Supporting Information

Modular Control of L-Tryptophan Isotopic Substitution via an Efficient Biosynthetic Cascade

Thompson, C. M.; McDonald, A. D.; Yang, H.; Cavagnero, S.*; Buller, A. R.*

Department of Chemistry, University of Wisconsin-Madison

1101 University Avenue Madison, WI 53706

*corresponding authors: Andrew R. Buller, email:arbuller@wisc.edu; Silvia Cavagnero, email: cavagnero@chem.wisc.edu

Table of contents

1. General information

The glassware used in the reactions carried out in this work was thoroughly washed, and all experiments were executed following necessary safety precautions. Evaporation of solvents was performed at reduced pressure using a rotary evaporator. Electronic-absorption measurements were done with a UV-2600 spectrophotometer (Shimadzu).

2. Experimental section

2.1. Materials

Chemicals and solvents were obtained from commercial suppliers and used without further purification: formaldehyde-D_2 (Cambridge Isotope Laboratories, Inc.); (2-^{13}C)glycine (Cambridge Isotope Laboratories, Inc.); D_2O 99.9% (Sigma-Aldrich); indole (Sigma-Aldrich); pyridoxal 5' monophosphate; (Sigma-Aldrich) L-serine (Sigma-Aldrich). Analytical-scale reaction products were analyzed by UPLC/LC-MS on an Acquity UHPLC apparatus equipped with an Acquity QDA MS detector (Waters), using an Acquity UPLC CSH Phenyl-Hexyl column (Waters). Preparative-scale flash chromatographic separations were performed on an Isolera One Flash Purification system (Biotage). NMR analysis was carried out on either a Bruker AVANCE 400 NMR spectrometer or Bruker AVANCE 600 NMR spectrometer.

2.2. Methods

2.2.1 Cloning, expression, and purification of *Pf*TrpB^{2B9} and *Tm*LTA

Codon-optimized copies of the engineered Pyrococcus furiosus tryptophan synthase β-subunit (*Pf*TrpB^{2B9}) and Thermotoga maritima L-threonine aldolase (*Tm*LTA, EC 4.1.2.5)

genes were purchased as gBlocks from Integrated DNA Technologies.[1] These DNA fragments were inserted into pET22b vectors via Gibson Assembly.[2] BL21 (DE3) E. coli cells were separately transformed with plasmids encoding each of the above proteins via electroporation. After 30 min of growth in Luria Bertani (LB) broth at 37 °C, cells were plated onto LB plates in the presence of 100 μg/mL ampicillin (amp) and incubated overnight. Single colonies were used to inoculate 2 x 5 mL Terrific Broth (TB) containing 100 μg/mL amp, followed by overnight growth at 37 °C. Expression cultures, typically consisting of 1 L of TB and 100 μg/mL amp, were inoculated with the above starter culture encoding the gene of the desired protein and shaken (180 rpm) at 37 °C. After 3 h (OD_{600} = 1.5), the cultures were chilled on ice for 60 min. Protein overexpression was induced upon addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cultures were expressed overnight at 23 °C upon shaking at 200 rpm. Cell were then harvested upon centrifugation at 4300 x g at 4 °C for 30 min. Cell pellets were then immediately subject to lysis or stored frozen at -20 °C until needed.

When stored frozen, cell pellets were thawed on ice. Cells were then resuspended in lysis buffer (50 mM potassium phosphate buffer pH = 8.0, 1 mg/mL hen egg white lysozyme (GoldBio), 0.2 mg/mL DNaseI (GoldBio), 1 mM MgCl2 and 400 μM pyridoxal 5′-phosphate (PLP). A volume of 4 mL of lysis buffer was used per each gram of wet cell pellet. After 30 min under shaking at 37 °C, the resuspended lysis mixture was heated for 10 min at 75 °C. The resulting lysate was spun down at 75,600 x g to pellet cell debris and other insoluble products. Ni/NTA beads (GoldBio) were added to the supernatant and incubated on ice for 40 min prior to purification by Ni-affinity chromatography. The column was washed with 3 column volumes of 20 mM imidazole, 50 mM potassium

phosphate adjusted to pH = 8.0, 3 column volumes of 40 mM imidazole, 50 mM potassium phosphate adjusted to pH = 8.0 and eluted with 250 mM imidazole, 50 mM potassium phosphate buffer pH = 8.0. Elution of the desired protein was monitored by eye inspection via the disappearance of the bright-yellow column band arising from the protein-associated PLP cofactor. Note that PLP associates to both *Pf*TrpB^{2B9} and *Tm*LTA. Solutions were dialyzed to < 1 μM imidazole, flash-frozen by dripping into liquid nitrogen and stored at -80 °C. Protein concentration was estimated via a Bradford assay. The above procedure yielded ~500 mg *Pf*TrpB^{2B9}per L culture and ~1000 mg *Tm*LTA per L culture. SDS-Page analysis of the expression mixture and purified enzymes is shown in Figure S1.

2.2.2 Optimization of *Tm*LTA-TrpB^{2B9} cascade

Analytical-scale reactions to determine optimal conditions for L-tryptophan (Trp) production for the *Tm*LTA-*Pf*TrpB^{2B9} cascade were performed in 5 mL glass vials in quadruplicate (final volume 500 μL). Reactions were performed in the presence of 5 mM Gly, 2.5 – 10 mM indole, 0.004 – 0.04 μM *Tm*LTA, 0.004 – 0.04 μM *Pf*TrpB^{2B9}, 0.02 mM PLP, 3.75 – 7.5 mM formaldehyde and 50 mM potassium phosphate adjusted to pH = 8.0 (Fig. 2). All reactants and catalysts were prepared as stock solutions in 50 mM potassium phosphate adjusted to pH = 8.0. Reactions were carried out for 16 h at 50 °C and quenched with 500 μL 1:1 MeOH: 1 M HCl, followed by vortexing and centrifugation at 20,000 x g for 10 min, to pellet down the enzymes. The amount of Trp in the supernatant was quantified via UPLC-MS by integrating the single-ion peak with m/z + 1 = 205.09. To assess the amount of generated product (Trp, see Supporting Table S2), peak intensities were compared to those of a Trp standard of known concentration via a calibration curve.

2.2.3 Preparative-scale synthesis of Trp isotopologs

(2-^{2}H)Trp

The reactants below were added in the following order to a capped high-pressure flask: indole (11.7 mg, 0.1 mmol), L-serine (32.0 mg, 0.3 mmol) and *Pf*TrpB^{2B9} enzyme (4.3 mg, 0.1 μmol). The total reaction volume was adjusted to 5 mL with 50 mM potassium phosphate in D_2O, with pD = 8.0. pD was determined by a pH electrode according to: pD = pH + 0.42.[3] The buffer pH was adjusted to ~7.6 with DCl. Likewise, all reagent stock solutions were in D_2O and contained 50 mM potassium phosphate at pD = 8.0. The reaction mixture was incubated overnight at 75 °C. The reaction solution was incubated at 90 °C for 15 min to heat-quench. The enzymes were then removed from the post-reaction mixture upon centrifugation at 20,000 x g for 10 min. The Trp isotopolog was purified via an Isolera One Flash Purification system (Biotage) under isocratic conditions (10% MeOH, 90% H_2O). Samples containing desired Trp isotopolog were pooled and evaporated to dryness. Dried samples were resuspended in water, frozen, and lyophilized to remove excess solvent. The lyophilized samples were weighed. Identity was assessed and purity was confirmed by ^{1}H- and ^{13}C-NMR.

(2-^{13}C, 2, 3, 3-^{2}H$_3$)Trp

The reactants below were added in the following order to a capped high-pressure flask: indole (4.1 mg, 0.035 mmol), 50 mM potassium phosphate (6.75 mL), (2-^{13}C) glycine (Gly, 70 μL of 0.5 M material, i,e., 0.035 mmol), PLP (7.00 μL of 20 mM PLP, i.e., 0.14 μmol), *Pf*TrpB^{2B9} (1.5 mg, 0.035 μmol) and *Tm*LTA (1.3 mg, 0.035 μmol) enzymes,

formaldehyde-D_2 (70 µL of 0.500 M material, 0.035 mmol). The final volume was adjusted to 7 mL with 50 mM potassium phosphate at pD = 8.0, respectively. All reagent stock solutions were in D_2O and contained 50 mM potassium phosphate at pD = 8.0. The reaction mixture was incubated overnight at 50 °C. The reaction solution was incubated at 90 °C for 15 min to heat-quench. The enzymes were then removed from the post-reaction mixture upon centrifugation at 20,000 x g for 10 min. The Trp isotopolog was purified via an Isolera One Flash Purification system (Biotage) under isocratic conditions (10% MeOH, 90% H_2O). Samples containing desired Trp isotopolog were pooled and evaporated to dryness. Dried samples were resuspended in water, frozen, and lyophilized to remove excess solvent. Lyophilized samples were weighed, and purity / identity were assessed by ^{1}H- and ^{13}C-NMR.

2.2.4 Analytical- and preparative-scale chromatography

Analytical-scale reaction products were analyzed by UPLC/LC-MS on an Acquity UHPLC apparatus equipped with an Acquity QDA MS detector (Waters), using an Acquity UPLC CSH Phenyl-Hexyl column (Waters). Preparative-scale flash chromatographic separations were performed on an Isolera One Flash Purification system (Biotage).

2.2.5 Determination of enantiopurity of Trp isotopologs

Enantiopurity of the Trp isotopologs was assessed upon treatment with the Marfey's reagent analog 2-(5-fluoro-2,4-dinitrophenyl)amino- L-valinamide (FDVA) as described.[4] Briefly, 1 mM of the newly synthesized Trp isotopolog (or a standard consisting of a D/L-Trp mixture of known concentration) were mixed with 5 mM FDVA, 400 mM sodium bicarbonate in a 100 µL solution of 20:80 acetonitrile:H2O (v/v) and incubated for 2 h at 40 °C. The reaction was quenched with 200 µL of 1:1 acetonitrile:1 M HCl (v/v), followed

by UPLC-MS analysis. Enantiopurity of the Trp isotopologs was assessed upon comparison with standards composed of a D-Trp/L-Trp mixture of known concentration (Fig. S2).

2.2.6 NMR characterization of synthesized compounds

The extent of ^{13}C- and ^{2}H-incorporation at the α- and β-positions of deuterated Trp isotopologs was determined by ^{1}H and ^{13}C nuclear magnetic resonance (^{1}H, ^{13}C NMR, pulse-acquire experiments) on a 400 MHz Bruker Avance spectrometer in D_2O, in the presence of trace amounts of 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), employed as an internal chemical-shift standard.

All other ^{1}H NMR data were collected on a Bruker 600 MHz spectrometer equipped with an Avance III HD console and a cryogenic $^{1}H/^{13}C/^{15}N$ TCI-F probe. NMR experiments were carried out in 90% H_2O and 10% D_2O with presaturation of the HDO-resonance (4.70 ppm).

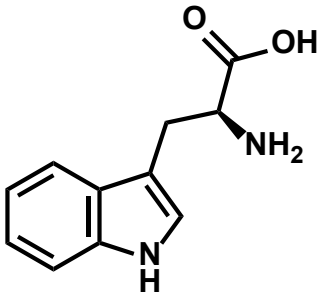
1D ^{1}H spectra were acquired with a 15.97 ppm sweepwidth, 38,312 total points and 8 scans. All ^{1}H spectra were zero-filled to 65,536 complex points and apodized with an exponential window function (0.5 Hz line-broadening).

1D ^{13}C spectra were acquired with ^{1}H and ^{2}H decoupling during acquisition, a 282.5 ppm sweepwidth centered at 110 ppm and 71,424 total points. A variable number of scans were acquired until a suitable S/N was achieved. All ^{13}C spectra were zero-filled to 131,072 complex points and apodized with an exponential window function (1 Hz linebroadening). $^{13}C^{\alpha}$ longitudinal relaxation times (T1) of Trp isotopologs were determined via the inversion recovery pulse sequence with ^{1}H and ^{2}H decoupling during acquisition. $^{13}C_{\alpha}$ transverse relaxation times (T2) were assessed with the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence in the presence of ^{1}H and ^{2}H decoupling during both evolution and

acquisition times. T1 and T2 data had the same acquisition parameters as the other ^{13}C experiments, except that the center frequency was set to 55 ppm (i.e., centered on C_α resonances) to prevent undesired off-resonance effects.

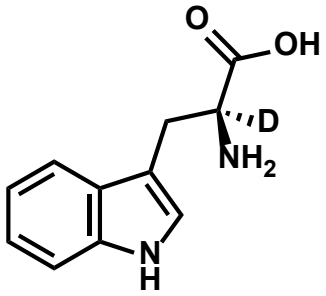
3. Characterization of Trp and Trp isotopologs

L-tryptophan



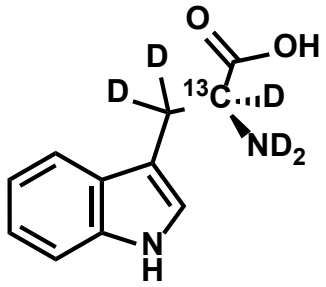
White solid (Sigma-Aldrich); ^1H NMR (600 MHz, 90% H_2O, 10% D_2O) δ 7.72 (d, J = 7.9 Hz, 1H), 7.53 (d, J = 8.2 Hz, 1H), 7.30 (d, J = 1.78 Hz, 1H), 7.27 (t, J = 7.6 Hz, 1H), 7.19 (t, *J* = 7.5 Hz, 1H), 4.04 (dd, J = 4.82, 8.1 Hz, 1H), 3.47 (dd, *J* = 15.4, 4.8 Hz, 1H), 3.28, (dd, J = 15.4, 8.1 Hz, 1H); ^{13}C NMR (151 MHz, 90% H_2O, 10% D_2O) δ 174.60, 136.50, 126.72, 125.25, 122.18, 119.53, 118.51, 112.06, 107.60, 55.28, 26.53 (Figs. S3 and S4).

(2-^{2}H)-L-tryptophan



Yield 89% (18.2 mg); white solid; ^{1}H NMR (600 MHz, 90% H_2O, 10% D_2O) δ 7.73 (d, J = 7.9 Hz, 1H), 7.53 (d, J = 8.1 Hz, 1H), 7.31 (s, 1H), 7.28 (t, J = 7.6 Hz, 1H), 7.20 (t, *J* = 7.5 Hz, 1H), 3.46 (d, *J* = 15.4, 1H), 3.28, (dd, J = 15.3, 1H); ^{13}C NMR (151 MHz, 90% H_2O, 10% D_2O) δ 174.50, 136.36, 126.58, 125.10, 122.04, 119.38, 118.37, 111.91, 107.47, 54.83, 26.30 (Figs S5 and S6).

(2-^{13}C, 2, 3, 3-$^{2}H_3$)-L-tryptophan



Yield 60% (4.8 mg); white solid; ^{1}H NMR (600 MHz, 90% H_2O, 10% D_2O) δ 7.73 (d, J = 7.9 Hz, 1H), 7.53 (d, J = 8.2 Hz, 1H), 7.29 (s, 1H), 7.27 (t, J = 7.6 Hz, 1H), 7.19 (t, *J* = 7.5 Hz, 1H); ^{13}C NMR (151 MHz, 90% H_2O, 10% D_2O) δ 136.45, 126.85, 125.08, 122.11, 119.46, 118.59, 112.01, 55.10 (Figs S7 and S8).

4. Supporting references

1 D. K. Romney, J. Murciano-Calles, J. E. Wehrmüller and F. H. Arnold, *J. Am. Chem. Soc.*, 2017, **139**, 10769–10776.

2 D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith, *Nat. Methods*, 2009, **6**, 343–345.

3 A. Krezel and W. Bal, *J. Inorg. Biochem.*, 2004, **98**, 161–166.

4 P. Marfey, *Carlsb. Res. Commun.*, 1984, **49**, 591–596.

5. Supporting tables and figures

Table S1. Reaction yields and enantiomeric excess of Trp isotopologs prepared in this work.[a]

Starting material	Product	Mass (mg Trp)	Yield (%)	ee[b] (%)
L-serine	(2-^{2}H)Trp	18.2	89	>99
(2-^{13}C) glycine	(2-^{13}C, 2, 3, 3-$^{2}H_3$)Trp	4.8	60	>99

[a]All reactions were carried out in 50 mM potassium phosphate in D_2O, at pD = 8.0.

[b]Enantiomeric excess was determined via UPLC-MS after derivatization of the reaction products with a Marfey's reagent analog (see section 2.2.5).

Table S2. Overview of tested experimental conditions (analytical scale) for the synthesis of Trp via the *Tm*LTA-*Pf*TrpB^{2B9}-mediated cascade introduced in this work.

Entry[a]	Eq./ Conc.	Yield[b] (%)	Yield Std. error (%)	[Trp] (mM) in post-reaction mixture	Std. error ± [Trp] (mM)
1	4 μM	29	5	6	1
2	8 μM	43	2	8.7	0.3
3	20 μM	78	4	16	1
4	40 μM	79	4	16	1
5	0.5 eq.	39	3	8	1
6	0.75 eq.	61	3	12	1
7	1.0 eq.	80	8	16	2
8	2.0 eq.	75	8	15	2
9	0.75 eq.	81	6	16	1
10	1.0 eq.	79	9	16	2
11	1.5 eq	18	5	4	1
12	5 mM	97	10	4.8	0.5
13	10 mM	91	4	9.1	0.4
14	20 mM	75	7	15	1
15	40 mM	48	5	19	2

[a] Entries 1-4: [enzymes], with: 1:1 *Pf*TrpB^{2B9} and *Tm*LTA; entries 4-8: indole equivalents (relative to glycine); entries 9-11: formaldehyde equivalents (relative to glycine); entries 12-15: reactant concentration (1: 1: 1 glycine: indole: formaldehyde) with 0.001 eq. of *Pf*TrpB^{2B9} and *Tm*LTA.
[b] Trp assessed via UPLC-MS, single-ion m/z+1 = 205.09.

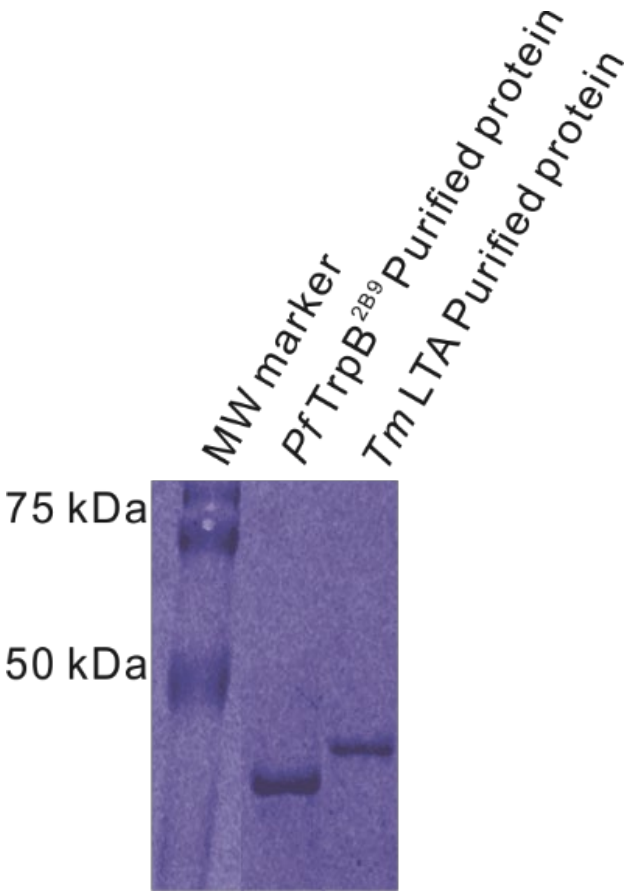
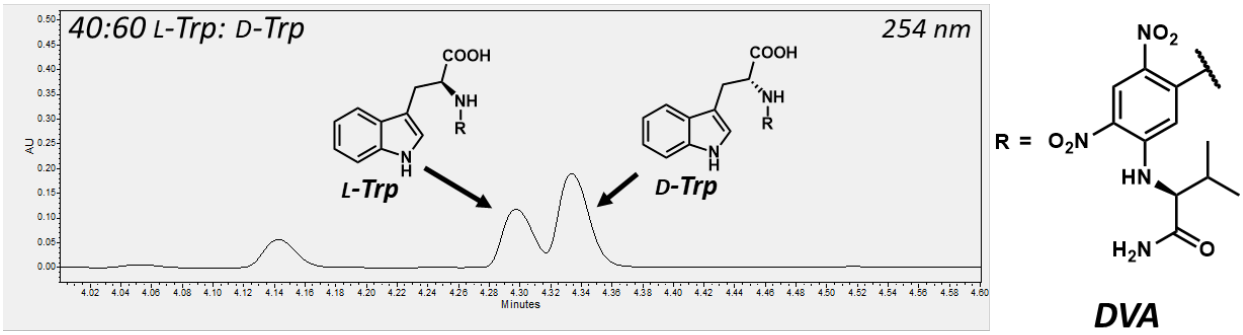
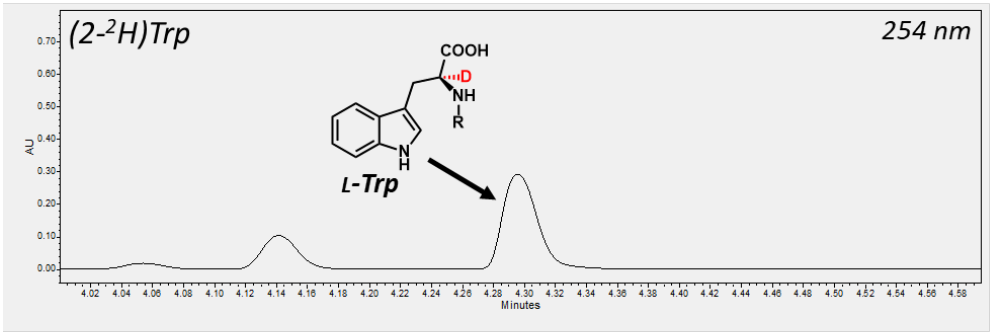


Figure S1. SDS-Page analysis of purified *Pf* TrpB^{2B9} and *Tm* LTA enzymes.





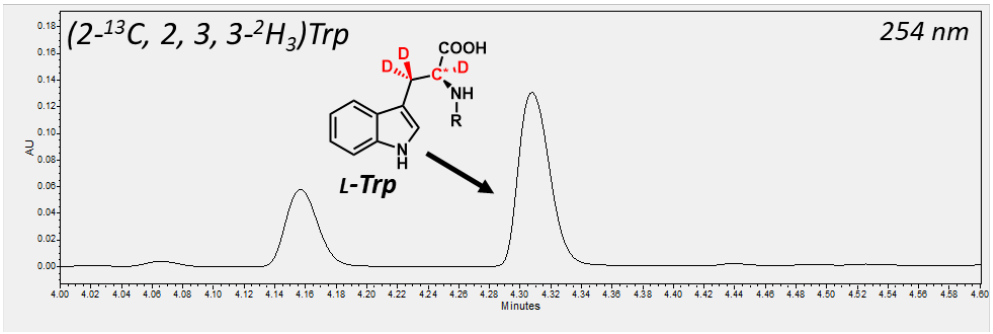


Figure S2. Chromatogram highlighting (2-^{13}C, 2, 3, 3-$^{2}H_3$)Trp isotopolog derivatized with FDVA (see section 2.2.6) to assess enantiomeric purity. The additional peak on the left hand side (retention time ~4.1 min) is due to 1,5-difluoro, 2,4-dinitrobenzene after the synthesis.

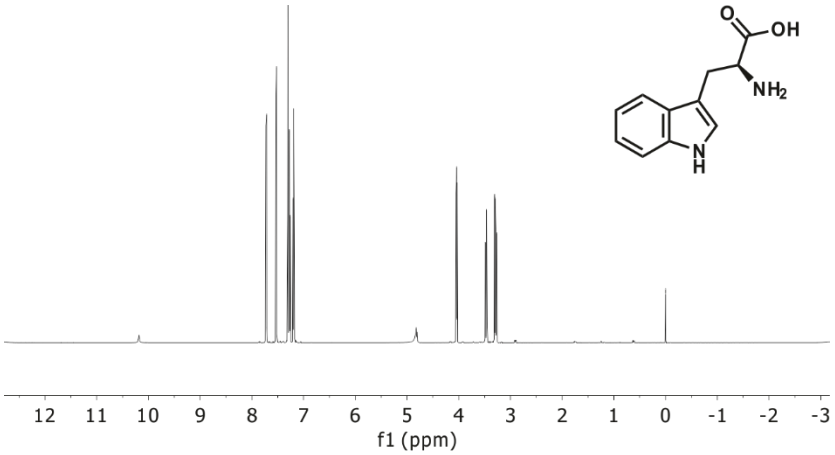


Figure S3. ^{1}H NMR spectrum of Trp (8 scans).

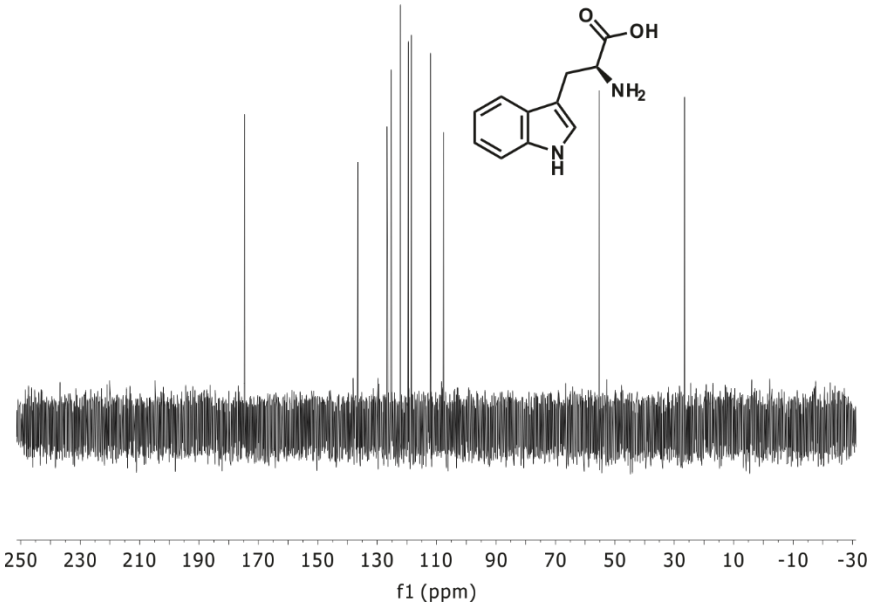


Figure S4. ^{13}C NMR spectrum of Trp (with ^{1}H-, ^{2}D-decoupling during acquisition, 8 scans).

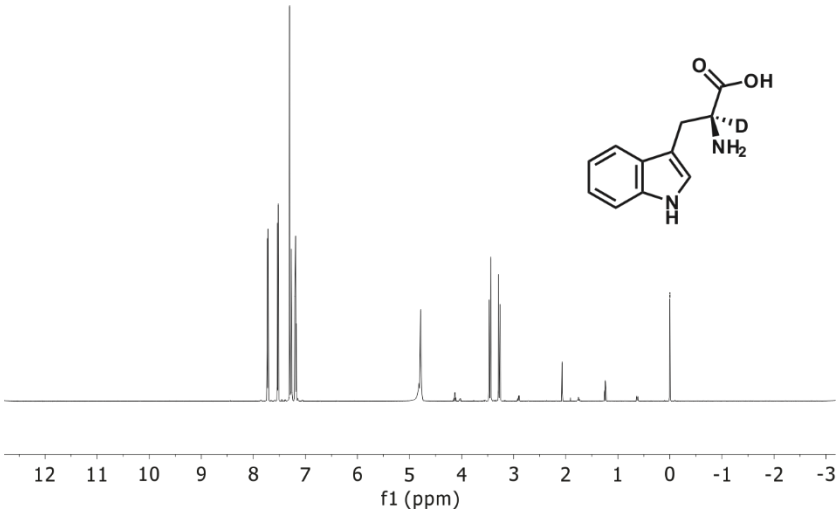


Figure S5. ^{1}H NMR spectrum of (2-^{2}H)Trp (8 scans).

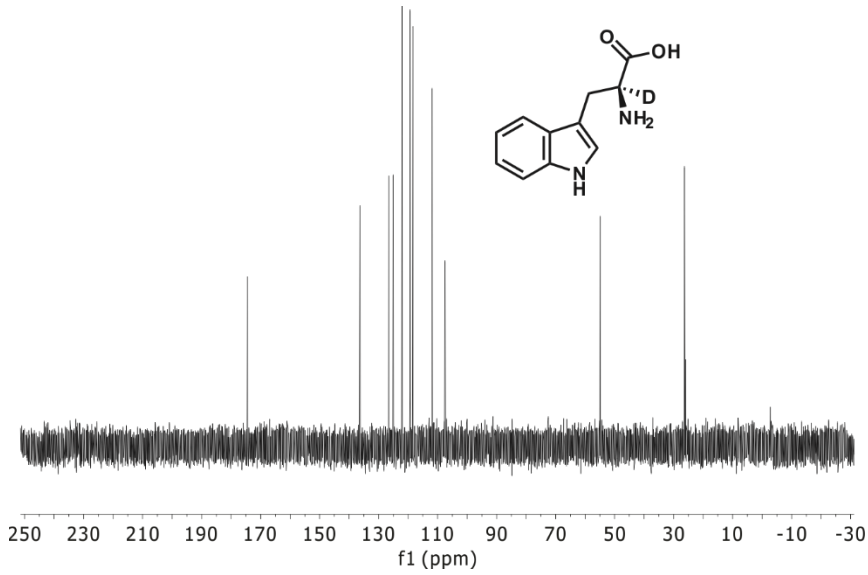


Figure S6. ^{13}C NMR spectrum of (2-^{2}H)Trp (512 scans).

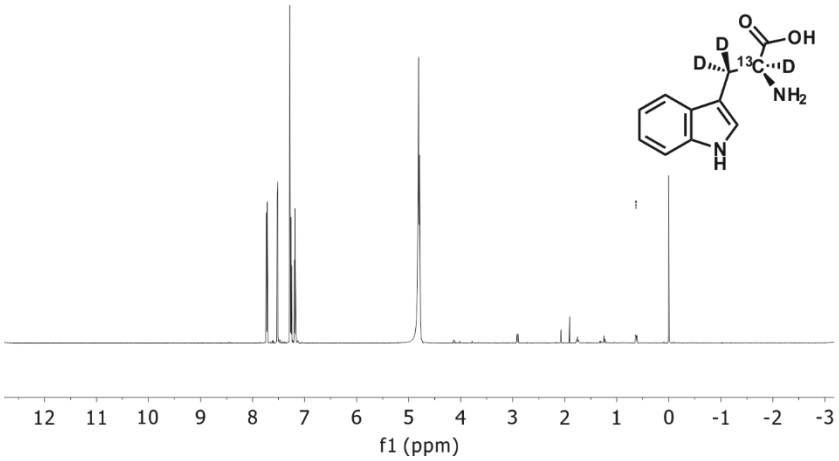


Figure S7. ^{1}H NMR spectrum of (2-^{13}C, 2, 3, 3-$^{2}H_3$)Trp (8 scans).

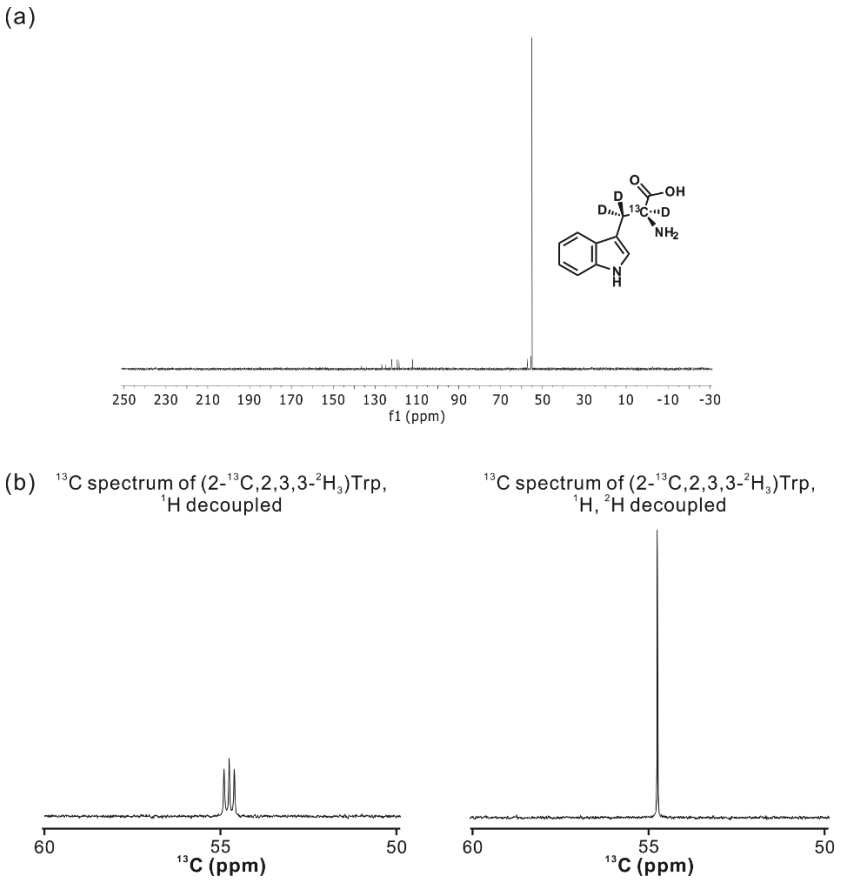


Figure S8. (a) ^{13}C NMR spectrum of (2-^{13}C, 2, 3, 3-$^{2}H_3$)Trp (512 scans), 1H and 2H decoupled. (b) Comparison ^{13}C NMR spectra of (2-^{13}C, 2, 3, 3-$^{2}H_3$)Trp (2 scans), in the absence and presence of ^{2}H decoupling.