JCTC Journal of Chemical Theory and Computation

Improved Description of Stereoelectronic Effects in Hydrocarbons Using Semilocal Density Functional Theory

Gábor I. Csonka,*,† Adrienn Ruzsinszky,‡
John P. Perdew,‡ and Stefan Grimme§

Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, H-1521 Budapest, Hungary, Department of Physics and Quantum Theory Group, Tulane University, New Orleans, Louisiana 70118, and Theoretical Organic Chemistry, Organic Chemistry Institute, University of Münster, Corrensstrasse 40, 48149 Münster, Germany

Received January 3, 2008

Abstract: Serious and systematic errors with popular density functionals occur for isodesmic stabilization energies of n-alkanes, isomerization, and dimerization energies of hydrocarbons and geometries of sterically overcrowded aromatic systems. These functionals are too biased toward the correct description of free atoms. Changing two parameters within the Perdew-Burke-Ernzerhof approximation leads to a new nonempirical functional, PBEsol, that improves the description of large organic systems.

Proper description of stereoelectronic (SE) effects[1] is desirable for any theoretical method to be used in organic chemistry. The SE design rules are frequently used in synthetic organic chemistry to design and explain new reactions by electron donating and withdrawing effects or steric interactions. It is shown in this communication that a new, nonempirical semilocal Kohn−Sham (KS) density functional theory (DFT)[2] approximation, called PBEsol,[3] describes more correctly the SE effects in many hydrocarbons than do the semilocal DFT functionals tested in previous reports.[4,5]

KS DFT approximations for the exchange-correlation energy can be classified into two major groups, empirically fitted or nonempirical. The nonempirical functionals form a hierarchy of semilocal approximations with increasing complexity and (usually) accuracy. The simplest approximation is the local spin density approximation (LSDA), while the next two levels are the generalized gradient approximation or GGA (e.g., PBE[6]) and the meta-GGA (e.g., TPSS[7]). The hybrid functionals define a fourth level that is fully nonlocal and semiempirical. In the assessment of density functional approximations, great weight is usually given to the accuracy of molecular atomization energies or the enthalpies of formation constructed from calculated atomization energies. Earlier studies have shown[8] that GGA and global hybrid GGA functionals (e.g., B3LYP,[9] or B3PW91[10]), that were parametrized for the enthalpies of formation of the relatively small molecules, can fail seriously for larger molecules. Due to the computational expense of the exact exchange, calculations using hybrid functionals can be very expensive compared to the pure GGA or meta-GGA functionals. Solid-state calculations are also quite inefficient with hybrid functionals, and thus development of good quality, reliable semilocal functionals is important.

We have shown[11] that the TPSS meta-GGA[7] achieves remarkably accurate atomization energies for 50 large hydrocarbons and substituted hydrocarbons. We also presented strong evidence that most of the error of previous nonempirical functionals resides in the energy of the free atom and so cancels out of typical reaction energies. Finally, we suggested that enthalpies of formation, calculated without any reference to the free atoms, would provide a fairer assessment of the performance of approximate density functionals. (See also refs 12 and 13.)

Recently examples of stereoelectronic effects in alkane isomers have been summarized.[4] Independent works have shown serious and systematic errors of several popular DFT methods (e.g., B3LYP or to a lesser extent PBE) for isodesmic stabilization energies of n-alkanes[5] and octane or $(CH)_{12}$ hydrocarbon isomer energy differences.[14]

More recently it has been argued that popular exchange-correlation functionals are biased toward the correct description of free atoms[3,15] by their gradient coefficients for exchange and thus often fail to improve upon LSDA for solids. These functionals also fail for larger organic molecules. It has been shown that the exact second-order gradient expansion for exchange is relevant for realistic densities of densely packed solids, while the similar expansion for correlation is less relevant. This suggests using the exact second-order gradient expansion coefficient ($\mu_{GE} = 10/81$) within the PBE exchange enhancement factor, $F_x(s)$ where s is a dimensionless density gradient, for such densities. Also adjusting the correlation functional to satisfy another condition leads to a new nonempirical GGA, PBEsol, that performs well for solids and their surfaces and could improve the description of large organic systems and reactions.

---

* Corresponding author e-mail: csonka@web.inc.bme.hu.
† Budapest University of Technology and Economics.
‡ Tulane University.
§ University of Münster.

10.1021/ct800003n CCC: $40.75

With increasing size, the electron densities of molecules become more similar to solid-state densities.

We note that PBE and PBEsol are both first-principles GGAs, but they are based upon different selections of exact constraints to satisfy. At the GGA level, but not at the higher meta-GGA level, one can at most satisfy two out of the following three constraints exactly: (I) second-order gradient expansion for exchange, (II) second-order gradient expansion for correlation, (III) LSD-like linear density response of a uniform electron gas. PBE satifies (II) and (III) but not (I). PBEsol satisfies (I) and compromises between (II) and (III). Any PBE GGA subroutine can be converted to PBEsol simply by changing two gradient coefficients[3] (μ from 0.21951 to 10/81 and β from 0.0667 to 0.0460). Further details of the new PBEsol functional and its performance can be found in ref 3.

In order to estimate the long-range dispersion energy we use the well established and tested DFT-D method,[16] employing damped atom-pairwise -C_6/R^6 terms. (For a recent review see 17.) For PBEsol-D we take the (asymptotically correct) s_6 scaling factor of unity (for PBE $s_6 = 0.75$) and merely readjust the R_{scal} scaling factor of the atomic van der Waals radii from the original value of 1.1 to 1.42 to fit the computed interaction energies of a few typical van der Waals complexes (S22 set, see the Supporting Information). The r_0 value used in the damping function is the sum of the two scaled atomic van der Waals radii.[16] The larger R_{scal} accounts for the more attractive character of the interatomic PBEsol interactions in the medium-range correlation regime (making the dispersion correction more long-ranged). While no GGA without a dispersion correction can account for long-range correlation, the PBEsol (after the error cancelation between exchange and correlation expected for local and semilocal approximations) can apparently account for medium-range interaction better than PBE. Moreover, for Ne_2 and Ar_2, PBE is reasonably good,[18,19] but PBEsol is even better[20] for the repulsive part of the binding energy curve.

Following Zhao and Truhlar,[21] we have used a small test set composed of the following components: experimentally derived zero-point exclusive chemical reaction energies of

$$\text{n-}C_6H_{14} + 4CH_4 \rightarrow 5C_2H_6 \tag{1}$$

and

$$\text{n-}C_8H_{18} + 6CH_4 \rightarrow 7C_2H_6 \tag{2}$$

experimentally derived[22] zero-point exclusive relative energies of 2,2,3,3-tetramethylbutane **3a** and *n*-octane **3b** and relative energies of three $(CH)_{12}$ isomers **4a**, **4b**, and **4c**. We added to the test set the anthracene dimerization reaction energy (**5a** and **5b**). The structures are shown in Figure 1.

We note that the serious discrepancies between computed and experimental isodesmic reaction energies of (1) and (2) have been attributed to the so-called "protobranching" effect, defined as the stabilizing interaction of geminal methyl (-CH_3) or methylene (-CH_2-) groups. It has been shown[4] that many popular exchange-correlation functionals fail for this effect. Later it was shown that the heavily parametrized M05-2X global hybrid functional (for organic chemistry only) performs well on these tests.[21] This was attributed to an improved correlation functional for the medium-range electron correlation. While we concur with the importance of medium-range correlation in these systems, it is interesting to note that the LSDA that is the most local functional performs extremely well for these problems, as observed by Wodrich et al.[5] and shown in Table 1. These good LSDA results can be attributed to the serious overbinding error of the Slater exchange that is imperfectly balanced by the LSDA correlation.[19]

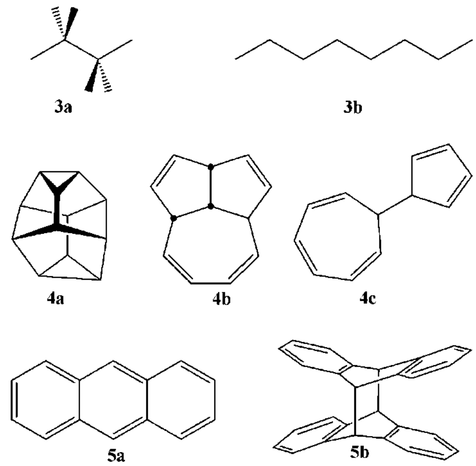

Figure 1. Structures of 2,2,3,3-tetramethylbutane **3a**, *n*-octane **3b**, and the $(CH)_{12}$ isomers, where **4a**, **4b**, and **4c** correspond to the structures 1, 22, and 31 in ref 14. Anthracene is **5a** and its dimer is **5b**.

Table 1. Zero-Point Exclusive Energies (kcal/mol) of Hexane ($\Delta E(1)$) and Octane ($\Delta E(2)$) Isodesmic Stabilization Energies of Reactions 1 and 2 and Relative Energies of Two Octane Isomers

method	ΔE (1) hexane	ΔE (2) octane	E(**3b**)-E(**3a**) octane isomers
expt	9.3[a]	14.0[a]	1.9[a]
CCSD(T)	8.6[b]		
SCS-MP2	8.0[c]	12.0[c]	1.9[c]
MP2	9.7[c]	14.6[c]	5.1[c]
M05-2X	8.2[d]	12.2[d]	2.0[d]
B3PW91	4.7[d]	7.0[d]	−7.0[d]
B3LYP	4.0[b]	5.9[b]	−8.4[d]
TPSS	3.9[e]	5.7[e]	−5.6[e]
PBE	5.6[e]	8.3[e]	−4.6[e]
PBEsol	7.4[e]	11.0[e]	−1.3[e]
PBEsol-D	8.3[e]	12.5[e]	0.3[e]
LSDA	9.9[e]	14.8[e]	3.7[e]

[a] For experimental values see ref 22. [b] Reference 9. [c] Present work, single point frozen core SCS-MP2 or MP2/TZV(*2df,2pd*)//PBEh/TZV(*2d,p*) results. [d] Reference 21. [e] Present work, single point 6-311+G(*2d,2p*)//M05-2X/6-311+G(*2df,2p*) results. PBEsol-D: an estimated dispersion correction was added to PBEsol values.

The results in Table 1 show the relatively good performance of M05-2X, PBEsol, and LSDA for the isodesmic stabilization energies of hexane and octane and for the isomer energy difference of octane. Good performance of PBEsol for $(CH)_{12}$-(E(**4c**)-E(**4b**)) can also be observed in Table 2 (taking CCSD(T) as the standard for **4**; see details below). In contrast, the 8.3 kcal/mol CCSD(T) energy difference between the **4c** and **4b** isomers is seriously underestimated by the PBE, TPSS, and especially the B3LYP functionals. (The B3LYP functional gives very poor results for structures with single bonds only and for bicyclic hydrocarbons, as pointed out in ref 14.) However, E(**4b**)-E(**4a**), involving a large change in the number of multiple bonds, is strongly overestimated by LSDA and PBEsol, while

Table 2. Relative Energies of Three $(CH)_{12}$ Isomers[e]

method	E(**4c**)-E(**4b**)	E(**4b**)-E(**4a**)
CCSD(T)	8.3[a]	12.2[a]
SCS-MP2	8.1[a]	17.3[a]
MP2	7.3[a]	24.4[a]
M05-2X	7.4[b]	14.0[b]
B3PW91	6.7[c]	16.8[c]
B3LYP	1.3[c]	−1.7[c]
TPSS	3.4[d]	10.8[d]
PBE	5.2[d]	13.2[d]
PBEsol	9.4[d]	25.9[d]
PBEsol-D	10.0[d]	26.5[d]
LSDA	11.4[d]	29.3[d]

[a] Present work, single point frozen core CCSD(T)/complete basis set TQ extrapolated// PBEh/TZV(*d,p*) and SCS-MP2 or MP2/TZV(*2df,2pd*)// PBEh/TZV(*2d,p*) results. [b] Reference 21. [c] Reference 14. [d] Present work, single point 6-311+G(*2d,2p*)//M05-2X/6-311+G(*2df,2p*) results. PBEsol-D: an estimated dispersion correction was added to PBEsol values. [e] The number of multiple bonds is zero for **4a**, four for **4b**, and five for **4c**.

PBE, TPSS, and M05-2X give energy differences close to that of CCSD(T). The M06-L functional[23] gives slightly worse results[24] for these problems than the M05-2X functional.

The photodimerization of anthracene (**5a**) is a reversible [4 + 4] cycloaddition which yields the covalently bound polycyclic dimer (**5b**). The condensed phase dimer is labile at elevated temperatures and exothermally dissociates to monomers. The dissociation energy of the gas phase dimer is well studied, and the highest levels of calculations give $D_e = 9 \pm 3$ kcal/mol (endothermic).[25] It can be observed that HF, B3LYP, PBE, and dispersion corrected PBE-D fail considerably for D_e, giving −19.6, −24.6, −13.2, and 3.2 kcal/mol, respectively (using TZV(*2p,d*) basis set and MP2/TZV(*d,p*) geometries; see also ref 26). The LSDA/6-311+G(2*d*,2*p*) D_e is too positive (13 kcal/mol). Our new PBEsol and dispersion corrected PBEsol-D/TZV(*2p,d*) single point energies are considerably better, giving for D_e −0.5 and 9.4 kcal/mol, respectively. (See the Supporting Information).

Longstanding problems for density functionals are the geometries of cyclophanes and related sterically overcrowded aromatic systems.[27] As analyzed in detail in ref 27, typical density functionals properly account for the Pauli-exchange repulsion between the clamped aromatic rings but do not describe accurately the electron correlation effects between the π-systems. These are at distances below the van der Waals radius of carbon, showing that this is not a typical dispersion problem. In effect this leads to too long inter-ring spacings and too strong stretching of the bridging single bonds (cf. Figure 2 for PBE results and ref 27 for even worse B3LYP data).

As can be seen from Figure 2, this problem is more or less solved with PBEsol, which yields almost perfect agreement with experiment for all important structural parameters. This holds in particular for the (chirality inducing) torsion angle ϕ which is too small with PBE and even zero with B3LYP but very accurate with PBEsol. The differences between the PBE and PBEsol geometries are much larger than the expected packing effects. PBEsol geometry is much closer to the SCS-MP2 or experimental geometry[27] than PBE is.

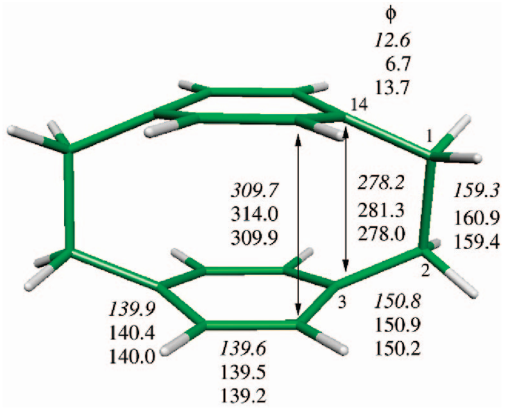

Figure 2. Optimized structure (PBEsol/TZV(*2df,2p*)) of [2.2]paracyclophane **6** (D_2-symmetry) and important geometrical parameters (bond lengths in pm, torsion angle ϕ(3−2−1−14) in °; top: experimental X-ray data,[27] middle: PBE, bottom: PBEsol).

The presented results suggest that a diminished gradient dependence makes PBEsol better than PBE and other standard semilocal functionals not only for solids (moderately varying densities with metallic or single bonds) but also for single-bonded hydrocarbon molecules and for E(**4c**)-E(**4b**), while PBE and TPSS remain better for energy differences between some multiply and singly bonded isomers (e.g., E(**4b**)-E(**4a**) or E(**4c**)-E(**4a**)). PBEsol is clearly better than PBE for the gas phase dissociation energy of the anthracene dimer (**5a** and **5b**). For **6**, where the medium-range electron correlation is important, the new PBEsol functional gives excellent agreement with the experimental geometry. It can be expected that PBEsol, which provides good lattice constants for solids, will also give good geometries for large molecules. Further studies are needed to confirm and refine these tentative conclusions. The stronger enhancement of exchange in PBE GGA and other standard semilocal density functionals is needed for free atoms, for hydrogen bonds, and perhaps for some multiple bonds. In independent work, Vela[28] has also found improvement for octane isomerization energies from a GGA that uses the exact gradient coefficient for exchange over a wide range of density gradients.

In summary, the simple PBEsol GGA, which was developed nonempirically for densely packed solids, is also useful for large organic molecules (in the absence of free atoms). It can be useful for quick evaluations of geometries and frequencies, to be followed possibly by single-point energy calculations at higher levels of approximation. Moreover, it may be possible to build the PBEsol construction principle (recovery of the gradient expansion for exchange over a wide range of slowly- or moderately varying densities) into meta-GGAs, which can be more widely useful than GGAs and are not much more expensive.

Acknowledgment. We thank Y. Zhao and D. G. Truhlar for the M06-L results. Support is acknowledged from OTKA, (PD-050014), NSF (DMR-0501588), and NSF-MTA-OTKA (98).

Appendix: Technical Details

The experimental values used in Table 1 were derived from experimental standard enthalpies of formation.[22] The values were corrected to 0 K, and an experimental zero-point vibration energy correction was subtracted. The error bars for the experimental values are about 0.5 kcal/mol. The frozen core CCSD(T) isomerization energies for **4** were obtained from complete basis set extrapolation using SCS-MP2/cc-pVXZ (X = T, Q) energies. Smaller, double-ζ quality basis sets used in ref 14 give 2−3 kcal/mol basis set error for the energy differences of the $(CH)_{12}$ isomers shown in Table 1. We have calculated the effect of the core-correlation on the relative energies of **4**, and it is comparable to the complete valence basis set extrapolation error (0.3−0.6 kcal/mol). Reference 14 shows that single-point CCSD(T) calculations that use DFT and MP2 geometries give similar isomerization energies within the range of 1 kcal/mol. The M05-2X/6-311+G(*2df,2p*), B3PW91 and B3LYP results obtained with triple-ζ quality basis sets are taken from the literature.[4,8,14,21] The M05-2X/cQZV3P//MP2/TZV(*d,p*) model gives slightly different values for ΔE (1), ΔE (2), and E(**3b**)-E(**3a**) (within 0.6 kcal/mol). For the TPSS, PBE, PBEsol, and LSDA (with the SVWN5 functional for LSDA) calculations, we use 6-311+G(*2d*,*2p*), cc-pVTZ(-*f*), and TZV(2*d*,*p*) or TZV(2*df*,2*p*) basis sets of Ahlrichs et al.[29] and PBE/6-311+G(*d,p*), M05-2X/6-311+G(*2df,2p*) and MP2/TZV(*d,p*) geometries. Using the complete cc-PVTZ basis set (adding *f* functions) changes the ΔE (1) by 0.2 kcal/mol, so using *f* functions has a small influence on the calculated results. The DFT calculations were performed with the modified GAUSSIAN 03 program[30] and with the slightly modified versions of the Turbomole suite of programs.[31] It was observed that the geometry differences result in a maximum of 0.6 kcal/mol change in the energies of ΔE (1), ΔE (2), and E(**3b**)-E(**3a**). The basis set effect is about 1 kcal/mol. We confirm the suggestion of ref 31: Triple-ζ quality basis sets are needed for such DFT calculations, and an increase of the polarization part to (*2d*,*2p*) is advantageous.

Supporting Information Available: Geometries and/or total energies of compounds **4**, **5**, and **6** and dissociation energies and statistical data for the S22 benchmark set. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) *Stereoelectronic Effects in Organic Chemistry: P. Delongchampsin, Organic Chemistry Series*; Baldwin, J. E., Ed.; Pergamon: Oxford, 1983; Vol. *1,* pp 5−20.

(2) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.

(3) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L.; Zhou, X.; Burke, K. *Phys. Rev. Lett.* **2008**, *100*, 136406.

(4) Grimme, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 4460.

(5) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2006**, *8*, 3631.

(6) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(7) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(8) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.

(9) Stephens, P. J.; Devlin, F. J.; Chablowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(10) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(11) Csonka, G. I.; Ruzsinszky, A.; Tao, J.; Perdew, J. P. *Int. J. Quantum Chem.* **2005**, *101*, 506.

(12) Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067.

(13) Brothers, E. N.; Scuseria, G. E. *J. Chem. Theory Comput.* **2006**, *2*, 1045.

(14) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A., Jr; de Meijere, A *Org. Lett.* **2006**, *8*, 3635.

(15) Perdew, J. P.; Constantin, L. A.; Sagvolden, E.; Burke, K. *Phys. Rev. Lett.* **2006**, *97*, 223002.

(16) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(17) Grimme, S.; Antony, J.; Schwabe, T.; Mück-Lichtenfeld, C. *Org. Biomol. Chem.* **2007**, *5*, 741.

(18) Tao, J. M.; Perdew, J. P *J. Chem. Phys.* **2005**, *122*, 114102.

(19) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I. *J. Phys. Chem. A.* **2005**, *109*, 11015.

(20) Csonka, G. I., private communication.

(21) Zhao, Y.; Truhlar, D. G. *Org. Lett.* **2006**, *8*, 5753.

(22) NIST Standard Reference Database. See http:// webbook.nist.gov/chemistry/ (accessed Feb 1, 2008).

(23) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.

(24) Zhao, Y.; Truhlar, D. G., private communication.

(25) Grimme, S.; Diedrich, C.; Korth, M. *Angew. Chem., Int. Ed.* **2006**, *45*, 625.

(26) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(27) Grimme, S. *Chem. Eur. J.* **2004**, *10*, 3423.

(28) Vela, A., private communication.

(29) For standard basis sets see http://www.emsl.pnl.gov/forms/basisform.html (accessed Feb 19, 2008). The basis sets for Turbomole are available from the Turbomole homepage via the FTP server button (in the subdirectories basen, jbasen, and cbasen). See http://www.turbomole.com (accessed Feb 19, 2008).

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria G. E. Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; 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.; Bakken, V.; 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 D.01; Gaussian, Inc.: Wallingford, CT, 2004.

(31) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169. Turbomole (Version 5.9): R. Ahlrichs et al., Universität Karlsruhe, 2003. See also: http://www.turbomole.com (accessed Feb 1, 2008)

(32) Grimme, S.; Steinmetz, M.; Korth, M. *J. Org. Chem.* **2007**, *72*, 2118.

CT800003N