Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.
This journal is © The Royal Society of Chemistry 2021

Chalcogen bonding mediates formation of supramolecular helices of azapeptides in crystals

Di Shi, Jinlian Cao, Peimin Weng, Xiaosheng Yan,* Zhao Li, and Yun-Bao Jiang*

Department of Chemistry, College of Chemistry and Chemical Engineering, The MOE Key Laboratory of Spectrochemical Analysis and Instrumentation, and Collaborative Innovation Center of Chemistry for Energy Materials (*i*ChEM), Xiamen University, Xiamen 361005, China

*E-mail: xshyan@xmu.edu.cn; ybjiang@xmu.edu.cn

Supporting Information

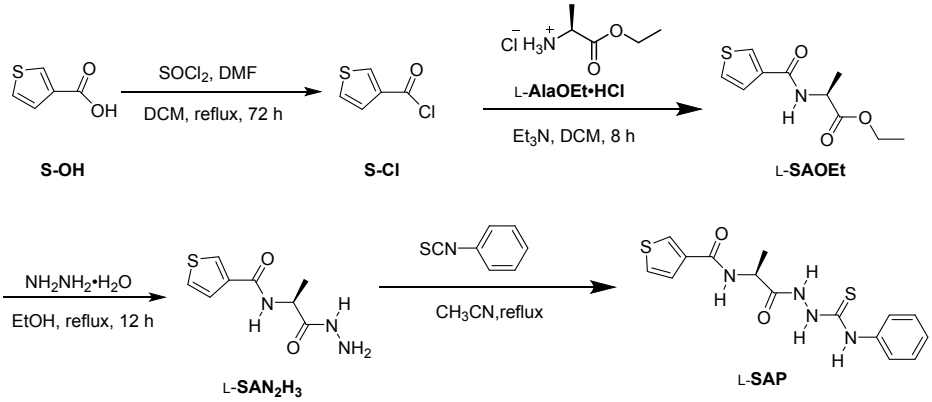
Contents

1. **Synthesis and characterization**……..…S2
2. **Crystallographic data**……..……...S6
3. **Spectral data**……..……S9
4. **^{1}H NMR and ^{13}C NMR spectra**…….…….….S11
5. **ORTEP diagrams**……...…….……S23

1. Synthesis and characterization

General

Reagents and solvents were purchased from commercial sources and used without further purification. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV850 MHz spectrometer in dimethyl sulfoxide-D_6 (DMSO-d_6), acetonitrile-D_3 (CD_3CN) or mixed solvents. Electrospray ionization high resolution mass spectra (ESI-HRMS) were recorded on a Bruker En Apex ultra 7.0 FT-MS. Absorption spectra were recorded on a Thermo Scientific Evolution 300 UV/Vis spectrophotometer. CD spectra were recorded with a JASCO J-1500 spectropolarimeter. X-ray crystallography data were collected on an Agilent SuperNova Dual system (CuKα or MoKα). The structure was solved by direct methods using OLEX2 program package, and non-hydrogen atoms were refined anisotropically unless otherwise stated.



Scheme S1. General procedures for the syntheses of L-**SAP**.

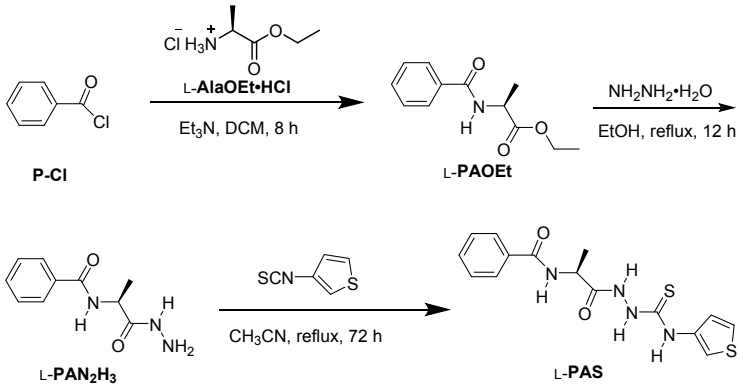
S-Cl: To a solution of 3-thiophenecarboxylic acid (1.28 g, 10 mmol) and 30 mL CH_2Cl_2, thionylchloride (20 mL) and DMF (0.2 mL) were added dropwise. The reaction mixture was refluxed for 72 hours. The solvent was removed by evaporated to give yellow solid **S-Cl**.

L-**SAOEt**: **AlaOEt·HCl** (L-configuration, 2.30 g, 15 mmol), Et_3N (5 mL) and 30 mL CH_2Cl_2 were added to a 100 ml round-bottom flask equipped with magnetic stirring. **S-Cl** in 15 mL CH_2Cl_2 was added dropwise. The reaction mixture was stirred at room temperature for 8 hours. The solvent was removed by evaporated *in vacuo*, and the crude product was washed by EtOH/water mixtures and Et_2O, producing white solid L-**SAOEt** (1.84 g, 81% yield).

L-**SAN_2H_3**: Excess aqueous hydrazine (85%, 4.0 mL) was added to L-**SAOEt** in EtOH (50 mL) and the mixture was refluxed for 12 hours. After filtration, the crude product was washed with EtOH, CH_3CN and Et_2O several times to afford 1.62g white solid product L-**SAN_2H_3** (88% yield).

L-**SAP**: Excess phenyl isothiocyanate was added to L-**SAN_2H_3** (0.11 g, 0.5 mmol) in CH_3CN (60 mL) and then refluxed for 72 hours. The solvent was removed by filtration, and the crude product

was washed with EtOH, CH_3CN and Et_2O several times, affording pure L-**SAP** (87% yield). D-**SAP** was synthesized according to the same procedures by using D-alanine.



Scheme S2. General procedures for the syntheses of L-**PAS**.

L-**PAOEt**: **AlaOEt·HCl** (L-configuration, 2.30 g, 15 mmol), Et_3N (5 mL) and 30 mL CH_2Cl_2 were added to a 100 ml round-bottom flask equipped with magnetic stirring. **P-Cl** in 15 mL CH_2Cl_2 was added dropwise. The reaction mixture was stirred at room temperature for 8 hours. The solvent was removed by evaporated *in vacuo*, and the crude product was washed by EtOH/water mixtures and Et_2O, producing white solid L-**PAOEt** (1.88 g, 85% yield).

L-**PAN$_2$H$_3$**: Excess aqueous hydrazine (85%, 4.0 mL) was added to L-**PAOEt** in EtOH (50 mL) and the mixture was refluxed for 12 hours. After filtration, the crude product was washed with EtOH, CH_3CN and Et_2O several times to afford 1.65g white solid product L-**PAN$_2$H$_3$** (88% yield).

L-**PAS**: Excess phenyl isothiocyanate was added to L-**PAN$_2$H$_3$** (0.11 g, 0.5 mmol) in CH_3CN (60 mL) and then refluxed for 72 hours. After filtration, the crude product was washed with EtOH, CH_3CN and Et_2O several times, affording pure white solid product L-**PAS** (89% yield). D-**PAS** was synthesized according to the same procedures by using D-alanine.

L-**SAOEt**: **^{1}H NMR** (850 MHz, DMSO-d_6): δ (ppm) 8.58 (d, J = 6.8 Hz, 1H), 8.20 (d, J = 1.8 Hz, 1H), 7.59 (dd, J = 4.9, 3.0 Hz, 1H), 7.53 (dd, J = 4.9, 0.7 Hz, 1H), 4.42 (p, J = 7.2 Hz, 1H), 4.10 (dd, J = 7.1, 4.6 Hz, 2H), 1.38 (d, J = 7.4 Hz, 3H), 1.19 (t, J = 7.1 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-d_6): δ (ppm) 172.62, 161.84, 136.93, 129.15, 126.86, 126.62, 60.32, 47.95, 16.66, 13.98; **HRMS** (ESI): calcd for $C_{10}H_{13}O_3NSNa$ $[M+Na]^+$: 250.0514, found: 250.0561.

D-**SAOEt**: **^{1}H NMR** (850 MHz, DMSO-d_6): δ (ppm) 8.58 (d, J = 6.8 Hz, 1H), 8.20 (d, J = 1.8 Hz, 1H), 7.59 (dd, J = 4.9, 3.0 Hz, 1H), 7.53 (d, J = 4.9 Hz, 1H), 4.41 (p, J = 7.2 Hz, 1H), 4.10 (dd, J = 7.1, 4.6 Hz, 2H), 1.38 (d, J = 7.4 Hz, 3H), 1.19 (t, J = 7.1 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-d_6): δ (ppm) 172.61, 161.83, 136.93, 129.15, 126.86, 126.62, 60.32, 47.94, 16.66, 13.98; **HRMS** (ESI): calcd for $C_{10}H_{13}O_3NSNa$ $[M+Na]^+$: 250.0514, found: 250.0513.

L-**SAN$_2$H$_3$**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 9.14 (s, 1H), 8.30 (d, *J* = 7.4 Hz, 1H), 8.22 (d, *J* = 1.0 Hz, 1H), 7.58 – 7.55 (m, 1H), 7.54 (d, *J* = 4.5 Hz, 1H), 4.45 – 4.40 (m, 1H), 4.24 (s, 2H), 1.30 (d, *J* = 7.2 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-*d$_6$*) *δ* (ppm) 171.58, 161.55, 137.31, 128.94, 127.04, 126.28, 47.17, 18.08; **HRMS** (ESI): calcd for $C_8H_{11}O_2N_3SNa$ $[M+Na]^+$: 236.0470, found: 236.0479.

D-**SAN$_2$H$_3$**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 9.14 (s, 1H), 8.31 (d, *J* = 7.5 Hz, 1H), 8.22 (d, *J* = 1.6 Hz, 1H), 7.59 – 7.55 (m, 1H), 7.54 (d, *J* = 4.7 Hz, 1H), 4.45 – 4.40 (m, 1H), 4.26 (s, 2H), 1.30 (d, *J* = 7.2 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-*d$_6$*) *δ* (ppm) 171.58, 161.55, 137.31, 128.94, 127.04, 126.28, 47.17, 18.08; **HRMS** (ESI): calcd for $C_8H_{11}O_2N_3SNa$ $[M+Na]^+$: 236.0470, found: 236.0483.

L-**SAP**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 10.42 (s, 1H), 9.73 (s, 1H), 9.30 (s, 1H), 8.78 (s, 1H), 8.23 (d, *J* = 1.6 Hz, 1H), 7.70 (s, 2H), 7.61 (s, 1H), 7.52 (d, *J* = 4.6 Hz, 1H), 7.35 (t, *J* = 7.6 Hz, 2H), 7.15 (t, *J* = 7.0 Hz, 1H), 4.27 (s, 1H), 1.39 (d, *J* = 6.7 Hz, 3H)；**^{13}C NMR** (214 MHz, DMSO-*d$_6$*): *δ* (ppm) 179.91, 171.85, 163.18, 138.94, 136.45, 129.68, 127.98, 126.86, 126.73, 124.58, 123.78, 48.83, 16.36; **HRMS** (ESI): calcd for $C_{15}H_{16}O_2N_4S_2Na$ $[M+Na]^+$: 371.0612, found: 371.0607.

D-**SAP**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 10.42 (s, 1H), 9.73 (s, 1H), 9.30 (s, 1H), 8.78 (s, 1H), 8.23 (d, *J* = 1.6 Hz, 1H), 7.70 (s, 2H), 7.61 (s, 1H), 7.52 (d, *J* = 4.6 Hz, 1H), 7.35 (t, *J* = 7.6 Hz, 2H), 7.15 (t, *J* = 7.0 Hz, 1H), 4.27 (s, 1H), 1.39 (d, *J* = 6.7 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-*d$_6$*): *δ* (ppm) 179.86, 171.84, 163.16, 138.93, 136.47, 129.69, 127.98, 126.85, 126.72, 124.58, 123.79, 48.78, 16.38; **HRMS** (ESI): calcd for $C_{15}H_{16}O_2N_4S_2Na$ $[M+Na]^+$: 371.0612, found: 371.0608.

L-**PAOEt**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 8.76 (d, *J* = 6.5 Hz, 1H), 7.88 (d, *J* = 7.5 Hz, 2H), 7.55 (t, *J* = 7.2 Hz, 1H), 7.48 (t, *J* = 7.5 Hz, 2H), 4.48 – 4.41 (m, 1H), 4.14 – 4.07 (m, 2H), 1.40 (d, *J* = 7.3 Hz, 3H), 1.19 (t, *J* = 7.0 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-*d$_6$*): *δ* (ppm) 172.58, 166.19, 133.66, 131.32, 128.16, 127.32, 60.31, 48.26, 16.60, 13.98; **HRMS** (ESI): calcd for $C_{12}H_{15}O_3NNa$ $[M+Na]^+$: 244.0950, found: 244.1018.

D-**PAOEt**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 8.76 (d, *J* = 6.5 Hz, 1H), 7.88 (d, *J* = 7.5 Hz, 2H), 7.55 (t, *J* = 7.2 Hz, 1H), 7.48 (t, *J* = 7.5 Hz, 2H), 4.48 – 4.41 (m, 1H), 4.14 – 4.07 (m, 2H), 1.40 (d, *J* = 7.3 Hz, 3H), 1.19 (t, *J* = 7.0 Hz, 3H); **^{13}C NMR** (214 MHz, DMSO-*d$_6$*): *δ* (ppm) 172.58, 166.21, 133.65, 131.34, 128.17, 127.32, 60.32, 48.27, 16.59, 13.97; **HRMS** (ESI): calcd for $C_{12}H_{15}O_3NNa$ $[M+Na]^+$: 244.0950, found: 244.1008.

L-**PAN$_2$H$_3$**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): *δ* (ppm) 9.14 (s, 1H), 8.45 (d, *J* = 7.0 Hz, 1H), 7.89 (d, *J* = 7.4 Hz, 2H), 7.53 (t, *J* = 7.1 Hz, 1H), 7.46 (t, *J* = 7.3 Hz, 2H), 4.48 – 4.42 (m, 1H), 4.21 (s, 2H), 1.32 (d, *J* = 7.0 Hz, 3H); **^{13}C NMR** (214 MHz, CD_3CN): *δ* (ppm) 171.61, 165.85, 133.99, 131.11, 128.00, 127.42, 47.58, 18.04; **HRMS** (ESI): calcd for $C_{10}H_{14}O_2N_3$ $[M+H]^+$: 208.1086, found: 208.0993.

D-**PAN$_2$H$_3$**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): δ (ppm) 9.14 (s, 1H), 8.45 (d, *J* = 7.0 Hz, 1H), 7.89 (d, *J* = 7.4 Hz, 2H), 7.53 (t, *J* = 7.1 Hz, 1H), 7.46 (t, *J* = 7.3 Hz, 2H), 4.48 – 4.42 (m, 1H), 4.21 (s, 2H), 1.32 (d, *J* = 7.0 Hz, 3H); **^{13}C NMR** (214 MHz, CD_3CN): δ (ppm) 171.61, 165.85, 133.99, 131.11, 128.00, 127.41, 47.58, 18.04; **HRMS** (ESI): calcd for $C_{10}H_{13}O_2N_3Na$ $[M+Na]^+$: 230.0905, found: 230.0955.

L-**PAS**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): δ (ppm) 10.41 (s, 1H), 9.72 (s, 1H), 9.53 (s, 1H), 8.98 (s, 1H), 8.04 (s, 1H), 7.92 (d, *J* = 7.6 Hz, 2H), 7.57 (t, *J* = 7.2 Hz, 1H), 7.50 (dd, *J* = 13.9, 5.6 Hz, 4H), 4.27 (s, 1H), 1.40 (d, *J* = 6.9 Hz, 3H); **^{13}C NMR** (214 MHz, CD_3CN): δ (ppm) 181.68, 172.90, 170.01, 138.15, 134.28, 133.19, 129.64, 128.59, 125.14, 124.92, 113.68, 51.46, 16.77; **HRMS** (ESI): calcd for $C_{15}H_{16}O_2N_4S_2Na$ $[M+Na]^+$: 371.0612, found: 371.0610.

D-**PAS**: **^{1}H NMR** (850 MHz, DMSO-*d$_6$*): δ (ppm) 10.41 (s, 1H), 9.72 (s, 1H), 9.53 (s, 1H), 8.98 (s, 1H), 8.04 (s, 1H), 7.92 (d, *J* = 7.6 Hz, 2H), 7.57 (t, *J* = 7.2 Hz, 1H), 7.50 (dd, *J* = 13.9, 5.6 Hz, 4H), 4.27 (s, 1H), 1.40 (d, *J* = 6.9 Hz, 3H); **^{13}C NMR** (214 MHz, CD_3CN): δ (ppm) 181.68, 172.90, 170.01, 138.15, 134.28, 133.19, 129.64, 128.59, 125.14, 124.92, 113.68, 51.46, 16.77; **HRMS** (ESI): calcd for $C_{15}H_{16}O_2N_4S_2Na$ $[M+Na]^+$: 371.0612, found: 371.0601.

2. Crystallographic data

Table S1. Crystal data and structure refinement for L-**SAP** and L-**PAS**

Compound reference	L-**SAP**	L-**PAS**
Empirical formula	$C_{15}H_{16}N_4O_2S_2$	$C_{15}H_{16}N_4O_2S_2$, C_2H_3N
Formula weight	348.44	389.49
Temperature/K	99.97(11)	100.01(10)
Crystal system	monoclinic	monoclinic
Space group	$P2_1$	$P2_1$
a/Å	10.1096(3)	9.0334(2)
b/Å	8.1557(2)	8.7125(2)
c/Å	10.7216(4)	12.1508(3)
α/°	90	90
β/°	110.545(4)	94.034(2)
γ/°	90	90
Volume/Å^3	827.78(5)	953.94(4)
Z	2	2
ρ_{calc}g/cm^3	1.398	1.356
μ/mm^{-1}	3.044	0.301
F(000)	364.0	408.0
Crystal size/mm^3	0.22 × 0.18 × 0.16	0.18 × 0.17 × 0.13
Radiation	CuKα (λ = 1.54184)	MoKα (λ = 0.71073)
2Θ range for data collection/°	8.808 to 142.772	7.176 to 61.432
Index ranges	$-12 \leqslant h \leqslant 11$, $-9 \leqslant k \leqslant 9$, $-12 \leqslant l \leqslant 12$	$-8 \leqslant h \leqslant 12$, $-10 \leqslant k \leqslant 12$, $-16 \leqslant l \leqslant 16$
Reflections collected	4847	5890
Independent reflections	2570 [R_{int} = 0.0401, R_{sigma} = 0.0510]	4193 [R_{int} = 0.0232, R_{sigma} = 0.0475]
Data/restraints/parameters	2570/54/217	4193/63/237
Goodness-of-fit on F^2	1.042	1.054
Final R indexes [I>=2σ (I)]	R_1 = 0.0379, wR_2 = 0.0952	R_1 = 0.0401, wR_2 = 0.1030
Final R indexes [all data]	R_1 = 0.0388, wR_2 = 0.0965	R_1 = 0.0452, wR_2 = 0.1091
Largest diff. peak/hole / e Å^{-3}	0.29/-0.31	0.59/-0.47
Flack parameter	0.001(16)	-0.02(4)
CCDC number	2084139	2084138

Table S2. Torsions, type, bond length and bond angle of β-turns revealed by the X-ray crystal structures

Crystal	φ_{i+1} / °	ψ_{i+1} / °	φ_{i+2} / °	ψ_{i+2} / °	Type	Length / Å	Angle / °
L-**SAP**	-61.63	131.27	-176.23	75.58	II	2.243	154.01
L-**PAS**	-66.80	128.90	-179.66	81.32	II	2.344	158.14

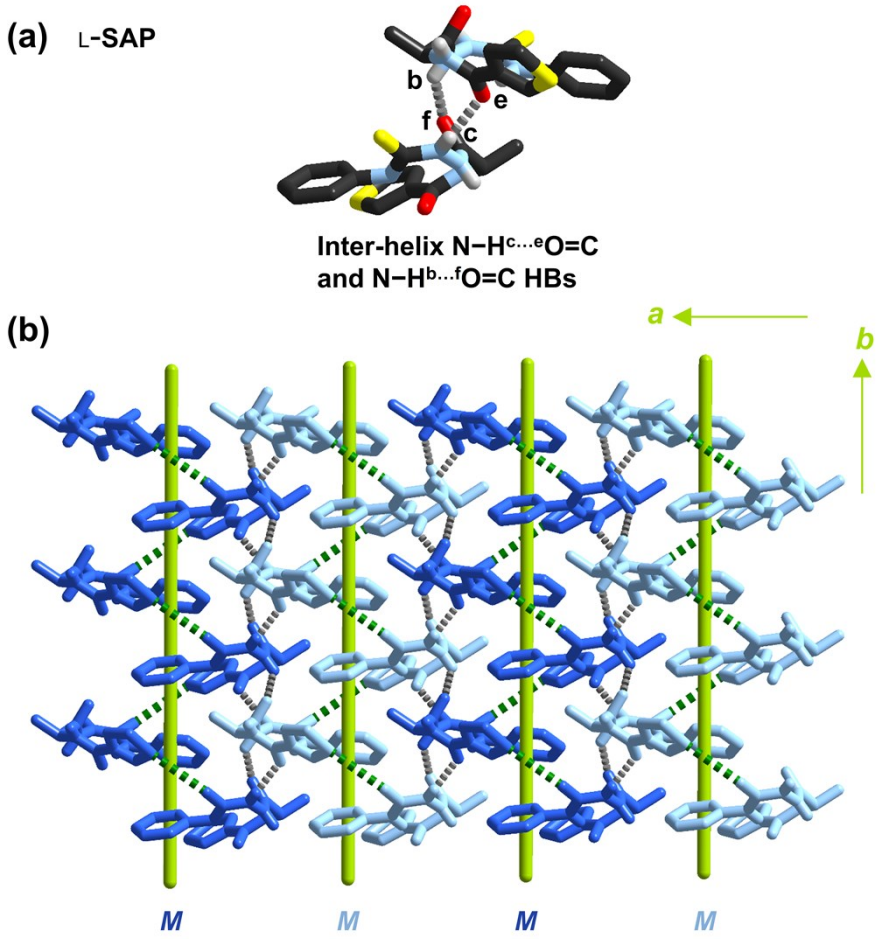


Figure S1. (a) Inter-helix N−H$^{c\cdots e}$O=C and N−H$^{b\cdots f}$O=C hydrogen bonds (HBs, dashed gray lines) between two adjacent L-**SAP** molecules. (b) Parallel chalcogen-bonded *M*-helices of L-**SAP** (see **Figure 1**) are contacted by inter-helix N−H$^{c\cdots e}$O=C and N−H$^{b\cdots f}$O=C hydrogen bonds (dashed gray lines) along *a*-axis, resulting in 2D supramolecular helical network. For clarity, all −CH protons are omitted.

Table S3. Parameters of inter-helix hydrogen bonds revealed in X-ray crystal structures

Crystal	Intra-helix interactions
L-**SAP**	N−H^{c}···^{e}O=C ($d_{H\cdots O}$ = 2.198 Å; ∠NHO = 149.1º)
	N−H^{b}··· ^{f}O=C ($d_{H\cdots O}$ = 2.109 Å; ∠NHO = 128.8º)
L-**PAS**	N−H^{c}···^{e}O=C ($d_{H\cdots O}$ = 2.172 Å; ∠NHO = 144.8º)
	N−H^{b}··· ^{f}O=C ($d_{H\cdots O}$ = 2.237 Å; ∠NHO = 127.2º)

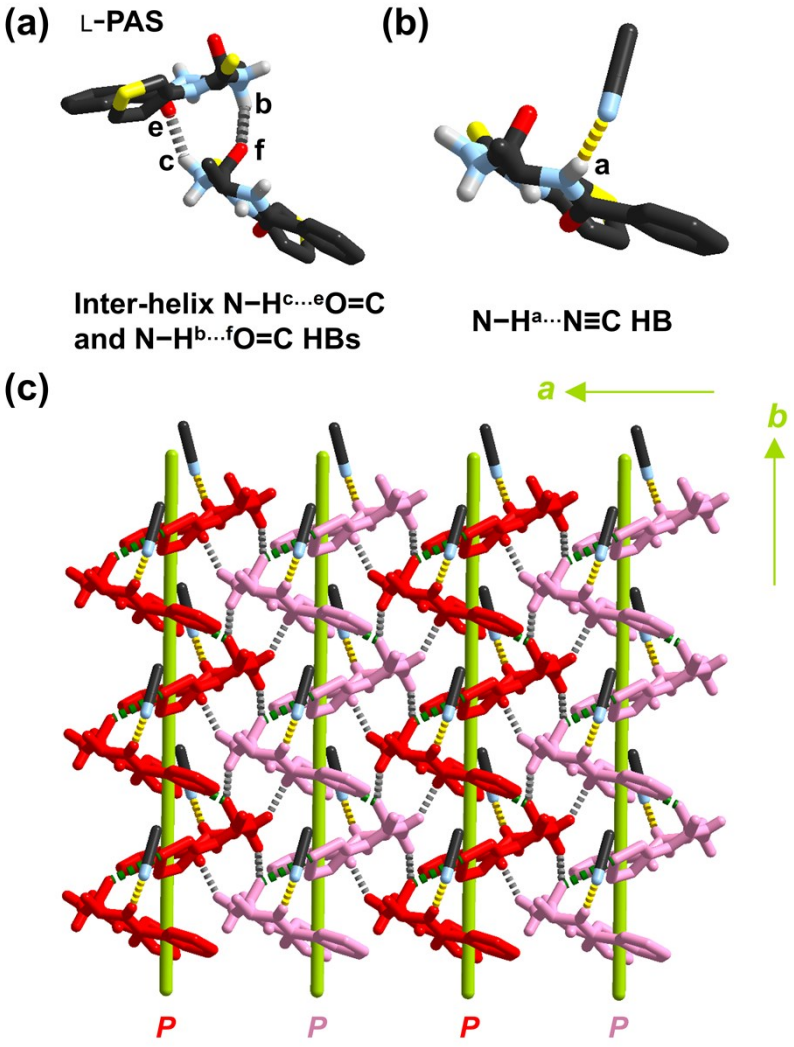


Figure S2. (a) Inter-helix N−H^{c}···^{e}O=C and N−H^{b}···^{f}O=C hydrogen bonds (HBs, dashed gray lines) between two adjacent L-**PAS** molecules. (b) Intermolecular N−H^{a}···N≡C hydrogen bond (HB, dashed yellow line) between L-**PAS** and solvent CH_3CN molecule. (c) Parallel chalcogen-bonded *P*-helices of L-**PAS** (see **Figure 2**) are contacted by inter-helix N−H^{c}···^{e}O=C and N−H^{b}···^{f}O=C hydrogen bonds (dashed gray lines) along *a*-axis, resulting in 2D supramolecular helical network. CH_3CN molecules are embedded into the 2D network through N−H^{a}···N≡C hydrogen bonds (dashed yellow lines). For clarity, all −CH protons are omitted.

3. Spectral data

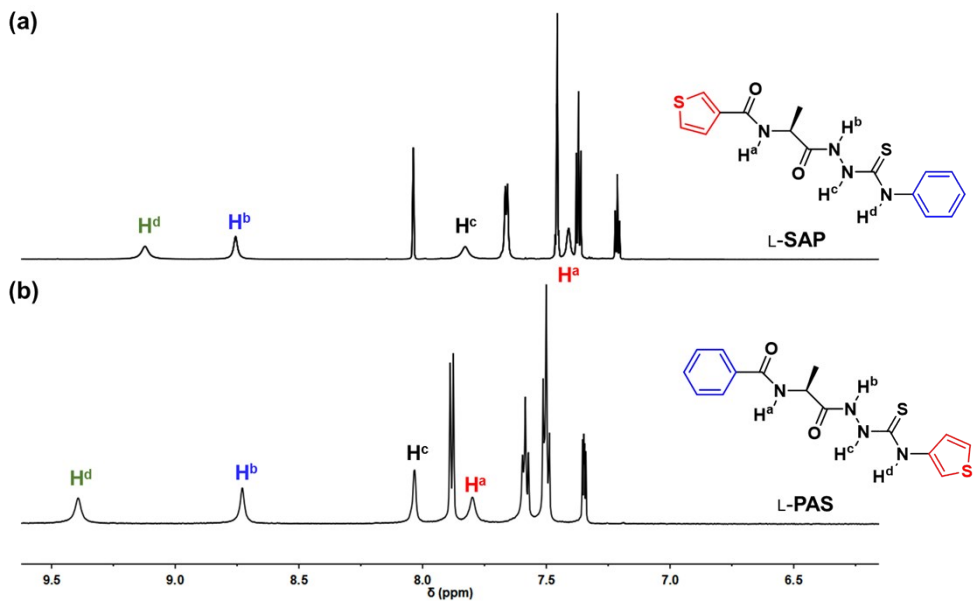


Figure S3. Partial 850 MHz ^{1}H NMR spectra of L-**SAP** (a) and L-**PAS** (b) in CD_3CN. [L-**SAP**] = [L-**PAS**] = 2 mM.

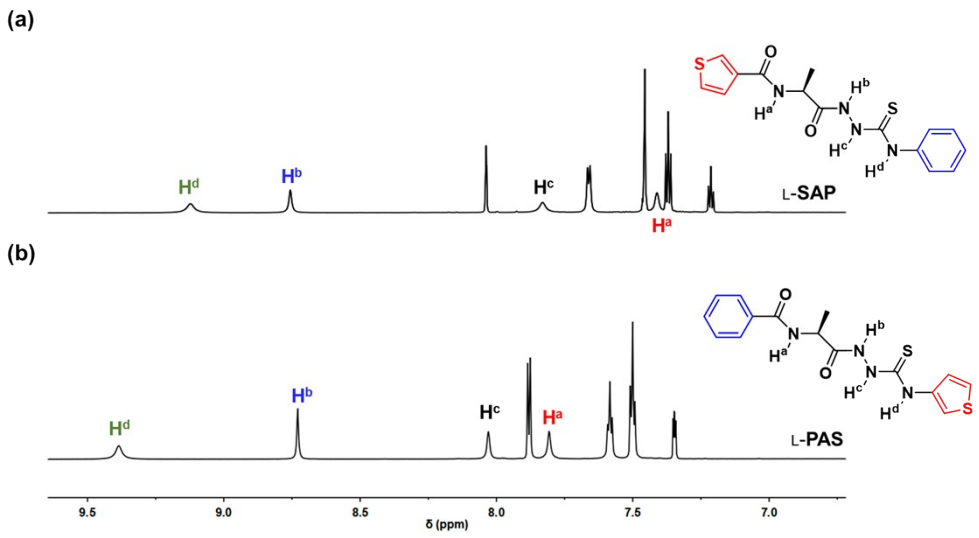


Figure S4. Partial 850 MHz ^{1}H NMR spectra of the crystals of L-**SAP** (a) and L-**PAS** (b) that dissolved in CD_3CN. The concentration is close to 1 mM.

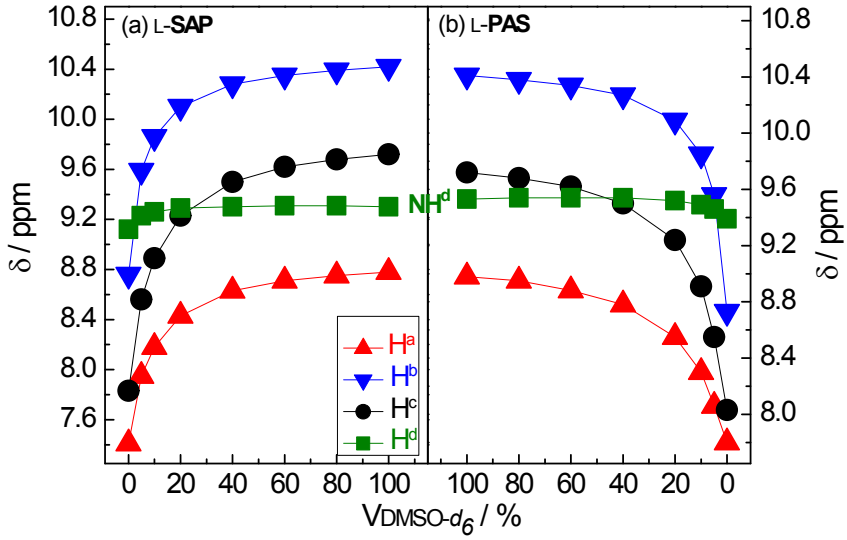


Figure S5. Influence on resonances of −NH proton of L-**SAP** (a) and L-**PAS** (b) in CD_3CN/DMSO-d_6 mixtures by volume fraction of DMSO-d_6. [L-**SAP**] = [L-**PAS**] =2 mM.

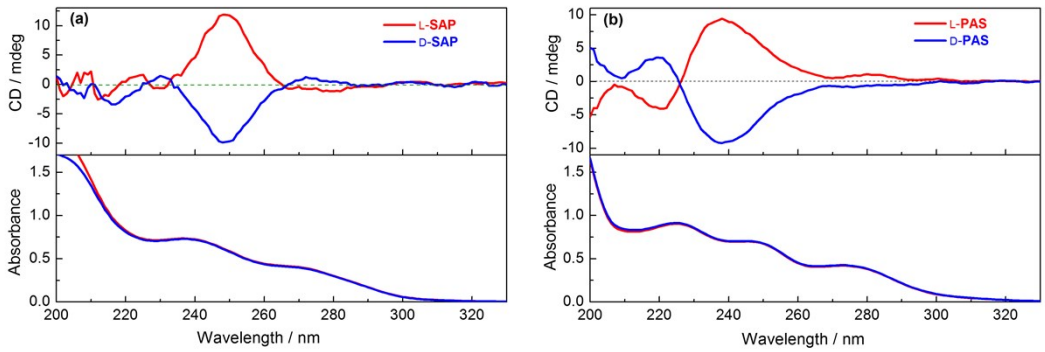


Figure S6. Absorption and CD spectra of L-/D-**SAP** (a) and L-/D-**PAS** (b) in CH_3CN. [L-/D-**SAP**] = [L-/D-**PAS**] = 40 μM.

4. ^{1}H NMR and ^{13}C NMR spectra

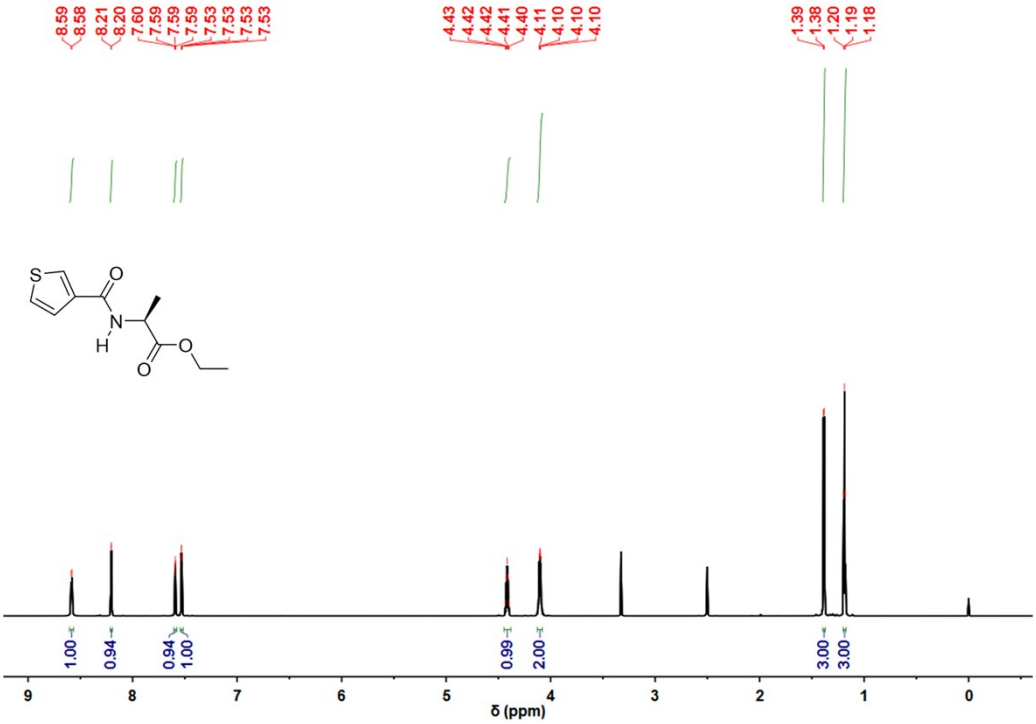


Figure S7. 850 MHz ^{1}H NMR spectrum of L-**SAOEt** in DMSO-d_6

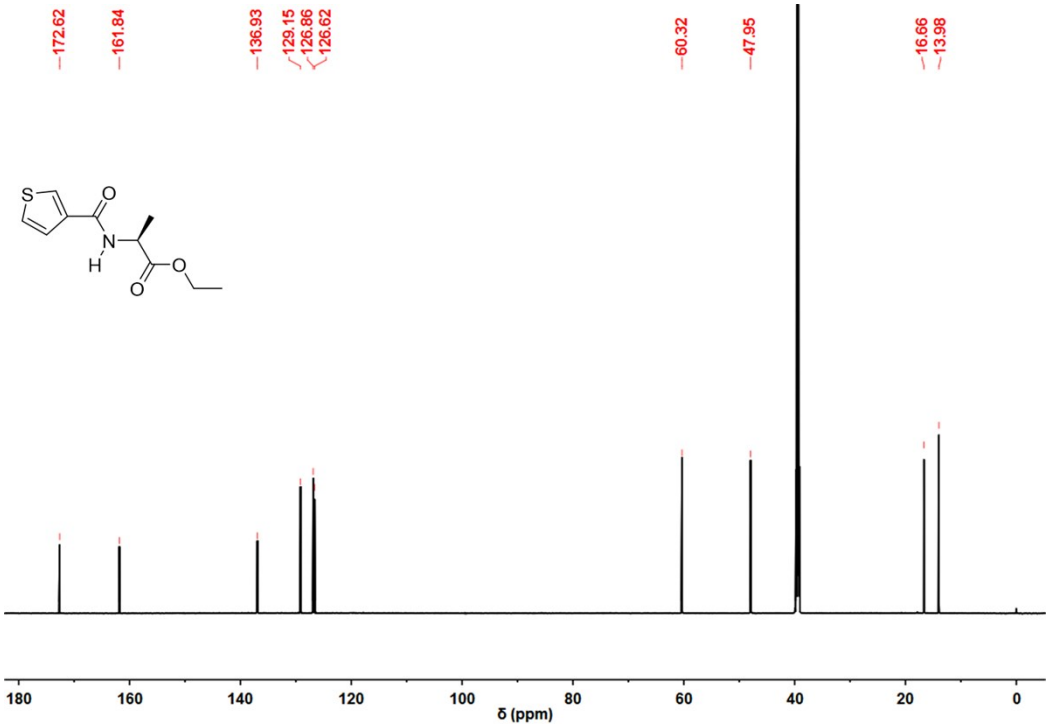


Figure S8. 214 MHz ^{13}C NMR spectrum of L-**SAOEt** in DMSO-d_6

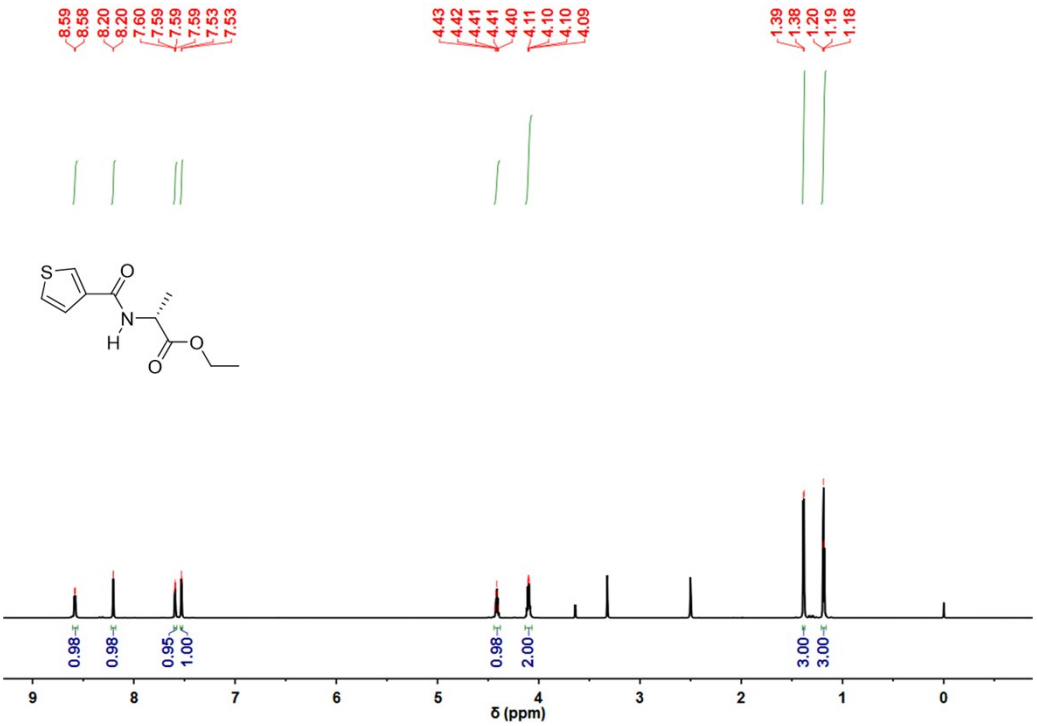


Figure S9. 850 MHz ^{1}H NMR spectrum of D-**SAOEt** in DMSO-d_6

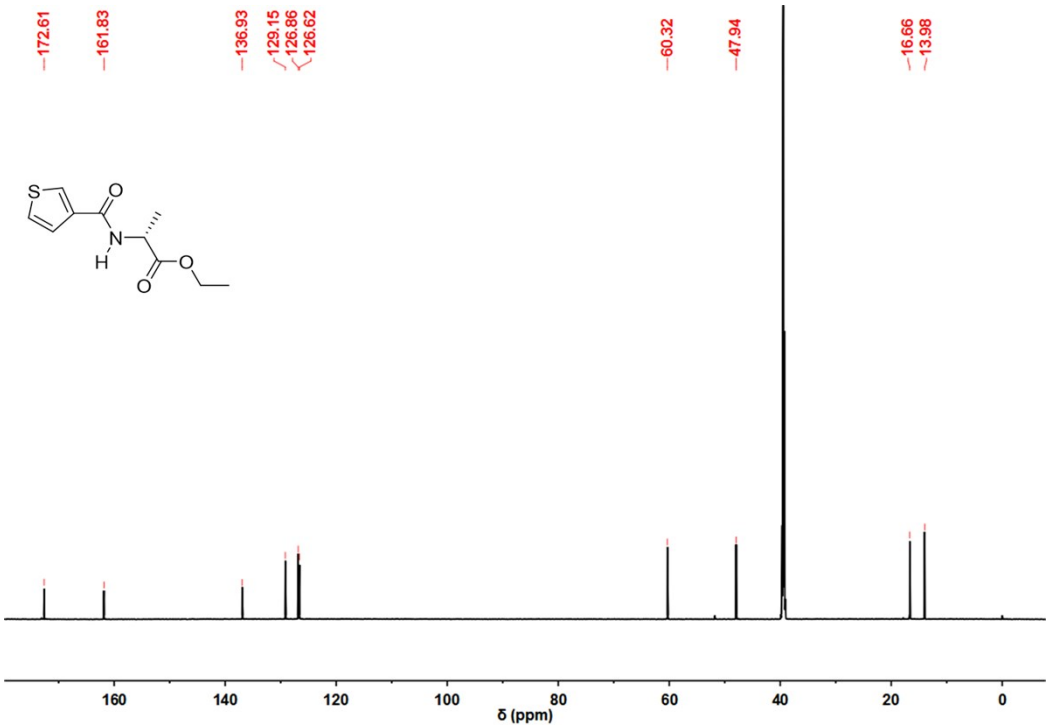


Figure S10. 214 MHz ^{13}C NMR spectrum of D-**SAOEt** in DMSO-d_6

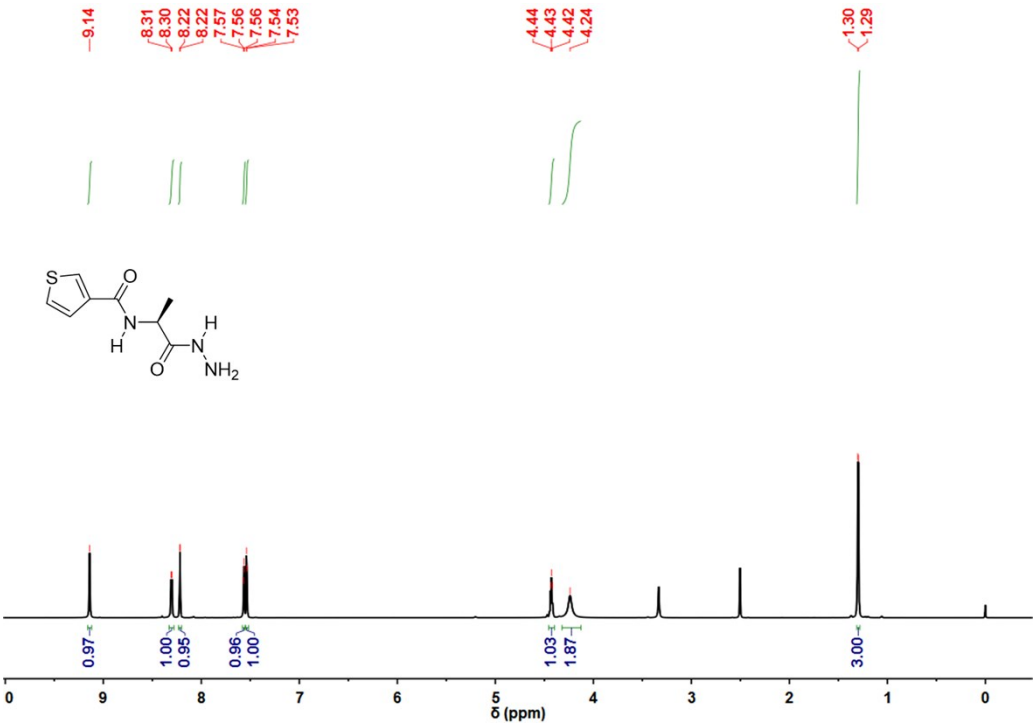


Figure S11. 850 MHz ^{1}H NMR spectrum of L-**SAN$_2$H$_3$** in DMSO-d_6

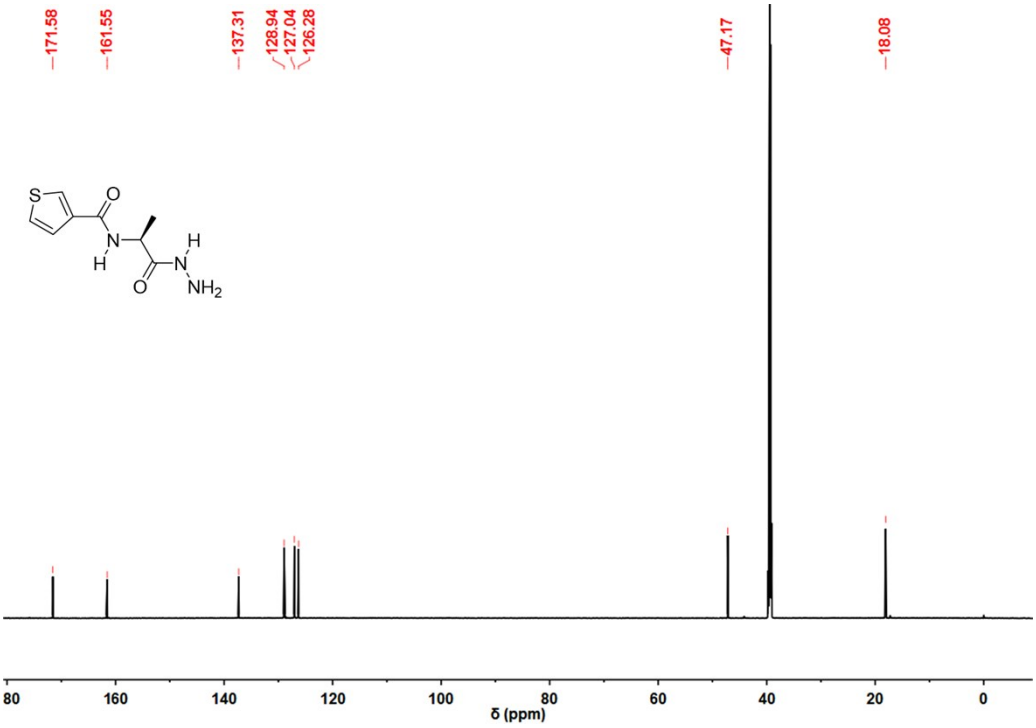


Figure S12. 214 MHz ^{13}C NMR spectrum of L-**SAN$_2$H$_3$** in DMSO-d_6

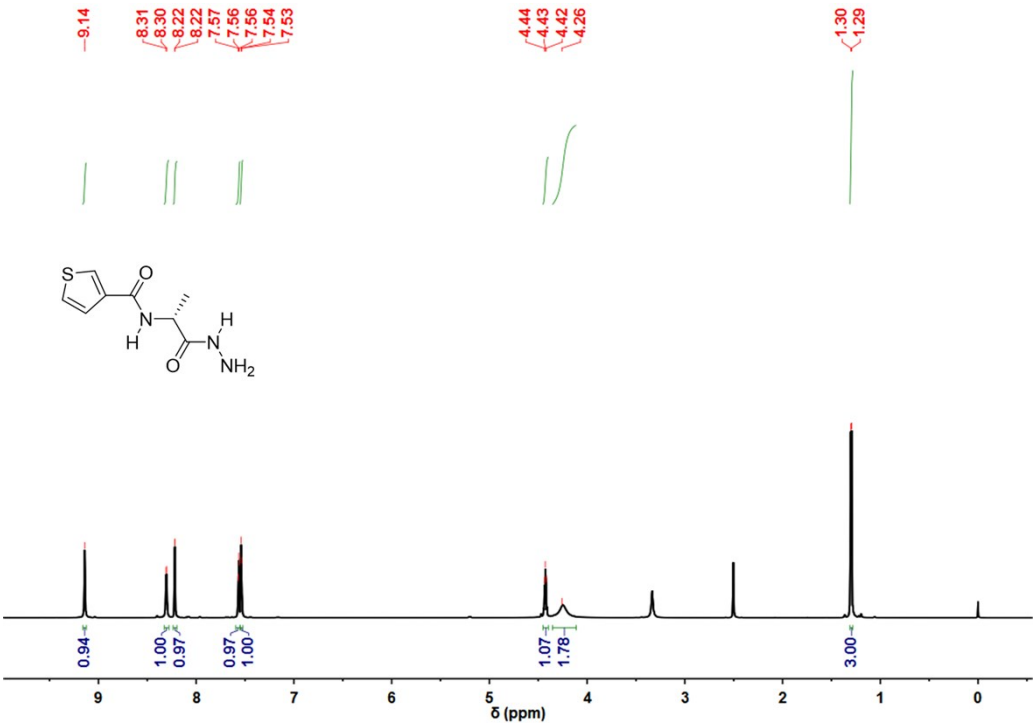


Figure S13. 850 MHz ^{1}H NMR spectrum of D-**SAN$_2$H$_3$** in DMSO-d_6

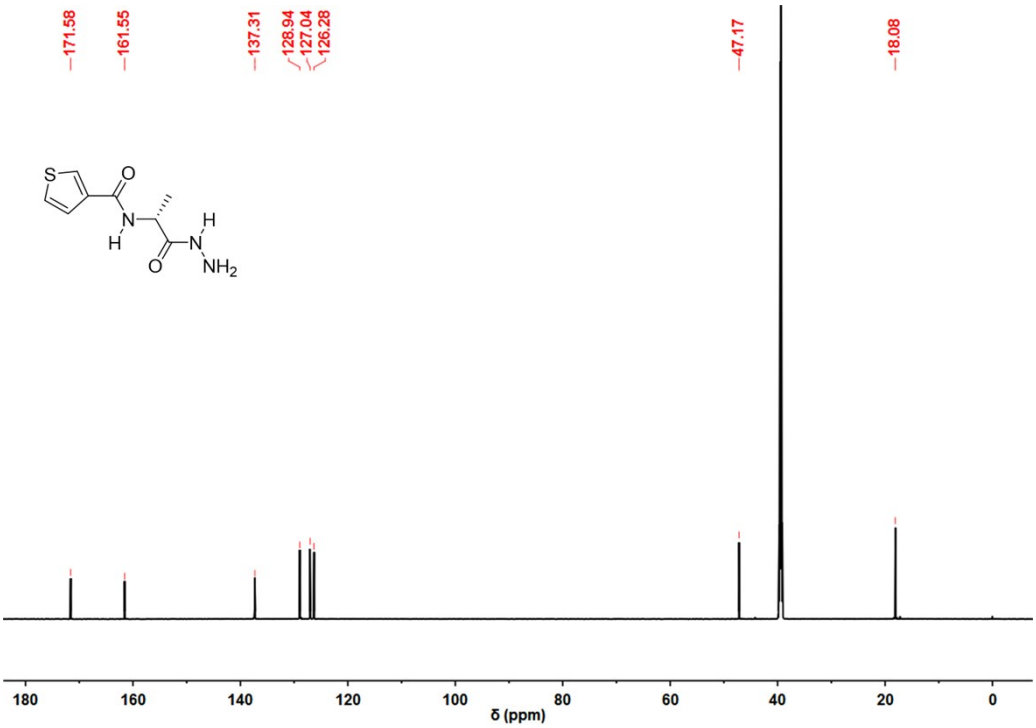


Figure S14. 214 MHz ^{13}C NMR spectrum of D-**SAN$_2$H$_3$** in DMSO-d_6

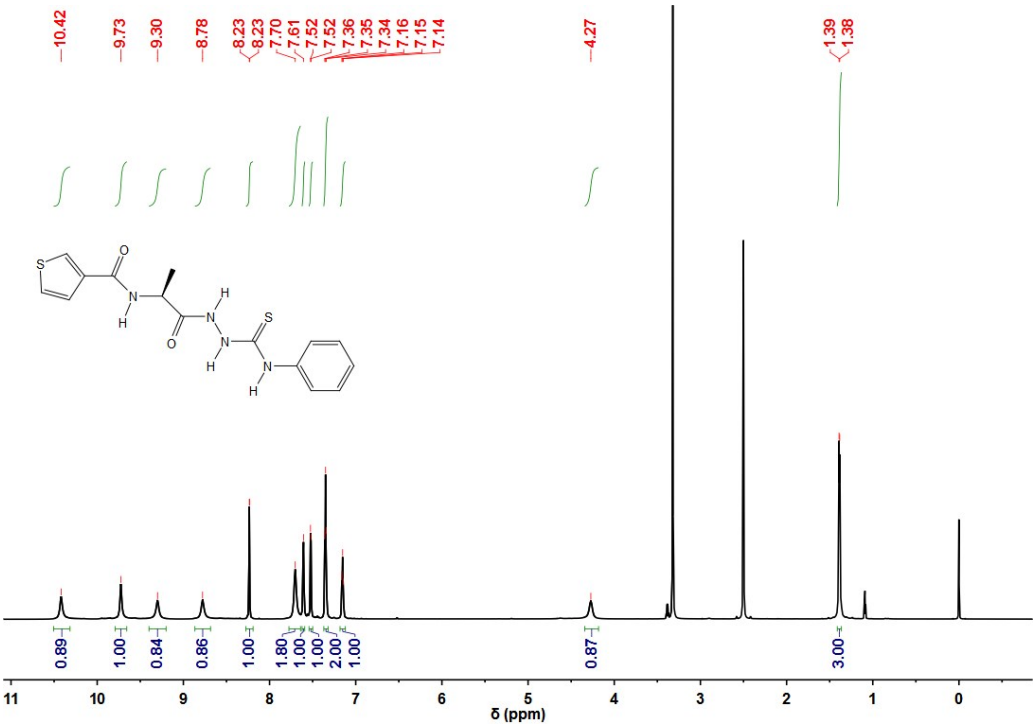


Figure S15. 850 MHz ^{1}H NMR spectrum of L-**SAP** in DMSO-d_6

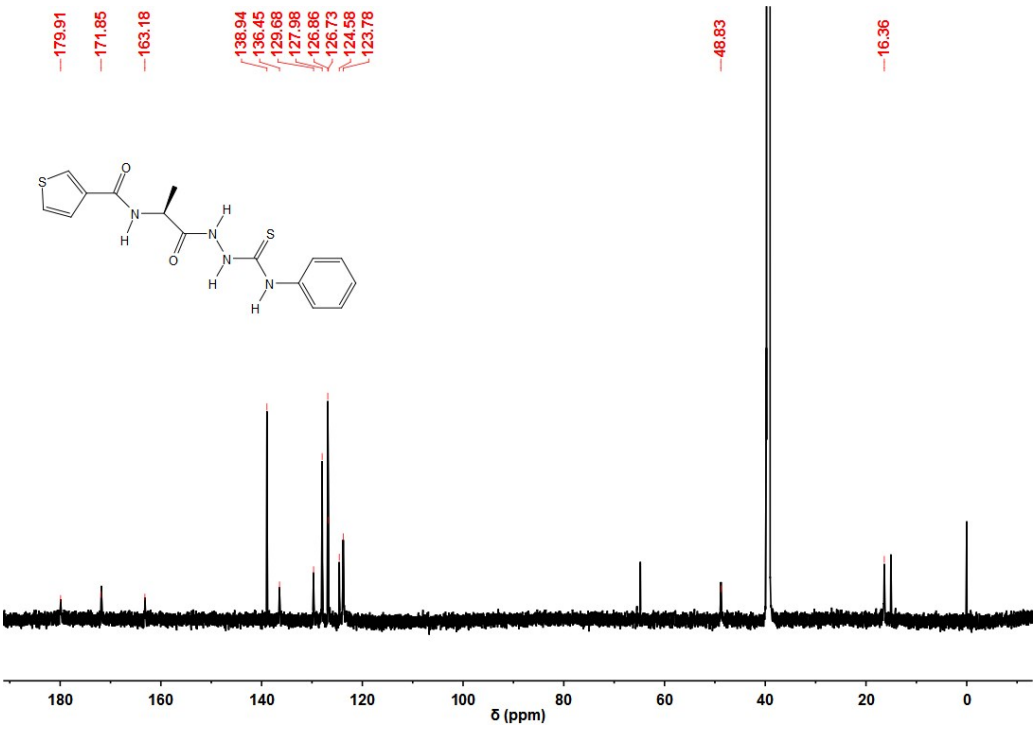


Figure S16. 214 MHz ^{13}C NMR spectrum of L-**SAP** in DMSO-d_6

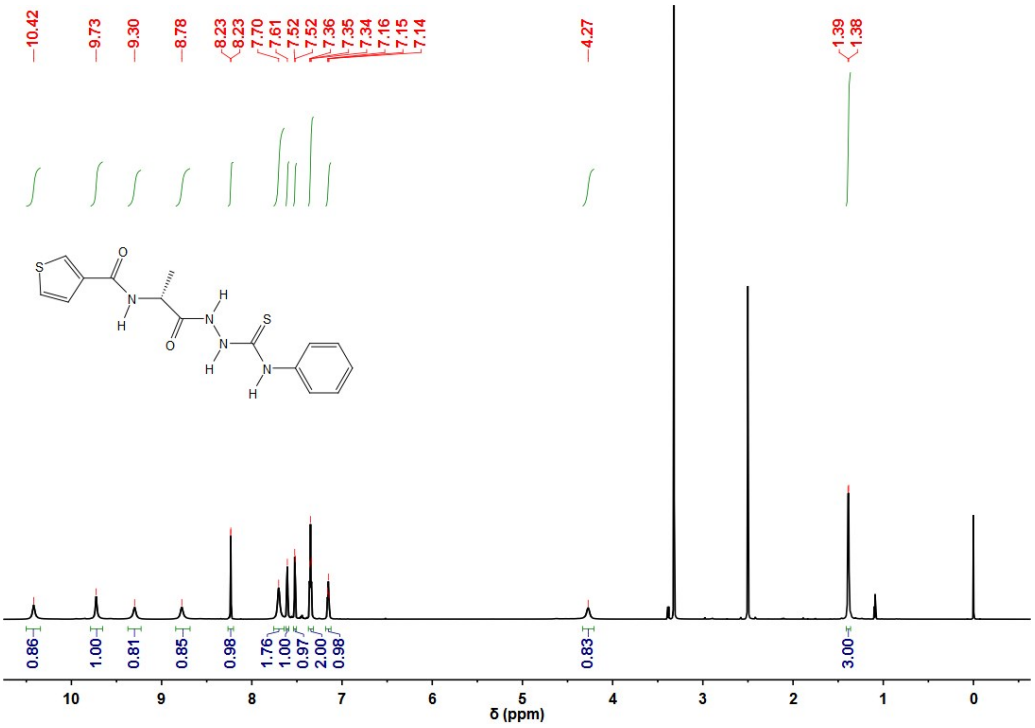


Figure S17. 850 MHz ^{1}H NMR spectrum of D-**SAP** in DMSO-d_6

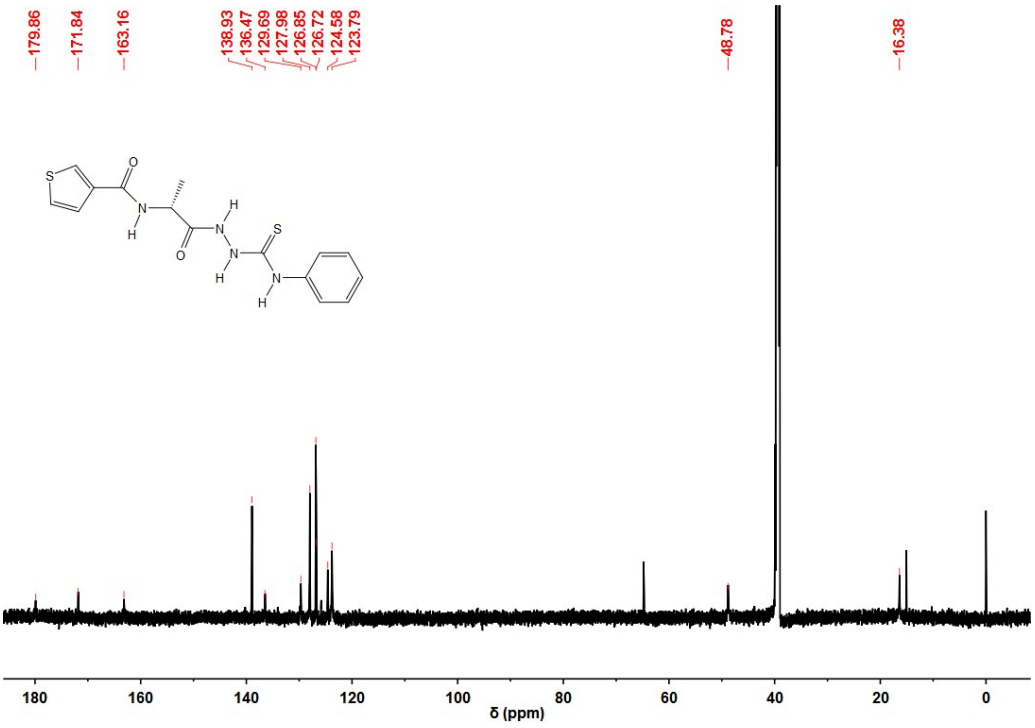


Figure S18. 214 MHz ^{13}C NMR spectrum of D-**SAP** in DMSO-d_6

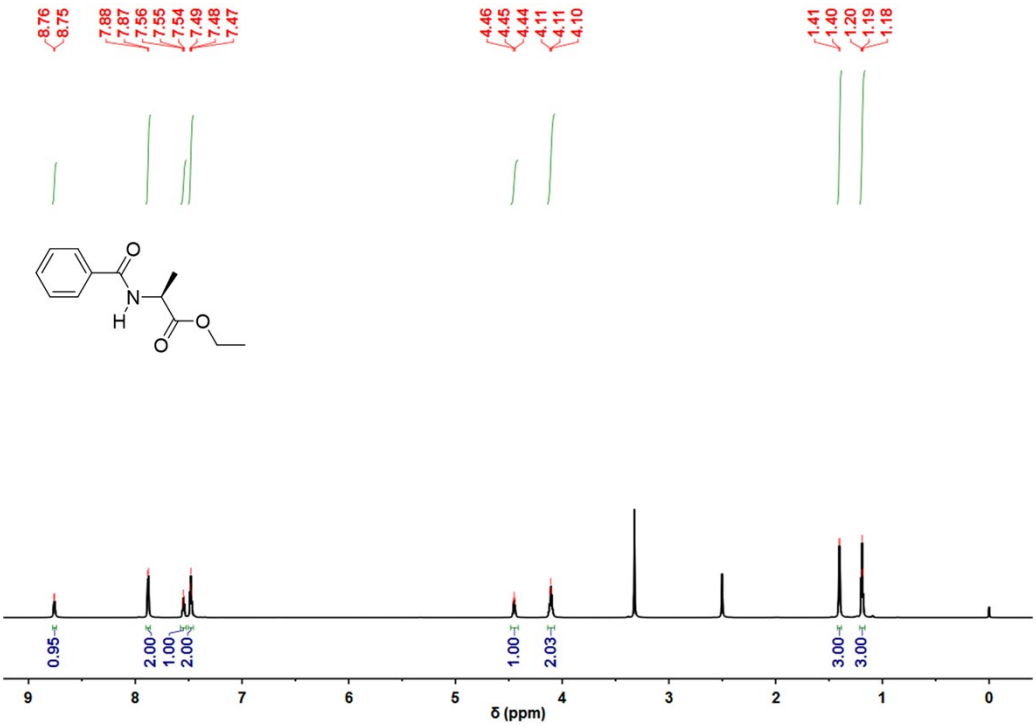


Figure S19. 850 MHz ^{1}H NMR spectrum of L-**PAOEt** in DMSO-d_6

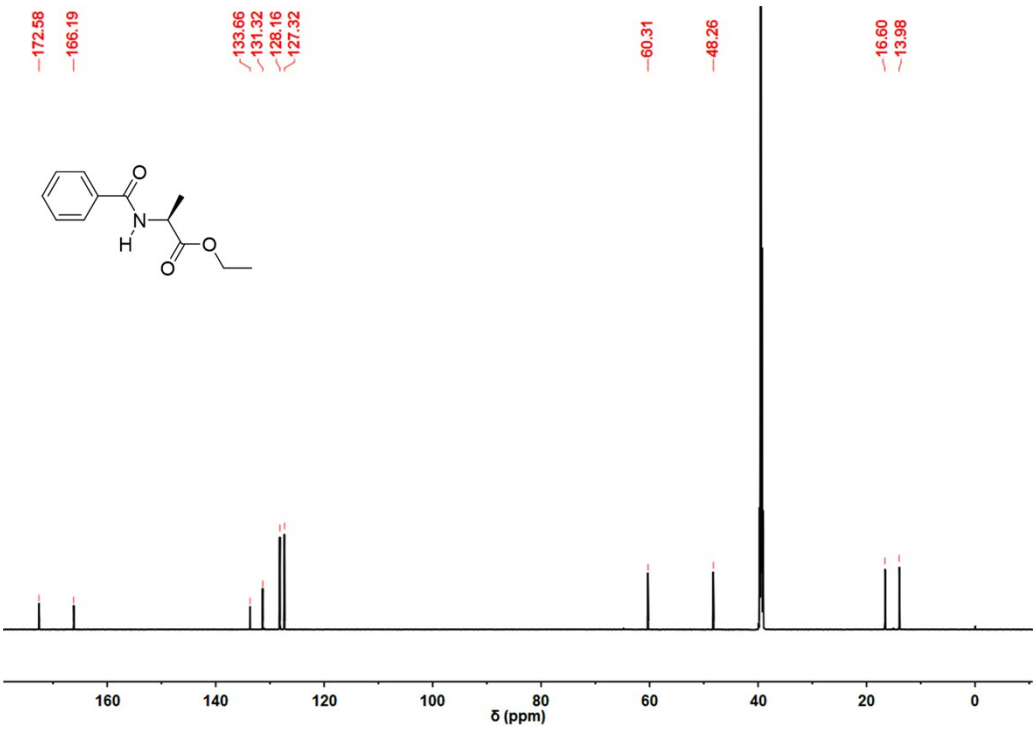


Figure S20. 214 MHz ^{13}C NMR spectrum of L-**PAOEt** in DMSO-d_6

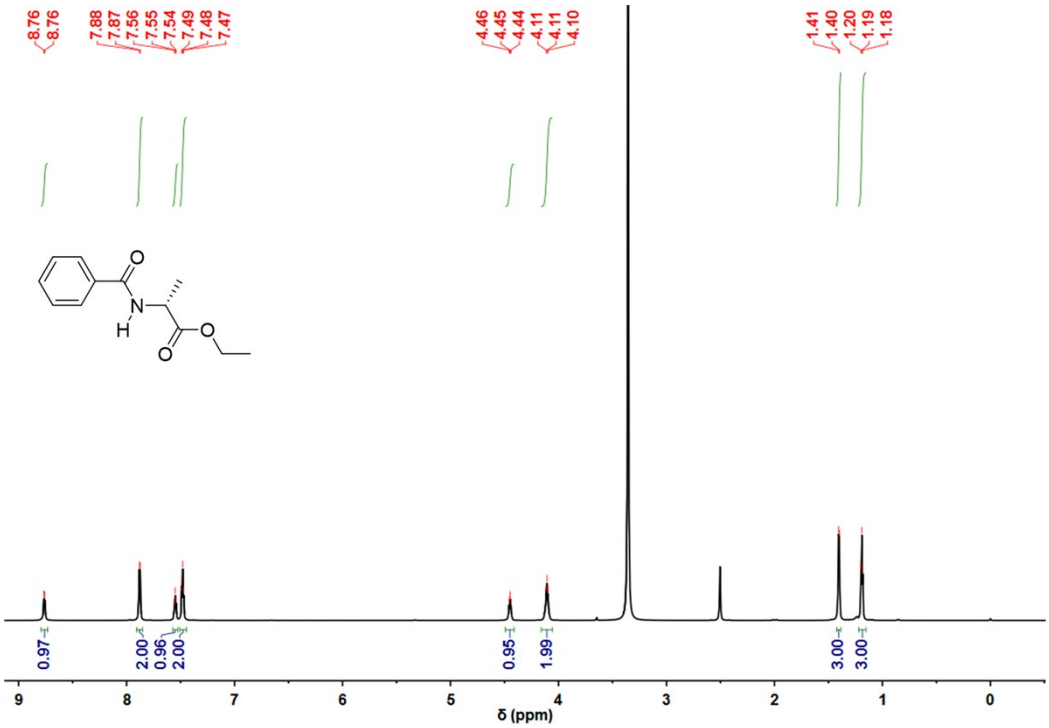


Figure S21. 850 MHz ^{1}H NMR spectrum of D-**PAOEt** in DMSO-d_6

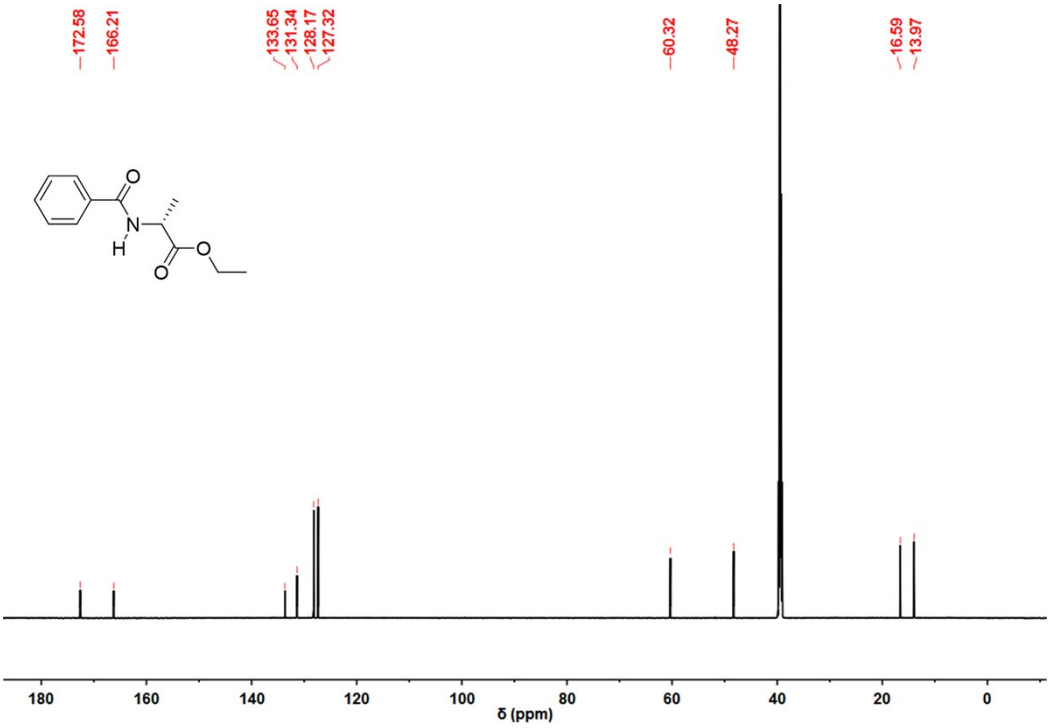


Figure S22. 214 MHz ^{13}C NMR spectrum of D-**PAOEt** in DMSO-d_6

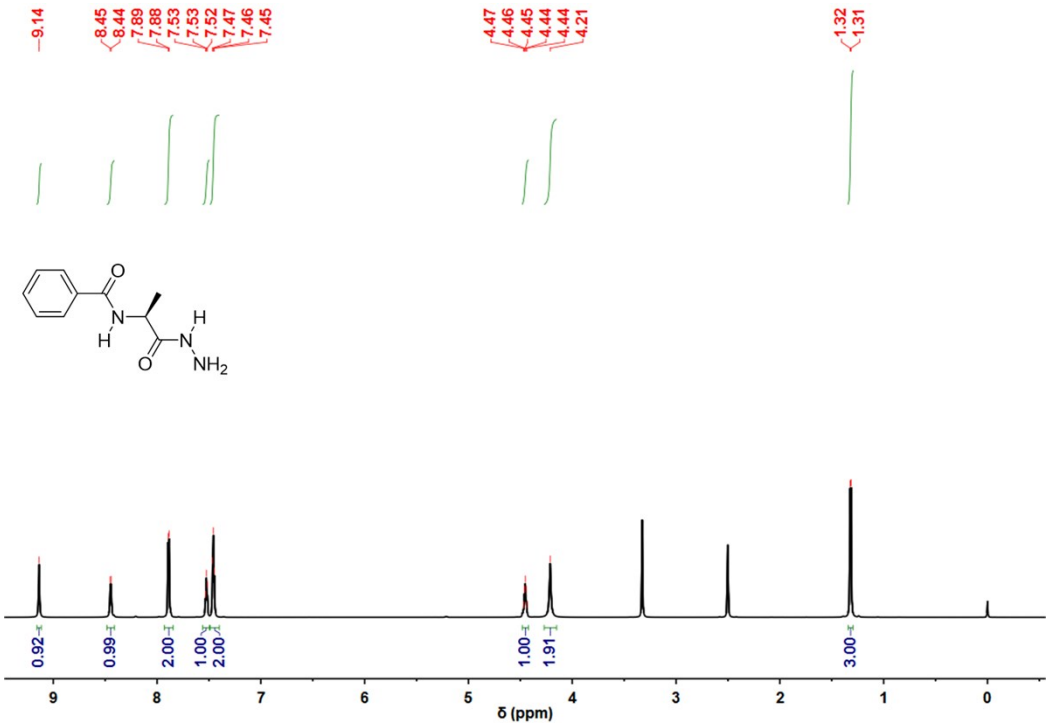


Figure S23. 850 MHz ^{1}H NMR spectrum of L-**PAN$_2$H$_3$** in DMSO-d_6

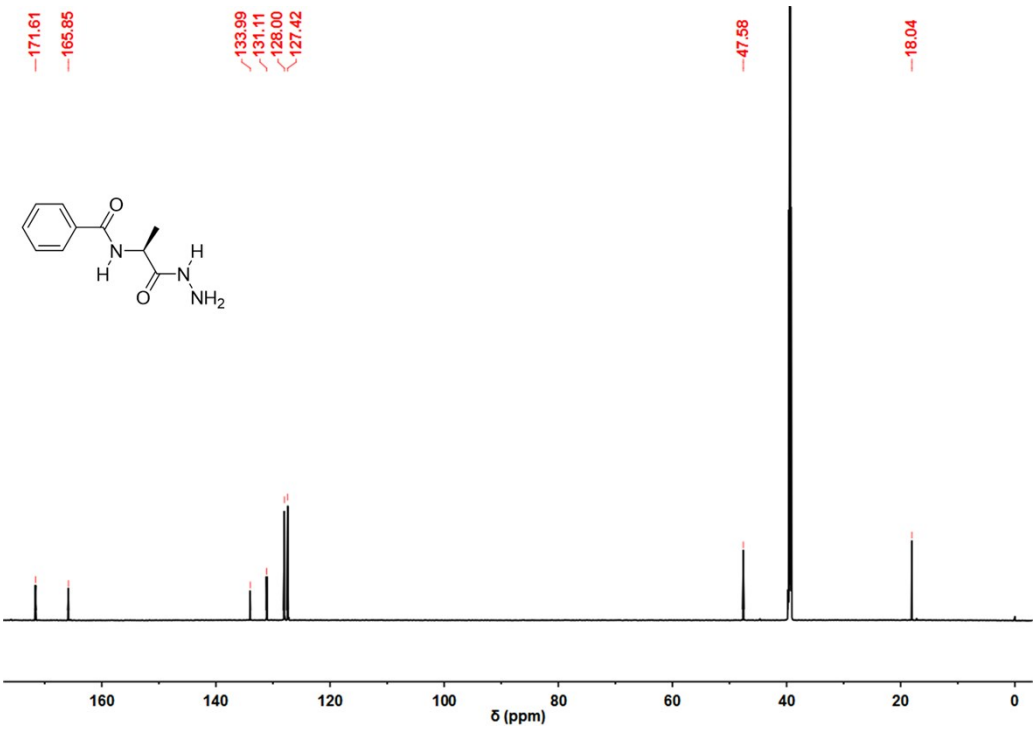


Figure S24. 214 MHz ^{13}C NMR spectrum of L-**PAN$_2$H$_3$** in DMSO-d_6

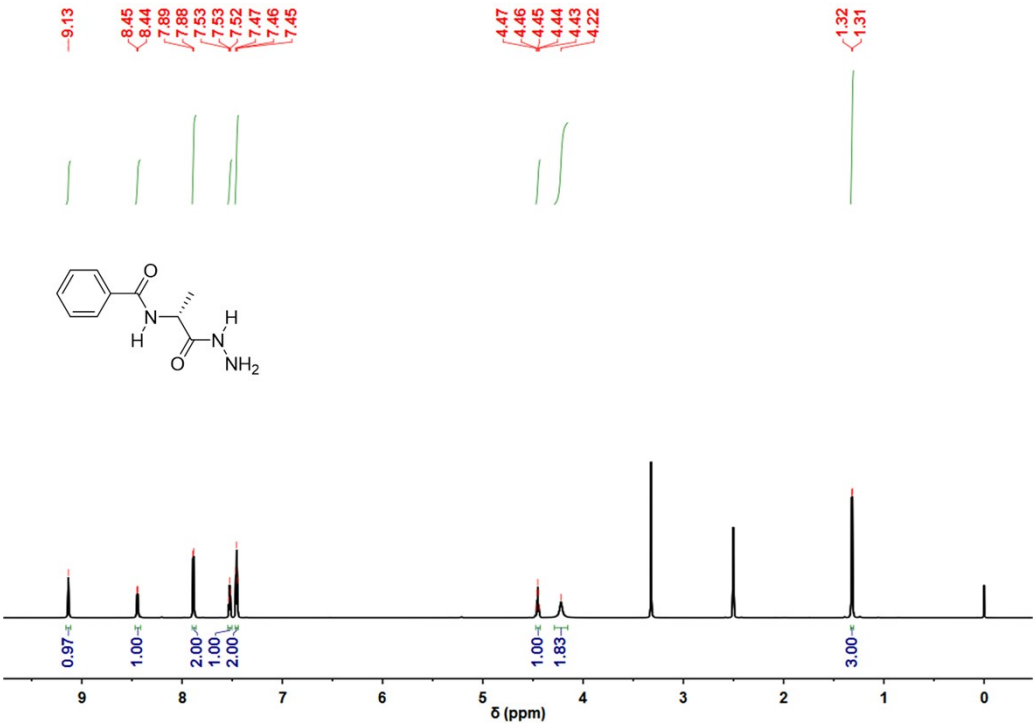


Figure S25. 850 MHz ^{1}H NMR spectrum of D-**PAN$_2$H$_3$** in DMSO-d_6

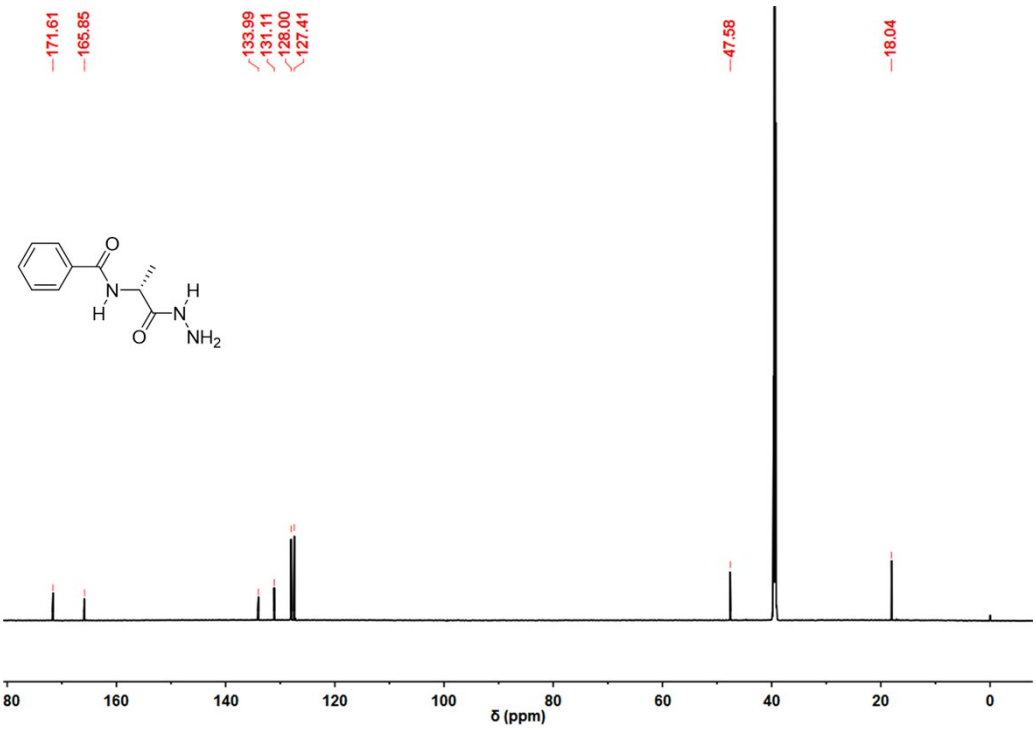


Figure S26. 214 MHz ^{13}C NMR spectrum of D-**PAN$_2$H$_3$** in DMSO-d_6

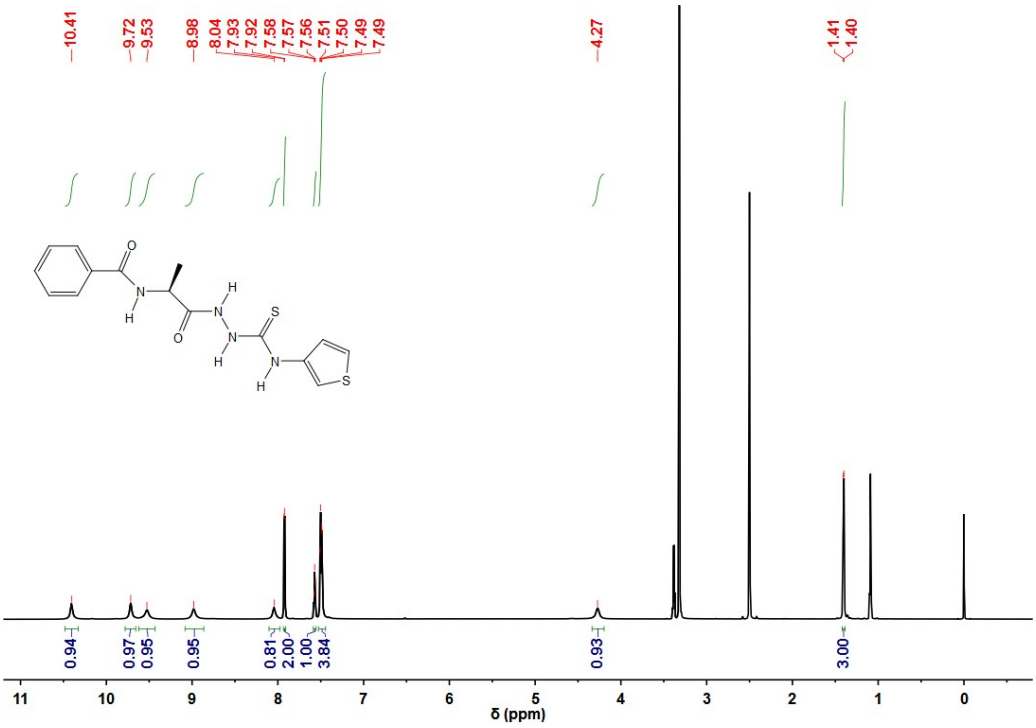


Figure S27. 850 MHz ^{1}H NMR spectrum of L-**PAS** in DMSO-d_6

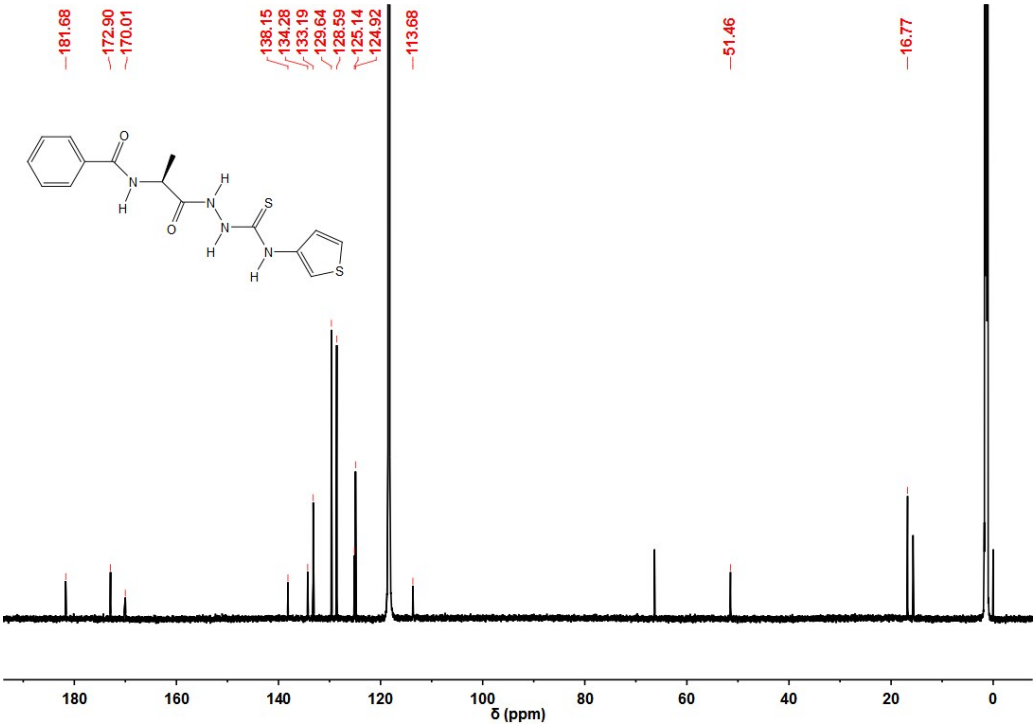


Figure S28. 214 MHz ^{13}C NMR spectrum of L-**PAS** in CD_3CN

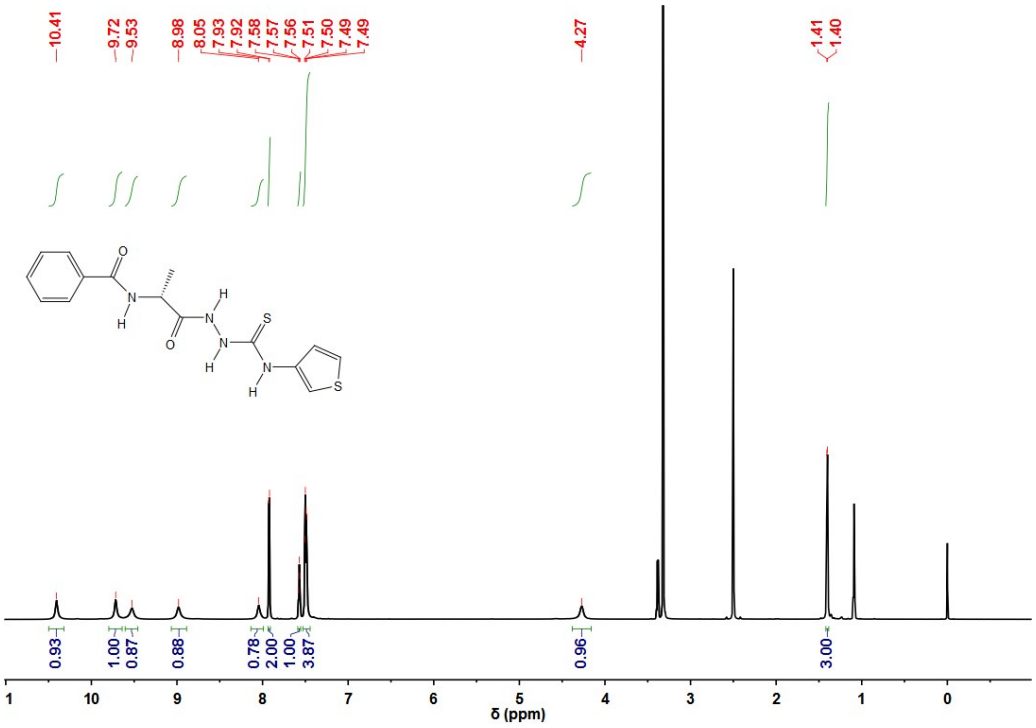


Figure S29. 850 MHz ^{1}H NMR spectrum of D-**PAS** in DMSO-d_6

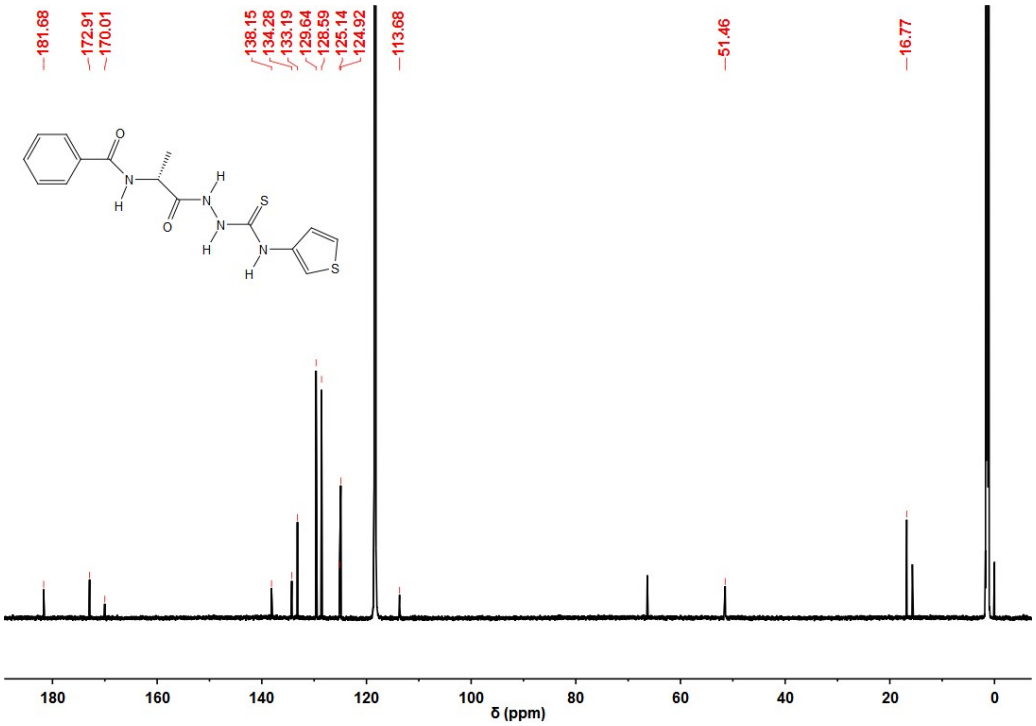


Figure S30. 214 MHz ^{13}C NMR spectrum of D-**PAS** in CD_3CN

5. ORTEP diagrams

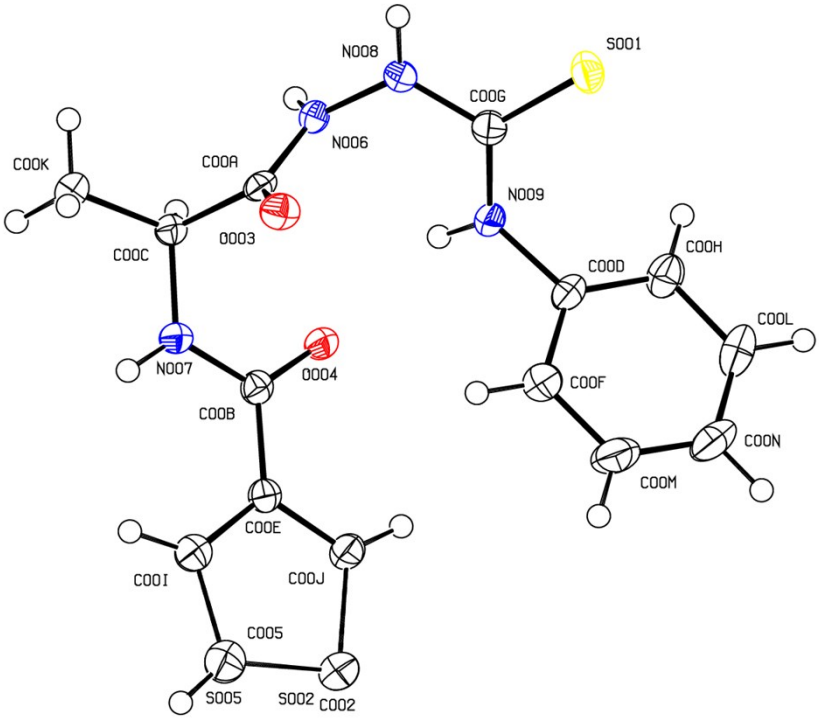


Figure S31. The ORTEP diagram for crystal structure of L-**SAP** with ellipsoids shown at the 50% probability level (arbitrary spheres for H atoms).

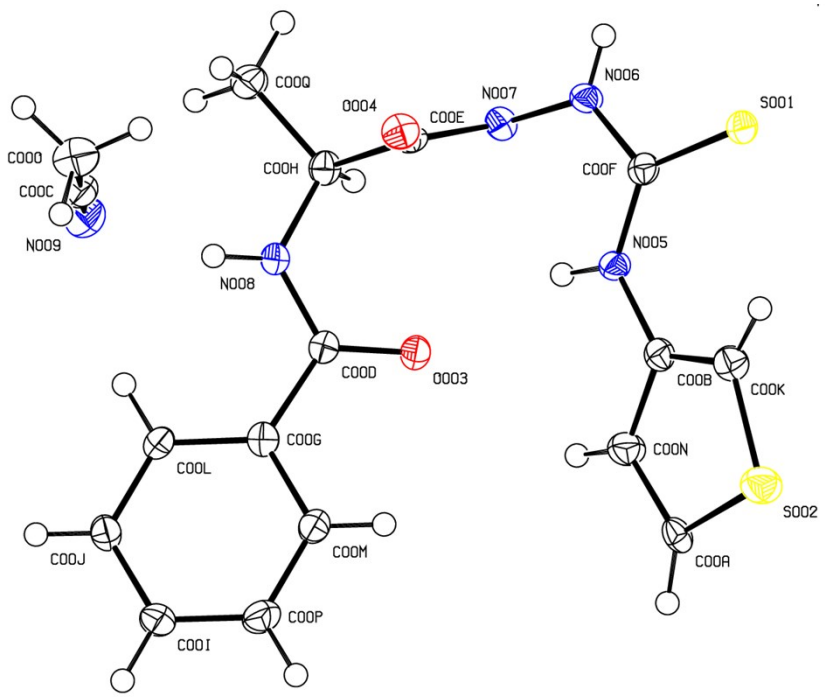


Figure S32. The ORTEP diagram for crystal structure of L-**PAS** with ellipsoids shown at the 50% probability level (arbitrary spheres for H atoms).