Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers.

Supplementary Information

Optimizing Chemical Reaction Conditions Using Deep Learning: a Case Study for Suzuki-Miyaura Cross-Coupling Reaction

Zunyun Fu,[ab] Xutong Li,[bd] Zhaohui Wang,[c] Zhaojun Li,[b] Xiaohong Liu,[b] Xiaolong Wu,[b] Jihui Zhao,[bd] Xiaoyu Ding,[bd] Xiaozhe Wan,[bd] Feisheng Zhong,[bd] Dingyan Wang,[bd] Xiaomin Luo,[bd] Kaixian Chen,[ab] Hong Liu,[cd] Jiang Wang,[*cd] Hualiang Jiang[*bde] and Mingyue Zheng [*bd]

[a.]School of Chinese Materia Medica, Nanjing University of Chinese Medicine, 138 Xianlin Road, Jiangsu, Nanjing 210023, China

[b.]Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China

[c.]CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China.

[d.]University of Chinese Academy of Sciences, No.19A Yuan Road, Beijing 100049, China

[e.]State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210023, Jiangsu, People's Republic of China.

Table of Contents

1. Data collection

The high-quality reaction data we used was collected from the remarkable work of Jensen et al., who explored and optimized the reaction performance of Pd-catalyzed Suzuki-Miyaura reaction in microfluidic system.[1] The structures of aryl halides Ar_1X, boron reagents $Ar_2B(OR)_2$, and precatalysts were shown in Fig. 1. All the experiments were conducted in solvent THF: H_2O (5:1) with the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). In the microfluidic system, how the catalyst types, catalyst loading, residence time, and reaction temperature affecting the reaction performance were investigated while other reaction conditions were fixed: the catalyst loading ranged from 0.5 mol% to 2.5 mol%, reaction temperature ranged from 30 ºC to 110 ºC, and residence time ranged from 1 minute to 10 minutes. For **1a-2a**, **1c-2a**, and **1a-2b**, 105 microfluidic experiments were conducted, while 106 microfluidic experiments for **1b-2c** were conducted. For the case of identical reaction conditions but different yields, we retained the reaction data with higher yield. After deleting 34 reactions (see S1.tex), we obtained a total of 387 reaction samples(see S2.tex).

2. Model comparison

2.1 Comparison to DNN models with One-hot encoded components

With the QM-based features of reagents and reaction condition as inputs, our DNN model was trained to learn the relationships between the structure of reagents, reaction conditions and yields, and achieved excellent performance on external set by extrapolating reasonably to unseen reaction conditions. For comparison, we also investigated how the deep learning model performed with the one-hot encoded components, where discrete variables such as aryl halides Ar_1X, boron reagents $Ar_2B(OR)_2$ and catalysts were one-hot encoded. Finally, a total of 17 descriptors were used to characterize each reaction sample, and the division of data set was consistent with the DNN model built with QM-based features. The one-hot encoded DNN model with the best performance achieved RMSE of 3.156, R2 of 0.993 on training set, RMSE of 9.036, R^2 of 0.943 on test set, RMSE of 9.782, R^2 of 0.928 on external set when layer-size was [1000, 1000], learning-rate was 0.001, dropout-1 was 0.3, dropout-2 was 0.1, batch-size was 20 (Fig. S1a, b and c).

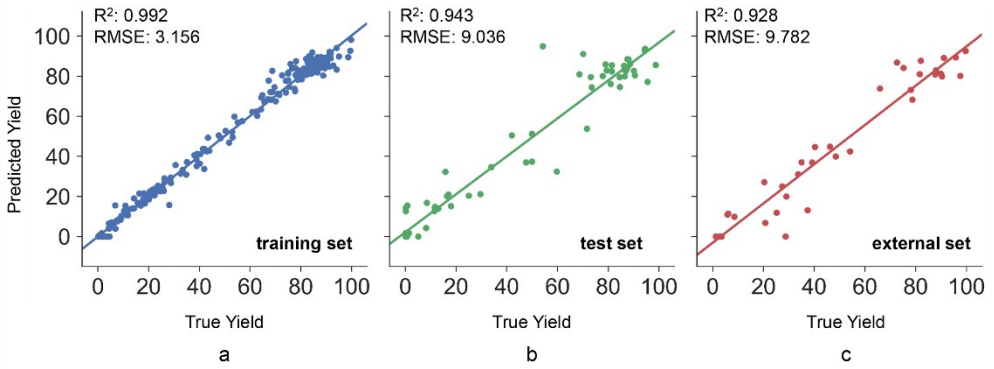


Fig. S1 The performance of DNN model trained with one-hot encoded descriptors (a) The performance of DNN model trained with one-hot encoded descriptors on the training set (RMSE: 3.156, R^2: 0.993), (b) The performance of DNN model trained with one-hot encoded descriptors on the test set (RMSE: 9.036, R^2: 0.943), (c) The performance of DNN model trained with one-hot encoded descriptors on the external validation set (RMSE: 9.782, R^2: 0.928).

As shown, the deep learning model built on one-hot encoded components yielded similar performance as compared to the model built on QM-based features. To avoid the same components appearing in both the training set and test set, another control experiment was performed with 283 reaction samples of **1a-2a**, **1c-2a**, **1a-2b** as the training set, and 104 reaction data of coupling of **1b-2c** as the test set. In this experiment, the predictive power of resulted models decreased significantly due to the heterogeneity between training and test data, and the model with QM-based descriptors showed a higher R^2 on the test set (R^2= 0.295) than that with one-hot encoded descriptors (R^2 =0.165). This result suggested that the QM-based encoding scheme indeed has better generalization capability.[2] Moreover, a Y-randomization test[2, 3] was performed to validate the structure-activity relationship present in the QM-based descriptors, in

which the Y- randomized DNN model yielded significantly deteriorated R^2 of 0.027 on the external test. This result indicated that DNN model with QM-based features as inputs encoded meaningful information related to reaction.

2.2 Comparison to k-nearest neighbours regression model

In addition to deep neural network (DNN) model, we also implemented k-nearest neighbour regression model. The k-nearest neighbour (KNN) regression model was developed with Scikit-Learn (Version 0.19.0) and the code was written with Python 3.6. Likewise, we firstly used data set characterized by QM-based features and reaction conditions to build the KNN model, and the division of data was also as same as the DNN model. In the training process, a grid search and 5-fold cross-validation within the training set was employed to optimize the n-neighbours, the one single tuneable parameter of KNN. The parameter n-neighbours was allowed to take a value from [1,2,3,4,5], and the best performing KNN model achieved RMSE of 10.656, R^2 of 0.918 on training set, RMSE of 11.295, R^2 of 0.911 on test set, RMSE of 10.988, R^2 of 0.927 on external set when the n-neighbours was 2 (Fig. S2a, b and c).

Apart from QM-based features, the KNN model with one-hot encoded descriptors as inputs was also established. As the procedures described above, and the KNN model performed best when the n-neighbours was 2. The optimal KNN model achieved RMSE of 10.018, R^2 of 0.925 on training set, RMSE of 20.319, R^2 of 0.949 on test set, RMSE of 13.548, R^2 of 0.828 on external set (Fig. S2d, e and f). According to these metrics, it could be clearly found that the DNN models were superior to KNN models.

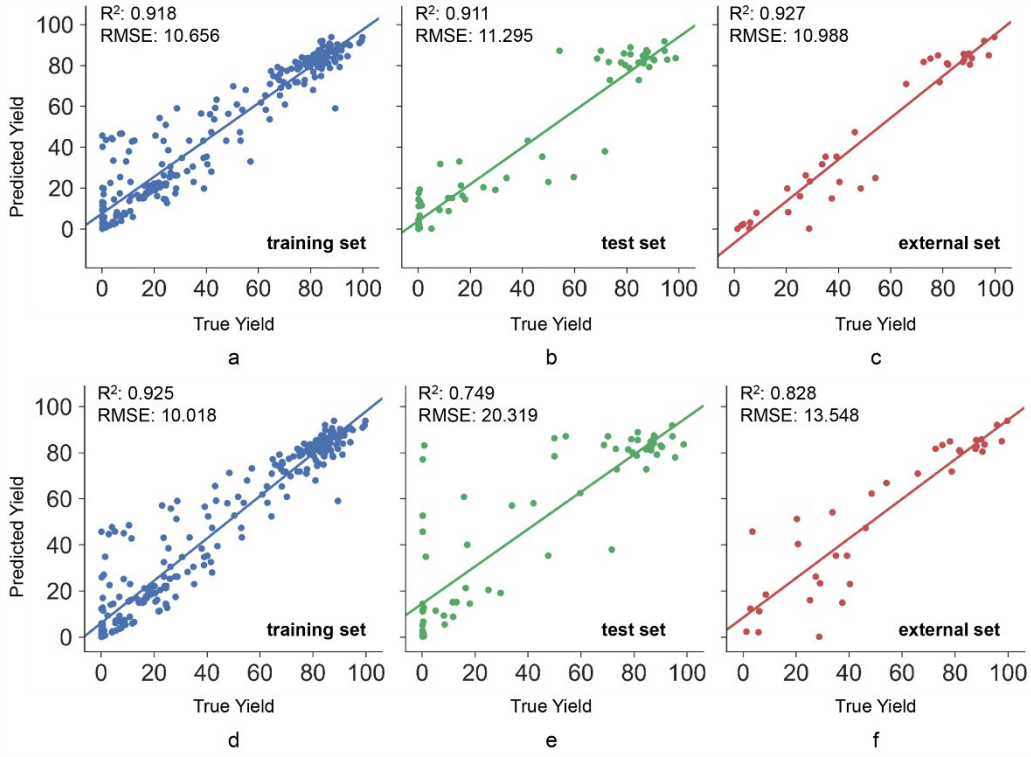
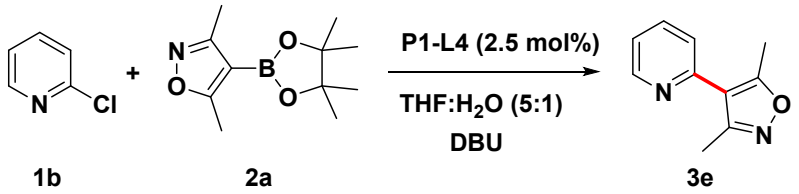


Fig. S2 The performance of KNN models (a) The performance of the optimal KNN model with QM-based features as inputs on training set (RMSE: 10.656, R^2: 0.918). (b) The performance of the optimal KNN model with QM-based features as inputs on the test set (RMSE: 11.295, R^2: 0.911). (c) The performance of the optimal KNN model with QM-based features as inputs on external validation set (RMSE: 10.988, R^2: 0.927). (d) The performance of the optimal KNN model with one-hot encoded features as inputs on training set (RMSE: 10.018, R^2: 0.925). (e) The performance of the optimal KNN model with one-hot encoded features as inputs on the test set (RMSE: 20.319, R^2: 0.749). (f) The performance of the optimal KNN model with one-hot encoded features as inputs on external validation set (RMSE: 13.548, R^2: 0.828).

3. Experimentation Procedures

General procedures for the products 3a-3i (compound 3e as example)



1b **2a** **3e**

A flask was charged with 2-chloropyridine (**1b**, 100 mg, 0.87 mmol), 3,5-dimethyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)isoxazole (**2a**, 318 mg, 1.4 mmol), P1-L4 (21.8 mg, 2.5 mol %), DBU (266 mg, 1.74 mmol) and THF : H_2O (5 : 1, 5 mL) . The reaction mixture was stirred at 110 °C for 10 min under argon condition. After that, the solvent was extracted with EtOAc. The organic layer was washed with brine, dried over anhydrous Na_2SO_4, and concentrated in vacuo. The residue was purified by silica gel chromatography using petroleum ether/EtOAc (15:1) to afford the product **3e** as a white solid (yield: 73%).

4. The Predicted results of DNN model

4.1 Prediction results for 4 modelled schemes

In the work of Jensen et al., the optimal reaction conditions and corresponding yields were determined under microfluidic system for each study case.[1] For the four couplings of reactants, the optimal DNN model was also employed to predict yield under experimentally determined favourable microfluidic reaction conditions. As summarized in Table S1, the predicted yields were very close to the results observed in microfluidic experiments. The largest discrepancy lied in **1b-2c**, which still fall within 10% difference. The good agreement between the predictions and the experimental observations suggested the reliability of our model.

Table S1. The optimal reaction conditions found by microfluidic system and corresponding yields from optimal DNN regression model and microfluidic experiments.

Scheme	Optimal Conditions[a]	Yield (%)[a]	Yield (%)[b]
1a-2a	10 min 110 ºC 2.1 mol % P1-L5	35.0	34.6
1c-2a	10 min 110ºC 1.2 mol % P1-L4	82.0	81.8
1a-2b	3.9 min 110 ºC 1.2 mol % P1-L1	88	84.4
1b-2c	4.7 min 97 ºC 1.0 mol % P1-L1	90.0	81.3

[a] microfluidic optimal conditions and yield [b] DNN model predicted yield

How precatalysts, catalyst loading, residence time, and reaction temperature influencing the reaction performance had been investigated by Jensen et al., by considering 8 precatalysts, catalyst loading from 0.5 mol% to 2.5 mol%, residence time from 1 minute to 10 minutes, and reaction temperature from 30 ºC to 110 ºC.[1] For each scheme, there were thousands of possible reaction conditions, and it would be burdensome to investigate reaction conditions thoroughly, even for the efficient microfluidic system. In contrast, computational model is more suitable to address such problems. For each coupling of reactants and each precatalyst, we fully mapped reaction space with interval of catalyst loading 0.5 mol%, the interval of residence time 30 s, and the interval of reaction temperature 20 ºC. A total of 475 reactions were obtained for each coupling of reactants and each precatalyst. The same procedure was applied to 4 couplings of reactants and 8 precatalysts, resulting a total of 15,200 candidate reactions. Using the optimal DNN regression model, all the computational predictions for 15,200 reactions can be completed within 1 minute on Nvidia GPUs (GeForce RTX 2080)(see S3.tex).

Based on the predicted results, we could clearly observe the performance of each coupling of reactants under the catalysis of each precatalyst (Fig. S3a). For **1a-2a**, the predicted yield was zero or close to zero under most reaction conditions, and the highest yield occurred under the catalysis of trialkylphosphine ligand precatalyst P1-L5, but was only 34.4%. Different from 1**a-2a**, the performance of **1c-2a** was much better. The electron-deficient bidentate ligand precatalyst P1-L4 outperformed other precatalysts with a maximum predicted yield of 91.6%. For **1a-2b**, dialkylbiarylphosphine ligands (L1, L2, L3) conferred significantly higher yields than trialkylphosphine ligands (L5, L6, L7) and bidentate ligand L4, in which bidentate ligand precatalyst P1-L4 could not catalyze the scheme. Among the four kinds of dialkylbiarylphosphine ligand precatalysts (P1-L1, P1-L2, P1-L3, P2-L1), P1-L3 performed best, and the predicted highest yield of target product reached 97.6%. Similar to the **1a-2b**, dialkylbiarylphosphine ligands (L1, L2, L3) also behaved well in **1b-2c**, especially P1-L1, showing a maximum yield of 100%.

In addition to predicted yield, how reaction temperature and catalyst loading affecting the reaction performance can also be clearly observed. As shown in Fig. S3b and Fig. S3c, the performance of most reactions would increase with the increase of reaction temperature and catalyst loading. However, the optimal predicted yield of **1a-2a** occurred at catalyst loading of 2.0 mol%, and the optimal temperature for **1b-2c** was 70 ºC.

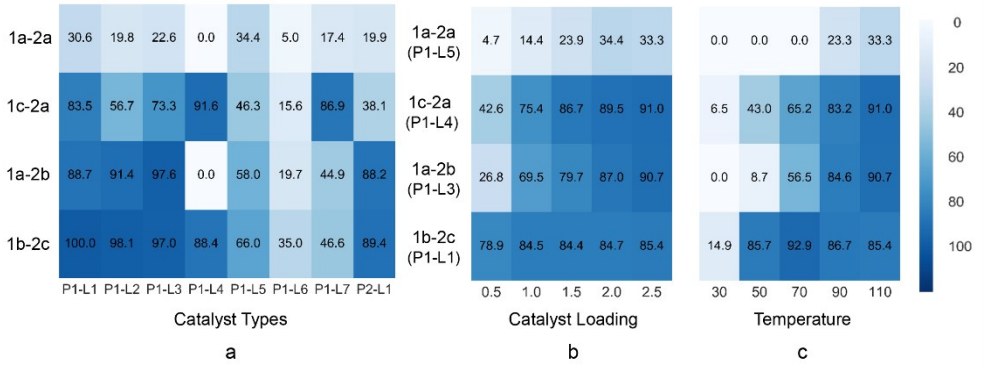
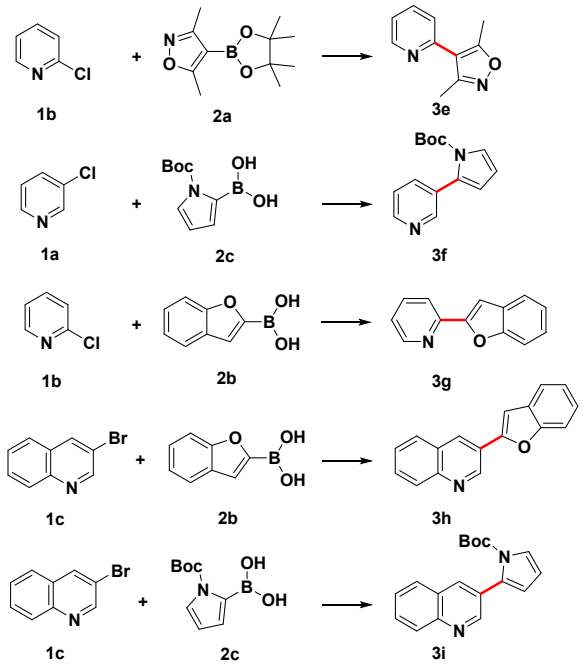


Fig. S3 The predicted results for 4 couplings of reactants used in modelling (a) The highest yields obtained when using 8 precatalysts under favourable conditions. (b) The effects of catalyst loading when using the optimal precatalyst (with the reaction temperature set to 110 ºC and residence time 10 minutes). (c) The effects of reaction temperature when using the optimal precatalysts (with the catalyst loading set to 2.5 mol% and residence time 10 minutes).

By analyzing the predicted results, the highest-yielding reaction conditions for each scheme were obtained. To experimentally validate the prediction results, both the reaction conditions found by microfluidic system and our DNN model were scale-up in conventional flask reactors to investigate their corresponding performance. For microfluidic system, reaction occurs more quickly due to the enhanced heat- and mass-transfer rate,[4] however, when reproduced in flaks, we set no time limit until the reaction was complete. The experimental results for the comparison were shown in Table S3. For **1a-2a**, **1c-2a**, **1b-2c**, the yields under the reaction conditions found by microfluidic system and our predictive model were highly consistent. However, for **1a-2b**, the predicted optimal reaction conditions found by our DNN model were completely different from those determined by microfluidic system, and the yield of **3c** was improved from 30.5% to 89.2%. By searching more extensively, our model did find more productive reaction conditions.

4.2 Prediction Results for External Schemes

In addition to predicting yield and finding out high-yielding reaction conditions for schemes used in modelling, it is of interest to investigate whether the model also has predictive power for new reactions. Because of the high-tolerance towards functional groups,[5-7] we firstly made new combinations of existing reactants, i.e., letting **1b** react with **2a**, **1a** react with **2c**, **1b** react with **2b**, **1c** react with **2b** and **2c** (Scheme S1). Following the same procedure used in Scheme 1, we designed a total of 19,000 reactions for the 5 new combinations of existing reactants (see S4.tex). The optimal DNN model was employed to predicted yields, determined high-yielding reaction conditions, and the predicted results were confirmed by experiments.



Scheme S1 5 new schemes obtained by crossover of existing reactants

The above predictions were made to new schemes between existing reactants. A more challenging task for evaluating the generalization ability of our model is to predict reactions for unseen reactants. To this aim, we made modifications to the structures of reactants and obtained 3 extra novel schemes (Scheme 2), which were used to examine whether our model could accurately predict the yields. We designed 11,400 reactions for 3 external schemes, used DNN model to predict yield and determine optimal reaction conditions(see S5.tex). Based on the predicted results, both precatalysts P1-L1 and P1-L4 were favourable for the 3 new reactions (Fig. S4a). Among them, the precatalyst P1-L1 performed best for schemes where 2a involved, and the predicted highest yields were 53.6% and 45.8%, respectively. For 1b-2d, precatalyst P1-L4 exhibited higher catalytic efficiency, with the predicted maximum yield of 80.6%. Under the catalysis of optimal precatalysts, the effects of reaction temperature and catalyst loading on reaction performance were shown in Fig. S4b and Fig. S4c.

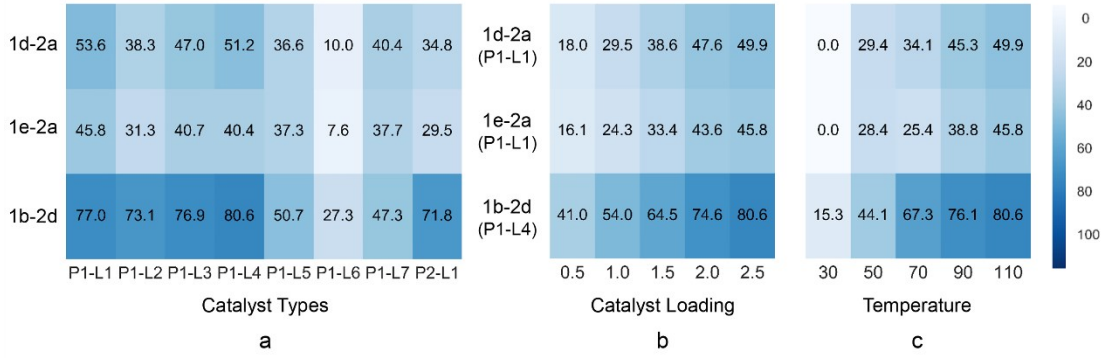
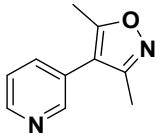


Fig. S4 The predicted results for 3 unseen schemes (a) The highest yields obtained when using 8 precatalysts under favourable conditions. (b) The effects of catalyst loading when using the optimal precatalyst (with the reaction temperature set to 110 ºC and residence time 10 minutes). (c) The effects of reaction temperature when using the optimal precatalysts (with the catalyst loading set to 2.5 mol% and residence time 10 minutes).

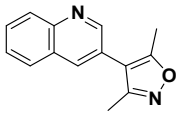
5. References

1. B. J. Reizman, Y.-M. Wang, S. L. Buchwald and K. F. Jensen, Suzuki–Miyaura cross-coupling optimization enabled by automated feedback, *React. Chem. Eng.*, 2016, **1**, 658-666.
2. J. G. Estrada, D. T. Ahneman, R. P. Sheridan, S. D. Dreher and A. G. Doyle, Response to Comment on "Predicting reaction performance in C-N cross-coupling using machine learning", *Science*, 2018, **362**, 3.
3. A. Tropsha, P. Gramatica and V. K. Gombar, The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models, *QSAR Comb. Sci.*, 2003, **22**, 69-77.
4. J. P. McMullen and K. F. Jensen, in *Annual Review of Analytical Chemistry, Vol 3*, eds. E. S. Yeung and R. N. Zare, Annual Reviews, Palo Alto, 2010, vol. 3, pp. 19-42.
5. A. Suzuki, Recent advances in the cross-coupling reactions of organoboron derivatives with organic electrophiles, 1995-1998, *J. Organomet. Chem.*, 1999, **576**, 147-168.
6. R. Martin and S. L. Buchwald, Palladium-Catalyzed Suzuki-Miyaura Cross-Coupling Reactions Employing Dialkylbiaryl Phosphine Ligands, *Acc. Chem. Res.*, 2008, **41**, 1461-1473.
7. I. Maluenda and O. Navarro, Recent Developments in the Suzuki-Miyaura Reaction: 2010-2014, *Molecules*, 2015, **20**, 7528-7557.

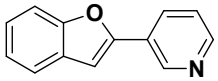
6. NMR Spectra



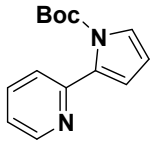
3,5-dimethyl-4-(pyridin-3-yl)isoxazole (3a): white solid, m.p. 43 – 44 ºC.
^{1}H NMR (400 MHz, DMSO-d_6): δ 8.61 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.57 (dd, *J* = 4.8, 1.7 Hz, 1H), 7.84 (ddd, *J* = 7.9, 2.3, 1.6 Hz, 1H), 7.49 (ddd, *J* = 7.9, 4.8, 0.9 Hz, 1H), 2.41 (s, 3H), 2.23 (s, 3H).
^{13}C NMR (125 MHz, DMSO-d_6): δ 166.5, 158.7, 149.9, 149.0, 136.8, 126.5, 124.3, 113.5, 11.8, 10.8.
LRMS (ESI) m/z: $[M+H]^+$ found 175.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{10}H_{10}N_2O$ 175.0866; found 175.0867.



3,5-dimethyl-4-(quinolin-3-yl)isoxazole (3b): white solid, m.p. 86 – 87 ºC.
^{1}H NMR (400 MHz, $CDCl_3$): δ 8.84 (d, *J* = 2.2 Hz, 1H), 8.16 (dt, *J* = 8.4, 0.9 Hz, 1H), 8.07 – 8.02 (m, 1H), 7.89 – 7.83 (m, 1H), 7.77 (ddd, *J* = 8.4, 6.9, 1.4 Hz, 1H), 7.61 (ddd, *J* = 8.1, 6.8, 1.2 Hz, 1H), 2.48 (s, 3H), 2.33 (s, 3H).
^{13}C NMR (125 MHz, DMSO-d_6): δ 166.8, 158.9, 151.1, 147.0, 135.9, 130.3, 129.2, 128.6, 128.0, 127.6, 123.8, 113.6, 11.9, 10.9.
LRMS (ESI) m/z: $[M+H]^+$ found 225.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{14}H_{12}N_2O$ 225.1022; found 225.1026.



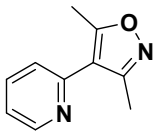
3-(benzofuran-2-yl)pyridine (3c): white solid, m.p. 74 – 76 ºC.
^{1}H NMR (400 MHz, DMSO-d_6): δ 9.14 (d, *J* = 2.3 Hz, 1H), 8.58 (dd, *J* = 4.9, 1.6 Hz, 1H), 8.27 (dt, *J* = 8.0, 1.9 Hz, 1H), 7.69 (d, *J* = 7.7 Hz, 1H), 7.65 (d, *J* = 8.2 Hz, 1H), 7.61 – 7.58 (m, 1H), 7.55 – 7.50 (m, 1H), 7.40 – 7.33 (m, 1H), 7.28 (t, *J* = 7.5 Hz, 1H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 154.9, 153.0, 150.0, 146.3, 132.3, 129.0, 126.3, 125.6, 124.5, 123.9, 121.9, 111.8, 104.0.
LRMS (ESI) m/z: $[M+H]^+$ found 196.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{13}H_9NO$ 196.0757; found 196.0757.



Tert-butyl 2-(pyridin-2-yl)-1*H*-pyrrole-1-carboxylate (3d): colorless liquid.
^{1}H NMR (400 MHz, $CDCl_3$): δ 8.60 (ddd, *J* = 4.9, 1.8, 0.9 Hz, 1H), 7.67 (td, *J* = 7.7, 1.8 Hz, 1H), 7.41 – 7.34 (m, 2H), 7.18 (ddd, *J* = 7.6, 4.9, 1.2 Hz, 1H), 6.40 (dd, *J* = 3.3, 1.7 Hz, 1H), 6.23 (t, *J* = 3.3 Hz, 1H), 1.34 (s, 9H).

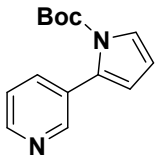
^{13}C NMR (125 MHz, $CDCl_3$): δ 152.9, 149.3, 148.8, 135.8, 134.0, 123.6, 123.5, 121.7 115.6, 110.5, 83.6, 27.6.

LRMS (ESI) m/z: $[M+H]^+$ found 245.0, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{14}H_{16}N_2O_2$ 245.1285; found 245.1282.

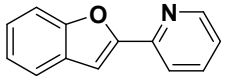


3,5-dimethyl-4-(pyridin-2-yl)isoxazole (3e): colorless liquid.
^{1}H NMR (400 MHz, DMSO-d_6): δ 8.65 (ddd, *J* = 4.9, 1.9, 1.0 Hz, 1H), 7.86 (td, *J* = 7.7, 1.9 Hz, 1H), 7.53 (dt, *J* = 7.9, 1.0 Hz, 1H), 7.32 (ddd, *J* = 7.6, 4.8, 1.1 Hz, 1H), 2.54 (s, 3H), 2.35 (s, 3H).

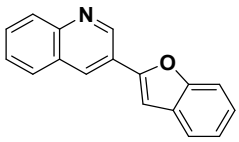
^{13}C NMR (125 MHz, DMSO-d_6): δ 167.7, 158.8, 150.4, 150.1, 137.5, 123.4, 122.5, 115.9, 12.6, 11.7.
LRMS (ESI) m/z: $[M+H]^+$ found 175.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{10}H_{10}N_2O$ 175.0866; found 175.0865.



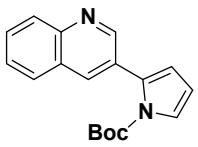
Tert-butyl 2-(pyridin-3-yl)-1*H*-pyrrole-1-carboxylate (3f): colorless liquid.
^{1}H NMR (400 MHz, $CDCl_3$): δ 8.66 – 8.59 (m, 1H), 8.54 (dd, *J* = 4.9, 1.7 Hz, 1H), 7.69 (dt, *J* = 7.8, 1.9 Hz, 1H), 7.39 (dd, *J* = 3.2, 2.0 Hz, 1H), 7.30 (ddd, *J* = 7.9, 4.9, 0.8 Hz, 1H), 6.27 – 6.20 (m, 2H), 1.37 (s, 9H).
^{13}C NMR (125 MHz, $CDCl_3$): δ 149.6, 149.0, 147.8, 137.0, 130.9, 130.7, 123.5, 122.7, 115.8, 110.9, 84.2, 27.7.
LRMS (ESI) m/z: $[M+H]^+$ found 245.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{14}H_{16}N_2O_2$ 245.1285; found 245.1277.



2-(benzofuran-2-yl)pyridine (3g): white solid, m.p. 81 – 83 ºC.
^{1}H NMR (400 MHz, DMSO-d_6): δ 8.67 (dd, *J* = 4.8, 1.6 Hz, 1H), 7.96 – 7.89 (m, 2H), 7.71 (d, *J* = 7.7 Hz, 1H), 7.66 (d, *J* = 8.3 Hz, 1H), 7.56 (s, 1H), 7.42 – 7.35 (m, 2H), 7.28 (t, *J* = 7.5 Hz, 1H).
^{13}C NMR (125 MHz, DMSO-d_6): δ 155.3, 155.1, 150.5, 148.7, 137.8 128.9, 125.9, 124.0, 123.9, 122.3, 120.0, 111.9, 105.3.
LRMS (ESI) m/z: $[M+H]^+$ found 196.1, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{13}H_9NO$ 196.0757; found 196.0756.

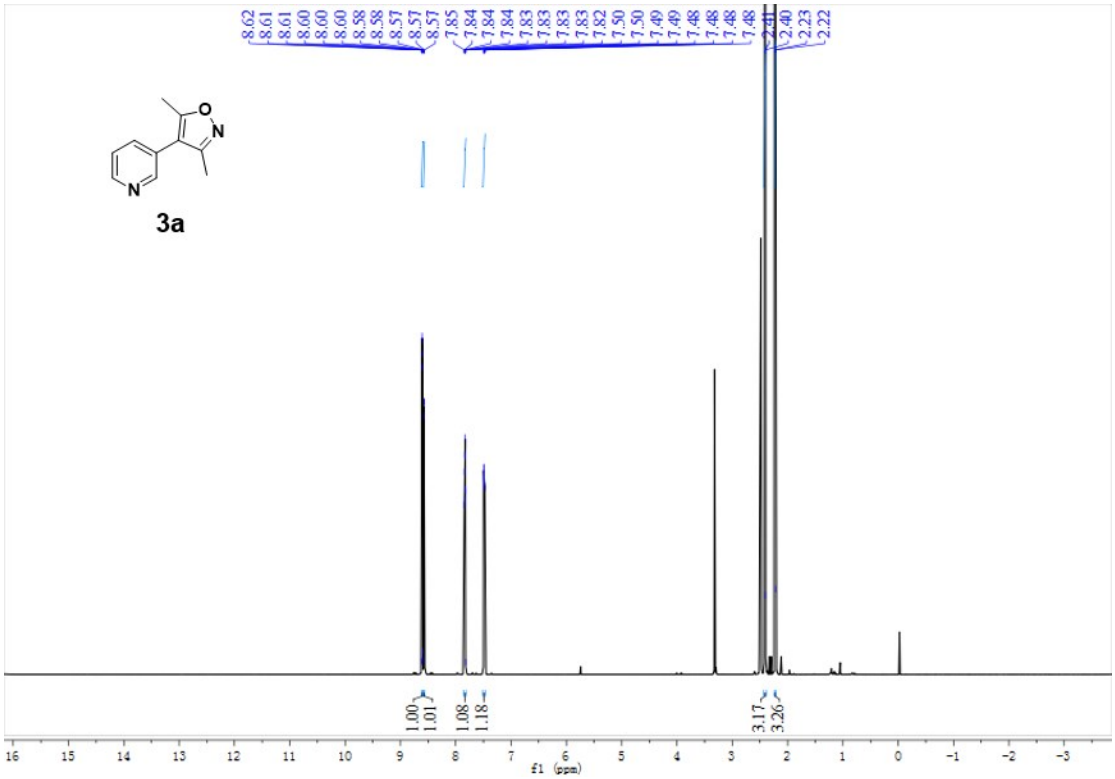


3-(benzofuran-2-yl)quinolone (3h): light yellow solid, m.p. 140 – 142 ºC.
^{1}H NMR (400 MHz, DMSO-d_6): δ 9.46 (d, *J* = 2.3 Hz, 1H), 8.79 (d, *J* = 2.3 Hz, 1H), 8.10 (dd, *J* = 8.2, 1.4 Hz, 1H), 8.06 – 8.02 (m, 1H), 7.77 (ddd, *J* = 8.4, 6.8, 1.4 Hz, 1H), 7.74 – 7.69 (m, 2H), 7.69 – 7.62 (m, 2H).
^{13}C NMR (125 MHz, DMSO-d_6): δ 155.0, 153.3, 148.0, 147.6, 130.9, 130.6, 129.3, 129.1, 128.0, 127.8, 125.7, 124.0, 123.6, 122.0, 111.7, 104.4.
LRMS (ESI) m/z: $[M+H]^+$ found 245.9, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{17}H_{12}NO$ 246.0913; found 246.0917.

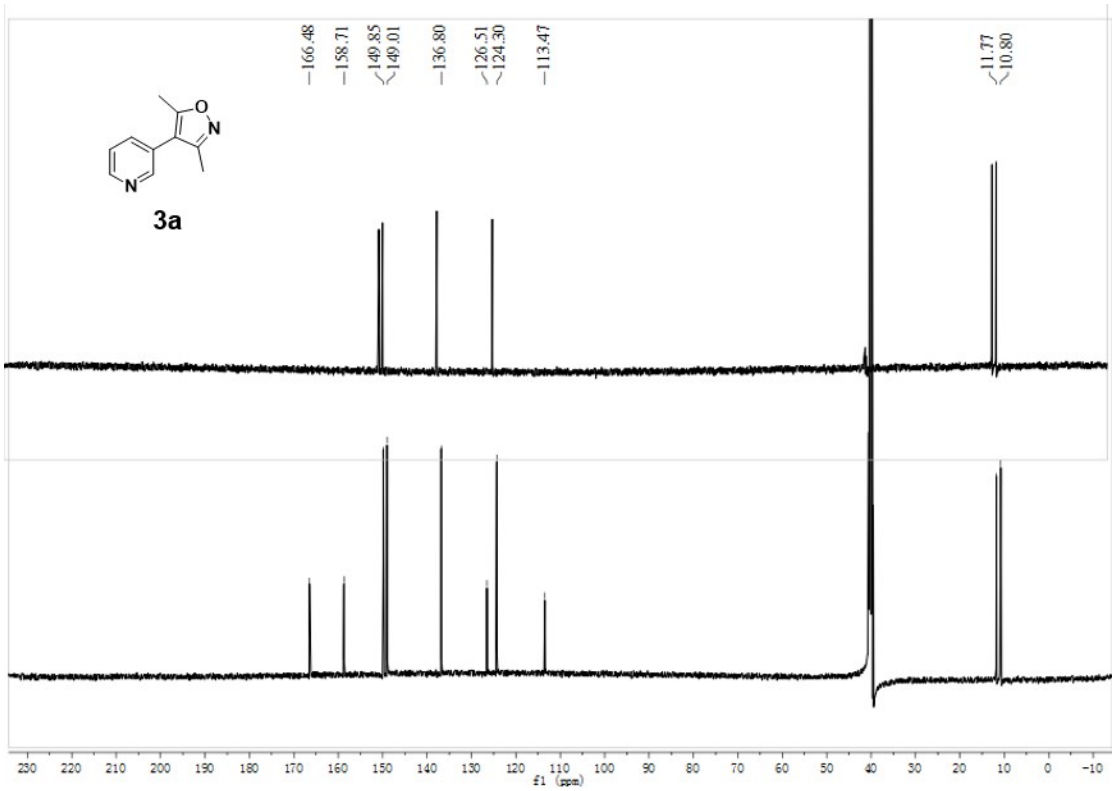


Tert-butyl 2-(quinolin-3-yl)-1*H*-pyrrole-1-carboxylate (3i): colorless liquid.
^{1}H NMR (400 MHz, $CDCl_3$): δ 8.93 – 8.89 (m, 1H), 8.14 – 8.09 (m, 2H), 7.82 (dd, *J* = 8.1, 1.5 Hz, 1H), 7.70 (ddd, *J* = 8.4, 6.9, 1.4 Hz, 1H), 7.56 (ddd, *J* = 8.0, 6.9, 1.1 Hz, 1H), 7.46 (dd, *J* = 3.4, 1.8 Hz, 1H), 6.35 (dd, *J* = 3.3, 1.8 Hz, 1H), 6.31 (t, *J* = 3.3 Hz, 1H), 1.35 (s, 9H).
^{13}C NMR (125 MHz, $CDCl_3$): δ 151.6, 149.1, 146.6, 134.6, 131.4, 129.4, 129.1, 127.8, 127.5, 126.9, 123.4, 116.1, 111.1, 84.4, 27.7.
LRMS (ESI) m/z: $[M+H]^+$ found 295.0, HRMS (ESI) m/z: $[M+H]^+$ calcd for $C_{18}H_{18}N_2O_2$ 295.1441; found 295.1442.

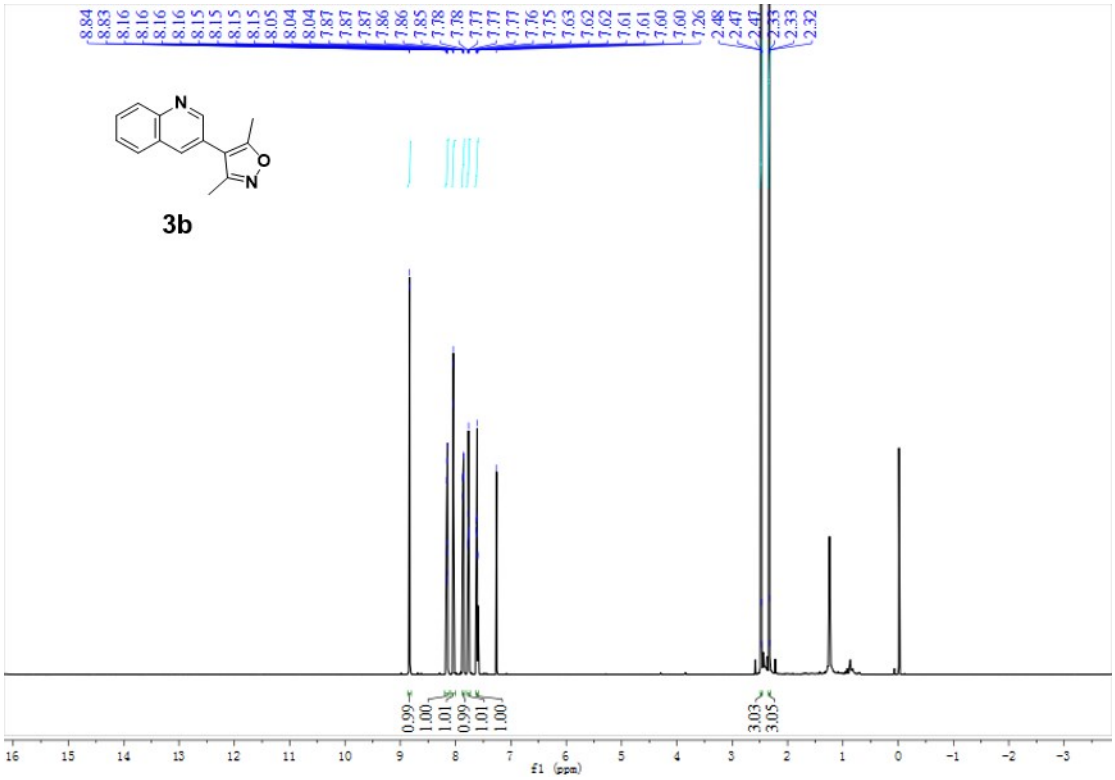
^{1}H NMR, 400 MHz, DMSO-d_6



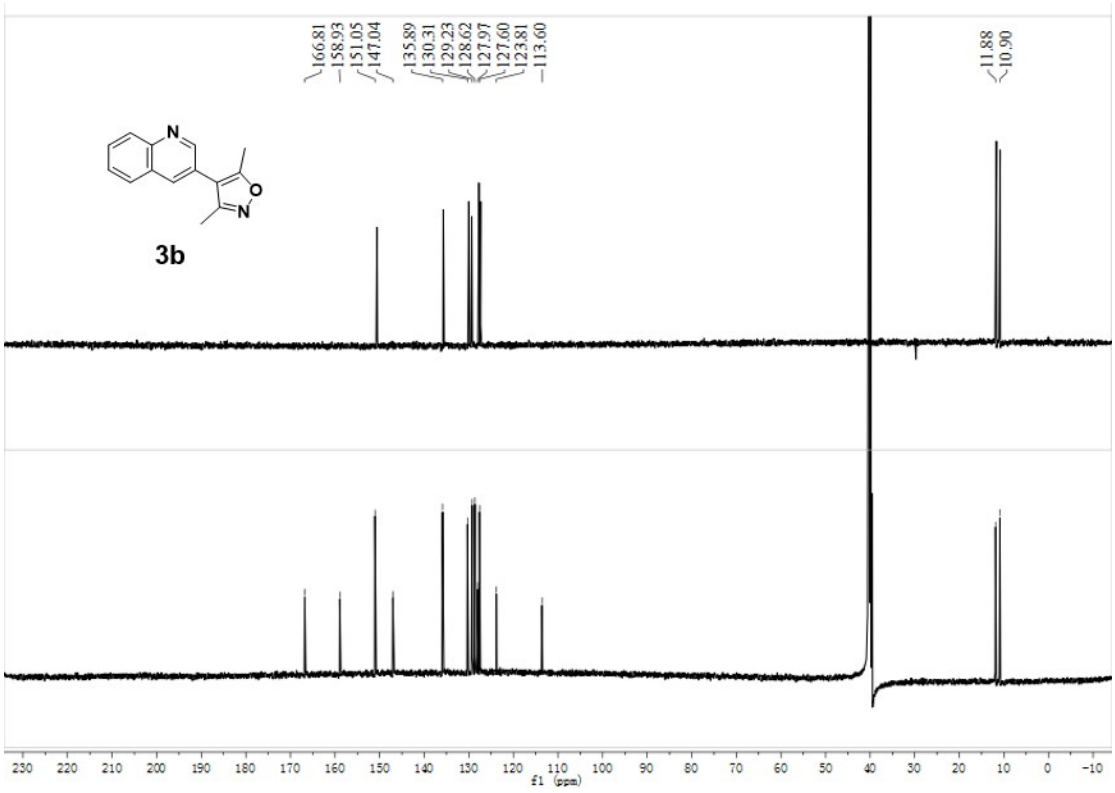
^{13}C NMR, 125 MHz, DMSO-d_6



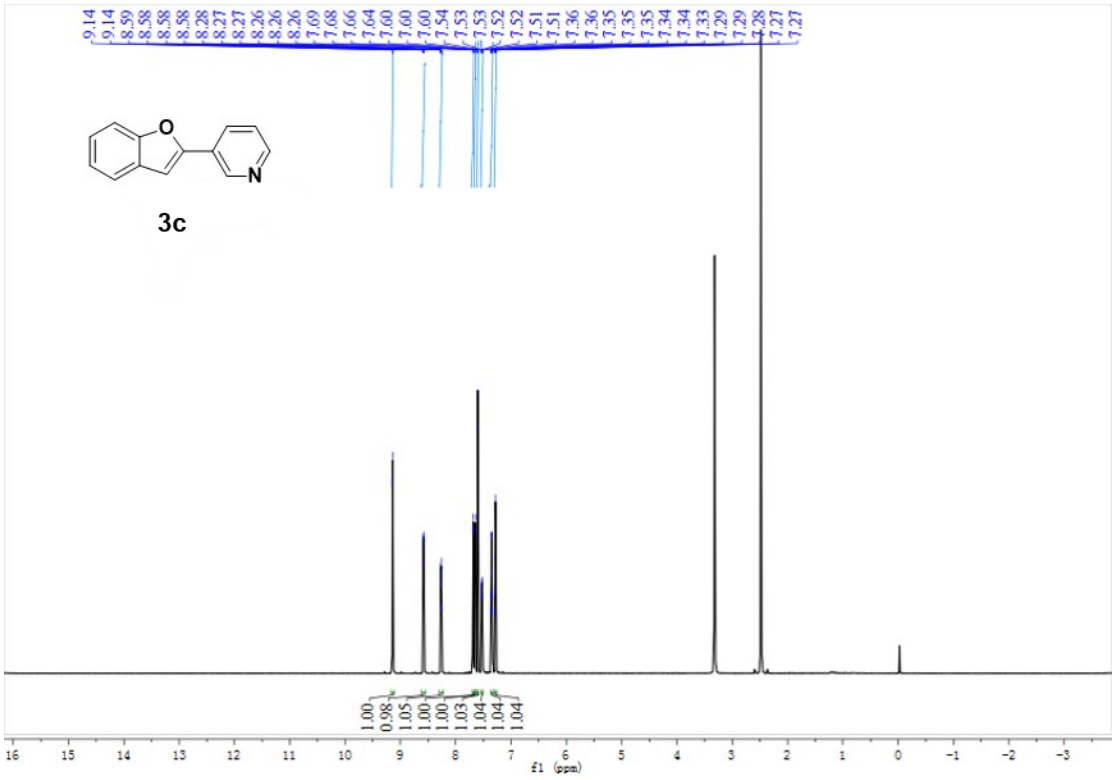
^{1}H NMR, 400 MHz, $CDCl_3$



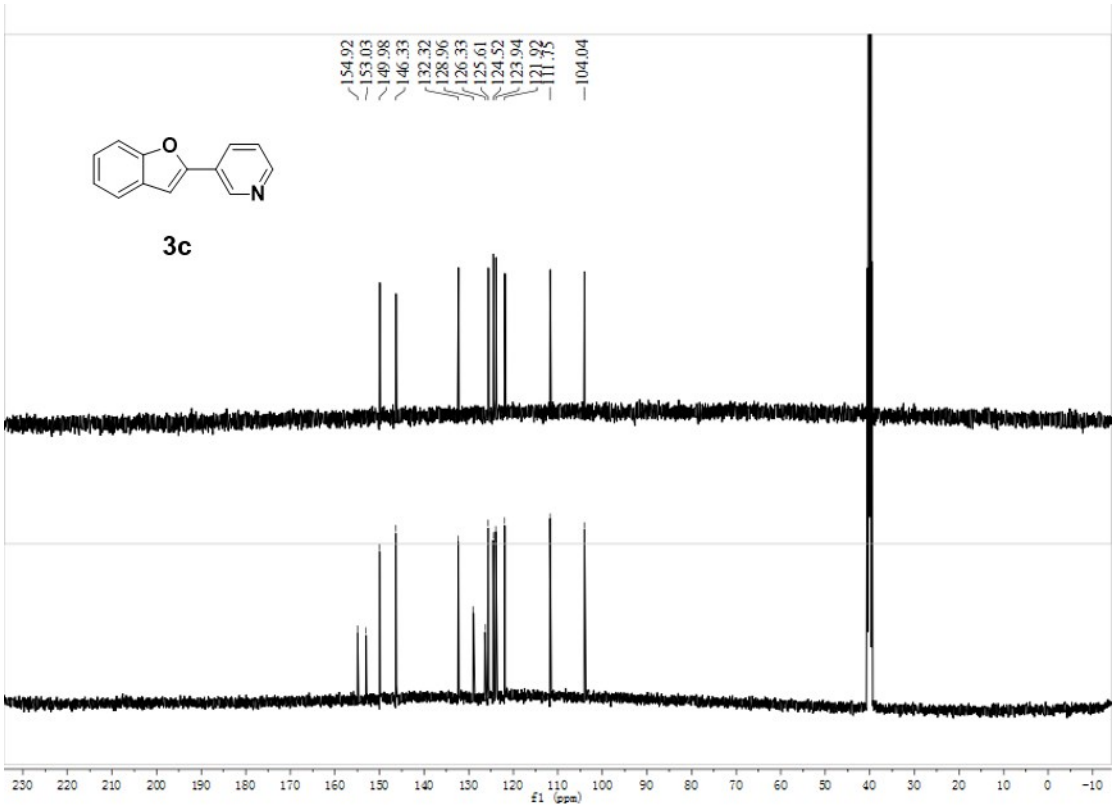
^{13}C NMR, 125 MHz, DMSO-d_6



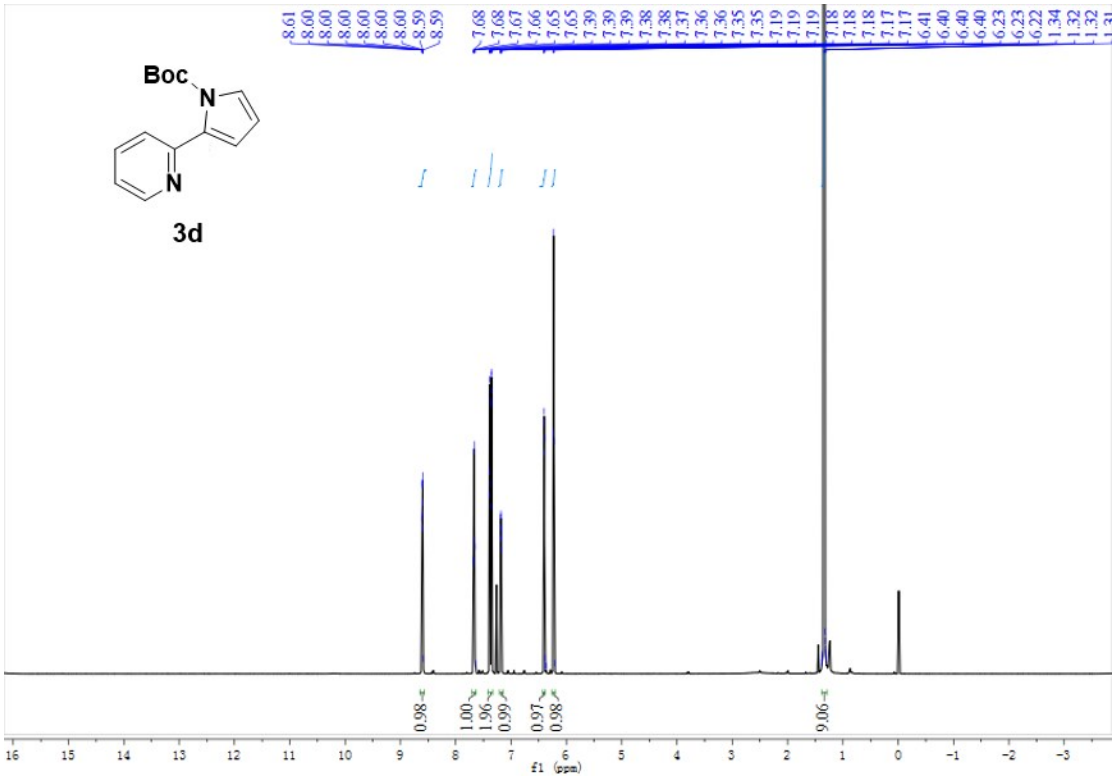
^{1}H NMR, 400 MHz, DMSO-d_6



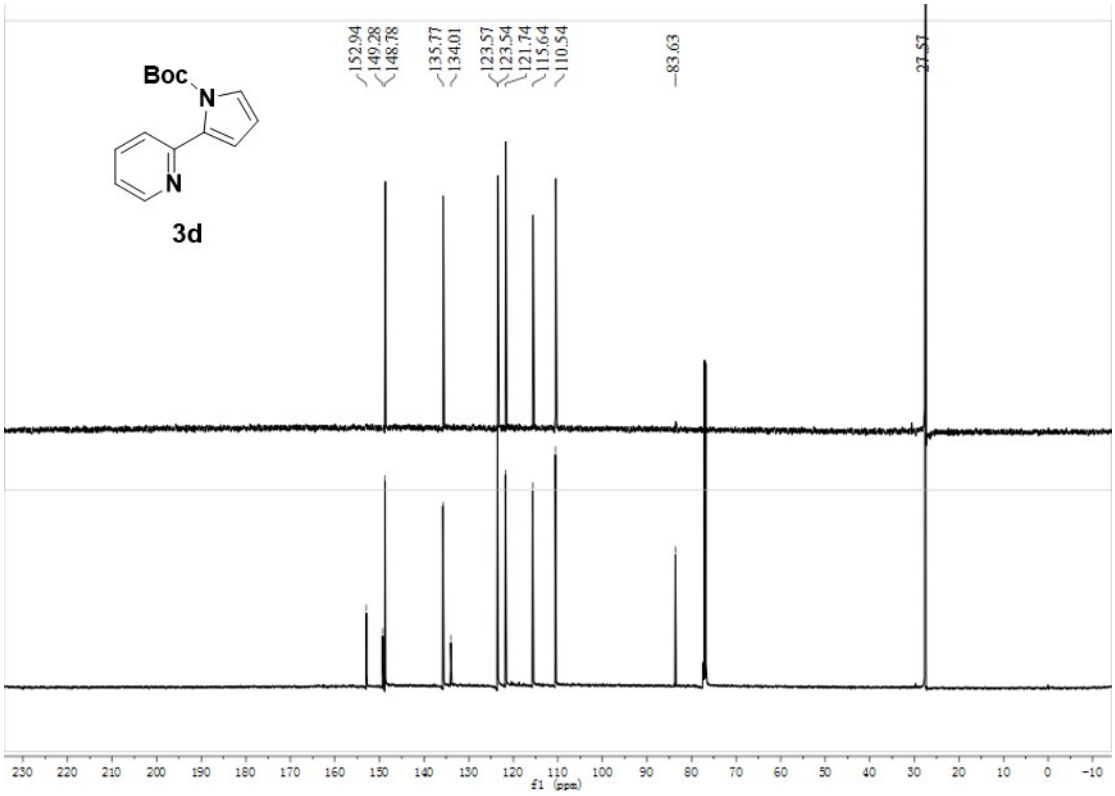
^{13}C NMR, 125 MHz, DMSO-d_6



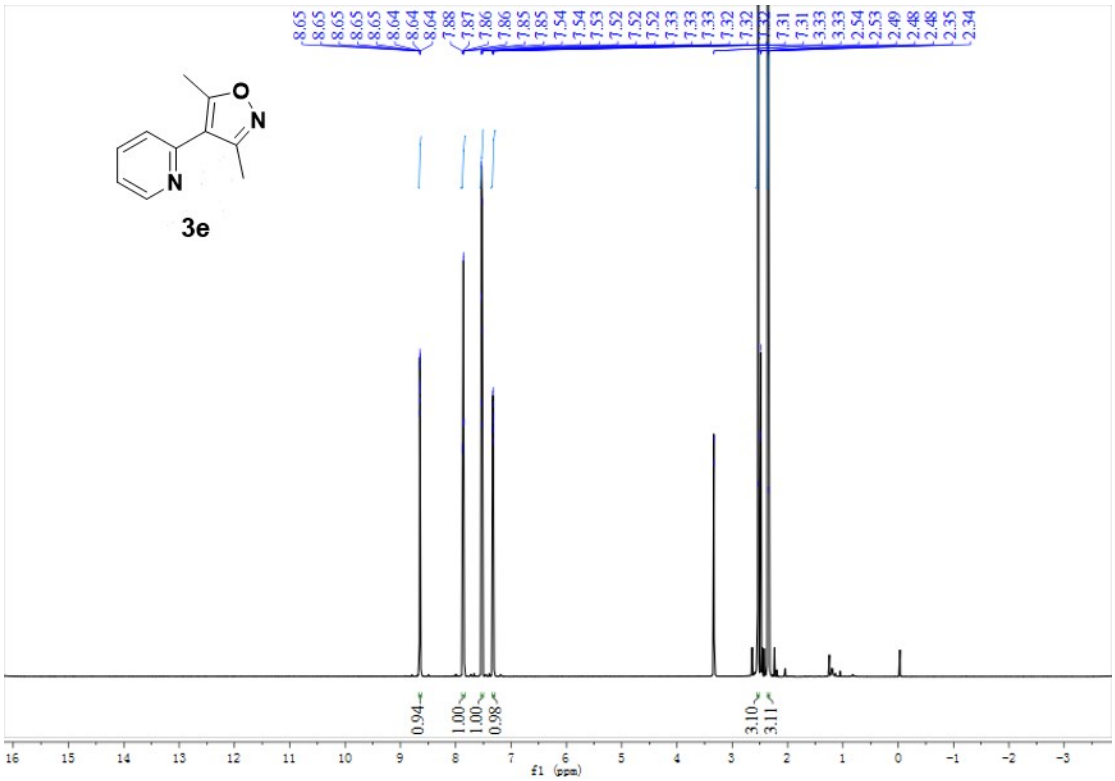
^{1}H NMR, 400 MHz, $CDCl_3$



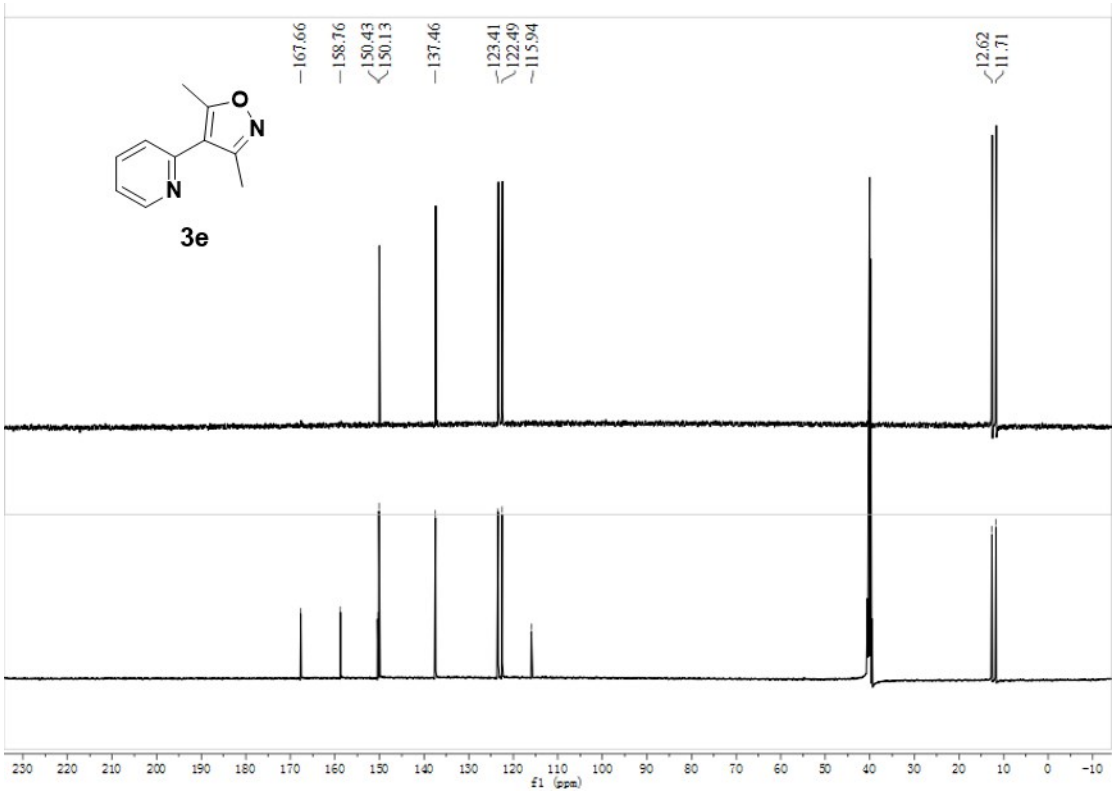
^{13}C NMR, 125 MHz, $CDCl_3$



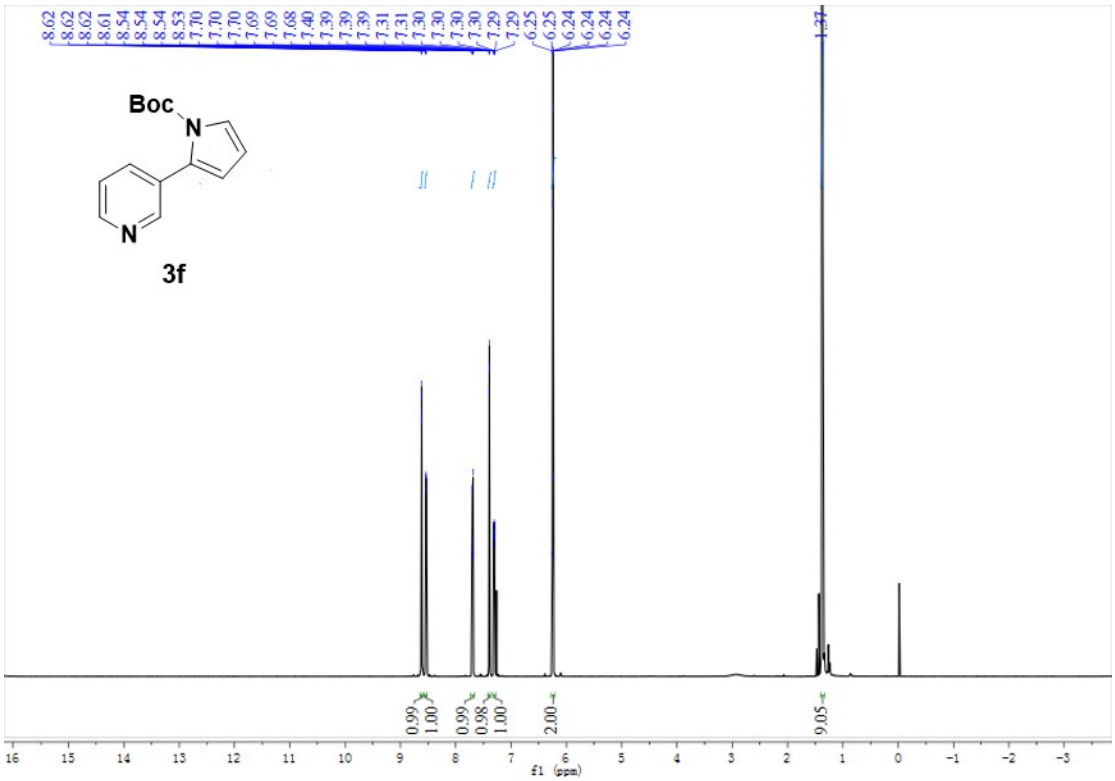
^{1}H NMR, 400 MHz, DMSO-d_6



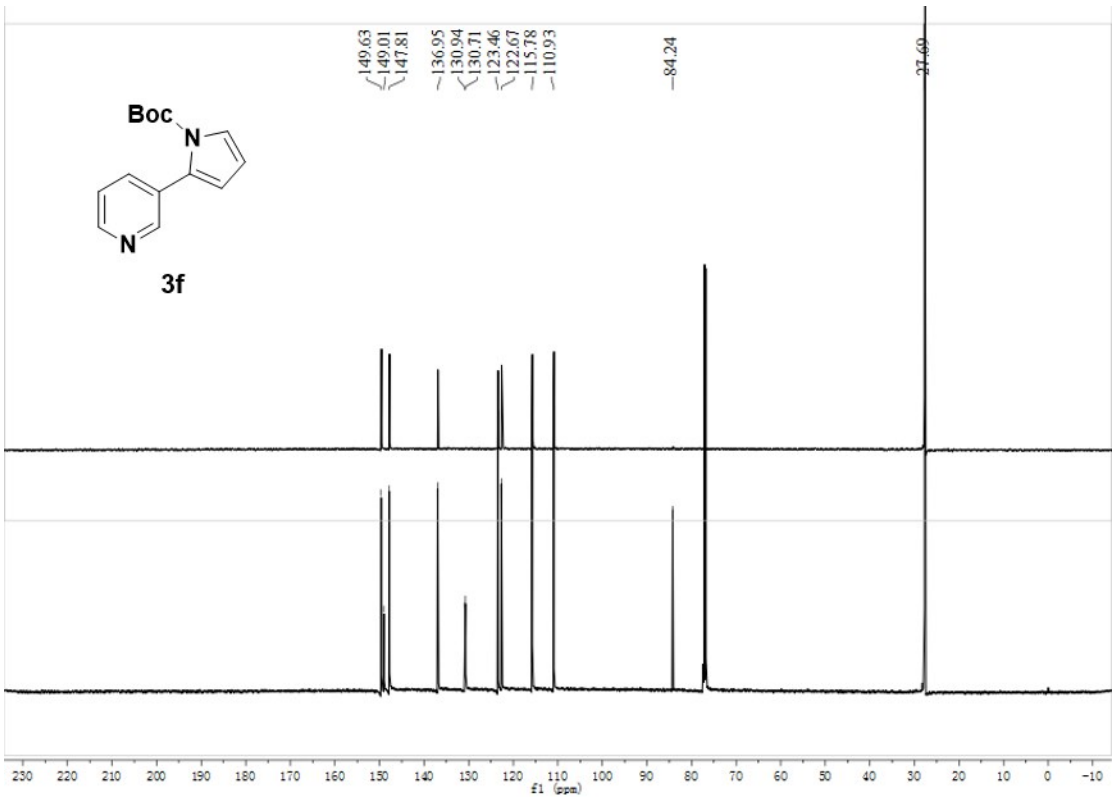
^{13}C NMR, 125 MHz, DMSO-d_6



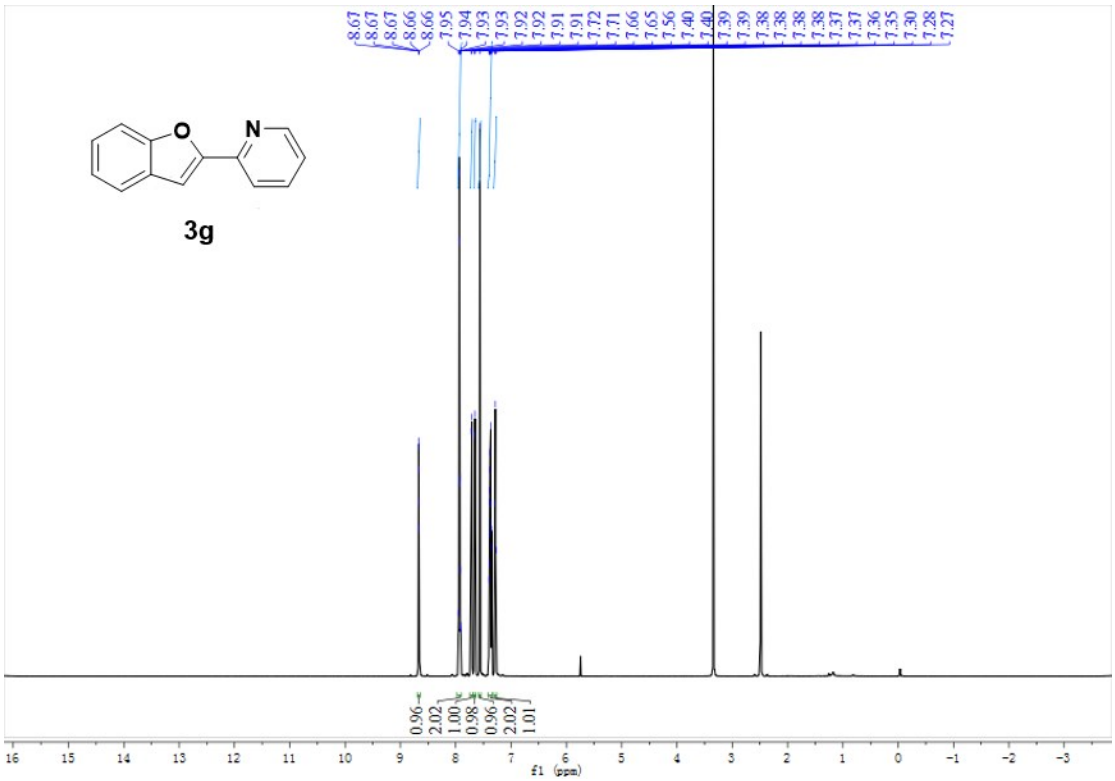
^{1}H NMR, 400 MHz, $CDCl_3$



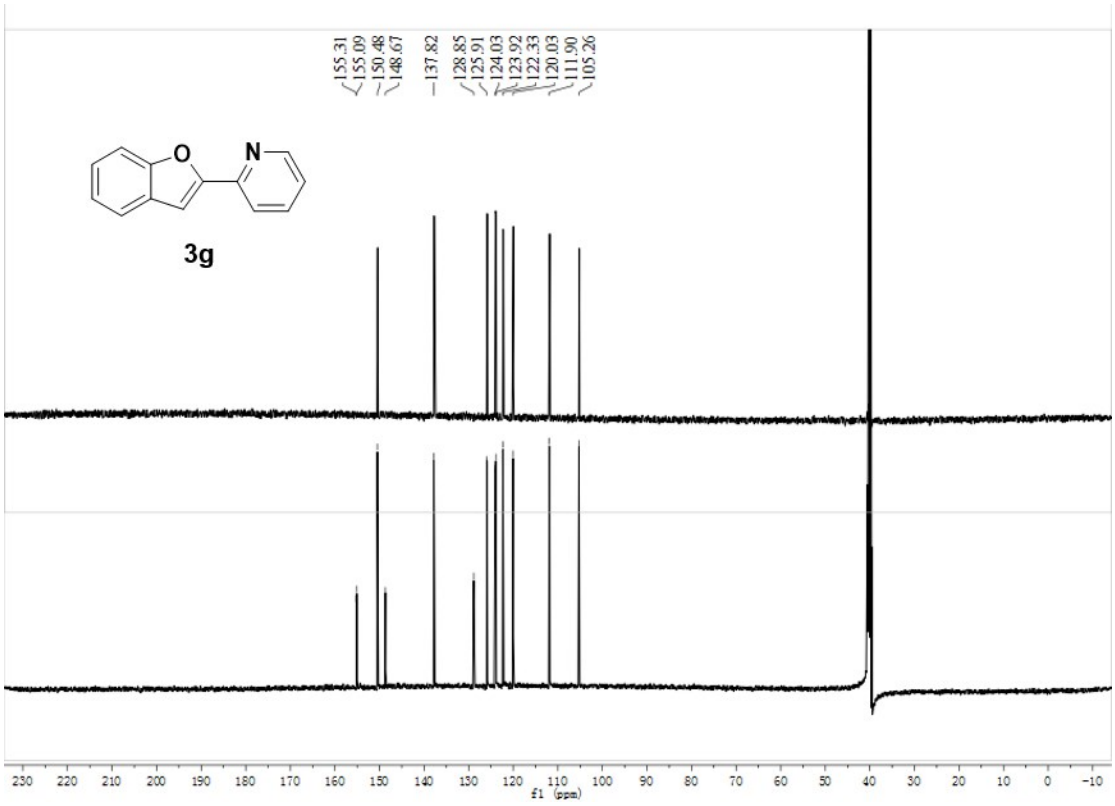
^{13}C NMR, 125 MHz, $CDCl_3$



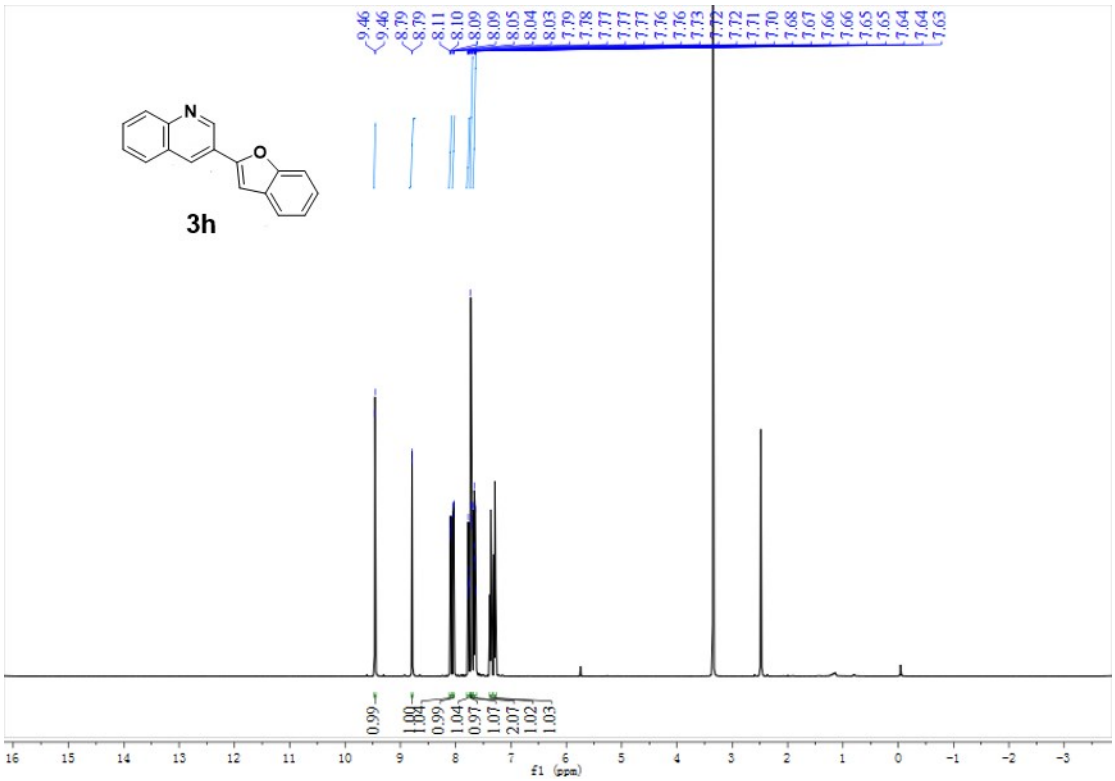
^{1}H NMR, 400 MHz, DMSO-d_6



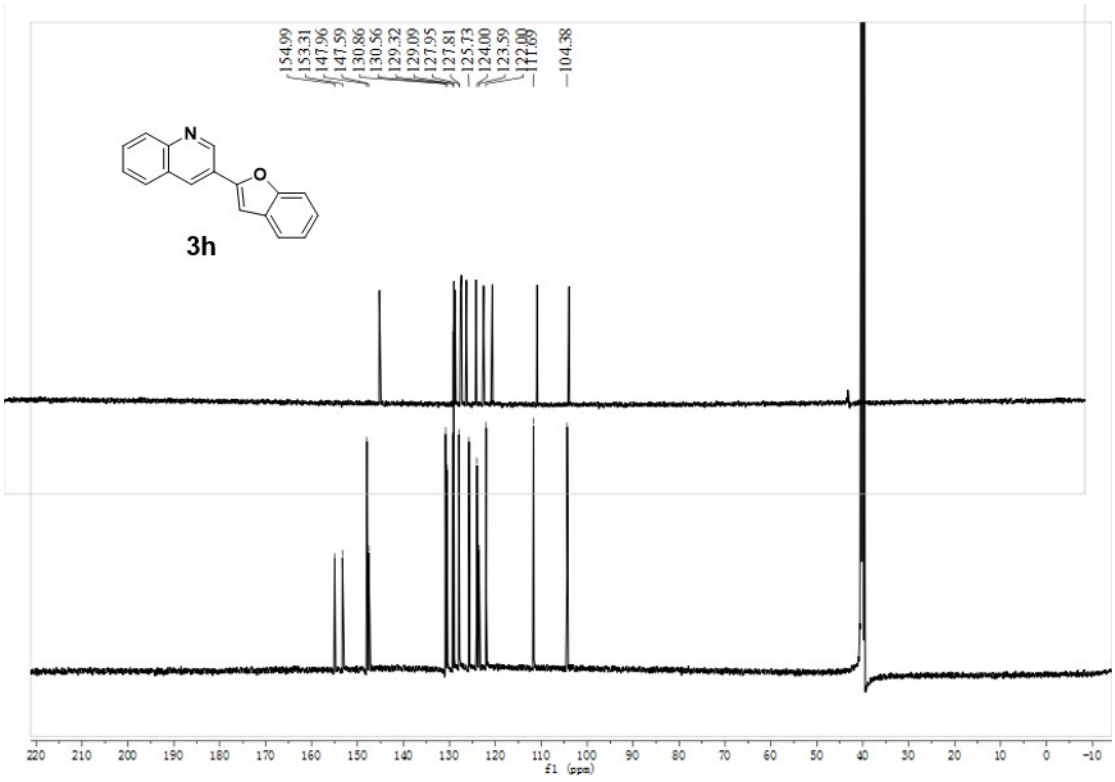
^{13}C NMR, 125 MHz, DMSO-d_6



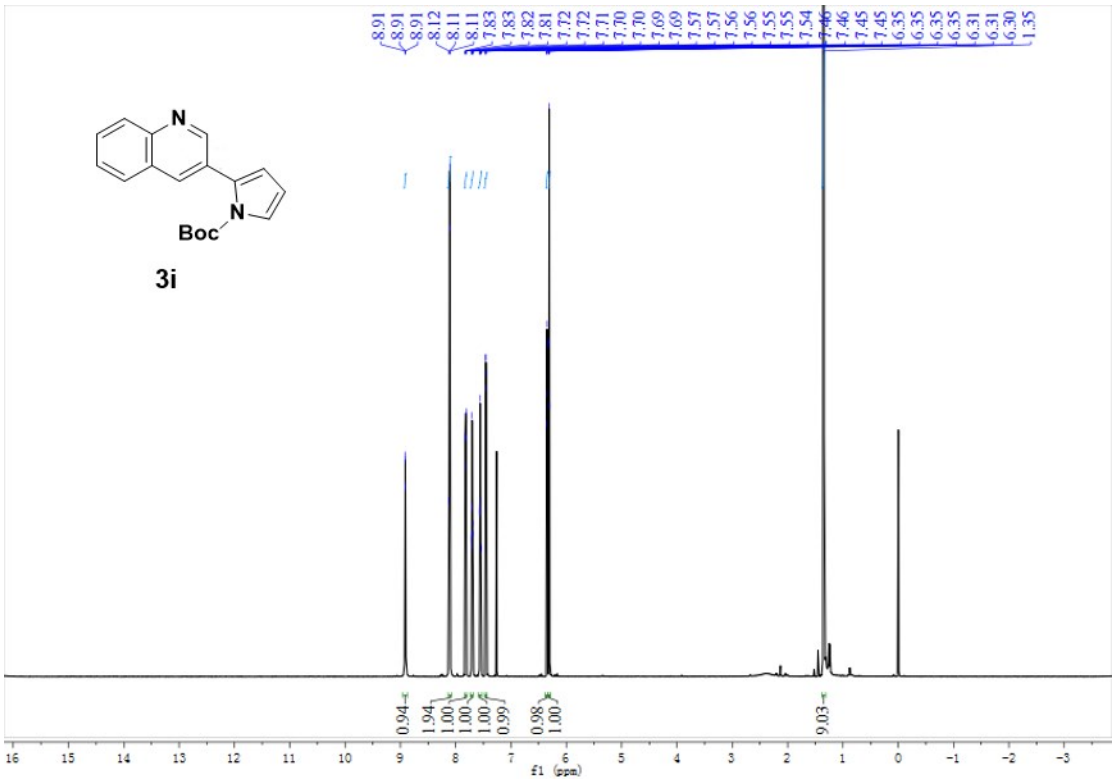
^{1}H NMR, 400 MHz, DMSO-d_6



^{13}C NMR, 125 MHz, DMSO-d_6



^{1}H NMR, 400 MHz, $CDCl_3$



^{13}C NMR, 125 MHz, $CDCl_3$