SUPPORTING INFORMATION

for

Machine Learning Directed Discovery and Optimisation of a Platinum-Catalysed Amide Reduction

Eleonora Casillo,[a] Benon P. Maliszewski,[a] César A. Urbina-Blanco,[b,c] Thomas Scattolin[d] Catherine S. J. Cazin*[a] and Steven P. Nolan*[a]

[a] *Department of Chemistry and Center for Sustainable Chemistry, Ghent University, Krijgslaan 281 (S-3), 9000 Ghent, Belgium. E-mail: steven.nolan@ugent.be; catherine.cazin@ugent.be.*
[b] *Laboratory for Chemical Technology (LCT), Department of Materials, Textiles and Chemical Engineering, University of Ghent, Ghent, 9052, Belgium.*
[c] *Prophecy Labs, Raymond Blyckaerts 13, Ixelles, 1050, Belgium*
[d] *Dipartimento di Scienze Chimiche, Università degli Studi di Padova, via Marzolo 1, 35131 Padova (Italy).*

Table of Contents

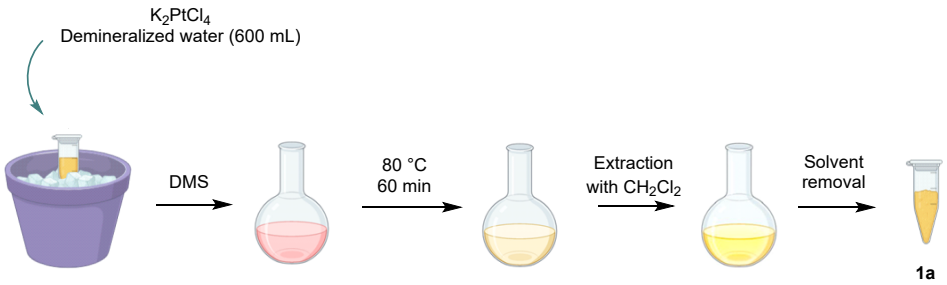
1. Synthetic procedures and characterisation data

1.1 General information

All chemical syntheses were performed in glass vials under air. Solvents and all other reagents were purchased and used as received without any additional purification. Platinum pre-catalysts were synthesised according to the literature procedures.[1,2] ^{1}H NMR spectra were recorded in $CDCl_3$ using a Magritek 80 MHz Spinsolve Phosphorus benchtop NMR spectrometer. All chemical shifts are quoted in parts per million referenced to the $CHCl_3$ solvent residue (δ_H = 7.26 ppm). Five different Pt-based catalysts were employed.

1.2 Synthesis of $[Pt(DMS)_2Cl_2]$ CAS 55449-91-7

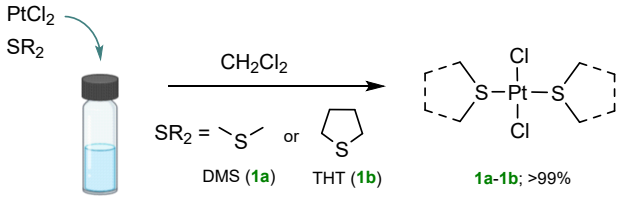
A 1000 mL round-bottom flask was charged with a solution of K_2PtCl_4 (15.500 g, 37.34 mmol, 1.00 equiv.) in demineralised water (600 mL) and placed in an ice bath. DMS (15.50 mL, 209.56 mmol, 5.61 equiv.) was added, while stirring vigorously, resulting in a rapid formation of a pale pink precipitate (Magnus-type complex of the formula $[Pt(DMS)_4][PtCl_4]$). The flask was then equipped with a reflux condenser and the resulted mixture was heated up to 80 °C and stirred at this temperature for 10 min, until the solution became clear yellow. After this time, the mixture was cooled down to the room temperature and extracted with CH_2Cl_2 (4x200 mL) until colourless. The combined organic fractions were dried over anhydrous $MgSO_4$ and evaporated to dryness, affording the desired complex **1a** as a yellow crystalline solid (14.570 g, >99%, Scheme S1).



Scheme S1: Synthesis of $Pt(DMS)_2Cl_2$.

1.3 Synthesis of $[Pt(THT)_2Cl_2]$ (1b) CAS 24940-43-0

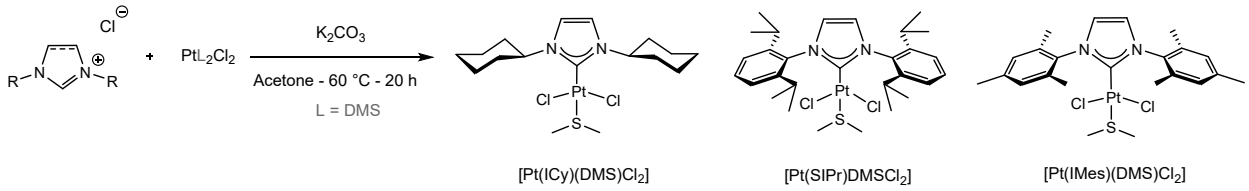
A 20.0 mL vial was charged with $PtCl_2$ (250.0 mg, 0.940 mmol, 1.0 equiv.) and CH_2Cl_2 (5.0 mL). THT (500 µL, 5.671 mmol, 6.0 equiv.) was added, the vial was closed with a septum screw cap and the resulted mixture was stirred at 20 °C for 60 min. After this time, a complete consumption of the dark $PtCl_2$ was observed, leading to the formation of a clear, orange solution. The volatiles were removed under vacuum, affording the desired complex **1b** as an orange powder (414.0 mg, >99%; Scheme S2). **^{1}H NMR** (300 MHz, $CDCl_3$): δ (ppm) = 3.67 (br, 2H, S(C**H**$_2$)), 2.83 (br, 2H, S(C**H**$_2$)), 2.24 (br, 2H, S(CH_2)(C**H**$_2$)), 2.00 (br, 2H, S(CH_2)(C**H**$_2$)). **^{13}C NMR** (75 MHz, $CDCl_3$): δ (ppm) = 38.87 (S(**C**H_2), minor), 37.40 (S(**C**H_2), major), 30.37 (S(CH_2)(**C**H_2), minor), 30.00 (S(CH_2)(**C**H_2), major).



*Scheme S2: Synthesis of Pt(II)-thioether complexes **1a**-**1b**.*

1.4 Synthesis of Pt(II)-NHC pre-catalysts

Three other catalysts based on carbene systems were also used in this project (Scheme S3). We try to explore the catalytic activity of a recently disclosed family of easily accessible well-defined Pt(II)-NHC pre-catalysts. The operationally simple protocol, involving a weak base-mediated reaction (a sustainable and facile synthetic route) between platinum precursors and azolium salts under mild and open-to-air conditions, allowed us to obtain a series of air- and moisture-stable complexes of general formula $[Pt(NHC)(L)Cl_2]$.[1]



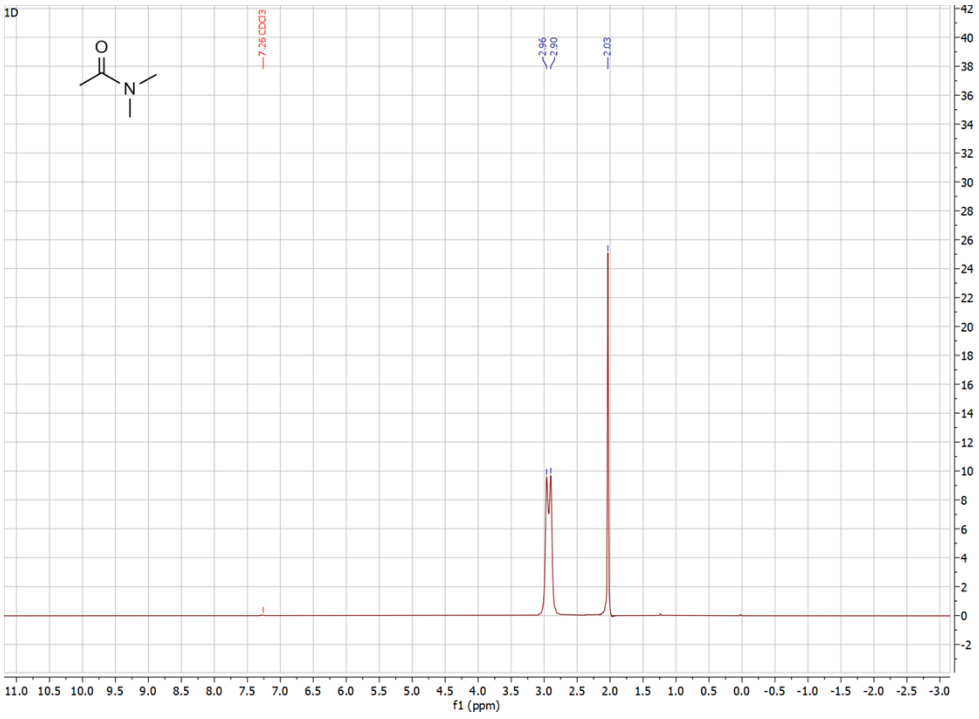
Scheme S3: Synthesis of $[Pt(NHC)(L)Cl_2]$-catalysts by weak base route.

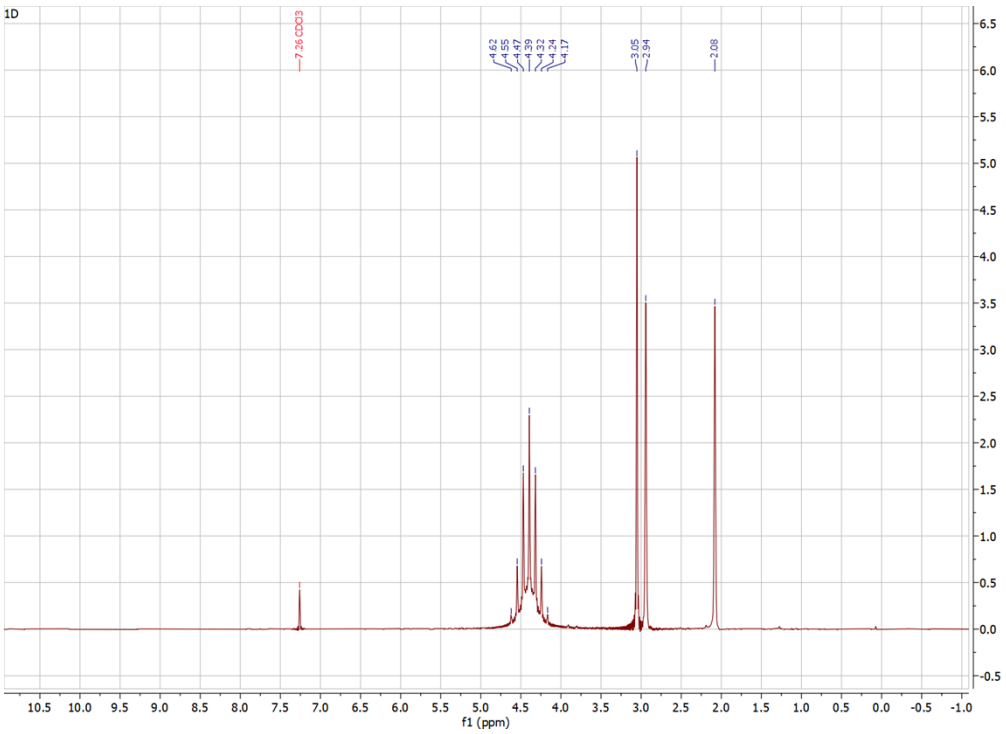
1.5 General reduction of amide procedure

A 4.0 mL vial was charged with a platinum catalyst, 1,1,3,3-tetramethyldisiloxane (265 µL, 1.5 mmol, 3 equiv.) and *N*,*N*-dimethylacetamide (46.5 µL, 0.5 mmol, 1.0 equiv.) were added. The vial was closed with a septum screw cap and the resulted mixture was stirred at the set temperature. The progress of the reaction was monitored by ^{1}H NMR spectroscopy.

2. NMR spectra

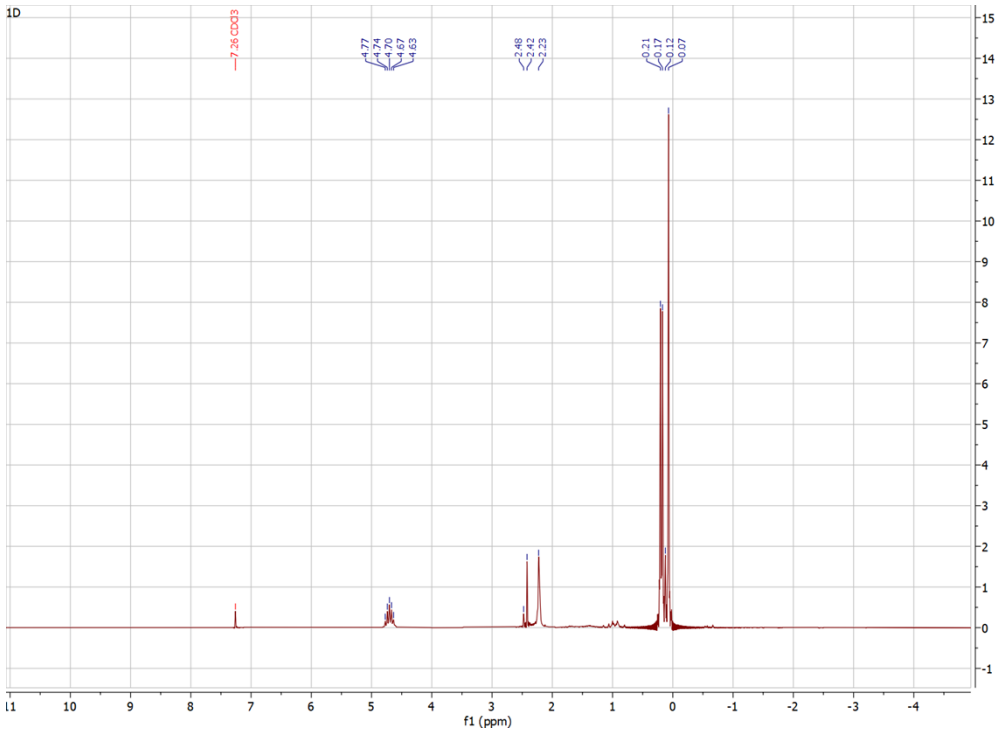
2.1. 1H-NMR of N,N-dimethylacetamide CAS 127-19-5 (substrate)



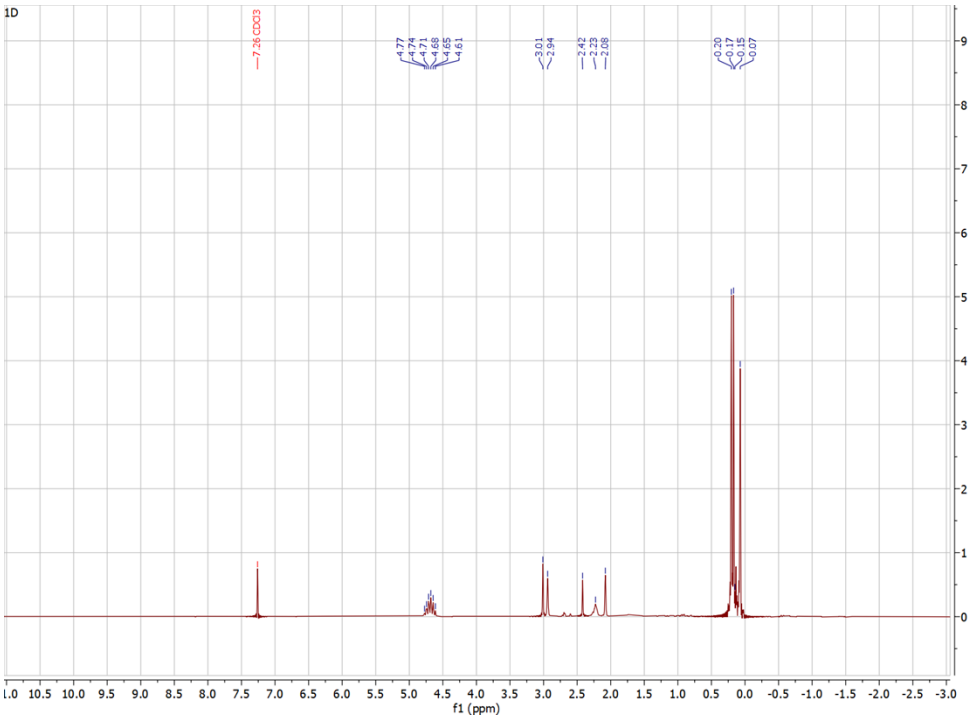


2.2. ^{1}H-NMR of Entry 1 Table 1

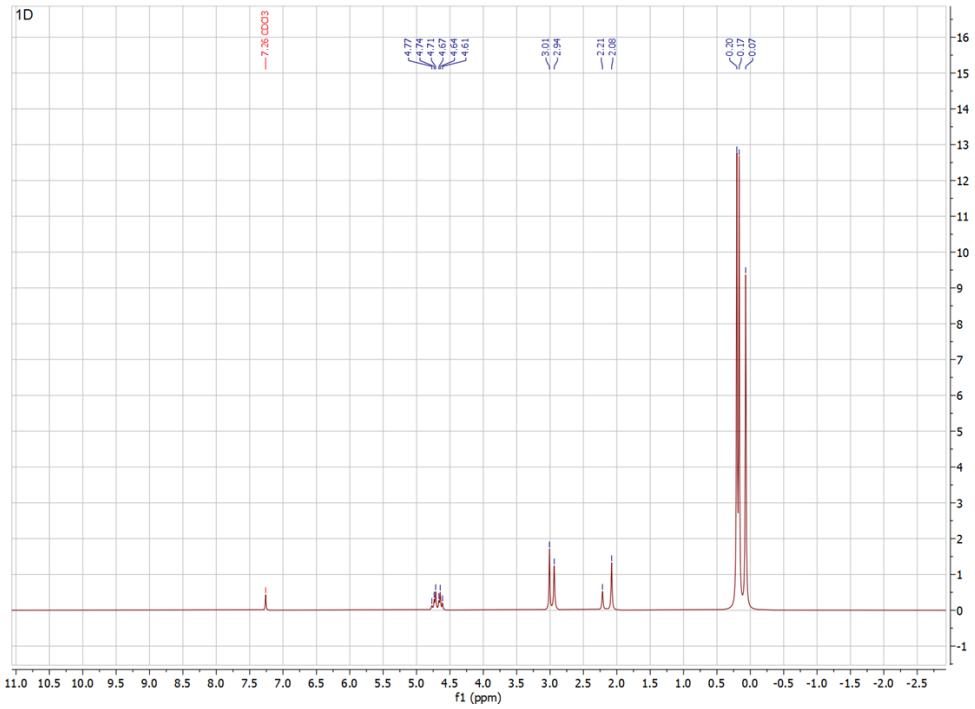
2.3. ^{1}H-NMR of Entry 2 Table 1



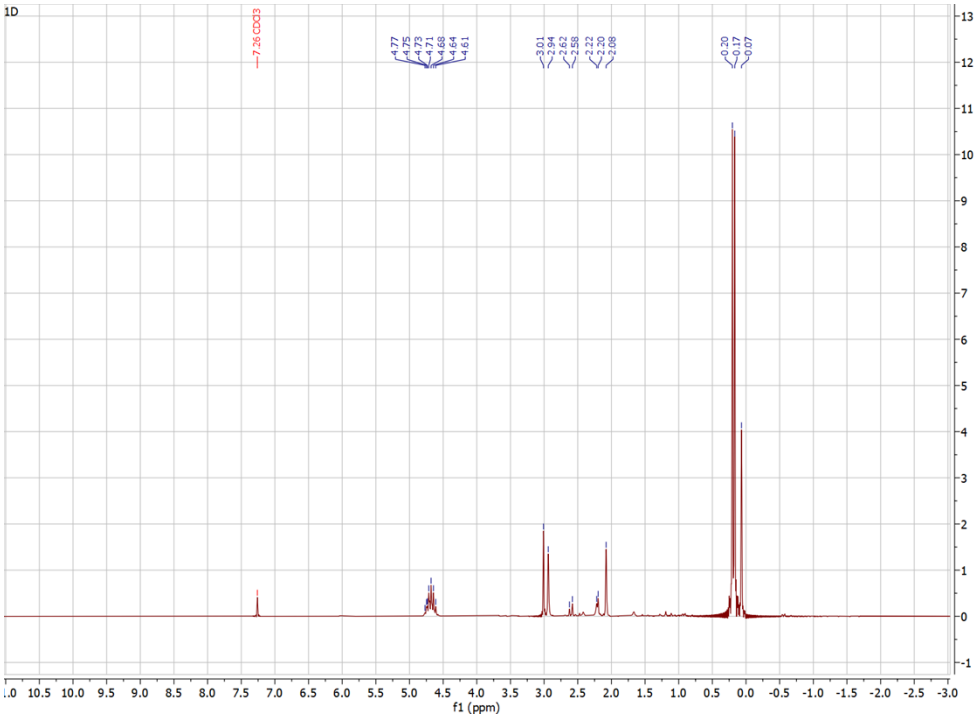
2.4. ^{1}H-NMR of Entry 3 Table 1



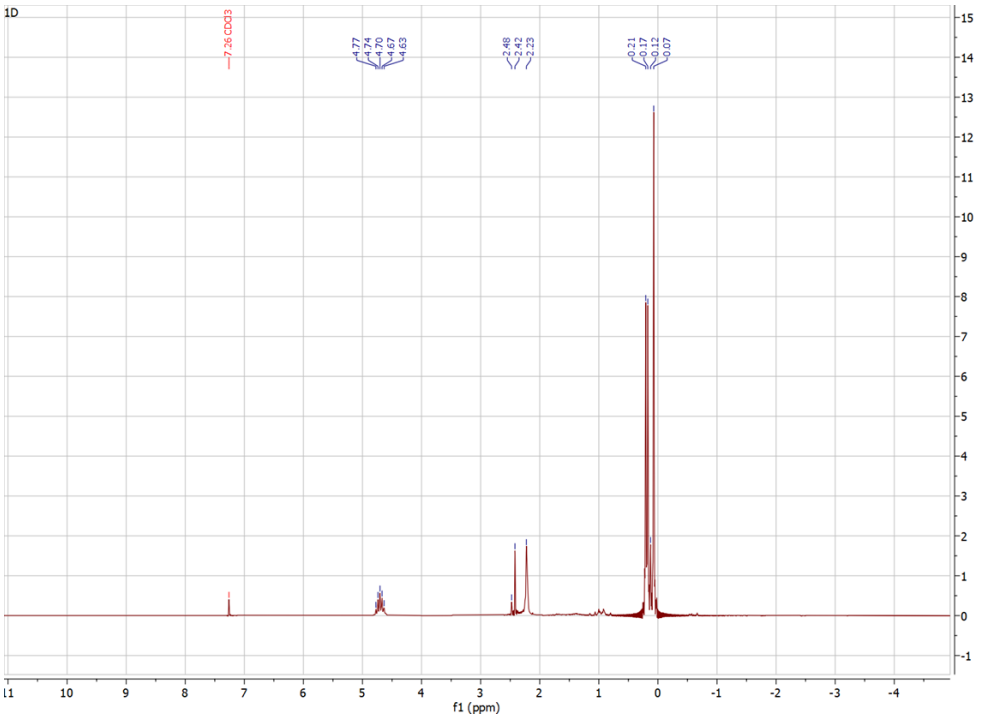
2.5. ^{1}H-NMR of Entry 4 Table 1



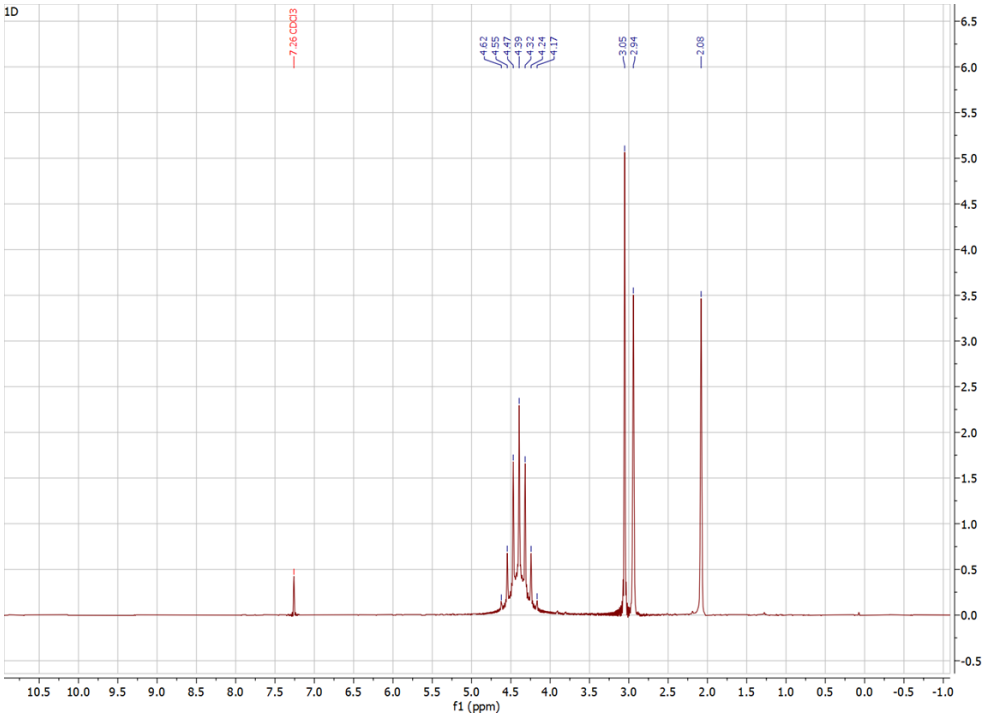
2.6. ^{1}H-NMR of Entry 5 Table 1



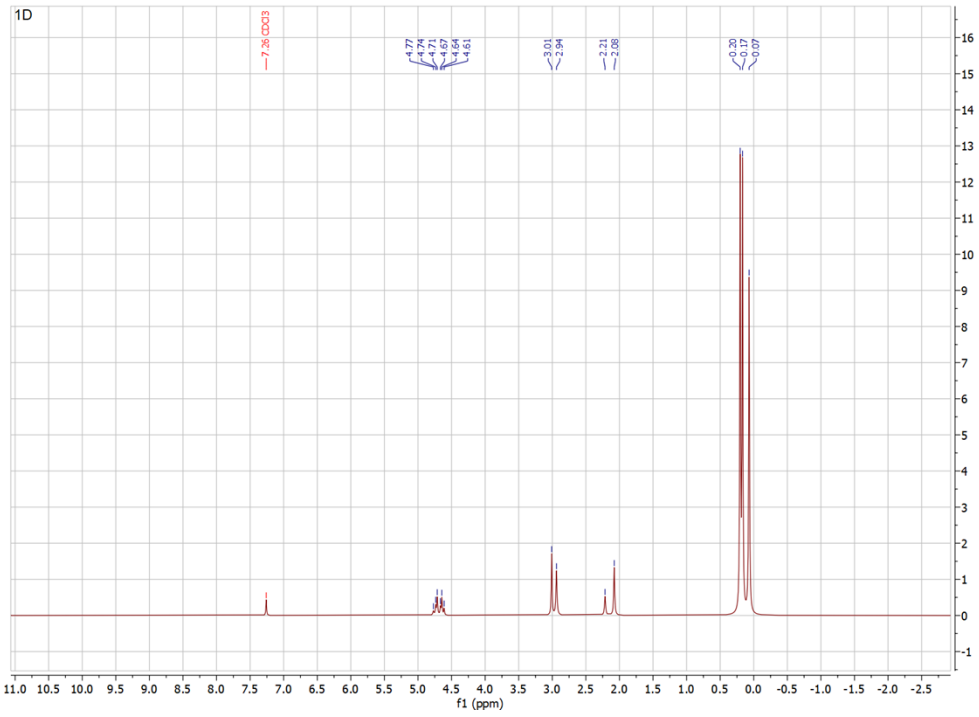
2.7. ^{1}H-NMR of Entry 6 Table 1



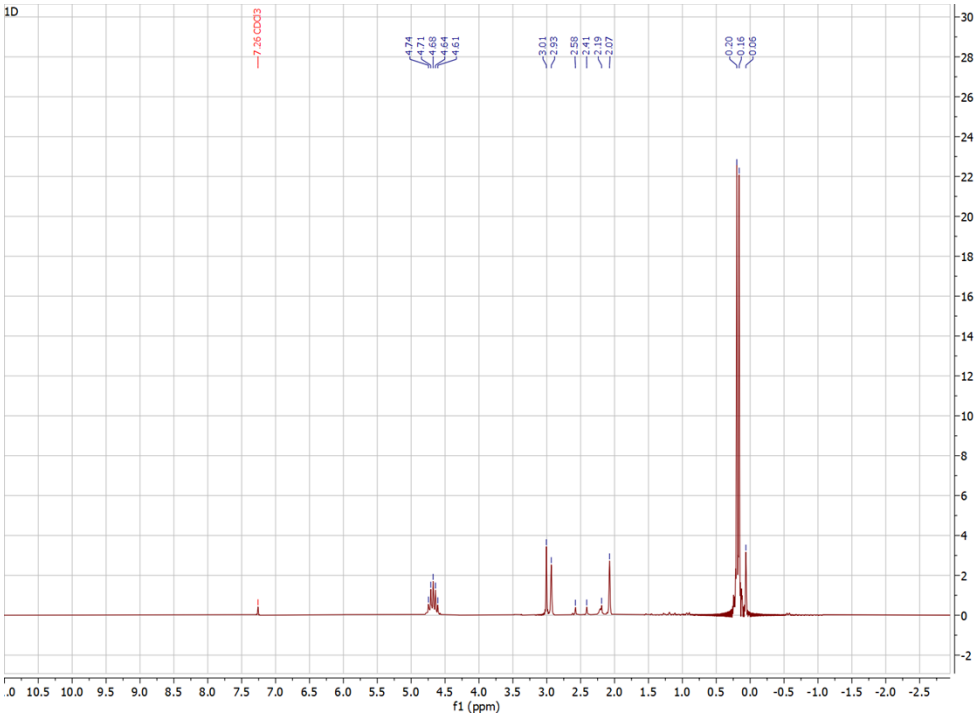
2.8. ^{1}H-NMR of Entry 7 Table 1



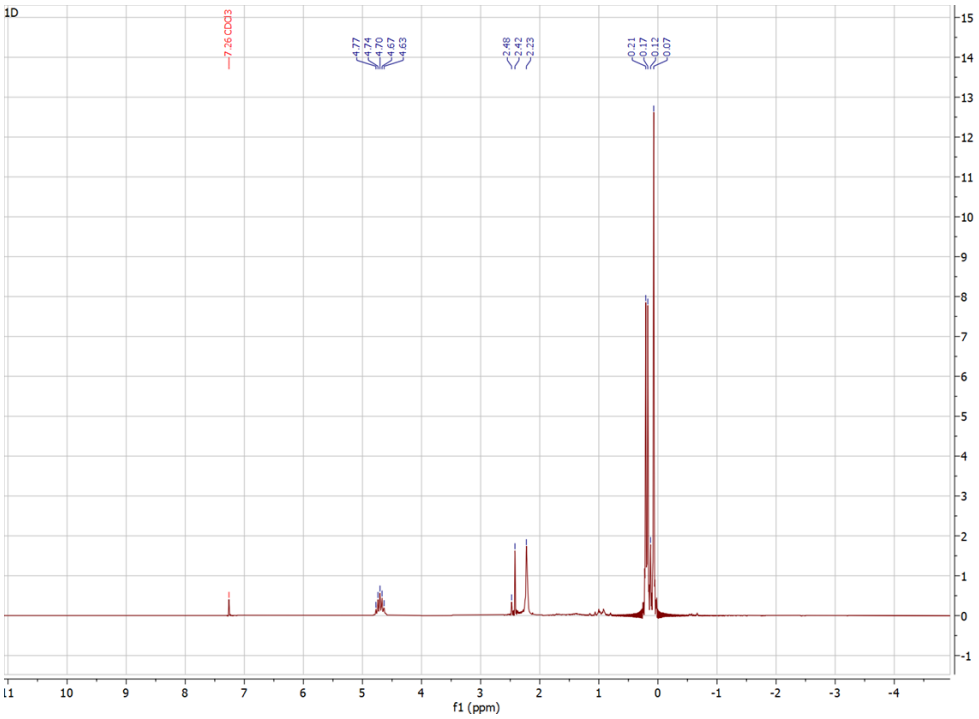
2.9. ^{1}H-NMR of Entry 8 Table 1



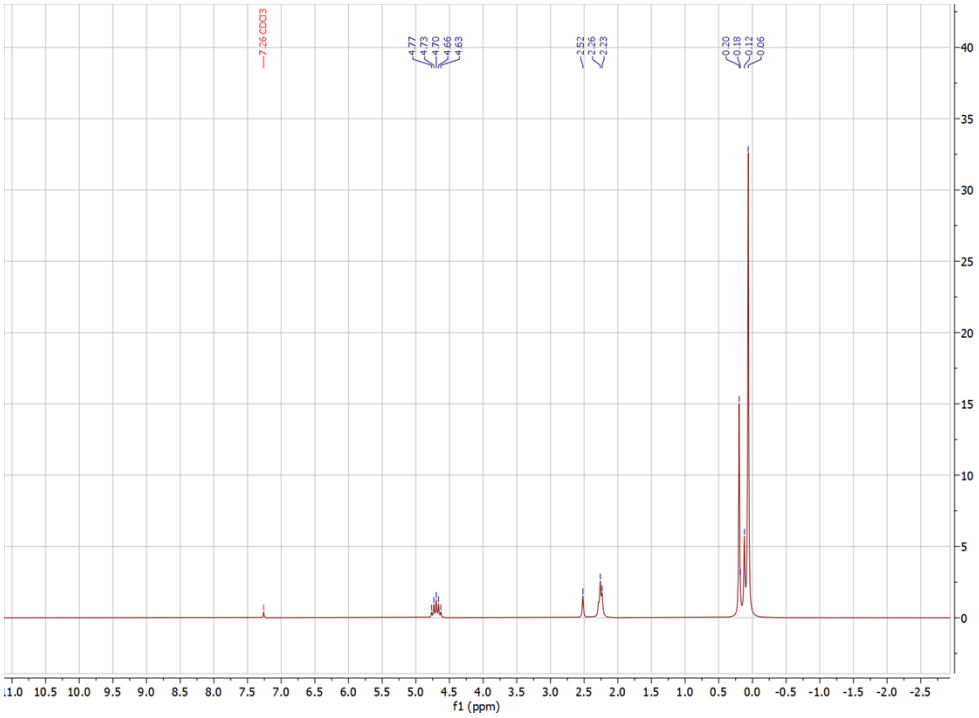
2.10. ^{1}H-NMR of Entry 9 Table 1



2.11. ^{1}H-NMR of Entry 10 Table 1

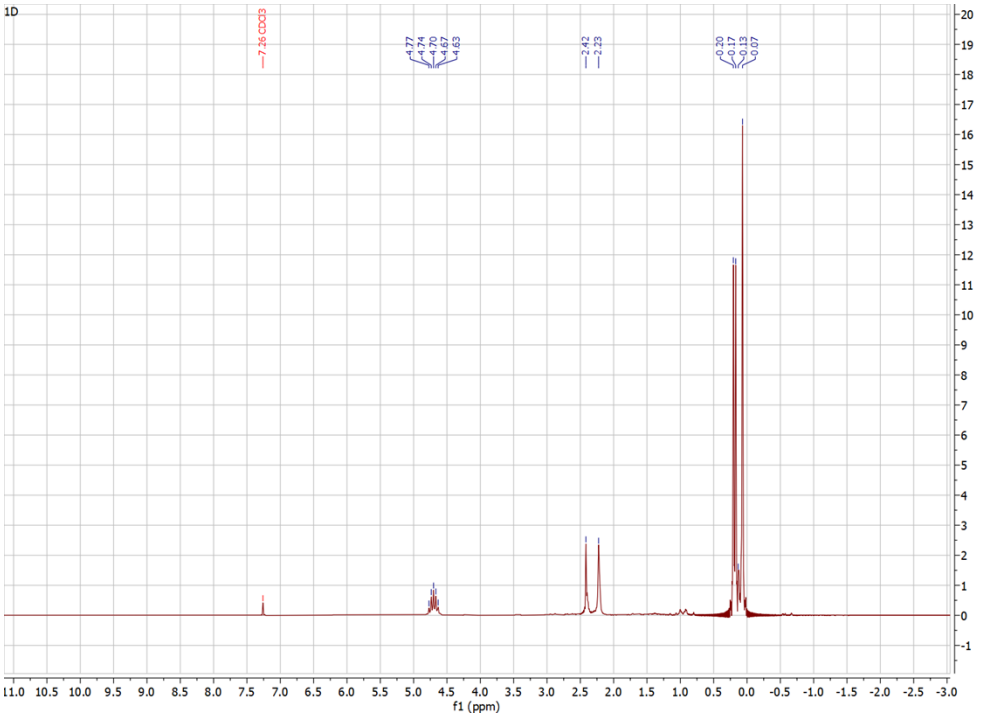


2.12. ^{1}H-NMR of Entry 1 Table 2 [a]



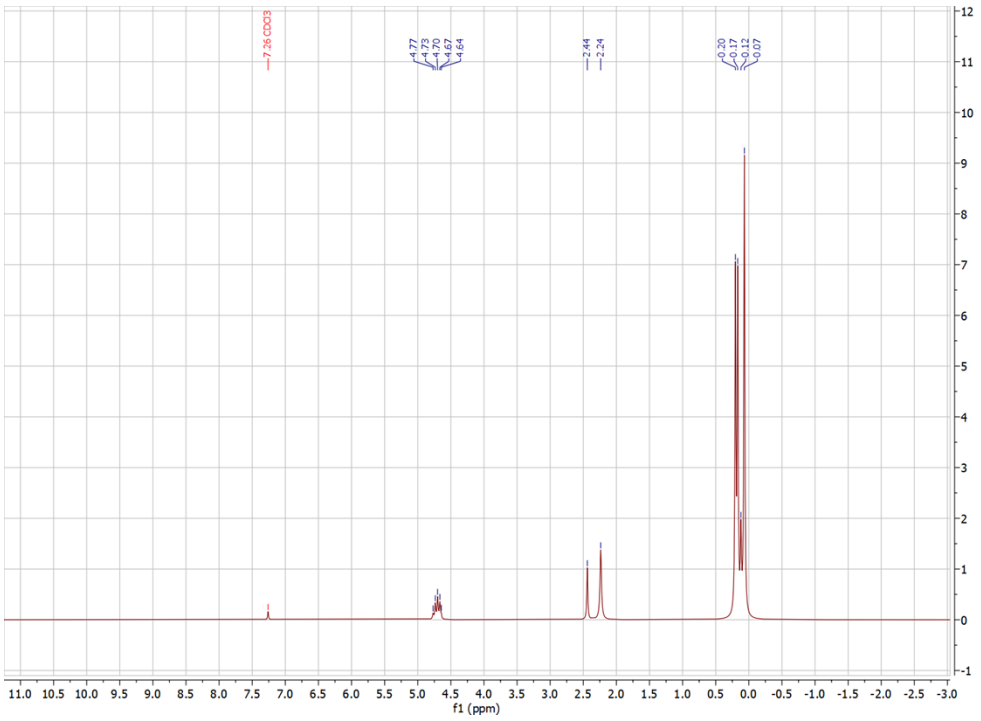
[a] = Monitored by 1H-NMR after 2 hours

2.13. ^{1}H-NMR of Entry 2 Table 2 [a]



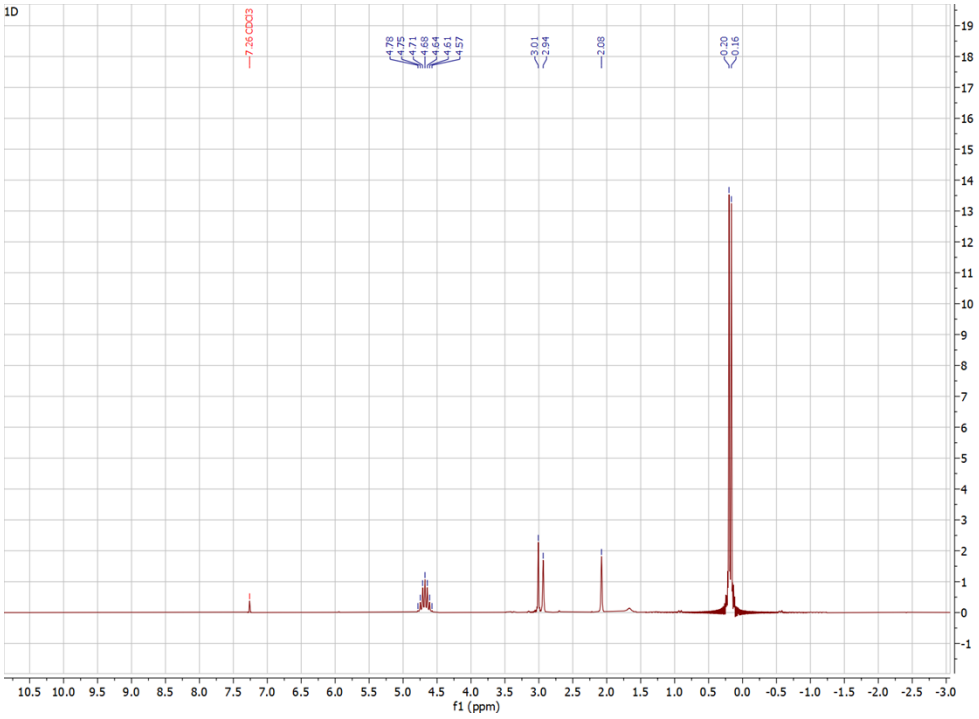
[a] Monitored by ^{1}H-NMR after 2 hours

2.14. ^{1}H-NMR of Entry 3 Table 2 [a]

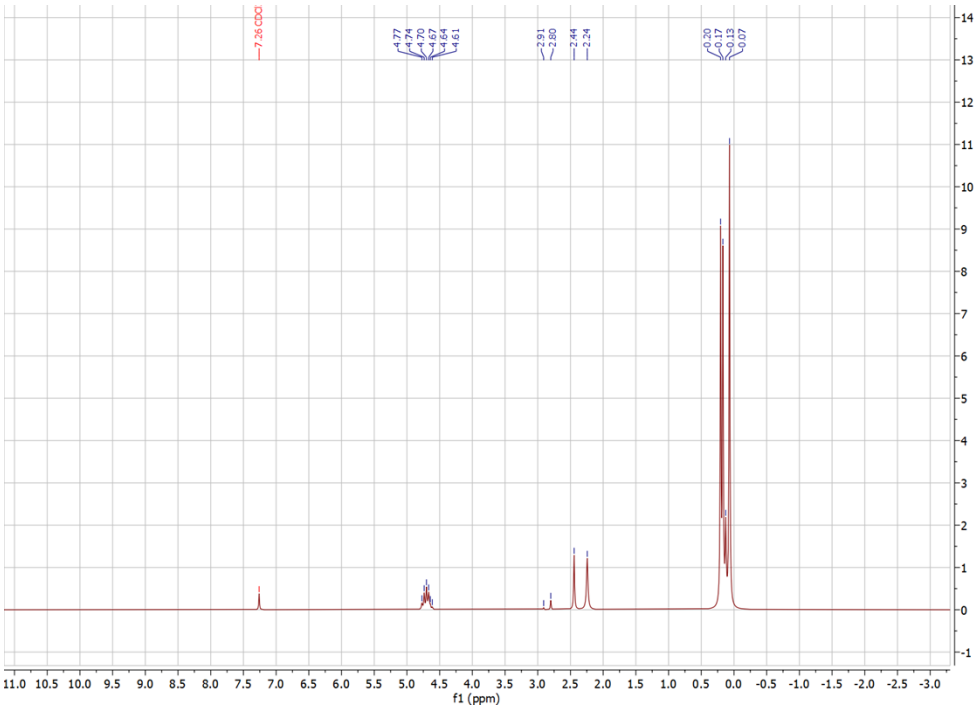


[a] Monitored by ^{1}H-NMR after 2 hours

2.15. ^{1}H-NMR of Entry 4 Table 2

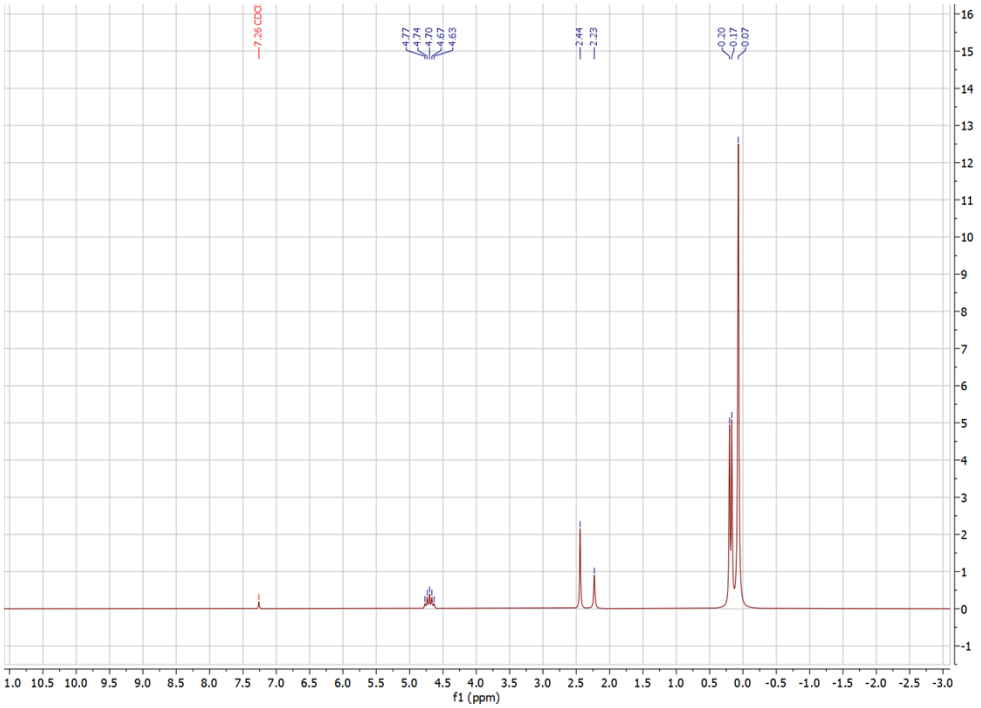


2.16. ^{1}H-NMR of Entry 5 Table 2 [a]

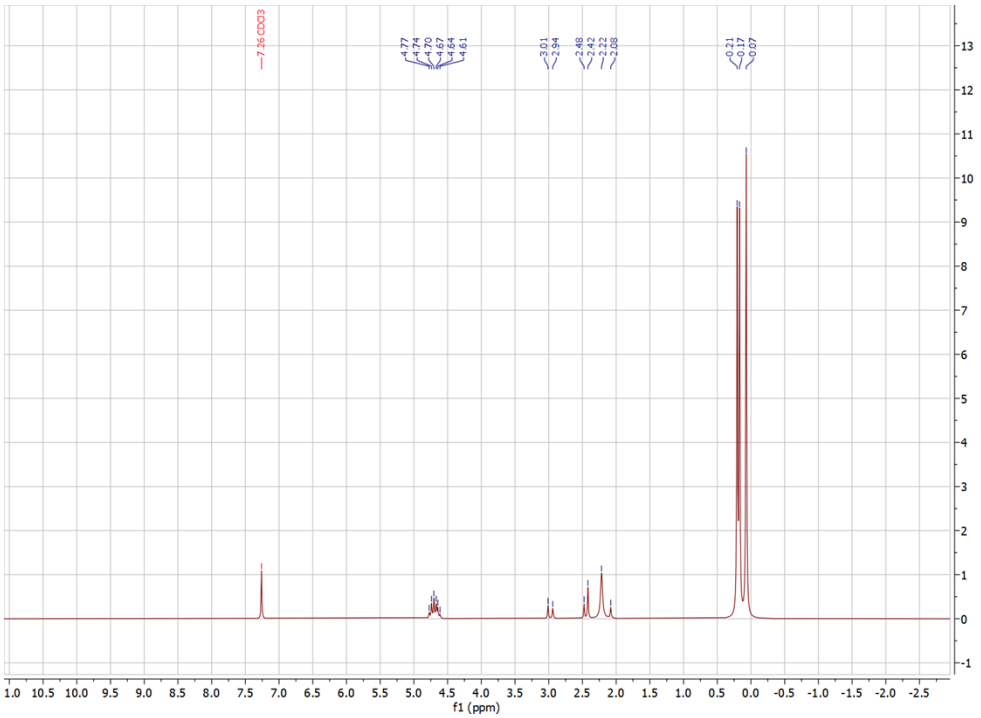


[a] Monitored by ^{1}H-NMR after 2 hours

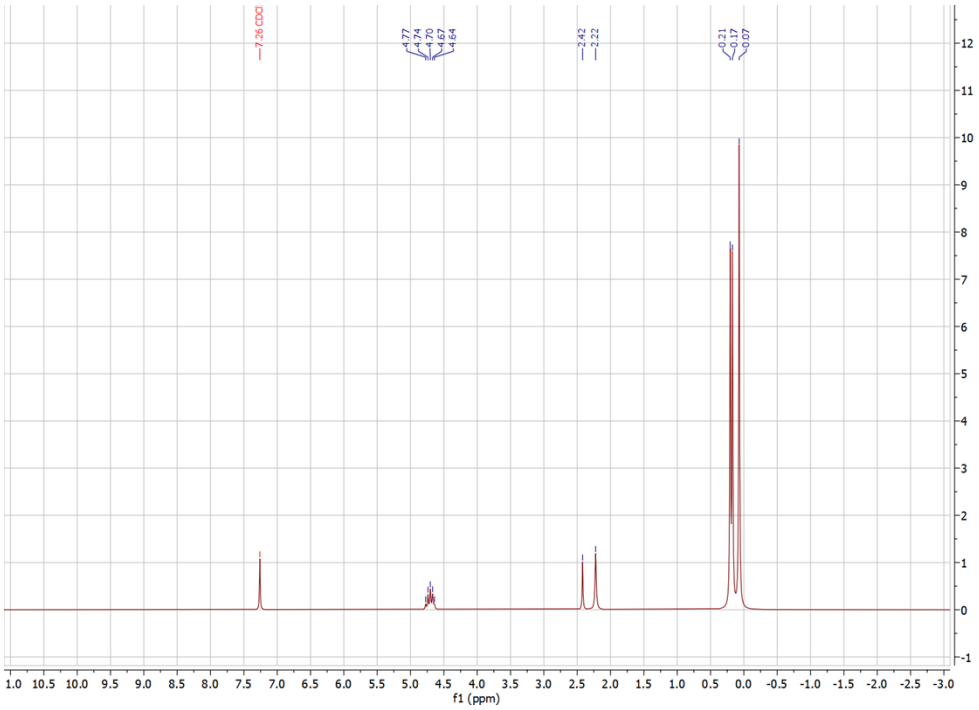
2.17. ^{1}H-NMR of Entry 5 Table 2 after 21 hours



2.18. ^{1}H-NMR of Entry 6 Table 2

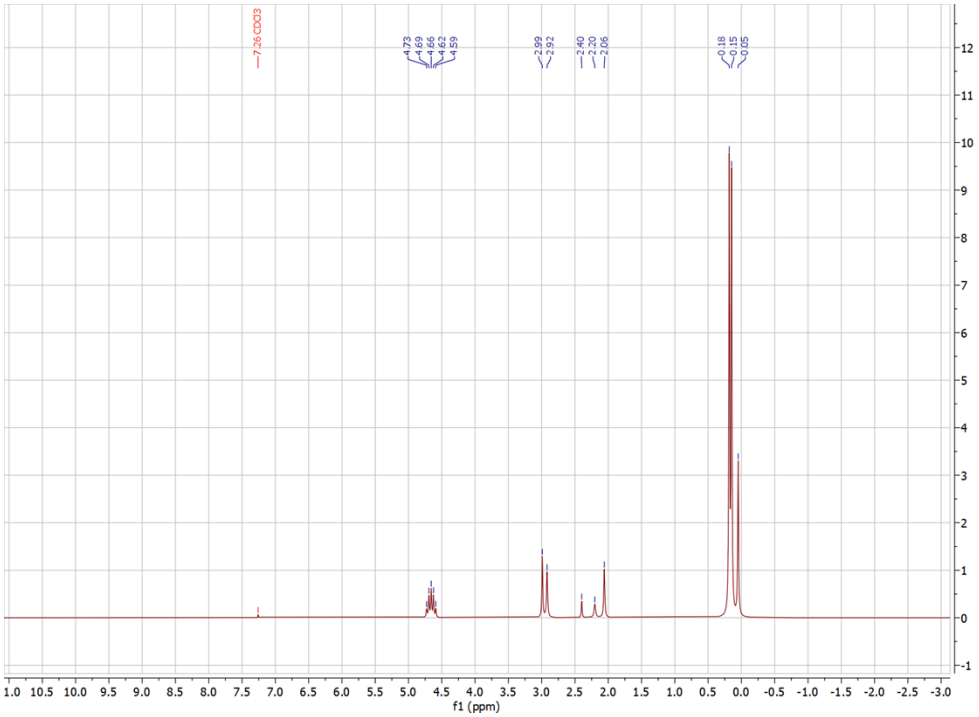


2.19. ^{1}H-NMR of Entry 7 Table 2 [b]

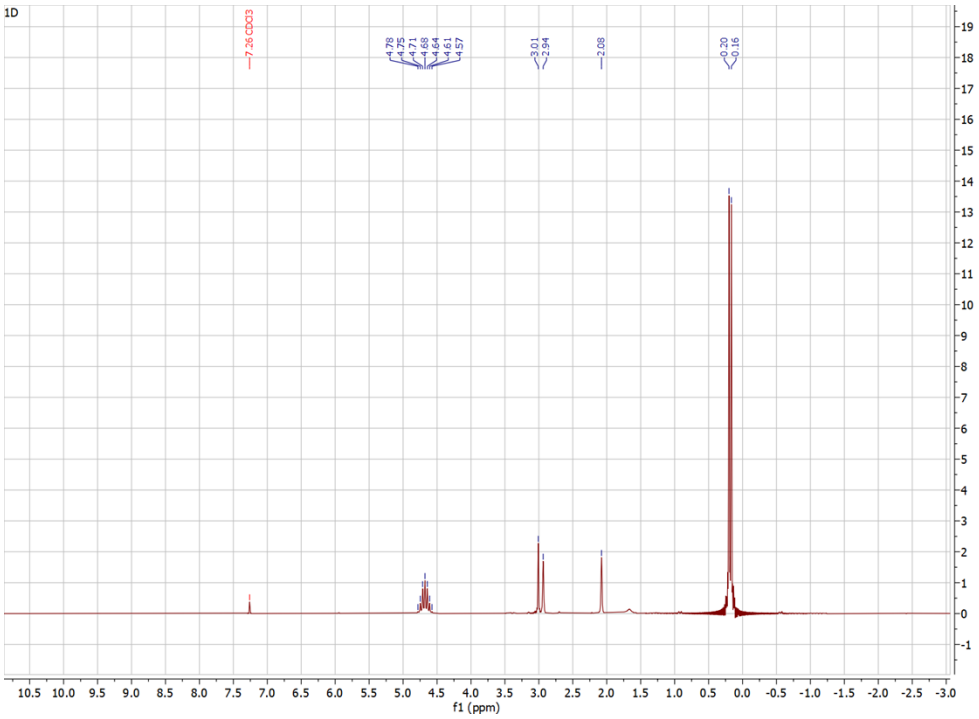


[b] Monitored by ^{1}H-NMR after 1 hours

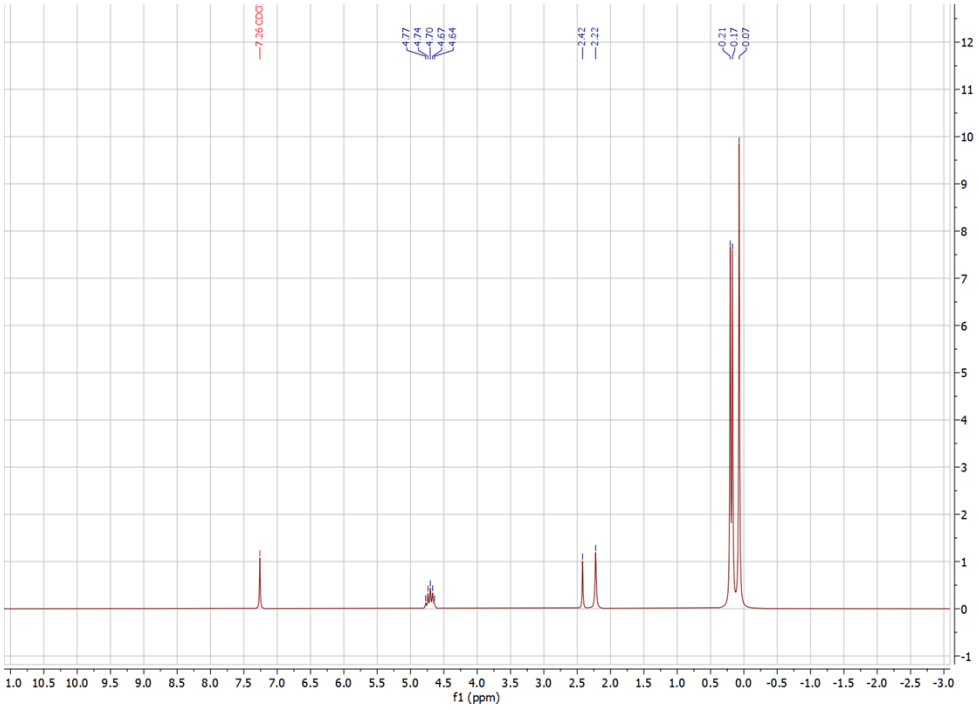
2.20. ^{1}H-NMR of Entry 8 Table 2



2.21. ^{1}H-NMR of Entry 9 Table 2

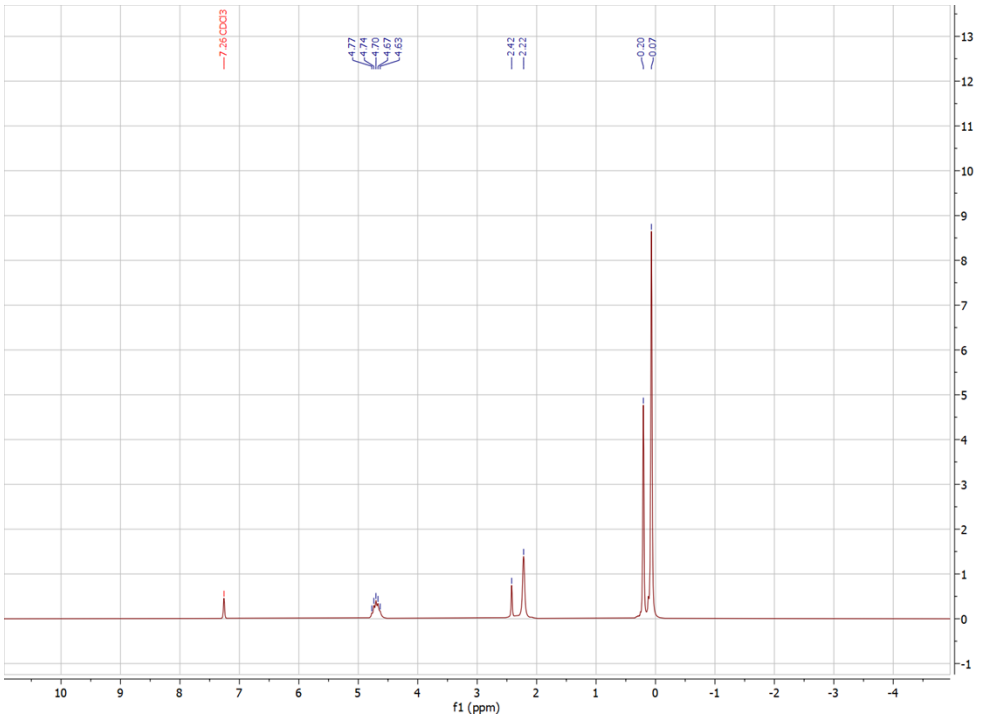


2.22. ^{1}H-NMR of Entry 10 Table 2 [a]

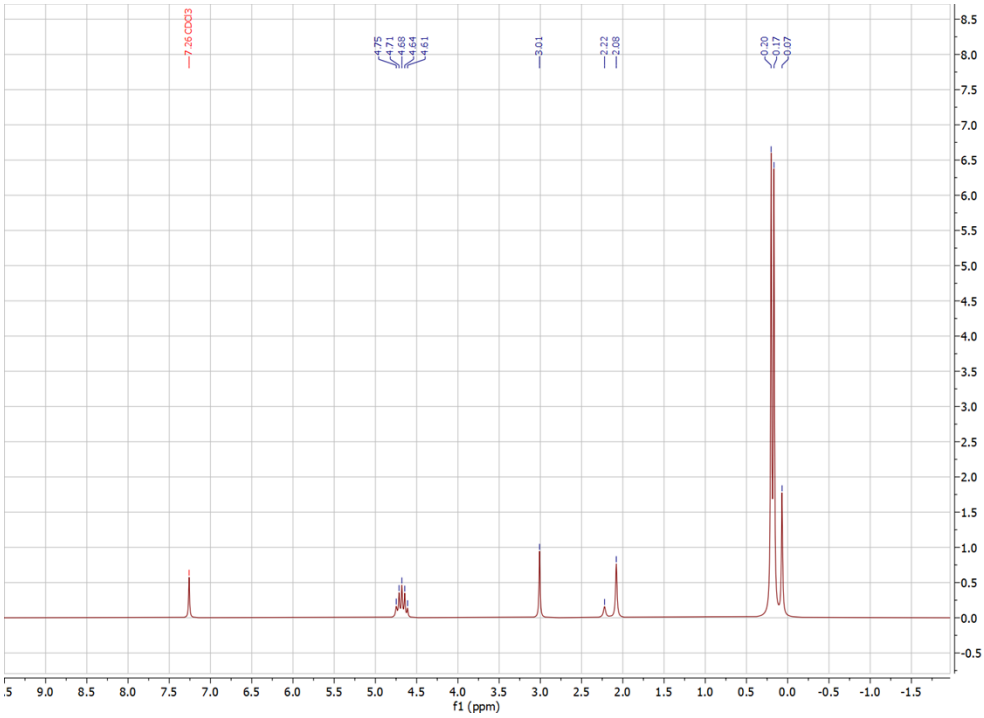


[b] Monitored by ^{1}H-NMR after 1 hours

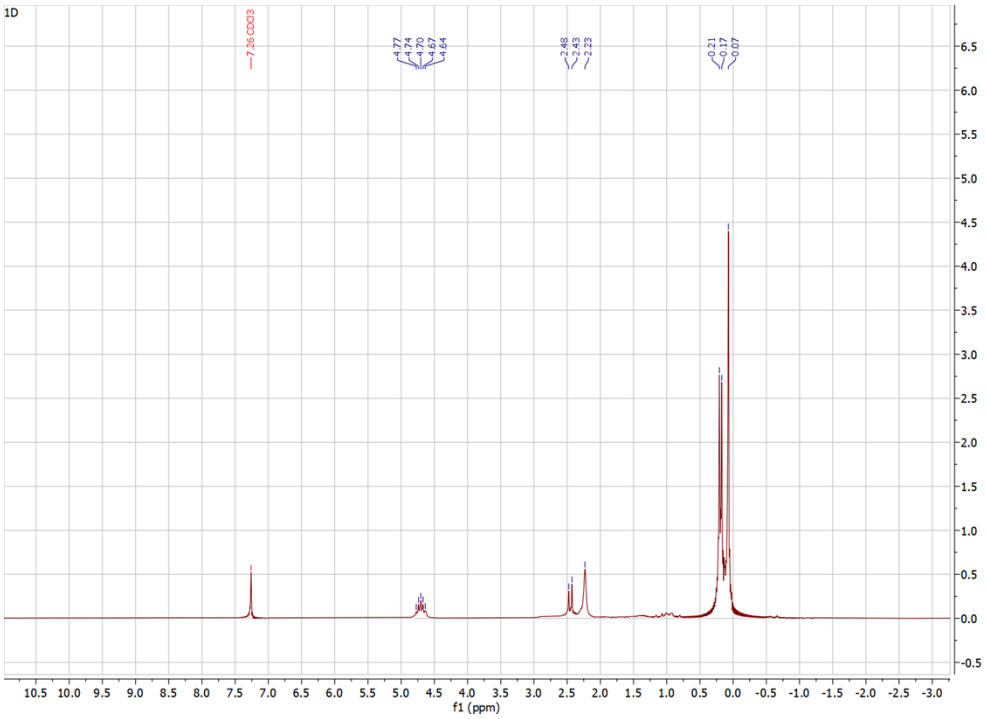
2.23. ^{1}H-NMR of Entry 1 Table 3



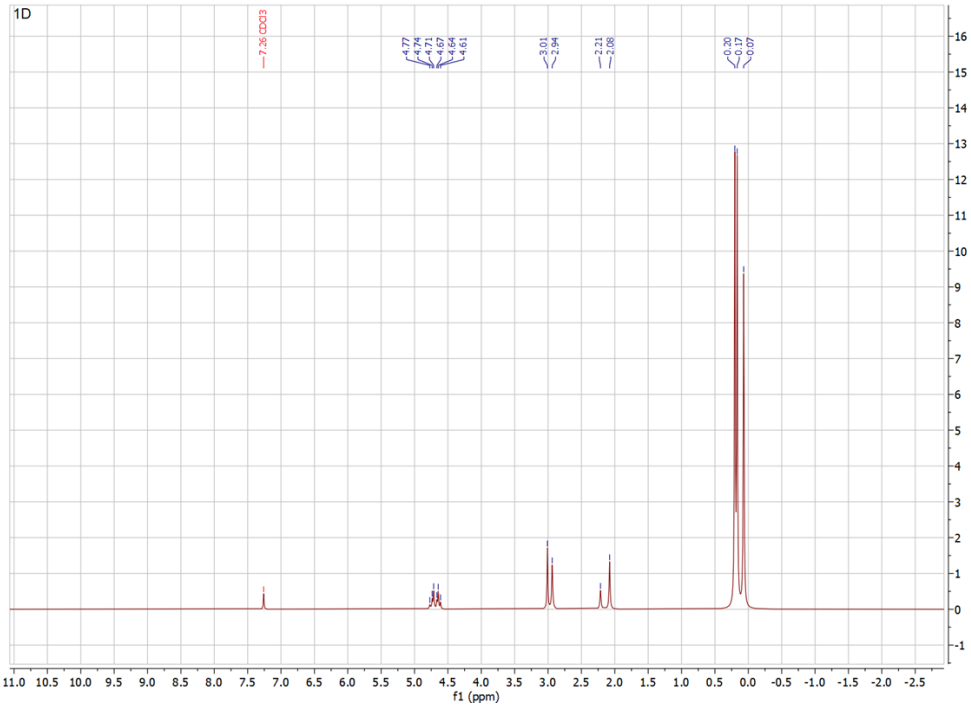
2.24. ^{1}H-NMR of Entry 2 Table 3



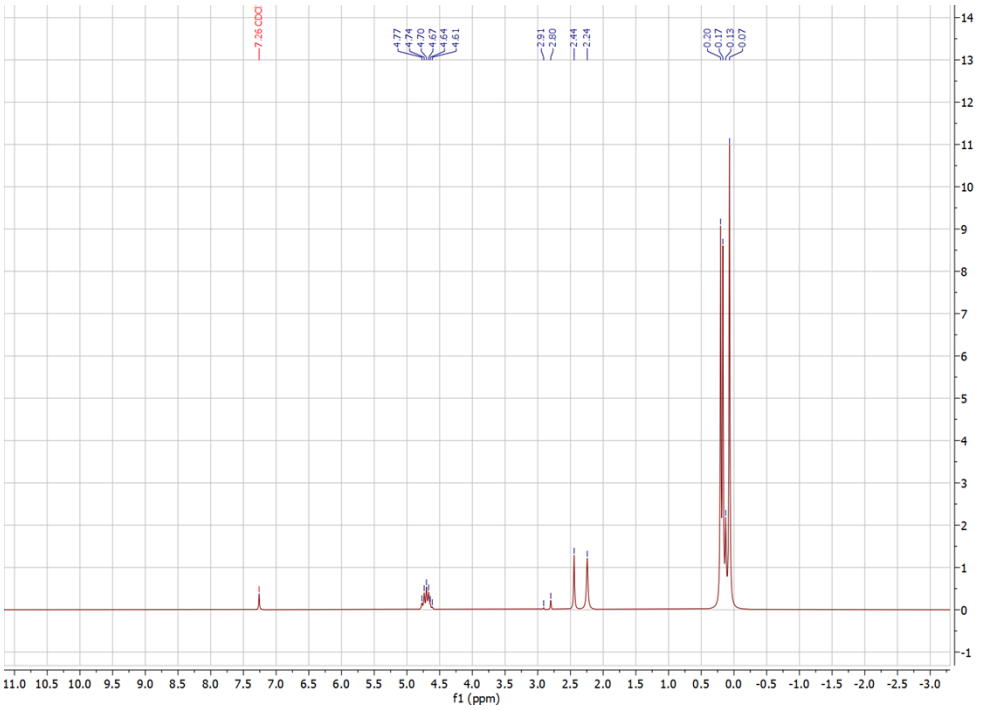
2.25. ^{1}H-NMR of Entry 3 Table 3



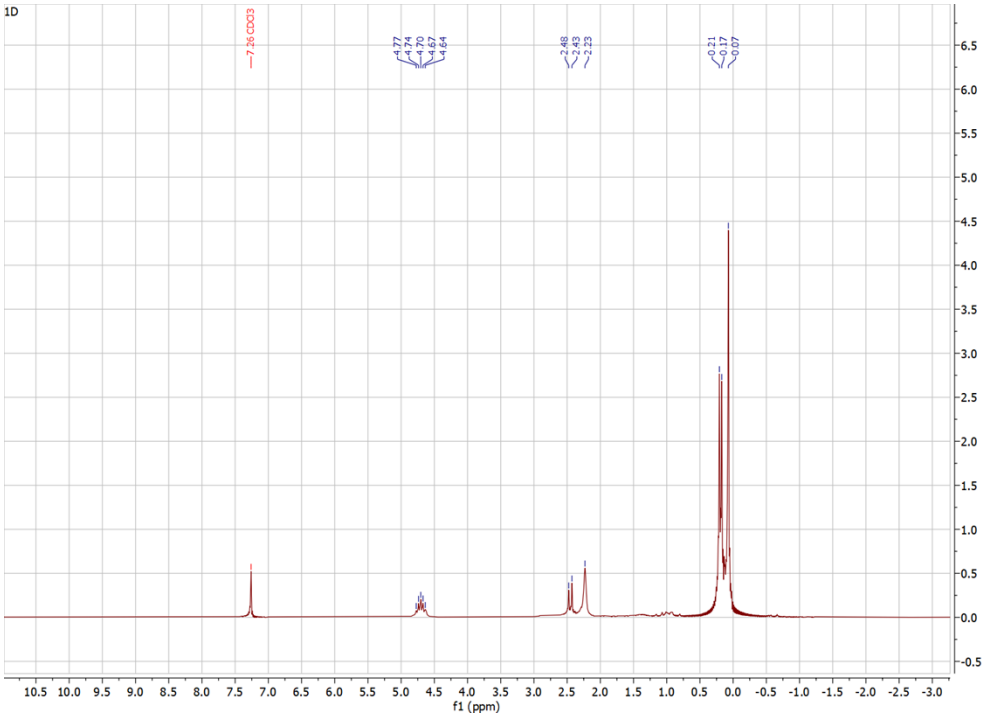
2.26. ^{1}H-NMR of Entry 4 Table 3



2.27. ^{1}H-NMR of Entry 5 Table 3



2.28. ^{1}H-NMR of Entry 6 Table 3



3. How the ML algorithm works

Machine learning (ML) prediction algorithms are only as accurate as the quality of the data provided for model training. However, Suntheticsʼ ML algorithms are self-correcting. They learn from the initial data and suggest subsequent experiments to confirm, adapt, and refine the predicted trends. SuntheticsML is designed to be independent of specific reactions. The tool learns from the reaction or system in question, the initial dataset provided, and the subsequent experiments suggested by the algorithms. It is specifically designed to leverage information from very small datasets, which sets it apart from other optimisation platforms (users can start utilising the algorithms with as few as 5 data points).

3.1 Model error and preliminary analysis

The model error is presented detailing RMSE and R-squared values for conversion, TON and TOF.

Table 1: model error with the average R-squared and RMSE for every target.

	Conversion	**TOF**	**TON**

R-squared	0.27 ± 0.39	0.60 ± 0.55	0.39 ± 0.56
RMSE	0.32 ± 0.09	27.01 ± 20.39	125.87 ± 25.52

R-squared (coefficient of determination) is another metric used in regression analysis to evaluate the goodness of fit of a model. The formula for *R-squared* is as follows:

$$R^2 = 1 - \frac{SSres}{SStot} = 1 - \frac{\sum_i (yi - \hat{y}i)^2}{\sum_i (yi - \bar{y}i)^2}$$

Where *SSR* (Sum of Squared Residuals) is the sum of the squared differences between the observed (actual) values and the predicted values and *SST* (Total Sum of Squares) is the sum of the squared differences between the observed values and the mean of the observed values. The *R-squared* measures the proportion of the variance in the dependent variable that is explained by the independent variables in the model. *R-squared* values range from 0 to 1, where:

- R^2=0: The model does not explain any of the variability in the dependent variable.
- R^2=1: The model explains all the variability in the dependent variable.

Root Mean Square Error (RMSE) measures the accuracy of a predictive model, particularly in the context of regression analysis. It quantifies the difference between the predicted values and the actual values of a variable. The formula for RMSE is as follows:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}i - yi)_2}{n}}$$

where N is the number of data points, y(*i*) is the *i*-th measurement, and $\hat{y}(i)$ is its corresponding prediction. Lower RMSE values indicate better agreement between predicted and observed values, with a value of 0 indicating a perfect fit. However, it is important to note that RMSE is sensitive to outliers, as the squared term magnifies large errors. In practice, RMSE is a widely used measure for assessing the performance of regression models.

3.2 Bootstrapping method

SuntheticsML utilises the *Bootstrapping* method for error estimation. Bootstrapping is a resampling technique used in statistics to estimate the variability or uncertainty of a statistic or model parameter by repeatedly drawing random samples (with replacement) from a given dataset. This involves creating multiple resampled datasets, analysing each one, and then using the distribution of results to assess the stability and reliability of the statistic, such as *R-squared* or *RMSE*. This technique improves the robustness and reliability of models by creating multiple datasets through random sampling with replacement and is particularly beneficial for assessing the stability of model performance and estimating the uncertainty associated with model predictions. When bootstrapping is employed to estimate error metrics, the variance reflects the stability and uncertainty associated with the model. The term "variance" here does not refer to statistical variance but rather to the spread or inconsistency in the model's predictions, so a lower variance is generally desirable, indicating more consistent and reliable model performance. In this case there is a lot of variance in the error metrics and this is probably because of the time series component. That is why in the second set of suggested reactions (Table 2) an ^{1}H-NMR spectrum was taken after 2 hours per each experiment.

3.3 Variable importance

In Figure S1 is reported the variable importance in descending order, with the most important variable at the top for the three targets: the TON, the Conversion (%) and the TOF, so it is easy to see the parameters that are affecting the system the most. From the analysis of the plot, it is evident that the most critical variable concerning the maximisation of the TON is the reaction time. Immediately following with a slightly lower value is the catalyst loading, which thus holds a similar significance. For the Conversion, the catalyst choice is of utmost importance, while for the TOF, the most crucial factor is the catalyst loading.

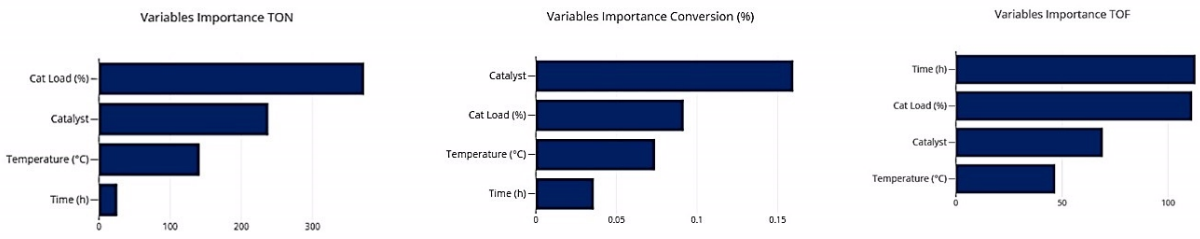


Figure S1: variable importance for the three targets of the reaction: TON, TOF and Conversion (%).

3.1 Prediction error graph

The prediction error was calculated, and it is reported in Figure S2. In the graph the x-axis representing the observed value and the y-axis representing the obtained value. It is an illustration of the relationship between observed values and corresponding obtained ones. Upon examining the distribution of obtained values, it is evident that they closely align with the trendline drawn by the predicted values. This observation suggests a strong correlation, indicative of a well-functioning model.

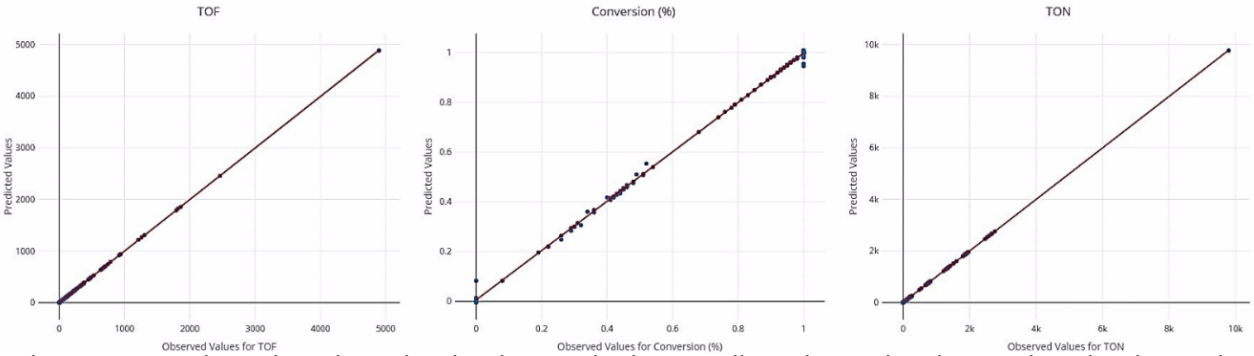


Figure S2: prediction error: deviation between the anticipated values of TOF, Conversion, and TON and the experimentally acquired values.

4. Author Contributions

Eleonora Casillo: conducted the experimental work; wrote the original draft.

Benon P. Maliszewski: conducted and supervised the experimental work; edited the manuscript.

César A. Urbina-Blanco: supervised project; edited the manuscript.

Thomas Scattolin: supervised project; edited the manuscript.

Catherine S. J. Cazin: supervised project; secured funding.

Steven P. Nolan: concept originator; supervised project; edited the manuscript; secured funding.

5. Notes and references

[1] B. P. Maliszewski, N. V. Tzouras, S. G. Guillet, M. Saab, M. Beliš, K. Van Hecke, F. Nahra and S. P. Nolan, *Dalton Trans.*, 2020, **49**, 14673–14679.

[2] B. P. Maliszewski, E. Casillo, P. Lambert, F. Nahra, C. S. J. Cazin and S. P. Nolan, *Chem. Commun.*, 2023, **59**, 14017–14020.