Supporting Information:

Iterative Supervised Principal Component Analysis-Driven Ligand Design for Regioselective Ti-Catalyzed Pyrrole Synthesis

Xin Yi See, Xuelan Wen, T. Alexander Wheeler, Channing K. Klein, Jason D. Goodpaster, Benjamin R. Reiner†*, Ian A. Tonks*

Contribution from the Department of Chemistry, University of Minnesota – Twin Cities, 207 Pleasant St SE, Minneapolis MN 55455.
E-mail: breiner@dow.com and itonks@umn.edu

†*Present address: Dow Chemical, 400 Arcola Road, Collegeville, PA 19426*

Contents

Experimental work

General considerations

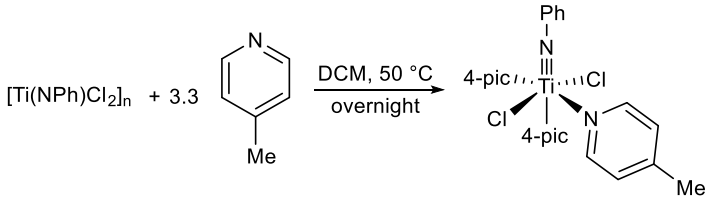
All air- and moisture-sensitive reactions were carried out in a nitrogen-filled glovebox. Solvents for air- and moisture-sensitive reactions were prepared in one of three ways: i) predried on a Pure Process Technology solvent purification system (C_6H_6, THF, PhMe, $PhCF_3$, hexanes, pentane, ether), ii) degassed, dried over CaH_2 and stored over molecular sieves prior to use (PhBr) and iii) vacuum transferred from sodium benzophenone ketyl (C_6D_6) or CaH_2 ($CDCl_3$). C_6D_5Br was synthesized following literature procedure,[1] degassed, dried over CaH_2, freeze-pump-thawed thrice , brought into the glovebox, filtered through basic alumina prior to use. Azobenzene was purchased from TCI America (100 g), extracted with hexanes/water to remove residual methanol and dried under *vacuo* before use. Commercial phenyl propyne was used dried over CaH_2, freeze-pump-thawed three times, brought into the glovebox and filtered through basic alumina prior to use.

Solid $[Ti(NPh)Cl_2]_n$[2] and $[Ti(NTol)Cl_2]_n$[3] were synthesized as previously described, dried under *vacuo* and stored in the freezer until use. $[Ti(NPh)Br_2]_n$ and $[Ti(NPh)I_2]_n$ was prepared following a modified procedure using $TiBr_4$ and TiI_4[2] instead. Catalysis results for $py_3Ti(NPh)Cl_2$ (**1**) were used directly for comparison from our previous work.[4]

^1H and ^{13}C NMR spectra were recorded on Bruker Avance III HD 400 and 500 MHz and Bruker Avance Neo 600 MHz spectrometers. Chemical shifts are referenced to residual solvent resonances: ^1H (s, 7.26 ppm for $CHCl_3$; s, 7.16 ppm for C_6H_6; s, 7.30 ppm, 7.02 ppm, and 6.94 ppm for C_6D_4HBr[5]), ^{13}C (t, 77.2 ppm for $CDCl_3$; t, 128.06 ppm for C_6D_6). Poor solubility of the following complexes required collection of the ^{13}C NMR spectrum via indirect detection (^1H – ^{13}C HSQC, HMBC): **3**, **6**, **10**, **12**, **13**, **16 – 18**, **20**, **22**, **23**, b, d and e. All expected resonances were not observed likely owing to fast T2 relaxation or dynamic ligand exchange occurring at the Larmor frequency. Complexes **5**, **9**, and **15** are virtually insoluble which precluded collection of ^{13}C NMR spectra entirely.

Synthesis of compounds

Synthesis of 2



(4-picoline)$_3$Ti(NPh)Cl$_2$ was synthesized with a slight modification to literature procedure by heating it overnight in CH_2Cl_2.[6]

^{1}H NMR (500 MHz, $CDCl_3$): δ 8.95 (d, $^3J_{HH}$ = 6.0 Hz, 4H, *o*-4-picoline-*H*), 8.63 (br s, 2H, axial *o*-4-picoline-*H*), 7.18 (d, $^3J_{HH}$ = 6.3 Hz, 4H, *o*-NPh-*H*), 7.06 – 7.03 (m, 4H, *m*-NPh-*H* and axial *m*-4-picoline-*H*), 6.91 (d, $^3J_{HH}$ = 7.6 Hz, 2H, *m*-4-picoline-*H*), 6.77 (t, $^3J_{HH}$ = 7.4 Hz, 1H, *p*-NPh-*H*), 2.41 (s, 6H, 4-picoline-C*H*$_3$), 2.33 (s, 3H, axial 4-picoline-C*H*$_3$).

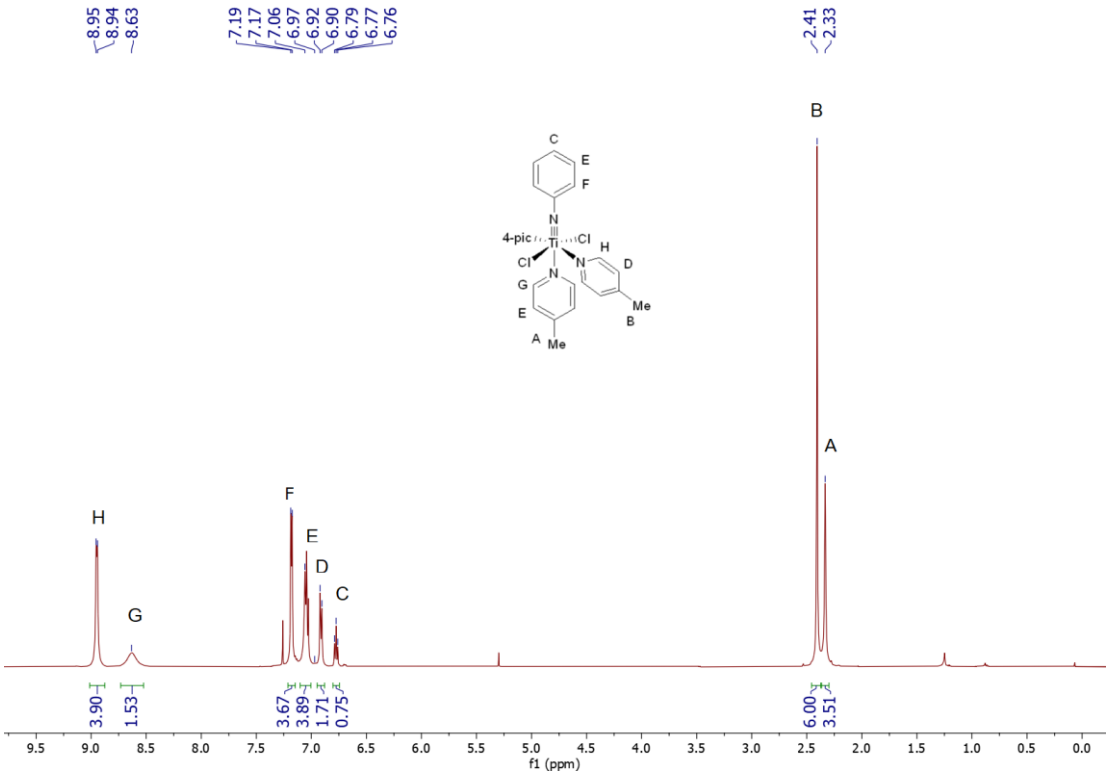
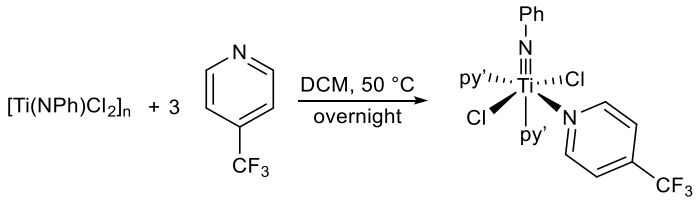


Figure S1. ^{1}H NMR spectrum of **2** in $CDCl_3$.

Synthesis of 3



$[Ti(NPh)Cl_2]_n$ (102 mg, 0.49 mmol, 1.0 equiv), 4-(trifluoromethyl)pyridine (214 mg, 1.46 mmol, 3.0 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. No visible residue was observed. The filtrate was dried *in vacuo* to give **3** as a tan powder (210 mg, 66 % yield).

^{1}H NMR (500 MHz, $CDCl_3$): δ 9.32 (br s, 4H, *o*-4-(trifluoromethyl)pyridine-*H*), 8.97 (br s, 2H, axial *o*-4-(trifluoromethyl)pyridine-*H*), 7.59 (br s, 4H, *m*-4-(trifluoromethyl)pyridine-*H*), 7.53 (br s, 2H, axial *m*-4-(trifluoromethyl)pyridine-*H*), 7.06 (t, $^3J_{HH}$ = 7.8 Hz, 2H, *o*-NPh-*H*), 6.88 (d, $^3J_{HH}$ = 8.4 Hz, 2H, *m*-NPh-*H*), 6.84 (t, $^3J_{HH}$ = 7.4 Hz, 1H, *p*-NPh-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (126 MHz, C_6D_6): 160.1, 152.1, 139.4 (q), 128.6, 128.0, 124.0, 123.6, 122.9, 120.7, 119.5, 119.5.
This is a partial line-list. Resonances associated with a quaternary C from the CF_3 fragment were not observed.

^{19}F NMR (471 MHz, $CDCl_3$): δ -65.1 (br s, 3F, axial-4-(trifluoromethyl)pyridine-*F*), -65.3 (s, 6F, 4-(trifluoromethyl)pyridine-*F*).

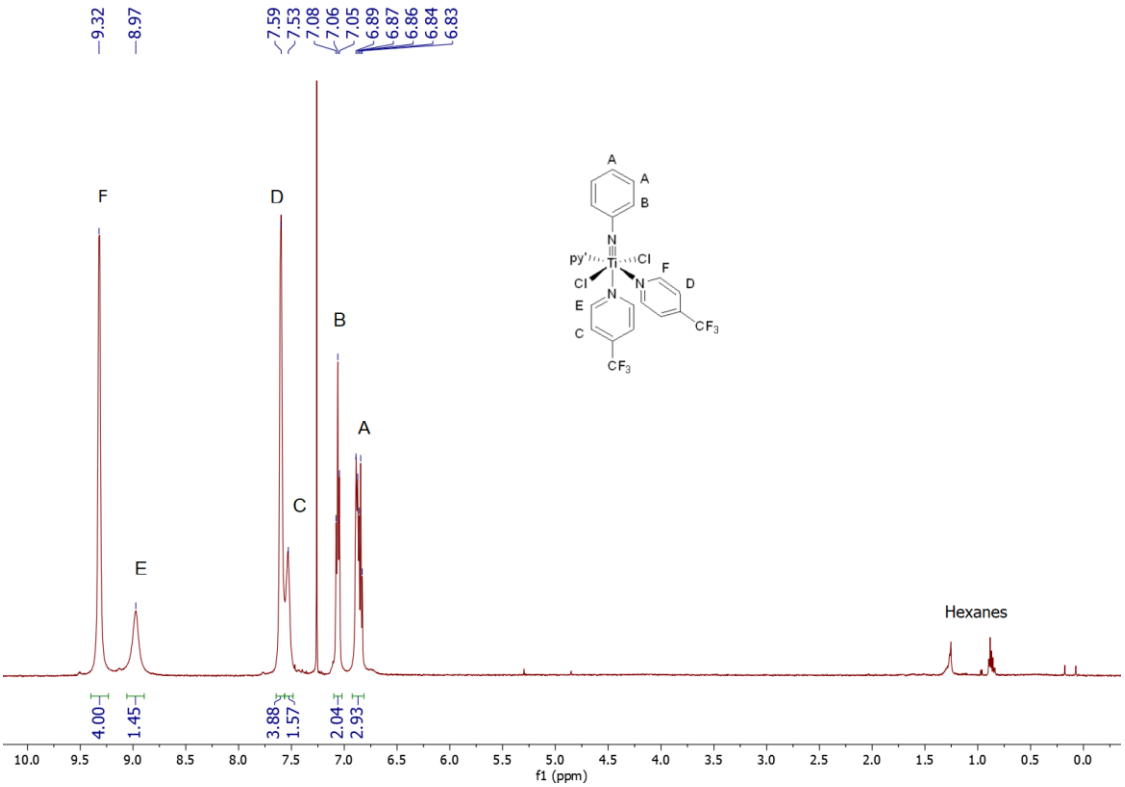


Figure S2. 1H NMR spectrum of **3** in $CDCl_3$.

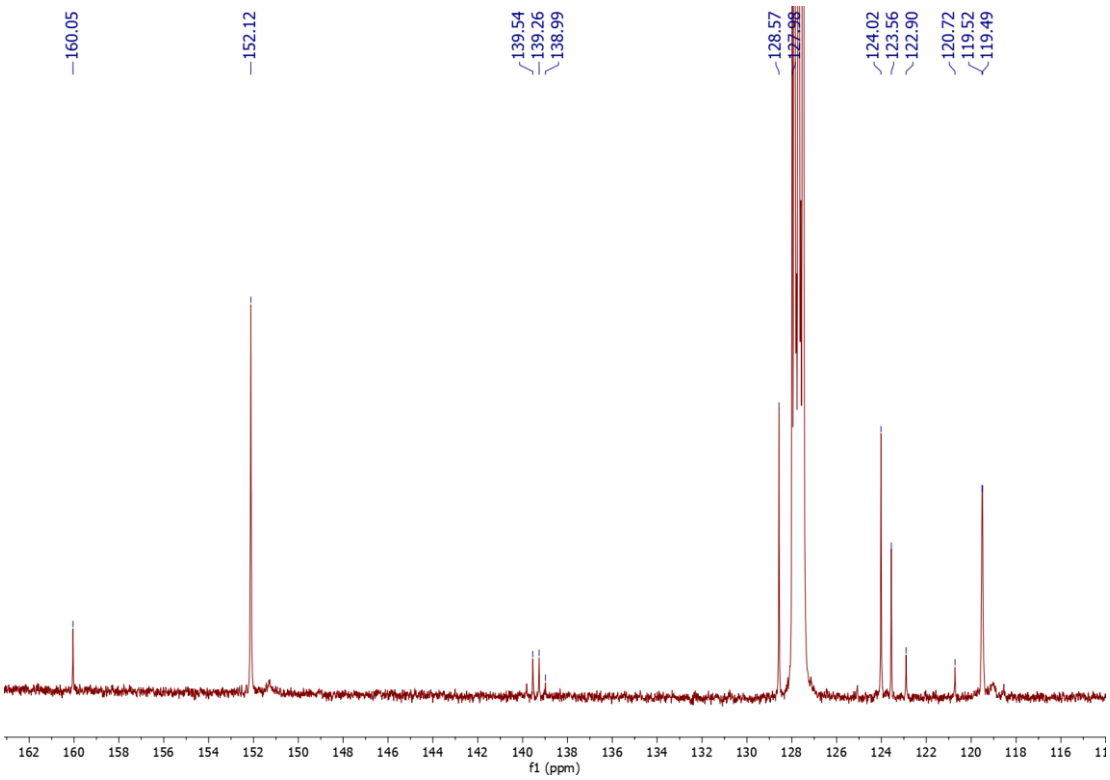


Figure S3. ^{13}C NMR spectrum of **3** in C_6D_6.

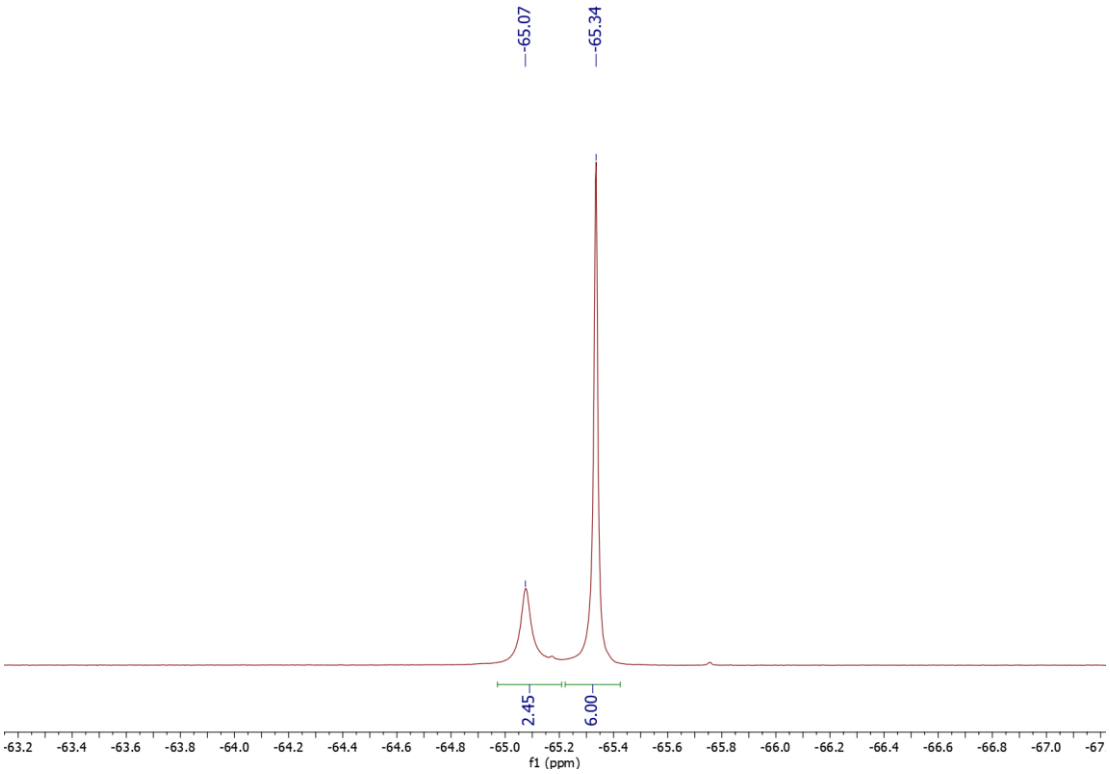
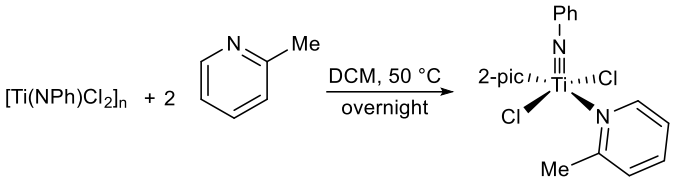


Figure S4. ^{19}F spectrum of **3** in $CDCl_3$.

Synthesis of 4



$[Ti(NPh)Cl_2]_n$ (88 mg, 0.42 mmol, 1.0 equiv), 2-picoline (78 mg, 0.84 mmol, 2.0 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried *in vacuo* to give **4** as a brown powder (125 mg, 75 % yield). Residual CH_2Cl_2 could not be removed despite heating product under *vacuo* at 50 °C for prolonged periods.

^{1}H NMR (500 MHz, $CDCl_3$): δ 9.23 (br s, 2H, *o*-(2-picoline)-*H*), 7.78 (br s, 2H, (2-picoline)-*H*), 7.33 – 7.28 (m, 2H, (2-picoline)-*H*), 7.20 – 7.13 (m, 6H, (2-picoline)-*H* and NPh-*H*), 6.94 (t, $^3J_{HH}$ = 7.2 Hz, 1H, *p*-NPh-*H*), 2.78 (s, 6H, Me-*H*).
Additional peaks in the spectrum belong to rotamers due to restricted rotation around the Ti-N_{py} bond. The peaks are observed to coalesce at 320 K (47 °C) upon heating. Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (101 MHz, $CDCl_3$): δ 161.0, 157.2, 152.8, 139.6, 130.1, 126.4, 124.9, 123.7, 122.4, 24.7.

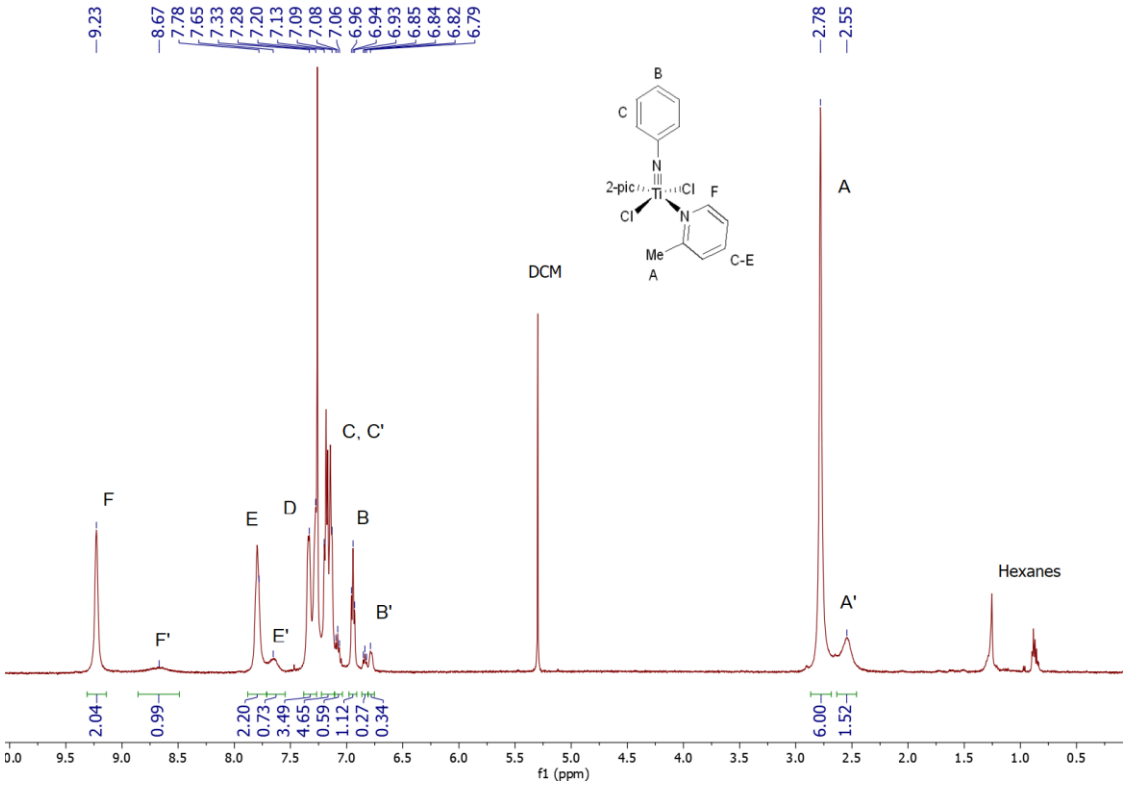


Figure S5. ^{1}H NMR spectrum of **4** in $CDCl_3$.

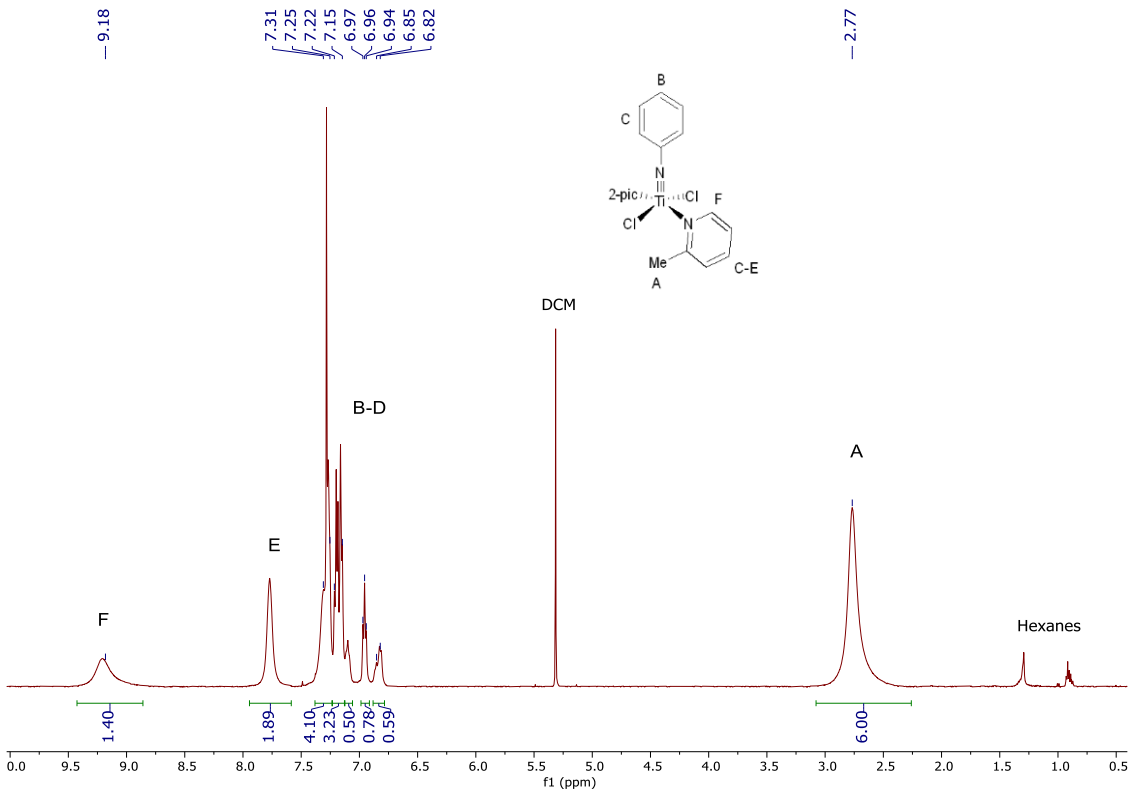


Figure S6. ^{1}H NMR spectrum of **4** in $CDCl_3$ at 320K.

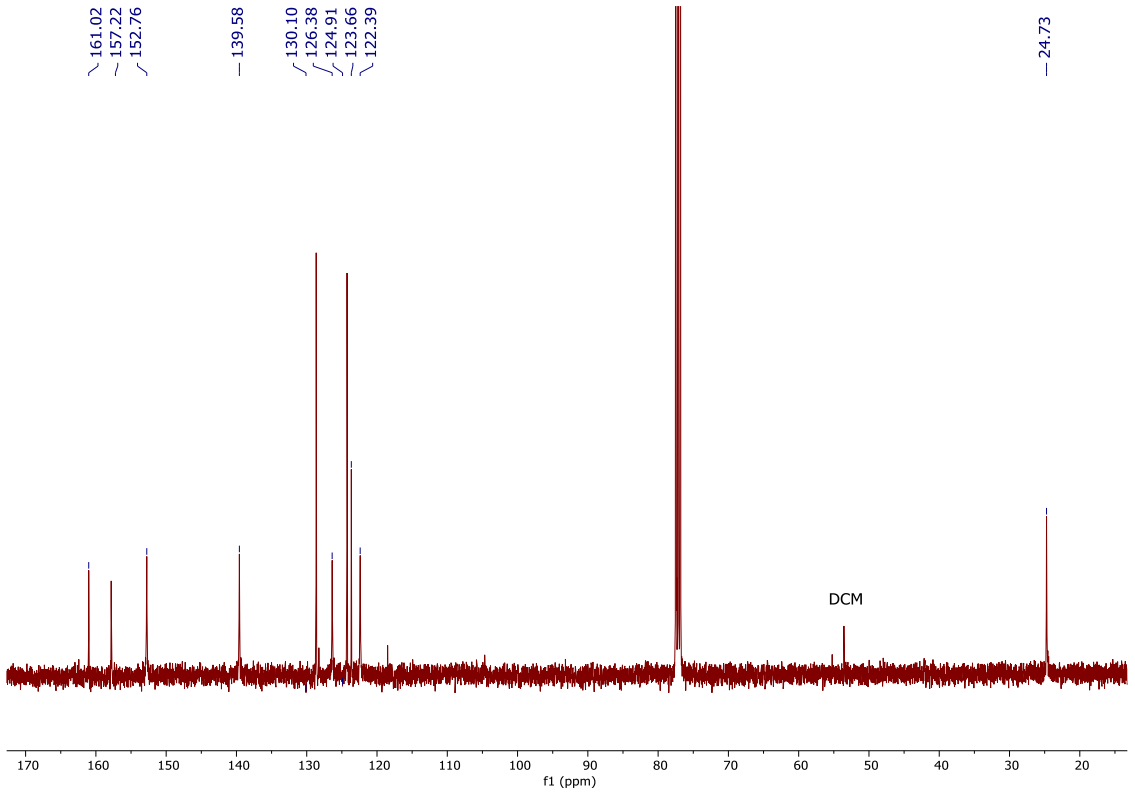
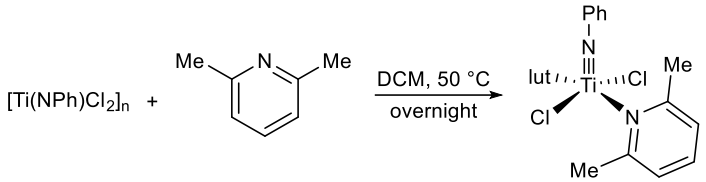


Figure S7. ^{13}C NMR spectrum of **4** in $CDCl_3$.

Synthesis of 5



$[Ti(NPh)Cl_2]_n$ (73 mg, 0.37 mmol, 1.0 equiv), 2,6-lutidine (75 mg, 0.7 mmol, 1.9 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried *in vacuo* to give **5** as a brown powder (97 mg, 75 % yield). **5** is highly insoluble in $CDCl_3$ and no ^{13}C NMR spectrum could be obtained.

^1H NMR (500 MHz, $CDCl_3$): δ 7.50 (br s, 2H, *o*-NPh-*H*), 7.21 (br s, 2H, *p*-lutidine-*H*), 6.95 (app d, 5H, *m*-lutidine-*H* and *p*-NPh-*H*), 6.85 (br s, 2H, *m*-NPh-*H*), 2.56 (s, 12H, Me-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

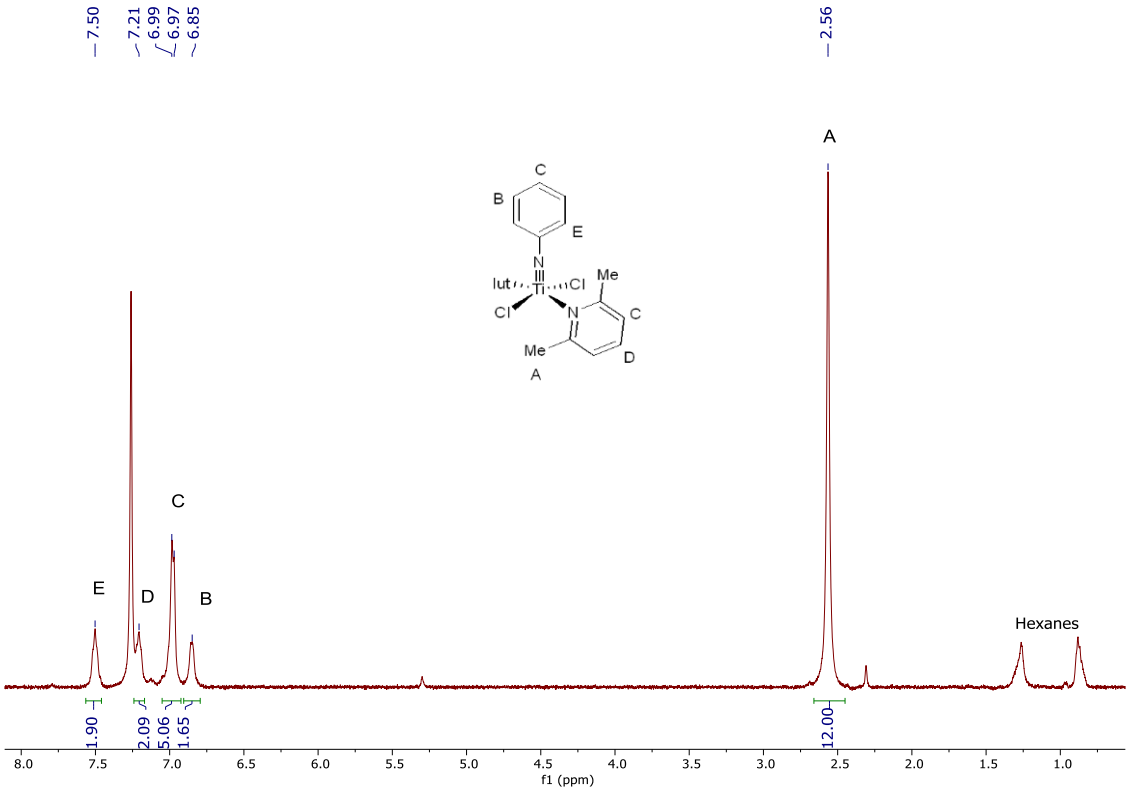
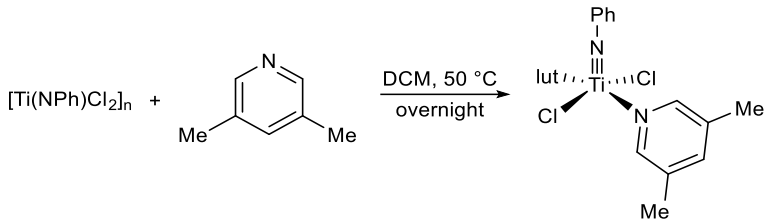


Figure S8. ^1H NMR spectrum of **5** in $CDCl_3$.

Synthesis of 6



$[Ti(NPh)Cl_2]_n$ (156 mg, 0.75 mmol, 1.0 equiv), 3,5-lutidine (170 mg, 1.65 mmol, 2.2 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit. The residue was dried *in vacuo* to give an impure brown powder that was re-suspended in 2 mL hexanes and stirred over the weekend. The suspension was filtered through a fine frit and washed with hexanes. The residue was dried *in vacuo* to give pure **6** as a brown powder (216 mg, 68 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.79 (s, 4H, *o*-lutidine-*H*), 7.52 (s, 2H, *p*-lutidine-*H*), 7.10 (t, $^3J_{HH}$ = 7.6 Hz, 2H, *o*-NPh-*H*), 6.91 (app d, $^3J_{HH}$ = 7.7 Hz, 2H, *m*-NPh-*H*), 6.84 (t, $^3J_{HH}$ = 7.8 Hz, 1H, *p*-NPh-*H*), 2.36 (s, 12H, Me-*H*).
Additional peaks in the spectrum belong to the mono-lutidine complex.

^{13}C NMR (126 MHz, $CDCl_3$): δ 148.2, 139.9, 134.1, 128.4, 123.6, 122.9, 18.7.
This is a partial line-list. Resonances associated with a quaternary C of the NPh fragment were not observed.

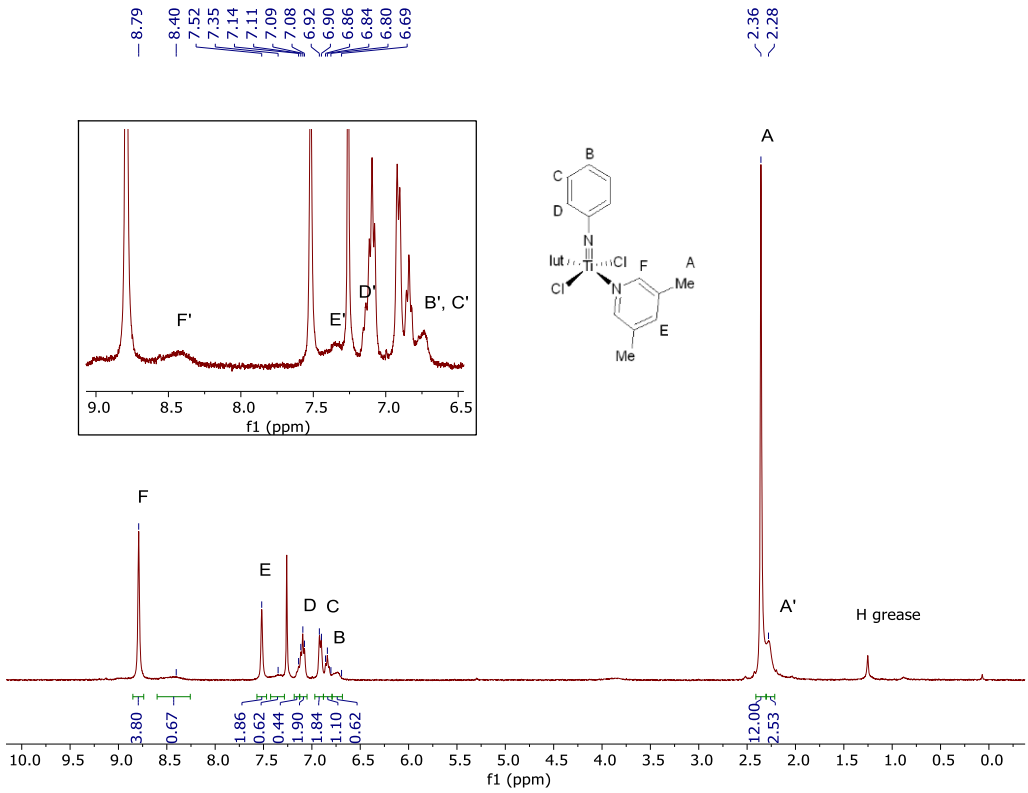


Figure S9. ^{1}H NMR spectrum of **6** in $CDCl_3$.

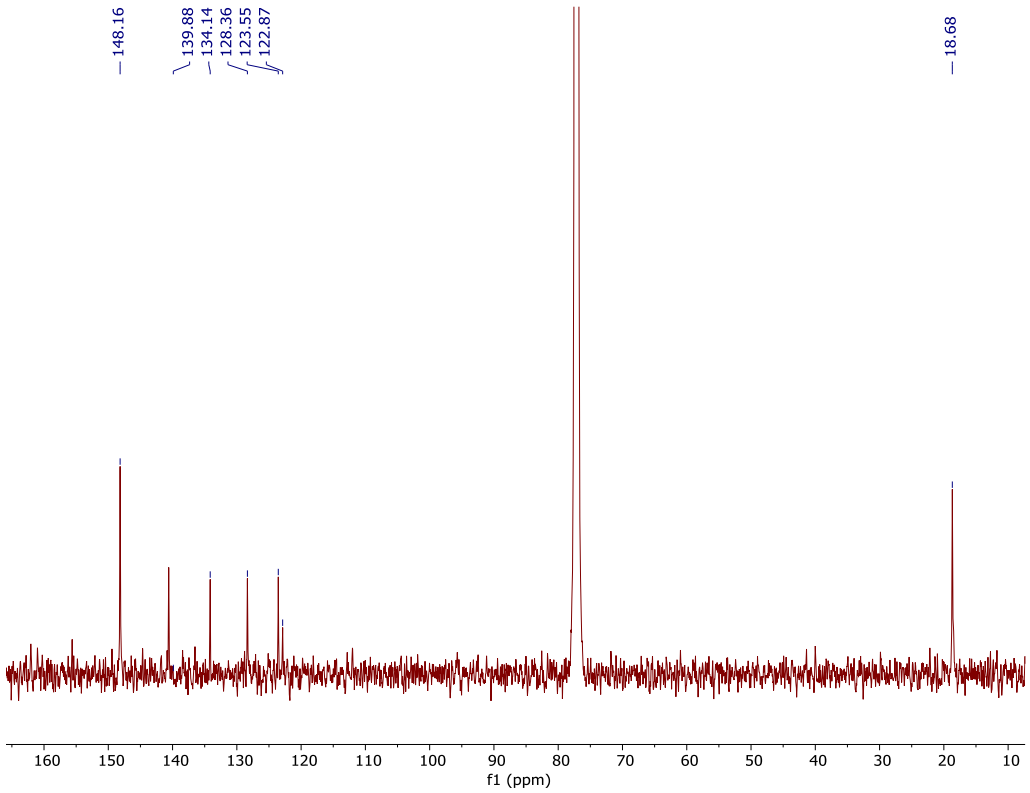
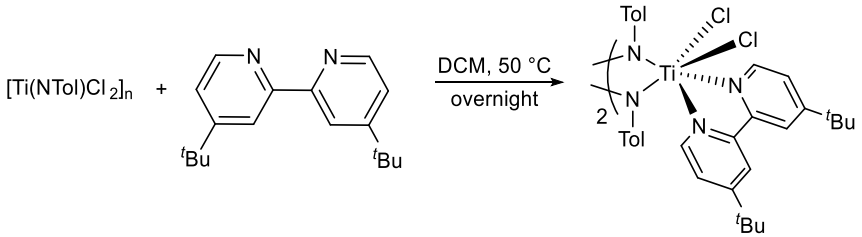


Figure S10. ^{1}H NMR spectrum of **6** in $CDCl_3$.

Synthesis of 7



$[Ti(NTol)Cl_2]_n$ (101 mg, 0.45 mmol, 1.0 equiv), 4,4′-di-tert-butyl-2,2′-bipyridine (121 mg, 0.45 mmol, 1.0 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, this reaction was dried *in vacuo*. The residue was re-dissolved in 2 mL CH_2Cl_2, layered with 2 mL hexanes and cooled in a -35 °C freezer to yield X-ray quality black crystals which were washed with hexanes (180 mg, 81 % yield). The crystalline material yielded a complex solution NMR spectrum with peaks appearing to belong to one molecule *via* various NMR experiments.

^{1}H NMR (500 MHz, $CDCl_3$): δ 9.44 – 9.41 (overlapping d, 3H, bipy*-*H*), 9.35 (d, $^3J_{HH}$ = 5.5 Hz, 1H, bipy*-*H*), 8.72 (d, $^3J_{HH}$ = 6.0 Hz, 2H, bipy*-*H*), 7.92 (s, 2H, bipy*-*H*), 7.80 (s, 2H, bipy*-*H*), 7.72 (app d, $^3J_{HH}$ = 5.8 Hz, 3H, bipy*-*H*), 7.62 (s, 2H, bipy*-H), 7.57 (d, $^3J_{HH}$ = 8.2 Hz, 2H, *o*-NTol-*H*), 7.46 (s, 1H, bipy*-*H*), 7.23 – 7.20 (overlapping d, 3H, *m*-NTol-*H* and bipy*-*H*), 6.71 (d, $^3J_{HH}$ = 6.1 Hz, 2H, bipy*-*H*), 6.55 (br s, 2H, bipy*-*H*), 6.36 (br s, 2H, bipy*-*H*), 6.04 (d, $^3J_{HH}$ = 7.9 Hz,

2H, *o*-NTol-*H*), 5.59 (d, $^3J_{HH}$ = 8.2 Hz, 2H, *m*-NTol-*H*), 2.36 (s, 3H, NTol-*H*), 2.07 (s, 3H, NTol-*H*), 2.59 (s, 3H, NC_6H_4-CH_3), 1.48 (s, 30H, ^{t}Bu-*H*), 1.28 (s, 11H, ^{t}Bu-*H*), 1.15 (s, 18H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, $CDCl_3$): δ 164.2, 163.9, 163.0, 154.86, 153.5, 152.4, 152.2, 151.96, 151.0, 134.6, 128.5, 127.5, 126.4, 123.8, 123.0, 122.2, 121.9, 117.3, 117.1, 114.9, 35.7, 35.2, 35.1, 30.6, 30.4, 30.3, 21.3, 21.0, 20.5.

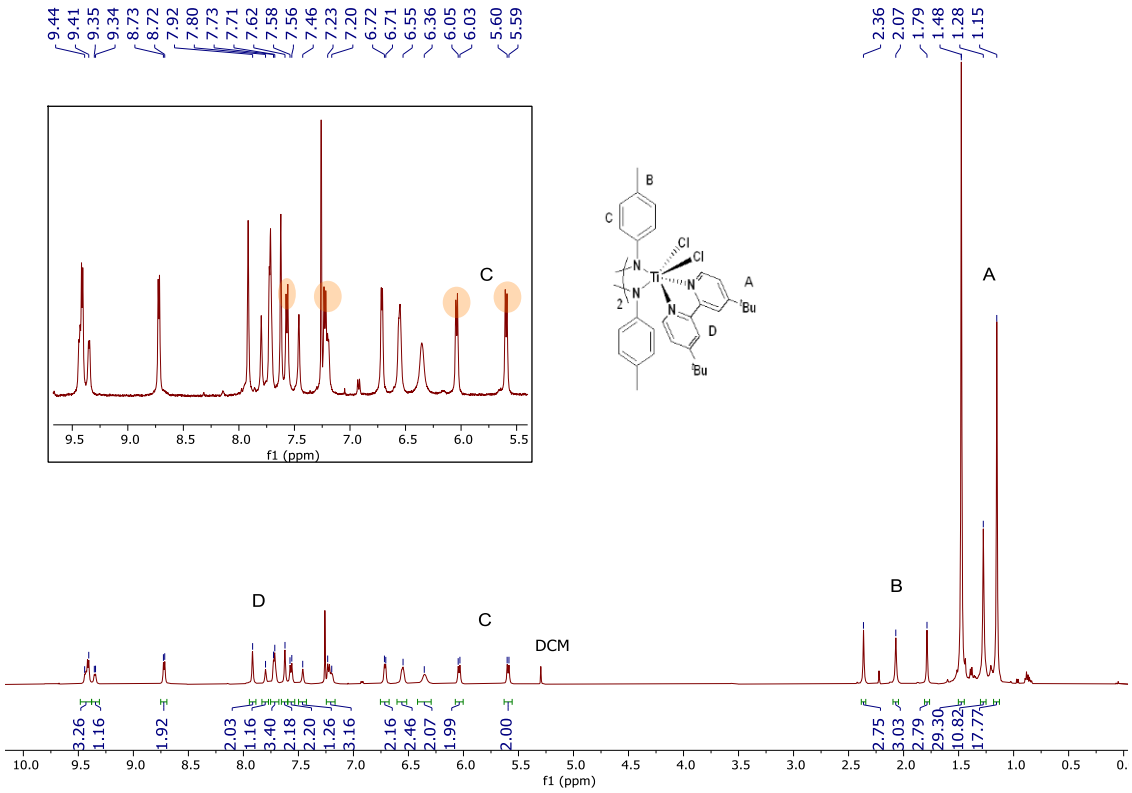


Figure S11. ^{1}H NMR spectrum of **7** in $CDCl_3$.

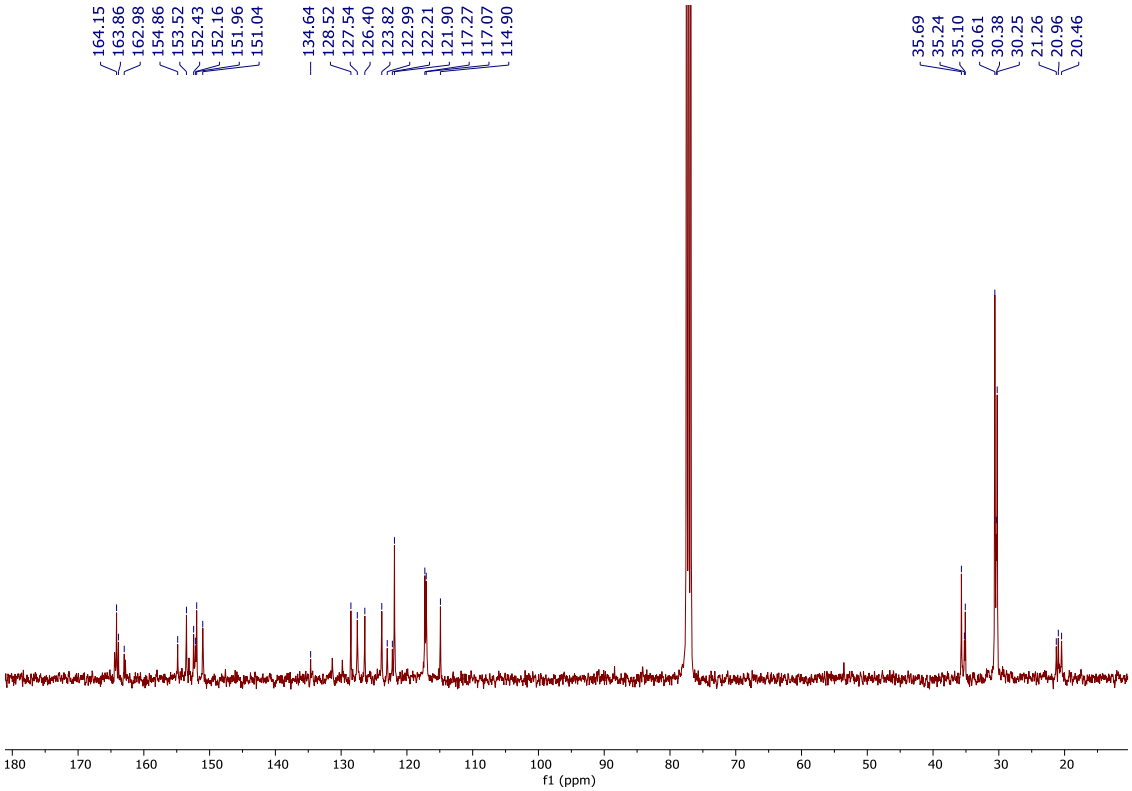


Figure S12. ^{13}C NMR spectrum of **7** in $CDCl_3$.

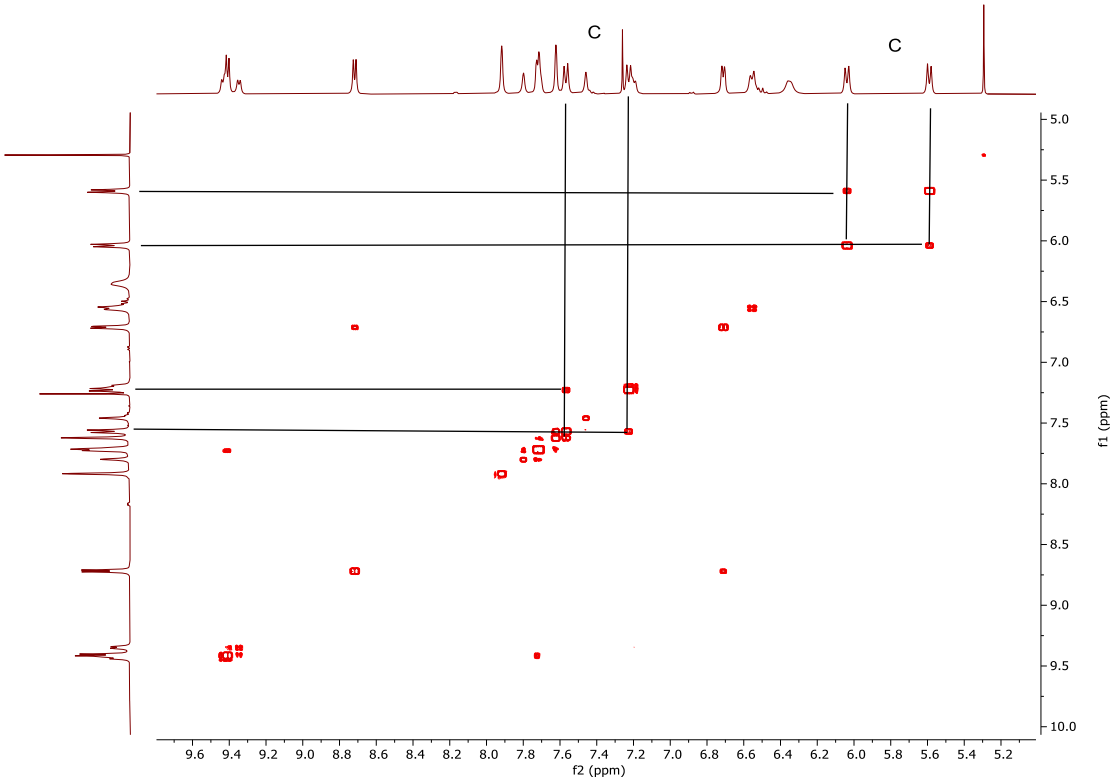


Figure S13. Zoom-in ^{1}H-^{1}H COSY spectrum of **7** in $CDCl_3$.

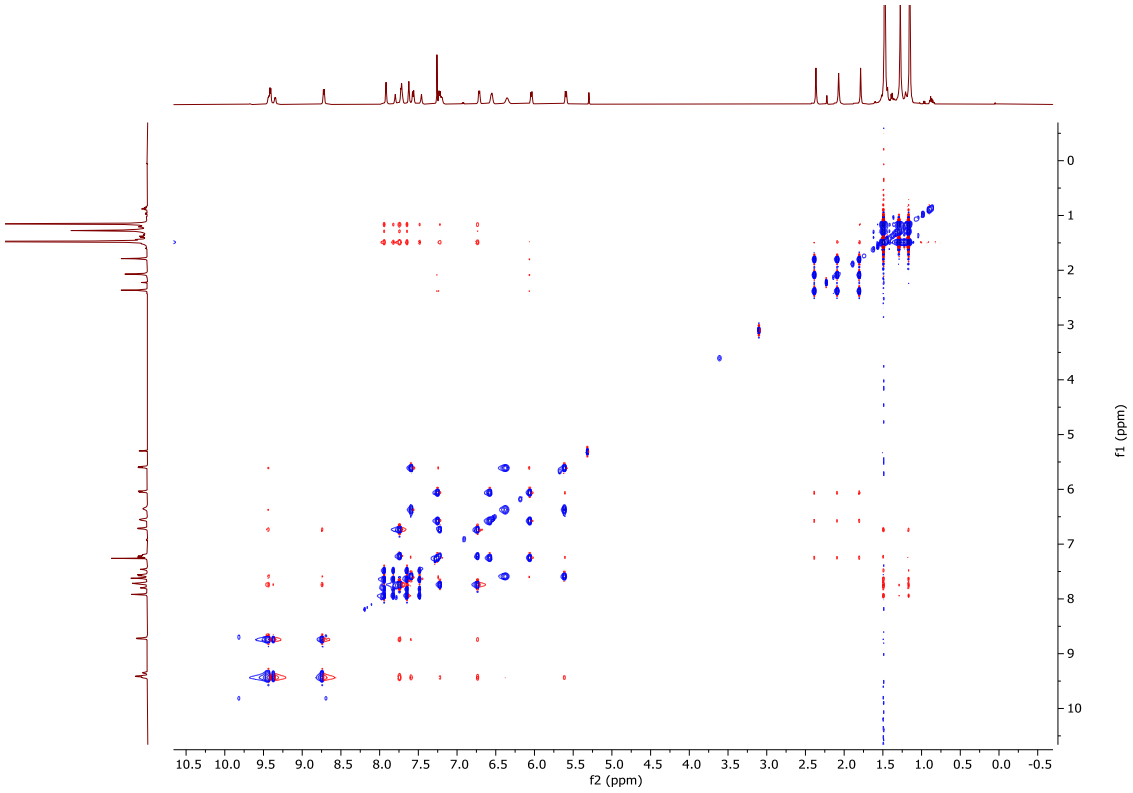


Figure S14. ^{1}H-^{1}H NOESY spectrum of **7** in $CDCl_3$.

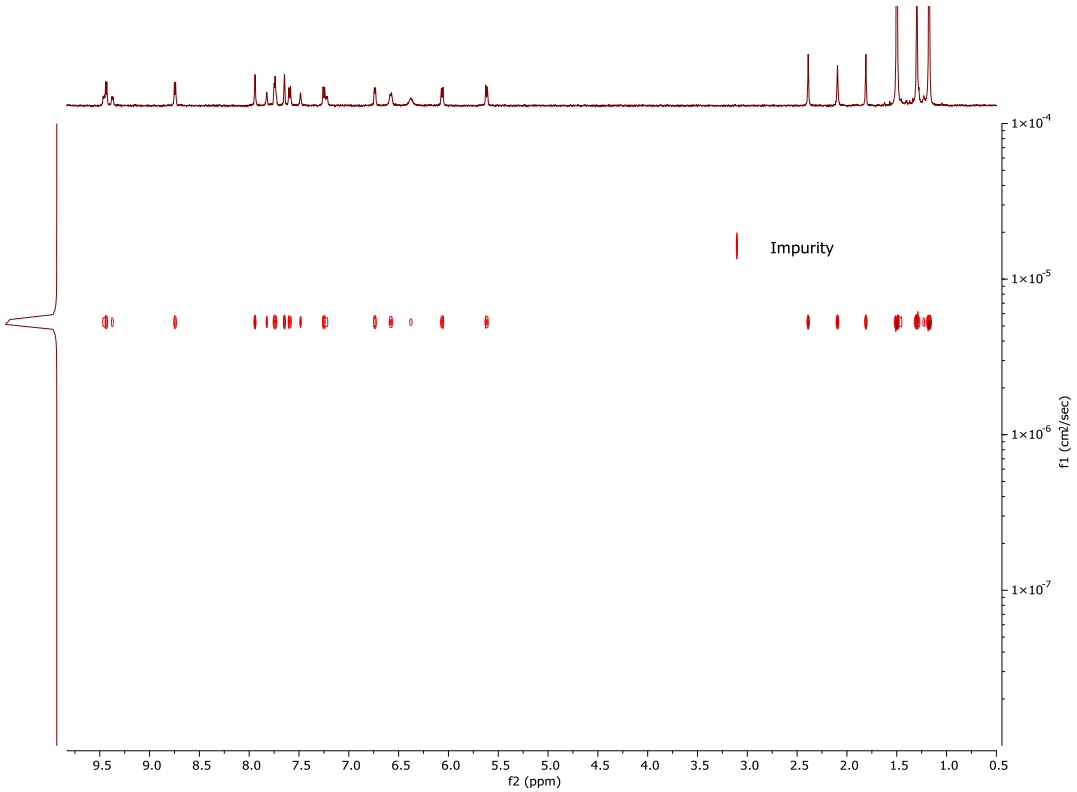


Figure S15. DOSY spectrum of **7** in $CDCl_3$.

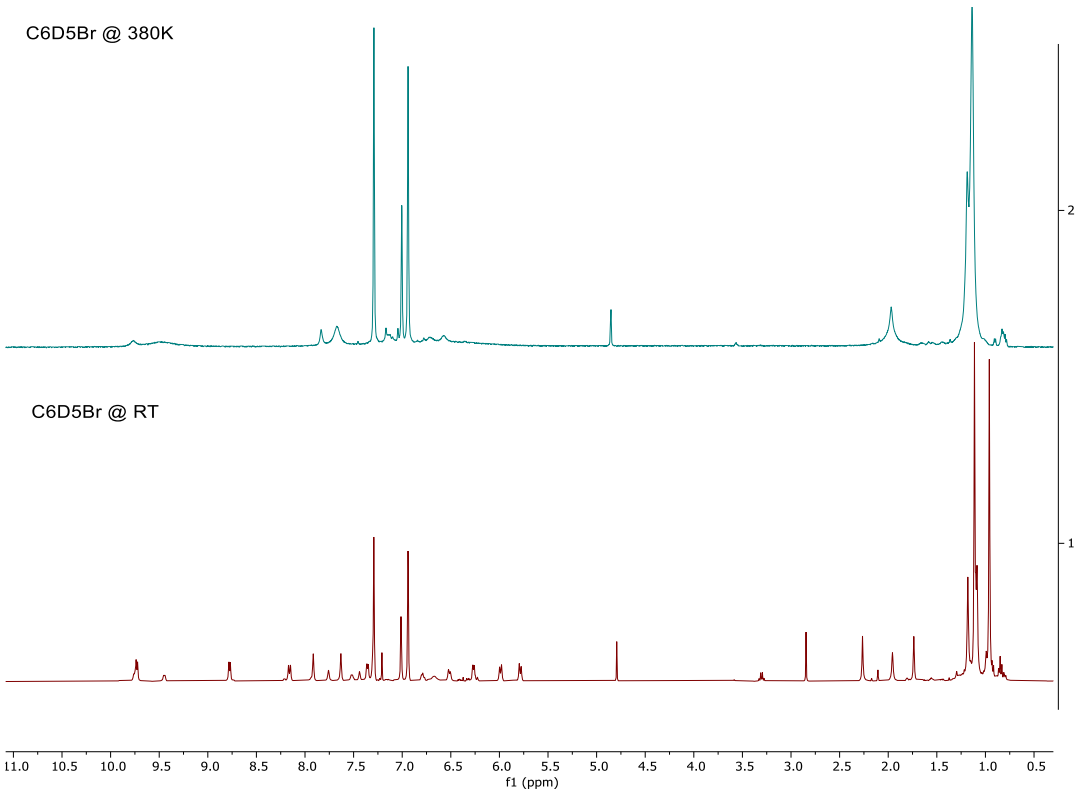


Figure S16. Stacked ^{1}H NMR spectrum of 7 at 380 K (2) and room temperature (1) in C_6D_5Br.

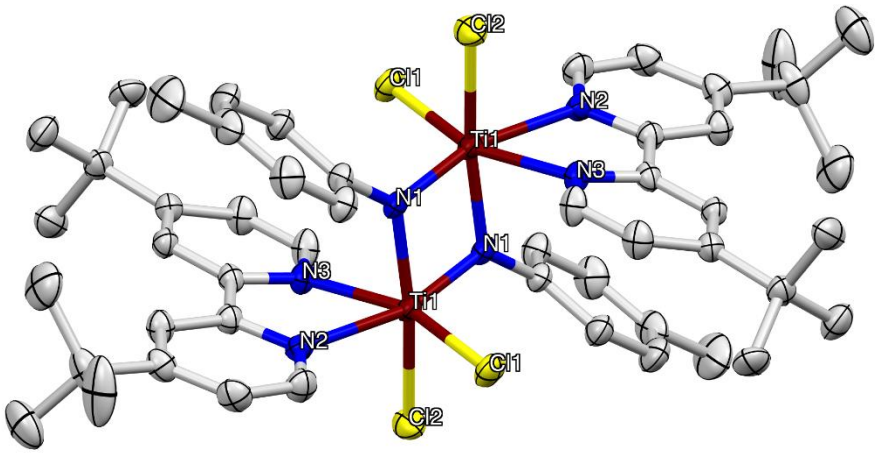
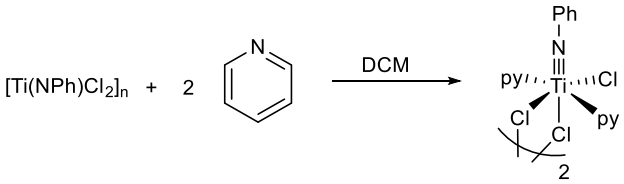


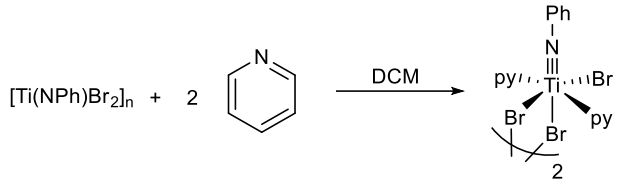
Figure S17. 50 % thermal ellipsoid drawing of **7**. Hydrogen atoms are omitted for clarity.

Synthesis of 8



$[py_2Ti(NPh)Cl_2]_2$ was synthesized with a slight modification to literature procedure by reacting $[Ti(NPh)Cl_2]_n$ with pyridine in the presence of CH_2Cl_2.[3] NMR of complex matches that of reported in literature.

Synthesis of 9



$[Ti(NPh)Br_2]_n$ (658 mg, 2.2 mmol, 1.0 equiv) and 40 mL CH_2Cl_2 were added to a 100 mL round-bottomed flask equipped with a small stir bar in a N_2-filled glovebox. Pyridine (340 mg, 4.4 mmol, 2.0 equiv) was then syringed in all at once into the stirred suspension. The $[Ti(NPh)Br_2]_n$ suspension dissolved and a yellow-brown powder precipitated out shortly after. ~ 40 mL of hexanes were layered to aid in the precipitation. The reaction mixture was filtered through a fine frit and washed with hexanes. The residue was dried *in vacuo* at 40 °C overnight to remove residual CH_2Cl_2 and afford pure **9** as a yellow-brown powder (458 g, 55 % yield). **9** is highly insoluble in $CDCl_3$ and the ^{13}C NMR spectrum was not obtained.

^1H NMR (400 MHz, $CDCl_3$): δ 9.20 (d, $^3J_{HH}$ = 4.9 Hz, 4H, *o*-pyridine-*H*), 7.93 (t, $^3J_{HH}$ = 8.0 Hz, 2H, *p*-pyridine-*H*), 7.51 (t, $^3J_{HH}$ = 6.9 Hz, 4H, *m*-pyridine-*H*), 7.13 (t, $^3J_{HH}$ = 7.6 Hz, 2H, *o*-NPh-*H*), 7.04 (d, $^3J_{HH}$ = 8.3 Hz, 2H, *m*-NPh-*H*), 6.90 (t, $^3J_{HH}$ = 7.0 Hz, 1H, *p*-NPh-*H*).

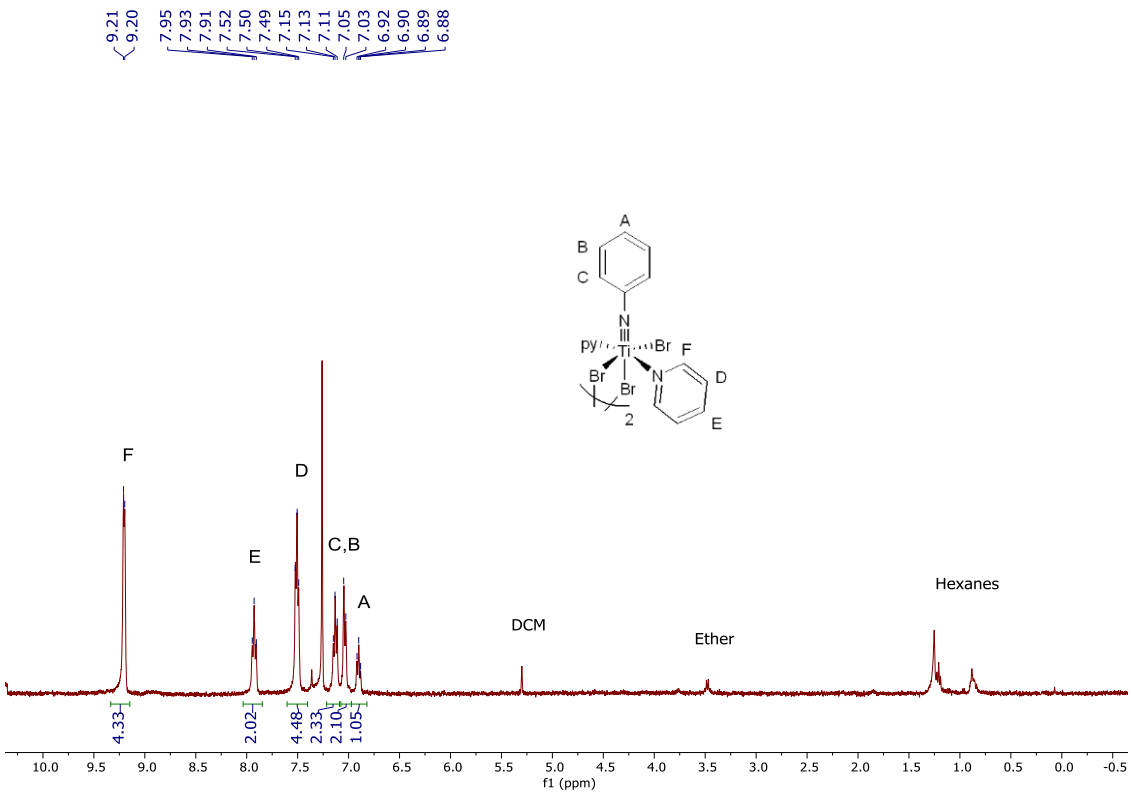
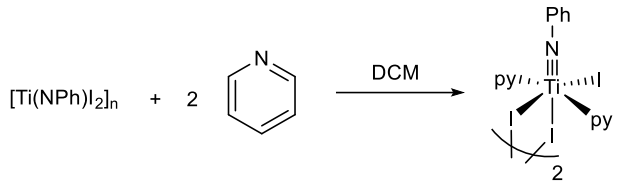


Figure S18. ^1H NMR spectrum of **9** in $CDCl_3$.

Synthesis of 10



$[Ti(NPh)I_2]_n$ (253 mg, 0.64 mmol, 1.0 equiv), 5 mL CH_2Cl_2 and pyridine (102 mg, 1.29 mmol, 2 equiv) were added to a 20 mL scintillation vial in that order in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and shaken by hand. The $[Ti(NPh)I_2]_n$ suspension dissolved and a brown powder precipitated out shortly after. The reaction mixture was left to cool in the freezer overnight to facilitate precipitation. Following which, the reaction mixture was filtered through a fine frit, washed with hexanes (2 x 5 mL) and dried *in vacuo* overnight at 50 °C to give pure **10** as a brown powder (316 mg, 89 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 9.24 (s, 4H, *o*-pyridine-*H*), 7.95 (s, 2H, *p*-pyridine-*H*), 7.50 (s, 4H, *m*-pyridine-*H*), 7.19 (s, 4H, *o*-NPh-*H* and *m*-NPh-*H*), 6.92 (s, 1H, *p*-NPh-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (101 MHz, $CDCl_3$): δ 151.4, 139.3, 129.2, 125.6, 124.4, 124.0.
This is a partial line-list. Resonances associated with a quaternary C of the NPh fragment were not observed.

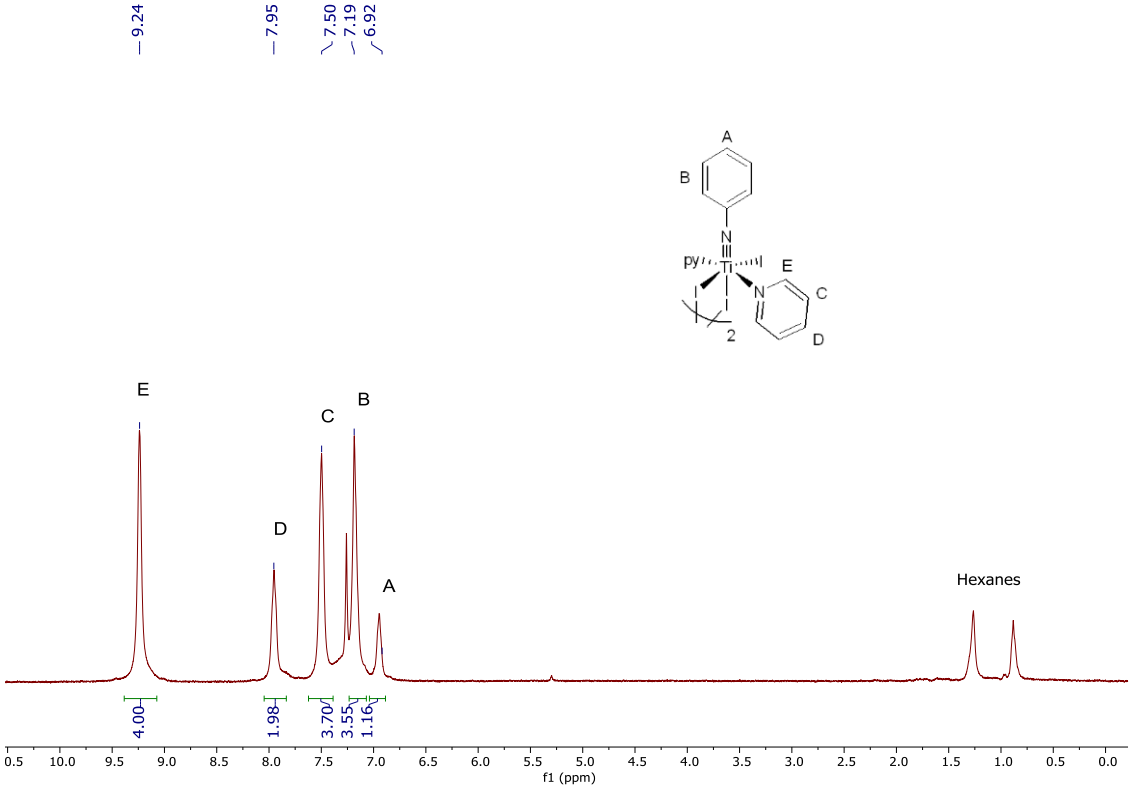


Figure S19. ^{1}H NMR spectrum of **10** in $CDCl_3$.

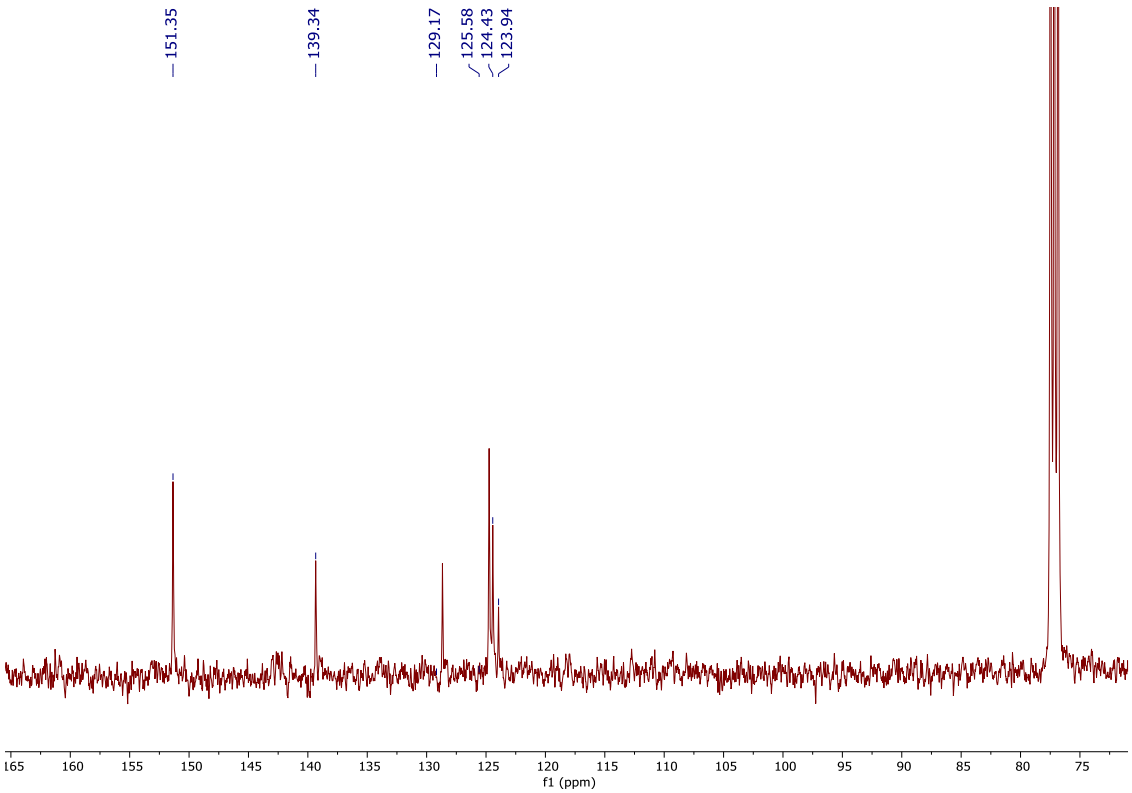
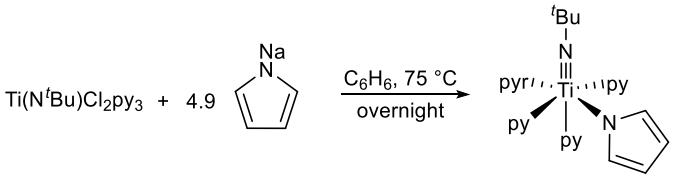


Figure S20. ^{13}C NMR spectrum of **10** in $CDCl_3$.

Synthesis of 11



Sodium pyrrolide was prepared beforehand by deprotonating pyrrole (377 mg, 5.6 mmol, 1 equiv) with excess NaH (400 mg, 16.6 mmol, 3 equiv) in 4 mL THF overnight. The suspension was filtered through celite, dried under vacuo to yield an off-white solid that was used without further purification.

py$_3$Ti(N^{t}Bu)Cl$_2$ (382 mg, 0.89 mmol, 1.0 equiv), sodium pyrrolide (391 mg, 4.3 mmol, 4.9 equiv) and 2 mL C_6H_6 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and heated at 75 °C overnight. After cooling to room temperature, the yellow suspension was filtered through a fine frit to remove a sticky brown residue and the yellow filtrate was dried *in vacuo*. The residue was re-dissolved in 5 mL ether, layered with 5 mL hexanes and left to cool in a -35 °C freezer to yield pure **11** as yellow crystalline material (130 mg, 30 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.53 (br s, 2H, axial *o*-pyridine-*H*), 8.01 (app d, $^3J_{HH}$ = 4.9 Hz, 4H, *o*-pyridine-*H*), 7.47 (s, 4H, *o*-pyr-*H*), 6.98 (t, $^3J_{HH}$ = 7.4 Hz, axial *p*-pyridine-*H*), 6.80 (s, 4H, *m*-pyr-*H*), 6.68 – 6.60 (two overlapping triplets, 4H, axial *m*-pyridine-H and *p*-pyridine-H), 6.33 (t, $^3J_{HH}$ = 7.0 Hz, 4H, *m*-pyridine-*H*), 1.12 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, C_6D_6): δ 151.2, 150.4, 138.1, 135.2, 126.0, 124.1, 123.5, 108.8, 72.8, 31.7.

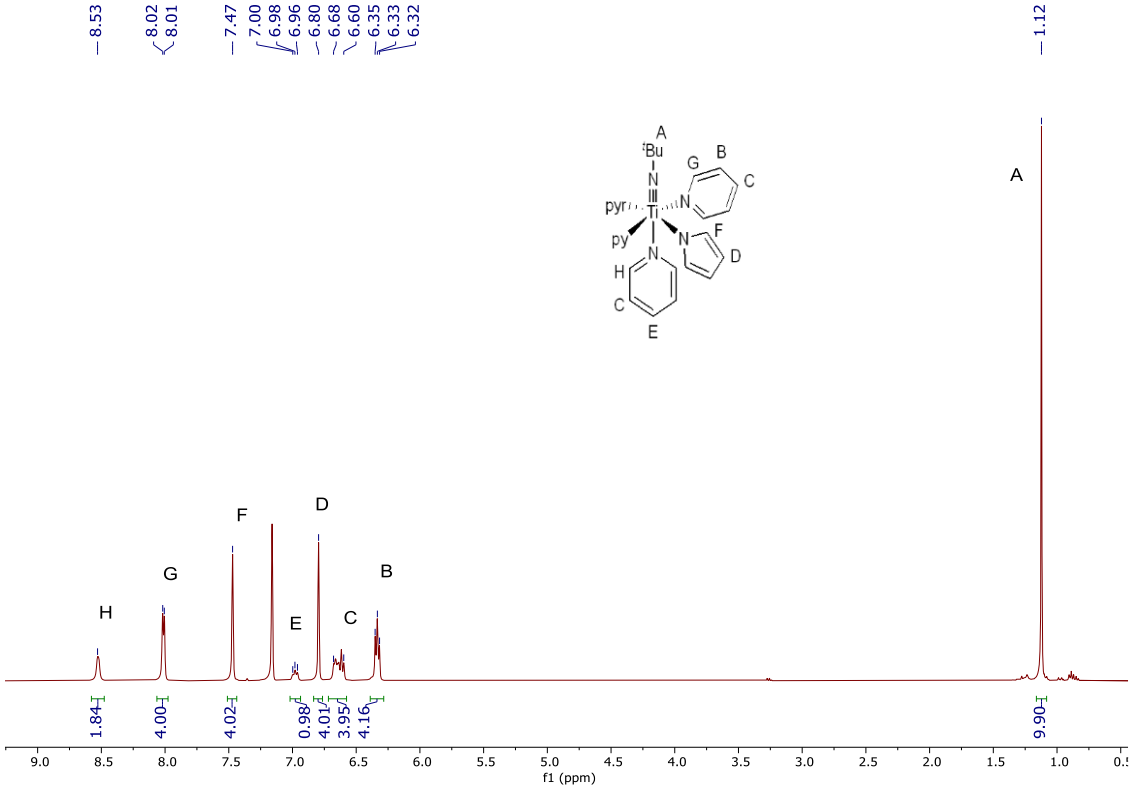


Figure S21. ^{1}H NMR spectrum of **11** in $CDCl_3$.

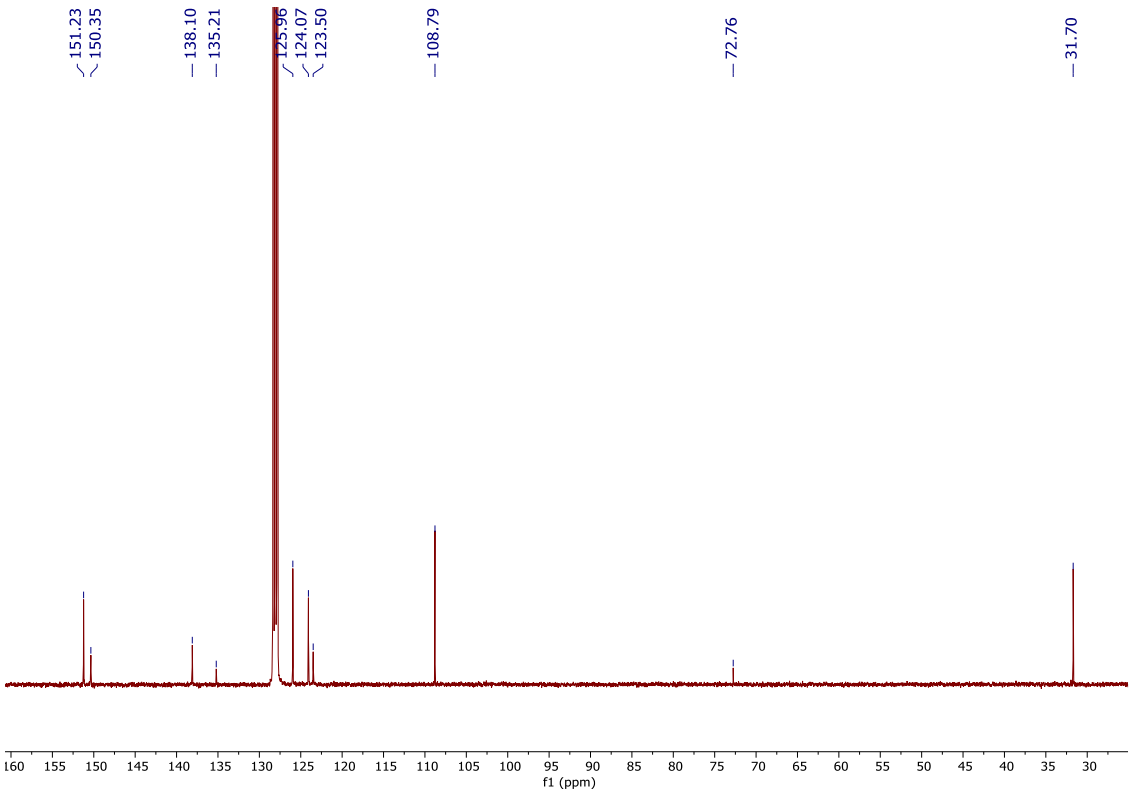
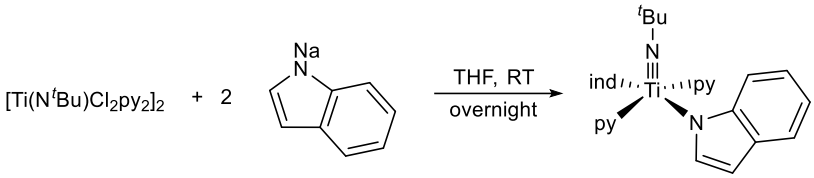


Figure S22. ^{13}C NMR spectrum of **11** in $CDCl_3$.

Synthesis of 12



Indole (270 mg, 2.3 mmol, 2.0 equiv), sodium hexamethyldisilazane (480 mg, 2.6 mmol, 2.2 mmol) and 3 mL THF were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred overnight at room temperature to generate sodium indolide. In a separate vial, [py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (401 mg, 1.15 mmol, 1.0 equiv) was dissolved in 3 mL THF and added to the colourless deprotonated sodium indolide solution. This mixture was stirred overnight at room temperature. Following which, the mixture was dried *in vacuo*, re-dissolved in C_6H_6, filtered through a celite plug and lyophilised *in vacuo*. The crude product was re-dissolved in 6 mL ether and cooled in a -35 °C freezer to yield orange crystalline blocks that contained residual ether. Dissolution of the blocks in C_6H_6 and lyophilisation *in vacuo* yields pure **12** as orange powder (176 mg, 15 % yield). X-ray quality crystals were grown from re-dissolving 20 mg of pure product in 2 mL PhMe, layering with 2 mL hexanes and cooling in a -35 °C freezer to yield yellow cubes.

^{1}H NMR (400 MHz, C_6D_6): δ 8.75 (br s, 4H, *o*-pyridine-*H*), 8.05 (app d, $^3J_{HH}$ = 4.9 Hz, 4H, indole-*H*), 7.99 (d, 2H, indole-*H*), 7.27 (br s, 4H, *m*-pyridine-*H*), 6.96 (br s, 2H, *p*-pyridine-H), 6.44 (app t, $^3J_{HH}$ = 7.6 Hz, 2H, indole-*H*), 6.03 (app t, $^3J_{HH}$ = 7.0 Hz, 4H, indole-*H*), 1.24 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (126 MHz, C_6D_6): δ 150.7, 138.2, 130.6, 128.6, 128.4, 124.2, 120.8, 120.5, 119.6, 103.0, 73. 8, 32.0.
This is a partial line-list. Resonances associated with a quaternary C were not observed.

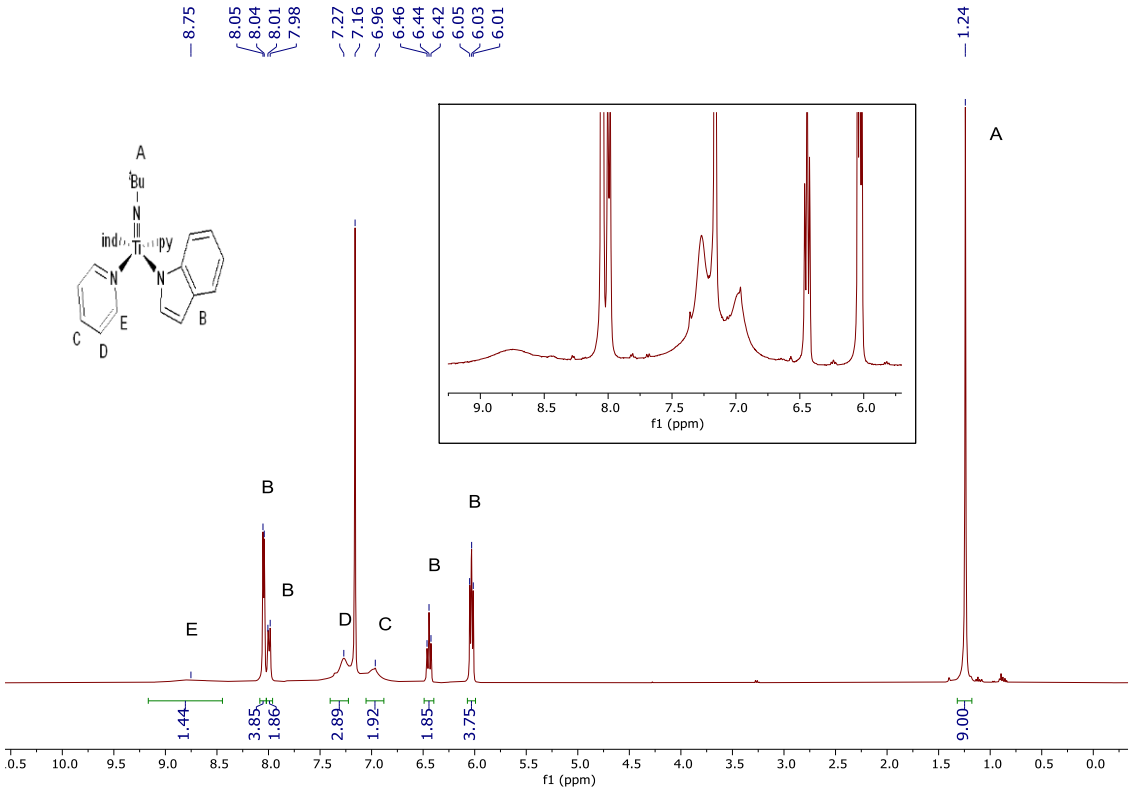


Figure S23. ^1H NMR spectrum of **12** in C_6D_6.

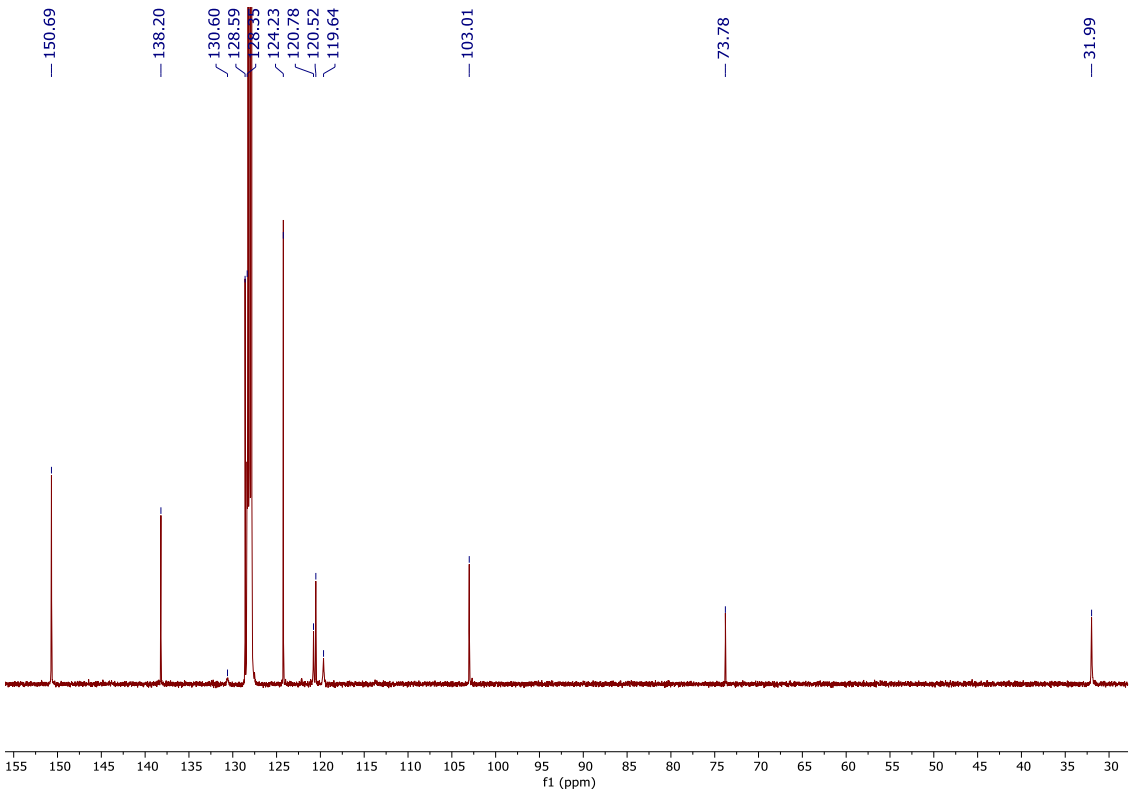


Figure S24. ^{13}C NMR spectrum of **12** in C_6D_6.

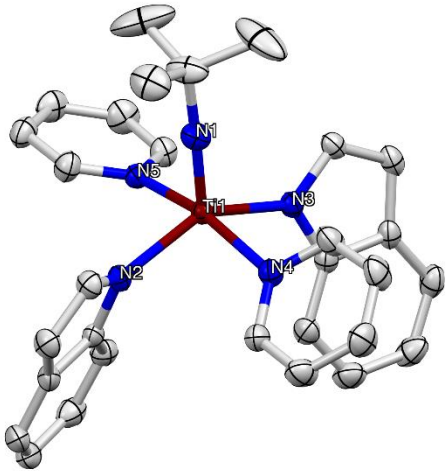
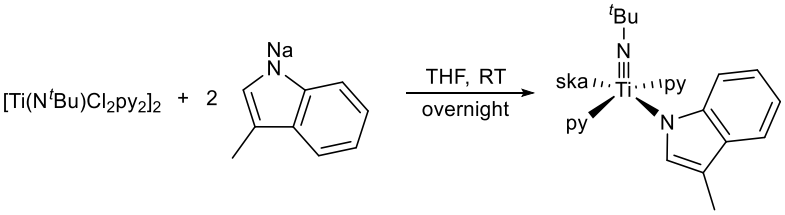


Figure S25. 50 % thermal ellipsoid drawing of **12**. Hydrogen atoms are omitted for clarity.

Synthesis of 13



Skatole (200 mg, 1.5 mmol, 2.0 equiv), sodium hexamethyldisilazane (280 mg, 1.5 mmol, 1.5 mmol) and 2 mL THF were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred overnight at room temperature to generate sodium skatolide. In a separate vial, $[py_2Ti(N^tBu)Cl_2]_2$ (266 mg, 0.76 mmol, 1.0 equiv) was dissolved in 2 mL THF and then added to the colourless sodium skatolide solution. This mixture was stirred overnight at room temperature. Following which, the mixture was dried *in vacuo*, re-dissolved in C_6H_6, filtered through a celite plug and lyophilised *in vacuo*. The crude product was stirred and washed in 10 mL of pentane overnight and filtered through a fine frit. The sticky solid was washed liberally with more pentane and dried *in vacuo* to yield pure **13** as a fluffy, brown powder (178 mg, 33 % yield).

^{1}H NMR (400 MHz, C_6D_6): δ 8.82 – 8.61 (m, 4H, *o*-pyridine-*H*), 8.13 – 8.11 (m, 4H, skatole-*H*), 7.89 – 7.87 (m, 2H, skatole-*H*), 7.29 (br s, 6H, *m,p*-pyridine-*H* (partial overlap with C_6H_6 peak)), 6.43 (app t, $^3J_{HH}$ = 7.6 Hz, 2H, skatole-*H*), 6.03 (app t, $^3J_{HH}$ = 6.9 Hz, 4H, skatole-*H*), 2.65 (br s, 3H, skatole-CH_3), 1.29 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, C_6D_6): δ 150.8, 138.1, 124.2, 120.8, 119.0, 118.6, 111.3, 73.4, 32.2, 10.6. This is a partial line-list. Four Cs could not be identified.

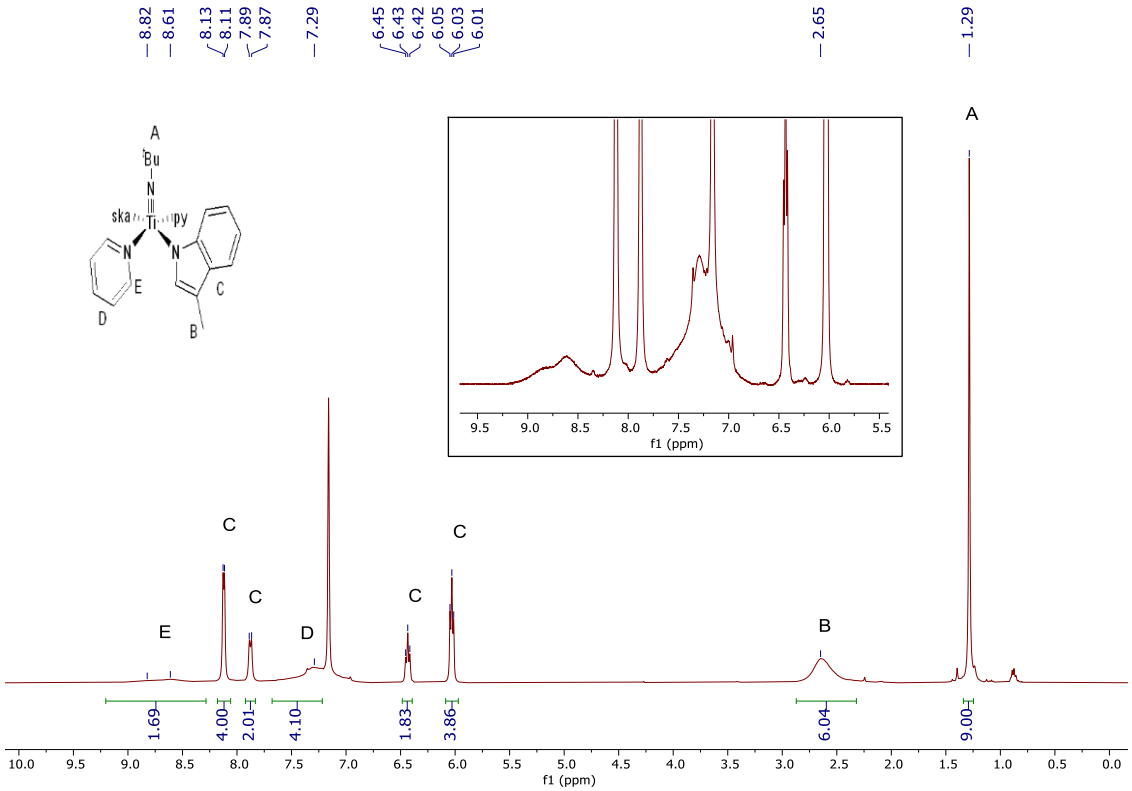


Figure S26. ^{1}H NMR spectrum of **13** in C_6D_6.

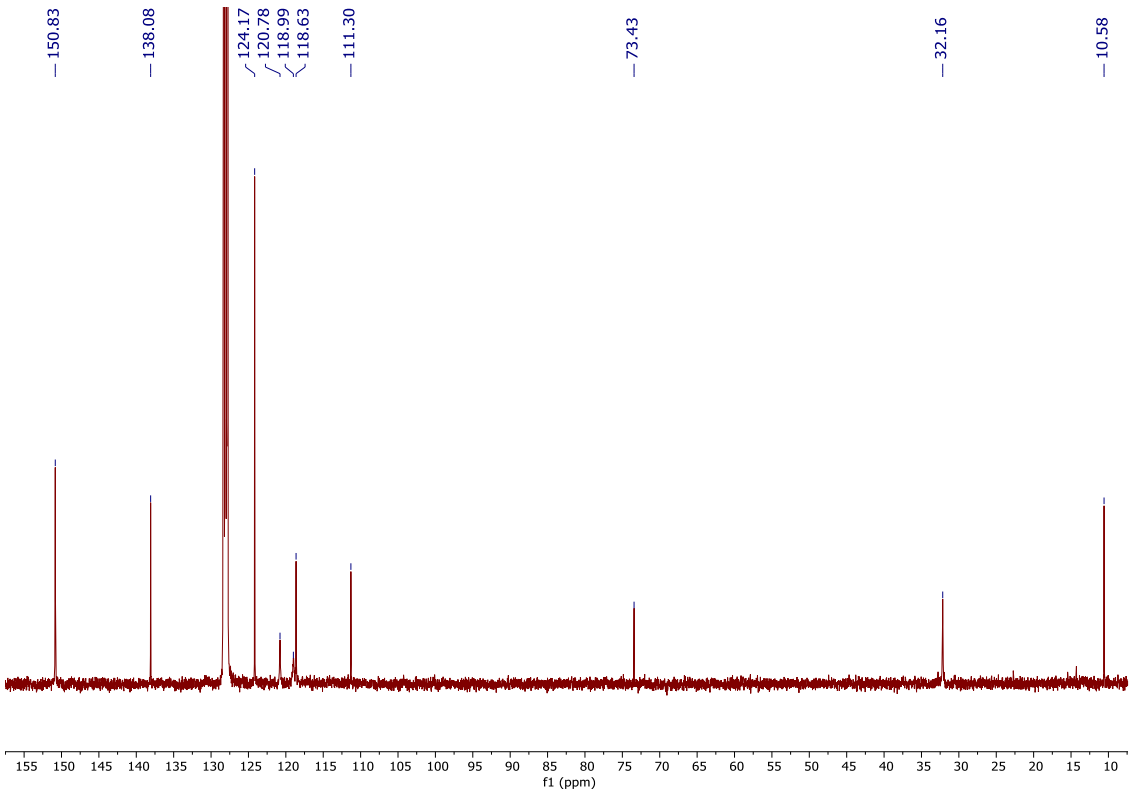
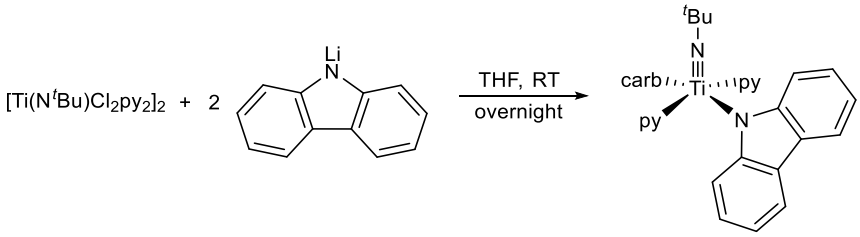


Figure S27. ^{13}C NMR spectrum of **13** in C_6D_6.

Synthesis of 14



[py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (1.01 g, 28.7 mmol, 1.0 equiv), lithium carbazolide (1.0 g, 57.8 mmol, 2.0 equiv) and 10 mL C_6H_6 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred overnight. The resulting suspension was filtered through a celite plug on a coarse frit and the filtrate was dried *in vacuo* to yield pure **14** as orange powder (1.50 g, 85 % yield). The powder was re-dissolved in 5 mL ether, layered with 5 mL hexanes and left to cool in a -35 °C freezer to yield pure **14** as yellow crystalline material (1.50 g, 86 % yield). X-ray quality crystals were grown from re-dissolving 100 mg of pure product in 3 mL C_6H_6, layering with 3 mL hexanes at room temperature to yield orange crystals.

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.47 (d, $^3J_{HH}$ = 8.2 Hz, 2H, Ar-*H*), 8.36 (app d, $^3J_{HH}$ = 4.9 Hz, 4H, Ar-*H*), 8.16 (d, $^3J_{HH}$ = 7.7 Hz, 2H, Ar-*H*), 8.06 (d, $^3J_{HH}$ = 7.0 Hz, 2H, Ar-*H*), 7.57 (t, $^3J_{HH}$ = 7.6 Hz, 2H, Ar-*H*), 7.40 (t, $^3J_{HH}$ = 7.6 Hz, 2H, Ar-*H*), 7.26 – 7.23 (m, 2H, Ar-*H*), 7.02 (t, $^3J_{HH}$ = 6.7 Hz, 4H, , Ar-*H*), 6.96 (t, $^3J_{HH}$ = 7.3 Hz, 2H, Ar-*H*), 6.82 (t, $^3J_{HH}$ = 7.6 Hz, 2H, Ar-*H*), 6.68 (d, $^3J_{HH}$ = 8.2 Hz, 2H, Ar-*H*), 1.18 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, $CDCl_3$): δ 150.9, 149.6, 147.5, 138.7, 126.0, 125.0, 124.9, 124.5, 124.5, 124.1, 120.5, 119.7, 119.6, 119.5, 118.3, 118.0, 117.5, 113.8, 110.7, 75.1, 32.2.

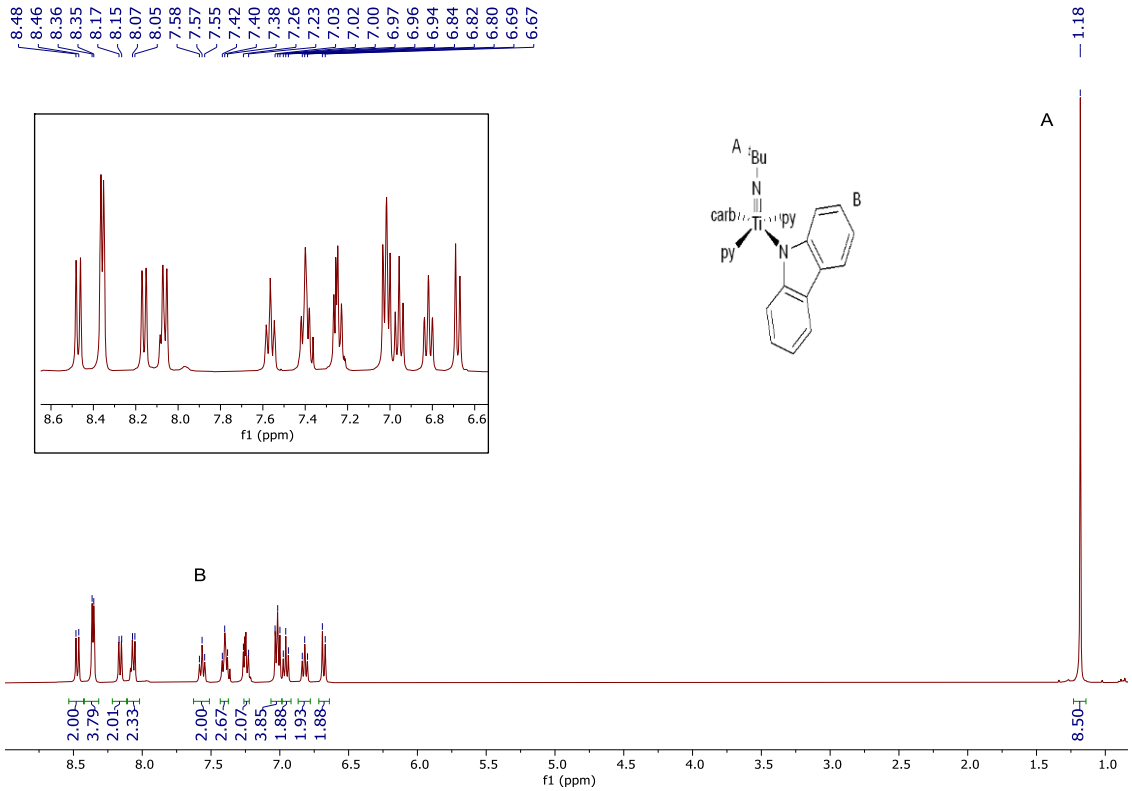


Figure S28. ^{1}H NMR spectrum of **14** in $CDCl_3$.

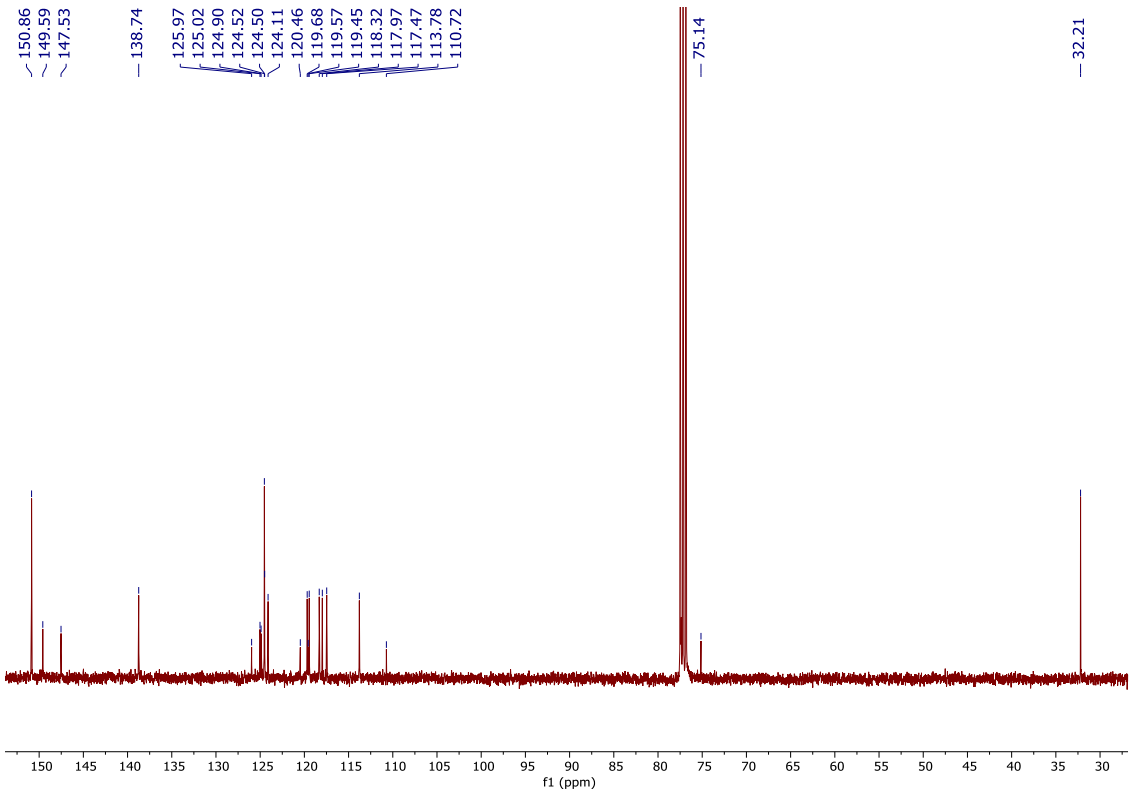


Figure S29. ^{13}C NMR spectrum of **14** in $CDCl_3$.

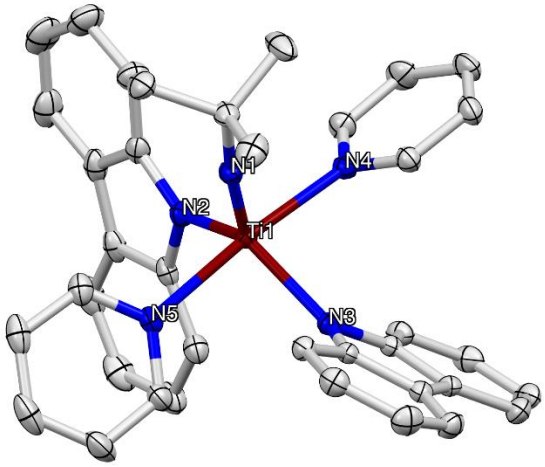
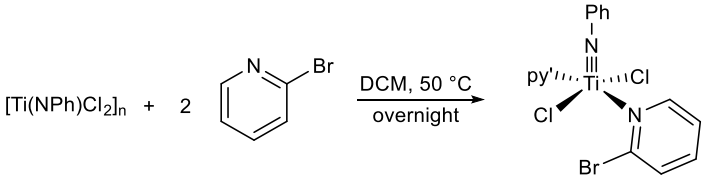


Figure S30. 50 % thermal ellipsoid drawing of **14**. Hydrogen atoms are omitted for clarity.

Synthesis of 15



$[Ti(NPh)Cl_2]_n$ (51 mg, 0.24 mmol, 1.0 equiv), 2-bromopyridine (89 mg, 0.56 mmol, 2.3 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried *in vacuo* to give **15** as a brown powder (70 mg, 55 % yield). Product is highly insoluble in $CDCl_3$ and the ^{13}C NMR spectrum was not obtained.

^1H NMR (500 MHz, $CDCl_3$): δ 8.67 (d, $^3J_{HH}$ = 8.2 Hz, 2H, 2-bromopyridine-6*H*), 7.62 – 7.61 (br t, 2H, 2-bromopyridine-4*H*), 7.47 (d, $^3J_{HH}$ = 8.7 Hz, 2H, 2-bromopyridine-3*H*), 7.37 (br s, 2H, 2-bromopyridine-5*H*), 7.17 (br s, 2H, NPh-*H*), 6.93 – 6.91 (m, 3H, NPh-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

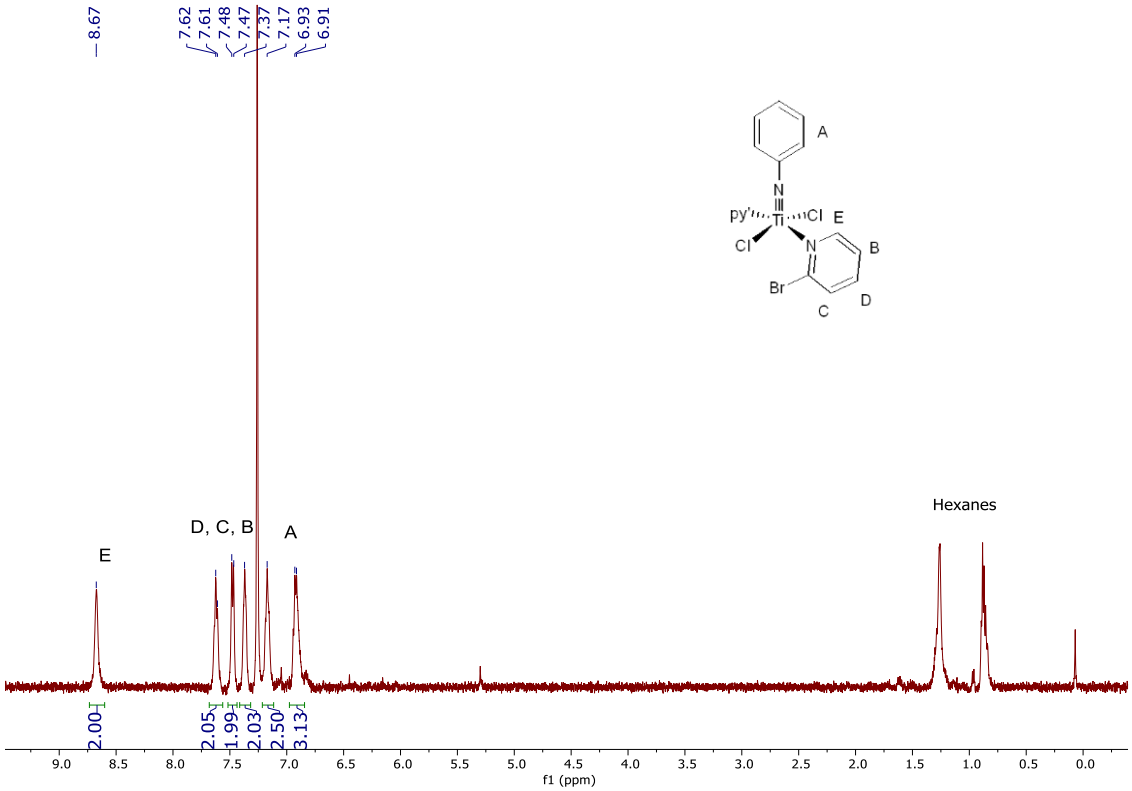
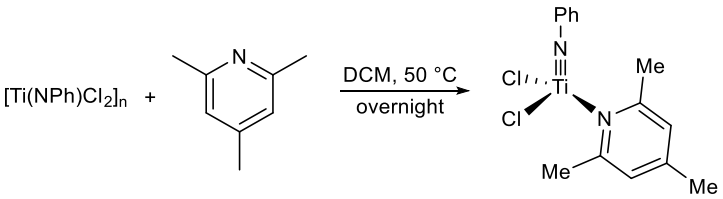


Figure S31. ^1H NMR spectrum of **15** in $CDCl_3$.

Synthesis of 16



$[Ti(NPh)Cl_2]_n$ (70 mg, 0.33 mmol, 1.0 equiv), 2,4,6-trimethylpyridine (45 mg, 0.37 mmol, 1.1 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and

stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried *in vacuo* to give **16** as a brown powder (83 mg, 75 % yield).

^{1}H NMR (500 MHz, $CDCl_3$): δ 7.16 (t, $^3J_{HH}$ = 7.8 Hz, 2H, *o*-NPh-*H*), 6.94 (d, $^3J_{HH}$ = 7.5 Hz, 1H, *p*-NPh-*H*), 6.82 – 6.80 (m, 4H, *m*-NPh-*H* and py-*H*), 2.56 (s, 6H, *o*-py-C*H*$_3$), 2.29 (s, 3H, *p*-py-C*H*$_3$). Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (101 MHz, $CDCl_3$): δ 156.2, 133.3, 128.2, 122.9, 117.9, 24.4, 20.8.
This is a partial line-list. Due to the insolubility of the material, two Cs could not be identified.

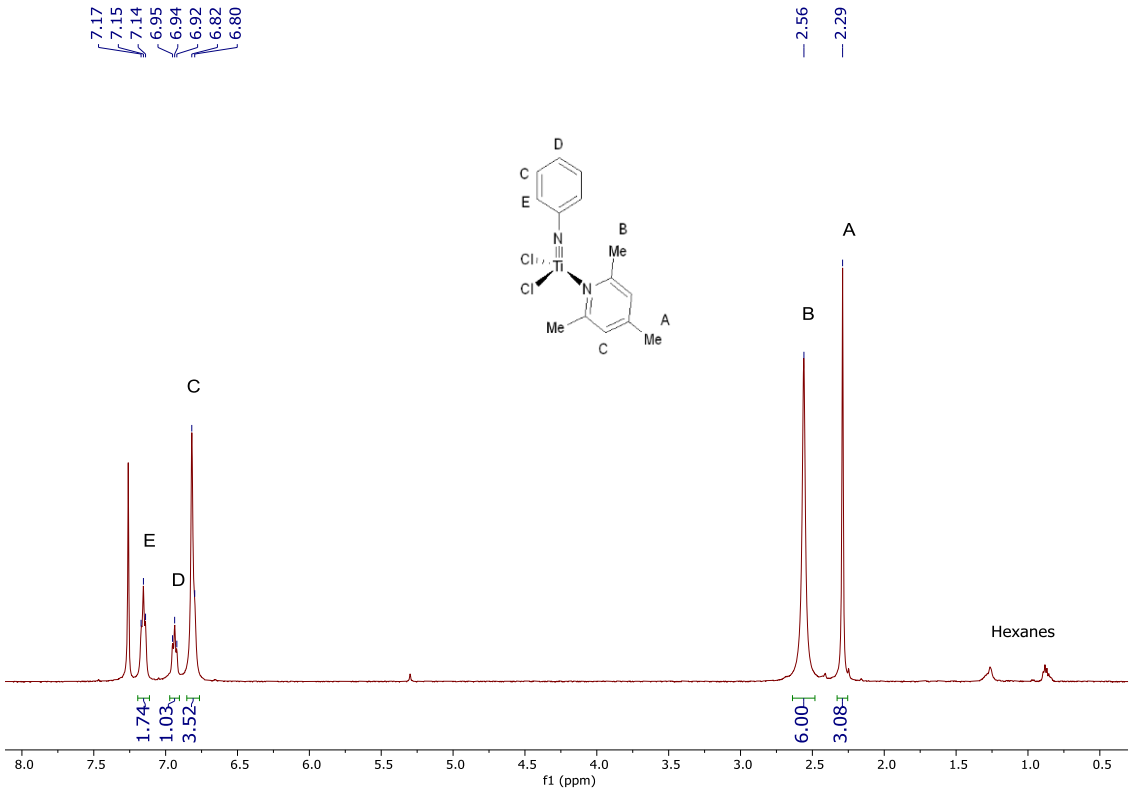


Figure S32. ^{1}H NMR spectrum of **16** in $CDCl_3$.

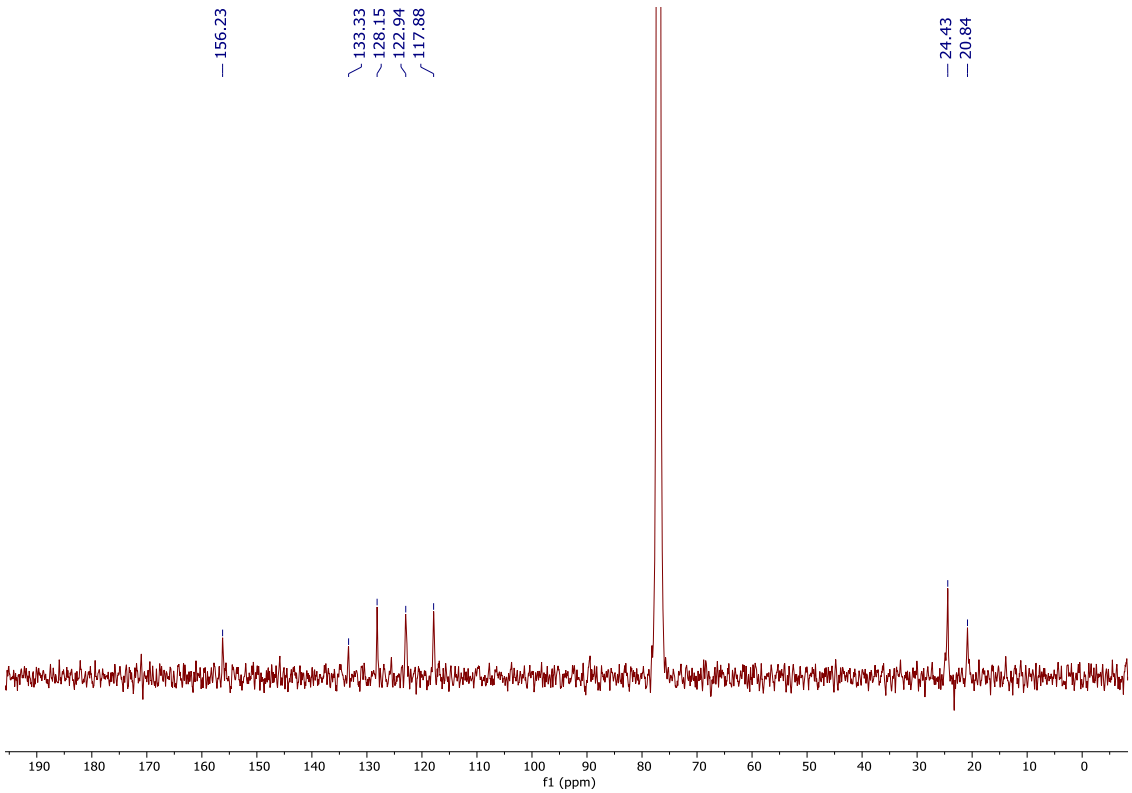
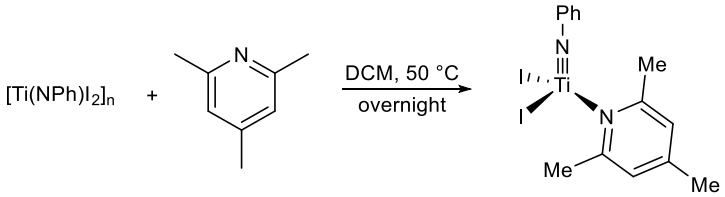


Figure S33. ^{13}C NMR spectrum of **16** in $CDCl_3$.

Synthesis of 17



$[Ti(NPh)I_2]_n$ (100 mg, 0.26 mmol, 1.0 equiv), 2,4,6-trimethylpyridine (34 mg, 0.28 mmol, 1.1 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried under *vacuo* to give **17** as a black powder (74 mg, 57 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.29 (br t, 4H, NPh-*H*), 7.11 (br s, 1H, *p*-NPh-*H*), 6.85 (s, 2H, py-*H*), 2.46 (s, 6H, *o*-py-C*H*$_3$), 2.32 (s, 3H, *p*-py-C*H*$_3$).
Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (101 MHz, $CDCl_3$): δ 129.1, 126.9, 120.4, 24.1, 21.2.
This is a partial line-list. Due to the insolubility of the material, four Cs could not be identified.

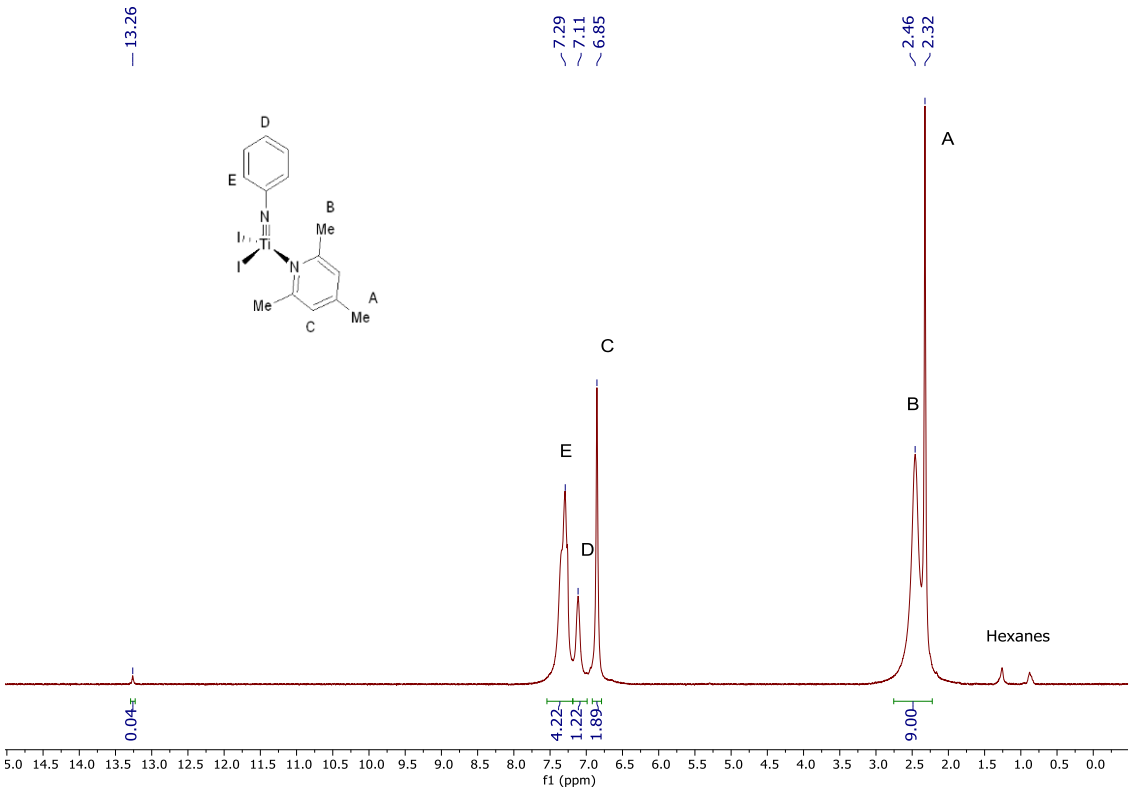


Figure S34. ^{1}H NMR spectrum of **17** in $CDCl_3$.

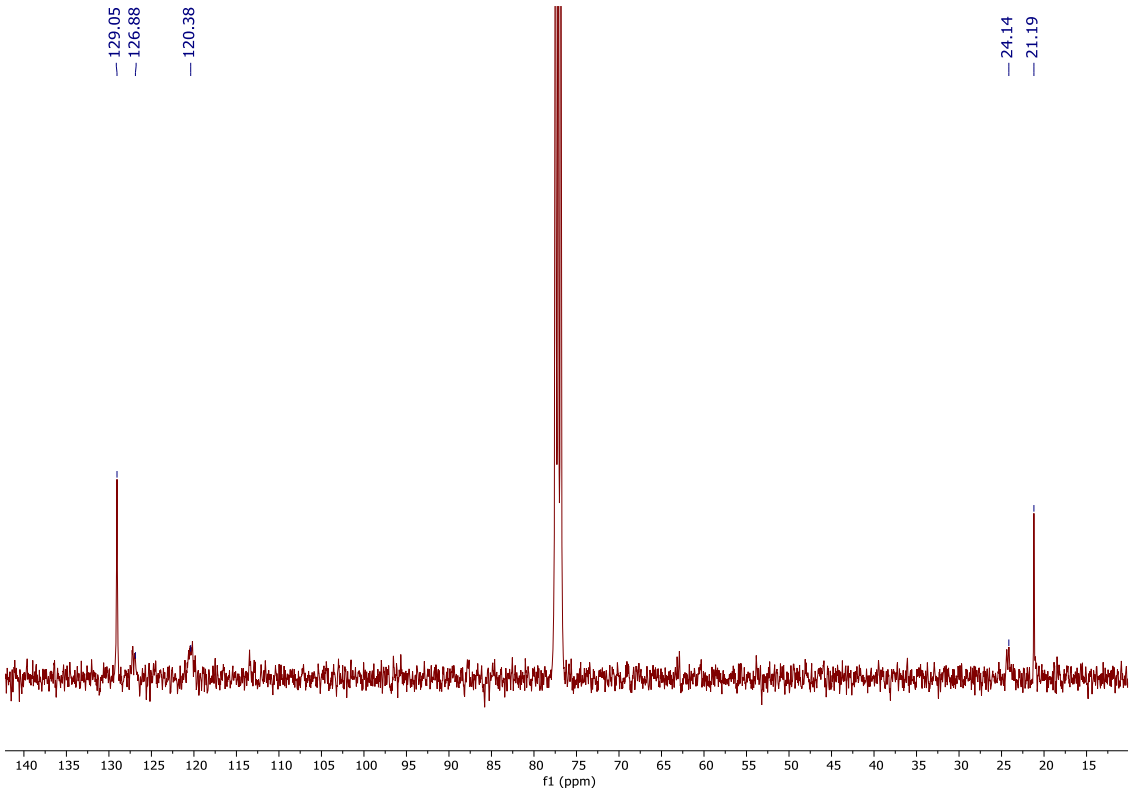
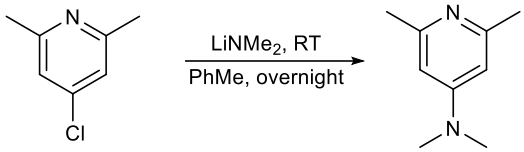


Figure S35. ^{13}C NMR spectrum of **17** in $CDCl_3$.

Synthesis of 2,6-dimethyl-4-dimethylaminopyridine



The synthesis of 4-chloro-2,6-lutidine was carried out with slight modifications from previously reported procedure:[7]

1) Post $POCl_3$ quench, the reaction mixture was extracted into $CHCl_3$ instead of CH_2Cl_2.
2) We consistently observe a mixture of oil and solid after the NEt_3 workup step, lowering the overall yield to 40 %. This was purified by filtering through a pipette to remove the solids before using.

$LiNMe_2$ (1.25 g, 24.4 mmol, 2.0 equiv.) was dissolved in PhMe (30 mL) in a 100 mL Schlenk flask in a N_2-filled glovebox. The Schlenk was sealed and removed from the glovebox, attached to a Schlenk line and cooled to -78 °C. 4-chloro-2,6-lutidine (1.73 g, 12.2 mmol, 1.0 equiv.) was syringed into the cooled solution. The resulting red solution was stirred for 30 minutes at -78 °C before warming to room temperature and stirring overnight. The unquenched solution was concentrated *in vacuo*, re-suspended in hexanes (75 mL) and filtered through a fine frit. The yellow filtrate was concentrated *in vacuo*, re-dissolved in CH_2Cl_2 (25 mL), passed through a basic alumina plug and concentrated *in vacuo* a final time to yield pure 2,6-dimethyl-4-dimethylaminopyridine as a thick brown oil (1.0 g, 55 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 6.22 (s, 2H, py-*H*), 2.97 (s, 6H, NMe_2-*H*), 2.42 (s, 6H, *o*-py-C*H*$_3$).

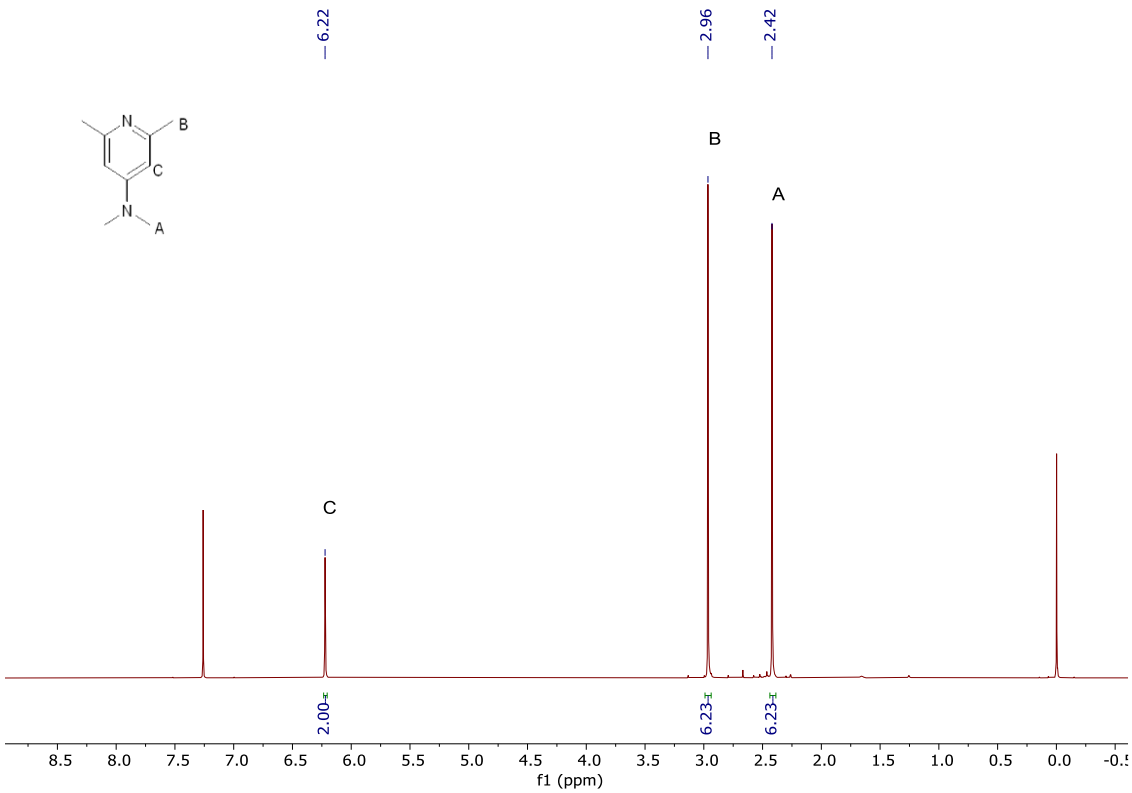
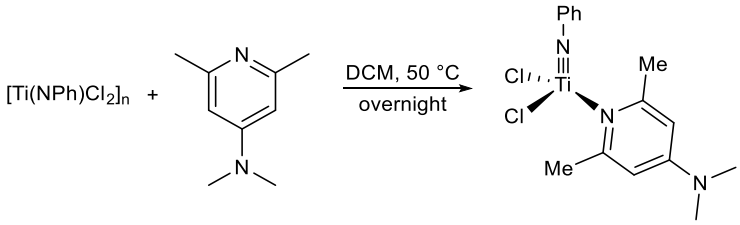


Figure S36. ^{1}H NMR spectrum of 2,6-dimethyl-4-dimethylaminopyridine in $CDCl_3$.

Synthesis of 18



$[Ti(NPh)Cl_2]_n$ (92 mg, 0.44 mmol, 1.0 equiv), 2,6-dimethyl-4-dimethylaminopyridine (66 mg, 0.44 mmol, 1.0 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried *in vacuo* to give **18** as a highly insoluble brown powder (137 mg, 87 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.10 – 6.72 (m, 4H, NPh-*H*), 6.00 (br s, 2H, py-*H*), 3.02 (s, 6H, *o*-py-C*H*$_3$), 2.48 (s, 6H, NMe_2-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (126 MHz, $CDCl_3$): δ 127.6, 127.9, 124.4, 119.9, 104.7, 39.5.
This is a partial line-list determined with ^{1}H-^{13}C HSQC & HMBC data. Three Cs could not be identified.

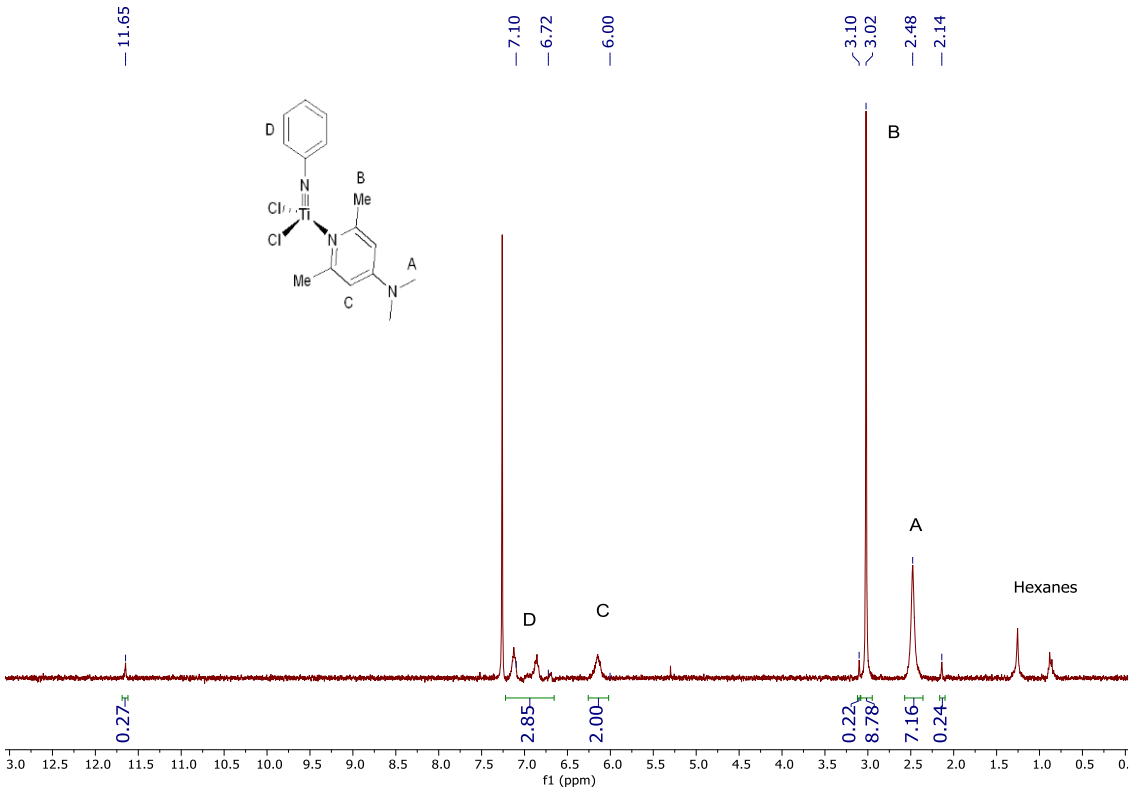


Figure S37. ^{1}H NMR spectrum of **18** in $CDCl_3$.

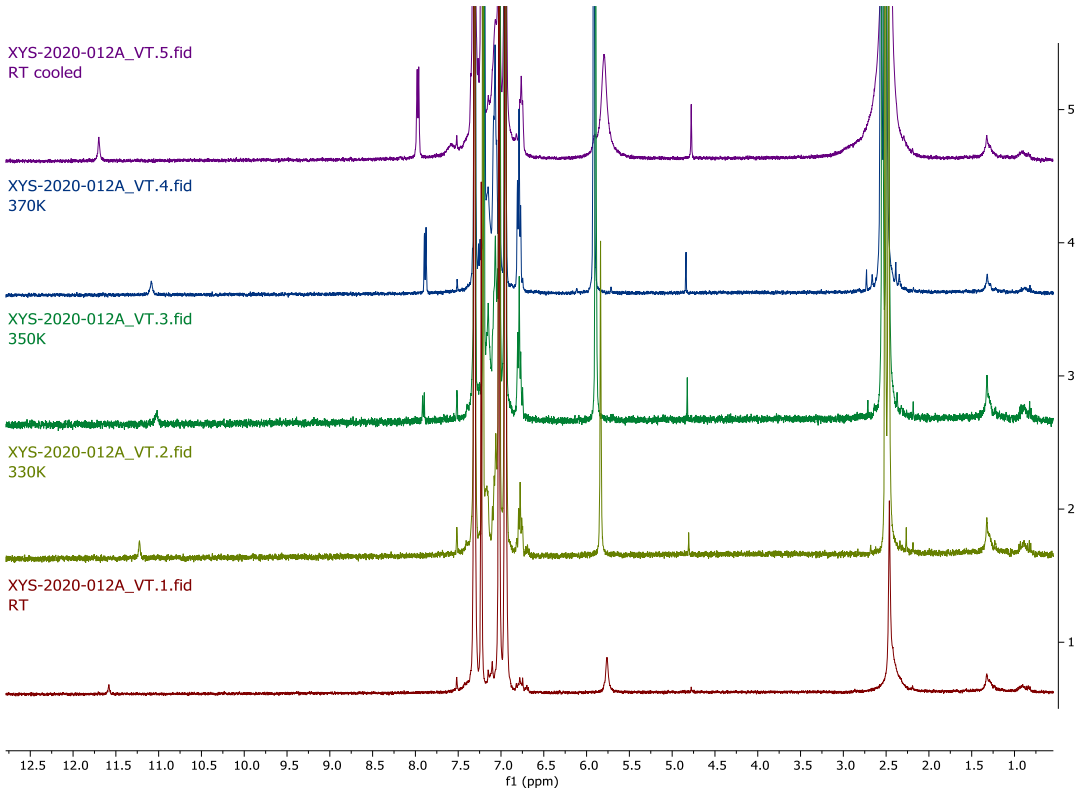


Figure S38. Stacked variable temperature ^{1}H NMR spectra of **18** in C_6D_5Br.

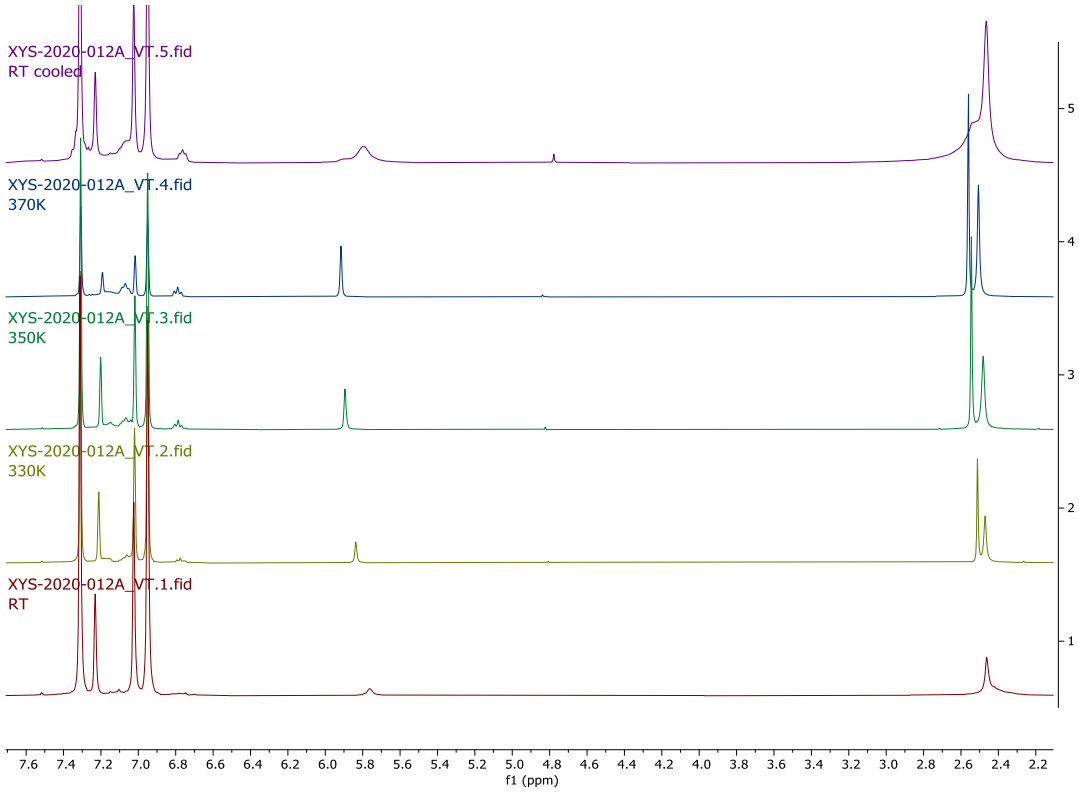


Figure S39. Zoom-in stacked variable temperature ^{1}H NMR spectra of **18** in C_6D_5Br.

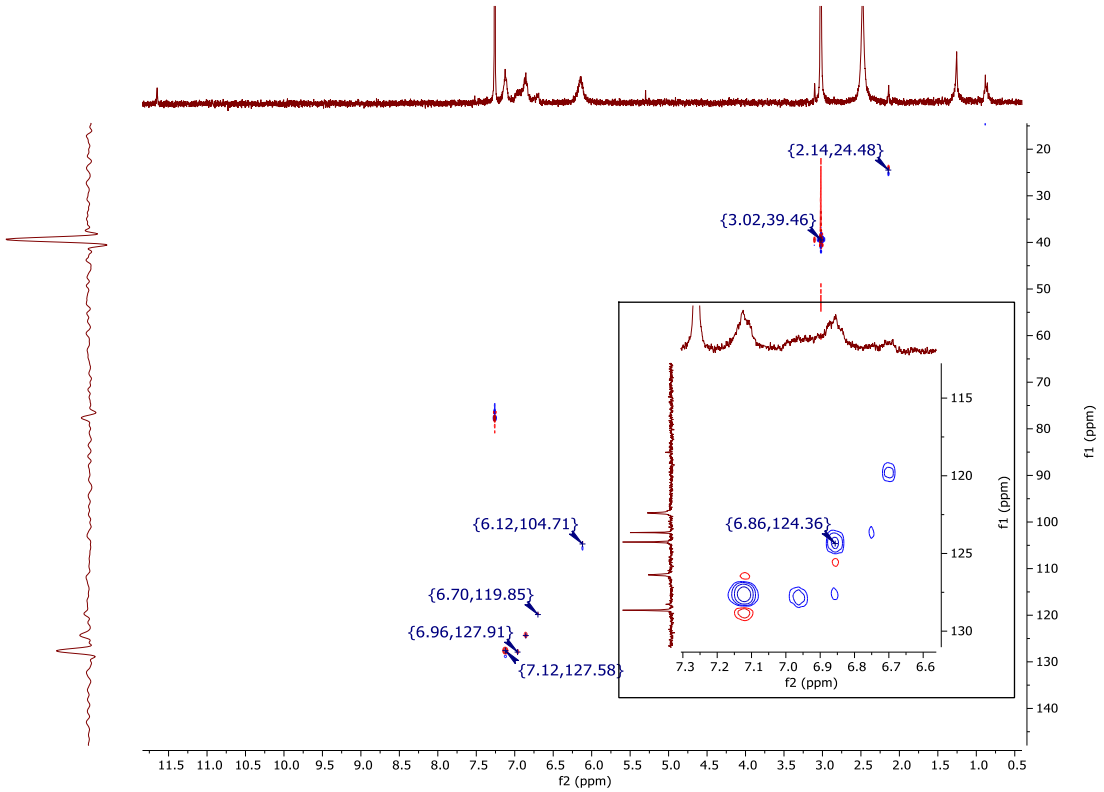


Figure S40. ^{1}H-^{13}C HSQC NMR spectrum of **18** in $CDCl_3$.

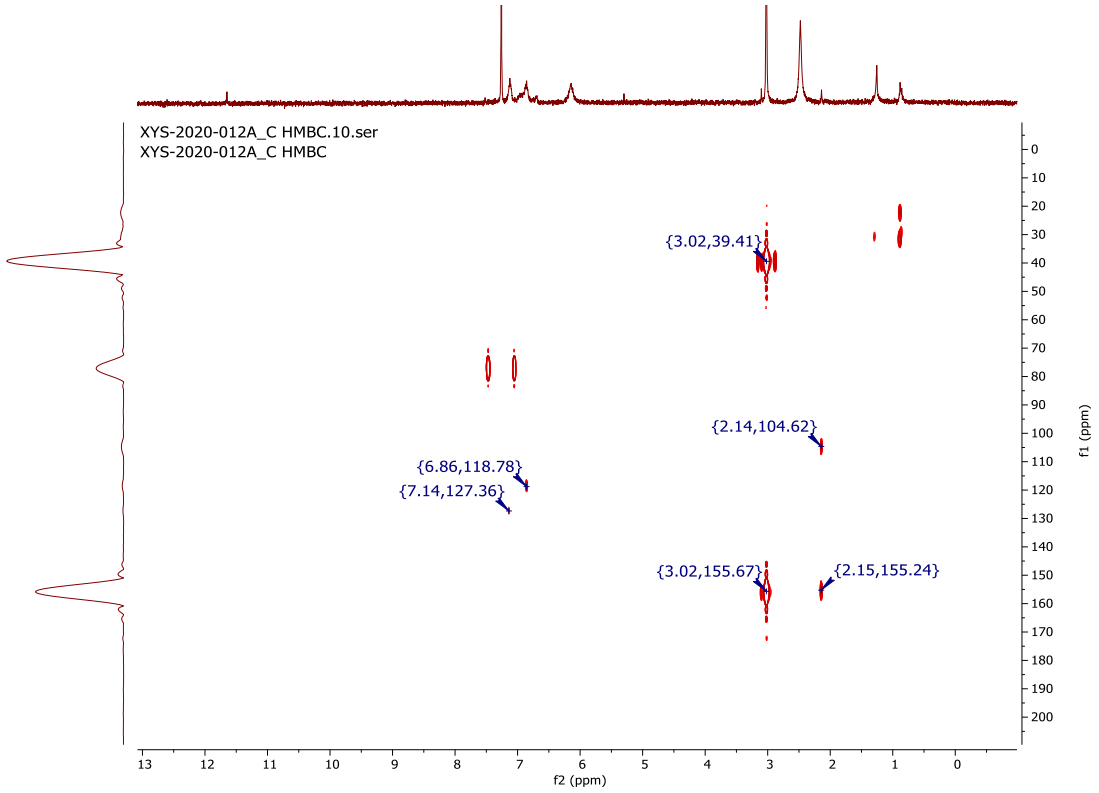
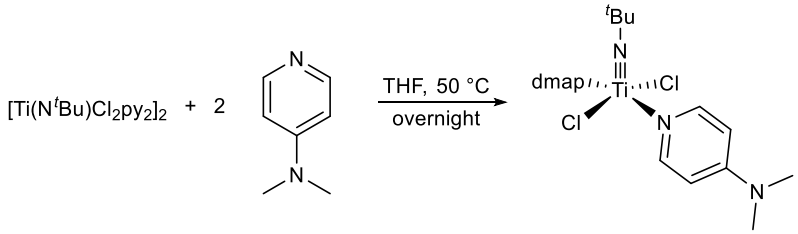


Figure S41. ^{1}H-^{13}C HMBC NMR spectrum of **18** in $CDCl_3$.

Synthesis of 19



[py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (137 mg, 0.39 mmol, 1.0 equiv), 4-dimethylaminopyridine (96 mg, 0.79 mmol, 2.0 equiv) and 2 mL THF were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, yellow suspension was filtered through a fine frit and washed with THF. The residue was dried under *vacuo* to give **19** as a yellow powder (120 mg, 71 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.83 (d, $^{3}J_{HH}$ = 7.3 Hz, 4H, *o*-dmap-H), 6.53 (d, $^{3}J_{HH}$ = 7.3 Hz, 4H, *m*-dmap-*H*), 3.08 (s, 12H, NMe$_2$-*H*), 1.03 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, $CDCl_3$): δ 154.9, 151.1, 105.7, 73.3, 39.3, 31.1.

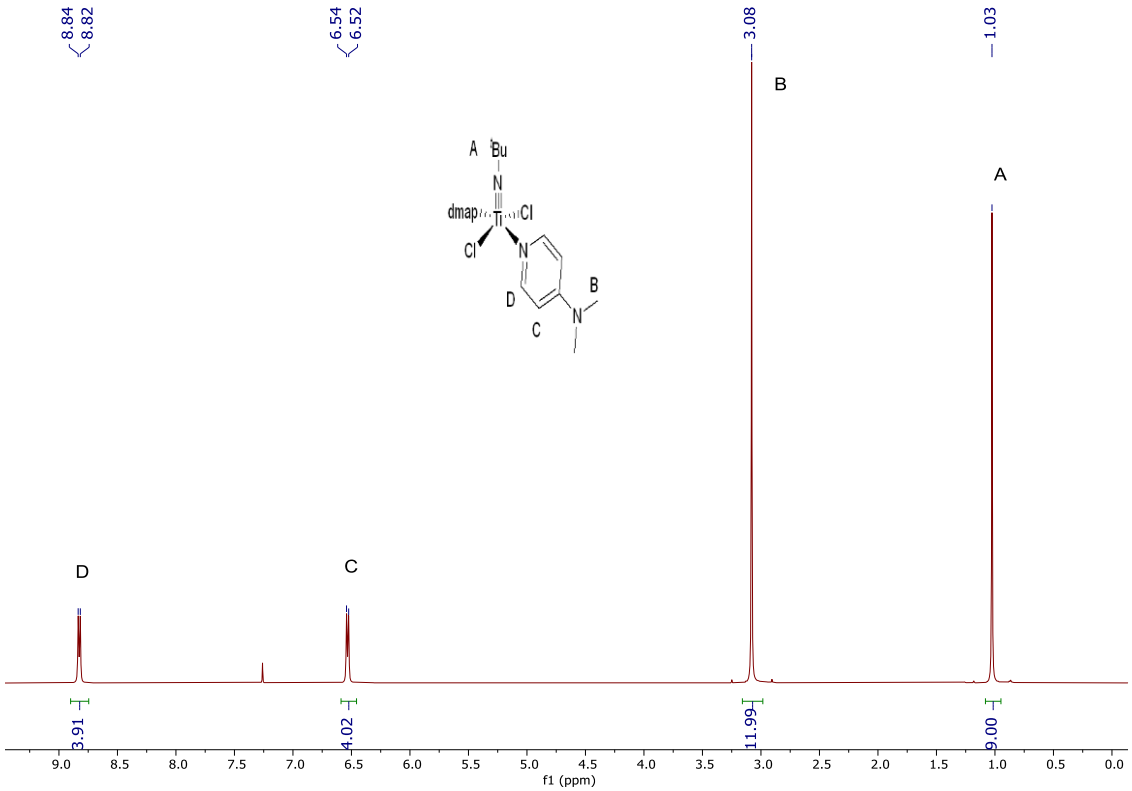


Figure S42. ^{1}H NMR spectrum of **19** in $CDCl_3$.

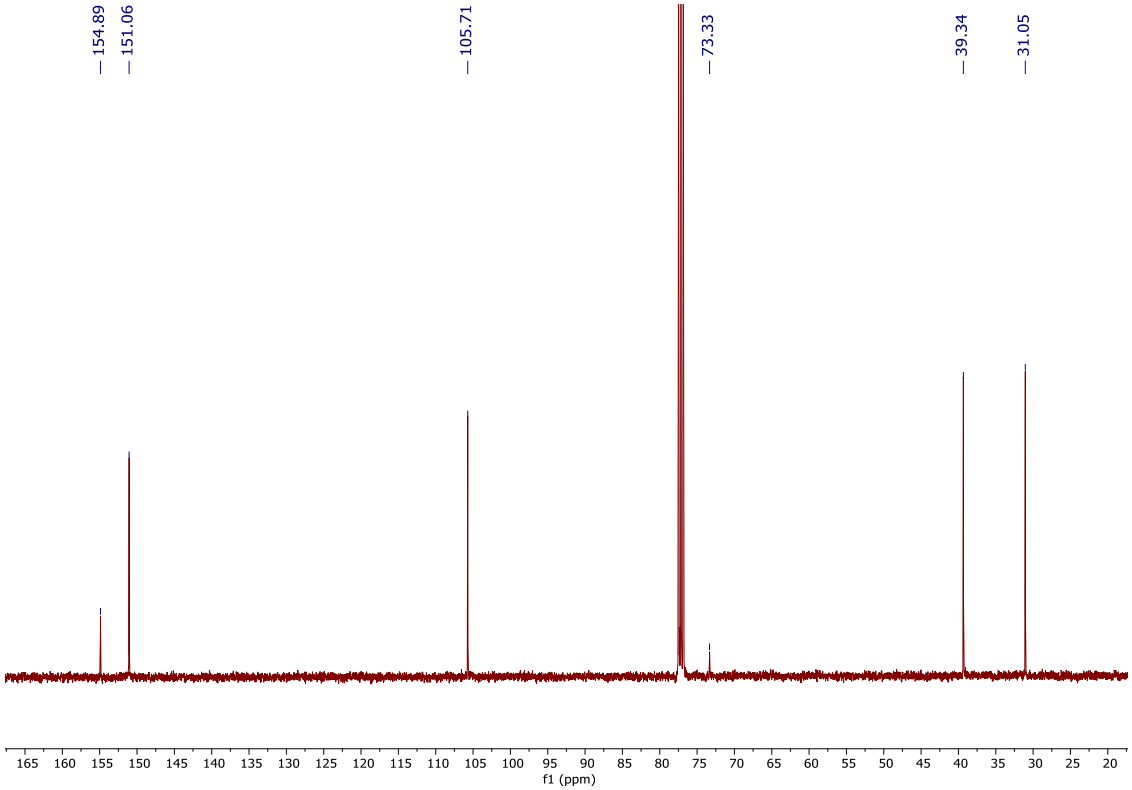
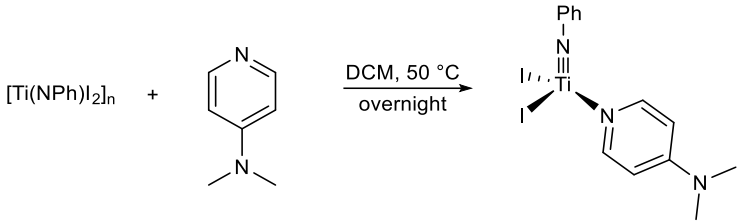


Figure S43. ^{13}C NMR spectrum of **19** in $CDCl_3$.

Synthesis of 20



$[Ti(NPh)I_2]_n$ (108 mg, 0.28 mmol, 1.0 equiv), 4-dimethylaminopyridine (37 mg, 0.30 mmol, 1.1 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was dried *in vacuo* to give a black sticky liquid. The crude product was taken up in 5 mL C_6H_6 to form a suspension. Suspension was filtered and washed with more C_6H_6 to give the filtrate as desired product with residual C_6H_6 present. 2 mL of hexanes was added to the product, stirred at room temperature, dried *in vacuo*. This process was repeated three times to give **20** as a brown powder (98 mg, 69 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.74 (s, 2H, *o*-dmap-*H*), 7.12 (br s, 4H, NPh-*H*), 6.88 (s, 1H, NPh-*H*), 6.52 (s, *m*-dmap-*H*), 3.10 (s, 6H, NMe_2-*H*).

^{13}C NMR (126 MHz, $CDCl_3$): δ 154.9, 150.4, 128.4, 124.1, 122.9, 105.6, 39.5.
This is a partial line-list. Resonances associated with a quaternary C were not observed.

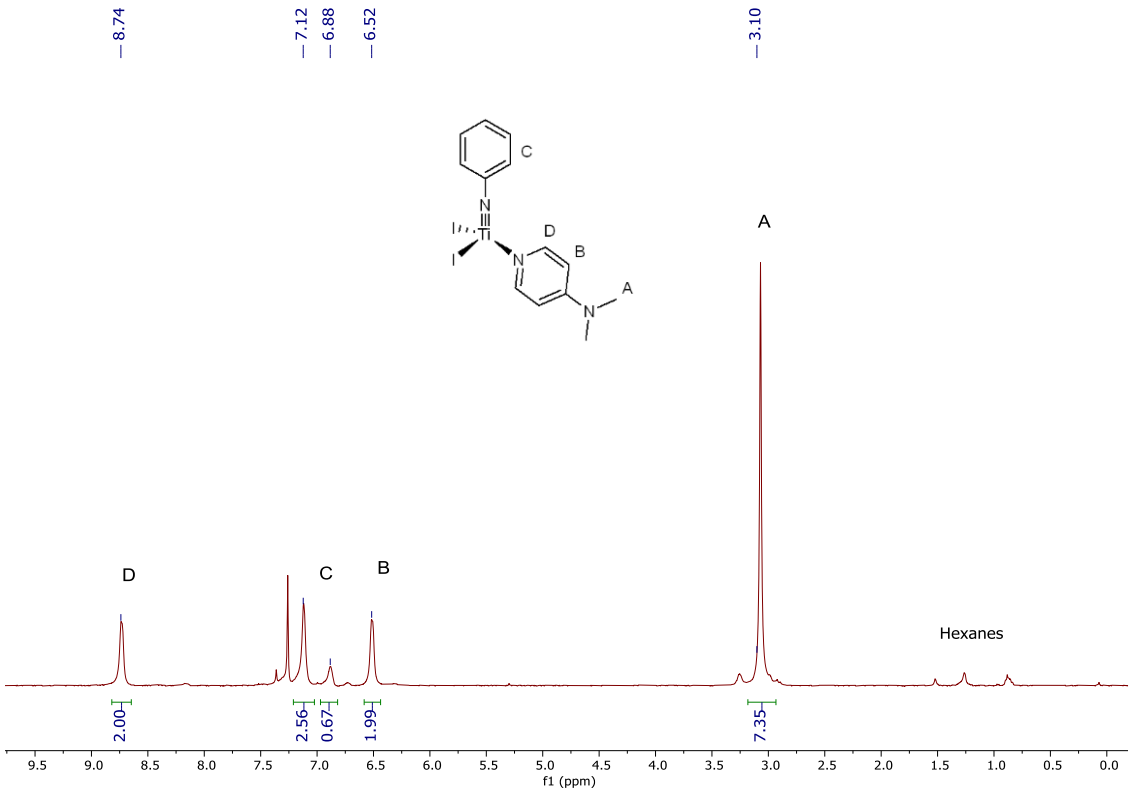


Figure S44. ^{1}H NMR spectrum of **20** in $CDCl_3$.

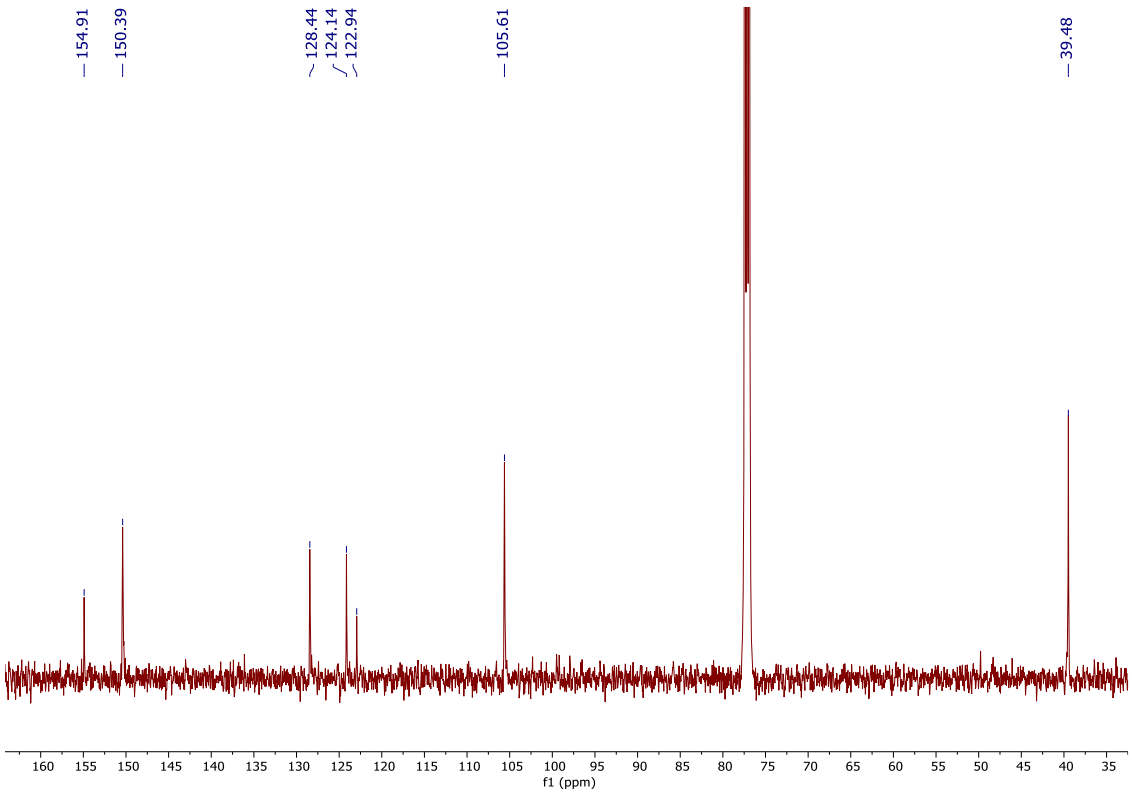
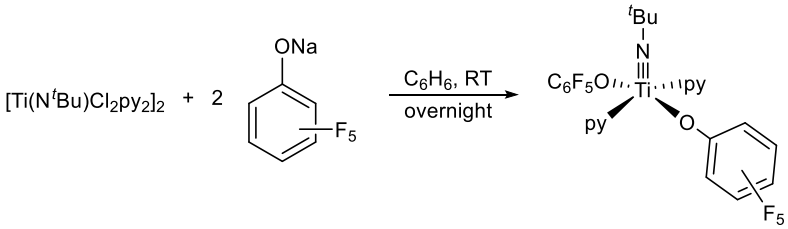


Figure S45. ^{13}C NMR spectrum of **20** in $CDCl_3$.

Synthesis of 21



[py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (100 mg, 0.29 mmol, 1.0 equiv), sodium pentafluorophenoxide (118 mg, 0.57 mmol, 2.0 equiv) and 2 mL C_6H_6 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred overnight. The suspension was then filtered through a celite plug on a coarse frit and the filtrate was lyophilised *in vacuo* to yield pure **21** as an orange powder (110 mg, 86 % yield).

^{1}H NMR (500 MHz, C_6D_6): δ 8.97 (s, 4H, *o*-py-H), 6.84 (t, $^3J_{HH}$ = 7.7 Hz, 2H, *p*-py-*H*), 6.59 (br t, 4H, *m*-py-*H*), 1.06 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (126 MHz, C_6D_6): δ 150.3, 141.9 (t), 140.0 – 139.7 (m), 138.7, 138.0 – 137.7 (m), 128.6, 124.2, 71.7, 31.2.

^{19}F NMR (471 MHz, C_6D_6): δ -163.3 (dd, $^3J_{FF}$ = 19.9 and $^4J_{FF}$ = 8.7 Hz, *o*-OPh-*F*), -167.6 (app t, $^3J_{FF}$ = 21.2 Hz, *m*-OPh-*F*), -176.6 (tt, $^3J_{FF}$ = 22.6 and $^4J_{FF}$ = 8.2 Hz, *p*-OPh-*F*).

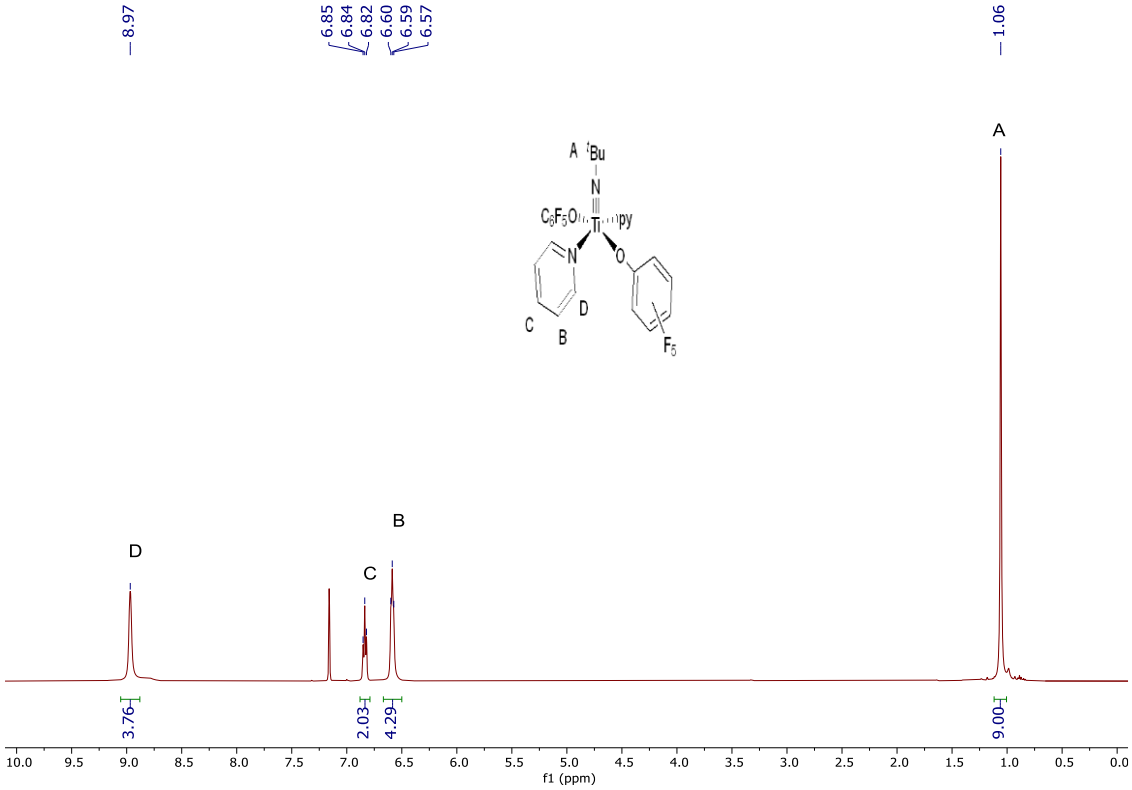


Figure S46. ^{1}H NMR spectrum of **21** in C_6D_6.

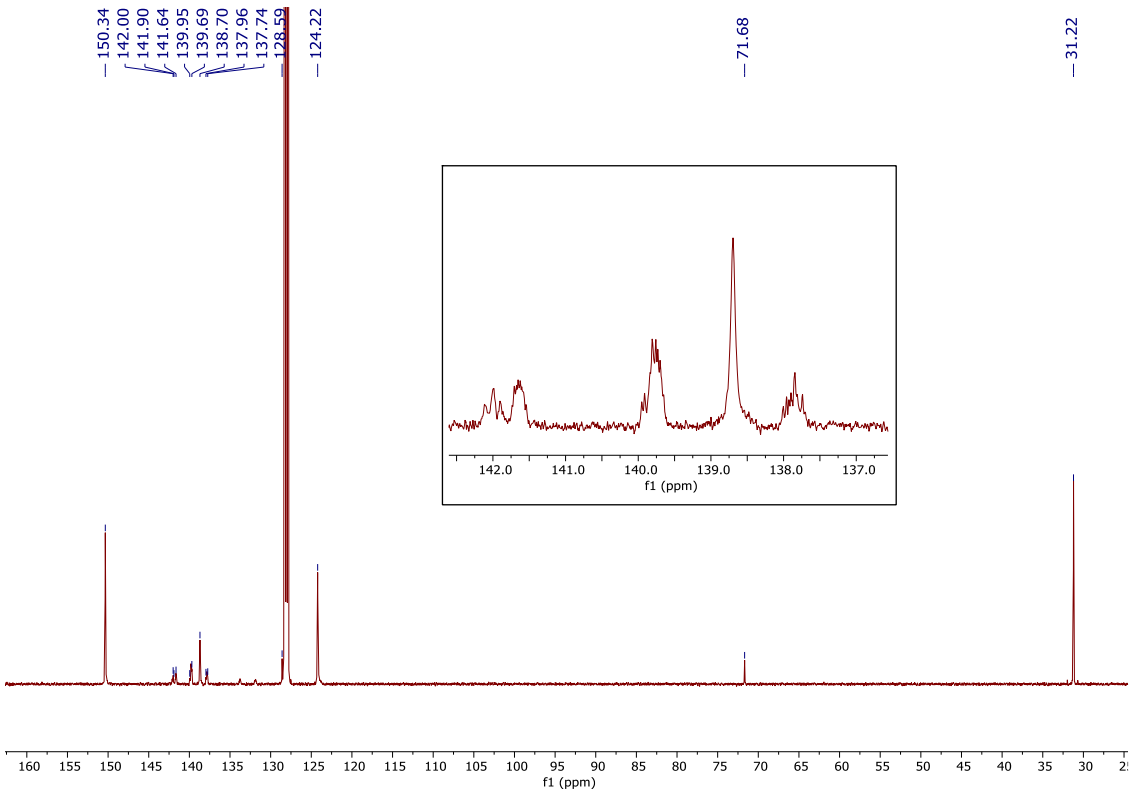


Figure S47. ^{13}C NMR spectrum of **21** in C_6D_6.

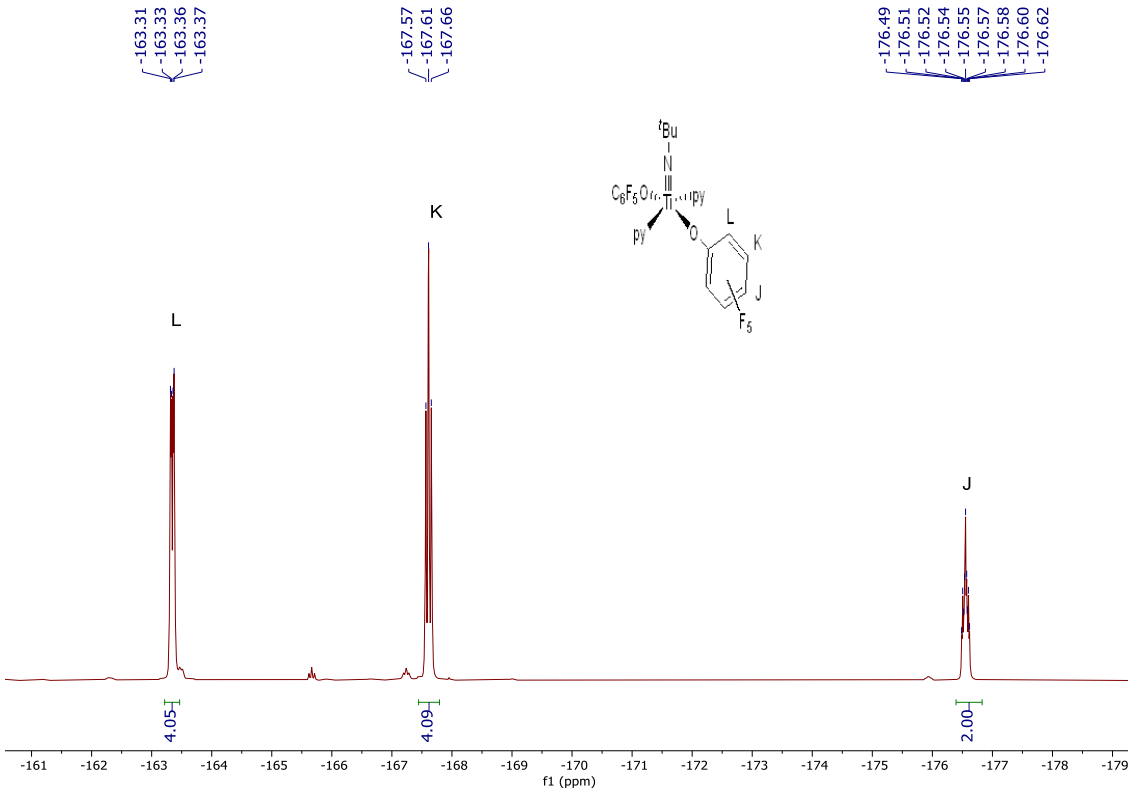
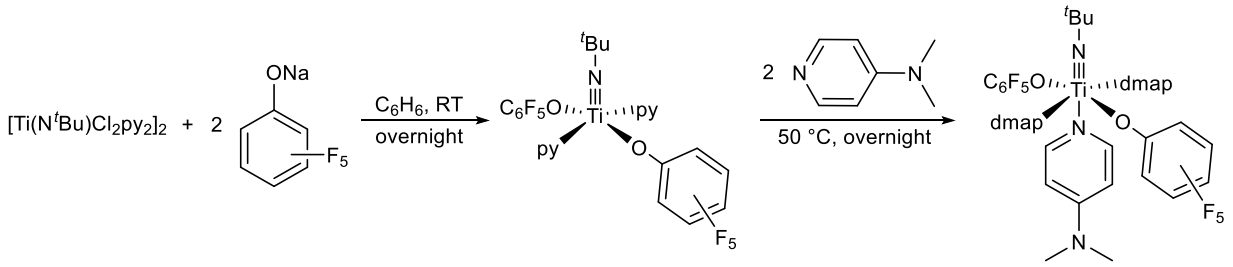


Figure S48. ^{19}F spectrum of **21** in C_6D_6.

Synthesis of 22



[py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (100 mg, 0.29 mmol, 1.0 equiv), sodium pentafluorophenoxide (118 mg, 0.57 mmol, 2.0 equiv) and 2 mL C_6H_6 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred overnight. The suspension was filtered through a celite plug on a coarse frit and washed with 1 mL C_6H_6. 4-dimethylaminopyridine (70 mg, 0.57 mmol, 2.0 equiv) was added to the yellow solution, sealed with a Teflon-lined cap and stirred overnight at 50 °C. The solution was cooled in the freezer, then lypophilised *in vacuo* to form a crude solid with residual free pyridine. The solid was heated *in vacuo* for several hours at 50 °C to yield pure **22** as a pale yellow solid (210 mg, 86 % yield).

^{1}H NMR (500 MHz, C_6D_6): δ 8.83 (d, $^3J_{HH}$ = 7.0 Hz, 4H, *o*-dmap-*H*), 8.68 (d, $^3J_{HH}$ = 6.0 Hz, 2H, *trans o*-dmap-*H*), 6.02 (d, $^3J_{HH}$ = 6.1 Hz, 2H, *trans m*-dmap-*H*), 5.93 (d, $^3J_{HH}$ = 7.2 Hz, 4H, *m*-dmap-*H*), 2.14 (s, 6H, *trans* NMe$_2$-*H*), 2.01 (s, 12H, NMe$_2$-*H*), 1.25 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (126 MHz, C_6D_6): δ 154.6, 150.7, 106.5, 105.7, 69.8, 38.1, 31.7.
This is a partial line-list. Due to the insolubility of the material, seven Cs could not be identified.

^{19}F NMR (471 MHz, C_6D_6): δ -162.7 (dd, $^3J_{FF}$ = 19.9 and $^4J_{FF}$ = 10.0 Hz, *o*-OPh-*F*), -168.7 (app t, $^3J_{FF}$ = 21.1 Hz, *m*-OPh-*F*), -179.3 (tt, $^3J_{FF}$ = 23.0 and $^4J_{FF}$ = 10.0 Hz, *p*-OPh-*F*).

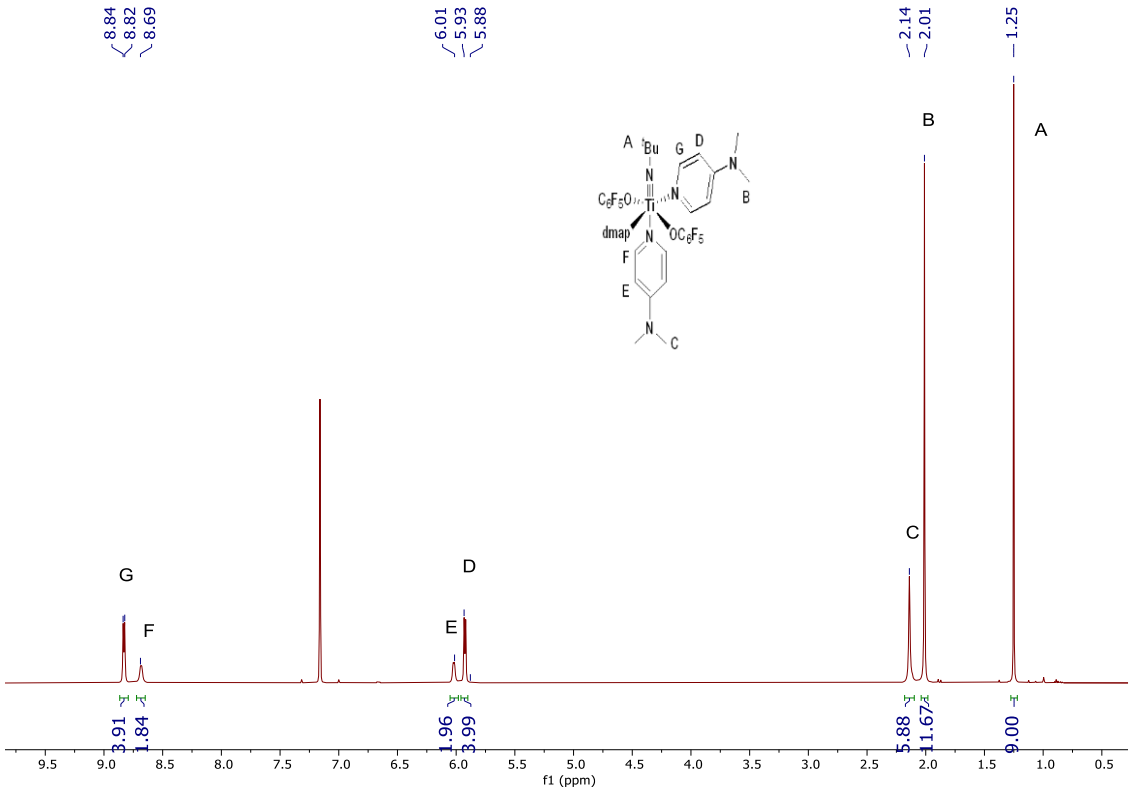


Figure S49. ^{1}H NMR spectrum of **22** in C_6D_6.

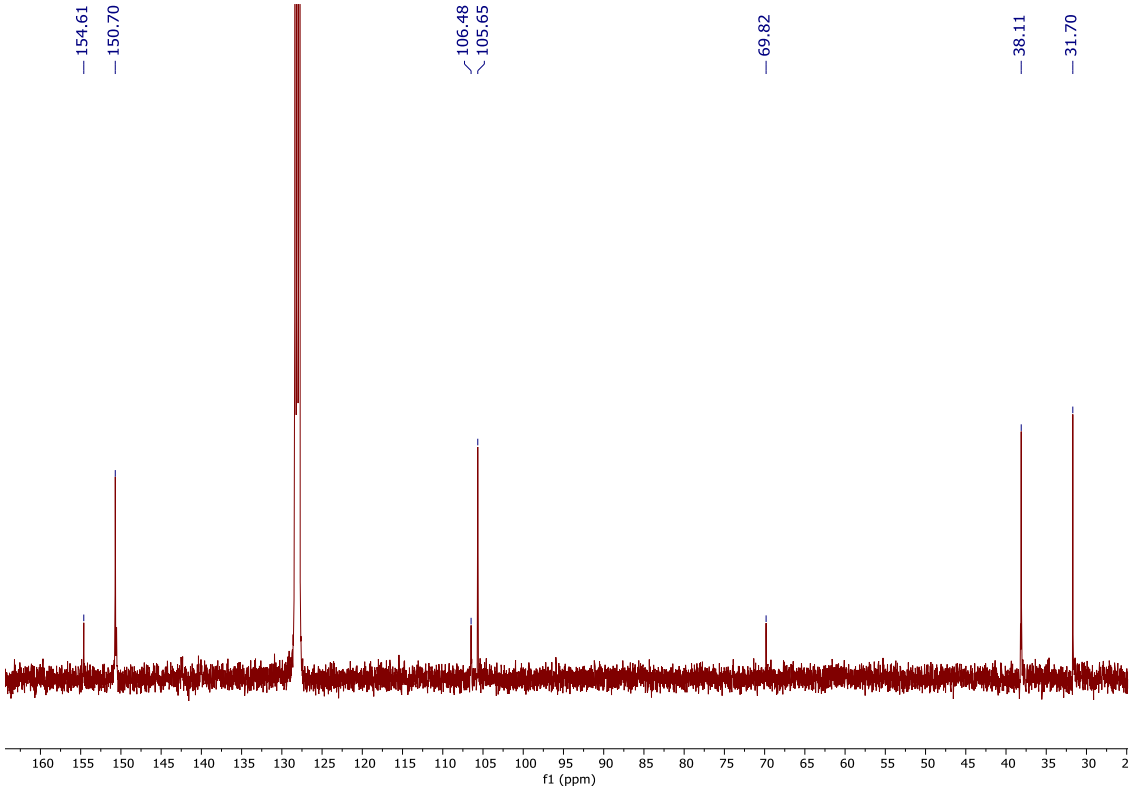


Figure S50. ^{13}C NMR spectrum of **22** in C_6D_6.

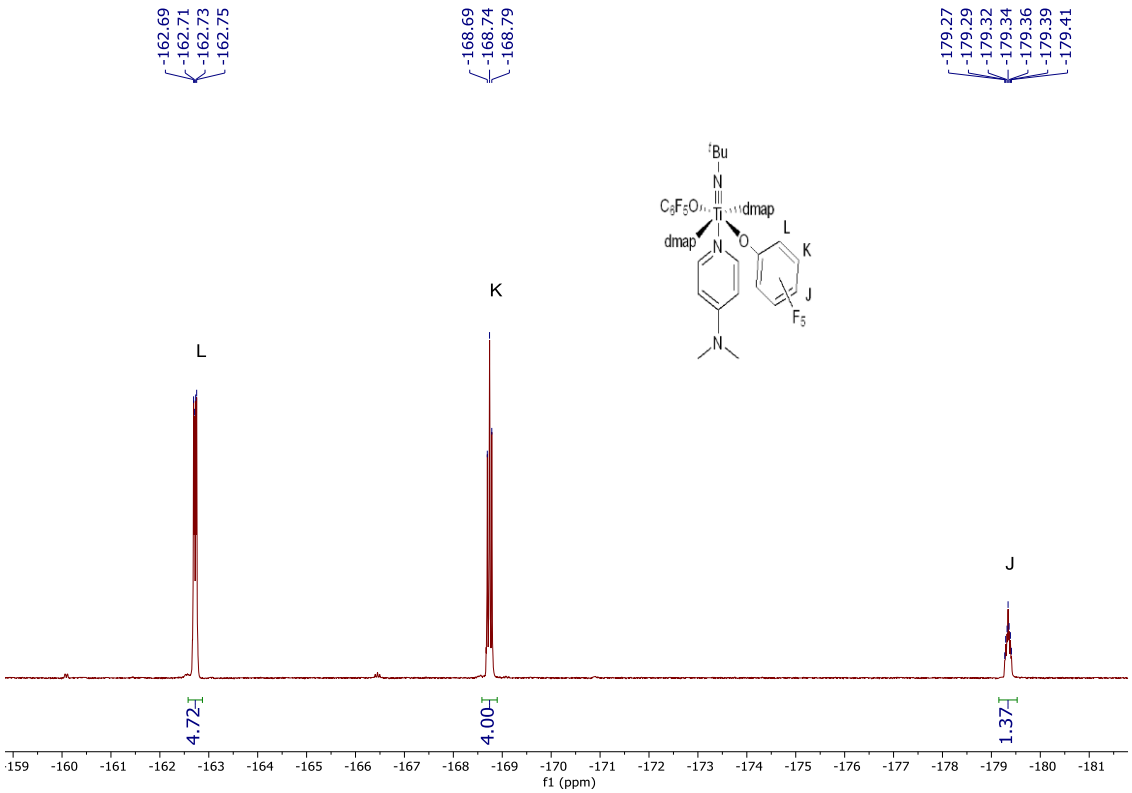
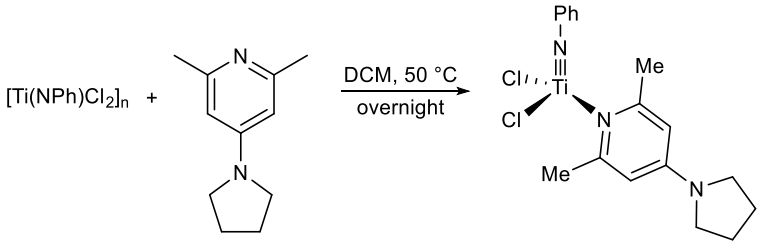


Figure S51. ^{19}F spectrum of **22** in C_6D_6.

Synthesis of 23



2,6-dimethyl-4-dimethylaminopyridine was synthesized according to literature procedure.[7]

$[Ti(NPh)Cl_2]_n$ (89 mg, 0.42 mmol, 1.0 equiv), 2,6-dimethyl-4-dimethylaminopyridine (78 mg, 0.45 mmol, 1.1 equiv) and 2 mL CH_2Cl_2 were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap, heated to 50 °C and stirred overnight. After cooling to room temperature, the reaction mixture was filtered through a fine frit and washed with CH_2Cl_2. The residue was dried under *vacuo* to give **23** as a highly insoluble brown powder (150 mg, 92 % yield).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.06 (br s, 1H, NPh-*H*), 6.89 (br s, 2H, NPh-*H*), 6.78 (br s, 1H, NPh-*H*), 6.67 (br s, 1H, NPh-*H*), 6.09 (br s, 2H, py-*H*), 3.26 (br s, 4H, *o*-prl-*H*), 2.41 (br s, 6H, *o*-py-C*H*$_3$), 2.04 (br s, 4H, *m*-prl-*H*).
Hexane impurities are present in the $CDCl_3$ solvent.

^{13}C NMR (151 MHz and 126 MHz, $CDCl_3$): δ 153.5, 148.4, 127.8, 127.5, 124.0, 123.7, 120.0, 104.2, 47.4, 25.4, 22.5.

This is a partial line-list pieced together from ^{13}C data (151 MHz, NS = 15360, AQ = 1 s, d1 = 5 s), ^{1}H-^{13}C HSQC & HMBC data. Resonances associated with a quaternary C of the NPh fragment were not observed.

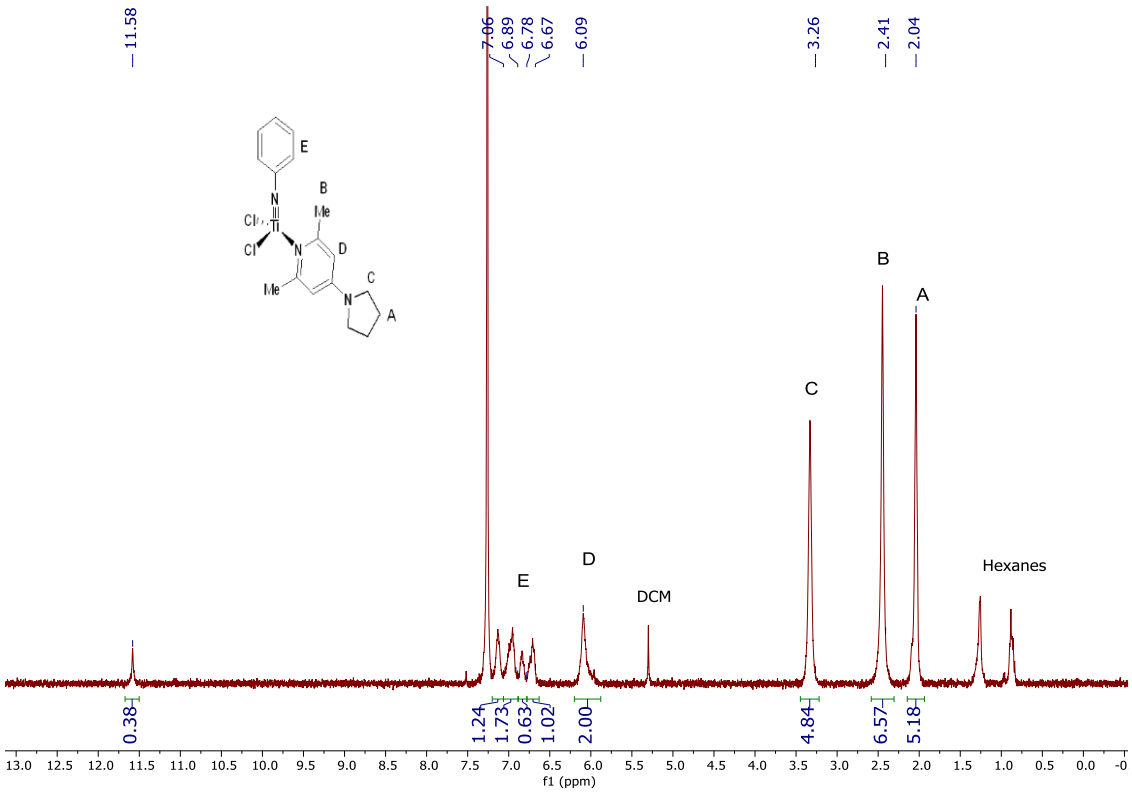


Figure S52. ^{1}H NMR spectrum of **23** in $CDCl_3$.

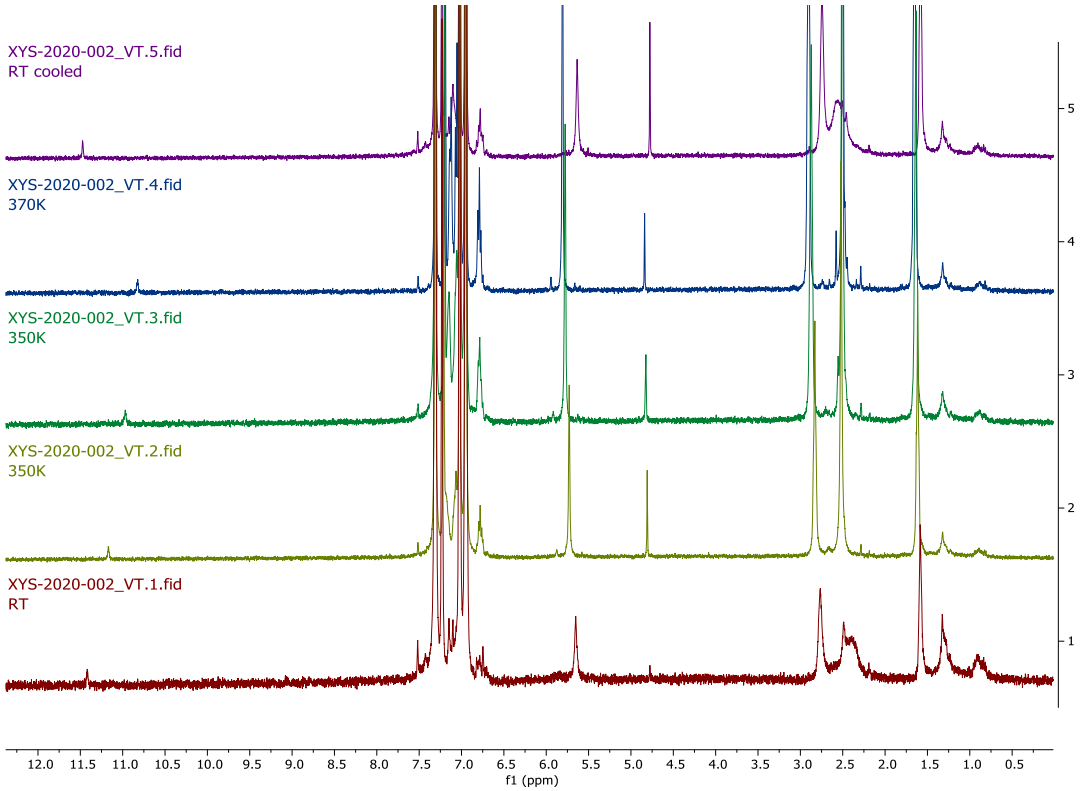


Figure S53. Stacked variable temperature ^{1}H NMR spectra of **23** in C_6D_5Br.

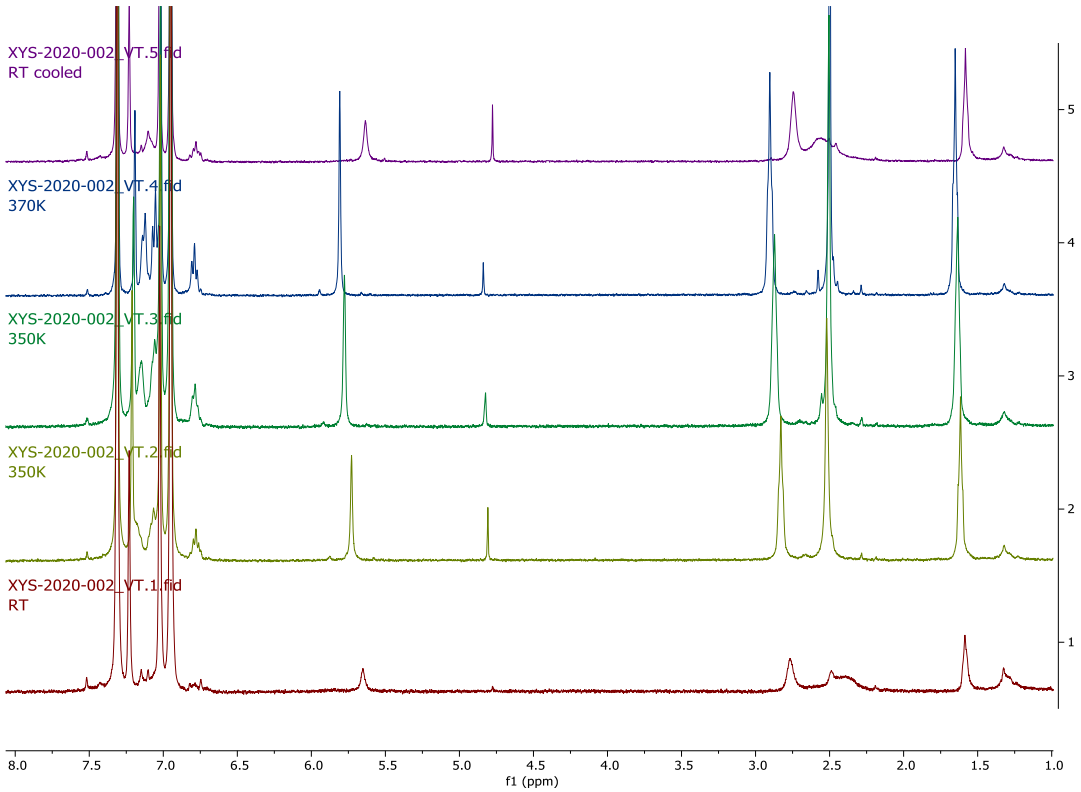


Figure S54. Zoom-in stacked variable temperature ^{1}H NMR spectrum of **23** in C_6D_5Br.

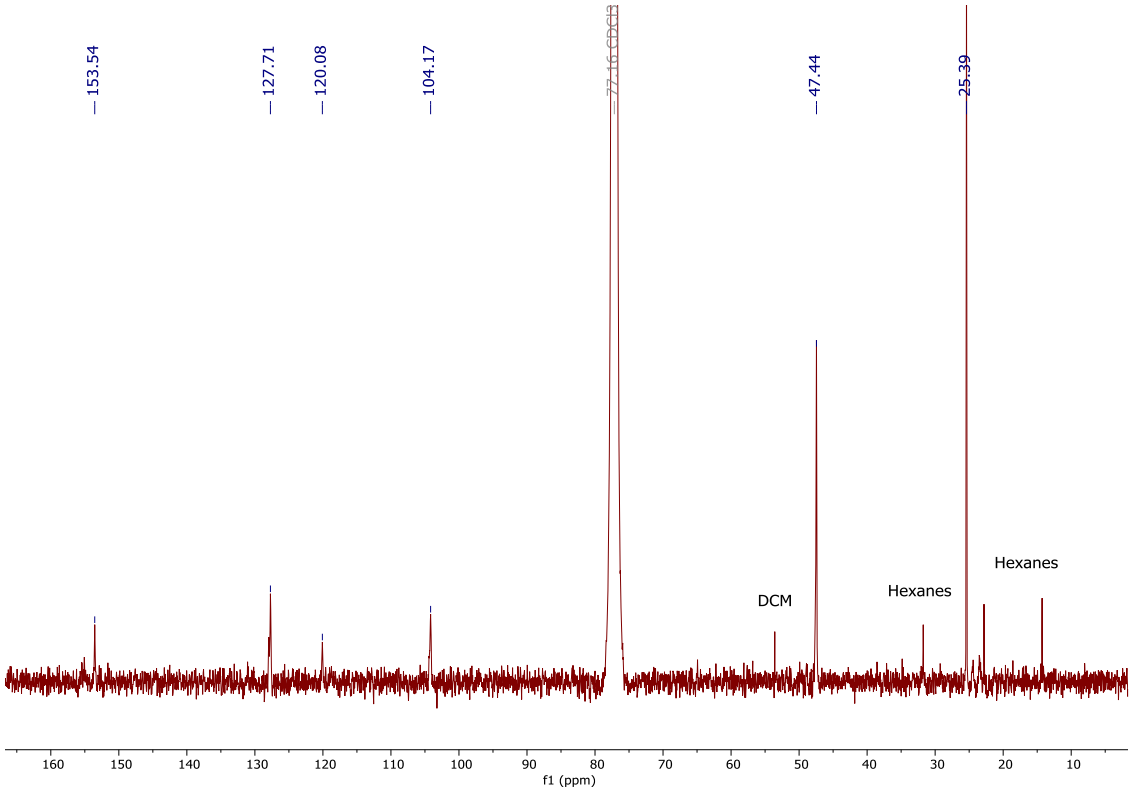


Figure S55. ^{13}C NMR spectrum of **23** in $CDCl_3$.

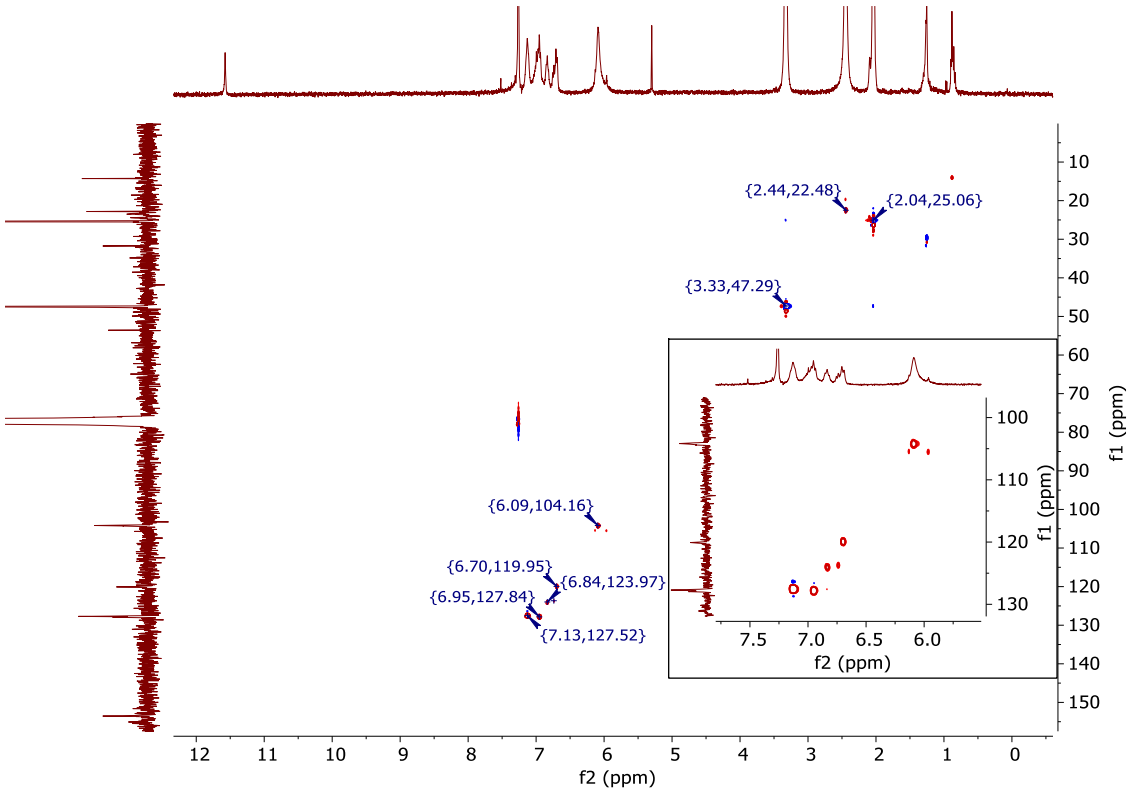


Figure S56. ^{1}H-^{13}C HSQC NMR spectrum of **23** in $CDCl_3$.

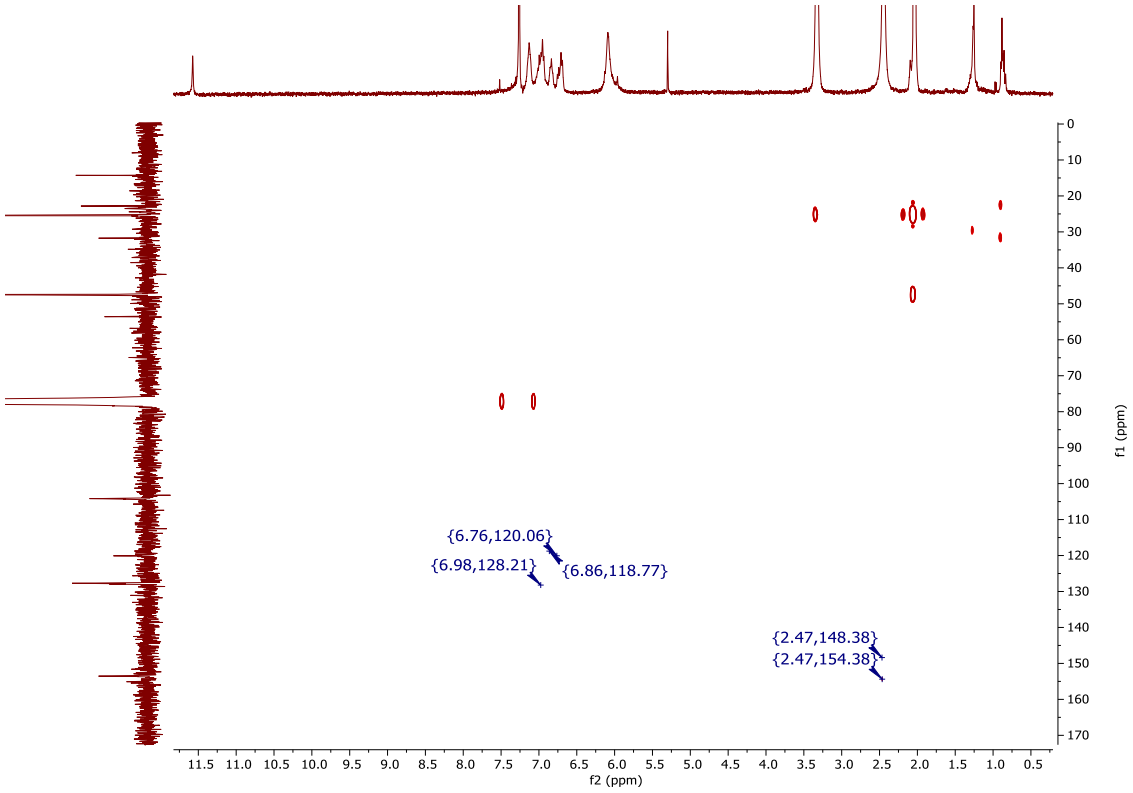
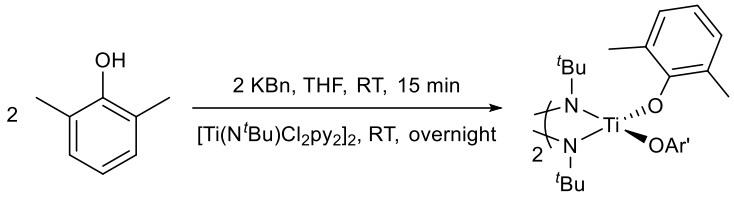


Figure S57. ^{1}H-^{13}C HMBC NMR spectrum of **23** in $CDCl_3$.

Synthesis of a



2,6-dimethylphenol (100 mg, 0.82 mmol, 2.0 equiv), KBn (107 mg, 0.82 mmol, 2.0 equiv) and 2 mL THF were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and stirred at room temperature for 15 mins to form a deep red solution. In a separate scintillation vial, [py$_2$Ti(N^{t}Bu)Cl$_2$]$_2$ (142 mg, 0.41 mmol, 1.0 equiv) was dissolved in 2 mL THF to form an orange solution. The [Ti] solution was added to the ligand solution and stirred overnight at room temperature. Following which, the reaction mixture was dried *in vacuo*, dissolved in benzene (5 mL), filtered through a celite plug and the filtrate dried *in vacuo*. The crude product was purified by dissolving in ether, layering with the same volume of hexanes and cooling in a -35 °C freezer to yield X-ray quality brown crystals of pure **a**. (14 mg, 9 % yield).

^{1}H NMR (400 MHz, C_6D_6): δ 7.02 (d, $^{3}J_{HH}$ = 7.6 Hz, 4H, *m*-Ph-*H*), 6.84 (t, $^{3}J_{HH}$ = 7.5 Hz, 2H, *p*-Ph-*H*), 2.49 (s. 12H, C*H*$_3$), 1.14 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (101 MHz, C_6D_6): δ 163.3, 128.9, 127.0, 121.5, 73.2, 33.1, 18.2.

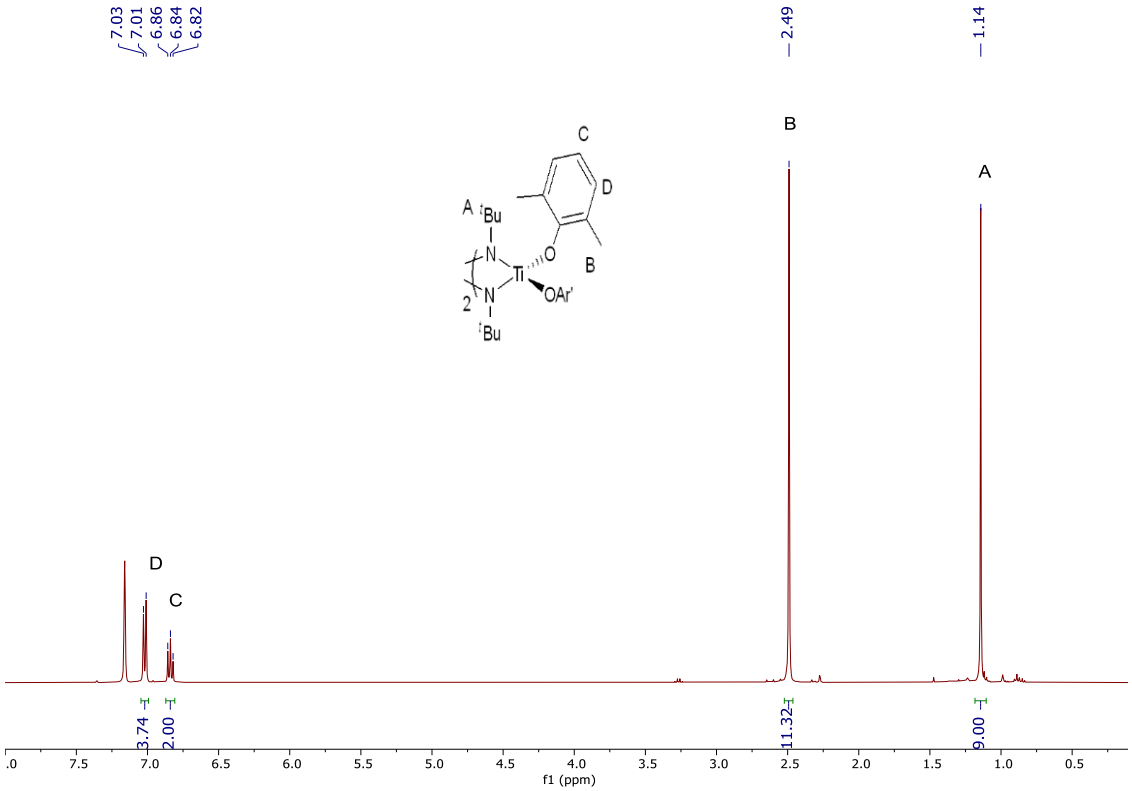


Figure S58. ^{1}H NMR spectrum of **a** in C_6D_6.

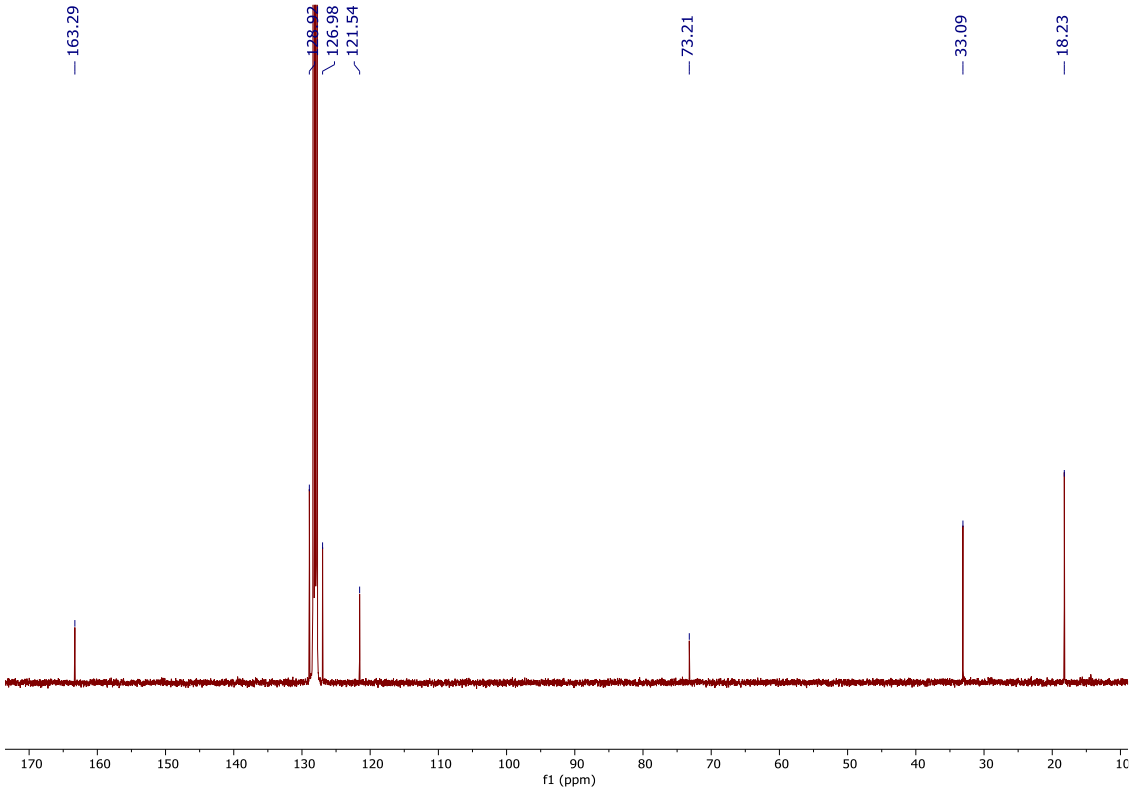
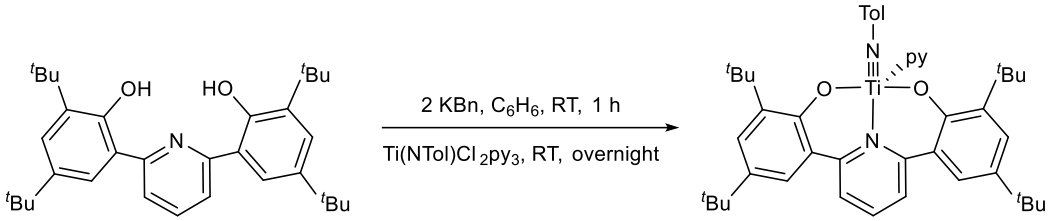


Figure S59. ^{13}C NMR spectrum of **a** in C_6D_6.

Synthesis of b



Complex **b** was synthesized by modified literature procedure[8] using $(ONO)H_2$ (200 mg, 0.41 mmol, 1.0 equiv), KBn (107 mg, 0.82 mmol, 2.0 equiv) and $py_3Ti(NTol)Cl_2$ (189 mg, 0.41 mmol, 1.0 equiv). Upon benzene lyophilization to form a sticky liquid, 5 mL of hexanes were added to the crude product to crash out an orange powder. The orange powder was dried *in vacuo* for several hours to yield pure **b** (100 mg, 34 %).

^{1}H NMR (500 MHz, C_6D_6): δ 8.74 (d, $^3J_{HH}$ = 4.7 Hz, 2H, *o*-py-*H*), 7.83 (app s, 2 H, *m*-OPh-*H*), 7.57 (app s, 2H, *m*-OPh-*H*), 7.40 (d, $^3J_{HH}$ = 7.9 Hz, 2H, *m*-NTol-*H*), 7.12 (t, $^3J_{HH}$ = 8.1 Hz, 1H, Ar-*H*), 6.93 (t, $^3J_{HH}$ = 7.7 Hz, 1H, Ar-*H*), 6.61 – 6.56 (m, 4H, Ar-*H*), 6.31 (d, $^3J_{HH}$ = 7.9 Hz, 2H, *o*-NTol-*H*), 1.92 (s, 3H, NC_6H_4-CH_3), 1.54 (s, 18H, o-^{t}Bu-*H*), 1.47 (s, 18H, *p*-^{t}Bu-*H*).

^{13}C NMR (101 MHz, C_6D_6): δ 159.8, 159.4, 158.9, 150.5, 140.2, 138.8, 138.5, 137.3, 129.2, 128.5, 127.1, 126.2, 125.6, 124.4, 123.2, 35.4, 34.6, 32.1, 30.1, 21.0.
This is a partial line-list. Resonances associated with a quaternary C were not observed.

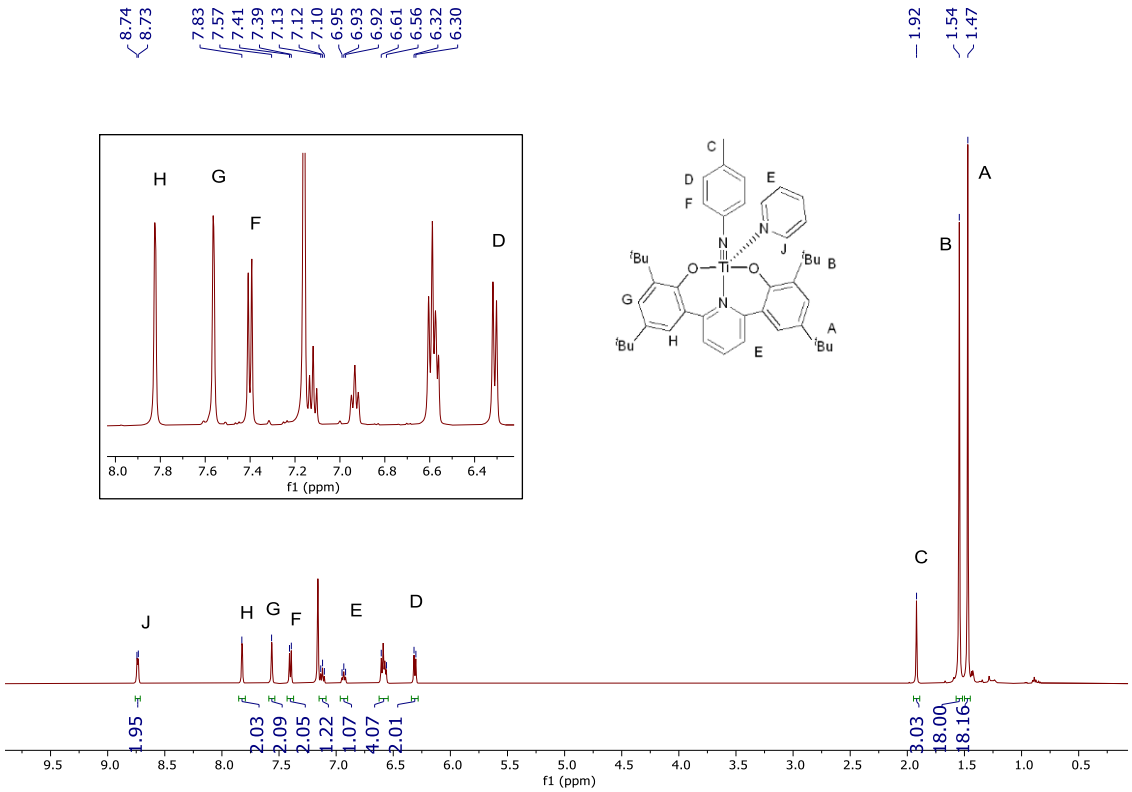


Figure S60. ^{1}H NMR spectrum of **b** in C_6D_6.

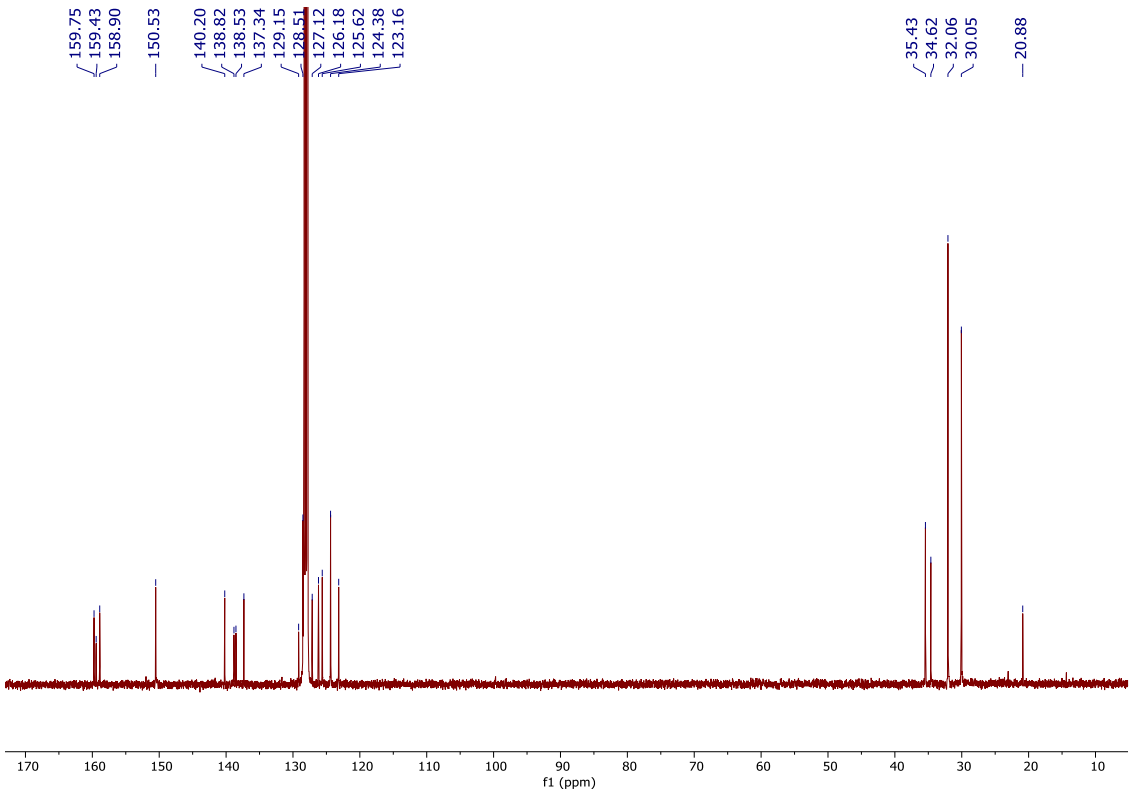
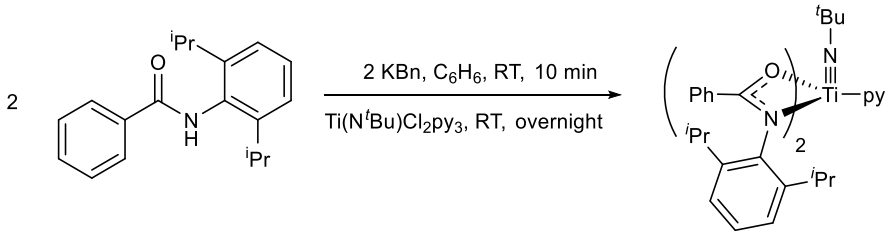


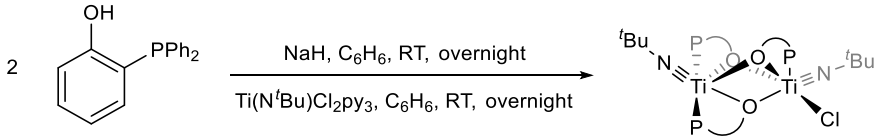
Figure S61. ^{13}C NMR spectrum of **b** in C_6D_6.

Synthesis of c



pyTi(N^{t}Bu)(amidate)$_2$ was synthesized with slight modifications to literature procedures.[3] NMR of complex matches that of reported in literature. [9]

Synthesis of d



(2-hydroxyphenyl)diphenylphosphine (296 mg, 1.1 mmol, 1 equiv), 5 mL hexanes and NaH (75 mg, 3.1 mmol, 3 equiv) were added in this order to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. Instant evolution of gas was observed. This was then sealed with a Teflon-lined cap about 5 min later and stirred at room temperature overnight. The reaction mixture was filtered through a fine frit and the residue was treated with 2 mL THF to give a fine suspension. The suspension was filtered through a pipette plug and washed with more THF. The filtrate was dried under under *vacuo* for several hours

before use to give the ligand Na-salt with one THF bound as a fluffy solid. (337 mg, 85 % yield).

In a separate scintillation vial equipped with a small stir bar, $py_2Ti(N^tBu)Cl_2$ (100 mg, 0.23 mmol, 1.0 equiv), Na-OAr' ligand (174 mg, 0.47 mmol, 2.0 equiv) and 5 mL C_6H_6 were added together. The mixture was stirred overnight at room temperature. Following which, the reaction mixture was filtered through a celite plug to remove gelatinous material and the filtrate was dried *in vacuo* to give a sticky orange liquid. The crude product was purified by dissolving in 2 mL C_6H_6 and slow diffusing with pentane to yield X-ray quality orange crystals of pure **d**. (34 mg, 13 % yield).

NMR spectra were taken of crystalline material, resulting in highly complex spectra indicative of fluxional species in solution.

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.66 – 8.59 (m, 2H, Ar-*H*), 8.26 (app t, $^3J_{HH}$ = 9.6 Hz, 1H, Ar-*H*), 8.04 (app t, $^3J_{HH}$ = 9.2 Hz, 1H, Ar-*H*), 7.72 – 7.29 (m, 18H, Ar-*H*), 7.23 – 7.08 (m, 5H, Ar-*H*), 6.96 – 6.42 (m, 14H, Ar-*H*), 6.07 – 5.92 (m, 1H, Ar-*H*), 5.44 – 5.34 (m, 1H, Ar-*H*), 0.87 – 0.79 (m, 15H, ^{t}Bu), 0.18 (s, 3H, ^{t}Bu).

^{13}C NMR (126 MHz, $CDCl_3$): δ 137.5, 137.4, 136.4, 136.3, 135.4, 135.2, 134.2 – 130.7 (m), 129.6, 128.7, 128.1, 127.1, 121.0 – 117.9, 69.6, 68.9, 32.8, 31.9, 30.8.
This is a partial line-list. Many C could not be positively be identified.

^{31}P NMR (162 MHz, $CDCl_3$): δ 7.62, 7.54, 5.44, 5.38, 3.47, 2.58, 2.52, -0.47, -2.73, -2.81.

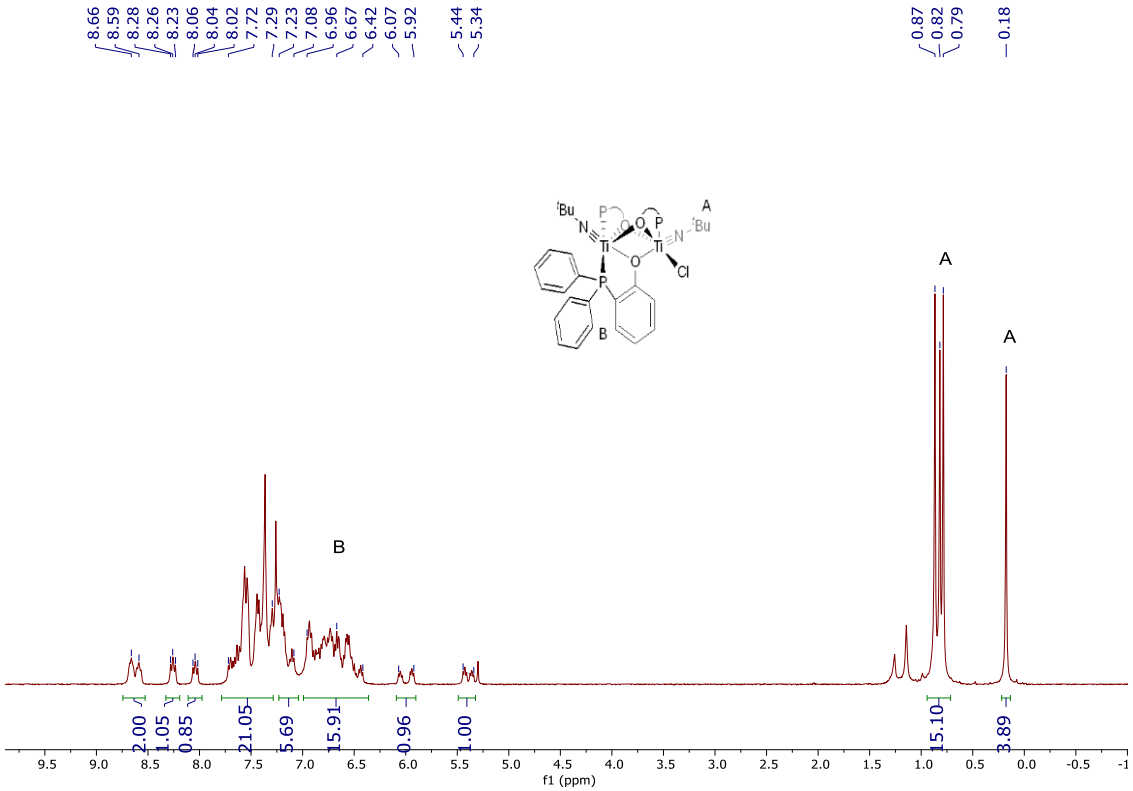


Figure S62. ^{1}H NMR spectrum of **d** in $CDCl_3$.

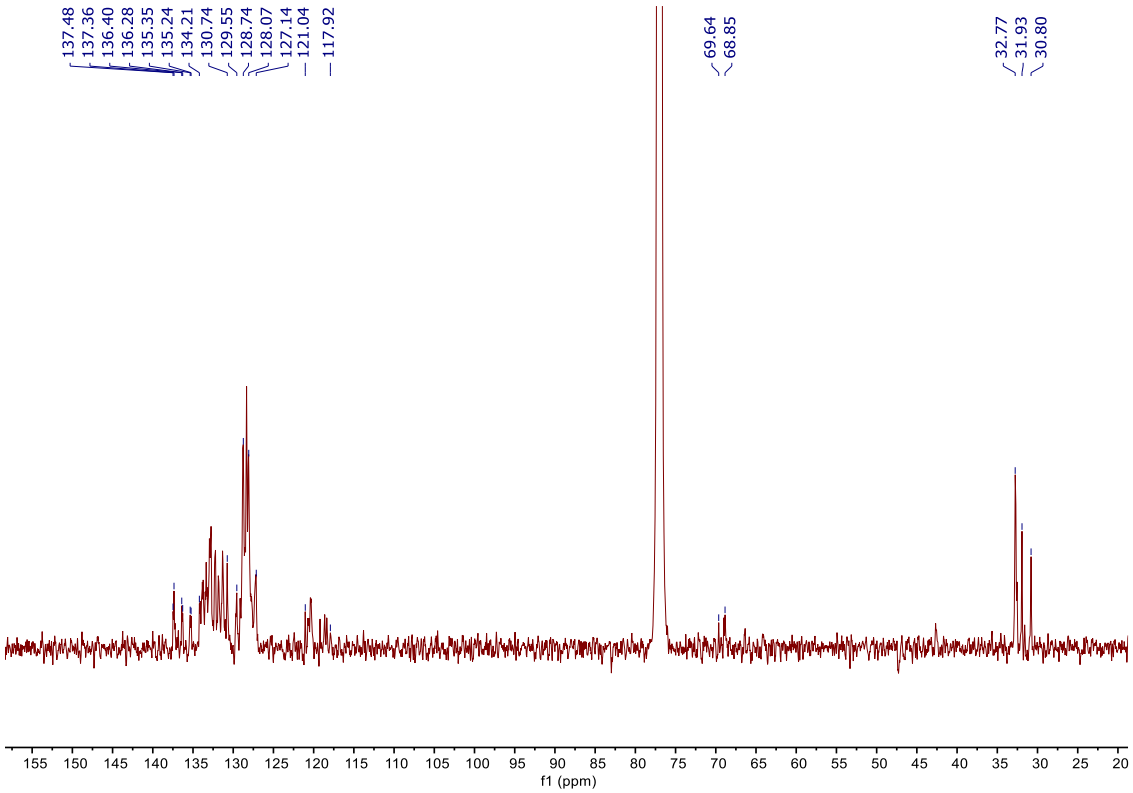


Figure S63. ^{13}C NMR spectrum of **d** in $CDCl_3$.

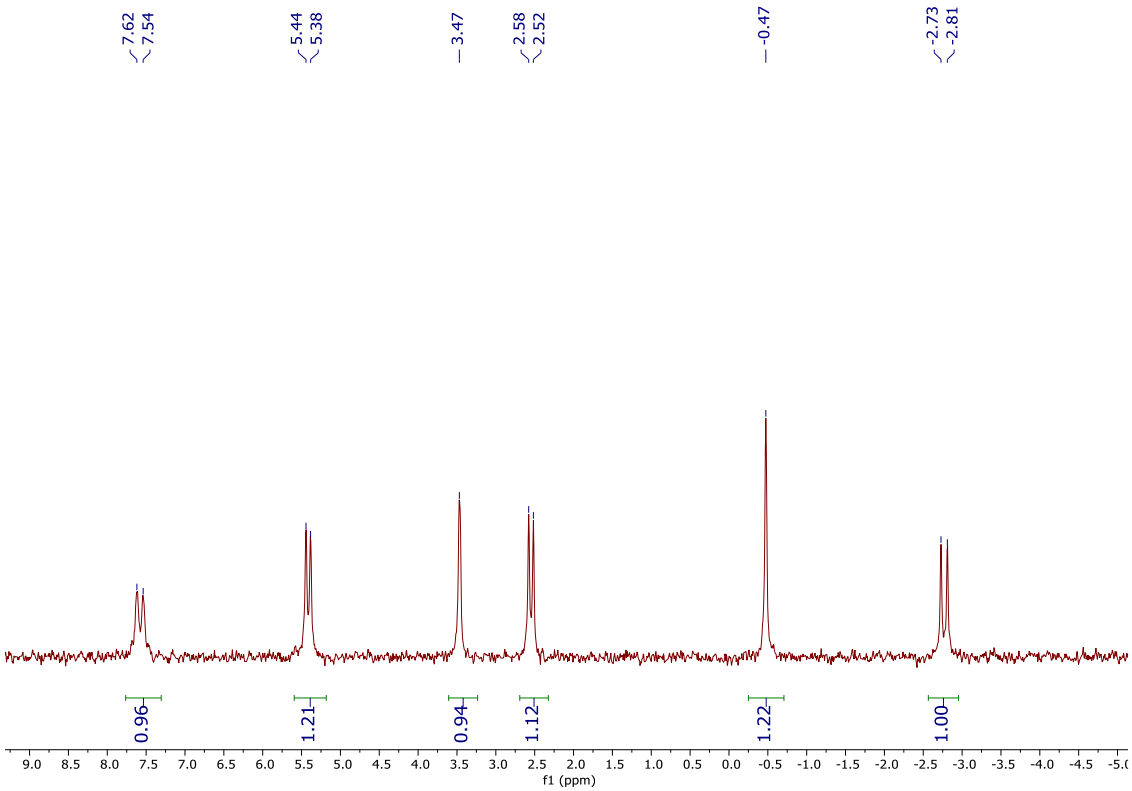


Figure S64. ^{31}P spectrum of **d** in $CDCl_3$.

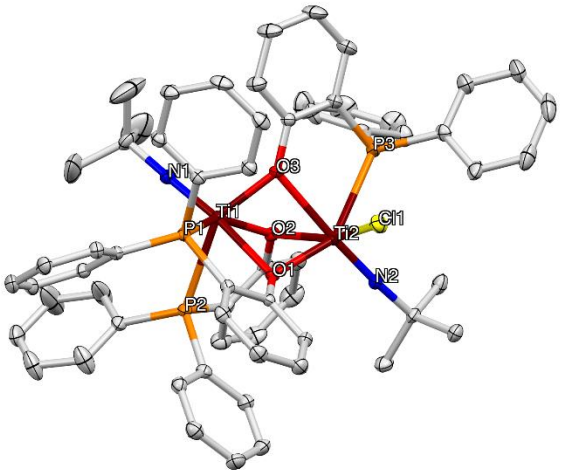
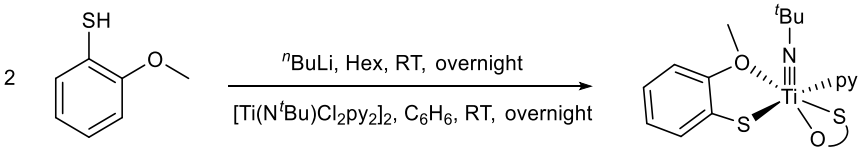


Figure S65. 50 % thermal ellipsoid drawing of **d**. Hydrogen atoms are omitted for clarity.

Synthesis of e



2-methoxythiophenol (1 g, 7.1 mmol, 1.0 equiv), 10 mL hexanes and ^{n}BuLi (3.4 mL, 2.5 M, 8.6 mmol, 1.2 equiv) were added in this order to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox in the order. This was then sealed with a Teflon-lined cap and stirred at room temperature overnight. The reaction mixture was filtered, washed with hexanes and the Li-salt residue was dried *in vacuo* for several hours before use (1.03 g, 99 % yield).

In a separate scintillation vial equipped with a small stir bar, $[py_2Ti(N^tBu)Cl_2]_2$ (206 mg, 0.48 mmol, 1.0 equiv), Li-salt residue (143 mg, 0.98 mmol, 2.0 equiv) and 2 mL C_6H_6 were added to form a deep red solution. The mixture was stirred overnight at room temperature. Following which, the reaction mixture was filtered through a celite plug and the filtrate was dried *in vacuo*. The crude product was purified by dissolving in 1 mL PhMe and cooling in a -35 °C freezer to yield X-ray quality yellow crystals of pure **e**. (75 mg, 33 % yield).

^{1}H NMR (500 MHz, C_6D_6): δ 9.00 (d, $^3J_{HH}$ = 4.9 Hz, 2H, *o*-py-*H*), 7.84 (br s, 2H, Ar-*H*), 6.85 (br t, 2H, Ar-*H*), 6.70 – 6.66 (m, 3H, Ar-*H*), 6.35 (app t, $^3J_{HH}$ = 7.1 Hz, 2H, Ar-*H*), 6.17 (d, $^3J_{HH}$ = 8.0 Hz, 2H), 3.94 (br s, 3H, O-C*H*$_3$), 3.24 (br s, 3H, O-C*H*$_3$), 1.31 (s, 9H, ^{t}Bu-*H*).

^{13}C NMR (126 MHz, C_6D_6): δ 151.2, 138.0, 137.9, 129.3, 128.6, 128.4, 125.7, 123.8, 70.3, 31.6, 21.4.
This is a partial line-list. Resonances associated with a quaternary C were not observed.

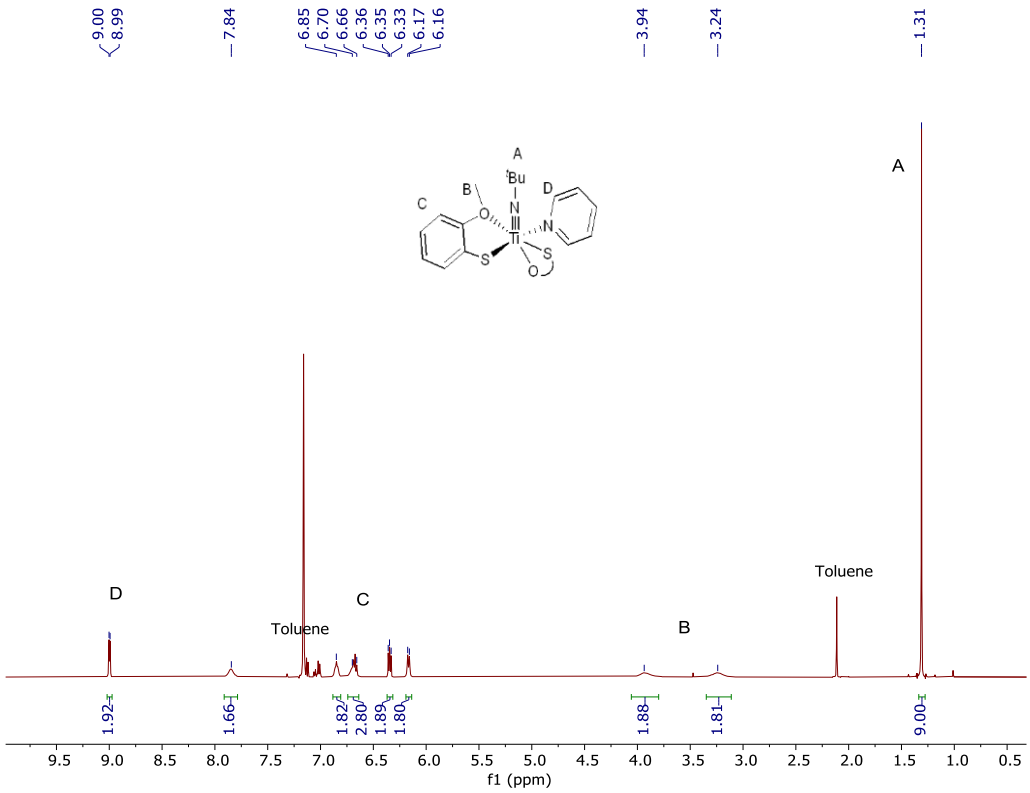


Figure S66. ^{1}H NMR spectrum of **e** in C_6D_6.

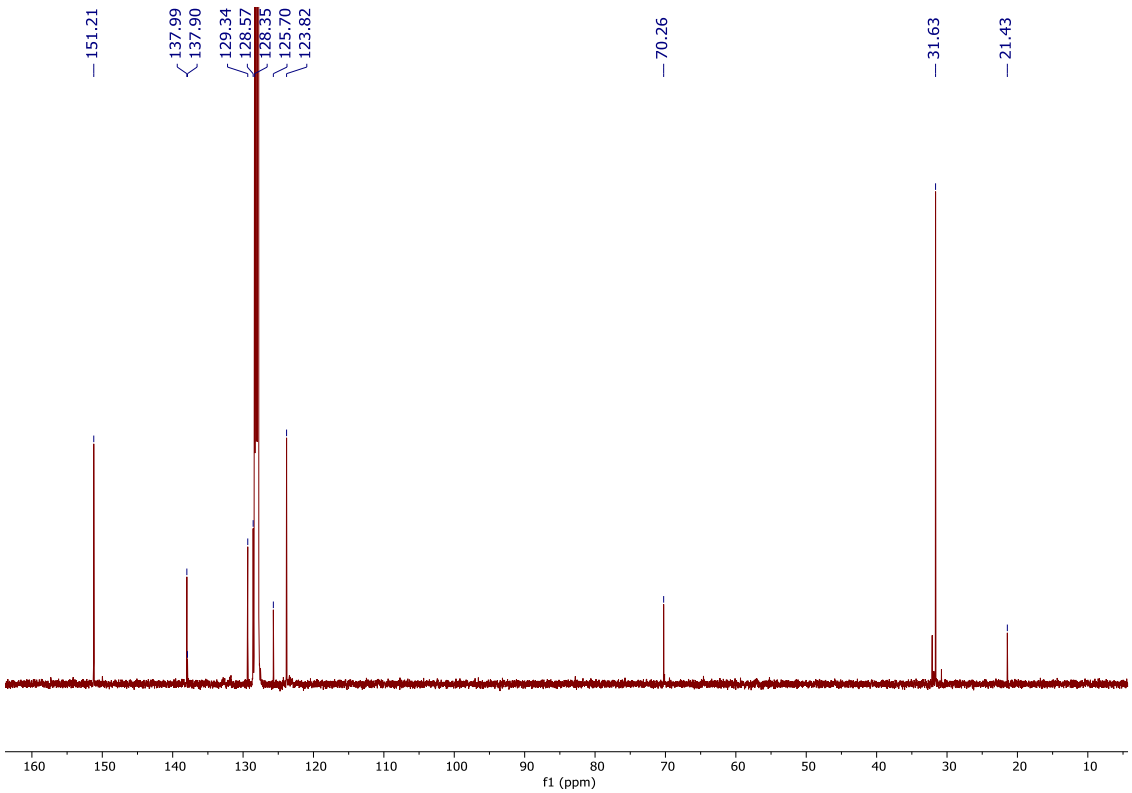
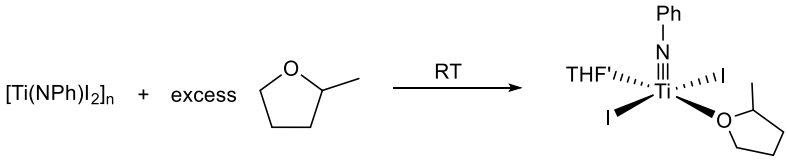


Figure S67. ^{13}C NMR spectrum of **e** in C_6D_6.

Synthesis of f



$[Ti(NPh)I_2]_n$ (170 mg, 0.43 mmol, 1.0 equiv) and 2-methyltetrahydrofuran (1 mL, 9.9 mmol, 23.0 equiv) were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox and stirred to dissolve all solids. 1 mL of hexanes was layered onto the mixture and the mixture was left to cool in a -35 °C freezer to yield solid material. The reaction mixture was filtered through a fine frit, washed with 5 mL hexanes and dried under *vacuo* for several hours to yield pure **f** (60 mg, 25 % yield).

^{1}H NMR (500 MHz, C_6D_6): δ 7.24 (d, $^3J_{HH}$ = 7.9 Hz, 2H, *o*-NPh-*H*), 6.92 (t, $^3J_{HH}$ = 7.7 Hz, 2H, *m*-NPh-*H*), 6.70 (t, $^3J_{HH}$ = 7.5 Hz, 1H, *p*-NPh-*H*), 5.09 (br s, 2H, 2-$C_5H_{10}O$-*H*), 4.24 (br s, 4H, 5-$C_5H_{10}O$-*H*), 1.66 (br s, 2H, $C_5H_{10}O$-*H*), 1.47 (br s, 2H, $C_5H_{10}O$-*H*), 1.37 – 1.29 (m, 2H, $C_5H_{10}O$-*H*), 1.24 – 1.22 (m, 6H, $C_5H_{10}O$-C*H*$_3$), 1.01 – 0.95 (m, 2H, $C_5H_{10}O$-*H*).

^{13}C NMR (101 MHz, C_6D_6): δ 162.1, 128.7, 124.1, 124.1, 86.9, 74.8, 32.2, 24.5, 21.3.

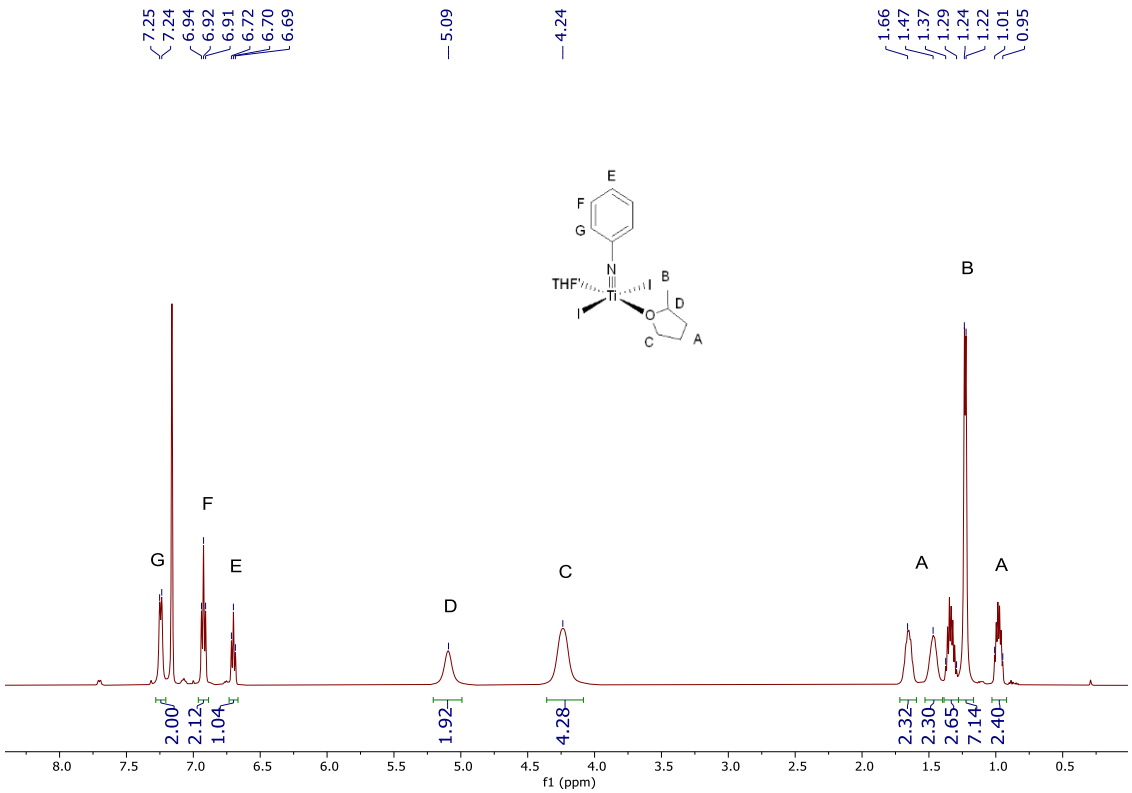


Figure S68. ^{1}H NMR spectrum of **f** in C_6D_6.

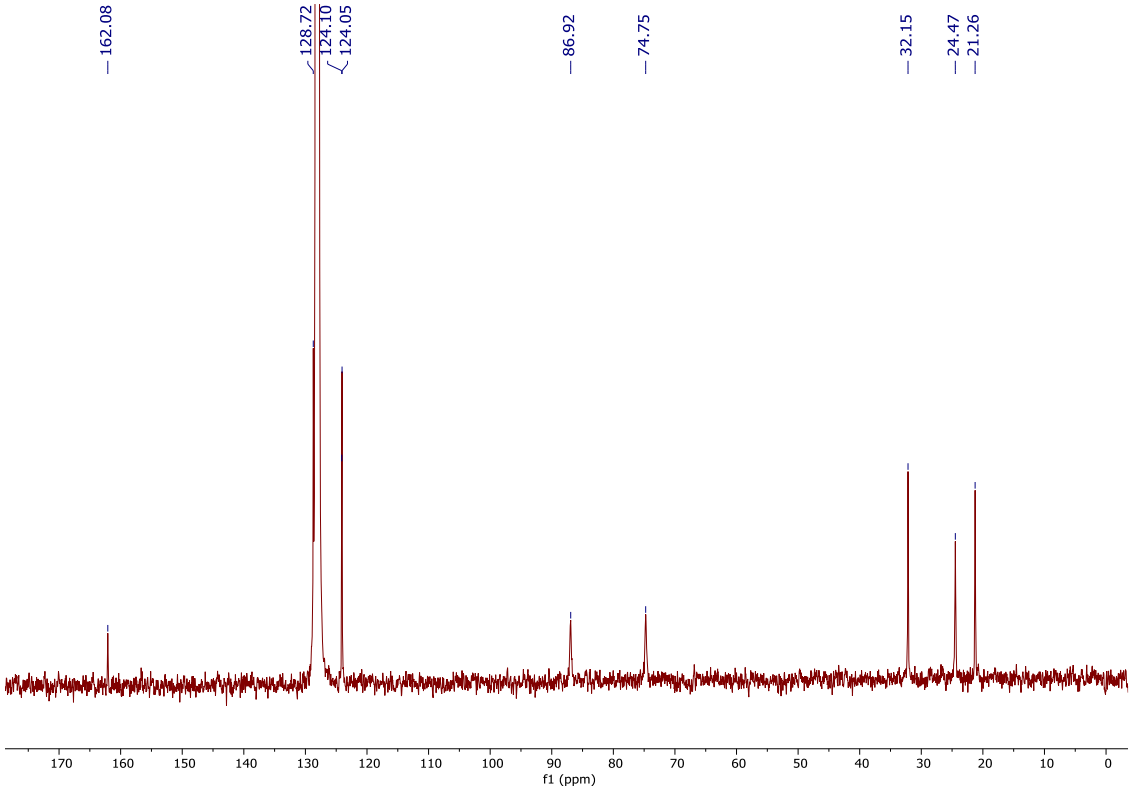


Figure S69. ^{13}C NMR spectrum of **f** in C_6D_6.

Representative Ti imido catalysts platforms tested

Over the years, we have also attempted catalysis with other more complex ligand classes that are found widely in Ti-catalyzed hydroamination and polymerisation literature, however they often gave poor to moderate yields and selectivity.

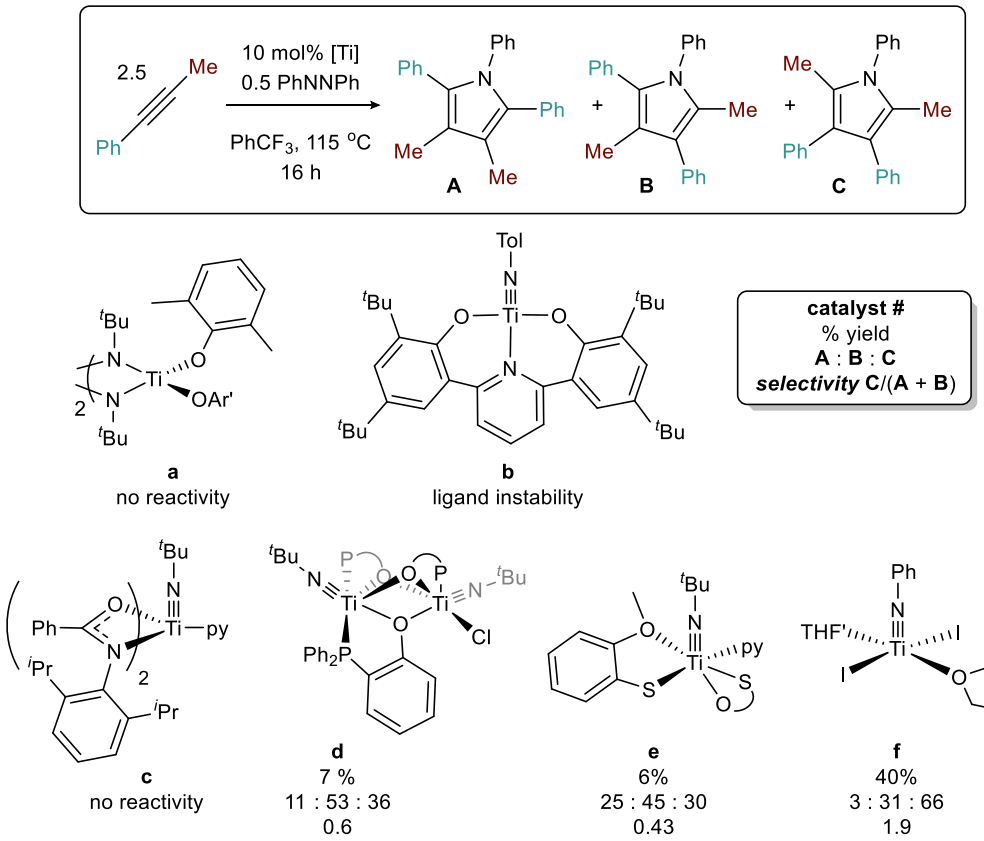


Figure S70. Representative ligand scaffolds common in literature tested in catalytic pyrrole formation with PhCCMe. See main text for more information.

Ligand scaffolds attempted

Represented below are ligands that we attempted but are unable to bind to a Ti imido moiety, presumably due to steric hindrance imposed by the ortho substituents.

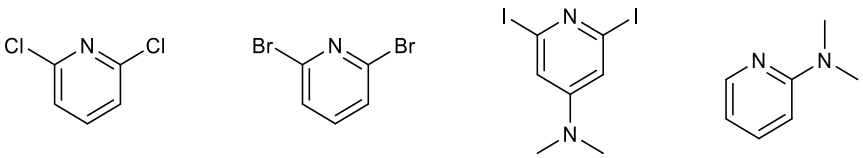


Figure S71. Ligand scaffolds that were unsuccessful in coordinating to Ti.

General procedure for catalytic pyrrole formation

Azobenzene (364 mg, 2 mmol), phenyl propyne (1.16 g, 10 mmol) and 1,3,5-trimethoxybenzene (336 mg, 2 mmol, as internal standard) were added to a 10 mL volumetric flask and diluted to 10 mL with $PhCF_3$ to make a stock solution. For each catalytic run, a precatalyst (10 mol % Ti, 0.01 mmol) and 0.5 mL of stock solution were added to a NMR tube in a N_2-filled glovebox. This was then sealed with a NMR cap and electrical tape before heating at 115 °C for 16 h.

Representative Catalytic Spectra

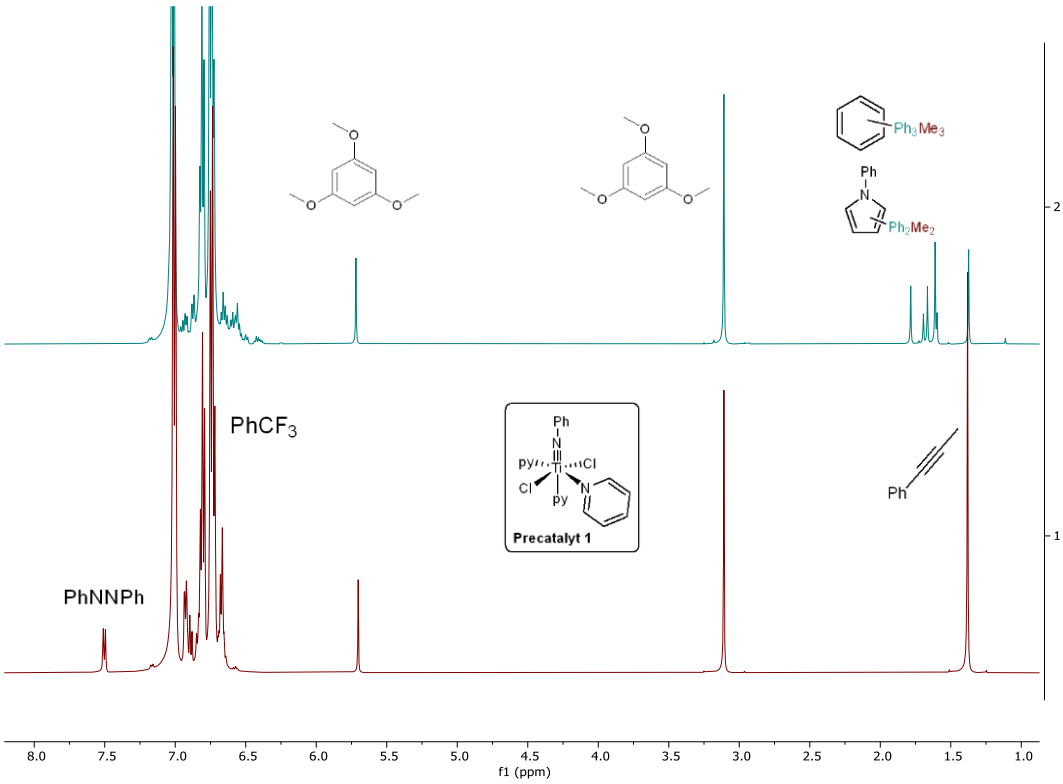


Figure S72. Representative stacked catalytic ^{1}H NMR spectra at t = 16 h (2) and t = 0 h (1) in $PhCF_3$ (115 °C/ Precatalyst **1**).

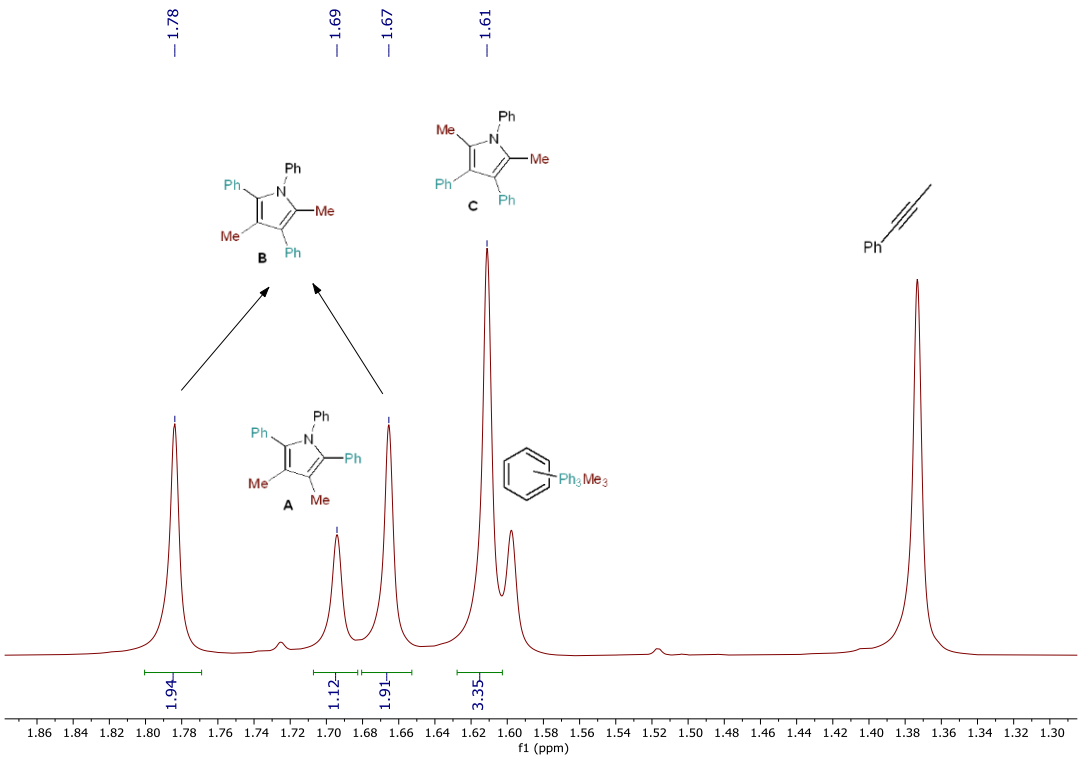


Figure S73. Representative zoom-in catalytic ^{1}H NMR spectra at t = 16 h in $PhCF_3$ (115 °C/ Precatalyst **1**).

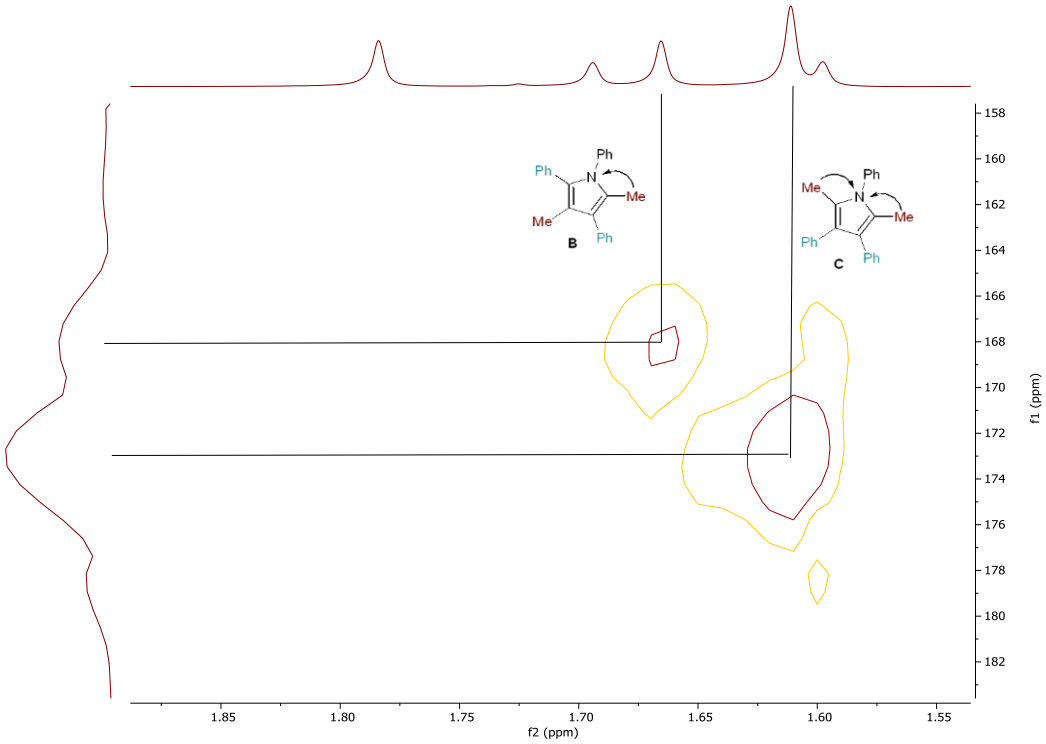


Figure S74. Representative catalytic ^{1}H-^{15}N HMBC NMR spectra at t = 16 h in $PhCF_3$ (115 °C/ Precatalyst **1**).

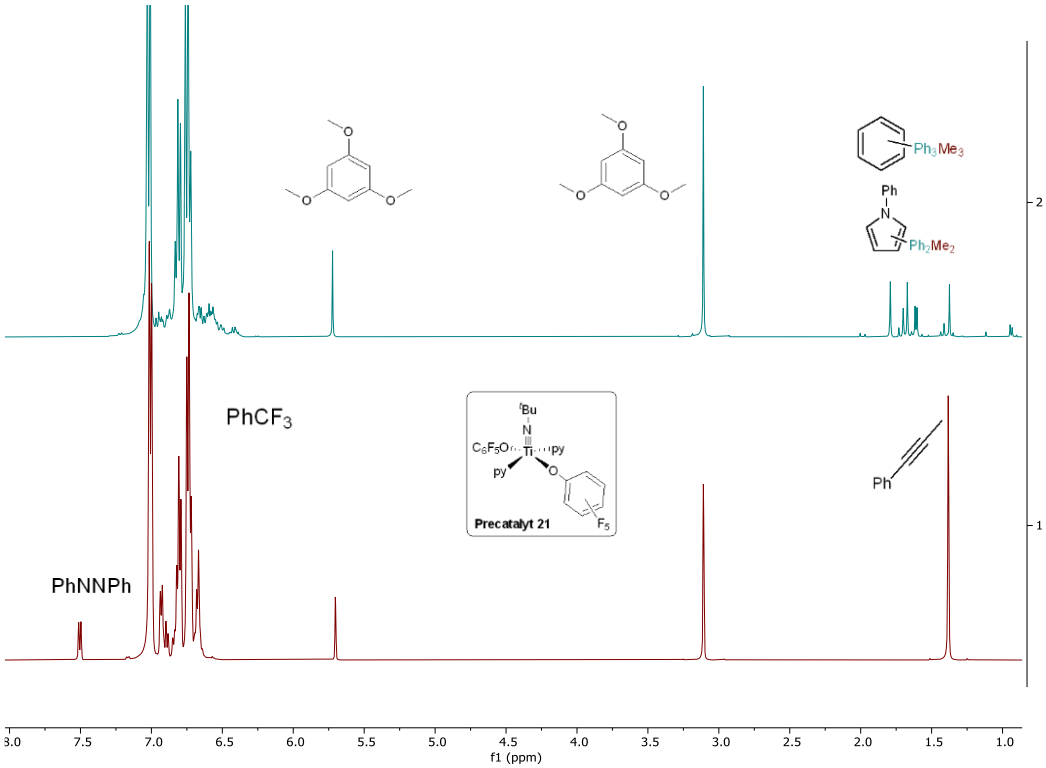


Figure S75. Representative stacked catalytic ^{1}H NMR spectra at t = 16 h (2) and t = 0 h (1) in $PhCF_3$ (115 °C/ Precatalyst **21**).

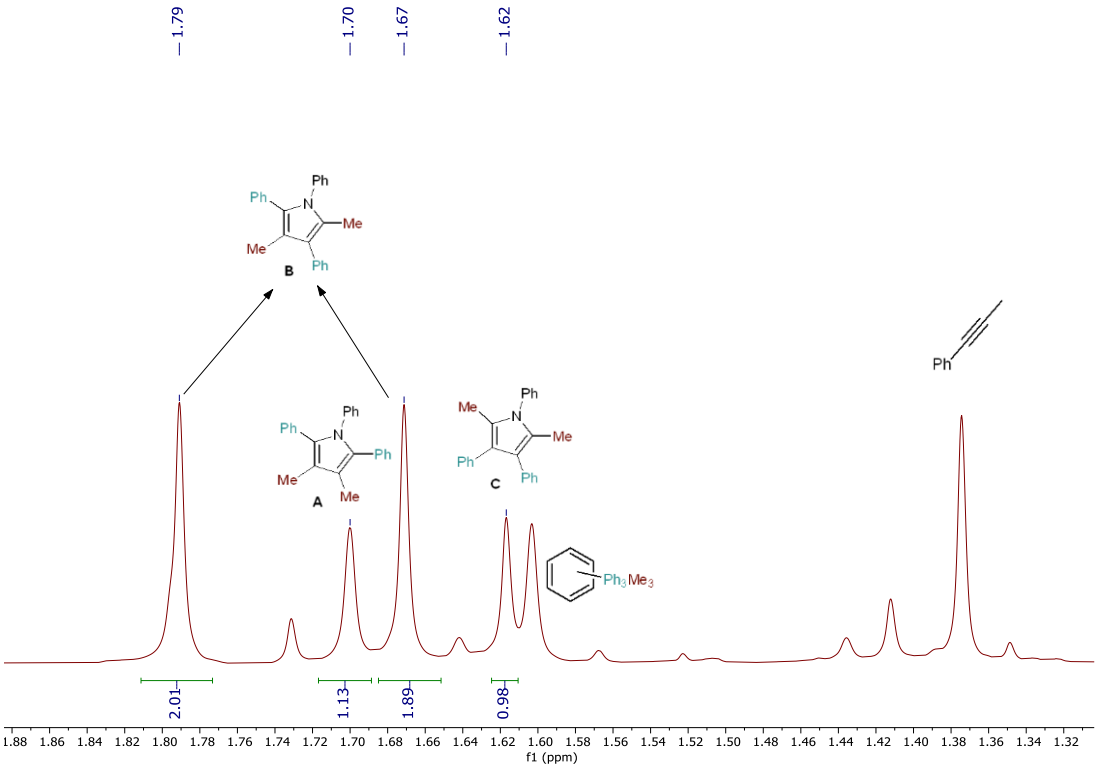


Figure S76. Representative zoom-in catalytic ^{1}H NMR spectra at t = 16 h in $PhCF_3$ (115 °C/ Precatalyst **21**).

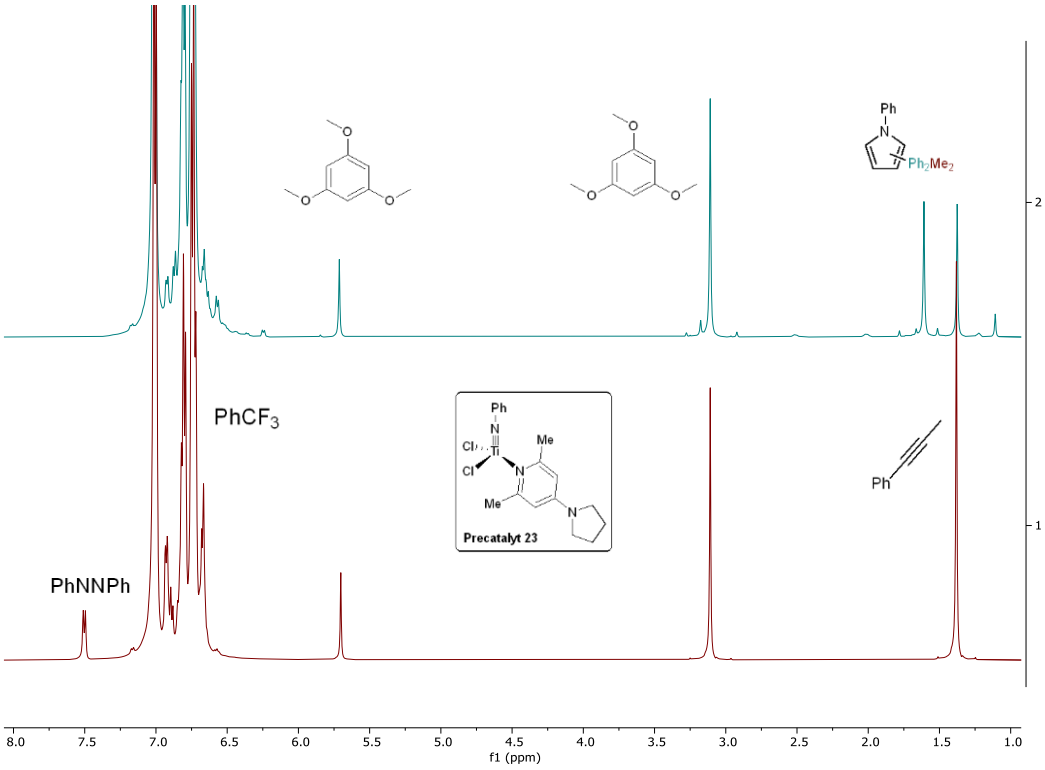


Figure S77. Representative stacked catalytic ^{1}H NMR spectra at t = 16 h (2) and t = 0 h (1) in $PhCF_3$ (115 °C/ Precatalyst **23**).

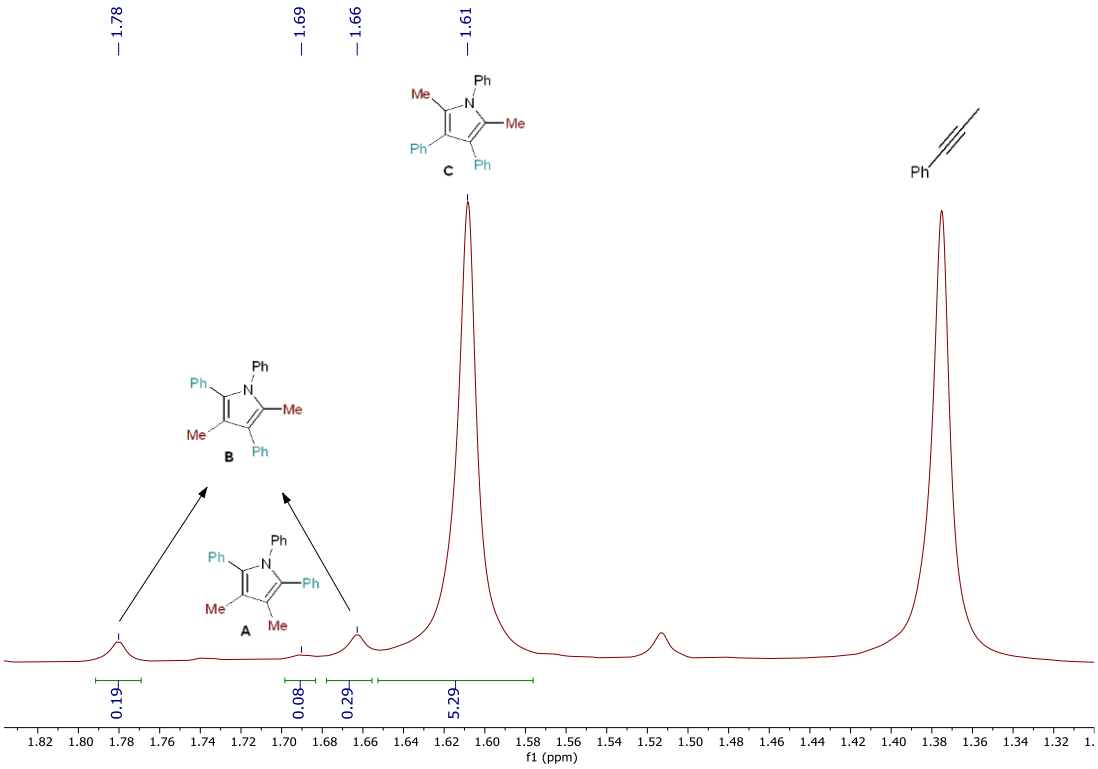


Figure S78. Representative zoom-in catalytic ^{1}H NMR spectra at t = 16 h in $PhCF_3$ (115 °C/ Precatalyst **23**).

Timepoint studies of catalytic reaction with Precatalyst 23

Precatalyst 23 (10 mol % Ti, 0.01 mmol, 0.02 M) and 0.5 mL of stock solution were added to a NMR tube in a N_2-filled glovebox. This was then sealed with a NMR cap and electrical tape before heating at 115 °C. Timepoints were taken at regular intervals (hourly for the first 7 hours and then at the end of reaction).

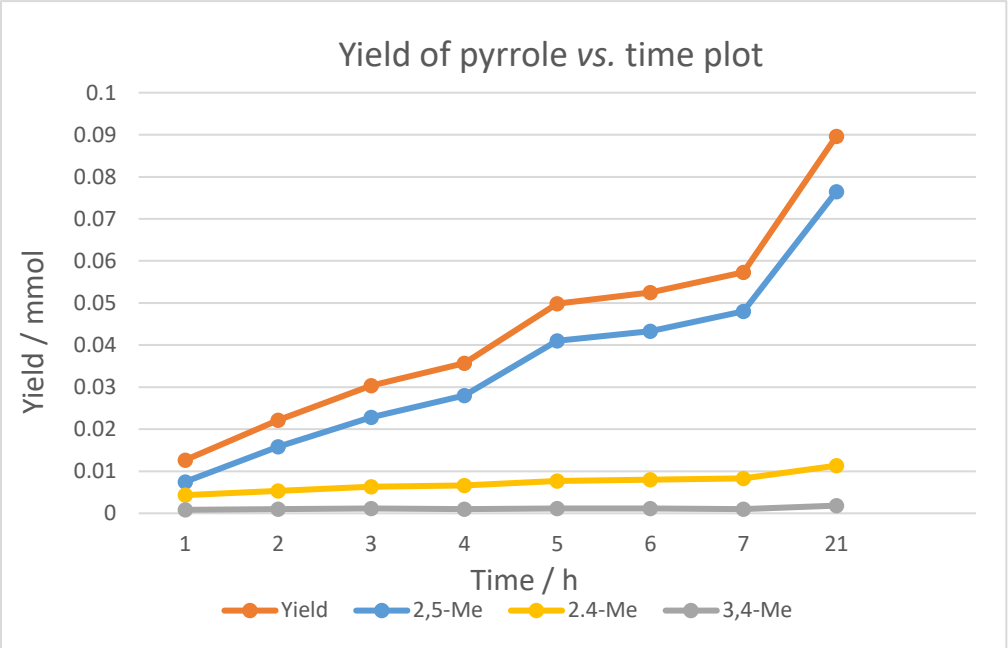


Figure S79. Yield of pyrrole *vs* time for a catalytic reaction with precatalyst 23.

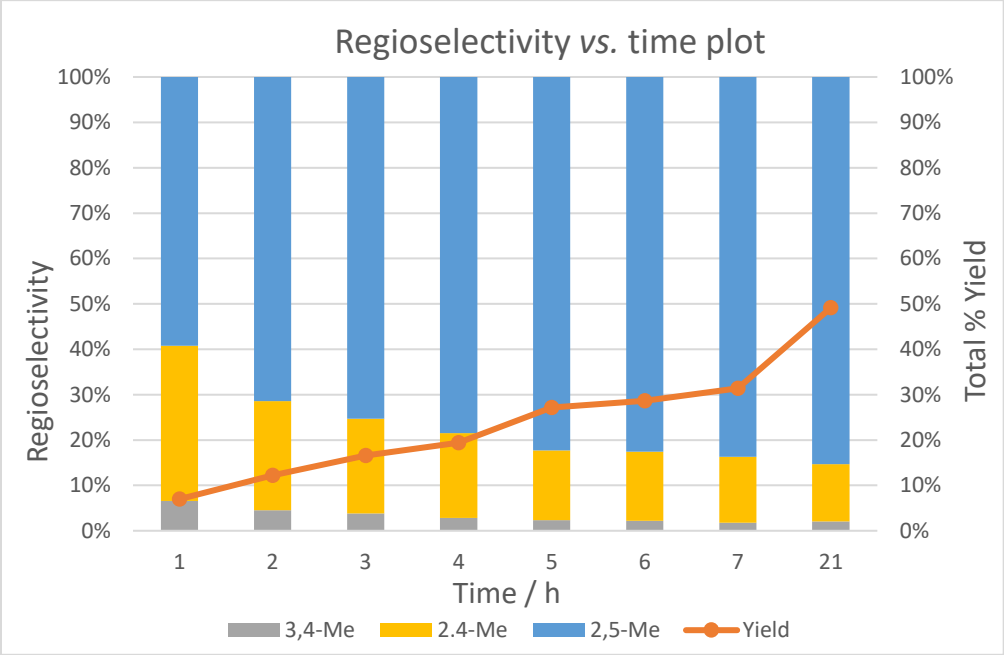
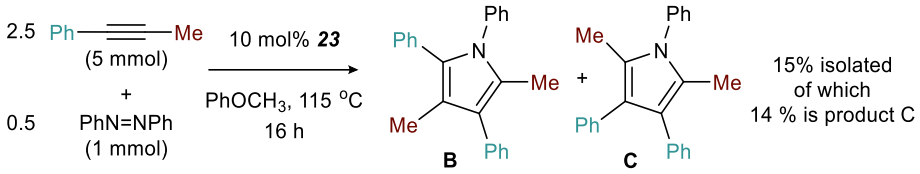


Figure S80. Regioselectivity of pyrroles (left axis) and total percent yield of pyrroles (right axis) *vs* time for a catalytic reaction with precatalyst 23.

Isolation of catalytic reaction



Precatalyst **23** (39 mg, 0.1 mmol, 0.1 equiv), azobenzene (182 mg, 1 mmol, 1 equiv), phenyl propyne (0.63 mL, 580 mg, 5 mmol, 5 equiv) and 4.5 mL of anisole were added to a 20 mL scintillation vial equipped with a small stir bar in a N_2-filled glovebox. This was then sealed with a Teflon-lined cap and electrical tape before heating at 115 °C for 16 h. The reaction mixture was dried under *vacuo* while heating at 80 °C for several hours. Crude product was dry loaded under neutral alumina and purified using a long column packed with neutral alumina using 0 % → 2 % ethyl acetate/hexanes eluent. A mixture of pyrrole regioisomers (Product **B** & **C**) was isolated (Mass = 100 mg, 15 % isolated yield). The low isolated yield for this reaction is ascribable to the difficulties in separating out the pyrrole regioisomers.

Product **C** (2,5-dimethyl-3,4-diphenyl-1-phenyl-1H-pyrrole)
^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.39 (m, 2H, *o*-NPh-*H*), 7.34 – 7.31 (m, 2H, *m*-NPh-*H*), 7.28 (br s, 1H, *p*-NPh-*H*), 7.16 – 7.12 (m, 4H, Ar-H), 7.05 – 7.03 (m, 6H, Ar-*H*), 2.00 (s, 6H, CH_3)
Peak at 2.05 ppm belongs to product B (2,4-dimethyl-3,5-diphenyl-1-phenyl-1H-pyrrole).

^{13}C NMR (101 MHz, $CDCl_3$): δ 139.0, 136.4, 130.5, 129.3, 128.6, 127.9, 126.0, 125.3, 120.1, 12.1.
Resonances for the last quaternary C could not be positively identified from the rest of the peaks.

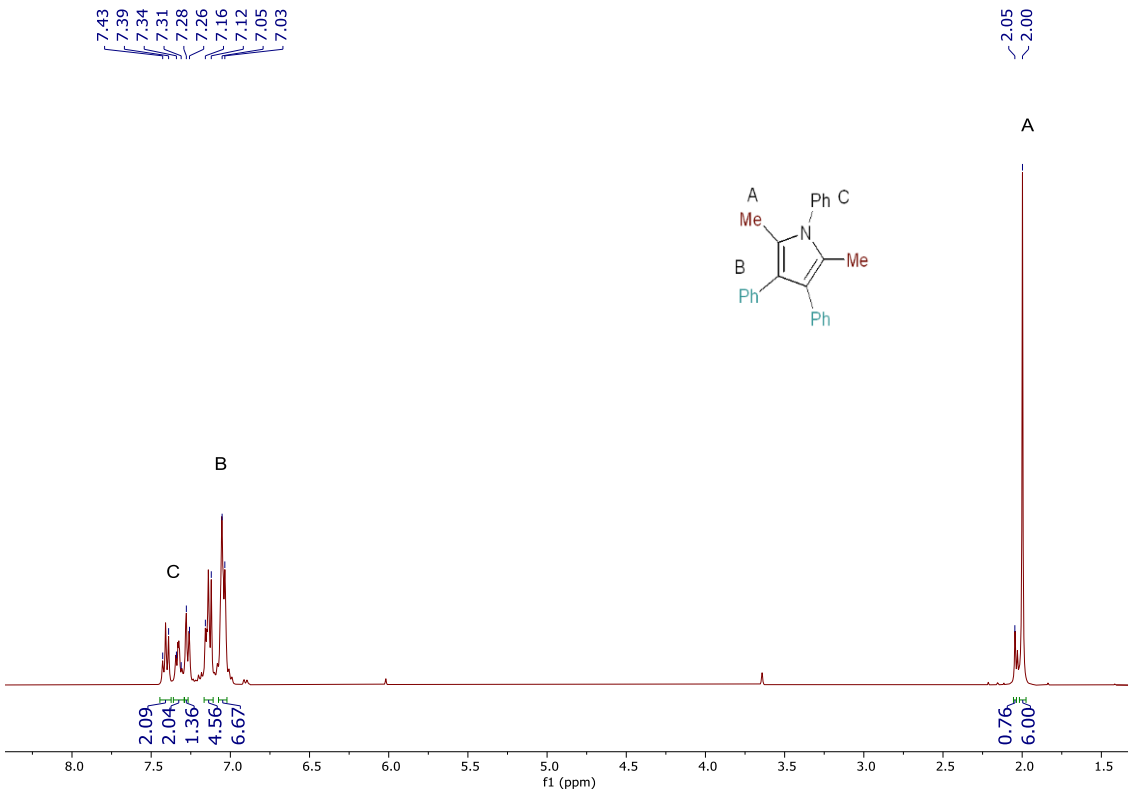


Figure S81. ^{1}H NMR spectrum of isolation attempt of product **C** in $CDCl_3$.

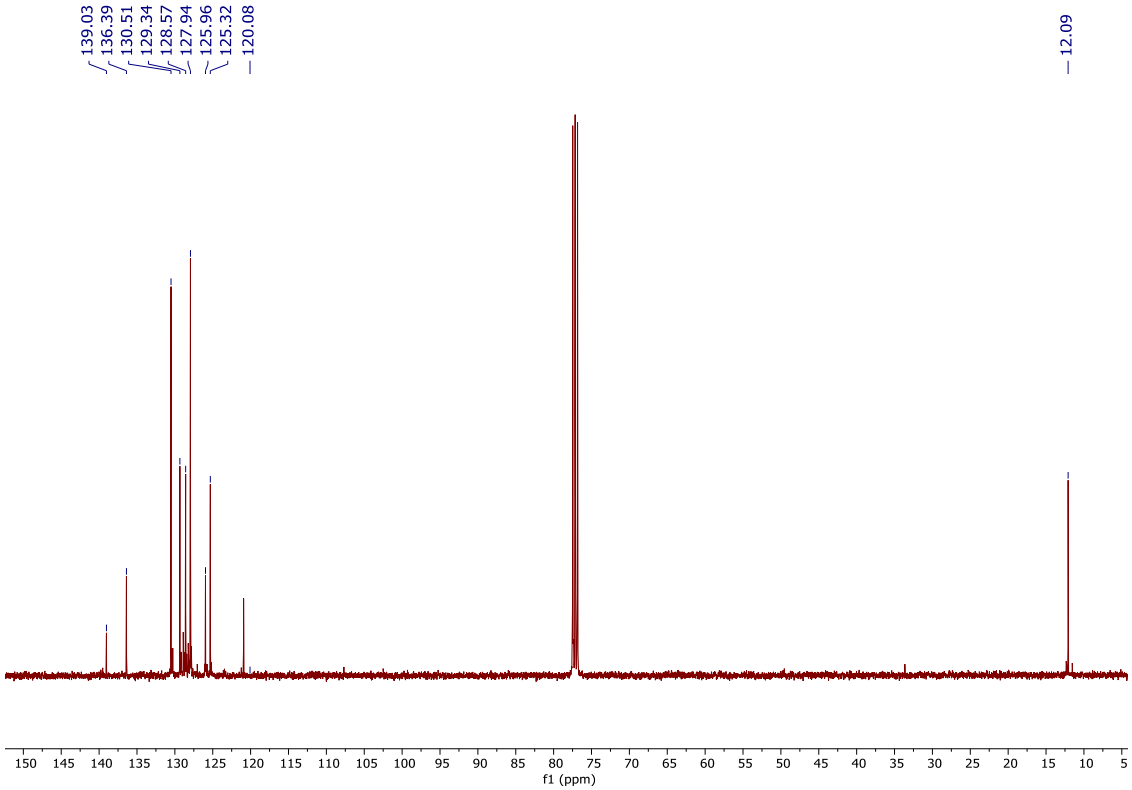


Figure S82. ^{13}C NMR spectrum of isolation attempt of product **C** in $CDCl_3$.

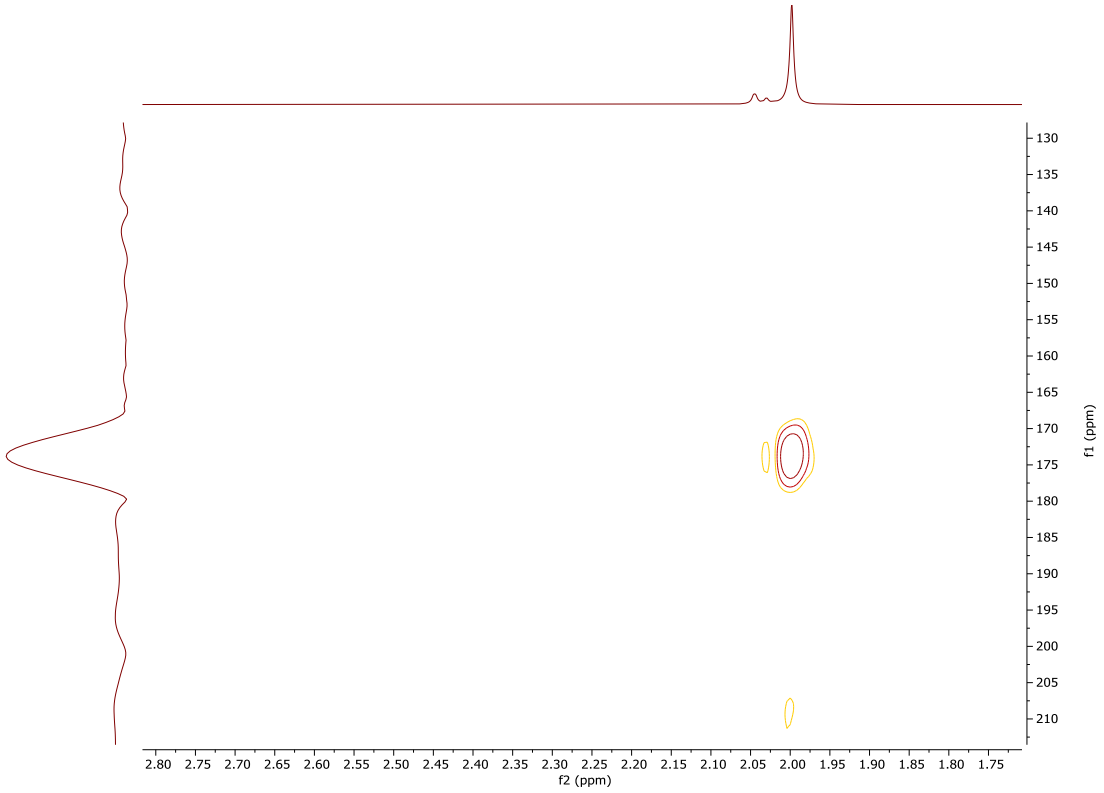


Figure S83. ^{1}H-^{15}N spectrum of isolation attempt of product **C** in $CDCl_3$.

Table S1. Refined data and cell parameters for X-ray Structures

	7	**12**	**14**	**e**
CCDC Number	1994540	1994543	1994542	1994541
Empirical Formula	$C_{50}H_{62}Cl_4N_6Ti_2$	$C_{30}H_{31}N_5Ti$	$C_{82.52}H_{77.04}N_{10}Ti_2$	$C_{80}H_{78}ClN_2O_3P_3Ti_2$
Formula weight	984.65	509.50	1304.64	1339.60
Temperature (K)	100(2)	125(2)	125(2)	125(2)
a, Å	9.8644(9)	12.9396(11)	9.4737(4)	24.117(3)
b, Å	12.5071(11)	12.9396(11)	36.2928(16)	15.0952(14)
c, Å	12.5128(10)	35.996(4)	19.9562(10)	20.943(2)
α, deg	116.587(2)	90	90	90
β, deg	93.947(3)	90	95.112(2)	114.010(4)
γ, deg	112.994(3)	90	90	90
Volume, $Å^3$	1213.89(18)	6026.9(12)	6834.2(5)	6964.5(12)
Z	1	8	4	4
Crystal System	triclinic	tetragonal	Monoclinic	Monoclinic

Space Group	P -1	P $4_3 2_1 2$	P $2_1/n$	P $2_1/c$
d_{calc}, gcm^{-3}	1.347	1.123	1.268	1.278
θ range, deg	2.35 to 26.13	2.30 to 26.06	2.37 to 29.93	2.15 to 30.24
μ, mm^{-1}	0.590	0.308	0.288	0.387
Abs. Correction	Multi-scan	Multi-scan	Multi-scan	Multi-scan
GOF	1.058	1.071	1.065	1.227
R_1 ,[a]	$R1$ = 0.0506	$R1$ = 0.0237	$R1$ = 0.1068	$R1$ = 0.1468
wR_2 [b] [I>2σ(I)]	$wR2$ = 0.1390	$wR2$ = 0.0617	$wR2$ = 0.1152	$wR2$ = 0.1525

[a] $R_1 = \sum||F_o| - |F_c||/\sum|F_o|$. [b] $wR_2 = [\sum[w(F_o^2-F_c^2)^2]/\sum[w(F_o^2)^2]]^{1/2}$.

Note: the .cif for **e** has a B-level alert in the .checkcif file. Reflections -2 4 5 (resolution 2.80 Å) and -9 1 4 (resolution 2.62 Å) have much less intensity than is calculated. Since these reflections are of low resolution, they are left in the final refinement and do not significantly affect the overall result.

Principal Component Analysis

Descriptor Set Details methodology

The 22 variables shown below were calculated at the M06/6-311g(d,p) level of theory for each catalyst in accordance with previous computational experiments.[10] Variables denoted with "*catalyst*" were calculated using the monomer of the corresponding N-Ph Ti imido complex. Variables denoted with "*free pyridine*" are calculated for the corresponding free pyridine. "*X*" and "*L*" denotations refer to X- and L-type ligands around the Ti-center.

Starting points for catalyst geometry optimization were either crystal structures if available or structures modified using the structure viewer, Avogadro. Geometric parameters were averaged when appropriate.

% buried volume was calculated using the online module, SambVca 2.1.[11]

NMR chemical shifts were calculated using gauge independent atomic orbitals and scaled according to the CHESHIRE CCAT website.[12]

All ortho atom related shifts were averaged across the relevant positions.

Ti-X donor BDE's were tabulated from the Lange's Handbook of Chemistry[13] and references therein.

Composite donor atom polarizability were calculated using the static dipole polarizability of the neutral atoms[14] surrounding the Ti center of each catalyst weighted by the number of atoms.

Mulliken electronegativities (EN) were tabulated from Bratsch.[15]

Free L ligand quadrupole moment refers to the Q_{zz} quadrupole vector.

Proton affinity was calculated as the sum of all thermal and free energies using the equation: Proton Affinity = LH – (L + H^+).

Table S2. List of all descriptors included in initial PCA basis set.

Descriptor Index	Descriptors
1	Catalyst HOMO
2	Catalyst LUMO
3	Free pyridine HOMO
4	Free pyridine LUMO
5	% Buried Volume[11]
6	Free pyridine *p*- ^{13}C NMR shift
7	Free pyridine *o*- ^{13}C NMR shift
8	Catalyst-bound pyridine *p*- ^{13}C NMR shift
9	Catalyst-bound pyridine *p*- ^{13}C NMR shift
10	Composite donor atom polarizability
11	Ti-X donor BDE
12	X donor Mulliken EN
13	Free pyridine quadrupole moment
14	Free pyridine N Mulliken charge
15	Catalyst-bound pyridine N Mulliken charge
16	Catalyst-bound X donor atom Mulliken charge
17	Free pyridine proton affinity
18	Imido-Ti-N_{py} angle
19	Ti-imido bond length
20	Ti- N_{py} bond length
21	Ti-X bond length
22	X-Ti-X angle

Automated Script Details

Overview

A Matlab script called *PCA4U2* was developed in-house in order to facilitate *de novo* catalyst design using principal component analysis (PCA). The script is designed for non-expert end users and uses a series of prompts to direct users. The script presents the option for regression *via* unsupervised principal component analysis (PCA), iterative supervised PCA (ISPCA), weighted PCA, or subspace specific PCA (determined by k-means clustering) and the associated computational costs as well as the option to use a previously determined basis set. Linear regression is conducted against the top 3 components. Data must be inputted as instructed. The catalyst variable matrix should be added as a numeric matrix with each column corresponding to a new variable and each row corresponding to a new catalyst. The script will automatically center and normalize each variable by standard deviation. Observables must be added as a single column vertical vector containing positive values.

Search algorithm details

Unsupervised PCA: Unsupervised PCA conducts PCA on the variable basis set without modification.

ISPCA: ISPCA is conducted with two search options: find the best possible fit or find the best possible fit with the lowest number of variables. Fit in all cases was determined by R^2. While Q^2 is the more desirable statistical parameter for predictive power, computational costs

precluded its general use (a 22-variable basis set with a 14-catalyst training set requires ~ 2.1 million calculations for R^2 and ~ 62 million calculation for Q^2). Searches are performed by conducting PCA on each combination of 2 variables and iteratively increasing the group size to the total number of variables. The best fit of each combination is collected and stored in a matrix. Optimization toward the best possible fit was accomplished by searching this matrix for the highest R^2 value. Optimizing toward the lowest number of variables was accomplished by terminating the search loop when R^2 of the current group became lower than the previous group.

Weighted PCA: Weighted PCA was performed by conducting unoptimized PCA on the full variable basis set and determining the pairwise distances between a test catalyst and the remaining training set. These distances were transformed into coefficients for each catalyst by applying the following function: $coefficient = e^{-distance/25}$. ISPCA or unsupervised PCA can be conducted with this weighted catalyst set.

Subspace selective ISPCA: ISPCA performed on a selected PCA subspace was accomplished by conducting unsupervised PCA on the full catalyst/variable set and using a dynamic k-means clustering algorithm (the number of clusters was increased iteratively until intracluster distances no longer decreased substantially) to identify which cluster contains the desired "test" catalyst. This cluster was considered the initial PCA subspace. ISPCA was then conducted on this subspace. The R^2 value of the optimal variable basis set was stored. The next closest catalyst to the subspace was identified using pairwise distances from the initial unsupervised PCA space. This nearest neighbor was added to the subspace and ISPCA was again conducted. This process was iterated until the R^2 of the current subspace fell below an arbitrary cutoff of $R^2 = 0.99$.

A*nalysisparalysis:* An automated script which accepts a sequence of catalysts to add to an existing training set. The catalysts can be added in one sequence or all possible permutations of the sequence. The resulting expanding training set is fit using either MLR or ISPCA analysis.

- ISPCA analysis was executed as described above.
- MLR analysis was conducted using the native Matlab function *stepwiselm*. The function was initiated using the "constant" option so it would operate as a forward stepwise regression. P-value entry and exit values were set at 0.20 and 0.21 respectively, and were determined empirically by finding the minimum values that would result in a model $R^2 > 0.9$.

Output details

All PCA was conducted using the native Matlab *pca* function. ISPCA was conducted using the native *nchoosek* function and a homebrewed algorithm. The script displays 7 figures for qualitative and quantitative multivariate analysis using PCA.

Figure 1 shows a variable Pearson correlation matrix on the left and univariate analysis on the right (See **Figure S84**). An additional figure will appear showing the best univariate fit if the maximum R^2 of the univariate analysis exceeds 0.9.

Figure 2 is a Scree-type plot showing each principal component's contribution to the total data variance (See **Figure S85**).

Figure 3 shows the data recast over the new PCA axes as a biplot and grouped using k-means clustering (See **Figure S86**). Dynamic k-means clustering is achieved using the native Matlab *ischange* function to determine the proper number of partitions in the corresponding Voronoi cells.

Figure 4 shows the model fit as a predicted vs experimental values plot. Script will output basic statistics (including last one out cross validation) and equation to linear least squares fit regression (See **Figure S87**). Note that the regression equation is related to the centered and standard deviation normalized data.

Figure 5 shows a color-coordinated PCA map in 3 dimensions where the coordinate space is correlated against the entered observable (See **Figure S88**). Color banding is set to the error in the regression and constructed using an interpolated patch object.

Figure 6 is a 2D projection of *Figure 5* against desired principal components for publication/presentation purposes (See **Figure S89**).

Figure 7 is a plot of pairwise PCA Euclidean distance versus pairwise observable change to demonstrate how different a new catalyst must be from the parent – this metric is an effective molecular ruler for catalyst design (See **Figure S90**). The coloration is based on successively increasing standard deviations (σ) in the pairwise observable change starting from zero (red) to $3\times\sigma$ (green).

Prediction of new data

The final portion of the script is designed to predict new data. The script outputs a variety of scaling vectors and regression coefficients.

groups contain the regression statistics for alternate variable groupings.

groupwinner contains the optimized variable basis set.

PCRreal contains the normalized coefficients of each variable in the chosen principal component regression.

PCR contains the regression coefficients associated with correlating PCA space and observable.

meann and *stdd* are the mean and standard deviation of each column in the supplied data set, respectively.

Loadings are the variable loadings in the displayed components.

Yreal are the model's predicted values for the observable.

Scores are the normalized scores of each observable in PCA space.

LOOpredicts are the prediction values for leave one out cross validation. The corresponding regression coefficients are denoted *LOOCV Qsquared*).

For subspace selective ISPCA the following additional values are generated:
Newtrainingset is the culled catalyst training set.

Newobservable are the observables for the catalysts found in the desired PCA subspace.

Neworder are the indices for *Newobservable*.

The value of any observable (e.g. catalytic selectivity) can be predicted using the following equation, where *predict* is the vector of variables and *observable* is the accompanying observable: ((((predict./stdd)-meann)*loadings)*PCR)+mean(observable)

The end of the script can compute these values for the user. Note the variable vector must be constructed from the appropriate basis set.

Output examples (for illustrative purposes—these examples use a different training set than the one(s) contained in the manuscript)

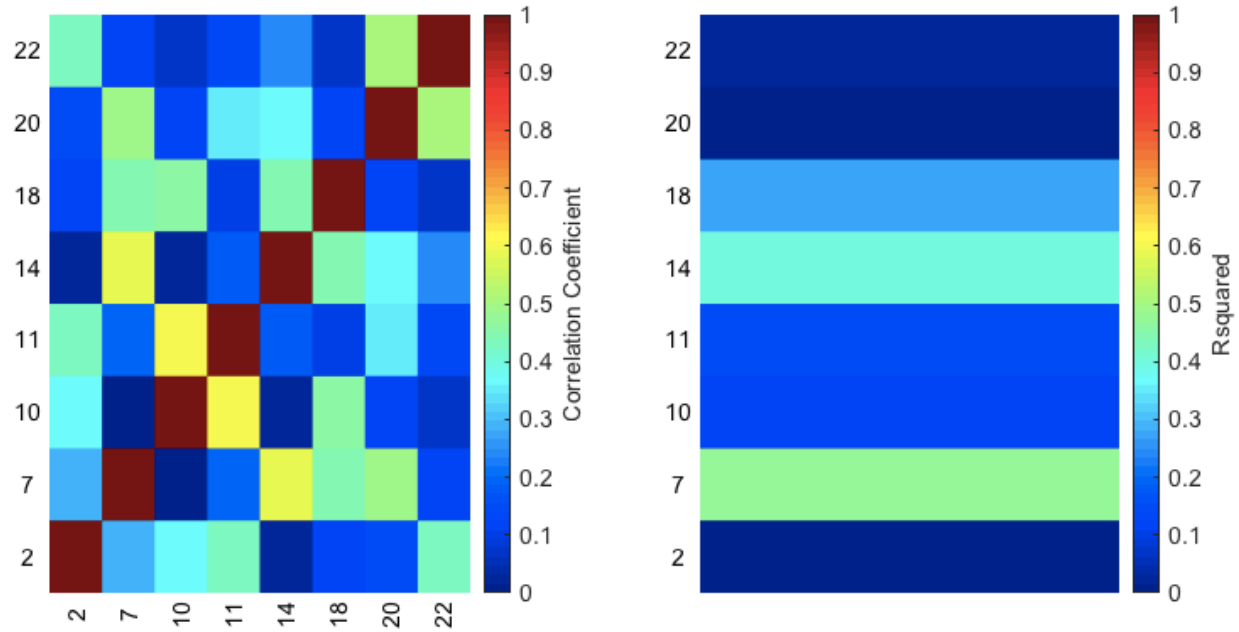


Figure S84. Pearson variable correlation matrix (left) and univariate analysis (right); from Figure 1 of *PCA4U2*.

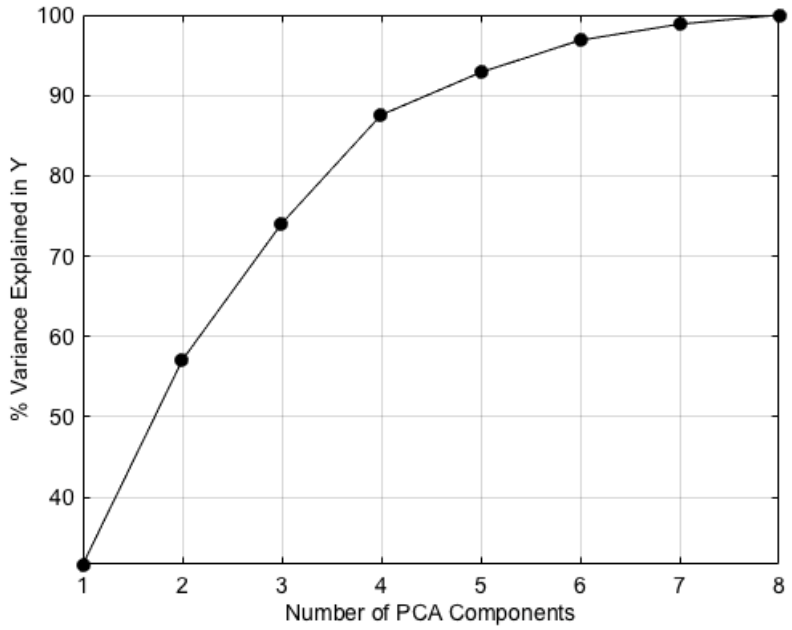


Figure S85. Variable Scree-type plot; from Figure 2 of *PCA4U2*.

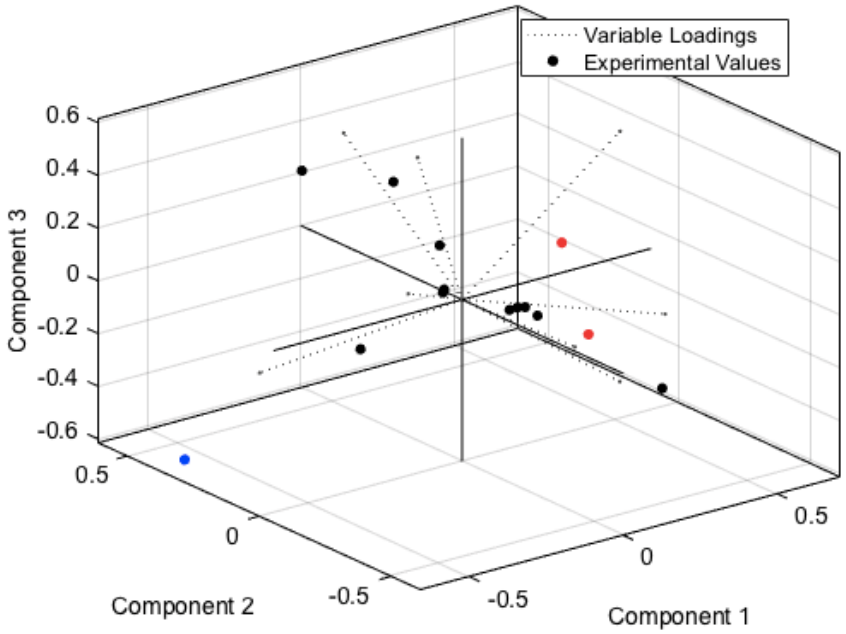


Figure S86. PCA-space biplot colored using dynamic k-means clustering; from Figure 3 of *PCA4U2*.

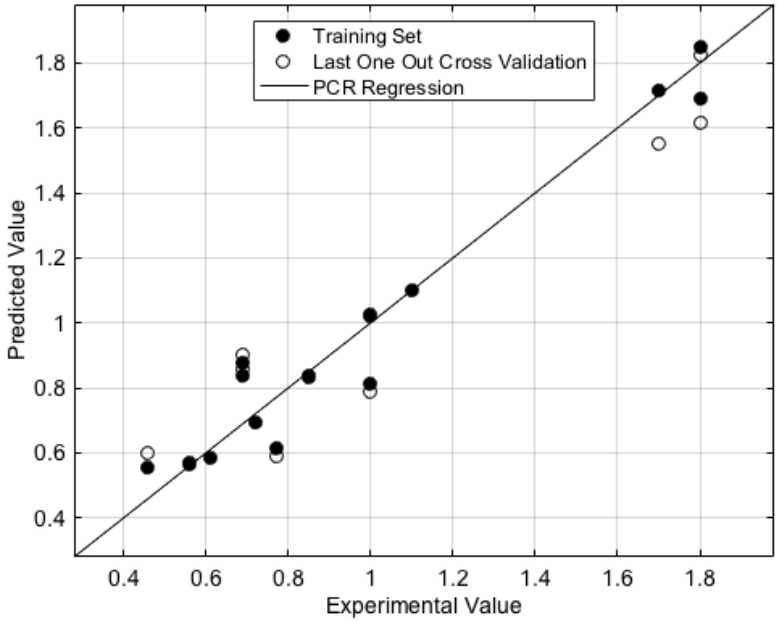


Figure S87. Regression fit using linear least squares fitting; from Figure 4 of *PCA4U2*.

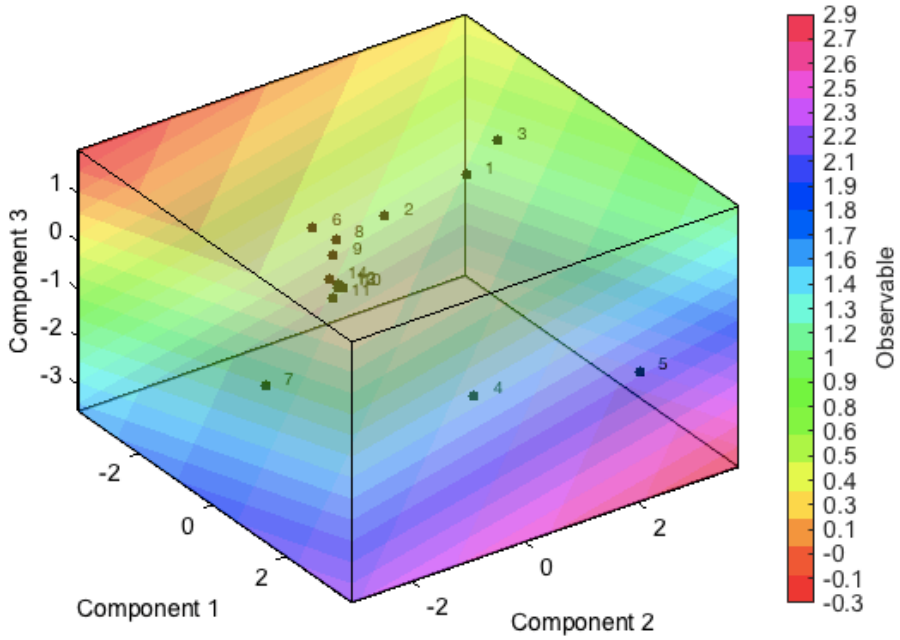


Figure S88. 3D PCA map colored according to linear regression between catalyst PCA scores and corresponding observable values; from Figure 5 of *PCA4U2*.

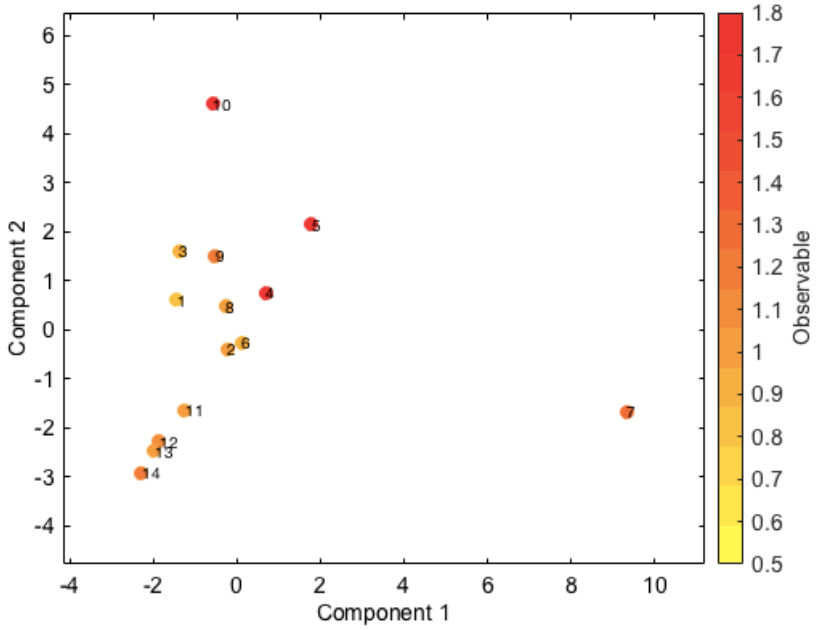


Figure S89. 2D PCA map projection showing principal component 1 *vs* 2, colored according to catalyst selectivity; from Figure 6 of *PCA4U2*.

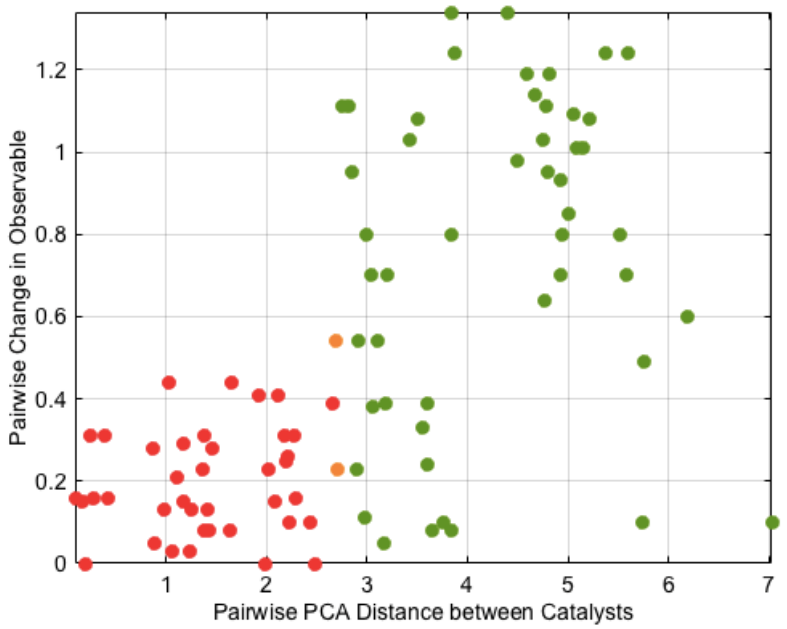


Figure S90. "Molecular ruler" plot constructed from pairwise Euclidean distances between catalyst PCA scores and catalyst selectivity; from Figure 7 of *PCA4U2*.

Additional PCA Modelling Data Details

Included below are supplementary principal-component based models and the corresponding fit statistics. Refer to PCAModellingData-SI.exe file for the complete data set.

PCA = Principle component analysis
ISPCA = Iterative supervised principle component analysis
PLSR = Partial least squares regression method reported by the Rothenberg lab[16]

PCA Model IA: Constructed using unoptimized PCA with catalysts **1 – 14** as the training set.

Table S3. Variable loadings in PC1 – PC3 and summed regression coefficients in model IA.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights(x 10^3)
1	Catalyst HOMO	-0.295	-0.178	0.053	-27.9
2	Catalyst LUMO	0.232	-0.266	0.033	-14.7
3	Free pyridine HOMO	0.323	-0.109	-0.041	-0.6
4	Free pyridine LUMO	-0.272	0.069	0.105	4.7
5	% Buried Volume	-0.222	-0.126	0.255	0.2
6	Free pyridine *p*-^{13}C NMR shift	0.305	-0.094	-0.058	-1.4
7	Free pyridine *o*-^{13}C NMR shift	0.198	0.087	0.360	52.4
8	Catalyst-bound pyridine *p*-^{13}C NMR shift	0.305	-0.106	-0.088	-5.5
9	Catalyst-bound pyridine *o*-^{13}C NMR shift	0.094	0.149	0.413	59.2
10	Composite donor atom polarizability	0.038	0.424	-0.174	31.0
11	Ti-X donor BDE	0.066	-0.231	0.246	1.4
12	X donor Mulliken EN	0.133	-0.002	0.205	25.2
13	Free pyridine quadrupole moment	-0.321	0.068	0.041	-3.7
14	Free pyridine N atom Mulliken charge	-0.012	-0.143	-0.407	-54.2
15	Catalyst-bound pyridine N atom Mulliken charge	-0.261	0.024	-0.202	-28.6
16	Catalyst X donor atom Mullken charge	0.162	0.366	-0.100	37.5
17	Free pyridine proton affinity	-0.317	0.089	-0.023	-7.3
18	Imido-Ti-N_{py} angle	0.010	-0.130	0.279	12.7
19	Ti-imido bond length	-0.127	-0.415	-0.137	-63.5
20	Ti-N_{py} bond length	0.018	0.171	0.315	48.8
21	Ti-X donor bond length	0.068	0.431	-0.121	38.1
22	X-Ti-X donor angle	-0.235	0.104	0.213	20.3
				R^2/Q^2	0.50/-0.16

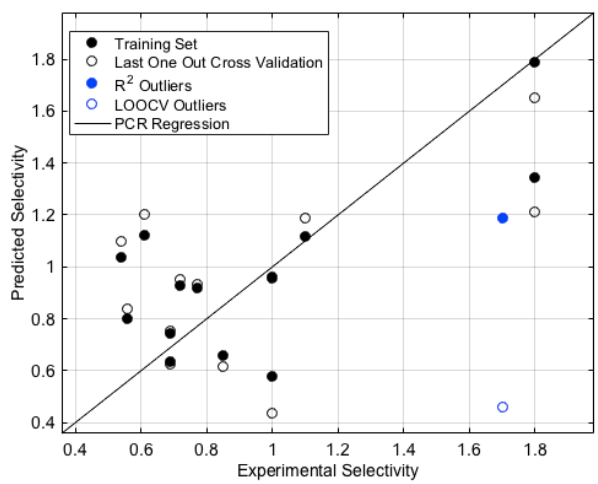


Figure S91. Predicted *vs* experimental selectivity determined by PCA model IA.

PCA Model IB: Constructed using ISPCA with catalysts **1** – **14** as the training set.

Table S4. Variable loadings in PC1 – PC3 and summed regression coefficients in model IB.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights(x 10^2)
2	Catalyst LUMO	0.348	-0.144	-0.218	-66.2
3	Free pyridine HOMO	0.419	0.072	-0.052	4.7
4	Free pyridine LUMO	-0.324	-0.139	0.230	58.6
5	% Buried Volume	-0.221	-0.349	0.126	17.7
7	Free pyridine *o*-^{13}C NMR shift	0.255	-0.068	0.514	187.1
10	Composite donor atom polarizability	-0.067	0.542	0.198	90.7
11	Ti-X donor BDE	0.144	-0.322	-0.308	-115.1
13	Free pyridine quadrupole moment	-0.394	-0.121	0.148	27.9
15	Catalyst-bound pyridine N atom Mulliken charge	-0.356	0.072	-0.380	-145.1
17	Free pyridine proton affinity	-0.414	-0.047	-0.031	-31.9
18	Imido-Ti-N_{py} angle	0.081	-0.359	0.523	168.8
21	Ti-X donor bond length	-0.028	0.530	0.204	93.8
				R^2/Q^2	0.95/0.93

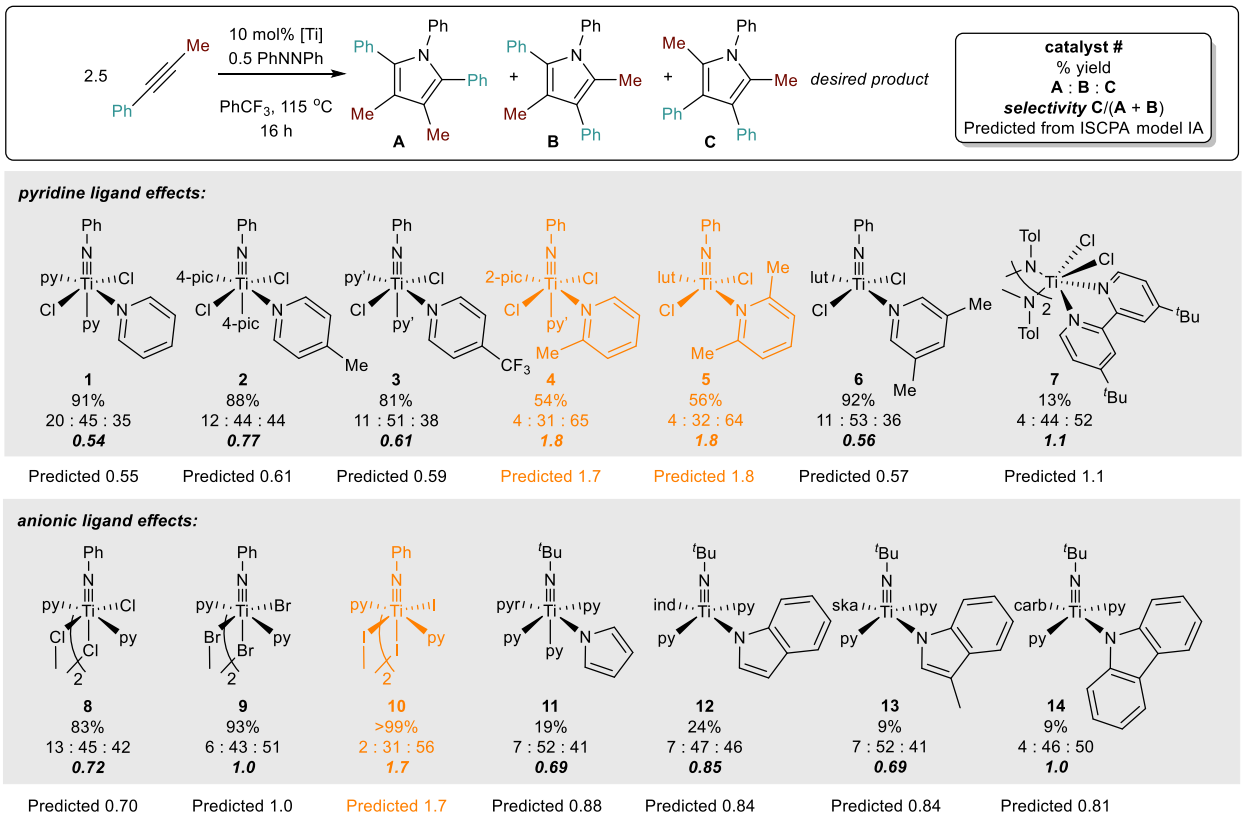


Figure S92. Initial catalyst screen for selective [2+2+1] pyrrole formation from PhCCMe with back-predicted selectivities from ISPCA Model IB. The most selective catalysts are drawn in orange. Conditions: 0.5 mmol PhCCMe, 0.1 mmol PhNNPh, 10 mol % [Ti], 0.5 mL $PhCF_3$, 115 °C, 16 h, average of 2 – 3 runs.

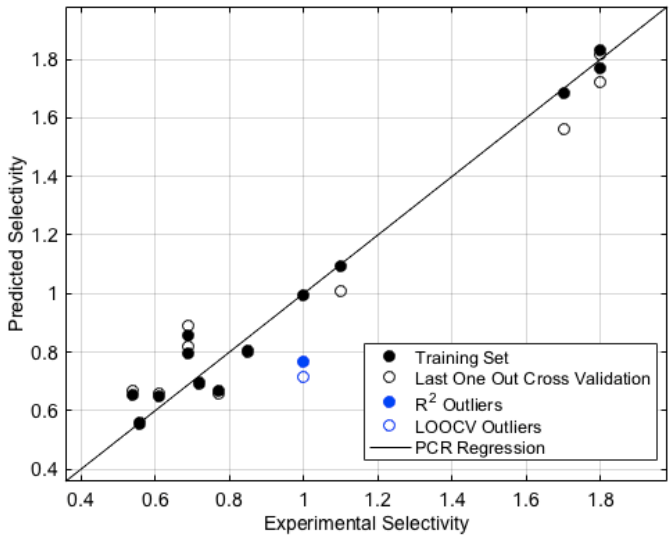


Figure S93. Predicted *vs* experimental selectivity determined by PCA model IB.

PLSR Model IC: Constructed using the PLSR model with catalysts **1** – **14** as the training set.

Table S5. Variable loadings in PC1 – PC3 in model IC.

Variable Index	Variables	Component 1	Component 2	Component 3
7	Free pyridine *o*-^{13}C NMR shift	-1.522	5.083	6.323
9	Catalyst-bound pyridine *o*-^{13}C NMR shift	-1.134	3.016	8.637
10	Composite donor atom polarizability	32.897	12.994	16.166
11	Ti-X donor BDE	-176.253	12.169	4.378
13	Free pyridine quadrupole moment	16.219	-79.770	16.224
18	Imido-Ti-N_{py} angle	-1.190	-0.700	3.298
			R^2/Q^2	0.72/0.22

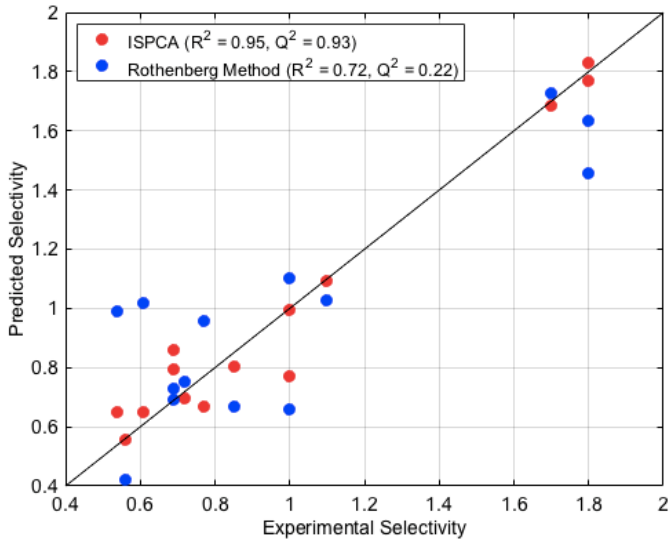


Figure S94. Comparison of predicted *vs* experimental selectivity plots calculated by ISPCA (red) and PLSR model (blue).

PCA Model IIA: Constructed using ISPCA with catalysts **1 – 22** as the training set.

Table S6. Variable loadings in PC1 – PC3 and summed regression coefficients in model IIA.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights
6	Free pyridine *p*-^{13}C NMR shift	0.599	0.416	-0.684	4.4
8	Catalyst-bound pyridine *p*-^{13}C NMR shift	0.609	0.319	0.727	-5.0
15	Catalyst-bound pyridine N atom Mulliken charge	-0.520	0.852	0.062	-2.0
				R^2/Q^2	0.72/0.13

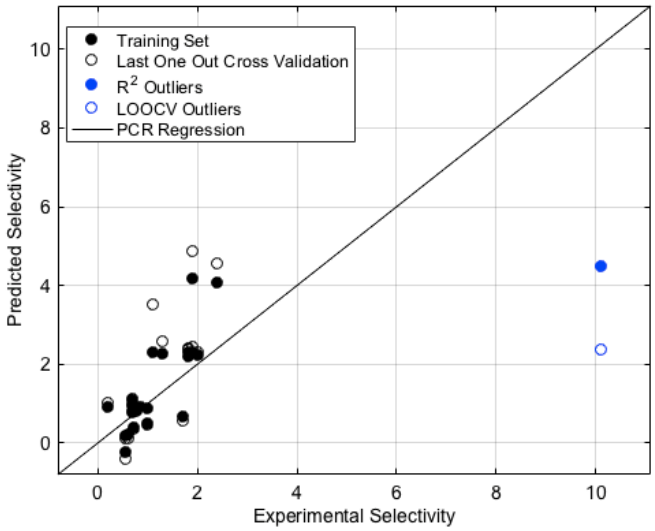


Figure S95. Predicted *vs* experimental selectivity determined by PCA model IIA.

PCA Model IIB: Constructed using ISPCA with catalysts **1 – 22**, sans **18**, as the training set.

Table S7. Variable loadings in PC1 – PC3 and summed regression coefficients in model IIB.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights(x 10^2)
1	Catalyst HOMO	-0.434	-0.210	0.386	-50.8
4	Free pyridine LUMO	-0.126	-0.046	0.731	112.1
7	Free pyridine *o*-^{13}C NMR shift	0.473	-0.349	0.036	140.0
9	Catalyst-bound pyridine *o*-^{13}C NMR shift	0.415	-0.394	0.146	145.0
10	Composite donor atom polarizability	0.193	0.518	0.314	128.2
11	Ti-X donor BDE	-0.286	-0.504	-0.295	-151.3
18	Imido-Ti-N_{py} angle	0.400	-0.274	0.263	166.5
22	X-Ti-X donor angle	-0.347	-0.281	0.196	-65.3
				R^2/Q^2	0.90/0.84

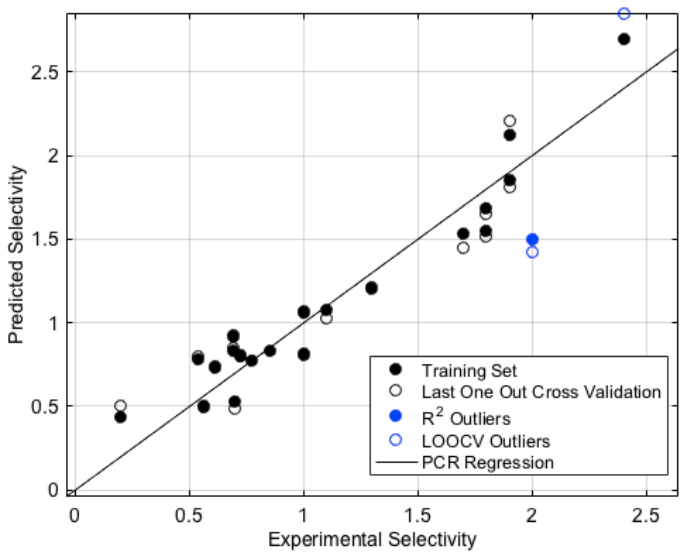


Figure S96. Predicted *vs* experimental selectivity determined by PCA model IIB.

PCA Model IIC: Constructed using ISPCA with catalysts **4**, **5**, **15**, **16**, **18 – 20** as the training set (determined by a dynamic k-means clustering search algorithm, *vide supra*).

Table S8: Variable loadings in PC1 – PC3 and summed regression coefficients in model IIC.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights
6	Free pyridine *p*-^{13}C NMR shift	0.566	-0.273	0.778	10.3
8	Catalyst-bound pyridine *p*-^{13}C NMR shift	0.581	-0.086	-0.430	-5.4
18	Imido-Ti-N_{py} angle	0.253	0.955	0.152	3.9
22	X-Ti-X donor angle	-0.527	0.072	0.433	5.4
				R^2/Q^2	0.99/-0.07

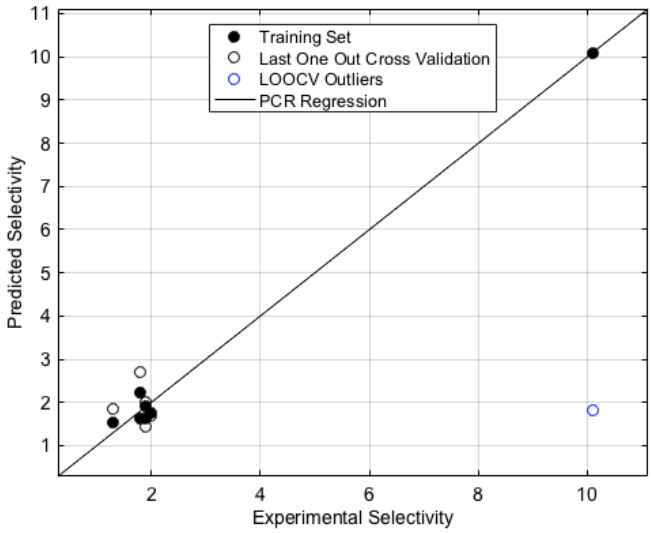


Figure S97. Predicted *vs* experimental selectivity determined by PCA model IIC.

PCA Model IID: Constructed using ISPCA with exponential weighting on catalysts **1** – **22** as the training set (weights were set according to pairwise distances to catalyst 18 in unoptimized PCA space, *vide supra*).

Table S9: Variable loadings in PC1 – PC3 and summed regression coefficients in model IID.

Variable Index	Variables	Component 1	Component 2	Component 3	Weights
3	Free pyridine HOMO	-0.407	0.877	-0.112	13.1
7	Free pyridine *o*-^{13}C NMR shift	0.409	0.170	-0.118	6.1
8	Catalyst-bound pyridine *p*-^{13}C NMR shift	0.408	0.318	0.840	-25.1
9	Catalyst-bound pyridine *o*-^{13}C NMR shift	0.409	0.155	-0.262	10.9
18	Imido-Ti-N_{py} angle	0.408	0.275	-0.420	17.6
20	Ti-N_{py} bond length	0.409	-0.043	-0.152	5.0
				R^2/Q^2	0.95/0.77

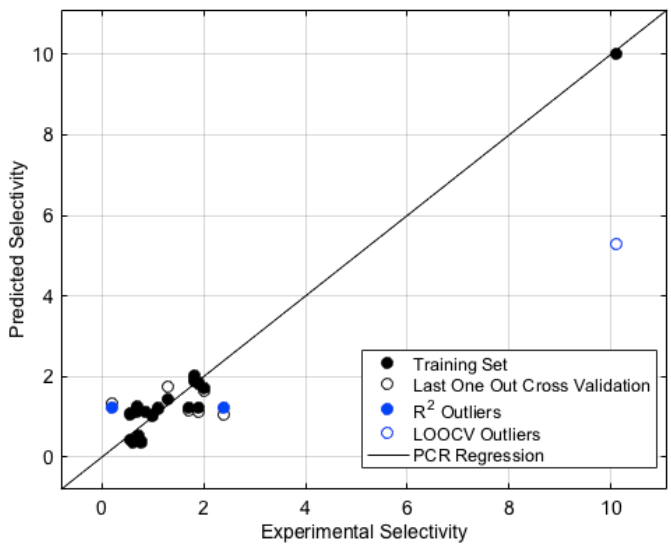


Figure S98. Predicted *vs* experimental selectivity determined by PCA model IID.

Molecular Ruler

A "molecular ruler" was used to cull potential test catalysts *in silico*. For example, catalysts bearing *o*-TMS groups, quinolines, or pendant alkoxides (**Figure S99**) were considered for steric and/or electronic modifications. These catalysts were found to be 1.2 to 4.1 units from the nearest training set neighbors in optimized PCA space. Given that the "molecular ruler" suggests a value of at least 1.7 units is necessary to expect a 0.46 increase in selectivity (derived from mapping the standard deviation of the selectivity of the training set comprised of catalysts **1** – **14** onto the Euclidean distance in the accompanying 3-component PCA space), in combination with the predicted selectivity of ~2.0 for the set, these catalysts did not warrant synthesis. It is important to note that the molecular ruler is a threshold value. Catalysts that are located below this threshold are unlikely to see a substantial increase in selectivity, however catalysts that are outside this threshold may or may not exhibit an increase in performance.

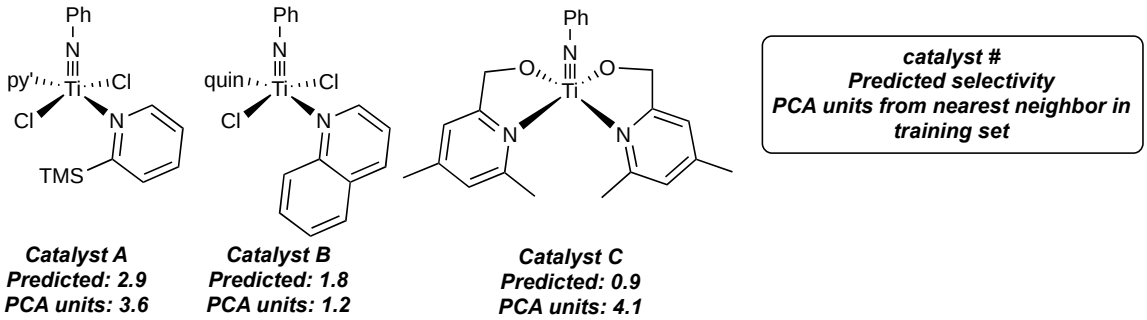


Figure S99. Example test catalysts excluded from synthesis via prediction from ISPCA model I.

Table S10. Pairwise Euclidean distances from catalyst 1 to all catalysts using different PCA models.

	Model Ia	Model Ib	Model IIa	Model IIb	Model IIc	Model IId
Catalyst ID	Pairwise Distance	Pairwise Distance	Pairwise Distance	Pairwise Distance	Pairwise Distance	Pairwise Distance
1	0.00	0.00	0.00	0.00	0.00	0.00
2	1.73	0.63	0.26	0.37	1.60	0.26
3	1.05	0.99	0.43	1.69	0.35	0.43
4	2.93	3.93	0.94	2.88	2.37	0.94
5	5.85	4.34	0.84	3.63	1.15	0.84
6	2.92	1.35	1.74	1.82	1.39	1.74
7	11.20	8.43	0.37	6.30	5.80	0.37
8	1.99	0.51	1.02	1.46	1.29	1.02
9	2.66	1.63	0.16	2.01	1.35	0.16
10	5.34	4.60	0.71	3.64	1.17	0.71
11	2.35	2.37	0.28	0.99	1.48	0.28
12	2.96	2.34	0.26	0.70	1.30	0.26
13	3.17	2.38	0.25	0.62	1.34	0.25
14	3.64	2.41	0.26	0.66	1.28	0.26
15	2.46	2.47	1.38	3.34	1.56	1.38
16	7.39	7.70	5.35	5.02	4.62	5.35
17	7.63	8.83	0.71	5.81	4.52	0.71
18	9.06	9.31	9.10	4.95	4.98	9.10
19	5.91	4.66	5.53	1.46	4.06	5.53
20	6.79	6.10	0.71	3.87	4.04	0.71
21	3.53	3.93	0.70	2.03	1.16	0.70
22	6.11	4.54	0.70	1.61	2.64	0.70
23	9.31	9.61	9.06	4.95	4.76	9.06
Catalyst A	5.41	7.83	9.36	6.85	4.62	9.36
Catalyst B	4.23	4.41	8.95	2.78	3.05	8.95
Catalyst C	8.89	6.00	9.17	4.54	1.64	9.17

* Boxes in grey refer to catalysts that were not included in the initial PCA training set.

Computational work

General considerations

Geometry optimizations were performed using Gaussian 16 program version c01.[17] All geometry optimizations and frequency calculations were performed using the M06 functional,[18] def2-SVP basis set,[19] and the SMD solvation model[20] with the experimentally used solvent $PhCF_3$ (ε = 9.18) as the solvent. All geometries were characterized by frequency analysis calculations to be local minima (without any imaginary frequency) or transition states (with only one imaginary frequency). All vibrational frequencies below 50 cm^{-1} were replaced with values of 50 cm^{-1} due to the breakdown of the harmonic oscillator model for low frequency vibrational modes. Zero-point vibrational energies and thermal contributions to electronic energy were calculated at 388.15 K and 1 atm. All computational geometries are given in XYZ format as supplementary materials (SelectivityGeometries.xyz, PCAGeometries.xyz). The information of each geometry is given at the line after atom numbers in the following format:
catalyst-configuration-intermediate: free energy before/after frequency correction.

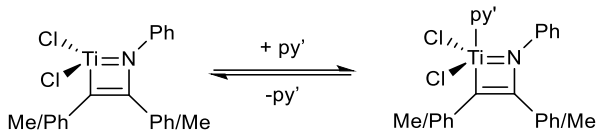
Limited by the size of the systems (78 atoms), a relatively small def2-SVP basis set for all geometry optimization and frequency calculations was used. In Table S11, $\Delta G^{\ddagger}$(TS2) was calculated by summing up the electronic energy at various DFT functionals with def2-TZVPP basis set (otherwise labelled) and thermal free energy correction at M06/def2-SVP level. Although the $\Delta G^{\ddagger}$(TS2) is vastly different depending on the functionals, the $\Delta\Delta G^{\ddagger}$($TS2_{BE}$ - $TS2_{BF}$) is very similar for all functionals. Therefore, the $\Delta\Delta G$(IM3) and $\Delta\Delta G^{\ddagger}$(TS2) values calculated by M06/def2-SVP are reliable.

Table S11. Gibbs free energies calculated by various DFT functionals with def2-TZVPP basis set.

$E_{high}+\Delta G_{low}^{corr}$	**$TS2_{AC}$**	**$TS2_{AD}$**	**$TS2_{BE}$**	**$TS2_{BF}$**
BP86	35.1	34.8	34.0	30.7
M06/def2SVP	44.2	43.4	43.3	39.8
MN15L	44.8	44.0	43.9	40.1
M06-D3	47.8	46.9	46.9	43.3
PW6B95D3/def2TZVP	47.5	47.0	47.3	43.9
B3LYP	50.3	49.4	49.7	46.2
M06L-D3	51.3	40.8	49.4	47.0
M06-L	55.4	55.0	53.6	51.3
M06	56.5	55.8	55.9	52.4

Computed reaction pathways for catalysts studied

Dissociation energy of pyridine at intermediate energy IM3 is defined as



$\Delta G_{diss} = \Delta G(IM3, L_n = Cl_2) + \Delta G(py') - \Delta G(IM3, L_n = Cl_2py')$

Six-coordinate TS2 has three possible configurations that were considered in the following reaction pathways. These configurations differ depending on the relative position of the pyridine and attacking alkyne.

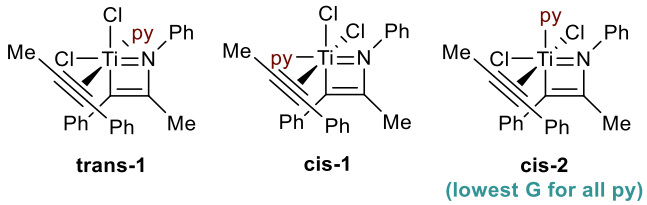


Figure S100. Possible isomers of TS2 ($TS2_{BF}$ shown) highlighting different positions of the pyridine ligand—either *trans* to the incoming alkyne (*trans-1*) or *cis* to the incoming alkyne (cis-1 and cis-2). In all cases the cis-2 isomer is lowest in free energy.

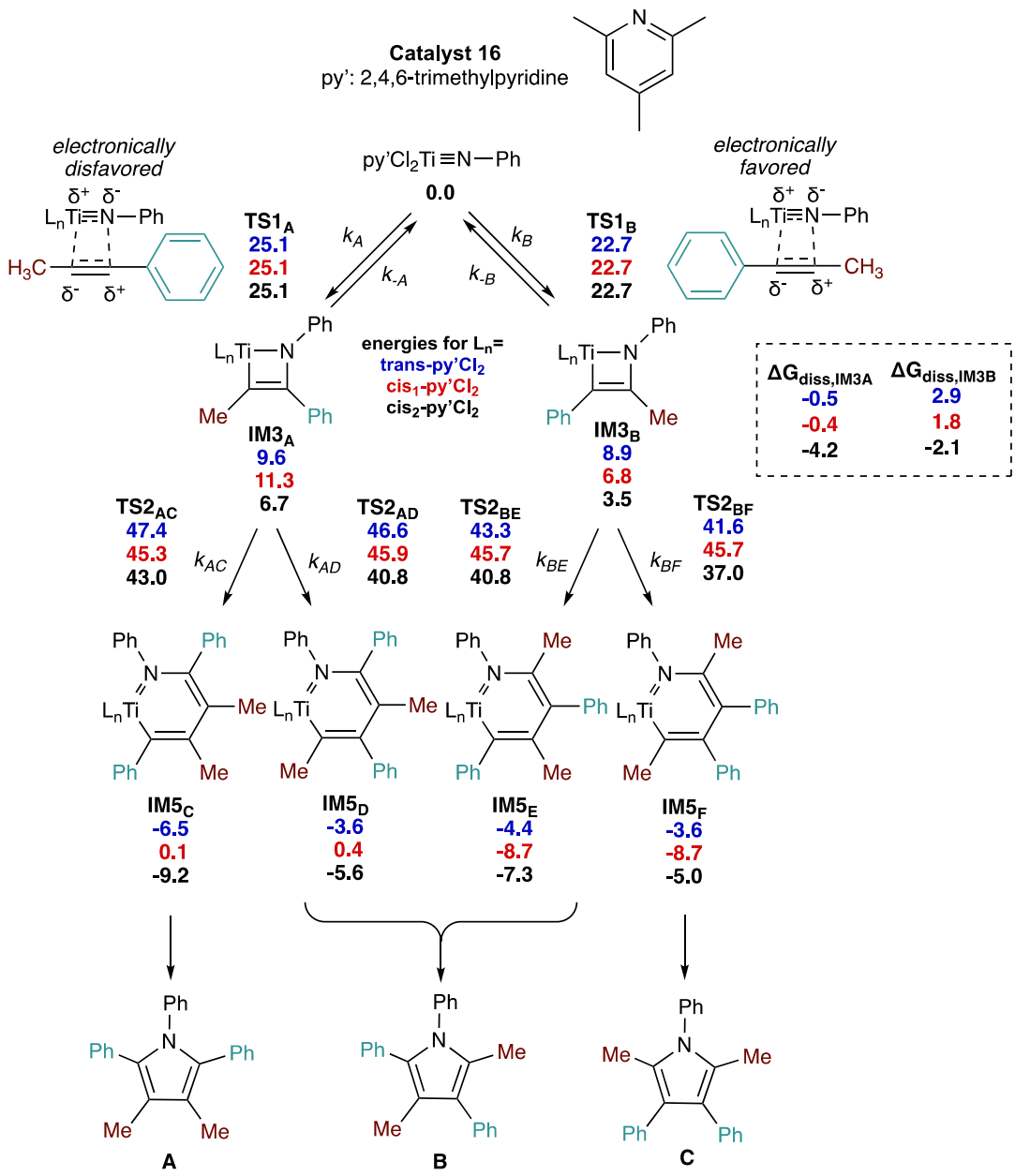


Figure S101. Calculated transition state barriers (TS1 and TS2) and intermediate free energies (IM3 and IM5) in kcal.mol^{-1} for catalyst **16** at trans- (blue), cis-1(red) and cis-2 (black) configurations. Dissociation energies of pyridine at IM3 are given in the dashed box.

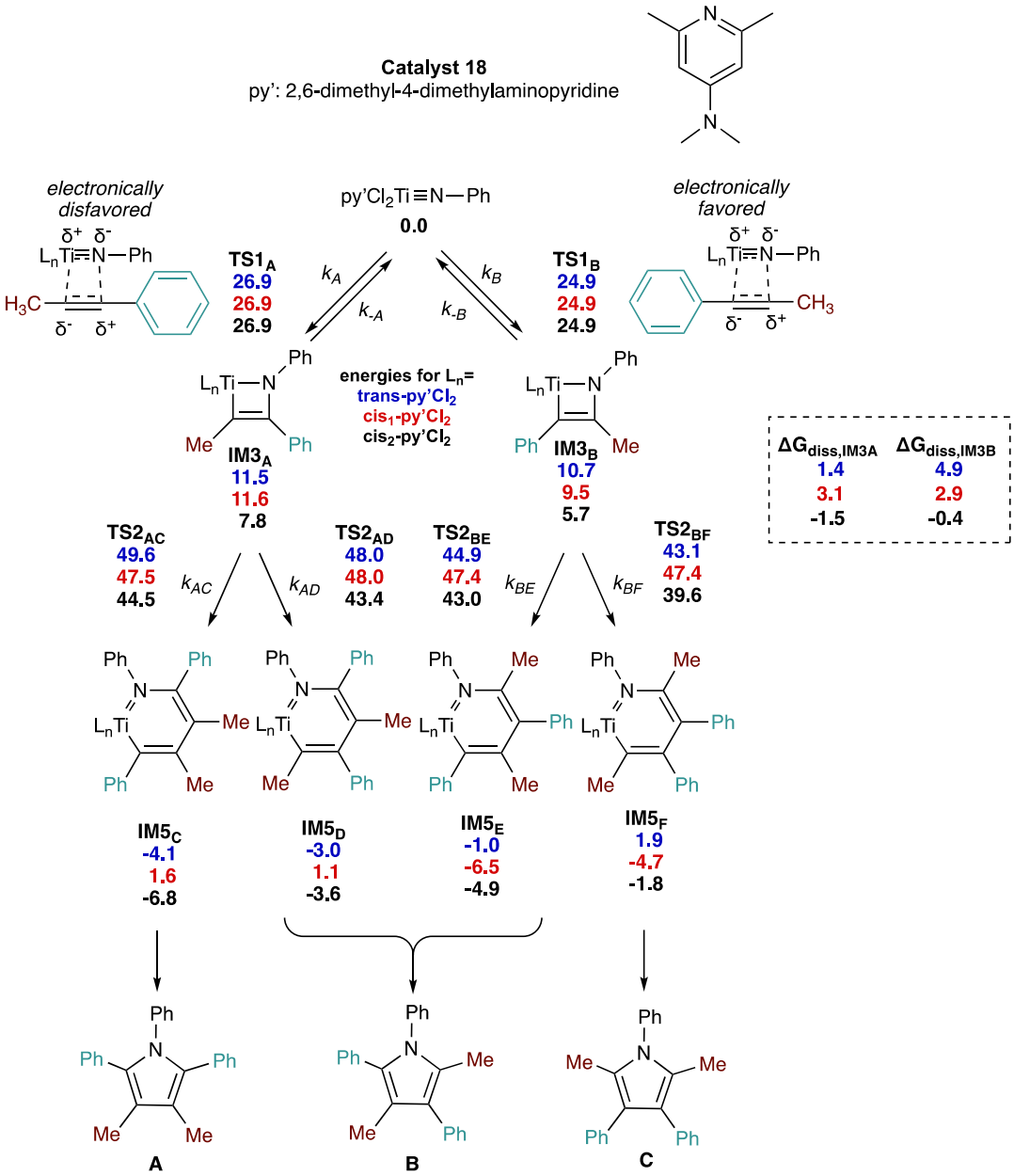


Figure S102. Calculated transition state barriers (TS1 and TS2) and intermediate free energies (IM3 and IM5) in kcal.mol^{-1} for catalyst **18** at trans- (blue), cis-1(red) and cis-2 (black) configurations. Dissociation energies of pyridine at IM3 are given in the dashed box.

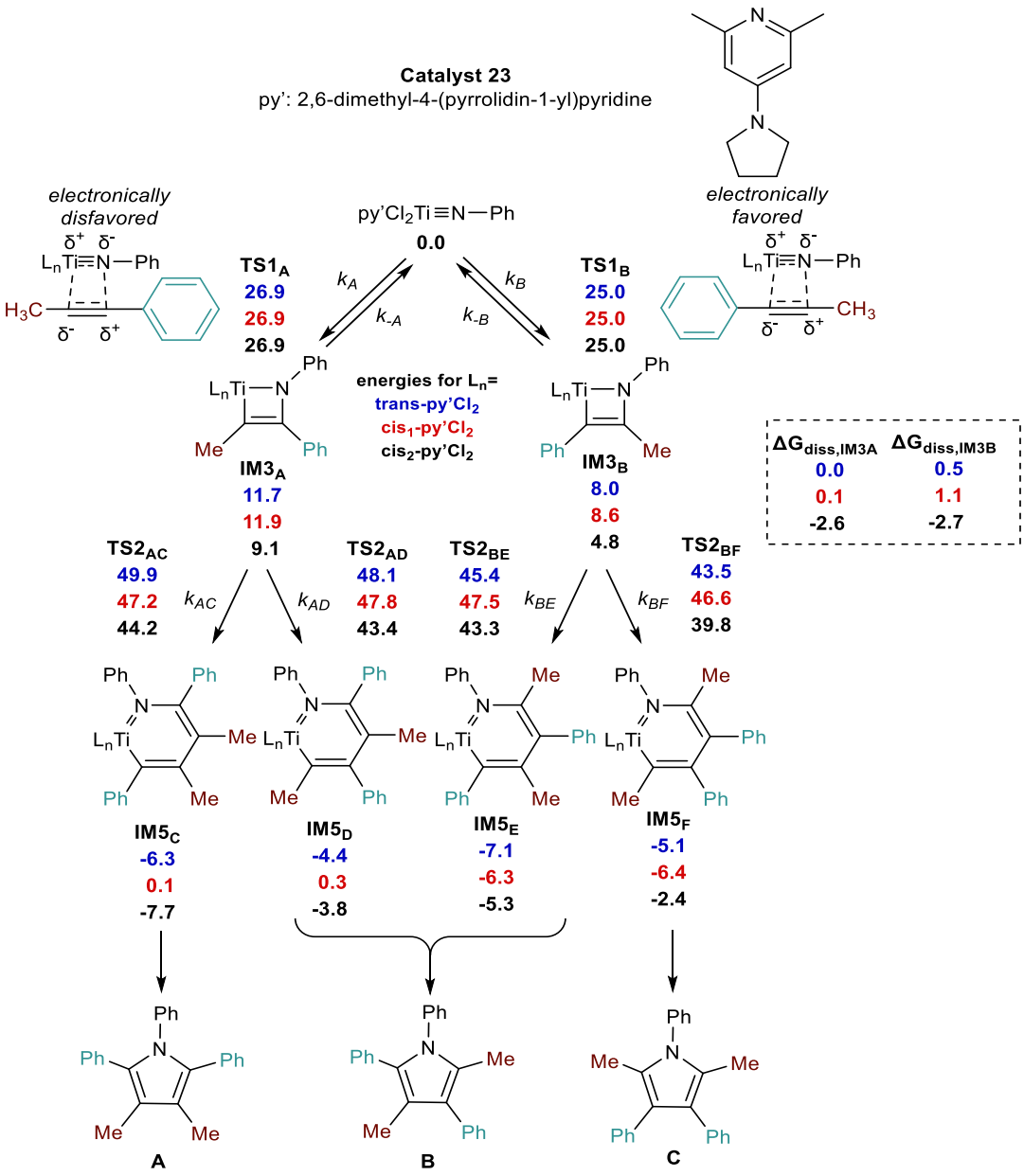


Figure S103. Calculated transition state barriers (TS1 and TS2) and intermediate free energies (IM3 and IM5) in kcal.mol^{-1} for catalyst **23** at trans- (blue), cis-1(red) and cis-2 (black) configurations. Dissociation energies of pyridine at IM3 are given in the dashed box.

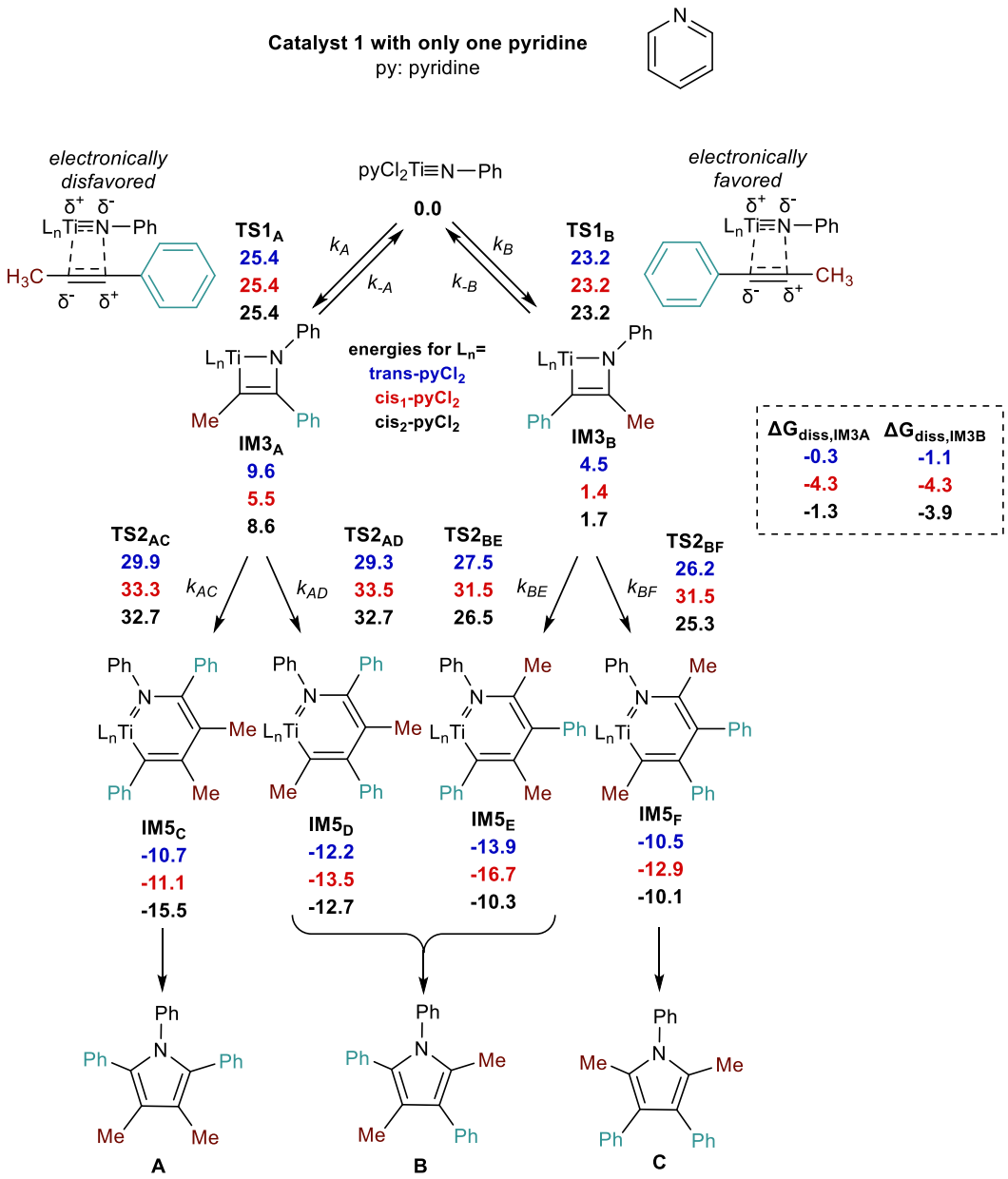


Figure S104. Calculated transition state barriers (TS1 and TS2) and intermediate free energies (IM3 and IM5) in kcal.mol^{-1} for catalyst **1** at trans- (blue), cis-1(red) and cis-2 (black) configurations. Dissociation energies of pyridine at IM3 are given in the dashed box.

Figure S104 differs from the computed reaction manifold pathway presented in our previous mechanistic work[10] in three aspects.

1. Our previous work was calculated using M06/6-311G(d,p)/SMD at 110 °C, while **Figure S104** was calculated using M06/def2-SVP/SMD at 115 °C .
2. Our previous work was calculated considering only the *trans*- configuration, while **Figure S104** considers *trans*-, *cis*-1 and *cis*-2 configurations.

3. Free energies in our previous work used the $py_3TiCl_2(NPh)$ catalyst as a starting point, while **Figure S104** used the $pyTiCl_2(NPh)$ catalyst in order to get directly comparable results to **Figures S101 – S103**.

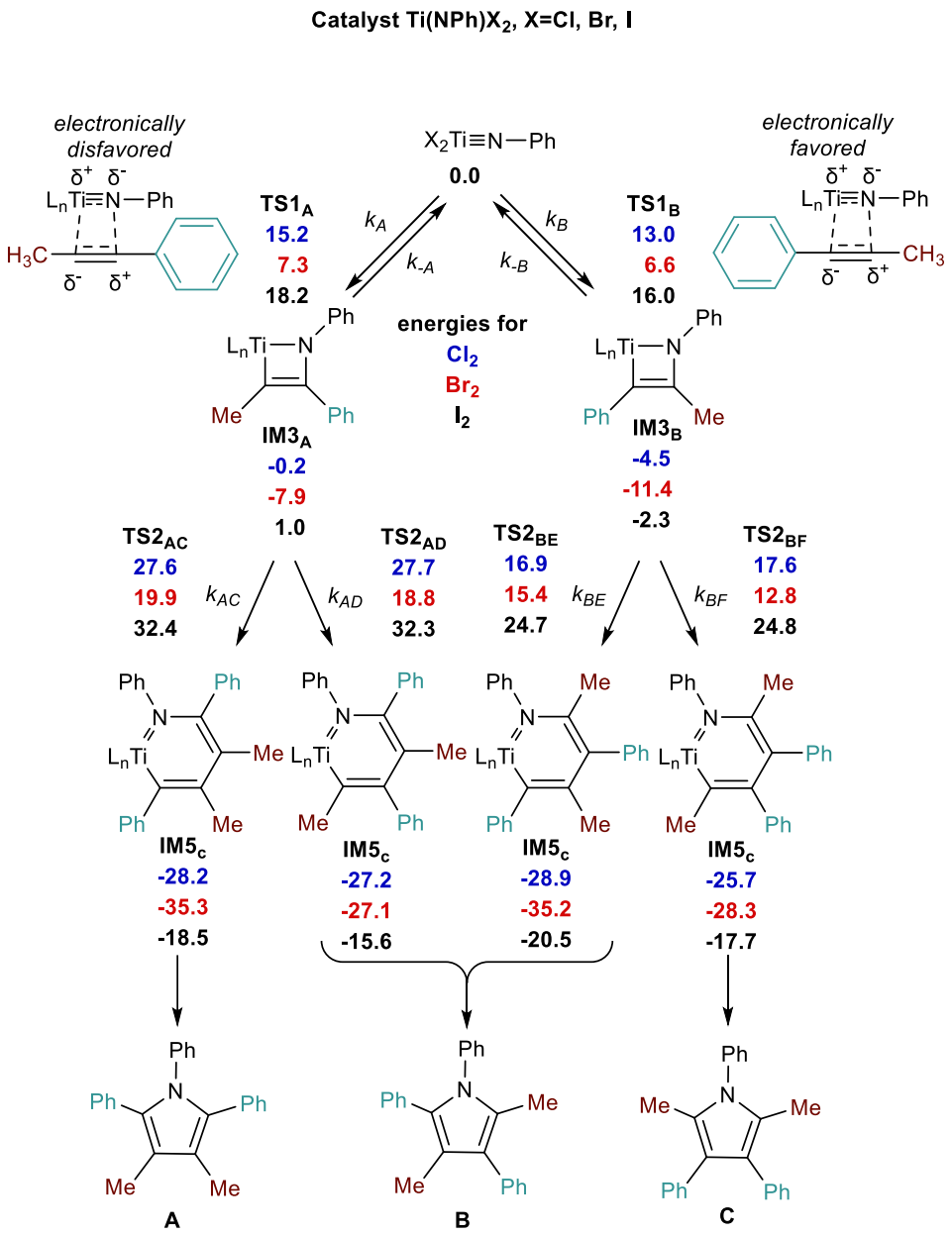


Figure S105. Calculated transition state barriers (TS1 and TS2) and intermediate free energies (IM3 and IM5) in kcal.mol^{-1} for $Ti(NPh)X_2$ (X = Cl, Br, I) at trans- (blue), cis-1(red) and cis-2 (black) configurations.

Computational Regioselectivity

The following procedure was used to obtain the computational product ratio. Since the [2+2] cycloaddition is reversible and not the rate-determining step, the concentration of [IM3$_A$] and [IM3$_B$] are thermodynamically controlled, and therefore follow the Boltzmann distribution:

$$\frac{[\mathrm{IM3_A}]}{[\mathrm{IM3_B}]} = \exp\left(\frac{-(\Delta \mathrm{G_{IM3_A}} - \Delta \mathrm{G_{IM3_B}})}{RT}\right) = \exp\left(\frac{-\Delta\Delta \mathrm{G(IM3_A - IM3_B)}}{RT}\right),$$

The following alkyne insertion step is irreversible as well as the rate-determining step, so the rate law is assumed to be:

$$\frac{d[\mathrm{TS2_{AC}}]}{dt} = k_{AC}[\mathrm{IM3_A}][\mathrm{MeCCPh}],$$

$$\frac{d[\mathrm{TS2_{AD}}]}{dt} = k_{AD}[\mathrm{IM3_A}][\mathrm{MeCCPh}],$$

$$\frac{d[\mathrm{TS2_{BE}}]}{dt} = k_{BE}[\mathrm{IM3_B}][\mathrm{MeCCPh}], \text{ and}$$

$$\frac{d[\mathrm{TS2_{BF}}]}{dt} = k_{BF}[\mathrm{IM3_B}][\mathrm{MeCCPh}].$$

where k is the reaction rate from the Eyring equation

$$k = \frac{k_B T}{h}\exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right).$$

therefore, the ratio of the products will be:

$$\frac{[\mathrm{TS2_{BE}}]}{[\mathrm{TS2_{BF}}]} = \frac{k_{BE}}{k_{BF}} = \exp\left(\frac{-(\Delta G^{\ddagger}_{\mathrm{TS2_{BE}}} - \Delta G^{\ddagger}_{\mathrm{TS2_{BF}}})}{RT}\right) = \exp\left(\frac{-\Delta\Delta G^{\ddagger}(\mathrm{TS2_{BE} - TS2_{BF}})}{RT}\right)$$

where R is the gas constant and T is the temperature. At experimental temperature 115 °C, RT=0.77 kcal.mol^{-1}. Therefore, a minor change in $\Delta\Delta G^{\ddagger}$ leads to a significant variation in product distribution.

Table S12 shows the computationally predicted selectivity. For catalyst **16** and **18**, we find that our computational prediction matches quantitatively with experimental data, predicting product C to be the dominant product.

Table S12. Predicted product ratio using M06/def2-SVP/SMD.

RT(388.15K): 0.77 kcal.mol^{-1}	$\Delta\Delta$G(IM3$_A$ − IM3$_B$)	$\Delta\Delta G^{\ddagger}$(TS2$_{AC}$ − TS2$_{AD}$)	$\Delta\Delta G^{\ddagger}$(TS2$_{BE}$ − TS2$_{BF}$)	Predicted ratio	Experimental ratio
Catalyst 1	5.77	0.00	1.23	0 : 17 : 83	20 : 45 : 35
Catalyst 16	3.16	2.20	3.78	0 : 2 : 98	3 : 32 : 65
Catalyst 18	2.11	1.12	3.48	1 : 6 : 93	1 : 8 : 91
Catalyst 23	4.30	0.77	3.41	0 : 1 : 99	1 : 7 : 92

References

1. Whitesides, G. M.; Ehmann, W. J., The Mechanism of Formation of 1,2,3,4-Tetramethylnaphthalene from 2-Butyne and Triphenyltris(tetrahydrofuran)chromium(III). *J. Am. Chem. Soc.* **1970,** *92*, 5625-5640.
2. Pennington, D. A.; Horton, P. N.; Hursthouse, M. B.; Bochmann, M.; Lancaster, S. J., Synthesis and catalytic activity of dinuclear imido titanium complexes: the molecular structure of $[Ti(NPh)Cl(\mu\text{-}Cl)(THF)_2]_2$. *Polyhedron* **2005,** *24*, 151-156.
3. See, X. Y.; Beaumier, E. P.; Davis-Gilbert, Z. W.; Dunn, P. L.; Larsen, J. A.; Pearce, A. J.; Wheeler, T. A.; Tonks, I. A., Generation of Ti^{II} Alkyne Trimerization Catalysts in the Absence of Strong Metal Reductants. *Organometallics* **2017,** *36*, 1383-1390.
4. Gilbert, Z. W.; Hue, R. J.; Tonks, I. A., Catalytic formal [2+2+1] synthesis of pyrroles from alkynes and diazenes via Ti^{II}/Ti^{IV} redox catalysis. *Nat. Chem.* **2016,** *8*, 63-68.
5. Hayes, P. G.; Piers, W. E.; Parvez, M., Arene Complexes of β-Diketiminato Supported Organoscandium Cations: Mechanism of Arene Exchange and Alkyne Insertion in Solvent Separated Ion Pairs. *Chem. Eur. J.* **2007,** *13*, 2632-2640.
6. Davis-Gilbert, Z. W.; Kawakita, K.; Blechschmidt, D. R.; Tsurugi, H.; Mashima, K.; Tonks, I. A., In Situ Catalyst Generation and Benchtop-Compatible Entry Points for Ti^{II}/Ti^{IV} Redox Catalytic Reactions. *Organometallics* **2018,** *37*, 4439-4445.
7. Obligacion, J. V.; Semproni, S. P.; Pappas, I.; Chirik, P. J., Cobalt-Catalyzed $C(sp^2)$-H Borylation: Mechanistic Insights Inspire Catalyst Design. *J. Am. Chem. Soc.* **2016,** *138*, 10645-10653.
8. Tonks, I. A.; Meier, J. C.; Bercaw, J. E., Alkyne Hydroamination and Trimerization with Titanium Bis(phenolate)pyridine Complexes: Evidence for Low-Valent Titanium Intermediates and Synthesis of an Ethylene Adduct of Titanium(II). *Organometallics* **2013,** *32*, 3451-3457.
9. Zhang, Z.; Leitch, D. C.; Lu, M.; Patrick, B. O.; Schafer, L. L., An Easy-To-Use, Regioselective, and Robust Bis(amidate) Titanium Hydroamination Precatalyst: Mechanistic and Synthetic Investigations toward the Preparation of Tetrahydroisoquinolines and Benzoquinolizine Alkaloids. *Chem. Eur. J.* **2007,** *13*, 2012-2022.
10. Davis-Gilbert, Z. W.; Wen, X.; Goodpaster, J. D.; Tonks, I. A., Mechanism of Ti-Catalyzed Oxidative Nitrene Transfer in [2+2+1] Pyrrole Synthesis from Alkynes and Azobenzene. *J. Am. Chem. Soc.* **2018,** *140*, 7267-7281.
11. Falivene, L.; Cao, Z.; Petta, A.; Serra, L.; Poater, A.; Oliva, R.; Scarano, V.; Cavallo, L., Towards the online computer-aided design of catalytic pockets. *Nat. Chem.* **2019,** *11*, 872-879.
12. Tantillo, D. J. Chemical Shift Repository for computed NMR scaling factors, with Coupling Constants Added Too. http://cheshirenmr.info/index.htm (accessed 2020).
13. Dean, J. A., Table 4.11 Bond Dissociation Energies. In *Lange's Handbook of Chemistry*, Fifteenth ed.; Esposito, R., Ed.; McGraw-Hill Education: 2005; Vol. 1, pp 4.41 - 4.53.
14. Schwerdtfeger, P.; Nagle, J. K., 2018 Table of static dipole polarizabilities of the neutral elements in the periodic table. *Mol. Phys.* **2019,** *117*, 1200-1225.
15. Bratsch, S. G., Revised Mulliken electronegativities: I. Calculation and conversion to Pauling units. *J. Chem. Educ.* **1988,** *65*, 34.
16. Landman, I. R.; Paulson, E. R.; Rheingold, A. L.; Grotjahn, D. B.; Rothenberg, G., Designing bifunctional alkene isomerization catalysts using predictive modelling. *Catal. Sci. Technol.* **2017,** *7*, 4842-4851.
17. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.;

Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

18. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008,** *120*, 215-241.

19. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005,** *7*, 3297-3305.

20. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009,** *113*, 6378-6396.