

Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules

Julian Tirado-Rives* and William L. Jorgensen*

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

Received September 24, 2007

Abstract: Testing of the commonly used hybrid density functional B3LYP with the 6-31G(d), 6-31G(d,p), and 6-31+G(d,p) basis sets has been carried out for 622 neutral, closed-shell organic compounds containing the elements C, H, N, and O. The focus is comparison of computed and experimental heats of formation and isomerization energies. In addition, the effect of an empirical dispersion correction term has been evaluated and found to improve agreement with the experimental data. For the 622 compounds, the mean absolute errors (MAE) in the heats of formation are 3.1, 2.6, 2.7, and 2.4 kcal/mol for B3LYP/6-31G(d), B3LYP/6-31G(d,p), B3LYP/6-31+G(d,p), and B3LYP/6-31+G(d,p) with the dispersion correction. A diverse set of 34 isomerizations highlights specific issues of general interest, such as performance on differences in steric effects, conjugation, and bonding. The corresponding MAEs for the isomerizations are 2.7, 2.4, 2.2, and 1.9 kcal/mol. Improvement is obtained for isomerizations of amines and alcohols when both polarization and diffuse functions are used, but the over stabilization of linear alkanes compared to branched isomers can be relieved only with the dispersion correction. Besides the insights on DFT methods, the study also aimed to quantify the gains in accuracy that can be achieved by replacing energetics from NDO-based semiempirical methods with DFT results. Since the MAEs obtained with the PDDG/PM3 method for the 622 heats of formation and 34 isomerizations are 2.8 and 2.3 kcal/mol, negligible advantage in accuracy for the B3LYP-based methods emerged in the absence of the dispersion corrections.

Introduction

Density functional methods have been gaining popularity in the past decade. They scale similarly to Hartree–Fock (HF) methods, though they account for some electron correlation effects.^{1,2} Among the ever increasing number of DFT methods, the hybrid functional B3LYP,³ first developed to study vibrational absorption and circular dichroism, has emerged as a good compromise between computational cost, coverage, and accuracy of results. It has become a standard method used to study organic chemistry in the gas phase. It is important, particularly for methods with empirical components, to assess performance on large sets of compounds for which reliable reference data are available. Results of such comparisons should be examined not only for overall

accuracy, usually expressed in statistical terms such as the mean average error (MAE or MAD) or root-mean-square (rms) deviation, but since these overall metrics can be dominated by a few large errors, performance should also be evaluated for specific cases of interest. The present study concentrates on comparisons for a diverse database of 622 neutral, closed shell organic compounds containing only C, H, N, and O, which was gathered in the development of the PDDG/PM3 semiempirical molecular orbital method.^{4–6} Particular comparisons extracted from this database include 34 isomerization reactions that highlight difference in bonding, conjugation, and steric effects of general interest to organic chemists. Besides the insights on the performance of DFT methods, this study was also aimed to address the limits of accuracy that can be expected from semiempirical quantum methods. Specifically, the goal was to quantify the possible gains in accuracy that can be achieved by replacing

* Corresponding author e-mail: julian.tirado-rives@yale.edu (J.T.-R.) and william.jorgensen@yale.edu (W.L.J.).

the semiempirical energetics with DFT results in the computation of heats of formation for the large data set and for the isomerization energies.

Recently, Riley et al.⁷ published an extensive evaluation of 37 DFT methods with 11 different basis sets in which they compared their overall performance on structures, heats of formation, ionization potentials, electron affinities, vibrational frequencies, conformational and hydrogen-bonding energies, and reaction barriers from the G2/97 test set.⁸ However, the size of the molecules was limited as the G2/97 test set contains only compounds with up to 6 non-hydrogen atoms. It was found that B3LYP ranks consistently well, although not at the top for any one category. Of particular interest is the fact that in calculating many of these properties the B3LYP/6-31G(d) combination gives better results than 6-31G+(d) or 6-31G++G(d) and results of comparable quality such as the larger Dunning-type basis sets, cc-pVDZ, cc-pVTz, aug-cc-pVDZ, or aug-cc-pVTZ. Similar results were noted in the original B3LYP publication.³ This characteristic may be interpreted as giving the right answer for the wrong reason or, more optimistically, as an economical, balanced description. Nevertheless, there are still some significant problems that have been receiving attention, particularly, decreasing accuracy with increasing system size, understabilization of branched vs linear alkanes, and underestimation of weak interactions.

The problem with increasing system size undermines one of the reasons for the popularity of DFT methods, i.e., the scaling of computational effort with the number of basis functions (N) is N^4 for B3LYP, while common alternatives such as MP2 scale as N^5 , and MP4 or G2 as N^7 . Curtiss et al. made a general comparison of B3LYP/6-311+G(3df,2p) single-point calculations using MP2(FULL)/6-31G(d) structures and 0.89 scaled HF/6-31G(d) frequencies when developing the G2/97 data set (148 compounds with 1–6 non-hydrogen atoms) and found that it gives a MAE of 3.11 kcal/mol (the G2 method gives 1.58).⁸ Further comparisons using the compounds added to create the newer G3/99 data set (75 additional compounds with 5–10 heavy atoms) using B3LYP/6-311+G(3df,2p) single-point calculations on B3LYP/6-31G(d) geometries with 0.96 scaled B3LYP/6-31G(d) frequencies gave a MAE of 8.21 kcal/mol (G3 gives 1.01).⁹ The net is a MAE of 4.8 kcal/mol with B3LYP for the full G3/99 data set of 223 heats of formation.

Similar evaluations and comparisons have been carried out in the development of other methodologies. For instance, in the development of the X3LYP extended functional, Xu et al.¹⁰ reported a MAE of 3.1 kcal/mol for B3LYP/6-311+G(3df,2p) on the G2/97 set, while the X3LYP functional yielded 2.8 kcal/mol. Kafafi developed the K2-BVWN functional, which includes the use of fitted atom equivalents; for a larger set of 350 atoms and molecules including the G2 set (nitrobenzene being the largest compound), B3LYP/6-311+G(3df,2p)/B3LYP/6-31G(d) yielded a MAE of 3.0 kcal/mol, which was reduced to 1.4 kcal/mol with K2-BVWN.¹¹ A significantly larger basis set was also used in the comparison of the spin-component scaled (SCS) variant of MP2 with B3LYP by Grimme on an extended G2/97' set that included 160 molecules with 6 or fewer heavy atoms as

a training set and 70 additional molecules, including some larger compounds such as naphthalene, azulene, and anthracene, as a test set.¹² The MAEs obtained using B3LYP/QZV3P for the training and test sets were 2.1 and 8.5 kcal/mol, while an extrapolation from TZV2P and QZV3P to obtain close to the basis-set-limit treatment for MP2 produced 1.7 and 4.6 kcal/mol and 1.2 and 2.8 kcal/mol for SCS-MP2. It should be noted that optimized atomic energy corrections were used in these calculations.¹² The system-size issue was also raised in the original report of the PDDG/PM3 method where it was found that the MAE with B3LYP/6-311+G(3df,2p) for heats of formation increases from 2.2 kcal/mol for 60 neutral molecules in the G2 set to 7.2 kcal/mol for 45 molecules in the G3 set, while the MAEs from PDDG/PM3 are 3.2 kcal/mol for both sets.⁴

Concerning linear hydrocarbons, Redfern et al. found that the errors in the calculated heats of formation for the alkanes in the G3 set increase steadily with system size from 0.6 to 30.3 kcal/mol for ethane to hexadecane at the B3LYP/6-311+G(3df,2p)/MP2/6-31G(d) level.¹³ They further noted that using isodesmic reactions to calculate the heats of formation decreases the errors significantly, but the errors become small only with the use of homoisodesmic reactions. For the largest case, hexadecane, the error from the direct calculation of 30.3 kcal/mol is reduced to 16.6 and 2.0 kcal/mol with the isodesmic and homoisodesmic approaches. Wodrich et al. compared experimental bond separation energies with values computed using several functionals, MP2 and CCSD(T), and they found similar systematic errors.¹⁴ The discrepancies for alkane homologation were attributed to deficiencies in the DFT description of “proto-branching” (1,3-interactions),^{14a} which is consistent with the improved performance with the homoisodesmic reactions.

Related problems have emerged in computing the energy differences between branched hydrocarbons and their linear isomers, with the additional characteristic that by their nature these are all isodesmic reactions. The energetic benefits of branching are significantly underestimated with B3LYP. Grimme provided an interesting analysis by comparing SCS-MP2, MP2, B3LYP, and B2PLYP results for isomers of butane, pentane, octane, and undecane that pointed to deficiencies in pair correlation energies between orbitals of the same type, e.g., 1,3-interactions.¹⁵ The same methodology was utilized to study C and Si based chains substituted with H, F, and Cl with similar results.¹⁶

The application of B3LYP to intermolecular interactions also reveals some problems. While hydrogen-bond energies of organic compounds with water are reproduced well,¹⁷ other weakly bonded complexes can be problematic. Tsuzuki and Lüthi found differences in the interaction energies of all three benzene dimers (T-shaped, stacked, or parallel-displaced) with reference CCSD(T)/aug-cc-pVDZ calculations that range from 3.7 to 5.5 kcal/mol; strikingly, the benzene dimer is not bound with B3LYP methods.¹⁸ This deficiency prompted the development of several new functionals. For instance, Zhao and Truhlar developed the PWB6K and PW6B95 alternatives¹⁹ and later M05-2X²⁰ based on a training set that included 115 atomization energies, 13 ionization potentials (IPs), 13 electron affinities (EAs), 76

barrier heights, and 51 complexation energies. The MAE per bond for the atomization energies was 0.92 kcal/mol with B3LYP and 0.41 kcal/mol with PWB6K. The MAEs for heats of formation from M05-2X for the G2 and G3 data sets were later calculated to be 3.2 and 4.2 kcal/mol, respectively.²¹ Development of the B2PLYP and mPW2PLYP functionals by Schwabe and Grimme followed a different approach by including a second-order Møller-Plesset type correction.²² Their training set, expanded from the G3/05 set, included 271 heats of formation, 105 IPs, 63 EAs, 10 proton affinities, and 6 hydrogen-bonded complexes. The MAE for the 271 heats of formation in this set with B3LYP/TZV2P is 5.3 kcal/mol, while their functionals gave 2.7 kcal/mol for B2PLYP/CQZV3P and 2.0 kcal/mol for mPW2PLYP/CQZV3P.

An alternative solution to the deficiencies in the description of π - π and other weakly bound complexes is to apply empirical dispersion corrections based on a scaled C_6/R_{ij}^6 interaction with a distance-dependent damping function. The scaling factors, parameters, combinatory rule, and damping function needed to improve the BLYP, B3LYP, and B97 functionals were determined by Grimme and co-workers based on the interaction energies and distances of increasing number of complexes (29,²³ 40,²⁴ and 161^{25,26}). A comparison of the 148 heats of formation in the G2/97 set using the modified B3LYP functional with TZV2P or TZVPP basis sets gave a MAE of 3.6 kcal/mol. This represents a slight decrease in accuracy for the isolated molecules, but it improves significantly the results for complexes. In a recent publication, addition of this correction to the semiempirical AM1 and PM3 methods was found to improve their performance on weakly bound complexes, when coupled with adjustments for the values of the parameters U_{ss} , β_s , and α for hydrogen and U_{ss} , U_{pp} , β_s , β_p , and α for carbon, nitrogen, and oxygen.²⁷ However, the impact of these modifications on computed heats of formation of neutral isolated molecules was not reported. A very similar method with different combinatory rules was optimized for 22 complexes and verified for 58 additional ones by Jurečka et al.²⁸ Interestingly, adding this correction is more beneficial when larger basis sets or counterpoise corrections are used. Both of these methods provide average improvements of ca. 1.0 kcal/mol for complexes.

Computational Methods

Heats of Formation. All B3LYP calculations were performed with Gaussian03 using the 6-31G(d), 6-31G(d,p), and 6-31+G(d,p) basis sets.²⁹ The PDDG/PM3 calculations were executed with BOSS³⁰ or a local version of MOPAC.^{31,32} Unless indicated otherwise, all energies were obtained using geometries optimized with SCC-DFTB³³ calculations in a prior assessment of it and other semiempirical methods.³⁴ As noted, the molecular structures from SCC-DFTB are generally in excellent agreement with MP2/cc-pVTZ results, showing average deviations for bond lengths, bond angles, and dihedral angles of 0.012 Å, 1.0°, and 2.9°.³⁴ Optimizations were subsequently carried out for the full set of 622 molecules at the B3LYP/6-31G(d) level to assess the effects of the choice of structures on MAE. Comparison with the

SCC-DFTB geometries showed few significant differences except in several cases mostly involving C=C angles in alkenes. The B3LYP/6-31G(d) geometries were used in these cases, as noted in the tables. Since the time of the earlier comparative study,³⁴ further optimization of the PDDG/PM3 parameters for C, H, N, and O was also performed on the full training set of molecular structures and properties;⁴ when the reoptimized parameters are applied to the 622 heats of formation, the MAE is reduced from 3.2 to 2.8 kcal/mol. Though the modifications were slight, both the old and the modified PDDG/PM3 parameters are provided in the Supporting Information. The present results use the modified parameters.

Heats of formation at 298 K were computed using the same procedure followed in the development of many semiempirical quantum methods including PDDG/PM3.⁴ Specifically, eq 1 is used where E_{el}^A (also known as *eisol*) and ΔH_f^A are the electronic energies and heats of formation of the

$$\Delta H_f^0 = E_{tot}^{mol} - \sum_A E_{el}^A + \sum_A \Delta H_f^A \quad (1)$$

atoms composing the molecule. The experimental values are used for ΔH_f^A , and the E_{el}^A values are treated as adjustable parameters, which were optimized to minimize the unsigned errors for the 622 heats of formation for the HF and DFT-based procedures that were examined here. This approach is used to remove systematic errors in the atomic component of the energy change, and it is expected to yield lower MAEs than alternative computation of heats of formation from the total energy with adjustments for the translational, rotational, vibrational, and $P\Delta V$ changes. The same method has been used in other recent studies to compute heats of formation with conventional DFT methods,^{35–37} SCC-DFTB,³⁴ and in a DFTB approach.³⁸ An additional point is that only the lowest-energy conformer for a molecule is considered in the calculation of the heat of formation (eq 1). As discussed elsewhere,^{34,35} this approximation normally yields errors below 1 kcal/mol for flexible molecules with fewer than 30 atoms. The resultant values for E_{el}^A are reported in Table 1 for all methods considered here.

Dispersion Correction. A promising approach to reduce the errors of DFT proposed by Grimme is the inclusion of an empirical dispersion correction based on eq 2, where $f(R_{ij})$ is an exponential damping function and s_6 is a functional- and basis set-dependent scaling factor.^{23,24} In order to test the possible effects of such a dispersion term, a modified version of the BOSS program³⁰ was used to calculate E_{disp} for all 622 structures, which was then added to the total energies.

$$E_{disp} = -s_6 \sum_i \sum_{j>i} f(R_{ij}) \frac{C_6^{ij}}{R_{ij}^6} \quad (2)$$

Several different variants of this method were tried. In all cases, reoptimization of the scaling factor and the atomic electronic energies, E_{el}^A , in eq 1 was carried out. The first attempt, using Grimme's original C_6 parameters and damping

Table 1. Atomic Heats of Formation and Fitted Electronic Energies for Eq 1

		H	C	N	O
ΔH_f^A (kcal/mol)	all	52.102	170.89	113.00	59.559
E_{el}^A (eV) ^a	HF/6-31G(d)	-13.267830	-1023.471108	-1477.085945	-2032.656503
E_{el}^A (eV) ^a	B3LYP/6-31G(d)	-13.841952	-1029.922686	-1485.365964	-2042.501380
E_{el}^A (eV) ^b	B3LYP/6-31G(d)	-13.846915	-1029.922295	-1485.368385	-2042.523239
E_{el}^A (eV) ^a	B3LYP/6-31G(d,p)	-13.879462	-1029.924616	-1485.414750	-2042.512214
E_{el}^A (eV) ^a	B3LYP/6-31+G(d,p)	-13.872636	-1029.970795	-1485.524958	-2042.646009
E_{el}^A (eV)	B3LYP/6-31+G(d,p)+OC6	-13.884916	-1029.966281	-1485.509729	-2042.662279
E_{el}^A (eV) ^c	PDDG/PM3	-13.248092	-112.796924	-157.693757	-294.712153

^a Using single-point energy evaluations on SCC-DFTB optimized geometries (ref 34). ^b Using B3LYP/6-31G(d) optimized geometries.

Table 2. Mean Absolute Errors in Calculated Heats of Formation for Neutral Molecules Containing the Elements C, H, N, and O (kcal/mol)^a

		HF			B3LYP			
	N	6-31G(d) ^b	6-31G(d) ^b	6-31G(d) ^c	6-31G (d,p) ^b	6-31+ G(d,p) ^b	6-31+G(d,p) +OC6 ^b	PDDG/PM3
all molecules	622	3.9	3.1	3.0	2.6	2.7	2.4	2.8
hydrocarbons	254	2.5	2.2	2.1	2.2	2.5	2.1	2.2
"training set" ^d	134	4.3	3.0	2.9	2.4	2.3	1.9	2.4
"test set" ^d	488	3.8	3.1	3.0	2.7	2.9	2.5	2.9

^a All data are in the Supporting Information. ^b Using single-point energy evaluations on SCC-DFTB optimized geometries (ref 34). ^c Using B3LYP/6-31G(d) optimized geometries. ^d "Training" and "test" sets refer to those used in the development of the PDDG/PM3 parameters (ref 4). All 622 data points were used here in optimizing the E_{el}^A values in Table 1.

function for all atom pairs, produced less than 0.1 kcal/mol improvements for the heats of formation. This lack of improvement may stem from the use of fitted atomic energies E_{el}^A in eq 1. Although a full comparison would require frequency calculations for all 622 molecules, the subset of 97 compounds common to the G3/99 set has a MAE of 4.38 kcal/mol using the published B3LYP/6-311+G(3df,2p) energies and frequencies, while our results at the B3LYP/6-31G(d) level using eq 1 have a MAE of 2.94 kcal/mol. This improvement is numerically similar to that found with previous implementations of the dispersion correction.^{24–26} Different partitions of the corrections depending on bonding arrangements (bonded pairs, pairs bonded to a common atom, pairs separated by two bonds, etc.) and different sets of parameters were tested. The best results were obtained using OPLS-AA generic C, H, N, and O Lennard-Jones parameters ($C_6^i = 4\epsilon_{ij}O_{ij}^6$),³⁹ no damping function ($f = 1$), a single scale factor s_6 of 0.985, and including only the interactions between atom pairs separated by more than three bonds. E.g., for butane, there would be no C–C term, 6 C–H terms, and 21 H–H terms. The resultant MAEs did improve with increasing basis-set size. Thus, the best results were obtained at the 6-31+G(d,p) level, and these are designated as B3LYP/6-31+G(d,p)+OC6 in the tables. Specifically, the ΔH_f MAE for the 622 compounds of 3.0 kcal/mol at the B3LYP/6-31G(d)+OC6 level decreases to 2.5 kcal/mol with the added p-functions in 6-31G(d,p) and to 2.4 kcal/mol with the addition of the diffuse functions. A full tabulation of results is presented in the Supporting Information.

The investigation of the present dispersion correction should be viewed as just a test. It is unsatisfactory from our perspective since it is not general in that it requires assigning atom pairs as bonded or not. This creates ambiguities or discontinuities for application to transition structures or

complexes. However, one could devise a closely related continuous version using an appropriate damping function in eq 2.

Results and Discussion

Heats of Formation. The mean absolute errors for the different methods are recorded in Table 2. The second column contains the number of structures in the molecule sets of the first column. The "training" and "test" sets refers to those used in the original PDDG/PM3 parametrization;⁴ this is an informative separation, since the 134 compounds in the "training" set contain 1–10 non-hydrogen atoms (the average is 5.0), while the 488 in the "test" set are somewhat larger, from 2 to 15 with an average of 7.4 non-hydrogen atoms.

There are several points to note. (1) The B3LYP-based results are uniformly better than those from HF/6-31G(d). (2) The 5–7 kcal/mol MAEs that are regularly found for B3LYP-based heats of formation are readily reduced to 3 kcal/mol using eq 1 with optimization of the atomic energies. The problems with larger systems, as evidenced by the consistency of the results for the "training" and "test" sets, are no longer striking. Substantial improvements with the use of optimized atomic energies have been noted in other studies.^{34–38} (3) All methods yield MAEs of 2.1–2.5 kcal/mol for the 254 hydrocarbons. Wodrich et al. computed bond separation energies, which can be converted to heats of formation, for 72 hydrocarbons with 9 different functionals and the 6-311+G(d,p) basis set.^{14b} Their MAE of ca. 4 kcal/mol for B3LYP is greater than the present result owing to a systematic (mean signed error) of ca. -2 kcal/mol that is relieved by use of eq 1. They obtained the best results with the M05-2X functional, which yields a MAE of 2.13 kcal/mol and a negligible mean signed error. Thus, it appears that current DFT methods and PDDG/PM3 are reaching a

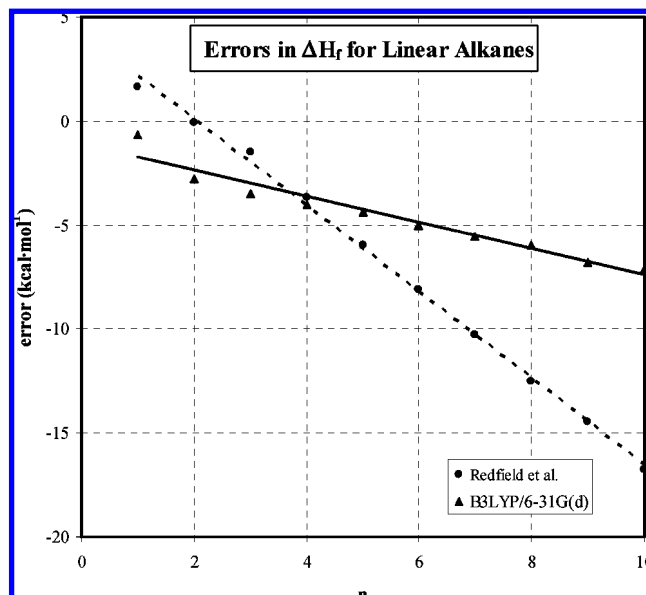


Figure 1. Errors in computed heats of formation for linear alkanes as a function of the number of carbon atoms, n . The present B3LYP/6-31G(d) results are compared with prior B3LYP-based results from ref 13.

MAE limit near 2.0 kcal/mol for hydrocarbons. (4) In comparing the results in columns 4 and 5, it is found that the use of SCC-DFTB or B3LYP/6-31G(d) optimized geometries has a negligible effect on the results. (5) The quality of the B3LYP-based results is not very sensitive to the choice of basis set. Without the dispersion correction, the best results here are obtained with B3LYP/6-31G(d,p). (6) However, none of these results is significantly better than that from the much less computationally demanding PDDG/PM3 method. This was unexpected and, actually, somewhat disappointing in that it does not point to an obvious path forward to a further improved semiempirical method, e.g., by increasing the basis set size or adding B3LYP-like electron repulsion terms. As noted previously, PDDG/PM3 does represent a significant improvement over AM1, PM3, and SCC-DFTB for the 622 heats of formation; these alternatives yield MAEs of 6.8, 4.4, and 5.8 kcal/mol, respectively.³⁴ (7) The lowest MAEs are obtained at the B3LYP/6-31+G(d,p) level with addition of the dispersion terms from eq 2. The overall MAE for the 622 heats of formation is 2.4 kcal/mol in this case. It appears to be difficult to go below this limit without resorting to more complex procedures such as G2 or G3, which better represent electron correlation, or to nongeneral methods using bond and group corrections.

Specific Problem Cases. An examination of the complete results indicates that some of the largest systematic errors in the B3LYP calculations occur in bridged polycyclic systems, particularly diamantane, diamantanone, adamantane, adamantanone, protoadamantane, bullvalene, cubane, camphor, and norbornadiene, which have errors in the 4–13 kcal/mol range at the different levels and contribute about 0.1 kcal/mol to the overall MAEs. Another problem case is molecules with multiple N–N bonds; nitrogen, tetramethyltetrazene, azidomethylbenzene, methyl azide, *trans*-diazene, azoisopropane, cyclopentyl azide, cyclohexyl azide, diazirine,

and azopropane yield errors ranging between 4 and 15 kcal/mol, and their contribution to the average MAE is up to 0.1 kcal/mol. Some of these problems have been reported in other studies, e.g., the heat of formation of norbornadiene calculated with B3LYP/6-311+G(3df,2p) is 66.5 kcal/mol, which is 10 kcal/mol greater than the experimental value or the result of a G2 calculation.⁴⁰ It was also previously noted that there is a 10 kcal/mol error in the isomerization energy for toluene going to norbornadiene using B3LYP/6-31G(d).³⁴ Indeed, Schreiner et al. have found numerous problems with relative energies from B3LYP calculations for polycyclic hydrocarbon isomers.⁴¹ Curiously, the computed heats of formation for norbornane with the present B3LYP-based methods are within 1–2 kcal/mol of experiment. As another example, although other calculated heats of formation for dimethyldiazene are not available, the energy change for the *trans* to *cis* isomerization was found to be 9.8 kcal/mol at the B3LYP/DZP level and 5.0 kcal/mol with G2.⁴² In addition, the error reported in the G2/97 data set for nitrogen (N_2) is 1.4 kcal/mol with B3LYP/6-311+G(3df,2p), so the present errors of ca. 5 kcal/mol may reflect basis-set insufficiency. The largest error for all levels of calculation is for isophthalamide (20–27 kcal/mol, accounting for ca. 0.04 of the total MAEs). It has been a consensus outlier; Stewart's analysis points to a possible experimental error of ca. 20 kcal/mol in this case.³⁵

Concerning hydrocarbon homology, there have been several attempts to address the scaling problem in B3LYP calculations. Already in 1996, Mole et al. fitted atom equivalents to the experimental values of heats of formation of 23 hydrocarbons and achieved rmsd values of 1.7 and 1.0 kcal/mol with the 6-31G(d) and 6-311G(d,p) basis sets.⁴³ The reasonable trend of increasing accuracy with the bigger basis set was obtained. However, Saeys et al. carried out a similar study with a larger set of 58 hydrocarbons with 10 or fewer carbon atoms.⁴⁴ After fitting the atomic energies, they were able to reduce the MAE for B3LYP/6-31G(d) from 7.4 to 2.2 kcal/mol, while the error with B3LYP/6-311G(d,p) decreased from 12.0 to only 3.1 kcal/mol. CBS-QB3 performed particularly well with the MAE declining from 2.2 to 0.6 kcal/mol. An apparently different approach was followed by Duan et al.⁴⁵ They based their corrections on the number of electrons for each atom at different levels and in different chemical environments (lone pairs, bonds). After fitting to experimental values this is equivalent to atom-based corrections. Recently, the smallest MAE for the full G3 set (222 molecules) reported so far has been given by Friesner et al.⁴⁶ Their method includes utilization of 22 optimized parameters that depend on atomic hybridization and bond types. This allows reduction of the MAE for B3LYP/6-311+G(3df,2p)-based calculations from 4.8 to 0.8 kcal/mol. However, methods that use such bonding corrections are not generally applicable, e.g., to nonequilibrium geometries or to bonding arrangements that have not been tested. Even with semiempirical MO methods, one can obtain average errors in heats of formation of 2 kcal/mol with the use of atom and bond corrections.⁴⁷

In the present case, the homology problem was not specifically addressed since the optimization of the atomic

Table 3. Isomerization Enthalpies for Hydrocarbons (kcal/mol)

	Exp. ^a	HF	B3LYP					PDDG/ PM3 ^b
		6-31G (d) ^b	6-31G (d) ^b	6-31G (d) ^c	6-31G (d,p) ^b	6-31+ G(d,p) _b	6-31+ G(d,p) _b +OC6	
	1.24	2.18	-2.70	-2.76	-2.83	-2.73	-2.74	6.12
	21.81	27.44	22.22	21.48	21.98	23.16	23.17	21.27
	7.85	8.47	8.29	7.75	8.33	9.04	9.20	7.89
	1.13	1.81	1.40	1.35	1.37	1.34	0.77	1.94
	1.27	1.12	0.75	0.23	0.71	0.86	1.04	1.18
	2.79	2.62	3.68	3.67	3.62	3.17	3.17	4.62
	11.34	12.94	11.95	11.85	12.26	13.90	14.18	8.99
	22.17	22.59	21.63	22.19	21.22	19.92	19.54	25.16
	7.13	8.47	10.38	9.78	9.54	9.52	9.54	5.36
	5.06	1.42	1.43	0.75	1.53	1.70	2.37	5.76
	4.09	-9.81	-6.83	-7.54	-6.64	-6.79	-0.36	4.03
	46.65	56.37	57.30	56.06	57.11	57.11	57.58	45.59
	35.83	39.23	41.06	40.61	41.00	40.04	39.62	30.52
MAE		3.2	3.2	3.1	3.1	3.4	2.9	1.7

^a From experimental heats of formation; see ref 4 and the Supporting Information. ^b Single-point energy evaluations on SCC-DFTB optimized geometries except B3LYP/6-31G(d) geometries were used for propene, 2-methylpropene, and penta-1,3-diene. ^c Using geometries optimized at the B3LYP/6-31G(d) level.

energies considered all 622 data points. Nevertheless, the MAE for the G3 linear alkane series (methane to decane) declines to 4.6 kcal/mol for the present B3LYP/6-31G(d) procedure from the 7.5 kcal/mol in the original B3LYP/6-311+G(3df,2p) study by Redfern et al.¹³ The results still reflect a cumulative error, although it is greatly decreased. As shown in Figure 1, the slope for the line fitted to the error vs the number of carbon atoms decreases from $-2.1 \text{ kcal} \cdot \text{mol}^{-1}$ per carbon atom with Redfern's data to -0.6 in the present calculations.

For a more general comparison of the effects of system size, it may be noted that the G2/97 set contains 148 heats of formation including 59 neutral, closed-shell, C, H, N, and O-containing molecules, which happen to have 1–6 heavy atoms.⁸ The MAE obtained for this subset using the published data is 2.2 kcal/mol,⁸ while the corresponding MAE from the present B3LYP/6-31G(d) calculations is 3.0 kcal/mol. However, for G3/99 heats of formation were added for an additional 38 neutral C, H, N, and O-containing molecules (5–10 heavy atoms).⁹ The original B3LYP/6-31G(d) calculations yielded an MAE of 7.8 kcal/mol for this set, while the present approach with eq 1 reduces the MAE for these larger molecules to 2.9 kcal/mol.

Isomerization Energies. Results for the same three sets of isomerization reactions, which were used in ref 34, are given in Tables 3–5. The changes in structure and bonding

for the 34 isomerizations are notably diverse and potentially challenging for the quantum methods. The experimental data and PDDG/PM3 results correspond to enthalpy changes at 25 °C, while the B3LYP results are the change in electronic energy. In the previous study,³⁴ it was found that including the zero-point corrections to the B3LYP/6-31G(d) results using B3LYP/6-31G(d) vibrational frequencies left the MAE for the 13 isomerizations in Table 3 unchanged at 3.1 kcal/mol, and inclusion of the thermal corrections to the vibrational energies only reduced the MAE to 2.9 kcal/mol. Consequently, these adjustments were neglected. Grimme et al. have also recently examined the 34 isomerizations with a variety of DFT procedures and their SCS-MP2 method.⁴⁸ They also performed reference calculations at the CCSD-(T)-cc-PV(DT)Z//B3LYP/TZV(d,p) level, which called into question four of the experimental enthalpy changes in ref 34. The corresponding revised differences in enthalpies of isomerization for the four cases have been adopted here, as noted in Tables 4 and 5.

For the hydrocarbons in Table 3, there are no striking problems with small ring compounds. However, the greatest errors with the B3LYP methods are for the largest molecules, the last four entries. The conversions of neopentane to pentane and tetramethylbutane to octane, with errors of 3.5 and 10.5 kcal/mol, reflect the now well-known branching error. As in most of the entries in this table, the increase in

Table 4. Isomerization Enthalpies (kcal/mol) for Nitrogen-Containing Molecules

	Exp. ^a	HF	B3LYP					PDDG/ PM3 ^b
		6-31G (d) ^b	6-31G (d) ^b	6-31G (d) ^c	6-31G (d,p) ^b	6-31+G (d,p) ^b	6-31+G (d,p) -OC6 ^b	
<chem>CH3CN -> CH3NC</chem>	24.2 ^d	21.16	24.94	24.07	25.17	23.12	23.12	25.78
<chem>CCN -> CN</chem>	6.92	5.33	4.26	4.51	5.65	6.82	6.69	7.11
<chem>C#N -> N#C</chem>	11.71	14.63	14.91	14.40	13.47	13.34	13.61	17.83
<chem>NCCN -> CNCN</chem>	26.06	22.53	19.29	20.31	21.98	24.54	25.23	30.58
<chem>C1CCNC1 -> C1CCN1</chem>	10.65	15.43	14.43	13.89	12.94	11.78	11.42	5.03
<chem>c1ccncc1 -> c1ccnc1</chem>	4.6 ^d	6.02	3.65	4.07	3.65	3.69	3.70	2.95
<chem>c1ccncc1 -> c1ccncc1</chem>	19.66	20.23	18.41	18.60	18.26	18.33	18.33	16.06
<chem>c1ccncc1 -> c1ccncc1</chem>	1.10	1.44	1.14	0.89	1.22	1.49	1.40	-0.27
<chem>Nc1ccccc1 -> Nc1ccccc1</chem>	2.89	-4.31	-2.95	-2.53	-0.18	1.66	1.79	1.12
<chem>C1CCNCC1 -> C1CCNCC1</chem>	4.88	2.01	1.00	1.26	2.30	3.21	3.26	5.64
MAE		2.8	2.9	2.5	1.9	1.1	1.0	2.7

^a From experimental heats of formation; see ref 4 and the Supporting Information. ^b Single-point energy evaluations on SCC-DFTB optimized geometries. ^c Using geometries optimized at the B3LYP/6-31G(d) level. ^d Revised value based on ref 48.

basis set seems to have mixed effects. It does improve slightly the agreement for the neopentane to pentane isomerization, but it has a negligible effect on the tetramethylbutane to octane case. However, strikingly, the dispersion correction reduces these errors by 0.7 and 6.4 kcal/mol, respectively. The 10 kcal/mol error for the toluene-norbornadiene conversion is dominated by the error in the heat of formation in the latter, as described above. The problem is not relieved with the larger basis sets. Owing to the compactness of the bicyclic molecule, the present dispersion correction also provides no help in this case, though overall it lowers the MAE by 0.5 kcal/mol at the B3LYP/6-31+G(d,p) level.

The isomerizations of nitrogen containing molecules reported in Table 4 show a general pattern of having large errors at the B3LYP/6-31G(d) level, e.g., 7 kcal/mol for ethylenediamine to 1,2-dimethylhydrazine and 6 for aniline to 2-methylpyridine. In fact, the problem is general to cases where the number of N–H bonds change; N–H bonds appear to not be energetically favorable enough. As noted previously, zero-point corrections are not the issue here; their addition actually worsens the accord by about 1 kcal/mol.³⁴ However, an increase in the size of the basis set in these cases has very favorable effects. The improvement is sufficient to lower the B3LYP MAE of the set from 2.5 at the 6-31G(d) level to 1.9 at 6-31G(d,p) and 1.1 kcal/mol at 6-31+G(d,p). It is noted that the experimental value for pyrimidine to pyrazine that was quoted previously, 0.03 kcal/

mol,³⁴ has been separately brought into question by Grimme et al.,⁴⁸ who obtained a value of 4.6 kcal/mol from CCSD-(T)-cc-PV(DT)Z//B3LYP/TZV(d,p), and by Cheng et al., who obtained 4.5 kcal/mol using G3 calculations.⁴⁹ The present B3LYP and PDDG/PM3 results support their higher estimates. The results of Grimme et al.⁴⁸ also indicate that the correct value for the acetonitrile to methyl isocyanide conversion should be 24.2 kcal/mol rather than the former 21.4 kcal/mol,³⁴ which is also supported by the B3LYP/6-31+G(d,p) and PDDG/PM3 estimates.

In Table 5, an analogous pattern is found as in Table 4 with the largest errors occurring for isomerizations that change the number of O–H bonds. The 12-kcal/mol error for the 1,2-ethanediol to dimethyl peroxide transformation is striking at the B3LYP/6-31G(d) level. Again, the problem is largely removed with an increase in the basis set size. There is also a 6-kcal/mol error for the acetic acid to methyl formate conversion with B3LYP/6-31G(d). However, the acid to ester conversion for hexanoic acid to methyl pivalate seems remarkably on target until one realizes that there is a cancellation with the alkyl branching error (ca. 4 kcal/mol for neopentane to pentane). So, as noted before,³⁴ a particularly bad case for B3LYP/6-31G(d) would be to consider the pivalic acid to ethyl propanoate isomerization, (CH₃)₃-CCOOH → CH₃CH₂COOCH₂CH₃. The experimental enthalpy change is +6.3 kcal/mol, while the B3LYP/6-31G(d) energy difference is −5.2 kcal/mol. For the entries in Table

Table 5. Isomerization Enthalpies (kcal/mol) for Oxygen-Containing Molecules

	Exp. ^a	HF/	B3LYP/					PDDG/ PM3 ^b
		6-31G (d) ^c	6-31G (d) ^b	6-31G (d) ^c	6-31G (d,p) ^b	6-31+G (d,p) ^b	6-31+G (d,p) +OC6 ^b	
<chem>C2H5OH -> (CH3)2O</chem>	12.25	7.90	5.81	5.50	8.55	10.59	10.27	10.01
<chem>CC=O -> C1CO1</chem>	27.15	32.49	27.52	27.52	27.49	27.64	27.64	32.27
<chem>CC(=O)O -> CC(=O)OC</chem>	18.30	15.98	13.05	11.79	15.67	17.21	16.56	16.12
<chem>OCCO -> COCO</chem>	62.57	64.04	50.50	50.12	55.62	59.86	60.63	61.36
<chem>CC(=O)C -> C1CCOC1</chem>	32.65	38.29	33.72	31.80	33.89	34.34	34.51	30.46
<chem>C1CCOC1 -> CC=CCO</chem>	10.39	17.44	13.21	12.26	12.61	11.32	10.93	15.60
<chem>CC(=O)N -> CC(=O)NC</chem>	9.4 ^d	10.09	7.30	7.10	8.72	9.80	9.51	10.57
<chem>C1CCCOC1 -> CC(=O)CC=O</chem>	12.33	14.16	12.59	12.61	11.99	11.12	10.07	15.90
<chem>C1CCCOC1 -> CC(=O)CC(=O)C</chem>	5.0 ^d	0.91	3.64	5.53	3.43	1.74	0.59	3.19
<chem>CCCCC(=O)O -> CC(C)(C)C(=O)OC</chem>	4.39	7.35	4.87	4.83	7.41	9.59	7.42	8.82
<chem>Cc1ccccc1O -> Oc1ccccc1</chem>	8.13	6.14	8.43	8.11	8.43	7.55	7.61	8.64
MAE		3.3	2.6	2.6	1.9	1.8	1.7	2.7

^a From experimental heats of formation; see ref 4 and the Supporting Information. ^b Single-point energy evaluations on SCC-DFTB optimized geometries except B3LYP/6-31G(d) geometries were used for THF and acetylacetone. ^c Using geometries optimized at the B3LYP/6-31G(d) level. ^d Revised value based on ref 48.

5, the results of Grimme et al.⁴⁸ provide revision of the value for the *N*-methacetamide to dimethylformamide conversion to 9.4 kcal/mol from 13.5 kcal/mol,³⁴ again closer to the B3LYP/6-31+G(d,p) and PDDG/PM3 results, and for the valerolactone to acetylacetone isomerization from 0.3 to 5.0 kcal/mol, in better agreement with the PDDG/PM3 result.

Overall, the B3LYP/6-31+G(d,p) results in Tables 4 and 5 are impressive, and for these relatively small molecules the dispersion correction has little impact. The greater impact for the hydrocarbons in Table 3 is dominated by the improvement for the tetramethylbutane isomerization. For the 34 isomerizations the overall MAEs in kcal/mol are 3.2 (HF/6-31G(d)), 2.7 (B3LYP/6-31G(d)), 2.4 (B3LYP/6-31G(d,p)), 2.2 (B3LYP/6-31+G(d,p)), 1.9 with the OC6 term, and 2.3 (PDDG/PM3). If the comparisons are made to the 34 energy changes from Grimme et al.⁴⁸ instead of the present enthalpy changes, the overall MAEs for the 34 isomerizations are 3.0, 2.9, 2.6, 2.4, 2.2, and 2.4 kcal/mol, respectively.

Grimme et al. also studied the 34 isomerization reactions with their modified SCS-MP2 procedure and with several functionals and many basis sets.⁴⁸ For B3LYP, starting from 6-31G(d) the rmsd improves from 4.33 to 3.32 kcal/mol by using the Dunning type *aug-cc-pVTZ* basis set, though it requires 211 times more CPU time. The corresponding SCS-MP2 calculations show rmsd values of 2.67 and 1.25 kcal/mol, respectively. However, the isomerization of tetrame-

thylbutane to octane still favors the linear isomer by 8.0 kcal/mol. In a very recent paper, Schwabe and Grimme described the implementation of their empirical dispersion term with the B3LYP, B2PLYP, and mPWPLYP functionals. When tested on the full G3/99 set of 223 heats of formation they obtained reductions in the MAEs from 5.6 to 3.1, 2.4 to 1.7, and 2.1 to 1.7 kcal/mol, respectively.⁵⁰

Conclusions

The present evaluation of the performance of B3LYP with different basis sets on 622 neutral, closed-shell organic compounds containing C, H, N, and O has provided a clearer assessment of its capabilities and limitations for organic molecules. When used with the popular 6-31G(d) basis set, the average errors are about 3.0 kcal/mol for both isomerization energies and heats of formation, when the latter are computed via eq 1. Optimization of the atomic energies in eq 1 helps alleviate the systematic homologation errors with B3LYP (Figure 1). Nevertheless, improvement with increasing basis-set size is modest for the heats of formation, though addition of the dispersion correction with 6-31+G(d,p) lowers the MAE to 2.4 kcal/mol (Table 2). There is greater improvement for the 34 isomerization reactions with a MAE of 2.2 kcal/mol for B3LYP/6-31+G(d,p) and a MAE of 1.9 kcal/mol with addition of the dispersion terms (Tables 3–5). The greatest benefit of the dispersion terms is in reducing

the underestimate of the stability of branched over linear alkyl groups; however, significant problems remain for bridged polycyclic molecules. It is clear that the 2.0-kcal/mol barrier in MAEs cannot be overcome without the use of a functional other than B3LYP and/or resorting to compound methods and empirical corrections.

A second key motivation for the present study was to ascertain the possible improvements in accuracy that could be obtained from the DFT calculations over the semiempirical PDDG/PM3 method when they are put on an even footing by use of eq 1 for computation of the 622 heats of formation. As it turned out, there is no significant difference in quality between the B3LYP-based results and those from PDDG/PM3 for the heats of formation and isomerization energies. Though this confirms that PDDG/PM3 is an efficient alternative for evaluation of the energetics of organic reactions, obvious guidance has not emerged on how to create a further-improved semiempirical method. There are no striking, general errors with PDDG/PM3, though gains could be sought in performance for small heterocycles (Tables 4 and 5) as well as for conformational energetics and hydrogen bonding.^{4–6,34} The recent scrutiny of the performance of quantum chemistry methods has demonstrated the importance of testing on large, diverse data sets. This does appear to be leading to the generation of improved DFT methodology such as the M05-2X functional.²⁰ In view of the need for rapid methods that can be applied to very large systems, as in the examination of enzymatic reactions or design of new materials, greater development efforts in the semiempirical arena also appear warranted.

Acknowledgment. The authors thank Dr. M. Elstner for the SCC-DFTB code used in this work and the National Science Foundation (CHE-0446920) for financial support.

Supporting Information Available: Computed and experimental data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140* (4A), 1133–1138.
- (2) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136* (3B), 864–871.
- (3) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (4) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601–1622.
- (5) Tubert-Brohman, I.; Guimaraes, C. R. W.; Repasky, M. P.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 138–150.
- (6) Tubert-Brohman, I.; Guimaraes, C. R. W.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2005**, *1*, 817–823.
- (7) Riley, K. E.; Op't Holt, B. T.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2007**, *3*, 407–433.
- (8) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063–1079.
- (9) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374–7383.
- (10) Xu, X.; Zhang, Q.; Muller, R. P.; Goddard, W. A., III. *J. Chem. Phys.* **2005**, *122*, 014105.
- (11) Kafafi, S. A. *J. Phys. Chem.* **1998**, *102*, 10404–10413.
- (12) Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067–3077.
- (13) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850–5854.
- (14) (a) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 3632–3634. (b) Wodrich, M. D.; Corminboeuf, C.; Schreiner, P. R.; Fokin, A. A.; Schleyer, P. v. R. *Org. Lett.* **2007**, *9*, 1851–1854.
- (15) Grimme, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460–4464.
- (16) Grimme, S.; Steinmetz, M.; Korth, M. *J. Chem. Theory Comput.* **2007**, *3*, 42–45.
- (17) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782–3797.
- (18) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949–3957.
- (19) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
- (20) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
- (21) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109–1 to 234109–9.
- (22) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
- (23) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (24) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (25) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287–5393.
- (26) Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. *Org. Biomol. Chem.* **2007**, *5*, 741–758.
- (27) McNamara, J. P.; Hillier, I. H. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2362–2370.
- (28) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.
- (29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.03*; Gaussian, Inc.: Pittsburgh, PA, 2003.
- (30) Jorgensen, W. L.; Tirado-Rives, J. *J. Comput. Chem.* **2005**, *26*, 1689–1700.

- (31) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–105.
- (32) Code and parameter files to modify MOPAC for PDDG/PM3 calculations are available at <http://www.jorgensenresearch.com>.
- (33) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (34) Sattelmeyer, J. W.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.
- (35) Stewart, J. J. P. *J. Mol. Model.* **2004**, *10*, 6–12.
- (36) Csonka, G. I.; Ruzinsky, A.; Tao, J.; Perdew, J. P. *Int. J. Quantum Chem.* **2005**, *101*, 506–511.
- (37) Brothers, E. N.; Scuseria, G. E. *J. Chem. Theory Comput.* **2006**, *2*, 1045–1049.
- (38) Voityuk, A. A. *J. Chem. Theory Comput.* **2006**, *2*, 1038–1044.
- (39) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. The generic parameters are (σ in Å, ϵ in kcal/mol) are as follows: C (3.55, 0.07), H (2.46, 0.03), N (3.25, 0.17), and O (3.00, 0.17).
- (40) Rogers, D. W.; McLafferty, F. J. *J. Phys. Chem. A* **1999**, *103*, 8733–8737.
- (41) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A., Jr.; de Meijere, A. *Org. Lett.* **2006**, *8*, 3635–3638.
- (42) Vrábel, I.; Biscupič, S.; Staško, A. *J. Phys. Chem. A* **1997**, *101*, 5805–5812.
- (43) Mole, S. J.; Zhou, X.; Liu, R. *J. Phys. Chem.* **1996**, *100*, 14665–14671.
- (44) Saeys, M.; Reyners, M.-F.; Marin, G.; Van Speybroeck, V.; Waroquier, M. *J. Phys. Chem. A* **2003**, *107*, 9147–9159.
- (45) Duan, X.-M.; Song, G.-L.; Li, Z.-H.; Wang, X.-J.; Chen, G.-H.; Fan, K.-N. *J. Chem. Phys.* **2004**, *121*, 7086–7095.
- (46) Friesner, R. A.; Knoll, E. H.; Cao, Y. *J. Chem. Phys.* **2006**, *125*, 124107-1–124107-24.
- (47) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 498–510.
- (48) Grimme, S.; Steinmetz, M.; Korth, M. *J. Org. Chem.* **2007**, *72*, 2118–2126.
- (49) Cheng, M.-F.; Ho, H.-O.; Lam, C.-S.; Lee, W.-K. *J. Serb. Chem. Soc.* **2002**, *67*, 257–264.
- (50) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.

CT700248K