

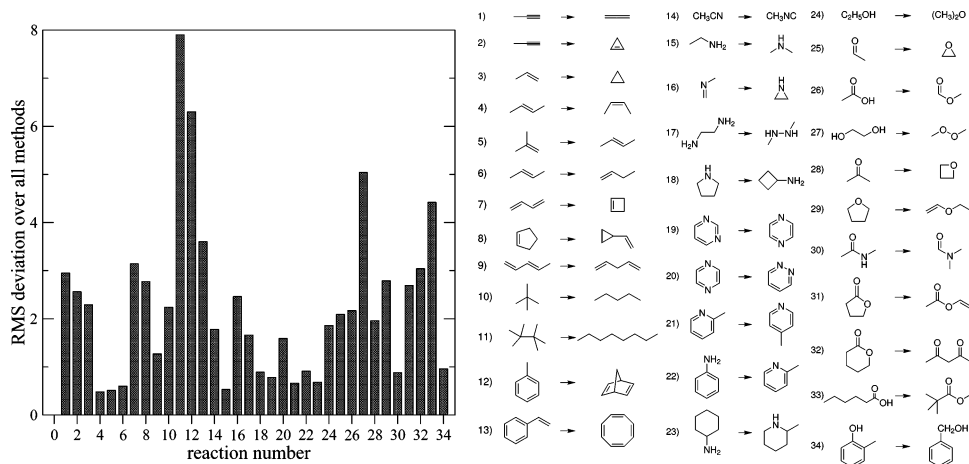
# How to Compute Isomerization Energies of Organic Molecules with Quantum Chemical Methods

Stefan Grimme,\* Marc Steinmetz, and Martin Korth

Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster,  
Corrensstrasse 40, D-48149 Münster, Germany

grimmes@uni-muenster.de

Received November 28, 2006



The reaction energies for 34 typical organic isomerizations including oxygen and nitrogen heteroatoms are investigated with modern quantum chemical methods that have the perspective of also being applicable to large systems. The experimental reaction enthalpies are corrected for vibrational and thermal effects, and the thus derived “experimental” reaction energies are compared to corresponding theoretical data. A series of standard AO basis sets in combination with second-order perturbation theory (MP2, SCS-MP2), conventional density functionals (e.g., PBE, TPSS, B3-LYP, MPW1K, BMK), and new perturbative functionals (B2-PLYP, mPW2-PLYP) are tested. In three cases, obvious errors of the experimental values could be detected, and accurate coupled-cluster [CCSD(T)] reference values have been used instead. It is found that only triple- $\zeta$  quality AO basis sets provide results close enough to the basis set limit and that sets like the popular 6-31G(d) should be avoided in accurate work. Augmentation of small basis sets with diffuse functions has a notable effect in B3-LYP calculations that is attributed to intramolecular basis set superposition error and covers basic deficiencies of the functional. The new methods based on perturbation theory (SCS-MP2, X2-PLYP) are found to be clearly superior to many other approaches; that is, they provide mean absolute deviations of less than 1.2 kcal mol<sup>-1</sup> and only a few (<10%) outliers. The best performance in the group of conventional functionals is found for the highly parametrized BMK hybrid meta-GGA. Contrary to accepted opinion, hybrid density functionals offer no real advantage over simple GGAs. For reasonably large AO basis sets, results of poor quality are obtained with the popular B3-LYP functional that cannot be recommended for thermochemical applications in organic chemistry. The results of this study are complementary to often used benchmarks based on atomization energies and should guide chemists in their search for accurate and efficient computational thermochemistry methods.

## 1. Introduction

There has been considerable interest and progress in the development of reliable quantum chemical methods for predic-

tion of thermochemical data of molecules.<sup>1</sup> The extent to which modern high-level calculations can be competitive with experiment in the precise determination of heats of formation strongly depends on the size of the investigated molecular species. For three- to four-atomic systems with up to about 20 electrons, there

\* To whom correspondence should be addressed. Phone: (+49)-251-8336512.

is little question that ab initio wave function theory (WFT) can provide very accurate energetic data (errors <1 kJ/mol).<sup>2</sup> For larger systems of practical relevance, however, one has to resort to more approximate methods because accurate coupled-cluster type treatments [CCSD(T) or QCISD(T)] have an unfavorable scaling behavior with system size and are thus computationally not feasible anymore.

Many of the electronic structure methods that are applicable to large systems cannot be systematically improved, and thus, empirical tests in general chemistry applications are mandatory. This holds in particular for methods based on Kohn–Sham density functional theory (DFT) that is, since the mid 1990s, the most widely used method in quantum chemistry.<sup>3,4</sup> However, in the last few years, the early euphoria about the performance of DFT in chemistry has given way to a more realistic view, and problem cases are more and more investigated in theoretical chemistry. Recently, for the most popular B3-LYP<sup>5</sup> hybrid density functional, huge and chemically very relevant errors for organic systems/reactions have been reported.<sup>6–11</sup> On the other hand, from the recent chemical literature that include electronic structure calculations for solving distinct problems, one can get the impression that the combination with a quite small AO basis set (B3-LYP/6-31G(d)) is the “swiss army knife” among quantum chemical methods. Although it is now quite well-known that the relatively good performance of B3-LYP/6-31G(d) stems mainly from error compensation<sup>12–15</sup> and more accurate alternatives for thermochemistry and kinetics are available,<sup>16–22</sup> the method is still amazingly popular especially

in organic chemistry applications (for recent work, see refs 23–26).

One major aim of the present work is to encourage (computational) chemists to use quantum chemical methods that are more accurate and reliable than B3-LYP/6-31G(d) in order to get “the right answer for the right reason”. We focus here on a benchmark set of organic isomerization reactions that recently has been used by Jorgensen et al.<sup>27,28</sup> in studies of lower-level (semiempirical) methods. Very often, quantum chemical methods are benchmarked on atomization energies (heat of formations), and well-established sets such as the large G3/05<sup>29</sup> with more than 450 entries are in common use. These tests, however, may provide somewhat biased conclusions because systematic errors in the description of free (open-shell) atoms compared to (mostly closed-shell) molecules enter very strongly. Although it is clear that atomization is some kind of worst-case scenario for quantum chemical methods, it is quite far away from typical chemistry applications. This is the reason why we<sup>17,30,31</sup> and other authors<sup>13,14,32–34</sup> use more and more reaction energies, barrier heights, or atom-equivalent schemes in energetic benchmark studies. Isomerization reactions are perfectly suited for this purpose because these are well-defined and accompanied often by big changes in electronic structure, but difficult higher-order effects (e.g., vibrational anharmonicity, relativity, core–core correlation) mostly cancel out and thus need not to be considered.

After a brief description of the benchmark set used (see Chart 1) and some considerations about the accuracy of the experimental reference data, we first discuss the AO basis set dependence of the results for selected methods. The choice of the AO basis is the most obvious but also most ignored source of error in quantum chemical computations. In a second step we evaluate a range of representative quantum chemical methods/density functionals (MP2, SCS-MP2, PBE/0, TPSS, B97-D, B-P86, B-LYP, B3-LYP, O3-LYP, MPW1K, MPWB1K, BMK) that have the perspective of being applicable routinely to large molecules. For convenience, only one reliable AO basis is used in this final assessment. We conclude with some considerations on computational efficiency, the cost-performance ratio, and some general methodical recommendations.

(1) (a) Irikura, K. K.; Frurip, D. J., Eds. *Computational Thermochemistry*; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998. (b) Cioslowski, J., Ed. *Quantum Mechanical Prediction of Thermochemical Data*; Kluwer: Dordrecht, The Netherlands, 2001.

(2) (a) Tajti, A.; Szalay, P. G.; Csaszar, A. G.; Kallay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vazquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599. (b) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108. (c) Bomble, Y. J.; Vazquez, J.; Kallay, M.; Michauk, C.; Szalay, P. G.; Csaszar, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.

(3) (a) Kohn, W.; Sham, L. J. *Phys. Rev. A* **1965**, *140*, 1133–1138. (b) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.

(4) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: New York, 2001.

(5) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(6) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Raghavachari, K. *J. Phys. Chem. A* **2000**, *104*, 5850.

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

(8) Grimme, S. *Angew. Chem.* **2006**, *118*, 4571; *Angew. Chem., Int. Ed.* **2006**, *45*, 4460.

(9) Wodrich, M. D.; Corminboeuf, C.; Schleyer, P. v. R. *Org. Lett.* **2005**, *8*, 3631.

(10) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A.; de Meijere, A. *Org. Lett.* **2005**, *8*, 3631.

(11) Grimme, S.; Steinmetz, M.; Korth, M. *J. Chem. Theory Comput.* **2007**, *3*, 42.

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

(13) Cioslowski, J.; Schimeczek, M.; Liu, G.; Stoyanov, V. *J. Chem. Phys.* **2000**, *113*, 9377.

(14) Winget, P.; Clark, T. *J. Comput. Chem.* **2004**, *25*, 725.

(15) Goumans, T. P. M.; Ehlers, A. W.; Lammertsma, K.; Würthwein, E.-U.; Grimme, S. *Chem.—Eur. J.* **2004**, *10*, 6468.

(16) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405.

(17) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(18) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.

(19) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2005**, *123*, 121103.

(20) Mori-Sanchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2006**, *124*, 091102.

(21) Grimme, S.; Schwabe, T. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398.

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

(23) Gordillo, R.; Houk, K. N. *J. Am. Chem. Soc.* **2006**, *128*, 3543.

(24) Northrop, B. H.; O'Malley, D. P.; Zografos, A. L.; Baran, P. S.; Houk, K. N. *Angew. Chem., Int. Ed.* **2006**, *45*, 4126.

(25) Galabov, B.; Ilieva, S.; Schaefer, H. F., III. *J. Org. Chem.* **2006**, *71*, 6382.

(26) Baldwin, J. E.; Raghavan, A. S.; Hess, B. A., Jr.; Smentek, L. J. *Am. Chem. Soc.* **2006**, *128*, 14854.

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

(28) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. A* **2006**, *110*, 13551.

(29) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(30) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

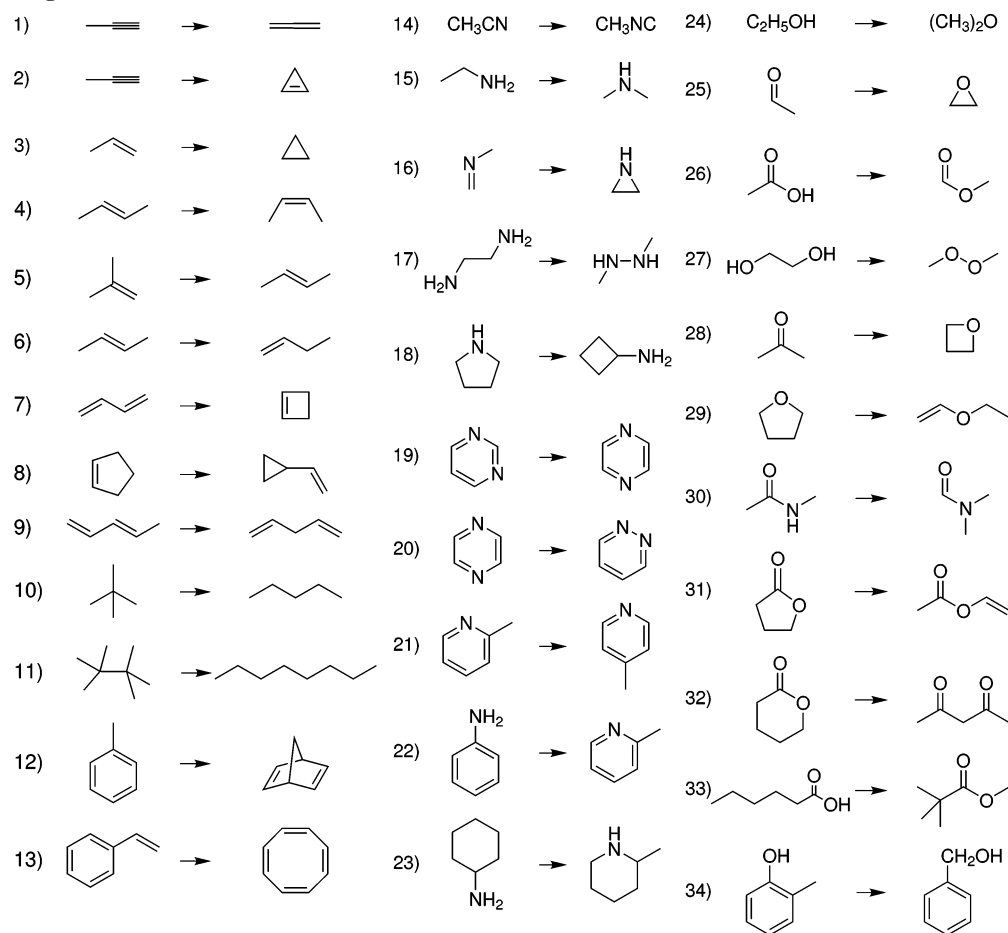
(31) Grimme, S. *J. Phys. Chem. A* **2005**, *109*, 3067.

(32) Baker, J.; Pulay, P. *J. Chem. Phys.* **2002**, *117*, 1441.

(33) Zhao, Y.; Gonzalez-Garcia, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012.

(34) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.

## CHART 1. Investigated Reactions



## 2. Methods and Technical Details

The SCS-MP2 and DFT calculations have been performed with slightly modified versions of the TURBOMOLE suite of programs.<sup>35</sup> The CCSD(T) computations were performed with MOLPRO,<sup>36</sup> and for the MPW1K/MPWB1K calculations, we used Gaussian03.<sup>37</sup> As AO basis, valence double- $\zeta$  (SV), triple- $\zeta$  (TZV), or quadruple- $\zeta$  (QZV) sets of Ahlrichs et al.,<sup>37–41</sup> Dunning's

correlation consistent sets cc-pVXZ<sup>42</sup> (X = D, T, Q; augmentations (aug) from ref 43) or the Pople sets 6-31G(d)<sup>44</sup> and 6-311G(d,p)<sup>45</sup> have been employed. The exponents of polarization functions for TZV and QZV were taken from the corresponding cc-pVXZ basis sets. Note, that TZV and QZV are particularly efficient for many programs due to the relatively small number of Gaussian primitives. Furthermore, they suffer less than the cc-pVXZ basis sets from basis set superposition error (BSSE).<sup>40</sup> Intentionally (and opposed to others<sup>46</sup>) we are not recommending basis sets including spatially diffuse functions for neutral organic molecules (see also section 3.2). Especially together with small valence basis sets (e.g., 6-31G+G(d) or similar) huge intramolecular BSSE can occur. For the effect of using unbalanced basis sets, see the recent example of artificial out-of-plane distortions in aromatic molecules.<sup>47</sup> Furthermore, diffuse functions significantly slow down SCF convergence for large molecules and more or less destroy efficient integral-screening techniques in the SCF procedure. The

(35) (a) Ahlrichs, R. *TURBOMOLE*, version 5.7; Universität Karlsruhe, 2003 (see <http://www.turbomole.com>). (b) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

(36) Werner, H.-J. *MOLPRO*, version 2006.1, a package of ab initio programs (see <http://www.molpro.net>).

(37) 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.; 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 C.01; Gaussian, Inc.: Wallingford, CT, 2004.

(38) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.

(39) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.

(40) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753.

(41) The basis sets are available from the TURBOMOLE homepage via the FTP server button (in the subdirectories basen, jbasen, and cbasen). See <http://www.turbomole.com>.

(42) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(43) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(44) (a) Ditchfeld, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724. (b) Hehre, W. J.; Ditchfeld, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(45) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(46) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.

(47) Moran, D.; Simmonett, A. C.; Leach, F. E., III; Allen, W. D.; Schleyer, P. v. R.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **2006**, *128*, 9342.



overall result is typically (when comparing, e.g., cc-pVXZ with aug-cc-pVXZ) an increase of computation time by a factor between 3 and 10.

Four groups of methods are tested: ab initio second-order perturbation theory in standard form (MP2<sup>48</sup>) or in the spin-scaled variant (SCS-MP2<sup>50</sup>) that improves in almost all cases on MP2; double hybrid density functionals that (besides including “exact” Hartree–Fock exchange) replace part of the local correlation by a perturbative MP2-type expression (X2-PLYP). These functionals contain only two empirical scaling parameters, and we employ here variants with Becke’s exchange GGA (X = B, B2-PLYP<sup>17</sup>) or with modified Perdew–Wang exchange (X = mPW, mPW2-PLYP<sup>21</sup>). At a lower level, we have hybrid functionals that only account for the nonlocality in the exchange part, namely, B3-LYP, O3-LYP,<sup>49</sup> PBE0,<sup>50</sup> MPW1K,<sup>51</sup> MPWB1K,<sup>52</sup> and BMK.<sup>16</sup> The last three functionals have been proposed especially for reaction barriers, and it seems interesting to find out their performance for reaction energies. The computationally cheapest methods (with resolution of the identity (RI, also called density-fitting) integral approximation they scale as  $O(N_{el}^3)$  with system size) are pure (non-hybrid) functionals such as B-LYP,<sup>53,54</sup> PBE,<sup>55</sup> B97-D,<sup>56</sup> and TPSS.<sup>57</sup> PBE and TPSS mainly differ by kinetic energy density ingredients (meta-GGA) in the latter, and this comparison should provide some insight into the success of this approach. The BMK and MPWB1K functionals are also of the meta-hybrid type although semiempirical in nature. The B97-D functional is a recent reparametrization of Becke’s GGA ansatz from 1997<sup>58</sup> but now including damped, atom-pairwise  $C_6 \times R^{-6}$  corrections to account for dispersion (van der Waals) effects. Because CCSD(T) with (too) small AO basis is sometimes used as reference method in the literature, we also include CCSD(T)/6-31G(d) values for comparison. Results obtained at the CCSD(T)/cc-pVTZ level are expected to yield on average errors of less than 1 kcal mol<sup>-1</sup> (maximum errors of 2–3 kcal mol<sup>-1</sup> at most) and are mainly used to check the consistency of the experimental reference data.

In all second-order perturbative treatments (MP2, SCS-MP2, and X2-PLYP), the RI approximation for the two-electron integrals has been used.<sup>59</sup> As RI auxiliary basis, the sets of Weigend et al.<sup>60</sup> that were optimized for the cc-pVXZ AO basis have been employed. The RI approximation has also been used in DFT treatments with pure GGA functionals<sup>61</sup> employing the corresponding optimized auxiliary basis sets.<sup>62</sup> In general, the RI approximation speeds up calculations by about a factor of 3–15 depending on system/basis size, albeit at insignificant loss of accuracy (errors for  $\Delta E < 0.05$  kcal mol<sup>-1</sup> compared to non-RI treatments<sup>60,63</sup>). In the (SCS)-MP2 and CCSD(T) computations, we employ the frozen-core (FC) approximation, while all electrons have been correlated for consistency in the X2-PLYP treatments. Note that correlating the core electrons in (SCS)-MP2 or CCSD(T) but not using appropriate

AO basis sets containing steep basis functions leads to larger errors due to BSSE effects. For the set of reactions investigated here, the MAD *increases* by about 0.2 kcal mol<sup>-1</sup> with the TZV(2df,2dp) basis when including the core electrons in the correlation treatment compared to FC computations.

We tested B3-LYP/TZV(d,p) and B97-D/TZV(d,p) for geometry optimization but found insignificant differences between these methods. This holds in general for not too “exotic” molecules and also for other functionals because DFT yields quite accurate structures<sup>4</sup> and remaining errors mostly cancel for isomerizations. In order to be consistent with most previous benchmark studies, we decided finally to employ the B3-LYP/TZV(d,p)-optimized structures in all subsequent single-point energy computations. The experimental reaction enthalpies were corrected to zero-point vibrational (ZPV) exclusive energies by employing computed harmonic vibrational frequencies (scale factors of 0.97 and 1.0, respectively, for B3-LYP/TZV(d,p) and B97-D/TZV(d,p)). The scaling factors are taken from ref 64, which also includes a discussion about the accuracy of the derived ZPV energies. For some molecules, thermal population of energetically higher-lying conformers contributes significantly to the enthalpy at 298 K, and in these cases, we also computed these corrections at the B97-D/TZV(d,p) level (see Supporting Information, Table S1). The thus derived “experimental” isomerization energies  $\Delta E$  differ at most by 0.5 kcal mol<sup>-1</sup> when B3-LYP/TZV(d,p) or B97-D/TZV(d,p) geometries/harmonic frequencies are used. The typical difference between both methods is about 0.1–0.2 kcal mol<sup>-1</sup>, which is much smaller than common errors for experimental heats of formation (about 0.5 kcal mol<sup>-1</sup>) from which the isomerization enthalpies are derived. Overall, the largest corrections from enthalpy to energy are about 2 kcal mol<sup>-1</sup> for reactions 11 and 30 (the average correction is only 0.6 kcal mol<sup>-1</sup>; see Table S1). Our final estimate for the accuracy of the reference data of typically  $\sim 0.5$  kcal mol<sup>-1</sup> is thus composed of an experimental part and two smaller uncertainties due to the use of inaccurate DFT structures and ZPV energies.

### 3. Results and Discussion

**3.1. General Considerations and Accuracy of the Experimental Data.** The benchmark set of organic isomerization reactions (see Chart 1) has been taken from the literature.<sup>27,28</sup> It consists of 13 reactions of hydrocarbons (entries 1–13), 10 reactions involving nitrogen (entries 14–23), and 11 with oxygen (entries 24–34) atoms. Common bonding situations for organic molecules are present in a well-balanced manner, although it should be noted that more complicated electronic structures (e.g., carbenes, radicals, and biradicals) as well as (transition) metals are not included (for recent benchmarks in this area, see refs 65–68). Transition states have been extensively used for benchmark purposes in recent years, see, for example, refs 15, 33, 69, and 70. As described above, the experimental isomerization enthalpies are derived from standard (gas phase) heat of formations<sup>71</sup> and corrected to energies (zero-

(48) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(49) How, W.-M.; Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2001**, *341*, 319.

(50) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

(51) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.

(52) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.

(53) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(54) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

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

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

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

(58) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554.

(59) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331.

(60) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175.

(61) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283.

(62) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119.

(63) Grimme, S.; Waletzke, M. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2075.

(64) Sinha, P.; Boesch, S. E.; Gu, C.; Wheeler, R. A.; Wilson, A. K. *J. Phys. Chem. A* **2004**, *108*, 9213.

(65) Hyla-Kryspin, I.; Grimme, S. *Organometallics* **2004**, *23*, 5581.

(66) Furche, F.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 044103.

(67) Quintal, M. M.; Karton, A.; Iron, M. A.; Boese, A.; Martin, J. M. L. *J. Phys. Chem. A* **2006**, *110*, 709.

(68) Schultz, N. E.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 11127.

(69) Grimme, S.; Mück-Lichtenfeld, C.; Würthwein, E.-U.; Ehlers, A. W.; Goumans, T. P. M.; Lammertsma, K. *J. Phys. Chem. A* **2006**, *110*, 2583.

(70) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936.

(71) NIST Standard Reference Database; see <http://webbook.nist.gov/chemistry/>.

**TABLE 1.** Comparison of Experimental<sup>a</sup> and Computed<sup>b</sup> Isomerization Energies  $\Delta E$  (in kcal mol<sup>-1</sup>)

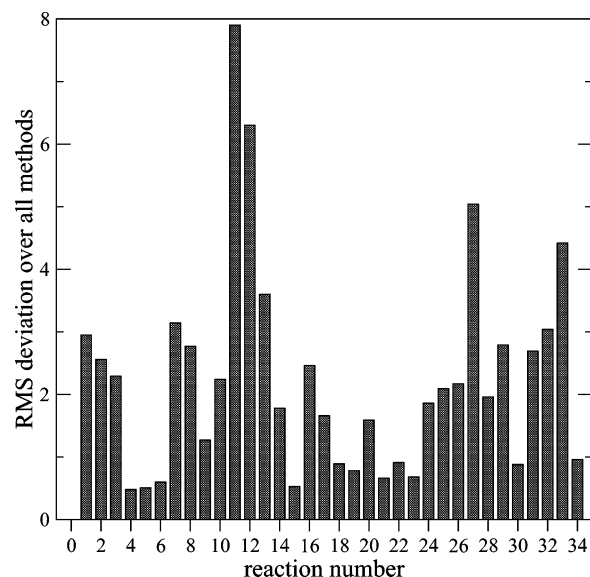
reaction	$\Delta E$				
	exptl	SCS-MP2	mPW2-PLYP	PBE	B3-LYP
1	1.6	3.9	0.7	-3.1	-2.0
2	21.9	24.4	24.3	19.2	23.9
3	7.2	6.4	7.7	5.2	9.0
4	1.0	1.2	1.2	1.1	1.3
5	0.9	1.1	0.8	0.4	0.3
6	2.6	2.4	2.9	3.6	3.2
7	11.1	10.5	13.1	10.4	15.4
8	22.9	22.1	21.3	20.1	19.8
9	6.9	6.5	7.5	9.0	8.2
10	3.6	4.0	2.5	1.5	0.8
11	1.9	1.6	-2.7	-5.3	-8.0
12	46.9	45.3	53.0	50.9	57.1
13	36.0	38.0	39.7	39.0	39.2
14 <sup>c</sup>	24.2 (21.3)	25.8	24.5	24.9	23.5
15	7.3	7.8	7.7	7.5	7.3
16	10.8	10.1	11.2	8.5	12.5
17	27.0	28.4	27.2	25.3	25.6
18	11.2	11.6	11.9	10.7	11.5
19 <sup>d</sup>	4.6 (0.0)	4.0	4.4	3.8	4.1
20	20.2	18.4	18.6	17.0	18.4
21	0.9	1.1	1.2	1.1	1.2
22	3.2	2.4	3.4	4.4	3.6
23	5.3	5.4	5.0	4.7	4.6
24	12.5	11.6	11.2	11.2	10.6
25	26.5	26.2	27.6	24.6	28.0
26	18.2	16.2	16.5	16.7	16.0
27	64.2	65.8	64.2	58.5	60.6
28	31.2	30.9	32.7	30.7	33.5
29	11.9	13.7	11.4	11.6	8.8
30 <sup>d</sup>	9.5 (13.6)	9.1	9.6	9.0	9.6
31	14.0	16.0	13.7	14.7	10.9
32 <sup>d</sup>	7.1 (2.4)	5.9	5.5	5.8	3.3
33	5.6	8.2	9.2	10.6	10.2
34	7.3	6.1	6.9	8.6	6.9

<sup>a</sup> From ref 71; see also section 2. <sup>b</sup> Single-point energy calculations on B3-LYP/TZV(d,p) optimized geometries and employing the TZV(2df,2pd) AO basis. <sup>c</sup> CCSD(T)/cc-pV(TQ)Z//B3-LYP/TZV(d,p); experimental value in parentheses. <sup>d</sup> CCSD(T)/cc-pV(DT)Z//B3-LYP/TZV(d,p); experimental value in parentheses.

point vibrational energy exclusive) denoted in the following as  $\Delta E$ . These values are given together with those from selected theoretical methods in Table 1.

Of crucial importance for any benchmark study is the accuracy of the reference data. During our work, it became evident that, for four out of the 34 reactions considered, the experimental data are doubtful. For reactions 14 (CH<sub>3</sub> → CH<sub>3</sub>-NC), 19 (pyrimidine → pyrazine), 30 (*N*-methylacetamide → dimethylformamide), and 32 (caprolactone → acetylacetone), all methods (and almost independent of basis set used) furnish large deviations with respect to the experimental values (about 3, 5, 4, and 5 kcal mol<sup>-1</sup> for reactions 14, 19, 30 and 32, respectively) always of the same sign and magnitude (see Table 1). In these cases, we performed CCSD(T) computations using AO basis set extrapolation<sup>72</sup> (cc-pVXZ, X = D–T for reactions 19, 30, and 32 and X = T–Q for reaction 14) that are expected to be accurate to about ±0.5 kcal mol<sup>-1</sup> for  $\Delta E$ . These results confirmed our view as these CCSD(T) data are in much better agreement with those from the lower-level methods than the experimental values. We thus replaced in these four cases the experimental values by the corresponding CCSD(T) reference data. Although we see no other apparent deviations, it cannot

(72) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

**FIGURE 1.** Root-mean-square deviation over 17 quantum chemical approaches (see Table 3) for all reactions considered.

be excluded that errors on the order of about 0.5 kcal mol<sup>-1</sup> still contaminate the reference values. These may only in part be related to problems with the combustion measurements on which the heat of formations are based but instead on the experimental phase change data which are often used to correct for condensed phase effects. Because experimental thermochemistry seems to be a dying out part of chemical science, the use of entirely theoretical reference data seems to be a solution to the problem. At present, however, such computations (i.e., CCSD(T)/cc-pV(TQ)Z) are not feasible for many molecules considered. We estimate that the achievable mean absolute deviation (MAD) for the test set (that has an average  $\Delta E$  of 14 kcal mol<sup>-1</sup>) is limited to about 0.5 kcal mol<sup>-1</sup> which should be kept in mind in the following discussion.

All reactions considered are formulated such that  $\Delta E$  is positive. A closer inspection of the data in Table 1 reveals that sign errors ( $\Delta E < 0$ ) only occur for the density functional methods. All functionals have problems with the isomerization of branched to linear alkanes (reactions 10, 11, and 33) as investigated in detail recently<sup>8,11</sup> as well as for alkyne → cummulene isomerization (reaction 1, see also refs 73 and 74). The degree of difficulty in the description of each isomerization has been analyzed statistically by taking for 17 different quantum chemical methods (see section 3.3, CCSD(T) omitted) the root-mean-square (rms) deviation. The result is shown graphically in Figure 1.

Clearly, the set is composed of about a third of very simple reactions (rms < 1 kcal mol<sup>-1</sup>), a bunch of systems of intermediate complexity (rms of about 2 kcal mol<sup>-1</sup>), and a few very problematic cases. Beside the ones mentioned above, these are in particular the butadiene → cyclobutene (7), toluene → norbornadiene (12), and ethanediol → dimethylperoxide (27) reactions.

**3.2. Basis Set Dependence.** As examples for conventional density functionals and perturbative wave function methods,

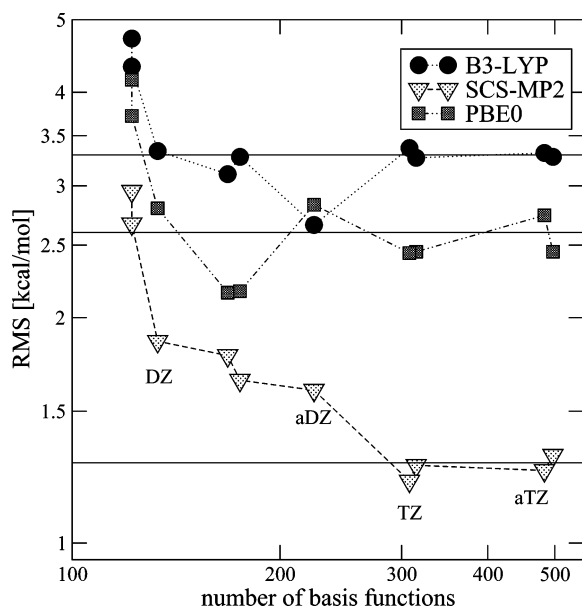
(73) Woodcock, H. L.; Schaefer, H. F., III; Schreiner, P. R. *J. Phys. Chem. A* **2002**, *106*, 11923.

(74) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 10478.

**TABLE 2.** Statistical Assessment of SCS-MP2 and B3-LYP Methods with Various AO Basis Sets (in kcal mol<sup>-1</sup>)

basis	cost <sup>a</sup>	SCS-MP2		B3-LYP	
		rms <sup>b</sup>	max <sup>c</sup>	rms <sup>b</sup>	max <sup>c</sup>
aug-cc-pVTZ	211.5 (483)	1.25	2.8	3.32	10.3
QZV(3d2f,3p2d)	143.3 (497)	1.31	2.9	3.28	10.3
cc-pVTZ	28.5 (308)	1.22	2.7	3.37	10.3
TZV(2df,2pd)	26.4 (315)	1.27	2.6	3.27	10.2
aug-cc-pVDZ	10.4 (224)	1.60	3.3	2.66	8.3
TZV(d,p)	3.6 (175)	1.65	3.6	3.28	9.8
6-311G(d,p)	3.1 (168)	1.78	4.6	3.11	9.2
cc-pVDZ	2.1 (133)	1.86	4.4	3.34	10.1
6-31G(d)	1.1 (122)	2.67	7.9	4.33	14.4
SV(d)	1.0 (122)	2.95	8.5	4.72	14.5

<sup>a</sup> Relative CPU time (SV(d) value set to unity) for a B3-LYP energy computation on 1-methylpyridine. Number of contracted AO basis functions in parentheses. <sup>b</sup> Mean absolute deviation. <sup>c</sup> Maximum absolute deviation.



**FIGURE 2.** Double-logarithmic plot of the rms deviation as a function of AO basis size (number of functions for 1-methylpyridine) for the various sets considered. The horizontal lines indicate the estimated basis set limit (taken as an average of the aug-cc-pVTZ and QZV3P results). The lines connecting the data points are merely used to guide the eye.

we first looked at B3-LYP, PBE0, and SCS-MP2 and investigated their basis set dependence. According to previous experience,<sup>17,21,69</sup> the X2-PLYP functionals behave similar to SCS-MP2. As will be discussed below, the B3-LYP functional is not very representative for hybrid GGAs as it is extremely sensitive with respect to diffuse functions in the basis set. We merely use it here because of its popularity in applications. The statistical assessment is given in Table 2 for B3-LYP and SCS-MP2, and rms deviations (also for PBE0) are graphically shown in Figure 2.

From top to bottom in Table 2, the AO basis sets are ordered according to decreasing computational cost. For valence-only sets, this is roughly proportional to the cube of the number of basis functions which also provides a measure of quality. The augmented sets (aug-cc-pVXZ, X = D, T) are more costly than their non-augmented counterparts by a factor of 5–7. The large QZV(3d2f,3p2d) set can be considered as being very close to the basis set limit for isomerization reactions. It is clearly seen

from Figure 2 that B3-LYP and SCS-MP2 behave differently with respect to AO basis set improvement. In the case of B3-LYP, except for valence double- $\zeta$  sets such as 6-31G(d,p) and SV(d), both the rms and maximum deviations are almost constant down the columns. Because this more or less also holds for the PBE0 functional (that shows a more oscillatorial behavior), we conclude that it thus makes little sense to go beyond the triple- $\zeta$  level in conventional DFT calculations. A notable exception is the aug-cc-pVDZ basis containing diffuse (spatially extended) functions which provides overall the best results with B3-LYP (but not with SCS-MP2). Note that augmentation has no effect when the valence basis is already quite saturated, that is, when going from cc-pVTZ to aug-cc-pVTZ. This indicates that the good aug-cc-pVDZ results (and this also holds for sets such as 6-31+G(d,p); see ref 46) are due to intramolecular BSSE which compensates basic deficiencies of the functional. As already mentioned, augmented basis sets are computationally quite demanding and are (despite giving the best B3-LYP results) not recommended also because they do not provide “the right answer for the right reason”. This view is supported by the results for the PBE0 hybrid functional where the rms values change only marginally with augmentation (from 2.80 to 2.83 kcal mol<sup>-1</sup> for X = D) and are even enlarged compared to triple- $\zeta$  type sets. The strange behavior of B3-LYP is best illustrated with reaction 11 where a small but balanced set (e.g., cc-pVDZ) yields a  $\Delta E$  of  $-8.2$  kcal mol<sup>-1</sup>, which is quite close to the estimated basis set limit ( $-8.0$  kcal mol<sup>-1</sup> with QZV3P or aug-cc-pVTZ). Due to intramolecular BSSE, the aug-cc-pVDZ set now stabilizes the spatially more compact branched isomer compared to the linear isomer which results in a putatively better value for  $\Delta E$  of  $-6.4$  kcal mol<sup>-1</sup> (exptl 1.9 kcal mol<sup>-1</sup>). Note that in neither of these B3-LYP computations the essential physics of the problem (electron correlation on medium length scales) is considered correctly; for details, see ref 8.

For B3-LYP (and this also holds for other functionals), even underpolarized but otherwise balanced sets such as 6-311G(d,p) or TZV(d,p) provide results reasonably close to the basis set limit. Note the poor performance of the popular B3-LYP/6-31G(d) combination; that is, the MAD and maximum deviations are larger by about 0.8 and 4 kcal mol<sup>-1</sup>, respectively, compared to most other sets.

The SCS-MP2 method is for all tested basis sets much more accurate than B3-LYP, and furthermore, SCS-MP2 benefits more from an increasing quality of the AO basis than B3-LYP (see Figure 2). This clearly reflects the more systematic behavior of wave function based compared to DFT methods. Both the MAD/rms as well as the maximum deviations decrease monotonically with the quality of the AO basis until at about the aug-cc-pVTZ AO or QZV(3d2f,3p2d) level, the methodical limit is reached. Changes in MAD/rms below 0.1 kcal mol<sup>-1</sup> should be taken not too seriously as long as more accurate reference data are unavailable.

To conclude this section, we recommend, in general, properly polarized triple- $\zeta$  basis for SCS-MP2 (and also other perturbative) and conventional DFT approaches and discourage from the use of double- $\zeta$  type sets. When aiming at high accuracy or when more complicated bonding situations are involved, increasing the polarization part to (2d,2p) or even (2df,2pd) is expected to be beneficial. Going beyond the triple- $\zeta$  level (e.g., cc-pVQZ or QZV3P) has little impact as long as only electronically quite “normal” bonding situations are involved.



**TABLE 3.** Statistical Assessment of Various Methods with the TZV(2df,2pd) Basis Set (in kcal mol<sup>-1</sup>)

method	type <sup>a</sup>	rms <sup>b</sup>	MAD <sup>c</sup>	max <sup>d</sup>	# outliers <sup>e</sup>
CCSD(T) <sup>f,g</sup>	WFT	0.95	0.68	2.3 (33)	0
SCS-MP2	WFT	1.27	1.03	2.6 (2)	0
BMK	hybrid meta-GGA (42%)	1.79	1.28	4.7 (7)	4
mPW2-PLYP	perturbative hybrid GGA (55%)	1.83	1.19	6.1 (12)	4
B2-PLYP	perturbative hybrid GGA (53%)	1.93	1.32	6.0 (12)	4
MP2	WFT	2.04	1.45	6.2 (13)	3
PBE0	hybrid GGA (25%)	2.45	1.79	7.0 (11)	7
MPWB1K	hybrid meta-GGA (44%)	2.47	1.89	4.9 (16)	11
PBE	GGA	2.54	1.89	7.3 (11)	6
BH-LYP	hybrid GGA (50%)	2.66	1.66	8.9 (11)	4
B-P86	GGA	2.70	1.88	8.4 (11)	6
CCSD(T) <sup>h</sup>	WFT	2.78	1.82	9.6 (27)	6
MPW1K	hybrid GGA (42.8%)	2.79	2.05	7.6 (11)	10
B97-D	GGA	3.00	2.07	9.3 (12)	10
B3-LYP	hybrid GGA (20%)	3.27	2.29	10.2 (12)	11
TPSS	meta-GGA	3.46	2.52	11.4 (27)	10
O3-LYP	hybrid GGA (16.6%)	3.48	2.13	14.4 (11)	8
HF-SCF	WFT	3.79	2.67	12.9 (11)	11
B-LYP	GGA	4.43	3.24	11.7 (12)	15

<sup>a</sup> WFT: wave function theory; GGA: generalized gradient approximation (non-hybrid) functional. The number in parentheses is the admixture of nonlocal HF exchange in a hybrid functional. Meta functionals include kinetic energy densities (occupied orbital-dependent quantities). <sup>b</sup> Root-mean-square deviation. <sup>c</sup> Mean absolute (unsigned) deviation. <sup>d</sup> Maximum absolute (unsigned) deviation. Corresponding reaction number in parentheses. <sup>e</sup> Number of absolute deviations > 3 kcal mol<sup>-1</sup>. <sup>f</sup> cc-pVTZ AO basis. <sup>g</sup> Excluding reaction 23 for which the calculations could not be performed. <sup>h</sup> 6-31G(d) AO basis.

**3.3. Assessment of Quantum Chemical Methods.** The various quantum chemical methods mentioned in the Introduction are used together with the TZVP(2df,2pd) basis to compute all isomerization energies, and the resulting statistical assessment is given in Table 3. The methods are ordered according to accuracy (rms deviation). Results from the Hartree-Fock (HF-SCF) method that entirely neglects electronic Coulomb correlation effects are included for comparison.

Perusing Table 3, one finds that the methods can be grouped according to accuracy into four classes. This is more or less independent of the four statistical measures as they all provide the same qualitative picture (exceptions are BH-LYP and MPWB1K that have about the same rms deviation but very different error distributions (e.g., 4 vs 11 outliers) and MAD, respectively). The ordering of the methods within one group should be taken not too seriously (at least in the first three) because our statistical analysis is based only on 34 data points.

The CCSD(T) as well as the SCS-MP2 methods furnish by far the most accurate isomerization energies<sup>75</sup> as judged not only from very small MADs of about 0.7 and 1.0 kcal mol<sup>-1</sup> but also from the absence of any outliers with errors > 3 kcal mol<sup>-1</sup>. This is very encouraging because the SCS-MP2 method can be applied routinely [opposed to CCSD(T)] to much larger molecules than investigated here. In the next group, we find the perturbative methods mPW2-PLYP, B2-PLYP, standard MP2, and the meta-hybrid BMK, which is the most accurate of the conventional functionals tested. These methods perform well enough according to common standards with MAD between about 1.3 and 1.5 kcal mol<sup>-1</sup> and only a few

(about 10%) outliers. However, MP2 can (considering the excellent SCS-MP2 results) be regarded as obsolete. For the X2-PLYP functionals, the largest error occurs for the reaction from toluene to norbornadiene (12) that is also very problematic for B3-LYP (error of 10.2 kcal mol<sup>-1</sup>). This (and the observation that this reaction is less problematic for other functionals) seems to indicate a problem with the LYP correlation functional in this case.

The methods in the next group (PBE0, MPWB1K, PBE, BH-LYP, B-P86, CCSD(T)/6-31G(d)) are significantly less accurate (MAD of 1.8–1.9 kcal mol<sup>-1</sup>), and in particular, the number of outliers and the maximum errors (7–9 kcal mol<sup>-1</sup>) are very significant. Note the dramatic decrease of the performance of CCSD(T) when a too small AO basis set is used. Especially, the good results with PBE, which is a pure, non-empirical GGA, are encouraging as GGAs are very efficient in RI treatments for the two-electron Coulomb terms. PBE also performs quite well for metal-containing systems.<sup>66</sup> Note that for PBE (opposed to the BX-LYP family of functionals) the amount of exact exchange admixture (i.e., when going from PBE to PBE0) has only a small effect on the performance. For the relatively new meta-hybrid functional MPWB1K, we notice many small errors but also a large number of outliers.

At the bottom of the list with MAD values > 2 kcal mol<sup>-1</sup> and ≥ 10 outliers, we find to our surprise methods that are commonly considered as being very accurate. Note that these methods perform in this test not much better than uncorrelated HF-SCF. TPSS that has been specially advocated<sup>57</sup> as an improved successor of PBE is in fact significantly worse. This confirms previous conclusions about its performance in main group thermochemistry applications<sup>31</sup> and shows how difficult the non-empirical route to better density functionals is.

The bad performance of B3-LYP with an MAD of 2.3 kcal mol<sup>-1</sup>, 11 large errors and a maximum deviation of about 10 kcal mol<sup>-1</sup> (for reactions 11 and 12), comes as no surprise when considering recent findings in the literature.<sup>6–10</sup> This functional even falls back behind pure GGAs that are computationally much less demanding. Particularly disappointing is

(75) As suggested by a reviewer, we also tested SCS-MP3 that scales, however, as  $O(N^6)$  with system size and is thus significantly more costly than SCS-MP2 (but much cheaper than CCSD(T)). We obtain at the RI-SCS-MP3(FC)/TZV(2df,2pd) level a rms (MAD) of 1.17 (0.92) kcal mol<sup>-1</sup>, which is slightly better than with SCS-MP2. For details about SCS-MP3 that seems to be beneficial for electronically more complicated cases see refs 65 and 77.

not only the size of the errors on average but even more the large number of outliers indicating missing robustness for organic reactions. Note that the effect of the amount of HF exchange mixing (that is usually considered as a major factor) is very pronounced only in the BX-LYP family of functionals (the rms decreases from 4.4 to 3.3 to 2.7 kcal mol<sup>-1</sup> when reducing it from 0 to 50% in the series B-LYP, B3-LYP, and BH-LYP). For our test set, O3-LYP (essentially replacing B88 exchange in B3-LYP with OPTX<sup>76</sup>) which has been specially recommended for reaction energies<sup>32</sup> is even worse than B3-LYP. This seems to indicate a general problem with the mixing procedure of the GGA, LDA, and HF exchange in these three-parameter hybrids.

Finally, we want to comment on the mean (signed) deviations (MD). In general, these are found to be more or less randomly distributed around  $\pm 0.5$  kcal mol<sup>-1</sup> with a tendency of being more negative (i.e., underestimating the isomerization energies) for the conventional functionals. A notable exception is TPSS that underestimates the  $\Delta E$  values on average by 1.7 kcal mol<sup>-1</sup>. Most accurate (i.e., close to zero MD) are CCSD(T)/cc-pVTZ, SCS-MP2, the X2-PLYP functionals, and MPW1K.

#### 4. Summary and Conclusions

Our work adheres to the general ambition of making quantum chemical methods applicable to large parts of chemistry at affordable computational costs. This requires new method developments as well as careful testing and calibration of existing methods. We studied here 34 organic isomerization reactions involving small to medium-sized (neutral) molecules. The systems are large enough to allow some general conclusions but already represent the limit of what routinely can be done using accurate coupled-cluster treatments. We developed a benchmark set (that has been selected by others) further by correcting for thermal and vibrational effects such that the purely electronic isomerization energies can be used in future benchmark studies. In four cases, we replaced suspicious experimental data by accurate basis set extrapolated CCSD(T) values, and we estimated a typical error of  $\pm 0.5$  kcal mol<sup>-1</sup> for each of the reference energies that are on average 14 kcal mol<sup>-1</sup>.

A series of standard AO basis sets in combination with second-order perturbation theory methods and representative density functionals have been tested. Note that our conclusions are based solely on organic isomerization reactions involving first-row elements. Good transferability of our results is expected for other closed-shell reaction types as well as for (nonmetallic) main group chemistry. In this sense (and when accuracy is the major concern), we can recommend and conclude the following:

(1) For closed-shell reactions, the SCS-MP2 results are very close to those from CCSD(T) but at a tiny fraction of the computational cost. SCS-MP2 with properly polarized triple- $\zeta$  basis sets is the recommended method for main parts of organic chemistry and should thus replace standard MP2.

(2) The perturbatively corrected (virtual orbital dependent) X2-PLYP density functionals (in particular, mPW2-PLYP) are almost as accurate as SCS-MP2 but have a broader range of application (i.e., for transition metal complexes or for open-shell systems suffering from spin-contamination problems). They are free of “over-fitting” problems and thus “robustness” can be expected.

(3) Quite surprisingly, the most popular B3-LYP hybrid density functional performs inferior to most other methods and is even outperformed by pure GGAs. Because isomerization often involves a change of shape of the molecules and a reorganization of the bonds around the atoms, this finding confirms recent studies<sup>8–11</sup> on peculiar B3-LYP problems. Because of clear evidence from our group and work from others, we strongly recommend omitting B3-LYP from thermochemical applications in general. For preliminary investigations or geometry optimizations, we can recommend pure GGAs such as PBE or B97-D that are very efficient together with the RI approximation. If nonlocal exchange effects are important, the non-empirical PBE0 hybrid functional can also be recommended.

(4) No definite conclusion is possible for the highly empirical BMK functional (17 fitting parameters) as the success found here at least in part contrasts recent findings for general main group reactions and some transition states.<sup>17</sup> As the other meta-functionals (MPWB1K and TPSS) are also not very accurate, the success of including kinetic energy density ingredients seems at least questionable and should be investigated further.

(5) Although this seems to be common knowledge in the quantum chemical community, we still want to comment on the choice of AO basis. Double- $\zeta$  type sets such as 6-31G(d) are very incomplete and cannot accurately describe even simple (closed-shell) electronic reorganizations as considered here. This in particular holds for wave function based correlation treatments, and thus, we consider a CCSD(T)/6-31G(d) computation in general as a waste of computation time. Properly polarized triple- $\zeta$  basis sets (e.g., 6-311G(2d,2p) or TZV(2d,2p)) can be recommended in many cases when electronically not too difficult situations are involved.

Less clear are general considerations of computational cost that have to be considered in practical work. The reason is that the computation time depends in a complicated way on software, hardware, special algorithms used, system size, technical parameters (convergence or integral neglect thresholds), and other issues. According to our experience that should be based on (a) computations using the RI approximation, (b) for systems with up to 50–100 non-hydrogen atoms, and (c) single-point energies, we can conclude the following.

The largest difference in computation time between the three groups: pure GGA functional (1), (meta)hybrid-GGA functional (2), or second-order perturbative method (3) occurs between 1 and 2 which is roughly a factor of 5–10. When using RI, the methods in groups 2 and 3 cause almost the same computational effort because the SCF step (that is necessary in both) is rate determining and the perturbative part matters only for very large systems.

In summary, from the perspective of the cost performance ratio, we recommend skipping computations with conventional hybrid density functionals altogether. Instead, cheap treatments employing pure (local) GGAs (mainly for geometry optimizations and vibrational frequency computations) should be combined with accurate SCS-MP2 or X2-PLYP treatments. It remains to be seen how the latest generation of highly parametrized meta-hybrid functionals such as the new M05-2X<sup>18</sup> or local versions<sup>22</sup> of it that claim high accuracy in many cases perform in practice.

(76) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.

(77) Grimme, S. *J. Comput. Chem.* **2003**, *24*, 1529.



**Acknowledgment.** This work was supported by the Deutsche Forschungsgemeinschaft in the framework of the SFB 424 (“Molekulare Orientierung als Funktionskriterium in chemischen Systemen”). The authors thank C. Mück-Lichtenfeld for technical assistance and helpful discussions.

**Supporting Information Available:** Tables with ZPV energy and thermal corrections, B3-LYP/TZV(d,p) optimized geometries of all molecules, and reaction energies for all methods in Table 3. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JO062446P