Electronic Supplementary Material (ESI) for Chemical Science.

Combining traditional 2D and modern physical organic-derived descriptors to predict enhanced enantioselectivity for the key aza-Michael conjugate addition in the synthesis of Prevymis™ (letermovir)

Toni T. Metsänen,[a,†] Katrina W. Lexa,[b,†,*] Celine B. Santiago,[a] Cheol K. Chung,[c] Yingju Xu,[c] Zhijian Liu,[c] Guy R. Humphrey,[c] Rebecca T. Ruck,[c] Edward C. Sherer[b] and Matthew S. Sigman[a,*]

[a]Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, United States

[b]Modeling and Informatics, MRL, Merck Sharpe & Dohme, Rahway, New Jersey 07065, United States

[c]Process Research and Development, MRL, Merck Sharpe & Dohme, Rahway, New Jersey 07065, United States

Table of Contents

1. General Considerations

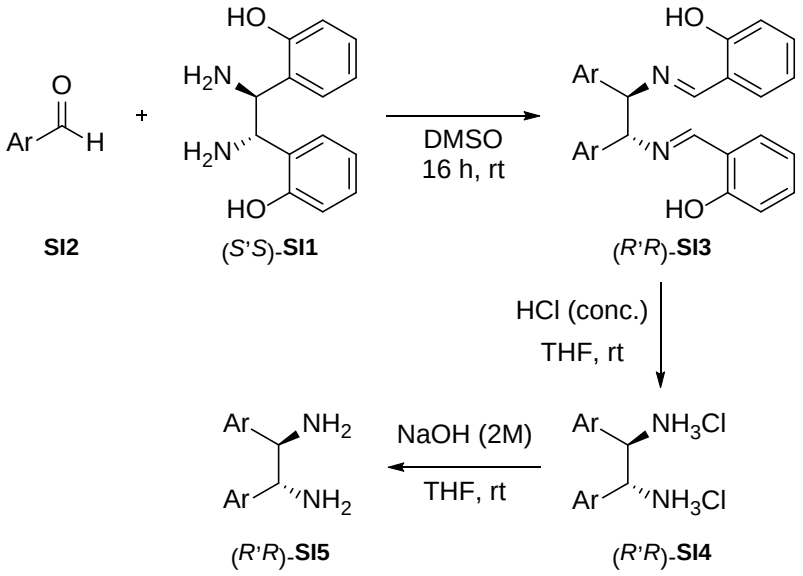
All reagents were purchased from commercial suppliers and used without further purification. Dichloromethane, tetrahydrofuran (THF), diethyl ether and acetonitrile were passed through a column of activated alumina immediately prior to use or anhydrous solvents were used from Sure/Seal™ bottles without further drying or purification. NMR spectra were recorded in acetone-d_6, $CDCl_3$, DMSO-d_6, or CD_3OD on Varian VXR 500 or Bruker Avance 500 spectrometers. ^{1}H NMR chemical shifts are reported in ppm and referenced to the acetone-d_5 signal at 2.05 ppm, $CHCl_3$ signal at 7.26 ppm, DMSO-d_5 signal at 2.50 ppm, or CHD_2OD signal at 3.31 ppm. ^{13}C NMR spectra reported in ppm referenced to the acetone-d_6 signal at 29.84 ppm, $CDCl_3$ signal at 77.16 ppm, DMSO-d_6 signal at 39.52 ppm, or CD_3OD signal at 49.00 ppm. The abbreviations s, br s, d, t, q, and m stand for the resonance multiplicities singlet, broad singlet, doublet, triplet, quartet, and multiplet, respectively. Thin-layer chromatography was performed with EMD silica gel 60 F254 plates eluting with solvents indicated, visualized by a 254 nm UV lamp. Flash chromatography was performed using EM reagent silica 60 (230–400 mesh). IR spectra were recorded using a Thermo Nicolet FT-IR. HRMS data were obtained on a Waters LCP Premier XE instrument or Micromass QTOF Ultima API US mass spectrometer by ESI/TOF. Optical rotations were measured on a Perkin Elmer Model 343 Polarimeter or Autopol V Plus Automatic Polarimeter from Rudolph Research Analytical at 589 nm with a micro cell with a 1 dm path length; concentrations are reported in g/100 mL. Enantiomeric excess was measured using Hewlett Packard 1100 Series HPLC; Column: Chiralpak IC-3 column (150 × 4.6 mm, 3 µm fused-core particle size); Eluent: A) 0.1% aq. $Na_2B_4O_7$ containing 0.04% TFA, B) MeCN; Isocratic at 72% B over 6 or 12 min; Flow rate: 1.0 mL/min; Sample injection volume: 5 µL; Column temperature: 35 °C; UV detection: 230 nm.

Catalysts **2c**, **2e**, **2f**, **2i**, **2l**, **2q**, and **2aa** were prepared according to literature procedures.[1]

[1] C. K. Chung, Z. Liu, K. W. Lexa, T. Andreani, Y. Xu, D. A. DiRocco, G. R. Humphrey and R. T. Ruck, *J. Am. Chem. Soc.* 2017, **139**, 10637.

2. General Procedures

2.1 General Procedure 1: Stereospecific Synthesis of C_2 Symmetric Diamines SI5



According to a literature procedure[2] motherdiamine (**SI1**, 1.0 equiv) was dissolved in DMSO (0.2 M) and aldehyde (**SI2**, 2.4 equiv) was added. The resulting mixture was stirred at room temperature overnight. The solution was diluted with water and extracted with Et_2O.[3] The combined organic phase was back extracted with water, brine, and dried using Na_2SO_4. Concentration under reduced pressure afforded the crude diimine **SI3**. The diimine **SI3** was dissolved in THF (0.1 M) and 38% aqueous HCl (6 equiv) was added. The solution was stirred for 2–16 h at room temperature until the desired salt precipitated.[4] The salt was filtered and washed with THF to give crude diamine salt **SI4**. The salt **SI4** was suspended in THF (0.1 M) and 2 M aqueous NaOH (20 equiv) was added. The mixture was stirred for 2 h to obtain a clear solution. The diamine was extracted with EtOAc, washed with brine, and dried using Na_2SO_4. The crude product was concentrated under reduced pressure and purified by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) to give diamine **SI5**.

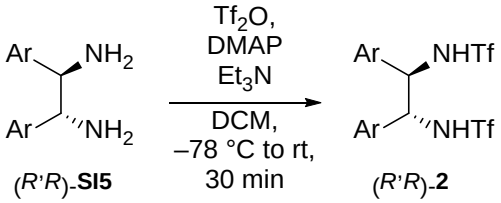
---

[2] H. Kim, Y. Nguyen, C. P.-H. Yen, L. Chagal, A. J. Lough, B. M. Kim and J. Chin, *J. Am. Chem. Soc.* 2008, **130**, 12184.

[3] Typically, a small amount of EtOAc was added to dissolve the product. In this case, the EtOAc:Et_2O ratio should not exceed ca. 20:80 to avoid DMSO extraction.

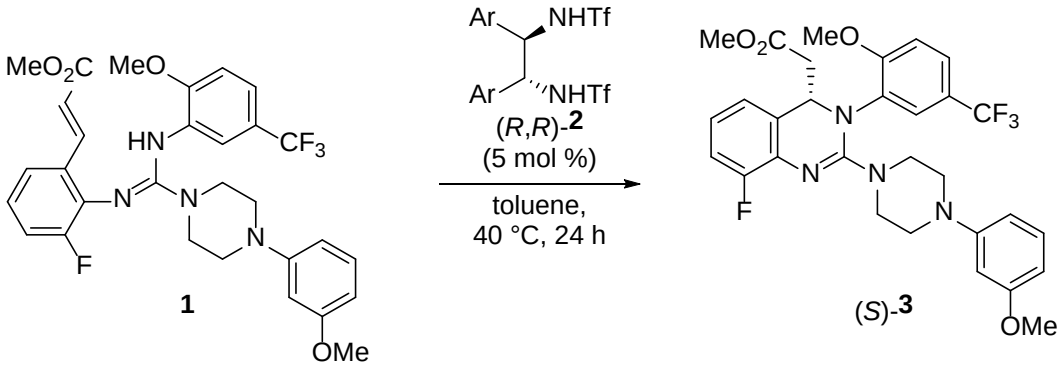
[4] If no precipitation occurs overnight, the solvent can be exchanged to Et_2O or MeCN.

2.2 General Procedure 2: Triflation Diamines



According to a literature procedure[5] diamine **SI5** (1.0 equiv), DMAP (0.02 equiv), and triethylamine[6] (3.0 equiv) were dissolved in DCM (0.2 M). The resulting clear solution was cooled to –78 °C (dry ice/acetone). After 30 min, triflic anhydride (2.5 equiv) was added dropwise over 5 min. The cooling bath was removed and the mixture was allowed to warm to room temperature over 30 min. The reaction was poured into 5% $NaHCO_3$ aqueous solution, the phases were separated, and the aqueous phase was extracted with DCM. Combined organic phases were washed with 1 M HCl and brine, dried using Na_2SO_4, and concentrated under reduced pressure to give crude product. Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4 or hexanes:EtOAc 100:0 to 50:50) gave desired bistriflimides **2**.

2.3 General Procedure 3: Asymmetric guanidine aza-conjugate addition reaction



According to a literature procedure[1] the guanidine substrate **1** (10 mg, 17 µmol, 1.0 equiv) was added at room temperature to a screw-capped vial containing the catalyst (5 mol %). Toluene (0.1 mL) was then added. The vial was sealed with a cap containing a PTFE septum and heated up to 40 °C for 24 hours.

HPLC retention times: Enantiomer A: 4.4 min; Enantiomer B: 5.4 min.

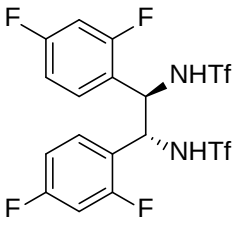
---

5 S. Pikul and E. J. Corey, *Org. Synth.* 1993, **71**, 30.

6 2,6-lutidine can be used to replace triethylamine as base.

3. Preparation of the bistriflimide catalysts 2

2,4-F_2-C_6H_3 bistriflimide 2a



(*R*,*R*)-**2a**
$C_{16}H_{10}F_{10}N_2O_4S_2$
M = 548.37

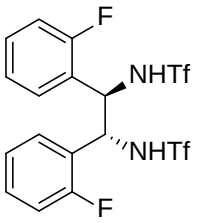
Prepared according to General Procedure 2.

Yield: 69%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.45 (dd, *J* = 14.8, 8.4 Hz, 2H), 6.92 (td, *J* = 8.5, 1.5 Hz, 2H), 6.75 (m, 2H), 5.13 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 164.4 (dd, *J* = 250.7, 12.6 Hz), 160.9 (dd, *J* = 250.1, 12.0 Hz), 131.0 (dd, *J* = 9.5, 4.4 Hz), 122.0 (dd, *J* = 13.9, 3.8 Hz), 120.9 (q, *J* = 320.9 Hz), 112.9 (dd, *J* = 21.4, 3.8 Hz), 104.5 (t, *J* = 26.5 Hz), 57.0 ppm.

Analytical data were found to match the literature data.[7]

2-F-C_6H_4 bistriflimide 2b



(*R*,*R*)-**2b**
$C_{16}H_{12}F_8N_2O_4S_2$
M = 512.39

Prepared according to General Procedure 2.

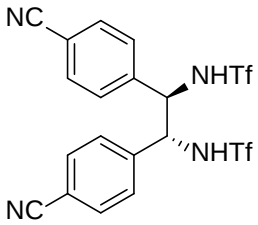
Yield: 46%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.41 (t, *J* = 7.3 Hz, 2H), 7.20 (dd, *J* = 13.1, 6.6 Hz, 2H), 7.08 (t, *J* = 7.5 Hz, 2H), 6.83 (t, *J* = 9.3 Hz, 2H), 5.21 (s, 2H) ppm, NH protons not detected;

[7] C. K. Chung, G. R. Humphrey, Z. Liu, M. Mclaughlin, Y. Xu and Y. Yu, WO2017091453 A1, June 1, 2017.

^{13}C NMR (126 MHz, CD_3OD) δ = 160.6 (d, *J* = 247.0 Hz), 131.6 (d, *J* = 8.8 Hz), 129.6, 125.7, 125.5 (d, *J* = 3.8 Hz), 120.9 (q, *J* = 320.9 Hz), 116.1 (d, *J* = 22.7 Hz), 57.5 ppm.

4-CN-C_6H_4 bistriflimide 2d



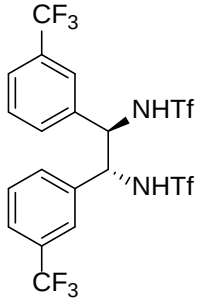
(*R*,*R*)-**2d**
$C_{18}H_{12}F_6N_4O_4S_2$
M = 526.43

Prepared according to General Procedure 2.

Yield: 79%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.59 (d, *J* =8.4 Hz, 4H), 7.34 (d, *J* = 8.3 Hz, 4H), 4.86 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 143.4, 133.6, 129.7, 120.8 (q, *J* = 320.5 Hz), 119.0, 113.5, 64.4 ppm.

3-CF_3-C_6H_4 bistriflimide 2g



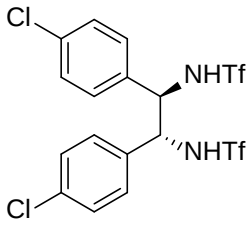
(*R*,*R*)-**2g**
$C_{18}H_{12}F_{12}N_2O_4S_2$
M = 612.40

Prepared according to General Procedure 2.

Yield: 34%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.50–7.47 (m, 2H), 7.43–7.40 (m, 4H), 7.37–7.35 (m, 2H), 4.85 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 139.8, 132.2, 132.0 (q, *J* = 32.5 Hz), 130.6 (s), 126.1 (q, *J* = 3.6 Hz), 125.6 (q, *J* = 3.7 Hz), 125.1 (q, *J* = 272.2 Hz), 120.9 (q, *J* = 320.5 Hz), 64.8 ppm.

4-Cl-C_6H_4 bistriflimide 2h



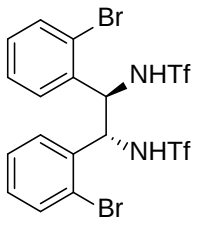
(*R*,*R*)-**2h**
$C_{16}H_{12}Cl_2F_6N_2O_4S_2$
M = 545.29

Prepared according to General Procedure 2.

Yield: 32%

^{1}H NMR (500 MHz, DMSO-d_6) δ = 10.35 (s, 2H), 7.29 (d, *J* = 8.6 Hz, 4H), 7.22 (d, *J* = 8.3 Hz, 4H), 4.7 (s, 2H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 136.1, 132.7, 129.3, 128.3, 119.0 (q, *J* = 321.6 Hz), 62.65 ppm.

2-Br-C_6H_4 bistriflimide 2j



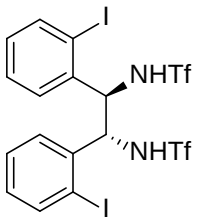
(*R*,*R*)-**2j**
$C_{16}H_{12}Br_2F_6N_2O_4S_2$
M = 634.20

Prepared according to General Procedure 2.

Yield: 70%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.78 (d, *J* = 6.1 Hz, 2H), 7.33–7.29 (m, 4H), 7.06 (t, *J* = 7.6 Hz, 2H), 5.49 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 137.9, 133.6, 131.0, 130.9, 128.7, 124.4, 120.9 (q, *J* = 321.2 Hz), 62.8 ppm.

2-I-C_6H_4 bistriflimide 2k



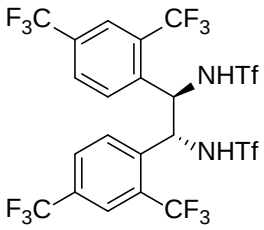
(*R*,*R*)-**2k**
$C_{16}H_{12}F_6I_2N_2O_4S_2$
M = 728.20

Prepared according to General Procedure 2.

Yield: 62%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.86 (d, *J* = 7.7. Hz, 2H), 7.59 (d, *J* = 7.9 Hz, 2H), 7.36 (t, 7.5 Hz, 2H), 6.89 (t, *J* = 7.6 Hz, 2H), 5.35 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 141.0, 140.8, 131.2, 131.0, 129.7, 120.9 (q, *J* = 321.1 Hz), 101.0, 67.7 ppm.

2,4-$(CF_3)_2$-C_6H_3 bistriflimide 2m



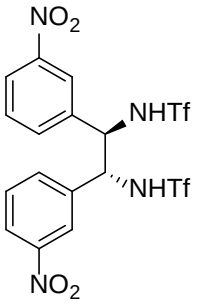
(*R*,*R*)-**2m**
$C_{20}H_{10}F_{18}N_2O_4S_2$
M = 748.40

Prepared according to General Procedure 2.

Yield: 32%

^{1}H NMR (500 MHz, CD_3OD) δ = 8.10–8.03 (m, 4H), 7.90 (s, 2H), 5.58 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 140.4, 133.3, 132.7 (q, *J* = 33.6 Hz), 130.6 (q, *J* = 2.5 Hz), 129.9 (q, *J* = 31.5 Hz), 124.43 (q, *J* = 272.2 Hz), 124.37 (q, *J* = 274.3 Hz), 124.7, 120.7 (q, *J* = 320.9 Hz), 59.7 ppm.

3-NO_2-C_6H_4 bistriflimide 2n



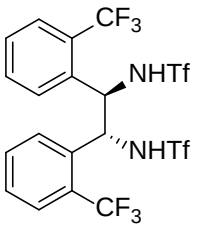
(*R*,*R*)-**2n**
$C_{16}H_{12}F_6N_4O_8S_2$
M = 566.40

Prepared according to General Procedure 2.

Yield: 27%

^{1}H NMR (500 MHz, CD_3OD) δ = 8.17 (s, 2H), 8.10 (dd, *J* = 8.2, 1.3 Hz, 2H), 7.59 (d, *J* = 7.7 Hz, 2H), 7.50 (t, *J* = 8.0 Hz, 2H), 4.99 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 149.6, 140.9, 135.0, 131.1, 124.3, 123.5, 121.0 (q, *J* = 320.7 Hz), 64.4 ppm.

2-CF_3-C_6H_4 bistriflimide 2o



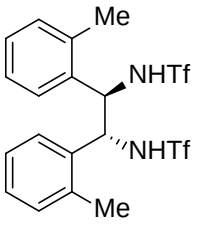
(*R*,*R*)-**2o**
$C_{18}H_{12}F_{12}N_2O_4S_2$
M = 612.40

Prepared according to General Procedure 2.

Yield: 34%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.80 (d, *J* = 7.2 Hz, 2H), 7.67 (t, *J* = 7.5 Hz, 2H), 7.54 (d, *J* = 7.8 Hz, 2H), 7.46 (t, *J* = 7.6 Hz, 2H), 5.54 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 136.3, 133.5, 131.6, 130.1, 128.8 (q, *J* = 30.1 Hz), 127.6 (q, *J* = 4.9 Hz), 125.2 (q, *J* = 273.7 Hz), 120.7 (q, *J* = 320.6 Hz), 60.1 ppm.

2-Me-C_6H_4 bistriflimide 2p



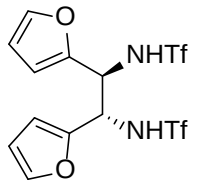
(*R*,*R*)-**2p**
$C_{18}H_{18}F_6N_2O_4S_2$
M = 504.46

Prepared according to General Procedure 2.

Yield: 55%

^{1}H NMR (500 MHz, DMSO-d_6) δ = 10.49 (br s, 2H), 7.69 (d, *J* = 6.4 Hz, 2H), 7.15 (t, *J* = 7.3 Hz, 2H), 7.02 (t, *J* = 7.4 Hz, 2H), 6.83 (d, *J* = 7.5 Hz, 2H), 5.09 (s, 2H), 1.91 (s, 6H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 136.1, 134.7, 129.7, 127.9, 127.1, 125.8, 119.1 (q, *J* = 322.4 Hz), 58.2, 18.6 ppm.

2-furyl bistriflimide 2r



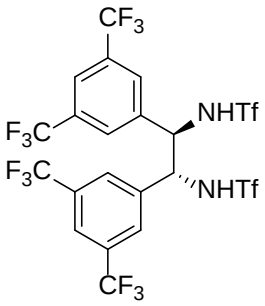
(*R*,*R*)-**2r**
$C_{12}H_{10}F_6N_2O_6S_2$
M = 456.33

Prepared according to General Procedure 2.

Yield: 44%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.41 (s, 2H), 6.25 (s, 2H), 6.15 (s, 2H), 4.95 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 150.5, 144.1, 121.1 (q, *J* = 320.6 Hz), 111.3, 109.7, 56.9 ppm.

3,5-$(CF_3)_2$-C_6H_3 bistriflimide 2s



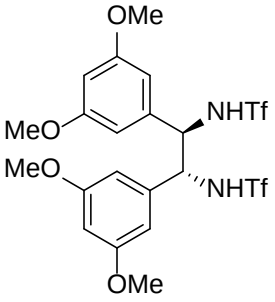
(*R*,*R*)-**2s**
$C_{20}H_{10}F_{18}N_2O_4S_2$
M = 748.40

Prepared according to General Procedure 2.

Yield: 78%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.82 (s, 6H), 5.03 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 142.0, 133.3 (q, *J* = 34.0 Hz), 129.2 (q, *J* = 3.8 Hz), 124.3 (q, *J* = 272.2 Hz), 120.9 (q, *J* = 320.9 Hz), 123.2, 64.2 ppm.

3,5-$(OMe)_2$-C_6H_3 bistriflimide 2t



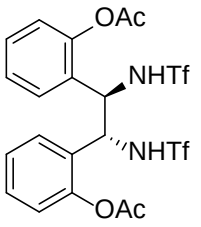
(*R*,*R*)-**2t**
$C_{20}H_{22}F_6N_2O_8S_2$
M = 596.51

Prepared according to General Procedure 2.

Yield: 42%

^{1}H NMR (500 MHz, $CDCl_3$) δ = 6.28 (s, 6H), 4.58 (s, 2H), 3.66 (s, 12H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 162.2, 140.6, 121.0 (q, *J* = 320.8 Hz), 106.8, 101.2, 65.5, 55.8 ppm.

2-OAc-C_6H_4 bistriflimide 2u



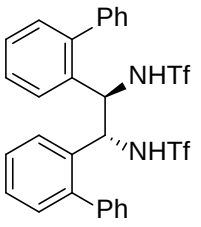
(*R*,*R*)-**2u**
$C_{20}H_{18}F_6N_2O_8S_2$
M = 592.48

Prepared according to General Procedure 2.

Yield: 72%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.52–7.49 (m, 2H), 7.21–7.16 (m, 4H), 6.95 (d, *J* = 7.3 Hz, 2H), 5.27 (s, 2H), 2.28 (s, 6H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 169.4, 148.5, 130.2, 129.0, 126.7, 123.5, 120.9 (q, *J* = 320.7 Hz), 116.3, 57.2, 21.1 ppm.

2-Ph-C_6H_4 bistriflimide 2v



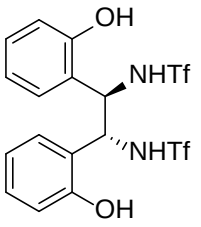
(*R*,*R*)-**2v**
$C_{28}H_{22}F_6N_2O_4S_2$
M = 628.60

Prepared according to General Procedure 2.

Yield: 29%

^{1}H NMR (500 MHz, DMSO-d_6) δ = 10.48 (s, 2H), 7.32–7.14 (m, 10H), 7.03 (d, *J* = 7.6 Hz, 2H), 6.91 (t, *J* = 7.5 Hz, 2H), 6.82 (d, *J* = 7.5 Hz, 2H), 6.68–6.58 (m, 2H), 5.13 (s, 2H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 140.6, 139.3, 134.1, 129.7, 128.9, 127.9, 127.8, 127.5, 126.6, 119.1 (q, *J* = 322.7 Hz), 58.4 ppm.

2-OH-C_6H_4 bistriflimide 2w



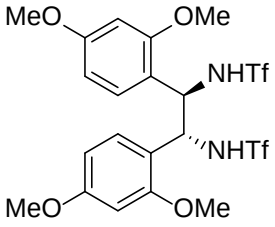
(*R*,*R*)-**2w**
$C_{16}H_{14}F_6N_2O_6S_2$
M = 508.41

Prepared according to General Procedure 2.

Yield: 2% (isolated as a byproduct during preparation of catalyst **2c**)

^{1}H NMR (500 MHz, CD_3OD) δ = 6.98 (t, *J* = 7.5 Hz, 2H), 6.95–6.91 (m, 2H), 6.68 (d, *J* = 8.0 Hz, 2H), 6.60–6.56 (m, 2H), 5.25 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 155.7, 130.3, 124.74, 124.68, 120.9 (q, *J* = 320.9 Hz), 120.2, 116.2, 61.6 ppm.

2,4-$(OMe)_2$-C_6H_3 bistriflimide 2x



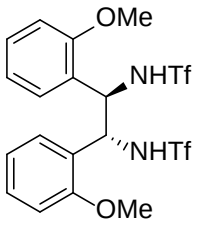
(*R*,*R*)-**2x**
$C_{20}H_{22}F_6N_2O_8S_2$
M = 596.51

Prepared according to General Procedure 2.

Yield: 46%

^{1}H NMR (500 MHz, $CDCl_3$) δ = 6.48–6.42 (m, 2H), 6.42 (s, 2H), 6.39–6.33 (m, 2H), 6.16 (d, *J* = 7.6 Hz, 2H), 4.87 (s, 2H), 3.92 (s, 6H), 3.73 (s, 6H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 161.3, 157.5, 130.9, 119.6 (q, *J* = 321.1 Hz), 116.8, 104.7, 99.2, 61.2, 55.8, 55.5 ppm.

2-OMe-C_6H_4 bistriflimide 2y



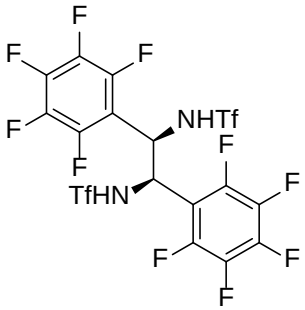
(*R*,*R*)-**2y**
$C_{18}H_{18}F_6N_2O_6S_2$
M = 536.46

Prepared according to General Procedure 2.

Yield: 29%

^{1}H NMR (500 MHz, CD_3OD) δ = 7.16–7.12 (m, 4H), 6.78–6.74 (m, 4H), 4.84 (s, 2H), 3.71 (s, 6H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 157.7, 130.5, 129.5, 126.7, 121.1, 121.0 (q, *J* = 320.8 Hz), 55.8 ppm.

C_6F_5 bistriflimide 2z



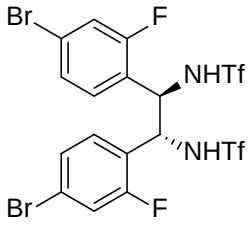
(*R*,*R*)-**2z**
$C_{16}H_4F_{16}N_2O_4S_2$
M = 656.31

Prepared according to General Procedure 2.

Yield: 27%

^{1}H NMR (500 MHz, CD_3OD) δ = 5.59 (s, 2H) ppm, NH protons not detected; **^{13}C NMR** (126 MHz, CD_3OD) δ = 146.0 (dm, *J* = 249.8 Hz), 143.3 (dtt, *J* = 256.5, 13.2, 5.1 Hz), 139.1 (dt, *J* = 252.6, 12.9 Hz), 120.9 (q, *J* = 320.4 Hz), 111.81 (t, *J* = 16.4 Hz), 52.6 ppm.

2-F-4-Br-C_6H_3 bistriflimide 2ab



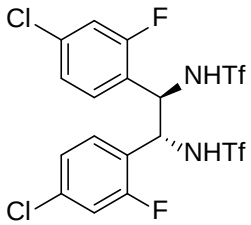
(*R*,*R*)-**2ab**
$C_{16}H_{10}Br_2F_8N_2O_4S_2$
M = 670.18

Prepared according to General Procedure 2.

Yield: 20%

[α]$^{20}_D$ = –42.3° (*c* = 1.0, DCM); **^{1}H NMR** (500 MHz, DMSO-d_6) δ = 10.76 (br s, 2H), 7.60 (t, *J* = 8.1 Hz, 2H), 7.46 (dd. *J* = 8.4, 1.7 Hz, 2H), 7.35 (dd, *J* = 9.6, 1.8 Hz, 2H), 5.08 (s, 2H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 158.2 (d, *J* = 250.8 Hz), 129.9, 128.2 (d, *J* = 3.0 Hz), 123.4 (d, *J* = 13.5 Hz), 122.4 (d, *J* = 9.8 Hz), 119.1 (q, *J* = 321.8 Hz), 118.3 (d, *J* = 25.9 Hz), 54.4 ppm; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_9Br_2F_8N_2O_4S_2]$ $(M–H)^+$: 666.8243, found 666.8256.

2-F-4-Cl-C_6H_3 bistriflimide 2ac

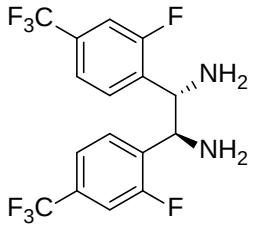


(*R*,*R*)-**X**
$C_{16}H_{10}Cl_2F_8N_2O_4S_2$
M = 581.2732

Prepared according to General Procedure 2 from diamine (*R*,*R*)-**SI5ac** (461 mg, 1.45 mmol, 1.00 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (215 mg, 0.369 mmol, 25%).

R_f = 0.29 (80:20 hexanes:EtOAc); **m.p.** (DCM): 141 °C; **[α]$^{20}_D$** = +27° (*c* = 0.1, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.08 (dd, *J* = 10.7, 1.8 Hz, 2H), 7.03 (dd, *J* = 8.3, 1.8 Hz, 2H), 6.94 (t, *J* = 8.0 Hz, 2H), 6.09 (br s, 2H), 5.00 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 160.0 (d, *J* = 249.8 Hz), 136.9 (d, *J* = 10.9 Hz), 130.0 (d, *J* = 3.1 Hz), 125.8 (d, *J* = 3.1 Hz), 121.3 (d, *J* = 13.2 Hz), 119.4 (q, *J* = 320.5 Hz), 117.3 (d, *J* = 25.1 Hz), 58.6 ppm; **IR** (neat): 3853, 2362, 2337, 1653, 1559, 1506, 1457, 1232, 1201, 1146, 668, 604 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_{10}Cl_2F_8N_2O_4S_2Na]$ $(M+Na)^+$: 602.9229, found 602.9237.

2-F-4-CF_3-C_6H_3 diamine SI5ad

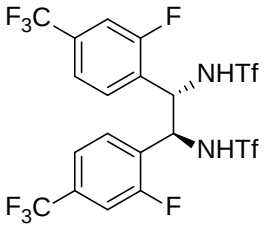


(*S*,*S*)-**SI5ad**
$C_{16}H_{12}F_8N_2$
M = 384.2732

Prepared according to General Procedure 1 from (*R*,*R*)-mother diamine (*R*,*R*)-**SI1** (500 mg, 2.0 mmol, 1.0 equiv) and 2-fluoro-4-(trifluoromethyl)benzaldehyde (**SI2ad**, 920 mg, 4.8 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as a white solid (405 mg, 1.05 mmol, 53%).

R_f = 0.20 (90:10 DCM:MeOH); **m.p.** (DCM): 69–70 °C; **$[\alpha]^{20}_D$** = –49° (*c* = 0.06, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.55 (t, *J* = 7.25 Hz, 2H), 7.37 (dd, *J* = 8.1, 1.7, 2H), 7.23 (dd, *J* = 10.3, 2.1 Hz, 2H), 4.51 (s, 2H), 1.82 (br s, 4H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 159.9 (d, *J* = 246.3 Hz), 134.3 (d, *J* = 12.5 Hz), 131.5 (dq, *J* = 33.1, 8.1 Hz), 129.3 (dq, *J* = 5.0, 1.7 Hz), 123.4 (q, *J* = 270.6 Hz), 121.2 (m), 113.1 (dq, *J* = 25.6, 3.8 Hz), 54.3 ppm; **IR** (neat): 1586, 1509, 1429, 1331, 1124, 909 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_{13}F_8N_2]$ $(M+H)^+$: 385.0951, found 385.0954.

2-F-4-CF_3-C_6H_3 bistriflimide 2ad



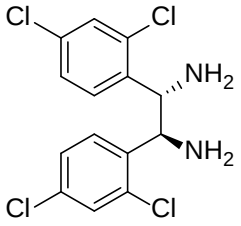
(*S*,*S*)-**2ad**
$C_{18}H_{10}F_{14}N_2O_4S_2$
M = 648.3856

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5ad** (190 mg, 0.50 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (260 mg, 0.40 mmol, 80%).

R_f = 0.05 (80:20 hexanes:EtOAc); **m.p.** (DCM): 138–143 °C; **$[\alpha]^{20}_D$** = –9° (*c* = 0.19, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.35–7.32 (m, 4H), 7.16 (t, *J* = 7.6 Hz, 2H), 6.25 (br s, 2H), 5.10 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 159.5 (d, *J* = 147.4 Hz), 134.4 (qd, *J* = 33.8, 8.8 Hz), 130.2 (d, *J* = 1.2 Hz), 126.5 (d, *J* = 12.5 Hz), 122.6 (qd, *J* = 271.3, 2.5 Hz), 119.4 (q, *J* = 318.8 Hz), 122.4 (q, *J* = 3.8 Hz), 114.2 (dq, *J* = 25.0, 3.8 Hz), 58.9 ppm; **IR**

(neat): 3291, 1457, 1436, 1372, 1331, 1215, 1172, 1139, 1068, 944 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{18}H_{10}F_{14}N_2O_4S_2Na]$ $(M+Na)^+$: 670.9756, found 670.9763.

2,4-Cl_2-C_6H_3 diamine SI5ae

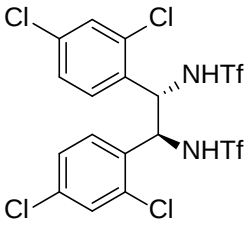


(*S*,*S*)-**SI5ae**
$C_{14}H_{12}Cl_4N_2$
M = 350.0640

Prepared according to General Procedure 1 from (*R*,*R*)-mother diamine (*R*,*R*)-**SI1** (500 mg, 2.0 mmol, 1.0 equiv) and 2,4-dichlorobenzaldehyde (**SI2ae**, 840 mg, 4.8 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as a white solid (488 mg, 1.40 mmol, 70%).

R_f = 0.26 (90:10 DCM:MeOH); **m.p.** (DCM): 88–89 °C; **$[\alpha]^{20}_D$** = –4° (*c* = 0.18, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.51 (d, *J* = 8.4 Hz, 2H), 7.32 (d, *J* = 2.2 Hz, 2H), 7.23 (dd, *J* = 8.4, 2.2 Hz, 2H), 4.64 (s, 2H), 1.62 (br s, 4H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 139.2, 138.8, 133.7, 129.8, 129.6, 127.4, 55.1 ppm; **IR** (neat): 3383, 3090, 2899 1586, 1559, 1469, 1383, 1102, 1045, 866 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{14}H_{13}Cl_4N_2]$ $(M+H)^+$: 348.9833, found 348.9838.

2,4-Cl_2-C_6H_3 bistriflimide 2ae



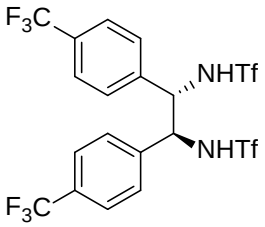
(*S*,*S*)-**2ae**
$C_{16}H_{10}Cl_4F_6N_2O_4S_2$
M = 614.1764

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5ae** (180 mg, 0.50 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (213 mg, 0.35 mmol, 69%).

R_f = 0.33 (80:20 hexanes:EtOAc); **m.p.** (DCM): 90–95 °C; **$[\alpha]^{20}_D$** = –54° (*c* = 0.18, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.38 (d, *J* = 2.2 Hz, 2H), 7.04 (dd, *J* = 8.2, 2.2 Hz, 2H), 6.85

(d, *J* = 8.3 Hz, 2H), 6.33 (s, 2H), 5.43 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 136.7, 132.5, 131.8, 131.3, 130.9, 128.4, 119.3 (q, *J* = 320.6 Hz), 61.1 ppm; **IR** (neat): 3295, 1591, 1564, 1451, 1371, 1196, 1140, 939, 739 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_{10}Cl_4F_6N_2O_4S_2Na]$ $(M+Na)^+$: 634.8638, found 634.8644.

4-CF_3-C_6H_4 bistriflimide 2af

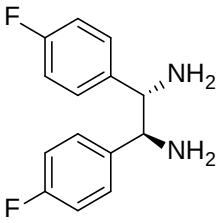


(*S*,*S*)-**2af**
$C_{18}H_{12}F_{12}N_2O_4S_2$
M = 612.4048

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5af** (367 mg, 1.05 mmol, 1.00 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (87.9 mg, 0.144 mmol, 14%).

R_f = 0.33 (80:20 hexanes:EtOAc); **m.p.** (DCM): 136 °C; **$[\alpha]^{20}_D$** = –16° (*c* = 0.1, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.57 (d, *J* = 7.8 Hz, 4H), 6.95 (d, *J* = 7.9 Hz, 4H), 4.82 (s, 2H), 1.26 (s, 2H) ppm; **^{13}C NMR** (126 MHz, CD_3OD) δ = 142.7, 131.6 (q, *J* = 32.5 Hz), 129.5, 126.6 (q, *J* = 5.6 Hz), 125.3 (q, *J* = 270.0 Hz), 120.9 (q, *J* = 318.1 Hz), 64.6 ppm; **IR** (neat): 3296, 2361, 2337, 1653, 1559, 1457, 1370, 1327, 1207, 1169, 1140, 1115, 1070 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{18}H_{12}F_{12}N_2O_4S_2Na]$ $(M+Na)^+$: 634.9958, found 634.9964.

4-F-C_6H_4 diamine SI5ag

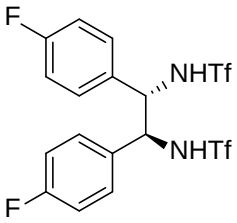


(*S*,*S*)-**SI5ag**
$C_{14}H_{14}F_2N_2$
M = 248.2768

Prepared according to General Procedure 1 from (*R*,*R*)-mother diamine (*R*,*R*)-**SI1** (730 mg, 3.0 mmol, 1.0 equiv) and 4-fluorobenzaldehyde (**SI2ag**, 890 mg, 7.2 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as viscous oil (543 mg, 1.41 mmol, 58%).

R_f = 0.14 (90:10 DCM:MeOH); **$[\alpha]^{20}_D$** = –48° (*c* = 0.26, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.16 (dd, *J* = 8.6, 5.5 Hz, 4H), 6.93 (t, *J* = 8.7 Hz, 4H), 4.00 (s, 4H), 1.61 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 162.1 (d, *J* = 245.0 Hz), 139.2 (d, *J* = 3.8 Hz), 128.7 (d, *J* = 8.8 Hz), 115.2 (d, *J* = 21.3 Hz), 61.9 (d, *J* = 1.3 Hz) ppm; **IR** (neat): 3373, 2916, 2850, 1602, 1499, 1218, 1156, 830 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{14}H_{15}F_2N_2]$ $(M+H)^+$: 249.1203, found 249.1200.

4-F-C_6H_4 bistriflimide 2ag



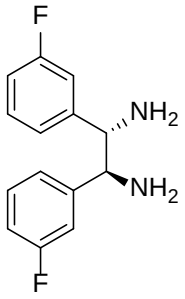
(*S*,*S*)-**2ag**
$C_{16}H_{12}F_8N_2O_4S_2$
M = 512.3892

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5ag** (110 mg, 0.45 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (160 mg, 0.31 mmol, 70%).

R_f = 0.15 (80:20 hexanes:EtOAc); **m.p.** (DCM): 209 °C; **$[\alpha]^{20}_D$** = –15° (*c* = 0.19, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.01–6.95 (m, 8H), 6.06 (br s, 2H), 4.73 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 164.0 (d, *J* = 248.8 Hz), 131.3 (d, *J* = 2.5 Hz), 129.0 (d, *J* = 8.8 Hz), 119.4 (q, *J* = 318.8 Hz), 116.6 (d, *J* = 22.5 Hz), 63.4 ppm; **IR** (neat): 3301, 2361, 2339, 1609, 1513, 1369, 1229, 1146, 734 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_{12}F_8N_2O_4S_2Na]$ $(M+Na)^+$: 535.0008, found 535.0005.

Analytical data were found to match the literature data.[1]

3-F-C_6H_4 diamine SI5ah

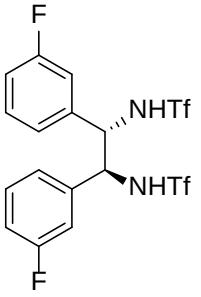


(*S*,*S*)-**SI5ah**
$C_{14}H_{14}F_2N_2$
M = 248.2768

Prepared according to General Procedure 1 from (*R*,*R*)-mother diamine (*R*,*R*)-**SI1** (500 mg, 2.0 mmol, 1.0 equiv) and 3-fluorobenzaldehyde (**SI2ah**, 600 mg, 4.8 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as a viscous oil (371 mg, 1.50 mmol, 75%).

R_f = 0.17 (90:10 DCM:MeOH); **$[\alpha]^{20}_D$** = –38° (*c* = 0.14, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.25–7.21 (m, 2H), 7.06–6.97 (m, 4H), 6.95–6.86 (m, 2H), 4.07 (s, 2H), 1.89 (s, 4H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 163.1 (d, *J* = 245.0 Hz), 146.1 (d, *J* = 6.3 Hz), 130.0 (d, *J* = 7.5 Hz), 122.8 (d, *J* = 2.5 Hz), 114.3 (d, *J* = 21.3 Hz), 114.1 (d, *J* = 21.3 Hz), 61.7 (d, *J* = 2.5 Hz) ppm; **IR** (neat): 3054, 1613, 1588, 1486, 1451, 1264, 1141, 896, 761 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{14}H_{15}F_2N_2]$ $(M+H)^+$: 249.1203, found 249.1205.

3-F-C_6H_4 bistriflimide 2ah



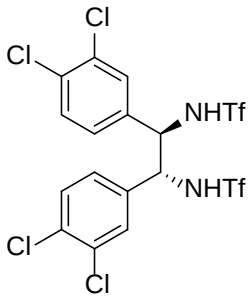
(*S*,*S*)-**X**
$C_{16}H_{12}F_8N_2O_4S_2$
M = 512.3892

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5ah** (370 mg, 1.5 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (136 mg, 0.27 mmol, 18%).

R_f = 0.25 (80:20 hexanes:EtOAc); **m.p.** (DCM): 160–170 °C; **$[\alpha]^{20}_D$** = +10° (*c* = 0.22, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.25 (td, *J* = 7.7, 5.7 Hz, 2H), 6.97 (td, *J* = 8.4, 2.4 Hz, 2H), 6.91 (d, *J* = 8.3 Hz, 2H), 6.77 (d, *J* = 9.1 Hz, 2H), 6.46 (br s, 2H), 4.79 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 162.9 (d, *J* = 248.8 Hz), 137.6 (d, *J* = 7.5 Hz), 131.3 (d, *J* = 7.5 Hz),

122.8 (d, *J* = 2.5 Hz), 119.4 (q, *J* = 318.8 Hz), 116.6 (d, *J* = 21.3 Hz), 114.5 (d, *J* = 22.5 Hz), 63.7 (d, *J* = 2.5 Hz) ppm; **IR** (neat): 3305, 1598, 1459, 1366, 1210, 1143, 1060, 925, 787 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for [$C_{16}H_{12}F_8N_2O_4S_2Na$] $(M+Na)^+$: 535.0008, found 535.0019.

3,4-Cl_2-C_6H_3 bistriflimide 2ak



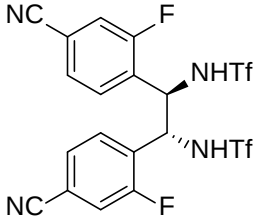
(*R*,*R*)-**2ak**
$C_{16}H_{10}Cl_4F_6N_2O_4S_2$
M = 614,18

Prepared according to General Procedure 2.

Yield: 60%

[α]$^{20}_D$ = –39.1° (*c* = 1.0, CH_2Cl_2); **^{1}H NMR** (500 MHz, DMSO-d_6) δ = 10.21 (br s, 2H), 7.68 (s, 2H), 7.54 (d, *J* = 8.3 Hz, 2H), 7.26 (d, *J* = 8.0 Hz, 2H), 4.81 (s, 2H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 137.8, 131.2, 131.0, 130.1, 130.0, 127.8, 119.0 (q, *J* = 321.9 Hz), 61.9 ppm; **HRMS** (ESI-TOF) *m/z* calculated for [$C_{16}H_9Cl_{14}F_6N_2O_4S_2$] $(M–H)^+$: 610.8662, found 610.8669.

2-F-4-CN-C_6H_3 bistriflimide 2aj



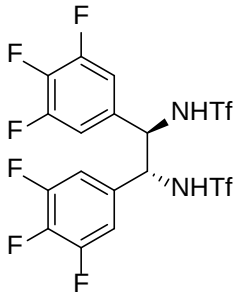
(*R*,*R*)-**2aj**
$C_{18}H_{10}F_8N_4O_4S_2$
M = 562.41

Prepared according to General Procedure 2.

Yield: 20%

[α]$^{20}_{D}$ = –51.5° (*c* = 1.0, CH_2Cl_2); **^{1}H NMR** (500 MHz, acetone-d_6) δ = 9.24 (br s, 2H), 8.02 (t, *J* = 7.7 Hz, 2H), 7.71 (d, *J* = 8.0 Hz, 2H), 7.56 (d, *J* = 9.7 Hz, 2H), 5.46 (s, 2H) ppm; **^{13}C NMR** (126 MHz, acetone-d_6) δ = 159.5 (d, *J* = 249.9 Hz), 130.9, 130.1 (d, *J* = 4.0 Hz), 129.9, 120.29 (d, *J* = 26.6 Hz), 120.26 (q, *J* = 320.4 Hz), 117.4 (d, *J* = 2.5 Hz), 115.6 (d, *J* = 10.2 Hz), 56.5 ppm; **HRMS** (ESI-TOF) *m/z* calculated for [$C_{18}H_9F_8N_4O_4S_2$] $(M–H)^+$: 560.9938, found 560.9941.

3,4,5-F_3-C_6H_2 bistriflimide 2ak



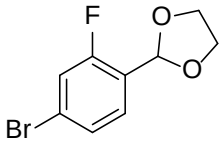
(*R*,*R*)-**2ak**
$C_{16}H_8F_{12}N_2O_4S_2$
M = 584.35

Prepared according to General Procedure 2.

Yield: 45%

[α]$^{20}_{D}$ = –20.1° (*c* = 1.0, CH_2Cl_2); **^{1}H NMR** (500 MHz, DMSO-d_6) δ = 10.37 (br s, 2H), 7.40 (s, 4H), 4.76 (s, 2H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 149.8 (ddd, *J* = 248.9, 9.5, 3.2 Hz), 138.3 (dt, *J* = 251.6 Hz), 134.0, 119.1 (q, *J* = 322.6 Hz), 112.8 (d, *J* = 21.4 Hz), 61.7 ppm; **HRMS** (ESI-TOF) *m/z* calculated for [$C_{28}H_7F_{12}N_2O_4S_2$] $(M–H)^+$: 582.9656, found 582.9667.

2-(4-bromo-2-fluorophenyl)-1,3-dioxolane (SI6)



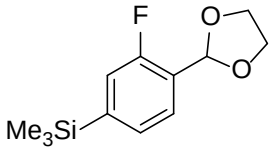
SI6
$C_9H_8BrFO_2$
M = 247.0634

In a 1 L round bottom flask equipped with a Dean–Stark apparatus, 2-fluoro-4-bromobenzaldehyde (**SI2ab**, 20 g, 100 mmol, 1.0 equiv), *p*-toluenesulfonic acid (860 mg,

5.0 mmol, 5.0 mol %) and ethylene glycol (6.8 g, 110 mmol, 1.1 equiv) were dissolved in toluene (500 mL) and the resulting mixture was stirred at reflux overnight. The solution was cooled to room temperature, neutralized with 1M NaOH solution, and the aqueous phase was extracted with EtOAc (3 × 50 mL). The combined organic phase was washed with brine, dried using Na_2SO_4, and concentrated under reduced pressure. Purification by distillation (ca. 0.1 mbar/105 °C) gave desired product as a colorless oil (2.2 g, 8.9 mmol, 89%).

R_f = 0.45 (80:20 hexanes:EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.41 (t, *J* = 7.9 Hz, 1H), 7.31 (dd, *J* = 8.3, 1.9 Hz, 1H), 7.26 (dd, *J* = 9.7, 1.9 Hz, 1H), 6.03 (s, 1H), 4.15–4.11 (m, 2H), 4.07–4.02 (m, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 161.0 (d, *J* = 252.5 Hz), 129.1 (d, *J* = 3.8 Hz), 127.6 (d, *J* = 3.8 Hz), 124.7 (d, *J* = 11.3 Hz), 123.5 (d, *J* = 8.8 Hz), 119.5 (d, *J* = 25.0 Hz), 98.6, 65.6 ppm; **IR** (neat): 2955, 2888, 1607, 1579, 1484, 1405, 1220, 1110, 873, 808 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_9H_9BrFO_2]$ $(M+H)^+$: 246.9770, found 246.9771.

(4-(1,3-dioxolan-2-yl)-3-fluorophenyl)trimethylsilane (SI7)

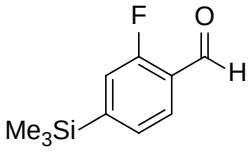


SI7
$C_{12}H_{17}FO_2Si$
M = 240.3494

In a flame-dried flask, acetal **SI6** (12 g, 50 mmol, 1.0 equiv) was dissolved in THF (250 mL) under nitrogen atmosphere. The solution was cooled to –78 °C (dry ice/acetone) and after 30 min *n*-BuLi (2.48 M in hexanes, 22 mL, 55 mmol, 1.1 equiv) was added dropwise. The solution was stirred at –78 °C for 1 h. TMS-Cl (7.6 mL, 60 mmol, 1.2 equiv) was added dropwise and the cooling bath was removed. The solution was allowed to warm to room temperature overnight. The reaction was quenched with saturated NH_4Cl (100 mL) and the aqueous phase was extracted with EtOAc (3 × 100 mL). The combined organic phase was washed with brine, dried using Na_2SO_4, and concentrated under reduced pressure. Purification by flash chromatography on silica gel (hexanes:EtOAc:Et_3N 95:4:1 to 80:19:1) gave desired product as a colorless oil (10 g, 41 mmol, 83%).

R_f = 0.53 (80:20 hexanes:EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.51 (t, *J* = 7.2 Hz, 1H), 7.29 (dd, *J* = 7.5, 1.0 Hz, 1H), 7.19 (dd, *J* = 10.4, 1.0 Hz, 1H), 6.10 (s, 1H), 4.16–4.13 (m, 2H), 4.06–4.03 (m, 2H), 0.26 (s, 9H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 160.9 (d, *J* = 250.0 Hz), 145.1 (d, *J* = 5.0 Hz), 128.9 (d, *J* = 3.8 Hz), 127.1 (d, *J* = 2.5 Hz), 125.6 (d, *J* = 12.5 Hz), 120.1 (d, *J* = 18.8 Hz), 99.1 (d, *J* = 3.8 Hz), 65.6, –1.2 ppm; **IR** (neat): 2955, 2892, 1565, 1387, 1070, 807 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{12}H_{18}FO_2Si]$ $(M+H)^+$: 241.1060, found 241.1060.

2-F-4-SiMe$_3$-C$_6$H$_3$ aldehyde SI2al

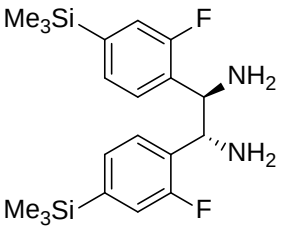


SI2al
$C_{10}H_{13}FOSi$
M = 196.2964

Acetal **SI7** (8.3 g, 34 mmol, 1.0) was dissolved in acetone (300 mL). *p*-Toluenesulfonic acid (260 mg, 1.5 mmol, 5.0 mol %) was added and the mixture was stirred at reflux for 8 h. The solution was cooled down to room temperature and neutralized with saturated $NaHCO_3$ solution (100 mL). The aqueous phase was extracted with EtOAc (5 × 100 mL). The combined organic phase was washed with brine (100 mL), dried using Na_2SO_4, and concentrated under reduced pressure. Purification by flash chromatography on silica gel (hexanes:EtOAc 99:1 to 95:5) gave desired product as a colorless oil (5.8 g, 29 mmol, 86%).

R$_f$ = 0.70 (80:20 hexanes:EtOAc); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 10.37 (s, 1H), 7.82 (t, *J* = 7.1 Hz, 1H), 7.40 (d, *J* = 7.6 Hz, 1H), 7.29 (d, *J* = 10.4 Hz, 1H), 0.3 (s, 9H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 187.6 (d, *J* = 7.5 Hz), 164.2 (d, *J* = 260.0 Hz), 152.9 (d, *J* = 5.0 Hz), 129.3 (d, *J* = 3.8 Hz), 127.8 (d, *J* = 1.1 Hz), 124.2 (d, *J* = 8.8 Hz), 121.0 (d, *J* = 17.5 Hz), –1.4 ppm; **IR** (neat): 2957, 2859, 2760, 1693, 1609, 1556, 1386, 1250, 822, 803, 754 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for [$C_{10}H_{14}FOSi$] $(M+H)^+$: 197.0798, found 197.0797.

2-F-4-SiMe$_3$-C$_6$H$_3$ diamine SI5al

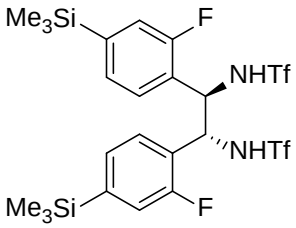


(*R*,*R*)-**SI5al**
$C_{20}H_{30}F_2N_2Si_2$
M = 392.6408

Prepared according to General Procedure 1 from (*S*,*S*)-mother diamine (*S*,*S*)-**SI1** (980 mg, 4.0 mmol, 1.0 equiv) and 2-fluoro-4-(trimethylsilyl)benzaldehyde (**SI2al**, 1.9 g, 9.6 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as a white solid (430 mg, 1.10 mmol, 27%).

R_f = 0.23 (90:10 DCM:MeOH); **m.p.** (DCM): 99–102 °C; **$[\alpha]^{20}_D$** = +51° (*c* = 0.22, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.40 (t, *J* = 7.2 Hz, 2H), 7.21 (dd, *J* = 7.4, 1.1 Hz, 2H), 7.08 (d, *J* = 10.5 Hz, 2H), 4.52 (s, 2H), 0.24 (s, 18H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 160.2 (d, *J* = 247.5 Hz), 143.3, 129.5 (d, *J* = 67.5 Hz), 129.1, 128.0 (d, *J* = 3.8 Hz), 119.9 (d, *J* = 18.8 Hz), 54.2, –1.1 ppm; **IR** (neat): 2956, 2361, 1554, 1494, 1387, 1264, 1249, 827 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{20}H_{31}F_2N_2Si_2]$ $(M+H)^+$: 393.1994, found 393.1999.

2-F-4-SiMe$_3$-C$_6$H$_3$ bistriflimide 2al

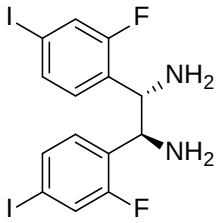


(*R*,*R*)-**2al**
$C_{22}H_{28}F_8N_2O_4S_2Si_2$
M = 656.7532

Prepared according to General Procedure 2 from diamine (*R*,*R*)-**SI5al** (196 mg, 0.50 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (200 mg, 0.30 mmol, 61%).

R_f = 0.65 (80:20 hexanes:EtOAc); **m.p.** (DCM): 114 °C; **$[\alpha]^{20}_D$** = +17° (*c* = 0.15, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.10 (dd, *J* = 11.2, 1.1 Hz, 2H), 7.02 (dd, *J* = 7.4, 1.1 Hz, 2H), 6.83 (t, *J* = 7.2 Hz, 2H), 6.00 (s, 2H), 4.96 (s, 2H), 0.20 (s, 18H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 160.1 (d, *J* = 247.5 Hz), 146.2 (d, *J* = 3.8 Hz), 129.6 (d, *J* = 3.8 Hz), 129.0 (d, *J* = 2.5 Hz), 122.9 (d, *J* = 12.5 Hz), 120.4 (d, *J* = 18.8 Hz), 119.5 (q, *J* = 318.8 Hz), 60.0, –1.3 ppm; **IR** (neat): 3305, 2958, 1561, 1379, 1193, 1144, 1058, 838, 755 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{22}H_{28}F_8N_2O_4S_2Si_2Na]$ $(M+Na)^+$: 679.0799, found 679.0799.

2-F-4-I-C_6H_3 diamine SI5am

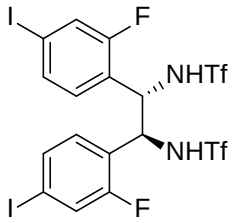


(*S*,*S*)-**SI5am**
$C_{14}H_{12}F_2I_2N_2$
M = 500.0697

Prepared according to General Procedure 1 from (*R*,*R*)-mother diamine (*R*,*R*)-**SI1** (980 mg, 4.0 mmol, 1.0 equiv) and 2-fluoro-4-iodobenzaldehyde (**SI2am**, 2,4 g, 9.6 mmol, 2.4 equiv). Purification by flash chromatography on silica gel (DCM:MeOH:Et_3N 96:2:2 to 90:5:5) gave desired product as a white solid (624 mg, 1.25 mmol, 31%).

R_f = 0.26 (90:10 DCM:MeOH); **m.p.** (DCM): 98 °C; **$[\alpha]^{20}_D$** = –85° (*c* = 0.26, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.41 (d, *J* = 8.0 Hz, 2H), 7.31 (d, *J* = 8.4 Hz, 2H), 7.09 (t, *J* = 7.7 Hz, 2H), 4.35 (s, 2H), 1.80 (s, 4H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 159.9 (d, *J* = 250.0 Hz), 133.5, 130.1 (d, *J* = 12.5 Hz), 130.0 (d, *J* = 3.8 Hz), 124.8 (d, *J* = 25.0 Hz), 91.9 (d, *J* = 8.8 Hz), 54.4 ppm; **IR** (neat): 3370, 3062, 1599, 1566, 1451, 1339, 1214, 859, 820 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{14}H_{13}F_2I_2N_2]$ $(M+H)^+$: 500.9136, found 500.9148.

2-F-4-I-C_6H_3 bistriflimide 2am



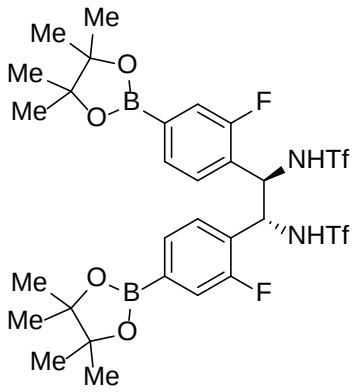
(*S*,*S*)-**2am**
$C_{16}H_{10}F_8I_2N_2O_4S_2$
M = 764.1822

Prepared according to General Procedure 2 from diamine (*S*,*S*)-**SI5am** (250 mg, 0.50 mmol, 1.0 equiv). Purification by flash chromatography on silica gel (hexanes:EtOAc:acetone 95:4:1 to 80:16:4) gave desired product as a white solid (144 mg, 0.19 mmol, 38%).

R_f = 0.35 (80:20 hexanes:EtOAc); **m.p.** (DCM): 101 °C; **$[\alpha]^{20}_D$** = –38° (*c* = 0.14, $CHCl_3$); **^{1}H NMR** (500 MHz, $CDCl_3$) δ = 7.43 (dd, *J* = 10.2, 1.4 Hz, 2H), 7.38 (dd, *J* = 8.1, 1.6 Hz, 2H), 6.69 (t, *J* = 7.8 Hz, 2H), 5.98 (br s, 2H), 4.94 (s, 2H) ppm; **^{13}C NMR** (126 MHz, $CDCl_3$) δ = 159.6 (d, *J* = 250.0 Hz), 134.7 (d, *J* = 3.8 Hz), 130.5 (d, *J* = 2.5 Hz), 125.9 (d, *J* = 23.8 Hz),

122.5 (d, *J* = 12.5 Hz), 119.4 (q, *J* = 319.2 Hz), 95.5 (d, *J* = 7.5 Hz), 59.1 ppm; **IR** (neat): 3298, 1604, 1485, 1451, 1376, 1232, 1201, 1144, 1059 934, 868 cm^{-1}; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{16}H_{10}F_8I_2N_2O_4S_2Na]$ $(M+Na)^+$: 786.7942, found 786.7957.

2-F-4-Bpin-C_6H_3 bistriflimide 2an



(*R*,*R*)-**2an**
$C_{28}H_{34}B_2F_8N_2O_8S_2$
M = 764.31

Prepared according to General Procedure 2.

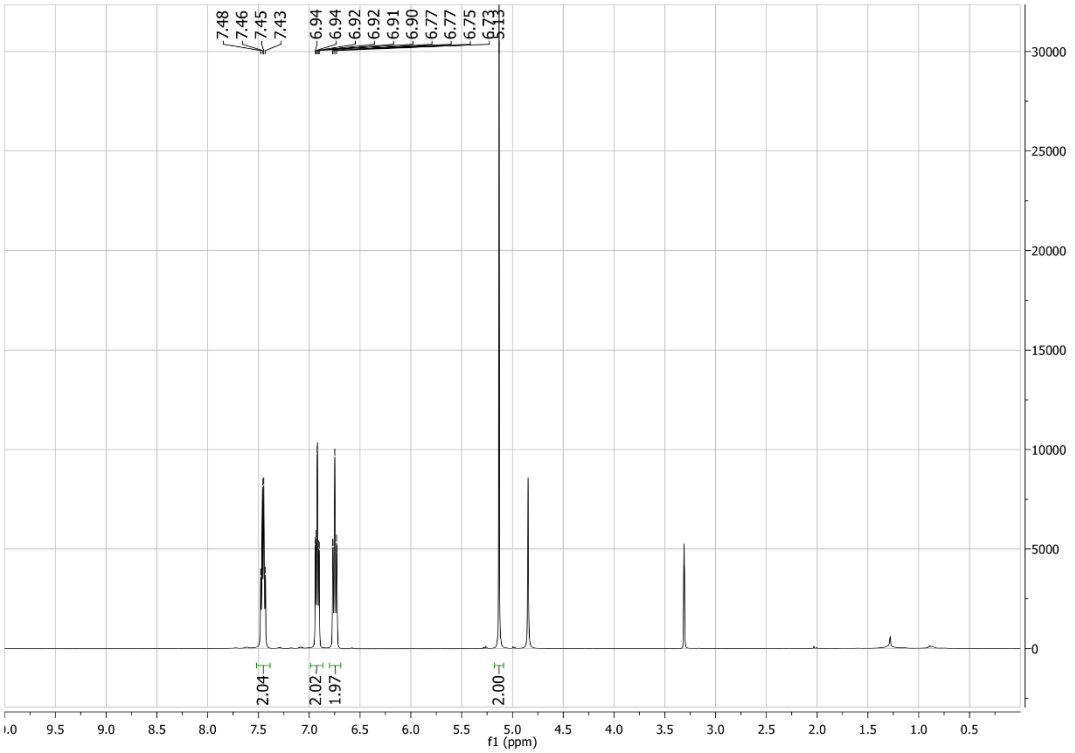
Yield: 33%

$[\alpha]^{20}_D$ = –34.1° (*c* = 1.0, CH_2Cl_2); **^{1}H NMR** (500 MHz, DMSO-d_6) δ = 10.7 (br s, 2H), 7.65 (t, *J* = 6.2 Hz, 2H), 7.42 (d, *J* = 7.6 Hz, 2H), 7.06 (d, *J* = 9.9 Hz, 2H), 5.18 (s, 2H), 1.25 (s, 24 H) ppm; **^{13}C NMR** (126 MHz, DMSO-d_6) δ = 158.0 (d, *J* = 158.0 Hz), 131.6, 130.4, 128.3, 126.9 (d, *J* = 11.3 Hz), 119.8 (d, *J* = 20.0 Hz), 119.1 (q, *J* = 320.0 Hz), 84.1, 54.7, 24.64, 24.58 ppm; **HRMS** (ESI-TOF) *m/z* calculated for $[C_{28}H_{33}B_2F_8N_2O_8S_2]$ $(M–H)^+$: 763.1747, found 763.1759.

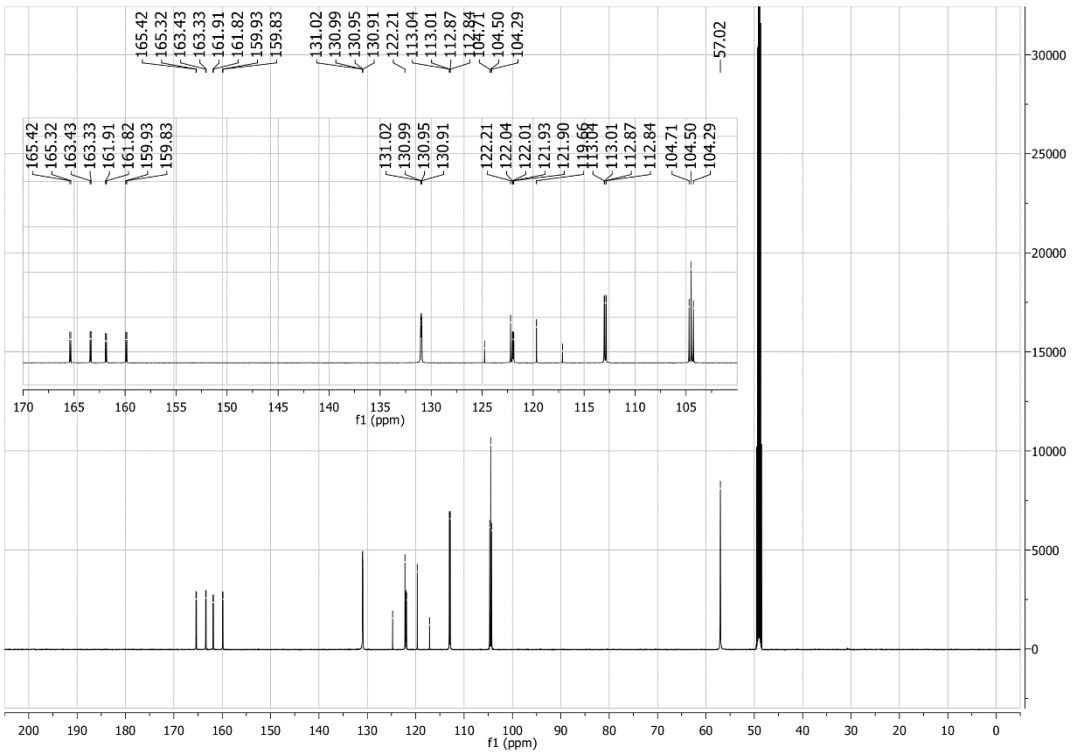
4. NMR Spectra

2,4-F_2-C_6H_3 bistriflimide **2a**

^{1}H NMR (500 MHz, CD_3OD)

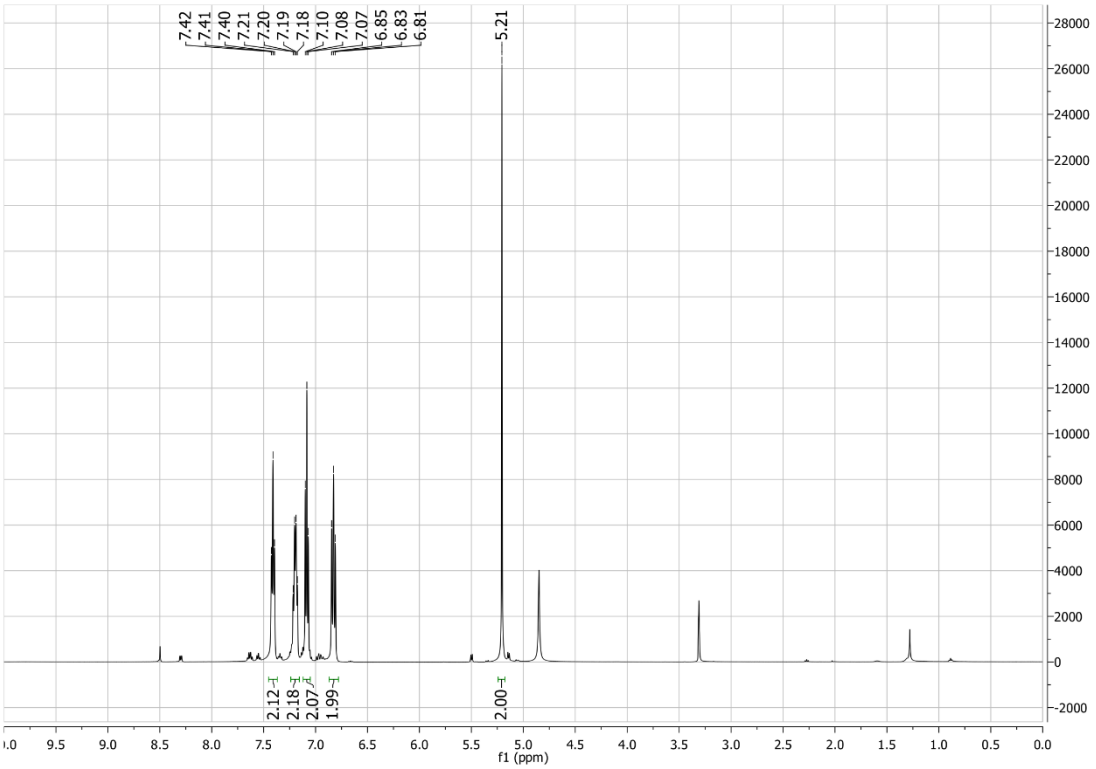


^{13}C NMR (126 MHz, CD_3OD)

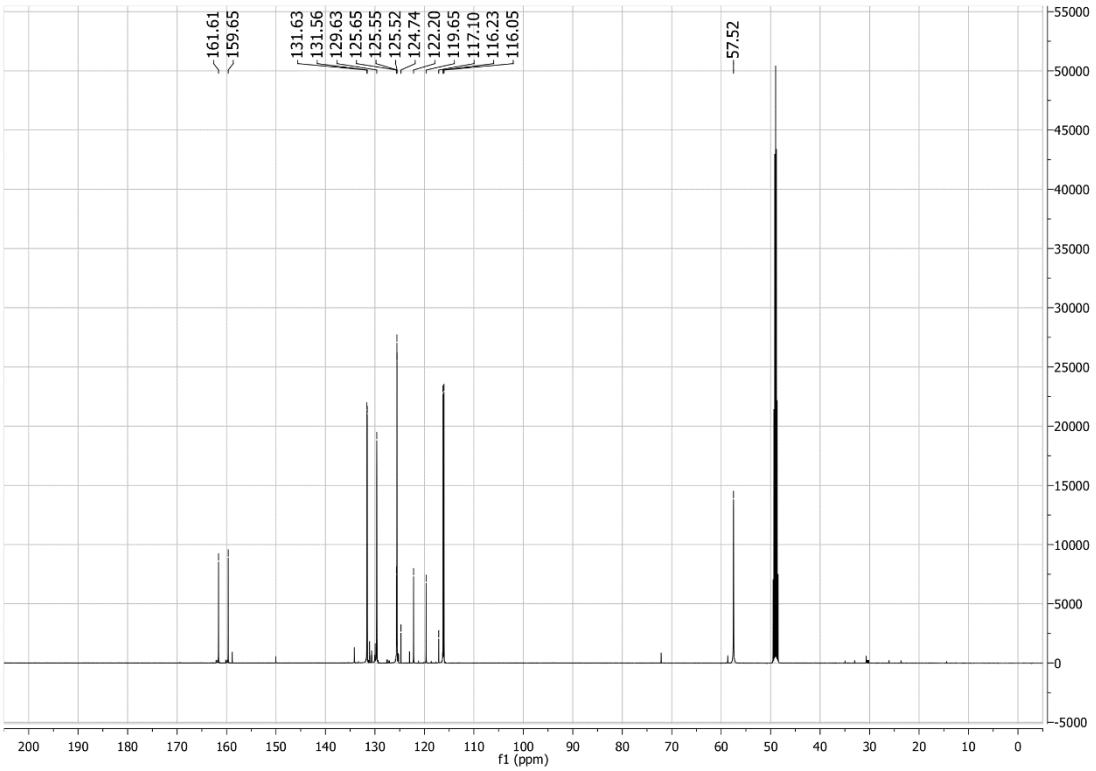


2-F_2-C_6H_4 bistriflimide **2b**

^{1}H NMR (500 MHz, CD_3OD)

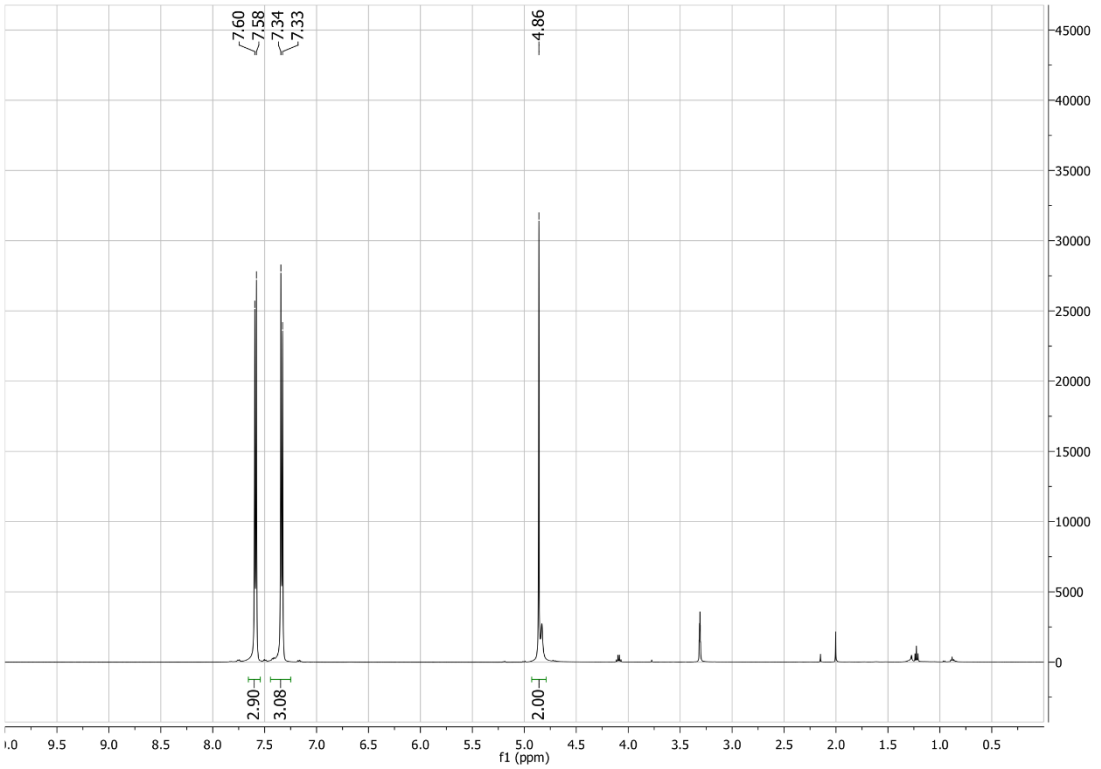


^{13}C NMR (126 MHz, CD_3OD)

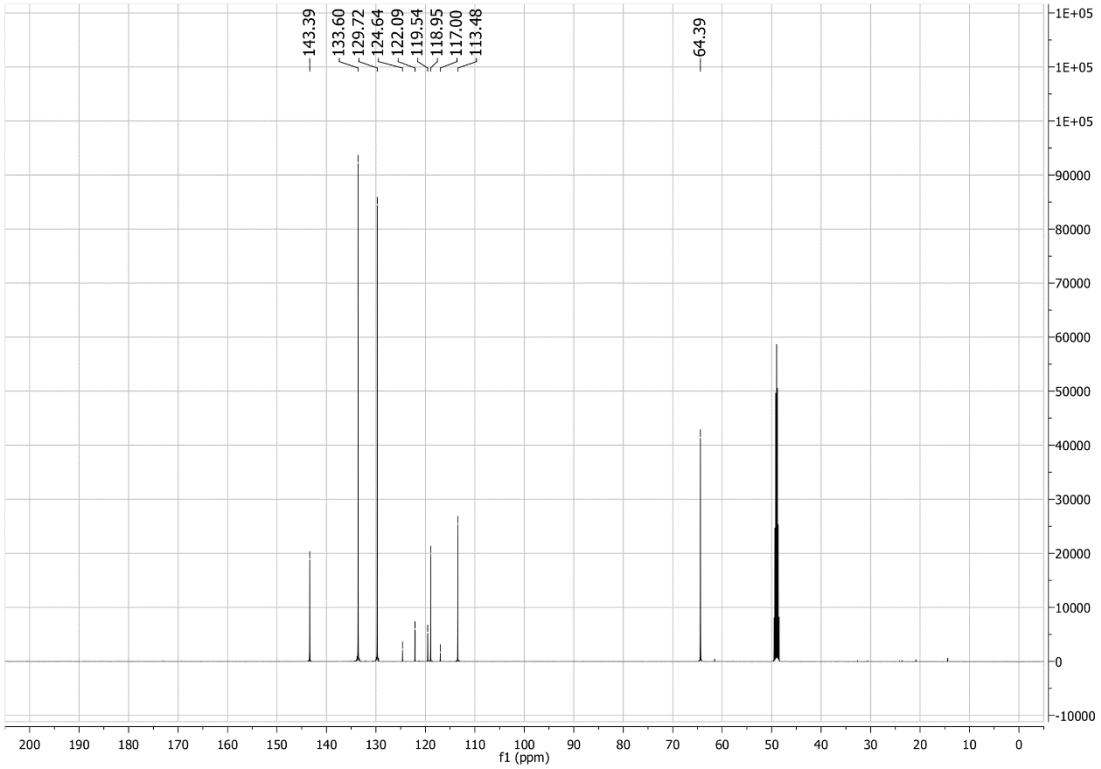


4-CN-C_6H_4 bistriflimide **2d**

^{1}H NMR (500 MHz, CD_3OD)

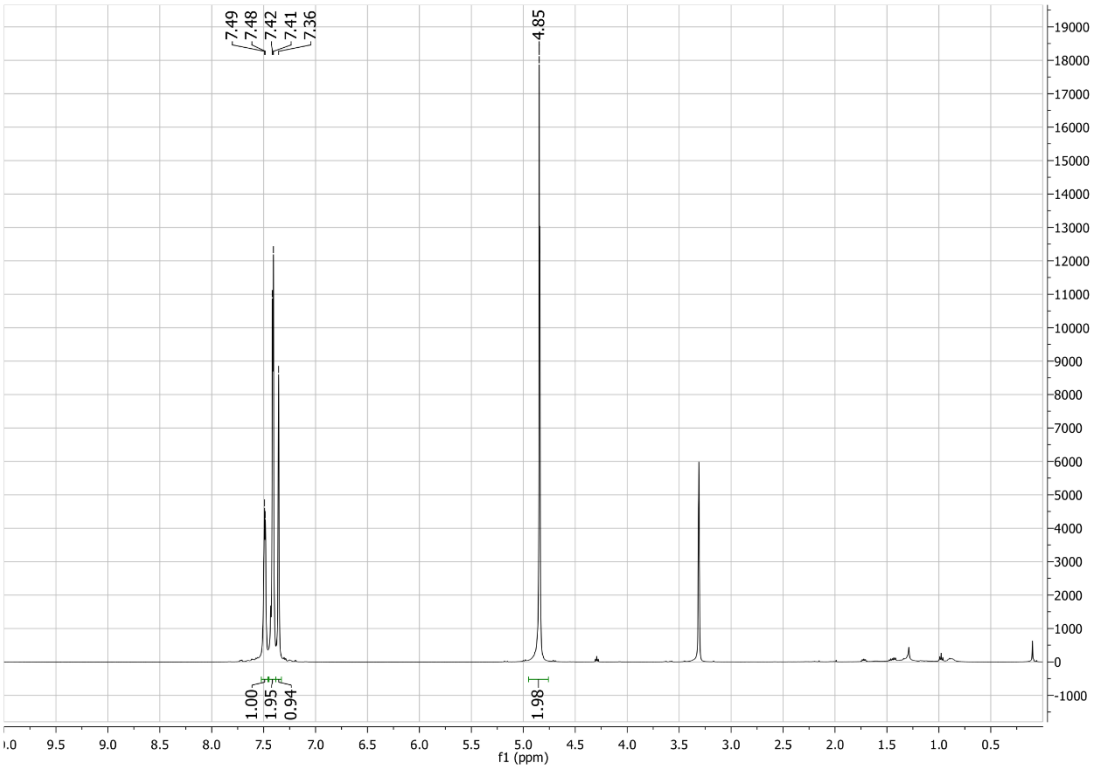


^{13}C NMR (126 MHz, CD_3OD)

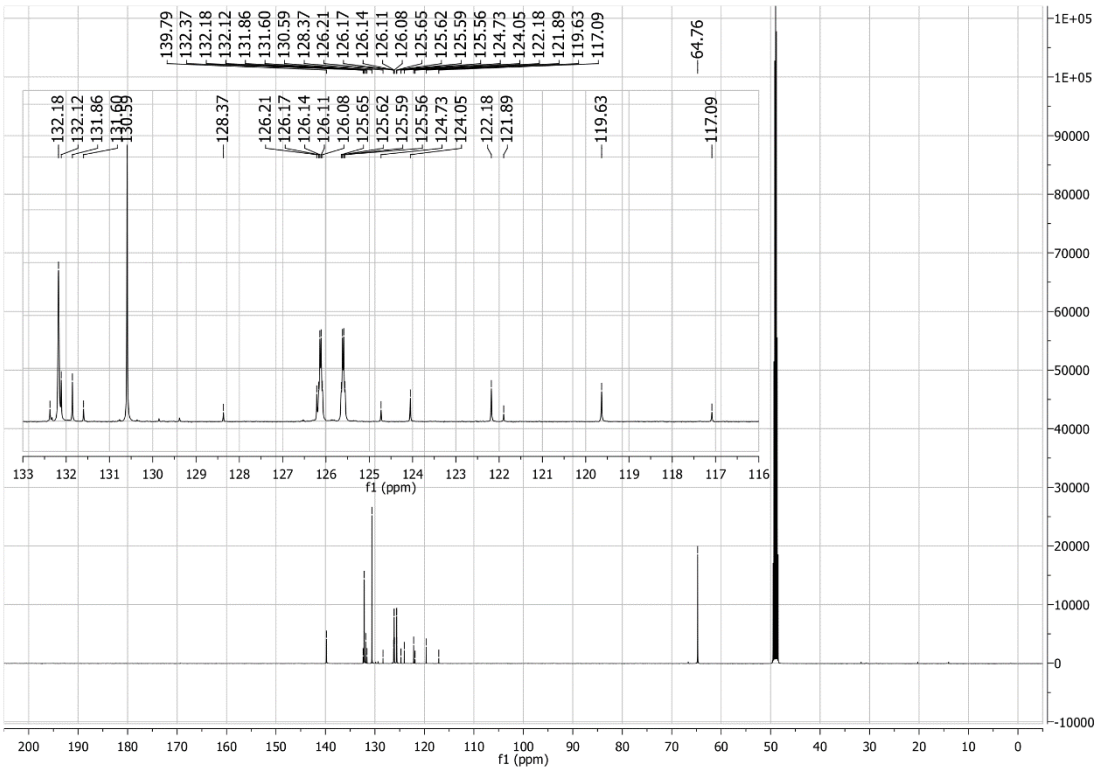


3-CF_3-C_6H_4 bistriflimide **2g**

^{1}H NMR (500 MHz, CD_3OD)

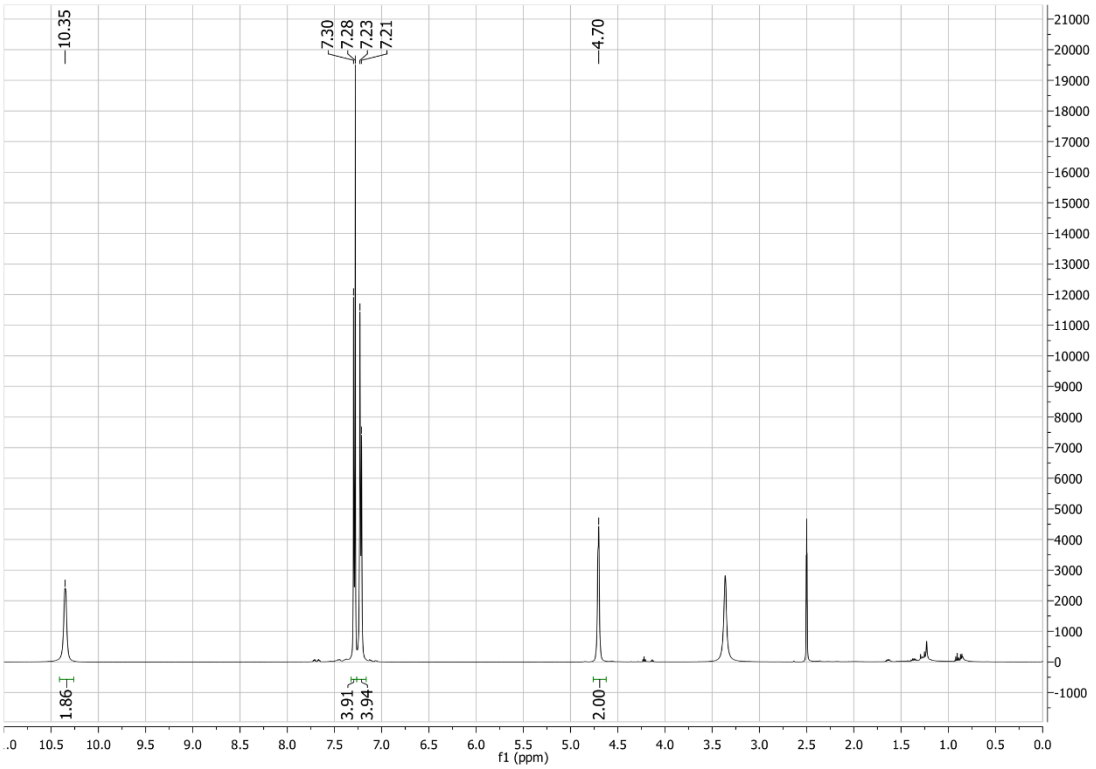


^{13}C NMR (126 MHz, CD_3OD)

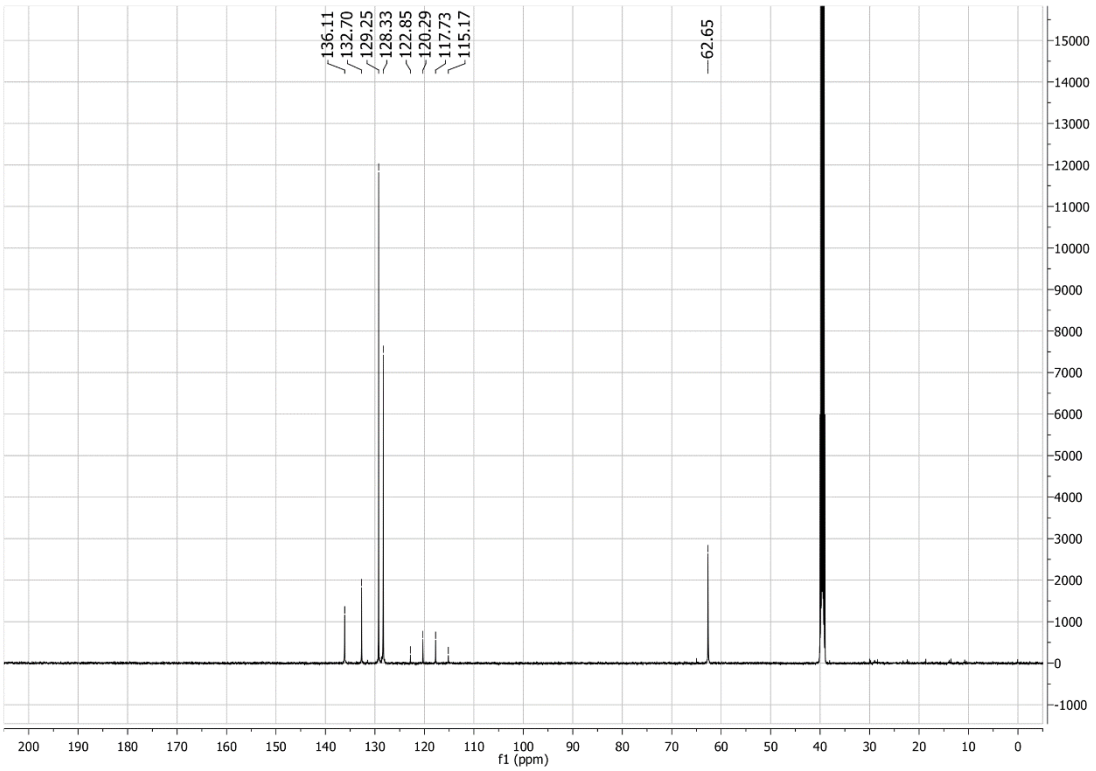


4-Cl-C_6H_4 bistriflimide **2h**

^{1}H NMR (500 MHz, DMSO-d_6)

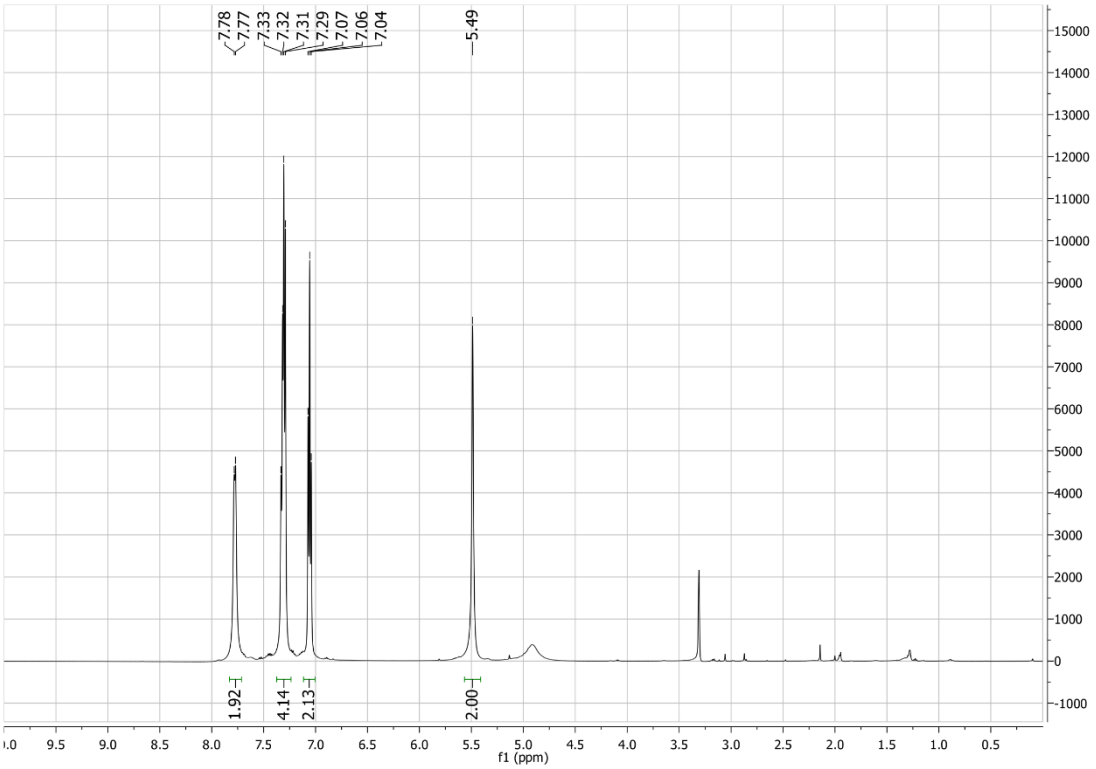


^{13}C NMR (126 MHz, DMSO-d_6)

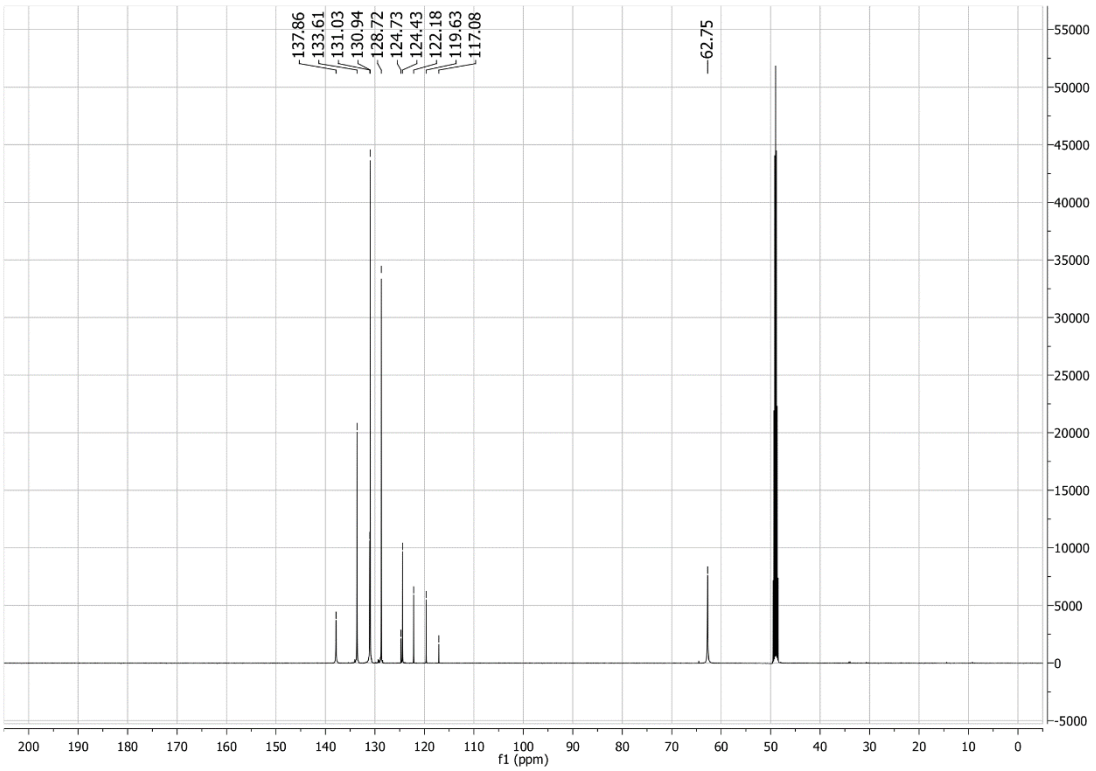


2-Br-C_6H_4 bistriflimide **2j**

^{1}H NMR (500 MHz, CD_3OD)

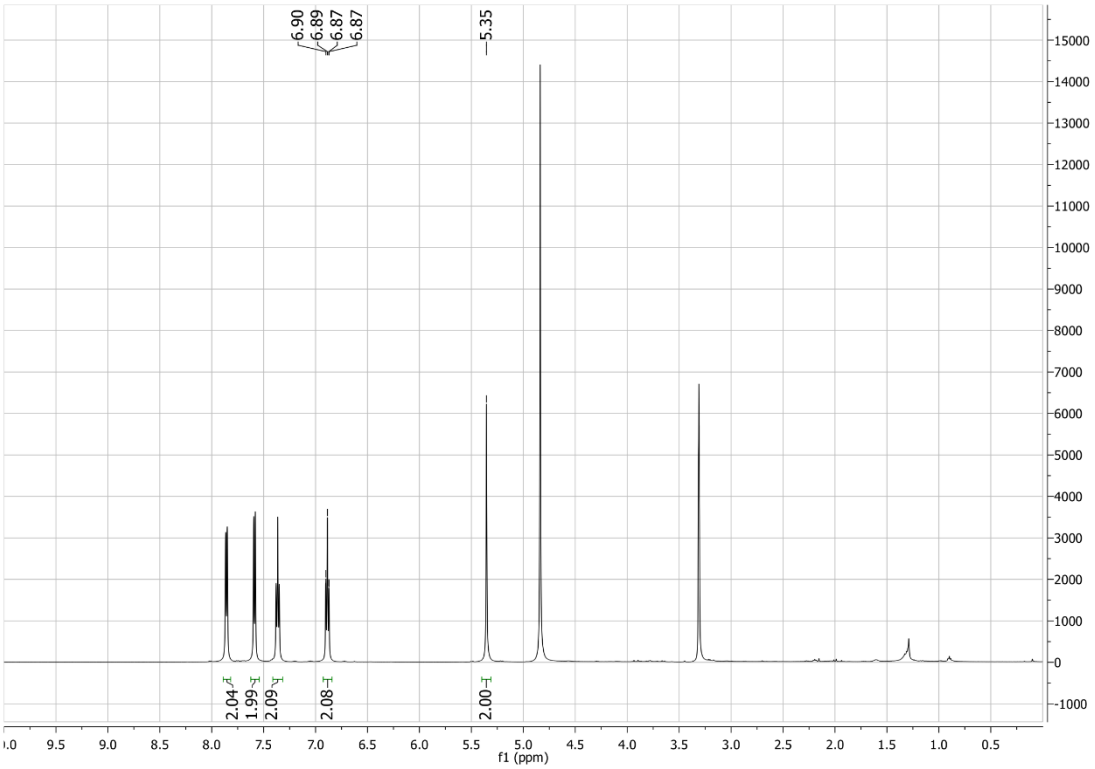


^{13}C NMR (126 MHz, CD_3OD)

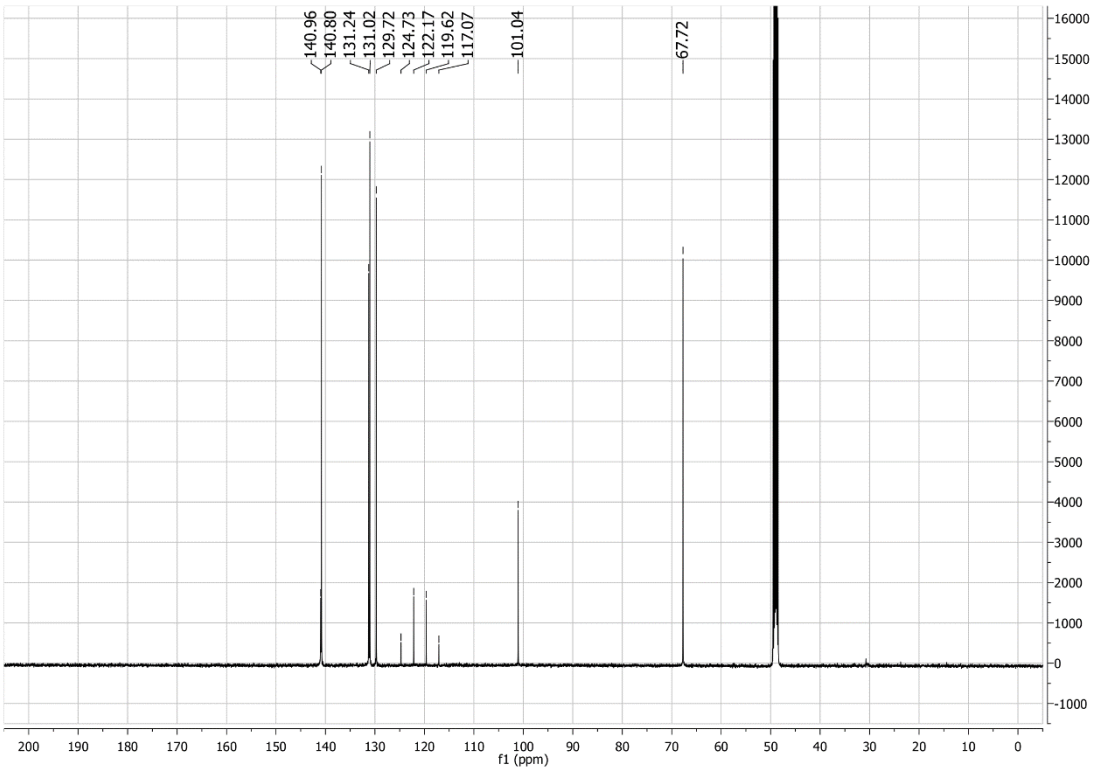


2-I-C_6H_4 bistriflimide **2k**

^{1}H NMR (500 MHz, CD_3OD)

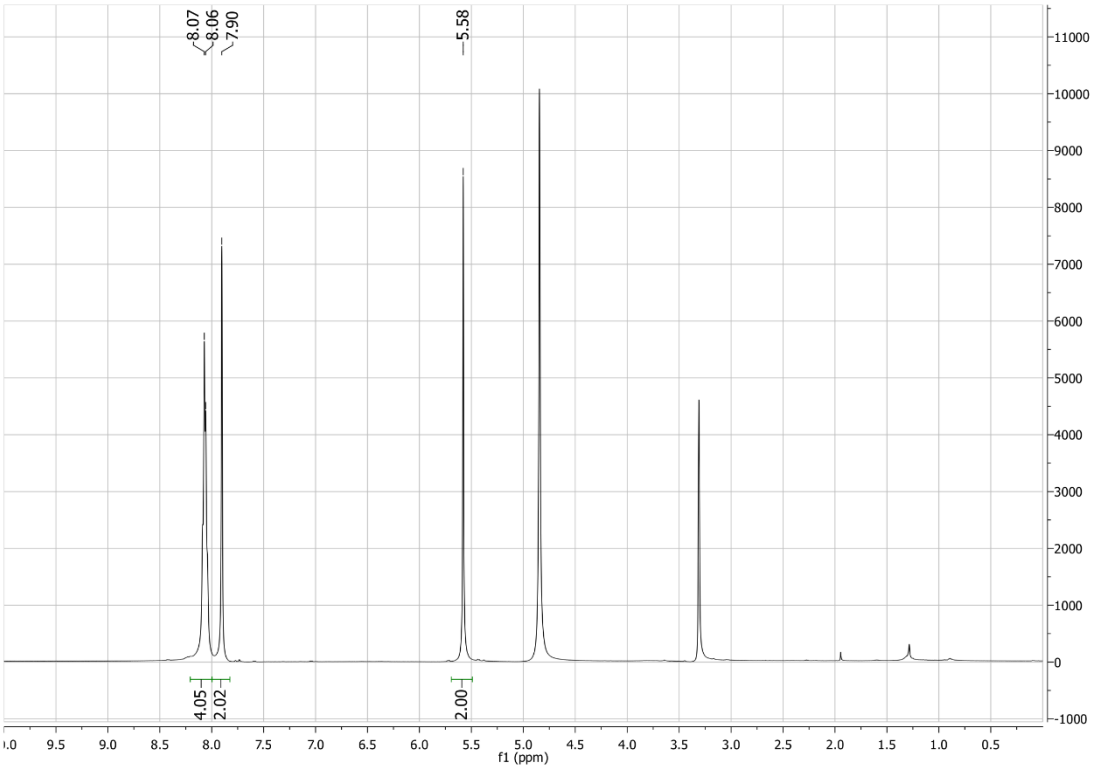


^{13}C NMR (126 MHz, CD_3OD)

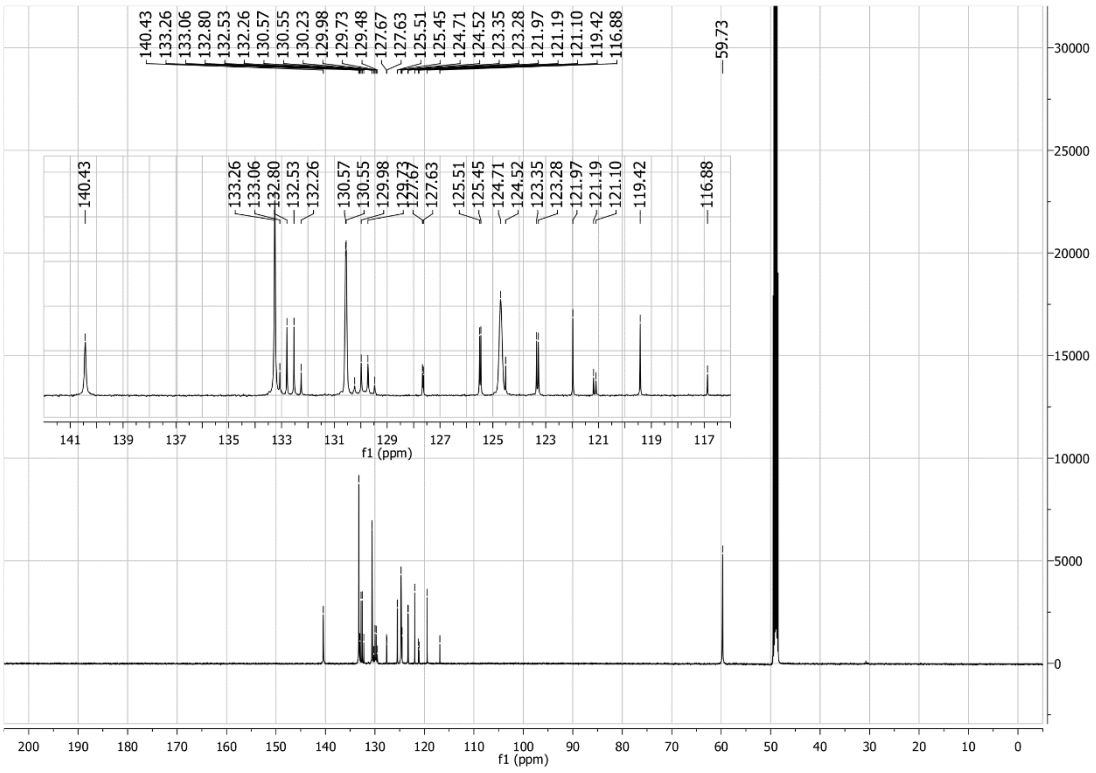


2,4-$(CF_3)_2$-C_6H_3 bistriflimide **2m**

^{1}H NMR (500 MHz, CD_3OD)

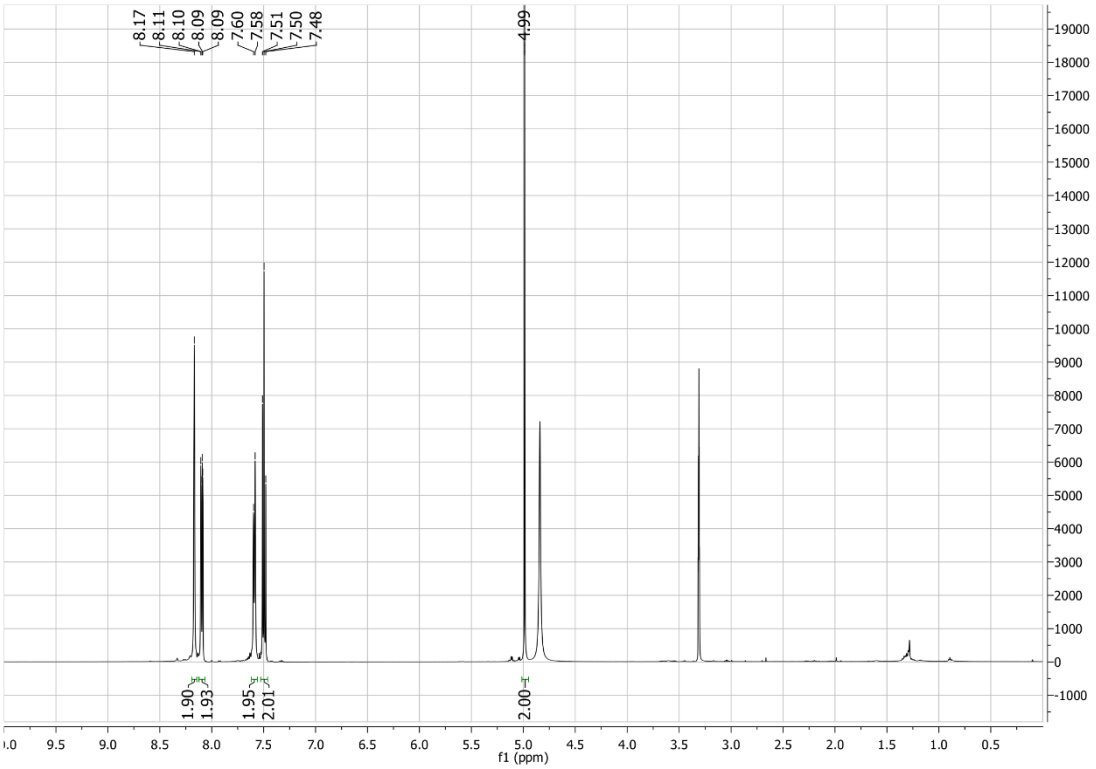


^{13}C NMR (126 MHz, CD_3OD)

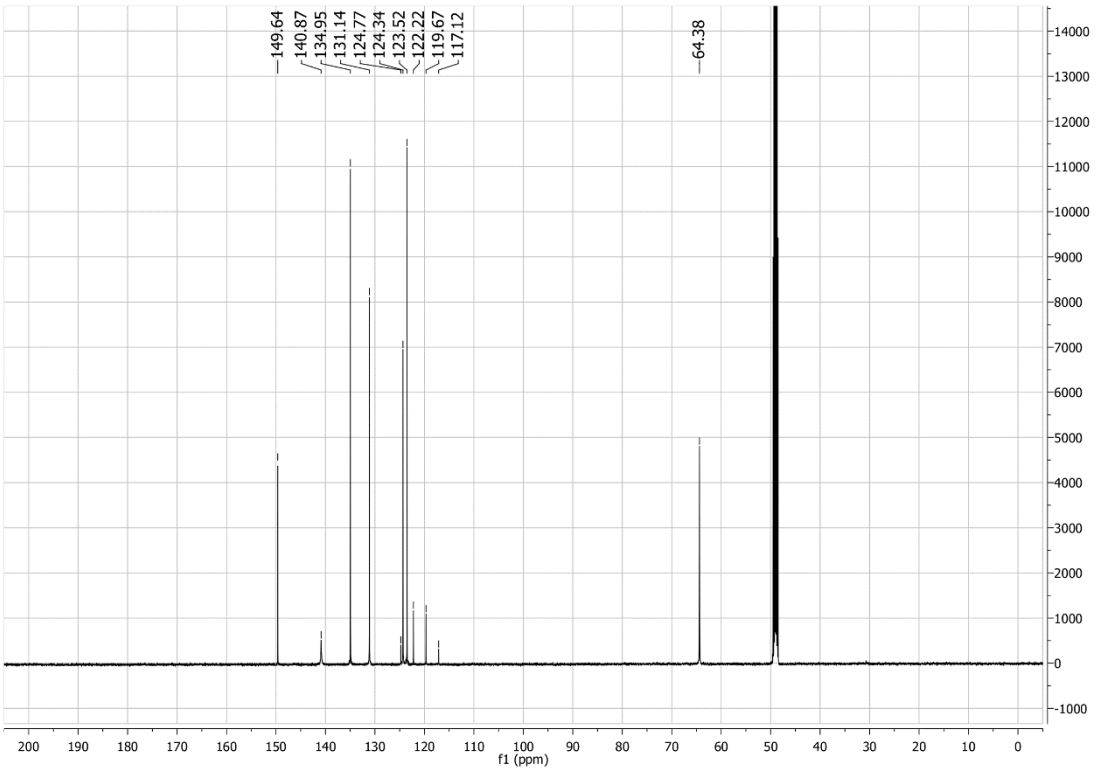


3-NO_2-C_6H_4 bistriflimide **2n**

^{1}H NMR (500 MHz, CD_3OD)

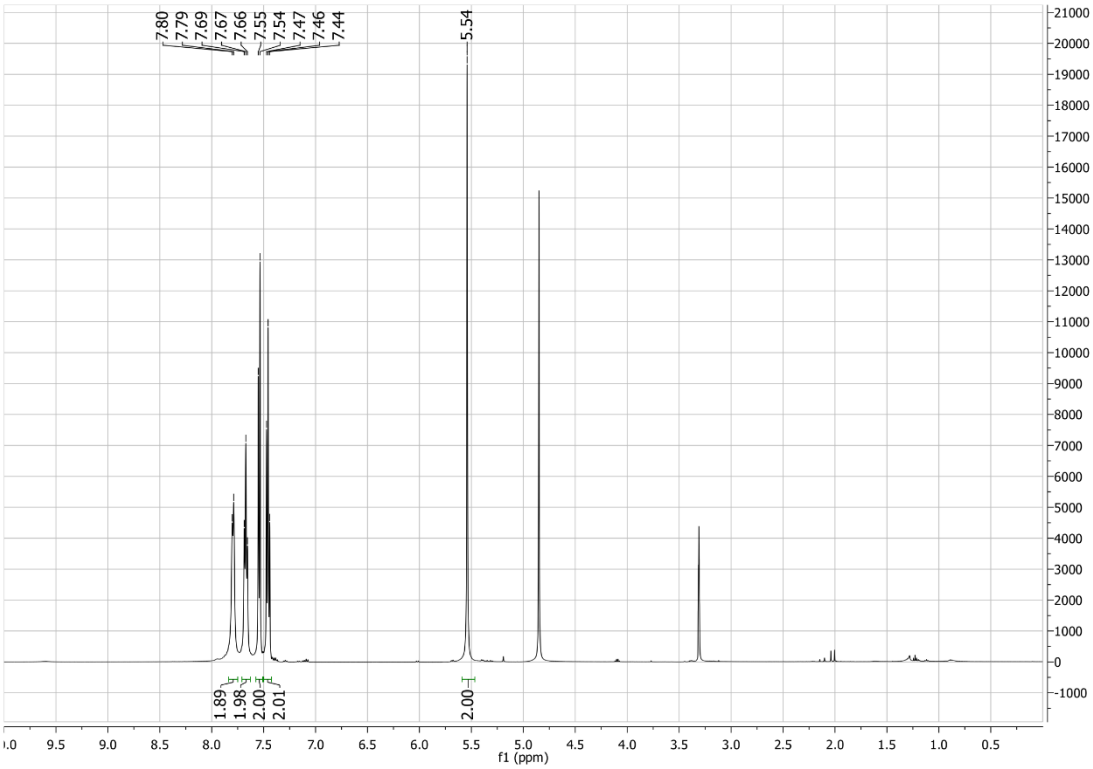


^{13}C NMR (126 MHz, CD_3OD)

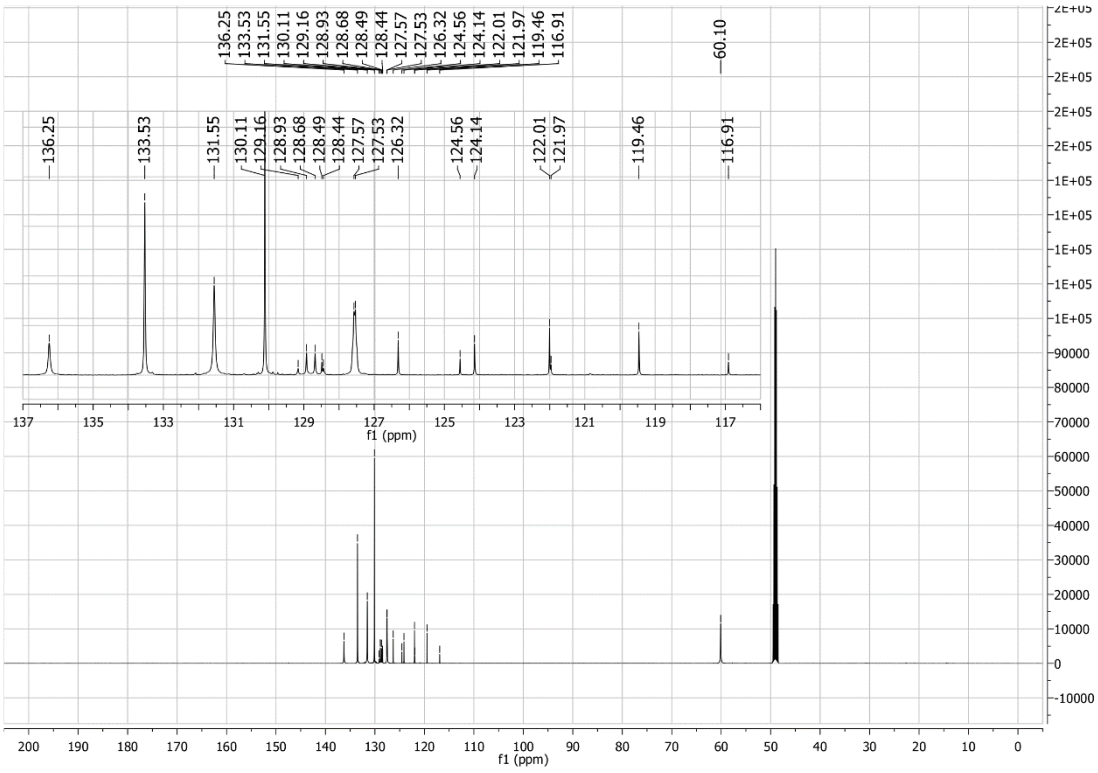


2-CF_3-C_6H_4 bistriflimide **2o**

^{1}H NMR (500 MHz, CD_3OD)

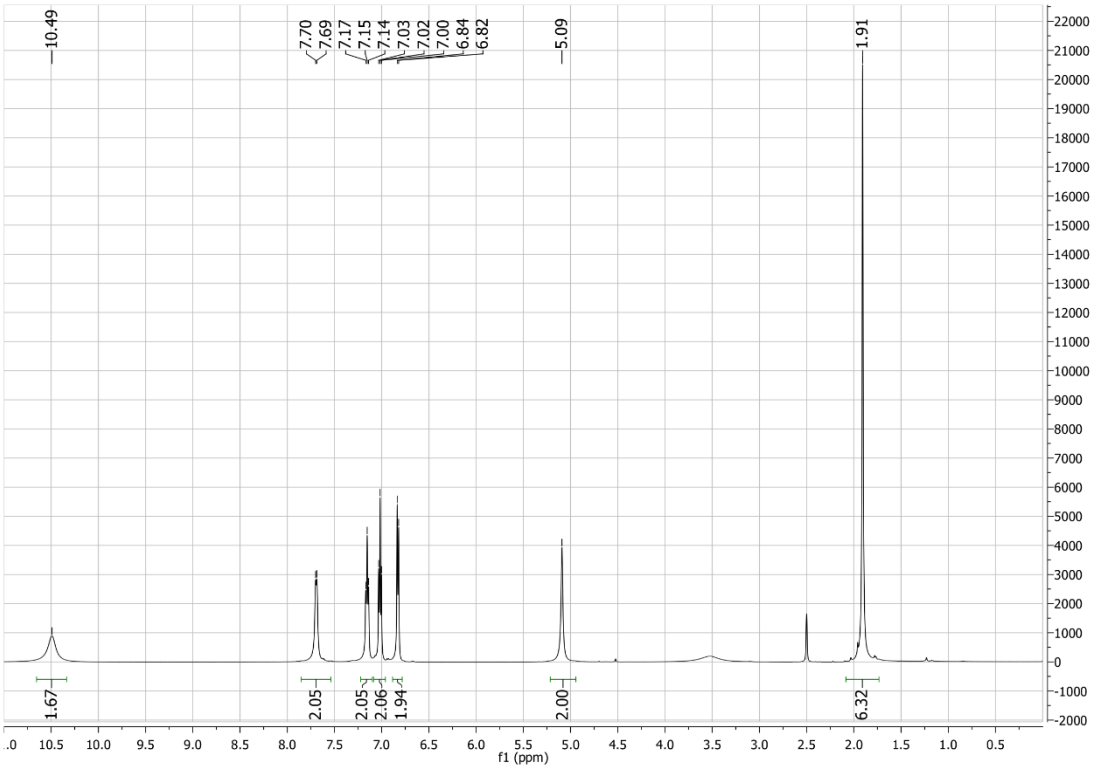


^{13}C NMR (126 MHz, CD_3OD)

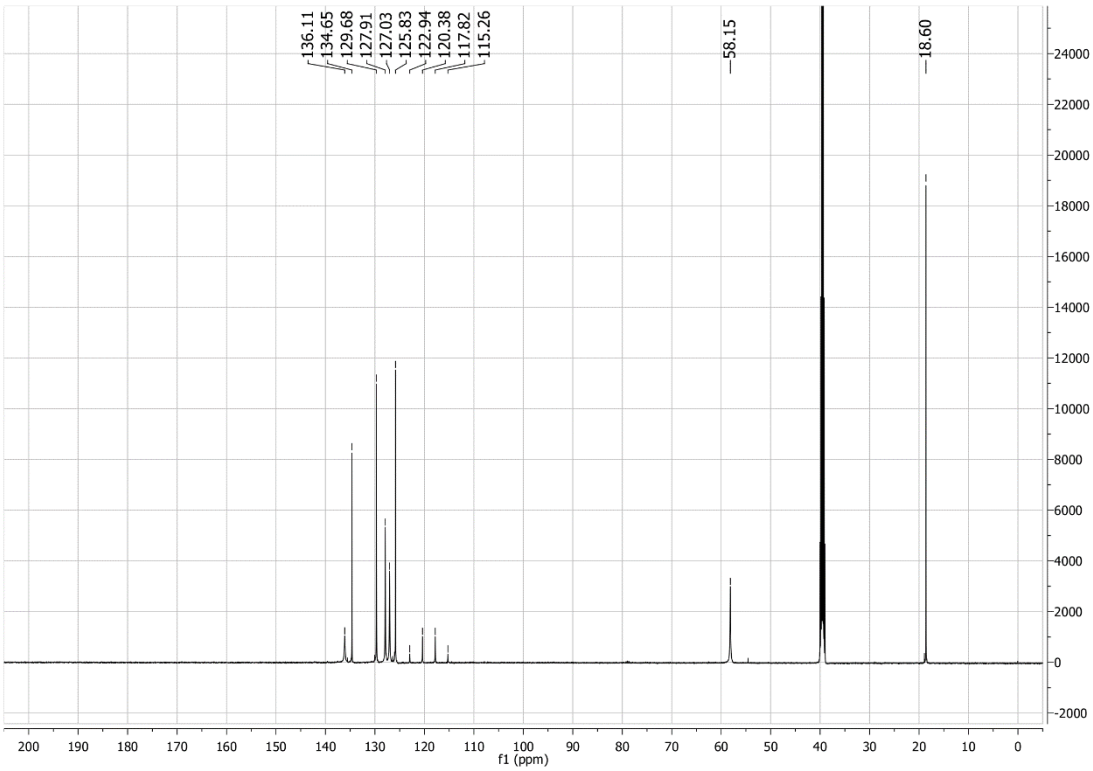


2-Me-C_6H_4 bistriflimide **2p**

^{1}H NMR (500 MHz, DMSO-d_6)

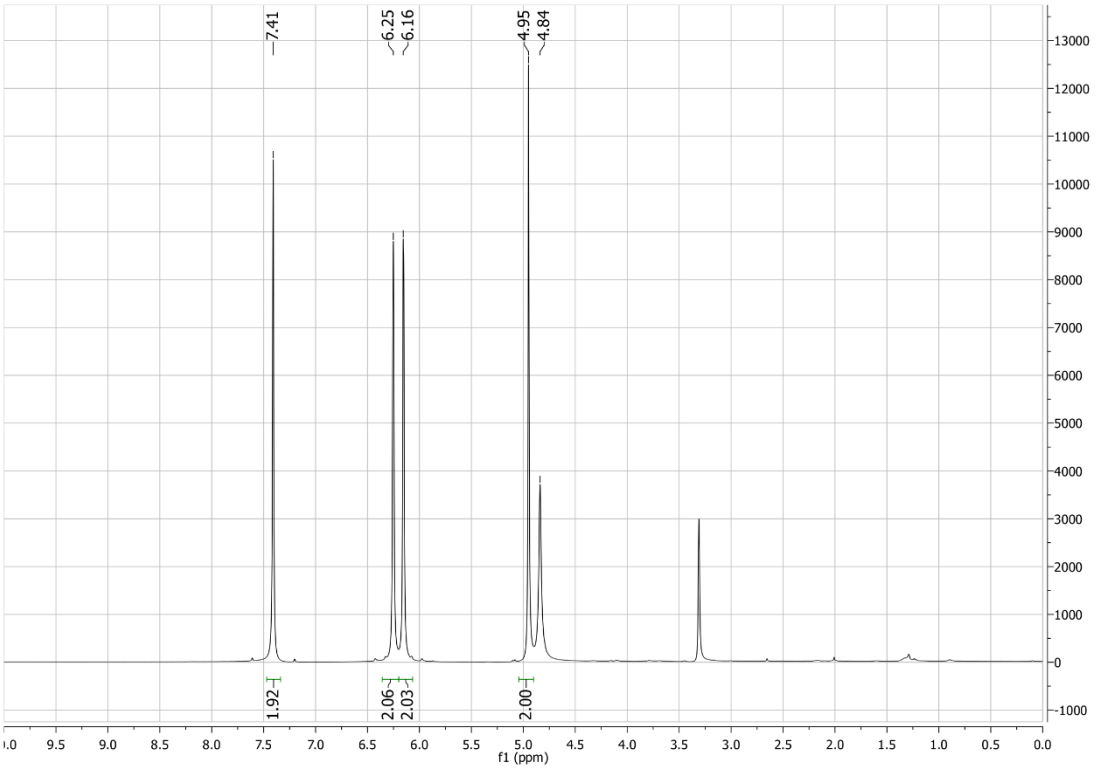


^{13}C NMR (126 MHz, DMSO-d_6)

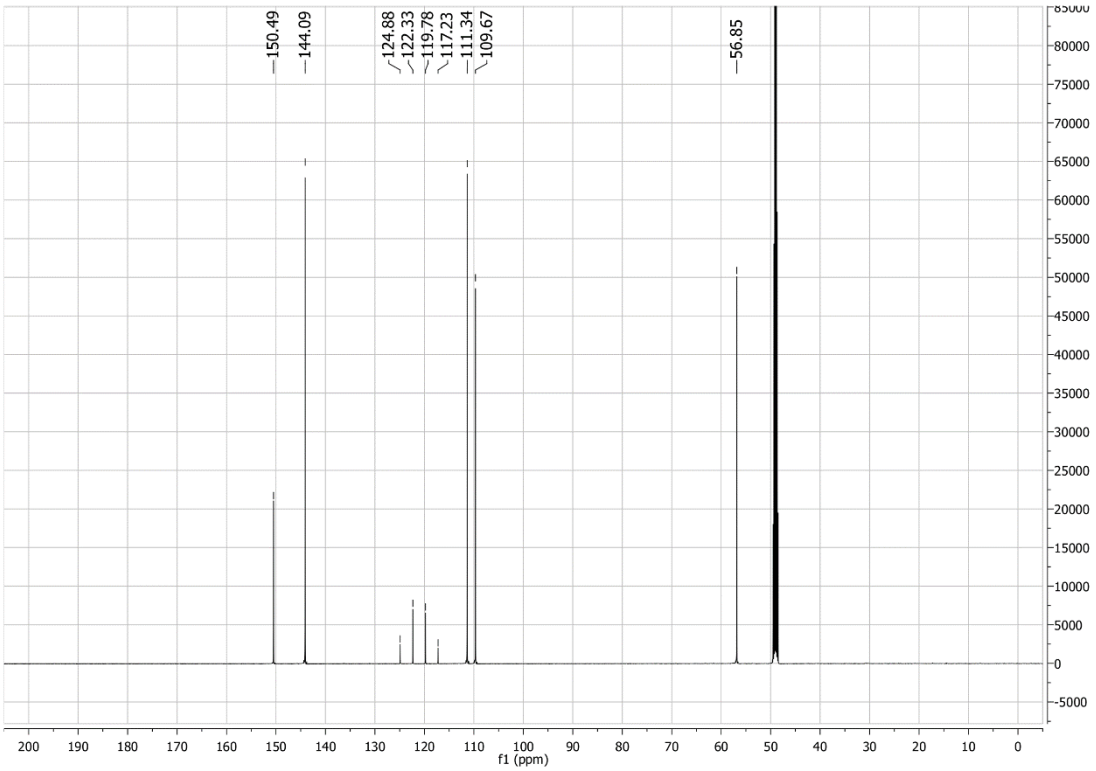


2-furyl bistriflimide **2r**

^{1}H NMR (500 MHz, CD_3OD)

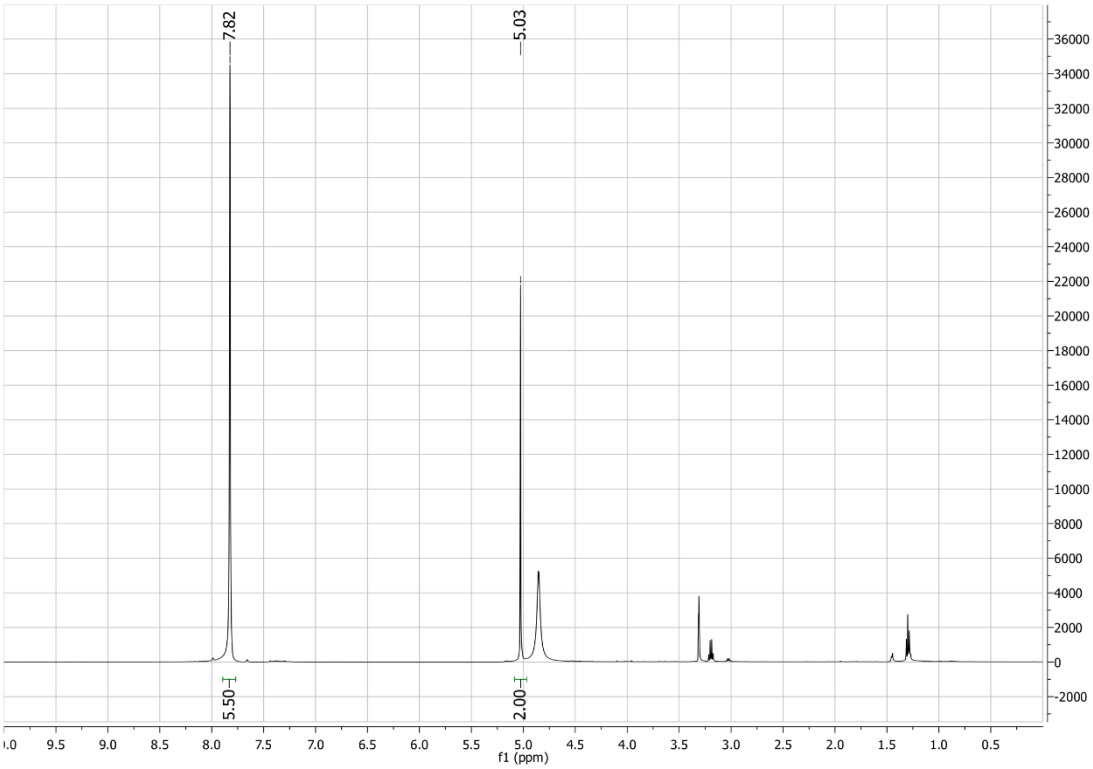


^{13}C NMR (126 MHz, CD_3OD)

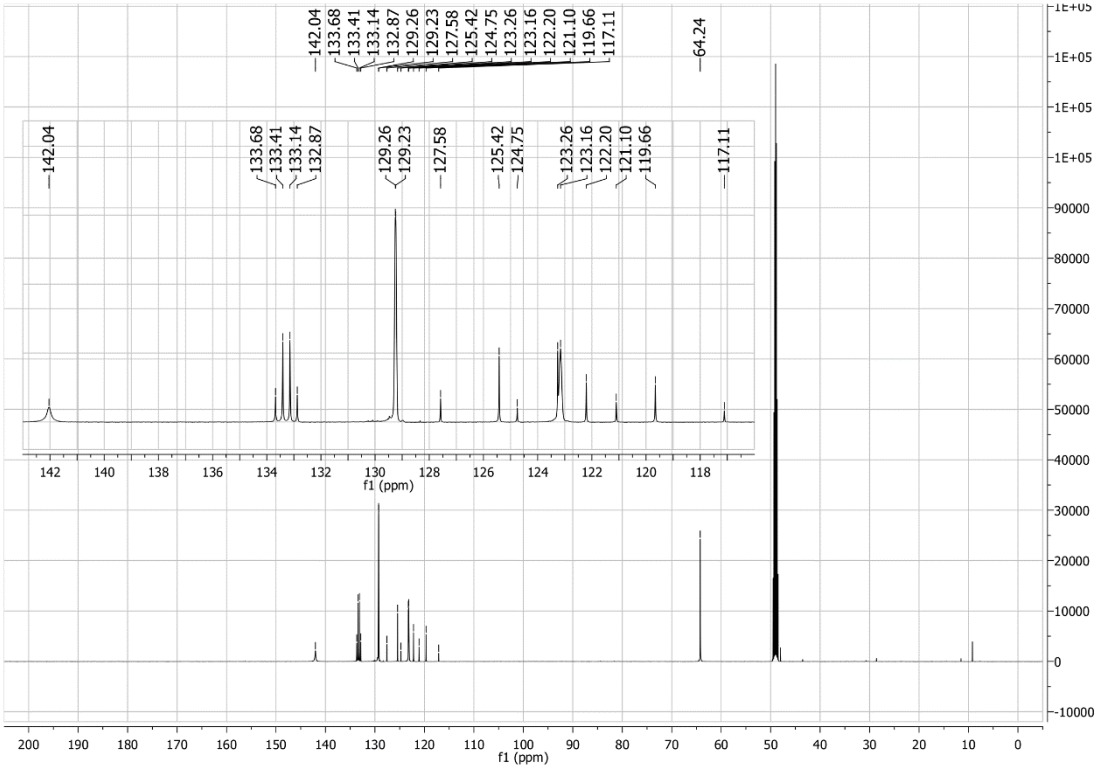


3,5-$(CF_3)_2$-C_6H_3 bistriflimide **2s**

^{1}H NMR (500 MHz, CD_3OD)

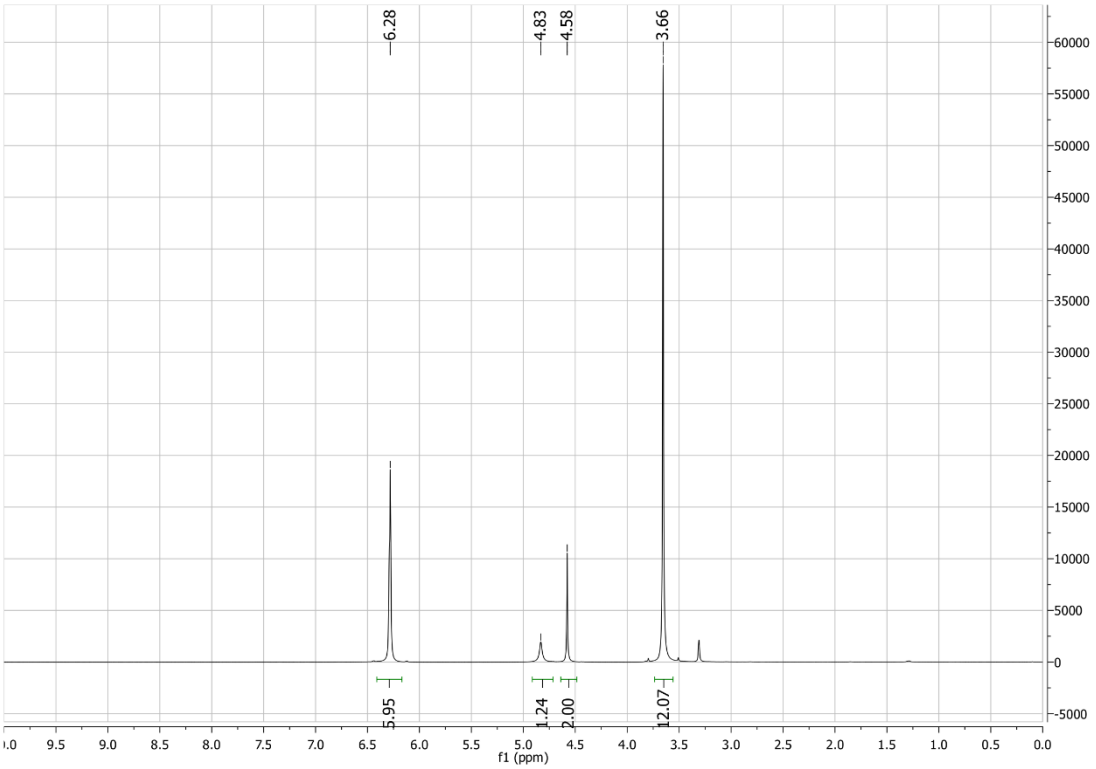


^{13}C NMR (126 MHz, CD_3OD)

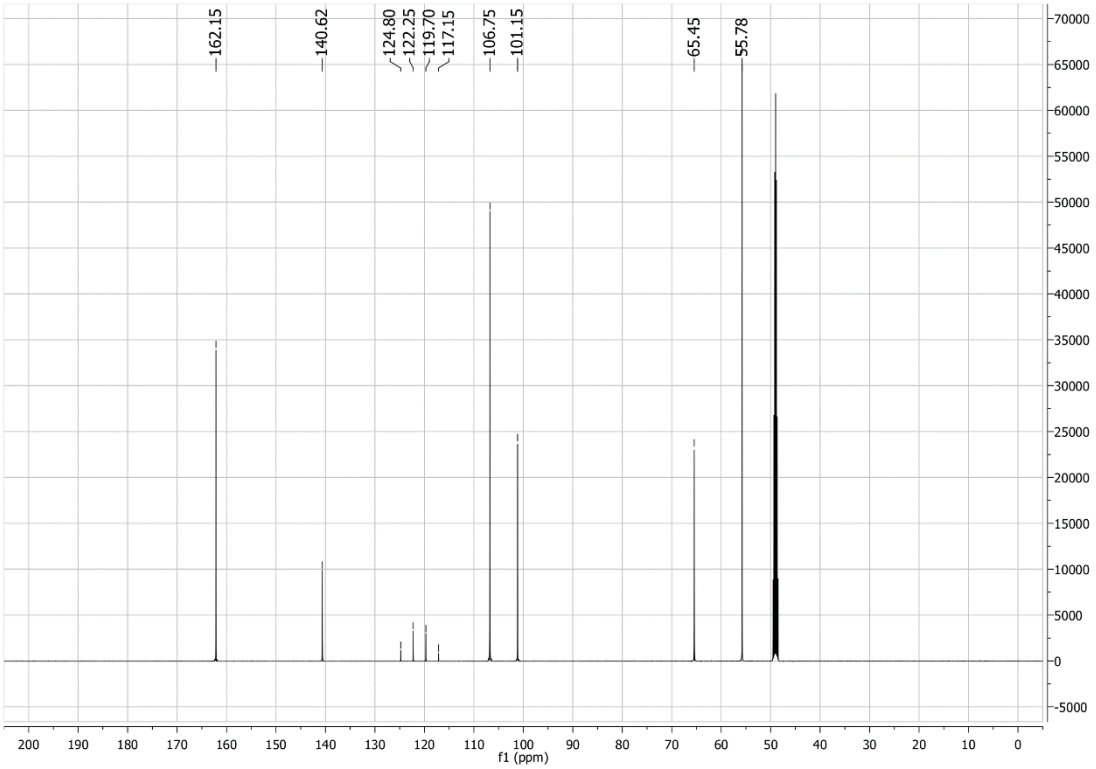


2,4-$(OMe)_2$-C_6H_3 bistriflimide **2t**

^{1}H NMR (500 MHz, CD_3OD)

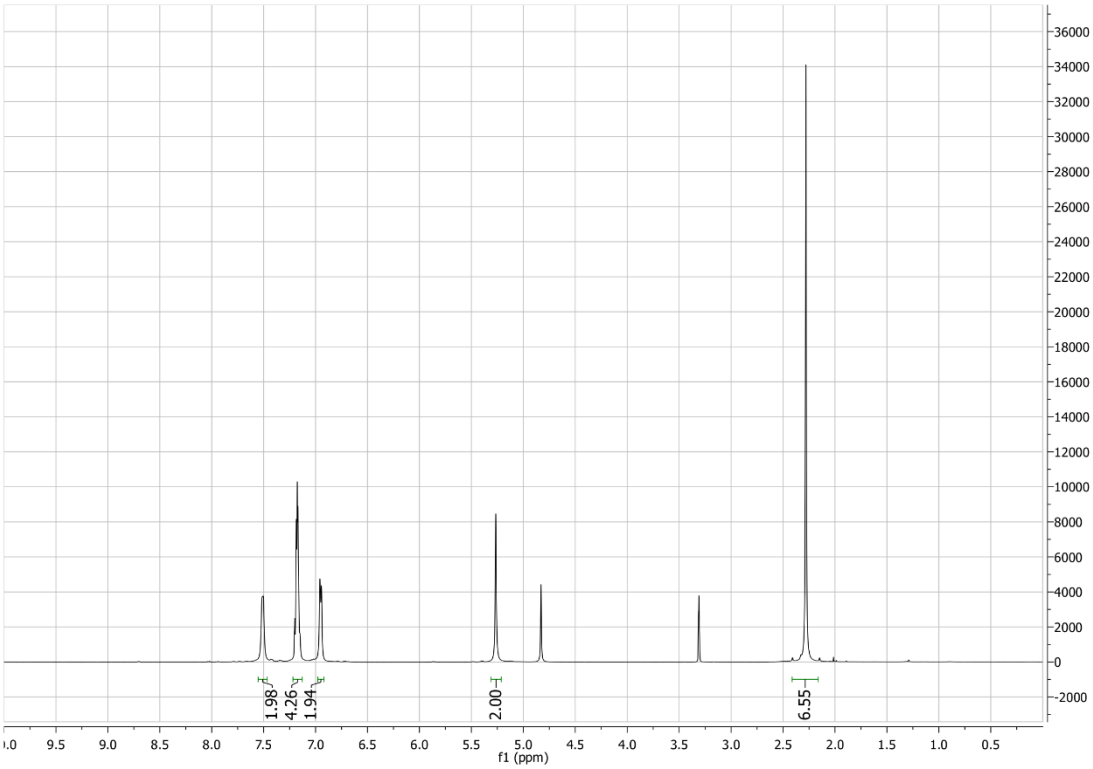


^{13}C NMR (126 MHz, CD_3OD)

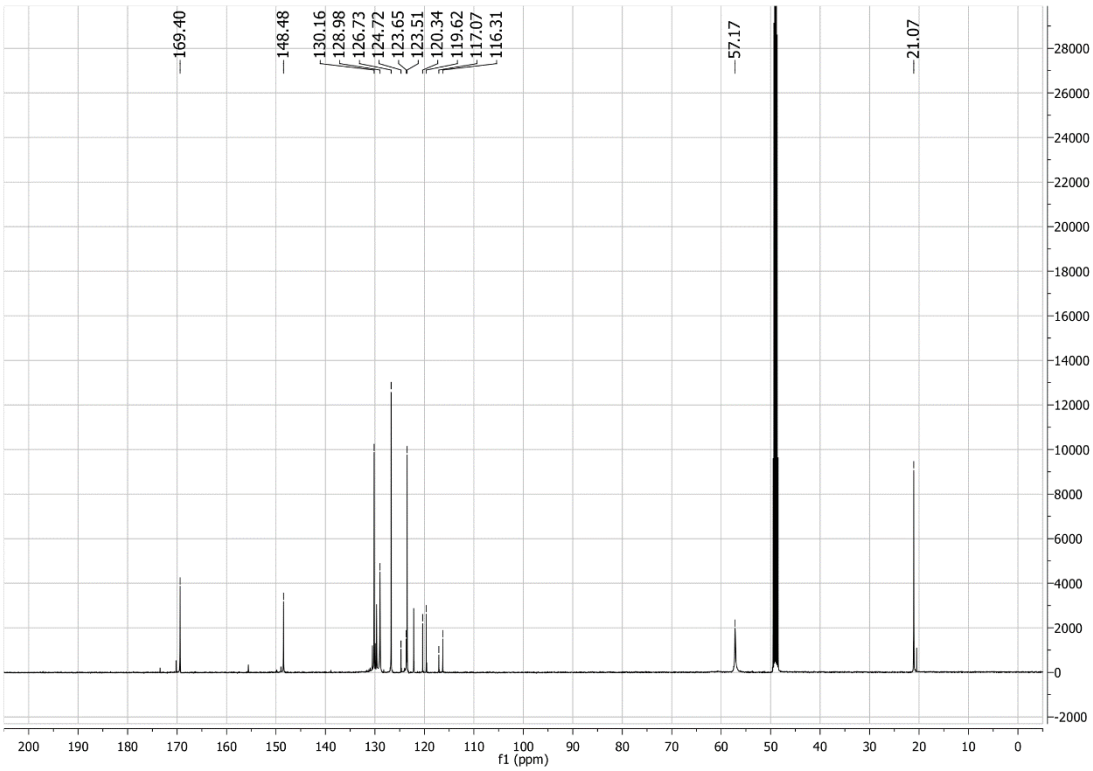


2-OAc-C_6H_4 bistriflimide **2u**

^{1}H NMR (500 MHz, CD_3OD)

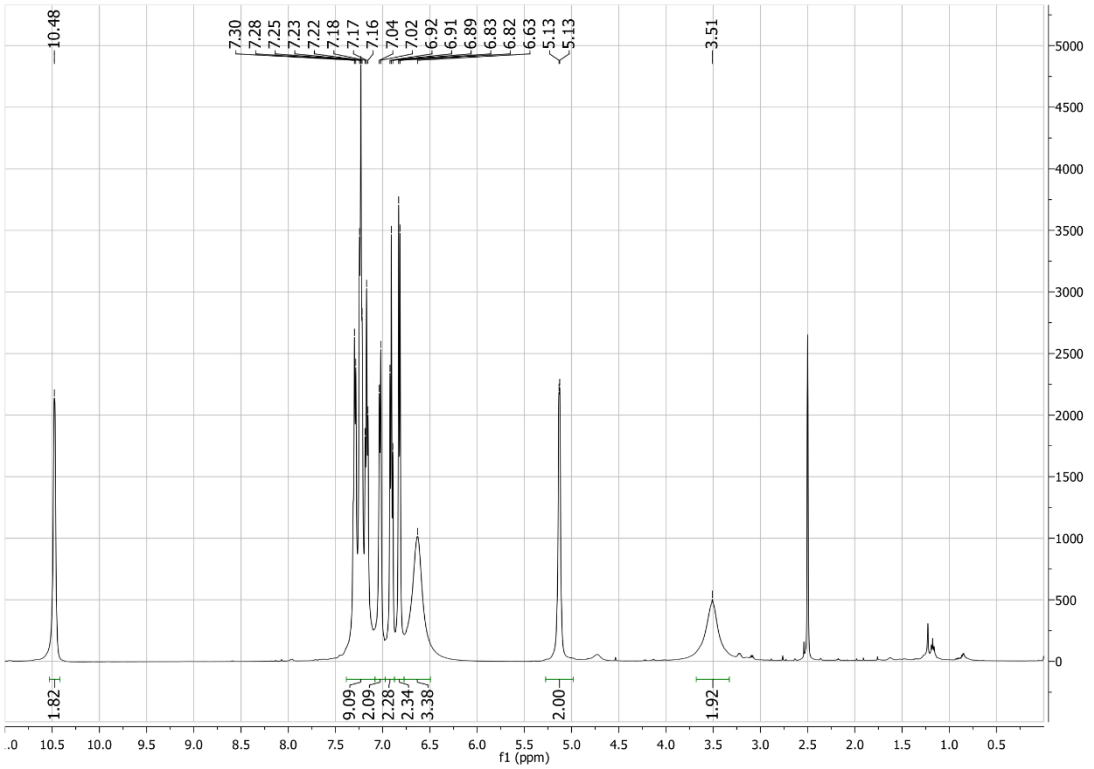


^{13}C NMR (126 MHz, CD_3OD)

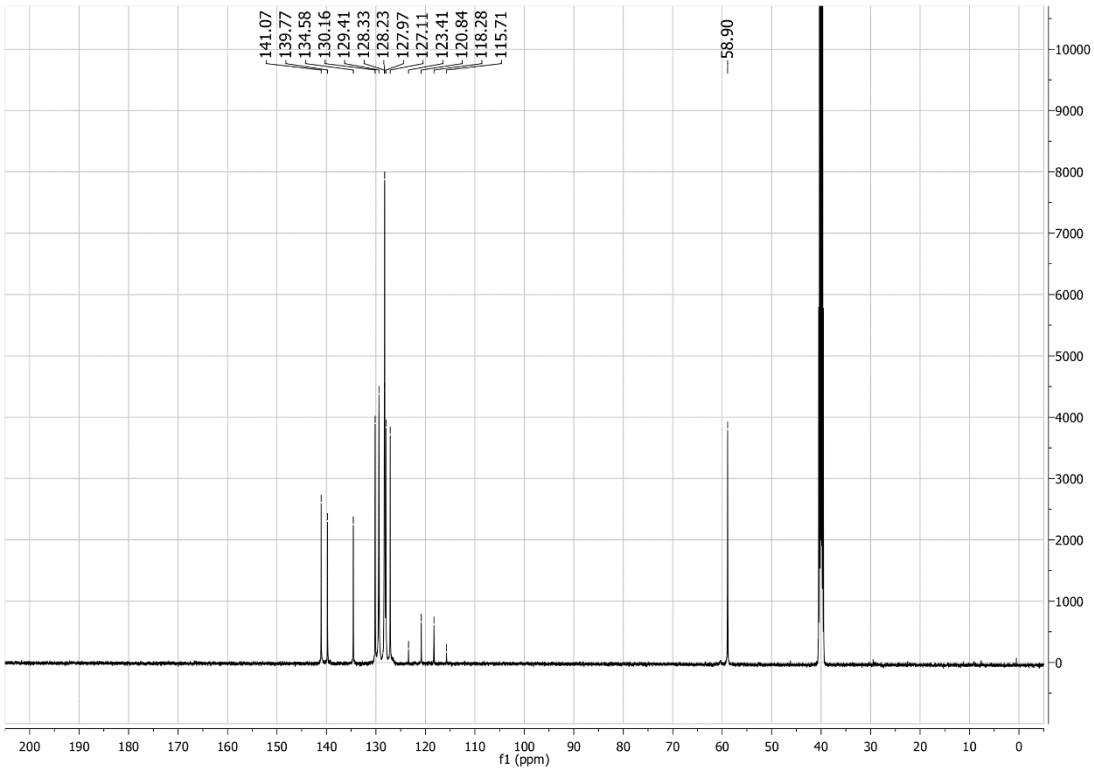


2-Ph-C_6H_4 bistriflimide **2v**

^{1}H NMR (500 MHz, DMSO-d_6)

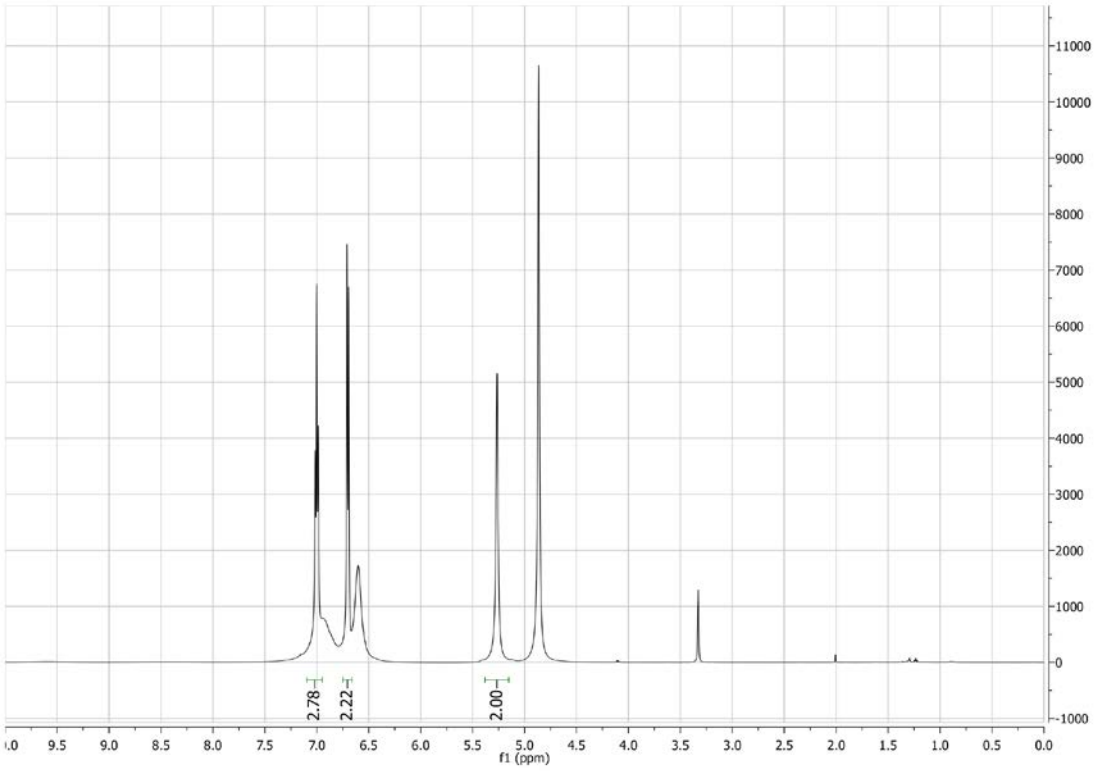


^{13}C NMR (126 MHz, DMSO-d_6)

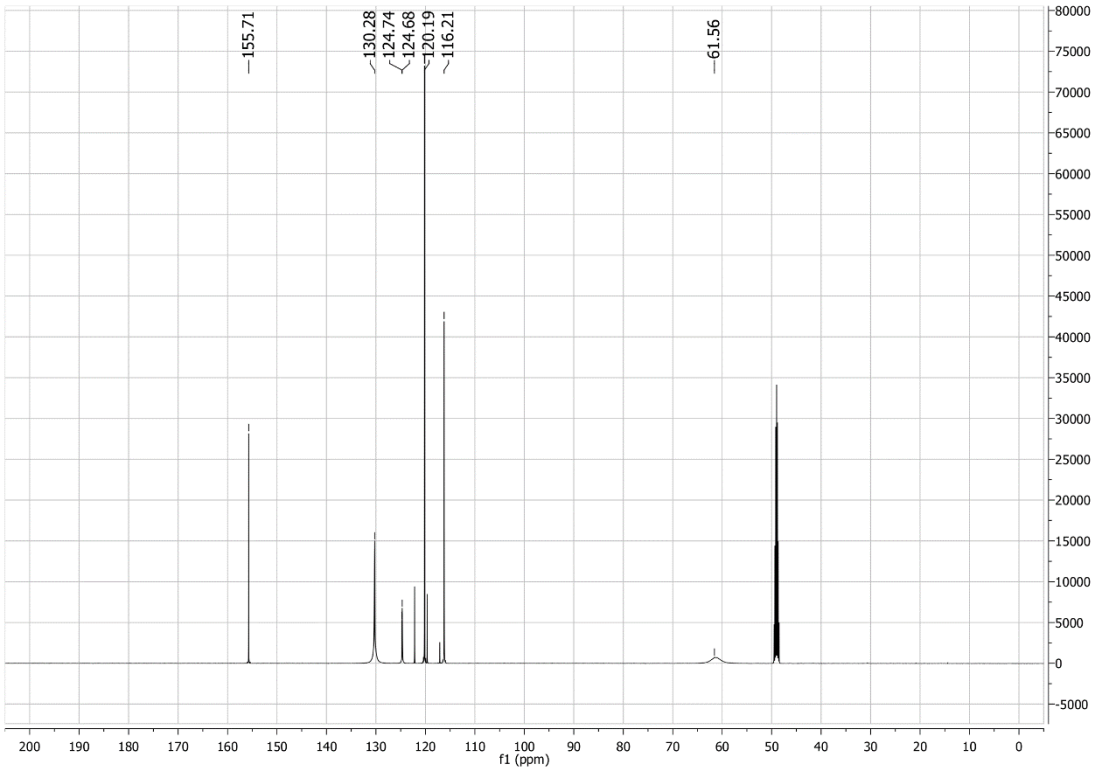


2-OH-C_6H_4 bistriflimide **2w**

^{1}H NMR (500 MHz, CD_3OD)

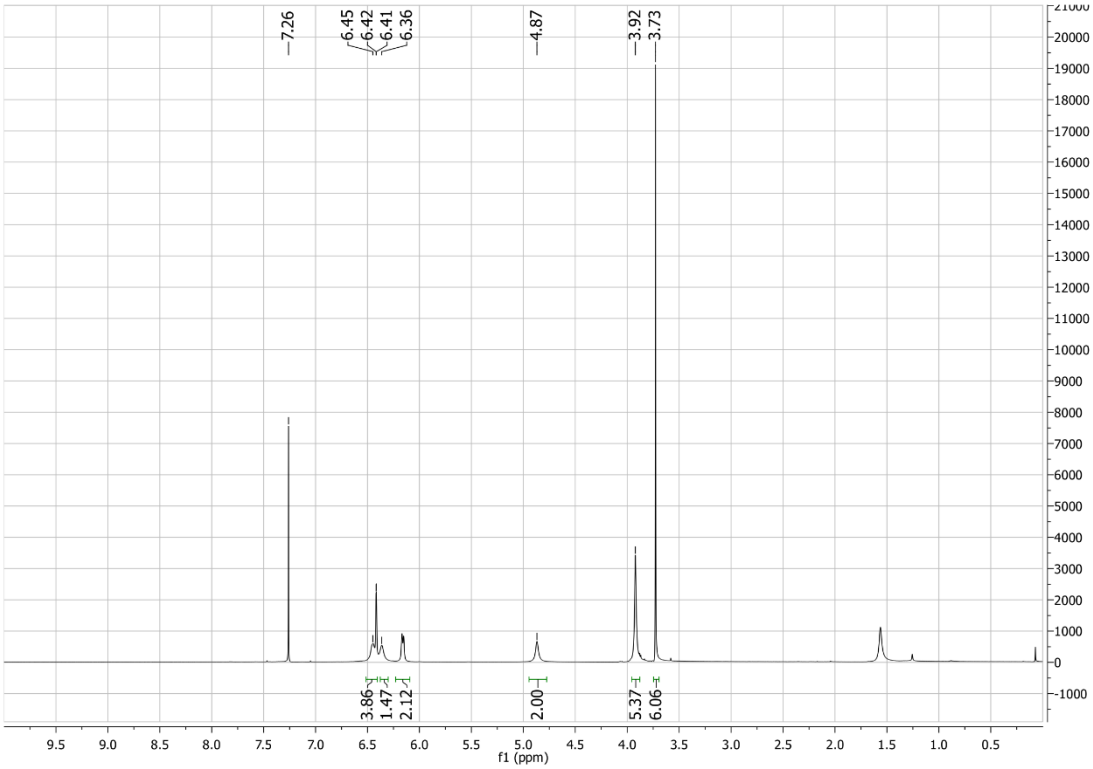


^{13}C NMR (126 MHz, CD_3OD)

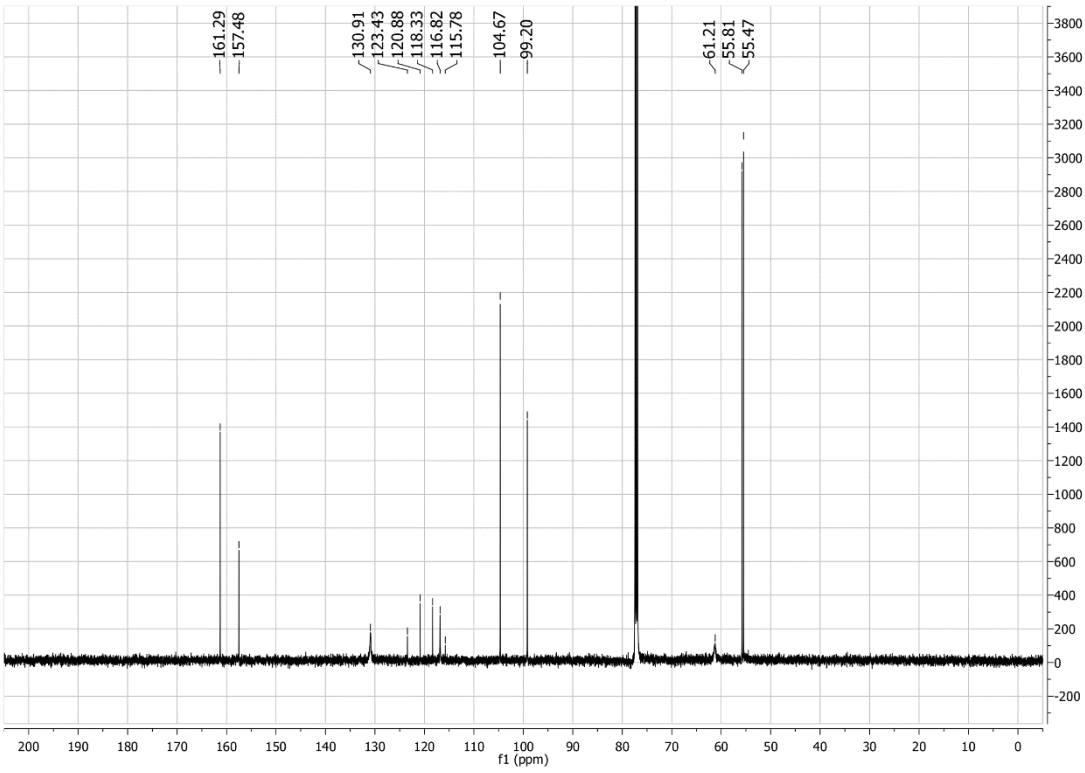


2,4-$(OMe)_2$-C_6H_3 bistriflimide **2x**

^{1}H NMR (500 MHz, $CDCl_3$)

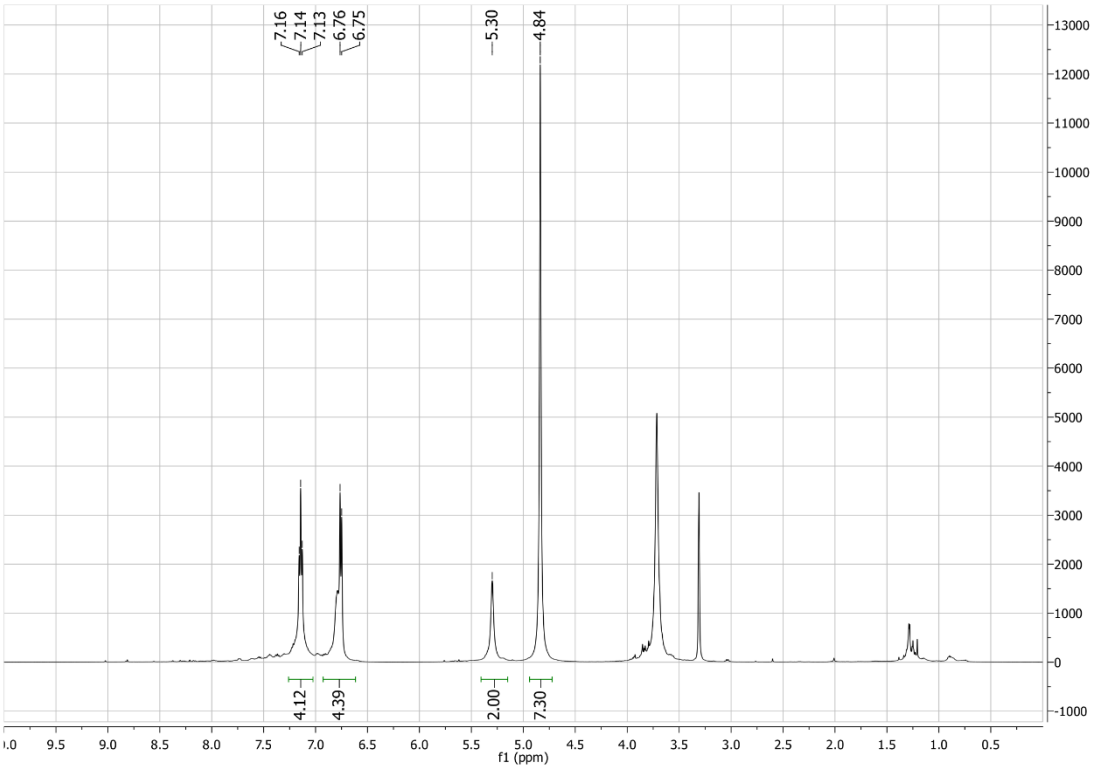


^{13}C NMR (126 MHz, $CDCl_3$)

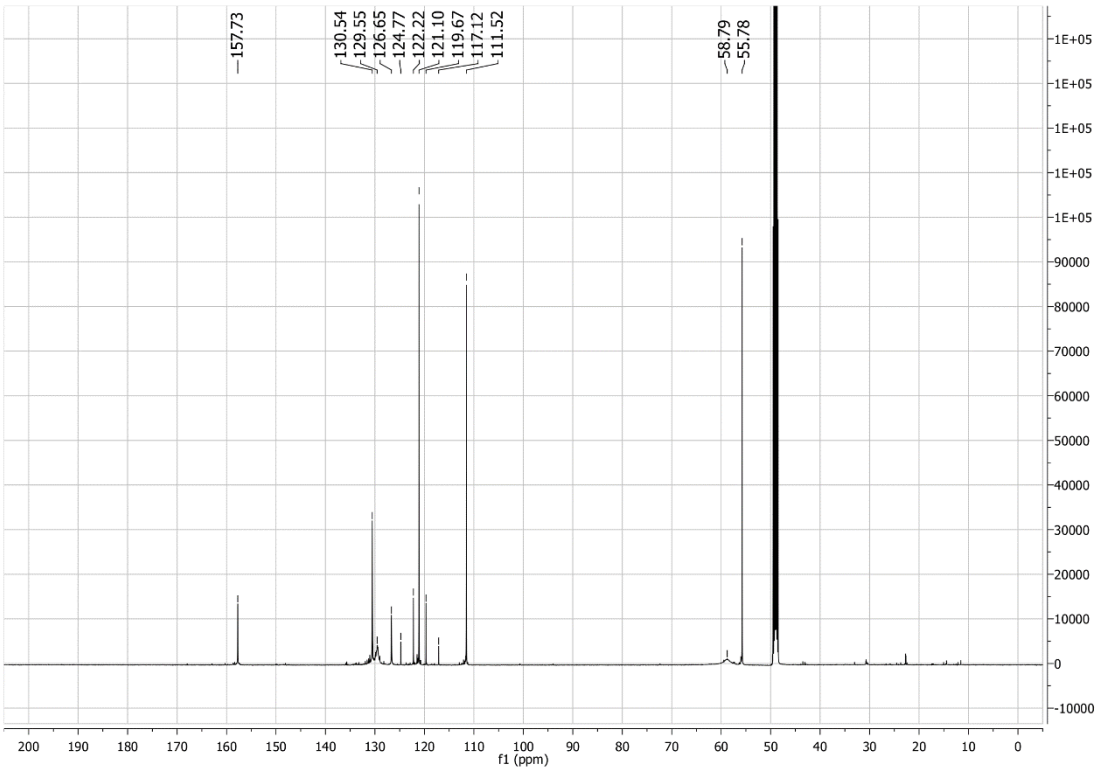


2-OMe-C_6H_4 bistriflimide **2y**

^{1}H NMR (500 MHz, CD_3OD)

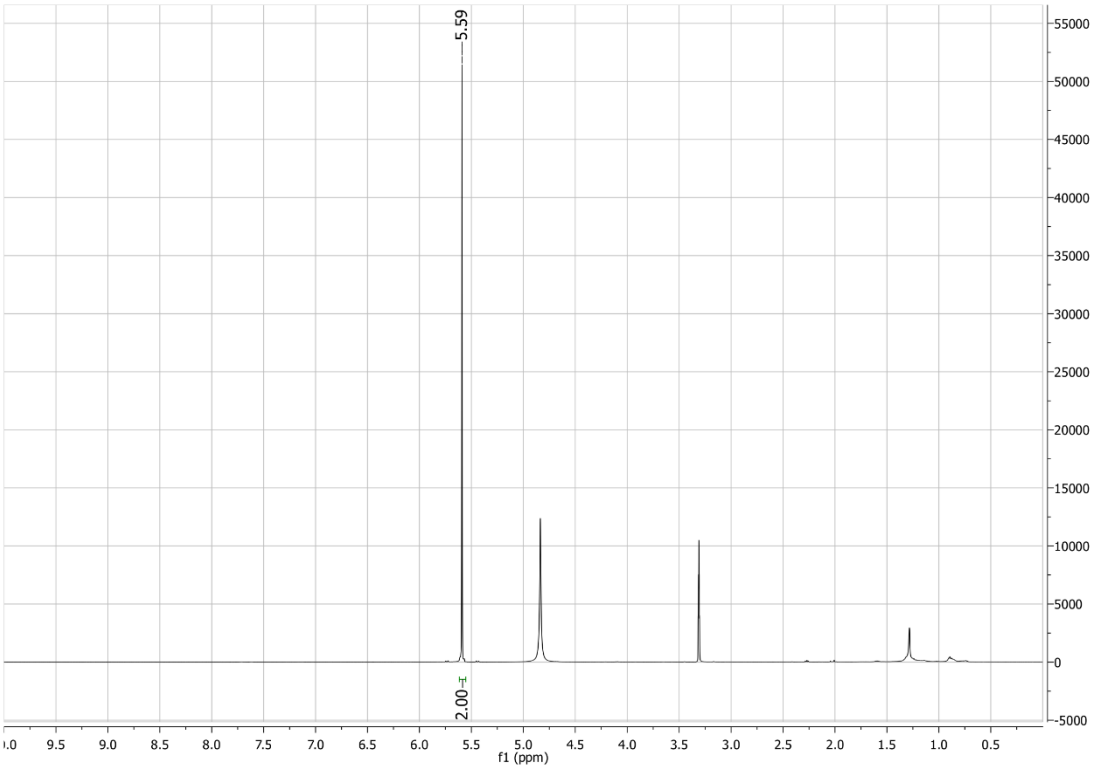


^{13}C NMR (126 MHz, CD_3OD)

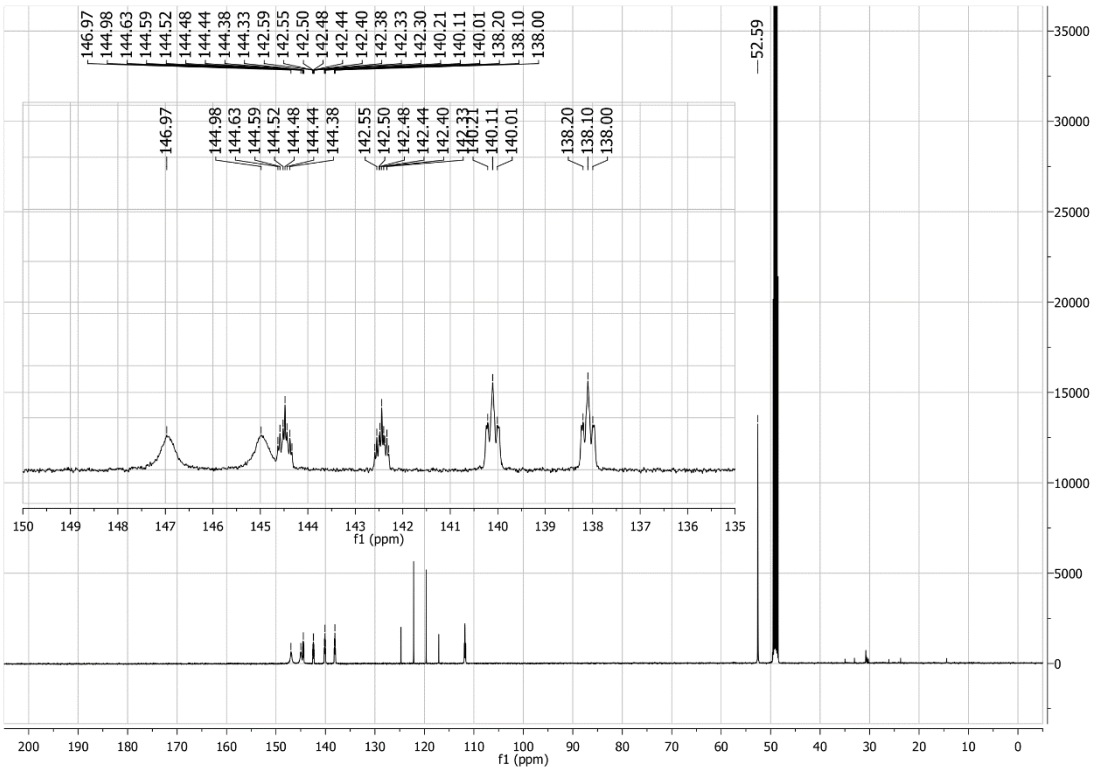


C_6F_5 bistriflimide **2z**

^{1}H NMR (500 MHz, CD_3OD)

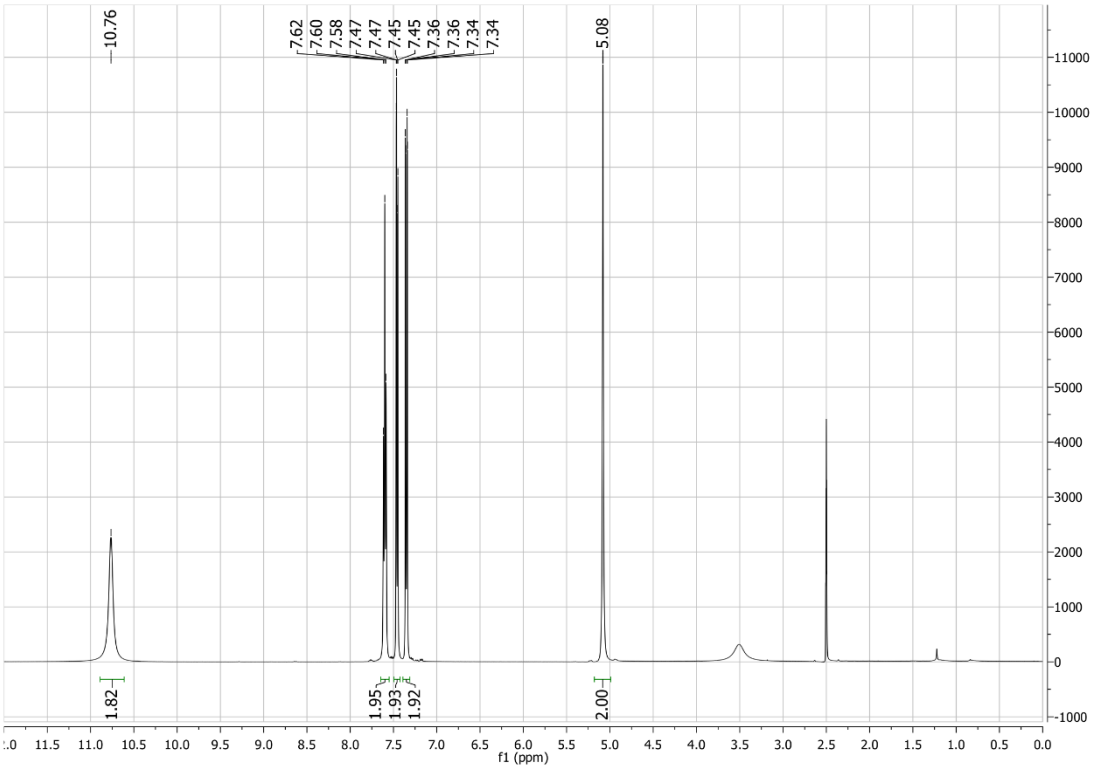


^{13}C NMR (126 MHz, CD_3OD)

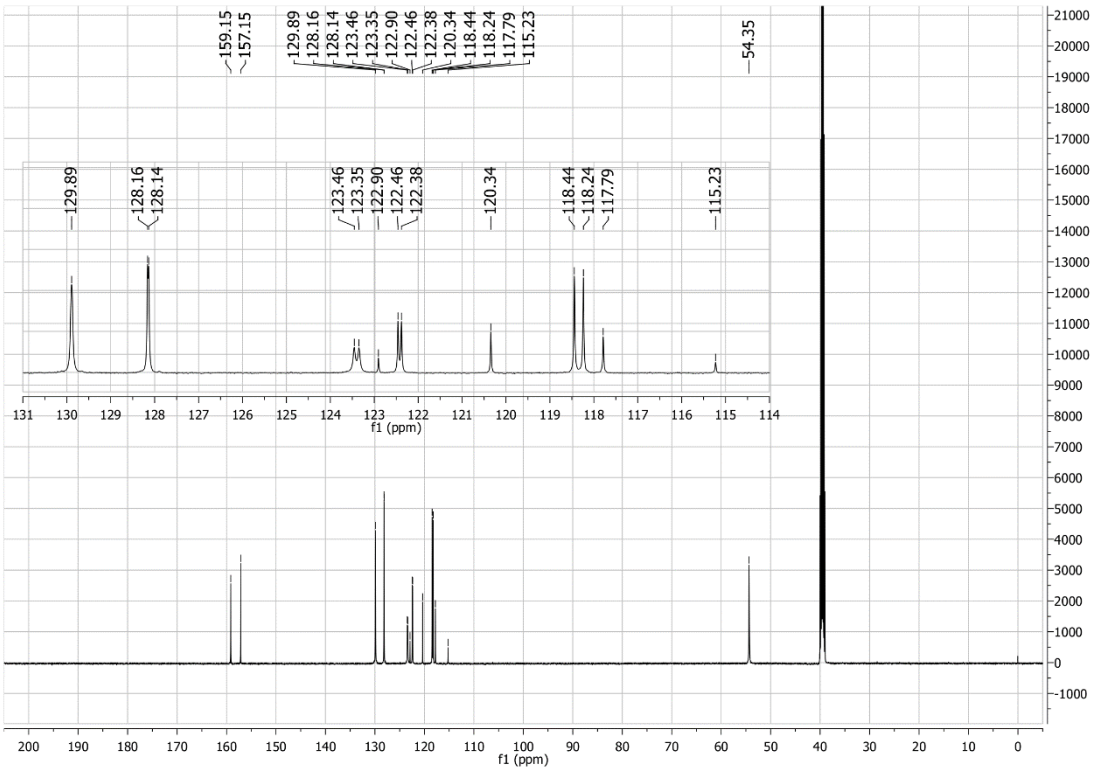


2-F-4-Br-C_6H_3 bistriflimide **2ab**

^{1}H NMR (500 MHz, DMSO-d_6)

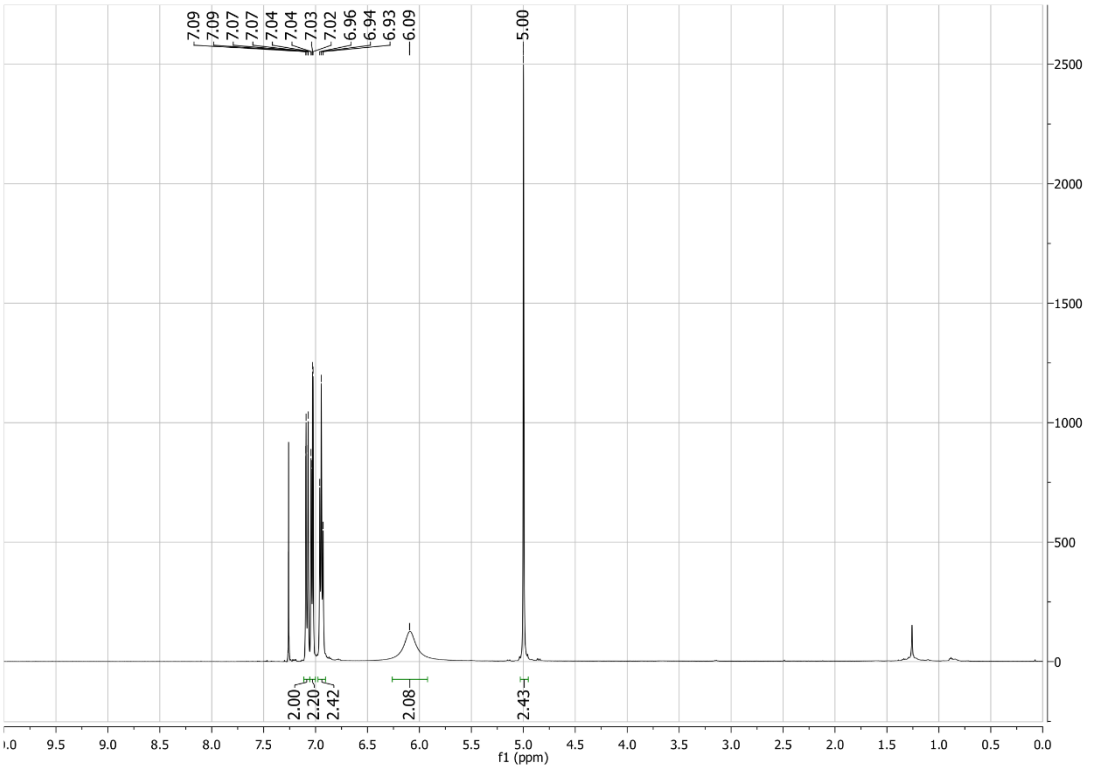


^{13}C NMR (126 MHz, DMSO-d_6)

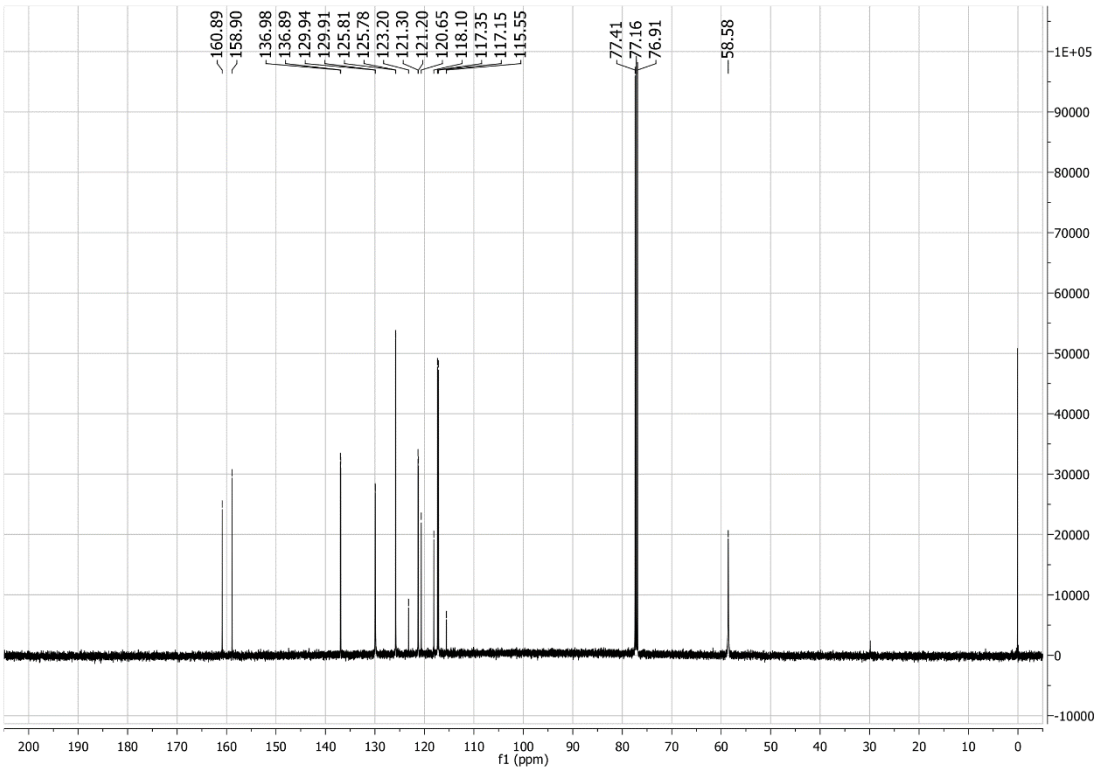


2-F-4-Cl-C_6H_3 bistriflimide **2ac**

^{1}H NMR (500 MHz, $CDCl_3$)

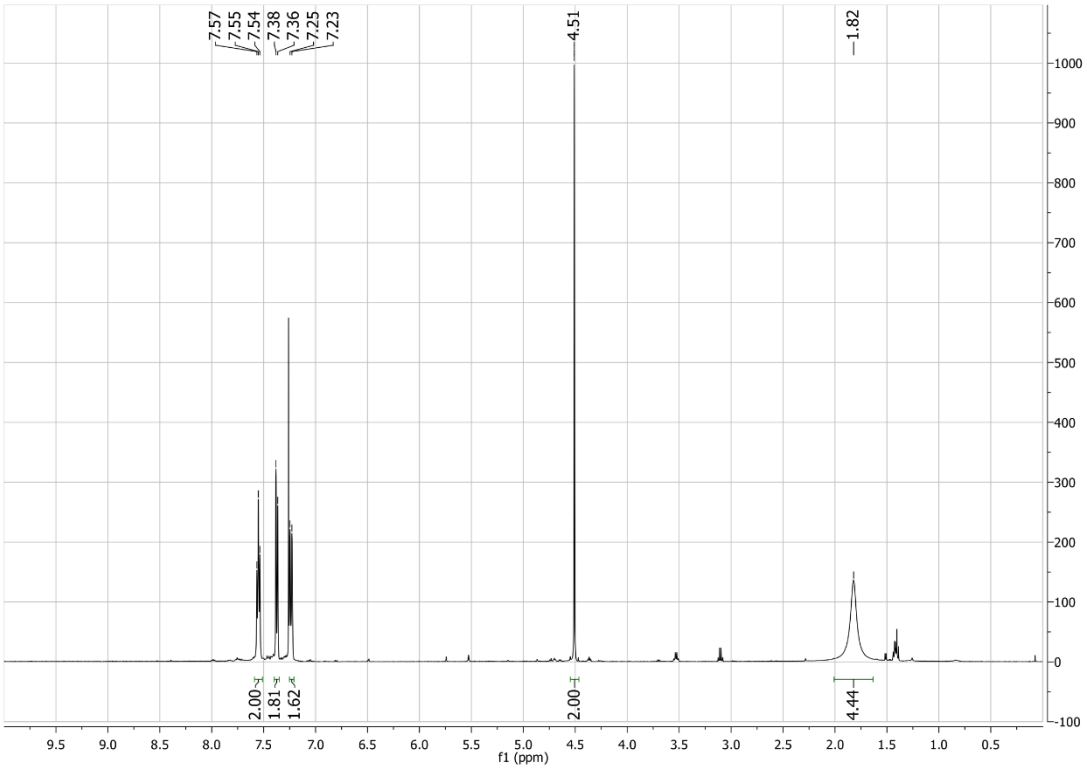


^{13}C NMR (126 MHz, $CDCl_3$)

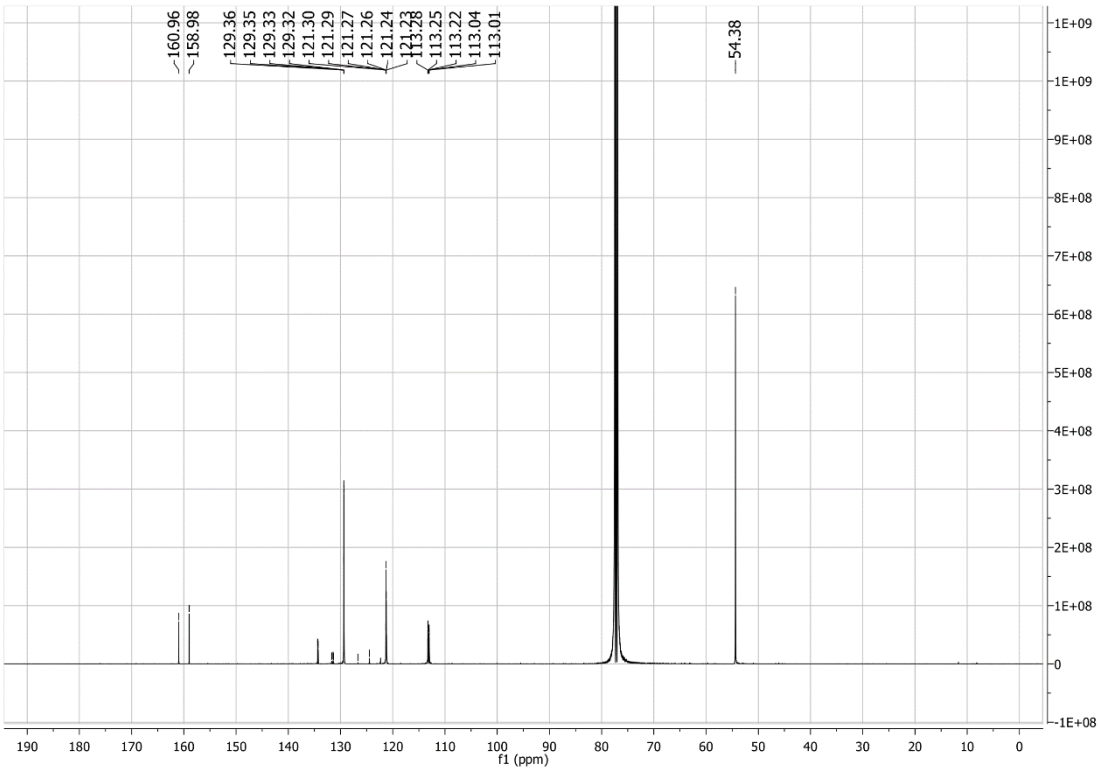


2-F-4-CF_3-C_6H_3 diamine **SI5ad**

^{1}H NMR (500 MHz, $CDCl_3$)

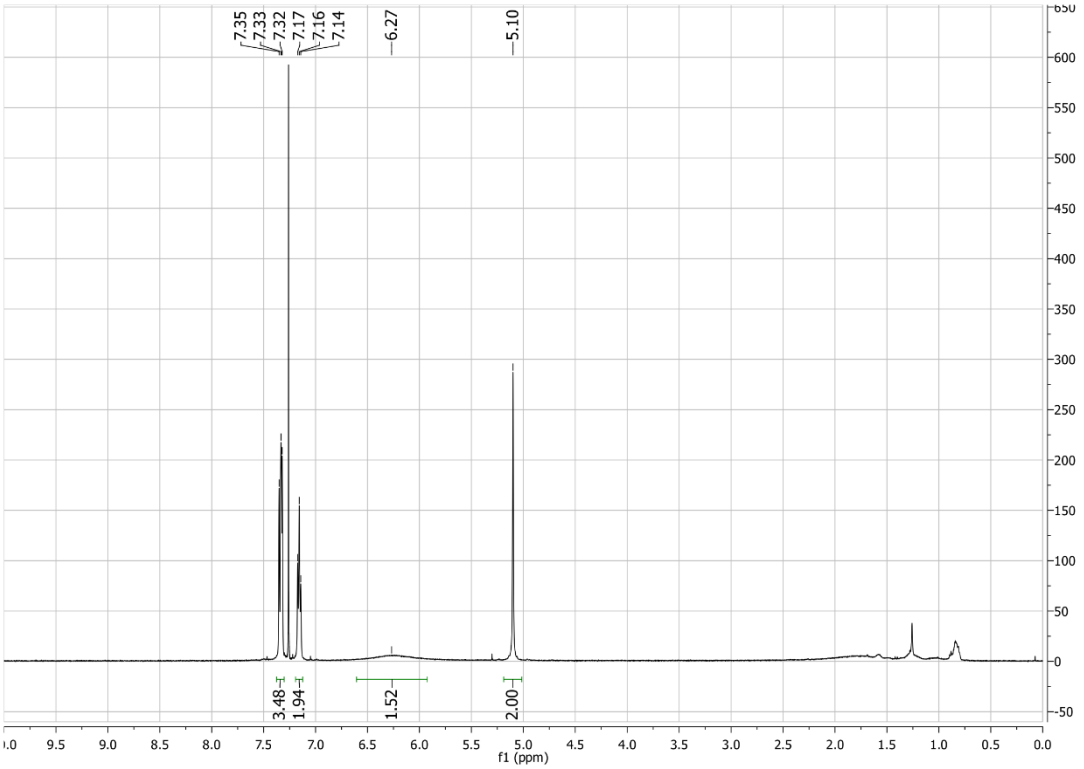


^{13}C NMR (126 MHz, $CDCl_3$)

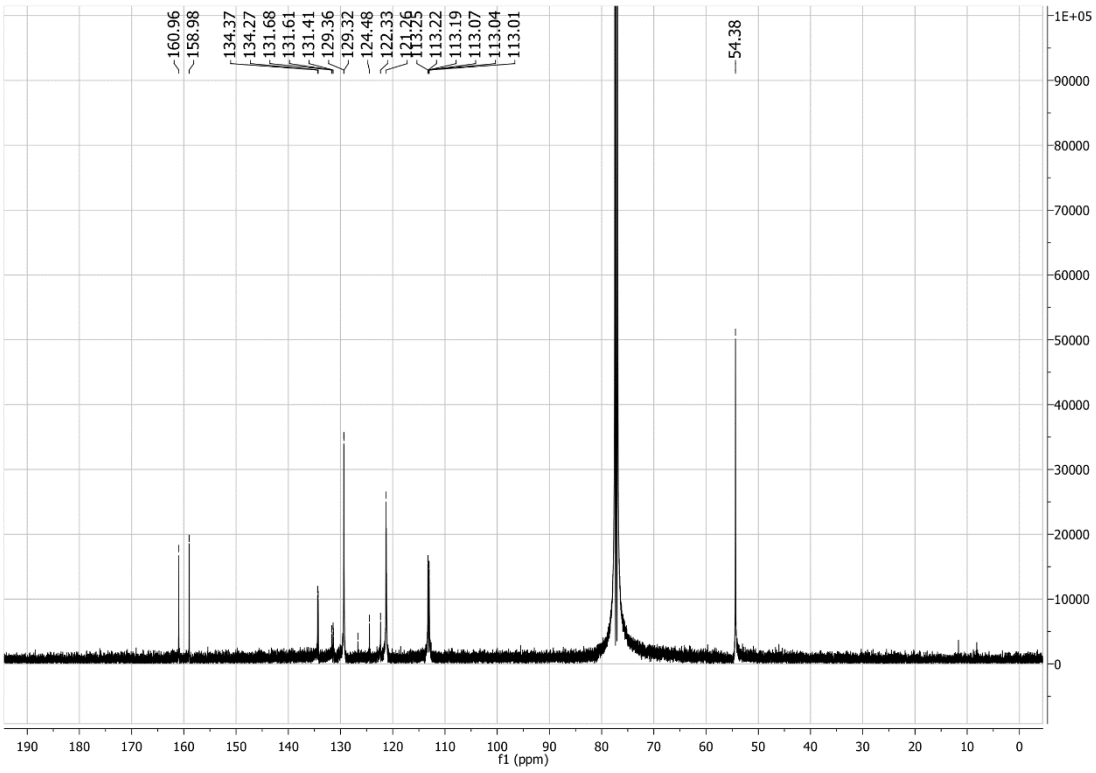


2-F-4-CF_3-C_6H_3 bistriflimide **2ad**

^{1}H NMR (500 MHz, $CDCl_3$)

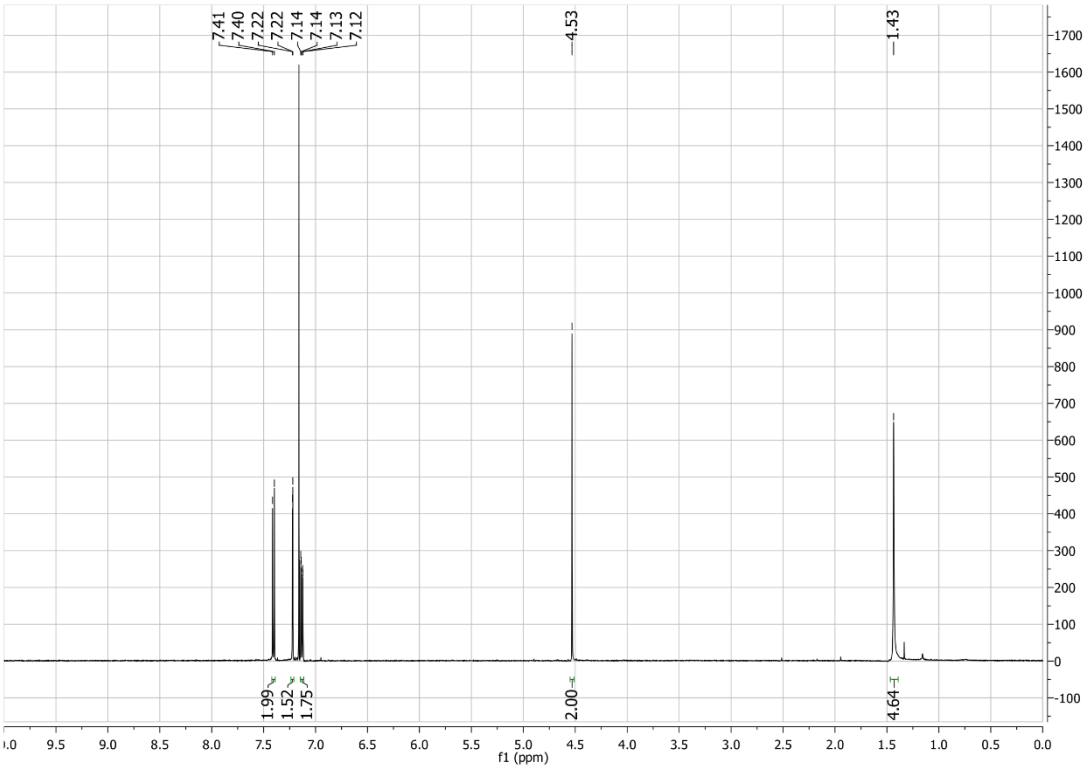


^{13}C NMR (126 MHz, $CDCl_3$)

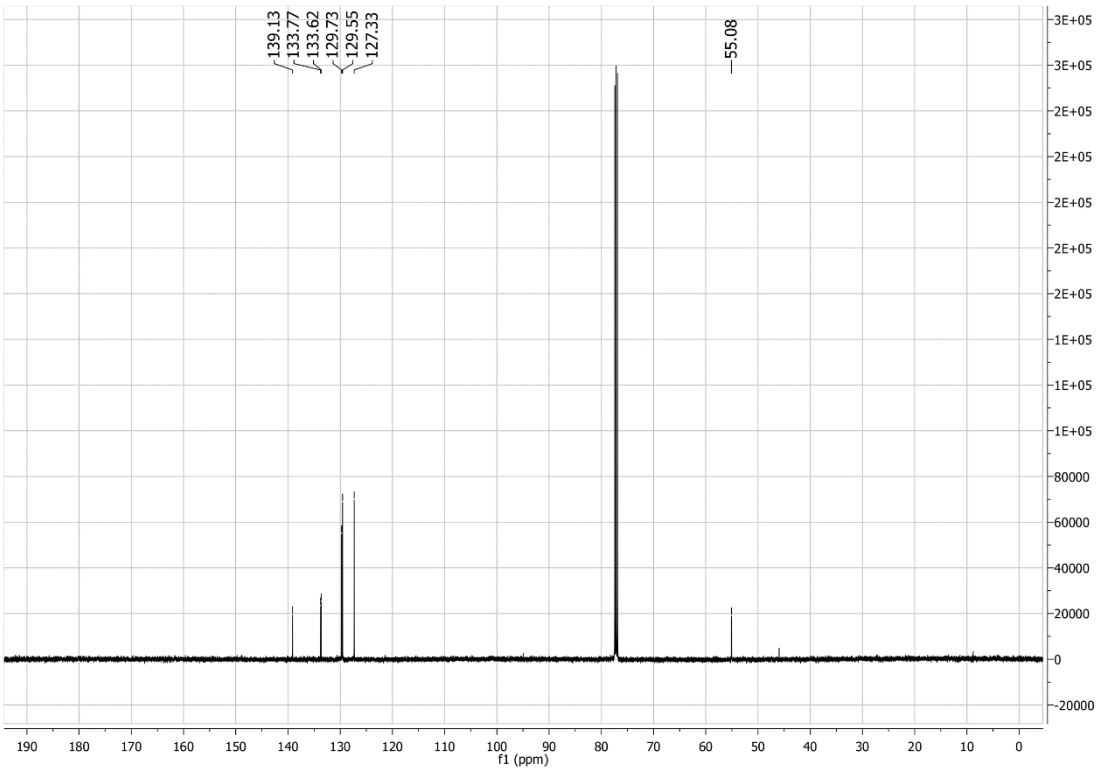


2,4-Cl_2-C_6H_3 diamine **SI5ae**

^{1}H NMR (500 MHz, $CDCl_3$)

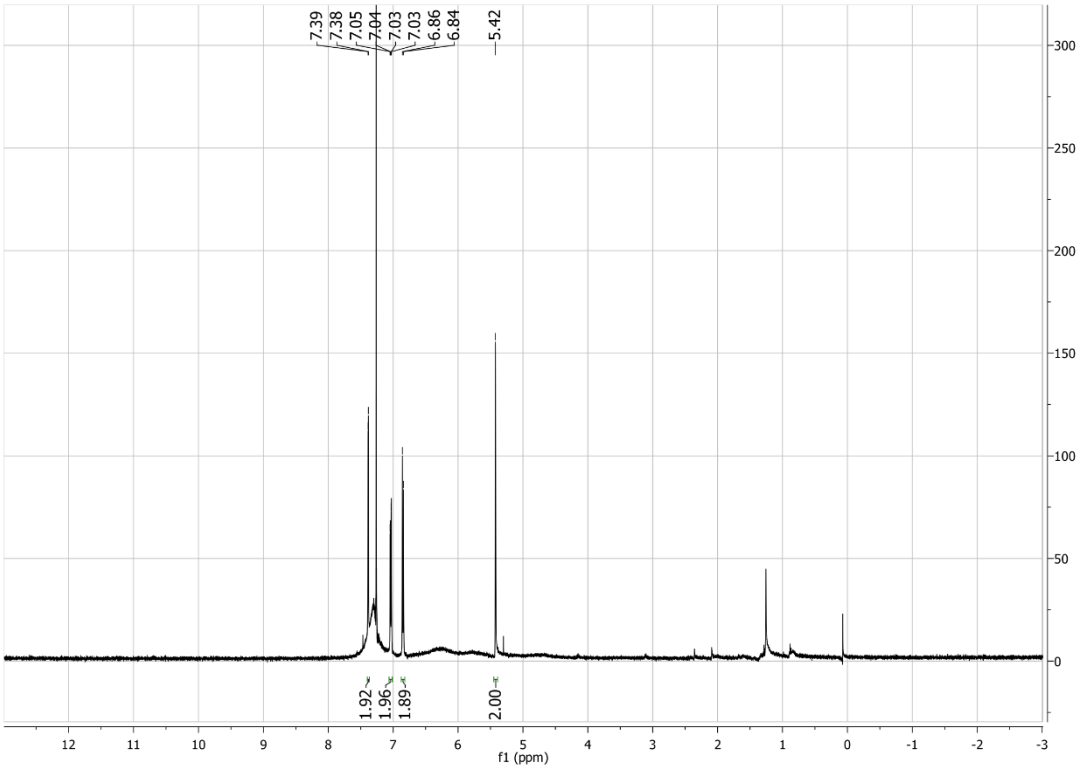


^{13}C NMR (126 MHz, $CDCl_3$)

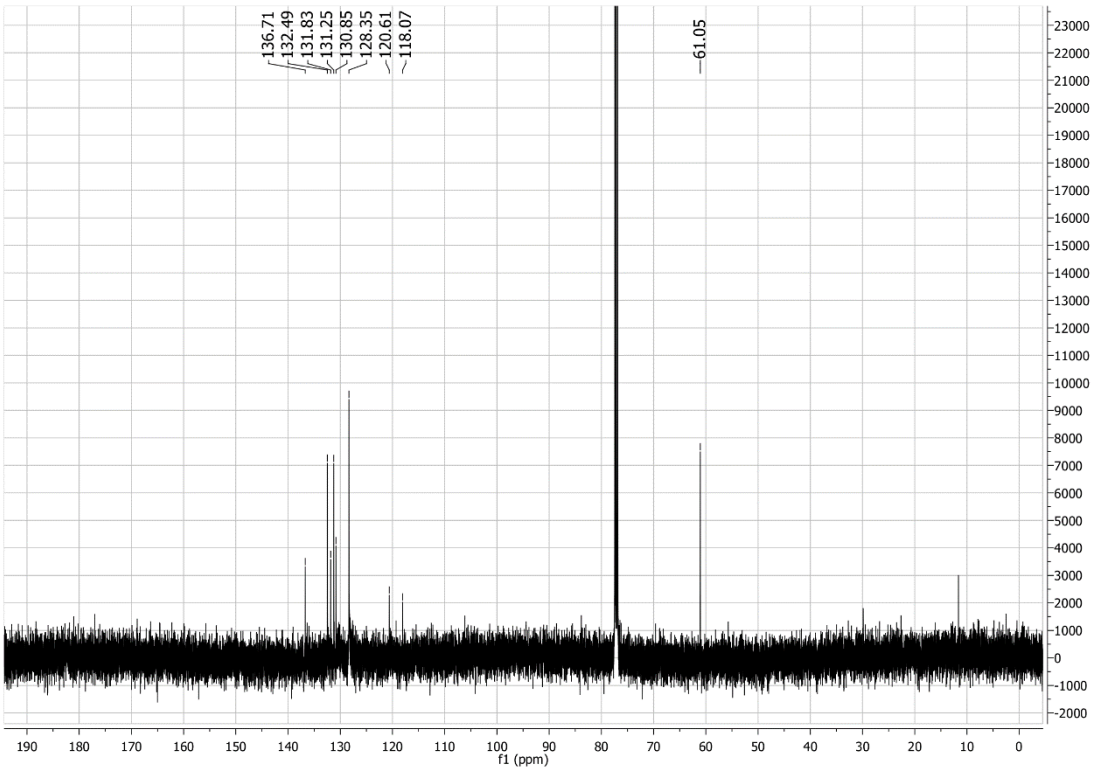


2,4-Cl_2-C_6H_3 bistriflimide **2ae**

^{1}H NMR (500 MHz, $CDCl_3$)

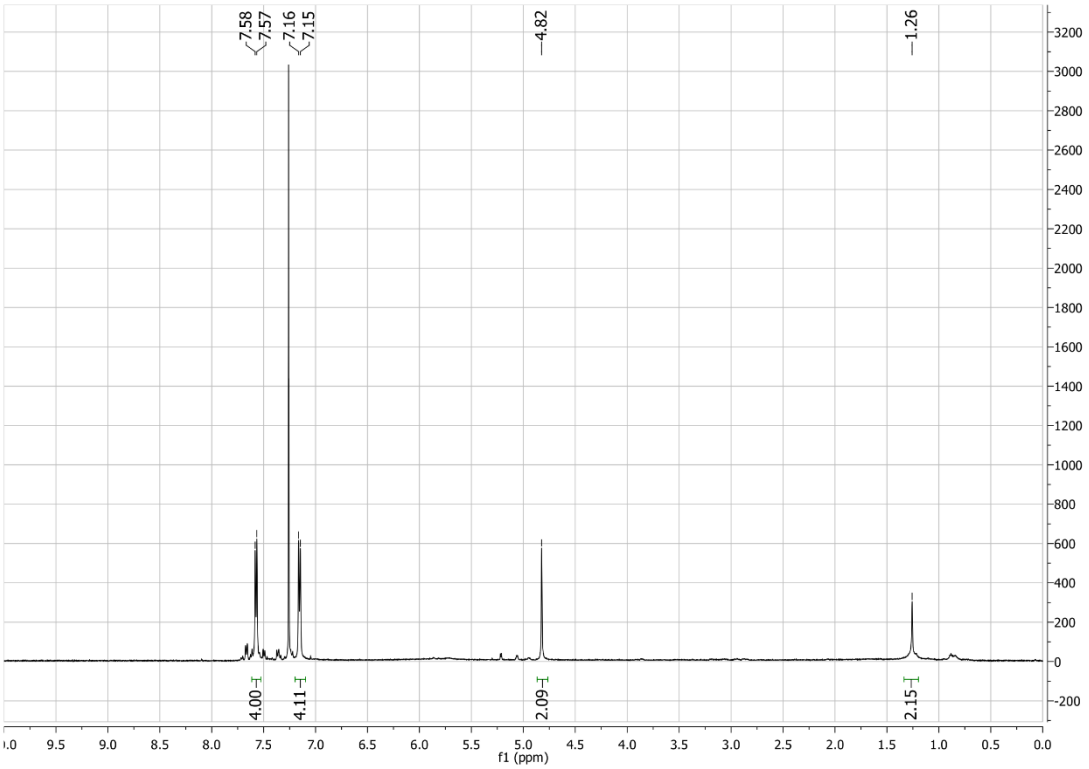


^{13}C NMR (126 MHz, $CDCl_3$)

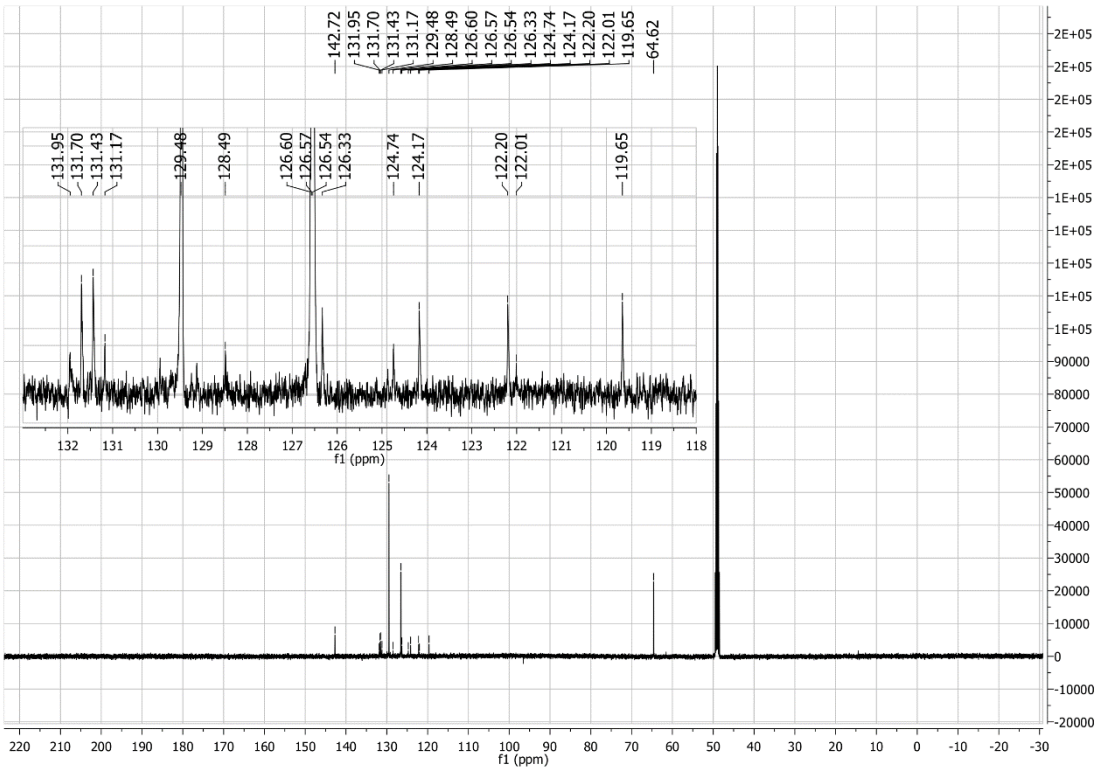


4-CF_3-C_6H_4 bistriflimide **2af**

^{1}H NMR (500 MHz, $CDCl_3$)

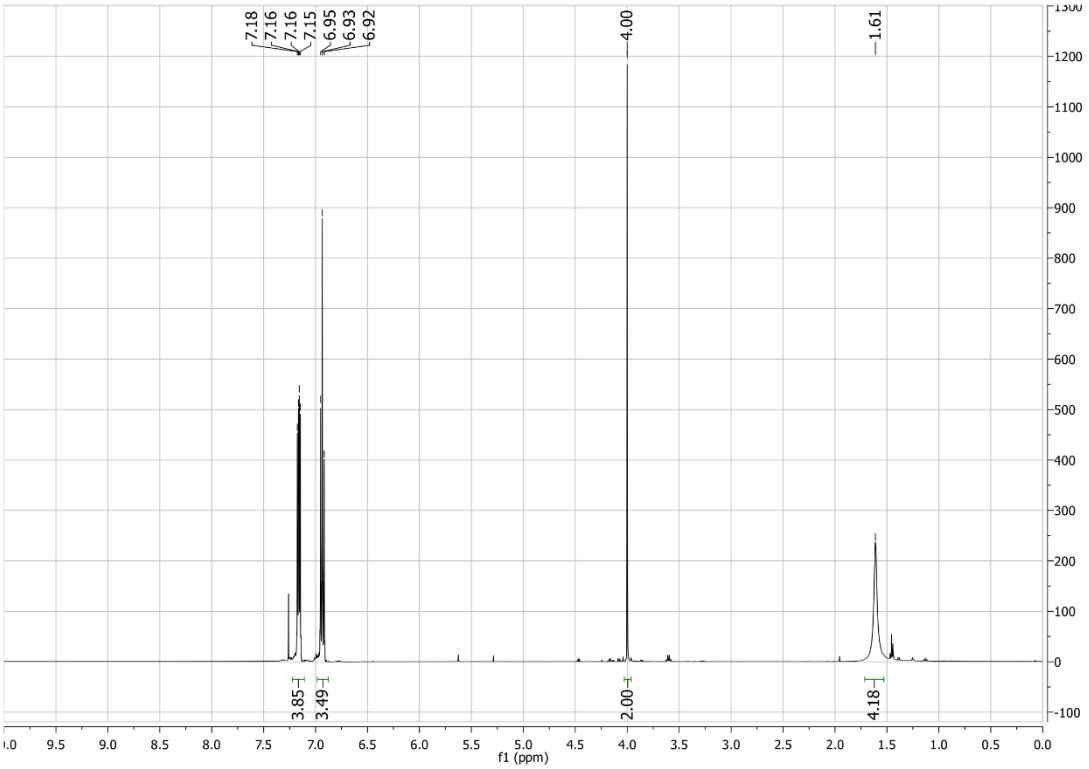


^{13}C NMR (126 MHz, $CDCl_3$)

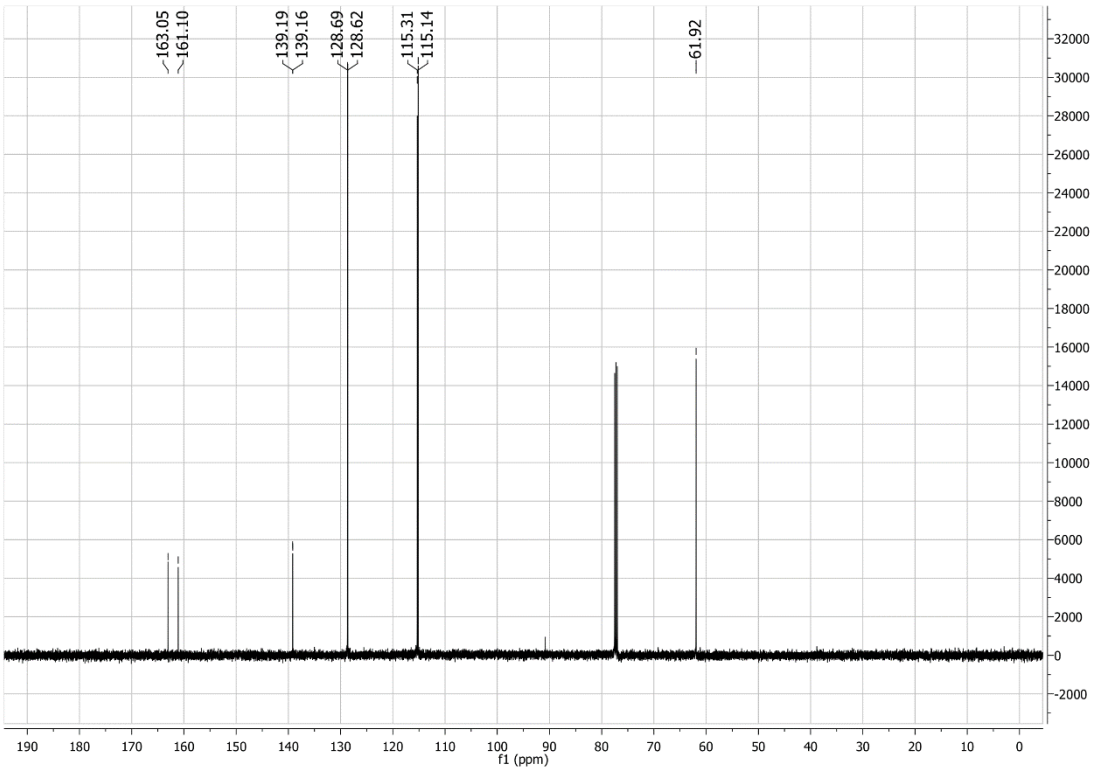


4-F-C_6H_4 diamine **SI5ag**

^{1}H NMR (500 MHz, $CDCl_3$)

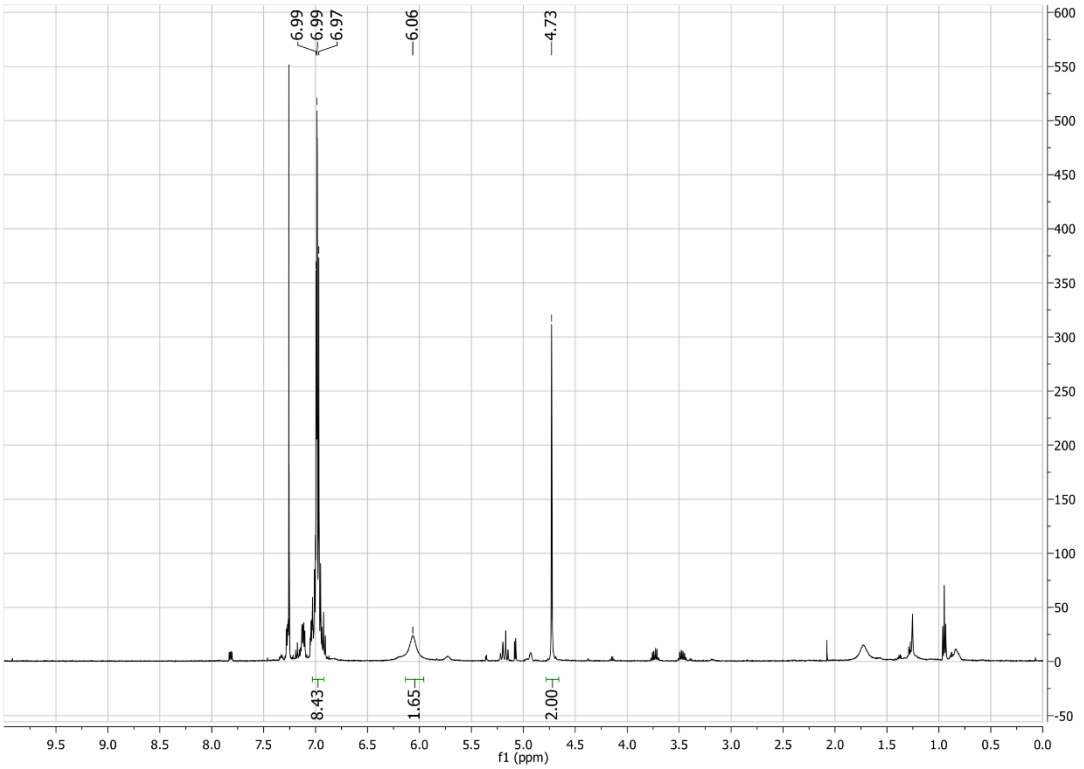


^{13}C NMR (126 MHz, $CDCl_3$)

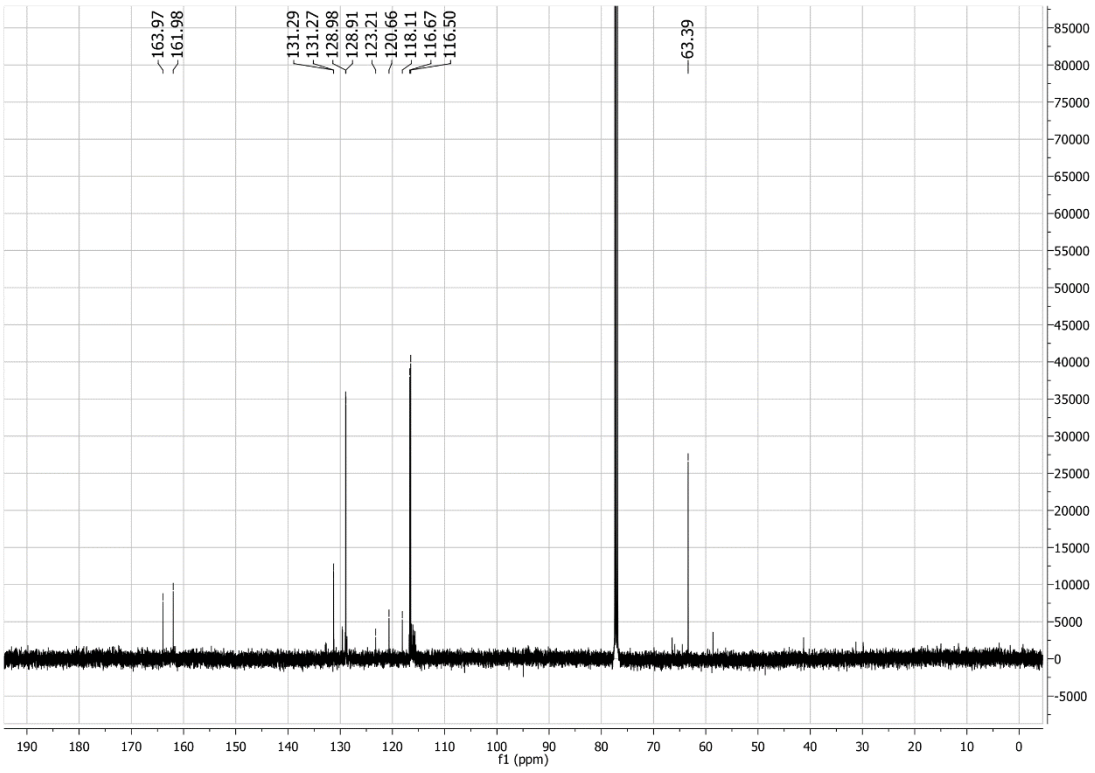


4-F-C_6H_4 bistriflimide **2ag**

^{1}H NMR (500 MHz, $CDCl_3$)

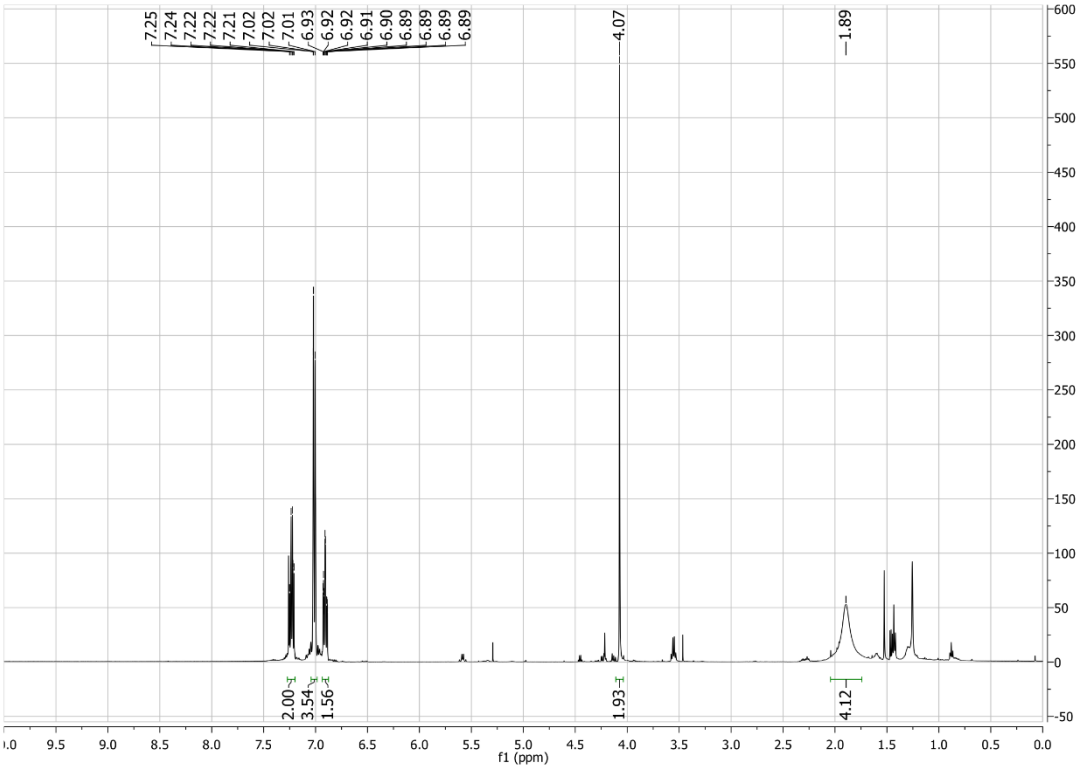


^{13}C NMR (126 MHz, $CDCl_3$)

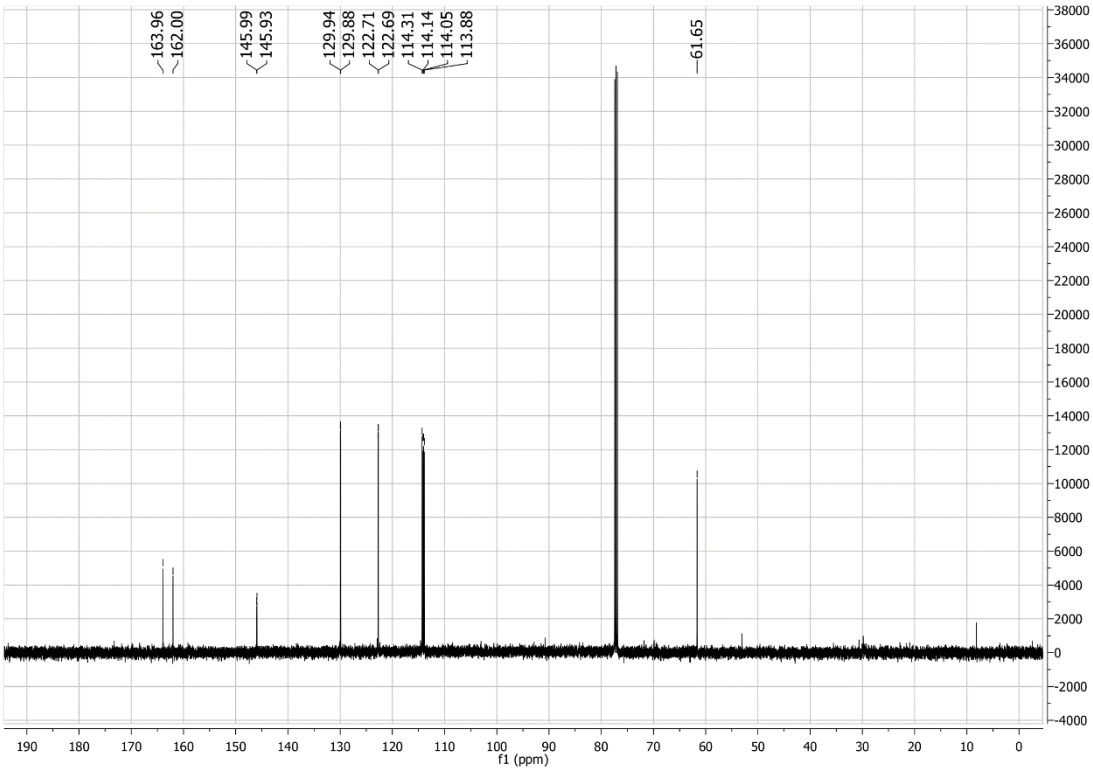


3-F-C_6H_4 diamine **SI5ah**

^{1}H NMR (500 MHz, $CDCl_3$)

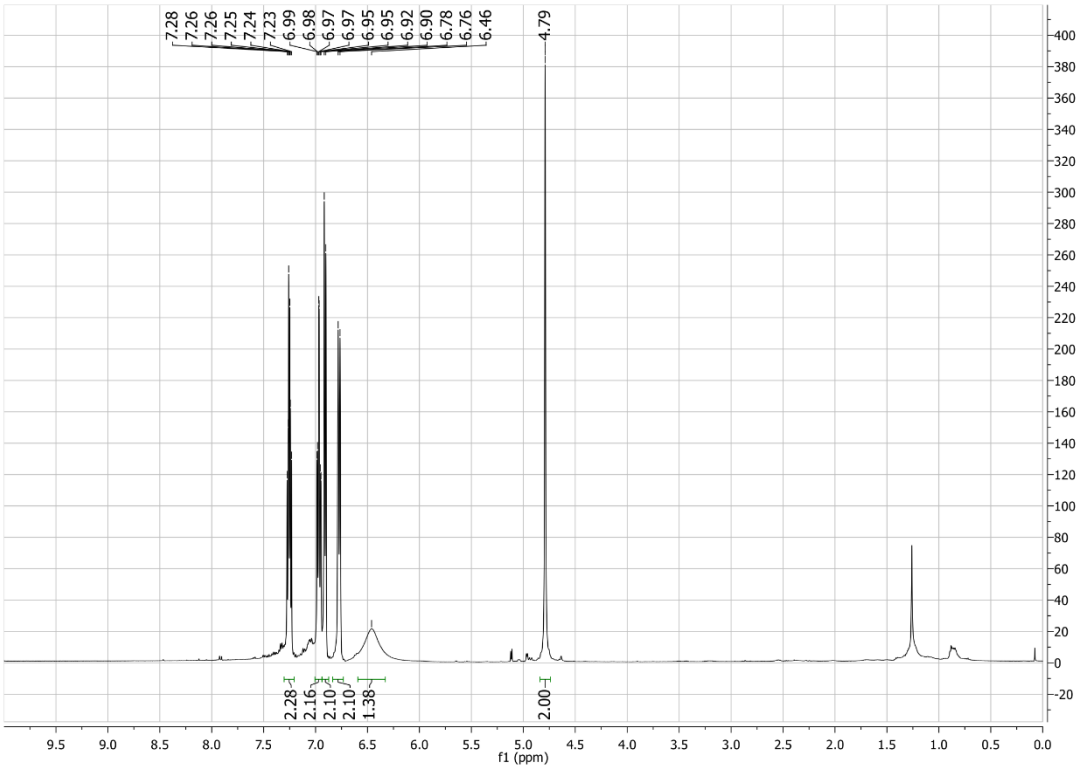


^{13}C NMR (126 MHz, $CDCl_3$)

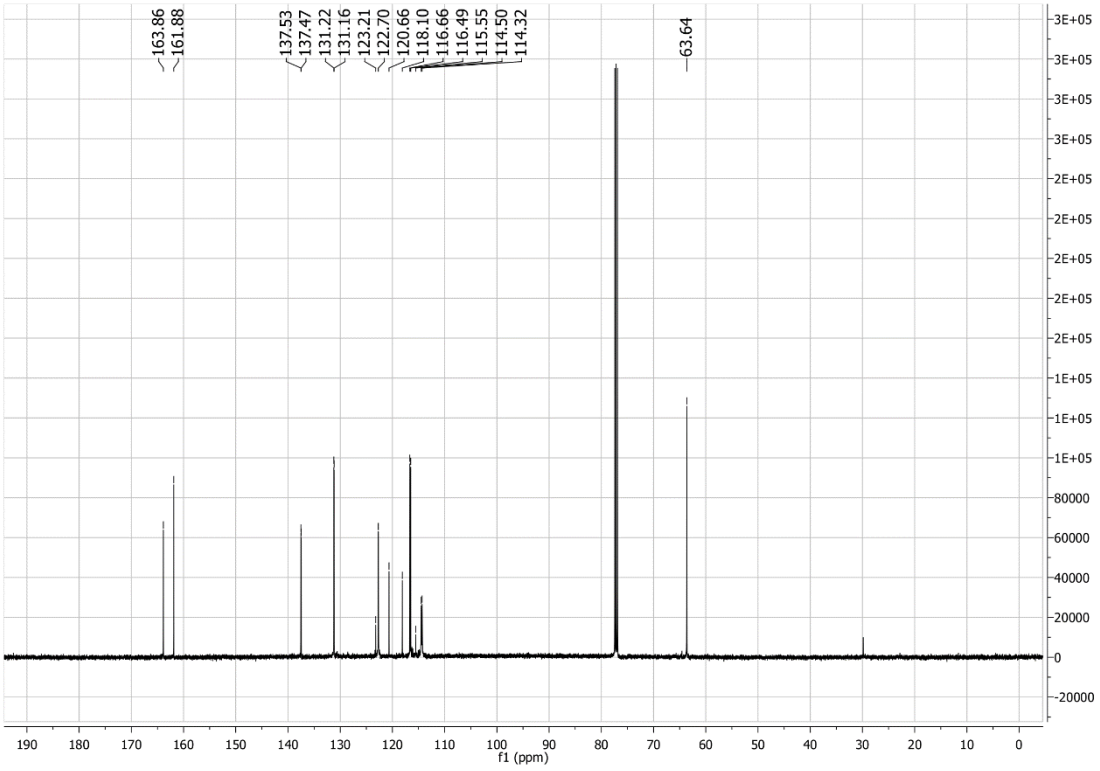


3-F-C_6H_4 bistriflimide **2ah**

^{1}H NMR (500 MHz, $CDCl_3$)

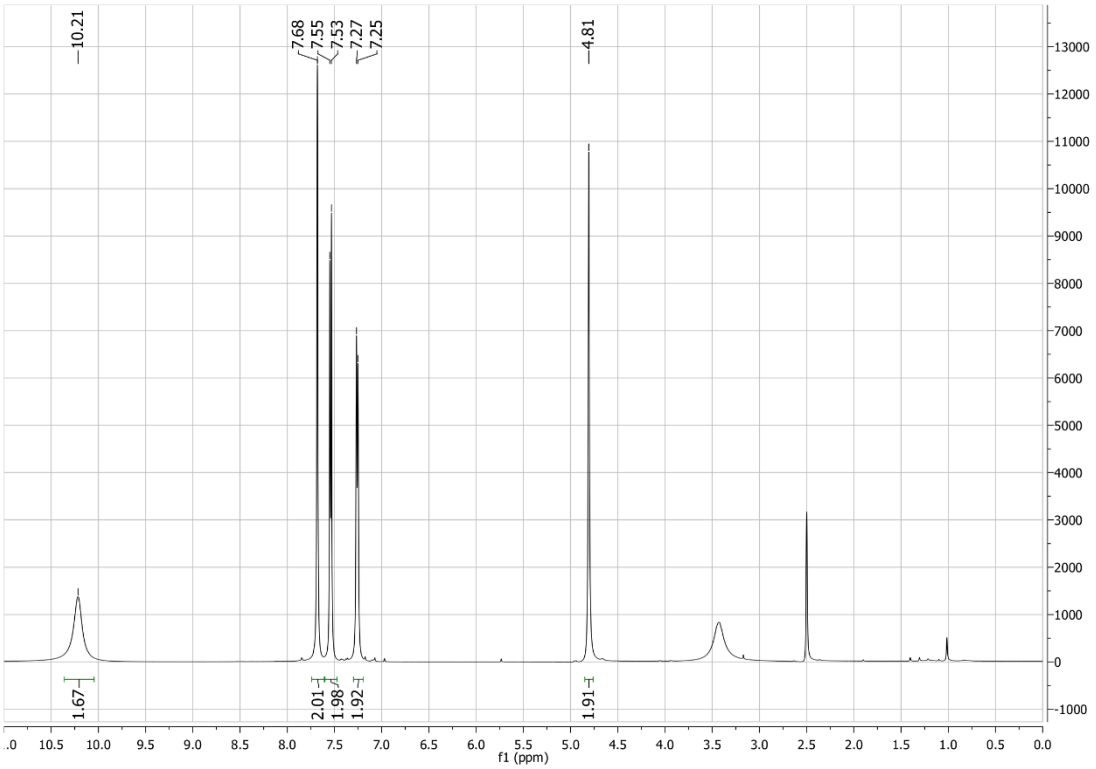


^{13}C NMR (126 MHz, $CDCl_3$)

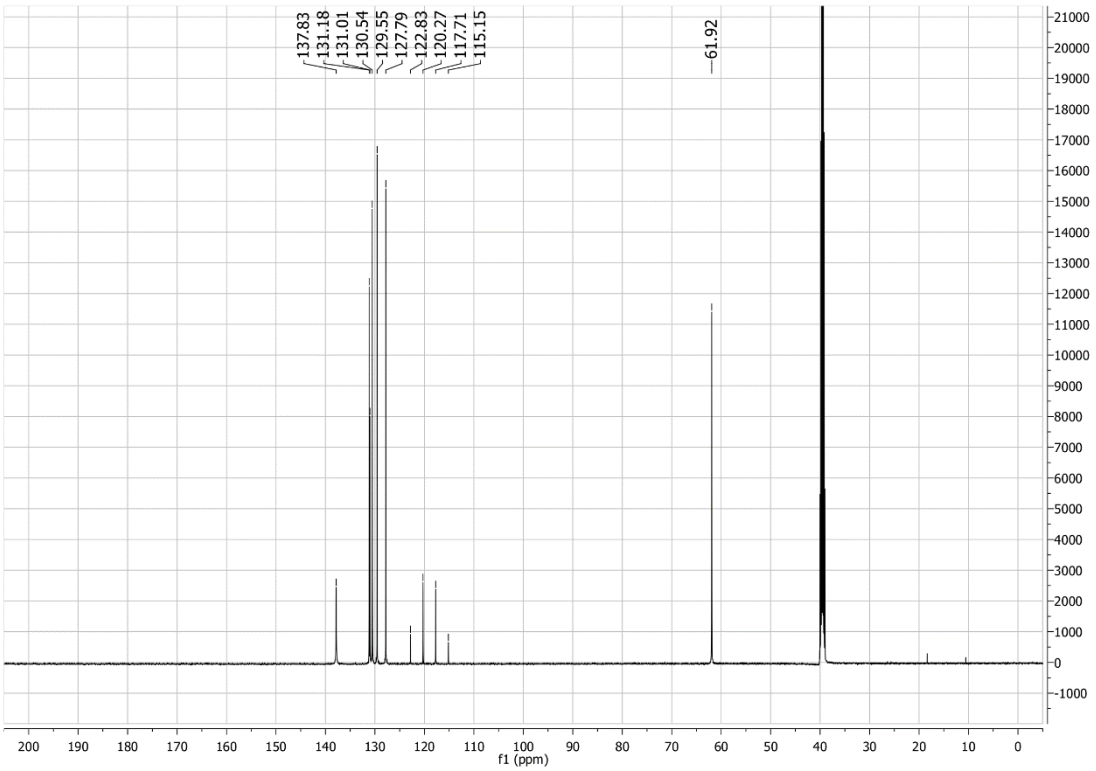


3,4-Cl_2-C_6H_3 bistriflimide **2ai**

^{1}H NMR (500 MHz, DMSO-d_6)

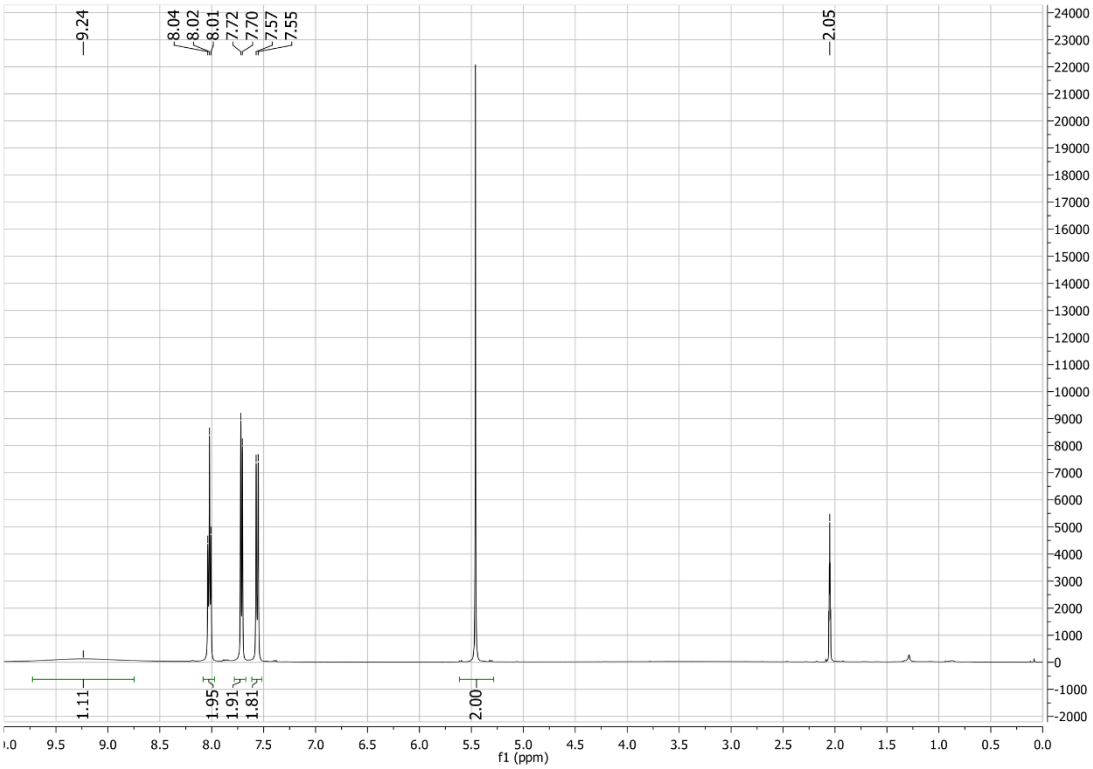


^{13}C NMR (126 MHz, DMSO-d_6)

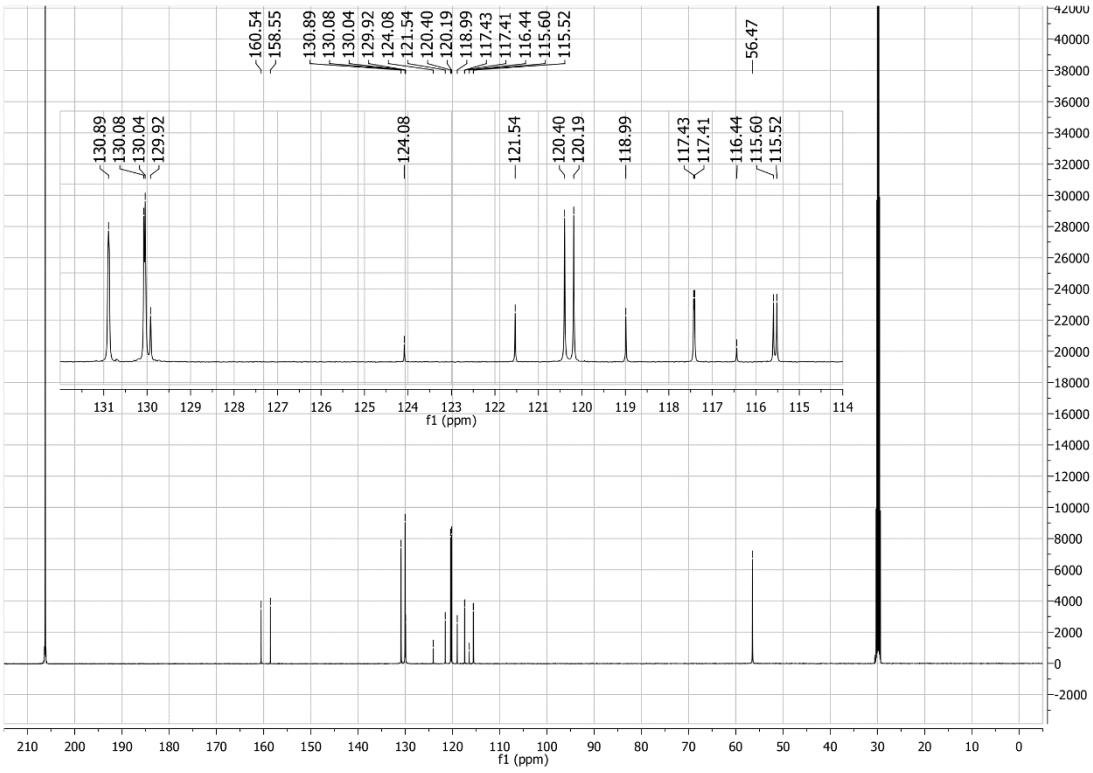


2-F-4-CN-C_6H_3 bistriflimide **2aj**

^{1}H NMR (500 MHz, acetone-d_6)

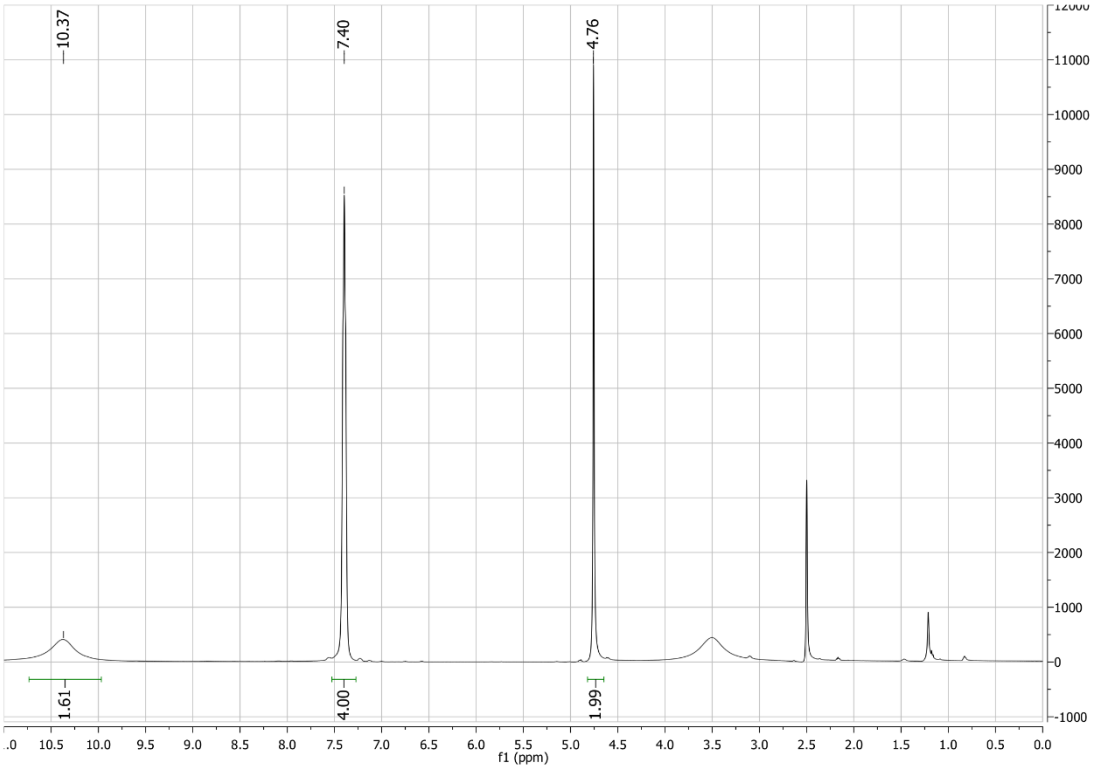


^{13}C NMR (126 MHz, acetone-d_6)

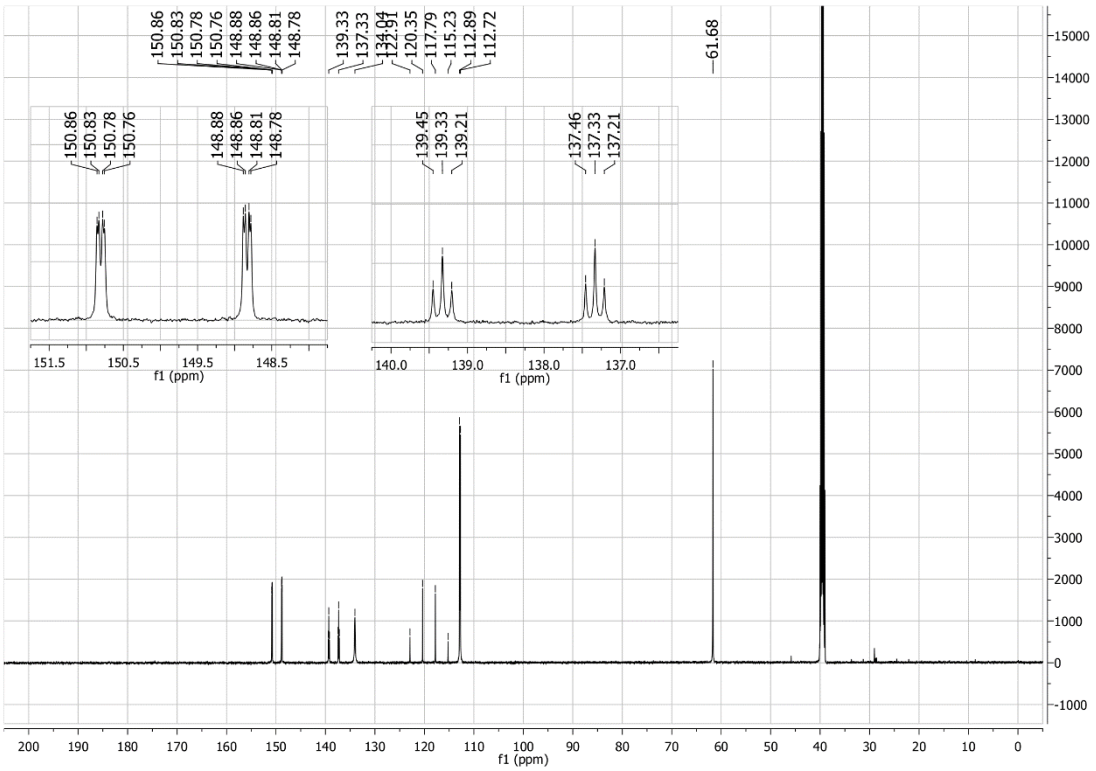


3,4,5-F_3-C_6H_2 bistriflimide **2ak**

^{1}H NMR (500 MHz, DMSO-d_6)

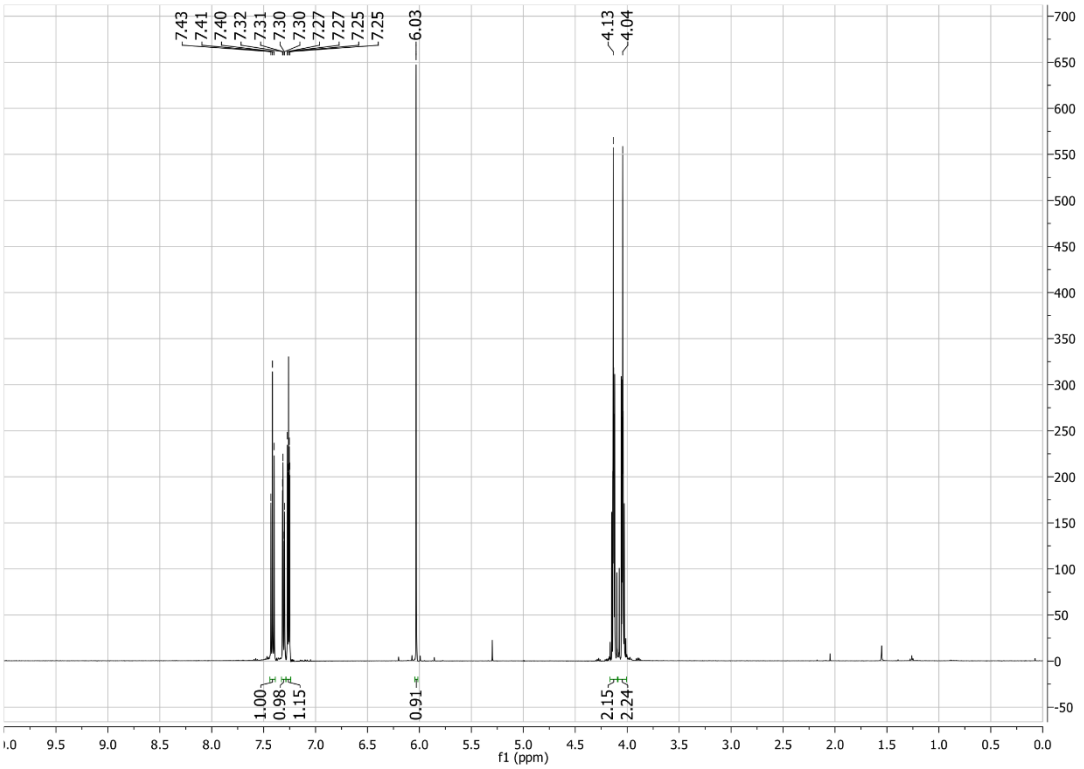


^{13}C NMR (126 MHz, DMSO-d_6)

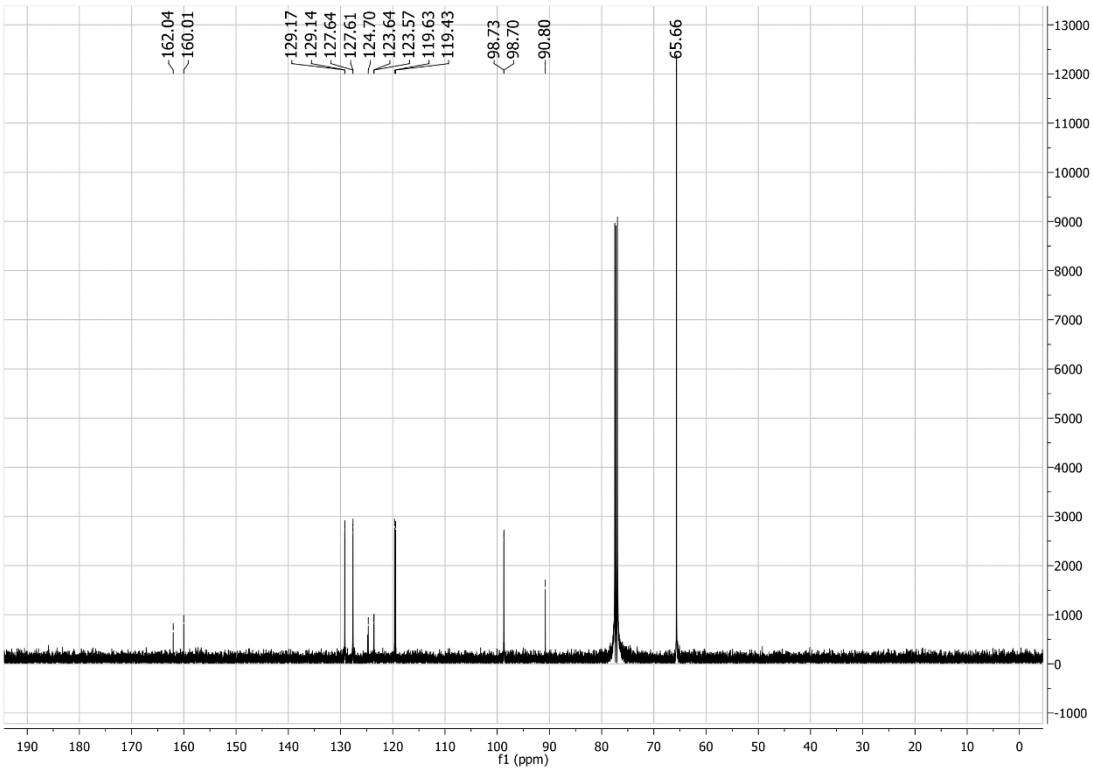


2-(4-bromo-2-fluorophenyl)-1,3-dioxolane (**SI6**)

^{1}H NMR (500 MHz, $CDCl_3$)

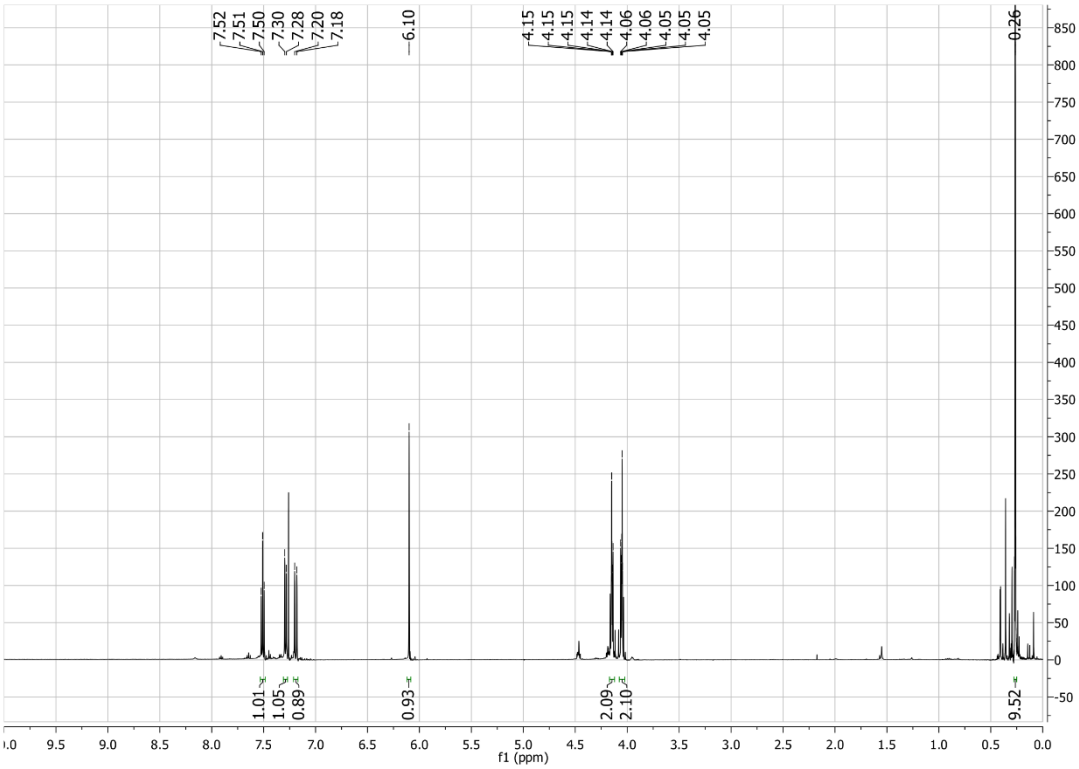


^{13}C NMR (126 MHz, $CDCl_3$)

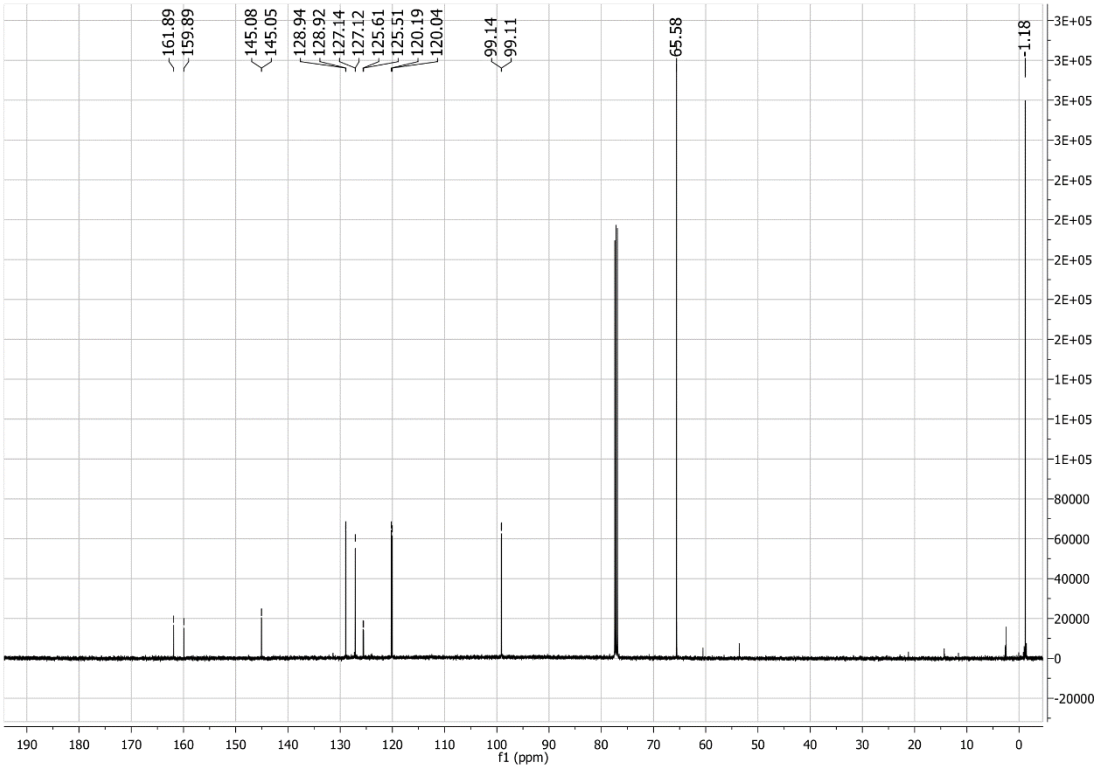


(4-(1,3-dioxolan-2-yl)-3-fluorophenyl)trimethylsilane (**SI7**)

^{1}H NMR (500 MHz, $CDCl_3$)

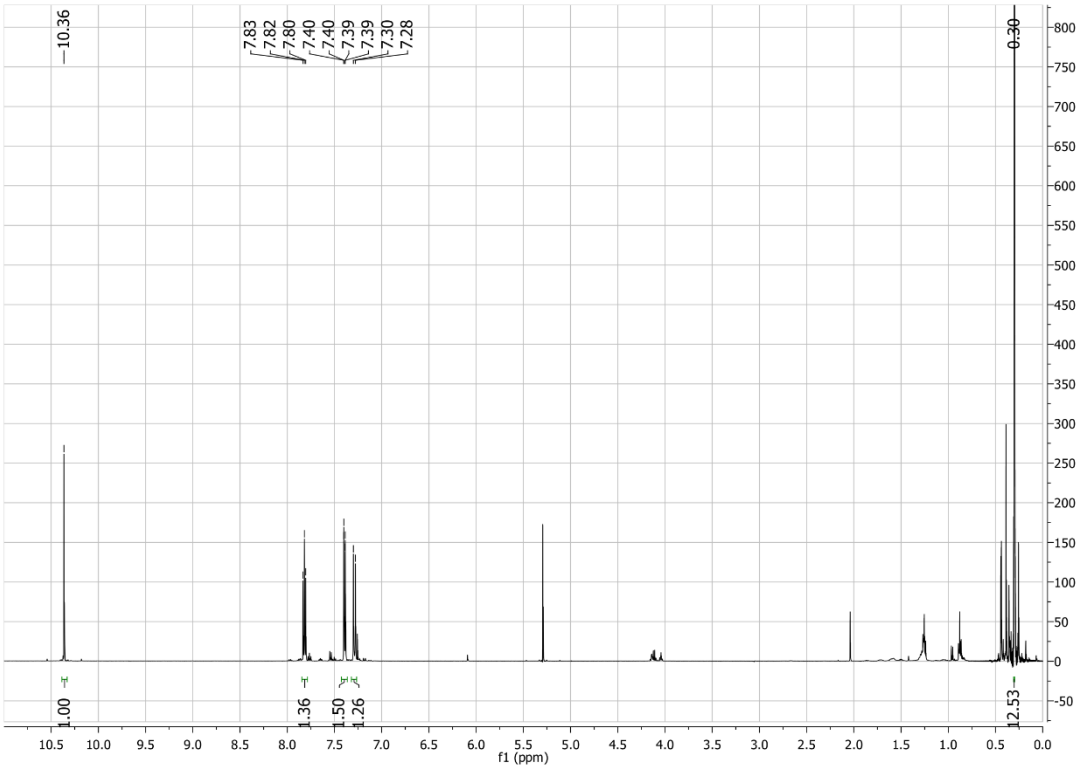


^{13}C NMR (126 MHz, $CDCl_3$)

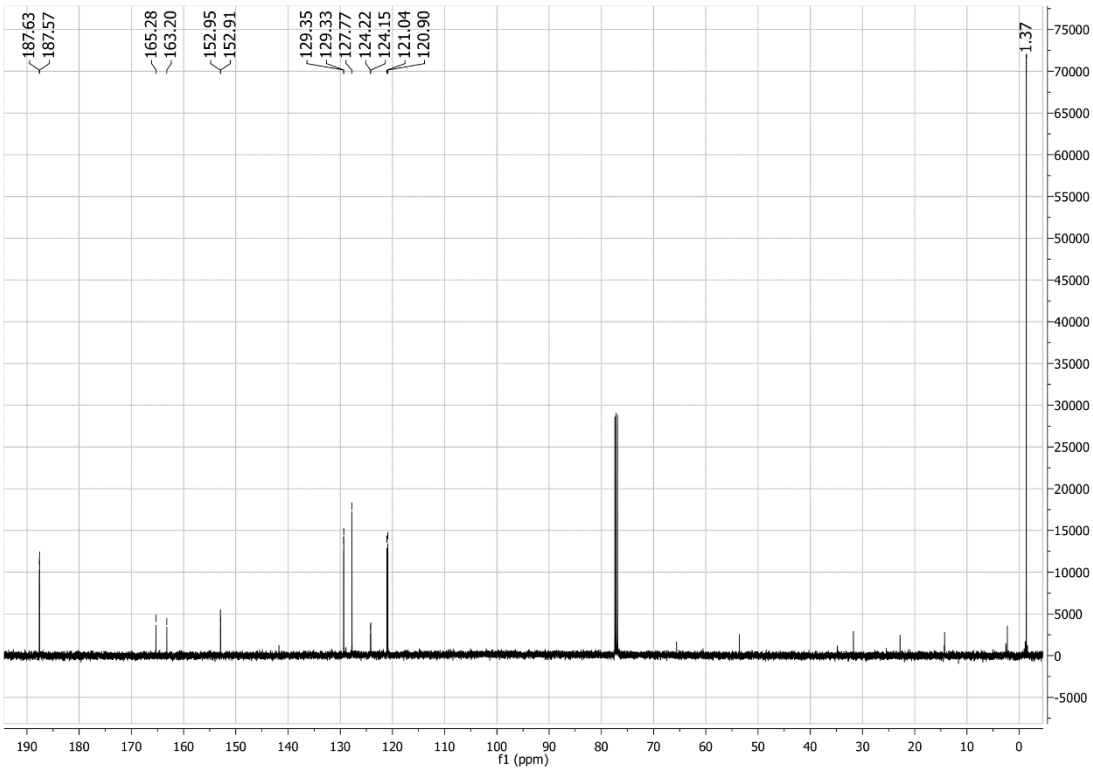


2-fluoro-4-(trimethylsilyl)benzaldehyde (**SI2al**)

^{1}H NMR (500 MHz, $CDCl_3$)

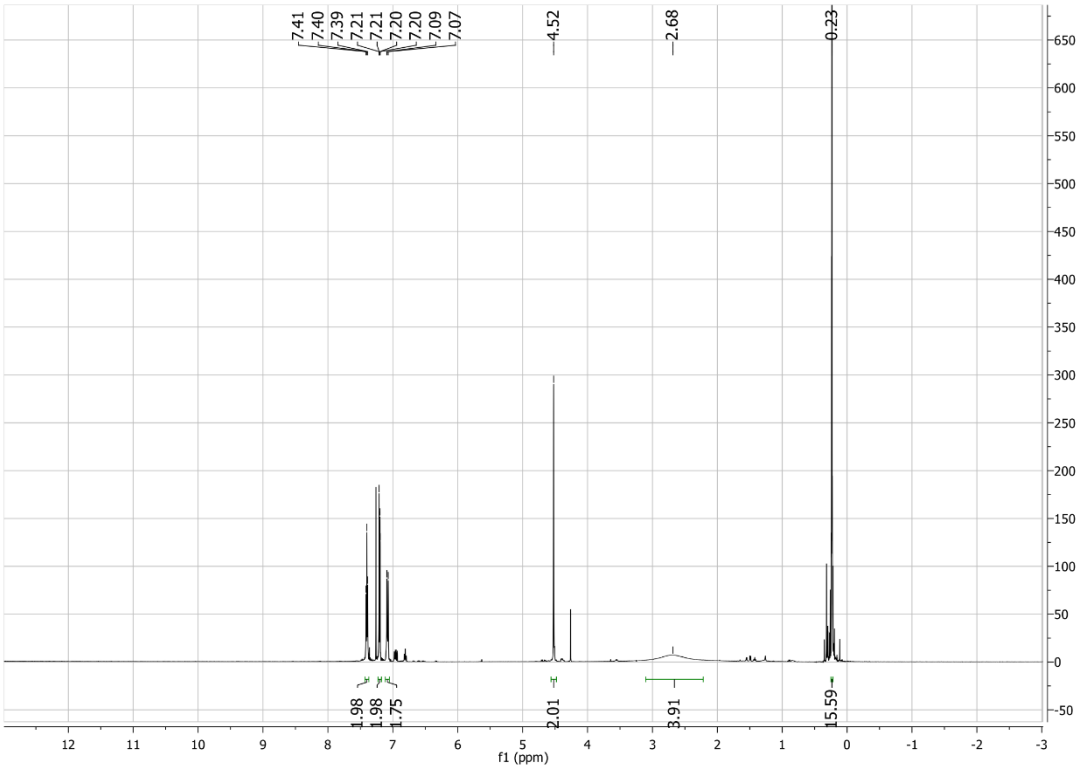


^{13}C NMR (126 MHz, $CDCl_3$)

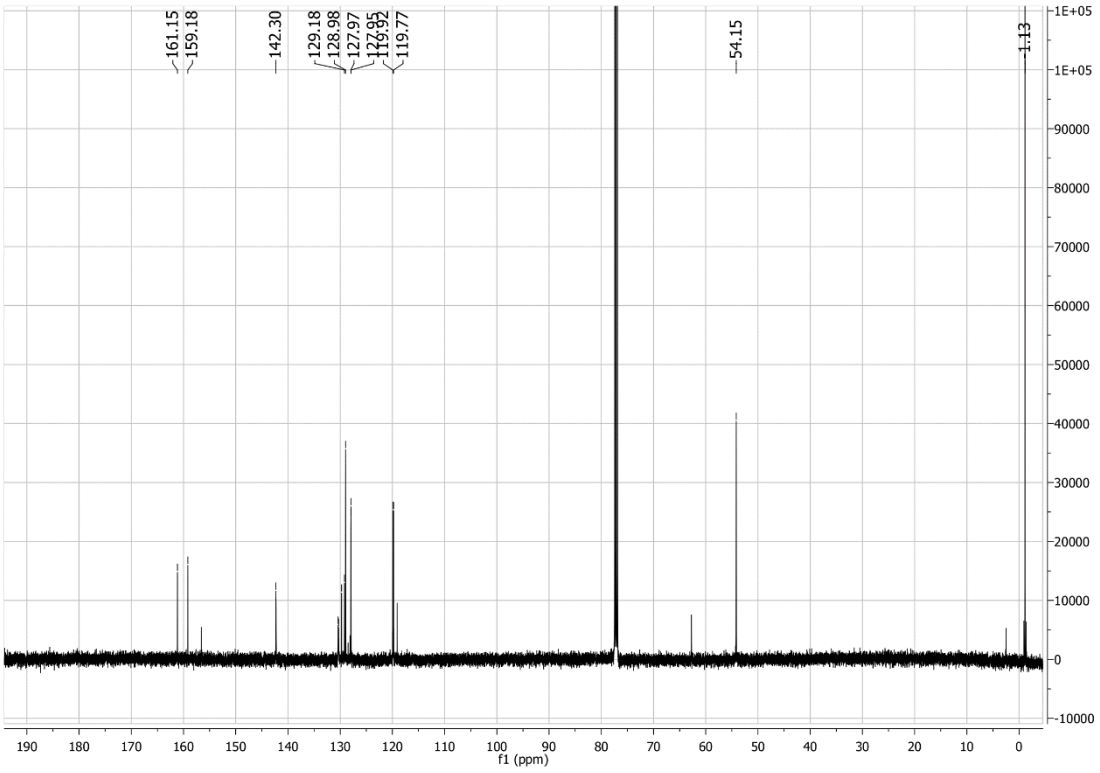


2-F-4-SiMe$_3$-C$_6$H$_3$ diamine **SI5al**

^{1}H NMR (500 MHz, $CDCl_3$)

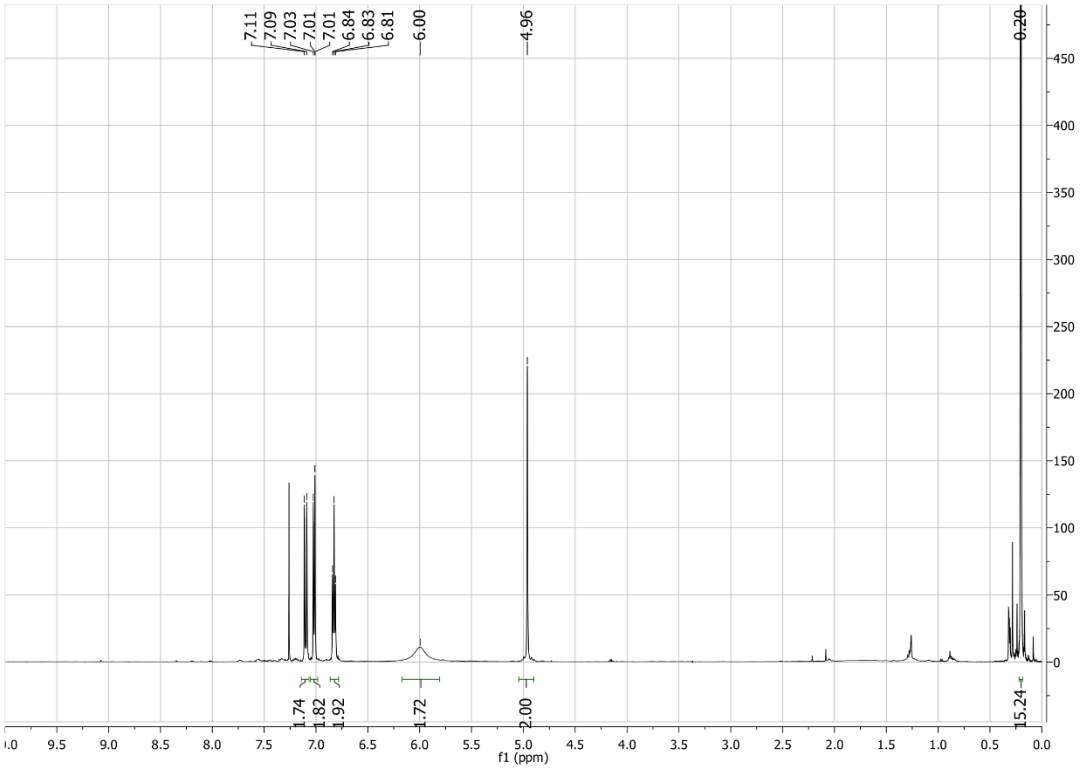


^{13}C NMR (126 MHz, $CDCl_3$)

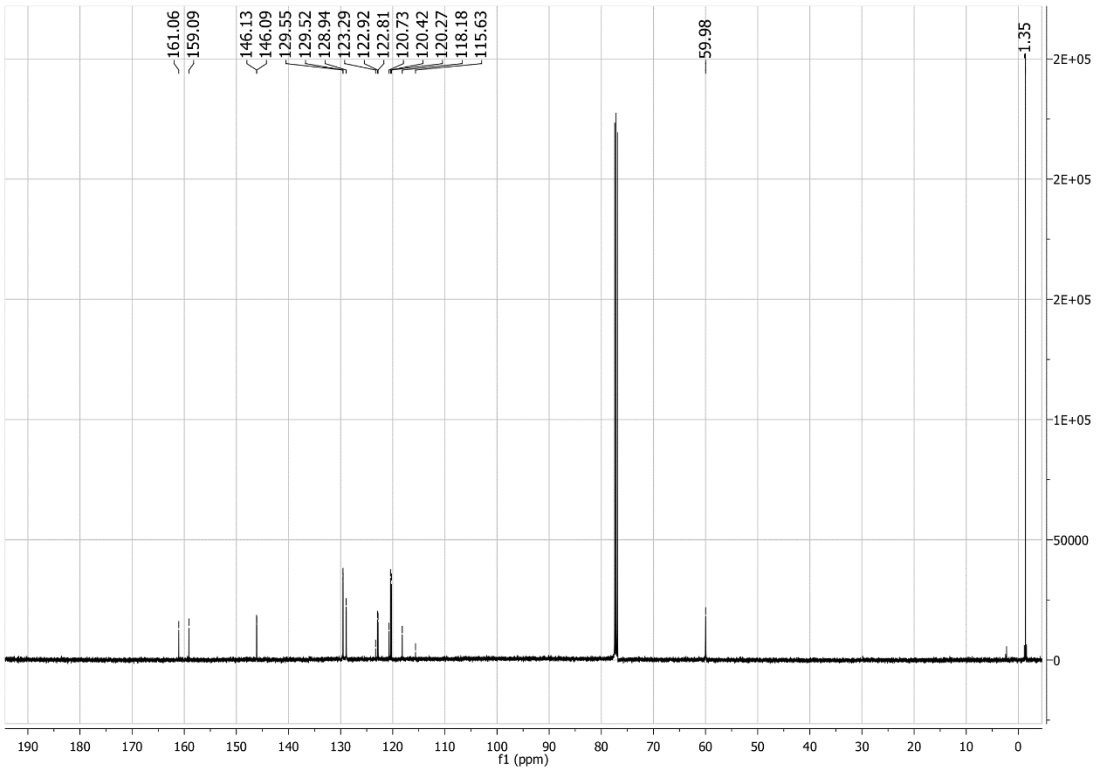


2-F-4-SiMe$_3$-C$_6$H$_3$ bistriflimide **2al**

^{1}H NMR (500 MHz, $CDCl_3$)

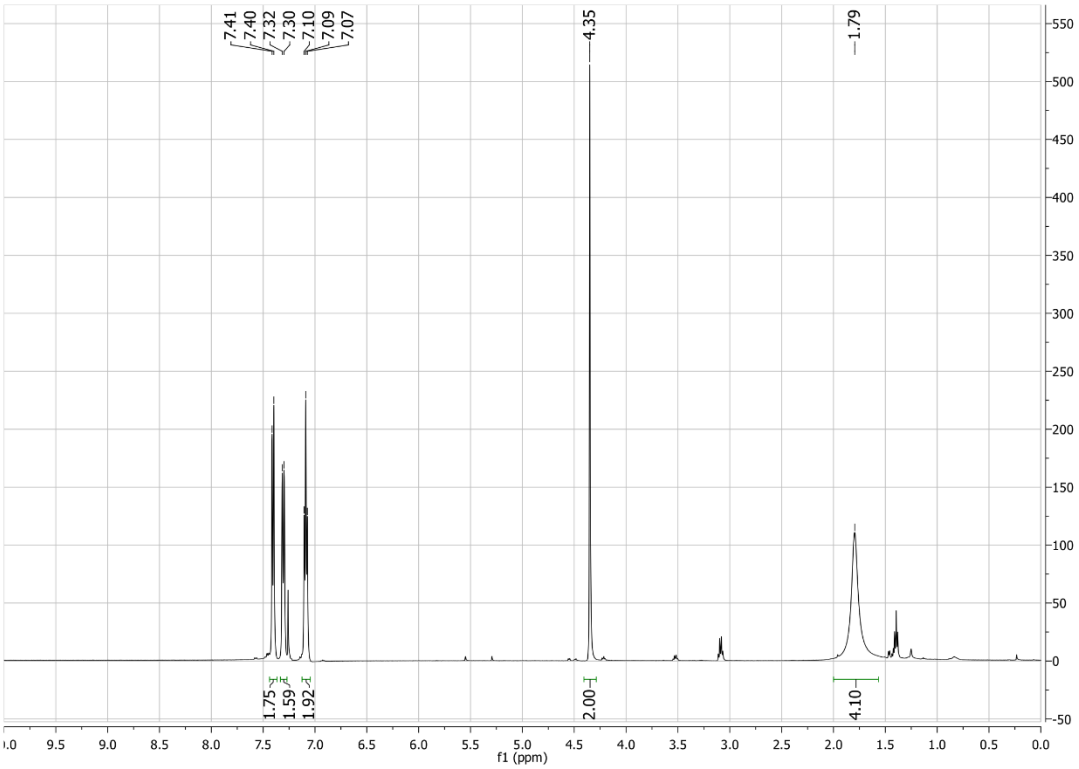


^{13}C NMR (126 MHz, $CDCl_3$)

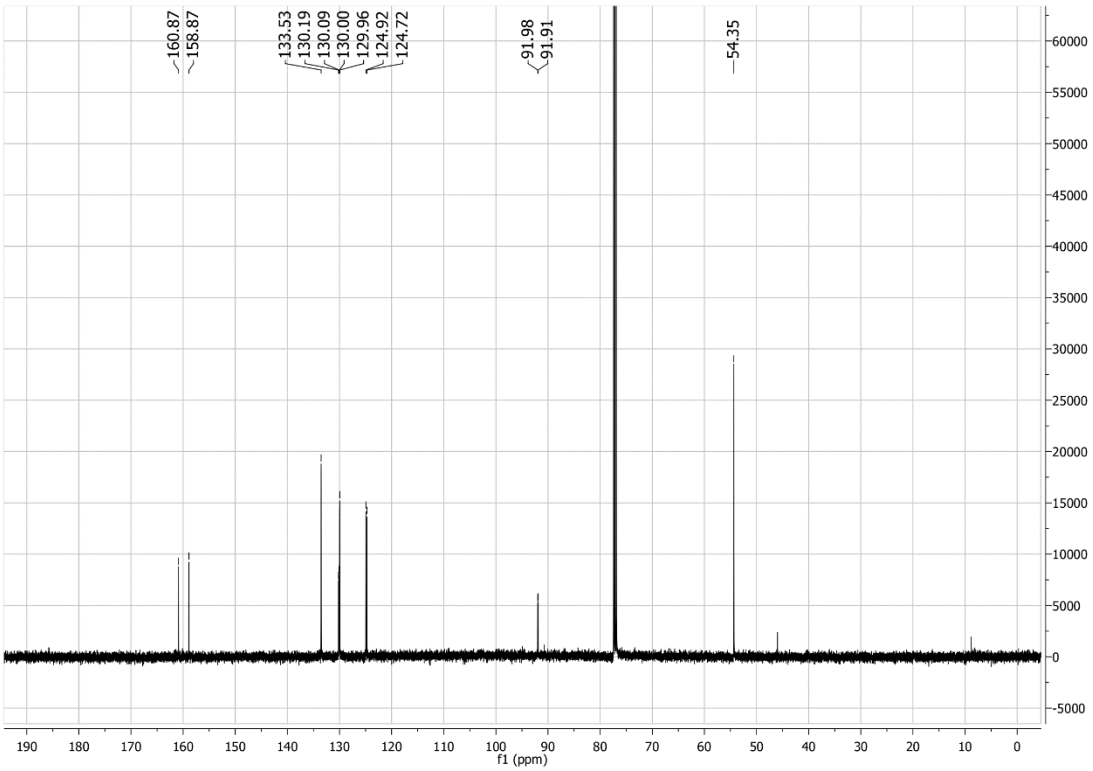


2-F-4-I-C_6H_3 diamine **SI5am**

^{1}H NMR (500 MHz, $CDCl_3$)

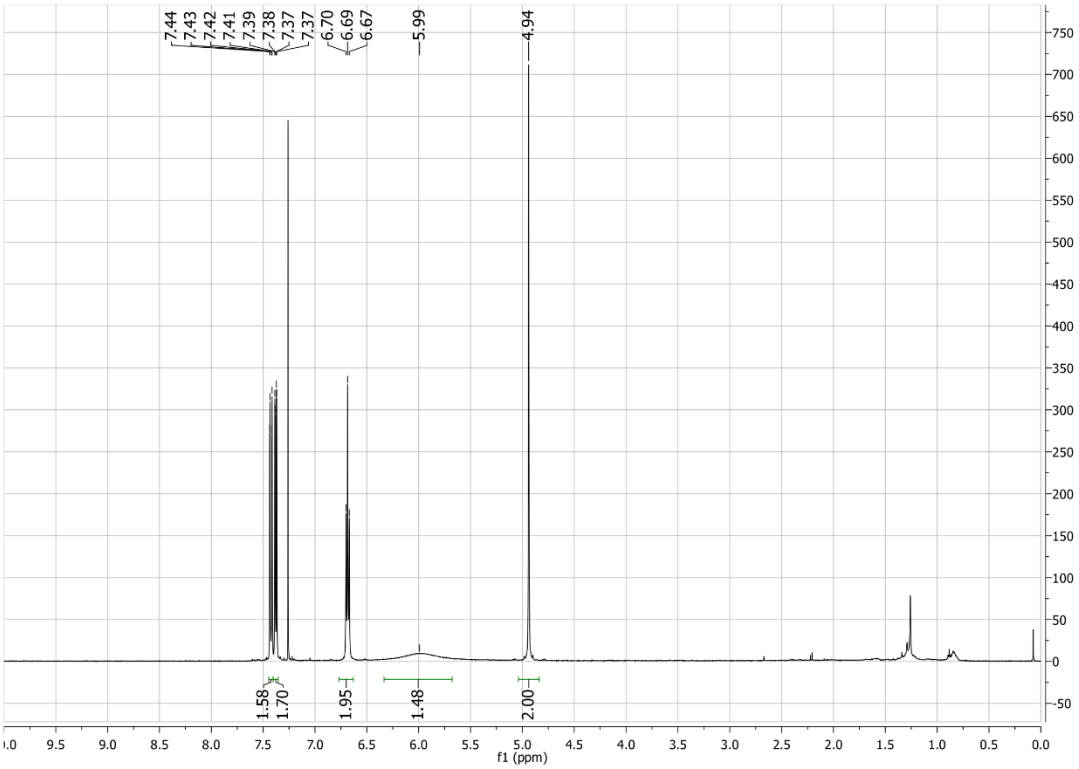


^{13}C NMR (126 MHz, $CDCl_3$)

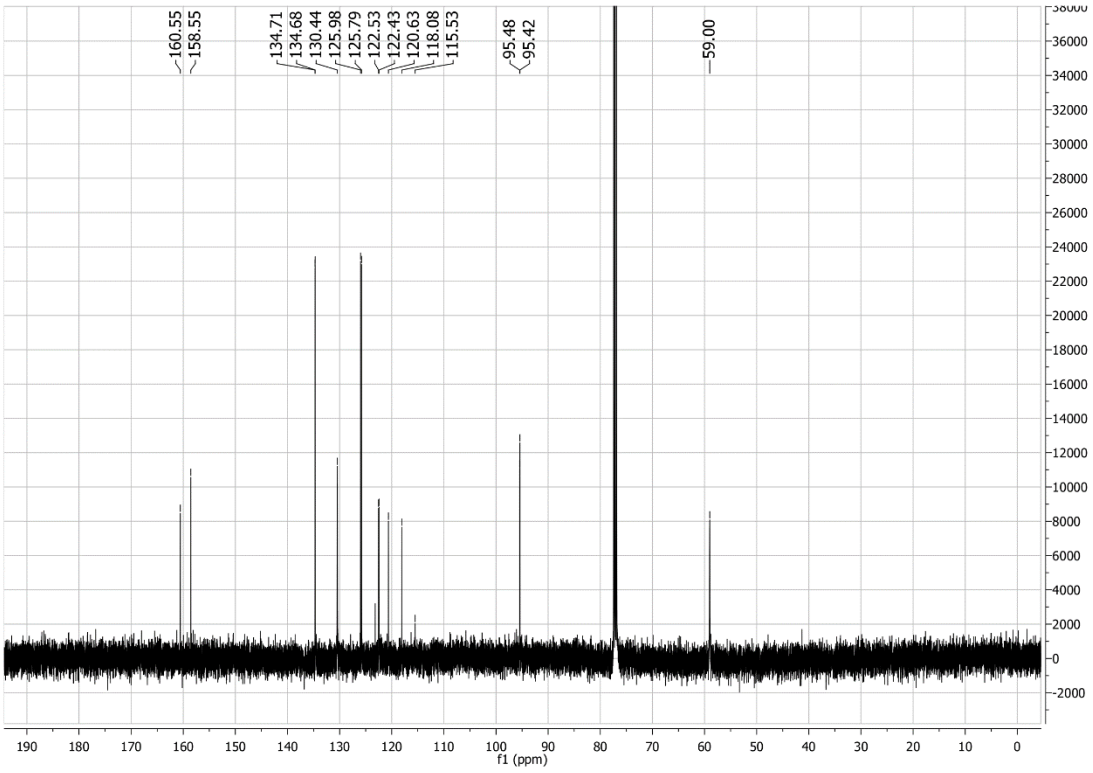


2-F-4-I-C_6H_3 bistriflimide **2am**

^{1}H NMR (500 MHz, $CDCl_3$)

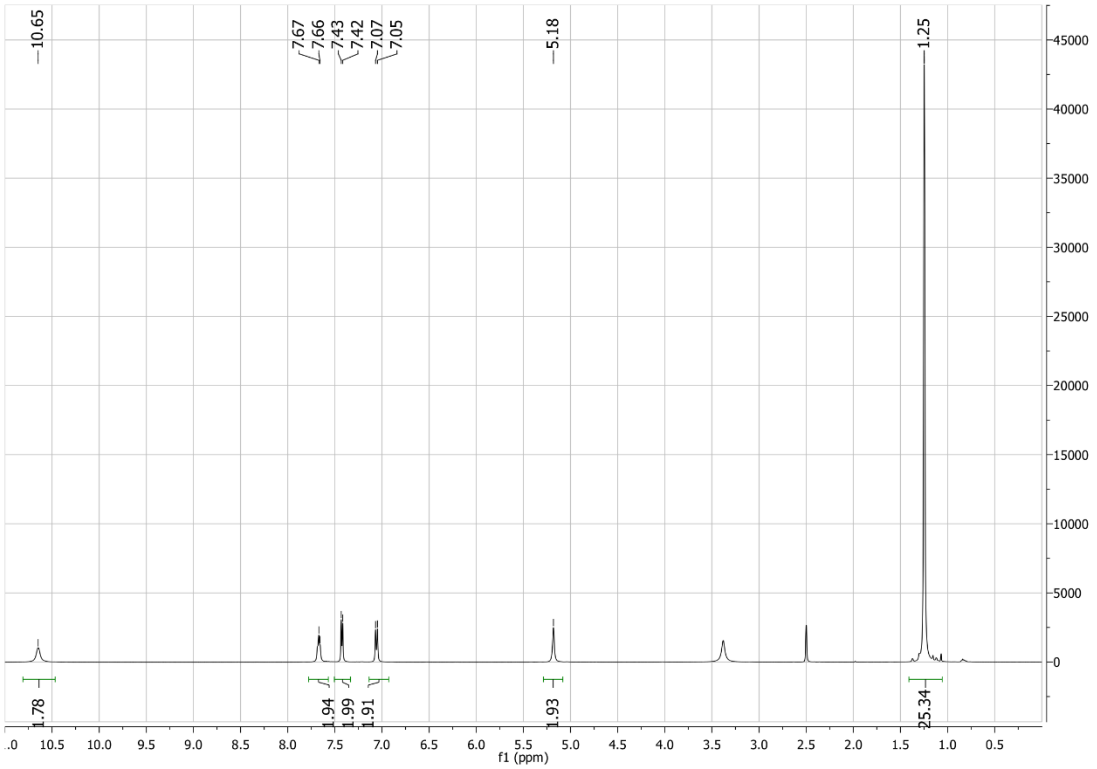


^{13}C NMR (126 MHz, $CDCl_3$)

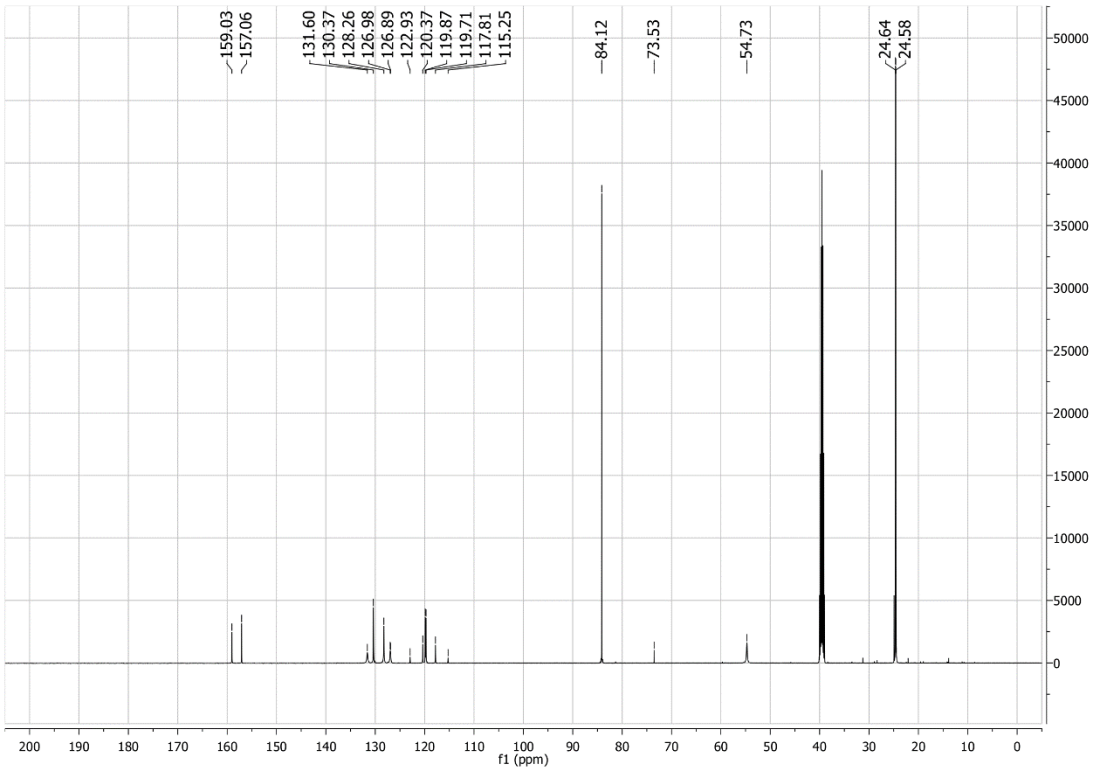


2-F-4-Bpin-C_6H_3 bistriflimide **2an**

^{1}H NMR (500 MHz, DMSO-d_6)



^{13}C NMR (126 MHz, DMSO-d_6)



5. Modeling

Since enantioselectivity cannot be reliably quantified for low-activity catalysts, compounds with a measured conversion below 25% were not included in initial model testing, model building, or analysis. Experimentally measured enantioselectivity was converted to the natural log, allowing a direct comparison to free energy ($\Delta\Delta G^{\ddagger}$) since:

$$\ln \frac{[\boldsymbol{A}_R]}{[\boldsymbol{A}_S]} = -\frac{\Delta\Delta G^{\ddagger}}{RT}$$

where [$\mathbf{A}_R$] is the concentration of (*R*)-enantiomer of compound **A**, [$\mathbf{A}_S$] is the concentration of (*S*)-enantiomer of coumpound **A**, $\Delta\Delta G^{\ddagger}$ is the Gibbs free energy difference between the two transition states leading to compounds $\mathbf{A}_S$ or $\mathbf{A}_R$, R is the ideal gas constant, and T is the temperature.

2.2 2D-QSAR

Initial efforts to build a QSAR model for these bistriflimide catalysts relied upon a set of 48 catalysts with a range of enantioselectivity and conversion, as measured during catalyst screening. As a result, these experimental values were determined under a variety of solvent conditions. By limiting the set of triflimides used to build the QSAR model to only those screened in our process chemistry solvent condition (toluene), we saw significant improvements in the prospective power of our QSAR models.

Our ultimate goal was to generate a reliable model with which we could assign an enantioselectivity score to each potential bistriflimide catalyst. To do this, we used a QSAR method to fit enantioselectivity against a relevant set of calculated descriptors. Based on our substantial experience developing reliable 2D QSAR models for medicinal chemistry programs, we tried the following descriptor sets:

AP(C): Carhart atom pairs,[8] which encode distances between detailed atom types, with (APC) or without (AP) chiral center information.

TT(C): Topological torsions,[9] which encode 4 contiguous atoms and their environment.

ECFP4: Extended-connectivity fingerprints with bond diameter 4,[10] which are circular fingerprints specifically developed to represent substructures, include chiral center information, and facilitate model interpretation.

SP3CARBONS: Count of all sp^3 hybridized carbons.

MOE_2D: A set of diverse physical and topological properties calculated by the CCG software package MOE.[11]

---

8 R. E. Carhart, D. H. Smith and R. Ventkataraghavan, *J. Chem. Inf. Comput. Sci.* 1985, **25**, 64.
9 R. Nilakantan, N. Bauman, J. S. Dixon and R. Venkataraghavan, *J. Chem. Inf. Comput. Sci.* 1987, **27**, 82.
10 D. Rogers and M. Hahn, *J. Chem. Inf. Model.* 2010, **50**, 742.
11 Molecular Operating Environment (MOE), Version 2016, Chemical Computing Group, Montreal, Canada, 2016. http://www. chemcomp.com/

These descriptors were generated for every compound in the initial data set, and all subsequent training and test sets. To determine the most relevant QSAR method, we evaluated random forest (RF)[12] and support vector machine (SVM)[13] methods. All 2D QSAR modeling was performed using our in-house software; MIX.

The standard approach for determining the best descriptor set is to perform cross-validation. Thus, we used a leave-many-out 10-fold cross-validation (CV) procedure. The initial data set was randomly divided in half 10 times to form the training and test sets. The average cv-R^2 from this process was used to determine the predictivity of each 2D QSAR model. We found the minimum descriptor type necessary for maximum cv-R^2 was also one of the simplest; APC descriptors were sufficient to predict enantioselectivity with a CV-$R^2 = 0.34$ using the RF method. Furthermore, APC descriptors alone significantly outperformed other single type descriptors, ECFP4 had a cv-$R^2 = 0.17$ using the RF method. Interestingly, including TT(C), ECFP4, SP3CARBONS, or MOE_2D descriptors provided no meaningful improvement in the CV-R^2. This implied that specific chemical groups and their organization relative to one another had the most meaningful impact on prediction of whether a bistriflimide catalyst would be enantioselective. The difference between cv-R^2 for RF and SVM methods was slight for the initial model building, however by following the two as we progressed our evaluation of 2D QSAR for prospective catalyst design, we found the RF method performed significantly better for some of the more divergent catalysts in VS2.

Following screening results for VS1, these eleven new bistriflimide catalysts were added to the original dataset and a new cross-validated QSAR model was generated. This model were used to predict activity of the bistriflamides in VS2.

[12] V. Svetnik, A. Liaw, A. C. Tong, J. C. Culberson, R. P. Sheridan and B. P. Feuston, *J. Chem. Inf. Comput. Sci.* 2003, **43**, 1947.

[13] C. Cortes and V. N. Vapnik, *Machine Learning* 1995, **20**, 273.

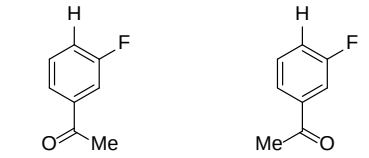
Table S1: The ten most important beneficial and most deleterious single APC-type descriptors from the original cross-validated RF model.

Descriptor	Weight (model coefficient)
OX1sp2CX3sp207	0.739
SX4sp3CX2sp205	0.483
CX3sp3RCX2sp203	0.462
FX1sp3CX2sp206	0.432
OX1sp2CX2sp206	0.363
CX3sp3RCX2sp202	0.244
SX4sp3CX2sp204	0.157
CX4sp3CX2sp206	0.141
FX1sp3CX2sp207	0.107
FX1sp3CX4sp306	0.103
OX1sp2CX3sp210	–0.204
CX3sp2CX3sp205	–0.248
CX2sp2CX2sp203	–0.254
CX3sp2CX2sp204	–0.27
CX3sp2CX2sp206	–0.294
CX3sp2CX2sp203	–0.362
CX3sp2CX3sp202	–0.388
CX3sp2CX1sp304	–0.75
CX3sp2CX3sp204	–0.854
CX3sp2CX3sp207	–1.956

2.3 3D-QSAR Modeling

All catalyst surrogate structures were optimized using the functional M06-2X[14] and a triple zeta basis set def2-TZVP,[15] a combination known for its high accuracy in optimization of organic molecules. All geometry optimizations were performed under ideal gas phase approximation and ultrafine integration grid using the Gaussian 09 software.[16] NBO charges were computed using NBO6 (accessed through Gaussian 09). Vibrational frequencies, intensities, and structural characteristics, were acquired from Gaussian 09 optimized structures using the same level of theory. Partition coefficient parameters were calculated based on the optimized structures using Molecular Modeling Pro.

In case two or more conformational isomers with small difference in minimum energy (<2 kcal mol^{-1}) were found, the parameters were averaged using Boltzmann distribution, e.g. 2-F-4-Ac-C_6H_3 bistriflimide **2bb** surrogates:



conformation A E_{rel} = 0 kcal mol$_{-1}$ [**A**] = 0.55 NBO_{C1} = –0.2601

conformation B E_{rel} = 0.123 kcal mol$_{-1}$ [**B**] = 0.45 NBO_{C1} = –0.2624

averaged NBO_{C1} = 0.55(–0.2624) + 0.45(–0.2601) = –0.2611

Boltzmann distribution

$$\frac{[\mathbf{A}]}{[\mathbf{B}]} = e^{\frac{E_B - E_A}{kT}}$$

where [**A**] is the concentration of conformer **A**, [**B**] is the concentration of conformer **B**, E_B is the relative free energy of conformer **B**, E_A is the relative free energy of conformer **A**, k is the Boltzmann constant, and T is the temperature.

[14] Y. Zhao and D. G. Truhlar, *Acc. Chem. Res*. 2008, **41**, 157.
[15] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* 2005, **7**, 3297.
[16] M. J. Frisch *et. al.* Gaussian 09, revision D. 01. Gaussian, Inc., Wallingford CT: 2009.

Table S2. Training set parameter data

	Catalyst	Parameter							
2	**Ar**	**NBO1**	**NBOC2**	**NBOC2avg**	**IRvArylSymm**	**P**	**polar i**	**polar a**	**FX**
a	2,4-F_2-C_6H_3	–0.304	–0.173	0.118	1694	1260	64.3	38.9	22
b	2-F-C_6H_4	–0.231	–0.197	–0.122	1690	442	63.9	38.2	22
c	2-OTf-C_6H_4	–0.235	–0.194	0.019	1684	314	104.1	54.0	24
d	4-CN-C_6H_4	–0.190	–0.208	–0.208	1684	9.30	80.2	61.0	24
e	Ph	–0.214	–0.214	–0.214	1671	148	63.8	37.9	24
f	1-naphtyl	–0.182	–0.212	–0.138	1653	1540	111.6	88.0	18
g	3-CF_3-C_6H_4	–0.206	–0.196	–0.191	1679	5550	76.4	41.7	18
h	4-Cl-C_6H_4	–0.220	–0.197	–0.197	1666	440	77.0	52.4	24
i	2-Cl-C_6H_4	–0.284	–0.192	0.097	1690	440	77.0	52.4	18
j	2-Br-C_6H_4	–0.236	–0.197	–0.145	1668	440	84.9	59.3	18
k	2-I-C_6H_4	–0.244	–0.198	–0.173	1663	189	95.9	71.6	18
l	4-OMe-C_6H_4	–0.230	–0.200	–0.200	1684	59.3	81.8	49.8	24
m	2,4-$(CF_3)_2$-C_6H_3	–0.169	–0.201	–0.172	1709	132000	89.2	45.1	24
n	3-NO_2-C_6H_4	–0.190	–0.207	–0.094	1672	45.3	80.0	56.0	18
o	2-CF_3-C_6H_4	–0.185	–0.206	–0.179	1679	5550	76.4	41.7	24
p	2-Me-C_6H_4	–0.214	–0.204	–0.115	1669	333	76.8	44.4	18
q	2-pyridyl	0.033	–0.259	–0.339	1665	1.73	58.7	34.9	18
r	2-furyl	0.084	–0.324	–0.360	1545	1.31	44.2	21.7	18
s	3,5-$(CF_3)_2$-C_6H_3	–0.201	–0.168	–0.169	1687	132000	89.2	45.1	12
t	3,5-$(OMe)_2$-C_6H_3	–0.170	–0.317	–0.317	1670	3.68	100.0	59.1	12
aa	C_6F_5	–0.340	0.341	0.341	1671	47100	65.8	41.5	0

Table S3. Virtual screen 1 parameter data

Catalyst		Parameter							
2	**Ar**	**NBO1**	**NBOC2**	**NBOC2avg**	**IRvArylSymm**	**P**	**polar i**	**polar a**	**FX**
ab	2-F-4-Br-C_6H_3	–0.289	–0.177	0.113	1669	1770	85.4	59.5	22
ac	2-F-4-Cl-C_6H_3	–0.291	–0.178	0.113	1676	1770	77.5	52.8	22
ad	2-F-4-CF_3-C_6H_3	–0.267	–0.186	0.104	1699	14400	76.9	42.4	22
ae	2,4-Cl_2-C_6H_3	–0.237	–0.183	–0.109	1660	208	90.9	63.9	18
af	4-CF_3-C_6H_4	–0.196	–0.206	–0.206	1694	5550	76.4	41.7	24
ag	4-F-C_6H_4	–0.234	–0.192	–0.192	1681	442	63.9	38.2	24
ah	3-F-C_6H_4	–0.192	–0.234	–0.259	1690	442	63.9	38.2	22
ai	3,4-Cl_2-C_6H_3	–0.205	–0.205	–0.209	1663	2470	89.6	60.1	18
aj	2-F-4-CN-C_6H_3	–0.260	–0.188	0.102	1690	65.6	80.7	61.1	22
ak	3,4,5-F_3-C_6H_2	–0.192	–0.287	–0.287	1704	4860	64.7	39.6	20

Table S4. Virtual screen 2 parameter data

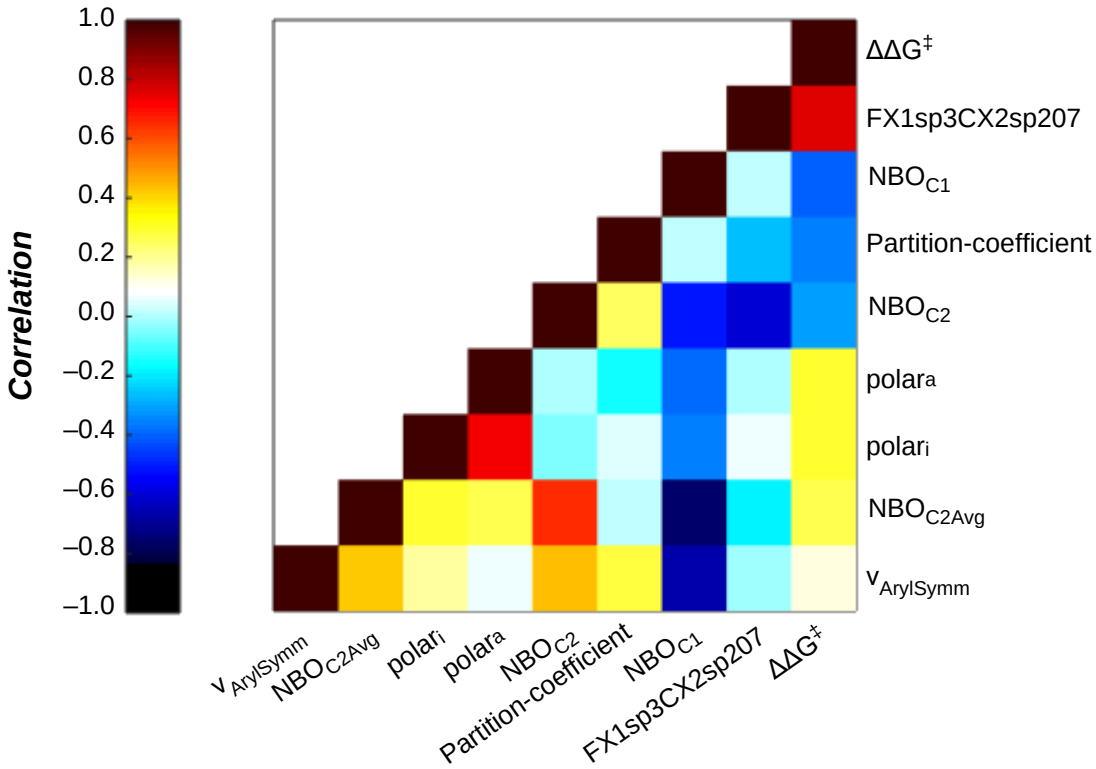
Catalyst		Parameter							
2	**Ar**	**NBO1**	**NBOC2**	**NBOC2avg**	**IRvArylSymm**	**P**	**polar i**	**polar a**	**FX**
al	2-F-4-$SiMe_3$-C_6H_3	–0.280	–0.194	0.097	1679	7370	126	51.6	22
am	2-F-4-I-C_6H_3	–0.287	–0.179	0.112	1667	291	96.1	71.5	22
an	2-F-4-Bpin-C_6H_3	–0.269	–0.201	0.089	1694	264	152	69.4	22
ao	2-F-4-$SnMe_3$-C_6H_3	–0.285	–0.193	0.098	1673	7370	141	53.0	22
ap	ferrocenyl	–0.305	–0.305	–0.305	1479	0.0000306	116	16.3	24
aq	2-OCF_3-C_6H_4	–0.237	–0.198	0.023	1681	117	80.5	44.6	24
ar	2-F-4-SF_5-C_6H_3	–0.263	–0.183	0.106	1682	19700	87.8	47.7	22
as	2-Cl-4-SF_5-C_6H_3	–0.209	–0.190	–0.117	1663	27500	101	60.0	18
at	2-F-5-SF_5-C_6H_3	–0.274	–0.171	0.117	1690	14100	87.6	48.8	14
au	2-Cl-5-SF_5-C_6H_3	–0.222	–0.177	–0.106	1668	19700	102	69.5	12
av	2-Cl-4-$SiMe_3$-C_6H_3	–0.226	–0.200	–0.122	1659	10300	139	63.7	18
aw	2-Cl-4-$SnMe_3$-C_6H_3	–0.231	–0.199	–0.121	1647	10300	155	65.5	18
ax	2-Cl-4-Bpin-C_6H_3	–0.215	–0.207	–0.130	1674	43.6	166	79.9	18
ay	2-F-4-OCF_3-C_6H_3	–0.284	–0.178	0.112	1690	993	81.1	45.3	28
az	2-Cl-4-OCF_3-C_6H_3	–0.230	–0.184	–0.110	1675	164	94.5	57.6	24
ba	2-F-4-OAc-C_6H_3	–0.261	–0.195	0.096	1689	136	91.2	58.6	22
bb	2-Cl-4-OAc-C_6H_3	–0.207	–0.200	–0.125	1676	161	105	70.1	18
bc	2-F-4-(ethynyl)-C_6H_3	–0.276	–0.190	0.100	1688	1420	89.2	74.9	22
bd	2-Cl-4-(ethynyl)-C_6H_3	–0.222	–0.196	–0.121	1677	433	103	83.3	18
be	2,3-F_2-C_6H_3	–0.269	–0.211	0.056	1695	1770	64.3	38.9	16
bf	2,5-F_2-C_6H_3	–0.262	–0.262	0.053	1701	1770	64.1	38.9	16
bg	2-F-3-Cl-C_6H_3	–0.270	–0.200	0.086	1670	2470	77.2	52.6	14
bh	2-F-3-Me-C_6H_3	–0.278	–0.203	0.093	1671	1330	76.9	44.6	14
bi	2-F-4-Me-C_6H_3	–0.295	–0.185	0.107	1700	953	77.1	44.8	22
bj	2-F-5-CF_3-C_6H_3	–0.276	–0.163	0.120	1690	1700	76.7	43.0	14
bk	2-F-5-Cl-C_6H_3	–0.268	–0.209	0.086	1682	1770	77.1	53.4	14
bl	3,4-F_2-Me-C_6H_3	–0.211	–0.211	–0.240	1694	1770	64.3	38.9	22
bm	3-Br-C_6H_4	–0.197	–0.218	–0.227	1668	442	84.9	59.3	18
bn	3-Cl-C_6H_4	–0.197	–0.220	–0.226	1690	442	77.0	52.4	18
bo	3-CN-C_6H_4	–0.208	–0.189	–0.175	1662	9.30	80.2	61.0	18
bp	2-CN-C_6H_4	–0.160	–0.208	–0.191	1662	9.30	80.2	61.0	18
bq	4-Br-C_6H_4	–0.218	–0.197	–0.197	1661	442	84.9	59.3	24
br	4-I-C_6H_4	–0.216	–0.198	–0.198	1656	189	95.9	71.6	24
bs	3-I-C_6H_4	–0.198	–0.216	–0.230	1663	189	95.9	71.6	18
bt	2-CHF_2-C_6H_4	–0.189	–0.207	–0.166	1674	2760	77.5	43.1	22

Table S5. Model predictions

Catalyst		predicted %ee (ΔΔG‡)												measured	
2	**Ar**	**Model A**		**Model B'**		**Model C**		**Model D**		**Model E**		**Model F**		**%ee (ΔΔG‡)**	
a	2,4-F_2-C_6H_3	87.9	1.71	86.5	1.64	86.6	1.64	87.6	1.69	89.7	1.81	88.1	1.72	89.7	1.81
b	2-F-C_6H_4	90.3	1.85	83.2	1.49	84.2	1.53	82.1	1.44	86.0	1.61	84.8	1.55	89.0	1.77
c	2-OTf-C_6H_4	91.1	1.91	83.6	1.50	86.7	1.65	89.1	1.78	87.5	1.68	84.3	1.53	88.3	1.73
d	4-CN-C_6H_4	90.6	1.87	81.5	1.42	85.3	1.58	82.6	1.46	80.2	1.37	82.2	1.45	87.1	1.66
e	Ph	88.1	1.72	80.5	1.38	86.1	1.61	83.3	1.49	84.5	1.54	84.2	1.53	84.7	1.55
f	1-naphtyl	77.4	1.28	80.8	1.39	75.7	1.23	81.2	1.41	84.4	1.54	85.2	1.57	84.5	1.54
g	3-CF_3-C_6H_4	86.4	1.62	83.2	1.49	76.7	1.26	74.4	1.19	78.3	1.31	79.2	1.34	83.7	1.51
h	4-Cl-C_6H_4	88.3	1.73	83.2	1.49	86.3	1.62	84.5	1.54	82.9	1.47	83.6	1.50	82.2	1.45
i	2-Cl-C_6H_4	89.0	1.77	84.0	1.52	80.7	1.39	83.6	1.50	83.2	1.49	82.3	1.45	81.9	1.44
j	2-Br-C_6H_4	88.0	1.71	83.2	1.49	78.5	1.32	78.5	1.32	79.9	1.36	80.0	1.37	81.3	1.41
k	2-I-C_6H_4	85.0	1.56	83.1	1.48	78.8	1.33	77.5	1.29	82.2	1.45	82.6	1.46	80.8	1.39
l	4-OMe-C_6H_4	85.9	1.60	82.8	1.47	86.6	1.64	82.8	1.47	83.0	1.48	84.0	1.52	80.7	1.39
m	2,4-$(CF_3)_2$-C_6H_3	90.1	1.84	78.6	1.32	78.8	1.33	81.0	1.40	79.9	1.36	81.5	1.42	77.7	1.29
n	3-NO_2-C_6H_4	84.6	1.54	81.7	1.43	76.2	1.25	80.4	1.38	76.1	1.24	76.7	1.26	77.27	1.28
o	2-CF_3-C_6H_4	91.8	1.97	81.8	1.43	84.9	1.56	84.1	1.52	80.4	1.38	82.5	1.46	71.4	1.11
p	2-Me-C_6H_4	82.5	1.46	82.1	1.44	77.4	1.28	80.0	1.37	78.3	1.31	78.7	1.32	69.57	1.07
q	2-pyridyl	86.1	1.61	70.7	1.10	62.1	0.90	65.1	0.97	59.3	0.85	60.0	0.86	64.7	0.96
r	2-furyl	54.1	0.75	50.1	0.68	58.1	0.83	59.2	0.85	63.7	0.94	59.1	0.84	58.9	0.84
s	3,5-$(CF_3)_2$-C_6H_3	80.6	1.39	54.3	0.76	51.9	0.72	60.9	0.88	60.4	0.87	62.1	0.90	56.3	0.79
t	3,5-$(OMe)_2$-C_6H_3	61.4	0.89	52.7	0.73	61.1	0.89	50.6	0.69	44.0	0.59	44.7	0.60	47.9	0.65
aa	C_6F_5	69.7	1.07	23.5	0.30	31.7	0.41	24.0	0.30	22.5	0.28	21.7	0.27	22.9	0.29
ab	2-F-4-Br-C_6H_3	86.7	1.64	85.9	1.60	86.1	1.62	92.0	1.98	86.9	1.65	86.7	1.64	90.2	1.85
ac	2-F-4-Cl-C_6H_3	87.0	1.66	85.8	1.60	86.2	1.62	90.9	1.89	87.1	1.66	86.5	1.63	89.3	1.79
ad	2-F-4-CF_3-C_6H_3	87.7	1.70	83.6	1.50	84.9	1.56	86.1	1.61	85.7	1.59	85.9	1.61	88.3	1.73
ae	2,4-Cl_2-C_6H_3	73.3	1.16	85.3	1.58	78.5	1.32	81.8	1.43	81.5	1.42	81.3	1.41	86.2	1.62
af	4-CF_3-C_6H_4	84.9	1.56	81.8	1.43	85.3	1.58	81.7	1.43	81.1	1.40	83.0	1.48	85.3	1.58
ag	4-F-C_6H_4	82.0	1.44	84.0	1.52	86.7	1.64	83.2	1.49	86.5	1.63	85.7	1.60	84.1	1.52
ah	3-F-C_6H_4	82.8	1.47	76.6	1.26	82.7	1.47	77.1	1.27	82.1	1.44	81.6	1.42	79.9	1.36
ai	3,4-Cl_2-C_6H_3	69.8	1.07	81.9	1.44	76.9	1.27	75.3	1.22	77.9	1.30	78.0	1.30	78.9	1.33
aj	2-F-4-CN-C_6H_3	87.1	1.66	84.6	1.54	85.2	1.57	88.1	1.72	85.0	1.56	85.3	1.58	78.5	1.32
ak	3,4,5-F_3-C_6H_2	62.3	0.91	66.3	0.99	79.6	1.35	71.6	1.12	77.2	1.28	77.0	1.27	75.9	1.24
al	2-F-4-$SiMe_3$-C_6H_3	81.8	1.43	83.3	1.49	85.6	1.59	90.0	1.83	95.2	2.30	85.5	1.59	90.4	1.86
am	2-F-4-I-C_6H_3	86.7	1.64	85.9	1.60	86.1	1.61	92.3	2.00	88.1	1.72	88.1	1.72	88.8	1.76
an	2-F-4-Bpin-C_6H_3	82.2	1.45	82.5	1.46	85.5	1.59	86.8	1.65	99.2	3.41	86.4	1.63	87.1	1.66

Table S5. continued

Catalyst 2	Ar	predicted %ee ($\Delta\Delta G^{‡}$) Model A		Model B'		Model C		Model D		Model E		Model F		measured %ee ($\Delta\Delta G^{‡}$)
ao	2-F-4-$SnMe_3$-C_6H_3	81.8	1.43	83.3	1.49	84.6	1.55	89.8	1.82	85.7	1.59	85.4	1.58	
ap	ferrocenyl	75.4	1.22	82.3	1.45	75.7	1.23	80.8	1.40	82.4	1.46	79.0	1.33	
aq	2-OCF_3-C_6H_4	80.7	1.39	85.0	1.56	82.6	1.46	86.5	1.63	85.7	1.59	84.9	1.56	
ar	2-F-4-SF_5-C_6H_3	87.0	1.66	83.8	1.51	63.5	0.93	69.9	1.08	67.7	1.02	66.0	0.99	
as	2-Cl-4-SF_5-C_6H_3	76.8	1.26	82.5	1.46	77.5	1.29	81.1	1.41	96.8	2.57	80.0	1.37	
at	2-F-5-SF_5-C_6H_3	71.2	1.11	83.5	1.50	85.8	1.60	91.2	1.91	98.2	2.91	85.7	1.60	
au	2-Cl-5-SF_5-C_6H_3	66.3	0.99	82.6	1.46	77.7	1.29	83.0	1.48	99.1	3.34	80.6	1.39	
av	2-Cl-4-$SiMe_3$-C_6H_3	60.5	0.87	81.6	1.43	77.5	1.29	78.4	1.31	99.7	3.95	83.6	1.50	
aw	2-Cl-4-$SnMe_3$-C_6H_3	60.5	0.87	85.9	1.60	86.0	1.61	88.4	1.73	86.6	1.64	86.6	1.64	
ax	2-Cl-4-Bpin-C_6H_3	63.2	0.93	85.2	1.57	78.2	1.31	79.1	1.34	81.8	1.43	80.1	1.37	
ay	2-F-4-OCF_3-C_6H_3	86.7	1.64	83.6	1.50	85.3	1.58	88.2	1.72	85.4	1.58	84.9	1.56	
az	2-Cl-4-OCF_3-C_6H_3	76.6	1.26	82.7	1.47	77.1	1.27	78.3	1.31	83.0	1.48	80.4	1.38	
ba	2-F-4-OAc-C_6H_3	80.5	1.39	84.2	1.53	85.7	1.60	88.5	1.74	86.1	1.61	88.1	1.72	
bb	2-Cl-4-OAc-C_6H_3	57.4	0.81	83.4	1.50	77.8	1.30	78.3	1.31	83.4	1.49	85.5	1.59	
bc	2-F-4-(ethynyl)-C_6H_3	86.7	1.64	56.9	0.80	88.7	1.75	83.2	1.49	90.4	1.86	90.3	1.85	
bd	2-Cl-4-(ethynyl)-C_6H_3	72.7	1.15	83.1	1.48	94.3	2.20	95.4	2.33	63.9	0.94	75.6	1.23	
be	2,3-F_2-C_6H_3	71.5	1.12	80.9	1.40	76.5	1.25	78.0	1.30	80.9	1.40	78.7	1.32	
bf	2,5-F_2-C_6H_3	66.8	1.00	70.7	1.10	76.1	1.24	75.7	1.23	77.3	1.28	75.7	1.23	
bg	2-F-3-Cl-C_6H_3	71.6	1.12	82.7	1.47	72.6	1.14	83.7	1.51	72.8	1.15	72.4	1.14	
bh	2-F-3-Me-C_6H_3	43.2	0.58	82.2	1.45	73.1	1.16	83.6	1.50	73.2	1.16	73.5	1.17	
bi	2-F-4-Me-C_6H_3	59.6	0.85	85.0	1.56	86.3	1.63	85.9	1.60	87.0	1.66	86.8	1.65	
bj	2-F-5-CF_3-C_6H_3	73.2	1.16	87.6	1.69	73.0	1.15	78.8	1.33	76.1	1.24	76.0	1.24	
bk	2-F-5-Cl-C_6H_3	68.6	1.05	81.3	1.41	72.5	1.14	80.1	1.37	71.8	1.12	71.7	1.12	
bl	3,4-F_2-Me-C_6H_3	78.8	1.33	80.9	1.40	83.4	1.50	77.5	1.29	84.2	1.53	83.3	1.49	
bm	3-Br-C_6H_4	76.8	1.33	79.7	1.36	76.6	1.26	73.5	1.17	75.8	1.23	76.8	1.26	
bn	3-Cl-C_6H_4	78.3	1.31	79.3	1.34	76.6	1.26	73.3	1.16	75.9	1.24	76.3	1.25	
bo	3-CN-C_6H_4	77.7	1.29	84.4	1.54	77.1	1.27	77.7	1.29	78.5	1.32	79.1	1.34	
bp	2-CN-C_6H_4	74.6	1.20	81.5	1.42	74.6	1.20	76.6	1.26	73.7	1.17	75.3	1.22	
bq	4-Br-C_6H_4	80.8	1.40	83.2	1.49	86.2	1.62	84.9	1.56	82.6	1.46	83.8	1.51	
br	4-I-C_6H_4	80.8	1.40	83.1	1.48	86.1	1.62	85.3	1.58	84.1	1.52	85.6	1.59	
bs	3-I-C_6H_4	76.7	1.26	80.1	1.37	76.6	1.26	73.8	1.18	78.3	1.31	79.6	1.35	
bt	2-CHF_2-C_6H_4	80.7	1.39	81.7	1.43	82.5	1.46	82.5	1.46	80.3	1.38	81.7	1.43	



Scheme S1. Correlation color map.

6. Optimized geometries

1,3-difluorobenzene (surrogate for **2a**)

C	-0.65624800	1.39311900	-0.00000100
C	0.53242600	0.69295300	-0.00001700
C	-0.65624300	-1.39311900	0.00000400
C	-1.85583500	-0.69361400	-0.00000600
C	-1.85583600	0.69361000	0.00000500
H	-0.62236100	2.47421300	0.00000500
H	-0.62235600	-2.47421200	0.00001000
H	-2.79014200	-1.23795900	-0.00000800
H	-2.79014600	1.23795100	0.00001500
C	0.53243000	-0.69294900	0.00000200
F	1.69893500	-1.34029200	0.00000100
F	1.69893500	1.34029200	0.00000500

fluorobenzene (surrogate for **2b**, **2ah**, **2ai**)

C	-1.12812800	1.20147200	0.00000100
C	0.25978500	1.21083500	0.00000100
C	0.25978900	-1.21082200	0.00000200
C	-1.12809800	-1.20149500	0.00000100
C	-1.82380700	-0.00000400	-0.00000200
H	-1.66640100	2.14029400	-0.00000400
H	0.82806400	2.13085200	0.00000600
H	0.82812500	-2.13080500	0.00000300
H	-1.66639100	-2.14030600	-0.00000400
H	-2.90529800	-0.00003300	0.00000000
C	0.92456200	0.00002100	-0.00000100

F	2.26636500	-0.00000500	-0.00000100

phenyl trifluoromethanesulfonate (surrogate for **2c**)

C	2.14729700	-1.20315500	-0.35085100
C	1.45740100	-0.00757100	-0.38692900
C	2.07149700	1.21921700	-0.22801300
C	3.44476200	1.24325900	-0.03267900
C	4.16714700	0.05784900	0.00192300
C	3.52061800	-1.16050200	-0.15547500
H	1.61325000	-2.13545600	-0.47338800
H	1.47905600	2.12343400	-0.25347000
H	3.94853100	2.19215900	0.09389400
H	5.23798700	0.08362200	0.15356200
H	4.08411300	-2.08344000	-0.12645500
O	0.06784100	-0.04001400	-0.63701900
S	-0.88805400	-0.08887600	0.61762600
C	-2.43654000	0.06130400	-0.36802200
F	-2.55584000	-0.96228600	-1.19207700
F	-3.44895100	0.06480500	0.48098800
F	-2.43603700	1.18643800	-1.05777300
O	-0.87014100	-1.37719900	1.21134700
O	-0.75016500	1.08954600	1.39615800

benzonitrile (surrogate for **2d**, **2bp**, **2bq**)

C	1.47499500	-1.20420500	0.00000000
C	0.08998500	-1.21017000	0.00000400
C	0.09003300	1.21020500	0.00000100

C	1.47500400	1.20419000	0.00000400
C	2.16635100	-0.00003400	-0.00000200
H	2.01581600	-2.14109400	-0.00000300
H	-0.46191800	-2.14022800	0.00000600
H	-0.46189000	2.14025100	0.00000200
H	2.01590100	2.14103600	-0.00000100
H	3.24848800	-0.00001600	-0.00000300
C	-0.60127800	0.00001100	-0.00000500
C	-2.03804000	0.00004700	-0.00000600
N	-3.18552700	-0.00003100	0.00000300

benzene (surrogate for **2e**)

C	-0.37246700	1.33747700	0.00000100
C	-1.34481500	0.34622000	-0.00001100
C	-0.97222100	-0.99123800	0.00001100
C	0.37252800	-1.33746000	-0.00000200
C	1.34479900	-0.34628100	-0.00000800
C	0.97217700	0.99128300	0.00000700
H	-0.66259800	2.38034900	-0.00000400
H	-2.39313100	0.61602700	-0.00000400
H	-1.72990600	-1.76432900	0.00000600
H	0.66251800	-2.38037100	0.00000600
H	2.39315100	-0.61595200	-0.00001300
H	1.72996200	1.76427500	0.00001900

Naphthalene (surrogate for **2f**)

C	2.41862300	-0.70647700	0.00000000

C	1.23888200	-1.39494800	-0.00000100
C	0.00000200	-0.70908300	0.00000000
C	0.00000900	0.70909000	0.00000000
C	1.23889100	1.39494600	0.00000000
C	2.41862800	0.70646900	0.00000000
H	-1.23331000	-2.47856900	0.00000100
H	3.35984200	-1.24086000	0.00000200
H	1.23331600	-2.47856700	-0.00000200
C	-1.23889900	-1.39494800	0.00000000
C	-1.23889500	1.39495300	0.00000000
H	1.23333100	2.47856700	-0.00000200
H	3.35985000	1.24084700	0.00000100
C	-2.41861800	0.70648300	-0.00000100
C	-2.41862400	-0.70648400	0.00000000
H	-1.23332200	2.47857200	0.00000100
H	-3.35984500	1.24085200	-0.00000200
H	-3.35985500	-1.24084500	0.00000000

trifluorotoluene (surrogate for **2g**, **2o**, **2ag**)

C	-2.12720800	1.20282400	0.00172300
C	-0.74057100	1.20690400	-0.01802400
C	-0.05444600	0.00097700	-0.02892100
C	-0.73934300	-1.20586300	-0.01811300
C	-2.12579400	-1.20344900	0.00165400
C	-2.81866000	-0.00062600	0.01189700
H	-2.66779200	2.13997900	0.00645700
H	-0.19215300	2.13940500	-0.03215200

H	-0.18975000	-2.13774600	-0.03233900
H	-2.66537200	-2.14118400	0.00631700
H	-3.90081500	-0.00126100	0.02521800
C	1.44748300	0.00039200	-0.00277000
F	1.96083900	1.08561200	-0.59351300
F	1.95973000	-1.07000800	-0.62108900
F	1.92022100	-0.01628700	1.25258300

chlorobenzene (surrogate for **2h**, **2i**, **2bo**)

C	-1.56127600	-1.20063300	-0.00000300
C	-0.17355800	-1.20784000	0.00002200
C	-0.17358100	1.20786100	0.00001600
C	-1.56119800	1.20067000	-0.00000100
C	-2.25751200	-0.00001100	-0.00001800
H	-2.09766400	-2.14034700	-0.00002000
H	0.38047200	-2.13622500	0.00001100
H	0.38059800	2.13615600	0.00001200
H	-2.09767100	2.14033800	-0.00000500
H	-3.33915600	0.00007200	-0.00001100
C	0.50563200	-0.00004000	0.00001400
Cl	2.24131600	-0.00000200	-0.00001000

bromobenzene (surrogate for **2j**, **2bn**, **2br)**

C	-2.16557500	-1.20057900	0.00000200
C	-0.77758600	-1.20834700	-0.00000700
C	-0.77758400	1.20834500	-0.00000200
C	-2.16552900	1.20060500	-0.00000300

C	-2.86149800	0.00000200	0.00000400
H	-2.70178400	-2.14069500	0.00000800
H	-0.22750000	-2.13912700	-0.00000500
H	-0.22742100	2.13908100	0.00000000
H	-2.70177200	2.14070300	0.00000200
H	-3.94330300	0.00004600	0.00000700
C	-0.09903600	-0.00002600	-0.00000200
Br	1.79664700	0.00000000	0.00000100

iodobenzene (surrogate for **2k**, **2bs**, **2bt)**

C	-2.62370900	-1.20034500	0.00000100
C	-1.23520300	-1.20780300	-0.00000600
C	-1.23520700	1.20780500	-0.00000300
C	-2.62367100	1.20036700	-0.00000200
C	-3.31965900	-0.00000100	0.00000400
H	-3.15977400	-2.14067600	0.00000800
H	-0.69215100	-2.14296100	-0.00000500
H	-0.69208500	2.14292300	-0.00000100
H	-3.15977200	2.14067700	0.00000500
H	-4.40155800	0.00003800	0.00000700
C	-0.55296400	-0.00002100	-0.00000300
I	1.54052500	0.00000000	0.00000100

anisole (surrogate for **2l**)

C	1.31913200	1.34510200	-0.00014800
C	-0.04149600	1.05217100	-0.00032100
C	-0.45259700	-0.27653700	-0.00005400

C	0.49704800	-1.29803500	0.00001900
C	1.84370500	-0.98985600	0.00015800
C	2.26595500	0.33639500	0.00026000
H	1.63306900	2.38138100	-0.00020700
H	-0.75985800	1.85894300	-0.00080000
H	0.14940700	-2.32259400	0.00000500
H	2.57148000	-1.79139400	0.00043400
H	3.32079200	0.57519500	0.00050800
O	-1.75026000	-0.67336500	-0.00080200
C	-2.74305400	0.32561800	0.00064100
H	-2.67459100	0.95555800	-0.89086700
H	-3.69769600	-0.19398200	0.00140800
H	-2.67268500	0.95466800	0.89260800

1,3-bis(trifluoromethyl)benzene (surrogate for **2m**, **2s**)

C	1.20441100	0.33274200	-0.00354600
C	0.01593000	-0.38678900	-0.01852500
C	-1.18415700	0.29915600	-0.02433600
C	-1.20745000	1.68957800	-0.01324000
C	-0.01743000	2.39482600	0.00148500
C	1.19555600	1.71680900	0.00643800
H	0.03151700	-1.46898800	-0.03041900
H	-2.15585400	2.21111300	-0.02203500
H	-0.02985300	3.47608400	0.00691000
H	2.12906000	2.26216600	0.01620800
C	-2.48811600	-0.45015300	-0.00217300
C	2.49741400	-0.43566200	0.00050800

F	-2.34402400	-1.72322000	-0.37564400
F	-3.02589000	-0.45598800	1.22484200
F	-3.39035600	0.10905300	-0.81647400
F	3.56331100	0.36763300	0.01007900
F	2.60075700	-1.22470000	-1.07589000
F	2.58822200	-1.23315900	1.07194100

nitrobenzene (surrogate for **2n**)

C	1.81024100	-1.20547200	-0.00000600
C	0.42420900	-1.21416500	-0.00002000
C	-0.23810700	-0.00000400	-0.00001500
C	0.42421500	1.21417000	0.00000000
C	1.81023100	1.20547700	0.00001300
C	2.50015500	-0.00000500	0.00001100
H	2.35152600	-2.14192700	-0.00001000
H	-0.14363600	-2.13279500	-0.00003000
H	-0.14365100	2.13278700	0.00000300
H	2.35153900	2.14191800	0.00002800
H	3.58234600	0.00001000	0.00002300
N	-1.71480500	-0.00000100	-0.00003000
O	-2.27375700	1.07266200	-0.00000700
O	-2.27376100	-1.07266000	0.00004500

toluene (surrogate for **2p**)

C	-1.19596100	1.19734100	0.00188800
C	0.19220800	1.19658400	-0.00784300
C	0.90738800	0.00276000	-0.01050500

C	0.19561500	-1.19444300	-0.00786900
C	-1.19133700	-1.19994300	0.00187200
C	-1.89296300	-0.00199900	0.00756300
H	-1.73363700	2.13703300	0.00148400
H	0.73066900	2.13733200	-0.01633000
H	0.73774200	-2.13328500	-0.01630400
H	-1.72619500	-2.14125300	0.00142700
H	-2.97511400	-0.00399600	0.01220400
C	2.41136000	0.00101400	0.00827000
H	2.78529900	-0.12064300	1.02735700
H	2.81151300	-0.81921500	-0.58801800
H	2.81185700	0.93613800	-0.38207200

pyridine (surrogate for **2q**)

C	-1.19185800	0.66815500	0.00000600
C	-1.13605100	-0.71901900	-0.00009400
C	1.13670200	-0.71802400	-0.00001300
C	1.19128600	0.66912700	-0.00002200
C	-0.00061100	1.37628600	0.00002200
H	-2.14723400	1.17515800	0.00007300
H	-2.04965200	-1.30417200	0.00009800
H	2.05074800	-1.30246400	0.00002500
H	2.14619400	1.17701200	-0.00002200
H	-0.00103100	2.45885800	0.00003500
N	0.00059600	-1.40907700	0.00005700

furan (surrogate for **2r**)

C	-1.08714000	-0.34679400	0.00014700
C	-0.71635800	0.95272000	-0.00012300
C	0.71618700	0.95283700	0.00004400
C	1.08720000	-0.34661300	0.00013400
O	0.00010400	-1.14990700	-0.00018500
H	-2.04137300	-0.84360700	0.00021700
H	-1.37172300	1.80647700	-0.00022500
H	1.37139300	1.80671100	0.00007500
H	2.04153500	-0.84322600	0.00021100

1,3-Dimethoxybenzene (surrogate for **2t**)

C	1.04508600	-0.21023600	0.00010000
C	-0.28218000	-0.64830500	0.00012800
C	-1.02020000	1.65077400	-0.00001400
C	0.29393800	2.05977200	-0.00001800
C	1.34607500	1.14360100	0.00003400
H	-0.46043900	-1.71312200	0.00028300
H	-1.84028100	2.35486400	-0.00007400
H	0.52022400	3.11848900	-0.00002900
H	2.36605400	1.49612700	0.00002600
C	-1.30755400	0.28156500	0.00001700
O	-2.62245000	-0.04759500	-0.00000900
O	1.97473100	-1.19928200	0.00016800
C	-2.95557600	-1.41700500	-0.00010300
H	-4.04119300	-1.46669800	-0.00082300
H	-2.56736400	-1.91769200	-0.89141800
H	-2.56857900	-1.91750400	0.89185500

C	3.33172500	-0.82029100	-0.00018400
H	3.58008100	-0.23789600	0.89124800
H	3.90583400	-1.74312600	-0.00048200
H	3.57952100	-0.23766700	-0.89163000

pentafluorobenzene (surrogate for **2aa**)

C	0.00000500	1.11117300	-0.00000100
C	-1.20184400	0.42203600	-0.00000400
C	-1.18722500	-0.96226900	-0.00000200
C	0.00002500	-1.66704300	0.00000000
C	1.18721300	-0.96227300	0.00000400
C	1.20181900	0.42209000	0.00000200
H	-0.00002000	-2.74738900	0.00000200
F	2.34795300	-1.60998100	-0.00000200
F	2.34624400	1.09099000	0.00000100
F	-0.00006400	2.43419100	-0.00000300
F	-2.34622600	1.09101600	0.00000400
F	-2.34790100	-1.61009300	-0.00000100

1-bromo-3-fluorobenzene (surrogate for **2ac**)

C	0.21428000	0.18740200	0.00000200
C	-0.69022100	-0.86211100	0.00000500
C	-2.03634700	-0.54860800	0.00000200
C	-2.49917500	0.75288800	-0.00000400
C	-1.56641900	1.77938300	0.00000400
C	-0.20489400	1.50874300	0.00000100
H	-0.37026800	-1.89409900	0.00001700

H	-3.56325800	0.94234400	-0.00000700
H	-1.90430500	2.80735000	-0.00000700
H	0.52030600	2.30968200	0.00000000
F	-2.92137900	-1.55253100	-0.00000300
Br	2.06590300	-0.20282000	-0.00000100

1-chloro-3-fluorobenzene (surrogate for **2ad**)

C	0.82615000	0.04925300	0.00000000
C	-0.16908400	-0.91498200	0.00000500
C	-1.48109000	-0.48248300	0.00000800
C	-1.82581900	0.85538300	0.00000000
C	-0.80520500	1.79458900	0.00000100
C	0.52573300	1.40293100	0.00000100
H	0.06133300	-1.97057400	0.00001200
H	-2.86867100	1.13993700	-0.00000400
H	-1.04933000	2.84872000	-0.00000800
H	1.32438500	2.13062200	0.00000000
F	-2.45339900	-1.40226700	-0.00000600
Cl	2.48169200	-0.45626200	-0.00000200

1-fluoro-3-(trifluoromethyl)benzene (surrogate for **2ae**)

C	-0.26468400	0.18419800	-0.03049800
C	0.66262800	-0.84820600	-0.02143000
C	2.00146400	-0.51555700	-0.00119100
C	2.43844800	0.79587100	0.01084700
C	1.49328600	1.80975500	0.00194800
C	0.13785900	1.51075200	-0.01795300

H	0.36040400	-1.88656400	-0.03677200
H	3.49998400	1.00167400	0.02370100
H	1.81722200	2.84179700	0.00812900
H	-0.60020000	2.30035500	-0.03102400
C	-1.72881300	-0.15896600	-0.00277300
F	2.90805600	-1.49928200	0.00349600
F	-2.48514100	0.82732000	-0.49289000
F	-2.15239800	-0.39272100	1.24724900
F	-1.99479900	-1.26024400	-0.71316000

1,3-dichlorobenzene (surrogate for **2af**)

C	0.00006200	2.06518700	-0.00000100
C	-1.20655700	1.38110000	0.00002000
C	-1.18987200	-0.00469000	0.00001900
C	-0.00000500	-0.71520200	0.00001200
C	1.18979000	-0.00481200	-0.00001300
C	1.20655400	1.38109000	-0.00001600
H	0.00003300	3.14686800	-0.00001800
H	-2.14908600	1.90952000	0.00001000
H	-0.00012800	-1.79546400	0.00001600
H	2.14918100	1.90933300	-0.00000800
Cl	2.68665200	-0.87606500	0.00000200
Cl	-2.68664200	-0.87607000	-0.00000900

1,2-dichlorobenzene (surrogate for **2aj**)

C	-1.17511400	1.38601100	0.00000100
C	0.02946200	0.69596500	-0.00001800

C	0.02946600	-0.69596600	-0.00000300
C	-1.17511000	-1.38601000	0.00000200
C	-2.37449700	-0.69385200	-0.00000500
C	-2.37449700	0.69384700	0.00000700
H	-1.15513400	2.46708600	0.00000500
H	-1.15512900	-2.46708700	0.00001000
H	-3.30701000	-1.24187800	-0.00000300
H	-3.30701300	1.24187300	0.00001600
Cl	1.50488500	-1.58736000	0.00000200
Cl	1.50488100	1.58736300	0.00000200

3-fluorobenzonitrile (surrogate for **2ak**)

C	-0.92639900	0.04904800	0.00000000
C	0.06573600	-0.92848800	-0.00000200
C	1.38138500	-0.51513400	0.00000000
C	1.73879200	0.82043300	0.00000300
C	0.73578300	1.77871300	0.00000400
C	-0.59798000	1.40311600	0.00000300
H	-0.17636100	-1.98192200	-0.00000400
H	2.78640400	1.08922400	0.00000400
H	0.99858700	2.82794000	0.00000700
H	-1.38505200	2.14393000	0.00000400
F	2.34488000	-1.44147800	-0.00000200
C	-2.30700500	-0.34924700	-0.00000200
N	-3.40991100	-0.66521600	-0.00000300

1,2,3-trifluorobenzene (surrogate for **2al**)

C	-1.19259600	-0.01436700	-0.00000200
C	-0.00001600	-0.71868200	0.00000000
C	1.19258600	-0.01449000	-0.00000700
C	1.20669200	1.36629800	-0.00000400
C	0.00009100	2.05065100	0.00000500
C	-1.20660800	1.36638100	0.00000000
H	2.15811600	1.87921300	-0.00000400
H	0.00011000	3.13161100	0.00001000
H	-2.15795600	1.87943000	0.00000200
F	2.33127100	-0.70510600	0.00000300
F	-0.00011400	-2.04603200	0.00000300
F	-2.33128600	-0.70497400	-0.00000100

(3-fluorophenyl)trimethylsilane (surrogate for **2am**)

C	-2.36020900	-0.53821900	-0.00000200
C	-1.96009800	1.80171600	0.00000500
H	-2.32844600	2.81964000	0.00002700
C	-1.00620200	-0.80379000	-0.00007400
H	-0.68449700	-1.83969500	-0.00012400
C	-0.09149100	0.25242600	-0.00010600
F	-3.22398100	-1.56673800	0.00000000
C	-0.59218800	1.55538200	-0.00006700
H	0.09138900	2.39601700	-0.00012500
Si	1.75587900	-0.11343100	0.00000300
C	2.69926000	1.50271500	-0.00196000
H	2.46688500	2.10051900	0.88160000
H	3.77449500	1.31367500	-0.00214500

H	2.46624800	2.09878500	-0.88651400
C	2.17250600	-1.10786200	1.53151700
H	1.60376600	-2.03946700	1.56168100
H	3.23397100	-1.36342200	1.55247400
H	1.93933000	-0.54545000	2.43711300
C	2.17199900	-1.11130800	-1.52941800
H	3.23335900	-1.36736800	-1.54984900
H	1.60289000	-2.04273800	-1.55768800
H	1.93909900	-0.55067200	-2.43618500
C	-2.86235000	0.74932400	0.00003900
H	-3.93232000	0.90654200	0.00008300

1-fluoro-3-iodobenzene (surrogate for **2am**)

C	2.48456900	-0.57802700	-0.00000400
C	2.09682100	1.76393400	0.00000400
H	2.46940300	2.77978100	0.00000300
C	1.12820300	-0.84440800	-0.00000400
H	0.78016300	-1.86749100	-0.00000700
C	0.72676200	1.53961500	0.00000500
H	0.03499200	2.36977200	0.00000800
F	3.33509700	-1.61223000	-0.00001200
C	2.99315000	0.70579400	0.00000000
H	4.06334200	0.85758500	-0.00000400
C	0.25671900	0.23400100	0.00000100
I	-1.80153100	-0.12368000	0.00000200

2-(3-fluorophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (surrogate for **2ao**)

C	-3.54189200	0.64082500	0.07540000
C	-3.31165300	-1.71512200	-0.18048700
H	-3.75692200	-2.69586200	-0.28741800
C	-2.17308900	0.80949100	0.09174800
H	-1.75548100	1.80290000	0.19763800
C	-1.34862500	-0.30991100	-0.02934100
C	-1.93193300	-1.57110000	-0.16493000
H	-1.29352200	-2.44085800	-0.25759300
F	-4.33391000	1.71934700	0.19177500
B	0.19693200	-0.14383100	-0.01383800
O	0.82457300	1.06017400	0.15837500
O	1.07378900	-1.18279900	-0.17155300
C	2.39141000	-0.67757100	0.13578700
C	2.22557700	0.85414800	-0.12434100
C	3.05251600	1.74547900	0.78024300
H	4.11676900	1.54646400	0.63822800
H	2.86371700	2.78984500	0.53156100
H	2.80207100	1.59287000	1.82795600
C	2.44283200	1.22984900	-1.58470400
H	2.12054800	2.26042900	-1.73213600
H	3.49399800	1.14858600	-1.86347300
H	1.85442200	0.58933600	-2.24340900
C	2.65866700	-1.00205700	1.59995800
H	2.55985400	-2.07795400	1.74296600
H	3.66393500	-0.70085500	1.89654000
H	1.93695100	-0.50223400	2.24767900
C	3.40829600	-1.36487000	-0.75327500

H	4.39731600	-0.92493300	-0.60871500
H	3.46273600	-2.42232600	-0.49421700
H	3.13696900	-1.28305900	-1.80379500
C	-4.13192600	-0.60181100	-0.05880800
H	-5.21079600	-0.68042700	-0.06667100

(3-fluorophenyl)trimethylstannane (surrogate for **2ap**)

C	-3.03466400	-0.56057400	0.00052000
C	-2.68168500	1.78688700	0.00128600
H	-3.07079900	2.79723700	0.00155200
C	-1.67402300	-0.79986200	0.00016900
H	-1.33649500	-1.83091700	-0.00079900
C	-0.78438500	0.27505000	0.00052300
C	-1.30895200	1.56813200	0.00088600
H	-0.64331900	2.42440800	0.00037800
Sn	1.34944500	-0.06053500	-0.00001100
C	1.86020800	-1.22752300	-1.73784000
H	2.92217500	-1.47132200	-1.74007900
H	1.29339400	-2.15840400	-1.74653900
H	1.62667100	-0.67504000	-2.64704500
C	1.88003900	-1.14492600	1.78462700
H	2.94599500	-1.37020700	1.79602100
H	1.63704500	-0.55905900	2.67018800
H	1.32916000	-2.08410700	1.83229200
C	2.29808200	1.87327600	-0.04984000
H	2.00299200	2.42148400	-0.94426300
H	2.01854100	2.45982800	0.82510600

H	3.38220000	1.76467400	-0.05671600
F	-3.87728200	-1.60678000	0.00002400
C	-3.56272700	0.71581400	0.00116000
H	-4.63561900	0.85153900	0.00131300

ferrocene (surrogate for **2aq)**

C	1.76290800	1.04463100	-0.60052300
C	1.76339000	0.89397200	0.80743200
C	1.76456700	-0.49178600	1.09960400
C	1.76512000	-1.19787000	-0.12814800
H	1.73510900	1.97864000	-1.13776700
H	1.73601800	1.69342400	1.52984000
H	1.73817700	-0.93179600	2.08310100
H	1.73941800	-2.26916700	-0.24260000
Fe	-0.00017000	-0.00034600	-0.00009800
C	-1.76313900	0.02633200	1.20484000
C	-1.76273500	1.15384200	0.34747400
C	-1.76446000	-1.13729700	0.39725300
H	-1.73680500	0.05020500	2.28217200
C	-1.76343800	0.68689000	-0.98989100
H	-1.73676400	2.18561700	0.65794800
C	-1.76463600	-0.72928800	-0.95877000
H	-1.73993400	-2.15474800	0.75235600
H	-1.73735200	1.30129100	-1.87510400
H	-1.74030200	-1.38186900	-1.81630800
C	1.76395300	-0.24813400	-1.17891400
H	1.73765700	-0.47035600	-2.23322600

(trifluoromethoxy)benzene (surrogate for **2ar**)

C	-2.45135600	-1.20244300	0.09110100
C	-1.11699000	1.20921500	-0.29050400
C	-3.11753500	0.00031000	0.28302900
H	-2.96978000	-2.14090600	0.23636300
H	-0.57572900	2.13170400	-0.45074200
H	-4.15829800	0.00063000	0.57894900
C	-2.45079700	1.20276600	0.09090700
H	-2.96886400	2.14146000	0.23591300
C	-0.47137900	-0.00031000	-0.47065700
O	0.85804800	-0.00084600	-0.90757400
C	1.80313700	-0.00001500	0.05184200
F	2.98464600	-0.00029400	-0.53809400
F	1.72555500	-1.07166700	0.84301100
F	1.72532000	1.07266400	0.84156400
C	-1.11752100	-1.20954600	-0.29033900
H	-0.57674500	-2.13230700	-0.45049800

1-pentafluorosulfanyl-3-fluorobenzene (surrogate for **2as**)

C	-2.60624700	-0.54593700	0.00021000
F	-3.46805800	-1.56791800	-0.00011600
C	-2.19344400	1.79573900	-0.00006400
H	-2.55589500	2.81465700	-0.00022600
C	-0.82660800	1.55586700	0.00001700
H	-0.12512500	2.37677400	0.00000500
C	-0.37795500	0.24604200	0.00004600

C	-1.25403500	-0.82610100	0.00003700
H	-0.91427900	-1.85158900	-0.00008900
S	1.39695400	-0.08243100	-0.00000900
F	1.24704700	-1.19510400	1.12163400
F	1.65527800	1.01090800	1.12029100
F	2.94635700	-0.36997300	-0.00010200
F	1.24702600	-1.19496200	-1.12178600
F	1.65499700	1.01116400	-1.12011500
C	-3.09731500	0.74504100	0.00009200
H	-4.16616400	0.90810000	0.00017400

1-pentafluorosulfanyl-3-chlorobenzene (surrogate for **2at**)

C	2.37609700	-0.04605100	-0.00017500
C	1.60187300	2.21635000	0.00006100
H	1.80836300	3.27804800	0.00029800
C	0.28740200	1.77394800	-0.00002800
H	-0.53013900	2.47978500	0.00002500
C	0.04333200	0.41096500	-0.00004500
C	1.07248100	-0.51468200	-0.00004900
H	0.87653200	-1.57661100	0.00009000
S	-1.66091600	-0.18519500	0.00004600
F	-1.34374400	-1.26242800	-1.12087500
F	-2.08191700	0.85513000	-1.12135000
F	-3.14862200	-0.70380000	-0.00000500
F	-1.34384300	-1.26230500	1.12112500
F	-2.08172600	0.85554400	1.12113200
Cl	3.67874000	-1.18385400	0.00005300

C	2.65225500	1.31241800	-0.00011800
H	3.67935100	1.65046600	-0.00018000

1-pentafluorosulfanyl-4-fluorobenzene (surrogate for **2au**)

C	-3.04429600	-0.00003200	0.00001500
C	-0.99416000	1.20797600	-0.00002700
H	-0.45045000	2.14172800	-0.00005700
C	-0.31536600	0.00004000	0.00002100
C	-2.37924300	-1.21111200	-0.00001100
H	-2.94371300	-2.13292300	0.00000300
C	-0.99411800	-1.20794500	0.00003800
H	-0.45035600	-2.14165200	0.00005900
S	1.48471000	0.00007700	-0.00000700
F	1.54279200	-1.12238200	1.12108900
F	1.54258200	-1.12219400	-1.12130700
F	3.06159800	-0.00017700	-0.00005300
F	1.54269500	1.12244400	-1.12121700
F	1.54268200	1.12221900	1.12145500
F	-4.37992600	-0.00005000	0.00000000
C	-2.37929300	1.21107900	0.00003100
H	-2.94381200	2.13285100	-0.00000600

1-pentafluorosulfanyl-4-chlorobenzene (surrogate for **2av**)

C	2.70419000	-0.00000500	0.00000500
C	0.63816700	-1.20665200	0.00006700
H	0.09673300	-2.14182600	0.00008800
C	-0.04116800	0.00006700	0.00002400

C	2.02350000	1.20755100	0.00006500
H	2.57407000	2.13763100	0.00009300
Cl	4.43273000	-0.00004600	-0.00000400
C	0.63821900	1.20676100	-0.00004600
H	0.09685300	2.14196800	-0.00007800
S	-1.84260900	0.00006900	0.00001000
F	-1.89917000	1.12282400	1.12043800
F	-1.89900100	1.12118000	-1.12206700
F	-3.41882300	-0.00023600	-0.00007100
F	-1.89898500	-1.12278300	-1.12029700
F	-1.89895300	-1.12115700	1.12194900
C	2.02344700	-1.20753600	-0.00005800
H	2.57394800	-2.13765500	-0.00010800

(3-chlorophenyl)trimethylsilane (surrogate for **2aw**)

C	-2.08966600	-0.03366600	0.00000800
C	-1.31682100	2.22361600	0.00000200
H	-1.52360600	3.28629900	0.00004300
C	-0.78483500	-0.49618100	-0.00016200
H	-0.61431300	-1.56709700	-0.00029700
C	0.28481300	0.40229900	-0.00024500
C	-0.00463600	1.76745000	-0.00019100
H	0.80297300	2.49022500	-0.00037200
Si	2.04889600	-0.25883700	-0.00006400
C	3.24479100	1.17980000	-0.00620100
H	3.11622900	1.80973900	0.87619500
H	4.27347100	0.81416000	-0.00629900

H	3.11389500	1.80398800	-0.89235000
C	2.29512800	-1.30601200	1.53198400
H	1.58710700	-2.13672700	1.55587900
H	3.30273300	-1.72593900	1.55976400
H	2.14721500	-0.71644400	2.43828300
C	2.29130400	-1.31649700	-1.52548100
H	3.29972400	-1.73439800	-1.55415800
H	1.58515200	-2.14901400	-1.54026500
H	2.13838800	-0.73408600	-2.43554100
Cl	-3.40048800	-1.17348400	0.00005000
C	-2.37119600	1.32420100	0.00011000
H	-3.39850900	1.66218700	0.00023200

(3-chlorophenyl)trimethylstannane (surrogate for **2ax**)

C	-2.80870600	-0.05794300	0.00196300
C	-2.14056700	2.23237700	0.00379400
H	-2.39675700	3.28454200	0.00429300
C	-1.48234300	-0.46094800	0.00131000
H	-1.26764800	-1.52413300	-0.00142200
C	-0.45888900	0.48667600	0.00219300
C	-0.80854500	1.83735400	0.00286500
H	-0.03577400	2.59892900	0.00131800
Cl	-4.06538300	-1.25789500	-0.00014700
Sn	1.60885900	-0.14011900	-0.00036500
C	1.91539500	-1.44921900	-1.68399800
H	2.93254300	-1.83972100	-1.68823300
H	1.22473800	-2.29100500	-1.63610900

H	1.74208100	-0.91616900	-2.61785000
C	2.01149400	-1.20660900	1.82787200
H	3.00005400	-1.66373400	1.79667900
H	1.96570900	-0.53161900	2.68146900
H	1.27170800	-1.99378400	1.97252600
C	2.81444600	1.63894600	-0.15234200
H	2.57914400	2.18220700	-1.06720500
H	2.63618500	2.29727300	0.69779600
H	3.87294700	1.38107900	-0.16893500
C	-3.15260100	1.28459000	0.00346400
H	-4.19445600	1.57494300	0.00368100

2-(3-chloroophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (surrogate for **2ay**)

C	3.28108800	-0.17615800	0.02220400
C	2.79303800	2.15817600	-0.16807400
H	3.13500900	3.18196700	-0.24780800
C	1.92948800	-0.47588800	0.03646800
H	1.60526300	-1.50593900	0.11562800
C	0.99034000	0.55234200	-0.05115700
C	1.43647700	1.87048800	-0.15335700
H	0.70944600	2.67041100	-0.22157900
B	-0.53116000	0.23326500	-0.03055700
O	-1.50126800	1.18998100	-0.15873100
O	-1.04212400	-1.02639600	0.12274000
C	-2.46119500	-0.94441300	-0.13513800
C	-2.76236600	0.56103800	0.15703500
C	-3.85101500	1.16842500	-0.70477900

H	-4.79468500	0.64104700	-0.55089600
H	-3.99411900	2.21299400	-0.42797200
H	-3.59210900	1.12634500	-1.76071800
C	-3.03600700	0.83392100	1.63053800
H	-3.03665300	1.91166100	1.79188600
H	-4.00380300	0.43409100	1.93492700
H	-2.26092100	0.39326200	2.25923300
C	-2.66858000	-1.31272700	-1.59864800
H	-2.25512500	-2.30620300	-1.77101800
H	-3.72794900	-1.32396300	-1.85766200
H	-2.15473400	-0.60880100	-2.25500200
C	-3.18908000	-1.92230800	0.76517300
H	-4.26932000	-1.80700300	0.65313700
H	-2.92275200	-2.94183300	0.48644100
H	-2.92584900	-1.77313200	1.81022700
C	3.72473400	1.13367600	-0.07973600
H	4.78628500	1.34052600	-0.08896200
Cl	4.44608400	-1.45881500	0.13221600

1-fluoro-3-(trifluoromethoxy)benzene (surrogate for **2az**)

C	2.29992300	-0.58575100	-0.00168700
C	1.90260900	1.75763700	0.11101800
H	2.25840700	2.76251100	0.29396200
C	0.99062000	-0.83162600	-0.36672200
H	0.64026400	-1.83651800	-0.55438300
C	0.14881800	0.25950900	-0.48384000
C	0.57966700	1.55347500	-0.25410700

H	-0.11577800	2.37340400	-0.36337300
O	-1.16338500	0.04454200	-0.90854300
C	-2.08376300	-0.15225500	0.05780600
F	-1.80363800	-1.21402000	0.81493200
F	-2.18860400	0.89469900	0.87718300
F	-3.25106100	-0.34761900	-0.52555800
F	3.13509800	-1.62383500	0.11669500
C	2.77654200	0.68835300	0.24113800
H	3.81152400	0.82518900	0.52123900

1-chloro-3-(trifluoromethoxy)benzene (surrogate for **2ba**)

C	2.05109500	-0.09023100	-0.04098500
C	1.26958300	2.16470500	0.12505700
H	1.46203500	3.20947500	0.32906400
C	0.78622400	-0.52438900	-0.40337700
H	0.58560000	-1.56513000	-0.61461400
C	-0.22329500	0.41641300	-0.49142900
C	-0.00614700	1.75814000	-0.23725300
H	-0.82426300	2.45897100	-0.32606700
Cl	3.33511000	-1.24361200	0.07676700
O	-1.48688600	-0.00426600	-0.91025500
C	-2.36183600	-0.33690600	0.06174700
F	-1.91557500	-1.33692300	0.82199800
F	-2.62561500	0.68607200	0.87614800
F	-3.48615900	-0.71502800	-0.51546900
C	2.30436500	1.24681200	0.22603600
H	3.30105200	1.55786100	0.50574600

3-fluorophenyl acetate, conformation A (surrogate for **2bb**)

E(RM062X) = –484.12428891 a.u.

C	-1.79479100	-0.44897500	-0.00000100
C	-1.11166300	1.83232500	-0.00012600
H	-1.35493500	2.88653900	-0.00014800
C	-0.48809200	-0.88092700	0.00004700
H	-0.24794400	-1.93560300	0.00005800
C	0.52967800	0.06950400	0.00001600
C	0.21606700	1.42544300	-0.00015800
H	1.00131700	2.16917200	-0.00028600
F	-2.78115500	-1.35532900	0.00016500
C	1.94677700	-0.42829200	0.00001300
C	3.06431100	0.58459900	0.00030700
H	2.99912600	1.22583000	-0.88037800
H	2.99901400	1.22547600	0.88127400
H	4.01339800	0.05652600	0.00025000
O	2.17616000	-1.61223200	-0.00034100
C	-2.13138600	0.89408600	-0.00002200
H	-3.17426600	1.18130600	0.00001800

3-fluorophenyl acetate, conformation B (surrogate for **2bb**)

E(RM062X) = –484.12409276 a.u.

C	-1.72605100	-0.54945300	-0.00002300
C	-1.27822900	1.78739300	0.00004000
H	-1.63086200	2.81026000	0.00006600
C	-0.37687100	-0.84421200	-0.00001400

H	-0.06831400	-1.88084800	-0.00003600
C	0.53654500	0.20578700	-0.00003500
C	0.08070500	1.52170600	-0.00005200
H	0.81047100	2.31983300	-0.00009700
F	-2.60543600	-1.56161700	0.00000600
C	2.01879600	-0.03775800	-0.00004400
C	2.51190400	-1.46344600	0.00003700
H	2.14603200	-1.99367100	0.88106300
H	2.14556900	-1.99388700	-0.88070500
H	3.59795400	-1.45810600	-0.00028100
O	2.78869800	0.89033600	0.00003600
C	-2.19934800	0.74726100	0.00002300
H	-3.26621300	0.92460900	0.00006200

3-chlorophenyl acetate, conformation A (surrogate for **2bc**)

E(RM062X) = –844.47427445 a.u.

C	-1.45775200	-0.01983900	-0.00001300
C	-0.49426000	2.16999400	0.00007700
H	-0.61632800	3.24510400	0.00019600
C	-0.19596300	-0.59252200	-0.00006800
H	-0.10676300	-1.67014400	-0.00002800
C	0.92658300	0.22928500	-0.00002600
C	0.77250700	1.61315000	0.00008900
H	1.66095100	2.23012000	0.00017700
C	2.31934200	-0.33378000	0.00002600
C	2.49013000	-1.83233300	0.00021500
H	2.01746600	-2.27097000	0.88073700

H	2.01812700	-2.27110500	-0.88058600
H	3.55149400	-2.06290500	0.00054700
O	3.27303600	0.40441600	-0.00028800
Cl	-2.85511600	-1.04419500	-0.00005500
C	-1.61998900	1.35577700	0.00005000
H	-2.61584000	1.77749700	0.00010200

3-chlorophenyl acetate, conformation B (surrogate for **2bc**)

E(RM062X) = –844.47411726 a.u.

C	1.48784000	0.00637700	-0.00000200
C	0.38724500	2.13138300	-0.00003300
H	0.43677500	3.21216000	-0.00005800
C	0.27034100	-0.64561800	-0.00002900
H	0.21224200	-1.72561600	-0.00008200
C	-0.90526700	0.10118400	-0.00001300
C	-0.84394300	1.49087400	0.00004300
H	-1.75107000	2.07996400	0.00012700
C	-2.20879800	-0.64560400	-0.00003100
C	-3.48975600	0.15131400	0.00005800
H	-3.53787600	0.79403100	0.88079900
H	-3.53802000	0.79397300	-0.88074500
H	-4.33045100	-0.53622700	0.00011300
O	-2.22163100	-1.85122300	-0.00004400
Cl	2.95349500	-0.91706900	0.00003700
C	1.56005500	1.39361200	-0.00004700
H	2.52572800	1.88053100	-0.00010300

1-ethynyl-3-fluorobenzene (surrogate for **2bd**)

C	1.39483400	-0.53239600	0.00001700
C	0.79559900	1.76861200	-0.00000500
H	1.07725600	2.81330100	0.00008000
C	0.07124500	-0.91833700	0.00001800
H	-0.18828900	-1.96772500	0.00004300
C	-0.91295200	0.07059700	0.00000100
C	-0.54445900	1.41741300	0.00001200
H	-1.31470800	2.17605900	0.00001400
C	-2.29655600	-0.29961100	-0.00002200
C	-3.45545600	-0.60691700	-0.00003100
H	-4.48347700	-0.87994100	0.00013600
F	2.34005100	-1.48239500	-0.00000400
C	1.78323600	0.79370600	-0.00001800
H	2.83581900	1.04146100	-0.00006500

1-ethynyl-3-chlorobenzene (surrogate for **2be**)

C	1.10728300	-0.05461100	0.00002500
C	-0.00497300	2.06076200	-0.00000800
H	0.03803300	3.14182900	0.00006400
C	-0.10716200	-0.71646000	0.00003300
H	-0.14125000	-1.79672700	0.00005700
C	-1.28935800	0.02526400	0.00001300
C	-1.23233300	1.41977400	0.00001100
H	-2.15276100	1.98713300	-0.00003200
Cl	2.57302600	-0.97841800	-0.00001000
C	-2.55222400	-0.64962100	-0.00001800

C	-3.60884100	-1.21578100	-0.00002900
H	-4.54566800	-1.71908200	0.00005600
C	1.17470800	1.33033300	-0.00001100
H	2.13761000	1.82200200	-0.00007600

1,2-difluorobenzene (surrogate for **2bf**, **2bm**)

C	-0.65624800	1.39311900	-0.00000100
C	0.53242600	0.69295300	-0.00001700
C	-0.65624300	-1.39311900	0.00000400
C	-1.85583500	-0.69361400	-0.00000600
C	-1.85583600	0.69361000	0.00000500
H	-0.62236100	2.47421300	0.00000500
H	-0.62235600	-2.47421200	0.00001000
H	-2.79014200	-1.23795900	-0.00000800
H	-2.79014600	1.23795100	0.00001500
C	0.53243000	-0.69294900	0.00000200
F	1.69893500	-1.34029200	0.00000100
F	1.69893500	1.34029200	0.00000500

1,4-difluorobenzene (surrogate for **2bg**)

C	1.36058200	0.00001700	-0.00000300
C	0.69372900	-1.20981700	0.00000300
C	-0.69372500	-1.20981600	-0.00000200
C	-1.36058200	0.00001200	-0.00000700
C	-0.69374400	1.20982000	0.00000400
C	0.69374100	1.20982100	0.00000100
H	1.25909900	-2.13137000	0.00002100

H	-1.25909500	-2.13137200	-0.00000400
H	-1.25906600	2.13140300	0.00000000
H	1.25906000	2.13140500	0.00000200
F	-2.70216000	-0.00001600	0.00000200
F	2.70216000	-0.00001500	-0.00000100

1-chloro-2-fluorobenzene (surrogate for **2bh**)

C	-0.40252000	-1.46042500	0.00000000
C	0.42411900	-0.34712800	-0.00000100
C	-0.14314300	0.91844400	-0.00001300
C	-1.51346300	1.08947600	-0.00000400
C	-2.33459200	-0.02811700	0.00000800
C	-1.77991200	-1.30024400	-0.00000200
H	0.04616500	-2.44440800	0.00000400
H	-1.91100700	2.09530400	-0.00000600
H	-3.40867600	0.09898500	0.00001500
H	-2.41816100	-2.17331300	-0.00000400
F	0.64833200	1.99072300	0.00000600
Cl	2.13845600	-0.51324200	0.00000100

1-fluoro-2-methylbenzene (surrogate for **2bi**)

C	-0.49330600	-1.44481900	0.00000200
C	0.61660200	-0.60699700	-0.00001000
C	0.36525500	0.75770000	-0.00001600
C	-0.90369000	1.29783200	-0.00000600
C	-1.99118300	0.43490400	0.00000500
C	-1.78572000	-0.93653700	0.00000300

H	-0.33442300	-2.51667100	0.00000100
H	-1.02305200	2.37268700	-0.00001000
H	-2.99530600	0.83763200	0.00001700
H	-2.63073600	-1.61182400	0.00000600
F	1.42055100	1.59192700	0.00001100
C	2.02783900	-1.11588800	0.00000000
H	2.57098900	-0.76079900	-0.87714100
H	2.57084900	-0.76117600	0.87738500
H	2.04192900	-2.20436400	-0.00022100

1-fluoro-3-methylbenzene (surrogate for **2bj**)

C	-1.19217400	-0.19555400	-0.00649600
C	-0.04323500	-0.98394900	-0.00524600
C	1.19469200	-0.37636400	0.00034500
C	1.34997600	0.99688400	0.00389100
C	0.20457300	1.77759200	0.00112900
C	-1.05324900	1.18857200	-0.00439600
H	-0.10189400	-2.06526800	-0.00911400
H	2.34258800	1.42505000	0.00689900
H	0.29472300	2.85607200	0.00188000
H	-1.93894900	1.81162500	-0.00785700
F	2.29161600	-1.15027500	0.00086200
C	-2.54833500	-0.84454900	0.00496700
H	-2.62810400	-1.60036600	-0.77741000
H	-2.72963500	-1.34122700	0.96010500
H	-3.33675600	-0.10919600	-0.14742800

1-fluoro-4-(trifluoromethyl)benzene (surrogate for **2bk**)

C	-0.35088500	0.01761500	-0.02176400
C	0.32924300	-1.19404300	-0.01511900
C	1.71281500	-1.21082100	-0.00350600
C	2.38587500	-0.00230700	0.00241100
C	1.73072800	1.21304700	-0.00336200
C	0.34407500	1.21728900	-0.01459400
H	-0.22477800	-2.12379600	-0.02429900
H	2.30408000	2.12947700	-0.00099600
H	-0.19535700	2.15427300	-0.02284100
C	-1.85173000	0.00455700	-0.00086300
F	-2.37206400	1.19539900	-0.31227100
F	-2.32758300	-0.32692400	1.20854800
F	-2.35201800	-0.88857000	-0.86447800
H	2.27397300	-2.13475700	-0.00134300
F	3.72292700	-0.01292900	0.01156400

1-chloro-4-fluorobenzene (surrogate for **2bl**)

C	-0.25982200	1.20689800	0.00000200
C	1.12696200	1.20912700	0.00000600
C	1.79453900	0.00001800	0.00000600
C	1.12692900	-1.20911700	0.00000400
C	-0.25981300	-1.20689200	0.00000500
C	-0.94180200	0.00002500	0.00000200
H	-0.81165700	2.13650300	-0.00001000
H	1.69025800	2.13204400	0.00000800
H	1.69028200	-2.13200100	0.00000000

H	-0.81170500	-2.13646300	-0.00000700
F	3.13429400	-0.00002900	-0.00001000
Cl	-2.67575200	-0.00001000	-0.00000300

(difluoromethyl)benzene (surrogate for **2bu**)

C	-1.73388100	1.27644900	0.21026600
C	-0.36152700	1.11649300	0.10428700
C	0.16793400	-0.14711000	-0.12734600
C	-0.67137300	-1.24426600	-0.25430500
C	-2.04534100	-1.08297000	-0.14513800
C	-2.57548000	0.17771400	0.08713000
H	-2.15019700	2.25975300	0.38650600
H	0.30310700	1.96611500	0.19173000
H	-0.25041700	-2.22627400	-0.43765600
H	-2.70053500	-1.93839200	-0.24418500
H	-3.64711200	0.30677800	0.16865200
C	1.65226200	-0.34412000	-0.19850700
H	1.93732600	-1.23159800	-0.76687700
F	2.26063500	0.73168900	-0.75935800
F	2.17406100	-0.47052500	1.05308100

7. References and footnotes

1. C. K. Chung, Z. Liu, K. W. Lexa, T. Andreani, Y. Xu, D. A. DiRocco, G. R. Humphrey and R. T. Ruck, *J. Am. Chem. Soc.* 2017, **139**, 10637.
2. H. Kim, Y. Nguyen, C. P.-H. Yen, L. Chagal, A. J. Lough, B. M. Kim and J. Chin, *J. Am. Chem. Soc.* 2008, **130**, 12184.
3. Typically, a small amount of EtOAc was added to dissolve the product. In this case, the EtOAc:Et_2O ratio should not exceed ca. 20:80 to avoid DMSO extraction.
4. If no precipitation occurs overnight, the solvent can be exchanged to Et_2O or MeCN.
5. S. Pikul and E. J. Corey, *Org. Synth.* 1993, **71**, 30.
6. 2,6-lutidine can be used to replace triethylamine as base.
7. C. K. Chung, G. R. Humphrey, Z. Liu, M. Mclaughlin, Y. Xu and Y. Yu, WO2017091453 A1, June 1, 2017
8. R. E. Carhart, D. H. Smith and R. Ventkataraghavan, *J. Chem. Inf. Comput. Sci.* 1985, **25**, 64.
9. R. Nilakantan, N. Bauman, J. S. Dixon and R. Venkataraghavan, *J. Chem. Inf. Comput. Sci.* 1987, **27**, 82.
10. D. Rogers and M. Hahn, *J. Chem. Inf. Model.* 2010, **50***,* 742.
11. Molecular Operating Environment (MOE), Version 2016, Chemical Computing Group, Montreal, Canada, 2016. http://www. chemcomp.com/
12. V. Svetnik, A. Liaw, A. C. Tong, J. C. Culberson, R. P. Sheridan and B. P. Feuston, *J. Chem. Inf. Comput. Sci.* 2003, **43**, 1947.
13. C. Cortes and V. N. Vapnik, *Machine Learning* 1995, **20**, 273.
14. Y. Zhao and D. G. Truhlar, *Acc. Chem. Res*. 2008, **41**, 157.
15. F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* 2005, **7**, 3297.
16. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian 09, revision D. 01. Gaussian, Inc., Wallingford CT, 2009.