JCTC Journal of Chemical Theory and Computation

Article

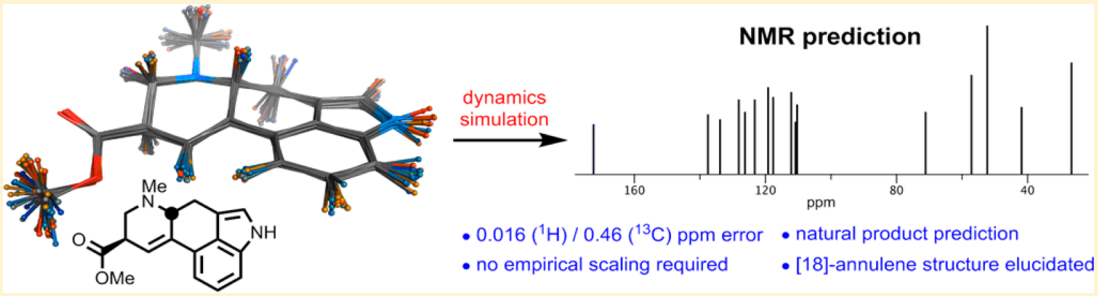
pubs.acs.org/JCTC

Enhancing NMR Prediction for Organic Compounds Using Molecular Dynamics

Eugene E. Kwan* and Richard Y. Liu

Department of Chemistry & Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

Supporting Information



ABSTRACT: NMR spectroscopy is a crucial tool in organic chemistry for the routine characterization of small molecules, structural elucidation of natural products, and study of reaction mechanisms. Although there is evidence that thermal motions strongly affect observed resonances, conventional predictions are performed only on stationary structures. Here we show that quasiclassical molecular dynamics provides a highly accurate and broadly applicable method for improving shielding predictions. Gas-phase values of the absolute shieldings of protons and carbons are predicted to nearly within experimental uncertainty, while the chemical shifts of large systems such as natural products are closely reproduced. Importantly, these results are obtained without the use of any empirical corrections. Our analysis suggests that the linear scaling factors currently employed are primarily a correction for vibrational effects. As a result, our method extends the reach of prediction methods to the study of molecules with unusual dynamics such as the iconic and controversial [18]annulene. Our predictions agree closely with experiment at both low and high temperatures and provide strong evidence that the equilibrium structure of [18]annulene is planar and aromatic.

1. INTRODUCTION

Since the 1960s, high-resolution nuclear magnetic resonance (NMR) spectroscopy has enabled the rapid determination and investigation of molecular structure in organic chemistry. This revolutionary technique depends on the sensitivity of the resonance frequency of an atomic nucleus (shielding) to its chemical environment when a magnetic field is applied. Today, shieldings and other NMR properties can be predicted using density functional theory (DFT) and gauge-including atomic orbital (GIAO) methods[1] with enough accuracy to distinguish between stereoisomers of natural products, identify reactive intermediates, and study fundamental concepts such as aromaticity.[2,3] Despite these successes, most prediction methods require the assistance of empirical corrections to achieve high accuracy. Using quasiclassical molecular dynamics,[4a,b] we demonstrate that these corrections implicitly account for the effects of molecular motions. The approach is highly accurate for small molecules in the gas phase, can be applied to relatively large systems such as natural products, and extends to molecules with unusual dynamics such as [18]annulene.

During an NMR experiment, rotations and vibrations cause each molecule to sample a large number of geometries. The observed time-averaged shielding is generally quite different from the shielding that would be expected for the minimum-energy geometry (stationary point). The difference in shieldings is called the "rovibrational correction" and is the result of two factors: geometric differences between the dynamic and static structures and asymmetry in the dependence of shielding on displacement (Figures S5−S7). The importance of rovibrational corrections has been clearly demonstrated in benchmark-level calculations of absolute shieldings in small systems. For example, Auer, Gauss, and Stanton showed that vibrational corrections to carbon absolute shieldings are large (several parts per million, or ppm) and that these corrections are essential for reproducing experimental observations.[5] However, because significant error cancellation results when only differences in shieldings are considered, it is generally assumed that rovibrational effects are small and only one of the many factors that contribute to error in NMR predictions.[2] Nevertheless, Ruud et al.[6] found that corrections to proton chemical shifts can still exceed 1 ppm in some cases. This may explain why conventional predictions are accurate only when they are combined with empirical scaling factors: linear scaling compensates for the neglect of molecular motions.

On the basis of this hypothesis, we investigated whether explicit modeling of molecular motions might lead to improved

Received: September 5, 2015
Published: October 7, 2015

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

NMR predictions in many systems, including those that linear scaling fails to correct. Although previous approaches have relied on the analytic computation of the expected shielding over molecular vibrations,[7] they have been applied only to very small molecules. For example, 1-adamantyl cation is one the largest systems that been studied rigorously in this manner.[8] In contrast, stochastic approaches employing molecular dynamics do not require technically demanding shielding derivatives, can implicitly account for anharmonicity, and have been applied to dihydrogen in a fullerene cage[9] and even protein chemical shifts.[10] However, purely classical simulations like these ignore zero-point energy and give unrealistic vibrational behavior. At room temperature, the equipartition energy of $RT/2$ per degree of freedom is 0.3 kcal/mol, which is much less than the zero-point energy of 4.2 kcal/mol for a typical C−H bond. As a result, displacements are systematically underestimated (Table S5). Our simple solution is to use quasiclassical dynamics (QCD),[4c−f] which has been successfully applied to the study of reaction dynamics for decades.[4b,d,e] In QCD, the position and momentum of each atom are initialized with quantum mechanics, but the system evolves classically. In the following, we show that the use of QCD dramatically enhances the prediction accuracy while maintaining the simplicity of a stochastic approach.

2. COMPUTATIONAL METHODS

Programs Used. Unless otherwise noted, all of the DFT calculations were performed using Gaussian 09. GIAO shieldings were obtained using Gaussian 09 (DFT methods) and CFOUR version 1 (coupled cluster methods). Quasiclassical dynamics calculations were performed using in-house software written in Java. DLPNO−CCSD(T) single-point energies were obtained with Orca 3.0.3 using normal cutoffs. Central composite design and response surface analysis were carried out using JMP version 12. Please see the Supporting Information for further details and citations of these programs.

General Procedure for Quasiclassical Calculation of the Raw Dynamic Correction. For the molecule of interest, the geometry of the lowest-energy conformation is optimized at the B3LYP/MIDI! level, and a frequency analysis is performed at the same level. For each vibrational mode, excluding those with wavenumbers below 50 cm^{-1}, an energy level is randomly selected from a Boltzmann distribution, with the assumption that the vibration can be modeled as a simple quantum harmonic oscillator (Figure 1a). A displacement is randomly made in each mode according to the displacement-space probability distribution of a simple quantum harmonic oscillator eigenstate of the appropriate level. The molecule is initialized with the classical amount of kinetic energy necessary in each mode on the basis of the displacements, with the direction of mode velocities sign-randomized. Rotation is imparted to the entire molecule using a classical initialization scheme. For each principal axis, an angular momentum is drawn from a Gaussian distribution, such that the average energy of rotation around each axis is equal to the classical equipartition energy of $kT/2$.

After initialization, trajectory propagation is performed using the velocity Verlet integration scheme in increments of 1.0 fs (Figure 1b). The required forces are calculated at each point at the B3LYP/MIDI! level, and 125 points are calculated in both the forward and reverse directions. For a typical natural-product-sized molecule, 25 trajectories are sufficient. After completion of the simulation, every eighth point is selected, and

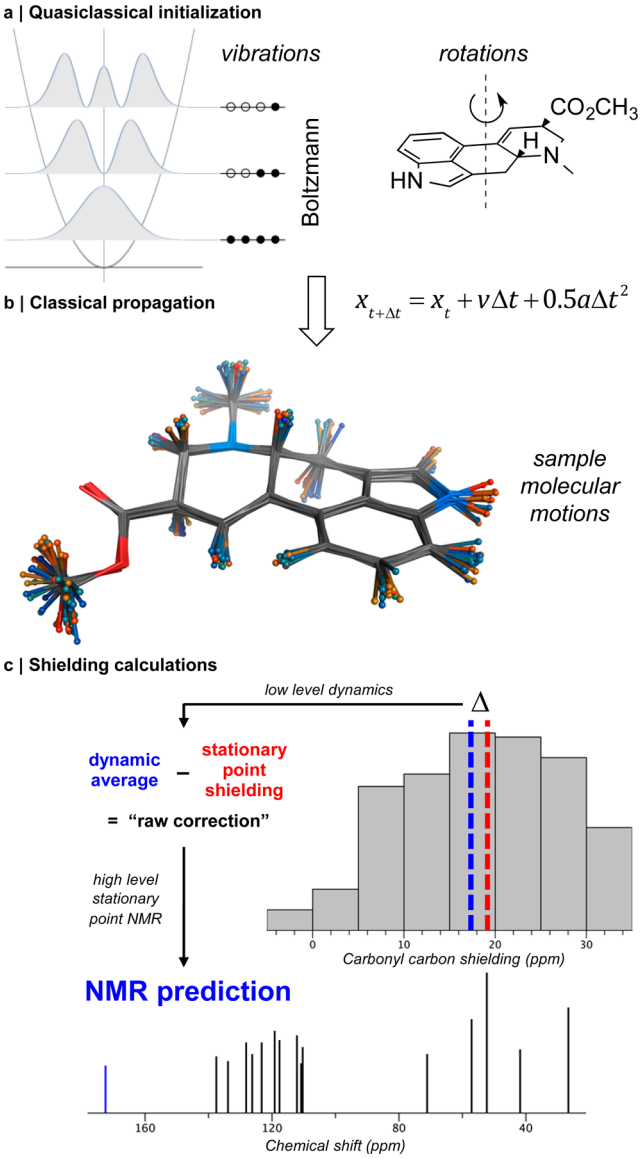


Figure 1. Correcting NMR predictions for molecular motions. (a) Thermal vibrational energy and displacements are initialized on the basis of the eigenstates of the quantum harmonic oscillator. Rotational energy is also added by sampling a classical thermal distribution for angular momentum. (b) Classical trajectories are propagated using the velocity Verlet scheme at a low level of theory for speed. NMR shieldings are calculated periodically (protons are colored by shielding). (c) The difference between the average shielding and the stationary-point shielding is the "raw correction." This is applied to a high-level stationary-point shielding calculation to obtain high-quality NMR predictions.

the shieldings are calculated using the GIAO scheme at the B3LYP/cc-pVDZ level. These shieldings are symmetry-averaged on an atom-by-atom basis. Finally, shieldings calculated using the same level of theory on the initial low-level stationary point are subtracted from these dynamically averaged shieldings to yield the raw corrections for the molecule. Statistical uncertainties are calculated by treating corrections from different trajectories as independent. The standard errors were below 0.01 ppm (proton) and 0.1 ppm (carbon) for all of the calculations.

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

3. RESULTS AND DISCUSSION

3.1. Method Development. We assumed that electron correlation and rovibrational motion can be separately addressed by static calculations at a high level of theory and quasiclassical dynamics at a low level of theory:

$$\langle\sigma\rangle = \sigma^{\text{high-level}}_{\text{stationary}} + (\langle\sigma^{\text{low-level}}_{\text{dynamic}}\rangle - \sigma^{\text{low-level}}_{\text{stationary}})$$

where $\langle\sigma\rangle$ is the rovibrationally corrected absolute shielding prediction, $\sigma^{\text{high-level}}_{\text{stationary}}$ is the static shielding at a high-level stationary point, $\langle\sigma^{\text{low-level}}_{\text{dynamic}}\rangle$ is the dynamically averaged shift using a low level of theory, and $\sigma^{\text{low-level}}_{\text{stationary}}$ is the static shielding at the low-level stationary point. This second bracketed term, which we call the "raw correction," allows for molecular motion (Figure 1c).

Our composite approach assumes that the shape of the potential energy surface (PES) at the low level of theory is similar to that of the true PES, at least in the accessible phase space. To evaluate this requirement, we compared the raw corrections (B3LYP/cc-pVTZ/GIAO) obtained from the PESs of common DFT methods to those from a highly accurate coupled cluster method and found B3LYP to be optimal (Table 1a). We also examined the effect of the dynamics basis set and NMR method on the accuracy of the predicted carbon absolute shieldings (Table 1b). We found the minimal basis set MIDI! to be optimal for dynamics.[11] While this is likely the result of fortuitous error cancellation, it allows large systems to be studied quickly. The pcS-3 basis set was optimal for NMR points, but the much smaller cc-pVDZ basis set resulted in only a small loss of accuracy. Interestingly, the addition of classical rotations to the quasiclassical vibrational initialization procedure resulted in uniformly lower regression intercepts (Table S6). Presumably, vibrational−rotational coupling and centrifugal distortion effects are more important for small molecules, which rotate quickly at temperatures of interest.

Table 1. Choice of Dynamics Method

(a) Effect of the Method Used To Obtain the PES[a]

	proton RMSE/ppm		carbon RMSE/ppm	
method	cc-pVDZ	cc-pVTZ	cc-pVDZ	cc-pVTZ
HF	0.08	0.08	0.74	0.52
MP2	0.11	not calcd	0.38	not calcd
B3LYP	0.03	0.04	0.19	0.28
PBE0	0.02	0.02	0.24	0.36
M06-2X	0.03	0.02	0.35	0.47

(b) Effect of the Dynamics Basis Set and NMR Method[b]

	NMR basis set				
dynamics basis set	6-31G*	cc-pVDZ	cc-pVTZ	pcS-2	pcS-3
MIDI!	0.57	0.50	0.44	0.40	0.40
6-31G*	0.61	0.49	0.49	0.47	0.46
cc-pVDZ	0.58	0.49	0.46	0.43	0.43

[a]Root-mean-square errors (RMSEs) compare quasiclassical raw corrections computed with various potential energy surfaces to those calculated with the CCSD/cc-pVTZ surface (B3LYP/cc-pVTZ NMR points) for methane, ethane, ethylene, acetylene, methyl fluoride, methanol, methylamine, hydrogen cyanide, carbon monoxide, carbon dioxide, and formaldehyde. [b]Values shown are RMSEs (in ppm) for quasiclassical predictions of carbon absolute shieldings for the molecules considered in Table 1a. All of the calculations were performed with the B3LYP functional.

Our interest in large molecules led us to seek further efficiencies. In the most straightforward procedure, each trajectory point requires energy, gradient, and shielding calculations, of which the last are by far the most expensive. However, because the geometries of consecutive trajectory points are highly correlated, shielding calculations are not necessary for each point. Our data indicate that an 8 fs NMR point interval leads to no loss of accuracy (Figure S3), allowing significant time savings. Furthermore, bootstrap resampling analysis of repeated trajectory studies on acetylene showed that the sampling distribution is nearly independent of the trajectory length beyond 250 fs (Figures S1 and S2). Because increased simulation times merely re-explore the same phase space, several shorter trajectories are more efficient than one longer trajectory.

3.2. Gas-Phase Absolute Shieldings. To test our method in the absence of confounding effects, we predicted the gas-phase absolute shieldings of the proton and carbon nuclei in a set of small, rigid molecules. For proton nuclei, we considered a diverse range of sp, sp^2, sp^3, and exchangeable protons selected from the high-quality experimental data set reported in 2012 by Garbacz et al.[12] (Table 2). Using B3LYP/MIDI! dynamics, B3LYP/pcS-3 NMR points, and CCSD(T)/pcS-3//CCSD(T)/cc-pVQZ stationary shieldings, we obtained a mean absolute error (MAE) of 0.016 ppm. Although no direct comparison is possible, previous estimates had random and systematic errors of approximately 0.1 and 0.3 ppm, respectively.[13] Our method also performs well for carbon shieldings. With the minor modification of using known CCSD(T)/13s9p4d3f//CCSD(T)/cc-pVTZ stationary-point shieldings and with the same test set as considered by Auer et al.,[5] we obtained an MAE of 0.46 ppm, compared with the previous value of 1.67 ppm. Importantly, the residuals are small even for challenging molecules like acetylene, allene, and carbon monoxide. The biggest outlier is CF_4, which may suffer from known deficiencies in DFT methods in the treatment of highly electronegative atoms with several lone pairs, such as fluorine.[14] Overall, these results suggest that QCD is highly accurate and should be applicable to more complex systems.

3.3. Natural Products. NMR spectroscopy is crucial for the structural elucidation of natural products, where errors can lead to significant delays in synthetic or medicinal investigations.[29] Although coupling constants and nuclear Overhauser effect correlations are adequate for routine cases, statistical tools[15] for isomer discrimination based on the comparison of predicted and observed chemical shifts have emerged as the method of choice for difficult cases. Because this process relies on the capacity to distinguish between nuclei in similar chemical environments, it is essential to choose an accurate shift prediction method.

In general, the performance of a prediction scheme depends on its ability to treat molecular motion as well as many other factors such as electron correlation, conformational equilibria, solvation, and relativistic effects. While increasingly accurate methods have been developed to treat these latter effects, we sought to create a benchmark that would most clearly reveal any beneficial effect of rovibrational corrections with minimal confounding influences. Accordingly, a high-quality data set was constructed from 14 natural products that are representative of a range of structural classes (Figure 2). Our choices were restricted to rigid[30] polycyclic structures that have no halogen atoms and have well-characterized spectral data in chloroform. Raw dynamic corrections[31] were obtained using the previously

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

Table 2. Predictions of Gas-Phase Absolute Shieldings (in ppm)[a]

atom (bold)	stationary	correction	prediction	expt	error
CH_4	31.26	−0.64	30.62	30.63	−0.01
C_2H_6	30.63	−0.74	29.89	29.89	0.00
C_2H_4	26.03	−0.55	25.48	25.46	+0.02
C_2H_2	30.02	−0.72	28.30	29.32	−0.02
CH_3F	27.31	−0.67	26.64	26.64	0.00
CH_3OH	28.05	−0.71	27.34	27.35	−0.01
CH_3NH_2	29.11	−0.76	28.35	28.31	+0.04
CH_3NH_2	31.21	−0.76	30.45	30.47	−0.02
$CH_3C{\equiv}N$	29.81	−0.64	29.17	29.22	−0.05
$H_2C{=}C{=}CH_2$	26.75	−0.44	26.30	26.28	+0.02
$(CH_3)_2O$	28.26	−0.74	27.54	27.54	0.00
H_2O	30.69	−0.57	30.12	30.10	+0.02
NH_3	31.43	−0.70	30.73	30.73	0.00
					MAE: 0.016
					slope: 0.995
CH_4	198.8	−4.13	194.7	195.0	−0.3
C_2H_6	185.7	−5.08	180.6	180.8	−0.2
C_2H_4	69.6	−5.70	63.9	64.4	−0.5
C_2H_2	122.6	−5.29	117.3	117.1	+0.2
CH_3F	122.4	−5.13	117.3	116.7	+0.6
CH_3OH	141.8	−5.01	136.8	136.5	+0.3
CH_3NH_2	163.3	−5.03	158.3	158.2	+0.1
$CH_3C{\equiv}N$	191.7	−3.61	188.1	187.6	+0.5
$CH_3C{\equiv}N$	76.2	−1.86	74.3	73.7	+0.6
CH_3CHO	161.2	−4.53	156.7	157.1	−0.4
CH_3CHO	−3.3	−3.54	−6.9	−6.8	−0.1
$(CH_3)_2CO$	163.0	−4.89	158.1	157.9	+0.2
$(CH_3)_2CO$	−10.7	−2.01	−12.7	−13.2	+0.5
$C{\equiv}O$	2.6	−2.46	0.1	0.9	−0.8
$O{=}C{=}O$	60.2	−1.95	58.2	58.7	−0.5
$H{-}C{\equiv}N$	84.9	−2.35	82.6	82.0	+0.6
$H_2C{=}C{=}CH_2$	118.9	−3.23	115.7	115.1	+0.6
$H_2C{=}C{=}CH_2$	−26.2	−2.91	−29.1	−29.4	+0.3
CF_4	64.7	−1.78	62.9	64.4	−1.5
					MAE: 0.46
					slope: 1.000

[a]Experimental absolute shieldings were taken from ref 12 (proton) and ref 5 (carbon). Uncertainties are ca. ±0.01 ppm for proton and ±0.1 ppm for carbon. Proton stationary shieldings: CCSD(T)/pcS-3/GIAO//CCSD(T)/cc-pVQZ. Carbon stationary shieldings: literature (ref 5) CCSD(T)/CBS//CCSD(T)/cc-pVTZ (for the CBS procedure, see the Supporting Information). QCD corrections: B3LYP/MIDI! surface, B3LYP/pcS-3/GIAO NMR, 250 fs, 200 trajectories per molecule, 298 K. "Slope" refers to the slope of the regression line of predicted values vs experimental values.

optimized combination of B3LYP/MIDI! dynamics and B3LYP/cc-pVDZ NMR points. To generate chemical shift predictions, we used PBE0/cc-pVTZ/PCM stationary shieldings and geometries and the following expression:

vibration-corrected chemical shift prediction for **natural product (NP)** = experimental chemical shift of **reference** +

[stationary point shielding of **reference** (PBE0/cc-pVTZ/PCM geometry and NMR) + dynamically averaged shielding of **reference** (B3LYP/MIDI! dynamics and B3LYP/cc-pVDZ NMR) − stationary point shielding of **reference** (B3LYP/cc-pVDZ NMR and B3LYP/MIDI! geometry)] −

raw correction (reference)

[stationary point shielding of **NP** (PBE0/cc-pVTZ/PCM geometry and NMR) + dynamically averaged shielding of **NP** (B3LYP/MIDI! dynamics and B3LYP/cc-pVDZ NMR) − stationary point shielding of **NP** (B3LYP/cc-pVDZ NMR and B3LYP/MIDI! geometry)]

raw correction (NP)

We call the difference between the raw corrections the "actual vibrational correction". The above expression allows an arbitrary compound to be chosen as the chemical shift reference. We selected the methyl group in methanol as the reference for all proton and carbon nuclei, which has been previously found to be advantageous.[16]

Our results show that the addition of quasiclassical corrections to stationary shifts significantly improves the prediction accuracy (Figure 2a). We obtained proton and carbon MAEs of 0.08 and 1.48 ppm, respectively, both with near unit slope and zero intercept. The elimination of the slope and intercept in the error distribution is significant because current methods require empirical scaling for accuracy.[17,18] That is, they apply slope and intercept factors derived from training data sets to stationary predictions for unknown molecules. When a dozen such methods were applied here (Figure 2b), only the best performed as well as unscaled QCD (Table 3). The reason for the variable performance of literature methods becomes clear when our actual dynamic corrections are plotted against literature scaling corrections (Figure 2c). The significant correlation indicates that linear scaling is largely an implicit correction for dynamic errors, with the residuals representing remaining errors from other sources such as

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

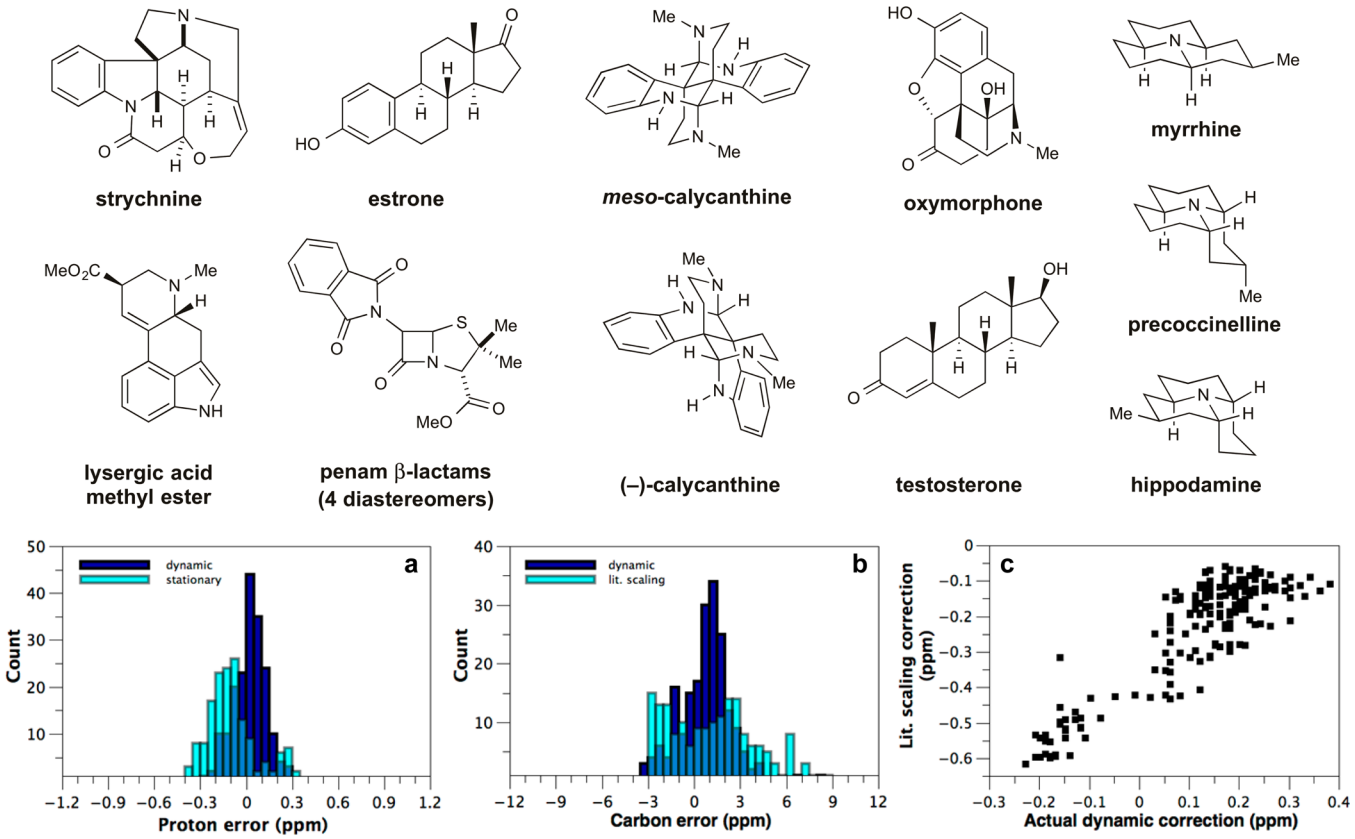


Figure 2. (top) Natural products data set (187 unique proton chemical shifts and 195 unique carbon chemical shifts after averaging; exchangeable protons ignored). (bottom) Results for this data set: (a) Dynamically corrected vs stationary-point residual errors for proton predictions. Dynamic corrections are a significant improvement over the stationary predictions. (b) Dynamically corrected vs literature scaling residuals for carbon predictions (see Cramer method 5 in Table 3). Dynamic corrections can be significantly superior to literature scaling methods. (c) Correlation between literature scaling corrections (see Tantillo method 1 in Table 3) and dynamic corrections. Linear scaling is largely a correction for dynamic effects.

Table 3. Dynamics versus Literature Scaling for Natural Products[a]

	proton (ppm)			carbon (ppm)		
method	MAE	slope	intercept	MAE	slope	intercept
QCD	0.08	0.999	+0.03	1.48	1.007	+0.18
Tantillo[2] method 1	0.09	0.983	−0.01	1.38	0.994	+0.51
Tantillo method 2	0.27	0.997	+0.27	2.83	0.993	+2.48
Bally[17] method 3	0.12	0.982	−0.05	no scaling factors provided		
Bally method 4	0.16	0.957	0.00	no scaling factors provided		
Cramer[18] method 5	0.14	1.057	−0.22	2.74	1.005	−1.12

[a]QCD = unscaled quasiclassical dynamics on the B3LYP/MIDI! surface with B3LYP/cc-pVDZ NMR points, the PBE0/cc-pVTZ/PCM stationary geometry, and high-level NMR shieldings, referenced to methanol. The literature methods used DFT and reported scaling factors (see Table S11 for method details). Experimental uncertainties are ca. 0.01 ppm (proton) and 0.20 ppm (carbon).

incomplete electron correlation or solvation. This suggests that that the performance of scaling methods is sensitive to the match in dynamic effects between the training set and the molecules being considered, explaining the variability in Table 3.

The full benefit of QCD is also partially hidden by noise in the natural products data set. When the natural products stationary DFT method (PBE0/cc-pVTZ) is used for the gas-phase data sets with direct scaling to experiment, the proton and carbon MAEs are reduced only to 0.095 and 2.00 ppm, respectively, compared to 0.016 and 0.46 ppm for unscaled quasiclassical predictions (Tables S7 and S8). The situation is not improved by the use of other DFT methods or even CCSD(T) shieldings (Figure 3), which suggests that dynamic errors have a significant nonlinear component and are better treated with QCD. Together, these findings show that dynamic effects are an important source of error in conventional predictions requiring the use of scaling.

3.4. Unusual Dynamics. Empirical scaling approaches depend on the fact that thermal motions are usually a small perturbation with respect to the shielding surface of typical molecules. However, linear approximations break down for molecules with unusual dynamics or pathological shielding surfaces. For example, all stationary predictions are at odds with experiment for the iconic molecule [18]annulene.[19,20] The shifts of the internal and external protons of its aromatic D_{6h} geometry (Figure 4a) were predicted to be approximately −10 and +11 ppm, respectively, in stark contrast to the experimental values of −3.0 and +9.3 (Figure 4b and Table S15). Schleyer and co-workers have used this inconsistency to argue that [18]annulene cannot possess D_{6h} symmetry.[21] Their alternative proposal of a bond-alternant structure with C_2 symmetry is controversial because other experimental evidence is ambiguous: X-ray crystallographic analysis[22] shows an essentially D_{6h}

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

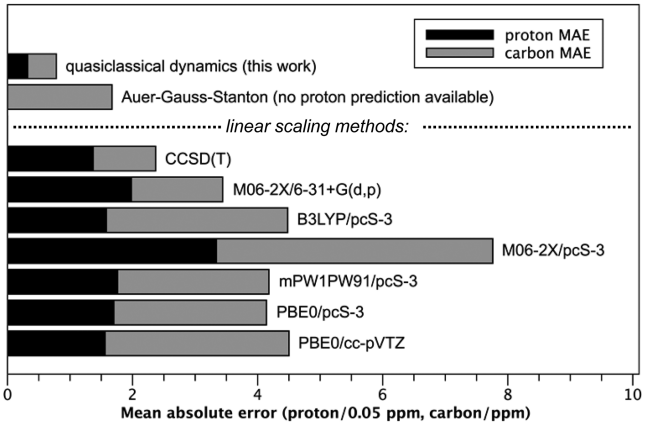
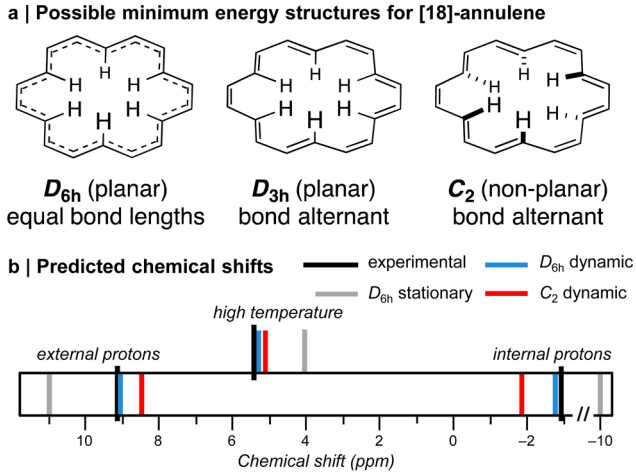


Figure 3. Dynamics vs linear scaling for absolute shielding predictions. Proton errors are shown at 20 times their actual size to account for the different sizes of the chemical shift scales. See Tables S7 and S8 for method descriptions. Methods above the dotted line are unscaled; methods below are linearly scaled to experiment.

structure with little bond alternation,[23] but [18]annulene can undergo addition chemistry that is more typical of olefins.[19]

[18]Annulene displays anomalous behavior because its low-frequency vibrational modes generate a large degree of out-of-plane motion that breaks aromaticity.[24] As a result, its chemical shifts are very sensitive to thermal motions and have a pronounced temperature dependence.[25,26] While these properties challenge existing scaling methods, they pose no problem for QCD, which requires only an accurate energy surface. In this case, B3LYP is unsuitable because of overdelocalization.[32] To identify a more appropriate surface, we generated 75 random quasiclassical initializations for each symmetry (D_{6h}, D_{3h}, and C_2) at 298 K and compared the single-point energies of various DFT methods to those obtained with DLPNO−CCSD(T).[27] The initializations were restricted to excitations in modes below 1300 cm^{-1}, so that the accuracy in these key out-of-plane modes could be measured without interference from high-energy displacements in stiffer modes, for which most DFT methods perform well. The results showed that M06-2X uniformly outperforms other methods such as KMLYP and BHandHLYP (Table S12).

We then used quasiclassical trajectories on the M06-2X/cc-pVDZ surface to correct the stationary-point chemical shifts of D_{6h}-, D_{3h}-, and C_2-symmetric [18]annulene. The stationary D_{6h} geometry was determined by minimizing single-point DLPNO−CCSD(T)/cc-pVTZ energies with a central composite experimental design and closely matches the X-ray structure (Table S14). For the lower-symmetry D_{3h} and C_2 structures, such an approach was impractical, so conventional DFT methods were used. The predicted chemical shifts for the D_{6h} structure at the B3LYP/jul-cc-pVTZ level[28] are in excellent agreement with experiment at both low and high temperatures (Figure 4c). In contrast, the C_2 structure does not match experiment regardless of the chosen stationary geometry (Table S18). Although the predicted D_{6h} shifts are closer to experiment than are the D_{3h} shifts, the possibility that the two structures may rapidly interconvert cannot be discounted.[26] These results demonstrate that quasiclassical corrections can treat unusual systems that are otherwise beyond the reach of conventional methods.



	Proton Shift (ppm)			Carbon Shifts (ppm)		
Symmetry	External	Internal	High Temp.	External	Internal	High Temp.
experimental	9.17	−2.96	5.45	128	121	126
D_{6h} (stationary)	*ca.* 11	−10	4	125	114	121
D_{6h} (dynamic)	**9.13**	**−2.86**	**5.27**	**129.6**	**121.1**	**127.6**
D_{3h} (dynamic)	8.81	−2.66	4.98	125.1	116.9	123.0
C_2 (dynamic)	8.52	−1.90	5.14	125.7	118.0	123.5

Experimental shifts (in tetrahydrofuran, ref. 21): 203 K (external and internal entries) and 393 K (high temp. entries) for protons, 213 K (external and internal entries) and 333 K (high temp. entries). D_{6h} stationary: various DFT methods using literature scaling (Table S15). Approximate values are given. Dynamic entries: M06-2X/cc-pVDZ surface, B3LYP/cc-pVDZ NMR points, and B3LYP/jul-cc-pVTZ stationary NMR (benzene reference), with cc-pVTZ stationary geometries (D_{6h}, CCSD(T); D_{3h}, BHandHLYP; C_2, KMLYP).

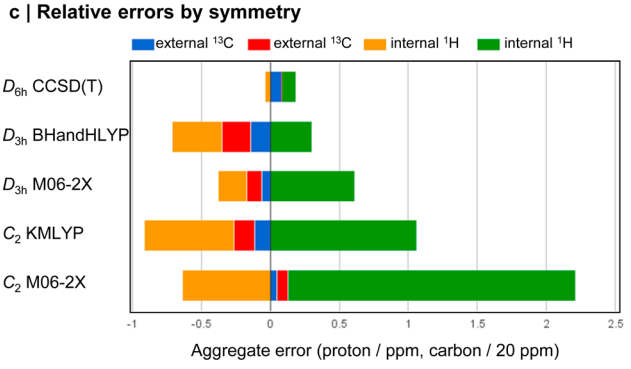


Figure 4. Quasiclassical corrections are essential for the prediction of [18]annulene chemical shifts. (a) Possible minimum-energy structures for [18]annulene. (b) Linearly scaled stationary predictions are 7 ppm in error for the internal protons, while dynamic predictions agree closely at both low and high temperatures. (c) The results indicate that D_{6h} symmetry gives the best agreement with experiment (see Table S18 for more non-D_{6h} geometries, all of which agree poorly with experiment).

4. CONCLUSION

Ultimately, the quantitative prediction of NMR shieldings for arbitrary organic molecules requires high accuracy in all aspects of the underlying physics. Our efforts take a step in this direction by demonstrating the accuracy of QCD in the explicit treatment of rovibrational effects. The above results suggest that the influence of electron correlation, basis set, and other effects should now be re-evaluated for many systems with the inclusion of QCD. At the moment, our method is the best available for both gas-phase and solution-phase predictions and extends the scope of predictions from typical molecules to any

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089

molecule for which an accurate energy surface is available. Because our approach is computationally efficient, we anticipate that it will find immediate application in large systems that were previously only tractable with linear scaling while providing a compelling depiction of motions at an atomic level.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00856.

Computational procedures, additional experiments, tabulated shielding calculation results, and geometries and energies of key structures (PDF)

Animations of selected trajectories in MP4 format (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ekwan@fas.harvard.edu.

Notes

The authors declare the following competing financial interest(s): A provisional patent application has been filed on this methodology (U.S. PTO 62/220,340).

■ ACKNOWLEDGMENTS

We are grateful to Professor Daniel A. Singleton (Texas A&M) for inspiration and performing preliminary calculations. We thank Professors E. J. Corey, David A. Evans, and Eric N. Jacobsen (Harvard), Dr. Gregory R. Maloney (Newcastle), and Professor William F. Reynolds (Toronto) for their support. Calculations were carried out on the Odyssey Cluster at Harvard University. We recognize Dr. Paul Edmon, Dr. Bob Freeman, Mr. Bradford Freeman, Dr. Aaron Kitzmiller, and Dr. Scott Yockel of the Research Computing Team at Harvard University for generous technical assistance.

■ REFERENCES

(1) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807.

(2) Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. *Chem. Rev.* **2012**, *112*, 1839−1862.

(3) Willoughby, P. H.; Jansma, M. J.; Hoye, T. R. *Nat. Protoc.* **2014**, *9*, 643−660.

(4) (a) Karplus, M.; Porter, R. N.; Sharma, R. D. *J. Chem. Phys.* **1965**, *43*, 3259−3287. (b) Bogle, X. S.; Singleton, D. A. *J. Am. Chem. Soc.* **2011**, *133*, 17172−17175. (c) Carpenter, B. K. *Acc. Chem. Res.* **1992**, *25*, 520−528. (d) Bowman, J. M.; Schatz, G. C. *Annu. Rev. Phys. Chem.* **1995**, *46*, 169−195. (e) Collins, M. A. *Theor. Chem. Acc.* **2002**, *108*, 313−324. (f) Fernandez-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 4518−4584.

(5) Auer, A. A.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2003**, *118*, 10407−10417.

(6) Ruud, K.; Åstrand, P.−O.; Taylor, P. R. *J. Am. Chem. Soc.* **2001**, *123*, 4826−4833.

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

(8) Harding, M. E.; Gauss, J.; Schleyer, P.v.R. *J. Phys. Chem. A* **2011**, *115*, 2340−2344.

(9) Jiménez-Osés, G.; García, J. I.; Corzana, F.; Elguero, J. *Org. Lett.* **2011**, *13*, 2528−2531.

(10) Li, D.-W.; Brüschweiler, R. *J. Phys. Chem. Lett.* **2010**, *1*, 246−248.

(11) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta.* **1996**, *93*, 281−301.

(12) Garbacz, P.; Jackowski, K.; Makulski, W.; Wasylishen, R. E. *J. Phys. Chem. A* **2012**, *116*, 11896−11904.

(13) (a) Chesnut, D. B. *Chem. Phys.* **1997**, *214*, 73−79. (b) Fukui, H.; Miura, K.; Matsuda, H. *J. Chem. Phys.* **1992**, *96*, 2039−2043.

(14) Vaara, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5399.

(15) Smith, S. G.; Goodman, J. M. *J. Am. Chem. Soc.* **2010**, *132*, 12946−12959.

(16) Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* **2009**, *74*, 7254−7260.

(17) Jain, R.; Bally, T.; Rablen, P. R. *J. Org. Chem.* **2009**, *74*, 4017−4023.

(18) Wiitala, K. W.; Hoye, T. R.; Cramer, C. J. *J. Chem. Theory Comput.* **2006**, *2*, 1085−1092.

(19) Spitler, E. L.; Johnson, C. A., III; Haley, M. M. *Chem. Rev.* **2006**, *106*, 5344−5386.

(20) Baldridge, K. K.; Siegel, J. S. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 745−748.

(21) Wannere, C. S.; Sattelmeyer, K. W.; Schaefer, H. F., III; Schleyer, P. v. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 4200−4206.

(22) Gorter, S.; Rutten-Keulemans, E.; Krever, M.; Romers, C.; Cruickshank, D. W. J. *Acta Crystallogr., Sect. B: Struct. Sci.* **1995**, *51*, 1036−1045.

(23) Ermer, O. *Helv. Chim. Acta* **2005**, *88*, 2262−2266.

(24) Baldridge, K. K.; Siegel, J. S. *J. Phys. Chem. A* **1999**, *103*, 4038−4042.

(25) Stevenson, C. D.; Kurth, T. L. *J. Am. Chem. Soc.* **2000**, *122*, 722−723.

(26) Hudson, B. S.; Allis, D. G. *J. Mol. Struct.* **2012**, *1023*, 212−215.

(27) Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.

(28) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.

(29) Nicolaou, K. C.; Snyder, S. A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1012.

(30) In principle, QCD can be applied to conformationally mobile molecules through Boltzmann averaging. However, vibrational corrections may not be constant between conformations, and we did not investigate those effects in this study.

(31) Rovibrational corrections can be efficiently calculated using any modern computing cluster. For strychnine, a typical natural product containing 25 heavy atoms, application of Tantillo method 1 (Table 3) required approximately 17 h using one computer (four cores) and scaled poorly with increased parallelization. In contrast, the rovibrational correction can be calculated in roughly the same time using 10 computers by propagating trajectories and computing NMR points in parallel. Although this comparison emphasizes wall-clock time over CPU time, we feel that it appropriately illustrates the practical nature of these calculations.

(32) (a) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289. (b) Yu, L.-J.; Karton, A. *Chem. Phys.* **2014**, *441*, 166.

DOI: 10.1021/acs.jctc.5b00856
J. Chem. Theory Comput. 2015, 11, 5083−5089