Comparison of SCC-DFTB and NDDO-Based Semiempirical Molecular Orbital Methods for Organic Molecules

Kurt W. Sattelmeyer, Julian Tirado-Rives, and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Received: July 18, 2006; In Final Form: October 10, 2006

Extensive testing of the SCC−DFTB method has been performed, permitting direct comparison to data available for NDDO-based semiempirical methods. For 34 diverse isomerizations of neutral molecules containing the elements C, H, N, and O, the mean absolute errors (MAE) for the enthalpy changes are 2.7, 3.2, 5.0, 5.1, and 7.2 kcal/mol from PDDG/PM3, B3LYP/6-31G(d), PM3, SCC−DFTB, and AM1, respectively. A more comprehensive test was then performed by computing heats of formation for 622 neutral, closed-shell H, C, N, and O-containing molecules; the MAE of 5.8 kcal/mol for SCC−DFTB is intermediate between AM1 (6.8 kcal/mol) and PM3 (4.4 kcal/mol) and significantly higher than for PDDG/PM3 (3.2 kcal/mol). Similarly, SCC−DFTB is found to be less accurate for heats of formation of ions and radicals; however, it is more accurate for conformational energetics and intermolecular interaction energies, though none of the methods perform well for hydrogen bonds with strengths under ca. 7 kcal/mol. SCC−DFTB and the NDDO methods all reproduce MP2/cc-pVTZ molecular geometries with average errors for bond lengths, bond angles, and dihedral angles of only ca. 0.01 Å, 1.5°, and 3°. Testing was also carried out for sulfur containing molecules; SCC−DFTB currently yields much less accurate heats of formation in this case than the NDDO-based methods due to the over-stabilization of molecules containing an SO bond.

Introduction

During the last several decades, semiempirical molecular orbital (SMO) methods aimed at reliably reproducing experimentally determined molecular properties, yet scaling only as N^3, have been developed; chief among these have been methods derived from the neglect of diatomic differential overlap (NDDO) approximation,[1−3] namely MNDO[4−5] and its successors AM1,[6] PM3,[7] and PDDG/PM3.[8−10] Although the latter methods differ only in their inclusion of extra terms in the core repulsion formula and the better parameter optimization procedures afforded by modern computational power,[11] the resulting enhancements to molecular properties are considerable. For example, the MAEs for heats of formation of a set of 622 neutral, closed-shell molecules containing H, C, N, and O for these methods are 8.3, 6.8, 4.4, and 3.2 kcal/mol, respectively. Much of the focus in recent development has been on lowering these errors while maintaining the generally good agreement for structural properties; e.g., PDDG/PM3 gives MAEs for bond lengths, bond angles, and dihedral angles of only 0.013 Å, 1.9°, 3.7°, respectively, versus MP2/cc-pVTZ values on a large test set.[12] Continued development of such fast quantum mechanical methods is also driven by the interests in applications to large molecular systems and in mixed quantum and molecular mechanics (QM/MM) calculations for organic and enzymatic reactions in solution.[13]

Mirroring the increased use of density functional theory (DFT) over the past decade has been the introduction and development of a semiempirical methodology, density-functional based tight binding (DFTB), which incorporates tight binding principles and is parametrized using results based on the local density approximation (LDA).[14,15] Here, the central approximation is that the initial molecular electron density can be constructed as a superposition of atomic densities, implying that a self-consistent solution is not needed. The energy expression (eq 1) is then fully determined after parametrization for

$$E_0^{\mathrm{TB}} = \sum_i^{\mathrm{occ}} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle + E_{\mathrm{rep}} \qquad (1)$$

the core repulsion formula, CRF or E_{rep}, which is typically composed of a Chebyshev polynomial fit such that E_0^{TB} approximates the energy of LDA calculations on the relevant diatomics or another suitable system over a range of interatomic distances.[14] The adjustable parameters are the coefficients of the polynomials, so with the use of 10 polynomials, as in ref 14, there are 10 parameters. These are pairwise terms by atom type for DFTB, so there are 10 parameters for H−H interactions, 10 for C−H, etc., whereas the terms are atom-based for PM3 and PDDG/PM3. Thus, for hydrocarbons, DFTB contains 30 (3 × 10) parameters in the CRFs, whereas PM3 and PDDG/PM3 have 14 (2 × 7) and 22 (2 × 11), respectively. Then, the number of parameters increases quadratically for DFTB and linearly for the NDDO methods with the addition of new atom types.

DFTB has shown promise for hydrocarbons and bulk systems, but polar molecules can be problematic; for example, the two CO bonds in carboxylic acids are computed to have nearly equal lengths.[16] Therefore, Elstner et al. extended the DFTB formalism to account for the charge redistribution in organic molecules. The resultant method (SCC−DFTB) requires an iterative solution according to eq 2.[16] The $H_{\mu\nu}$ elements are defined by eq 3, where S is the overlap matrix, the γs correspond to Coulomb interactions, and $\Delta q_\xi = q_\xi - q_\xi^0$ represents the change in the charge distribution. The presence of the second

---

* Corresponding author. E-mail: william.jorgensen@yale.edu.

10.1021/jp064544k CCC: $33.50

term in eq 3 and the need for an SCF cycle provide the principal differentiation from DFTB. The energy expression is then given by eq 4

$$\sum_{\nu}^{M} c_{\nu i}\,(H_{\mu n} - \epsilon_i S_{\mu n}) = 0 \qquad (2)$$

$$H_{\mu n} = \langle \varphi_\mu | \hat{H}_0 | \varphi_\nu \rangle + \frac{1}{2} S_{\mu n} \sum_{\xi}^{N} (\gamma_{\alpha x} + \gamma_{\beta x}) \Delta q_\xi \qquad (3)$$

$$E = \sum_{i}^{occ} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle + \frac{1}{2} \sum_{\alpha,\beta}^{N} \gamma_{a\beta} \Delta q_\alpha \Delta q_\beta + E_{\mathrm{rep}} \qquad (4)$$

Unlike the NDDO methods, SCC−DFTB explicitly incorporates the overlap matrix into the Roothaan−Hall equations,[17] the solution of which necessitates one additional matrix diagonalization at the start of the procedure and several additional matrix−matrix multiplications per iterative step. Therefore, SCC−DFTB is somewhat slower than methods which utilize the NDDO approximation, although still scaling as N^3. Using similar compiler options and the same machine, it appears that SCC−DFTB calculations require about 50% more computer time than NDDO-based methods. For example, the consecutive calculation of 100 single-point energies of tri-*t*-butylmethane with SCC−DFTB and PDDG/PM3 gave a timing ratio of 1.4. This timing difference is typical, as similar comparisons for a number of large molecules give ratios between 1.2 and 2.0.

There are similarities between SCC−DFTB and extended Hückel theory including use of a valence-only basis set and inclusion of the overlap matrix in the secular equations.[18] In addition, the diagonal elements of eq 3 are similar to those of the first approximation suggested by Streitwieser in his description of the ω-Hückel method in 1960.[19] However, as with MNDO and its NDDO variants, the expanded two-center elements of SCC−DFTB are better able to represent the electronic structure of molecules. Additionally, both genres of semiempirical methods are similar in their deviation from their predecessors by the inclusion of core repulsion terms (E_{rep} in eq 4), which repair deficiencies in the underlying methods associated with, for example, the minimal basis sets and simplified treatment of electron repulsion.

Thus far, benchmarking of SCC−DFTB for organic molecules has not been extensive; it has focused on reaction energies and structures of small, acyclic molecules with fewer than five nonhydrogen atoms.[16,20,21] Thus, many questions remain concerning the performance for larger molecules and the treatment of basic issues such as ring strain, steric effects, conjugated systems, conformational energetics, and intermolecular interactions. The present paper presents a much more comprehensive comparison of results from SCC−DFTB and the NDDO schemes focusing on these issues in the standard manner that has been characteristic of the development of SMO methods.[4−12] The SCC−DFTB code used here was obtained from Dr. M. Elstner, the other SMO calculations were performed with BOSS[22] or a local version of MOPAC,[23] and the B3LYP/6-31G(d) results are from Gaussian 03[24] or the NIST database.[25] The provided SCC−DFTB code has parameters to treat the elements C, H, N, O, and S, so the present comparisons have been restricted accordingly.

Results and Discussion

Isomerization Energies. A good place to start is the energetics of isomerization reactions since this allows direct comparisons on the performance for differences in bonding, conjugation, and steric effects. Results are reported in Tables 1−3 from geometry optimizations with AM1, PM3, PDDG/PM3, and SCC−DFTB. For comparison with a well-established DFT method, B3LYP/6-31G(d) results have also been included using B3LYP/6-31G(d) optimized geometries. The experimental data and the AM1, PM3, and PDDG-PM3 results correspond to enthalpy changes at 25 °C. In some prior studies, SCC−DFTB electronic energies have been corrected for vibrational energy changes to yield enthalpy differences.[20,21] Besides being inconvenient for an SMO method, inclusion of the zero-point vibrational energy changes using frequencies from the B3LYP/6-31G(d) optimizations is found to have little effect on the MAE for the SCC−DFTB results in Table 1. Inclusion of the zero-point vibrational energy change lowers the MAE from 5.0 to 4.9 kcal/mol, further adding the thermal correction for 0 to 298 K makes the MAE 4.7 kcal/mol. It also has little effect on the B3LYP/6-31G(d) results in Table 1; the MAE is unchanged at 3.1 kcal/mol with the zero-point correction and goes to 2.9 kcal/mol including the thermal term. Consequently, the vibrational energy corrections were not made for Tables 2 and 3.

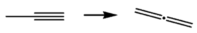
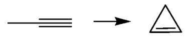
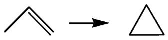
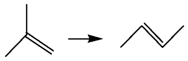
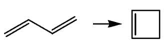
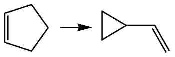
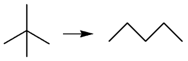
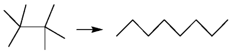
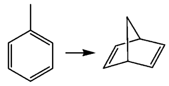
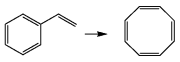
A total of 34 isomerizations is represented in Tables 1−3. For the hydrocarbons in Table 1, the MAEs increase in the order PDDG/PM3 (2.4 kcal/mol), B3LYP/6-31G(d) (3.1 kcal/mol), PM3 (4.3), SCC/DFTB (5.0), and AM1 (6.6). The sequence of MAEs is similar in all three tables. If the 34 results are combined, the overall order is PDDG/PM3 (2.7 kcal/mol), B3LYP/6-31G(d) (3.2), PM3 (5.0), SCC−DFTB (5.1), and AM1 (7.2). If the vibrational energy change including thermal corrections is included for the SCC−DFTB results, the MAE may drop to ca. 4.7 kcal/mol based on the results for Table 1. The newly reparameterized version of AM1 (RM1)[11] was also tried for the 34 isomerizations and it yielded an overall MAE of 4.2 kcal/mol.

For the hydrocarbons, the isomerizations cover diverse processes including various pericyclic transformations. For PDDG/PM3, the only errors larger than 4 kcal/mol are for the methylacetylene to allene and styrene rearrangements. Errors with B3LYP/6-31G(d) are small until one gets to the larger molecules, e.g., the last four entries in Table 1. This pattern was pointed out previously,[8] and branching problems for alkanes with DFT calculations in general have recently been highlighted.[26] Prior branching problems with MNDO and AM1 were resolved well with PDDG/PM3.[8] For example, the errors with numerous DFT methods are 6−12 kcal/mol all overly disfavoring tetramethylbutane compared to octane, and the error with B3LYP/cQZV3P is 10 kcal/mol, similar to what is found here with the much smaller 6-31G(d) basis set.[26b] The errors with multiple DFT methods such as B88−PW91/TZVP for octamethylhexane compared to tetradecane exceed 30 kcal/mol,[26a] whereas the error is 46 kcal/mol for AM1, 9 kcal/mol for PM3, and 6 kcal/mol for both PDDG/PM3 and SCC−DFTB.

B3LYP/6-31G(d) also has significant trouble with the toluene to norbornadiene isomerization with a nearly 10 kcal/mol error. Since the treatment of ring strain does not appear to be problematic for smaller molecules, it is likely that the nonbonded interactions in the more congested bicyclic ring system are not well represented. For SCC−DFTB, the principal deficiencies are for the examples involving cyclopropene, vinylcyclopropane, and cyclooctatetraene, whereas the 1,3-butadiene to cyclobutene transformation is handled well. AM1 and PM3 also have some difficulties with small rings and the branching examples.

Tables 2 and 3 report the isomerization results for molecules containing nitrogen and oxygen. A noticeable item in Table 2

TABLE 1: Isomerization Enthalpies for Hydrocarbons (kcal/mol)

	Exp.[a]	B3LYP/6-31G(d)	AM1	PM3	PDDG/PM3	SCC-DFTB	SCC-DFTB + ZPVE[b]
	1.24	-2.76	2.73	6.84	5.87	6.18	6.03
	21.81	21.48	31.41	27.95	21.58	38.67	39.02
	7.85	7.75	10.66	9.20	8.14	12.20	13.23
	1.13	1.35	1.14	0.21	1.66	0.96	1.08
	1.27	0.23	-3.33	-1.65	-0.64	-1.19	-1.21
	2.79	3.67	3.60	5.25	4.97	4.11	4.44
	11.34	11.85	15.85	6.64	7.83	12.84	13.75
	22.17	22.19	31.97	30.85	24.90	34.95	33.51
	7.13	9.78	4.38	5.19	5.36	7.37	7.42
	5.06	0.75	-5.20	1.35	7.15	1.81	2.55
	4.09	-7.54	-18.65	-3.48	7.23	-1.13	-0.16
	46.65	56.06	53.28	44.67	43.92	49.65	50.11
	35.83	40.61	25.68	27.56	30.91	44.42	43.94
MAE		3.1[c]	6.6	4.3	2.4	5.0	4.9

[a] Values from ref 8 and provided in the Supporting Information. [b] Including zero-point vibrational energy (ZPVE) from B3LYP/6-31G(d) frequencies. MAE is 4.7 kcal/mol including thermal corrections. [c] With ZPVE, the MAE is 3.1 kcal/mol; with ZPVE and thermal corrections, it is 2.9 kcal/mol.

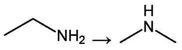
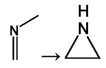
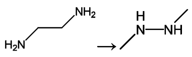
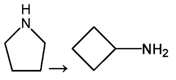
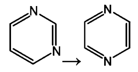
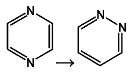
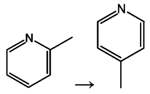
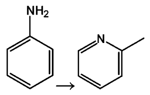
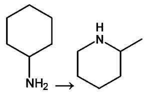
is the 6 kcal/mol error with B3LYP/6-31G(d) for the ethylenediamine to 1,2-dimethylhydrazine case. The pattern worsens to a 12 kcal/mol error for the isoelectronic 1,2-ethanediol to dimethylperoxide isomerization in Table 3. Zero-point corrections are not the issue; adding the zero-point vibrational energy change actually worsens the accord by 0.7 and 1.0 kcal/mol in these two cases, respectively. The ethanol to dimethyl ether isomerization with a nearly 7 kcal/mol error appears to reflect a related problem for B3LYP/6-31G(d). PDDG/PM3 performs consistently; the largest errors are 6−7 kcal/mol and involve cyclobutylamine and oxirane. With SCC−DFTB, the errors are ca. 10 kcal/mol for the nitrile to isonitrile case and for the examples with aziridine and the two lactones.

At first glance, remarkably, the branching example of hexanoic acid to methyl pivalate is not problematic for B3LYP/6-31G(d) in Table 3. However, it just reflects a cancellation of errors; the acetic acid to methyl formate error (−6.5 kcal/mol in Table 3) is off-setting the error for the pentane to neopentane component (+4.3 kcal/mol in Table 1). So, if one was seeking a really poor result with B3LYP/6-31G(d), consider pivalic acid to ethyl propanoate, $(CH_3)_3CCOOH \rightarrow CH_3CH_2COOCH_2CH_3$. The experimental enthalpy change is +6.3 kcal/mol, whereas the B3LYP/6-31G(d) result is −5.2 or −4.3 kcal/mol including the zero-point and thermal corrections.

General Issues for Comparisons of Methods. Though these isomerization examples are diverse and provide good tests of the treatment of variations in bonding, steric effects, conjugation, and ring strain, they are still limited to ca. 60 unique molecules. They were selected quite randomly by going through the database of SMO results[8] looking for isomers and variety in the processes. Nevertheless, by either under- or over-representation of a problem class for any one method, the MAE results for that method can be distorted.[7] Addition of the pivalic acid to ethyl propanoate example alone would increase the MAE for B3LYP/6-31G(d) in Table 3 by 1 kcal/mol, and another 3 kcal/mol could likely be obtained with the octamethylhexane example. The issue also arises in repeated tests for one reaction type for which a particular method may poorly represent one key molecule (e.g., hydrogen or water for hydrogenations or hydrolyzes). A systematic error can result that distorts impressions.

For example, in ref 20 on the parametrization of SCC−DFTB for sulfur, energetic results are compared for 13 reactions using SCC−DFTB, B3LYP/6-31G(d), and several SMO methods

TABLE 2: Isomerization Enthalpies (kcal/mol) for Nitrogen-Containing Molecules

	Exp.[a]	B3LYP/6-31G(d)	AM1	PM3	PDDG/PM3	SCC-DFTB
$CH_3CN \rightarrow CH_3NC$	21.38	24.07	31.10	31.40	26.72	10.48
	6.92	4.51	7.87	3.25	6.57	8.99
	11.71	14.40	15.77	17.05	16.78	27.99
	26.06	20.31	34.26	19.97	27.08	25.24
	10.65	13.89	10.66	10.49	3.38	12.75
	0.03	4.07	0.31	1.35	1.57	3.94
	19.66	18.60	11.10	16.68	16.29	11.58
	1.10	0.89	-1.49	-0.37	-0.43	1.26
	2.89	-2.53	5.15	-1.81	0.86	3.46
	4.88	1.26	8.72	1.99	5.21	6.96
MAE		3.1	4.0	3.9	2.8	4.7

[a] Values from ref 8.

including PM3. PM3 appears to perform poorly with a surprisingly large rms error, 36.6 kcal/mol; however, 8 of the 13 reactions are hydrogenations involving 1−4 equiv of H_2. Since the error in heat of formation (ΔH_f) of H_2 with PM3 is known to be 13.4 kcal/mol,[7,8] the poor outcome in this test is preordained. It would be easy to fix the H_2 error for PM3 by a trivial addition to the H−H CRF, but the philosophy in its development emphasized larger molecules, whereas much attention has been paid to small molecules in the development of the CRFs for SCC−DFTB. Nevertheless, it turns out that SCC−DFTB does not represent carbon monoxide, and carbon dioxide well, so errors for carbonylations and carboxylations are substantial, e.g., the SCC/DFTB results for eqs 5−6 are too endothermic by 10−15 kcal/mol. Clearly, it would be inappropriate to judge the performance of SCC−DFTB by emphasizing the energetics of, e.g., carbonylations

$$C_2H_6 + CO \rightarrow (CH_3)_2C{=}O \tag{5}$$

$$C_2H_6 + CO_2 \rightarrow CH_3COOCH_3 \tag{6}$$

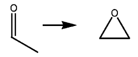
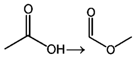
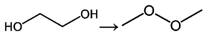
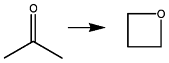
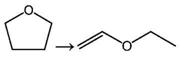
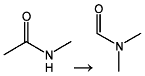
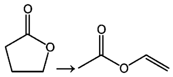
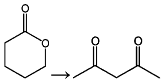
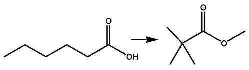
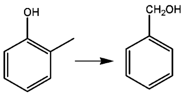
In order to minimize such problems for overall comparisons of molecular energetics, a very large collection of molecules needs to be treated, and it is advisable to consider either heats of atomization or heats of formation that have been adjusted to minimize systematic errors. This has been the standard procedure in the development of the NDDO-based methods[4−12] and the G2 and G3 procedures.[27]

Heats of Formation. One approach to estimate heats of formation at 298 K for SCC−DFTB is to use the same procedure as for the NDDO methods. For the compounds containing C, H, N, and O, the identical procedure was followed as in the development of PDDG/PM3.[9] Specifically, the energies of optimized structures for a training set of 134 molecules were calculated. Heats of formation are computed via eq 7, where E_{el}^{A} (also known as *eisol*) and ΔH_f^A are the electronic energies and heats of formation of the atoms composing the molecule.

$$\Delta H_f = E_{tot}^{mol} - \sum_A E_{el}^{A} + \sum_A \Delta H_f^A \tag{7}$$

The heats of formation of the atoms are taken as the experimental values, and the electronic energies of the atoms are treated as adjustable parameters, which are optimized to minimize the errors for the heats of formation of the molecules in the training set.[8] This removes systematic errors in the atomic component of the energy change. The same method has been used in other recent studies to compute heats of formation with

TABLE 3: Isomerization Enthalpies (kcal/mol) for Oxygen-Containing Molecules

	Exp.[a]	B3LYP/6-31G(d)	AM1	PM3	PDDG/PM3	SCC-DFTB
$C_2H_5OH \rightarrow (CH_3)_2O$	12.25	5.50	9.49	8.54	9.09	12.30
	27.15	27.52	32.60	36.07	33.07	36.37
	18.30	11.79	11.94	14.97	16.21	19.58
	62.57	50.12	83.44	63.37	59.44	62.32
	32.65	31.80	23.64	26.61	30.37	28.40
	10.39	12.26	27.07	22.99	15.14	18.95
	13.50	7.10	11.38	7.16	11.90	12.01
	12.33	12.61	23.61	22.98	16.19	21.93
	0.26	5.53	16.71	8.28	3.05	11.14
	4.39	4.83	20.90	15.20	7.61	10.50
	8.13	8.11	-2.17	4.51	8.89	11.58
MAE		3.4	10.7	6.8	3.1	5.5

[a] Values from ref 8.

TABLE 4: Atomic Heats of Formation and Fitted Electronic Energies for Computation of Heats of Formation with SCC−DFTB

	H	C	N	O	S
ΔH_f^A (kcal/mol)	52.102	170.89	113.00	59.559	66.4
E_{el}^A (eV)[a]	−7.7196445	−39.779913	−59.990944	−85.978530	
E_{el}^A (eV)[b]	−7.7215000	−39.778071	−59.987014	−85.967125	−63.649332

[a] Including training molecules with N−O bonds. [b] Excluding training molecules with N−O bonds.

conventional DFT methods[26a,28] and in a DFTB approach.[29] The resultant values for E_{el}^A are reported in Table 4 for SCC−DFTB.

An alternative would be to compute heats of formation for SCC−DFTB from the total computed energy (eq 4) with adjustments for the translational, rotational, vibrational and $P\Delta V$ changes. In view of the optimization of the E_{el}^A values here and prior experience,[26a,28,29] it is extremely unlikely that the latter procedure would yield a lower MAE. However, it should be kept in mind that the errors with the present approach for SCC−DFTB could be lowered if all parameters for the method were optimized simultaneously along with the E_{el}^A values to minimize the errors in heats of formation. It should also be noted that the atomic terms in eq 8 cancel in computing the ΔH of any reaction in view of the conservation of atoms, so for the SCC−DFTB results this is equivalent to equating the ΔH^{298} with the total energy change at 0 K. For the unimolecular processes in Table 1, this approximation had a 6% effect on the MAEs. A final point is that only the lowest-energy conformer for a molecule is considered in the present ΔH_f calculations, as usual.[26a] If there was uncertainty, a conformational search was performed. This approximation has been analyzed and normally yields errors of less than 1 kcal/mol for flexible molecules with fewer than 30 atoms and under 2 kcal/mol for cases with 50−60 atoms.[26a]

The resultant ΔH_f MAEs for AM1, PM3, PDDG/PM3, and SCC−DFTB on the combined training and test sets are given in Table 5, and complete documentation of the results for the 622 molecules is provided in the Supporting Information (SI). The MAE of 5.8 kcal/mol for SCC−DFTB is between those for AM1 (6.8 kcal/mol) and PM3 (4.4 kcal/mol) and significantly larger than for PDDG/PM3 (3.2 kcal/mol). It is known

TABLE 5: Mean Absolute Errors in Calculated Heats of Formation for Neutral Molecules Containing the Elements C, H, N, and O (kcal/mol)[a]

	N	AM1	PM3	PDDG/ PM3	SCC−DFTB
hydrocarbons	254	5.6	3.6	2.6	4.8
all molecules	622	6.8	4.4	3.2	5.8
training set	134	5.9	4.1	2.6	6.9
test set	488	7.0	4.4	3.4	5.5

[a] All data are in the Supporting Information.

that PM3 generally performs better for heats of formation than many DFT methods[28,30] and that PDDG/PM3 represents a significant improvement over PM3.[8−10]

An inspection of the principal sources of error reveals that SCC−DFTB predicts heats of formation for molecules containing one or more NO bonds to be too low. In the training set, the errors for nitromethane, nitroethane, 1-nitropropane, 2-nitropropane, methyl nitrate, and ethyl nitrate are 40−80 kcal/mol. When there is just a solitary N—O bond as in isoxazoles, the underestimate is less, 15−30 kcal/mol, so an additive effect is indicated. As discussed below, the computed structures of the molecules with N—O bonds are fine. It is likely that the SCC−DFTB E_{rep} could be better parametrized for N—O, so a second ΔH_{f} fit and testing were performed, which excluded all systems that contained both nitrogen and oxygen atoms. This lowered the MAE for the reduced set of 581 molecules to 4.4 kcal/mol, a result comparable to that from PM3 (4.2 kcal/mol on this reduced set). The MAE for the 254 hydrocarbons only changes from 4.8 to 4.7 kcal/mol. Owing to this issue, molecules with N—O bonds were also not included in Tables 2 or 3.

As with most SMO methods and as indicated in Tables 1−3, a problem class for SCC−DFTB is highly strained compounds, e.g., the ΔH_{f} values are too positive by 10−20 kcal/mol for most mono- and bicyclic molecules containing three-membered rings. Four-membered rings are handled comparatively well, and remarkably the ΔH_{f} error is only 8 kcal/mol for cubane. There are also some difficulties with another standard problem class, molecules with adjacent heteroatoms (e.g., azides, tetrazole, and ozone). Notably, the homologation and branching errors that are problematic for some NDDO methods are less severe with SCC−DFTB. For example, the ΔH_{f} error only increases from 1 to 7 kcal/mol in going from propane to decane with SCC−DFTB, whereas it is 1 to 13 kcal/mol with AM1. Also, the ΔH_{f} errors are only 0.5 and 1.8 kcal/mol for neopentane and 2,3,4-trimethylpentane with SCC−DFTB, whereas they are 7.4 and 6.1 with AM1, 4.3 and 3.4 with PM3, and 1.6 and 1.3 with PDDG/PM3. SCC−DFTB has relatively higher errors for nonaromatic molecules with multiple bonds, e.g., ethylene (7.0) and 1,3-butadiene (9.6), than the other SMO methods.

As mentioned above, SMO methods all have difficulties with some small molecules owing presumably to their often unique bonding characteristics including bond lengths. For example, H_2 is a problem for PM3 with an error of 13.4 kcal/mol and for PDDG/PM3 (22.2 kcal/mol), whereas it is handled comparatively well by AM1 (5.2) and SCC−DFTB (6.2). N_2 is a problem case for PM3 (17.6), PDDG/PM3 (9.5), and AM1 (11.2), whereas it is fine with SCC−DFTB (2.9). On the other hand, CO is not problematic for PDDG/PM3 (0.3 kcal/mol), whereas it is a problem for MNDO (20.5), AM1 (20.7), PM3 (6.7), and SCC−DFTB (12.2). Special cases like H_2 can, in general, be easily remedied by an addition to the CRFs, but overall the motivation for such finely targeted adjustments is not compelling.

TABLE 6: Heats of Formation of Ions and Radicals (kcal/mol)

	exp.[a]	AM1	PM3	PDDG/ PM3	SCC−DFTB
methyl cation	262.9	252.4	256.6	256.7	284.2
ethyl cation	215.6	216.8	222.5	220.3	235.4
isopropyl cation	190.9	191.9	197.3	192.9	202.0
tert-butyl cation	176.0	173.9	177.8	171.5	177.3
allyl cation	226.0	226.2	232.7	231.9	247.4
tropylium cation	203.0	210.5	221.0	222.4	222.2
benzyl cation	215.0	222.1	227.4	226.7	231.1
H_3O^+	139.0	143.5	159.1	157.1	134.5
CHO^+	199.0	187.5	176.9	175.4	186.1
NH_4^+	155.0	150.6	153.4	149.6	167.0
$CH_2NH_2^+$	178.0	176.3	185.3	182.4	192.0
methyl radical	34.8	31.3	29.8	25.9	59.0
ethyl radical	25.0	18.2	17.3	15.6	45.2
propyl radical	16.8	11.5	12.2	10.6	39.4
isopropyl radical	22.3	6.8	5.5	4.5	32.8
sec-butyl radical	17.0	0.2	0.3	−0.4	27.1
tert-butyl radical	11.0	−2.8	−5.9	−6.4	20.6
allyl radical	40.0	38.6	39.6	36.9	56.8
CHO radical	10.4	−1.0	−9.3	−11.6	16.6
methoxy radical	−0.5	−3.7	−6.8	1.3	13.2
methoxy anion	−33.2	−38.5	−37.9	−28.3	−23.5
ethoxy anion	−47.5	−45.5	−44.8	−38.4	−31.4
phenoxide anion	−40.5	−41.0	−44.1	−45.3	−34.9
methyl anion	33.2	57.7	51.5	43.3	57.2
ethyl anion	35.1	34.5	31.7	31.3	31.4
isopropyl anion	28.2	16.9	15.1	18.6	22.0
tert-butyl anion	16.0	2.7	−0.4	4.9	12.1
allyl anion	29.9	27.6	25.2	20.6	34.0
hydroxide ion	−32.7	−14.1	−17.5	−14.6	−3.5
methylamine anion	30.5	33.1	21.7	19.8	58.6
MAE		7.0	9.8	10.0	13.9

[a] Exptl. values as compiled in ref 8. Computed results here for radicals are from RHF calculations as in ref 7; the results in ref 8 for radicals are from UHF calculations.

The heats of formation for some ions and radicals, none containing N—O bonds, are listed in Table 6. All methods perform less well than on the set of neutral, closed-shell species. This arises from the increased variations in bonding and electrostatics and especially from the exclusion of radicals and ions in the training set. The MAE from SCC−DFTB, 13.9 kcal/mol, is 4−7 kcal/mol higher than from the other SMO methods [e.g., PDDG/PM3 (10.0 kcal/mol)]. It may be noted that the MAEs of 4.7 or 4.8 kcal/mol obtained here with SCC−DFTB for 254 hydrocarbons are very similar to the result of 4.6 kcal/mol obtained for 83 hydrocarbons by the DFTB method of Voityuk.[29] This is coincidental as he also reports MAEs excluding C_{60} of 7.0 and 5.6 for AM1 and PM3, which are 2−3 kcal/mol higher than for the hydrocarbon results in Table 5. The differences stem from Voityuk's inclusion of 15 radicals among the 83 hydrocarbons that were used in optimizing his DFTB method. As one can see in Table 6, there is a known[7b,8] systematic error with hydrocarbon radicals such that their heats of formation are underestimated with MNDO, AM1, PM3, and PDDG/PM3 by ca. 10 kcal/mol. So, (15 × 10 kcal/mol)/83 adds 2 kcal/mol to the MAEs. This comes under the discussion of the last section of overemphasizing one type of molecule or reaction, which is known to be a weak spot for a particular method.

Other Energetics. Some conformational energy differences and intermolecular interaction energies have also been compiled in Tables 7 and 8. These should be viewed as initial surveys in view of the limited sizes of the data sets. Molecules with N—O bonds were again excluded.

Fifteen prototypical conformational energy changes are considered in Table 7. In this case, SCC−DFTB with an MAE

TABLE 7: Conformational Energetics (kcal/mol)

	ref ΔE[a]	B3LYP/6-31G(d)	AM1	PM3	PDDG/PM3	SCC−DFTB
butane anti vs gauche	0.7	0.8	0.7	0.6	0.3	0.5
butane anti vs cis	5.7	5.7	3.3	4.0	3.9	3.8
ethane anti vs eclip.	2.8	1.6	1.2	1.4	1.1	2.3
methylcyclohexane eq. vs ax.	1.8	2.1	1.4	1.1	0.9	0.9
cis-1,3-dimethylcyclohexane eq,eq vs ax,ax	5.5	5.9	5.2	2.2	3.2	3.4
propene eclip. vs anti	2.0	2.1	0.6	0.7	0.7	1.1
1,3-butadiene trans vs skew	2.49	3.5	0.8	0.7	0.7	1.5
piperidine eq vs ax	0.53	0.3	−2.7	−2.2	−3.0	0.3
ethanol trans vs gauche	0.12	−0.3	−1.6	−1.8	−2.1	−0.2
ethyl methyl ether trans vs gauche	1.5	1.4	−0.4	−1.0	−1.8	0.8
propanal cis vs skew	0.95	1.2	−0.6	−0.7	−1.1	−4.4
butanone cis vs skew	1.07	1.6	0.1	0.1	0.3	0.3
acrolein trans vs cis	2.0	1.7	0.2	0.4	0.8	0.8
methyl formate Z vs E	4.75	5.1	5.6	1.9	1.8	2.7
N-methylacetamide Z vs E	2.3	2.5	1.6	0.4	1.9	2.2
MAE		0.4	1.4	1.8	1.8	1.2

[a] Values from ref 12 and 31.

TABLE 8: Intermolecular Interaction Energies (kcal/mol)

	ab initio[a]	AM1	PM3	PDDG/PM3	SCC−DFTB
$H_2O\cdots H_2O$	−4.80	−5.4	−3.5	−3.7	−3.3
$H_2O\cdots MeOH$	−4.90	−5.1	−1.7	−2.6	−3.1
$H_2O\cdots Me_2O$	−5.51	−2.2	−1.7	−3.4	−2.6
$MeOH\cdots MeOH$	−5.45	−2.7	−1.7	−3.2	−3.4
$H_2O\cdots H_2CO$	−5.17	−2.7	−1.5	−1.8	−2.4
$HCOOH\cdots HCOOH$	−13.93	−6.4	−8.6	−4.0	−13.4
$NH_3\cdots NH_3$	−2.94	−1.4	−1.4	−1.3	−1.9
$NH_3\cdots H_2O$	−6.36	−2.0	−1.5	−1.8	−3.4
$HCONH_2\cdots H_2O$	−8.88	−7.5	−2.2	−3.6	−5.4
$HCONH_2\cdots HCONH_2$	−13.55	−8.1	−3.3	−5.4	−11.8
$H_2O\cdots CH_3NH_3^+$	−18.76[b]	−13.1	−12.7	−13.2	−17.5
$H_2O\cdots CH_3CO_2^-$	−19.22[b]	−15.9	−15.8	−16.2	−19.8
MAE		3.2	4.5	4.1	1.9

[a] CCSD(T) results from ref 32. [b] RHF/6-31G(d) BSSE result from ref 33.

of 1.2 kcal/mol performs better than the NDDO schemes by 0.2−0.6 kcal/mol. RM1 was also checked and it yields results very similar to AM1 including an MAE of 1.5 kcal/mol for this data set. Some improvement with SCC−DFTB could be expected since its inclusion of the overlap matrix in the secular equations reintroduces four-electron "Pauli repulsions", which should improve the description of rotational barriers.[12,32] Nevertheless, all four SMO methods generally err by underestimating the energy of the less stable form. There are also some qualitative errors, particularly for propanal and piperidine. Propanal is problematic for all of the SMO methods since they uniformly find the cis conformation to be erroneously higher in energy than the skew form. The greater error with SCC−DFTB may reflect a specific problem in the E_{rep} for C−O or H−O. For conformational energetics, the B3LYP/6-31G(d) results are clearly superior to the SMO methods with only one minor qualitative error (ethanol) and an MAE of only 0.4 kcal/mol.

The prediction of intermolecular interaction energies is generally a weak point for the NDDO-based methods,[7,8] though PDDG/PM3 does well for halide ion−molecule affinities.[9] It has been pointed out previously that PM3 describes poorly intermolecular Coulombic interactions and relies on the deformation of the electronic charge density to provide net binding.[35,36] The results in Table 8 illustrate the pattern for neutral systems. For the weak complexes with interaction energies less than 7 kcal/mol, the NDDO-based methods and SCC−DFTB typically underestimate the interactions by ca. 50%. SCC−DFTB does significantly better for the cyclic dimers of formic acid and formamide, which primarily accounts for its improved MAE (1.9 kcal/mol) over the MAEs of 3.2, 4.5, and 4.1 kcal/

TABLE 9: Mean Absolute Errors for Additional Molecular Properties of CHNO-Containing Species

	N	AM1[a]	PM3[a]	PDDG/PM3[a]	SCC−DFTB
bond lengths (Å)	218	0.017	0.012	0.013	0.012
bond angles (deg.)	126	1.5	1.7	1.9	1.0
dihedral angles (deg.)	30	2.8	3.2	3.7	2.9
dipole moments (Debye)	47	0.23	0.25	0.23	0.39

[a] Values from ref 12.

mol for AM1, PM3, and PDDG/PM3. RM1 was also checked for the complexes in Table 8; its MAE is 5.0 kcal/mol, and now the weak complexes are all too well bound, e.g., the ΔE for methanol dimer is −16.4 kcal/mol. In this area, all SMO methods perform far less well than classical force fields that describe intermolecular interactions using Coulombic interactions between atoms with point charges plus, typically, Lennard-Jones potentials.[33,37] SMO methods are not suitable for simulations of condensed-phase systems such as liquid water or alcohols. Thus, a standard approach for QM/MM simulations of reactions in solution is to use SMO calculations to describe the reacting system and to treat the intermolecular interactions with a force field using partial charges for the reacting system derived from the SMO wavefunction.[38]

Structures and Dipole Moments. In addition to the energetic data, a comparison of molecular geometries and dipole moments has been compiled in Table 9. For the molecular structures, the comparisons are made to the recent collection of MP2/cc-pVTZ results.[12] The SMO methods all yield fine molecular structures with average errors of 0.01−0.02 Å for bond lengths, 1−2° for bond angles, and ca. 3° for dihedral angles. Overall, the SCC−DFTB results best reproduce the reference MP2/cc-pVTZ structural parameters. A detailed inspection shows no major problems. Of the 218 bond lengths considered here, none deviate by more than 0.05 Å from the reference values. Concerning N—O bonds, the results are within 0.02 Å for all relevant molecules, nitromethane, ethyl nitrate, and nitric acid. However, CN triple bonds are consistently 0.030−0.040 Å too short, whereas CH bonds in aldehyde groups are 0.035−0.045 Å too long.

Bond angles and dihedral angles determined by SCC−DFTB are similarly excellent. Only three out of 126 bond angles (HCC in bicyclobutane, NNH in hydrazine, and HON in nitric acid) differ from the reference structures by more than 5° and all are within 10°. As with NDDO methods, particularly MNDO, the largest errors arise when the central atom is oxygen, but SCC−DFTB still gets these correct to within 2−3°. The CCCO dihedral angle of skew propanal is the only outlier among the 34 dihedral angles tested; SCC−DFTB finds it to be 149.5° versus 122.3° from MP2/cc-pVTZ. This structure is also an outlier for the NDDO methods, with AM1, PM3, and PDDG/PM3 yielding 131.0°, 130.6°, and 131.3°, respectively.

Finally, the MAEs for the dipole moments for 47 molecules in comparison with gas-phase experimental values are listed in Table 9. As with the structures, the SMO results are all reasonable, though the MAE is 0.15 D larger with SCC−DFTB than the other methods. The largest errors for SCC−DFTB occur for nitrogen-containing molecules, where the calculated dipole moments are substantially smaller than the experimental ones, e.g., the errors are 1.1 D for acetonitrile and pyridine and ca. 0.5 D for amines. These results are consistent with the more extensive study of dipole moments by Kalinowski et al. in the course of their development of a CM3 charge model for SCC−DFTB.[39]

TABLE 10: Mean Absolute Errors for Sulfur-Containing Species

	N	AM1	PM3	PDDG/PM3	SCC−DFTB
training set MAE (kcal/mol)	81	7.5	7.1	5.6	13.7
test set MAE (kcal/mol)	143	11.3	12.3	5.6	20.8
overall MAE (kcal/mol)	224	9.9	10.4	5.6	18.2
bond lengths (Å)	61	0.06	0.08	0.05	0.07
bond angles (deg.)	42	5.1	8.0	5.2	6.3
dipole moments (Debye)	24	0.34	0.43	0.49	0.38

Performance for Sulfur. Since SCC−DFTB has also been parametrized for sulfur,[20] additional comparisons were pursued using the sulfur-containing set of ref 10. Halogen-containing species were excluded in view of the lack of SCC−DFTB parameters, and transition states were excluded owing to the absence of an automated procedure for their location in the SCC−DFTB code. For heats of formation, the electronic energy for sulfur was determined by minimizing the errors for heats of formation as described above using a training set of 81 sulfur-containing molecules, whereas the $E_{\mathrm{el}}^{\mathrm{A}}$ values for H, C, N, and O were kept fixed at the results from the second fit, which excluded compounds with both N and O atoms (Table 4). The resulting ΔH_f MAEs for the SMO methods are compared in Table 10. The results are substantially worse for SCC−DFTB with an MAE of 18.2 kcal/mol than for any of the NDDO methods.

Similar to the problem with N—O bonds, S—O bonds yield energies which are too low. Geometries are not an issue; the structural results reported in Reference 20 were reproduced here including for SO_2 and dimethylsulfate. The energy errors for S—O containing molecules typically range from 30 to 60 kcal/mol, except, curiously, for ethynesulfenic acid, benzenesulfenic acid, and methanesulfenic acid, which have differences from the experimental heats of formation of only 4.3, 7.3, and 2.8 kcal/mol. The heats of formation of molecules containing both sulfur and oxygen but without S—O bonds, however, are adequate. For example, the errors for thioacetic acid and S-ethyl thioacetate are less than 5.0 kcal/mol. Summaries of the structural results and dipole moments are also included in Table 10. The performance of the SMO methods is similar, and there are no striking problems. The somewhat larger errors with SCC−DFTB and PM3 for bond lengths are largely due to contraction of the SN distance for the ammonia−SO_2 and dimethylamine−SO_2 complexes by about 0.6 to near 2.0 Å.

Conclusions

The present comparisons of results from SCC−DFTB and the NDDO-based SMO alternatives have provided a clearer picture of the strengths and weaknesses of the methods. The focus was on energetics of reactions, which are well represented by isomerization energies and heats of formation. The energetic errors for reactions of CHNO-containing molecules with SCC−DFTB are similar to those from PM3 and substantially greater than those from PDDG/PM3, which also outperforms B3LYP/6-31G(d) calculations. Besides known branching problems for DFT methods, the present results with B3LYP/6-31G(d) indicate that there are additional deficiencies for bridged, bicyclic molecules and numerous simple bond-switching processes such as carboxylic acid to ester, and diol to peroxide. As a final comparison among the NDDO methods including the newly reparameterized version of AM1 (RM1[11]).

Table 11 lists the MAEs for heats of formation for the full dataset[8−10] of 1356 molecules, radicals, ions, and complexes

TABLE 11: Mean Absolute Errors for ΔH_f from NDDO-Based Methods for All Molecules, Radicals, Ions, and Complexes

		ΔH_f MAE (kcal/mol)			
	N	AM1	PM3	RM1	PDDG/PM3
All	**1356**	**9.2**	**7.2**	**5.5**	**5.0**
CH	254[a]	5.6	3.6	3.8	2.6
CHNO	622[a]	6.8	4.4	4.0	3.2
+ FClBrI	442[b]	11.1	8.1	5.6	5.6
+ S	249[c]	10.6	10.5	6.4	6.4
+ P	43[c]	18.2	21.5	20.9	17.9

[a] Dataset from ref 8. [b] Dataset from ref 9. [c] Dataset from ref 10.

containing the elements C, H, N, O, P, S, F, Cl, Br and I; silicon-containing systems were excluded since RM1 parameters for Si have not been reported. PDDG/PM3 gives the lowest overall MAE (5.0 kcal/mol), and it is not out-performed for any subset, whereas RM1 does represent a significant improvement over AM1 and PM3.

The quality of molecular structures from the SMO methods including SCC−DFTB is similar and high. Intermolecular energetics need more analyses, but the description of hydrogen bonds weaker than ca. 7 kcal/mol is poor from all of the SMO methods. Conformational energy differences are notably improved with SCC−DFTB over the NDDO-based methods, though B3LYP/631G(d) greatly outperforms all SMO methods in this case. Evaluations for sulfur-containing molecules and to a lesser extent for nitrogen-containing ones were hampered by apparent errors in the treatment of N−O and S−O bonds with SCC−DFTB. Other issues that make SCC−DFTB less attractive at this time than PM3 or PDDG/PM3 are the lack of parameters for additional atom types including halogens, the larger number of parameters in the core repulsion formulas, and their quadratic rather than linear growth with addition of new atom types. However, the present results emphasize that further improvements can be made to all of the SMO methods, and it is auspicious that their development is seeing resurgence and added variety.

Acknowledgment. The authors thank Dr. M. Elstner for the SCC−DFTB code used in this work, Drs. G. Zheng and I. Tubert-Brohman for discussions and assistance, and the National Science Foundation (CHE-0446920) for financial support.

Supporting Information Available: Excel file for the computed and experimental data, and full citation for ref 24. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Zerner, M. C. In *Semiempirical Molecular Orbital Methods*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991; Vol. 2, p 313−365.

(2) Stewart, J. J. P. In *Semiempirical Molecular Orbital Methods*; Lipkowitz, K. B., Boyd; D. B., Eds.; VCH Publishers: New York, 1991; Vol. 1, p 45−81.

(3) Pople, J. A.; Beveridge, D. L. In *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970.

(4) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.

(5) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907−4917.

(6) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(7) (a) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220. (b) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.

(8) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601−1622.

(9) Tubert-Brohman, I.; Guimarães, C. R. W.; Repasky, M. P.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 138−150.

(10) Tubert-Brohman, I.; Guimarães, C. R. W.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2005**, *1*, 817−823.

(11) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101−1111.

(12) Sattelmeyer, K. W., Tubert-Brohman, I.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2006**, *2*, 413−419.

(13) (a) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298−305. (b) Monard, G.; Merz, K. M., Jr. *Acc. Chem. Res.* **1999**, *32*, 904−911. (c) Warshel, A. *Ann. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443. (d) Ridder, L.; Mulholland, A. J. *Curr. Top. Med. Chem.* **2003**, *3*, 1241−1256.

(14) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947−12957.

(15) Seifert, G.; Porezag, D.; Frauenheim, T. *Int. J. Quantum Chem.* **1996**, *58*, 185−192.

(16) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260−7268.

(17) (a) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69−89. (b) Hall, G. G. *Proc. R. Soc. London Ser. A* **1951**, *205*, 541−552.

(18) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397−1412.

(19) Streitwieser, A. *J. Am. Chem. Soc.* **1960**, *82*, 4123−4135.

(20) Niehaus, T. A.; Elstner, M.; Frauenheim, T.; Suhai, S. *Theochem.* **2001**, *541*, 185−194.

(21) Krüger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. *J. Chem. Phys.* **2005**, *122*, 114110.

(22) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689−1700.

(23) (a) Stewart, J. J. P. *J. Comput. Aided Mol. Des.* **1990**, *4*, 1−105. (b) Code and parameter files to modify MOPAC for PDDG/PM3 calculations are available at http://www.jorgensenresearch.com

(24) Frisch, M. J.; Gaussian03, Revision B.03; Gaussian, Inc.: Pittsburgh PA, 2003 [Full reference given in the Supporting Information].

(25) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 12; Ed.: Johnson, R. D., III, Ed.; NIST: Gathersburg, MD, 2005; http://srdata.nist.gov/cccbdb

(26) (a) Stewart, J. J. P. *J. Mol. Model.* **2004**, *10*, 6−12. (b) Grimme, S. *Angew. Chem. Int. Ed.* **2006**, *45*, 4460−4464.

(27) (a) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063−1079. (b) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374−7383.

(28) (a) Csonka, G. I.; Ruzinsky, A.; Tao, J.; Perdew, J. P. *Int. J. Quantum Chem.* **2005**, *101*, 506−511. (b) Brothers, E. N.; Scuseria, G. E. *J. Chem. Theory Comput.* **2006**, *2*, 1045−1049.

(29) Voityuk, A. A. *J. Chem. Theory Comput.* **2006**, *2*, 1038−1044.

(30) Brothers, E. N.; Merz, K. M., Jr. *J. Phys. Chem. A* **2004**, *108*, 2904−2911.

(31) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730−748.

(32) Tsuzuki, S.; Lüthi, H.-P. *J. Chem. Phys.* **2001**, *114*, 3949−3957.

(33) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520−552.

(34) (a) Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495−506. (b) Mohle, K.; Hofmann, H.-J.; Thiel, W. *J. Comput. Chem.* **2001**, *22*, 509−520.

(35) Cummins, P. L.; Titmuss, S. J.; Jayatilaka, D.; Bliznyuk, A. A.; Rendell, A. P.; Gready, J. E. *Chem. Phys. Lett.* **2002**, *352*, 245−251.

(36) (a) van der Vaart, A.; Merz, K. M. J. *J. Phys. Chem. A* **1999**, *103*, 3321−3329. (b) van der Vaart, A.; Merz, K. M. J. *J. Am. Chem. Soc.* **1999**, *121*, 9182−9190.

(37) Hobza, P.; Kabelac, M.; Sponer, J.; Mejzlik, P.; Vondrasek, J. *J. Comput. Chem.* **1997**, *18*, 1136−1150.

(38) For recent examples and more references, see: (a) Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2005**, *127*, 8829−8834. (b) Acevedo, O.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2006**, *128*, 6141−6146.

(39) Kalinowski, J. A.; Lesyng, B.; Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2545−2549.