Electronic Supplementary Material (ESI) for Polymer Chemistry.

Supporting Information

Manipulating the helix-coil transition profile of synthetic polypeptides through leveraging side-chain molecular interactions

Ziyuan Song, Zhengzhong Tan, Xuetao Zheng, Zihuan Fu, Ettigounder Ponnusamy and Jianjun Cheng*

Corresponding email: jianjunc@illinois.edu

Table of Contents

Materials

All chemicals are purchased from MilliporeSigma (St. Louis, MO, USA) unless otherwise specified. Amino acids were purchased from Chem-Impex Inc. (Wood Dale, IL, USA). Dialysis membrane with a molecular-weight cut-off (MWCO) of 1 kDa was purchased from Repligen Corporation (Waltham, MA, USA). Anhydrous dimethylformamide (DMF) was treated with isocyanate beads (MilliporeSigma, St. Louis, MO, USA) to remove nucleophilic impurities, and stored in a glovebox. Anhydrous tetrahydrofuran (THF) and hexane was dried in a column using alumina. The monomer γ-chloropropyl-$_L$-glutamate NCA was synthesized following literature procedures.[1]

Instrumentation

^{1}H and ^{13}C NMR spectra were obtained on a Varian U500 spectrometer in NMR laboratory, University of Illinois. The chemical shifts of the spectra were referenced to the solvent proton and recorded in ppm. Gel permeation chromatography (GPC) characterizations were carried out on an instrument equipped with an isocratic pump (1260 Infinity II, Agilent, Santa Clara, CA, USA), a multi-angle static light scattering (MALS) detector with the detection wavelength at 658 nm (DAWN HELEOS-II, Wyatt Technology, Santa Barbara, CA, USA), and a differential refractometer (DRI) detector (Optilab T-rEX, Wyatt Technology, Santa Barbara, CA, USA). Separations were performed using serially connected size exclusion columns (three PLgel MIXED-B columns, 10 μm, 7.5 × 300 mm, Agilent, Santa Clara, CA, USA) using DMF containing 0.1 M LiBr at 40 °C as the mobile phase at a flow rate of 0.7 mL/min. The MALS detector was calibrated using pure toluene and can be used for the determination of the absolute molecular weights (MWs). All sample solutions were filtered using a 0.45 μm PTFE filter before injection. The dn/dc value of each sample was calculated offline by using the internal calibration system processed by the ASTRA 7 software (version 7.1.3.15, Wyatt Technology, Santa Barbara, CA, USA) and was used to determine the MWs of polypeptides. Circular dichroism (CD) measurements were carried out on a JASCO J-815 CD spectrometer (JASCO, Easton, MD, USA). The mean residue molar ellipticity of each polypeptide was calculated on the basis of the measured apparent ellipticity by the following formulas: Ellipticity ([θ] in deg cm^2 dmol^{-1}) = (milidegrees × mean residue molecular weight)/(path length in millimeters × concentration of polypeptide in mg/mL).[2, 3] Helicity was calculated by the following equation: helicity = (-[$\theta_{222\ nm}$] + 3000)/39000 × 100%.[4] Lyophilization was performed on a FreeZone lyophilizer (Labconco, Kansas City, MO, USA).

Synthesis of azide-bearing polypeptides

The azide-containing polypeptide, poly(γ-azidopropyl-$_L$-glutamate) (PAPLG), was synthesized as previously reported.[1]

Synthesis of alkyne-functionalized ammonium salts

Synthesis of alkyne-functionalized ammonium salts was achieved through the reaction of 3-dimethylamino-1-propyne with electrophiles. Typically, 3-dimethylamino-1-propyne (500 μL, 4.64 mmol) was dissolved in DMF (1.0 mL), into which the DMF solution of 1-bromopropane (422 μL, 4.64 mmol, 1.0 equiv.) was added. The mixture was stirred at 50 °C overnight. The crude product was purified by precipitation in ether (10 mL) and isolated by centrifugation, which was washed by ether for 3 times and dried under vacuum. The product *N*-propyl-*N*,*N*-dimethylpropargylammonium bromide (PrDA) was obtained as white powder (827 mg, 87% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.26 (s, 2H, HC≡CCH_2-), 3.41 (m, 2H, $CH_3CH_2CH_2N^+$-), 3.18 (s, 7H, -(CH_3)$_2N^+$- and HC≡C-), 1.80 (m, 2H, $CH_3CH_2CH_2N^+$-), 1.00 (t, J = 7.32, 3H, $CH_3CH_2CH_2N^+$-). ^{13}C NMR (D_2O, δ, 125 MHz): 81.6, 70.9, 66.0, 54.4, 50.6, 16.1, 10.0. HR-MS (ESI, *m*/*z*): [M^+] Calcd. For $C_8H_{16}N$ 126.1283; Found: 126.1288.

Synthesis of *N*-pentyl-*N*,*N*-dimethylpropargylammonium bromide (PeDA) was performed following the above procedure, but using 1-bromopentane as the electrophile. The product was obtained as viscous solid (81% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.43 (s, 2H, HC≡CCH_2-), 3.58 (m, 2H, $CH_3(CH_2)_2CH_2CH_2N^+$-), 3.48 (s, 1H, HC≡C-), 3.32 (s, 6H, -(CH_3)$_2N^+$-), 1.92 (m, 2H, $CH_3(CH_2)_2CH_2CH_2N^+$-), 1.50 (m, 4H, $CH_3(CH_2)_2CH_2CH_2N^+$-), 1.04 (t, J = 6.04, 3H, $CH_3(CH_2)_2CH_2N^+$-). ^{13}C NMR (D_2O, δ, 125 MHz): 82.2, 71.6, 64.7, 54.5, 51.0, 28.0, 22.2, 22.0, 13.8. HR-MS (ESI, *m*/*z*): [M^+] Calcd. For $C_{10}H_{20}N$ 154.1596; Found: 154.1594.

Synthesis of *N*-hexyl-*N*,*N*-dimethylpropargylammonium bromide (HexDA) was performed following the above procedure, but using 1-bromohexane as the electrophile. The product was obtained as viscous solid (84% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.28 (s, 2H, HC≡CCH_2-), 3.46 (m, 2H, $CH_3(CH_2)_3CH_2CH_2N^+$-), 3.30 (t, J = 2.53, 1H, HC≡C-), 3.20 (s, 6H, -(CH_3)$_2N^+$-), 1.81 (m, 2H, $CH_3(CH_2)_3CH_2CH_2N^+$-), 1.42-1.32 (m, 6H, $CH_3(CH_2)_3CH_2CH_2N^+$-), 0.90 (t, J = 7.12, 3H, $CH_3(CH_2)_3CH_2N^+$-). ^{13}C NMR (D_2O, δ, 125 MHz): 81.7, 71.3, 64.6, 54.2, 50.7, 30.6, 25.3, 22.2, 22.0, 13.6. HR-MS (ESI, *m*/*z*): [M^+] Calcd. For $C_{11}H_{22}N$ 168.1752; Found: 168.1745.

Synthesis of *N*-heptyl-*N*,*N*-dimethylpropargylammonium bromide (HeptDA) was performed following the above procedure, but using 1-bromoheptane as the electrophile. The product was obtained as viscous solid (84% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.39 (s, 2H, HC≡CCH_2-), 3.54 (m, 2H, $CH_3(CH_2)_4CH_2CH_2N^+$-), 3.44 (t, J = 2.40, 1H, HC≡C-), 3.29 (s, 6H, -(CH_3)$_2N^+$-), 1.87 (m, 2H, $CH_3(CH_2)_4CH_2CH_2N^+$-), 1.49-1.36 (m, 8H, $CH_3(CH_2)_4CH_2CH_2N^+$-), 0.97 (t, J = 6.84, 3H, $CH_3(CH_2)_4CH_2N^+$-). ^{13}C NMR (D_2O, δ, 125 MHz): 82.0, 71.5, 64.5, 54.3, 51.0, 31.4, 28.4, 25.9, 22.5, 22.5, 14.1. HR-MS (ESI, *m*/*z*): [M^+] Calcd. For $C_{12}H_{24}N$ 182.1909; Found: 182.1916.

Synthesis of *N*-octyl-*N*,*N*-dimethylpropargylammonium bromide (OctDA) was performed following the above procedure, but using 1-bromooctane as the electrophile. The product was obtained as viscous solid (84% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.42 (s, 2H, HC≡CCH_2-), 3.51 (m, 2H, $CH_3(CH_2)_5CH_2CH_2N^+$-), 3.40 (t, J = 2.40, 1H, HC≡C-), 3.26 (s, 6H, -(CH_3)$_2N^+$-), 1.84 (m, 2H, $CH_3(CH_2)_5CH_2CH_2N^+$-), 1.46-1.32 (m, 10H, $CH_3(CH_2)_5CH_2CH_2N^+$-), 0.93 (t, J = 6.88, 3H, C$H_3(CH_2)_5CH_2N^+$-). ^{13}C NMR (D_2O, δ, 125 MHz): 81.9, 71.4, 64.3, 54.2, 51.0, 31.6, 28.9, 28.8, 26.0, 22.6, 22.5, 14.0. HR-MS (ESI, m/z): [M^+] Calcd. For $C_{13}H_{26}N$ 196.2065; Found: 196.2068.

Synthesis of *N*-*tert*-butoxycarbonylmethyl-*N*,*N*-dimethylpropargylammonium bromide (tBuCMDA) was performed following the above procedure, but using *tert*-butyl bromoacetate as the electrophile. The product was obtained as white powder (88% yield). ^{1}H NMR (D_2O, δ, 500 MHz): 4.46 (s, 2H, $(CH_3)_3COOCCH_2$-), 4.30 (s, 2H, HC≡CCH_2-), 3.32 (s, 6H, -(CH_3)$_2N^+$-), 3.28 (t, J = 2.51, 1H, HC≡C-), 1.49 (s, 9H, $(CH_3)_3COOCCH_2$-). ^{13}C NMR (D_2O, δ, 125 MHz): 164.1, 86.9, 82.7, 70.8, 61.9, 55.5, 52.0, 27.4. HR-MS (ESI, m/z): [M^+] Calcd. For $C_{11}H_{20}NO_2$ 198.1494; Found: 198.1501.

Synthesis of triazole polypeptides

Triazole polypeptides were synthesized with the copper-catalyzed, Huisgen click reaction between azide-bearing polypeptides and alkyne-containing molecules. Generally, PAPLG (40 mg, 0.19 mmol of side-chain azide groups), *N,N,N',N'',N''*-pentamethyldiethylenetriamine (PMDETA) (19.7 μL, 0.09 mmol) and alkyne-containing molecules (0.28 mmol, 1.5 equiv.) were dissolved in DMF (1 mL) in a glovebox. CuBr (2.7 mg, 0.02 mmol) was then added into the solution and the mixture was stirred at room temperature for 24 h. The reaction was then quenched by exposing to air. The resulting solution was dialyzed in a dialysis membrane (MWCO = 1 kDa) against EDTA solution for 1 h to remove the copper salts, and then against DI water for 6 h (water changed every hour). The final triazole polypeptides were obtained after lyophilization (65-82% yield).

P1-Pr was synthesized from the reaction between PAPLG and PrDA. ^{1}H NMR (TFA-*d*, δ, 500 MHz): 8.55 (s, 1H, triazole-H), 5.01-4.76 (m, 5H, α-H, -COOCH_2CH$_2$CH$_2$-, and -C$H_2N^+(CH_3)_2$-), 4.41 (s, 2H, -COOCH$_2$CH$_2$CH_2-), 3.48 (s, 2H, $CH_3CH_2CH_2(CH_3)_2N^+$-), 3.26 (s, 6H, $CH_3CH_2CH_2(CH_3)_2N^+$-), 2.76 (s, 2H, -CH$_2$CH_2COO-), 2.54 (s, 2H, -COOCH$_2$CH_2CH$_2$-), 2.48-2.17 (m, 2H, -CH_2CH$_2$COO-), 2.05 (s, 2H, $CH_3CH_2CH_2(CH_3)_2N^+$-), 1.18 (s, 3H, C$H_3CH_2CH_2(CH_3)_2N^+$-).

P1-Pe was synthesized from the reaction between PAPLG and PeDA. ^{1}H NMR (TFA-*d*, δ, 500 MHz): 8.63 (s, 1H, triazole-H), 5.11-4.75 (m, 5H, α-H, -COOCH_2CH$_2$CH$_2$-, and -C$H_2N^+(CH_3)_2$-), 4.46 (s, 2H, -COOCH$_2$CH$_2$CH_2-), 3.58 (s, 2H, $CH_3(CH_2)_2CH_2CH_2(CH_3)_2N^+$-), 3.30 (s, 6H, $CH_3(CH_2)_2CH_2CH_2(CH_3)_2N^+$-), 2.83 (s, 2H, -CH$_2CH_2$COO-), 2.58 (s, 2H, -COOCH$_2$CH_2CH$_2$-), 2.54-2.23

(m, 2H, -C*H*$_2$CH$_2$COO-), 2.07 (s, 2H, CH$_3$(CH$_2$)$_2$C*H*$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 1.56 (m, 4H, CH$_3$(C*H*$_2$)$_2$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 1.08 (s, 3H, C*H*$_3$(CH$_2$)$_2$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-).

P1-Hex was synthesized from the reaction between PAPLG and HexDA. ^{1}H NMR (TFA-*d*, *δ*, 500 MHz): 8.45 (s, 1H, triazole-H), 4.90-4.57 (m, 5H, α-H, -COOC*H*$_2$CH$_2$CH$_2$-, and -C*H*$_2$N$^+$(CH$_3$)$_2$-), 4.26 (s, 2H, -COOCH$_2$CH$_2$C*H*$_2$-), 3.38 (s, 2H, CH$_3$(CH$_2$)$_3$CH$_2$C*H*$_2$(CH$_3$)$_2$N$^+$-), 3.10 (s, 6H, CH$_3$(CH$_2$)$_3$CH$_2$CH$_2$(C*H*$_3$)$_2$N$^+$-), 2.60 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.39 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.33-2.03 (m, 2H, -C*H*$_2$CH$_2$COO-), 1.86 (s, 2H, CH$_3$(CH$_2$)$_3$C*H*$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 1.31 (m, 6H, CH$_3$(C*H*$_2$)$_3$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 0.84 (s, 3H, C*H*$_3$(CH$_2$)$_3$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-).

P1-Hept was synthesized from the reaction between PAPLG and HeptDA. ^{1}H NMR (TFA-*d*, *δ*, 500 MHz): 8.51 (s, 1H, triazole-H), 5.00-4.65 (m, 5H, α-H, -COOC*H*$_2$CH$_2$CH$_2$-, and -C*H*$_2$N$^+$(CH$_3$)$_2$-), 4.36 (s, 2H, -COOCH$_2$CH$_2$C*H*$_2$-), 3.47 (s, 2H, CH$_3$(CH$_2$)$_4$CH$_2$C*H*$_2$(CH$_3$)$_2$N$^+$-), 3.19 (s, 6H, CH$_3$(CH$_2$)$_4$CH$_2$CH$_2$(C*H*$_3$)$_2$N$^+$-), 2.73 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.48 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.40-2.12 (m, 2H, -C*H*$_2$CH$_2$COO-), 1.95 (s, 2H, CH$_3$(CH$_2$)$_4$C*H*$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 1.50-1.28 (m, 8H, CH$_3$(C*H*$_2$)$_4$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 0.92 (s, 3H, C*H*$_3$(CH$_2$)$_4$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-).

P1-Oct was synthesized with the reaction between PAPLG and OctDA. ^{1}H NMR (TFA-*d*, *δ*, 500 MHz): 8.69 (s, 1H, triazole-H), 4.98-4.47 (m, 5H, α-H, -COOC*H*$_2$CH$_2$CH$_2$-, and -C*H*$_2$N$^+$(CH$_3$)$_2$-), 4.27 (s, 2H, -COOCH$_2$CH$_2$C*H*$_2$-), 3.43 (s, 2H, CH$_3$(CH$_2$)$_5$CH$_2$C*H*$_2$(CH$_3$)$_2$N$^+$-), 3.14 (s, 6H, CH$_3$(CH$_2$)$_5$CH$_2$CH$_2$(C*H*$_3$)$_2$N$^+$-), 2.60 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.42 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.35-2.04 (m, 2H, -C*H*$_2$CH$_2$COO-), 1.86 (s, 2H, CH$_3$(CH$_2$)$_5$C*H*$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 1.45-1.10 (m, 10H, CH$_3$(C*H*$_2$)$_5$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-), 0.81 (s, 3H, C*H*$_3$(CH$_2$)$_5$CH$_2$CH$_2$(CH$_3$)$_2$N$^+$-).

P2-CA was synthesized with the reaction between PAPLG and 4-pentynoic acid. ^{1}H NMR (TFA-*d*, *δ*, 500 MHz): 8.30 (s, 1H, triazole-H), 4.70 (s, 3H, α-H and -COOC*H*$_2$CH$_2$CH$_2$-), 4.28 (s, 2H, -COOCH$_2$CH$_2$C*H*$_2$-), 3.26 (s, 2H, -C*H*$_2$CH$_2$COOH), 2.95 (s, 2H, -CH$_2$C*H*$_2$COOH), 2.64 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.43 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.34-2.02 (m, 2H, -C*H*$_2$CH$_2$COO-).

P2-ZW was synthesized with the reaction between PAPLG and tBuCMDA. HCl was added during the dialysis to facilitate the hydrolysis of *t*Bu esters to generate side-chain zwitterions. ^{1}H NMR (TFA-*d*): 8.43 (s, 1H, triazole-H), 5.12 (s, 2H, -C*H*$_2$N$^+$(CH$_3$)$_2$CH$_2$COOH), 4.65 (m, 3H, α-H and -COOC*H*$_2$CH$_2$CH$_2$-), 4.26 (s, 4H, -C*H*$_2$COOH and -COOCH$_2$CH$_2$C*H*$_2$-), 3.41 (s, 6H, -N$^+$(C*H*$_3$)$_2$CH$_2$COOH), 2.63 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.37 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.34-2.03 (m, 2H, -C*H*$_2$CH$_2$COO-).

P3 was synthesized with the reaction between PAPLG and 3-dimethylamino-1-propyne. ^{1}H NMR (TFA-*d*, *δ*, 500 MHz): 8.84 (s, 1H, triazole-H), 4.84 (m, 5H, (CH$_3$)$_2$NC*H*$_2$-, α-H and -COOC*H*$_2$CH$_2$CH$_2$-), 4.33 (s,

2H, -COOCH$_2$CH$_2$C*H*$_2$-), 3.13 (s, 6H, (C*H*$_3$)$_2$N-), 2.69 (s, 2H, -CH$_2$C*H*$_2$COO-), 2.48 (s, 2H, -COOCH$_2$C*H*$_2$CH$_2$-), 2.40-2.07 (m, 2H, -C*H*$_2$CH$_2$COO-).

Supporting Figures

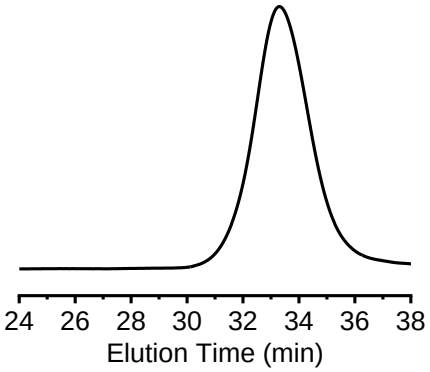

Fig. S1. GPC-dRI trace of poly(γ-chloropropyl-$_L$-glutamate) (PCPLG) (M_n = 9.47 kDa. $Đ = M_w/M_n$ = 1.18).

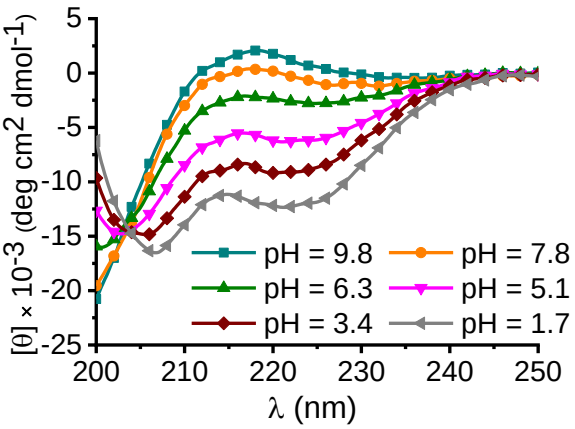

Fig. S2. Overlaid CD spectra of **P1**-Pr at various aqueous pH values.

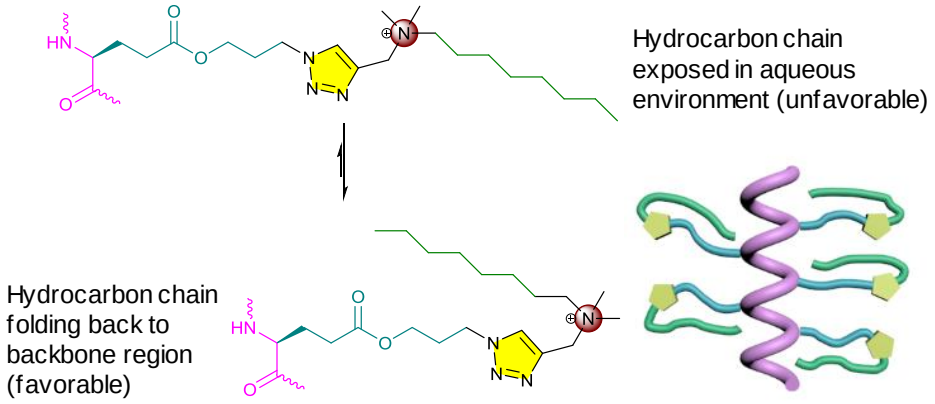

Fig. S3. Proposed mechanism for the formation of stable α-helical structure of **P1**-Oct under neutral and basic conditions.

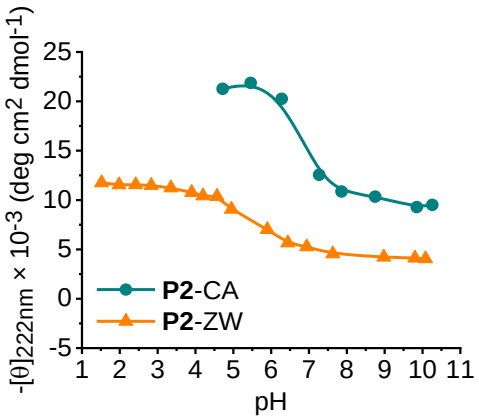

Fig. S4. pH-ellipticity plot of **P2**-CA and **P2**-ZW. The ellipticity at 222 nm was selected to indicate the helicity of polypeptides. **P2**-CA showed decreased water-solubility under acidic conditions, and the CD spectra were not collected at pH < 4.5.

NMR Spectra

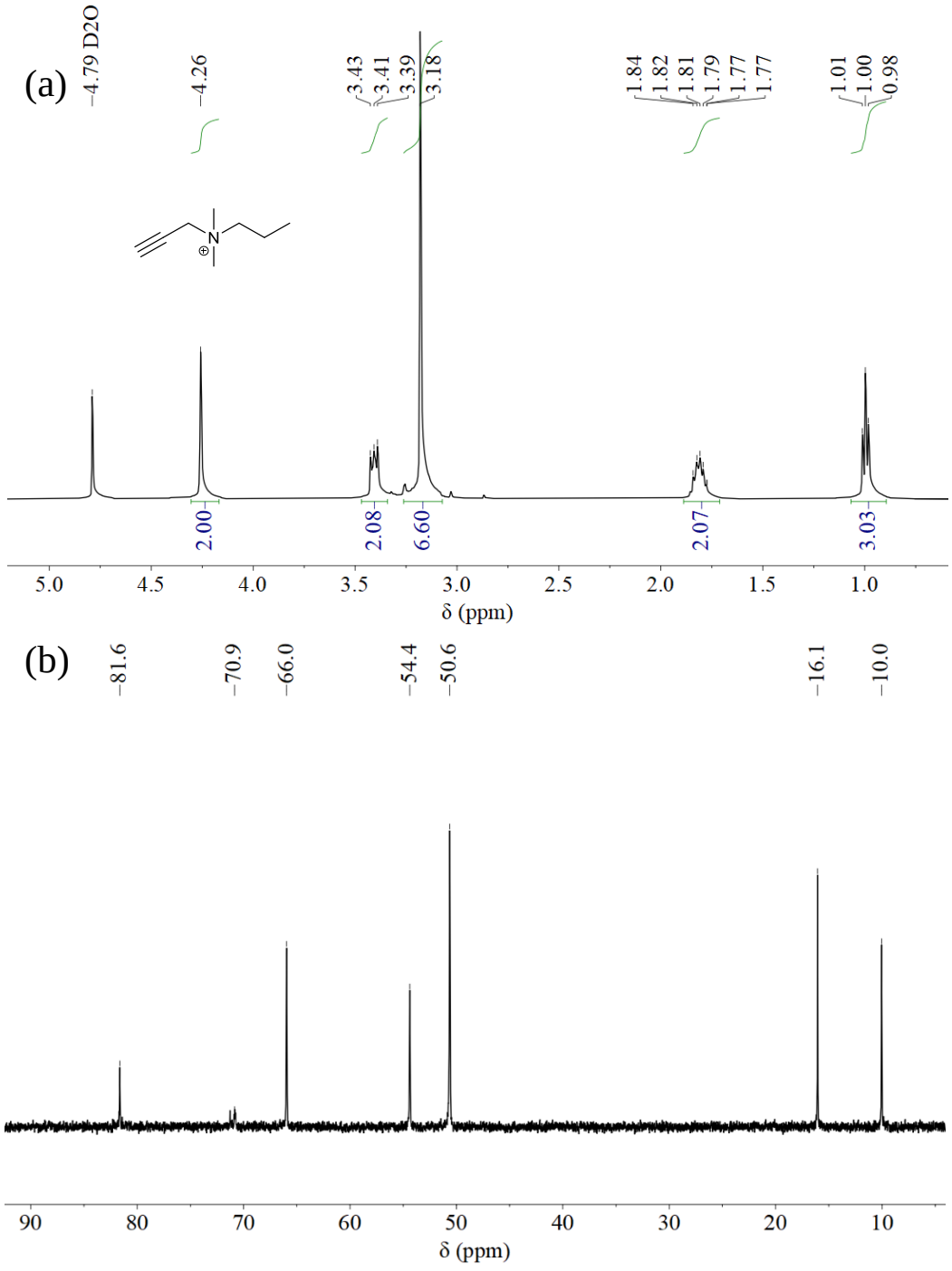

Fig. S5. ^{1}H (a) and ^{13}C (b) NMR spectra of PrDA in D_2O.

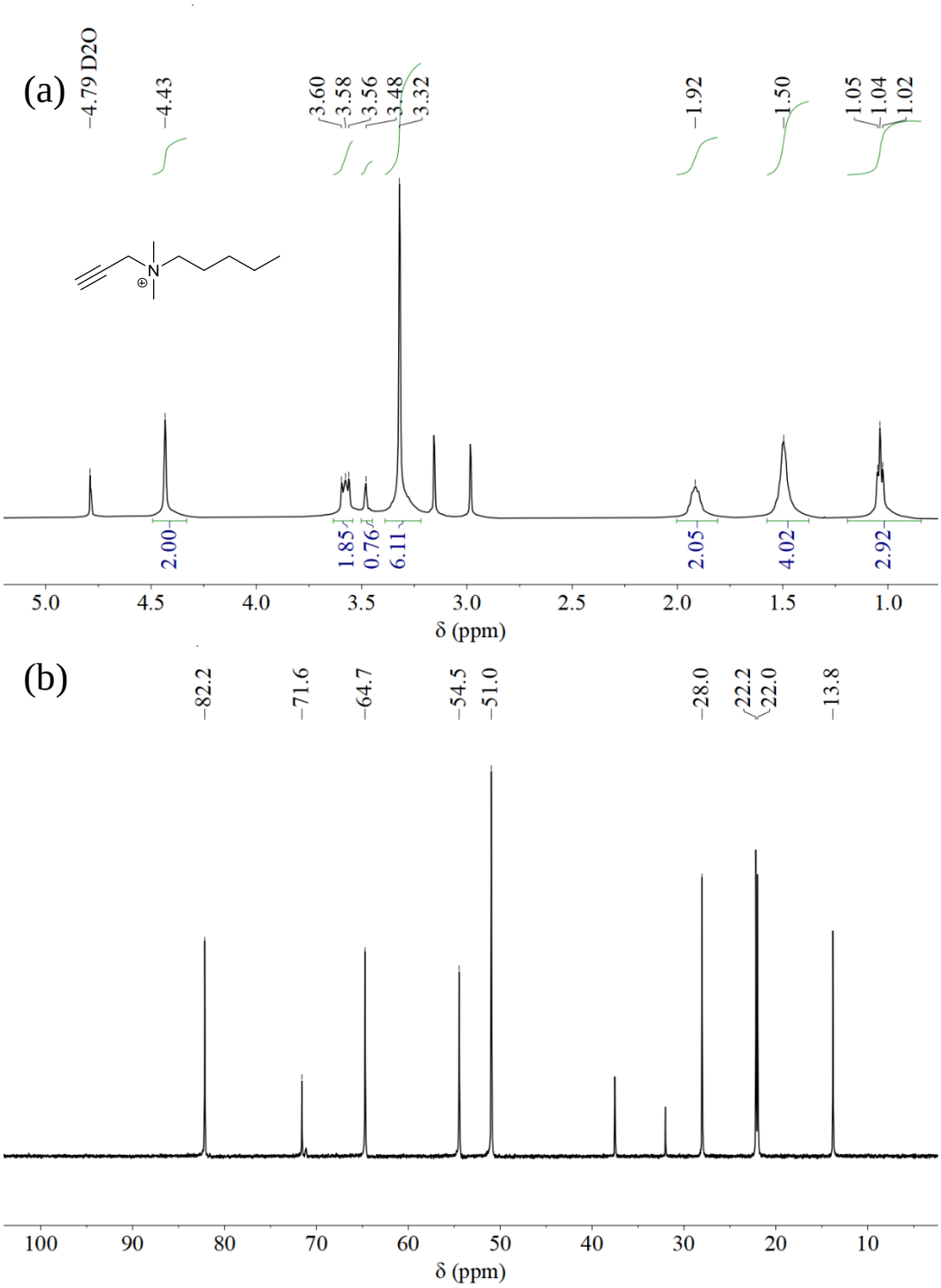

Fig. S6. ^{1}H (a) and ^{13}C (b) NMR spectra of PeDA in D_2O.

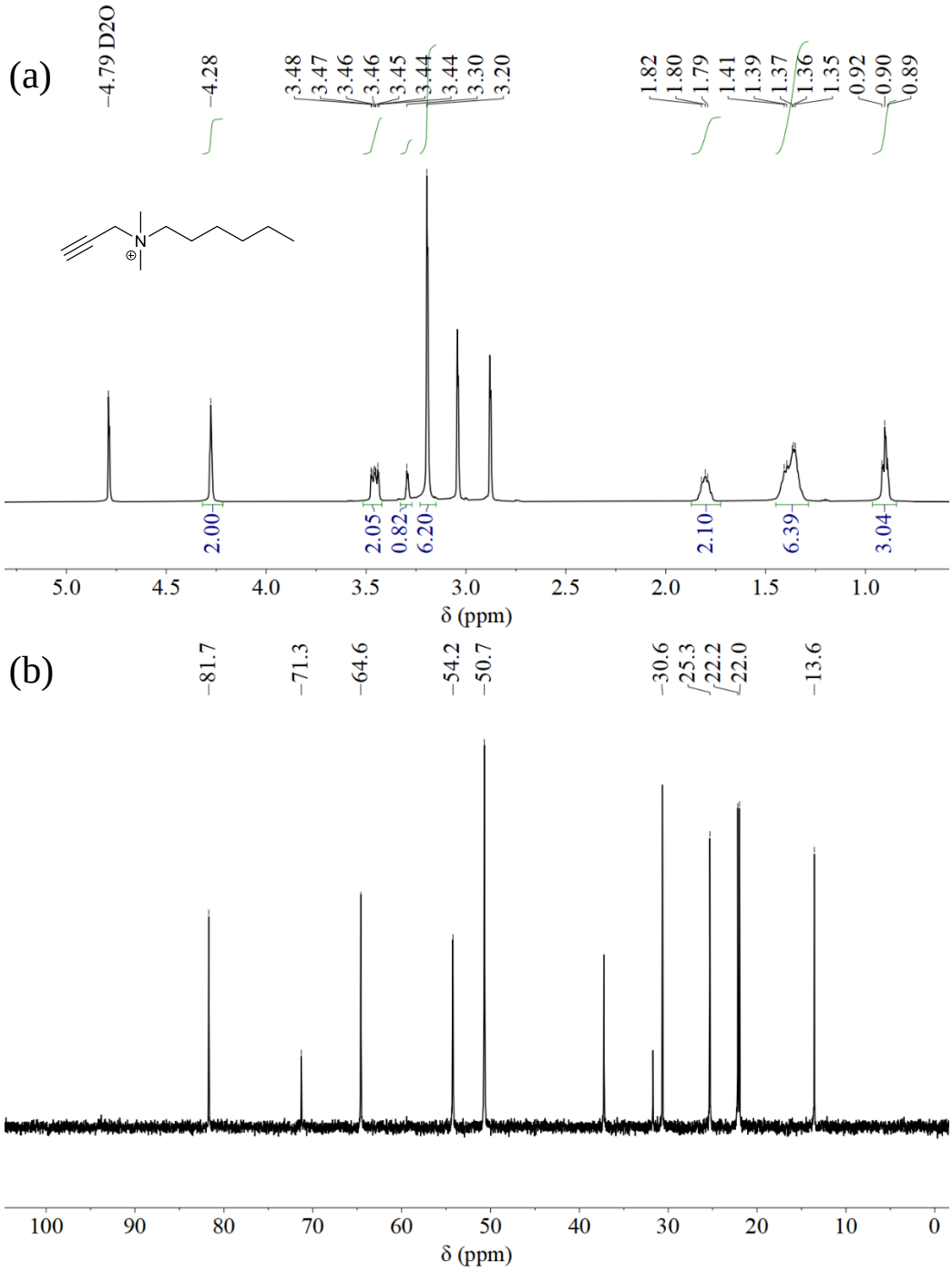

Fig. S7. ^{1}H (a) and ^{13}C (b) NMR spectra of HexDA in D_2O.

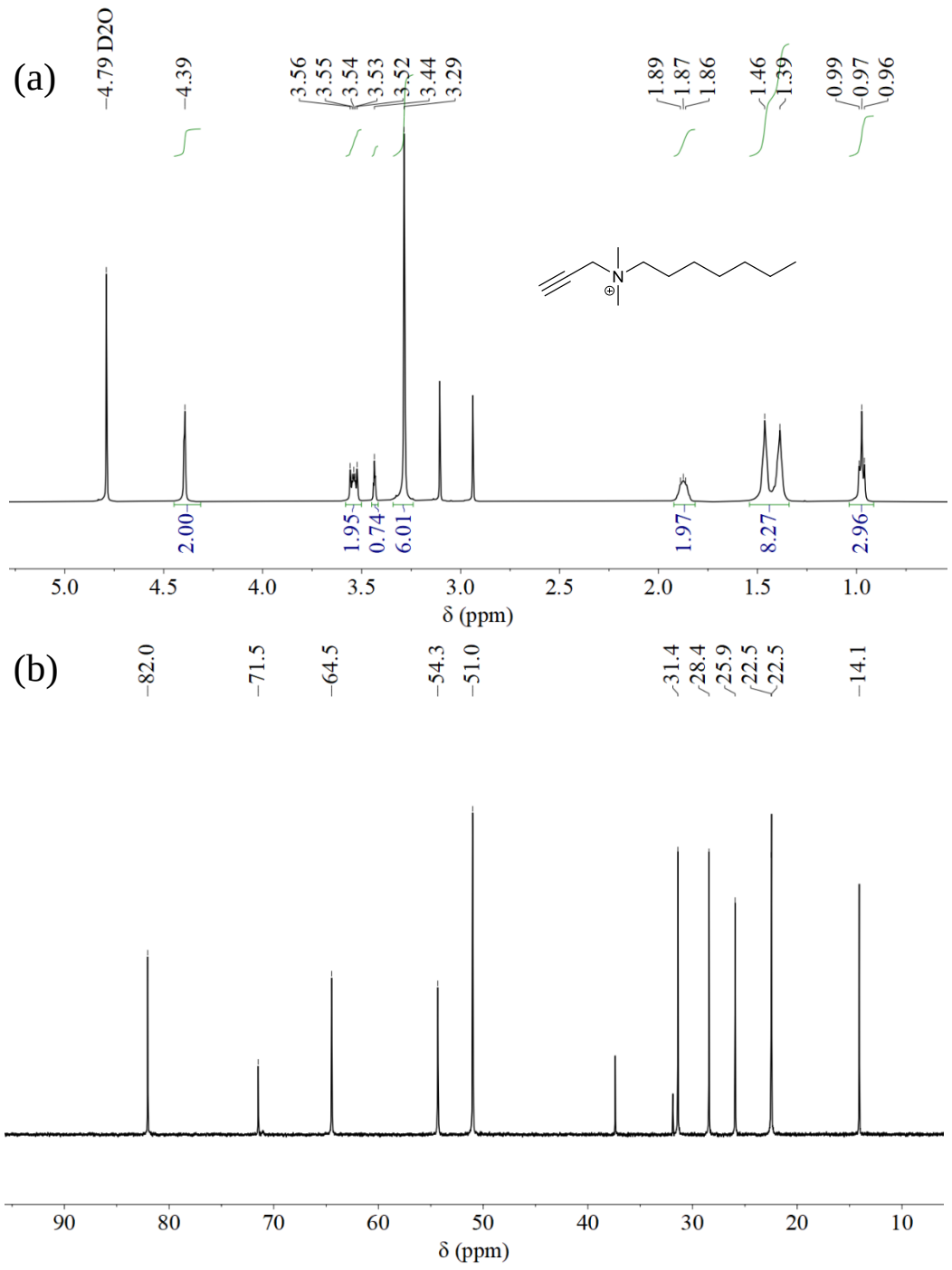

Fig. S8. ^{1}H (a) and ^{13}C (b) NMR spectra of HeptDA in D_2O.

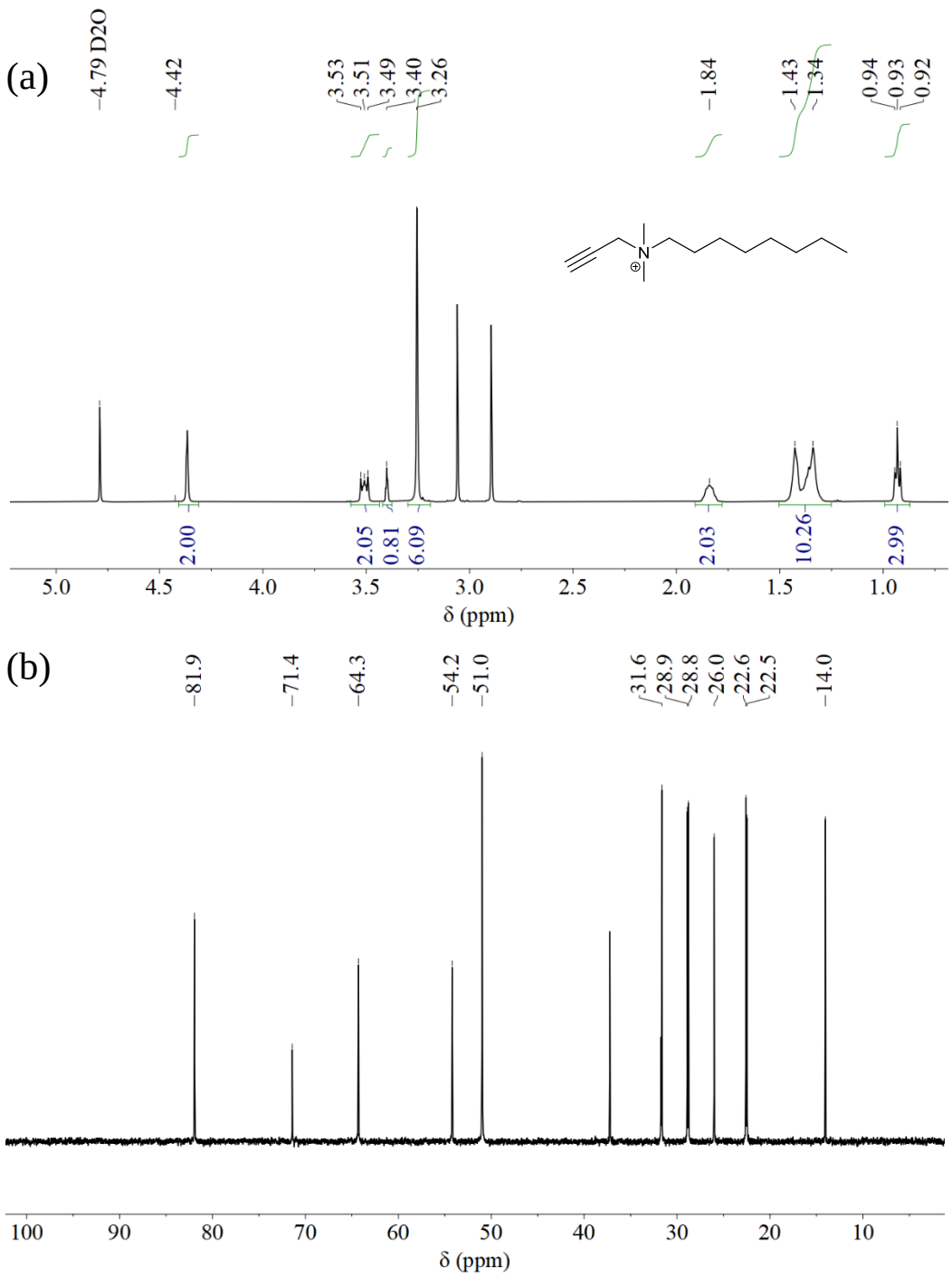

Fig. S9. ^{1}H (a) and ^{13}C (b) NMR spectra of OctDA in D_2O.

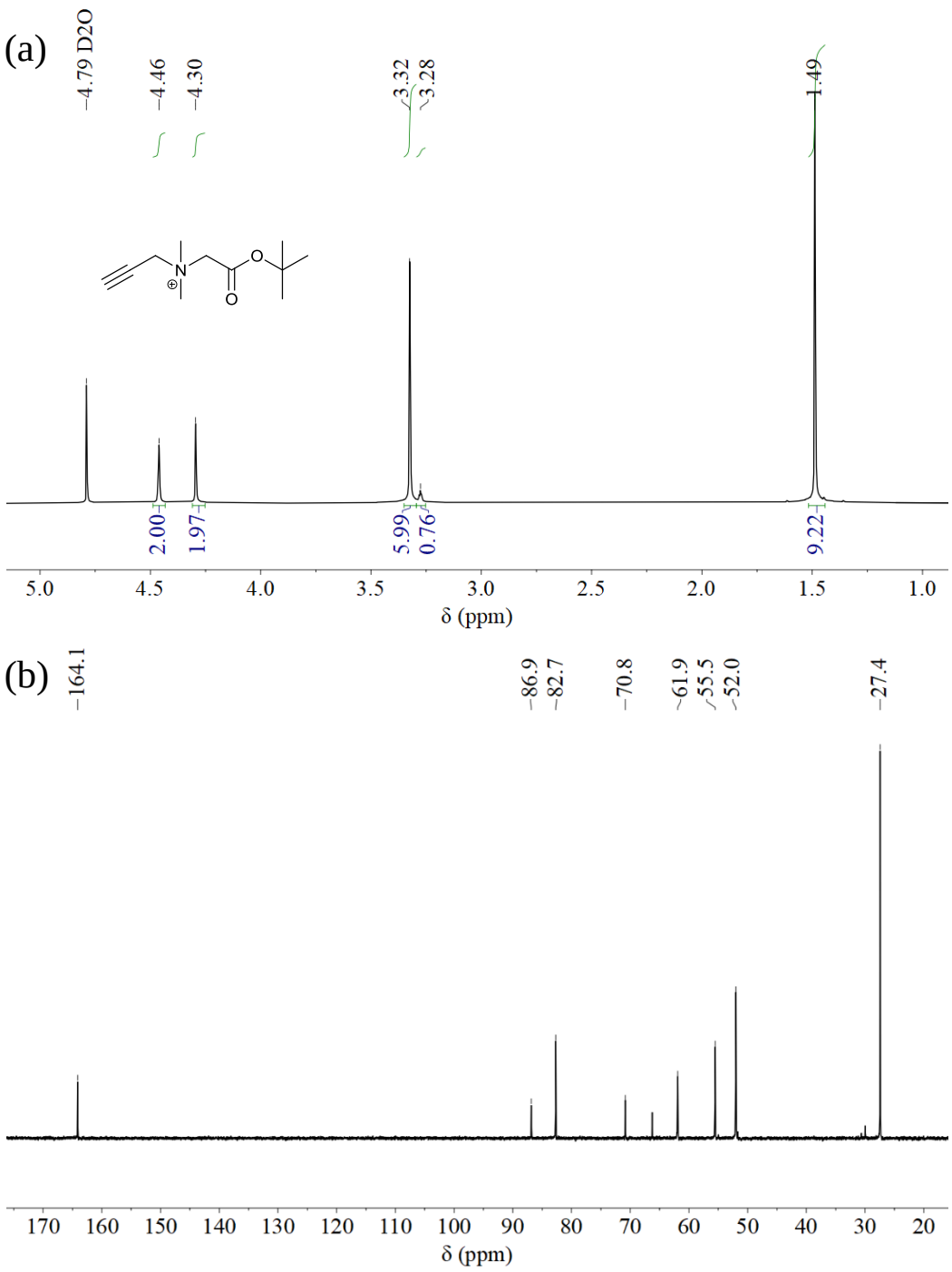

Fig. S10. ^{1}H (a) and ^{13}C (b) NMR spectra of tBuCMDA in D_2O.

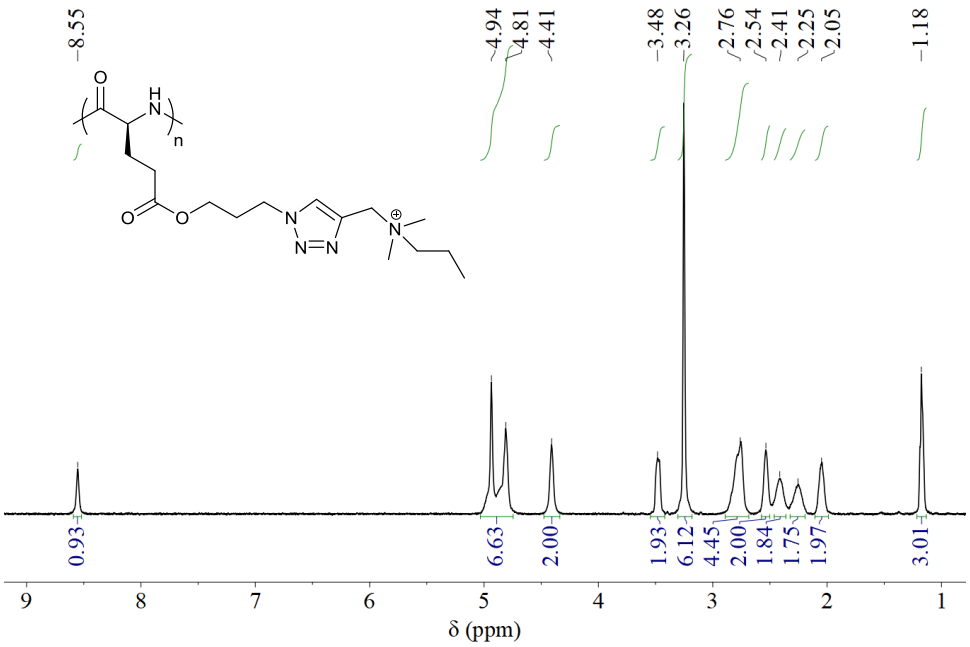

Fig. S11. ^{1}H NMR spectrum of **P1**-Pr in TFA-*d*.

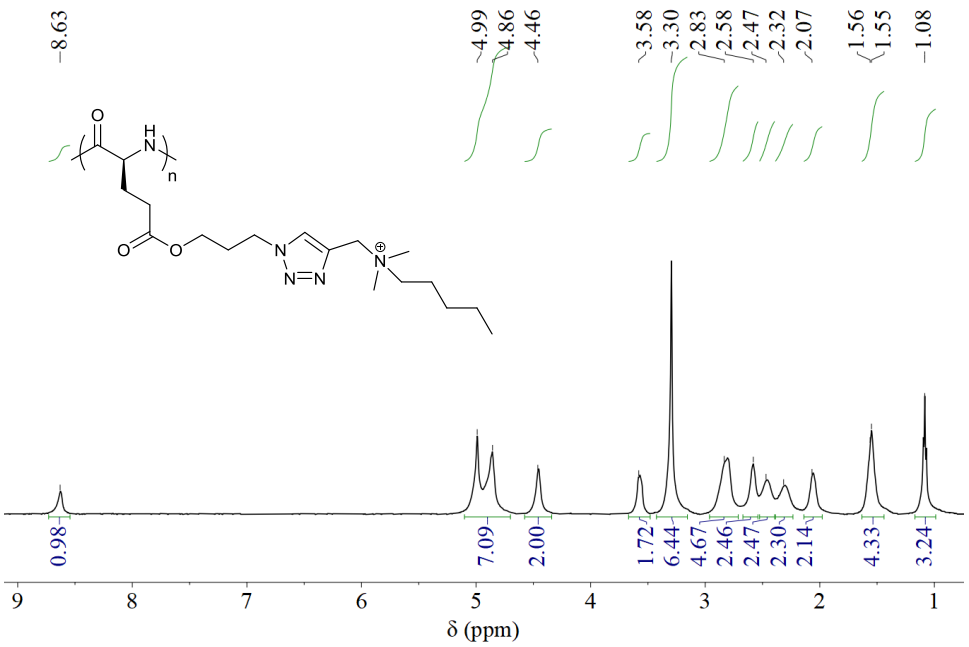

Fig. S12. ^{1}H NMR spectrum of **P1**-Pe in TFA-*d*.

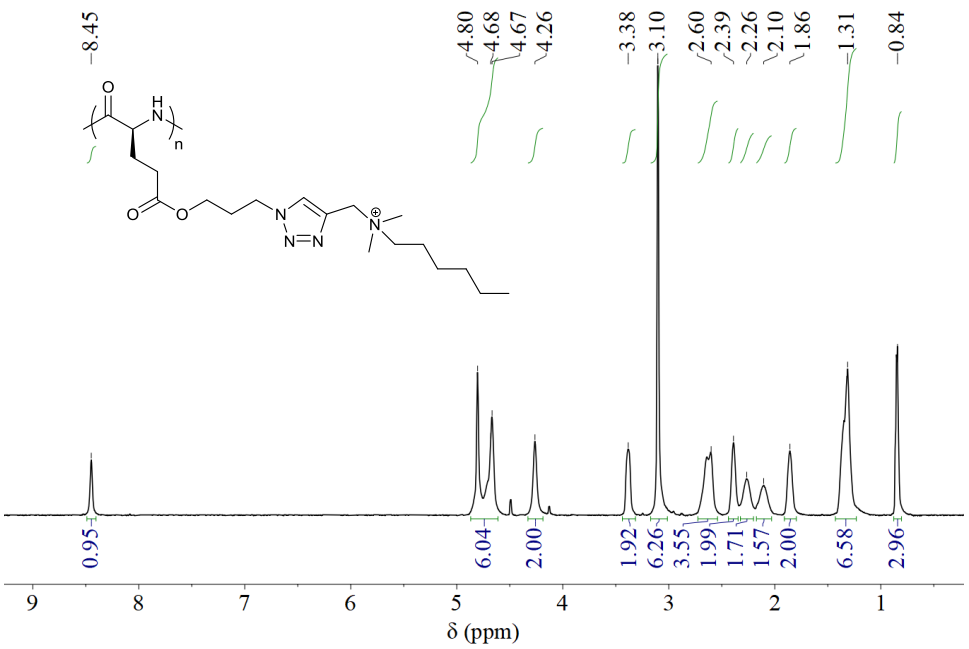

Fig. S13. ^{1}H NMR spectrum of **P1**-Hex in TFA-*d*.

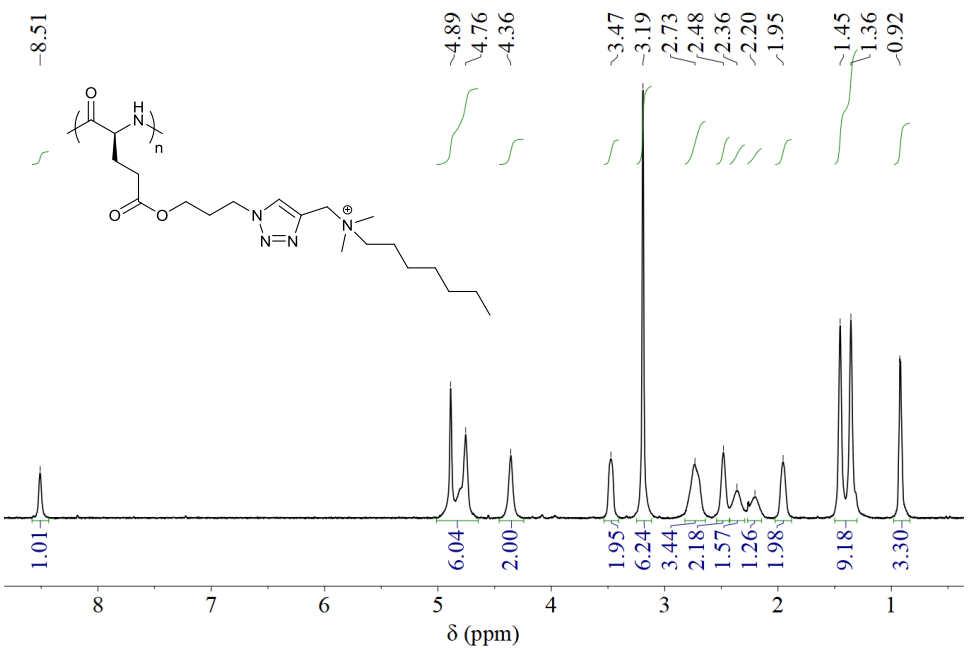

Fig. S14. ^{1}H NMR spectrum of **P1**-Hept in TFA-*d*.

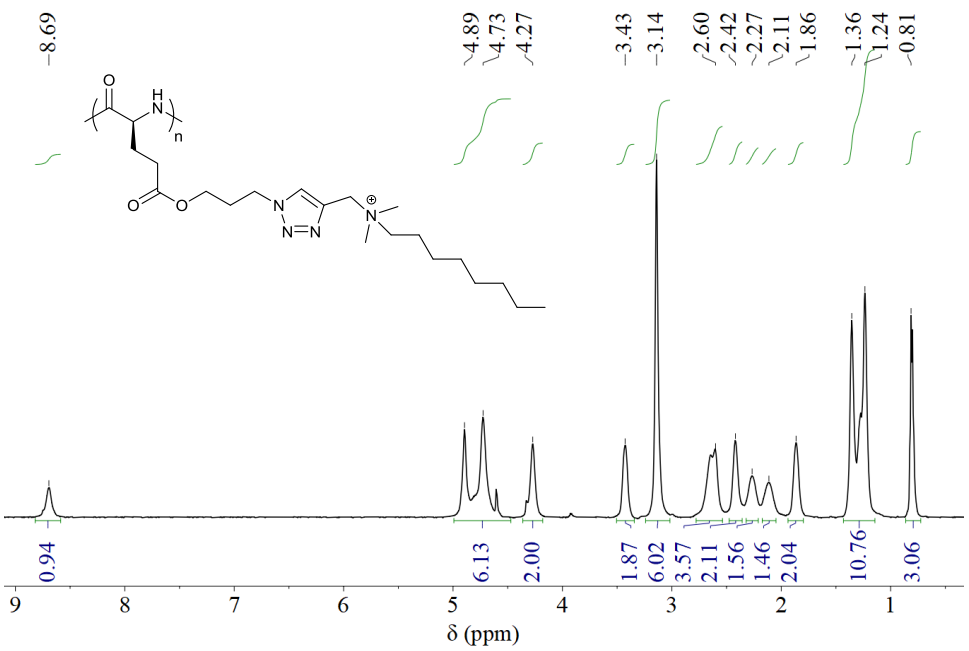

Fig. S15. ^{1}H NMR spectrum of **P1**-Oct in TFA-*d*.

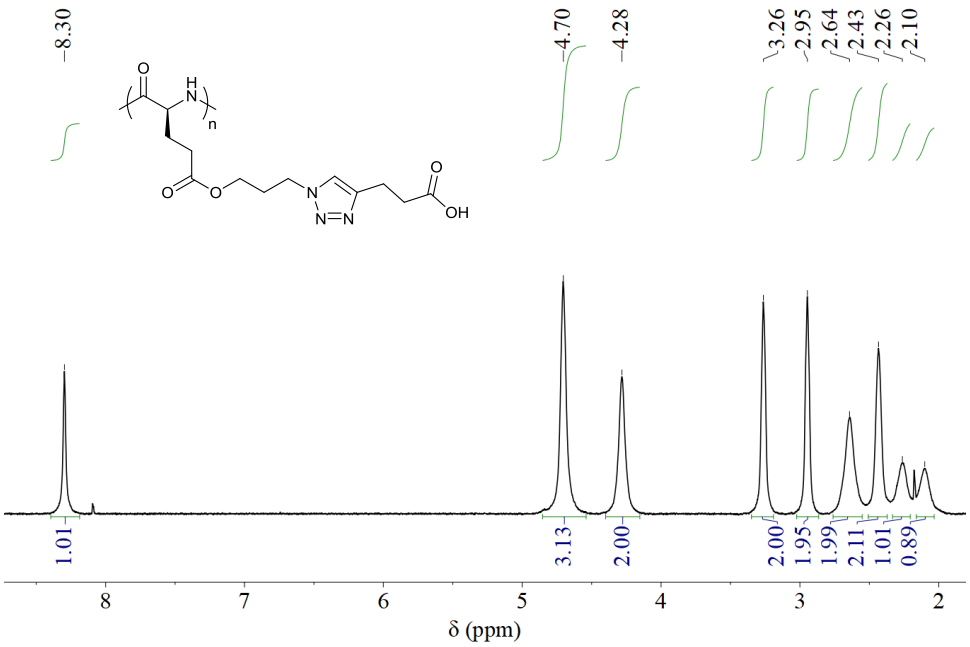

Fig. S16. ^{1}H NMR spectrum of **P2**-CA in TFA-*d*.

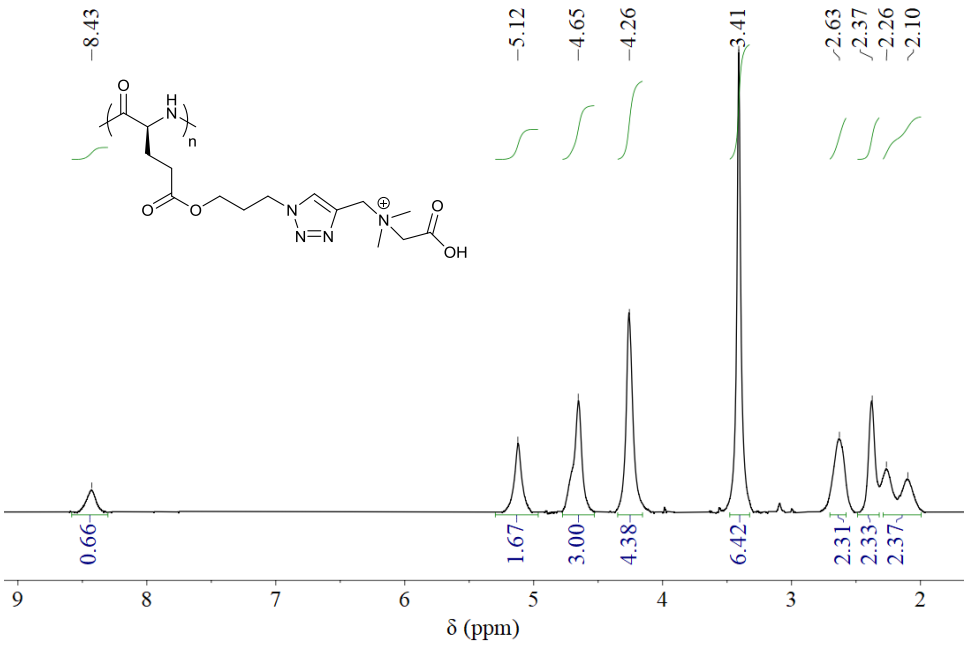

Fig. S17. ^{1}H NMR spectrum of **P2**-ZW in TFA-*d*.

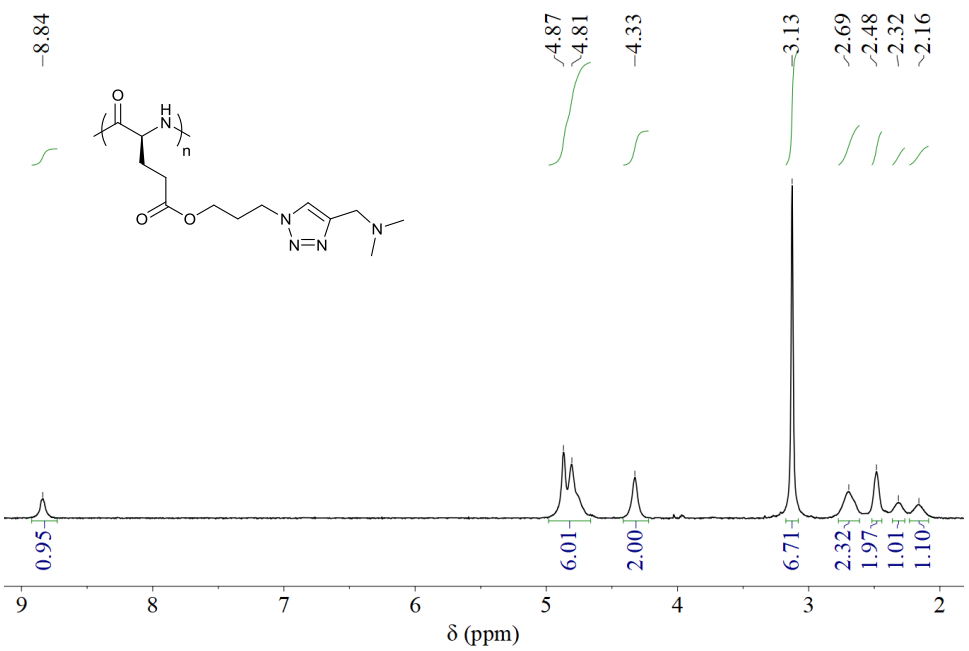

Fig. S18. ^{1}H NMR spectrum of **P3** in TFA-*d*.

Supporting References

1. H. Tang, L. Yin, K. H. Kim and J. Cheng, *Chem. Sci.*, 2013, **4**, 3839-3844.
2. A. J. Adler, N. J. Greenfield and G. D. Fasman, in *Part D: Enzyme Structure*, Academic Press, 1973, vol. 27, pp. 675-735.
3. N. J. Greenfield, *Nat. Protoc.*, 2006, **1**, 2876-2890.
4. J. A. Morrow, M. L. Segall, S. Lund-Katz, M. C. Phillips, M. Knapp, B. Rupp and K. H. Weisgraber, *Biochemistry*, 2000, **39**, 11657-11666.