JCTC Journal of Chemical Theory and Computation

Hybrid Density Functional Methods Empirically Optimized for the Computation of ^{13}C and ^{1}H Chemical Shifts in Chloroform Solution

Keith W. Wiitala, Thomas R. Hoye, and Christopher J. Cramer*

Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431

Received March 20, 2006

Abstract: Two hybrid generalized-gradient approximation density functionals, WC04 and WP04, are optimized for the prediction of ^{13}C and ^{1}H chemical shifts, respectively, using a training set of 43 molecules in chloroform solution. Tests on molecules not included in the training set, namely six stereoisomeric methylcyclohexanols and a β-lactam antibiotic, indicate the models to be robust and moreover to provide results more accurate than those from equivalent B3LYP, PBE1, or *m*PW1PW91 calculations, particularly for the prediction of downfield resonances in nuclear magnetic resonance spectra. However, linear regression of the B3LYP, PBE1, and *m*PW1PW91 predicted values on the experimental data improves the accuracy of those models so that they are comparable to WC04 and WP04.

Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for the determination of molecular structure.[1] Its utility in the pharmaceutical discovery process has been emphasized,[2−4] and recent developments in the field include taking advantage of increasingly high-field magnets, novel pulse sequences, and the design of multidimensional spectroscopic experiments.

A fundamental observable in NMR spectroscopy is the nuclear chemical shift, δ. The chemical shift is sensitive to molecular environment, thereby providing insight into local functionality and stereochemistry. Many predictive models have been advanced in order to assist in the interpretation of experimental chemical shifts. The oldest such models are purely empirical and typically adopt a fragment substitution approach,[5] although more modern variations are increasingly sophisticated in their ability to account for local fields and anisotropies.[6] However, fragment models are limited in their ability to account for chemical shift differences associated with nonlinear interactions between multiple fragments or with stereoisomerism. Thus, there has been substantial interest in the use of quantum chemical models to predict chemical shift values from first principles for use in spectral interpretation.[7−12]

The theory associated with the computation of chemical shifts is well developed,[13] and it has been demonstrated that highly correlated electronic structure methods using very large basis sets are capable of achieving high accuracy. However, such models are not computationally practical for large molecules or for databases containing a very large number of molecules. For molecules of moderate to large size, density functional theory (DFT) arguably provides the best combination of accuracy and efficiency among quantum chemical models, and the utility of a number of functionals for the prediction of chemical shift values has been evaluated.[13−23] In general, modern functionals with large basis sets provide results of reasonable quantitative accuracy, but improvements are possible. For example, correcting the energy separation between occupied and virtual Kohn−Sham eigenvalues has been observed to provide improved chemical shift predictions.[24−26] Model exchange-correlation potentials uncoupled from the self-consistent field procedure have also demonstrated good accuracy.[27,28] Allen et al.[23] have recently compared results from some of these approaches with those from the KT2 functional,[29] which was developed specifically to be a generalized gradient approximation (GGA) functional useful for the direct prediction of chemical shifts.

* Corresponding author phone: (612)624-0859; fax: (612)626-2006; e-mail: cramer@chem.umn.edu.

10.1021/ct6001016 CCC: $33.50 © 2006 American Chemical Society
Published on Web 06/01/2006

From a more statistical standpoint, a number of authors have explored linear regression approaches to correct systematic errors associated with smaller basis sets and/or inaccurate functionals.[30−34] In this work we consider an alternative somewhat along those lines. Many density functionals include one or more parameters whose values are chosen either on the basis of theoretical arguments associated with ideal model systems or by optimization against experimental data. In the latter instance, the data of choice have tended to be dominated by thermodynamic quantities (e.g., atomization energies), structural features (e.g., bond lengths), and in some instances molecular vibrational frequencies. In this paper, we present two new functionals optimized specifically to predict ^{13}C and ^{1}H chemical shifts in chloroform solution, and we demonstrate that they have substantially improved accuracy compared to popular, current "off-the-shelf" functionals. Such optimized functionals should facilitate interpretation of NMR spectra of moderate to highly complex structures containing these two nuclei. Our approach is similar in spirit to prior work by Patchkovskii and Thiel,[35] who reparametrized the semiempirical modified neglect of differential overlap (MNDO) model to create a model named MB3, which is designed to give improved accuracy for ^{1}H, ^{13}C, ^{15}N, and ^{17}O chemical shifts.

Computational Methods

A total of 160 conformers spanning the 43 molecules in the NMR training set (described below) were fully optimized at the B3LYP level[36−39] using the 6-31G(d) basis set.[40] In addition to gas-phase geometries, geometries taking account of chloroform as solvent were optimized using the integral equation formalism[41] of the polarized continuum model[42] (IEFPCM). The molecular cavity for these calculations was constructed as a sum of atom-centered spheres using the radii of Bondi.[43]

For each individual geometry, atomic chemical shielding tensors σ were computed[17] using the gauge independent atomic orbital (GIAO) formalism[44−46] and including the effects of chloroform solvation via the PCM model[47,48] (this inclusion is at the level of the electronic structure irrespective of whether solvated geometries are employed). Isotropic atomic chemical shifts δ in units of ppm were computed as differences between atomic isotropic shieldings in solutes and corresponding reference atoms in tetramethylsilane (TMS). When more than a single conformer merits consideration, δ values are reported as an average over a Boltzmann-weighted population of conformers according to[9,49]

$$\delta = \sum_i \left(\frac{\delta_i e^{-G_i^{\circ}/RT}}{\sum_j e^{-G_j^{\circ}/RT}} \right) \quad (1)$$

where i and j run over conformers, G is the free energy of the conformer in solution, R is the universal gas constant, and T is 298 K. The free energy in solution is taken as the sum of the electronic energy and solvation free energy computed at the B3LYP level using the 6-311+G(2d,p) basis set[40] and IEFPCM chloroform solvation free energies. This same level of theory was used for the computation of the chemical shifts.

We define the energy E for a general hybrid exchange-correlation (xc) functional as

$$E_{xc}^{BxLYP} = P_2 E_x^{HF} + P_3 \Delta E_x^{B} + P_4 E_x^{LSDA} + P_5 \Delta E_c^{LYP} + P_6 E_c^{LSDA} \quad (2)$$

where P_2-P_6 are weighting parameters ranging from 0 to 1, and the terms on the right-hand side correspond, respectively, to the Hartree−Fock (HF) exchange energy,[49] the Becke[36] (B) gradient correction to the local spin-density approximation (LSDA) exchange energy, and the Lee, Yang, and Parr[37] (LYP) correction to the local spin-density approximation (LSDA) correlation energy of Vosko, Wilk, and Nusair[50] (VWN). The popular B3LYP functional[36−39] is defined by $P_2 = 0.20$, $P_3 = 0.72$, $P_4 = 0.80$, $P_5 = 0.81$, and $P_6 = 1.00$.

For the weighted carbon (WC04) and proton (WP04) functionals, the 5 weighting parameters were optimized using a central composite design.[51] Separate optimizations were done for ^{13}C chemical shifts and for ^{1}H chemical shifts of protons bonded to carbon (protons bonded to heteroatoms were not considered because of the extreme sensitivity of their chemical shifts to the purity of experimental NMR solvents, solute concentration, and variations in pH). The design response was the total absolute error over all chemical shifts in each data set defined as

$$|\Delta\delta| = \sum_i \sum_j |\delta_{\mathrm{exp},ij} - \delta_{\mathrm{calc},ij}| \quad (3)$$

where i runs over the 43 molecules in the training set and j runs over the number of carbon or hydrogen atoms in molecule i. Separate ^{13}C- and ^{1}H NMR response surfaces were generated with the parameters permitted to range from 0.0001 to 0.9999. The numbers of unique ^{13}C- and ^{1}H chemical shifts are 141 and 255, respectively.

The 43 molecules in the NMR training set (including Chemical Abstracts Service numbers) were as follows: acetaldehyde (75-07-0), acetamide (60-35-5), acetic acid (64-19-7), acetic anhydride (108-24-7), acetone (67-64-1), acetone oxime (127-06-0), acetonitrile (75-05-8), acetyl chloride (75-36-5), acrolein (107-02-8), 1-bromopropane (106-94-5), 3-buten-2-one (78-94-4), 1-chloropropane (540-54-5), cyclohexane (110-82-7), diacetamide (625-77-4), diethyl ether (60-29-7), dimethyl carbonate (616-38-6), 2,2-dimethyl-1,3-dioxolane (2916-31-6), dimethyl sulfate (77-78-1), dimethyl sulfide (75-18-3), dimethyl sulfite (616-42-2), dimethyl sulfone (67-71-0), dimethyl sulfoxide (67-68-5), 1,3-dimethylurea (96-31-1), ethanethiol (75-08-1), ethylbenzene (100-41-4), ethyl carbamate (51-79-6), ethyl isocyanate (109-90-0), furan (110-00-9), methanol (67-56-1), methyl acetate (79-20-9), methyl acrylate (96-33-3), methyl isothiocyanate (556-61-6), methyl thiocyanate (556-64-9), nitrobenzene (98-95-3), nitromethane (75-52-5), *N*-nitrosodimethylamine (62-75-9), *n*-propylamine (107-10-8), *n*-pentane (109-66-0), 1-pentene (109-67-1), 1-pentyne (627-19-0), phenol (108-95-2), 2-methyloxirane (75-56-9), and pyridine (110-86-1). Experimental reference data in deuteriochloroform solution were taken from the Spectral

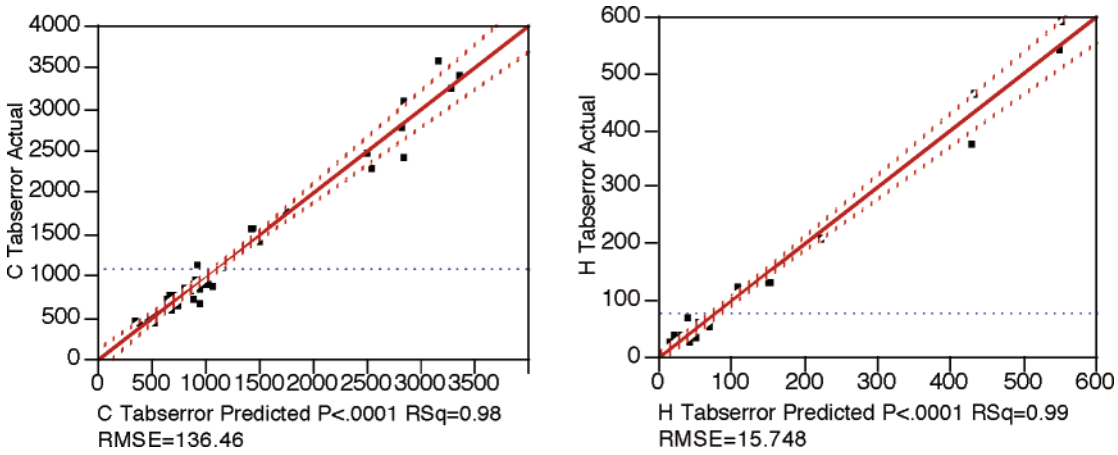

Figure 1. Response-surface predicted vs observed $|\Delta\delta|$ plots (ppm) with 99.99% confidence limits for ^{13}C (left) and ^{1}H (right) chemical shifts over the 59 parameter set choices.

Database for Organic Compounds, SDBS, organized by the National Institute of Advanced Industrial Science and Technology (AIST) of Japan (http://www.aist.go.jp/RIODB/SDBS).

Central composite design optimizations were performed using JMP 5.1.[52] Electronic structure calculations were carried out using the Gaussian 03 suite of programs.[53] Parameter settings (eq 2) were enforced in Gaussian 03 by use of the keywords BLYP, IOp(3/76 = 10000*nnnnn*), which sets P_2 to *nnnnn*/10000, IOp(3/77 = *mmmmmnnnnn*), which sets P_3 to *mmmmm*/10000 and P_4 to *nnnnn*/10000, and IOp(3/78 = *mmmmmnnnnn*), which sets P_5 to *mmmmm*/10000 and P_6 to *nnnnn*/10000.

Results and Discussion

Parameter Optimization. An initial selection of 28 points on the parameter response surface was made by JMP 5.1 and corresponding ^{13}C and ^{1}H $|\Delta\delta|$ values were computed. Subsequently, 31 additional points were selected in order to improve the polynomial fit of the response surfaces. The 59 parameter sets and their associated errors are provided as Supporting Information.

Full second-order polynomial response surfaces (11 degrees of freedom) were fit to the ^{13}C and ^{1}H data. Terms, coefficients, and term *t* ratios for the fitted surfaces are provided as Supporting Information. The fits for the ^{13}C and ^{1}H surfaces provided Pearson correlation coefficient R^2 of 0.9753 and 0.9866 and *F* ratios of 169 and 256, respectively; these are reasonable levels of statistical significance. Observed vs predicted values for $|\Delta\delta|$ are plotted in Figure 1 with 99.99% confidence limits. Analysis of the term *t* ratios for the ^{13}C surface indicates only modest sensitivity to coefficients of the correlation functional (absolute *t* ratios of 1.44 and 0.97 for P_5 and P_6, respectively). All other terms have absolute *t* ratios ranging from 3.3 to 31.7 with the exception of P_2, which has an absolute *t* ratio of 1.74. Interestingly, the ^{1}H surface exhibits different sensitivity to the primary terms. In the case of ^{1}H, the least important terms are the gradient corrections to the exchange and correlation functionals (absolute *t* ratios of 0.06 and 0.48 for P_3 and P_5, respectively). All other terms have absolute *t* ratios ranging from 2.3 to 36.0 with the exception of P_2, which has an absolute *t* ratio of 1.63. The lower sensitivity of the ^{1}H surface to gradient corrections likely reflects a lack of variation in reduced density gradients at the hydrogen nucleus (which is only a single proton), at least in carbon-bound environments. The relatively modest sensitivity of the two surfaces to the percentage of HF exchange in the functional suggests that this term plays more of a role in affecting the bonding in interatomic regions than it does at the nucleus, since it is certainly well-known that inclusion of HF exchange in hybrid functionals dramatically improves bond energies, for example.[49]

Table 1. Functionals Optimized for Prediction of ^{13}C and ^{1}H Chemical Shifts

functional	P_2	P_3	P_4	P_5	P_6
WC04	0.7400	0.9999	0.0001	0.0001	0.9999
WP04	0.1189	0.9614	0.9999	0.0001	0.9999

Table 2. Mean (ME), Mean Unsigned (MUE), and Root-Mean Square (RMSE) Errors (ppm) in Predicted Absolute ^{13}C and ^{1}H Chemical Shifts[a]

	^{13}C			^{1}H		
theory	ME	MUE	RMSE	ME	MUE	RMSE
WC04	0.7	3.1	3.8	0.06	0.13	0.20
WP04	6.4	6.4	7.6	0.01	0.09	0.13
HF	5.2	5.8	8.3	0.05	0.17	0.27
B3LYP	6.4	6.4	7.7	0.08	0.12	0.19
PBE1	5.5	5.5	6.9	0.07	0.13	0.22
*m*PW1PW91	5.6	5.6	7.0	0.07	0.13	0.21

[a] All calculations used PCM(chloroform)/B3LYP/6-31G(d) geometries and chemical shifts were computed at the PCM(chloroform)/method/6-311+G(2d,p) level.

Based on the fitted surfaces, global minimum parameter values were identified (Table 1). The performance of the optimized WC04 and WP04 functionals for chemical shift prediction over the training set is compared to four other methods in Table 2. The other methods are Hartree−Fock theory (which is generally regarded as insufficiently accurate for first-principles calculations),[49] the one-parameter hybrid generalized-gradient approximation (GGA) functionals PBE1 and *m*PW1PW91, and the three-parameter hybrid GGA functional B3LYP.

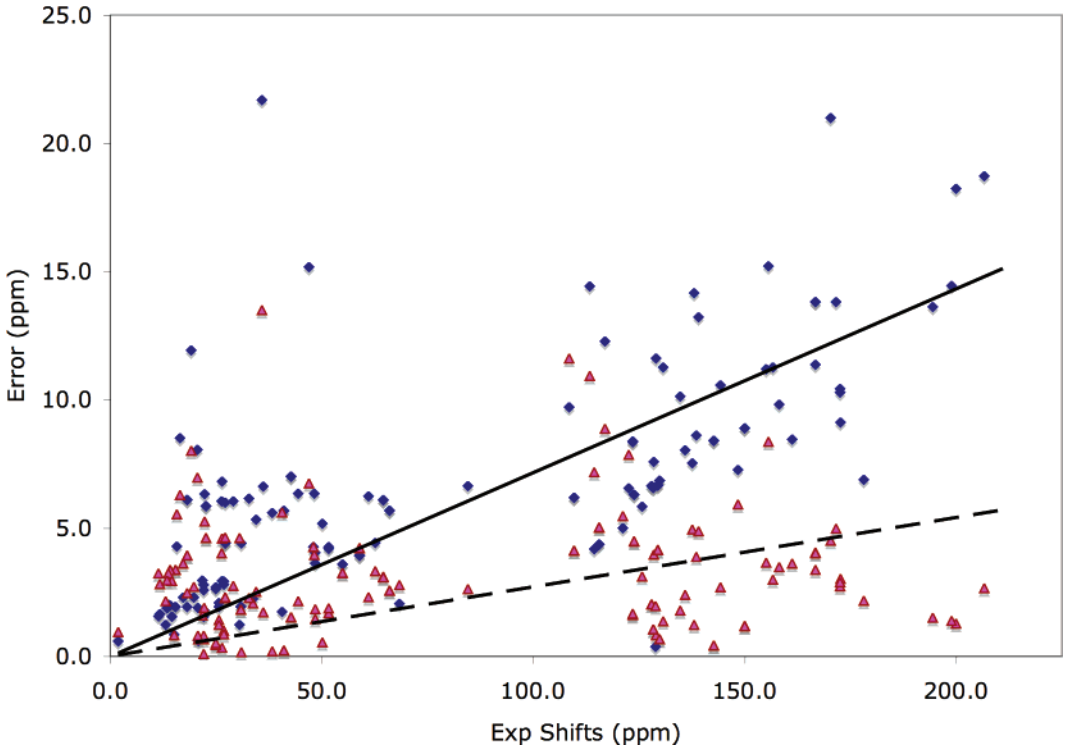

Figure 2. Absolute errors in 141 ^{13}C chemical shift predictions for B3LYP (blue diamonds) and WC04 (magenta triangles) as a function of chemical shift. The trendlines are linear fits to the errors forced to include the origin.

The optimized parameter values in Table 1 are quite different for WC04 and WP04, and both sets of parameters are themselves significantly different from the popular B3LYP functional. On the one hand, it is not particularly desirable to have to perform two separate calculations in order to obtain chemical shifts for the two different magnetically active nuclei. On the other hand, as Table 2 makes clear, the differences in parameter values lead to substantial differences in predictive accuracy. For ^{13}C chemical shifts, WC04 is more than twice as accurate as WP04. It is also substantially more accurate than either HF or any of the 3 hybrid GGA functionals. The nearest competitor is PBE1, which has an error that is 79% larger than WC04.

The variation in model accuracy over the ^{1}H training set data is smaller compared to the ^{13}C data set but still substantial. The WP04 method shows the highest accuracy, with the next nearest competitor, B3LYP, having an error that is 28% larger. The errors for WC04, PBE1, and *m*PW1PW91 are all within a few tenths of a ppm of one another in magnitude, while that for HF is substantially higher.

An analysis of errors as a function of the experimental chemical shift values indicates that the optimized functionals are more robust than B3LYP especially in the downfield regions of the spectrum (Figures 2 and 3). For ^{13}C data, WC04 remains more accurate in the upfield region by a statistically significant amount, but the magnitude is smaller. For ^{1}H data, WP04 and B3LYP are both accurate to within 0.1 ppm for most chemical shifts between 0 and 4 ppm.

Improvements from Linear Regression. Errors in chemical shift predictions from standard functionals have previously been shown to be reasonably systematic in various test sets,[30−34] so that substantial improvements in accuracy may be obtained from linear regression of the predicted data on experimental data. We have examined this approach for the various functionals and our training set, and the results are summarized in Table 3. Not surprisingly, since they have been optimized by design, WC04 and WP04 have regression slopes and intercepts most near unity for their respective nuclei. However, the performance of the other approaches improves substantially following regression. While WP04 continues to have the best accuracy for ^{1}H (albeit by a very

Table 3. Mean (ME), Mean Unsigned (MUE), and Root-Mean Square (RMSE) Errors (ppm) in Predicted ^{13}C and ^{1}H Chemical Shifts Following Linear Regression[a,b]

	^{13}C			^{1}H		
theory	ME	MUE	RMSE	ME	MUE	RMSE
WC04	0.0	3.0	3.8	0.00	0.11	0.14
WP04	0.0	2.3	3.5	0.00	0.07	0.10
HF	0.0	2.8	3.9	0.00	0.12	0.17
B3LYP	0.0	2.1	3.0	0.00	0.07	0.10
PBE1	0.1	1.8	2.8	0.00	0.08	0.11
*m*PW1PW91	0.0	1.8	2.8	0.00	0.08	0.11
regression data	*m*	*b*	*R*	*m*	*b*	*R*
WC04	1.0032	−0.9647	0.9958	0.9451	0.1157	0.9943
WP04	0.9601	−3.0273	0.9964	0.9587	0.1127	0.9969
HF	0.9164	1.7078	0.9955	0.9077	0.2318	0.9925
B3LYP	0.9488	−2.1134	0.9973	0.9333	0.1203	0.9974
PBE1	0.9486	−1.257	0.9977	0.9169	0.1895	0.9969
*m*PW1PW91	0.9487	−1.3423	0.9977	0.9191	0.1834	0.9977

[a] All calculations used PCM(chloroform)/B3LYP/6-31G(d) geometries and chemical shifts were computed at the PCM(chloroform)/method/6-311+G(2d,p) level. [b] Regression data are slopes (*m*), intercepts (*b*), and Pearson correlation coefficients (*R*).

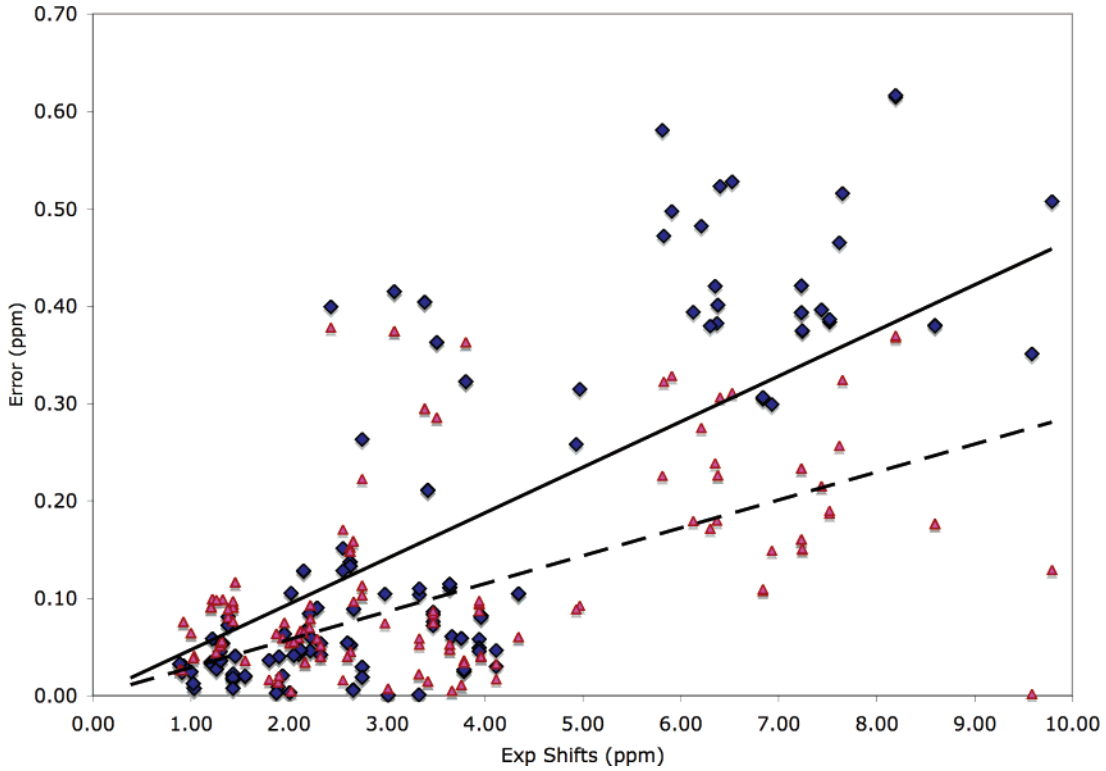

Figure 3. Absolute errors in 255 ^{1}H chemical shift predictions for B3LYP (blue diamonds) and WC04 (magenta triangles) as a function of chemical shift. The trendlines are linear fits to the errors forced to include the origin.

small margin), WC04 shows decreased accuracy compared to the other density functionals.

Applications to Molecules Not Included in the Training Set. To further assess the range of applicability of the WC04 and WP04 models, we applied them to the six isomeric *n*-methylcyclohexanols (n = 2, 3, or 4) and also to a pharmaceutically relevant β-lactam. Results for the methylcyclohexanols are provided in Table 4, which also summarize predictions using B3LYP. In those stereoisomers where both ring substituents must be either simultaneously axial or equatorial in the chair conformer, only the latter was considered because of its much greater stability. In those stereoisomers where one substituent must be axial, the lower-energy conformer always had the hydroxyl group axial, consistent with its lower *A*,[54] but the population of the other chair conformer was accounted for using a Boltzmann weighting. In every chair that was considered, a Boltzmann weighting over all hydroxyl rotamers was also applied.

For every isomer, the WC04 functional is two to three times more accurate than the B3LYP functional for ^{13}C chemical shifts prior to any correction through linear regression. After linear regression, the two models are about equally accurate. For ^{1}H chemical shifts, raw B3LYP is roughly twice as accurate as WP04, although both methods are accurate to within 0.1 ppm for these particular molecules. Except for the proton attached to the hydroxyl-substituted carbon, all of the resonances in the methylcyclohexanols are in the far upfield region where the performances of WP04 and B3LYP are generally similar. Linear regression in this case has fairly little effect on the B3LYP predictions but improves the WP04 predictions so that the two models are comparable in accuracy.

Table 4. Mean Unsigned Errors (MUEs, ppm) in Predicted ^{13}C[a] and ^{1}H[b] Chemical Shifts for Methylcyclohexanol Stereoisomers[c]

	^{13}C		^{1}H	
compound	WC04	B3LYP	WP04	B3LYP
cis-2-methylcyclohexanol	2.10[d]	4.75	0.056	0.041
	(1.65)	(1.21)	(0.032)	(0.050)
trans-2-methylcyclohexanol	2.03	4.82	0.072	0.037
	(1.69)	(1.18)	(0.059)	(0.042)
cis-3-methylcyclohexanol	1.60	4.78	0.084	0.051
	(1.29)	(1.14)	(0.067)	(0.052)
trans-3-methylcyclohexanol	1.73	4.96	0.080	0.033
	(1.23)	(1.32)	(0.045)	(0.052)
cis-4-methylcyclohexanol	1.50	4.73	0.095	0.031
	(0.87)	(1.15)	(0.055)	(0.026)
trans-4-methylcyclohexanol	1.57	4.77	0.081	0.041
	(1.08)	(1.04)	(0.054)	(0.033)
average MUE	1.76	4.80	0.078	0.039
	(1.30)	(1.17)	(0.052)	(0.042)

[a] MUEs are $|\Delta\delta|$ as defined in eq 3 divided by 7. [b] MUEs are $|\Delta\delta|$ as defined in eq 3 divided by 13; the hydroxyl proton is not included. [c] All calculations used PCM(chloroform)/B3LYP/6-31G(d) geometries and chemical shifts were computed at the PCM(chloroform)/method/6-311+G(2d,p) level. [d] Raw errors are listed above, and errors after linear regression (using regression equations derived from training set data) are listed below in parentheses.

To more comprehensively examine the utility of the WP04 functional, we next consider the pharmaceutically relevant, heteroatom-rich molecule (+)-(2*S*,5*R*,6*R*)-3,3-dimethyl-7-oxo-6-phthalimido-4-thia-1-azabicyclo(3.2.0)heptane-2-carboxylic acid (**1**, Figure 4). We averaged chemical shifts for **1** over a family of 18 conformers according to eq 1 and compared them to experimental ^{13}C- and ^{1}H NMR data in

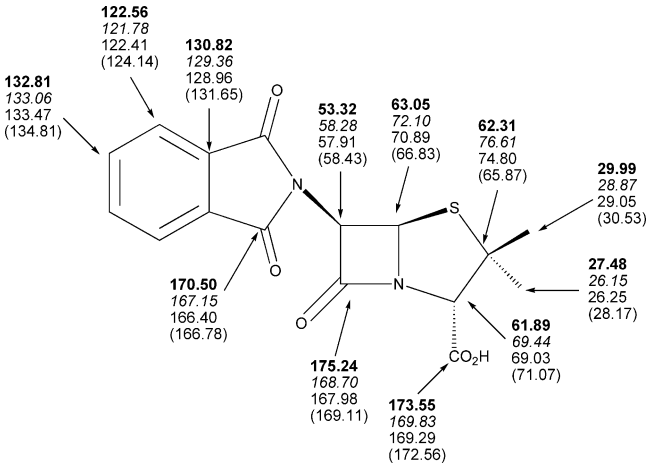

Figure 4. Predicted (after linear regression using regression equations developed on training set data for B3LYP and PBE1) and observed carbon chemical shifts for **1** in chloroform (from top to bottom: WC04 bold; B3LYP italic; PBE1 roman; experiment in parentheses).

Table 5. Mean and Mean Unsigned Errors (ppm) in Predicted Chemical Shifts for 1[a]

	^{13}C		^{1}H	
level of theory	ME	MUE	ME	MUE
WC04	−1.1[b]	2.9	−0.02	0.23
	(−1.7)	(3.3)	(−0.15)	(0.16)
WP04	8.0	8.0	0.01	0.10
	(0.3)	(3.1)	(−0.06)	(0.06)
HF	5.9	7.5	0.10	0.30
	(−4.7)	(4.8)	(−0.09)	(0.14)
B3LYP	7.9	7.9	0.11	0.15
	(−0.3)	(2.4)	(−0.07)	(0.07)
PBE1	6.7	6.7	0.14	0.19
	(−0.6)	(**2.2**)	(−0.05)	(**0.05**)
*m*PW1PW91	6.8	6.8	0.13	0.18
	(−0.6)	(**2.2**)	(−0.05)	(0.06)

[a] MUEs are $|\Delta\delta|$ as defined in eq 3 divided by the number of relevant nuclei. All calculations used PCM(chloroform)/B3LYP/6-31G(d) geometries and chemical shifts were computed at the PCM(chloroform)/method/6-311+G(2d,p) level. [b] Raw errors are listed above, and errors after linear regression (using regression equations derived from training set data) are listed below in parentheses.

$CDCl_3$.[55] The performances of the WC04, WP04, B3LYP, PBE1, and *m*PW1PW91 functionals are summarized in Table 5, as are results from HF theory. Specific ^{13}C chemical shifts are also provided in Figure 4 from experiment, WC04, B3LYP, and PBE1. This molecule, substantially decorated with polar functionality (giving rise to more downfield chemical shifts), demonstrates the superior performances of the WC04 and WP04 functionals in terms of mean and mean unsigned errors prior to linear regression. After linear regression, however, the PBE1 functional is overall the most accurate for both nuclei, and it is the only one of those listed in Figure 4 that correctly predicts the relative positions of the ^{13}C chemical shifts for the carboxyl carbons of the carboxylic acid and the β-lactam.

Solvation Effects. As a point of technical as well as practical interest, we examined the degree to which chloroform solvation, as implemented via the PCM continuum solvation model, influenced the predicted chemical shifts. In particular, for the training set we computed WC04 and WP04 chemical shifts for gas-phase densities at gas-phase geometries, for gas-phase densities at solvated geometries, and for solvated densities at solvated geometries (as already discussed above). The corresponding $|\Delta\delta|$ values are provided in Table 6.

Table 6. Mean Unsigned Errors (MUE, ppm) in Predicted WC04 ^{13}C and WP04 ^{1}H Chemical Shifts as a Function of Computational Protocol[a]

model	^{13}C	^{1}H
gas//gas	3.3	0.14
gas//chloroform	3.2	0.14
chloroform//chloroform	3.1	0.10

[a] Errors $|\Delta\delta|$ are defined in eq 3. Calculations used either B3LYP(gas)/6-31G(d) or PCM(chloroform)/B3LYP/6-31G(d) geometries (indicated after the double solidus), and chemical shifts were computed at either the W*x*04/6-311+G(2d,p) or PCM(chloroform)/W*x*04/6-311+G(2d,p) levels (indicated before the double solidus).

In the case of the ^{13}C data set, there is a reduction in error of about 4% when the geometries are relaxed in solution and an additional 4% when solvated densities are used for the NMR calculations. In the ^{1}H data set, on the other hand, the use of gas-phase densities with relaxed geometries actually increases the total error by a small amount. However, the use of solvated densities leads to a substantial improvement in the predictive accuracy.

In principle, there might be some value in optimizing parameter sets designed to predict chemical shifts in solution from gas-phase densities at gas-phase geometries.[34] Gas-phase calculations are efficient and, as long as solvation effects are systematic, the statistical approach might be expected to absorb deviations into the parameter set. However, the cost of including a continuum solvent model into a self-consistent reaction field model is typically no more than 15% or so of the total computational time,[56,57] so we consider it worthwhile to adopt this approach in order to more accurately capture the physics of solvation when the goal is to predict data for solutes in solution.

Klein et al.[58] have pointed out that a continuum model alone is generally *not* sufficient for the computation of the ^{17}O chemical shift of liquid water because of the strong, nonisotropic interactions between water oxygen atoms and the protons of neighboring water molecules. In that case, explicit supermolecular clusters surrounded by a continuum are required to accurately model the polarization of the system. However, interactions of this magnitude are unlikely to be associated with carbon atoms and nonheteroatom bound hydrogen atoms, so a pure continuum approach to account for solvation is within the spirit of the WC04 and WP04 models, which are designed to balance accuracy and efficiency for the interpretation of NMR spectra.

Acknowledgment. This work was supported in part by the National Science Foundation (C.J.C., CHE-0203346). We thank Dr. Vadim Dvornikov and Ziyad Al-Rashid for providing ^{13}C- and ^{1}H NMR chemical shift data for the methylcyclohexanols and discussions. Loren Greenman,

Corey Stotts, and Dr. Jason Thompson are thanked for programming assistance.

Supporting Information Available: Initial and subsequent design worksheets and computed responses, terms, and coefficients for the ^{13}C and ^{1}H response surfaces, Cartesian coordinates and electronic energies for all conformers of all molecules, and experimental chemical shift data for the methylcyclohexanols. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Claridge, T. *High-resolution NMR Techniques in Organic Chemistry*; Elsevier: London, 1999.

(2) Lipinski, C. A.; Lombardo, F.; Domin, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.

(3) Uccello-Barretta, G.; Balzano, F.; Sicoli, G.; Friglola, C.; Aldana, I.; Monge, A.; Paolino, D.; Guccione, S. *Bioorg. Med. Chem.* **2004**, *12*, 447−458.

(4) Lipinski, C. A. *Pharmacol. Toxicol. Methods* **2001**, *44*, 235−249.

(5) Pretsch, E.; Bühlmann, P.; Affolter, C. *Structure Determination of Organic Compounds: Tables of Spectral Data*; Springer: Berlin, 2003.

(6) Abraham, R. J.; Byrne, J. J.; Griffiths, L.; Koniotou, R. *Magn. Reson. Chem.* **2005**, *43*, 611−624.

(7) Klemp, C.; Bruns, M.; Gauss, J.; Haussermann, U.; Stosser, G.; van Wullen, L.; Jansen, M.; Schnockel, H. *J. Am. Chem. Soc.* **2001**, *123*, 9099−9106.

(8) Ochsenfeld, C.; Brown, S. P.; Schnell, I.; Gauss, J.; Spiess, H. W. *J. Am. Chem. Soc.* **2001**, *123*, 2597−2606.

(9) Barone, G.; Duca, D.; Silvestri, A.; Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Chem.-- Eur. J.* **2002**, *8*, 3240−3245.

(10) Barone, G.; Gomez-Paloma, L.; Duca, D.; Silvestri, A.; Riccio, R.; Bifulco, G. *Chem.-- Eur. J.* **2002**, *8*, 3233−3239.

(11) Price, D. R.; Stanton, J. F. *Org. Lett.* **2002**, *4*, 2809−2811.

(12) Balandina, A.; Mamedov, V.; Franck, X.; Figadère, B.; Latypov, S. *Tetrahedron Lett.* **2004**, *45*, 4003−4007.

(13) *Calculation of NMR and EPR Parameters: Theory and Applications*; Kaupp, M., Bühl, M., Malkin, V. G., Eds.; John Wiley & Sons: New York, 2004.

(14) Wiberg, K. B. *J. Comput. Chem.* **1999**, *20*, 1299−1303.

(15) Wiberg, K. B.; Hammer, J. D.; Zilm, K. W.; Keith, T. A.; Cheeseman, J. R.; Duchamp, J. C. *J. Org. Chem.* **2004**, *69*, 1086−1096.

(16) Adamo, C.; Barone, V. *Chem. Phys. Lett.* **1998**, *298*, 113−119.

(17) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497−5509.

(18) Malkin, V. G.; Malkina, O. L.; Eriksson, L. A.; Salahub, D. R. In *Modern Density Functional Theory; A Tool for Chemistry*; Politzer, P., Seminario, J., Eds.; Elsevier: Amsterdam, 1995; Vol. 2, p 273−374.

(19) Bühl, M.; Kaupp, M.; Malkina, O. L.; Malkin, V. G. *J. Comput. Chem.* **1999**, *20*, 91−105.

(20) Schreckenbach, G.; Ziegler, T. *Theor. Chem. Acc.* **1998**, *99*, 71−82.

(21) Adamo, C.; Cossi, M.; Barone, V. *J. Mol. Struct. (THEOCHEM)* **1999**, *493*, 145−147.

(22) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233−9241.

(23) Allen, M. J.; Keal, T. W.; Tozer, D. J. *Chem. Phys. Lett.* **2003**, *380*, 70−77.

(24) Malkin, V. G.; Malkina, O. L.; Casida, M. E.; Salahub, D. R. *J. Am. Chem. Soc.* **1994**, *116*, 5898−5908.

(25) Fadda, E.; Casida, M. E.; Salahub, D. R. *Int. J. Quantum Chem.* **2003**, *91*, 67−83.

(26) Wilson, P. J.; Amos, R. D.; Handy, N. C. *Chem. Phys. Lett.* **1999**, *312*, 475−484.

(27) Patchkovskii, S.; Autschbach, J.; Ziegler, T. *J. Chem. Phys.* **2001**, *115*, 26−42.

(28) Poater, J.; van Lenthe, E.; Baerends, E. J. *J. Chem. Phys.* **2003**, *118*, 8584−8593.

(29) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 3015−3024.

(30) Rablen, P. R.; Pearlman, S. A.; Finkbiner, J. *J. Phys. Chem. A* **1999**, *103*, 7357−7363.

(31) Sebag, A. B.; Forsyth, D. A.; Plante, M. A. *J. Org. Chem.* **2001**, *66*, 7967−7973.

(32) Giesen, D. J.; Zumbulyadis, N. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5498−5507.

(33) Wang, B.; Fleischer, U.; Hinton, J. F.; Pulay, P. *J. Comput. Chem.* **2001**, *22*, 1887−1895.

(34) Wang, B.; Hinton, J. F.; Pulay, P. *J. Comput. Chem.* **2002**, *23*, 492−497.

(35) Patchkovskii, S.; Thiel, W. *J. Comput. Chem.* **1999**, *20*, 1220−1245.

(36) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(37) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(38) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(39) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(40) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(41) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032−3041.

(42) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.

(43) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.

(44) London, F. *J. Phys. Radium (Paris)* **1937**, *8*, 397−409.

(45) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.

(46) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629−3643.

(47) Cammi, R. *J. Chem. Phys.* **1998**, *109*, 3185−3196.

(48) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1999**, *110*, 7627−7638.

(49) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons: Chichester, 2004.

(50) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.

(51) Khuri, A. I. *Response Surfaces: Designs and Analyses*; Marcel Dekker: New York, 1996.

(52) *JMP version 5.1*; SAS Institute Inc.: Cary, NC, 2005.

(53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; 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.; 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 (Revision B.05)*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(54) Eliel, E. L.; Wilen, S. H. *Stereochemistry of Organic Compounds*; John Wiley and Sons: New York, 1994.

(55) Fekner, T.; Baldwin, J. E.; Adlington, R. M.; Jones, T. W.; Prout, C. K.; Schofield, C. J. *Tetrahedron* **2000**, *56*, 6053−6074.

(56) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.

(57) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999−3093.

(58) Klein, R. A.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2004**, *108*, 5851−5863.

CT6001016