**Supporting Information**

# Synthesis and Characterization of an Unnatural Boron and Nitrogen-containing Tryptophan Analogue and its Incorporation into Proteins

Katherine Boknevitz, James S. Italia, Bo Li, Abhishek Chatterjee*, and Shih-Yuan Liu*

Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467, United States

## Table of Contents

General Considerations……S-2

Synthesis of BN-tryptophan **8**……..S-3

Chiral Resolution of BN-tryptophan ester **7**……S-7

Determination of Enantiopurity and Absolute Configuration of BN-Tryptophan **8**……S-9

Quantum Yield Measurement of (L)-BN-Tryptophan **8** and (L)-Tryptophan **9**…….....S-12

Solvatochromism of BN-tryptophan ester **7** and L-tryptophan ester **S1**……S-13

NMR Assignment of BN-tryptophan **8**…….…….S-14

Cell Fermentation and Protein Isolation……..……...S-16

Oxidation Study……...……...S-21

X-Ray Crystallographic Data for BN-trp **6**……S-25

References……...……...S-26

NMR Spectra……...……...S-27

***General Considerations***

All oxygen- and moisture-sensitive manipulations were carried out under $N_2$ using either standard Schlenk techniques or a nitrogen-filled glovebox. THF, $Et_2O$, $CH_2Cl_2$, toluene, and pentane were dried with a solvent purification system consisting of columns of molecular sieves under argon. Silica gel (230-400 mesh) was dried for 12 hours at 180 °C under high vacuum. Flash chromatography was performed with this silica gel under an inert atmosphere. All other chemicals and solvents were purchased and used as received.

NMR spectra were recorded on a Varian VNMRS 600 MHz, VNMRS 500 MHz, INOVA 500 MHz, or VNMRS 400 MHz spectrometer. Deuterated solvents were purchased from Cambridge Isotope Labs. $^{11}B$ NMR spectra were externally referenced to $BF_3 \bullet Et_2O$ ($\delta$ 0). All NMR chemical shifts are reported in ppm relative to residual solvent for $^{1}H$ and $^{13}C$ NMR.

Infrared spectroscopy was performed on a Bruker ALPHA-Platinum FT-IR Spectrometer with ATR-sampling module.

High-resolution mass spectra were collected by Marek Domin on a JEOL AccuTOF instrument (JEOL USA, Peabody, MA), equipped with a DART ion source (IonSense, Inc., Danvers, MA) in positive ion mode at the Boston College Center for Mass Spectrometry.

UV-vis absorption spectra were collected on an Agilent Cary 100 UV-Vis spectrometer. Emission spectra were collected on a Photon Technology International spectrometer. Quantum yield ($\Phi$) was determined using a PTI K-Sphere "Petite" integrating sphere.

Analytical HPLC spectra were recorded on an Agilent 1200 Infinity Series HPLC equipped with a DAD detector. Recycling HPLC was performed on a Japan Analytical Co., Ltd. Next Recycling Preparative HPLC with UV/vis detector.

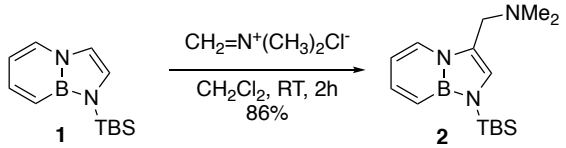
Specific rotations measured on a Rudolph Research Analytical Autopol IV.

Circular dichroism spectra were recorded on an Aviv Biomedical Inc. Model 420 Circular Dichroism Spectrometer.

***Synthesis of BN-tryptophan 8***

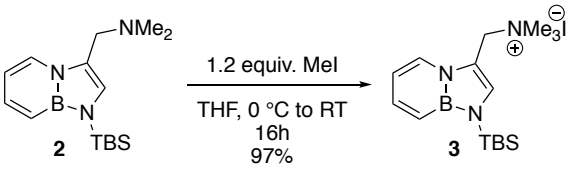
BN-indole **1** was prepared according to Abbey, et al[1].

**Synthesis of 2:**



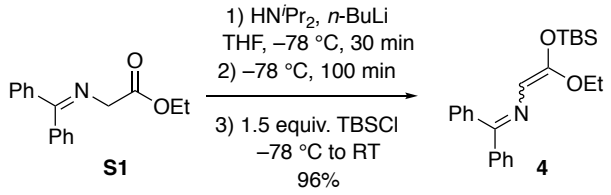

In a glove box, a 500 mL round bottom flask (RBF) equipped with a stir bar was charged with TBS-BN-indole **1** (2.0 g, 8.6 mmol) and N,N'-dimethylmethylene-iminium chloride (967 mg, 10.3 mmol) and dissolved in $CH_2Cl_2$ (100 mL). The reaction was allowed to stir for 3 hours. Solvent was removed under reduced pressure and the resulting yellow solids were suspended in 15 mL triethylamine and 15 mL diethyl ether and passed through a glass frit. After the solvents were removed under reduced pressure, the product was purified by silica gel chromatography under a nitrogen atmosphere with 15% triethylamine in diethyl ether as eluent. TBS-BN-gramine **2** was concentrated under reduced pressure to yield an opaque, light yellow oil (2.15 g, 7.43 mmol, 86.3% yield). $^1$H NMR (500 MHz, $CD_2Cl_2$) δ 8.01 (d, $J$ = 6.5 Hz, 1H), 7.57 (dd, $J$ = 11.4, 6.5 Hz, 1H), 6.76 (d, $J$ = 11.5 Hz, 1H), 6.62 (s, 1H), 6.38 (t, $J$ = 7.0 Hz, 1H), 3.52 (s, 2H), 2.22 (s, 6H), 0.91 (s, 9H), 0.47 (s, 6H). $^{13}$C NMR (126 MHz, $CD_2Cl_2$) δ 138.69, 128.51, 123.57, 123.56, 123.40, 108.47, 55.08, 45.38, 26.51, 19.16, -4.13 (B-C signal not observed). $^{11}$B NMR (160 MHz, $CD_2Cl_2$) δ 26.5. IR (ATR) 2950, 2928, 2855, 2813, 2763, 1618, 1587, 1499, 1426, 1347, 1310, 1173, 1162, 1014, 947, 834, 806, 781, 728, 698, 678, 429 cm$^{-1}$. HRMS (DART+) calcd. for $C_{15}H_{28}B_1N_3Si_1$ $[M+H]^+$ 289.21455, found 289.21513.

**Synthesis of 3:**



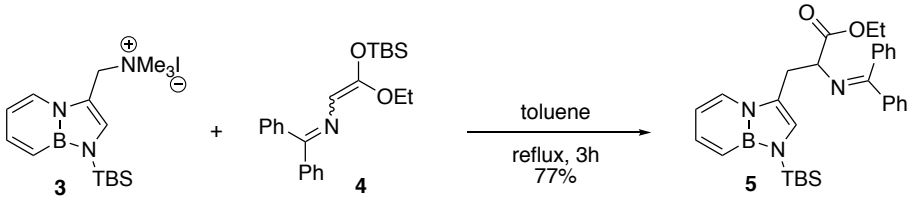

In a glove box, to a 250 mL RBF equipped with a stir bar was added TBS-BN-gramine **2** (1.70 g, 5.88 mmol) and THF (75 mL). The flask was cooled to 0 °C in an ice bath under nitrogen and iodomethane (1.00 g, 7.05 mmol, 439 μL) was added dropwise via syringe. The reaction was allowed to warm to 25 °C for 16 hours. The solids were allowed to settle in the bottom of the flask and the supernatent was removed via cannula under nitrogen pressure. When most of the solvent was removed, the solids were resuspended in $CH_2Cl_2$ and precipitated out with THF and the supernatent was removed by cannula. The residual solvent was removed under reduced pressure. TBS-BN-gramine methiodide salt **3** was obtained as a white, powdery solid (2.45 g, 5.68 mmol, 96.7% yield). $^1$H NMR (500 MHz, DMSO-$d_6$) δ 8.34 (d, $J$ = 6.7 Hz, 1H), 7.57 (dd, $J$ = 11.5, 6.5 Hz, 1H), 7.28 (s, 1H), 6.76 (d, $J$ = 11.5 Hz, 1H), 6.50 (t, $J$ = 6.6 Hz, 1H), 4.83 (s, 2H), 3.06 (s, 9H), 0.86 (s, 9H), 0.47 (s, 6H). $^{13}$C NMR (151 MHz, dmso) δ 138.77, 130.74, 126.81, 114.32, 108.88, 58.26, 51.32, 25.72, 17.99, -4.82 (B-C signal not observed). $^{11}$B NMR (160 MHz, DMSO-$d_6$) δ 26.2. IR (ATR) 3000, 2950, 2928, 2854, 1605, 1578, 1500, 1471, 1428, 1364, 1324, 1296, 1252, 1154, 1119, 999, 973, 940, 838, 807, 787, 728, 700, 592, 430 cm$^{-1}$.

**Synthesis of 4:**




To a 1 L RBF equipped with a stir bar, diisopropyl amine (3.15 g, 31.1 mmol, 4.39 mL) and THF (300 mL) were added and cooled to –78 °C in a dry ice/acetone bath under an atmosphere of nitrogen. A solution of *n*-butyllithium (2.5 M in hexane, 28.7 mmol, 11.5 mL) was added dropwise via syringe and the reaction was allowed to stir for 30 min. N-(Diphenylmethylene)glycine ethyl ester **S1** (6.40 g, 24.0 mmol) was dissolved in THF (10 mL) and added to the reaction mixture dropwise via syringe. The resulting red solution was allowed to stir for 100 min at –78 °C, after which time a solution of tert-Butyldimethylchlorosilane (5.05 g, 33.5 mmol) in THF (30 mL) was added dropwise via cannula. The reaction was allowed to stir and attain room temperature for 16 hours. Solvent was removed under reduced pressure and the crude mixture was suspended in hexane and passed through a glass frit. The filtrate was concentrated in vacuo. The desired product **4** was isolated as a red viscous oil (8.8 g, 96% yield). The product was used without further purification. $^1$H NMR (500 MHz, $CD_2Cl_2$) δ 7.64 – 7.59 (m, 1H), 7.48 (ddd, *J* = 7.8, 6.6, 2.9 Hz, 2H), 7.42 (dt, *J* = 8.5, 2.0 Hz, 2H), 7.29 (dd, *J* = 4.5, 2.3 Hz, 3H), 7.22 – 7.17 (m, 2H), 5.62 (d, *J* = 9.9 Hz, 1H), 4.45 (q, *J* = 7.1 Hz, 1H), 3.69 (q, *J* = 7.0 Hz, 2H), 1.40 (t, *J* = 7.1 Hz, 1H), 1.21 (t, *J* = 7.0 Hz, 2H), 1.05 (s, 6H), 0.89 (s, 3H), 0.31 (s, 4H), 0.09 (s, 2H). $^{13}$C NMR (126 MHz, $CD_2Cl_2$) δ 158.65, 155.83, 141.24, 141.07, 137.87, 137.66, 129.35, 129.30, 129.17, 129.15, 128.85, 128.63, 128.50, 128.40, 128.18, 127.97, 101.20, 98.19, 66.05, 64.06, 54.00, 26.10, 25.95, 18.95, 18.62, 15.86, 14.85, -3.90, -4.33. IR (ATR) 3059, 2953, 2958, 2884, 2855, 1624, 1470, 1444, 1376, 1315, 1294, 1240, 1070, 1017, 941, 826, 782, 695, 466 cm$^{-1}$. HRMS (DART+) calcd. for $C_{23}H_{32}N_1O_2Si_1$ 382.2202, found 382.2203.

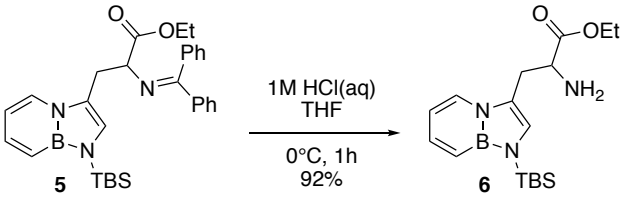
**Synthesis of 5:**




In a glove box, a 1 L RBF equipped with a stir bar was charged with TBS-BN-gramine methiodide salt **3** (2.65 g, 6.20 mmol), silylketeneacetal **4** (7.10 g, 18.6 mmol), and toluene (400 mL). The reaction was brought to reflux under nitrogen and allowed to heat at 140 °C. After 3 hours, the reaction was cooled to room temperature and concentrated under reduced pressure. The product was purified by silica gel chromatography with a gradient of ether in pentane (starting at 10% and then 40% ether) in the glove box to afford an orange goo (2.23 g, 76.9 %). $R_f$ (2:3 ether:pentane) = 0.4 $^1$H NMR (500 MHz, $CD_2Cl_2$) δ 7.60 – 7.55 (m, 2H), 7.51 (dd, *J* = 11.5, 6.5 Hz, 1H), 7.42 – 7.25 (m, 5H), 7.15 (t, *J* = 7.5 Hz, 2H), 6.76 – 6.72 (m, 1H), 6.58 (s, 2H), 6.51 (s, 1H), 6.18 (t, *J* = 6.5 Hz, 1H), 4.31 (ddd, *J* = 9.1, 4.2, 1.0 Hz, 1H), 4.19 (qdd, *J* = 7.1, 4.5, 1.0 Hz, 2H), 3.40 (dd, *J* = 14.9, 4.2 Hz,

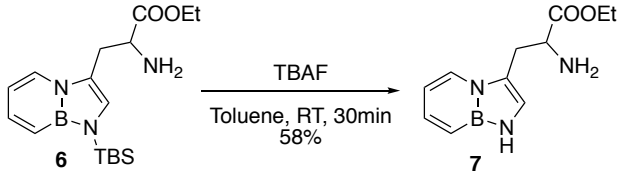
1H), 3.26 (dd, $J$ = 15.0, 9.1 Hz, 1H), 1.27 (t, $J$ = 7.1 Hz, 3H), 0.81 (s, 9H), 0.37 (d, $J$ = 15.3 Hz, 6H). $^{13}$C NMR (126 MHz, $CD_2Cl_2$) δ 172.08, 171.38, 139.81, 138.40, 136.30, 130.87, 129.27, 128.76, 128.48, 127.85, 126.55, 123.15, 121.78, 108.69, 64.51, 61.61, 29.25, 26.40, 18.99, 14.60, -4.10 (B-C signal not observed). $^{11}$B NMR (160 MHz, $CD_2Cl_2$) δ 25.75. IR (ATR) 3063, 2953, 2927, 2855, 1735, 1617, 1586, 1499, 1470, 1444, 1427, 1314, 1288, 1252, 1175, 1137, 1118, 1030, 945, 834, 806, 780, 728, 694, 428 cm$^{-1}$. HRMS (DART+) calcd. for $C_{30}H_{39}B_1N_3O_2Si_1$ $[M+H]^+$ 512.2905, found 512.2921.

**Synthesis of 6:**




To a 250 mL round bottom flask equipped with a stir bar was added a solution of **5** (1.4 g, 2.7 mmol) in THF (50 mL) and cooled to 0 °C under nitrogen in an ice bath. An aqueous solution of HCl (20 mL, 1M) was added dropwise via syringe and the solution was stirred at 0 °C for 30 minutes. The solution was diluted with 100 mL $CH_2Cl_2$ and washed thrice with 200 mL saturated sodium bicarbonate. The organic layer was separated and the aqueous layer was extracted twice more with 100 mL $CH_2Cl_2$. The combined organic layer was washed with brine and dried over $MgSO_4$. The solvent was removed under reduced pressure. The crude reaction mixture was purified by silica gel column, first eluted with $CH_2Cl_2$ and then 20% $Et_3N$ in $CH_2Cl_2$, to yield **6** (0.867 g, 92%) as a brown oil. $R_f$ (100 % ether) = 0.3 $^1$H NMR (500 MHz, $CD_2Cl_2$) δ 7.81 (d, $J$ = 6.6 Hz, 1H), 7.57 (dd, $J$ = 11.5, 6.5 Hz, 1H), 6.77 (d, $J$ = 12.4 Hz, 1H), 6.60 (s, 1H), 6.41 (t, $J$ = 6.6 Hz, 1H), 4.17 – 4.09 (m, 2H), 3.77 (dd, $J$ = 7.6, 5.5 Hz, 1H), 3.22 (dd, $J$ = 15.1, 5.5 Hz, 1H), 2.97 (dd, $J$ = 15.1, 7.7 Hz, 1H), 1.52 (s, 2H), 1.22 (t, $J$ = 7.1 Hz, 3H), 0.90 (s, 9H), 0.46 (s, 6H). $^{13}$C NMR (126 MHz, $CD_2Cl_2$) δ 175.53, 138.55, 129.55, 126.76, 122.67, 121.87, 108.88, 61.46, 31.34, 26.51, 19.16, 14.54, -4.10 (B-C signal not observed). $^{11}$B NMR (160 MHz, $CD_2Cl_2$) δ 26.0. IR (ATR) 2953, 2928, 2856, 1733, 1615, 1586, 1450, 1470, 1443, 1427, 1309, 1283, 1251, 1184, 1139, 1119, 1027, 945, 835, 820, 807, 782, 729, 699, 677, 428 cm$^{-1}$. HRMS (DART+) calcd. for $C_{17}H_{31}B_1N_3O_2Si_1$ $[M+H]^+$ 348.22786, found 348.22637.

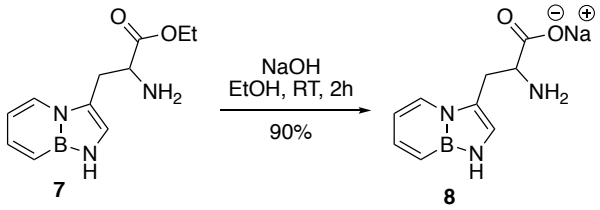
**Synthesis of 7:**




A 250 mL RBF was charged with **6** (473 mg, 1.36 mmol) and toluene (50 mL) with a stir bar and put under a nitrogen atmosphere. While stirring, tetrabutylammonium fluoride (1 M in THF, 1.36 mL) was added dropwise via syringe. After 30 minutes, the solvent was removed under reduced pressure. The resulting brown oil was suspended in 150 mL toluene. The organic layer was washed four times with 200 mL water followed by one wash with brine. The suspension was passed through filter paper and the solvent was removed under reduced pressure to afford compound 7 as a brown oil (185 mg, 58%). $^1$H

NMR (500 MHz, $CD_2Cl_2$) δ 7.81 (d, $J$ = 6.6 Hz, 1H), 7.57 (dd, $J$ = 11.4, 6.5 Hz, 1H), 6.82 – 6.50 (m, 3H), 6.41 (td, $J$ = 6.6, 1.1 Hz, 1H), 4.15 (q, $J$ = 7.1 Hz, 2H), 3.76 (dd, $J$ = 7.9, 5.2 Hz, 1H), 3.27 – 3.19 (m, 1H), 2.96 (dd, $J$ = 15.1, 7.9 Hz, 1H), 1.86 (s, 2H), 1.23 (t, $J$ = 7.1 Hz, 3H). $^{13}$C NMR (126 MHz, $CD_2Cl_2$) δ 175.10, 138.12, 126.03, 118.62, 118.15, 108.45, 61.06, 53.35, 30.71, 14.05 (B-C signal not observed). $^{11}$B NMR (160 MHz, $CD_2Cl_2$) δ 23.44. IR (ATR) 3360, 2978, 2933, 1725, 1620, 1584, 1497, 1447, 1372, 1347, 1306, 1264, 1188, 1095, 1077, 1023, 857, 802, 730, 694, 586, 549, 519 cm$^{-1}$. HRMS (DART+) calcd. for $C_{11}H_{17}B_1N_3O_2Si_1$ $[M+H]^+$ 234.1414, found 234.1405.

**Synthesis of 8:**




In a glove box, to a 20 mL vial was added a stir bar, **7** (284 mg, 1.22 mmol), sodium hydroxide (97.6 mg, 2.44 mmol), and ethanol (3.00 mL) and stirred for 2 hours. Most of the ethanol was removed under reduced pressure. 1 mL of ether was added to the solution, followed by 2 mL pentane, after which a brown and white solid precipitated. The solid was triterated twice in pentane and ether, solution removed by pipette in between triterations. BN-tryptophan **8** was obtained as a powdery tan solid (247 mg, 89.2%). $^1$H NMR (500 MHz, DMSO-$d_6$) δ 8.34 (s, 1H), 7.91 (d, $J$ = 5.9 Hz, 1H), 7.42 (dd, $J$ = 10.8, 6.5 Hz, 1H), 6.73 (s, 1H), 6.61 (d, $J$ = 11.3 Hz, 1H), 6.30 (t, $J$ = 5.5 Hz, 1H), 3.24 (d, $J$ = 13.4 Hz, 2H), 2.66 – 2.53 (m, 1H), 1.49 (s, 2H). $^{13}$C NMR (126 MHz, DMSO-$d_6$) δ 178.81, 136.84, 126.46, 120.40, 117.85, 107.47, 54.37, 32.02 (B-C signal not observed). $^{11}$B NMR (160 MHz, DMSO-$d_6$) δ 23.22. IR (ATR) 3445, 3370, 3350, 1561, 1495, 1467, 1449, 1406, 1346, 1326, 1306, 1074, 1051, 927, 878, 860, 728, 692, 608, 557, 541 cm$^{-1}$. HRMS (DART+) calcd. for $C_9H_{13}B_1N_3O_2$ 206.11008, found 206.11063.

***Chiral resolution of BN-tryptophan ester 7***

Enantiomers of ester **7** were separated by chiral recycling HPLC (Chiralpak AD-H, 25°C, 78:19.5:2.0:0.5 Hexane:Isopropanol:Methanol:*n*-butylamine, 5 mL/min). The product was cycled through 4-6 times before collection of each peak.

Collected fractions were analyzed by chiral analytical HPLC (Chiralpak AD-H, 25 °C, 78:19.5:2.0:0.5 Hexane:Isopropanol: Methanol:*n*-butylamine, 0.5 mL/min).

**Figure S1:** HPLC trace of racemic BN-tryptophan ester **7**.

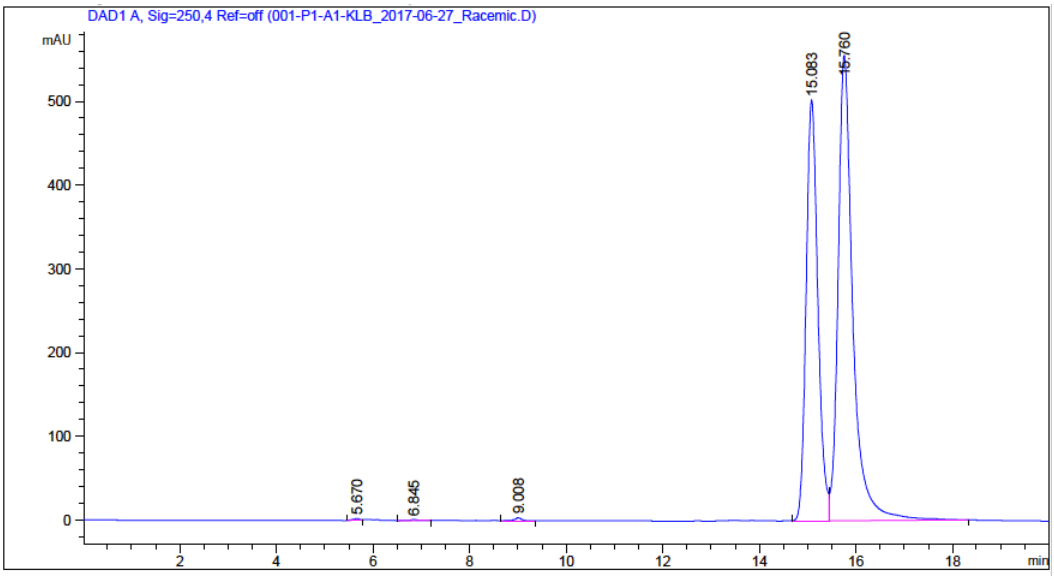



**Figure S2:** HPLC trace of (D)-BN-Tryptophan ester **7**.

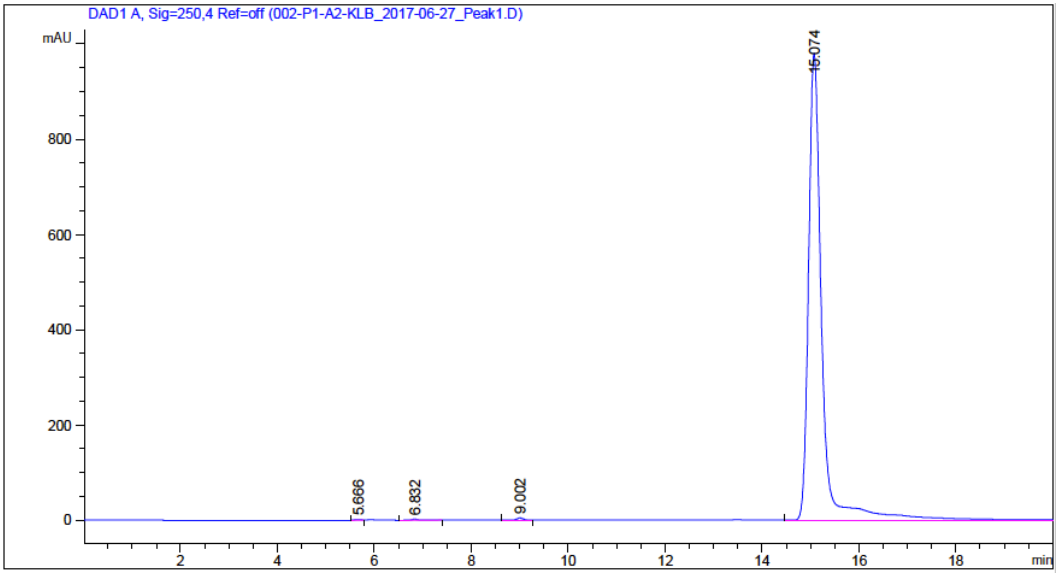
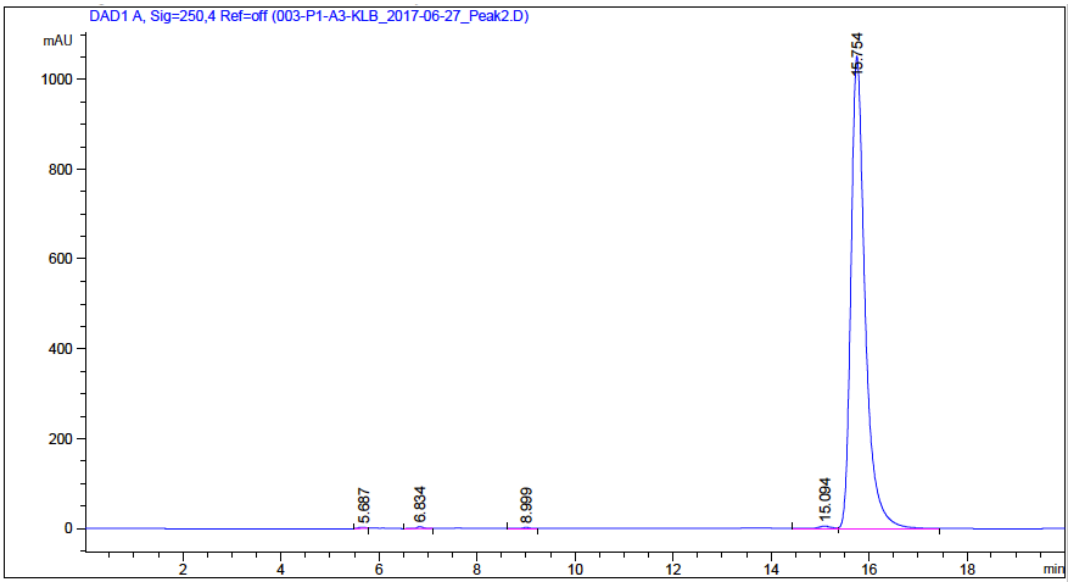
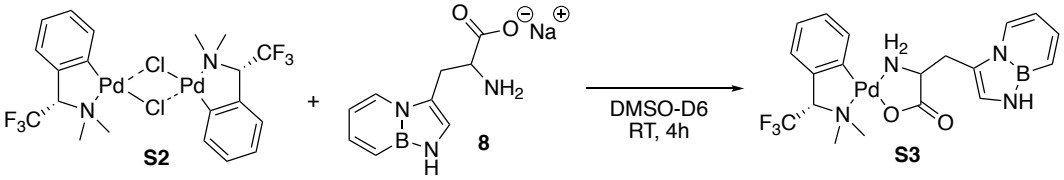


**Figure S3:** HPLC trace of (L)-BN-tryptophan ester **7**.



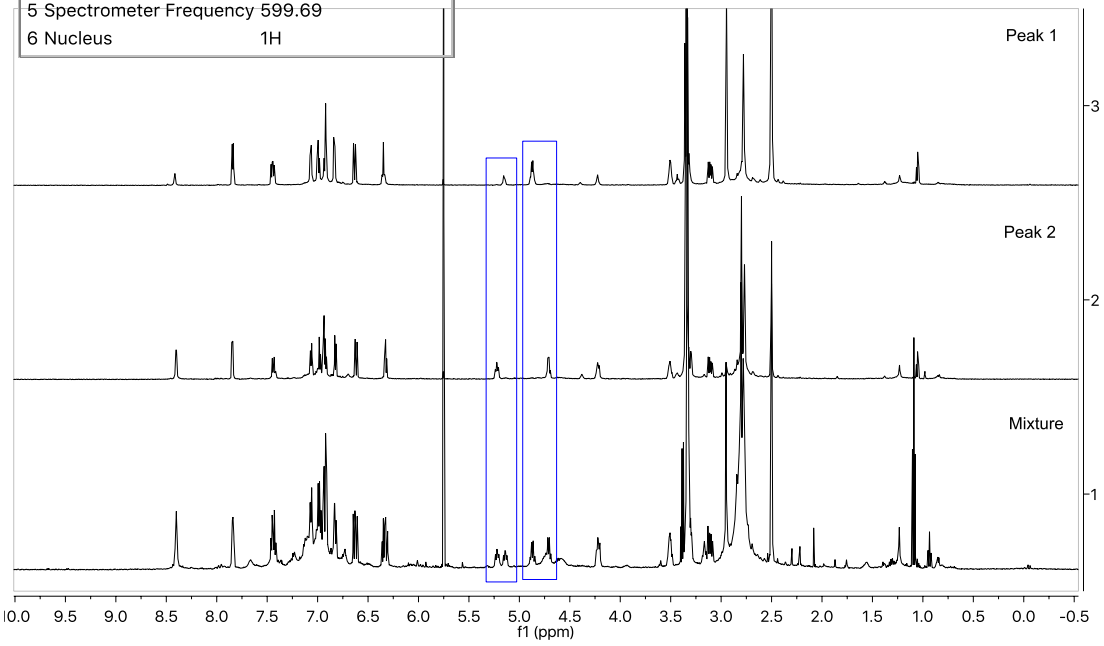
***Chiral analysis of BN-tryptophan (D)-8 and (L)-8***

*NMR Chiral Analysis*




A 4mL vial was charged with Pd-complex[2] **S2** (7.7 mg, 0.011 mmol), BN-tryptophan **8** (5.0 mg, 0.022 mmol), and a stir bar. DMSO-D6 (0.5 mL) was added and the solution was allowed to stir under nitrogen for 4 hours, after which the NMR spectra were collected. The NMR showed quantitative formation of **S3**. The spectra of the resolved compounds, **(D)-8** and **(L)-8** are as shown in Figure **S4**.

| Parameter | Value |
|---|---|
| 1 Solvent | dmso |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2017-06-19T18:05:52 |
| 5 Spectrometer Frequency | 599.69 |
| 6 Nucleus | 1H |



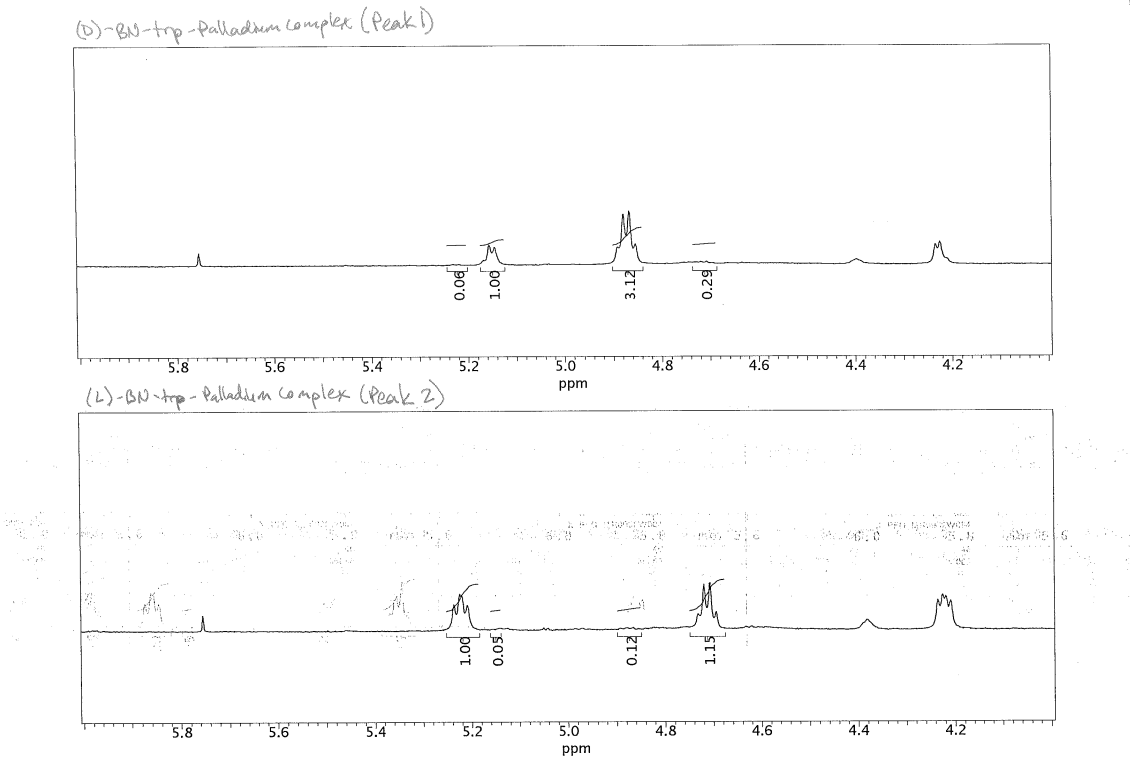



**Figure S4:** NMR analysis of enantiopurity of resolved BN-trp **8**.

*e.e. calculation:*

d.e. Peak 1 complex (= e.e. (D)-BN-trp **8**) = (1.00-0.06)/(1.00+0.06)= 0.8868x100% = 90%
d.e. Peak 2 complex (= e.e. (L)-BN-trp **8**) = (1.00-0.05)/(1.00+0.05)= 0.9048x100% = 90%

*Specific Rotation*

**(D)-BN-tryptophan 8** $[\alpha]^{25}_D = -7.600$ (c = 0.50, $H_2O$, *l* = 0.5 dm)

**(L)-BN-tryptophan 8** $[\alpha]^{25}_D = +7.600$ (c = 0.50, $H_2O$, *l* = 0.5 dm)

**(D)-tryptophan 9** $[\alpha]^{25}_D = -3.600$ (c = 0.50, $H_2O$, *l* = 0.5 dm)

**(L)-tryptophan 9** $[\alpha]^{25}_D = +3.200$ (c = 0.50, $H_2O$, *l* = 0.5 dm)

*CD-spectroscopy*

A 0.1 mm quartz cuvette was used for the measurements. Methanol was used as solvent and concentrations of 0.1 mg/mL were used. Step size was 0.5 nm and averaging time was 6 seconds. Methanol was used as the blank and the CD spectra were corrected according to it (baseline subtraction). Spectra are an average of 3 scans.

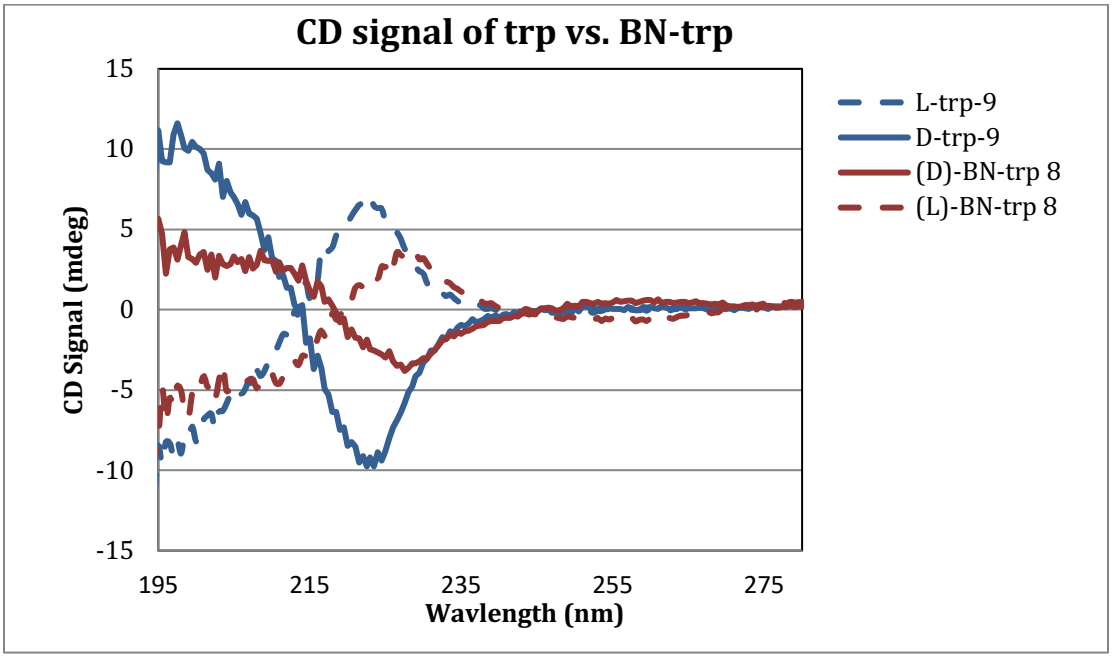



**Figure S5:** CD-signal of tryptophan **9** compared to BN-trp **8**.

***Quantum Yield Measurement of (L)-BN-Tryptophan 8 and (L)-Tryptophan 9***

A dilute solution of each (L)-BN-Trp **8** and (L)-Trp **9** was prepared in water in a 1 cm quartz cuvette. Each solution was diluted to have an absorbance intensity between 0.5 and 0.7 a.u. The same solution was then used to measure the quantum yield with the "petite" integrating sphere.

For (L)-BN-Trp **8**:

Settings:
293:265-570 nm, excitation slit width = 3.0 nm, emission slit width: 5.0 nm

Results:
1) Abs = 0.493, $\Phi = 0.297$
2) Abs = 0.528, $\Phi = 0.288$
3) Abs = 0.643, $\Phi = 0.307$
4) Abs = 0.605, $\Phi = 0.286$
5) Abs = 0.558, $\Phi = 0.275$
Average: $\Phi = 0.291 \pm 0.011$

For (L)-Trp **9**:

Settings:
279:250-550 nm, excitation slit width = 3.0 nm, emission slit width: 5.0 nm

Results:
1) Abs = 0.567, $\Phi = 0.251$
2) Abs = 0.628, $\Phi = 0.248$
3) Abs = 0.577, $\Phi = 0.247$
Average: $\Phi = 0.249 \pm 0.0017$

***Solvatochromism of BN-tryptophan ester 7 and L-tryptophan ester S1***

All solvents were degassed through at least 3 freeze-pump-thaw cycles and solutions were prepared in a nitrogen filled dry box.

A dilute solution (~1-1.4x10$^{-4}$ M) of each BN-Trp ester **7** or (L)-Trp ester **S1** was prepared in the specified solvent in a 1 cm quartz cuvettes and the abosorbance and fluorescence were measured.

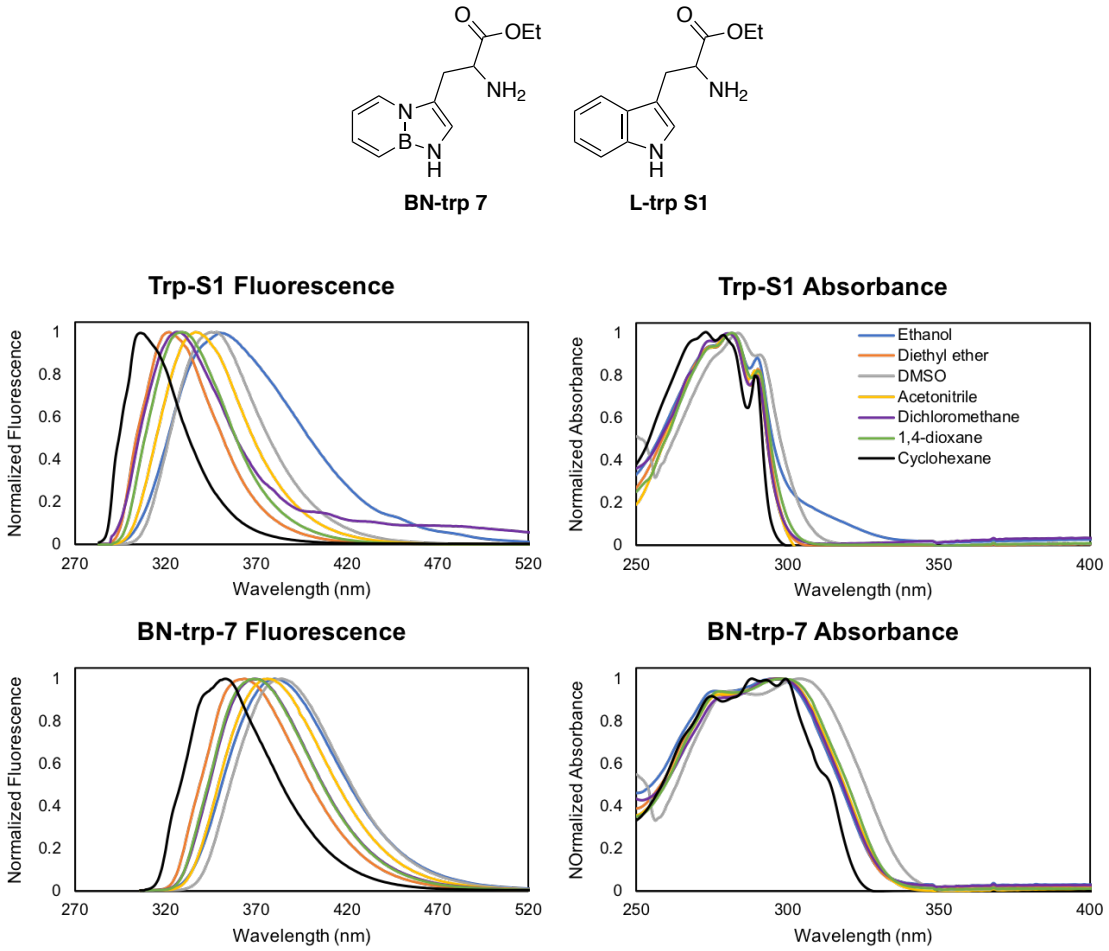



**Figure S6:** Absorbance and Emission spectra of Trp-**S1** and BN-trp **7**.

*NMR Assignment of BN-tryptophan 8*

**Figure S7:** gCOSY spectrum of BN-trp **8**.

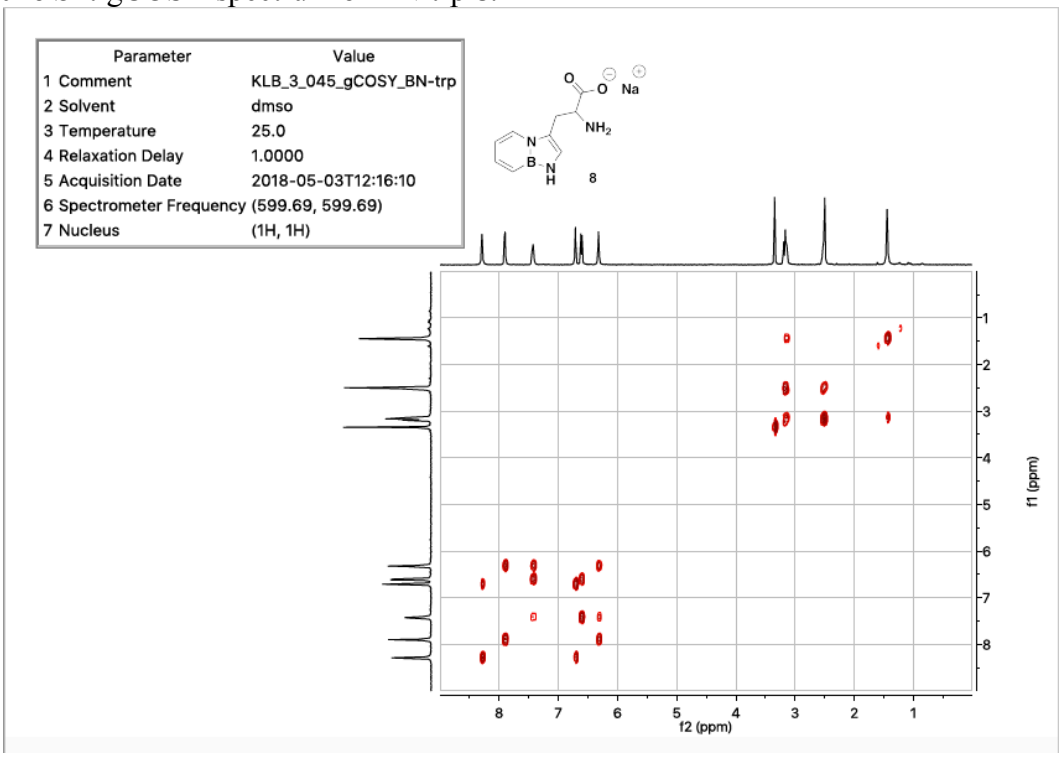



**Figure S8:** gHSQC spectrum of BN-trp **8**.

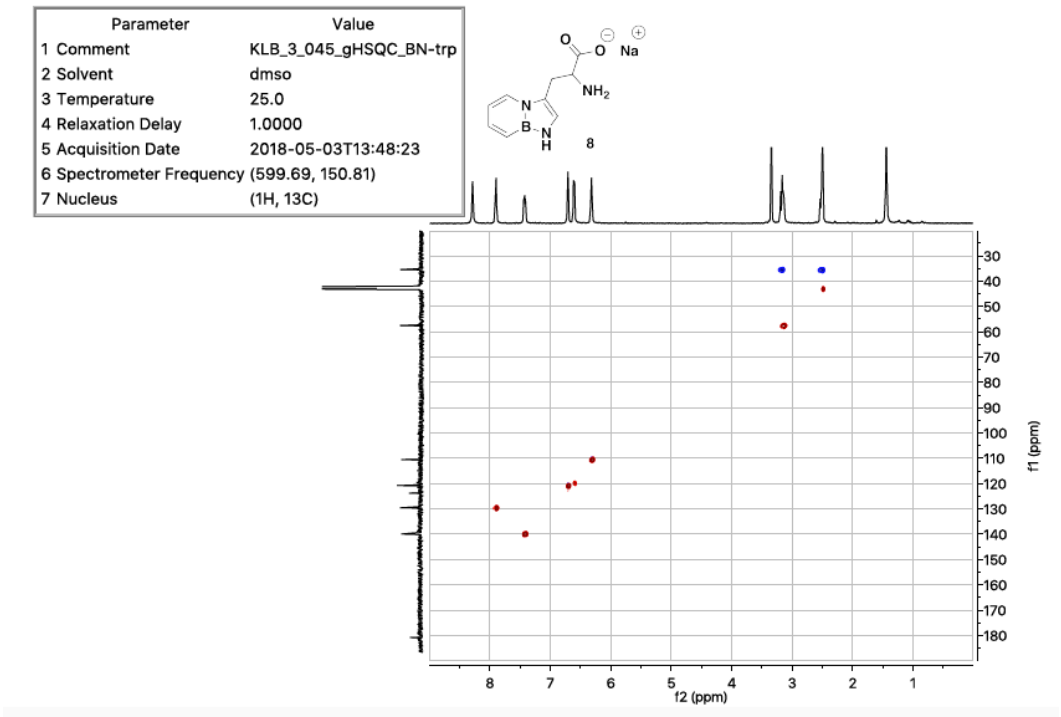


**Figure S9:** $^{1}$H NMR spectrum of BN-trp **8** with proton assignments.

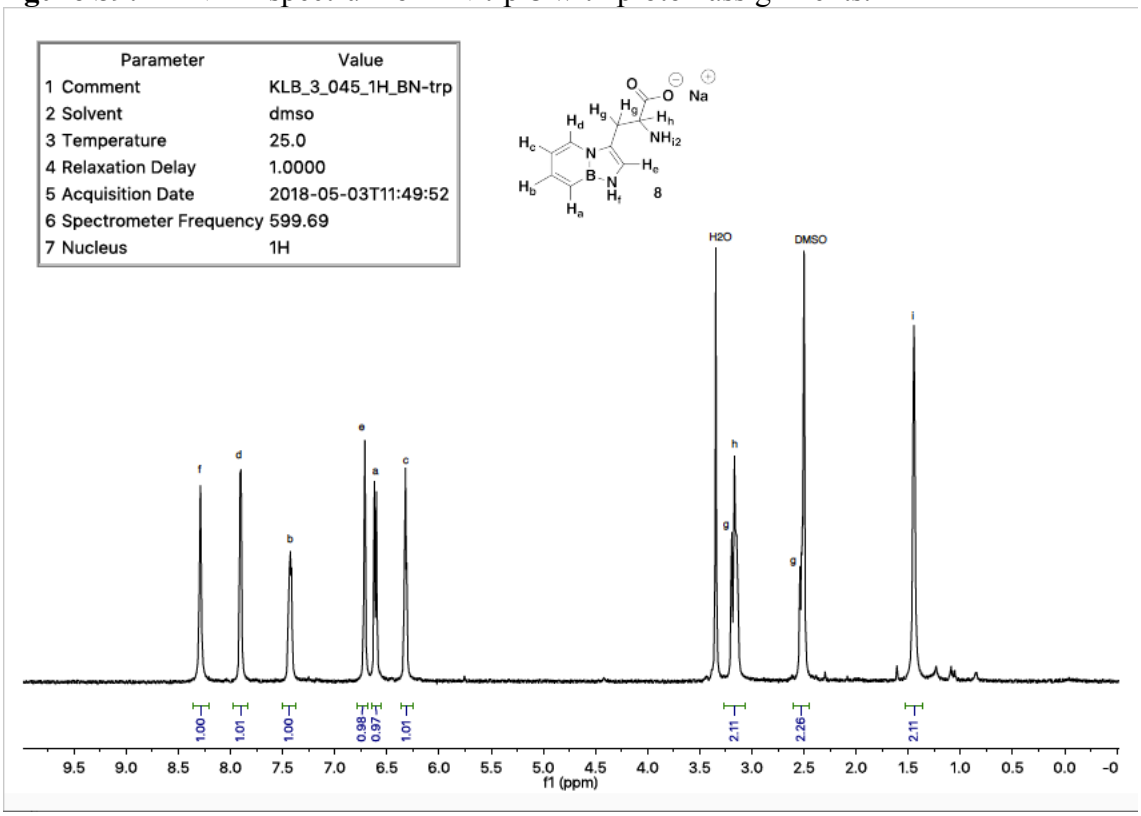
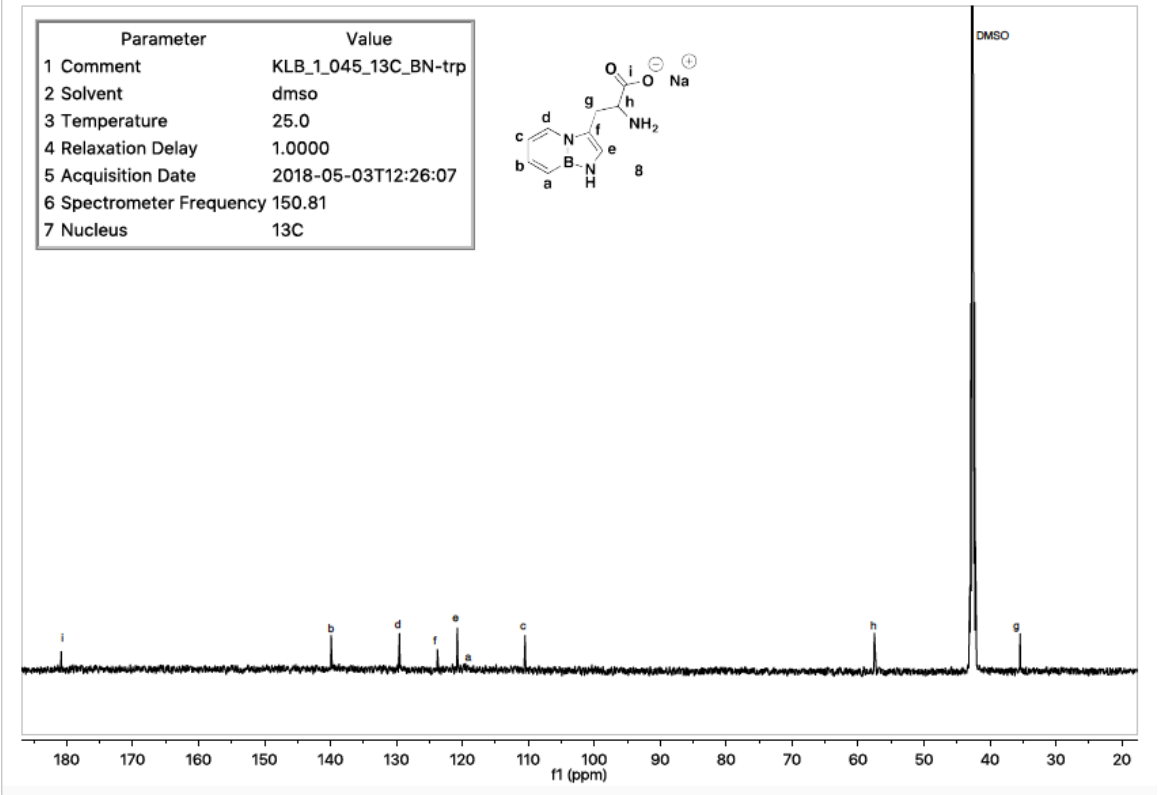



**Figure S10:** $^{13}$C NMR spectrum of BN-trp **8** with carbon assignments.



***Cell Fermentation and Protein Isolation***

Trp-auxotrophic *Escherichia coli* strain was purchased from ATCC (catalog number 49980 [genotype *WP2 uvrA*]). All fermentation and expression experiments were performed in GMML supplemented with 200 μM glucose (minimal media) (see recipe below). The expression host *E. coli* ATCC49980 was routinely transformed with either plasmid: pET22b-sfGFPwt, pET22b-sfGFP151TGG (subcloned via directed mutagenesis from pET22b-sfGFPwt using standard cloning methods, sequence below), or pET22b-Llac-KSI-7TGA.[3]

The incorporation experiments were performed using cultures grown in minimal media in the presence of 0.0075 mM tryptophan (the optimal limiting concentration of the native substrate at 30 °C) and 2 mM Ampicillin (in the transformed cell cultures). Cell cultures were grown to ~0.5 $OD_{600}$, at which time tryptophan was depleted and the desired tryptophan analogue (1 or 2 mM) was added along with 1 mM IPTG. End-point optical densities and fluorescence were measured after incubation at 30 °C overnight.

Overnight expression cultures were purified as described in Italia et al.[3]

*GMML Recipe:*

## Minimal Media Recipe

Glycerol based minimal media supplemented with leucine (GMML)

For every liter culture, make:
- 12.8 g Na2HPO4•7H2O (or 6 g Na2HPO4)
- 3 g KH2PO4
- 0.5 g NaCl
- 1 g NH4Cl

Bring to approximately 450 ml with **Tap H2O**, and adjust pH to 7.4, and then add **Tap H2O** to 0.5L
Autoclave the above solution
Also autoclave 10 ml glycerol in 500 ml **Tap H2O**.
**If you need to make agar plates, add 15 g agar into the above 500 mL glycerol solution before autoclaving.**

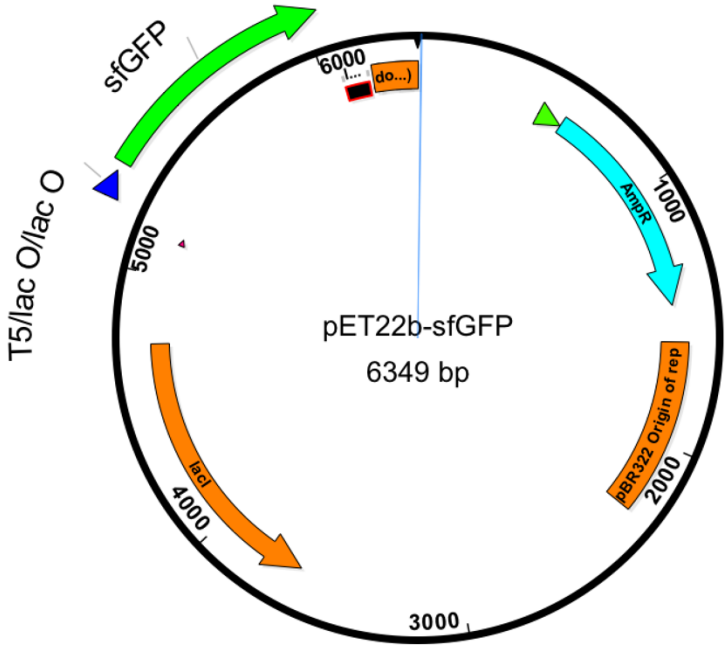
**GMML media are made by mixing the above two 500 ml solutions and the following ingredient.**

Make other stock solutions (~10 mL each), sterile them separately through 0.22 uM filters:
1. 300 mM L-leucine (1000x stock; adjust with minimal NaOH for solubilization)
2. 4 mM d-biotin (1000X stock)
3. 330 uM Thiamine (1000X stock)
4. 1 M MgSO4 (1000X stock)
5. 0.1 M CaCl2 (1000X stock)
6. Misc. heavy metal solution (1000X stock): To make 10 mL stock, add 10 mg of CuSO4·5H2O, 40 mg MnCl2·4H2O, 40 mg ZnCl2, 12 mg FeSO4·5H2O; store at -20C for stability

Above 6 components have to be added before use (cannot be added when the autoclaved solutions are hot).

3.6 g of glucose added per 1 L of media after autoclaving before a final sterile filtration.

*Plasmid Sequences:*




tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacactt
gccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctc
taaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatgg
ttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactc
ttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcggcct
attggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcaggtgg
cacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaga
caataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagtt
gggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgtttt
ccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtc
gccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagt
aagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccg
aaggagctaaccgctttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaag
ccataccaaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaaactattaactggcgaact
acttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcc
cttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggc
cagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagat
cgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgattta
aaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagt
tttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctgcgcgtaatctg
ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaag
gtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaact
ctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttac
cgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaa
aggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggta
tctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagccta
tggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgt
tatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcg
cagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacac
cgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgt
gactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccg

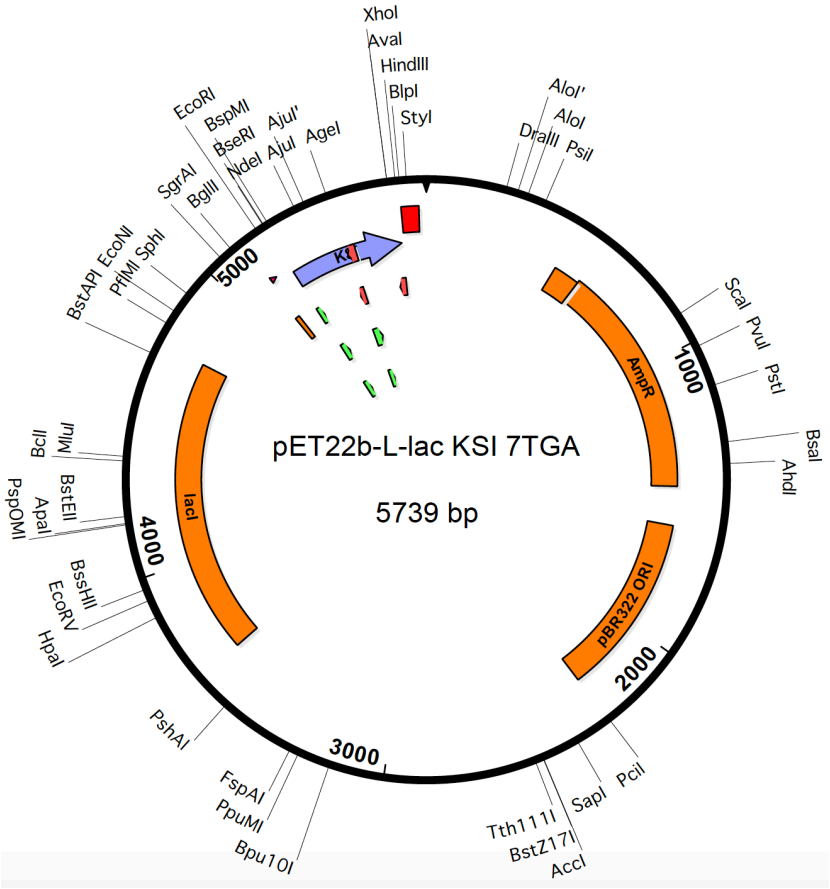
cttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggca
gctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagt
ttctccagaagcgttaatgtctggcttctgataaagcgggccatgttaagggcggttttttcctgtttggtcactgatg
cctccgtgtaaggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttac
tgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaa
atcactcagggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgc
agatccggaacataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaagaccattca
tgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaa
ccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggggccgccatgc
cggcgataatggcctgcttctcgccgaaacgtttggtggcgggaccagtgacgaaggcttgagcgagggcgtgcaagat
tccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggtcctcgccgaaaatgacccagagcgct
gccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccacc
ggaaggagctgactgggttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacatt
aattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcg
gggagaggcggtttgcgtattgggcgccagggtggttttctttttcaccagtgagacgggcaacagctgattgcccttc
accgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtgg
ttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagccc
ggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctca
ttcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgat
tgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgat
ttgctggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatg
ggtgtctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcat
ccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgcc
gcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaatttgcgac
ggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggt
tgggaatgtaattcagctccgccatcgccgcttccactttttcccgcgttttcgcagaaacgtggctggcctggttcac
cacgcgggaaacggtctgataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccacc
ctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcga
cgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaagga
atggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgct
catgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcg
ccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactatagg
ggaattgtgagcggataacaattcccctctagagtttgacagcattatcatcgatctcgagaaatcataaaaaatttat
ttgctttgtgagcggataacaattatataagattcaattgtgagcggataacaatttcacacagaattcattaaagagg
agaaattacatATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAA
TGGGCACAAATTTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACT
ACTGGAAAACTACCTGTTCCG**TGG**CCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGG
ATCACATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGA
TGACGGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGGGTATT
GATTTTAAAGAAGATGGAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTA**TAT**ATCACGGCAG
ACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCA
TTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGTCCTT
TCGAAAGATCCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGG
ATGAGCTCTACAAAGGATCCCACCACCACCACCACCACTAAaagcttaattagctgagcttggactcctgttgatagat
ccagtaatgacctcagaactccatctggatttgttcagaacgctcggttgccgccgggcgtttttattggtgagaatc
caagctagcttggcggcggccgcactcgagcaccaccaccaccaccactgagatccggctgctaacaaagcccgaaagg
aagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttt
tttgctgaaaggaggaactatatccggat

**pET22b-sfGFP-151TGG**

Same plasmid backbone sequence as above, but contains sfGFP Y151W (TAT->TGG) mutation.

ATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAAT
TTTCTGTCCGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACTACTGGAAAACT
ACCTGTTCCG**TGG**CCAACACTTGTCACTACTCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGGATCACATGAAA
CGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGACCT
ACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGGGTATTGATTTTAAAGA
AGATGGAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTA**TGG**ATCACGGCAGACAAACAAAAG
AATGGAATCAAAGCTAACTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCATTATCAACAAA
ATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGTCCTTTCGAAAGATCC
CAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGATGAGCTCTAC
AAAGGATCCCACCACCACCACCACCACTAA

**pET22b-Llac-KSI-7TGA**




tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacactt
gccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctc
taaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgattagggtgatgg
ttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactc
ttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattttgccgatttcggcct
attggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatattaacgtttacaatttcaggtgg
cacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgaga
caataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattc
ccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagtt
gggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgtttt
ccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtc
gccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagt
aagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccg
aaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaag
ccataccaaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaaactattaactggcgaact
acttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcc
cttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggc
cagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagat
cgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgattta
aaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagt
tttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctg
ctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaag
gtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaact
ctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttac
cgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagc
ttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaa

aggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggta
tctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagccta
tggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgt
tatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcg
cagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacac
cgcatatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgt
gactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccg
cttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggca
gctgcggtaaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgtccagctcgttgagt
ttctccagaagcgttaatgtctggcttctgataaagcgggccatgttaagggcggttttttcctgtttggtcactgatg
cctccgtgtaagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttac
tgatgatgaacatgcccggttactggaacgttgtgagggtaaacaactggcggtatggatgcggcgggaccagagaaaa
atcactcagggtcaatgccagcgcttcgttaatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgc
agatccggaacataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaagaccattca
tgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgctcgcgtatcggtgattcattctgctaa
ccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgcacccgtggggccgccatgc
cggcgataatggcctgcttctcgccgaaacgtttggtggcgggaccagtgacgaaggcttgagcgagggcgtgcaagat
tccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaaagcggtcctcgccgaaaatgacccagagcgct
gccggcacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgcccacc
ggaaggagctgactgggttgaaggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacatt
aattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcg
gggagaggcggtttgcgtattgggcgccagggtggtttttctttcaccagtgagacgggcaacagctgattgcccttc
accgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcgaaaatcctgtttgatggtgg
ttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatatccgcaccaacgcgcagccc
ggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctca
ttcagcatttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgat
tgcgagtgagatatttatgccagccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgat
ttgctggtgacccaatgcgaccagatgctccacgcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatg
ggtgtctggtcagagacatcaagaaataacgccggaacattagtgcaggcagcttccacagcaatggcatcctggtcat
ccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgctttacaggcttcgacgcc
gcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaatttgcgac
ggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggt
tgggaatgtaattcagctccgccatcgccgcttccactttttcccgcgttttcgcagaaacgtggctggcctggttcac
cacgcgggaaacggtctgataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattcaccacc
ctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcga
cgctctcccttatgcgactcctgcattaggaagcagcccagtagtaggttgaggccgttgagcaccgccgccgcaagga
atggtgcatgcaaggagatggcgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgct
catgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcg
ccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactatagg
ggaattgtgagcggataacaattgacattgtgagcggataacaagatactgagcacagaattcattaaagaggaggaat
tacatATGAACCTGCCGACCGCCtatGAAGTCCAGGGTCTGATGGCGCGTTTTATTGAACTGGTCGATGTGGGCGACAT
TGAAGCAATTGTGCAGATGTACGCGGACGATGCGACCGTTGAAAGCCCGTTTGGTCAACCACCGATTCATGGCCGTGAA
CAGATTGCGGCGCACTACCGTCAGTGGCTGGGCGGGGGTAAACTGCGTGTTTGCCTGACCGGTCCAGTGCGTGCTAGTc
atAACGGCTGCGGCGCGATGCCGTTGCGTAAAGAGTGGGTTTGGAATGGTCAGCCTTGCGCAACGGATGTTATTCTGGT
TATGCGCTTCGATGAACACGGTCGCATCCAGACCGAACAGCGCTATTGGAGCGAAGTGAATCTGAGTGTACGCGAACCG
CAGGGCAGTCTCGAGCACCACCACCACCACCACTAAgcttaattagctgagcaataactagcataacccctttggggcct
ctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccggat

***Oxidation Study***

Protein oxidation studies were performed on the L-trp and BN-trp containing sfGFP151TGG proteins. The isolated proteins were incubated in 1 mM $H_2O_2$ for 5 minutes and then quenched with 1 mM DTT (dithiolthreitol). Samples were analyzed by LC/MS (Agilent Technologies, 1260 Inifinity ESI–TOF) with a water/acetonitrile gradient (95% water to 5% water) on a C18 column.

The expected mass is 28620 g/mol, which is seen in Figure S11.

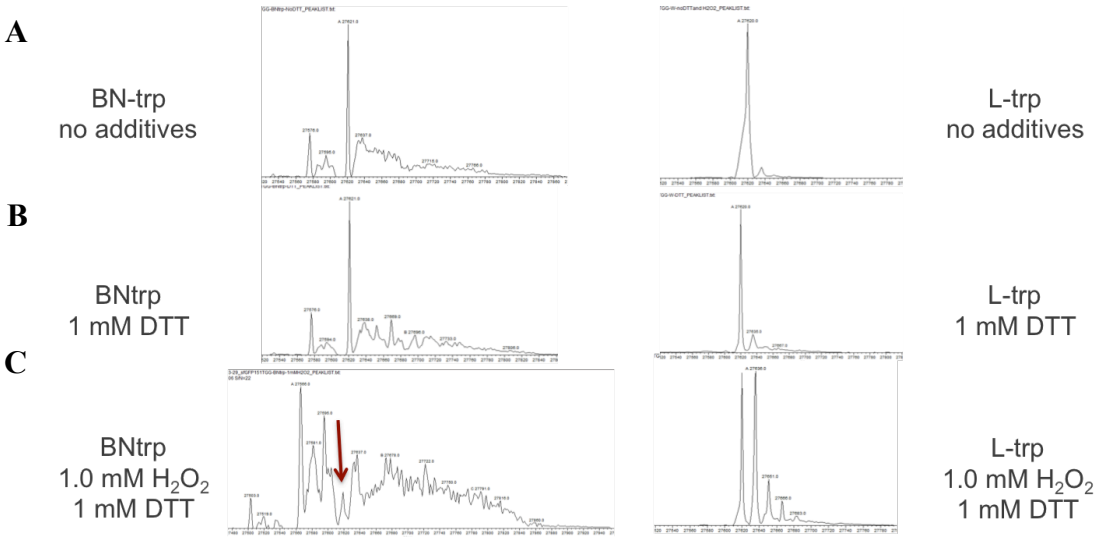
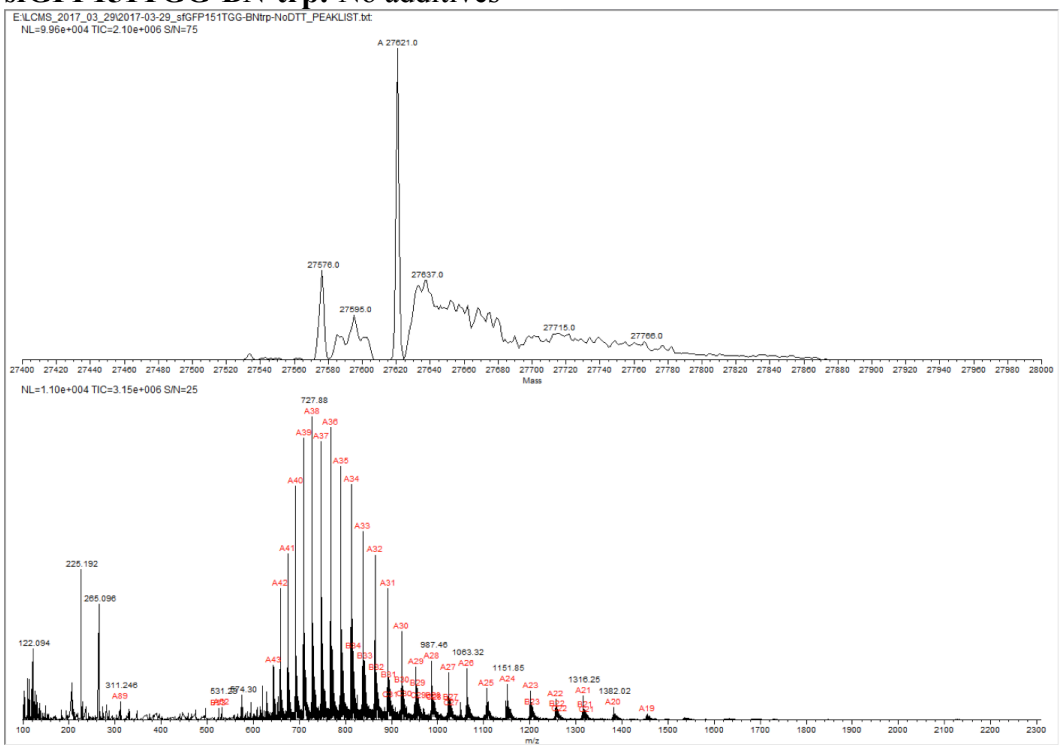


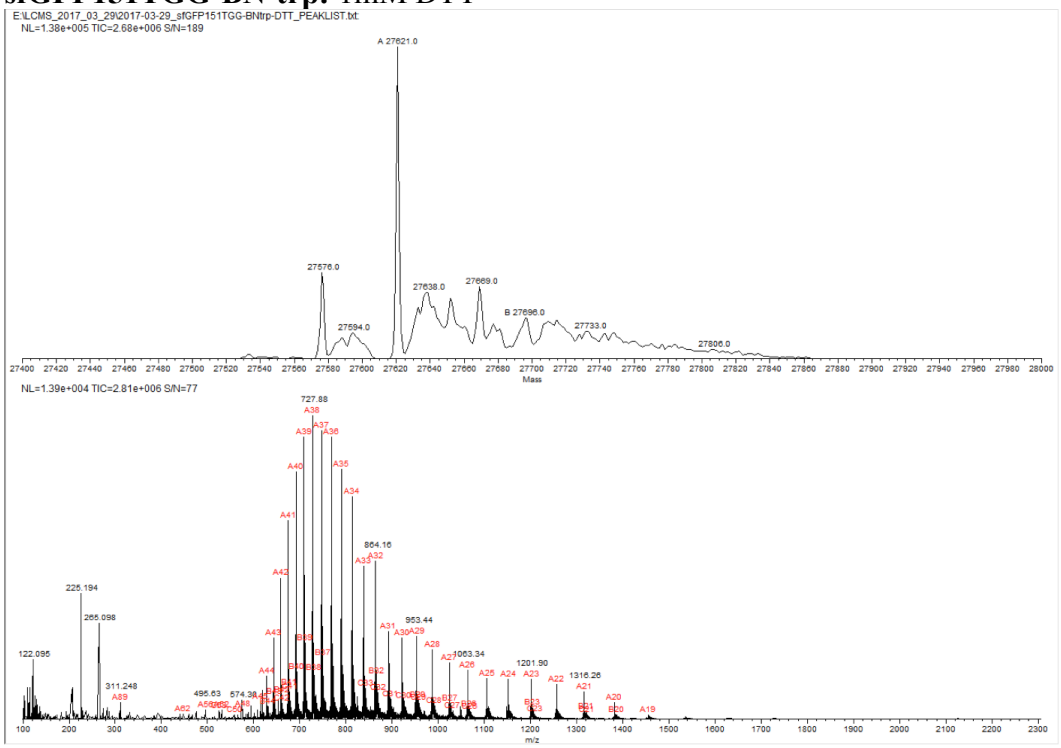

**Figure S11:** Mass profiles of Oxidation experiments. A) Control with no additives. B) Control with DTT. C) Oxidation with $H_2O_2$ and quench with DTT.

*LC/MS Spectra of Oxidation Study*

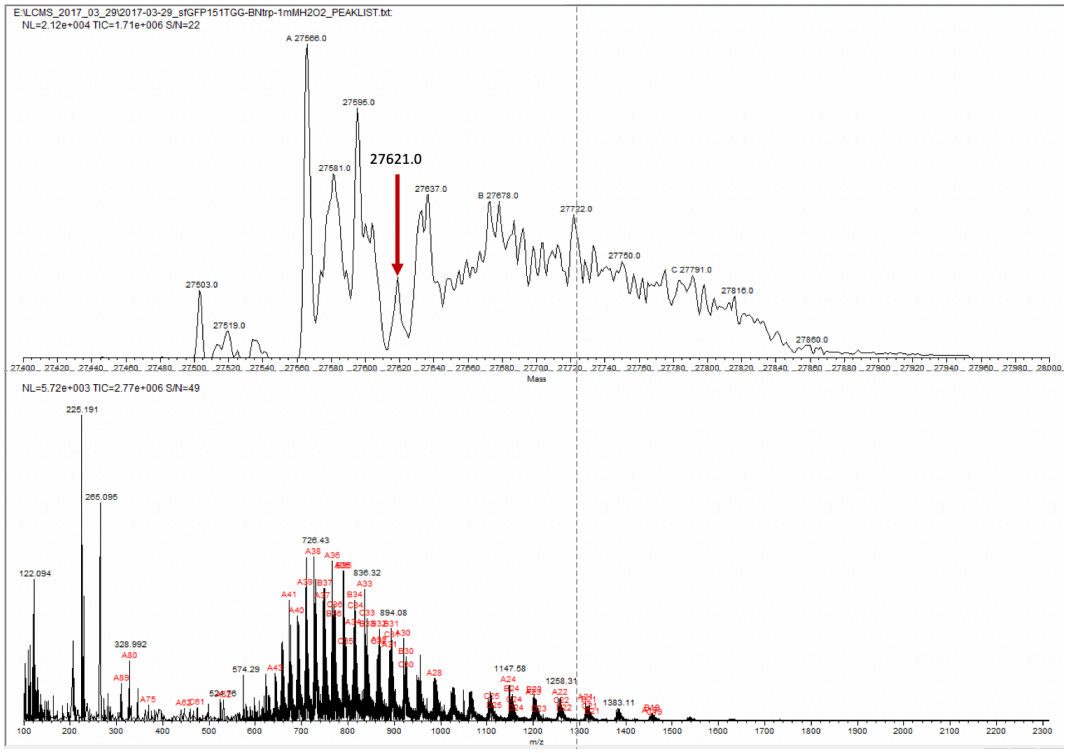
**sfGFP151TGG-BN-trp:** No additives



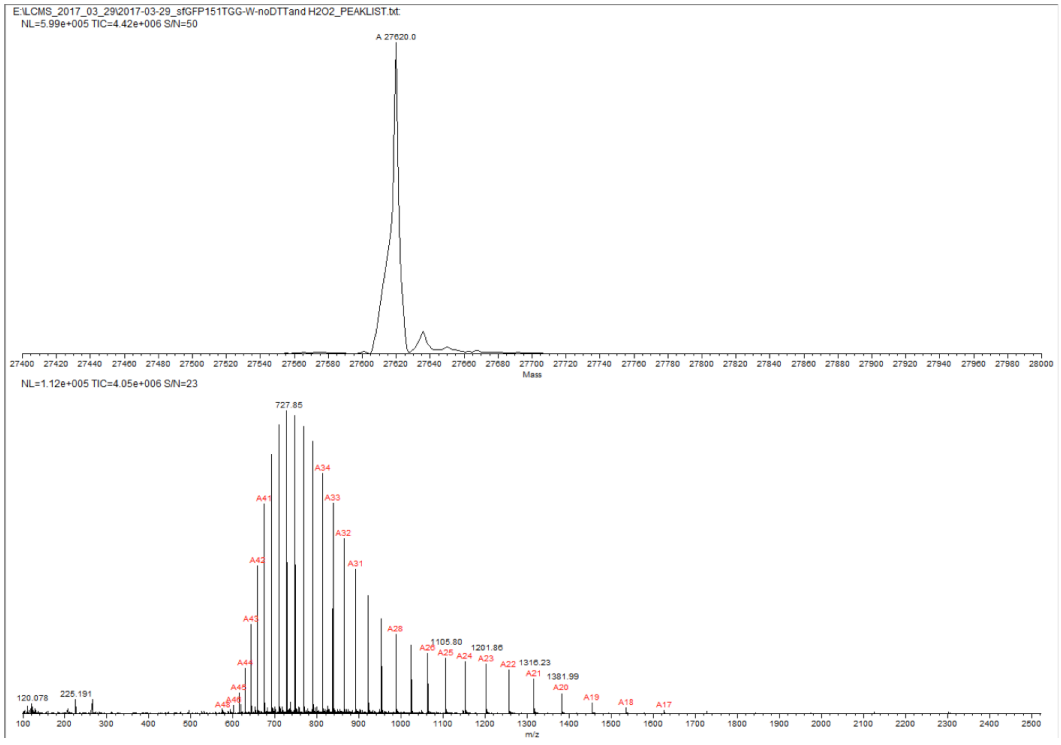

**sfGFP151TGG-BN-trp:** 1mM DTT



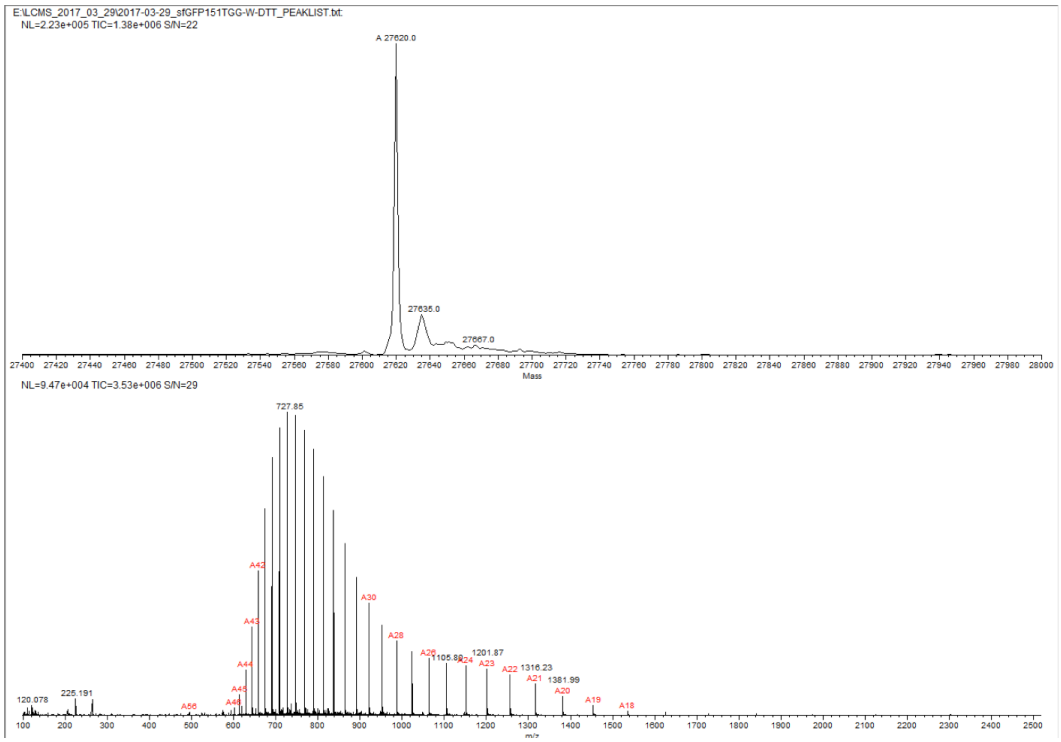
**sfGFP151TGG-BN-trp:** 1 mM $H_2O_2$ and 1mM DTT



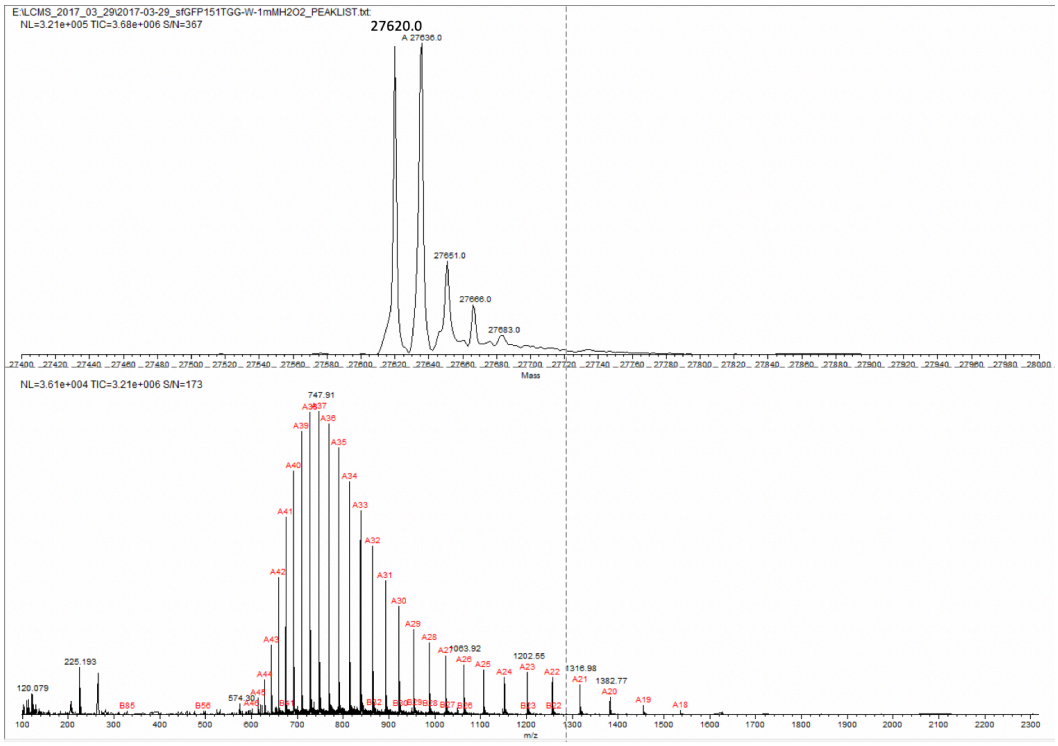

**sfGFP151TGG-L-trp:** No additives



**sfGFP151TGG-L-trp:** 1mM DTT




**sfGFP151TGG-L-trp:** 1 mM $H_2O_2$ and 1mM DTT




***X-Ray Crystallographic Data for BN-trp ester 6***

*Crystal Data and Structure Refinement*

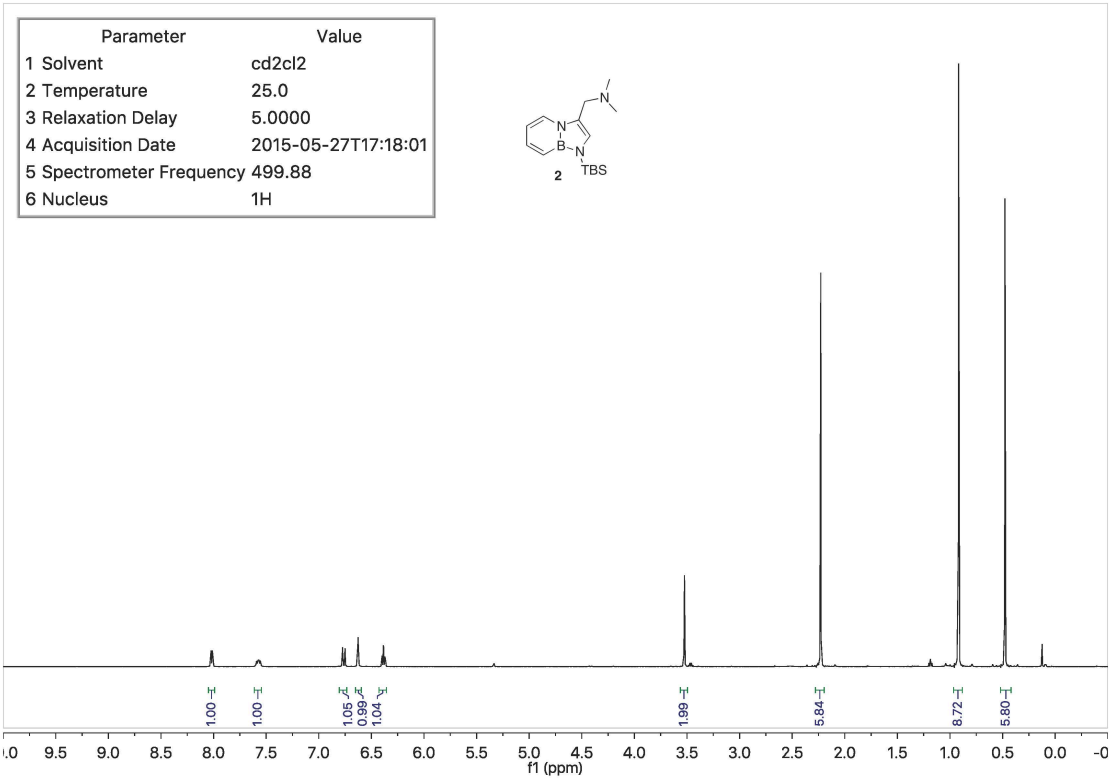
| | | |
|---|---|---|
| Identification code | (C17H31BN3O2Si)+(CF3O3S)-(H2O)0.5 | |
| Empirical formula | C18 H32 B F3 N3 O5.50 S Si | |
| Formula weight | 506.42 | |
| Temperature | 100(2) K | |
| Wavelength | 0.71073 Å | |
| Crystal system | Monoclinic | |
| Space group | $P2_1/n$ | |
| Unit cell dimensions | a = 15.3865(16) Å | α= 90°. |
| | b = 6.5658(7) Å | β= 90.703(3)°. |
| | c = 26.361(3) Å | γ = 90°. |
| Volume | 2662.9(5) $Å^3$ | |
| Z | 4 | |
| Density (calculated) | 1.263 $Mg/m^3$ | |
| Absorption coefficient | 0.220 $mm^{-1}$ | |
| F(000) | 1068 | |
| Crystal size | 0.300 x 0.130 x 0.050 $mm^3$ | |
| Theta range for data collection | 1.524 to 28.296°. | |
| Index ranges | -20<=h<=20, -8<=k<=8, -35<=l<=0 | |
| Reflections collected | 12647 | |
| Independent reflections | 6615 [R(int) = 0.0426] | |
| Completeness to theta = 25.242° | 99.9 % | |
| Absorption correction | Semi-empirical from equivalents | |
| Max. and min. transmission | 0.7457 and 0.6264 | |
| Refinement method | Full-matrix least-squares on $F^2$ | |
| Data / restraints / parameters | 6615 / 4 / 325 | |
| Goodness-of-fit on $F^2$ | 1.110 | |
| Final R indices [I>2sigma(I)] | R1 = 0.0532, wR2 = 0.1440 | |
| R indices (all data) | R1 = 0.0836, wR2 = 0.1562 | |
| Extinction coefficient | na | |
| Largest diff. peak and hole | 0.575 and -0.580 $e.Å^{-3}$ | |

***References:***

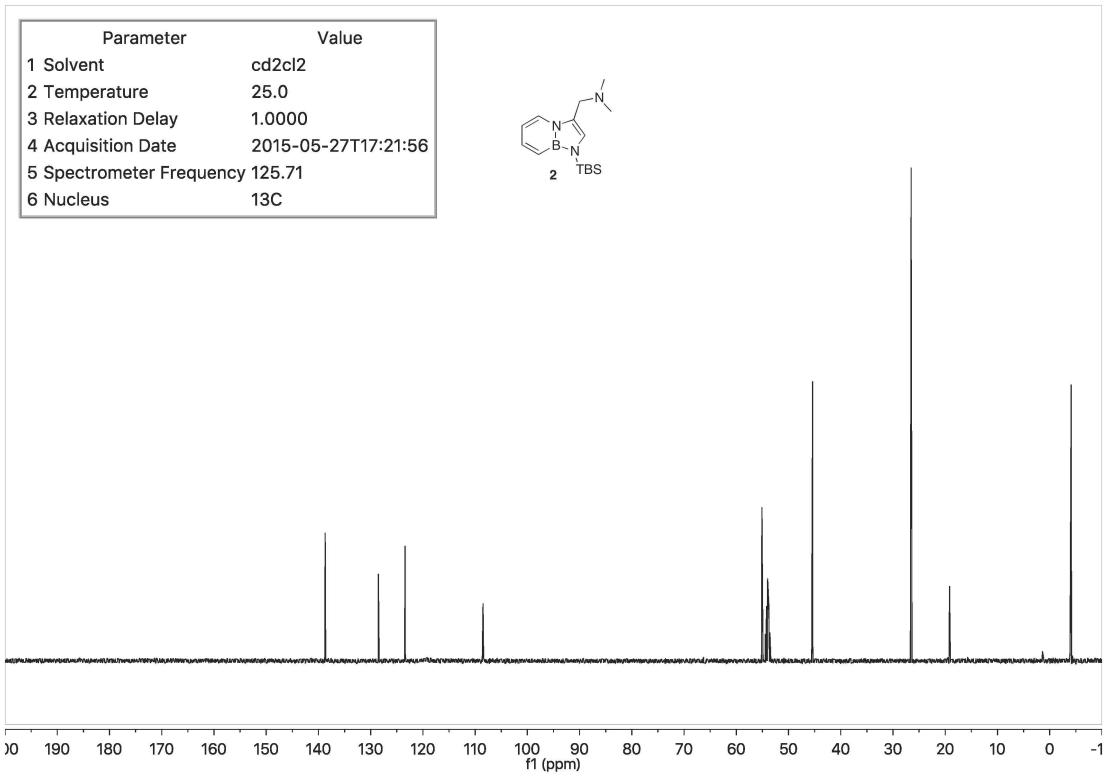

1. E. R. Abbey, L. N. Zakharov, S.-Y. Liu *J. Am. Chem. Soc.* **2011**, *133*, 11508-11511.
2. Palladium complex was synthesized via the procedure described in: F. Levrat, H. Stoeckli-Evans, N. Engel *Tetrahedron: Asymmetry* **2002**, *13*, 2335–2344.
3. J. S. Italia, P. S. Addy, C. J. J. Wrobel, L. A. Crawford, M. J. Lajoie, Y. Zheng, A. Chatterjee *Nat. Chem. Biol.* **2017**, *13*, 446–450.

## *NMR Spectra*

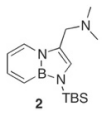
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 5.0000 |
| 4 Acquisition Date | 2015-05-27T17:18:01 |
| 5 Spectrometer Frequency | 499.88 |
| 6 Nucleus | 1H |



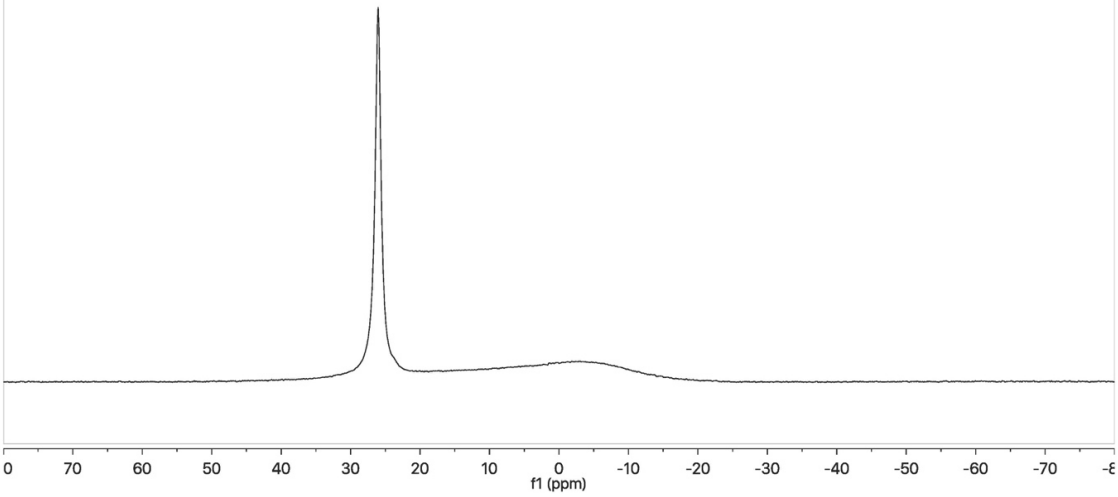

| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2015-05-27T17:21:56 |
| 5 Spectrometer Frequency | 125.71 |
| 6 Nucleus | 13C |



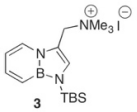
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2015-05-27T17:14:05 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



2

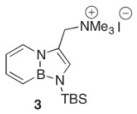



| Parameter | Value |
|---|---|
| 1 Solvent | dmso |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2017-06-12T15:07:30 |
| 5 Spectrometer Frequency | 499.89 |
| 6 Nucleus | 1H |

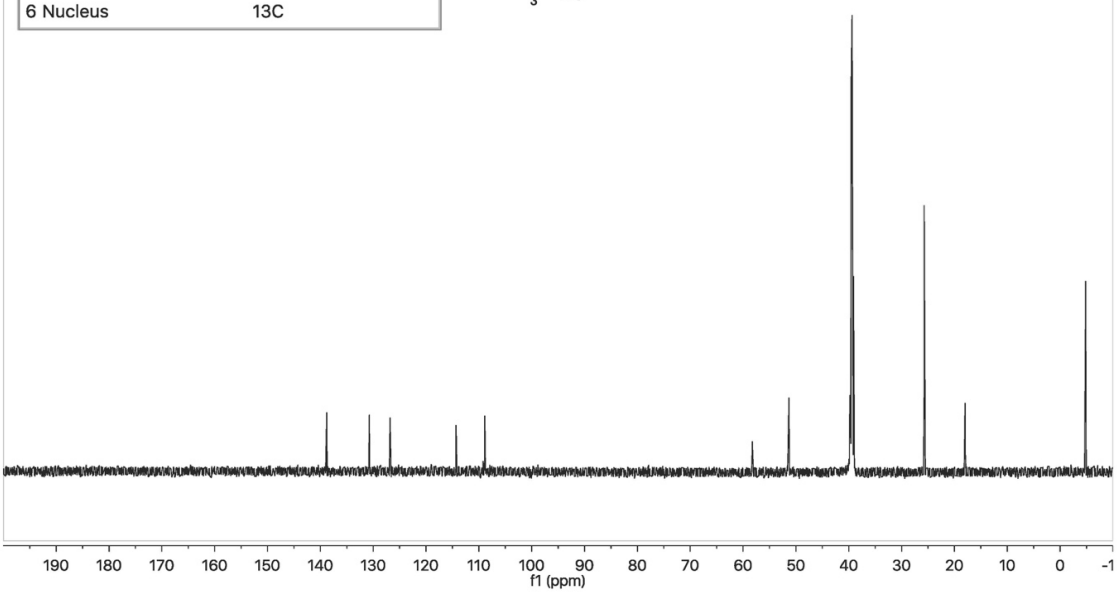


3

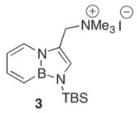
| Parameter | Value |
|---|---|
| 1 Solvent | dmso |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2016-10-14T12:49:58 |
| 5 Spectrometer Frequency | 150.81 |
| 6 Nucleus | 13C |



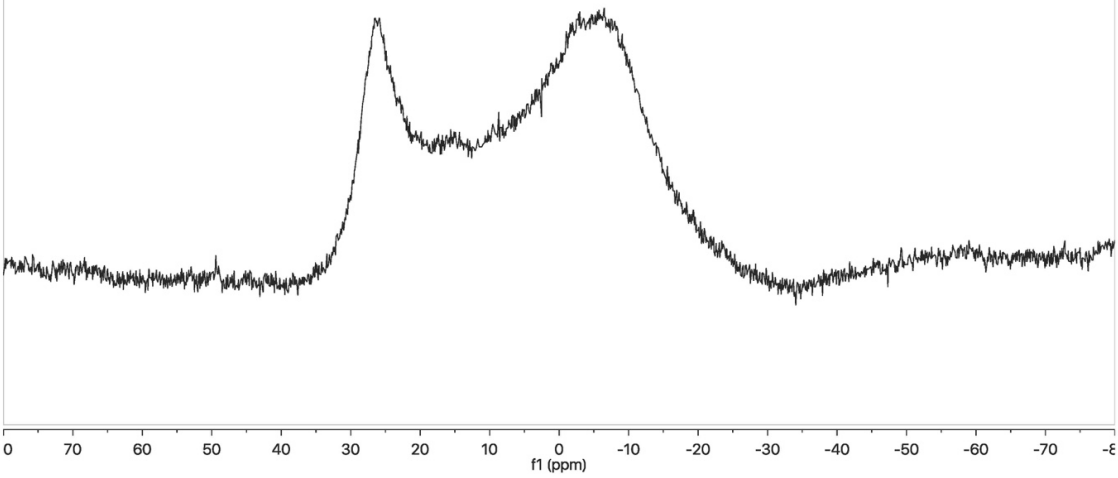
3



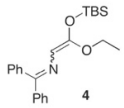

| Parameter | Value |
|---|---|
| 1 Solvent | dmso |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2017-06-12T15:04:40 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



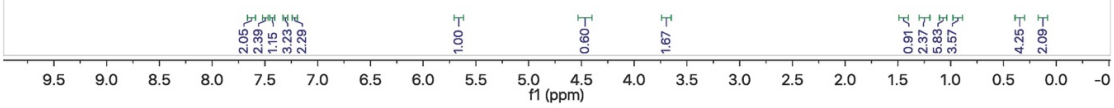
3



| Parameter | Value |
|---|---|
| 1 Comment | KLB_3_044_1H_B |
| 2 Solvent | cd2cl2 |
| 3 Temperature | 25.0 |
| 4 Relaxation Delay | 5.0000 |
| 5 Acquisition Date | 2018-05-01T14:24:16 |
| 6 Spectrometer Frequency | 499.88 |
| 7 Nucleus | 1H |



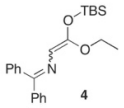

4



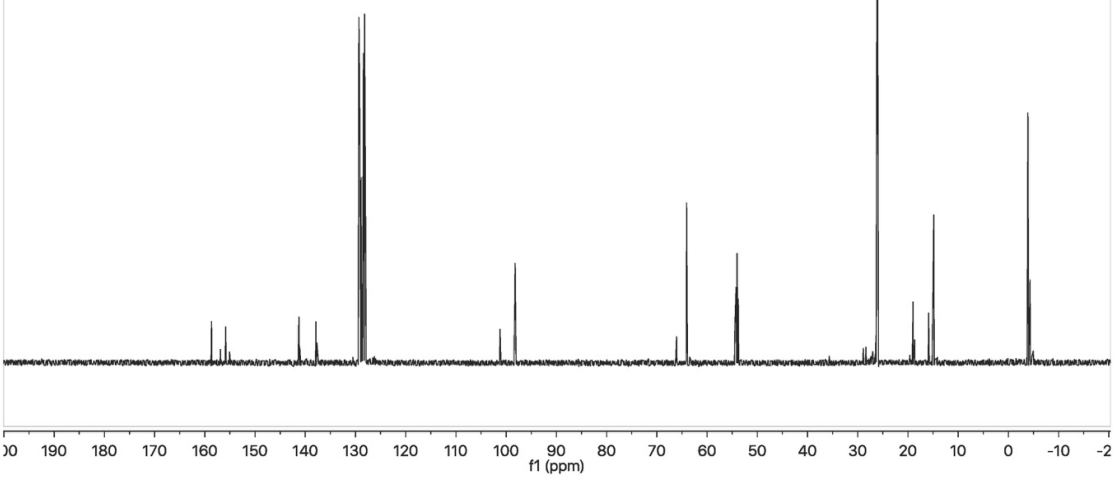




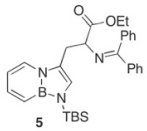
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2018-05-01T14:29:05 |
| 5 Spectrometer Frequency | 125.71 |
| 6 Nucleus | 13C |




4



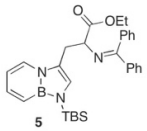
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 5.0000 |
| 4 Acquisition Date | 2016-10-16T12:14:09 |
| 5 Spectrometer Frequency | 499.88 |
| 6 Nucleus | 1H |



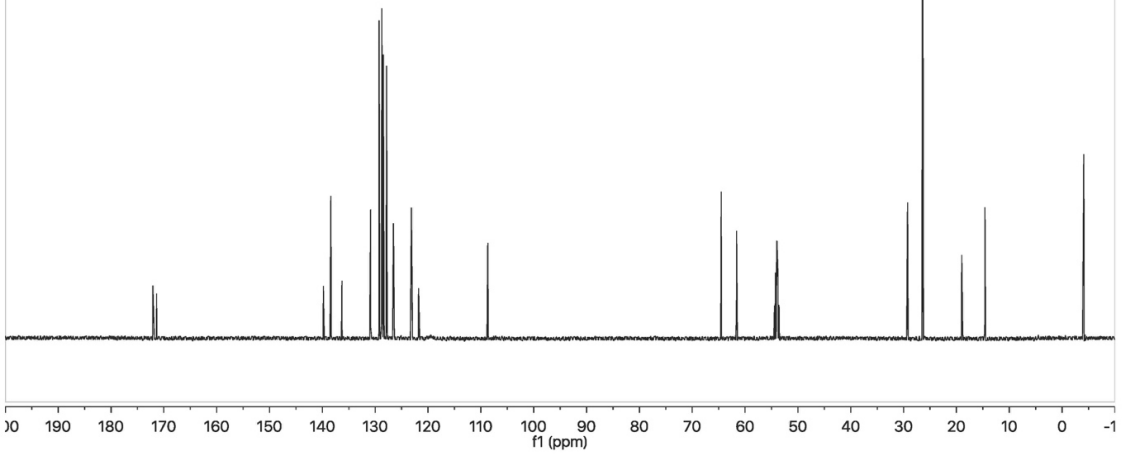
5



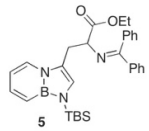

| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2016-10-16T12:18:26 |
| 5 Spectrometer Frequency | 125.71 |
| 6 Nucleus | 13C |



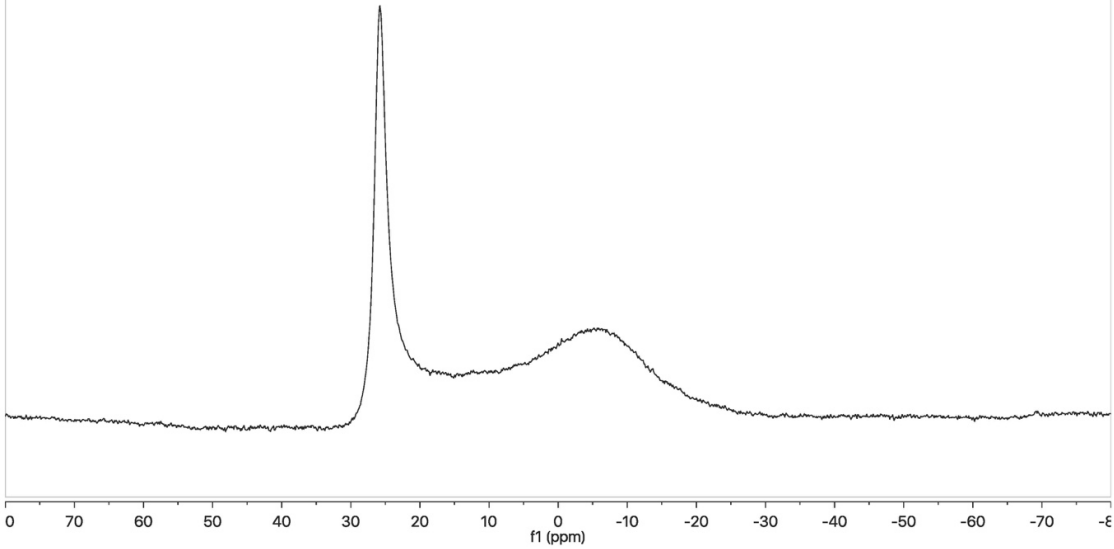
5



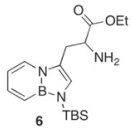
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2016-10-16T12:07:46 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



5

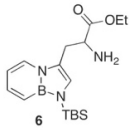



| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 5.0000 |
| 4 Acquisition Date | 2015-05-27T14:11:13 |
| 5 Spectrometer Frequency | 499.88 |
| 6 Nucleus | 1H |

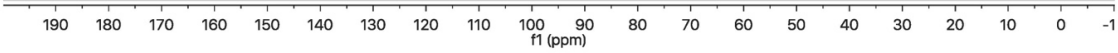


6

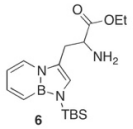
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 1.0000 |
| 4 Acquisition Date | 2015-05-27T14:16:05 |
| 5 Spectrometer Frequency | 125.71 |
| 6 Nucleus | 13C |



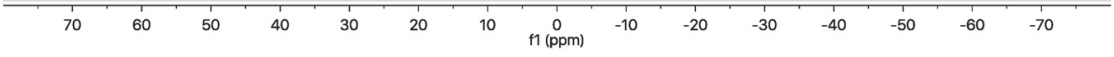
6

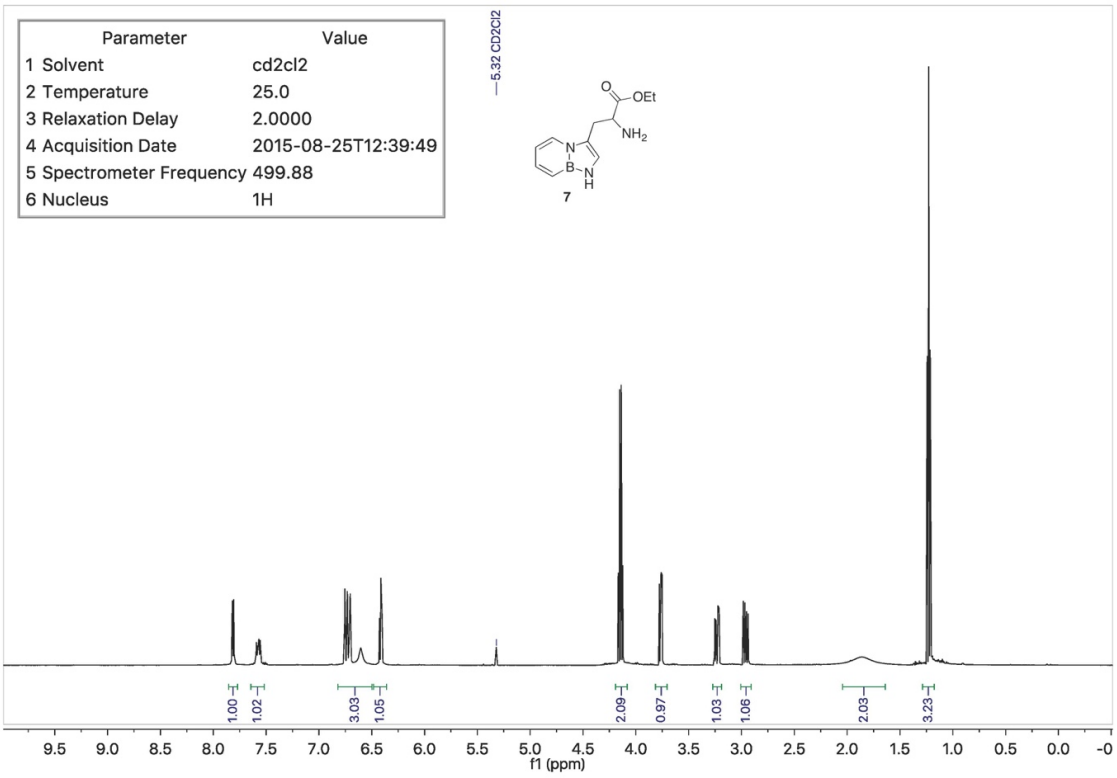


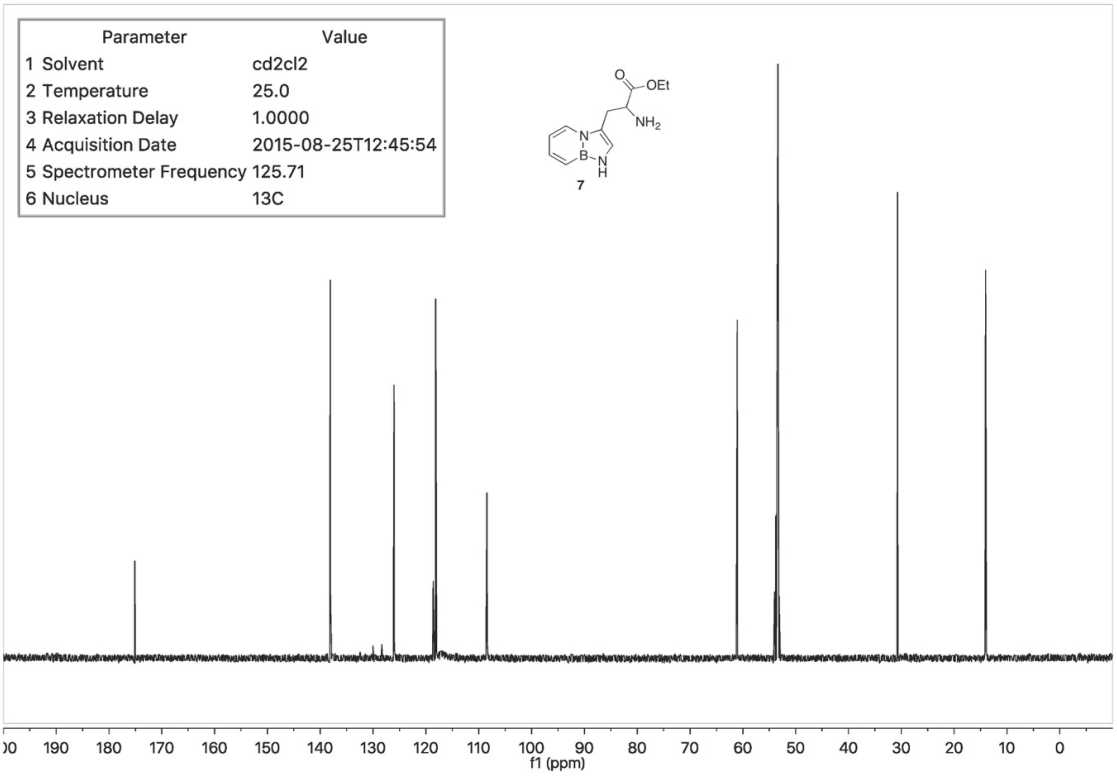

| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2015-05-27T14:06:45 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



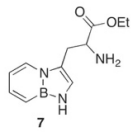
6



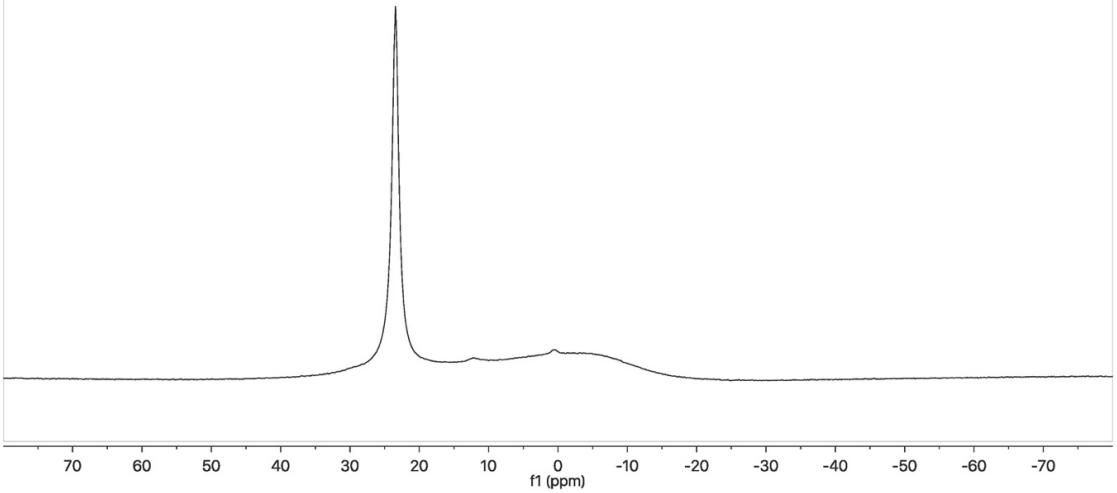

Parameter Value
1 Solvent cd2cl2
2 Temperature 25.0
3 Relaxation Delay 2.0000
4 Acquisition Date 2015-08-25T12:39:49
5 Spectrometer Frequency 499.88
6 Nucleus 1H
5.32 CD2Cl2
OEt
NH2
N
B
N
H
7
1.00
1.02
3.03
1.05
2.09
0.97
1.03
1.06
2.03
3.23
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0
f1 (ppm)


Parameter Value
1 Solvent cd2cl2
2 Temperature 25.0
3 Relaxation Delay 1.0000
4 Acquisition Date 2015-08-25T12:45:54
5 Spectrometer Frequency 125.71
6 Nucleus 13C
OEt
NH2
N
B
N
H
7
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

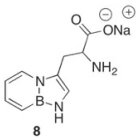
| Parameter | Value |
|---|---|
| 1 Solvent | cd2cl2 |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2015-08-25T12:35:22 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



7




| Parameter | Value |
|---|---|
| 1 Solvent | dmso |
| 2 Temperature | 25.0 |
| 3 Relaxation Delay | 0.1000 |
| 4 Acquisition Date | 2016-10-14T11:28:05 |
| 5 Spectrometer Frequency | 160.38 |
| 6 Nucleus | 11B |



8