

Theor Chem Acc (2010) 127:429–442  
DOI 10.1007/s00214-010-0818-3

FEATURE ARTICLE

# Overcoming systematic DFT errors for hydrocarbon reaction energies

Stephan N. Steinmann · Matthew D. Wodrich ·  
Clemence Corminboeuf

Received: 3 August 2010 / Accepted: 6 September 2010 / Published online: 28 September 2010  
© Springer-Verlag 2010

**Abstract** Despite the widespread use and numerous successful applications of density functional theory, descriptions of hydrocarbon reaction energies remain problematic. Illustrative examples include large underestimation of energies associated with alkane bond separation reactions and poor general description of intramolecular dispersion in hydrocarbons (e.g., B3LYP, MAD = 14.1 kcal mol<sup>-1</sup>). More recent, but not readily available functionals, along with efficient *posteriori* corrections, not only show considerable improvement in the energy description of hydrocarbons but also help identify the sources of error in traditional DFT. Interactions in branched alkanes and compact hydrocarbons are adequately mimicked by systems compressed below their typical van der Waals distances. At these distances, standard DFT exchange functionals are overly repulsive for non-bonded density overlaps, and significant improvement is offered by the long-range corrected exchange functionals (e.g., LC-BLYP0.33, MAD = 5.5 kcal mol<sup>-1</sup>). For those systems, the neglect of long-range dispersion is found to be a critical shortcoming, as well as “overlap dispersion”, for which non-negligible amounts are

captured by the correlation functional. Accounting for the missing dispersion interactions is of key importance. Accordingly, most noteworthy improvements over standard functionals are obtained by using non-local van der Waals density functionals (e.g., LC-S-VV09, MAD = 3.6 kcal mol<sup>-1</sup>, rPW86-VV09, MAD = 5.8 kcal mol<sup>-1</sup>), a dispersion corrected double hybrid (B2PLYP-D, MAD = 2.5 kcal mol<sup>-1</sup>), or by the addition of an atom pairwise density-dependent dispersion correction to a standard functional (e.g., PBE-dDXDM, MAD = 0.8 kcal mol<sup>-1</sup>). To a lesser extent, the reduction of the delocalization error (e.g., MCY3, MAD = 6.3 kcal mol<sup>-1</sup>) or careful parameter fitting (e.g., M06-2X, MAD = 5.6 kcal mol<sup>-1</sup>) also lowers the errors.

**Keywords** Density functional theory · Weak intramolecular interactions · Hydrocarbon energies

## 1 Introduction

Kohn–Sham density functional theory (DFT) [1] is a powerful framework for many aspects of electronic structure theory and has become the preferred method for modeling the energy and structural properties of large molecules. Despite overwhelming popularity, common semi-local and hybrid density functional approximations are affected by well-known deficiencies. In addition to the inability of the most popular exchange–correlation functionals to accurately model long-range (dispersion) interactions in van der Waals complexes [2–11], recent studies have also noted failures to describe intramolecular energies in seemingly simple hydrocarbons [12–15].

Alkanes represent the simplest examples of organic molecules for which the energetic description remains challenging for density functional approximations. The

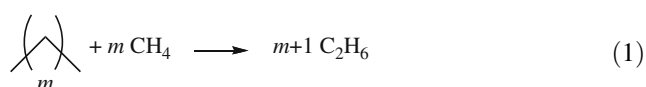
**Electronic supplementary material** The online version of this article (doi:10.1007/s00214-010-0818-3) contains supplementary material, which is available to authorized users.

S. N. Steinmann · M. D. Wodrich · C. Corminboeuf (✉)  
Laboratory for Computational Molecular Design,  
Ecole Polytechnique Fédérale de Lausanne,  
Institut des Sciences et Ingénierie Chimiques,  
CH-1015 Lausanne, Switzerland  
e-mail: clemence.corminboeuf@epfl.ch

*Present Address:*

M. D. Wodrich  
School of Chemistry and Biochemistry, University of Geneva,  
Geneva, Switzerland

DFT accuracy for alkane energies is generally benchmarked through computed heats of formation [16–20] or reaction energies [21–23]. Disturbing failures have, for instance, been noted for the evaluation of isodesmic bond separation reactions [21–23] of *n*-alkanes (Eq. 1; Fig. 1), which are commonly used to determine the total sum of the (de)stabilizing interactions within molecules. In the bond separation equation (BSE) procedure, all bonds between heavy (non-hydrogen) atoms are split into their simplest (or parent) molecular fragments preserving the heavy atom bond types. Reactions are balanced by inclusion of the necessary number of simple hydrides (methane, ammonia, water, etc.). The BSE of propane (Eq. 1) has been used to quantify the 1,3-methyl-methyl stabilizing interaction, 2.83 kcal mol<sup>-1</sup>, termed protobranching by Schleyer and coworkers [24]. Most functionals systematically underestimate the BSE of propane and larger alkanes (Fig. 1) [13, 25, 26].



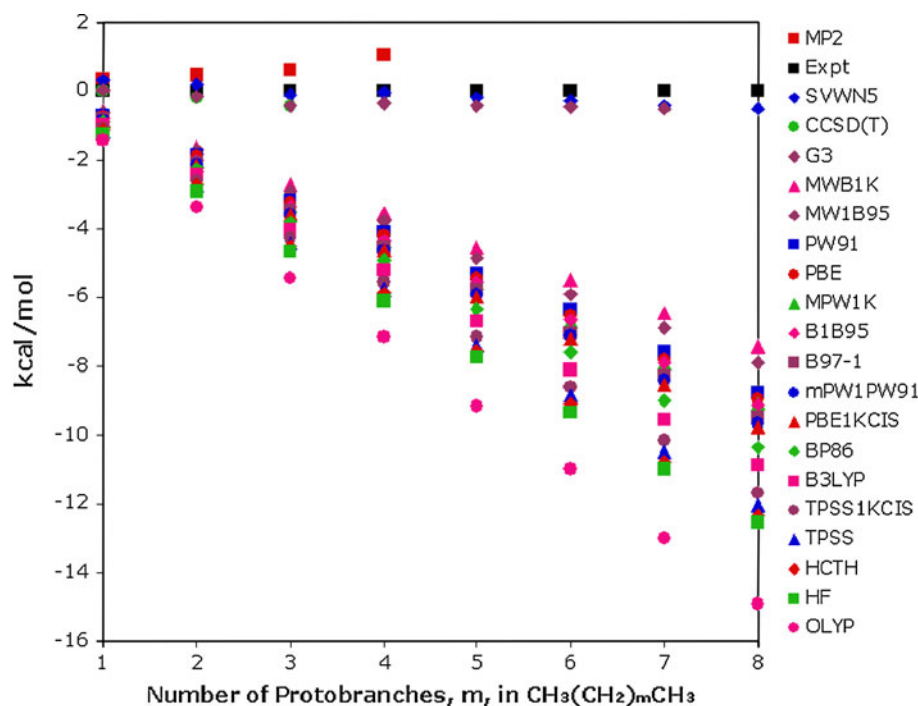
While the physical origin of the branching stabilization remains uncertain, its roots trace to both Allen's [27, 28] and Pitzer's [29, 30] 1950s studies of alkane stability, where van der Waals type (London dispersion [31]) interactions were invoked to explain the enhanced stability of branched over linear species. For this reason, the poor treatment of non-bonded intramolecular interactions between 1,3-disposed methyl/methylene groups have been

proposed to explain the failure of conventional density functionals [26, 32–38].

Numerous related problems have been reported for the computation of hydrocarbon bond energies [39–44]. Redfern et al. [19] noted large per bond B3LYP deviations for alkane heats of formation, which grew with increasing alkane size. Grimme's analysis of the alkane isomerization reactions showed that many density functionals do not reproduce the correct ordering for heats of formation, preferring *n*-alkanes over their more highly branched (and more stable) counterparts [12]. Feng et al. [39] first noted considerable underestimation of C–C bond energies using B3LYP [45, 46]. Later, Check and Gilbert [47] showed increasing errors in C–C bond energies as the peripheral hydrogen atoms were replaced with methyl groups, resulting in an error of over 20 kcal mol<sup>-1</sup> for cleavage of the central C–C in tetramethylbutane. Note that the cleavage of these C–C bonds necessarily involves changes in the number of 1,3-alkyl-alkyl (protobranching) interactions, thus, similar shortcomings are observed for bond cleavage energies as for alkane BSEs.

Organic systems possessing structural features other than C–C and C–H single bonds are also susceptible to DFT failures. Schreiner and coworkers showed that gradient corrected functionals over stabilize cummulene structures when examining cummulene/acetylene energy differences in small organic compounds [48]. Different functionals yielded widely varying energy differences for sets of hydrocarbon structural isomers (over 40 kcal mol<sup>-1</sup> for C<sub>12</sub>H<sub>12</sub>) [49]. For the same set of compounds, bond

**Fig. 1** Deviations of various DFT functionals from experimental (OK) protobranching stabilization energies. Negative values denote underestimation. Stabilization energies are based on Eq. 1. CCSD(T) and MP2 refer to CCSD(T)/aug-cc-pVTZ//MP2/6-311+G(d,p) and MP2/aug-cc-pVTZ//MP2/6-311+G(d,p), respectively, and include MP2/6-311+G(d,p) zero-point corrections. All other computations employed the 6-311+G(d,p) basis set. Copyright American Chemical Society



separation reactions give generally smaller errors [50] that increase with the system size.

Over the past decades, the shortcomings of conventional semi-local and hybrid density functionals have motivated developments “beyond” the realm of traditional DFT. These more sophisticated and/or more accurate formalisms generally aim at (1) improving the treatment of long-range dispersive interactions, which, by construction, is missing in semi-local density functionals and (2) reducing the intrinsic self-interaction error, [51–53] generalized to “delocalization error” in many electron systems [54–57]. In this article, we discuss the current state of density functional approximations for describing energies associated with weak intramolecular interactions present in four test sets featuring hydrocarbons. Our primary focus is to evaluate the performance of the most recent but not widely used DFT methods and test their ability to overcome known deficiencies. The latest approximations are often implemented in developmental versions of codes, frequently unavailable to users. A comprehensive benchmarking of “modern” functionals on the reactions energies of hydrocarbons will not only enable straightforward comparisons of their performance, but also serve to supplement the debate regarding the origin of the DFT errors [12, 26, 58, 59]. Our classification and applications of “modern” density functionals distinguish “pure” from atom pairwise dispersion-corrected density functionals. The “pure” class includes vdW-density functionals (e.g., vdW-DF04 [60] and VV09 [61, 62]), long-range corrected exchange functionals (e.g., LC-BLYP [26, 63], LC-PBE [26], LC- $\omega$ PBE [64, 65], LC- $\omega$ PBEh [65]), and functionals designed specifically to minimize the delocalization errors (e.g., MCY3) [66]. For comparisons, this category also includes the increasingly popular highly parameterized hybrid Minnesota functional, M06-2X [67]. The second class of methods considers Grimme’s B97-D [68] and B2PLYP-D [32, 69], along with our recent density-dependent dispersion correction dDXDM [37], in which an explicit atom pairwise dispersion correction is added *a posteriori*.

## 2 Computational methods

### 2.1 Test sets

The performance of a series of recent functionals (described below) is compared with traditional semi-local and hybrid functionals for four test sets representative of the intramolecular weak interactions in hydrocarbons (Figs. 2 and 3). Three of these test sets assess Pople’s isodesmic bond separation equation reactions [21–23] of alkanes (chains, rings and cages, see Fig. 2). The geometries and thermal corrections are taken at the B3LYP/6-311+G\*\* level from

ref [35, 37] except for R14, which has been updated. Reference values are derived from experimental heats of formation [70]. Whereas the use of these experimental reference data has been disputed [71], our recent density-dependent dispersion correction (-dDXDM) [37] was fit to experimental energy differences, which compared well with composite approaches (see Figure S1 in the Supporting Information). For the sake of comparisons, results based on computed ab initio benchmark values (i.e., CCSD(T)/CBS), which are best reproduced by B2PLYP-D, are given in the Supporting Information. Note that the general trends and analysis discussed in this work, however, remain unaffected by the choice of reference values (experimental vs. CCSD(T)/CBS, see Supporting Information). The “intramolecular dispersion interactions in hydrocarbons” (IDHC) [32] set contains two isomerization reactions (*n*-octane and *n*-undecane to the fully branched isomer), two folding reactions of large hydrocarbon chains ( $C_{14}H_{30}$  and  $C_{22}H_{46}$ ), the dimerization of anthracene, and the hydrogenation reaction of [2.2]paracyclophane to *p*-xylene (Fig. 3). Geometries and reference values are taken from ref [32]. Finally, the geometries and CCSD(T\*)-F12a/aug-cc-pVTZ [72, 73] reference values for the methane dimer, used as a model system, are taken from reference [37] and the interaction energy profile completed at the same level using Molpro2009.1 [74]. SAPT0 computations were performed with the aug-cc-pVTZ basis set [75] in Molpro2009.1.

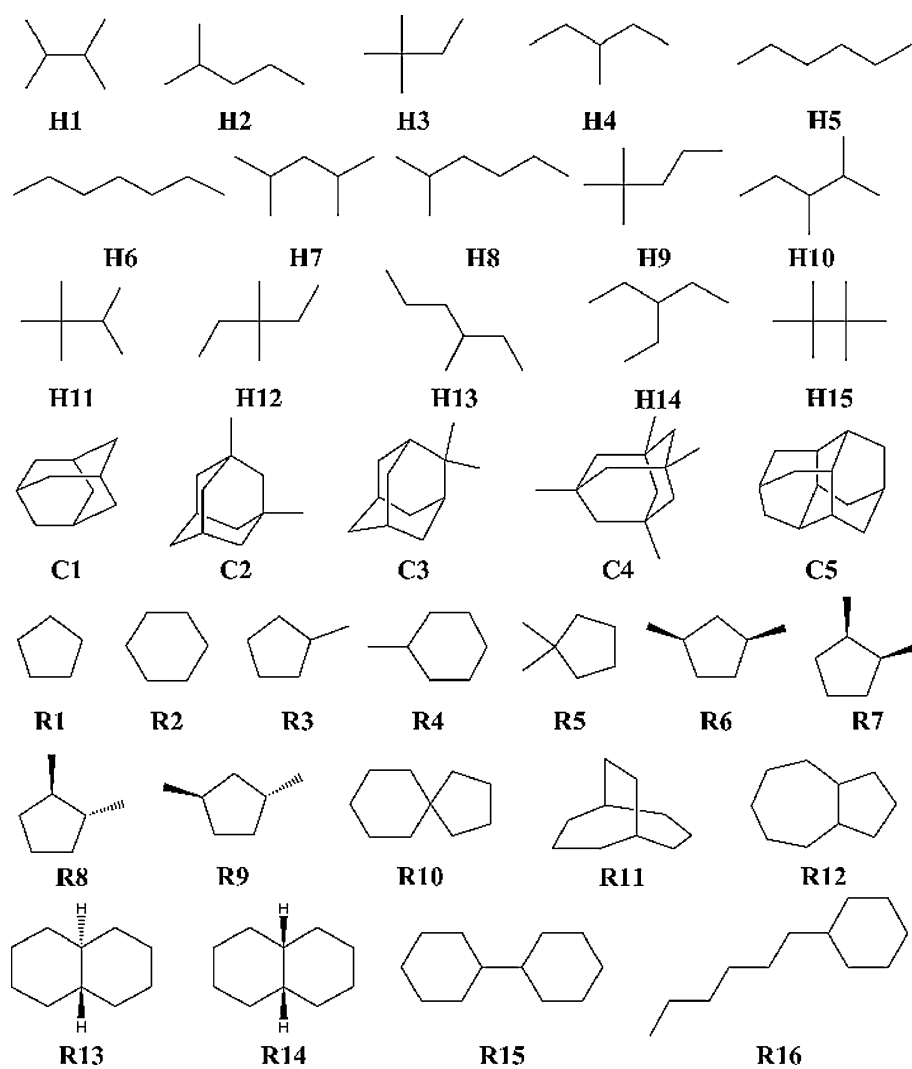
### 2.2 Functionals

The energy data for HF, PBE [76], PBE0 [77, 78], BP86 [79–81], BLYP [79, 82], B3LYP [46, 83], B2LYP, and their dDXDM [37] corrected versions as well as M06-2X [67], B2PLYP-D [32, 69], and B97-D [68] are also taken from ref [37]. Data for SVWN5 [84, 85] are taken from ref [35].

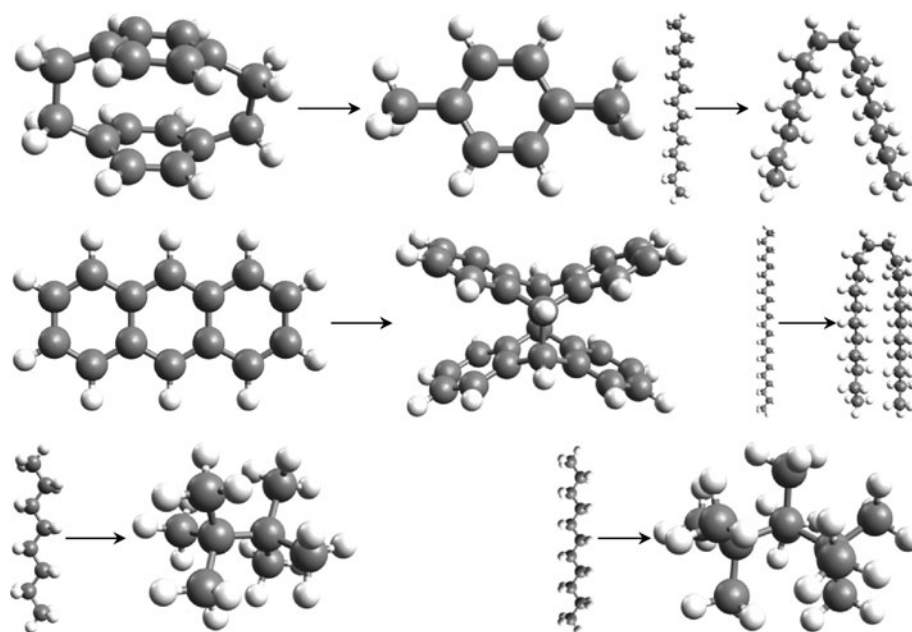
LC-BLYP [26, 63], LC-PBE [26], LC- $\omega$ PBE [64, 65], LC- $\omega$ PBEh [65], HFLYP, HFPBE, S [84], B88 [79], PBEx [76], rPW86 [80, 86], VV09 [61, 62], and vdW-DF04 [60] energies are computed with a developmental version of Q-Chem 3.2 [87]. HSE06 [88, 89] computations were performed in Gaussian 09 [90], while CAMB3LYP [91], rCAMB3LYP [66], LC-BLYP(0.33), MCY2 [92], and MCY3 [66] were computed with a version of CADPAC 6.5 [93, 94], kindly provided by Aron Cohen. The cc-pVTZ [75] basis set was used for all test sets, but interaction energies for the methane dimer were computed with the aug-cc-pVTZ basis set.

The specificities of the “modern” functionals listed above are briefly summarized below. M06-2X [67] is a fitted hybrid meta-GGA functional (about 30 parameters), designed to describe main group elements and weak interactions accurately. B97-D [68] is a GGA fitted

**Fig. 2** Schematic representation of the 36 saturated hydrocarbons in the chain (H1–H15), cages (C1–C5), and rings (R1–R16) set



**Fig. 3** The six reactions of the IDHC test set



together with the dispersion correction in order to minimize the double-counting of DFT correlation and the empirical dispersion correction. The double hybrid B2PLYP-D [32, 69] contains 27% MBPT2 correlation energy, 53% “exact” exchange, and a *posteriori* dispersion correction. vdW-DF04 [60] and VV09 [61, 62] are two fully non-local vdW-density functionals that are supplemented by an exchange and a local correlation functional. There is some freedom in the choice of the exchange component, but functionals that bind van der Waals complexes are obviously unsuitable (revPBE [95] and recently PW86 refits [80, 86] are popular options). The PW92 [96] parameterization is usually chosen for the local correlation.

LC-BLYP [26, 63], LC-PBE [26], LC- $\omega$ PBE [64, 65], and LC- $\omega$ PBEh [65] are long-range corrected functionals (labeled LC or LCR): the long-range is described by “exact” exchange and the short-range by DFT exchange (Eq. 2). In the range-separation scheme, pioneered by Savin et al. for the local density approximation [97, 98] and extended to general functionals by Tsuneda and coworkers [99], the electron repulsion operator  $\frac{1}{r_{12}}$  is partitioned into two ranges (long and short) with the most common choice being an Ewald-style partition based on the error function:

$$\frac{1}{r_{12}} = \underbrace{\frac{\operatorname{erfc}(\mu r_{12})}{r_{12}}}_{\text{SR}} + \underbrace{\frac{\operatorname{erf}(\mu r_{12})}{r_{12}}}_{\text{LR}} \quad (2)$$

where the  $\mu$  parameter is selected empirically and controls the definition of the two ranges (for other forms of Eq. 2 see ref [97, 100–103]). The LC scheme is motivated by the incorrectly decaying potential of standard DFT functionals (the  $x_c$  potential of semi-local functionals decays exponentially along with the density, while the asymptotic form of the exact potential is  $-1/r$ ). Applying the range separation and introducing HF exchange for the long-range corrects this error. The ‘long-range’ is considered especially important in the asymptotic region (i.e., surface) of atoms and molecules. Given a GGA or hybrid functional, the corresponding LRC functional is

$$E_{xc}^{\text{LRC}} = E_c + (1 - C_{\text{HF}})E_{x,\text{GGA}}^{\text{SR}} + C_{\text{HF}}E_{x,\text{HF}}^{\text{SR}} + E_{x,\text{HF}}^{\text{LR}} \quad (3)$$

The components labeled “LR” and “SR” are evaluated using the long- and short-range Coulomb potential, respectively, while  $C_{\text{HF}}$  denotes the coefficient of the HF exchange present in the original functional ( $E_{x,\text{HF}}$ ). Hybrid LRC functionals therefore contain some fraction of short-range HF exchange, but all LRC (CAMB3LYP excluded) functionals contain full HF exchange in the long-range limit (Eq. 3). The construction of the short-range exchange functional ( $E_{x,\text{GGA}}^{\text{SR}}$ ) requires an expression for the exchange-hole not given by conventional semi-local GGAs. The LC schemes thus mostly vary by the construction of the

short-range functional. An illustrative example is PBE, for which four short-range parameterizations exist in the literature ranging from using a pseudo-LDA exchange hole [99] (LC-PBE), applying the range separation to the enhancement factor (sr-PBE or  $\mu$ -PBE) [104], taking the model PBE-exchange-hole [105] (LC- $\omega$ PBE) [106] or using a more general exchange-hole [64] parameterized to reproduce PBE-results (called LC- $\omega$ PBE as well) [107]. HSE06 [88, 89], which has been motivated mainly for use in solid-state computations, is a screened hybrid that is the inverse of a long-range corrected functional; the short-range is described by “exact” exchange and the long-range by DFT exchange, this avoids the computationally expensive full-range exact exchange. CAMB3LYP [91] uses a different partitioning than Eq. 3, but can be seen as an extension of LC-BLYP (similar to the B3LYP extension to BLYP) fitted to atomization energies, ionization potentials, and total atomic energies, by varying the fraction of global and long-range exact exchange. rCAMB3LYP [66] is a re-parameterization of CAMB3LYP containing about twice the amount of the long-range correction and aims to improve the fractional charge behavior of a carbon atom. MCY2 [92] was constructed to be one-electron self-interaction free and to give good thermochemistry and reaction barriers, while MCY3 [66] uses long-range corrected components and has been, akin to rCAMB3LYP, fitted to improve the fractional charge behavior. Finally, LC-S-vdW-DF04 and LC-S-VV09 [61] pair the long-range corrected Slater exchange with the fully non-local vdW-density functionals with (i.e., VV09 [61]) or without (i.e., vdW-DF04 [60]) refitting the long-range separation parameter.

### 2.3 Dispersion correction

Very recently we introduced a dispersion correction [37] that combines Becke and Johnson’s exchange-hole dipole moment (XDM) formalism [108–114] with an extended Tang and Toennies damping function [115], in which the damping parameter depends on atomic (overlap) populations. As results obtained with the DFT-dDXDM are presented and analyzed hereafter, the most important ingredients of the correction are briefly reviewed below. The general form of dDXDM is the Tang and Toennies (TT) damping function [115]

$$E_{\text{disp}} = - \sum_{i=2}^{N_{\text{at}}} \sum_{j=1}^{i-1} f_{2n}(bR_{ij}) \frac{C_{2n}^{ij}}{R_{ij}^{2n}} \quad (4)$$

where  $N_{\text{at}}$  is the number of atoms in the system and  $b$  the TT-damping factor.  $f_{2n}(bR_{ij})$  are the “universal damping functions” [115], which are specific to each dispersion coefficient and that serve to attenuate the correction at short internuclear distances.



$$f_{2n}(x) = 1 - \exp(-x) \sum_{k=0}^{2n} \frac{x^k}{k!} \quad (5)$$

Dispersion coefficients are computed according to Becke and Johnson's XDM formalism [108–114], as efficiently implemented in Q-Chem by Kong and coworkers [116, 117]. The central quantity is the dipole moment arising from the reference electron and its exchange-hole (XDM). The XDM is evaluated using the Becke–Roussel model exchange hole [118]. The dispersion coefficients are partitioned among the atoms using the iterative Hirshfeld scheme (Hirshfeld-I) [119], to properly account for polarization (and overall charge). A key component of the dDXDM correction is the damping factor  $b$ , which is modeled by

$$b(x) = F(x) \cdot b_{ij,asym} \quad (6)$$

$F(x)$  and  $x$  are, respectively, the damping function and argument for  $b_{ij,asym}$ , that is the TT-damping factor associated with well-separated atoms  $i$  and  $j$ .  $b_{ij,asym}$  is obtained from  $b_{ii,asym}$  and  $b_{jj,asym}$  according to the recommended combination rule [120, 121].  $b_{ii,asym}$  is estimated [122, 123] by the square root of the atomic ionization energy  $\sqrt{I_i}$  taken from ref [124]. The atom in a molecule character is described by a cubic root scaling of the ratio between the free atom and the atomic volume in the molecule [125, 126]. The adjustable  $b_0$  parameter is introduced to scale the strength of correction in the medium range:

$$b_{ii,asym} = b_0 \cdot \sqrt{2I_i} \cdot \sqrt[3]{\frac{V_{i,free}}{V_{i,AIM}}} \quad (7)$$

The form of the  $b_{ij,asym}$  damping function is chosen as:

$$F(x) = 1 - \frac{2\arctan(a_0 \cdot x)}{\pi} \quad (8)$$

with the fitted  $a_0$  parameter adjusting the short-range behavior of the correction. Finally, the damping argument  $x$  is given by

$$x = \text{abs} \left( q_{ij} + q_{ji} - \frac{(Z_i - N_i) \cdot (Z_j - N_j)}{r_{ij}} \right) \frac{N_i + N_j}{N_i \cdot N_j} \quad (9)$$

where  $Z_i$  and  $N_i$  are the nuclear charge and Hirshfeld population of atom  $i$   $N_i = \int w_i(r)\rho(r)dr$ , respectively. The overlap population [127]  $q_{ij} = \int w_i(r)w_j(r)\rho(r)dr$  is a covalent bond index, whereas the fractional term in the parenthesis is an ionic bond index [128].

### 3 Results and discussions

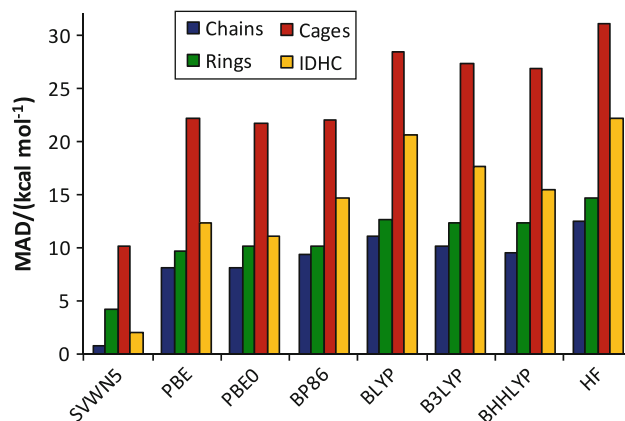
#### 3.1 General performance

Simple (closed shell molecules) hydrocarbon reactions such as People's bond separation equations (BSEs) of linear

alkanes (chains), cycloalkanes (rings), or cages (such as adamantane) show highly characteristic and systematic DFT errors [13] (see Fig. 4) of the same magnitude as HF (except for SVWN5). SVWN5 outperforms all the others common density functionals for alkane reaction energies (vide infra).

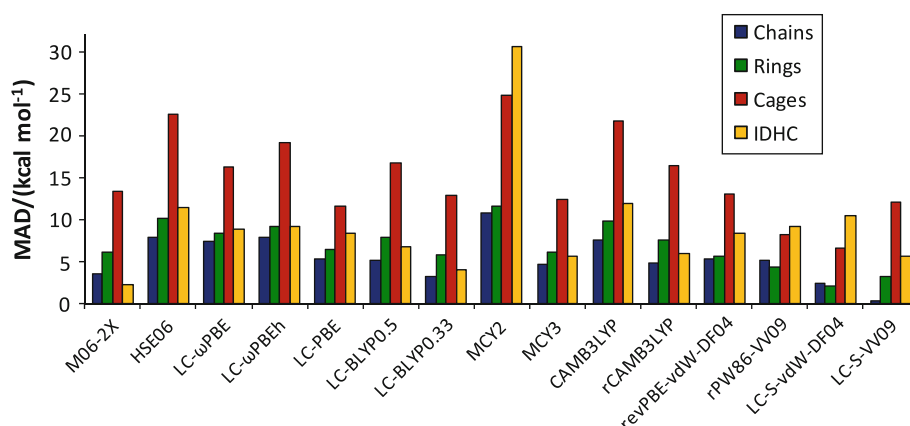
With the exception of SVWN5 (LDA), the contrast between “standard” (Fig. 4) and “modern” density functionals (Fig. 5) is striking, with the average mean absolute deviation (MAD) over the four test sets being 40% lower for the latter (MAD(standard) = 13.3 kcal mol<sup>-1</sup>, MAD(modern) = 7.7 kcal mol<sup>-1</sup>). Two general tendencies emerge from the comparisons between Figs. 4 and 5: (1) the inclusion of long-range corrected exchange energy terms improves the general performance, while the incorporation of a fraction of global exchange does not and (2) the accurate treatment of weak interactions is essential. The most illustrative examples are the superior performance of LC-BLYP0.33 (MAD = 5.5 kcal mol<sup>-1</sup>) and of the VV09-based functionals, rPW86-VV09 and LC-S-VV09 (MAD of 5.8 and 3.6 kcal mol<sup>-1</sup>, respectively). On the other hand, the remarkable performance of the hybrid meta-GGA, M06-2X (MAD = 5.6 kcal mol<sup>-1</sup>) highlights the valuable success of semi-empirical fitting for improving the performance of conventional DFT.

Inconveniently, the approaches that are best for general thermochemistry (i.e., atomization energies and barrier heights) tend to perform worst for the weak intramolecular interactions discussed herein, and vice versa. For instance, the inclusion of global HF exchange in conjunction with the long-range correction (i.e., (r)CAMB3LYP and LC- $\omega$ PBEh [65]) increases the MADs for each of our individual sets when compared to the “pure” long-range corrected functional (e.g., MAD = 9.9 and 9.0 kcal mol<sup>-1</sup> for LC- $\omega$ PBEh [65] and LC- $\omega$ PBE [64, 107], respectively).



**Fig. 4** Performance for standard functionals: Mean absolute deviations for bond separation energies over hydrocarbon chains, rings and cages, and for reaction energies of the test set “intramolecular dispersion interactions” (IDHC) using the cc-pVTZ basis set

**Fig. 5** Performance for “modern” functionals: Mean absolute deviations for bond separation energies over hydrocarbon chains, rings and cages, and for reaction energies of the test set “intramolecular dispersion interactions” (IDHC) using the cc-pVTZ basis set



These results contrast with other thermochemical properties for which the additional empirical parameter associated with the fraction of exact exchange improves the functional performance [65, 91]. Similarly, the LC-BLYP long-range separation parameter optimized for atomization energies (0.5), [26, 129] is significantly less accurate ( $MAD(0.5) = 7.8 \text{ kcal mol}^{-1}$ ) than LC-BLYP0.33 ( $MAD(0.33) = 5.5 \text{ kcal mol}^{-1}$ ) [63], which has been shown to give considerably lower delocalization errors [54, 66], but poor atomization energies [91]. MCY3 ( $MAD = 6.3 \text{ kcal mol}^{-1}$ ), which has been designed to minimize the delocalization error and benefits from the inclusion of long-range corrected exchange energy terms and from the same parameterization as rCAMB3LYP [66] also performs very well for our test sets, but less satisfactory for general thermochemistry [66]. This performance contrasts with MCY2 ( $MAD = 15.6 \text{ kcal mol}^{-1}$ , full-range exact exchange, one-electron self-interaction free), which has been parameterized against “general” thermochemistry and gives poor results for weak intramolecular interactions. Screened hybrid density functionals, which perform similar to PBE0 for thermochemistry, [130] also do not outperform the long-range corrected functionals in the present study cases (e.g.,  $MAD$  HSE06 = 11.0 kcal/mol). The sets of hydrocarbon reaction energies tested herein seemingly benefit from an earlier switching to the long-range interactions in exchange, but at this stage, no simple rationalization is possible as it is not unequivocally clear from where the improvement arises. In a recent study, Tsuneda and coworkers claimed that the lack of long-range interactions in exchange functionals is the major cause for the underestimation of alkane isodesmic reaction energies [26], but this on-going question [12, 26, 58, 59] will be thoroughly discussed in the next section.

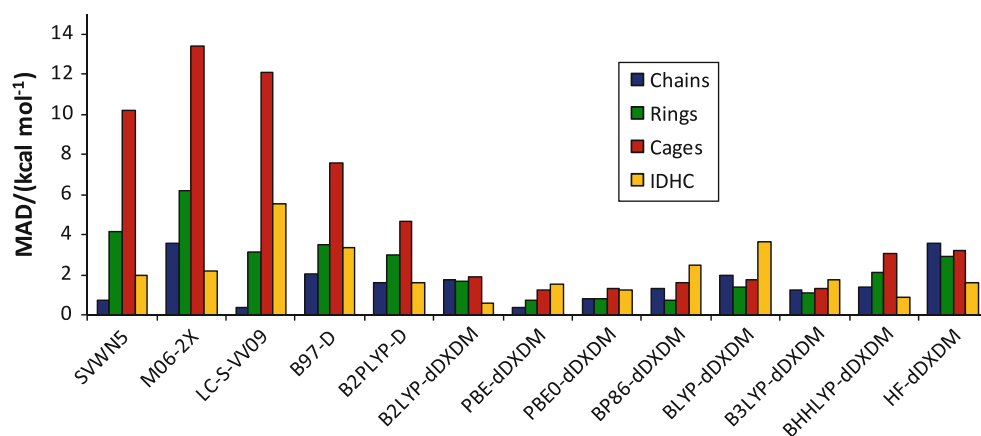
A much less debated origin for these DFT errors is the absence of non-local correlation effects to account for long-range intramolecular interactions such as dispersion [131]. This dependency is demonstrated by the good performance of the vdW-density functionals such as revPBE-vdW-DF04

and rPW86-VV09 [62] ( $MAD = 6.8$  and  $5.8 \text{ kcal mol}^{-1}$ , respectively). Combining VV09 with a long-range corrected functional, as suggested in ref [61], further stabilize the hydrocarbons ( $MAD = 3.6 \text{ kcal mol}^{-1}$ ). LC-S-vdW-DF04 performs similarly ( $MAD = 4.0 \text{ kcal mol}^{-1}$ ). The relevance of such combinations has also been demonstrated for rare-gas dimers and other systems [132–137].

An alternate and computationally cheaper solution to the intramolecular dispersion problem is to add a damped, atom pairwise energy correction to the standard KS-DFT energy as originally proposed by Yang et al. and others [11, 138–140]. Our recently proposed system-dependent dispersion correction (-dDXDM), based on Becke and Johnson’s exchange-hole dipole formalism, has shown to reduce both these “intramolecular errors” as well as the errors on typical (intermolecular) van der Waals complexes [37]. Figure 6 illustrates the performance of -dDXDM along with that of the popular B97-D and B2PLYP-D [32, 68]. Note that Grimme’s latest DFT-D3 correction [141] is not considered herein, but its performance on isodesmic reactions for linear alkane chains has been demonstrated in ref [59]. The success of dDXDM is due to a flexible, density-dependent, damping function that adapts well to a given functional (vide infra) together with accurate, density-dependent dispersion coefficients. To enable comparisons with Fig. 5, the best performing “standard” and “modern” functionals (SVWN5,  $MAD = 3.4 \text{ kcal mol}^{-1}$ , M06-2X,  $MAD = 5.6 \text{ kcal mol}^{-1}$  and LC-S-VV09,  $MAD = 3.6 \text{ kcal mol}^{-1}$ ) are included in Fig. 6.

The explicitly dispersion corrected functionals outperform the best functionals tested in Fig. 5, LC-BLYP0.33 and LC-S-VV09. In particular, PBE-dDXDM and PBE0-dDXDM ( $MAD = 0.8$  and  $0.9 \text{ kcal mol}^{-1}$ , respectively) give the best results with respect to experiment, while B2PLYP-D is the best for reproducing CCSD(T)/CBS reference values ( $MAD = 0.7 \text{ kcal mol}^{-1}$ , see Supporting Information). The dDXDM correction also lowers the error, which generally increases considerably going from chains, to rings, to cages. B2LYP-dDXDM (same

**Fig. 6** Performance for dispersion corrected functionals: Mean absolute deviations for bond separation energies over hydrocarbon chains, rings and cages, and for reaction energies of the test set “intramolecular dispersion interactions” (IDHC) using the cc-pVTZ basis set. SVWN5, M06-2X, and LC-S-VV09 are shown for comparison with Figs. 4 and 5



functional contributions as in B2PLYP, but without PT2 correlation energy) is not intended for real world applications, but provides insights into the good performance of B2PLYP-D. dDXDM alone is able to reproduce the combined role of the PT2 energy and the empirical dispersion correction for the present test sets.

The main outcome resulting from this preliminary investigation of the general performance of various functionals is (1) accounting for dispersion interactions is indispensable, (2) the improved treatment of “medium-range correlation” decreases the errors significantly (e.g., M06-2X), and (3) a long-range corrected exchange improves performance. Whereas (1) can be efficiently solved by the use of a dispersion correction that offers an attractive alternative to the computationally more expensive non-local vdW functionals, the importance of the exchange is uncertain [12, 26, 58, 59]. In the next section, we analyze the errors for these simple systems in greater detail to shed greater light on the origin of the DFT failures.

### 3.2 Detailed analysis of the functional performance

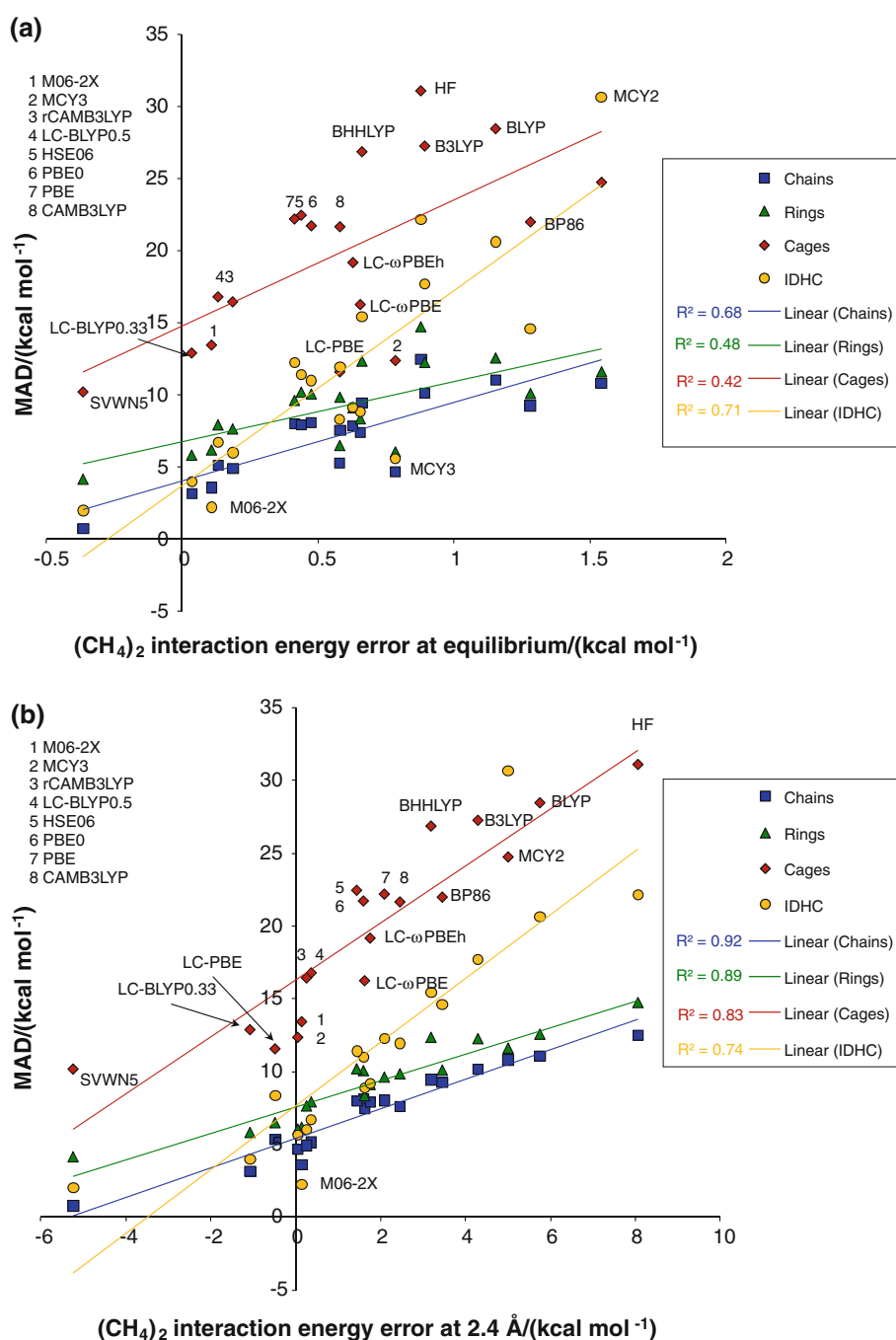
Our recent dispersion corrections, which are calibrated on alkane reaction energies [33, 35], aimed at accounting for “intramolecular errors” efficiently, without altering the long-range correction. Interestingly, our empirical variants of the correction [33, 35] show a minimum between 2.3 and 2.6 Å for the carbon–carbon interaction, a distance that corresponds roughly to the 1,3 C–C distance in propane (2.536 Å) [142]. This distance range is also similar to the compressed methane dimer as originally chosen by Yang and coworkers [55] as a model for probing delocalization errors [54, 55, 57] in Diels–Alder reaction energies. Comparisons between the errors in our test sets with those of the methane dimer interactions at the highly repulsive distance of 2.4 Å and at the equilibrium distance [143] is

instructive, as a correlation would be indicative of a common source of error.

The correlation of MADs is vastly superior with the error in the repulsive methane interaction region (Fig. 7b) than with the error at the equilibrium distance (Fig. 7a). The poorer correlation between the errors in our test sets and the methane interaction at the equilibrium distance is in line with the strongly varying results given by the density functionals for describing vdW-interactions [144–146]. The treatment of long-range dispersive interactions is missing, and the various performances strongly depend on the high-reduced density gradient  $s \propto \frac{|\nabla\rho|}{\rho^{4/3}}$  (low density, high gradient) behavior of the exchange functional. As demonstrated hereafter, the situation differs drastically in the repulsive range, which is somewhat more adapted to the description of branching in alkanes and compact hydrocarbons (e.g., anthracene dimer). M06-2X gives small errors for both the methane interaction energies and the IDHC test set. The errors are the largest for the BSE of the alkane cages, which result in a large y-axis intercept for this series. A reasonable interpretation for the non-zero intercept is the missing long-range correlation energy (dispersion) that can, by no means, be recovered by a local correlation functional (not even by extensive fitting like in M06-2X) and that increases with system size. Figures 8 and 9 give valuable insights on the improved performance of PBE0 when compared to its parent ingredients, HF and PBE, which both overestimate the repulsion of the compressed methane dimer. PBEx is considerably more repulsive than HF in the highly repulsive region modeled by the non-bonded compressed methane dimer (Fig. 8)! Whereas density functionals do not reproduce the repulsive wall of rare-gas dimers, most studies have focused on the slightly repulsive region, where exchange-only computations give various trends from too soft (e.g., PBEx) to too repulsive (e.g., B88) [86, 144, 146, 147]. One thus argues that in the highly repulsive range, the improved

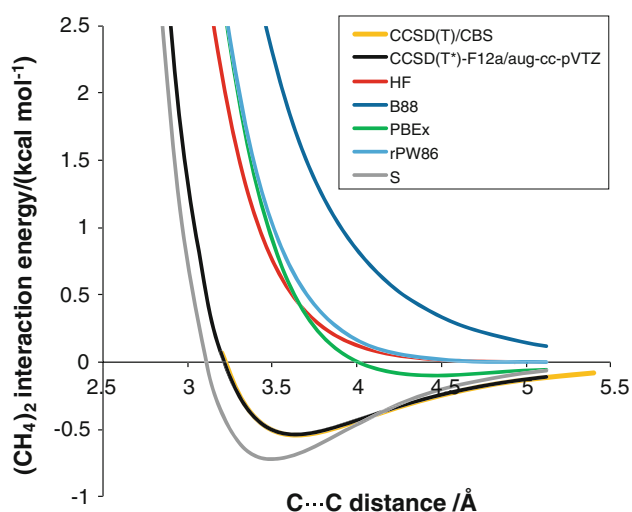


**Fig. 7** Correlations of “standard” and “modern” functionals between the methane interaction energy and the mean absolute deviations for, bond separation energies over hydrocarbon chains, rings and cages, and for reaction energies of the test set “intramolecular dispersion interactions” (IDHC)



performance of PBE0 compared to PBE is due to the smaller amount of overly repulsive PBE exchange (75% instead of 100%) in favor of the HF exchange. The improvement compared to HF is obviously due to the correlation functional. This interpretation corroborates with the more accurate interactions given by HF exchange together with PBE correlation, HFPBE, used here as a probe in line with Brittain et al. [57]. This reasoning in terms of over repulsive DFT exchange is, however, only valid for non-bonded interactions and not for semi-local DFT exchange treatments of covalent bonds [114]. HF

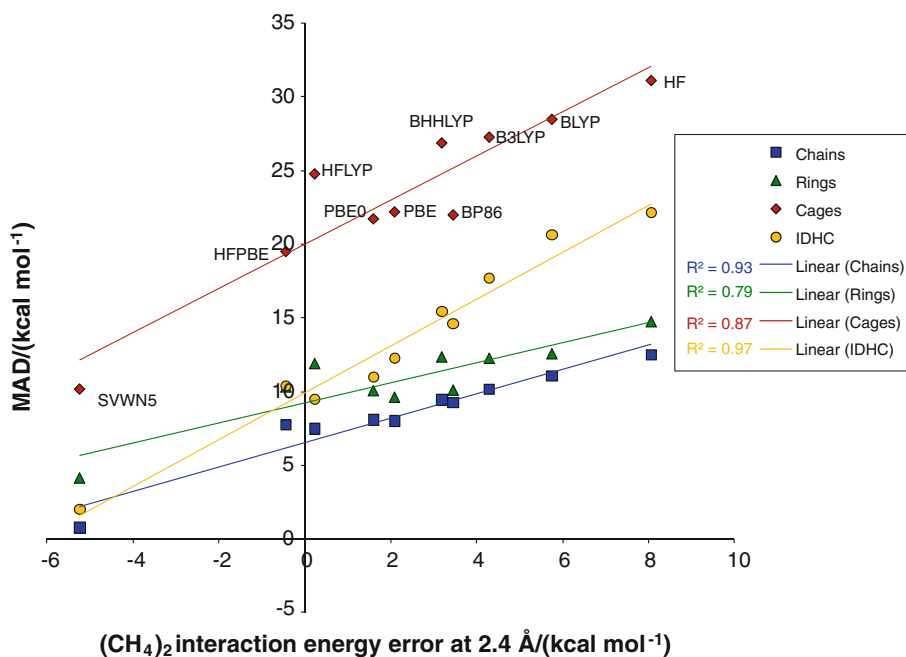
supplemented with semi-local correlation is, of course, not recommended for general purposes. The present interpretation of the overly repulsive nature of standard DFT exchange functionals at compressed distances are also responsible, potentially, for the “surprising and somewhat alarming” larger DFT errors recently noticed by Hobza et al. [148] for non-covalent energy computations of the compressed S22 geometries when compared to those at equilibrium (0.9 shift when compared to the equilibrium distance). At the equilibrium, PBE exchange is, on the contrary, under repulsive. Thus, the combination of HF



**Fig. 8** Methane dimer interaction profiles for exchange-only computations. CCSD(T\*)-F12a/aug-cc-pVTZ serves as a reference for the true interaction energy. The corresponding CCSD(T)/CBS benchmark values are also shown for the region considered in ref [149]

exchange with classic semi-local DFT correlation worsens the description of vdW-complexes [9] (e.g., HFPBE error at the equilibrium is in between HF and PBE). To a lesser extent, similar trends are observed for the BSE of the alkane cages and the IDHC test set (i.e., highly branched and compact system), whereas the alkane chains and rings (less branching, not as compact) remain weakly affected by the incorporation of global exchange. The improved performance of the long-range corrected exchange functionals for all our test sets and the compressed methane dimer illustrate, somewhat, that the description of non-bonded

**Fig. 9** Correlations of “standard” semi-local and hybrid density functionals between the methane interaction energy and the mean absolute deviations for, bond separation energies over hydrocarbon chains, rings and cages, and for reaction energies of the test set “intramolecular dispersion interactions” (IDHC)



interactions in the increasingly repulsive range benefits more from having full HF exchange at the surface of the molecule (in the long-range) than from containing some fraction of HF exchange at all ranges or full HF exchange at the short-range (e.g., screened hybrid functionals). As demonstrated by Yang and coworkers, the overly repulsive nature of the functional at compressed distances decreases considerably if the delocalization error is reduced (e.g., MCY3).

Apart from the exchange functional, correlation and especially dispersion play a major role in these errors, as confirmed by the dramatic improvement obtained when accounting for dispersion interactions [26, 32–37]. We here distinguish long-range dispersion from “overlap” dispersion (that is shorter-range dispersion in the region of overlapping density partially accounted for by semi-local functionals), both of which accumulate with system size. The notion of “overlap” dispersion is in line with a recent DFT study of the description of water hexamer interactions that recover a large part of the dispersion energy, dominated by the short-range [150]. It is also closely related to Grimme’s “overlap dispersive” interactions [151]. We, however, regard “overlap” dispersion as a particular case of medium-range correlation ([12, 38]) and not as an alternate terminology [141]. The importance of dispersion interactions is directly illustrated by the Hartree–Fock error for the compressed methane dimer ( $8.1 \text{ kcal mol}^{-1}$ ). The individual contributions to the energy interaction of the methane dimer are analyzed in Table 1 by symmetry-adapted perturbation theory (SAPT0) [152]. The difference between HF and SAPT first-order interaction energies plus second-order induction (-exchange) is given by  $\delta\text{HF}$

**Table 1** Interaction energy contributions for the methane dimer at equilibrium (C...C distance = 3.7 Å) and a repulsive distance (C...C distance = 2.4 Å), computed with SAPT0 using the aug-cc-pVTZ basis set

	Repulsive methane dimer	Equilibrium distance
$E_{\text{pol}}^{(1)}$	-14.11	-0.14
$E_{\text{exch}}^{(1)}$	45.55	0.53
$E_{\text{ind}}^{(2)}$	-11.17	-0.06
$E_{\text{ind-exch}}^{(2)}$	9.82	0.06
$E_{\text{disp}}^{(2)}$	-13.14	-0.98
$E_{\text{disp-exch}}^{(2)}$	2.99	0.07
$E_{\text{el}} = E_{\text{pol}}^{(1)} + E_{\text{exch}}^{(1)}$	31.44	0.39
$E_{\text{ind}}^{(2)} + E_{\text{ind-exch}}^{(2)}$	-1.35	0.00
$E_{\text{disp}}^{(2)} + E_{\text{disp-exch}}^{(2)}$	-10.15	-0.91
$E_{\text{tot}}$	19.94	-0.52
HF (BSSE corrected)	29.21	0.37
$\delta\text{HF}$	-0.88	-0.02
CCSD(T*)-F12a	21.11	-0.53

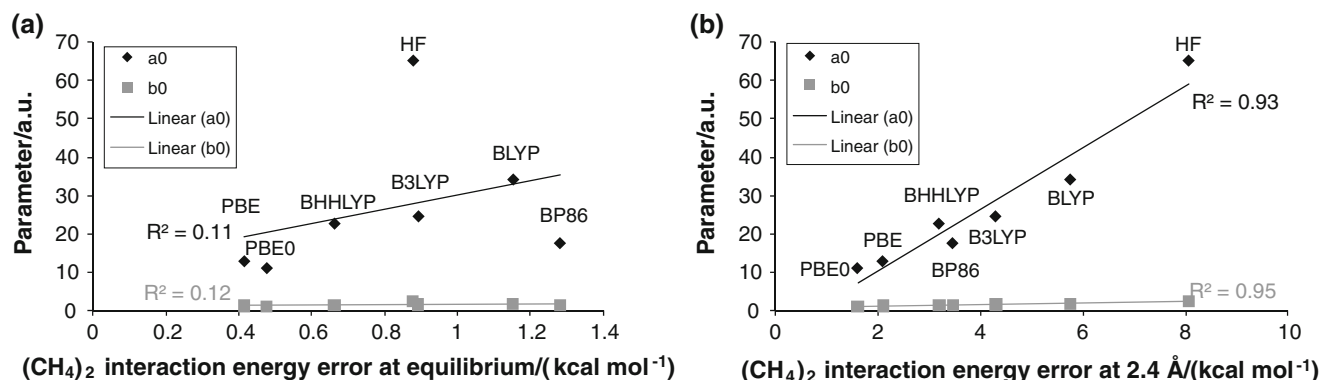
CCSD(T\*)-F12a/aug-cc-pVTZ is given as a reference. All values in kcal mol<sup>-1</sup>

[153, 154]. The HF error exactly matches the second-order dispersion energy. However, note that SAPT does not distinguish between “long-range” and “overlap” (i.e., shorter-range) dispersion that are, to a certain extent, captured by the DFT correlation functionals. At the highly repulsive distance, the correlation functional accounts for some “overlap” dispersion (due to overlapping density, see HFPBE and HFLYP large improvement when compared to HF in Fig. 9), but there are severe shortcomings (over repulsive) in the exchange functional. In line with the compressed methane dimer, the IDHC test set and the BSEs of branched alkanes (cages > rings > chains) follow

the same interpretation regarding the origin of the DFT errors. The error at the equilibrium distance is, however, interpreted in terms of the missing dispersion that can be compensated by an under repulsive DFT exchange (e.g., PBE). The good overall performance of SVWN5 for alkane BSEs is easily explained by error compensation between the short-range (overly attractive) and the missing long-range weak interactions. It has also been explained in terms of the surface energies [155] (LDA inherently favors compact over extended systems, and thereby correctly describes alkane branching).

Given that the dDXDM correction efficiently reduces the MADs of our test sets, it is reasonable to expect a correlation between the parameters adjusting the correction for a given functional and the error in the methane dimer interaction energies for the parent functional as illustrated by Fig. 10. As discussed in the Computational Details, the dDXDM correction contains only two fitted parameters:  $a_0$ , which controls how fast the correction decays to zero for short interatomic distances and  $b_0$ , which determines the strength of the correction at intermediate interatomic distances (i.e., in the medium range).

The physical interpretation of the good correlation for  $b_0$  in Fig. 10b is straightforward: the higher the  $b_0$  value, the stronger the dispersion correction at intermediate distances. At long ranges, the damping vanishes exponentially and  $b_0$  becomes unimportant. At short ranges, the second damping function, with the steepness  $a_0$ , dominates:  $a_0 = 0$  turns off the additional damping and restores the “conventional” Tang and Toennies damping, whereas  $a_0 = \infty$  turns the entire correction off. Intermediary, the additional damping responds to  $b_0$ . A high  $b_0$  value means large corrections at short (even covalent) distances that are prevented by increasing  $a_0$ . For a given density overlap, a high  $a_0$  value turns down the TT-dispersion correction more than a low  $a_0$ . This behavior justifies the poorer performance of dDXDM when combined with highly repulsive functionals such as BLYP (or HF) when compared to softer



**Fig. 10** Correlation between best fit  $a_0$  and  $b_0$  values for dDXDM against the absolute interaction error for the uncorrected functional of the methane dimer at equilibrium (a) and a compressed distance (b)

functionals, such as PBE or even B3LYP. The high  $b_0$  values required for correcting the medium range increases  $a_0$  as well, thereby diminishing the correction at distances with considerable density overlap whether the overlap is bonding or not. dDXDM is, nevertheless, a powerful correction, as even HF and BLYP are corrected to an extent that they compete with M06-2X. The fact that dDXDM benefits from the “overlap dispersion” intrinsic to the DFT functional advises against combining the correction with the dispersionless functional of Pernal et al. [156]

#### 4 Conclusions

The benchmarking of “modern” density functionals for overcoming common DFT errors in hydrocarbon reaction energies provides valuable insight into the origin of the errors given by traditional semi-local and hybrid density functionals. The most illustrative example for these DFT shortcomings is the large B3LYP energy underestimation ( $MAD = 14.1 \text{ kcal mol}^{-1}$ ) over four test sets of hydrocarbons featuring weak intramolecular interactions. Our comprehensive analysis demonstrates that for increasingly branched alkanes and most compact hydrocarbons, the “intramolecular errors” strongly correlate with the error for the compressed methane dimer interaction energy. At these compressed distances, the shortcomings can be partially attributed to the general overly repulsive nature of DFT exchange in the treatment of non-bonded density overlaps. The significant improvement offered by the long-range corrected exchange functionals stems from the substitution of the long-range DFT exchange by a less repulsive “exact” exchange (e.g., LC-BLYP0.33,  $MAD = 5.5 \text{ kcal mol}^{-1}$ ). The overly repulsive nature of the DFT exchange at compressed distances sharply contrasts with the various trends (e.g., PBEx under repulsive to B88 over repulsive) characteristic of the near equilibrium region. Our study also emphasizes the essential role played by the correlation functional for lowering the error of these hydrocarbon reactions. At regions of non-bonded density overlap, the correlation functionals account for a non-negligible extent of “overlap dispersion” (Fig. 9). Improving the treatment of long-range dispersive interactions leads to enhanced performances, as illustrated by both the impressive results of the non-local van der Waals density functionals (e.g., rPW86-VV09,  $MAD = 5.8$  and LC-S-VV09,  $MAD = 3.6 \text{ kcal mol}^{-1}$ ) and the atom pairwise density-dependent corrected PBE (e.g., PBE-dDXDM,  $MAD = 0.8 \text{ kcal mol}^{-1}$ ). The overall repulsive nature of the functionals for treating these weak intramolecular interactions also decreases considerably if the delocalization error is reduced (e.g., MCY3,  $MAD = 6.3 \text{ kcal mol}^{-1}$ ) or by developing improved and flexible hybrid

meta functional forms coupled with careful parameter fitting (e.g., M06-2X,  $MAD = 5.6 \text{ kcal mol}^{-1}$ ).

**Acknowledgments** C.C. acknowledges the Sandoz family foundation, the Swiss NSF Grant 200021\_121577/1, and EPFL for financial support. We are grateful to Aron Cohen for providing the binary version of CADPAC 6.5.

#### References

1. Kohn W, Sham LJ (1965) *Phys Rev* 140(4):A1133–A1138
2. Jones RO, Gunnarsson O (1989) *Rev Mod Phys* 61:689–746
3. Kristyan S, Pulay P (1994) *Chem Phys Lett* 229(3):175–180
4. Perez-Jorda JM, Becke AD (1995) *Chem Phys Lett* 233(1–2):134–137
5. Zhang YK, Pan W, Yang WT (1997) *J Chem Phys* 107:7921–7925
6. Hobza P, Sponer J, Reschel T (1995) *J Comput Chem* 16(11):1315–1325
7. Patton DC, Pederson MR (1997) *Phys Rev A* 56(4):R2495–R2498
8. Milet A, Korona T, Moszynski R, Kochanski E (1999) *J Chem Phys* 111(17):7727–7735
9. Perez-Jorda JM, San-Fabian E, Perez-Jimenez AJ (1999) *J Chem Phys* 110(4):1916–1920
10. Tsuzuki S, Luthi HP (2001) *J Chem Phys* 114(9):3949–3957
11. Wu X, Vargas MC, Nayak S, Lotrich V, Scoles G (2001) *J Chem Phys* 115(19):8748–8757
12. Grimme S (2006) *Angew Chem Int Ed* 45(27):4460–4464
13. Wodrich MD, Corminboeuf C, Schleyer PvR (2006) *Org Lett* 8(17):3631–3634
14. Schreiner PR (2007) *Angew Chem Int Ed* 46(23):4217–4219
15. Gonthier JF, Wodrich MD, Steinmann SN, Corminboeuf C (2010) *Org Lett* 12(13):3070–3073
16. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) *J Chem Phys* 106(3):1063–1079
17. Raghavachari K, Stefanov BB, Curtiss LA (1997) *Mol Phys* 91(3):555–559
18. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (2000) *J Chem Phys* 112(17):7374–7383
19. Redfern PC, Zapol P, Curtiss LA, Raghavachari K (2000) *J Phys Chem A* 104(24):5850–5854
20. Saeys M, Reyniers MF, Marin GB, Van Speybroeck V, Waroquier M (2003) *J Phys Chem A* 107(43):9147–9159
21. Hehre WJ, Ditchfield R, Radom L, Pople JA (1970) *J Am Chem Soc* 92(16):4796–4801
22. Pople JA, Radom L, Hehre WJ (1971) *J Am Chem Soc* 93(2):289–300
23. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) *Ab initio molecular orbital theory*. Wiley, New York
24. Wodrich MD, Wannere CS, Mo Y, Jarowski PD, Houk KN, Schleyer PvR (2007) *Chem Eur J* 13(27):7731–7744
25. Karton A, Gruzman D, Martin JML (2009) *J Phys Chem A* 113:8434–8447
26. Song J-W, Tsuneda T, Sato T, Hirao K (2010) *Org Lett* 12(7):1440–1443
27. Allen TL (1958) *J Chem Phys* 29(4):951–952
28. Allen TL (1959) *J Chem Phys* 31(4):1039–1049
29. Pitzer KS, Catalano E (1956) *J Am Chem Soc* 78(19):4844–4846
30. Pitzer KS (1959) *Adv Chem Phys* 2:59–83
31. London F (1937) *Trans Faraday Soc* 33:8b–26

32. Schwabe T, Grimme S (2007) *Phys Chem Chem Phys* 9(26):3397–3406
33. Wodrich MD, Jana DF, Schleyer PvR, Corminboeuf C (2008) *J Phys Chem A* 112(45):11495–11500
34. Karton A, Tarnopolsky A, Lamere J-F, Schatz GC, Martin JML (2008) *J Phys Chem A* 112(50):12868–12886
35. Steinmann SN, Csonka G, Corminboeuf C (2009) *J Chem Theory Comput* 5(11):2950–2958
36. Shamov GA, Budzelaar PHM, Schreckenbach G (2010) *J Chem Theory Comput* 6(2):477–490
37. Steinmann SN, Corminboeuf C (2010) *J Chem Theory Comput* 6(7):1990–2001
38. Zhao Y, Truhlar DG (2006) *Org Lett* 8(25):5753–5755
39. Feng Y, Liu L, Wang J-T, Huang H, Guo Q-X (2003) *J Chem Inf Comput Sci* 43:2005
40. Yao F, Xiao-Yu D, Yi-Min W, Lei L, Qing-Xiang G (2005) *Chin J Chem* 23(5):474–482
41. Cote ML, Pross A, Radom L (2004) In: Brandas EJ, Kryachko ES (eds) *Fundamental world of quantum chemistry*, vol III. Kluwer, The Netherlands, p 563
42. Izgorodina EI, Cote ML, Radom L (2005) *J Phys Chem A* 109(33):7558–7566
43. Check CE, Gilbert TM (2005) *J Org Chem* 70(24):9828–9834
44. Zhao Y, Truhlar DG (2008) *J Phys Chem A* 112(6):1095–1099
45. Becke AD (1993) *J Chem Phys* 98(2):1372–1377
46. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) *J Phys Chem* 98(45):11623–11627
47. Check CE, Gilbert TM (2005) *J Org Chem* 70:9828–9834
48. Woodcock HL, Schaefer HF, Schreiner PR (2002) *J Phys Chem A* 106(49):11923–11931
49. Schreiner PR, Fokin AA, Pascal RA, de Meijere A (2006) *Org Lett* 8(17):3635–3638
50. Wodrich MD, Corminboeuf C, Schreiner PR, Fokin AA, Schleyer PvR (2007) *Org Lett* 9(10):1851–1854
51. Fermi E, Amaldi E (1934) *Accad Ital Rome* 6:119
52. Perdew JP (1979) *Chem Phys Lett* 64(1):127–130
53. Perdew JP, Zunger A (1981) *Phys Rev B* 23(10):5048–5079
54. Mori-Sanchez P, Cohen AJ, Yang W (2006) *J Chem Phys* 125(20):201102
55. Ruzsinszky A, Perdew JP, Csonka GI, Vydrov OA, Scuseria GE (2007) *J Chem Phys* 126(10):104102–104108
56. Cohen AJ, Mori-Sanchez P, Yang W (2008) *Science* 321(5890):792–794
57. Mori-Sanchez P, Cohen AJ, Yang W (2008) *Phys Rev Lett* 100(14):146401
58. Brittain DRB, Lin CY, Gilbert ATB, Izgorodina EI, Gill PMW, Cote ML (2009) *Phys Chem Chem Phys* 11(8):1138–1142
59. Grimme S (2010) *Org Lett*. doi:10.1021/ol1016417
60. Dion M, Rydberg H, Schröder E, Langreth DC, Lundqvist BI (2004) *Phys Rev Lett* 92(24):246401
61. Vydrov OA, Van Voorhis T (2009) *Phys Rev Lett* 103(6):063004
62. Vydrov OA, Voorhis TV (2010) *J Chem Phys* 132(16):164113
63. Tawada Y, Tsuneda T, Yanagisawa S, Yanai T, Hirao K (2004) *J Chem Phys* 120(18):8425–8433
64. Henderson TM, Janesko BG, Scuseria GE (2008) *J Chem Phys* 128(19):194105
65. Rohrdanz MA, Martins KM, Herbert JM (2009) *J Chem Phys* 130(5):054112
66. Cohen AJ, Mori-Sanchez P, Yang W (2007) *J Chem Phys* 126(19):191109
67. Zhao Y, Truhlar D (2008) *Theor Chem Acc* 120(1):215–241
68. Grimme S (2006) *J Comput Chem* 27(15):1787–1799
69. Grimme S (2006) *J Chem Phys* 124(3):034108
70. Afeefy HY, Liebman JF, Stein SE (2010) Neutral thermochemical data. In: Linstrom PJ, Mallard WG (eds) *Nist chemistry webbook, nist standard reference database number 69*. National Institute of Standards and Technology, Gaithersburg MD, 20899
71. Kreig H, Grimme S (2010) *Mol Phys*. doi:10.1080/00268976.2010.519729
72. Adler TB, Knizia G, Werner H-J (2007) *J Chem Phys* 127(22):221106
73. Marchetti O, Werner H-J (2009) *J Phys Chem A* 113(43):11580–11585
74. Werner H-J, Knowles PJ, Lindh R, Manby FR, Schütz M, Celani P, Korona T, Mitrushenkov A, Rauhut G, Adler TB, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F, Goll E, Hampel C, Hetzer G, Hrenar T, Knizia G, Köppl C, Liu Y, Lloyd AW, Mata RA, May AJ, McNicholas SJ, Meyer W, Mura ME, Nicklass A, Palmieri P, Pflüger K, Pitzer R, Reiher M, Schumann U, Stoll H, Stone AJ, Tarroni R, Thorsteinsson T, Wang M, Wolf A (2009) *Molpro*, version 2009.1, a package of ab initio programs
75. Thom H, Dunning J (1989) *J Chem Phys* 90(2):1007–1023
76. Perdew JP, Burke K, Ernzerhof M (1996) *Phys Rev Lett* 77(18):3865–3868
77. Perdew JP, Ernzerhof M, Burke K (1996) *J Chem Phys* 105(22):9982–9985
78. Adamo C, Barone V (1999) *J Chem Phys* 110(13):6158–6170
79. Becke AD (1988) *Phys Rev A* 38(6):3098
80. Perdew JP, Wang Y (1986) *Phys Rev B* 33(12):8800–8802
81. Perdew JP, Yue W (1989) *Phys Rev B* 40(5):3399
82. Lee C, Yang W, Parr RG (1988) *Phys Rev B* 37(2):785
83. Becke AD (1993) *J Chem Phys* 98(7):5648–5652
84. Slater JC (1951) *Phys Rev* 81(3):385
85. Vosko SH, Wilk L, Nusair M (1980) *Can J Phys* 58(8):1200–1211
86. Murray EaD, Lee K, Langreth DC (2009) *J Chem Theory Comput* 5(10):2754–2762
87. Shao Y, Molnar LF, Jung Y, Kussmann J, Ochsenfeld C, Brown ST, Gilbert ATB, Slipchenko LV, Levchenko SV, O'Neill DP, DiStasio Jr RA, Lochan RC, Wang T, Beran GJO, Besley NA, Herbert JM, Lin CY, Voorhis TV, Chien SH, Sodt A, Steele RP, Rassolov VA, Maslen PE, Korambath PP, Adamson RD, Austin B, Baker J, Byrd EFC, Dachsels H, Doerksen RJ, Dreuw A, Dunietz BD, Dutoi AD, Furlani TR, Gwaltney SR, Heyden A, Hirata S, Hsu C-P, Kedziora G, Khalliulin RZ, Klunzinger P, Lee AM, Lee MS, Liang W, Lotan I, Nair N, Peters B, Proynov EI, Pieniazek PA, Rhee YM, Ritchie J, Rosta E, Sherrill CD, Simmonett AC, Subotnik JE, Woodcock HL III, Zhang W, Bell AT, Chakraborty AK, Chipman DM, Keil FJ, Warshel A, Hehre WJ, Schaefer HF III, Kong J, Krylov AI, Gill PMW, Head-Gordon M (2006) *Phys Chem Chem Phys* 8(27):3172–3191
88. Heyd J, Scuseria GE, Ernzerhof M (2003) *J Chem Phys* 118(18):8207–8215
89. Heyd J, Scuseria GE, Ernzerhof M (2006) *J Chem Phys* 124(21):219906
90. Revision A.1, Gaussian, Inc., Wallingford CT, 2009
91. Yanai T, Tew DP, Handy NC (2004) *Chem Phys Lett* 393(1–3):51–57
92. Mori-Sanchez P, Cohen AJ, Yang W (2006) *J Chem Phys* 124(9):091102
93. CADPAC, The Cambridge Analytic Derivatives Package
94. f functions were omitted in CADPAC computations of the anthracene dimerization reaction and the folding of C<sub>22</sub>H<sub>46</sub> for technical reasons
95. Zhang Y, Yang W (1998) *Phys Rev Lett* 80(4):890
96. Perdew JP, Wang Y (1992) *Phys Rev B* 45(23):13244–13249
97. Savin A, Flad H-J (1995) *Int J Quantum Chem* 56(4):327–332
98. Savin A (1996) On degeneracy, near-degeneracy and density functional theory. In: Seminario JM (ed) *Theoretical and computational chemistry*, vol 4. Elsevier, Amsterdam, pp 327–357



99. Iikura H, Tsuneda T, Yanai T, Hirao K (2001) *J Chem Phys* 115(8):3540–3544
100. Toulouse J, Colonna F, Savin A (2004) *Phys Rev A* 70(6):062505
101. Henderson TM, Izmaylov AF, Scuseria GE, Savin A (2007) *J Chem Phys* 127(22):221103
102. Song J-W, Tokura S, Sato T, Watson MA, Hirao K (2007) *J Chem Phys* 127(15):154109
103. Song J-W, Watson MA, Hirao K (2009) *J Chem Phys* 131(14):144108
104. Toulouse J, Colonna F, Savin A (2005) *J Chem Phys* 122(1):014110
105. Ernzerhof M, Perdew JP (1998) *J Chem Phys* 109(9):3313–3320
106. Vydrov OA, Scuseria GE (2006) *J Chem Phys* 125(23):234109
107. Weintraub E, Henderson TM, Scuseria GE (2009) *J Chem Theory Comput* 5(4):754–762
108. Johnson ER, Becke AD (2005) *J Chem Phys* 123(2):024101
109. Becke AD, Johnson ER (2005) *J Chem Phys* 123(15):154101
110. Becke AD, Johnson ER (2005) *J Chem Phys* 122(15):154104
111. Becke AD, Johnson ER (2006) *J Chem Phys* 124(1):014104
112. Johnson ER, Becke AD (2006) *J Chem Phys* 124(17):174104
113. Becke AD, Johnson ER (2007) *J Chem Phys* 127(15):154108
114. Becke AD, Johnson ER (2007) *J Chem Phys* 127(12):124108
115. Tang KT, Toennies JP (1984) *J Chem Phys* 80(8):3726–3741
116. Proynov E, Gan Z, Kong J (2008) *Chem Phys Lett* 455(1–3):103–109
117. Kong J, Gan Z, Proynov E, Furlani TR (2009) *Phys Rev A* 79(4):042510
118. Becke AD, Roussel MR (1989) *Phys Rev A* 39(8):3761
119. Bultinck P, Alsenoy CV, Ayers PW, Carbo-Dorca R (2007) *J Chem Phys* 126(14):144111
120. Bohm H-J, Ahlrichs R (1982) *J Chem Phys* 77(4):2028–2034
121. Sheng XW, Li P, Tang KT (2009) *J Chem Phys* 130(17):174310
122. Douketis C, Scoles G, Marchetti S, Zen M, Thakkar AJ (1982) *J Chem Phys* 76(6):3057–3063
123. Tang KT, Toennies JP, Yiu CL (1995) *Phys Rev Lett* 74(9):1546
124. Martin WC, Musgrove A, Kotochigova S, Sansonetti JE (2003) Ground levels and ionization energies for the neutral atoms. In: Drake GWF (ed) *Physical reference data, nist standard reference database number 111*. National Institute of Standards and Technology, Gaithersburg MD, 20899
125. Tkatchenko A, Scheffler M (2009) *Phys Rev Lett* 102(7):073005
126. Tkatchenko A, Robert A, DiStasio J, Head-Gordon M, Scheffler M (2009) *J Chem Phys* 131(9):094106
127. Mayer I, Salvador P (2004) *Chem Phys Lett* 383(3–4):368–375
128. Mulliken RS (1955) *J Chem Phys* 23(10):1841–1846
129. Song J-W, Hirose T, Tsuneda T, Hirao K (2007) *J Chem Phys* 126(15):154105–154107
130. Vydrov OA, Heyd J, Krukau AV, Scuseria GE (2006) *J Chem Phys* 125(7):074106
131. Zhao Y, Lynch BJ, Truhlar DG (2004) *J Phys Chem A* 108(21):4786–4791
132. Kamiya M, Tsuneda T, Hirao K (2002) *J Chem Phys* 117(13):6010–6015
133. Sato T, Tsuneda T, Hirao K (2005) *Mol Phys* 103(6):1151–1164
134. Sato T, Tsuneda T, Hirao K (2005) *J Chem Phys* 123(10):104307
135. Angyan JG, Gerber IC, Savin A, Toulouse J (2005) *Phys Rev A* 72(1):012510
136. Gerber IC, Angyan JG (2005) *Chem Phys Lett* 416(4–6):370–375
137. Goll E, Werner HJ, Stoll H (2005) *Phys Chem Chem Phys* 7(23):3917–3923
138. Wu Q, Yang W (2002) *J Chem Phys* 116(2):515–524
139. Elstner M, Hobza P, Frauenheim T, Suhai S, Kaxiras E (2001) *J Chem Phys* 114(12):5149–5155
140. Meijer EJ, Sprik M (1996) *J Chem Phys* 105(19):8684–8689
141. Grimme S, Antony J, Ehrlich S, Krieg H (2010) *J Chem Phys* 132(15):154104
142. NIST Computational Chemistry Comparison and Benchmark Database, <http://srdata.nist.gov/cccbdb>
143. Jurecka P, Sponer J, Cerny J, Hobza P (2006) *Phys Chem Chem Phys* 8(17):1985–1993
144. Lacks DJ, Gordon RG (1993) *Phys Rev A* 47(6):4681
145. Wesolowski TA, Parisel O, Ellinger Y, Weber J (1997) *J Phys Chem A* 101(42):7818–7825
146. Zhang Y, Pan W, Yang W (1997) *J Chem Phys* 107(19):7921–7925
147. Kannemann FO, Becke AD (2009) *J Chem Theory Comput* 5(4):719–727
148. Grafova L, Pitonak M, Rezac J, Hobza P (2010) *J Chem Theory Comput*:null-null
149. Sherrill CD, Takatani T, Hohenstein EG (2009) *J Phys Chem A* 113(38):10146–10159
150. Wang F-F, Jenness G, Al-Saidi WA, Jordan KD (2010) *J Chem Phys* 132(13):134303
151. Grimme S (2004) *Chem Eur J* 10(14):3423–3429
152. Jeziorski B, Moszynski R, Szalewicz K (1994) *Chem Rev* 94(7):1887–1930
153. Jeziorska M, Bogumil J, Jiri C (1987) *Int J Quantum Chem* 32(2):149–164
154. Moszynski R, Heijmen TGA, Jeziorski B (1996) *Mol Phys* 88(3):741–758
155. Stroppa A, Kresse G (2008) *New J Phys* 10(6):063020
156. Pernal K, Podeszwa R, Patkowski K, Szalewicz K (2009) *Phys Rev Lett* 103(26):4