Chiral bicyclic [2.2.2] octadiene ligands for Rh-catalysed catalytic asymmetric conjugate additions to acyclic enones: A quantitative structure-property relationship

Yunfei Luo, Neil G. Berry and Andrew J. Carnell

1. General

All anaerobic and moisture-sensitive manipulations were carried out with standard Schlenk techniques under nitrogen. NMR spectra were recorded on a Bruker AMX-400 MHz spectrometer. Chemical shifts are reported in δ ppm referenced to an internal $SiMe_4$ standard for ^1H NMR and chloroform-*d* (δ 77.05) for ^{13}C NMR. Enantiomeric excess was measured by normal phase HPLC on a Waters 2695 separation module equipped with a Waters 996 photo diode array detector. Separations were carried out using Chiracel AD, AD-H, AS-H and OD-H chiral column provided by Diacel company (columns and conditions under each compound later). The optical rotation data were recorded on Perkin Elmer Polarimeter 343 Plus.

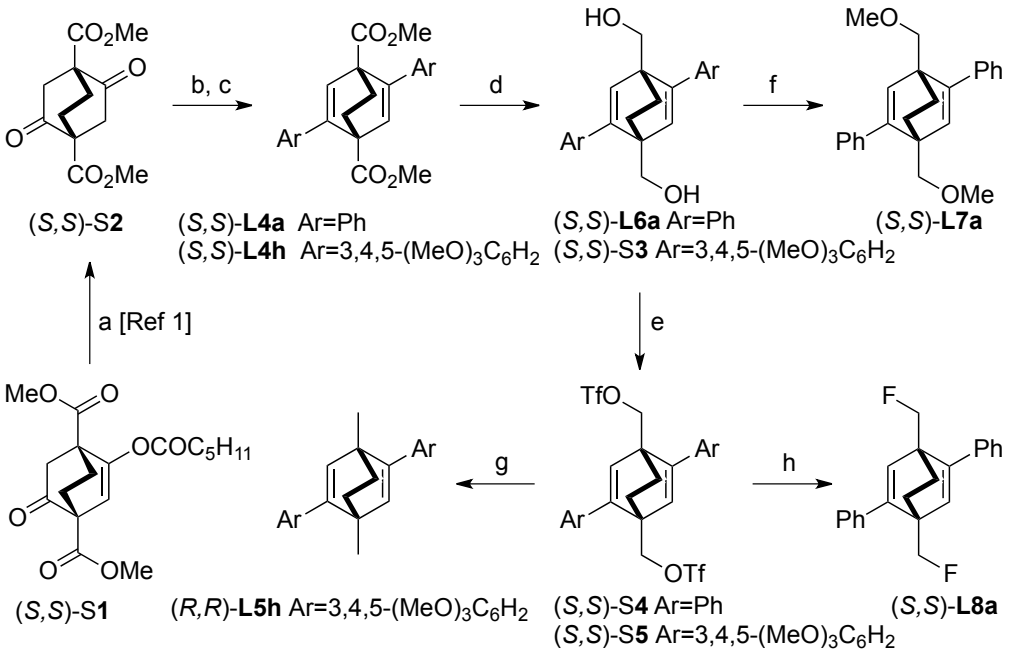
All solvents and reagents were used without further purification if not specified in the procedures. Ligands **4a-g**, **5a**, **5f, 6a** and **9a** were made according to previously published procedures.[1,2]

2. Rh-diene-catalyzed asymmetric conjugate additions reactions

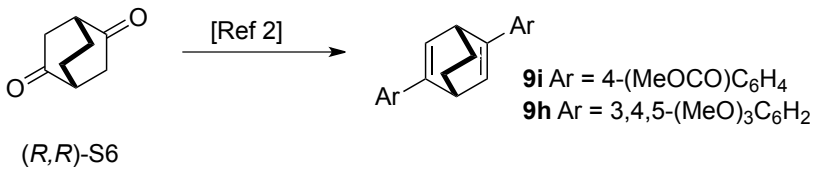
These reactions were carried out according to a modification of the published procedure[2]:

To a reaction tube was added $[RhCl(C_2H_4)_2]_2$ (1.8mg, 9µmol Rh), 10µmol chiral diene (1.1eq) and dioxane (0.3mL). This mixture was stirred at room temperature for 30min and then extra dioxane (0.7mL), phenylboronic acid (72mg, 0.6mmol) and 3-nonen-2-one (42mg, 0.3mmol) were added. To the resulting mixture KOH solution (0.1mL, 1.5M) was added in one portion. The reaction was stirred at room temperature for 3hrs and then filtered through a short silica pad and washed with diethyl ether. After removal of solvents and the product was purified by using preparative TLC (Hexane : EtOAc 10: 1) to give the pure product.

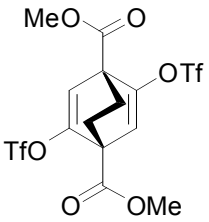
3. Ligand Synthesis



Scheme S1: Preparation of 1,4-disubstututed bicyclo[2.2.2]ligands from (S,S)-enol ester S1. a) Na_2CO_3, MeOH (99%); b) LHMDS, Tf_2O, THF, -78°C (70%); c) $ArB(OH)_2$, $Pd(PPh_3)_4$, Tol/EtOH/aq, Na_2CO_3, rt. (95%); d) LiAlH4, THF, (99%); e) Tf_2O, CH_2Cl_2/Pyr, -78°C - rt (99%); f) NaH, MeI, THF (95%); g) $LiHBEt_3$, THF, 40°C (93-99%); h) TBAF, THF (95%).



(*S*,*S*)-1,4-di(methoxycarbonyl)-2,5-di(trifluromethysulfonyloxy)bicyclo [2.2.2]octan-2,5-diene (bis enol triflate intermediate, after step b, Scheme S1)[1]

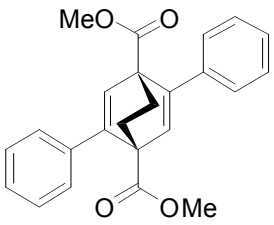


Diketone (*S*,*S*)-**S2** (254 mg, 1 mmol) was added to a 25 mL flask containing anhydrous THF (5 mL). The solution was stirred, cooled to -78°C and LHMDS (2.2, mL, 1.06M, 2.33 mmol) was added. The reaction mixture was stirred for 30 min and triflic anhydride (648 mg, 2.3mmol) was added in one portion. The reaction finished instantaneously and was quenched with saturated $NaHCO_3$ aqueous solution (10 mL) then transferred to a separating funnel. The mixture was extracted with EtOAc (3 × 15 mL), extracts were combined and washed with water and brine then dried over Na_2SO_4. After filtration and evaporation the crude product was purified by flash column chromatography (Hexane:EtOAc; 9:1) to give pure product (*S,S*)-bisenol triflate (362 mg, 70% yield) as colorless oil.

^{1}H NMR δ 6.62 (s, 2H), 3.92 (s, 6H), 2.32-2.38 (m, 2H), 1.90-1.97 (m, 2H); ^{13}C NMR δ 168.1, 152.2, 121.0, 118.7 (q, J = 320Hz), 54.9, 53.8, 31.0.

HRMS ESI+: $C_{14}H_{12}F_6O_{10}S_2Na^+$ {$[M+Na]^+$}, Calc.: 540.9674, Found: 540.9695.

(*S*,*S*)-(+)-1,4-di(methoxycarbonyl)-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (L4a)[1]

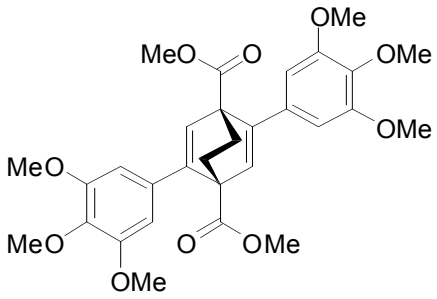


(*S*,*S*)-bis enol triflate (570 mg, 1.1 mmol) was dissolved in toluene (20 mL) and EtOH (7mL) in a 100 ml flask. To this solution, phenylboronic acid (360 mg, 3.0mmol), of Na_2CO_3 aqueous solution (4.5 mL, 2.0M, 9mmol) and tetrakis-

(triphenylphosphine)palladium (63mg, 55 μmol, 5 mol%) were added in turn. The resulting mixture was degassed, charged with nitrogen and then stirred at room temperature for 6h for full conversion. The reaction mixture was extracted with Et_2O (20 mL) and washed with water and brine. After being dried over Na_2SO_4, the organic solution was filtered and evaporation of the solvents gave crude product which was purified by flash column chromatography (hexane:EtOAc; 5:1) to give pure product (*S,S*)-(+)-**L4a** as white solid (390 mg, 95%).

^{1}H NMR δ 7.25-7.32(m, 6H), 7.15-7.17(m, 4H), 6.64 (s, 2H), 3.51 (s, 6H), 2.12-2.18 (m, 2H), 1.87-1.92 (m, 2H), ^{13}C NMR δ 174.4, 148.0, 138.4, 131.9, 128.5, 127.7, 126.9, 57.4, 52.3, 30.6, HRMS ESI+: $C_{24}H_{22}O_4Na^+$ {$[M+Na]^+$}, Calc.: 397.1416, Found: 397.1418; $[\alpha]_D^{20}$ = +60.8 (*c* 0.44, $CHCl_3$).

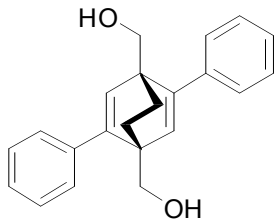
(*S,S*)-(+)-1,4-di(methoxycarbonyl)-2,5-di{3',4',5'-tri(methoxyl)phenyl}bicyclo [2.2.2]octan -2,5-diene (L4h)



This compound was made according to the procedure for compound (*S,S*)-(+)-**L4a** above to give pure product (*S,S*)-(+)-**L4h** as a white solid (152mg of **S2** gave 200 mg of **4h**, 60% overall yield for 2 steps).

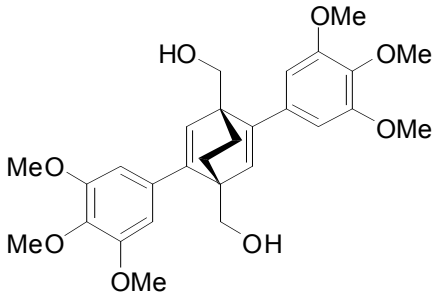
^{1}H NMR δ 6.64 (s, 2H), 6.39 (s, 4H), 3.85 (s, 12H), 3.84 (s, 6H), 3.58 (s, 6H), 2.13-2.18 (m, 2H), 1.88-1.94 (m, 2H); ^{13}C NMR δ 174.2, 153.0, 147.6, 137.3, 133.5, 131.1, 103.4, 60.9, 57.1, 56.1, 52.2, 30.4; HRMS ESI+: $C_{30}H_{34}O_{10}Na^+$ {$[M+Na]^+$}, Calc.: 577.2050, Found: 577.2048; $[\alpha]_D^{20}$ = +36 (*c* 0.37, $CHCl_3$)

(*S*,*S*)-(+)-1,4-di(hydroxymethyl)-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (L6a)[1]



The (*S,S*)-(+)-**L4a** (171mg, 0.45 mmol) was dissolved in THF (5 mL) and Et_2O (5 mL). To this solution $LiAlH_4$ (38 mg, 1 mmol) was added under nitrogen. The reaction finished in 5 min after the addition of $LiAlH_4$. The reaction was quenched by adding EtOAc (5mL) and then water (ca.30 mL) until there was no further effervescence. The resultant mixture was transferred to a separating funnel and extracted with dichloromethane (3 x 20 mL). All extracts were combined and washed with 1N HCl aq. solution, water, sat. $NaHCO_3$ solution and brine then dried over anhydrous Na_2SO_4. Filtration and removal of the solvent gave a colorless oil. Flash column chromatography (hexane:EtOAc; 3:1) gave pure product as sticky oil (145 mg, 99%). ^{1}H NMR δ 7.20-7.40 (m, 10H), 6.27 (s, 2H), 4.15 (dd, J = 11.5Hz, J = 6Hz, 2H), 4.00 (dd, J = 11.5, J = 7Hz, 2H), 1.70-1.77 (m, 2H), 1.58-1.65 (m, 2H), 1.17-1.26 (m, 2H, -OH); ^{13}C NMR δ 149.6, 139.6, 136.0, 128.8, 128.1, 127.6, 65.8, 52.1, 31.0; HRMS ESI+: $C_{22}H_{22}O_2Na^+$ {$[M+Na]^+$}, Calc.: 341.1517, Found: 341.1513; $[\alpha]_D^{20}$ = +44.4 (*c* 0.53, $CHCl_3$)

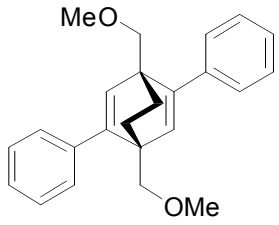
(*S*,*S*)-1,4-di(hydroxymethyl)-2,5-di{3',4',5'-tri(methoxyl)phenyl}bicyclo [2.2.2]octan -2,5-diene (S3)



This compound was made using the same procedure as used for (*S,S*)-(+)-**L6a**. 150

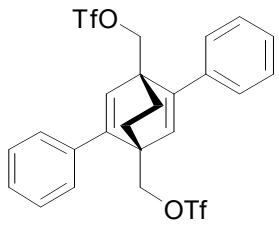
mg starting material gave product (*S,S*)-(+)-**S3** (134mg 99%).
^{1}H NMR δ 6.43 (s, 4H), 6.29 (s, 2H), 4.19 (d, *J* = 11.6Hz, 2H), 4.03 (d, *J* = 11.6Hz, 2H), 3.87 (s, 12H), 3.86 (s, 6H), 1.63-1.75 (m, 2H), 1.56-1.62(m, 2H), 1.39 (brs, 2H); ^{13}C NMR δ 153.5, 149.5, 137.7, 135.8, 135.0, 105.1, 66.9, 61.3, 56.6, 52.2, 31.1; HRMS ESI+: $C_{28}H_{34}O_8Na^+$ {$[M+Na]^+$}, Calc.: 521.2151, Found: 521.2160.

(*S,S*)-(+)-1,4-Di(methoxyl)-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (L7a)



The (*S,S*)- diol **L6a** (300 mg, 0.94 mmol) was dissolved in anhydrous THF (10ml) in a Schlenck reaction tube and stirred at room temperature. To this solution, NaH (200 mg, 60 wt% in mineral oil, 5 mmol) was added followed by addition of MeI (1.42 g, 0.62 mL, 10 mmol) and then stirred at 40^{o}C for 1hr. The reaction was quenched with water then extracted with Et_2O (20 mL). The organic solution was washed with brine and dried over Na_2SO_4. Filtration and removal of solvents by rotary evaporation gave crude product which was purified by flash column chromatography (hexane:EtOAc; 4:1) to give product **L7a (**308 mg, 95%).
$[\alpha]_D^{20}$ = +71.8^o (*c* 0.46, $CHCl_3$); ^{1}H NMR δ 7.11-7.33 (m, 10H), 6.25 (s, 2H), 3.73 (d, *J* = 9.6Hz, 2H), 3.60 (d, *J* = 9.6Hz, 2H), 3.21 (s, 6H), 1.69-1.76 (m, 2H), 1.54-1.61 (m, 2H), ^{13}C NMR δ 149.7, 139.8, 135.3, 128.6, 128.0, 127.0, 75.8, 59.5, 50.0, 31.1; HRMS ESI+: $C_{24}H_{26}O_2Na^+$ {$[M+Na]^+$}, Calc.: 369.1831, Found: 369.1827.

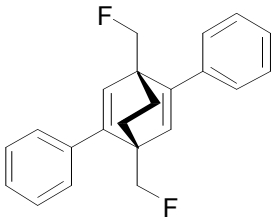
(*S,S*)-1,4-di(trifluoromethanesulfonyloxy)-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (S4)[1]



Diol **L6a** (572 mg, 0.93 mmol) was added to a 50ml flask containing DCM (10 mL) and pyridine (2mL). The solution was cooled to -78°C and triflic anhydride (787 mg, 2.8 mmol) was added. The reaction was then allowed to warm to room temperature and transferred to a separating funnel to which ice-water (40 mL) was subsequently added. The mixture was extracted with DCM (3 X 40 mL). The combined organic extracts were washed with 1N HCl aq. solution, water and brine and dried over Na_2SO_4. Filtration and removal of solvents by rotary evaporation gave crude product, which was purified by flash column chromatography (Hexane:EtOAc 5:1) to give product (*S,S*)- **S4** (1.02 g, 99%) as a colourless oil.

^{1}H NMR δ 7.32-7.40 (m, 6H), 7.10-7.12 (m, 4H), 6.24 (s, 2H), 4.89 (d, J = 10Hz, 2H), 4.73 (d, J = 10Hz, 2H), 1.74-1.90 (m, 4H); ^{13}C NMR δ 149.1, 136.9, 128.9, 128.4, 128.2, 78.3, 49.1, 30.6; HRMS ESI+: $C_{24}H_{20}F_6O_6S_2Na^+$, {$[M+Na]^+$}, Calc.: 605.0503, Found: 605.0530. (Due to the instability of this compound the $[\alpha]_D$ value was not measured.)

(*S*,*S*)-(+)-1,4-di(fluoromethyl)-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (L8a)

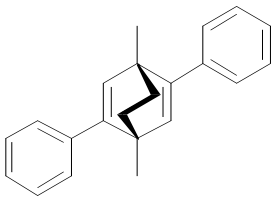


The (*S*,*S*)-**S4** (50 mg, 86μmol) was dissolved in anhydrous THF (5ml) and stirred at room temperature. To this solution TBAF (tetra-*n*butylammonium fluoride THF solution, 0.43 mL, 430μmol) was added in one portion and the reaction mixture allowed to warm 40°C for 3hrs. To this reaction mixture ca. 2g of silica gel was added and resultant mixture was subjected to rotary evaporation to remove all solvent to give silica powder with product absorbed. This was loaded onto a short silica pad and washed with pure hexane to give pure product (*S,S*)-(+)-**L8a** (26mg, 95% yield).

^{1}H NMR δ 7.17-7.35 (m, 10H), 6.30 (s, 2H), 4.62-4.91 (doublet of a AB quartets, J_{H-F} = 47Hz, J = 9.5Hz, 4H), 1.63-1.78 (m, 4H); ^{13}C NMR δ 149.3, 138.5, 134.0 (d, J = 26Hz), 138.5, 127.7, 85.8 (d, J = 678Hz), 50.3 (d, J = 73Hz), 29.9 (d, J = 23Hz);

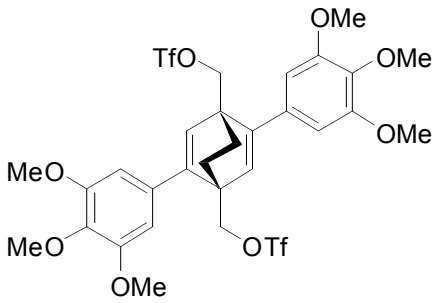
HRMS EI: $C_{22}H_{21}F_2$ {$[M+H]^+$}, Calc. 323.1606, Found: 323.1601; $[\alpha]_D^{20}$ = +125.9 (*c* 0.37, $CHCl_3$)

(*R*,*R*)-(+)1,4-dimethyl-2,5-diphenylbicyclo[2.2.2]octan-2,5-diene (L5h)[1]



The bis triflate (*S*,*S*)- **S4** (582 mg, 1.0 mmol) was dissolved in anhydrous THF (20ml) and cooled to -78°C. To this a stirred solution $LiHBEt_3$ (5 mL, 5.0mmol) was added in one portion and the reaction mixture allowed to warm to room temperature. The reaction was stirred for a further 30min and around 10 g of silica gel was added which was pre-cooled in an ice-water bath. Removal of the solvent gave a silica gel powder with product absorbed, which was loaded onto a short silica column and washed with pure hexane to give pure product (*R*,*R*)- **L5h** as colorless oil (370 mg, 99%).
^{1}H NMR δ 7.13-7.32 (m, 10H), 5.97 (s, 2H), 1.58-1.66 (m, 2H), 1.46-1.54 (m, 2H), 1.42 (s, 6H); ^{13}C NMR δ 151.6, 140.2, 139.0, 128.7, 128.1, 126.8, 45.2, 37.5, 22.6; HRMS: $C_{22}H_{23}$, {$[M+H]^+$}Calc.: 287.1794, Found: 287.1794; $[\alpha]_D^{20}$ = +127.2 (*c* 0.46, $CHCl_3$).

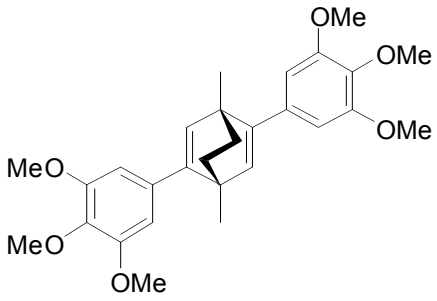
(*S*,*S*)-1,4-di(trifluoromethanesulfonyloxy)-2,5-di{3',4',5'-tri(methoxyl) phenyl}bicyclo [2.2.2]octan -2,5-diene (S5)



The triflate derivative was made using the same procedures as described for compound **S4**. The resulting bistriflate was unstable and used without

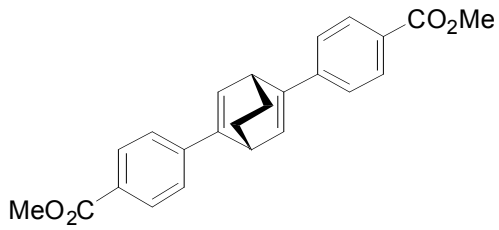
characterization for the subsequent reaction

(*R*,*R*)-(+)-1,4-di(hydroxymethyl)-2,5-di{3',4',5'-tri(methoxyl)phenyl}bicyclo [2.2.2]octan -2,5-diene (L5h)



The bis triflate (*S*,*S*)-**S5** (75 mg, 0.16 mmol) was dissolved in anhydrous THF (3ml) and cooled to -78°C. To this a stirred solution $LiHBEt_3$ (0.16 mL, 0.8mmol, 5eq) was added in one portion and the reaction mixture allowed to warm to room temperature. The reaction was stirred for a further 30min and around 2 g of silica gel was added which was pre-cooled in an ice-water bath. Removal of the solvent gave a silica gel powder with product absorbed, which was loaded onto a short silica column and washed with pure hexane to give pure product 66 mg (94% yield) (*R,R*)-**L5h** as white solid. ^{1}H NMR δ 6.34 (s, 4H), 5.99 (s, 2H), 3.87 (s, 12H), 3.86 (s, 6H), 1.59-1.65 (m, 2H), 1.49-1.54 (m, 2H), 1.47 (s, 6H); ^{13}C NMR δ 152.9, 151.6, 138.6, 137.2, 135.6, 105.8, 61.3, 56.5, 45.3, 37.6, 22.7; $[\alpha]_D^{20}$ = +95.5 (*c* 0.69, $CHCl_3$)

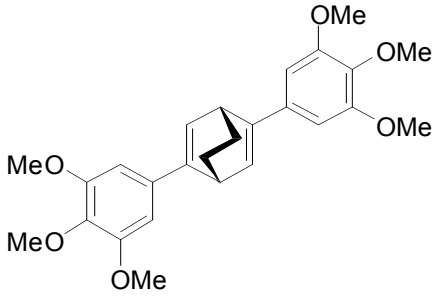
(*R*,*R*)-(+)-2,5-di[4'-(methoxycarbonyl)phenyl]bicyclo[2.2.2]octan-2,5-diene (9i)



This compound was prepared according to ref. [4b] from (*R,R*)-(-)-bicyclo[2.2.2]octan-2,5-dione. ^{1}H NMR δ 8.00 (d, *J* = 8.3Hz, 4H), 7.49 (d, *J* = 8.3Hz, 4H), 6.79 (dd, *J* = 6.4Hz, *J* = 2.0Hz, 2H), 4.30 (d, *J* = 6.4Hz, 2H), 3.92 (s, 6H), 1.59 (s, 4H); ^{13}C NMR δ 166.9, 145.9, 142.2, 131.6, 139.9, 128.3, 124.5, 52.1, 39.9, 25.5;

HRMS ESI+: $C_{24}H_{22}O_4Na^+$ {$[M+Na]^+$}, Calc.: 397.1416, Found: 397.1418; $[\alpha]_D^{20}$ = +22.5 (*c* 0.28, $CHCl_3$),

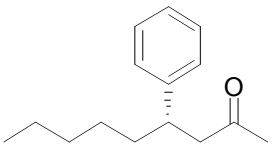
(*R*,*R*)-(-)-2,5-di{3',4',5'-tri(methoxyl)phenyl}bicyclo[2.2.2]octan-2,5-diene (9h)



This compound was prepared according to ref. [4b] from (*R*,*R*)-(-)-bicyclo[2.2.2]octan-2,5-dione. ^{1}H NMR δ 6.65(s, 4H), 6.59 (dd, *J* = 6.4Hz, *J* = 2.0Hz, 2H), 4.17 (d, *J* = 6.0Hz, 2H), 3.90 (s, 12H), 3.84 (s, 6H), 1,58 (s, 4H); ^{13}C NMR δ 153.2, 146.9, 137.3, 134.0, 128.8, 101.9, 60.9, 56.1, 40.3, 25.8; HRMS ESI+: $C_{26}H_{30}O_6Na^+$ {$[M+Na]^+$}, Calc.: 461.1940, Found: 461.1944; $[\alpha]_D^{20}$ = -16.5 (*c* 0.35, $CHCl_3$)

4. Chiral Analysis of Asymmetric Conjugate Addition Products

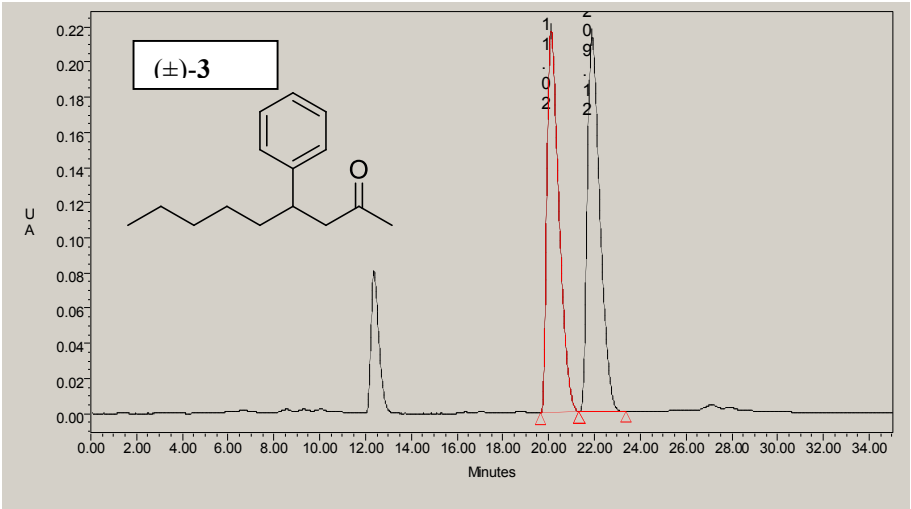
(*S*)-3-phenylnonan-2-one (3)[1]



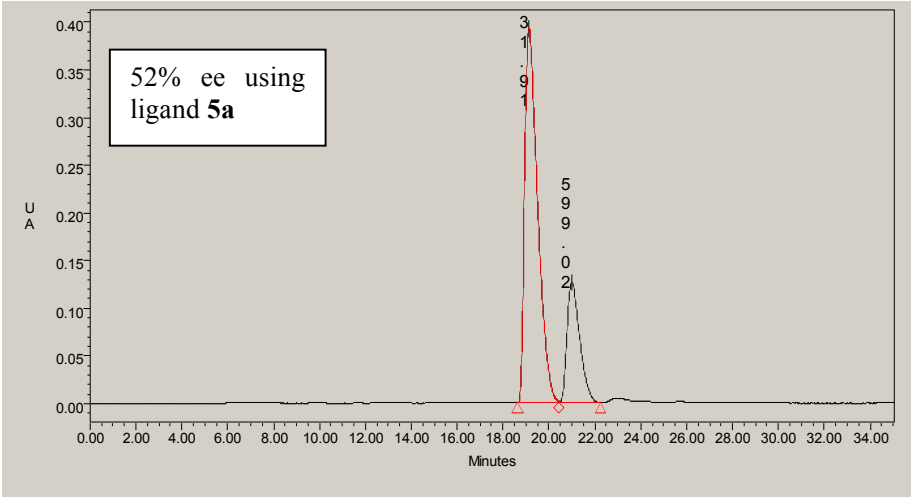
The ee was determined on a Daicel Chiralcel OD-H column with hexane: 2-propanol = 99.9 : 0.1, flow = 1.0 mL/min. Retention times: 20.1 min [(*S*)-enantiomer], 21.9 min [(*R*)-enantiomer]. **L5a** gave 52% ee, $[\alpha]_D^{20}$ = +9.90° (c 0.72, $CHCl_3$); lit[3] ($[\alpha]_D^{20}$ = +17.15° (c 1.30, $CHCl_3$) 92% ee).

^{1}H NMR δ 7.27-7.29 (m, 2H), 7.15-7.20 (m, 3H), 3.06-3.14 (m, 1H), 2.71 (d, *J* = 2.0Hz, 1H), 2.70 (d, *J* = 1.6Hz, 1H), 2.00 (s, 3H), 1.53-1.53 (m, 2H), 1.09-1.21 (m, 7H), 0.82 (t, *J* = 7.0Hz, 3H); ^{13}C NMR: 208.49, 145.02, 128.85, 127.87, 126.70,

51.36, 41.72, 36.84, 32.13, 31.05, 27.44, 22.89, 14.20; HRMS CI: $C_{15}H_{22}O_1^+N_1$ {$[M+NH_4]^+$}, Calc.: 236.2009, Found: 236.2012;

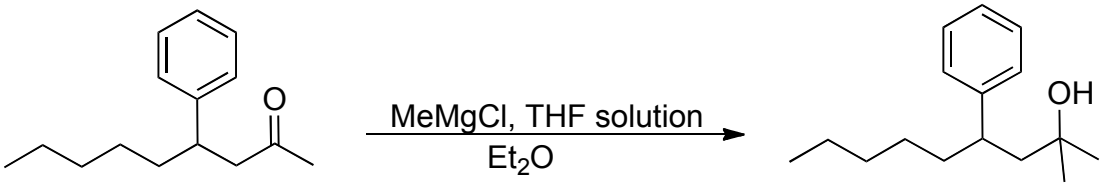


	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		20.110	7936424	49.08	217472	bb	
2		21.902	8233182	50.92	213077	bb	



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		19.131	15511309	75.99	392281	BV	
2		20.995	4899859	24.01	126613	Vb	

Due to the poor reproducibility of the resolution of this compound, the remaining entries were resolved by reacting the compound with methylmagnesium chloride to give the alcohol derivative (**S7**) and subjecting to chiral HPLC with chiralcel OD-H column.



3 **S7**

Protocol for alcohol **S7**:

To a 10ml reaction flask, a diethyl ether solution of ketone **3** (2ml, 5mgs / ml, 46μmol) was added and stirred at room temperature. To this solution, methyl magnesium chloride THF solution (22 μL, 68.8 μmol, 3mol / L) was added and stirred for 5mins. The reaction was quenched by adding 1 ml water to the solution and extracted with diethyl ether (10ml), then dried over anhydrous magnesium sulphate. Filtration followed by evaporation gave the crude product which was purified by preparative TLC (5mg, 93%).

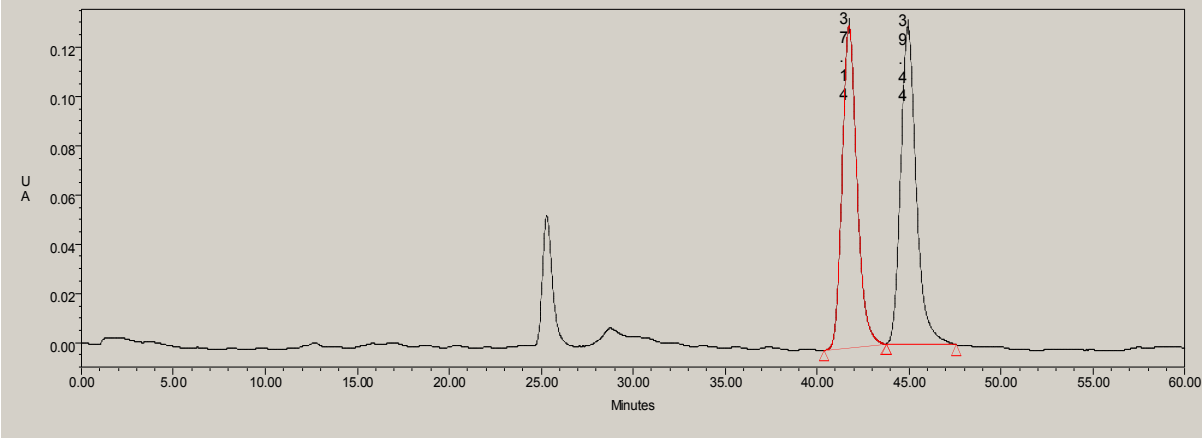
This compound can be resolved on a chiralcel OD-H column with very good reproducibility. Hexane : IPA = 99.5 : 0.5, 0.5ml / min, Retention times: 41 min [(*R*)-enantiomer], 44 min [(*S*)-enantiomer].

Data for compound **S7**

^{1}H NMR δ 7.32-7.17 (m, 3H), 7.23-7.17 (m, 2H), 2.78-2.68 (m, 1H), 1.98 (dd, *J* = 14.2Hz, *J* = 10.0Hz, 1H), 1.83 (dd, *J* = 14.2Hz, *J* = 3.5Hz, 1H), 1.64-1.51 (m, 4H), 1.36-1.17 (m, 4H), 1.13 (s, 3H), 1.11 (s, 3H), 1.03 (brs, 1H), 0.83 (t, *J* = 6.80Hz, 3H); ^{13}C NMR: 146.4, 128.5, 127.8, 126.1, 71.5, 50.0, 42.2, 39.1, 31.8, 30.1, 29.7, 27.1, 22.5, 14.0; HRMS CI: $C_{15}H_{22}O_1^+N_1$ {$[M+NH_4]^+$}, Calc.: 236.2009, Found: 236.2012;

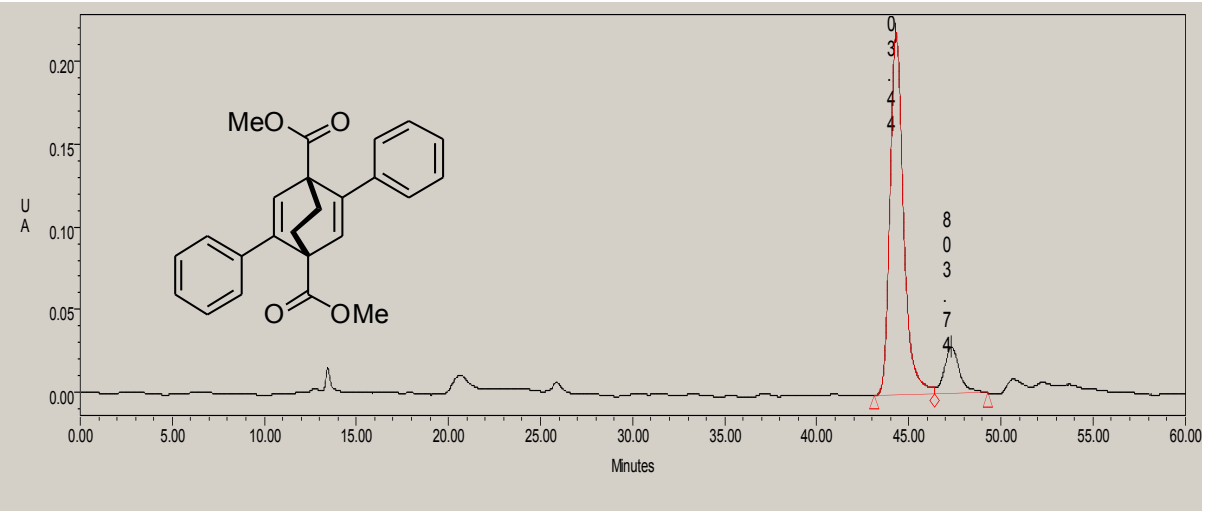
HPLC spectra for **S7** from Ligands tested

Racemic **S7**



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		41.739	7313383	49.07	130817	bb	
2		44.938	7589528	16.32	128815	bb	

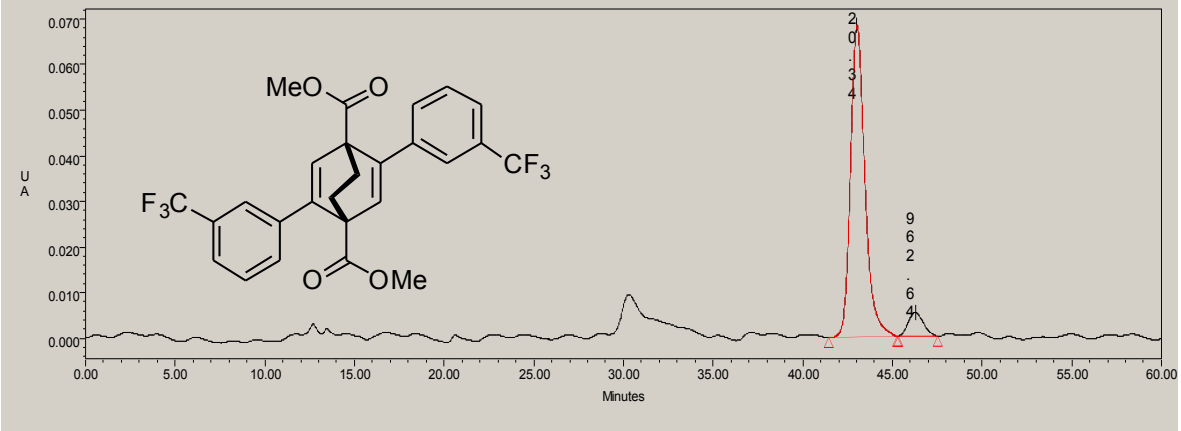
Ligand 4a



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		44.303	11301914	87.21	219050	Bv	
2		47.308	1656787	12.79	28522	vb	

e.e. 74%

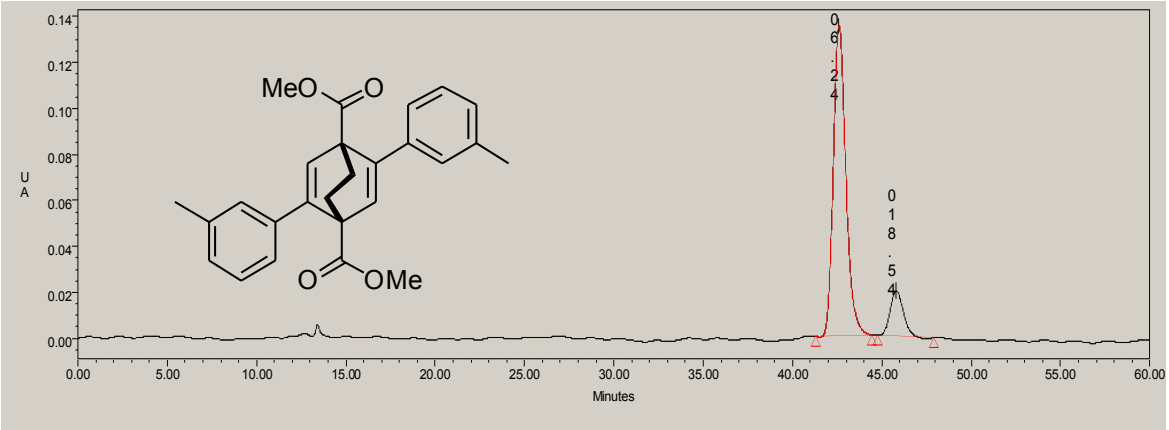
Ligand 4b



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		43.028	3819922	92.67	68429	bb	
2		46.269	302303	7.33	5220	bb	

e.e. 85%

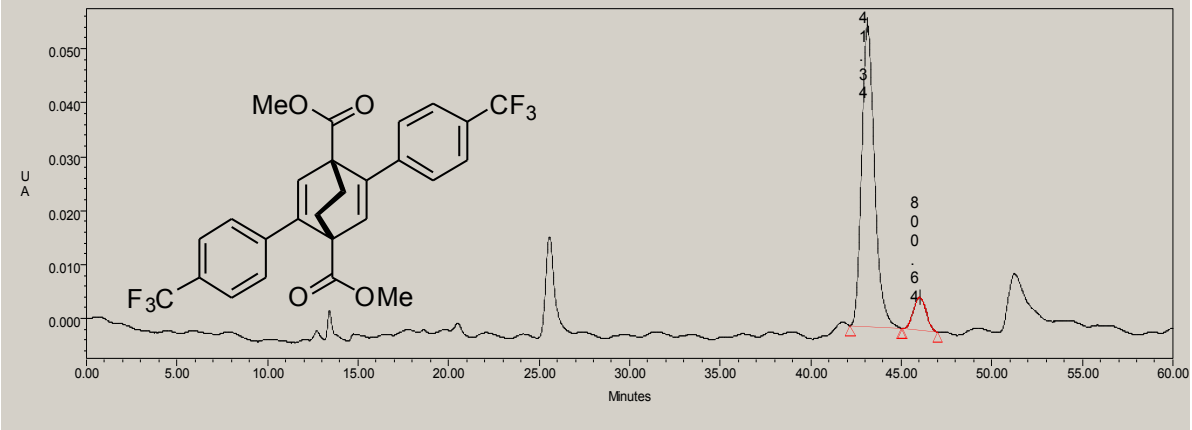
Ligand 4c



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		42.605	6580837	86.54	134892	bb	
2		45.810	1023385	13.46	19701	bb	

e.e. 73%

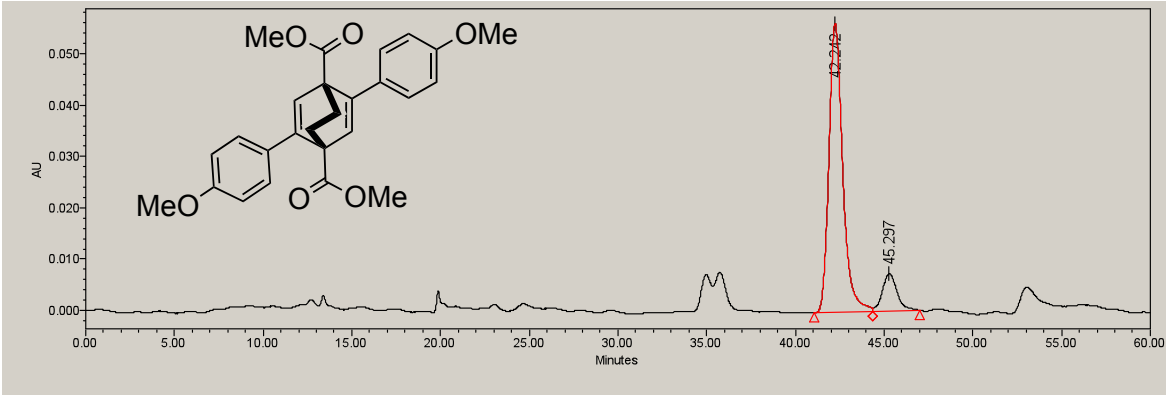
Ligand 4d



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		43.140	2755837	90.31	55920	bb	
2		46.008	295553	9.69	6048	bb	

e.e. 81%

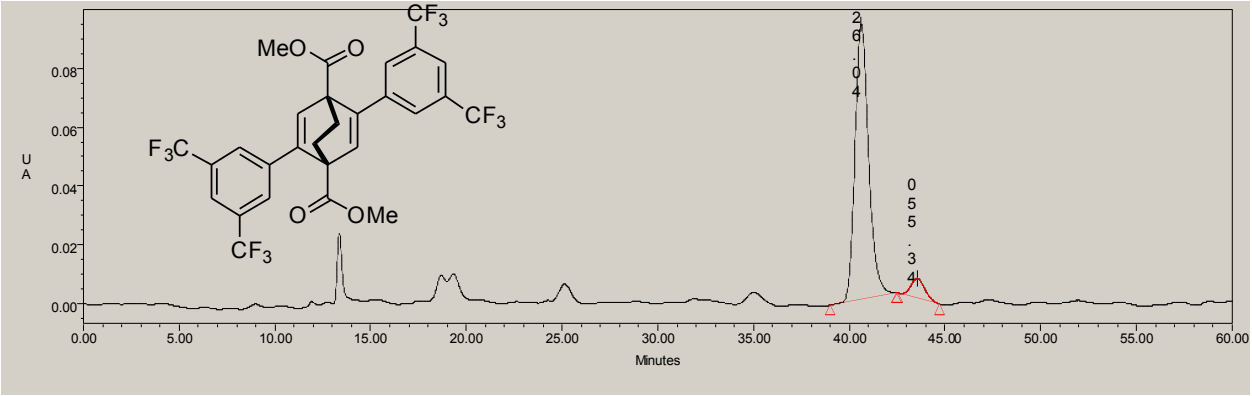
Ligand 4e



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		42.242	3038425	87.79	56340	Bv	
2		45.297	422451	12.21	7301	vB	

e.e. 76%

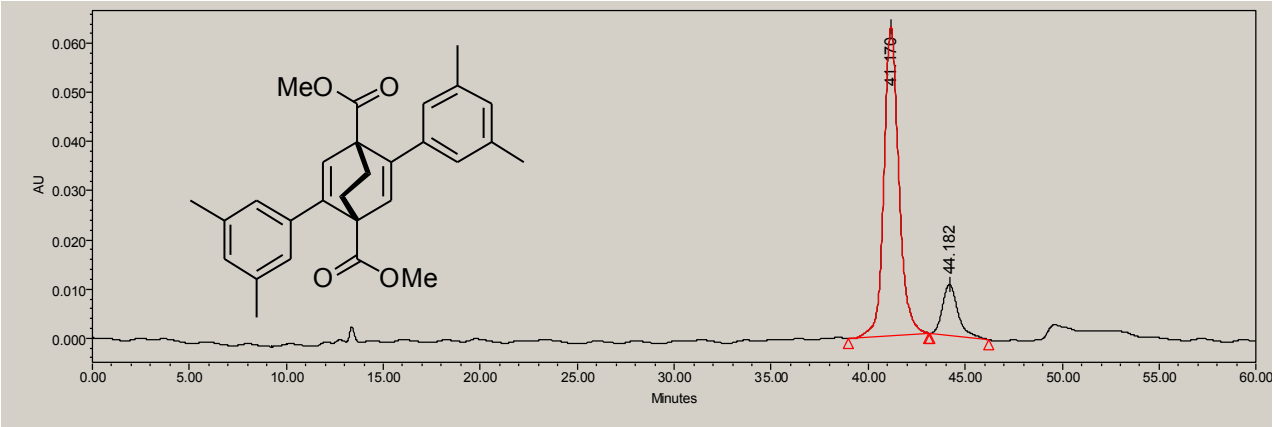
Ligand 4f



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		40.624	4631373	93.09	93545	bb	
2		43.550	343618	6.91	6733	bb	

e.e. 86%

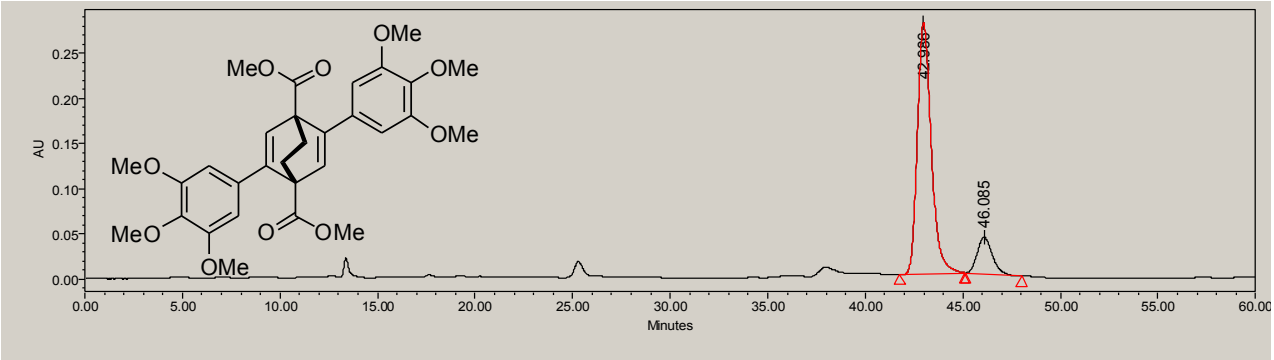
Ligand 4g



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		41.170	3342165	85.34	62947	bb	
2		44.182	574192	14.66	10318	bb	

e.e. 71%

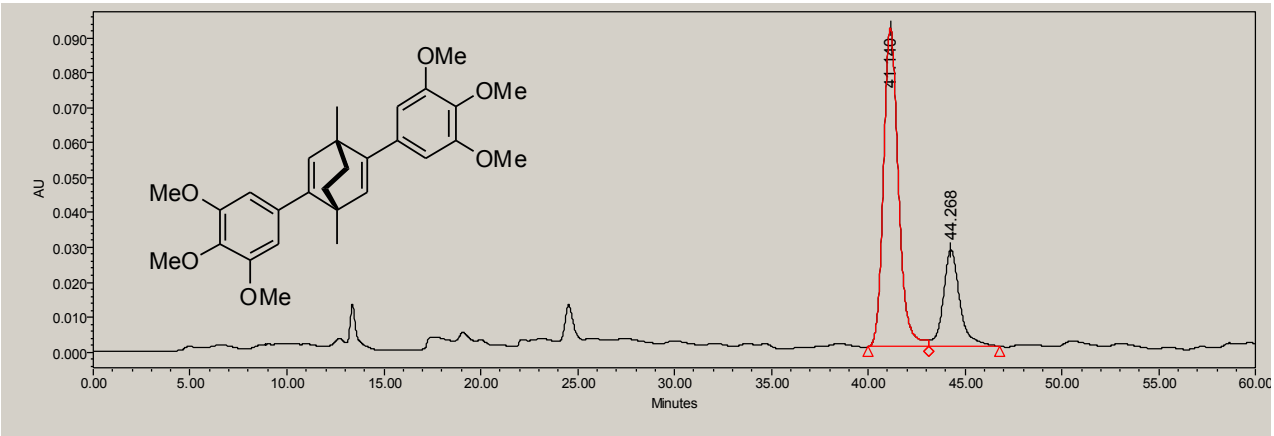
Ligand 4h



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		42.986	13688000	86.10	278389	bb	
2		46.085	2210640	13.90	41003	bb	

e.e. 72%

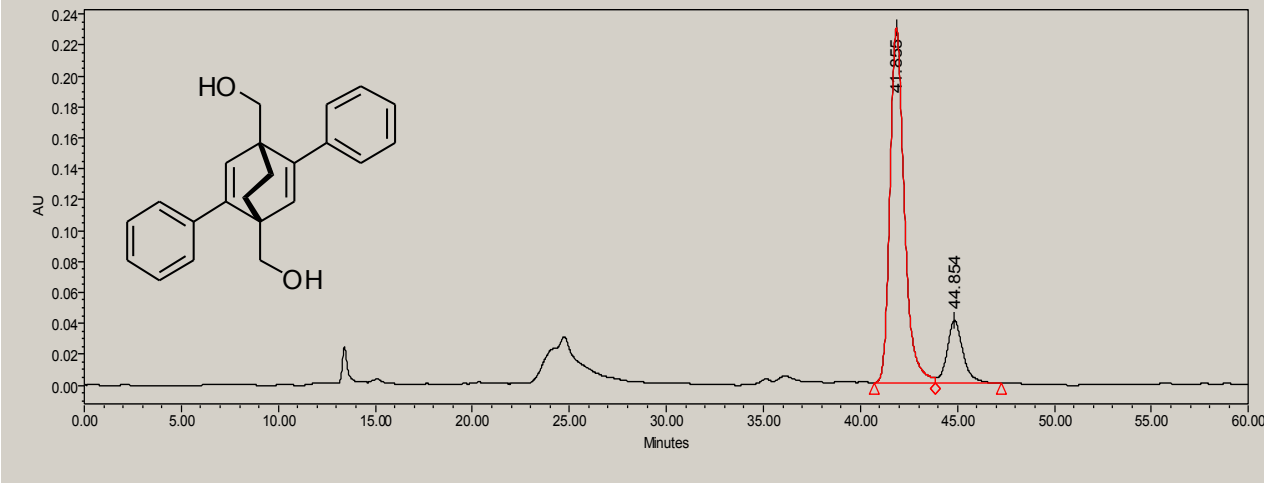
Ligand 5h



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		41.140	4900524	74.67	91320	Bv	
2		44.268	1662761	25.33	27805	vb	

e.e. 49%

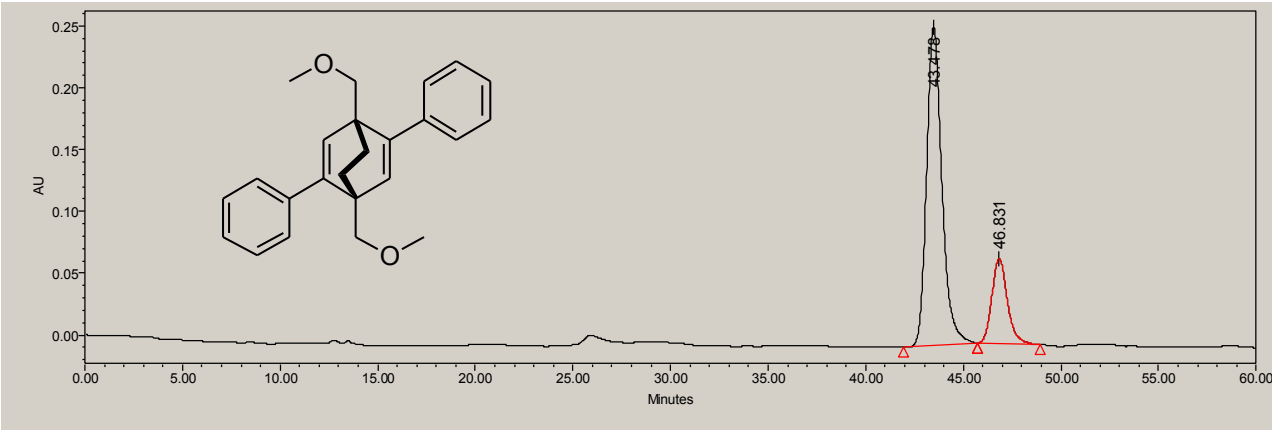
Ligand 6a



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		41.855	11852962	83.68	229801	Bv	
2		44.854	2311387	16.32	40371	vB	

e.e. 67%

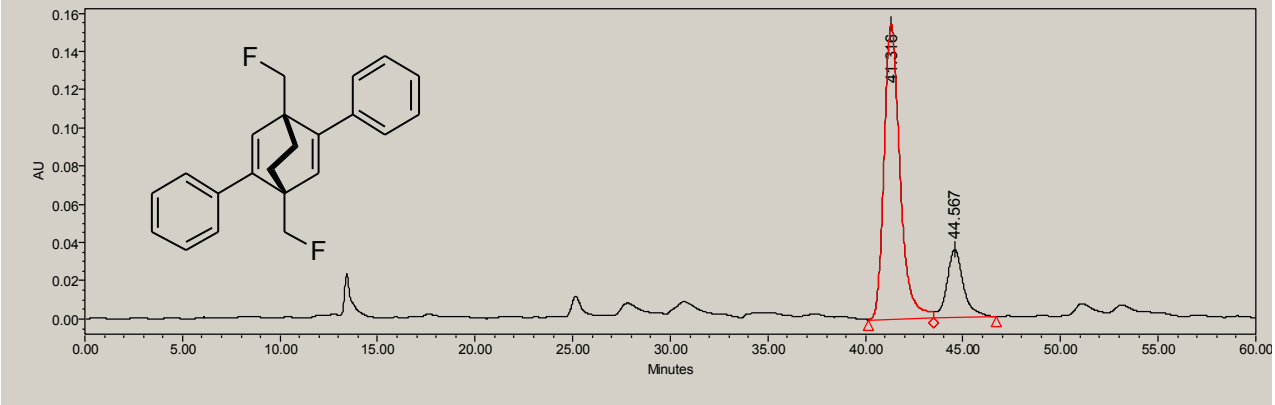
Ligand 7a



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		43.478	14072619	78.55	257816	bb	
2		46.831	3842010	21.45	69042	bb	

e.e. 57%

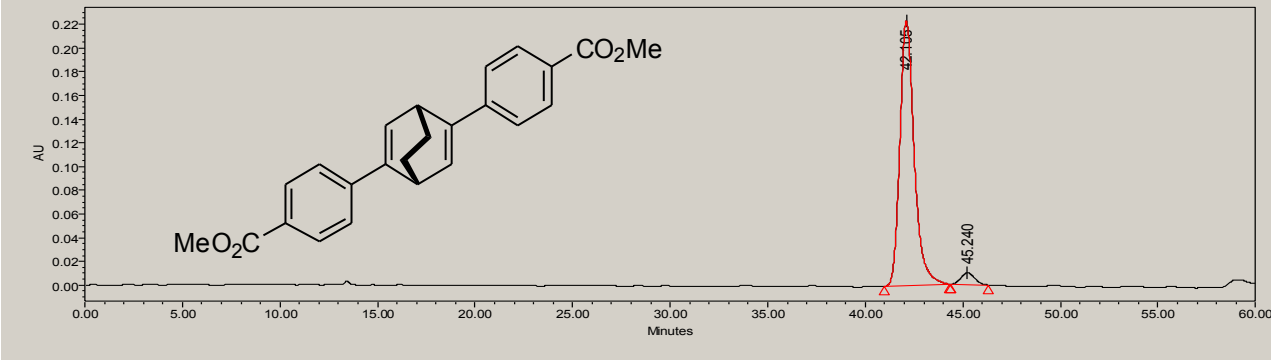
Ligand 8a



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		41.316	8783616	80.59	154685	Bv	
2		44.567	2115200	19.41	35749	vB	

e.e. 61%

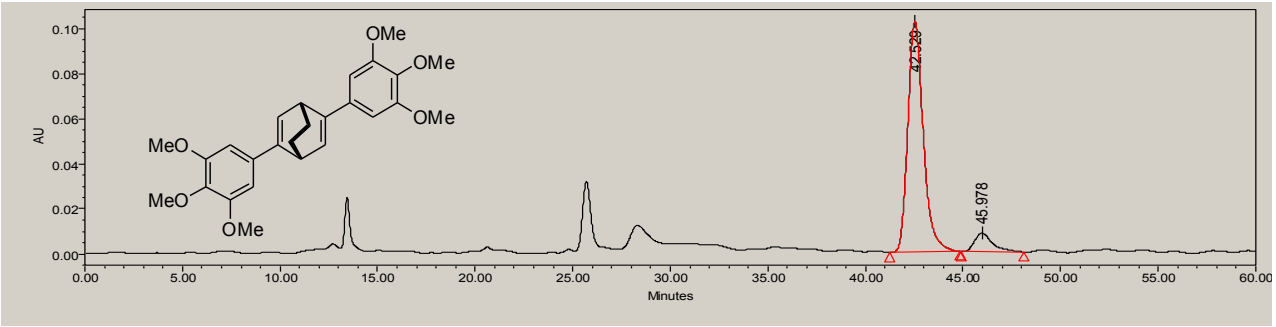
Ligand 9i



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		42.105	11462220	96.02	223582	bb	
2		45.240	475632	3.98	9773	bb	

e.e. 92%

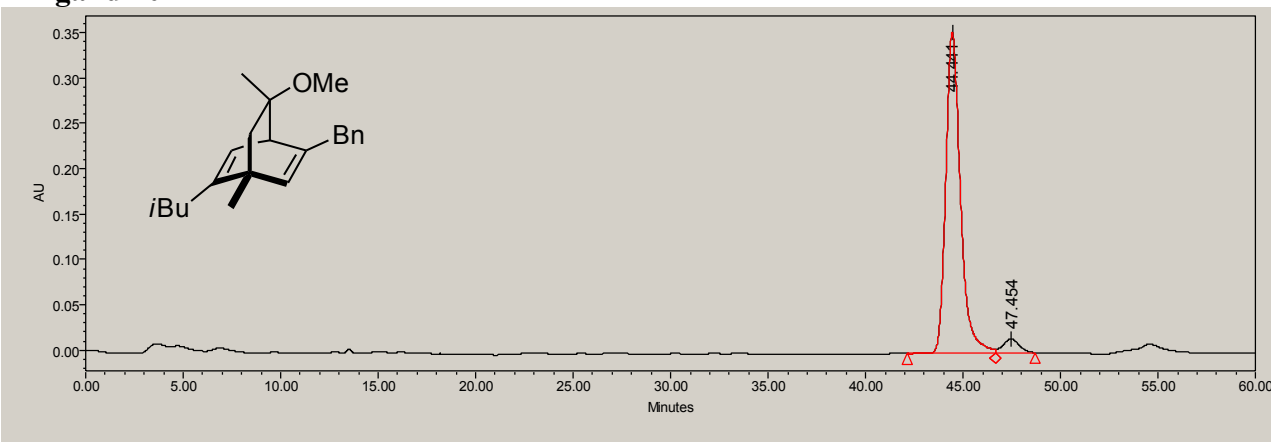
Ligand 9h



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		42.529	5644446	91.44	102506	bb	
2		45.978	528220	8.56	8213	bb	

e.e. 83%

Ligand 10



	Name	Retention Time	Area	% Area	Height	Int Type	Peak Codes
1		44.441	18583584	95.29	353881	bV	
2		47.454	918510	4.71	15688	VB	

e.e. 91%

5. QSPR Modelling

Quantitative structure-property relationship (QSAR) models were developed for the experimentally measured e.e.s for the ligands described in the paper. Three dimensional structures of the ligands were generated using Spartan08[3] and energy minimised using the MMFF94 forcefield. In total 3763 0, 1 and 2-dimensional molecular descriptors/properties were calculated for the set of compounds using DRAGON Version 6.0.[4] To reduce redundancy and remove low-information descriptors, descriptors were removed that had a standard deviation less than 0.001, contained at least one missing value or excluded a descriptor from a pair that had a correlation larger than or equal to 0.95. Descriptors were removed that contained the same values for 80% or more of the training set and the Chorchop procedure was performed to reduce the number of descriptors whilst retaining the intrinsic information content.[5] The multiple linear regression machine learning method coupled with genetic algorithm subjective descriptor selection (GA-MLR) as implemented in the PHAKISO program was used to relate the activities (Y) of a set compounds to their molecular descriptors (X) using a linear equation.[6] The genetic algorithm was set to have population size 50, replacement rate 0.6, cross-over rate 1.0 and maximum number of generations 100. The maximum number of descriptors allowed was set to 3, in order to follow the recommended 5:1 to ratio of number of descriptors to molecule so minimise the occurrence of chance correlations.[7] The subjective fitness function for descriptor selection in this case was chosen to be the adjusted r^2 for the training set. Internal validation statistics were calculated using R[8] with libraries boot[9] and DAAG.[10]

The QPSR model is reliable as confirmed by various diagnostic plots (Figure S1) that indicate that the assumptions on which MLR is based are true for this data set. The plot of residual vs fitted indicates that the residuals do not have a significant trend. The normal Q-Q plot is close to linear indicating the residuals are normally distributed. None of the standardized residuals are greater than 2. None of the data points have a Cook's distance >1.

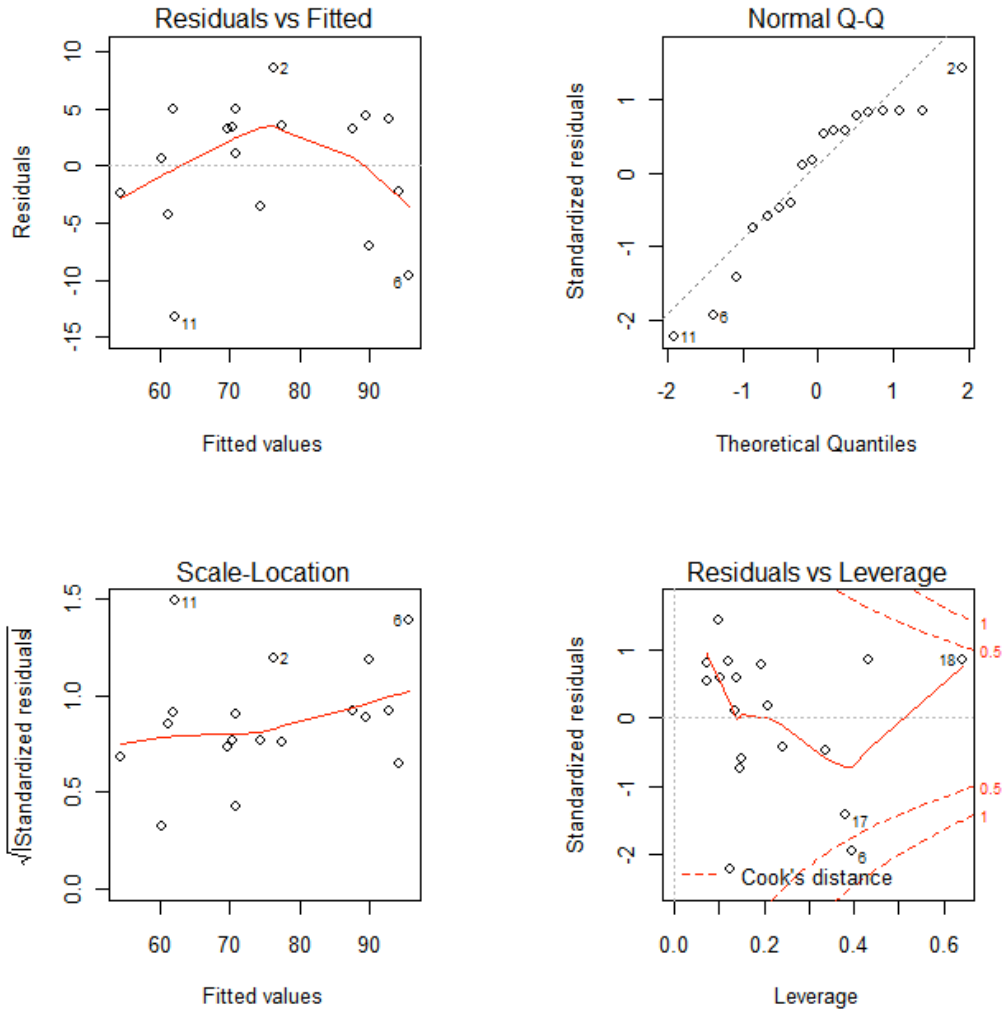
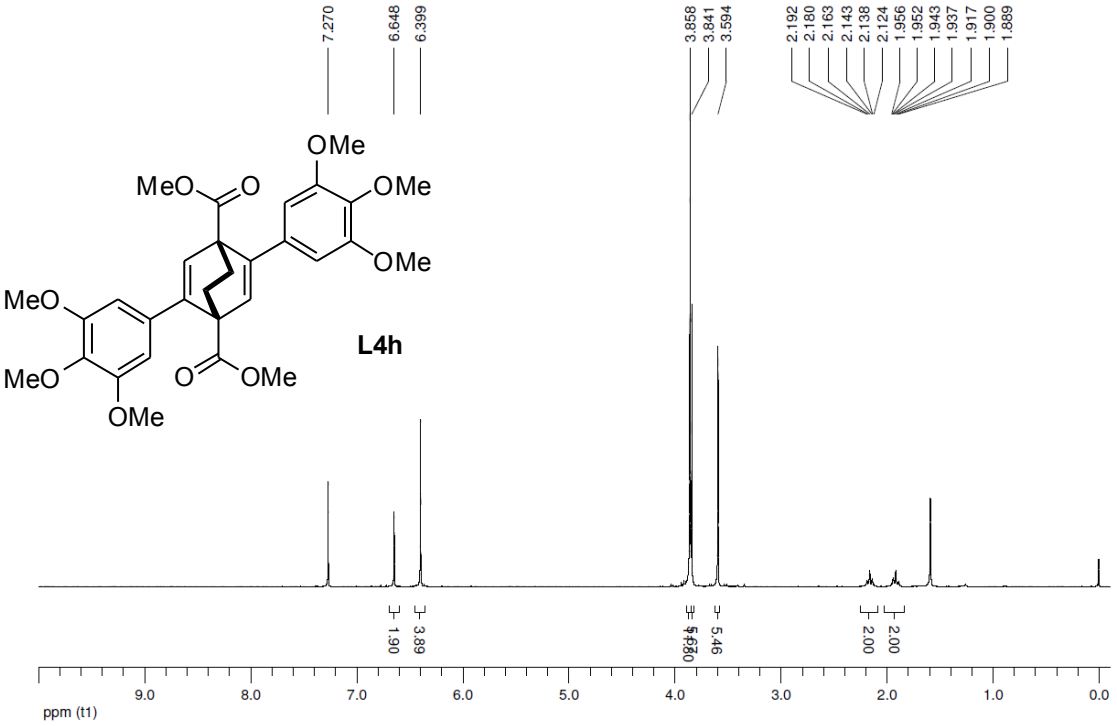
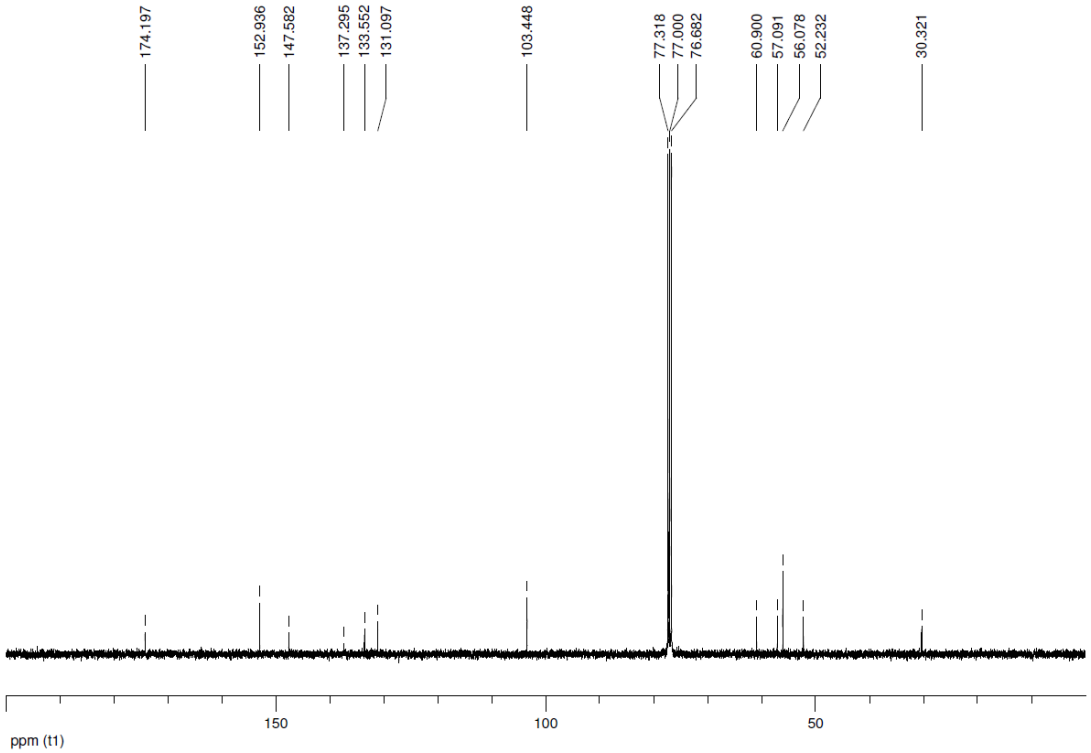


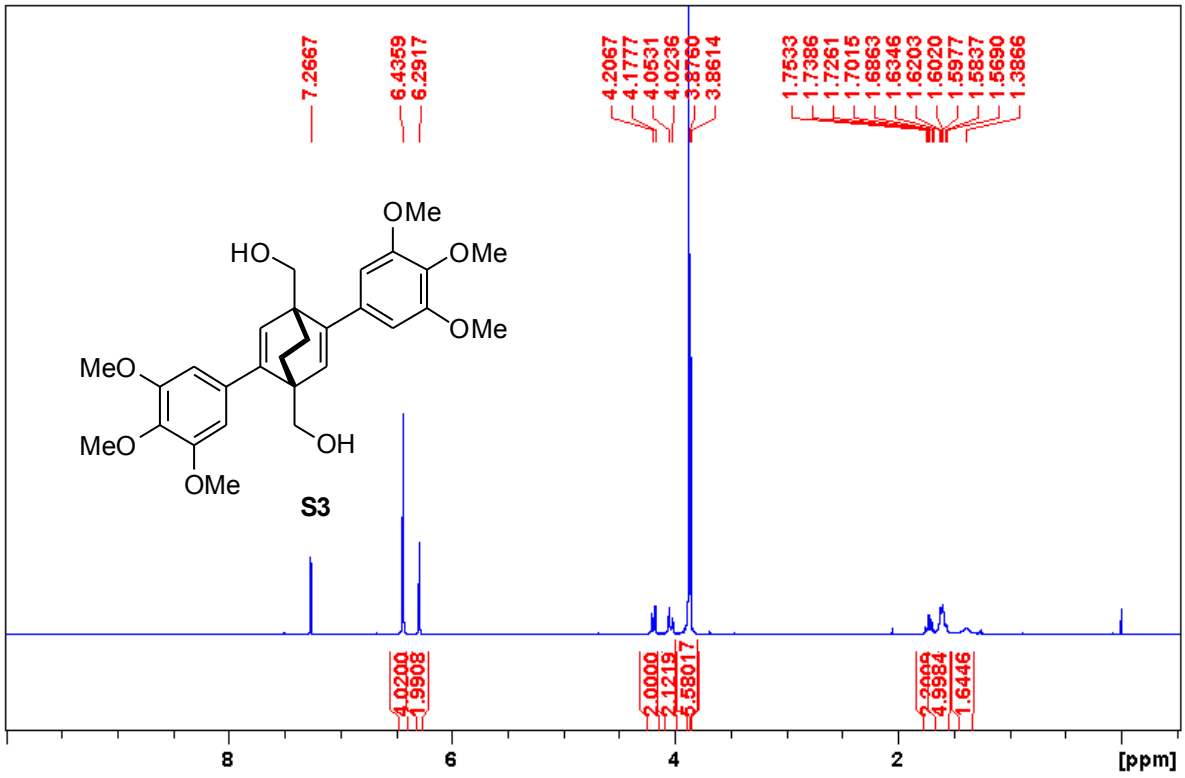
Figure S1 Plots in support of the validity of QSPR model

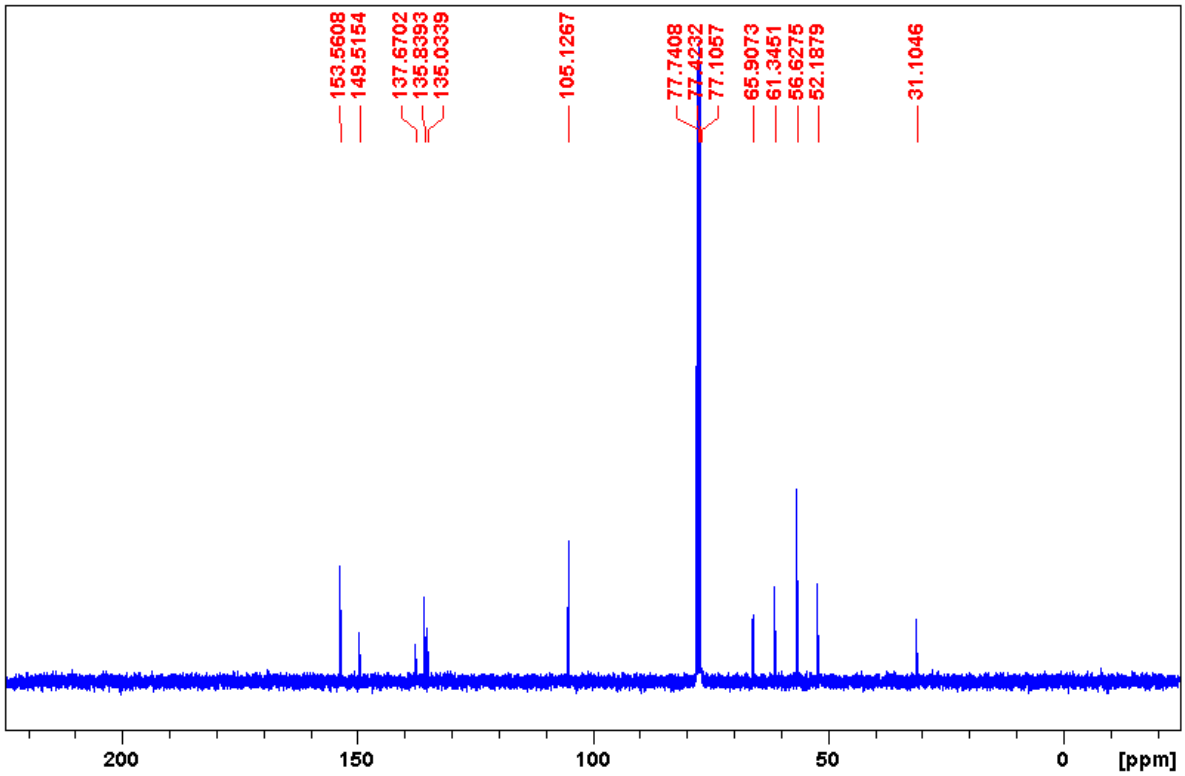
References

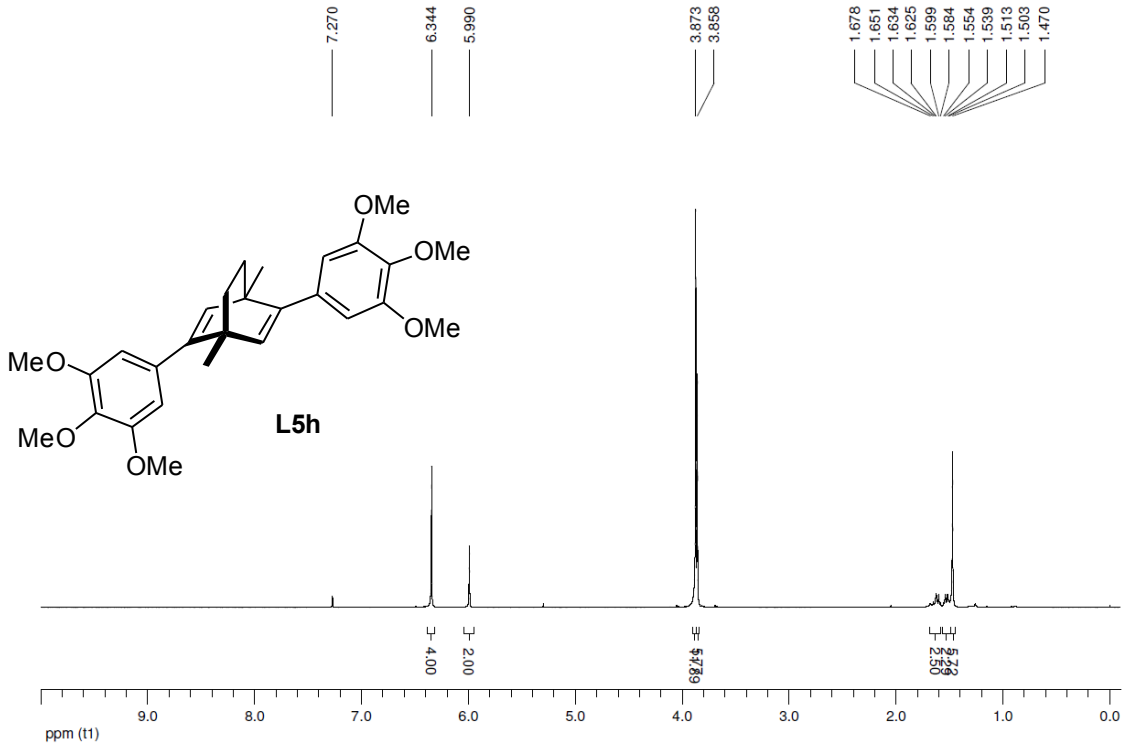
1. Y. Luo and A. J. Carnell, *Angew. Chem. Int. Ed.*, 2010, **49**, 2750.
2. Y. Otomaru, K. Okamoto, R. Shintani and T. Hayashi, *J. Org. Chem.* 2005, **70**, 2503.
3. http://www.wavefun.com/
4. Talete srl, DRAGON (Software for Molecular Descriptor Calculation) Version 6.0 - 2010 - http://www.talete.mi.it/
5. D. J. Livingstone, E. Rahr, *Quant. Struct.-Act. Relat.*, 1989, **8**, 103.
6. http://www.phakiso.com/
7. J. G. Topliss, R. P. Edwards, *J. Med. Chem.*1979, 22, 1238.
8. http://cran.r-project.org/
9. http://cran.r-project.org/web/packages/boot/index.html
10. http://cran.r-project.org/web/packages/DAAG/index.html

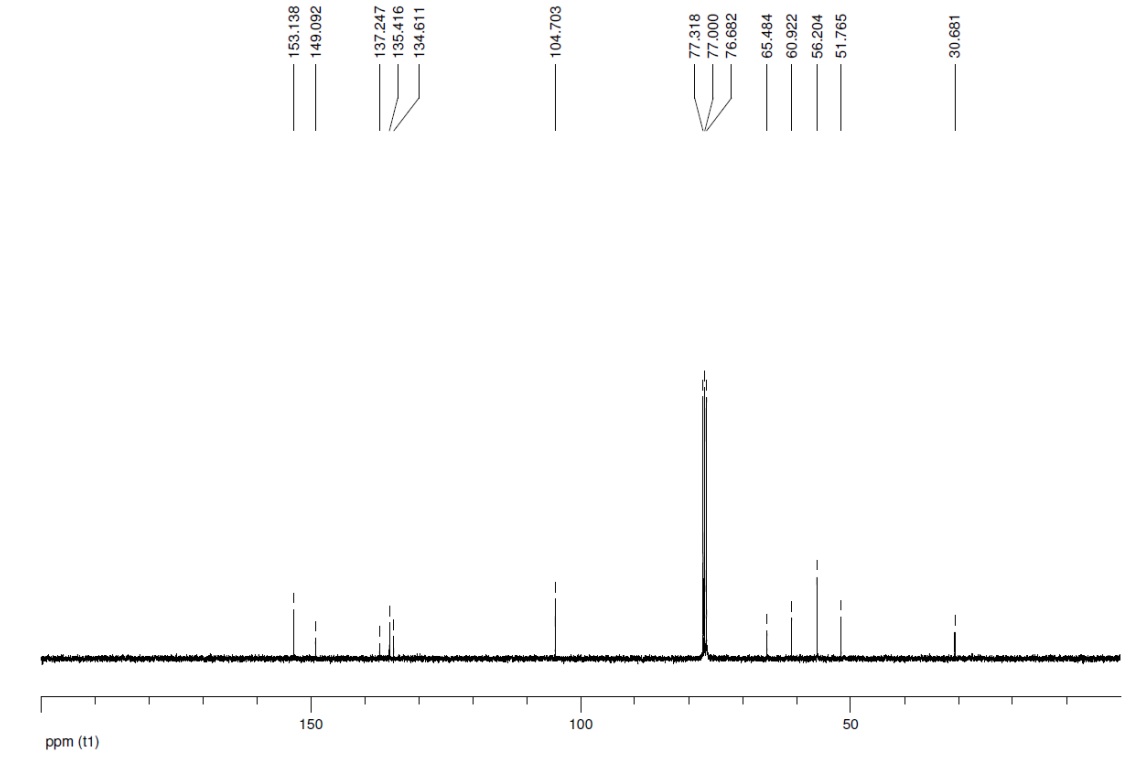

7.270
6.648
6.399
3.858
3.841
3.594
2.192
2.180
2.163
2.143
2.138
2.124
1.956
1.952
1.943
1.937
1.917
1.900
1.889
OMe
MeO
O
OMe
OMe
MeO
MeO
O
OMe
OMe
L4h
1.90
3.89
5.46
2.00
2.00
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (t1)

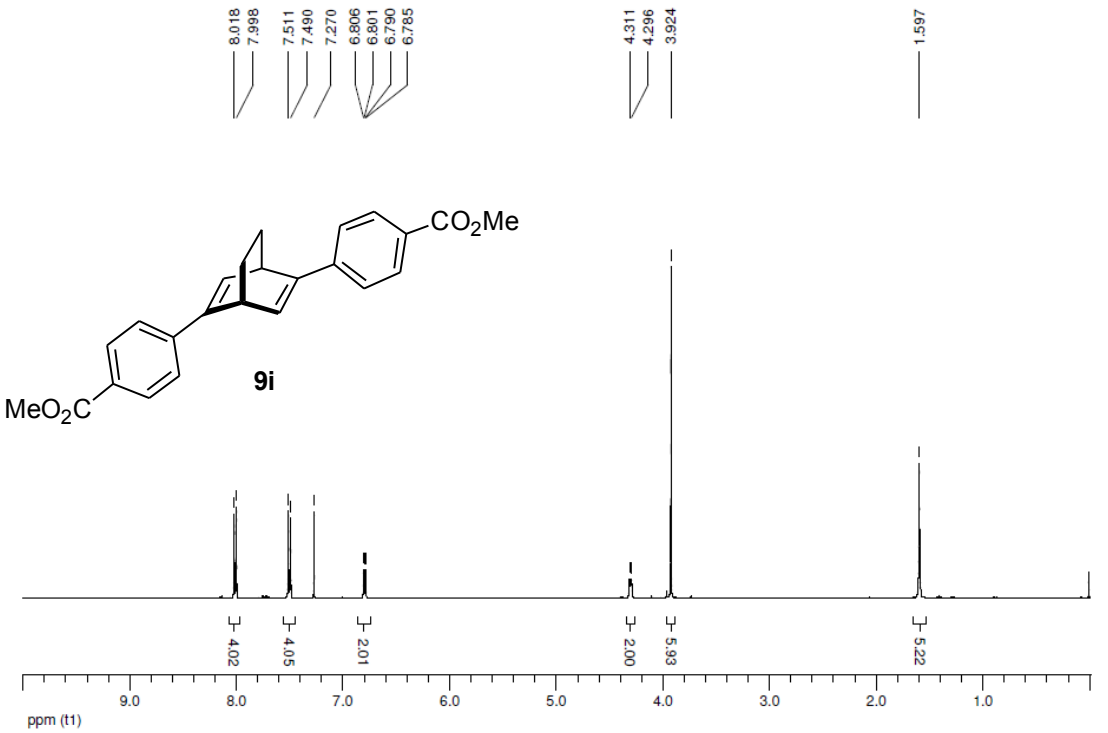

174.197
152.936
147.582
137.295
133.552
131.097
103.448
77.318
77.000
76.682
60.900
57.091
56.078
52.232
30.321
150
100
50
ppm (t1)

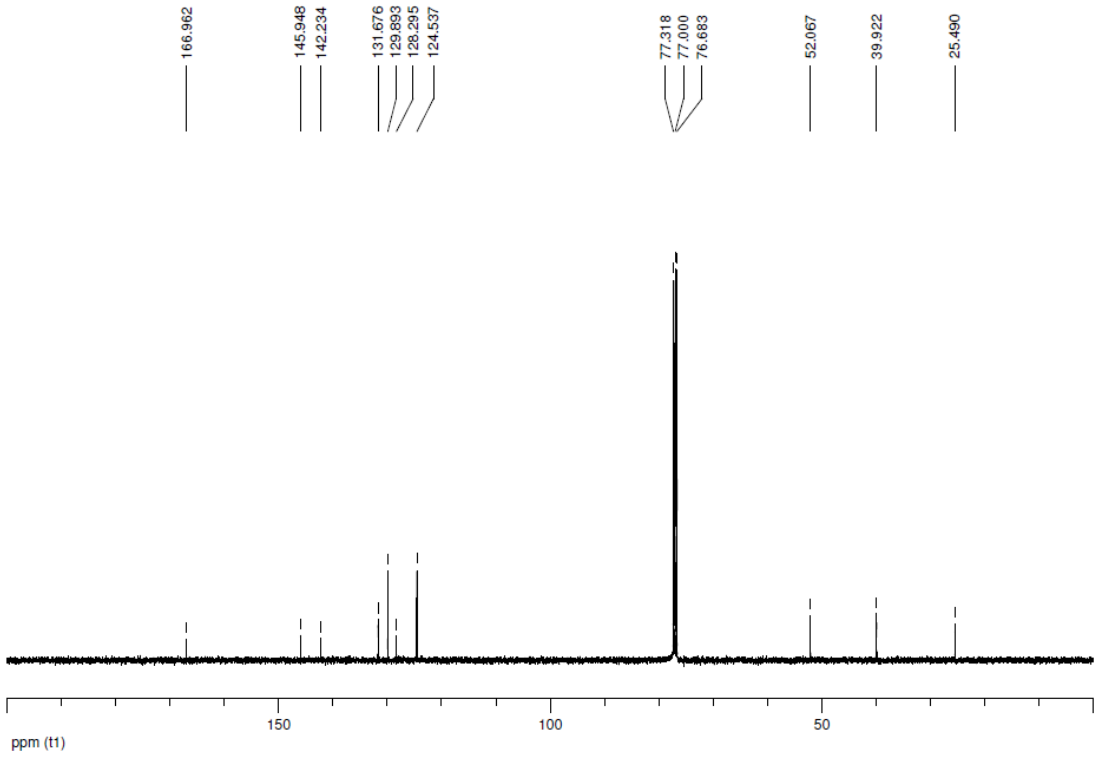

7.2667
6.4359
6.2917
4.2067
4.1777
4.0531
4.0236
3.8760
3.8614
1.7533
1.7386
1.7261
1.7015
1.6863
1.6346
1.6203
1.6020
1.5977
1.5837
1.5690
1.3866
OMe
HO
OMe
OMe
MeO
MeO
OMe
OH
S3
4.0200
1.9908
2.0000
2.1219
5.5801
2.0009
4.9984
1.6446
8
6
4
2
[ppm]

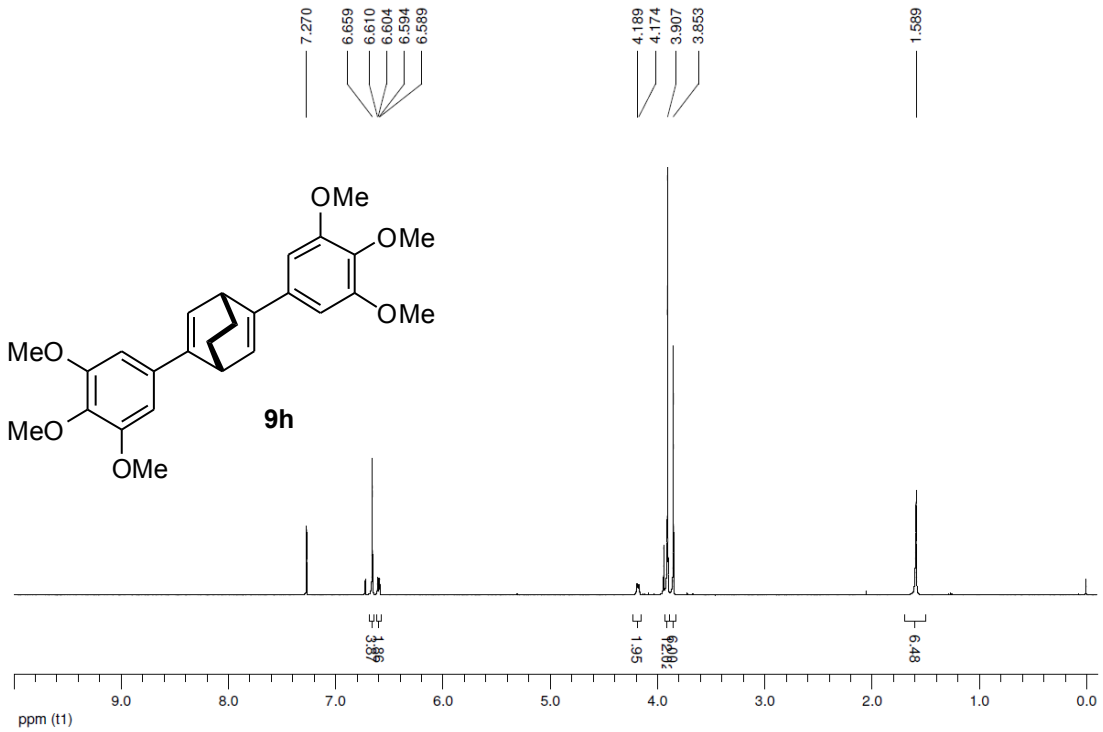

153.5608
149.5154
137.6702
135.8393
135.0339
105.1267
77.7408
77.4232
77.1057
65.9073
61.3451
56.6275
52.1879
31.1046
200
150
100
50
0
[ppm]

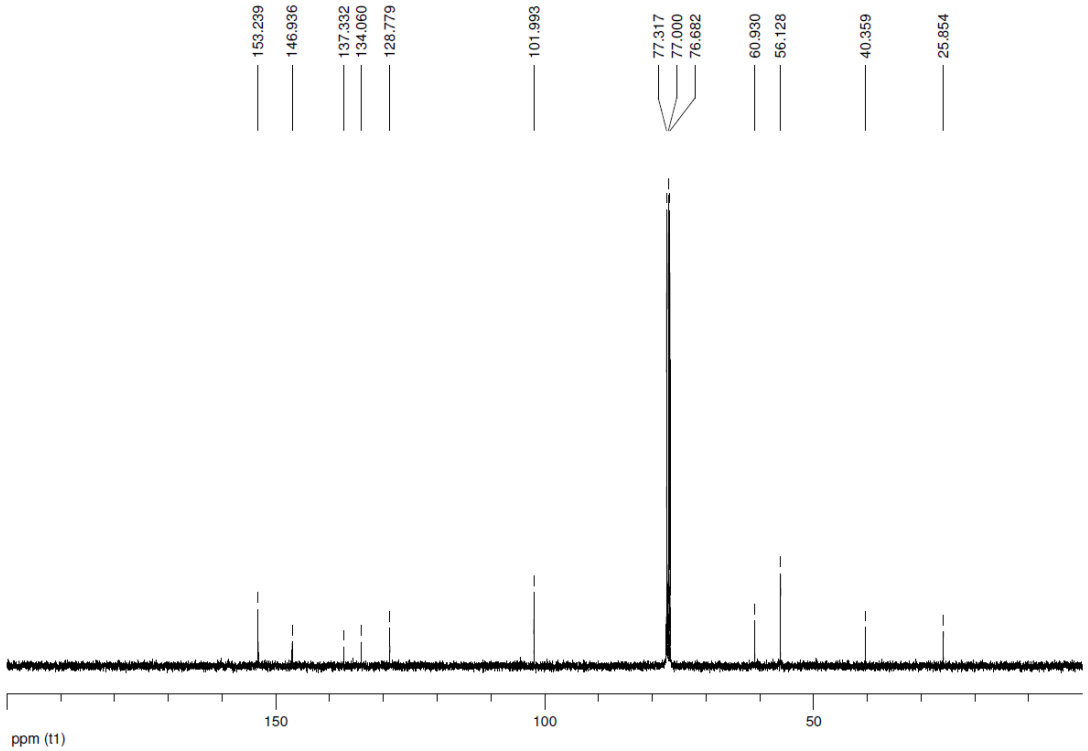

7.270
6.344
5.990
3.873
3.858
1.678
1.651
1.634
1.625
1.599
1.584
1.554
1.539
1.513
1.503
1.470
OMe
OMe
OMe
MeO
MeO
OMe
L5h
4.00
2.00
6.89
2.30
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (t1)

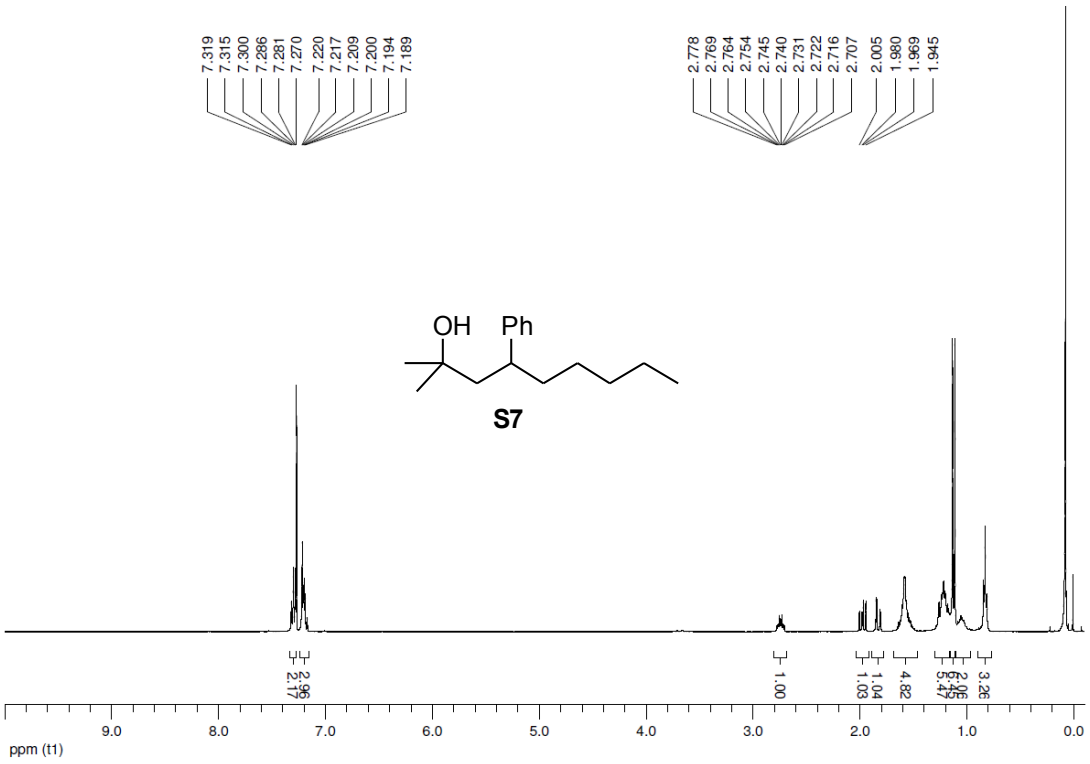

153.138
149.092
137.247
135.416
134.611
104.703
77.318
77.000
76.682
65.484
60.922
56.204
51.765
30.681
150
100
50
ppm (t1)

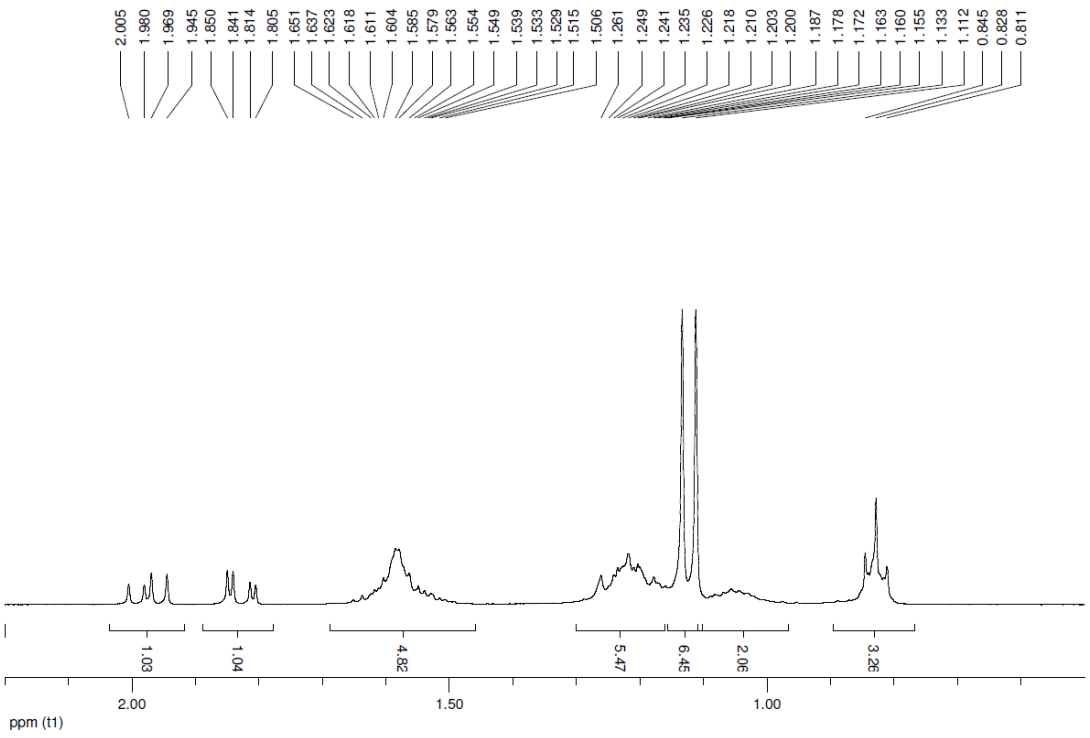




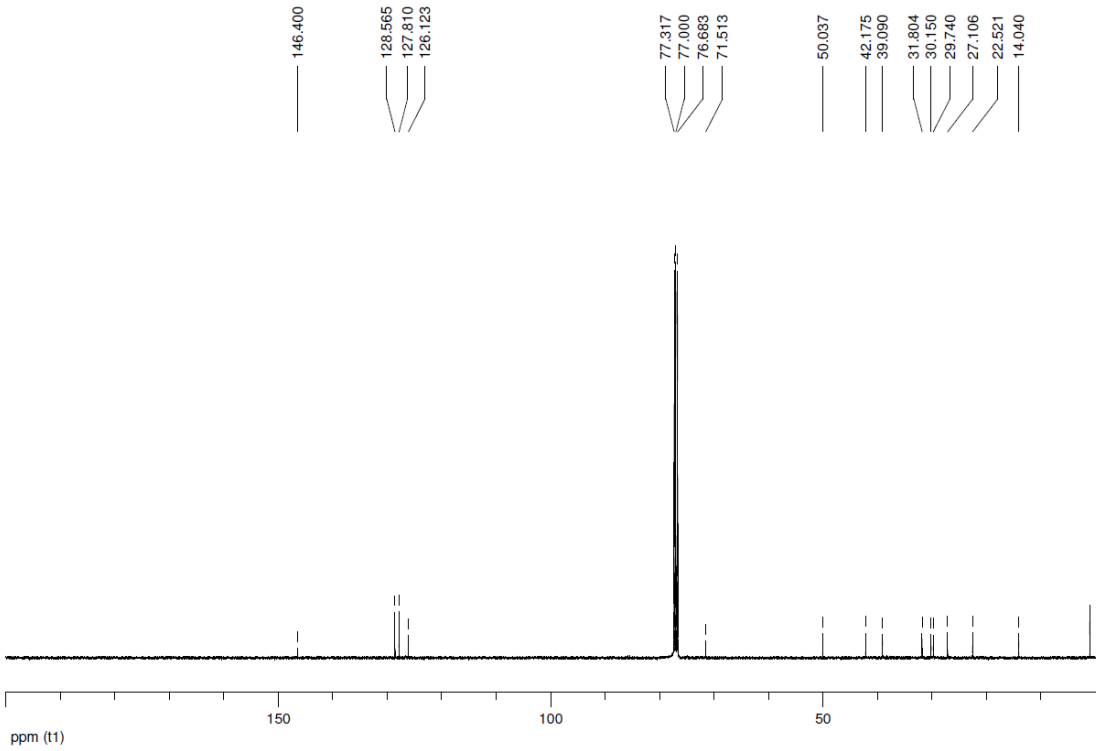

7.270
6.659
6.610
6.604
6.594
6.589
4.189
4.174
3.907
3.853
1.589
OMe
OMe
OMe
MeO
MeO
OMe
9h
3.00
1.95
6.48
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
ppm (t1)


153.239
146.936
137.332
134.060
128.779
101.993
77.317
77.000
76.682
60.930
56.128
40.359
25.854
150
100
50
ppm (t1)





Expanded spectrum of **S7** from 0.5-2.2 ppm


146.400
128.565
127.810
126.123
77.317
77.000
76.683
71.513
50.037
42.175
39.090
31.804
30.150
29.740
27.106
22.521
14.040
150
100
50
ppm (t1)