Angewandte Chemie

Density Functional Theory

DOI: 10.1002/anie.200700386

Relative Energy Computations with Approximate Density Functional Theory—A Caveat!**

*Peter R. Schreiner**

Keywords:
computational chemistry · correlation energy · density functional calculations · hydrocarbons

Approximate density functional theory (DFT) is now a common computational chemistry tool to examine a broad variety of structures and reactions involving increasingly larger molecules.[1] DFT performs exceptionally well for molecular structures with much reduced computational effort than traditional ab initio methods, which it is often on par with, especially for difficult electronic situations—although this is not to say that right answers are given for the right reasons. While it has long been recognized and appreciated that there is no obvious way to improve DFT systematically, and although it is considered an approximate (short of saying empirical) method within its current implementations, it is silently assumed that the lack of quantitative accuracy is by far outweighed by the high qualitative agreement of structures and energies computed with this method. This may be true for small molecules such as those contained in the so-called G2 (Gaussian 2) test set[2] of molecules for which good experimental data are available for validation, but there is growing and convincing evidence that this is not the case for increasingly larger (organic) molecules. The most popular B3LYP (Becke-3-Lee-Yang-Parr) functional combination was, amongst others, the first to show its shortcomings in studies on, for example, 1) the enthalpies of formation for chain hydrocarbons,[3] 2) the energy difference of propyne versus allene (propyne is more stable by 1.4 kcal mol^{-1} under standard conditions, but all established DFT methods give the opposite energy ordering with considerable error bars;[4] newer ones perform better[5,6]), 3) some electrocyclic reactions,[7] 4) hydrocarbon reaction energies,[8] and 5) C–C bond energies (root-mean-square (rms) deviations from experiment for reaction (1) amount to 15.0 kcal mol^{-1} with maximum errors of up to 21.1 kcal mol^{-1}).[9]

$$\mathrm{Alk-Alk'} \rightarrow \mathrm{Alk}^{\cdot} + {}^{\cdot}\mathrm{Alk'} \qquad (1)$$

As with any systematic error, it may guide the way to improve the underlying method. Hence, it has long been recognized that the lack of van der Waals interactions in most current DFT implementations will be problematic when dealing with weakly bound molecular complexes and structures where dispersion forces are critical (e.g., alkanes). Unfortunate choices in the terms describing the kinetic energy density (for meta-GGAs[10]), the amount of Hartree–Fock exchange (for hybrid functionals), and the self-interaction error[11] (the spurious interaction of an electron with itself) in the correlation functionals are likely to be contributors to some of the observed DFT deficiencies. The use of the G2 reference data set with only small molecules may also be insufficient for accurate parameterizations of DFT implementations. This is evident from the fact that the errors with B3LYP for the G3 test set,[12] which additionally contains a number of larger molecules, are double those for the G2 set.[13]

There are three recent studies[14–16] pointing out that the relative energies of simple hydrocarbon isomers are not reproduced well with a variety of common DFT methods; B3LYP turns out to be particularly poor. These conclusions are in line with earlier[4,6,8] as well as with follow-up studies.[17] The errors are cumulative and can become so large that the results for systems involving just eight carbon atoms may be meaningless. What is particularly important about these recent studies is the realization that the inclusion of dispersion interactions will *not* be sufficient to alleviate the deficiencies in the energy computations.

It is usually assumed that isodesmic equations will largely cancel systematic errors. As this is not the case for the stabilization energies of higher linear alkanes (computed with isodesmic equation (2)), systematic errors can be identified.[14]

$$\mathrm{CH_3(CH_2)}_m\mathrm{CH_3} + m\,\mathrm{CH_4} \rightarrow (m+1)\,\mathrm{C_2H_6} \qquad (2)$$

Equation (2) evaluates the stabilization of linear alkanes through stabilizing 1,3-methyl (or methylene) interactions. This "protobranching" also offers a sound explanation for the thermodynamic preference of branched over linear alkanes as the number of protobranches in the branched alkanes is larger. Figure 1 reveals that Hartree–Fock (HF) theory and common DFT approaches perform poorly and that the errors are cumulative.

---

[*] Prof. Dr. P. R. Schreiner
Institute of Organic Chemistry
Justus-Liebig University
Heinrich-Buff-Ring 58
35392 Giessen (Germany)
Fax: (+49) 641-9934-309
E-mail: prs@org.chemie.uni-giessen.de
Homepage: http://www.chemie.uni-giessen.de/home/international/schreiner

[**] This work was supported by the Justus-Liebig University Giessen. I thank Stefan Grimme, Donald Truhlar, and Paul von R. Schleyer for helpful discussions and Matthew D. Wodrich for the data for Figure 1.

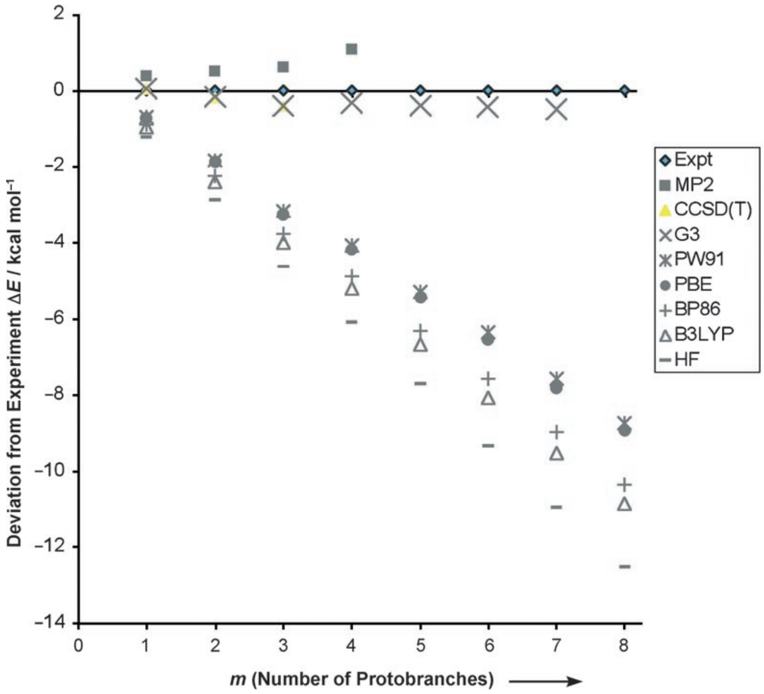
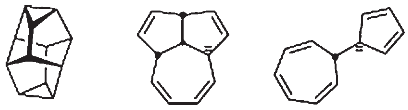


Figure 1. Deviations of computational approaches from experimental (0 K) protobranching stabilization energies (see Equation (2), including zero-point vibrational energies).[14] Levels of theory: CCSD(T)/aug-cc-pVTZ//MP2/6–311+G(d,p), MP2/aug-cc-pVTZ//MP2/6–311+G(d,p); all other computations were performed with a 6-311+G(d,p) basis set.

The DFT errors are even larger when relative isomer energies of a variety of hydrocarbons with largely different bonding situations are compared.[15] At first glance, it is rather surprising that a highly strained hydrocarbon such as octahedrane (**1**) is the most stable $(CH)_{12}$ structure.[18] Indeed, common DFT approaches such as BLYP and B3LYP favor **2** or **3** (Scheme 1), while high-level CCSD(T)/cc-pVDZ//MP2/aug-cc-pVDZ computations strongly favor **1** (by 14.3 and 25.0 kcal mol^{-1}, respectively).

Systematic studies on larger hydrocarbons including structures with strained rings and unsaturation revealed also that the errors increase with the size of the system and that computations on structures with only single bonds are more error-prone than others.[15]



1 **2** **3**

Scheme 1. Octahedrane (**1**), the thermodynamically most stable $(CH)_{12}$ hydrocarbon, and the next energetically higher isomers **2** and **3**.

But where do these energetic errors come from? In all likelihood not from basis-set incompleteness, and although basis-set deficiencies can play an important role in ab initio quantum theory, they are less pronounced for DFT.[19] Note, for instance, that there are several basis sets for which the MP2 and other explicitly correlated levels of theory give a negative eigenvalue for D_{6h}-benzene, while B3LYP and other functionals correctly give all real frequencies.[20]

The neglect of long-range correlation (mostly van der Waals dispersion) also cannot be solely responsible for the observed accumulation of errors.[16] This is evident from the insufficient improvement of the isomer energy differences when dispersion interactions are explicitly taken into account (DFT-D).[21] Figure 2 displays the isomerization energy difference for *n*-octane (**4**) to 2,2,3,3-tetramethylbutane (**5**) at selected levels of theory relative to experiment.

It is obvious that B3LYP performs only slightly better than HF theory, although it is usually assumed that, apart from dispersion, all other important electron correlation effects are taken into account in DFT. Hence, it is likely that medium-range correlation is a decisive factor that determines the stability of alkanes, and that this part of the electron correlation is not appropriately included in current DFT implementations.[16] A careful computational analysis of the pair correlation energies (at the spin-component-scaled MP2 level of theory)[22] shows that indeed non-local medium-range interactions (1.5–3.5 Å)

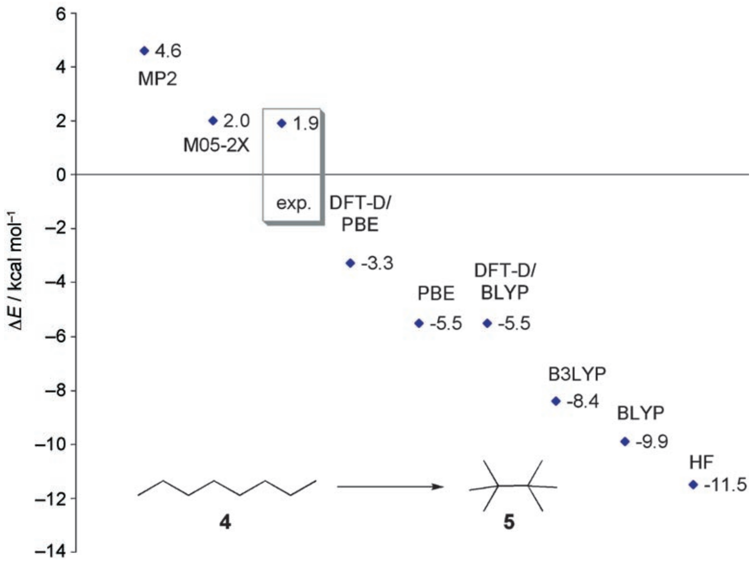


Figure 2. Energies of isomerization from **4** to **5** at various levels of theory versus experiment. Structures were optimized at the MP2/TZV(d,p) level; energies were computed with a cQZV3P basis set.[16,21]

are key to rationalizing the alkane isomer energy differences. These findings are supported by the excellent performance of new functionals (e.g., M05-2X) that account better for medium-range correlation energy by using new functional forms for the dependence of the functional on the density gradient and kinetic energy density in both the exchange and correlation functionals.[23]

The solid quantum mechanical explanation of the importance of stabilizing geminal interactions nicely falls into place with the aforementioned importance of 1,3 interactions ("protobranching") and also makes the success of simple additivity rules to determine relative hydrocarbon energies plausible. The DFT medium-correlation energy problems described herein are not limited to hydrocarbons and also arise for other atomic compositions and structures.[24]

The statement by Paul von R. Schleyer at a recent ACS meeting that *"…the happy days of 'black box' DFT usage are over, at least for energy evaluations. Confidence has been undermined. No available density functional is generally reliable for larger molecules. Efforts to correct for DFT dispersion errors will not be sufficient…"* could not better summarize some of the disappointing as well as alarming results of most current DFT implementations to seemingly simple organic structures. On the positive side, this is likely to pave the way for *systematic* improvements of this highly efficient and useful way to compute the structures and energies of molecules. In the meantime, a sound recommendation therefore is to utilize (ideally various) DFT approaches for structure optimizations but to use higher levels for energy comparisons.

---

[1] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed., Wiley-VCH, Weinheim, **2001**.

[2] L. A. Curtiss, K. Raghavachari, G. W. Trucks, J. A. Pople, *J. Chem. Phys.* **1991**, *94*, 7221–7230.

[3] P. C. Redfern, P. Zapol, L. A. Curtiss, K. Raghavachari, *J. Phys. Chem. A* **2000**, *104*, 5850–5854; L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, *J. Chem. Phys.* **2000**, *112*, 7374–7383.

[4] H. L. Woodcock, H. F. Schaefer III, P. R. Schreiner, *J. Phys. Chem. A* **2002**, *106*, 11 923–11 931.

[5] J. K. Kang, C. B. Musgrave, *J. Chem. Phys.* **2001**, *115*, 11 040–11 051.

[6] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2006**, *110*, 10 478–10 486.

[7] A. G. Leach, E. Goldstein, K. N. Houk, *J. Am. Chem. Soc.* **2003**, *125*, 8330–8339; S. M. Bachrach, J. C. Gilbert, *J. Org. Chem.* **2004**, *69*, 6357–6364.

[8] C. E. Check, T. M. Gilbert, *J. Org. Chem.* **2005**, *70*, 9828–9834.

[9] X.-Q. Yao, X.-J. Hou, H. Jiao, H.-W. Xiang, Y.-W. Li, *J. Phys. Chem. A* **2003**, *107*, 9991–9996; Y. Feng, L. Liu, T.-J. Wang, H. Huang, Q.-X. Guo, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2005–2013.

[10] J. P. Perdew, J. M. Tao, V. N. Staroverov, G. E. Scuseria, *J. Chem. Phys.* **2004**, *120*, 6898–6911.

[11] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, *J. Chem. Phys.* **2006**, *125*, 19 411.

[12] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, J. A. Pople, *J. Chem. Phys.* **1998**, *109*, 7764–7776.

[13] L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, *J. Chem. Phys.* **2000**, *112*, 7374–7383.

[14] M. D. Wodrich, C. Corminboeuf, P. von R. Schleyer, *Org. Lett.* **2006**, *8*, 3631–3634.

[15] P. R. Schreiner, A. A. Fokin, R. A. Pascal, A. de Meijere, *Org. Lett.* **2006**, *8*, 3635–3638.

[16] S. Grimme, *Angew. Chem.* **2006**, *118*, 4571–4575; *Angew. Chem. Int. Ed.* **2006**, *45*, 4460–4464.

[17] Y. Zhao, D. G. Truhlar, *Org. Lett.* **2006**, *8*, 5753–5755.

[18] A. de Meijere, C. H. Lee, M. A. Kuznetsov, D. V. Gusev, S. I. Kozhushkov, A. A. Fokin, P. R. Schreiner, *Chem. Eur. J.* **2005**, *11*, 6175–6184.

[19] F. Weigend, F. Furche, R. Ahlrichs, *J. Chem. Phys.* **2003**, *119*, 12 753–12 762.

[20] D. Moran, A. C. Simmonett, F. E. Leach III, W. D. Allen, P. von R. Schleyer, H. F. Schaefer III, *J. Am. Chem. Soc.* **2006**, *128*, 9342–9343.

[21] S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463–1473.

[22] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[23] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput.* **2006**, *2*, 364–382.

[24] S. Grimme, M. Steinmetz, M. Korth, *J. Chem. Theory Comput.* **2007**, *3*, 42–45.