Supplementary Information

Synthesis and Application of the Blue Fluorescent Amino Acid L-4-Cyanotryptophan to Assess Peptide-Membrane Interactions†

Kui Zhang[a, ‡], Ismail A. Ahmed[b,‡], Huong T. Kratochvil[a] , William F. DeGrado[a], Feng Gai[c,*], Hyunil Jo[a,*]

[a]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158, United States; [b]Department of Biochemistry and Biophysics and [c]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

[‡]These authors contributed equally

*Corresponding authors: hyunil.jo@ucsf.edu (H.J.) and gai@sas.upenn.edu (F.G.); Tel.: +1-215-913-7838 (H.J.); +1-215-573-6256 (F.G.)

Table of Contents

Experimental Section

General. All reagents were used as received from commercial sources. Reactions were monitored through thin layer chromatography (TLC) on 0.25-mm SiliCycle silica gel plates and visualized under UV light. Flash column chromatography (FCC) was performed using Combiflash®Rf + Lumen UV-VIS/ELSD. NMR spectra were recorded using Bruker Avance-300 calibrated to CD_3OD (3.31 and 49.0 ppm for ^1H and ^{13}C NMR spectra, respectively), $(CD_3)_2CO$ (2.05 ppm for ^1H and 29.8, 206.3 ppm ^{13}C NMR spectra, respectively), D_2O (4.79 ppm for ^1H) as the internal reference. ^1H NMR spectral data are reported in terms of chemical shift (δ, ppm), multiplicity, coupling constant (Hz), and integration. ^{13}C NMR spectral data are reported in terms of chemical shift (δ, ppm). The following abbreviations indicate the multiplicities: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Optical rotation was measured on a Jasco P2000 polarimeter and CD experiment was performed on a Jasco J-810 spectropolarimeter.

Synthesis of N-acetyl amino ester 5. To a solution of amino ester **4**[1] (10.3 g, 40.0 mmol) in dry DCM (300 mL) under inert atmosphere at 0 °C were added dropwise and successively Et_3N (11.1 mL, 3.0 equiv) and acetyl chloride (3.2 mL, 1.1 equiv). The reaction mixture was stirred at 0 °C for 1 h and then allowed to stir for overnight at rt. The solution was quenched with a solution of sat. aq. $NaHCO_3$ and the resulting mixture was extracted with ethyl acetate. The combined organic layers were washed with brine, dried over Na_2SO_4, and concentrated under reduced pressure. The residue was purified by the column chromatography using CombiFlash®Rf + Lumen UV-VIS/ELSD (Teledyne ISCO) and a RediSep column [silica 120 g (60-Å pore size, 20–40 μm)] to provide the N-acetyl amino ester **5** (11.0 g, 92%). Ethyl 2-acetamido-3-(4-cyano-1H-indol-3-yl)propanoate **5**: Pale yellow solid; R_f = 0.52 (hexane/EtOAc, 3:1); ^1H NMR (300 MHz, CD_3OD) δ 7.68–7.65 (m, 1H), 7.46–7.43 (m, 1H), 7.32 (s, 1H), 7.24–7.19 (m, 1H), 4.81 (dd, J = 6.7, 8.3 Hz, 1H), 4.10 (q, J = 7.1 Hz, 2H), 3.55 (dd, J = 6.4, 14.7 Hz, 1H), 3.36 (dd, J = 8.2, 14.7 Hz, 1H), 1.95 (s, 3H), 1.14 (t, J = 7.1 Hz, 3H); ^{13}C NMR (75 MHz, CD_3OD) δ 172.1, 171.8, 136.9, 127.0, 126.4, 125.5, 120.7, 119.2, 116.4, 109.4, 100.6, 60.8, 53.7, 26.7, 20.9, 12.9.

Synthesis of N-acetyl amino acid 6: A mixture of the ester **5** (4.8 g, 16 mmol) in ethanol (50 mL) and LiOH (3.0 equiv, 1 M solution) stirred at rt overnight. The ethanol was evaporated in vacuo. The residue was diluted with water (50 mL), acidified to pH 2 using aqueous HCl (1N), and extracted with ethyl acetate. The

combined organic phases were washed with $NaHCO_3$, brine and concentrated under reduced pressure to yield the acid (3.86g, 89%). 2-acetamido-3-(4-cyano-1H-indol-3-yl)propanoic acid **6:** Pale yellow solid; R_f = 0.30 (DCM/Methanol, 10:1); ^{1}H NMR (300 MHz, CD_3OD) δ 7.68–7.65 (m, 1H), 7.46–7.38 (m, 1H), 7.38 (s, 1H), 7.23–7.18 (m, 1H), 4.81 (dd, J = 5.1, 8.3 Hz, 1H), 3.63 (dd, J = 5.1, 15.1 Hz, 1H), 3.36 (dd, J = 8.4, 15.1 Hz, 1H), 1.93 (s, 3H); ^{13}C NMR (75 MHz, CD_3OD) δ 173.7, 171.8, 136.9, 126.9, 126.4, 125.5, 120.7, 119.2, 116.4, 109.7, 100.5, 53.2, 26.8, 21.0.

Synthesis of L-4CN-Trp 7 and Fmoc-L-4CN-Trp 8: Compound **6** (0.51 g, 2.0 mmol) was dissolved in PBS buffer solution (0.1 M, 25 mL) containing $CoCl_2 \bullet H_2O$ (0.125 mM) at pH 8.0. Amano acylase (0.5 g, >30,000 U/g) was added with stirring. The mixture was placed in a shaker at 37 ºC for 48 h and then quenched by adjusting the pH to 5 with 1M HCl solution. After centrifugation, the precipitate was isolated and lyophilized to obtain crude L-4CN-Trp **7**, which was subsequently used in the next step without purification. A small amount of sample was isolated by RP-HPLC for analytical analysis, whose spectroscopic data is in good agreement with literature values.[2] (S)-2-amino-3-(4-cyano-1H-indol-3-yl)propanoic acid **7,** white solid; ^{1}H NMR (300 MHz, D_2O + 100 mM DCl) δ 7.42 (d, J = 8.01 Hz, 1H), 7.18 (d, J = 7.2 Hz, 1H), 7.11 (s, 1H), 6.94 (t, J = 7.89 Hz, 1H), 4.06 (dd, J = 5.6, 9.6 Hz, 1H), 3.31 (dd, J = 5.6, 15.3 Hz, 1H), 2.97 (dd, J = 9.6, 15.3 Hz, 1H); ^{13}C NMR (75 MHz, D_2O + 100 mM DCl) δ 170.9, 136.2, 128.5, 126.3, 125.2, 121.3, 119.8, 117.4, 105.7, 98.9, 53.4, 25.4; HRMS (ESI) m/z 230.0909 [$(M^+H)^+$; calcd for $C_{12}H_{13}N_3O_2^+$ $(M^+H)^+$: 230.0930]; $[\alpha]_D^{25}$=-209.64 (c= 2.5 mg/mL, in 1 M HCl).

The aforementioned crude **7** was dissolved in an aqueous Na_2CO_3 solution (0.1 M, 40 mL) followed by the addition of 9-fluorenylmethyl succinimidyl carbonate (0.67 g, 2.0 mmol) in THF (20 mL). The mixture was stirred for 2 h at rt. THF was removed under vacuo and the crude mixture was washed with Et_2O. The pH of the aqueous layer was adjusted to 2 using 3 M aqueous HCl and was extracted with DCM. The organic extracts were combined and washed with brine, dried over Na_2SO_4 and concentrated *in vacuo*. The residue was purified by the column chromatography to yield the product (280 mg, 62% of the theoretical yield over two steps). (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-cyano-1H-indol-3-yl)propanoic acid **8** Pale yellow solid; R_f = 0.20 (DCM/methanol, 10:1); ^{1}H NMR (300 MHz, $(CD_3)_2CO$) δ 10.74 (s, 1H), 7.86–7.84 (m, 2H), 7.78–7.75 (m, 1H), 7.67–7.64 (m, 2H), 7.55–7.50 (m, 2H), 7.43–7.38 (m, 2H), 7.32–7.24 (m, 3H), 6.88 (d, J = 8.0 Hz, 1H), 4.81–

4.73 (m, 1H), 4.28–4.18 (m, 3H), 3.77–3.71 (m, 1H), 3.50–3.42 (m, 1H); ^{13}C NMR (75 MHz, $(CD_3)_2CO$) δ 172.8, 156.1, 144.1, 141.1, 136.9, 127.6, 127.4, 127.0, 126.7, 125.7, 125.3 (2C), 121.1, 119.9, 119.3, 116.7, 110.5, 101.4, 66.3, 54.6, 47.0, 27.0. HRMS (ESI) m/z 452.1605 [$(M^+H)^+$; calcd for $C_{27}H_{22}N_3O_4^+$ $(M^+H)^+$: 452.1610].

Determination of optical purity: Marfey's reagent (1-fluoro-2-4-dinitrophenyl-5-L-alanine amide) was used as a solution in acetone (33 mM). In a 2 mL vial, the amino acid (racemic amino acid that was obtained by the hydrolysis (1M LiOH in THF-H_2O) of compound **4** or crude L-amino acid from the above resolution (0.5 μmol) was dissolved in 1 M aqueous $NaHCO_3$ (140 μL). Marfey's reagent (60 μL, 2 μmol) was added, then the vial was placed in incubator (rpm 900) at 37 ºC. After 1 h, the reaction mixture was allowed to cool to rt and then diluted with 1:1 water /acetonitrile (550 μL) and analyzed by HPLC [solvent A (0.1% TFA in water) and B (0.1% TFA, 1% water in acetonitrile), Vydac® 214TP, 5μm, C4, 300Å, 4.6 mm i.d. x 250 mm), wavelength 330 nm with a gradient of 5–95% B over 35 min].

Solid-phase peptide synthesis of L-4CN-Trp labeled pHLIP: The sequence synthesized: NH_2-GGEQNPIYW$_{CN}$ARYADWLFTTPLLLLDLALLVDADEGT-CO_2H, where W_{CN} stands for L-4CN-Trp. The peptide was synthesized on a 0.1 mmol scale on the preloaded Thr(OtBu)-HMPB-ChemMatrix resin (0.5 meq/g) using a Biotage Initiator + Alstra peptide synthesizer. The deprotection was carried out for 5 min at 70 ºC with 4.5 mL 20% 4-methylpiperidine in DMF. A standard coupling step (for all amino acids except for Fmoc-protected 4-CN-Trp) was done for 5 min at 75 ºC with 5 equivalents of Fmoc-protected amino acids, 4.98 equivalents of HCTU, and 10 equivalents of DIPEA (relative to the amino groups on resin) in DMF. For L-Fmoc-protected 4-CN-Trp, a coupling reaction was done for 5 min at 75 ºC with 1.5 equivalents of Fmoc-protected amino acids, 1.49 equivalents of HCTU, and 3 equivalents of DIPEA (relative to the amino groups on resin) in DMF. Peptide cleavage was carried out in the presence of TFA//TIS/H_2O (95:2.5:2.5, v/v) for 2 h at rt. The crude peptide was obtained after precipitation in cold diethyl ether and purified by RP-HPLC. (Vydac C4 214TP1022) using solvent A (0.1% TFA in water) and B (0.1% TFA, 1% water in acetonitrile). After 5 min equilibration with 5% B at a flow rate of 5 mL/min, a gradient of 5–70% B in 35 min was used. The chemical entity and purity of synthesized peptides were verified by a Shimazu AXIMA MALDI-TOF mass spectrometer and an HP 1100 analytical HPLC system, respectively. MS (MALDI-TOF): m/z 4077.73 (calcd $[M^+H]^+$ = 4077.51).

Solid-phase peptide synthesis of Ac-Gly-L-4CN-Trp-Gly: The peptide was synthesized on a 0.05 mmol scale using a Biotage Initiator+ Alstra peptide synthesizer. A typical solid-phase peptide synthesis reaction cycle includes Fmoc deprotection, washing, coupling, and post-coupling washing steps. The deprotection step was carried out for 5 min at 70 ºC with 4.5 mL 20% 4-methylpiperidine in DMF. A standard coupling step (for all amino acids except for Fmoc-protected L-4CN-Trp) was done for 5 min at 75 ºC with 5 equivalents of Fmoc-protected amino acids, 4.98 equivalents of HCTU, and 10 equivalents of DIPEA (relative to the amino groups on resin) in DMF. For Fmoc-protected L-4CN-Trp, a coupling reaction was done for 5 min at 75 ºC with 1.5 equivalents of Fmoc-protected amino acids, 1.49 equivalents of HCTU, and 3 equivalents of DIPEA (relative to the amino groups on resin) in DMF. Acetylation was done by using 5 equivalents of Ac_2O and 10 equivalents of DIPEA. Peptide cleavage was carried out in the presence of TFA/TIS/H_2O (95:2.5:2.5, v/v) for 2 h at rt. The crude peptide was obtained after precipitation in cold diethyl ether. The chemical entity and purity of synthesized peptides were verified by a Shimazu mass spectrometer and an HP 1100 analytical HPLC system, respectively. MS (ESI): m/z 385.3 (calcd $[M^+H]^+$ = 385.4).

Cell culture: HeLa cells were grown overnight at 37 ºC on slides to about 60% confluency in DMEM medium containing 10% FBS and 4 mM L-Glutamine. At that point the supernatant was aspirated and 0.5 mL of a stock 4CN-Trp-pHLIP solution (in PBS buffer at the desired pH) was added to each slide with a final peptide concentration of 10 μM. Then the peptide treated cells were incubated at rt for 1 h, followed by washing with the corresponding PBS buffer 3× and fixed for 30 minutes with 2% formaldehyde in the corresponding PBS. Subsequently, each slide was again washed 3× with the corresponding PBS buffer and air-dried. Finally, two drops of 50% glycerol were then added to each slide, which was covered with a cover slip and the edges were sealed with clear nail polish.

Cell imaging: Cell images were acquired on a commercial widefield fluorescence microscope (Leica DM6000) equipped with a 20X dry objective using a standard DAPI filter (excitation bandwidth: 325 – 375 nm, emission bandwidth: 435 nm – 485 nm). The integration time for each image was 50 ms and data/image processing was carried out using ImageJ 1.5 software.[2]

Preparation of large unilamellar vesicles (LUVs): The 100-nm LUVs used in the peptide-membrane binding experiments were composed of 99% POPC (purchased from Avanti Polar Lipids Inc., AL) and 1% of

3,3'-dioctadecyloxacarbocyanine perchlorate (DiO, purchased from Thermo Fisher Scientific), which were prepared following previously published procedures.[3] Briefly, an appropriate amount of DiO was added to a stock lipid solution in chloroform and the resultant mixture was put to shake at rt for 30 minutes. This lipid solution was then allowed to dry under a flow of nitrogen to form a lipid film, which was followed by a 30-minute lyophilization to remove any remaining solvent. The resultant lipid film was then rehydrated in the desired buffers. This sample was then subjected to 7 rounds of slow vortexing, freezing and thawing. The resulting vesicle solution was then extruded 11 times through an extruder (Avanti Polar Lipids Inc., AL) equipped with a 100 nm membrane. After extrusion, the LUV solution was diluted to 1.0 mM (lipid concentration) with desired buffers.

Fluorescence measurements: All fluorescence spectra were collected on a Jobin Yvon Horiba Fluorolog 3.10 spectrofluorometer using a 1 cm quartz cuvette with a 1.0 nm resolution, 1 nm excitation/emission slit, an integration time of 1.0 nm/s, and an excitation wavelength of 320 nm at 20°C. Fit for fluorescence binding curve is detailed in the SI.

Förster distance calculation: The Förster distance (R_0) of the FRET pair, L-4CN-Trp (donor) and DiO (acceptor), is calculated using the following equation:[4]

$$R_0^6 = \left(\frac{9000(ln10)\kappa^2 Q_D}{128\pi^5 N\eta^4}\right)J(\lambda), \quad (1)$$

where κ^2 is the orientation factor, assumed to be 2/3, Q_D is the fluorescence quantum yield of the donor in the absence of the acceptor which is 0.85 for L-4CN-Trp, N is the Avogadro's number, η is the refractive index of the medium (1.33 for MeOH), and $J(\lambda)$ is the overlap integral, determined by the following equation: [1]

$$J(\lambda) = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda, \quad (2)$$

where $F_D(\lambda)$ is the area-normalized emission spectrum of the donor and $\varepsilon_A(\lambda)$ is the wavelength-dependent molar absorption coefficient of the acceptor.

Data fitting: The fluorescence binding curve (i.e., Figure 3), obtained from the intensity at 512 nm of the fluorescence spectrum of every peptide-membrane mixture (i.e., Figure S3), was fit to the following membrane binding model:[3]

$$P + \mathrm{n}L \rightleftharpoons PL_{\mathrm{n}}\,.$$

where P and L represent peptide and lipid, respectively. Following Engelman and coworkers,[5] we assumed that n = 50 (i.e., every membrane-bound peptide is solvated by 50 lipids). For a given set of initial peptide ($[P_0]$ and lipid ($[L_0]$) concentrations, one can easily show that the equilibrium concentration of PL_n ($[PL_n]_{eq}$ is:

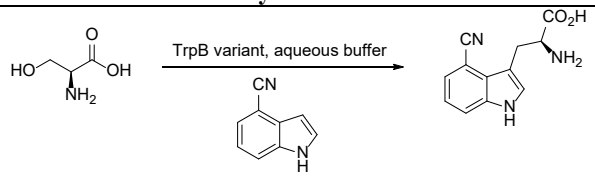
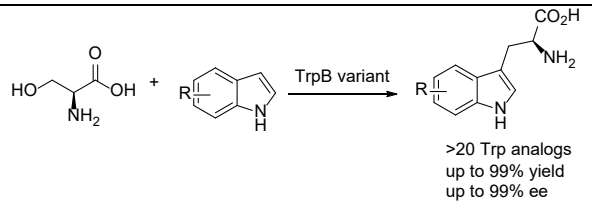
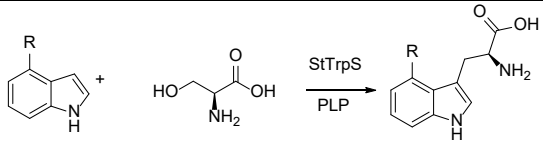
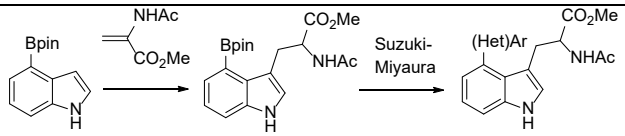
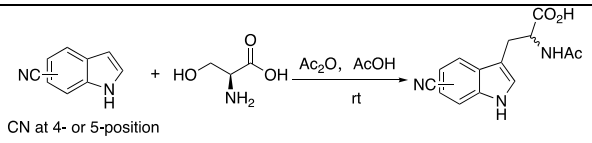
$$[\mathrm{PL_n}]_{\mathrm{eq}} = \frac{(K_{\mathrm{d}}+[\mathrm{P_0}]+[\mathrm{L_0}]^*)-\sqrt{(K_{\mathrm{d}}+[\mathrm{P_0}]+[\mathrm{L_0}]^*)^2-4[\mathrm{P_0}][\mathrm{L_0}]^*}}{2}, \qquad (3)$$

where K_d is the dissociation equilibrium constant and $[L_0]^* = [L_0]/50$. Because only the membrane-bound peptides can contribute to the observed FRET signal, the binding curve can be described by the following equation:

$$I = I_{\mathrm{m}} \times [\mathrm{PL_n}]_{\mathrm{eq}}. \qquad (4)$$

In the fitting, K_D and I_m were treated as variables.

Table S1. Summary of reported syntheses of 4CN-Trp and its derivatives.

Reference	Key reaction	Advantages
J. Org. Chem., 2018, **83,** 7447.		Green chemistry, use of readily available starting material, high enantiomeric purity
J. Am. Chem. Soc., 2017, **139,** 10769.		Green chemistry, use of readily available starting material, synthesis of Trp analogues with various substitution
Angew. Chem., Int. Ed., 2018, **57,** 6830.		Green chemistry, use of readily available starting material, high enantiomeric purity
Org. Biomol. Chem., 2016, **14,** 10095.		Gram-scale and rapid synthesis
Phys. Chem. Chem. Phys., 2018, **20,** 19906.		Facile method to synthesis a variety of Trp derivatives

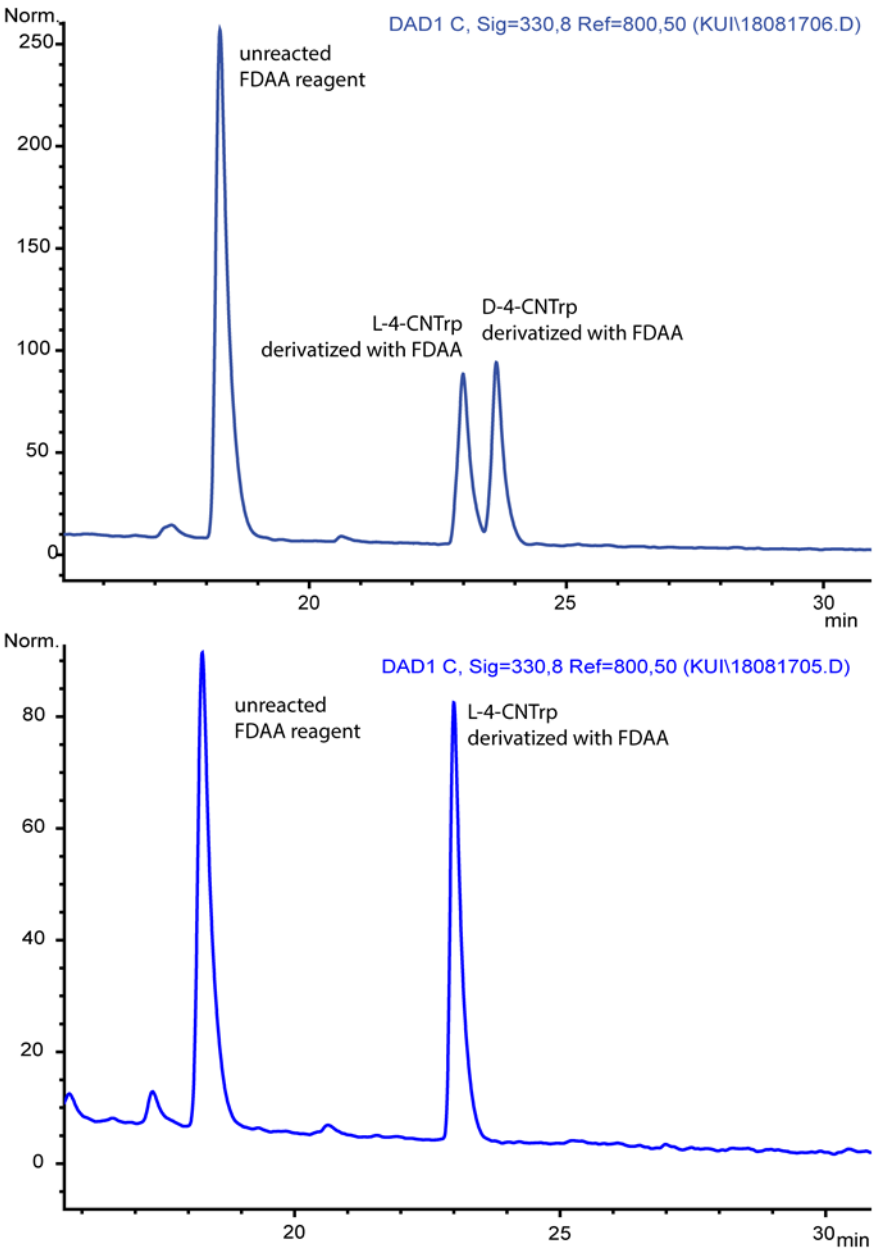


Figure S1. HPLC profile of FDAA derivatized racemic 4CN-Trp (Top) and L-4CN-Trp (Bottom).

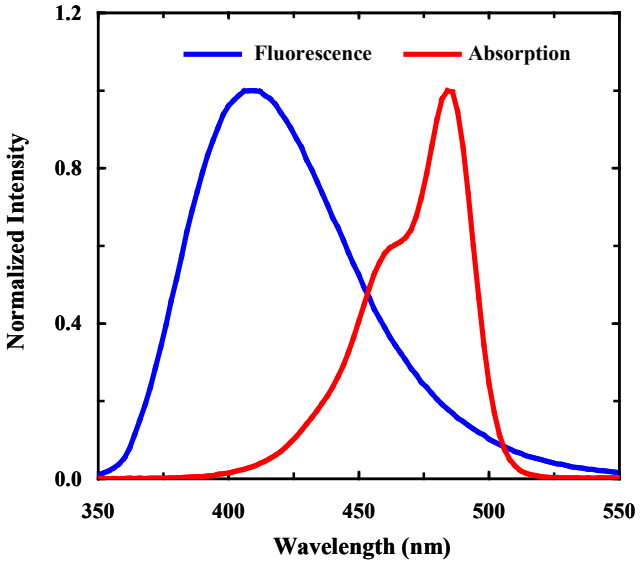


Figure S2. Normalized fluorescence spectrum of L-4CN-Trp (blue) and absorption spectrum of DiO (red) in methanol.

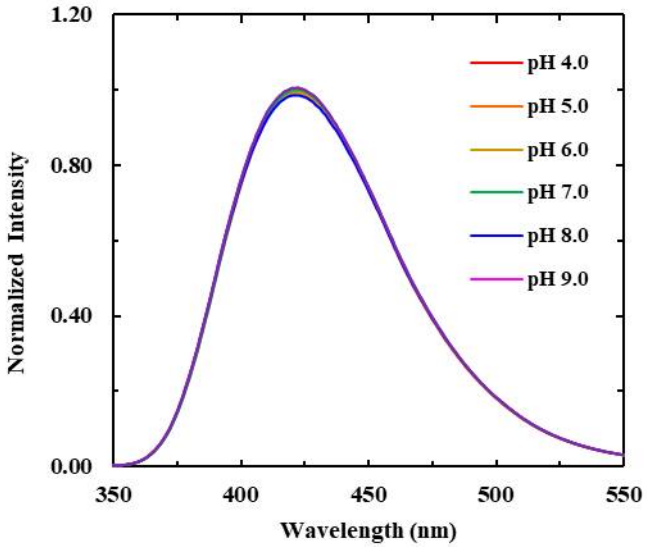


Figure S3. Normalized fluorescence spectra of Ac-G-L-4CN-Trp-G at different pH values (as indicated), using the spectrum obtained at pH 7.0 as the reference. The concentration was 10 μM for each sample and the excitation wavelength was 320 nm.

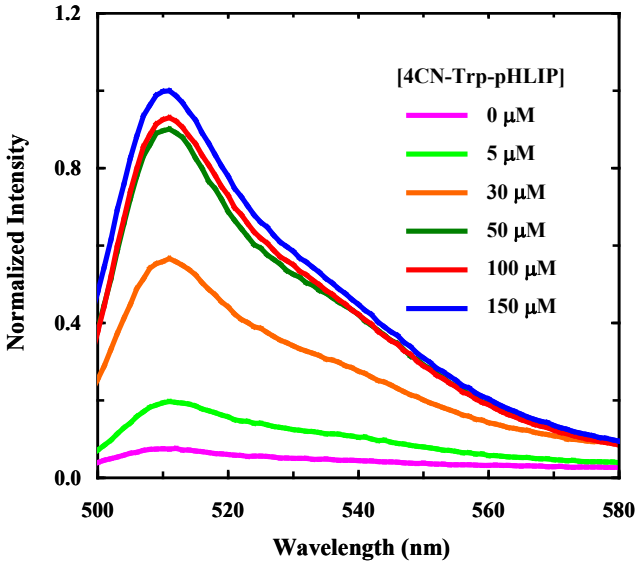


Figure S4. Normalized fluorescence spectra (in the DiO emission wavelength region) of mixtures of DiO-stained POPC (1.0 mM) LUVs and 4CN-Trp-pHLIP of different concentrations, as indicated. The excitation wavelength was 320 nm.

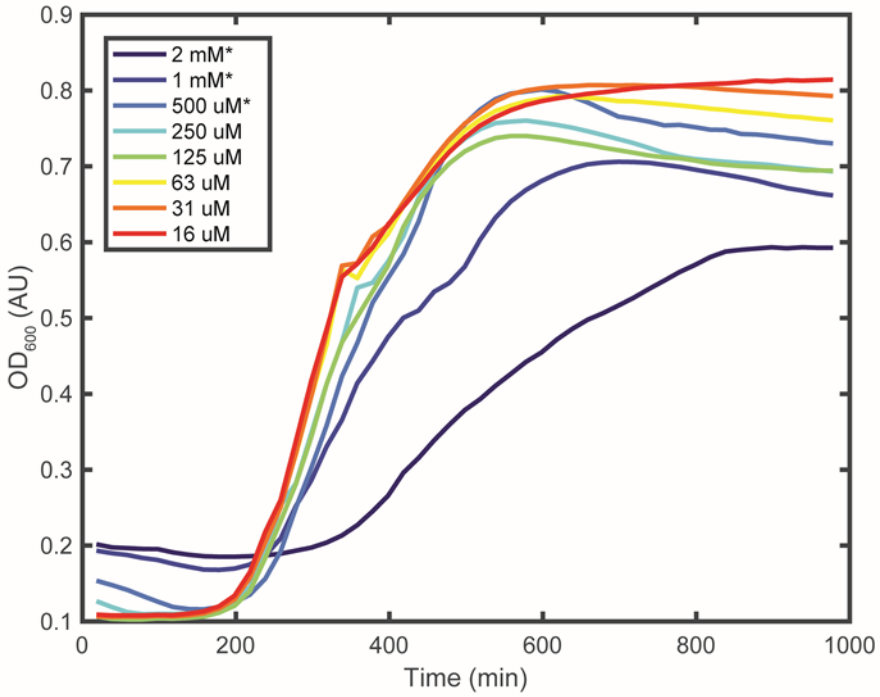


Figure S5. *E. coli* growth curves in the presence of different concentrations of L-4CN-Trp, which suggest that at concentrations of below 250 µM, L-4CN-Trp has no significant effect on the growth of *E. coli* cells. It is worth noting that for experiments carried out at higher concentrations, marked with an asterisk (*), the results are not conclusive because the poor solubility of the compound at these concentrations.

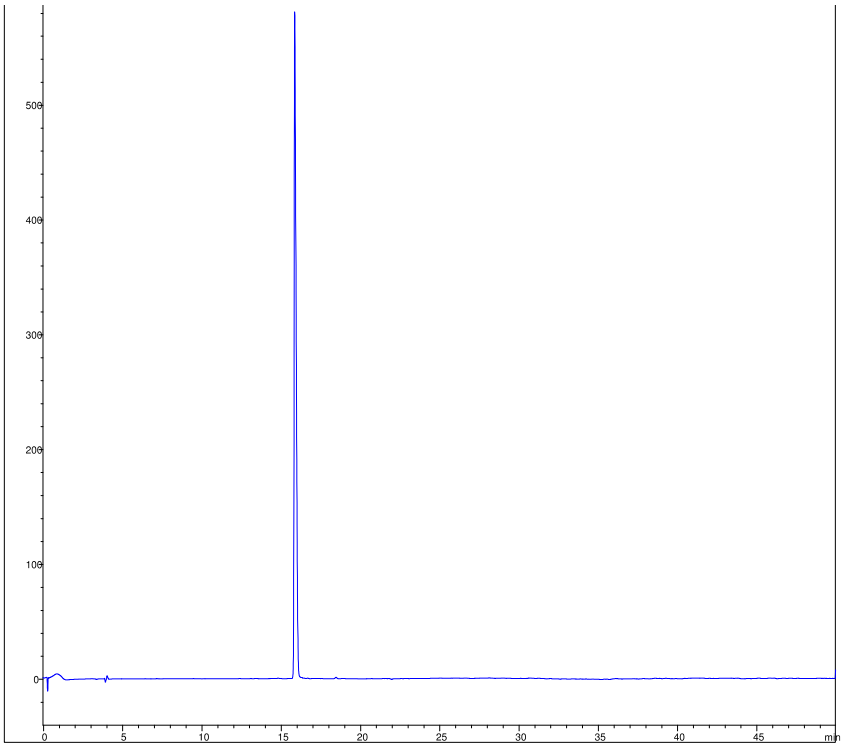


Figure S6. HPLC profile of purified L-4CN-Trp **7**.

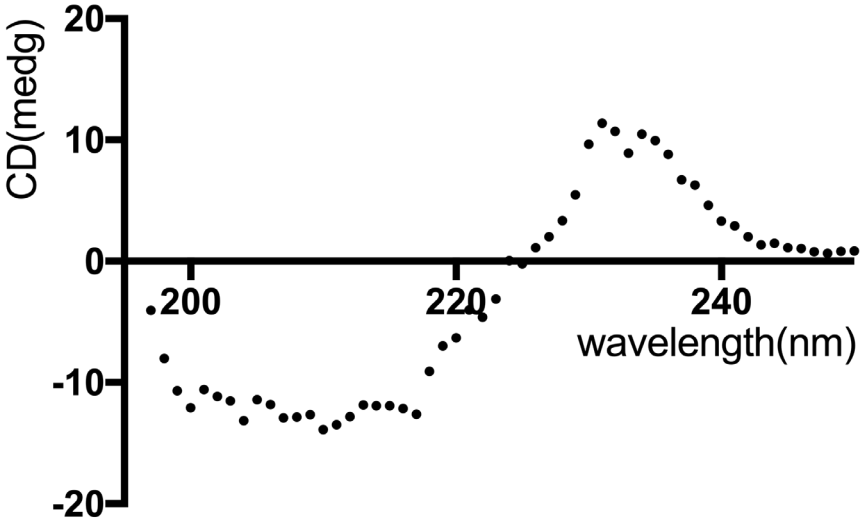


Fig S7. CD spectrum of L-4CN-Trp **7** (2 mM, in 0.1 M HCl)

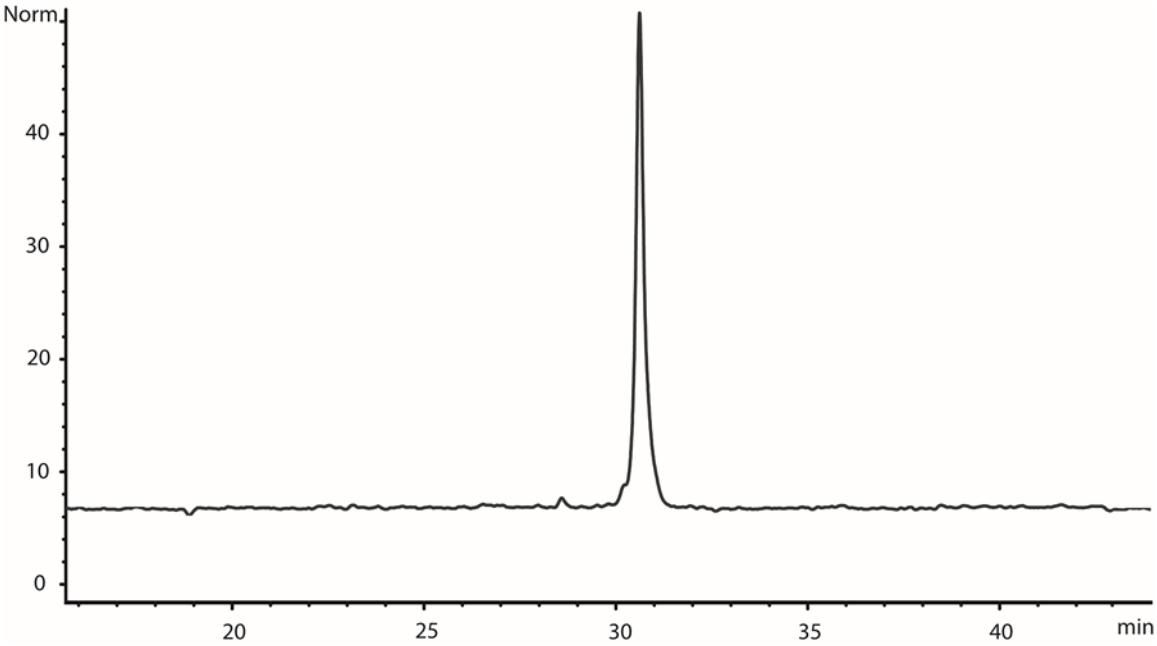


Figure 8. HPLC profile of purified 4CN-Trp-pHLIP

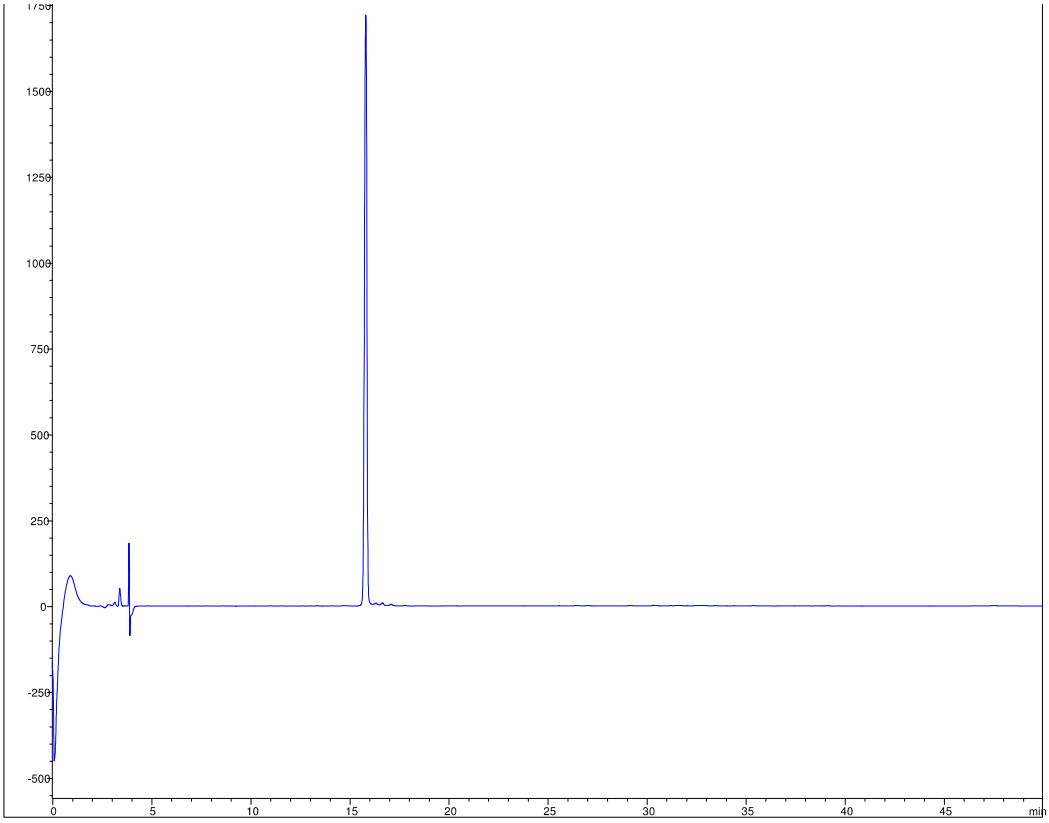
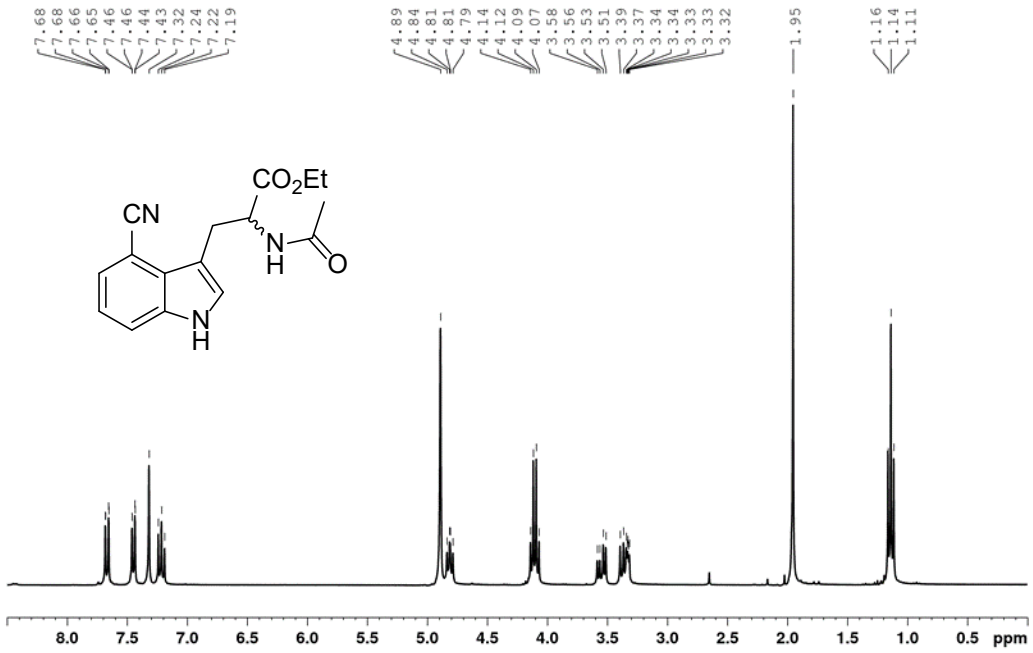
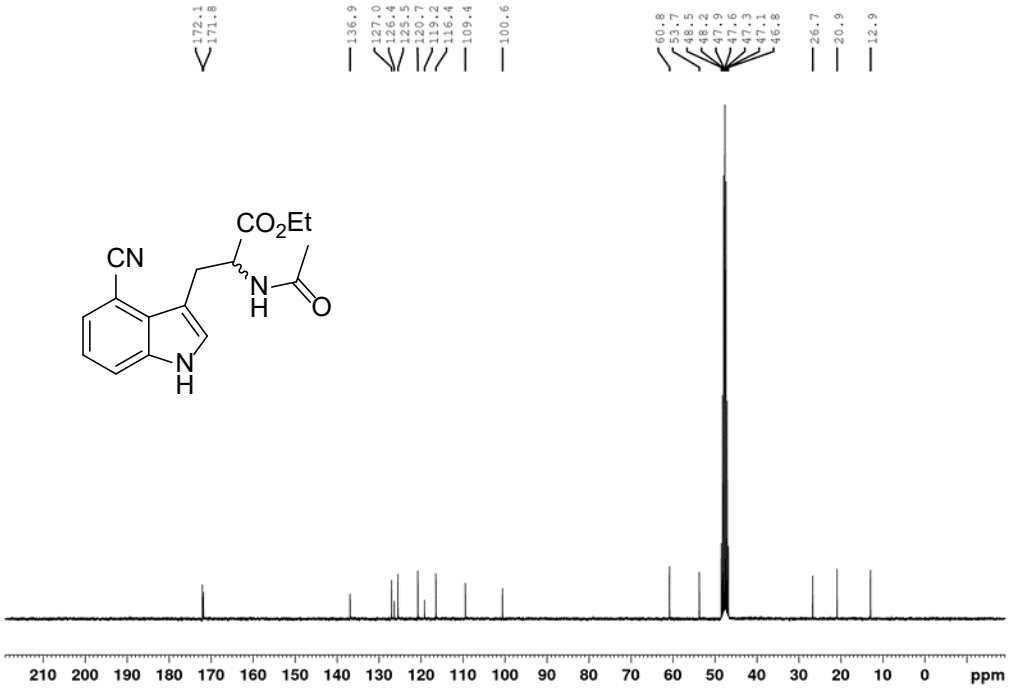


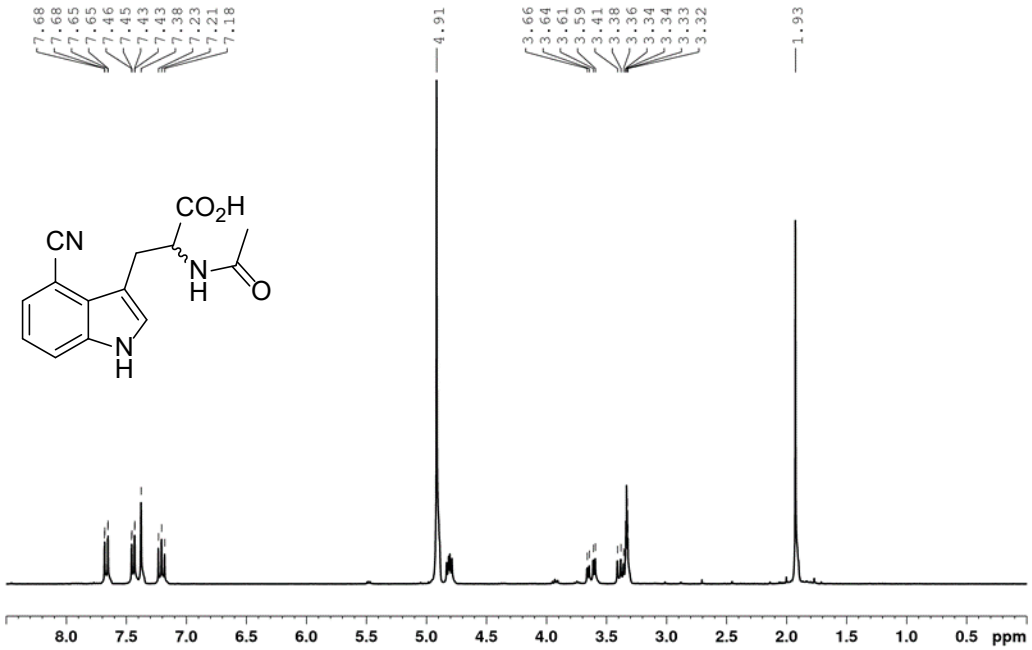
Figure 9. HPLC profile of purified Ac-G-L-4CN-Trp-G.



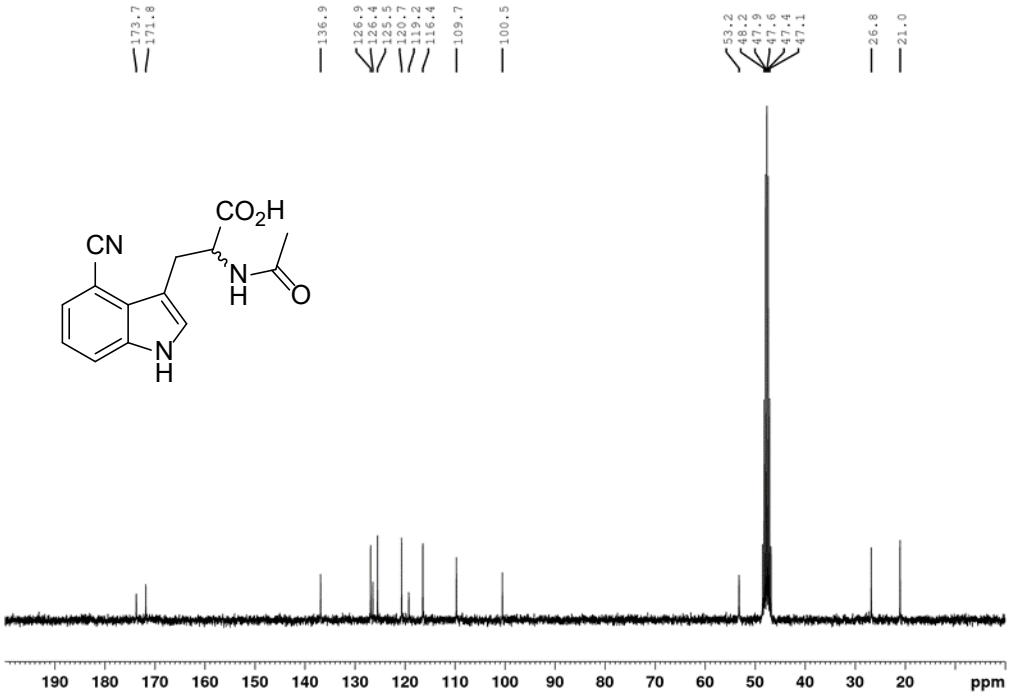
^{1}H NMR spectrum of **5**



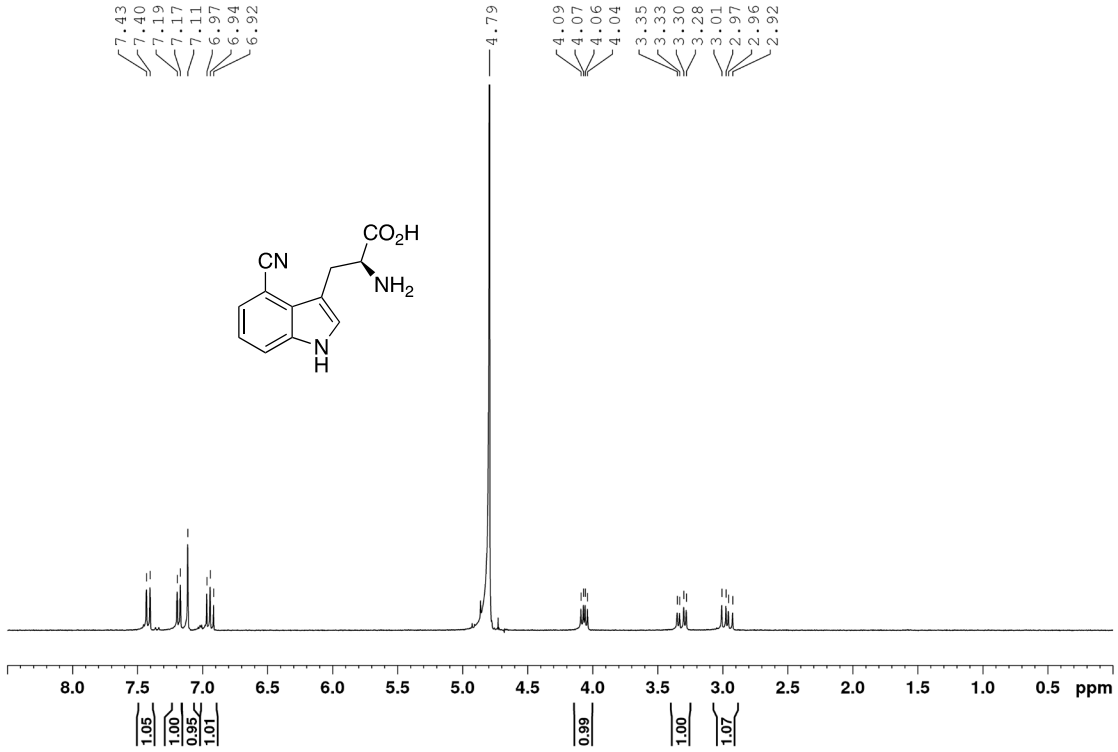
^{13}C NMR spectrum of **5**



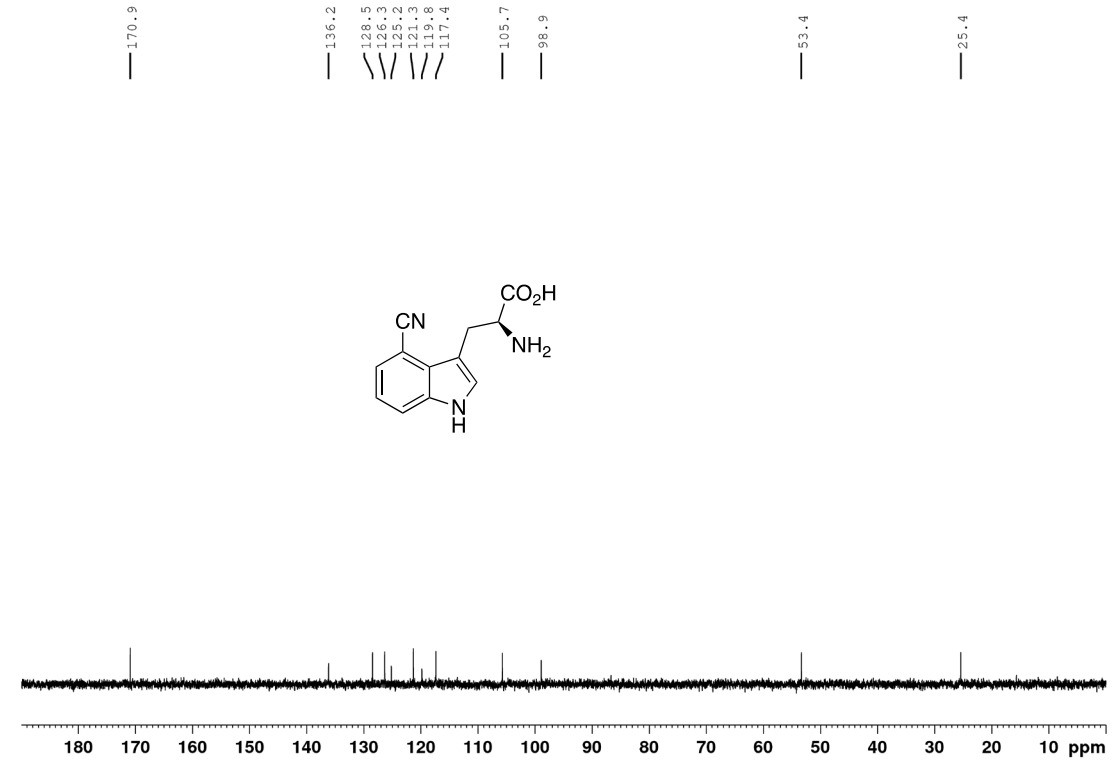
^{1}H NMR spectrum of **6**



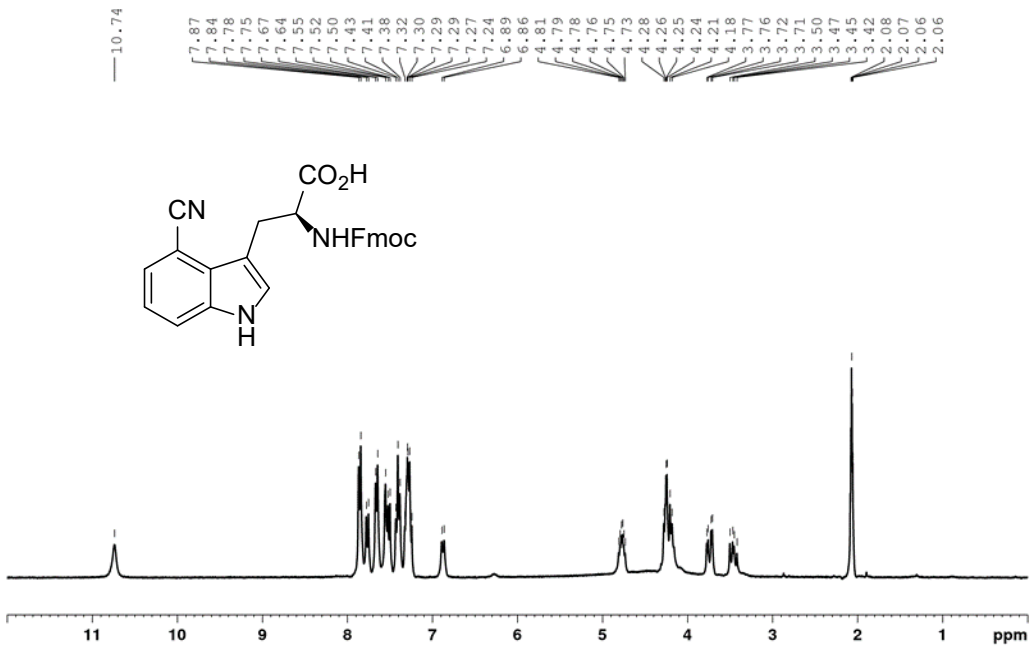
^{13}C NMR spectrum of **6**



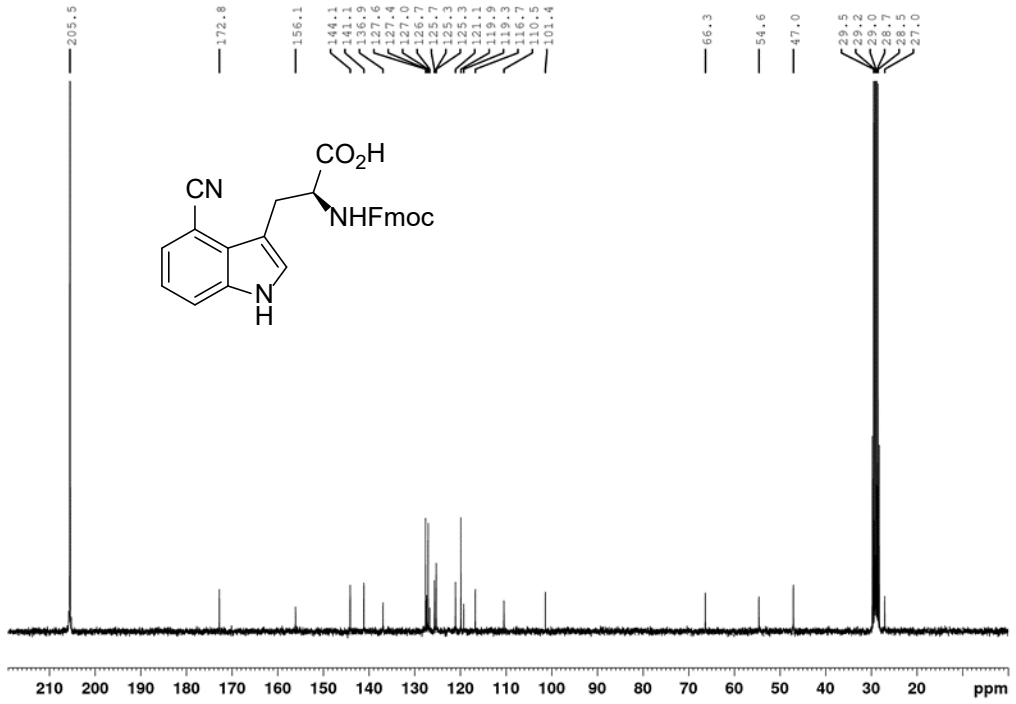
^{1}H NMR spectrum of **7**



^{13}C NMR spectrum of **7**



^{1}H NMR spectrum of **8**



^{13}C NMR spectrum of **8**

References

1. L. L. Guo, Jian; Nargund, Ravi P.; Pasternak, Alexander; Yang, Lihu; Ye, Zhixiong, *PCT Int. Appl., 2010051177, 06 May 2010*
2. M. Winn, D. Francis and J. Micklefield, *Angew. Chem., Int. Ed.*, 2018, **57,** 6830.
3. C. A. Schneider, W. S. Rasband and K. W. Eliceiri, *Nat Methods*, 2012, **9**, 671.
4. T. S. Aurora, W. Li, H. Z. Cummins and T. H. Haines, *Biochim Biophys Acta*, 1985, **820**, 250.
5. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer: New York, USA 2nd edn., 1999.
6. Y. K. Reshetnyak, O. A. Andreev, M. Segala, V. S. Markin and D. M. Engelman, *Proc Natl Acad Sci U S A*, 2008, **105**, 15340.