Electronic Supplementary Material (ESI) for Chemical Communications

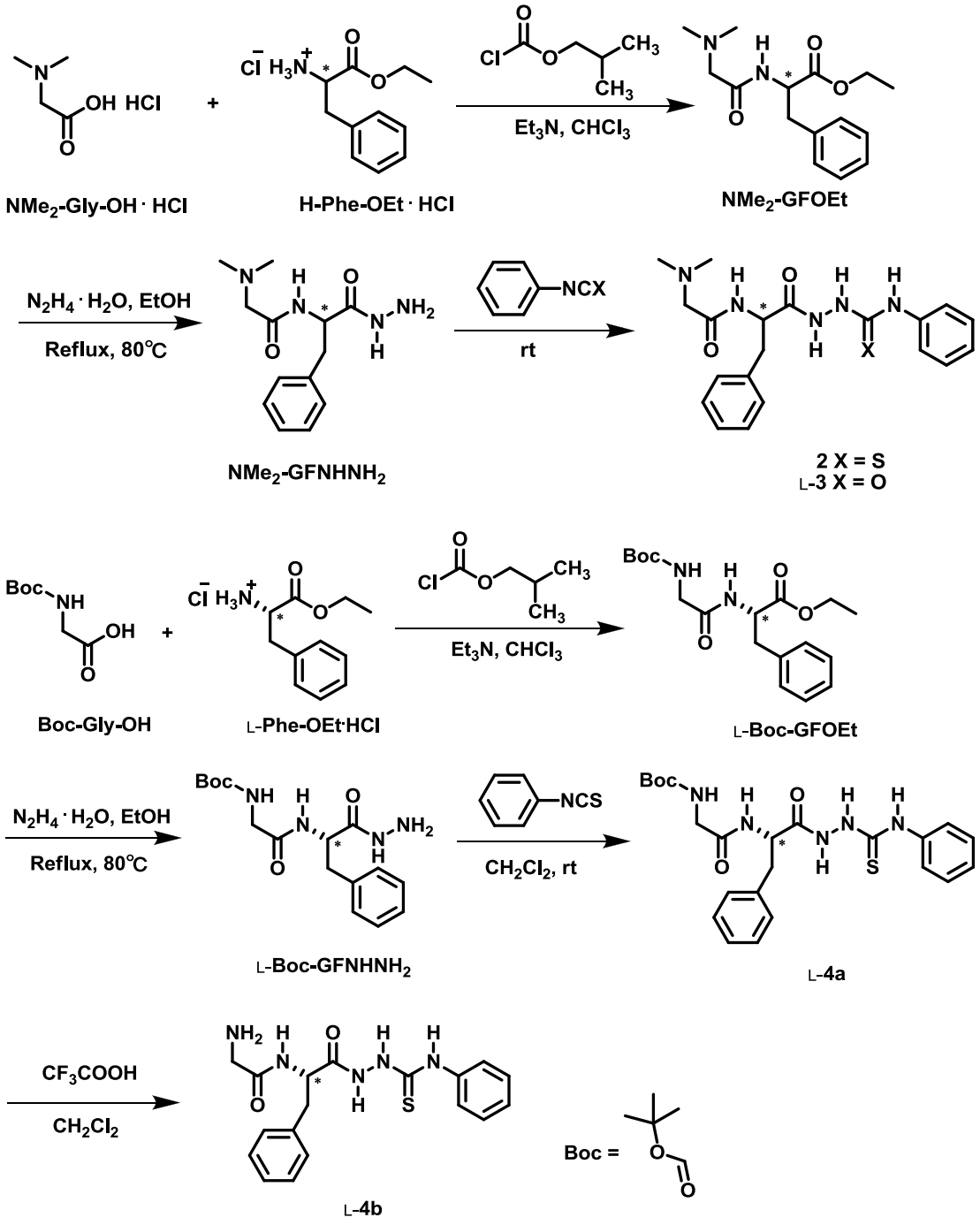
β-Turn Structure in Glycinylphenylalanine Dipeptide Based *N*-Amidothioureas

Xiao-Sheng Yan, Kun Wu, Yuan Yuan, Ying Zhan, Jin-He Wang, Zhao Li and Yun-Bao Jiang*

Department of Chemistry, College of Chemistry and Chemical Engineering, and the MOE Key Laboratory of Analytical Sciences, Collaborative Innovation Centre of Chemistry for Energy Materials (iChEM), Xiamen University, Xiamen 361005, China.

Tel: +86 592 218 8372; Fax: +86 592 218 5662; E-mail: ybjiang@xmu.edu.cn

Electronic Supplementary Information (ESI)



Scheme S1. Syntheses of **2**, L-**3**, L-**4a** and L-**4b**

NMe_2-GFOEt: To a chilled solution of NMe_2-Gly-OH·HCl (0.56 g, 4.0 mmol) and Et_3N (0.62 mL, 4.4 mmol) in $CHCl_3$ (10 mL) was added isobutyl chloroformate (0.65 mL, 5.0 mmol) at 0 ˚C. After 30 min, a solution of (L or D) H-Phe-OEt·HCl (0.69g, 3.0 mmol) and Et_3N (0.46 mL, 3.3 mmol) in $CHCl_3$ (20 mL) was added. The mixture was left to stand at room temperature for 4 hours, evaporated *in vacuo*, and the solid residue was dissolved in AcOEt. The solution was washed successively with 1% $NH_3·H_2O$, saturated NH_4Cl and water, dried over anhydrous Na_2SO_4

and concentrated under reduced pressure, to obtain 0.76 g oily product **NMe_2-GFOEt**, 91.0%.

NMe_2-GFNHNH$_2$: Excess aqueous hydrazine (80%) was added to **NMe_2-GFOEt** in ethanol (15.0 mL) and then refluxed for 24 hours, evaporated *in vacuo*, and the viscous liquid was dissolved in water. The solution was extracted by CH_2Cl_2, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure, to obtain 0.56 g **NMe_2-GFNHNH$_2$**, 77.8%.

2: **NMe_2-GFNHNH$_2$** then reacted with excess phenyl isothiocyanate in CH_2Cl_2. The solution was stirred at room temperature for 24 hours, after which the solvent was removed under reduced pressure. The product was washed by petroleum ether, water and hot diethyl ether to obtain 0.59 g **2**, 69.7%.

L-**2:** ^{1}H NMR (400 MHz, DMSO-d_6): δ (ppm) 10.30 (s, 1H), 9.78 (s, 1H), 9.33 (s, 1H), 7.98 (s, 1H), 7.53 (d, J = 7.7 Hz, 2H), 7.37 – 7.11 (m, 8H), 4.59 (s, 1H), 3.20 (dd, J = 13.9, 4.5 Hz, 1H), 3.05 – 2.88 (m, 2H), 2.78 (d, J = 15.5 Hz, 1H), 2.07 (s, 6H); ^{13}C NMR (101 MHz, CD_3CN): δ (ppm) 182.00, 172.27, 170.68, 138.73, 136.77, 129.28, 128.65, 128.36, 126.94, 125.74, 124.98, 62.37, 54.01, 45.04, 36.31; HRMS (ESI): calcd for $[C_{20}H_{26}N_5O_2S]^+$: 400.1802, found: 400.1807. Crystal data for compound L-**2**: C20H25N5O2S, M_r = 399.51, T = 293(2) K, Orthorhombic, space group $P2_12_12_1$, a = 8.7524(4), b = 15.1376(7), c = 15.8697(8) Å, V= 2102.58(17) Å^3, ρ_c = 1.262 Mg.m^{-3}, $\mu(Mo_{K\alpha})$ = 0.179 mm^{-1}, Z = 4, reflections collected: 5437, independent reflections: 3517 (R_{int} = 0.0211), S = 0.814, final R indices [I > 2σ(I)]: R_1 = 0.0360, wR_2 = 0.1040, R indices (all data): R_1 = 0.0411, wR_2 = 0.1088, Flack parameter = -0.03(8), CCDC No. 914552.

D-**2**: ^{1}H NMR (400 MHz, DMSO-d_6): δ (ppm) 10.31 (s, 1H), 9.78 (s, 1H), 9.33 (s, 1H), 7.97 (s, 1H), 7.52 (d, J = 7.7 Hz, 2H), 7.41 – 7.11 (m, 8H), 4.59 (s, 1H), 3.20 (dd, J = 13.9, 4.6 Hz, 1H), 3.00 – 2.89 (m, 2H), 2.78 (d, J = 15.5 Hz, 1H), 2.07 (s, 6H); ^{13}C NMR (101 MHz, CD_3CN): δ (ppm) 182.00, 172.26, 170.71, 138.73, 136.77, 129.29, 128.65, 128.36, 126.94, 125.75, 125.00, 62.36, 53.99, 45.04, 36.33; HRMS (ESI): calcd for $[C_{20}H_{26}N_5O_2S]^+$: 400.1802, found: 400.1809.

L-**3:** L-**NMe_2-GFNHNH$_2$** (0.53 g, 2.0 mmol) then reacted with equivalent phenyl isocyanate (0.24 g, 2.0 mmol) in 50 mL CH_3CN. The solution was stirred at room temperature for 3 hours, after which the solvent was removed under reduced pressure. The product was washed by anhydrous ether and recrystallized in CH_3CN to obtain 0.61 g L-**3**, 79.6%.

L-**3:** ^{1}H NMR (500 MHz, CD_3CN) : δ (ppm) 8.41 (s, 1H), 7.86 (s, 1H), 7.60 (d, J = 4.8 Hz, 1H), 7.53 (dd, J = 8.6, 0.9 Hz, 2H), 7.33 (dd, J = 10.0, 4.6 Hz, 2H), 7.31 – 7.21 (m, 5H), 7.07 – 6.96 (m, 1H), 6.67 (s, 1H), 4.42 (dt, J = 9.0, 6.0 Hz, 1H), 3.21 (dd, J = 14.0, 5.7 Hz, 1H), 3.02 (dd, J = 13.9, 9.1 Hz, 1H), 2.94 (d, J = 16.3 Hz, 1H), 2.81 (d, J = 16.3 Hz, 1H), 2.13 (s, 5H); ^{13}C NMR (126 MHz, MeOD): δ (ppm) 172.40, 171.88, 156.62, 138.60, 136.58, 129.97, 128.32, 128.24, 126.60, 122.88, 119.65, 61.91, 53.22, 44.50, 37.05; HRMS (ESI): calcd for $[C_{20}H_{26}N_5O_3]^+$: 384.2036,

found: 384.2036. Crystal data for compound L-**3**: C20H25N5O3, M_r = 383.45, T = 273(2) K, Orthorhombic, space group $P2_12_12_1$, a = 8.6444(15), b = 15.088(3), c = 15.404(3) Å, V= 2009.1(6) $Å^3$, ρ_c = 1.268 $Mg.m^{-3}$, $\mu(Mo_{K\alpha})$ = 0.088 mm^{-1}, Z = 4, reflections collected: 11532, independent reflections: 4677 (R_{int} = 0.0451), S = 0.1033, final R indices [I > 2σ(I)]: R_1 = 0.0443, wR_2 = 0.1172, R indices (all data): R_1 = 0.0470, wR_2 = 0.1193, CCDC No. 926794.

L-**Boc-GFOEt**: To a chilled solution of Boc-Gly-OH (1.05 g, 6.0 mmol) and Et_3N (0.84 mL, 6.0 mmol) in $CHCl_3$ (10 mL) was added isobutyl chloroformate (1.0 mL, 7.5 mmol) at 0 °C. After 30 min, a solution of *L*-Phe-OEt·HCl (1.05 g, 4.5 mmol) and Et_3N (0.63 mL, 4.5 mmol) in $CHCl_3$ (20 mL) was added. The mixture was left to stand at room temperature for 12 hours, evaporated *in vacuo*, and the residue was dissolved in AcOEt. The solution was washed successively with 1% $NH_3·H_2O$, saturated NH_4Cl, 1% HCl and water, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure, to obtain 1.35 g L-**Boc-GFOEt**, 85.6%.

L-**Boc-GFNHNH$_2$:** Excess aqueous hydrazine (80%) was added to L-**Boc-GFOEt** in ethanol (20.0 mL) and then refluxed for 24 hours, evaporated *in vacuo*, and the residue was washed by AcOEt and recrystallized in ethanol to obtain 0.98 g L-**Boc-GFNHNH$_2$**, yield 76.0%.

L-**4a:** L-**Boc-GFNHNH$_2$** then reacted with phenyl isothiocyanate in ethanol and then refluxed for 24 hours, evaporated *in vacuo*. The oily residue was washed by petroleum ether and then obtain solid product, which was washed by a little CH_2Cl_2 to obtain 1.08 g L-**4a**, yield 78.8%.

L-**4b:** 10 mL CH_2Cl_2 and 10 mL CF_3COOH was added in 0.5 g L-**4a** and then the solution was stirred at room temperature for 5 hours, evaporated *in vacuo*. The oily residue was dissolved in water. Then adjust the pH value to 8-9 by saturated $NaHCO_3$ solution. The solution was extracted by CH_2Cl_2, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure, to obtain 0.30 g L-**4b**, yield 76.1%.

L-**4a:** ^{1}H NMR (400 MHz, CD_3CN): δ (ppm) 8.84 (s, 1H), 7.61 (d, J = 7.8 Hz, 2H), 7.44 – 7.18 (m, 8H), 7.13 (s, 1H), 5.72 (s, 1H), 4.37 (dt, J = 8.7, 5.8 Hz, 1H), 3.70 – 3.53 (m, 2H), 3.21 (dd, J = 14.0, 5.4 Hz, 1H), 3.04 (dd, J = 14.0, 8.9 Hz, 1H), 1.37 (s, 9H); ^{13}C NMR (101 MHz, CD_3CN): δ (ppm) 181.85, 171.84, 170.45, 156.65, 138.73, 136.84, 129.29, 128.59, 128.35, 126.89, 125.64, 124.73, 79.63, 54.83, 43.92, 36.01, 27.53; HRMS (ESI): calcd for $[C_{23}H_{29}N_5O_4SNa]^+$: 494.1832, found: 494.1834.

L-**4b**: ^{1}H NMR (500 MHz, CD_3CN): δ (ppm) 9.08 (s, 1H), 7.91 (s, 1H), 7.57 – 7.55 (m, 2H), 7.35 – 7.17 (m, 8H), 4.41 (s, 1H), 3.21 – 3.13 (m, 3H), 3.01 (dd, J = 13.9, 8.7 Hz, 1H); ^{13}C NMR (101 MHz, CD_3CN): δ (ppm) 181.37, 174.47, 170.91, 138.61, 136.56, 129.18, 128.54, 128.31, 126.89, 125.65, 124.92, 54.06, 43.79, 36.65; HRMS (ESI): calcd for $[C_{18}H_{22}N_5O_2S]^+$: 372.1489, found: 372.1489.

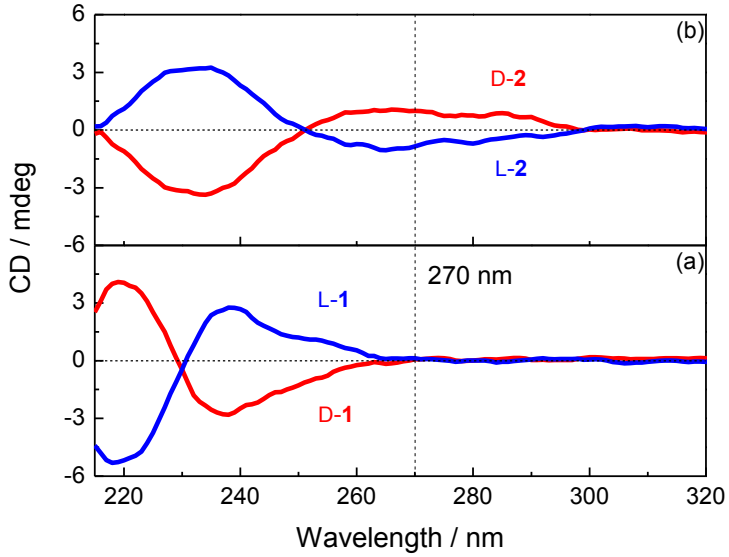


Figure S1. CD spectra of **1** (a) and **2** (b) in CH_3CN. [**1**] = [**2**] = 40 μM.

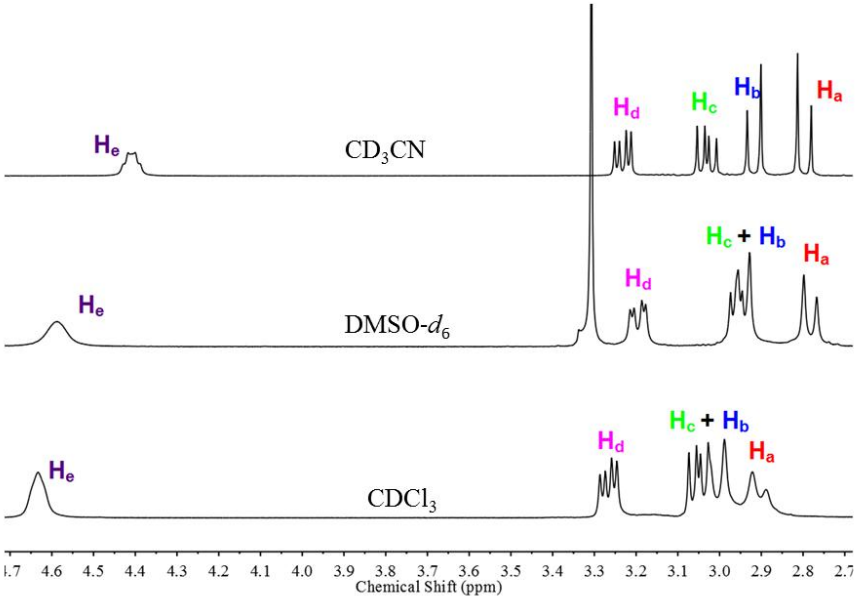
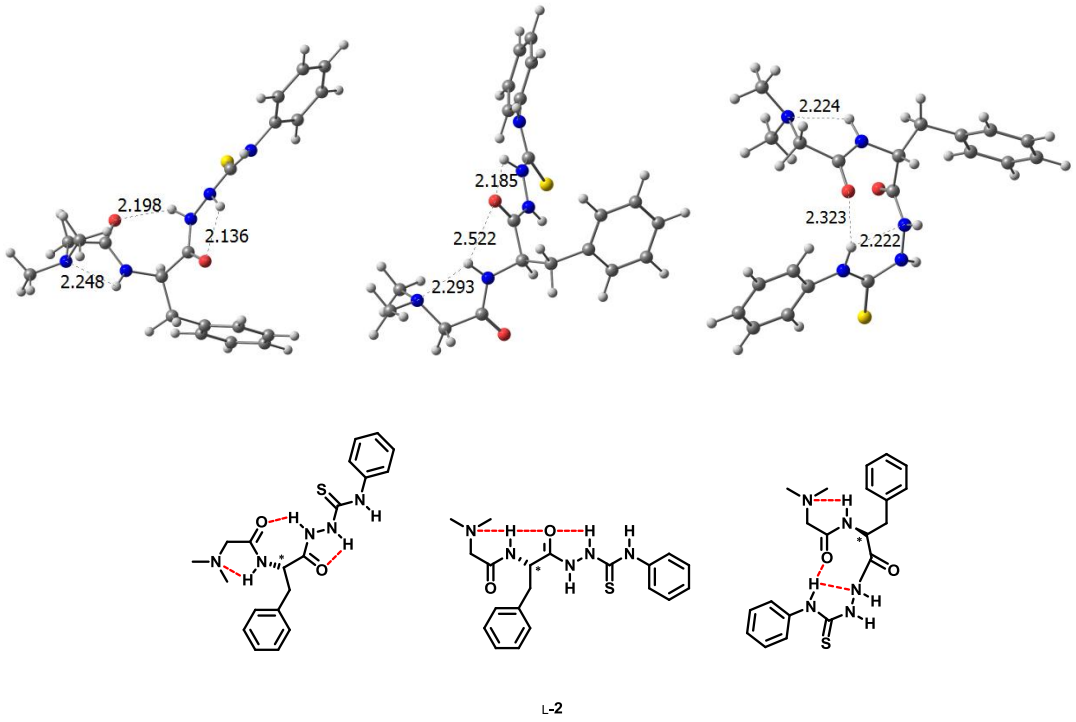


Figure S2. ^{1}H NMR spectra of L-**2** in CD_3CN, DMSO-d_6 and $CDCl_3$. [L-**2**] = 10 mM.

Table S1. Torsions in the X-ray crystal of L-**2**

ω_i	φ_{i+1}	ψ_{i+1}	ω_{i+1}	φ_{i+2}	ψ_{i+2}	ω_{i+2}
176.63°	**-73.47°**	**117.51°**	**-173.39°**	**77.92°**	**14.84°**	**-178.10°**



L-2

3.2(4.1) 1.5(3.1) 0.0(0.0)

ΔG (ΔE), kcal/mol

Figure S3. Calculated structures of L-**2** at B3LYP/6-31++G** level.

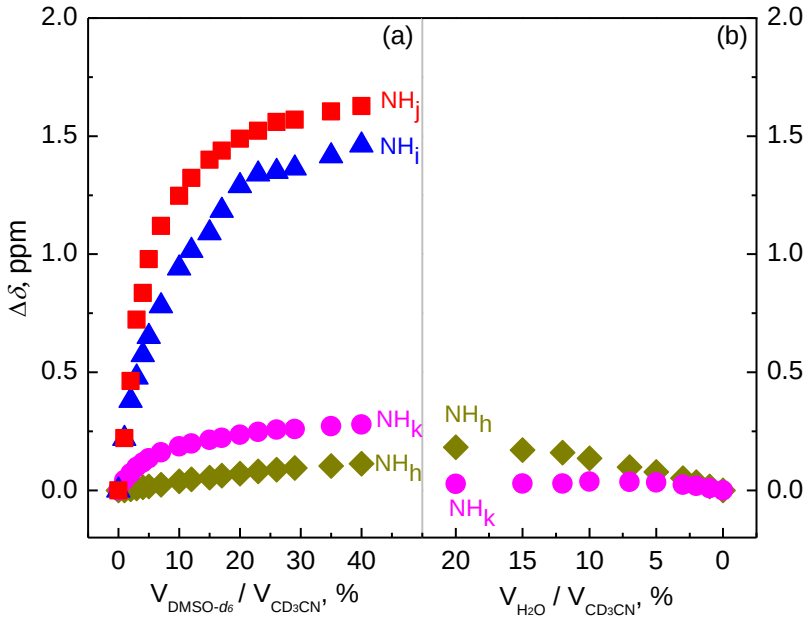


Figure S4. Influence on NH proton resonances of L-**2** in CD_3CN/DMSO-d_6 and CD_3CN/H_2O mixture by volume ratio of (a) DMSO-d_6 and (b) H_2O. Signals of NH_i and NH_j are invisible in CD_3CN/H_2O mixture. [L-**2**] = 10 mM. For NH proton numbering see Figure 2 in the text.

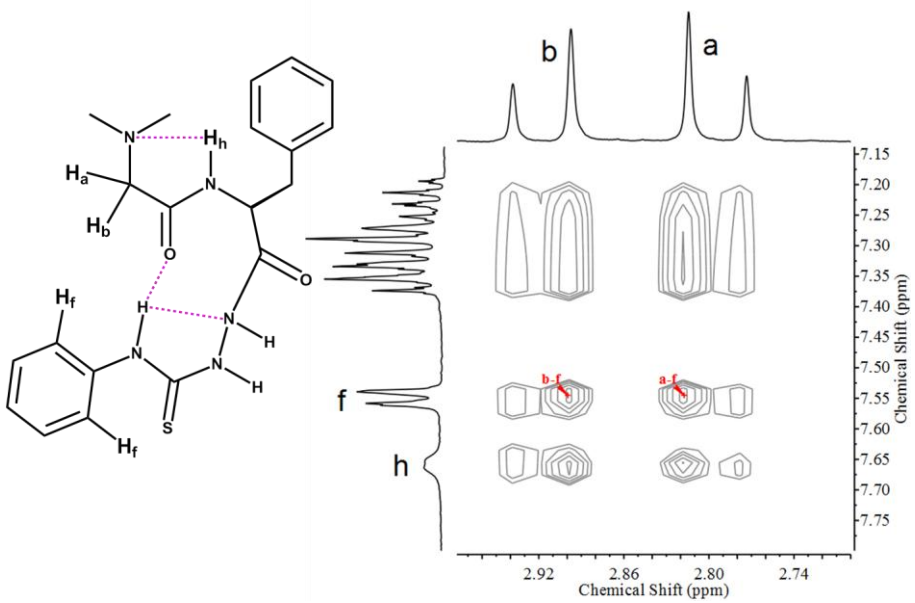


Figure S5. ^{1}H-^{1}H NOESY spectrum of L-**2** in CD_3CN. The nuclear Overhauser effect (NOE) signals between H_a or H_b and H_f are in consistent with the existence of β-turn in L-**2**. [L-**2**] = 10 mM.

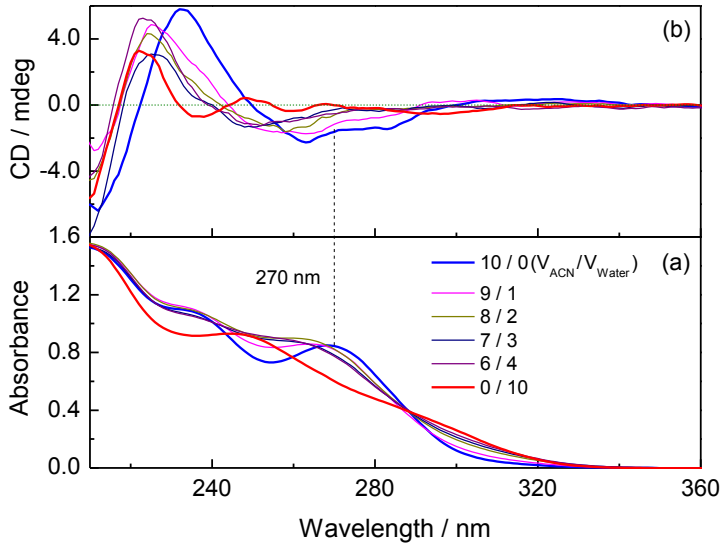


Figure S6. Absorption (a) and CD (b) spectra of L-**2** in CH_3CN-H_2O binary solvents. [L-**2**] = 80 μM. Using CD signal at 270 nm as an indication of the β-turn structure, it is assumed that it exists in solution containing up to *ca.*15% by volume water.

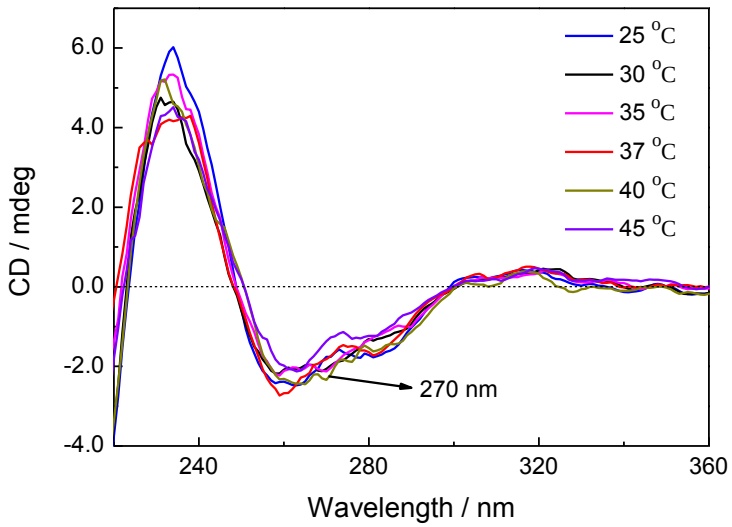


Figure S7. Temperature dependent CD spectra of L-**2** in CH_3CN. [L-**2**] = 80 μM. The CD spectrum does not change very much, suggesting the β-structure in **2** is stable over 25 - 45 °C.

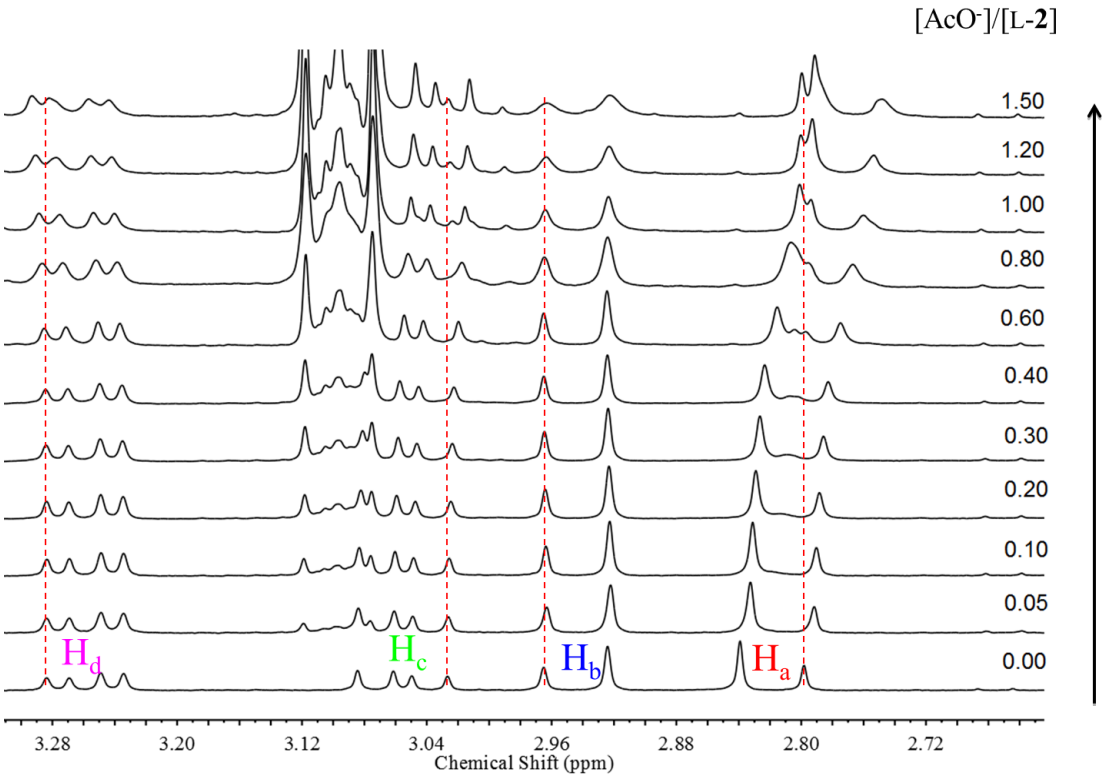


Figure S8. ^1H NMR spectra of L-**2** in the presence of AcO^- in CD_3CN. [L-**2**] = 10 mM.

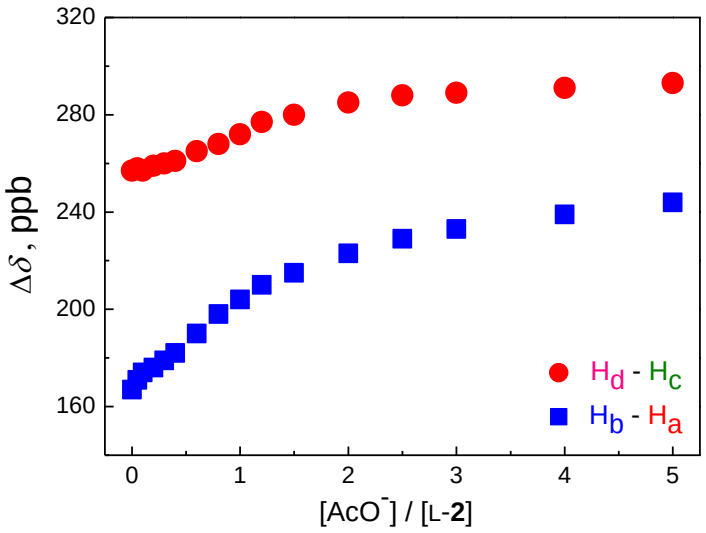
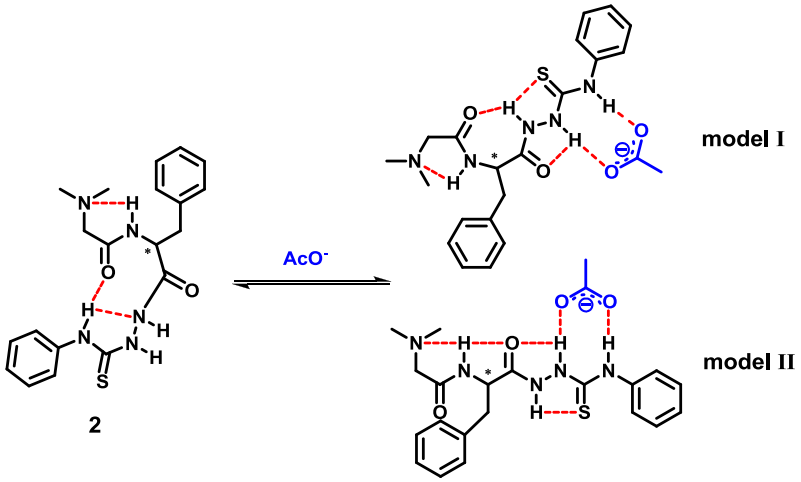


Figure S9. Splitting of NMR signals of protons H_b, H_a and H_d, H_c in L-**2** in the presence of AcO^- in CD_3CN. [L-**2**] = 10 mM.



Scheme S2. AcO^- binding with **2** and two possible models of hydrogen bonding networks in the anion binding complex

ω_i	φ_{i+1}	ψ_{i+1}	ω_{i+1}	φ_{i+2}	ψ_{i+2}	ω_{i+2}
177.30°	-72.22°	122.88°	-176.15°	71.07°	23.35°	-179.72°

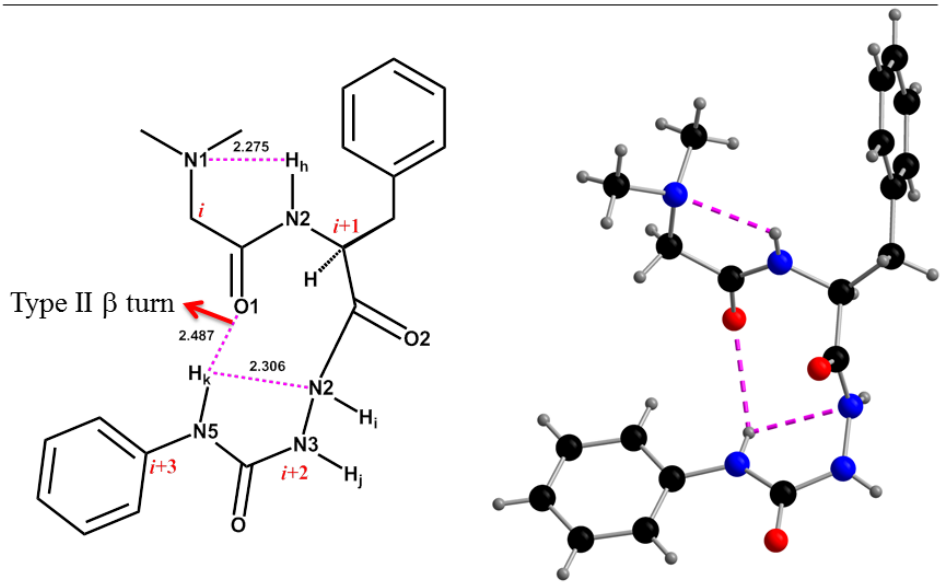


Figure S10. X-ray crystal structure of L-**3** and torsion angles of L-**3**. Dashed pink lines highlight the intramolecular hydrogen bonds.

Table S2. Bond lengths and angles of β-turns in the crystal structures of L-**2** and L-**3**

	N5-H$_k$···O1	**∠N5H$_k$O1**
L-**2**	2.401Å	152.72°
L-**3**	2.487Å	146.53°

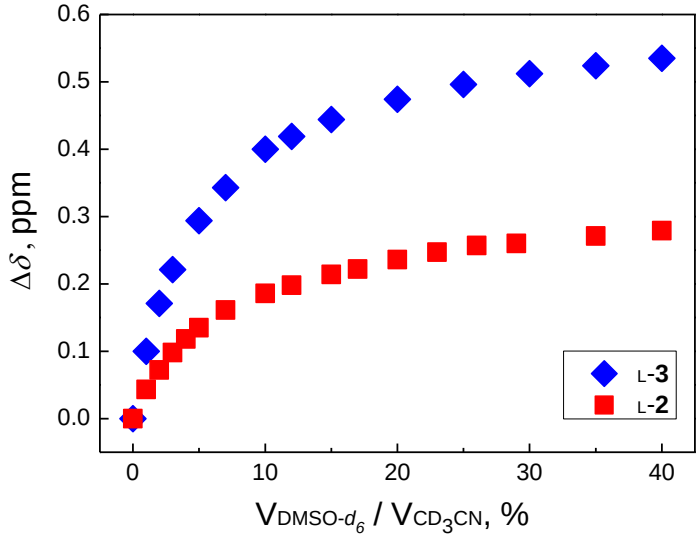


Figure S11. Influence of DMSO-d_6 volume ratio in CD_3CN/DMSO-d_6 mixture on the resonance of NH_k proton in L-**2** and L-**3**. [L-**2**] = [L-**3**] = 10 mM.

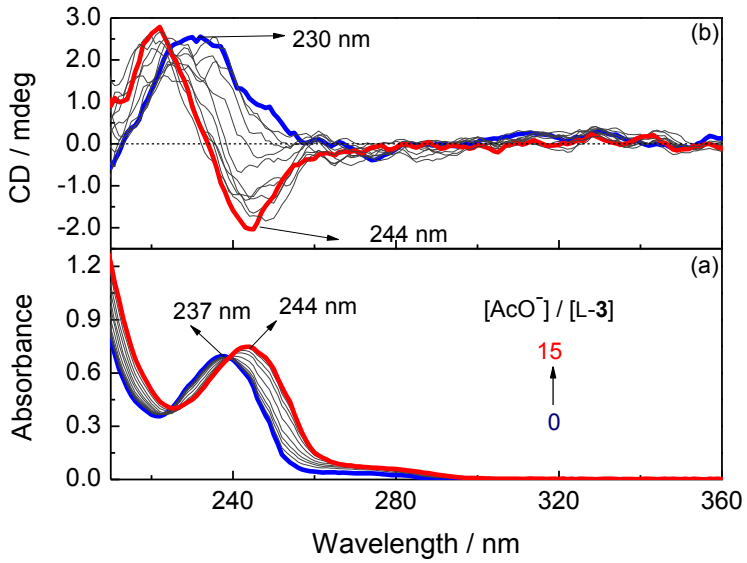


Figure S12. Absorption (a) and CD (b) spectra of L-**3** in CH_3CN in the presence of AcO^-. [L-**3**] = 40 μM.

Table S3. Binding constants of L-**2** and L-**3** with AcO^- in CH_3CN [a]

	$K_a \pm$ sd, M^{-1}	R^2
L-**2**	$(8.75 \pm 5.41)*10^6$	0.9973
L-**3**	$(5.18 \pm 0.39)*10^4$	0.9983

a. The constants were determined from the absorption titration at 296 nm (L-**2**) or 244 nm (L-**3**) as shown in Fig. 3 and Fig. S12. The following equation was used for fitting the data according to 1:1 binding model,

$$A = A_0 + \frac{A_{limit} - A_0}{2c_0}\left[x + 1/K_a + c_0 - \sqrt{(x + 1/K_a + c_0)^2 - 4c_0 x}\right]$$

in which A_0, A and A_{limit} denote respectively the absorbance of host (L-**2** or L-**3**), the bound complex and the limit value, c_0 is the molar concentration of host, x is the molar concentration of guest (AcO^-), and K_a is the binding constant.

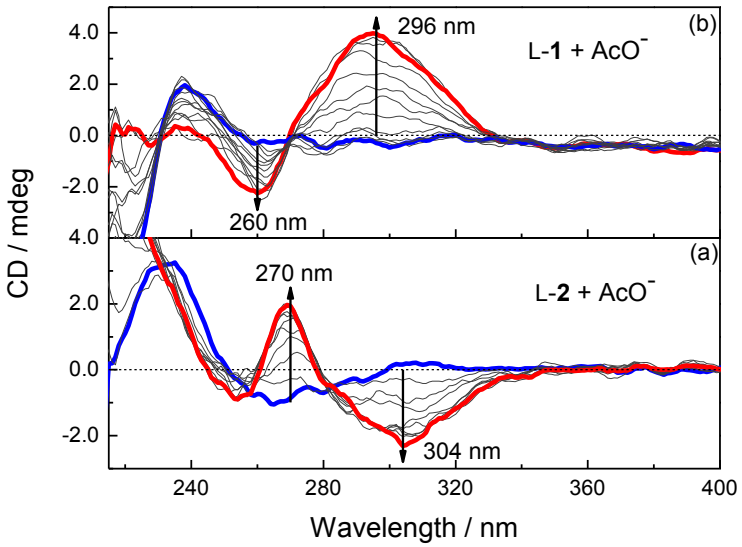


Figure S13. CD spectra of L-**2** (a) and L-**1** (b) in CH_3CN in the presence of AcO^-. [L-**2**] = [L-**1**] = 40 μM.

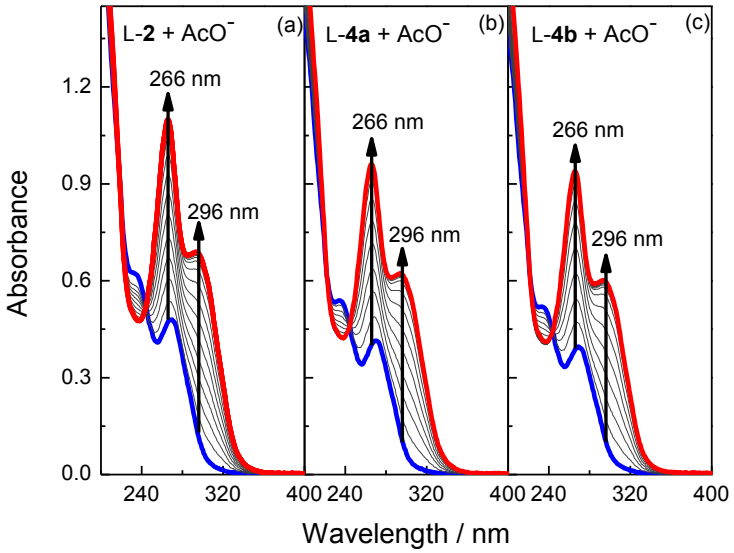


Figure S14. Absorption spectra of L-**2** (a), L-**4a** (b) and L-**4b** (c) in CH_3CN in the presence of AcO^-. [L-**2**] = [L-**4a**] = [L-**4b**] = 40 μM.

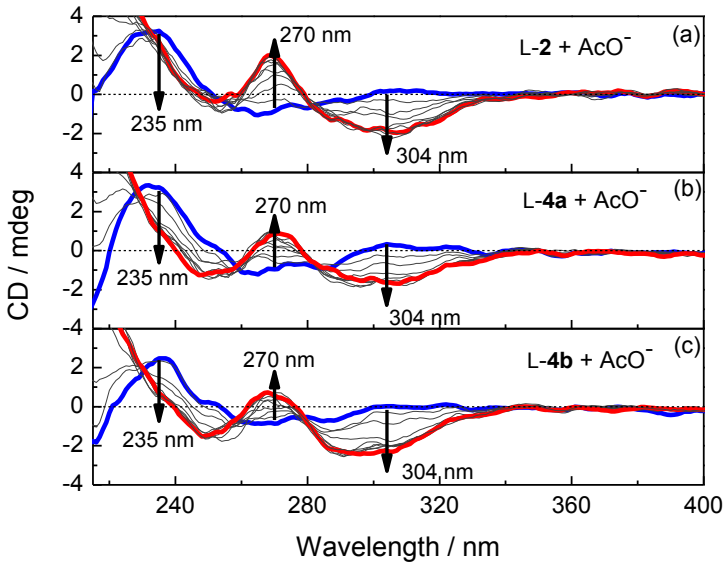
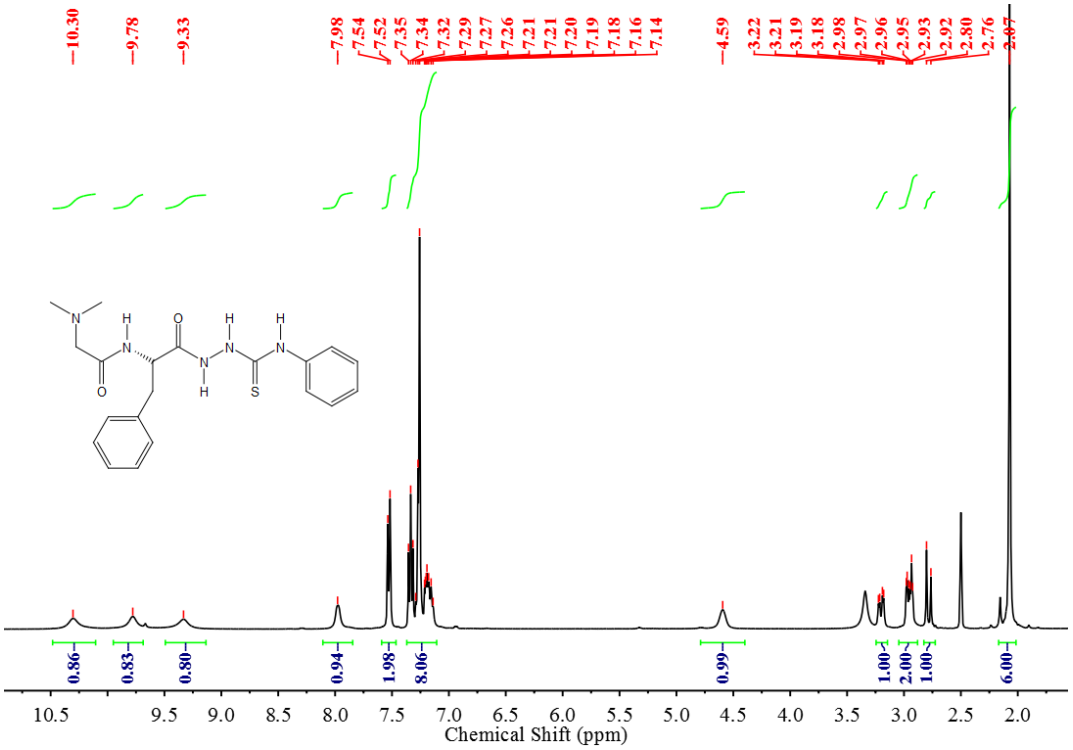


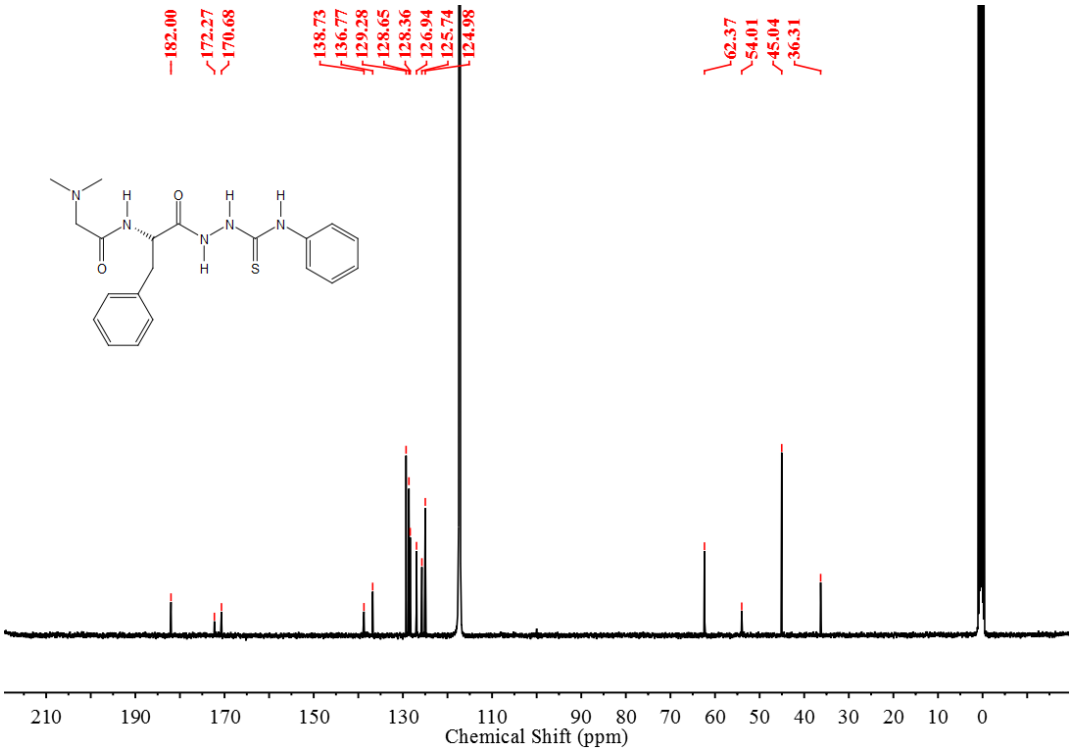
Figure S15. CD spectra of L-**2** (a), L-**4a** (b) and L-**4b** (c) in CH_3CN in the presence of AcO^-. [L-**2**] = [L-**4a**] = [L-**4b**] = 40 μM.

^{1}H NMR and ^{13}C NMR spectra of 2, L-3, L-4a and L-4b

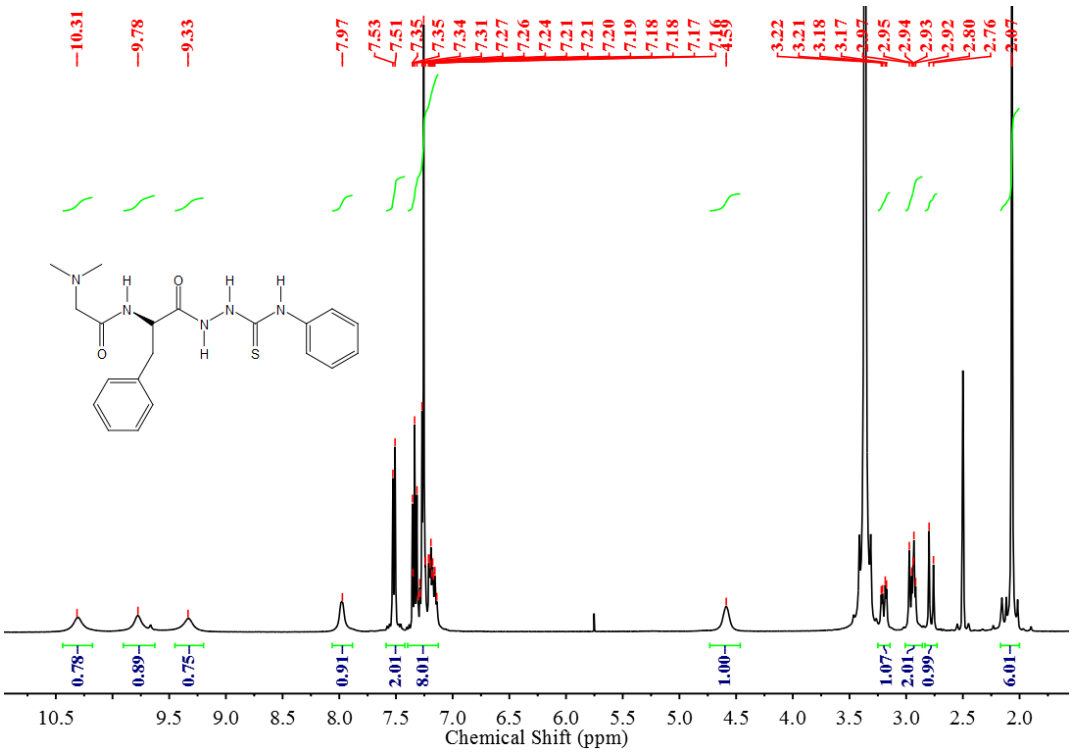
^{1}H NMR of L-**2** (400 MHz, DMSO-d_6)



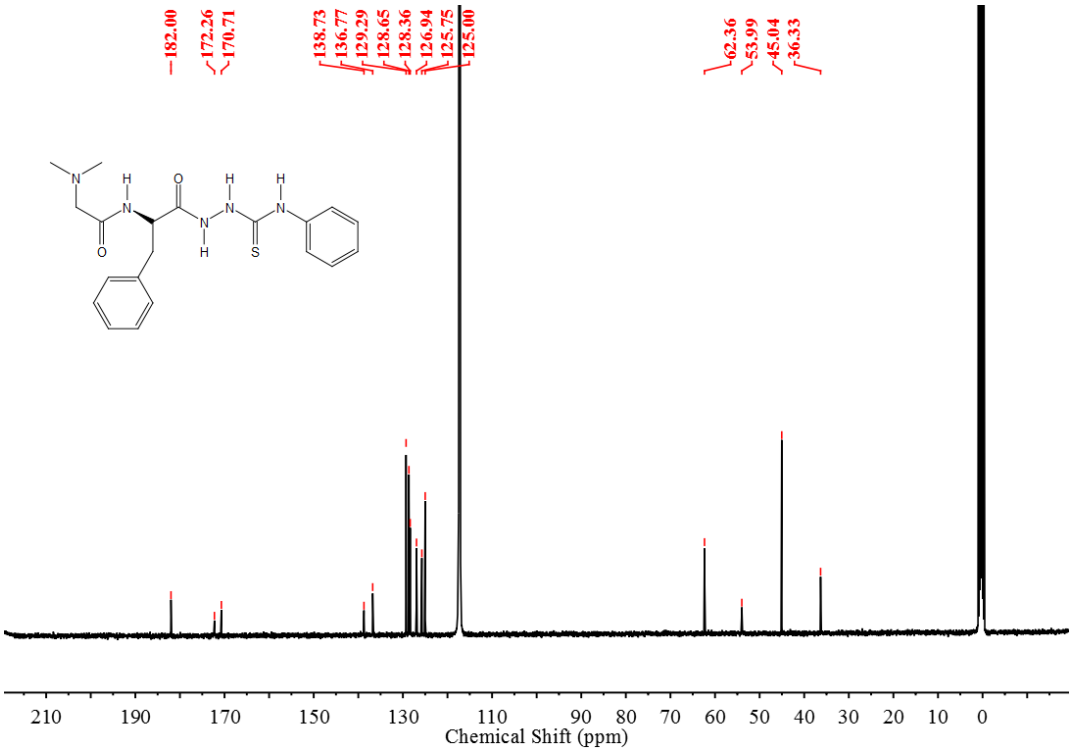
^{13}C NMR of L-**2** (CD_3CN)



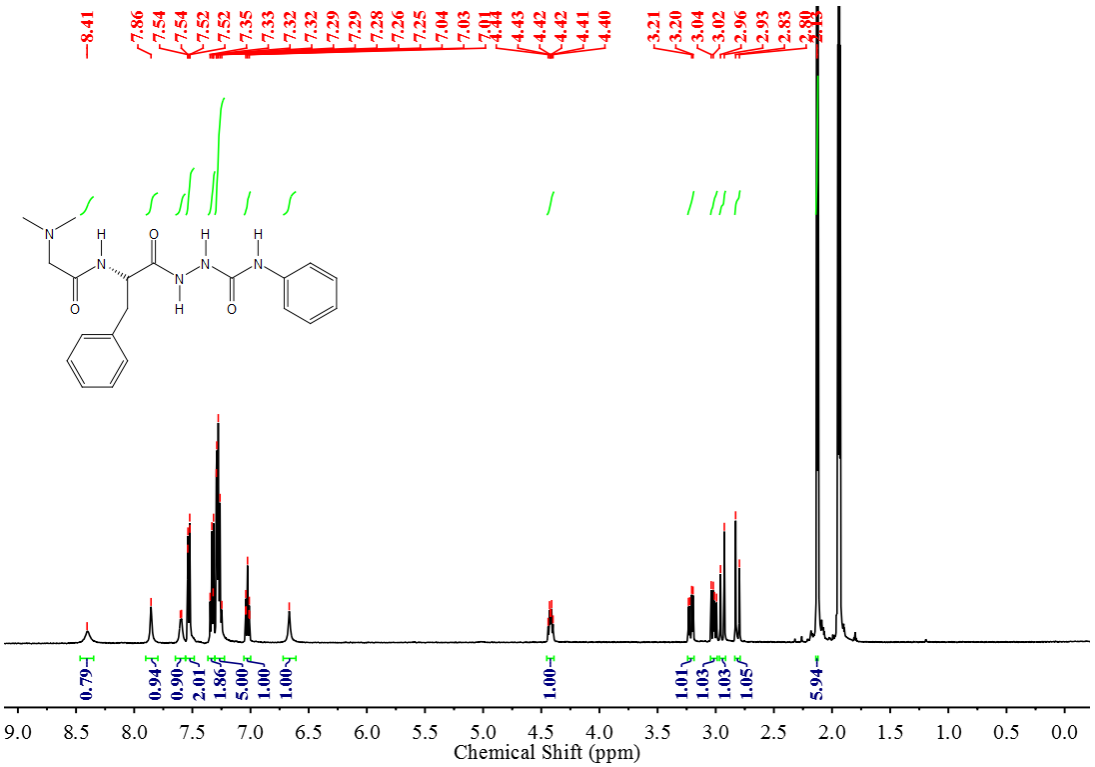
^{1}H NMR of D-**2** (400 MHz, DMSO-d_6)



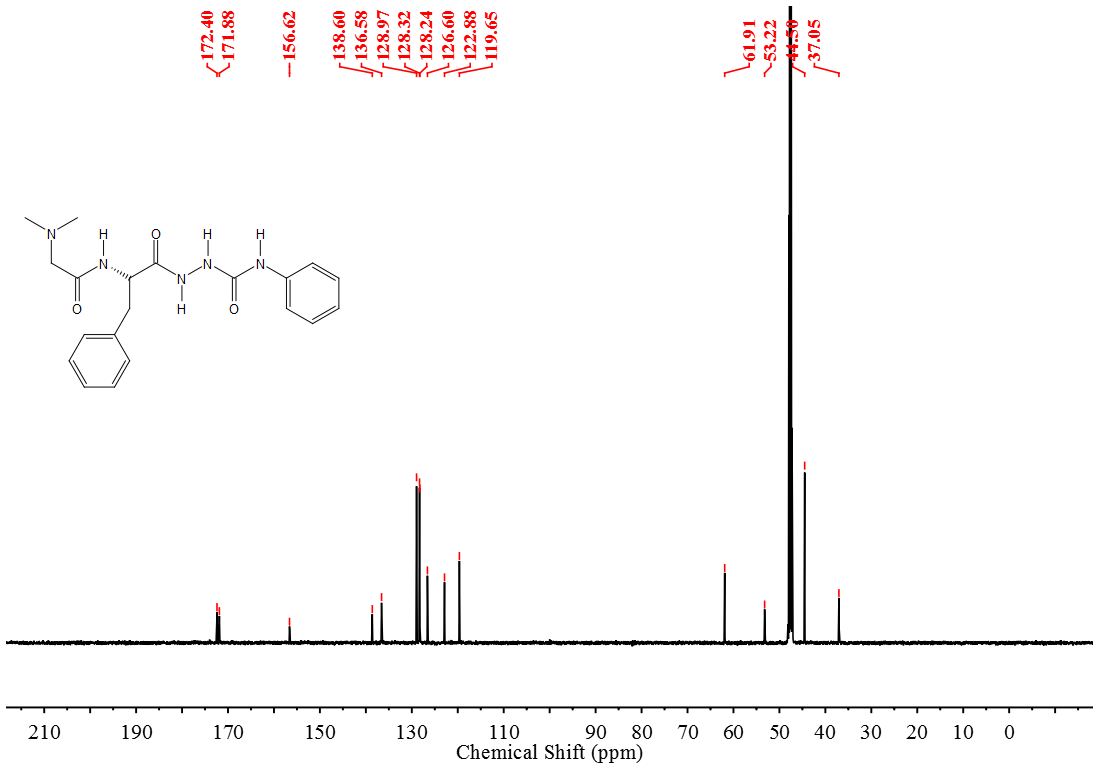
^{13}C NMR of D-**2** (101 MHz, CD_3CN)



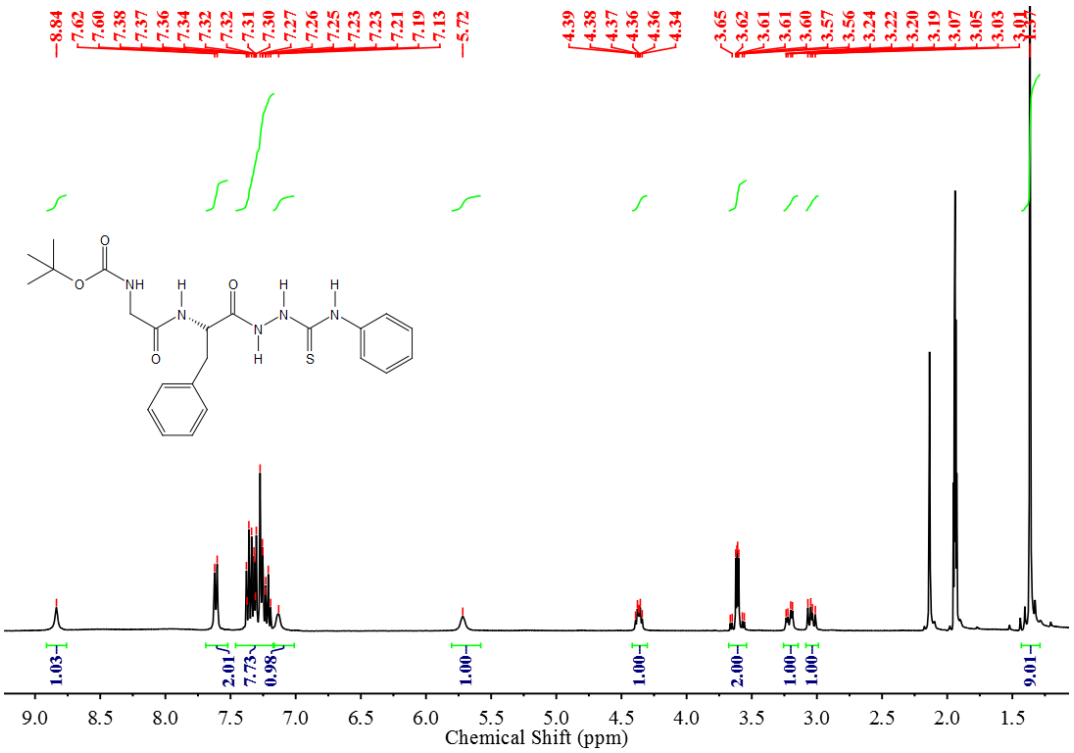
^{1}H NMR of L-**3** (500 MHz, CD_3CN)



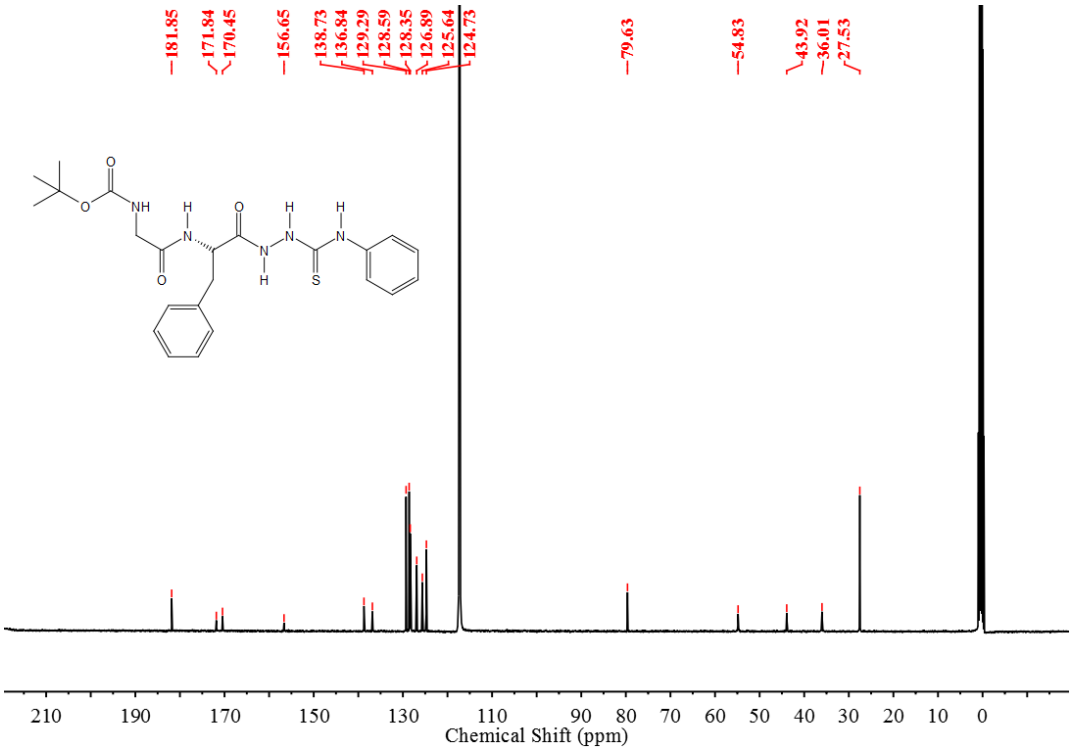
^{13}C NMR of L-**3** (126 MHz, MeOD)



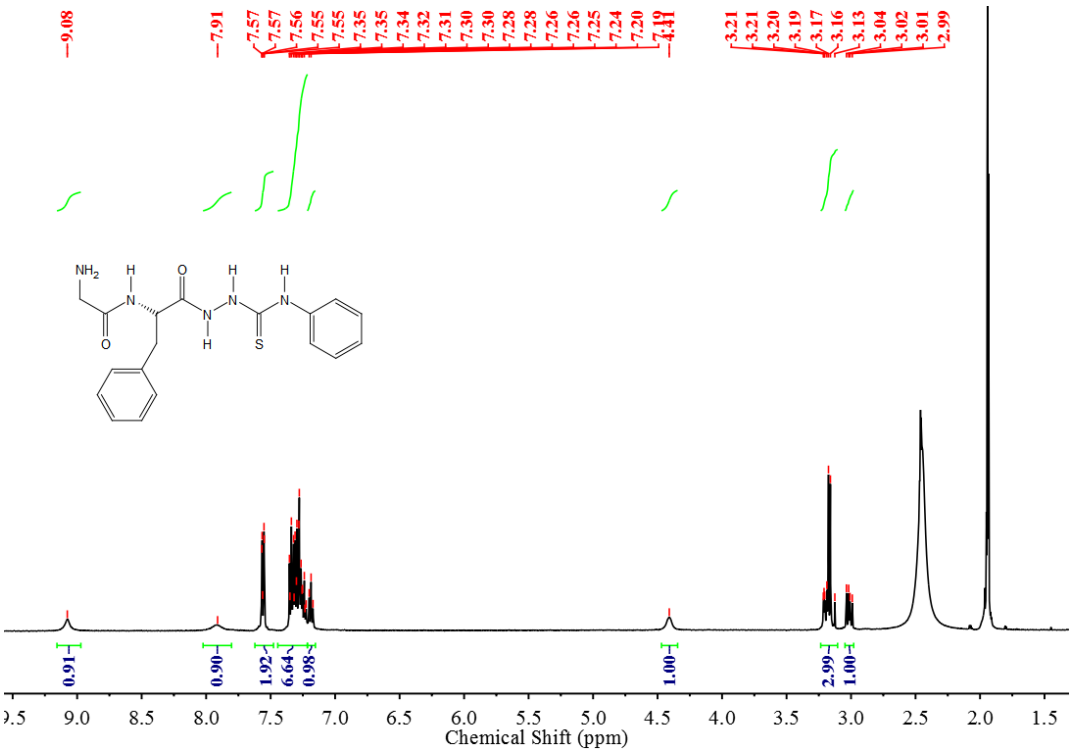
^{1}H NMR of L-**4a** (400 MHz, CD_3CN)



^{13}C NMR of L-**4a** (101 MHz, CD_3CN)



^{1}H NMR of L-**4b** (500 MHz, CD_3CN)



^{13}C NMR of L-**4b** (101 MHz, CD_3CN)