

Research



Cite this article: Nguyen TT. 2021 $^1\text{H}/^{13}\text{C}$ chemical shift calculations for biaryls: DFT approaches to geometry optimization. *R. Soc. Open Sci.* **8**: 210954.
<https://doi.org/10.1098/rsos.210954>

Received: 1 June 2021

Accepted: 1 September 2021

Subject Category:

Chemistry

Subject Areas:

computational chemistry/organic chemistry/
spectroscopy

Keywords:

biaryl, NMR, chemical shift, DFT,
functionals, basis sets

Author for correspondence:

Thien T. Nguyen

e-mail: nguyentrongthien@duytan.edu.vn

This article has been edited by the Royal Society of Chemistry, including the commissioning, peer review process and editorial aspects up to the point of acceptance.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5628899>.



$^1\text{H}/^{13}\text{C}$ chemical shift calculations for biaryls: DFT approaches to geometry optimization

Thien T. Nguyen^{1,2}

¹Institute of Research and Development, and ²Faculty of Pharmacy, College of Medicine and Pharmacy, Duy Tan University, Da Nang 550000, Vietnam

TTN, 0000-0002-2080-786X

Twelve common density functional methods and seven basis sets for geometry optimization were evaluated on the accuracy of $^1\text{H}/^{13}\text{C}$ NMR chemical shift calculations for biaryls. For these functionals, ^1H shifts calculations for gas phase optimized geometries were significantly less accurate than those for in-solution optimized structures, while ^{13}C results were not strongly influenced by geometry optimization methods and solvent effects. B3LYP, B3PW91, mPW1PW91 and ωB97XD were the best-performing functionals with lowest errors; among seven basis sets, DGDZVP2 and 6-31G(d,p) outperformed the others. The combination of these functionals and basis sets resulted in high accuracy with $\text{CMAE}_{\min} = 0.0327$ ppm (0.76%) and 0.888 ppm (0.58%) for ^1H and ^{13}C , respectively. The selected functionals and basis set were validated when consistently producing optimized structures with high accuracy results for ^1H and ^{13}C chemical shift calculations of two other biaryls. This study highly recommends the IEFPCM/B3LYP, B3PW91, mPW1PW91 or ωB97XD /DGDZVP2 or 6-31G(d,p) level of theory for the geometry optimization step, especially the solvent incorporation, which would lead to high accuracy $^1\text{H}/^{13}\text{C}$ calculation. This work would assist in the fully structural assignments of biaryls and provide insights into in-solution biaryl conformations.

1. Introduction

Biaryls (figure 1) are prominent substructures found in important, chiral ligands and organocatalysts, which are useful for asymmetric transformations [1,2]. They are also present in biologically active natural products and biopolymer lignins [3,4]. Biaryls are used for sequencing lignin oligomers [5], oxidative C–C cleavage [6] and pyrolysis behaviour studies [7] and are found to be anti-oxidant and anti-inflammatory agents [8]. Owning biaryl

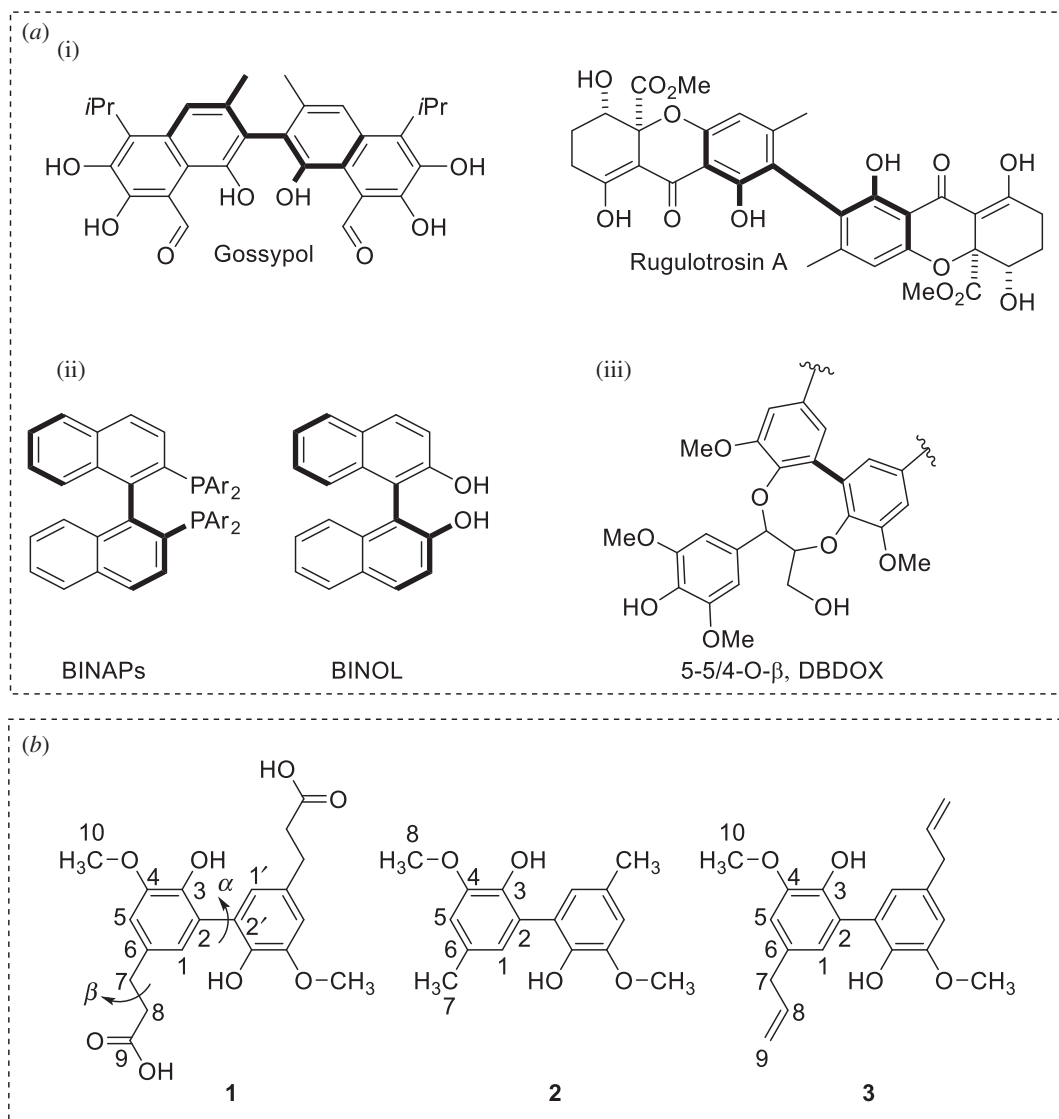


Figure 1. (a) Biaryls in natural products (i), ligands and catalysts (ii) and lignin structure (iii), and (b) Biaryls with numbering labels and the dihedral angles α (C1C2C2'C1') and β (C6C7C8C9) of biaryl 1.

linkages, they can also be employed for biaryl conformation analysis, which is intrinsically difficult due to the absence of proton NMR spin–spin coupling information. An accurate method for predicting NMR spectra would contribute valuable insights into the conformations of biaryl structures and the local electron environment of each NMR active nucleus. Gauge-independent atomic orbital (GIAO)-DFT NMR calculations have effectively supported the structural assignments and validations of biaryl compounds with accurate predictions at affordable computational costs [9,10]. In general, the accuracy is impacted by geometry optimizations, functionals, basis sets, solvation models and NMR methods. [11–14]. For the two common nuclei of organic molecules, ^1H shift predictions are more challenging than ^{13}C shift ones due to significant impacts of solvation effects on protons [15].

Previous studies reported how the use of different density functional methods and basis sets for NMR calculations affected $^1\text{H}/^{13}\text{C}$ results on a variety of organic structures [11,13,16,17]. Among those, few reports investigated the impact of optimized geometry on the NMR results. In 1997, Rablen *et al.* [18] reported the influence of molecular geometry on the calculated chemical shifts using four different levels of theory, including MP2/6-31 + G(d), B3LYP/6-31 + G(d), B3LYP/6-311 + +G(d,p) and B3LYP/6-311 + +G(2df,p) and found that such methods have small impacts on ^1H NMR calculations for small organic molecules. In 2008, Xu *et al.* [19] reported the use of functionals B3LYP, PBE0 and OPBE on 23 small, simple molecules and found that OPBE was the best-performing one. Herein, this present study shows how optimized geometries using 12 density functional methods and seven basis sets, which have been commonly employed for ^1H and ^{13}C NMR calculations of organic compounds, influence

the ^1H and ^{13}C NMR shift calculation accuracy for biaryls. Its aim is to reveal how sensitive such optimized geometries are on the chemical shift predictions of such important structure motifs and to selectively recommend density functional methods and basis sets for geometry optimization with high accuracy and an economical amount of CPU time. The broad differences of results would mean that the functional and basis set selections for geometry optimization methods are significant, while the small ones should provide confidence in carrying out such predictions.

2. Computational methods

Structures of **1**, **2** and **3** were initially optimized using MMFF94/MM2, which are implemented in Chem3D. The resulting geometries were used as the starting point for the individual optimizations using Gaussian09 [20], Integral equation formalism variant of the polarized continuum model (IEFPCM) was used when solvent effects were incorporated during geometry optimization [21]. Subsequent frequency calculations ensured that a potential energy surface local minimum was attained during the energy minimization. The size of integral was set at default because energy differences between optimized gas phase and in-solution structures were considered. Cartesian coordinates of such optimized geometries were given in the electronic supplementary material. For the performance evaluation of density functional methods and basis sets, local minima were employed in this investigation for economical computation. The following 12 functionals and seven basis sets were evaluated:

- Functionals: B3LYP (Becke's three-parameter hybrid functional using B exchange and LYP correlation) [22], B3PW91 (Perdew and Wang's 1991 gradient-corrected correlation functional) [23], BPV86 (Perdew's 1986 functional) [24], CAM-B3LYP (Handy and co-workers' long-range corrected version of B3LYP using the Coulomb-attenuating method) [25], HCTH (Hamprecht-Cohen-Tozer-Handy GGA functional) [26], HSEH1PB (The exchange part of the screened Coulomb potential of Heyd, Scuseria and Ernzerhof) [27], LSDA (Local spin-density approximation) [28], mPW1PW91 (mPW exchange and PW91 correlation) [29], PBEPBE (The functional of Perdew, Burke, and Ernzerhof) [30], TPSSPTSS (The exchange component of the Tao-Perdew-Staroverov-Scuseria) [31] and ω B97XD (Head-Gordon and co-workers' dispersion corrected long-range corrected hybrid functional) [32]; M06-2X (a high-non-locality functional with double the amount of nonlocal exchange) [33].
- Basis sets: Pople's 6-31G, 6-31G(d,p), 6-31 + G(d,p), and 6-311G [34]; Dunning's cc-pVDZ correlation consistent basis set [35]; DGDZVP2 [36]; SDD (Stuttgart-Dresden effective core potential) [37].

Unless specified otherwise, single-point NMR GIAO calculations were carried out at IEFPCM (DMSO or CHCl_3)/ ω B97XD/6-31G(d,p) level of theory, which was found to consistently provide ^1H and ^{13}C chemical shifts with a high level of accuracy [10]. The GIAO NMR results were observed and extracted using GaussView05. Each optimized structure was used for computing the corresponding isotropic shielding constants (σ_{cal}). The chemical shifts (δ_{cal}) given in the electronic supplementary material were obtained using equation (2.1). For both the ^1H and ^{13}C NMR calculations, an average of values of equivalents atoms was assumed. For example, a single proton/carbon signal is observed for the two methoxy groups of biaryl **1** due to fast rotations of biaryl linkage relative to the NMR measurement time scale. To reduce the systematic error of the calculations, the linear regression analysis of the calculated chemical shifts versus the experimental ones (δ_{exp}) (equation (2.2)) were performed, and the scaled chemical shifts (δ_{scal}) were computed according to equation (2.3). As the reference had a negligible impact on the linear regression analysis, the fixed values of 197 ppm and 31 ppm were chosen as TMS shielding constants for ^{13}C and for ^1H , respectively. Computed results were evaluated using mean absolute value ($|\Delta\delta|$ /ppm, equation (2.4)), corrected mean absolute error (CMAE/ppm, equation (2.5)), corrected root mean squared error (CRMSE/ppm, equation (2.6) and the Pearson correlation coefficient (r^2). The smaller values of CMAE and CRMSE indicate smaller errors and the larger value of r^2 means a stronger correlation between theoretical and experimental data. Error calculations and linear correlations were performed using Microsoft Excel 2013.

$$\delta_{\text{cal}} = \sigma_{\text{TMS}} - \sigma_{\text{cal}}, \quad (2.1)$$

$$\delta_{\text{cal}} = a\delta_{\text{exp}} + b, \quad (2.2)$$

$$\delta_{\text{scal}} = \frac{\delta_{\text{cal}} - b}{a}, \quad (2.3)$$

$$|\Delta\delta| = |\delta_{\text{scal}} - \delta_{\text{exp}}|, \quad (2.4)$$

$$\text{CMAE} = \sum_1^n |\delta_{\text{scal}} - \delta_{\text{exp}}|/n \quad (2.5)$$

$$\text{CRMSE} = \sqrt{\sum_1^n (\delta_{\text{scal}} - \delta_{\text{exp}})^2/n}. \quad (2.6)$$

Figure 1*b* shows the numbered biaryls used for the proton and carbon atoms in this study. Due to the axial symmetry of biaryls **1**, **2** and **3**, only one side of the structures was labelled. The experimental ^1H and ^{13}C NMR spectra of **1** [10,38], **2** [39,40] and **3** [41,42] were reported. These compounds contain phenolic and carboxylic protons which typically did not appear in the NMR spectra due to rapid exchanges in $\text{DMSO}-d_6$ or CDCl_3 solvent. Therefore, this work excluded these protons in the calculations. This also excluded the main interacting atoms of the molecules leading to basis set superposition errors (BSSE). In addition, the BSSE should have a similar contribution to each NMR calculation on the same molecule. Therefore, these errors were insignificant for the comparison of performances between density functional methods and basis sets in this study.

3. Results and discussion

3.1. Functional impact

Initially, optimized geometries of biaryl **1** in gas phase and in DMSO using 12 functionals coupled with 6-31G(d,p) basis set were tested on compound **1**, which was previously prepared by a two-stepped synthesis from ferulic acid [10]. The optimized structures were obtained with energies and different dihedral angles α (C1C2C2'C1') and β (C6C7C8C9) (figure 2). Because the energy differences between gas phase and in-solution geometries were considered, the fine grid was used during the optimization step. Aliphatic angle β was not altered much, while biaryl angle α showed a significant discrepancy when HCTH (table 1, entry 5) and LSDA (table 1, entry 7) were employed. In-solution structures were about 19 kcal/mol more stable than gas phase ones. These structures were subjected to the ^1H and ^{13}C chemical shift calculations and the results are shown in table 2 and figure 2.

The ^1H NMR calculations of these geometries were calculated at the level of IEFPCM(DMSO)/ $\omega\text{B97XD}/6\text{-}31\text{G(d,p)}$ [10], and their statistical parameters and mean absolute values are shown in table 2 and figure 2, respectively. Compared to other functionals, HCTH (Entry 5) and LSDA (Entry 7) showed discrepancy results with highest errors. Therefore, we excluded these two functionals in figure 2 and further discussions. These results were expected as the optimized gas phase and in-solution geometries obtained from these functionals were too different from other functionals (table 1). For other functionals, the statistical parameters for gas phase structures $(0.0711 \text{ ppm (1.66\%)} \leq \text{CMAEs} \leq 0.0991 \text{ ppm (2.31\%)})$, $0.0770 \text{ ppm (1.79\%)} \leq \text{CRMSEs} \leq 0.113 \text{ ppm (2.63\%)}$ and $0.9963 \leq r^2 \leq 0.9983$ were significantly less accurate than those for in-solution structures $(0.0327 \text{ ppm (0.76\%)} \leq \text{CMAE} \leq 0.0549 \text{ ppm (1.28\%)})$, $0.0412 \text{ ppm (0.96\%)} \leq \text{CRMSE} \leq 0.0585 \text{ ppm (1.36\%)}$, $0.9990 \leq r^2 \leq 0.9995$). This could be explained by the high exposure of hydrogen atoms to solvent molecules. The solvent effects can be clearly observed in figure 2; except for calculated **H10** shifts of the methoxy group, the mean absolute values of aromatic protons **H1** and **H5**, and methylene protons **H7** and **H8** shifts were significantly decreased when going from gas phase to in-solution structures. For in-solution geometries, noticeable deviations were consistently observed for methylene protons **H7** and **H8** ($|\Delta\delta|_{\text{max}} = 0.0764 \text{ ppm}$, ωB97XD) were observed with the noticeable deviations ranging from 0.0715 to 0.166 ppm. Therefore, solvent incorporation during the optimization process is essential for achieving the high accuracy of ^1H calculations. The three best performers, those with lowest errors, were B3PW91 (CRMSE = 0.0412 ppm, 0.96%), mPW1PW91 (CRMSE = 0.0432 ppm, 1.01%) and B3LYP (CRMSE = 0.0437 ppm, 1.02%).

^{13}C results for structures optimized in gas phase and in DMSO were in high agreement with the experimental data for all tested functionals (table 3). CMAE and CRMSE values of all 12 tested functionals were in the narrow ranges of 1.09 (0.71%) to 1.50 ppm (0.98%) and 1.30 (0.85%) to 1.83 ppm (1.19%), respectively. The ^{13}C results were obtained with high coefficients of determination ($0.9985 \leq r^2 \leq 0.9992$). Compared to the ^1H results, ^{13}C chemical shift calculations were not strongly influenced by molecular geometries and solvation effects. This could be explained by the fact that carbon nuclei are less exposed to solvent molecules than protons. Optimized geometries using B3LYP, B3PW91, BPV86, CAM-B3LYP, HSEH1PBE, LSDA, M06-2X, mPW1PW91 and PBEH1PBE produced high correlated results. Among those, B3LYP is economical from the view of computational cost. The

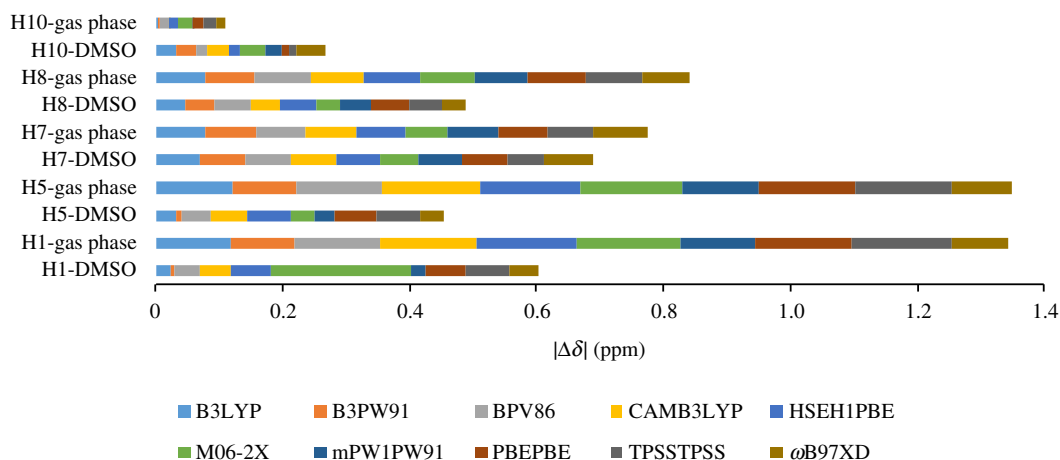


Figure 2. Mean absolute values of ^1H chemical shifts for geometries optimized in gas phase and in DMSO using 12 functionals coupled with 6-31G(d,p).

Table 1. Dihedral angles α and β and the energy differences of optimized geometries ($\Delta E = E_{\text{gas phase}} - E_{\text{in-solution}}$).

entry	functional	gas phase		in-solution		ΔE (kcal/mol)
		α	β	α	β	
1	B3LYP	−132.89	−179.06	−128.07	−179.76	19.07
2	B3PW91	−131.98	−178.94	−126.84	−179.62	19.80
3	BPV86	−134.82	−178.99	−130.24	−179.09	18.46
4	CAM-B3LYP	−134.07	−179.15	−129.03	−179.94	19.58
5	HCTH	−125.29	−178.93	−114.71	−179.92	18.43
6	HSEH1PBE	−134.85	−179.24	−130.42	−179.51	19.73
7	LSDA	−150.22	−173.91	−146.66	−174.71	18.66
8	M06-2X	−134.48	−179.42	−130.62	−179.79	19.95
9	mPW1PW91	−132.88	−179.14	−127.97	−179.7	19.97
10	PBEPBE	−135.51	−179.29	−131.39	−179.14	18.19
11	TPSSTPSS	−135.88	−179.16	−131.33	−179.57	18.53
12	ωB97XD	−131.06	−179.68	−123.05	−179.39	20.21

noticeable deviations of carbon atoms **C8** and **C10**, which are sp^3 -hybridized centres and among the most shielded carbons, were consistently observed in the range of 1.58 to 2.67 ppm and 2.11 to 3.87 ppm, respectively (figure 3).

In addition, when considering only sp^2 -hybridized carbons for the NMR calculations with a much narrower spectral window of ^{13}C chemical shifts (63.1 ppm), slight decreases in correlations ($0.9943 \leq r^2 \leq 0.9984$) and CRMSE values ($0.94 \text{ ppm (1.49\%)} \leq \text{CRMSEs} \leq 1.44 \text{ ppm (2.28\%)}$) for all tested density functional methods were consistently observed (figure 4). These lower correlations were expected due to the narrower spectral window. The CRMSE values were smaller in all cases but their relative percentages (CRMSE values over the spectral window) were much higher. Even though the statistical parameters for all carbon of the molecule were more meaningful than those for a carbon type due to the intrinsically electronic environments of a specific skeleton leading to characteristic magnetic properties of the active NMR carbon nucleus, the results would clarify the impact of narrowing the spectral window on calculation accuracy.

3.2. Basis set effects

Seven common basis sets, including 6-31G, 6-31G(d,p), 6-31G + (d,p), 6-311G, cc-pVDZ, DGDZVP2 and SDD, were coupled with B3LYP for in-solution geometry optimization steps, and NMR calculations of

Table 2. Accuracy evaluation of ^1H chemical shift calculations for optimized geometries of biaryl 1 in gas phase and in DMSO using 12 functionals coupled with 6-31G(d,p) basis set.

entry	functional	$\delta(^1\text{H})$ -gas phase			$\delta(^1\text{H})$ -DMSO		
		r^2	CMAE	CRMSE	r^2	CMAE	CRMSE
1	B3LYP	0.9976	0.0794	0.0901	<i>0.9994</i>	<i>0.0406</i>	<i>0.0437</i> ^a
2	B3PW91	0.9981	0.0729	0.081	<i>0.9995</i>	<i>0.0327</i>	<i>0.0412</i> ^a
3	BPV86	0.9970	0.0898	0.101	0.9993	0.0465	0.0499
4	CAM-B3LYP	0.9965	0.0940	0.110	0.9992	0.0519	0.0535
5	HCTH	0.9990	0.0448	0.058	0.9968	0.101	0.105
6	HSEH1PBE	0.9963	0.0989	0.113	0.9990	0.0549	0.0584
7	LSDA	0.9670	0.288	0.341	0.9820	0.205	0.250
8	mPW1PW91	0.9975	0.0808	0.0917	<i>0.9995</i>	<i>0.0399</i>	<i>0.0432</i> ^a
9	M06-2X	0.9962	0.100	0.114	0.9989	0.0790	0.106
10	PBEPBE	0.9965	0.0975	0.110	0.9990	0.0543	0.0585
11	TPSSTPSS	0.9964	0.0991	0.112	0.9991	0.0543	0.0585
12	ω B97XD	0.9983	0.0711	0.077	0.9992	0.049	0.0510

^aThe data of three best functionals are in italics.**Table 3.** Accuracy evaluation of ^{13}C chemical shift calculations for optimized geometries of biaryl 1 in gas phase and in DMSO using 12 functionals coupled with 6-31G(d,p) basis set.

entry	functional	$\delta(^{13}\text{C})$ -gas phase			$\delta(^{13}\text{C})$ -DMSO		
		r^2	CMAE	CRMSE	r^2	CMAE	CRMSE
1	B3LYP	0.9992	1.09	1.34	0.9992	1.09	1.36
2	B3PW91	0.9989	1.19	1.57	0.9991	1.13	1.40
3	BPV86	0.9990	1.12	1.46	0.9992	1.18	1.30
4	CAM-B3LYP	0.9990	1.16	1.50	0.9992	1.09	1.34
5	HCTH	0.9986	1.36	1.75	0.9988	1.35	1.62
6	HSEH1PBE	0.9989	1.19	1.55	0.9992	1.09	1.35
7	LSDA	0.9990	1.18	1.51	0.9990	1.35	1.50
8	M06-2X	0.9988	1.27	1.63	0.9991	1.18	1.42
9	mPW1PW91	0.9989	1.23	1.59	0.9991	1.16	1.42
10	PBEPBE	0.9990	1.15	1.48	0.9992	1.20	1.33
11	TPSSTPSS	0.9991	1.10	1.41	0.9985	1.50	1.83
12	ω B97XD	0.9987	1.31	1.68	0.9989	1.27	1.53

optimized structures were carried out. Table 4 displays the dihedral angles α and β , and table 5 shows the statistical parameters. While adding polarity functions for heavy atoms and hydrogen atoms (table 5, entry 1 and 2) significantly increased the calculation accuracy, a set of the diffusion function to Pople's basis sets (table 5, entry 2 and 3) created opposite effects and cost much more computation time. 6-31G + (d,p) (table 5, entry 3), generating the most discrepancy geometries, provided ^1H results with low correlations and large errors. Other basis sets produced highly correlated results with CMAEs ≤ 0.0957 ppm, CRMSEs ≤ 0.111 ppm and $r^2 \geq 0.9987$ for ^1H and CMAEs ≤ 1.23 ppm, CRMSEs ≤ 1.62 ppm, and $r^2 \geq 0.9988$ for ^{13}C . These results would allow a meaningful prediction of ^1H and ^{13}C chemical shifts. The best-performing basis set was DGDZVP2 (CMAE = 0.0361 ppm for ^1H ; CMAE = 0.888 ppm for ^{13}C). In terms of absolute deviations, the totals of absolute deviations for protons ($\sum |\delta|$)

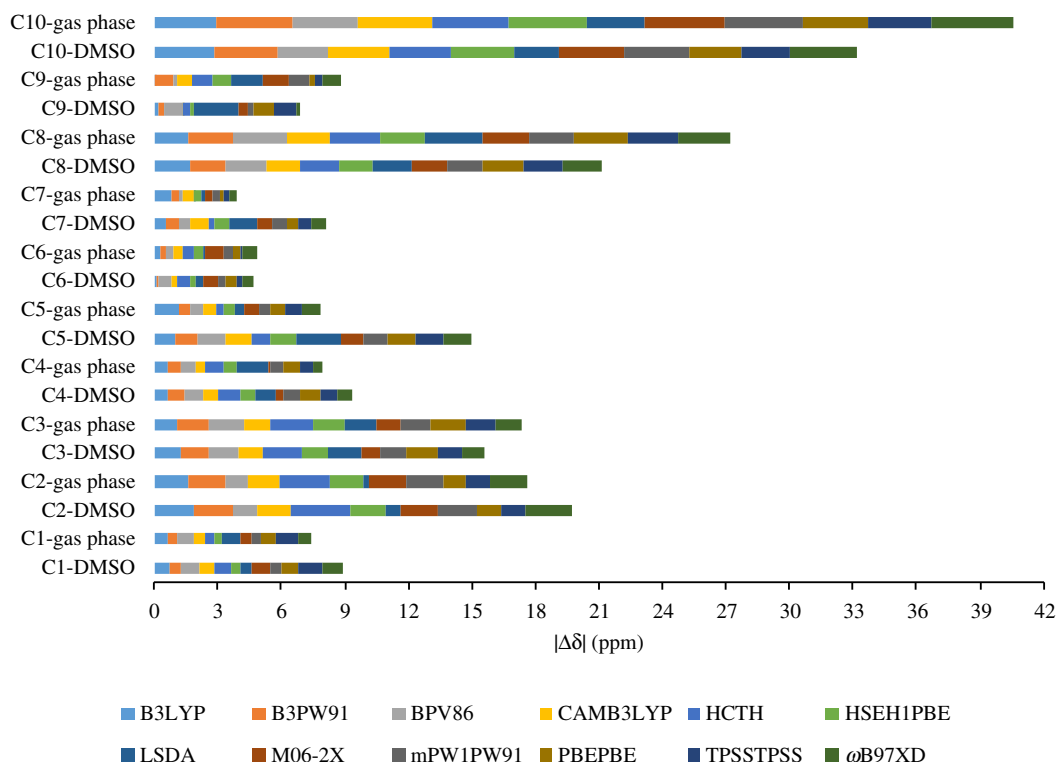


Figure 3. Mean absolute values of ^{13}C chemical shifts for geometries optimized in gas phase and in DMSO using 12 functionals coupled with 6-31G(d,p).

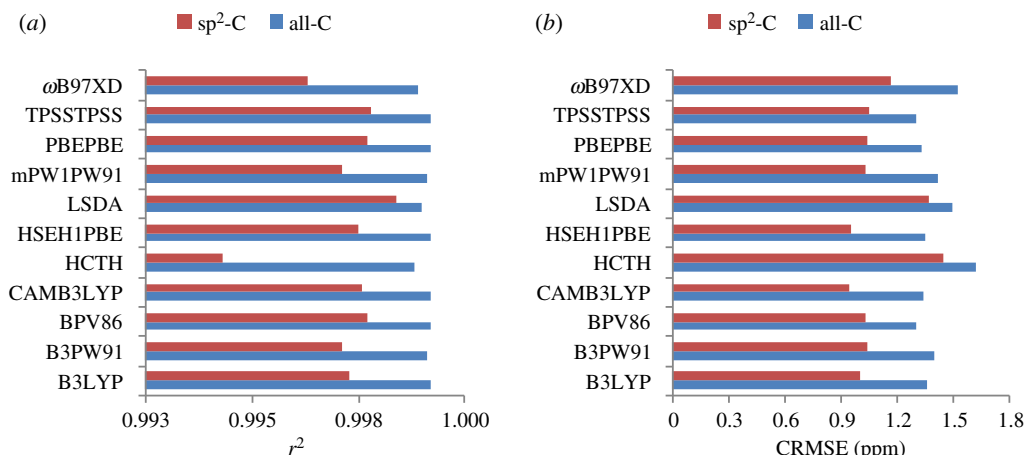


Figure 4. r^2 Values (a) and CRMSE (b) of the ^{13}C chemical shift calculations for all carbons (All-C) and sp^2 -hybridized ones ($\text{sp}^2\text{-C}$).

were evenly distributed while those for C8, C9 and C10 were relatively large. DGDZVP2 was more effective than the other basis sets, except for methoxy protons H10, aromatic carbon C2 and methoxy carbon C10 (figure 5).

The $^1\text{H}/^{13}\text{C}$ calculations using the other combination of functionals and basis sets were carried out in order to assert the best-performing functionals and basis set. LSDA and HCTH functionals were excluded from these calculations due to their relatively low performances. As one of the Pople' basis sets were tested for all tested functionals (tables 2 and 3), the remaining sets would provide similar results. Therefore, cc-pVDZ, DGDZVP and SDD were coupled with 10 functionals for the calculations, and the r^2 values are summary in table 6 (full statistical parameters Are given in the electronic supplementary material). When using the same basis sets, ^{13}C results were relatively less affected by the choice of functionals. The r^2 values were ranged from 0.9990 to 0.9994, 0.9991 to 0.9995 and 0.9980 to 0.9988 for cc-pVDZ, DGDZVP and SDD, respectively (table 6). ^1H results showed strong

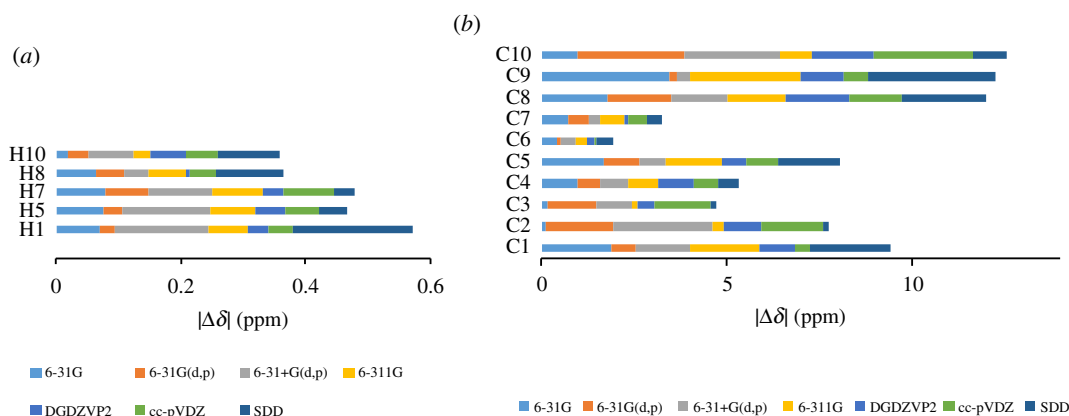


Figure 5. Mean absolute values of ^1H (a) and ^{13}C (b) chemical shifts for geometries optimized in DMSO using seven basis sets coupled with B3LYP.

Table 4. Dihedral angles α and β of optimized geometries using seven basis sets.

entry	basis set	α	β
1	6-31G	−130.29	−179.62
2	6-31G(d,p)	−132.89	−179.06
3	6-31G + (d,p)	−111.62	−179.88
4	6-311G	−129.63	−179.90
5	cc-pVDZ	−129.28	−178.97
6	DGDZVP2	−124.66	−179.69
7	SDD	−130.77	−179.55

Table 5. Accuracy evaluation of ^{13}C and ^1H chemical shift calculations for optimized geometries in DMSO using seven basis sets coupled with B3LYP functional.

entry	basis set	$\delta(^1\text{H})$			$\delta(^{13}\text{C})$		
		r^2	CMAE	CRMSE	r^2	CMAE	CRMSE
1	6-31G	0.9988	0.0616	0.0652	0.9989	1.21	1.54
2	6-31G(d,p)	0.9994	0.0406	0.0437	0.9992	1.09	1.36
3	6-31G + (d,p)	0.9965	0.101	0.109	0.9991	1.17	1.44
4	6-311G	0.9988	0.0611	0.064	0.9991	1.10	1.38
5	cc-pVDZ	0.9991	0.0525	0.0543	0.9993	1.03	1.27
6	<i>DGDZVP2^a</i>	<i>0.9995</i>	<i>0.0361</i>	<i>0.0404</i>	<i>0.9995</i>	<i>0.888</i>	<i>1.03</i>
7	SDD	0.9988	0.0957	0.111	0.9988	1.22	1.62

^aThe best-performing basis set and its data are in italics.

dependence on the choices of functionals and basis sets. For instance, the r^2 value of B3LYP/DGDZVP was 0.9995, while that of M06-2X/SDD was 0.9844. In addition to show good results obtained from ω B97XD (Entry 9), the calculations confirmed that B3LYP, B3PW91 and mPW1PW91 functionals and DGDZVP basis set were consistently provided high accuracy results.

3.3. The use of recommended methods for biaryls 2 and 3

For validation of optimization methods, selected functionals B3LYP, B3PW91 and mPW1PW91 and basis set DGDZVP2 were employed for optimizing biaryls 2 and 3. The biaryl conformations of optimized

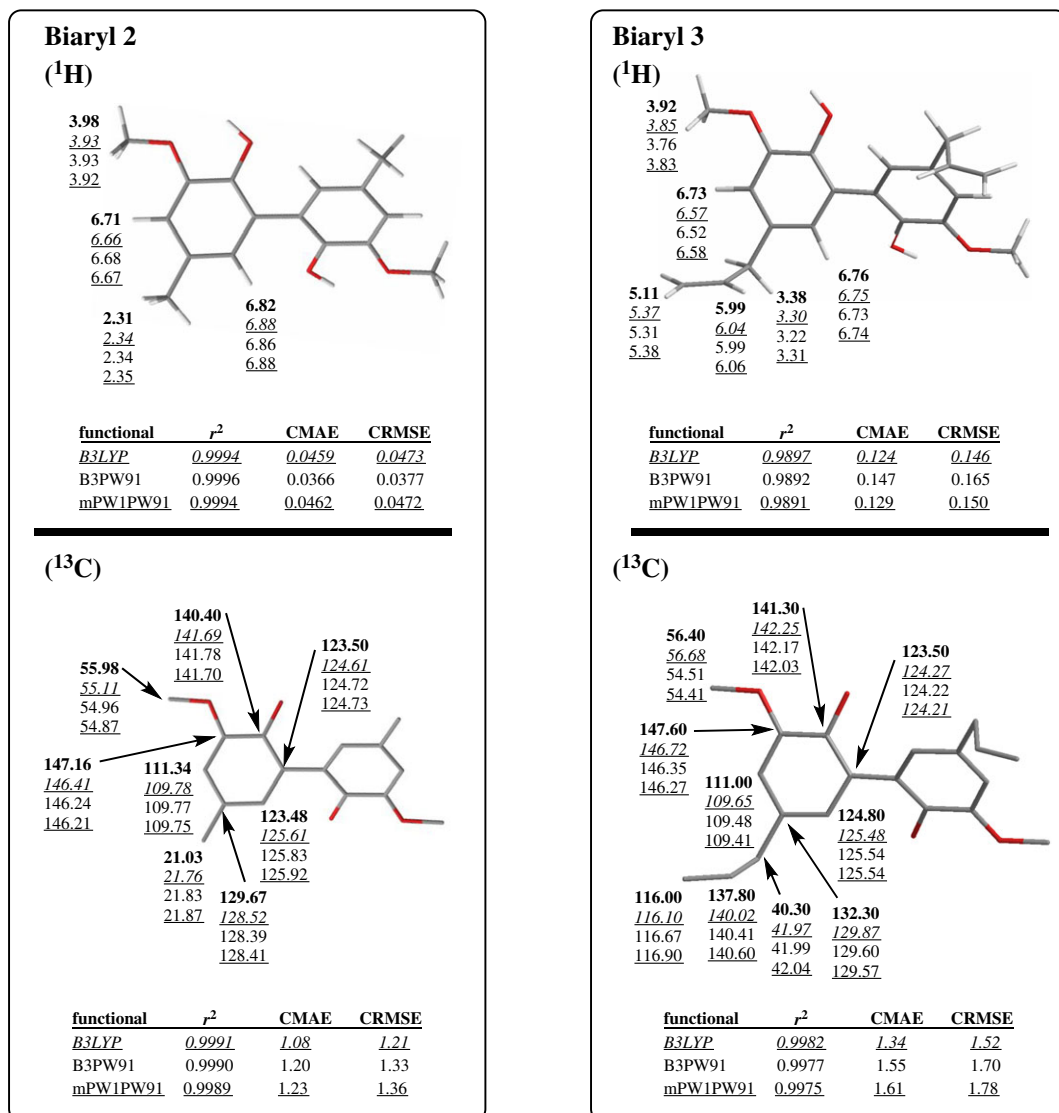


Figure 6. Two biaryls **2** and **3** with chemical shift data and statistical parameters for geometries optimized at IEFPCM/B3LYP/DGDZVP2 displayed in underlined italics, IEFPCM/B3PW91/DGDZVP2 displayed in normal text, IEFPCM/mPW1PW91/DGDZVP2 displayed in underlined normal text, and experimental chemical shifts displayed in **bold**.

Table 6. ¹H/¹³C Results using the combination of 10 functionals and three basis sets. Highest accurate data for ¹H results were in italics.

entry	functional	cc-pVDZ		DGDZVP		SDD	
		r ² (¹ H)	r ² (¹³ C)	r ² (¹ H)	r ² (¹³ C)	r ² (¹ H)	r ² (¹³ C)
1 ^a	B3LYP	<i>0.9994</i>	0.9992	<i>0.9995</i>	0.9995	<i>0.9988</i>	0.9988
2	B3PW91	<i>0.9932</i>	0.9992	<i>0.9975</i>	0.9992	<i>0.9908</i>	0.9988
3	BPV86	0.9848	0.9993	0.9943	0.9992	0.9892	0.9981
4	CAM-B3LYP	0.9894	0.9993	0.9945	0.9993	0.9881	0.9988
5	HSEH1PBE	0.9898	0.9993	0.9941	0.9993	0.9875	0.9988
6	mPW1PW91	<i>0.9922</i>	0.9992	<i>0.9959</i>	0.9992	<i>0.9898</i>	0.9988
7	PBEPBE	0.9890	0.9993	0.9942	0.9991	0.9857	0.9980
8	TPSSTPSS	0.9895	0.9994	0.9938	0.9992	0.9885	0.9980
9	ωB97XD	<i>0.9949</i>	0.9990	<i>0.9979</i>	0.9992	<i>0.9938</i>	0.9987
10	M06-2X	0.9718	0.9992	0.9922	0.9992	0.9844	0.9987

^aThese previous calculation results were included for comparison purposes. Highest accurate data for ¹H results are in italics.

structures were adopted with the biaryl dihedral angle values (α) of $-124^\circ \pm 2^\circ$ (electronic supplementary material, table S3). NMR calculations were performed and the chemical shift data and statistical parameters are shown in figure 6. In general, the calculated results were observed with low associated errors (CMAEs ≤ 0.147 ppm and CRMSEs ≤ 0.165 ppm for ^1H ; CMAEs ≤ 1.61 ppm and CRMSEs ≤ 1.78 ppm for ^{13}C) and strong linear correlations ($r^2 \geq 0.9891$ for ^1H and $r^2 \geq 0.9975$ for ^{13}C). ^1H chemical shifts of biaryl **3** were observed with higher errors than biaryl **2**. This could be a result of the broad, overlapping signals of an allyl proton NMR spectrum affecting the accurate assignments of experimental data. Large deviations were consistently observed for atoms **H9** ($|\Delta\delta_{\text{max}}| = 0.269$ ppm), **H5** ($|\Delta\delta_{\text{max}}| = 0.212$ ppm), **C8** ($|\Delta\delta_{\text{max}}| = 2.80$ ppm) and **C6** ($|\Delta\delta_{\text{max}}| = 2.44$ ppm) of biaryl **3** and **C1** ($|\Delta\delta_{\text{max}}| = 2.44$ ppm) of biaryl **2**, which are sp^2 -hybridized centres or are bound to these ones.

4. Conclusion

The evaluation of 12 different density functional methods and seven basis sets for geometry optimization on the accuracy of ^1H and ^{13}C chemical shift calculations for biaryls have been performed. The tested functionals showed relatively small impacts on ^1H and ^{13}C results in comparison to the basis sets. Solvent incorporation during the geometry optimization step was necessary for significantly improving ^1H chemical shift calculation accuracy, while it has little effect on ^{13}C chemical shifts. This result was expected as carbon nuclei are less exposed to solvent molecules than protons. The B3LYP, B3PW91, mPW1PW91 and ω B97XD of 12 density functional methods were the most effective ones; among seven basis sets, DGDZVP2 and 6-31G(d,p) generated the optimized geometry with excellent ^1H and ^{13}C results. The combinations of these functionals and basis sets resulted in high accuracy with CMAE values as low as 0.0329 ppm (0.76%) and 0.888 ppm (0.58%) for ^1H and ^{13}C , respectively, which would allow meaningful predictions. The consistency of these functionals and basis set were validated on the NMR calculations of biaryls **2** and **3**. Computed $^1\text{H}/^{13}\text{C}$ results were well correlated with the experimental data. Given the high degrees of accuracy achieved in the $^1\text{H}/^{13}\text{C}$ chemical shift calculations, these results have provided insights into the conformation of the above biaryl linkage systems, which may be unavailable or not forthcoming using experimental NMR techniques. This work can also be useful for assisting in the full $^1\text{H}/^{13}\text{C}$ NMR assignments of similar biaryls.

Data accessibility. The data are provided in the electronic supplementary material [43].

Competing interests. I declare I have no competing interests.

Funding. This project was supported by the International Foundation for Science (IFS), Stockholm, Sweden, through a grant no. 3-I-E-6576-1 to TTN.

References

1. Nguyen TT. 2019 Traceless point-to-axial chirality exchange in the atropselective synthesis of biaryls/heterobiaryls. *Org. Biomol. Chem.* **17**, 6952–6963. (doi:10.1039/c9ob01304k)
2. Nguyen TT. 2020 The use of chiral ortho-auxiliaries/substituents and remote stereogenic centers in atropselective biaryl synthesis. *Eur. J. Org. Chem.* **2020**, 147–155. (doi:10.1002/ejoc.201901557)
3. Ralph J, Lapierre C, Boerjan W. 2019 Lignin structure and its engineering. *Curr. Opin. Biotechnol.* **56**, 240–249. (doi:10.1016/j.copbio.2019.02.019)
4. Ragauskas AJ *et al.* 2014 Lignin valorization: improving lignin processing in the biorefinery. *Science* **344**, 1246843. (doi:10.1126/science.1246843)
5. Sheng H *et al.* 2017 (–)ESI/CAD MSn procedure for sequencing lignin oligomers based on a study of synthetic model compounds with β -O-4 and 5-5 linkages. *Anal. Chem.* **89**, 13 089–13 096. (doi:10.1021/acs.analchem.7b01911)
6. Salam MA, Arora P, Ojagh H, Cheah YW, Olsson L, Creaser D. 2020 Sustainable energy & fuels interdisciplinary research for the development of sustainable energy technologies rsc.li/sustainable-energy NiMoS on alumina-USY zeolites for hydrotreating lignin dimers: effect of support acidity and cleavage of C–C bonds †. (doi:10.1039/c9se00507b)
7. Nakamura T, Kawamoto H, Saka S. 2008 Pyrolysis behavior of Japanese cedar wood lignin studied with various model dimers. *J. Anal. Appl. Pyrolysis* **81**, 173–182. (doi:10.1016/j.jaap.2007.11.002)
8. Santana FPR *et al.* 2020 Dehydrodieugenol improved lung inflammation in an asthma model by inhibiting the STAT3/SOCS3 and MAPK pathways. *Biochem. Pharmacol.* **180**, 114175. (doi:10.1016/j.bcp.2020.114175)
9. Watts HD, Mohamed MNA, Kubicki JD. 2011 Comparison of multistandard and TMS-standard calculated NMR shifts for coniferyl alcohol and application of the multistandard method to lignin dimers. *J. Phys. Chem. B* **115**, 1958–1970. (doi:10.1021/jp110330q)
10. Nguyen TT, Le PQ, Helminen J, Sipilä J. 2021 The ^1H and ^{13}C chemical shifts of 5–5 lignin model dimers: an evaluation of DFT functionals. *J. Mol. Struct.* **1226**, 129300. (doi:10.1016/j.molstruc.2020.129300)
11. Iron MA. 2017 Evaluation of the factors impacting the accuracy of ^{13}C NMR chemical shift predictions using density functional theory - the advantage of long-range corrected functionals. *J. Chem. Theory Comput.* **13**, 5798–5819. (doi:10.1021/acs.jctc.7b00772)
12. Lodewyk MW, Siebert MR, Tantillo DJ. 2012 Computational prediction of ^1H and ^{13}C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem. Rev.* **112**, 1839–1862. (doi:10.1021/cr200106v)
13. Benassi E. 2017 Benchmarking of density functionals for a soft but accurate prediction

- and assignment of ^1H and ^{13}C NMR chemical shifts in organic and biological molecules. *J. Comput. Chem.* **38**, 87–92. (doi:10.1002/jcc.24521)
14. Kupka T, Stachów M, Nieradka M, Kaminsky J, Pluta T. 2010 Convergence of nuclear magnetic shieldings in the Kohn–Sham limit for several small molecules. *J. Chem. Theory Comput.* **6**, 1580–1589. (doi:10.1021/CT100109J)
 15. Casabianca LB. 2020 Calculating nuclear magnetic resonance chemical shifts in solvated systems. *Magn. Reson. Chem.* **58**, 611–624. (doi:10.1002/MRC.4994)
 16. Jain R, Bally T, Rablen PR. 2009 Calculating accurate proton chemical shifts of organic molecules with density functional methods and modest basis sets. *J. Org. Chem.* **74**, 4017–4023. (doi:10.1021/jo900482q)
 17. Toomsalu E, Burk P. 2015 Critical test of some computational methods for prediction of NMR ^1H and ^{13}C chemical shifts. *J. Mol. Model.* **21**, 1–21. (doi:10.1007/s00894-015-2787-x)
 18. Rablen PR, Pearlman SA, Finkbiner J. 1999 A comparison of density functional methods for the estimation of proton chemical shifts with chemical accuracy. *J. Phys. Chem. A* **103**, 7357–7363. (doi:10.1021/jp9916889)
 19. Zhang Y, Xu X, Yan Y. 2008 Systematic investigation on the geometric dependence of the calculated nuclear magnetic shielding constants. *J. Comput. Chem.* **29**, 1798–1807. (doi:10.1002/jcc.20941)
 20. Frisch MJ *et al.* 2013 Gaussian 09, Revision D.01.
 21. Tomasi J, Mennucci B, Cancès E. 1999 The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct.* **464**, 211–226. (doi:10.1016/S0166-1280(98)00553-3)
 22. Becke AD. 1993 Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652. (doi:10.1063/1.464913)
 23. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C. 1993 Erratum: atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **48**, 4978. (doi:10.1103/PhysRevB.48.4978.2)
 24. Perdew JP. 1986 Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822–8824. (doi:10.1103/PhysRevB.33.8822)
 25. Yanai T, Tew DP, Handy NC. 2004 A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **393**, 51–57. (doi:10.1016/j.cplett.2004.06.011)
 26. Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC. 1998 Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **109**, 6264–6271. (doi:10.1063/1.477267)
 27. Heyd J, Scuseria GE, Ernzerhof M. 2003 Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207–8215. (doi:10.1063/1.1564060)
 28. Perdew JP, Wang Y. 1992 Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **45**, 13 244–13 249. (doi:10.1103/PhysRevB.45.13244)
 29. Adamo C, Barone V. 1998 Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mPW and mPW1PW models. *J. Chem. Phys.* **108**, 664–675. (doi:10.1063/1.475428)
 30. Perdew JP, Burke K, Ernzerhof M. 1996 Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868. (doi:10.1103/PhysRevLett.77.3865)
 31. Tao J, Perdew JP, Staroverov VN, Scuseria GE. 2003 Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **91**, 146401. (doi:10.1103/PhysRevLett.91.146401)
 32. Chai J Da, Head-Gordon M. 2008 Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **128**, 084106. (doi:10.1063/1.2834918)
 33. Zhao Y, Truhlar DG. 2008 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.* **120**, 215–241. (doi:10.1007/s00214-007-0310-x)
 34. Binkley JS, Pople JA, Hehre WJ. 1980 Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.* **102**, 939–947. (doi:10.1021/ja00523a008)
 35. Dunning TH. 1989 Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90**, 1007–1023. (doi:10.1063/1.456153)
 36. Sosa C, Andzelm J, Elkin BC, Wimmer E, Dobbs KD, Dixon DA. 1992 A local density functional study of the structure and vibrational frequencies of molecular transition-metal compounds. *J. Phys. Chem.* **96**, 6630–6636. (doi:10.1021/j100195a022)
 37. Andrae D, Häußermann U, Dolg M, Stoll H, Preuß H. 1990 Energy-adjusted ab initio pseudopotentials for the second and third row transition elements. *Theor. Chim. Acta* **77**, 123–141. (doi:10.1007/BF01114537)
 38. Russell WR, Scobbie L, Chesson A. 2005 Structural modification of phenylpropanoid-derived compounds and the effects on their participation in redox processes. *Bioorg. Med. Chem.* **13**, 2537–2546. (doi:10.1016/j.bmc.2005.01.047)
 39. Marques FA, Simonelli F, Oliveira ARM, Gohr GL, Leal PC. 1998 Oxidative coupling of 4-substituted 2-methoxy phenols using methyltributylammonium permanganate in dichloromethane. *Tetrahedron Lett.* **39**, 943–946. (doi:10.1016/S0040-4039(97)10665-7)
 40. Albrecht M, Schneider M. 2000 The preparation of rigid linear Di- (and Tri-)catechol derivatives. *Synthesis (Stuttg.)* **2000**, 1557–1560. (doi:10.1055/s-2000-7604)
 41. Kumar N *et al.* 2019 Integrative natural medicine inspired graphene nanovehicle-benzoxazine derivatives as potent therapy for cancer. *Mol. Cell. Biochem.* **454**, 123–138. (doi:10.1007/s11010-018-3458-x)
 42. Bortolomeazzi R, Verardo G, Liessi A, Callea A. 2010 Formation of dehydrodiisoeugenol and dehydrodieugenol from the reaction of isoeugenol and eugenol with DPPH radical and their role in the radical scavenging activity. *Food Chem.* **118**, 256–265. (doi:10.1016/j.foodchem.2009.04.115)
 43. Nguyen TT. 2021 $^1\text{H}/^{13}\text{C}$ Chemical shift calculations for biaryls: DFT approaches to geometry optimization. Figshare.