JOC The Journal of Organic Chemistry

Article

pubs.acs.org/joc

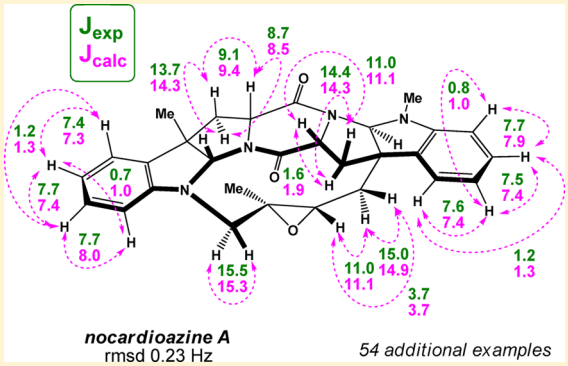
Minimalist Relativistic Force Field: Prediction of Proton−Proton Coupling Constants in ^{1}H NMR Spectra Is Perfected with NBO Hybridization Parameters

Andrei G. Kutateladze* and Olga A. Mukhina

Department of Chemistry and Biochemistry, University of Denver, Denver, Colorado 80208, United States

Supporting Information

ABSTRACT: We previously developed a reliable method for multiparametric scaling of Fermi contacts to achieve fast and accurate prediction of proton−proton spin−spin coupling constants (SSCC) in ^{1}H NMR. We now report that utilization of NBO hybridization coefficients for carbon atoms in the involved C−H bonds allows for a significant simplification of this parametric scheme, requiring only four general types of SSCCs: geminal, vicinal, 1,3-, and long-range constants. The method is optimized for inexpensive B3LYP/6-31G(d) molecular geometries. A new DU8 basis set, based on a training set of 475 experimental spin−spin coupling constants, is developed for hydrogen and common non-hydrogen atoms (Li, B, C, N, O, F, Si, P, S, Cl, Se, Br, I) to calculate Fermi contacts. On a test set of 919 SSCCs from a diverse collection of natural products and complex synthetic molecules the method gave excellent accuracy of 0.29 Hz (rmsd) with the maximum unsigned error not exceeding 1 Hz.



INTRODUCTION

Fast and accurate computations of the proton−proton spin−spin coupling constants (SSCC) for predicting NMR spectra have been challenging, yet there has been growing consensus that the Fermi contact (FC) mechanism dominates nuclear spin scalar couplings.[1] As FCs are easy to compute, Bally and Rablen[2] have developed a simple single parameter linear scaling of Fermi contacts, which can be used to get reasonably accurate estimates of SSCCs. Recently, we have proposed that parametric scaling of Fermi contacts should be based on a more detailed selection and individual parametrization of a representative set of SSCC "types" defined by connectivity and hybridization (for example, geminal sp^2, or vicinal sp^2−sp^3, etc.), i.e., not unlike the parametrization (force fields) in molecular mechanics.[3] We termed this approach relativistic force field. It was shown that with a modest Linux cluster, NMR spectra can be predicted within 1 h for the majority of organic molecules with accuracy fully adequate for an unambiguous stereochemical assignment in most cases.

Further analysis indicated that the scaling of FCs in various representative SSCC types is generally affected by the hybridization state[4] of the carbon atoms bearing the involved hydrogens. This observation is not unprecedented: a paper by Weinhold, Markley, and co-workers[5] on natural *J*-coupling analysis, i.e., the interpretation of scalar *J*-couplings in terms of natural bond orbitals (NBOs), provided additional understanding of the nuances of applying Weinhold's NBO analysis[6] to the evaluation of Fermi contacts. We hypothesized that the range of SSCC types in the relativistic force field parametrization scheme can be limited to ^2J, ^3J, ^4J, and ^NJ (i.e., geminal, vicinal, 1,3-, and other long-distance) if the parametric equation for each of the four types were to include corrections for the hybridization state of the carbon atoms in the involved C−H NBOs. In this paper, we report the development of such streamlined relativistic force field for parametrization of Fermi contacts, which involves only four general SSCC types scaled using both empirical scaling parameters and the NBO hybridization coefficients. All computations are carried out at a minimalist level of the density functional theory (DFT) to accelerate NMR predictions without compromising the accuracy.

RESULTS AND DISCUSSION

Parametric Scheme and Training Set. As in the previous work,[3] our training set of experimental proton spin−spin coupling constants was assembled using data available in the literature and the spectral database of organic compounds by AIST, Japan.[7] Optimization of the basis sets for hydrogen and non-hydrogen (Li, B, C, N, O, F, Si, P, S, Cl, Se, Br, I) atoms was carried out in an iterative procedure to optimize the accuracy of SSCC computations. Each of the four SSCC types was calculated by scaling of the computed FCs via the following parametric equation

Received: March 18, 2015
Published: April 17, 2015

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225

$$J = c_1HF + c_2F + c_3H + c_4 \quad (1)$$

where $H = h_A\, h_B$, i.e., the product of the (p- to s-) hybridization ratios for the involved carbon orbitals, computed with the NBO program[8] as implemented in Gaussian 09 computational package; F is the computed Fermi contact; and c_1−c_4 are empirical coefficients optimized via multivariate regression analysis. In addition to the hybridization p-coefficients, h_A and h_B, we initially considered including the NBO second-order perturbation energy, $E(2)$, into the multivariate regression analysis as a variable for scaling of FCs. The second-order perturbation energy was central to Weinhold's discussion of "how J-coupling, or rather the transfer of spin density, is related to spin hyperconjugative delocalization (by means of second order perturbation analysis)...".[5] However, we found little evidence for improvement of FCs scaling with $E(2)$ inclusion: the Student's t test values for the respective coefficient in the multivariate expression were very low (<1). This, in fact, made perfect sense because, according to Weinhold, Fermi contacts are dominated by these interactions whereas our task was to identify the NBO elements which do not dominate the FCs but rather help correct for other (missing) contributors to SSCCs, such as dia- and paramagnetic components of spin−orbit coupling.[9]

The optimization employing the four-parameter multivariate expression (eq 1) for each of the four types was carried out on the training set of 475 experimental constants until the accuracy (rmsd) of the calculations reached 0.2 Hz. The optimized basis set, which we term DU8, is described in the Supporting Information. The objective of this work was to develop a *fast* and reliable method for SSCC computations. As a starting point in the basis set optimization for heavy atoms we used a very inexpensive 4-21G basis set for carbons; 3-21G* for Si, S, P, Cl, Se, Br, and I; and finally, 6-31G(d) for Li, Be, B, N, O, and F. The use of a light basis set resulted in a considerable acceleration of Fermi contact calculations while maintaining the accuracy of their scaling to model SSCCs. For example, calculations of Fermi contacts for a large molecule such as ginkgolide B[10] ($C_{20}H_{24}O_{10}$, for the structure see Figure 3) took less than 15 min on a single 16-core Linux node resulting in an rmsd of 0.17 Hz for the set of six available experimental constants.

Table 1 summarizes the optimized c_1−c_4 scaling coefficients for the four SSCC types. Because this parametrization approach is inherently associated with the optimized basis set, we use the term DU8 interchangeably as a reference to both the basis set and the entire method of computing the SSCCs.

This 13-parameter relativistic force field (rff) performs very well as we will demonstrate below using an extensive test set of polycyclic natural and synthetic products. Experimental 1D proton NMR spectra of nine natural products from our original test set[3] were augmented in this study with an additional 55 spectra of complex natural products for which accurate experimental SSCCs were published. We have requested and received raw digital NMR data files for 14 of these structures directly from the colleagues who reported their synthesis or isolation. This allowed for careful line fitting to extract accurate spin−spin coupling constants.[11] Eleven in-house-synthesized polyheterocycles were added to ensure that each of the four SSCC types is represented by a large subset of experimental constants, bringing the total number of experimental test SSCCs to 919. The range of the magnitudes of these SSCCs in the experimental test set is also of importance. The average value for a constant in the test set was 7.3 Hz, with the size distribution histogram shown in Figure 1. With the B3LYP/6-31G(d) molecular structures, DU8 basis set, and the parametric coefficients listed in Table 1 we obtained promising results, with an rms value of 0.285 Hz (standard deviation 0.283 Hz). All deviations from the experimental data were within the interval of −0.83 to +0.98 Hz. Figure 2 shows a tight Gaussian-like distribution of the unsigned errors, attesting to the high fidelity of the method.

Table 1. Parametric Coefficients c_1−c_4 for Each of the Four Types of SSCCs

Type	c_1 *(F H)*	c_2 *(F)*	c_3 *(H)*	c_4 *(const)*
H–C–H ^2J (geminal)	-0.020	1.676	-0.151	0.300
H–C–C–H ^3J (vicinal)	-0.004	1.550	0.030	-0.204
H–C–X–C–H ^4J (1,3-)	–[a]	1.505	-0.027	0.381
^NJ (other long range)	–[a]	0.95	–[a]	0.2

[a]Terms weighted by c_1 (for ^4J and ^NJ) and c_3 (for ^NJ) were eliminated from the parametric scaling scheme due to the low Student's t test values.

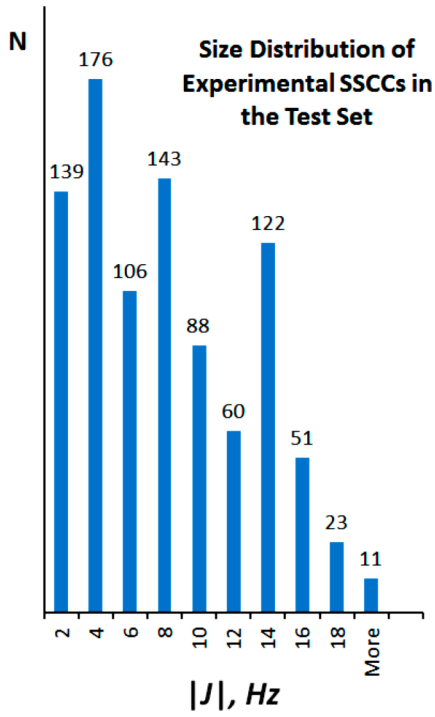


Figure 1. Size distribution of 919 experimental SSCCs in the test set.

Test Set Performance. Table 2 compares the performance of our first-generation multiparametric relativistic force field (with DU4 basis) and the second-generation rff described in this work, i.e., the streamlined 13-parameters DU8 method. It is clear that the second-generation method performs well if not better than the previous one while using a smaller set of parameters for only four parametric types. This is reassuring as it implies that the hybridization correction is likely to be a

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225

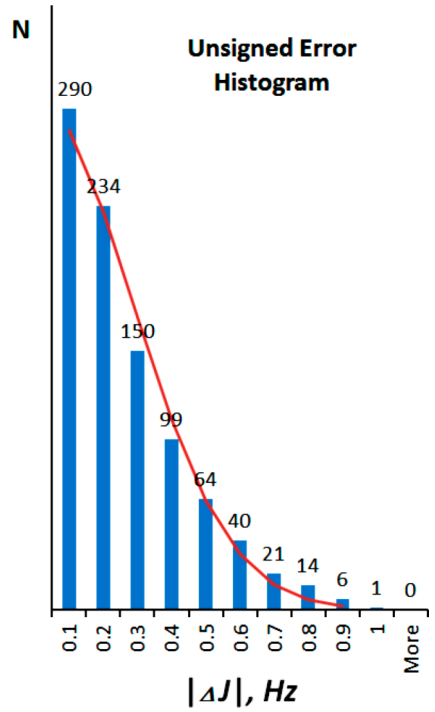


Figure 2. Unsigned error histogram for the test set. Red line shows a Gaussian distribution with σ = 0.285 Hz for comparison.

Table 2. Comparison of DU4[3] and DU8 (This Work) for a Subset of Nine Natural Products

	this work (DU8 basis)		ref 3 (DU4 basis)			
	rmsd	mue	rmsd	mue	N^a	ref[b]
ginkgolide B	0.17	0.22	0.21	0.34	6	10
oxachamigrene	0.22	0.48	0.27	0.44	16	12
chloranthalactone A	0.32	0.56	0.32	0.67	10	13
strychnine	0.16	0.41	0.19	0.48	27	14
aquatolide	0.28	0.59	0.23	0.52	9	15
perforatol	0.27	0.46	0.29	0.50	8	16
isopalhinine A	0.24	0.53	0.24	0.44	27	17
morphine	0.19	0.43	0.24	0.51	16	18
dichomitol[c]	0.23	0.40	0.24	0.57	17	19

[a]Number of experimental SSCCs. [b]References for experimental SSCCs. [c]rmsd and mue for dichomitol are improved due to the revision of one experimental *syn*-vicinal cyclobutyl constant based on second-rder simulation of the spectrum provided by Wei.

general approach to improve the accuracy of scaling of the Fermi contacts without the need for numerous parametric corrections for a variety of new structural and hybridization subtypes.

Table 3 shows the rmsd's and maximum errors for an additional 55 natural and synthetic products in the test set. For details, refer to the Supporting Information.

Conformationally Flexible Systems, J-Fitting. The treatment of conformationally fluxional systems, where an ensemble of rapidly interconverting conformers contribute to the NMR spectrum, normally involves conformer averaging based on their relative computed energies. This approach is generally applicable to molecules weakly interacting with solvents and also special cases where conformational changes do not cause significant changes in SSCCs. However, the accurate relative energies, obtained with high levels of theory, ZPE corrections, and various PCM models, exact heavy toll on the computational time while often delivering very little help in obtaining accurate prediction of SSCCs, especially in the cases of intramolecular hydrogen bonds competing with H-bonding to solvent. An error of 1 kcal/mol in relative energy calculations may alter the conformer ratio of 70:30 into 30:70. We assert that a practical compromise is not to waste time on the high end energy computations but rather fit the experimental spectra by mixing the computed SSCCs of individual conformers to reproduce experimental NMR spectra (*J*-fitting) and subsequently compare the weighting coefficients to those obtained with the inexpensive DFT energies. Only egregious energy outliers should be subjected to additional scrutiny. In our experience, such *J*-fitting of two or more potential candidate structures, each represented by several conformers to reproduce experimental spectra, in most cases resulted in the identification of the actual structure based both on the overall quality of prediction as expressed by the rmsd value and on the analysis of individual mismatched constants.

Table 4 illustrates this point with the examples of complex natural products which are well described by only two equilibrating conformers. Two conformers of tetrapetalone A-Me aglycon, first entry in the table (for structure see Figure 3), are approximately 0.8 kcal/mol apart by B3LYP/6-31G(d), which corresponds to a 78:22 ratio resulting in a 0.24 Hz rmsd match of SSCCs with experimental spectrum. Direct *J*-fitting of the calculated spectrum requires a 85:15 ratio and gives a slightly better value of rmsd (0.23 Hz). The fitted ratio corresponds to a 1.0 kcal/mol conformational energy difference between the two conformers, which is within a small $\Delta\Delta E \sim$ 0.2 kcal/mol energy adjustment required to reconcile the two approaches. Clearly, this is the case where both schemes for weighing the conformers give very similar results.

Nonetheless, this energy "discrepancy," $\Delta\Delta E$, increases down the table, reaching >2 kcal/mol for hawaiinolide A (the proton NMR spectrum is taken in acetone-d_6). When the conformers were mixed using *J*-fitting we obtained an excellent rmsd of 0.22 Hz for the nearly 1:1 (i.e., energy degenerate) conformers content. In contrast, the computed DFT energy difference of 2.1 kcal/mol corresponded to a 97:3 conformer ratio and resulted in a poor rmsd value exceeding 1 Hz, with an unacceptably high maximum deviation of 3.3 Hz. As there are spin−spin coupling constants which change considerably in this conformational equilibrium, we are confident that the mixing of accurately computed SSCCs gives a more realistic picture of the ratio of conformers in acetone, where hydrogen bonding of hawaiinolide's hydroxyl group to the solvent could be reasonably expected to change the computed gas-phase energetics by a few kcal/mol.

Error associated with the *J*-fitting procedure depends on the dynamic range of SSCCs. The error increases for the conformers with very similar SSCCs which do not vary much as a result of conformational change. We assert that for systems exhibiting reasonably large changes in SSCCs, the *J*-fitting procedure could offer significantly more accurate indirect energy estimates for conformers equilibrating in solutions, than any of the existing high-end direct energy computations.

Relative Conformer Ratios by Higher Level of DFT Theory. The relative content of equilibrating conformers is an important general issue because conformational flexibility manifests itself even in the cases of rigid polycyclic structures. One instructive example of this is penicilleremophilane B, a sesquiterpene from the soil fungus *Penicillium copticola* reported

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225

Table 3. Additional Natural Products in the Test Set

compd	RMSD[a]	MUE[a]	N^b	ref[c]	compd	RMSD[a]	MUE[a]	N^b	ref[c]
affinisine oxindole	0.41	0.81	15	20	oxocrinine	0.39	0.79	11	21
alstoscholarisine A	0.29	0.58	11	22	oxycodone	0.47	0.53	9	23
alstoscholarisine E	0.17	0.35	10	22	pannosane	0.27	0.48	5	24
Banwell–cmpd 24[d]	0.36	0.86	25	25	penifulvin A[d]	0.23	0.52	12	26
bielschowskysin	0.32	0.56	14	27	physangulidine D	0.35	0.69	26	28
botryosphaerin F	0.39	0.56	7	29	physangulidine G	0.43	0.75	9	28
cornolactone B[d]	0.17	0.35	10	30	pyrenolide D	0.24	0.47	7	31
cyanosporaside A	0.31	0.56	5	32	quassinoid compd **10**	0.40	0.63	9	33
dehydroleuconoxine[d]	0.41	0.82	25	34	quassinoid compd **6**	0.23	0.31	3	33
echinopine A	0.29	0.54	15	35	salvileucalin D[d]	0.24	0.56	10	36
eurifoloid C	0.35	0.70	6	37	salvipuberulin[d]	0.37	0.98	9	36
eurifoloid F	0.35	0.83	9	37	securinine compd **20**	0.34	0.58	8	38
eurifoloid M	0.31	0.61	10	37	Shea compd **33**[d]	0.23	0.59	7	39
ganodermalactone G	0.15	0.27	8	40	Sieburth compd **18**	0.18	0.41	13	41
gelsemine	0.16	0.27	15	42	sorazinone A	0.24	0.32	4	43
hawaiinolide A	0.22	0.50	10	44	sinensilactam A	0.41	0.91	11	45
hawaiinolide B	0.39	0.79	9	44	sporol	0.26	0.49	4	46
hexacyclinol	0.36	0.79	9	47	strepsesquitriol	0.11	0.15	9	48
isosalvipuberulin	0.13	0.34	11	36	striatoid B[d]	0.17	0.39	11	49
isoschizogamine	0.32	0.73	26	50	striatoid C[d]	0.29	0.61	14	49
lepistine (first rotamer)[d]	0.29	0.41	6	51	striatoid E[d]	0.26	0.48	13	49
leuconodine B	0.33	0.81	15	52	talaperoxide C	0.28	0.45	16	53
lundurine B	0.37	0.66	20	54	tetrapetalone A Me aglicon[d]	0.24	0.42	8	55
maoecrystal V	0.34	0.71	6	56	trichocladinol D	0.28	0.57	6	57
N,O-dimethyloxostephine[d]	0.19	0.46	7	58	trichocladinol E	0.35	0.58	6	57
nardoaristolone B	0.20	0.22	7	59	tronoharine[d]	0.31	0.99	29	60
nocardioazine A	0.23	0.55	20	61	wenyujinin H[d]	0.30	0.61	12	62
nominine	0.21	0.40	16	63					

[a]Values are in hertz. [b]No. of experimental SSCCs. [c]References for the experimental SSCC data. [d]Digital NMR data files are provided by authors.

Table 4. Two-Conformer Mixing by *J* Fit vs DFT Energy

	mixing	% ratio	RMSD	ΔE	$\Delta\Delta E$
tetrapetalone A-Me aglycon	*J* fit	85:15	0.23	1.0	0.2
	DFT	78:22	0.24	0.8	
alstoscholarisine A	*J* fit	77:23	0.29	0.7	0.5
	DFT	57:43	0.36	0.2	
alstoscholarisine E	*J* fit	91:9	0.17	1.4	0.9
	DFT	70:30	0.52	0.5	
hawaiinolide A	*J* fit	46:54	0.22	0.1	2.0
	DFT	97:3	1.06	2.1	

by the Rukachaisirikul group[64] (see Figure 3). Even in its octahydronaphthalene core, which is additionally rigidified by a two atom oxamethylene bridge, the chair conformation of the ring A is in equilibrium with two contributing twisted boat conformations. Furthermore, the rotatable isopropenyl group adds three conformations resulting in a total of 3 × 3 = 9 conformers. Normally we identify most stable conformers by force field (Sybyl) energy estimate and limit their number to a few which lie within ∼5 kcal/mol of the lowest energy conformer. In this particular example, for completeness we have performed geometry optimizations at two levels of theory:

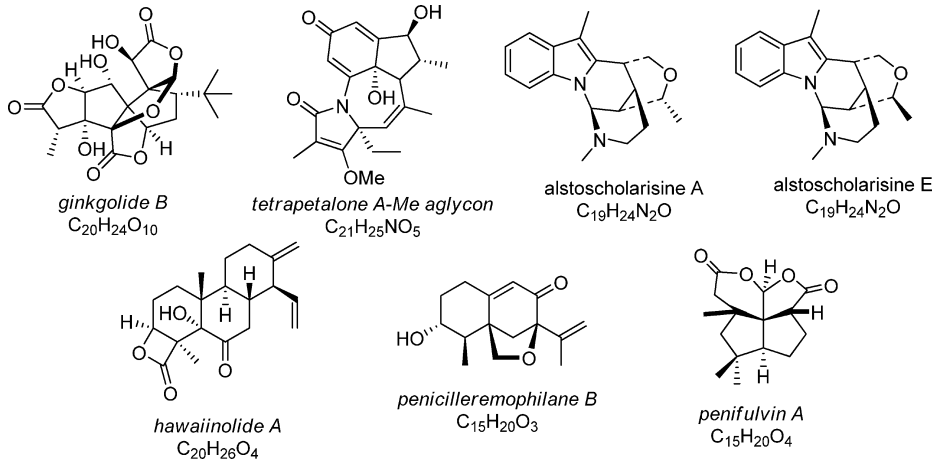


Figure 3. Structures of selected natural products (see Tables 4 and 5).

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218–5225

B3LYP/6-31G(d) and B3LYP/6-311++G(2d,2p) for all nine conformers and augmented these results with vibrational analysis to introduce thermal free energy corrections. Conformers are then "mixed" according to their calculated fractions. The rmsd values for SSCCs computed using all these weighting coefficients derived from increasingly sophisticated levels of DFT theory are summarized in Table 5.

The *J*-fitting procedure for penicilleremophilane B, $C_{15}H_{20}O_3$, using DU8-calculated SSCCs gave an rmsd of 0.31 Hz (mue = 0.58 Hz). Geometry optimization [B3LYP/6-31G(d)], computations of Fermi contacts (DU8), and isotropic components of the magnetic shielding tensors [GIAO, mPW1PW91/6-311+G(d,p)] took on average 4.3 h of CPU time per conformer, which corresponds to less than 17 min of computational (wall) time on a 16-core Linux node.[65] However, the quality of predicted SSCCs plunged when the relative conformer weighting was based on their B3LYP/6-31G(d)//B3LYP/6-31G(d) energies: rmsd ~0.67 Hz (mue = 1.32 Hz).

Table 5. Quality of SSCC Prediction As a Function of Conformer Ratios Computed at Different Levels of DFT Theory

DFT theory level for rel energy	rmsd[a]	mue[b]	<ΔΔ*E*>[c]	ΔΔ*E*max[d]	CPU time[e]
		penicilleremophilane B, $C_{15}H_{20}O_3$			
b3lyp/6-31G(d) geometry:					
(*J*-fitting)	**0.31**	**0.58**			**4.3**
B3LYP/6-31G(d)	0.67	1.32	0.5	1.1	
B3LYP/6-31G(d) + TFE[f]	0.66	1.44	0.6	1.1	(+2)
B3LYP/6-311++G(2d,2p)	0.52	1.06	0.5	1.2	
B3LYP/6-311++G(2d,2p) + TFE[f]	0.50	1.11	0.6	1.4	(+33)
b3lyp/6-311G++(2d,2p) geometry:					
B3LYP/6-311+G(2d,2p)	0.51	1.06	0.5	1.1	(+24)
B3LYP/6-311+G(2d,2p) + TFE[f]	0.47	1.06	0.6	1.3	(+33)
		penifulvin A, $C_{15}H_{20}O_4$			
b3lyp/6-31G(d) geometry:					
(*J*-fitting)	**0.23**	**0.52**			**8.4**
B3LYP/6-31G(d)	0.47	1.57	0.6	1.7	
B3LYP/6-31G(d) + TFE[f]	0.68	2.16	0.8	2.2	(+2)
b3lyp/6-311G++(2d,2p) geometry:					
B3LYP/6-311+G(2d,2p)	0.54	1.73	0.6	1.8	(+32)
B3LYP/6-311+G(2d,2p) + TFE[f]	0.26	0.6	0.7	2.1	(+35)
		hawaiinolide A, $C_{20}H_{26}O_4$			
b3lyp/6-31G(d) geometry:					
(*J*-fitting)	**0.22**	**0.50**			**11.7**
B3LYP/6-31G(d)	0.37	0.71	1.9	1.9	
B3LYP/6-31G(d) + TFE[f]	0.35	0.83	1.0	1.0	(+5)
b3lyp/6-311G++(2d,2p) geometry:					
B3LYP/6-311+G(2d,2p)	0.39	0.73	2.2	2.2	(+85)
B3LYP/6-311+G(2d,2p) + TFE[f]	0.34	0.77	0.8	1.4	(+93)

[a]rmsd on calculated and experimental SSCCs, Hz. [b]Maximum unsigned error on SSCCs, Hz. [c]Average energy adjustment per conformer in kcal/mol required to reconcile the conformers ratios obtained from *J*-fitting vs the DFT energies. [d]maximum energy adjustment. [e]CPU time in hours, average per conformer; (extra time, needed for the higher level computationsand frequency jobs are shown in parentheses). [f]thermal free energy correction from the frequency job.

By using increasingly more expensive levels of theory, the quality of conformer mixing was somewhat improved to rmsd ~0.47 Hz (mue 1.06 Hz). However, this was achieved with additional 24 h of CPU time required to obtain B3LYP/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) energy and another 33 h of CPU time to introduce the thermal free energy correction. Thus, the total running wall time per conformer was increased from under 17 min to more than 4 h, and yet, the rmsd of 0.31 Hz obtained by *J*-fitting was not achieved.

Another instructive example is penifulvin A ($C_{15}H_{20}O_4$)[26] with an ostensibly inflexible [5.5.5.6]fenestrane core and no rotatable bonds, Figure 3. B3LYP/6-31G(d) computations show that it has four interconverting conformers lying within 1.8 kcal/mol and that several large spin−spin coupling constants vary significantly as a result of this conformational flexibility. The largest variation, in the interval from 0.6 to 13.9 Hz, is exhibited by one of the *trans* constants in the unsubstituted dimethylene moiety. This large dynamic range requires the inclusion of all four conformers into the averaging scheme in order to predict the SSCCs accurately. The *J*-fitting procedure for penifulvin A yielded an excellent rmsd of 0.23 Hz (mue = 0.52 Hz). This accuracy was not reproduced when the ratio of conformers was calculated based on their relative DFT energies. Even with a better B3LYP/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) geometry and thermal free energy correction, the SSCCs were not quite as accurate (rmsd = 0.26 Hz). The value of 0.26 Hz is, of course, perfectly acceptable. However, there was a heavy computational cost in addition to the original 8.4 h of CPU time needed for DU8 computations: the higher level of theory required extra 67 h of CPU time per conformer.

Similar results were obtained for a larger molecule, hawaiinolide A, $C_{20}H_{26}O_4$, Figure 3. DU8 with *J*-fitting resulted in an excellent rmsd of 0.22 Hz (mue = 0.5 Hz) in less than 45 min of computational (wall) time for each of the two lowest energy conformers. However, with much more expensive B3LYP/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) energies and thermal free energy corrections, obtained with additional ~11 h of computational (wall) time, we nonetheless failed to achieve rmsd of 0.22 Hz. As the last entry in Table 5 illustrates, these much more accurate energies for conformer mixing yielded inferior rmsd of 0.34 Hz (mue = 0.77 Hz). Application of a polarized continuum model (IPCM, not shown) to account for solvent did not improve the quality of conformational mixing either.

What is very instructive in these examples is that the energy corrections, ΔΔ*E*, needed to reconcile the *J*-fitting procedure and the conformational averaging based on DFT energies are rather small, 1−2 kcal/mol, Table 5. These values are barely within the accuracy of gas-phase computations and beyond the accuracy of solution PCM models. Again, we would argue that a more practical approach to conformational averaging is to use *J*-fitting as the primary tool, with a subsequent cursory analysis of the energy corrections needed to reconcile the inexpensive B3LYP/6-31G(d)//B3LYP/6-31G(d) energies with the *J*-fitting results. This post factum inspection should be calibrated

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225

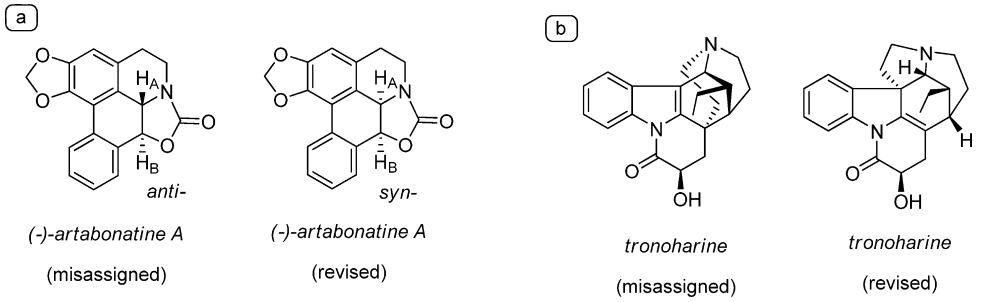


Figure 4. Artabonatine A and tronoharine.

for the reported accuracy of the energy method, triggering concerns only for the obvious $\Delta\Delta E$ outliers.

Spin–spin Coupling Constants Are Critical for Reliable Structure Assignments. Spin–spin coupling constants carry a wealth of structural information and thus are by far more important than the proton chemical shifts in matters of structural assignment (and misassignment). Arguably, there have been quite a few reported structural misassignments which could have been avoided if the spectra of the isolated natural products were analyzed using the accurate computed SSCC values. As a recent example, the originally proposed structure of artabonatine A[66] was corrected by Cuny[67] (Figure 4a), who synthesized both *syn*- and *anti*-isomers of it and showed that the critical spin–spin coupling constant J_{AB} for the misassigned *anti* structure was 13.2 Hz, whereas in the actual natural product (*syn*) this constant was 7.5 Hz. Our computations matched both structures very well (computed values for J_{AB} were 13.6 Hz for the *anti* and 7.4 Hz for the *syn* isomer).

Another example is the recent structural correction of tronoharine by Kam[60] (Figure 4b). This is a particularly difficult case for the initial assignment as both misassigned and the corrected structures have the same set of geminal and vicinal proton fragments, forming separate spin groups. Our computations resulted in an excellent match with the experimental spectrum of the revised structure, graciously provided by Kam, with rmsd of 0.31 Hz over 29 accurately measured experimental SSCCs, Table 3. Attempts to reproduce the experimental data with the misassigned structure yielded unsatisfactory rmsd > 3 Hz.

As computations for each individual conformer of tronoharine, $C_{21}H_{24}N_2O_2$, took under 45 min of running (wall) time on a single 16-core Linux node, the DU8 method offered a very practical and expeditious access to tronoharin's accurate proton NMR spectrum. Such computations could be readily justified as a sensible alternative to the synthetic attempts toward the misassigned structure of tronoharine, which seem to continue well into 2015.[68]

CONCLUSIONS

The second-generation relativistic force field parametrization method DU8 is developed for scaling of Fermi contacts to compute accurate proton spin–spin coupling constants. It utilizes a smaller (13) parameter set and NBO hybridization corrections to reliably achieve rmsd ~0.2–0.4 Hz in predicting SSCCs. As the basis set optimized for computing Fermi contacts evolved from an inexpensive collection of 3-21G*, 4-31G, and 6-31G(d) bases, short computational (wall) times of 30–45 min per conformer are achieved on a 16-core node of a Linux cluster for organic molecules containing 15–20 carbon atoms. This computational time includes geometry optimization and calculations of proton chemical shifts (approximately 70% of total computational time). Shortening of the computational time per conformer without sacrificing the accuracy of computations is critical for fast and accurate predictions of NMR spectra in conformationally flexible organic structures where averaging is necessarily performed over an ensemble of conformers.

ASSOCIATED CONTENT

Supporting Information

Computational details, structures, and computed SSCCs (62 pages). The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.joc.5b00619.

AUTHOR INFORMATION

Corresponding Author

*E-mail: akutatel@du.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Support of this work by the National Science Foundation, CHE-1362959 is gratefully acknowledged. We thank Martin Banwell (Australian National University), Tushar Chakraborty (Indian Institute of Science, Bangalore), Yong-Xian Cheng (Kunming Institute of Botany and Henan College of Traditional Chinese Medicine), Hanfeng Ding (Zhejiang University), Alison Frontier (University of Rochester), Seth Herzon (Yale University), Toh-Seok Kam (University of Malaya), Vatcharin Rukachaisirikul (Prince of Songkla University, Thailand), Scott Sieburth (Temple University), Xiao-Yi Wei (Chinese Academy of Sciences), Lindon West (Florida Atlantic University), Satoshi Yokoshima (Nagoya University), Qin-Shi Zhao (Chinese Academy of Sciences), and Jing-Jing Zhu (Institute of Chinese Materia Medica) for sharing experimental NMR data with us.

REFERENCES

(1) (a) Onak, T.; Jaballas, J.; Barfield, M. *J. Am. Chem. Soc.* **1999**, *121*, 2850. (b) Scheurer, C.; Bruschweiler, R. *J. Am. Chem. Soc.* **1999**, *121*, 8661. (c) Del Bene, J. E.; Bartlett, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 10480. (d) Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 3560. (e) Del Bene, J. E.; Perera, S. A.; Bartlett, R. J. *J. Phys. Chem. A* **2001**, *105*, 930. (f) Diez, E.; Casanueva, J.; San Fabian, J.; Esteban, A. L.; Galache, M. P.; Barone, V.; Peralta, J. E.; Contreras, R. H. *Mol. Phys.* **2005**, *103*, 1307.

(2) Bally, T.; Rablen, P. R. *J. Org. Chem.* **2011**, *76*, 4818.

(3) Kutateladze, A. G.; Mukhina, O. A. *J. Org. Chem.* **2014**, *79*, 8397.

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218–5225

(4) For a recent mini-review on orbital hybridization as a key electronic factor in control of structure and reactivity, see: Alabugin, I. V.; Bresch, S.; Gomes, G. P. *J. Phys. Org. Chem.* **2014**, *28*, 147.
(5) Wilkens, S. J.; Westler, W. M.; Markley, J. L.; Weinhold, F. *J. Am. Chem. Soc.* **2001**, *123*, 12026.
(6) (a) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211. (b) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066. (c) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.
(7) National Institute for Advanced Industrial Science and Technology, Japan, http://sdbs.db.aist.go.jp/sdbs/cgi-bin/cre_index.cgi.
(8) Latest version of the NBO program, NBO 6.0: Glendening, E. D.; Badenhoop, K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Landis, C. R.; Weinhold, F. *Theoretical Chemistry Institute*; University of Wisconsin, Madison, 2013.
(9) We suggest that the corrections accounting for these "missing" contributions to SSCCs should correlate with the p-character of the hybrid orbitals (for example, a non-zero contribution from spin−orbit coupling necessarily requires orbital angular momentum). On the contrary, if parametric corrections were needed primarily to refine the dominating contributor, i.e., the Fermi contacts, one would expect these corrections to correlate with the s-character of the hybrid orbitals, not p.
(10) Crimmins, M. T.; Pace, J. M.; Nantermet, P. G.; Kim-Meade, A. S.; Thomas, J. B.; Watterson, S. H.; Wagman, A. S. *J. Am. Chem. Soc.* **2000**, *122*, 8453.
(11) (a) Cornolactone B: ref 30. (b) Dehydroleuconoxine: ref 34. (c) Dichomitol: ref 19c. (d) Isopalhinine A: ref 17. (e) Salvileucalin D, salvipuberulin, and isosalvipuberulin: ref 36. (f) Lepistine: ref 51. (g) *N,O*-Dimethyloxostephine: ref 58. (h) Penifulvin A: ref 26. (i) Tronoharine: ref 60a. (j) Wenyujinin H: ref 62. (k) Compound **24**, ABCDE-ring system of the *Strychnos*-type alkaloids: ref 25.
(12) Brito, I.; Cueto, M.; Diaz-Marrero, A. R.; Darias, J.; San Martin, A. S. *J. Nat. Prod.* **2002**, *65*, 946.
(13) Okamura, H.; Iwagawa, T.; Nakatani, M. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 3465.
(14) Strychnine SSCCs are averaged from a compilation in: Cobas, J. C.; Constantino-Castillo, V.; Martin-Pastor, M.; del Rio-Portilla, F. *Magn. Reson. Chem.* **2012**, *50*, S86.
(15) Lodewyk, M. W.; Soldi, C.; Jones, P. B.; Olmstead, M. M.; Rita, J.; Shaw, J. T.; Tantillo, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 18550.
(16) Findlay, J. A.; Li, G. *Can. J. Chem.* **2002**, *80*, 1697.
(17) Dong, L.-B.; Gao, X.; Liu, F.; He, J.; Wu, X.-D.; Li, Y.; Zhao, Q.-S. *Org. Lett.* **2013**, *15*, 3570.
(18) Neville, G. A.; Ekiel, I.; Smith, I. C. P. *Magn. Reson. Chem.* **1987**, *25*, 31.
(19) (a) Original report of a triquinane structure of dichomitol (misassigned): Huang, Z.; Dan, Y.; Huang, Y.; Lin, L.; Li, T.; Ye, W.; Wei, X. *J. Nat. Prod.* **2004**, *67*, 2121. (b) Synthesis of the originally proposed structure: Mehta, G.; Pallavi, K. *Tetrahedron Lett.* **2006**, *47*, 8355. (c) Structure correction: Xie, H.-H.; Xu, X.-Y.; Dan, Y.; Wei, X.-Y. *Helv. Chim. Acta* **2011**, *94*, 868.
(20) Kam, T.-S.; Choo, Y.-M. *Phytochemistry* **2004**, *65*, 603.
(21) Ali, A. A.; El Sayed, H. M.; Abdallah, O. M.; Steglich, W. *Phytochemistry* **1986**, *25*, 2399.
(22) Yang, X.-W.; Yang, C.-P.; Jiang, L.-P.; Qin, X.-J.; Liu, Y.-P.; Shen, Q.-S.; Chen, Y.-B.; Luo, X.-D. *Org. Lett.* **2014**, *16*, 5808.
(23) Kimishima, A.; Umihara, H.; Mizoguchi, A.; Yokoshima, S.; Fukuyama, T. *Org. Lett.* **2014**, *16*, 6244.
(24) Suzuki, M.; Daitoh, M.; Vairappan, C. S.; Abe, T.; Masuda, M. *J. Nat. Prod.* **2001**, *64*, 597.
(25) Reekie, T. A.; Banwell, M. G.; Willis, A. C. *J. Org. Chem.* **2012**, *77*, 10773.
(26) Das, D.; Kant, R.; Chakraborty, T. K. *Org. Lett.* **2014**, *16*, 2618.
(27) Marrero, J.; Rodrıguez, A. D.; Baran, P.; Raptis, R. G.; Sanchez, J. A.; Ortega-Barria, E.; Capson, T. L. *Org. Lett.* **2004**, *6*, 1661.
(28) Casero, C. N.; Oberti, J. C.; Orozco, C. I.; Cárdenas, A.; Brito, I.; Barboza, G. E.; Nicotra, V. E. *Phytochemistry* **2015**, *110*, 83.
(29) Deng, C.; Huang, C.; Wu, Q.; Pang, J.; Lin, Y. *Nat. Prod. Res.* **2013**, *27*, 1882.
(30) He, Y.; Peng, J.; Hamann, M. T.; West, L. M. *J. Nat. Prod.* **2014**, *77*, 2138.
(31) Zhang, C.; Liu, J.; Du, Y. *Tetrahedron Lett.* **2013**, *54*, 3278.
(32) Oh, D.-C.; Williams, P. G.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *Org. Lett.* **2006**, *8*, 1021.
(33) Miyake, K.; Tezuka, Y.; Awale, S.; Li, F.; Kadota, S. *J. Nat. Prod.* **2009**, *72*, 2135.
(34) Low, Y. Y.; Hong, F.-J.; Lim, K.-H.; Thomas, N.-F.; Kam, T.-S. *J. Nat. Prod.* **2014**, *77*, 327.
(35) Dong, M.; Cong, B.; Yu, S.-H.; Sauriol, F.; Huo, C.-H.; Shi, Q.-W.; Gu, Y.-C.; Zamir, L. O.; Kiyota, H. *Org. Lett.* **2008**, *10*, 701.
(36) Fu, C.; Zhang, Y.; Xuan, J.; Zhu, C.; Wang, B.; Ding, H. *Org. Lett.* **2014**, *16*, 3376.
(37) Zhao, J.-X.; Liu, C.-P.; Qi, W.-Y.; Han, M.-L.; Han, Y.-S.; Wainberg, M. A.; Yue, J.-M. *J. Nat. Prod.* **2014**, *77*, 2224.
(38) Li, J.-Y.; Zhao, B.-X.; Zhang, W.; Li, C.; Huang, X.-J.; Wang, Y.; Sun, P.-H.; Ye, W.-C.; Chen, W.-M. *Tetrahedron* **2012**, *68*, 3972.
(39) Cleary, L.; Pitzen, J.; Brailsford, J. A.; Shea, K. J. *Org. Lett.* **2014**, *16*, 4460.
(40) Lakornwong, W.; Kanokmedhakul, K.; Kanokmedhakul, S.; Kongsaeree, P.; Prabpai, S.; Sibounnavong, P.; Soytong, K. *J. Nat. Prod.* **2014**, *77*, 1545.
(41) Finn, P. B.; Kulyk, S.; Sieburth, S. McN. *Tetrahedron Lett.* **2015**, DOI: 10.1016/j.tetlet.2015.01.145.
(42) Ng, F. W.; Lin, H.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 9812.
(43) Jansen, R.; Sood, S.; Mohr, K. I.; Kunze, B.; Irschik, H.; Stadler, M.; Müller, R. *J. Nat. Prod.* **2014**, *77*, 2545.
(44) Chen, S.; Zhang, Y.; Niu, S.; Liu, X.; Che, Y. *J. Nat. Prod.* **2014**, *77*, 1513.
(45) Luo, Q.; Tian, L.; Di, L.; Yan, Y.-M.; Wei, X.-Y.; Wang, X.-F.; Cheng, Y.-X. *Org. Lett.* **2015**, *17*, 1565.
(46) Ziegler, F. E.; Metcalf, C. A., III; Nangia, A.; Schulte, G. *J. Am. Chem. Soc.* **1993**, *115*, 2581.
(47) Porco, J. A., Jr.; Su, S.; Lei, X.; Bardhan, S.; Rychnovsky, S. D. *Angew. Chem.* **2006**, *118*, 5922.
(48) Yang, X.-W.; Peng, K.; Liu, Z.; Zhang, G.-Y.; Li, J.; Wang, N.; Steinmetz, A.; Liu, Y. *J. Nat. Prod.* **2013**, *76*, 2360.
(49) Bai, R.; Zhang, C.-C.; Yin, X.; Wei, J.; Gao, J.-M. *J. Nat. Prod.* **2015**, *78*, 783.
(50) Hajicek, J.; Taimr, J.; Budesinsky, M. *Tetrahedron Lett.* **1998**, *39*, 505.
(51) Kitabayashi, Y.; Yokoshima, S.; Fukuyama, T. *Org. Lett.* **2014**, *16*, 2862.
(52) Umehara, A.; Ueda, H.; Tokuyama, H. *Org. Lett.* **2014**, *16*, 2526.
(53) Chen, H.-J.; Wu, Y. *Org. Lett.* **2015**, *17*, 592.
(54) Hoshi, M.; Kaneko, O.; Nakajima, M.; Arai, S.; Nishida, A. *Org. Lett.* **2014**, *16*, 768.
(55) Carlsen, P. N.; Mann, T. J.; Hoveyda, A. H.; Frontier, A. J. *Angew. Chem.* **2014**, *53*, 9334.
(56) Lu, P.; Gu, Z.; Zakarian, A. *J. Am. Chem. Soc.* **2013**, *135*, 14552.
(57) Chen, S.; Ren, F.; Niu, S.; Liu, X.; Che, Y. *J. Nat. Prod.* **2014**, *77*, 9.
(58) Calandra, N. A.; King, S. M.; Herzon, S. B. *J. Org. Chem.* **2013**, *78*, 10031.
(59) Homs, A.; Muratore, M. E.; Echavarren, A. M. *Org. Lett.* **2015**, *17*, 461.
(60) (a) Structure revision: Sim, D. S.-Y.; Chong, K.-W.; Nge, C.-E.; Low, Y.-Y.; Sim, K.-S.; Kam, T.-S. *J. Nat. Prod.* **2014**, *77*, 2504. (b) Original structure: Kam, T.-S.; Sim, K.-M.; Lim, T.-M. *Tetrahedron Lett.* **1999**, *40*, 5409.
(61) (a) Raju, R.; Piggott, A. M.; Huang, X.; Capon, R. J. *Org. Lett.* **2011**, *13*, 2770. (b) Note that proton *1′′a* in Capon's paper is listed as 2.79 ppm *dd* 11.0, 3.7 Hz, with 11.0 being a typo. The correct larger constant of 15.0 Hz is taken from: (c) Wang, H.; Reisman, S. E. *Angew. Chem., Int. Ed.* **2014**, *53*, 6206.

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225

(62) Yin, G.-P.; Li, L.-C.; Zhang, Q.-Z.; An, Y.-W.; Zhu, J.-J.; Wang, Z.-M.; Chou, G.-X.; Wang, Z.-T. *J. Nat. Prod.* **2014**, *77*, 2161.
(63) Peese, K. M.; Gin, D. Y. *Chem.—Eur. J.* **2008**, *14*, 1654.
(64) Daengrot, C.; Rukachaisirikul, V.; Tansakul, C.; Thongpanchang, T.; Phongpaichit, S.; Bowornwiriyapan, K.; Sakayaroj, J. *J. Nat. Prod.* **2015**, *78*, 615.
(65) It is important to note that in the DU8 approach the actual computations of Fermi contacts take less than a third of the total computational time. Nearly two-thirds of computational time is used for computing the isotropic components of the magnetic shielding tensors at the GIAO, mPW1PW91/6-311+G(d,p) level of theory to obtain chemical shifts, with the reminder of the time spent for geometry optimization. Without chemical shifts, accurate SSCCs for penicilleremophilane B are obtained in less than 6 min of computational (wall) time.
(66) Hsieh, T.-J.; Chen, C.-Y.; Kuo, R.-Y.; Chang, F.-R.; Wu, Y.-C. *J. Nat. Prod.* **1999**, *62*, 1192.
(67) Ku, A. F.; Cuny, G. D. *Org. Lett.* **2015**, *17*, 1134.
(68) Zhong, X.; Li, Y.; Zhang, J.; Han, F.-S. *Org. Lett.* **2015**, *17*, 720.

DOI: 10.1021/acs.joc.5b00619
J. Org. Chem. 2015, 80, 5218−5225