MNDO Parameters for the Prediction of ^{19}F NMR Chemical Shifts in Biologically Relevant Compounds

Duane E. Williams, Martin B. Peters, Bing Wang, and Kenneth M. Merz, Jr.*

Department of Chemistry and Quantum Theory Project, 2328 New Physics Building, P.O. Box 118435, University of Florida, Gainesville, Florida 32611-8435

Received: February 25, 2008; Revised Manuscript Received: May 20, 2008

The semiempirical MNDO methodology for qualitative description NMR chemical shifts has now been extended with the addition of NMR-specific parameters for the fluorine atom. This approach can be employed using semiempirical (AM1/PM3) geometries with good accuracy and can be executed at a fraction of the cost of *ab initio* and DFT methods, providing an attractive option for the computational studies of ^{19}F NMR for much larger systems. The data set used in the parametrization is large and diverse and specifically geared toward biologically relevant compounds. The new parameters are applicable to fluorine atoms involved in carbon−fluorine bonds. These parameters yield results comparable to NMR calculations performed at the DFT (B3LYP) level using the 6−31++G(d,p) basis set. The average R^2 and rms error for this data set is 0.94 and 13.85 ppm, respectively, compared to 0.96 and 10.45 ppm when DFT methods are used.

Introduction

Fluorinated compounds with impressive biological activity have drawn significant attention to the important role that fluorine plays in the world of biochemistry.[1] The utility of fluorine (^{19}F) NMR in the study of such biological compounds has been discussed in detail in previous publications.[1–5] A few advantages that fluorine NMR provides relative to the more common ^{1}H and ^{13}C NMR methods derive from the natural properties of the fluorine atom. ^{19}F is a spin 1/2 nucleus and is 100% abundant, making it highly amenable to NMR studies. In biological systems the fluorine atom serves as a relatively small steric "footprint." In many instances it can replace a hydrogen atom in a ligand with influence on an event of molecular recognition, and on the subsequent biological response which can range from minimal to beneficial.[1] The fluorine nucleus can be incorporated into proteins and other biological compounds by routine methods including chemical synthesis and biosynthesis using a living organism.[1] The ^{19}F nucleus is highly sensitive to its chemical environment because of the lone pairs of electrons it possesses. Consequently, in protein systems ^{19}F spin-labeled amino acids provide a useful probe for determining conformational change in a specific area of a protein.[6]

Given that endogenous fluorine is found in biological systems quite infrequently,[7] the NMR spectra can typically be attributed solely to ^{19}F labeling; the resulting spectra contain more distinct signals and are easier to interpret than proton NMR. The simplified ^{19}F NMR spectra can provide valuable information about protein−drug complexes, thereby aiding in the characterization of potential lead compounds at a reduced opportunity cost. Because of the positive effects of fluorine labeling and the additional value of the fluorine atom as a critical component of several highly effective drug compounds currently on the market,[8,9] NMR screening monitoring the ^{19}F nucleus of fluorine containing ligands as selective markers has gained popularity.[10–13]

The ability to make quick, routine, and accurate predictions of fluorine NMR chemical shifts for biological molecules (ranging in size from a few hundred atoms to many thousands) can aid in structure elucidation, give insight into the binding modes of ligands in proteins, and add valuable information on dynamics of biological systems. The ideal method of calculating fluorine chemical shifts in the biological environment should be versatile and sufficiently rapid to enable it to be applied to large systems. Empirical models used to estimate fluorine chemical shifts for protein have been well established and are often very useful.[14,15] Modern QM approaches offer several advantages over empirical or classically based approaches to the prediction of NMR chemical shifts. A few of the main advantages are given below:

One advantage is that the electrostatic representation of QM approaches is more accurate because environmental, conformational, polarization, and charge transfer effects are explicitly included, in contrast to the process used in simplified models. A second is that the ring current effect (the total magnetic effect felt as a result of the magnetic fields in the system) is an inherent part of the QM model and does not have to be built into the model in a parametric sense. A third advantage is that the ways in which quantum chemical methods can be improved have been thoroughly documented (e.g., inclusion of correlation and improvements in the basis set). Furthermore, structural biologists have generated an immense amount of experimental data that can be used to validate quantum chemical approaches to the study of NMR chemical shifts in biomolecules.

In principle, QM calculations are able to include each of the effects that influence a given chemical shift. This lends QM methods the beneficial quality that molecular mechanics (MM) methods lack, that of being more generalized and applicable to the variety of organic molecules of interest as ligands in biochemistry studies. Unfortunately, the expense of *ab initio* and density functional theory (DFT) computation prohibits their application to protein systems containing thousands of atoms.[16]

We are currently unaware of any of *ab initio* techniques that have been shown to routinely and quickly calculate the NMR chemical shifts of large systems with significant nonbonded interactions (such as proteins) which are of interest to NMR spectroscopists. The finite perturbation theory (FPT) developed

* Corresponding author: merz@qtp.uf.edu.

10.1021/jp801649f CCC: $40.75
Published on Web 08/23/2008

in the framework of the MNDO[17] Hamiltonian using gauge-including atomic orbitals (GIAO) has shown promising results for the calculation of ^{1}H and ^{13}C chemical shifts.[18] We have recently implemented the FPT-MNDO GIAO method by using a divide and conquer strategy for the diagonalization of the Fock matrix.[19] This method has previously been implemented using NMR-specific parameters developed for ^{1}H,^{13}C, ^{15}N, and ^{17}O and has been shown to give fast and accurate results that can be applied to large protein−ligand complexes.[19,20] Here we expand the scope of our approach to handle ^{19}F chemical shift calculations.

Methods

Parameterization. The systematic underestimation of excitation energies using the standard MNDO parameters results in an overestimation of the variation of the paramagnetic contribution to the NMR chemical shifts.[21] Patchkovskii and Thiel[21] have generated NMR-specific MNDO parameters resulting in significant improvement in the agreement between experimental and calculated NMR chemical shifts for ^{1}H,^{13}C, ^{15}N, and ^{17}O. In the same work, the authors also provided a full explanation of the choice of parameters to be optimized. Our method for the fast semiempirical QM NMR calculations using these parameters has previously been published.[20] Here the first addition to this set of MNDO-NMR parameters is presented.

In the previous MNDO-NMR parametrization,[21] the authors ran the initial tests avoiding the alteration of the one center/one electron terms ($U_{ss/pp}$). These parameters affect the core energies and heats of formation of single atoms, and large alterations can significantly change the charges, dipole moments, and electronic structure. Since these parameters in the standard MNDO formalism were optimized to reproduce the aforementioned quantities, and are based on the ionization potentials and electron affinities,[22] only the Slater atomic orbital exponents ($\xi_{s/p}$) and atomic orbital two center/one electron resonance parameters ($\beta_{s/p}$) were changed, leaving $U_{ss/pp}$ terms at the MNDO value. While only the ξ and β parameters were directly changed, the derived MNDO parameters were recalculated as appropriate.

Chemical shifts are calculated as the difference between the calculated shielding constant and some reference value as illustrated in eq 1.

$$\delta_{\text{calculated}} = \sigma_{\text{reference}} - \sigma_{\text{calculated}} \tag{1}$$

Previously, the MNDO-NMR parameters for H, C, N, and O were generated by minimizing the difference between the experimental and calculated chemical shifts via the goal function shown in eq 2.

$$G = \sum_{\text{X=H,C,N,O}} W^{\text{X}} \times \sqrt{\frac{1}{N^{\text{X}}} \sum_{i=1}^{N^{\text{X}}} (\delta_i^{\text{X}} + \sigma_i^{\text{X}} - \sigma_{\text{ref}}^{\text{X}})^2} + P \tag{2}$$

Because the parameters were generated simultaneously for all four atoms, a weighting factor, W^{X}, was used to ensure that the difference in chemical shift ranges for each atom X was taken into account. In this equation, N^{X} is the number of chemical shifts for atom X. The reference shielding constants, $\sigma_{\text{ref}}^{\text{X}}$, were values chosen to minimize the overall rms deviations from experiment for each atom, and δ_i^{X} and σ_i^{X} are the experimental chemical shifts and calculated absolute shielding respectively. This equation includes a penalty function, shown in eq 3, which was introduced in order to avoid searching in areas of parametric space that would cause large deviations from the known properties of the single atom. This penalty function specifically prevents the calculated energy, e_i, of a single atom from straying too far from the atom's experimental energy, E_i. In the current method, the one center/one electron terms were kept at the standard values, resulting in no variation of the atomic energies, thereby rendering the penalty function unnecessary.

$$P = \sum_{\text{X=H,C,N,O}} w^{\text{X}} \times \sum_{E_i \le E_{\max}} (E_i^{\text{X}} - e_i^{\text{X}})^2 \tag{3}$$

In the present parametrization, the σ_{ref} value used in eq 1 was initially set to the ^{19}F shielding constant that was calculated for $CFCl_3$ for each new set of parameters tested. This was done in an effort to be more consistent with experimental studies, which often use $CFCl_3$ as the standard. However, in many cases parameters performed well for most other compounds but did not perform as well for $CFCl_3$. In these instances the large rms error was a false indication that the parameters yielded poor results for all compounds, when in fact a more appropriate σ_{ref} was needed (cf. Tables 2 and 3). This problem was addressed by calculating the average signed error (see eq 4) between the experimental chemical shifts and the calculated shielding constants and using this average signed error as the σ_{ref} value.

$$\text{average signed error} = \frac{\sum (\delta_{\text{exp}} - \sigma_{\text{calcd}})}{N} \tag{4}$$

For a given set of parameters, this method essentially sets the average signed error to zero and minimizes the rms error for the data set. The scoring function then optimized the parameters by minimizing the rms error between experimental and resulting calculated chemical shifts via eq 5. All reported chemical shifts are given using the σ_{ref} value chosen to minimize the average signed and rms error.

$$G = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\delta_i + \sigma_i - \sigma_{\text{ref}})^2} \tag{5}$$

Since the goal of this project was to provide parameters that could be used for large biological systems such as ligands bound to proteins, semiempirical geometries were used in the parametrization because the large size of proteins and other biomol-

TABLE 1: Comparison of Standard and NMR-Optimized MNDO Parameters

atom	parameter	MNDO	MNDO-NMR
H	ξ_s (a.u)	1.3319670	1.1782700[a]
	β_s (eV)	−6.9890640	−15.2092800[a]
C	ξ_s (a.u)	1.7875370	1.6544500[a]
	ξ_p (a.u)	1.7875370	1.6544500[a]
	β_s (eV)	−18.9850440	−18.5418100[a]
	β_p (eV)	−7.9341220	−12.8114400[a]
N	ξ_s (a.u)	2.2556140	2.0265800[a]
	ξ_p (a.u)	2.2556140	2.0265800[a]
	β_s (eV)	−20.4957580	−23.5315500[a]
	β_p (eV)	−20.4957580	−23.5315500[a]
O	ξ_s (a.u)	2.6999050	2.2027400[a]
	ξ_p (a.u)	2.6999050	2.2027400[a]
	β_s (eV)	−32.6880820	−19.8048500[a]
	β_p (eV)	−32.6880820	−19.8048500[a]
F	ξ_s (a.u)	2.8484870	4.98552600
	ξ_p (a.u)	2.8484870	3.75447400
	β_s (eV)	−48.2904660	−32.70953400
	β_p (eV)	−36.5085400	−87.01227727

[a] Values currently used for MNDO-NMR calculations in the Wang−Merz method; these values were parametrized by Patchkovskii et al.[21]

TABLE 2: Comparison of Errors Associated with Each Method for 100 Compounds Using the Shielding Constant Calculated for $CFCl_3$ as the σ_{ref} value in Eq 1

		NMR method		
	geometry	B3LYP/6−31++G(d,p)	MNDO	MNDO-NMR
RMS error	AM1	16.42	92.41	90.21
	PM3	22.51	93.72	92.90
	B3LYP/6−31G(d,p)	13.79	94.29	89.47
R^2	AM1	0.94	0.38	0.94
	PM3	0.95	0.36	0.94
	B3LYP/6−31G(d,p)	0.96	0.33	0.93
average signed error	AM1	6.81	−81.78	−89.14
	PM3	−15.04	−82.83	−91.94
	B3LYP/6−31G(d,p)	8.26	−82.96	−88.23
average unsigned error	AM1	12.79	81.92	89.14
	PM3	19.82	83.95	91.94
	B3LYP/6−31G(d,p)	10.59	83.12	88.23
reference (σ_{ref}) value	AM1	172.29	69.07	466.66
	PM3	188.91	75.10	470.65
	B3LYP/6−31G(d,p)	179.19	74.68	464.54

TABLE 3: Comparison of Errors Associated with Each Method for 100 Compounds Using the Average Signed Error as the σ_{ref} Value in Eq 1

		NMR method		
	geometry	B3LYP/6−31++G(d,p)	MNDO	MNDO-NMR
rms error	AM1	14.95	43.03	13.85
	PM3	16.75	43.85	13.35
	B3LYP/6−31G(d,p)	10.45	44.80	14.80
R^2	AM1	0.94	0.38	0.94
	PM3	0.95	0.36	0.94
	B3LYP/6−31G(d,p)	0.96	0.33	0.93
average signed error	AM1	0.00	0.00	0.00
	PM3	0.00	0.00	0.00
	B3LYP/6−31G(d,p)	0.00	0.00	0.00
average unsigned error	AM1	11.77	35.90	10.73
	PM3	12.96	36.69	10.34
	B3LYP/6−31G(d,p)	6.75	37.65	11.38
reference (σ_{ref}) value	AM1	178.93	−12.71	377.51
	PM3	173.87	−7.73	378.77
	B3LYP/6−31G(d,p)	187.57	−8.28	376.31

ecules precludes the use of structures obtained via higher level calculations. Although our current procedure for NMR chemical shift calculations uses the MNDO Hamiltonian, geometry optimization was carried out at the AM1[23] and PM3[24] levels because they are better able to capture some important geometric features, including hydrogen bonding, and are consequently more likely to be used for biological systems.[23,24] As discussed below, our results indicate that the choice of Hamiltonian for the geometry optimization has minimal effect on the quality of the NMR calculation. It is important to note that in all instances the calculations using new parameters are single point NMR calculations using the standard geometries noted (AM1, PM3, B3LYP[25,26]), since the new parameters are ^{19}F NMR specific and have not been tested for their ability to provide realistic geometries. In addition, the starting point for the parametrization incorporated parameters developed by Patchkovskii and Thiel for C, H, N, and O and standard MNDO parameters for all other atoms. Changes were made only to the fluorine parameters. Gauge-including atomic orbitals were used for all calculations.[16,27]

The original MNDO parameters[17] were not optimized for a high level of performance in NMR calculations. As a result, it was necessary to search over a large parameter space in order to find parameters that would yield a more accurate description of the NMR chemical shifts. Additionally, some of the parameters in the MNDO method are derived from others and have other features of interdependence. This results in a poorly defined parameter space with multiple minima. Minimizing the goal function (shown in eq 5) via more routine optimization schemes, such as BFGS or conjugate gradient, may keep the parameters within a local minimum. Finding the global minimum was not our specific goal, as that may have involved straying too far from parameters with any physical meaning; however, a genetic algorithm (GA) was used because it had the potential to find parameters covering a broader range of parameter space. Furthermore, a GA has been successfully used previously in several semiempirical parametrizations.[28–30] A full description of the type of GA used in this work in the generation of semiempirical parameters has been previously published.[28] The GA and all handling of the data were performed using an in-house molecular tools package MTK++ (Molecular Tool Kit [written in C++]). Semiempirical geometry optimizations and NMR calculations were performed using our DivCon program.[31]

The results obtained using the MNDO-NMR ^{19}F parameters are compared to those of DFT calculations performed at the B3LYP level. This comparison required the selection of the most appropriate basis set for the DFT calculations. There have been several investigations into the optimal geometries and basis sets for accurate NMR chemical shift calculations at the DFT level.[32–34]

Recently, a study of ^{19}F NMR chemical shift calculations highlighted the importance of diffuse functions to accurately reproduce experimental values.[34] The use of diffuse functions

TABLE 4: Comparison of NMR Chemical Shifts (ppm) Grouped to Facilitate Analysis

Table 4-A. Fluorinated aliphatic chains

	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
1		-267.90	-271.75	-197.86	-231.38	CH_2Cl_2	37-39
2		-219.02	-220.76	-186.13	-198.94	$CFCl_3$	40,41
3		-211.50	-217.43	-187.28	-197.75	CH_2Cl_2	37,42
4		-143.40	-143.09	-168.78	-140.72	$CFCl_3$	43,44
5†		-164.00	-176.82	-178.27	-175.06	$CFCl_3$	37,39,45
6		-78.60	-78.84	-146.23	-71.17	Neat	44,46,47
7		-63.50	-67.14	-134.81	-65.89	CCl_4	48-50
8		a. -74.60 b. -189.20	a. -71.33 b. -186.82	a. -126.48 b. -139.85	a. -78.00 b. -205.43	Not Reported	51,52
9†		-64.60	-61.96	-134.42	-35.61	Not Reported	44,46
	R^2		0.99	0.74	0.95		
	RMSE		5.04	52.21	17.98		
	Signed		1.96	12.49	-7.64		
	Unsigned		3.68	47.82	14.25		

Table 4-B. Chains containing heteroatoms

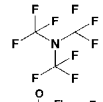
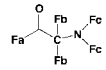
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
10		-226.00	-249.70	-189.67	-207.57	$CFCl_3$	53
11		-130.42	-128.97	-152.99	-138.25	Not Reported	54
12		a. -83.60	a. -82.63	a. -131.79	a. -80.14	$CDCl_3$	55
13†		-76.57	-74.45	-136.09	-78.92	CD_2Cl_2	56,57
14		-74.21	-76.38	-139.12	-77.90	DMSO	58,59
15		-232.30	-233.89	-179.73	-210.83	Not reported	60
16		a. -125.20	a. -123.72	a. -131.98	a. -136.45	$CFCl_3$	61
17		-55.80	-55.12	-122.43	-66.46	Not reported	62
18		a. 23.80	a. 36.92	a. -32.01	a. 9.92	Not Reported	63
19		33.24	46.76	-26.06	14.60	$CFCl_3$	13
20		-68.00	-68.17	-126.51	-69.10	Not Reported	54
	R^2		0.99	0.83	0.98		
	RMSE		10.67	53.43	16.04		
	Signed		-2.94	30.39	-0.06		
	Unsigned		7.28	49.96	12.00		

Table 4-C. Aliphatic rings

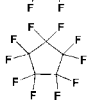
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
21		-165.81	-177.74	-177.81	-176.43	$CFCl_3$	64,37,65
22		-184.47	-194.05	-178.16	-176.70	$CFCl_3$	64,37,65
23		-135.15	-133.93	-133.05	-132.94	Not Reported	66,67
24†		-132.90	-129.64	-130.60	-138.64	Not Reported	66

Table 4-C. Aliphatic rings: Continued

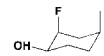
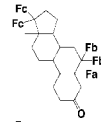
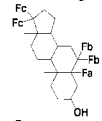
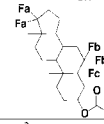
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
25		-142.30	-146.72	-159.62	-133.28	Not Reported	68-70
26		-197.50	-191.42	-180.14	-180.14	$CFCl_3$	71
27		-178.50	-195.46	-180.01	-179.64	$CFCl_3$	71
28		-196.10	-196.82	-174.40	-179.55	$CFCl_3$	71
29		-180.10	-196.44	-178.99	-182.61	$CFCl_3$	71
30		-194.00	-159.07	-155.95	-159.25	$CFCl_3$	72
31†		-191.00	-152.65	-161.29	-167.42	$CFCl_3$	72
32		-186.00	-154.56	-161.87	-168.27	$CFCl_3$	72
33		a. -164.08	a. -160.04	a. -130.78	a. -173.23	$CFCl_3$	54,73
34†		a. -167.87	a. -160.06	a. -130.24	a. -175.56	$CFCl_3$	54,73
35		a. -108.86 b. -107.06	a. -123.21 b. -115.49	a. -127.85 b. -123.87	a. -107.88 b. -109.42	$CFCl_3$	54,73
	R^2		0.59	0.56	0.81		
	RMSE		17.48	21.42	13.93		
	Signed		-2.75	-9.13	-5.65		
	Unsigned		13.11	17.46	10.59		

Table 4-D. Non-aromatic double bonds

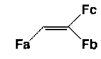
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
36		-113.00	-108.44	-119.90	-123.74	Not Reported	42
37		-81.30	-78.30	-110.23	-70.51	Not Reported	42
38		a. -205.00 b. -100.00 c. -126.00	a. -197.30 b. -93.36 c. -122.71	a. -135.07 b. -106.82 c. -121.74	a. -207.37 b. -87.59 c. -133.88	Not Reported	42,74
39†		-134.00	-126.63	-118.71	-139.69	Not Reported	42,75
40		a. -150.00 b. -117.80 c. -130.10	a. -142.83 b. -115.64 c. -127.90	a. -89.61 b. -135.50 c. -134.62	a. -128.35 b. -124.16 c. -136.25	Not Reported	66,76
41		a. -111.50 b. -109.10	a. -97.07 b. -104.83	a. -129.39 b. -75.06	a. -114.50 b. -122.54	Not Reported	77
42†		a. -108.40	a. -107.56	a. -138.06	a. -119.44	Not Reported	78
43		a. -114.10	a. -109.20	a. -133.84	a. -125.53	Not Reported	78
44		b. -120.20	b. -118.04	b. -133.73	b. -125.45	Not Reported	77

TABLE 4: Continued

Table 4-D. Non-aromatic double bonds: Continued

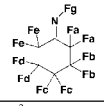
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
45		a. -114.50	a. -111.25	a. -131.99	a. -126.73	$CFCl_3$	79
	R^2		0.99	0.02	0.87		
	RMSE		5.92	30.07	10.49		
	Signed		-4.93	-1.38	3.38		
	Unsigned		4.93	23.48	9.36		

Table 4-E. Bicyclic compounds

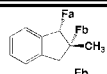
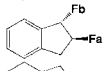
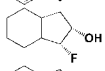
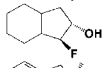
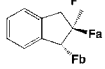
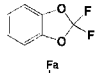
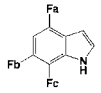
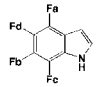
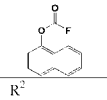
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
46		a. -190.40 b. -163.30	a. -208.38 b. -171.05	a. -168.40 b. -161.96	a. -187.69 b. -171.91	Not Reported	80
47		a. -200.50 b. -196.30	a. -198.15 b. -190.40	a. -171.60 b. -170.04	a. -182.73 b. -179.71	Not Reported	80
48		-177.60	-195.82	-177.35	-190.16	$CFCl_3$	71
49		-188.70	-195.10	-174.68	-184.63	$CFCl_3$	71
50†		a. -146.30	a. -156.70	a. -162.48	a. -158.24	Not Reported	80
51		-57.20	-51.05	-131.44	-96.43	Neat	81
52		a. -124.70 b. -146.10 c. -166.70	a. -126.28 b. -144.22 c. -169.47	a. -111.70 b. -115.02 c. -125.71	a. -130.02 b. -138.90 c. -155.24	CCl_4	82
53†		a. -149.40 b. -166.50 c. -163.10 d. -170.30	a. -147.39 b. -163.81 c. -164.07 d. -167.60	a. -110.74 b. -113.90 c. -121.74 d. -116.54	a. -145.89 b. -155.13 c. -157.28 d. -156.50	CCl_4	83
54†		-16.60	-12.26	-67.12	-13.16	CCl_4	84
	R^2		0.98	0.49	0..92		
	RMSE		7.15	37.52	14.11		
	Signed		2.01	-14.35	-1.62		
	Unsigned		5.51	31.96	11.33		

Table 4-F. Five member heterocycles

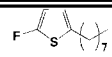
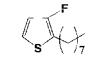
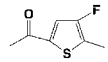
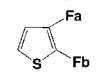
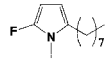
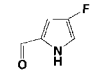
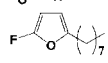
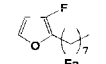
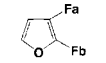
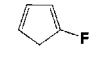
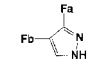
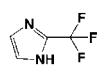
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
55		-133.70	-127.33	-95.22	-135.91	$CDCl_3$	85
56		-136.60	-141.45	-114.43	-130.56	$CDCl_3$	85
57		-131.30	-137.82	-113.34	-130.17	$CDCl_3$	85
58		a. -149.60 b. -156.70	a. -152.47 b. -147.59	a. -109.94 b. -92.30	a. -135.35 b. -146.93	$CDCl_3$	85
59		-142.70	-148.61	-124.50	-142.15	CD_3COCD_3	85
60†		-130.30	-162.71	-129.90	-160.46	$CDCl_3$	85
61		-118.00	-120.32	-96.36	-120.72	$CDCl_3$	85
62		-176.10	-175.22	-130.33	-165.14	$CDCl_3$	85
63†		a. -187.80 b. -130.10	a. -188.18 b. -129.62	a. -127.61 b. -89.46	a. -165.59 b. -124.45	$CDCl_3$	85
64		-134.70	-119.68	-112.85	-130.68	$CDCl_3$	85
65		a. -143.50 b. -189.50	a. -143.95 b. -190.54	a. -112.18 b. -127.50	a. -148.32 b. -170.00	$CDCl_3$	86
66		-62.50	-67.02	-138.34	-68.76	CD_3COCD_3	87

Table 4-F. Five member heterocycles: Continued

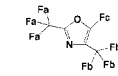
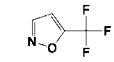
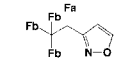
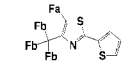
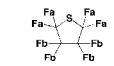
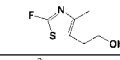
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
67		a. -67.70 b. -64.40 c. -115.40	a. -63.80 b. -61.20 c. -113.19	a. -135.87 b. -135.14 c. -66.01	a. -77.38 b. -70.63 c. -109.87	Not reported	88
68		-64.60	-67.77	-137.55	-72.48	$CDCl_3$	89
69		a. -200.30 b. -80.10	a. -192.44 b. -80.78	a. -163.58 b. -137.30	a. -201.98 b. -73.29	$CDCl_3$	90
70		a. -139.10 b. -62.40	a. -131.85 b. -66.19	a. -74.66 b. -138.13	a. -134.11 b. -68.26	$CDCl_3$	91
71		a. -87.04	a. -84.84	a. -128.34	a. -122.81	$CFCl_3$	92,93
72†		-84.50	-76.07	-69.65	-97.11	$CDCl_3$	94
	R^2		0.96	0.00	0.91		
	RMSE		8.40	49.47	12.86		
	Signed		0.56	-7.58	0.58		
	Unsigned		5.33	44.46	9.49		

Table 4-G. 6-member heterocycles

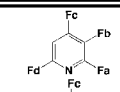
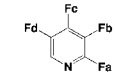
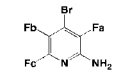
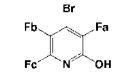
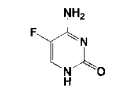
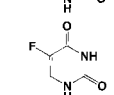
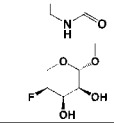
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
73		b. -113.10	b. -165.25	b. -110.96	b. -150.76	Not Reported	95
74†		b. -157.90	b. -155.84	b. -107.93	b. -145.81	Not Reported	95
75		a. -137.80 b. -152.40	a. -140.88 b. -147.14	a. -119.78 b. -112.61	a. -155.50 b. -144.88	Not Reported	96
76		a. -135.70	a. -145.85	a. -118.32	a. -151.12	Not Reported	96
77		a. -51.80	a. -40.69	a. -54.62	a. -46.44	$CFCl_3$	79
78		-171.00	-178.86	-140.03	-173.65	DMSO-D6	97
79†		-170.19	-165.13	-124.39	-171.77	C_6D_6	98,99
80		-203.00	-210.08	-179.10	-195.16	$CDCl_3$	100
	R^2		0.84	0.84	0.86		
	RMSE		18.63	30.48	15.89		
	Signed		6.31	-25.04	4.69		
	Unsigned		11.53	25.67	11.98		

Table 4-H. Benzene Derivatives

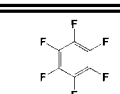
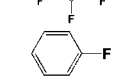
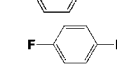
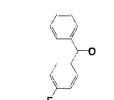
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
81		-164.90	-159.73	-102.81	-150.44	Not Reported	101, 102
82		-113.41	-114.75	-117.01	-123.39	$CDCl_3$	103, 104
83†		-119.67	-121.23	-114.67	-126.39	CCl_4	104
84		-106.46	-108.42	-111.68	-120.47	Not Reported	54
85†		-105.94	-106.58	-110.92	-120.22	Not Reported	54

TABLE 4: Continued[a]

Table 4-II. Benzene Derivatives

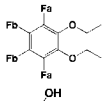
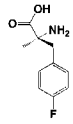
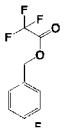
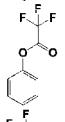
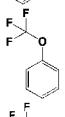
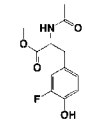
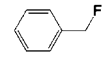
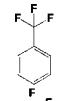
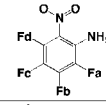
	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
86		a. -160.10	a. -155.97	a. -115.61	a. -153.15	$CFCl_3$	105
87		-115.80	-118.53	-118.83	-125.27	D_2O	106
88		b. -159.30	b. -156.03	b. -110.85	b. -156.20	$CFCl_3$	105
89		-75.02	-76.70	-139.34	-77.35	CH_2Cl_2	58
90†		-73.85	-76.99	-138.77	-77.18	Not reported	58
91		-58.40	-62.93	-139.12	-63.99	Not reported	54
92		-43.50	-47.24	-135.54	-52.60	C_6F_6	107, 108
93		18.20	10.89	-57.39	2.24	Not reported	75

Table 4-II. Benzene Derivatives : Continued

	Structure	Exp.	DFT	MNDO	MNDO-NMR	Medium	Ref
94		-16.80	-12.09	-65.49	-10.96	Not Reported	109
95		-140.30	-150.30	-129.19	-143.77	Not Reported	110
96		-207.00	-192.96	-181.50	-191.20	Neat	111, 104
97		a. -183.35	a. -185.32	a. -147.38	a. -188.35	Not Reported	54
98		-64.00	-67.86	-139.69	-63.95	CH_2Cl_2	37,112
99†		-89.20	-100.43	-152.26	-103.45	Not Reported	78,37
100		a. -160.70 d. -148.00	a. -167.16 d. -136.68	a. -124.93 d. -92.14	a. -167.75 d. -138.45	CD_3COCD_3	113
	R^2		0.99	0.20	0.98		
	RMSE		6.19	51.30	9.64		
	Signed		0.93	11.86	3.08		
	Unsigned		4.99	43.00	8.39		

[a] σ_{ref} values used were optimized to minimize the average signed error for the complete data set of 100 compounds and are given in Table 3 denotes compounds used in test set.

for ^{19}F NMR calculations has yielded calculated values within ~10 ppm of experimental results, which is acceptable accuracy for a theoretical chemical shift range of ~500 ppm (~2%). The MNDO-NMR results are therefore compared to DFT calculations performed at the B3LYP/6−31G++(d,p)//B3LYP/6−31G(d,p) level as the previously mentioned study showed this to have the best agreement with experimental ^{19}F chemical shifts for a range of compounds. Our own tests on the complete data set of 100 compounds supported this choice. The DFT calculations were performed using the Gaussian03 program.[35]

Experimental Data. The data set comprised a training set of 81 compounds and a test set of 19 compounds containing 100 and 23 unique ^{19}F chemical shifts respectively. The structures are shown in Table 4. The test set was chosen to represent each class of chemical environments (as they are grouped in Table 4.) While the goal was to ensure the robustness of the model, it seemed counterproductive to deliberately exclude the data from a particular chemical environment for the sole purpose of observing the predictability of the method. Therefore, all chemical environments in the test set were closely represented in the training set.

The data set is diverse in nature and contains C, H, N, O, F, Cl, Br, and S atoms. The compounds contained a variety of small cyclic (three- to eight-membered rings), fused-cyclic, aliphatic, and aromatic functional groups and a range of compounds from singly fluorinated to perfluorinated. Particular effort was made to include compounds that have characteristics relevant to the study of biological systems (drugs, amino acids and bases). A relevant factor here is that some molecules of interest were large with many degrees of freedom and would require a conformation search to determine the appropriate conformation(s) to be used. In these instances, an alternate molecule was used that contained the particular functional group of interest. Included in the data set are fluorinated derivatives of three amino acids: tyrosine (TYR), phenylalanine (PHE), and tryptophan (TRP). Additionally, steroids, sugars, pyrrole, pyrazole, imidazole, oxazole, and thiazole rings, as well as two DNA bases (cytosine, uracil), were incorporated into the set. Excluded

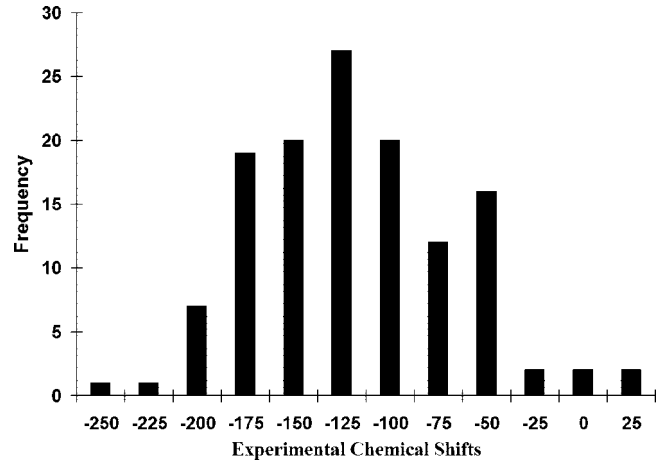


Figure 1. Histogram of the distribution of experimental chemical shifts in the full data set (training set and test set of 123 unique chemical shifts for 100 compounds.)

from the data set were the chemical shifts of fluorine atoms that were involved in bonds other than a single C−F bond. This was done because C−F is the most common bonding arrangement seen in the biological systems in which we are interested. Of the 100 compounds included in the statistics, the experimental chemical shifts ranged from −267.9 ppm to +33.24 ppm, covering a span of roughly 300 ppm. Figure 1 illustrates the distribution of the experimental chemical shifts for the combined training and test set of 100 compounds.

One problem that arose was the discovery of discrepancies among different reports of experimental chemical shifts for the same compounds. In many instances this was found even when the solvent was the same. It is known that subjectivity in the methods of interpreting and validating raw NMR data limits its reproducibility.[36] Fortunately many of the discrepancies were sufficiently small (<1 ppm, well within the error of higher level calculations) to not have a significant effect on our results.

As with the previous MNDO-NMR parametrization, the reference data were preferentially chosen from experiments run under conditions that minimize association and solvent effects. Additionally, in most cases the most recently reported data were used as the reference, as they were calculated using higher resolution instruments. In Tables 4A−H the actual references for the values used are those listed first in column 8; subsequent references listed were in close agreement (off by no more than 2 ppm) and may not have been run using the solvent listed in column 7.

Results and Discussion

The new fluorine parameters [ξ_s, ξ_p, β_s, β_p] are listed in Table 1. The chosen parameters were those that yielded the closest agreement with experiment while straying minimally from the original MNDO parameters. To reiterate, this is in an effort to preserve any physical meaning that the original parameters were designed to reproduce, to the extent that this is possible without detracting from the correlation for the NMR chemical shifts. For consistency, all chemical shifts listed were calculated using a σ_{ref} value that minimizes the average error for this data set. As mentioned previously, this procedure was also implemented in the first MNDO-NMR parametrization by Patchkovskii and Thiel.[17] The σ_{ref} values used in the implementation of eq 1 are listed in Tables 2 and 3. With the assumption that our training set was sufficiently large and diverse in this parametrization, these reported σ_{ref} values should be applicable to future calculations.

The final parameter set resulting from this work includes just 4 optimized parameters to be added to the 16 optimized MNDO-NMR parameters previously published for H, C, N, and O.[21] The ξ_p and β_p were critical terms for obtaining good agreement with the experimental NMR data. The larger values of the Slater orbital exponent ($\xi_{s/p}$) terms indicate a preference for less diffuse orbitals. Because the fluorine parameters are being added to pre-existing parameters, little knowledge can be gained about the shortcomings of the MNDO method solely by examining the changes made for fluorine parameters in this instance. This is because it is difficult to discern whether the changes deemed necessary would be to compensate for a shortcoming in the pre-existing parameters or one in the method itself. Although AM1 geometries were used in the parametrization, it is unlikely that the geometries used in the parametrization were a significant source of error since the new parameters yield very similar results when DFT geometries were used.

As illustrated in Table 3, it is apparent that the choice of geometry does not significantly impact the overall quality of fit between the experimental and calculated chemical shifts when the new parameters are used. This suggests that the new parameters can be used to calculate NMR chemical shifts for structures using either the AM1 or PM3 Hamiltonian, and that the more computationally expensive *ab initio* geometries are not necessary. This is important because our semiempirical QM approach for NMR calculations on large molecules will rely only on semiempirical optimized geometries and not on a more expensive method for geometry optimization.

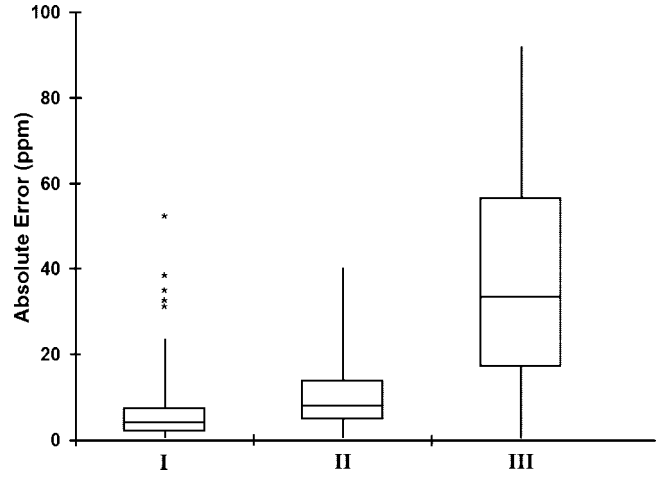


Figure 2. Box and whiskers plot of the average unsigned error for ^{19}F NMR chemical shifts calculated by I. B3LYP//B3LYP, II. MNDO-New ^{19}F Parameters//AM1, and III. MNDO-Standard Fluorine Parameters//AM1. All MNDO calculations were performed using NMR-specific parameters for C, H, N, O given in the fourth column of Table 1. (For interpretation of box plots, see: Bernstein, S. *Schaum's Outline of Elements of Statistics I: Descriptive Statistics and Probability*; McGraw-Hill: New York, 1999.)

Unlike the chemical shift calculations, the absolute shielding tensors do appear to be influenced by the geometry used. This can be inferred by the deviations in the σ_{ref} values used in the different instances. Because of this difference, when comparing results for different molecules it is of course advisable to use geometries that were optimized using the same level of theory.

The ^{19}F chemical shifts of our data set (aimed at capturing many ^{19}F environments that are of biological interest) ranged from −267.9 ppm to +33.24 ppm. Within this range, calculations using the new parameters on the complete data set of 100 compounds yielded an average rms error of 13.85 ppm which is just below 5% of the total range of chemical shifts used in this data set. Furthermore, these values yielded an R^2 value of 0.94. The new parameters performed particularly well for some fluorinated derivatives of DNA bases and amino acids including uracil, cytosine, phenylalanine, glycine, and tyrosine. For some amino acids that are not included in the data set the functional groups of interest are represented and accurately calculated using our new parameters. For example, the parameters do well for prediction of the chemical shifts of fluorine at various positions on indole rings, the functional group present in tryptophan.

The complete list of errors, including average signed and unsigned errors calculated using $CFCl_3$ and the signed error as the σ_{ref} values, is given in Tables 2 and 3. For the training set of 100 ^{19}F chemical shifts, the optimized parameters yielded an rms error of 13.47 ppm and an R^2 value of 0.95. The results were very promising for the test set of 23 ^{19}F chemical shifts, yielding an rms error of 13.61 and an R^2 of 0.90.

Tables 4A−H contain the experimental chemical shifts and those calculated using the new parameters. The tables are grouped by chemical functionality so as to facilitate analysis

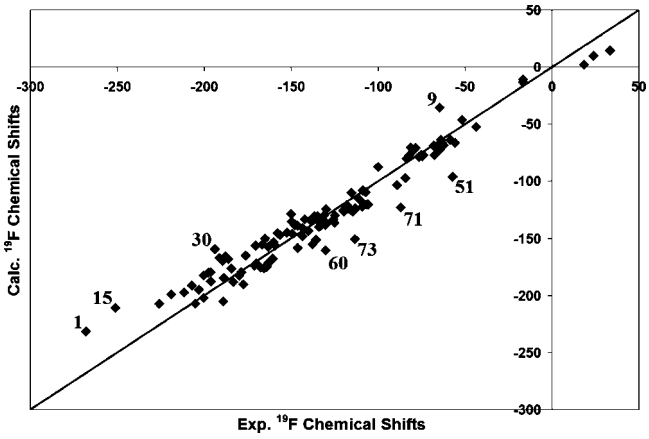


Figure 3. Correlation between experimental and calculated (MNDO-NMR//AM1) ^{19}F chemical shifts using the newly generated fluorine parameters. Compounds with chemical shifts that deviate from experiment by greater than 27.2 ppm are indicated by the compound number used in Table 4.

of our results, and the errors associated with each group of compounds are given at the bottom of each table. This method of examining the errors by groups enables us to highlight some of the areas in which the parameters yield favorable results. In particular, the benzene derivatives and the five-membered heterocycles show good agreement with experiment. This appears to indicate that the chemical shifts of aromatic systems are well described using the new parameters.

For comparison purposes, the chemical shifts calculated at the B3LYP level and at the MNDO level using the MNDO-NMR parameters with the original or "old" fluorine parameters are also shown in Table 4. The compounds included in the test set are denoted with the dagger symbol in the Table 4 series; all other compounds were included in the training set. Since the parameters were optimized using AM1 geometries, the NMR data listed in Table 4 were calculated using AM1 geometries for the semiempirical NMR calculations and B3LYP geometries for the DFT NMR calculations.

When compared with the original MNDO-NMR parameters, the final parameters show significant improvement in agreement with experimental results for our data set. Figure 2 illustrates via box plot the magnitude of the error that resulted from calculations using our new parameters versus those using the original fluorine parameters and the DFT calculations. The extreme outliers (marked by asterisks) were greater than 3 × IQR (interquartile range) above the value of the third quartile. These outliers are discussed below, along with possible areas in which the new method may be limited. The results are significantly better than those using the standard fluorine parameters and compare very well to DFT results calculated at the B3LYP-GIAO/6−31++G(d,p) level for our data set. As shown in Figure 2, 75% of the calculations using our new parameters yielded an absolute error less than 13.8 ppm, as compared to the standard fluorine parameters (56.4 ppm), and DFT (7.4 ppm). To reiterate, this basis set has been shown to provide the closest agreement with experiments for ^{19}F NMR chemical shift calculations using the B3LYP-GIAO method.[34]

The correlation between the experimental and theoretical chemical shifts using the new fluorine parameters is shown in Figure 3. Mild outliers are highlighted, chemical environments in which the deviations between the calculated and experimental chemical shifts exceeded 27.2 ppm (1.5 × IQR above the third quartile). There were three main conditions under which use of the new parameters resulted in chemical shifts that were not in close agreement with the experimental results. These are summarized in the paragraphs below.

The first difficult chemical environment involved systems in which the fluorine was bonded to an α carbon adjacent to a carbon involved in multiple bonds with a heteroatom. The α carbon problem is also seen in the B3LYP calculations accounting for all of the outliers (compounds **30**, **31**, **32**, **60**, and **73** in Table 4.) This may be due to an effect of the solvent forming hydrogen bonds with the heteroatom as well as the fluorine atom and deshielding the ^{19}F nucleus to a greater extent than can be accounted for by the *in vacuo* calculations.

Second, rather large errors were noted in CH_3F and CF_4, where the fluorine atoms were too deshielded by ~30 ppm. In these very small systems it is possible that the contributions of the core electrons are more significant than in the larger systems. Systematic problems of the core electrons were addressed by using a chosen σ_{ref} value. However, a different σ_{ref} value may be more appropriate for these particular systems, as the core electrons may be playing a more important role. In the previous MNDO-NMR parametrization it was found that different σ_{ref} values were required to capture the relative chemical shifts for ^{1}H under different circumstances as well (in this instance hydrogen atoms involved in N−H, C−H, and O−H bonds used unique σ_{ref} values) Since this problem occurred in so few cases in our data set, there were not sufficient data points to determine a more appropriate σ_{ref} value to use in these cases. These errors may stem from insufficient flexibility in the single-ζ basis sets being used in the MNDO approximation. The source of this error may have also been the source of larger errors found in some small ring systems crowded with polar atoms. Polarization functions may be necessary in order to simultaneously describe the shielding felt by the fluorine atom in these and other systems.

Third, important large errors were also found for several compounds with chlorine or sulfur atoms. In this case the errors may be a result of the standard MNDO parameters for these atoms, possibly making the electron density of chlorine and sulfur too diffuse, and consequently excessively shielding the fluorine atom. Because of the large deviation seen in the calculation of $CFCl_3$ (too shielded by ~90 ppm) it was not appropriate to use this as the reference value.

Conclusions

The addition of NMR-specific MNDO parameters for fluorine has now extended the semiempirical QM NMR-based methodology for qualitative (to near quantitative) description of chemical shifts. Our approach can be employed using semiempirical (AM1/PM3) geometries with good accuracy and can be executed at a fraction of the cost of *ab initio* methods, thereby assisting in the computational studies of ^{19}F NMR for much larger systems including DNA and protein−ligand complexes. The results are comparable to calculations performed at the B3LYP-GIAO level for our data set. There is a marked improvement over the standard MNDO parameters in the agreement between experimental and calculated chemical shifts for our data set. This improvement in the correlation is clearly indicated by the increased R^2 value from 0.38 to 0.94 using the standard and new MNDO fluorine parameters, respectively. Our parameters work well for a variety of functional groups frequently found in biological molecules where ^{19}F is present. Furthermore, the results seen in the test set are promising for the extendability of the method. Although the parameters were not generated by a method that allows us to state conclusively that they correspond to a global minimum, the genetic algorithm has allowed us to search a broad range of parameter space and

there has been no evidence suggesting that a significantly lower minimum exists.

Halogens present a peculiar problem for the semiempirical methods in the Neglect of Diatomic Differential Overlap (NDDO) family of approximations in that certain calculations may require polarization or multizeta representations that the minimal basis sets do not offer. As discussed previously, this inflexibility appears to prevent us from capturing the relative chemical shifts for compounds including $CFCl_3$. For this reason we have found that the most effective method of obtaining qualitative accuracy for the ^{19}F NMR chemical shifts is through the use of an optimized reference value and not through the use of the fluorine chemical shift of $CFCl_3$ as the reference. When $CFCl_3$ is used as the reference value, large signed errors are found in spite of the high correlations, as illustrated by the R^2 values.

The limitations of these parameters that we have found have been outlined in Results and Discussion. These limitations include, but are not limited to, the fact that the parameters were generated for fluorine atoms involved in C−F bonds only. Nonetheless, the present parameters allow one to study the effect of the environment on ^{19}F chemical shifts that are seen in many biological systems. Application of the new ^{19}F parameters to biological problems will be illustrated in future publications.

Acknowledgment. We thank the NSF-SEAGEP and NSF (MCB-0211639) for financial support. We also thank Valerie Williams and Mike Weaver for their help in proofreading this document.

References and Notes

(1) Gerig, J. T. *Biophysical Society Biophysics Textbook*; 2001.
(2) Lau, E. Y.; Gerig, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4408–4417.
(3) Lepre, C. A.; Moore, J. M.; Peng, J. W. *Chem. Rev.* **2004**, *104*, 3641–3675.
(4) Opella, S. J.; Marassi, F. M. *Chem. Rev.* **2004**, *104*, 3587–3606.
(5) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, *104*, 3519–3540.
(6) Leone, M. R.; Rodriguez-Mias, R. A.; Pellecchia, M. *ChemBioChem* **2003**, *4*, 649–650.
(7) O'Hagan, D.; Harper, D. B. *J. Fluorine Chem.* **1999**, *100*, 127–133.
(8) Gerig, J. T. *Biophysical Society Biophysics Textbook*; 2001.
(9) Isanbor, C.; O'Hagan, D. *J. Fluorine Chem.* **2006**, *127*, 303–319.
(10) Dalvit, C.; Ardini, E.; Flocco, M.; Fogliatto, G. P.; Mongelli, N.; Veronesi, M. *J. Am. Chem. Soc.* **2003**, *125*, 14620–14625.
(11) Dalvit, C.; Ardini, E.; Fogliatto, G. P.; Mongelli, N.; Veronesi, M. *Drug Discovery Today* **2004**, *9*, 595–602.
(12) Shikii, K.; Sakurai, S.; Utsumi, H.; Seki, H.; Tashiro, M. *Anal. Sci.* **2004**, *20*, 1475–1477.
(13) Haas, A.; Reinke, H. *Angew. Chem., Int. Ed. Engl.* **1967**, *6*, 705–706.
(14) Gregory, D. H.; Gerig, J. T. *Biopolymers* **1991**, *31*, 845–858.
(15) Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. *J. Am. Chem. Soc.* **1993**, *115*, 6851–6862.
(16) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.
(17) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
(18) Wu, W.; You, X.; Dai, A. *Sci. Sin., Ser. B (Engl. Ed.)* **1988**, *31*, 1048–1061.
(19) Wang, B.; Merz, K.M., Jr *J. Am. Chem. Soc.* **2005**, *127*, 5310–5311.
(20) Wang, B.; Brothers, E. N.; van der Vaart, A.; Merz, K. M. *J. Chem. Phys.* **2004**, *120*, 11392–11400.
(21) Patchkovskii, S.; Thiel, W. *J. Comput. Chem.* **1999**, *20*, 1220–1245.
(22) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
(23) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
(24) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
(25) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(26) Lee, C.; Yang, W.; Parr, R. G. *Phys. Chem. Rev. B* **1987**, *37*, 785–789.
(27) Ditchfield, R. *J. Chem. Phys.* **1972**, *56*, 5688–5691.
(28) Brothers, E. N.; Merz, K. M. *J. Phys. Chem. B* **2002**, *106*, 2779–2785.
(29) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *233*, 231–236.
(30) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998**, *102*, 8080–8090.
(31) Wang, B.; Raha, K.; Liao, N.; Peters, M. B.; Kim, H.; Westerhoff, L. M.; Wollacott, A. M.; van der Vaart, A.; Gogonea, V.; Suarez, D.; Dixon, S. L.; Vincent, J. J.; Brothers, E. N.; Merz, K. M., Jr. Div Con.
(32) Tanuma, T.; Irisawa, J. *J. Fluorine Chem.* **1999**, *99*, 157–160.
(33) Ying, Z.; Wu, A.; Xu, X.; Yan, Y. *J. Phys. Chem. A* **2007**, *111*, 9431–9437.
(34) Fukaya, H.; Ono, T. *J. Comput. Chem.* **2003**, *25*, 51–60.
(35) 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, revision C.02; Gaussian, Inc., Wallingford, CT, 2004.
(36) Baran, M. C.; Huang, Y. J.; Mosely, H. N. B.; Montelione, G. T. *Chem. Rev.* **2004**, *104*, 3541–3555.
(37) Weigert, F. J. *J. Org. Chem.* **1980**, *45*, 3476–3483.
(38) Singer, R. J.; Eisenhut, M.; Schmutzler, R. *J. Fluorine Chem.* **1971**, *1*, 193–202.
(39) Dmowski, W.; Kaminski, M. *J. Fluorine Chem.* **1983**, *23*, 219–228.
(40) Filipovich, G.; Tiers, G. V. D. *J. Phys. Chem.* **1959**, *63*, 761–763.
(41) Vcelak, J.; Chvalovsky, V.; Voronkov, M. G.; Pukhnarevich, V. B.; Pestunovich, V. A. *Collect. Czech. Chem. Commun.* **1976**, *41*, 386–390.
(42) Weigert, F. J. *J. Fluorine Chem.* **1990**, *46*, 375–384.
(43) Harris, R. K. *J. Mol. Spectrosc.* **1963**, *10*, 309–319.
(44) Sartori, P.; Habel, W. *J. Fluorine Chem.* **1980**, *16*, 265–276.
(45) Schmutzler, R. *J. Chem. Soc* **1964**, (Nov), 4551–4557.
(46) Naumann, D.; Kischkewitz, J. *J. Fluorine Chem.* **1990**, *47*, 283–299.
(47) Christe, K. O.; Wilson, W. W. *J. Fluorine Chem.* **1990**, *47*, 117–120.
(48) Solovev, D. V.; Rodin, A. A.; Zenkevich, I. G.; Lavrentev, A. N. *Zhurn. Obshch. Khim.* **1988**, *58*, 1544–1550 (in Russian).
(49) Bloshchitsa, F. A.; Burmakov, A. I.; Kunshenko, B. V.; Alekseeva, L. A.; Yagupol'skii, L. M. *Zh. Obshch. Khim.* **1985**, *21*, 1414–1420 (in Russian).
(50) Aktaev, N. P.; Il'in, G. F.; Sokol'skii, G. A.; Knunyants, I. L. *Izv. Akad. Nauk SSSR Ser. Khim.* **1977**, *5*, 1112–1117.
(51) Schreider, V. A.; Rozhkov, I. N. *Izv. Akad. Nauk SSSR, Ser. Khim.* **1979**, *3*, 673–675.
(52) Burdon, J.; Huckerby, T. N.; Stephens, R. *J. Fluorine Chem.* **1977**, *10*, 523–540.
(53) Jullien, J.; Martin, J. A.; Ramanadin, R. *Bull. Soc. Chim. Fr.* **1964**, *171*, 172.
(54) Dungan, C. H.; Van Wazer, J. R. *Compilation of Reported ^{19}F NMR Chemical Shifts (1951−1967)*; John Wiley & Sons, Inc.: New York, 1970.
(55) Hemer, I.; Havlicek, J.; Dedek, V. *J. Fluorine Chem.* **1986**, *34*, 241–250.
(56) Mironova, A. A.; Maletina, I. I.; Iksanova, S. V.; Orda, V. V.; Yagupolsky, L. M. *Zh. Org. Khim.* **1989**, *25*, 306–311.
(57) Bogachev, Y. S.; Serebryanskaya, A. I.; Khutsishvili, V. G.; Korenkova, V. M.; Shapet'ko, N. N. *Zh. Obsch. Khim.* **1986**, *56*, 909–915.
(58) Manatt, S. L. *J. Am. Chem. Soc.* **1966**, *88*, 1323–1324.
(59) Pellerite, M. J. *J. Fluorine Chem.* **1990**, *49*, 43–66.
(60) Burdon, J.; Knights, J. R.; Parsons, I. W.; Tatlow, J. C. *J. Chem. Soc., Perkin Trans. 1* **1976**, *18*, 1930–1933.
(61) Pitcher, E.; Buckingham, A. D.; Stone, F. G. A. *J. Chem. Phys.* **1961**, *36*, 124–129.
(62) Burger, H.; Niepel, H.; Pawelke, G.; Frohn, H. J.; Satori, P. *J. Fluorine Chem.* **1980**, *15*, 231–237.
(63) Lustig, M.; Ruff, K. J. *Inorg. Chem.* **1965**, *4*, 1441–1443.

(64) Bovey, F. A.; Anderson, E. W.; Hood, F. P.; Kornegay, R. L. *J. Chem. Phys.* **1963**, *40*, 3099–3109.
(65) Schneider, H. J.; Gschwendtner, W.; Heiske, D.; Hoppen, V.; Thomas, F. *Tetrahedron* **1977**, *33*, 1769–1773.
(66) Gash, V. W.; Bauer, D. J. *J. Org. Chem.* **1966**, *31*, 3602–3607.
(67) Feeney, J.; Sutcliffe, L. H.; Walker, S. M. *Mol. Phys.* **1966**, *11*, 117–128.
(68) Mitsch, R. A. *J. Am. Chem. Soc.* **1965**, *87*, 758–761.
(69) Cullen, W. R.; Waldman, M. C. *J. Fluorine Chem.* **1971**, *1*, 151–163.
(70) Wheaton, G. A.; Burton, D. J. *J. Fluorine Chem.* **1977**, *1*, 25–44.
(71) Jullien, J.; Stahl-Lariviere, H. *Bull. Soc. Chim. Fr.* **1967**, *1*, 99–104.
(72) Cantacuzene, J.; Ricard, D. *Bull. Soc. Chim. Fr.* **1967**, *5*, 1587–1593.
(73) Boswell, G. A. J. *J. Org. Chem.* **1966**, *31*, 991–1000.
(74) Koroniak, H.; Palmer, K. W.; Dolbier, W. R., Jr.; Zhang, H. Q. *Magn. Reson. Chem.* **1993**, *31*, 748–751.
(75) Krause, L. J.; Morrison, J. A. *J. Am. Chem. Soc.* **1981**, *103*, 2995–3001.
(76) Chambers, R. D.; Edwards, A. R. *J. Chem. Soc., Perkin Trans. 1* **1997**, 3623–3628.
(77) Campbell, S. F.; Hudson, A. G.; Mooney, E. F.; Pedler, A. E.; Stevens, R.; Wood, K. N. *Spectrochim. Acta, Part A* **1967**, *23*, 2119–2125.
(78) Olah, G. A.; Chambers, R. D.; Comisarow, M. B. *J. Am. Chem. Soc.* **1967**, *89*, 1268–1269.
(79) Mitsch, R. A. *J. Am. Chem. Soc.* **1965**, *87*, 328–333.
(80) Merritt, R. F.; Johnson, F. A. *J. Org. Chem.* **1966**, *31*, 1859–1863.
(81) Bystrov, V. F.; Utyanskaya, E. Z.; Yagupol'skii, L. M. *Opt. Spektrosk.* **1961**, *10*, 138–141.
(82) Petrova, T. D.; Savchenko, T. I.; Kukovinets, O. S.; Yakobson, G. G. *Izv. Sib. Otdel. Akad. Nauk SSSR Ser. Khim* **1974**, *2*, 117–123.
(83) Petrova, T. D.; Savchenko, T. I.; Kukovinets, O. S.; Yakobson, G. G. *Izv. Sib. Otdel. Akad. Nauk SSSR Ser. Khim* **1973**, *2*, 104–108.
(84) Christe, K. O.; Pavlath, A. E. *J. Org. Chem.* **1965**, *30*, 4104–4107.
(85) Dvornikova, E.; Bechcicka, M.; Kamienska-Trela, K.; Krowczynski, A. *J. Fluorine Chem.* **2003**, *124*, 159–168.
(86) Fabra, F.; Fos, E.; Vilarrasa, J. *Tetrahedron Lett.* **1979**, *34*, 3179–3180.
(87) Owen, D.; Plevey, R. G.; Tatlow, J. C. *J. Fluorine Chem.* **1981**, *17*, 179–186.
(88) Koshelev, V. M.; Barsukov, I. N.; Vasilev, N. V.; Gontar, A. F. *Chem. Heterocycl. Compd. (N.Y., NY, U.S.)* **1989**, *12*, 1699–1700.
(89) Gerus, I. I.; Gorbunova, M. G.; Vdovenko, S. I.; Yagupol'sky, Y. L.; Kukhar, V. P. *Zh. Org. Khim.* **1990**, *26*, 1877–1883.
(90) Vasil'ev, N. V.; Savostin, V. S.; Kolomiets, A. F.; Sokolsky, G. A. *Khim. Geterotsikl. Soedin.* **1989**, *5*, 663–667.
(91) Burger, K.; Geith, K.; Norbert, S. *J. Fluorine Chem.* **1990**, *46*, 105–122.
(92) Tiers, G. V. D. *J. Phys. Chem.* **1962**, *66*, 764–765.
(93) Abe, T.; Shreeve, J. M. *J. Fluorine Chem.* **1973**, *3*, 17–26.
(94) Lowe, G.; Potter, B. V. L. *J. Chem. Soc., Perkin Trans. 1* **1980**, 2026–2028.
(95) Chambers, R. D.; Drakesmith, F. G.; Musgrave, W. K. R. *J. Chem. Soc.* **1965**, 5045–5048.
(96) Chambers, R. D.; Hutchinson, J.; Musgrave, W. K. R. *J. Chem. Soc.* **1965**, 5040–5045.
(97) Robins, M. J.; Maccoss, M.; Naik, S. R.; Ramani, G. *J. Am. Chem. Soc.* **1976**, *98*, 7381–7389.
(98) Ellermann, J.; Schamberger, J.; Knock, F. A.; Moll, M.; Bauer, W. *Monatsh. Chem.* **1997**, *128*, 399–410.
(99) Robins, M. J.; MacCoss, M.; Naik, S. R.; Ramani, G. *J. Am. Chem. Soc.* **1976**, *98*, 7381–7389.
(100) Nakai, K.; Takagi, Y.; Tsuchiya, T. *Carbohydr. Res.* **1999**, *316*, 47–57.
(101) Dean, P. A.W.; Ibbott, D. G. *Can. J. Chem.* **1976**, *54*, 177–187.
(102) Sheppard, W. A.; Foster, S. S. *J. Fluorine Chem.* **1972**, *2*, 53–62.
(103) Kitching, W.; Adcock, W.; Aldous, G. *J. Org. Chem.* **1979**, *44*, 2652–2658.
(104) Zweig, A.; Fischer, R. G.; Lancaster, J. E. *J. Org. Chem.* **1980**, *45*, 3597–3603.
(105) Cavalli, L. *J. Chem. Soc. B* **1967**, 384–387.
(106) Soloshonok, V. A.; Belokon, Y. N.; Kukhar, V. P.; Chernoglazova, N. I.; Saporovskaya, M. B.; Bakhmutov, V. I.; Kolycheva, M. T.; Belikov, V. M. *Izv. Akad. Nauk SSSR Ser. Khim.* **1990**, *7*, 1630–1636.
(107) Haas, A.; Hellwig, V. *J. Fluorine Chem.* **1975**, *6*, 521–532.
(108) Clark, J. H.; Jones, C. W.; Kybett, A. P.; McClinton, M. A.; Miller, J. M.; Bishop, D.; Blade, R. J. *J. Fluorine Chem.* **1990**, *48*, 249–253.
(109) Christe, K. O.; Pavlath, A. E. *J. Org. Chem.* **1965**, *30*, 3170–3173.
(110) Hebel, D.; Kirk, K. L. *J. Fluorine Chem.* **1990**, *47*, 179–183.
(111) Muller, N.; Carr, D. T. *J. Phys. Chem.* **1963**, *67*, 112–115.
(112) Kobayashi, Y.; Kumadaki, I. *J. Chem. Soc., Perkin Trans. 1* **1980**, *3*.
(113) Homer, J.; Thomas, L. F. *J. Chem. Soc.* **1966**, 141–144.

JP801649F