View Online / Journal Homepage / Table of Contents for this issue
Dynamic Article Links

PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670–6688

www.rsc.org/pccp

PAPER

A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions†

Lars Goerigk[ab] **and Stefan Grimme***[a]

Received 28th December 2010, Accepted 10th February 2011
DOI: 10.1039/c0cp02984j

A thorough energy benchmark study of various density functionals (DFs) is carried out with the new GMTKN30 database for general main group thermochemistry, kinetics and noncovalent interactions [Goerigk and Grimme, *J. Chem. Theor. Comput.*, 2010, **6**, 107; Goerigk and Grimme, *J. Chem. Theor. Comput.*, 2011, **7**, 291]. In total, 47 DFs are investigated: two LDAs, 14 GGAs, three meta-GGAs, 23 hybrids and five double-hybrids. Besides the double-hybrids, also other modern approaches, *i.e.*, the M05 and M06 classes of functionals and range-separated hybrids, are tested. For almost all functionals, the new DFT-D3 correction is applied in order to consistently test the performance also for important noncovalent interactions; the parameters are taken from previous works or determined for the present study. Basis set and quadrature grid issues are also considered. The general aim of the study is to work out which functionals are generally well applicable and robust to describe the energies of molecules. In summary, we recommend on the GGA level the B97-D3 and revPBE-D3 functionals. The best meta-GGA is oTPSS-D3 although meta-GGAs represent in general no clear improvement compared to numerically simpler GGAs. Notably, the widely used B3LYP functional performs worse than the average of all tested hybrids and is also very sensitive to the application of dispersion corrections. We discourage its usage as a standard method without closer inspection of the results, as it still seems to be often done nowadays. Surprisingly, long-range corrected exchange functionals do in general not perform better than the corresponding standard hybrids. However, the ωB97X-D functional seems to be a promising method. The most robust hybrid is Zhao and Truhlar's PW6B95 functional in combination with DFT-D3. If higher accuracy is required, double-hybrids should be applied. The corresponding DSD-BLYP-D3 and PWPB95-D3 variants are the most accurate and robust functionals of the entire study. Additional calculations with MP2 and and its spin-scaled variants SCS-MP2, S2-MP2 and SOS-MP2 revealed that double-hybrids in general outperform those. Only SCS-MP2 can be recommended, particularly for reaction energies. We suggest its usage when a large self-interaction error is expected that prohibits usage of double-hybrids. Perdews' metaphoric picture of Jacob's Ladder for the classification of density functionals' performance could unbiasedly be confirmed with GMTKN30. We also show that there is no statistical correlation between a functional's accuracy for atomization energies and the performance for chemically more relevant reaction energies.

1. Introduction

The aim of computational thermochemistry is to describe the energetic properties of chemical processes with an accuracy of 1 kcal mol^{-1} or less (0.1–0.2 kcal mol^{-1} for the relative energy of conformers). At the same time, the methods applied should not be too demanding in terms of necessary computation times and hardware resources, which rules highly accurate *ab initio* wavefunction based methods out if larger, chemically relevant systems are considered. The Kohn–Sham density functional theory [(KS-)DFT] offers an ideal solution to this dilemma.[1,2] However, the number of proposed exchange-correlation functionals is immense and most of them suffer from severe problems. Very prominent examples are the self-interaction-error[3–6] (SIE, also called delocalization-error in many-electron systems) and the lack of adequately describing long-range correlation effects, *e.g.* London-dispersion.[7–10] Moreover, the applicability of many functionals to various

[a] *Theoretische Organische Chemie, Organisch-Chemisches Institut der Universität Münster, Corrensstraße 40, D-48149 Münster, Germany. E-mail: grimmes@uni-muenster.de*
[b] *NRW Graduate School of Chemistry, Wilhelm-Klemm-Straße 10, D-48149 Münster, Germany*

† Electronic supplementary information (ESI) available: Parameters for the DFT-D3(BJ) correction (Table S1); best functionals for a certain benchmark set (Table S2); all results for GMTKN30 with the (aug-)def-QZVP bases (Tables S3–S49); WTMADs for various basis sets (Table S50); MADs for HF and various MP2 variants (Tables S51–S55). See DOI: 10.1039/c0cp02984j

problems is not broad but rather specialized (see *e.g.* ref. 11), which makes their application sometimes very difficult for 'non-experts'.

Benchmarking is, consequently, a crucial process that helps better understanding a quantum chemical method's performance. This was first realized by Pople and co-workers, who developed the G1 test set,[12] which later evolved to the G2/97,[13] G3/99[14] and G3/05 sets.[15] However, these works mainly focused on atomization energies (or, equivalently, heats of formation). Electron affinities, ionization potentials, and proton affinities of small molecules played an additional minor role. Truhlar and co-workers extended the idea of benchmarking by introducing databases covering a wide variety of different physico-chemical properties, including reaction energies, barrier heights and noncovalent interactions.[16–24] Important and popular benchmark sets introduced by other groups are for example the S22 set for noncovalent systems[25] (recently revised with new reference values[26,27]), the ISO34[28] set for isomerization energies of organic molecules and the 'mindless-benchmark' set (MB08-165)[29] for testing decomposition energies of randomly created molecular systems.

Having access to a plethora of different published benchmark sets, the question is, which combination of these is the best to obtain a thorough insight into a quantum chemical method's performance. In 2010, we published the so-called GMTKN24 database as an answer to this question. It is a collection of 24 previously published or newly developed benchmark sets for general main group thermochemistry, kinetics and noncovalent interactions.[30] Very recently, it was extended by six additional sets and dubbed GMTKN30.[31] In total, GMTKN30 comprises 1218 single point calculations and 841 data points (relative energies). Its 30 subsets can be divided into three sections, which are depicted in Fig. 1. These sections cover basic properties (*e.g.* atomization energies, electron affinities, ionization potentials, proton affinities, SIE related problems, barrier heights), various reaction energies (*e.g.* isomerizations, Diels–Alder reactions, ozonolyses, reactions involving alkaline metals), and noncovalent interactions (water clusters, conformational energies, and inter- and intra-molecular London-dispersion interactions). Reference values for all subsets are based on highly accurate theoretical or experimental data; for more details see ref. 30 and 31. More information on the subsets is also given in Table 1.

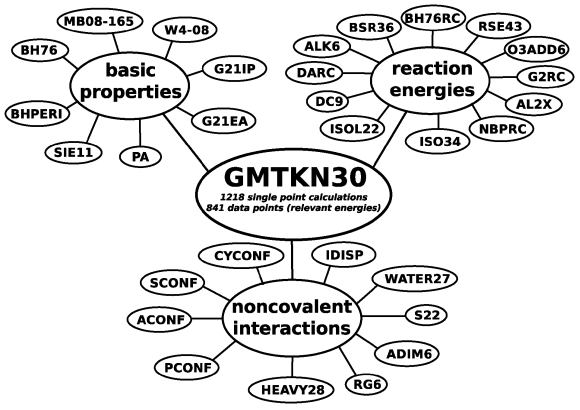

Fig. 1 Schematic overview of GMTKN30 with its three major sub-sections and their respective benchmark sets.

GMTKN30 covers a large cross section of chemically relevant properties. The systems are mostly dominated by electronic wavefunctions of single-reference character and mainly contain early row elements. Nevertheless, they reflect everyday problems many chemists are challenged with. Therefore, GMTKN30 is ideal for thoroughly evaluating existing methods and developing new density functionals (DFs). In the original publication of its precursor GMTKN24 we demonstrated this by investigating four common (meta-)GGA functionals and tried to answer the question how beneficial a reparameterization of these is.[30] The general outcome of this study was that none of the investigated functionals was equally applicable to every kind of problem. Lifting physical boundary conditions improves the description of some processes, but worsens the results for others. The only exception was a reparameterized version of TPSS, dubbed oTPSS, where 'o' stands for optimized. The performance of oTPSS was even comparable to some hybrid DFs. In the following GMTKN30 publication, two new double-hybrid meta-GGA functionals were developed and compared to other existing double-hybrids.[31] This study revealed that the newly proposed PWPB95 functional is more robust, accurate and less basis-set dependent than the other double-hybrids. PWPB95 is based on reparameterized Perdew–Wang[32] exchange and Becke95[33] correlation and incorporates a spin-opposite scaled perturbative correlation.[34,35] Because of this, it is possible to employ a Laplace algorithm for the MP2 part[36] which brings the formal scaling behavior of $\mathcal{O}(N^5)$ down to $\mathcal{O}(N^4)$, with N being the system size. GMTKN30 was also very useful for demonstrating the importance of adding a London-dispersion correction (DFT-D[37]/DFT-D3[38]) to improve the results, not only for noncovalent interactions but also for reaction energies.[30,31,39,40]

The promising results for GMTKN30 encouraged us to focus on the huge number of published DFs. Thus, in this work we want to give a new and extended perspective on DFT performance for GMTKN30 that was lacking in our previous publication. We will shed light on the questions which many users of DFT ask themselves regularly: Which functional has to be used? Which functional is robust? Which functional is most reliable for the investigation of hitherto unknown problems? We think that GMTKN30 provides useful answers to these questions. We try to show, which functionals can be considered as accurate (on the specific rung on Jacob's Ladder[41]) and broadly applicable. With 'robust' we mean any DF that has a reliable and quick SCF-convergence, a small grid dependence, and that yields often relatively accurate results for various properties without showing any severe outliers.

In total, 47 functionals will be investigated: two LDA, 14 GGA, three meta-GGA, 23 hybrid and five double-hybrid (DHDF) density functionals. Besides the DHDFs, we will also consider other modern approaches, like the Minnesota classes of functionals and range-separated hybrids. Basis set and quadrature grid dependences will also be discussed. For almost all functionals the new dispersion correction DFT-D3 will be applied. This is absolutely necessary in order to get a

Table 1 Description of the subsets within the GMTKN30 database

Set	Description	#	av. $\|\Delta E\|^a$	Ref. method	Ref.
MB08-165	Decomposition energies of artificial molecules	165	117.2	est. CCSD(T)/CBS	*b*
W4-08	Atomization energies of small molecules	99	237.5	W4	*c*
G21IP	Adiabatic ionization potentials	36	250.8	exp.	*d*
G21EA	Adiabatic electron affinities	25	33.6	exp.	*d*
PA	Adiabatic proton affinities	12	174.9	est. CCSD(T)/CBS and W1	*e,f*
SIE11	Self-interaction error related problems	11	34.0	est. CCSD(T)/CBS	*g*
BHPERI	Barrier heights of pericyclic reactions	26	19.4	W1 and CBS-QB3	*c,h,i,j,k*
BH76	Barrier heights of hydrogen transfer, heavy atom transfer, nucleophilic substitution, unimolecular and association reactions	76	18.5	W1 and theor. est.	*l,m*
BH76RC	Reaction energies of the BH76 set	30	21.5	W1 and theor. est.	*l,m*
RSE43	Radical stabilization energies	43	7.5	est. CCSD(T)/CBS	*n*
O3ADD6	Reaction energies, barrier heights, association energies for addition of O_3 to C_2H_4 and C_2H_2	6	22.7	est. CCSD(T)/CBS	*o*
G2RC	Reaction energies of selected G2/97 systems	25	50.6	exp.	*p*
AL2X	Dimerization energies of AlX_3 compounds	7	33.9	exp.	*q*
NBPRC	Oligomerizations and H_2 fragmentations of NH_3/BH_3 systems; H2 activation reactions with PH_3/BH_3 systems	12	27.3	est. CCSD(T)/CBS	*g,r*
ISO34	Isomerization energies of small and medium-sized organic molecules	34	14.3	exp.	*s*
ISOL22	Isomerization energies of large organic molecules	22	18.3	SCS-MP3/CBS	*t*
DC9	Nine difficult cases for DFT	9	35.7	Theor. and exp.	*g,j,u,v,w,x,y,z*
DARC	Reaction energies of Diels–Alder reactions	14	32.2	est. CCSDT/CBS	*q*
ALK6	Fragmentation and dissociation reactions of alkaline and alkaline-cation–benzene complexes	6	44.6	est. CCSD(T)/CBS	*aa*
BSR36	Bond separation reactions of saturated hydrocarbons	36	16.7	est. CCSD(T)/CBS	*bb*
IDISP	Intramolecular dispersion interactions	6	13.5	Theor. and exp.	*s,cc,dd,ee*
WATER27	Binding energies of water, $H^+(H_2O)_n$ and $OH^-(H_2O)_n$ clusters	27	82.0	est. CCSD(T)/CBS; MP2/CBS	*ff*
S22	Binding energies of noncovalently bound dimers	22	7.3	est. CCSD(T)/CBS	*gg,hh*
ADIM6	Interaction energies of *n*-alkane dimers	6	3.3	est. CCSD(T)/CBS	*aa*
RG6	Interaction energies of rare gas dimers	6	0.46	exp.	*aa,ii,jj,kk,ll*
HEAVY28	Noncovalent interaction energies between heavy element hydrides	28	1.3	est. CCSD(T)/CBS	*aa*
PCONF	Relative energies of phenylalanyl–glycyl–glycine tripeptide conformers	10	1.5	est. CCSD(T)/CBS	*mm*
ACONF	Relative energies of alkane conformers	15	1.8	W1h-val	*nn*
SCONF	Relative energies of sugar conformers	17	4.9	est. CCSD(T)/CBS	*g,oo*
CYCONF	Relative energies of cysteine conformers	10	2.1	est. CCSD(T)/CBS	*pp*

[a] Averaged absolute energies in kcal mol^{-1}, excluding ZPVEs. [b] Ref. 29. [c] Ref. 42. [d] Ref. 43. [e] Ref. 44. [f] Ref. 17. [g] Ref. 30. [h] Ref. 45. [i] Ref. 46. [j] Ref. 47. [k] Ref. 48. [l] Ref. 20. [m] Ref. 21. [n] Ref. 49. [o] Ref. 22. [p] Ref. 13. [q] Ref. 50. [r] Ref. 51. [s] Ref. 28. [t] Ref. 52. [u] Ref. 53. [v] Ref. 54. [w] Ref. 55. [x] Ref. 56. [y] Ref. 57. [z] Ref. 58. [aa] Ref. 38. [bb] Ref. 59. [cc] Ref. 60. [dd] Ref. 61. [ee] Ref. 31. [ff] Ref. 62. [gg] Ref. 25. [hh] Ref. 26. [ii] Ref. 63. [jj] Ref. 64. [kk] Ref. 65. [ll] Ref. 66. [mm] Ref. 67. [nn] Ref. 68. [oo] Ref. 69. [pp] Ref. 70.

reliable picture of the 'true' functional performance that would otherwise be contaminated by the results for the important noncovalent interactions.

GMTKN30 allows also a thorough comparison of the DFs with MP2. MP2 is still used in many applications and also for deriving allegedly accurate theoretical reference values (see for *e.g.* ref. 69). As the data in GMTKN30 are based on more accurate methods, it is possible to reliably benchmark MP2 and various spin-scaled variants. Particularly the comparison with DHDFs is of interest.

In Section 2 the computational details are explained. Section 3 discusses the DFT-D3 correction in more detail and the parameters for all tested functionals are presented. For those methods, for which no parameters had yet been published, we will present these for the first time. These values are determined for the original DFT-D3 method,[38] which uses a 'zero-damping' function, and alternatively with the finite 'Becke–Johnson' (BJ)[71–73] damping as recently discussed by us.[40] Section 4 focuses on the interpretation of the results obtained for GMTKN30. Although it is in principle possible to conclude which functional is the best for a specific subset, we will mainly carry out overall statistical analyses to investigate the general accuracy, robustness and broad applicability of the tested methods. During the review process of this manuscript, Zhao and Truhlar published a DFT benchmark study based on four subsets taken from GMTKN30 (SIE11, DC9, DARC, and ALK6).[74] They also tested the Minnesota classes of functionals, some range-separated methods and the B2PLYP functional and gave conclusions for about 30 DFs. Dispersion corrections were used in the old form (although it had been shown that DFT-D3 performs better particularly for alkaline metals) and not applied consistently to all methods.

We want again to emphasize that we regard our present analysis as novel in several aspects. The range of considered chemical properties and the variety of DFs is, to the best of our knowledge, larger than in preceding benchmark studies. Also, for the first time, dispersion effects are considered consistently and results close to the basis set limit (*i.e.*, the 'true'

functional performance) are discussed. A thorough comparison between DHDFs and other modern approaches (Minnesota, range-separated functionals, spin-scaled MP2 methods) is provided.

2. Computational details

All calculations were either carried out with the Turbomole suite of programs (a modified version of Turbomole 5.9 and the original version of Turbomole 6.0.[75–78]) or with Gaussian09[79] (DFs applied with Gaussian are marked in Table 2). In total, two LDAs, 14 GGAs, three meta-GGAs, 23 hybrids and five DHDFs were tested. These are in the LDA case SPW92[80,81] and SVWN[80,82] (in the VWN5-parameterization). The GGA-level comprises B97-D[37] (but with DFT-D3 applied instead of the original DFT-D), BP86,[83–85] BOP,[83,86] various combinations with LYP[87,88] correlation (BLYP, mPWLYP,[89] OLYP[90]), PBE[91] and its reparameterized variants PBEsol,[92] revPBE[93] and combinations of various exchange functionals with PBE correlation (OPBE, BPBE, rPW86PBE[94]). The new SSB[95] and revSSB[96] functionals of Swart *et al.*, which are based on OPTX and PBE, were also investigated. On the meta-GGA level the original TPSS[97] functional and its reparameterized version oTPSS[30] were tested. Also, Zhao's and Truhlar's M06L[98] functional was applied. Tested global hybrid functionals are the popular B3LYP,[99,100] B3PW91,[99] BHLYP[101] and PBE0[102,103] functionals. Also other PBE variants were tested. PBE38[38] (which was used in the development of the DFT-D3 correction) is basically like PBE0, but instead of 25% Fock-exchange (E_X^{HF}) it contains 37.5% (3/8). revPBE0 and revPBE38 are based on the reparameterized revPBE functional and include 25% or 37.5% Fock-exchange, respectively. They are tested here for the first time. TPSS0[104] is based on TPSS and includes 25% E_X^{HF}, in contrast to the also investigated TPSSh[105] (10% E_X^{HF}). Various additional functionals are based on combinations of PW[32] exchange and B95[33] correlation, and differ by either the amount of Fock-exchange or the values of the inherent functional parameters. These functionals were proposed by Zhao and Truhlar and are called PW6B95,[106] MPW1B95,[107] PWB6K[106] and MPWB1K.[107] Also the combination of B95 with B88 is tested (B1B95).[33] Other investigated global hybrids are BMK[108] and Truhlar's Minnesota classes of functionals (M05,[109] M052X,[110] M06,[23] M062X[23] and M06HF[111]). Additionally, three range-separated functionals are also applied: CAM-B3LYP,[112] LC-ωPBE[113] and ωB97X-D.[114] The tested double-hybrids are B2PLYP,[58] B2GPPLYP,[42] DSD-BLYP,[115] XYG3[116] and PWPB95.[31] The first two differ by their amounts of Fock-exchange and perturbative correlation. DSD-BLYP uses the same DFT ingredients as B2PLYP, but adds a spin-component-scaled perturbative contribution.[34] In common DHDFs, the SCF is solved for the hybrid part of the functional, and the resulting orbitals are used for the perturbative treatment. XYG3, however, is based on B3LYP densities, which are then used for the evaluation of the hybrid part (non-selfconsistently) and the perturbative correction. For a closer critical analysis of this approach, see ref. 31. PWPB95 is, as explained in the introduction, based on reoptimized PW and B95 ingredients and makes use of SOS-PT2 correlation.[35]

For most of the methods, the DFT-D3[38] correction was applied with our group's own program *dftd3* (see Section 3 for more details). Exceptions were the LDA functionals (because of their overbinding tendency, the usage of a dispersion correction does not make any sense), the ωB97X-D functional (which technically could only be used with the DFT-D correction from 2006) and XYG3 (for which double-counting effects are expected when combined with a dispersion correction similar to LDA). Note that some functionals were explicitly optimized in combination with the previous DFT-D[37] correction. These are B97-D, SSB, revSSB, oTPSS, and DSD-BLYP. Preliminary investigations on B97-D showed that the parameters did not change much, if the new DFT-D3 correction was applied. Therefore, all mentioned

Table 2 DFT-D3 parameters of all investigated functionals with corresponding reference

Functional	s_6	$s_{r,6}$	s_8	Ref.
B97-D3	1.0	0.892	0.909	[a]
BP86	1.0	1.139	1.683	[a]
BOP	1.0	0.929	1.975	This work
BLYP	1.0	1.094	1.682	[a]
mPWLYP	1.0	1.239	1.098	This work
OLYP	1.0	0.806	1.764	This work
PBE	1.0	1.217	0.722	[a]
PBEsol	1.0	1.345	0.612	This work
revPBE	1.0	0.923	1.010	[a]
BPBE	1.0	1.087	2.033	This work
OPBE	1.0	0.837	2.055	This work
rPW86PBE	1.0	1.224	0.901	[b]
SSB	1.0	1.215	0.663	This work
revSSB	1.0	1.221	0.560	This work
TPSS	1.0	1.166	1.105	[a]
oTPSS	1.0	1.128	1.494	This work
M06-L[c]	1.0	1.581	0.0	This work
B3LYP	1.0	1.261	1.703	[a]
B3PW91[c]	1.0	1.176	1.775	This work
BHLYP	1.0	1.370	1.442	This work
PBE0	1.0	1.287	0.928	[a]
revPBE0	1.0	0.949	0.792	This work
PBE38	1.0	1.333	0.998	This work
revPBE38	1.0	1.021	0.862	This work
TPSSh	1.0	1.223	1.219	This work
TPSS0	1.0	1.252	1.242	[a]
PW6B95	1.0	1.523	0.862	[a]
MPW1B95	1.0	1.605	1.118	This work
PWB6K	1.0	1.660	0.550	This work
MPWB1K	1.0	1.671	1.061	This work
B1B95	1.0	1.613	1.868	This work
BMK[c]	1.0	1.931	2.168	This work
M05[c]	1.0	1.373	0.595	This work
M052X[c]	1.0	1.417	0.0	This work
M06[c]	1.0	1.325	0.0	This work
M062X[c]	1.0	1.619	0.0	This work
M06HF[c]	1.0	1.446	0.0	This work
CAM-B3LYP[c]	1.0	1.378	1.217	This work
LC-ωPBE[c]	1.0	1.355	1.279	This work
B2PLYP	0.64	1.427	1.022	[d]
B2GPPLYP	0.56	1.586	0.760	[d]
PWPB95	0.82	1.557	0.705	[d]
DSD-BLYP	0.50	1.569	0.705	[d]

[a] Ref. 38. [b] Ref. 40. [c] Applied with Gaussian09. ωB97X-D was also applied with Gaussian09, but is not mentioned in this list, as it had to be combined with the older DFT-D correction. [d] Ref. 31.

functionals are kept electronically in their original form and combined with DFT-D3 in the standard way.

For all DFs, the large Ahlrichs' type quadruple-ζ basis sets def2-QZVP were applied.[117] For the calculations of electron affinities (G21EA), diffuse s- and p-functions (for hydrogen only an s-function) were added from the Dunning aug-cc-pVQZ basis sets;[118] the resulting set is denoted by aug-def2-QZVP. As discussed previously,[30,31] one diffuse s- and one diffuse p-function (taken from aug-cc-pVQZ) were added to oxygen in the case of WATER27. In some cases we also carried out calculations with the def2-TZVPP and def2-SV(P) sets and used diffuse functions from aug-cc-pVTZ/aug-cc-pVDZ where necessary (see Section 4.2). To account for scalar relativistic effects, the heavy atoms in HEAVY28 and RG6 were treated with the effective core potentials ECP-28 (for Sb-Xe) and ECP-60 (for Pb, Bi, and Rn),[119,120] which were slightly modified by Weigend and Ahlrichs and called 'def2-ecp' in Turbomole.[117] In all cases, SCF-convergence criteria were set to $10^{-7}E_h$. All open-shell calculations were done within the unrestricted Kohn–Sham formalism (UKS).

Additional calculations were carried out for HF, MP2 and its spin-scaled versions SCS-MP2,[34] S2-MP2[121] and SOS-MP2.[35] S2-MP2 differs from SCS-MP2 by the scaling factors for the same and opposite spin. Extrapolations to the complete basis set (CBS) limit[122] were carried out with the same Ahlrichs type basis sets as described above.

In all Turbomole calculations, the hybrid functionals and the hybrid-(meta-)GGA parts of the double-hybrids were treated within the resolution of the identity (RI-JK) approximation.[123] This is also the case for the HF calculations. For the perturbative parts of the DHDFs and the MP2 variants, the RI approximation was used as well.[78] All electrons were fully correlated for the DHDF perturbative calculations. For the MP2 methods, the frozen core approximation was applied. The RI-J approximation was applied to LDA and (meta-)GGA functionals. All auxiliary basis functions were taken from the Turbomole basis set library.[124,125] For GMTKN30 DFT calculations, the Turbomole grid *m*4 was used, whereas the larger *m5* grid was chosen in the DFT-D3 fitting procedure.[125] The 'm-options' apply smaller grids for the actual SCF and finally evaluate the exchange-correlation potential with the converged KS-orbitals (see ref. 125 for more details). In some cases, a grid dependence study was carried out with the large 'reference' grid, which is comparable to Gaussian's large 'ultrafine' grid (see Section 4.1.3.2).

GAUSSIAN09 calculations were carried out without any density fitting approximations. The 'fine grid' was used in all cases (a pruned grid with 75 radial shells and 302 angular points per shell). Preliminary investigations with PBE showed no major differences between the Turbomole and Gaussian set-ups (the statistical results differed by less than 0.015 kcal mol^{-1}). The 'ultrafine grid' (99 radial shells and 590 angular points per shell) was used for some functionals in the grid dependence study (see Section 4.1.3.2).

In our recent studies on GMTKN24 and GMTKN30, we defined a weighted total mean absolute deviation (WTMAD) to combine all obtained mean absolute deviations (MADs) for each subset into one final number for a tested method. We also discussed that such a procedure can be defined in several ways and that there is in general no clear right or wrong. After having tested several schemes, we had found that the overall ranking of methods was not altered. In the scheme which we finally presented (see eqn (1)), each of the 30 MAD values was weighted by the number of entries (N_i) of each subset to take into account its size. Furthermore, each subset was weighted by an additional factor that was calculated as the ratio between the MADs of BLYP and B2PLYP-D [*i.e.* MAD(BLYP)/MAD(B2PLYP-D)] to take into account the difficulty of a certain subset.

$$\mathrm{WTMAD} = \frac{1}{3091.4} \cdot \sum_{i}^{30} N_i \cdot \frac{\mathrm{MAD}_i^{\mathrm{BLYP}}}{\mathrm{MAD}_i^{\mathrm{B2PLYP\text{-}D}}} \cdot \mathrm{MAD}_i \quad (1)$$

In order to be consistent and although the new DFT-D3 correction is applied in the present study, the WTMAD is still defined with the older version DFT-D. The pre-factors for each subset underlie the constraint that the product of system size and scale factor of a certain set should not be larger than one half of the corresponding value for MB08-165, *i.e.*, 222.75. Therefore, it is guaranteed that smaller sets with a large scale factor enter not too strongly. The actual values for the weighting factors of all 30 subsets are given in the ESI.† In the following discussions, WTMADs are shown for the complete GMTKN30 set and for its three major sub-classes.

3. The DFT-D3 London-dispersion correction

Although common functionals sometimes seem to qualitatively give correct interaction-potentials for some weakly bound compounds at equilibrium distances (*e.g.* BLYP or B3LYP) and although highly parameterized DFs (like the Minnesota classes of functionals) have been designed to account for medium-range dispersion effects, it is nowadays clear that all semi-local DFs and conventional hybrid functionals (that include nonlocal Fock-exchange) asymptotically cannot provide the correct $-C_6/R^6$ dependence of the dispersion interaction energy on the inter-atomic(molecular) distance R. Various approaches to tackle this problem were proposed in the literature; for a recent review see *e.g.* ref. 39. In the following, only the DFT-D approach will be discussed in detail. It provides a dispersion energy $E_{\mathrm{disp}}^{\mathrm{DFT\text{-}D}}$, which can be added to the result of a standard DFT calculation without any additional computational cost. After the first two versions in 2004[126] and 2006,[37] our group proposed very recently a new variant of this correction called DFT-D3,[38] of which we will make use in the present study. Compared to the previously published versions, DFT-D3 contains more '*ab initio*' ingredients and is characterized by less empiricism. It covers more elements (H-Pu) than the first DFT-D versions and performs in general better, in particular for metal systems. The general form for the dispersion energy $E_{\mathrm{disp}}^{\mathrm{DFT\text{-}D3}}$ is

$$E_{\mathrm{disp}}^{\mathrm{DFT\text{-}D3}} = -\frac{1}{2}\sum_{A\neq B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{\mathrm{damp},n}(R_{AB}), \quad (2)$$

where the sum is over all atom pairs in the system, C_n^{AB} denotes the averaged (isotropic) nth-order dispersion coefficient (orders n = 6,8) for atom pair AB, and R_{AB} is their

internuclear distance. In DFT-D3 these coefficients depend on the coordination sphere of each atom within a molecule (see ref. 38 for more details). The s_n's are global (DF dependent) scaling factors. For common DFs, s_6 is set to unity to assure that the DFT-D3 correction has a physically correct asymptotic behavior. For double-hybrids, an s_6 value smaller than unity has to be chosen, because of the presence of the nonlocal PT2 contribution (note that a dispersion correction is still necessary, as the PT2 contribution is scaled down). Recently, we introduced a scheme to determine the s_6 values for DHDFs consistently.[31] s_8 is adjusted specifically for each functional. Note that this term is more short-ranged and rather strongly interferes with the (short-ranged) DF description of electron correlation. In order to avoid near-singularities for small R and double-counting effects of correlation at intermediate distances, damping functions $f_{\mathrm{damp},n}$ are used which determine the range of the dispersion correction. The original expression for DFT-D3 is[37]

$$f_{\mathrm{damp},n}(R_{AB}) = \frac{1}{1+\mathrm{e}^{-\gamma(R_{AB}/s_{r,n}R_0^{AB}-1)}}, \qquad (3)$$

where R_0^{AB} is a cut-off radius for atom pair AB. $s_{r,6}$ is a DF-dependent (global) scaling factor (as introduced in ref. 127) that has to be fitted; $s_{r,8}$ is set to unity for all DFs. γ (which is set to 14 for the R^{-6} part and to 16 for the R^{-8} part) is a global constant that determines the steepness of the functions for small R. The damping function was the recently dubbed 'zero-damping' function, as it goes to zero for small R.[40] Recently, we tested the new DFT-D3 with Becke's and Johnson's rational damping function[71–73] and dubbed this variant DFT-D3(BJ).[40] This so-called BJ-damping leads to a constant contribution of E_{disp} to the total correlation energy from each spatially close pair of atoms (*i.e.*, directly bonded). In this variant the dispersion energy is given by

$$E_{\mathrm{disp}}^{\mathrm{DFT\text{-}D3(BJ)}} = -\frac{1}{2}\sum_{A\neq B}\sum_{n=6,8} s_n \frac{C_n^{AB}}{R_{AB}^n + f(R_{AB}^0)^n} \qquad (4)$$

with

$$f(R_{AB}^0) = a_1 R_{AB}^0 + a_2. \qquad (5)$$

Here, the sum runs over all atom pairs in the system and a_1 and a_2 are free fit parameters introduced by BJ. R_{AB}^0 is defined as[73]

$$R_{AB}^0 = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}. \qquad (6)$$

Thus, DFT-D3 rquires for each functional two fit parameters ($s_{r,6}$ and s_8), whereas DFT-D3(BJ) has three (a_1, s_8 and a_2). The parameters are determined in a least squares fit to a set of 130 noncovalent interaction energies. During the fitting procedure problems arose with the Minnesota classes of functionals. These highly parameterized DFs were developed to describe middle-range correlation effects and noncovalent interactions correctly. In combination with DFT-D3, this leads to double-counting effects and deteriorating results. Therefore, s_8 was set to zero for all functionals except for M05, which did not show this behavior. Nevertheless, also these functionals usually benefit from including a dispersion correction and furthermore become also asymptotically correct. The Minnesota functionals, however, proved to be incompatible with DFT-D3(BJ), which in general provides more 'double-counting' than DFT-D3.

The D3 parameters of all functionals are given in Table 2. All functional values are implemented into the *dftd3* program which is available on our website.[128] In this paper, only the DFT-D3 correction will be applied. The DFT-D3(BJ) correction leads, however, to the same conclusions, regarding the best and most robust methods. DFT-D3(BJ) parameters of all functionals are shown in the ESI.† The benefit from including a dispersion correction to the GMTKN subsets has also been discussed elsewhere.[30,31,39,40] It was shown that it does not only improve the description of noncovalent interactions but also of the basic properties and particularly the reaction energies. Therefore, only results with DFT-D3 will be discussed in the following.

4. Results and discussion

In the following sections, the results for all investigated functionals will be shown. Only where necessary, MADs for specific benchmark sets will be discussed. In general, we do not want to recommend a definite DF for each subset. However, the interested reader is referred to Table S2 of the ESI†, which lists the best DFs for each benchmark set and for each rung of Jacob's Ladder. Tables S3–S49 show all MADs and RMSDs for the DFs with and without dispersion correction. All results are also published on our website.[128] Our main discussion focuses on an assessment of the general applicability of DFs and the following analyses are based on WTMADs for the complete GMTKN30 database and for its three sub-sections. Moreover, an estimate of the robustness of a DF is given by counting how often it yields the best or worst result. This procedure in combination with the WTMADs allows a good estimate of a DF's performance. For example a DF which gives at the same time very good and very bad MADs cannot be regarded as reliable. On the other hand, a DF which gives less times the best and the worst MAD, but has also a very good WTMAD, is considered to be better and more robust. The discussion sheds light on the different rungs of Jacob's Ladder. Also the quadrature grid dependences of some functionals are discussed. An overview of all functionals is given and the basis set dependence is studied. Finally, the results for the MP2 variants are briefly discussed and compared to DHDFs.

4.1 Results with (aug-)def2-QZVP

4.1.1 LDA, GGA and meta-GGA functionals. Fig. 2 shows the WTMADs of LDA, GGA and meta-GGA functionals for the complete GMTKN30 set and its three sub-sections. The LDA functionals yield, as expected, the worst results of all DFs. Particularly the basic properties, which describe severe changes in chemical structures (*e.g.* atomization), are described very badly (WTMAD of 21.7 kcal mol^{-1}). Due to error compensation, the difference between LDA and GGA is smaller for the reaction energies (WTMAD of 6.9 kcal mol^{-1}). The total WTMADs for the complete database [part (d) of Fig. 2]

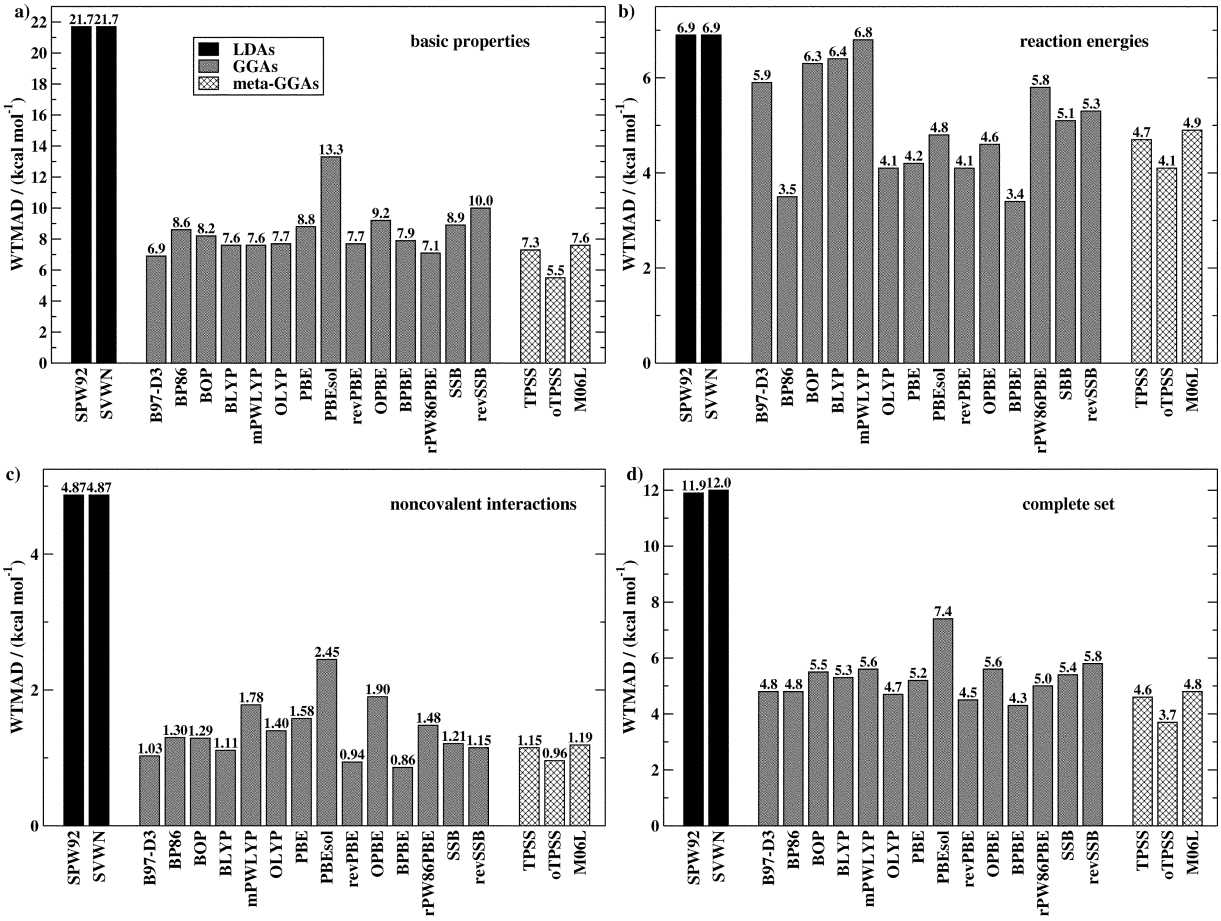

Fig. 2 Weighted total mean absolute deviations (WTMADs) of LDAs, GGAs and meta-GGAs in kcal mol^{-1} for the basic properties (a), the reaction energies (b), the noncovalent interactions (c) and the complete GMTKN30 set (d). All values are based on (aug-)def2-QZVP calculations. All functionals were combined with the DFT-D3 correction, except for the LDAs. The suffix '-D3' was skipped.

are 11.9 (SPW92) and 12.0 kcal mol^{-1} (SVWN). Although they differ substantially in their absolute energies, the relative energies for SPW92 and SVWN are basically identical. There seems not to be a chemically significant difference between the two correlation functionals.

Before continuing with the statistical analysis for the GGAs, one comment has to be made on the absolute energies of atoms and molecules. In most cases only relative energies between similar species are considered and DFs can profit from a fortuitous error-compensation. Nevertheless, a proper DF should also be able to compute total electronic energies within a reasonable accuracy. As an example the water molecule is discussed. The estimated experimental value is about $-76.432E_h$.[129] Most functionals have an error of $\pm 0.15E_h$ and double-hybrids lie closest to this value. However, we note some DFs with larger deviations. The LDAs and PBEsol-D3 have higher energies of about -75.9 and $-76.1E_h$, respectively. SSB-D3 and revSSB-D3 yield unusually low energies of about $-77E_h$. These trends are also observed for other systems. This has in particular a strong influence for the atomization energies of the W4-08 set for the LDAs, PBEsol-D3 and the two SSB variants: the errors are about 47 kcal mol^{-1} for the LDAs, 24.7 kcal mol^{-1} for PBEsol-D3, 12.1 kcal mol^{-1} for SSB-D3 and 16.0 kcal mol^{-1} for revSSB-D3. The other GGAs are in a range of 4 to 8 kcal mol^{-1}, with the exception of PBE, which is known for its deficiency in describing atomization energies (13 kcal mol^{-1}).[93]

In the original GMTKN24 study, only the GGAs BLYP, mPWLYP and PBE were discussed and the general outcome was that none of them is particularly better than the other. This herein presented and extended study with 14 GGAs gives a clearer picture and reveals that indeed some GGA functionals are in general better than others.

B97-D3 and rPW86PBE-D3 have the lowest WTMADs for the basic properties [6.9 and 7.1 kcal mol^{-1}, part (a) of Fig. 2]. The three functionals with LYP correlation, revPBE-D3 and BPBE-D3 follow with WTMADs of 7.6 to 7.9 kcal mol^{-1}. BOP-D3, PBE-D3, SSB-D3 and OPBE-D3 are in a range of 8.2 to 9.2 kcal mol^{-1}. revSSB-D3 and PBEsol-D3 have the highest WTMADs with 10.0 and 13.3 kcal mol^{-1}.

A different picture is seen for the reaction energies in part (b) of Fig. 2. mPWLYP-D3 and BLYP-D3 have the worst WTMADs (6.8 and 6.4 kcal mol^{-1}), but OLYP-D3 is much better (4.1 kcal mol^{-1}), which indicates the influence of the

exchange functional for the considered types of reactions. revPBE-D3 has the same WTMAD as OLYP-D3; PBE-D3 lies very close with 4.2 kcal mol^{-1}. The other PBE-like functionals (including the SSB variants) are by up to 1 kcal mol^{-1} higher. The only exception is BPBE-D3, which yields the best WTMAD with 3.4 kcal mol^{-1}, followed by BP86-D3 with 3.5 kcal mol^{-1}.

The investigation of the noncovalent interactions revealed a surprising flaw in some DFs. Some methods are not able to bind certain systems even when dispersion corrected. This is observed for DFT-D3 and DFT-D3(BJ) and thus seems to be more an inherent problem of the DF and not of the correction. Test calculations with various quadrature grids also ruled out the choice of grid as a possible error-source. BPBE-D3 and BP86-D3 do not bind the lighter rare gas dimers Ne_2 and Ar_2 in the RG6 set, whereas the heavier dimers are bound. BOP-D3 does not bind Ne_2 and the PbH_4–water complex in HEAVY28. OPBE-D3 shows the same behavior for this complex.

In previous investigations, BLYP-D3 already was found to perform very well for noncovalent interactions.[30] Particularly, the low MAD of 0.24 kcal mol^{-1} for the S22 set is astonishing as it is the second best value of all 47 functionals. BLYP-D3 is also the best GGA functional for the PCONF test set for relative energies of tripeptide conformers. For the complete subclass of noncovalent interaction benchmarks, BLYP-D3 has a WTMAD of only 1.11 kcal mol^{-1} [part (c) of Fig. 2]. Substituting B88 by OPTX exchange worsens this result to 1.40 kcal mol^{-1}, which is the opposite behavior compared with the reaction energies. Three DFs show an overall better performance than BLYP-D3 for the considered test sets. These are B97-D3 (1.03 kcal mol^{-1}), revPBE-D3 (0.94 kcal mol^{-1}) and BPBE-D3 (0.86 kcal mol^{-1}). However, the latter functional is not recommended, due to the described inability of binding the lighter rare gas dimers. Previously reported findings that the SSB variants work particularly well for noncovalent interactions[96] cannot be confirmed for the present test sets (WTMADs of 1.21 for SSB-D3 and 1.15 for revSSB-D3). The worst GGA functionals for noncovalent interactions are mPWLYP-D3, OPBE-D3 and PBEsol-D3.

Part (d) of Fig. 2 shows the total WTMADs for the complete GMTKN30 set. Five GGA functionals have a WTMAD below 5 kcal mol^{-1}: BPBE-D3 (4.3 kcal mol^{-1}), revPBE-D3 (4.5 kcal mol^{-1}), OLYP-D3 (4.7 kcal mol^{-1}), B97-D3 and BP86-D3 (4.8 kcal mol^{-1} both). However, considering the WTMADs of the three main sections and the mentioned problems of some functionals, we recommend revPBE-D3 and B97-D3 as robust and well applicable methods. Considering also the number of empirical parameters in the electronic part of these two GGAs (one and nine, respectively) one can finally nominate revPBE-D3 as a 'winner' in the GGA category.

This finding is additionally supported by an analysis of how many times a GGA functional yields the best and the worst MAD. The results are depicted in part (a) of Fig. 3. OPBE-D3 yields six times the best MAD, but also three times the worst, which indicates that it is not very robust. This is in accordance with the high WTMAD of 5.6 kcal mol^{-1} for the complete set. PBEsol-D3 is the worst GGA functional (worst MADs in 16 cases) and is not recommended for molecular chemistry. The recommended B97-D3 and revPBE-D3 behave also in this analysis well. B97-D3 provides in five cases the best MAD (mainly for the basic properties) and revPBE-D3 in two cases. Results for other test sets also lie often close to the best result for that certain set. In no cases large outliers are observed and they never yield the worst MAD.

In the original GMTKN24 study, also the TPSS and oTPSS methods were considered. Here this investigation is repeated for GMTKN30 with the DFT-D3 correction and additionally including the M06L functional. Initially, the VSXC[130] functional was also tested. However, it is not considered here due to frequent SCF-convergence problems and an incompatibility with the DFT-D and DFT-D3 corrections. Comparing just these three meta-GGAs with each other and counting how many times they are the best and worst functionals reveals that TPSS-D3 and oTPSS-D3 both yield the best MADs in 12 cases [part (b) of Fig. 3]. M06L-D3

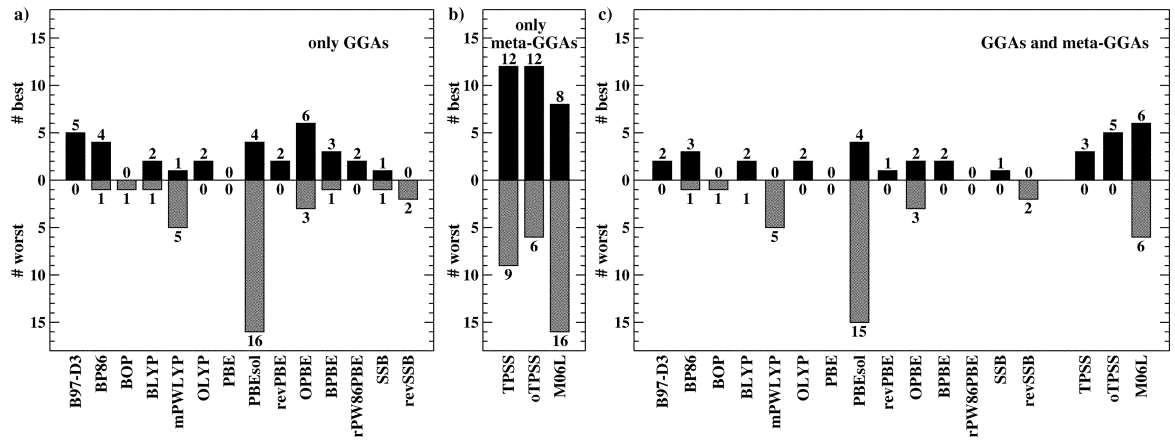

Fig. 3 Analysis of how many times a DF yields the best and the worst MAD for the test sets of GMTKN30. The analysis is first separately carried out for GGAs (a) and for meta-GGAs (b) and then combined for both functional classes (c). All functionals were combined with the DFT-D3 correction. The suffix '-D3' was skipped.

performs best for eight subsets. In 16 cases, however, it yields the worst MAD. oTPSS-D3 is the most robust meta-GGA (worst MAD in only six cases, compared to TPSS-D3 with nine cases). For all three sub-sections and the complete GMTKN30, oTPSS-D3 yields the lowest WTMADs (Fig. 2). M06L-D3 has the highest WTMADs of all three meta-GGAs. One M06L-calculation for the MB08-165 test set did not converge, although various convergence techniques were applied. Therefore, only 164 entries of MB08-165 were considered in the calculation of the WTMADs. Moreover, an inability of binding the Kr_2 and Xe_2 dimers in RG6 and the PbH_4–water and $(TeH_2)_2$ complexes in HEAVY28 was observed. Again, this finding is independent from the choice of quadrature grid or whether a dispersion correction is used or not.

Carrying out the analysis of the best and worst MADs for the GGAs and meta-GGAs together, we find that M06L-D3 yields six times the best MADs [part (c) of Fig. 3], closely followed by oTPSS (five times). However, M06L-D3 is also in six cases worse than all other functionals, whereas the other meta-GGAs are more robust in this sense. The figure also shows that meta-GGAs are not necessarily better than GGAs. This is also indicated by the WTMADs. TPSS-D3 and M06L-D3 have WTMADs that are in the range of typical GGA functionals. In the cases of the noncovalent interactions, no meta-GGA can be recommended, as the above mentioned GGAs are also computationally simpler. In a nutshell, the only meta-GGA showing an improvement compared to GGAs is oTPSS-D3 with an overall WTMAD of 3.7 kcal mol^{-1}, which also competes with some hybrid functionals, as will be seen in the next section.

4.1.2 Global hybrid functionals. The WTMADs for the conventional hybrid functionals are shown in Fig. 4. The by far best conventional hybrids for the basic properties are PW6B95-D3, MPW1B95-D3, B1B95-D3 and BMK-D3. Particularly, the MAD for the MB08-165 set is extraordinary low for PW6B95-D3 with 4.7 kcal mol^{-1} (Table S31, ESI†). Other hybrids usually lie within a range of 5.7 to 8 kcal mol^{-1}, sometimes higher than 10 kcal mol^{-1}. The WTMAD of B3LYP-D3 is 5.0 kcal mol^{-1}, which is only an average value. PBE- and TPSS-based hybrids have higher WTMADs. The by far worst hybrid for the basic properties is BHLYP, which is with 9.5 kcal mol^{-1} also worse than most (meta-)GGAs.

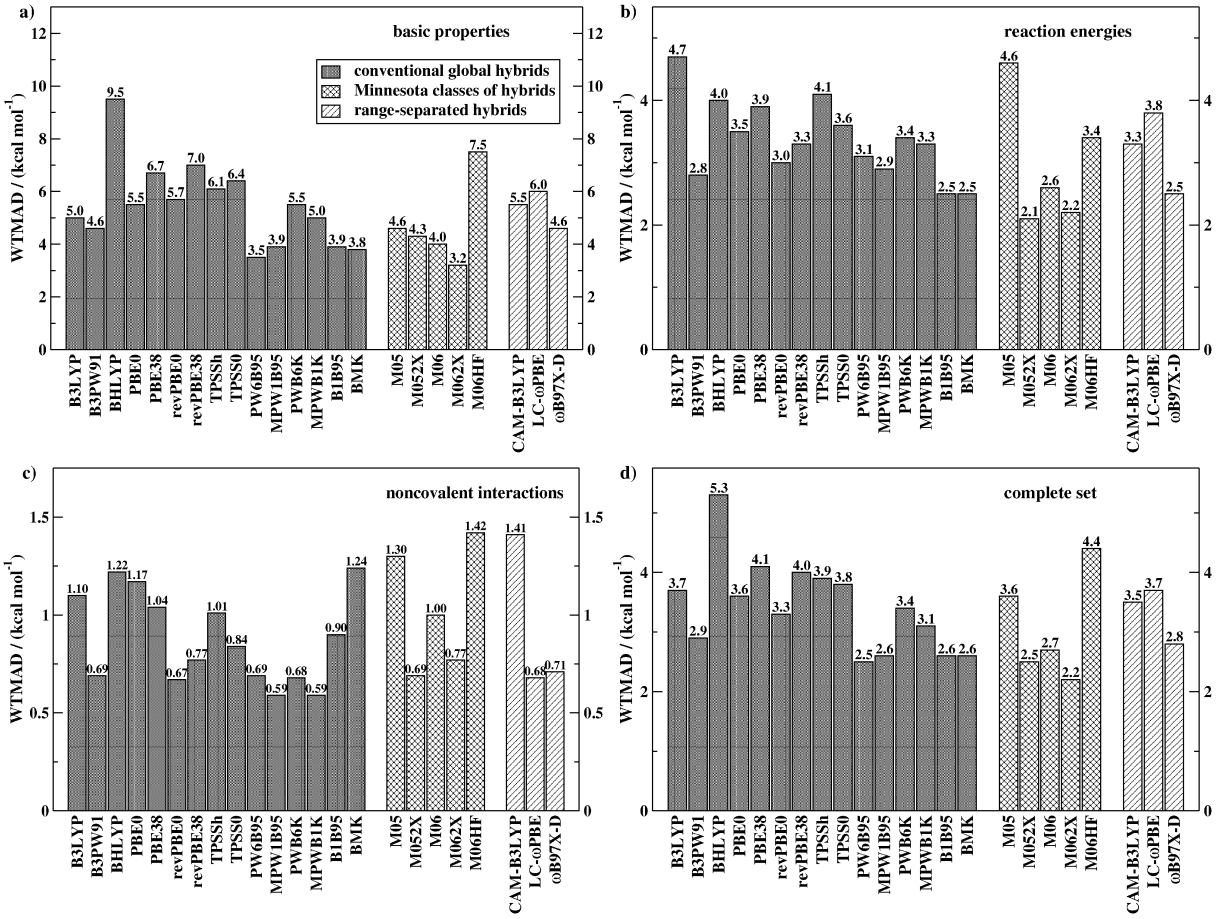

Fig. 4 Weighted total mean absolute deviations (WTMADs) of all tested hybrid functionals in kcal mol^{-1} for the basic properties (a), the reaction energies (b), the noncovalent interactions (c) and the complete GMTKN30 set (d). All values are based on (aug-)def2-QZVP calculations. All functionals were combined with the DFT-D3 correction, except for ωB97X-D (older DFT-D version). The suffix '-D3' was skipped.

The analysis of reaction energies reveals a very important fact. B3LYP-D3, the most popular functional, which is often used in everyday calculations for similar reactions, has the worst WTMAD of all hybrids with 4.7 kcal mol^{-1}. Its precursor B3PW91-D3 is one of the best functionals, though, with 2.8 kcal mol^{-1}. The best WTMADs are obtained for BMK-D3 and B1B95-D3 (2.5 kcal mol^{-1} both). MPW1B95-D3 (2.9 kcal mol^{-1}), revPBE0-D3 (3.0 kcal mol^{-1}), and PW6B95-D3 (3.1 kcal mol^{-1}) follow closely.

An analysis of noncovalently bound complexes shows again that some DFs have problems here. BMK-D3 does not bind the neon, argon and krypton dimers (RG6), the PbH_4–water complex (HEAVY28) and the methane dimer (S22). B1B95-D3 does not bind the PbH_4–water complex and the neon and argon dimers. B3PW91-D3 shows problems with the neon dimer. Having said that the best GGA functional has a WTMAD of 0.94 kcal mol^{-1} for noncovalent interactions, only those hybrids can be recommended, which are far below this value. These are in particular MPW1B95-D3, MPWB1K-D3 (0.59 kcal mol^{-1} both), revPBE0-D3 (0.67 kcal mol^{-1}), PWB6K-D3 (0.68 kcal mol^{-1}), and PW6B95-D3 (0.69 kcal mol^{-1}). Finally, the total WTMADs for the complete GMTKN30 set are considered in part (d) of Fig. 4. The overall best hybrid functional is PW6B95-D3 (2.5 kcal mol^{-1}), followed closely by MPW1B95-D3 (2.6 kcal). B1B95-D3 and BMK-D3 have identical values. B3PW91-D3 is with 2.9 kcal mol^{-1} better than B3LYP (3.7 kcal mol^{-1}). All other conventional hybrids provide values >3–4 kcal mol^{-1}. BHLYP-D3 is statistically similar to many GGAs (5.3 kcal mol^{-1}) and cannot be recommended. Also the PBE-based functionals do not show an overall good performance. With TPSS as ingredient, the results even worsen.

The values shown in part (a) of Fig. 5 underline the outcome of the WTMAD-analysis. BHLYP-D3, B3LYP-D3 and BMK-D3 cannot be considered as robust, as they yield in many cases the worst MADs. The by far best functional according to this analysis is PW6B95-D3, which has the best MAD in ten cases and the worst just in one case (the radical stabilization energies in RSE43, together with MPW1B95-D3 and B1B95-D3).

4.1.3 The M05 and M06 classes of hybrid-meta-GGA functionals. Very recently, Zhao and Truhlar carried out an analysis of the Minnesota classes of functionals for four test sets of GMTKN30.[74] Herein, we will give a broader overview of these methods and also investigate the effects of the dispersion correction and the quadrature grids.

4.1.3.1 Results for the 'finegrid' option. The analysis of the M05 and M06 classes of functionals is carried out in the same way as for the conventional hybrids. The WTMADs are also shown in Fig. 4. For these functionals a sometimes very slow SCF-convergence was observed. Even for small systems like H_2^+, C^-, F^+ and the sodium and sulfur atoms, special convergence procedures (*e.g.* quadratic-SCF[131]) had to be applied sometimes. In some cases, the calculations did not converge at all (*e.g.* for some rather difficult systems in the MB08-165 set). M06, M062X and M052X did not converge in one case, M06HF in five cases. Also the C_2 molecule could not be brought to convergence with M06HF. All these systems were left out of the statistical analyses.

For the MB08-165 set, M062X-D3 has, averaged over 164 systems, the same MAD as PW6B95-D3 (averaged over all 165 systems). The WTMAD for basic properties is with 3.2 kcal mol^{-1} lower than for PW6B95-D3. The other functionals behave like conventional hybrids. M06HF-D3 represents an outlier with an WTMAD of 7.5 kcal mol^{-1} [part (a) of Fig. 4].

For reaction energies, only M052X-D3 (WTMAD = 2.1 kcal mol^{-1}), M062X-D3 (2.2 kcal mol^{-1}) and M06-D3 (2.6 kcal mol^{-1}) can compete with the other hybrids. M06HF-D3 is with 3.4 kcal mol^{-1} comparable to PBE0-D3 (3.5 kcal mol^{-1}), M05-D3 (4.6 kcal mol^{-1}) behaves similarly to B3LYP-D3. Although the functionals were specifically designed to describe noncovalent interactions, problems arose for some systems (with and without dispersion correction). M06 does

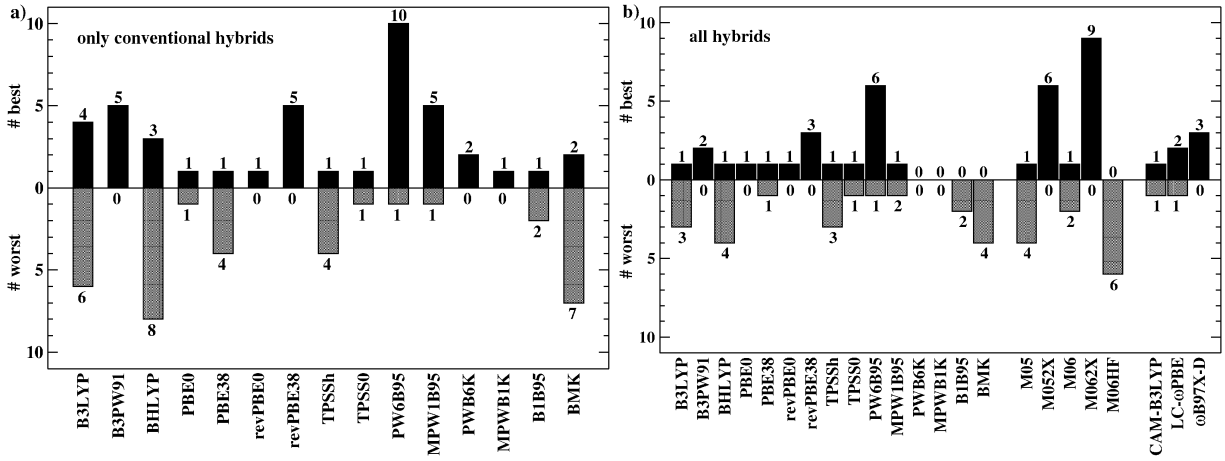

Fig. 5 Analysis of how many times a DF yields the best and the worst MAD for the test sets of GMTKN30. The analysis is first separately carried out for conventional, global hybrids (a) and then combined for all tested hybrids (b). All functionals were combined with the DFT-D3 correction, except for ωB97X-D (older DFT-D version). The suffix '-D3' was skipped.

not bind the krypton dimer. M06HF does not bind five of the 28 heavy element complexes in HEAVY28. The WTMADs for the noncovalent interactions show that only M052X-D3 is competitive to the best conventional hybrids (0.69 kcal mol^{-1}). The other methods have higher WTMADs; M05-D3 and M06HF-D3 have, in fact, worse WTMADs than all other conventional hybrids (1.30 and 1.42 kcal mol^{-1}, respectively).

The total WTMADs in part (d) of Fig. 4 show again that M062X-D3, M052X-D3 and M06-D3 are the best functionals of the whole class of Minnesota functionals. M062X-D3 yields, in fact, the lowest WTMAD of all hybrids with 2.2 kcal mol^{-1}. M052X-D3 is similar to PW6B95-D3 (2.5 kcal mol^{-1}), M06-D3 has a WTMAD of 2.7 kcal mol^{-1}. M05-D3 is with 3.6 kcal mol^{-1} comparable to B3LYP-D3. M06HF-D3 has a relatively large WTMAD of 4.4 kcal mol^{-1}. Part (b) of Fig. 5 shows how many times the functionals perform best or worst. This analysis is now based on all tested hybrids and allows a direct comparison between conventional and Minnesota functionals. M062X-D3 outperforms in this analysis all other hybrids and is the best method in nine cases. M052X-D3 follows with six cases. PW6B95-D3 remains the only conventional hybrid competing with these methods (six cases), although it still yields the worst MAD in one case. M05-D3, M06-D3, and M06HF-D3 are not competitive with the other hybrids and yield often the worst MADs. Thus, although, M06-D3 has a low total WTMAD, the other conventional methods with similar WTMADs are in general more robust.

4.1.3.2 Grid dependence. Quantum chemistry packages offer various quadrature grids for the numerical integration of the exchange-correlation potential and energy. The size of the grid has a significant impact on computation times, especially when the Coulomb interactions are treated efficiently as *e.g.* in density-fitting (RI) schemes. However, several authors realized that results for certain functionals and properties might oscillate for different grid sizes.[132–135] This effect is typically larger for meta-GGA functionals. This was first shown for noncovalent interactions, but later also for reaction energies.[136–138] Particularly, the Minnesota classes of functionals were reported to show these oscillations, both in geometry optimizations and single-point energy calculations.[69,139–142] Wheeler and Houk[143] did a thorough analysis of these functionals for the ISO34 set of isomerization energies and found strong oscillations particularly for the SG-1 grid,[144] which is a standard option in QCHEM. They concluded that this is caused by the sometimes unusually high values of the numerous fit parameters in the kinetic energy density enhancement factors of the exchange parts of these functionals. We decided to investigate the grid dependence for these functionals also for the GMTKN30 set. We start the study by closely investigating the M06HF functional (here without DFT-D3), as it was shown to be most affected by this problem. The study is carried out with the 'fine' and 'ultrafine' grid options. Also here SCF-convergence problems were observed and calculations of C_2 and five systems in MB08-165 did not converge. Interestingly, these were not always the same systems that had not converged for the fine grid.

Fig. 6 shows the ratios between the MADs for the larger and the smaller grid for all 30 subsets. Usually the ratios are close

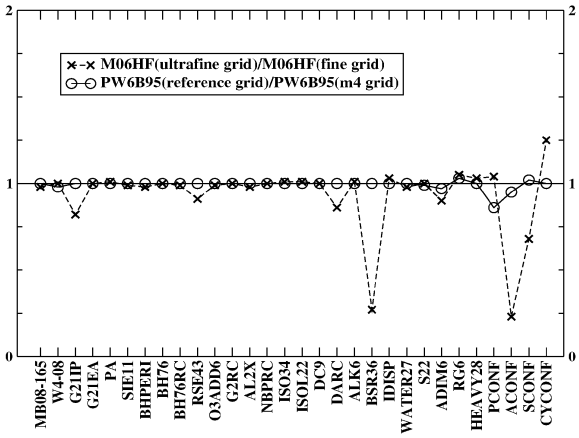

Fig. 6 Ratios between MADs for M06HF with the ultrafine and fine grid options (Gaussian09) and between the MADs for PW6B95 with the reference and *m*4 grids (TURBOMOLE).

to unity and the difference between the MADs is at the most 0.1 kcal/mol, but usually less for the majority of systems. However, strong outliers are observed for G21IP, BH76RC, DARC, BSR36, ADIM6, ACONF, SCONF and CYCONF. Note, that this behavior for the sugar conformers has already been mentioned by Csonka *et al.*[69] The MADs for both grids for the four most affected test sets are shown in Table 3. The MAD for the ionization potentials (G21IP) improves by 1 kcal mol^{-1} for the larger grid. The alkane bond separation energies in BSR36 show a huge difference between both grids of 3.4 kcal mol^{-1}. The MAD for ACONF increases by more than a factor of four with the ultrafine grid. SCONF is also described better by 0.12 kcal mol^{-1}.

Fig. 6 also shows the ratios for the less parameterized PW6B95 functional for the smaller TURBOMOLE grid *m*4 and the large reference grid, which is basically Gaussian's ultrafine grid. Less differences are found for this hybrid-meta-GGA functional. Only the tripeptides show a difference of 0.2 kcal mol^{-1} between both grids. Table 3 also shows the MADs for PW6B95 for the four sets that were particularly

Table 3 Influence of the quadrature grid on the MADs for four selected test sets. All MADs are in kcal mol^{-1}

Functional	Grid	G21IP	BSR36	ACONF	SCONF
M06HF	Fine	5.5	4.7	0.47	0.38
	Ultrafine	4.5	1.3	0.11	0.26
M062X	Fine	2.5	3.3	0.23	0.33
	Ultrafine	2.5	3.6	0.24	0.35
M06	Fine	3.1	2.9	0.34	0.56
	Ultrafine	3.1	3.0	0.35	0.58
M06L	Fine	4.5	6.0	0.46	0.39
	Ultrafine	4.5	6.2	0.48	0.39
M052X	Fine	3.2	2.5	0.12	0.20
	Ultrafine	3.2	2.3	0.08	0.19
M05	Fine	3.4	10.3	0.71	1.64
	Ultrafine	3.4	10.4	0.71	1.65
oTPSS	*m*4	4.5	9.2	0.97	0.55
	Reference	4.5	9.2	0.96	0.55
PW6B95	*m*4	2.8	6.6	0.21	0.51
	Reference	2.8	6.7	0.20	0.52
PWPB95	*m*4	2.0	4.1	0.11	0.40
	Reference	2.0	4.1	0.10	0.41

problematic for M06HF. Basically no differences are found between both grids.

Having determined the most difficult sets for M06HF, additional investigations for these sets were then undertaken for all other Minnesota functionals, including M06L, and as a comparison with oTPSS and PWPB95 (Table 3). M062X, M06 and M06L work well for three sets but have differences of 0.3, 0.4 and 0.2 kcal mol^{-1} for BSR36. In contrast to M06HF, the MADs for the larger grid are higher than for the fine grid. The MAD of BSR36 for M052X is by 0.2 kcal mol^{-1} lower with the larger grid; for ACONF it is lowered by 0.04 kcal mol^{-1}. oTPSS and PWPB95 show no significant differences between the two quadrature grids.

4.1.4 Range-separated hybrid functionals. In the recent years, range-separated functionals have gained a lot of interest in the DFT community. Although their success for the description of electronic excitation energies is unquestionable,[145–148] a thorough investigation of their performance for ground state thermochemistry has never been carried out. Herein, three range-separated methods are investigated: CAM-B3LYP-D3, LC-ωPBE-D3 and ωB97X-D. Due to the current implementation of DFT-D and of the latter functional in Gaussian09 it was not possible to carry out the analysis for systems containing elements heavier than xenon. Whenever an element was not implemented in the D-correction, the whole SCF procedure failed. Using the functional with its correct parameterization without D-correction was not possible. Thus, RG6 had to be evaluated for only four systems and HEAVY28 for only 14.

WTMADs for the three range-separated DFs are also shown in Fig. 4 and can be directly compared to those of the other hybrids. For the basic properties none of these range-separated DFs can compete with the best global hybrids. Particularly CAM-B3LYP-D3 and LC-ωPBE-D3 are worse than most of the other global hybrids (WTMAD = 5.5 and 6.0 kcal mol^{-1}). These two functionals perform only moderately for reaction energies. With a WTMAD of 3.3 (CAM-B3LYP-D3) and 3.8 kcal mol^{-1} (LC-ωPBE-D3) they are better than B3LYP-D3 and in the range of the PBE-based global hybrids. ωB97X-D, however, seems to be very promising and has a WTMAD of only 2.5 kcal mol^{-1}, similar to B1B95-D3 and BMK-D3. CAM-B3LYP-D3 is completely inadequate for the description of noncovalent interactions (WTMAD = 1.41 kcal mol^{-1}). LC-ωPBE-D3 and ωB97X-D are better with 0.68 and 0.71 kcal mol^{-1}. ωB97X-D is the best of all 47 functionals for WATER27 (MAD = 1.3 kcal/mol) and S22 (MAD = 0.23 kcal/mol). However, it does not bind the neon dimer. Maybe this could be remedied by using this functional together with the newer dispersion corrections. Note that the reasonable performance of *e.g.* LC-ωPBE-D3 for noncovalent interactions mainly results from the dispersion correction and that the long-range exchange correction does not improve for typical van der Waals systems compared to PBE. The total WTMADs are for CAM-B3LYP-D3 3.5 kcal mol^{-1} and for LC-ωPBE-D3 3.7 kcal mol^{-1}. Thus, no overall improvement over global hybrids is observed. ωB97X-D is with 2.8 kcal mol^{-1} again promising. It is also relatively robust; in three cases it gives the best MAD of all hybrids and in none the worst [part (b) of Fig. 5]. In general, these findings are surprising. We had expected that the range-separated hybrids would outperform the global ones, as the first encompass the latter ones. We doubt that the dispersion correction is the reason for the missing improvement compared to global hybrids, as it is supposed to describe different physics than the long-range exchange part. Applying *e.g.* CAM-B3LYP without a dispersion-correction would not make any sense and does not help in reaching the desired accuracy for inter- and intramolecular noncovalent interactions (see Table S42, ESI†). Furthermore, the results for ωB97X-D are very promising and this functional is also combined with such a correction. ωB97X-D seems to point in the 'right' direction for range-separated DFs and further improvements are encouraged.

4.1.5 Double-hybrid functionals. The results for GMTKN30 with double-hybrid functionals have already been discussed in detail in ref. 31. Herein, just a short summary is given. Fig. 7 shows the WTMADs of the tested DHDFs. In general, DSD-BLYP-D3 and PWPB95-D3 outperform the other DHDFs. However, this comparison happens already on a high level of accuracy. Most MADs for noncovalent interactions lie in the error-range of the theoretical reference method (usually estimated CCSD(T)/CBS).

The DHDFs outperform all other DFs, including the Minnesota functionals (such a direct comparison has rarely been done before). PWPB95-D3 works extraordinarily well for the MB08-165 set with an MAD of 2.5 kcal mol^{-1}, which is in the range similar to that of CCSD(T)/cc-pVQZ (2.6 kcal mol^{-1}).[29] This already indicates the high robustness of this method. For comparison, DSD-BLYP-D3 yields 3.3 kcal mol^{-1}, B2PLYP-D3 and B2GPPLYP-D3 provide values around 4 kcal mol^{-1}. Only XYG3 is worse than the best hybrids with 5.2 kcal mol^{-1}. In general, PWPB95-D3 and DSD-BLYP-D3 can be recommended as accurate general purpose functionals. They have the lowest total WTMADs of all 47 tested functionals (1.6 and 1.5 kcal mol^{-1}, respectively). B2PLYP-D3 (2.0 kcal mol^{-1}) and XYG3 (1.9 kcal mol^{-1}) are much less robust, B2GPPLYP-D3 (1.7 kcal mol^{-1}) lies in between. For further analyses, including their performance for transition metal compounds, see ref. 31. Note that for systems with small HOMO–LUMO gaps a break-down of the perturbative treatment is expected. The fifth-rung random-phase approximation (RPA) based methods seem to be a promising ansatz for these systems (for recent publications of RPA see ref. 149–151 and references therein).

4.1.6 Overall comparison. To have a final overview over all tested DFs, all total WTMADs for the complete GMTKN30 set are shown in part (a) of Fig. 8. The results nicely show the systematically decreasing WTMADs when going from LDAs to DHDFs. Exceptions have already been discussed above. This clear representation of Jacob's Ladder is even better seen when the total WTMADs of all functionals on a rung on this ladder are averaged. Part (b) of Fig. 8 shows these averaged WTMADs. Again we want to emphasize, particularly for

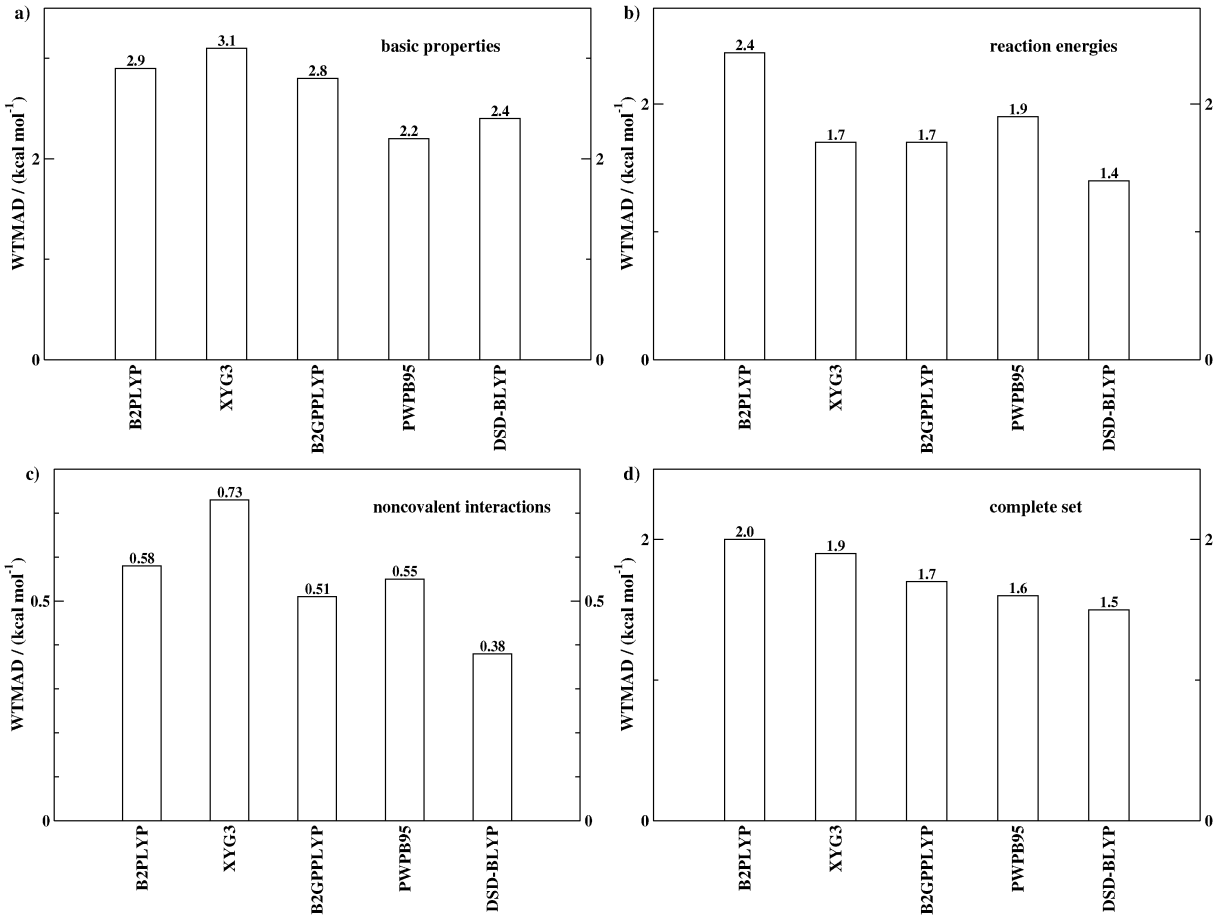

Fig. 7 Weighted total mean absolute deviations (WTMADs) of the double-hybrids in kcal mol^{-1} for the basic properties (a), the reaction energies (b), the noncovalent interactions (c) and the complete GMTKN30 set (d). All values are based on (aug-)def2-QZVP calculations. All functionals were combined with the DFT-D3 correction, except for XYG3. The suffix '-D3' was skipped.

'non-experts' of DFT, that B3LYP-D3 is worse than the average of all hybrid functionals.

While our final analyses were carried out and shortly before submitting this work, Csonka, Perdew and Ruzsinszky published a detailed study on the effect of Fock-exchange admixture on the thermochemical performance of various density functional approximations.[152] In this (and many similar studies) mainly atomization energies are taken as a reasonable measure of performance. We think that atomization is rather unimportant in chemistry and that atomization energy DF benchmarks can cause misleading impressions regarding the description of more realistic chemical reaction energies. The data generated here can be used to shine more light on this issue. In Fig. 9 we plot the WTMAD for all reaction subsets as a function of the MAD for the W4-08 atomization energy subset. As can be seen clearly there is overall only a very weak correlation between the errors for atomizations and other reactions (correlation coefficient of 0.4). Statistically closer relations are found for the modern methods but the relations are still weak. Based on this solid evidence we very strongly recommend not to discuss fundamental issues of DF performance in chemistry solely based on atomization energy benchmarks.

4.2 Basis set dependence

Very recently, we investigated the basis set dependence of B3LYP-D3, PW6B95-D3 and the DHDFs by comparing the total WTMADs of GMTKN30 for the (aug-)def2-QZVP and (aug-)def2-TZVPP bases (for other investigations of DHDFs and their basis set dependence, see *e.g.* ref. 42, 153 and 154).[31] Herein, we will extend this investigation to (meta-)GGAs and will also present results for a double-ζ basis. This is particularly important, as many users prefer *e.g.* the small 6-31G*[155] Pople-basis for their applications. As 6-31G* is not available for all elements included in GMTKN30, we made use of the Ahlrichs' basis sets. However, the results with the herein chosen (aug-)def2-SV(P) valence double-zeta basis are usually comparable to 6-31G*. An investigation on a double-ζ level seems also appropriate for DHDFs that, due to the perturbative contribution, intrinsically require larger basis sets.

Fig. 10 shows the WTMADs of some chosen functionals from different rungs on Jacob's Ladder. The actual values are given in Table S50 in the ESI.† For the (meta-)GGAs and the hybrids it is observed that the WTMADs only increase by about 0.1 to 0.2 kcal mol^{-1} when going from the quadruple- to the triple-ζ level. In the case of oTPSS-D3, it does not change

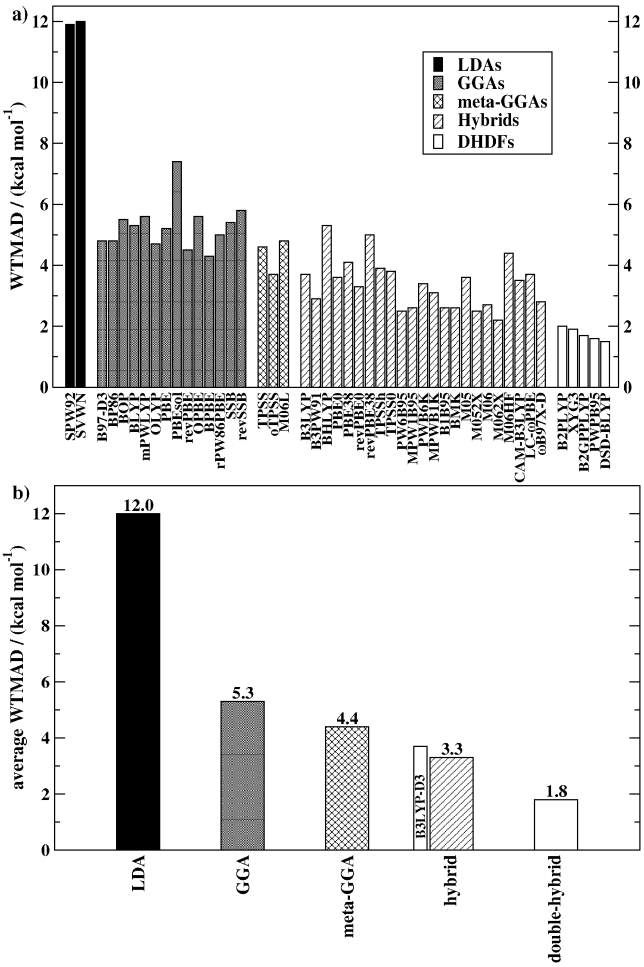

Fig. 8 (a) WTMADs for the complete GMTKN30 for all 47 DFs. The values are based on (aug-)def2-QZVP calculations. All functionals were combined with the DFT-D3 correction, except for the LDAs, ωB97X-D (older DFT-D version), and XYG3. The suffix '-D3' was skipped. (b) WTMADs for all five rungs on Jacob's Ladder, averaged over all functionals that belong to a specific rung.

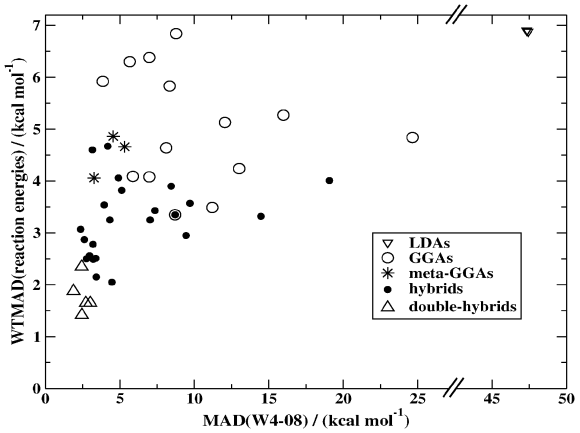

Fig. 9 WTMADs for the reaction energies *versus* MADs of the W4-08 set of total atomization energies.

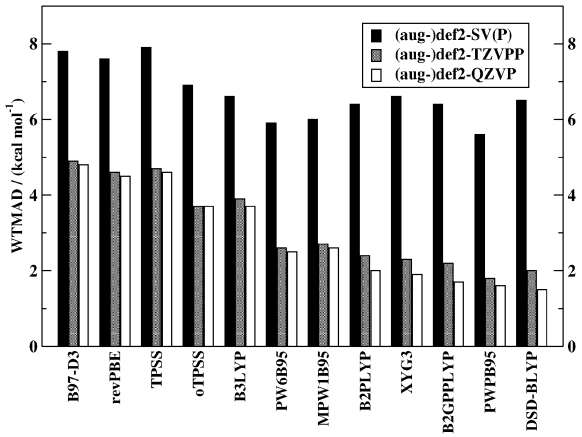

Fig. 10 WTMADs in kcal mol^{-1} for the complete GMTKN30 set for various DFs and three different basis sets. All functionals were combined with the DFT-D3 correction, except for XYG3. The suffix '-D3' was skipped.

at all. This shows that these methods are already at the KS-limit with a large triple-ζ basis. When going to the very small basis, the WTMADs worsen drastically for these methods (by about 3 kcal mol^{-1}). No counter-poise corrections were applied for the noncovalent systems to make the analysis unbiased. A very large BSSE is observed for the water clusters in WATER27. The energies of these clusters are already too low for the pure functional and are even more lowered by adding the DFT-D3 correction (*e.g.* the MAD for TPSS increased from 20.7 to 29.5 kcal mol^{-1} when adding the dispersion correction). However, this has only a minor impact on the total WTMAD. For all three basis sets the relative order of the functionals does not change and is the same as for the large quadruple-ζ basis.

This picture completely changes for the DHDFs. As already discussed, the basis set effect is larger than for the other functional classes. The effect increases with increasing effective perturbative contribution. Usually WTMADs worsen by 0.4 kcal mol^{-1} and they are very close to those of the best hybrids [when going from (aug-)def2-QZVP to (aug-)def2-TZVPP]. The only exception is PWPB95-D3. Its basis set dependence is comparable to those of the (meta-)GGAs and hybrids. It remains the only functional with a WTMAD lower than 2 kcal mol^{-1} (1.8 kcal mol^{-1}). At this triple-ζ level it is thus the best DF for GMTKN30. When going to (aug-)def2-SV(P), the WTMADs worsen by about 4 to 4.5 kcal mol^{-1}. The WTMADs for the DHDFs are now by about 0.5 kcal mol^{-1} worse than for PW6B95-D3 and MPW1B95-D3. Only PWPB95-D3 is still by 0.3 kcal mol^{-1} better. However, in general we cannot recommend application of double-ζ basis sets and only want to illustrate the possible effects here.

4.3 Performance of MP2 methods

As double-hybrids include a perturbative correction and as MP2 (particularly at the CBS limit) is a widely used method, we finally analyze MP2 and its spin-component scaled versions SCS-MP2, S2-MP2 and SOS-MP2 for GMTKN30. Additionally, also the HF method is shortly discussed.

All WTMADs at the CBS limit (based on triple/quadruple-ζ basis set extrapolations with the Ahlrichs type basis sets also used for the DFs) are shown in Fig. 11. For the three sub-classes and for the complete set, HF performs worse than

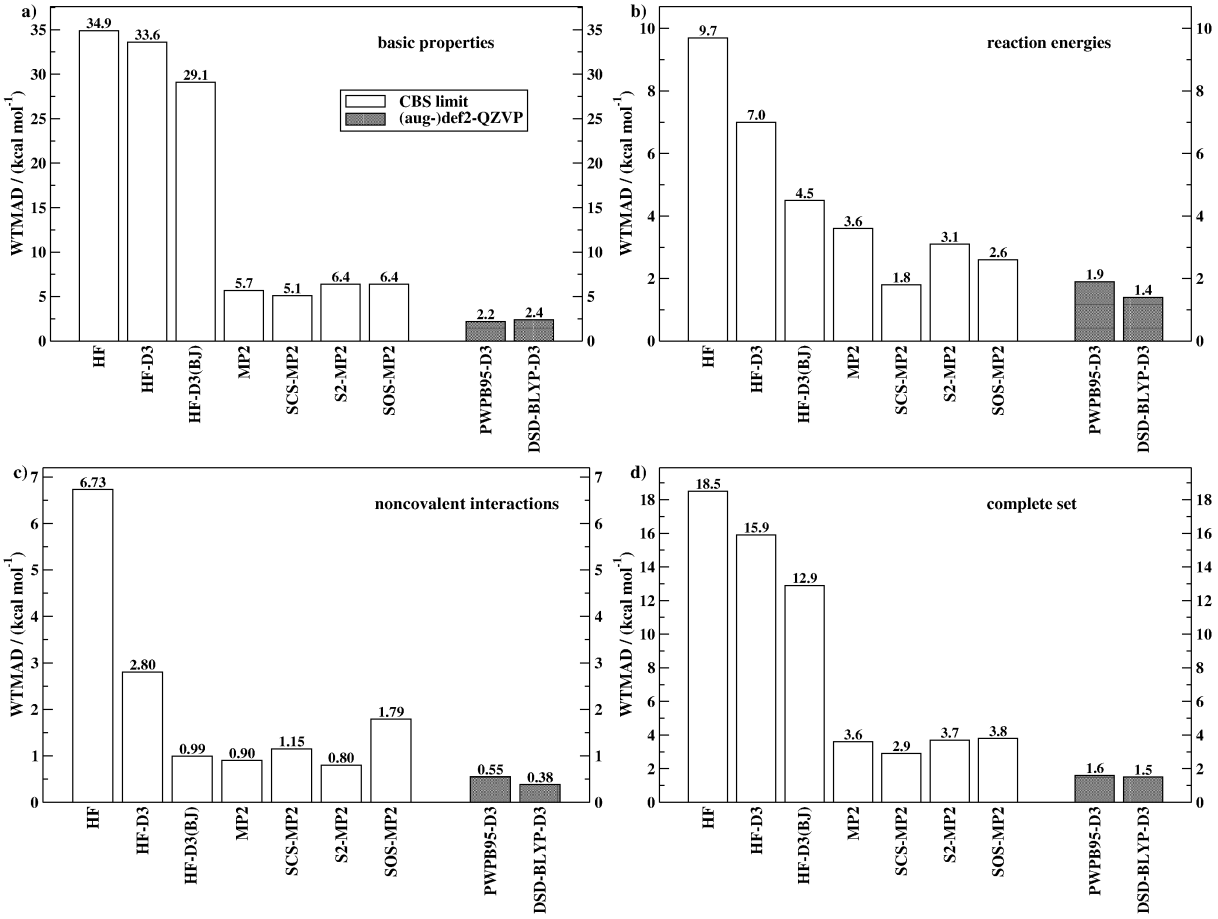

Fig. 11 Weighted total mean absolute deviations (WTMADs) of HF, MP2 variants and two DHDFs in kcal mol^{-1} for the basic properties (a), the reaction energies (b), the noncovalent interactions (c) and the complete GMTKN30 set (d). For the wave function based methods, the values at the CBS limit are shown [(aug-)def2-TZVPP → (aug-)def2-QZVP extrapolations]. The DF values are based on (aug-)def2-QZVP calculations. The DFs were combined with the DFT-D3 correction. Results are also shown for HF-D3 and HF-D3(BJ).

LDA (compare with Fig. 2 in Section 4.1). When dispersion-corrected, HF becomes competitive to GGAs at least for non-covalent interactions (this is particularly the case when using the BJ-damping function, as already discussed elsewhere[40]). The results for all benchmark sets for the two basis sets and the CBS limit are shown in Table S51, ESI.† The results in Table S51 (ESI†) show that the HF-limit seems to be almost reached at the triple-ζ level. The WTMAD for the complete set is 18.5 kcal mol^{-1} for all tested basis set levels without dispersion-correction. It improves to 15.9 kcal mol^{-1} for HF-D3 and to 12.9 kcal mol^{-1} for HF-D3(BJ).

The basis set dependence of MP2 (and this holds also for the spin-component scaled variants) is, as expected, larger and also with the large quadruple-ζ basis the results for some subsets are far away from the CBS limit (see Table S52, ESI†). Particularly for MB08-165, W4-08, and ALK6 the basis set effect is large.

The WTMAD for MP2/CBS for the complete GMTKN30 database is 3.6 kcal mol^{-1}. The results for noncovalent interactions confirm previous findings that MP2 works well for hydrogen bonds (the MAD for WATER27 is 1.4 kcal mol^{-1} at the CBS limit), but usually overestimates dispersion dominated complexes (see *e.g.* ref. 156–159). The MD for S22 for MP2/CBS is 0.97 kcal mol^{-1}, the MAD is 0.98 kcal mol^{-1}. The WTMAD for this sub-class is 0.90 kcal mol^{-1}.

The WTMADs of the three spin-scaled MP2 versions are also shown in Fig. 11 (see also Tables S53–S55, ESI†). SCS-MP2/CBS is only slightly better than MP2/CBS for the basic properties (WTMAD = 5.1 kcal mol^{-1}). For the reaction energies it performs much better with an WTMAD of 1.8 kcal mol^{-1}, which is in the range of the DHDFs. For noncovalent interactions, previous findings are confirmed that SCS-MP2 underestimates hydrogen bonds (MD = −6.15 kcal mol^{-1}; MAD = 6.25 kcal mol^{-1} for WATER27) but corrects for the overestimation of MP2 for dispersion-dominated unsaturated complexes. The MAD for S22 is 0.66 kcal mol^{-1} for SCS-MP2/CBS. Intramolecular-dispersion effects are better described by SCS-MP2 than by MP2 (MAD = 2.5 kcal mol^{-1} *vs.* 4.6 kcal mol^{-1}). The higher WTMAD, compared to MP2 (1.15 kcal mol^{-1} *vs.* 0.90 kcal mol^{-1}), can be explained by the large MAD for WATER27. For the complete set, SCS-MP2/CBS has a WTMAD of 2.9 kcal mol^{-1}, which is the range of the better hybrid DFs.

S2-MP2/CBS is worse than MP2 for the basic properties (WTMAD = 6.4 kcal mol^{-1}). For reaction energies, the WTMAD lies between that of MP2 and SCS-MP2 (3.1 kcal mol^{-1}). Noncovalent interactions are on average better described than by (SCS-)MP2 with a WTMAD of 0.80 kcal mol^{-1}. For the complete benchmark database, though, S2-MP2 is comparable with MP2 (WTMAD = 3.7 kcal mol^{-1}) and cannot compete with SCS-MP2.

SOS-MP2 has higher WTMADs than SCS-MP2. The value at the CBS limit for the complete set is 3.8 kcal mol^{-1}. Head-Gordon and co-workers reported in 2005 that long-range interactions are not fully covered by SOS-MP2 (see ref. 160 for their modified version MOS-MP2). GMTKN30 contains relatively large molecules, which is a possible explanation for the performance of SOS-MP2.

If the size of an investigated system allows a perturbative treatment, we suggest to apply DHDFs instead of MP2 methods. SCS-MP2 can compete with DFT for reaction energies, though, and it is recommended when SIE related problems are expected.

5. Summary and conclusion

A thorough study of various density functionals (DFs) was carried out on the new GMTKN30 database for general main group thermochemistry, kinetics and noncovalent interactions. In total, 47 DFs were investigated: two LDAs, 14 GGAs, three meta-GGAs, 23 hybrids and five double-hybrids. To the best of our knowledge this is the by far largest and most comprehensive DFT benchmark regarding the range of considered systems and density functionals.

Besides double-hybrids, also other modern approaches, *i.e.*, the M05 and M06 classes of functionals and range-separated hybrids, were tested. For almost all functionals, the new DFT-D3 correction was applied; the parameters were taken from previous works or determined for the present study. This allows a consistent evaluation of the performance also including noncovalently bound systems. The calculations were carried out with a large quadruple-ζ basis; additional calculations were done with smaller bases for some functionals. MP2 and three spin-scaled variants are also tested to allow a comparison with the DHDFs. Moreover, the influence of the quadrature grid on the accuracy of the results was tested for hybrid-, double-hybrid-, and meta-GGA functionals. The aim of the various analyses was to organize the plethora of available DFs and to find out which of these can be considered as broadly applicable and robust. As GMTKN30 covers 841 relative energies from a large cross-section of 30 different chemical processes, we think that our findings are based on very solid grounds. We could show that there is no direct correlation between a DF's performance for atomization energies and the chemically more important reaction energies, contrary to what is often assumend in the literature. This confirms the necessity of developing and using large and diverse benchmark sets like GMTKN30. In general, we strongly recommend to use a dispersion correction with any functional, also for reaction energies, as intermolecular dispersion interactions already occur in medium-sized molecules.

The 'simulation' of (usually stabilizing) noncovalent interactions by a basis set superposition error as *e.g.* in many B3LYP/6-31G* applications (see *e.g.* ref. 161) is not recommended. Except for LDAs, which are not consistently applicable to molecular chemistry, the following conclusions can be drawn and recommendations can be given on the various rungs of Jacob's Ladder:

- **GGA-level:** for basic properties, like *e.g.* atomization energies, B97-D3 performs best. It was noted that some empirical functionals yield unusually low or high total electronic energies, which worsens the result for the basic properties dramatically (*e.g.* the SSB-D3 and revSSB-D3 methods yield about $-77E_h$ for water instead of about $-76.4E_h$). Reaction energies are well described by BPBE-D3 and BP86-D3. However, these methods fail in correctly binding small

noncovalently bound systems. BLYP-D3 is the second best of all 47 functionals for the S22 test set (MAD = 0.24 kcal mol^{-1}) and the best GGA for the relative energies of tripeptide conformers (PCONF). However, averaged over all considered systems (inter-, intra-molecular and hydrogen-bonded interactions) it is outperformed by B97-D3 and revPBE-D3. Averaged over the complete GMTKN30 set, B97-D3 and revPBE-D3 can be considered as the currently most robust GGA functionals.

• **Meta-GGA-level:** the three meta-GGAs TPSS-D3, oTPSS-D3 and M06L-D3 were tested. The VSXC functional caused many problems and had to be left out of the analysis. oTPSS-D3 resulted to be the most robust meta-GGA, M06L-D3 the least robust one. Compared to the GGA functionals, TPSS-D3 and M06L-D3 represent no clear improvement. For noncovalent interactions, no meta-GGA is recommended, but GGAs should be preferred if no (double-)hybrid can be applied. Overall, oTPSS-D3 is the only meta-GGA, which is an improvement over GGAs for the description of energetics. It is even comparable to some hybrid functionals. However, previous investigations showed that TPSS yields slightly better geometries.[30] Nevertheless, if geometries are already available (and no (double-)hybrid can be used), we recommend oTPSS-D3 for subsequent single-point calculations.

• **Hybrid-level:** one of the important findings of our study is that B3LYP-D3 is not the overall applicable functional, as many users may still believe. Surprisingly it is even worse than the average hybrid. Particularly for reaction energies it was found to be the worst of all tested 23 hybrids. M062X-D3 and M052X-D3 were the best functionals of the M05 and M06 classes and in fact, M062X-D3 is statistically the best of all hybrids. However, we also observed for all Minnesota functionals SCF-convergence problems, even for simple atomic systems. Furthermore, problems with some non-covalently bound complexes occurred. In accordance to other works, we sometimes found a strong dependence of the results on the quadrature grid that occurs randomly for various test sets and functionals. This makes the M05 and M06 classes of functionals less robust than other hybrid-GGAs and hybrid-meta-GGAs and we strongly recommend careful consideration of this point in hitherto unexplored systems. Another important (and also surprising) finding is that the range separated hybrids CAM-B3LYP-D3 and LC-ωPBE-D3 are on average not better than global ones. ωB97X-D seems promising and it is worthwhile to reparameterize it with the new DFT-D3 correction. As the most robust and very accurate general purpose hybrid-functional, we recommend Zhao and Truhlar's PW6B95, which was published already in 2005.

• **Double-hybrid-level:** on the quadruple- and triple-ζ level the tested double-hybrids outperform all other functionals. DSD-BLYP-D3 is the best functional on the quadruple-ζ level, closely followed by PWPB95-D3, which has, however, a better formal scaling behavior with system size. It is also expected to be useful for cases in which LYP correlation is known to fail. Although B2PLYP-D3, which was the first published DHDF, is outperformed by its successors, the overall performance is not bad and it may still be a reasonable option. Due to the perturbative contribution in double-hybrids, their basis set dependence is higher than for the other DFs. Only PWPB95-D3 has a comparable basis set dependence as (meta-)GGAs and hybrids and is the best DF on a triple-ζ level. In another publication it was also shown that its performance for transition metals is promising.[31] Further investigations on transition metal compounds are currently being carried out in our laboratories. In general, DHDFs cannot be recommended for small double-ζ bases sets. DHDFs overall outperform MP2 and its spin-scaled methods. SCS-MP2 competes with DHDFs for the reaction energies, only. We suggest general usage of DHDFs. SCS-MP2 should be used whenever the self-interaction error plays a role.

We hope that these insights are useful for the development of new DFs, that program developers are encouraged to include the recommended functionals into their codes and that the community of DFT users will benefit from the above given recommendations.

Acknowledgements

This work was carried out within the framework of the SFB 858 ("Synergetische Effekte in der Chemie – Von der Additivität zur Kooperativität") granted by the Deutsche Forschungsgemeinschaft. L. G. was supported with scholarships by the "Fonds der Chemischen Industrie" and the "Research School NRW – Molecules and Materials A Common Design Principle" within the framework of the "NRW Graduate School of Chemistry". We thank Marcel Swart for his assistance regarding the coding of the SSB-D functional. We are also grateful to C. Mück-Lichtenfeld for his technical support.

References

1 P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864–871.
2 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
3 Y. Zhang and W. Yang, *J. Chem. Phys.*, 1998, **109**, 2604–2608.
4 O. Gritsenko, B. Ensing, P. R. T. Schipper and E. J. Baerends, *J. Phys. Chem. A*, 2000, **104**, 8558–8565.
5 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2007, **126**, 104102.
6 P. Mori-Sanchez, A. J. Cohen and W. Yang, *J. Chem. Phys.*, 2006, **125**, 201102.
7 S. Kristyán and P. Pulay, *Chem. Phys. Lett.*, 1994, **229**, 175–180.
8 P. Hobza, J. Sponer and T. Reschel, *J. Comput. Chem.*, 1995, **16**, 1315–1325.
9 J. M. Pérez-Jordá and A. D. Becke, *Chem. Phys. Lett.*, 1995, **233**, 134–137.
10 J. M. Pérez-Jordá, E. San-Fabián and A. J. Pérez-Jiménez, *J. Chem. Phys.*, 1999, **110**, 1916–1920.
11 D. Rappoport, N. R. M. Crawford, F. Furche and K. Burke, in *Computational Inorganic and Bioinorganic Chemistry*, Wiley-VCH, New York, 2009, pp. 159–172.
12 J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari and L. A. Curtiss, *J. Chem. Phys.*, 1989, **90**, 5622–5629.
13 L. A. Curtiss, K. Raghavachari, P. C. Redfern and J. A. Pople, *J. Chem. Phys.*, 1997, **106**, 1063–1079.
14 L. A. Curtiss, K. Raghavachari, P. C. Redfern and J. A. Pople, *J. Chem. Phys.*, 2000, **112**, 7374–7383.
15 L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 2005, **123**, 124107.
16 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 8996–8999.
17 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2006, **110**, 10478–10486.

18 Y. Zhao and D. G. Truhlar, *J. Chem. Theor. Comput.*, 2005, **1**, 415–432.
19 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. C*, 2008, **112**, 6860–6868.
20 Y. Zhao, B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 2715–2719.
21 Y. Zhao, N. González-García and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 2012–2018.
22 Y. Zhao, O. Tishchenko, J. R. Gour, W. Li, J. J. Lutz, P. Piecuch and D. Truhlar, *J. Phys. Chem. A*, 2009, **113**, 5786–5799.
23 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
24 Y. Zhao and D. G. Truhlar, *J. Chem. Theor. Comput.*, 2009, **5**, 324–333.
25 P. Jurečka, J. Sponer, J. Cerny and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.
26 T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall and C. D. Sherrill, *J. Chem. Phys.*, 2010, **132**, 144104.
27 R. Podeszwa, K. Patkowski and K. Szalewicz, *Phys. Chem. Chem. Phys.*, 2010, **12**, 5974–5979.
28 S. Grimme, M. Steinmetz and M. Korth, *J. Org. Chem.*, 2007, **72**, 2118–2126.
29 M. Korth and S. Grimme, *J. Chem. Theor. Comput.*, 2009, **5**, 993–1003.
30 L. Goerigk and S. Grimme, *J. Chem. Theor. Comput.*, 2010, **6**, 107–126.
31 L. Goerigk and S. Grimme, *J. Chem. Theor. Comput.*, 2011, **7**, 291–309.
32 J. P. Perdew, in *Proceedings of the 21st Annual International Symposium on the Electronic Structure of Solids*, ed. P. Ziesche and H. Eschrig, Akademie Verlag, Berlin, 1991, p. 11.
33 A. D. Becke, *J. Chem. Phys.*, 1996, **104**, 1040–1046.
34 S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095–9102.
35 Y. Jung, R. C. Lochan, A. D. Dutoi and M. Head-Gordon, *J. Chem. Phys.*, 2004, **121**, 9793–9802.
36 J. Almlöf, *Chem. Phys. Lett.*, 1991, **181**, 319–320.
37 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
38 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
39 S. Grimme, *Density functional theory with London dispersion corrections*, in *Wiley Interdisciplinary Reviews: Computational Molecular Science (WIREs:CMS)*, Wiley, Hoboken, NJ, 2011.
40 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem*, 2010, DOI: 10.1002/jcc.21759.
41 J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria and G. Csonka, *J. Chem. Phys.*, 2005, **123**, 62201.
42 A. Karton, A. Tarnopolsky, J. F. Lamère, G. C. Schatz and J. M. L. Martin, *J. Phys. Chem. A*, 2008, **112**, 12868–12886.
43 L. A. Curtiss, K. Raghavachari, G. W. Trucks and J. A. Pople, *J. Chem. Phys.*, 1991, **94**, 7221–7230.
44 S. Parthiban and J. M. L. Martin, *J. Chem. Phys.*, 2001, **114**, 6014–6029.
45 V. Guner, K. S. Khuong, A. G. Leach, P. S. Lee, M. D. Bartberger and K. N. Houk, *J. Phys. Chem. A*, 2003, **107**, 11445–11459.
46 D. H. Ess and K. N. Houk, *J. Phys. Chem. A*, 2005, **109**, 9542–9553.
47 S. Grimme, C. Mück-Lichtenfeld, E.-U. Würthwein, A. W. Ehlers, T. P. M. Goumans and K. Lammertsma, *J. Phys. Chem. A*, 2006, **110**, 2583–2586.
48 T. C. Dinadayalane, R. Vijaya, A. Smitha and G. N. Sastry, *J. Phys. Chem. A*, 2002, **106**, 1627–1633.
49 F. Neese, T. Schwabe, S. Kossmann, B. Schirmer and S. Grimme, *J. Chem. Theor. Comput.*, 2009, **5**, 3060–3073.
50 E. R. Johnson, P. Mori-Sánchez, A. J. Cohen and W. Yang, *J. Chem. Phys.*, 2008, **129**, 204112.
51 S. Grimme, H. Kruse, L. Goerigk and G. Erker, *Angew. Chem., Int. Ed.*, 2010, **49**, 1402–1405.
52 R. Huenerbein, B. Schirmer, J. Moellmann and S. Grimme, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6940–6948.
53 M. Piacenza and S. Grimme, *J. Comput. Chem.*, 2004, **25**, 83–99.
54 H. L. Woodcock, H. F. Schaefer III and P. R. Schreiner, *J. Phys. Chem. A*, 2002, **106**, 11923–11931.
55 P. R. Schreiner, A. A. Fokin, R. A. Pascal and A. de Meijere, *Org. Lett.*, 2006, **8**, 3635–3638.
56 C. Lepetit, H. Chermette, M. Gicquel, J.-L. Heully and R. Chauvin, *J. Phys. Chem. A*, 2007, **111**, 136–149.
57 J. S. Lee, *J. Phys. Chem. A*, 2005, **109**, 11927–11932.
58 S. Grimme, *J. Chem. Phys.*, 2006, **124**, 034108.
59 H. Krieg and S. Grimme, *Mol. Phys.*, 2010, **108**, 2655–2666.
60 T. Schwabe and S. Grimme, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3397–3406.
61 S. Grimme, *Angew. Chem., Int. Ed.*, 2006, **45**, 4460–4464.
62 V. S. Bryantsev, M. S. Diallo, A. C. T. van Duin and W. A. Goddard III, *J. Chem. Theor. Comput.*, 2009, **5**, 1016–1026.
63 E. Goll, H.-J. Werner and H. Stoll, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3917–3923.
64 J. F. Ogilvie and F. J. H. Wang, *J. Mol. Struct.*, 1992, **273**, 277–290.
65 J. F. Ogilvie and F. J. H. Wang, *J. Mol. Struct.*, 1993, **291**, 313–322.
66 N. Runeberg and P. Pyykö, *Int. J. Quantum Chem.*, 1998, **66**, 131–140.
67 D. Řeha, H. Valdes, J. Vondrasek, P. Hobza, A. Abu-Riziq, B. Crews and M. S. de Vries, *Chem.–Eur. J.*, 2005, **11**, 6803–6817.
68 D. Gruzman, A. Karton and J. M. L. Martin, *J. Phys. Chem. A*, 2009, **113**, 11974–11983.
69 G. I. Csonka, A. D. French, G. P. Johnson and C. A. Stortz, *J. Chem. Theor. Comput.*, 2009, **5**, 679–692.
70 J. J. Wilke, M. C. Lind, H. F. Schaefer III, A. G. Császár and W. D. Allen, *J. Chem. Theor. Comput.*, 2009, **5**, 1511–1523.
71 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2005, **122**, 154101.
72 E. R. Johnson and A. D. Becke, *J. Chem. Phys.*, 2005, **123**, 024101.
73 E. R. Johnson and A. D. Becke, *J. Chem. Phys.*, 2006, **124**, 174104.
74 Y. Zhao and D. G. Truhlar, *J. Chem. Theor. Comput.*, 2011, DOI: 10.1021/ct1006604, published online.
75 TURBOMOLE: R. Ahlrichs *et al.*, Universität Karlsruhe 2008 and 2009. See http://www.turbomole.com.
76 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165–169.
77 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283–289.
78 C. Hättig and F. Weigend, *J. Chem. Phys.*, 2000, **113**, 5154–5161.
79 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *GAUSSIAN 09 (Revision A.02)*, Gaussian, Inc., Wallingford, CT, 2009.
80 J. C. Slater, *Phys. Rev.*, 1951, **81**, 385–390.
81 J. P. Perdew and Y. Wang, *Phys. Rev. B: Condens. Matter*, 1992, **45**, 13244–13249.
82 S. J. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.
83 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
84 J. P. Perdew, *Phys. Rev. B: Condens. Matter*, 1986, **33**, 8822–8824.
85 J. P. Perdew, *Phys. Rev. B: Condens. Matter*, 1986, **34**, 7406.
86 T. Tsuneda, T. Suzumura and K. Hirao, *J. Chem. Phys.*, 1999, **110**, 10664.
87 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, **37**, 785–789.
88 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200–206.
89 C. Adamo and V. Barone, *J. Chem. Phys.*, 1998, **108**, 664–675.
90 N. C. Handy and A. J. Cohen, *Mol. Phys.*, 2001, **99**, 403–412.

91 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
92 J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou and K. Burke, *Phys. Rev. Lett.*, 2008, **100**, 136406.
93 Y. Zhang and W. Yang, *Phys. Rev. Lett.*, 1998, **80**, 890–890.
94 E. D. Murray, K. Lee and D. C. Langreth, *J. Chem. Theor. Comput.*, 2009, **5**, 2754–2762.
95 M. Swart, M. Solà and F. M. Bickelhaupt, *J. Chem. Phys.*, 2009, **131**, 094103.
96 M. Swart, M. Solà and F. M. Bickelhaupt, *J. Comput. Chem.*, 2010, DOI: 10.1002/jcc.21693, published online.
97 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
98 Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 194101.
99 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
100 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
101 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
102 M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.*, 1999, **110**, 5029–5036.
103 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
104 S. Grimme, *J. Phys. Chem. A*, 2005, **109**, 3067–3077.
105 V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2003, **119**, 12129–12137.
106 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2005, **109**, 5656–5667.
107 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 6908.
108 A. D. Boese and J. M. L. Martin, *J. Chem. Phys.*, 2004, **121**, 3405–3416.
109 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Phys.*, 2005, **123**, 161103.
110 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theor. Comput.*, 2006, **2**, 364–382.
111 Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2007, **110**, 13126–13130.
112 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
113 O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 234109.
114 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
115 S. Kozuch, D. Gruzman and J. M. L. Martin, *J. Phys. Chem. C*, 2010, **114**, 20801–20808.
116 Y. Zhang, X. Xu and W. A. Goddard III, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 4963–4968.
117 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
118 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.
119 B. Metz, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2000, **113**, 2563–2569.
120 K. A. Peterson, D. Figgen, E. Goll, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2003, **119**, 11113–11123.
121 R. Fink, *J. Chem. Phys.*, 2010, **133**, 174113.
122 A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chem. Phys. Lett.*, 1998, **286**, 243–252.
123 F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285–4291.
124 F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143–152.
125 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Acc.*, 1997, **97**, 119–124.
126 S. Grimme, *J. Comput. Chem.*, 2004, **25**, 1463–1473.
127 P. Jurečka, J. Cerny, P. Hobza and D. R. Salahub, *J. Comput. Chem.*, 2007, **28**, 555–569.
128 Prof. Stefan Grimme Research Web Site. http://www.uni-muenster.de/Chemie.oc/grimme/en/index.html.
129 B. L. Hammond, J. W. A. Lester and P. J. Reynolds, *Monte Carlo Methods in Ab initio Quantum Chemistry*, World Scientific, Singapore, 1994.
130 T. V. Voorhis and G. E. Scuseria, *J. Chem. Phys.*, 1998, **109**, 400.
131 G. A. Bacskay, *Chem. Phys.*, 1981, **61**, 385.
132 J. M. L. Martin, C. W. Bauschlicher, Jr and A. Ricca, *Comput. Phys. Commun.*, 2001, **133**, 189.
133 B. N. Papas and H. F. Schaefer III, *THEOCHEM*, 2006, **768**, 175.
134 S. Dressler and W. Thiel, *Chem. Phys. Lett.*, 1997, **273**, 71.
135 V. Termath and J. Sauer, *Chem. Phys. Lett.*, 1996, **255**, 187.
136 E. R. Johnson, R. A. Wolkow and G. A. DiLabio, *Chem. Phys. Lett.*, 2004, **394**, 334.
137 J. Gräfenstein, D. Izotov and D. Cremer, *J. Chem. Phys.*, 2007, **127**, 164113.
138 J. Gräfenstein, D. Izotov and D. Cremer, *J. Chem. Phys.*, 2007, **127**, 214103.
139 E. R. Johnson, A. Becke, C. D. Sherrill and G. A. DiLabio, *J. Chem. Phys.*, 2009, **131**, 034111.
140 E. R. Johnson, A. Becke, C. D. Sherrill and G. A. DiLabio, *J. Chem. Phys.*, 2009, **131**, 034111.
141 L. Fusti-Molnar, X. He, B. Wang, J. Merz and K. M., *J. Chem. Phys.*, 2009, **131**, 065102.
142 C. A. Jiménez-Hoyos, B. G. Janesko and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6621.
143 S. E. Wheeler and K. N. Houk, *J. Chem. Theor. Comput.*, 2010, **6**, 395.
144 P. M. W. Gill, B. G. Johnson and J. A. Pople, *Chem. Phys. Lett.*, 1993, **209**, 506.
145 D. Jacquemin, E. A. Perpète, G. E. Scuseria, I. Ciofini and C. Adamo, *J. Chem. Theor. Comput.*, 2008, **4**, 123–135.
146 D. Jacquemin, E. A. Perpète, I. Ciofini and C. Adamo, *Theor. Chem. Acc.*, 2011, **128**, 127–136.
147 B. Tian, E. S. E. Eriksson and L. A. Eriksson, *J. Chem. Theor. Comput.*, 2010, **6**, 2086–2094.
148 L. Goerigk and S. Grimme, *J. Chem. Phys.*, 2010, **132**, 184103.
149 D. C. Langreth and J. P. Perdew, *Phys. Rev. B: Solid State*, 1977, **15**, 2884–2901.
150 J. Paier, B. G. Janesko, T. M. Henderson, G. E. Scuseria, A. Grüneis and G. Kresse, *J. Chem. Phys.*, 2010, **132**, 094103.
151 H. Eshuis, J. Yarkoni and F. Furche, *J. Chem. Phys.*, 2010, **132**, 234114.
152 G. I. Csonka, J. P. Perdew and A. Ruzsinszky, *J. Chem. Theor. Comput.*, 2010, **6**, 3688–3703.
153 A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman and J. M. L. Martin, *J. Phys. Chem. A*, 2008, **112**, 3–8.
154 D. C. Graham, A. S. Menon, L. Goerigk, S. Grimme and L. Radom, *J. Phys. Chem. A*, 2009, **113**, 9861–9873.
155 W. J. Hehre, R. Ditchfeld and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257.
156 H. Kruse and S. Grimme, *J. Phys. Chem. C*, 2009, **113**, 17006–17010.
157 O. Hübner, A. Glöss and W. Klopper, *J. Phys. Chem. A*, 2004, **108**, 3019.
158 A. Donchev, *J. Chem. Phys.*, 2007, **126**, 124706.
159 M. Rubs and O. Bludský, *ChemPhysChem*, 2009, **10**, 1868–1873.
160 R. C. Lochan, Y. Jung and M. Head-Gordon, *J. Phys. Chem. A*, 2005, **109**, 7598–7605.
161 D. Tzeli, I. D. Petsalakis and G. Theodorakopoulos, *Phys. Chem. Chem. Phys.*, 2011, **13**, 954–965.

