Geometric Dependence of the B3LYP-Predicted Magnetic Shieldings and Chemical Shifts†

Ying Zhang,‡ Anan Wu,‡ Xin Xu,*,‡ and Yijing Yan£

State Key Laboratory of Physical Chemistry of Solid Surfaces, Center for Theoretical Chemistry, College for Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China, and Department of Chemistry, Hong Kong University of Science and Technology, Kowloon, Hong Kong

Received: May 25, 2007

We perform a systematic investigation of how the B3LYP/6-311+G(2d,p) calculated ^{13}C nuclear magnetic shielding constants depend on the 6-31G(d)-optimized geometries for a set of 18 molecules with various chemical environments. For absolute shieldings, the Hartree−Fock (HF)-optimized geometries lead to a mean absolute deviation (MAD) of 5.65 ppm, while the BLYP- and B3LYP-optimized geometries give MADs of 13.07 and 10.14 ppm, respectively. For chemical shifts, the HF, BLYP and B3LYP geometries lead to MADs of 2.36, 5.80, and 4.43 ppm, respectively. We find that the deshielding tendency of B3LYP can be effectively compensated by using the HF-optimized geometries. When we apply the B3LYP//HF protocol to versicolorin A and 5α-androstan-3,17-dione, MADs of 1.86 and 1.41 ppm, respectively, are obtained for chemical shifts, in satisfactory agreement with the experiment.

1. Introduction

In the past few decades, theoretical prediction of NMR (nuclear magnetic resonance) shielding constants has gained more and more attention.[1−4] In order to obtain calculated results with good accuracy for a correlated system, it is necessary to introduce the post Hartree−Fock (HF) methods such as the Møller−Plesset (MP_n) perturbation methods,[5−7] the coupled cluster methods,[7,8] the multiconfiguration self-consistent-field methods,[9] and so forth. However, calculations with such methods are too expensive to be applicable routinely to the chemically interesting large molecules.

Density function theory (DFT) offers an alternative way to take electron correlation into account, signifying itself with a reasonably high accuracy and a favorable cost-to-benefit ratio. As a salutary supplement to the wave-function-based methods, DFT, in particular B3LYP, has made great success in predicting the ground-state electronic structures, reaction energetics, molecular geometries, and so forth.[2] However, in NMR calculations with current DFT methods, some fundamental problems remain unsolved, that is, the current density dependency is not included,[10] and it is not free from self-interaction error.[11]

For NMR calculations, there exists the notorious gauge problem.[4,5,12] While several methods, for example, the gauge including atomic orbital (GIAO) method,[13] the individual gauge for localized orbital (IGLO) method,[14] the continuous set of gauge transformation (CSGT) method,[15] and so forth, have been proposed, it was concluded that no less than the triple-ζ basis set is necessary to reach the gauge invariance condition.[16,17] Unfortunately, it was found that B3LYP at the triple-ζ basis set always underestimates the NMR shieldings and overestimates the chemical shifts.[16,18,19] Hence, Pulay et al. concluded that the commonly used B3LYP method is significantly inferior to the HF method for general organic molecules.[18]

As NMR spectroscopy is an effective tool to determine the molecular geometry, it is not surprising that NMR constants are very sensitive to the changes in the molecular geometries. In this contribution, we have examined how the optimized geometries influence the calculated NMR constants by using HF, BLYP, and B3LYP with the 6-31G(d) basis set. Even through all methods lead to reasonably accurate geometries, there exists a systematic underestimation of bond lengths for the HF method and a clear overestimation for the BLYP and B3LYP methods. We find that the deshielding tendency of B3LYP can be effectively compensated by using the HF-optimized geometries. When we apply the B3LYP//HF protocol to versicolorin A and 5α-androstan-3,17-dione, MADs of 1.86 and 1.41 ppm, respectively, are obtained for chemical shifts, in satisfactory agreement with the experiment.

2. Computational Details

In this work, we have examined the performance of a commonly used hybrid functional, B3LYP.[20] It is based on Becke's three parameter scheme, consisting of the Slater exchange,[21] the exchange functional of Becke88,[22] and the HF exchange, as well as a mixture of the correlation functionals of Vosco−Wilk−Nusair[23] and Lee−Yang−Par.[24] In all NMR calculations, the GIAO method[13] was employed to circumvent the gauge problem. We adopted the 6-311+G(2d,p) basis set as recommended by Cheeseman et al.[16]

A set of 18 molecules was employed as a testing set in the present work.[5−8] For the whole set of molecules, reliable experimental NMR data (σ_0 or δ_0) are available in the gas phase at the zero-pressure limit, such that the experimental measurements were taken under conditions as close as possible to the isolated molecules.[16,25−27]

To investigate how the B3LYP-predicted NMR properties depend on the chosen geometries, we have adopted the optimized geometries obtained at the levels of HF, BLYP, and B3LYP by using the basis set 6-31G(d). This basis set has been used for geometry optimization in the well-known Gn method developed by Curtiss and co-workers for highly accurate

† Part of the "Sheng Hsien Lin Festschrift".

* To whom correspondence should be addressed. E-mail: xinxu@xmu.edu.cn.

‡ Xiamen University.

£ Hong Kong University of Science and Technology.

10.1021/jp0740503 CCC: $37.00
Published on Web 08/14/2007

TABLE 1: Theoretical versus Experimental Bond Distances (in Angstroms); Theoretical Optimization Values Were Obtained with the 6-31G(d) Basis Set, and Experimental Geometries Are the Equilibrium or Near-Equilibrium Geometries[31−33]

molecule	distance	HF	BLYP	B3LYP	expt[c]
CH_4	CH	1.0834	1.1003	1.0935	1.086
CH_3COCH_3[a]	CH_s	1.0810	1.0990	1.0914	1.085
	CH_a	1.0863	1.1051	1.0967	1.085
	CO	1.1921	1.2281	1.2118	1.222
	CC	1.5139	1.5333	1.5127	1.507
C_6H_6	CH	1.0755	1.0944	1.0870	1.0857
	CC	1.3862	1.4067	1.3966	1.3929
C_2H_2	CH	1.0568	1.0727	1.0672	1.06215
	CC	1.1855	1.2151	1.2051	1.20257
C_2H_4	CH	1.0760	1.0948	1.0874	1.081
	CC	1.3169	1.3408	1.3309	1.334
C_2H_6	CH	1.0856	1.1037	1.0963	1.089
	CC	1.5272	1.5408	1.5307	1.522
H_2CCCH_2	CH	1.0756	1.0960	1.0885	1.076
	CC	1.2959	1.3156	1.3070	1.3082
$CH_3CH_2CH_3$ [b]	C_mH	1.0872	1.1062	1.0986	1.0929
	C_tH_a	1.0865	1.1046	1.0973	1.0907
	C_tH_s	1.0857	1.1036	1.0962	1.0877
	CC	1.5281	1.5432	1.5323	1.5202
CF_4	CF	1.3018	1.3469	1.3290	1.3151
CH_3CN	CH	1.0821	1.1024	1.0946	1.087
	CC	1.4676	1.4688	1.4615	1.457
	CN	1.1347	1.1727	1.1603	1.156
CH_3F	CH	1.0818	1.1050	1.0964	1.095
	CF	1.3646	1.3977	1.3832	1.382
CH_3NH_2[a]	CH_s	1.0909	1.1134	1.1099	1.0929
	CH_a	1.0840	1.1031	1.0998	1.0929
	CN	1.4534	1.4794	1.4579	1.471
CH_3OH[a]	CH_s	1.0811	1.1010	1.0935	1.093
	CH_a	1.0875	1.1099	1.1013	1.093
	CO	1.3997	1.4347	1.4191	1.421
CHF_3	CH	1.0747	1.1013	1.0928	1.098
	CF	1.3166	1.3595	1.3420	1.332
CO	CO	1.1138	1.1505	1.1379	1.12832
CO_2	CO	1.1433	1.1828	1.1692	1.1615
H_2CO	CH	1.0913	1.1209	1.1100	1.1005
	CO	1.1846	1.2182	1.2069	1.2033
HCN	CH	1.0588	1.0775	1.0709	1.06501
	CN	1.1325	1.1691	1.1570	1.15324

[a] H_s refers to the syn mode with respect to the CO, NH_2, or OH group. [b] C_m is the middle carbon; C_t is the terminal carbon. The conformation of the methyl groups is staggered with respect to the CH_2 group. H_a refers to the anti mode; H_s refers to the syn mode. [c] Refs 31−33. We have updated the experimental geometries, which differ slightly from those used in ref 19.

prediction of molecular energies.[28,29] Indeed, this is also the basis set for geometry optimization used by Cheeseman et al.[16] for their NMR calculations. 6-31G(d) would be the basis set that is affordable, yet accurate, enough for geometry optimization of larger molecules.

All calculations are performed by using the GAUSSIAN 03 program suite.[30]

3. Results and Discussion

3.1. Geometric Parameters for the Testing Set. Tables 1 and 2 list the optimized bond lengths and bond angles, as well as the corresponding experimental data.[31−33] The mean absolute deviations (MADs) and the mean deviations (MDs, theory − expt) of the optimized structures from the experimental values are summarized in Table 3. Altogether, we have 39 comparisons for the bond distances and 19 comparisons for the bond angles. Our calculations give MDs of −0.009 and +0.015 Å, with MADs of 0.011 and 0.015 Å for HF and BLYP, respectively. Hence, the HF method consistently predicts bond lengths which are too short, and the BLYP method gives bond lengths which are consistently too long (see Table 1). Previous studies[19,34−36] showed that the MP2/6-31G(d) results, on average, were similar to those of the HF/6-31G(d) method. The MAD from MP2 for bond lengths of this set of molecules is 0.010 Å. For the HF method, the maximal error happens at the CO of CH_3COCH_3, for which the error is up to −0.030 Å. As the correlation effect is not included in the HF method, the predicted CC or CX bond lengths are less accurate than the C−H bond lengths. For the BLYP method, a CF in CF_4 contributes the maximal error (0.032 Å). While the CX bonds are described slightly better by BLYP, MADs for all other bonds are more than doubled as compared to those from HF. B3LYP also shows a consistent overestimation of bond lengths with 6-31G(d), similar to BLYP. B3LYP behaves less satisfactory in predicting the C−H bonds. Error in the C−H bonds of CH_3NH_2 is as high as 0.017 Å. Nevertheless, B3LYP is most satisfactory for the CC and CX bonds, leading to a MAD and MD of 0.007 and 0.005 Å, respectively, for the whole set.

TABLE 2: Theoretical versus Experimental Bond Angels (in Degrees); Theoretical Optimization Values Were Obtained with the 6-31G(d) Basis Set, and Experimental Geometries Are the Equilibrium or Near-Equilibrium Geometries[31−33]

molecule	bond angle	HF	BLYP	B3LYP	expt[c]
CH_3COCH_3[a]	CCC	116.6	116.5	116.4	117.2
	H_aCH_s	109.6	109.5	109.6	108.8
	OCC	121.7	121.8	121.8	121.4
C_2H_4	HCH	116.4	116.2	116.3	117.37
C_2H_6	HCH	107.6	107.5	107.5	107.69
H_2CCCH_2	HCH	117.8	116.6	117.0	118.2
$CH_3CH_2CH_3$[b]	CCC	112.8	113.1	112.9	112.35
	HC_mH	106.3	106.0	106.1	106.13
	$H_aC_tH_a$	107.6	107.4	107.5	107.04
	$H_aC_tH_s$	107.8	107.7	107.7	108.41
	$C_mC_tH_s$	111.4	111.5	111.6	111.6
	$C_mC_tH_a$	111.1	111.1	111.1	110.2
CH_3CN	HCC	109.9	110.5	110.3	110.1
CH_3F	HCH	109.8	109.2	109.3	110.45
CH_3NH_2[a]	H_aCN	109.2	109.0	109.1	108.80
	H_aCH_s	108.0	107.6	107.5	108.4
CH_3OH[a]	H_aCH_s	108.4	108.1	108.1	107.8
CHF_3	FCF	108.5	108.6	108.6	108.8
H_2CO	HCH	115.8	114.9	115.4	116.3

[a] H_s refers to the syn mode with respect to the CO, NH_2, or OH group. [b] C_m is the middle carbon; C_t is the terminal carbon. The conformation of the methyl groups is staggered with respect to the CH_2 group. H_a refers to the anti mode; H_s refers to the syn mode. [c] Refs 31−33. We have updated the experimental geometries, which differ slightly from those used in ref 19.

For the bond angles, the deviations from the experiment are all around 0.6°. For ∠HCH, the HF, BLYP, and B3LYP methods lead to MDs of −0.14, −0.53, and −0.42°, respectively. Hence, this bond angle can be calculated as too small for all three methods. For other angles like HCC, HCN, HCO, OCC, and FCF, the optimized angles can be too large, giving MDs of 0.08, 0.20, and 0.15°, for HF, BLYP, and B3LYP, respectively. The maximum errors are −0.99 (HF), −1.63 (BLYP), and −1.19° (B3LYP).

The above comparison demonstrates that the errors of the calculated geometries by the HF and BLYP methods are highly systematic. Omitting the correlation effects in the HF method usually gives too short of a bond distance, while the BLYP method consistently leads to bond distances which are too long. B3LYP shares the same tendency with BLYP, showing an overestimation of bond lengths with this basis set. For ordinary organic molecules, all three methods are able to predict geometries with reasonable accuracy. Nevertheless, as the NMR

TABLE 3: Mean Absolute Deviations (MADs) and Mean Deviations, (MDs, Theory − Expt) of the Optimized Structures from the Experimental Values;[31−33] Bond Distances Are in Angstroms, and Bond Angels Are in Degrees

structural type	number of comparisons	HF		BLYP		B3LYP	
		MAD	MD	MAD	MD	MAD	MD
CH	20	0.006	−0.006	0.014	0.014	0.007	0.006
CC[a]	8	0.011	−0.003	0.015	0.015	0.005	0.004
CX (X ≠ C)[b]	11	0.019	−0.019	0.018	0.018	0.007	0.003
bond distance	39	0.011	−0.009	0.015	0.015	0.007	0.005
max. dev.		−0.030 (CO, CH_3COCH_3)		0.032 (CF, CF_4)		0.017 (CH,CH_3NH_2)	
HCH	11	0.52	−0.14	0.79	−0.53	0.69	−0.42
other angles[c]	8	0.42	0.08	0.44	0.20	0.43	0.15
bond angles	19	0.48	−0.05	0.64	−0.23	0.58	−0.18
max. dev.		−0.99 (∠HCH, C_2H_4)		−1.63 (∠HCH, CH_2CCH_2)		−1.92 (∠HCH, CH_2CCH_2)	

[a] CC single, double, and triple bonds. [b] CN triple bonds, CF single bonds, and CO single and double bonds. [c] HCC, HCN, HCO, OCC, and FCF angles.

TABLE 4: Calculated[a] and Experimental[b] Isotropic Magnetic Shieldings (σ, in ppm)

	Expt σ_0[b]	B3LYP//HF err[c]	B3LYP//BLYP err[c]	B3LYP//B3LYP err[c]
$\underline{C}H_4$	195.1	−3.54	−7.53	−5.91
$\underline{C}_2H_2$	117.2	−4.88	−7.79	−6.79
$\underline{C}_2H_4$	64.5	−7.84	−12.97	−10.85
$\underline{C}_2H_6$	180.9	−4.85	−9.41	−7.45
$H_2C\underline{C}CH_2$	−28.9	−14.35	−19.38	−17.00
$H_2C\underline{C}\underline{C}CH_2$	115.2	−4.67	−10.31	−8.00
$CH\underline{C}H_2CH_3$	169.3	−6.05	−10.94	−8.62
$CH_3\underline{C}H_2CH_3$	170.9	−3.97	−8.20	−6.46
$\underline{C}_6H_6$	57.2	−5.79	−11.03	−8.51
$\underline{C}H_3F$	116.8	−3.50	−10.30	−7.43
$\underline{C}HF_3$	68.4	−4.74	−17.63	−12.33
$\underline{C}F_4$	64.5	−7.31	−21.48	−15.75
$\underline{C}O_2$	58.5	−2.00	−12.34	−8.61
$\underline{C}O$	1.0	−5.76	−22.89	−16.93
$H_2\underline{C}O$	−1.0	−10.19	−20.68	−17.10
$C\underline{H}_3OH$	136.6	−4.31	−10.68	−8.08
$\underline{C}H_3COCH_3$	−13.1	−8.65	−20.87	−16.44
$\underline{C}H_3COCH_3$	158.0	−4.13	−8.67	−6.54
$HC\underline{N}$	82.1	−4.22	−11.98	−9.33
$C\underline{H}_3CN$	187.7	−3.49	−8.48	−6.03
$\underline{C}H_3CN$	73.8	−5.11	−13.38	−10.65
$\underline{C}H_3\underline{N}H_2$	158.3	−4.88	−10.58	−8.19

[a] Method_1//method_2 stands for NMR calculation at the level of method_1 with the basis set of 6-311+G(2d, p) using the geometry obtained by method_2 with the basis set of 6-31G(d). [b] Experimental values for ^{13}C are taken from refs 16 and 25−27. [c] Errors are calculated with (theory − expt).

shielding constants are very sensitive to the geometric changes, we will see that small differences in the geometry can have large effects on the calculated magnetic constants.

3.2. Magnetic Shieldings and Chemical Shifts for the Testing Set. Table 4 reports the NMR shieldings at the level of B3LYP/6-311+G(2d,p) using the geometries optimized by HF, BLYP, or B3LYP with basis set 6-31G(d) (denoted as B3LYP//HF, B3LYP//BLYP, or B3LYP//B3LYP, respectively). MADs with respect to the experimental σ_0 are summarized in Table 5. At this point, we would like to point out that the calculations of NMR shieldings are performed at fixed equilibrium geometries such that the rovibrational effects are not included.[18,19,25,27]

From Table 4, we see that B3LYP has a heavy tendency of deshielding at the B3LYP geometries. The MAD from B3LYP//B3LYP for 22 comparisons of ^{13}C shieldings is 10.14 ppm. Interestingly, by adopting the HF geometry, the performance for B3LYP to predict ^{13}C shieldings impressively improves. As decreasing bond lengths always result in increasing shieldings, the tendency of systematic underestimation of the bond lengths by the HF method corrects, to some extent, the systematic deshielding tendency by the B3LYP method. For example, as compared to the B3LYP geometry of C_2H_4, the HF-calculated C−H and C−C bond lengths are too short by 0.006 and 0.017 Å, respectively; an improvement of 3.01 ppm is observed for this molecule with B3LYP//HF. This kind of improvement can be as high as 11.17 ppm in CO with a 0.015 Å shrinking of the bond length (see Table 4). Overall, B3LYP//HF leads to a MAD of 5.65 ppm for ^{13}C. Our testing set contains carbon nuclei in various environments, but the two sources of errors are so systematic and comparative that the B3LYP//HF protocol works well for all 22 comparisons examined here (see Figure 1a).

TABLE 5: Mean Absolute Deviations (MADs) of the Theoretical Magnetic Shieldings and Shifts from the Experimental Values (σ_0 and δ_0);[16,25−27] the Maximum Errors (Theory − Expt) for Each Method Are Also Included, and Numbers in Boldface Represent the Best Method in Each Entry

methods[a]	shielding		shift	
	MAD	max. err.	MAD	max. err.
B3LYP//HF	**5.65**	**−14.35** ($H_2C\underline{C}CH_2$)	**2.36**	**10.81** ($H_2C\underline{C}CH_2$)
B3LYP//BLYP	13.07	−22.89 ($\underline{C}O$)	5.80	15.36 ($\underline{C}O$)
B3LYP//B3LYP	10.14	−17.10 ($H_2\underline{C}O$)	4.43	11.19($H_2\underline{C}O$)

[a] Method_1//method_2 stands for the NMR calculation at the level of method_1 with the basis set of 6-311+G(2d, p) using the geometry obtained by method_2 with the basis set of 6-31G(d).

Since the BLYP geometries are systematically too long, it is not surprising to see that combining with such geometries deteriorates the performance for the B3LYP prediction of the NMR shieldings. The MAD for ^{13}C shieldings is as high as 13.07 ppm with B3LYP//BLYP.

Quantum chemistry calculates the magnetic shielding constants, which, however, cannot be directly measured in practice. Because the magnetic field strength and the nuclear magnetic moments are not known to parts per million accuracy,[4] what is measured experimentally is the chemical shift, defined as the difference between the magnetic shielding of a reference compound and that of the probe. Therefore, it is more interesting to check how the B3LYP-predicted chemical shifts depend on the chosen geometries.

As chemical shift is a relative quantity, errors in the prediction of chemical shifts could be better controlled for a method with systematic errors in the shielding prediction. Other effects, such as relativistic effects, intermolecular interactions, solvent effects, and rovibrational effects, are also expected to affect the chemical shift calculations less severely. Indeed, the widely observed success of BLYP for chemical shift predictions benefits from this systematic error cancellation.[6,16,19,37] Hence, we found that the MP2//MP2 protocol leads to a MAD of 10.02 ppm for ^{13}C magnetic shieldings but only a MAD of 2.80 ppm for chemical

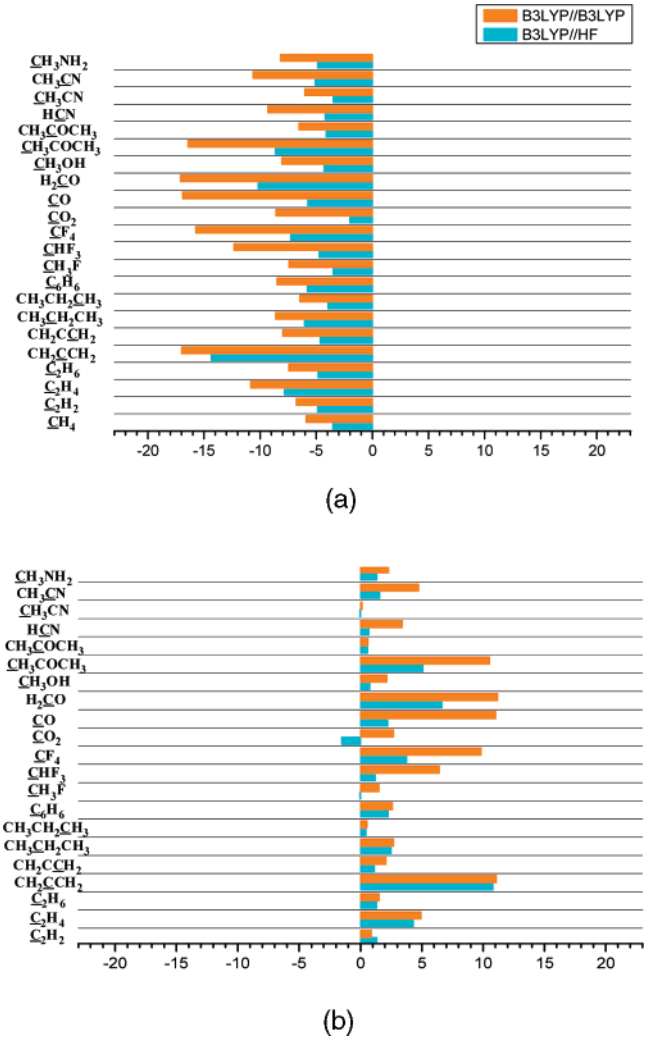

Figure 1. Errors (theory − expt) for the prediction of the ^{13}C NMR constants at the level of B3LYP/6-311+G(2d,p) by using the B3LYP-optimized or the Hartree−Fock-optimized geometries; (a) magnetic shieldings, (b) chemical shifts.

shifts.[19] It is therefore expected that B3LYP//B3LYP for chemical shifts works better than it does for magnetic shieldings[19] (see Table 6 for details). For example, the error associated with B3LYP//B3LYP for ^{13}C shielding in C_2H_2 is −6.79 ppm, which reduces to 0.88 ppm for the chemical shift, as B3LYP//B3LYP has an error of −5.91 ppm for shielding of the reference ^{13}C in CH_4. For B3LYP//B3LYP, the MAD for 19 comparisons of ^{13}C chemical shifts is 4.43 ppm, improved on average by 5.71 ppm with respect to the shielding prediction. We agree with Cheeseman to recommend B3LYP/6-311+G(2d,p) at the B3LYP/6-31G(d)-optimized geometry for the prediction of ^{13}C chemical shifts.[16]

From Table 6, we notice that the B3LYP//B3LYP results still show a systematic overestimation of the chemical shifts. The situation may be further improved if the B3LYP//HF protocol is applied. For example, B3LYP//B3LYP has errors of 1.52, 6.42, and 9.84 ppm for the ^{13}C shifts in CH_3F, CHF_3, and CF_4, respectively, whereas B3LYP//HF has only errors of −0.04, 1.20, and 3.78 ppm for the corresponding molecules. The improvement is most significant for CO. For this molecule, B3LYP//B3LYP has an error of 11.02 ppm, which reduces to 2.22 ppm with B3LYP//HF. Figure 1b demonstrates the consistent improvement for ^{13}C shifts with the B3LYP//HF protocol over B3LYP//B3LYP, leading to a MAD of only 2.36 ppm for 19 comparisons (cf. Table 5).

TABLE 6: Calculated[a] and Experimental[b] Chemical Shifts δ_0 in ppm

	expt δ_0[a]	B3LYP// HF err[c]	B3LYP// BLYP err[c]	B3LYP// B3LYP err[c]
C_2H_2	77.9	1.34	0.26	0.88
C_2H_4	130.6	4.30	5.43	4.94
C_2H_6	14.2	1.31	1.87	1.54
H_2CCCH_2	224.0	10.81	11.84	11.09
H_2CCCH_2	79.9	1.13	2.78	2.09
$CH_3CH_2CH_3$	25.8	2.51	3.40	2.71
$CH_3CH_2CH_3$	24.2	0.44	0.66	0.55
C_6H_6	137.9	2.26	3.49	2.60
CH_3F	78.3	−0.04	2.77	1.52
CHF_3	126.7	1.20	10.09	6.42
CF_4	130.6	3.78	13.95	9.84
CO_2	136.6	−1.54	4.80	2.70
CO	194.1	2.22	15.36	11.02
H_2CO	196.1	6.65	13.15	11.19
CH_3OH	58.5	0.77	3.15	2.17
CH_3COCH_3	208.2	5.11	13.33	10.53
CH_3COCH_3	37.1	0.59	1.13	0.63
HCN	113.0	0.68	4.45	3.42
CH_3CN	7.4	−0.05	0.94	0.12
CH_3CN	121.3	1.57	5.84	4.74
CH_3NH_2	36.8	1.34	3.05	2.28

[a] Method_1//method_2 stands for NMR calculation at the level of method_1 with the basis set of 6-311+G(2d, p) using the geometry obtained by method_2 with the basis set of 6-31G(d). [b] Experimental values for ^{13}C are taken from refs 16 and 25−27. [c] Errors are calculated with (theory − expt).

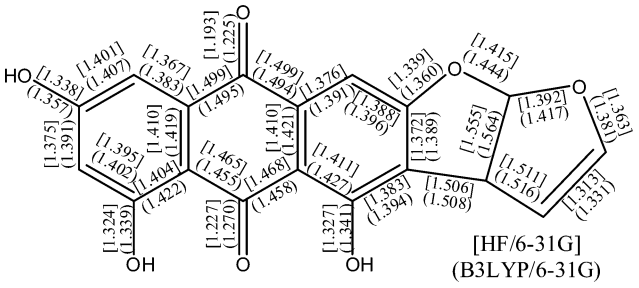

Figure 2. Geometries optimized at HF/6-31G(d) and B3LYP/6-31G(d) for versicolorin A.

3.3. Application of the B3LYP//HF Protocol. *3.3.1. Versicolorin A.* We have applied the B3LYP//HF protocol to the calculation of the ^{13}C chemical shifts of versicolorin A, a key intermediate for the biosynthesis of aflatoxin B_1.[38] The HF/6-31G(d)- and B3LYP/6-31G(d)-optimized geometries are represented in Figure 2. Generally, the HF geometry agrees with the B3LYP geometry but shows a consistent underestimation tendency. The MAD for the HF bond lengths with respect to the B3LYP geometry is around 0.015 Å, with the largest difference at two CO bonds. For the benzoquinone ring, the lack of electron correlation may lead to less electron delocalization, which may make the CC bonds too long for the HF geometry.

Table 7 lists the experimentally measure chemical shifts. As the experimental chemical shifts are usually reported with respect to liquid TMS, we need a scheme to convert the calculated shieldings in the gas phase to the chemical shifts in solutions. Taking the experimental shieldings[27] of 184.1 and 195.1 ppm for ^{13}C of liquid TMS and gaseous CH_4, respectively, we may have

$$\delta_A^{Calc}(\text{w.r.t. TMS(l)}) = (\sigma_{CH_4(g)}^{Calc} - \sigma_A^{Calc}) + (\sigma_{TMS(l)}^{Expt} - \sigma_{CH_4(g)}^{Expt}) \quad (1)$$

TABLE 7: Versicolorin A Calculated[a] and Experimental[b] Chemical Shifts δ_0 in ppm

	expt δ_0[b]	B3LYP//HF err[c]	B3LYP//B3LYP err[c]	ChemNMR[d] err[c]
1	108.8	−1.08	0.87	−0.5
2	165	−0.24	1.56	−1.7
3	107.8	−3.15	−1.91	−0.8
4	164	2.23	5.27	0.5
5	108.4	1.67	2.01	0.5
6	134.6	3.86	3.02	1.8
7	188.7	0.28	4.00	−0.7
8	111.3	0.87	1.02	−3.3
9	135.1	3.35	2.38	−2.2
10	180.4	−1.04	4.08	1.7
11	158.2	2.16	5.10	3.1
12	120.4	0.61	2.01	−5.3
13	163.4	3.65	5.75	2.3
14	101.6	1.11	3.29	7.2
15	47.2	4.12	3.33	−2.6
16	112.8	−2.05	2.83	−16.2
17	101.2	0.97	2.09	−1.0
18	145.3	1.04	3.48	3.5
MAD		1.86	3.00	3.05

[a] Method_1//method_2 stands for NMR calculation at the level of method_1 with the basis set of 6-311+G(2d, p) using the geometry obtained by method_2 with the basis set of 6-31G(d). [b] Experimental values are with respect liquid TMS.[38] [c] The theoretical values are obtained by using eq 1. Errors are calculated with (theory − expt). [d] The theoretical values are obtained with ChemNMR.[39] Errors are calculated with (theory − expt).

That is, we may first calculate the chemical shift with respect to the gaseous CH_4, which may then be converted to the chemical shift with respect to the liquid TMS by using the difference between experimental shieldings of gaseous CH_4 and liquid TMS.

Errors (theory − expt) from B3LYP//HF and B3LYP//B3LYP are summarized in Table 7. From Table 7, it is clear that B3LYP//HF agrees with the experiment better than B3LYP//B3LYP. For 18 comparisons of ^{13}C chemical shifts, B3LYP//HF leads to a MAD of 1.86 ppm, while that for B3LYP//B3LYP amounts to 3.00 ppm. For B3LYP//HF, the maximum error (4.11 ppm) occurs at C_{15}, with the other four large errors (3.15−3.86 ppm) being at C_3, C_6, C_9, and C_{13}. For B3LYP//B3LYP, there are five values with errors larger than 4.0 ppm. The maximum error is 5.75, occurring at C_{13}. We noticed that most of the carbon atoms close to oxygens are erroneous and are particularly improved by the B3LYP//HF protocol.

For comparison, we also present the results with the fragment-based model using ChemNMR.[39] ChemNMR predicts the shifts in a new molecule by comparing its fragments with the prestored experimental ^{13}C shifts of similar fragments in a library. Generally, this model works well. There, however, exists an exception for versicolorin A. The error at C_{16}, which connects two oxygen atoms, is as high as 16.2 ppm. The final MAD for versicolorin A is 3.05 ppm.

3.3.2. 5α-Androstan-3,17-dione. Figure 3 depicts two conformers of 5α-androstan-3,17-dione, which is a kind of steroid characterized by a carbon skeleton with four fused rings. The structural differences between conformers A and B originate from the difference in the stereo structures of C_{10} and C_{14}. Here, A adopts (10*S*) and (14*S*), whereas B takes (10*R*) and (14*R*). On the basis of HF/6-31G, A is 1.99 kcal/mol more stable than B. B3LYP/6-311+G(2d,p)// HF/6-31G(d) also favors conformer A 1.93 kcal/mol more than conformer B.

Reich et al. have measured the ^{13}C NMR chemical shifts with respect to carbon disulfide.[40] In a way similar to eq 1, we may take the experimental shieldings[27] of −8.3 and 195.1 ppm for the ^{13}C of liquid CS_2 and gaseous CH_4, respectively, such that

$$\delta_A^{Calc}(\text{w.r.t. } CS_2(l)) = (\sigma_{CH_4(g)}^{Calc} - \sigma_A^{Calc}) + (\sigma_{CS_2(l)}^{Expt} - \sigma_{CH_4(g)}^{Expt}) \quad (2)$$

In Table 8, we lists the experimental data, as well as errors from ChemNMR[39] and B3LYP//HF. As is obvious, ChemNMR[39] works very well in this case, leading to a MAD of only 0.93 ppm. The maximum error is 3.2 ppm, occurring at C_{18}. Hence, the ^{13}C chemical environment for this molecule is typical

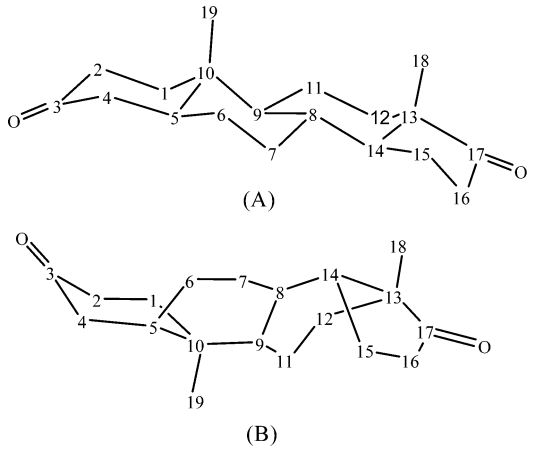

Figure 3. Conformation change of 5α-androstan-3,17-dione; (A) (5*S*,8*R*,9*S*,10*S*,13*S*,14*S*) and (B) (5*S*,8*R*,9*S*,10*R*,13*S*,14*R*).

TABLE 8: Conformation Change of 5α-Androstan-3,17-dione with Calculated[a] and Experimental[b] Chemical Shifts δ_0 in ppm; Conformer A Corresponds to (5*S*,8*R*,9*S*,10*S*,13*S*,14*S*) and Conformer B Corresponds to (5*S*,8*R*,9*S*,10*R*,13*S*,14*R*)

	expt	ChemNMR[d]	B3LYP//HF err[c]	
	δ_0[b]	err[c]	conformer A[e]	conformer B[e]
1	154.0	−0.3	0.30	−13.00
2	154.7	0.1	0.71	−0.63
3	−16.5	1.9	−1.17	−0.32
4	148.1	0.3	0.67	−0.70
5	145.8	−0.1	2.09	1.83
6	163.7	−0.5	1.02	1.66
7	160.6	−1.3	−0.75	−0.14
8	157.4	1.5	1.27	−2.13
9	138.2	−1.1	1.33	−10.29
10	156.6	−0.5	3.86	4.77
11	171.7	−0.3	0.05	−1.44
12	161.8	0.7	1.33	−0.53
13	145.1	0.8	2.14	0.82
14	141.1	−0.2	2.08	−1.70
15	170.7	0.0	−0.72	−1.63
16	157.1	0.4	−0.76	1.10
17	−25.9	2.5	−0.04	4.44
18	181.5	3.2	−0.76	14.24
19	179.0	−1.9	−5.75	7.43
MAD		0.93	1.41	3.62

[a] Method_1//method_2 stands for NMR calculation at the level of method_1 with the basis set of 6-311+G(2d, p) using the geometry obtained by method_2 with the basis set of 6-31G(d). [b] Experimental values are with respect liquid CS_2.[40] [c] The theoretical values are obtained by using eq 2. Errors are calculated with (theory − expt). [d] The theoretical values are obtained with ChemNMR[39] with respect to liquid TMS, which are then converted to data with respect to liquid CS_2 by using the experimental values of shieldings.[27] Errors are calculated with (theory − expt). [e] Based on HF/6-31G and B3LYP/6-311+G(2d,p)//HF/6-31G, conformer A is 1.99 and 1.93 kcal/mol more stable than conformer B, respectively.

to be well predicted in the library. Nevertheless, ChemNMR[39] is unable to predict the NMR difference for conformers. B3LYP//HF reveals the dramatic difference for NMR chemical shifts of conformers A and B. Changing structure from A to B makes C_1 and C_9 be deshielded by about 13 ppm and C_{18} and C_{19} to be overshielded by about 13 ppm. As compared with the observed ^{13}C chemical shifts, B3LYP//HF leads to a MAD of 1.41 ppm for conformer A and 3.62 ppm for conformer B. Hence, B3LYP//HF concludes that conformer A is the experimentally observed structure, which is in line with the conclusion based on the energy criterion that A is more stable than B.

4. Concluding Remarks

A typical study using ab initio methods starts with a tentative molecular structure for a system that may or may not be known. Through a geometry optimization, a stable system is obtained. Then, all of its properties can be abstracted from the wave function. While experimental determination of the geometry can be a nontrivial task for a polyatomic molecule, geometry optimization from first principles is now routinely applied in computational chemistry. Even though theory may suffer from the uncertainty introduced by the deficiencies in the basis set and the incomplete treatment of the correlation effects, the trend of a given method can be well understood and predicted. Using a set of carbons in various chemical environments, we show that optimizations at the levels of HF, BLYP, and B3LYP with 6-31G(d) can all lead to reasonably good results. While B3LYP gives the best results for CC and CX bonds, it does not always improve over HF for CH bonds. HF gives CH bonds in good quality, whereas errors from BLYP are always the largest among these three. Consistently, HF has a tendency to underestimate the bond length (MD = −0.009 Å), while BLYP and B3LYP have a tendency to overestimate the bond length (MD = 0.015 and 0.005 Å, respectively). Hence, the exact geometry could be in between those from HF and B3LYP.

NMR spectroscopy has continued to be the most valuable tool for structure elucidation. By comparison of the experimental NMR data with the theoretical ones, structures of unknown compounds may be better identified. For example, Cremer et al. determined the C−C bond distance by minimizing the difference between the calculated and the experimental ^{13}C chemical shifts.[41] Barbosa et al. compared the calculated chemical shifts with the NMR measurements to identify the best conformer, as opposed to using the energy criterion.[42] To fulfill such purpose, we believe that we shall understand the geometric dependence well for the calculated NMR data by a given method.

Cheeseman et al.[16] found that, while HF/6-31G(d)//B3LYP/6-31G(d) is better than B3LYP/6-31G(d)//B3LYP/6-31G(d) for ^{13}C chemical shifts (MAD: 9.3 versus 11.4 ppm), their recommended method is B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d), with MAD being 3.6 ppm for the ^{13}C chemical shifts. We agree with Cheeseman[16] but point out here that there could be another option for the prediction of ^{13}C NMR constants. Using our testing test, we find that B3LYP/6-311+G(2d,p)//HF/6-31G(d) can work better than B3LYP/6-311+G(2d,p)//B3LYP/6-31G(d), leading to MADs of 2.36 and 4.34 ppm, respectively.

As most of the ^{13}C chemical shifts are reported with respect to liquid TMS or CS_2, we may convert the calculated chemical shifts with respect to gaseous CH_4 by using the experimental difference of liquid TMS/CS_2 and gaseous CH_4. When we apply the B3LYP//HF protocol to versicolorin A and 5α-androstan-3,17-dione in this way, MADs of 1.86 and 1.41 ppm, respectively, are obtained, in satisfactory agreement with the experiment.

Acknowledgment. This work is supported by NSFC (20525311, 20533030, 20423002), the Ministry of Science and Technology (2004CB719902), and the Department of Science and Technology of Fujian Province (2006F3120).

References and Notes

(1) Nakatsuji, H. *J. Chem. Phys.* **1974**, *61*, 3728.
(2) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley: New York, 2000.
(3) Helgaker, T.; Jaszunski, M.; Ruud, K. *Chem. Rev.* **1999**, *99*, 293.
(4) Kaupp, M., Bühl, M., Malkin, V. C., Eds. *Calculation of NMR and EPR Parameters. Theory and Applications*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany 2004.
(5) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629.
(6) Kollwitz, M.; Gauss, J. *Chem. Phys. Lett.* **1996**, *260*, 639.
(7) Cybulski, S. M.; Bishop, D. M. *J. Chem. Phys.* **1997**, *106*, 4082.
(8) Christiansen, O.; Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1997**, *266*, 53.
(9) van Wullen, C.; Kutzelnigg, W. *J. Chem. Phys.* **1996**, *104*, 2330.
(10) Lee, A. M.; Handy, N. C.; Colwell, S. M. *J. Chem. Phys.* **1995**, *103*, 10095.
(11) Patchkovskii, S.; Autschbach, J.; Ziegler, T. *J. Chem. Phys.* **2001**, *115*, 26.
(12) Hameka, H. F. *Advanced Quantum Chemistry*; Addison-Wesley: New York, 1963.
(13) Ditchfield, R. *J. Chem. Phys.* **1972**, *56*, 5688.
(14) Schindler, M.; Kutzelnigg, W. *Mol. Phys.* **1983**, *48*, 781.
(15) Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1993**, *210*, 223.
(16) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497.
(17) Olsson, L.; Cremer, D. *J. Phys. Chem.* **1996**, *100*, 16881.
(18) Magyarfalvi, G.; Pulay, P. *J. Chem. Phys.* **2003**, *119*, 1350.
(19) Zhang, Y.; Wu, A.; Xu, X.; Yan, Y. J. *Chem. Phys. Lett.* **2006**, *421*, 383.
(20) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(21) Slater, J. C. *Quantum Theory of Molecules and Solids, Vol. 4: The Self-Consistent Field for Molecules and Solids*; McGraw Hill: New York, 1974.
(22) Becke, A. D. *Phys. Rev. B* **1988**, *38*, 3098.
(23) Vosco, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(24) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(25) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1996**, *104*, 2574.
(26) Malkin, V. G.; Malkina, O. L.; Casida, M. E.; Salahub, D. R. *J. Am. Chem. Soc.* **1994**, *116*, 5898.
(27) Jameson, A. K.; Jameson, C. J. *Chem. Phys. Lett.* **1987**, *134*, 461.
(28) Curtiss, L. A.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 1293.
(29) Bauschlicher, C. W., Jr.; Partridge, H. *J. Chem. Phys.* **1995**, *103*, 1788.
(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.
(31) Callonion, J. H.; Hirota, E.; Kuchitsu, K.; Lafferty, W. J.; Maki, A. G.; Pote, C. S.; in *Landolt-Bornstein: Numerical Data and Functional Relationships in Science and Technology, Vol. 7, New Series, Structure Data on Free Polyatomic Molecules*; Hellwege, K. H., Ed.; Springer-Verlag: West Berlin, Germany, 1976.
(32) Huber, K. P.; Herzberg, G. H.; *Molecular Spectra and Molecular Structure, IV. Constants of Diatomic Molecules*; Van Nostrand-Reinhold: New York, 1979.
(33) *Landolt-Bornstein: Numerical Data and Functional Relationships in Science and Technology, Vol. 23, New Series, Structure Data of Free*

Polyatomic Molecules. Kuchitsu, K., Ed.; Springer-Verlag: West Berlin, Germany, 1995.

(34) DeFrees, D. J.; Levi, B. A.; Pollack, S. K.; Hehre, W. J.; Binkley, J. S.; Pople, J. A. *J. Am. Chem. Soc.* **1979**, *101*, 4085.

(35) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.

(36) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430.

(37) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629.

(38) Gorts-Allman, C. P.; Steyn, P. S.; Wessels, P. L.; Scott, D. B. *J. Chem. Soc., Perkin Trans.* **1978**, *1*, 961.

(39) *ChemNMR*, in ChemDraw 7.0; Cambridgesoft: Cambridge, MA, 2007.

(40) Reich, H. J.; Jautelat, M.; Messe, M. T.; Weigert, F. J.; Roberts, J. D. *J. Am. Chem. Soc.* **1969**, *91*, 9483.

(41) Szabo, K. J.; Kraka, E.; Cremer, D. *J. Org. Chem.* **1996**, *61*, 2783.

(42) Alcântara, A. F. C.; Piló-Veloso, D.; De Almeida, W. B.; Maltha, C. R. A.; Barbosa, L. C. A. *J. Mol. Struct.* **2006**, *791*, 180.