KFAGL)

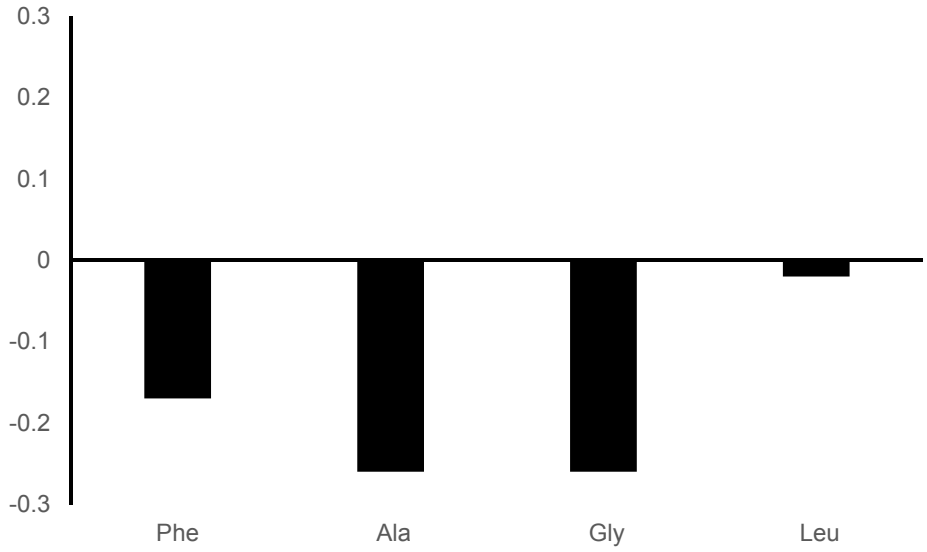
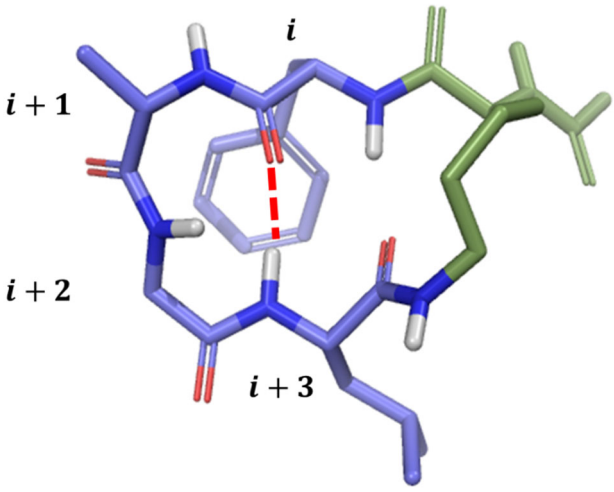


Figure S27. Plot of ∆δHα values for c-(KFAGL). $\Delta\delta H^{\alpha}$ [δH^{α} (cyclic) - δH^{α} (linear)]. In case of glycine, chemical shift for low field methylene proton has been plotted. c-(KFAGL) shows negative $\Delta\delta H^{\alpha}$ values consistent with a pattern of type II β-turn structure.

Energy minimized model of c-(KFAGL)

Energy minimization using AMBER* forcefield with $CHCl_3$, GB/SA implied solvent parameter, was performed using MacroModel utility after building the model in Schrödinger Maestro suite. A quick search of conformational space is performed through 2000 steps Monte Carlo conformational search.

Figure S28. The energy minimized model of c-(KFAGL) demonstrates features of a type II β-turn. The glycine phi torsional angle, 158, is deviated from ideal type II β-turn.



Summary of 1H NMR spectroscopic features of Ac-FAGL-OH and c-(KFAGL)

Table S1. Chemical Shift and Temperature dependencies of the ^{1}H NMR chemical shifts (ppb/K) of various NH protons for Ac-FAGL-OH (4 mM in a mixture of 20% DMSO-d_6 and 80% CD_2Cl_2)[a]

	Phe	**Ala**	**Gly**	**Leu**
N_H (ppm)	7.51	7.85	7.53	7.43
Hα (ppm)	4.54	4.22	3.83, 3.ff	4.38
Hβ (γ, δ)	3.07, 2.90	1.87	N/A	1.59 (1.66, 0.91,0.88)
Δ $\delta N_H/\Delta T$ (ppb/K)	-8.6	-7.6	-4.0	-5.5

[a] ^{1}H NMR spectra recorded at 5 °C intervals from -5 °C to 35 °C.

Table S2. Chemical Shift and Temperature dependencies of the ^{1}H NMR chemical shifts (ppb/K) of various NH protons for c-(KFAGL) (4 mM in a mixture of 20% DMSO-d_6 and 80% CD_2Cl_2)[a]

	Lys	**Phe**	**Ala**	**Gly**	**Leu**
N_H (ε N_H)	7.22 (6.96)	8.11	8.16	8.18	7.69
α	4.33	4.37	3.96	4.09, 3.43	4.36
β (γ, δ)	1.66	3.02, 2.91	1.26	N/A	(1.66,0.9,0.85)
Δ $\delta N_H/\Delta T$ (ppb/K)	-5.6 (-2.5)	-8.28	-7.72	-5.57	-3.8

Variable Temperature 1H NMR spectrum of peptide I in 20% DMSO-d6 in CD_2Cl_2

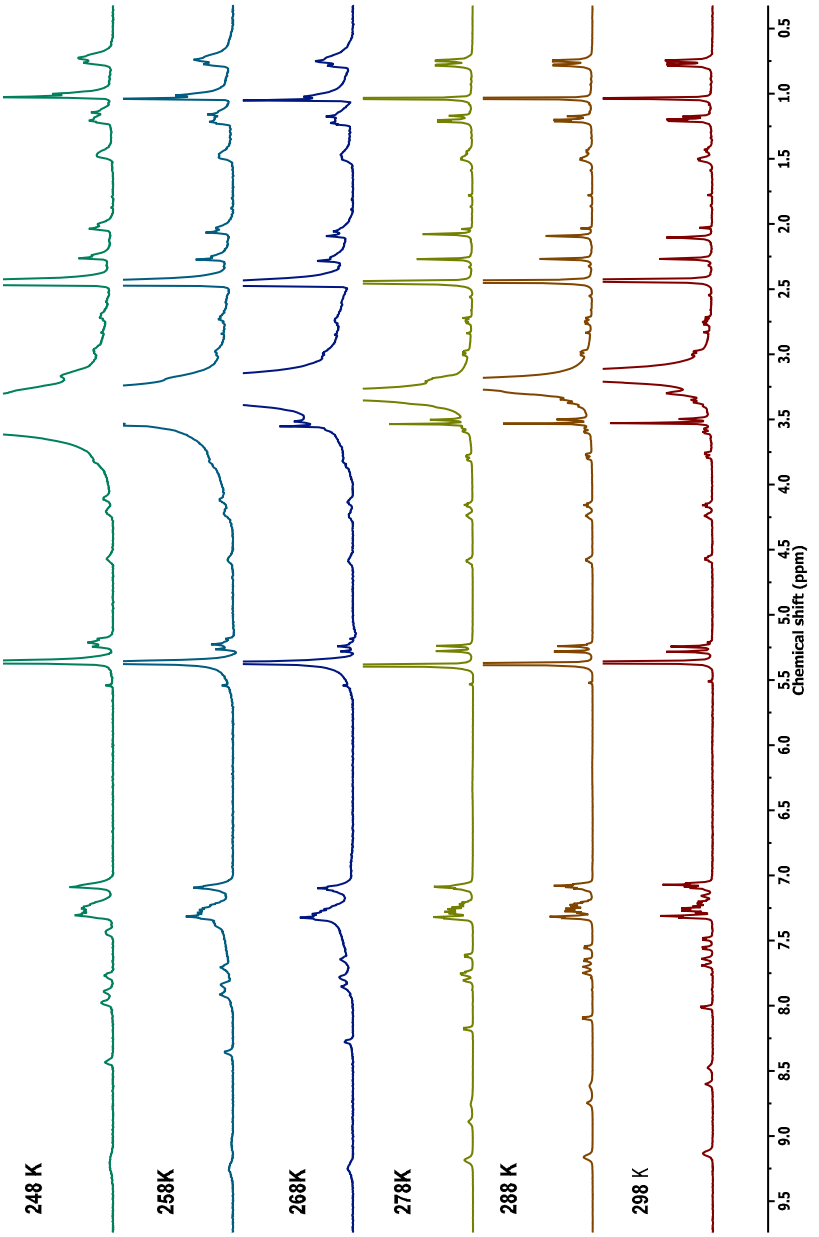


Figure S29. Variable Temperature ^{1}H NMR spectrum of peptide **I** in 20% DMSO-d_6 in CD_2Cl_2. The first three spectra were recorded on a GN500 machine capable of reaching lower temperature to ensure linearity of amide-temperature response.

Variable Temperature ^{1}H NMR spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2

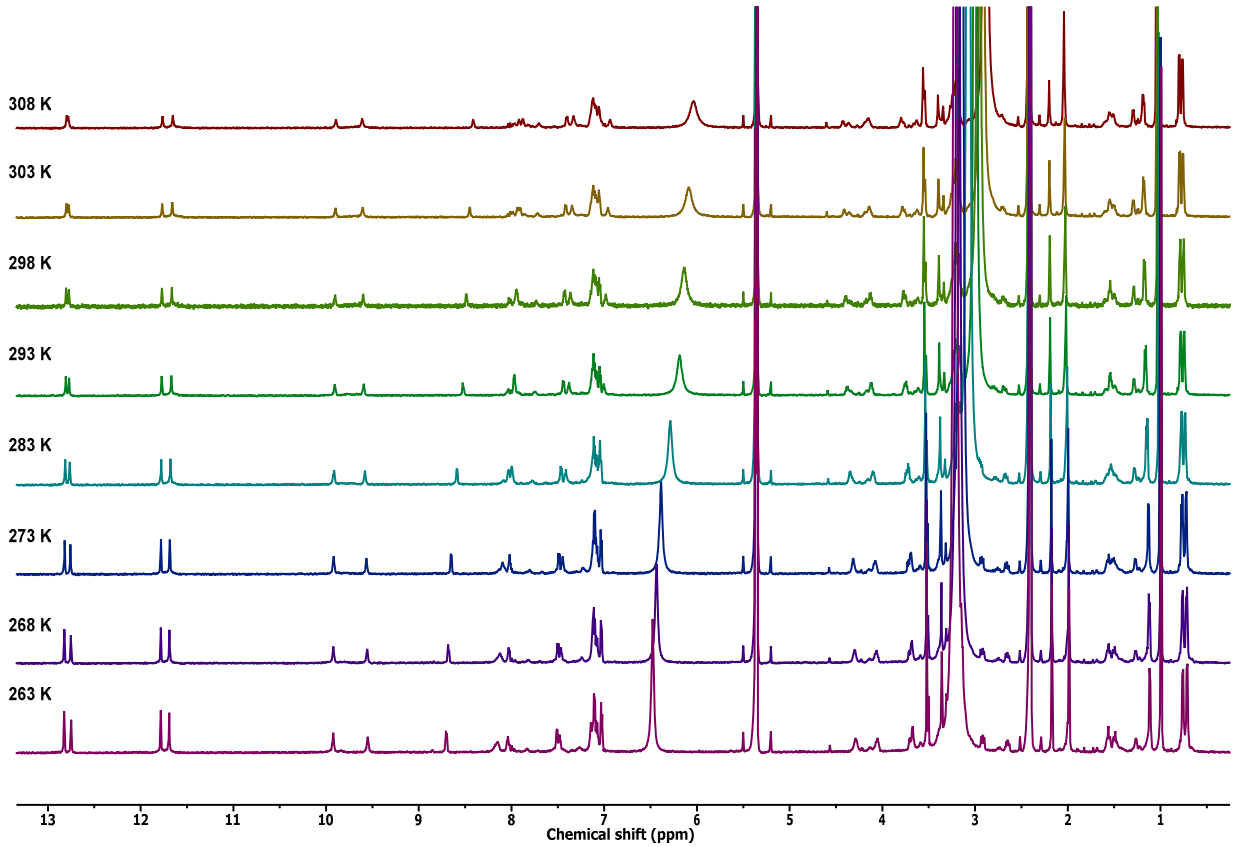


Figure S30. Variable Temperature ^{1}H NMR spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2.

Variable Temperature ^{1}H NMR spectrum of peptide III in 20% DMSO-d_6 in CD_2Cl_2

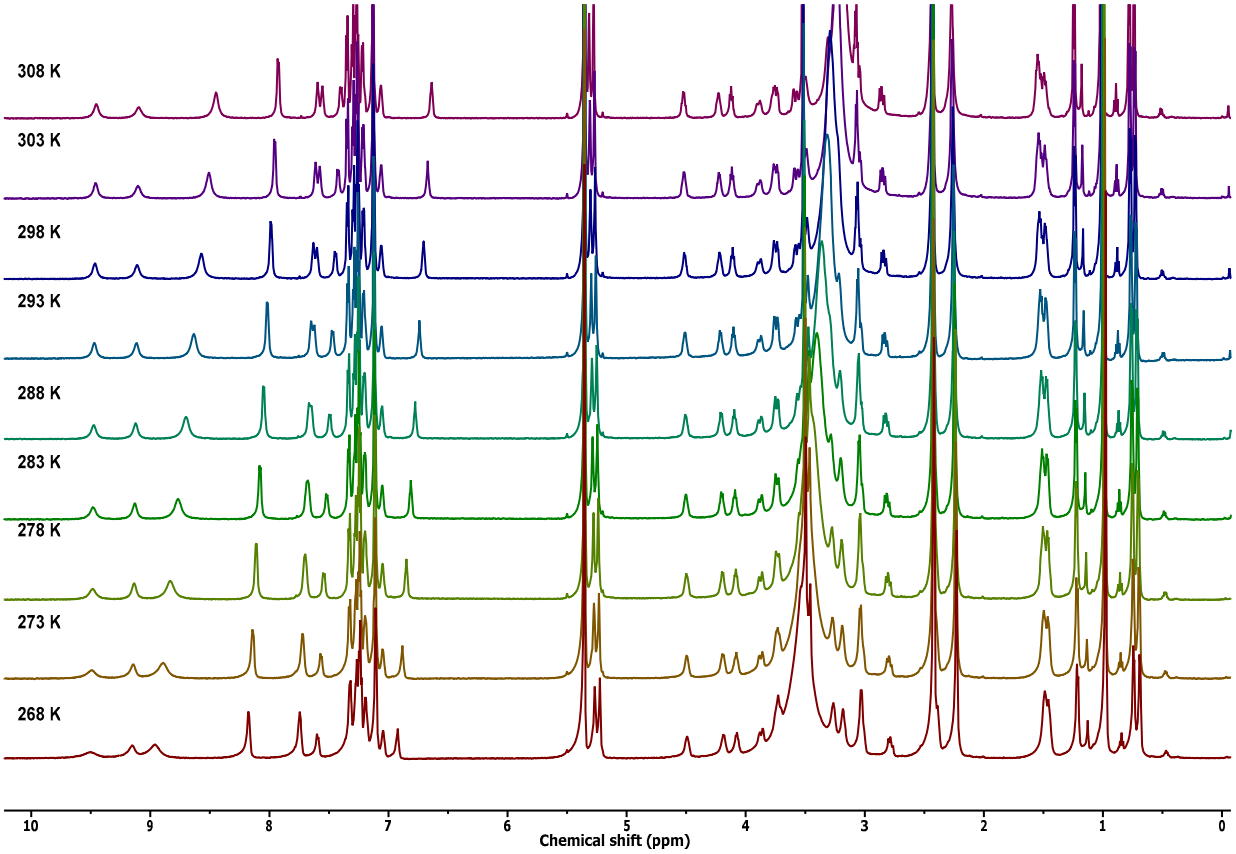


Figure S31. Variable Temperature ^{1}H NMR spectrum of peptide **III** in 20% DMSO-d_6 in CD_2Cl_2.

Variable Temperature ^{1}H NMR spectrum of peptide IV in 20% DMSO-d_6 in CD_2Cl_2

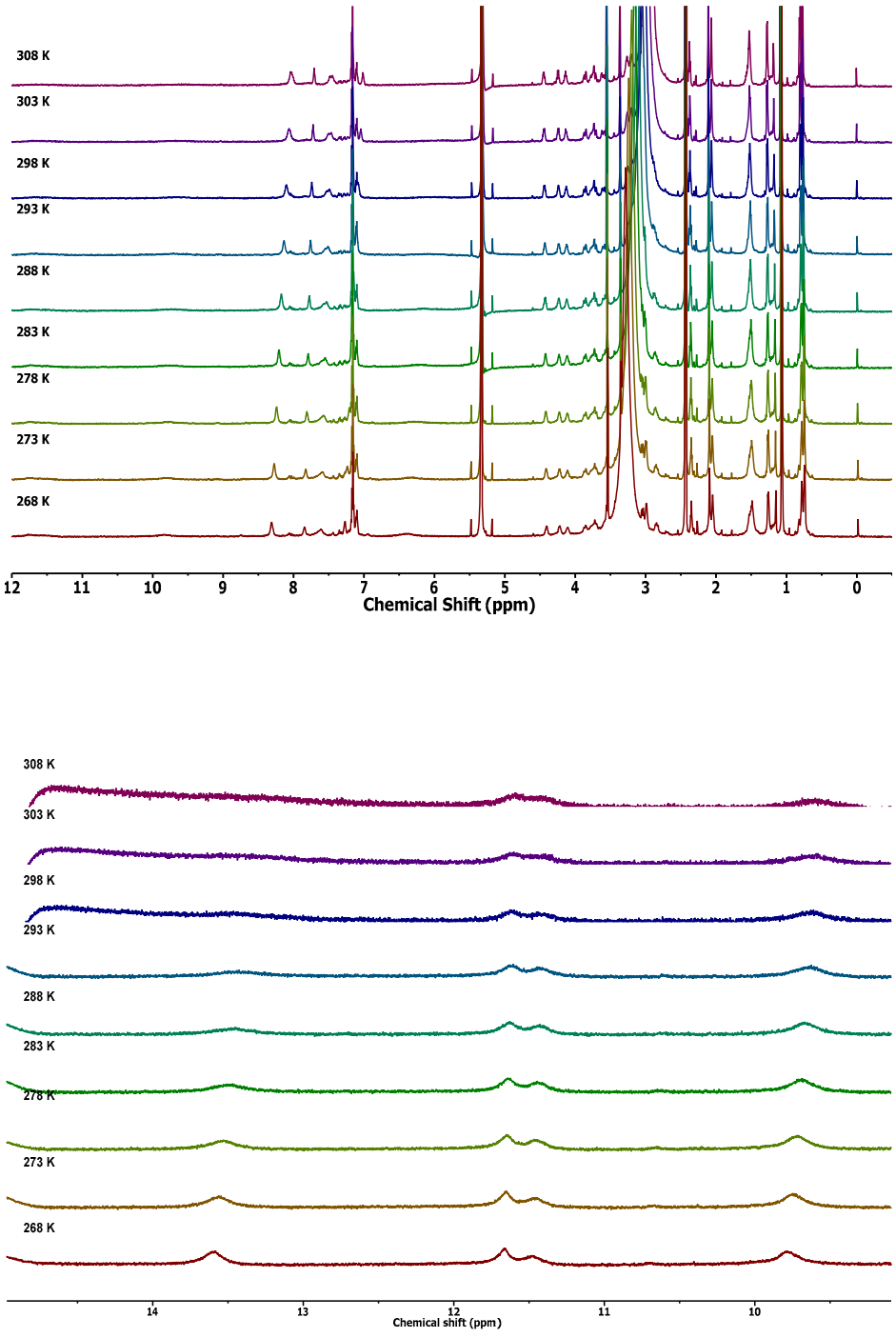


Figure S32. Variable Temperature ^{1}H NMR spectrum of peptide **IV** in 20% DMSO-d_6 in CD_2Cl_2.

Temperature dependence of the ^{1}H NMR chemical shifts (ppb/K) of various NH protons for peptides I – IV

Table S3. Temperature dependence of the ^{1}H NMR chemical shifts (ppb/K) of various NH protons for peptides **I** - **IV** (4 mM in a mixture of 20% DMSO-d_6 and 80% CD_2Cl_2)[a]

Peptide	NH(Phe)	NH(Ala)	NH(Gly)	NH(Leu)	UPy (amide) [b]		UPy (urea) [b]		UPy (urea) [b]	
I	-9.5	-8.5	-6.0	-6.5	n.d.[c]	-6	-14.5	-14.5	-2.5	-2.5
II	-6.5	-7.0	-3.5	-3.0	-6.0	-3.5	-1.0	+1.0	-1.0	-1.5
III	-6.0	-6.0	-3.7	-4.7	-7.0	n.d.[c]	-12.9	-12.9	-1.2	-1.5
IV	-5.6	-5.6	-3.5	-3.1	-3.0	n.d.[c]	-7.0		-1.0	-1.0

[a] ^{1}H NMR spectra recorded at 5 °C intervals from -5 °C to 35 °C. [b] Due to non-equivalence of the two UPy units, two sets of peaks are observed in most cases. [c] The value was not determined due to signal overlap

We investigated the temperature dependence of amide ^{1}H NMR chemical shifts of peptide **I**-**IV** to obtain additional insight into the effect of SAA association on the amide NH protons (See SI Fig. S29-S32, and Table S3 for summary). The changes in temperature chemical shift dependence between **I** and **II** is consistent with an intramolecularly folded conformation for peptide **II**. No significant difference between **III** and **IV** were observed, agreeing with a random conformation for peptide **IV**.

Section of 500 MHz ROESY spectrum of peptide II in 20% DMSO-d6 in CD_2Cl_2

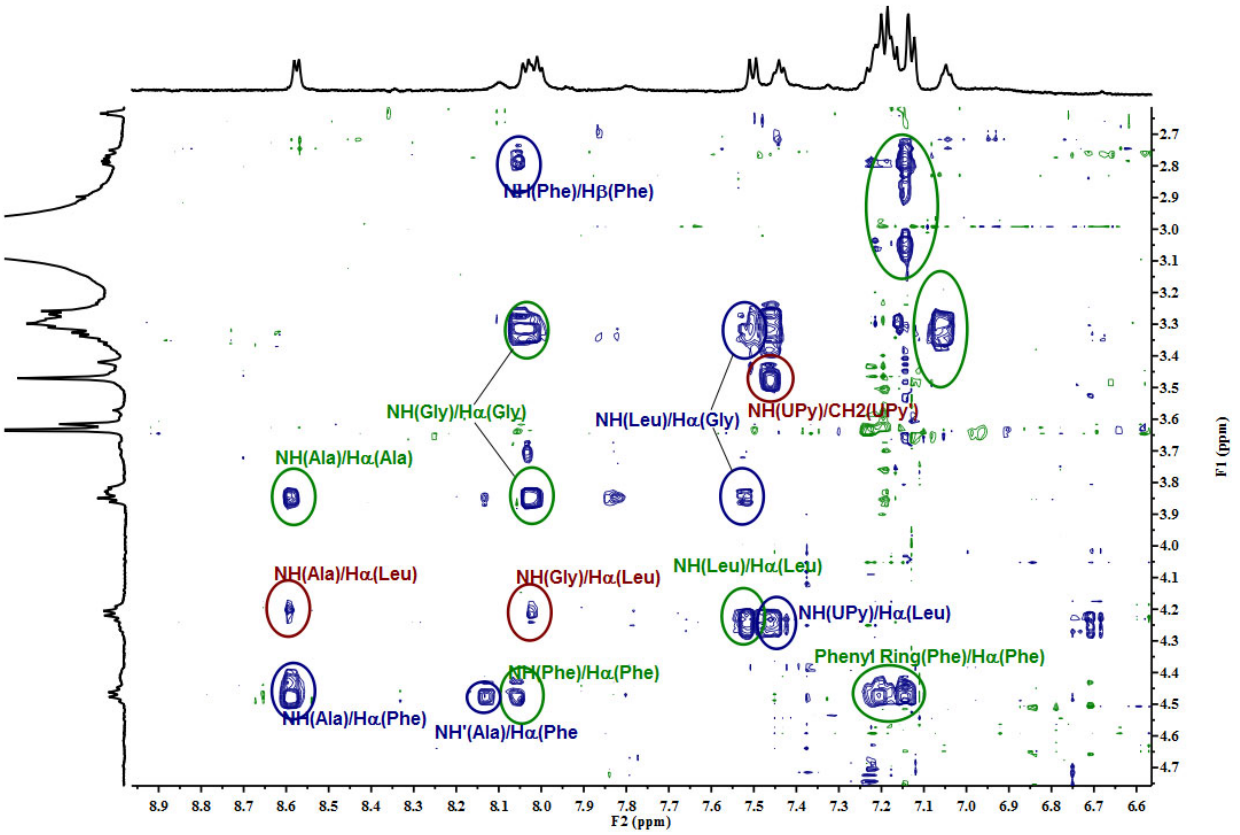


Figure S33. Section of 500 MHz ROESY spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2 (400ms mixing time). Green: intraresidue, Blue: sequential interresidue, Red: nonsequential interresidue.

Overlay of ROESY and TOCSY spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2

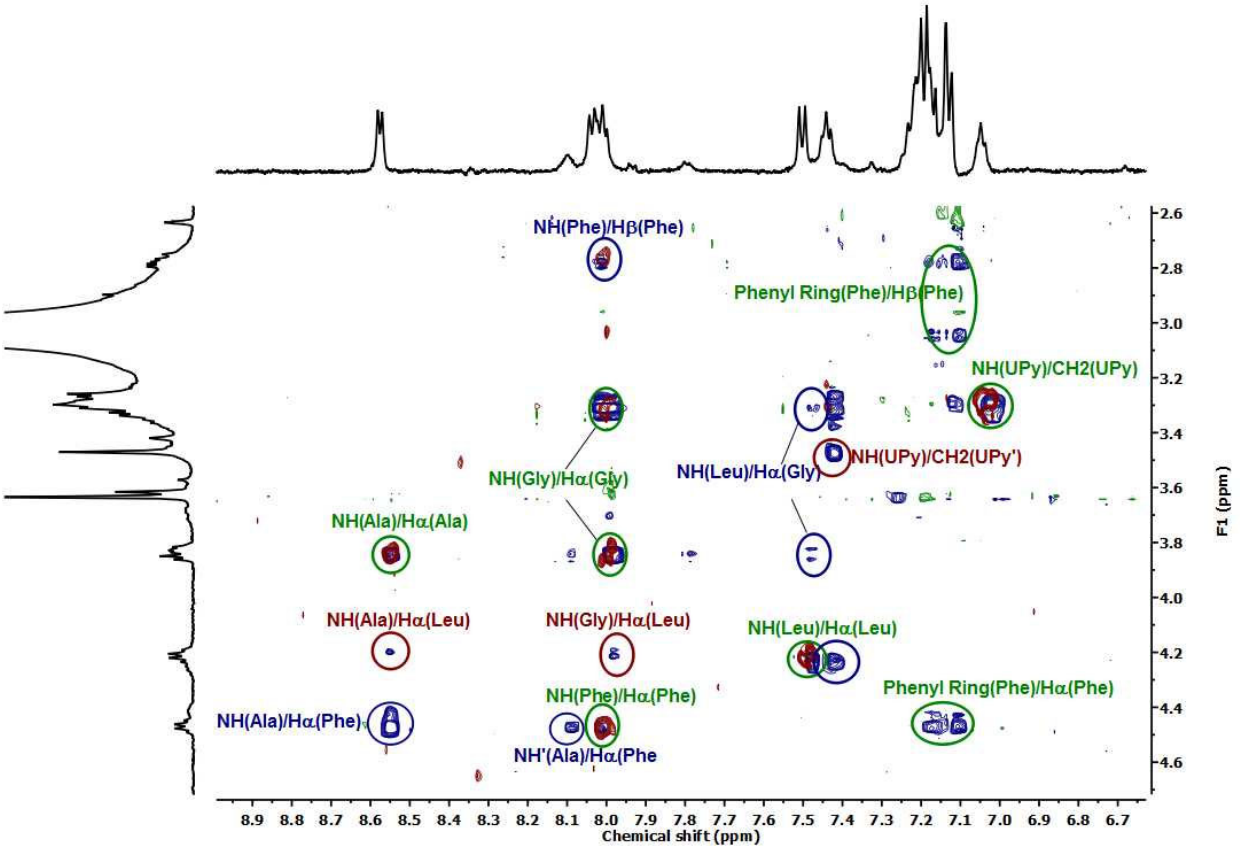


Figure S34. Overlay of ROESY and TOCSY spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2 **.** Green: intraresidue, Blue: sequential interresidue, Red: nonsequential interresidue.

Sections of 500 MHz ROESY spectrum of peptide II in 20% DMSO-d_6 in CD_2Cl_2 showing the EXSY cross-peaks

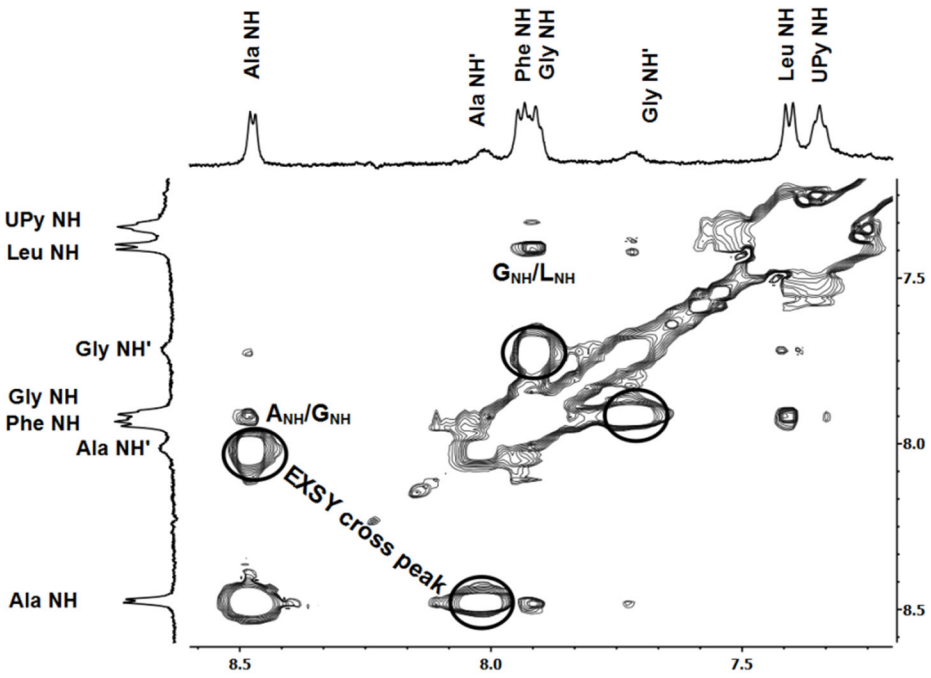


Figure S35. Sections of 500 MHz EXSY spectrum of peptide **II** in 20% DMSO-d_6 in CD_2Cl_2 (400ms mixing time). The EXSY cross peaks are circled, showing the connection between minor peaks and major amide resonances.

The observed chemical exchange cross-peaks (EXSY) in ROESY spectra for peptide II (Fig. S35), suggesting that the minor peaks observed in the ^{1}H NMR spectrum (Fig. 3b) originate from an alternative conformation, possibly an unfolded minor conformer. Previous studies have shown that polar solvents such as DMSO can compete for UPy hydrogen bonding and break up the resulting dimer. (Ref 14) It is therefore possible that this secondary conformer is the result of disruption of UPy dimerization by the DMSO in the solution.

Section of 500 MHz ROESY spectrum of peptide IV in 20% DMSO-d_6 in CD_2Cl_2

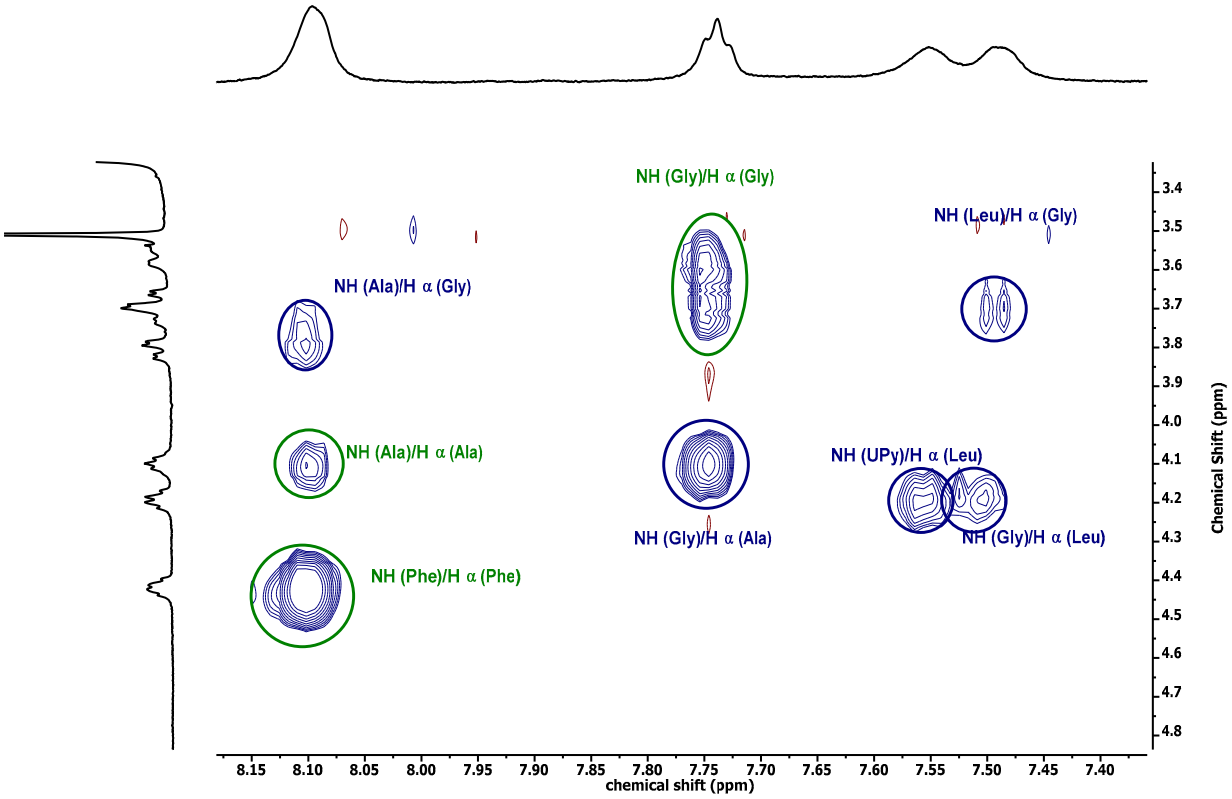


Figure S36. Section of 500 MHz ROESY spectrum of peptide **IV** in 20% DMSO-d_6 in CD_2Cl_2 (400ms mixing time). Green: intraresidue, Blue: sequential interresidue, Red: nonsequential interresidue.

^{1}H and ^{13}C NMR of Fmoc-CUPy(Bn)-OH and Fmoc-NUPy(Bn)-OH.

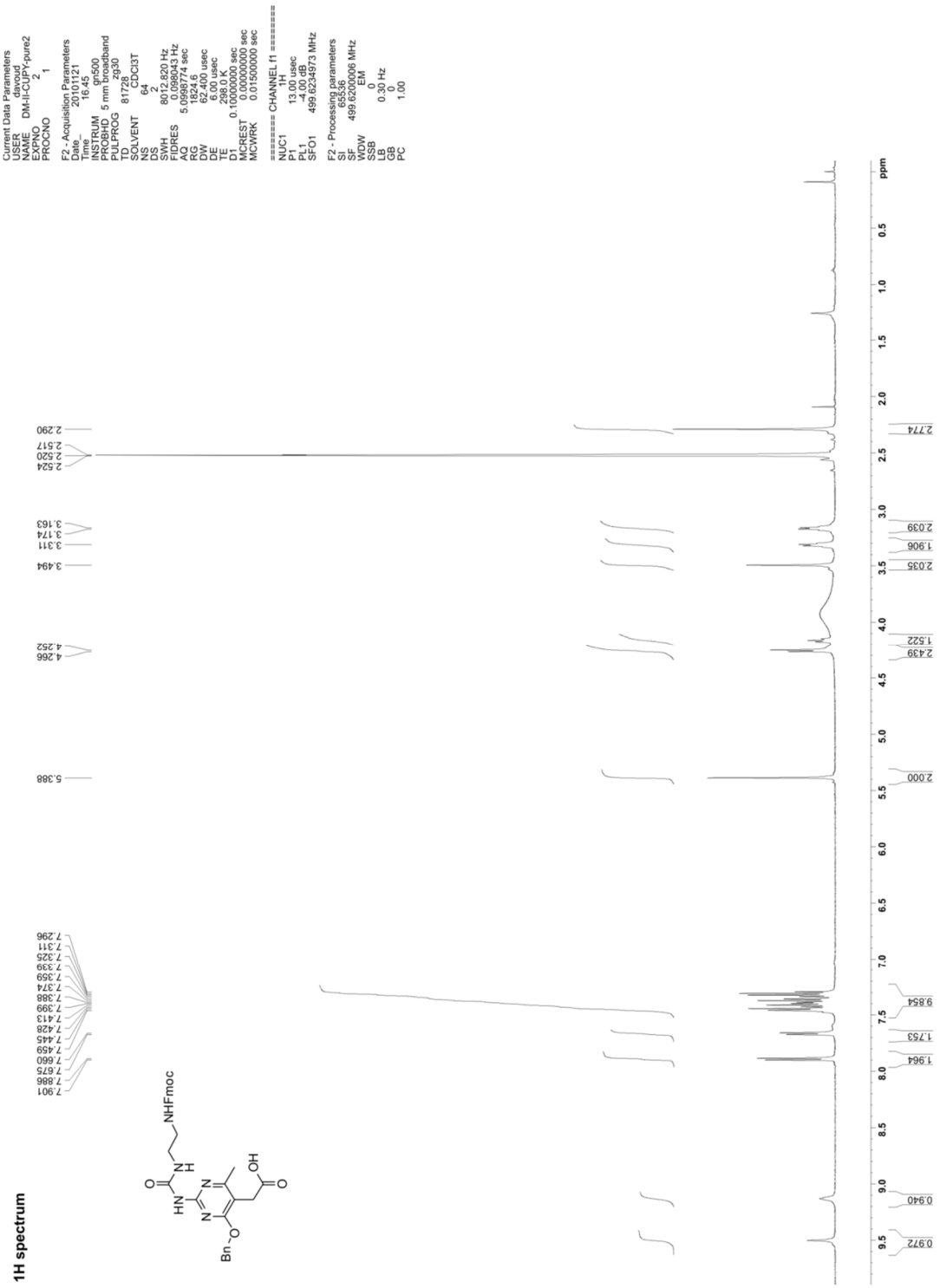


Figure S37. ^{1}H NMR of Fmoc-CUPy(Bn)-OH in DMSO-d_6

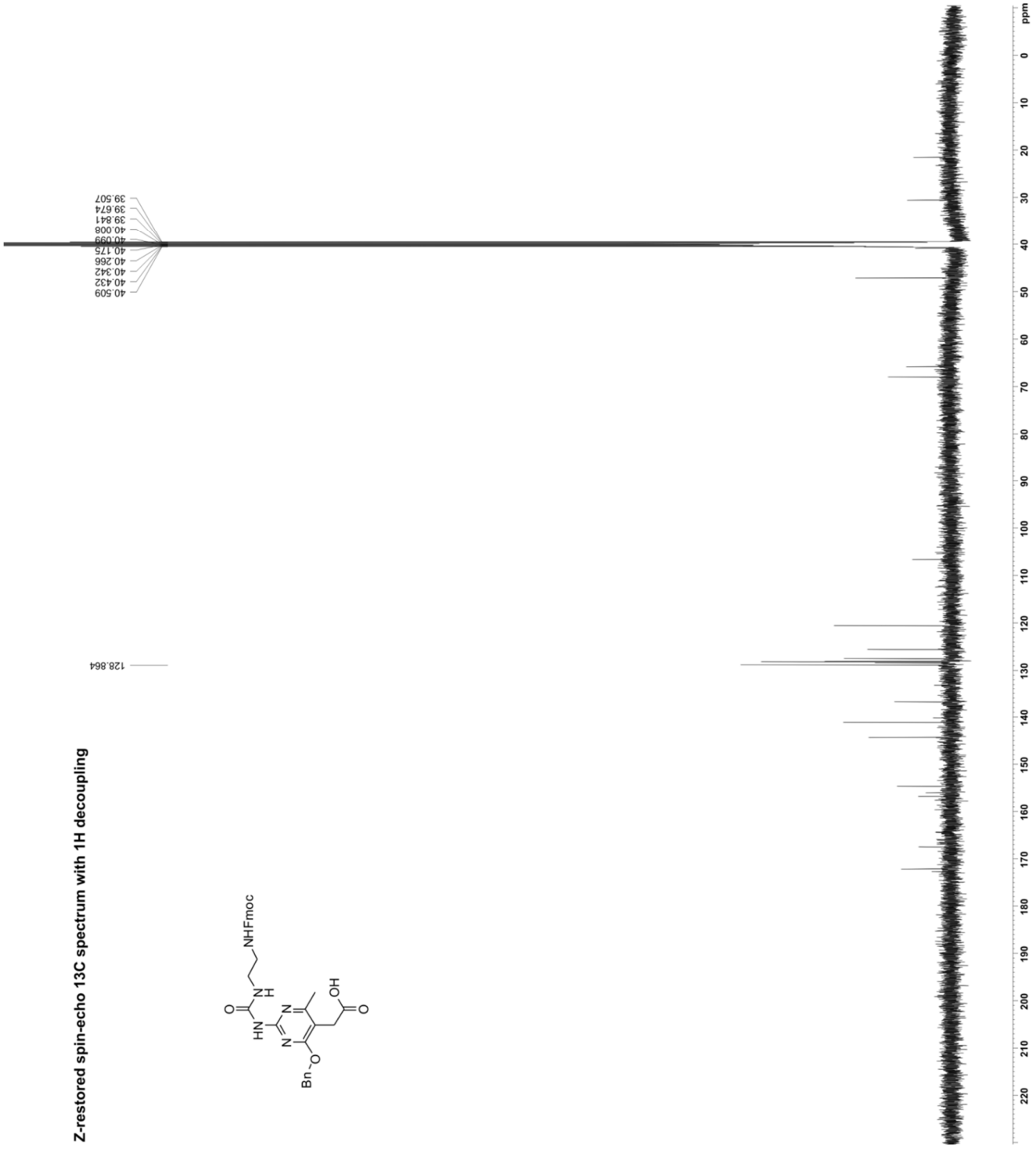


Figure S38. ^{13}C NMR of Fmoc-CUPy(Bn)-OH in DMSO-d_6

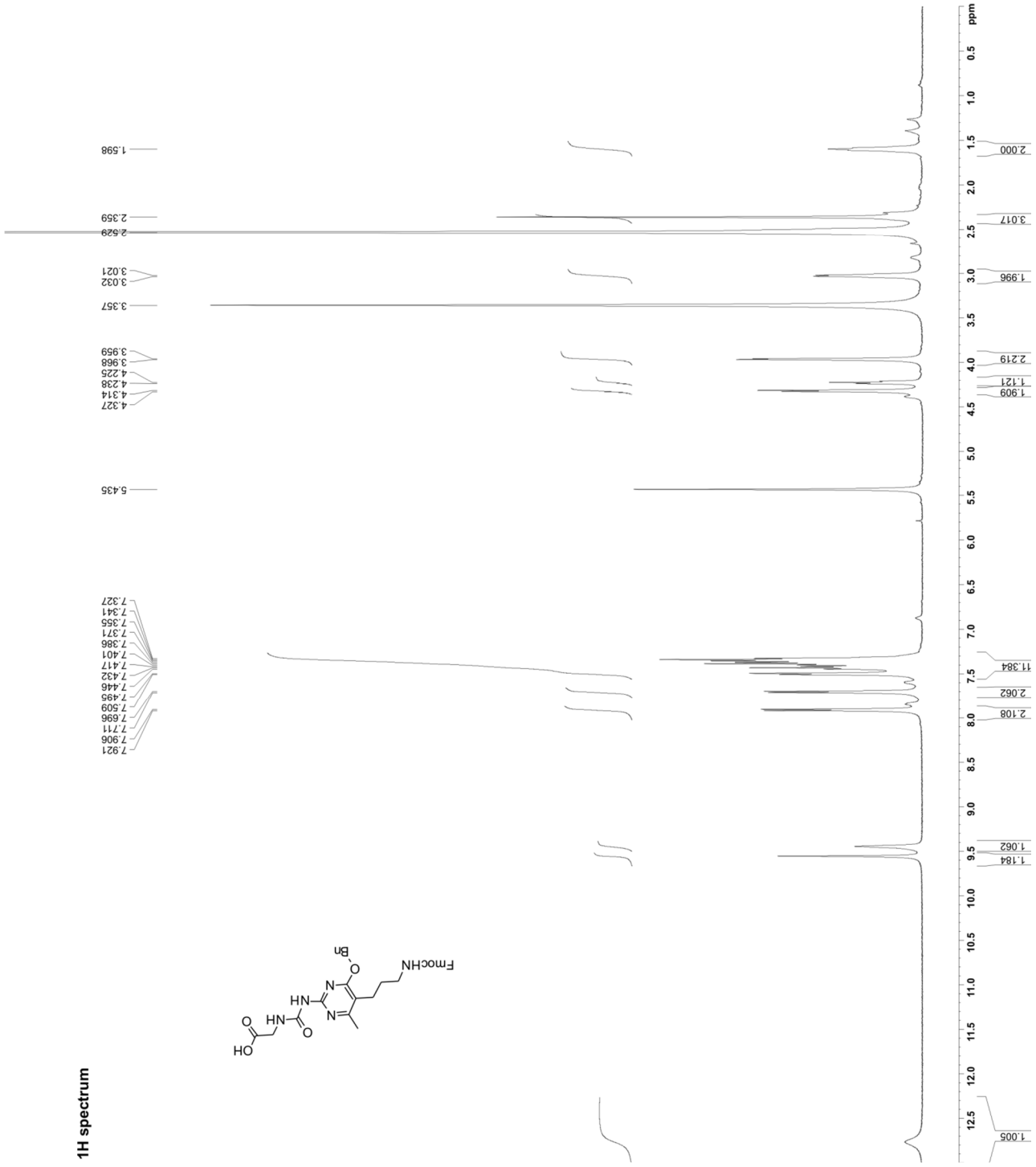


Figure S39. ^{1}H NMR of Fmoc-NUPy(Bn)-OH in DMSO-d_6.

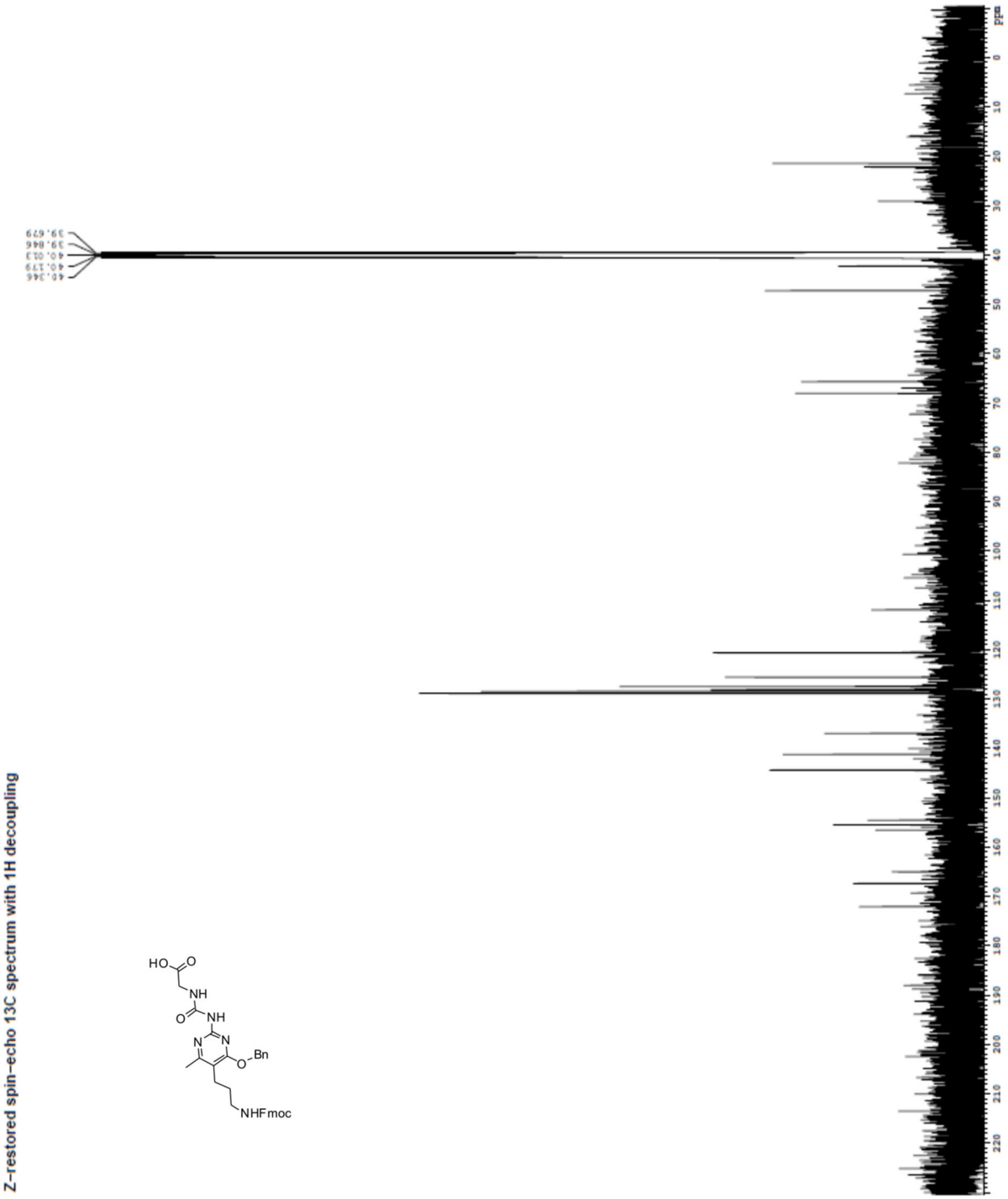


Figure S40. ^{13}C NMR of Fmoc-NUPy(Bn)-OH in DMSO-d_6

References

1 H. M. Keizer, R. P. Sijbesma and E. M. Meijer, *Eur. J. Org. Chem.*, 2004, 2553-2555.
2 R. A. Turner, R. J. Weber and R. S. Lokey, *Org. Lett.*, 2010, **12**, 1852-1855.