Supporting Information

Structural Insight into Hybrid Peptide ε-Helices

Rajkumar Misra,[a] Gijo George,[b] Rahi M. Reja [a], Sanjit Dey [a] Srinivasarao Raghothama*[b] Hosahudya N. Gopi*[a]

[a]Department of Chemistry, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pune 411 008, India.[b]NMR Research Center, Indian Institute of Science, Bangalore-560012, India

E-mail: hn.gopi@iiserpune.ac.in

[1] 2D NMR details of the peptide P1, P2 and P3 in $CDCl_3$:

2D NMR details of the peptide P1:

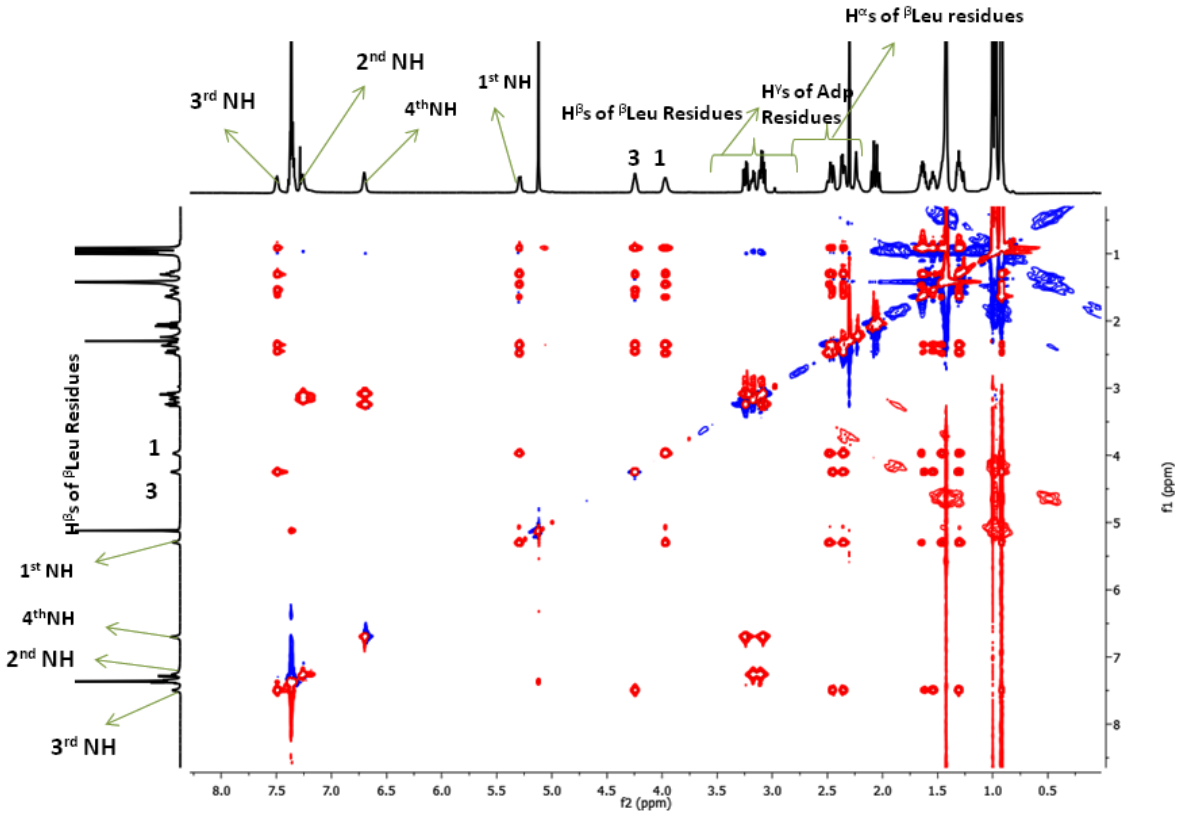


Figure S1: Full TOCSY spectrum of **P1** (5 mM) in $CDCl_3$ at room temperature.

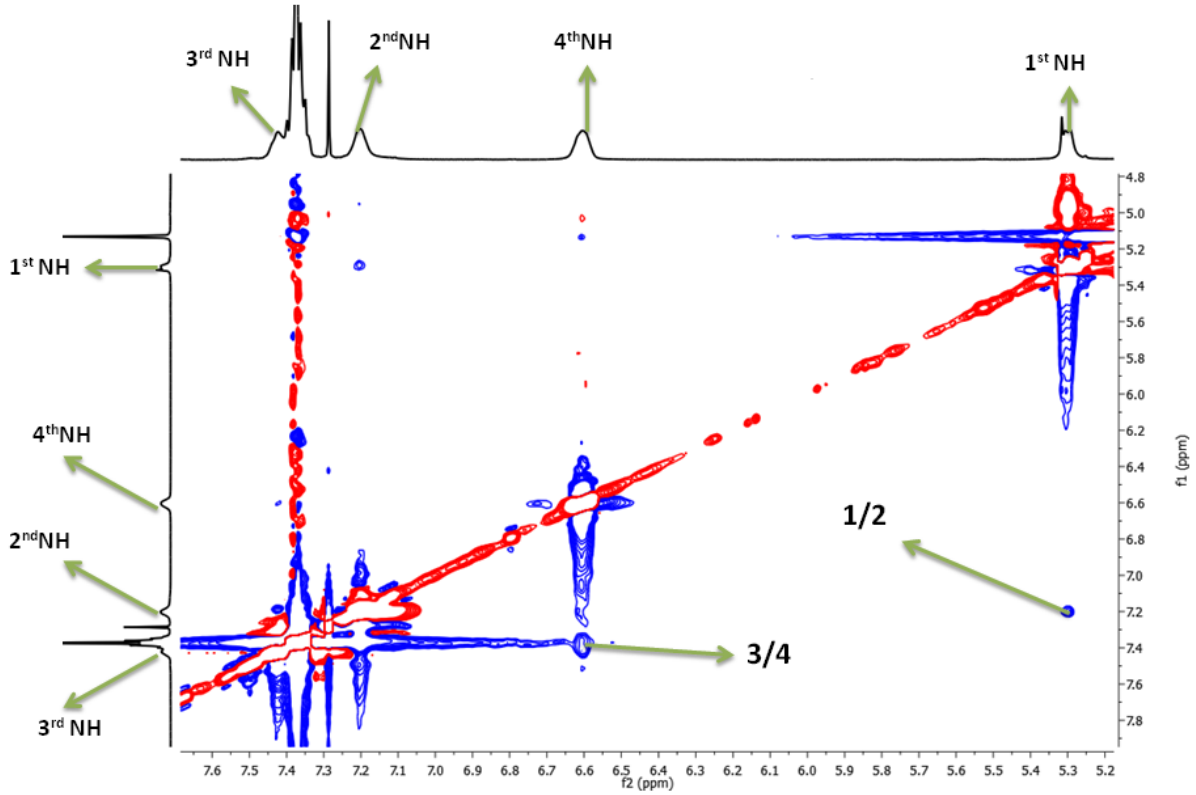


Figure S2: Partial ROESY spectrum of **P1** (5 mM) in $CDCl_3$showing NH↔NH NOEs.

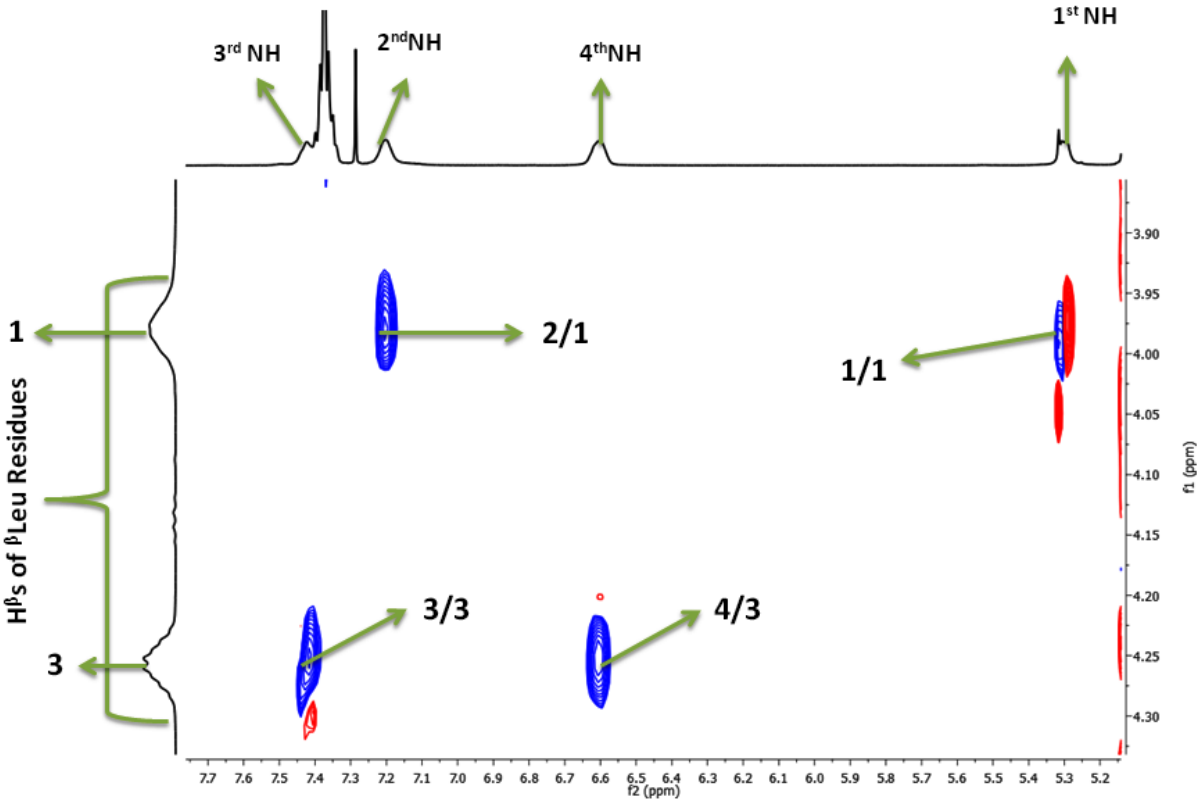


Figure S3: Partial ROESY spectrum of **P1** (5 mM) in $CDCl_3$ showing NH↔$^{\beta}$CH NOEs of β-Leucine.

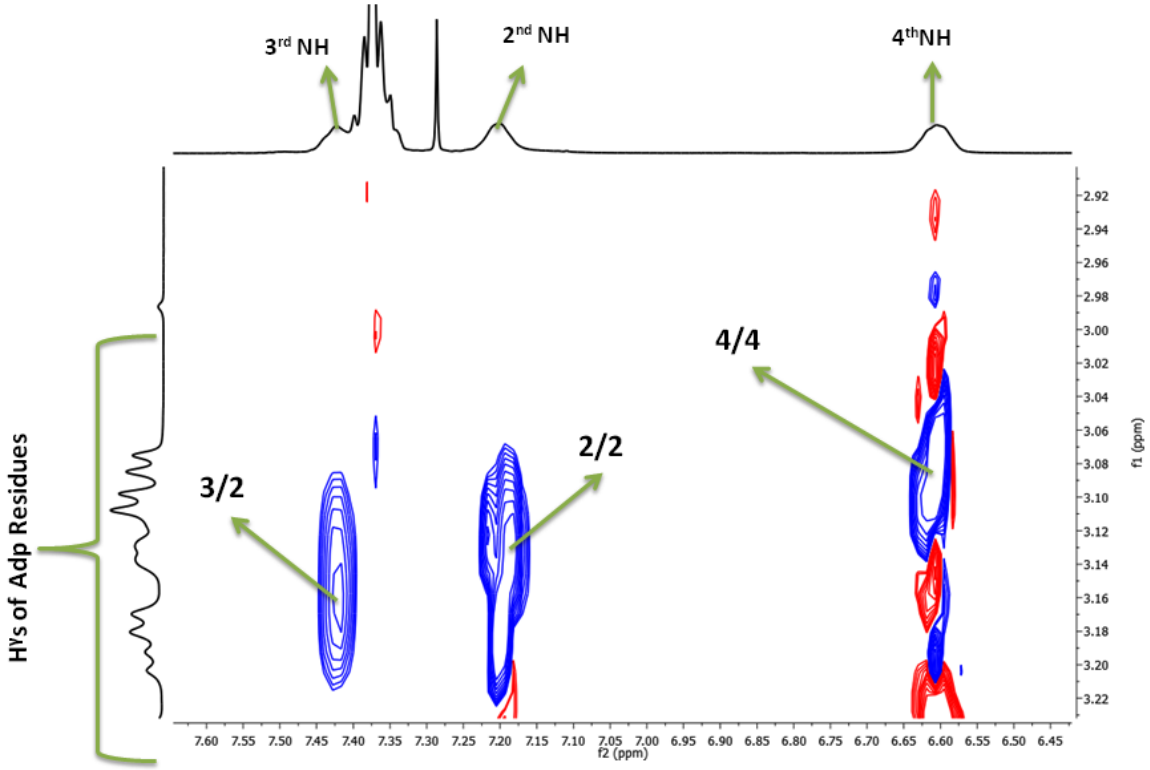


Figure S4: Partial ROESY spectrum of **P1** (5 mM) in $CDCl_3$ showing NH↔$^{\gamma}CH_2$ NOEs of Adb.

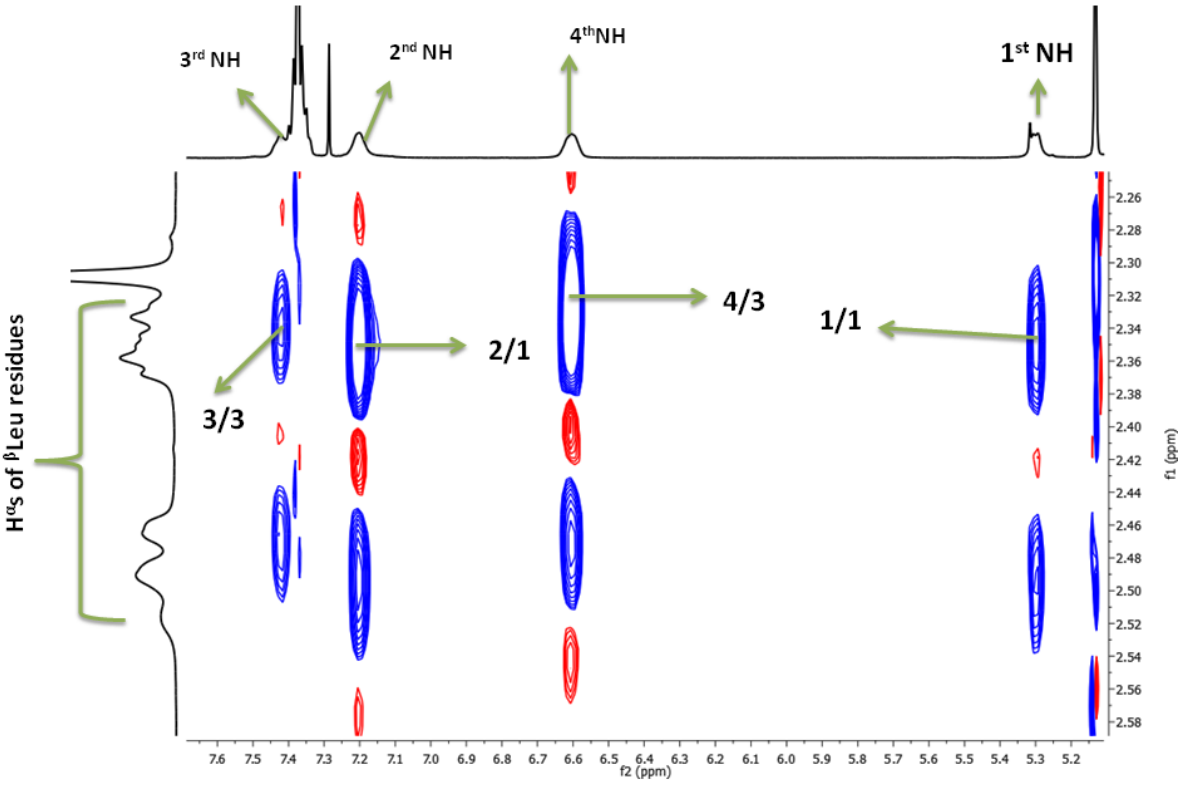


Figure S5: Partial ROESY spectrum of **P1** (5 mM) in $CDCl_3$ showing NH↔$^{\alpha}CH_2$ NOEs of β-Leucine**.**

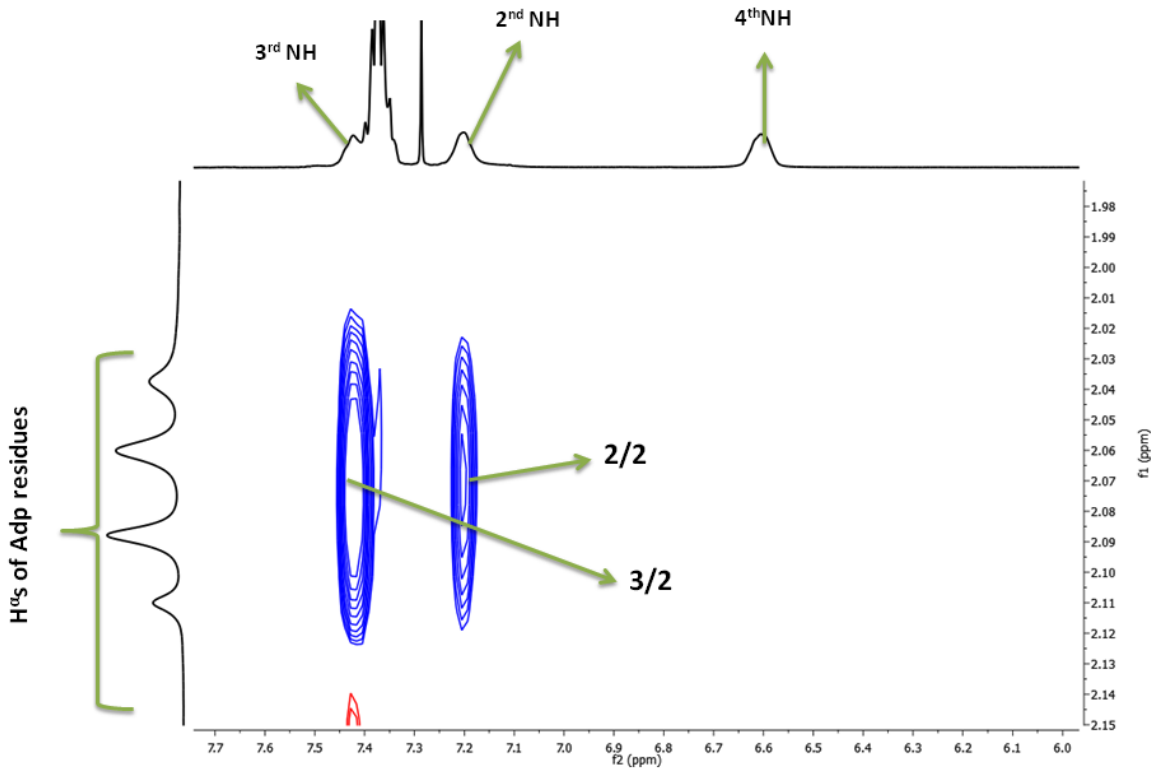


Figure S6: Partial ROESY spectrum of **P1** (5 mM) in $CDCl_3$ showing NH↔$^{\alpha}$CH2 NOEs of Adb**.**

2D NMR details of peptide P2:

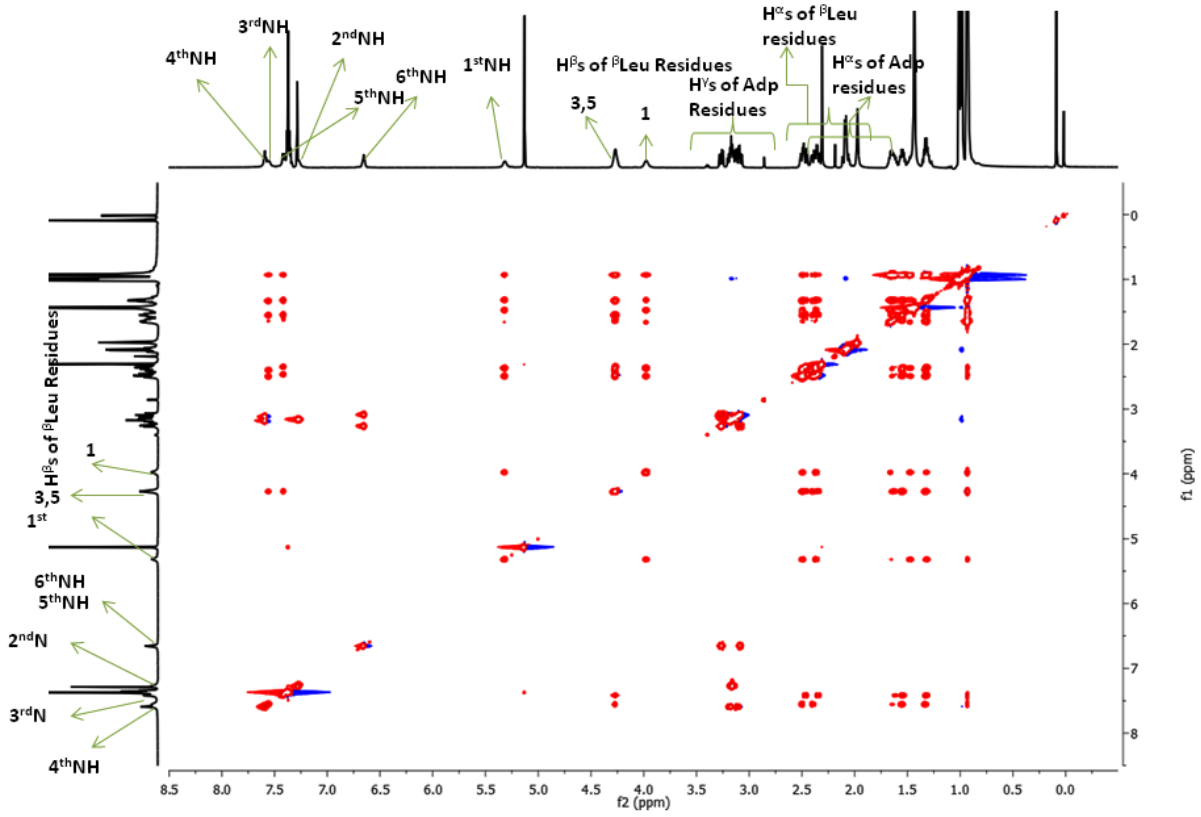


Figure S7: Full TOCSY spectrum of **P2** (5 mM) in $CDCl_3$ at 303K

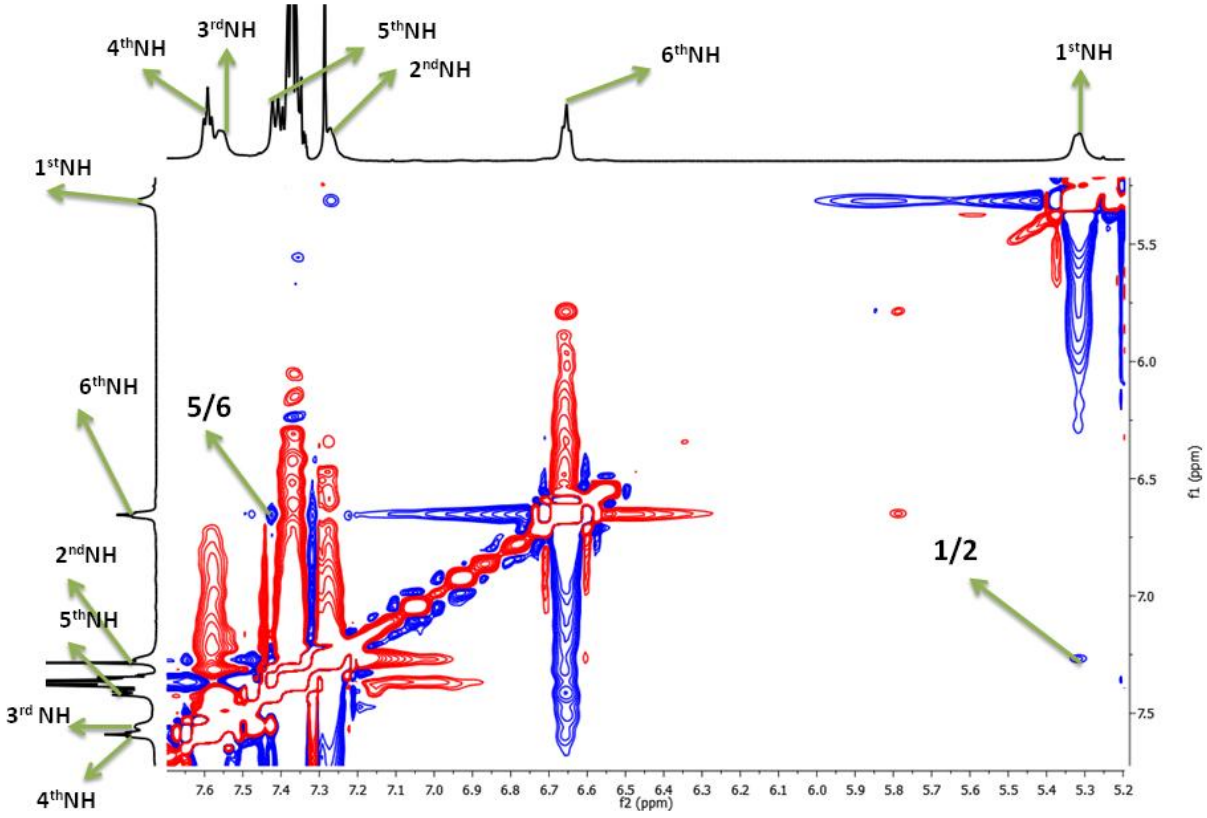


Figure S8: Partial ROESY spectrum of **P2** (5 mM) in $CDCl_3$ showing NH↔NH NOEs at 303K.

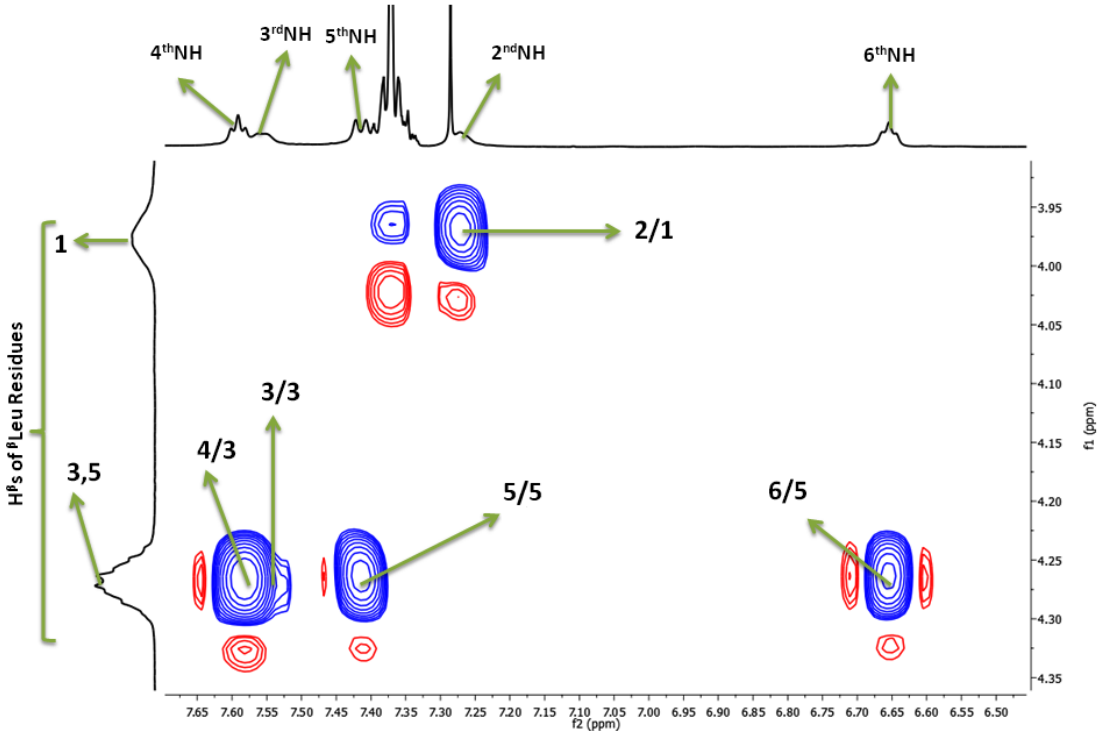


Figure S9: Partial ROESY spectrum of **P2** (5 mM) in $CDCl_3$ showing Adb NH↔$^{\beta}$CH NOEs of β-Leucine at 303K.

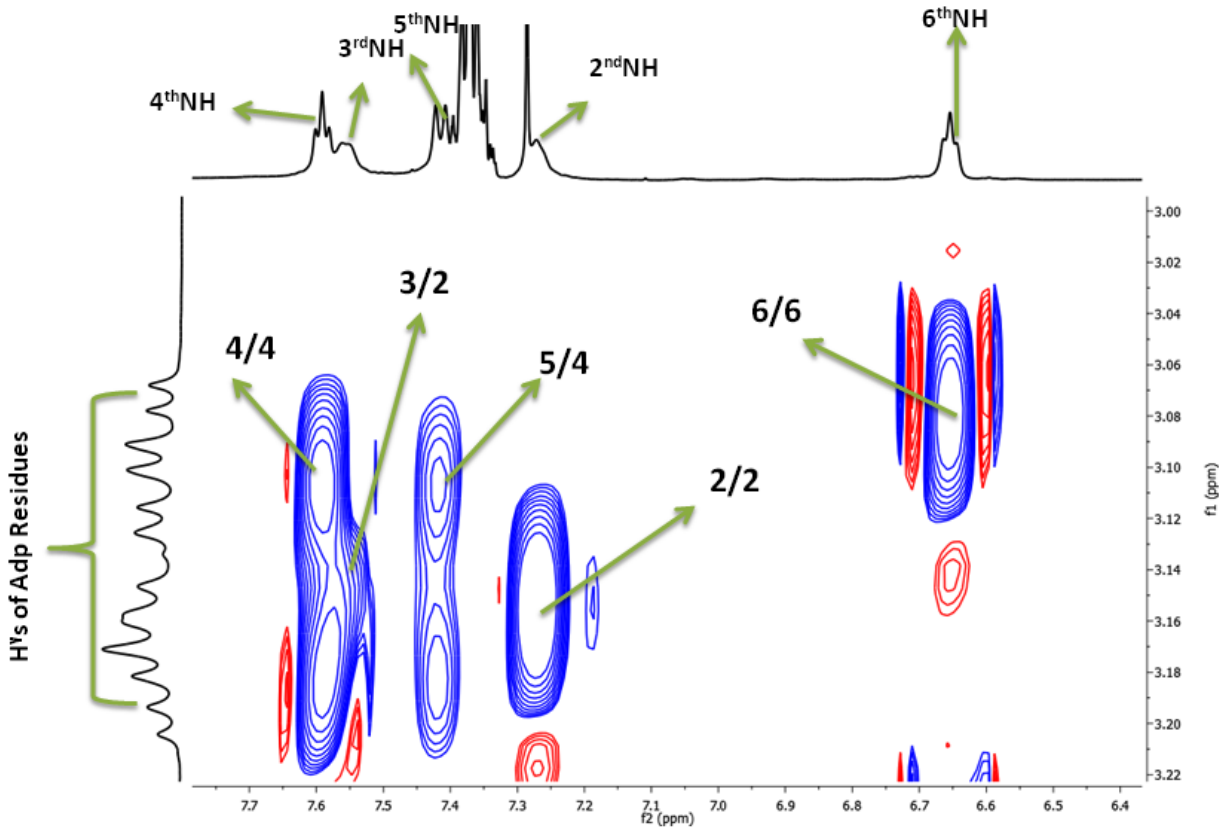


Figure S10: Partial ROESY spectrum of **P2** (5 mM) in $CDCl_3$ showing NH↔$^{\gamma}CH_2$ NOEs of Adb at 303K.

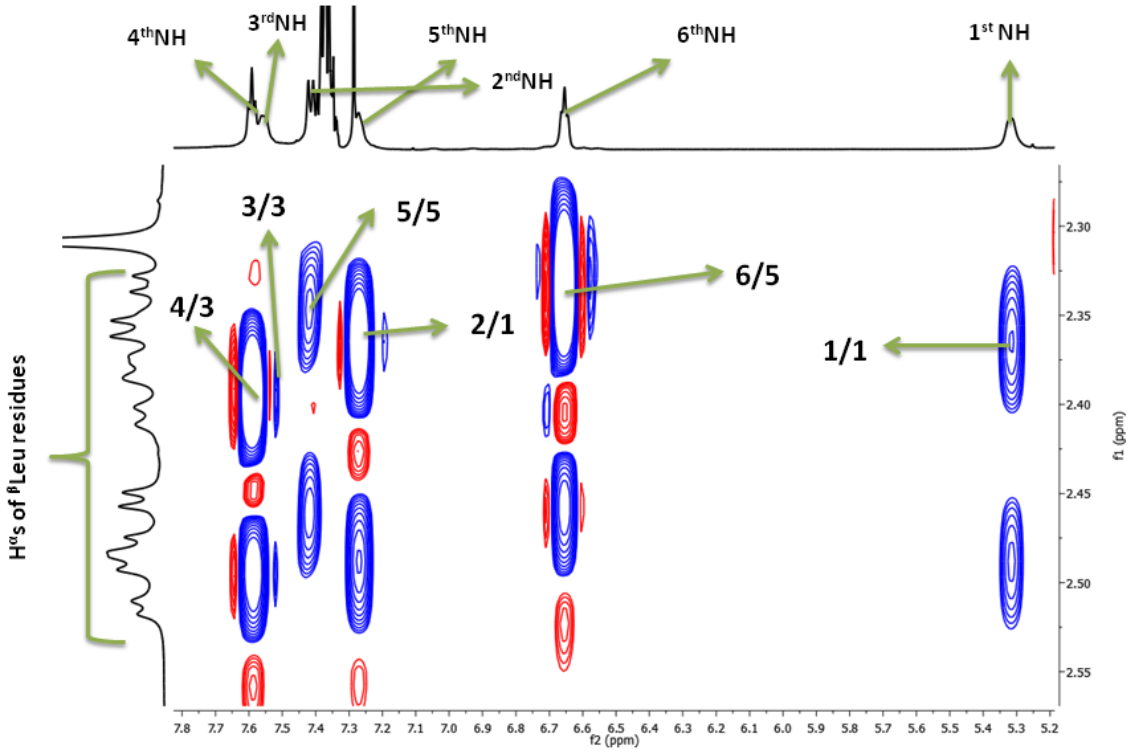


Figure S11: Partial ROESY spectrum of **P2** (5 mM) in $CDCl_3$ showing NH↔$^{\alpha}CH_2$ NOEs of β-Leucineat 303K.

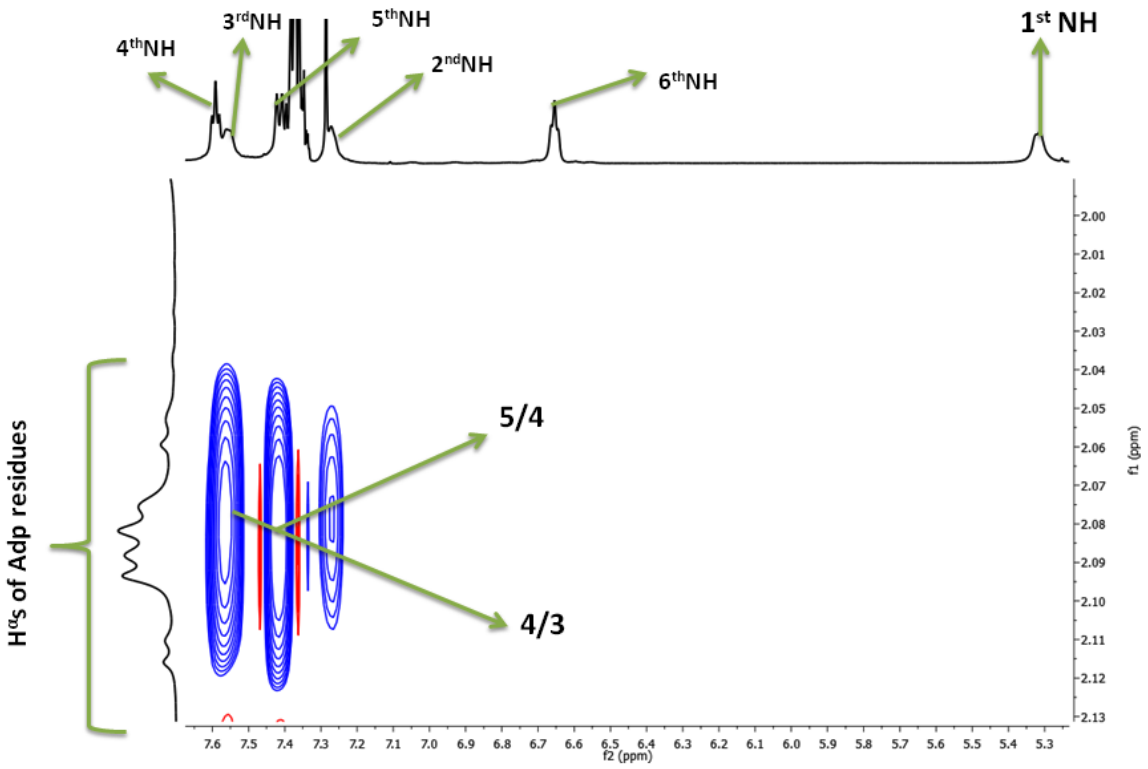


Figure S12: Partial ROESY spectrum of **P2** (5 mM) in $CDCl_3$ showing NH↔$^{\alpha}CH_2$ NOEs of Adb.

2D NMR details of peptide P3:

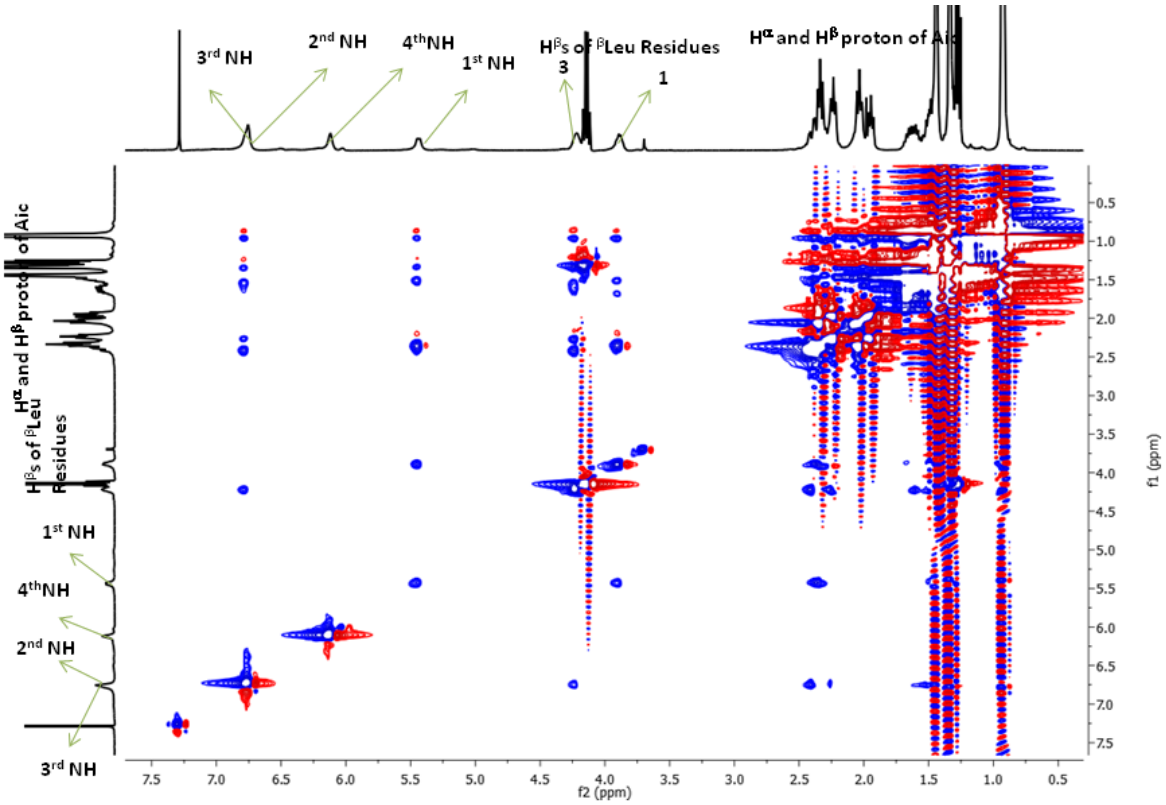


Figure S13: Full TOCSY spectrum of **P3** (5 mM) in $CDCl_3$ at room temperature.

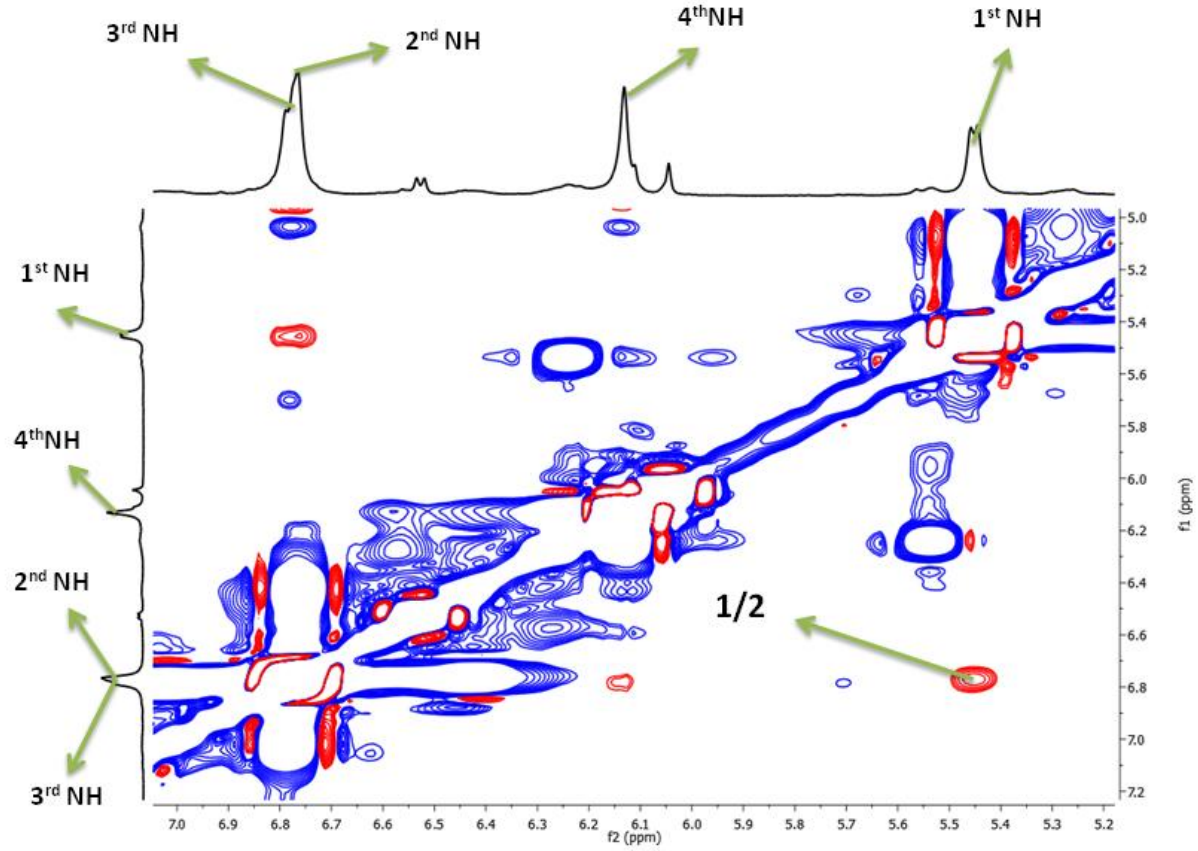


Figure S14: Partial ROESY spectrum of **P3** (5 mM) in $CDCl_3$showing NH↔NH NOEs.

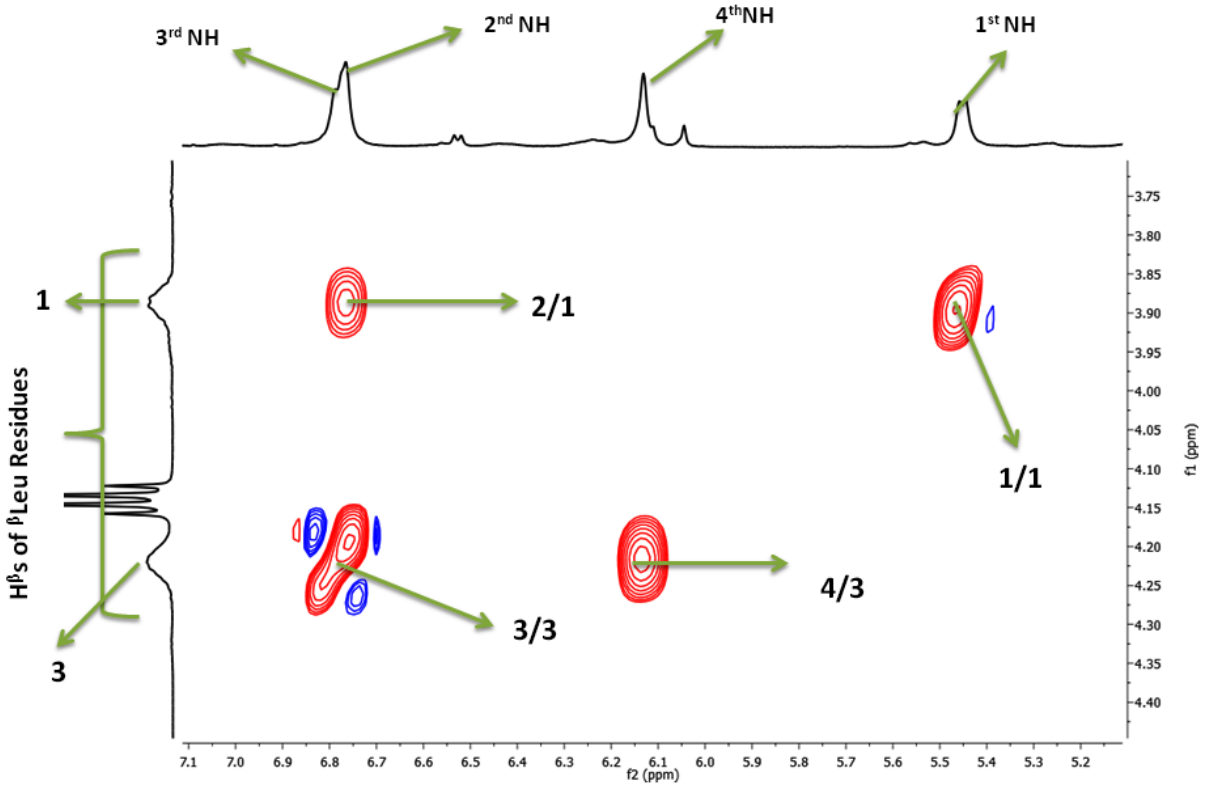


Figure S15: Partial ROESY spectrum of **P3** (5 mM) in $CDCl_3$ showing NH↔ $^{\beta}$CH of β-Leucine

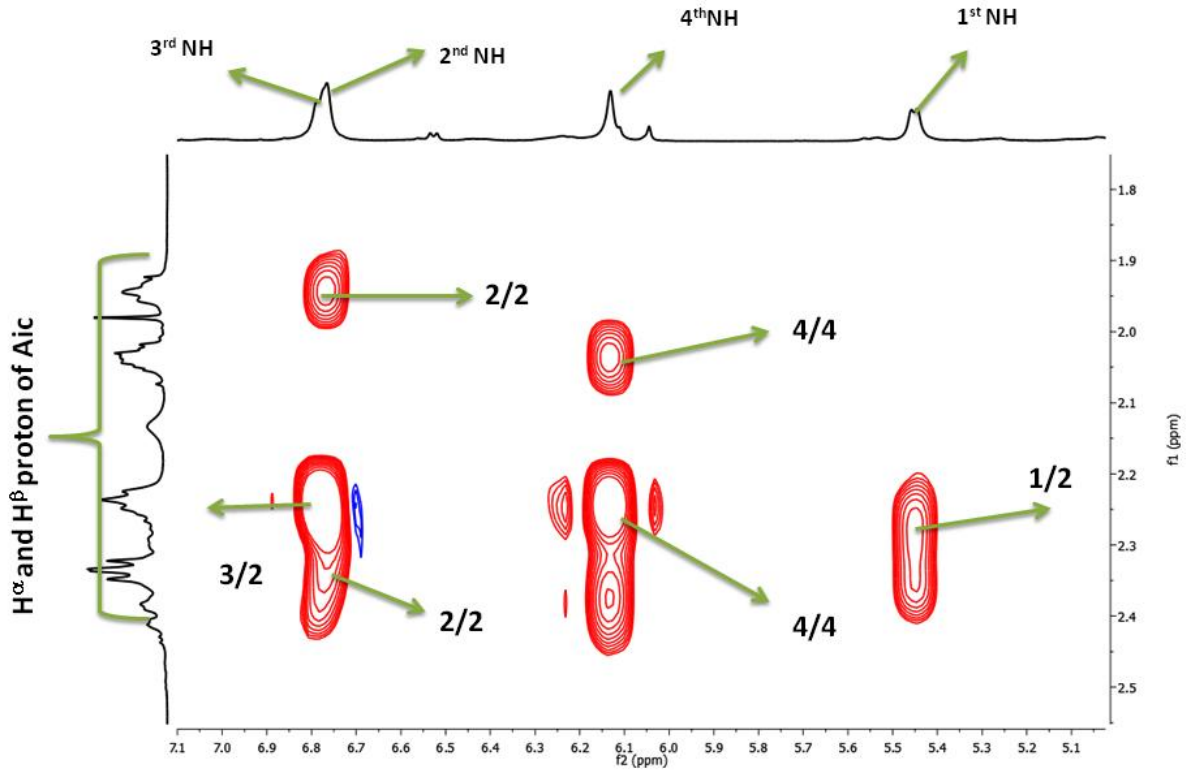
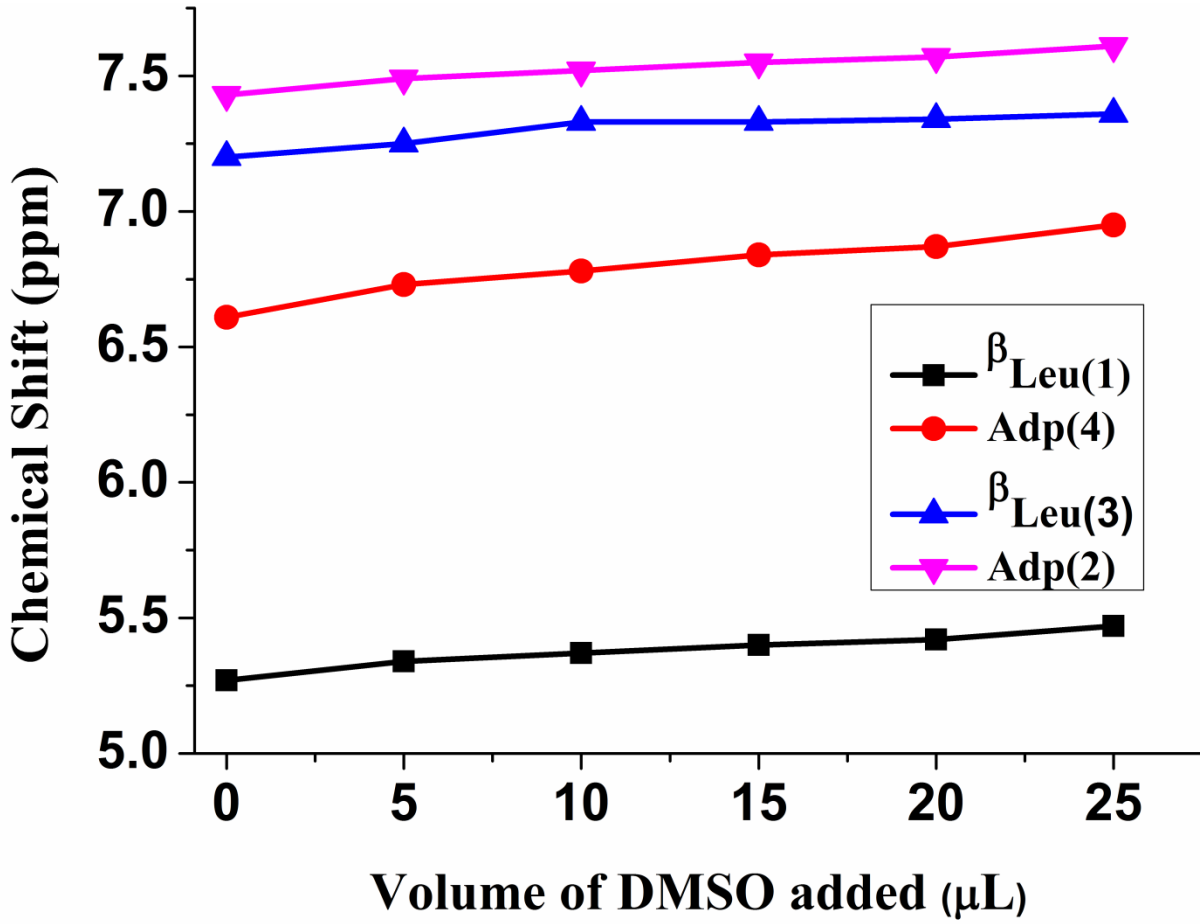


Figure S16: Partial ROESY spectrum of **P3**(5 mM) in $CDCl_3$ showing NH↔ $^{\alpha}CH_2$ and NH↔ $^{\beta}CH_2$ NOEs of Aic.

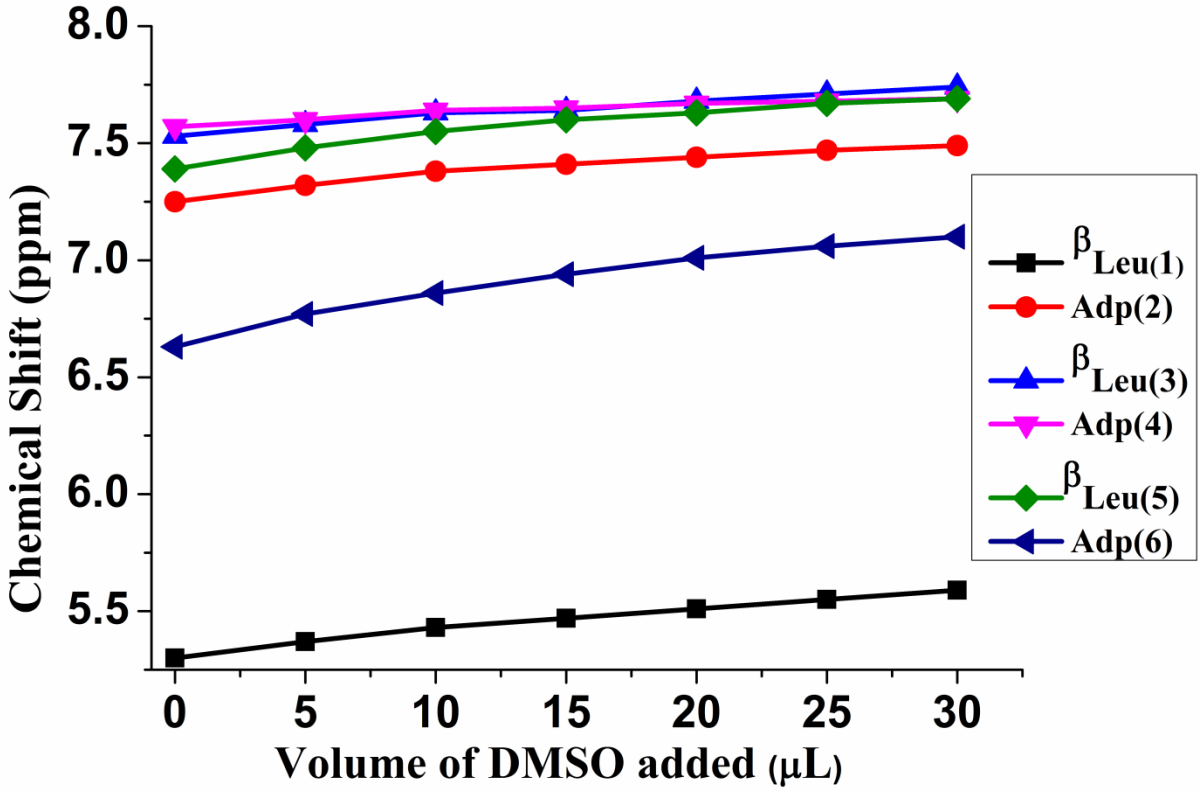
[2]. DMSO titration of the peptide P1, P2 and P3

DMSO titration of peptide P1:



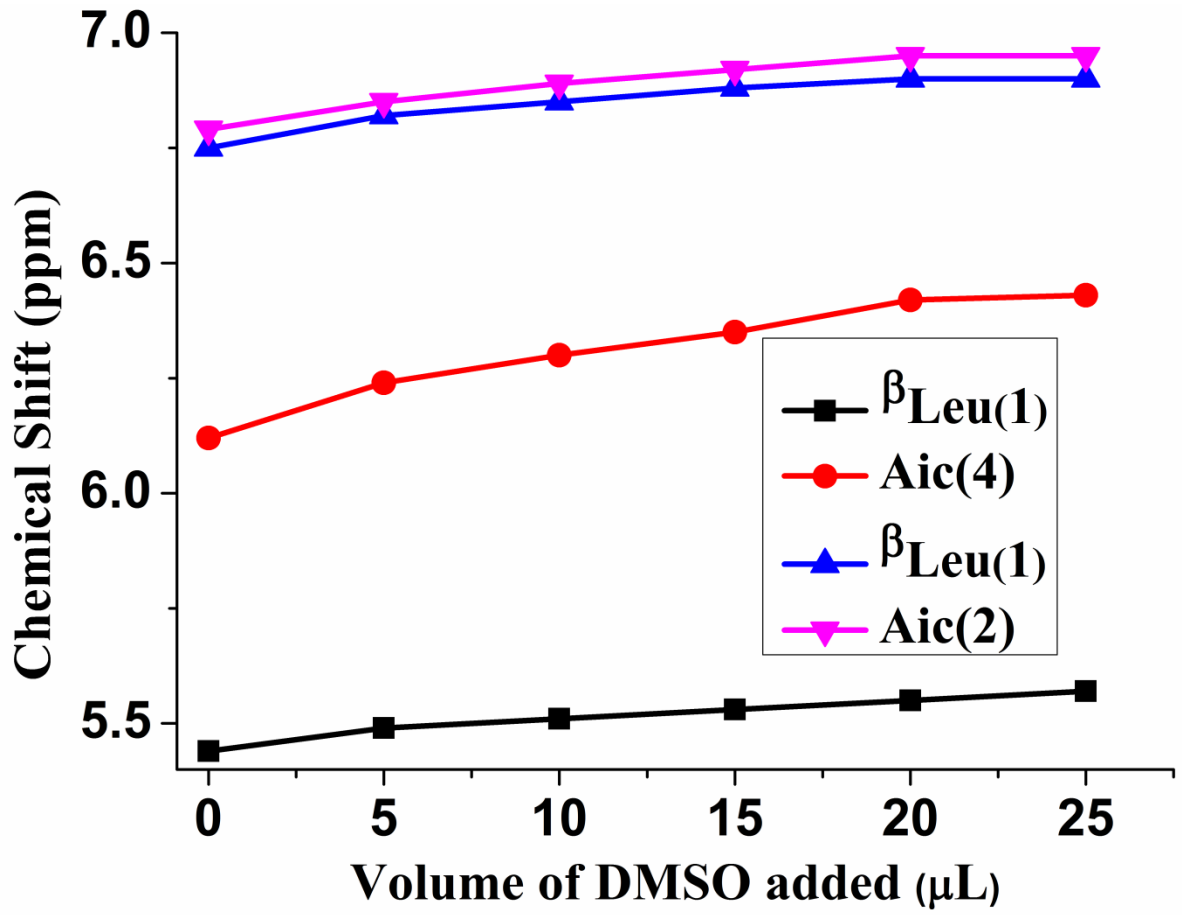
FigureS17. Solvent dependent NH chemical shifts of the peptide **P1**at varying concentrations of $(CD_3)_2SO$.

DMSO titration of peptide P2:



FigureS18. Solvent dependent NH chemical shifts of the peptide**P2**at varying concentrations of $(CD_3)_2SO$.

DMSO titration of peptide P3:



FigureS19. Solvent dependent NH chemical shifts of the peptide**P3**at varying concentrations of $(CD_3)_2SO$.

[3] NMR structure calculation of peptides P1,P2 and P3:

NOEs used for Boc-($^{\beta}$Leu-Adb)$_2$-OCH$_2$Ph(P1):

NH(1)-NH(2) – Weak

NH(3)-NH(4) – Weak

Hβ(1)-NH(1) – Strong

Hβ(1)-NH(2) – Strong

Hβ(3)-NH(3) – Strong

Hβ(3)-NH(4) – Strong

NOEs used for Boc-($^{\beta}$Leu-Aic)$_2$-OEt (P3):

NH(1)-NH(2) – Weak

Hβ(1)-NH(1) – Strong

Hβ(1)-NH(2) – Medium

Hβ(3)-NH(3) – Strong

Hβ(3)-NH(4) – Medium

NOEs used for Boc-($^{\beta}$Leu-Adb)$_3$-OCH$_2$Ph (P2):

NH(1)-NH(2) – Weak

NH(5)-NH(6) – Weak

Hβ(1)-NH(1) – Medium

Hβ(1)-NH(2) – Strong

Hβ(3)-NH(3) – Strong

Hβ(3)-NH(4) – Strong

Hβ(5)-NH(5) – Strong

Hβ(5)-NH(6) – Strong

Strong ≤ 2.5 Å

Medium ≤ 3.5 Å

Weak ≤ 5 Å

Hydrogen bond constraints used for the peptide P1andP2:

NH(1) → CO(2)

NH(2) → CO(3)

NH(3) → CO(4)

Hydrogen bond constraints used for the peptide P3:

NH(1) → CO(2)

NH(2) → CO(3)

NH(3) → CO(4)

NH(4) → CO(5)

NH(5) → CO(6)

Structure calculation was done using a simulated annealing protocol in vacuum using DESMOND and OPLS 2005 force field with NOE and hydrogen bonding constraints.[1] A peptide molecule was kept in orthorhombic simulation cell. Upper limit for distance was kept at 2.5 Å, 3.5 Å and 5 Å for strong, medium and weak NOEs respectively. All the lower distance limits were taken to be 1.8 Å. For Hydrogen bonding constraints, an upper bound of 2.1 Å and lower bound of 1.8 Å was used. A force constant of 3.5 Kcal/Mol and 2.0 Kcal/Mol were used for all the constraints used for tetrapeptides and hexapeptide. NOE potentials (appropriate for treating ambiguous NOE assignments) used are having the following form,

E_{NOE} = **fc*** (**lower** - *d*)^2, if *d* <**lower**;

E_{NOE} = 0, if **lower** <= *d* <= **upper**;

E_{NOE} = **fc*** (**upper** - *d*)^2, if **upper** <*d* <= **upper** + **sigma**;

E_{NOE} = **fc*** (a + **beta** * (*d* - **upper**) + c / (*d* - **upper**)), if *d* >**upper** + **sigma**;

Where *d* is the distance and fc is the force constant.

Values of sigma and beta used in the calculation are 0.5 and 1.5 respectively. The values a and c are determined automatically such that potential is continuous and differential everywhere.

Before production run simulation, a default NVT relaxation was done as implemented in DESMOND. NVT ensemble was used for the production run simulation. Berendsen thermostat with a relaxation time of 1 ps was used. A RESPA integrator was used in which for all the bonded interactions and near nonbonded interactions a time step of 1 fs was used and far nonbonded interaction time step of 3 fs was used. A cutoff of 9 Å was used for short range electrostatic interactions. A smooth particle mesh ewald method was used for treating long range electrostatic interactions. Simulated annealing was done in 6 stages. First stage consist simulation for 30 ps at 10 K. In the second stage, temperature was linearly increased to 100 K till 100 ps. In the third stage, temperature was linearly increased to 300 K till 200 ps. In the fourth stage, temperature was linearly increased to 400 K till 300 ps. In the fifth stage, temperature was maintained at 400 k till 500 ps. In the sixth stage, temperature was linearly decreased to 300 K till 1000 ps and maintained at 300 k till 1200 ps. 20 minimum energy structures were taken from the trajectory between 1000 ps and 1200 ps. The lowest energy structure from the trajectory between 1000 ps and 1200 ps was taken and minimized using a steepest descent method using a convergence gradient threshold of 0.05 kcal/mol/Å.

[4]. Superimposition of 20 lowest energy structures of peptide P1, P2 and P3 determined by NMR in $CDCl_3$:

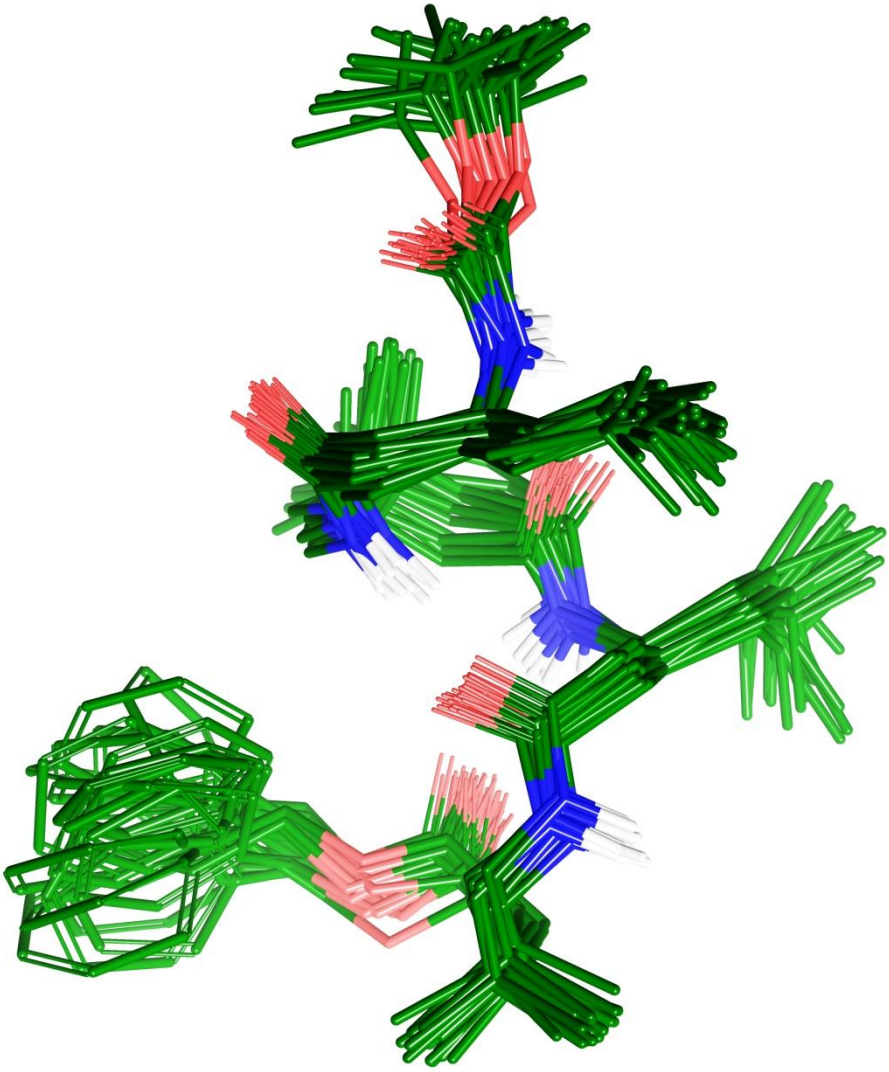


Figure S20. Backbone superposition of 20 NMR-determined structures of peptide **P1**

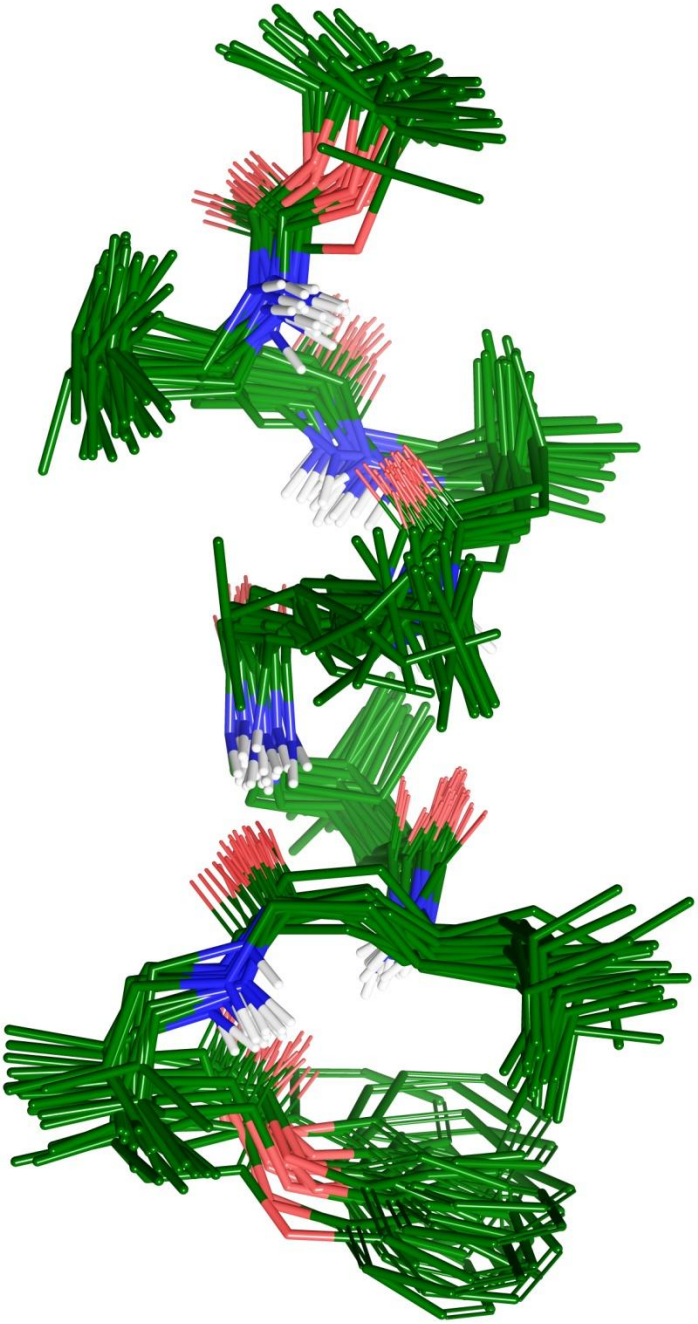


Figure S21: Backbone superposition of 20 NMR determined structures of peptide **P2.**

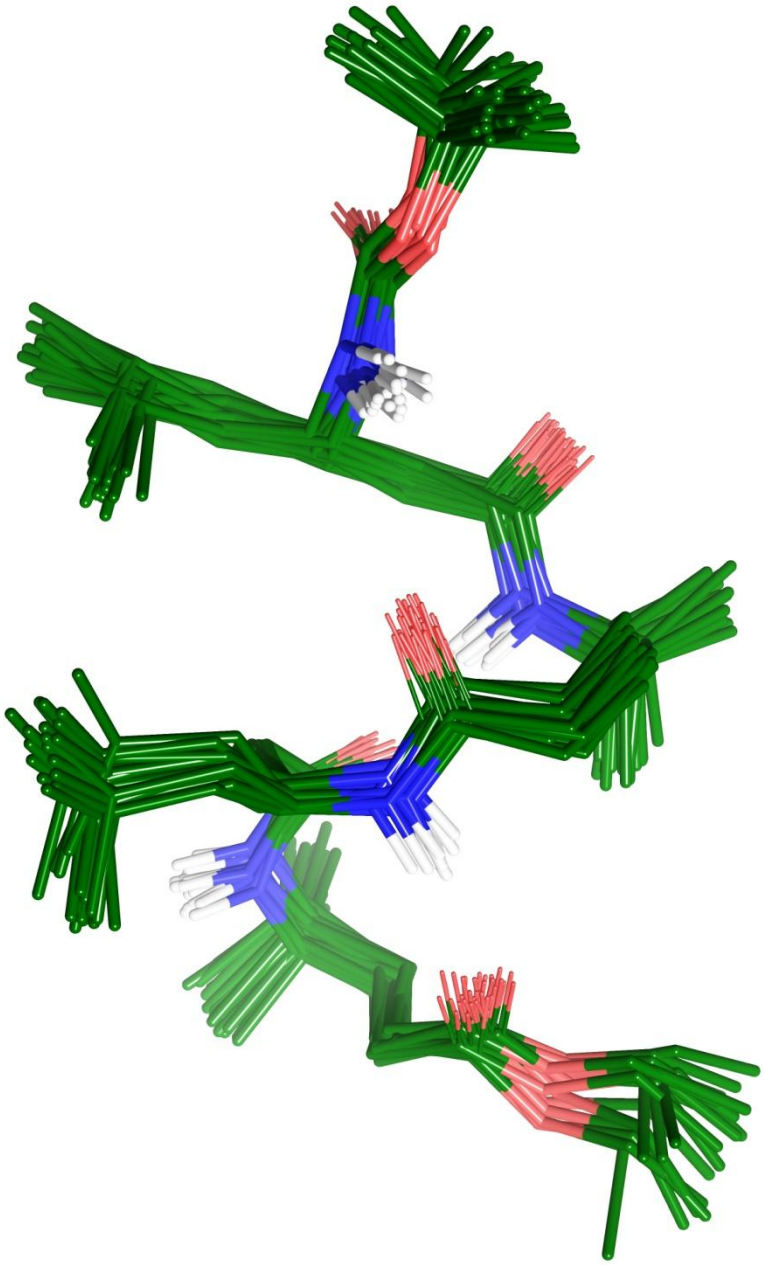


Figure S22: Backbone superposition of 20 NMR determined structures of peptide **P3**

[5] Torsion angle Table of P1, P2 and P3 derived from NMR solution spectroscopy in $CDCl_3$

Table S1: Dihedral angles (in degree) measured from the minimized lowest energy conformation of peptide **P1**from simulation

Residue	φ	θ1	θ2	ψ
$^{\beta}$Leu(1)	63.07	45.94		60.59
Adp(2)	94.24	79.20	-59.49	126.61
$^{\beta}$Leu(3)	64.03	58.50		81.45
Adp(4)	146.39	-52.54	-58.23	-85.48

Table S2: Dihedral angles (in degree) measured from the minimized lowest energy conformation of peptide **P2** from simulation

Residue	φ	θ1	θ2	ψ
$^{\beta}$Leu(1)	68.09	52.22		33.85
Adp(2)	101.40	79.05	-61.90	144.99
$^{\beta}$Leu(3)	65.29	46.70		62.51
Adp(4)	90.68	75.22	-59.94	135.10
$^{\beta}$Leu(5)	62.78	53.64		81.92
Adp(6)	135.61	-77.04	70.73	79.02

Table S3: Dihedral angles (in degree) measured from the minimized lowest energy conformation of peptide **P3** from simulation

Residue	φ	θ1	θ2	ψ
$^{\beta}$Leu(1)	69.87	53.90		97.88
Aic(2)	68.35	76.51	-63.82	131.00
$^{\beta}$Leu(3)	72.02	50.79		91.10
Aic(4)	61.55	55.48	-136.08	-68.86

[6] 2D NMR details of the peptide P2 and in CD_3OH:

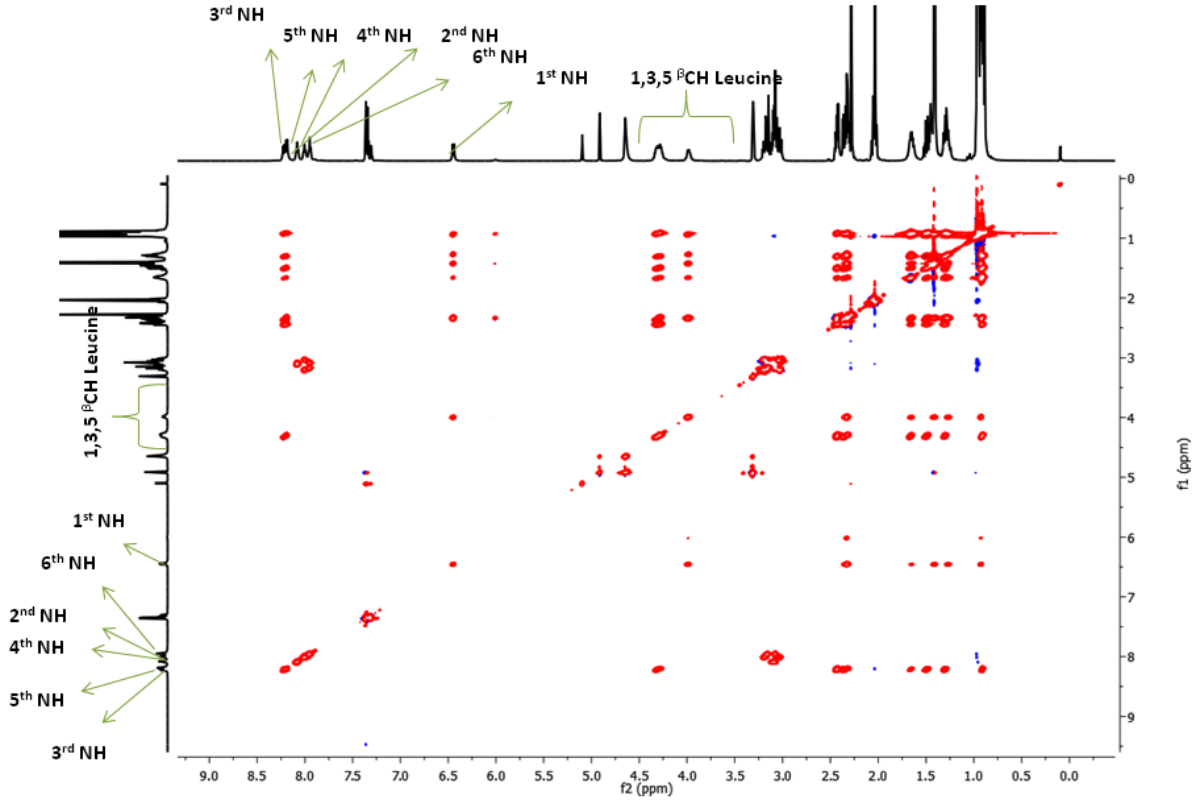


Figure S23: TOCSY spectrum of **P2** (5 mM) in CD_3OH at 298K

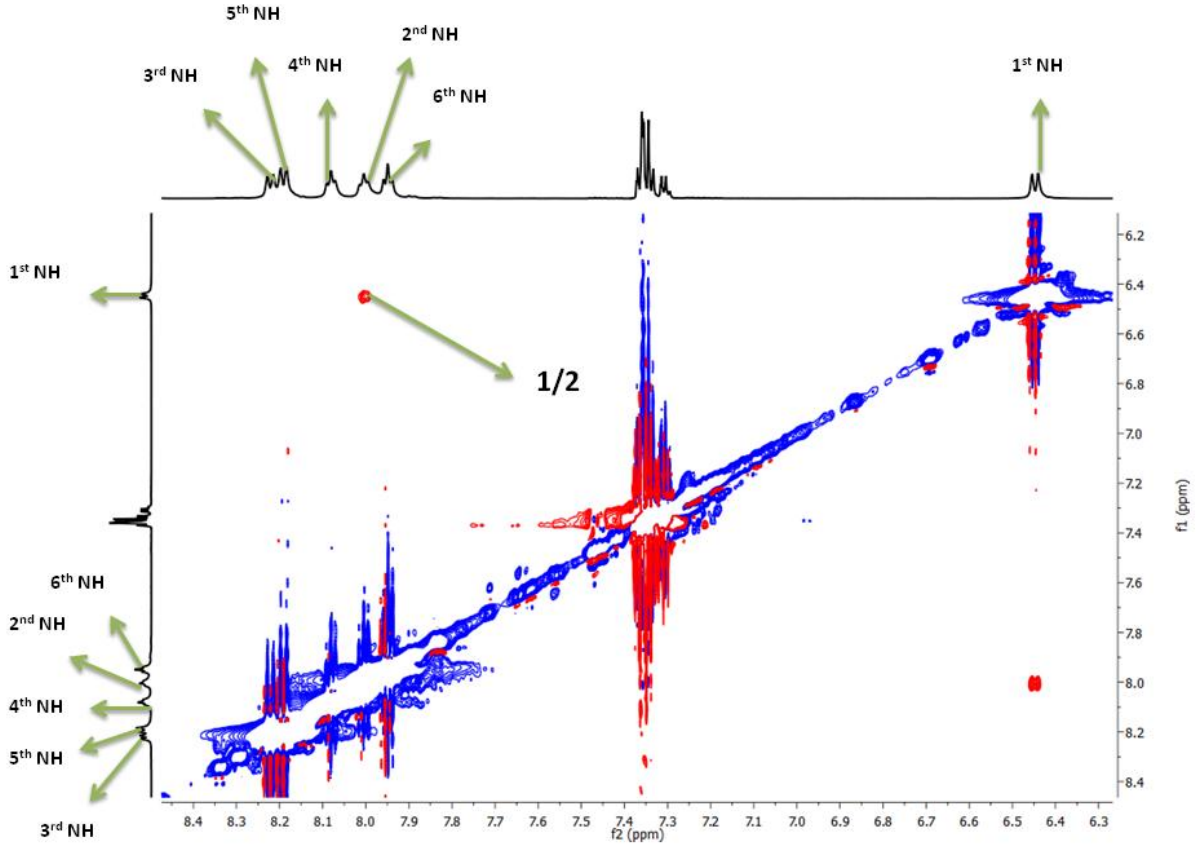


Figure S24: Partial ROESY spectrum of **P2** (5 mM) in CD_3OHshowing NH↔NH NOEs at 298K.

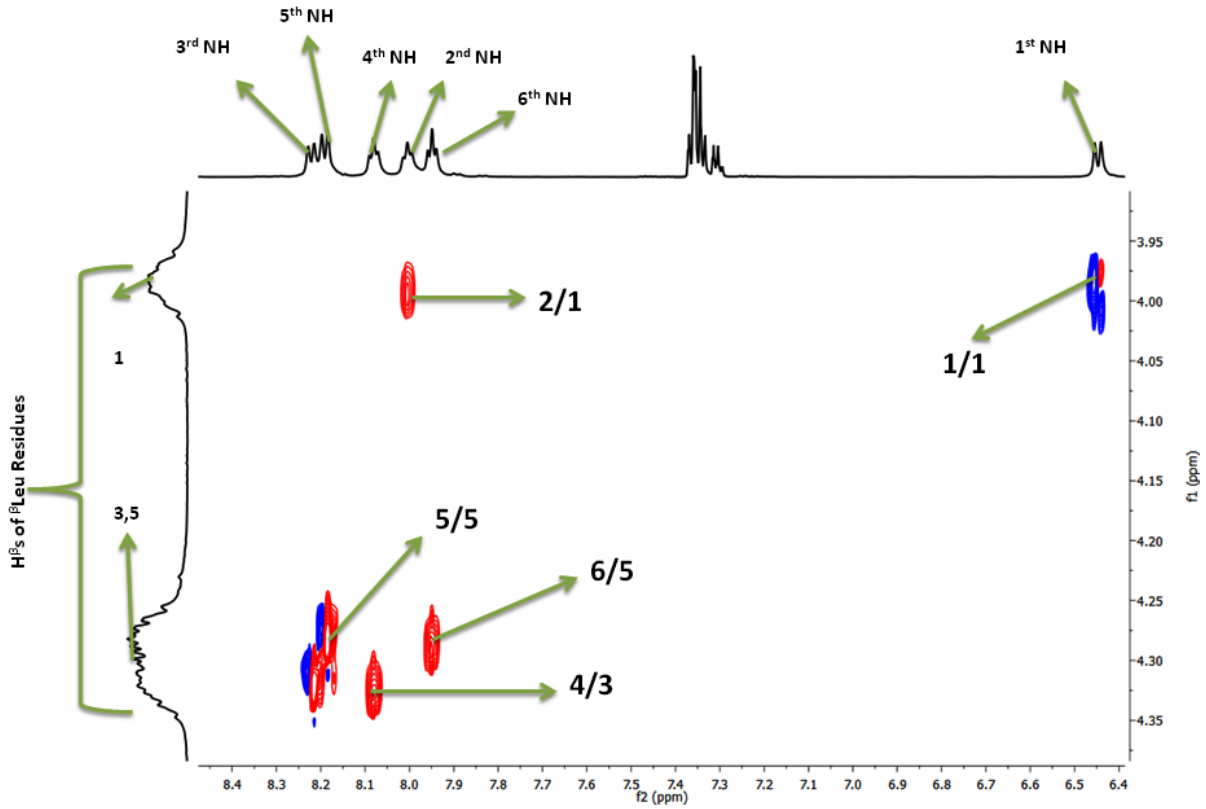


Figure S25: Partial ROESY spectrum of **P2** (5 mM) in CD_3OH showing NH↔$^{\beta}$CH NOEs of β-Leucine at 298K.

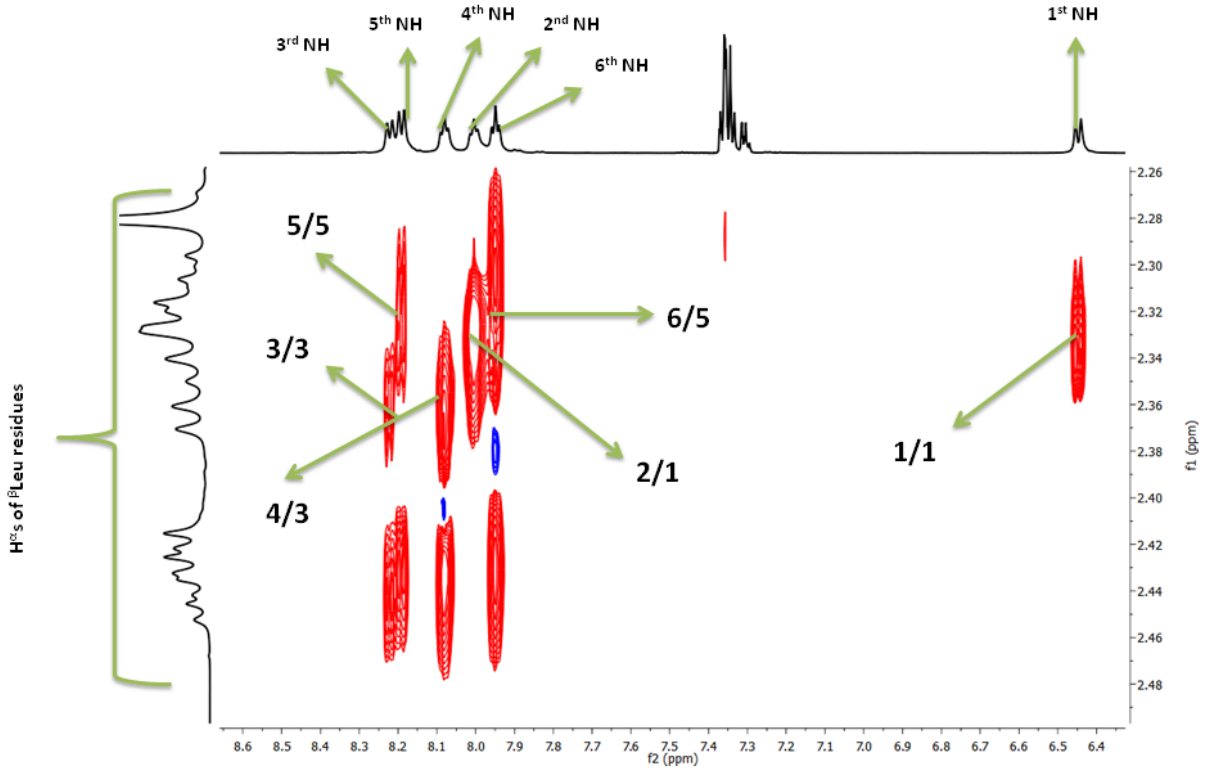


Figure S26: Partial ROESY spectrum of **P2** (5 mM) in CD_3OH showing $NH \leftrightarrow {}^{\alpha}CH_2$ NOEs of β-Leucine at 298K.

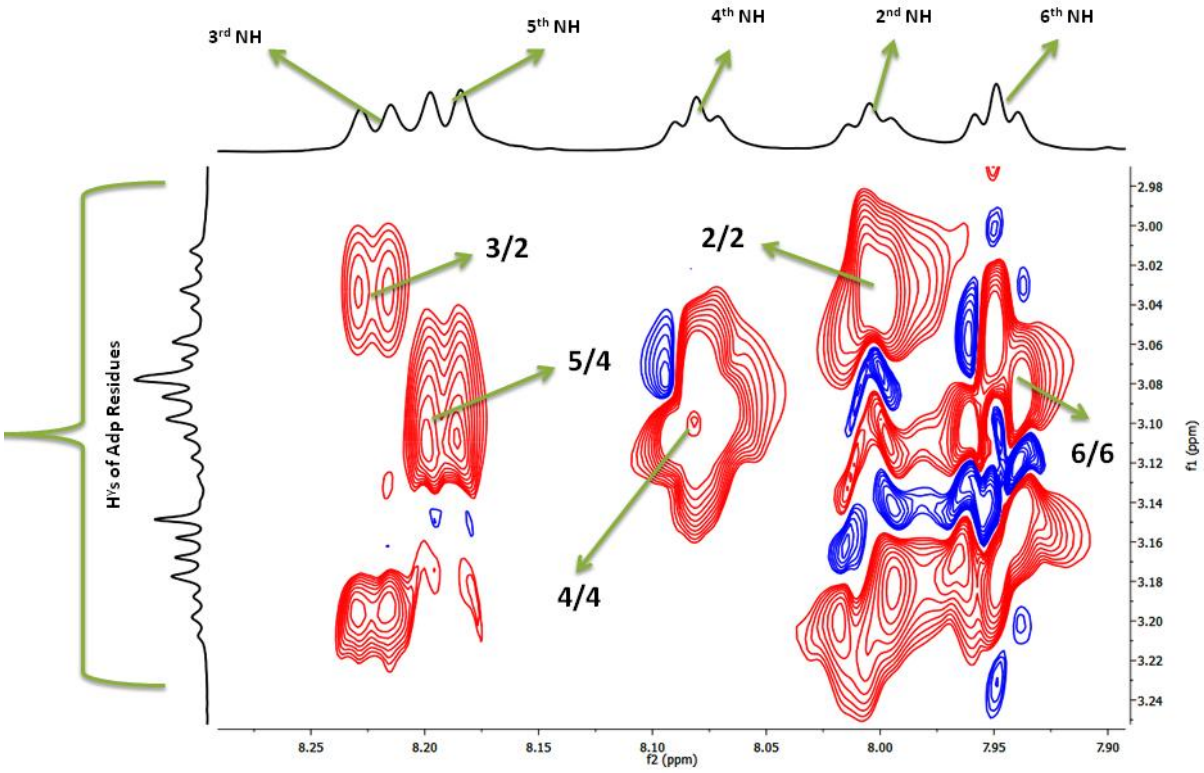


Figure S27: Partial ROESY spectrum of **P2** (5 mM) in CD_3OH howing NH↔$^{\gamma}CH_2$ NOEs of Adb at 298K.

[7]. Chemical shifts and dδ/dT values of amide NH's of P2

Table S4. Chemical shifts and dδ/dT values of amide NHs of P2 with respect to the temperature.

Residue	Temperature								dδ/dT PPBK
	275K	280K	285K	290K	295K	300K	305K	310K	
NH(1)	6.60	6.57	6.53	6.50	6.47	6.43	6.40	6.36	6.81
NH(2)	8.14	8.12	8.08	8.05	8.02	7.99	7.96	7.93	6.12
NH(3)	8.37	8.34	8.31	8.27	8.24	8.21	8.17	8.14	6.14
NH(4)	8.23	8.20	8.16	8.13	8.10	8.07	8.03	8.00	6.57
NH(5)	8.36	8.32	8.28	8.25	8.21	8.18	8.14	8.10	7.29
NH(6)	8.13	8.09	8.05	8.01	7.97	7.93	7.89	7.86	7.83

[8] Superimposition of 10 structures of peptide P2 determined by NMR in CD_3OH

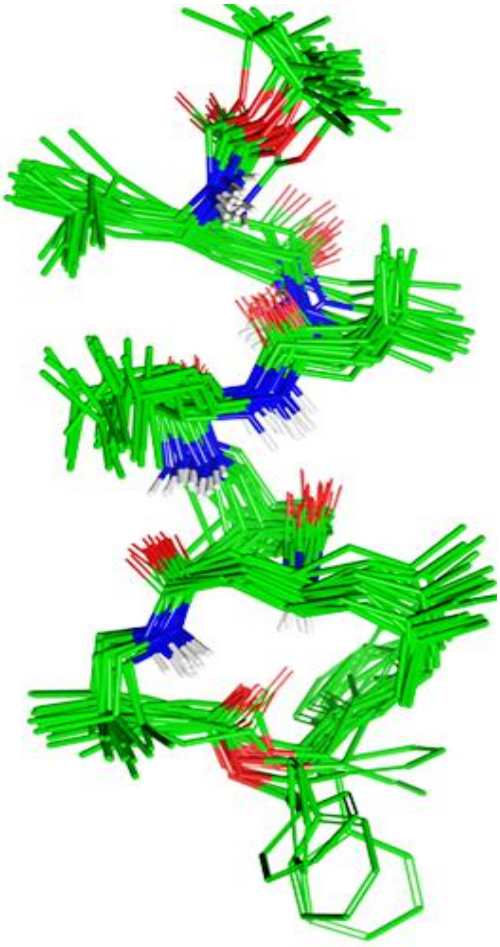


Figure S28: Backbone superposition of 10 NMR determined structures of peptide **P2**.

[9] 2D NMR details of the peptide P3 and in CD_3OH:

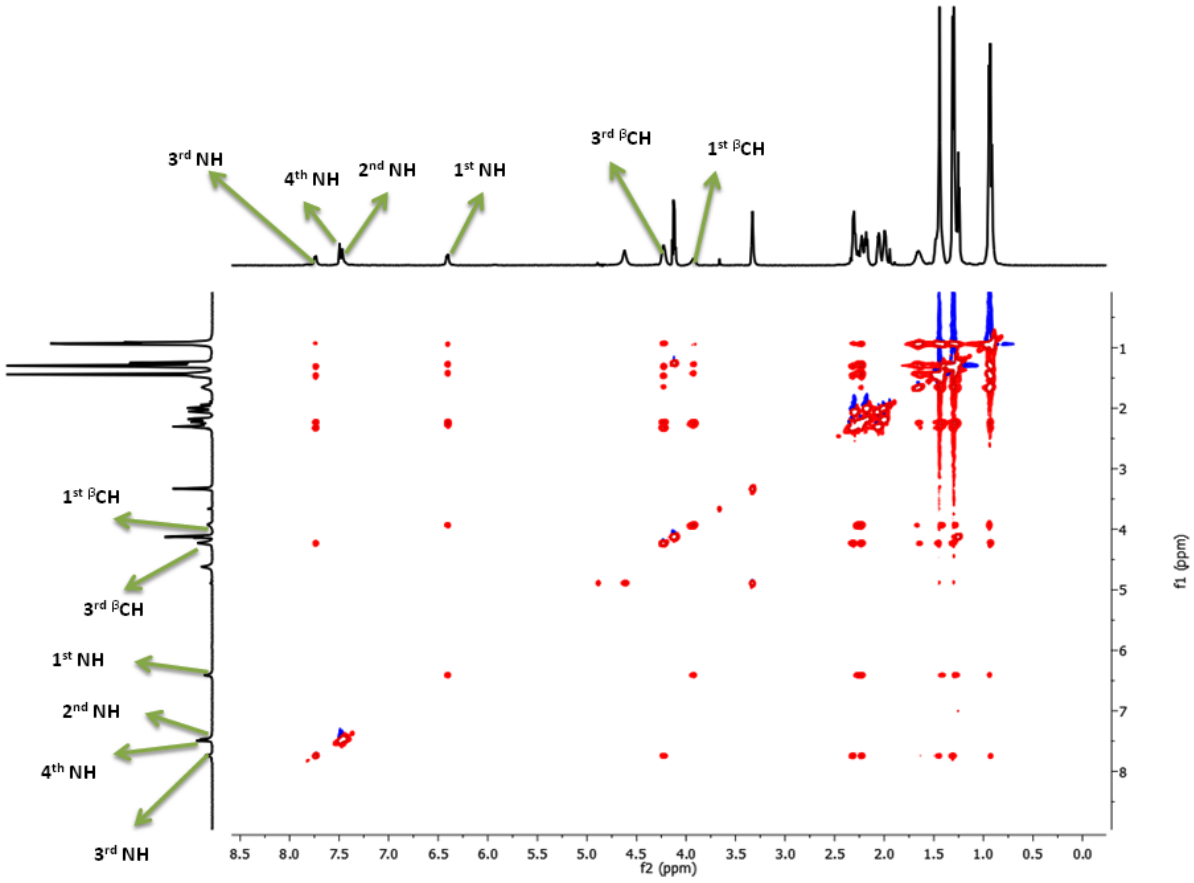


Figure S29:TOCSY spectrum of **P3** (5 mM) in CD_3OH at 298K

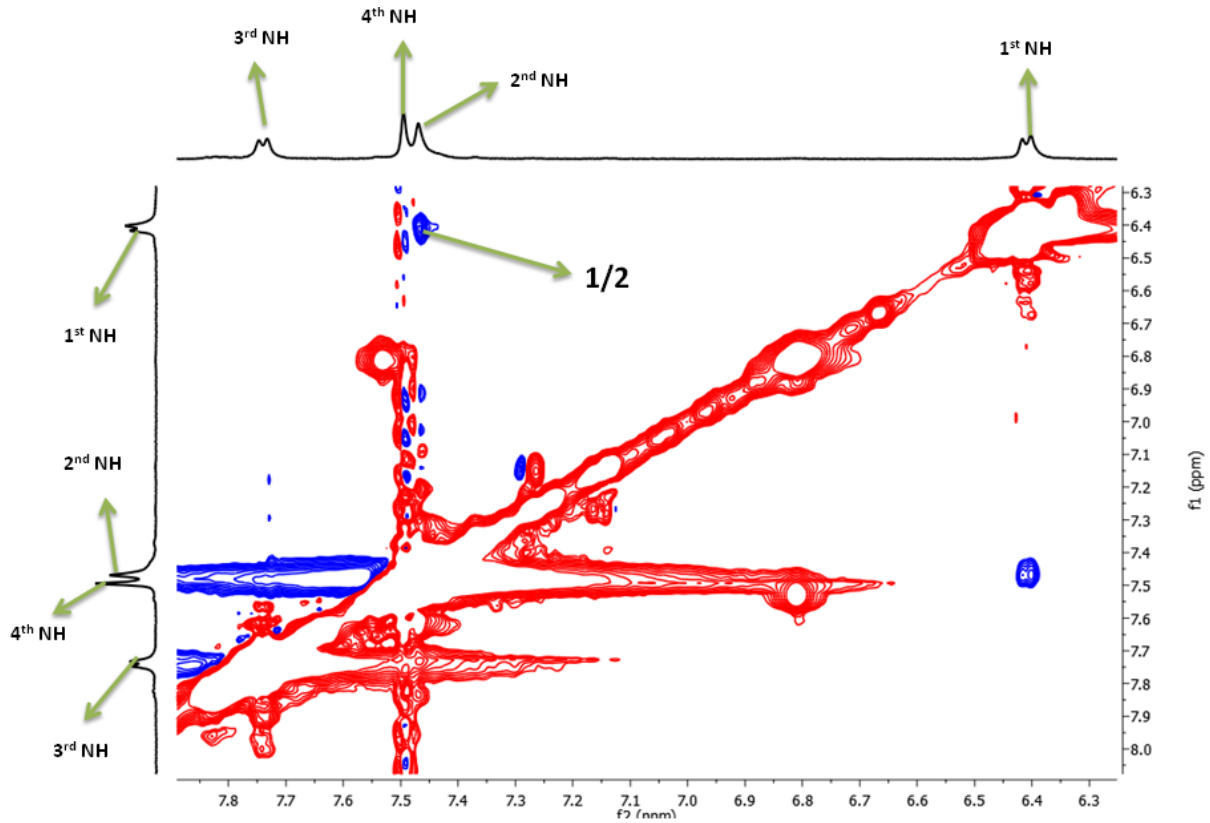


Figure S30: Partial ROESY spectrum of**P3** (5 mM) in CD_3OH showing NH↔NH NOEs.

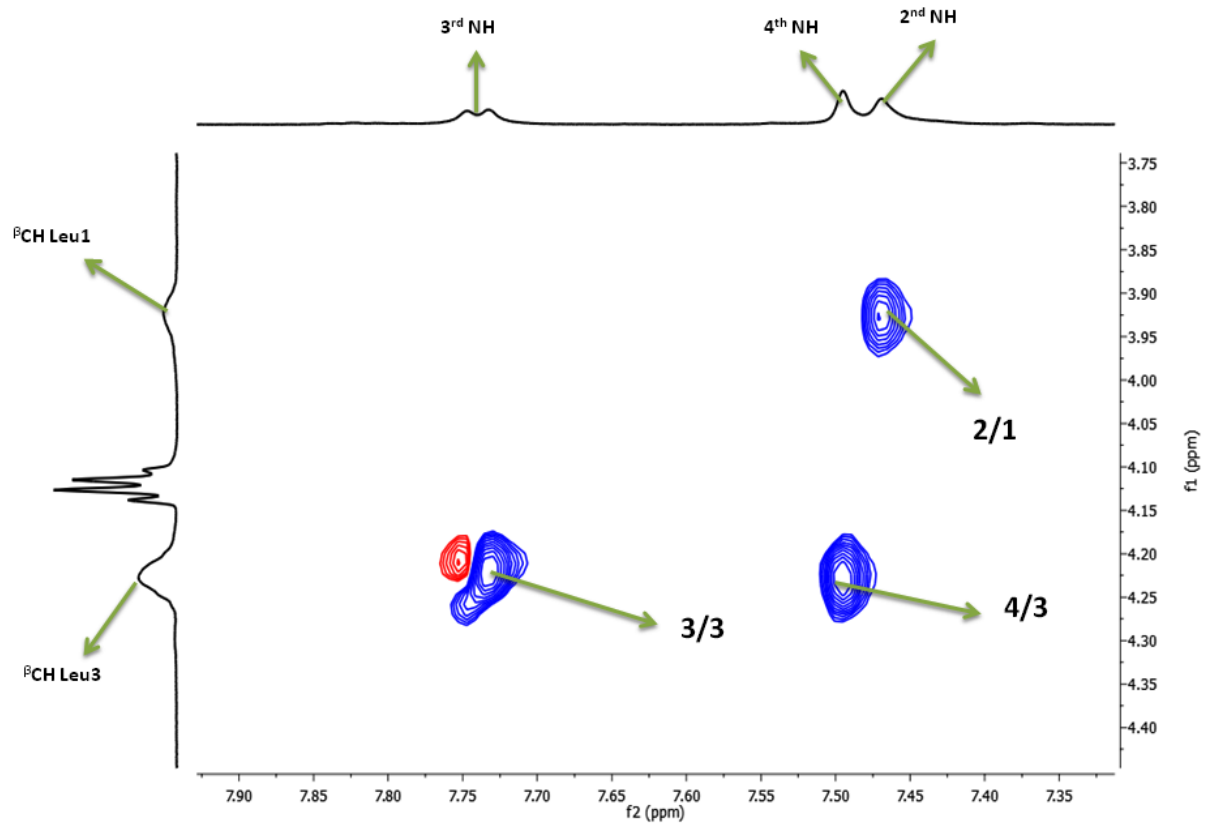


Figure S31: Partial ROESY spectrum of **P3** (5 mM) in CD_3OH showing NH↔$^{\beta}$CH of β-Leucine

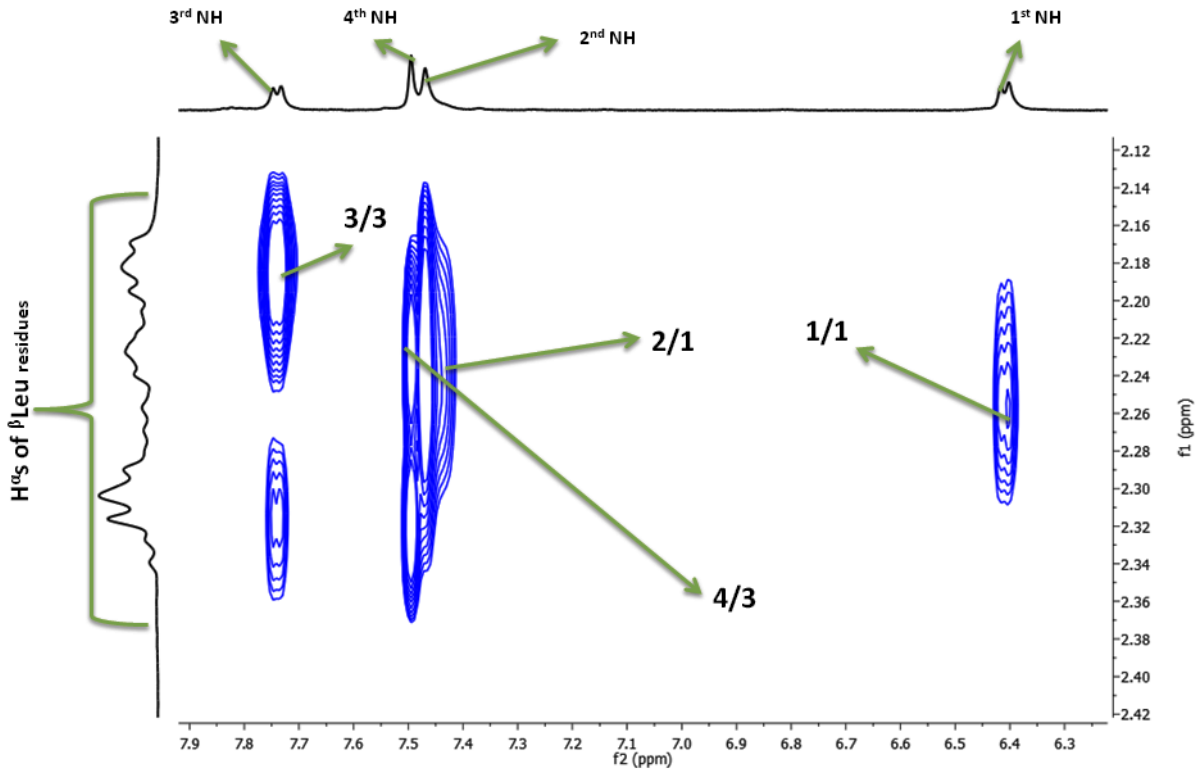


Figure S32: Partial ROESY spectrum of **P3** (5 mM) in CD_3OH showing NH↔$^{\alpha}CH_2$ of β-Leucine

[9] General Experimental Details:

All amino acids, ethyl 3,3-dimethylacrylate, nitro methane, Pd/C, TFA, EDC, HOBt, DIEPA, were commercially available. DCM, DMF, ethyl acetate and pet-ether (60-80 °C) were distilled prior to use.Column chromatography was performed on silica gel (120-200 mesh). Final peptides were purified on reverse phase HPLC (C18 column, MeOH/H_2O 70:30-95:5 as gradient with flow rate 2.5 mL/min) as well as in Column chromatography .^{1}H NMR and ^{13}CNMR spectra were recorded on 400 MHz , 600 MHz and on 100 MHz,125 MHz, respectively, using residual solvent signal as internal standards ($CDCl_3$). Chemical shifts (δ) reported in parts per million (*ppm*) and coupling constants (*J*) reported in Hz. Mass spectra were recorded using MALDITOF/TOF and HRMS Electron Spray Ionization (ESI).

Synthesis of N- Boc protected 3,3-dimethylbutanoic acid (Adb):

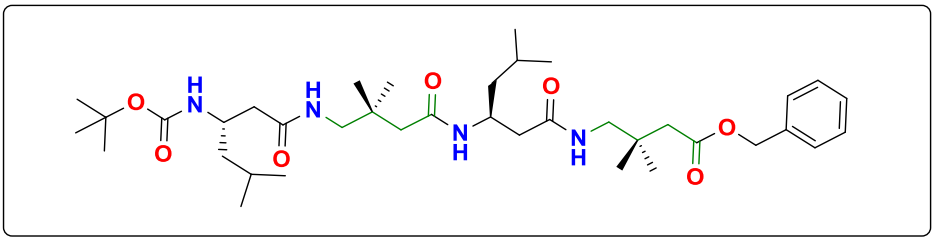
Ethyl 3,3-dimethylacrylate (6.4 g, 50 mmol) was dissolved in neat nitromethane (13.5 ml, 250 mmol) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 11.2 mL, 75 mmol) were added. The mixture was heated at 60 °C overnight, and the nitromethane was evaporated under reduced pressure. The residue was dissolved in ethyl acetate followed by 1M HCl, (200 ml) and the organic phase was separated. The acidic aqueous layer was washed twice with ethyl acetate, the combined organic phase was dried over Na_2SO_4, and the solvent was evaporated under reduced pressure. The product, 3, 3-dimethyl-4-nitro-butyric acid ethyl ester was collected (7.56 g, 80 % yield) as colorless oil. The suspension of activated Pd/C (20% by weight) and 3,3-dimethyl-4-nitro-butyric acid ethyl ester (3.78 g, 20 mmol) in MeOH (40 mL) and acetic acid(5 mL) was stirred at room temperature in the presence of hydrogen. After completion of the reaction (TLC, ~36 hrs), Pd/C was filtered through the bed of celite and the filtrate was evaporated to dryness under vacuum to get gummy 4,4-dimethyl-2-pyrrolidinone (2.14 g, 95 % yield) as oil. The amide NH group of 4,4-dimethyl-2-pyrrolidinone was further protected with Boc group and then hydrolyzed using NaOH (2.0 M) in MeOH to get final product *N*-Boc protected 3,3-dimethylbutanoic acid (3.21 g, 80 % yield in two step)

Synthesis of peptide P1-P3:

Synthesis of all the peptides was carried out by conventional solution phase methods using fragment condensation strategy. The Boc group was used for the N-terminal protection and the C-terminus was protected as a benzyl ester or ethyl esters.The Boc-(*S*)-β-Leu was synthesized by Arndt–Eistert homologation of Boc-Leu and 4-amino isocaproic acid (Aic)was synthesized reported earlier. Couplings were carried out using *N*-Ethyl-*N′*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) and 1- hydroxybenzotriazole (HOBt). The C-terminal methyl or ethyl ester group was deprotected by using aqueous sodium hydroxide. The final peptides were purifiedby column chromatography(ethyl acetate: hexane) as well as by reverse phase HPLC on C18 column using methanol/water gradient. All three peptides were characterized by ^{1}H NMR and mass spectrometry.

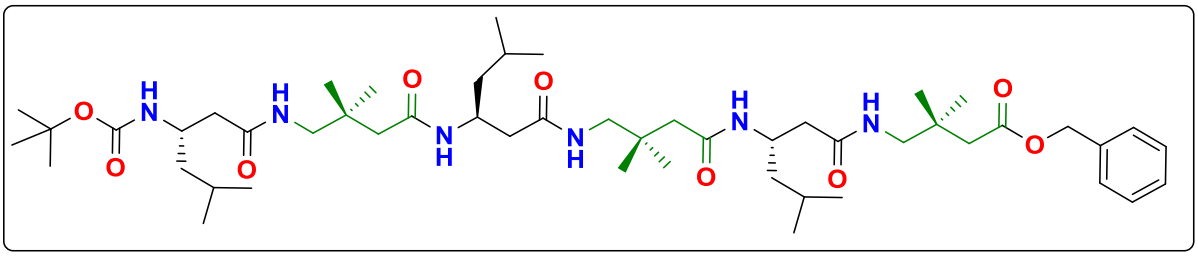
Structural data of the peptide P1, P2 and P3:

Peptide P1:



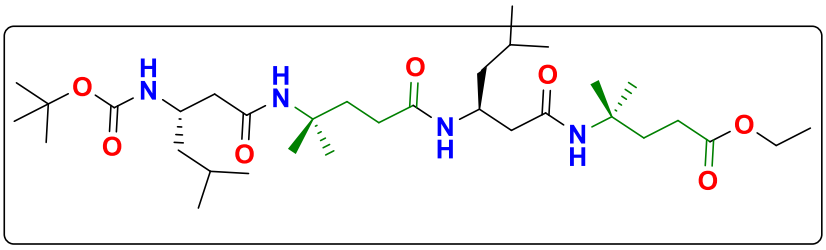
^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.42 (d, *J* = 9.4 Hz, 1H), 7.339-7.341(m,5H) , 6.61 (m, 1H), 5.37 – 5.23 (m, 1H), 4.26 (m, 1H), 3.98 (m, 1H), 3.26 (dd, *J* = 13.7, 6.8 Hz, 1H), 3.19 (dd, *J* = 13.8, 6.6 Hz, 1H), 3.16 – 3.12 (m, 1H), 3.11 (s, 1H), 2.53 – 2.44 (m, 2H), 2.35 (m, 2H), 2.31 (s, 2H), 2.15 – 2.01 (m, 2H), 1.70 – 1.59 (m, 2H), 1.55 (m, 1H), 1.37 – 1.25 (m, 3H), 1.06 – 0.95 (m, 13H), 0.96 – 0.88 (m, 13H).

Peptide P2:



^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.56 (t, *J* = 6.3 Hz, 1H), 7.54 – 7.48 (d, 6 Hz, 1H), 7.42 – 7.33 (m, 7H), 7.25 (d, *J* = 7.2 Hz, 1H), 6.62 (t, *J* = 6.4 Hz, 1H), 5.31 (d, *J* = 9.4 Hz, 1H), 4.27 (m, 2H), 3.98 (d, *J* = 9.8 Hz, 1H), 3.27 (m, 1H), 3.14 (m, 6H), 2.49 (m, 3H), 2.43 – 2.32 (m, 3H), 2.31 (s, 2H), 2.08 (m, 4H), 1.71 – 1.49 (m, 7H), 1.32 (m, 4H), 1.03 – 0.96 (m, 20H), 0.93 (m, 20H).

Peptide P3:



^{1}H NMR (600 MHz, Chloroform-*d*) δ 6.83 – 6.72 (m, 2H), 6.13 (s, 1H), 5.45 (d, *J* = 8.8 Hz, 1H), 4.26 – 4.18 (m, 2H), 3.88 (hept, *J* = 5.4, 4.4 Hz, 1H), 2.43 – 2.19 (m, 12H), 2.08 – 1.90 (m, 7H), 1.70 – 1.56 (m, 4H), 1.51 – 1.38 (m, 16H), 1.37 – 1.30 (m, 19H), 0.93 (dd, *J* = 6.6, 4.2 Hz, 20H).

[10] FT-IR spectroscopy of peptide P1, P2 and P3in $CDCl_3$:

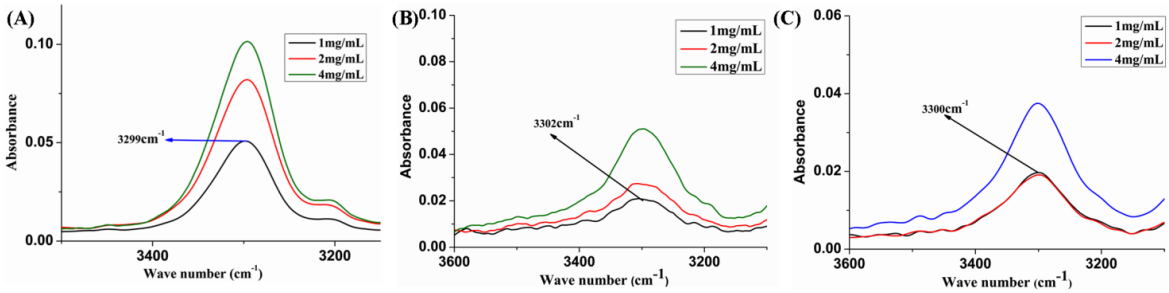
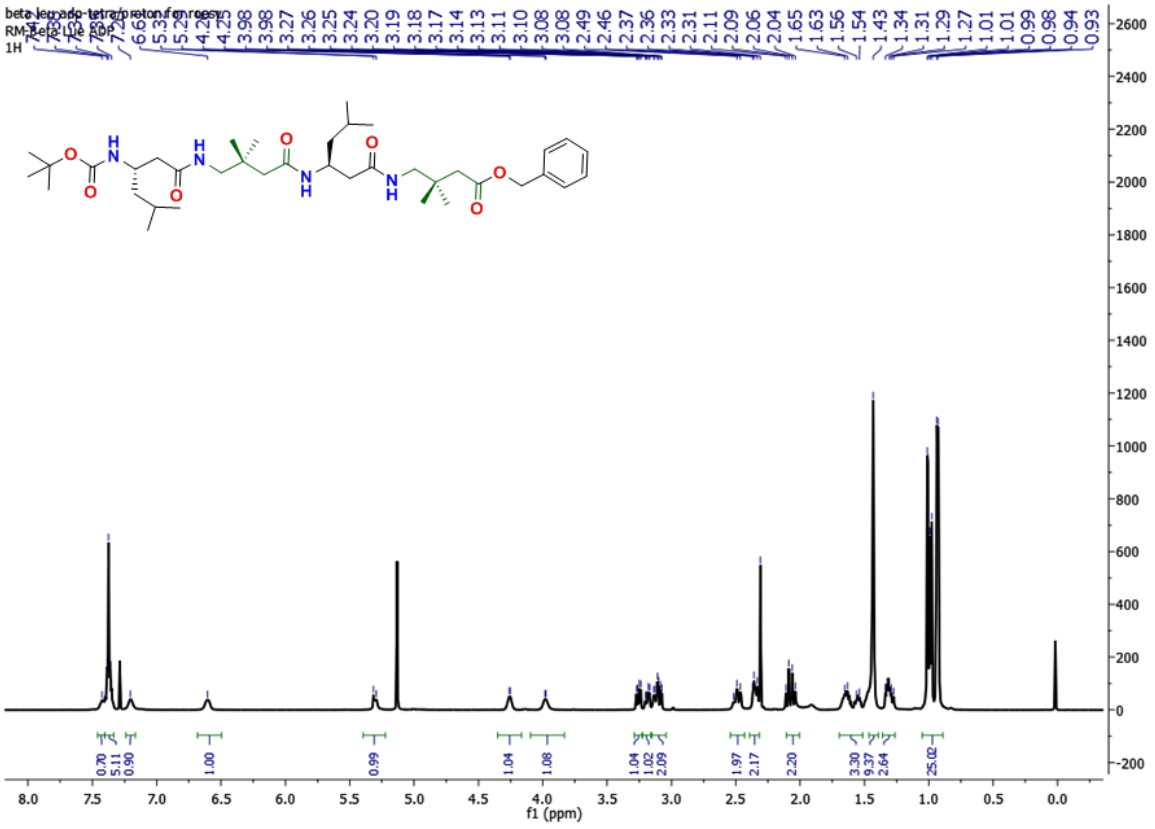
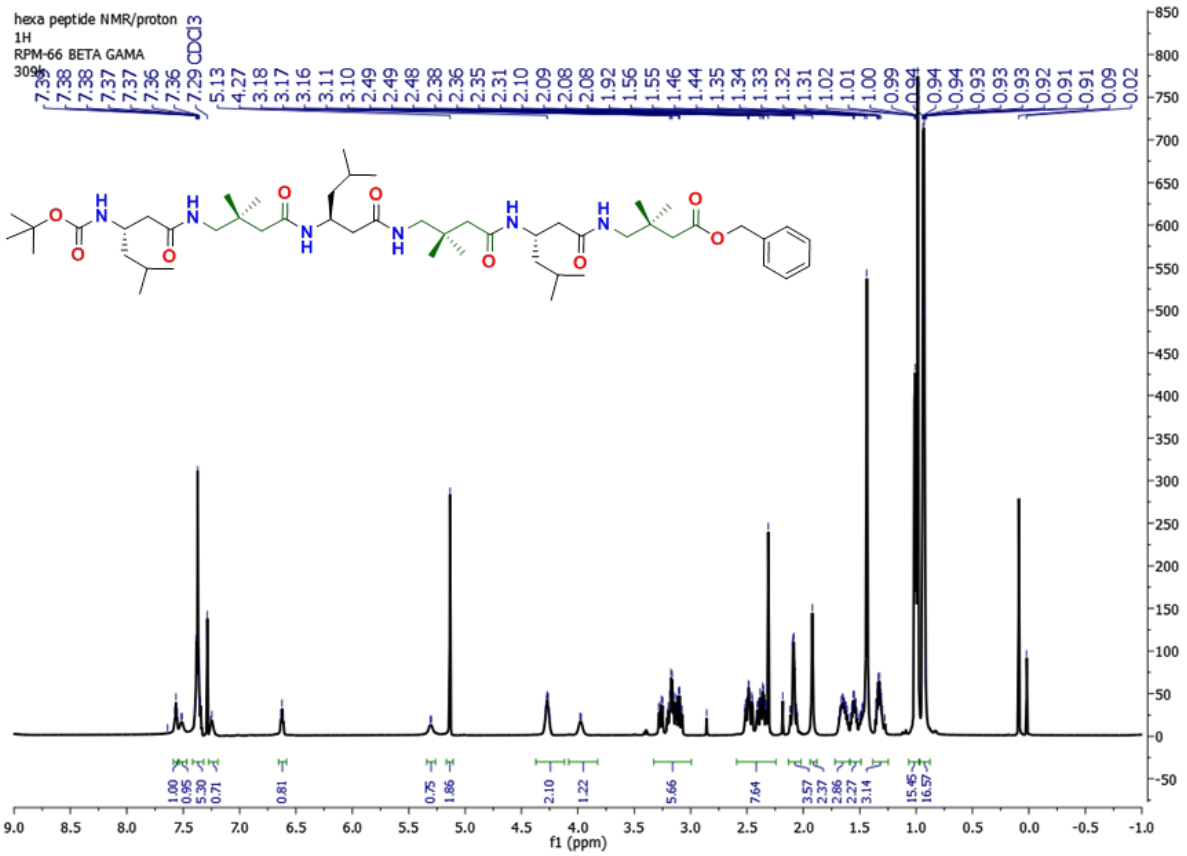


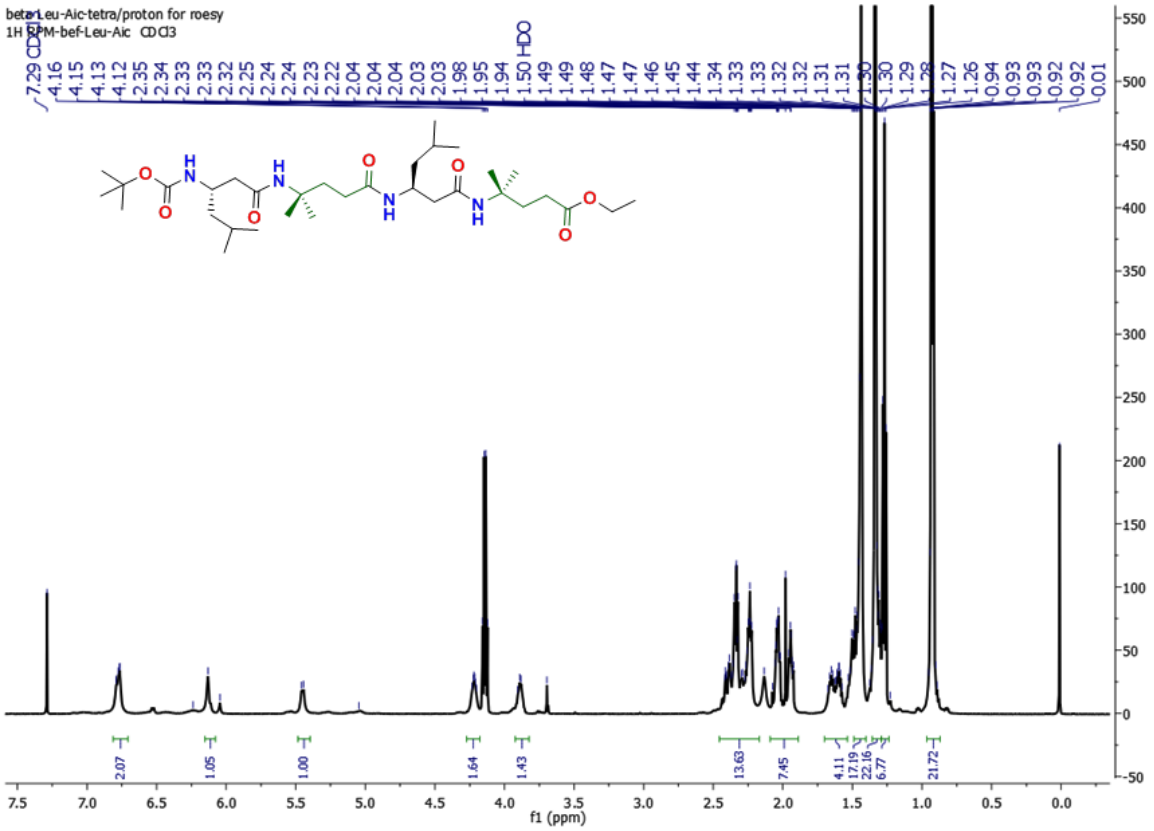
Figure S33: (A, B, C) FT-IR spectroscopy of peptide **P1**, **P2** and **P3** at different concentration in $CDCl_3$

[

[16]^{1}H and Mass Spectra of the peptide P1, P2 and P3 :

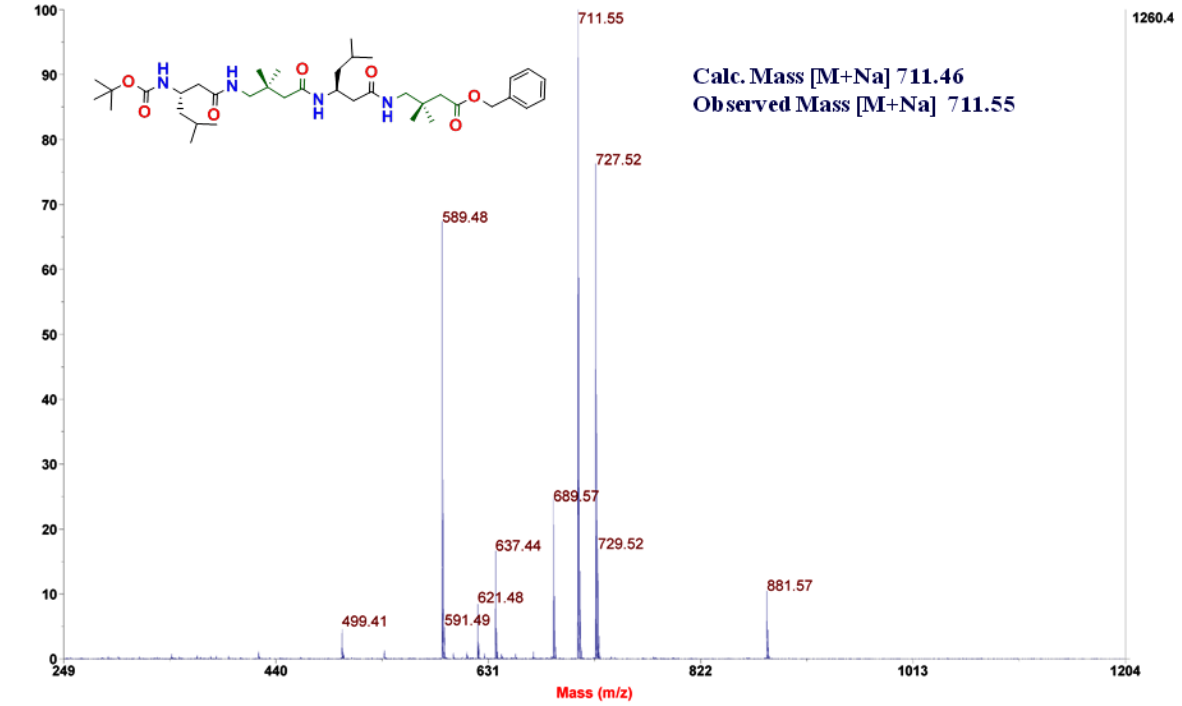



hexa peptide NMR/proton
1H
RPM-66 BETA GAMA
309k
CDCl3
f1 (ppm)


beta-Leu-Aic-tetra/proton for roesy
1H NMM-bef-Leu-Aic CDCl3
7.29 CDCl3
1.50 HDO
f1 (ppm)

Spectrum Report

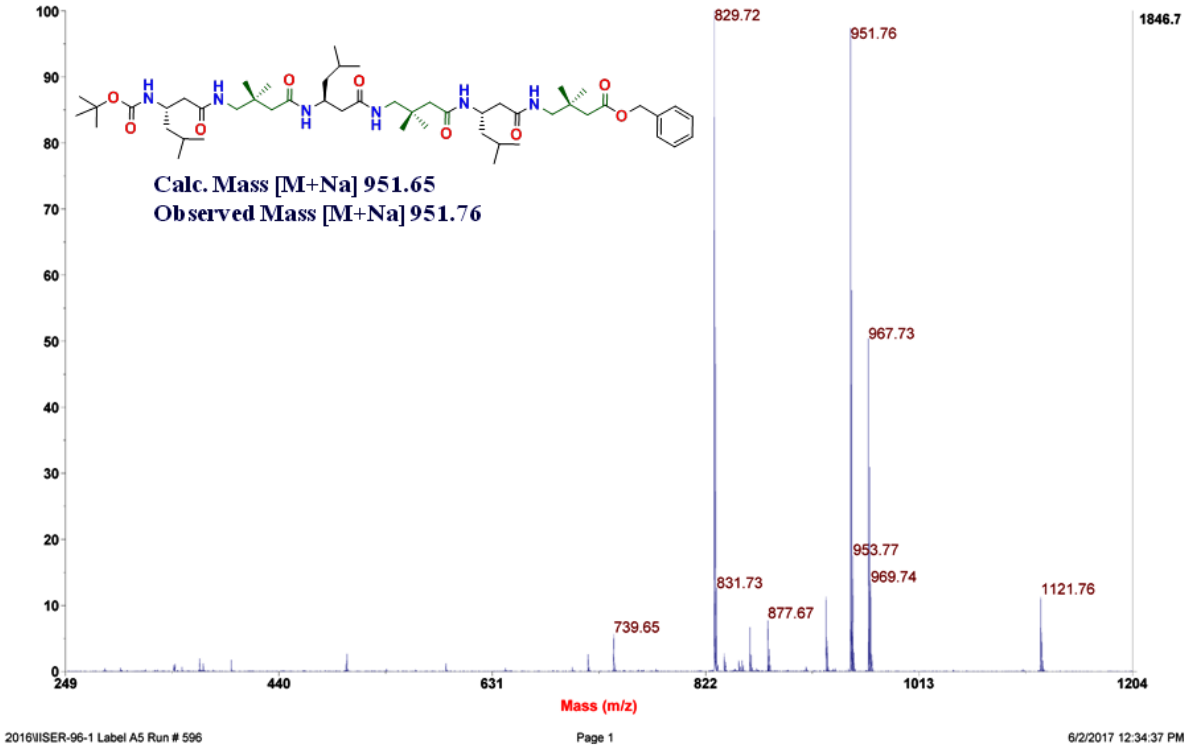
Final - Shots 1000 - IISER-96-1; Run #596; Label A4



2016\IISER-96-1 Label A4 Run # 596 Page 1 6/2/2017 12:33:26 PM

Spectrum Report

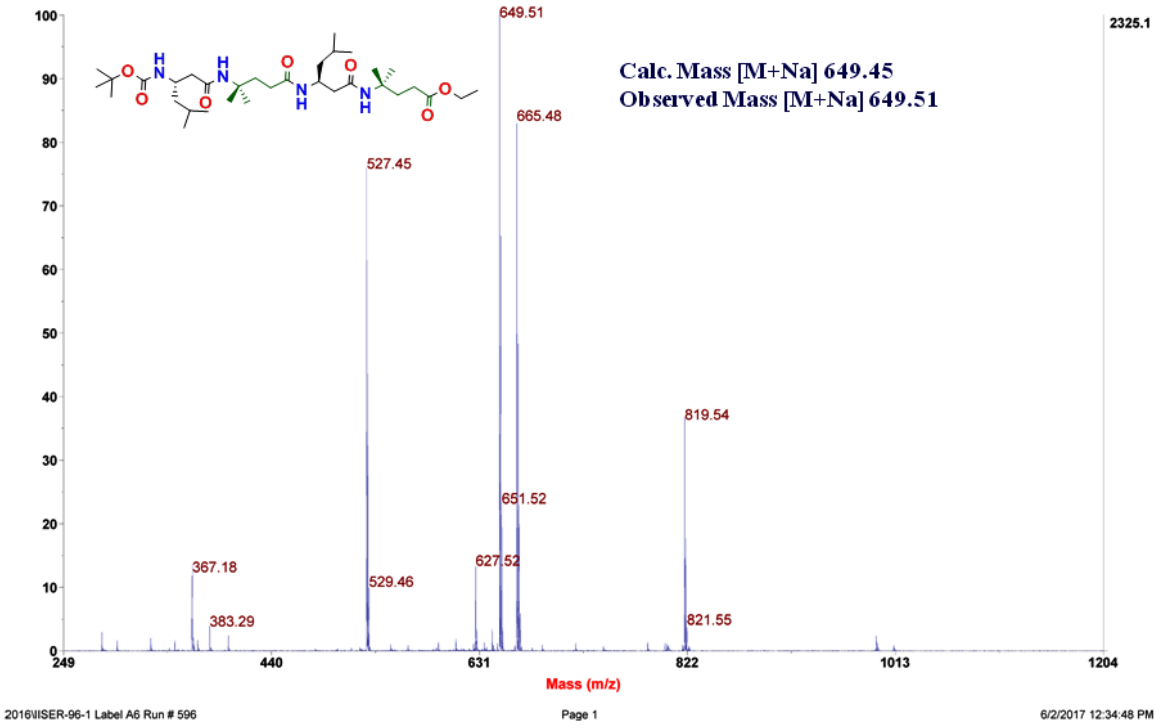
Final - Shots 1000 - IISER-96-1; Run #596; Label A5



2016\IISER-96-1 Label A5 Run # 596 Page 1 6/2/2017 12:34:37 PM

Spectrum Report

Final - Shots 1000 - IISER-96-1; Run #596; Label A6



2016\IISER-96-1 Label A6 Run # 596 Page 1 6/2/2017 12:34:48 PM

[17] References:

1. Desmond Molecular Dynamics System, version 4.8, D. E. Shaw Research, New York, NY, 2016