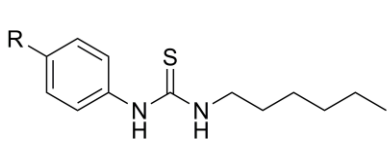
Supplementary Information:

Towards Predictable Transmembrane Transport: QSAR Analysis of the Anion Binding and Anion Transport Properties of Thioureas

Nathalie Busschaert, Samuel J. Bradberry, Marco Wenzel, Cally J. E. Haynes, Jennifer R. Hiscock, Isabelle L. Kirby, Louise Karagiannidis, Stephen J. Moore, Neil J. Wells, Julie Herniman, G. John Langley, Peter N. Horton, Mark E. Light, Igor Marques, Paulo J. Costa, Vítor Félix, Jeremy G. Frey, Philip A. Gale

Chemistry, University of Southampton, Southampton, SO17 1BJ, UK and Departamento de Química, CICECO and Secção Autónoma de Ciências da Saúde, Universidade de Aveiro, 3810-193 Aveiro, Portugal. Corresponding author: philip.gale@soton.ac.uk

S1. OVERVIEW OF COMPOUNDS



1	R = Br	**7**	R = COOMe	**13**	R = OCF_3	**19**	R = CH_2CH_3
2	R = CF_3	**8**	R = F	**14**	R = OEt	**20**	R = $(CH_2)_2CH_3$
3	R = Cl	**9**	R = H	**15**	R = OMe	**21**	R = $(CH_2)_3CH_3$
4	R = CN	**10**	R = I	**16**	R = SMe	**22**	R = $(CH_2)_4CH_3$
5	R = $COCF_3$	**11**	R = NO_2	**17**	R = SO_2Me		
6	R = COMe	**12**	R = O(CO)Me	**18**	R = CH_3		

S2. SYNTHESIS

General. ^{1}H NMR (300 MHz), ^{19}F NMR (282 MHz) and ^{13}C{^{1}H} NMR (75 MHz) spectra were determined on a Bruker AV300 spectrometer. ^{1}H NMR (400 MHz) and ^{13}C{^{1}H} NMR (100 MHz) spectra were determined on a Bruker DPX400 spectrometer. Chemical shifts (δ) are reported in parts per million (ppm) and calibrated to the residual protio solvent peak in DMSO-d_6 (δ = 2.50 (^{1}H) and 39.51 ppm (^{13}C)). The following abbreviations are used for spin multiplicity: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quartet, m = multiplet, br = broad. Infrared (IR) spectra were recorded on a Matterson Satellite (ATR) and are reported in wavenumbers (cm^{-1}). High resolution electron spray (ES) mass spectra were recorded on a Bruker maXis ESI. All mass spectra are reported as *m/z* (relative intensity). Melting points were determined by a Barnstead Electrothermal 9100 melting point apparatus and were not corrected. All reactions were performed using oven-dried glassware. Dichloromethane was distilled over calcium hydride under nitrogen prior to use and pyridine was dried with KOH prior to use. All other solvents and reagents were used as provided by the supplier. The synthesis and characterization of compound **2**[1] and **9**[2] were performed as previously reported. The majority of the compounds was obtained as a crystalline solid and further characterized by single crystal X-ray diffraction (see section S4).

3-hexyl-1-(4-bromo-phenyl)thiourea (1). A solution of 4-bromoaniline (0.41 g, 2.38 mmol) in 3 mL pyridine was treated dropwise with a solution of hexyl isothiocyanate (0.39 mL, 2.5 mmol) in 3 mL pyridine under an inert atmosphere. The reaction was stirred overnight at room temperature and concentrated *in vacuo* to give a clear, viscous oil. Trituration under hexane afforded an off-white solid that was isolated via filtration and washed with excess hexane. The resulting white solid was subsequently dried under high vacuum (0.53 g, 1.67 mmol). Yield: 85%; Mp: 102-104 °C (lit. 189 °C)[3]; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.2 Hz, 3 H), 1.28 (m, 6 H), 1.52 (m, 2 H), 3.44 (q, *J*=4.4 Hz, 2 H), 7.44 (m, 4 H), 7.81 (br. s, 1 H), 9.49 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.1, 26.1, 28.3, 31.0, 43.8, 115.6, 124.6, 131.2, 138.9, 180.2; IR (solid): ν= 3240, 3060, 2930, 2870, 1550, 1480, 1310, 1070, 1010, 825 cm^{-1}; LRMS (ESI+): *m/z*= 315.3, 317.3 [M+H]$^{+}$, 337.3, 339.4 [M+Na]$^{+}$; HRMS (ES) for $C_{13}H_{19}BrSN_2Na$

$[M+Na]^+$: *m*/*z*= 337.0345 (calcd), 337.0341 (found). The structure of compound **1** was confirmed by single crystal X-ray diffraction (see section S4.1).

1-(4-chlorophenyl)-3-hexylthiourea (3). 4-Chloroaniline (0.50 g, 3.92 mmol) and hexyl isothiocyanate (0.6 mL, 3.91 mmol) were dissolved in 10 mL dry pyridine. The mixture was stirred overnight at 80°C under nitrogen atmosphere. The solvents were removed under reduced pressure and the residue was recrystallized from chloroform:hexane. A white crystalline solid was formed, filtered off and washed with hexane. The white solid was dried overnight *in vacuo* (0.98 g, 3.62 mmol). Yield: 92%; Mp: 91-92 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=7.3 Hz, 3 H), 1.28 (m, 6 H), 1.52 (m, 2 H), 3.44 (q, *J*=4.8 Hz, 2 H), 7.34 (d, *J*=8.4 Hz, 2 H), 7.47 (d, *J*=9.1 Hz, 2 H), 7.80 (br. s, 1 H), 9.48 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.1, 26.1, 28.4, 31.0, 43.8, 124.4, 127.6, 128.3, 138.5, 180.3; IR (solid): ν= 3250, 3060, 2930, 2850, 1540, 1490, 1090, 825 cm^{-1}; LRMS (ESI-): *m*/*z*= 269.1 $[M-H]^-$; HRMS (ES) for $C_{13}H_{20}SN_2Cl$ $[M+H]^+$: *m*/*z*= 271.1030 (calcd), 271.1032 (found). The structure of compound **3** was confirmed by single crystal X-ray diffraction (see section S4.3).

1-(4-cyanophenyl)-3-hexylthiourea (4). Hexylamine (0.22 mL, 1.66 mmol) was dissolved in 8 mL dry dichloromethane under nitrogen. 4-Cyanophenyl isothiocyanate (0.27 g, 1.68 mmol) in 8 mL dichloromethane was added dropwise to the reaction mixture over the course of 30 minutes. The reaction mixture was refluxed overnight under nitrogen atmosphere (40 °C). The reaction was quenched by the addition of 10 mL water, the organic phase was washed with water (2 x 10 mL) and dried over $MgSO_4$. $MgSO_4$ was then removed by filtration, the solvents were removed under reduced pressure and the residue was recrystallized from dichloromethane:hexane. A white solid was formed, filtered off and dried on air (0.37 g, 1.42 mmol). Yield: 85%; Mp: 104-105 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.2 Hz, 3 H), 1.28 (m, 6 H), 1.54 (m, 2 H), 3.46 (q, *J*=5.5 Hz, 2 H), 7.73 (m, 4 H), 8.14 (br. s, 1 H), 9.85 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.1, 28.1, 30.9, 43.8, 104.4, 119.1, 121.0, 132.7, 144.3, 179.9; IR (solid): ν= 3340, 3175, 2930, 2850, 2230, 1525, 1310, 849 cm^{-1}; LRMS (ESI+): *m*/*z*= 262.4 $[M+H]^+$, 284.3 $[M+Na]^+$; HRMS (ES) for $C_{14}H_{19}SN_3Na$ $[M+Na]^+$: *m*/*z*= 284.1192 (calcd), 284.1193 (found). The structure of compound **4** was confirmed by single crystal X-ray diffraction (see section S4.4).

1-hexyl-3-(4-(2,2,2-trifluoroacetyl)phenyl)thiourea (5). A solution of 1-(4-aminophenyl)-2,2,2-trifluoroethanone (0.50 g, 2.64 mmol) and hexyl isothiocyanate (0.41 mL, 2.64 mmol) in 10 mL pyridine was heated to 70 °C overnight under nitrogen atmosphere. The mixture was then concentrated *in vacuo* to give a clear, viscous oil. The crude oil was dissolved in methanol and purified using an ion exchange column for bases (Isolute SCX-2, Biotage). The methanol fraction was allowed to evaporate to afford a white crystalline solid which was subsequently dried under high vacuum (0.23 g, 0.69 mmol). Yield: 26%; Mp: 88 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.8 Hz, 3 H), 1.29 (m, 6 H), 1.56 (m, 2 H), 3.48 (q, *J*=6.0 Hz, 2 H), 7.85 (d, *J*=9.0 Hz, 2 H), 7.99 (d, *J*=8.7 Hz, 2 H), 8.32 (br. s, 1 H), 10.09 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.1, 26.1, 28.0, 31.0, 43.9, 116.7 (q, $J_{C\text{-}F}$=291.9 Hz), 120.0, 122.7,

131.1, 147.4, 177.9 (q, $J_{C\text{-}F}$=33.7 Hz), 179.8; ^{19}F NMR (282 MHz, DMSO-d_6) δ ppm -69.82; IR (solid): ν= 3240, 3060, 2930, 2850, 1700, 1600, 1530, 1140, 941, 845 cm^{-1}; LRMS (ESI+): *m/z*= 333.4 [M+H]$^+$, 355.4 [M+Na]$^+$; HRMS (ES) for $C_{15}H_{19}F_3N_2OSNa$ [M+Na]$^+$: *m/z*= 355.1062 (calcd), 355.1058 (found). The structure of compound **5** was confirmed by single crystal X-ray diffraction (see section S4.5).

1-(4-acetylphenyl)-3-hexylthiourea (6). 4-Aminoacetophenone (0.3 g, 2.22 mmol) was dissolved in 10 mL dry pyridine. Hexyl isothiocyanate (0.34 mL, 2.22 mmol) was added to this mixture under nitrogen atmosphere. The reaction was stirred overnight under nitrogen atmosphere at 80°C. Pyridine was removed under reduced pressure and the crude oil was recrystallized from 1:1 dichloromethane:hexane to give a white solid. This solid was subsequently dissolved in a minimum amount of hot methanol and recrystallized from a two phase solution of methanol and hexane in equal amounts. The resulting off-white crystalline solid was subsequently dried under high vacuum (0.13 g, 0.46 mmol). Yield: 21%; Mp: 102 °C (lit. 118-119 °C)[4]; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.85 (t, *J*=6.2 Hz, 3 H), 1.27 (m, 7 H), 1.52 (m, 2 H), 3.44 (br q, *J*=4.4 Hz, 2 H), 7.64 (d , *J*=8.4 Hz, 2 H), 7.87 (d, *J*=8.4 Hz, 2 H), 8.02 (br. s, 1 H), 9.75 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.1, 26.1, 26.5, 28.2, 31.0, 43.8, 120.5, 129.0, 131.5, 144.3, 180.0, 196.5; IR (solid): ν= 3290, 2920, 2850, 1660, 1600, 1540, 957, 825 cm^{-1}; LRMS (ESI+): *m/z*= 279.3 [M+H]$^+$, 301.3 [M+Na]$^+$; HRMS (ES) for $C_{15}H_{22}SON_2Na$ [M+Na]$^+$: *m/z*= 301.1345 (calcd), 301.1345 (found). The structure of compound **6** was confirmed by single crystal X-ray diffraction (see section S4.6).

Methyl 4-(3-hexylthioureido)benzoate (7). Methyl-4-aminobenzoate (0.60 g, 3.97 mmol) was dissolved in 30 mL dichloromethane and thiophosgene (0.4 mL, 5.25 mmol) was added, followed by 30 mL of a saturated aqueous $NaHCO_3$ solution. The mixture was stirred gently for 12 hours at room temperature. The aqueous layer was removed and the dichloromethane phase was washed 20 mL saturated $NaHCO_3$ solution. The dichloromethane layer was subsequently dried with $MgSO_4$ for 1 hour, then $MgSO_4$ was removed by filtration and hexylamine (0.55 mL, 4.16 mmol) was added to the dried reaction mixture. The mixture was refluxed overnight under nitrogen atmosphere (40 °C). The solvents were removed under reduced pressure and the residue was recrystallized from chloroform:hexane. A white crystalline solid was formed, filtered off and washed with hexane. The white solid was then dried overnight *in vacuo* (0.64 g, 2.17 mmol). Yield: 55%; Mp: 127-129 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.4 Hz, 3 H), 1.28 (m, 6 H), 1.55 (m, 2 H), 3.46 (q, *J*=5.5 Hz, 2 H), 3.82 (s, 3 H), 7.67 (d, *J*=8.4 Hz, 2 H), 7.88 (d, *J*=8.8 Hz, 2 H), 8.05 (br. s, 1 H), 9.79 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.1, 28.2, 31.0, 43.8, 51.8, 120.6, 123.6, 129.8, 144.4, 165.8, 180.0; IR (solid): ν= 3240, 3060, 2920, 2860, 1715, 1600, 1550, 1270, 1110 cm^{-1}; LRMS (ESI-): *m/z*= 293.2 [M-H]$^-$; HRMS (ES) for $C_{15}H_{23}SN_2O_2$ [M+H]$^+$: *m/z*= 295.1475 (calcd), 295.1479 (found). HRMS (ES) for $C_{15}H_{22}SN_2O_2Na$ [M+Na]$^+$: *m/z*= 317.1294 (calcd), 317.1299 (found). The structure of compound **7** was confirmed by single crystal X-ray diffraction (see section S4.7).

3-hexyl-1-(4-fluoro-phenyl)thiourea (8). A solution of hexylamine (0.30 g, 2.96 mmol) in dichloromethane (5 mL) was treated dropwise with a solution of 4-fluorophenyl isothiocyanate (0.48 g, 3.13

mmol) in dichloromethane (5mL) under an inert atmosphere. The reaction was stirred overnight at room temperature. The solvents were removed under reduced pressure to yield a viscous yellow oil. Trituration under excess hexane (50 mL) afforded a white solid that was isolated via filtration and washed with hexane. The resulting solid was dried under high vacuum (0.57 g, 2.24 mmol). Yield: 76%; Mp: 77-78 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.6 Hz, 3 H), 1.28 (m, 6 H), 1.52 (m, 2 H), 3.43 (m, 2 H), 7.14 (t, *J*=8.6 Hz, 2 H), 7.39 (dd, *J*=8.6, 5.1 Hz, 2 H), 7.69 (br. s, 1 H), 9.37 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.0, 28.4, 31.0, 43.8, 115.0 (d, *J*=22.0 Hz), 125.5, 135.6, 158.8 (d, *J*=241.5 Hz), 180.6; ^{19}F NMR (282 MHz, DMSO-d_6) δ ppm -118.24; IR (solid): ν= 3210, 2940, 2870, 1560, 1500, 1320, 1110, 1060, 837 cm^{-1}; LRMS (ESI-): *m*/*z*= 253.2 [M-H]$^{-}$; HRMS (ES) for $C_{13}H_{20}SN_2F$ [M+H]$^{+}$: *m*/*z*= 255.1326 (calcd), 255.1323 (found); HRMS (ES) for $C_{13}H_{19}SFN_2Na$ [M+Na]$^{+}$: *m*/*z*= 277.1145 (calcd), 277.1143 (found). The structure of compound **8** was confirmed by single crystal X-ray diffraction (see section S4.8).

1-hexyl-3-(4-iodophenyl)thiourea (10). A solution of 4-(iodophenyl) isothiocyanate (0.30 g, 1.15 mmol) and hexylamine (0.15 mL, 1.15 mmol) in 2 mL dichloromethane was shaken in a sealed vial for 10 seconds and left at room temperature overnight. Hexane was subsequently added to the solution until a white precipitate formed. The precipitate was collected, washed with hexane (10 mL) and dried under high vacuum (0.41 g, 1.14 mmol). Yield: 99%; Mp: 89 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.6 Hz, 3 H), 1.28 (m, 6 H), 1.53 (m, 2 H), 3.43 (br q, *J*=4.5 Hz, 2 H), 7.28 (d, *J*=8.6 Hz, 2 H), 7.62 (d, *J*=8.6 Hz, 2 H), 7.81 (br. s, 1 H), 9.47 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.0, 28.3, 31.0, 43.8, 87.6, 124.8, 137.1, 139.4, 180.1; IR (solid): ν= 3260, 3050, 2920, 2850, 1620, 1540, 1480, 1320, 822 cm^{-1}; LRMS (ESI+): *m*/*z*= 363.3 [M+H]$^{+}$, 385.3 [M+Na]$^{+}$; HRMS (ES) for $C_{13}H_{19}SIN_2Na$ [M+Na]$^{+}$: *m*/*z*= 385.0206 (calcd), 385.0210 (found). The structure of compound **10** was confirmed by single crystal X-ray diffraction (see section S4.9).

1-hexyl-3-(4-nitrophenyl)thiourea[5] **(11).** 4-Nitroaniline (0.72 g, 5.21 mmol) was dissolved in 40 mL dichloromethane and thiophosgene (0.4 mL, 5.25 mmol) was added, followed by 30 mL of a saturated aqueous $NaHCO_3$ solution. The mixture was stirred gently for 12 hours at room temperature. The aqueous layer was removed and the dichloromethane phase was subsequently dried with $MgSO_4$ for 1 hour, then $MgSO_4$ was removed by filtration and hexylamine (0.69 mL, 5.22 mmol) was added to the dried reaction mixture. The mixture was stirred for 3 days under nitrogen atmosphere. The solvents were removed under reduced pressure and the residue was recrystallized from dichloromethane:hexane. The product was further purified via an ion exchange column for bases (Isolute SCX-2, Biotage) using methanol as eluent. After removing the solvents under reduced pressure a yellow solid was obtained and dried overnight *in vacuo* (0.74 g, 2.50 mmol). Yield: 48%; Mp: 103-104 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.3 Hz, 3 H), 1.30 (m, 6 H), 1.56 (m, 2 H), 3.48 (q, *J*=6.1 Hz, 2 H), 7.83 (d, *J*=9.1 Hz, 2 H), 8.18 (d, *J*=9.1 Hz, 2 H), 8.27 (br. s, 1 H), 10.07 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.1, 28.0, 31.0, 43.9, 120.2, 124.5, 141.6, 146.5, 179.9; IR (solid): ν= 3330, 2920, 1590, 1525, 1495, 850 cm^{-1}; LRMS (ESI-):

m/z= 280.2 $[M-H]^-$; HRMS (ES) for $C_{13}H_{19}SN_3O_2Na$ $[M+Na]^+$: *m/z*= 304.1090 (calcd), 304.1088 (found). The structure of compound **11** was confirmed by single crystal X-ray diffraction (see section S4.10).

4-(3-hexylthioureido)phenyl acetate (12). A suspension of 10 % Pd/C (0.03 g) in a 4-nitrophenyl acetate (0.30 g, 1.66 mmol) and ethanol solution (10 mL) was vigorously stirred under a hydrogen atmosphere for 3 hours. The mixture was subsequently filtered through celite and concentrated *in vacuo* to afford a white crystalline solid. This solid was dissolved in pyridine (8 mL), hexyl isothiocyanate (0.25 mL, 1.66 mmol) was added and the solution was heated to 70 °C overnight under inert atmosphere. The solvents were removed under reduced pressure to yield a viscous oil that was subsequently dissolved in methanol and purified via an ion exchange column for bases (Isolute SCX-2, Biotage). The methanol fraction was taken to dryness and the product was recrystallized from ethyl acetate (2 mL), hexane (3 mL) and diethyl ether (10 mL) to give a white solid which was dried under high vacuum (0.32 g, 1.09 mmol). Yield: 65%; Mp: 72 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.6 Hz, 3 H), 1.29 (m, 6 H), 1.53 (m, 2 H), 2.26 (s, 3 H), 3.44 (br q, *J*=5.5 Hz, 2 H), 7.05 (d, *J*=8.8 Hz, 2 H), 7.41 (d, *J*=8.8 Hz, 2 H), 7.75 (br. s, 1 H), 9.42 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 20.8, 22.0, 26.1, 28.4, 31.0, 43.8, 121.7, 124.0, 136.9, 146.6, 169.3, 180.4; IR (solid): ν= 3340, 2920, 1720, 1500, 1240, 845 cm^{-1}; LRMS (ESI+): *m/z*= 295.3 $[M+H]^+$, 317.3 $[M+Na]^+$; HRMS (ES) for $C_{15}H_{22}O_2SN_2Na$ $[M+Na]^+$: *m/z*= 317.1294 (calcd), 317.1298 (found). The structure of compound **12** was confirmed by single crystal X-ray diffraction (see section S4.11).

1-hexyl-3-(4-(trifluoromethoxy)phenyl)thiourea (13). p-(Trifluoro-methoxy)phenyl isothiocyanate (222 µL, 1.37 mmol) and hexylamine (199 µL, 1.51 mmol) were dissolved in 10 mL dry dichloromethane and the reaction mixture was stirred overnight at room temperature under nitrogen atmosphere. The solvents were removed under reduced pressure and the crude mixture was recrystallized from 1:1 chloroform:hexane. A white crystalline solid was obtained which was filtered off, washed with hexane and dried under high vacuum (0.35 g, 1.09 mmol). Yield: 80 %; Mp: 76-78 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.8 Hz, 3 H), 1.29 (m, 6 H), 1.53 (m, 2 H), 3.44 (br. q, *J*=6.4 Hz, 2 H), 7.29 (d, *J*=8.7 Hz, 2 H), 7.54 (d, *J*=8.7 Hz, 2 H), 7.86 (br. s., 1 H), 9.54 (br. s., 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 14.0, 22.1, 26.2, 28.4, 31.1, 43.9, 120.18 (q, $J_{C\text{-}F}$=256.5 Hz), 121.3, 124.1, 138.8, 144.1, 180.4; ^{19}F NMR (282 MHz, DMSO-d_6) δ ppm -56.66; IR (solid): ν= 3270, 3080, 2930, 2870, 1560, 1500, 1210, 1160, 849 cm^{-1}; LRMS (ESI+): *m/z*= 321.3 $[M+H]^+$, 343.4 $[M+Na]^+$; HRMS (ES) for $C_{14}H_{19}F_3OSN_2Na$ $[M+Na]^+$: *m/z*= 343.1062 (calcd), 343.1068 (found).

1-(4-ethoxyphenyl)-3-hexylthiourea (14). 4-Ethoxyphenyl isothiocyanate (0.29 g, 1.62 mmol) and hexylamine (0.25 mL, 1.89 mmol) were dissolved in 5 mL dichloromethane and stirred overnight at room temperature under nitrogen atmosphere. The solvents were reduced under reduced pressure and the crude mixture was recrystallized from chloroform hexane to yield a wite solid, which was filtered off and washed with hexane. This white solid was then dissolved in methanol and purified via an ion exchange column for bases (Isolute SCX-2, Biotage). The methanol fraction was taken to dryness and the product was

recrystallized from diethyl ether (10 mL). The resulting white crystalline solid was subsequently dried under high vacuum (0.41 g, 1.46 mmol). Yield: 90%; Mp: 61 °C (lit. 73-74 °C)[6]; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.6 Hz, 3 H), 1.28 (m, 9 H), 1.50 (m, 2 H), 3.41 (br q, *J*=6.1 Hz, 2 H), 4.00 (q, *J*=6.7 Hz, 2 H), 6.87 (d, *J*=8.6 Hz, 2 H), 7.20 (d, *J*=9.1 Hz, 2 H), 7.46 (br. s, 1 H), 9.19 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 14.7, 22.0, 26.0, 28.5, 31.0, 43.9, 63.1, 114.4, 125.7, 131.7, 155.7, 180.6; IR (solid): ν= 3210, 3040, 2920, 1540, 1250 cm^{-1}; LRMS (ESI+): *m*/*z*= 281.3 [M+H]$^{+}$, 303.3 [M+Na]$^{+}$; HRMS (ES) for $C_{15}H_{24}SON_2Na$ [M+Na]$^{+}$: *m*/*z*= 303.1502 (calcd), 303.1502 (found). The structure of compound **14** was confirmed by single crystal X-ray diffraction (see section S4.12).

1-Hexyl-3-(4-methoxyphenyl)thiourea (15). Hexylamine (0.40 mL, 3.03 mmol) was dissolved in 20 mL dry dichloromethane and 4-methoxyphenyl isothiocyanate (0.42 mL, 3.03 mmol) was added dropwise. The mixture was stirred overnight at room temperature under nitrogen atmosphere. The solvents were removed under reduced pressure and the residue was recrystallized from chloroform:hexane. A white crystalline solid was formed, filtered off and washed with hexane. The white solid was then dried overnight *in vacuo* (0.67 g, 2.52 mmol). Yield: 83%; Mp: 59-60 °C (lit. 81-82 °C)[6]; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=7.0 Hz, 3 H), 1.27 (m, 6 H), 1.50 (m, 2 H), 3.42 (q, *J*=5.9 Hz, 2 H), 3.74 (s, 3 H), 6.89 (d, *J*=8.8 Hz, 2 H), 7.22 (d, *J*=8.8 Hz, 2 H), 7.48 (br. s, 1 H), 9.20 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.0, 28.5, 31.0, 43.8, 55.2, 113.8, 125.7, 131.8, 156.4, 180.5; IR (solid): ν= 3240, 3070, 2920, 1560 1510, 1230, 1030 cm^{-1}; LRMS (ESI-): *m*/*z*= 265.2 [M-H]$^{-}$; HRMS (ES) for $C_{14}H_{23}SN_2O$ [M+H]$^{+}$: *m*/*z*= 267.1526 (calcd), 267.1528 (found), HRMS (ES) for $C_{14}H_{22}SN_2ONa$ [M+Na]$^{+}$: *m*/*z*= 289.1345 (calcd), 289.1347 (found). The structure of compound **15** was confirmed by single crystal X-ray diffraction (see section S4.13).

1-hexyl-3-(4-(methylthio)phenyl)thiourea (16). A solution of 4-(methylthio)phenyl isothiocyanate (0.30 g, 1.65 mmol) and hexylamine (0.22 mL, 1.65 mmol) in 2 mL dichloromethane was shaken in a sealed vial for 10 seconds. The solution was left at room temperature overnight. Hexane was then added to the solution until a white precipitate had formed. This was collected and washed with hexane (10 mL) and subsequently dried under high vacuum (0.37 g, 1.32 mmol). Yield: 80%; Mp: 82 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=7.0 Hz, 3 H), 1.27 (m, 6 H), 1.52 (m, 2 H), 2.46 (s, 3 H), 3.43 (br q, J=5.1 Hz, 2 H), 7.22 (d, *J*=8.6 Hz, 2 H), 7.35 (d, *J*=8.6 Hz, 2 H), 7.69 (br. s, 1 H), 9.39 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 15.4, 22.0, 26.1, 28.4, 31.0, 43.8, 123.8, 126.7, 133.0, 136.6, 180.3; IR (solid): ν= 3240, 3040, 2920, 2850, 1540, 1310, 818 cm^{-1}; LRMS (ESI+): *m*/*z*= 283.3 [M+H]$^{+}$, 305.3 [M+Na]$^{+}$; HRMS (ES) for $C_{14}H_{22}S_2N_2Na$ [M+Na]$^{+}$: *m*/*z*= 305.1117 (calcd), 305.1119 (found).

1-hexyl-3-(4-(methylsulfonyl)phenyl)thiourea (17). A solution of 4-(methylsulfonyl)aniline (0.30 g, 1.75 mmol) and hexyl isothiocyanate (0.27 mL, 1.75 mmol) in 8 mL pyridine was heated to 70 °C overnight under inert atmosphere, then concentrated *in vacuo* to give a clear, viscous oil. The oil was dissolved in dichloromethane (2 mL) and a white precipitate was collected upon the addition of hexane. This white solid was then dissolved in methanol and purified via an ion exchange column for bases (Isolute SCX-2, Biotage).

The methanol fraction was reduced to a volume of 5 mL and slow evaporation at room temperature resulted in a white crystalline product that was dried under high vacuum (0.18 g, 0.57 mmol). Yield: 33%; Mp: 120 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.88 (t, *J*=6.4 Hz, 3 H), 1.29 (m, 6 H), 1.56 (m, 2 H), 3.17 (s, 3H), 3.47 (br q, *J*=5.1 Hz, 2 H), 7.76 (d, *J*=8.8 Hz, 2 H), 7.82 (d, *J*=9.1 Hz, 2 H), 8.12 (br. s, 1 H), 9.86 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 22.0, 26.1, 28.1, 31.0, 43.8, 121.0, 127.7, 134.3, 144.5, 180.1 (one peak underneath DMSO peak); IR (solid): ν= 3230, 3040, 2930, 2850, 1600, 1540, 1290, 1140, 968 cm^{-1}; LRMS (ESI+): *m*/*z*= 315.3 [M+H]$^{+}$, 337.4 [M+Na]$^{+}$; HRMS (ES) for $C_{14}H_{22}S_2O_2N_2Na$ [M+Na]$^{+}$: *m*/*z*= 337.1015 (calcd), 337.1018 (found).

1-Hexyl-3-(*p*-tolyl)thiourea (18). *p*-Toluidine (0.50 g, 4.67 mmol) was dissolved in 10 mL dry pyridine and hexyl isothiocyanate (0.70 mL, 4.56 mmol) was added dropwise. The mixture was stirred overnight at 80°C under nitrogen atmosphere. The solvents were removed under reduced pressure and the residue was recrystallized from chloroform:hexane. A white crystalline solid was formed, filtered off and washed with hexane. The white solid was then dried overnight *in vacuo* (1.06 g, 4.23 mmol). Yield: 93%; Mp: 70-72 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.6 Hz, 3 H), 1.27 (m, 6 H), 1.51 (m, 2 H), 2.27 (s, 3 H), 3.43 (br q, *J*=5.5 Hz, 2 H), 7.11 (d, *J*=8.0 Hz, 2 H), 7.24 (d, *J*=8.0 Hz, 2 H), 7.59 (br. s, 1 H), 9.32 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9, 20.4, 22.0, 26.1, 28.4, 31.0, 43.8, 123.4, 129.0, 133.3, 136.6, 180.2; IR (solid): ν= 3240, 3070, 2930, 2850, 1550, 1510, 1320, 818 cm^{-1}; LRMS (ESI-): *m*/*z*= 249.0 [M-H]$^{-}$; HRMS (ES) for $C_{14}H_{23}SN_2$ [M+H]$^{+}$: *m*/*z*= 251.1576 (calcd), 251.1578 (found). The structure of compound **18** was confirmed by single crystal X-ray diffraction (see section S4.14).

1-(4-ethylphenyl)-3-hexylthiourea (19). A solution of 4-(ethylphenyl) isothiocyanate (0.28 g, 1.84 mmol) and hexylamine (0.24 mL, 1.84 mmol) in 2 mL dichloromethane was shaken in a sealed vial for 10 seconds and then left at room temperature overnight. Hexane (5 mL) was added and the solution was left to evaporate giving off white crystals which were collected, washed with hexane (10 mL) and recrystallized from diethyl ether (10 mL). This afforded white crystals which were collected, washed with hexane (10 mL) and subsequently dried under high vacuum (0.40 g, 1.53 mmol). Yield: 82%; Mp: 64 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 0.87 (t, *J*=6.6 Hz, 3 H), 1.17 (t, *J*=7.6 Hz, 3 H), 1.28 (m, 6 H), 1.52 (m, 2 H), 2.57 (q, *J*=7.6 Hz, 2 H), 3.43 (br q, *J*=5.6 Hz, 2 H), 7.15 (d, *J*=8.1 Hz, 2 H), 7.26 (d, *J*=8.1 Hz, 2 H), 7.60 (br. s, 1 H), 9.31 (br. s, 1 H); ^{13}C{^{1}H} NMR (100 MHz, DMSO-d_6) δ ppm 13.9, 15.6, 22.0, 26.1, 27.6, 28.4, 31.0, 43.9, 123.4, 127.8, 136.8, 139.7, 180.3; IR (solid): ν= 3210, 2930, 2860, 1550, 1500, 1270, 829 cm^{-1}; LRMS (ESI+): *m*/*z*= 265.3 [M+H]$^{+}$, 287.3 [M+Na]$^{+}$; HRMS (ES) for $C_{15}H_{24}SN_2Na$ [M+Na]$^{+}$: *m*/*z*= 287.1552 (calcd), 287.1555 (found). The structure of compound **19** was confirmed by single crystal X-ray diffraction (see section S4.15).

1-(4-propylphenyl)-3-hexylthiourea (20). A solution of 4-propylaniline (0.32 mL, 2.22 mmol) and hexyl isothiocyanate (0.34 mL, 2.22 mmol) in 8 mL pyridine was heated to 70 °C overnight under inert atmosphere, then concentrated under reduced pressure to give a clear, viscous oil. This oil was dissolved in

methanol and purified via an ion exchange column for bases (Isolute SCX-2, Biotage). The methanol fraction was taken to dryness and the product was recrystallized from diethyl ether (2 mL). The resulting off-white crystalline solid was subsequently dried under high vacuum (0.51 g, 1.83 mmol). Yield: 83 %; Mp: 46 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.84-0.91 (m, 6 H), 1.27 (m, 6 H), 1.46-1.63 (m, 4 H), 2.48-2.53 (m, 2 H, overlap with DMSO), 3.42 (q, *J*=5.9 Hz, 2 H), 7.12 (d, *J*=8.4 Hz, 2 H), 7.25 (d, *J*=8.4 Hz, 2 H), 7.60 (br. s, 1 H), 9.30 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.6, 13.9, 22.0, 24.1, 26.1, 28.4, 31.0, 36.7, 43.8, 123.2, 128.4, 136.8, 138.0, 180.2; IR (solid): ν= 3220, 2930, 2860, 1550, 1500, 1270, 837 cm^{-1}; LRMS (ESI+): *m*/*z*= 279.3 [M+H]$^{+}$, 301.3 [M+Na]$^{+}$; HRMS (ES) for $C_{16}H_{26}SN_2Na$ [M+Na]$^{+}$: *m*/*z*= 301.1709 (calcd), 301.1714 (found). The structure of compound **20** was confirmed by single crystal X-ray diffraction (see section S4.16).

1-(4-butylphenyl)-3-hexylthiourea (21). 4-Butylaniline (0.3 g, 2.01 mmol) was dissolved in 10 mL dry pyridine, hexyl isothiocyanate (0.3 mL, 2.01 mmol) was added dropwise and the solution was stirred overnight at 60 °C under nitrogen atmosphere. The solvents were removed under reduced pressure and the crude mixture was purified via an ion exchange column for bases (Isolute SCX-2, Biotage) to yield a yellow viscous oil. This oil was then purified by flash chromatography (100% diethyl ether) to give a waxy yellow solid, which was subsequently dried under high vacuum (0.54 g, 1.84 mmol). Yield: 92 %; Mp: 51 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.86-0.92 (m, 6 H), 1.28 (m, 8 H), 1.49-1.58 (m, 4 H), 2.54 (t, *J*=7.6 Hz, 2 H, overlap with DMSO), 3.44 (br q, *J*=5.6 Hz, 2 H), 7.12 (d, *J*=8.1 Hz, 2 H), 7.26 (d, *J*=8.6 Hz, 2 H), 7.59 (br. s, 1 H), 9.30 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.7, 13.9, 21.7, 22.0, 26.0, 28.4, 31.0, 33.1, 34.2, 43.8, 123.2, 128.3, 136.8, 138.2, 180.2; IR (solid): ν= 3210, 2930, 2860, 1550, 1500, 1270, 837 cm^{-1}; LRMS (ESI+): *m*/*z*= 293.4 [M+H]$^{+}$, 315.4 [M+Na]$^{+}$; HRMS (ES) for $C_{17}H_{28}SN_2Na$ [M+Na]$^{+}$: *m*/*z*= 315.1865 (calcd), 315.1870 (found).

1-(4-pentylphenyl)-3-hexylthiourea (22). A solution of 4-pentylaniline (0.54 mL, 3.06 mmol) and hexyl isothiocyanate (0.47 mL, 3.06 mmol) in 10 mL pyridine was heated to 70 °C overnight under inert atmosphere and then concentrated *in vacuo* to give a clear, viscous oil. This oil was dissolved in methanol and purified via an ion exchange column for bases (Isolute SCX-2, Biotage). The methanol fraction was taken to dryness and the product was recrystallized from diethyl ether (2 mL). The product was then further purified via flash chromatography (diethylether:hexane 1:1) and subsequently dried under high vacuum to yield a waxy solid (0.72 g, 2.35 mmol). Yield: 77 %; Mp: 48 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 0.85-0.89 (m, 6 H), 1.28 (m, 10 H), 1.48-1.61 (m, 4 H), 2.53 (t, *J*=7.3 Hz, 2 H, overlap with DMSO), 3.44 (q, *J*=5.9 Hz, 2 H), 7.12 (d, *J*=8.4 Hz, 2 H), 7.26 (d, *J*=8.4 Hz, 2 H), 7.60 (br. s, 1 H), 9.31 (br. s, 1 H); ^{13}C{^{1}H} NMR (75 MHz, DMSO-d_6) δ ppm 13.9 (2 peaks), 21.9, 22.0, 26.1, 28.4, 30.6, 30.9, 31.0, 34.5, 43.8, 123.2, 128.3, 136.8, 138.2, 180.2; IR (solid): ν= 3190, 2920, 2850, 1540 cm^{-1}; LRMS (ESI+): *m*/*z*= 307.4 [M+H]$^{+}$, 329.4 [M+Na]$^{+}$; HRMS (ES) for $C_{18}H_{30}SN_2Na$ [M+Na]$^{+}$: *m*/*z*= 329.3027 (calcd), 329.2027 (found).

S3. NMR SPECTRA

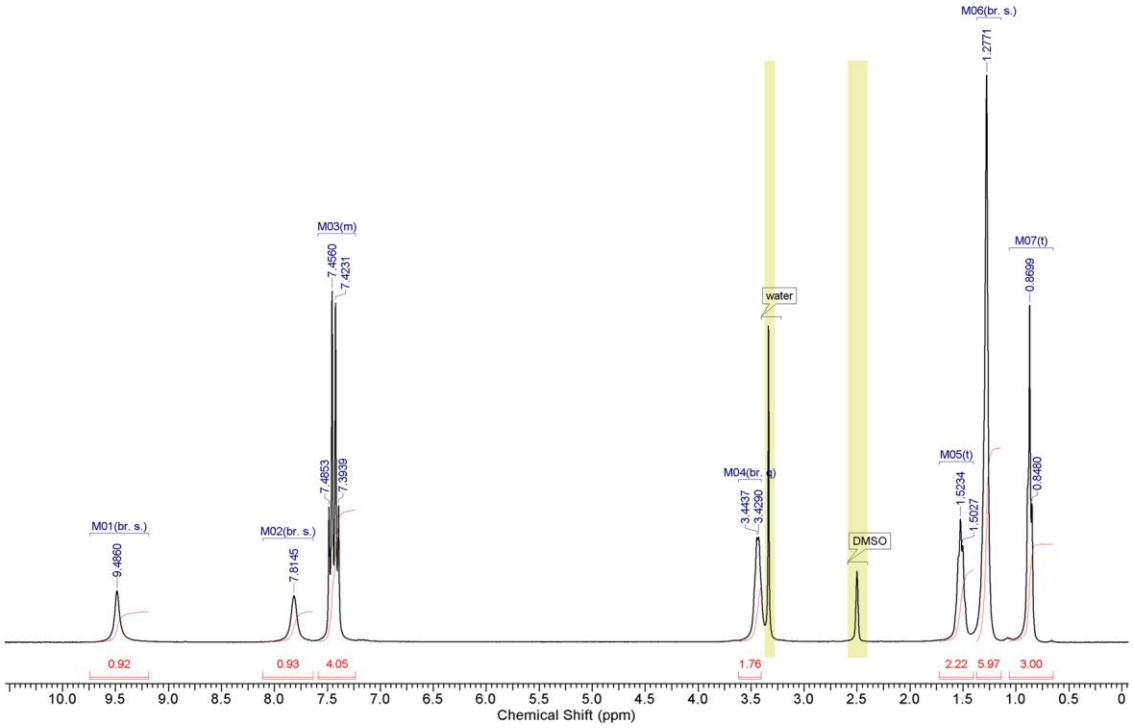


Figure S1. [1]H NMR spectrum of compound **1** (Br) in DMSO-d_6 at 298K.

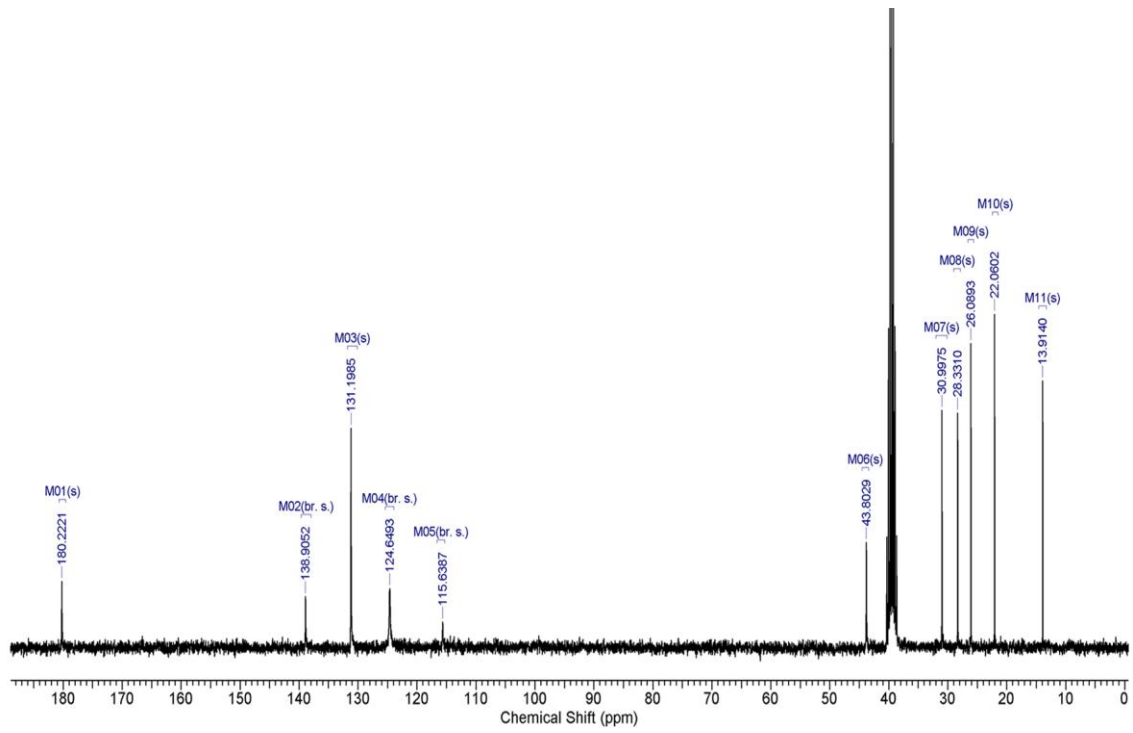


Figure S2. [13]C NMR spectrum of compound **1** (Br) in DMSO-d_6 at 298K.

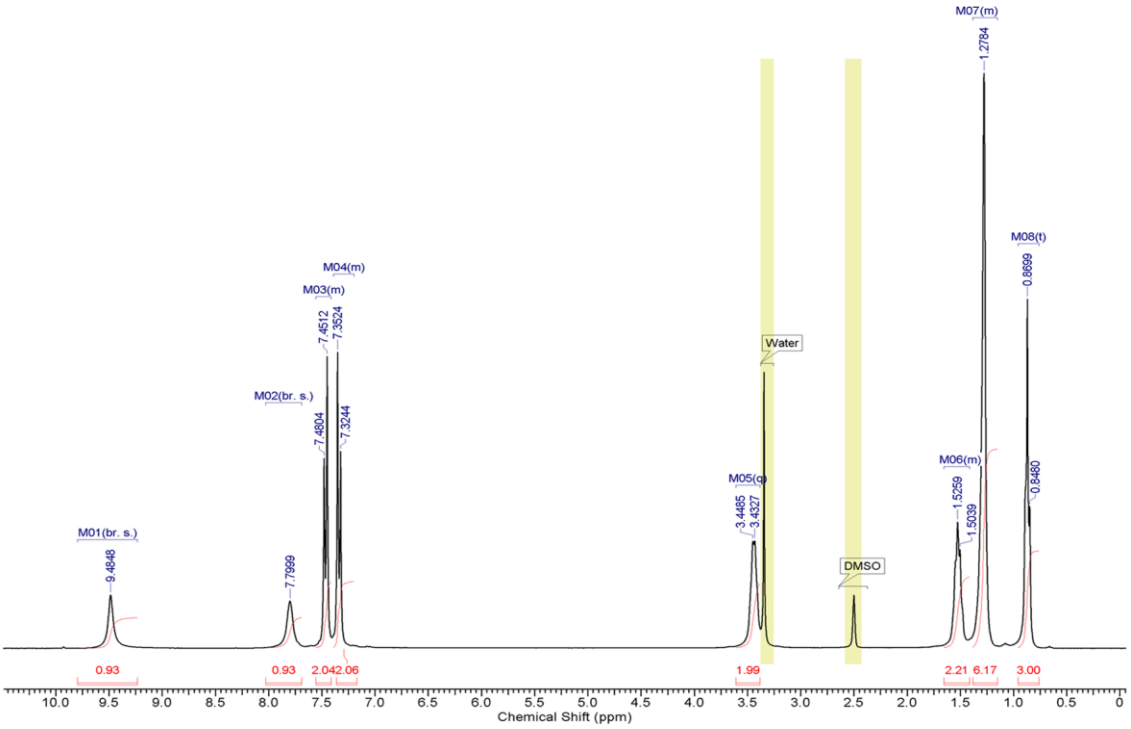


Figure S3. [1]H NMR spectrum of compound **3** (Cl) in DMSO-d_6 at 298K.

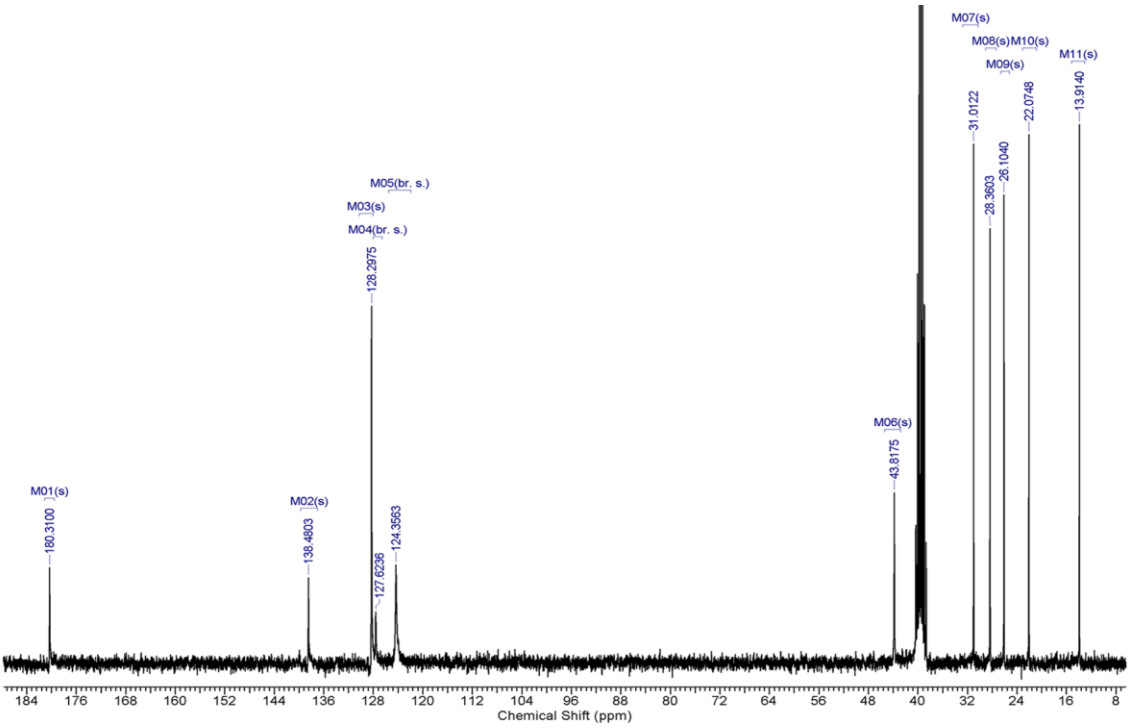


Figure S4. [13]C NMR spectrum of compound **3** (Cl) in DMSO-d_6 at 298K.

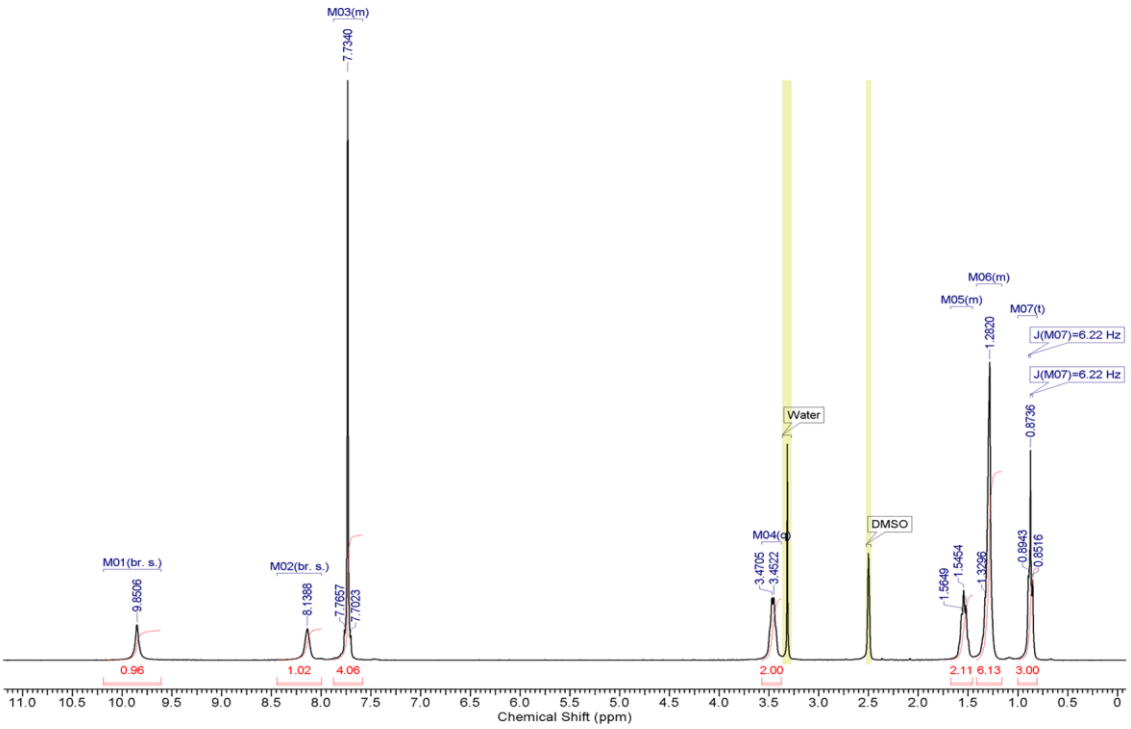


Figure S5. [1]H NMR spectrum of compound **4** (CN) in DMSO-d_6 at 298K.

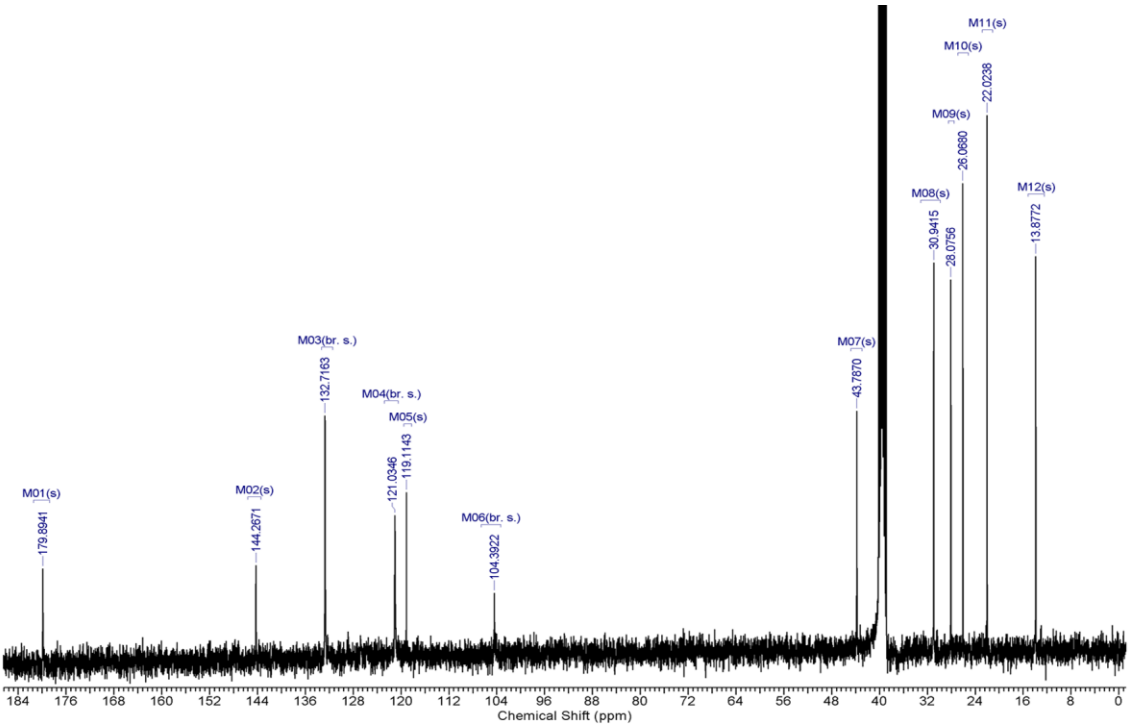


Figure S6. [13]C NMR spectrum of compound **4** (CN) in DMSO-d_6 at 298K.

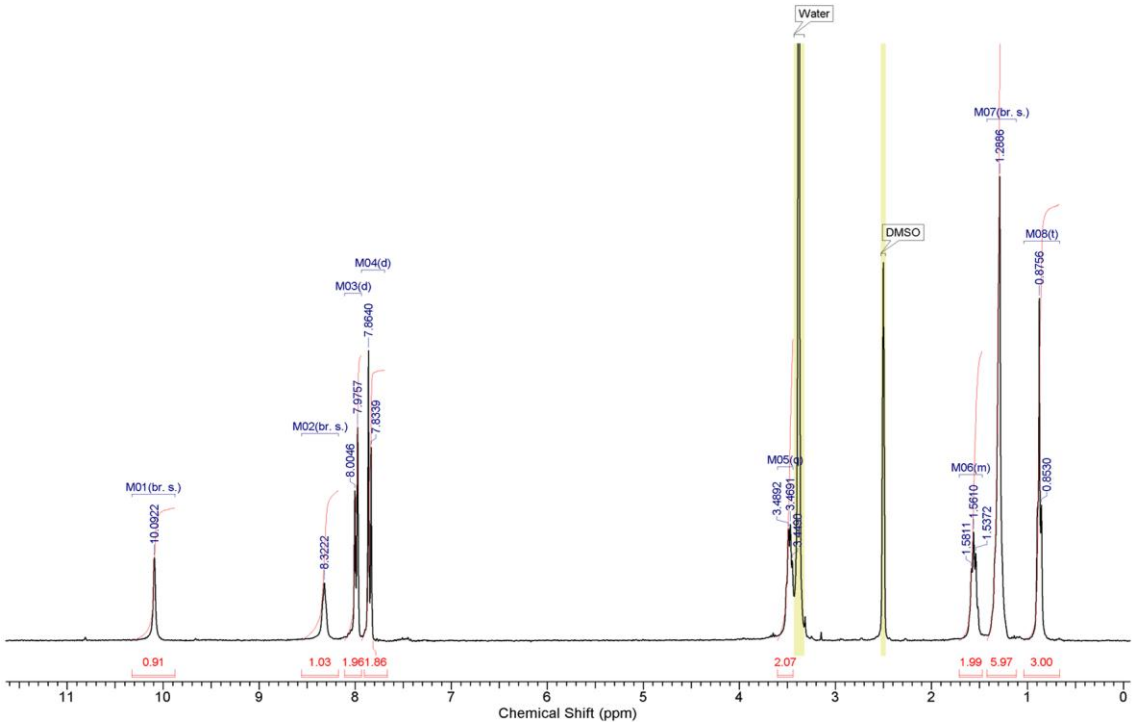


Figure S7. ^{1}H NMR spectrum of compound **5** ($COCF_3$) in DMSO-d_6 at 298K.

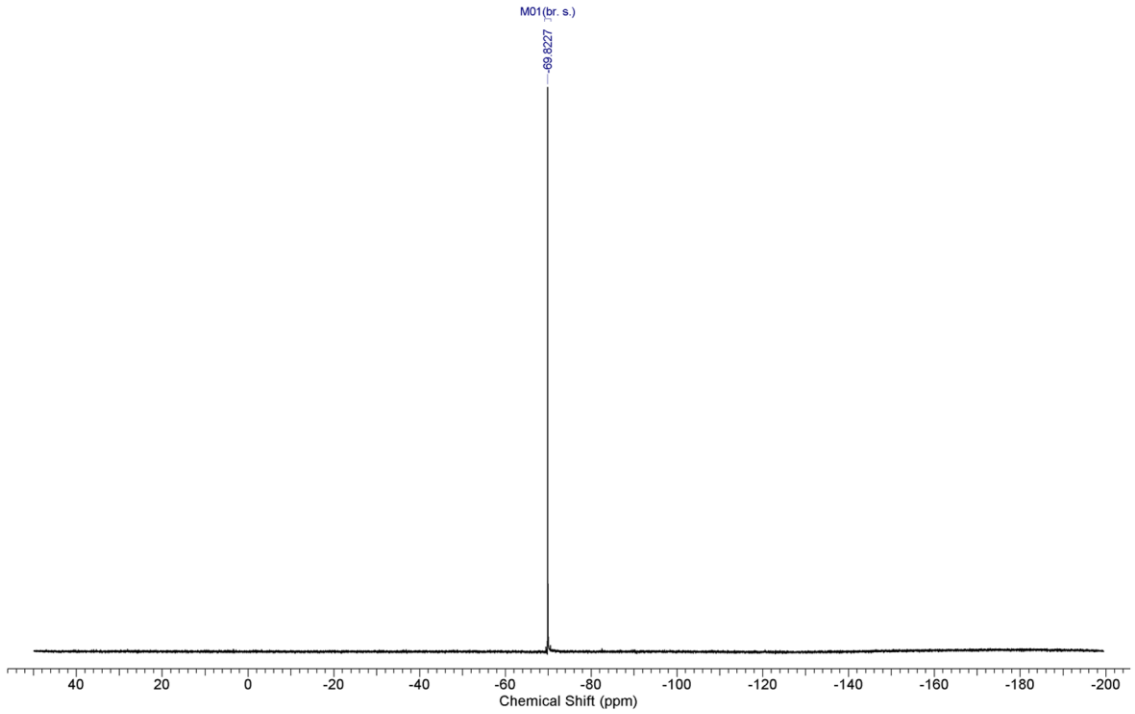


Figure S8. ^{19}F NMR spectrum of compound **5** ($COCF_3$) in DMSO-d_6 at 298K.

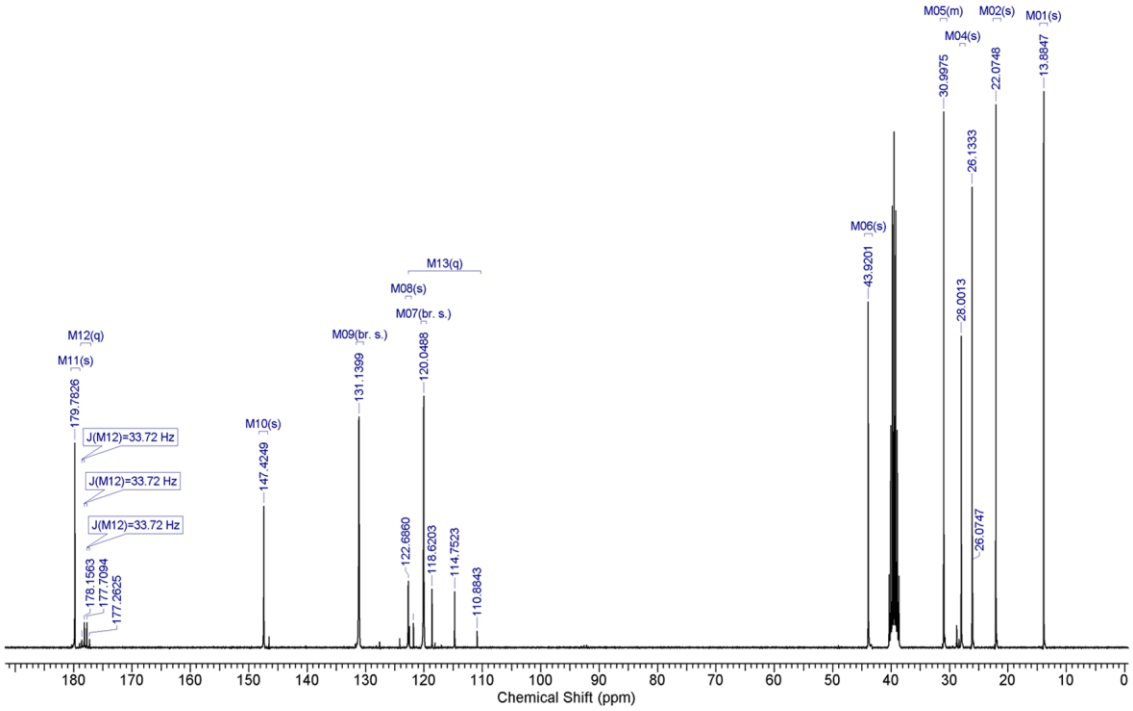


Figure S9. ^{13}C NMR spectrum of compound **5** ($COCF_3$) in DMSO-d_6 at 298K.

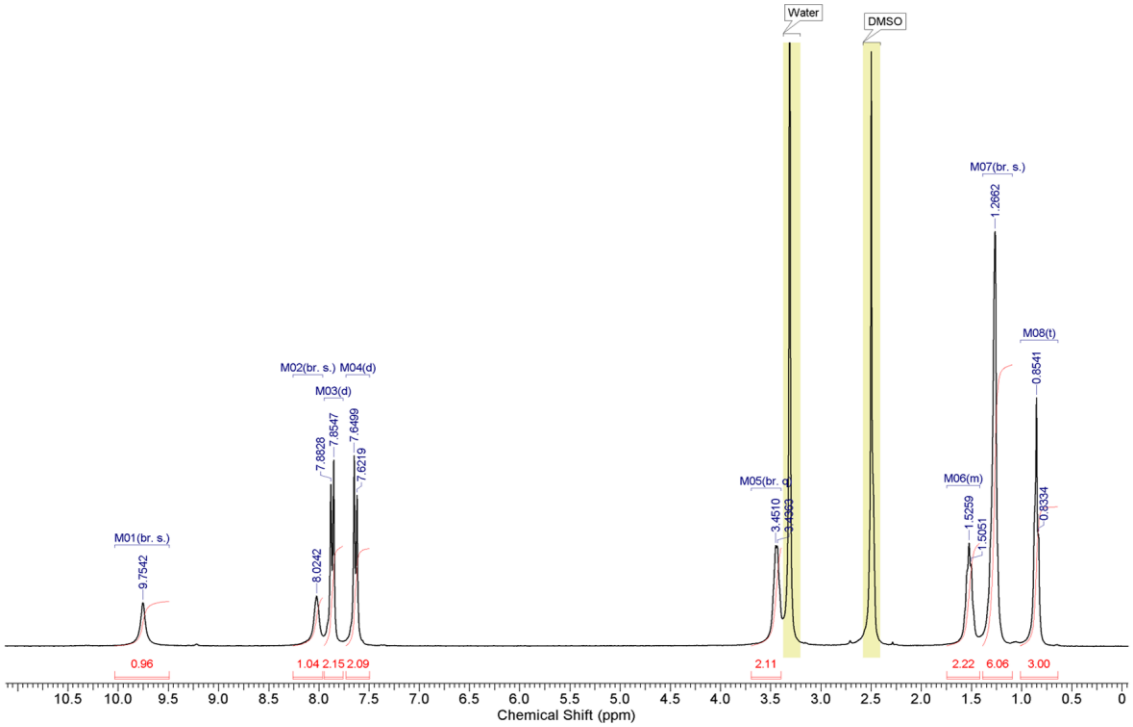


Figure S10. ^{1}H NMR spectrum of compound **6** (COMe) in DMSO-d_6 at 298K.

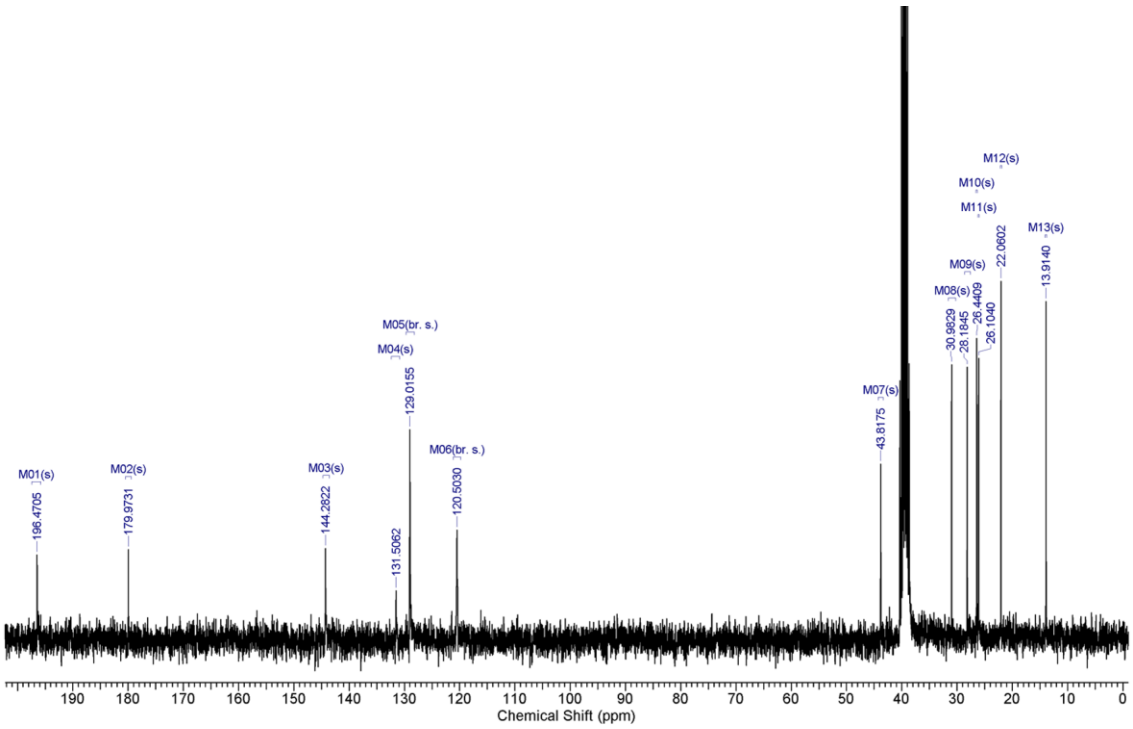


Figure S11. ^{13}C NMR spectrum of compound **6** (COMe) in DMSO-d_6 at 298K.

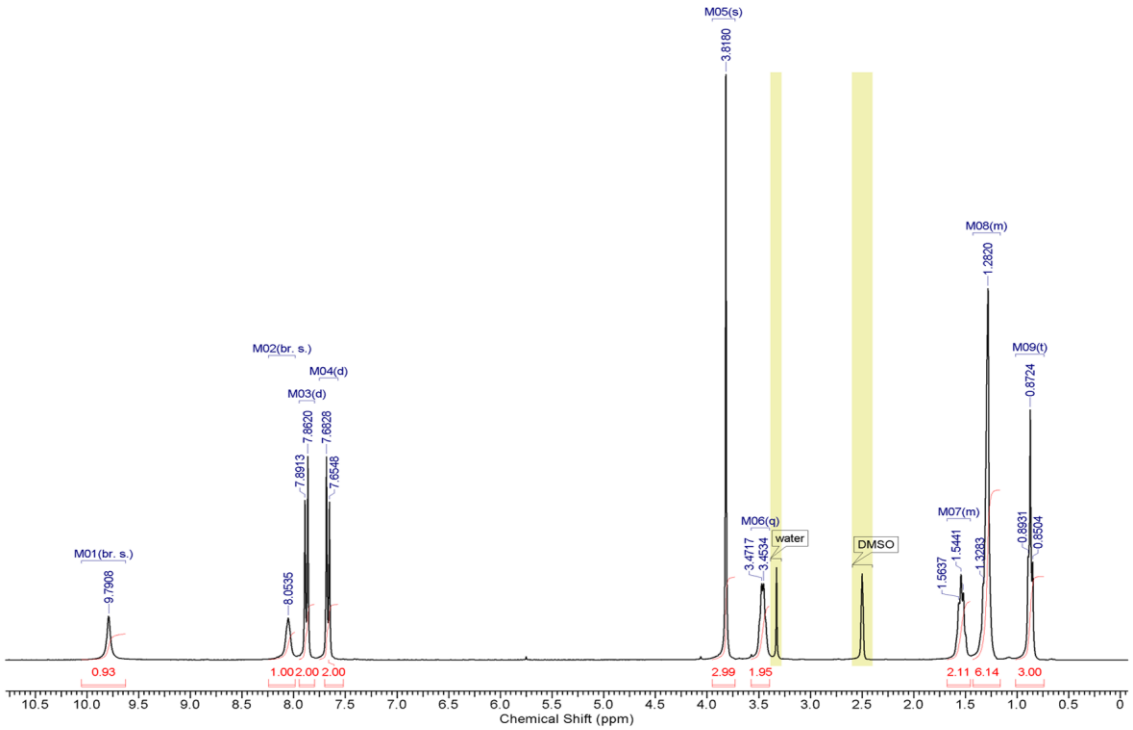


Figure S12. ^{1}H NMR spectrum of compound **7** (COOMe) in DMSO-d_6 at 298K.

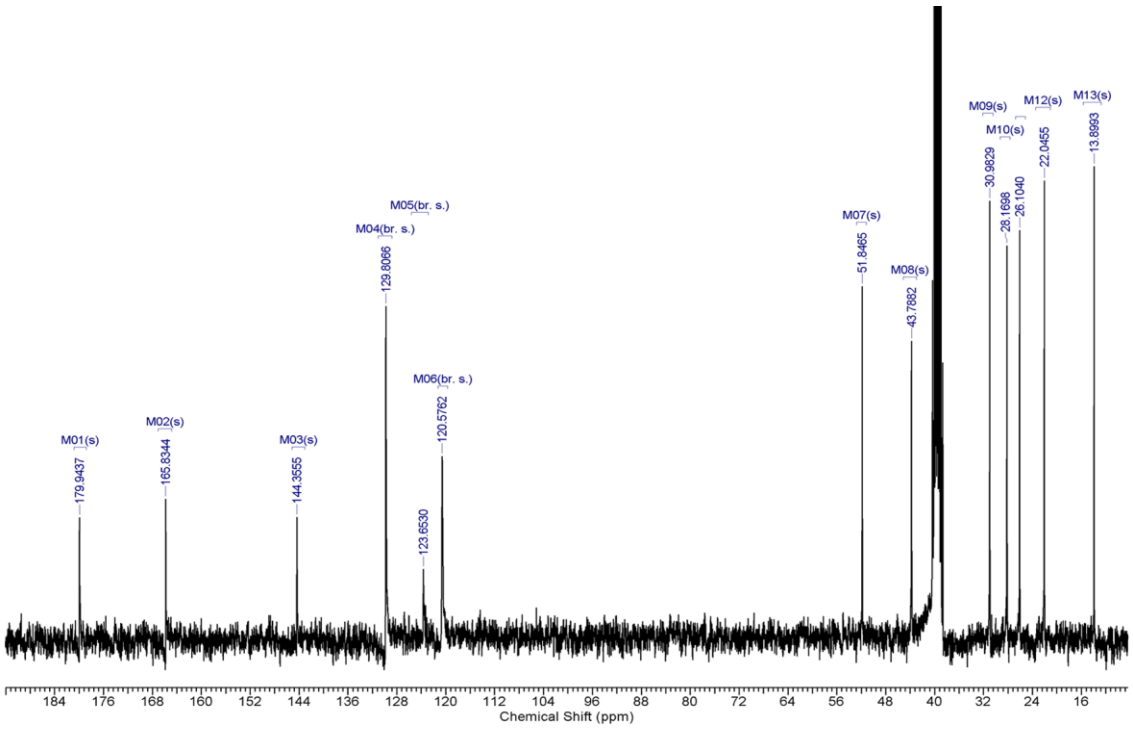


Figure S13. ^{13}C NMR spectrum of compound **7** (COOMe) in DMSO-d_6 at 298K.

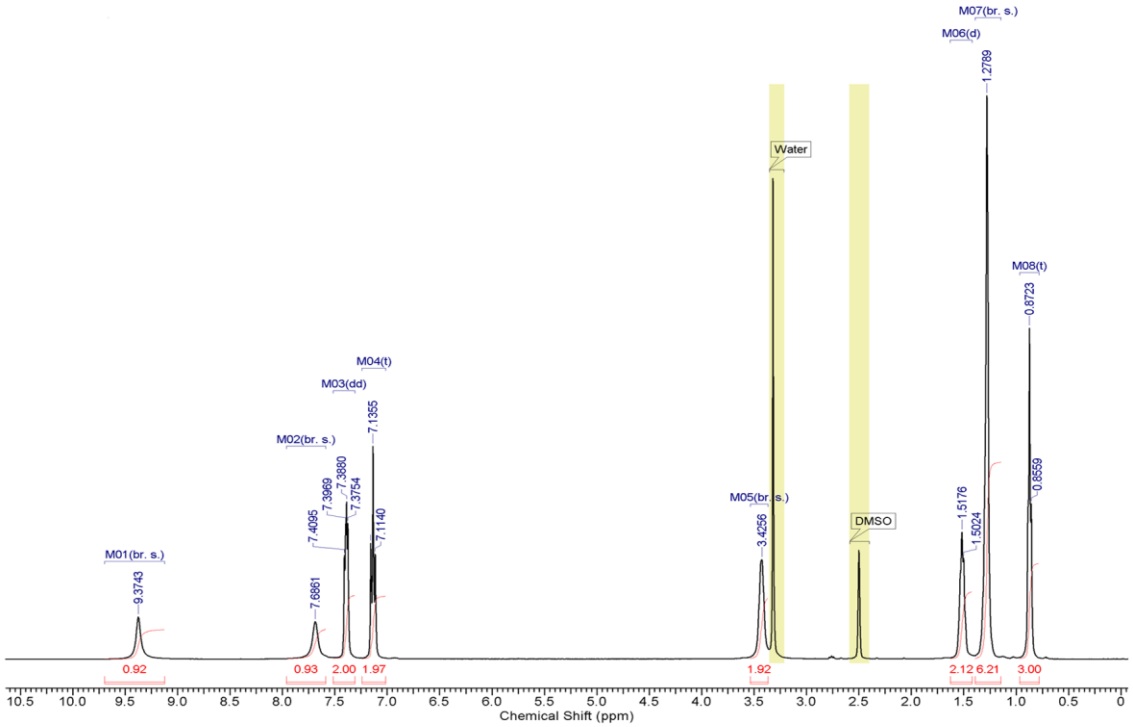


Figure S14. ^{1}H NMR spectrum of compound **8** (F) in DMSO-d_6 at 298K.

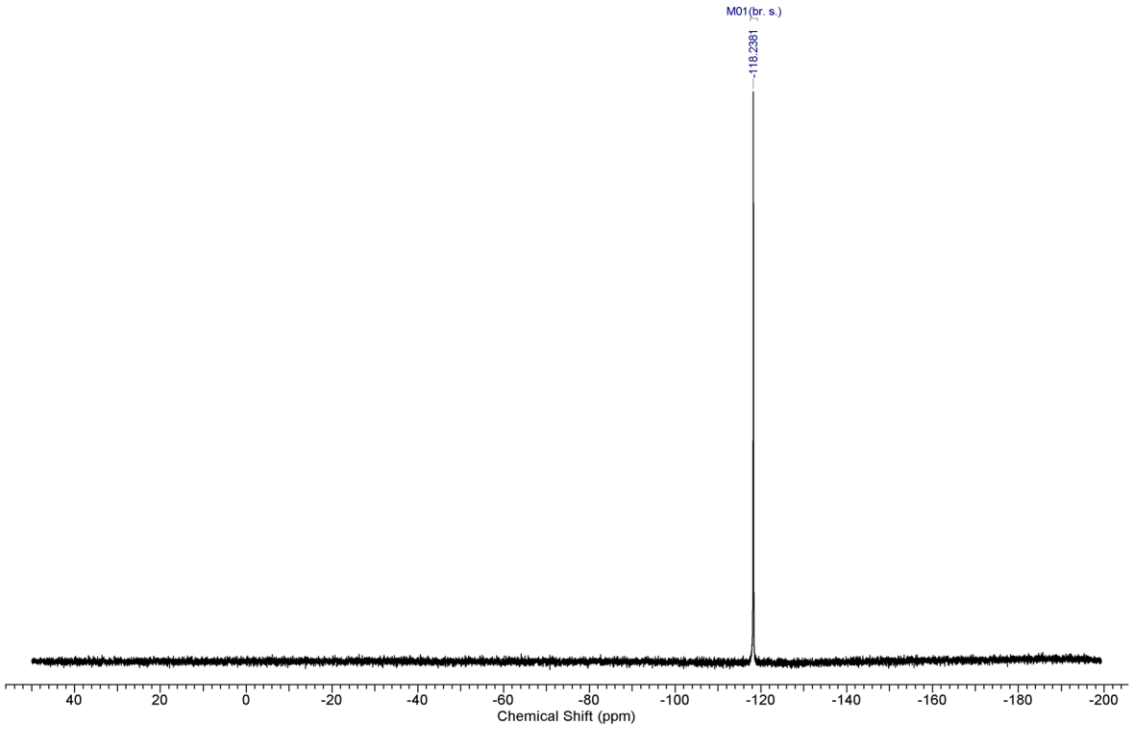


Figure S15. [19]F NMR spectrum of compound **8** (F) in DMSO-d_6 at 298K.

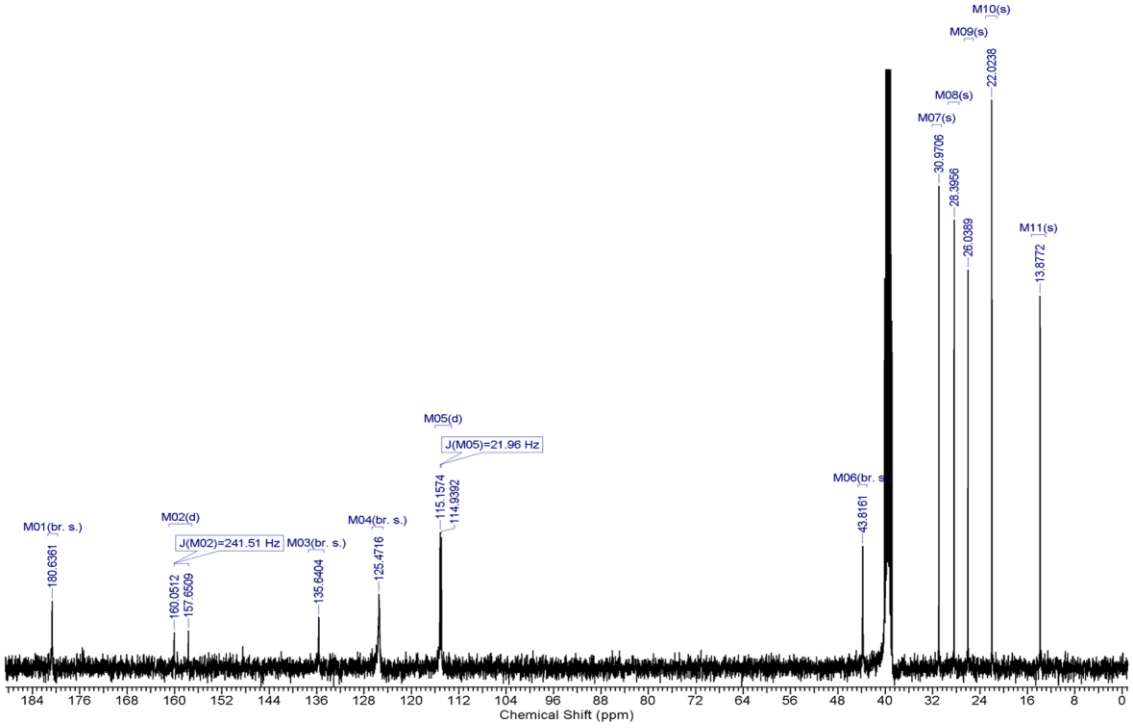


Figure S16. [13]C NMR spectrum of compound **8** (F) in DMSO-d_6 at 298K.

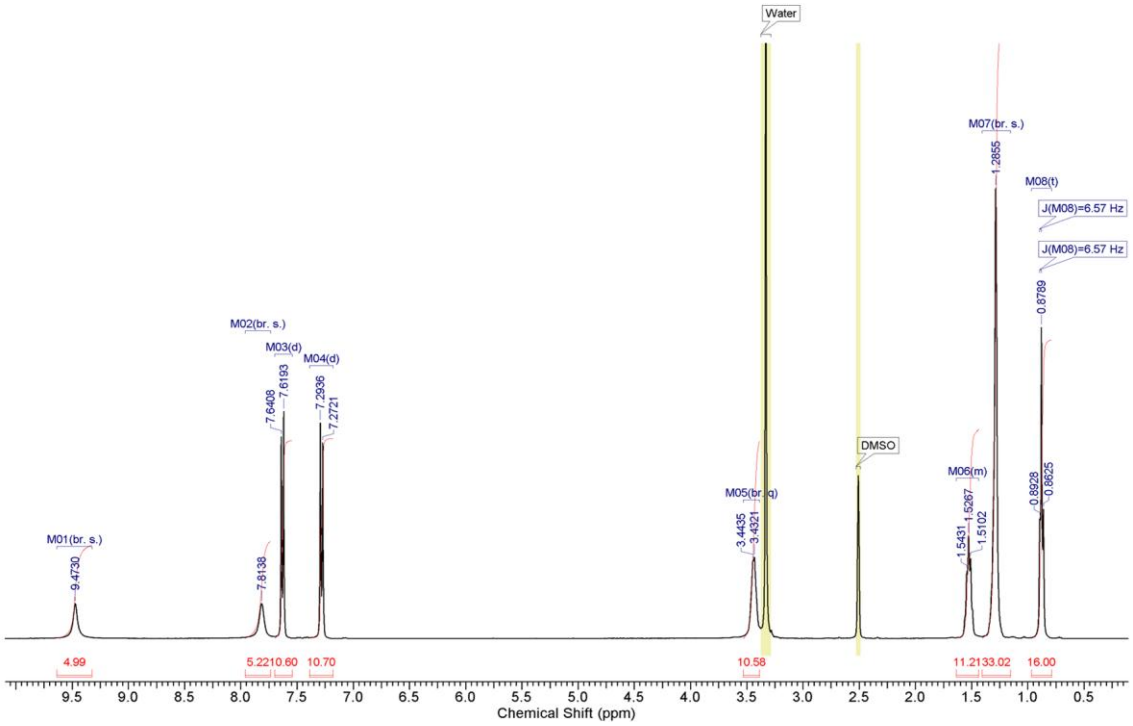


Figure S17. ^{1}H NMR spectrum of compound **10** (I) in DMSO-d_6 at 298K.

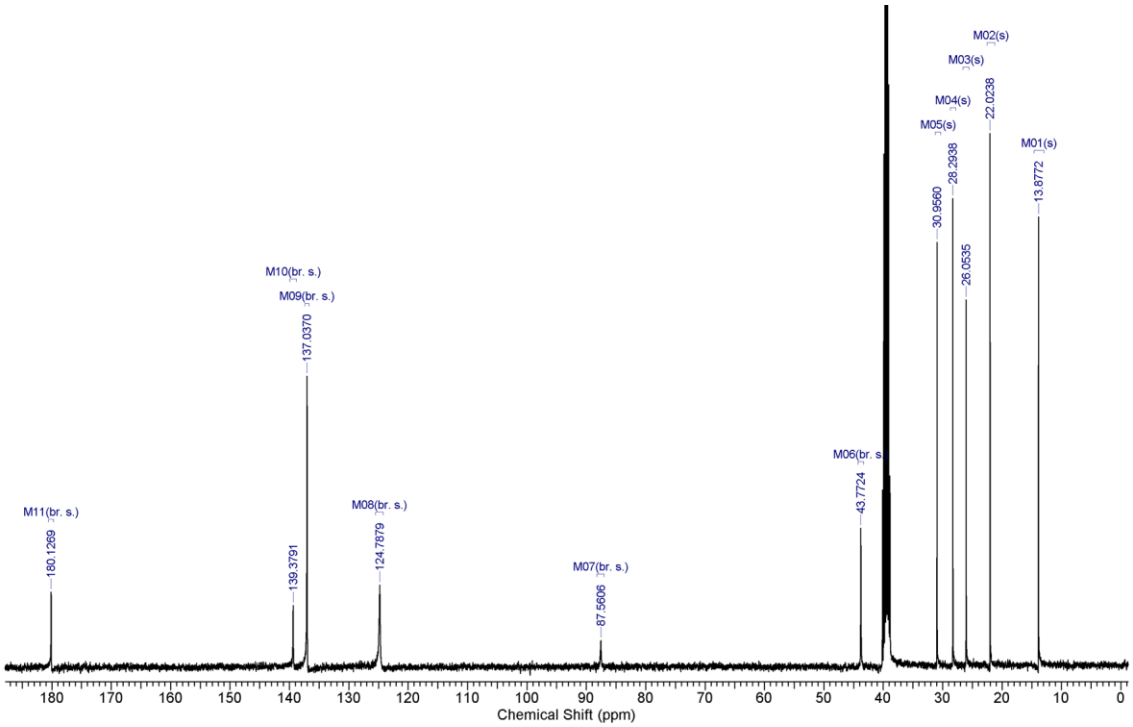


Figure S18. ^{13}C NMR spectrum of compound **10** (I) in DMSO-d_6 at 298K.

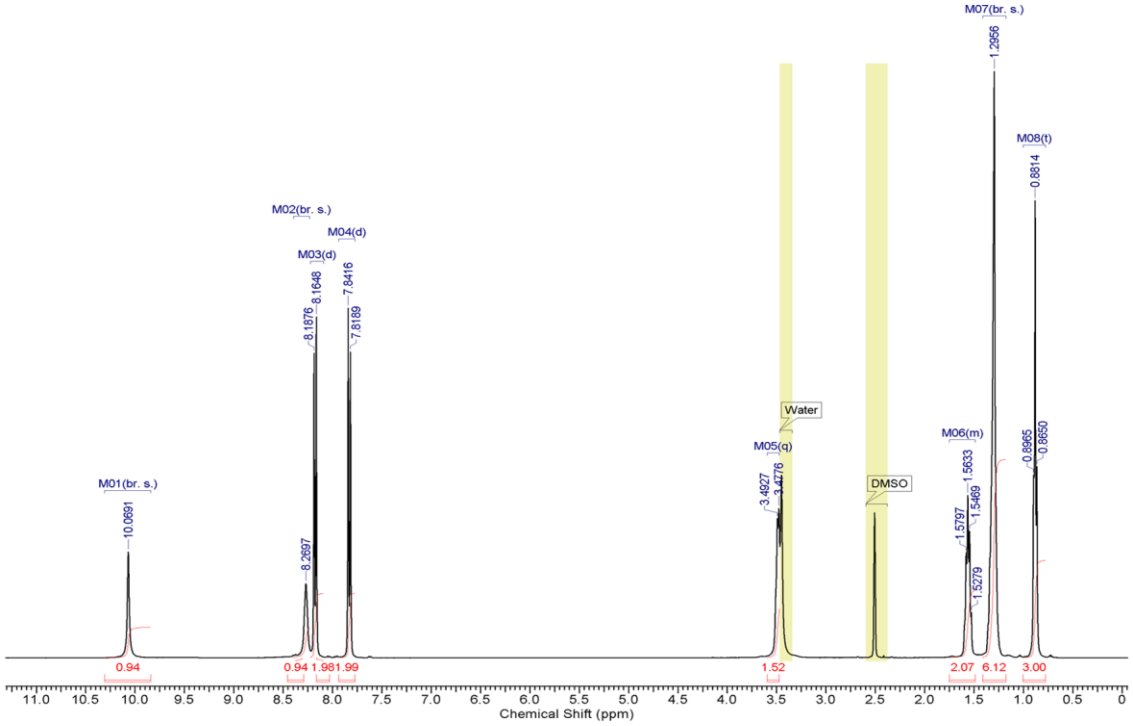


Figure S19. ^{1}H NMR spectrum of compound **11** (NO_2) in DMSO-d_6 at 298K.

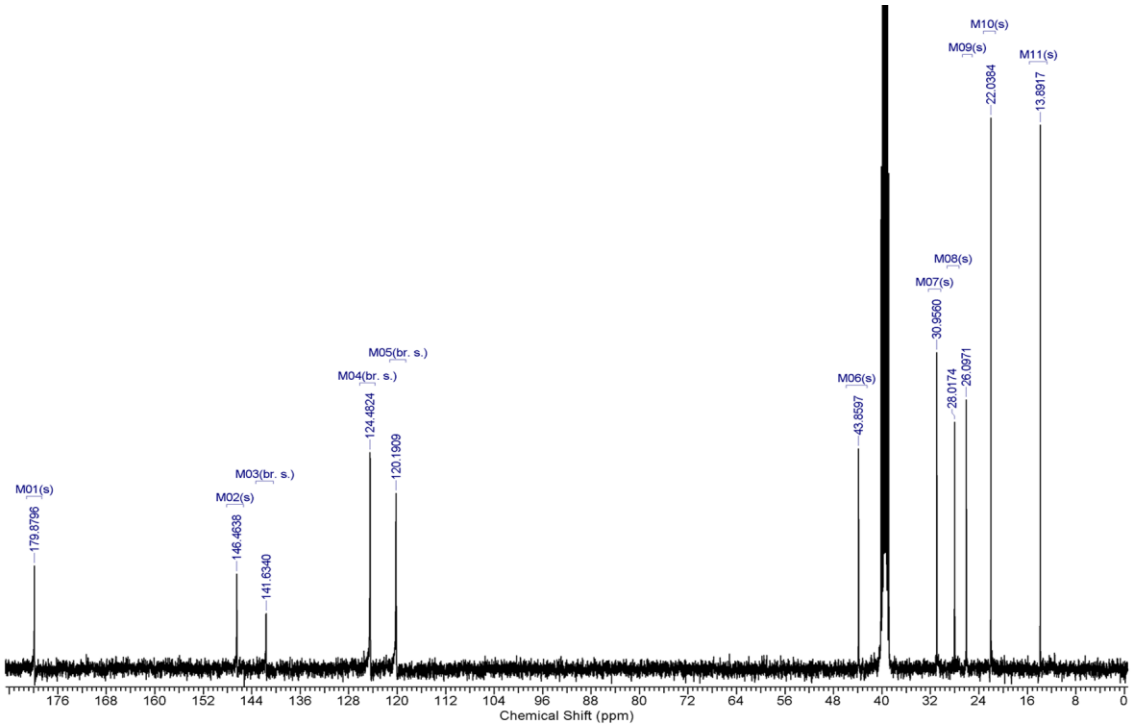


Figure S20. ^{13}C NMR spectrum of compound **11** (NO_2) in DMSO-d_6 at 298K.

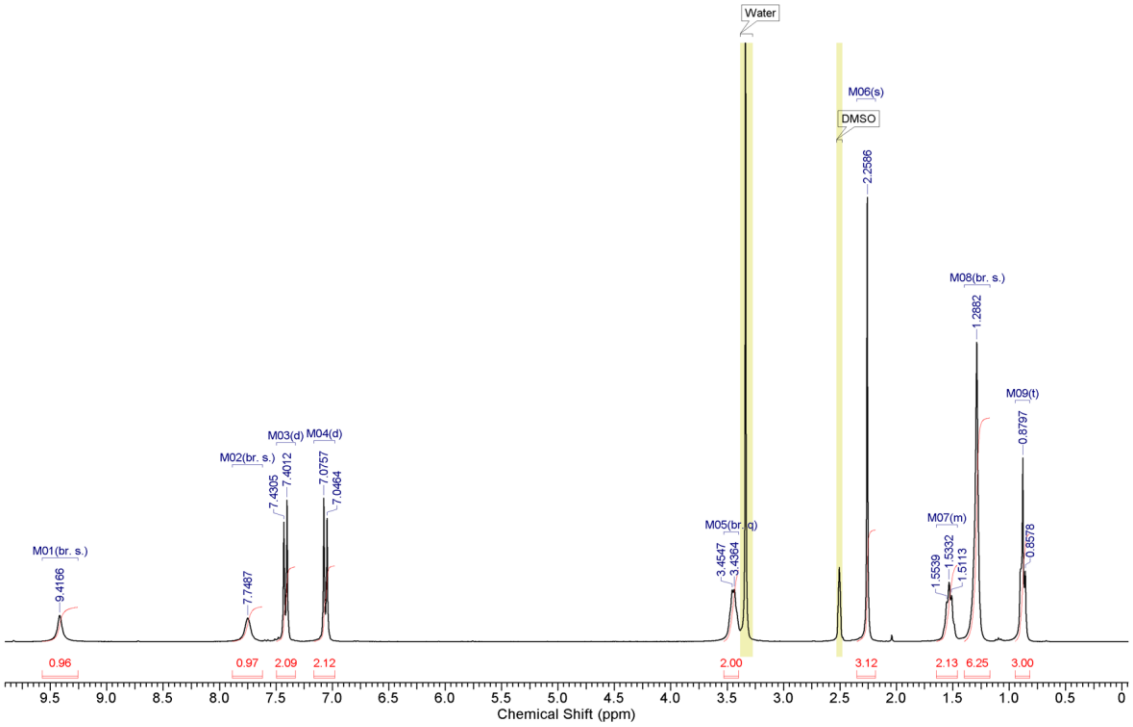


Figure S21. ^{1}H NMR spectrum of compound **12** (OCOMe) in DMSO-d_6 at 298K.

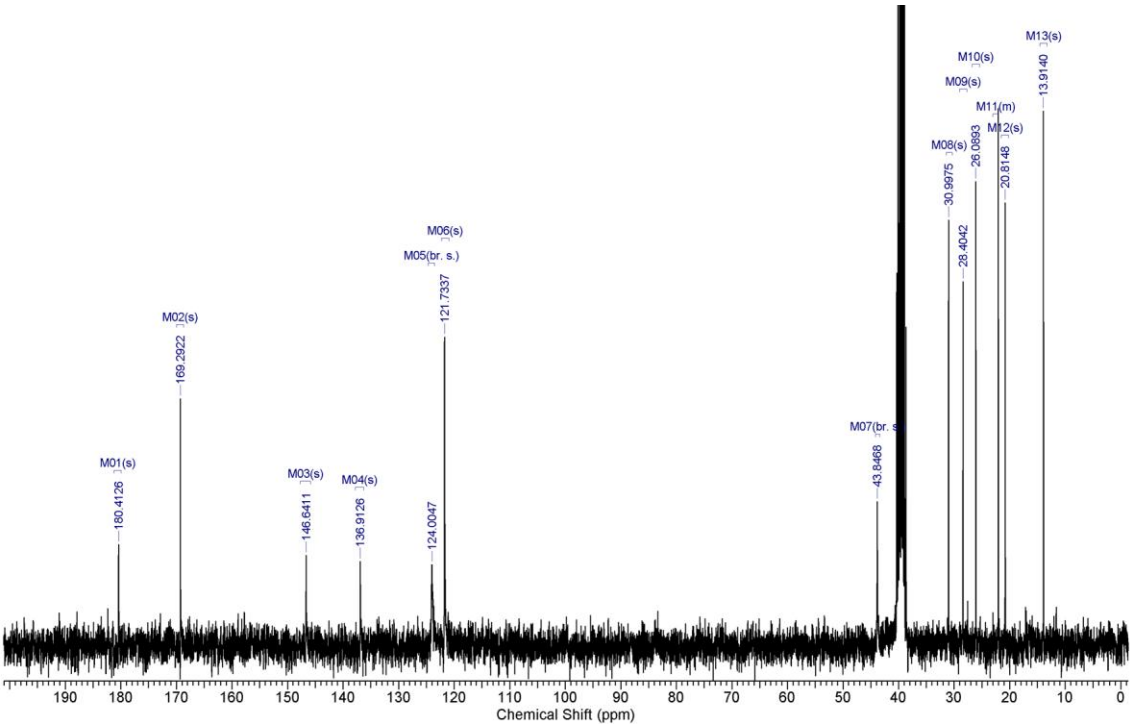


Figure S22. ^{13}C NMR spectrum of compound **12** (OCOMe) in DMSO-d_6 at 298K.

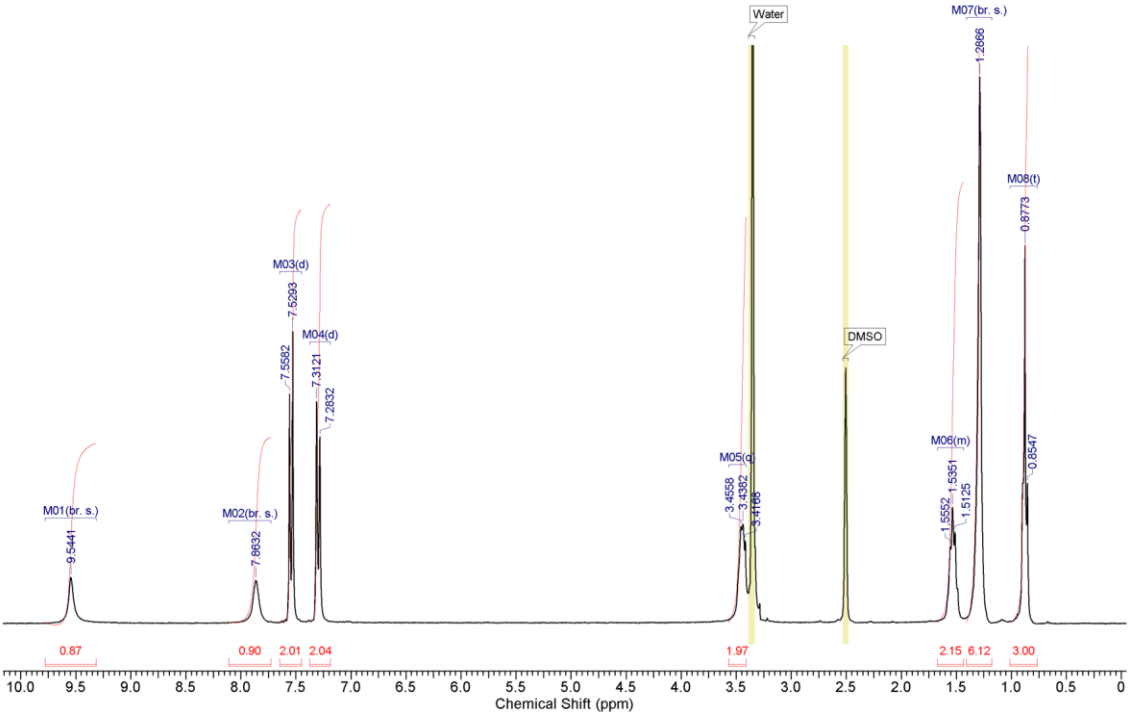


Figure S23. ^{1}H NMR spectrum of compound **13** (OCF_3) in DMSO-d_6 at 298K.

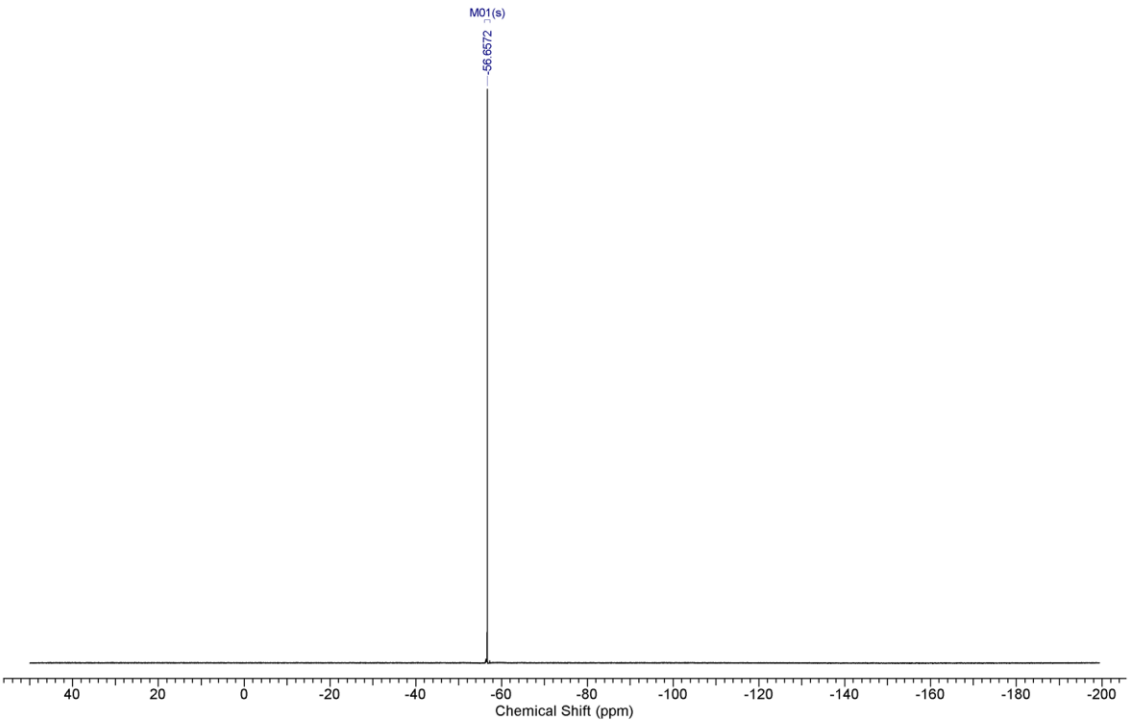


Figure S24. ^{19}F NMR spectrum of compound **13** (OCF_3) in DMSO-d_6 at 298K.

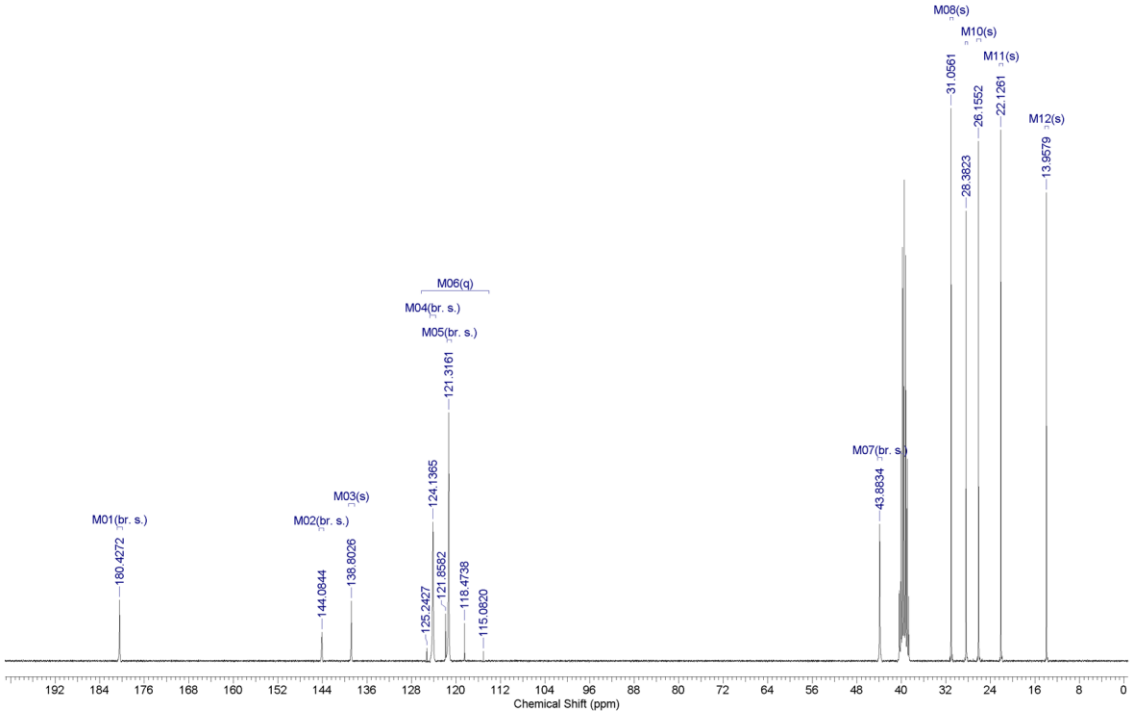


Figure S25. ^{13}C NMR spectrum of compound **13** (OCF_3) in DMSO-d_6 at 298K.

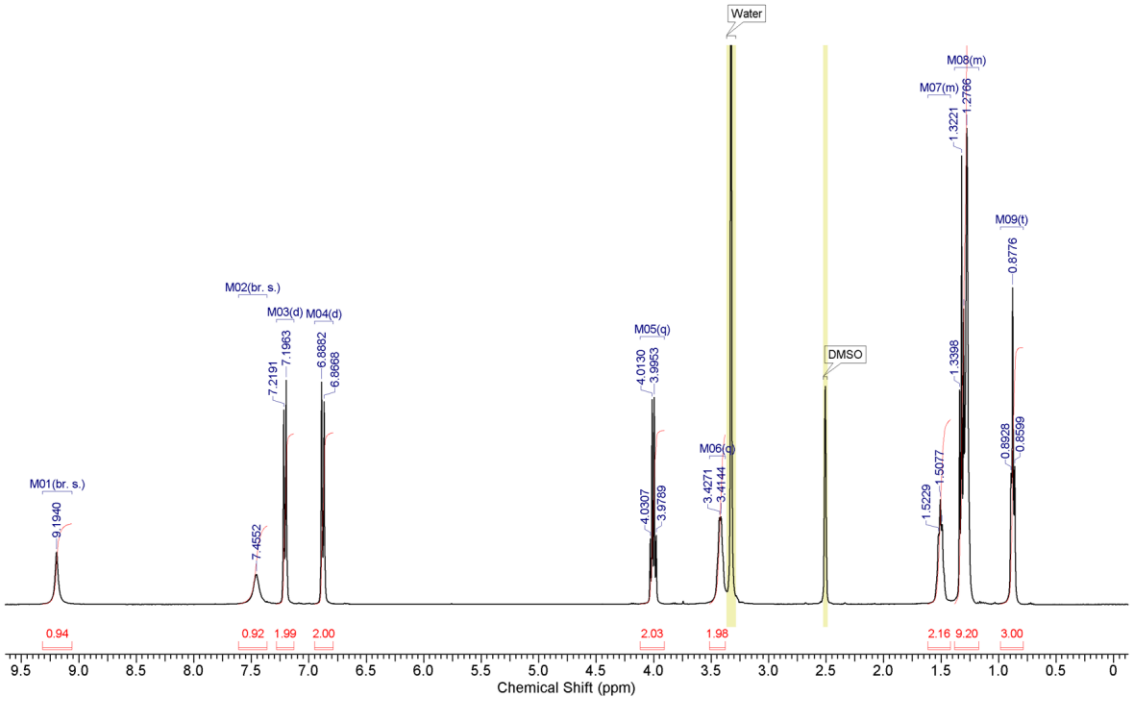


Figure S26. ^{1}H NMR spectrum of compound **14** (OEt) in DMSO-d_6 at 298K.

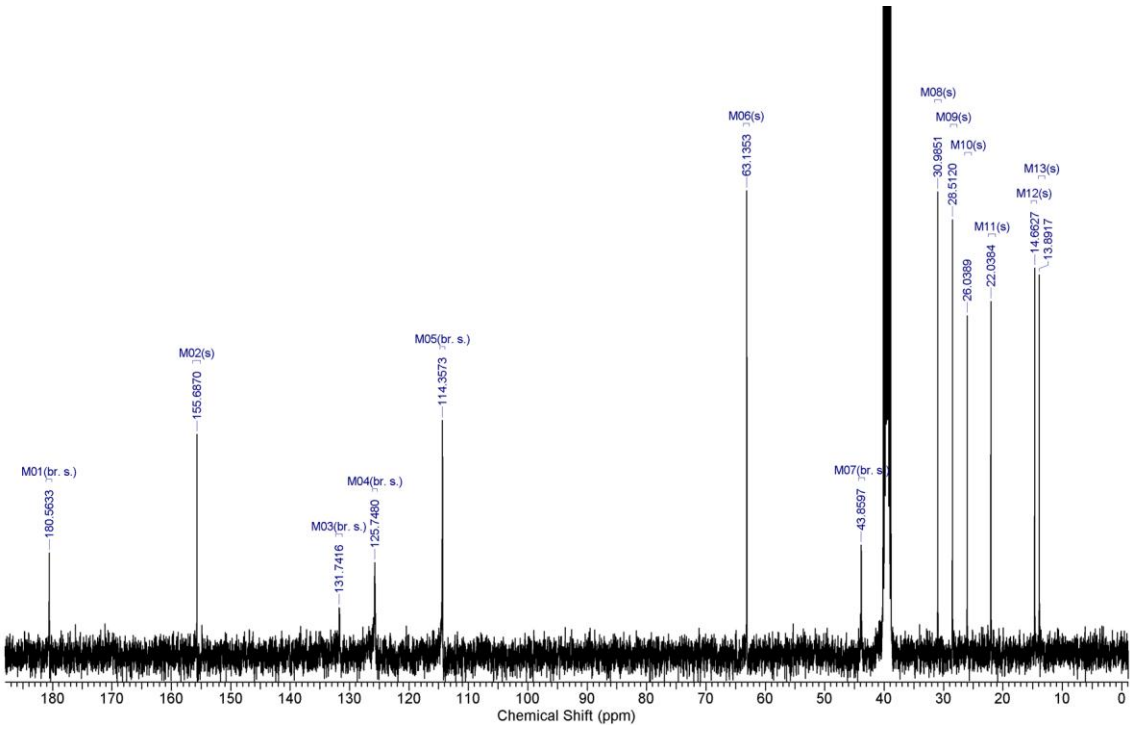


Figure S27. ^{13}C NMR spectrum of compound **14** (OEt) in DMSO-d_6 at 298K.

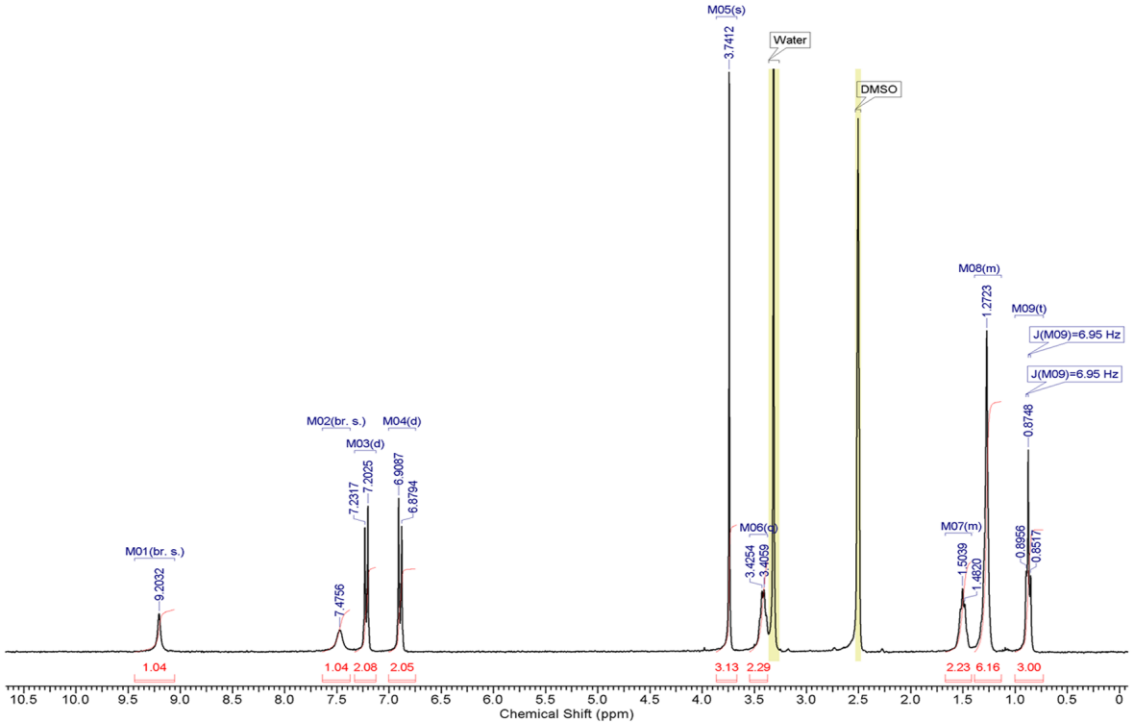


Figure S28. ^{1}H NMR spectrum of compound **15** (OMe) in DMSO-d_6 at 298K.

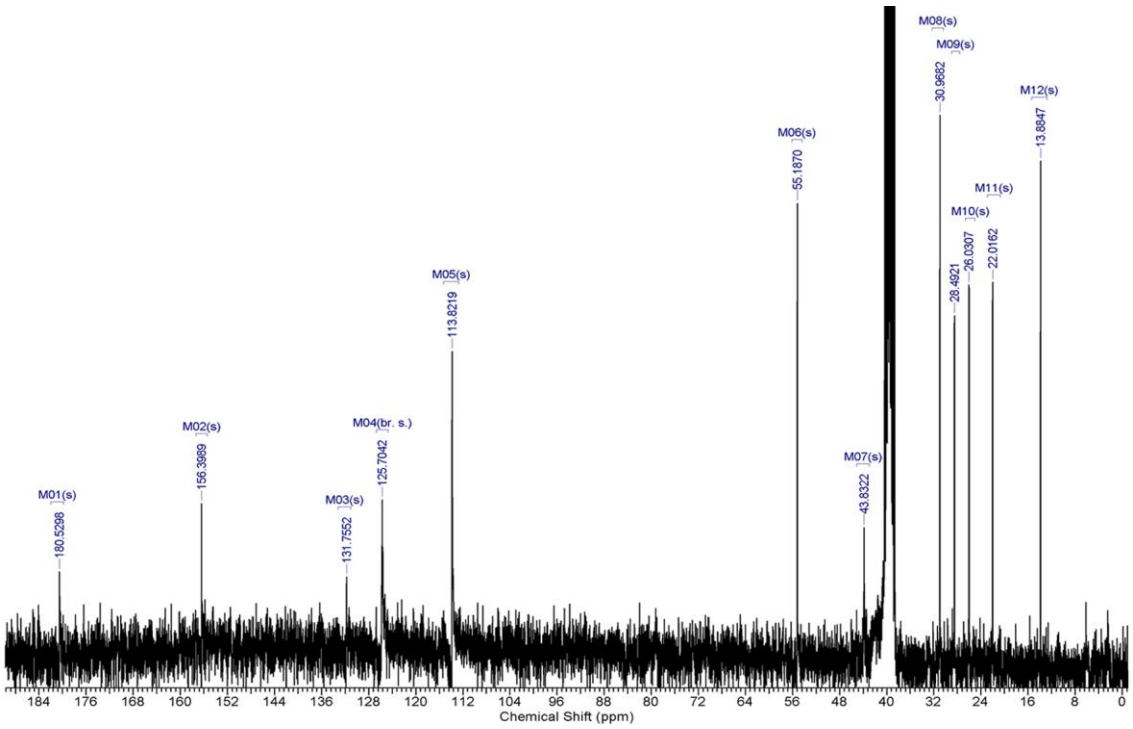


Figure S29. ^{13}C NMR spectrum of compound **15** (OMe) in DMSO-d_6 at 298K.

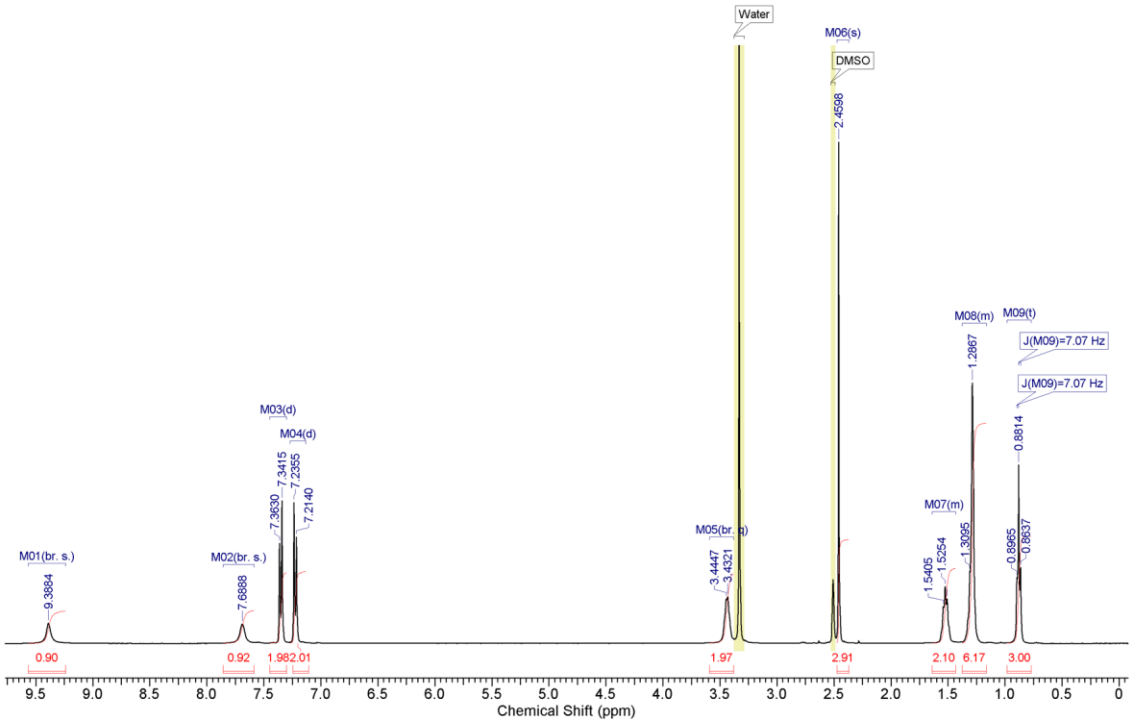


Figure S30. ^{1}H NMR spectrum of compound **16** (SMe) in DMSO-d_6 at 298K.

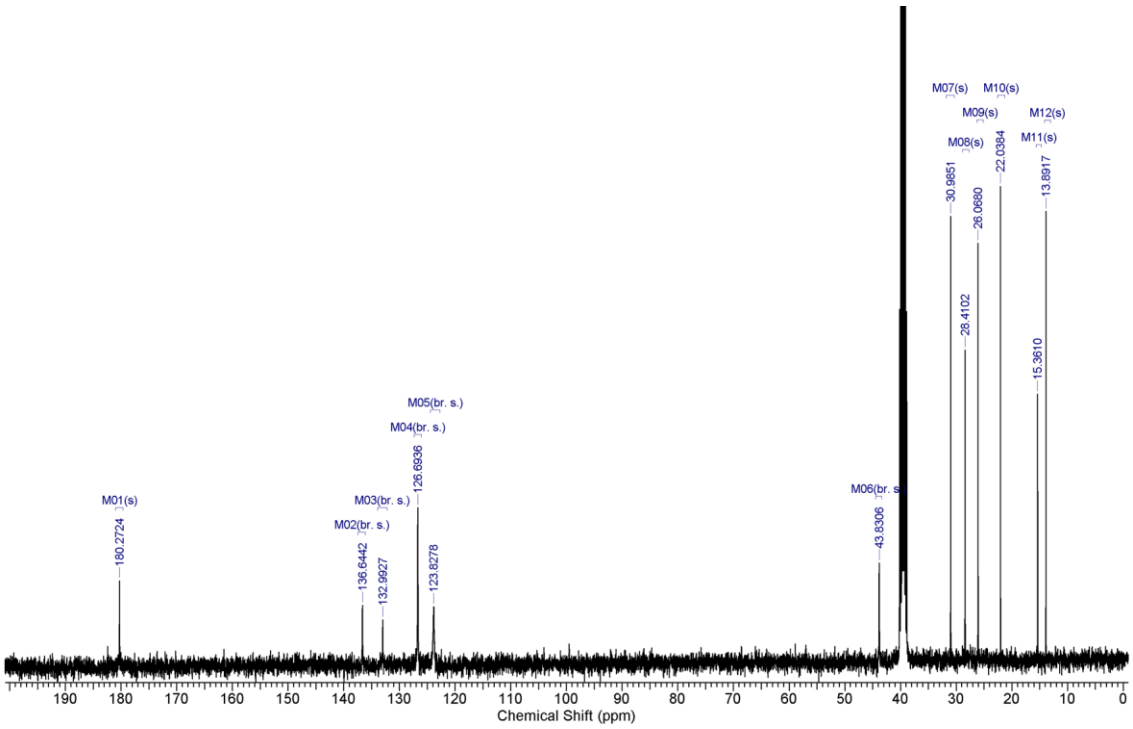


Figure S31. ^{13}C NMR spectrum of compound **16** (SMe) in DMSO-d_6 at 298K.

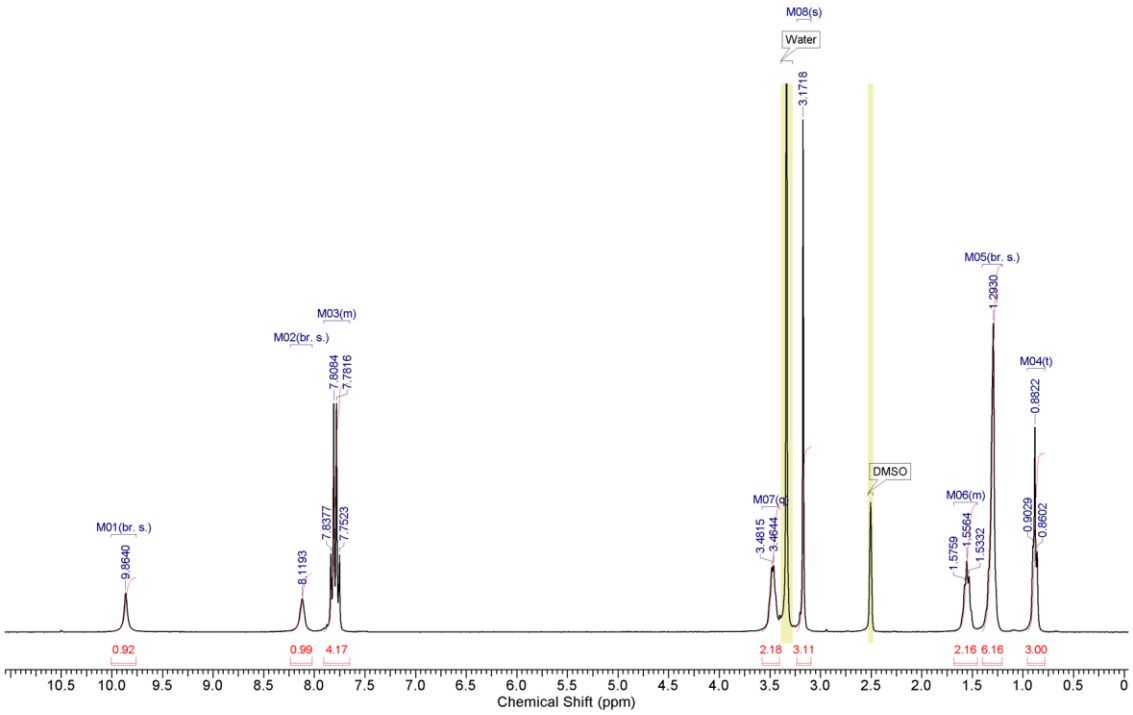


Figure S32. ^{1}H NMR spectrum of compound **17** (SO_2Me) in DMSO-d_6 at 298K.

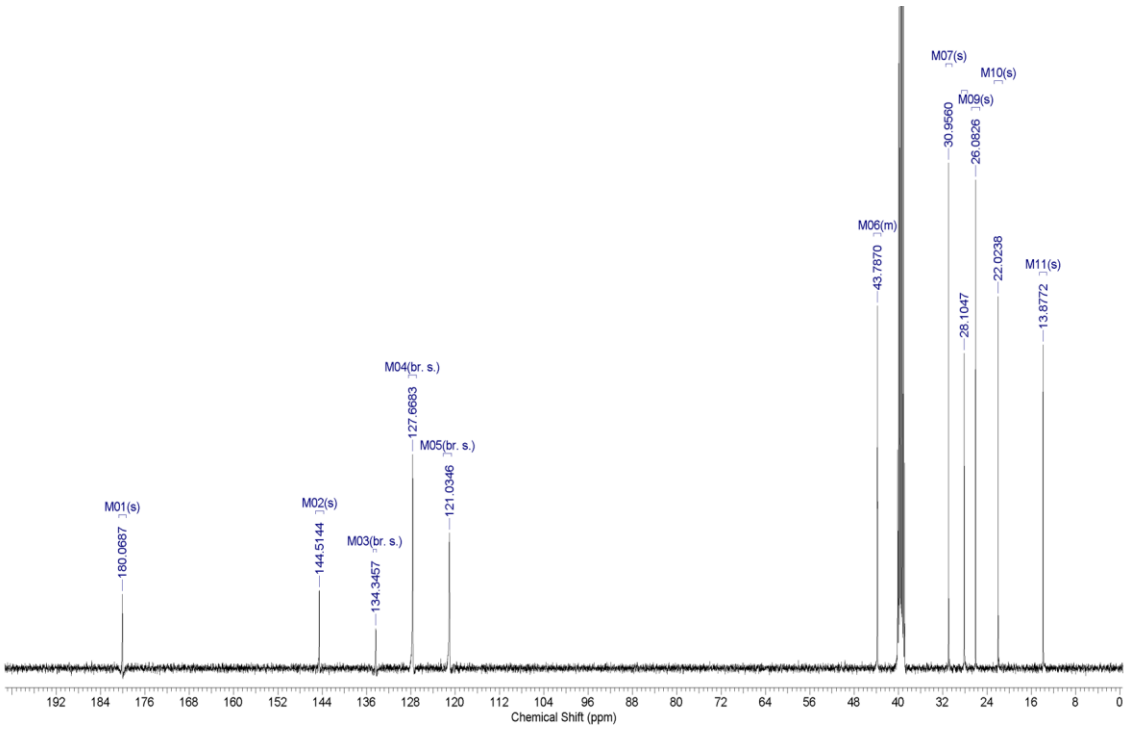


Figure S33. ^{13}C NMR spectrum of compound **17** (SO_2Me) in DMSO-d_6 at 298K.

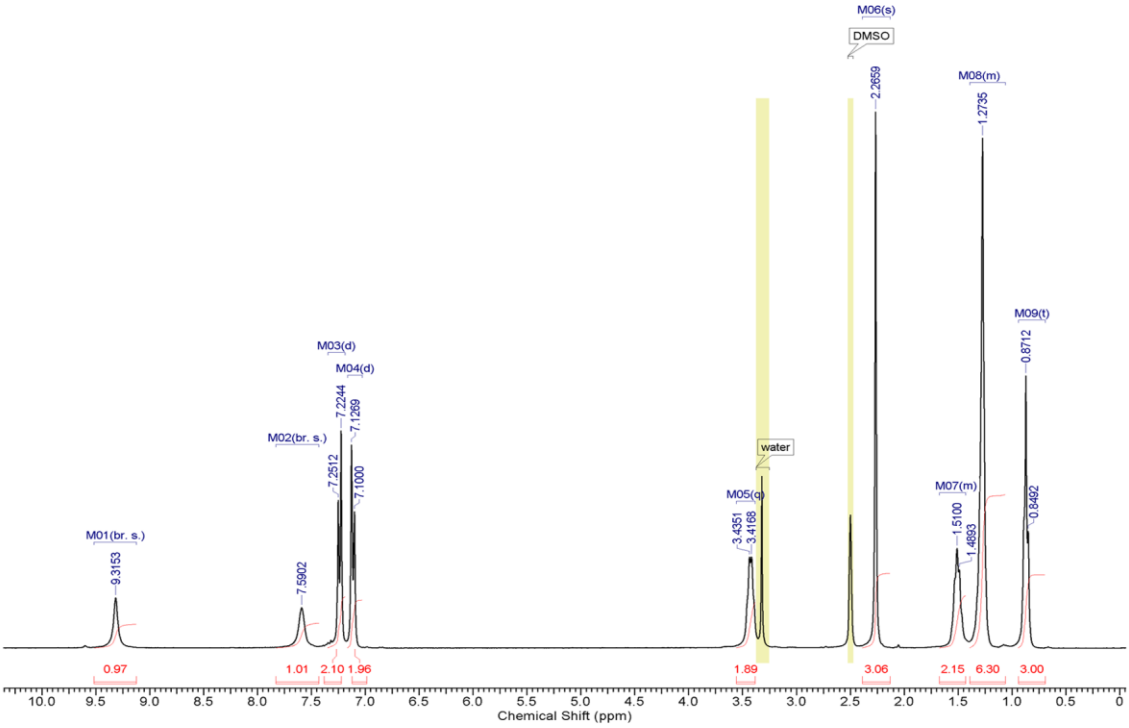


Figure S34. ^{1}H NMR spectrum of compound **18** (Me) in DMSO-d_6 at 298K.

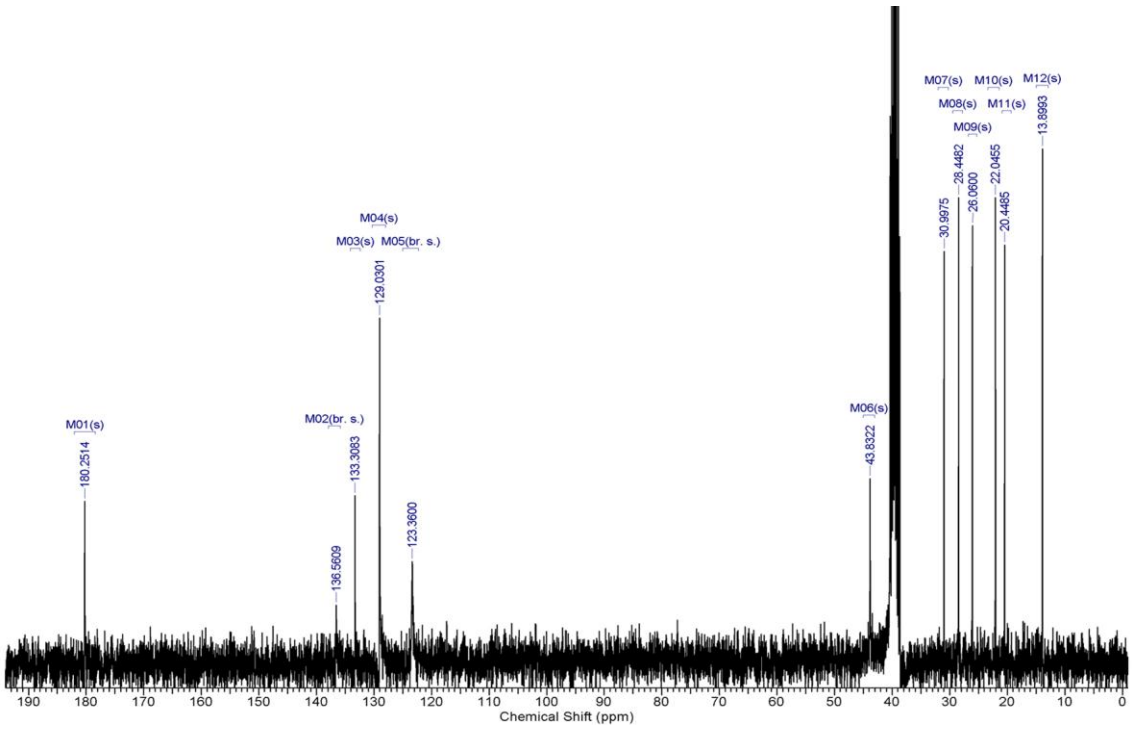


Figure S35. ^{13}C NMR spectrum of compound **18** (Me) in DMSO-d_6 at 298K.

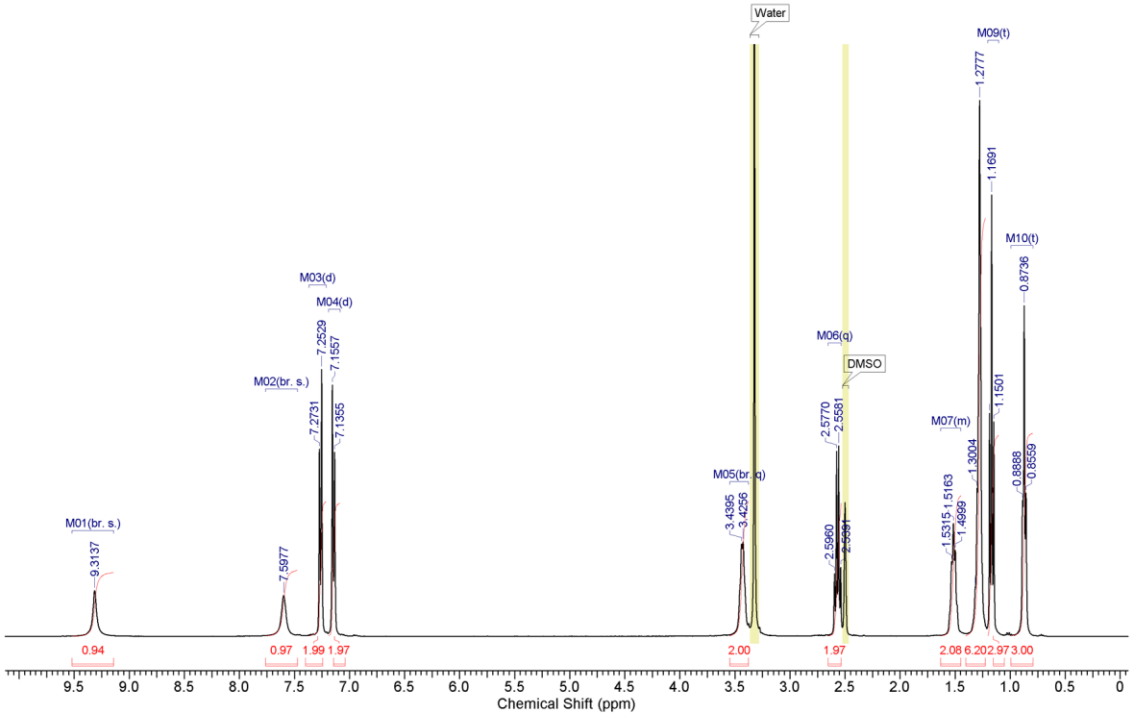


Figure S36. ^{1}H NMR spectrum of compound **19** (Et) in DMSO-d_6 at 298K.

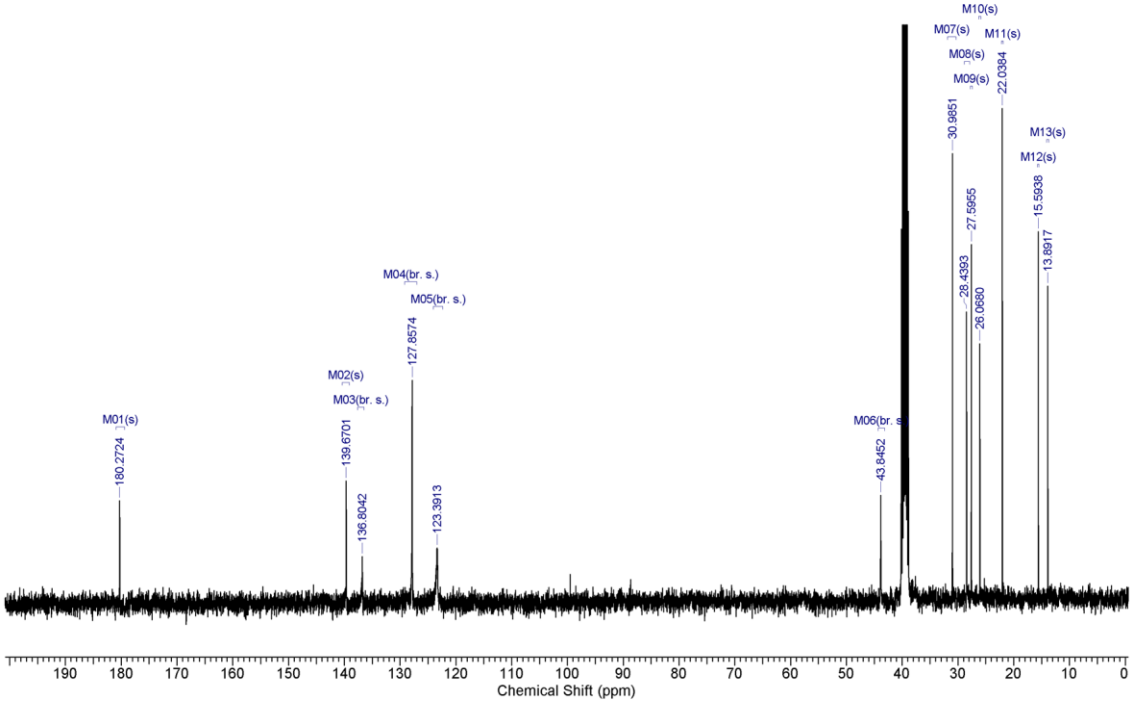


Figure S37. ^{13}C NMR spectrum of compound **19** (Et) in DMSO-d_6 at 298K.

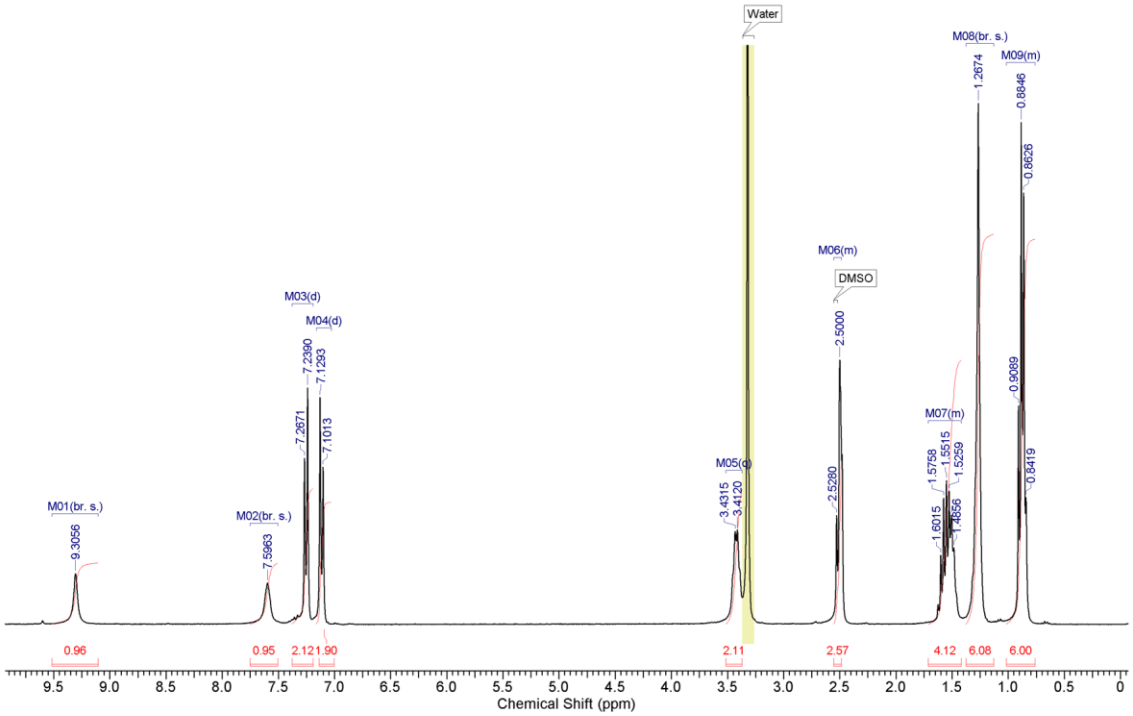


Figure S38. ^{1}H NMR spectrum of compound **20** (Pr) in DMSO-d_6 at 298K.

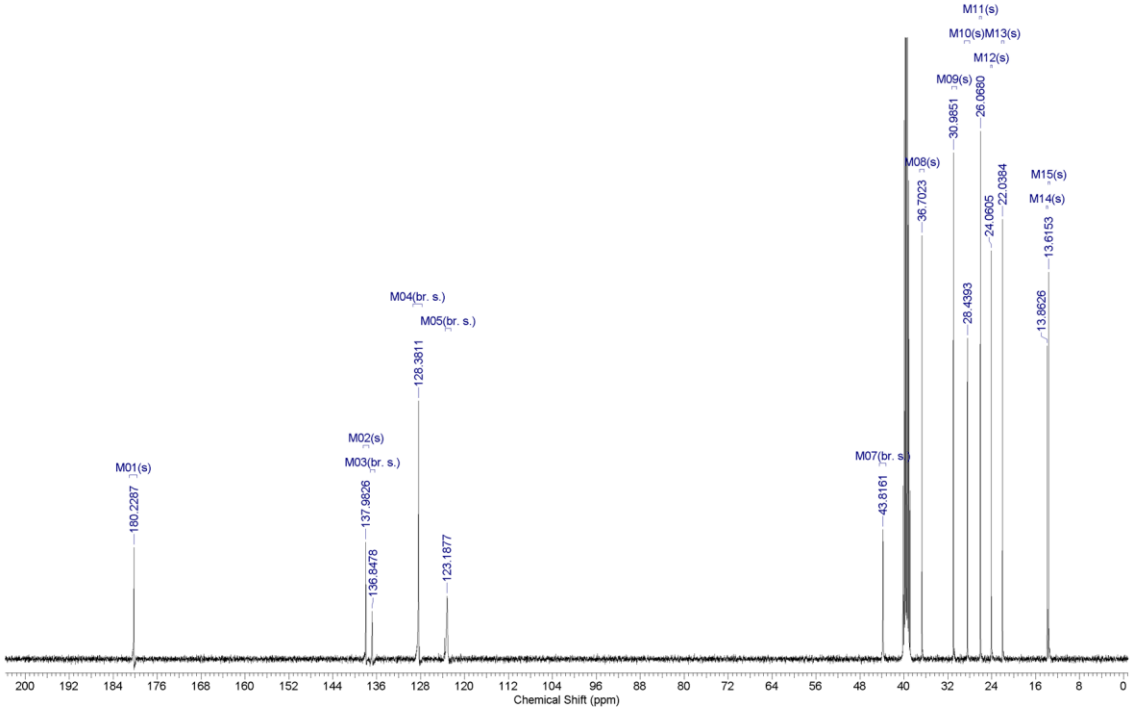


Figure S39. ^{13}C NMR spectrum of compound **20** (Pr) in DMSO-d_6 at 298K.

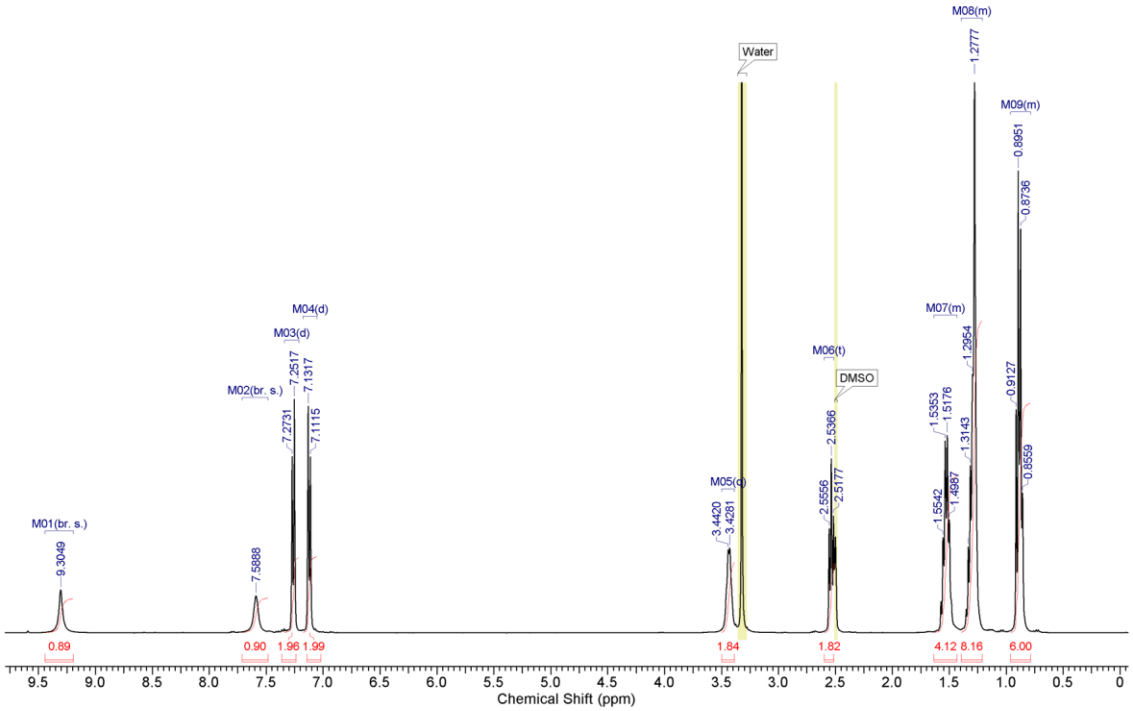


Figure S40. ^{1}H NMR spectrum of compound **21** (Bu) in DMSO-d_6 at 298K.

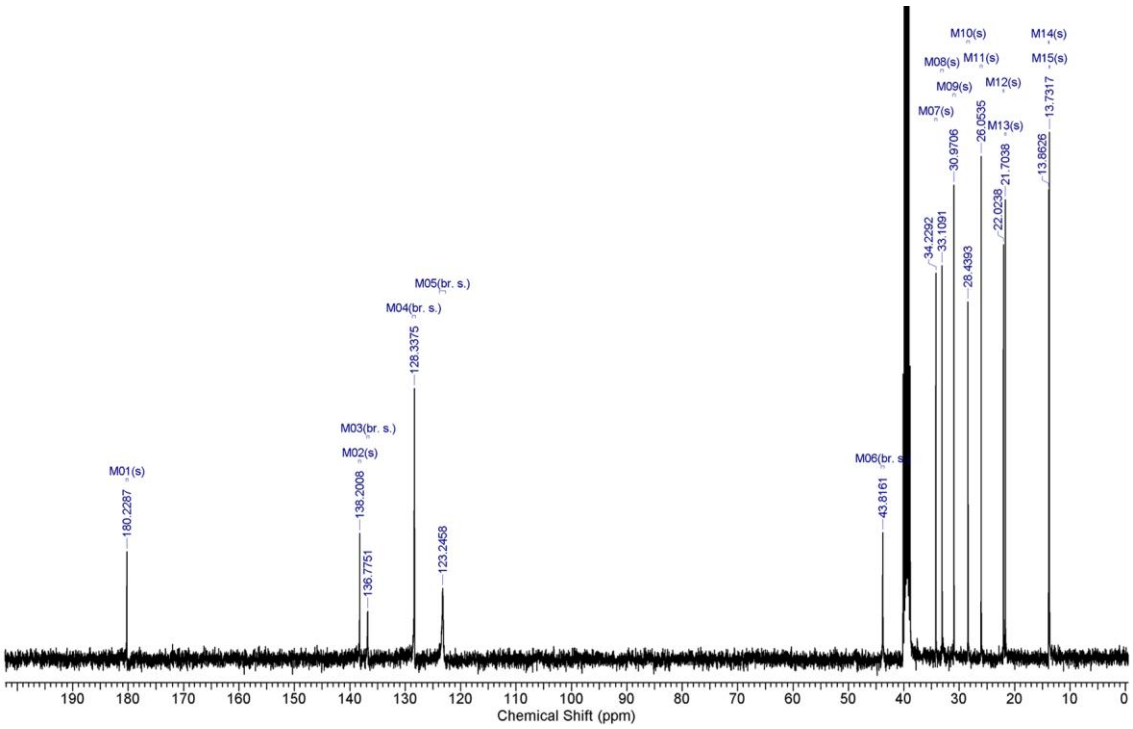


Figure S41. ^{13}C NMR spectrum of compound **21** (Bu) in DMSO-d_6 at 298K.

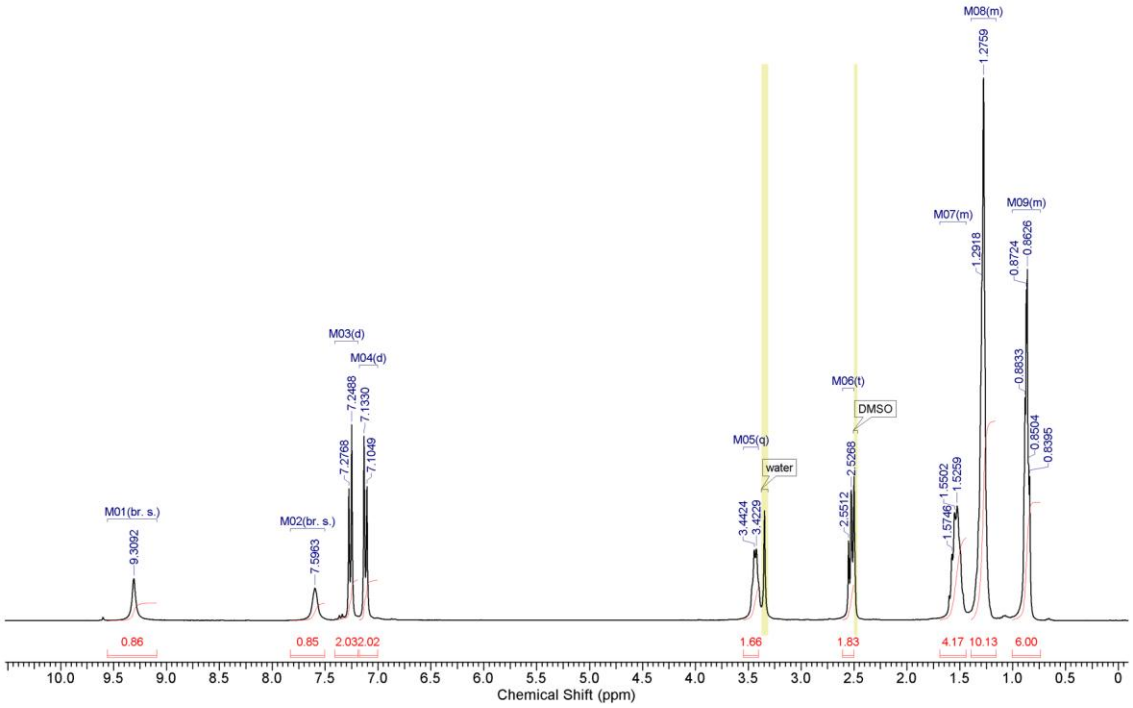


Figure S42. ^{1}H NMR spectrum of compound **22** (Pe) in DMSO-d_6 at 298K.

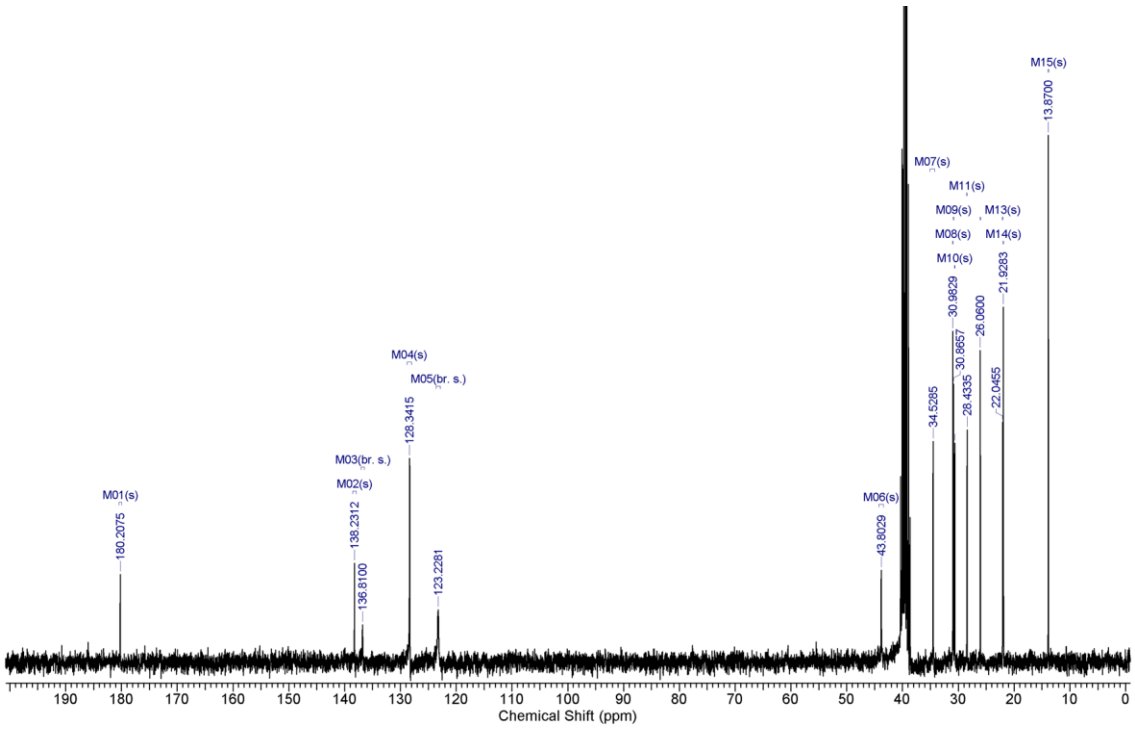


Figure S43. ^{13}C NMR spectrum of compound **22** (Pe) in DMSO-d_6 at 298K.

S4. SINGLE CRYSTAL X-RAY DIFFRACTION

Data were collected on various diffractometers at the University of Southampton. The crystal structure of compound **9** has been previously reported.[2] The structures were solved using SHELXS-97 (G. M. Sheldrick, Acta Cryst. (2008) A**64** 112-122)[7] and refined on *F2* by the full-matrix least-squares technique using the SHELXL-97 program package (G. M. Sheldrick (1997), University of Göttingen, Germany)[8]. Graphics are generated using ORTEP-III, MERCURY 3.0 or ViewerLite and Pov-Ray. In all cases the non-hydrogen atoms are refined anisotropically till convergence. Hydrogen atoms were stereochemically fixed at idealized positions and then refined isotropically. Hydrogen bonds are calculated using HTAB command in SHELXL-97. Structures were deposited with the Cambridge Crystallographic Database Centre (CCDC).

For **2** (-CF_3, CCDC 927450), **3** (-Cl, CCDC 927445), **4** (-CN, CCDC 927447), **5** (-$COCF_3$, CCDC 927456), **10** (-I, CCDC 927455), **11** (-NO_2, CCDC 927446), **12** (-OCOMe, CCDC 927453), **14** (-OEt, CCDC 927454) and **18** (-Me, CCDC 927448): **Diffractometer:** *Rigaku AFC12* goniometer equipped with an enhanced sensitivity (HG) *Saturn724+* detector mounted at the window of an *FR-E+ SuperBright* molybdenum rotating anode generator with VHF *Varimax* optics (70 μm focus). **Cell determination, Data collection, Data reduction and cell refinement & Absorption correction**: CrystalClear-SM Expert 2.0 r7 (Rigaku, 2011).

For **1** (-Br, CCDC 927451), **6** (-COMe, CCDC 927460), **7** (-COOMe, CCDC 927449), **8** (-F, CCDC 927452) and **20** (-Pr, CCDC 927459): **Diffractometer:** *Rigaku AFC12* goniometer equipped with an enhanced sensitivity (HG) *Saturn724+* detector mounted at the window of an *FR-E+ SuperBright* molybdenum rotating anode generator with HF *Varimax* optics (100 μm focus). **Cell determination, Data collection, Data reduction and cell refinement & Absorption correction**: CrystalClear-SM Expert 2.0 r7 (Rigaku, 2011).

For **15** (-OMe, CCDC 927457) and **19** (-Et, CCDC 927458): **Diffractometer:** Beamline I19 situated on an undulator insertion device with a combination of double crystal monochromator, vertical and horizontal focussing mirrors and a series of beam slits (primary white beam and either side of the focussing mirrors). The experimental hutch (EH1) is equipped with a Crystal Logic 4-circle kappa geometry goniometer with a Rigaku Saturn 724 CCD detector and an Oxford Cryosystems Cryostream plus cryostat (80-500K). For conventional service crystallography the beamline operates at a typical energy of 18 keV (Zr K absorption edge) and a Rigaku ACTOR robotic sample changing system is available. **Cell determination, Data collection, Data reduction and cell refinement & Absorption correction**: CrystalClear-SM Expert 2.0 r7 (Rigaku, 2011).

S4.1 X-ray data for compound 1 (-Br), CCDC 927451

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **1** in 1:1 chloroform:hexane. Crystal data for compound **1**: $C_{13}H_{19}BrN_2S$, M_r = 315.27 g/mol, crystal size = 0.26 x 0.03 x 0.01 mm , colourless needle, monoclinic, space group *P21/c*, a = 17.3563(8) Å, b = 7.8428(3) Å, c = 21.0265(15) Å, α = 90 º, β = 97.599(7) º, γ = 90 º, V = 2837.0(3) Å^3, Z = 8, ρ_c = 1.476 g cm^{-3}, μ = 3.026 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.48, reflections collected: 26506, independent reflections: 6491 (R_{int} = 0.0555), 307 parameters, R indices (all data): R_1 = 0.0523, wR_2 = 0.0739, final R indices [$I > 2\sigma I$]: R_1 = 0.0328, wR_2 = 0.0662, *GOOF* = 1.068, largest diff. peak and hole = 0.397 and -0.502 e Å .

Table S1. Hydrogen bond properties for **1**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N2-H2···S2	0.88	2.48	3.3124(17)	158.3
N3-H3···S1	0.88	2.58	3.3703(18)	150.0
N4-H4···S1	0.88	2.61	3.3423(16)	141.1
N1-H1···S2	0.88	2.75	3.5491(18)	152.0

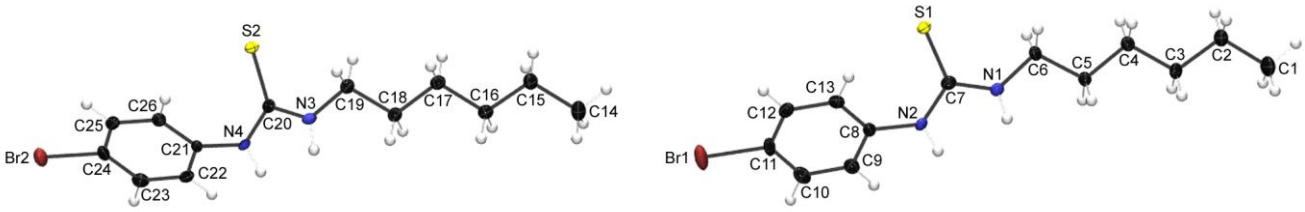


Figure S44. ORTEP diagram of **1** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

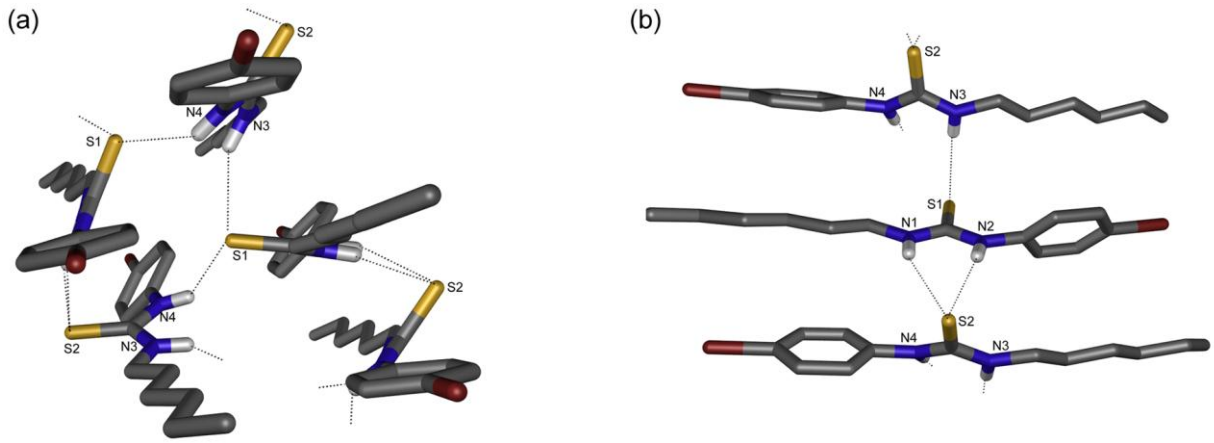


Figure S45. Schematic representation of the intermolecular hydrogen bonds in the crystal of **1** (two different views (a) and (b)). For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.2 X-ray data for compound 2 (-CF_3), CCDC 927450

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **2** in 1:1 chloroform:hexane. Crystal data for compound **2**: $C_{14}H_{19}F_3N_2S$, M_r = 304.37 g/mol, crystal size = 0.39 x 0.03 x 0.03 mm , colourless needle, monoclinic, space group *P21/c*, a = 13.298(8) Å, b = 13.278(6) Å, c = 8.673(4) Å, α = 90 º, β = 91.912(18) º, γ = 90 º, V = 1530.5(14) Å^3, Z = 4, ρ_c = 1.321 g cm^{-3}, μ = 0.235 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.47, reflections collected: 8095, independent reflections: 3436 (R_{int} = 0.0772), 181 parameters, R indices (all data): R_1 = 0.0888, wR_2 = 0.1325, final R indices [$I > 2\sigma I$]: R_1 = 0.0549, wR_2 = 0.1192, *GOOF* = 1.053, largest diff. peak and hole = 0.590 and -0.452 e Å .

Table S2. Hydrogen bond properties for **2**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.61	3.465(2)	164.8
N2-H2···S1	0.88	2.92	3.714(2)	150.7

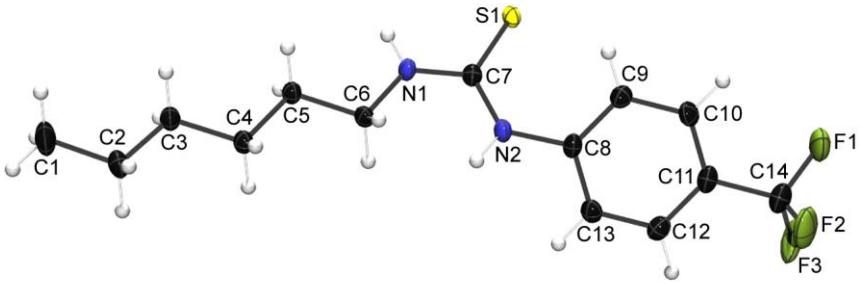


Figure S46. ORTEP diagram of **2** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

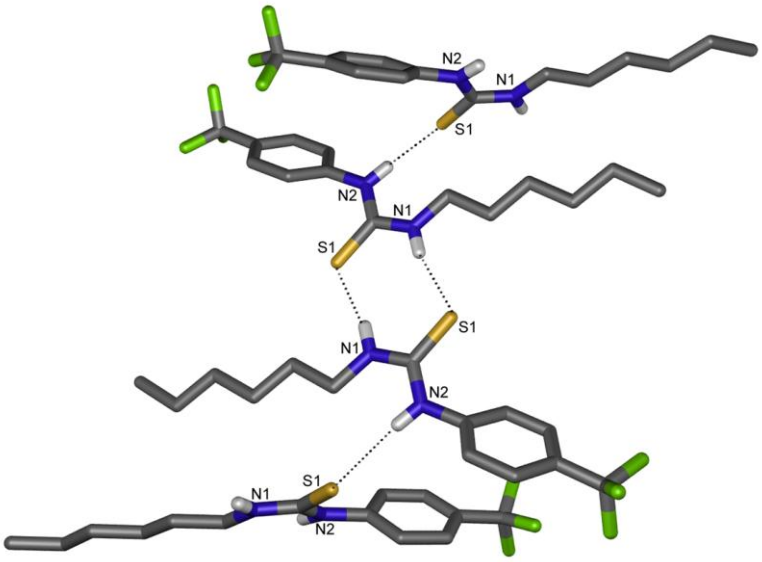


Figure S47. Schematic representation of the intermolecular hydrogen bonds in the crystal of **2**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.3 X-ray data for compound 3 (-Cl), CCDC 927445

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **3** in 1:1 chloroform:hexane. Crystal data for compound **3**: $C_{13}H_{19}ClN_2S$, M_r = 270.82 g/mol, crystal size = 0.70 x 0.12 x 0.02 mm , colourless needle, orthorhombic, space group *Pbcn*, a = 13.987(7) Å, b = 8.170(4) Å, c = 24.959(12) Å, α = 90 º, β = 90 º, γ = 90 º, V = 2852(2) Å^3, Z = 8, ρ_c = 1.262 g cm^{-3}, μ = 0.396 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 13636, independent reflections: 2526 (R_{int} = 0.0758), 176 parameters, R indices (all data): R_1 = 0.1033, wR_2 = 0.1342, final R indices [$I > 2\sigma I$]: R_1 = 0.0914, wR_2 = 0.1303, *GOOF* = 1.324, largest diff. peak and hole = 0.271 and -0.266 e Å . The hexyl chain was found to be disorder and was modeled with 2 separate orientations (occupancies of 1/3 and 2/3), thermal parameter constraints and geometrical restraints were applied (9 restraints).

Table S3. Hydrogen bond properties for **3**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.65	3.500(4)	163.1
N2-H2···S1	0.88	2.61	3.442(4)	158.9

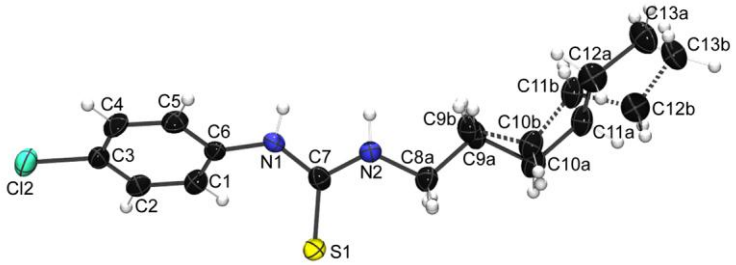


Figure S48. ORTEP diagram of **3** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

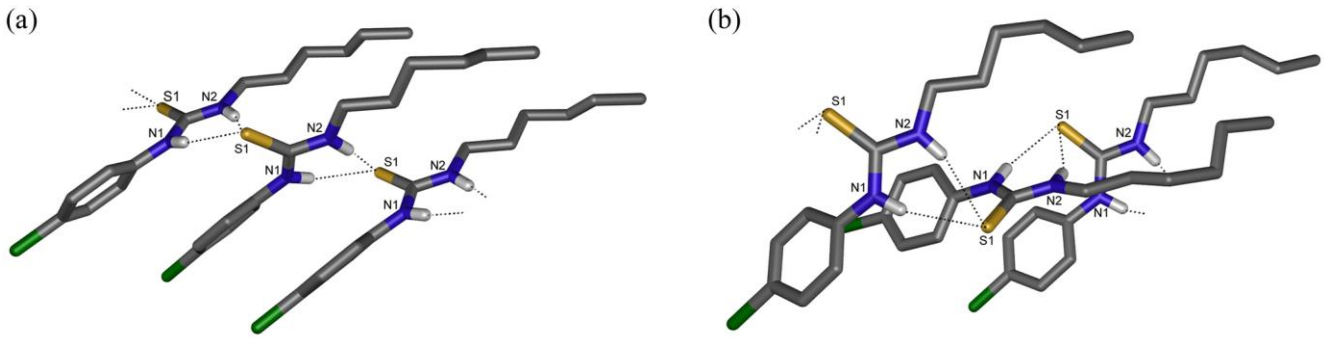


Figure S49. Schematic representation of the intermolecular hydrogen bonds in the crystal of **3** (two different views (a) and (b)). For clarity, only atoms involved in hydrogen bonding are labeled and disorder is omitted. Hydrogen bonds are represented by dashed lines.

S4.4 X-ray data for compound 4 (-CN), CCDC 927447

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **4** in 1:1 dichloromethane:hexane. Crystal data for compound **4**: $C_{14}H_{19}N_3S$, M_r = 261.38 g/mol, crystal size = 0.18 x 0.08 x 0.01 mm , colourless platelet, triclinic, space group *P-1*, a = 7.466(3) Å, b = 8.049(3) Å, c = 13.472(5) Å, α = 82.276(9) º, β = 82.154(10) º, γ = 63.734(8) º, V = 716.7(5) Å^3, Z = 2, ρ_c = 1.211 g cm^{-3}, μ = 0.213 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.46, reflections collected: 6649, independent reflections: 3271 (R_{int} = 0.0820), 179 parameters, R indices (all data): R_1 = 0.1143, wR_2 = 0.2781, final R indices [$I > 2\sigma I$]: R_1 = 0.1018, wR_2 = 0.2625, *GOOF* = 1.064, largest diff. peak and hole = 0.929 and -1.303 e Å . The hexyl chain was found to be disordered and was modeled with two separate orientations (occupancies of 1/3 and 2/3), thermal parameter and geometrical restraints were applied (195 restraints).

Table S4. Hydrogen bond properties for **4**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···N3	0.88	2.27	3.029(3)	144.0
N2-H2···S1	0.88	2.52	3.327(3)	152.0

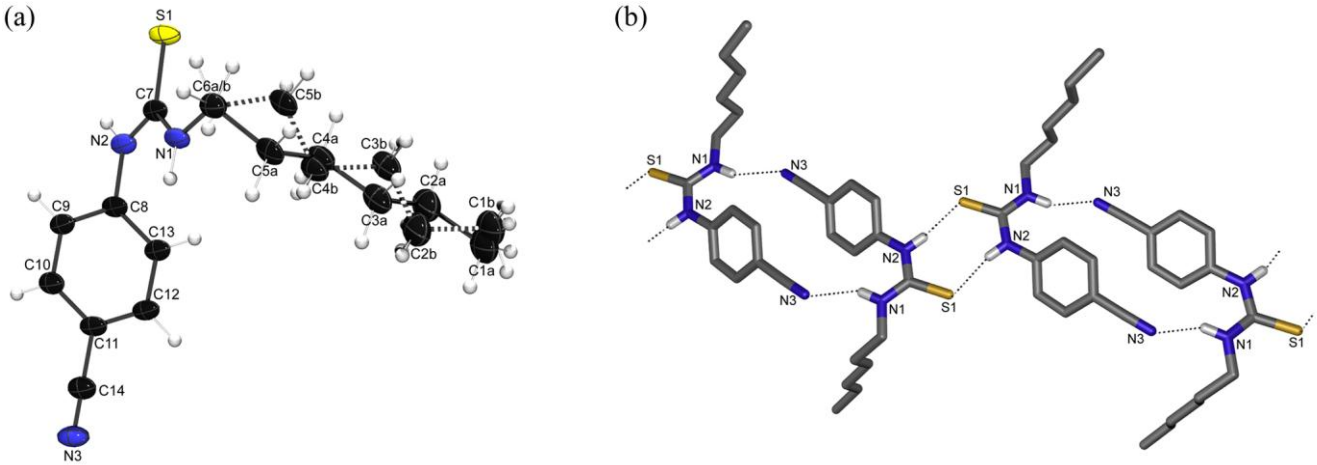


Figure S50. (a) ORTEP diagram of **4** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Schematic representation of the intermolecular hydrogen bonds in the crystal of **4**. For clarity, only atoms involved in hydrogen bonding are labeled and disorder is omitted. Hydrogen bonds are represented by dashed lines.

S4.5 X-ray data for compound 5 (-$COCF_3$), CCDC 927456

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **5** in 1:1 chloroform:hexane. Crystal data for compound **5**: $C_{15}H_{19}F_3N_2OS$, M_r = 332.38 g/mol, crystal size = 0.27 x 0.10 x 0.02 mm , colourless plate, triclinic, space group *P-1*, a = 5.2392(4) Å, b = 11.0200(9) Å, c = 14.5788(11) Å, α = 75.946(5) º, β = 79.884(6) º, γ = 79.501(6) º, V = 795.20(11) Å^3, Z = 2, ρ_c = 1.388 g cm^{-3}, μ = 0.237 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.48, reflections collected: 7651, independent reflections: 3632 (R_{int} = 0.0283), 199 parameters, R indices (all data): R_1 = 0.0376, wR_2 = 0.0952, final R indices [$I > 2\sigma I$]: R_1 = 0.0331, wR_2 = 0.0878, *GOOF* = 1.086, largest diff. peak and hole = 0.387 and -0.219 e Å .

Table S5. Hydrogen bond properties for **5**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.48	3.3380(11)	165.4
N2-H2···O1	0.88	2.40	3.0444(15)	130.4

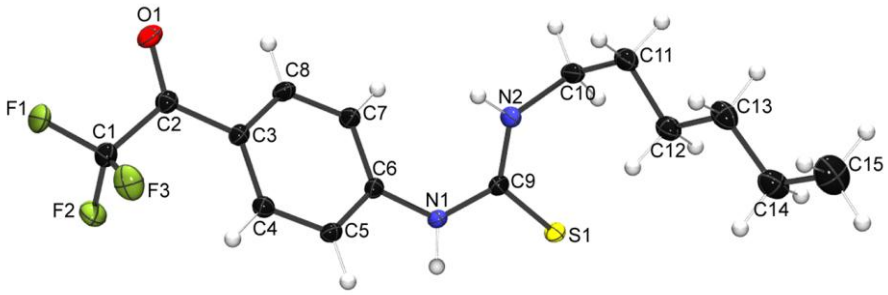


Figure S51. ORTEP diagram of **5** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

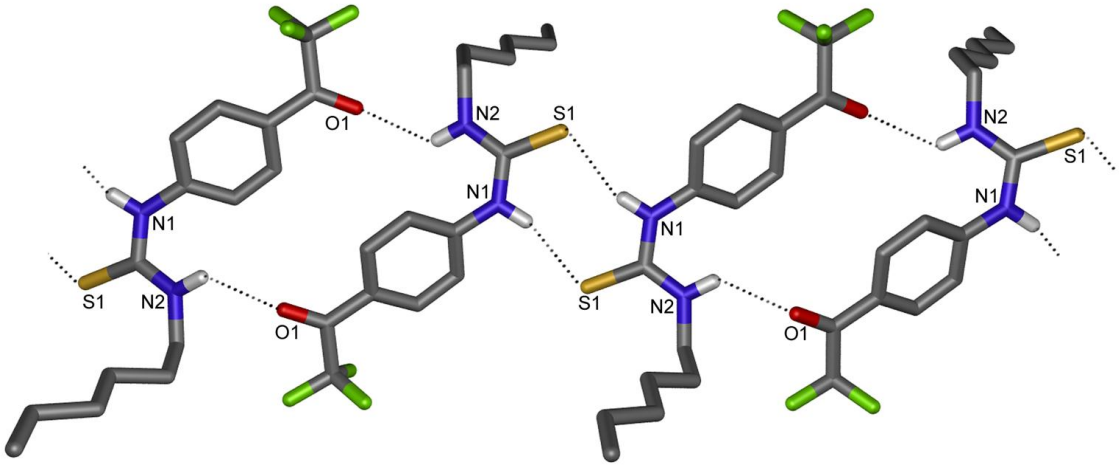


Figure S52. Schematic representation of the intermolecular hydrogen bonds in the crystal of **5**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.6 X-ray data for compound 6 (-COMe), CCDC 927460

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **6** in 1:1 methanol:hexane. Crystal data for compound **6**: $C_{15}H_{22}N_2OS$, M_r = 278.42 g/mol, crystal size = 0.20 x 0.20 x 0.20 mm , colourless needle (fragment), orthorhombic, space group *Pbca*, a = 7.756(6) Å, b = 10.120(7) Å, c = 37.38(3) Å, α = 90 º, β =90 º, γ = 90 º, V = 2934(4) Å^3, Z = 8, ρ_c = 1.261 g cm^{-3}, μ = 0.215 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.68, reflections collected: 12353, independent reflections: 2507 (R_{int} = 0.0793), 172 parameters, R indices (all data): R_1 = 0.1142, wR_2 = 0.2504, final R indices [$I > 2\sigma I$]: R_1 = 0.0970, wR_2 = 0.2373, *GOOF* = 1.111, largest diff. peak and hole = 0.384 and -0.331 e Å . Data completeness is only 89.8%, but this is sufficient for proof of structure and hydrogen bonding network.

Table S6. Hydrogen bond properties for **6**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···O1	0.88	2.50	3.125(6)	129.1
N1-H1···S1	0.88	2.67	3.401(4)	141.2
N2-H2···S1	0.88	2.78	3.576(4)	151.3

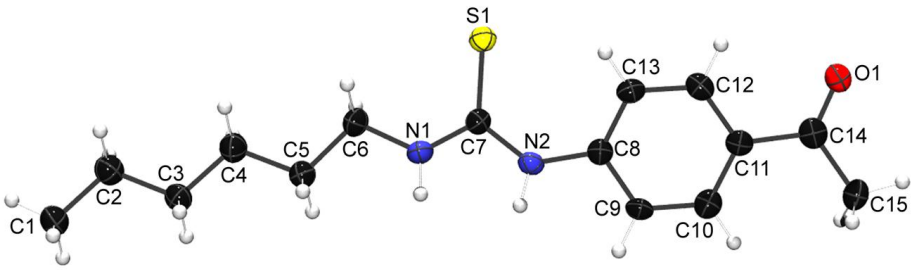


Figure S53. ORTEP diagram of **6** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

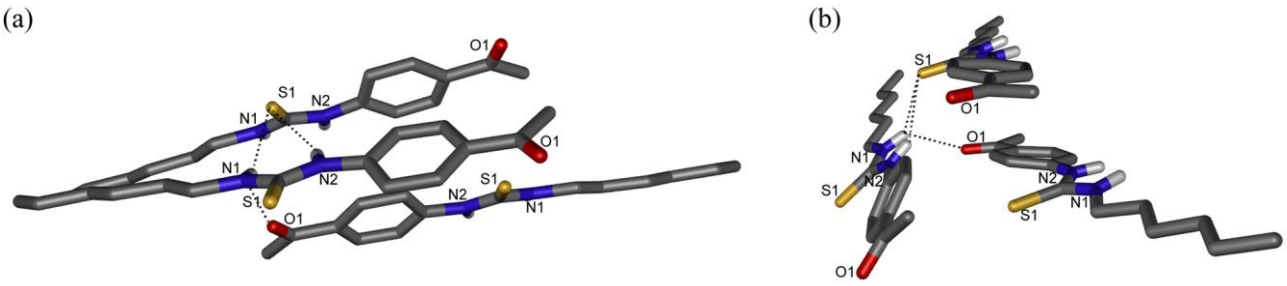


Figure S54. Schematic representation of the intermolecular hydrogen bonds in the crystal of **6** (two different views (a) and (b)). For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.7 X-ray data for compound 7 (-COOMe), CCDC 927449

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **7** in 1:1 chloroform:hexane. Crystal data for compound **7**: $C_{15}H_{22}N_2O_2S$, M_r = 294.41 g/mol, crystal size = 0.04 x 0.04 x 0.01 mm , colourless plate, monoclinic, space group *P21/c*, a = 11.0875(8) Å, b = 7.4703(6) Å, c = 37.579(3) Å, α = 90 º, β = 100.356(7) º, γ = 90 º, V = 3061.8(4) Å^3, Z = 8, ρ_c = 1.277 g cm^{-3}, μ = 0.215 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 10229, independent reflections: 5045 (R_{int} = 0.0867), 365 parameters, R indices (all data): R_1 = 0.1527, wR_2 = 0.2537, final R indices [$I > 2\sigma I$]: R_1 = 0.0864, wR_2 = 0.2149, *GOOF* = 1.042, largest diff. peak and hole = 0.821 and -0.515 e Å . Data completeness is only 93.5%, but this is sufficient for proof of structure and hydrogen bonding network.

Table S7. Hydrogen bond properties for **7**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.48	3.321(5)	159.9
N2-H2···S1	0.88	2.60	3.440(5)	159.9
N3-H3···S2	0.88	2.50	3.334(5)	157.7
N4-H4A···S2	0.88	2.63	3.462(5)	159.1

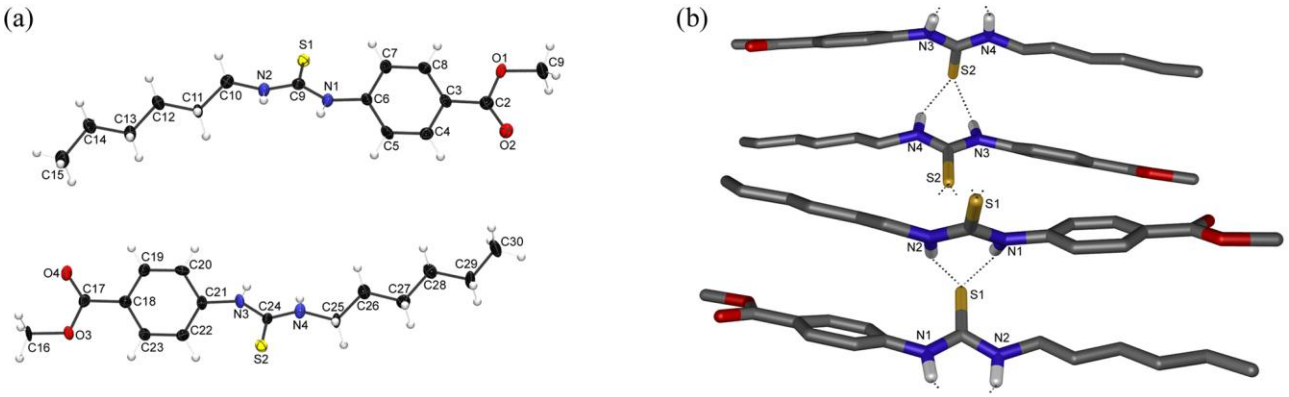


Figure S55. (a) ORTEP diagram of **7** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Schematic representation of the intermolecular hydrogen bonds in the crystal of **7**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.8 X-ray data for compound 8 (-F), CCDC 927452

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **8** in 1:1 chloroform:hexane. Crystal data for compound **8**: $C_{13}H_{19}FN_2S$, M_r = 254.36 g/mol, crystal size = 0.55 x 0.50 x 0.01 mm , colourless plate, orthorhombic, space group *Pca21*, a = 17.405(12) Å, b = 8.732(6) Å, c = 17.612(12) Å, α = 90 º, β =90 º, γ = 90 º, V = 2934(4) Å^3, Z = 8, ρ_c = 1.262 g cm^{-3}, μ = 0.234 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 19097, independent reflections: 4720 (R_{int} = 0.1075), 307 parameters, R indices (all data): R_1 = 0.0822, wR_2 = 0.1367, final R indices [$I > 2\sigma I$]: R_1 = 0.0693, wR_2 = 0.1303, *GOOF* = 1.099, largest diff. peak and hole = 0.416 and -0.293 e Å .

Table S8. Hydrogen bond properties for **8**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H2···S2	0.86	2.49	3.321(4)	161.5
N1-H2···S2	0.86	2.49	3.321(4)	161.5
N3-H3B···S1	0.88	2.56	3.383(4)	155.4
N4-H4···S1	0.88	2.56	3.373(4)	154.3

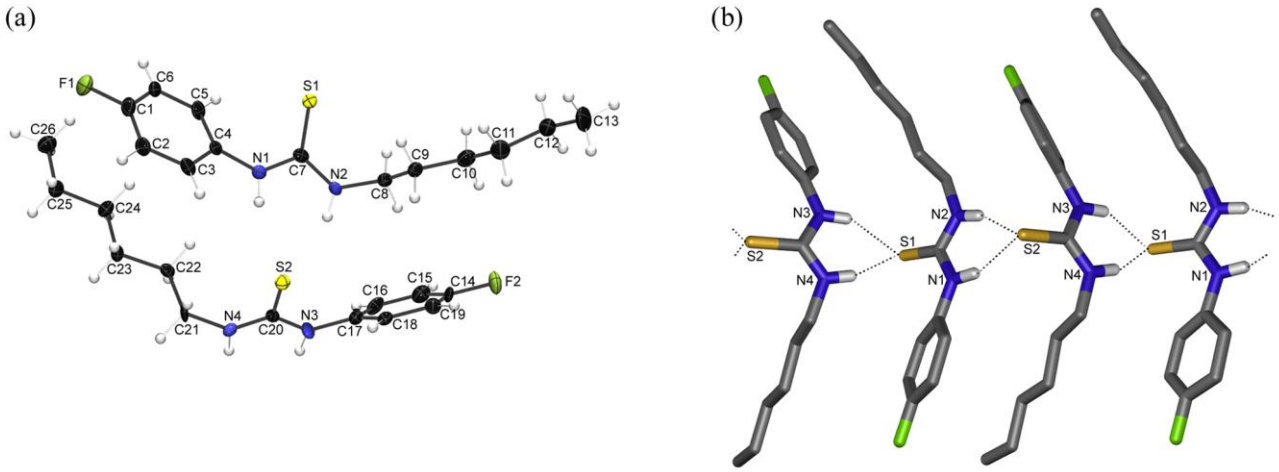


Figure S56. (a) ORTEP diagram of **8** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Schematic representation of the intermolecular hydrogen bonds in the crystal of **8**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.9 X-ray data for compound 10 (-I), CCDC 927455

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **10** in 1:1 methanol:hexane. Crystal data for compound **10**: $C_{13}H_{19}IN_2S$, M_r = 362.26 g/mol, crystal size = 0.10 x 0.01 x 0.01 mm , colourless needle, monoclinic, space group *P21/n*, a = 11.4798(8) Å, b = 8.4051(6) Å, c = 15.6294(11) Å, α = 90 º, β = 102.550(7) º, γ = 90 º, V = 1472.03(18) Å^3, Z = 4, ρ_c = 1.635 g cm^{-3}, μ = 2.300 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.48, reflections collected: 13677, independent reflections: 3362 (R_{int} = 0.0483), 154 parameters, R indices (all data): R_1 = 0.0327, wR_2 = 0.0562, final R indices [$I > 2\sigma I$]: R_1 = 0.0250, wR_2 = 0.0541, *GOOF* = 1.021, largest diff. peak and hole = 0.488 and -0.571 e Å .

Table S9. Hydrogen bond properties for **10**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.49	3.328(2)	159.1
N2-H2···S1	0.88	2.54	3.380(2)	159.5

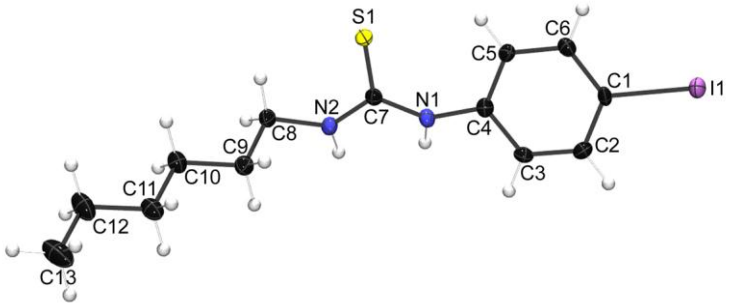


Figure S57. ORTEP diagram of **10** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

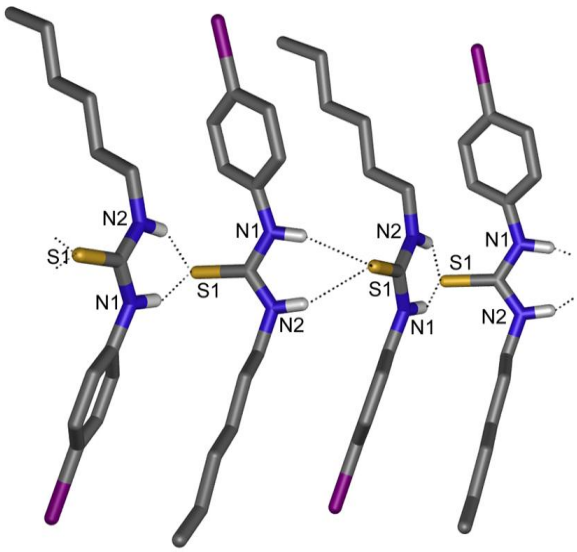


Figure S58. Schematic representation of the intermolecular hydrogen bonds in the crystal of **10**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.10 X-ray data for compound 11 (-NO_2), CCDC 927446

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **11** in 1:1 dichloromethane:hexane. Crystal data for compound **11**: $C_{13}H_{19}N_3O_2S$, M_r = 281.37 g/mol, crystal size = 0.20 x 0.08 x 0.03 mm , yellow block, triclinic, space group *P-1*, a = 5.2833(5) Å, b = 14.9019(15) Å, c = 18.0119(18) Å, α = 91.957(7) º, β = 95.987(7) º, γ = 92.116(7) º, V = 1408.4(2) Å^3, Z = 4, ρ_c = 1.327 g cm^{-3}, μ = 0.232 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.02, reflections collected: 11029, independent reflections: 4960 (R_{int} = 0.0211), 343 parameters, R indices (all data): R_1 = 0.0359, wR_2 = 0.0874, final R indices [$I > 2\sigma I$]: R_1 = 0.0321, wR_2 = 0.0849, *GOOF* = 1.079, largest diff. peak and hole = 0.313 and -0.358 e Å .

Table S10. Hydrogen bond properties for **11**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.83	3.3296(13)	117.9
N2-H2···S1	0.88	2.56	3.4079(13)	161.4
N4-H4···O2	0.88	2.42	3.2625(17)	159.2
N4-H4···S2	0.88	2.90	3.4009(14)	117.9
N5-H5···O1	0.88	2.17	2.9811(16)	153.0

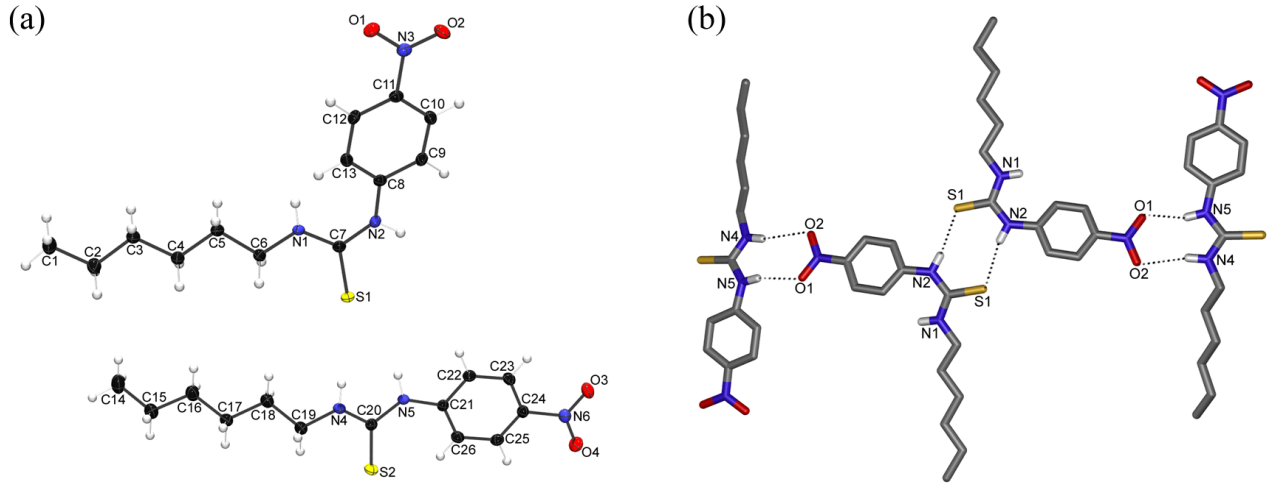


Figure S59. (a) ORTEP diagram of **11** with atom numbering, showing 50 % probability factor for the thermal ellipsoids. (b) Schematic representation of the intermolecular hydrogen bonds in the crystal of **11**.

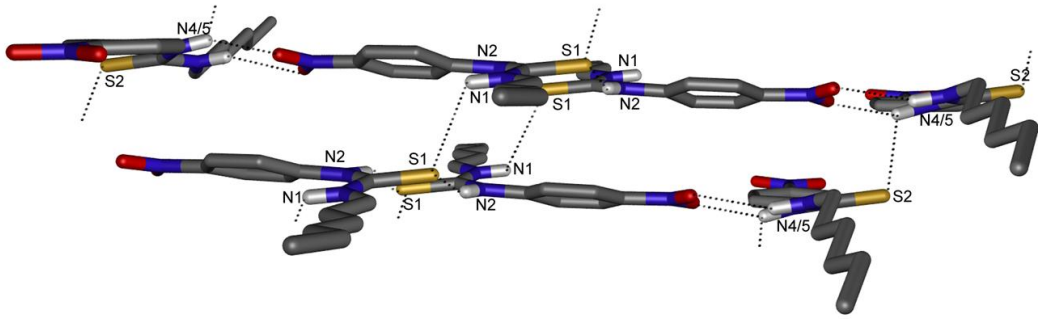


Figure S60. Second schematic representation of the intermolecular hydrogen bonds in the crystal of **11**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.11 X-ray data for compound 12 (-OCOMe), CCDC 927453

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **12** in 1:1 dichloromethane:hexane. Crystal data for compound **12**: $C_{15}H_{22}N_2O_2S$, M_r = 294.41 g/mol, crystal size = 0.11 x 0.10 x 0.02 mm , colourless platelet, monoclinic, space group *P21/n*, a = 5.553(4) Å, b = 9.157(8) Å, c = 30.91(3) Å, α = 90 º, β = 93.215(8) º, γ = 90 º, V = 1569(2) Å^3, Z = 4, ρ_c = 1.246 g cm^{-3}, μ = 0.210 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 27.49, reflections collected: 24561, independent reflections: 3586 (R_{int} = 0.0661), 181 parameters, R indices (all data): R_1 = 0.0848, wR_2 = 0.1805, final R indices [$I > 2\sigma I$]: R_1 = 0.0724, wR_2 = 0.1729, *GOOF* = 1.187, largest diff. peak and hole = 0.449 and -0.438 e Å .

Table S11. Hydrogen bond properties for **12**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···O2	0.88	2.08	2.928(3)	160.4
N2-H2···O2	0.88	2.35	3.120(4)	146.7

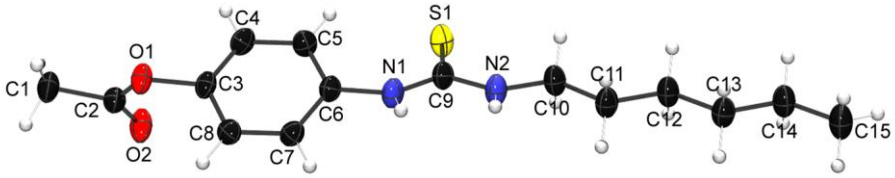


Figure S61. ORTEP diagram of **12** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

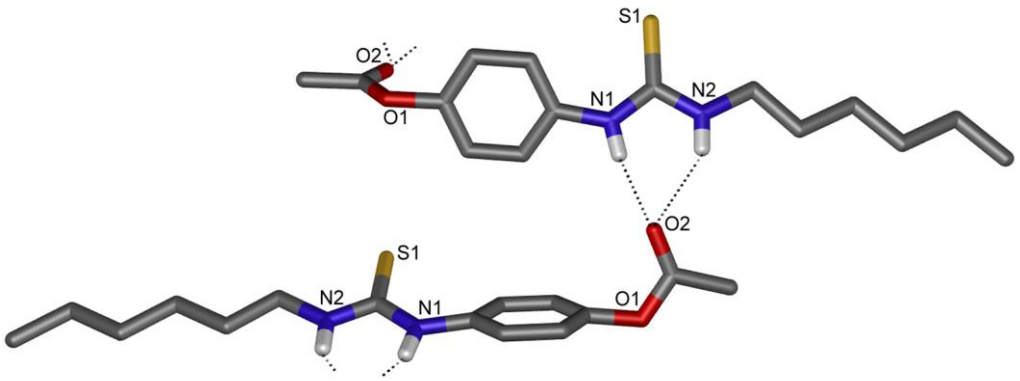


Figure S62. Schematic representation of the intermolecular hydrogen bonds in the crystal of **12**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.12 X-ray data for compound 14 (-OEt), CCDC 927454

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **14** in 1:1 chloroform:hexane. Crystal data for compound **14**: $C_{15}H_{24}N_2OS$, M_r = 280.42 g/mol, crystal size = 0.27 x 0.03 x 0.02 mm , colourless needle, triclinic, space group *P-1*, a = 8.39(1) Å, b = 9.689(11) Å, c = 19.62(2) Å, α = 93.19(2) º, β = 101.167(18) º, γ = 90.09(3) º, V = 1562(3) Å^3, Z = 4, ρ_c = 1.192 g cm^{-3}, μ = 0.203 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 12058, independent reflections: 5451 (R_{int} = 0.0876), 343 parameters, R indices (all data): R_1 = 0.1467, wR_2 = 0.2111, final R indices [$I > 2\sigma I$]: R_1 = 0.0970, wR_2 = 0.1868, *GOOF* = 1.151, largest diff. peak and hole = 0.375 and -0.346 e Å .

Table S12. Hydrogen bond properties for **14**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S2	0.87	2.54	3.390(5)	165.7
N2-H2···S2	0.88	2.67	3.502(5)	157.9
N3-H3···S1	0.88	2.55	3.397(5)	161.2
N4-H4···S1	0.88	2.62	3.453(5)	157.7

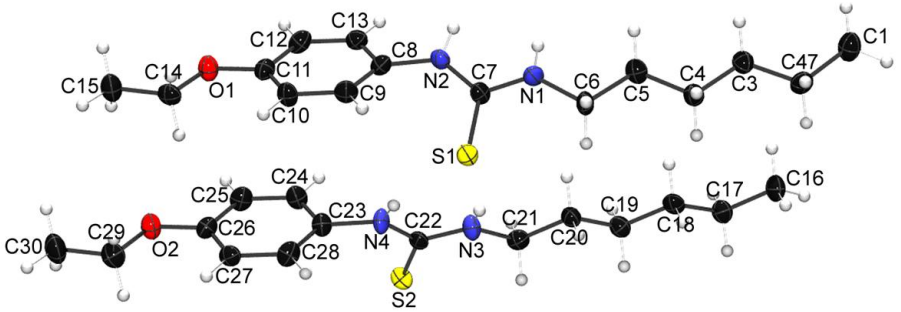


Figure S63. ORTEP diagram of **14** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

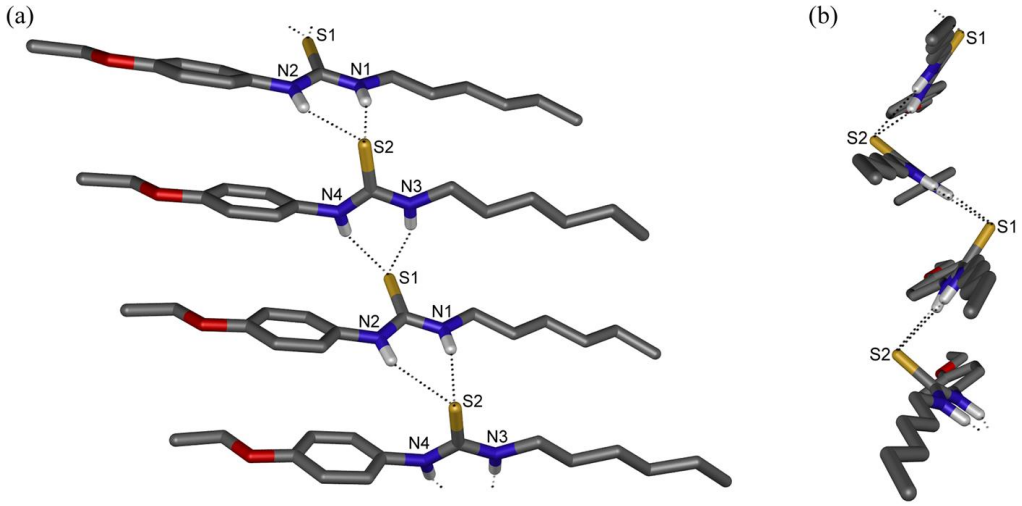


Figure S64. Schematic representation of the intermolecular hydrogen bonds in the crystal of **14** (two different views (a) and (b)). For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.13 X-ray data for compound 15 (-OMe), CCDC 927457

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **15** in 1:1 chloroform:hexane. Crystal data for compound **15**: $C_{14}H_{22}N_2OS$, M_r = 266.40 g/mol, crystal size = 0.12 x 0.08 x 0.01 mm , colourless plate, monoclinic, space group *P21/c*, a = 36.040(5) Å, b = 4.347(5) Å, c = 9.590(5) Å, α = 90 º, β = 95.300(5) º, γ = 90 º, V = 1496.0(19) Å^3, Z = 4, ρ_c = 1.183 g cm^{-3}, μ = 0.208 mm^{-1}, radiation and wavelength = synchrotron (0.68890 Å), T = 100(2) K, θ_{max} = 24.21, reflections collected: 8777, independent reflections: 2402 (R_{int} = 0.1318), 163 parameters, R indices (all data): R_1 = 0.1509, wR_2 = 0.2973, final R indices [$I > 2\sigma I$]: R_1 = 0.0993, wR_2 = 0.2443, *GOOF* = 0.953, largest diff. peak and hole = 1.085 and -0.782 e Å . Data completeness is only 90.8%, but this is sufficient for proof of structure and hydrogen bonding network.

Table S13. Hydrogen bond properties for **15**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.71	3.482(5)	147.6
N2-H2···S1	0.88	2.65	3.434(5)	149.4

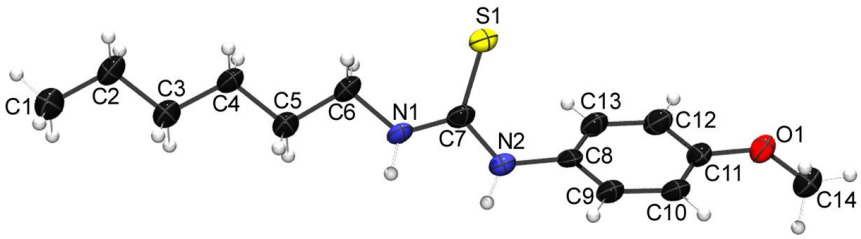


Figure S65. ORTEP diagram of **15** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

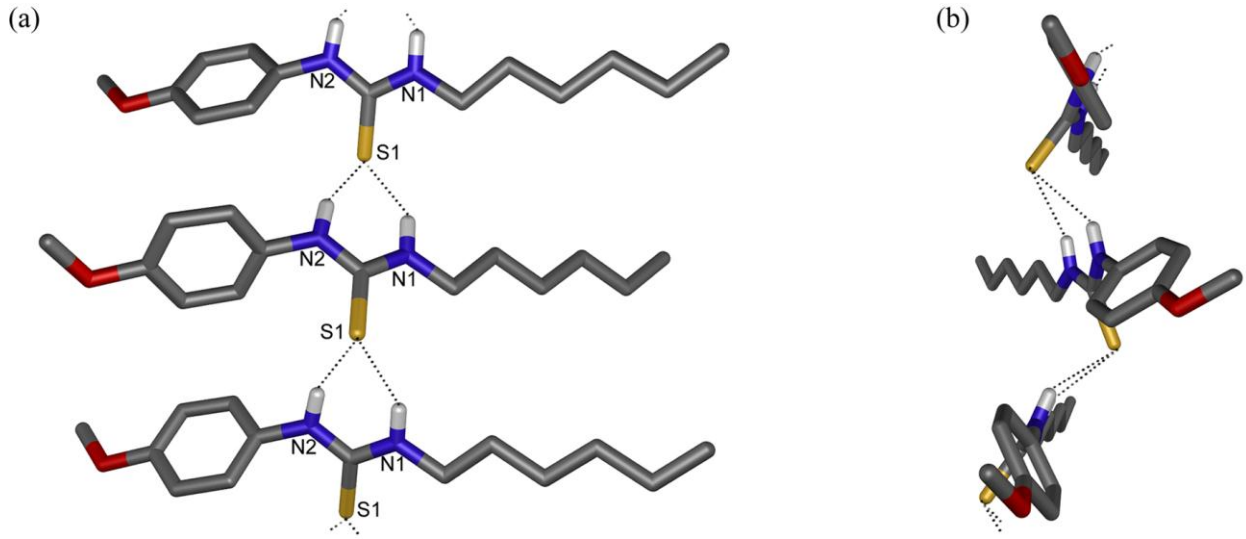


Figure S66. Schematic representation of the intermolecular hydrogen bonds in the crystal of **15** (two different views (a) and (b)). For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.14 X-ray data for compound 18 (-Me), CCDC 927448

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **18** in 1:1 chloroform:hexane. Crystal data for compound **18**: $C_{14}H_{22}N_2S$, M_r = 250.40 g/mol, crystal size = 0.32 x 0.02 x 0.01 mm , colourless needle, monoclinic, space group *P21/n*, a = 10.731(3) Å, b = 8.475(2) Å, c = 16.092(4) Å, α = 90 º, β = 102.020(7) º, γ = 90 º, V = 1431.5(6) Å^3, Z = 4, ρ_c = 1.162 g cm^{-3}, μ = 0.208 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 5504, independent reflections: 2520 (R_{int} = 0.0514), 154 parameters, R indices (all data): R_1 = 0.1123, wR_2 = 0.3108, final R indices [$I > 2\sigma I$]: R_1 = 0.0980, wR_2 = 0.3033, *GOOF* = 1.150, largest diff. peak and hole = 0.904 and -0.594 e Å .

Table S14. Hydrogen bond properties for **18**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.50	3.344(6)	160.2
N2-H2···S1	0.88	2.54	3.371(6)	156.9

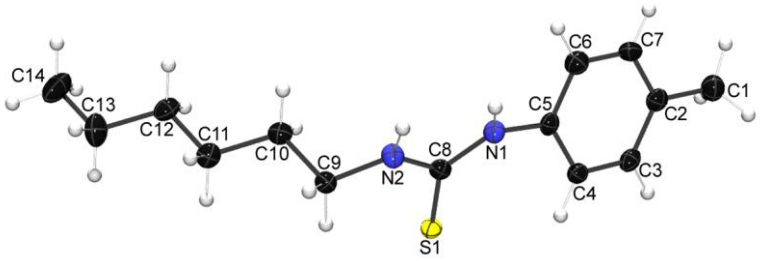


Figure S67. ORTEP diagram of **18** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

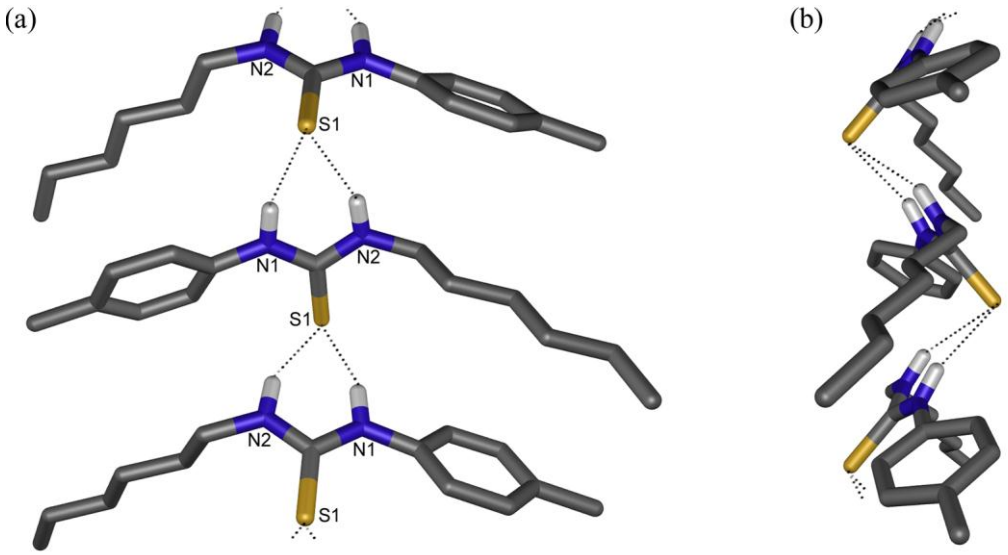


Figure S68. Schematic representation of the intermolecular hydrogen bonds in the crystal of **18**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.15 X-ray data for compound 19 (-Et), CCDC 927458

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **19** in diethylether. Crystal data for compound **19**: $C_{15}H_{24}N_2S$, M_r = 264.42 g/mol, crystal size = 0.03 x 0.03 x 0.01 mm , colourless blade, monoclinic, space group *P21/c*, a = 13.07(2) Å, b = 13.28(3) Å, c = 8.885(18) Å, α = 90 º, β = 93.56(2) º, γ = 90 º, V = 1539(5) Å^3, Z = 4, ρ_c = 1.141 g cm^{-3}, μ = 0.197 mm^{-1}, radiation and wavelength = synchrotron (0.68890 Å), T = 100(2) K, θ_{max} = 24.20, reflections collected: 10433, independent reflections: 2637 (R_{int} = 0.2114), 164 parameters, R indices (all data): R_1 = 0.1794, wR_2 = 0.3229, final R indices [$I > 2\sigma I$]: R_1 = 0.1039, wR_2 = 0.2559, *GOOF* = 1.075, largest diff. peak and hole = 0.432 and -0.614 e Å .

Table S15. Hydrogen bond properties for **19**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N2-H2···S001	0.88	2.57	3.437(7)	167.8

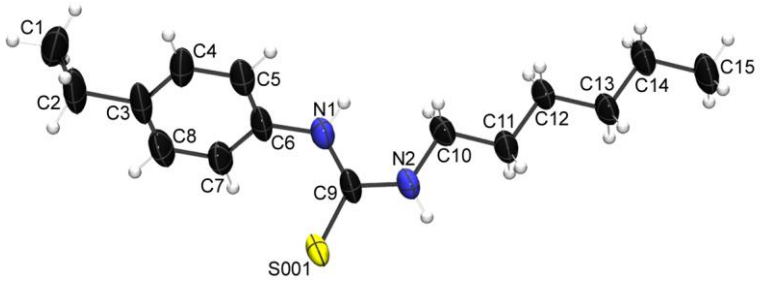


Figure S69. ORTEP diagram of **19** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

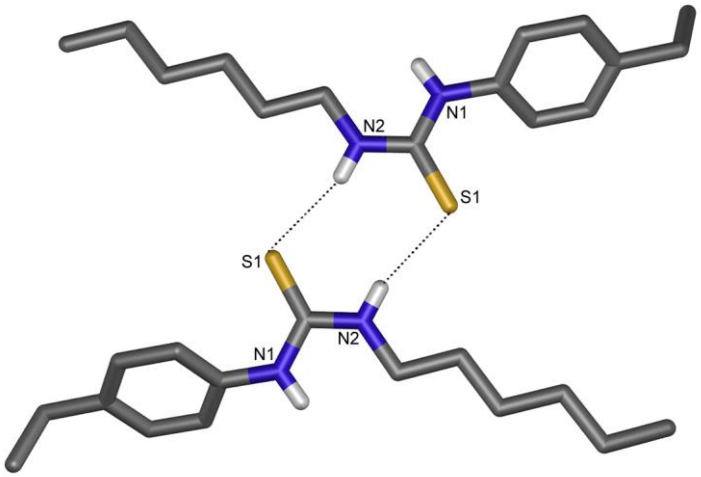


Figure S70. Schematic representation of the intermolecular hydrogen bonds in the crystal of **19**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S4.16 X-ray data for compound 20 (-Pr), CCDC 927459

Single crystals suitable for X-ray diffraction were obtained by slow evaporation of a solution of compound **20** in diethylether. Crystal data for compound **20**: $C_{16}H_{26}N_2S$, M_r = 278.45 g/mol, crystal size = 0.32 x 0.05 x 0.01 mm , colourless needle, monoclinic, space group *P21/c*, a = 22.536(5) Å, b = 7.7393(16) Å, c = 9.223(2) Å, α = 90 º, β = 98.488(7) º, γ = 90 º, V = 1591.0(6) Å^3, Z = 4, ρ_c = 1.162 g cm^{-3}, μ = 0.194 mm^{-1}, radiation and wavelength = MoKα (0.71075 Å), T = 100(2) K, θ_{max} = 25.03, reflections collected: 11058, independent reflections: 2808 (R_{int} = 0.1524), 174 parameters, R indices (all data): R_1 = 0.1934, wR_2 = 0.1688, final R indices [$I > 2\sigma I$]: R_1 = 0.0788, wR_2 = 0.1361, *GOOF* = 0.935, largest diff. peak and hole = 0.264 and -0.306 e Å .

Table S16. Hydrogen bond properties for **20**

Donor--H···Acceptor	D-H (Å)	H···A (Å)	D···A (Å)	D-H···A (°)
N1-H1···S1	0.88	2.83	3.595(4)	146.6
N2-H2···S1	0.88	2.54	3.346(4)	153.4

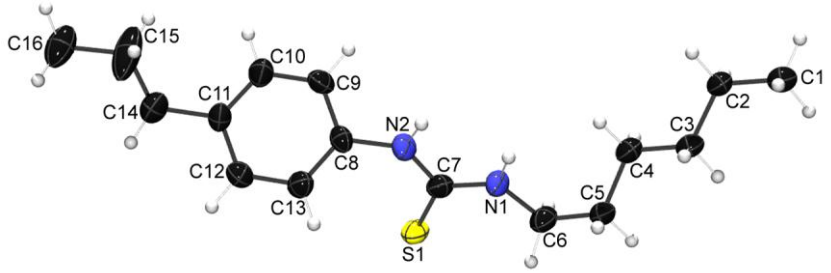


Figure S71. ORTEP diagram of **20** with atom numbering, showing 50 % probability factor for the thermal ellipsoids.

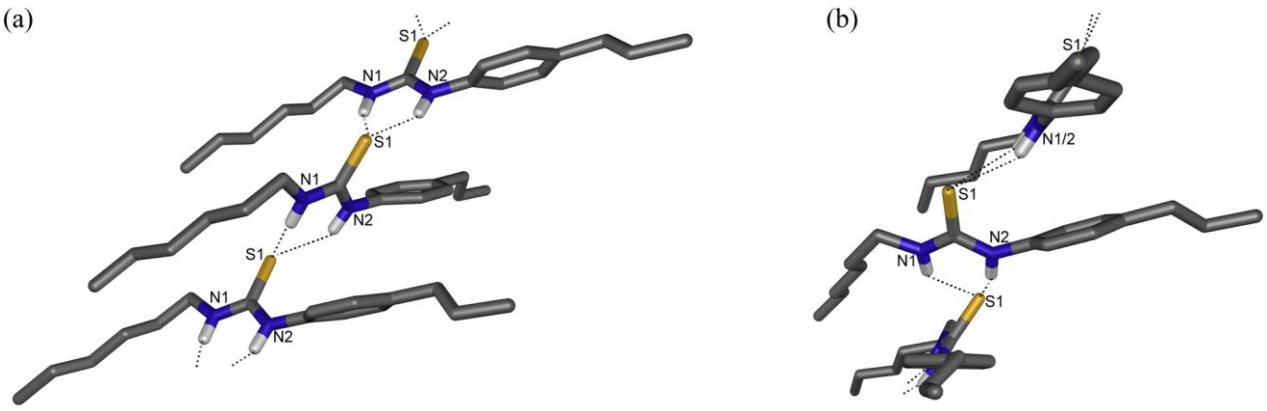


Figure S72. Schematic representation of the intermolecular hydrogen bonds in the crystal of **20**. For clarity, only atoms involved in hydrogen bonding are labeled. Hydrogen bonds are represented by dashed lines.

S5. NMR BINDING STUDIES

S5.1 Experimental procedure of binding studies

NMR titrations were performed by addition of aliquots of the putative anionic guest as the tetrabutylammonium (TBA) or tetraethylammonium (TEA) salt (0.15 M), in a solution of the receptor (0.01 M) in DMSO-d_6/0.5% H_2O, to a 0.01 M solution of the receptor in DMSO-d_6/0.5% H_2O. Both salt and receptor were dried under high vacuum prior to use. ^{1}H NMR spectra were recorded on a Bruker AV300 spectrometer and calibrated to the residual protio solvent peak in DMSO-d_6 (δ = 2.50 ppm). In most cases a change in the chemical shift of both (thio)urea NH protons was observed, as well as a shift in the *ortho*-CH proton of the aromatic ring (*ortho* with respect to thiourea functionality). Where possible, the WinEQNMR2 computer program[9] was used to curve-fit the data and to obtain binding constants (using a 1:1 model). Stack plots and fit plots can be found in figures S75-S162, whereas an overview of the obtained binding constants can be found in Sections 5.3 and 5.4. Data for compounds **2** and **9** has been previously reported[1-2] and is given for comparison.

Job plot analyses were performed in a separate experiment. 10 NMR tubes were filled with 0.5 mL of a DMSO-d_6/water solution containing 0.01 M of an anion-receptor mixture in different ratios (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1 molar fraction of the receptor). Both salt (TBA or TEA salts of various anions) and receptor were dried under high vacuum prior to use. ^{1}H NMR spectra were recorded on a Bruker AV300 spectrometer and calibrated to the residual protio solvent peak in DMSO-d_6 (δ = 2.50 ppm). Job plots were obtained by plotting the molar fraction of the receptor as a function of the relative change in chemical shift.

S5.2 Hammett constant, pK_a and $V_{S,max}$ calculations

Hammett constants were taken from ref 10 and pKa values were calculated using ACD Ilabs 2.0[11] and their values are given in Section S5.4.

The DFT and $V_{S,max}$ calculations were preceded by conformational analyses on chloride complexes, which were performed with AMBER12 suite.[12] The receptors **1-22** were described with General Amber Force Field (GAFF) parameters[13-14] and AM1-BCCatomic charges[15-16]. The anion was described with a net charge set to -1 and van der Waals parameters[17] developed for the TIP3P water model[18]. The Gaussian '09 package was used for all electronic structure calculations.[19]

Conformational analysis

The molecular mechanics (MM) optimized structures of **1-22** chloride complexes were submitted to a 1 ns MD quenched run in the gas phase at 500 K. 10,000 structures were saved and further full MM minimized until the convergence criterion of 0.0001 kcal mol^{-1} was achieved. Afterwards, the conformations of each complex were energy sorted and the lowest energy conformation complex was selected for DFT geometry optimizations and determination of $V_{S,max}$.

DFT calculations

The structures of the chloride complexes with **1-22** previously determined by quenched MD were re-optimized by DFT methods using the B3LYP functional and 6-311+G** basis set for all elements, except for iodine in **10**, which was described with 6-311G* basis set. The DFT optimized structures of chloride complexes are represented in Figures S73.

Electrostatic potential calculations

Afterwards, single point energy calculations at the same level of theory were performed on the free receptors, after removal of the chloride anion and keeping the conformation observed in the complex. This calculation enabled the determination of the electrostatic potential maxima on these molecules. The electrostatic potential at any point *r* created by the molecule's nuclei and electrons is given by

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')dr'}{|r' - r|}$$

Where Z_A is the charge of nucleus A located at R_A and $\rho(r')$ is the molecule's electronic density. In this work, the $V(r)$ was evaluated on the 0.001 electrons $Bohr^{-3}$contour of $\rho(r)$and is henceforth labeled $V_s(r)$. The electrostatic potential surface maxima points, $V_{S,max}$, the electrostatic potential surface minima points, $V_{S,min}$, and the average absolute deviation values, PI, were computed using the Wavefunction Analysis Program gently provided by Bulat.[20-21]

The strength of $V_{S,max}$ can be related with the hydrogen bond capacity as reported by Clark *et al*.[22] Additionally, $V_{S,max}$ is also considered an excellent descriptor for the acidity.[23] The computed $V_{S,max}$, $V_{S,min}$ and PI values for **1-22** are in given in Section S5.4, while the electrostatic potential mapped on the molecular electron density surfaces are represented in Figures S74. All compounds' $V_{S,max}$ are represented as pink dots in Figure S74, and they are located at the anion binding pocket, which is the most positive region of the molecules.

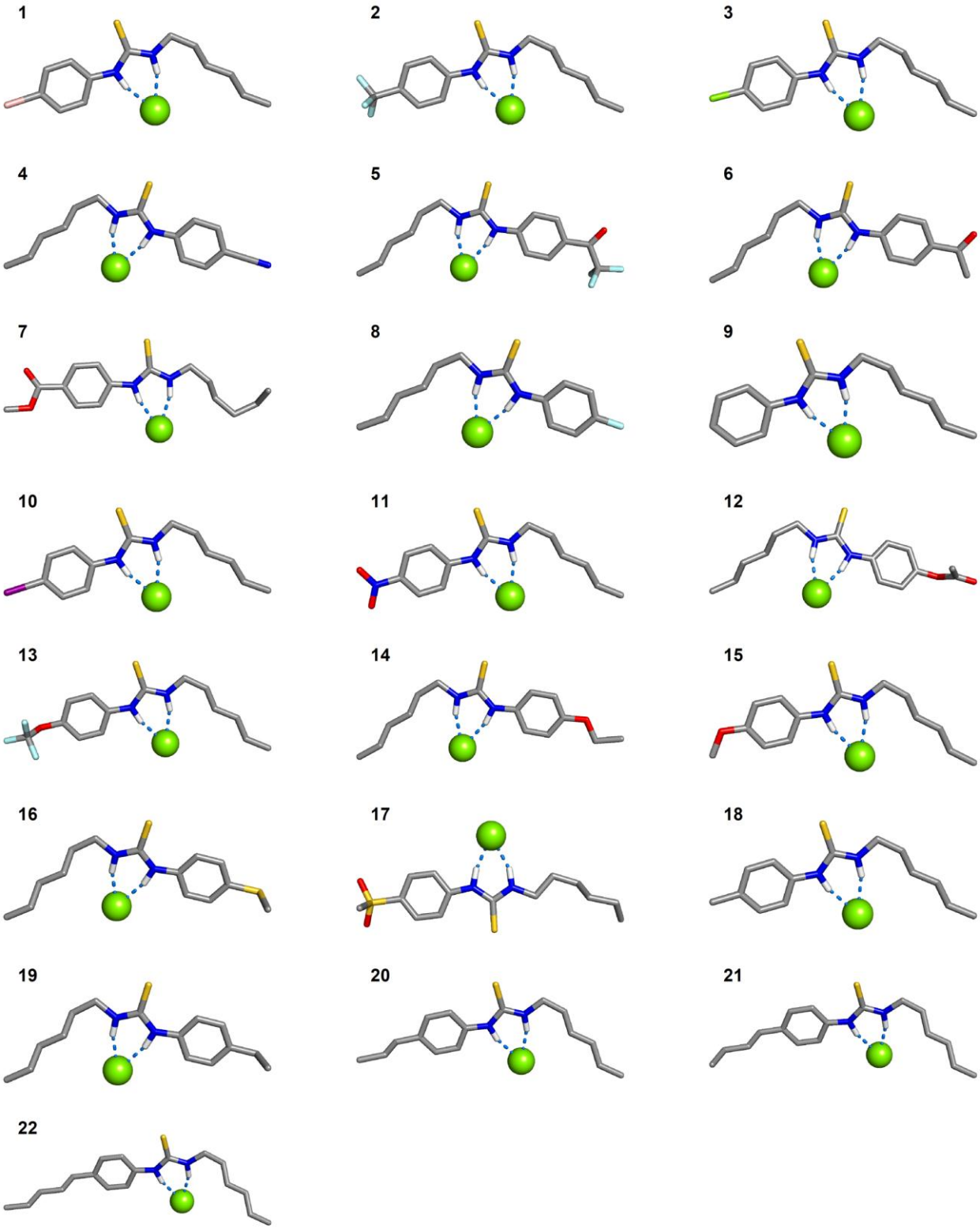


Figure S73. DFT optimized structures of **1**-**22** chloride complexes. Hydrogen atoms apart from the ones involved in hydrogen bonds were omitted for clarity. Hydrogen bonds are shown in blue dashes.

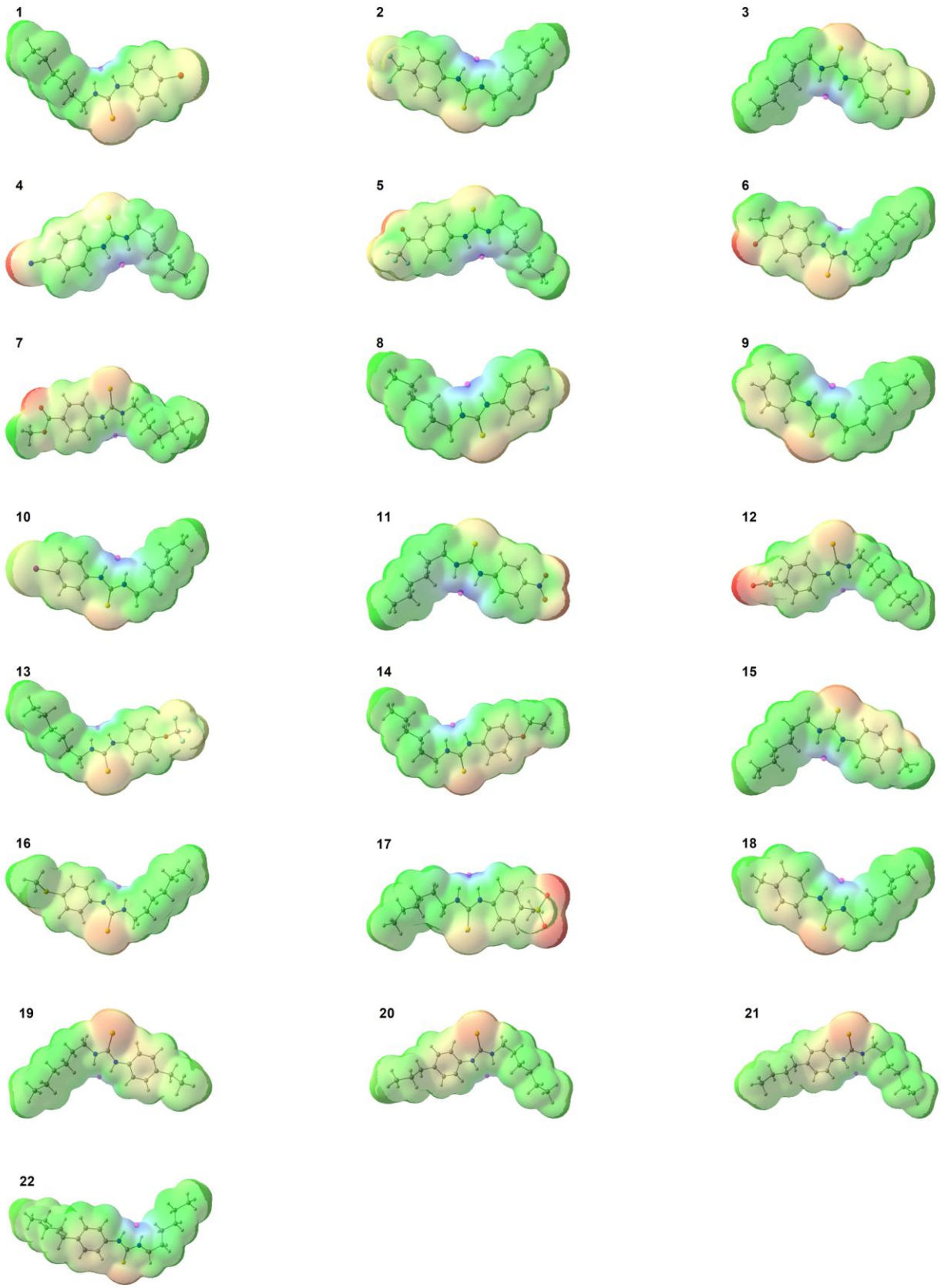


Figure S74. Electrostatic potential mapped on the molecular electron density surface (0.001 electrons Bohr^{-3}) for **1**-**22**. The colour scale ranges from blue (+0.11 a.u.) to red (-0.07 a.u.). The pink dots correspond to the location of the $V_{S,max}$.

S5.3 Overview of ^{1}H NMR titrations

*Interactions of compound **1** (Br) with various anions*

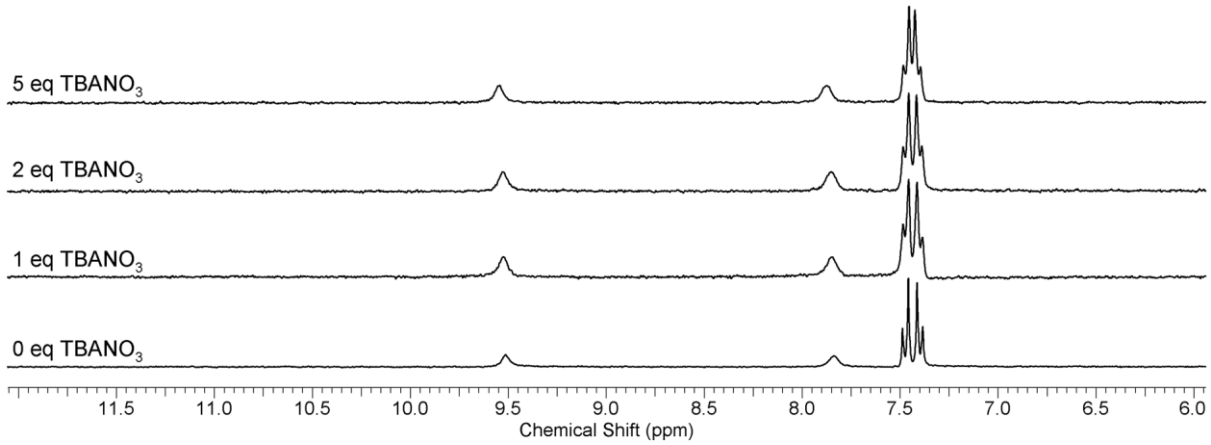


Figure S75. ^{1}H NMR titration of compound **1** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

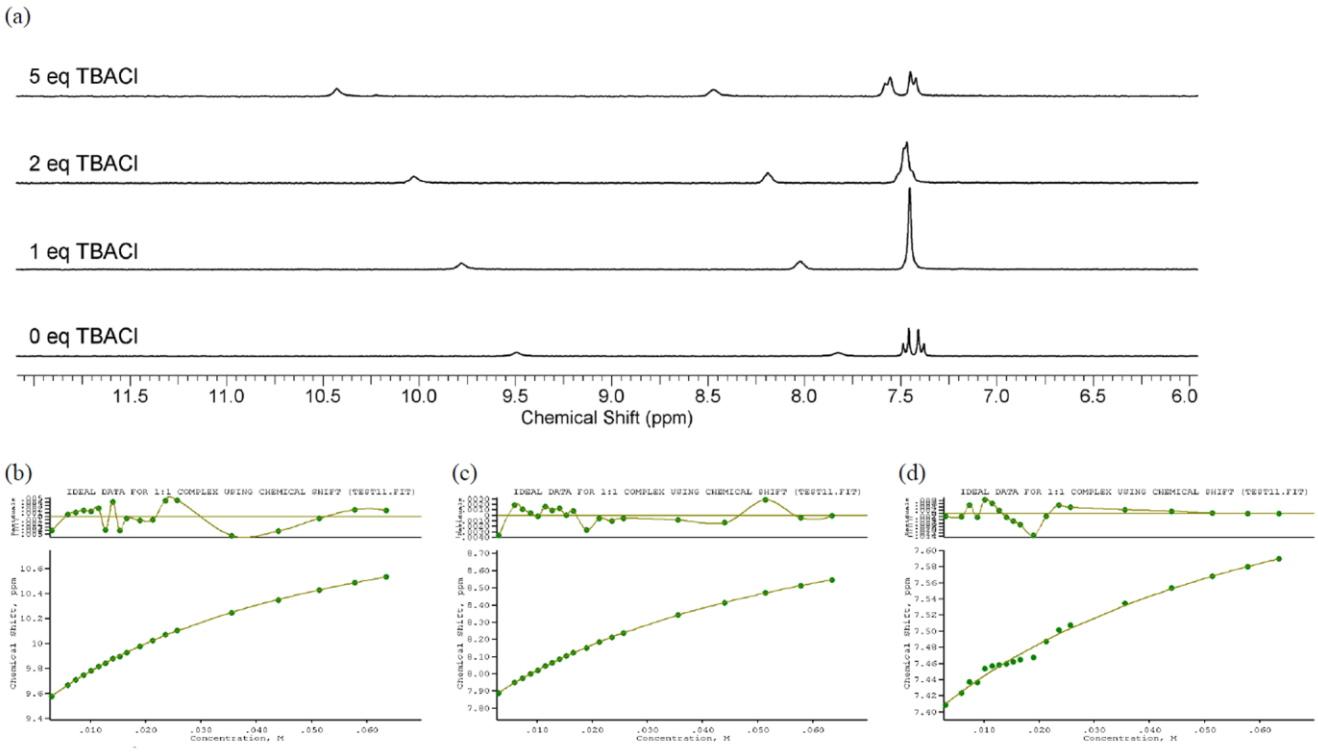


Figure S76. ^{1}H NMR titration of compound **1** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.49 ppm. K_a = 22.0 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 7.81 ppm. K_a = 18.6 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.38 ppm. K_a = 16.5 M^{-1} (error 2 %).

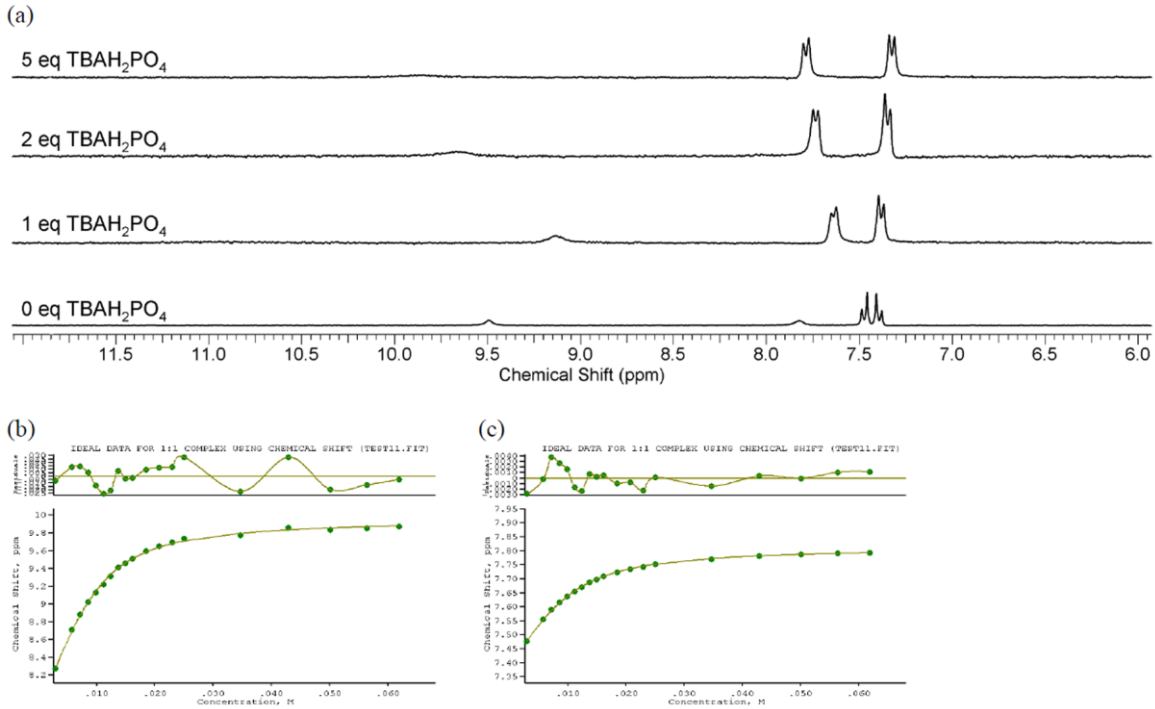


Figure S77. ^{1}H NMR titration of compound **1** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.81 ppm. K_a = 465 M^{-1} (error 5 %). (c) Fit plot for CH proton at δ = 7.38 ppm. K_a = 347 M^{-1} (error 2 %).

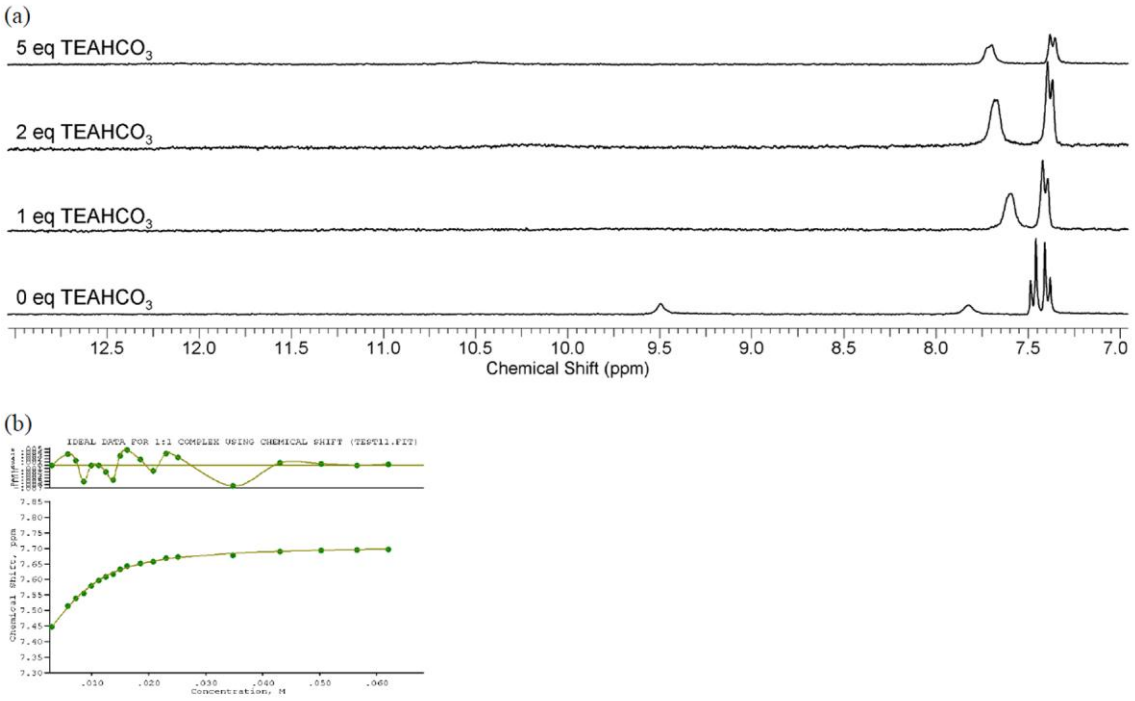


Figure S78. ^{1}H NMR titration of compound **1** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.38 ppm. K_a = 447 M^{-1} (error 3 %).

*Interactions of compound **2** (CF_3) with various anions*

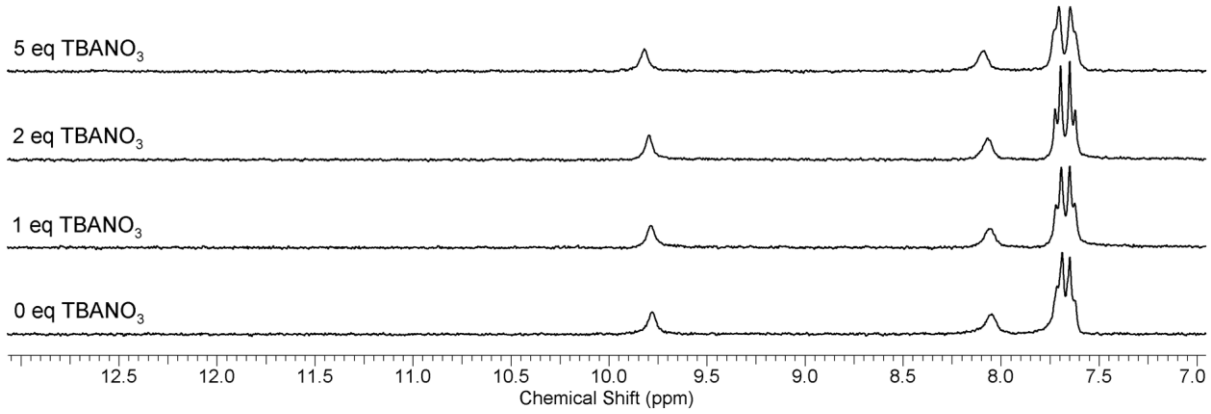


Figure S79. ^{1}H NMR titration of compound **2** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

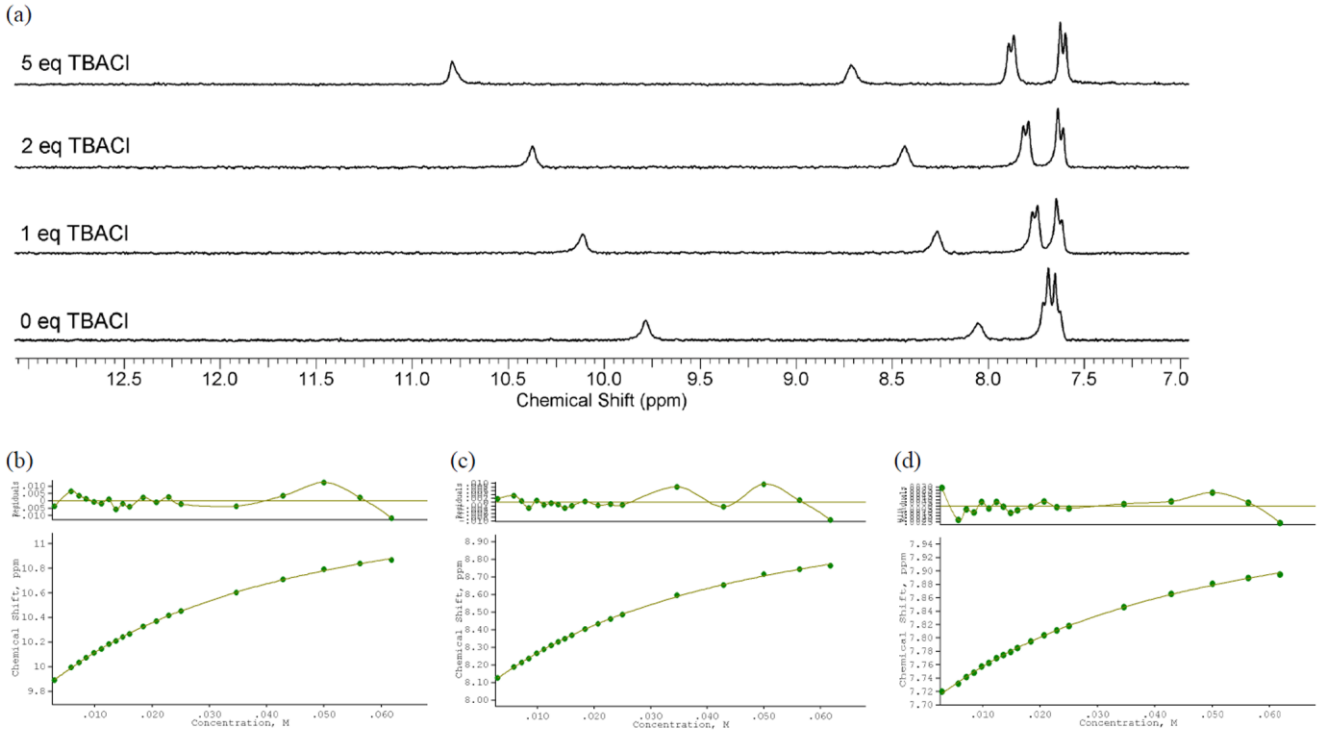


Figure S80. ^{1}H NMR titration of compound **2** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.80 ppm. K_a = 27.9 M^{-1} (error 3 %). (c) Fit plot for NH proton at δ = 8.05 ppm. K_a = 26.1 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.71 ppm. K_a = 25.9 M^{-1} (error 5 %).

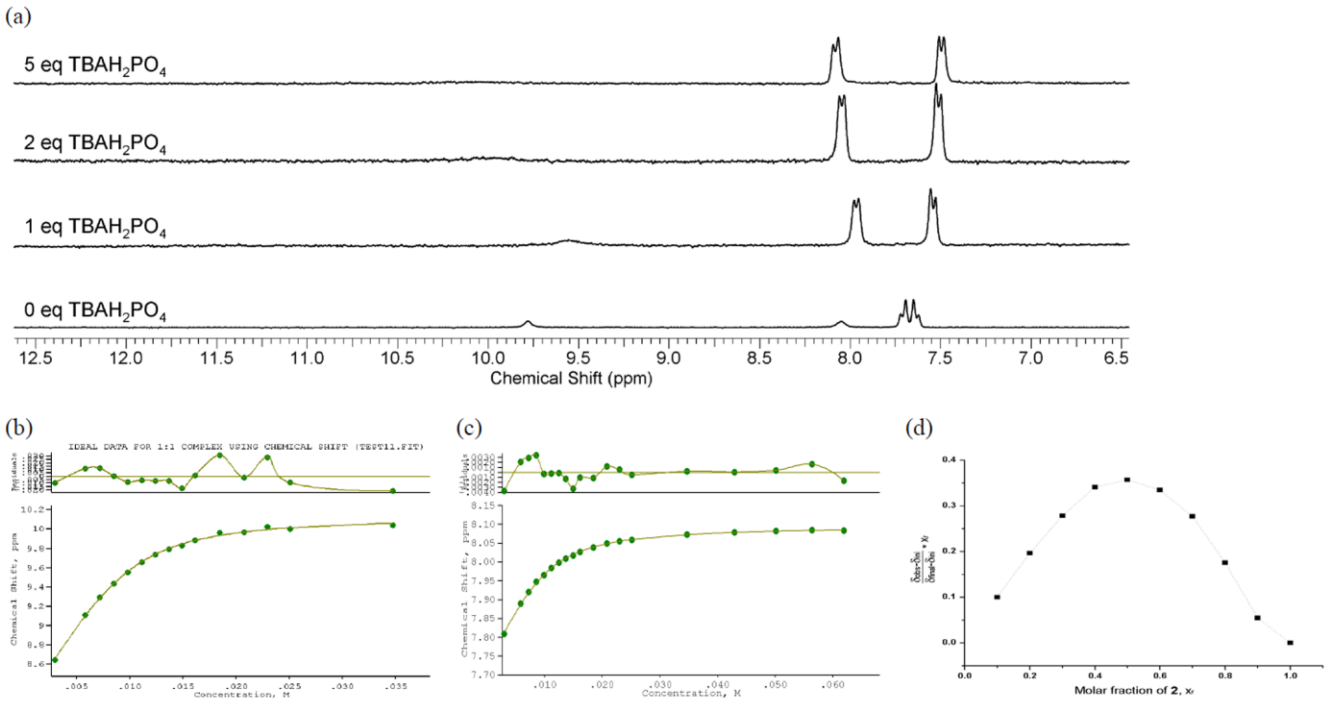


Figure S81. ^{1}H NMR titration of compound **2** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.05 ppm. K_a = 852 M^{-1} (error 8 %). (c) Fit plot for CH proton at δ = 7.71 ppm. K_a = 532 M^{-1} (error 3 %). (d) Job plot for CH proton at δ = 7.71 ppm.

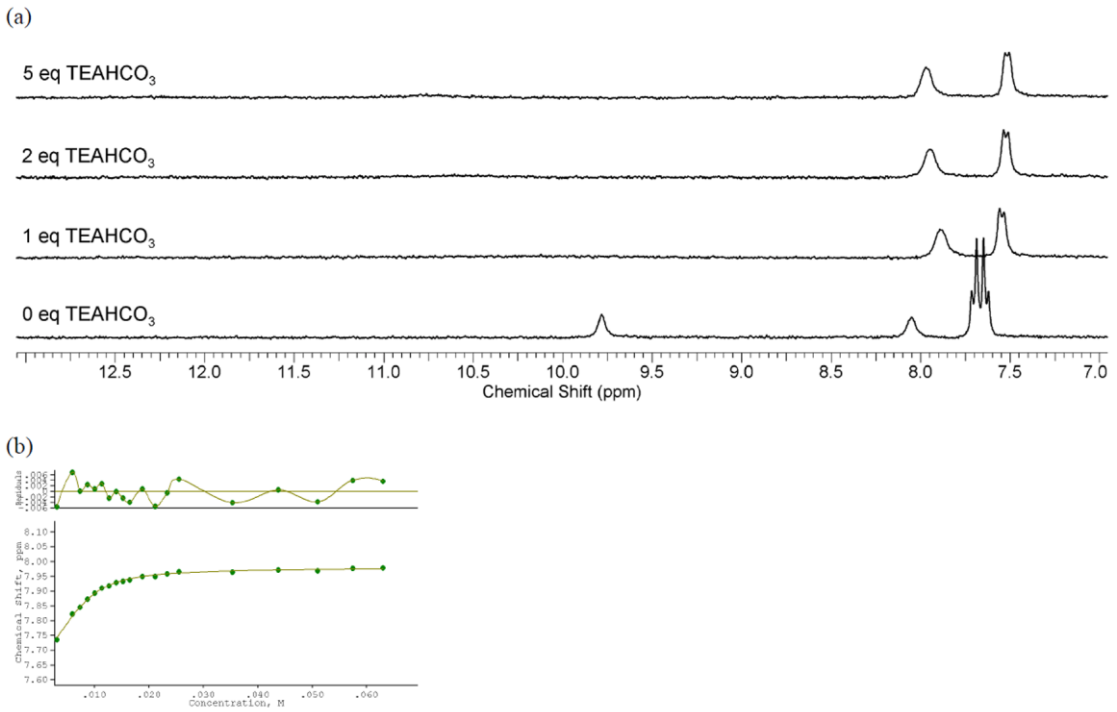


Figure S82. ^{1}H NMR titration of compound **2** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.71 ppm. K_a = 931 M^{-1} (error 7.5 %).

*Interactions of compound **3** (Cl) with various anions*

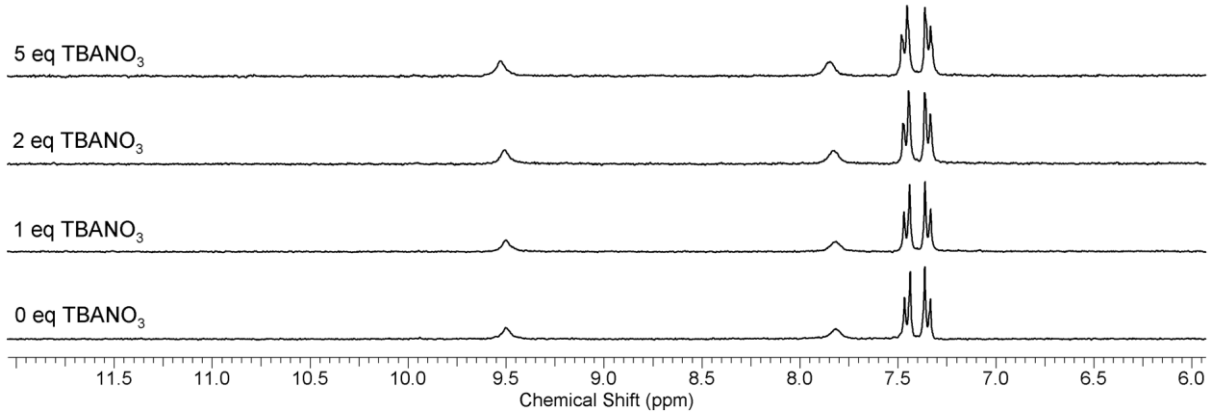


Figure S83. ^{1}H NMR titration of compound **3** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

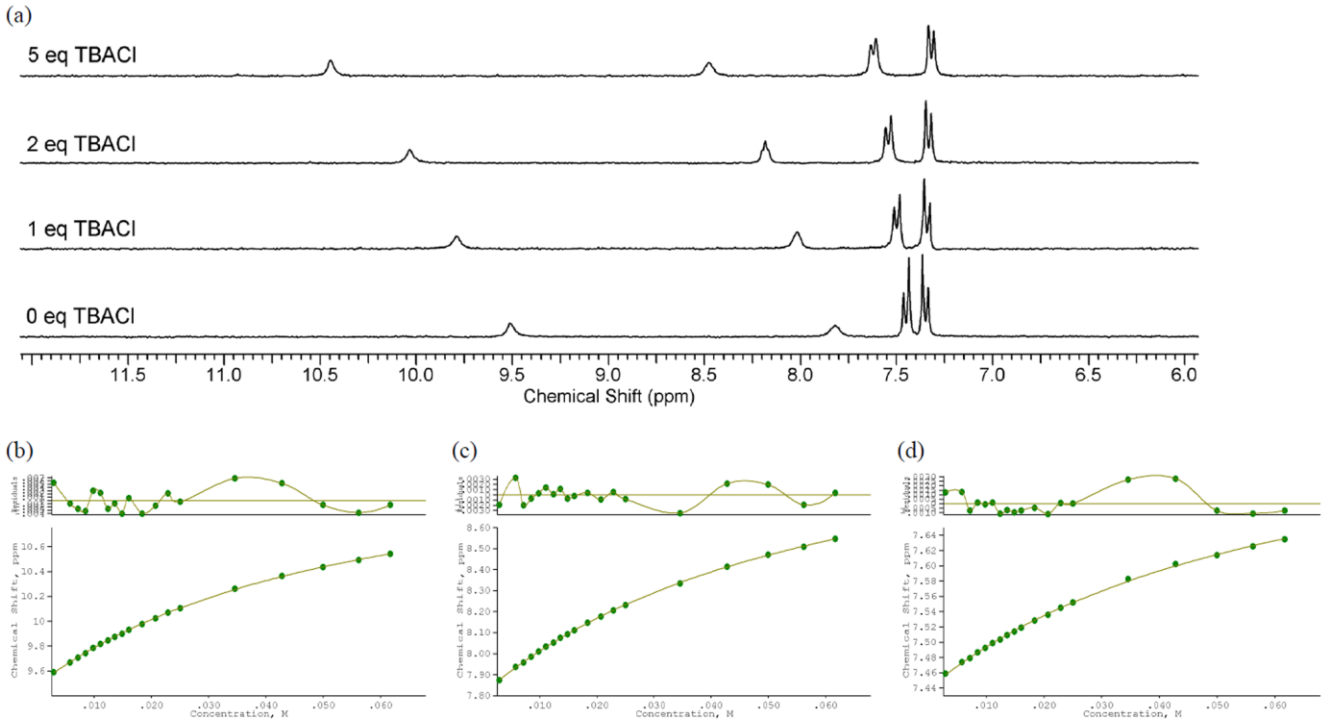


Figure S84. ^{1}H NMR titration of compound **3** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.48 ppm. K_a = 22.2 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 7.80 ppm. K_a = 19.9 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.47 ppm. K_a = 19.1 M^{-1} (error 5 %).

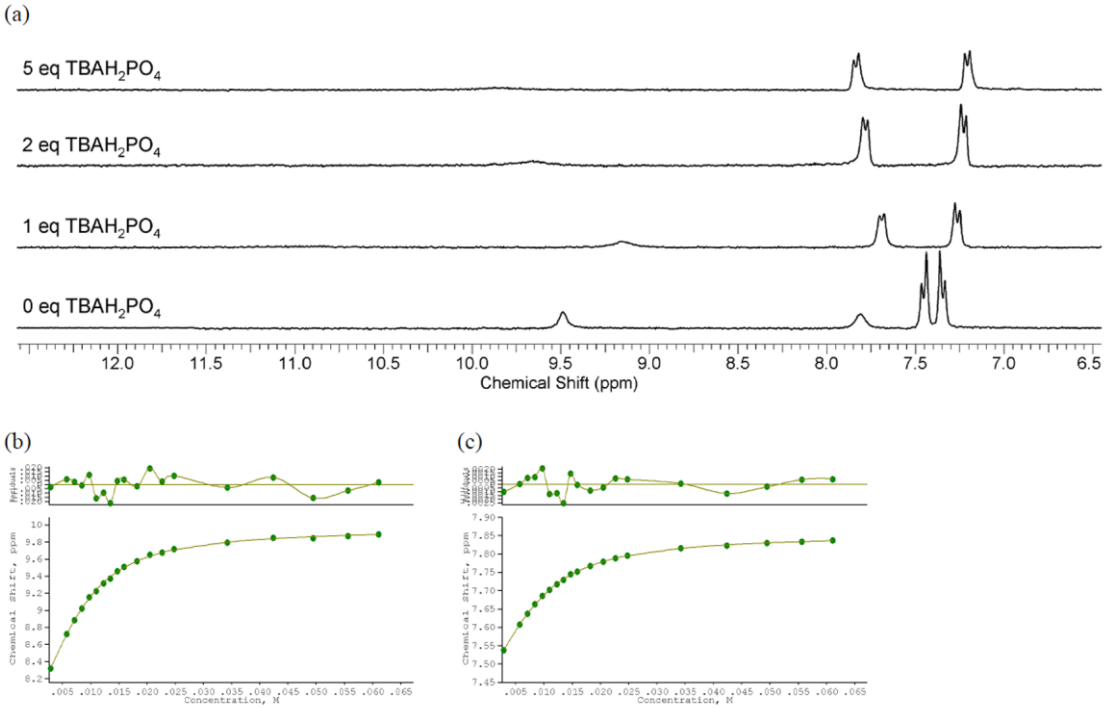


Figure S85. ^{1}H NMR titration of compound **3** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.80 ppm. K_a = 428 M^{-1} (error 3 %). (c) Fit plot for CH proton at δ = 7.47 ppm. K_a = 331 M^{-1} (error 2 %).

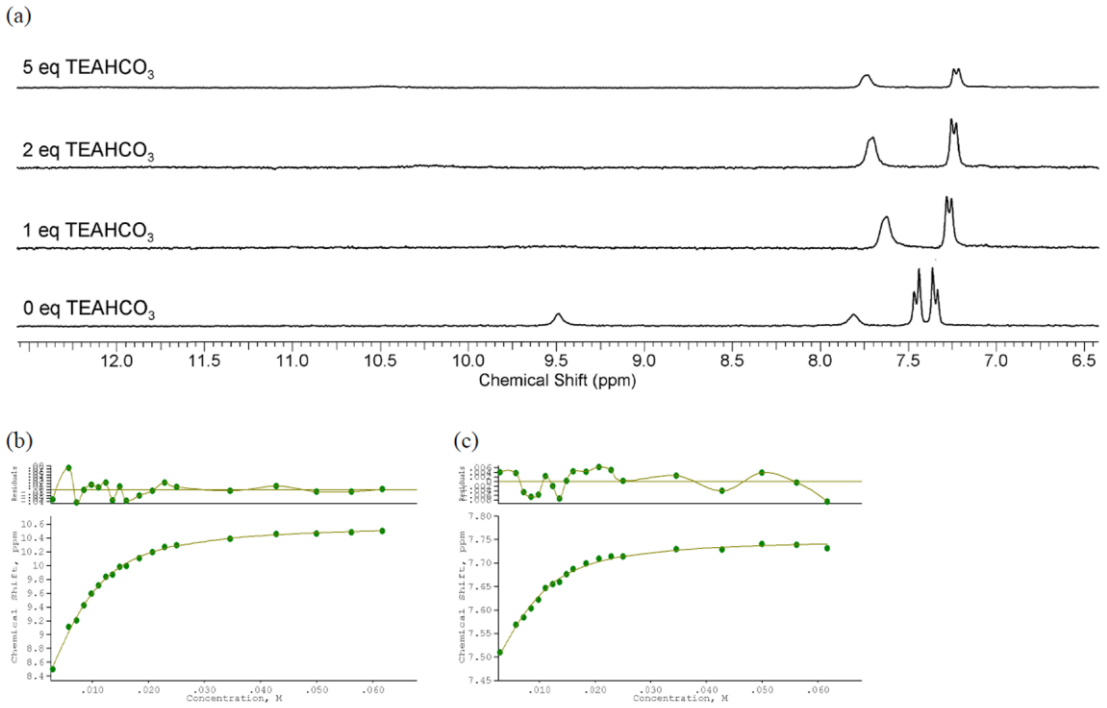


Figure S86. ^{1}H NMR titration of compound **3** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.80 ppm. K_a = 452 M^{-1} (error 6 %). (c) Fit plot for CH proton at δ = 7.47 ppm. K_a = 428 M^{-1} (error 9 %).

*Interactions of compound **4** (CN) with various anions*

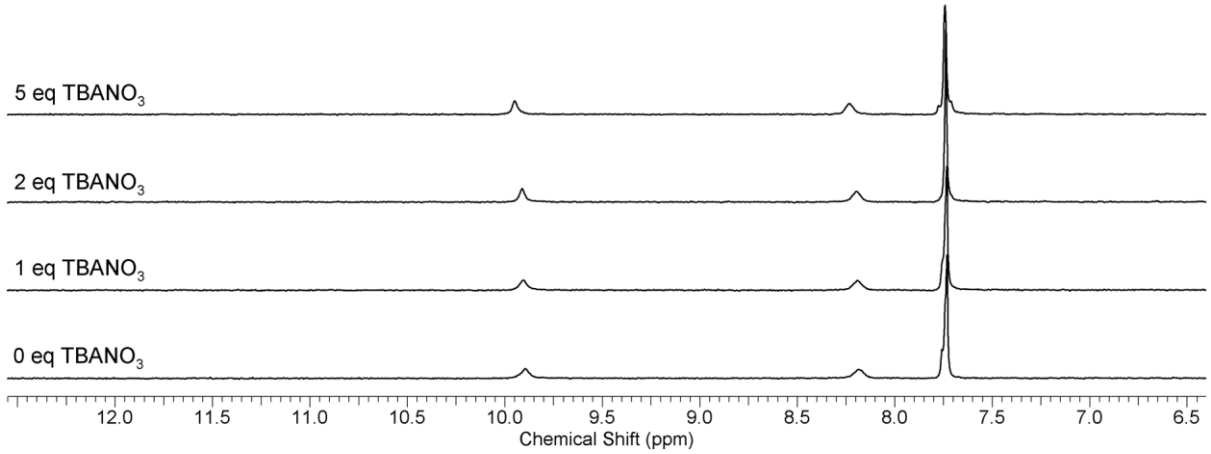


Figure S87. ^{1}H NMR titration of compound **4** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

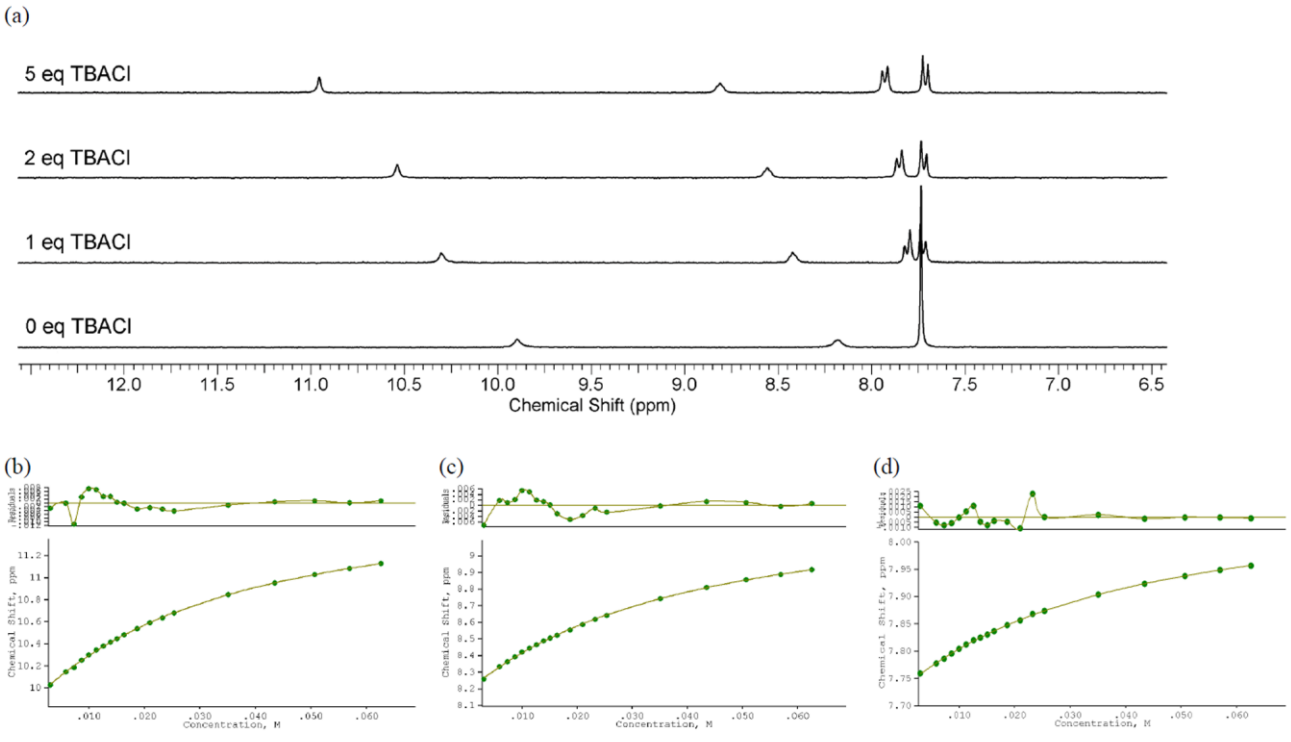


Figure S88. ^{1}H NMR titration of compound **4** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.85 ppm. K_a = 34.6 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 8.14 ppm. K_a = 31.4 M^{-1} (error 3 %). (d) Fit plot for CH proton at δ = 7.73 ppm. K_a = 30.5 M^{-1} (error 2.5 %).

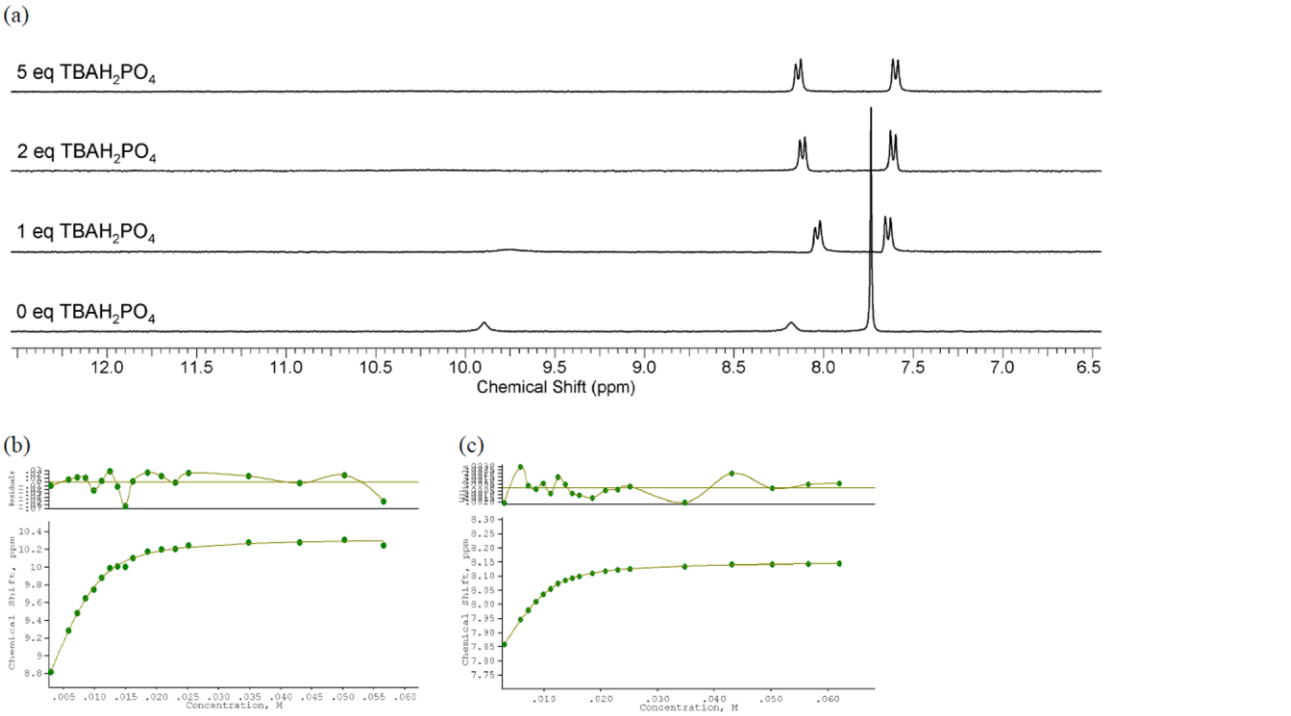


Figure S89. ^{1}H NMR titration of compound **4** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.14 ppm. K_a = 1103 M^{-1} (error 10 %). (c) Fit plot for CH proton at δ = 7.73 ppm. K_a = 921 M^{-1} (error 2 %).

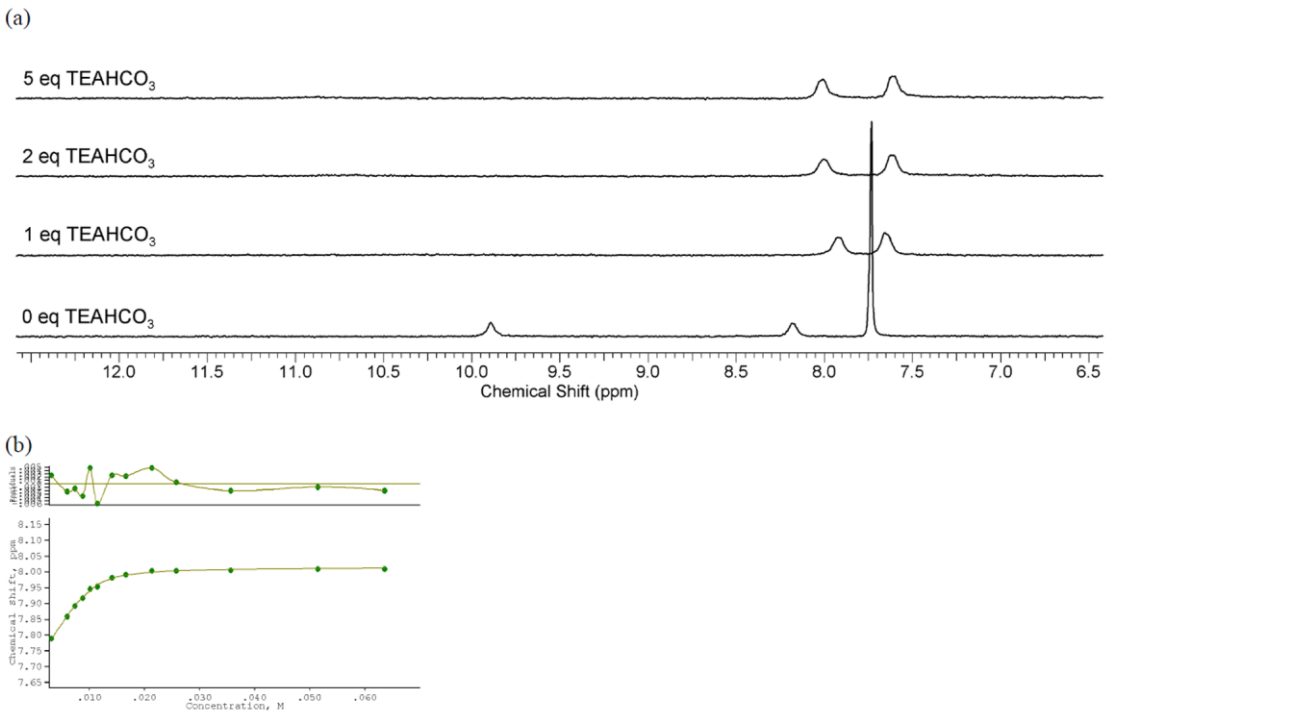


Figure S90. ^{1}H NMR titration of compound **4** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.73 ppm. K_a = 1550 M^{-1} (error 11 %).

*Interactions of compound **5** ($COCF_3$) with various anions*

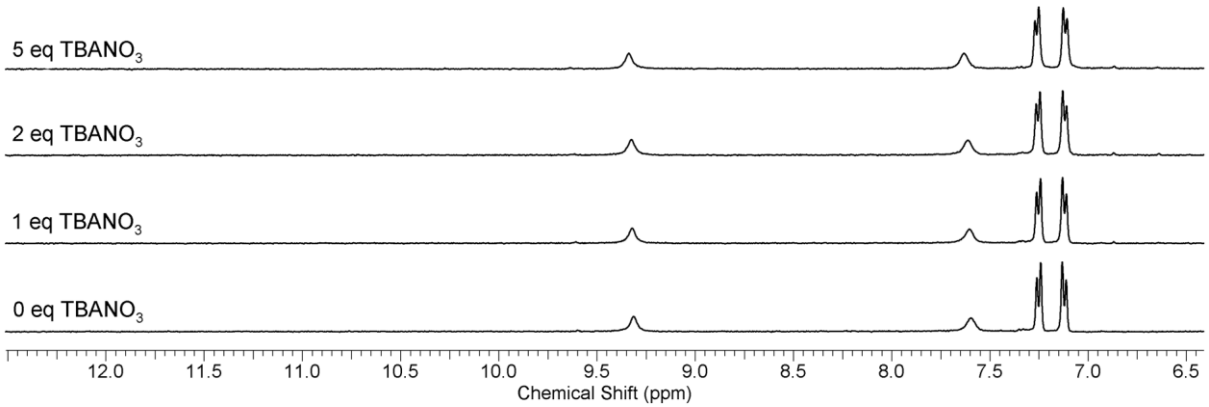


Figure S91. ^{1}H NMR titration of compound **5** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

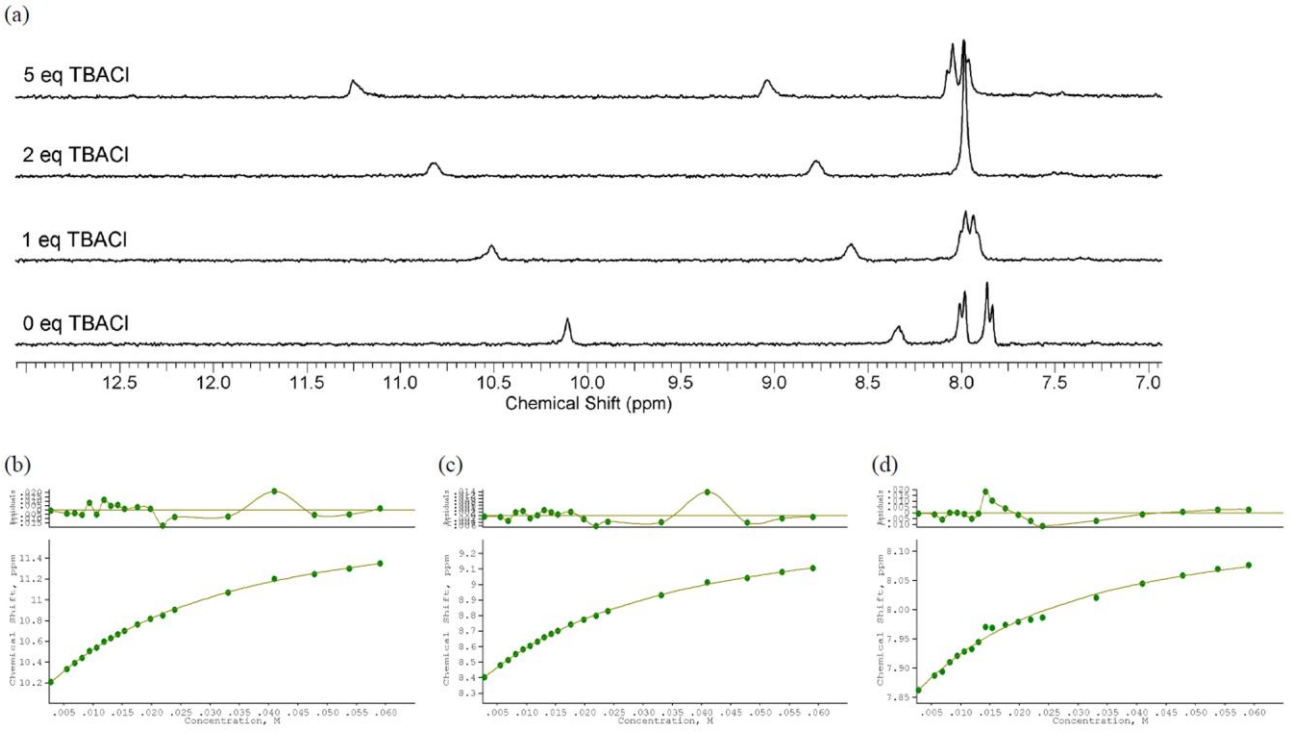


Figure S92. ^{1}H NMR titration of compound **5** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 10.09 ppm. K_a = 43.2 M^{-1} (error 4 %). (c) Fit plot for NH proton at δ = 8.32 ppm. K_a = 41.3 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.85 ppm. K_a = 51.8 M^{-1} (error 17 %).

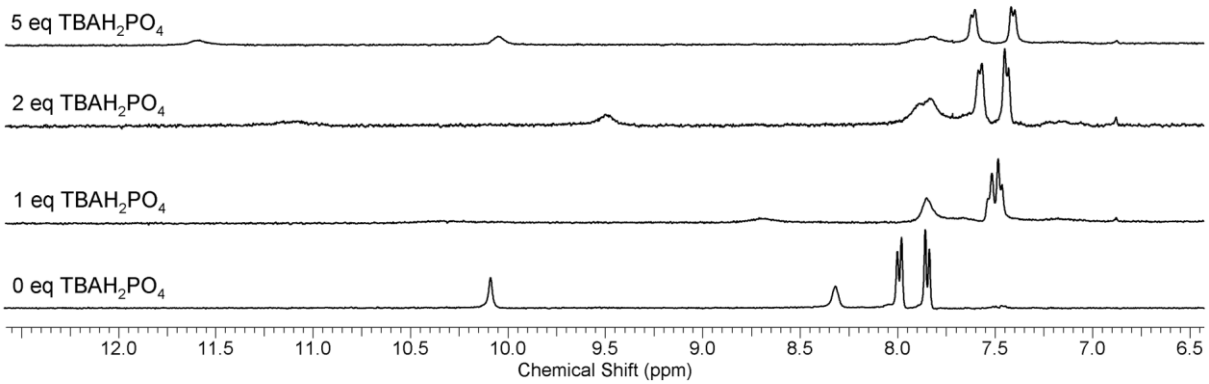


Figure S93. ^{1}H NMR titration of compound **5** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Compound decomposed.

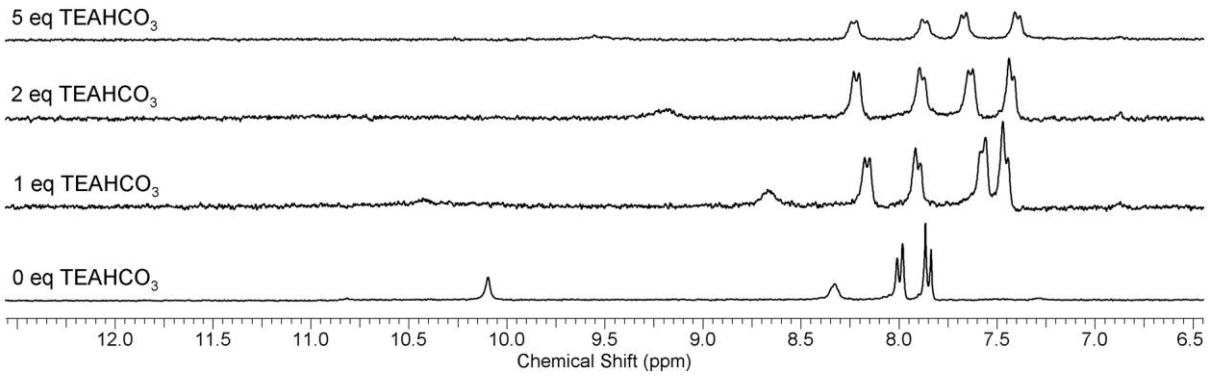


Figure S94. ^{1}H NMR titration of compound **5** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Compound decomposed.

*Interactions of compound **6** (COMe) with various anions*

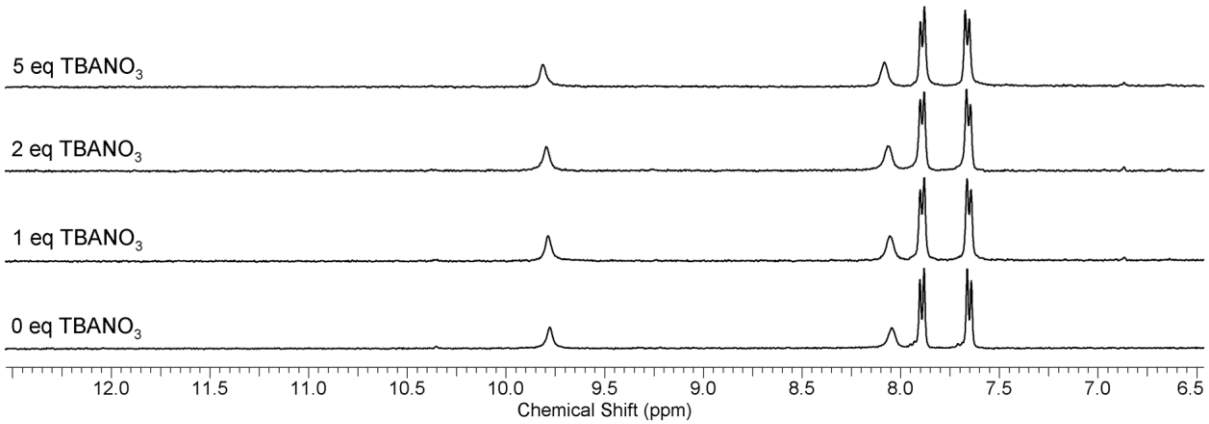


Figure S95. ^{1}H NMR titration of compound **6** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

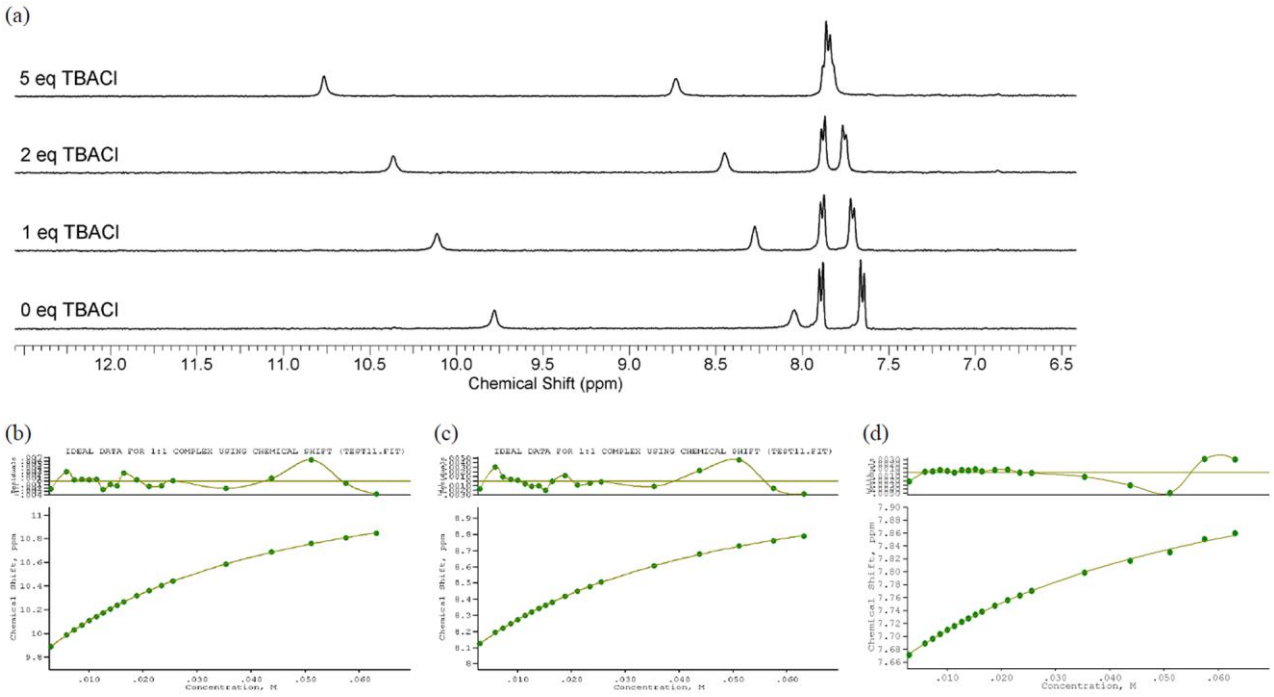


Figure S96. ^{1}H NMR titration of compound **6** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.75 ppm. K_a = 28.6 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 8.02 ppm. K_a = 26.8 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.64 ppm. K_a = 18.8 M^{-1} (error 7 %).

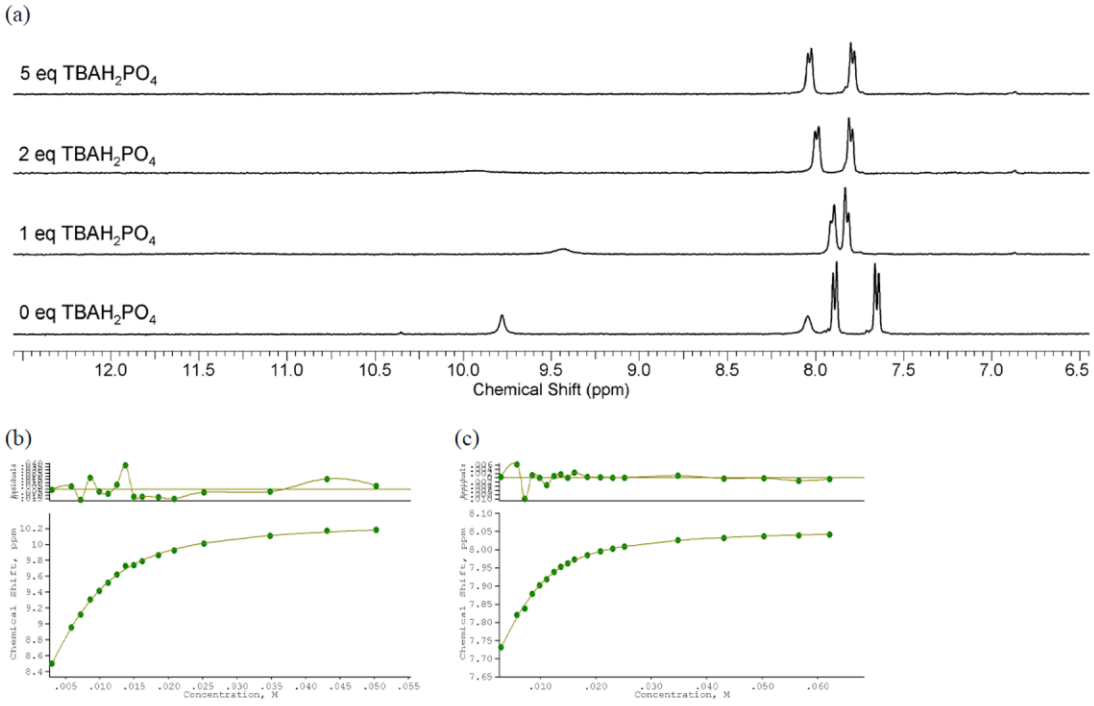


Figure S97. ^{1}H NMR titration of compound **6** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.02 ppm. K_a = 417 M^{-1} (error 4.5 %). (c) Fit plot for CH proton at δ = 7.64 ppm. K_a = 444 M^{-1} (error 4 %).

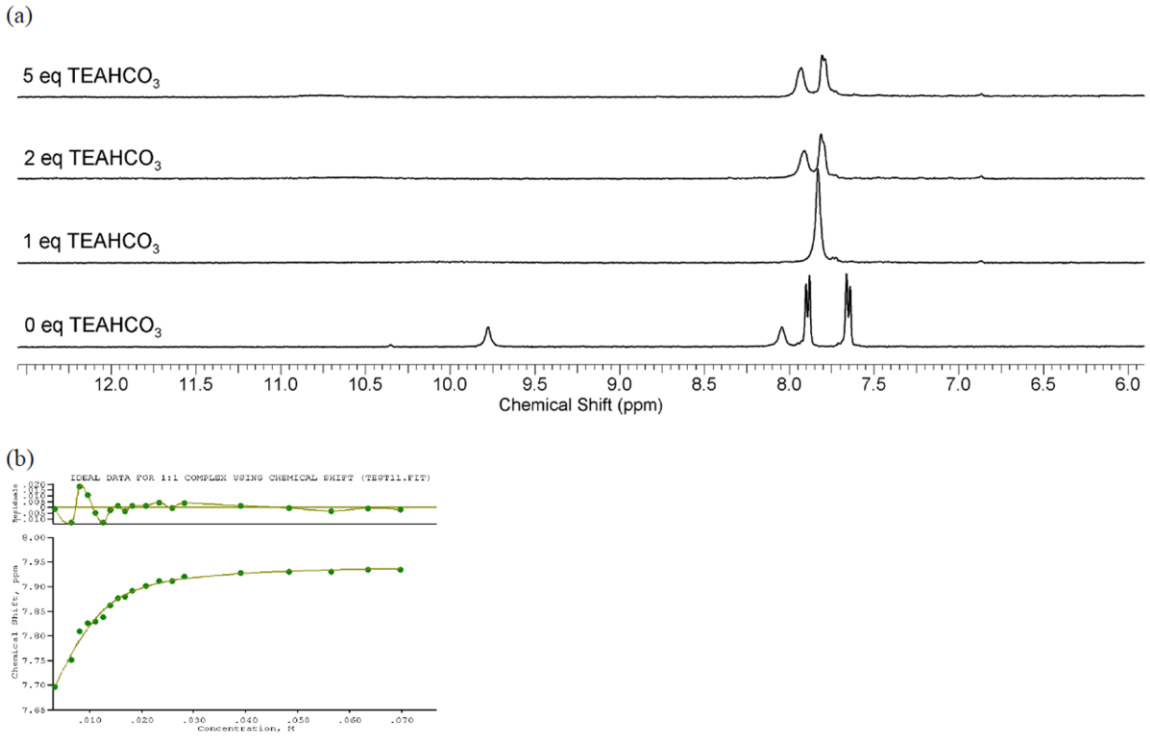


Figure S98. ^{1}H NMR titration of compound **6** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.64 ppm. K_a = 553 M^{-1} (error 13.5 %).

*Interactions of compound **7** (COOMe) with various anions*

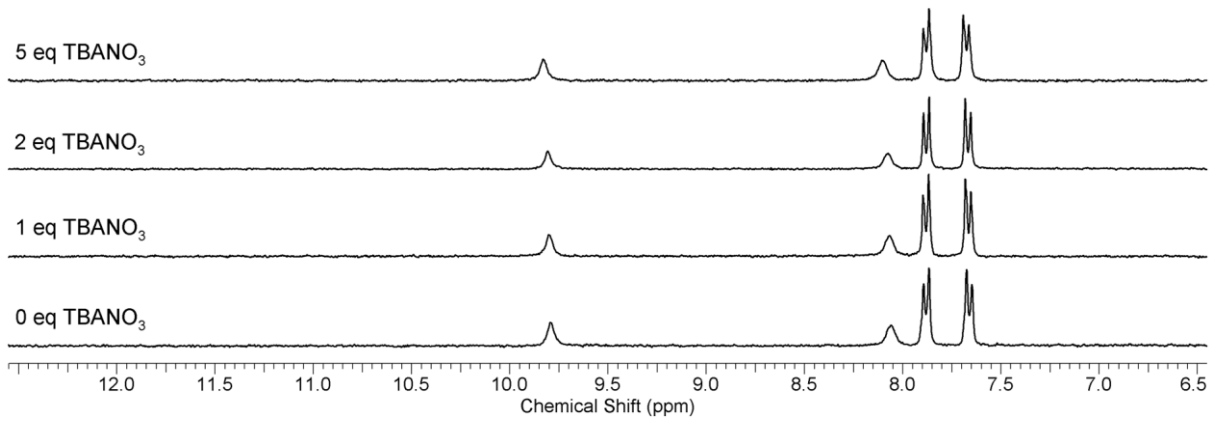


Figure S99. ^{1}H NMR titration of compound **7** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

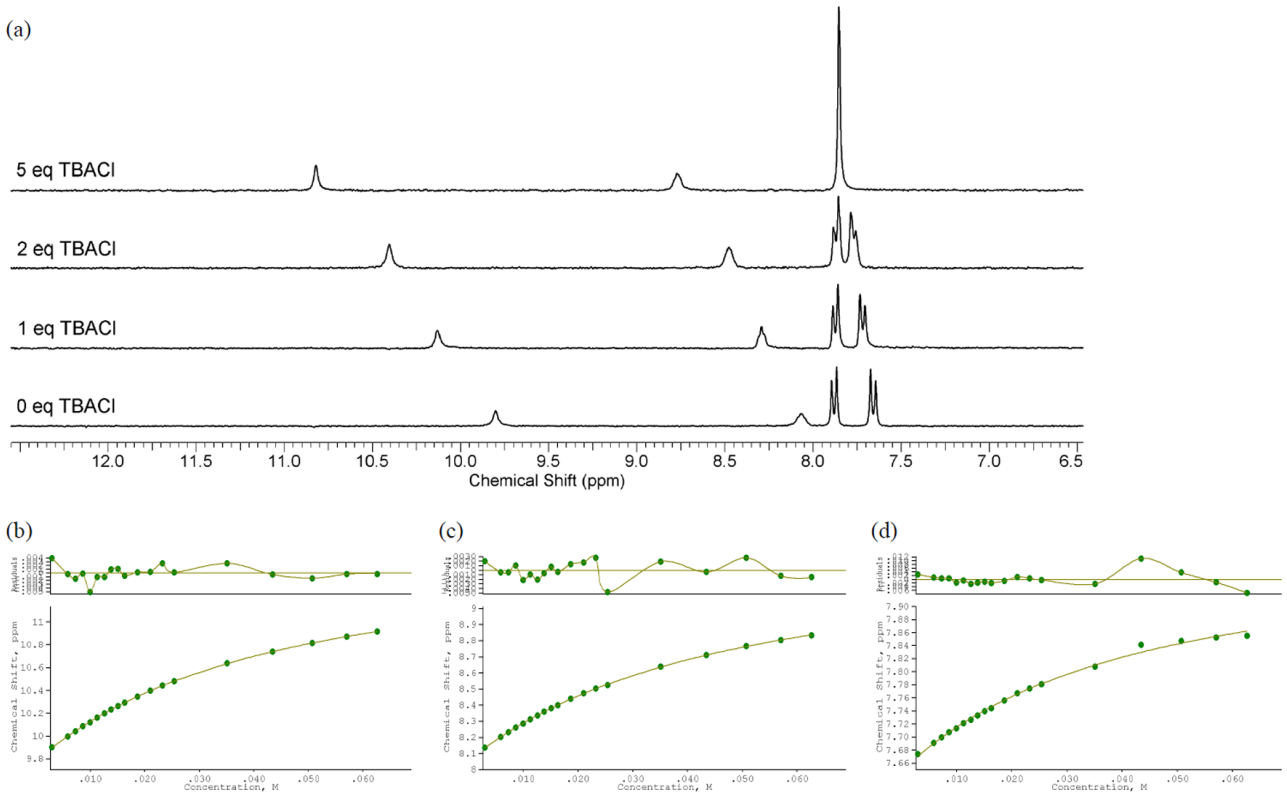


Figure S100. ^{1}H NMR titration of compound **7** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.79 ppm. K_a = 28.1 M^{-1} (error 1 %). (c) Fit plot for NH proton at δ = 8.05 ppm. K_a = 25.6 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.67 ppm. K_a = 29.2 M^{-1} (error 12 %).

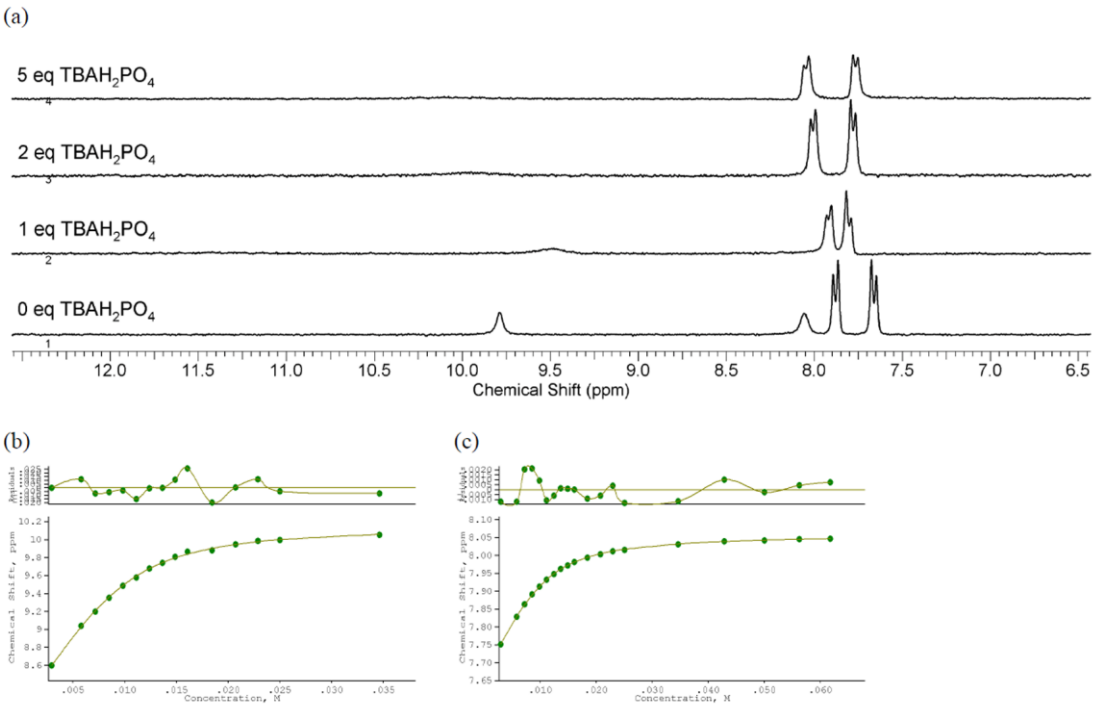


Figure S101. ^{1}H NMR titration of compound **7** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.05 ppm. K_a = 700 M^{-1} (error 6 %). (c) Fit plot for CH proton at δ = 7.67 ppm. K_a = 511 M^{-1} (error 2 %).

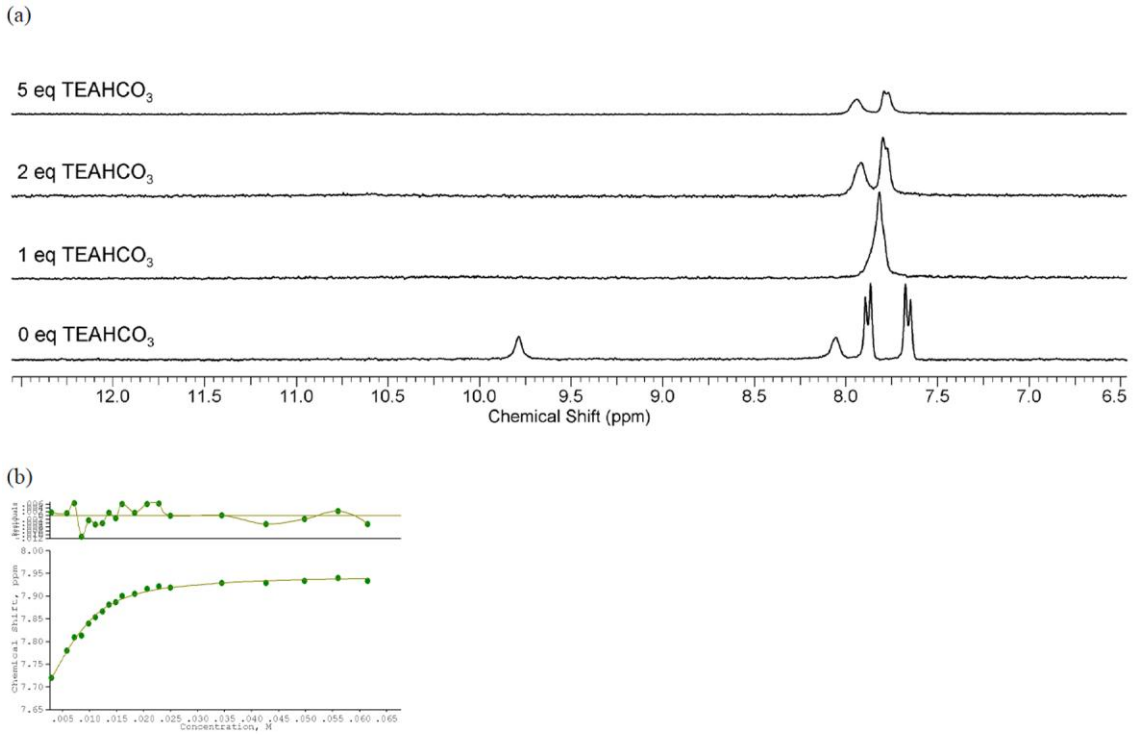


Figure S102. ^{1}H NMR titration of compound **7** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.67 ppm. K_a = 589 M^{-1} (error 9.5 %).

*Interactions of compound **8** (F) with various anions*

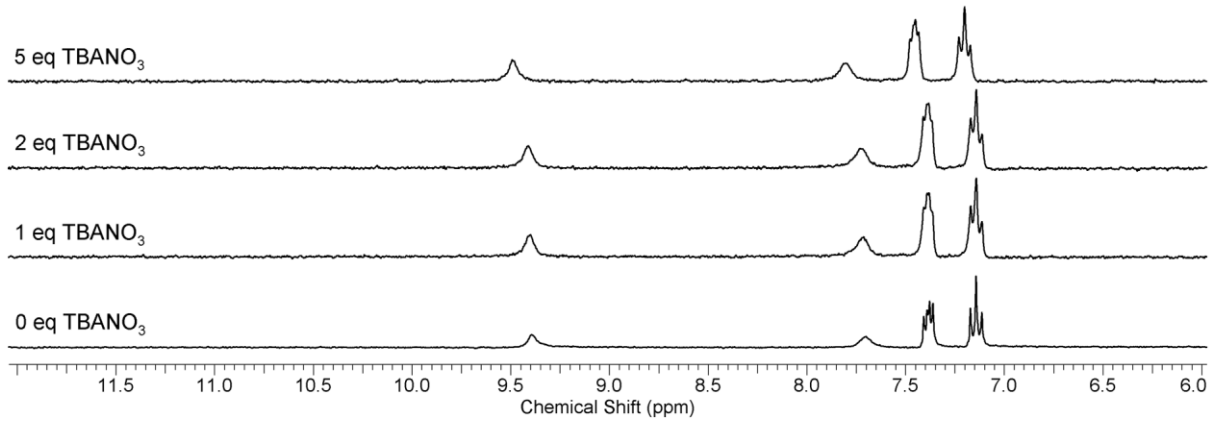


Figure S103. ^{1}H NMR titration of compound **8** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

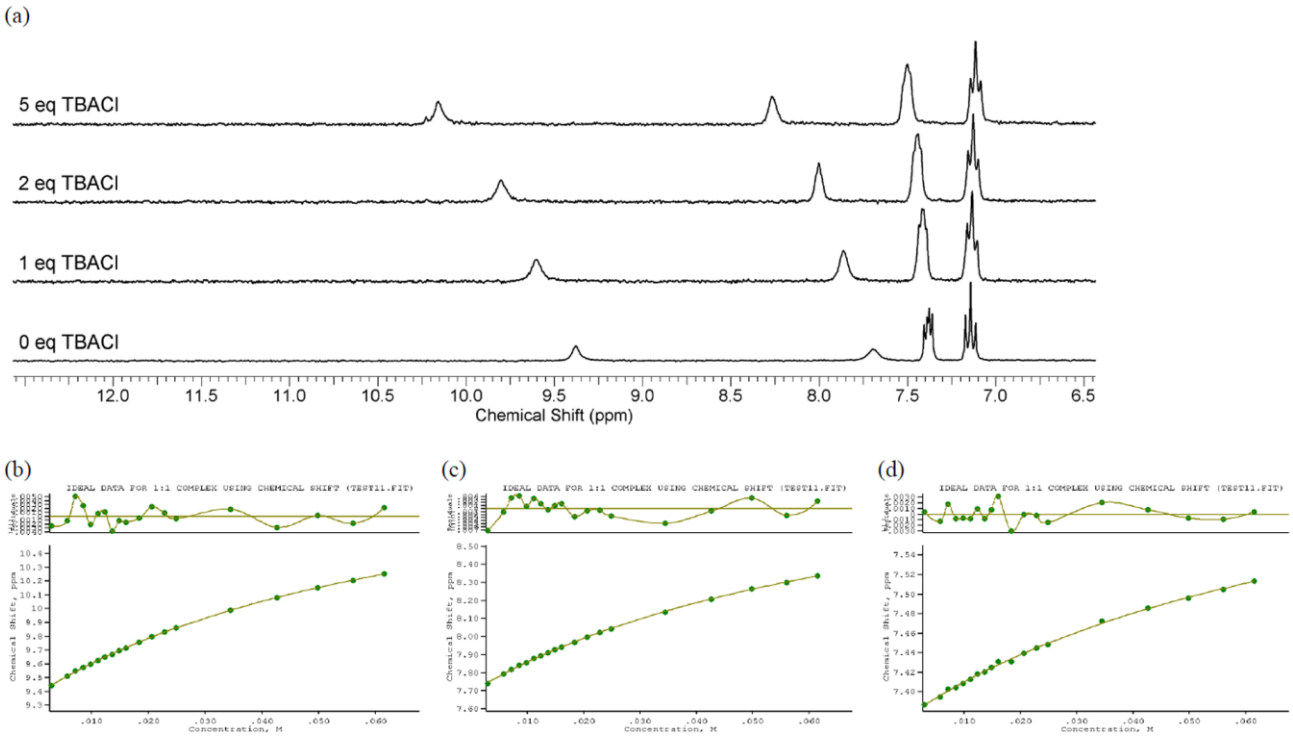


Figure S104. ^{1}H NMR titration of compound **8** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.37 ppm. K_a = 16.8 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 7.69 ppm. K_a = 15.5 M^{-1} (error 4.5 %). (d) Fit plot for CH proton at δ = 7.39 ppm. K_a = 13.8 M^{-1} (error 9 %).

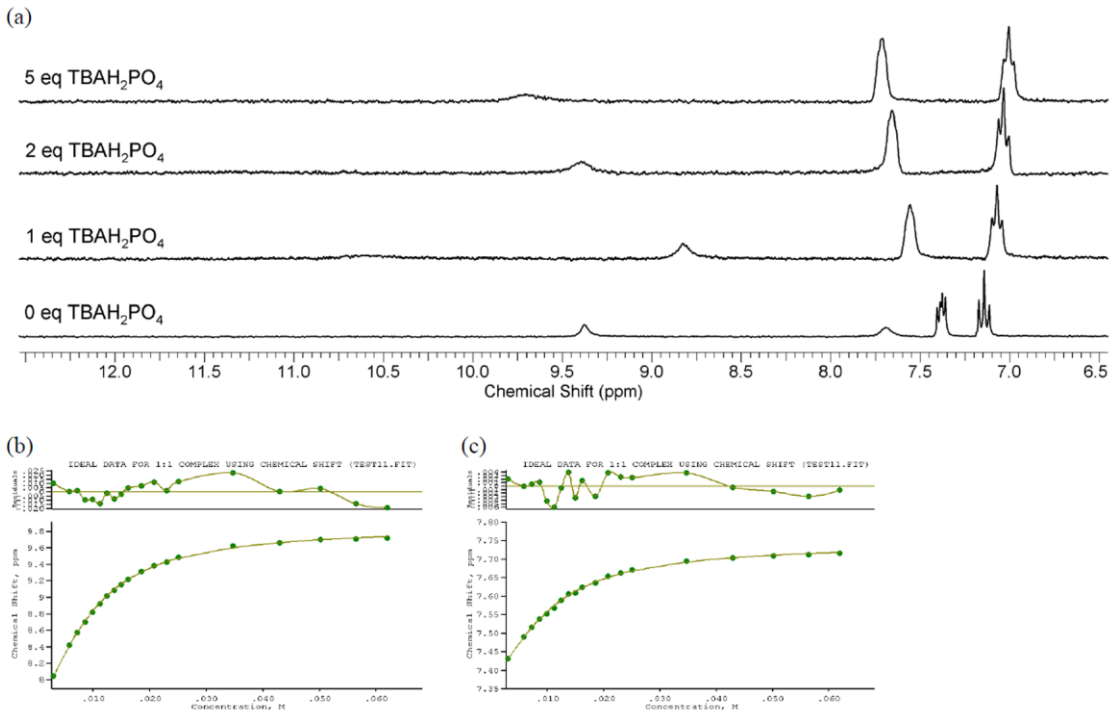


Figure S105. ^{1}H NMR titration of compound **8** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.69 ppm. K_a = 272 M^{-1} (error 2.5 %). (c) Fit plot for CH proton at δ = 7.39 ppm. K_a = 229 M^{-1} (error 4 %).

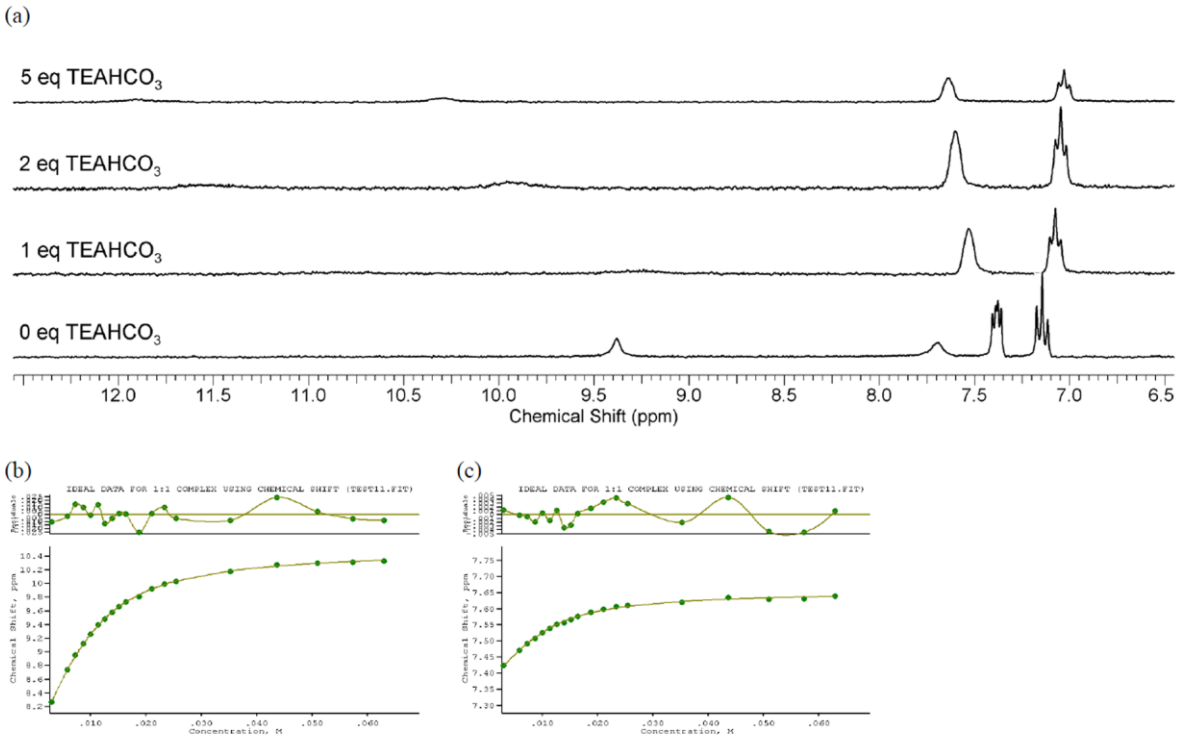


Figure S106. ^{1}H NMR titration of compound **8** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.69 ppm. K_a = 292 M^{-1} (error 3 %). (c) Fit plot for CH proton at δ = 7.39 ppm. K_a = 303 M^{-1} (error 5.5 %).

*Interactions of compound **9** (H) with various anions*

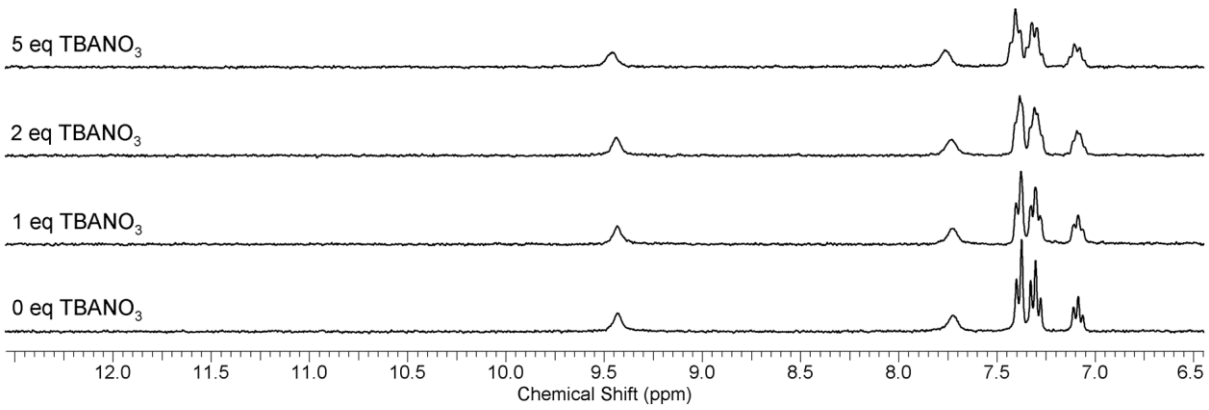


Figure S107. ^{1}H NMR titration of compound **9** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

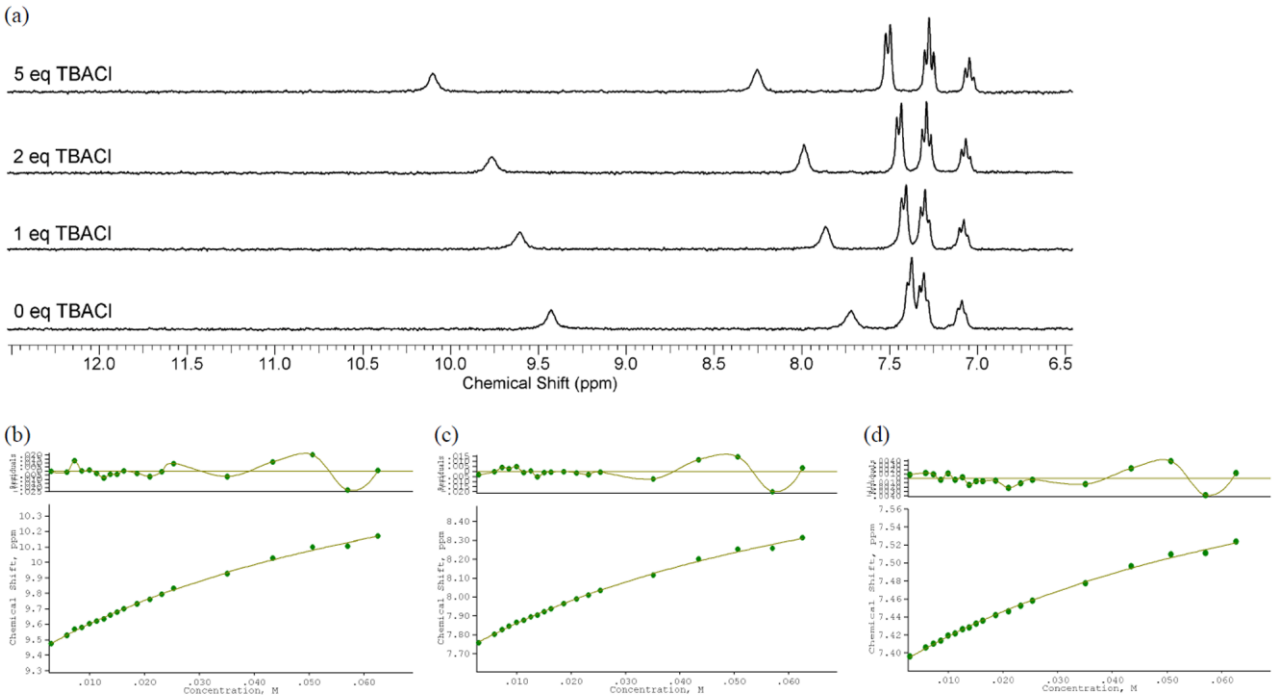


Figure S108. ^{1}H NMR titration of compound **9** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.45 ppm. K_a = 13.7 M^{-1} (error 12 %). (c) Fit plot for NH proton at δ = 7.71 ppm. K_a = 13.5 M^{-1} (error 12 %). (d) Fit plot for CH proton at δ = 7.39 ppm. K_a = 13.2 M^{-1} (error 10 %).

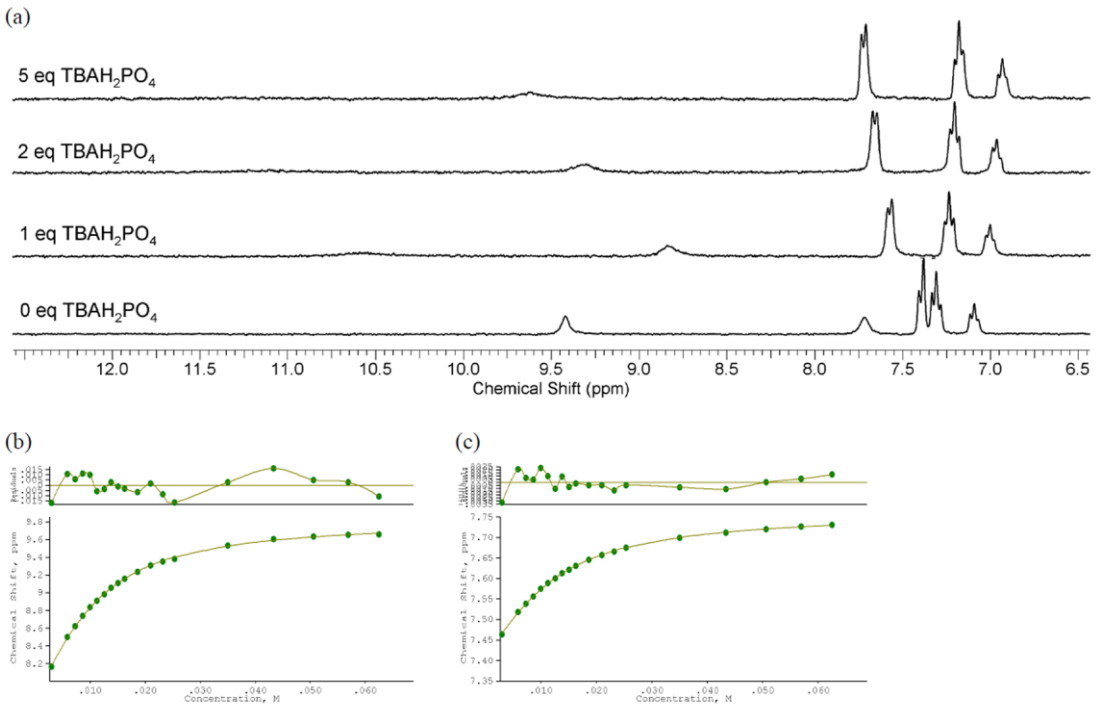


Figure S109. ^{1}H NMR titration of compound **9** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.71 ppm. K_a = 211 M^{-1} (error 3 %). (c) Fit plot for CH proton at δ = 7.39 ppm. K_a = 169 M^{-1} (error 2 %).

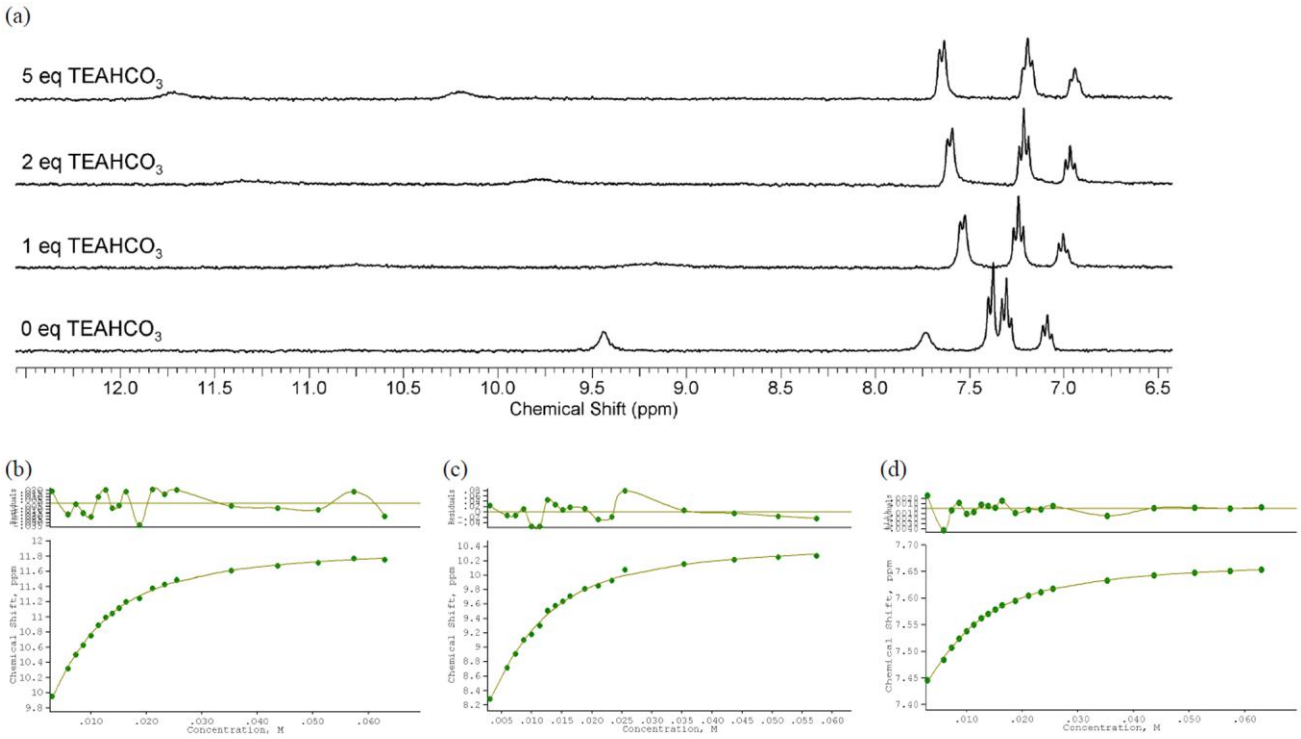


Figure S110. ^{1}H NMR titration of compound **9** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.45 ppm. K_a = 230 M^{-1} (error 4 %). (c) Fit plot for NH proton at δ = 7.71 ppm. K_a = 262 M^{-1} (error 8 %). (d) Fit plot for CH proton at δ = 7.39 ppm. K_a = 222 M^{-1} (error 2 %).

*Interactions of compound **10** (I) with various anions*

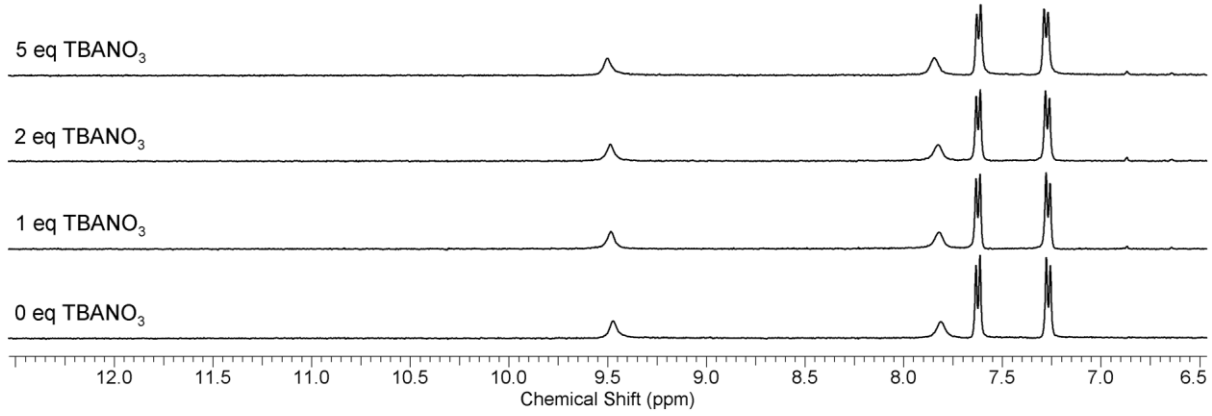


Figure S111. ^{1}H NMR titration of compound **10** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

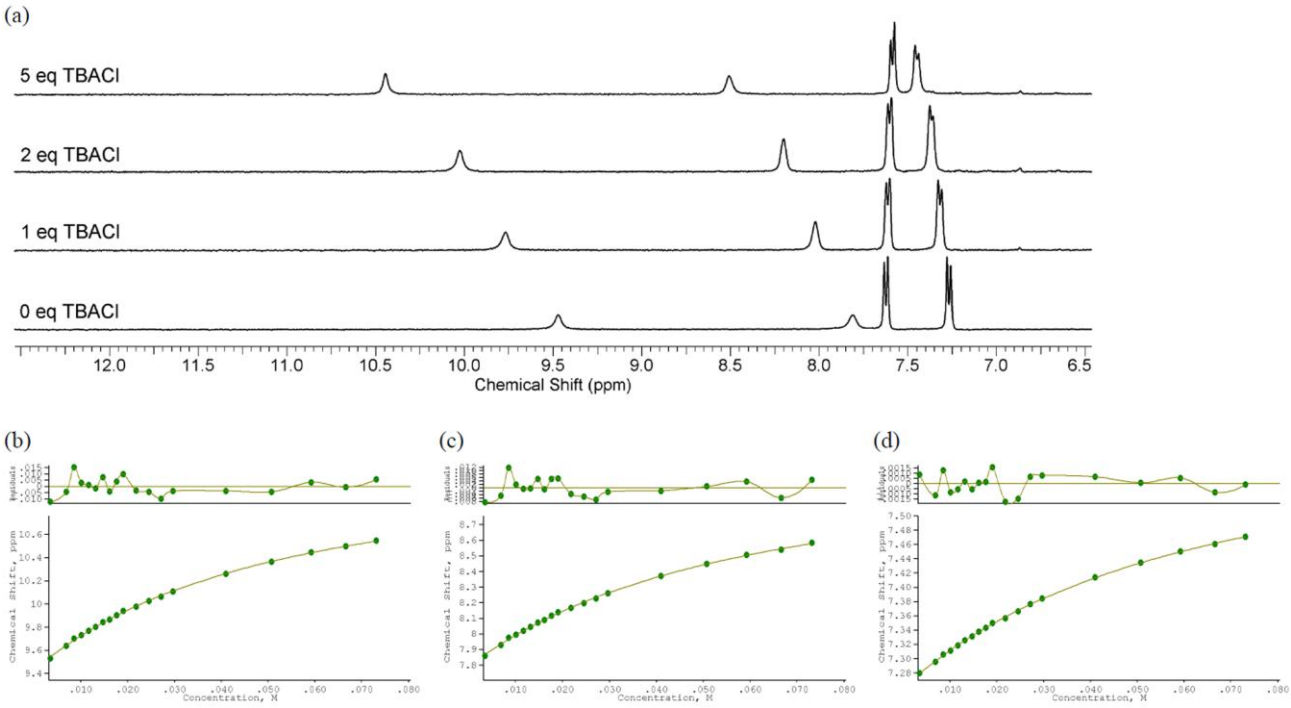


Figure S112. ^{1}H NMR titration of compound **10** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.47 ppm. K_a = 22.7 M^{-1} (error 4.5 %). (c) Fit plot for NH proton at δ = 7.81 ppm. K_a = 19.9 M^{-1} (error 5 %). (d) Fit plot for CH proton at δ = 7.28 ppm. K_a = 17.8 M^{-1} (error 3 %).

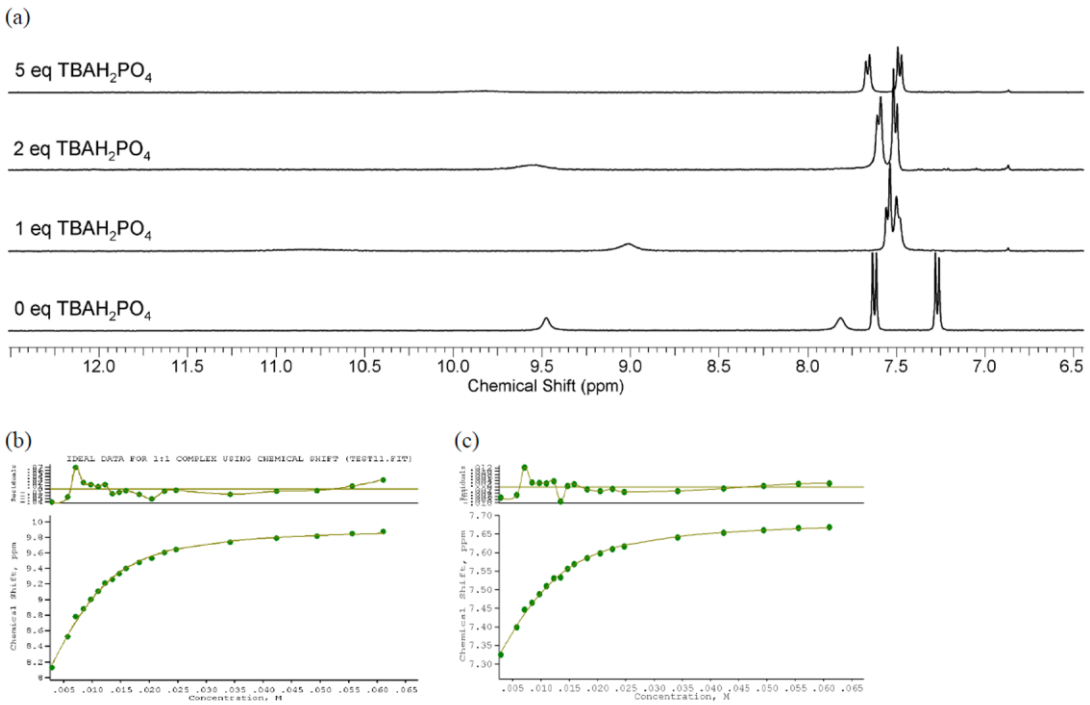


Figure S113. ^{1}H NMR titration of compound **10** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.81 ppm. K_a = 448 M^{-1} (error 7.5 %). (c) Fit plot for CH proton at δ = 7.28 ppm. K_a = 354 M^{-1} (error 7 %).

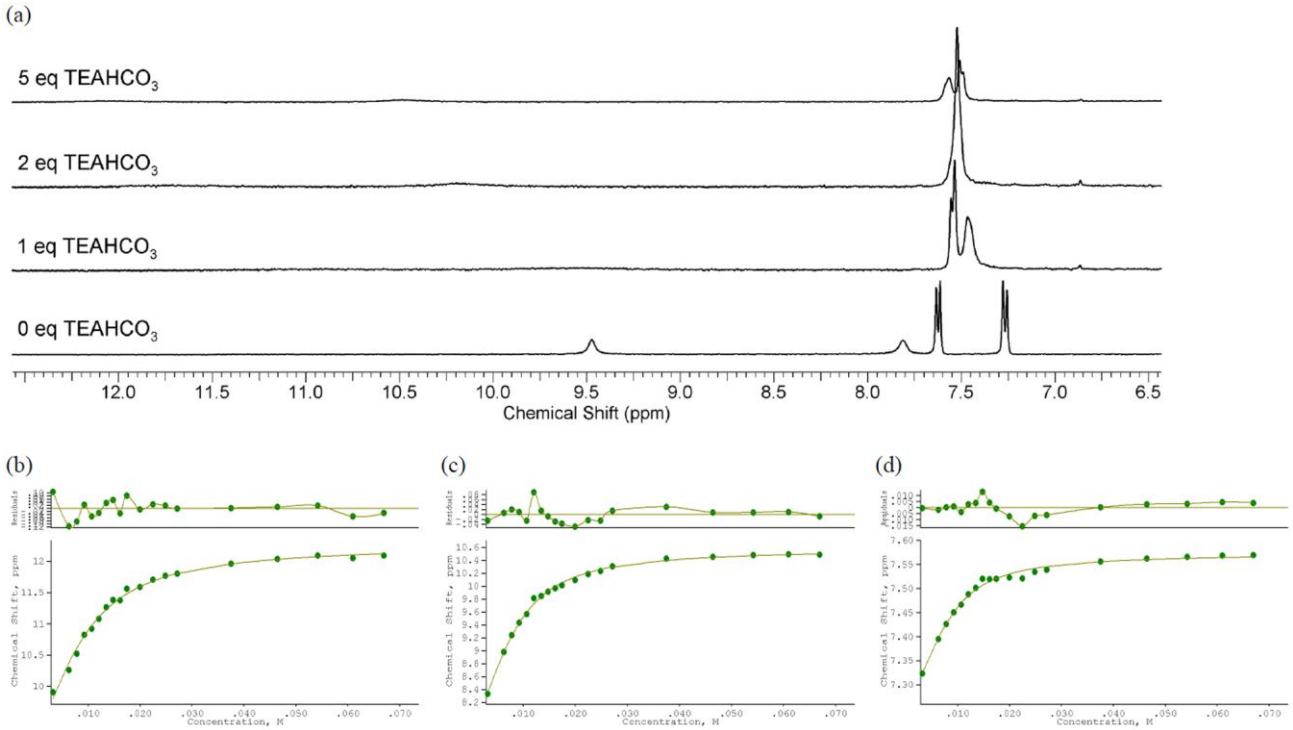


Figure S114. ^{1}H NMR titration of compound **10** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.47 ppm. K_a = 287 M^{-1} (error 10 %). (c) Fit plot for NH proton at δ = 7.81 ppm. K_a = 468 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.28 ppm. K_a = 564 M^{-1} (error 11 %).

*Interactions of compound **11** (NO_2) with various anions*

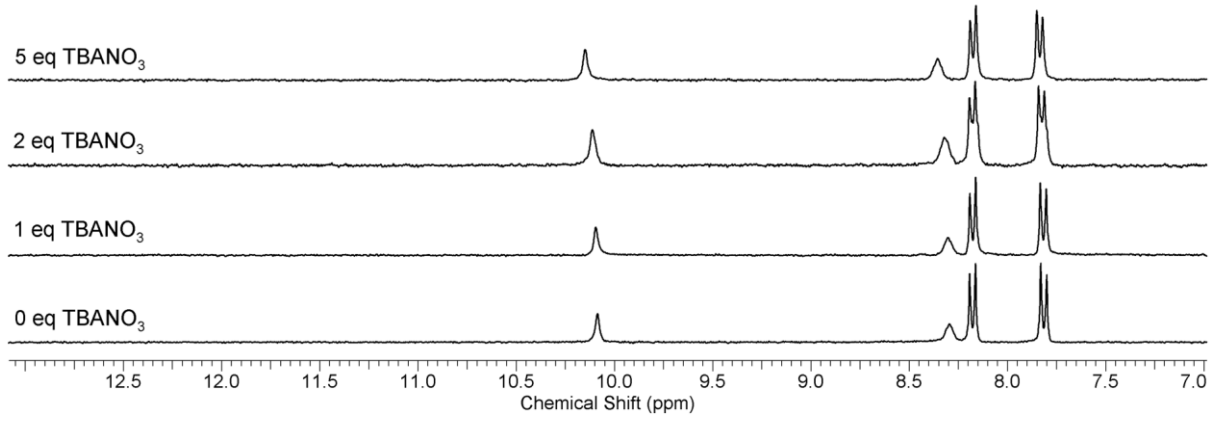


Figure S115. ^{1}H NMR titration of compound **11** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

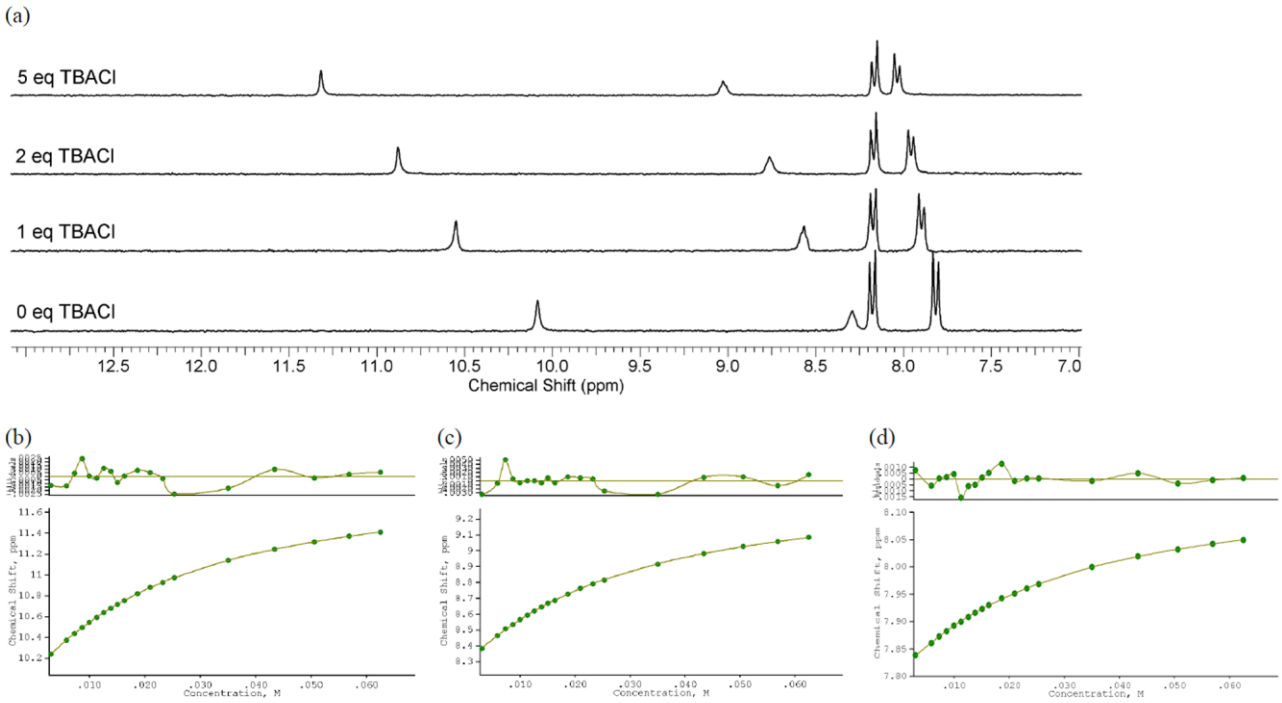


Figure S116. ^{1}H NMR titration of compound **11** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 10.07 ppm. K_a = 45.1 M^{-1} (error 5 %). (c) Fit plot for NH proton at δ = 8.27 ppm. K_a = 42.7 M^{-1} (error 1.5 %). (d) Fit plot for CH proton at δ = 7.83 ppm. K_a = 42.3 M^{-1} (error 1.5 %).

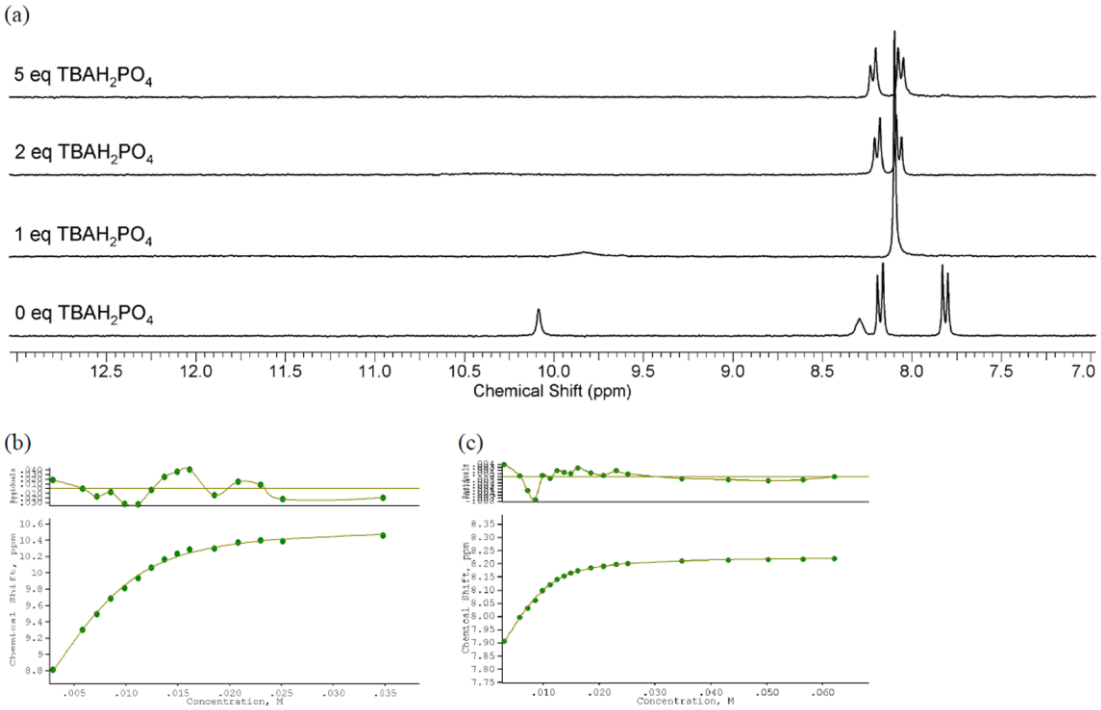


Figure S117. ^{1}H NMR titration of compound **11** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.27 ppm. K_a = 827 M^{-1} (error 10 %). (c) Fit plot for CH proton at δ = 7.83 ppm. K_a = 912 M^{-1} (error 4 %).

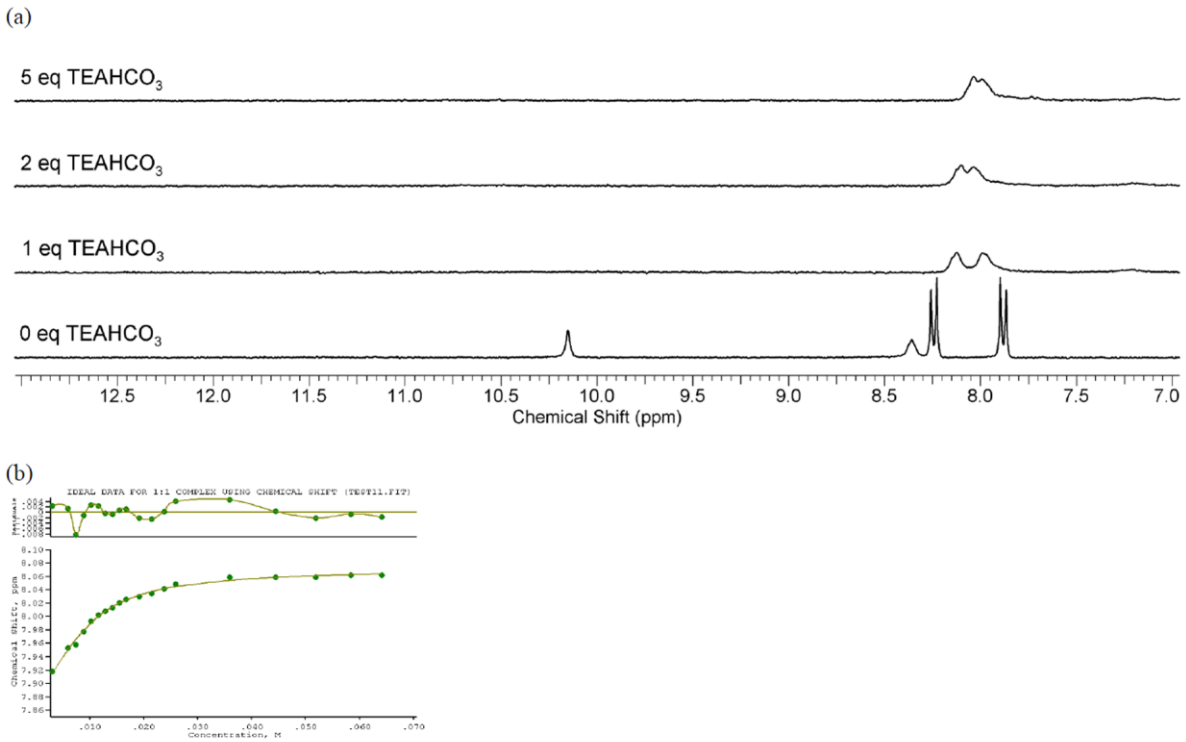


Figure S118. ^{1}H NMR titration of compound **11** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.83 ppm. K_a = 345 M^{-1} (error 8.5 %), estimate, data uncertain due to peak broadening and peak overlap.

*Interactions of compound **12** (OCOMe) with various anions*

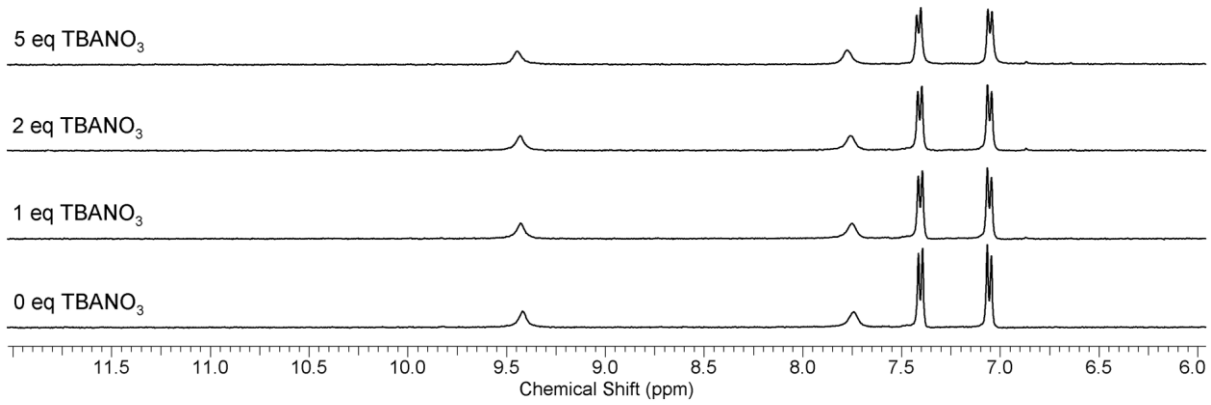


Figure S119. ^{1}H NMR titration of compound **12** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

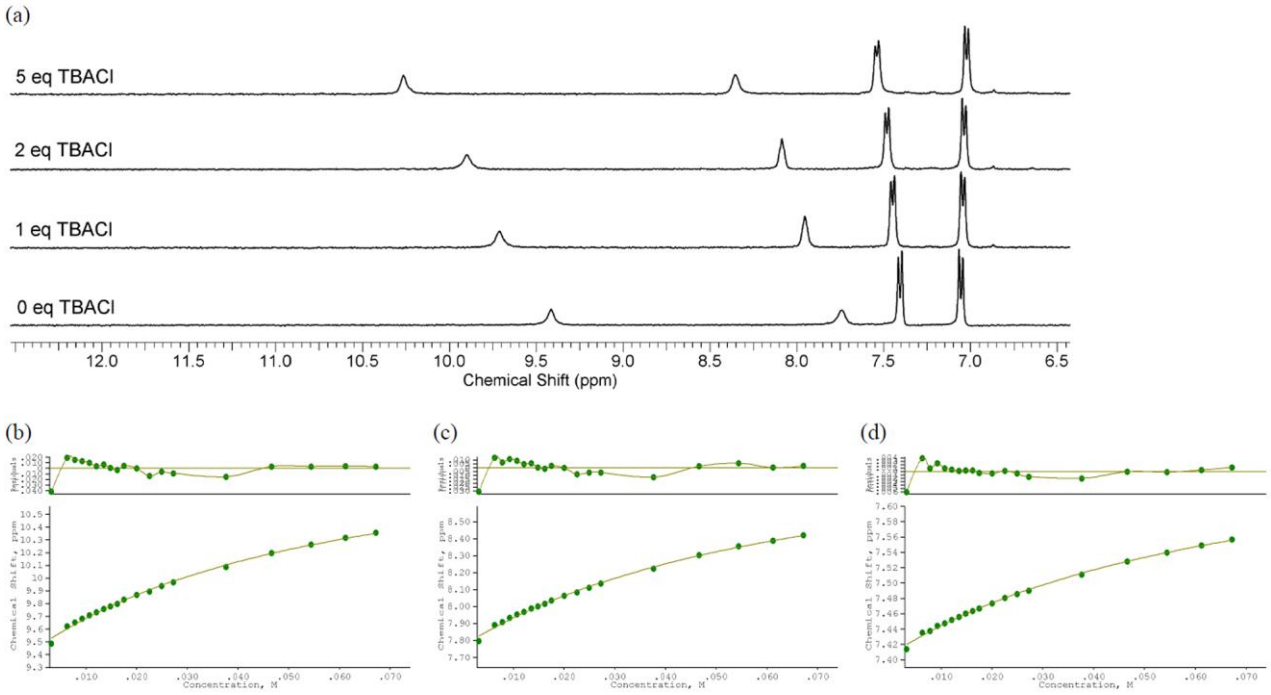


Figure S120. ^{1}H NMR titration of compound **12** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.42 ppm. K_a = 19.4 M^{-1} (error 12 %). (c) Fit plot for NH proton at δ = 7.75 ppm. K_a = 17.3 M^{-1} (error 13 %). (d) Fit plot for CH proton at δ = 7.41 ppm. K_a = 16.6 M^{-1} (error 9 %).

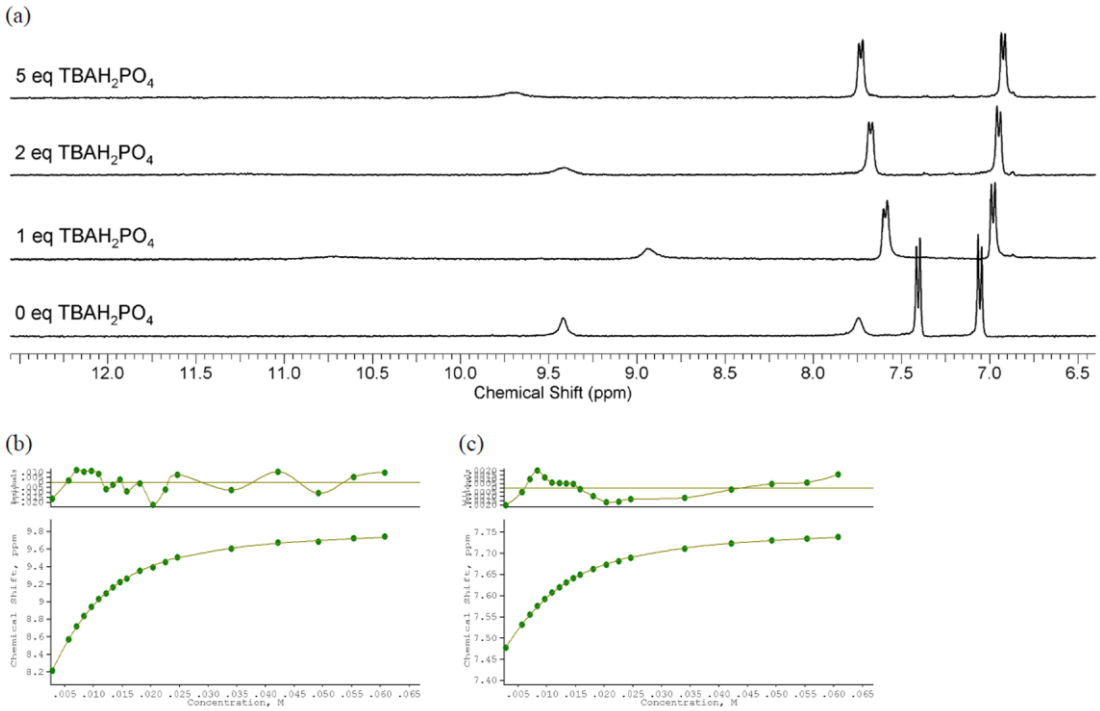


Figure S121. ^{1}H NMR titration of compound **12** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.75 ppm. K_a = 277 M^{-1} (error 3 %). (c) Fit plot for CH proton at δ = 7.41 ppm. K_a = 217 M^{-1} (error 2 %).

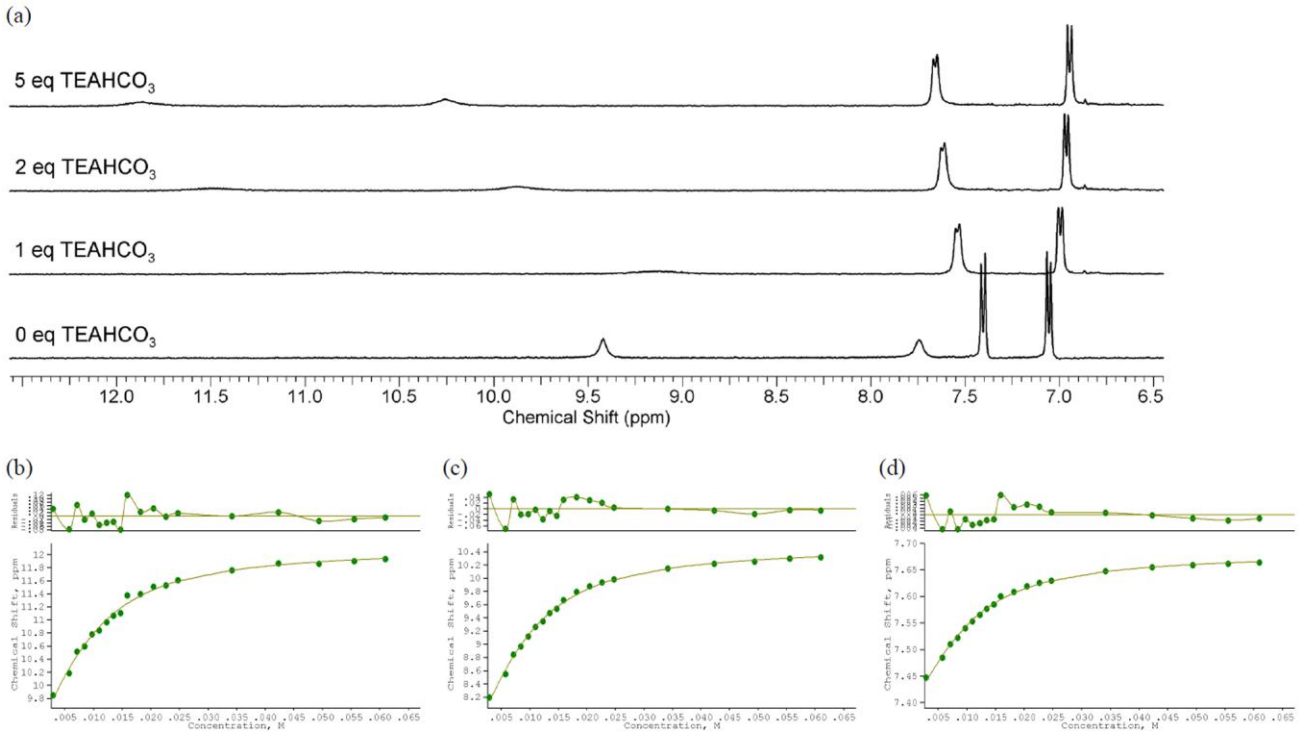


Figure S122. ^{1}H NMR titration of compound **12** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. ((b) Fit plot for NH proton at δ = 9.42 ppm. K_a = 262 M^{-1} (error 10.5 %). (c) Fit plot for NH proton at δ = 7.75 ppm. K_a = 271 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.41 ppm. K_a = 266 M^{-1} (error 4 %).

*Interactions of compound **13** (OCF_3) with various anions*

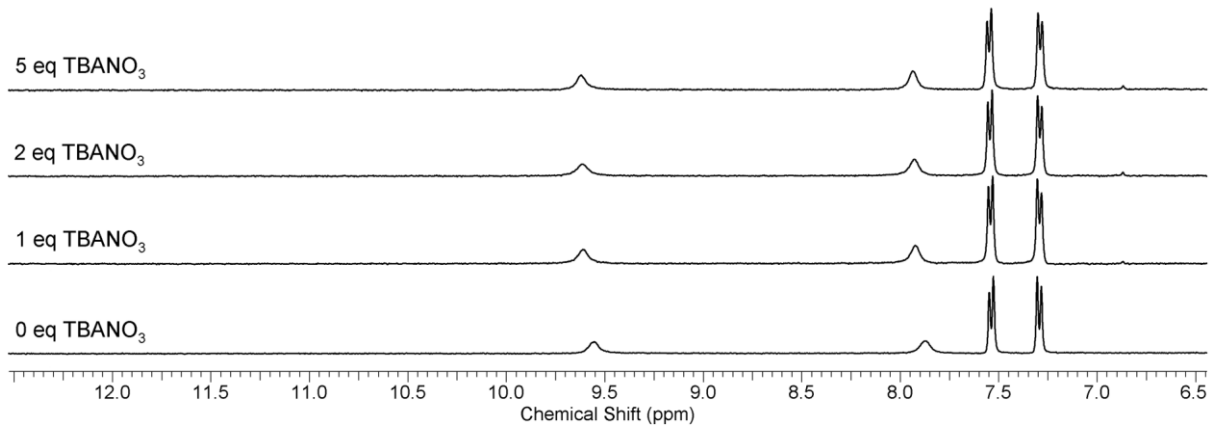


Figure S123. ^{1}H NMR titration of compound **13** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

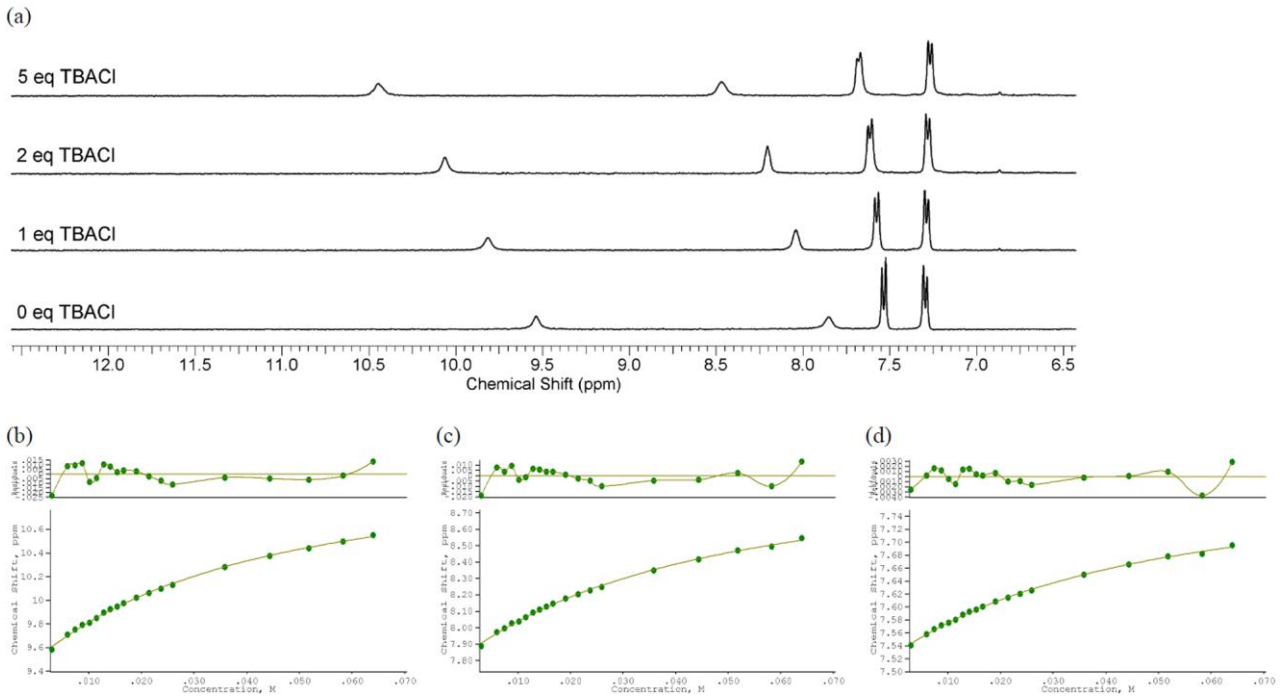


Figure S124. ^{1}H NMR titration of compound **13** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.54 ppm. K_a = 28.1 M^{-1} (error 7 %). (c) Fit plot for NH proton at δ = 7.86 ppm. K_a = 24.8 M^{-1} (error 8.5 %). (d) Fit plot for CH proton at δ = 7.54 ppm. K_a = 24.8 M^{-1} (error 7 %).

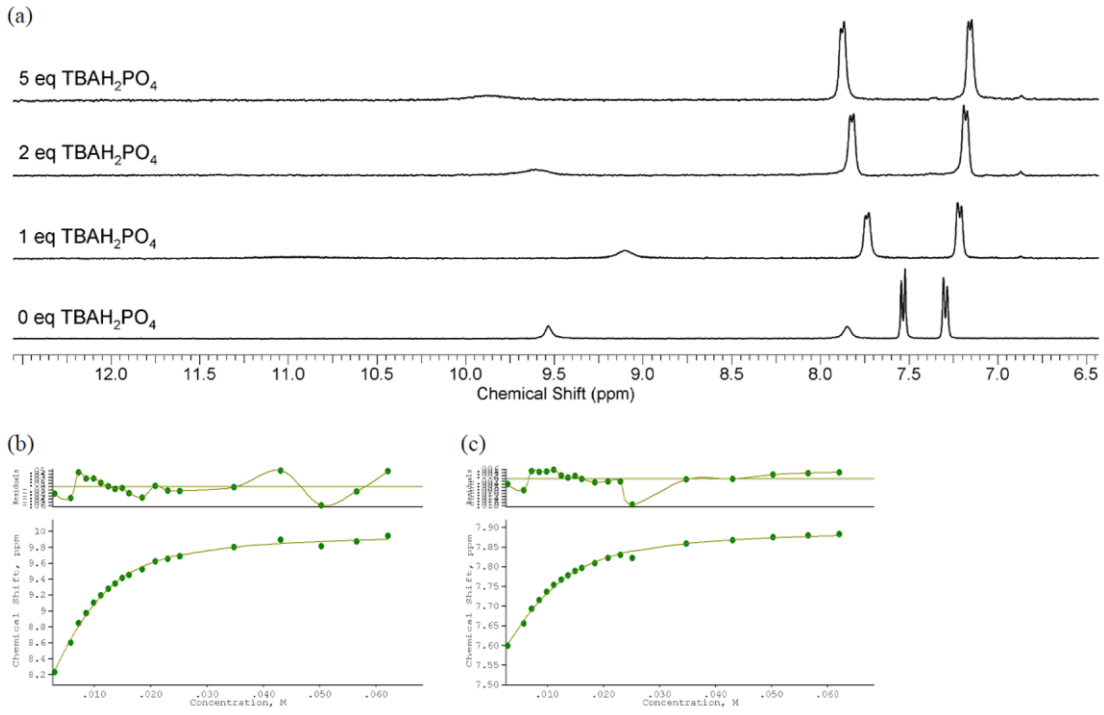


Figure S125. ^{1}H NMR titration of compound **13** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.86 ppm. K_a = 426 M^{-1} (error 9 %). (c) Fit plot for CH proton at δ = 7.54 ppm. K_a = 327 M^{-1} (error 9 %).

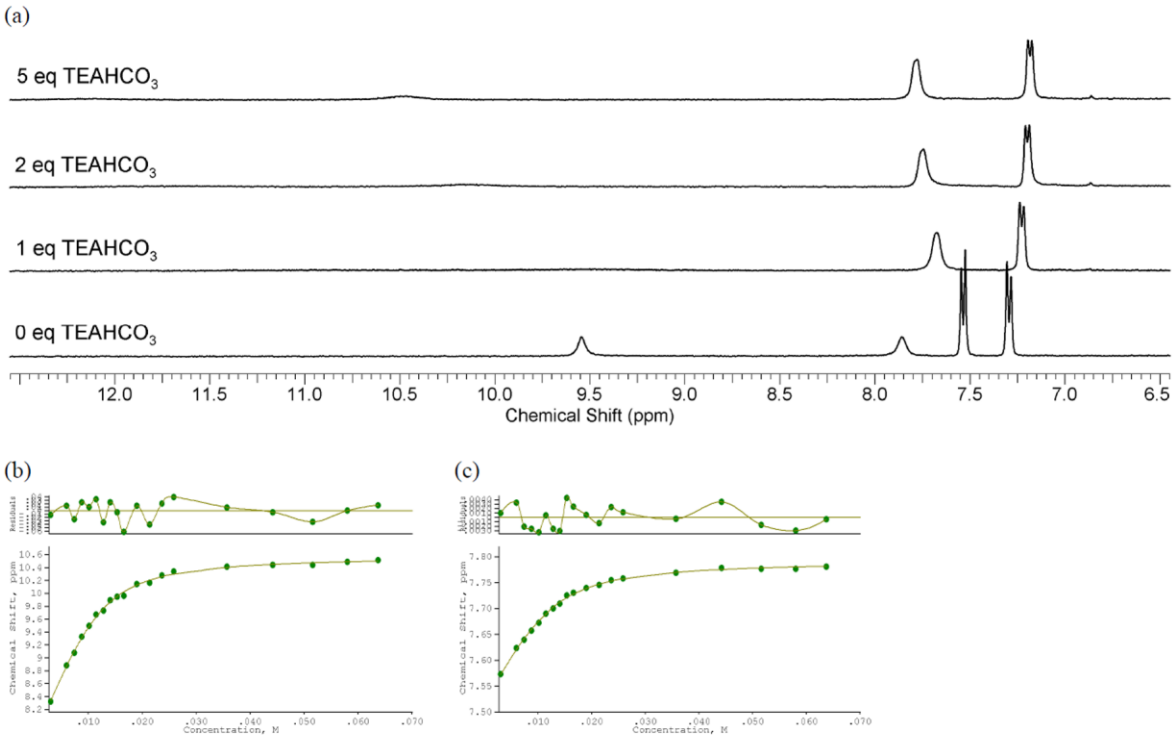


Figure S126. ^{1}H NMR titration of compound **13** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.86 ppm. K_a = 592 M^{-1} (error 5 %). (c) Fit plot for CH proton at δ = 7.54 ppm. K_a = 439 M^{-1} (error 3 %).

*Interactions of compound **14** (OEt) with various anions*

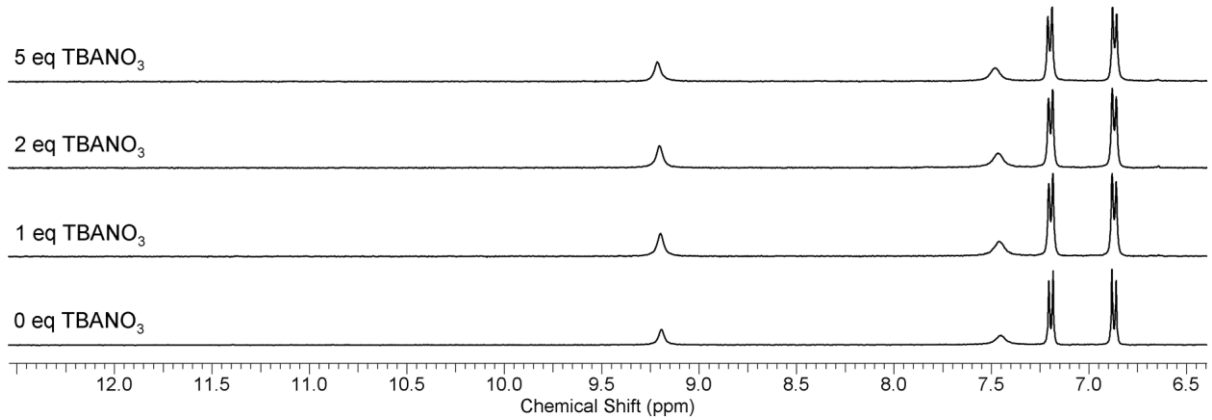


Figure S127. ^{1}H NMR titration of compound **14** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

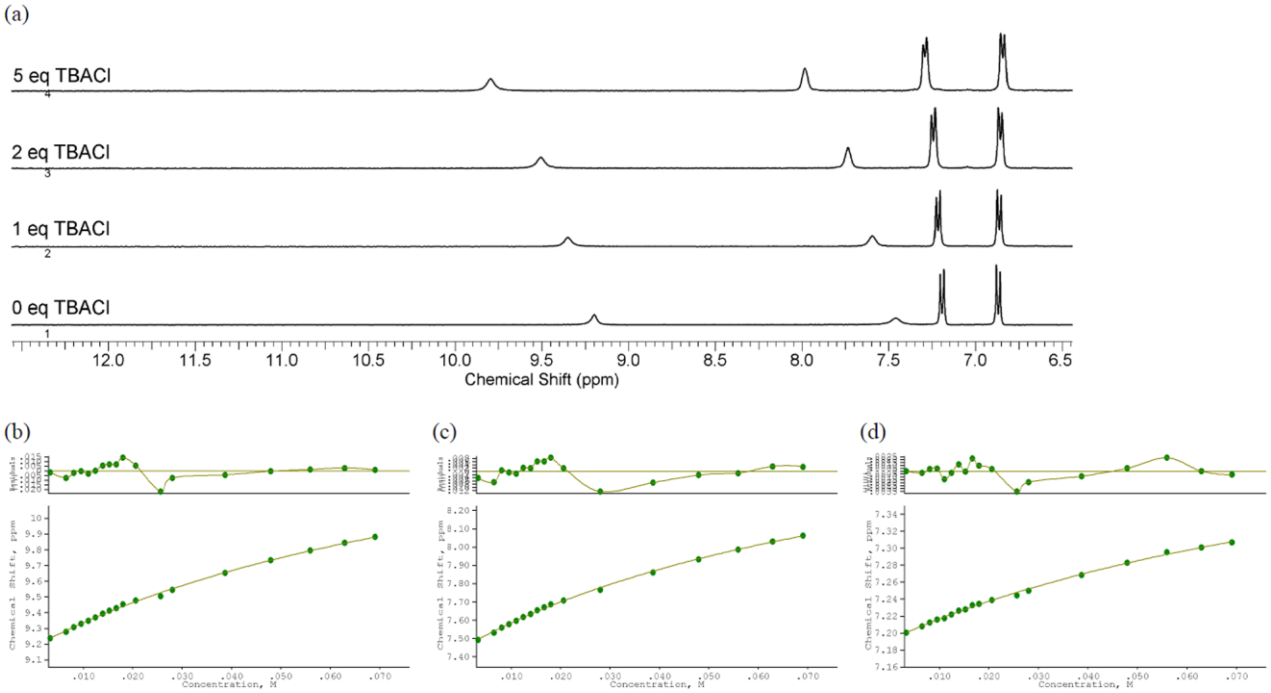


Figure S128. ^{1}H NMR titration of compound **14** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.19 ppm. K_a = 10.8 M^{-1} (error 12 %). (c) Fit plot for NH proton at δ = 7.46 ppm. K_a = 12.3 M^{-1} (error 9.5 %). (d) Fit plot for CH proton at δ = 7.20 ppm. K_a = 9.5 M^{-1} (error 10.5 %).

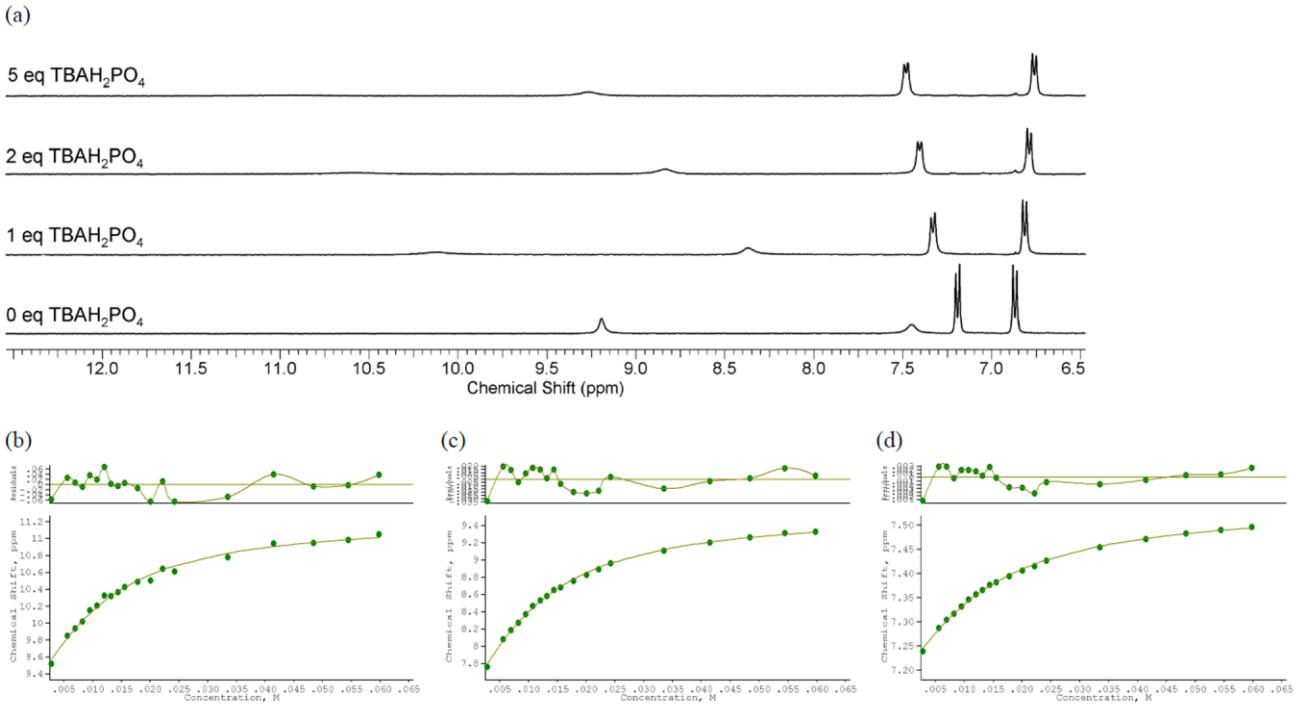


Figure S129. ^{1}H NMR titration of compound **14** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.19 ppm. K_a = 144 M^{-1} (error 12.5 %). (c) Fit plot for NH proton at δ = 7.46 ppm. K_a = 141 M^{-1} (error 5 %). (d) Fit plot for CH proton at δ = 7.20 ppm. K_a = 109 M^{-1} (error 5 %).

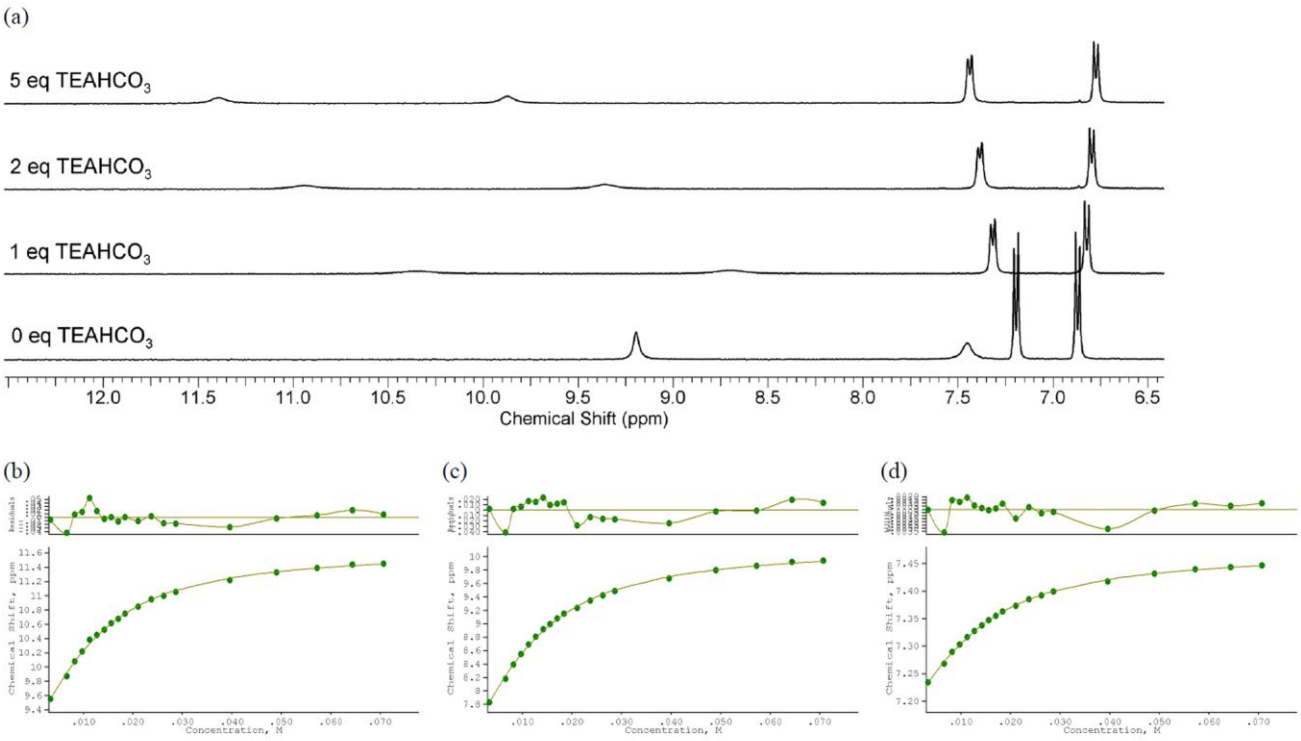


Figure S130. ^{1}H NMR titration of compound **14** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.19 ppm. K_a = 172 M^{-1} (error 5 %). (c) Fit plot for NH proton at δ = 7.46 ppm. K_a = 163 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.20 ppm. K_a = 141 M^{-1} (error 3 %).

*Interactions of compound **15** (OMe) with various anions*

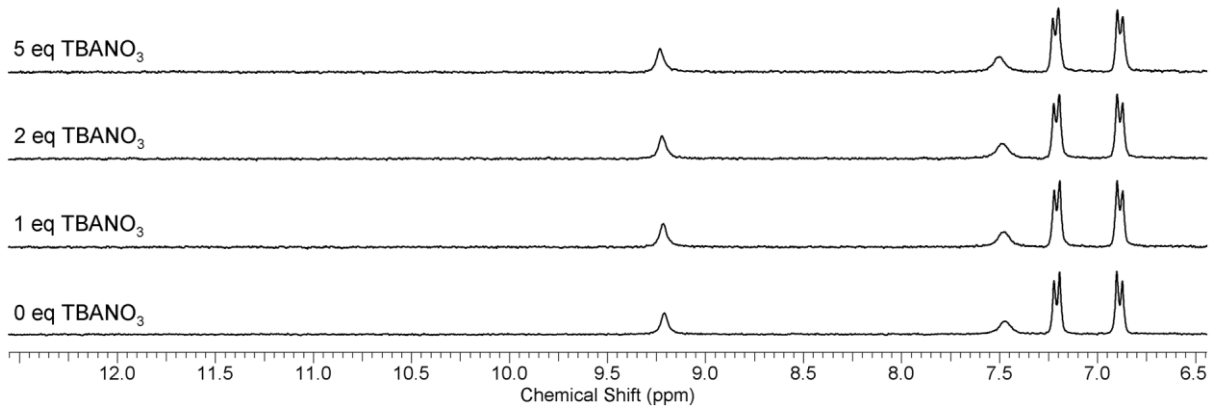


Figure S131. ^{1}H NMR titration of compound **15** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

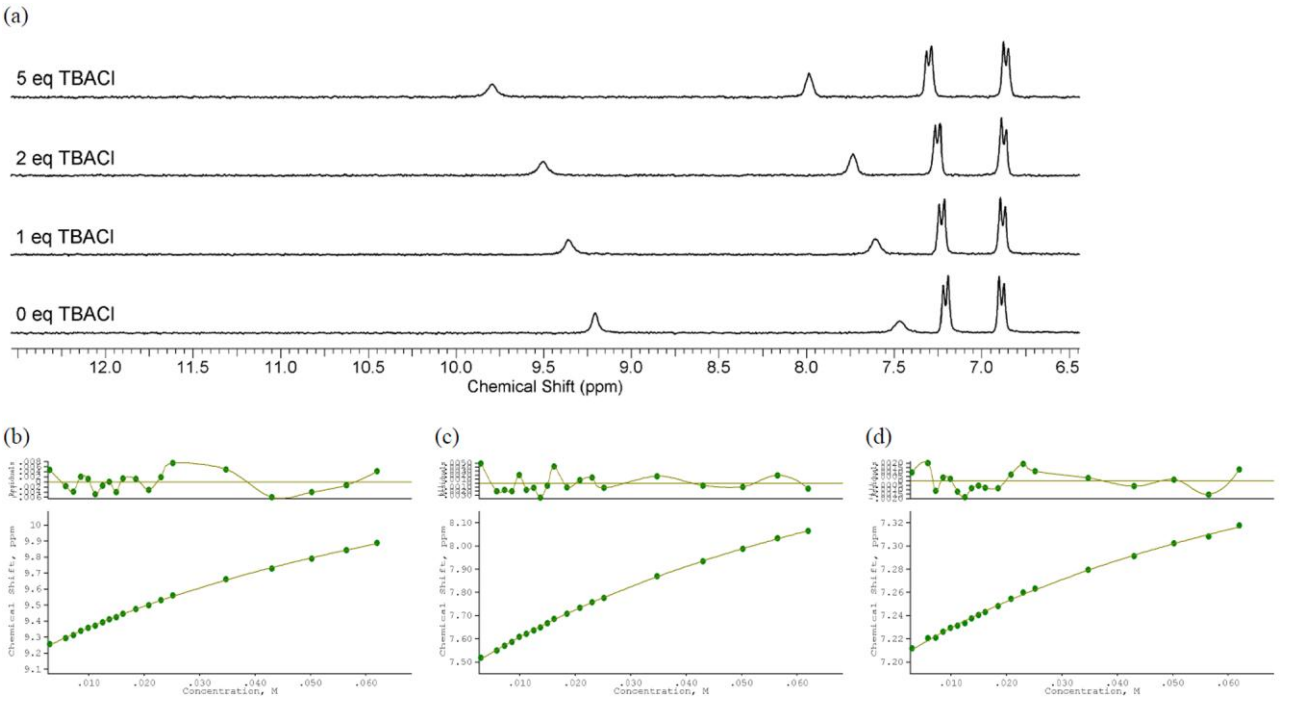


Figure S132. ^{1}H NMR titration of compound **15** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.20 ppm. K_a = 10.6 M^{-1} (error 6 %). (c) Fit plot for NH proton at δ = 7.48 ppm. K_a = 10.7 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.22 ppm. K_a = 11.4 M^{-1} (error 9 %).

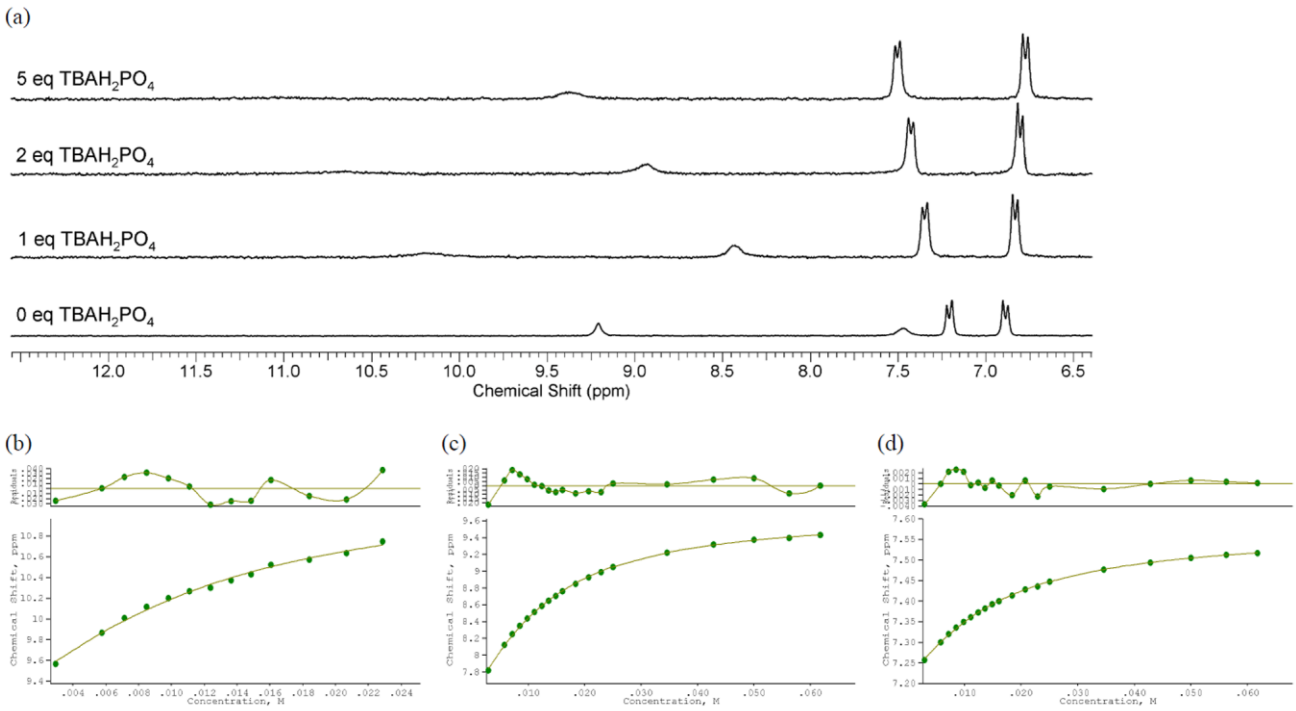


Figure S133. ^{1}H NMR titration of compound **15** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.20 ppm. K_a = 162 M^{-1} (error 21 %). (c) Fit plot for NH proton at δ = 7.48 ppm. K_a = 131 M^{-1} (error 3 %). (d) Fit plot for CH proton at δ = 7.22 ppm. K_a = 102 M^{-1} (error 2 %).

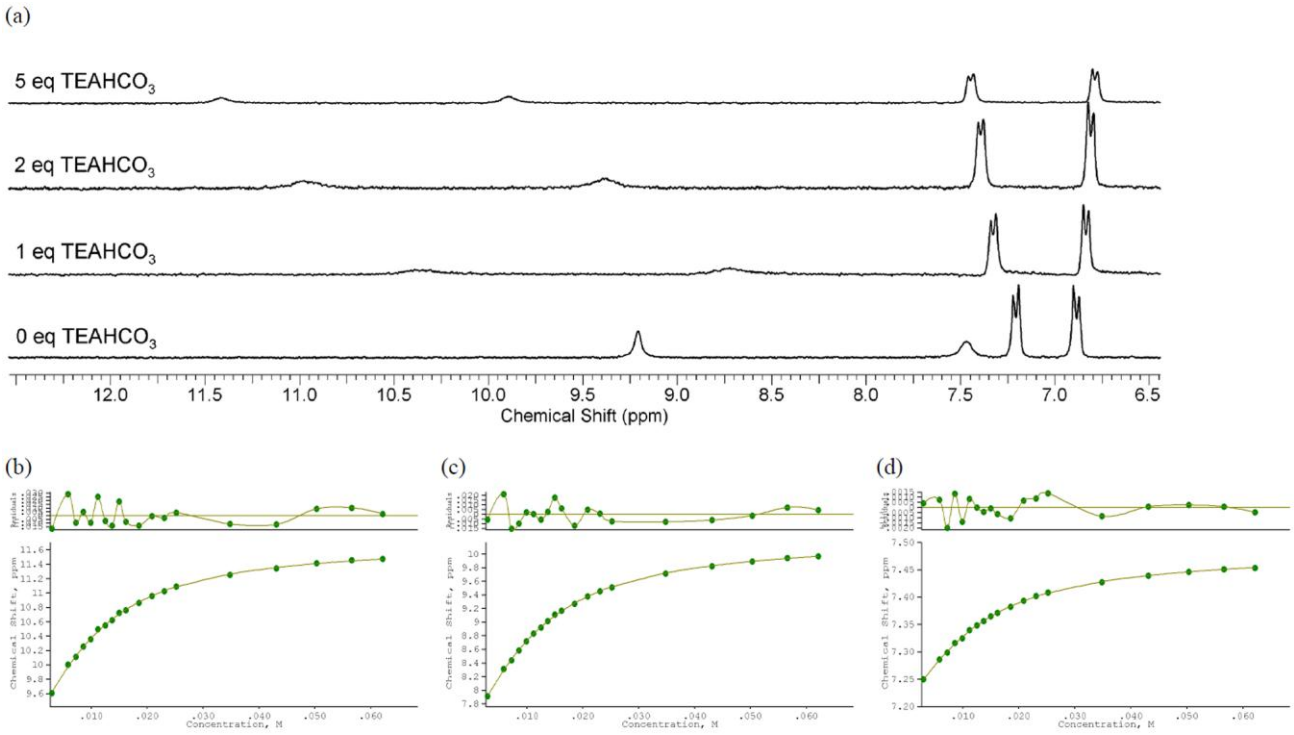


Figure S134. ^{1}H NMR titration of compound **15** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.20 ppm. K_a = 171 M^{-1} (error 3 %). (c) Fit plot for NH proton at δ = 7.48 ppm. K_a = 157 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.22 ppm. K_a = 145 M^{-1} (error 2 %).

*Interactions of compound **16** (SMe) with various anions*

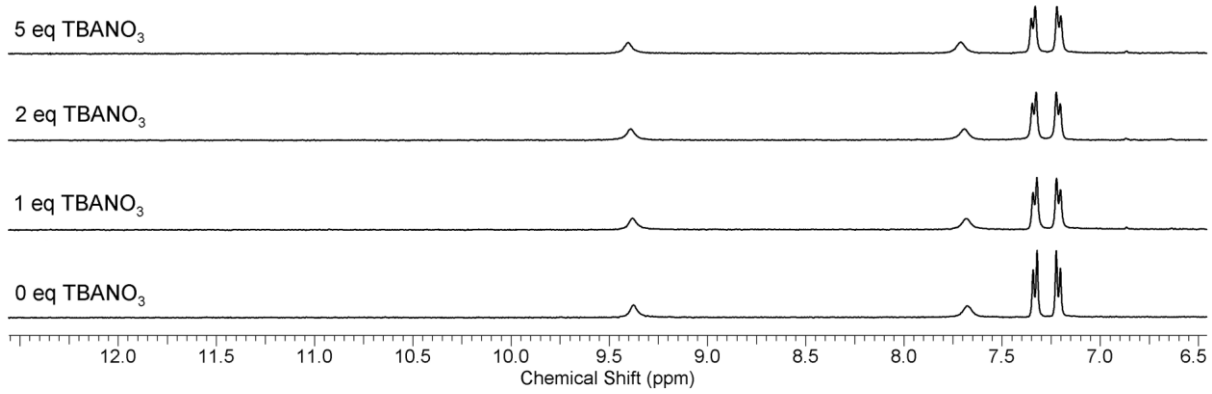


Figure S135. ^{1}H NMR titration of compound **16** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

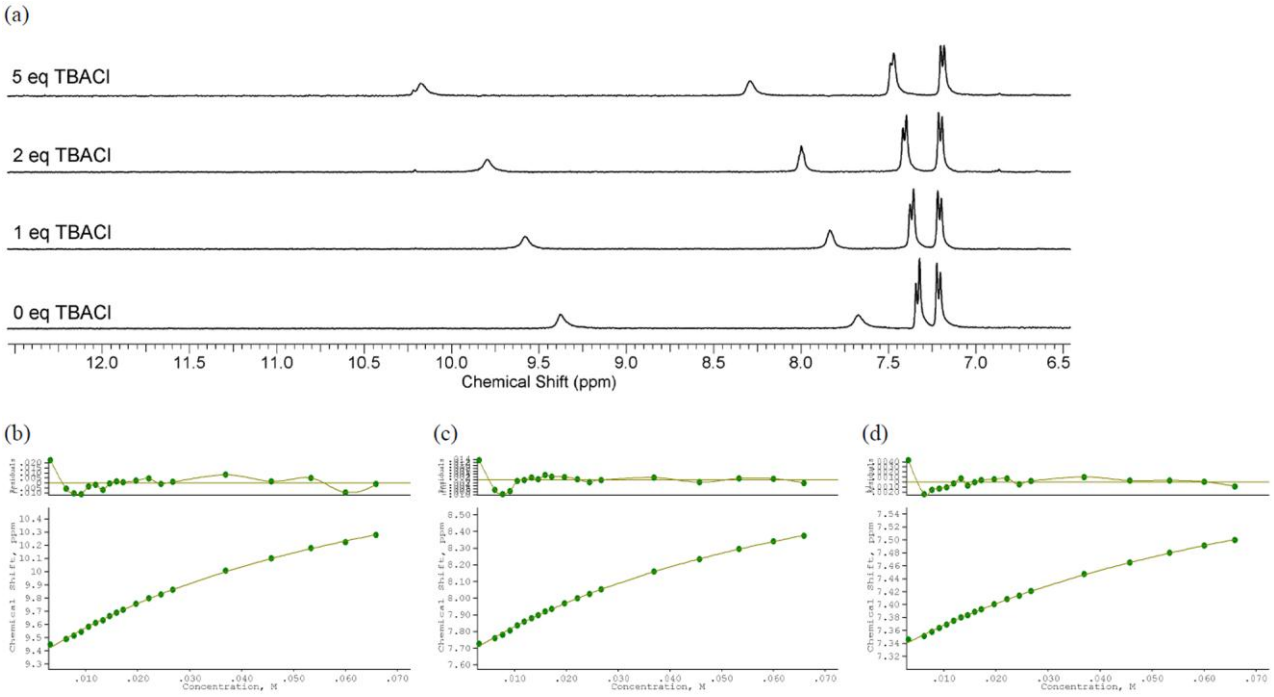


Figure S136. ^{1}H NMR titration of compound **16** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.39 ppm. K_a = 15.0 M^{-1} (error 7 %). (c) Fit plot for NH proton at δ = 7.69 ppm. K_a = 14.8 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.35 ppm. K_a = 12.2 M^{-1} (error 8 %).

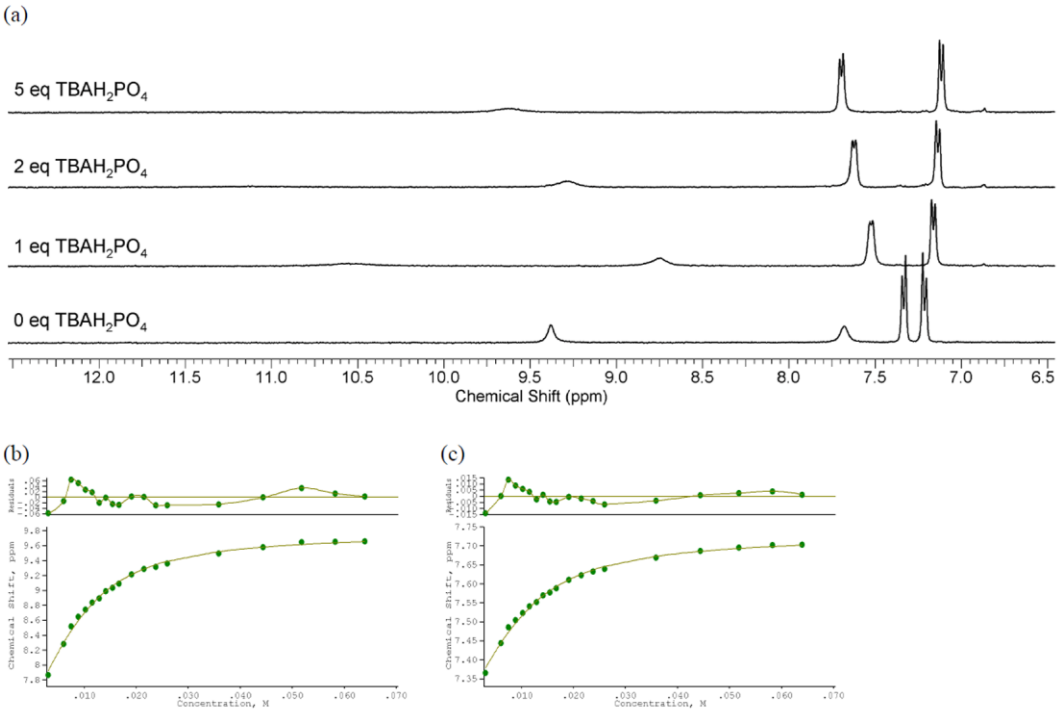


Figure S137. ^{1}H NMR titration of compound **16** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.69 ppm. K_a = 285 M^{-1} (error 8 %). (c) Fit plot for CH proton at δ = 7.35 ppm. K_a = 233 M^{-1} (error 6 %).

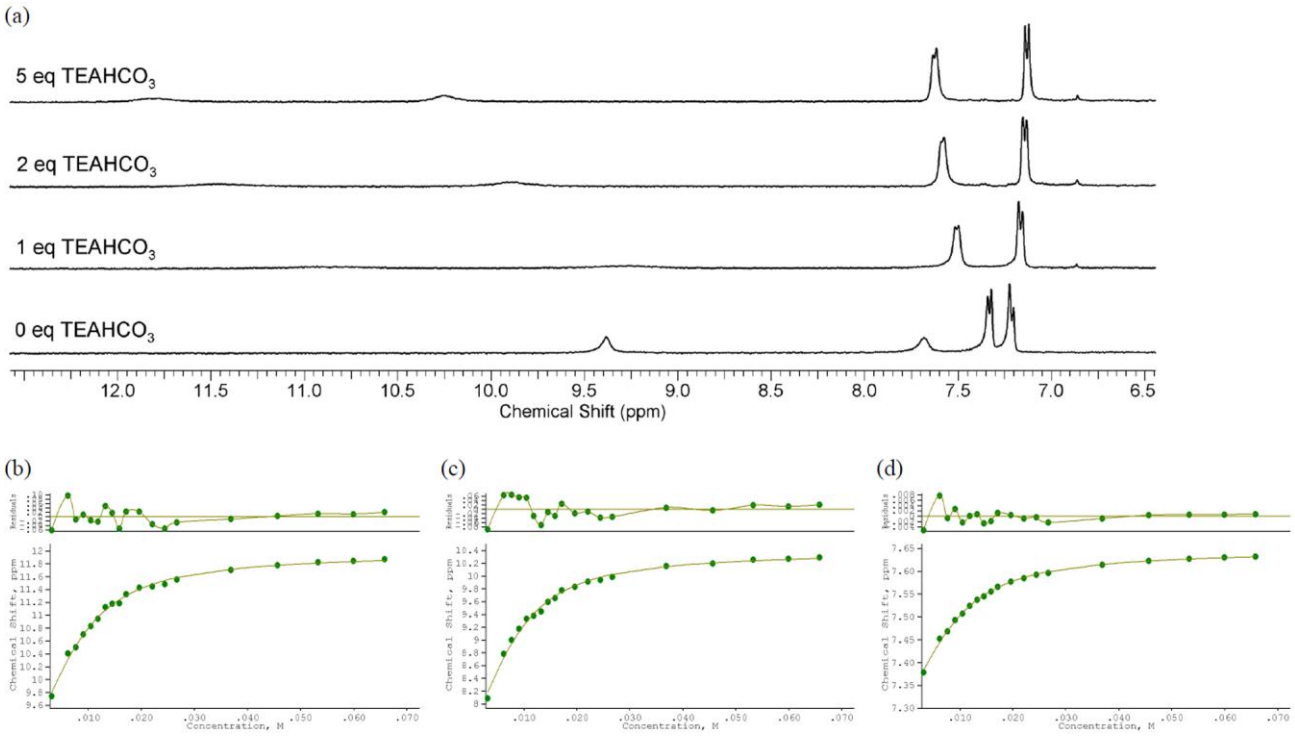


Figure S138. ^{1}H NMR titration of compound **16** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.39 ppm. K_a = 324 M^{-1} (error 8 %). (c) Fit plot for NH proton at δ = 7.69 ppm. K_a = 364 M^{-1} (error 9 %). (d) Fit plot for CH proton at δ = 7.35 ppm. K_a = 316 M^{-1} (error 4 %).

*Interactions of compound **17** (SO_2Me) with various anions*

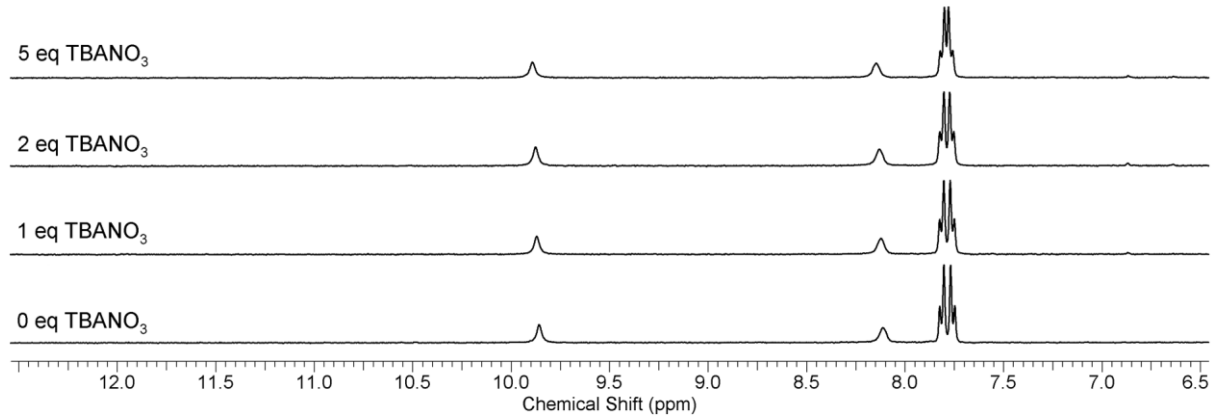


Figure S139. ^{1}H NMR titration of compound **17** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

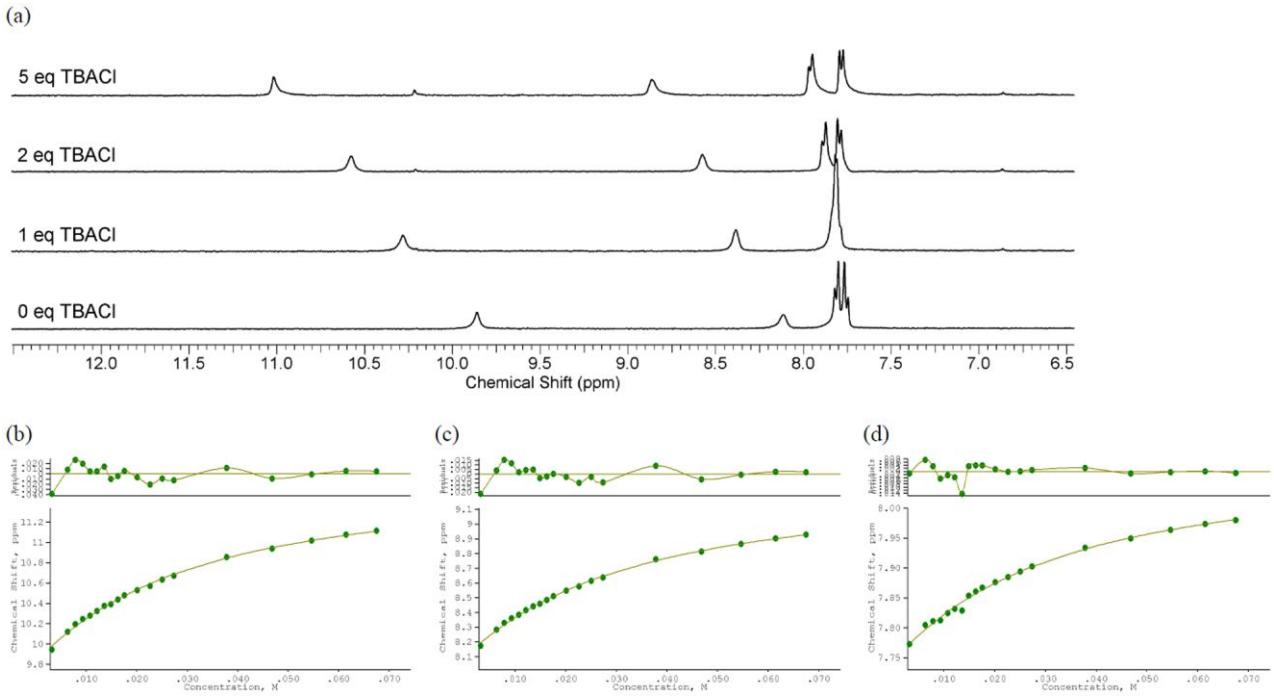


Figure S140. ^{1}H NMR titration of compound **17** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.86 ppm. K_a = 40.3 M^{-1} (error 8 %). (c) Fit plot for NH proton at δ = 8.12 ppm. K_a = 37.8 M^{-1} (error 6.5 %). (d) Fit plot for CH proton at δ = 7.76 ppm. K_a = 38.1 M^{-1} (error 10.5 %).

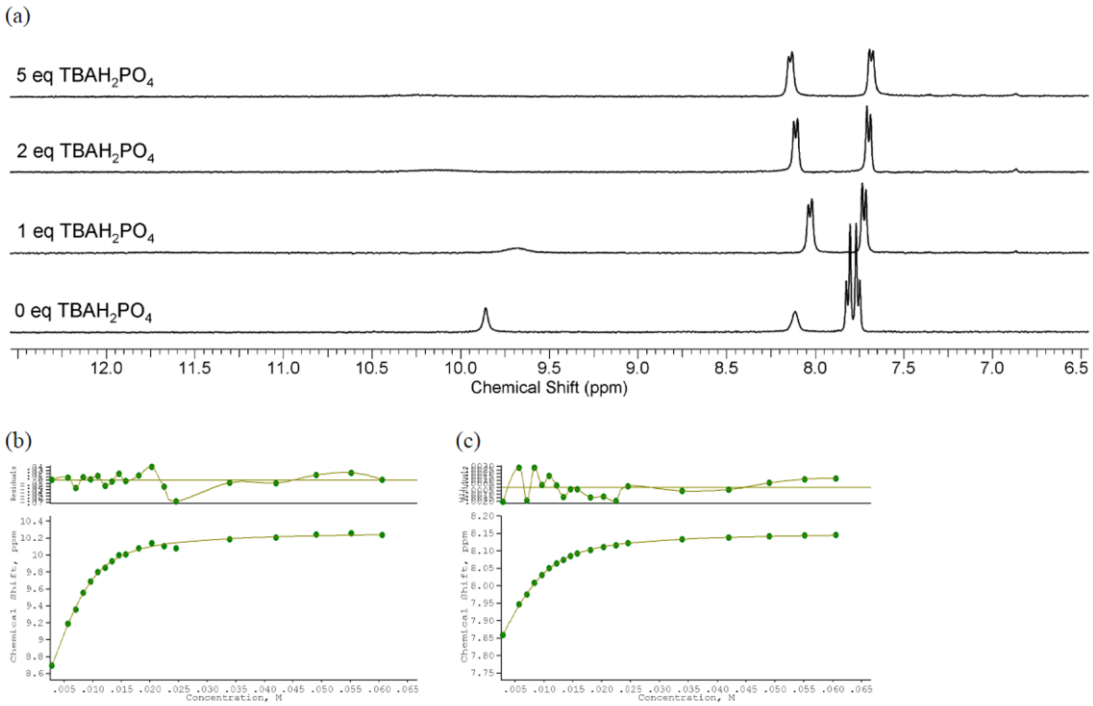


Figure S141. ^{1}H NMR titration of compound **17** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 8.12 ppm. K_a = 943 M^{-1} (error 7 %). (c) Fit plot for CH proton at δ = 7.76 ppm. K_a = 684 M^{-1} (error 2.5 %).

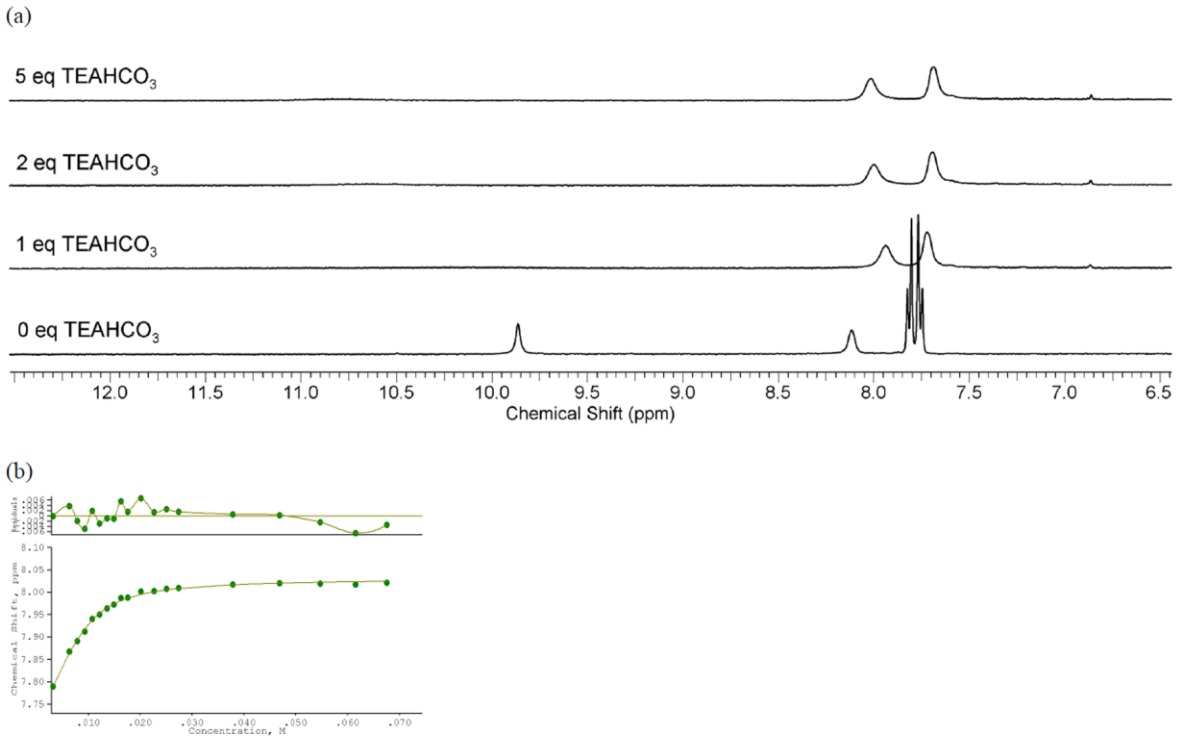


Figure S142. ^{1}H NMR titration of compound **17** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for CH proton at δ = 7.76 ppm. K_a = 702 M^{-1} (error 2 %).

*Interactions of compound **18** (Me) with various anions*

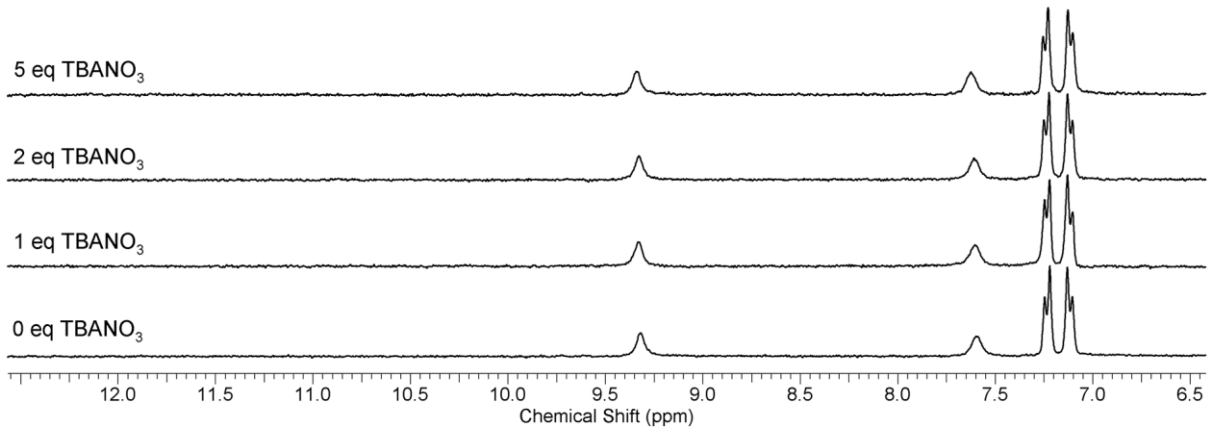


Figure S143. ^{1}H NMR titration of compound **18** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

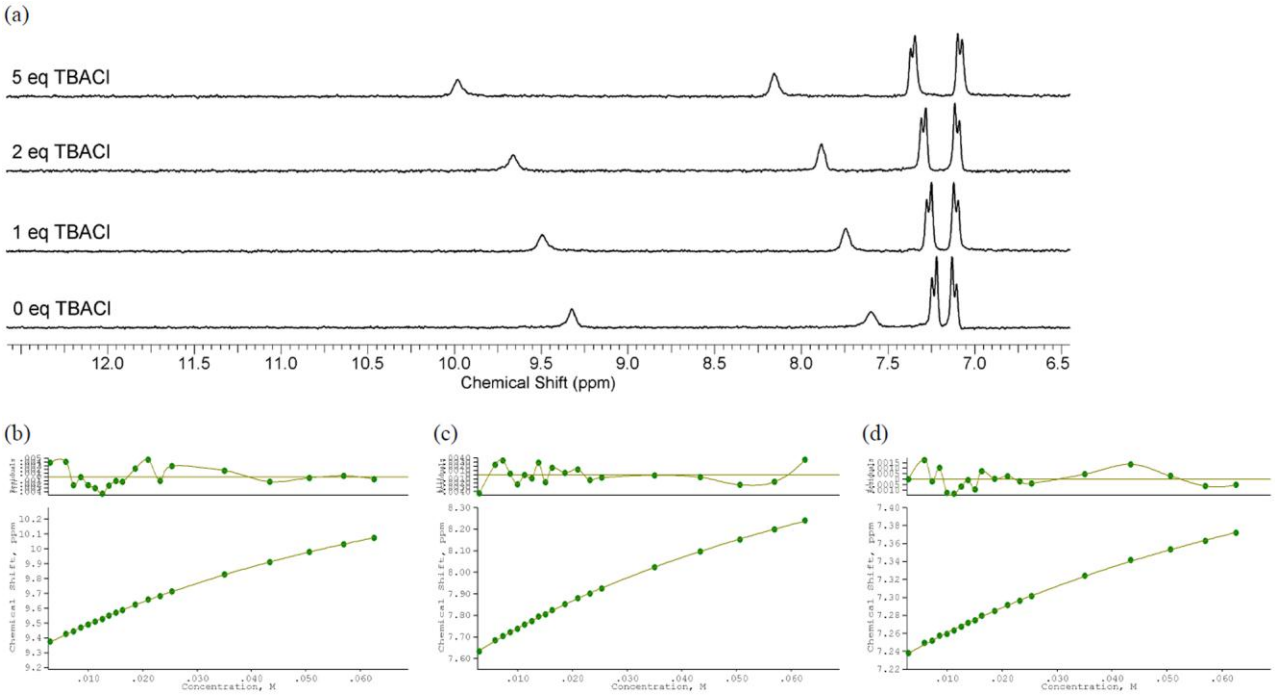


Figure S144. ^{1}H NMR titration of compound **18** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.32 ppm. K_a = 11.1 M^{-1} (error 4 %). (c) Fit plot for NH proton at δ = 7.59 ppm. K_a = 10.9 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.24 ppm. K_a = 9.8 M^{-1} (error 5 %).

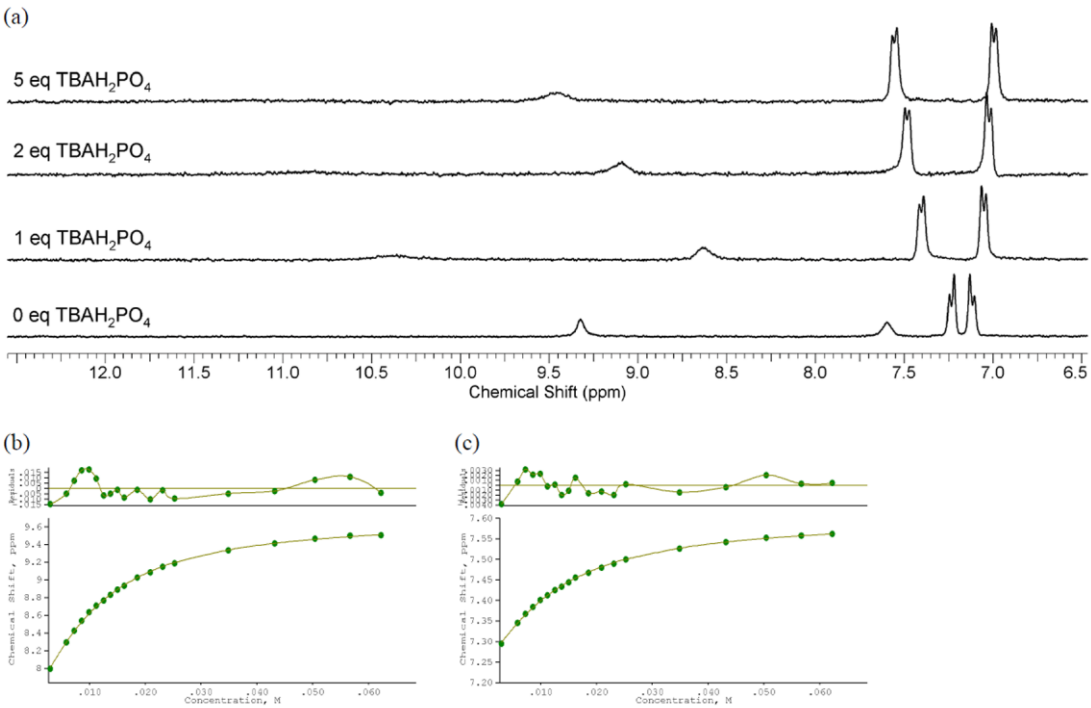


Figure S145. ^{1}H NMR titration of compound **18** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 7.59 ppm. K_a = 164 M^{-1} (error 3 %). (c) Fit plot for CH proton at δ = 7.24 ppm. K_a = 134 M^{-1} (error 3 %).

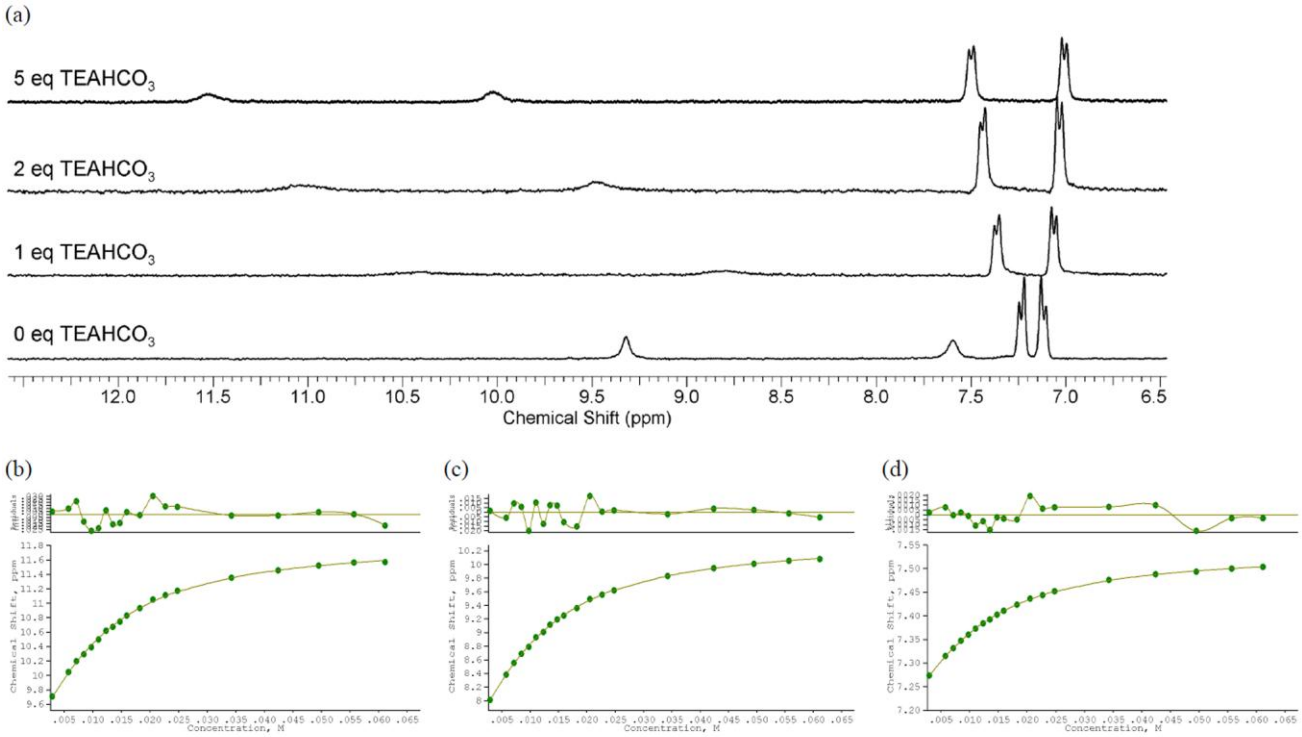


Figure S146. ^{1}H NMR titration of compound **18** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.32 ppm. K_a = 147 M^{-1} (error 4 %). (c) Fit plot for NH proton at δ = 7.59 ppm. K_a = 152 M^{-1} (error 2 %). (d) Fit plot for CH proton at δ = 7.24 ppm. K_a = 147 M^{-1} (error 1.5 %).

*Interactions of compound **19** (Et) with various anions*

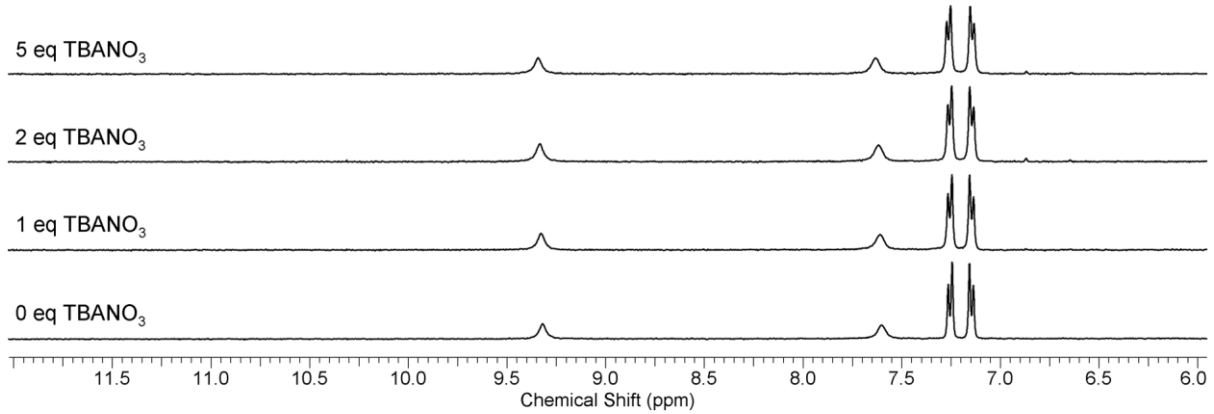


Figure S147. ^{1}H NMR titration of compound **19** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

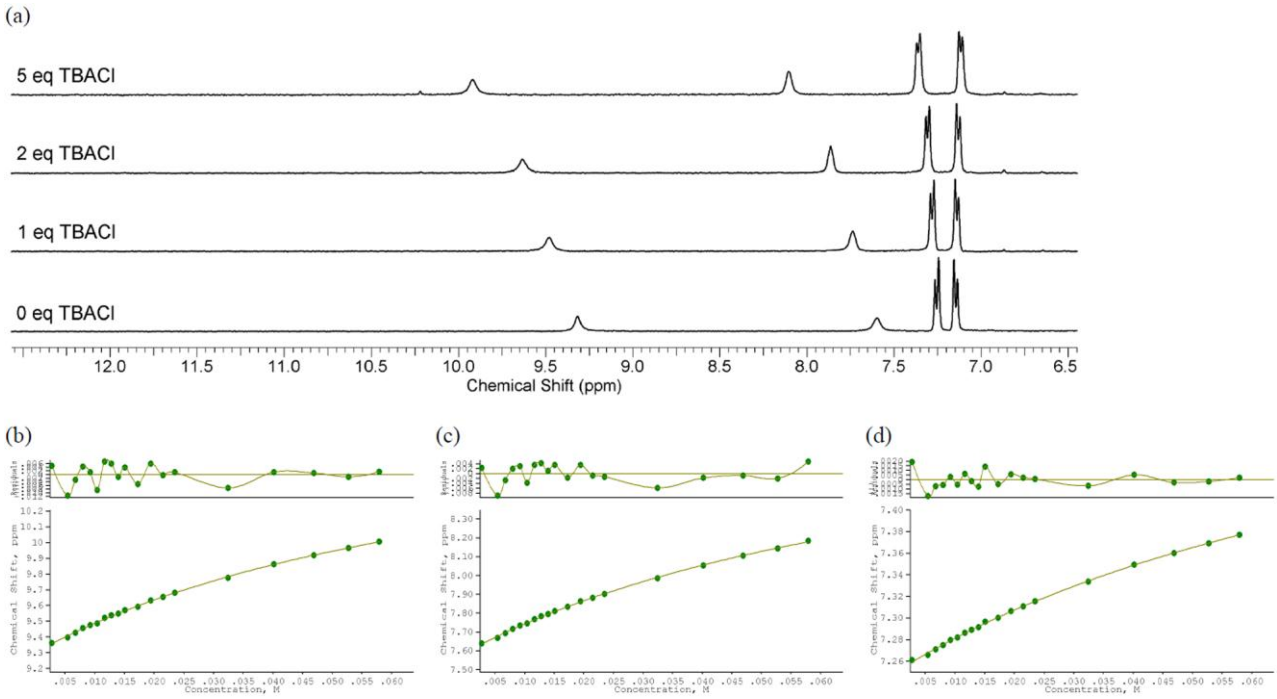


Figure S148. ^{1}H NMR titration of compound **19** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 14.4 M^{-1} (error 7.5 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 13.4 M^{-1} (error 6.5 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 10.7 M^{-1} (error 7.5 %).

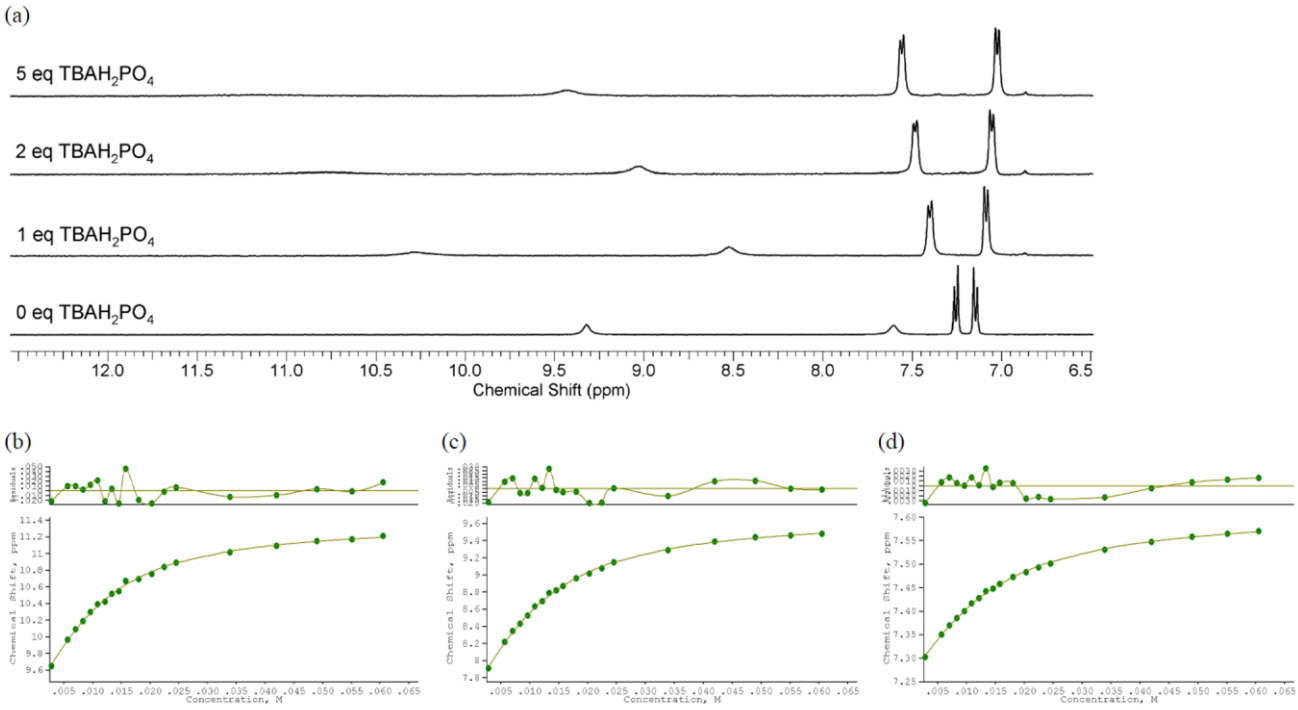


Figure S149. ^{1}H NMR titration of compound **19** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 191 M^{-1} (error 6 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 169 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 135 M^{-1} (error 3 %).

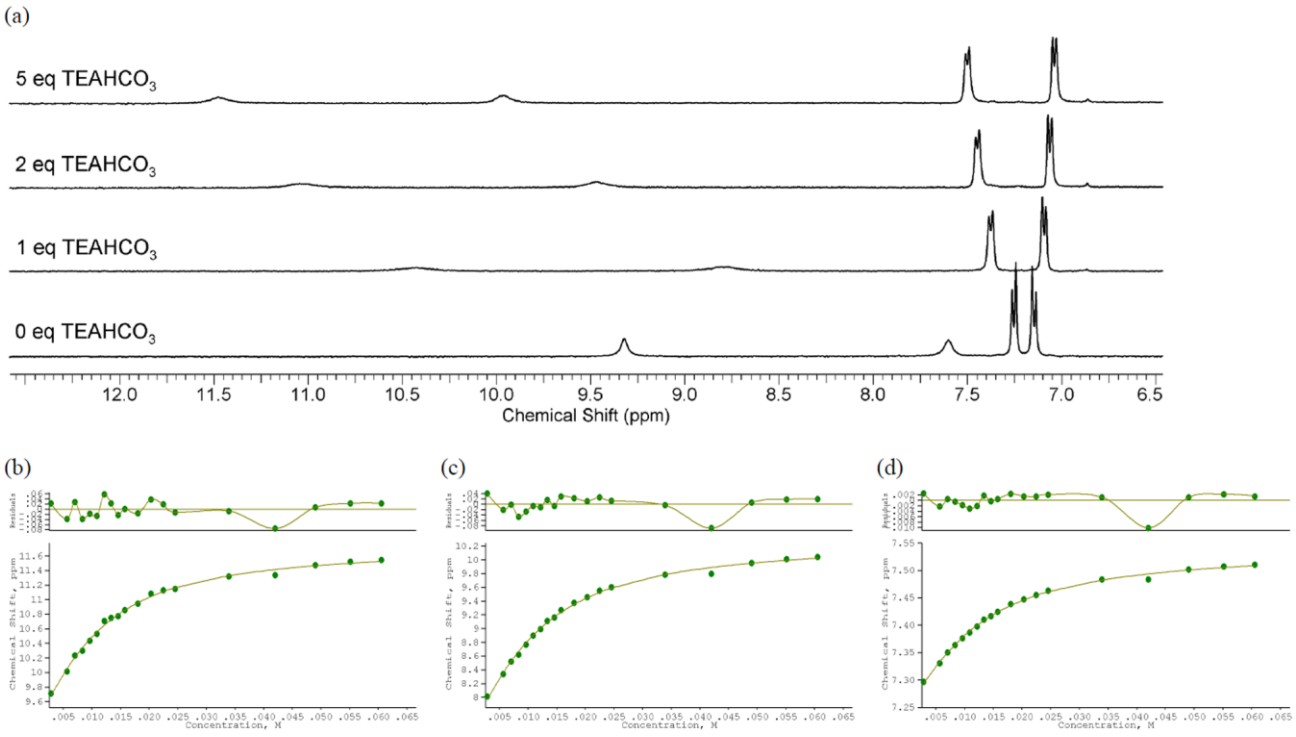


Figure S150. ^{1}H NMR titration of compound **19** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 188 M^{-1} (error 8 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 166 M^{-1} (error 7 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 150 M^{-1} (error 6 %).

*Interactions of compound **20** (Pr) with various anions*

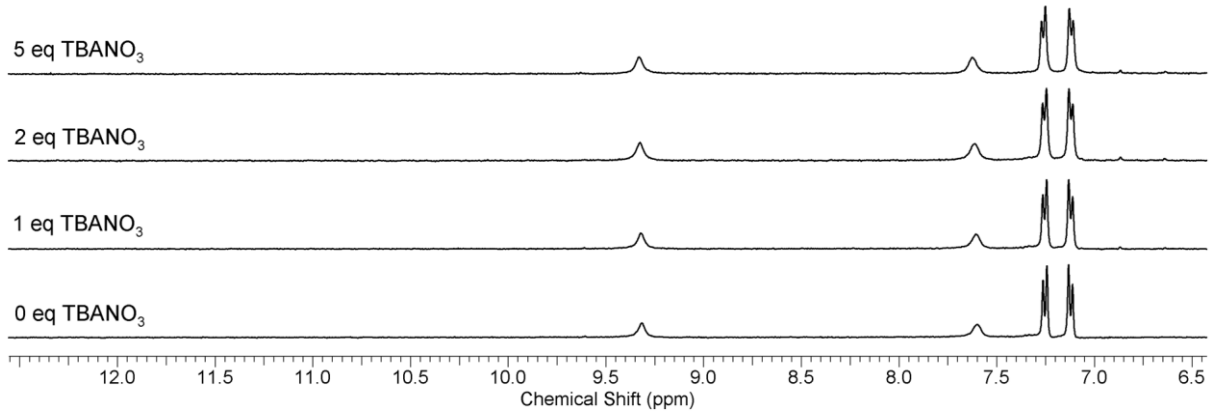


Figure S151. ^{1}H NMR titration of compound **20** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

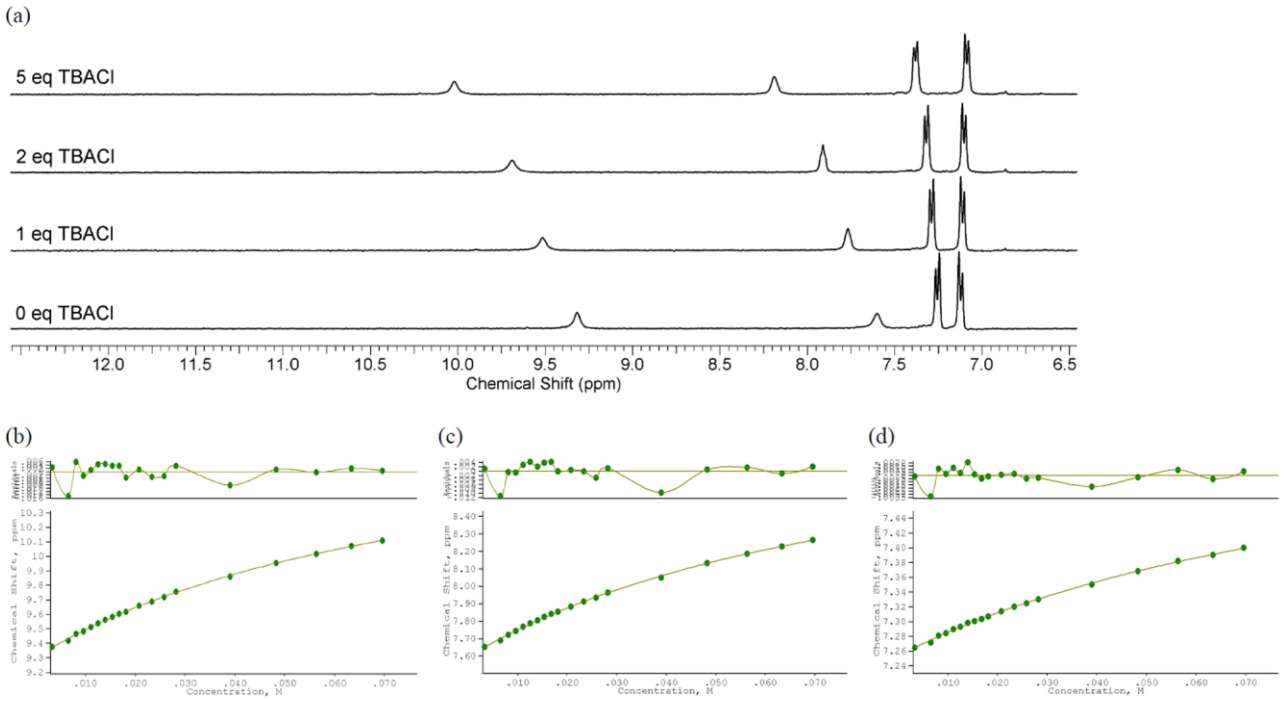


Figure S152. ^{1}H NMR titration of compound **20** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 13.9 M^{-1} (error 6 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 13.1 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.25 ppm. K_a = 10.4 M^{-1} (error 7.5 %).

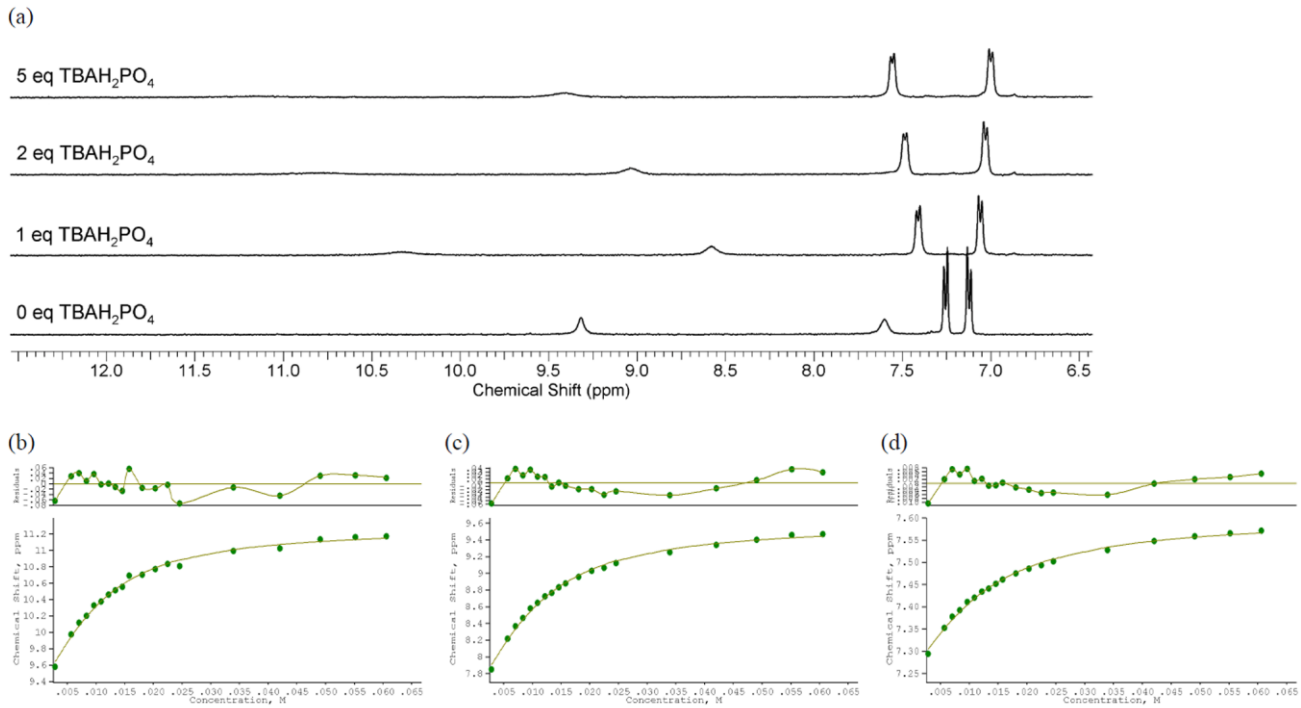


Figure S153. ^{1}H NMR titration of compound **20** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 240 M^{-1} (error 10.5 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 208 M^{-1} (error 8.5 %). (d) Fit plot for CH proton at δ = 7.25 ppm. K_a = 158 M^{-1} (error 8 %).

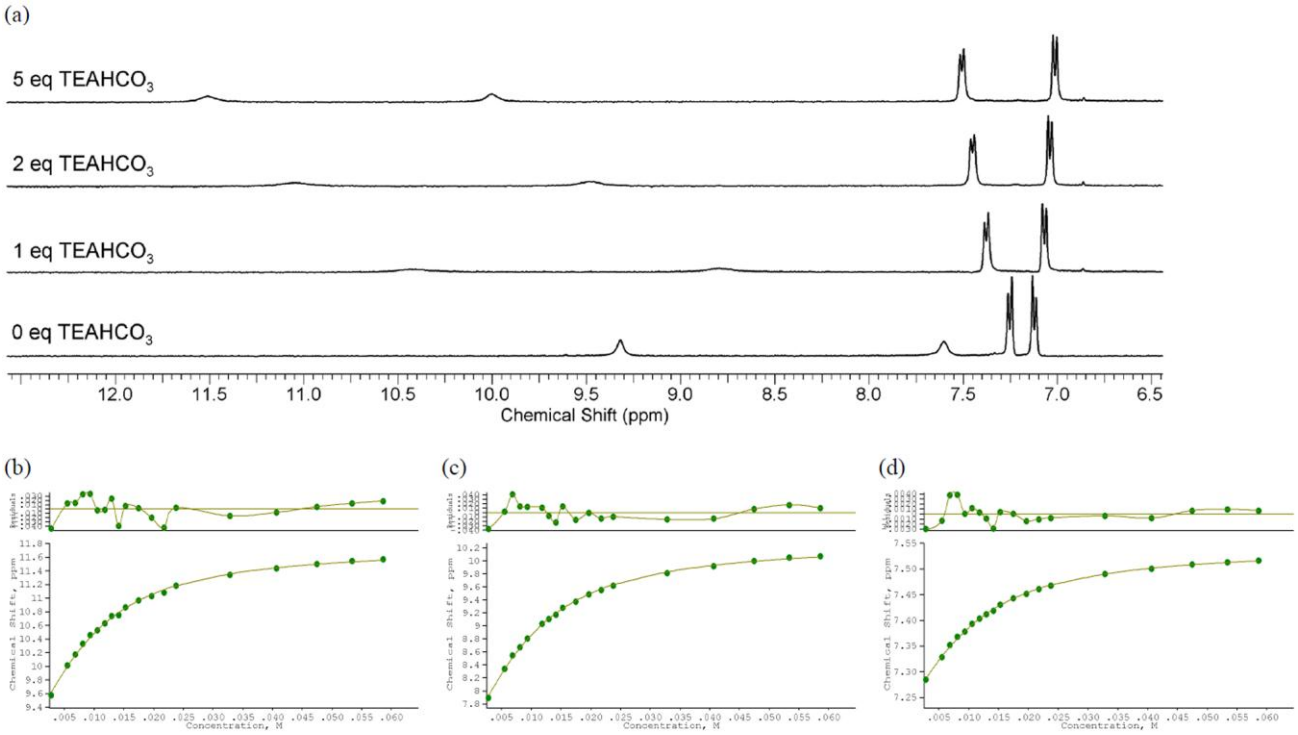


Figure S154. ^{1}H NMR titration of compound **20** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 210 M^{-1} (error 6 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 196 M^{-1} (error 4 %). (d) Fit plot for CH proton at δ = 7.25 ppm. K_a = 188 M^{-1} (error 3.5 %).

*Interactions of compound **21** (Bu) with various anions*

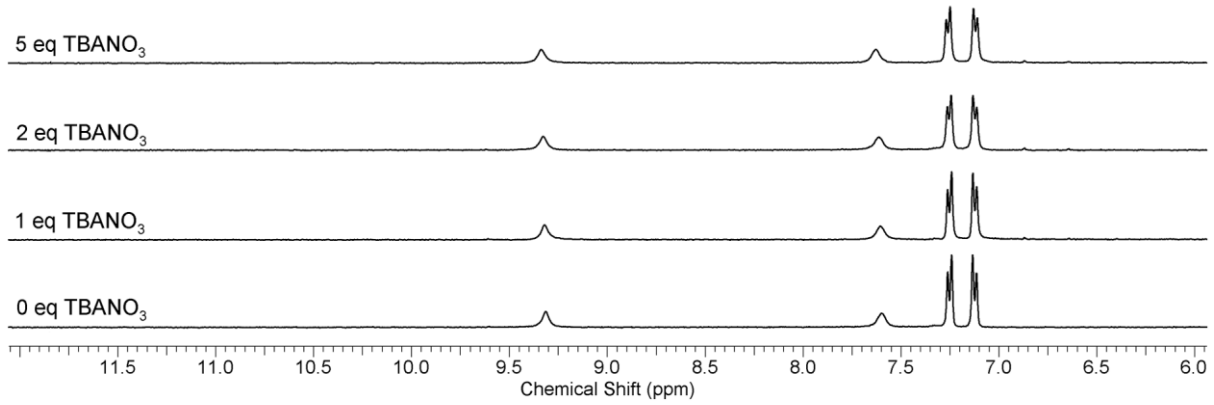


Figure S155. ^{1}H NMR titration of compound **21** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

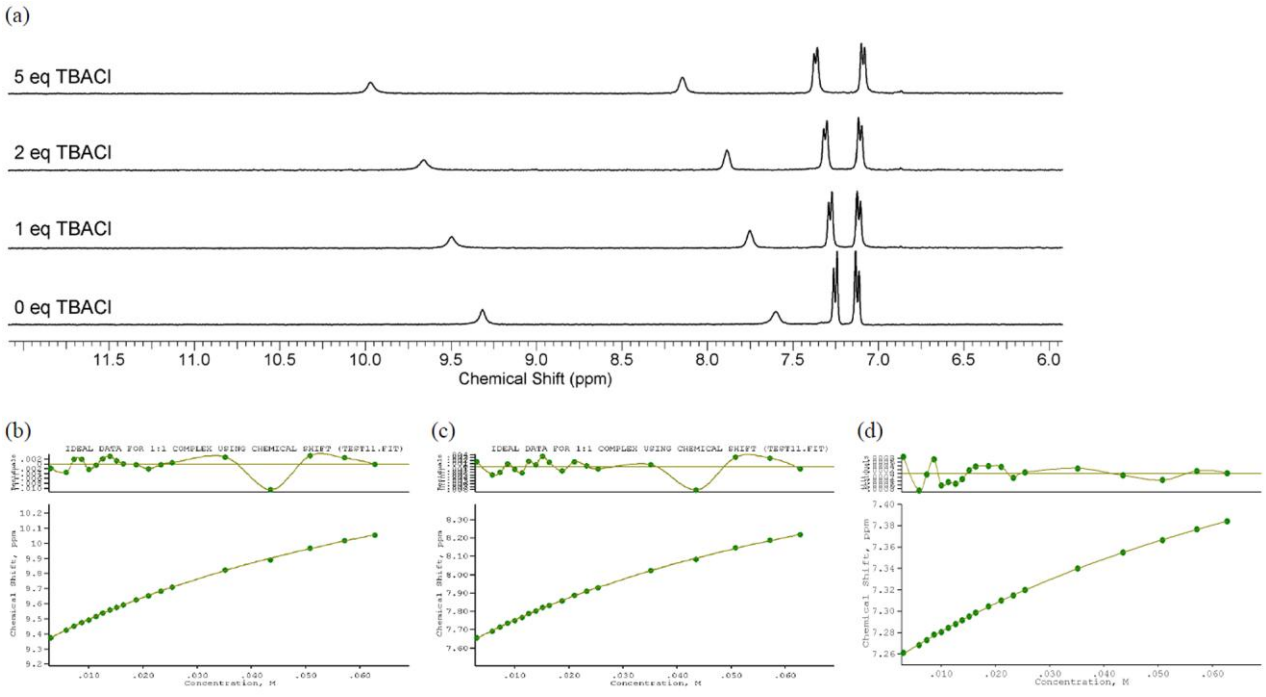


Figure S156. ^{1}H NMR titration of compound **21** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 12.8 M^{-1} (error 4.5 %). (c) Fit plot for NH proton at δ = 7.59 ppm. K_a = 12.6 M^{-1} (error 4.5 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 10.7 M^{-1} (error 3.5 %).

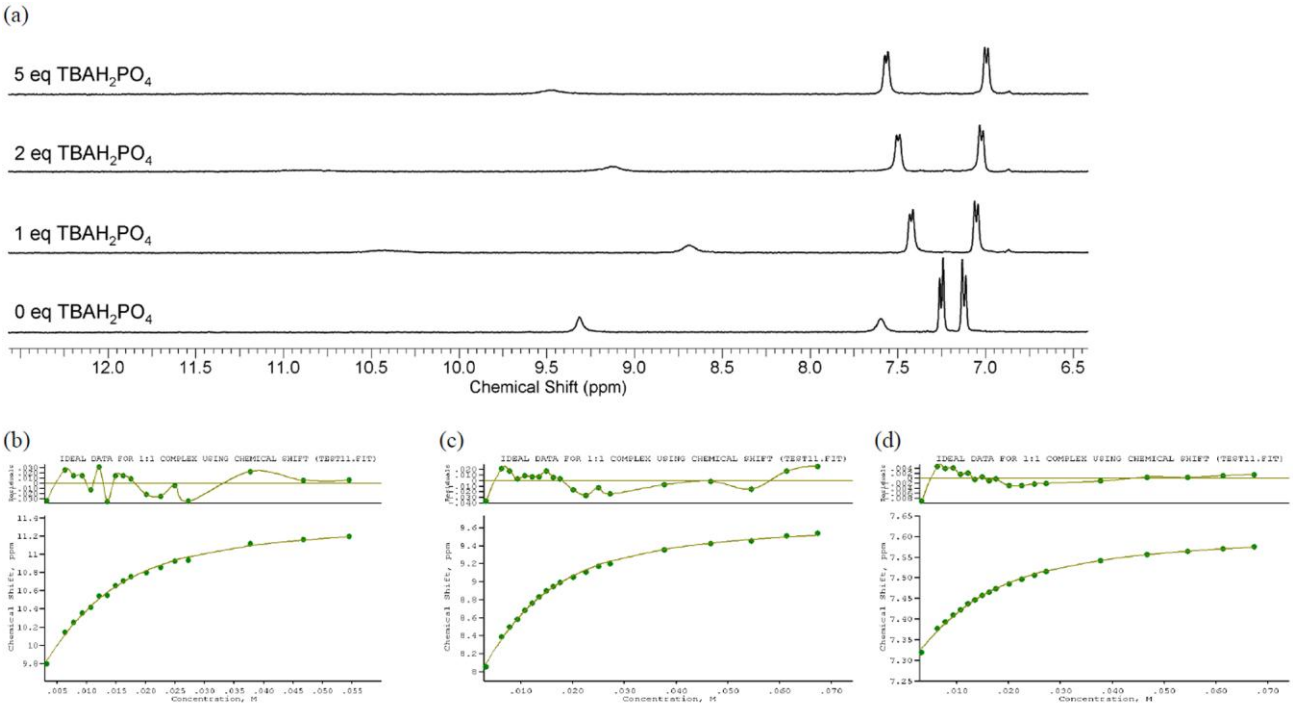


Figure S157. ^{1}H NMR titration of compound **21** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 187 M^{-1} (error 9 %). (c) Fit plot for NH proton at δ = 7.59 ppm. K_a = 163 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 125 M^{-1} (error 6 %).

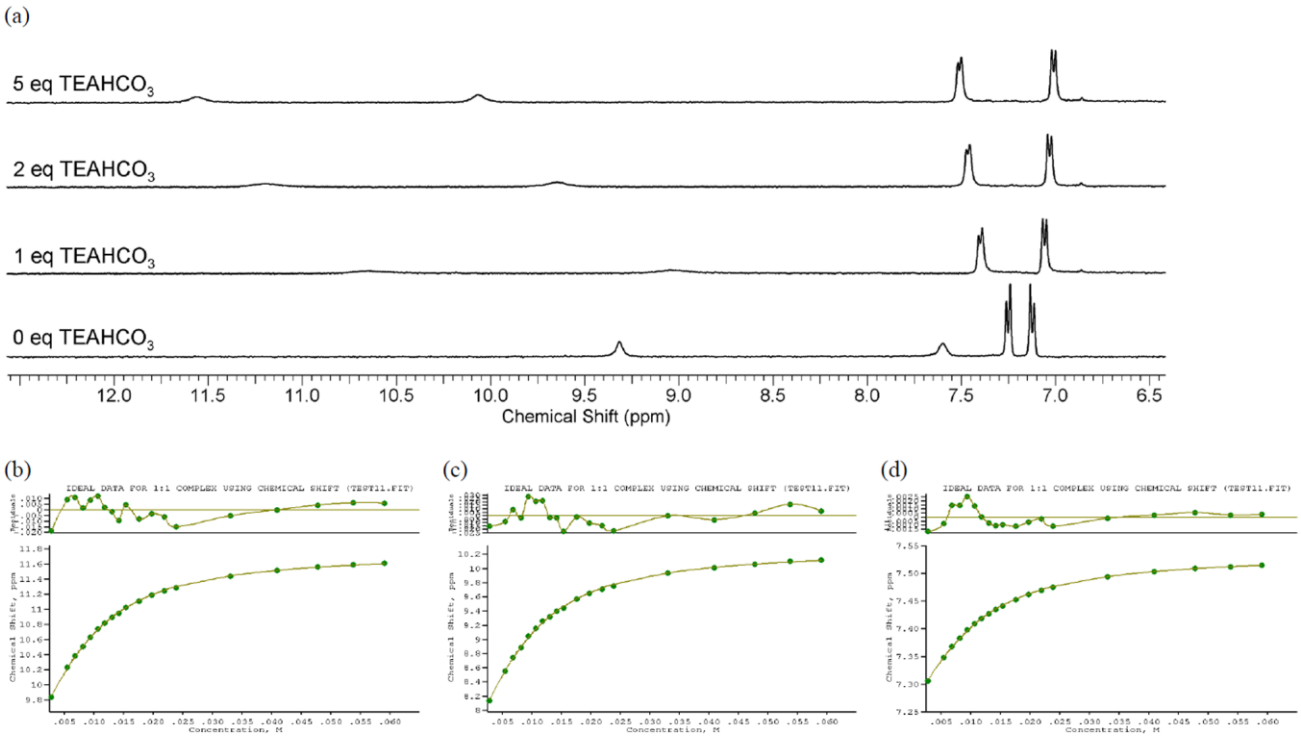


Figure S158. ^{1}H NMR titration of compound **21** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.30 ppm. K_a = 243 M^{-1} (error 2 %). (c) Fit plot for NH proton at δ = 7.59 ppm. K_a = 249 M^{-1} (error 3.5 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 217 M^{-1} (error 2 %).

*Interactions of compound **22** (Pe) with various anions*

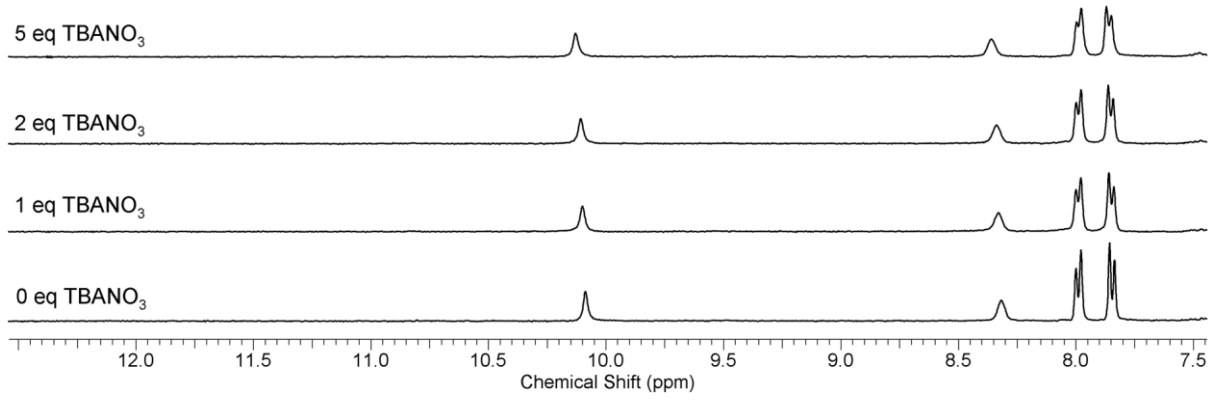


Figure S159. ^{1}H NMR titration of compound **22** with $TBANO_3$ in DMSO-d_6 with 0.5 % water at 298 K. No interaction observed.

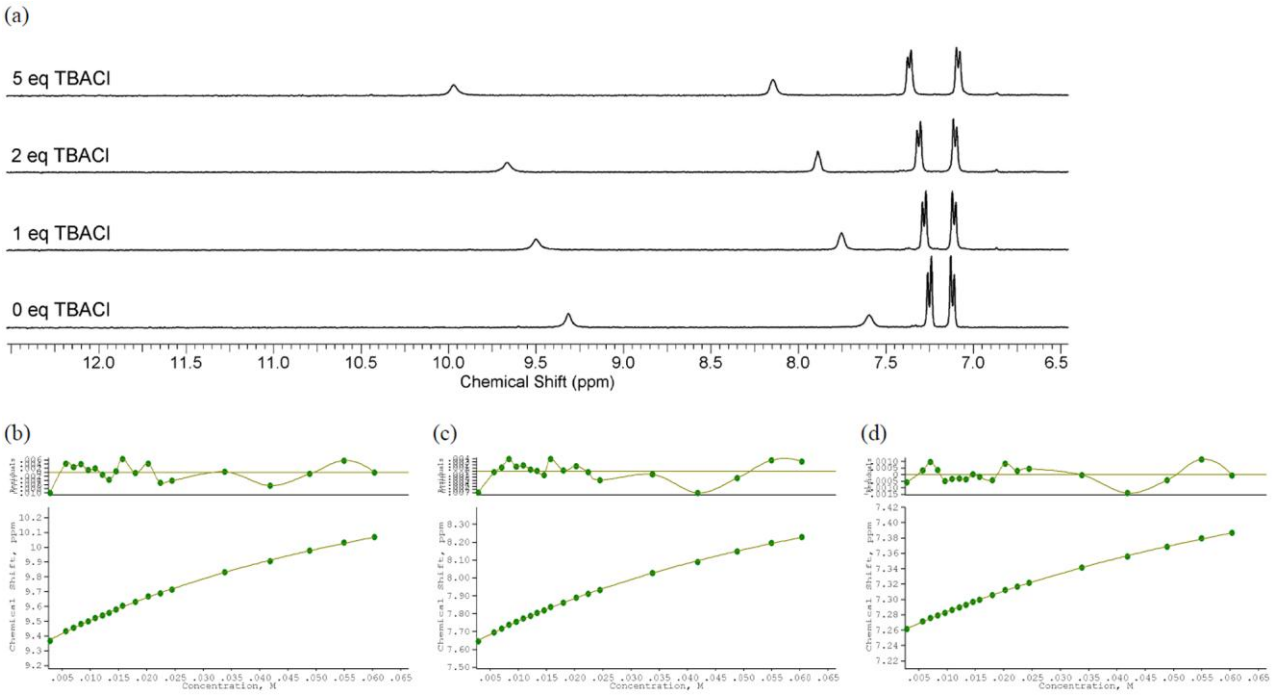


Figure S160. ^{1}H NMR titration of compound **22** with TBACl in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 13.3 M^{-1} (error 6 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 12.8 M^{-1} (error 5 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 10.3 M^{-1} (error 4.5 %).

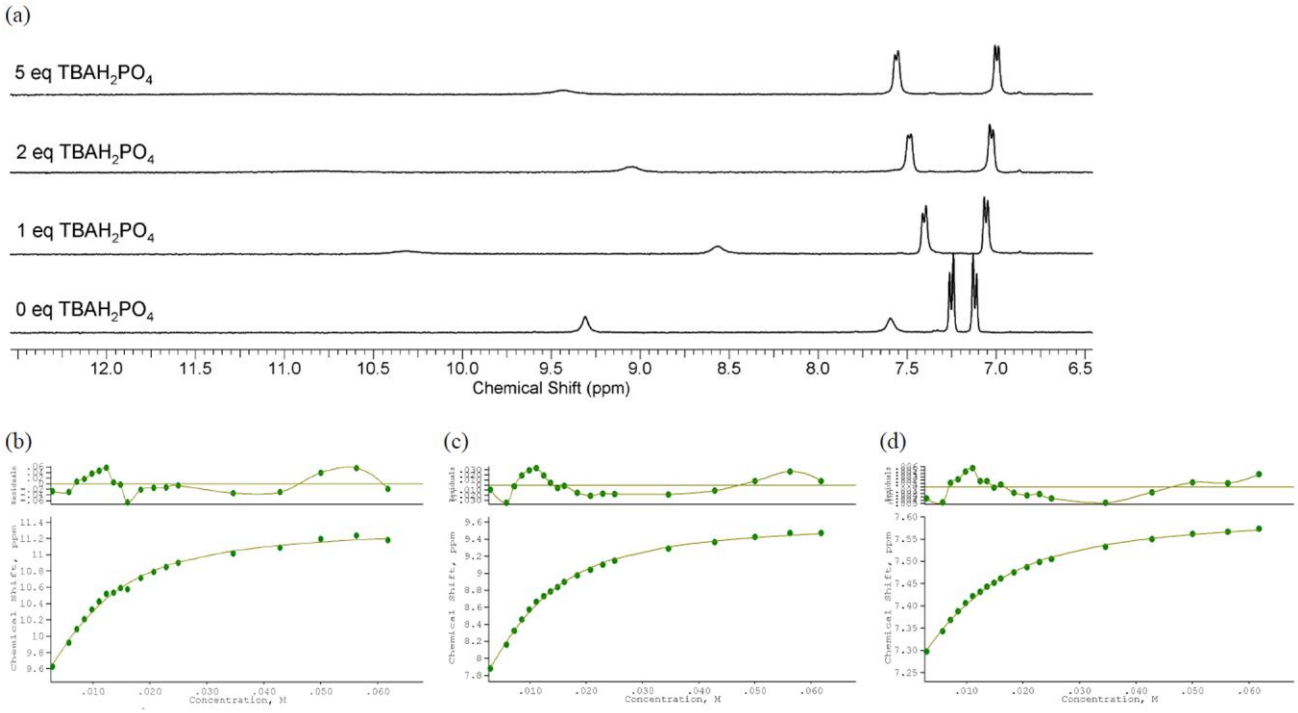


Figure S161. ^{1}H NMR titration of compound **22** with $TBAH_2PO_4$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 217 M^{-1} (error 11 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 218 M^{-1} (error 6 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 157 M^{-1} (error 5.5 %).

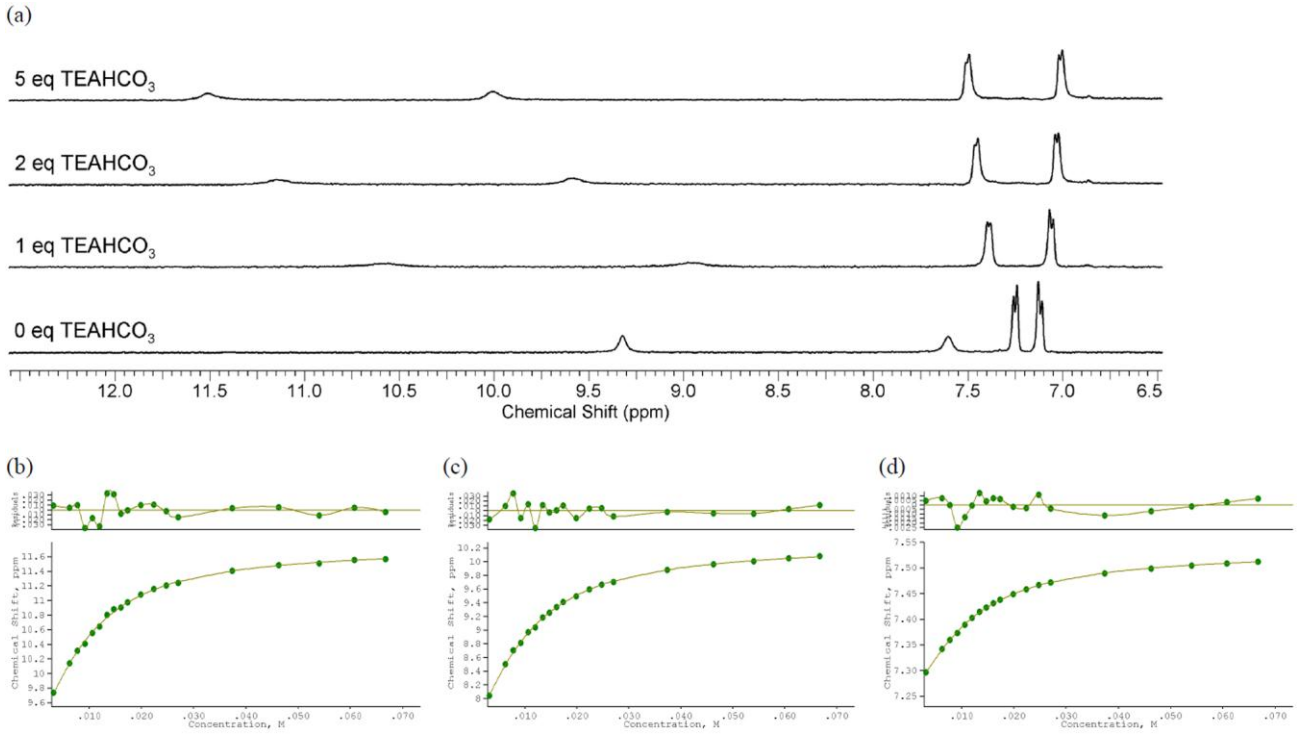


Figure S162. ^{1}H NMR titration of compound **22** with $TEAHCO_3$ in DMSO-d_6 with 0.5 % water at 298 K. Fit plots show chemical shift (ppm) versus anion concentration (M). (a) Stack plot. (b) Fit plot for NH proton at δ = 9.31 ppm. K_a = 208 M^{-1} (error 4 %). (c) Fit plot for NH proton at δ = 7.60 ppm. K_a = 197 M^{-1} (error 3.5 %). (d) Fit plot for CH proton at δ = 7.26 ppm. K_a = 182 M^{-1} (error 2 %).

S5.4 QSAR analysis of anion binding

Table S17 gives an overview of the Hammett constants (σ, for substituents in meta and in para position), pK_a values (for both NH functions of the thiourea moiety) and the $V_{S,max}$, $V_{S,min}$ and PI values obtained through DFT calculations as described in Section S5.2. Tables S18-S20 give an overview of the association constants with chloride, dihydrogen phosphate and bicarbonate respectively, as described in Section S5.3. Values are given for each of the protons that are affected during association with the anion and Figure S163 represents how these protons are labeled.

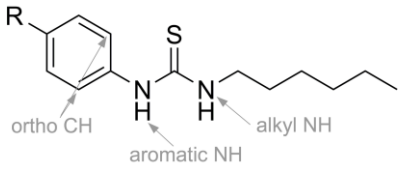


Figure S163. Labeling of the various protons involved in anion binding (hydrogen bonding).

Table S17. Overview of the Hammett constant (σ), pK_as and electrostatic surface descriptors.

Substituent	σ-para	σ-meta	pKa (arom NH)	pKa (alkyl NH)	$V_{S,max}$	$V_{S,min}$	PI
			training set				
2 (CF_3)	0.54	0.43	11.96	15.56	66.575	-23.2602	12.62661
3 (Cl)	0.23	0.37	12.44	15.7	63.2372	-25.9107	11.869437
4 (CN)	0.66	0.56	11.71	15.59	68.111	-42.1146	13.299468
5 ($COCF_3$)	0.80	0.63	11.53	15.61	70.3536	-36.9306	13.108429
7 (COOMe)	0.45	0.37	12.1	15.75	64.1963	-39.1565	12.392093
8 (F)	0.06	0.34	12.67	15.79	62.8197	-26.5471	12.033459
9 (H)	0.00	0.00	12.75	15.84	59.7756	-29.0791	11.570452
10 (I)	0.18	0.35	12.41	15.7	63.3237	-25.8099	11.63142
11 (NO_2)	0.78	0.71	11.35	15.34	70.1887	-39.8408	13.750149
12 (OCOMe)	0.31	0.39	12.46	15.8	65.4778	-43.5784	13.12108
13 (OCF_3)	0.35	0.38	12.16	15.69	65.3214	-24.3503	12.227026
15 (OMe)	-0.27	0.12	12.76	15.88	59.1301	-30.5202	12.16239
16 (SMe)	0.00	0.15	12.8	15.76	62.4402	-26.724	12.033459
17 (SO_2Me)	0.72	0.6	11.63	15.6	68.1084	-43.951	13.96438
18 (Me)	-0.17	-0.07	13.02	15.93	58.7091	-30.0245	10.924613
19 (Et)	-0.15	-0.07	13.02	15.93	58.8355	-29.8674	10.426436
21 (Bu)	-0.16	-0.08	13.04	15.93	58.438	-30.2971	9.606557
22 (Pentyl)	-0.15	-0.08	13.03	15.93	58.6686	-30.3497	9.250343
			test set				
1 (Br)	0.23	0.39	12.44	15.7	63.2297	-25.7824	11.788309
6 (COMe)	0.5	0.38	12.02	15.76	65.3441	-42.2773	12.930252
14 (OEt)	-0.24	0.1	13.1	15.89	58.7365	-30.8415	11.512424
20 (Pr)	-0.13	-0.06	13.05	15.93	58.5161	-30.1247	9.994914

Table S18. Overview of the association constants (K_a in M^{-1}) and their log function ($\log K_a$) for the interaction with TBA chloride in DMSO-d_6/0.5% water at 298 K.

Substituent	K_a aromatic NH	K_a Alkyl NH	K_a *o*-CH	$\log K_a$ aromatic NH	$\log K_a$ alkyl NH	$\log K_a$ *o*-CH
			training set			
2 (CF_3)	27.9	26.1	25.9	1.45	1.42	1.41
3 (Cl)	22.2	19.9	19.1	1.35	1.30	1.28
4 (CN)	34.6	31.4	30.5	1.54	1.50	1.48
5 ($COCF_3$)	43.2	41.3	51.8	1.64	1.62	1.71
7 (COOMe)	28.1	25.6	29.2	1.45	1.41	1.47
8 (F)	16.8	15.5	13.8	1.23	1.19	1.14
9 (H)	13.7	13.5	13.2	1.14	1.13	1.12
10 (I)	22.7	19.9	17.8	1.36	1.30	1.25
11 (NO_2)	45.1	42.7	42.3	1.65	1.63	1.63
12 (OCOMe)	19.4	17.3	16.6	1.29	1.24	1.22
13 (OCF_3)	28.1	24.8	24.8	1.45	1.39	1.39

15 (OMe)	10.6	10.7	11.4	1.03	1.03	1.06
16 (SMe)	15.0	14.8	12.2	1.18	1.17	1.09
17 (SO_2Me)	40.3	37.8	38.1	1.61	1.58	1.58
18 (Me)	11.1	10.9	9.8	1.05	1.04	0.99
19 (Et)	14.4	13.4	10.7	1.16	1.13	1.03
21 (Bu)	12.8	12.6	10.7	1.11	1.10	1.03
22 (Pentyl)	13.3	12.8	10.3	1.12	1.11	1.01
test set						
1 (Br)	22.0	18.6	16.5	1.34	1.27	1.22
6 (COMe)	28.6	26.8	18.8	1.46	1.43	1.27
14 (OEt)	10.8	12.3	9.5	1.03	1.09	0.98
20 (Pr)	13.9	13.1	10.4	1.14	1.12	1.02

Table S19. Overview of the association constants (K_a in M^{-1}) and their log function ($\log K_a$) for the interaction with TBA dihydrogen phosphate in DMSO-d_6/0.5% water at 298 K.

Substituent	**K_a aromatic NH**	**K_a Alkyl NH**	**K_a *o*-CH**	**$\log K_a$ aromatic NH**	**$\log K_a$ alkyl NH**	**$\log K_a$ *o*-CH**
training set						
2 (CF_3)	-	852	532	-	2.93	2.73
3 (Cl)	-	428	331	-	2.63	2.52
4 (CN)	-	1103	912	-	3.04	2.96
5 ($COCF_3$)	-	-	-	-		
7 (COOMe)	-	700	511	-	2.85	2.71
8 (F)	-	272	229	-	2.43	2.36
9 (H)	-	211	169	-	2.32	2.23
10 (I)	-	448	354	-	2.65	2.55
11 (NO_2)	-	827	912	-	2.92	2.96
12 (OCOMe)	-	277	217	-	2.44	2.34
13 (OCF_3)	-	426	327	-	2.63	2.51
15 (OMe)	162	131	102	2.21	2.12	2.01
16 (SMe)	-	285	233	-	2.45	2.37
17 (SO_2Me)	-	943	684	-	2.97	2.84
18 (Me)	-	164	134	-	2.21	2.13
19 (Et)	191	169	135	2.28	2.23	2.13
21 (Bu)	187	163	125	2.27	2.21	2.10
22 (Pentyl)	217	218	157	2.34	2.34	2.20
test set						
1 (Br)	-	465	347	-	2.67	2.54
6 (COMe)	-	417	444	-	2.62	2.65
14 (OEt)	144	141	109	2.16	2.15	2.04
20 (Pr)	240	208	158	2.38	2.32	2.20

Table S20. Overview of the association constants (K_a in M^{-1}) and their log function (logK_a) for the interaction with TEA bicarbonate in DMSO-d_6/0.5% water at 298 K.

Substituent	K_a aromatic NH	K_a Alkyl NH	K_a *o*-CH	logK_a aromatic NH	logK_a alkyl NH	logK_a *o*-CH
			training set			
2 (CF_3)	-	-	931	-	-	2.97
3 (Cl)	-	452	428	-	2.66	2.63
4 (CN)	-	-	1550	-	-	3.19
5 ($COCF_3$)	-		-	-	-	-
7 (COOMe)	-	-	589	-	-	2.77
8 (F)	-	292	303	-	2.47	2.48
9 (H)	230	262	222	2.36	2.42	2.35
10 (I)	287	468	564	2.46	2.67	2.75
11 (NO_2)	-	-	(345)[a]	-	-	(2.54)[a]
12 (OCOMe)	262	271	266	2.42	2.43	2.42
13 (OCF_3)	-	592	439	-	2.77	2.64
15 (OMe)	171	157	145	2.23	2.20	2.16
16 (SMe)	324	364	316	2.51	2.56	2.50
17 (SO_2Me)	-	-	702	-	-	2.85
18 (Me)	147	152	147	2.17	2.18	2.17
19 (Et)	188	166	150	2.27	2.22	2.18
21 (Bu)	243	249	217	2.39	2.40	2.34
22 (Pentyl)	208	197	182	2.32	2.29	2.26
			test set			
1 (Br)	-	-	447	-	-	2.65
6 (COMe)	-	-	553	-	-	2.74
14 (OEt)	172	163	141	2.24	2.21	2.15
20 (Pr)	210	196	188	2.32	2.29	2.27

[a] Peak broadening and overlap of CH peaks make determination uncertain.

QSAR for interaction with TBA chloride

The anion binding properties of thioureas are most likely governed by electronic contributions of the substituents, which can be described by the Hammett constant, pK_a or $V_{S,max}$. Figures S164-S170 represent the correlation between the chloride binding abilities of the test set (logK_a) and the various electronic descriptors given in Table S17 calculated using Origin 8.1. It can be seen that the best correlation is obtained for the Hammett constant for substituent in the *para* position, the pK_a of the aromatic NH, or $V_{S,max}$. The correlation with the Hammett constant for substituent in the *meta* position is less good, indicating that the effect of the substituents on anion binding is due to their influence on the acidity of the aromatic NH and not due to the increased acidity of the *ortho*-CH. This is further confirmed by the fact that there is a good correlation with the pK_a of the aromatic NH and not the alkyl NH (which is not strongly influenced by the electronics of the substituent). The correlation with $V_{S,min}$ and PI are also less good because these descriptors tend to correlate with hydrogen bond acceptor ability, while anion binding is regulated by anion bond donor

ability ($V_{S,max}$). It can also be seen that the association constants obtained from the change in chemical shift of the aromatic NH, alkyl NH and *ortho*-CH give similar results, showing that all of them are a good measure for the association constant, which can also be seen from the fact that the values for the aromatic NH, alkyl NH and *ortho*-CH are all close to each other.

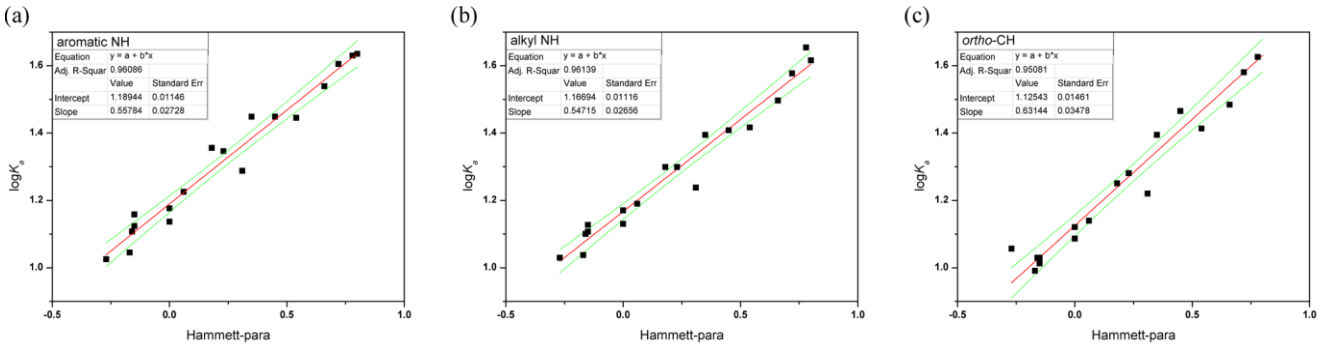


Figure S164. Correlation between logK_a for the association with TBA chloride and the Hammett constant for substituents in the *para* position. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

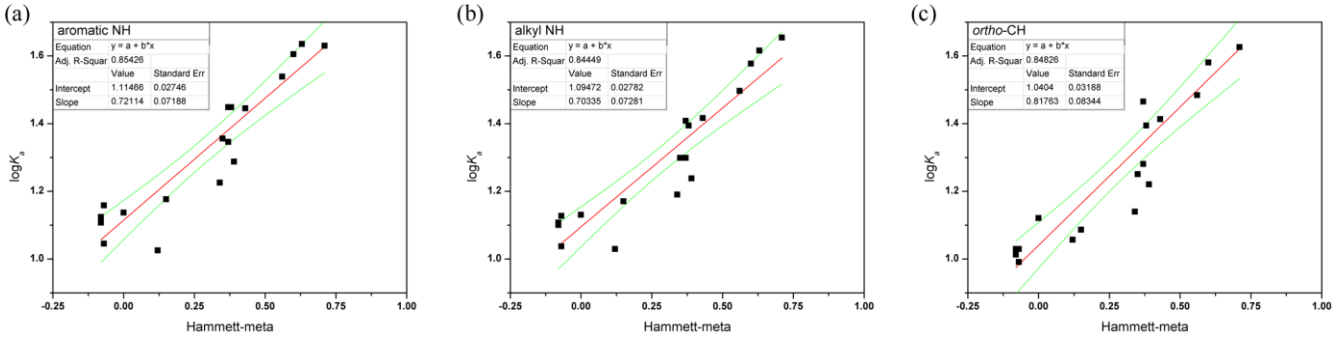


Figure S165. Correlation between logK_a for the association with TBA chloride and the Hammett constant for substituents in the *meta* position. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

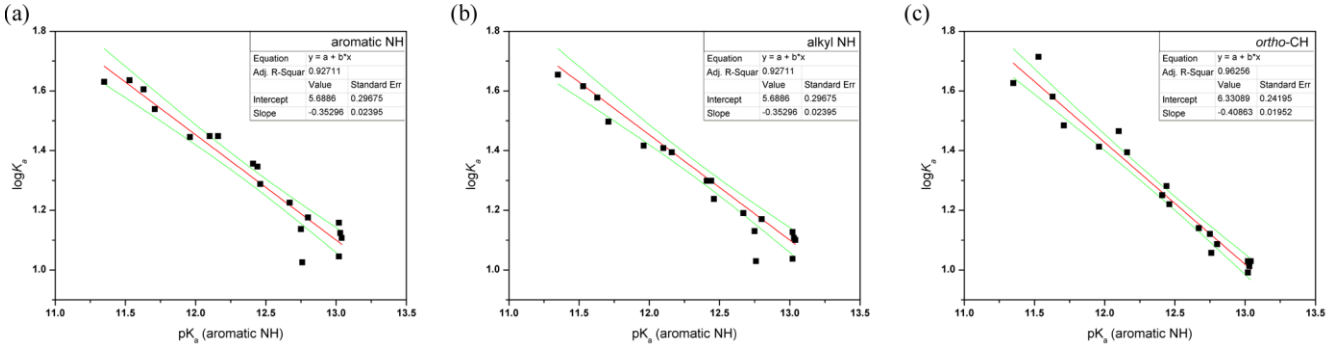


Figure S166. Correlation between logK_a for the association with TBA chloride and the pK_a of the aromatic NH. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

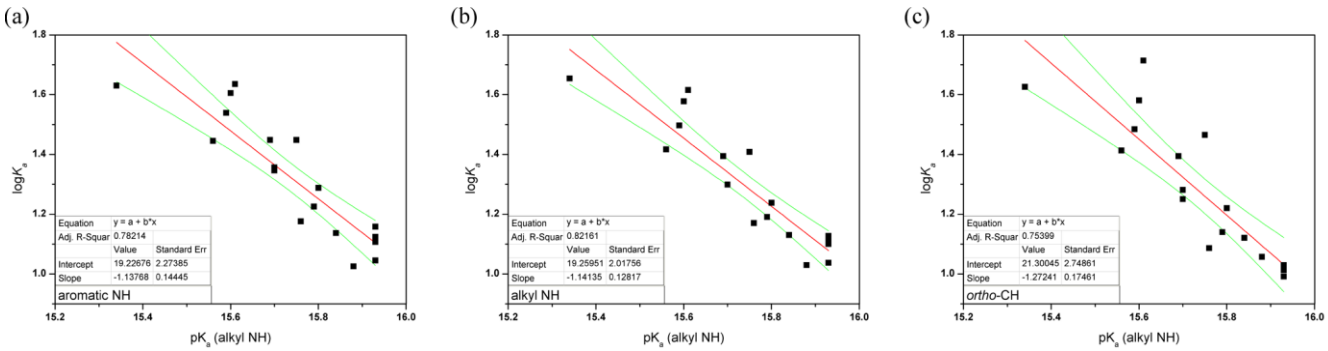


Figure S167. Correlation between logK_a for the association with TBA chloride and the pK_a of the alkyl NH. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

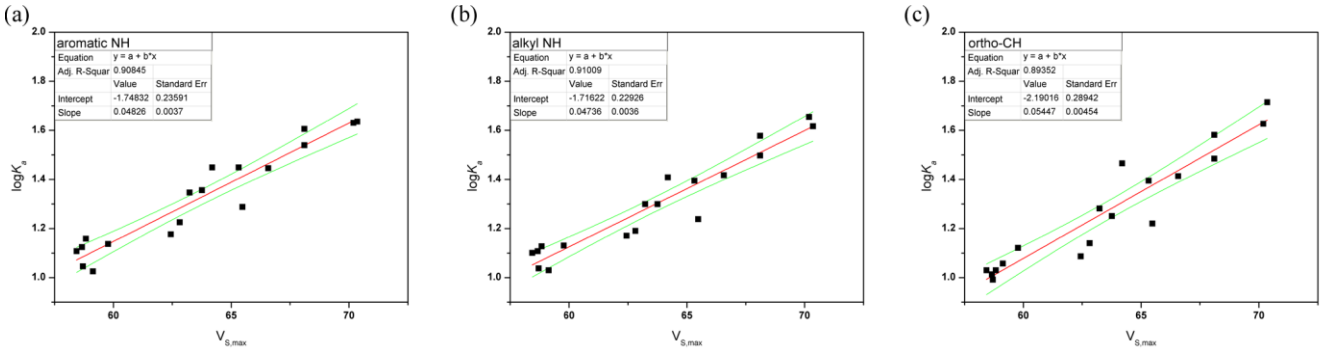


Figure S168. Correlation between logK_a for the association with TBA chloride and $V_{S,max}$. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

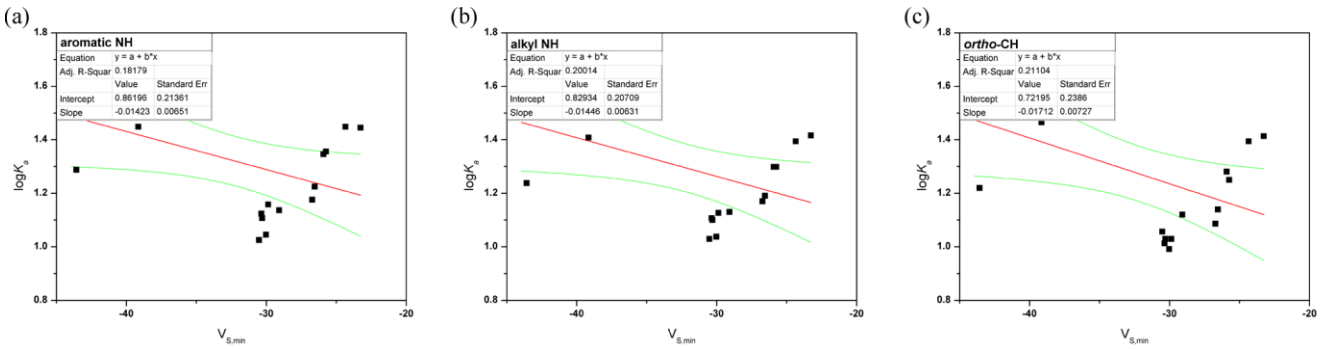


Figure S169. Correlation between logK_a for the association with TBA chloride and $V_{S,min}$. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

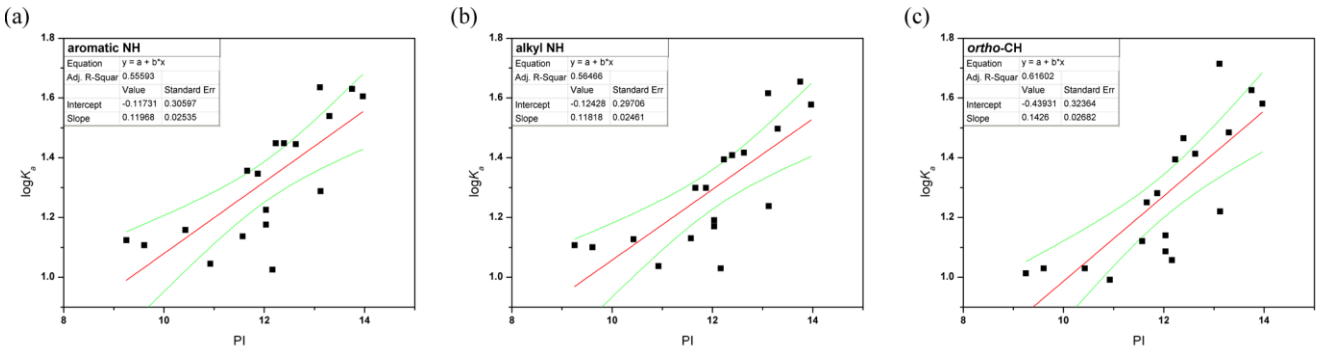


Figure S170. Correlation between logK_a for the association with TBA chloride and PI. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

QSAR for interaction with TBA dihydrogen phosphate

Figures S171-S177 represent the correlation between the phosphate binding abilities of the test set (logK_a) and the various electronic descriptors given in Table S17 calculated using Origin 8.1. The same correlations are seen as for the association with chloride, with the best correlations observed with the Hammett constant for *para* substituents, pK_a for the aromatic NH and $V_{S,max}$. Similar to chloride, it can be seen that the association constants obtained from the change in chemical shift of the aromatic NH, alkyl NH and *ortho*-CH give similar results. However due to peak broadening of the NH peaks during NMR titrations, the association constants could not always be calculated from the aromatic NH signals. The data sets obtained for the alkyl NH and the *ortho*-CH are therefore the most complete. As the correlation found for both the alkyl NH and the *ortho*-CH are similar, both association constants can be used for building a QSAR model. The correlation with the *ortho*-CH are a bit lower than for the alkyl NH, which is probably due to the higher uncertainty of the association constants due to a smaller change in chemical shift for the *ortho*-CH signal. LogK_a could not be obtained for **5** (–$COCF_3$) due to decomposition of the compound during the titration.

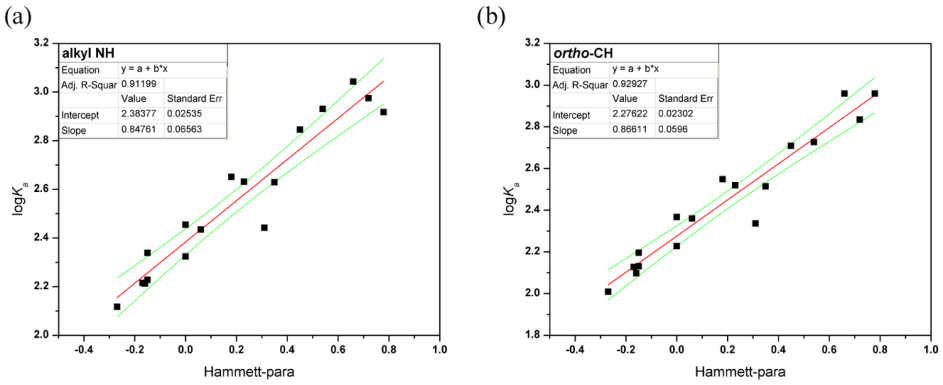


Figure S171. Correlation between logK_a for the association with TBA dihydrogen phosphate and the Hammett constant for substituents in the *para* position. LogK_a was calculated using the change in chemical shift of (a) alkyl NH, (b) *ortho* CH.

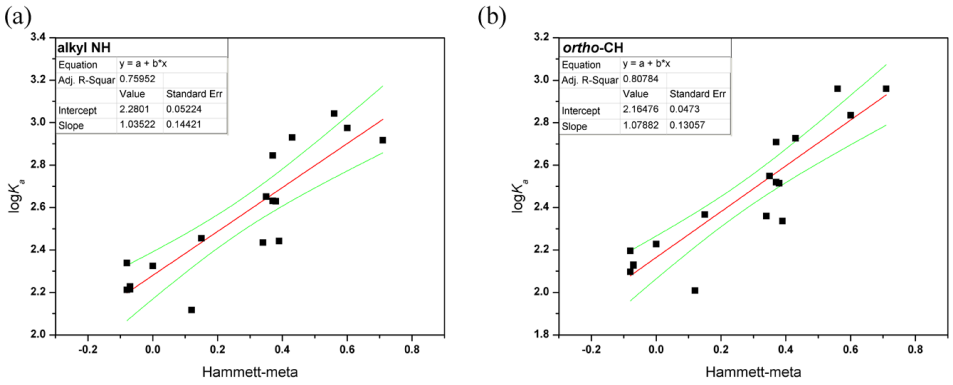


Figure S172. Correlation between logK_a for the association with TBA dihydrogen phosphate and the Hammett constant for substituents in the *meta* position. LogK_a was calculated using the change in chemical shift of (a) alkyl NH, (b) *ortho* CH.

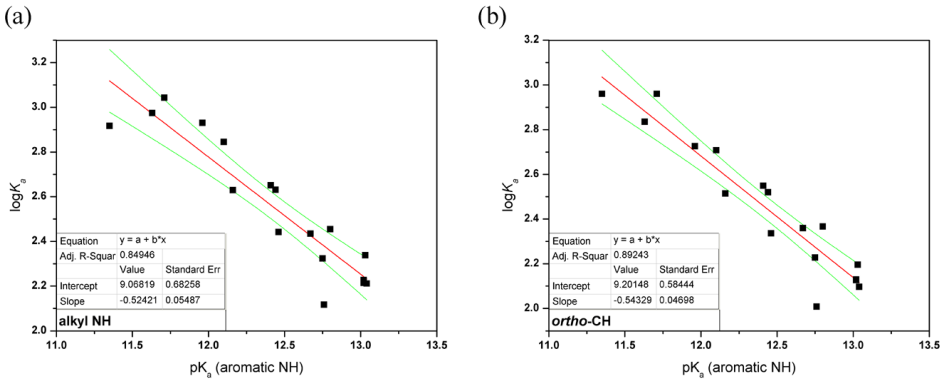


Figure S173. Correlation between logK_a for the association with TBA dihydrogen phosphate and the pK_a of the aromatic NH. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

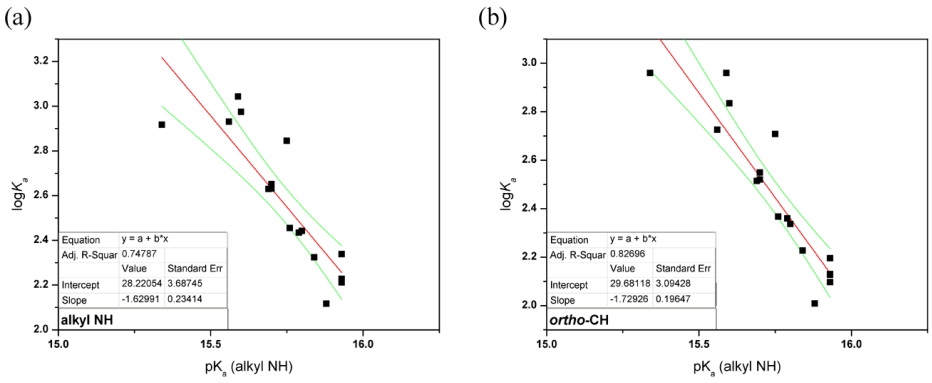


Figure S174. Correlation between logK_a for the association with TBA dihydrogen phosphate and the pK_a of the alkyl NH. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

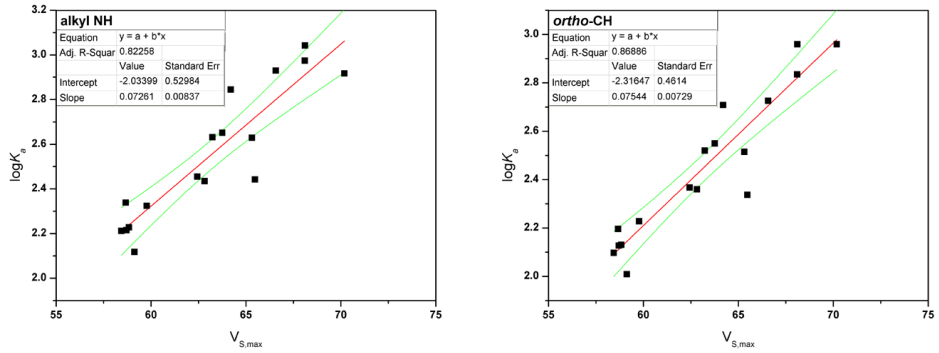


Figure S175. Correlation between logK_a for the association with TBA dihydrogen phosphate and $V_{S,max}$. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

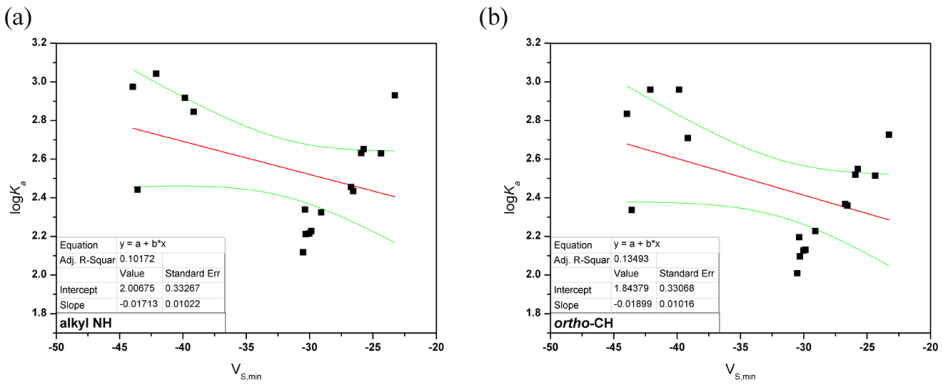


Figure S176. Correlation between logK_a for the association with TBA dihydrogen phosphate and $V_{S,min}$. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

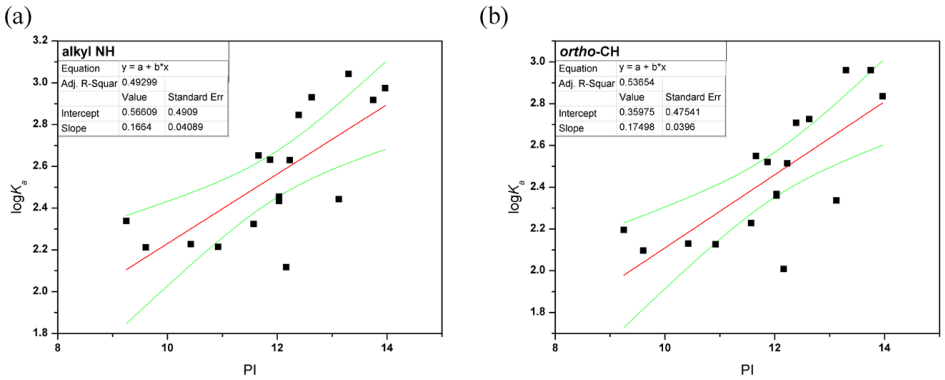


Figure S177. Correlation between logK_a for the association with TBA dihydrogen phosphate and PI. LogK_a was calculated using the change in chemical shift of (a) aromatic NH, (b) alkyl NH, (c) *ortho* CH.

QSAR for interaction with TEA bicarbonate

Figures S178-S184 represent the correlation between the bicarbonate binding abilities of the test set (logK_a) and the various electronic descriptors given in Table S17 calculated using Origin 8.1. The same correlations are seen as for the association with chloride and phosphate, with the best correlations observed with the Hammett constant for *para* substituents, pK_a for the aromatic NH and $V_{S,max}$. However, the correlations are less good than for chloride and phosphate, which might be due to competing deprotonation of the receptor. Similar to chloride, it can be seen that the association constants obtained from the change in chemical shift of the aromatic NH, alkyl NH and *ortho*-CH give similar results. However due to peak broadening of the NH peaks during NMR titrations, the association constants could not always be calculated from the aromatic and/or alkyl NH signals. The data obtained for the *ortho*-CH is therefore the most complete data set and it is clear from the data that the *ortho*-CH data is a good measure for the association constant due to the high correlation with the NH data (where it could be obtained, see interaction with chloride). The data set for the aromatic NH could only be obtained for the lowest values of the Hammett constant and the range is therefore not sufficient for a meaningful QSAR analysis. Furthermore, the data set for the alkyl NH suffered from the same problem to a lesser extend and was hindered by large inaccuracies of the logK_a values obtained from these NHs due to significant peak broadening. LogK_a could not be obtained for **5** (–$COCF_3$) and **11** (-NO_2) due to decomposition of the compound during the titration (**5**) or broadening/overlap of the *ortho*-CH peak (**11**).

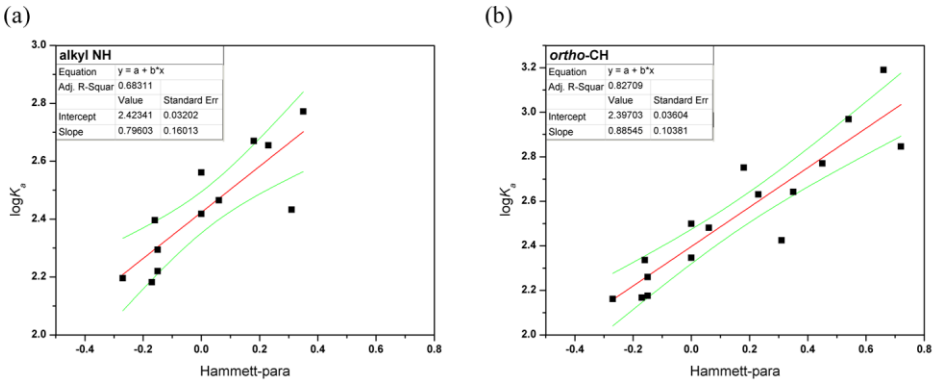


Figure S178. Correlation between logK_a for the association with TEA bicarbonate and the Hammett constant for substituents in the *para* position. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

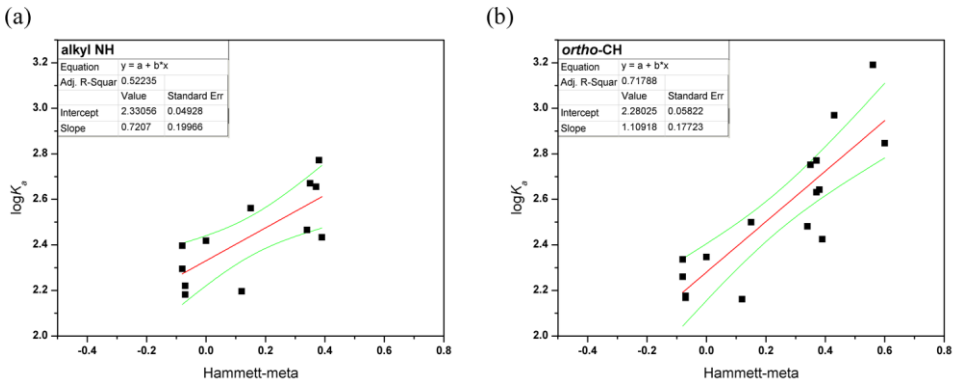


Figure S179. Correlation between logK_a for the association with TEA bicarbonate and the Hammett constant for substituents in the *meta* position. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

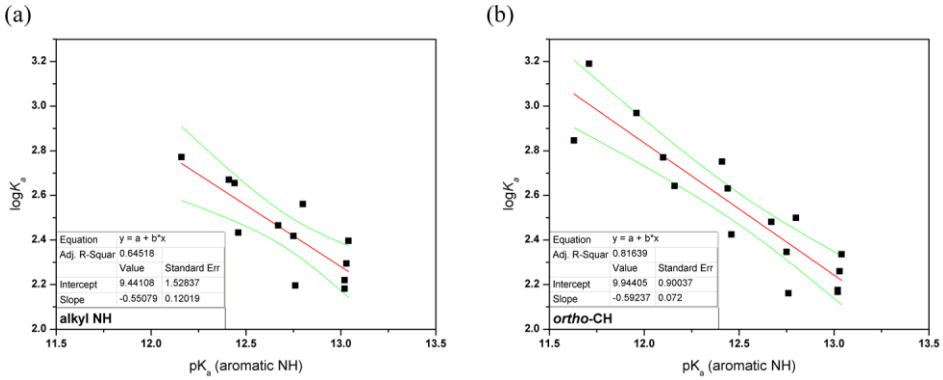


Figure S180. Correlation between logK_a for the association with TEA bicarbonate and the pK_a of the aromatic NH. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

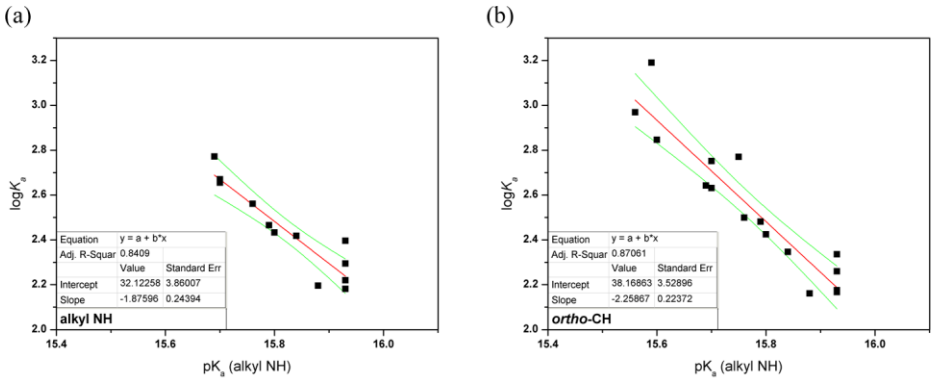


Figure S181. Correlation between logK_a for the association with TEA bicarbonate and the pK_a of the alkyl NH. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

(a) (b)

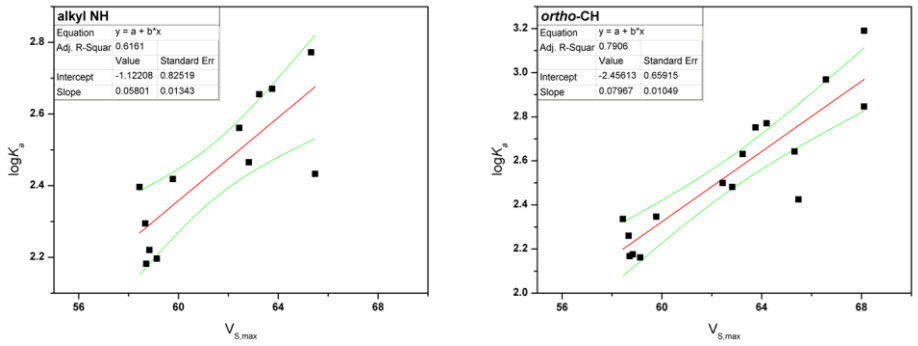


Figure S182. Correlation between logK_a for the association with TEA bicarbonate and $V_{S,max}$. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

(a) (b)

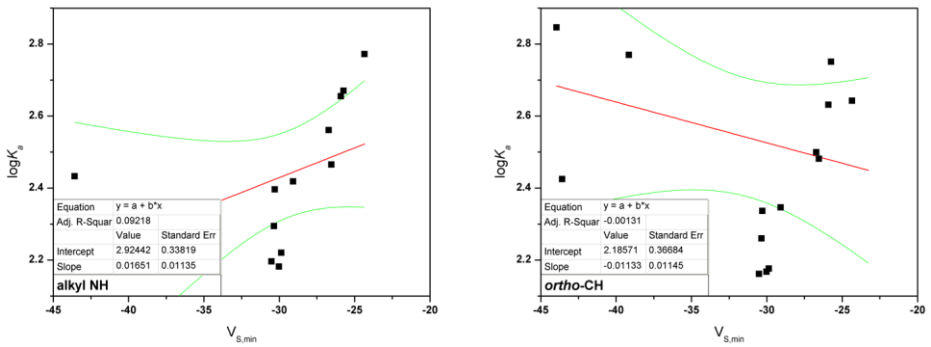


Figure S183. Correlation between logK_a for the association with TEA bicarbonate and $V_{S,min}$. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

(a) (b)

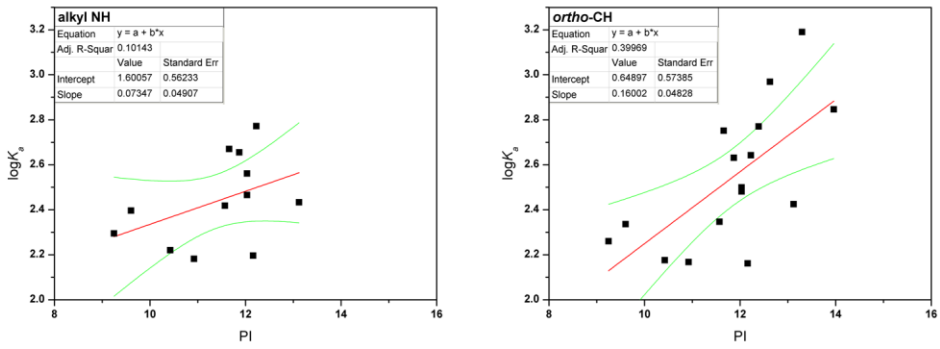


Figure S184. Correlation between logK_a for the association with TEA bicarbonate and PI. LogK_a was calculated using the change in chemical shift of (a) alkyl NH (incomplete data set), (b) *ortho* CH.

Correlation between chloride, phosphate and bicarbonate binding

Figures S185-S187 represent the correlation between the chloride binding abilities of the test set (logK_a) and the phosphate or bicarbonate binding abilities of the test set (logK_a) calculated using Origin 8.1. It can be seen that the associations with the different anions are highly correlated to one another. This means that the association with chloride, phosphate and bicarbonate are governed by similar factors (in this case mainly electronic factors). The largest deviation from a perfect correlation are seen for bicarbonate and for the most electron withdrawing substituents (highest values of the Hammett constant), which can indicate competing deprotonation of the receptor.

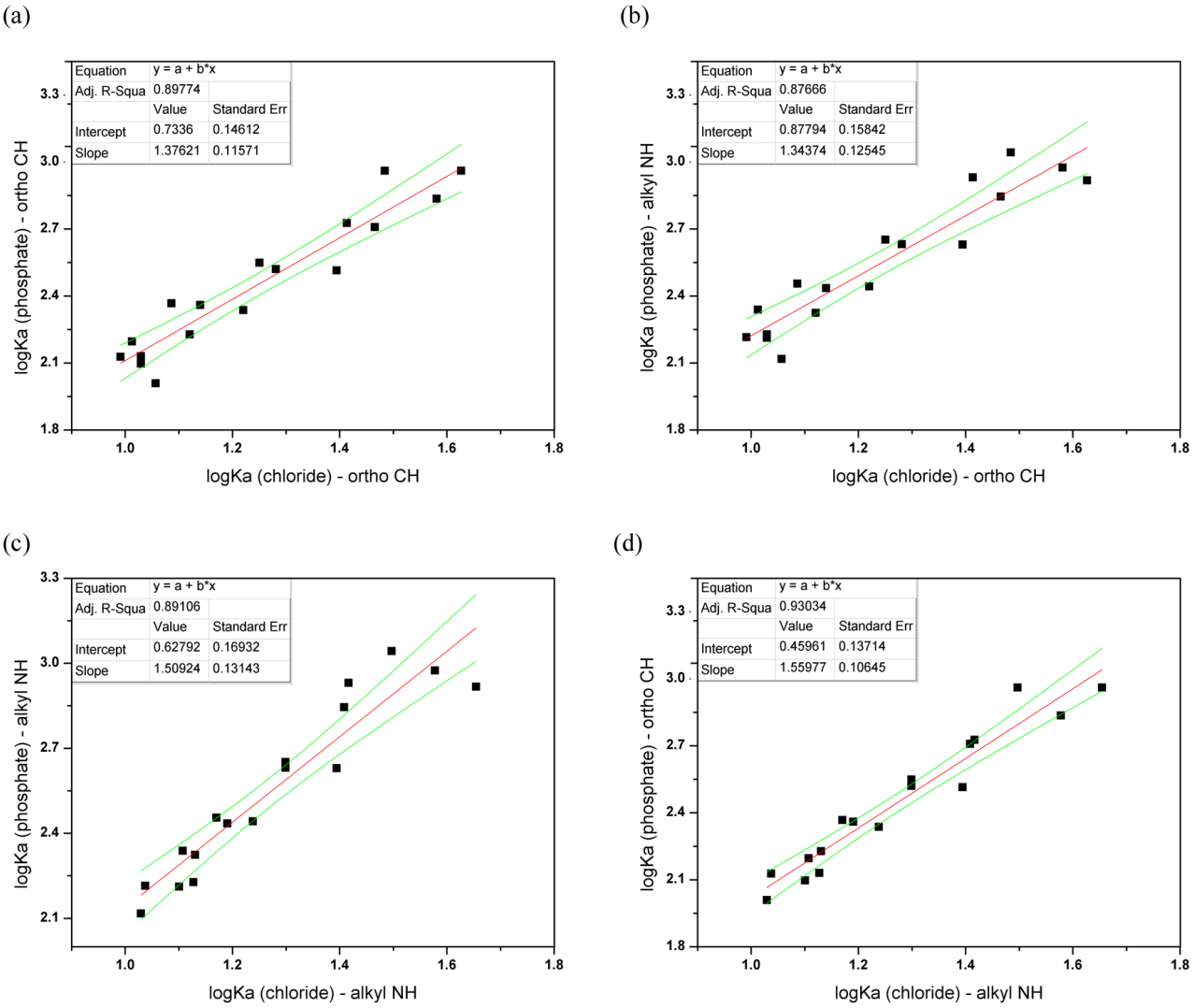


Figure S185. Correlation between logK_a for the association with TBA chloride and logK_a for the association with TBA dihydrogen phosphate. LogK_a values were calculated from the change in chemical shift of either the alkyl NH or the *ortho* CH.

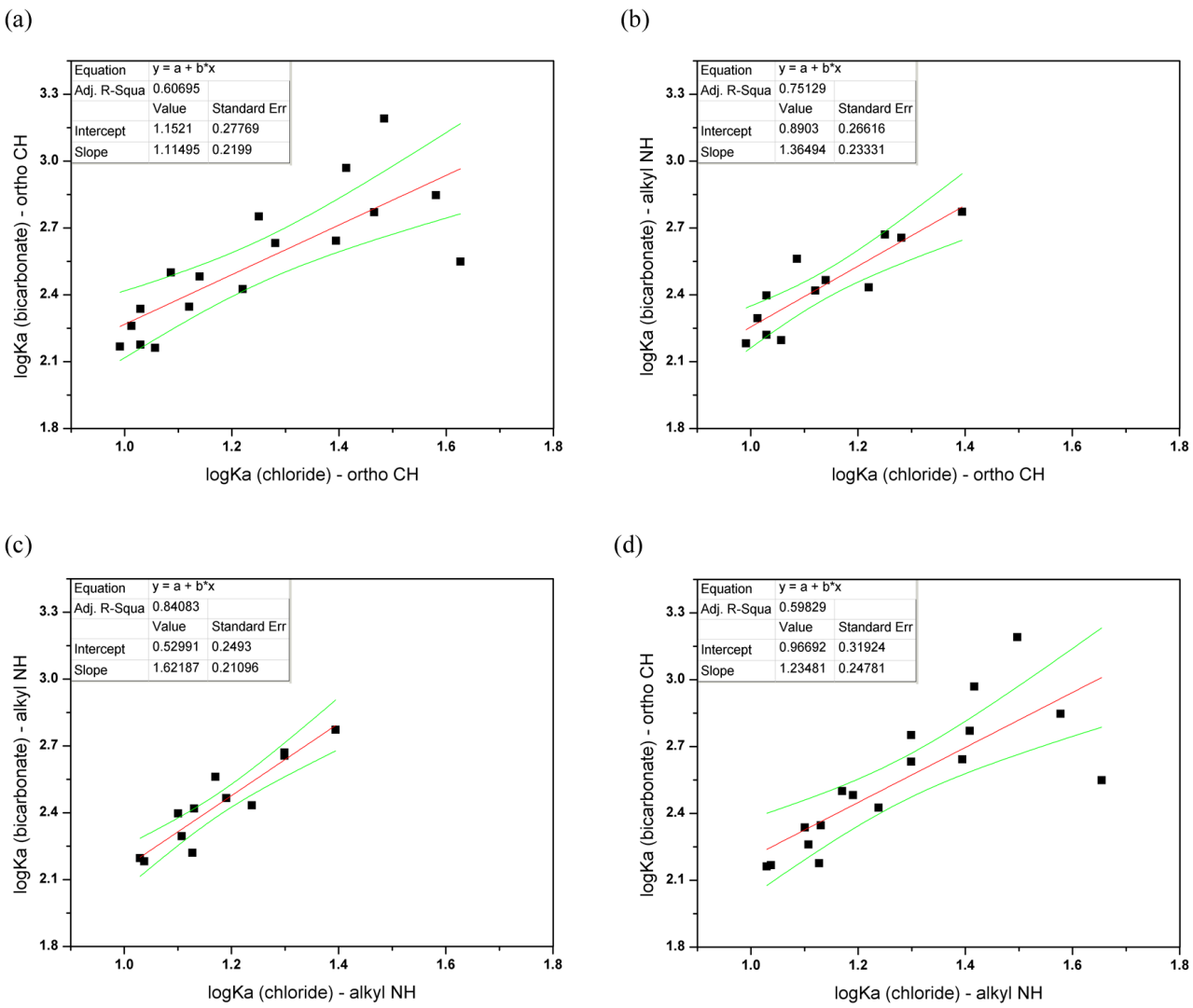
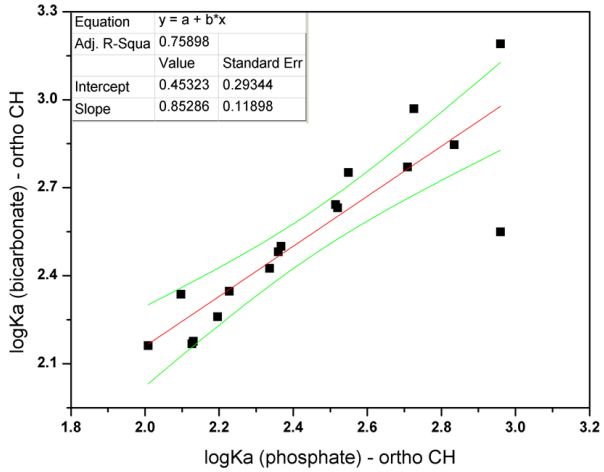
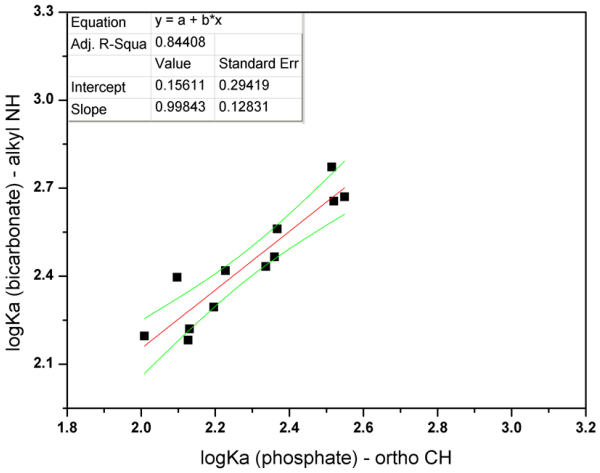


Figure S186. Correlation between logK_a for the association with TBA chloride and logK_a for the association with TEA bicarbonate. LogK_a values were calculated from the change in chemical shift of either the alkyl NH or the *ortho* CH.

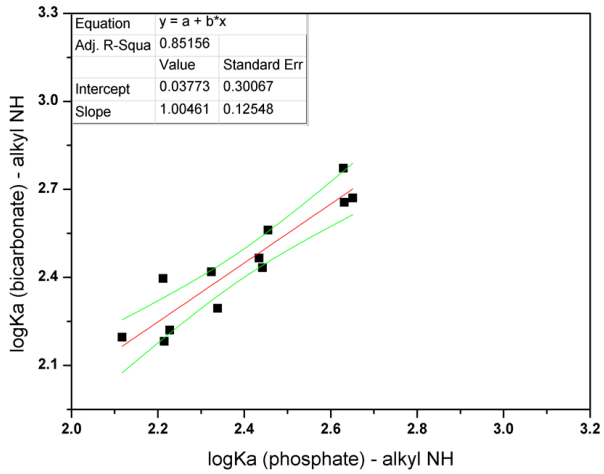
(a)



(b)



(c)



(d)

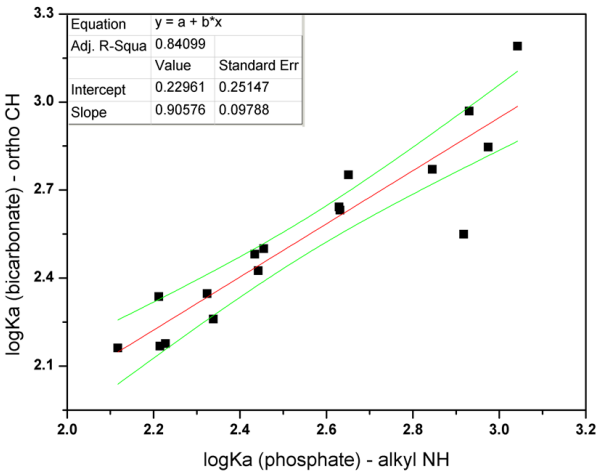


Figure S187. Correlation between logK_a for the association with TBA dihydrogen phosphate and logK_a for the association with TEA bicarbonate. LogK_a values were calculated from the change in chemical shift of either the alkyl NH or the *ortho* CH.

S6. TRANSPORT STUDIES

S6.1 Experimental procedures

These procedures describe typical membrane transport tests as referred to in the article. Internal and external solutions can vary (see caption of figures). Chloride concentrations during transport experiments were determined using an *Accumet* chloride-selective electrode (our lab posses 4 chloride selective electrodes and all of them were used during the preparation of this manuscript). POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) was supplied by *Genzyme* and was stored at –20°C as a solution in chloroform (1 g POPC in 35 mL chloroform). Polyoxyethylene(8)lauryl ether was used as detergent and was supplied by *TCI*.

Preparation of Vesicles

A lipid film of POPC was formed from a chloroform solution under reduced pressure and dried under vacuum for at least 4 hours. The lipid film was rehydrated by vortexing with a metal chloride (MCl) salt solution (489 mM MCl, 5 mM phosphate buffer at pH 7.2). The lipid suspension was then subjected to nine freeze-thaw cycles, where the suspension was alternatingly allowed to freeze in a liquid nitrogen bath, followed by thawing in a water bath. The lipid suspension was allowed to age for 30 min at room temperature and was subsequently extruded 25 times through a 200 nm polycarbonate membrane. The resulting unilamellar vesicles were dialyzed against the external medium to remove unencapsulated MCl salts.

Chloride/Nitrate Transport Assay

Unilamellar POPC vesicles containing NaCl, prepared as described above, were suspended in the external medium consisting of a 489 mM $NaNO_3$ solution buffered to pH 7.2 with sodium phosphate salts (5 mM buffer). The lipid concentration per sample was 1 mM. A DMSO solution of the carrier molecule was added to start the experiment and the chloride efflux was monitored using a chloride sensitive electrode. At 5 min, the vesicles were lysed with 50 µl of polyoxyethylene(8)lauryl ether (0.232 mM in 7:1 water:DMSO v/v) and a total chloride reading was taken at 7 min. The initial value was set at 0 % chloride efflux and the final chloride reading (at 7 minutes) was set as 100 % chloride efflux. All other data points were calibrated to these points.

Chloride/Bicarbonate Transport Assay

Unilamellar POPC vesicles containing 489 mM NaCl solution buffered to pH 7.2 with 20 mM sodium phosphate salts, prepared as described above, were suspended in the external medium consisting of a 162 mM Na_2SO_4 solution buffered to pH 7.2 with sodium phosphate salts (20 mM buffer). The lipid

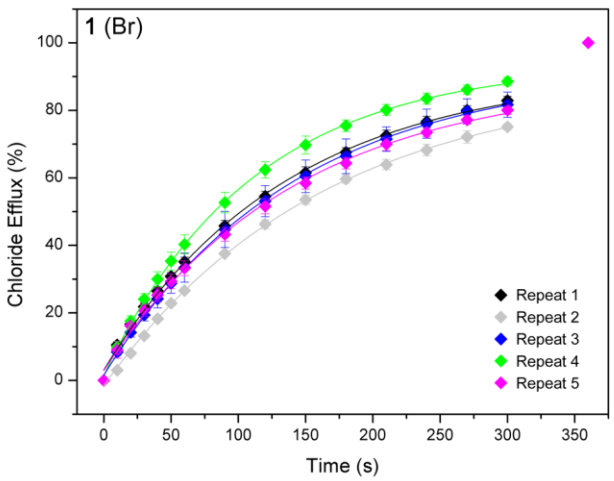
concentration per sample was 1 mM. A DMSO solution of the carrier molecule (10 mM) was added to start the experiment and chloride efflux was monitored using a chloride sensitive electrode. At 2 min, a $NaHCO_3$ solution (1 M in 162 mM Na_2SO_4 buffered to pH 7.2 with 20 mM sodium phosphate salts) was added so that the outer solution contained 40 mM $NaHCO_3$. At 7 min, the vesicles were lysed with 50 µl of polyoxyethylene(8)lauryl ether (0.232 mM in 7:1 water:DMSO v/v) and a total chloride reading was taken at 9 min. The initial value was set at 0 % chloride efflux and the final chloride reading (at 9 minutes) was set as 100 % chloride efflux. All other data points were calibrated to these points.

S6.2 Initial rate of chloride efflux (k_{ini}) determination

The initial rate of chloride release (k_{ini}) was obtained from chloride/nitrate transport experiments. For this a series of unilamellar 1-palmitoyl-2-oleoylphospatidylcholine (POPC) vesicles of defined size (200 nm in diameter) were prepared. The vesicles were loaded with a buffered sodium chloride solution (489 mM in 5 mM phosphate buffer at pH 7.2) and suspended in an isotonic sodium nitrate solution. A thiourea was then added as a solution in a small amount of DMSO (2 mol% thiourea to lipid) and the resultant transport of chloride out of the vesicles was monitored using an ion selective electrode (ISE). At the end of the experiment, the vesicles were lysed by addition of detergent and the final reading was used to calibrate the ISE to 100 % chloride release. It is possible to fit the obtained chloride efflux with the following asymptotic function using Origin 8.1, where y is the chloride efflux (%) and x is time (s):

$$y = a - b \cdot c^x$$

The initial rate of chloride release (k_{ini}) is then given by $k_{ini} = -b \cdot ln(c)$ and is obtained in % s^{-1}. To ensure repeatability and precise data, the whole procedure was repeated a minimum of 3 times and was conducted each time with a newly prepared set of vesicles, a different experimenter and a different chloride selective electrode (out of 4 *Accumet* electrodes in total). The data given in the article represents the average initial rate of chloride release (k_{ini}) between each repeat and the errors are given by standard deviations. Within each repeat, the chloride efflux by 2% thiourea was monitored a further 3 times to ensure stability of the data during that repeat. Figures S188-S209 show all of the obtained transport data for 2% thiourea to lipid and the corresponding asymptotic fit (both training and test set). An overview of the obtained initial rates of chloride release (k_{ini}) can be found in Table S21 in Section S6.5.



Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	1.45617	
Adj. R-Square	0.99806	
	Value	Standard Error
a	91.57293	1.85028
b	88.48559	1.60709
c	0.99262	3.72111E-4

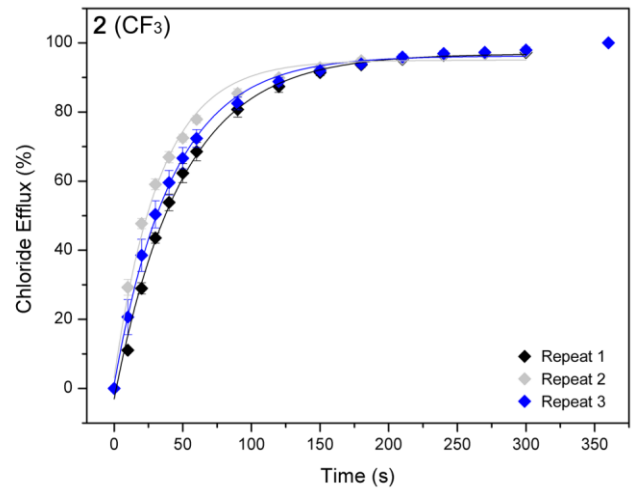
Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.44558	
Adj. R-Square	0.99937	
	Value	Standard Error
a	89.09452	1.39743
b	90.68496	1.22649
c	0.99382	2.04427E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.44875	
Adj. R-Square	0.99942	
	Value	Standard Error
a	92.38799	1.10717
b	90.8356	0.9617
c	0.99294	2.01661E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	0.38344	
Adj. R-Square	0.99957	
	Value	Standard Error
a	94.04883	0.68679
b	92.58841	0.61864
c	0.991	1.84127E-4

Repeat 5	y = a-b*c^x	
Reduced Chi-Sqr	1.59605	
Adj. R-Square	0.99771	
	Value	Standard Error
a	89.45931	2.09146
b	86.51267	1.81672
c	0.99294	3.99345E-4

Figure S188. Overview of the initial rate of chloride release (k_{ini}) for compound **1**. For experimental details, see main text.

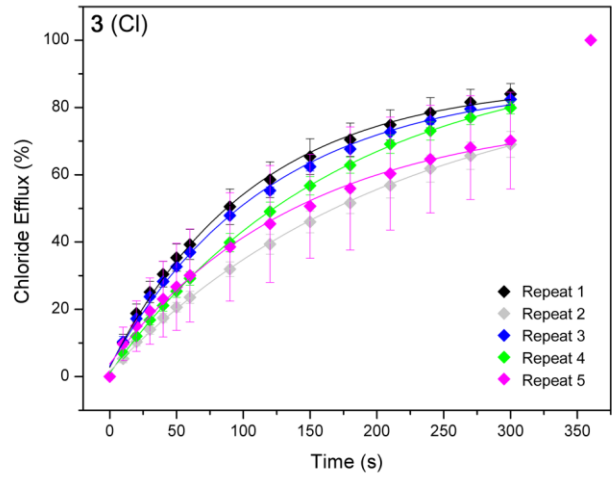


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	3.13833	
Adj. R-Square	0.99711	
	Value	Standard Error
a	96.85573	0.81835
b	99.85959	1.48127
c	0.97979	6.9795E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	5.65975	
Adj. R-Square	0.99321	
	Value	Standard Error
a	95.01391	0.93363
b	91.91593	2.13859
c	0.97057	0.0014

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	1.9397	
Adj. R-Square	0.99796	
	Value	Standard Error
a	96.22905	0.60005
b	95.00415	1.19613
c	0.97691	6.38873E-4

Figure S189. Overview of the initial rate of chloride release (k_{ini}) for compound **2**. For experimental details, see main text.



Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	2.06074	
Adj. R-Square	0.99729	
	Value	Standard Error
a	87.5332	1.51802
b	84.72564	1.38488
c	0.99072	4.68451E-4

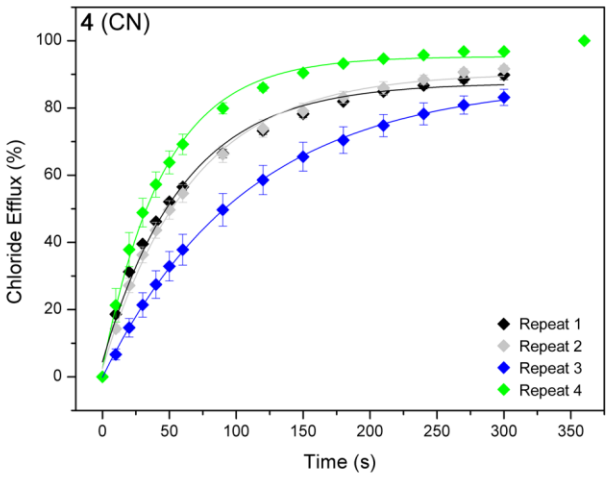
Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.48752	
Adj. R-Square	0.99909	
	Value	Standard Error
a	90.5622	2.44617
b	89.01267	2.23327
c	0.99533	2.34124E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	2.04752	
Adj. R-Square	0.99718	
	Value	Standard Error
a	87.65141	1.78979
b	84.52845	1.5765
c	0.99166	4.62755E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	0.19281	
Adj. R-Square	0.99974	
	Value	Standard Error
a	96.42004	1.01356
b	95.40441	0.89568
c	0.99415	1.29097E-4

Repeat 5	y = a-b*c^x	
Reduced Chi-Sqr	2.46141	
Adj. R-Square	0.99522	
	Value	Standard Error
a	78.07671	2.61259
b	74.30166	2.2696
c	0.99297	5.77562E-4

Figure S190. Overview of the initial rate of chloride release (k_{ini}) for compound **3**. For experimental details, see main text.



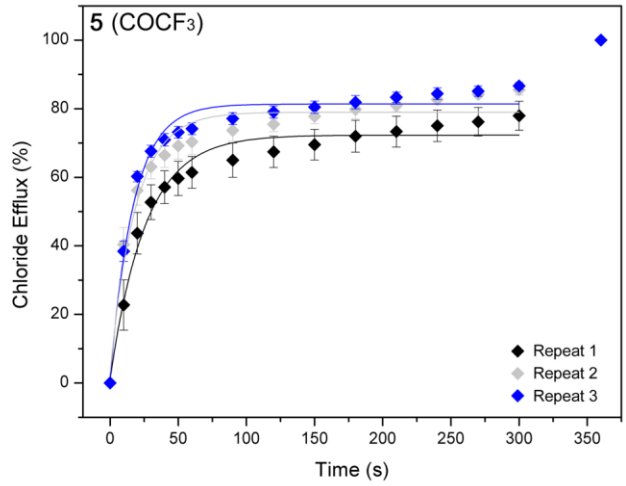
Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	6.4327	
Adj. R-Square	0.99173	
	Value	Standard Error
a	87.45684	1.34613
b	83.00552	2.04462
c	0.98343	0.00105

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	3.13075	
Adj. R-Square	0.99643	
	Value	Standard Error
a	90.3975	1.03124
b	87.94269	1.40918
c	0.98509	6.53471E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.28232	
Adj. R-Square	0.99965	
	Value	Standard Error
a	87.89094	0.57035
b	88.04334	0.51812
c	0.99081	1.66618E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	3.2144	
Adj. R-Square	0.99653	
	Value	Standard Error
a	95.32561	0.79907
b	92.85264	1.51877
c	0.97844	7.97244E-4

Figure S191. Overview of the initial rate of chloride release (k_{ini}) for compound **4**. For experimental details, see main text.

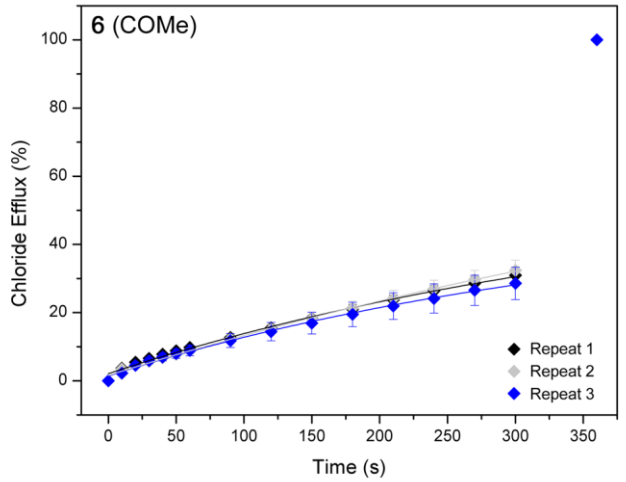


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	12.89562	
Adj. R-Square	0.97254	
	Value	Standard Error
a	72.30686	1.31757
b	71.10119	3.3598
c	0.96229	0.00351

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	20.1464	
Adj. R-Square	0.95863	
	Value	Standard Error
a	78.96818	1.53525
b	75.70279	4.38439
c	0.94786	0.00598

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	11.69147	
Adj. R-Square	0.97757	
	Value	Standard Error
a	81.35658	1.14917
b	79.96535	3.3726
c	0.94315	0.00479

Figure S192. Overview of the initial rate of chloride release (k_{ini}) for compound **5**. For experimental details, see main text.

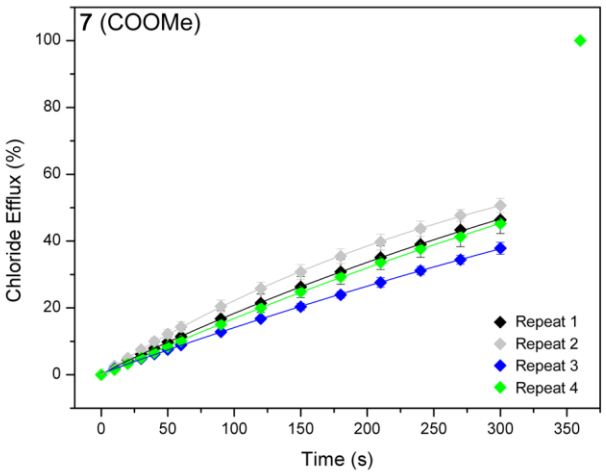


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.59828	
Adj. R-Square	0.99369	
	Value	Standard Error
a	58.90852	10.47671
b	56.81079	10.21027
c	0.99769	5.84302E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.29727	
Adj. R-Square	0.99725	
	Value	Standard Error
a	87.56504	18.18791
b	86.01582	17.99655
c	0.99854	3.79006E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.39325	
Adj. R-Square	0.99532	
	Value	Standard Error
a	49.94053	6.44597
b	48.54565	6.23226
c	0.99734	5.06814E-4

Figure S193. Overview of the initial rate of chloride release (k_{ini}) for compound **6**. For experimental details, see main text.



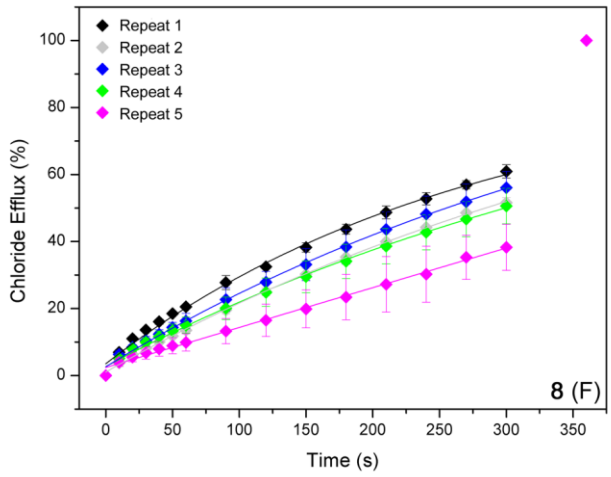
Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.04271	
Adj. R-Square	0.99983	
	Value	Standard Error
a	119.984	5.58558
b	119.61987	5.51325
c	0.99837	9.51757E-5

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.02398	
Adj. R-Square	0.99985	
	Value	Standard Error
a	141.10716	10.42351
b	140.75453	10.36882
c	0.99897	8.78022E-5

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.03052	
Adj. R-Square	0.9999	
	Value	Standard Error
a	85.58573	1.44817
b	85.40963	1.38931
c	0.99703	7.51876E-5

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	0.0343	
Adj. R-Square	0.99986	
	Value	Standard Error
a	143.63787	8.18689
b	143.97175	8.12169
c	0.99873	8.61327E-5

Figure S194. Overview of the initial rate of chloride release (k_{ini}) for compound **7**. For experimental details, see main text.



Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	1.87416	
Adj. R-Square	0.99504	
	Value	Standard Error
a	88.44557	7.69346
b	84.92103	7.24602
c	0.99636	5.33287E-4

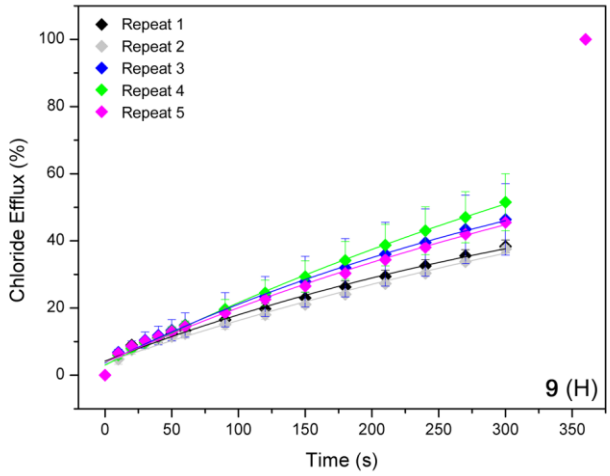
Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	0.98724	
Adj. R-Square	0.99623	
	Value	Standard Error
a	103.65425	16.13553
b	100.95039	15.79143
c	0.99789	4.49574E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.22654	
Adj. R-Square	0.99925	
	Value	Standard Error
a	111.18451	8.05569
b	110.13245	7.89067
c	0.99793	2.00298E-4

Repeat 5	y = a-b*c^x	
Reduced Chi-Sqr	0.79486	
Adj. R-Square	0.99449	
	Value	Standard Error
a	2210.0874	21263.10579
b	2207.6871	21262.78836
c	0.99995	5.26765E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.92416	
Adj. R-Square	0.9972	
	Value	Standard Error
a	110.57679	13.609
b	107.94517	13.27738
c	0.99774	3.8795E-4

Figure S195. Overview of the initial rate of chloride release (k_{ini}) for compound **8**. For experimental details, see main text.



Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	2.3056	
Adj. R-Square	0.98259	
	Value	Standard Error
a	71.3336	20.84396
b	67.14353	20.32068
c	0.9977	9.74911E-4

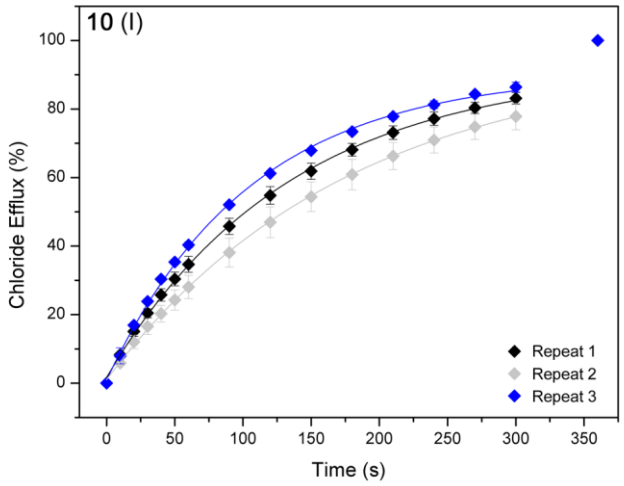
Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	1.10174	
Adj. R-Square	0.99587	
	Value	Standard Error
a	129.6713	29.89994
b	126.68938	29.53233
c	0.99841	4.6574E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	2.29005	
Adj. R-Square	0.98153	
	Value	Standard Error
a	94.39358	49.29485
b	90.59597	48.76389
c	0.99852	9.89437E-4

Repeat 5	y = a-b*c^x	
Reduced Chi-Sqr	1.82875	
Adj. R-Square	0.99069	
	Value	Standard Error
a	99.58265	28.09677
b	95.77825	27.62572
c	0.99814	7.04603E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	1.63723	
Adj. R-Square	0.99224	
	Value	Standard Error
a	87.97707	17.15001
b	84.38335	16.70941
c	0.99767	6.48455E-4

Figure S196. Overview of the initial rate of chloride release (k_{ini}) for compound **9**. For experimental details, see main text.

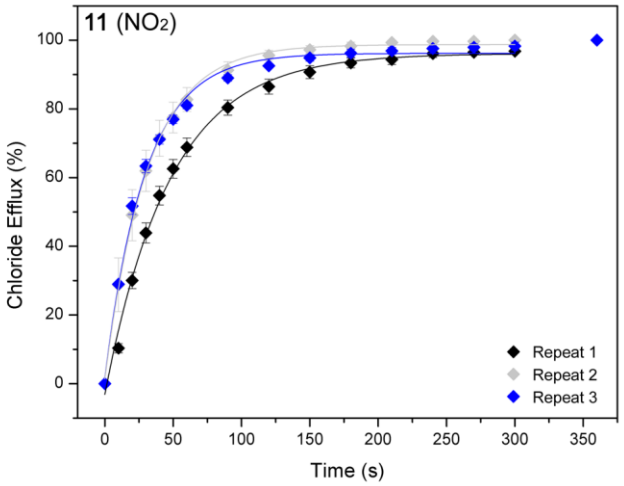


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.52274	
Adj. R-Square	0.99933	
	Value	Standard Error
a	92.03397	1.08573
b	90.38499	0.9434
c	0.99253	2.18191E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.28265	
Adj. R-Square	0.9996	
	Value	Standard Error
a	95.05326	1.31652
b	94.01172	1.16963
c	0.99437	1.59941E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.75832	
Adj. R-Square	0.99911	
	Value	Standard Error
a	90.57216	0.89865
b	89.69467	0.82596
c	0.99057	2.69102E-4

Figure S197. Overview of the initial rate of chloride release (k_{ini}) for compound **10**. For experimental details, see main text.

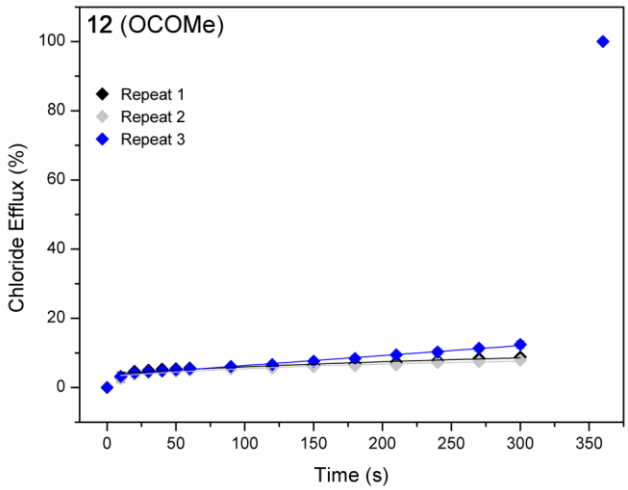


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	4.19209	
Adj. R-Square	0.99607	
	Value	Standard Error
a	96.12292	0.93377
b	99.22318	1.71968
c	0.97933	8.25234E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	1.66225	
Adj. R-Square	0.99818	
	Value	Standard Error
a	98.7222	0.49722
b	97.26696	1.17045
c	0.96883	7.56546E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	4.21524	
Adj. R-Square	0.99503	
	Value	Standard Error
a	96.1911	0.77873
b	94.46681	1.88217
c	0.96693	0.00132

Figure S198. Overview of the initial rate of chloride release (k_{ini}) for compound **11**. For experimental details, see main text.

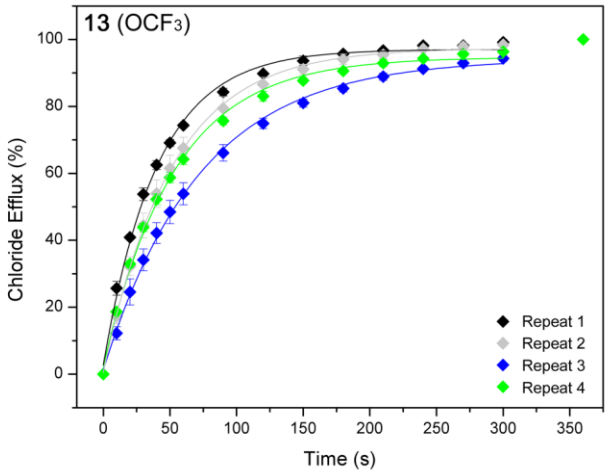


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.23719	
Adj. R-Square	0.90498	
	Value	Standard Error
a	12.01132	4.90516
b	8.02886	4.6748
c	0.99719	0.00255

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.15553	
Adj. R-Square	0.92972	
	Value	Standard Error
a	8.771	1.27332
b	5.58805	1.09895
c	0.99489	0.00231

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.07546	
Adj. R-Square	0.9908	
	Value	Standard Error
a	2330.4660	136036.67086
b	2327.0035	136036.54366
c	0.99999	7.26816E-4

Figure S199. Overview of the initial rate of chloride release (k_{ini}) for compound **12**. For experimental details, see main text.



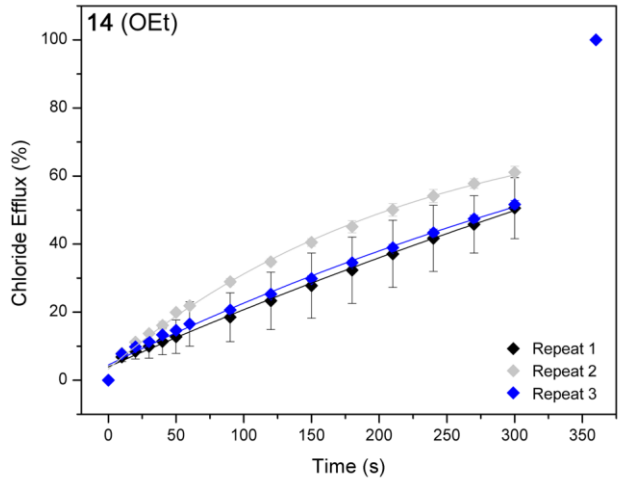
Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	3.44702	
Adj. R-Square	0.99627	
	Value	Standard Error
a	97.04677	0.78083
b	94.27635	1.61165
c	0.97563	8.97542E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.93299	
Adj. R-Square	0.99909	
	Value	Standard Error
a	97.40085	0.4585
b	96.80877	0.80051
c	0.98067	3.80239E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	1.59422	
Adj. R-Square	0.99837	
	Value	Standard Error
a	94.25781	0.80274
b	92.98839	1.00382
c	0.98632	4.22765E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	2.52081	
Adj. R-Square	0.99733	
	Value	Standard Error
a	94.70258	0.77283
b	92.39665	1.30605
c	0.98139	6.37776E-4

Figure S200. Overview of the initial rate of chloride release (k_{ini}) for compound **13**. For experimental details, see main text.

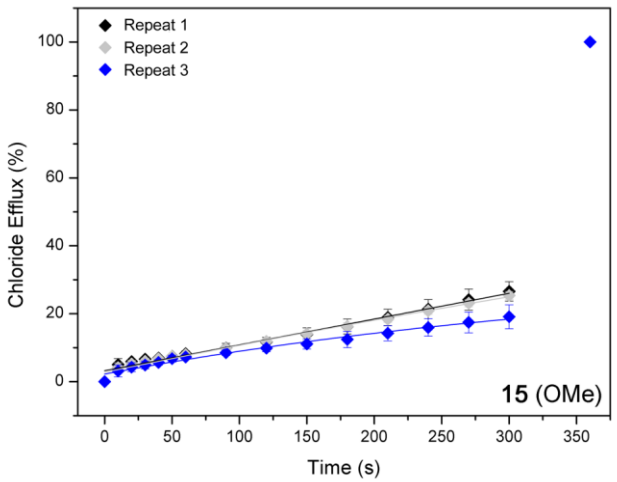


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	2.23809	
Adj. R-Square	0.99081	
	Value	Standard Error
a	188.0056	116.15126
b	184.08596	115.62241
c	0.99904	6.89487E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	1.53101	
Adj. R-Square	0.99606	
	Value	Standard Error
a	80.2468	4.6433
b	76.96329	4.26088
c	0.9955	4.86127E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	2.47935	
Adj. R-Square	0.99016	
	Value	Standard Error
a	120.67738	39.06567
b	116.22242	38.51539
c	0.9983	7.22375E-4

Figure S201. Overview of the initial rate of chloride release (k_{ini}) for compound **14**. For experimental details, see main text.

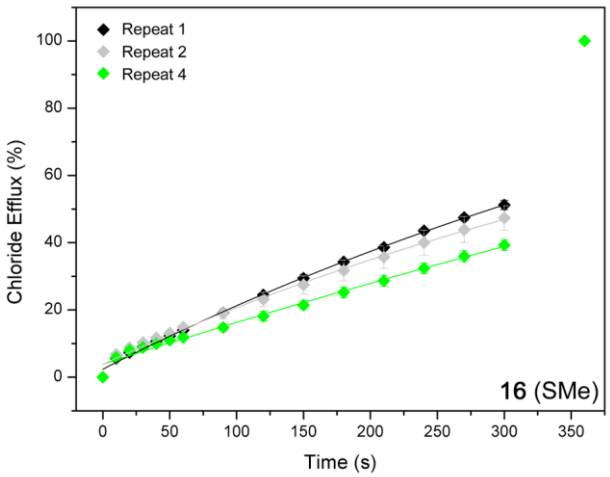


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	1.42698	
Adj. R-Square	0.97618	
	Value	Standard Error
a	3848.7197	213551.16931
b	3845.5050	213550.74395
c	0.99998	0.0011

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.92586	
Adj. R-Square	0.98383	
	Value	Standard Error
a	114.43059	124.49849
b	111.69684	124.15756
c	0.99926	9.14524E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.78535	
Adj. R-Square	0.97489	
	Value	Standard Error
a	33.71554	11.13714
b	31.50257	10.83265
c	0.9976	0.00118

Figure S202. Overview of the initial rate of chloride release (k_{ini}) for compound **15**. For experimental details, see main text.

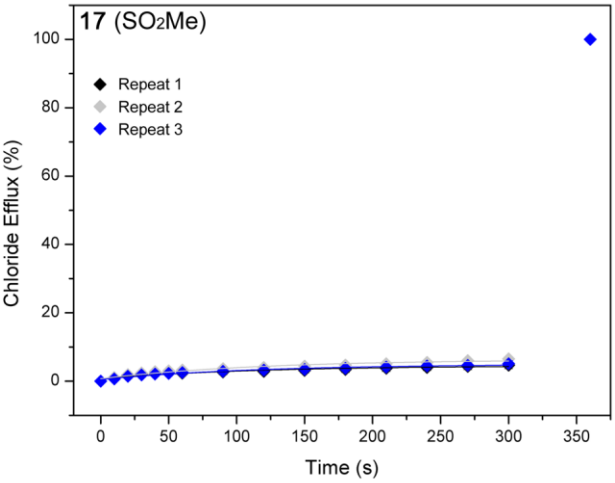


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.75755	
Adj. R-Square	0.99726	
	Value	Standard Error
a	133.84032	25.98823
b	131.47904	25.6832
c	0.99845	3.78984E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	1.81652	
Adj. R-Square	0.99161	
	Value	Standard Error
a	112.05894	33.80148
b	108.24063	33.33039
c	0.99831	6.66638E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	1.88732	
Adj. R-Square	0.98663	
	Value	Standard Error
a	211.94068	257.84643
b	208.06351	257.3591
c	0.99939	8.28984E-4

Figure S203. Overview of the initial rate of chloride release (k_{ini}) for compound **16**. For experimental details, see main text.

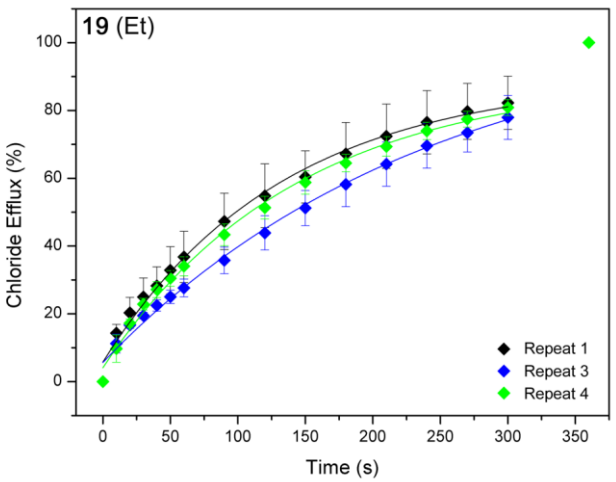


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	0.12584	
Adj. R-Square	0.9224	
	Value	Standard Error
a	4.52251	0.41139
b	3.82516	0.36695
c	0.99125	0.00255

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	0.18081	
Adj. R-Square	0.94626	
	Value	Standard Error
a	6.33395	0.51671
b	5.60336	0.45707
c	0.99151	0.00208

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	0.08265	
Adj. R-Square	0.9621	
	Value	Standard Error
a	4.96412	0.33075
b	4.51147	0.2955
c	0.99121	0.00175

Figure S204. Overview of the initial rate of chloride release (k_{ini}) for compound **17**. For experimental details, see main text.

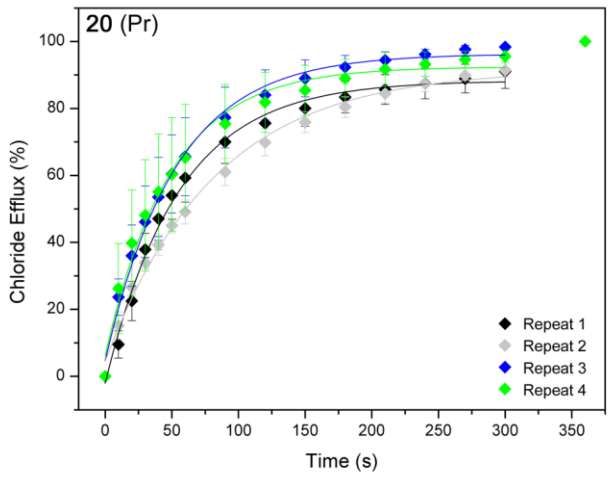


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	5.16392	
Adj. R-Square	0.99247	
	Value	Standard Error
a	89.62244	3.32686
b	83.87078	2.8929
c	0.99242	7.38848E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	4.57914	
Adj. R-Square	0.99253	
	Value	Standard Error
a	106.67516	9.51097
b	101.02215	8.83108
c	0.99589	6.63288E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	3.40959	
Adj. R-Square	0.995	
	Value	Standard Error
a	89.57087	3.07189
b	85.48609	2.66856
c	0.99296	5.90804E-4

Figure S205. Overview of the initial rate of chloride release (k_{ini}) for compound **19**. For experimental details, see main text.



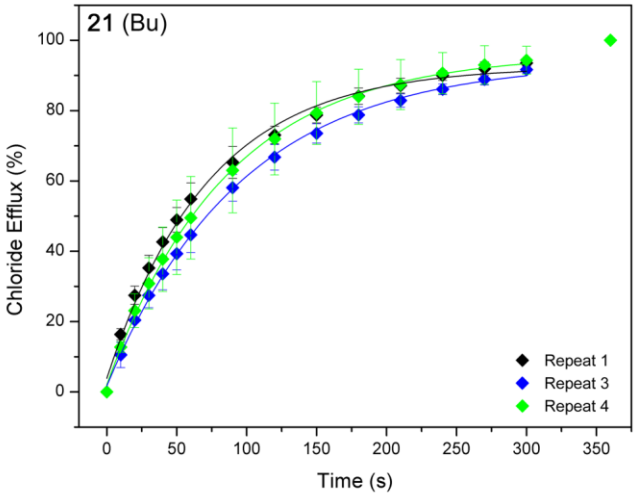
Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	5.66023	
Adj. R-Square	0.99378	
	Value	Standard Error
a	88.31347	1.18347
b	90.38729	1.94563
c	0.98197	9.56798E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	4.75617	
Adj. R-Square	0.99446	
	Value	Standard Error
a	91.67989	1.59306
b	87.22356	1.76485
c	0.98788	7.4061E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	5.87272	
Adj. R-Square	0.99369	
	Value	Standard Error
a	96.29858	1.17927
b	91.61105	1.99362
c	0.98139	9.82077E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	11.57988	
Adj. R-Square	0.98582	
	Value	Standard Error
a	92.43302	1.57283
b	86.20108	2.84483
c	0.9798	0.00155

Figure S206. Overview of the initial rate of chloride release (k_{ini}) for compound **20**. For experimental details, see main text.

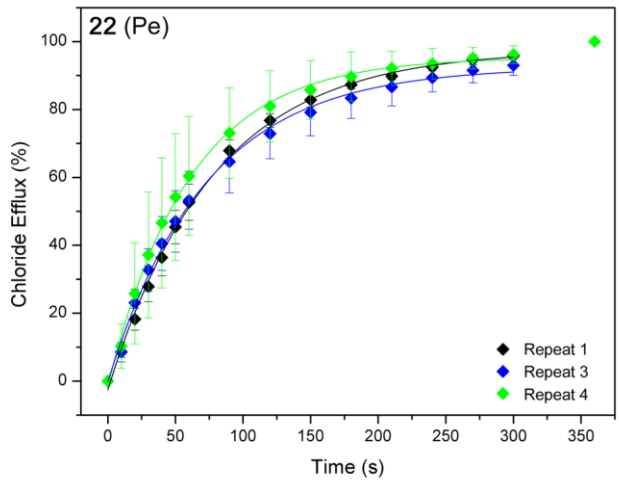


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	4.34416	
Adj. R-Square	0.99515	
	Value	Standard Error
a	92.60213	1.32408
b	88.76889	1.657
c	0.98631	7.31296E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	1.06202	
Adj. R-Square	0.99886	
	Value	Standard Error
a	94.02652	0.93115
b	92.33127	0.90242
c	0.98967	3.1501E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	0.99151	
Adj. R-Square	0.999	
	Value	Standard Error
a	96.15659	0.76587
b	94.14145	0.81662
c	0.98836	3.08832E-4

Figure S207. Overview of the initial rate of chloride release (k_{ini}) for compound **21**. For experimental details, see main text.

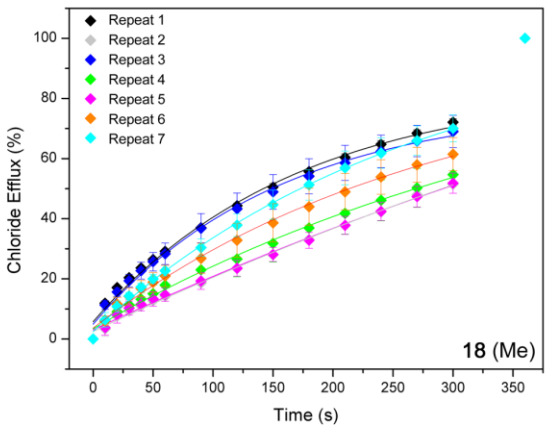


Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	1.90959	
Adj. R-Square	0.99831	
	Value	Standard Error
a	97.58913	0.95106
b	100.12444	1.10626
c	0.98726	4.16731E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	2.81074	
Adj. R-Square	0.99711	
	Value	Standard Error
a	92.52287	1.06714
b	92.66696	1.33294
c	0.98634	5.63016E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	1.90524	
Adj. R-Square	0.99819	
	Value	Standard Error
a	95.37667	0.73815
b	96.73518	1.11122
c	0.98358	4.89104E-4

Figure S208. Overview of the initial rate of chloride release (k_{ini}) for compound **22**. For experimental details, see main text.



Repeat 1	y = a-b*c^x	
Reduced Chi-Sqr	4.75671	
Adj. R-Square	0.99062	
	Value	Standard Error
a	83.95442	5.01465
b	78.22005	4.43015
c	0.99414	7.81742E-4

Repeat 2	y = a-b*c^x	
Reduced Chi-Sqr	1.28895	
Adj. R-Square	0.9953	
	Value	Standard Error
a	181.61229	71.94254
b	179.12895	71.5417
c	0.99894	4.93023E-4

Repeat 3	y = a-b*c^x	
Reduced Chi-Sqr	3.5057	
Adj. R-Square	0.99261	
	Value	Standard Error
a	78.4342	3.71671
b	73.52653	3.2514
c	0.99364	7.03996E-4

Repeat 4	y = a-b*c^x	
Reduced Chi-Sqr	1.86341	
Adj. R-Square	0.99372	
	Value	Standard Error
a	103.29167	18.17284
b	99.89139	17.70289
c	0.99767	5.83072E-4

Repeat 5	y = a-b*c^x	
Reduced Chi-Sqr	1.79354	
Adj. R-Square	0.99328	
	Value	Standard Error
a	176.41486	81.31451
b	173.39722	80.84181
c	0.99892	5.90044E-4

Repeat 6	y = a-b*c^x	
Reduced Chi-Sqr	2.39348	
Adj. R-Square	0.99383	
	Value	Standard Error
a	91.20064	9.24287
b	87.57745	8.73421
c	0.99648	5.93485E-4

Repeat 7	y = a-b*c^x	
Reduced Chi-Sqr	0.95762	
Adj. R-Square	0.99821	
	Value	Standard Error
a	102.12132	5.2028
b	99.65706	4.88483
c	0.99626	3.20442E-4

Figure S209. Overview of the initial rate of chloride release (k_{ini}) for compound **18**. For experimental details, see main text.

S6.3 Evidence for antiport and mobile carrier mechanisms

Symport vs. antiport: Cesium

To test whether the transporters function as symporters or antiporters, the chloride/nitrate tests were repeated with CsCl encapsulated in the vesicles instead of NaCl. Figure S210 shows that there is little difference (within error) between the chloride efflux observed in the case of CsCl and NaCl, indicating an antiport process.

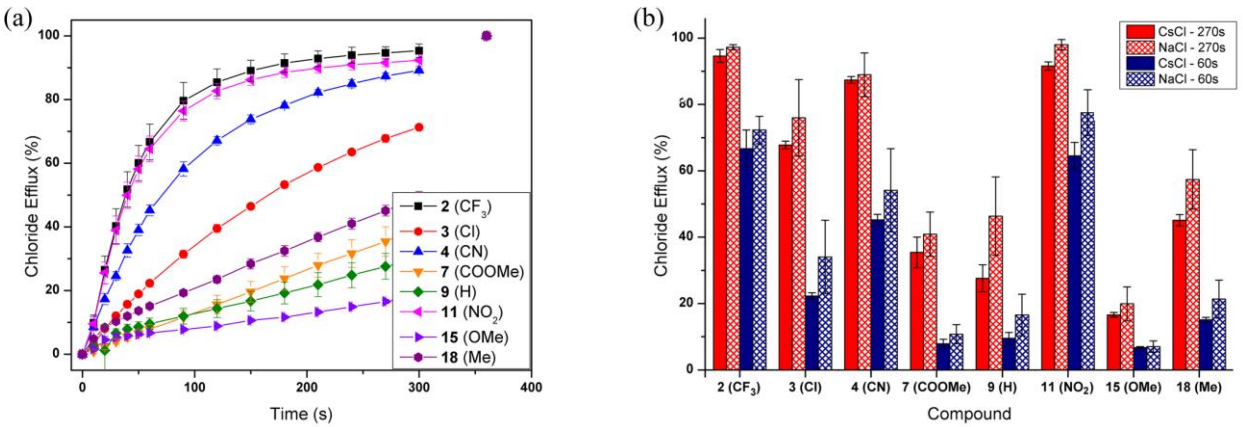


Figure S210. (a) Chloride efflux promoted by a selection of **1-22** (2 % molar carrier to lipid) from unilamellar POPC vesicles loaded with 489 mM CsCl buffered to pH 7.2 with 5 mM sodium phosphate salts. The vesicles were dispersed in 489 mM $NaNO_3$ buffered to pH 7.2 with 5 mM sodium phosphate salts. At the end of the experiment, detergent was added to lyse the vesicles and calibrate the ISE to 100 % chloride efflux. Each point represents the average of three trials. (b) Comparison between the chloride efflux from vesicles loaded with CsCl and with NaCl for different time intervals (60s and 270s after the addition of transporter).

Symport vs. antiport: Bicarbonate

To test whether the transporters function as symporters or antiporters, tests were conducted where the extravesicular anion was either nitrate, sulfate or bicarbonate. Figure S211 represent the results of these chloride/bicarbonate tests (experimental procedure in Section S6.1) and show a profound difference between the chloride efflux observed in the presence of sulfate and bicarbonate, indicating an antiport process.

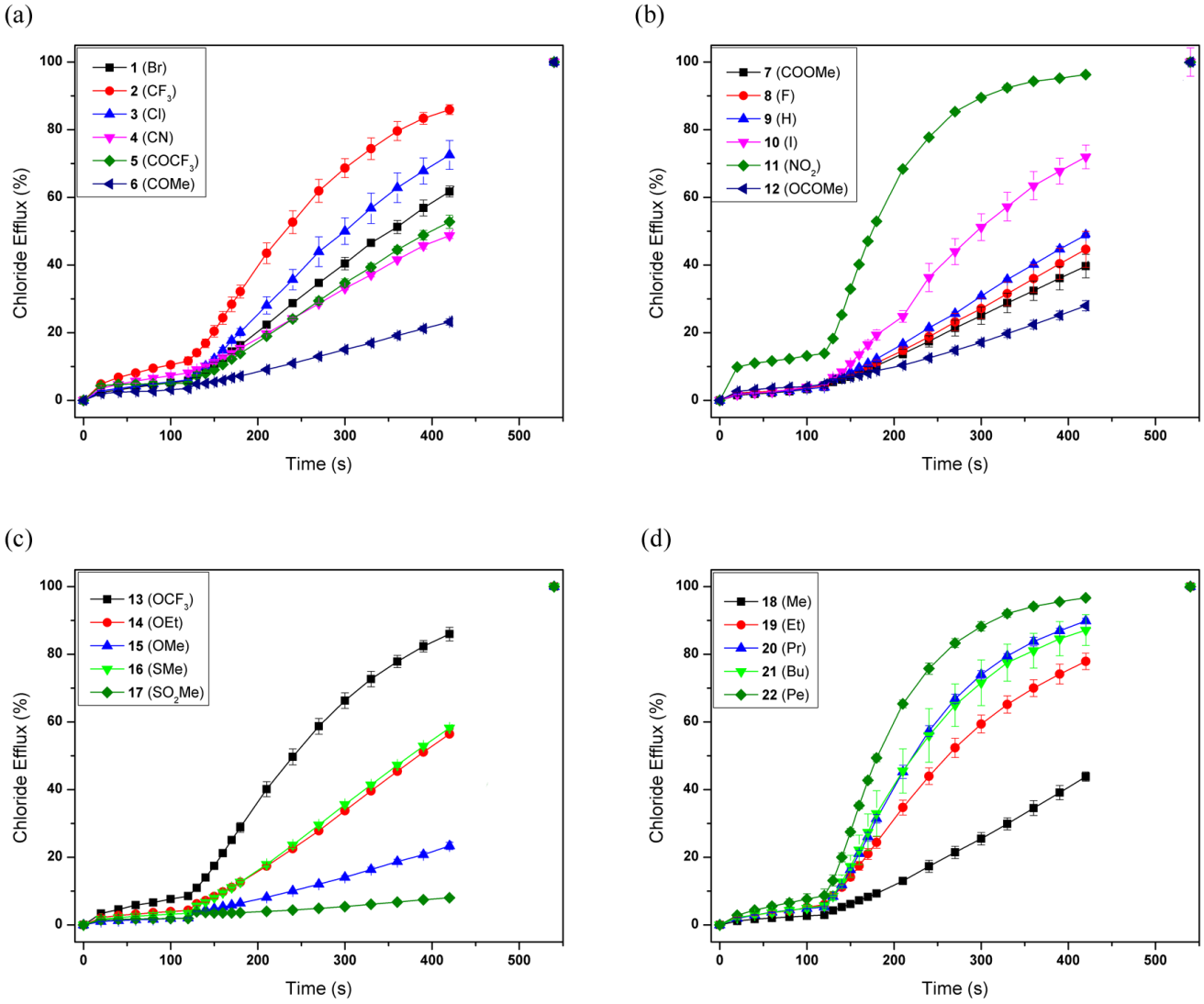


Figure S211. Chloride efflux promoted by **1-22** (2 % molar carrier to lipid) from unilamellar POPC vesicles loaded with 450 mM NaCl buffered to pH 7.2 with 20 mM sodium phosphate salts. The vesicles were dispersed in 162 mM Na2SO4 buffered to pH 7.2 with 20 mM sodium phosphate salts. At t = 120s, a solution of NaHCO3 was added to give a 40 mM external concentration. At the end of the experiment, detergent was added to lyse the vesicles and calibrate the ISE to 100 % chloride efflux. Each point represents the average of three trials. (a) compounds **1-6**. (b) compounds **7-12**. (c) compounds **13-17**. (d) compounds **18-22**.

Mobile carrier vs. ion channel: U-tube experiment

In a U-tube experiment the lipid bilayer is substituted with a bulk organic phase. In these conditions ion channel formation is virtually impossible. The organic phase consisted of 20 mL nitrobenzene (for solubility reasons) and contained 1mM of carrier and 1mM of $TBAPF_6$. A control experiment was executed with neat nitrobenzene (no carrier). The same aqueous phases were used as for the vesicle experiments. The donating phase contained 500 mM NaCl and was buffered to pH 7.2 with 5 mM phosphate salts (10 mL). The receiving phase contained 500 mM $NaNO_3$ and was buffered to pH 7.2 with 5 mM phosphate salts (10 mL). The change in chloride concentration of the receiving phase was monitored with a chloride-selective electrode (the electrode was calibrated to convert the potential readings (mV) to chloride concentrations (M)). The experiments were conducted at room temperature and the results are shown in Figure S212 and show that chloride transport through a bulk organic layer is possible, indicating that the compounds function as mobile carriers.

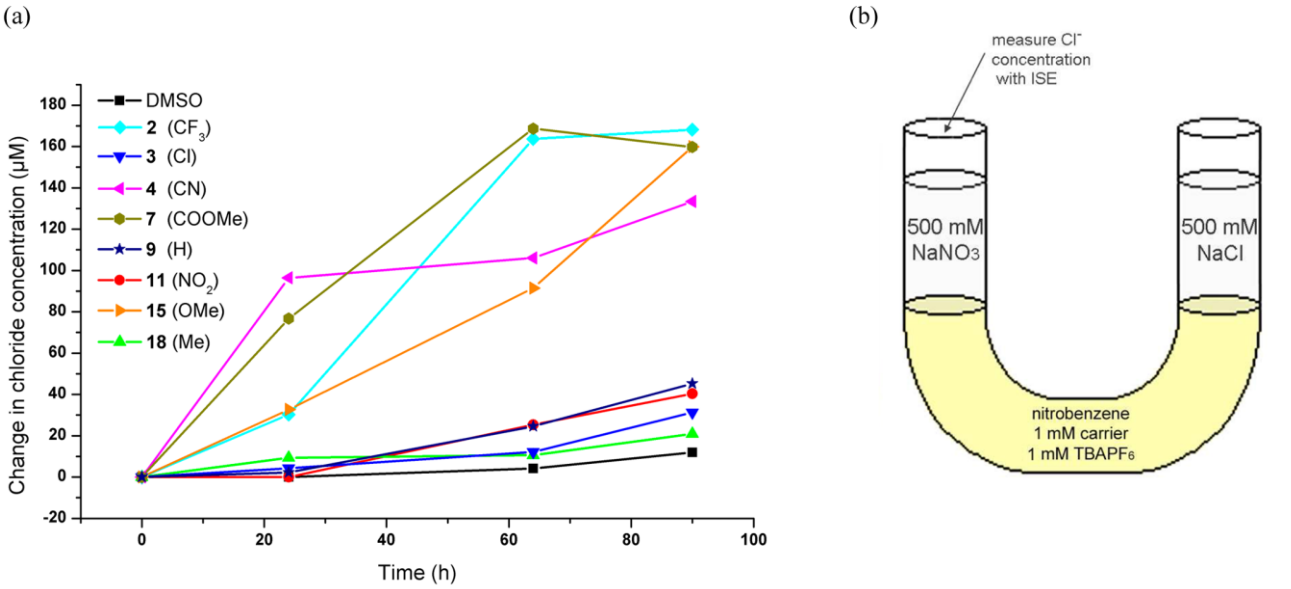


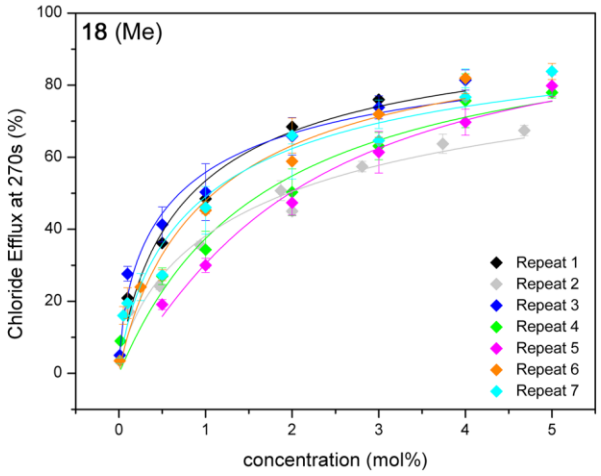
Figure S212. U-tube experiment in nitrobenzene. (a) Graph showing the change in the chloride concentration of the receiving aqueous phase as a function of time for a selection of compounds. (b) Experimental set-up.

S6.4 Hill plots

During the Hill plots the chloride/nitrate transport assays were performed as described above (see section S6.1) for various concentrations of carrier. The chloride efflux (%) 270 s after the addition of carrier was plotted as a function of the carrier concentration. Data points were fitted to the Hill equation using Origin 8.1:

$$y = V_{\max}\frac{x^n}{k^n + x^n} = 100\%\frac{x^n}{(EC_{50})^n + x^n}$$

where y is the chloride efflux at 270 s (%) and x is the carrier concentration (mol% carrier to lipid). V_{max}, k and n are the parameters to be fitted. V_{max} is the maximum efflux possible (this was fixed to 100%, as this is physically the maximum chloride efflux possible), n is the Hill coefficient and k is the carrier concentration needed to reach $V_{max}/2$ (when V_{max} is fixed to 100%, k equals EC_{50}). From the Hill plot it is therefore possible to directly obtain $EC_{50,270s}$ values, defined as the carrier concentration (molar % carrier to lipid) needed to obtain 50 % chloride efflux after 270 s. To ensure repeatability and precise data, each Hill plot was repeated a minimum of 3 times and was conducted each time with a newly prepared set of vesicles, a different experimenter and a different chloride selective electrode (out of 4 *Accumet* electrodes in total). The data given in the article represents the average EC_{50} values between each Hill plot and the errors are given by standard deviations. Within each Hill plot repeat, the chloride efflux for every concentration was monitored a further 3 times to ensure stability of the data during that repeat. Figures S213-S234 show all of the obtained Hill plots (including the results of the fit) and an overview of the obtained EC_{50}, n and the converted $\log(1/EC_{50})$ values can be found in Tables S21 and S22 in Section S6.5.



Repeat 1	Hill	
Reduced Chi-Sqr	22.47133	
Adj. R-Square	0.96098	
	Value	Standard Error
Vmax	100	0
EC50	0.8476	0.10254
n	0.83025	0.09979

Repeat 2	Hill	
Reduced Chi-Sq	13.16523	
Adj. R-Square	0.96057	
	Value	Standard Error
Vmax	100	0
EC50	1.95082	0.15587
n	0.72397	0.07148

Repeat 3	Hill	
Reduced Chi-Sq	21.79835	
Adj. R-Square	0.97028	
	Value	Standard Error
Vmax	100	0
EC50	0.70318	0.09909
n	0.65952	0.07576

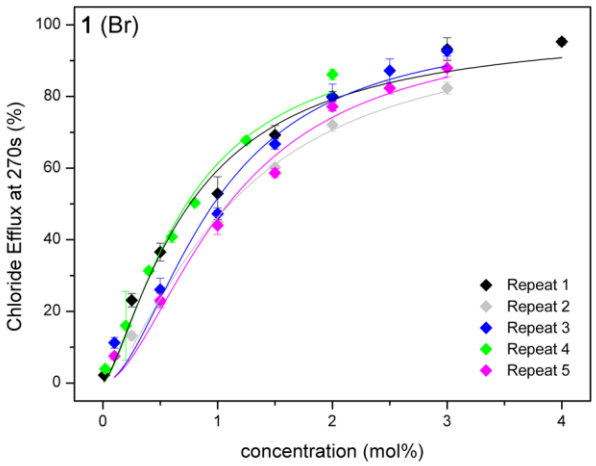
Repeat 4	Hill	
Reduced Chi-Sqr	27.75873	
Adj. R-Square	0.95885	
	Value	Standard Error
Vmax	100	0
EC50	1.65662	0.1684
n	1.02	0.13799

Repeat 5	Hill	
Reduced Chi-Sq	9.90926	
Adj. R-Square	0.98207	
	Value	Standard Error
Vmax	100	0
EC50	1.97474	0.10235
n	1.21669	0.09441

Repeat 6	Hill	
Reduced Chi-Sq	18.0445	
Adj. R-Square	0.97952	
	Value	Standard Error
Vmax	100	0
k	1.09726	0.11801
n	0.90317	0.10671

Repeat 7	Hill	
Reduced Chi-Sq	43.51925	
Adj. R-Square	0.93835	
	Value	Standard Error
Vmax	100	0
k	1.05162	0.1662
n	0.78528	0.11174

Figure S213. Overview of the Hill plots for compound **18**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	23.52585	
Adj. R-Square	0.97951	
	Value	Standard Error
Vmax	100	0
EC50	0.75836	0.05999
n	1.37717	0.14276

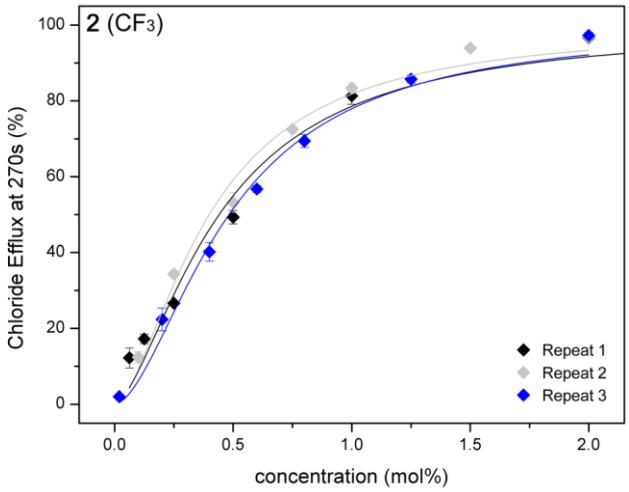
Repeat 4	Hill	
Reduced Chi-S	12.23156	
Adj. R-Square	0.98507	
	Value	Standard Error
Vmax	100	0
EC50	0.72949	0.03534
n	1.44834	0.12494

Repeat 2	Hill	
Reduced Chi-S	4.44309	
Adj. R-Square	0.99412	
	Value	Standard Error
Vmax	100	0
EC50	1.10833	0.03362
n	1.48631	0.07485

Repeat 5	Hill	
Reduced Chi-Sq	17.63359	
Adj. R-Square	0.98154	
	Value	Standard Error
Vmax	100	0
EC50	1.09329	0.06125
n	1.74428	0.17389

Repeat 3	Hill	
Reduced Chi-Sq	29.26261	
Adj. R-Square	0.97037	
	Value	Standard Error
Vmax	100	0
EC50	0.96568	0.07235
n	1.80077	0.2297

Figure S214. Overview of the Hill plots for compound **1**. For experimental details, see main text.

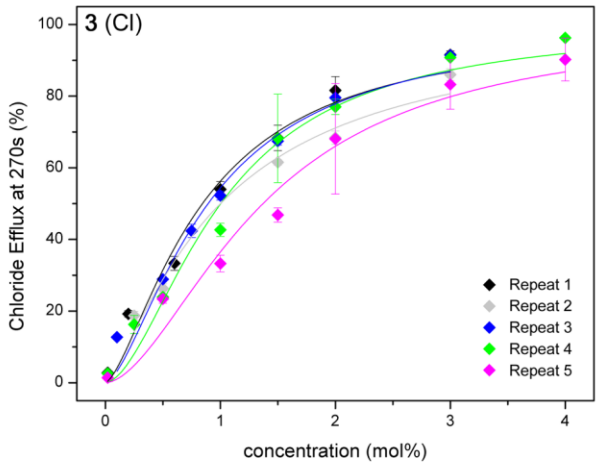


Repeat 1	Hill	
Reduced Chi-Sqr	34.3006	
Adj. R-Square	0.97564	
	Value	Standard Error
Vmax	100	0
EC50	0.44103	0.04174
n	1.58575	0.20978

Repeat 2	Hill	
Reduced Chi-S	16.84204	
Adj. R-Square	0.98339	
	Value	Standard Error
Vmax	100	0
EC50	0.39948	0.02477
n	1.6369	0.14901

Repeat 3	Hill	
Reduced Chi-Sq	12.65942	
Adj. R-Square	0.98918	
	Value	Standard Error
Vmax	100	0
EC50	0.48265	0.02267
n	1.73018	0.15052

Figure S215. Overview of the Hill plots for compound **2**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	28.12622	
Adj. R-Square	0.97416	
	Value	Standard Error
Vmax	100	0
EC50	0.84434	0.06998
n	1.48378	0.19462

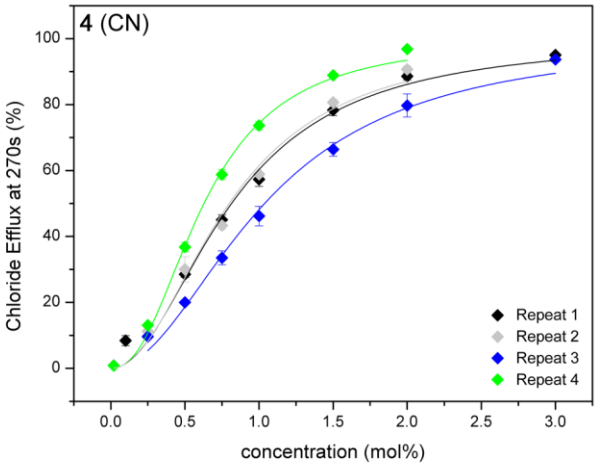
Repeat 4	Hill	
Reduced Chi-S	26.62903	
Adj. R-Square	0.97916	
	Value	Standard Error
Vmax	100	0
EC50	1.00342	0.06914
n	1.75498	0.19757

Repeat 2	Hill	
Reduced Chi-S	21.69176	
Adj. R-Square	0.97358	
	Value	Standard Error
Vmax	100	0
EC50	1.00122	0.08684
n	1.30051	0.14539

Repeat 5	Hill	
Reduced Chi-Sq	33.86961	
Adj. R-Square	0.96828	
	Value	Standard Error
Vmax	100	0
EC50	1.37104	0.09953
n	1.75472	0.24448

Repeat 3	Hill	
Reduced Chi-Sq	24.50429	
Adj. R-Square	0.96875	
	Value	Standard Error
Vmax	100	0
EC50	0.89312	0.05818
n	1.55864	0.1903

Figure S216. Overview of the Hill plots for compound **3**. For experimental details, see main text.



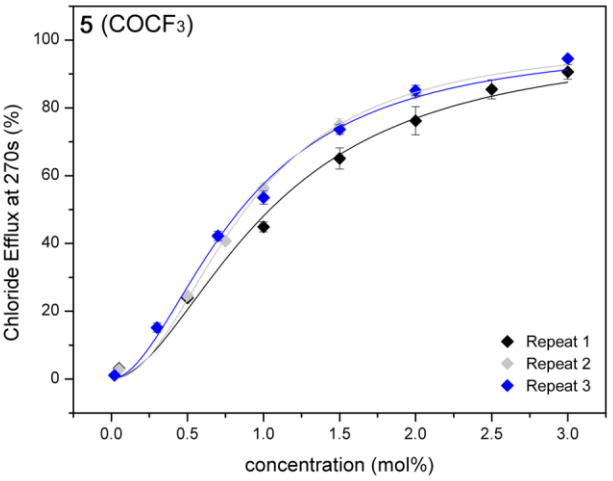
Repeat 1	Hill	
Reduced Chi-Sqr	13.94813	
Adj. R-Square	0.98654	
	Value	Standard Error
Vmax	100	0
EC50	0.81448	0.03403
n	2.03342	0.18797

Repeat 3	Hill	
Reduced Chi-Sq	8.83254	
Adj. R-Square	0.99099	
	Value	Standard Error
Vmax	100	0
EC50	1.03773	0.03245
n	2.01754	0.13153

Repeat 2	Hill	
Reduced Chi-S	10.33814	
Adj. R-Square	0.99095	
	Value	Standard Error
Vmax	100	0
EC50	0.79848	0.02766
n	2.07354	0.15954

Repeat 4	Hill	
Reduced Chi-S	4.05616	
Adj. R-Square	0.99705	
	Value	Standard Error
Vmax	100	0
EC50	0.62887	0.01376
n	2.29862	0.11386

Figure S217. Overview of the Hill plots for compound **4**. For experimental details, see main text.

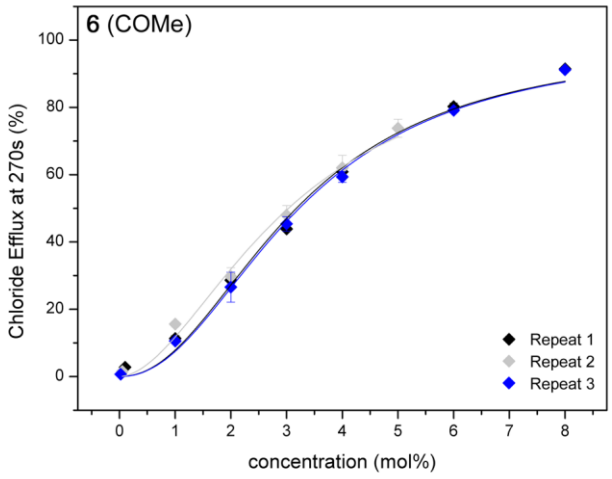


Repeat 1	Hill	
Reduced Chi-Sqr	9.0214	
Adj. R-Square	0.99167	
	Value	Standard Error
Vmax	100	0
EC50	1.04253	0.0414
n	1.84765	0.1318

Repeat 2	Hill	
Reduced Chi-S	2.55536	
Adj. R-Square	0.9977	
	Value	Standard Error
Vmax	100	0
EC50	0.88265	0.01514
n	2.08063	0.08042

Repeat 3	Hill	
Reduced Chi-Sq	7.95498	
Adj. R-Square	0.99356	
	Value	Standard Error
Vmax	100	0
EC50	0.84907	0.03084
n	1.85836	0.12831

Figure S218. Overview of the Hill plots for compound **5**. For experimental details, see main text.

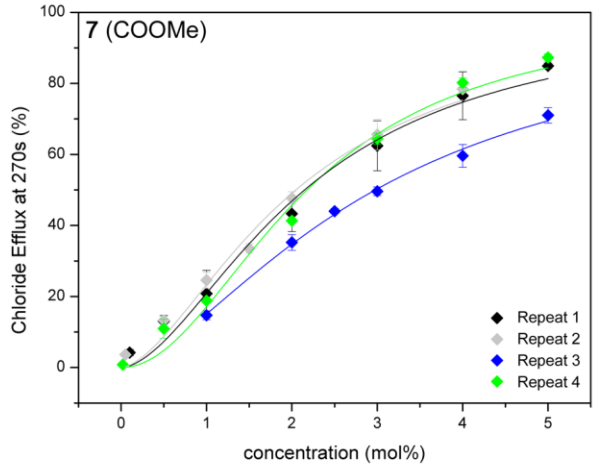


Repeat 1	Hill	
Reduced Chi-Sqr	9.32994	
Adj. R-Square	0.99182	
	Value	Standard Error
Vmax	100	0
EC50	3.17626	0.10284
n	2.13142	0.15603

Repeat 2	Hill	
Reduced Chi-S	7.60232	
Adj. R-Square	0.99015	
	Value	Standard Error
Vmax	100	0
EC50	3.03405	0.09876
n	1.8178	0.14523

Repeat 3	Hill	
Reduced Chi-Sq	5.56211	
Adj. R-Square	0.99523	
	Value	Standard Error
Vmax	100	0
EC50	3.22785	0.08002
n	2.14879	0.12149

Figure S219. Overview of the Hill plots for compound **6**. For experimental details, see main text.



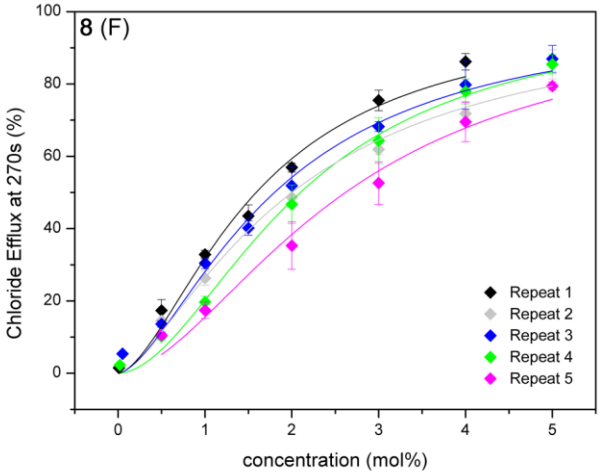
Repeat 1	Hill	
Reduced Chi-Sqr	15.62329	
Adj. R-Square	0.98476	
	Value	Standard Error
Vmax	100	0
EC50	2.14963	0.11083
n	1.74147	0.163

Repeat 3	Hill	
Reduced Chi-Sq	1.89055	
Adj. R-Square	0.99508	
	Value	Standard Error
Vmax	100	0
EC50	2.97868	0.04795
n	1.58652	0.06401

Repeat 2	Hill	
Reduced Chi-S	11.9807	
Adj. R-Square	0.98396	
	Value	Standard Error
Vmax	100	0
EC50	2.04797	0.08698
n	1.66209	0.14463

Repeat 4	Hill	
Reduced Chi-S	15.51831	
Adj. R-Square	0.98707	
	Value	Standard Error
Vmax	100	0
EC50	2.17791	0.10299
n	2.03094	0.19213

Figure S220. Overview of the Hill plots for compound **7**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	12.79997	
Adj. R-Square	0.98622	
	Value	Standard Error
Vmax	100	0
EC50	1.58999	0.0714
n	1.64459	0.14281

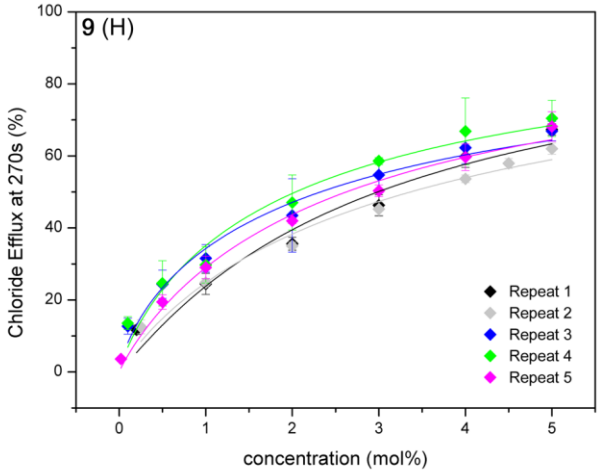
Repeat 4	Hill	
Reduced Chi-S	5.61441	
Adj. R-Square	0.99502	
	Value	Standard Error
Vmax	100	0
EC50	2.09965	0.06323
n	1.85625	0.10371

Repeat 2	Hill	
Reduced Chi-S	14.17189	
Adj. R-Square	0.98075	
	Value	Standard Error
Vmax	100	0
EC50	2.0026	0.1098
n	1.47667	0.13097

Repeat 5	Hill	
Reduced Chi-Sq	17.38847	
Adj. R-Square	0.97761	
	Value	Standard Error
Vmax	100	0
EC50	2.619	0.13431
n	1.76083	0.18785

Repeat 3	Hill	
Reduced Chi-Sq	10.40392	
Adj. R-Square	0.98846	
	Value	Standard Error
Vmax	100	0
EC50	1.80128	0.07149
n	1.59683	0.11293

Figure S221. Overview of the Hill plots for compound **8**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	22.72042	
Adj. R-Square	0.94919	
	Value	Standard Error
Vmax	100	0
EC50	2.9904	0.26116
n	1.06302	0.16343

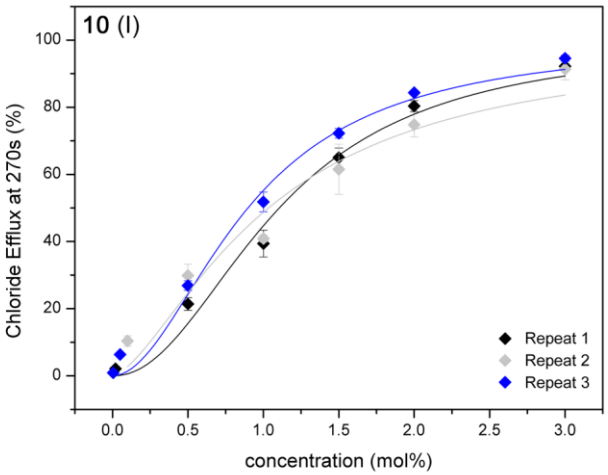
Repeat 4	Hill	
Reduced Chi-S	18.63907	
Adj. R-Square	0.96208	
	Value	Standard Error
Vmax	100	0
EC50	2.04062	0.18369
n	0.86411	0.09768

Repeat 2	Hill	
Reduced Chi-S	8.58247	
Adj. R-Square	0.9744	
	Value	Standard Error
Vmax	100	0
EC50	3.36915	0.18901
n	0.91684	0.08525

Repeat 5	Hill	
Reduced Chi-Sq	6.65262	
Adj. R-Square	0.98733	
	Value	Standard Error
Vmax	100	0
EC50	2.62692	0.13517
n	0.9287	0.07009

Repeat 3	Hill	
Reduced Chi-Sq	10.55709	
Adj. R-Square	0.97435	
	Value	Standard Error
Vmax	100	0
EC50	2.32862	0.17422
n	0.77137	0.06978

Figure S222. Overview of the Hill plots for compound **9**. For experimental details, see main text.

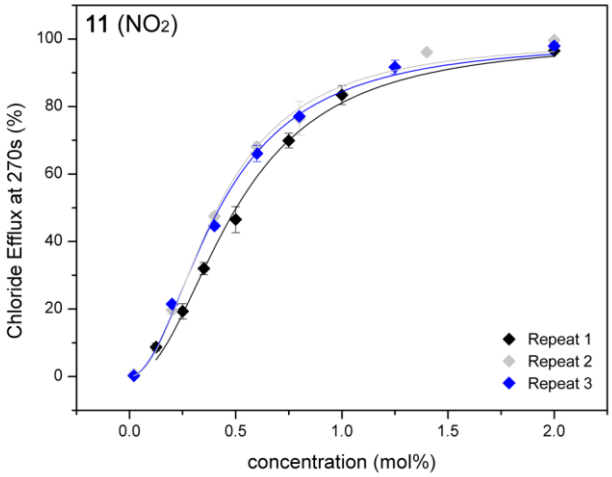


Repeat 1	Hill	
Reduced Chi-Sqr	20.10481	
Adj. R-Square	0.98366	
	Value	Standard Error
Vmax	100	0
EC50	1.1057	0.05935
n	2.12832	0.24773

Repeat 2	Hill	
Reduced Chi-S	53.37101	
Adj. R-Square	0.94074	
	Value	Standard Error
Vmax	100	0
EC50	1.03446	0.11345
n	1.53406	0.29574

Repeat 3	Hill	
Reduced Chi-Sq	13.97739	
Adj. R-Square	0.99013	
	Value	Standard Error
Vmax	100	0
EC50	0.89791	0.04621
n	1.94511	0.18896

Figure S223. Overview of the Hill plots for compound **10**. For experimental details, see main text.

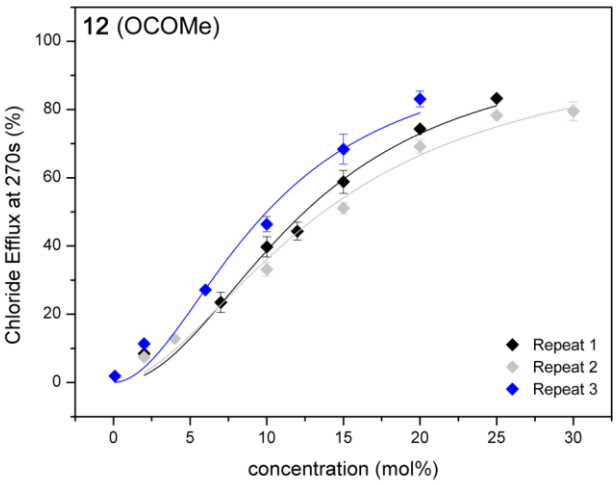


Repeat 1	Hill	
Reduced Chi-Sqr	6.51979	
Adj. R-Square	0.99412	
	Value	Standard Error
Vmax	100	0
EC50	0.50322	0.01342
n	2.11788	0.12136

Repeat 2	Hill	
Reduced Chi-S	7.26196	
Adj. R-Square	0.99487	
	Value	Standard Error
Vmax	100	0
EC50	0.4159	0.01401
n	2.06972	0.14152

Repeat 3	Hill	
Reduced Chi-Sq	5.27854	
Adj. R-Square	0.99604	
	Value	Standard Error
Vmax	100	0
EC50	0.42484	0.01243
n	1.97259	0.11269

Figure S224. Overview of the Hill plots for compound **11**. For experimental details, see main text.

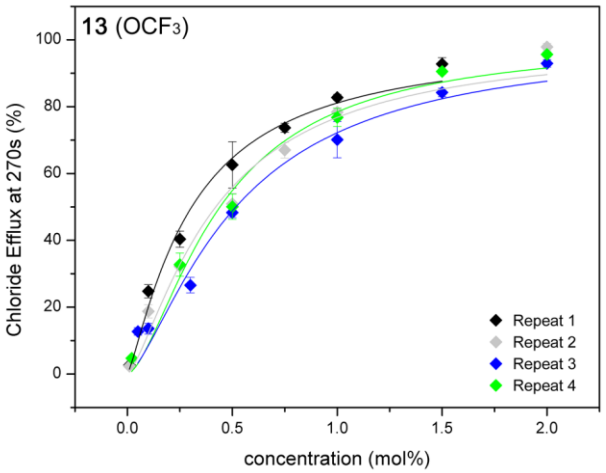


Repeat 1	Hill	
Reduced Chi-Sqr	12.36549	
Adj. R-Square	0.98273	
	Value	Standard Error
Vmax	100	0
EC50	12.47449	0.40479
n	2.09785	0.1849

Repeat 2	Hill	
Reduced Chi-S	13.59141	
Adj. R-Square	0.98507	
	Value	Standard Error
Vmax	100	0
EC50	13.69816	0.6092
n	1.82066	0.16638

Repeat 3	Hill	
Reduced Chi-Sq	20.29503	
Adj. R-Square	0.98026	
	Value	Standard Error
Vmax	100	0
EC50	10.00333	0.56826
n	1.91498	0.24279

Figure S225. Overview of the Hill plots for compound **12**. For experimental details, see main text.



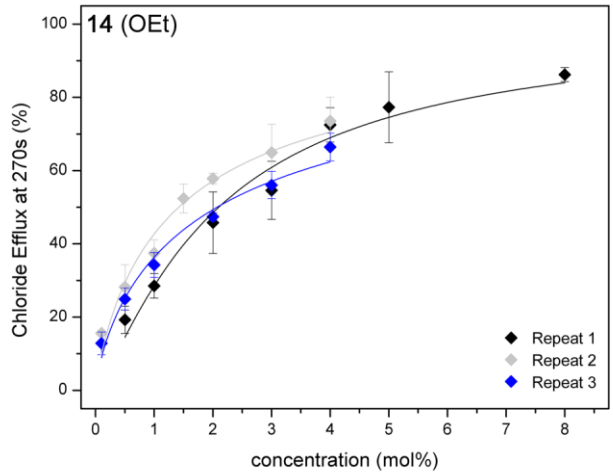
Repeat 1	Hill	
Reduced Chi-Sqr	14.0439	
Adj. R-Square	0.98697	
	Value	Standard Error
Vmax	100	0
EC50	0.30645	0.02185
n	1.23127	0.1071

Repeat 3	Hill	
Reduced Chi-Sq	34.53995	
Adj. R-Square	0.96902	
	Value	Standard Error
Vmax	100	0
EC50	0.51693	0.04919
n	1.44657	0.19047

Repeat 2	Hill	
Reduced Chi-S	29.59201	
Adj. R-Square	0.97445	
	Value	Standard Error
Vmax	100	0
EC50	0.42075	0.03768
n	1.381	0.1788

Repeat 4	Hill	
Reduced Chi-S	20.73682	
Adj. R-Square	0.98372	
	Value	Standard Error
Vmax	100	0
EC50	0.44153	0.03452
n	1.57805	0.18596

Figure S226. Overview of the Hill plots for compound **13**. For experimental details, see main text.

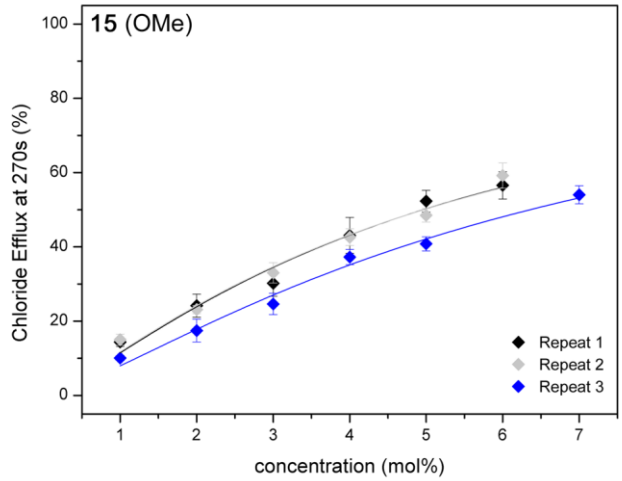


Repeat 1	Hill	
Reduced Chi-Sqr	18.93757	
Adj. R-Square	0.97041	
	Value	Standard Error
Vmax	100	0
EC50	2.09995	0.14519
n	1.23746	0.11995

Repeat 2	Hill	
Reduced Chi-S	14.97329	
Adj. R-Square	0.96548	
	Value	Standard Error
Vmax	100	0
EC50	1.4194	0.11445
n	0.84151	0.08898

Repeat 3	Hill	
Reduced Chi-Sq	10.32268	
Adj. R-Square	0.97432	
	Value	Standard Error
Vmax	100	0
EC50	2.07245	0.17044
n	0.76846	0.07534

Figure S227. Overview of the Hill plots for compound **14**. For experimental details, see main text.

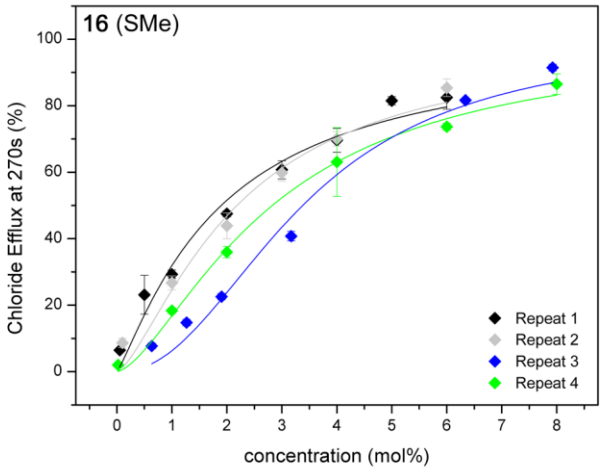


Repeat 1	Hill	
Reduced Chi-Sqr	7.92789	
Adj. R-Square	0.97123	
	Value	Standard Error
Vmax	100	0
EC50	4.96132	0.2427
n	1.27735	0.12314

Repeat 2	Hill	
Reduced Chi-S	7.12941	
Adj. R-Square	0.97359	
	Value	Standard Error
Vmax	100	0
EC50	4.9659	0.23366
n	1.25944	0.11563

Repeat 3	Hill	
Reduced Chi-Sq	4.26999	
Adj. R-Square	0.9839	
	Value	Standard Error
Vmax	100	0
EC50	6.35896	0.27667
n	1.32367	0.09711

Figure S228. Overview of the Hill plots for compound **15**. For experimental details, see main text.



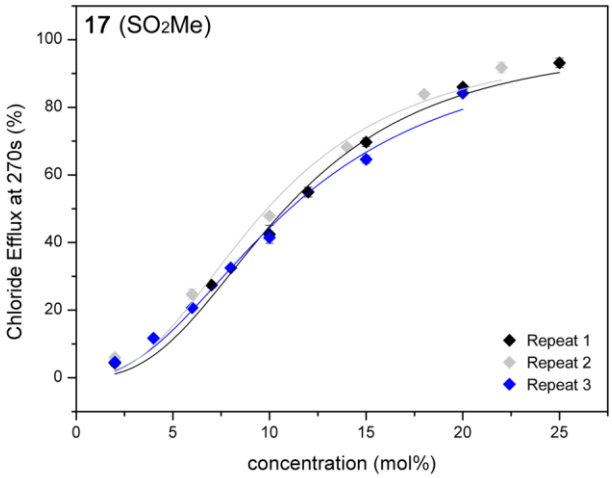
Repeat 1	Hill	
Reduced Chi-Sqr	22.16112	
Adj. R-Square	0.97229	
	Value	Standard Error
Vmax	100	0
EC50	1.89783	0.14798
n	1.18453	0.12401

Repeat 3	Hill	
Reduced Chi-Sq	26.3484	
Adj. R-Square	0.97912	
	Value	Standard Error
Vmax	100	0
EC50	3.37972	0.2105
n	2.21811	0.26035

Repeat 2	Hill	
Reduced Chi-S	22.60857	
Adj. R-Square	0.97198	
	Value	Standard Error
Vmax	100	0
EC50	2.18415	0.16097
n	1.43707	0.18989

Repeat 4	Hill	
Reduced Chi-S	6.0279	
Adj. R-Square	0.99452	
	Value	Standard Error
Vmax	100	0
EC50	2.82807	0.10864
n	1.53948	0.09244

Figure S229. Overview of the Hill plots for compound **16**. For experimental details, see main text.

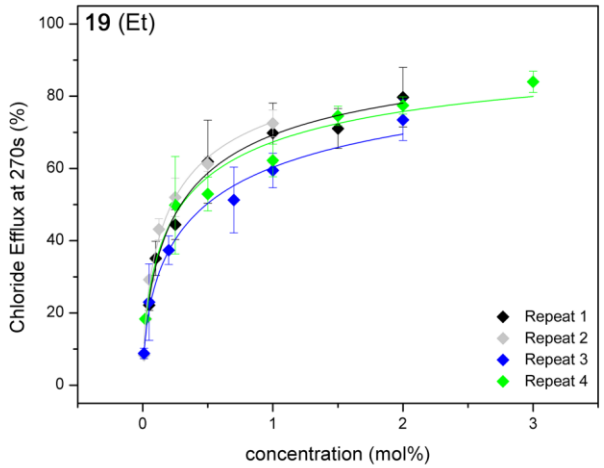


Repeat 1	Hill	
Reduced Chi-Sqr	9.21456	
Adj. R-Square	0.99095	
	Value	Standard Error
Vmax	100	0
EC50	10.81821	0.26931
n	2.66331	0.19986

Repeat 2	Hill	
Reduced Chi-S	13.66449	
Adj. R-Square	0.98805	
	Value	Standard Error
Vmax	100	0
EC50	9.88344	0.37514
n	2.5002	0.23095

Repeat 3	Hill	
Reduced Chi-Sq	10.03333	
Adj. R-Square	0.98792	
	Value	Standard Error
Vmax	100	0
EC50	11.05229	0.33377
n	2.27557	0.16617

Figure S230. Overview of the Hill plots for compound **17**. For experimental details, see main text.



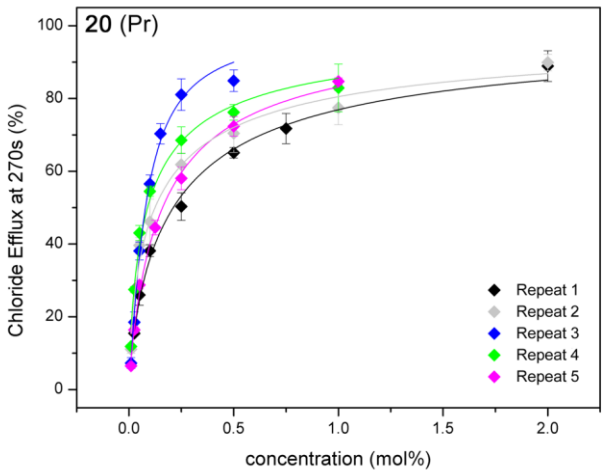
Repeat 1	Hill	
Reduced Chi-Sqr	8.26789	
Adj. R-Square	0.98177	
	Value	Standard Error
Vmax	100	0
EC50	0.29424	0.02337
n	0.66329	0.04404

Repeat 3	Hill	
Reduced Chi-Sq	9.18821	
Adj. R-Square	0.98396	
	Value	Standard Error
Vmax	100	0
EC50	0.48961	0.05053
n	0.58521	0.04685

Repeat 2	Hill	
Reduced Chi-S	4.95796	
Adj. R-Square	0.99084	
	Value	Standard Error
Vmax	100	0
EC50	0.22205	0.01429
n	0.65626	0.03387

Repeat 4	Hill	
Reduced Chi-S	16.35144	
Adj. R-Square	0.96712	
	Value	Standard Error
Vmax	100	0
EC50	0.29986	0.04403
n	0.60002	0.06172

Figure S231. Overview of the Hill plots for compound **19**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	6.47025	
Adj. R-Square	0.99058	
	Value	Standard Error
Vmax	100	0
EC50	0.21135	0.01329
n	0.77386	0.04194

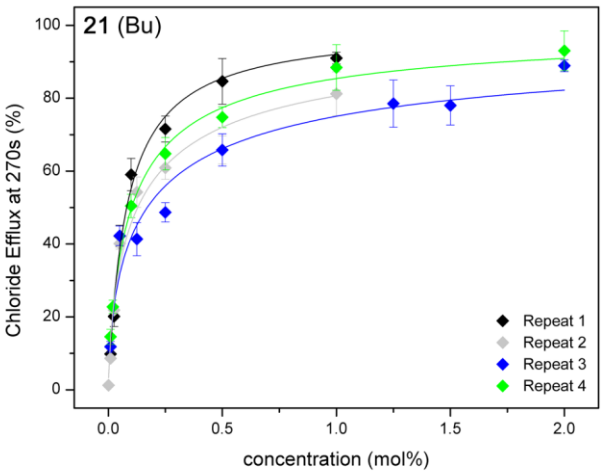
Repeat 4	Hill	
Reduced Chi-S	10.48407	
Adj. R-Square	0.98471	
	Value	Standard Error
Vmax	100	0
EC50	0.08888	0.00752
n	0.73528	0.05091

Repeat 2	Hill	
Reduced Chi-S	11.97757	
Adj. R-Square	0.98301	
	Value	Standard Error
Vmax	100	0
EC50	0.12398	0.01263
n	0.67858	0.05201

Repeat 5	Hill	
Reduced Chi-Sq	2.54307	
Adj. R-Square	0.99698	
	Value	Standard Error
Vmax	100	0
EC50	0.1604	0.00598
n	0.87195	0.02951

Repeat 3	Hill	
Reduced Chi-Sq	7.59464	
Adj. R-Square	0.99185	
	Value	Standard Error
Vmax	100	0
EC50	0.07874	0.00387
n	1.18561	0.07068

Figure S232. Overview of the Hill plots for compound **20**. For experimental details, see main text.



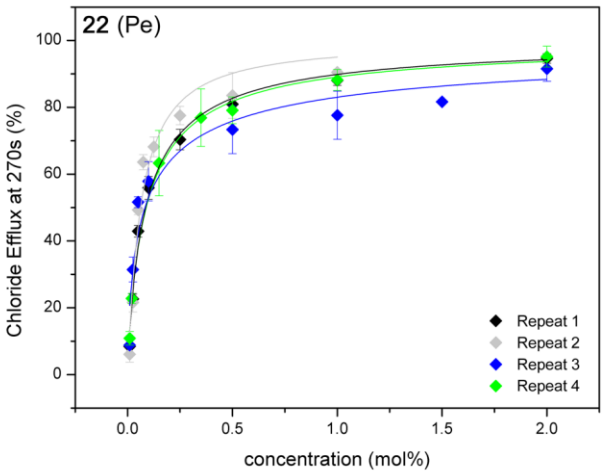
Repeat 1	Hill	
Reduced Chi-Sqr	12.35476	
Adj. R-Square	0.98739	
	Value	Standard Error
Vmax	100	0
EC50	0.0806	0.00626
n	0.96986	0.07188

Repeat 3	Hill	
Reduced Chi-Sq	38.62325	
Adj. R-Square	0.94091	
	Value	Standard Error
Vmax	100	0
EC50	0.17043	0.03173
n	0.62182	0.08066

Repeat 2	Hill	
Reduced Chi-S	23.45328	
Adj. R-Square	0.97167	
	Value	Standard Error
Vmax	100	0
EC50	0.12852	0.01678
n	0.69564	0.07127

Repeat 4	Hill	
Reduced Chi-S	5.04258	
Adj. R-Square	0.99466	
	Value	Standard Error
Vmax	100	0
EC50	0.10182	0.00662
n	0.77304	0.03558

Figure S233. Overview of the Hill plots for compound **21**. For experimental details, see main text.



Repeat 1	Hill	
Reduced Chi-Sqr	10.5675	
Adj. R-Square	0.98922	
	Value	Standard Error
Vmax	100	0
EC50	0.08574	0.00645
n	0.88741	0.05765

Repeat 3	Hill	
Reduced Chi-Sq	53.82184	
Adj. R-Square	0.93097	
	Value	Standard Error
Vmax	100	0
EC50	0.08252	0.01834
n	0.63603	0.09092

Repeat 2	Hill	
Reduced Chi-S	42.6502	
Adj. R-Square	0.95275	
	Value	Standard Error
Vmax	100	0
EC50	0.06173	0.0076
n	1.05882	0.14769

Repeat 4	Hill	
Reduced Chi-S	4.16898	
Adj. R-Square	0.99612	
	Value	Standard Error
Vmax	100	0
EC50	0.08801	0.00537
n	0.86343	0.03574

Figure S234. Overview of the Hill plots for compound **22**. For experimental details, see main text.

S6.5 Overview of anion transport results

Table S21 gives an overview of all of the obtained EC_{50}, n and k_{ini} values for each repeat, while Table S22 gives an overview of the average values and standard deviations obtained from the data in Table S21.

Table S21a. Overview of EC_{50}, n and k_{ini} values for each repeat. Part 1.

compound	repeat	EC_{50} (mol%)	$\log(1/EC_{50})$	n	k_{ini} (%s^{-1})
1 (Br)	repeat 1	0.758	0.1203	1.38	0.655
	repeat 2	1.108	-0.0445	1.49	0.562
	repeat 3	0.966	0.0150	1.80	0.644
	repeat 4	0.729	0.1373	1.45	0.837
	repeat 5	1.093	-0.0386	1.74	0.613
2 (CF_3)	repeat 1	0.441	0.3556	1.58	2.039
	repeat 2	0.399	0.3990	1.64	2.746
	repeat 3	0.483	0.3161	1.73	2.219
3 (Cl)	repeat 1	0.844	0.0737	1.48	0.790
	repeat 2	1.001	-0.0004	1.30	0.417
	repeat 3	0.893	0.0491	1.56	0.708
	repeat 4	1.003	-0.0013	1.75	0.560
	repeat 5	1.371	-0.1370	1.75	0.524
4 (CN)	repeat 1	0.814	0.0894	2.03	1.387
	repeat 2	0.798	0.0980	2.07	1.321
	repeat 3	1.038	-0.0162	2.02	0.813
	repeat 4	0.629	0.2013	2.30	2.024
5 ($COCF_3$)	repeat 1	1.043	-0.0183	1.85	2.733
	repeat 2	0.883	0.0540	2.08	4.054
	repeat 3	0.849	0.0711	1.86	4.680
6 (COMe)	repeat 1	3.176	-0.5019	2.13	0.131
	repeat 2	3.034	-0.4820	1.82	0.126
	repeat 3	3.228	-0.5089	2.15	0.129
7 (COOMe)	repeat 1	2.150	-0.3324	1.74	0.195
	repeat 2	2.048	-0.3113	1.66	0.254
	repeat 3	2.979	-0.4741	1.59	0.145
	repeat 4	2.178	-0.3380	2.03	0.183
8 (F)	repeat 1	1.590	-0.2014	1.64	0.310
	repeat 2	2.003	-0.3017	1.48	0.228
	repeat 3	1.801	-0.2555	1.60	0.244
	repeat 4	2.100	-0.3222	1.86	0.213
	repeat 5	2.619	-0.4181	1.76	0.110
9 (H)	repeat 1	2.990	-0.4757	1.06	0.155
	repeat 2	3.369	-0.5275	0.92	0.134
	repeat 3	2.329	-0.3672	0.77	0.197
	repeat 4	2.041	-0.3098	0.86	0.202
	repeat 5	2.627	-0.4195	0.93	0.178
10 (I)	repeat 1	1.106	-0.0438	2.13	0.678
	repeat 2	1.034	-0.0145	1.53	0.531
	repeat 3	0.898	0.0467	1.95	0.850

Table S21b. Overview of EC_{50}, n and k_{ini} values for each repeat. Part 2.

compound	repeat	EC_{50} (mol%)	$\log(1/EC_{50})$	n	k_{ini} (%s^{-1})
11 (NO_2)	repeat 1	0.503	0.2984	2.12	2.072
	repeat 2	0.416	0.3809	2.07	3.080
	repeat 3	0.425	0.3716	1.97	3.177
12 (OCOMe)	repeat 1	12.474	-1.0960	2.10	0.022
	repeat 2	13.698	-1.1367	1.82	0.029
	repeat 3	10.003	-1.0001	1.91	0.023
13 (OCF_3)	repeat 1	0.306	0.5143	1.23	2.326
	repeat 2	0.421	0.3757	1.38	1.890
	repeat 3	0.517	0.2865	1.45	1.281
	repeat 4	0.442	0.3546	1.58	1.736
14 (OEt)	repeat 1	2.100	-0.3222	1.24	0.177
	repeat 2	1.419	-0.1520	0.84	0.347
	repeat 3	2.072	-0.3164	0.77	0.198
15 (OMe)	repeat 1	4.961	-0.6956	1.28	0.076
	repeat 2	4.966	-0.6960	1.26	0.083
	repeat 3	6.359	-0.8034	1.32	0.070
16 (SMe)	repeat 1	1.898	-0.2783	1.18	0.204
	repeat 2	2.184	-0.3393	1.44	0.183
	repeat 3	3.380	-0.5289	2.22	-
	repeat 4	2.828	-0.4515	1.54	0.127
17 (SO_2Me)	repeat 1	10.818	-1.0341	2.66	0.034
	repeat 2	9.883	-0.9949	2.50	0.048
	repeat 3	11.052	-1.0434	2.28	0.040
18 (Me)	repeat 1	0.848	0.0716	0.83	0.460
	repeat 2	1.951	-0.2903	0.72	0.190
	repeat 3	0.703	0.1530	0.66	0.469
	repeat 4	1.657	-0.2192	1.02	0.233
	repeat 5	1.975	-0.2956	1.22	0.187
	repeat 6	1.097	-0.0402	0.90	0.297
	repeat 7	1.052	-0.0220	0.78	0.373
19 (Et)	repeat 1	0.294	0.5317	0.66	0.638
	repeat 2	0.222	0.6536	0.66	-
	repeat 3	0.490	0.3098	0.58	0.416
	repeat 4	0.300	0.5229	0.60	0.604
20 (Pr)	repeat 1	0.212	0.6737	0.77	1.644
	repeat 2	0.124	0.9066	0.68	1.064
	repeat 3	0.079	1.1024	1.18	1.721
	repeat 4	0.089	1.0506	0.74	1.759
	repeat 5	0.160	0.7959	0.87	-
21 (Bu)	repeat 1	0.081	1.0915	0.97	1.224
	repeat 2	0.128	0.8928	0.70	-
	repeat 3	0.170	0.7696	0.62	0.959
	repeat 4	0.102	0.9914	0.77	1.102
22 (Pe)	repeat 1	0.086	1.0655	0.89	1.284
	repeat 2	0.062	1.2076	1.06	-
	repeat 3	0.082	1.0862	0.64	1.275
	repeat 4	0.088	1.0555	0.86	1.602

Table S22. Overview of the average EC_{50}, n and k_{ini} values and the respective standard deviation (error).

Compound	EC_{50} (mol%)	Error EC_{50}	n	Error n	$log(1/EC_{50})$	Error $log(1/EC_{50})$	k_{ini} (%s^{-1})	Error k_{ini}
1 (Br)	0.93	0.18	1.57	0.19	0.031	0.093	0.66	0.10
2 (CF_3)	0.44	0.04	1.65	0.08	0.356	0.043	2.33	0.37
3 (Cl)	1.02	0.21	1.57	0.19	-0.011	0.100	0.60	0.15
4 (CN)	0.82	0.17	2.11	0.13	0.085	0.101	1.39	0.50
5 ($COCF_3$)	0.93	0.10	1.93	0.13	0.034	0.052	3.82	0.99
6 (COMe)	3.15	0.10	2.03	0.19	-0.4832	-0.0004	0.129	0.003
7 (COOMe)	2.34	0.43	1.76	0.19	-0.372	0.091	0.19	0.05
8 (F)	2.02	0.39	1.67	0.15	-0.304	0.090	0.22	0.07
9 (H)	2.67	0.53	0.91	0.11	-0.427	0.095	0.17	0.03
10 (I)	1.01	0.11	1.87	0.31	-0.005	0.048	0.69	0.16
11 (NO_2)	0.45	0.05	2.05	0.08	0.349	0.049	2.78	0.61
12 (OCOMe)	12.06	1.88	1.94	0.14	-1.080	0.073	0.025	0.004
13 (OCF_3)	0.42	0.09	1.41	0.15	0.375	0.101	1.81	0.43
14 (OEt)	1.86	0.39	0.95	0.25	-0.278	0.109	0.24	0.09
15 (OMe)	5.43	0.81	1.29	0.03	-0.738	0.073	0.076	0.007
16 (SMe)	2.57	0.66	1.60	0.44	-0.410	0.130	0.17	0.04
17 (SO_2Me)	10.58	0.62	2.48	0.19	-1.025	0.026	0.041	0.007
18 (Me)	1.33	0.53	0.88	0.19	-0.123	0.220	0.32	0.12
19 (Et)	0.33	0.11	0.63	0.04	0.482	0.192	0.55	0.12
20 (Pr)	0.13	0.05	0.85	0.20	0.877	0.230	1.55	0.33
21 (Bu)	0.12	0.04	0.77	0.15	0.917	0.169	1.10	0.13
22 (Pe)	0.08	0.01	0.86	0.17	1.100	0.070	1.39	0.19

S7. QSAR ANALYSIS OF ANION TRANSPORT

S7.1 Correlation EC_{50}, n and k_{ini}

Using JMP 9.0.0 we tried to find a correlation between log(1/EC_{50}), *n* and k_{ini}, by using any of the descriptors given in Table S23 (only data of the training set was used). A stepwise multiple linear regression was first executed, where the k-fold cross validation was set to 2 and all possible models with a maximum of three terms were calculated and subsequently ranked according to best fit (highest R). The results are given in Figure S235.

Table S23. Overview of combination between EC_{50}, *n* and k_{ini} for the training set, used as descriptors to find a correlation between them.

Compound	EC_{50} [a]	log(1/EC_{50})	n [b]	n-sq	k_{ini} [c]	k_{ini}/n	n*log(1/k_{ini})	log(1/k_{ini})/n	log(1/k_{ini})
21 (Bu)	0.12	0.917	0.77	0.59	1.10	1.431	-0.030	-0.052	-0.039
2 (CF_3)	0.44	0.356	1.65	2.72	2.33	1.415	-0.608	-0.223	-0.368
3 (Cl)	1.02	-0.011	1.57	2.46	0.60	0.383	0.348	0.142	0.222
4 (CN)	0.82	0.085	2.11	4.43	1.39	0.659	-0.299	-0.067	-0.142
5 ($COCF_3$)	0.93	0.034	1.93	3.72	3.82	1.980	-1.124	-0.302	-0.582
7 (COOMe)	2.34	-0.372	1.76	3.08	0.19	0.111	1.249	0.405	0.712
19 (Et)	0.33	0.482	0.63	0.39	0.55	0.884	0.161	0.412	0.258
8 (F)	2.02	-0.304	1.67	2.78	0.22	0.132	1.094	0.393	0.656
9 (H)	2.67	-0.427	0.91	0.82	0.17	0.191	0.691	0.839	0.761
10 (I)	1.01	-0.005	1.87	3.50	0.69	0.367	0.306	0.087	0.163
18 (Me)	1.33	-0.123	0.88	0.77	0.32	0.360	0.439	0.572	0.501
11(NO_2)	0.45	0.349	2.05	4.22	2.78	1.352	-0.911	-0.216	-0.443
12 (OCOMe)	12.06	-1.080	1.94	3.78	0.025	0.013	3.076	0.847	1.608
13 (OCF_3)	0.42	0.375	1.41	1.99	1.81	1.282	-0.363	-0.182	-0.257
15 (OMe)	5.43	-0.738	1.29	1.66	0.076	0.059	1.438	0.868	1.117
22 (Pentyl)	0.08	1.100	0.86	0.74	1.39	1.608	-0.123	-0.165	-0.142
16 (SMe)	2.57	-0.410	1.60	2.54	0.17	0.107	1.222	0.480	0.766
17 (SO_2Me)	10.58	-1.025	2.48	6.15	0.041	0.016	3.449	0.561	1.391

[a] concentration (in mol% transporter to lipid) needed to obtain 50% chloride efflux in 270 s. Obtained through Hill plot (see Section S6.4.). [b] Hill coefficient. Obtained through Hill plot (see Section S6.4.). [c] initial rate constant (in % per second) of chloride efflux when 2 mol% receptor was added to the vesicles.

All Possible Models

Ordered up to best 10 models up to 3 terms per

Model	Number	RSquare	RMSE	AICc	BIC
log(1/k(ini))	1	0,6922	0,3398	17,8212	18,7780
k(ini)/n	1	0,6449	0,3650	20,3975	21,3544
n*log(1/kini)	1	0,6409	0,3671	20,5960	21,5529
log(1/k(ini))/n	1	0,5846	0,3948	23,2203	24,1772
k(ini)	1	0,2837	0,5184	33,0254	33,9823
n	1	0,2525	0,5296	33,7932	34,7501
n-sq	1	0,2415	0,5335	34,0567	35,0135
n,log(1/k(ini))/n	2	0,9507	0,1404	-11,797	-11,313
n-sq,log(1/k(ini))/n	2	0,9195	0,1795	-2,9563	-2,4718
n,log(1/k(ini))	2	0,8580	0,2384	7,2599	7,7445
n-sq,log(1/k(ini))	2	0,8336	0,2580	10,1094	10,5939
k(ini),log(1/k(ini))	2	0,7992	0,2835	13,5000	13,9846
k(ini),k(ini)/n	2	0,7968	0,2852	13,7113	14,1958
n,k(ini)/n	2	0,7724	0,3018	15,7540	16,2386
n-sq,k(ini)/n	2	0,7699	0,3035	15,9476	16,4322
n,n*log(1/kini)	2	0,7270	0,3305	19,0233	19,5079
k(ini)/n,log(1/k(ini))	2	0,7181	0,3359	19,6016	20,0862
k(ini),k(ini)/n,log(1/k(ini))	3	0,9790	0,0948	-23,241	-23,789
n,k(ini),log(1/k(ini))/n	3	0,9648	0,1229	-13,914	-14,462
n,n-sq,log(1/k(ini))/n	3	0,9606	0,1301	-11,876	-12,424
n,k(ini)/n,log(1/k(ini))/n	3	0,9558	0,1376	-9,8440	-10,392
n,n*log(1/kini),log(1/k(ini))/n	3	0,9548	0,1392	-9,4396	-9,9878
n,log(1/k(ini))/n,log(1/k(ini))	3	0,9547	0,1394	-9,3849	-9,9331
k(ini),k(ini)/n,n*log(1/kini)	3	0,9395	0,1611	-4,1702	-4,7184
n,n*log(1/kini),log(1/k(ini))	3	0,9377	0,1634	-3,6603	-4,2084
n-sq,k(ini),log(1/k(ini))/n	3	0,9355	0,1663	-3,0340	-3,5822
n-sq,n*log(1/kini),log(1/k(ini))	3	0,9302	0,1730	-1,6120	-2,1602

Figure S235. Overview of best 10 models with up to 3 term correlating $\log(1/EC_{50})$ with n and k_{ini}.

It seems that $\log(1/EC_{50})$ can be described best by a model with two parameters, namely n and $\log(1/k_{ini})/n$. A multiple linear regression (standard least-squares method) was then run using JMP 9.0.0 where $\log(1/EC_{50})$ was modeled as a function of n and $\log(1/k_{ini})/n$. The results of this regression are in Figure S236.

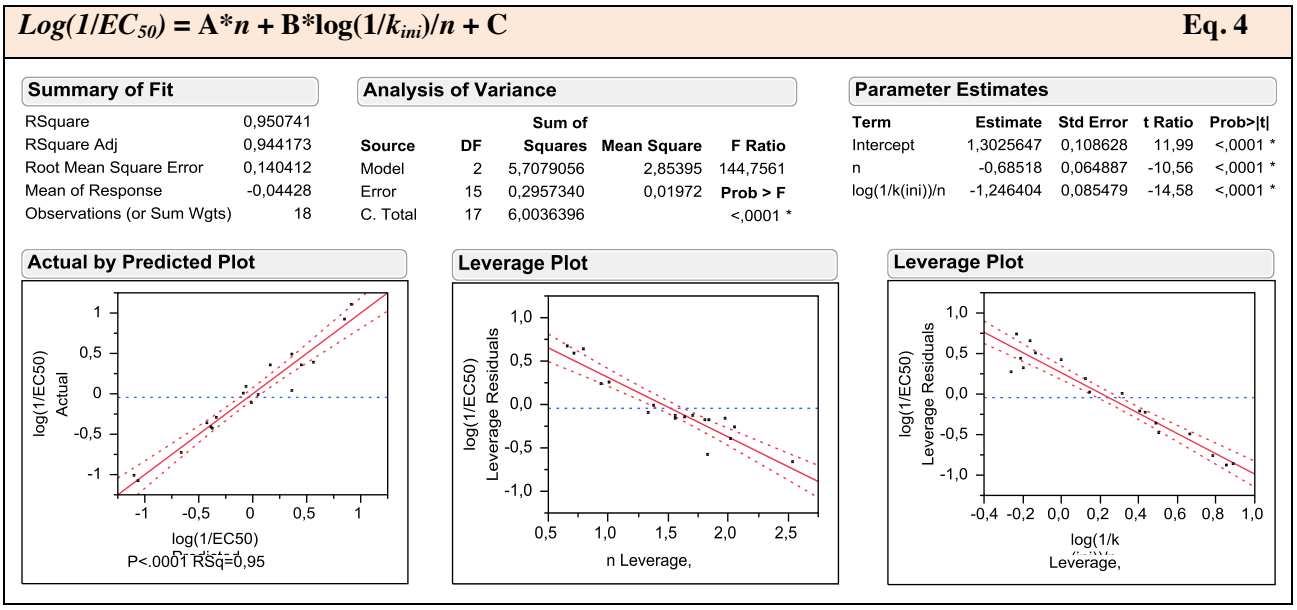


Figure S236. Overview of best model correlating $\log(1/EC_{50})$ with n and k_{ini}.

The best transporter is the receptor with the highest value of $\log(1/EC_{50})$. From the obtained equation, $log(1/EC_{50}) = -0.68(\pm0.06)*n\ -1.25(\pm0.08)*log(1/k_{ini})/n + 1.3(\pm0.1)$, it is clear that a high *n* value will decrease the transport ability (which is logical because a high *n* value will make the transporter more concentration depended and so the transport activity will decrease rapidly with decreasing receptor concentration). Also, very negative values of $\log(1/k_{ini})$ will result in better transport activity (which is expected because negative $log(1/k_{ini})$ means fast chloride release and hence good transport), but this effect can be counter-balanced by a high *n* value (therefore $\log(1/k_{ini})/n$).

S7.2 Correlation retention time and logP – choosing correct logP model

In order to assess the relative lipophilicity of the receptors studied, we utilised a previously reported HPLC experiment using a reverse phase column, as the retention time on the reverse phase column is related to its lipophilicity.[24] The HPLC mobile phase was prepared with LC-MS grade methanol and water (Fisher Scientific UK, Loughborough, UK) containing 0.1% HCOOH each. Samples were prepared as a solution in LC-MS grade methanol at a concentration of 50 µg/mL. HPLC separations were performed on a HP1050 Series system (Agilent Technologies, Palo Alto, CA, USA). Samples were injected (10 µL) directly onto a Xbridge C18 Column (50 mm X 2.1 mm 5 µm particle size; Waters, Milford, MA, USA) thermostatically controlled at 40°C. The separation was achieved using 20% methanol in water for 2 minutes followed by a linear gradient to 100% methanol over 10 minutes and held at 100% methanol for 4 minutes at a flow rate of 0.3 mL/min. UV data were recorded at 254 nm and mass spectra were recorded using a Platform LC single quadrupole mass spectrometer (Waters, Milford, MA, USA) using positive ion electrospray ionisation (120-1000 m/z) in order to assign to retention time to the respective receptor.

logP values were calculated with a variety of programs, including VCCLAB, Fieldview 2.0.2, Daylight version 4.73, ACD/I-Labs 2.0 and Accelrys Diamond Descriptors.[11,25-28] The average logP (average of all the calculated logP values) is also included. Tables S24 and S25 give an overview of the obtained retention times and calculated logP values. They also show the R values obtained by linear fitting of the (RT, logP)-plots with the function $RT = A*logP + B$ using Microsoft excel 2007. From these R values it can be seen that the logP values calculated using Daylight version 4.73 method are the most suitable for this series of compounds.

Table S24. Overview of the reversed-phase HPLC retention times (minutes) and the various logP values calculated using VCCLAB ALOGPS 2.1 (reference 25).

Compound	Retention time [a]	AlogPs	AC logP	AlogP	MlogP	KOWWIN	XlogP2	XlogP3	average
21 (Bu)	13.88	5.19	5.57	5.94	4.11	5.89	5.69	5.74	5.45
2 (CF_3)	13.28	4.56	4.73	5.03	4.01	4.84	4.58	4.75	4.64
3 (Cl)	12.79	4.09	4.58	4.75	3.62	4.52	4.28	4.49	4.33
4 (CN)	12.23	3.31	3.78	3.96	2.7	3.97	3.38	3.58	3.53
5 ($COCF_3$)	12.47	4.48	4.37	4.91	3.33	3.97	4.19	4.66	4.27
7 (COOMe)	12.44	3.34	3.94	3.94	2.96	3.71	3.59	3.71	3.60
19 (Et)	13.00	4.36	4.64	5.03	3.61	4.91	4.56	4.66	4.54
8 (F)	12.08	3.58	4.02	4.29	3.49	4.07	3.82	3.96	3.89
9 (H)	11.91	3.43	3.96	4.09	3.09	3.87	3.66	3.39	3.64
10 (I)	13.18	4.75	4.9	4.66	3.88	5.04	4.72	4.51	4.64
18 (Me)	12.51	3.69	4.28	4.57	3.36	4.42	4.09	4.23	4.09
11(NO_2)	12.75	3.61	3.83	3.98	3.03	3.69	3.55	3.69	3.63
12 (OCOMe)	11.77	3.74	3.9	3.85	2.96	3.47	3.44	3.61	3.57
13 (OCF_3)	13.24	4.73	4.79	6.21	3.17	4.92	4.82	5.04	4.81
15 (OMe)	11.95	3.52	3.86	4.07	2.79	3.95	3.57	3.83	3.66
22 (Pentyl)	14.23	5.71	6.03	6.4	4.34	6.38	6.26	6.28	5.95
16 (SMe)	12.61	4.25	4.46	4.63	3.62	4.47	4.4	4.37	4.32
17 (SO_2Me)	11.42	3.35	3.03	3.61	2.62	2.44	3.18	3.10	3.05
R (RT vs logP)[b]	-	0.80	0.89	0.78	0.72	0.87	0.87	0.86	0.89

[a] Retention Time (in minutes) on reverse-phase C18 HPLC column.

Table S25. Overview of the reversed-phase HPLC retention times (minutes) and the various logP values calculated using Daylight v4.73, ACD/I-Labs 2.0, Accelrys Diamond Descriptor and Fieldview 2.0.2. (references 26-28)

Compound	Retention time [a]	Daylight ClogP [b]	ACD/labs AlogP [c]	ACD/labs ABlogP [c]	Accelrys AlogP [d]	Fieldview wclogP [e]	average of all [f]
21 (Bu)	13.88	5.612	5.49	5.52	5.928	4.9	5.506
2 (CF_3)	13.28	4.938	4.04	5.18	5.154	4.57	4.722
3 (Cl)	12.79	4.541	4.22	4.71	4.79	4.23	4.426
4 (CN)	12.23	3.661	3.28	3.91	4.137	3.42	3.620
5 ($COCF_3$)	12.47	4.075	3.75	4.68	4.691	4.3	4.326
7 (COOMe)	12.44	4.046	4.14	3.72	4.002	3.34	3.732
19 (Et)	13.00	4.554	4.47	4.93	5.135	4.12	4.620
8 (F)	12.08	3.971	3.68	4.09	4.411	3.96	3.970
9 (H)	11.91	3.526	3.43	3.96	4.272	3.55	3.719
10 (I)	13.18	4.951	4.71	4.92	5.529	4.43	4.769
18 (Me)	12.51	4.025	3.96	4.43	4.739	3.86	4.176
11(NO_2)	12.75	3.917	3.79	3.92	4.225	3.36	3.740
12 (OCOMe)	11.77	3.059	2.84	4.00	3.77	3.48	3.561
13 (OCF_3)	13.24	4.738	4.40	4.94	5.764	4.45	4.887
15 (OMe)	11.95	3.629	3.36	4.06	4.019	3.56	3.709
22 (Pentyl)	14.23	6.141	6.00	6.25	6.324	5.29	5.990
16 (SMe)	12.61	4.085	4.29	4.88	4.362	4.28	4.375
17 (SO_2Me)	11.42	2.641	2.52	3.17	3.417	2.96	3.009
R (RT vs logP)	-	0.95	0.90	0.84	0.88	0.78	0.91

[a] Retention Time (in minutes) on reverse-phase C18 HPLC column.

S7.3 QSAR modelling using retention time or logP only

JMP 9.0.0 was used to perform a standard least-squares fitting of the log(1/EC_{50}) values as a function of the experimental retention times (RT). The results are given in Figure S237 and show a good correlation between log(1/EC_{50}) and retention times, indicating the importance of lipophilicity in anion transport. Similarly, a standard least-square fitting was performed for log(1/EC_{50}) versus the calculated logP values (calculated using method Daylight v4.73, see Section S7.2.) using JMP 9.0.0, showing the same results as for the retention time, although the correlation is less good. The results are given in Figure S238.

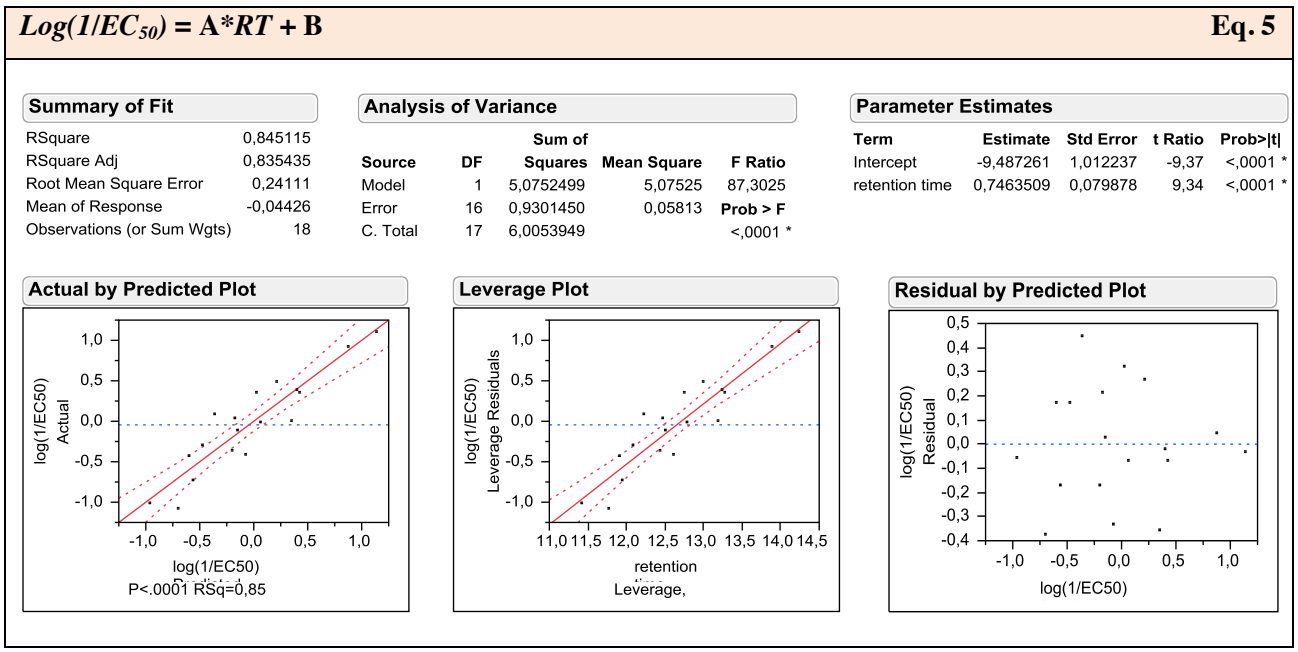


Figure S237. Overview of model correlating log(1/EC_{50}) with retention time (RT) in a linear fashion.

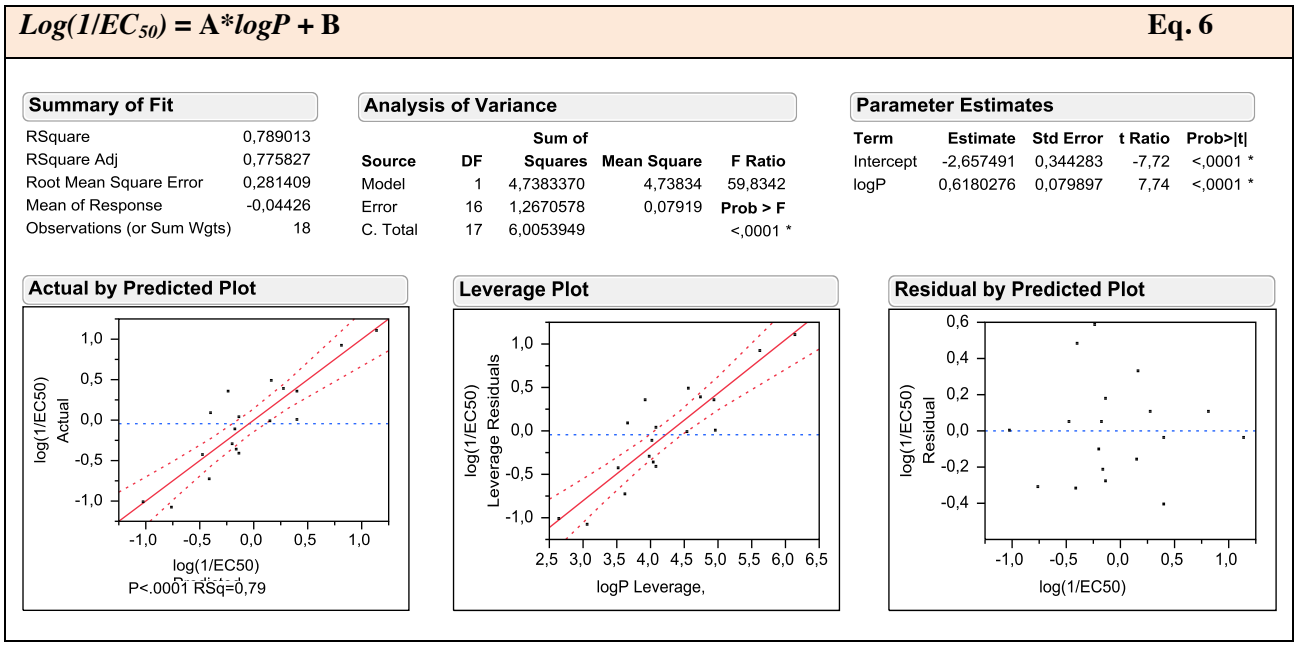


Figure S238. Overview of models correlating log(1/EC_{50}) with calculated logP in a linear fashion.

Often, a parabolic function is observed between anion transport ability and lipophilicity, where the anion transport initially increases with increasing lipophilicity, but then reaches an optimum lipophilicity. Increasing the lipophilicity even higher results in decreased anion transport ability, possibly due to low solubility or due to the fact that highly lipophilic compounds cannot move to the interphase anymore to pick up a suitable ion. In order to test whether that is the case for this set of compounds a squared term was included in the modelling to obtain a parabolic function (RT-sq is the squared form of the retention time, and logP-sq is the squared form of the calculated logP values). The results are shown in Figure S239 and indicated that the addition of the squared forms, does not significantly improve the model. Furthermore, t-values and prob|t| values show that the squared terms are not statistically significant. It therefore seems that the optimum logP value is not yet reached in this series of compounds.

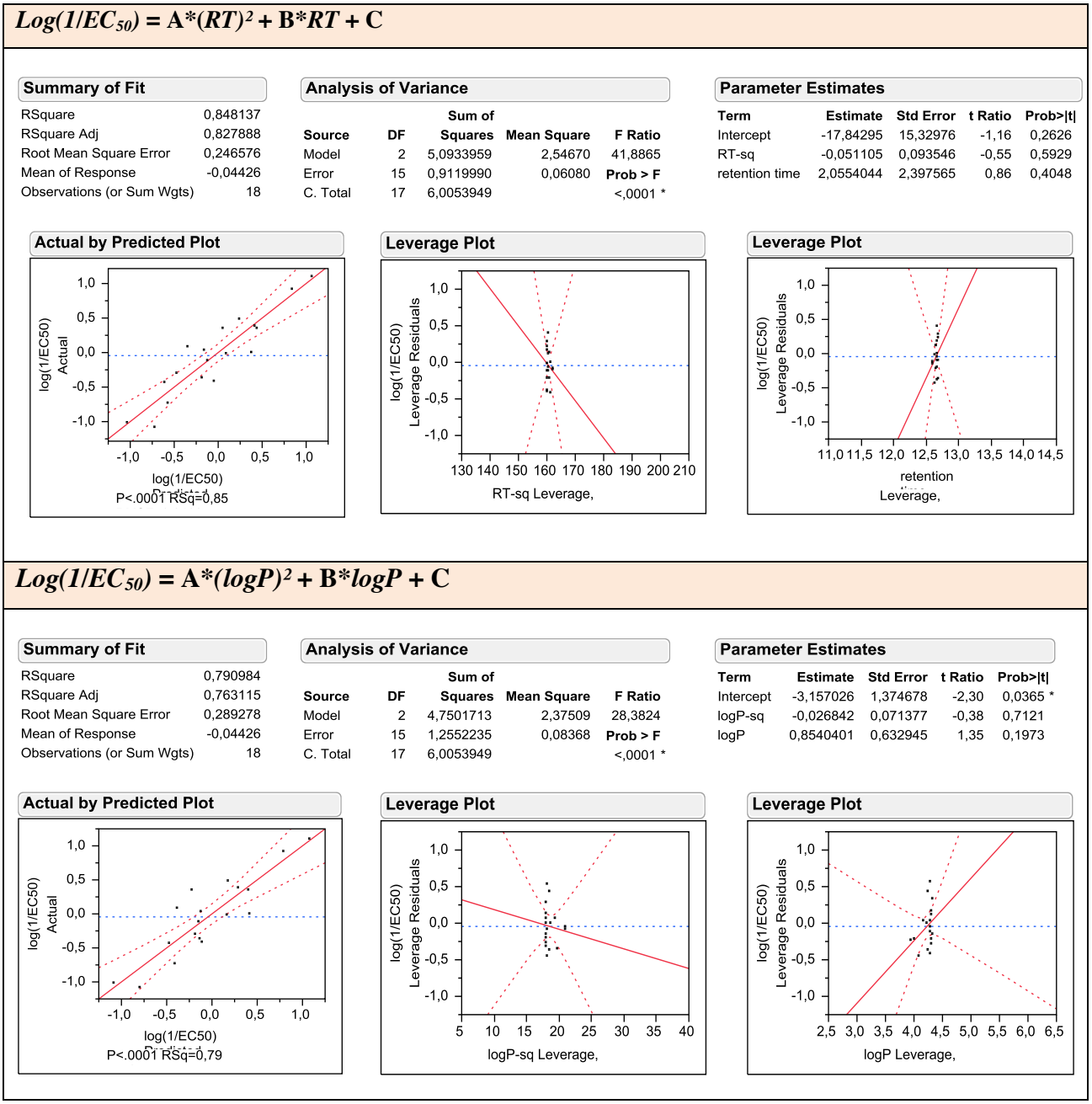


Figure S239. Overview of models correlating log(1/EC_{50}) with retention time (RT) or calculated logP in a parabolic fashion.

S7.4 List of descriptors

A total of 286 molecular descriptors were calculating using various programs. The obtained values for the training set are given in Tables S26-S48. Hammett constants came from ref 10, all others from:

- *Chemdraw 12.0*: molecular weight "Mr(g/mol)".
- *Daylight v4.73* (ref 26): logP (see Section S7.2.)
- *ChemAxon* (ref 29): PSA (polar surface area), solvent accesible suface area, (molecular) volume.
- *DFT calculations* (see Section S5.2): Surface Area, Molecular Volume, $V_{S,max}$, $V_{S,min}$, PI (average absolute deviation values).
- *ACD/I-Labs 2.0* (ref 11): Molar refractivity, Molecular volume, Parachor Index of refraction, TPSA, Bp, logD (pH 1.7), logD (pH 4.6), logD (pH 6.5), logD (pH 7.4), logD (pH 8), pKa (arom NH), pKa (alkyl NH), pKa2 (NH), logS (pH 1.7), logS (pH 4.6), logS (pH 6.5), logS (pH 7.4), logS (pH 8), Vx, logPS, Vd, polarizability.
- *e-Dragon 1.0 (VCCLAB)* (ref 25): all other descriptors were calculated using e-Dragon 1.0. Only the constitutional descriptors (48 descriptors), topological descriptors (119), topological charge indices (21), geometrical descriptors (74), WHIM descriptors (99) and molecular properties (30) were calculated. Descriptors that had an identical value for all of the compounds were deleted (or only 1 or 2 values were different), any logP values were deleted and also atom, functional group and path counts. For the meaning of descriptors: http://www.vcclab.org/lab/indexhlp/dragon_descr.html.

Table S26. Overview of the descriptors calculated using ChemDraw, Daylight, ChemAxon and DFT calculations.

Compound	Mr (g/mol)	Ham-mett	logP	PSA	solv acc suface area	volume	Surface Area	Molecular Vol.	Vsmax	Vsmin	PI
21 (Bu)	292.48	-0.16	5.612	24.06	515.03	303.83	400.10	431.64	58.44	-30.30	9.61
2 (CF_3)	304.37	0.54	4.938	24.06	441.55	267.89	349.87	374.15	66.58	-23.26	12.62
3 (Cl)	270.82	0.23	4.541	24.06	407.83	249.89	335.20	358.80	63.234	-25.91	11.87
4 (CN)	261.39	0.66	3.661	47.85	408.4	252.47	339.86	361.01	68.11	-42.11	13.30
5 ($COCF_3$)	332.38	0.80	4.075	41.13	467.58	287.1	370.14	398.95	70.35	-36.93	13.11
7 (COOMe)	294.41	0.45	4.046	50.36	464.91	281.29	364.46	394.31	64.20	-39.16	12.39
19 (Et)	264.43	-0.15	4.554	24.06	454.11	269.76	357.49	384.23	58.84	-29.87	10.43
8 (F)	254.37	0.06	3.971	24.06	398.43	240.88	321.18	341.41	62.82	-26.55	12.03
9 (H)	236.38	0.00	3.526	24.06	391.34	235.99	316.06	336.99	59.79	-29.08	11.57
10 (I)	362.27	0.18	4.951	24.06	416.96	260.33	350.66	378.91	63.32	-25.81	11.63
18 (Me)	250.4	-0.17	4.025	24.06	423.65	252.8	336.91	360.39	58.71	-30.02	10.92
11(NO_2)	281.37	0.78	3.917	69.88	432.43	258.07	342.18	365.96	70.19	-39.84	13.75
12 (OCOMe)	294.41	0.31	3.059	41.13	480.16	289.03	368.70	396.17	65.48	-43.58	13.12
13 (OCF_3)	320.37	0.35	4.738	33.29	457.22	276.98	361.43	385.94	65.32	-24.35	12.23
15 (OMe)	266.4	-0.27	3.629	33.29	439.56	262.09	346.38	370.90	59.13	-30.52	12.16
22 (Pentyl)	306.51	-0.15	6.141	24.06	545.57	320.95	421.48	455.40	58.67	-30.35	9.25
16 (SMe)	282.47	0.00	4.085	24.06	447.32	271.69	361.47	389.16	62.44	-26.72	12.03
17 (SO_2Me)	314.47	0.72	2.641	58.2	485.51	288.75	371.42	405.02	68.11	-43.95	13.96

Table S27. Overview of the descriptors calculated using ACD/I-Labs – part 1.

Compound	Molar refrac-tivity	Molec. volume	Parachor	Index refraction	TPSA	Bp	logD (pH1.7)	logD (pH4.6)	logD (pH6.5)	logD (pH7.4)	logD (pH8)
21 (Bu)	93.35	287.5	735.5	1.562	56.15	390.65	5.34	5.52	552	5.52	5.52
2 (CF_3)	79.52	255.2	636.5	1.535	56.15	341.06	5.17	5.18	5.18	5.18	5.18
3 (Cl)	79.43	233.6	615.2	1.595	56.15	357.38	4.66	4.71	4.71	4.71	4.71
4 (CN)	78.14	235.3	634.9	1.578	79.94	391.61	3.91	3.91	3.91	3.91	3.91
5 ($COCF_3$)	85.02	269.6	688.1	1.543	73.22	389.86	4.63	4.68	4.68	4.68	4.68
7 (COOMe)	86.31	259.6	684.2	1.579	82.45	400.69	4.13	4.14	4.14	4.14	4.14
19 (Et)	84.09	254.5	655.9	1.574	56.15	361.06	4.75	4.92	4.93	4.93	4.93
8 (F)	74.53	225.9	586.5	1.573	56.15	330.44	3.96	4.09	4.09	4.09	4.09
9 (H)	74.54	221.7	579.3	1.587	56.15	331.21	3.79	3.96	3.96	3.96	3.96
10 (I)	87.45	243.8	654.9	1.636	56.15	387.31	4.86	4.92	4.92	4.92	4.92
18 (Me)	79.36	238	617	1.581	56.15	347.6	4.25	4.43	4.43	4.43	4.43
11(NO_2)	81.08	233.5	634.8	1.611	104.98	402.44	3.92	3.92	3.92	3.92	3.92
12 (OCOMe)	85.88	259.3	680.8	1.576	82.45	397.27	3.97	4.00	4.00	4.00	4.00
13 (OCF_3)	81.67	262.1	659.6	1.535	65.38	342.6	4.89	4.94	4.94	4.94	4.94
15 (OMe)	81.22	245.7	636	1.575	65.38	368.81	3.46	4.06	4.06	4.06	4.06
22 (Pentyl)	97.99	304	775.3	1.557	56.15	404.97	6.07	6.25	6.25	6.25	6.25
16 (SMe)	86.21	255.5	677.9	1.589	81.45	390.46	4.71	4.88	4.88	4.88	4.88
17 (SO_2Me)	87.02	263.5	705.2	1.574	98.67	461.48	3.17	3.17	3.17	3.17	3.17

Table S28. Overview of the descriptors calculated using ACD/I-Labs – part 2.

Compound	pKa (arom NH)	pKa (alkyl NH)	pKa2 (NH)	logS (pH1.7)	logS (pH4.6)	logS (pH6.5)	logS (pH7.4)	logS (pH8)	Vx	logPS	Vd	polarizability
21 (Bu)	13.04	15.93	13	-4.81	-4.83	-4.83	-4.83	-4.83	2.59	-1.1	3.33	37.01
2 (CF_3)	11.96	15.56	12.5	-4.59	-4.61	-4.61	-4.61	-4.61	2.22	-1.1	2.73	31.52
3 (Cl)	12.44	15.7	12.8	-4.16	-4.19	-4.19	-4.19	-4.19	2.15	-1.1	2.45	31.49
4 (CN)	11.71	15.59	12.4	-4.05	-4.08	-4.08	-4.08	-4.08	2.18	-1.2	2.12	30.98
5 ($COCF_3$)	11.53	15.61	12.8	-4.7	4.72	-4.72	-4.72	-4.72	2.37	-1.1	2.4	33.7
7 (COOMe)	12.1	15.75	12.6	-3.99	-4.01	-4.01	-4.01	-4.01	2.38	-1.2	2.11	34.21
19 (Et)	13.02	15.93	13	-4.12	-4.14	-4.14	-4.14	-4.14	2.3	-1.1	2.7	33.33
8 (F)	12.67	15.79	12.9	-3.66	-3.68	-3.68	-3.68	-3.68	2.04	-1.1	2.2	29.54
9 (H)	12.75	15.84	13	-3.4	-3.41	-3.41	-3.41	-3.41	2.02	-1.1	2.17	29.55
10 (I)	12.41	15.7	12.8	-4.09	-4.11	-4.11	-4.11	-4.11	2.28	-1.1	2.64	34.66
18 (Me)	13.02	15.93	13	-3.77	-3.79	-3.79	-3.79	-3.79	2.16	-1.1	2.48	31.46
11(NO_2)	11.35	15.34	12.3	-4.12	-4.14	-4.14	-4.14	-4.14	2.2	-1.2	2.11	32.14
12 (OCOMe)	12.46	15.8	12.7	-3.45	-3.48	-3.48	-3.48	-3.48	2.38	-1.2	2.1	34.04
13 (OCF_3)	12.16	15.69	12.8	-4.89	-4.91	-4.91	-4.91	-4.91	2.28	-1.1	2.59	32.38
15 (OMe)	12.76	15.88	13.2	-3.55	-3.57	-3.57	-3.57	-3.57	2.22	-1.1	2.13	32.19
22 (Pentyl)	13.03	15.93	13	-5.15	-5.16	-5.16	-5.16	-5.16	2.73	-1.3	4.36	38.84
16 (SMe)	12.8	15.76	13	-4.27	-4.29	-4.29	-4.29	-4.29	2.33	-1.1	2.42	34.17
17 (SO_2Me)	11.63	15.6	12.3	-3.59	-3.61	-3.61	-3.61	-3.61	2.44	-1.4	1.43	34.49

Table S29. Overview of the descriptors calculated using VCCLAB e-Dragon – part 1.

Compound	MW	AMW	Sv	Se	Sp	Ss	Mv	Me	Mp	Ms	SCBO	ARR	RBN
21 (Bu)	292.54	6.09	27.85	46.77	30.56	37.67	0.58	0.97	0.64	1.88	24	0.300	9
2 (CF_3)	304.42	7.81	23.39	39.65	25.09	56.42	0.60	1.02	0.64	2.82	24	0.300	6
3 (Cl)	270.86	7.52	22.16	35.56	24.37	35.28	0.62	0.99	0.68	2.08	21	0.353	6
4 (CN)	261.43	7.07	22.85	36.45	24.76	39.67	0.62	0.99	0.67	2.20	24	0.333	6
5 ($COCF_3$)	332.43	8.11	24.9	41.98	26.54	65.08	0.61	1.02	0.65	2.96	27	0.273	7
7 (COOMe)	294.46	7.01	25.08	41.77	27.18	45.33	0.60	0.99	0.65	2.27	25	0.300	8
19 (Et)	264.48	6.30	24.65	41.00	27.03	34.67	0.59	0.98	0.64	1.93	22	0.333	7
8 (F)	254.41	7.07	21.57	35.75	23.45	39.17	0.60	0.99	0.65	2.30	21	0.353	6
9 (H)	236.42	6.57	21.46	35.23	23.51	31.5	0.60	0.98	0.65	1.97	20	0.375	6
10 (I)	362.31	10.06	22.88	35.30	26.17	33.29	0.64	0.98	0.73	1.96	21	0.353	6
18 (Me)	250.45	6.42	23.05	38.12	25.27	33.17	0.59	0.98	0.65	1.95	21	0.353	6
11(NO_2)	281.42	7.41	22.88	38.11	24.67	47.17	0.60	1.00	0.65	2.48	25	0.316	7
12 (OCOMe)	294.46	7.01	25.08	41.77	27.18	45.33	0.60	0.99	0.65	2.27	25	0.300	8
13 (OCF_3)	320.42	8.01	23.9	40.98	25.54	59.92	0.60	1.02	0.64	2.85	25	0.286	7
15 (OMe)	266.45	6.66	23.57	39.44	25.73	36.67	0.59	0.99	0.64	2.04	22	0.333	7
22 (Pentyl)	306.57	6.01	29.45	49.65	32.32	39.17	0.58	0.97	0.63	1.87	25	0.286	10
16 (SMe)	282.52	7.06	24.14	39.19	26.92	35.00	0.60	0.98	0.67	1.94	22	0.333	7
17 (SO_2Me)	314.52	7.49	25.17	41.85	27.83	48.08	0.60	1.00	0.66	2.40	26	0.300	7

Table S30. Overview of the descriptors calculated using VCCLAB e-Dragon – part 2.

Compound	RBF	ZM 1	ZM1V	ZM 2	ZM2V	Qindex	SNar	HNar	GNar	Xt	Dz	Ram	Pol
21 (Bu)	0.188	86	150.444	91	160.667	6	13	1.818	1.916	0.277	41	3	22
2 (CF_3)	0.154	92	300.444	99	242.667	9	12.307	1.702	1.85	0.285	45.5	5	24
3 (Cl)	0.167	74	138.049	78	145.778	6	10.92	1.789	1.901	0.303	35.33	3	18
4 (CN)	0.162	78	178.444	83	178.667	6	11.614	1.8	1.906	0.293	37.5	3	20
5 ($COCF_3$)	0.171	102	352.444	111	282.667	10	13.405	1.682	1.839	0.273	50.5	6	28
7 (COOMe)	0.19	88	226.444	95	212.667	7	12.712	1.765	1.888	0.28	43	4	24
19 (Et)	0.167	78	142.444	83	152.667	6	11.614	1.8	1.906	0.293	37	3	20
8 (F)	0.167	74	186.444	78	170.667	6	10.92	1.789	1.901	0.303	36.5	3	18
9 (H)	0.167	68	130.444	71	136.667	5	10.515	1.846	1.929	0.308	33	2	16
10 (I)	0.167	74	137.469	78	143.289	6	10.92	1.789	1.901	0.303	34.4	3	18
18 (Me)	0.154	74	138.444	78	146.667	6	10.92	1.789	1.901	0.303	35	3	18
11(NO_2)	0.184	84	234.444	90	222.667	7	12.019	1.754	1.882	0.288	41.5	4	22
12 (OCOMe)	0.19	88	226.444	93	218.667	7	12.712	1.765	1.888	0.28	43	4	22
13 (OCF_3)	0.175	96	336.444	101	274.667	9	13	1.714	1.857	0.277	48.5	5	23
15 (OMe)	0.175	78	174.444	83	172.667	6	11.614	1.8	1.906	0.293	38	3	20
22 (Pentyl)	0.196	90	154.444	95	164.667	6	13.693	1.826	1.919	0.27	43	3	23
16 (SMe)	0.175	78	138.889	83	146	6	11.614	1.8	1.906	0.293	37	3	20
17 (SO_2Me)	0.167	92	210.889	99	154	9	12.307	1.702	1.85	0.285	43	5	24

Table S31. Overview of the descriptors calculated using VCCLAB e-Dragon – part 3.

Compound	LPRS	VDA	MSD	SM TI	SMTIV	GM TI	GMTIV	Xu	SPI	W	WA	Har
21 (Bu)	93.112	108.1	0.346	4246	5218.444	3982	5798.667	21.002	10.126	1081	5.689	31.518
2 (CF_3)	92.241	103.2	0.329	4050	7366.444	3774	10986.667	20.754	12.611	1032	5.432	32.441
3 (Cl)	73.742	78.353	0.349	2649	3406.16	2475	3886.407	18.022	8.817	666	4.897	26.395
4 (CN)	80.008	87.333	0.346	3108	4665.778	2904	6095.333	19.003	9.27	786	5.137	28.16
5 ($COCF_3$)	105.05	121.545	0.319	5222	9863.111	4876	15371	22.535	14.7	1337	5.788	36.345
7 (COOMe)	92.548	104.9	0.335	4116	6365.111	3846	8537.667	20.84	11.469	1049	5.521	31.982
19 (Et)	80.008	87.333	0.346	3108	3941.778	2904	4464.667	19.003	9.27	786	5.137	28.16
8 (F)	73.742	78.353	0.349	2649	4095.444	2475	5377.667	18.022	8.817	666	4.897	26.395
9 (H)	67.701	70.375	0.355	2256	2991.778	2114	3504	17.073	6.852	563	4.692	24.443
10 (I)	73.742	78.353	0.349	2649	3341.491	2475	3737.281	18.022	8.817	666	4.897	26.395
18 (Me)	73.742	78.353	0.349	2649	3429.444	2475	3939.667	18.022	8.817	666	4.897	26.395
11(NO_2)	86.166	95.579	0.339	3575	5978.111	3337	8430.333	19.905	11.036	908	5.31	30.176
12 (OCOMe)	92.822	106.4	0.34	4178	6467.778	3910	8708.333	20.912	11.557	1064	5.6	31.801
13 (OCF_3)	99.083	114.857	0.332	4725	8922.778	4419	13792.333	21.767	13.203	1206	5.743	33.996
15 (OMe)	80.008	87.333	0.346	3108	4389.778	2904	5427.333	19.003	9.27	786	5.137	28.16
22 (Pentyl)	99.909	119.81	0.347	4933	5990.778	4639	6615.667	22.003	10.537	1258	5.99	33.162
16 (SMe)	80.008	87.333	0.346	3108	3799.556	2904	4143.778	19.003	9.27	786	5.137	28.16
17 (SO_2Me)	92.241	103.2	0.329	4050	5931.556	3774	7647.778	20.754	12.611	1032	5.432	32.441

Table S32. Overview of the descriptors calculated using VCCLAB e-Dragon – part 4.

Compound	Har2	QW	TI2	HyDp	RHyDp	w	ww	Rww	D/D	Wap	WhetZ	Whetm
21 (Bu)	57.918	967.167	5.367	4854	36.916	1285	6264	26.515	161.592	2140	921.134	921.175
2 (CF_3)	59.611	918.167	5.212	4430	38.16	1236	5804	27.36	160.934	2046	853.134	851.205
3 (Cl)	46.231	606.667	4.969	2585	30.676	834	3761	20.976	112.362	1293	550.933	550.736
4 (CN)	50.191	708.5	5.068	3197	32.837	966	4439	22.771	127.553	1542	650.759	650.793
5 ($COCF_3$)	68.789	1186.83	5.422	6089	43.023	1565	7631	31.971	199.029	2654	1106.24	1104.12
7 (COOMe)	58.864	935.167	5.251	4568	37.562	1253	5954	26.923	161.172	2078	863.509	863.591
19 (Et)	50.191	708.5	5.068	3197	32.837	966	4439	22.771	127.553	1542	662.902	662.935
8 (F)	46.231	606.667	4.969	2585	30.676	834	3761	20.976	112.362	1293	555.952	555.429
9 (H)	42.075	521.833	4.886	2093	28.283	719	3215	19.355	98.496	1076	476.67	476.696
10 (I)	46.231	606.667	4.969	2585	30.676	834	3761	20.976	112.362	1293	547.097	546.83
18 (Me)	46.231	606.667	4.969	2585	30.676	834	3761	20.976	112.362	1293	561.286	561.316
11(NO_2)	54.651	812.333	5.15	3812	35.332	1100	5120	24.899	143.743	1793	735.232	735.287
12 (OCOMe)	58.45	950.167	5.32	4700	37.298	1268	6098	26.819	161.41	2108	863.509	863.629
13 (OCF_3)	63.329	1074	5.427	5456	40.029	1422	6932	29.343	179.838	2394	970.75	968.816
15 (OMe)	50.191	708.5	5.068	3197	32.837	966	4439	22.771	127.553	1542	650.652	650.717
22 (Pentyl)	61.758	1126	5.587	5938	38.903	1474	7450	28.416	180.348	2491	1079.75	1079.795
16 (SMe)	50.191	708.5	5.068	3197	32.837	966	4439	22.771	127.553	1542	632.277	632.285
17 (SO_2Me)	59.611	918.167	5.212	4430	38.16	1236	5804	27.36	160.934	2046	787.603	787.599

Table S33. Overview of the descriptors calculated using VCCLAB e-Dragon – part 5.

Compound	Whetv	Whete	Whetp	J	Jhet Z	Jhet m	Jhetv	Jhete	Jhetp	MAX DN	MA XDP	DELS
21 (Bu)	1143.937	928.24	1201.166	2.012	2.393	2.393	1.895	2.375	1.798	0.942	1.628	8.303
2 (CF_3)	1177.029	861.428	1274.298	2.103	2.565	2.569	1.872	2.541	1.736	5.559	4.392	31.251
3 (Cl)	727.409	563.731	766.573	1.983	2.404	2.405	1.794	2.357	1.693	0.996	1.693	9.48
4 (CN)	839.073	657.023	887.225	2.004	2.433	2.433	1.855	2.41	1.748	1.031	2.689	11.356
5 ($COCF_3$)	1491.205	1115.566	1605.264	2.167	2.638	2.642	1.978	2.618	1.845	6.107	4.277	38.64
7 (COOMe)	1163.963	870.834	1237.08	2.07	2.533	2.533	1.878	2.513	1.766	2.006	4.294	16.848
19 (Et)	847.918	669.138	895.159	2.004	2.400	2.400	1.841	2.378	1.737	0.948	1.586	7.636
8 (F)	750.453	562.088	803.938	1.983	2.388	2.389	1.753	2.363	1.636	1.908	4.669	14.933
9 (H)	623.9	482.037	661.152	1.946	2.309	2.309	1.729	2.284	1.625	0.956	1.525	6.232
10 (I)	720.669	566.915	758.919	1.983	2.417	2.418	1.806	2.347	1.706	0.956	1.556	7.076
18 (Me)	727.409	567.087	769.655	1.983	2.371	2.371	1.794	2.347	1.688	0.953	1.558	6.991
11(NO_2)	1006.114	742.259	1078.829	2.045	2.534	2.534	1.846	2.511	1.721	2.423	3.509	19.558
12 (OCOMe)	1236.196	871.039	1324.093	2.041	2.534	2.534	1.765	2.513	1.646	1.991	3.800	16.526
13 (OCF_3)	1497.035	980.046	1637.862	2.085	2.612	2.615	1.714	2.588	1.574	5.926	4.011	29.717
15 (OMe)	894.659	657.055	953.969	2.004	2.437	2.436	1.764	2.413	1.652	0.989	1.600	9.447
22 (Pentyl)	1321.446	1087.29	1383.669	2.005	2.369	2.368	1.907	2.352	1.815	0.94	1.644	8.505
16 (SMe)	843.947	665.66	875.897	2.004	2.495	2.495	1.848	2.388	1.768	0.955	1.581	7.195
17 (SO_2Me)	1084.325	848.647	1106.91	2.103	2.760	2.760	1.995	2.574	1.943	4.064	4.335	23.077

Table S34. Overview of the descriptors calculated using VCCLAB e-Dragon – part 6.

Compound	TIE	S0K	S1K	S2K	S3K	PHI	BLI	PW2	PW3	PW4	PW5	PJI2	CSI
21 (Bu)	38.298	82.439	17.286	10.955	8.21	9.468	1.246	0.525	0.277	0.14	0.079	1	466
2 (CF_3)	58.461	77.684	17.091	8.353	6.868	7.139	1.059	0.564	0.286	0.128	0.066	1	426
3 (Cl)	27.776	63.487	14.58	8.587	6.737	7.365	1.202	0.537	0.277	0.135	0.075	0.857	332
4 (CN)	32.436	71.059	14.787	8.765	6.307	7.201	1.12	0.531	0.288	0.14	0.074	1	372
5 ($COCF_3$)	73.596	89.353	18.763	8.989	7.027	7.666	1.025	0.569	0.297	0.131	0.066	0.875	503
7 (COOMe)	43.22	82.439	16.923	9.724	6.749	8.228	1.097	0.54	0.299	0.141	0.071	0.875	445
19 (Et)	30.193	71.059	15.292	9.202	6.682	7.817	1.218	0.531	0.288	0.14	0.074	1	372
8 (F)	33.815	65.487	14.231	8.288	6.467	6.938	1.135	0.537	0.277	0.135	0.075	0.857	332
9 (H)	22.838	57.245	13.299	8.393	6.444	6.977	1.188	0.523	0.272	0.14	0.072	1	294
10 (I)	26.436	65.487	15.02	8.966	7.081	7.922	1.326	0.537	0.277	0.135	0.075	0.857	332
18 (Me)	26.44	65.487	14.295	8.343	6.517	7.016	1.19	0.537	0.277	0.135	0.075	0.857	332
11(NO_2)	44.621	74.711	15.861	8.829	6.5	7.371	1.079	0.546	0.289	0.135	0.07	1	399
12 (OCOMe)	41.221	82.439	16.923	9.724	7.927	8.228	1.104	0.541	0.27	0.14	0.075	0.875	447
13 (OCF_3)	57.693	83.484	18.05	9.107	8.786	7.827	1.031	0.559	0.261	0.133	0.072	0.875	476
15 (OMe)	29.099	71.059	15.253	9.168	6.653	7.769	1.143	0.531	0.288	0.14	0.074	1	372
22 (Pentyl)	42.603	86.239	18.283	11.846	8.971	10.31	1.258	0.524	0.275	0.138	0.075	0.889	518
16 (SMe)	29.511	71.059	15.641	9.506	6.945	8.26	1.347	0.531	0.288	0.14	0.074	1	372
17 (SO_2Me)	56.552	80.439	17.247	8.475	6.978	7.308	1.362	0.564	0.286	0.128	0.066	1	426

Table S35. Overview of the descriptors calculated using VCCLAB e-Dragon – part 7.

Compound	ECC	AECC	DECC	MDDD	UNIP	CENT	VAR	BAC	Lop	ICR	D/Dr06	GGI1	GGI2
21 (Bu)	239	11.95	2.05	22.32	79	582	86	27	2.704	3.084	77.792	3	1.111
2 (CF_3)	218	10.9	1.91	20.04	76	544	86	37	2.561	2.909	77.134	5	2
3 (Cl)	169	9.941	1.599	14.138	57	363	63	18	2.606	2.778	64.562	3	1.111
4 (CN)	190	10.556	1.778	16.741	64	420	70	21	2.636	2.933	68.753	3	1.111
5 ($COCF_3$)	258	11.727	2.091	23.686	89	716	104	51	2.559	2.959	86.229	5.5	2.444
7 (COOMe)	228	11.4	1.94	20.88	77	558	86	31	2.641	2.971	77.372	3.5	1.778
19 (Et)	190	10.556	1.778	16.741	64	420	70	21	2.636	2.933	68.753	3	1.111
8 (F)	169	9.941	1.599	14.138	57	363	63	18	2.606	2.778	64.562	3	1.111
9 (H)	149	9.313	1.563	12.547	51	310	56	13	2.625	2.733	60.696	2	0.667
10 (I)	169	9.941	1.599	14.138	57	363	63	18	2.606	2.778	64.562	3	1.111
18 (Me)	169	9.941	1.599	14.138	57	363	63	18	2.606	2.778	64.562	3	1.111
11(NO_2)	204	10.737	1.856	18.731	70	486	78	28	2.604	2.931	72.943	3.5	1.556
12 (OCOMe)	229	11.45	1.995	21.48	78	568	86	31	2.641	2.971	77.61	4	1.111
13 (OCF_3)	244	11.619	2.077	22.884	84	648	95	40	2.603	2.951	82.038	5.5	1.333
15 (OMe)	190	10.556	1.778	16.741	64	420	70	21	2.636	2.933	68.753	3	1.111
22 (Pentyl)	266	12.667	2.222	25.288	87	689	95	30	2.739	3.118	82.548	3	1.111
16 (SMe)	190	10.556	1.778	16.741	64	420	70	21	2.636	2.933	68.753	3	1.111
17 (SO_2Me)	218	10.9	1.91	20.04	76	544	86	37	2.561	2.909	77.134	5	2

Table S36. Overview of the descriptors calculated using VCCLAB e-Dragon – part 8.

Compound	GGI3	GGI4	GGI5	GGI6	GGI7	GGI8	GGI9	GGI10	JGI1	JGI2	JGI3	JGI4	JGI5
21 (Bu)	0.625	0.622	0.319	0.283	0.165	0.082	0.066	0.037	0.15	0.048	0.028	0.031	0.017
2 (CF_3)	1.125	0.702	0.528	0.264	0.262	0.107	0.086	0.07	0.25	0.077	0.047	0.033	0.031
3 (Cl)	0.625	0.582	0.229	0.223	0.106	0.057	0.046	0.037	0.176	0.056	0.035	0.039	0.016
4 (CN)	0.625	0.622	0.292	0.223	0.137	0.057	0.046	0.037	0.167	0.053	0.031	0.037	0.019
5 ($COCF_3$)	1.375	0.982	0.493	0.424	0.231	0.206	0.106	0.087	0.25	0.084	0.049	0.039	0.022
7 (COOMe)	0.875	0.662	0.438	0.263	0.2	0.107	0.066	0.054	0.175	0.074	0.036	0.032	0.024
19 (Et)	0.625	0.622	0.292	0.223	0.137	0.057	0.046	0.037	0.167	0.053	0.031	0.037	0.019
8 (F)	0.625	0.582	0.229	0.223	0.106	0.057	0.046	0.037	0.176	0.056	0.035	0.039	0.016
9 (H)	0.563	0.391	0.174	0.141	0.075	0.033	0.026	0.021	0.125	0.037	0.035	0.028	0.013
10 (I)	0.625	0.582	0.229	0.223	0.106	0.057	0.046	0.037	0.176	0.056	0.035	0.039	0.016
18 (Me)	0.625	0.582	0.229	0.223	0.106	0.057	0.046	0.037	0.176	0.056	0.035	0.039	0.016
11(NO_2)	0.875	0.662	0.41	0.223	0.2	0.082	0.066	0.054	0.184	0.068	0.04	0.035	0.026
12 (OCOMe)	0.875	0.782	0.347	0.344	0.137	0.132	0.066	0.054	0.2	0.046	0.04	0.037	0.018
13 (OCF_3)	1.125	0.942	0.375	0.424	0.168	0.181	0.086	0.07	0.262	0.049	0.049	0.041	0.018
15 (OMe)	0.625	0.622	0.292	0.223	0.137	0.057	0.046	0.037	0.167	0.053	0.031	0.037	0.019
22 (Pentyl)	0.625	0.622	0.319	0.283	0.181	0.103	0.066	0.054	0.143	0.046	0.027	0.03	0.016
16 (SMe)	0.625	0.622	0.292	0.223	0.137	0.057	0.046	0.037	0.167	0.053	0.031	0.037	0.019
17 (SO_2Me)	1.125	0.702	0.528	0.264	0.262	0.107	0.086	0.07	0.25	0.077	0.047	0.033	0.031

Table S37. Overview of the descriptors calculated using VCCLAB e-Dragon – part 9.

Compound	JGI6	JGI7	JGI8	JGI9	JGI 10	JGT	W3D	J3D	H3D	AGDD	DDI	ADDD
21 (Bu)	0.018	0.012	0.006	0.006	0.004	0.32	8463.403	3.44	244.862	352.642	1143.54	47.648
2 (CF_3)	0.018	0.019	0.007	0.009	0.008	0.498	4774.758	3.275	177.915	244.859	762.547	39.105
3 (Cl)	0.019	0.01	0.006	0.007	0.006	0.369	3798.21	3.239	158.763	211.012	649.62	36.09
4 (CN)	0.017	0.011	0.005	0.006	0.005	0.352	4128.441	3.239	164.81	223.159	686.428	37.104
5 ($COCF_3$)	0.025	0.014	0.013	0.007	0.008	0.512	5488.705	3.308	190.867	267.742	838.046	40.88
7 (COOMe)	0.018	0.014	0.008	0.006	0.006	0.393	6020.361	3.253	196.13	286.684	882.718	42.034
19 (Et)	0.017	0.011	0.005	0.006	0.005	0.352	5821.687	3.348	200.9	277.223	874.67	41.651
8 (F)	0.019	0.01	0.006	0.007	0.006	0.369	3785.711	3.246	159.297	210.317	647.261	35.959
9 (H)	0.013	0.008	0.004	0.004	0.004	0.272	3776.732	3.252	159.793	209.818	645.566	35.865
10 (I)	0.019	0.01	0.006	0.007	0.006	0.369	3809.957	3.232	158.361	211.664	651.865	36.215
18 (Me)	0.019	0.01	0.006	0.007	0.006	0.369	4757.331	3.283	179.438	243.966	758.394	38.892
11(NO_2)	0.016	0.015	0.006	0.007	0.007	0.404	4447.192	3.255	171.477	234.063	723.05	38.055
12 (OCOMe)	0.023	0.01	0.01	0.005	0.006	0.396	6033.28	3.246	196.203	287.299	879.506	41.881
13 (OCF_3)	0.027	0.011	0.013	0.006	0.007	0.483	5142.346	3.281	184.938	257.117	793.19	39.659
15 (OMe)	0.017	0.011	0.005	0.006	0.005	0.352	5067.614	3.318	186.899	253.381	784.712	39.236
22 (Pentyl)	0.016	0.011	0.007	0.005	0.005	0.307	10057.21	3.476	267.515	394.401	1294.89	50.78
16 (SMe)	0.017	0.011	0.005	0.006	0.005	0.352	5198.882	3.249	183.989	259.944	800.303	40.015
17 (SO_2Me)	0.018	0.019	0.007	0.009	0.008	0.498	5885.456	3.316	197.804	280.26	886.475	42.213

Table S38. Overview of the descriptors calculated using VCCLAB e-Dragon – part 10.

Compound	G1	G2	RGyr	SPAN	SPA M	SPH	ASP	FDI	PJI3	L/B w	SEig	HOM A	RCI
21 (Bu)	25.735	12.377	5.217	10.412	0.466	0.942	0.875	0.997	0.91	36.32	98.213	0.994	1.389
2 (CF_3)	28.682	12.884	4.942	10.415	0.517	0.95	0.88	1	0.965	31.2	80.56	0.994	1.389
3 (Cl)	23.112	10.996	4.692	9.678	0.519	0.947	0.894	1	0.927	31.94	74.203	0.994	1.389
4 (CN)	23.022	11.394	4.656	9.475	0.506	0.951	0.883	1	0.912	28.56	76.198	0.98	1.398
5 ($COCF_3$)	32.622	14.25	5.107	10.911	0.516	0.955	0.849	1	0.846	22.29	85.423	0.697	1.401
7 (COOMe)	27.09	12.945	5.009	10.289	0.495	0.962	0.869	1	0.909	24.76	86.417	0.696	1.398
19 (Et)	23.191	11.24	4.681	9.55	0.477	0.947	0.868	0.997	0.852	28.14	85.779	0.994	1.389
8 (F)	22.305	10.881	4.492	9.24	0.507	0.947	0.877	1	0.883	27.46	73.744	0.995	1.389
9 (H)	20.301	10.091	4.265	8.748	0.493	0.947	0.854	1	0.866	23.03	73.46	0.994	1.389
10 (I)	28.575	12.426	5.056	11.409	0.563	0.947	0.942	1	0.966	58.28	74.712	0.994	1.389
18 (Me)	21.77	10.671	4.472	9.15	0.484	0.952	0.865	1	0.968	24.94	79.814	0.994	1.389
11(NO_2)	26.027	12.646	4.813	9.943	0.512	0.954	0.885	1	0.959	29.36	78.175	0.994	1.389
12 (OCOMe)	27.318	13.037	4.996	10.29	0.495	0.956	0.878	1	0.941	34.93	86.364	0.996	1.389
13 (OCF_3)	31.555	13.809	5.081	10.8	0.52	0.95	0.898	0.997	0.869	40.15	83.432	0.996	1.389
15 (OMe)	23.909	11.597	4.623	9.53	0.488	0.955	0.857	0.987	0.857	23	81.868	0.996	1.389
22 (Pentyl)	26.956	12.946	5.528	10.869	0.462	0.948	0.884	1	0.942	40.58	104.61	0.994	1.389
16 (SMe)	24.687	11.892	4.843	9.995	0.5	0.952	0.886	1	0.886	30.2	83.132	0.995	1.389
17 (SO_2Me)	30.758	15.426	5.003	10.625	0.503	0.941	0.872	1	0.88	30.38	87.454	0.994	1.389

Table S39. Overview of the descriptors calculated using VCCLAB e-Dragon – part 11.

Compound	AROM	HOMT	DISPm	QXXm	QYYm	QZZm	DISPv	QXXv	QYYv	QZZv	DISPe	QXXe
21 (Bu)	0.993	5.961	5.075	28.705	647.94	653.376	6.037	39.827	905.059	915.052	0.039	88.303
2 (CF_3)	0.993	5.961	27.637	31.687	707.208	722.189	12.948	28.894	572.655	587.712	1.008	68.102
3 (Cl)	0.993	5.966	26.953	23.374	569.069	583.554	19.559	26.057	510.19	524.495	0.358	56.807
4 (CN)	0.982	5.881	21.176	22.032	507.608	521.075	19.636	26.931	553.879	568.919	0.3	58.587
5 ($COCF_3$)	0.969	5.578	29.254	49.425	819.456	850.996	15.448	37.958	655.28	678.886	1.164	88.582
7 (COOMe)	0.981	5.567	16.128	29.869	628.289	648.718	11.326	37.922	700.274	725.132	0.454	82.938
19 (Et)	0.993	5.961	9.54	22.493	474.965	484.582	9.375	30.788	647.205	660.187	0.087	69.677
8 (F)	0.993	5.97	21.537	21.129	463.488	475.983	14.737	25.139	466.292	479.833	0.482	56.641
9 (H)	0.993	5.961	15.34	19.451	366.195	377.173	14.13	24.954	458.009	471.384	0.151	55.852
10 (I)	0.993	5.961	48.384	36.808	1166.16	1192.58	25.118	27.541	571.102	586.777	0.204	56.763
18 (Me)	0.993	5.961	11.456	20.238	415.188	427.022	10.804	27.421	549.487	564.288	0.114	62.233
11(NO_2)	0.993	5.961	24.387	25.638	600.039	616.395	14.288	27.661	547.296	562.986	0.712	62.535
12 (OCOMe)	0.993	5.974	16.211	27.18	630.283	638.234	11.192	32.054	717.892	732.945	0.417	71.015
13 (OCF_3)	0.993	5.974	29.022	29.06	801.485	810.94	13.235	27.898	614.958	627.443	1.145	64.408
15 (OMe)	0.993	5.974	14.067	25.081	478.59	494.403	11.128	34.656	567.175	589.192	0.305	76.617
22 (Pentyl)	0.993	5.961	3.182	31.25	762.65	768.089	4.853	43.395	1067.11	1078.18	0.051	95.278
16 (SMe)	0.992	5.97	19.41	25.353	577.311	592.638	14.858	28.76	637.432	651.445	0.154	61.683
17 (SO_2Me)	0.993	5.961	24.299	34.95	707.313	717.158	12.91	34.454	687.352	699.56	0.556	77.266

Table S40. Overview of the descriptors calculated using VCCLAB e-Dragon – part 12.

Compound	QYYe	QZZe	DISPp	QXXp	QYYp	QZZp	L1u	L2u	L3u	P1u	P2u	G1u
21 (Bu)	1691.44	1712.73	0.451	47.078	1005.68	1016.21	35.972	1.189	0.73	0.949	0.031	0.152
2 (CF_3)	1080.17	1111.48	0.678	33.151	594.473	611.398	25.642	1.26	0.458	0.937	0.046	0.185
3 (Cl)	807.23	835.045	1.551	31.491	564.251	580.971	22.119	1.21	0.421	0.931	0.051	0.177
4 (CN)	877.28	906.521	1.329	31.847	586.118	603.167	23.599	1.224	0.415	0.935	0.048	0.161
5 ($COCF_3$)	1221.38	1272.3	0.76	41.432	676.72	701.311	27.502	1.599	0.444	0.931	0.054	0.182
7 (COOMe)	1290.84	1340.68	0.622	43.712	757.939	785.469	30.52	1.62	0.41	0.938	0.05	0.18
19 (Et)	1177.29	1204.03	0.708	36.786	715.275	729.894	28.619	1.194	0.536	0.943	0.039	0.156
8 (F)	812.05	839.85	0.997	30.153	497.519	513.092	21.956	1.204	0.419	0.931	0.051	0.177
9 (H)	774.95	802.06	1.058	30.156	500.148	515.717	21.843	1.198	0.418	0.931	0.051	0.177
10 (I)	794.92	822.82	2.553	34.814	708.421	728.142	22.278	1.22	0.42	0.931	0.051	0.162
18 (Me)	974.89	1005.05	0.818	32.818	604.603	621.635	25.521	1.232	0.432	0.939	0.045	0.185
11(NO_2)	982.67	1015.43	0.852	32.346	573.921	591.401	24.672	1.26	0.407	0.937	0.048	0.16
12 (OCOMe)	1311.94	1342.34	0.729	37.469	778.486	795.429	31.092	1.251	0.48	0.947	0.038	0.169
13 (OCF_3)	1191.56	1215.65	0.625	32.485	633.87	648.718	27.209	1.12	0.483	0.944	0.039	0.183
15 (OMe)	1027.97	1072.73	0.681	40.606	618.231	642.896	25.764	1.576	0.416	0.928	0.057	0.158
22 (Pentyl)	2006.17	2029.51	0.299	51.066	1187.42	1198.87	40.235	1.211	0.737	0.954	0.029	0.15
16 (SMe)	1098.76	1124.24	1.32	35.204	728.016	745.146	27.909	1.128	0.474	0.946	0.038	0.158
17 (SO_2Me)	1247.41	1272.76	1.077	40.216	770.35	783.965	29.127	1.23	0.613	0.941	0.04	0.156

Table S41. Overview of the descriptors calculated using VCCLAB e-Dragon – part 13.

Compound	G2u	G3u	E1u	E2u	E3u	L1m	L2m	L3m	P1m	P2m	G1m	G2m	G3m
21 (Bu)	0.172	0.276	0.587	0.36	0.382	26.125	0.719	0.457	0.957	0.026	0.152	0.181	0.152
2 (CF_3)	0.185	0.279	0.58	0.424	0.38	27.63	0.886	0.306	0.959	0.031	0.173	0.179	0.204
3 (Cl)	0.191	0.255	0.572	0.441	0.339	25.124	0.787	0.161	0.964	0.03	0.162	0.162	0.184
4 (CN)	0.161	0.233	0.562	0.451	0.342	23.156	0.811	0.168	0.959	0.034	0.161	0.202	0.195
5 ($COCF_3$)	0.17	0.234	0.584	0.42	0.345	29.427	1.32	0.321	0.947	0.042	0.17	0.17	0.182
7 (COOMe)	0.217	0.21	0.566	0.4	0.35	25.437	1.027	0.187	0.954	0.039	0.169	0.169	0.204
19 (Et)	0.186	0.204	0.593	0.413	0.378	21.278	0.756	0.263	0.954	0.034	0.156	0.21	0.21
8 (F)	0.237	0.255	0.578	0.437	0.335	21.712	0.791	0.171	0.958	0.035	0.177	0.205	0.177
9 (H)	0.276	0.237	0.582	0.434	0.333	18.393	0.799	0.184	0.949	0.041	0.177	0.177	0.177
10 (I)	0.177	0.237	0.565	0.446	0.339	38.916	0.668	0.12	0.98	0.017	0.162	0.191	0.184
18 (Me)	0.204	0.259	0.583	0.434	0.372	19.708	0.79	0.18	0.953	0.038	0.173	0.159	0.204
11(NO_2)	0.2	0.246	0.572	0.435	0.333	25.455	0.867	0.181	0.96	0.033	0.16	0.174	0.174
12 (OCOMe)	0.169	0.246	0.557	0.386	0.324	25.317	0.725	0.381	0.958	0.027	0.169	0.163	0.281
13 (OCF_3)	0.238	0.238	0.572	0.41	0.339	29.691	0.739	0.333	0.965	0.024	0.158	0.158	0.183
15 (OMe)	0.209	0.223	0.593	0.435	0.333	21.372	0.929	0.192	0.95	0.041	0.158	0.171	0.196
22 (Pentyl)	0.217	0.229	0.578	0.34	0.393	29.373	0.724	0.499	0.96	0.024	0.15	0.229	0.182
16 (SMe)	0.196	0.238	0.563	0.399	0.353	24.407	0.808	0.191	0.961	0.032	0.158	0.183	0.183
17 (SO_2Me)	0.192	0.263	0.587	0.406	0.377	26.634	0.877	0.354	0.956	0.031	0.156	0.21	0.192

Table S42. Overview of the descriptors calculated using VCCLAB e-Dragon – part 14.

Compound	E1m	E2m	E3m	L1v	L2v	L3v	P1v	P2v	G1v	G2v	G3v	E1v	E2v
21 (Bu)	0.31	0.154	0.129	31.966	0.894	0.535	0.957	0.027	0.152	0.172	0.177	0.464	0.204
2 (CF_3)	0.674	0.208	0.179	24.191	0.94	0.295	0.951	0.037	0.173	0.159	0.159	0.516	0.239
3 (Cl)	0.74	0.186	0.054	22.767	0.906	0.264	0.951	0.038	0.177	0.184	0.177	0.606	0.251
4 (CN)	0.541	0.206	0.06	23.982	0.915	0.259	0.953	0.036	0.161	0.161	0.202	0.580	0.253
5 ($COCF_3$)	0.67	0.237	0.18	26.03	1.236	0.288	0.945	0.045	0.17	0.157	0.206	0.524	0.252
7 (COOMe)	0.394	0.164	0.079	27.668	1.248	0.26	0.948	0.043	0.156	0.18	0.198	0.466	0.237
19 (Et)	0.328	0.192	0.086	25.894	0.888	0.36	0.954	0.033	0.156	0.204	0.169	0.485	0.231
8 (F)	0.566	0.207	0.062	21.353	0.897	0.268	0.948	0.04	0.162	0.255	0.162	0.547	0.247
9 (H)	0.413	0.228	0.071	21.077	0.893	0.269	0.948	0.04	0.177	0.255	0.162	0.542	0.246
10 (I)	1.746	0.097	0.03	24.719	0.925	0.256	0.954	0.036	0.162	0.177	0.177	0.695	0.254
18 (Me)	0.348	0.200	0.069	23.561	0.916	0.273	0.952	0.037	0.185	0.159	0.173	0.497	0.243
11(NO_2)	0.609	0.203	0.073	23.664	0.948	0.261	0.951	0.038	0.16	0.255	0.174	0.526	0.249
12 (OCOMe)	0.369	0.139	0.209	28.288	0.939	0.339	0.957	0.032	0.169	0.231	0.18	0.461	0.219
13 (OCF_3)	0.681	0.212	0.158	25.411	0.844	0.322	0.956	0.032	0.183	0.177	0.189	0.499	0.238
15 (OMe)	0.409	0.161	0.071	23.804	1.198	0.268	0.942	0.047	0.158	0.177	0.196	0.507	0.253
22 (Pentyl)	0.308	0.129	0.17	35.693	0.925	0.548	0.96	0.025	0.15	0.206	0.196	0.455	0.199
16 (SMe)	0.434	0.217	0.057	26.107	0.878	0.304	0.957	0.032	0.158	0.171	0.158	0.493	0.25
17 (SO_2Me)	0.495	0.220	0.088	26.875	0.929	0.437	0.952	0.033	0.156	0.231	0.18	0.500	0.235

Table S43. Overview of the descriptors calculated using VCCLAB e-Dragon – part 15.

Compound	E3v	L1e	L2e	L3e	P1e	P2e	G1e	G2e	G3e	E1e	E2e	E3e	L1p
21 (Bu)	0.205	35.451	1.172	0.717	0.949	0.031	0.152	0.152	0.208	0.57	0.35	0.367	32.314
2 (CF_3)	0.162	26.779	1.251	0.464	0.94	0.044	0.173	0.166	0.191	0.633	0.417	0.394	23.375
3 (Cl)	0.139	22.296	1.19	0.408	0.933	0.05	0.177	0.255	0.177	0.581	0.429	0.321	22.861
4 (CN)	0.139	23.665	1.205	0.403	0.936	0.048	0.161	0.291	0.175	0.565	0.438	0.326	23.381
5 ($COCF_3$)	0.145	28.654	1.652	0.449	0.932	0.054	0.157	0.157	0.194	0.634	0.441	0.351	25.185
7 (COOMe)	0.143	30.507	1.589	0.396	0.939	0.049	0.169	0.169	0.174	0.565	0.385	0.327	27.588
19 (Et)	0.171	28.191	1.176	0.524	0.943	0.039	0.156	0.156	0.156	0.575	0.401	0.362	26.048
8 (F)	0.143	22.315	1.181	0.403	0.934	0.049	0.162	0.255	0.162	0.597	0.423	0.315	20.907
9 (H)	0.143	21.588	1.177	0.408	0.932	0.051	0.177	0.162	0.205	0.569	0.42	0.32	20.964
10 (I)	0.134	22.109	1.199	0.409	0.932	0.051	0.162	0.162	0.177	0.556	0.432	0.324	26.841
18 (Me)	0.153	25.156	1.212	0.421	0.939	0.045	0.185	0.179	0.185	0.567	0.421	0.356	23.61
11(NO_2)	0.142	25.399	1.249	0.391	0.939	0.046	0.16	0.167	0.174	0.606	0.427	0.309	22.967
12 (OCOMe)	0.161	30.922	1.214	0.486	0.948	0.037	0.156	0.174	0.156	0.551	0.362	0.334	28.262
13 (OCF_3)	0.151	28.585	1.08	0.491	0.948	0.036	0.171	0.158	0.209	0.631	0.383	0.356	24.475
15 (OMe)	0.139	25.658	1.539	0.404	0.93	0.056	0.158	0.158	0.196	0.588	0.415	0.314	23.722
22 (Pentyl)	0.221	39.681	1.195	0.724	0.954	0.029	0.15	0.177	0.196	0.563	0.331	0.38	36.129
16 (SMe)	0.146	27.573	1.112	0.462	0.946	0.038	0.158	0.171	0.183	0.55	0.39	0.337	26.722
17 (SO_2Me)	0.186	29.19	1.226	0.618	0.941	0.04	0.156	0.174	0.169	0.59	0.405	0.38	27.206

Table S44. Overview of the descriptors calculated using VCCLAB e-Dragon – part 16.

Compound	L2p	L3p	P1p	P2p	G1p	G2p	G3p	E1p	E2p	E3p	L1s	L2s	L3s
21 (Bu)	0.943	0.598	0.955	0.028	0.152	0.225	0.197	0.474	0.23	0.251	25.993	0.623	0.339
2 (CF_3)	1.000	0.322	0.946	0.04	0.173	0.159	0.185	0.482	0.272	0.195	32.077	0.833	0.341
3 (Cl)	0.983	0.299	0.947	0.041	0.177	0.177	0.177	0.611	0.299	0.183	23.41	0.658	0.113
4 (CN)	0.987	0.296	0.948	0.04	0.161	0.189	0.189	0.551	0.298	0.185	26.861	0.682	0.106
5 ($COCF_3$)	1.241	0.316	0.942	0.046	0.17	0.176	0.17	0.491	0.253	0.175	33.944	1.398	0.343
7 (COOMe)	1.309	0.297	0.945	0.045	0.156	0.169	0.204	0.463	0.262	0.19	27.395	0.93	0.131
19 (Et)	0.953	0.407	0.95	0.035	0.156	0.169	0.18	0.491	0.27	0.221	21.649	0.618	0.18
8 (F)	0.977	0.308	0.942	0.044	0.162	0.255	0.177	0.524	0.299	0.194	24.615	0.638	0.107
9 (H)	0.974	0.307	0.942	0.044	0.177	0.177	0.177	0.536	0.298	0.193	19.412	0.628	0.121
10 (I)	0.991	0.279	0.955	0.035	0.162	0.198	0.184	0.821	0.28	0.162	22.716	0.67	0.115
18 (Me)	0.988	0.311	0.948	0.04	0.185	0.179	0.198	0.499	0.285	0.202	20.394	0.643	0.114
11(NO_2)	1.012	0.299	0.946	0.042	0.16	0.174	0.174	0.496	0.288	0.191	29.59	0.811	0.128
12 (OCOMe)	1.002	0.377	0.954	0.034	0.169	0.231	0.174	0.46	0.252	0.2	26.72	0.614	0.375
13 (OCF_3)	0.927	0.343	0.951	0.036	0.183	0.177	0.171	0.463	0.292	0.171	33.93	0.617	0.368
15 (OMe)	1.267	0.309	0.938	0.05	0.158	0.158	0.183	0.503	0.284	0.186	22.36	0.807	0.118
22 (Pentyl)	0.967	0.613	0.958	0.026	0.15	0.168	0.212	0.466	0.218	0.273	29.043	0.658	0.357
16 (SMe)	0.959	0.334	0.954	0.034	0.158	0.165	0.189	0.517	0.302	0.179	22.818	0.643	0.144
17 (SO_2Me)	0.971	0.471	0.95	0.034	0.156	0.18	0.174	0.513	0.26	0.216	27.487	0.913	0.351

Table S45. Overview of the descriptors calculated using VCCLAB e-Dragon – part 17.

Compound	P1s	P2s	G1s	G2s	G3s	E1s	E2s	E3s	Tu	Tm	Tv	Te
21 (Bu)	0.964	0.023	0.188	0.188	0.235	0.307	0.101	0.079	37.89	27.301	33.395	37.339
2 (CF_3)	0.965	0.025	0.188	0.188	0.188	0.91	0.17	0.225	27.36	28.821	25.426	28.493
3 (Cl)	0.968	0.027	0.197	0.197	0.217	0.641	0.136	0.027	23.751	26.072	23.937	23.893
4 (CN)	0.971	0.025	0.193	0.193	0.193	0.728	0.133	0.025	25.237	24.135	25.156	25.272
5 ($COCF_3$)	0.951	0.039	0.183	0.183	0.198	0.895	0.236	0.199	29.546	31.068	27.555	30.755
7 (COOMe)	0.963	0.033	0.188	0.188	0.205	0.458	0.13	0.036	32.551	26.651	29.176	32.493
19 (Et)	0.964	0.028	0.193	0.23	0.193	0.339	0.118	0.043	30.349	22.297	27.142	29.891
8 (F)	0.971	0.025	0.197	0.197	0.197	0.728	0.126	0.024	23.579	22.674	22.517	23.899
9 (H)	0.963	0.031	0.2	0.2	0.2	0.46	0.131	0.031	23.46	19.375	22.24	23.173
10 (I)	0.967	0.029	0.197	0.197	0.217	0.587	0.141	0.028	23.918	39.704	25.9	23.717
18 (Me)	0.964	0.03	0.197	0.197	0.217	0.373	0.126	0.028	27.184	20.679	24.751	26.789
11(NO_2)	0.969	0.027	0.191	0.208	0.208	0.824	0.161	0.034	26.339	26.504	24.872	27.039
12 (OCOMe)	0.964	0.022	0.188	0.188	0.188	0.411	0.089	0.189	32.823	26.423	29.566	32.622
13 (OCF_3)	0.972	0.018	0.185	0.185	0.185	0.889	0.148	0.187	28.813	30.763	26.577	30.155
15 (OMe)	0.96	0.035	0.193	0.193	0.193	0.448	0.117	0.027	27.757	22.494	25.27	27.601
22 (Pentyl)	0.966	0.022	0.185	0.215	0.185	0.301	0.101	0.093	42.183	30.596	37.166	41.6
16 (SMe)	0.967	0.027	0.193	0.193	0.193	0.378	0.141	0.032	29.511	25.406	27.289	29.147
17 (SO_2Me)	0.956	0.032	0.188	0.188	0.205	0.53	0.218	0.072	30.97	27.864	28.24	31.035

Table S46. Overview of the descriptors calculated using VCCLAB e-Dragon – part 18.

Compound	Tp	Ts	Au	Am	Av	Ae	Ap	As	Gu	Gm	Gs	Ku
21 (Bu)	33.854	26.955	69.886	31.066	46.17	67.777	50.331	25.227	0.193	0.161	0.202	0.924
2 (CF_3)	24.696	33.252	44.631	33.188	30.164	46.495	31.208	37.966	0.212	0.185	0.188	0.906
3 (Cl)	24.143	24.18	36.6	23.933	26.873	36.099	29.595	18.117	0.205	0.169	0.203	0.897
4 (CN)	24.663	27.65	39.177	22.792	28.392	38.518	30.282	21.251	0.182	0.185	0.193	0.903
5 ($COCF_3$)	26.742	35.685	56.92	48.712	40.03	60.936	39.61	59.575	0.193	0.174	0.188	0.896
7 (COOMe)	29.194	28.456	62.64	31.086	42.067	61.202	44.693	29.177	0.202	0.18	0.193	0.906
19 (Et)	27.409	22.448	50.158	21.886	32.644	48.525	35.821	17.398	0.181	0.19	0.205	0.914
8 (F)	22.192	25.359	36.138	21.018	25.114	35.822	27.166	18.401	0.22	0.186	0.197	0.897
9 (H)	22.245	20.16	35.816	18.222	24.744	34.702	27.157	14.609	0.226	0.177	0.2	0.897
10 (I)	28.111	23.501	37.05	30.74	29.424	36.04	34.359	17.903	0.189	0.178	0.203	0.897
18 (Me)	24.909	21.151	42.983	19.268	28.271	41.583	30.966	15.508	0.214	0.178	0.203	0.908
11(NO_2)	24.279	30.529	41.635	26.839	28.853	42.133	30.421	27.904	0.199	0.169	0.202	0.905
12 (OCOMe)	29.641	27.709	54.426	28.281	36.469	53.156	39.326	26.647	0.191	0.198	0.188	0.921
13 (OCF_3)	25.745	34.915	44.167	32.085	29.903	45.427	31.399	33.649	0.218	0.166	0.185	0.917
15 (OMe)	25.298	23.285	52.003	24.149	35.218	50.46	37.775	20.772	0.194	0.174	0.193	0.892
22 (Pentyl)	37.708	30.058	79.269	36.284	53.092	77.013	57.663	29.708	0.195	0.184	0.195	0.931
16 (SMe)	28.015	23.605	45.247	24.54	31.131	43.908	34.863	18.05	0.195	0.174	0.193	0.919
17 (SO_2Me)	28.648	28.75	54.42	33.074	37.108	54.604	39.682	35.059	0.199	0.185	0.193	0.911

Table S47. Overview of the descriptors calculated using VCCLAB e-Dragon – part 19.

Compound	Km	Kv	Ke	Kp	Ks	Du	Dm	Dv	De	Dp	Ds	Vu	Vm
21 (Bu)	0.935	0.936	0.924	0.932	0.946	0.443	0.198	0.291	0.429	0.318	0.162	138.987	66.961
2 (CF_3)	0.938	0.927	0.91	0.92	0.947	0.462	0.354	0.306	0.481	0.316	0.435	86.789	69.49
3 (Cl)	0.945	0.927	0.9	0.92	0.952	0.451	0.327	0.332	0.444	0.364	0.268	71.628	53.184
4 (CN)	0.939	0.93	0.905	0.922	0.957	0.452	0.269	0.324	0.443	0.345	0.295	76.399	50.072
5 ($COCF_3$)	0.921	0.917	0.898	0.913	0.927	0.45	0.362	0.307	0.475	0.306	0.443	106.016	92.241
7 (COOMe)	0.932	0.922	0.908	0.917	0.944	0.439	0.212	0.282	0.426	0.305	0.208	115.479	62.631
19 (Et)	0.931	0.931	0.915	0.926	0.947	0.461	0.202	0.296	0.446	0.327	0.167	98.825	48.417
8 (F)	0.936	0.922	0.901	0.913	0.956	0.45	0.278	0.312	0.445	0.339	0.293	70.79	46.629
9 (H)	0.924	0.922	0.897	0.914	0.944	0.45	0.237	0.311	0.436	0.342	0.207	70.227	40.302
10 (I)	0.97	0.932	0.898	0.932	0.95	0.45	0.625	0.361	0.438	0.421	0.252	72.382	73.565
18 (Me)	0.93	0.928	0.909	0.922	0.946	0.463	0.206	0.298	0.448	0.329	0.176	83.735	42.752
11(NO_2)	0.941	0.927	0.909	0.919	0.954	0.447	0.295	0.306	0.447	0.325	0.34	80.622	57.341
12 (OCOMe)	0.937	0.935	0.922	0.93	0.946	0.422	0.239	0.281	0.416	0.304	0.23	105.925	61.703
13 (OCF_3)	0.948	0.934	0.922	0.926	0.958	0.44	0.351	0.296	0.457	0.309	0.408	87.706	70.157
15 (OMe)	0.925	0.913	0.894	0.907	0.94	0.453	0.214	0.3	0.439	0.325	0.197	96.677	50.465
22 (Pentyl)	0.94	0.941	0.931	0.937	0.949	0.437	0.202	0.292	0.425	0.319	0.165	157.354	77.494
16 (SMe)	0.941	0.935	0.919	0.931	0.95	0.438	0.236	0.297	0.426	0.333	0.184	89.676	53.712
17 (SO_2Me)	0.934	0.927	0.911	0.925	0.934	0.457	0.268	0.307	0.458	0.33	0.273	107.342	69.193

Table S48. Overview of the descriptors calculated using VCCLAB e-Dragon – part 20.

Compound	Vv	Ve	Vp	Vs	nH Ac	Ui	Hy	AMR	TPSA(NO)	TPSA (Tot)	BLTF96	BLTD48	BLTA96
21 (Bu)	94.863	134.87	102.38	57.676	2	3	0.262	92.624	24.06	56.15	-4.88	-5.22	-5.33
2 (CF_3)	62.306	90.521	63.423	80.341	5	3	0.41	79.753	24.06	56.15	-4.79	-5.13	-5.24
3 (Cl)	56.25	70.804	60.451	44.032	2	3	0.411	78.585	24.06	56.15	-4.47	-4.76	-4.85
4 (CN)	59.23	75.265	61.775	50.851	3	3.17	0.374	79.517	47.85	79.94	-3.69	-3.89	-3.93
5 ($COCF_3$)	76.85	112.92	76.234	111.54	6	3.17	0.393	85.193	41.13	73.22	-4.22	-4.48	-4.56
7 (COOMe)	80.234	112.90	84.611	60.964	4	3.17	0.361	85.307	50.36	82.45	-3.91	-4.13	-4.19
19 (Et)	68.071	95.773	73.342	42.259	2	3	0.32	83.422	24.06	56.15	-4.46	-4.75	-4.84
8 (F)	52.765	70.346	55.65	45.434	3	3	0.411	73.996	24.06	56.15	-4.36	-4.64	-4.72
9 (H)	52.051	68.242	55.674	36.239	2	3	0.39	73.78	24.06	56.15	-4.02	-4.26	-4.32
10 (I)	61.177	70.594	69.893	43.152	2	3	0.411	86.188	24.06	56.15	-4.69	-5.01	-5.11
18 (Me)	58.918	81.199	63.12	38.156	2	3	0.353	78.821	24.06	56.15	-4.24	-4.51	-4.59
11(NO_2)	59.581	81.565	61.658	61.513	4	3.32	0.445	81.104	69.88	101.9	-3.96	-4.2	-4.26
12 (OCOMe)	75.04	104.02	79.625	60.502	4	3.17	0.361	84.912	50.36	82.45	-3.91	-4.13	-4.19
13 (OCF_3)	63.387	90.734	64.924	76.266	6	3	0.425	81.354	33.29	65.38	-4.08	-4.33	-4.4
15 (OMe)	68.129	94.001	72.358	46.184	3	3	0.374	80.243	33.29	65.38	-3.76	-3.97	-4.02
22 (Pentyl)	108.35	152.95	116.78	66.584	2	3	0.237	97.225	24.06	56.15	-5.08	-5.45	-5.57
16 (SMe)	65.395	87.216	71.428	43.763	2	3	0.374	86.686	24.06	81.45	-4.47	-4.76	-4.85
17 (SO_2Me)	76.255	107.77	80.764	72.615	4	3.32	0.41	87.534	58.2	98.67	-3.62	-3.81	-3.85

S7.5 QSAR Modelling Using Retention Time

With a dataset of 18 compounds, it is possible to have a QSAR model containing up to 3 descriptors. Therefore a stepwise multiple regression analysis was performed using JMP 9.0.0. Log(1/EC_{50}) was modeled against a total of 286 descriptors (see Section S7.5., only experimental retention times included and not calculated logP values), the k-fold cross validation was set to 2 and all possible models with a maximum of three terms were calculated (3,899,181 models in total) and subsequently ranked according to best fit (highest R), as shown in Figure S240.

All Possible Models

Ordered up to best 20 models up to 3 terms per

Model	Number	RSquare	RMSE	AICc	BIC
retention time	1	0,8451	0,2411	5,4659	6,4227
logS (pH1,7)	1	0,6979	0,3367	17,4930	18,4498
logS (pH6,5)	1	0,6957	0,3380	17,6223	18,5792
logS (pH7,4)	1	0,6957	0,3380	17,6223	18,5792
logS (pH8)	1	0,6957	0,3380	17,6223	18,5792
logD (pH1,7)	1	0,6717	0,3511	18,9907	19,9475
Vd	1	0,6644	0,3549	19,3858	20,3427
logD (pH7,4)	1	0,6415	0,3668	20,5721	21,5289
logD (pH8)	1	0,6415	0,3668	20,5721	21,5289
logD (pH4,6)	1	0,6409	0,3671	20,6020	21,5588
BLTD48	1	0,5070	0,4301	26,3053	27,2621
BLTA96	1	0,5059	0,4307	26,3475	27,3043
BLTF96	1	0,5035	0,4317	26,4340	27,3909
PI	1	0,3776	0,4833	30,5012	31,4580
J3D	1	0,3456	0,4956	31,4035	32,3603
E3u	1	0,2702	0,5234	33,3673	34,3241
G3m	1	0,2601	0,5270	33,6153	34,5721
PHI	1	0,2365	0,5353	34,1802	35,1371
S2K	1	0,2360	0,5355	34,1911	35,1479
Hy	1	0,2196	0,5412	34,5753	35,5321
retention time,E2m	2	0,8982	0,2018	1,2678	1,7523
retention time,L/Bw	2	0,8865	0,2132	3,2388	3,7234
retention time,P2m	2	0,8822	0,2172	3,8998	4,3844
retention time,SPAN	2	0,8788	0,2203	4,4208	4,9054
logS (pH1,7),Gs	2	0,8782	0,2208	4,5015	4,9861
logS (pH6,5),Gs	2	0,8771	0,2218	4,6640	5,1486
logS (pH7,4),Gs	2	0,8771	0,2218	4,6640	5,1486
logS (pH8),Gs	2	0,8771	0,2218	4,6640	5,1486
retention time,QYYm	2	0,8760	0,2228	4,8293	5,3139
retention time,MW	2	0,8759	0,2229	4,8390	5,3235
retention time,Mr (g/mol)	2	0,8759	0,2229	4,8394	5,3240
retention time,Du	2	0,8757	0,2231	4,8663	5,3508
retention time,Molar refractivity	2	0,8756	0,2232	4,8806	5,3651
retention time,polarizability	2	0,8756	0,2232	4,8905	5,3750
retention time,L1m	2	0,8753	0,2234	4,9279	5,4125
retention time,QZZm	2	0,8749	0,2238	4,9772	5,4618
retention time,P1p	2	0,8748	0,2238	4,9923	5,4769
retention time,Tm	2	0,8742	0,2244	5,0830	5,5676
retention time,P1m	2	0,8730	0,2255	5,2511	5,7356
retention time,Kp	2	0,8716	0,2268	5,4577	5,9422
retention time,SCBO,RGyr	3	0,9476	0,1499	-6,7706	-7,3188
retention time,Har2,S1K	3	0,9407	0,1596	-4,5156	-5,0637
retention time,pKa (arom NH),SPAN	3	0,9390	0,1617	-4,0300	-4,5781
retention time,S1K,VAR	3	0,9385	0,1624	-3,8678	-4,4159
retention time,Rww,S1K	3	0,9379	0,1632	-3,7101	-4,2583
retention time,Har,S1K	3	0,9371	0,1643	-3,4621	-4,0102
retention time,D/D,S1K	3	0,9368	0,1647	-3,3735	-3,9217
retention time,S2K,PHI	3	0,9360	0,1656	-3,1703	-3,7185
retention time,RHyDp,S1K	3	0,9357	0,1661	-3,0695	-3,6176
retention time,Hammett,SPAN	3	0,9356	0,1662	-3,0432	-3,5913
retention time,pKa (arom NH),MW	3	0,9351	0,1669	-2,8922	-3,4404
retention time,Mr (g/mol),pKa (arom NH)	3	0,9351	0,1669	-2,8916	-3,4397
retention time,pKa (arom NH),Tm	3	0,9328	0,1698	-2,2739	-2,8220
retention time,pKa (arom NH),L1m	3	0,9323	0,1704	-2,1397	-2,6879
retention time,Vsmin,Molar refractivity	3	0,9320	0,1707	-2,0740	-2,6222
retention time,Vsmin,polarizability	3	0,9320	0,1708	-2,0692	-2,6173
PI,Index of refraction,pKa (alkyl NH)	3	0,9304	0,1728	-1,6396	-2,1877
retention time,E1m,E2e	3	0,9301	0,1731	-1,5776	-2,1257
retention time,Hammett,MW	3	0,9296	0,1738	-1,4431	-1,9912
retention time,Mr (g/mol),Hammett	3	0,9296	0,1738	-1,4425	-1,9906

Figure S240. Overview of best 20 models with up to 3 term correlating log(1/EC_{50}) with various descriptors.

One descriptor

From the list of all possible models, it is clear that the single term that has the highest contribution to the model is the retention times (R = 0.84). Other descriptors that can explain the variation in log(1/EC_{50}) by a single term are logS, logD and Vd values. However, logD (pH-dependent distribution coefficient), logS (aqueous solubility) and Vd (volume of distribution) are other types of descriptors to describe lipophilicity. The correlation matrix in Figure S241 (obtained using JMP 9.0.0) clearly shows that logS, logD and Vd are highly correlated with the retention time and therefore the most important descriptor is the lipophilicity (described by the retention time).

Correlations

	retention time	logD (pH1,7)	logD (pH4,6)	logD (pH6,5)	logD (pH7,4)	logD (pH8)	logS (pH1,7)	logS (pH4,6)	logS (pH6,5)	logS (pH7,4)	logS (pH8)	Vd
retention time	1,0000	0,9408	0,9326	0,4229	0,9326	0,9326	-0,8685	-0,2755	-0,8673	-0,8673	-0,8673	0,9017
logD (pH1,7)	0,9408	1,0000	0,9805	0,3232	0,9805	0,9805	-0,8658	-0,1362	-0,8649	-0,8649	-0,8649	0,9170
logD (pH4,6)	0,9326	0,9805	1,0000	0,3457	1,0000	1,0000	-0,8288	-0,1500	-0,8273	-0,8273	-0,8273	0,9464
logD (pH6,5)	0,4229	0,3232	0,3457	1,0000	0,3453	0,3453	-0,3283	-0,1407	-0,3288	-0,3288	-0,3288	0,3552
logD (pH7,4)	0,9326	0,9805	1,0000	0,3453	1,0000	1,0000	-0,8284	-0,1501	-0,8269	-0,8269	-0,8269	0,9463
logD (pH8)	0,9326	0,9805	1,0000	0,3453	1,0000	1,0000	-0,8284	-0,1501	-0,8269	-0,8269	-0,8269	0,9463
logS (pH1,7)	-0,8685	-0,8658	-0,8288	-0,3283	-0,8284	-0,8284	1,0000	-0,0390	0,9999	0,9999	0,9999	-0,7550
logS (pH4,6)	-0,2755	-0,1362	-0,1500	-0,1407	-0,1501	-0,1501	-0,0390	1,0000	-0,0393	-0,0393	-0,0393	-0,2129
logS (pH6,5)	-0,8673	-0,8649	-0,8273	-0,3288	-0,8269	-0,8269	0,9999	-0,0393	1,0000	1,0000	1,0000	-0,7523
logS (pH7,4)	-0,8673	-0,8649	-0,8273	-0,3288	-0,8269	-0,8269	0,9999	-0,0393	1,0000	1,0000	1,0000	-0,7523
logS (pH8)	-0,8673	-0,8649	-0,8273	-0,3288	-0,8269	-0,8269	0,9999	-0,0393	1,0000	1,0000	1,0000	-0,7523
Vd	0,9017	0,9170	0,9464	0,3552	0,9463	0,9463	-0,7550	-0,2129	-0,7523	-0,7523	-0,7523	1,0000

Figure S241. Correlation matrix for retention time.

Two descriptors

From the 20 best models to contain only 2 terms, the majority of the models contain one term for lipophilicity (retention time or logS) and one term for the size of the molecule (SPAN, QYYm, MW, Mr(g/mol), L1m, QZZm, Tm, L/Bw, P2m, E2m, P1m). Some of the models consist of lipophilicity + polarizability (polarizability, molar refractivity, Kp or P1p), but also the polarizability terms show some correlation with the size terms and can be seen as equivalent. The correlation matrix in Figure S242 (JMP 9.0.0) clearly shows that all the different molecular size descriptors are highly correlated and therefore interchangeable. The two-term model suggests that the two most important factors that govern anion transport are lipophilicity and molecular size.

Correlations

	retention time	E2m	L/Bw	P2m	SPAN	Gs	QYYm	MW	Mr (g/mol)	Du	Molar refractivity	polarizability	L1m	QZZm	P1p	Tm	P1m	Kp
retention time	1,0000	-0,3932	0,5225	-0,5535	0,4416	0,1847	0,4081	0,3376	0,3375	-0,0753	0,5583	0,5589	0,4537	0,3962	0,5509	0,4524	0,3907	0,5631
E2m	-0,3932	1,0000	-0,6867	0,6105	-0,4213	-0,2263	-0,4693	-0,3748	-0,3748	0,3877	-0,5760	-0,5757	-0,4929	-0,4595	-0,5106	-0,4802	-0,5281	-0,4881
L/Bw	0,5225	-0,6867	1,0000	-0,9486	0,6735	0,1063	0,7954	0,6768	0,6768	-0,2679	0,4692	0,4693	0,8213	0,7824	0,7232	0,8030	0,9040	0,7118
P2m	-0,5535	0,6105	-0,9486	1,0000	-0,6394	-0,0499	-0,6867	-0,5961	-0,5961	0,3377	-0,5093	-0,5096	-0,7247	-0,6671	-0,8237	-0,7080	-0,8635	-0,8130
SPAN	0,4416	-0,4213	0,6735	-0,6394	1,0000	-0,3666	0,9389	0,9773	0,9773	-0,3202	0,6969	0,6968	0,9271	0,9345	0,5321	0,9374	0,4978	0,5150
Gs	0,1847	-0,2263	0,1063	-0,0499	-0,3666	1,0000	-0,2131	-0,3448	-0,3448	0,3743	-0,0353	-0,0352	-0,1859	-0,2131	0,0844	-0,2056	0,1602	0,0978
QYYm	0,4081	-0,4693	0,7954	-0,6867	0,9389	-0,2131	1,0000	0,9731	0,9731	-0,2030	0,5231	0,5230	0,9934	0,9995	0,4697	0,9946	0,6707	0,4539
MW	0,3376	-0,3748	0,6768	-0,5961	0,9773	-0,3448	0,9731	1,0000	1,0000	-0,2240	0,5676	0,5674	0,9522	0,9729	0,4321	0,9603	0,5289	0,4153
Mr (g/mol)	0,3375	-0,3748	0,6768	-0,5961	0,9773	-0,3448	0,9731	1,0000	1,0000	-0,2240	0,5675	0,5673	0,9522	0,9730	0,4321	0,9603	0,5289	0,4152
Du	-0,0753	0,3877	-0,2679	0,3377	-0,3202	0,3743	-0,2030	-0,2240	-0,2240	1,0000	-0,4322	-0,4324	-0,2226	-0,1936	-0,4328	-0,2252	-0,1664	-0,3974
Molar refractivity	0,5583	-0,5760	0,4692	-0,5093	0,6969	-0,0353	0,5231	0,5676	0,5675	-0,4322	1,0000	1,0000	0,5187	0,5096	0,7239	0,5297	0,1936	0,7156
polarizability	0,5589	-0,5757	0,4693	-0,5096	0,6968	-0,0352	0,5230	0,5674	0,5673	-0,4324	1,0000	1,0000	0,5186	0,5095	0,7243	0,5296	0,1937	0,7160
L1m	0,4537	-0,4929	0,8213	-0,7247	0,9271	-0,1859	0,9934	0,9522	0,9522	-0,2226	0,5187	0,5186	1,0000	0,9923	0,4923	0,9993	0,7231	0,4754
QZZm	0,3962	-0,4595	0,7824	-0,6671	0,9345	-0,2131	0,9995	0,9729	0,9730	-0,1936	0,5096	0,5095	0,9923	1,0000	0,4486	0,9937	0,6630	0,4325
P1p	0,5509	-0,5106	0,7232	-0,8237	0,5321	0,0844	0,4697	0,4321	0,4321	-0,4328	0,7239	0,7243	0,4923	0,4486	1,0000	0,4829	0,5763	0,9979
Tm	0,4524	-0,4802	0,8030	-0,7080	0,9374	-0,2056	0,9946	0,9603	0,9603	-0,2252	0,5297	0,5296	0,9993	0,9937	0,4829	1,0000	0,6975	0,4660
P1m	0,3907	-0,5281	0,9040	-0,8635	0,4978	0,1602	0,6707	0,5289	0,5289	-0,1664	0,1936	0,1937	0,7231	0,6630	0,5763	0,6975	1,0000	0,5623
Kp	0,5631	-0,4881	0,7118	-0,8130	0,5150	0,0978	0,4539	0,4153	0,4152	-0,3974	0,7156	0,7160	0,4754	0,4325	0,9979	0,4660	0,5623	1,0000

Figure S242. Correlation matrix for retention time and molecular size descriptors.

Three descriptors

From the 20 best models that contain 3 terms, three classes of models can be discerned. The first class consists of the combinations {retention time, SCBO, RGyr}, {retention time, Har2, S1K}, {retention time, S1K, VAR}, {retention time, Rww, S1K}, {retention time, Har, S1K}, {retention time, D/D, S1K}, {retention time, S2K, PHI} and {retention time, RHyDp, S1K}. The correlation matrix in Figure S243 clearly shows that this models consist of a term for lipophilicity (retention time) + two highly correlated terms (usually size or shape indices such as the Kier shape index S1K). Because the two other terms are highly correlated (>0.8), these models are discarded and will not be considered further.

Correlations

	retention time	SCBO	Har2	Har	RHyDp	D/D	Rww	S1K	S2K	PHI	VAR	RGyr
retention time	1,0000	0,0796	0,2597	0,2541	0,2485	0,3041	0,2558	0,3758	0,6242	0,6593	0,3039	0,5976
SCBO	0,0796	1,0000	0,9363	0,9378	0,9379	0,9248	0,9315	0,8855	0,3428	0,3249	0,9257	0,6988
Har2	0,2597	0,9363	1,0000	0,9998	0,9997	0,9963	0,9994	0,9795	0,4215	0,4268	0,9967	0,7811
Har	0,2541	0,9378	0,9998	1,0000	0,9999	0,9948	0,9989	0,9785	0,4107	0,4167	0,9953	0,7781
RHyDp	0,2485	0,9379	0,9997	0,9999	1,0000	0,9939	0,9989	0,9772	0,4024	0,4086	0,9945	0,7739
D/D	0,3041	0,9248	0,9963	0,9948	0,9939	1,0000	0,9962	0,9847	0,4877	0,4904	0,9999	0,8068
Rww	0,2558	0,9315	0,9994	0,9989	0,9989	0,9962	1,0000	0,9776	0,4135	0,4190	0,9966	0,7740
S1K	0,3758	0,8855	0,9795	0,9785	0,9772	0,9847	0,9776	1,0000	0,5385	0,5619	0,9848	0,8775
S2K	0,6242	0,3428	0,4215	0,4107	0,4024	0,4877	0,4135	0,5385	1,0000	0,9901	0,4826	0,7396
PHI	0,6593	0,3249	0,4268	0,4167	0,4086	0,4904	0,4190	0,5619	0,9901	1,0000	0,4858	0,7854
VAR	0,3039	0,9257	0,9967	0,9953	0,9945	0,9999	0,9966	0,9848	0,4826	0,4858	1,0000	0,8067
RGyr	0,5976	0,6988	0,7811	0,7781	0,7739	0,8068	0,7740	0,8775	0,7396	0,7854	0,8067	1,0000

Figure S243. Correlation matrix for 3 descriptors – class I.

The second class consists of the combinations {retention time, Vsmin, molar refractivity}, {retention time, Vsmin, polarizability}. The correlation matrix in Figure S244 shows that this models consist of a term for lipophilicity (retention time) + Vsmin + polarizability (molar refractivity, polarizability). The correlation matrix also shows a small inverse correlation between $V_{S,min}$ (the minimum of the surface electrostatic potential, usually correlated with hydrogen bond acceptor ability) and the Hammett constant (anion binding) and some degree of correlation between polarizability and size (molecular weight, Mr). This class of models thus describes a combination of lipophilicity, size and some form of binding.

Correlations

	retention time	Vsmin	Molar refractivity	polarizability	Hammett	Mr (g/mol)
retention time	1,0000	0,5075	0,5583	0,5589	-0,2763	0,3375
Vsmin	0,5075	1,0000	-0,1472	-0,1471	-0,5434	-0,0185
Molar refractivity	0,5583	-0,1472	1,0000	1,0000	-0,1551	0,5675
polarizability	0,5589	-0,1471	1,0000	1,0000	-0,1551	0,5673
Hammett	-0,2763	-0,5434	-0,1551	-0,1551	1,0000	0,3855
Mr (g/mol)	0,3375	-0,0185	0,5675	0,5673	0,3855	1,0000

Figure S244. Correlation matrix for 3 descriptors – class II.

The third class consists of the combinations {retention time, pKa (arom NH), SPAN}, {retention time, Hammett, SPAN}, {retention time, pKa (arom NH), MW}, {retention time, Mr (g/mol), pKa (arom NH)}, {retention time, pKa (arom NH), Tm}, {retention time, Hammett, MW}, {retention time, Mr(g/mol), Hammett} and {retention time, pKa (arom NH), L1M}. The correlation matrix in Figure S245 clearly shows that this models consist of a term for lipophilicity (retention time) + molecular size (SPAN, molecular weight (MW or Mr), Tm and L1M) + anion binding (Hammett constant or pKa of the aromatic NH).

Correlations

	retention time	Mr (g/mol)	Hammett	pKa (arom NH)	Vsmax	MW	SPAN	L1m	Tm
retention time	1,0000	0,3375	-0,2763	0,3048	-0,3121	0,3376	0,4416	0,4537	0,4524
Mr (g/mol)	0,3375	1,0000	0,3855	-0,3646	0,3961	1,0000	0,9773	0,9522	0,9603
Hammett	-0,2763	0,3855	1,0000	-0,9756	0,9750	0,3855	0,3381	0,3141	0,3274
pKa (arom NH)	0,3048	-0,3646	-0,9756	1,0000	-0,9643	-0,3645	-0,3054	-0,2965	-0,3091
Vsmax	-0,3121	0,3961	0,9750	-0,9643	1,0000	0,3960	0,3396	0,3287	0,3406
MW	0,3376	1,0000	0,3855	-0,3645	0,3960	1,0000	0,9773	0,9522	0,9603
SPAN	0,4416	0,9773	0,3381	-0,3054	0,3396	0,9773	1,0000	0,9271	0,9374
L1m	0,4537	0,9522	0,3141	-0,2965	0,3287	0,9522	0,9271	1,0000	0,9993
Tm	0,4524	0,9603	0,3274	-0,3091	0,3406	0,9603	0,9374	0,9993	1,0000

Figure S245. Correlation matrix for 3 descriptors – class III.

Due to the ease of interpretation of the third class of combinations and due to the fact that the combination retention time + SPAN was one of the highest ranking 2-term combination, the final models that are selected include {retention time, SPAN, Hammett}, {retention time, SPAN, pKa (arom NH)} and {retention time, SPAN, Vsmax} as it was shown previously that the Hammett constant, pK_a (arom NH) and $V_{s,max}$ are all suitable descriptors for anion binding.

Final model

Firstly, a standard least-square fitting was performed for $\log(1/EC_{50})$ versus the retention time, pK_a and SPAN using JMP 9.0.0. Secondly, other equally valid models can be obtained by replacing the pK_a with the Hammett constant for substituents in the *para* position or $V_{s,max}$ (it was shown in Section S5.4 that both pK_a, Hammett constants and $V_{s,max}$ correlate with the anion binding ability). Therefore a standard least-square fitting was performed for $\log(1/EC_{50})$ versus the retention time, Hammett constant and SPAN; and for $\log(1/EC_{50})$ versus the retention time, $V_{s,max}$ and SPAN using JMP 9.0. The results are summarized in Figure S246.

$$Log(1/EC_{50}) = A*RT + B*pK_a + C*SPAN + D$$

Summary of Fit

RSquare	0,939031
RSquare Adj	0,925966
Root Mean Square Error	0,16172
Mean of Response	-0,04426
Observations (or Sum Wgts)	18

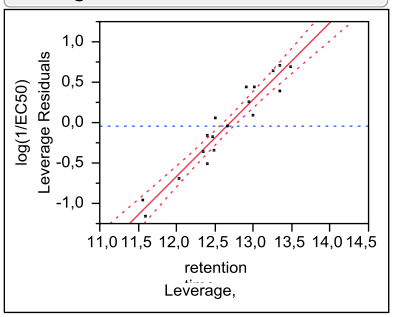
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	5,6392499	1,87975	71,8745
Error	14	0,3661450	0,02615	Prob > F
C. Total	17	6,0053949		<,0001 *

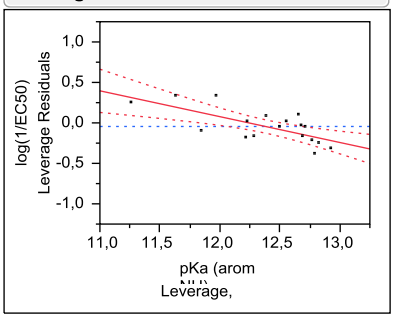
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-5,060439	1,224326	-4,13	0,0010 *
retention time	0,9529615	0,069642	13,68	<,0001 *
pKa (arom NH)	-0,319229	0,085807	-3,72	0,0023 *
SPAN	-0,306621	0,07134	-4,30	0,0007 *

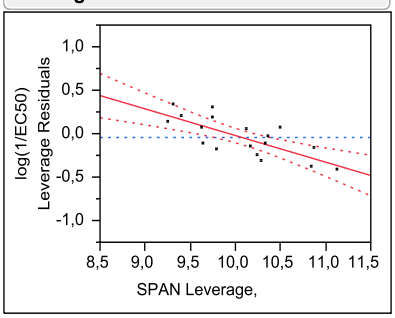
Leverage Plot



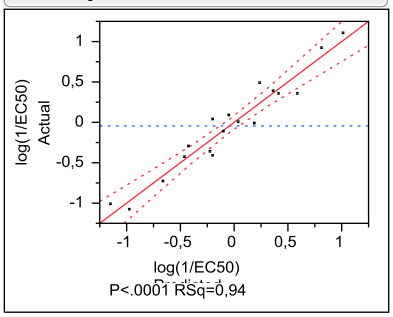
Leverage Plot



Leverage Plot



Actual by Predicted Plot



Scaled Estimates

Continuous factors centered by mean, scaled by

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0,044264	0,038118	-1,16	0,2650
retention time	1,3389109	0,097848	13,68	<,0001 *
pKa (arom NH)	-0,269749	0,072507	-3,72	0,0023 *
SPAN	-0,407959	0,094918	-4,30	0,0007 *

$Log(1/EC_{50}) = A*RT + B*Hammett + C*SPAN + D$ Eq. 7

Summary of Fit

RSquare	0,935595
RSquare Adj	0,921794
Root Mean Square Error	0,166214
Mean of Response	-0,04426
Observations (or Sum Wgts)	18

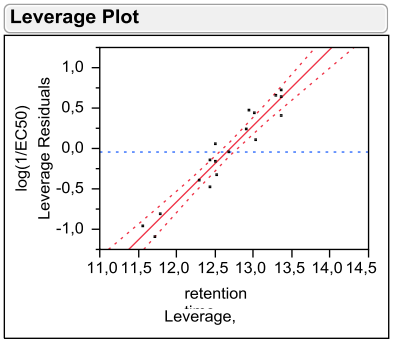
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	5,6186167	1,87287	67,7913
Error	14	0,3867782	0,02763	Prob > F
C. Total	17	6,0053949		<,0001 *

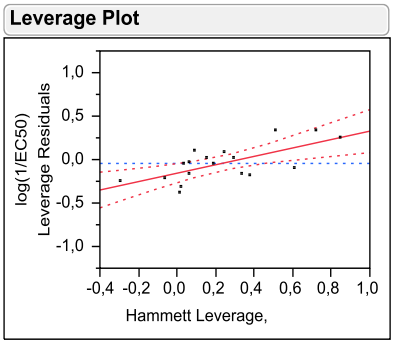
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-9,003356	0,760624	-11,84	<,0001 *
retention time	0,9455587	0,071064	13,31	<,0001 *
Hammett	0,4839938	0,137693	3,52	0,0034 *
SPAN	-0,309388	0,074324	-4,16	0,0010 *

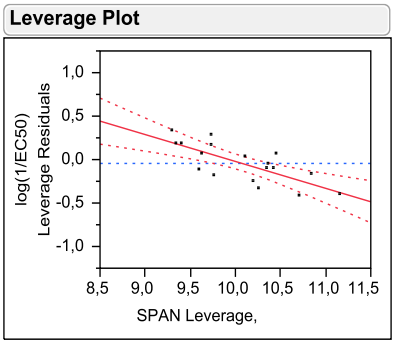
Leverage Plot



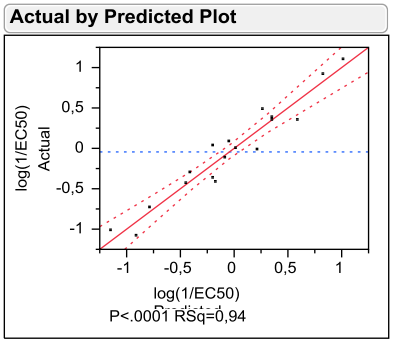
Leverage Plot



Leverage Plot



Actual by Predicted Plot



Scaled Estimates

Continuous factors centered by mean, scaled by

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0,044264	0,039177	-1,13	0,2775
retention time	1,3285099	0,099844	13,31	<,0001 *
Hammett	0,2589367	0,073666	3,52	0,0034 *
SPAN	-0,41164	0,098889	-4,16	0,0010 *

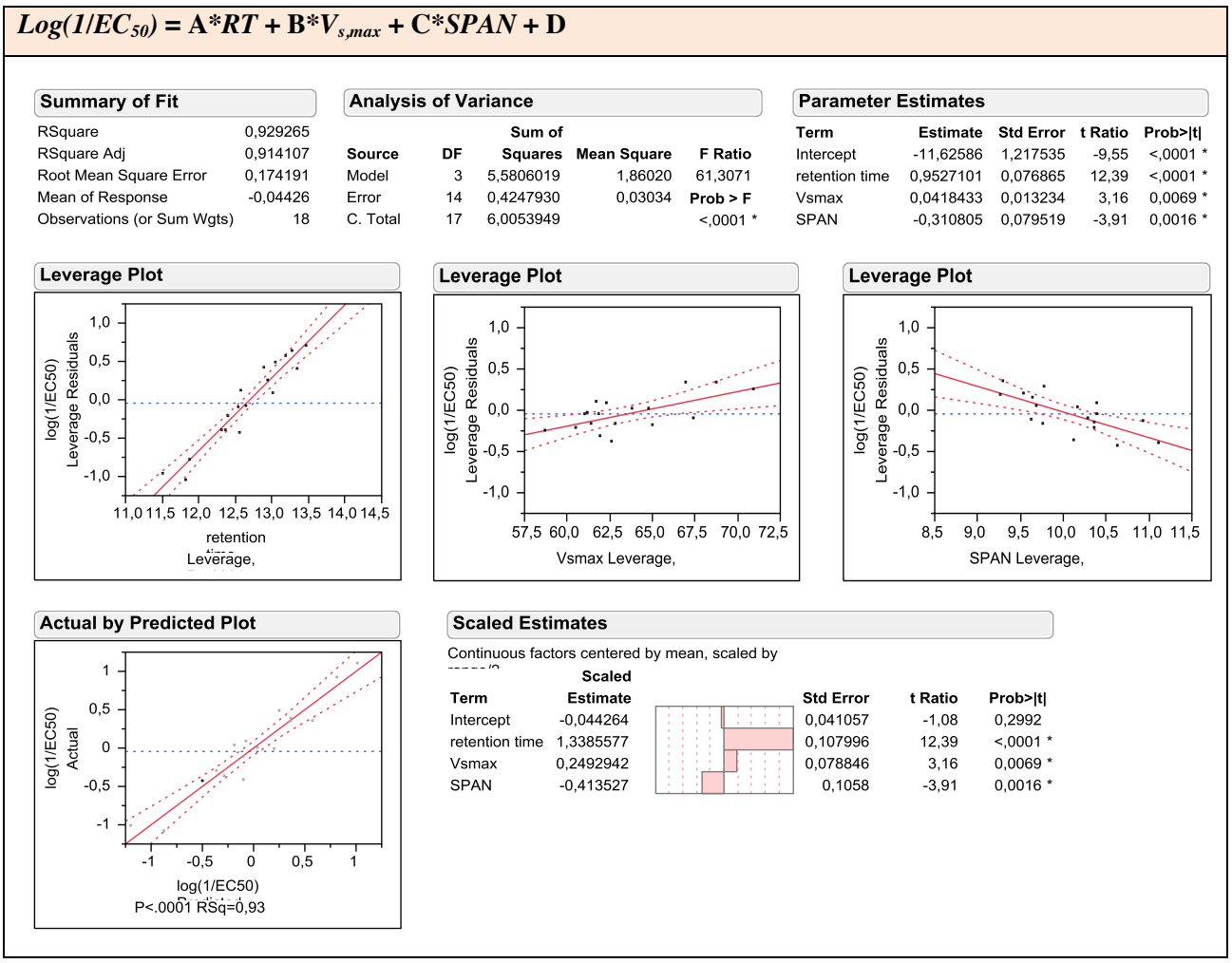


Figure S246. Overview of models correlating $\log(1/EC_{50})$ with retention time (RT), anion binding (Hammett, pK_a or $V_{S,max}$) and SPAN.

S7.6 QSAR Modelling Using logP

It is also possible to perform the same analysis using calculated logP values instead of experimental retention times. Therefore a stepwise multiple regression analysis was performed using JMP 9.0.0. $\log(1/EC_{50})$ was modeled against a total of 286 descriptors (see Section S7.4, only calculated logP values included, not retention times), the k-fold cross validation was set to 2 and all possible models with a maximum of three terms were calculated (3,899,181 models in total) and subsequently ranked according to best fit (highest R). The results are shown in Figure S247 (left). Once again, the best model with only one term is observed for logP and the second best models with logS. However, logP and logS are highly correlated descriptors and are both measures for lipophilicity. In the subsequent 2-term and 3-term models logS is preferred over logP. As we are more interested in the effect of changing retention time to logP, the analysis was repeated but the logS

descriptors were excluded from the stepwise regression (total of 281 descriptors and 3,737,863 models). The results of the best models are given in Figure S247 (right).

All Possible Models

Ordered up to best 20 models up to 3 terms per model

Model	Number	RSquare	RMSE	AICc	BIC
logP	1	0,7890	0,2814	11,0299	11,9868
logS (pH1,7)	1	0,6979	0,3367	17,4930	18,4498
logS (pH6,5)	1	0,6957	0,3380	17,6223	18,5792
logS (pH7,4)	1	0,6957	0,3380	17,6223	18,5792
logS (pH8)	1	0,6957	0,3380	17,6223	18,5792
logD (pH1,7)	1	0,6717	0,3511	18,9907	19,9475
Vd	1	0,6644	0,3549	19,3858	20,3427
logD (pH7,4)	1	0,6415	0,3668	20,5721	21,5289
logD (pH8)	1	0,6415	0,3668	20,5721	21,5289
logD (pH4,6)	1	0,6409	0,3671	20,6020	21,5588
BLTD48	1	0,5070	0,4301	26,3053	27,2621
BLTA96	1	0,5059	0,4307	26,3475	27,3043
BLTF96	1	0,5035	0,4317	26,4340	27,3909
PI	1	0,3776	0,4833	30,5012	31,4580
J3D	1	0,3456	0,4956	31,4035	32,3603
E3u	1	0,2702	0,5234	33,3673	34,3241
G3m	1	0,2601	0,5270	33,6153	34,5721
PHI	1	0,2365	0,5353	34,1802	35,1371
S2K	1	0,2360	0,5355	34,1911	35,1479
Hy	1	0,2196	0,5412	34,5753	35,5321
logS (pH1,7),Gs	2	0,8782	0,2208	4,5015	4,9861
logS (pH6,5),Gs	2	0,8771	0,2218	4,6640	5,1486
logS (pH7,4),Gs	2	0,8771	0,2218	4,6640	5,1486
logS (pH8),Gs	2	0,8771	0,2218	4,6640	5,1486
PI,pKa (alkyl NH)	2	0,8602	0,2366	6,9822	7,4668
logS (pH1,7),Jhetm	2	0,8564	0,2398	7,4663	7,9508
logS (pH1,7),JhetZ	2	0,8563	0,2399	7,4794	7,9639
logS (pH6,5),Jhetm	2	0,8558	0,2403	7,5473	8,0319
logS (pH7,4),Jhetm	2	0,8558	0,2403	7,5473	8,0319
logS (pH8),Jhetm	2	0,8558	0,2403	7,5473	8,0319
logS (pH6,5),JhetZ	2	0,8556	0,2404	7,5613	8,0459
logS (pH7,4),JhetZ	2	0,8556	0,2404	7,5613	8,0459
logS (pH8),JhetZ	2	0,8556	0,2404	7,5613	8,0459
logS (pH1,7),G1	2	0,8523	0,2432	7,9768	8,4613
logP,E2m	2	0,8516	0,2438	8,0622	8,5468
logS (pH6,5),G1	2	0,8506	0,2446	8,1833	8,6679
logS (pH7,4),G1	2	0,8506	0,2446	8,1833	8,6679
logS (pH8),G1	2	0,8506	0,2446	8,1833	8,6679
logS (pH1,7),G2s	2	0,8435	0,2503	9,0180	9,5025
logS (pH6,5),G2s	2	0,8433	0,2505	9,0379	9,5225
PI,Index of refraction,pKa (alkyl NH)	3	0,9304	0,1728	-1,6396	-2,1877
PI,pKa (alkyl NH),ZM2V	3	0,9257	0,1785	-0,4764	-1,0246
logS (pH6,5),H3D,SEig	3	0,9246	0,1798	-0,2106	-0,7587
logS (pH7,4),H3D,SEig	3	0,9246	0,1798	-0,2106	-0,7587
logS (pH8),H3D,SEig	3	0,9246	0,1798	-0,2106	-0,7587
logS (pH1,7),H3D,SEig	3	0,9244	0,1801	-0,1570	-0,7051
PI,pKa (alkyl NH),BLI	3	0,9239	0,1806	-0,0462	-0,5943
logP,SCBO,RGyr	3	0,9222	0,1826	0,3467	-0,2014
PI,pKa (alkyl NH),GMTIV	3	0,9200	0,1853	0,8613	0,3132
PI,pKa (alkyl NH),L1s	3	0,9199	0,1854	0,8840	0,3358
Molecular volume,logS (pH1,7),E3p	3	0,9197	0,1855	0,9167	0,3685
logS (pH1,7),BLI,Gs	3	0,9191	0,1863	1,0629	0,5148
Molecular volume,logS (pH6,5),E3p	3	0,9190	0,1864	1,0778	0,5296
Molecular volume,logS (pH7,4),E3p	3	0,9190	0,1864	1,0778	0,5296
Molecular volume,logS (pH8),E3p	3	0,9190	0,1864	1,0778	0,5296
PI,pKa (alkyl NH),Ts	3	0,9184	0,1871	1,2203	0,6721
PI,pKa (alkyl NH),ZM1V	3	0,9175	0,1881	1,4039	0,8557
logS (pH6,5),BLI,Gs	3	0,9173	0,1883	1,4508	0,9027
logS (pH7,4),BLI,Gs	3	0,9173	0,1883	1,4508	0,9027
logS (pH8),BLI,Gs	3	0,9173	0,1883	1,4508	0,9027

All Possible Models

Ordered up to best 20 models up to 3 terms per model

Model	Number	RSquare	RMSE	AICc	BIC
logP	1	0,7890	0,2814	11,0299	11,9868
logD (pH1,7)	1	0,6717	0,3511	18,9907	19,9475
Vd	1	0,6644	0,3549	19,3858	20,3427
logD (pH7,4)	1	0,6415	0,3668	20,5721	21,5289
logD (pH8)	1	0,6415	0,3668	20,5721	21,5289
logD (pH4,6)	1	0,6409	0,3671	20,6020	21,5588
BLTD48	1	0,5070	0,4301	26,3053	27,2621
BLTA96	1	0,5059	0,4307	26,3475	27,3043
BLTF96	1	0,5035	0,4317	26,4340	27,3909
PI	1	0,3776	0,4833	30,5012	31,4580
J3D	1	0,3456	0,4956	31,4035	32,3603
E3u	1	0,2702	0,5234	33,3673	34,3241
G3m	1	0,2601	0,5270	33,6153	34,5721
PHI	1	0,2365	0,5353	34,1802	35,1371
S2K	1	0,2360	0,5355	34,1911	35,1479
Hy	1	0,2196	0,5412	34,5753	35,5321
L3u	1	0,2186	0,5416	34,5971	35,5539
H3D	1	0,2154	0,5427	34,6716	35,6285
E3p	1	0,2146	0,5429	34,6885	35,6453
E3v	1	0,2119	0,5439	34,7501	35,7069
PI,pKa (alkyl NH)	2	0,8602	0,2366	6,9822	7,4668
logP,E2m	2	0,8516	0,2438	8,0622	8,5468
logP,PSA	2	0,8432	0,2506	9,0553	9,5399
logP,Ui	2	0,8377	0,2549	9,6734	10,1580
logP,TPSA(NO)	2	0,8325	0,2590	10,2400	10,7246
PI,pKa (arom NH)	2	0,8308	0,2603	10,4169	10,9015
logP,pKa (arom NH)	2	0,8285	0,2621	10,6677	11,1523
logP,pKa2 (NH)	2	0,8278	0,2626	10,7340	11,2186
Hammett,logP	2	0,8246	0,2650	11,0711	11,5556
logP,pKa (alkyl NH)	2	0,8244	0,2651	11,0880	11,5726
logP,PJI2	2	0,8227	0,2664	11,2630	11,7476
logP,Vsmin	2	0,8186	0,2695	11,6727	12,1573
logP,BLTF96	2	0,8185	0,2696	11,6853	12,1698
logP,BLTA96	2	0,8178	0,2700	11,7476	12,2321
logP,BLTD48	2	0,8175	0,2703	11,7824	12,2670
logP,Vsmax	2	0,8166	0,2710	11,8744	12,3590
logP,De	2	0,8154	0,2718	11,9845	12,4690
logP,L/Bw	2	0,8147	0,2724	12,0575	12,5420
logP,E1e	2	0,8136	0,2732	12,1603	12,6449
logP,G2s	2	0,8132	0,2735	12,1983	12,6829
PI,Index of refraction,pKa (alkyl NH)	3	0,9304	0,1728	-1,6396	-2,1877
PI,pKa (alkyl NH),ZM2V	3	0,9257	0,1785	-0,4764	-1,0246
PI,pKa (alkyl NH),BLI	3	0,9239	0,1806	-0,0462	-0,5943
logP,SCBO,RGyr	3	0,9222	0,1826	0,3467	-0,2014
PI,pKa (alkyl NH),GMTIV	3	0,9200	0,1853	0,8613	0,3132
PI,pKa (alkyl NH),L1s	3	0,9199	0,1854	0,8840	0,3358
PI,pKa (alkyl NH),Ts	3	0,9184	0,1871	1,2203	0,6721
PI,pKa (alkyl NH),ZM1V	3	0,9175	0,1881	1,4039	0,8557
logP,Vsmin,E2m	3	0,9167	0,1890	1,5904	1,0423
PI,pKa (alkyl NH),nHAcc	3	0,9164	0,1894	1,6534	1,1053
PI,pKa (alkyl NH),SMTIV	3	0,9159	0,1899	1,7493	1,2011
logP,RGyr,Ui	3	0,9154	0,1904	1,8559	1,3078
PI,pKa (alkyl NH),GGI4	3	0,9146	0,1913	2,0257	1,4775
PI,pKa (alkyl NH),Ss	3	0,9140	0,1921	2,1605	1,6123
logP,pKa (arom NH),Tm	3	0,9139	0,1922	2,1884	1,6403
logP,pKa (arom NH),L1m	3	0,9126	0,1937	2,4589	1,9107
PI,pKa (arom NH),Ks	3	0,9121	0,1942	2,5550	2,0069
logP,SPAN,Ui	3	0,9113	0,1951	2,7194	2,1713
logP,E2m,TPSA(NO)	3	0,9109	0,1955	2,8067	2,2586
logP,SCBO,S1K	3	0,9108	0,1956	2,8213	2,2731

Figure S247. Overview of best 20 models with up to 3 term correlating $\log(1/EC_{50})$ with various descriptors (left) logS included in the analysis. (right) logS values omitted from the analysis.

In the previous table the combinations containing PI+pKa can be discarded because these descriptors are highly correlated to each other and to the logP value (see correlation matrix in Figure S248). The models that are left are comparable to the ones obtained for the retention time and consist of logP + pKa (arom NH) + size (Tm, L1m). This shows that logP and retention time can be interchanged and the same models will be

built using logP as were built using retention time. The final models were calculated using a standard least-square fitting with JMP 9.0.0 for log(1/EC_{50}) versus the retention time, pKa (or Hammett or $V_{S,max}$) and SPAN. The results are summarized in Figure S249.

Correlations

	logP	PI	pKa (alkyl NH)	pKa (arom NH)
logP	1,0000	-0,7580	0,2853	0,3963
PI	-0,7580	1,0000	-0,7989	-0,8511
pKa (alkyl NH)	0,2853	-0,7989	1,0000	0,9287
pKa (arom NH)	0,3963	-0,8511	0,9287	1,0000

Figure S248. Correlation matrix with logP.

Log(1/EC₅₀) = A*logP + B*pKₐ + C*SPAN + D

Summary of Fit

RSquare	0,896893
RSquare Adj	0,874799
Root Mean Square Error	0,210306
Mean of Response	-0,04426
Observations (or Sum Wgts)	18

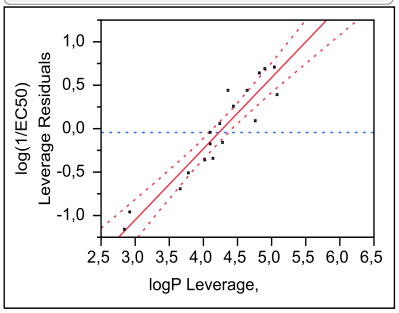
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	5,3861962	1,79540	40,5937
Error	14	0,6191987	0,04423	Prob > F
C. Total	17	6,0053949		<,0001 *

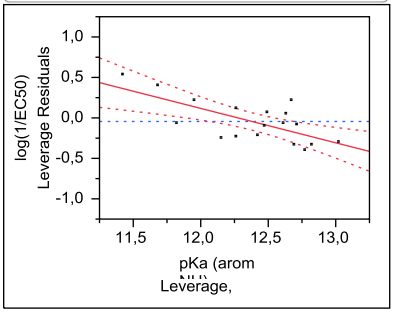
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	4,5822556	1,90928	2,40	0,0309 *
logP	0,8196256	0,079988	10,25	<,0001 *
pKa (arom NH)	-0,425023	0,118133	-3,60	0,0029 *
SPAN	-0,280964	0,092167	-3,05	0,0087 *

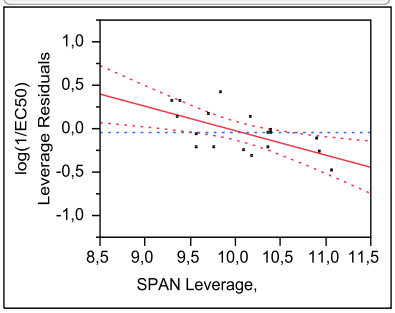
Leverage Plot



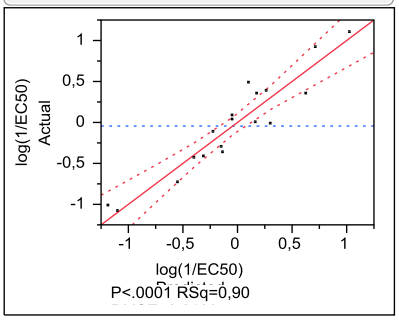
Leverage Plot



Leverage Plot



Actual by Predicted Plot



Scaled Estimates

Continuous factors centered by mean, scaled by

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0,044264	0,04957	-0,89	0,3870
logP	1,4343449	0,139979	10,25	<,0001 *
pKa (arom NH)	-0,359144	0,099822	-3,60	0,0029 *
SPAN	-0,373823	0,122629	-3,05	0,0087 *

Log(1/EC₅₀) = A*logP + B*Hammett + C*SPAN + D **Eq. 8**

Summary of Fit

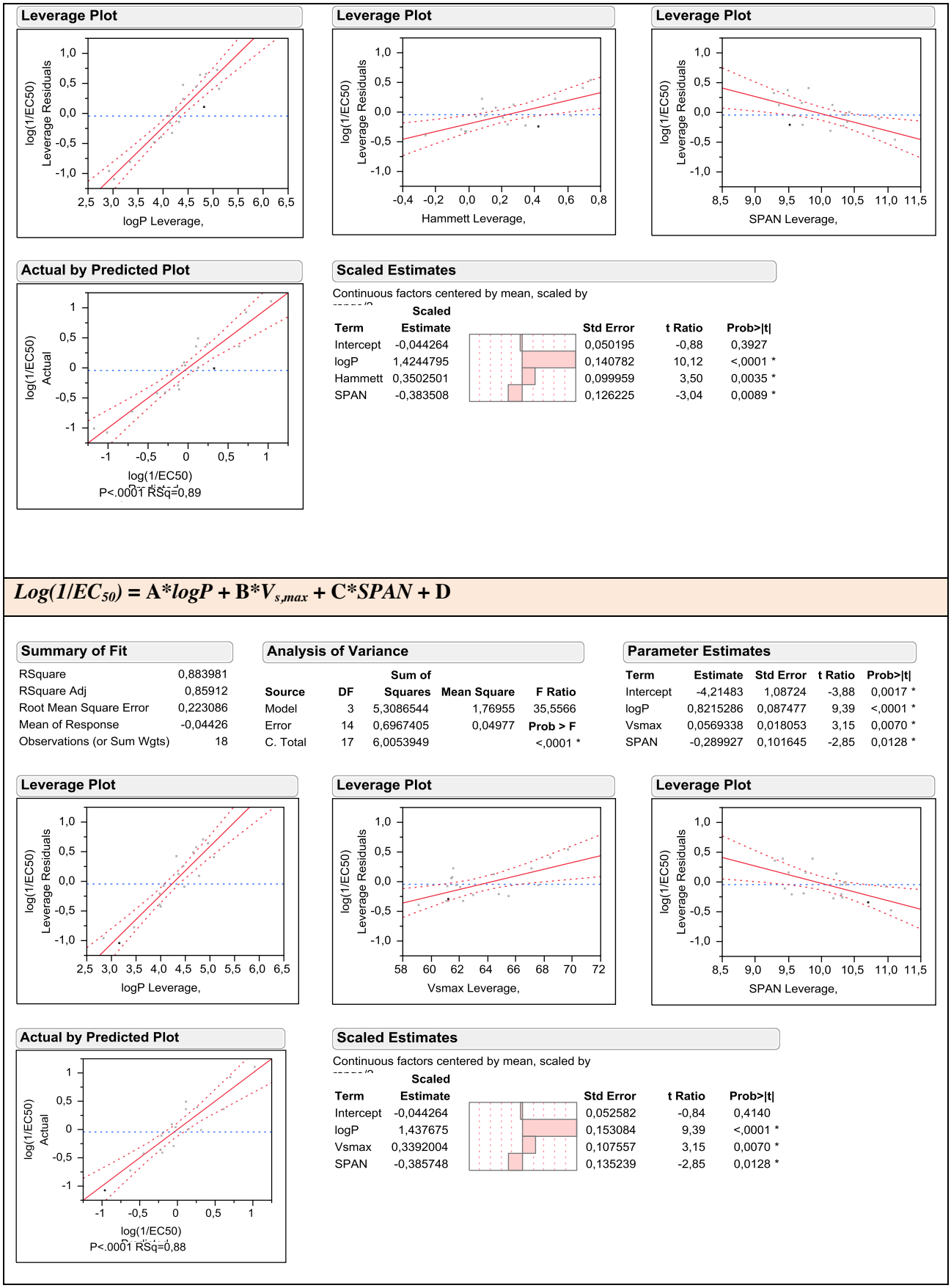
RSquare	0,894276
RSquare Adj	0,871621
Root Mean Square Error	0,212957
Mean of Response	-0,04426
Observations (or Sum Wgts)	18

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	5,3704834	1,79016	39,4736
Error	14	0,6349115	0,04535	Prob > F
C. Total	17	6,0053949		<,0001 *

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0,734393	0,788642	-0,93	0,3675
logP	0,8139883	0,080447	10,12	<,0001 *
Hammett	0,654673	0,186839	3,50	0,0035 *
SPAN	-0,288243	0,09487	-3,04	0,0089 *



$Log(1/EC_{50}) = A*logP + B*V_{s,max} + C*SPAN + D$

Summary of Fit

RSquare	0,883981
RSquare Adj	0,85912
Root Mean Square Error	0,223086
Mean of Response	-0,04426
Observations (or Sum Wgts)	18

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	5,3086544	1,76955	35,5566
Error	14	0,6967405	0,04977	Prob > F
C. Total	17	6,0053949		<,0001 *

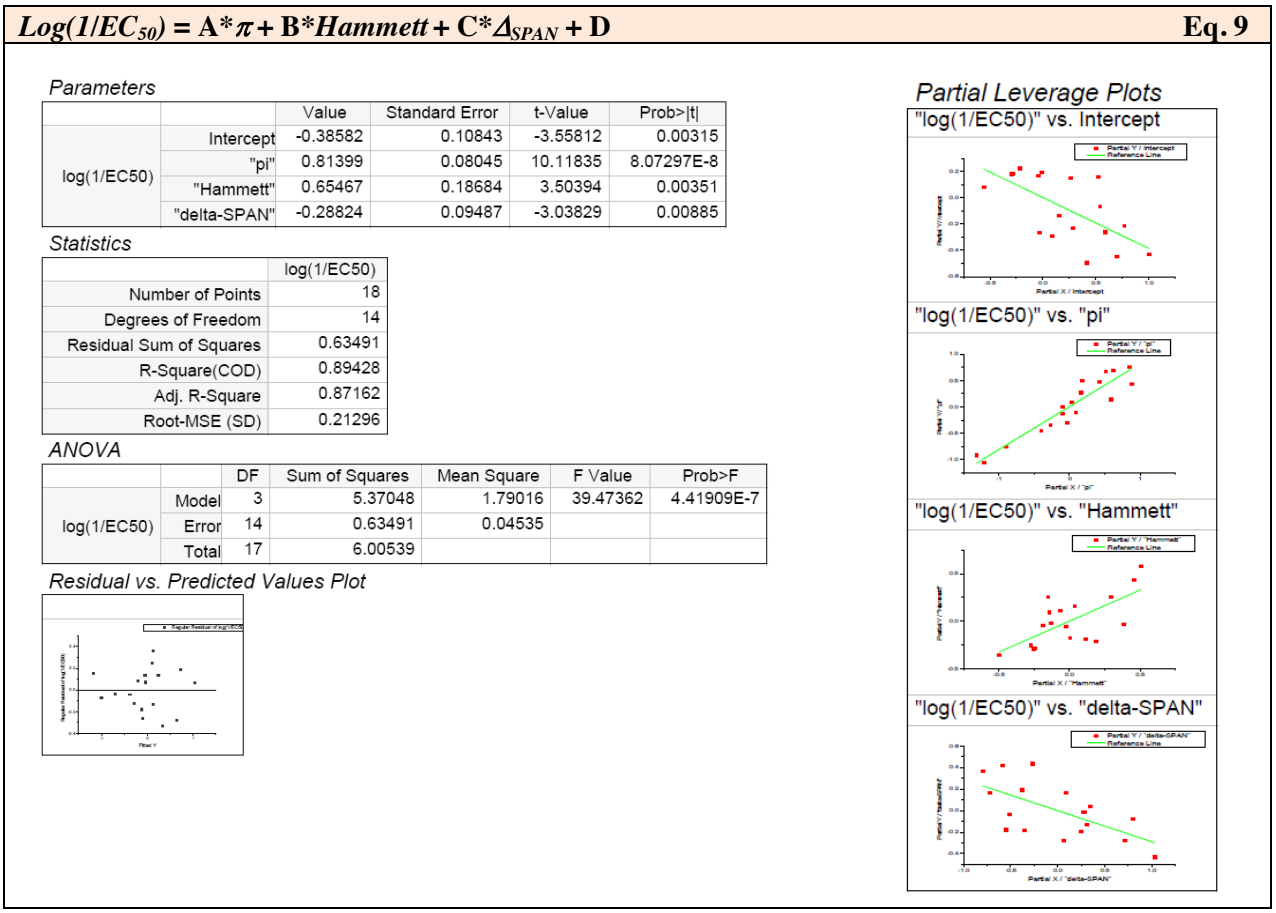
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-4,21483	1,08724	-3,88	0,0017 *
logP	0,8215286	0,087477	9,39	<,0001 *
Vsmax	0,0569338	0,018053	3,15	0,0070 *
SPAN	-0,289927	0,101645	-2,85	0,0128 *

Figure S249. Overview of models correlating log(1/EC_{50}) withlogP, anion binding (Hammett, pKa or $V_{S,max}$) and SPAN.

S7.7 QSAR Modelling Using Relative Values

The Hammett constant is defined as a relative value against an unsubstituted benzene ring. It is also possible to define the lipophilicity of a given substituent (π) by taking the logP value relative to the unsubstituted compound (π = logP(x)-logP(**9**, -H)). Similarly we can define Δ_{SPAN} as the value of the SPAN descriptor relative to the parent, unsubstituted compound (Δ_{SPAN} = SPAN(x)-SPAN(**9**, -H)). Origin 8.1 was used to perform a standard least-squares multiple linear regression for log(1/EC_{50}) versus the relative descriptors π, Hammett and Δ_{SPAN}. The results are shown in Figure S250.



		Value	Standard Error	t-Value	Prob>|t|
log(1/EC50)	Intercept	-0.38582	0.10843	-3.55812	0.00315
	"pi"	0.81399	0.08045	10.11835	8.07297E-8
	"Hammett"	0.65467	0.18684	3.50394	0.00351
	"delta-SPAN"	-0.28824	0.09487	-3.03829	0.00885

	log(1/EC50)
Number of Points	18
Degrees of Freedom	14
Residual Sum of Squares	0.63491
R-Square(COD)	0.89428
Adj. R-Square	0.87162
Root-MSE (SD)	0.21296

		DF	Sum of Squares	Mean Square	F Value	Prob>F
log(1/EC50)	Model	3	5.37048	1.79016	39.47362	4.41909E-7
	Error	14	0.63491	0.04535		
	Total	17	6.00539			

Figure S250. Overview of model correlating log(1/EC_{50}) with π, Hammett and Δ_{SPAN}.

In principle when all of the descriptors are defined as a relative value against parent compound **9** (unsubstituted), the intercept of the model should correspond to the experimentally observed log(1/EC_{50}) value of the parent compound. Indeed, the optimized intercept from figure S250 (-0.38±0.10) corresponds well to the experimental log(1/EC_{50}) value for **9** (-0.43±0.10). We can reduce the amount of optimized parameters and alter the degrees of freedom by fixing the intercept to this experimental value. Origin 8.1 was used to obtain the multiple linear regressed model with the intercept fixed at -0.43, as shown in Figure S251

$$Log(1/EC_{50}) = A*\pi + B*Hammett + C*\Delta_{SPAN} - 0.43 \qquad \textbf{Eq. 10}$$

Parameters

		Value	Standard Error	t-Value	Prob>|t|
log(1/EC50)	Intercept	-0.43	--	--	--
	"pi"	0.81632	0.07798	10.46846	2.72693E-8
	"Hammett"	0.65837	0.18136	3.63023	0.00247
	"delta-SPAN"	-0.26395	0.0717	-3.68112	0.00222

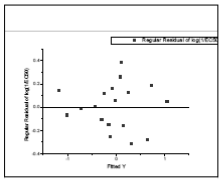
Statistics

	log(1/EC50)
Number of Points	18
Degrees of Freedom	15
Residual Sum of Squares	0.64244
R-Square(COD)	0.89365
Adj. R-Square	0.87238
Root-MSE (SD)	0.20695

ANOVA

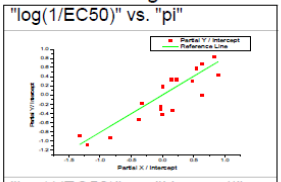
		DF	Sum of Squares	Mean Square	F Value	Prob>F
log(1/EC50)	Model	3	5.39822	1.79941	42.0135	1.54875E-7
	Error	15	0.64244	0.04283		
	Total	18	6.04066			

Residual vs. Predicted Values Plot

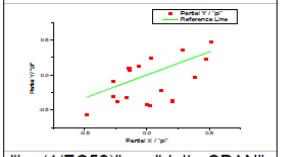


Partial Leverage Plots

"log(1/EC50)" vs. "pi"



"log(1/EC50)" vs. "Hammett"



"log(1/EC50)" vs. "delta-SPAN"

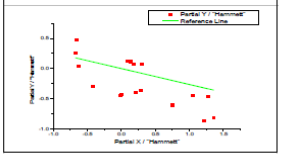


Figure S251. Overview of model correlating $\log(1/EC_{50})$ with π, Hammett and Δ_{SPAN} and intercept fixed at -0.43.

S8. PREDICTING ANION BINDING AND TRANSPORT

The best way to check the credibility of a QSAR model is to investigate its predictive power. For this reason, four receptors were initially excluded from the training set and used to test the predictability (they form the test set). The anion binding and the anion transport ability of the test set were assessed in the same way as the training set and an overview is given in Table S49.

Table S49. Overview of anion transport ($\log(1/EC_{50})$) and anion binding ($\log K_a$ in M^{-1}) of the test set.

Compound	$\log(1/EC_{50})$	$\log K_a(Cl^-)$	$\log K_a(HCO_3^-)$	$\log K_a(H_2PO_4^-)$
1 (Br)	0.03 (±0.09)	1.27	2.65	2.67
6 (COMe)	-0.4832 (±0.0004)	1.43	2.74	2.62
14 (OEt)	-0.28 (±0.11)	1.09	2.15	2.04
20 (Propyl)	0.88 (±0.23)	1.12	2.27	2.20

In order to predict the transport and binding ability, the molecular descriptors used to built equations 1-10 were calculated (including the retention times on a reversed-phase HPLC column, which were obtained experimentally after synthesis) and are shown in Table S50. With these descriptors, the anion transport and anion binding ability were calculated using equations 1-10 and these values are shown in Table S51.

Table S50. Overview of the descriptors of the test set used for predicting anion transport and binding abilities.

Compound	RT [a]	logP [b]	Hammett [c]	SPAN [d]	π [e]	Δ_{SPAN} [f]
1 (Br)	12.96	4.691	0.23	10.615	1.165	1.867
6 (COMe)	12.01	3.516	0.50	9.877	-0.010	1.129
14 (OEt)	12.51	4.158	-0.24	10.006	0.632	1.258
20 (Propyl)	13.49	5.083	-0.13	9.983	1.557	1.235

[a] RT: retention time in min on reversed phase HPLC column; [b] logP calculated using Daylight v4.73; [c] Hammett constants for substituents in *para* position; [d] SPAN calculated using e-Dragon; [e] π = logP(x)-3.526: [f] Δ_{SPAN} = SPAN(x) – 8.748.

Table S51. Overview of the anion transport and binding abilities of the test set as calculated using Eq. 1-10.

	Anion binding ($\log K_a$)			Anion transport ($\log(1/EC_{50})$)					
Compound	Eq1	Eq2	Eq3	Eq5	Eq6	Eq7	Eq8	Eq9	Eq10
1 (Br)	1.30	2.57	2.60	0.22	0.31	0.002	0.14	0.17	0.19
6 (COMe)	1.44	2.80	2.84	-0.49	-0.42	-0.53	-0.42	-0.39	-0.40
14 (OEt)	1.04	2.18	2.19	-0.12	-0.02	-0.46	-0.42	-0.39	-0.40
20 (Propyl)	1.10	2.27	2.28	0.62	0.55	0.52	0.41	0.44	0.44

A good way to visualize the predictability of a model is to plot the actual versus the predicted values. These plots are given in Figures S252-S260 for all of the obtained QSAR models (equations 1-10). A linear fit of these plots was also performed using Origin 8.1, the R values of which provide a quantitative measure for predictability.

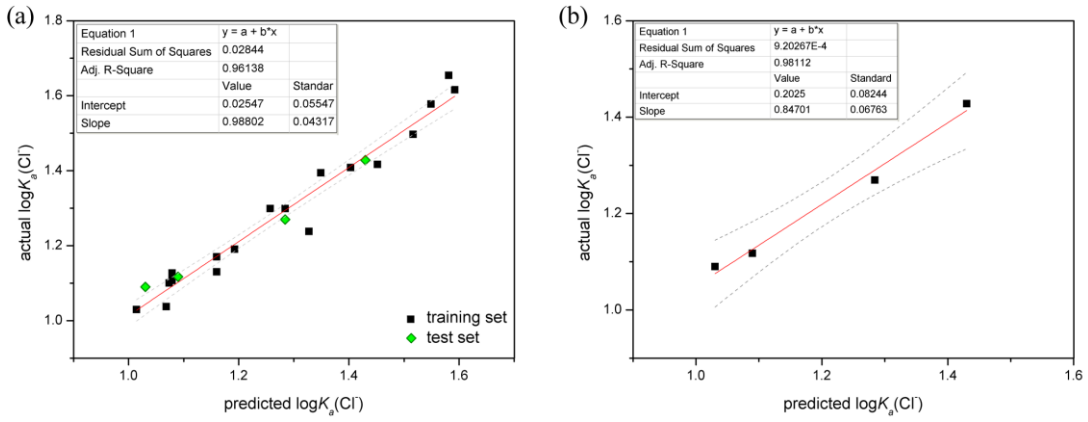


Figure S252. Actual vs. predicted plots for the anion binding constants (logK_a) for chloride obtained with equation 1: logK_a(Cl$^-$) = 0.55(±0.03)*σ_p + 1.17(±0.01). (a) Training set + test set. (b) Test set only.

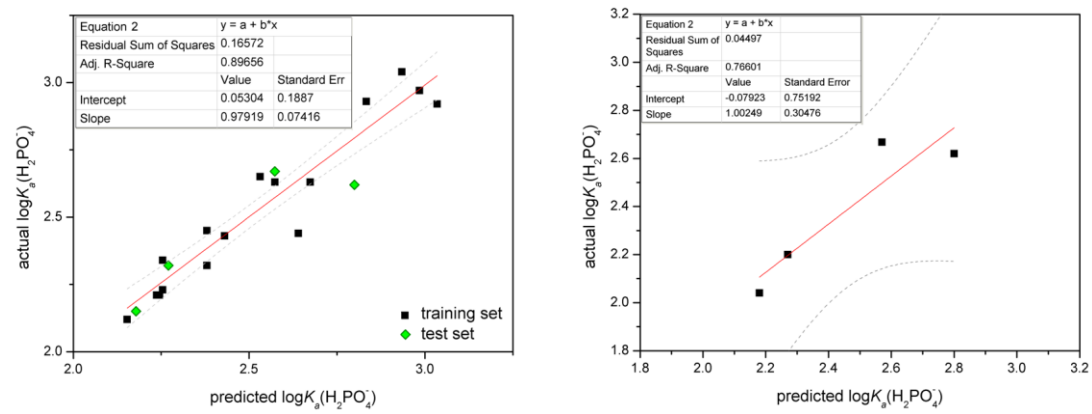


Figure S253. Actual vs. predicted plots for the anion binding constants (logK_a) for dihydrogen phosphate obtained with equation 2: logK_a(H$_2$PO$_4^-$) = 0.84(±0.06)*σ_p + 2.38(±0.02). (a) Training set + test set. (b) Test set only.

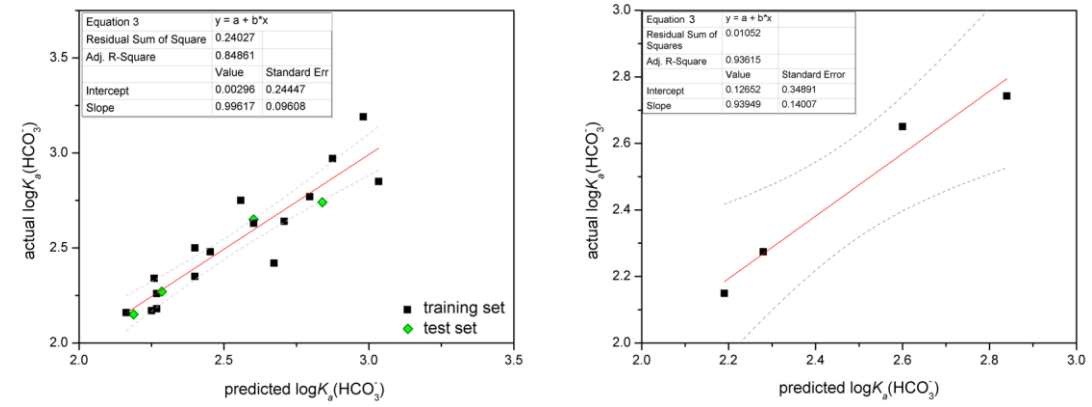


Figure S254. Actual vs. predicted plots for the anion binding constants (logK_a) for bicarbonate obtained with equation 3: logK_a(HCO$_3^-$) = 0.88(±0.10)*σ_p + 2.40(±0.04). (a) Training set + test set. (b) Test set only.

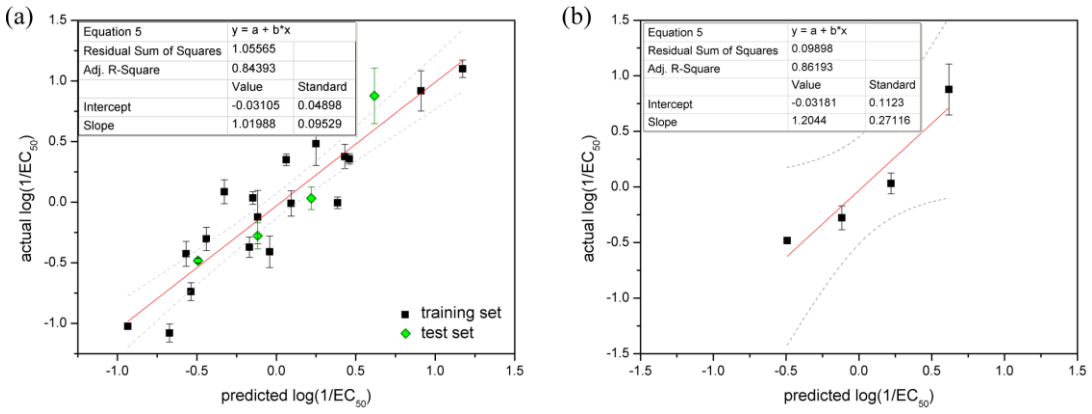


Figure S255. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 5 (anion transport): $\log(1/EC_{50}) = 0.75(\pm0.08)*RT - 9.5(\pm1.0)$. (a) Training set + test set. (b) Test set only.

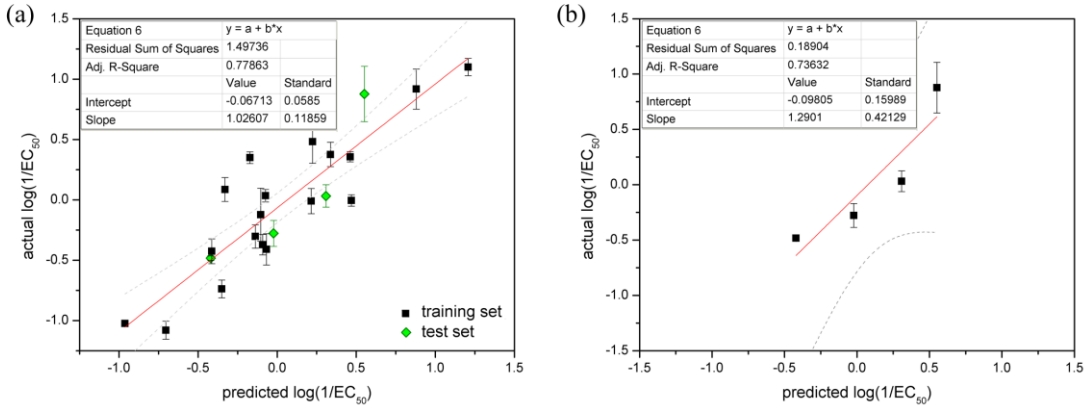


Figure S256. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 6 (anion transport): $\log(1/EC_{50}) = 0.62(\pm0.08)*logP - 2.6(\pm0.3)$. (a) Training set + test set. (b) Test set only.

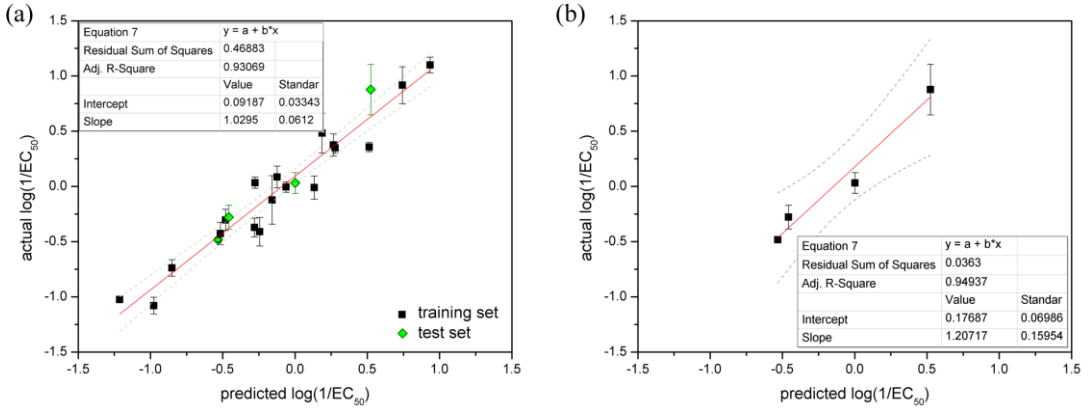


Figure S257. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 7 (anion transport): $\log(1/EC_{50}) = 0.94(\pm0.07)*RT + 0.48(\pm0.14)*\sigma_p - 0.31(\pm0.07)*SPAN - 9.0(\pm0.8)$. (a) Training set + test set. (b) Test set only.

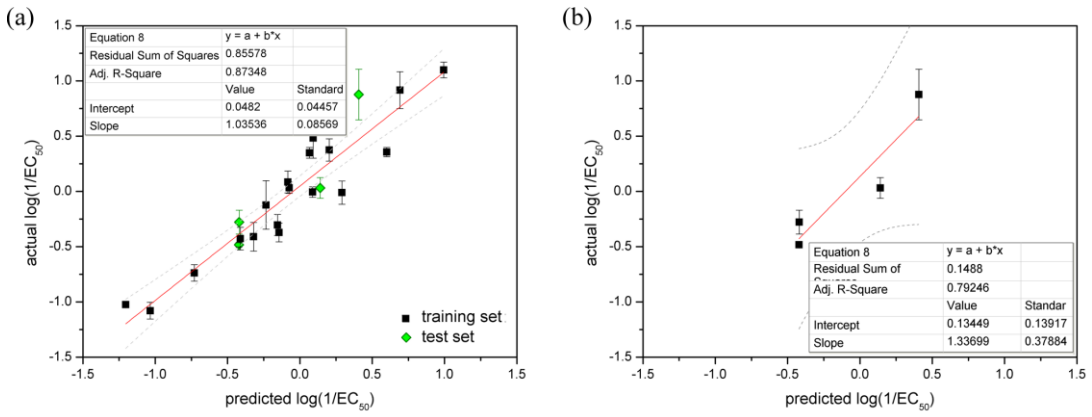


Figure S258. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 8 (anion transport): $\log(1/EC_{50}) = 0.81(\pm0.08)*logP + 0.65(\pm0.19)*\sigma_p - 0.29(\pm0.09)*SPAN - 0.73(\pm0.79)$. (a) Training set + test set. (b) Test set only.

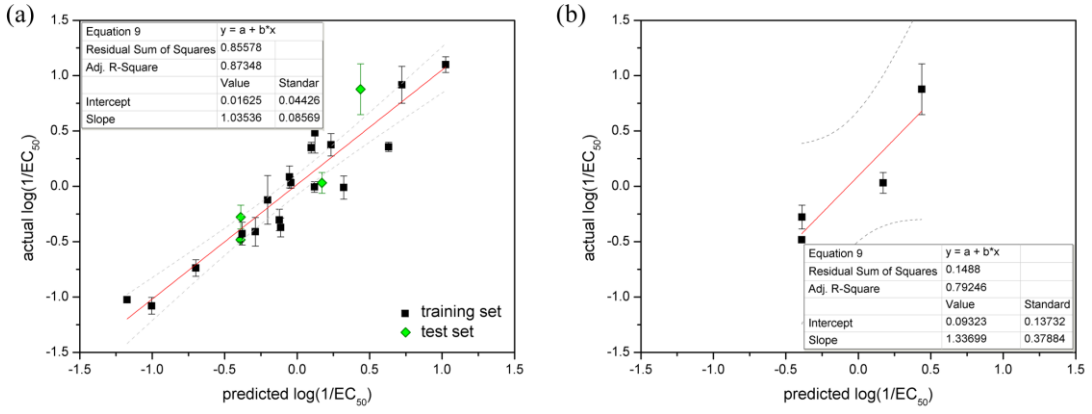


Figure S259. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 9 (anion transport): $\log(1/EC_{50}) = 0.81(\pm0.08)*\pi + 0.65(\pm0.19)*\sigma_p - 0.29(\pm0.09)*\Delta_{SPAN} - 0.38(\pm0.11)$. (a) Training set + test set. (b) Test set only.

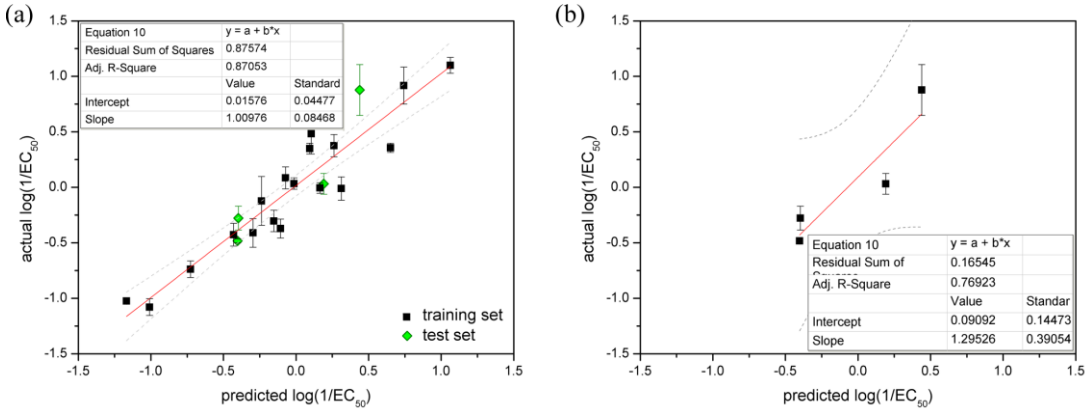


Figure S260. Actual vs. predicted plots for the $\log(1/EC_{50})$ values obtained with equation 10 (anion transport): $\log(1/EC_{50}) = 0.82(\pm0.08)*\pi + 0.66(\pm0.18)*\sigma_p - 0.26(\pm0.07)*\Delta_{SPAN} - 0.43$. (a) Training set + test set. (b) Test set only.

S9. REFERENCES AND NOTES

(1) S.J. Moore, M. Wenzel, M.E. Light, R. Morley, S.J. Bradberry, P. Gomez-Iglesias, V. Soto-Cerrato, R Perez-Tomas and P.A. Gale, *Chem. Sci.* 2012, **3**, 2501.

(2) M. Wenzel, M.E. Light, A.P. Davis, and P.A. Gale, *Chem. Commun.* 2011, **47**, 7641.

(3) R.F. Hunter and C. Soyka, *J. Chem. Soc.* 1926, 2958.

(4) L. Doub, L.M. Richardson, D.R. Herbst, M.L. Black, O.L. Stevenson, L.L. Bambas, G.P. Youmans and A.S. Youmans, *J. Am. Chem. Soc.* 1958, **80**, 2205.

(5) T. Hayashita, T. Onodera, R. Kato, S. Nishizawa and N. Teramae, *Chem. Commun.* 2000, **9**, 755.

(6) M. Lipp, F. Dallacker and I.M. zu Kocker, *Monatsh. Chem.* 1959, **90**, 41.

(7) G.M. Sheldrick, *Acta Cryst.* 1990, **A46,** 467.

(8) G.M. Sheldrick, University of Göttingen, Germany, **1997**.

(9) M.J. Hynes, *J. Chem. Soc., Dalton Trans.* 1993, 311.

(10) C. Hansch, A. Leo, and R.W. Taft, *Chem. Rev.* **1991**, *91*, 165.

(11) *ACD/I-Lab 2.0*; https://ilab.acdlabs.com/iLab2/ (accessed August 9, 2012); ACD/Labs 2010-2013.

(12) *AMBER 12*; University of California: San Francisco, USA, 2012.

(13) J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman and D.A. Case, *J. Comput. Chem.* 2004, **25**, 1157.

(14) J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman and D.A. Case, *J. Comput. Chem.* 2005, **26**, 114.

(15) A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, *J. Comput. Chem.* 2000, **21**, 132.

(16) A. Jakalian, D.B. Jack and C.I. Bayly, *J. Comput. Chem.* 2002, **23**, 1623.

(17) I.S. Joung and T.E. Cheatham, *J. Phys. Chem. B* 2008, **112**, 9020.

(18) W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.* 1983, **79**, 926.

(19) *Gaussian 09*, Revision **A.1**; Gaussian Inc.: Wallingford, CT, USA, 2009.

(20) F.A. Bulat, A. Toro-Labbe, WFA: A suite of programs to analyse wavefunctions (unpublished).

(21) F.A. Bulat, A. Toro-Labbe, T. Brinck, J.S. Murray, P. Politzer, *J. Mol. Model* 2010, **16**, 1679.

(22) J.S. Murray, K.E. Riley, P. Politzer, and T. Clark, *Aust. J. Chem.* 2010, **63**, 1598.

(23) Y.G. Ma, K.C. Gross, C.A. Hollingsworth, P.G. Seybold and J.S. Murray, *J. Mol. Model* 2004, **10**, 235.

(24) T. Braumann, *J. Chromatogr. A* 1986, **373**, 191.

(25) VCCLAB, Virtual Computational Chemistry Laboratory, ALOGPS 2.1 and e-Dragon 1.0; http://www.vcclab.org (accessed January 14, 2013); VCCLAB 2005.

(26) *Daylight ClogP*, version 4.73; Daylight Chemical Information Systems Inc.: Laguna Niguel, CA, USA, 2001.

(27) *Accelrys Diamond Descriptors (ALogP)*, version 1.5; Accelrys software Inc.: San Diego, CA, USA.

(28) *Fieldview*, version 2.0.2; Cresset BMD Ltd.: Hertfordshire, UK, 2011.

(29) *Chemicalize*; http://www.chemicalize.org/ (accessed August 9, 2012); ChemAxon.