JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Benchmarking Hydrogen and Carbon NMR Chemical Shifts at HF, DFT, and MP2 Levels

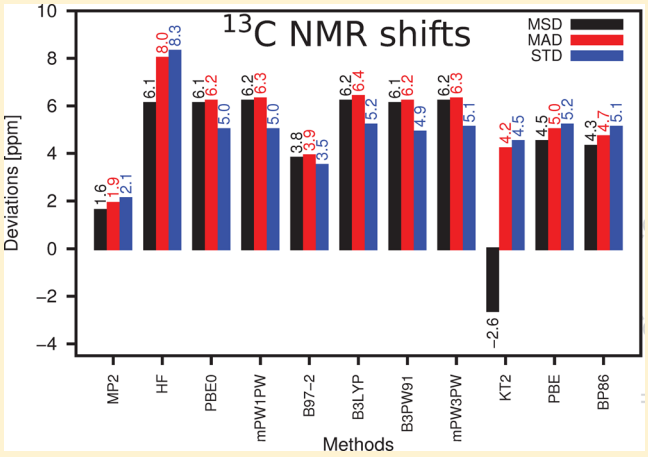
Denis Flaig, Marina Maurer, Matti Hanni, Katharina Braunger, Leonhard Kick, Matthias Thubauville, and Christian Ochsenfeld*

Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 Munich, Germany

Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstr. 5-13, D-81377 Munich, Germany

Supporting Information

ABSTRACT: An extensive study of error distributions for calculating hydrogen and carbon NMR chemical shifts at Hartree−Fock (HF), density functional theory (DFT), and Møller−Plesset second-order perturbation theory (MP2) levels is presented. Our investigation employs accurate CCSD(T)/cc-pVQZ calculations for providing reference data for 48 hydrogen and 40 carbon nuclei within an extended set of chemical compounds covering a broad range of the NMR scale with high relevance to chemical applications, especially in organic chemistry. Besides the approximations of HF, a variety of DFT functionals, and conventional MP2, we also present results with respect to a spin component-scaled MP2 (GIAO-SCS-MP2) approach. For each method, the accuracy is analyzed in detail for various basis sets, allowing identification of efficient combinations of method and basis set approximations.



1. INTRODUCTION

Over the past decades, much progress has been made in the ab initio calculation of nuclear-magnetic resonance (NMR) data of molecular systems (for an overview see, e.g., refs 1−5). Therefore, quantum-chemical calculations have become highly important tools in the often difficult assignment of experimental NMR spectra and can be performed today at various levels of theory ranging from Hartree−Fock (HF)[6−8] or density-functional theory (DFT)[9] to wave function-based correlation methods like, e.g., Møller−Plesset second-order perturbation theory (MP2)[10−12] and coupled-cluster (CC) methods such as CC singles doubles (CCSD)[13] or CCSD including perturbative triples (CCSD(T)).[14] While in particular the high-accuracy schemes such as CCSD or CCSD(T) are confined to rather small molecules, also the more approximate quantum-chemical schemes are severely hampered in their applicability to larger molecules by the steep scaling of the computational effort with molecular size. To overcome these limitations, linear-scaling methods for calculating NMR shieldings have been devised at the HF and DFT levels (instead of their conventional cubic scaling), opening the way to calculate molecules with more than 1000 atoms on workstation computers.[15−20] In addition, often just a few NMR shieldings are of interest (e.g., for a molecule within a solvent environment, where the shieldings of each solvent molecule are irrelevant), so that most recently nuclei-selected NMR methods have been introduced that allow exploitation of the locality of the perturbation and reduction of the computational effort even beyond linear to sublinear, $O(M^0)$, i.e., asymptotically independent of molecular size.[21] Also, for the simplest wave function-based correlation theory, MP2, such linear- or sublinear-scaling methods (instead of the conventional $O(M^5)$ scaling) have recently been formulated.[22] While the latter schemes are still in a pilot-implementation stage, much of the method progress made for the calculation of MP2 energies for large molecules[23] is expected to be transferable.

Besides progress in developing fast methods, it is crucial for an efficient calculation of NMR shieldings to establish by extensive benchmarks the reliability of both the various approximation methods for solving the Schrödinger equation and the commonly used incomplete basis sets. Here, extensive and reliable information on the error distribution of the existing methods is required, which is the goal of our present work. In our present work, we benchmark NMR data, where we build upon earlier studies of the accuracy of NMR shieldings calculations[24−40] and aim to go beyond the former studies in the following aspects: (1) We enlarge an earlier test set used by

Received: September 2, 2013
Published: January 23, 2014

   dx.doi.org/10.1021/ct400780f | J. Chem. Theory Comput. 2014, 10, 572−578

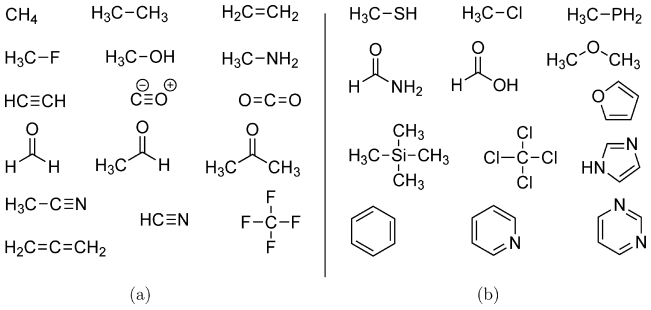


Figure 1. Molecular benchmark set of the current study: (a) Original molecular structures of ref 30. (b) Additional molecular structures of the present work.

Gauss and co-workers[24,30] (see Figure 1a) by adding new molecules (see Figure 1b) with structures optimized at the CCSD(T)/cc-pVTZ level and with high relevance to organic chemistry (or biochemistry), covering a broad range of the NMR scale for proton and carbon shifts. In detail, the compounds benzene, furan, imidazole, pyridine, and pyrimidine were added to the original set representing aromatic systems (for example, their derivatives are typically comprised of proteins or nucleic acids etc.). Dimethyl ether, formic acid, and formamide serve as representatives for ethers, carboxylic acids, and carboxamides, respectively. As a further important system, the standard reference molecule tetramethylsilane (TMS) was included in the selection, allowing for calculations of standard relative shieldings for hydrogen and carbon nuclei. Additionally, the compounds CH_3PH_2, CH_3SH, CCl_4, and CH_3Cl are employed for considering further hetero atoms. (2) Besides carbon (or hetero) shieldings, we also consider hydrogen shieldings. (3) Our error analysis is based on rather accurate reference calculations at the CCSD(T)/cc-pVQZ level. (4) In addition, MP2 as well as spin-component-scaled-(SCS)-MP2 results are provided. MP2 theory represents the most cost-effective wave-function-based correlation theory and is known to provide often reliable NMR chemical shifts. (5) Results for a broad selection of different DFT functionals are presented. (6) The accuracy is analyzed for a variety of basis sets.

While for the original set of structures introduced by Gauss and co-workers also highly accurate gas phase NMR experimental measurements are available,[41] our present study aims to extend the test set also to molecules for which no gas phase experiments are available. Therefore, we restrain ourselves to compare the accuracies to the most reliable theoretical NMR data available as computed at the CCSD(T)/cc-pVQZ level. In this way, influences of the structure, vibrations, etc. are eliminated, and we get the information on how to obtain the most reliable theoretical data in a most cost-efficient way (method/basis).

After describing some methodological aspects, we focus first on the accuracy of various quantum-chemical methods by largely eliminating basis set influences in using a large basis. Then cost-efficient pathways are discussed by studying the accuracy of a variety of smaller and medium-sized basis sets.

2. METHODOLOGICAL ASPECTS

The global minimum structures for all molecules of the benchmark set (see Figure 1) were obtained at the CCSD(T)/cc-pVTZ[42] level by the program package CFOUR[43] and are available via the Web site http://www.cup.uni-muenchen.de/pc/ochsenfeld/download.html.

The GIAO-(SCS)-MP2 NMR calculations were performed with a development version of the program package Q-Chem[44] based on an AO-MP2 NMR implementation.[22] The DFT calculations were performed with the newly developed QM package FermiONs++,[45] which includes DFT functionals from the Libxc database.[46] The selection of DFT methods comprises hybrid GGA-functionals PBE0,[47] mPW1PW,[48] B97-2,[49] B3LYP,[50,51] B3PW91,[52] mPW3PW,[48] and pure GGA functionals BP86,[53] PBE,[54] and KT2.[55] The reference calculations at the CCSD(T) level were performed with the program package CFOUR.[43]

The employed basis sets (STO-3G,[56,57] 3-21G,[58,59] 6-31G**,[60,61] 6-311G**,[62] pcS-0, pcS-1, pcS-2,[32] def2-SVP, def2-TZVP,[63] tz2p, qz2p,[30,64] cc-pVDZ, cc-pVTZ, cc-pVQZ[42]) were partly transferred from the basis set exchange database.[65,66]

3. ACCURACIES OF VARIOUS QUANTUM-CHEMICAL APPROXIMATIONS

In the following, we first aim at eliminating basis set influences by using a large basis set (cc-pVQZ[42]) and focus on the expected errors of HF, different DFT functionals, and MP2 as compared to the most reliable data obtained at the CCSD(T) level. By comparing to the rather accurate theoretical results at the CCSD(T) level, we focus on the accuracies for the same structure and avoid influences of conformational changes (in particular molecular vibrations) or the chemical environment (e.g., a host system or solvent etc.). For a comparison to experimental data, see refs 24, 30, and 41.

Figures 2 and 3 summarize the effective errors for all tested methods for hydrogen and carbon nuclei, respectively. The

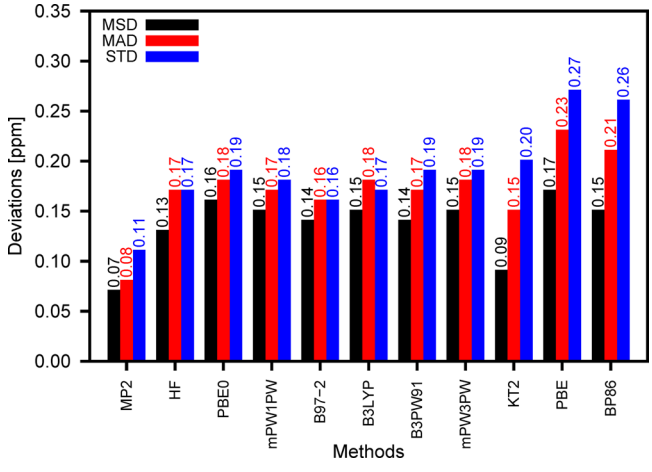


Figure 2. MSD and MAD with respect to TMS and the STD (in ppm) for ^{1}H NMR shifts at MP2, DFT, and HF levels with respect to CCSD(T) results (basis set cc-pVQZ and CCSD(T)/cc-pVTZ geometries; the GIAO approach is always employed).

explicit values of the underlying NMR shifts for all considered nuclei and all levels of theory are provided in the Supporting Information. For evaluating the accuracy of methods in calculating NMR shifts (relative shieldings of nuclei A with respect to the reference nucleus, e.g., in TMS: $\delta_A = \sigma_{TMS} - \sigma_A$), three distinctive error criteria are listed: the mean signed deviation (MSD), the mean absolute deviation (MAD), and the standard deviation (STD), which are computed as follows over N nuclei:

dx.doi.org/10.1021/ct400780f | J. Chem. Theory Comput. 2014, 10, 572−578

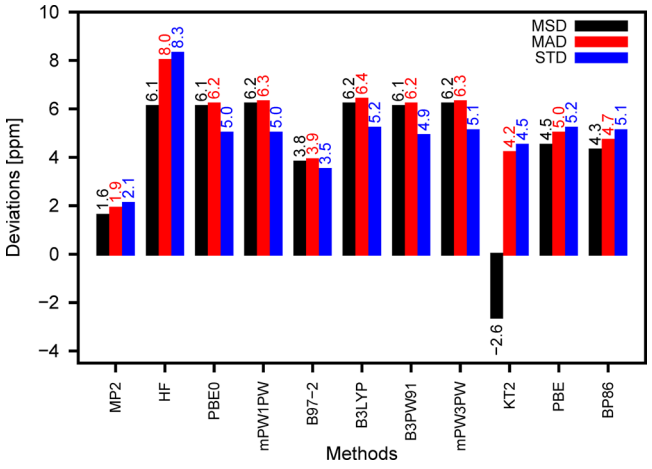


Figure 3. MSD and MAD with respect to TMS and the STD (in ppm) for ^{13}C NMR shifts at MP2, DFT, and HF levels with respect to CCSD(T) results (basis set cc-pVQZ and CCSD(T)/cc-pVTZ geometries; the GIAO approach is always employed).

$$\mathrm{MSD} = \frac{1}{N}\sum_{A}^{N}\overbrace{\delta_A - \delta_A^{\mathrm{CCSD(T)/cc-pVQZ}}}^{\Delta\delta_A}$$

$$\mathrm{MAD} = \frac{1}{N}\sum_{A}^{N}|\Delta\delta_A|$$

$$\mathrm{STD} = \sqrt{\frac{1}{N-1}\sum_{A}^{N}(\mathrm{MSD} - \Delta\delta_A)^2},$$

Here, the standard deviation STD is invariant with respect to the analogous formulation with absolute shielding values σ instead of relative shifts δ (the MSD then also needs to be calculated over σ). Thus, the STD is independent of the selection of the reference nucleus, in contrast to the MSD or MAD, and therefore represents the most meaningful general criterion for judging the accuracy of NMR shifts. Because of the common use of TMS as the reference compound, the MSD and MAD for shifts with respect to TMS are also provided.

The data plotted in Figures 2 and 3 indicate that the MP2 method achieves the lowest standard deviations of 0.11/2.1 ppm for ^{1}H/^{13}C nuclei with respect to the CCSD(T) reference. In contrast, HF and DFT results range between 0.16/3.5 and 0.27/8.3 ppm for ^{1}H/^{13}C. The order according to the accuracies differs to some extent for ^{1}H and ^{13}C nuclei: For ^{1}H shifts, HF proves to be on the same level of accuracy as compared to DFT, while for ^{13}C shifts all DFT functionals are better by more than 3.1 ppm. For both ^{1}H and ^{13}C shifts, the functional B97-2[49] performs best, followed by the functional KT2[55] in the case of ^{13}C shifts.

The good performance of B97-2 and KT2 is also reflected by the MAD values of shifts with respect to TMS. Although in general the MSD could be regarded to be a less meaningful error criterion for calculating relative shieldings (shifts), it shows an interesting behavior: The KT2 functional is the only method that systematically underestimates the carbon shifts calculated with respect to TMS (negative MSD), whereas all the other methods overestimate the standard shift (positive MSDs).

The Supporting Information gives further insights into the error distribution and the relationship to chemical structure. Generally, one can conclude that chemically similar structures (small absolute values of shifts) are better described than more different structures. Thus, if TMS is chosen as the reference compound, the problematic cases accumulate mainly in the regime of aromatic and carbonylic groups. Moreover, calculating carbon shifts for the compounds CO, CCl_4, and the allene CH_2CCH_2 is shown to be especially demanding for the DFT methods. An inverse effect is observed, if for instance CO is chosen as the reference compound: now the sp^3 hybridized structures are the most deviating cases (see Section 15 of the Supporting Information and section 6 for a more detailed discussion).

4. BASIS SET INFLUENCES

While the use of the rather large basis set cc-pVQZ[42] largely eliminates basis set effects for the data presented above, we focus in the following on the influence of basis set deficiencies. Clearly, the size of the basis set plays a central role for the computing time and may effectively limit the size of treatable systems. Therefore, the STD is listed in Tables 1 and 2 for hydrogen and carbon shifts as computed using a variety of basis sets for all different methods employed above.

Similar to the discussion above, the STD for all methods and basis sets is referenced to CCSD(T)/cc-pVQZ data. Thus, the

Table 1. Standard Deviation (STD) of H Shifts for Different Basis Sets and Methods with Respect to the Reference Calculation CCSD(T)/cc-pVQZ

basis set	$\bar{N}_{\mathrm{bas}}$[a]	CCSD(T)	MP2	HF	PBE0	mPW1PW	B97-2	B3LYP	B3PW91	mPW3PW	KT2	PBE	BP86
STO-3G	3.0	0.77	0.79	1.02	0.86	0.86	0.83	0.85	0.85	0.85	0.77	0.82	0.81
3-21G	5.4	0.56	0.50	0.65	0.62	0.61	0.60	0.62	0.61	0.61	0.62	0.66	0.64
6-31G**	9.8	0.19	0.17	0.22	0.23	0.22	0.21	0.23	0.23	0.23	0.26	0.31	0.30
6-311G**	11.9	0.18	0.17	0.21	0.23	0.22	0.21	0.23	0.23	0.23	0.25	0.31	0.30
pcS-0	5.4	0.47	0.41	0.66	0.54	0.53	0.51	0.52	0.52	0.52	0.46	0.54	0.52
pcS-1	10.8	0.16	0.14	0.21	0.17	0.17	0.16	0.18	0.18	0.18	0.21	0.25	0.24
pcS-2	23.2	0.03	0.10	0.19	0.17	0.16	0.15	0.15	0.17	0.17	0.18	0.25	0.24
def2-SVP	9.4	0.23	0.21	0.22	0.21	0.21	0.20	0.22	0.21	0.22	0.25	0.29	0.28
def2-TZVP	17.9	0.11	0.13	0.18	0.23	0.22	0.20	0.23	0.23	0.23	0.26	0.32	0.31
tz2p	16.3	0.12	0.12	0.17	0.19	0.18	0.17	0.19	0.19	0.19	0.22	0.28	0.27
qz2p	18.7	0.09	0.10	0.17	0.19	0.18	0.16	0.18	0.19	0.19	0.22	0.28	0.26
cc-pVDZ	9.4	0.25	0.23	0.23	0.27	0.26	0.26	0.28	0.27	0.27	0.31	0.36	0.35
cc-pVTZ	21.6	0.07	0.11	0.16	0.19	0.18	0.16	0.18	0.19	0.19	0.22	0.28	0.27
cc-pVQZ	41.8	0.00	0.11	0.17	0.19	0.18	0.16	0.17	0.19	0.19	0.20	0.27	0.26

[a]Average number of basis functions per atom (determined for the molecular benchmark set).

dx.doi.org/10.1021/ct400780f | J. Chem. Theory Comput. 2014, 10, 572−578

Table 2. Standard Deviation (STD) of C Shifts for Different Basis Sets and Methods with Respect to the Reference Calculation CCSD(T)/cc-pVQZ

basis set	$\bar{N}_{bas}$ [a]	CCSD(T)	CCSD(T)[b]	MP2	MP2[b]	HF	PBE0	mPW1PW	B97-2	B3LYP	B3PW91	mPW3PW	KT2	PBE	BP86
STO-3G	3.0	25.8	8.6	27.0	9.9	16.5	20.3	20.4	21.1	21.0	20.8	20.7	23.6	22.2	22.4
3-21G	5.4	11.9	4.4	13.0	6.4	7.4	7.6	7.5	8.6	8.2	7.9	7.9	12.4	10.0	10.1
6-31G**	9.8	7.2	2.2	7.5	2.7	5.5	3.3	3.4	4.4	4.5	3.8	3.8	8.6	5.9	6.2
6-311G**	11.9	3.5	1.8	3.9	2.9	8.2	3.4	3.3	2.5	3.5	3.3	3.4	4.9	3.7	3.7
pcS-0	5.4	8.1	6.8	6.7	9.4	19.2	11.2	11.1	9.8	10.4	10.7	10.8	7.3	9.6	9.4
pcS-1	10.8	2.0	1.4	2.5	2.6	9.0	4.6	4.5	3.2	4.7	4.4	4.5	4.2	4.5	4.3
pcS-2	23.2	1.0	0.5	2.5	2.2	9.4	6.2	6.2	4.5	6.4	6.1	6.2	4.1	6.0	5.8
def2-SVP	9.4	5.7	2.8	6.3	3.4	5.5	3.5	3.6	4.2	4.4	4.0	4.0	7.7	5.7	5.9
def2-TZVP	17.9	1.0	0.4	2.4	2.2	7.8	4.2	4.1	3.2	4.4	4.1	4.3	5.2	4.8	4.7
tz2p	16.3	1.7	0.8	2.6	2.2	8.1	3.9	3.8	2.7	4.0	3.7	3.9	4.6	4.2	4.1
qz2p	18.7	0.7	0.8	2.1	2.3	8.7	4.9	4.9	3.4	5.2	4.8	5.0	4.2	4.9	4.8
cc-pVDZ	9.4	6.0	2.9	6.7	3.5	5.7	3.8	3.8	4.4	4.6	4.1	4.2	7.6	5.8	6.0
cc-pVTZ	21.6	1.5	0.5	2.5	2.1	7.5	3.8	3.7	2.8	4.1	3.7	3.8	5.2	4.5	4.4
cc-pVQZ	41.8	0.0	0.0	2.1	2.1	8.3	5.0	5.0	3.5	5.2	4.9	5.1	4.5	5.2	5.1

[a]Average number of basis functions per atom (determined for the molecular benchmark set). [b]The HF part of the C shifts is calculated with the basis set cc-pVQZ (see Section 5 for a detailed discussion).

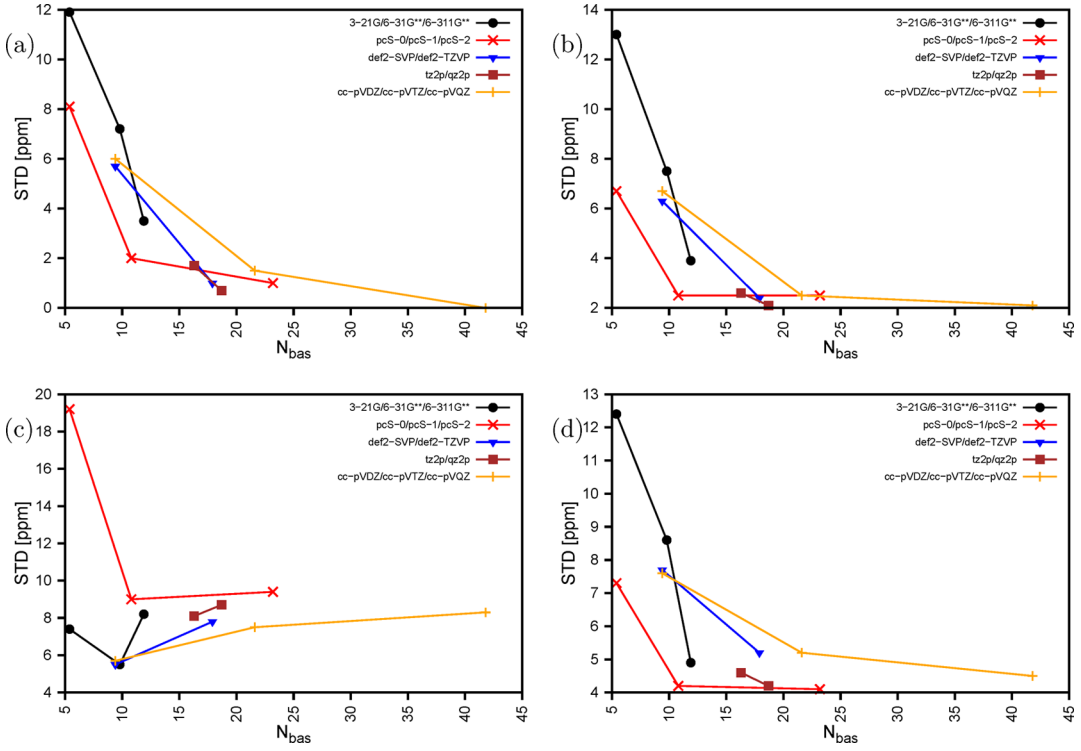


Figure 4. Dependence of the STD for C shifts with respect to the reference calculation CCSD(T)/cc-pVQZ on the mean number of basis functions per atom N_{bas} for (a) CCSD(T), (b) MP2, (c) HF, and (d) KT2. Here, five different basis set series are distinguished.

first column of Tables 1 and 2 lists deviations for CCSD(T) calculations caused by smaller basis sets, while the other columns list deviations caused by the interplay of method and basis set approximations (in addition, by referring to the results obtained for a cc-pVQZ basis within the respective method, the pure effects of smaller basis sets for each method can be extracted; see the Supporting Information). The STD for the cc-pVQZ basis visualized in Figure 3 is listed again in the last row. The size of the basis sets is specified by a single measure, that is, the average number $\bar{N}_{bas}$ of basis functions per atom computed over all atoms of the molecular benchmark set (the details for the basis set composition are provided in the Supporting Information in the form of contraction schemes).

The data for hydrogen shieldings in Table 1 show only weak influences by basis set changes (only for the rather poor basis sets STO-3G, 3-21G, and pcS-0, larger changes occur). Therefore, we focus in the following on carbon NMR shieldings as listed in Table 2: To begin with, the most pronounced increase of deviations for all methods occurs below the levels of double-ζ basis sets with polarization functions (DZP), e.g., 6-31G**, pcS-1 (as one might expect). The smallest basis set usable for wave-function-based correlation methods (CCSD(T), MP2) appears to be the pcS-1 basis introduced by Jensen.[32] Here, the wave-function-based correlation methods reach a strikingly good performance as compared to DFT with a particularly good compromise between accuracy and cost: the

dx.doi.org/10.1021/ct400780f | J. Chem. Theory Comput. 2014, 10, 572−578

average number of basis functions per atom for the present benchmark set, abbreviated by $\bar{N}_{bas}$, is 10.8. Also for Hartree−Fock and the various DFT functionals, the pcS-1 basis set clearly comes closest to the corresponding STDs obtained with the cc-pVQZ basis set (among the basis sets of similar size 6-31G**, pcS-1, def2-SVP, or cc-pVDZ). A further analysis (see the Supporting Information) shows that standard deviations of only 1.4−1.6 ppm are reached for HF and DFT compared to the results obtained with the respective method at the cc-pVQZ level (as compared to 3.9−5.9 ppm, e.g., for def2-SVP).

For higher accuracies than obtained with the cost-effective pcS-1 basis, the data in Table 2 indicate the very good performance of def2-TZVP[63] and, in particular, qz2p basis sets[30,64] with a size $\bar{N}_{bas}$ of 17.9 and 18.7, respectively. The latter is only slightly larger than the def2-TZVP basis but still noticeably better performing for the calculation of NMR chemical shifts. These basis sets are clearly a much better high-accuracy compromise than the more than twice as large cc-pVQZ basis[42] ($\bar{N}_{bas}$ = 41.8).

While the error reduces consistently with the size of the basis set for the wave-function-based correlation methods CCSD(T) and MP2, as well as the KT2 functional, deviating behavior can be observed for the Hartree−Fock method and other considered DFT functionals. Here, the data indicate that method and basis set errors to some extent cancel statistically, especially in the regime of DZP quality (while of course there is no guarantee that for a specific nucleus an error cancellation occurs). The different behavior with increasing basis set size is visualized in Figure 4 for the four representative examples of CCSD(T), MP2, HF, and KT2. For the larger basis sets, the B97-2 functional performs best of the tested DFT functionals, followed by the KT2 functional.

More data are listed in the Supporting Information that indicates that the general aspects listed above are also reflected by other error criteria (e.g., mean absolute and maximum deviations).

5. ERROR DISTRIBUTION FOR SPIN COMPONENT-SCALED MP2 (GIAO-SCS-MP2) APPROACHES

In addition to the MP2 NMR shieldings, we also list data for our newly introduced GIAO-SCS-MP2 shieldings method, which employs scaling coefficients for the same- and opposite-spin components and for which details of the fitting procedure and extensive accuracy studies will be presented elsewhere.[67] The latter scheme follows the lines of scaled MP2 methods developed for ground state energetics by Grimme[68] and others (see, e.g., refs 69 and 70 or ref 71 for a recent review). Analogously, one can split the NMR shielding tensor at the MP2 level into the HF contribution and the opposite spin (OS) and same spin (SS) terms of the perturbative second-order correction:

$$\sigma_{MP2} = \sigma_{HF} + c_{OS}\,\sigma_{OS} + c_{SS}\,\sigma_{SS}$$

The scaling factors c_{OS} and c_{SS} have been optimized by a fitting procedure with respect to the NMR shieldings at the CCSD(T)/cc-pVQZ level for each single basis set. The fitting procedure itself employs only one-half of the nuclei of the total benchmark set, while for analyzing the error distribution we include all nuclei. Furthermore, in the fitting procedure, we allow for systematic deviations of (absolute) shieldings, because our focus is on relative shieldings (shifts) and include for this purpose a constant offset (denoted as MSD_{opt}):

$$\sigma_{SCS/SOS\text{-}MP2} + MSD_{opt} \overset{!}{=} \sigma_{CCSD(T)}$$

Table 3 lists the results for spin component-scaled MP2 approaches for NMR shifts with respect to the analogously computed TMS values:

Table 3. Standard Deviations (STD) of C and H Shifts for the GIAO-SCS-MP2 and GIAO-SOS-MP2 Methods Introduced in This Work Employing a Variety of Basis Sets[a]

basis set	$\bar{N}_{bas}$[b]	CCSD(T)	GIAO-SCS-MP2	GIAO-SOS-MP2	MP2	B97-2
C shifts						
STO-3G	3.0	25.8	13.7	15.4	27.0	21.1
3-21G	5.4	11.9	5.1	7.4	13.0	8.6
6-31G**	9.8	7.2	2.4	4.4	7.5	4.4
6-311G**	11.9	3.5	2.6	3.7	3.9	2.5
pcS-0	5.4	8.1	6.5	6.7	6.7	9.8
pcS-1	10.8	2.0	2.2	2.9	2.5	3.2
pcS-0[c]		6.8	3.1	3.4	9.4	
pcS-1[c]		1.4	1.6	2.0	2.6	
H shifts						
pcS-0	5.4	0.47	0.36	0.45	0.41	0.51
pcS-1	10.8	0.16	0.14	0.18	0.14	0.16

[a]As the reference, CCSD(T)/cc-pVQZ results are used. For comparison, CCSD(T), MP2, and B97-2 results are again listed (see Table 2). More details for the GIAO-SCS-MP2 and GIAO-SOS-MP2 results will be presented elsewhere.[67] [b]Average number of basis functions per atom (determined for the moclecular benchmark set). [c]The HF part of the C shifts is calculated with the basis set cc-pVQZ.

•In particular, for the smaller basis sets like STO-3G, 3-21G, and 6-31G**, the scaling of the MP2 values reduces the STDs strongly. For example, the STD using the 6-31G** basis is close to the STDs with larger basis sets like 6-311G** and pcS-1.

• The STD of GIAO-SCS-MP2/pcS-1 is close to the value of CCSD(T)/pcS-1.

The proposed GIAO-SCS-MP2 method introduced above employs the same basis set for the HF and the perturbative second-order correction terms. A pragmatic alternative and less common approach in striving for a good compromise between accuracy and cost would be to employ a larger basis for the HF term to reduce the HF error to a minimum: We tested here the cc-pVQZ basis set. The results for this additional approach are listed in Table 3: The STD for GIAO-SCS-MP2/pcS-0 reduces from 6.5 to 3.1 ppm and for pcS-1 from 2.2 to 1.6 ppm. Remarkably, the mixed basis set GIAO-SCS-MP2 approach results for the latter case in smaller STDs than obtained by conventional CCSD(T)/pcS-1 calculations. In contrast to the good benefit in case of the GIAO-SCS-MP2 method, an analogous mixed approach for nonscaled MP2 shows to be far less beneficial. Actually, for some cases one obtains even larger STDs. Table 2 gives an overview for the mixed basis set approach for all basis sets and also in combination with the CCSD(T) method. Especially, the improvements for CCSD(T) in combination with the basis sets pcS-2, def2-TZVP, and cc-pVTZ to STDs below 1 ppm are remarkable.

The relative benefits of the GIAO-SCS-MP2 method to the nonscaled MP2 method were also investigated for H shifts for the basis sets pcS-0 and pcS-1 (see Table 3): The STD of

dx.doi.org/10.1021/ct400780f | J. Chem. Theory Comput. 2014, 10, 572−578

GIAO-SCS-MP2/pcS-0 drops from 0.41 to 0.36 ppm, whereas the STD of GIAO-SCS-MP2/pcS-1 remains unchanged.

Further details for the GIAO-SCS-MP2 approach are given in the Supporting Information and in a future publication.[67]

6. ERROR DISTRIBUTION FOR MOLECULAR SUBSETS

So far, we considered standard deviations for the complete molecular test set and reported mean absolute deviations with respect to TMS only. However, as concluded by several previous studies, the errors in calculating NMR shifts can be decreased substantially, if the shifts of the probed structure are computed with respect to a closely related structure. This refers to the concept of intermediate references,[72,73] or multistandards (see, e.g., refs 74 and 75), where the total shift with respect to TMS, $\delta^{high/low}_{TMS-sample}$, is determined as the sum of a highly accurate theoretical or experimental shift value, $\delta^{high}_{TMS-Int.ref.}$, for the intermediate reference (or second standard compound in a multistandard approach) and an incremental shift, $\delta^{low}_{Int.ref-sample}$, computed at a lower level:

$$\delta^{high/low}_{TMS-sample} = \overbrace{\sigma^{high}_{TMS} - \sigma^{high}_{Int.Ref.}}^{\delta^{high}_{TMS-Int.Ref.}} + \overbrace{\sigma^{low}_{Int.Ref.} - \sigma^{low}_{sample}}^{\delta^{low}_{Int.Ref-sample}}$$

To appraise the possible accuracy gains by employing intermediate references (or multistandards), Figure 5 depicts

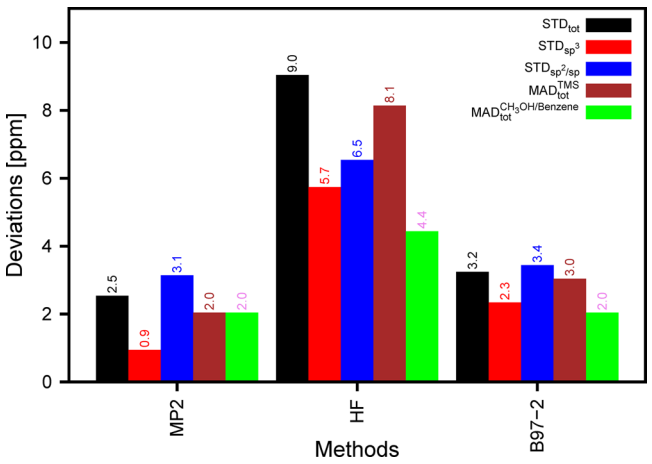


Figure 5. STDs (in ppm) for subsets of sp^3 or sp^2/sp hybridized carbon atoms and MADs with respect to TMS, and for an approach with two intermediate references (CH_3OH, benzene) at GIAO MP2, B97-2, and HF levels with respect to CCSD(T) results (basis set pcS-1 and CCSD(T)/cc-pVTZ geometries).

the standard deviations calculated for subsets of sp^3 or sp^2/sp hybridized carbon atoms for the example of the (well-performing) pcS-1 basis set and the MP2, HF, and B97-2 method. On the one hand, the figure indicates that electron correlation plays a major role within the sp^2/sp subset compared to the sp^3 subset (for all methods the $STD_{sp2/sp}$ is larger than the STD_{sp^3}). Especially for MP2, the sp^3 subset is very well described with STDs of only 0.9 ppm. On the other hand, the results reflect the benefits from the separation into subsets and referring to a chemically related nucleus. For example, if all the sp^3 carbon atoms are referred to $\underline{C}H_3OH$ and the sp^2/sp carbon atoms are referred to benzene, an overall $MAD^{CH_3OH/benzene}_{tot}$ of 2.0/4.4/2.0 can be reached for MP2/HF/B97-2 ($\underline{C}H_3OH$ and benzene are chosen here as an example, motivated by refs 74 and 75). Here, the overall MAD for using the two standards is computed as the average of the subset MADs weighted by the number of nuclei per subset (15 sp^3 and 25 sp^2/sp nuclei).

7. CONCLUSION

Within the present work, we have benchmarked the errors for calculating NMR shifts by MP2, HF, and DFT approaches based on a broad molecular set with large importance for computations of hydrogen and carbon NMR shifts in organic compounds. At the DZP level and above, the assigned standard deviations for hydrogen vary in between 0.10−0.23/0.16−0.23/0.15−0.36 ppm for MP2/HF/DFT and for carbon in between 2.1−7.4/5.5−9.4/2.5−8.6 ppm depending on the basis set. Here, among the DFT functionals, the considered hybrid-GGA functional B97-2 reaches rather constantly the lowest standard deviations. A further reduction of the standard deviation beyond the DFT level can be reached by MP2, where the study identifies the Jensen basis set pcS-1 as a remarkably well performing cost-efficient alternative to larger (TZP or QZP) basis sets. Furthermore, our new scaled MP2 approach for NMR shieldings is shown to be a helpful extension that can considerably reduce the deviation with respect to the CCSD(T)/cc-pVQZ reference values.

■ ASSOCIATED CONTENT

Supporting Information

Explicit NMR shift values for all considered nuclei and all levels of theory, additional error criteria, contraction schemes for all basis sets, and further details for the GIAO-SCS-MP2 approach. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: christian.ochsenfeld@cup.uni-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. J. Kussmann (LMU) for support concerning calculations of NMR shielding tensors, J. Glänzer (student at LMU) for participating in geometry optimizations and NMR calculations during a practical course, and A. Lünser (LMU) for further helpful discussions. C.O. acknowledges financial support by the Volkswagen Stiftung within the funding initiative "New Conceptual Approaches to Modeling and Simulation of Complex Systems," by the SFB 749 "Dynamik und Intermediate molekularer Transformationen" (DFG), and the DFG cluster of excellence EXC 114 "Center for Integrative Protein Science Munich" (CIPSM).

■ REFERENCES

(1) Bühl, M.; Malkin, V. G.; Kaupp, M. *Calculation of NMR and EPR Paramaters, Theory and Applications*; Wiley-VCH: Weinheim, Germany, 2004.

(2) Gauss, J. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; John von Neumann Institute for Computing: Julich, 2000; Vol. 3, pp 541−592.

(3) Vaara, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5399−5418.

(4) Helgaker, T.; Jaszuński, M.; Ruud, K. *Chem. Rev.* **1999**, *99*, 293−352.

(5) Kussmann, J.; Beer, M.; Ochsenfeld, C. *WIREs Comput. Mol. Sci.* **2013**, *3*, 614−636.

(6) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789.
(7) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(8) Häser, M.; Ahlrichs, R.; Baron, H.; Weis, P.; Horn, H. *Theor. Chim. Acta* **1992**, *83*, 455−470.
(9) Wolff, S. K.; Ziegler, T. *J. Chem. Phys.* **1998**, *109*, 895−905.
(10) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629−3643.
(11) Kollwitz, M.; Gauss, J. *Chem. Phys. Lett.* **1996**, *260*, 639−646.
(12) Kollwitz, M.; Häser, M.; Gauss, J. *J. Chem. Phys.* **1998**, *108*, 8295−8301.
(13) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *103*, 3561−3577.
(14) Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1997**, *276*, 70−77.
(15) Ochsenfeld, C.; Head-Gordon, M. *Chem. Phys. Lett.* **1997**, *270*, 399−405.
(16) Ochsenfeld, C.; Kussmann, J.; Koziol, F. *Angew. Chem., Int. Ed.* **2004**, *43*, 4485−4489.
(17) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2007**, *127*, 054103.
(18) Beer, M.; Ochsenfeld, C. *J. Chem. Phys.* **2008**, *128*, 221102.
(19) Larsen, H.; Helgaker, T.; Olsen, J.; Jorgensen, P. *J. Chem. Phys.* **2001**, *115*, 10344−10352.
(20) Niklasson, A. M. N.; Challacombe, M. *Phys. Rev. Lett.* **2004**, *92*, 193001.
(21) Beer, M.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2011**, *134*, 074102.
(22) Maurer, M.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 174104.
(23) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 014101.
(24) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629−3643.
(25) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497−5509.
(26) Olsson, L.; Cremer, D. *J. Chem. Phys.* **1996**, *105*, 8995−9006.
(27) Adamo, C.; Barone, V. *Chem. Phys. Lett.* **1998**, *298*, 113−119.
(28) Rablen, P. R.; Pearlman, S. A.; Finkbiner, J. *J. Phys. Chem. A* **1999**, *103*, 7357−7363.
(29) Sun, H.; Sanders, L. K.; Oldfield, E. *J. Am. Chem. Soc.* **2002**, *124*, 5486−5495.
(30) Auer, A. A.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2003**, *118*, 10407−10417.
(31) Manninen, P.; Vaara, J. *J. Comput. Chem.* **2006**, *27*, 434−445.
(32) Jensen, F. *J. Chem. Theory Comput.* **2008**, *4*, 719−727.
(33) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2008**, *112*, 6794−6799.
(34) Kongsted, J.; Aidas, K.; Mikkelsen, K. V.; Sauer, S. P. A. *J. Chem. Theory Comput.* **2008**, *4*, 267−277.
(35) Mulder, F. A. A.; Filatov, M. *Chem. Soc. Rev.* **2010**, *39*, 578−590.
(36) Gregušõova, A.; Perera, S. A.; Bartlett, R. J. *J. Chem. Theory Comput.* **2010**, *6*, 1228−1239.
(37) Kupka, T.; Stachów, M.; Nieradka, M.; Kaminsky, J.; Pluta, T. *J. Chem. Theory Comput.* **2010**, *6*, 1580−1589.
(38) Teale, A. M.; Lutnaes, O. B.; Helgaker, T.; Tozer, D. J.; Gauss, J. *J. Chem. Phys.* **2013**, *138*, 024111.
(39) Kupka, T.; Stachów, M.; Kaminsky, J.; Sauer, S. P. A. *Magn. Reson. Chem.* **2013**, *51*, 482−489.
(40) Kupka, T.; Stachów, M.; Nieradka, M.; Kaminsky, J.; Pluta, T.; Sauer, S. P. A. *Magn. Reson. Chem.* **2011**, *49*, 231−236.
(41) Jameson, A. K.; Jameson, C. J. *Chem. Phys. Lett.* **1987**, *134*, 461−466.
(42) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(43) CFOUR, a quantum chemical program package written by Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusélius, J., Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Ruud, K.; Schiffmann, F.; Stopkowicz, S.; Tajti, A.; Vazquez, J.; Wang, F.; Watts, J. D. and the integral packages MOLECULE (Almlöf, J.; Taylor, P. R.), *PROPS* (Taylor, P. R.), *ABACUS* (Helgaker, T.; Jensen, H. J.; Jorgensen, P.; Olsen, J.), and ECP routines by Mitin, A. V.; van Wullen, C. For the current version, see http://www.cfour.de.
(44) Development version of Q-Chem. www.q-chem.com.
(45) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 134114.
(46) Library of exchange and correlation functionals for DFT (libxc). www.tddft.org/programs/octopus/wiki/index.php/Libxc.
(47) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889−2899.
(48) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664−675.
(49) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233−9242.
(50) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(51) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(52) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(53) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(54) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(55) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 3015−3024.
(56) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657−2664.
(57) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769−2773.
(58) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939−947.
(59) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797−2803.
(60) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(61) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(62) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.
(63) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(64) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571−2577.
(65) Feller, D. *J. Comput. Chem.* **1996**, *17*, 1571−1586.
(66) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, *47*, 1045−1052.
(67) Maurer, M.; Ochsenfeld, C. In preparation.
(68) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095−9102.
(69) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793−9802.
(70) Fink, R. F. *J. Chem. Phys.* **2010**, *133*, 174113.
(71) Grimme, S.; Goerigk, L.; Fink, R. F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 886−906.
(72) Ochsenfeld, C.; Koziol, F.; Brown, S. P.; Schaller, T.; Seelbach, U. P.; Klärner, F.-G. *Solid State Nucl. Magn. Reson.* **2002**, *22*, 128−153.
(73) Zienau, J.; Kussmann, J.; Ochsenfeld, C. *Mol. Phys.* **2010**, *108*, 333−342.
(74) Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* **2009**, *74*, 7254−7260.
(75) Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* **2012**, *77*, 6059−6065.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on January 23, 2014. Tables 36 and 37 in the Supporting Information have been modified. The correct version was published on January 31, 2014.