1-1-2019

Towards an Accurate Prediction of Nitrogen Chemical Shifts by Density Functional Theory and Gauge-Including Atomic Orbital

Peng Gao
University of Wollongong, pg177@uowmail.edu.au

Xingyong Wang
University of Wollongong, xingyong@uow.edu.au

Haibo Yu
University of Wollongong, hyu@uow.edu.au

Towards an Accurate Prediction of Nitrogen Chemical Shifts by Density Functional Theory and Gauge-Including Atomic Orbital

Abstract

An effcient, yet accurate, computational protocol for predicting nitrogen NMR chemical shifts based on density functional theory and the gauge-including atomic orbital approach has been proposed. A database of small and relatively rigid compounds containing nitrogen atoms was compiled. Scaling factors for the linear correlation between experimental ^{15}N chemical shifts and calculated isotropic shielding constants have been systematically investigated with seven different levels of theory in both chloroform and dimethyl sulfoxide, two commonly used solvents for NMR experiments. The best method yields a root-mean-square deviation of about 5.30 ppm and 7.00 ppm in $CHCl_3$ and DMSO, respectively. Moreover, another set of scaling factors for $-NH_2$ chemical shifts was also proposed based on a separate database with three levels of theory. Furthermore, it is encouraging that a reasonable transferability for the linear correlation has been found between these two solvents. This finding will enable broader applications of the developed empirical scaling factors to other commonly used solvents in NMR experiments. The consistency between theoretical predictions and experimental results for structural elucidations was illustrated for selected examples including regioisomers, tautomers, oxidation states, and protonated structures.

Disciplines

Medicine and Health Sciences

Publication Details

Gao, P., Wang, X. & Yu, H. (2019). Towards an Accurate Prediction of Nitrogen Chemical Shifts by Density Functional Theory and Gauge-Including Atomic Orbital. Advanced Theory and Simulations, 2 (2), 1800148-1-1800148-8.

Towards an Accurate Prediction of Nitrogen Chemical Shifts by Density Functional Theory and Gauge-Including Atomic Orbital

Peng Gao, Xingyong Wang, Haibo Yu*

October 25, 2018

Peng Gao, Dr Xingyong Wang, Dr Haibo Yu
School of Chemistry and Molecular Bioscience, University of Wollongong, NSW 2522, Australia
Molecular Horizons, University of Wollongong, NSW 2522, Australia
* Corresponding author: hyu@uow.edu.au

Dr Haibo Yu
Illawarra Health and Medical Research Institute, Northfields Avenue, Wollongong NSW 2522, Australia

Keywords

Nuclear Magnetic Resonance Spectroscopy, Density Functional Theory, Gauge-Including Atomic Orbital, Nitrogen Chemical Shift, Linear Regression, Solvent Effects

Abstract

An efficient, yet accurate, computational protocol for predicting nitrogen NMR chemical shifts based on density functional theory and the gauge-including atomic orbital approach has been proposed. A database of small and relatively rigid compounds containing nitrogen atoms was compiled. Scaling factors for the linear correlation between experimental ^{15}N chemical shifts and calculated isotropic shielding constants have been systematically investigated with seven different levels of theory in both chloroform and dimethyl sulfoxide, two commonly used solvents for NMR experiments. The best method yields a root-mean-square deviation of about 5.30 ppm and 7.00 ppm in $CHCl_3$ and DMSO, respectively. Moreover, another set of scaling factors for -NH_2 chemical shifts was also proposed based on a separate database with three levels of theory. Furthermore, it is encouraging that a reasonable transferability for the linear correlation has been found between these two solvents. This finding will enable broader applications of the developed empirical scaling factors to other commonly used solvents in NMR experiments. The consistency between theoretical predictions and experimental results for structural elucidations was illustrated for selected examples including regioisomers, tautomers, oxidation states, and protonated structures.

1 Introduction

Theoretical predictions of nuclear magnetic resonance (NMR) chemical shifts have found an increasing number of applications for structural elucidation and mechanistic studies in modern chemistry research.[1] Development of accurate yet practically affordable computational methods is crucial to improve the reliability and accuracy of such predictions and can help to narrow the structural possibilities. Though quantum mechanical calculations of chemical shifts and coupling constants date back as early as to 1950s by Ramsey,[2] routine calculations of isotropic shielding constants and chemical shifts have been made practically accessible due to methodology developments, especially the introduction of the gauge-including atomic orbital (GIAO) approach.[3] Various methodological developments aiming at reducing errors have been carried out. These include accounting for electron correlations,[4] accurate modeling of solvation effects,[5] conformational averaging,[6] vibrational averaging,[7] heavy atom effects,[8] linear regression[1,9] and empirically parameterized quantum mechanical models.[10] Among them, the linear regression method, namely the application of corrections derived from linear regression procedures, is arguably the most general and straightforward approach for error reduction. It aims at achieving high accuracy with low to moderate computational costs by applying an empirical scaling to minimize the systematic errors in the adopted models. The success of such empirical scaling is demonstrated by the superior performance of predictions of chemical shifts for ^{1}H and ^{13}C, which has been recently reviewed by Lodewyk *et al.*[1] It has been noted that the major benefit for this empirical scaling is that the slope can be used as an empirical correction to correct the computed chemical shifts for systematic errors. Such a procedure can reduce error from various sources such as solvation effects, rovibratory effects, and other methodological limitations. The intercept values provide a convenient alternative for a reference value from which the calculated isotropic values can be converted to chemical shifts. Therefore, via the approach of linear regression, the obtained prediction value can not only be scaled to decrease the systematic errors but can also avoid the specific errors associated with certain reference compounds.

It has been suggested that the high sensitivity of the nitrogen lone pair to changes in the molecular environment makes ^{15}N NMR an exceptionally useful tool.[11–13] However, com-

pared to ^{1}H and ^{13}C, the range of ^{15}N chemical shifts is much larger (up to 1200 ppm), making it very challenging to develop an empirical scaling that works satisfactorily across all the ^{15}N chemical shift range. Moreover, oxidation states, conformational variations and protonation states pose additional challenges for theoretical predictions of ^{15}N chemical shifts.

In the current work, we aim at developing efficient protocols for computational prediction of ^{15}N chemical shifts, in order to further improve the accuracy of structural elucidation. We compiled a database with small and relatively rigid molecules covering most common nitrogen-containing functional groups and developed empirical scaling parameters for linear regression approaches with seven levels of theory. As a complement to previous experimental and computational work on ^{1}H and ^{13}C,[1,14,15] our protocol will assist in the full realization of predictions of chemical shifts in structural elucidations of many important natural products and pharmaceutical related NMR compounds. Furthermore, to the best of our knowledge, the transferability for the linear regression models between different solvents was investigated for the first time. The encouraging results enable a broader application of the developed protocol for other commonly used NMR solvents.

2 Results and Discussions

2.1 Performance of the adopted methods

The fitted empirical scaling factors are listed in Table 1, and their performance is summarised in Table 2. For all seven methods in the two solvents, R^2s were estimated to be close to 1.0 (Table 1) indicating that the developed linear regression model fits well with the applied data. On the other hand, some systematic errors were observed as indicated by the deviation of slope from the ideal values (up to 4%–5% in both $CHCl_3$ and DMSO). For the test sets, all seven methods performed reasonably well and they predicted ^{15}N chemical shifts with root-mean-square deviations (RMSDs) of 5.30–6.67 ppm and 7.00–8.04 ppm to the experimental values in $CHCl_3$ and DMSO with relative RMSDs around 5% and 7%, respectively (Table 2, and the linear regression figures can be found in ESI Figure S1 and S2). When the fitted empirical scaling factors were applied to the probe sets (Table 2), similar performances were

observed for both solvents.

We excluded the -NH_2 group containing molecules with the corresponding chemical shifts larger than 300 ppm from the linear regression. For this particular group, a significant deviation between predicted and experimental values was noted, as reported previously by Xin *et al.*[16] To address this issue, we built another separate database of 12 -NH_2 containing molecules, and based on this database new scaling factors specifically for -NH_2 were fitted for Method 5, 6 and 7 (see Figure S3 and Table S4 in ESI for more details). These fitted scaling factors gave an RMSD of 1.18 to 1.99 ppm.

Based on our data, Methods 5 and 6 with geometry optimization at M062X/6-31+G(d,p) or M062X/6-31+G(2d,p) in the gas phase and NMR GIAO calculations at mPW1PW91/6-311+G(2d,p) with the SMD solvent model and Method 7 with geometry optimization and NMR GIAO calculation at B3LYP/cc-pVDZ with the CPCM solvent model provide the best performance for predictions. The advantage of Methods 5 and 6 lies in providing the chemical shift prediction for ^{1}H, ^{13}C and ^{13}N in one set of calculations taking advantage of previous work.[17]

We also note that the fitted empirical scaling factors provide a better prediction for ^{15}N chemical shifts in $CHCl_3$ than those in DMSO in terms of RMSDs (Table 2). Two possible factors might contribute to the difference in their performances. Firstly, the chemical shift range covered in DMSO (∼-284 ppm to ∼35 ppm) is larger than that in $CHCl_3$ (∼-217 ppm to ∼36 ppm). Thus, the absolute RMSDs in DMSO is larger than those in $CHCl_3$. This has been observed in previous studies on ^{1}H and ^{13}C chemical shifts.[17] Secondly, the dielectric constant of DMSO (ϵ=45.80) is much larger than that of $CHCl_3$ (ϵ=4.55), and it is expected that DMSO might have a larger effect on the molecular structures[18] compared to those *in vacuo*. To further improve the prediction accuracy, including solvation effects in the geometric optimization may be helpful. However, the corresponding computational time will be increased significantly.

2.2 Transferability for scaling factors between $CHCl_3$ and DMSO

Experimentally, more than a dozen of solvents have been used in ^{15}N NMR measurements with a dielectric constant from 1.87 (cyclohexane) to 76.70 (water). It has been shown that

the solvent may significantly alter the observed shifts.[19] One obvious question is whether there is any transferability for the fitted scaling factors between different solvents with the same level of theory for calculation. Can the fitted empirical scaling factors in one solvent apply to other solvents without re-parameterization? This has been investigated by swapping the empirical scaling factors between $CHCl_3$ and DMSO and quantified their performances in terms of RMSDs (Table 3). For Method 5, the RMSDs change from 5.30 to 8.19 ppm for $CHCl_3$ and from 7.26 to 9.22 ppm for DMSO. For Method 6, the RMSDs increase from 5.35 to 8.03 ppm for $CHCl_3$ and from 7.12 to 8.91 ppm for DMSO. Based on such analysis on two solvents, we predict that these empirical scaling factors can achieve an RMSD of less than $\sim$10 ppm in other solvents; however, more accurate predictions will require solvent specific scaling factors. Alternatively, we also studied the transferability for all the molecules that have experimental data in both $CHCl_3$ and DMSO. Similar performances were observed for Method 5–7. More details about prediction errors can be seen in ESI Table S5–S8. We recommend using scaling factors of DMSO set for predictions of ^{15}N NMR chemical shift in polar solvents, and scaling factors of $CHCl_3$ set for predictions in non-polar solvents.

2.3 Applications of ^{15}N Chemical Shifts in Structural Elucidations

We have applied the developed protocol to illustrate the potential applications of ^{15}N chemical shifts in structural elucidations for selected case studies.

2.3.1 Regioisomers

The distinction between different regioisomers can be made by NMR chemical shift analyses. In order to differentiate various heterocycles containing molecules, current applications of ^{1}H and ^{13}C NMR spectra are sometimes insufficient[20], especially for small molecules, whose structures are difficult to ascertain. The broader range of nitrogen chemical shifts enables a more sensitive analysis than ^{1}H and ^{13}C chemical shifts analyses. Investigation of the regiochemistry of oxazole based molecules has drawn considerable attention due to their wide application as bioactive natural products.[21–23] At the same time, it is challenging for ^{15}N NMR predictions without linear regression, as the corresponding error will be significantly

larger. We applied our protocol to oxazoles, isoxazoles, and oxadiazoles (Figure 1 and Table 4), which have been systematically investigated recently by Xin *et al.*.[16] For the molecules listed, we carried out our calculations using Method 5-7 with the same procedure as above and only considered the conformer with the lowest energy obtained through the geometry optimization step in the gas phase (the same procedure was applied for **5**, **6** and **7**, discussed in next sections). For oxazole **1**, all three methods are able to predict the correct isomers through providing chemical shifts comparable with experimental data (most of the errors are within 1σ or at most 2σ). For isoxazoles (**2** and **3**), all three methods can distinguish the correct structure. What needs to be highlighted is that the substantial systematic error for the chemical shifts of the -NH_2 group has largely been reduced. As can be seen from Table 4, at Pos.2 of isoxazoles (**2**) via the new scaling factors, the deviations are merely 2.0 ppm, 2.2ppm and 1.0 ppm for Method 5, 6 and 7, respectively. For the molecules containing halogen atoms (**3** and **4**), there is an effect of spin-orbit coupling on halogen-substituted carbon atoms. Having no direct bond with halogen atoms ^{15}N NMR chemical shift prediction thus has an advantage as this effect will have a minor impact on ^{15}N NMR chemical shifts. Therefore, ^{15}N NMR chemical shift calculation can be an excellent complement to ^{1}H and ^{13}C NMR chemical shifts studies in structural elucidation. For oxadiazole **4**, all three methods provided robust results for structural differentiation among possible isomers.

2.3.2 Tautomers

We also applied our protocol to the structural elucidation of a tautomeric system - adenine (Figure 2). Previous studies have concluded that tautomer **5a** is favored over **5b** (Figure 2 and Table 5). Our calculations predicted that the deviations of ^{15}N NMR chemical shift for **5a** are all within 8.5 ppm with Method 5, and 7.0 ppm with Method 6. For Pos.3 of tautomer **5a**, the deviations are 1.2 ppm, 1.6 ppm and 0.3 ppm obtained with the new scaling factors based on Method 5, 6 and 7. A slightly larger deviation was observed for Pos.5 with Method 7 for tautomer **5a**. For the case of tautomer **5b**, an apparent deviation larger than 20 ppm was observed at Pos.5 with all three methods. Thus, we conclude that tautomer **5a** is favored over **5b** in DMSO, which is also consistent with previous studies.[16,24]

2.3.3 Oxidation states

For structural elucidation of nitrogen oxidation states, we tested our methods with **6** (Figure 3) - a second generation HIV-1 non-nucleoside reverse transcriptase inhibitor.[25] As we can see from Table 6, for **6a**, both Method 5 and 6 performed well. The predicted ^{15}N NMR chemical shifts are consistent with the experimental data, with deviations less than 9.5 ppm, with somewhat slightly larger deviations observed with Method 7. While for the other two possible nitrogen oxidation sites **6b**, **6c**, and the non-oxidized compound **6d**, the corresponding deviations of ^{15}N NMR chemical shift at Pos.5 are larger than 17.9 ppm. From the averaged deviation (Table 6), we can see that for **6a** the overall deviations for all three methods are within 1σ. Therefore, we can conclude that **6a** is the most likely oxidation state of this compound and that the most favorable nitrogen oxidation site is located at Pos.5.

2.3.4 Protonation states

For structural elucidation of different nitrogen protonation states, we tested our methods with molecule **7**, which belongs to the pyrrole-imidazole alkaloids family. **7a**, **7b**, and **7c** are the alternative tautomers, and **7d** is the protonated state of **7a** (Figure 4). As we can see from Table 7, for the protonated structure **7d**, all three methods performed well. The predicted ^{15}N NMR chemical shifts are in good agreement with the experimental data[26]. For **7a**, **7b**, and **7c**, there exist significant deviations up to 80 ppm, indicating that a protonated structure is the dominant population under the NMR experimental condition.[16,26] This was also supported by the averaged deviations for Method 5, 6 and 7, which are all less than 4.2 ppm (Table 7).

3 Conclusions

Linear regression parameters mapping calculated isotropic shielding constants based on seven different levels of theory to observable NMR chemical shifts have been developed for ^{15}N. The procedure proposed by Tantillo and co-workers[1,17] for ^{1}H and ^{13}C chemical shifts predictions had been proved useful for ^{15}N chemical shift predictions. Among the seven levels of theory in this study, Method 5 and 6 with geometry optimization at M062X/6-31+G(d,p) or M062X/6-

31+G(2d,p) and NMR GIAO calculation at mPW1PW91/6-311+G(2d,p) with SMD solvent model, provide reliable predictions. Moreover, the fitted scaling factors for -NH_2 chemical shifts based on Method 5 and 6 provide an improved estimation for this group of compounds. For general applications, we recommend the use of the linear regression parameters of Method 5 and 6 to predict ^{15}N chemical shifts together with those for ^{1}H and ^{13}C. One interesting observation is that the linear regression parameters can provide a reasonable prediction for the chemical shifts in a solvent for which they were not parameterized. Alternatively, one can choose the set of linear regression parameters based on the polarity of the solvent. Future work will include solvent specific parameterization for other commonly used NMR solvents and potential further refinement by taking into account solvation effects during the geometry optimization step. We expect that the current work on ^{15}N chemical shift and together with previous work on ^{1}H and ^{13}C chemical shifts prediction will serve as a robust tool in structural elucidation.

4 Experimental Section

4.1 Database of nitrogen-containing molecules

In total, 24 and 46 molecules were included in our test sets for $CHCl_3$ and DMSO, respectively (see Table S1 and S2 in ESI for more details). Two criteria for selection of molecules include a) reliable experimental data available, preferably in both $CHCl_3$ and DMSO; b) being small and relatively rigid (*i.e.*, with only one dominant potential energy minimum). The reason for the second selection criterion is to avoid dealing with conformational averaging in isotropic shielding constant calculations during the fitting process. Theoretically, for flexible molecules with multiple minima, the isotropic constant (σ) can be calculated from those of individual conformers via the Boltzmann averaging of the isotropic shielding constants with respect to the relative potential energies for different conformers.[6] However, the potential errors in the description of the energetic properties with a particular adopted quantum mechanical method (including the implicit solvent model) might contaminate the fitting process. Instead, in this work, we focused on small and rigid molecules to avoid such an

averaging. To study the performance of our fitted linear regressions, we also compiled a probe set with 12 rigid molecules (see Table S3 in ESI for more details) for both $CHCl_3$ and DMSO sets. Additionally, to address the challenge posed by the predictions of various -NH_2 chemical shifts, a database of 12 -NH_2 containing rigid molecules (see Table S4 in ESI for more details) was also built.

4.2 Computational details

We followed the procedures proposed by Tantillo and co-workers.[1,17] First, geometry optimization calculations were carried out *in vacuo* to locate the minimum on the potential energy surface, considering the fact that the addition of solvent model in this step may be computationally expensive.[27] The optimized structures were verified by vibrational frequency calculations. To avoid the complexity in dealing with the effects of multiple conformations on chemical shifts, we decided to select small and rigid molecules into our database for linear regression. In the second step, NMR single-point calculations were performed with an implicit solvent model, which has been shown to significantly improve the accuracy compared to NMR calculations *in vacuo*.[1] The implicit solvent model SMD[28] was applied. The gauge-independent atomic orbital (GIAO) approach was applied to calculate the ^{15}N isotropic shielding constants.[3] All the calculations were performed with Gaussian 09.[29] To be consistent with Tantillo and co-workers' previous work on ^{1}H and ^{13}C chemical shifts,[17] six different methods for the geometric optimization and NMR calculations were adopted for ^{15}N chemical shift calculations (Table 1, Method 1 to 6). This will facilitate ^{15}N chemical shift predictions without additional calculations in practical applications. Additionally, a recently proposed method B3LYP/cc-pVDZ, which was applied by Xin *et al.*,[16] was included for comparison (Table 1, Method 7). However, what needs to be underscored here is that there are two differences between our procedure and the procedure proposed by Xin *et al.*:[16] Firstly, the databases are different, so the fitted scaling factors are different; secondly, the solvation models are different, we applied the SMD model,[28] while they used the CPCM model.[30] The linear regression was then fitted based on experimental chemical shifts (δ)

and computed isotropic shielding constants (σ) for the test set with the following equation

$$\delta = \frac{\text{intercept} - \sigma}{-\text{slope}}. \quad (1)$$

Once these scaling factors are determined and verified accordingly, they can be applied for prediction of unknown ^{15}N chemical shifts in other molecules.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgments

H.Y. is the recipient of an Australian Research Council Future Fellowship (Project number FT110100034) and X.W. is the recipient of a University of Wollongong Vice-Chancellor's Postdoctoral Fellowship. We thank Andrew Montgomery for helpful discussions. We wish to acknowledge the Australian Government for an Australian International Postgraduate Award scholarship for P.G. This research was in part supported under the Australian Research Council's Discovery Projects funding scheme (project number DP170101773). We wish to acknowledge that this research was undertaken with the assistance of resources provided at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government (Project id: v15).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, *Chemical Reviews* **2012**, *112*, 1839–1862.

[2] P. Pyykkö, *Theoretical Chemistry Accounts* **2000**, *103*, 214–216.

[3] R. Ditchfield, *Molecular Physics* **1974**, *27*, 789–807.

[4] J. Gauss, J. F. Stanton in *Advances in Chemical Physics* (Eds.: I. Prigogine, S. A. Rice), John Wiley & Sons, Inc., **2003**, pp. 355–422.

[5] M. Dracinsky, P. Bour, *Journal of Chemical Theory and Computation* **2010**, *6*, 288–299.

[6] G. Barone, D. Duca, A. Silvestri, L. Gomez-Paloma, R. Riccio, G. Bifulco, *Chemistry - A European Journal* **2002**, *8*, 3240–3245.

[7] T. Helgaker, M. Jaszuński, K. Ruud, *Chemical Reviews* **1999**, *99*, 293–352.

[8] M. Kaupp, O. L. Malkina, V. G. Malkin, P. Pyykkö, *Chemistry - A European Journal* **1998**, *4*, 118–126.

[9] P. R. Rablen, S. A. Pearlman, J. Finkbiner, *The Journal of Physical Chemistry A* **1999**, *103*, 7357–7363.

[10] K. W. Wiitala, T. R. Hoye, C. J. Cramer, *Journal of Chemical Theory and Computation* **2006**, *2*, 1085–1092.

[11] M. M. King, H. J. C. Yeh, G. O. Dudek, *Organic Magnetic Resonance* **1976**, *8*, 208–212.

[12] L. Stefaniak, M. Witanowski, B. Kamieński, G. A. Webb, *Organic Magnetic Resonance* **1980**, *13*, 274–276.

[13] A. Puszko, K. Laihia, E. Kolehmainen, Z. Talik, *Structural Chemistry* **2013**, *24*, 333–337.

[14] J. Jaźwiński, L. Stefaniak, *Magnetic Resonance in Chemistry* **1993**, *31*, 447–450.

[15] A. Barszczewicz, M. Jaźwiński, L. Stefaniak, *Chemical Physics Letters* **1991**, *186*, 313 – 318.

[16] D. Xin, C. A. Sader, U. Fischer, K. Wagner, P.-J. Jones, M. Xing, K. R. Fandrick, N. C. Gonnella, *Organic Biomolecular Chemistry* **2017**, *15*, 928–936.

[17] *CHESHIRE CCAT, the Chemical Shift Repository for computed NMR scaling factors, with Coupling Constants Added Too.*, **2017**, `http://cheshirenmr.info/index.htm`.

[18] M. Witanowski, Z. Biedrzycka, W. Sicinska, Z. Grabowski, G. Webb, *Journal of Magnetic Resonance* **1997**, *124*, 127 – 131.

[19] H. Andersson, A.-C. C. Carlsson, B. Nekoueishahraki, U. Brath, M. Erdélyi in *Annual Reports on NMR Spectroscopy*, Elsevier, **2015**, pp. 73–210.

[20] K. C. Nicolaou, S. A. Snyder, *Angewandte Chemie International Edition* **2005**, *44*, 1012–1044.

[21] V. S. Yeh, *Tetrahedron* **2004**, *60*, 11995 – 12042.

[22] Z. Jin, *Natural Product Reports* **2011**, *28*, 1143–1191.

[23] E. Vitaku, D. T. Smith, J. T. Njardarson, *Journal of Medicinal Chemistry* **2014**, *57*, 10257–10274.

[24] A. Laxer, D. T. Major, H. E. Gottlieb, B. Fischer, *The Journal of Organic Chemistry* **2001**, *66*, 5463–5481.

[25] F. Huang, M. Koenen-Bergmann, T. R. MacGregor, A. Ring, S. Hattox, P. Robinson, *Antimicrobial Agents and Chemotherapy* **2008**, *52*, 4300–4307.

[26] S. W. Meyer, M. Köck, *Journal of Natural Products* **2008**, *71*, 1524–1529.

[27] C. Maurizio, R. Nadia, S. Giovanni, B. Vincenzo, *Journal of Computational Chemistry* **2003**, *24*, 669–681.

[28] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *The Journal of Physical Chemistry B* **2009**, *113*, 6378–6396.

[29] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09 Revision E.01*, Gaussian Inc. Wallingford CT 2009.

[30] V. Barone, M. Cossi, *The Journal of Physical Chemistry A* **1998**, *102*, 1995–2001.

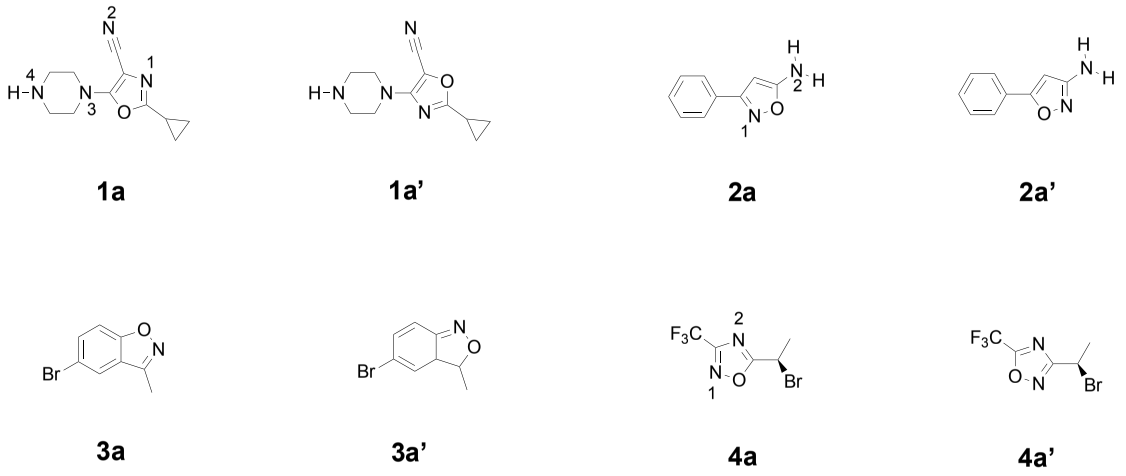


Figure 1: Regioisomers for selected oxazoles (**1**), isoxazoles (**2** and **3**) and oxadiazoles (**4**).

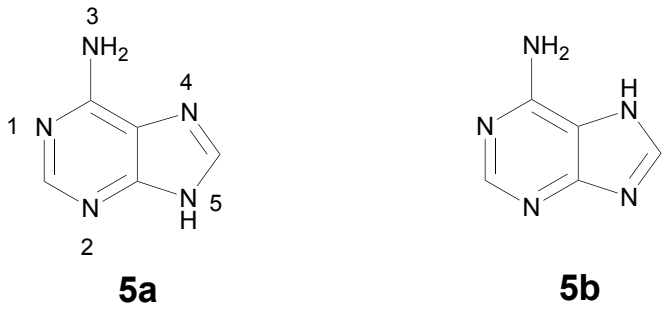


Figure 2: Two different tautomers for adenine.

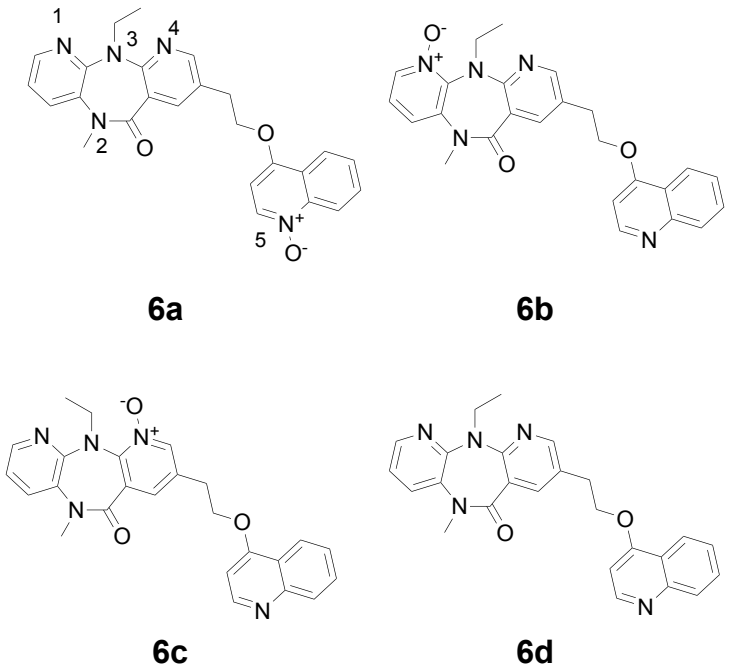


Figure 3: Molecule **6** in different oxidation states.

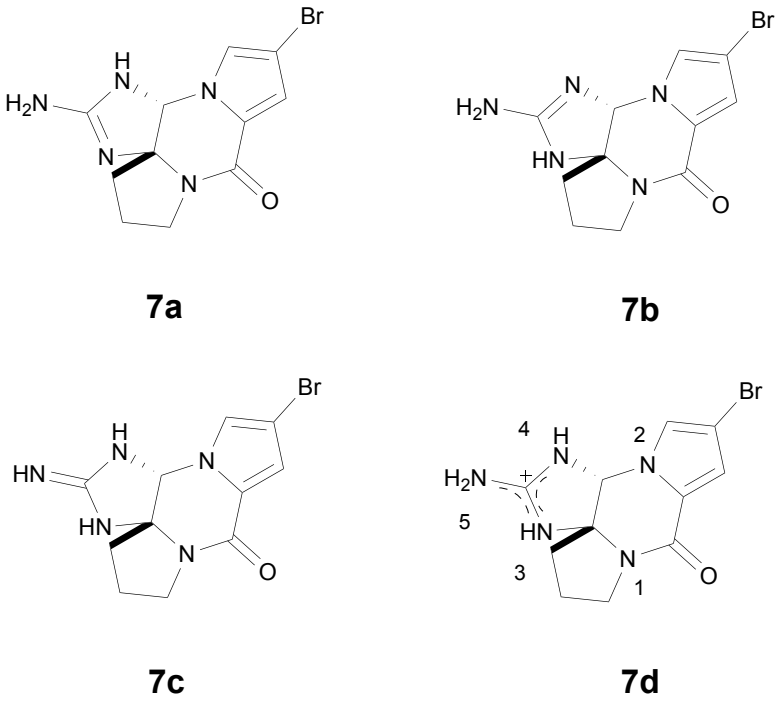


Figure 4: Molecule **7** and related structures.

Table 1: The adopted seven methods for calculating ^{15}N isotropic shielding constants and the fitted empirical scaling parameters (slope and intercept) in $CHCl_3$ and DMSO.

Method	Geometry[a)]	NMR[b)]	$CHCl_3$[c)]		DMSO[d)]	
	(opt & freq)	(GIAO, SMD)	slope	intercept	slope	intercept
1	B3LYP/6-31+G(d,p)	mPW1PW91/6-311+G(2d,p)	-1.0180	-152.60	-0.9981	-145.37
2	B3LYP/6-311+G(2d,p)	mPW1PW91/6-311+G(2d,p)	-1.0139	-148.67	-0.9917	-141.63
3	B3LYP/6-31+G(d,p)	PBE0/6-311+G(2d,p)	-1.0216	-150.45	-0.9928	-143.23
4	B3LYP/6-311+G(2d,p)	PBE0/6-311+G(2d,p)	-1.0086	-145.56	-0.9874	-139.46
5	M062X/6-31+G(d,p)[f]	mPW1PW91/6-311+G(2d,p)	-0.9915	-144.01	-0.9687	-136.01
					-0.9768[e)]	-120.21[e)]
6	M062X/6-311+G(2d,p)[f]	mPW1PW91/6-311+G(2d,p)	-0.9909	-141.49	-0.9653	-133.57
					-0.9589[e)]	-113.34[e)]
7	B3LYP/cc-pVDZ	B3LYP/cc-pVDZ	-0.9766	-129.06	-0.9578	-122.28
					-0.9222[e)]	-93.26[e)]

[a)] The geometry optimization and vibrational frequency calculations were performed with this method. [b)] The GIAO calculations[3] were performed with this method together with the SMD solvation model in Gaussian 09.[28] [c)] The fitted empirical scaling factors (slope and intercept in Eq. 1) for the chemical shift calculations in $CHCl_3$. The linear fitting can be seen in ESI Figure S1. [d)] The fitted empirical scaling factors (slope and intercept in Eq. 1) for the chemical shift calculations in DMSO. The linear fitting can be seen in ESI Figure S2. [e)] The fitted empirical scaling factors (slope and intercept in Eq. 1) for -NH_2 chemical shift calculations in DMSO based on the separate database of -NH_2 containing molecules. The linear fitting can be seen in ESI Figure S3. Such a fitting was not feasible for $CHCl_3$ due to lack of experimental data. [f] #int=ultrafine was included in all calculations involving M06 functionals.

Table 2: The performance of the fitted empirical scaling factors for the seven methods in Table 1.

Method	$CHCl_3$				DMSO			
	Test[a]			Probe[b]	Test[a]			Probe[b]
	R^{2}[c]	RMSD[d]	Rel. RMSD (%)[g]	RMSD[d]	R^{2}[c]	RMSD[d]	Rel. RMSD (%)[g]	RMSD[d]
1	0.9912	6.11	6.0	5.65	0.9877	8.04	7.5	7.54
2	0.9921	5.97	5.8	5.27	0.9886	7.74	7.2	7.64
3	0.9923	5.91	5.8	5.28	0.9885	7.77	7.2	7.48
4	0.9926	5.86	5.7	5.05	0.9891	7.54	7.0	7.48
5	0.9938	5.30	5.2	4.77	0.9899	7.26	6.9	7.33
					0.9913[e]	1.18[f]	0.4[g]	
6	0.9937	5.35	5.2	4.78	0.9903	7.12	6.6	7.49
					0.9982[e]	1.38[f]	0.4[g]	
7	0.9902	6.67	6.5	4.40	0.9906	7.00	6.5	6.15
					0.9758[e]	1.99[f]	0.6[g]	

[a] The performance of the empirical scaling for the test set. [b] The performance of the empirical scaling for the probe set. [c] R^2 is the coefficient of determination for the linear regression. [d] RMSD for the predicted chemical shifts with respect to their respective experimental values (in ppm) for the database of non -NH_2 containing molecules. [e] R^2 is the coefficient of determination for the linear regression of the independent database of -NH_2 containing molecules. [f] RMSD for the predicted chemical shifts with respect to their respective experimental values (in ppm) for the independent database of -NH_2 containing molecules. [g] Relative RMSD for the independent database of -NH_2 containing molecules.

Table 3: Transferability of the fitted scaling factors between $CHCl_3$ and DMSO.

Method	DMSO to $CHCl_3$[a)]	$CHCl_3$ to DMSO[b)]
1	8.27	9.46
2	7.96	9.01
3	8.14	9.25
4	7.91	8.45
5	8.19	9.22
6	8.03	8.91
7	8.62	8.56

[a)] The RMSDs were obtained for the test set when applying the fitted scaling factor of DMSO to predict the chemical shifts in $CHCl_3$. [b)] The RMSDs were obtained for the test set when applying the fitted scaling factor of $CHCl_3$ to predict the chemical shifts in DMSO.

Table 4: Experimental and predicted ^{15}N NMR chemical shifts (in ppm) for selected oxazoles (**1**), isoxazoles (**2**-**3**), and oxadiazoles (**4**).

	Pos.[a]	Exp.[b]	$\delta_{calc.}$ - δexp.			
			5[c]	6[d]	7[e]	Ref[i]
1a	1	-139.3	-4.0	-3.9	-2.9	-7.0
	2	-114.1	-2.0	-4.8	-4.9	-1.5
	3	-307.8	3.2	1.8	9.7	8.4
	4	-349.7	-10.3	-9.5	-5.6	-2.2
	Avg[j]		5.8	5.7	6.3	5.6
1a'			-17.0	-16.3	-22.8	-17.5
			18.4	14.4	16.5	19.3
			3.1	2.1	7.6	7.1
			-9.7	-9.0	-5.0	-0.7
	Avg[j]		13.5	11.8	14.8	13.5
2a	1	-34.5	-9.2	-7.0	-2.6	-1.7
	2	-322.5	-2.0[f]	-2.2[g]	-1.0[h]	0.4
	Avg[j]		6.6	5.2	2.0	1.2
2a'			-25.1	-23.2	-18.5	-17.1
			-13.7[f]	-13.8[g]	-14.1[h]	-11.2
	Avg[j]		20.2	19.1	16.5	15.7
3a	1	-2.8	-9.2	-8.1	-4.2	-1.7
3a'			-22.4	-20.0	-16.0	-12.2
4a	1	-13.8	-0.1	0.4	4.0	3.1
	2	-143.0	0.9	1.3	-0.5	0.7
	Avg[j]		0.6	0.9	2.8	2.2
4a'			-5.4	-3.5	0.9	0.9
			13.0	12.8	10.3	10.6
	Avg[j]		10.0	9.3	7.3	7.5

[a] Positions for the nitrogen of interest. [b] Experimental data taken from Ref.[16]. [c] *d*) [e] The calculated chemical shifts by Method 5, 6 and 7, respectively. [f] The calculated chemical shift for Pos.2 by -NH_2 group scaling factors based on Method 5 (slope: -0.9768, intercept: -120.21). [g] The calculated chemical shift for Pos.2 by -NH_2 group scaling factors based on Method 6 (slope: -0.9589, intercept: -113.34). [h] The calculated chemical shift for Pos.2 by -NH_2 group scaling factors based on Method 7 (slope: -0.9222, intercept: -93.26). [i] The difference between calculated chemical shifts and experimental values published by Xin *et al.*[16] [j] Root-mean-square deviation of the prediction for each structure.

Table 5: Experimental and predicted ^{15}N NMR chemical shifts (in ppm) for adenine.

	Pos.[a]	Exp.[b]	$\delta_{calc.}$ - $\delta_{exp.}$			
			5[c]	6[d]	7[e]	Ref[f]
5a	1	-149.8	-7.2	-5.1	-6.4	-4.2
	2	-151.5	-8.5	-7.0	-7.2	-4.6
	3	-301.2	-1.2[f]	1.6[g]	0.3[h]	-1.2
	4	-140.3	-7.1	-6.7	-7.2	-3.9
	5	-222.6	-7.4	-6.9	-11.4	-14.0
	Avg[j]		6.8	5.8	7.4	7.1
5b			3.7	5.0	6.1	2.7
			15.6	16.5	17.4	16.4
			-1.9[f]	-1.6[g]	-3.5[h]	-3.1
			0.6	1.3	1.0	2.3
			-20.7	-20.5	-23.4	-26.6
	Avg[j]		11.8	12.0	13.4	14.2

[a] Positions for the nitrogen of interest. [b] Experimental data taken from Ref.[16,24]. [c] [d] [e] The calculated chemical shifts by Method 5, 6 and 7, respectively. [f] The calculated chemical shift for Pos.3 by -NH_2 group scaling factors based on Method 5 (slope: -0.9768, intercept: -120.21) . [g] The calculated chemical shift for Pos.3 by -NH_2 group scaling factors based on Method 6 (slope: -0.9589, intercept: -113.34). [h] The calculated chemical shift for Pos.3 by -NH_2 group scaling factors based on Method 7 (slope: -0.9222, intercept: -93.26). [i] The difference between calculated chemical shifts and experimental values published by Xin *et al.*[16] [j] Root-mean-square deviation of the prediction for each structure.

Table 6: Experimental and predicted ^{15}N NMR chemical shifts for molecule **6** in different oxidation states.

	Pos.[a]	Exp.[b]	$\delta_{calc.}$ - $\delta_{exp.}$			
			5[c]	6[d]	7[e]	Ref[f]
6a	1	-86.5	2.2	2.3	-1.9	-1.1
	2	-256.9	1.7	1.9	0.1	1.3
	3	-273.3	0.5	1.1	4.9	4.0
	4	-91.2	-5.2	-4.8	-6.4	-3.5
	5	-113.3	9.5	9.2	13.6	8.1
	Avg[g]		5.0	4.9	7.1	4.4
6b			1.5	1.1	-3.2	0.6
			6.2	6.0	6.5	6.3
			-25.4	-25.1	-24.6	-25.2
			8.4	8.6	9.9	5.6
			18.3	23.5	21.0	22.5
	Avg[g]		12.6	12.5	12.6	12.2
6c			7.4	7.1	8.3	4.6
			1.4	1.1	1.1	1.6
			-19.9	-19.4	-17.6	-18.7
			-1.4	-1.4	-7.0	-3.3
			17.9	18.1	20.7	22.4
	Avg[g]		12.4	12.3	13.1	13.3
6d			-2.9	-3.2	-3.7	-1.0
			2.2	2.1	1.0	1.4
			1.2	1.3	6.1	4.0
			1.3	1.4	-3.6	-2.6
			18.0	18.2	20.7	22.1
	Avg[g]		8.2	8.4	10.0	10.1

[a] Positions for the nitrogen of interest. [b] Experimental data taken from Ref.[16]. [c] [d] [e] The calculated chemical shifts by Method 5, 6 and 7, respectively. [f] The difference between calculated chemical shifts and experimental values published by Xin *et al.*[16] [g] Root-mean-square deviation of the prediction for each structure.

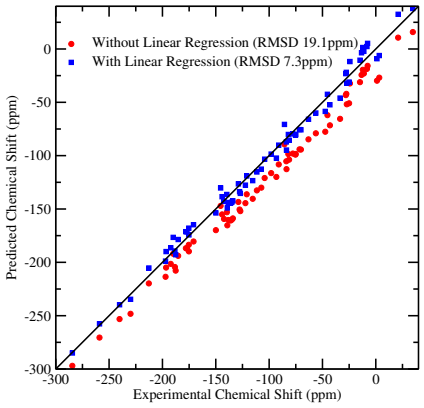
Table 7: Experimental and predicted ^{15}N NMR chemical shifts for molecule **7a** and related structures.

	Pos.[a]	Exp.[b]	$\delta_{calc.}$ - $\delta_{exp.}$			
			5[c]	6[d]	7[e]	Ref[f]
7a			6.1	6.0	4.8	5.6
			5.2	5.5	4.6	3.1
			74.6	78.3	76.1	71.0
			-4.5	-3.0	-4.0	-8.2
			-18.0[f]	-17.8[g]	-20.5[h]	-18.7
	Avg[j]		34.6	36.1	35.4	33.2
7b			0.6	0.9	0.9	0.3
			10.8	10.5	10.1	10.0
			-15.9	-14.0	-12.7	-13.3
			80.2	82.7	81.8	78.2
			-10.4[f]	-10.3[g]	-14.1[h]	-14.3
	Avg[j]		37.2	38.1	37.8	36.3
7c			1.6	2.1	1.1	1.2
			5.9	5.6	5.8	4.3
			-22.7	-21.5	-24.7	-25.0
			-9.8	-8.2	-10.3	-11.1
			51.8	57.0	58.7	61.3
	Avg[j]		25.8	27.6	29.0	30.1
7d	1	-259.2	-3.5	-3.0	-3.8	-4.5
	2	-211.2	-1.6	-1.3	-2.5	-5.4
	3	-271.2	-4.4	-2.8	-3.0	-9.4
	4	-285.2	-6.0	-4.9	-5.4	-8.4
	5	-303.2	-4.1[f]	-3.3[g]	-4.9[h]	-1.7
	Avg[j]		4.2	3.3	4.1	6.5

[a] Positions for the nitrogen of interest. [b] Experimental data taken from Ref.[16,26]. [c] [d] [e] The calculated chemical shifts by Method 5, 6 and 7, respectively. [f] The calculated chemical shift for Pos.5 by -NH_2 group scaling factors based on Method 5 (slope: -0.9768, intercept: -120.21). [g] The calculated chemical shift for Pos.5 by -NH_2 group scaling factors based on Method 6 (slope: -0.9589, intercept: -113.34). [h] The calculated chemical shift for Pos.5 by -NH_2 group scaling factors based on Method 7 (slope: -0.9222, intercept: -93.26). [i] The difference between calculated chemical shifts and experimental values published by Xin *et al.*[16] [j] Root-mean-square deviation of the prediction for each structure.

ToC Text: ^{15}N chemical shift calculations based on DFT/GIAO with linear regression significantly increase the accuracy of the prediction.

ToC figure:



ToC Keywords: Nuclear Magnetic Resonance Spectroscopy, Density Functional Theory, Gauge-Including Atomic Orbital, Nitrogen Chemical Shift, Linear Regression, Solvent Effects