Density Functional Calculations

DOI: 10.1002/anie.200801843

Sources of Error in DFT Computations of C–C Bond Formation Thermochemistries: π→σ Transformations and Error Cancellation by DFT Methods**

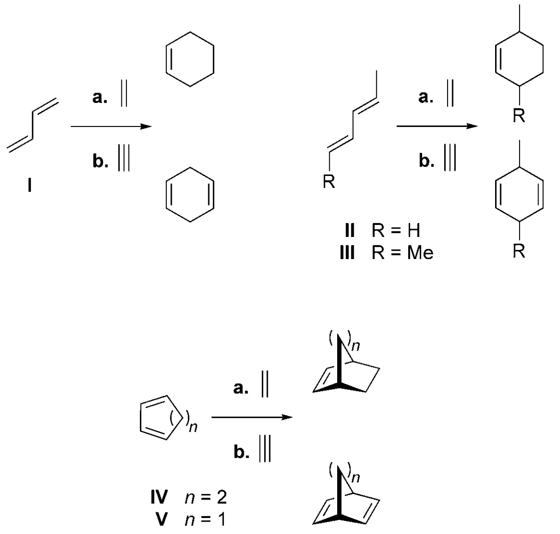
*Susan N. Pieniazek, Fernando R. Clemente, and Kendall N. Houk**

Quantum mechanical methods based on Kohn–Sham Density Functional Theory (DFT)[1] are currently the most used methods in computational chemistry. In particular, the hybrid density functional B3LYP[2] enjoys a vast popularity in the chemistry community, but there is increasing evidence about the limitations of this functional. The good performance of the ubiquitous B3LYP/6-31G(d) model chemistry has been suggested to result from error cancellation.[3,4] In a recent Highlight in *Angewandte Chemie*,[5] Schreiner reviewed current research about limitations of DFT methods. Examples of problematic areas include calculations of enthalpies of formation of long-chain and branched hydrocarbons,[6] hydrocarbon reaction energies,[7] and calculation of electrocyclic reactions.[8] In our study of Diels–Alder reactions of halofurans,[9] errors in B3LYP caused us to resort to high accuracy CBS-QB3 calculations.

Despite the alarming reports[4,10] about large errors in DFT energetics, little effort has been made to identify sources of error. Studies have shown that B3LYP is insufficient in the prediction of isomer relative energies of hydrocarbon and other main group element containing molecules.[10–12] The work of Schleyer et al.[13] identified systematic errors of DFT methods in the evaluation of the stabilizing 1,3-alkyl-alkyl interactions (protobranching) that exist in linear, branched, and most cycloalkanes, but not in methane and ethane. Grimme has provided a theoretical explanation based on the inability of DFT to compute medium-range correlation.[11] In related examples, increasing DFT errors have been found with increasing hydrocarbon system size for the isomers with the largest number of single bonds and small rings.[14]

As a result of the reported problems in energy evaluations, and the plethora of, apparently similar, existing DFT methods, chemists are facing uncertainty in their application to research problems, often relying on a fortunate error cancellation. Method development, on the other hand, should obviously benefit from the identification of problematic areas in fundamental transformations of interest in chemistry. To bridge the chemistry community and the field of functional development, the aim of this work is to provide insight into the combined sources of errors present in C–C bond forming reactions. In particular, Diels–Alder reactions are decomposed into contributions from the basic structural features involved in these C–C bond-forming reactions.

The energetics of the ten Diels–Alder reactions **I**–**V** in Scheme 1 have been computed[15] with the high-accuracy CBS-QB3[16] and G3[17] methods. The results are given in Table 1. These reactions involve acyclic and cyclic dienes, and ethylene and acetylene dienophiles. All reactions involve the conversion of π to σ bonds from either alkene or alkyne reactants, as well as changes in branching, hyperconjugation interactions, and strain.



Scheme 1. Cycloadditions involving ethylene and acetylene as dienophiles.

The errors in these hydrocarbon Diels–Alder reaction energetics using a number of popular hybrid DFT functionals[18] and the wave-function-based SCS-MP2[19] method are shown in Figure 1. Double-ζ (6-31+G(d,p)) and triple-ζ (6-311+G(2df,2p)) basis sets have been used (Figure 1 a and b, respectively). The double-ζ results emphasize the practical aspect of the study as this type of basis sets is widely used in calculations of large systems. For the set of reactions in Scheme 1, all methods have mean absolute deviations (MADs) ranging from 2–8 kcal mol^{-1} (double-ζ) or 1–12 kcal

---

[*] Dr. S. N. Pieniazek, Dr. F. R. Clemente, Prof. Dr. K. N. Houk
Department of Chemistry and Biochemistry, University of California
Los Angeles, CA 90095-1569 (USA)
Fax: (+1) 310-206-1843
E-mail: houk@chem.ucla.edu

[**] We are grateful to the National Science Foundation and the Partnerships for Advanced Computational Infrastructure (PACI) for the financial support of this research. The computations were performed on the UCLA Academic Technology Services (ATS) Hoffman Beowulf cluster.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200801843.

Table 1: Enthalpies (0 K, in kcal mol^{-1}) for reactions **I**–**V** computed with G3 and CBS-QB3 compound model chemistries and compared with the experimental values, when available.

	G3	CBS-QB3	Exp.[a]
Ia	−37.8	−38.3	−37.6 ± 0.5
IIa	−37.2	−37.6	N/A
IIIa	−36.2	−36.4	N/A
IVa	−31.6	−32.1	−30.9 ± 0.4
Va	−24.2	−24.6	−22.8 ± 8
Ib	−53.2	−53.9	−53.2 ± 0.5
IIb	−52.2	−52.8	N/A
IIIb	−51.4	−51.8	N/A
IVb	−43.5	−44.1	N/A
Vb	−28.5	−28.9	−25.4 ± 8

[a] Experimental reaction enthalpies computed from NIST ΔH_f at 298 K and corrected with B3LYP/6-31+G(d,p) thermal enthalpy corrections. N/A = not available.

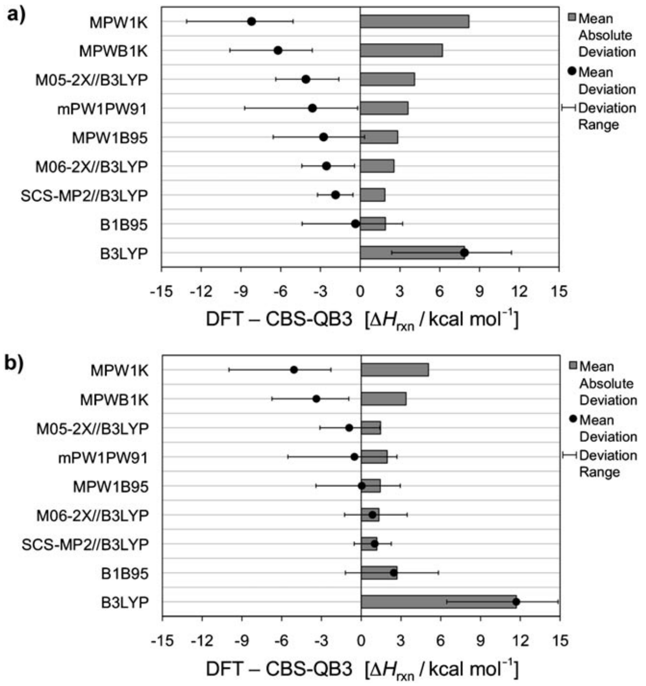
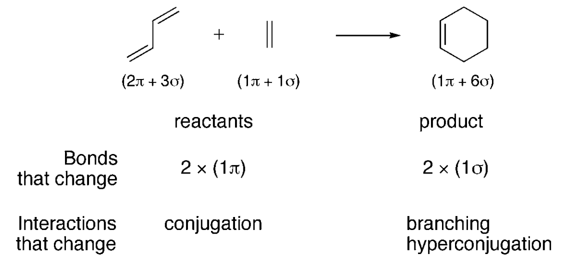


Figure 1. DFT and SCS-MP2 deviations from CBS-QB3 for reactions **I**–**V**: a) double-ζ basis sets results; b) triple-ζ basis sets results.

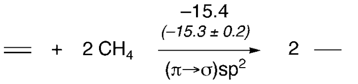
mol^{-1} (triple-ζ) from the CBS-QB3 ΔH_{rxn}. Increasing the basis set size produces a decrease of ca. 3 kcal mol^{-1} in the exothermicity of the reaction for all the computational methods considered. For methods that, on average, overestimate the exothermicities of the reactions, this means that the values would also be closer, on average, to the CBS-QB3 values with the larger basis set. However, much information is lost if only average quantities are considered. As it is obvious from looking at Figure 1, for either basis set, maximum deviations approach ±15 kcal mol^{-1}. More importantly, the errors typically span 7–10 kcal mol^{-1} regardless of the basis set used. For this particular set of reactions, B1B95 (with the double-ζ basis set) and M06-2X (with the triple-ζ basis set) have the lowest MAD, the errors are distributed around 0 albeit in an interval of 5–8 kcal mol^{-1}. Other functionals either consistently underestimate (B3LYP) or overestimate the reaction exothermicities (see Figure 1). Even though the recent M05-2X and M06-2X functionals generally show lower mean errors than most of the other functionals (MAD 3–4 kcal mol^{-1} with the double-ζ basis set and 1–2 kcal mol^{-1} with the triple-ζ basis set), the error ranges are only slightly improved to ca. 5 kcal mol^{-1}. In comparison with the DFT methods, SCS-MP2 has a relatively low MAD, 1–2 kcal mol^{-1}, and the errors distribute in a narrower (3 kcal mol^{-1}) range. Standard MP2 reaction enthalpies **I**–**V**, however, are more similar to those of the DFT methods (deviations span a ca. 4 kcal mol^{-1} range, and MADs of 6 and 3 kcal mol^{-1}).[19b]

In order to determine the origins of these deviations, calculations were performed to assess the errors that occur upon transformations of π to σ bonds, as well as changes in conjugation, branching, or hyperconjugation, all of which also change in a Diels–Alder reaction (Scheme 2).

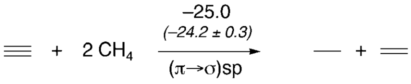


Scheme 2. Bond and interaction changes in Diels–Alder reactions.

There are two essential transformations in C–C bond forming reactions involving hydrocarbons, depending upon whether the π bond of the reactant belongs to an alkene [Eq. (1)] or an alkyne [Eq. (2)], respectively. CBS-QB3 enthalpies ($\Delta H_{0\mathrm{K}}$) for these π to σ transformations are shown in Equations (1) and (2) (experimental values are shown in parentheses).[20]

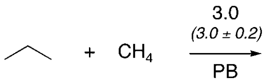


(1)

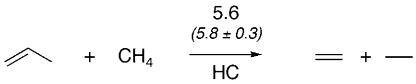


(2)

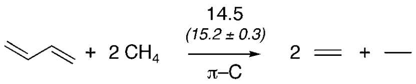
The simplest branching motif, protobranching (PB),[13] is used for the estimation of branching stabilization in reactants and products. The protobranching interaction energy ($\Delta H_{0\mathrm{K}} = 3.0$ kcal mol^{-1} at CBS-QB3 level) corresponds with the bond separation energy[21] of propane as shown in isodesmic Equation (3). Similarly, hyperconjugation (HC) and conjugation (π-C) interaction energies are defined as the bond separation energies of propene and 1,3-butadiene, respectively [isodesmic Equations (4) and (5)]. The CBS-QB3 enthalpies ($\Delta H_{0\mathrm{K}}$) for hyperconjugation and conjugation interactions are 5.6 and 14.5 kcal mol^{-1}.



(3)



(4)



(5)

Figure 2 shows a graphical comparison of these sources of systematic errors given as deviations of DFT from CBS-QB3 energies (ΔH_{0K}) for reactions (1)–(5). The greatest deviations from CBS-QB3 appear in the energies associated with the conversion of π bonds into σ bonds. These energies are highly overestimated by DFT (up to 9 kcal mol^{-1}), with the exception of B3LYP, which underestimates alkene to alkane transformation by 1 kcal mol^{-1} (double-ζ) or 3 kcal mol^{-1} (triple-ζ), and switches from overestimation by 2 kcal mol^{-1} (double-ζ) to underestimation by 1 kcal mol^{-1} (triple-ζ) for alkyne to alkene transformation. These π to σ transformations appear to be the most sensitive to basis set size. For all methods, going from double- to triple-ζ, the exothermicity of the alkene to alkane transformation [Eq. (1)] decreases by ca. 1.5 kcal mol^{-1}, while that of the alkyne to alkene transformation [Eq. (2)] decreases by ca. 2.5 kcal mol^{-1}. Given that the Diels–Alder reaction involves the transformation of two π bonds into two σ bonds, the basis set effect on the π to σ transformations [Eqs. (1) and (2)] acccounts for the observed 3–4 kcal mol^{-1} shift in Diels–Alder reaction energetics (cf. mean deviations in Figure 1 a vs. b).

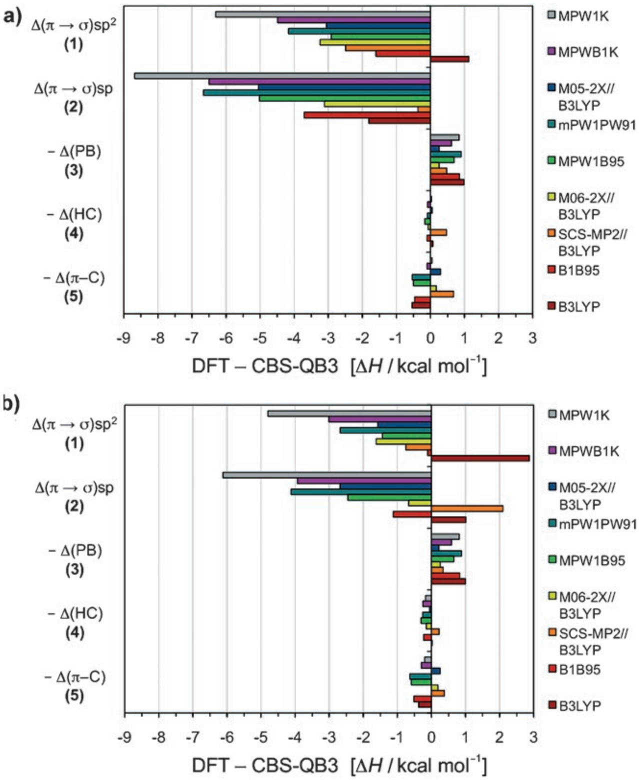


Figure 2. Comparison of DFT and SCS-MP2 deviations from CBS-QB3 values for reaction (1)–(5): a) double-ζ basis sets results; b) triple-ζ basis sets results.

DFT methods satisfactorily estimate hyperconjugation (HC); the deviations from CBS-QB3 are typically ±0.1 kcal mol^{-1} (largest deviation 0.3 kcal mol^{-1} for MPW1B95 with triple-ζ). However, deviations notably increase for the estimation of conjugation (π-C) and protobranching (PB). The former (π-C) is overestimated by ca. 0.5 kcal mol^{-1} by most DFT methods and the deviations increase by 0.1–0.2 kcal mol^{-1} with increasing basis set (only decreases for B3LYP). M05-2X and M06-2X underestimate conjugation by 0.3 and 0.2 kcal mol^{-1}, respectively. Protobranching (PB) components are underestimated by 0.6 to 1.0 kcal mol^{-1} with most DFT methods, except M05-2X and M06-2X (0.3 kcal mol^{-1}), regardless of basis set. M05-2X and M06-2X are the closest to CBS-QB3 in the estimation of protobranching, with B3LYP being the worst in this category. M05-2X and M06-2X, however, are only average in their ability to estimate the energetic requirement to convert π into σ bonds, for which B3LYP (double-ζ) and B1B95 (triple-ζ) are generally the closest to CBS-QB3. The ability of a method to reliably estimate the conversion of π into σ is fundamental and of great importance given the error cancellation that results from the potential underestimation of protobranching (PB).

A startling overestimation of the energy of conversion of π bonds into σ bonds is found with most methods. This is a serious problem, since these bond transformations are involved in reactions such as electrocyclic, ene, and aldol. While these errors may be partially or accidentally cancelled by different sources of errors in the opposite direction, giving a reasonable prediction of overall reaction energies (cf. Figure 1 to Figure 2), all of the DFT functionals (and even low-cost wave function methods) still harbor significant errors for transformations of π into σ bonds.

Received: April 21, 2008

Keywords: addition reactions · computational chemistry · density functional calculations · hydrocarbons · thermochemistry

[1] a) W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed., Wiley-VCH, Weinheim, **2001**; b) A. Chatterjee, *Int. J. Mol. Sci.* **2002**, *3*, 234–236.
[2] a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652; b) P. J. Stephens, F. J. Devlin, C. F. Chablowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627; c) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[3] a) L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, *J. Chem. Phys.* **2000**, *112*, 7374–7383; b) J. Cioslowski, M. Schimeczek, G. Liu, V. Stoyanov, *J. Chem. Phys.* **2000**, *113*, 9377–9389; c) P. Winget, T. Clark, *J. Comput. Chem.* **2004**, *25*, 725–733.

[4] T. P. M. Goumans, A. W. Ehlers, K. Lammertsma, E.-U. Würthwein, S. Grimme, *Chem. Eur. J.* **2004**, *10*, 6468 – 6475.

[5] P. R. Schreiner, *Angew. Chem.* **2007**, *119*, 4295 – 4297; *Angew. Chem. Int. Ed.* **2007**, *46*, 4217 – 4219.

[6] P. C. Redfern, P. Zapol, L. A. Curtiss, K. Raghavachari, *J. Phys. Chem. A* **2000**, *104*, 5850 – 5854.

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

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

[9] a) S. N. Pieniazek, K. N. Houk, *Angew. Chem.* **2006**, *118*, 1470 – 1473; *Angew. Chem. Int. Ed.* **2006**, *45*, 1442 – 1445; b) A. Padwa, K. R. Crawford, C. S. Straub, S. N. Pieniazek, K. N. Houk, *J. Org. Chem.* **2006**, *71*, 5432 – 5439.

[10] S. Grimme, M. Steinmetz, M. Korth, *J. Org. Chem.* **2007**, *72*, 2118 – 2126.

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

[12] a) H. L. Woodcock, H. F. Schaefer III, P. R. Schreiner, *J. Phys. Chem. A* **2002**, *106*, 11923 – 11931; b) Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2006**, *110*, 10478 – 10486; c) Y. Zhao, D. G. Truhlar, *Org. Lett.* **2006**, *8*, 5753 – 5755.

[13] a) M. D. Wodrich, C. Corminboeuf, P. von R. Schleyer, *Org. Lett.* **2006**, *8*, 3631 – 3634; b) M. D. Wodrich, C. S. Wannere, Y. Mo, P. D. Jarowski, K. N. Houk, P. von R. Schleyer, *Chem. Eur. J.* **2007**, *13*, 7731 – 7744.

[14] a) 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; b) P. R. Schreiner, A. A. Fokin, R. A. Pascal, A. de Meijere, *Org. Lett.* **2006**, *8*, 3635 – 3638; c) M. D. Wodrich, C. Corminboeuf, P. R. Schreiner, A. A. Fokin, P. von R. Schleyer, *Org. Lett.* **2007**, *9*, 1851 – 1854.

[15] a) All calculations were performed with the Gaussian 03 suite of programs: M. J. Frisch et al. *Gaussian 03*, revision C.02; Gaussian Inc.: Wallingford, CT, **2004**; b) single point energy calculations with the M05-2X and M06-2X functionals were run with NWChem. 5.0: E. J. Bylaska et al., *NWChem*, version 5.0, Pacific Northwest National Laboratory, Richland, WA, **2006**; c) geometries and unscaled zero point energy corrections of all the structures were obtained with each DFT method and the 6-31+G(d,p) basis set. For M05-2X, M06-2X and SCS-MP2, the electronic energies were computed by single point energy calculations on the B3LYP/6-31+G(d,p) geometries and added zero point energy corrections also computed at the latter level of theory (see Supporting Information for full computational details).

[16] a) M. R. Nyden, G. A. Petersson, *J. Chem. Phys.* **1981**, *75*, 1843 – 1862; b) G. A. Petersson, M. A. Al-Laham, *J. Chem. Phys.* **1991**, *94*, 6081 – 6090; c) G. A. Petersson, T. Tensfeldt, J. A. Montgomery, Jr., *J. Chem. Phys.* **1991**, *94*, 6091 – 6101; d) J. A. Montgomery, Jr., J. W. Ochterski, G. A. Petersson, *J. Chem. Phys.* **1994**, *101*, 5900 – 5909.

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

[18] a) B3LYP: ref. [2]; b) B1B95: A. D. Becke, *J. Chem. Phys.* **1996**, *104*, 1040 – 1046; c) mPW1PW91: C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664 – 675; d) MPW1K: B. J. Lynch, P. L. Fast, M. Harris, D. G. Truhlar, *J. Phys. Chem. A* **2000**, *104*, 4811 – 4815; e) MPW1B95 and MPWB1K: Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 6908 – 6918; f) M05-2X: Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput.* **2006**, *2*, 364 – 382; g) M06-2X: Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215 – 241.

[19] a) S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095 – 9102; b) standard MP2 results are given in the Supporting Information.

[20] Experimental values are computed from NIST (http://webbook.nist.gov/chemistry/) ΔH_f at 298 K and corrected to 0 K using the B3LYP/6-31+G(d,p) thermal energy corrections.

[21] a) W. J. Hehre, R. Ditchfield, L. Radom, J. A. Pople, *J. Am. Chem. Soc.* **1970**, *92*, 4796 – 4801; b) L. Radom, W. J. Hehre, J. A. Pople, *J. Am. Chem. Soc.* **1971**, *93*, 289 – 300.