

N. J. R. van Eikema Hommes · Timothy Clark

Regression formulae for *ab initio* and density functional calculated chemical shifts

Received: 15 October 2004 / Accepted: 5 November 2004 / Published online: 24 March 2005
© Springer-Verlag 2005

Abstract Linear regression formulae are given for converting ^1H and ^{13}C magnetic shielding constants calculated at common *ab initio* and density functional theory levels of calculation into chemical shifts relative to tetramethylsilane. Accuracies of roughly ± 2.2 ppm (^{13}C) and ± 0.15 ppm (^1H) or better are found for the training set for most levels. The highest level calculations do not always give better results than economical standard calculations.

Keywords NMR chemical shifts · Density functional calculations · Linear regression formulae

Introduction

Once the gauge problem in calculating magnetic shielding within *ab initio* or density functional theory (DFT) had been solved, [1, 2] rapid progress was made in developing techniques such as gauge invariant (or including) atomic orbitals (GIAO) [3, 4] or individual gauge for localized orbitals (IGLO) [5] that are able to calculate magnetic properties efficiently and relatively accurately. Extensions of the original Hartree-Fock (HF) formalisms to second-order Møller-Plesset (MP2) [6] and DFT [4, 7] calculations improved the accuracy of the calculated values. Magnetic properties have not only been used for calculating nuclear magnetic resonance (NMR) chemical shifts, but have also become very popular in the so-called nucleus-independent chemical

shift (NICS) technique, [8] which has been used extensively to investigate the aromaticity of organic [9] and inorganic [10] molecules. However, determining chemical shifts for direct comparison with experimental spectra remains a major use of *ab initio* or DFT magnetic shielding calculations. Generally, chemical shifts on the δ -scale are calculated by taking the difference between the calculated shielding and that found for a reference molecule such as tetramethylsilane (TMS). However, as Chesnut [11] as well as Forsyth and Sebag [12] have pointed out, this is not the best procedure and better results can be obtained by setting up a linear regression equation between calculated shieldings and experimental chemical shifts. We now report parameters for such regression equations for 18 calculational levels commonly used with the Gaussian series of programs [13] and give root mean-square deviations for each level based on our training dataset. The equations allow quite accurate chemical shifts to be calculated at modest levels of theory.

Methods

All calculations used Gaussian 98 [13] Geometries were optimized at the B3LYP [14] hybrid DFT level using the Becke three-parameter exchange treatment [15] with the Lee–Yang–Parr correlation functional [16] The 6-31G(d) [17, 18] and 6-311 + G(d,p) [19–21] basis sets were used for geometry optimizations. For the shielding calculations, B3LYP, HF and MP2 calculations were performed with the two basis sets used for the geometry optimizations and with D95* [22] and TZVP. [23] The large dataset (120 compounds, 381 ^{13}C and 316 ^1H chemical shifts) was taken from a compilation of experimental spectra recorded in CDCl_3 solution [24] and the small dataset (16 compounds, 20 gas phase ^{13}C chemical shifts) was taken from Jameson and Jameson [25]. Regression equations were calculated both with respect to gas-phase chemical shifts as well as to those determined in chloroform solution. For molecules with

Electronic Supplementary Material Supplementary material is available for this article at <http://dx.doi.org/10.1007/s00894-004-0223-8>

N. J. R. van Eikema Hommes · T. Clark (✉)
Computer-Chemie-Centrum, Universität Erlangen-Nürnberg,
Nägelsbachstraße 25, 91052 Erlangen, Germany
E-mail: clark@chemie.uni-erlangen.de
Tel.: +49-09131-8522948
Fax: +49-09131-8526565

internal rotations, only the most stable conformations were considered unless otherwise noted in the Tables.

Results

The experimental and calculated chemical shifts obtained at B3LYP/6-31G(d)//6-31G(d) are shown in Table 1. The detailed results for the other methods are available as supplementary material and will only be given here as summaries. The B3LYP/6-31G(d)//6-31G(d) results are typical and suffice to show the effects that will be discussed below.

These results were derived from simple linear regression equations of the form

$$\delta = \delta^0 + \kappa \cdot \sigma, \quad (1)$$

where δ is the calculated chemical shift on the δ -scale relative to TMS, σ is the calculated shielding and κ and δ^0 are the slope and the intercept obtained from the linear regression, respectively. This large dataset gives the regression equations, mean signed and unsigned errors (MSE and MUE, respectively) shown in Table 2 using B3LYP with the 6-31G(d) and 6-311+G(d,p) basis sets for ^{13}C and ^1H chemical shifts.

There is little to choose between the two levels of calculation and, perhaps surprisingly, the extra polarization functions on hydrogen do not improve the performance for ^1H chemical shifts. Thus, we can conclude that B3LYP/6-31G(d)//B3LYP/6-31G(d) is an adequate level for calculating chemical shifts for both ^{13}C and ^1H economically and reliably.

All δ^0 and κ values for ^{13}C chemical shifts are shown in Table 3. For comparison, these values for chemical shifts calculated as a simple difference to TMS are also given.

Discussion

Trends at B3LYP/6-31G(d)

The linear regression equation given in Table 2 for B3LYP/6-31G(d)//B3LYP/6-31G(d) gives remarkably small (0.14 and 2.16 ppm for ^1H and ^{13}C , respectively) mean unsigned errors in the calculated chemical shifts. The largest errors in each direction for ^{13}C are 13.56 ppm (bromoethane) and -12.05 ppm (N-methylacetamide). Figure 1 shows a plot of the experimental versus calculated values for the large dataset at B3LYP/6-31G(d)//B3LYP/6-31G(d) and Fig. 2 shows a histogram of the residuals between experiment and the regression-corrected calculated values.

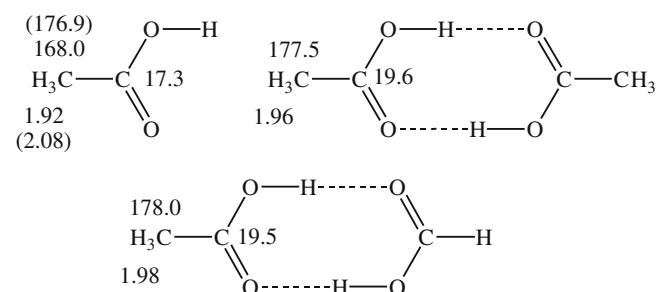
The standard deviation is calculated to be 3.29 ppm (0.44 ppm for ^1H) and Figure 2 suggests that the distribution of errors is roughly normal, so that our results suggest that B3LYP/6-31G(d) results for ^{13}C chemical shifts are accurate to ± 3.3 ppm to a 67% level of

certainty. Correspondingly, ^1H chemical shifts are accurate to ± 0.4 ppm to the same level of certainty.

Problem molecules

Carboxylic acids

The calculated chemical shifts for acetic acid (experimental values shown in parentheses for the monomer) are shown in Chart 1.



The agreement between experiment and the calculated values for the monomer is poor, as might be expected. The calculations for the acetic-acid dimer, however, give good agreement with experiment. Using dimers would, however, lead to very extensive calculations for large carboxylic acids, so that we investigated the alternative of using the “mixed dimer”, the complex between acetic acid and formic acid, which also gives good results for acetic acid in comparison with experiment. We have reported calculations for dimers and mixed dimers in Table 1, which demonstrate that the mixed dimer model is a good approximation that can be used usefully to calculate the chemical shifts of carboxylic acids.

Purine: the tautomer problem

The results for purine, calculated both as the $\text{N}^7\text{-H}$ and as the $\text{N}^9\text{-H}$ tautomer, show larger than average deviations from experiment. However, the exact nature of the purine tautomeric equilibrium in solution has still

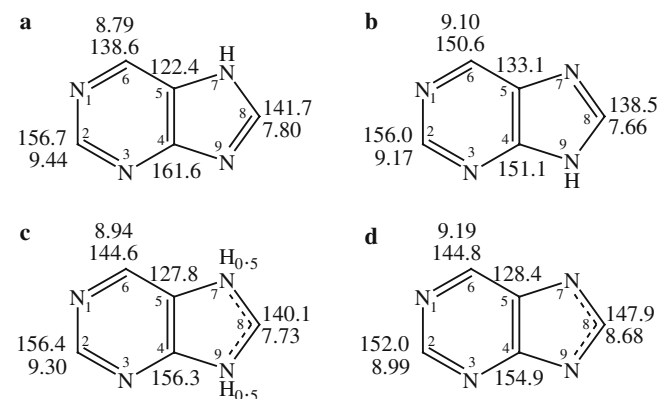


Table 1 Experimental NMR chemical shifts, calculated (B3LYP/6-31G(d)) magnetic shieldings and chemical shifts given by the regression equations for the large training dataset

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
Methane	-2.3	0.23	193.02	31.89	-6.17	-3.87	0.36	0.13
Ethane	5.7	0.86	180.37	31.21	7.39	1.69	1.05	0.19
Propane	15.4	0.91	172.43	31.18	15.90	0.50	1.08	0.17
	15.9	1.33	170.33	30.68	18.15	2.25	1.58	0.25
<i>n</i> -Butane	13.0	0.90	173.40	31.18	14.86	1.86	1.08	0.18
	24.8	1.23	162.12	30.74	26.94	2.14	1.52	0.29
Isobutane	24.3	0.88	165.36	31.19	23.47	-0.83	1.07	0.19
	25.0	1.77	163.67	30.24	25.28	0.28	2.02	0.25
Neopentane	27.4	0.93	158.99	31.13	30.30	2.90	1.13	0.20
	31.4		160.08		29.13	-2.27		
Ethylene	123.3	5.28	73.62	26.72	121.77	-1.53	5.56	0.28
1-Butene	112.1	4.87	83.49	27.18	111.19	-0.91	5.10	0.23
	139.0	4.95	57.98	27.17	138.52	-0.48	5.11	0.16
	26.0	5.79	163.28	26.02	25.70	-0.30	6.27	0.48
		2.00	178.63	30.06	9.25		2.20	0.20
		1.01		31.10			1.16	0.15
<i>cis</i> -2-Butene	123.4	5.45	70.44	26.66	125.17	1.77	5.62	0.17
	10.9	1.60	175.89	30.62	12.19	1.29	1.64	0.04
<i>trans</i> -2-Butene	124.8	5.43	69.12	26.66	126.59	1.79	5.62	0.19
	16.4	1.63	170.17	30.59	18.32	1.92	1.67	0.04
1,3-Butadiene	117.5	5.06	77.13	27.14	118.01	0.51	5.14	0.08
	138.1	5.16	59.24	27.08	137.17	-0.93	5.20	0.04
		6.27		25.92			6.37	0.10
2,3-Dimethylbutadiene	111.3	4.90	79.63	27.29	115.33	4.03	4.99	0.09
	142.1	4.98	53.56	27.24	143.26	1.16	5.04	0.06
		1.87	168.26	30.36	20.36		1.90	0.03
Acetylene	71.9	1.80	128.80	30.98	62.64	-9.26	1.28	-0.52
Propyne	66.9	1.80	128.83	30.88	62.61	-4.29	1.38	-0.42
	79.2	1.80	120.88	30.59	71.13	-8.07	1.67	-0.13
			185.60		1.78			
2-Butyne	73.6	1.68	121.97	30.61	69.96	-3.64	1.65	-0.03
			184.88		2.56			
Cyclopropane	-3.8	0.22	189.24	31.86	-2.12	1.68	0.39	0.17
Cyclobutane	22.1	1.96	165.12	30.18	23.73	1.63	2.08	0.12
Cyclopentane	25.3	1.51	161.72	30.51	27.37	2.07	1.75	0.24
Cyclohexane	26.6	1.44	161.91	30.65	27.17	0.57	1.61	0.17
Cyclopentene	130.8	5.60	64.21	26.46	131.85	1.05	5.83	0.23
	32.8	2.28	155.66	29.88	33.86	1.06	2.38	0.10
	23.3	1.90	164.14	30.44	24.78	1.48	1.82	-0.08
Cyclopentadiene	133.4	6.50	62.97	25.78	133.18	-0.22	6.51	0.01
	133.0	6.41	62.29	25.90	133.91	0.91	6.39	-0.02
	42.2	2.90	149.42	29.52	40.55	-1.65	2.75	-0.15
Cyclohexene	127.4	5.59	67.49	26.46	128.34	0.94	5.83	0.24
	25.4	1.96	162.85	30.19	26.16	0.76	2.07	0.11
	23.0	1.65	165.48	30.57	23.34	0.34	1.69	0.04
Cycloheptatriene	134.1	6.50	64.86	25.65	131.15	-2.95	6.64	0.14
	129.8	6.09	69.52	25.94	126.16	-3.64	6.35	0.26
	123.3	5.26	78.80	27.13	116.22	-7.08	5.15	-0.11
	28.8	2.22	160.28	29.92	28.91	0.11	2.34	0.12
Cyclooctatetraene	131.5	5.69	60.82	26.64	135.48	3.98	5.64	-0.05
Norbornane	29.2	1.21	150.96	30.86	38.90	9.70	1.40	0.19
	30.6	1.49	157.99	30.59	31.37	0.77	1.67	0.18
	37.3	1.18	150.69	30.88	39.19	1.89	1.38	0.20
		2.20		30.01			2.25	0.05
Adamantane	39.7	1.80	151.06	30.33	38.79	-0.91	1.93	0.13
	28.5	1.90	158.71	30.38	30.60	2.10	1.88	-0.02
Benzene	128.5	7.26	68.62	24.97	127.12	-1.38	7.33	0.07
Naphthalene	128.0	7.66	67.76	24.56	128.05	0.05	7.74	0.08
	126.0	7.30	70.54	24.79	125.07	-0.93	7.51	0.21
	133.7		63.67		132.43	-1.27		
Azulene	135.2	7.92	58.86	24.37	137.58	2.38	7.93	0.01
	119.7	7.39	78.82	24.97	116.20	-3.50	7.33	-0.06
	137.4	8.32	59.70	24.04	136.68	-0.72	8.26	-0.06
	123.9	7.11	73.95	25.08	121.41	-2.49	7.21	0.10
	137.4	7.57	59.64	24.50	136.75	-0.65	7.80	0.23
	140.1		58.03		138.47	-1.63		
Anthracene	128.1	7.91	67.11	24.34	128.74	0.64	7.96	0.05

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
Phenanthrene	125.3	7.39	70.64	24.77	124.96	-0.34	7.53	0.14
	126.2	8.31	68.25	24.04	127.52	1.32	8.26	-0.05
	131.8		65.22		130.77	-1.03		
	122.4	8.93	72.88	23.66	122.56	0.16	8.64	-0.29
	126.3	7.88	70.05	24.59	125.59	-0.71	7.71	-0.17
	126.3	7.82	70.19	24.69	125.44	-0.86	7.61	-0.21
	128.3	8.12	66.99	24.54	128.87	0.57	7.76	-0.36
	126.6	7.71	68.03	24.68	127.76	1.16	7.62	-0.09
	130.1		64.39		131.66	1.56		
Toluene	131.9		65.42		130.55	-1.35		
	137.7	7.17	58.99	25.13	137.44	-0.26	7.16	-0.01
	129.3	7.17	67.62	25.04	128.20	-1.10	7.25	0.08
	128.5	7.17	68.41	25.16	127.35	-1.15	7.13	-0.04
	125.6	2.32	71.01	29.98	124.56	-1.04	2.28	-0.04
Benzocyclohexane	21.3		167.43		21.25	-0.05		
	125.5	7.01	67.16	25.29	128.69	3.19	7.00	-0.01
	129.0	6.93	70.68	25.21	124.92	-4.08	7.08	0.15
	136.8	2.85	59.75	29.46	136.63	-0.17	2.81	-0.04
	29.5	1.60	158.51	30.44	30.81	1.31	1.82	0.22
Styrene	23.6		164.48		24.41	0.81		
	136.1	7.32	60.30	24.99	136.04	-0.06	7.31	-0.01
	126.7	7.23	69.68	25.01	125.99	-0.71	7.29	0.06
	126.7	7.16	68.20	25.13	127.57	0.87	7.16	0.00
	125.0	6.26	68.85	25.81	126.88	1.88	6.48	0.22
	135.5	4.80	58.96	27.03	137.47	1.97	5.25	0.45
	112.0	5.30	83.29	26.44	111.41	-0.59	5.85	0.55
Phenylacetylene	122.4	7.41	72.90	24.86	122.54	0.14	7.44	0.03
	132.3	7.24	62.90	25.01	133.25	0.95	7.29	0.05
	128.9	7.25	68.60	25.02	127.15	-1.75	7.28	0.03
	128.3	3.10	68.54	29.50	127.21	-1.09	2.77	-0.33
	84.6		115.05		77.38	-7.22		
Oxirane Tetrahydrofuran	78.3		115.83		76.54	-1.76		
	39.5	2.54	149.93	29.75	40.00	0.50	2.51	-0.03
	68.4	3.75	122.82	28.39	69.05	0.65	3.88	0.13
Furan	26.5	1.85	160.97	30.46	28.17	1.67	1.80	-0.05
	143.0	7.38	55.95	25.01	140.70	-2.30	7.29	-0.09
Dioxane	109.9	6.30	85.86	26.10	108.65	-1.25	6.19	-0.11
	67.6	3.71	127.70	28.28	63.82	-3.78	3.99	0.28
	28.7	1.61	170.72	30.75	17.73	-10.97	1.51	-0.10
Aziridine Pyrrolidine	47.1	2.75	141.87	29.43	48.64	1.54	2.84	0.09
	25.7	1.59	164.29	30.59	24.62	-1.08	1.67	0.08
Pyrrole	118.0	6.62	81.82	25.89	112.98	-5.02	6.40	-0.22
	107.7	6.05	87.51	26.25	106.88	-0.82	6.04	-0.01
	47.7	2.13	144.16	30.10	46.19	-1.51	2.16	0.03
N-Methylpiperidine	57.2	2.26	135.44	29.90	55.53	-1.67	2.36	0.10
	26.4	1.44	162.70	30.62	26.32	-0.08	1.64	0.20
	26.4	1.40	164.60	30.79	24.29	-2.11	1.47	0.07
	149.8	8.60	46.33	23.55	151.01	1.21	8.75	0.15
	123.6	7.25	73.24	25.19	122.17	-1.43	7.10	-0.15
Pyridine	135.7	7.64	62.41	24.81	133.78	-1.92	7.49	-0.15
	149.1	8.62	46.39	23.59	150.94	1.84	8.71	0.09
	120.7	7.23	75.75	25.37	119.48	-1.22	6.92	-0.31
	136.1	7.76	61.64	24.90	134.60	-1.50	7.40	-0.36
	121.8	7.29	73.84	25.31	121.53	-0.27	6.98	-0.31
2-Ethylpyridine	163.4	2.86	32.15	29.52	166.20	2.80	2.75	-0.11
	31.4	1.26	155.62	30.86	33.91	2.51	1.40	0.14
	13.8		172.15		16.20	2.40		
	158.0	9.26	35.38	22.84	162.74	4.74	9.47	0.21
	156.4	8.78	40.31	23.59	157.46	1.06	8.71	-0.07
Thiirane Tetrahydrothiophene	121.4	7.36	74.84	25.31	120.46	-0.94	6.98	-0.38
	18.9	2.30	168.01	29.98	20.63	1.73	2.28	-0.02
	31.2	2.80	153.35	29.48	36.34	5.14	2.79	-0.01
Sulfolane	31.4	1.90	157.22	30.37	32.19	0.79	1.89	-0.01
	51.2	3.00	134.18	29.66	56.88	5.68	2.61	-0.39
Thiophene	22.8	2.23	164.66	30.27	24.22	1.42	1.99	-0.24
	124.9	7.20	66.13	25.29	129.79	4.89	7.00	-0.20
	126.4	6.96	71.66	25.31	123.87	-2.53	6.98	0.02

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
2-Methylthiophene	122.6	7.03	66.39	25.46	129.51	6.91	6.83	-0.20
	126.4	6.87	72.28	25.63	123.20	-3.20	6.66	-0.21
	124.7	6.72	72.00	25.71	123.50	-1.20	6.58	-0.14
	139.0	2.48	49.69	29.82	147.41	8.41	2.44	-0.04
	14.7		172.71		15.60	0.90		
3-Methylthiophene	124.7	7.01	65.37	25.37	130.61	5.91	6.92	-0.09
	128.8	6.74	69.87	25.45	125.79	-3.01	6.84	0.10
	136.8	6.70	60.72	25.75	135.59	-1.21	6.54	-0.16
	120.1	2.18	69.08	30.01	126.63	6.53	2.25	0.07
			173.08		15.20			
Quinoline	150.0	8.81	46.92	23.28	150.38	0.38	9.03	0.22
	120.8	7.26	75.45	25.04	119.81	-0.99	7.25	-0.01
	135.7	8.00	61.20	24.40	135.07	-0.63	7.90	-0.10
	128.0	7.68	68.19	24.57	127.59	-0.41	7.73	0.05
	127.6	7.43	68.15	24.71	127.63	0.03	7.59	0.16
	126.3	7.61	69.82	24.53	125.84	-0.46	7.77	0.16
	129.2	8.05	67.55	24.16	128.27	-0.93	8.14	0.09
	129.2		64.67		131.36	2.16		
	148.1		47.82		149.41	1.31		
Isoquinoline	152.2	9.15	43.77	23.03	153.75	1.55	9.28	0.13
	142.7	8.45	51.64	23.53	145.32	2.62	8.77	0.32
	120.2	7.50	75.18	24.87	120.10	-0.10	7.43	-0.07
	135.5	7.71	61.48	24.57	134.77	-0.73	7.73	0.02
	126.2	7.57	69.20	24.62	126.50	0.30	7.68	0.11
	130.1	7.50	66.91	24.70	128.96	-1.14	7.60	0.10
	127.0	7.87	69.75	24.44	125.91	-1.09	7.86	-0.01
	127.3		67.71		128.10	0.80		
	128.5		68.20		127.57	-0.93		
N ⁷ -H purine	2							
	4							
	5							
	6							
	8							
N ⁹ -H purine	2							
	4							
	5							
	6							
	8							
N ⁷ -H and N ⁹ -H purine, 1:1	2							
	4							
	5							
	6							
	8							
Fluoroethane	78.0	4.36	113.29	27.63	79.26	1.26	4.65	0.29
	13.3	1.24	172.33	30.96	16.00	2.70	1.30	0.06
Chloroethane	38.7	3.47	144.78	28.76	45.52	6.82	3.51	0.04
	17.5	1.33	170.39	30.87	18.08	0.58	1.39	0.06
Bromoethane	26.9	3.37	149.50	28.79	40.46	13.56	3.48	0.11
	19.0	1.66	169.22	30.74	19.33	0.33	1.52	-0.14
1,1-Dibromoethane		5.86	108.76	26.12	84.12		6.17	0.31
	34.8	2.47	155.16	30.19	34.40	-0.40	2.07	-0.40
Chloroethylene	124.9	6.33	62.95	26.15	133.20	8.30	6.14	-0.19
	116.0	5.38	78.50	26.88	116.54	0.54	5.40	0.02
		5.43		26.87			5.41	-0.02
Chlorobenzene	134.9	7.25	53.26	25.13	143.58	8.68	7.16	-0.09
	128.7	7.33	67.61	25.04	128.21	-0.49	7.25	-0.08
	129.5	7.31	67.44	25.15	128.39	-1.11	7.14	-0.17
	126.5		70.62		124.98	-1.52		
Nitromethane	57.1	4.28	129.79	28.27	61.58	4.48	4.00	-0.28
	70.7	4.29	122.15	28.13	69.77	-0.93	4.15	-0.14
	12.3	1.48	178.22	30.87	9.69	-2.61	1.39	-0.09
Nitrobenzene	148.0	8.20	48.26	23.99	148.94	0.94	8.31	0.11
	123.2	7.50	71.47	24.81	124.07	0.87	7.49	-0.01
	128.3	7.60	67.32	24.68	128.52	0.22	7.62	0.02
	134.5		62.55		133.63	-0.87		

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
Methanol	50.2	3.39	139.31	28.70	51.38	1.18	3.57	0.18
Ethanol	57.0	3.59	131.28	28.35	59.99	2.99	3.92	0.33
	17.6	1.18	172.13	31.02	16.22	-1.38	1.24	0.06
Ethanol (<i>gauche</i>)	57.0	3.59	131.32	28.48	59.94	2.94	3.79	0.20
	17.6	1.18	169.23	31.10	19.32	1.72	1.16	-0.02
2-Propanol	63.4	3.94	125.80	28.17	65.86	2.46	4.10	0.16
	25.1	1.16	163.76	31.08	25.19	0.09	1.18	0.02
2-Propanol (<i>gauche</i>)	63.4	3.94	125.95	28.00	65.70	2.30	4.28	0.34
	25.1	1.16	164.67	31.07	24.21	-0.89	1.19	0.03
2-Methyl-2-Propanol	68.4	1.22	122.35	30.96	69.55	1.15	1.30	0.08
	31.3		159.12		30.16	-1.14		
1-Methylcyclohexanol	26.0	1.50	162.48	30.71	26.56	0.56	1.55	0.05
	22.8	1.50	165.84	30.45	22.96	0.16	1.81	0.31
	39.7	1.50	150.70	30.69	39.18	-0.52	1.57	0.07
	69.0	1.20	121.31	31.13	70.67	1.67	1.13	-0.07
	29.5		157.11		32.31	2.81		
Phenol	155.1	6.70	41.95	25.79	155.70	0.60	6.50	-0.20
	115.7	7.14	82.34	25.10	112.42	-3.28	7.19	0.05
	130.1	6.81	67.03	25.48	128.83	-1.27	6.81	0.00
	121.4		76.86		118.30	-3.10		
Benzyl alcohol	140.8	7.30	56.32	24.88	140.30	-0.50	7.42	0.12
	126.8	7.30	70.75	24.98	124.84	-1.96	7.32	0.02
	128.2	7.30	68.65	25.08	127.09	-1.11	7.21	-0.09
	127.2	4.58	69.85	27.32	125.81	-1.39	4.96	0.38
	64.5		124.89		66.83	2.33		
Diethyl ether	65.2	3.36	129.07	28.63	62.35	-2.85	3.64	0.28
	14.6	1.16	174.87	31.18	13.28	-1.32	1.08	-0.08
Anisole	160.2	6.80	39.14	25.66	158.71	-1.49	6.63	-0.17
	114.1	7.17	82.74	25.05	112.00	-2.10	7.24	0.07
	129.5	6.84	67.36	25.43	128.47	-1.03	6.86	0.02
	120.7	3.78	76.51	28.54	118.67	-2.03	3.73	-0.05
	54.7		137.61		53.20	-1.50		
Methylamine	26.9	2.47	159.13	29.79	30.15	3.25	2.47	0.00
Ethylamine	35.9	2.74	150.52	29.50	39.37	3.47	2.77	0.03
	17.7	1.10	167.18	31.25	21.52	3.82	1.00	-0.10
Ethylamine (<i>gauche</i>)	35.9	2.74	150.60	29.37	39.29	3.39	2.90	0.16
	17.7	1.10	171.36	31.13	17.04	-0.66	1.13	0.03
Diethylamine	44.1	2.51	143.45	29.56	46.95	2.85	2.71	0.20
	15.4	1.03	172.67	31.08	15.64	0.24	1.18	0.15
Benzylamine	143.4	7.20	54.82	24.73	141.91	-1.49	7.57	0.37
	126.9	7.20	68.87	25.02	126.86	-0.04	7.28	0.08
	128.3	7.20	68.65	25.09	127.09	-1.21	7.20	0.00
	126.5	3.72	69.66	28.31	126.01	-0.49	3.96	0.24
	46.3		141.09		49.48	3.18		
N-Methylaniline	150.2	6.50	50.21	25.99	146.85	-3.35	6.30	-0.20
	112.3	7.00	84.50	25.20	110.11	-2.19	7.09	0.09
	129.2	6.60	67.52	25.68	128.30	-0.90	6.61	0.01
	116.7	2.78	79.29	29.54	115.69	-1.01	2.73	-0.05
	30.2		159.27		30.00	-0.20		
2-Aminopyridine	158.9	6.70	37.53	26.17	160.44	1.54	6.12	-0.58
	108.5	7.44	89.48	25.08	104.77	-3.73	7.21	-0.23
	137.5	6.60	60.31	25.90	136.03	-1.47	6.39	-0.21
	113.3	8.11	82.54	24.15	112.21	-1.09	8.15	0.04
	147.7		47.18		150.10	2.40		
4-Aminopyridine	149.9	8.00	45.95	24.01	151.41	1.51	8.29	0.29
	109.4	6.46	87.16	26.19	107.26	-2.14	6.10	-0.36
	154.9		47.43		149.83	-5.07		
Ethaneithiol		2.44	164.45	29.56	24.45		2.71	0.27
		1.31	174.88	30.99	13.27		1.27	-0.04
Ethaneithiol (<i>gauche</i>)		2.44	163.14	29.80	25.85		2.46	0.02
		1.31	168.27	30.93	20.35		1.33	0.02
Diethyl sulfide	26.5	2.49	162.74	29.79	26.28	-0.22	2.47	-0.02
	15.8	1.25	174.73	31.02	13.43	-2.37	1.24	-0.01
Dimethyl sulfoxide	40.8	2.62	148.49	30.17	41.55	0.75	2.09	-0.53
Thiophenol	130.5	7.20	56.96	25.39	139.62	9.12	6.90	-0.30
	129.1	7.10	70.59	25.15	125.01	-4.09	7.14	0.04
	128.7	7.00	67.90	25.35	127.90	-0.80	6.94	-0.06

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
Thioanisole	125.2		72.87		122.57	-2.63		
	138.4	7.20	52.11	25.39	144.81	6.41	6.90	-0.30
	126.5	7.10	73.08	25.08	122.35	-4.15	7.21	0.11
	128.6	7.00	68.12	25.31	127.66	-0.94	6.98	-0.02
	124.8	2.47	73.24	30.06	122.17	-2.63	2.20	-0.27
Benzenesulfonic acid	15.6		172.42		15.91	0.31		
	143.5	8.00	55.69	24.39	140.98	-2.52	7.91	-0.09
	126.3	7.50	67.59	24.79	128.23	1.93	7.51	0.01
	129.8	7.60	68.07	24.71	127.71	-2.09	7.59	-0.01
Ethanal	132.3		63.20		132.93	0.63		
	199.7	9.80	4.26	22.17	196.08	-3.62	10.14	0.34
Benzaldehyde	30.7	2.20	159.92	30.26	29.30	-1.40	2.00	-0.20
	136.4	7.85	61.07	24.44	135.21	-1.19	7.86	0.01
	129.5	7.49	65.29	24.79	130.69	1.19	7.51	0.02
	128.9	7.56	67.84	24.74	127.96	-0.94	7.56	0.00
	134.2	10.00	62.87	22.27	133.29	-0.91	10.04	0.04
<i>trans</i> -2-Butenal	192.0		9.14		190.86	-1.14		
	193.4	9.48	9.57	22.62	190.39	-3.01	9.69	0.21
	134.9	6.13	62.15	26.30	134.06	-0.84	5.99	-0.14
	153.7	6.87	43.71	25.62	153.81	0.11	6.67	-0.20
2-Formylthiophene	18.2	2.03	169.86	30.26	18.65	0.45	2.00	-0.03
	143.3	7.78	45.51	24.98	151.89	8.59	7.32	-0.46
	136.4	7.22	62.25	25.27	133.95	-2.45	7.02	-0.20
	128.1	7.78	70.27	24.85	125.36	-2.74	7.45	-0.33
	134.6	9.92	55.90	22.34	140.75	6.15	9.97	0.05
3-Formylthiophene	182.8		20.43		178.76	-4.04		
	137.1	7.94	53.59	24.82	143.23	6.13	7.48	-0.46
	142.6	7.41	53.78	24.64	143.03	0.43	7.66	0.25
	124.9	7.18	71.45	25.32	124.09	-0.81	6.97	-0.21
	127.3	9.83	64.24	22.29	131.82	4.52	10.02	0.19
Acetone	184.7		17.21		182.21	-2.49		
	206.0	2.09	-2.99	30.25	203.85	-2.15	2.01	-0.08
Butanone	28.1		161.44		27.67	-0.43		
	27.5	2.13	161.58	30.27	27.52	0.02	1.99	-0.14
	206.3	2.47	-6.14	29.85	207.23	0.93	2.41	-0.06
	35.2	1.05	154.54	31.21	35.06	-0.14	1.05	0.00
	7.0		180.15		7.62	0.62		
Cyclopentanone	213.6	2.06	-16.65	30.19	218.49	4.89	2.07	0.01
	36.7	2.02	152.11	30.34	37.67	0.97	1.92	-0.10
	22.0		164.23		24.68	2.68		
Cyclohexanone	208.5	2.22	-8.73	29.98	210.00	1.50	2.28	0.06
	40.4	1.80	148.75	30.32	41.27	0.87	1.94	0.14
	26.5	1.80	160.78	30.46	28.38	1.88	1.80	0.00
	23.8		163.20		25.79	1.99		
Acetophenone	136.3	7.91	61.02	24.26	135.27	-1.03	8.04	0.13
	128.1	7.40	66.93	24.87	128.94	0.84	7.43	0.03
	128.1	7.45	68.35	24.81	127.41	-0.69	7.49	0.04
	131.3	2.55	64.06	29.86	132.01	0.71	2.40	-0.15
	195.7		3.33		197.08	1.38		
Cyclohex-2-Enone	24.6		163.78		25.16	0.56		
	188.5	6.00	3.48	26.35	196.92	8.42	5.94	-0.06
	129.9	7.00	64.47	25.39	131.57	1.67	6.90	-0.10
	150.8		48.05	29.97	149.16	-1.64	2.29	
	25.7		162.46	30.27	26.58	0.88	1.99	
	22.7		164.40	29.94	24.50	1.80	2.32	
	38.1		152.12		37.66	-0.44		
Acetylacetone (keto)	30.2	2.17	161.67	30.10	27.42	-2.78	2.16	-0.01
	201.9	3.62	-0.51	28.87	201.20	-0.70	3.40	-0.22
	58.2		129.77		61.60	3.40		
Acetylacetone (enol)	24.3	2.00	165.06	30.37	23.79	-0.51	1.89	-0.11
	191.4	5.57	9.80	27.24	190.15	-1.25	5.04	-0.53
	100.3		95.10		98.75	-1.55		
Acetylacetone (<i>c</i> _{2v})	24.3	2.00	166.14	30.37	22.64	-1.66	1.89	-0.11
	191.4	5.57	4.68	27.11	195.63	4.23	5.17	-0.40
	100.3		94.66		99.22	-1.08		
<i>p</i> -Benzoquinone	187.0	6.78	8.95	25.82	191.06	4.06	6.47	-0.31

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
<i>o</i> -Benzoquinone	136.4		61.14		135.14	-1.26		
	180.2	6.34	18.15	26.30	181.20	1.00	5.99	-0.35
	140.0	7.09	65.25	25.65	130.74	-9.26	6.64	-0.45
Acetic acid	130.4		58.39		138.09	7.69		
	176.9	2.08	30.52	30.34	167.95	-8.95	1.92	-0.16
	20.8		171.15		17.27	-3.53		
Acetic acid, dimer	176.9	2.08	21.57	30.30	177.54	0.64	1.96	-0.12
	20.8		169.01		19.56	-1.24		
Acetic acid, mixed dimer	176.9	2.08	21.16	30.28	177.98	1.08	1.98	-0.10
	20.8		169.03		19.54	-1.26		
Propanoic acid	180.1	2.36	26.79	29.90	171.94	-8.16	2.36	0.00
	27.5	1.16	163.75	31.12	25.20	-2.30	1.14	-0.02
	8.7		179.28		8.56	-0.14		
Propanoic acid, dimer	180.1	2.36	18.40	29.82	180.93	0.83	2.44	0.08
	27.5	1.16	161.69	31.13	27.40	-0.10	1.13	-0.03
	8.7		179.59		8.22	-0.48		
Propanoic acid, mixed dimer	180.1	2.36	18.03	29.82	181.33	1.23	2.44	0.08
	27.5	1.16	161.67	31.13	27.42	-0.08	1.13	-0.03
	8.7		179.74		8.06	-0.64		
Butanoic acid	179.3	2.31	27.35	29.94	171.34	-7.96	2.32	0.01
	36.0	1.68	155.89	30.46	33.62	-2.38	1.80	0.12
	18.2	1.00	169.36	31.25	19.18	0.98	1.00	0.00
Butanoic acid, dimer	13.1		174.80		13.36	0.26		
	179.3	2.31	18.89	29.87	180.41	1.11	2.39	0.08
	36.0	1.68	153.90	30.47	35.75	-0.25	1.79	0.11
Butanoic acid, mixed dimer	18.2	1.00	169.73	31.25	18.79	0.59	1.00	0.00
	13.1		174.87		13.28	0.18		
	179.3	2.31	18.46	29.86	180.87	1.57	2.40	0.09
Methacrylic acid	36.0	1.68	153.90	30.46	35.75	-0.25	1.80	0.12
	18.2	1.00	169.72	31.26	18.80	0.60	0.99	-0.01
	13.1		174.96		13.18	0.08		
Methacrylic acid, mixed dimer	179.3	2.31	18.46	29.86	180.87	1.57	2.40	0.09
	36.0	1.68	153.90	30.46	35.75	-0.25	1.80	0.12
	18.2	1.00	169.72	31.26	18.80	0.60	0.99	-0.01
Benzoic acid	13.1		174.96		13.18	0.08		
	171.9	6.30	33.57	25.96	164.68	-7.22	6.33	0.03
	135.2	5.72	59.39	26.53	137.01	1.81	5.76	0.04
Methacrylic acid, mixed dimer	126.3	1.97	66.93	30.35	128.94	2.64	1.91	-0.06
	16.5		169.08		19.48	2.98		
	171.9	6.30	26.31	25.78	172.46	0.56	6.51	0.21
Benzoic acid	135.2	5.72	58.26	26.47	138.22	3.02	5.82	0.10
	126.3	1.97	65.09	30.35	130.91	4.61	1.91	-0.06
	16.5		169.70		18.82	2.32		
Benzoic acid, mixed dimer	129.4	8.10	67.65	24.09	128.16	-1.24	8.21	0.11
	130.2	7.40	64.78	24.87	131.24	1.04	7.43	0.03
	128.4	7.50	68.30	24.77	127.47	-0.93	7.53	0.03
Methyl acetate	133.7		63.66		132.44	-1.26		
	172.6		33.44		164.82	-7.78		
	129.4	8.10	66.61	24.04	129.28	-0.12	8.26	0.16
Ethyl acetate	130.2	7.40	64.69	24.85	131.34	1.14	7.45	0.05
	128.4	7.50	68.21	24.73	127.56	-0.84	7.57	0.07
	133.7		62.96		133.19	-0.51		
Methyl benzoate	172.6		26.21		172.57	-0.03		
	170.7	2.01	29.11	30.40	169.46	-1.24	1.86	-0.15
	19.6	3.67	169.76	28.67	18.76	-0.84	3.60	-0.07
Methyl acetate	50.7		139.22		51.48	0.78		
	170.0	2.03	29.19	30.38	169.37	-0.63	1.88	-0.15
	20.0	4.12	169.60	28.20	18.93	-1.07	4.07	-0.05
Ethyl acetate	59.8	1.25	129.64	30.99	61.74	1.94	1.27	0.02
	13.8		175.25		12.87	-0.93		
	130.3	7.97	66.06	24.16	129.87	-0.43	8.14	0.17
Methyl benzoate	129.5	7.37	65.43	24.92	130.54	1.04	7.38	0.01
	128.3	7.47	68.45	24.82	127.31	-0.99	7.48	0.01
	132.8	3.88	64.25	28.45	131.81	-0.99	3.82	-0.06
Vinyl acetate	166.8		32.33		166.01	-0.79		
	51.8		139.04		51.67	-0.13		
	167.6	2.12	31.62	30.31	166.77	-0.83	1.95	-0.17
N-Methylacetamide	20.2	7.25	169.72	24.60	18.80	-1.40	7.70	0.45
	141.8	4.55	54.68	27.86	142.06	0.26	4.42	-0.13
	96.8	4.85	99.97	27.58	93.53	-3.27	4.70	-0.15
	174.9	2.00	34.90	30.64	163.25	-11.65	1.62	-0.38

Table 1 (Contd.)

Compound	Chemical shift (δ relative to TMS in CDCl ₃)		B3LYP/6-31G* shielding (ppm)		δ (ppm), ¹³ C		δ (ppm), ¹ H	
	¹³ C	¹ H	¹³ C	¹ H	pred.	error	pred.	error
Acetanilide	22.7	2.71	167.15	29.47	21.55	-1.15	2.80	0.09
	26.9		165.06		23.79	-3.11		
	138.2	7.40	57.59	25.45	138.94	0.74	6.84	-0.56
	120.4	7.20	74.79	24.99	120.51	0.11	7.31	0.11
	128.7	7.00	67.29	25.24	128.55	-0.15	7.05	0.05
Acetic anhydride	124.1	2.10	72.83	30.27	122.61	-1.49	1.99	-0.11
	169.5		32.79		165.52	-3.98		
	24.1		169.47		19.07	-5.03		
	166.0	2.20	33.17	30.32	165.11	-0.89	1.94	-0.26
	20.7		168.66		19.93	-0.77		
Maleic anhydride	164.3	7.10	31.95	25.80	166.42	2.12	6.49	-0.61
	136.6		60.71		135.60	-1.00		
Succinimide	183.6	2.73	25.26	29.86	173.58	-10.02	2.40	-0.33
	30.3		160.25		28.95	-1.35		
Acetonitrile	117.7	1.98	86.87	30.58	107.57	-10.13	1.68	-0.30
	1.3		187.81		-0.58	-1.88		
Propanonitrile	120.8	2.35	82.73	30.14	112.01	-8.79	2.12	-0.23
	10.6	1.31	176.97	30.99	11.03	0.43	1.27	-0.04
	10.8		176.98		11.02	0.22		
Methyl isocyanide	158.2	2.85	27.20	29.25	171.50	13.30	3.02	0.17
	26.8		164.19		24.72	-2.08		
Acrylonitrile	117.2	5.52	84.88	26.71	109.70	-7.50	5.57	0.05
	107.8	5.80	86.25	26.31	108.23	0.43	5.98	0.18
	137.5	6.00	59.28	26.10	137.13	-0.37	6.19	0.19
Benzonitrile	112.5	7.60	80.88	24.73	113.99	1.49	7.57	-0.03
	132.0	7.40	63.13	24.88	133.01	1.01	7.42	0.02
	129.2	7.50	67.83	24.79	127.97	-1.23	7.51	0.01
	132.8		65.10		130.90	-1.90		
	118.7		82.26		112.51	-6.19		
Fluoromethane	75.0	4.27	121.86	27.94	70.08	-4.92	4.34	0.07
Carbon monoxide	182.2		10.22		189.70	7.50		
Carbon dioxide	124.2		74.69		120.62	-3.58		
Formaldehyde	197.0		8.20	22.36	191.86	-5.14	9.95	
Hydrogen cyanide	110.9		95.01	29.64	98.85	-12.05	2.63	
Allene	73.5		119.51	27.73	72.60	-0.90	4.55	
	212.6		-12.70		214.26	1.66		
Tetrafluoromethane	118.6		65.90		130.04	11.44		
Tetramethylsilane	0.0	0.00	189.75	32.18	-2.66	-2.66	0.07	0.07

Table 2 Regression equations and errors found for the large dataset (Table 1) with B3LYP (all shift values in ppm on the δ -scale relative to TMS)

Level	Nucleus	CDCl ₃					
		κ	δ^0	R^2	MUE	E_{\max}	E_{\min}
B3LYP/6-31G(d)//6-31G(d)	¹ H	-1.0065	32.46	0.9974	0.15	0.55	-0.61
	¹³ C	-1.0715	200.65	0.9984	2.16	13.56	-12.05
B3LYP/6-311+G(d,p)//6-31G(d)	¹ H	-0.9607	30.71	0.9981	0.12	0.65	-0.55
	¹³ C	-0.9683	175.13	0.9985	2.07	14.66	-11.88

not been established accurately. NMR experiments suggest a 1:1 tautomer mixture in aqueous solution and a slight preference (70%) for the N⁹-H tautomer in DMSO [26]. These findings are corroborated by the comparison of computed and experimental electronic spectra (see Chart 2) [27].

In line with the above, taking the mean of the shielding values computed for N⁷-H and N⁹-H purine (thus assuming a 1:1 mixture of both tautomers) clearly improves the predicted ¹³C chemical shifts for C⁴, C⁵, and C⁶. However, the errors remain large for C² and C⁸.

Halogen compounds

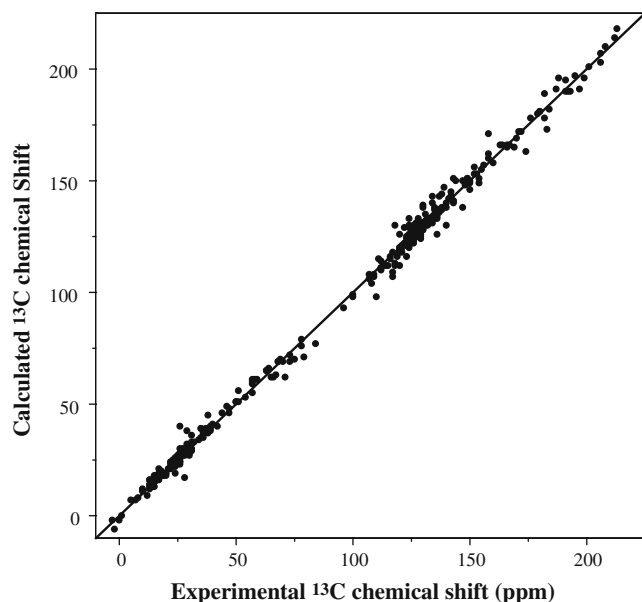
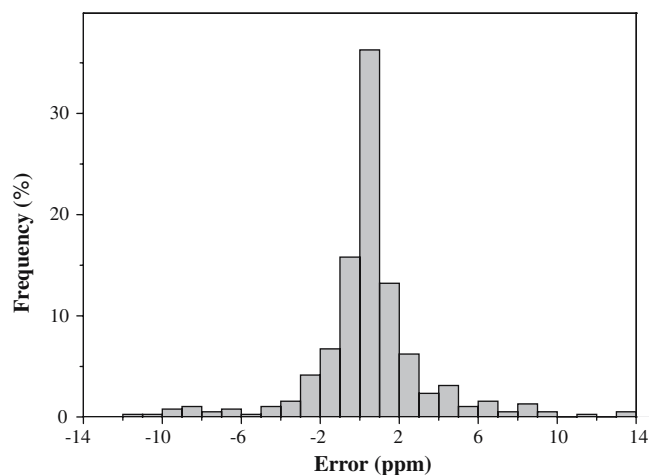
Quite generally, carbons bonded to chlorine or bromine show calculated ¹³C chemical shifts that are too high, more so for bromine than for chlorine. This trend is found for alkyl, vinyl and aryl systems.

Acetylenes

Calculated ¹³C chemical shifts for *sp*-carbon atoms are generally too negative. This effect is largest for CH

Table 3 Regression equations and errors found for ^{13}C chemical shifts using the small dataset at 18 different levels of calculation (all shift values in ppm on the δ -scale relative to TMS)

Method	Gas phase				CDCl_3					
	Difference formula		Linear regression formula		Linear regression formula		Linear regression formula			
	MSE	MUE	κ	δ^0	R^2	MUE	κ	δ^0	R^2	MUE
HF/6-31G(d)//B3LYP/6-31G(d)	-8.78	3.91	-1.0007	207.93	-0.9963	3.90	-1.0065	201.97	-0.9952	3.43
HF/D95*//B3LYP/6-31G(d)	-6.07	4.01	-0.9705	198.51	-0.9962	3.81	-0.9746	192.27	-0.9950	3.21
HF/6-311 + G(d,p)//B3LYP/6-31G(d)	0.36	6.78	-0.9207	187.67	-0.9940	5.26	-0.9280	181.91	-0.9927	4.99
HF/TZVP//B3LYP/6-31G(d)	1.80	7.63	-0.9067	185.73	-0.9943	5.13	-0.9144	180.05	-0.9928	4.94
B3LYP/6-31G(d)//B3LYP/6-31G(d)	-9.05	5.12	-1.0653	208.65	-0.9985	2.98	-1.0644	201.66	-0.9971	3.51
B3LYP/D95*//B3LYP/6-31G(d)	-4.47	5.03	-1.0258	198.41	-0.9962	4.62	-1.0233	191.22	-0.9949	4.55
B3LYP/6-311 + G(d,p)//B3LYP/6-31G(d)	4.13	3.00	-0.9547	181.88	-0.9997	1.21	-0.9554	175.10	-0.9983	2.91
B3LYP/TZVP//B3LYP/6-31G(d)	4.62	3.40	-0.9503	182.62	-0.9996	1.44	-0.9520	176.01	-0.9980	3.02
MP2/6-31G(d)//B3LYP/6-31G(d)	-12.81	7.83	-1.1292	236.07	-0.9985	2.36	-1.1315	229.62	-0.9966	3.89
MP2/D95*//B3LYP/6-31G(d)	-9.49	8.51	-1.1054	227.79	-0.9944	4.59	-1.1058	221.02	-0.9929	5.74
MP2/TZVP//B3LYP/6-31G(d)	-4.05	2.85	-1.0427	209.89	-0.9995	1.60	-1.0480	203.88	-0.9971	3.32
MP2/TZVP//B3LYP/6-31G(d)	-3.57	2.74	-1.0353	209.76	-0.9993	1.78	-1.0415	203.91	-0.9966	3.59
HF/6-311 + G(d,p)//B3LYP/6-311 + G(d,p)	-0.38	6.31	-0.9271	189.42	-0.9953	4.66	-0.9345	183.66	-0.9942	4.58
HF/TZVP//B3LYP/6-311 + G(d,p)	1.01	7.07	-0.9135	187.56	-0.9956	4.50	-0.9214	181.88	-0.9943	4.54
B3LYP/6-311 + G(d,p)//B3LYP/6-311 + G(d,p)	3.55	2.96	-0.9557	183.06	-0.9997	1.28	-0.9567	176.33	-0.9986	2.85
B3LYP/TZVP//B3LYP/6-311 + G(d,p)	4.00	3.35	-0.9515	183.87	-0.9996	1.47	-0.9536	177.30	-0.9982	3.05
MP2/6-311 + G(d,p)//B3LYP/6-311 + G(d,p)	-4.41	2.87	-1.0409	210.57	-0.9992	1.91	-1.0464	204.60	-0.9971	3.42
MP2/TZVP//B3LYP/6-311 + G(d,p)	-3.96	2.88	-1.0337	210.51	-0.9990	2.12	-1.0402	204.70	-0.9966	3.70

**Fig. 1** Experimental and calculated ^{13}C chemical shifts obtained using the fitted regression equation from Table 2 at B3LYP/6-31G(d)**Fig. 2** Histogram of the errors in calculated ^{13}C chemical shifts obtained using the fitted regression equation from Table 2 at B3LYP/6-31G(d)

groups, for which the calculated ^1H chemical shifts are also too negative.

Other levels of theory

Table 3 shows the performance for ^{13}C chemical shifts for the other levels of theory investigated and the constants determined for the regression equation using the small training dataset. The agreement for B3LYP/6-31G(d) between the large ($\delta^0 = -1.0715$, $\kappa = 200.65$) and the small ($\delta^0 = -1.0644$, $\kappa = 201.66$) datasets and the corresponding data ($\delta^0 = -0.9683$, $\kappa = 175.13$ vs. $\delta^0 = -0.9554$, $\kappa = 175.10$ for the large and small datasets, respectively) for B3LYP/6-311 + G(d,p) suggest that regression equa-

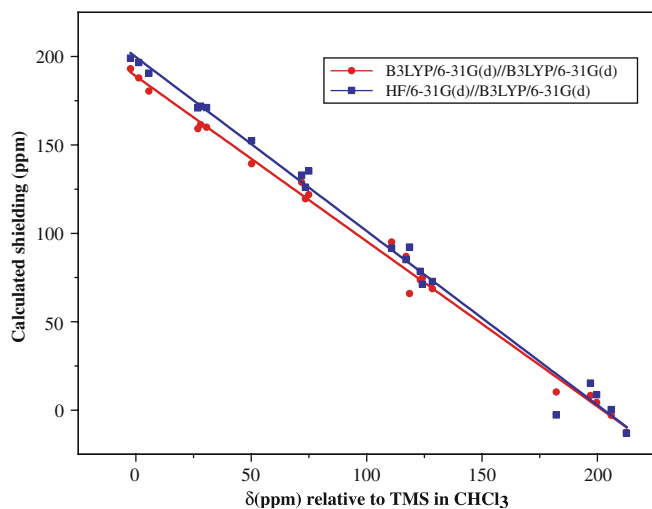


Fig. 3 Correlation between calculated shieldings and ^{13}C chemical shifts B3LYP/6-31G(d) and HF/6-31G(d) using the B3LYP optimized geometries

tions obtained for the small dataset will be of similar accuracy to those calculated with the large one.

A comparison of the correlations obtained with B3LYP/6-31G(d) and HF/6-31G(d) at the B3LYP/6-31G(d) optimized geometries is shown in Fig. 3. The slopes of the regression lines are significantly different but the correlations are similar. Thus, although the absolute performance of the different levels of calculation varies significantly, the regression-corrected results remain similar.

The general conclusions concerning the accuracy of the different levels of calculation is that both the regression constants obtained and the errors are unpredictable, especially for chemical shifts in chloroform. However, the results obtained with all the methods investigated are very satisfactory, so that the linear regression technique can be concluded to work well. The gas-phase data suggest that there is a slight advantage in calculating geometries with larger basis sets, but this trend is not reproduced in CDCl_3 solution.

Summary and conclusions

The simple linear regression technique results in significant improvements in the accuracy of ^{13}C and ^1H chemical shifts calculated both with DFT and ab initio methods. Chemical shifts calculated using the appropriate regression equation are accurate enough to allow assignment of spectra and even identification, for instance, of tautomers in many cases.

Similar performance is found at the relatively inexpensive B3LYP/6-31G(d)//B3LYP/6-31G(d) level of theory to more computationally intensive methods.

Some problem cases have been identified but in general performance is very good. Carboxylic acids can be

treated successfully by calculating their mixed dimer with formic acid.

References

- Ditchfield R, Miller DP, Pople JA (1970) *Chem Phys Lett* 6:573–575
- Gelius U, Roos B, Siegbahn P (1970) *Chem Phys Lett* 4:471–475
- Ditchfield R (1972) *Chem Phys Lett* 15:203–206
- See, for instance, Kaupp M, Malkin VG, Malkina OL, Salahub DR (1996) *Chem Eur J* 2:24–30
- Schindler M, Kutzelnigg W (1983) *J Am Chem Soc* 105:1360–1370
- Gauss J (1993) *J Chem Phys* 99:3629–3643
- van Wüllen CJ (1995) *Chem Phys* 102:2806–2811
- Schleyer PvR, Maerker C, Dransfield A, Jiao H, van Eikema Hommes NJR (1996) *J Am Chem Soc* 118:6317–6318
- Schleyer PvR, Manoharan M, Wang Z-X, Kiran B, Jiao H, Puchta R, van Eikema Hommes NJR (2001) *Org Lett* 3:2465–2468
- Schleyer PvR, Jiao H, Subramanian G, van Eikema Hommes NJR, Malkin VG, Malkina OL (1997) *J Am Chem Soc* 119:12669–12670
- Chesnut DB (1996) In: Lipkowitz LB, Boyd DB (eds) *Reviews in Computational chemistry*, vol 8, Chapter 5, VCH, New York.
- Forsyth DA, Sebag AB (1997) *J Am Chem Soc* 119:9483–9494
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Rega N, Salvador P, Dannenberg JJ, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2002) *Gaussian 98, Revision A11.3*. Gaussian Inc, Pittsburgh PA
- Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) *J Phys Chem* 98:11623–11627
- Becke AD (1993) *J Chem Phys* 97:5648–5652
- Lee C, Yang W, Parr RG (1988) *Phys Rev B* 37:785–789
- Hehre WJ, Ditchfield R, Pople JA (1972) *J Chem Phys* 56:2257–2261
- Francl MM, Pietro WJ, Hehre WJ, Binkley JS, DeFrees DJ, Pople JA, Gordon MS (1982) *J Chem Phys* 77:3654–3665
- Krishnan R, Binkley JS, Seeger R, Pople JA (1980) *J Chem Phys* 72:650–654
- McLean AD, Chandler CS (1980) *J Chem Phys* 72:5639–5648
- Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PvR (1983) *J Comp Chem* 4:294–301
- Dunning TH Jr, Hay PJ (1976) In: Schaefer III HF (ed) *Modern theoretical chemistry*, vol 3, Plenum, New York, pp 1–28
- Schäfer A, Horn H, Ahlrichs R (1992) *J Chem Phys* 97:2571–2577
- Hesse M, Meier H, Zeeh B (1979) *Spektroskopische Methoden in der organischen Chemie*. Thieme, Stuttgart, pp 254–262
- Jameson AK, Jameson C (1987) *J Chem Phys Lett* 134:461–466
- Schumacher M, Günther H (1982) *J Am Chem Soc* 104:4167–4173
- Borin AC, Serrano-Andrés L, Fülcher MP, Roos BO (1999) *J Phys Chem A* 103:1838–1845