Spin-Component-Scaled Double Hybrids: An Extensive Search for the Best Fifth-Rung Functionals Blending DFT and Perturbation Theory

ARTICLE *in* JOURNAL OF COMPUTATIONAL CHEMISTRY · JULY 2013

Impact Factor: 3.59 · DOI: 10.1002/jcc.23391 · Source: PubMed

2 AUTHORS:

Sebastian Kozuch
Ben-Gurion University of the Negev
47 PUBLICATIONS **1,547** CITATIONS

SEE PROFILE

Jan M L Martin
Weizmann Institute of Science
352 PUBLICATIONS **14,361** CITATIONS

SEE PROFILE

Spin-Component-Scaled Double Hybrids: An Extensive Search for the Best Fifth-Rung Functionals Blending DFT and Perturbation Theory

Sebastian Kozuch,*[a,b] and Jan M.L. Martin*[a,c]

[a] *Department of Organic Chemistry, Weizmann Institute of Science, IL-76100 Rechovot, Israel. E-mail: kozuchs@mail.com , gershom@weizmann.ac.il*

[b] *Present address: Center for Advanced Scientific Computing and Modeling (CASCAM), Department of Chemistry, University of North Texas, Denton, TX 76203-5017, USA.*

[c] *On leave of absence at: Center for Advanced Scientific Computing and Modeling (CASCAM), Department of Chemistry, University of North Texas, Denton, TX 76203-5017, USA.*

ABSTRACT

Following up on an earlier preliminary communication [S. Kozuch and J. M. L. Martin, *Phys. Chem. Chem. Phys.* **13**, 20104-20107 (2011)] we report here in detail on an extensive search for the most accurate spin-component-scaled double hybrid functionals (of which conventional double hybrids are a special case). Such fifth-rung functionals approach the performance of composite ab initio methods like G3 theory at a fraction of their computational cost, and with analytical derivatives available. In this paper, we provide a critical analysis of the variables and components that maximize the accuracy of double hybrids. These include the selection of the exchange and correlation functionals, the coefficients of each component (DFT, exact exchange and perturbative correlation in both the same spin and opposite spin terms), and the addition of an ad-hoc dispersion correction; we have termed these parametrizations "DSD-DFT". Somewhat surprisingly, the quality of DSD-DFT is only mildly dependent on the underlying DFT exchange and correlation components, with even DSD-LDA yielding respectable performance. Simple, nonempirical GGAs appear to work best, while meta-GGAs offer no advantage (with the notable exception of B95c). The best correlation components appear to be, in that order, B95c, P86, and PBEc, while essentially any good GGA exchange yields nearly identical results. Upon further validation with a wider variety of thermochemical, weak interaction, kinetic, and spectroscopic benchmarks, we find that the best functionals are, roughly in that order, DSD-PBEhB95, DSD-PBEP86, DSD-PBEPW91 and DSD-PBEPBE. In addition, DSD-PBEP86 and DSD-PBEPBE can be used without source code modifications in a wider variety of electronic structure codes. Sample job decks for several commonly used such codes are supplied as electronic supporting information.

Introduction

For wavefunction *ab initio* methods, the form of the exact solution to the electronic Schrödinger equation is well-known (full configuration interaction) and well-defined paths for systematically approaching it exist (many-body perturbation theory, limited configuration interaction, coupled cluster theory,…), leaving computational tractability as the main limiting factor.

In density functional theory, on the other hand, the form of the exact exchange-correlation functional is unknown, and possibly unknowable. Different functionals represent different "educated guesses" at this functional. While no systematic route for improving on a given DFT approximation exists, Perdew has identified a hierarchy[1] — known by a Biblical metaphor as "Jacob's Ladder" — based on the types of information entailed in the functionals' construction.

In this metaphor, Hartree theory (no exchange, no correlation) corresponds to the "Earthly Vale of Tears" (or rung zero) and the exact solution to Heaven:

• On the first rung, only the local density ρ is admitted as a variable, leading to various LDA and LSDA approximations.

• On the second rung, the reduced density gradient is introduced as an additional variable, leading to a variety of GGA (generalized gradient approximation) functionals such as BLYP, PBE,…

• On the third rung, higher derivatives of the density enter the picture or so does, more commonly, the kinetic energy density which contains equivalent information. Functionals such as TPSS and M06L are seen here.

• On the fourth rung, the occupied orbitals are also considered. By far the most important category of occupied orbital dependent functionals are the so-called "hybrids" in which the exchange functional is a linear combination of a GGA or meta-GGA part and a Hartree-Fock-like exchange term. (For alternative orbital-dependent approaches, see the review by Kronik and Kümmel.[2]) In hybrid GGAs, the DFT exchange part is a GGA, and such functionals might be considered "imperfect fourth-rung" or "incomplete fourth-rung", unlike hybrid meta-GGAs.

• Finally, on the fifth-rung, virtual (unoccupied) orbitals make their entry. A special case corresponds to double hybrids (DH), in which not only the exchange is a hybrid of (meta)GGA and HF-like exchange, but the correlation is another hybrid, this time of DFT and MP2-like correlation, the latter *in the basis of the Kohn-Sham orbitals*.[3] Such functionals are the subject of the present contribution. (A different approach to combining semilocal DFT correlation with wavefunction-based correlation is the DFT-RPA approach,[4–6] which has recently been reviewed in Ref. 7. It has been shown[8] that RPA can be viewed as an approximation — retaining only ring diagrams — to the CCSD correlation method.)

Double Hybrids

The idea of carrying out perturbation theory in a basis of Kohn-Sham orbitals was first proposed by Görling and Levy almost two decades ago,[9,10] yet the first successful integration of this idea into a practical functional design was Grimme's B2-PLYP[3,11] (see table 1 for a historical overview of double hybrids and some related developments). Functionally, the following steps are carried out: (a) the Kohn-Sham equations are solved for a hybrid GGA with 100 c_x % "exact" exchange and with the correlation damped by a factor 100 $(1–c_c)$ %; (b) An MP2 calculation in the space of these orbitals is carried out, neglecting single excitation ("non-Brillouin") terms; (c) scaled by 100 c_c %, this term is then added to the total energy:

$$E_{xc}=(1-c_x)E_{x,DFT}+c_xE_{x,HF}+(1-c_c)E_{c,DFT}+c_cE_{c,MP2} \quad (1)$$

Leaving aside the form of the exchange and correlation functionals, such a (conventional) double hybrid has two adjustable parameters, namely c_x and c_c. For a more comprehensive summary of the mathematical basis of DHs, see Ref. 12.

Table 1 Some significant DHs and important events, in semi-chronological order (updated from table 2 in Ref. 13 and Ref. 12).

Name	Base Functional	Characteristics	Ref.
GLPT		PT2 based on KS orbitals.	9,10
MC3BB, MC3MPW	B-B95, mPW-PW91	Additivity approximation of separate DFT and MP2 calculations, for which the term "Double Hybrid" was originally coined.	14
RSH+MP2		Range-separated hybrid (RSH) + MP2	15
B2-PLYP	B-LYP	First practical DH. PT2 based on the hybrid DFT KS orbitals at the defined parameters (the "default" method from now on).	3
mPW2-PLYP	mPW-LYP		16
B2K-PLYP, mPW2K-PLYP B2T-PLYP	B-LYP, mPW2K-PLYP	Optimized for kinetics (K) *or* thermochemistry (T)	17
B2-P3LYP, B2-OS3LYP	B-LYP	RSH, and RSH with scaled opposite spin MP2 (SOS-MP2)	18
B2GP-PLYP	B-LYP	"General Purpose" (Opt. for kinetics *and* thermochemistry)	19
XYG3	B-LYP	PT2 based on B3LYP KS orbitals	20
UB2-PLYP, ROB2-PLYP	B-LYP	Reoptimizations for comparative purpose of unrestricted and restricted open shell	21
B2π-PLYP	*B-LYP*	Opt. for π conjugated systems	22
*ω*B97X-2	*ω*B97X	RSH, first viable DH not based on LYP	23
DSD-BLYP	B-LYP	With SCS-MP2 and D2 dispersion correction fitted together with the other parameters (not as an addition over the parametrized DH).	24
PWPB95, PTPSS	PW-B95 TPSS	Reoptimized internal DFT parameters, SOS-PT2 and D3 fitted together with the other parameters.	13
1DH-BLYP	B-LYP	Only one empirical parameter	25
PBE0-DH	PBE-PBE	Non-empirical	26
DuT, PoT	B-LYP	DSD-BLYP optimized for triple zeta basis sets	27
LS1-DH	PBE-PBE	Non-empirical	28
XYGJ-OS	LDA-LYP	SOS-MP2 version of XYG3	29
DSD-PBEP86 (and others)	PBE-P86 (and others)	A search for the best functionals for DSD-DFT	30 & this work
B2-PPW91	B-PW91	Optimized for electric response properties	31
xDH-PBE0	PBE-PBE	Similar to XYGJ-OS	32
PBE0-2	PBE-PBE	Non-empirical	33
DSD-PBEhB95 (and others)	PBEh-B95 (and others)	A search for the best functionals for DSD-DFT, including SOS-MP2, D2, D3 and no dispersion correction options	This work

In B2-PLYP, B88 exchange[34] and LYP correlation[35] were used, and c_x and c_c were adjusted by minimizing the error in the G2/97 set of total atomization energies at the B2-PLYP/[7s4p3d2f] level: the resulting optimized parameters were c_x=0.53 and c_c=0.27. In subsequent work using the W4-08 reference dataset,[19] Martin and coworkers found[19] not only that double hybrid energetics (not surprisingly) exhibit greater basis set sensitivity than lower-rung DFT functionals (and hence c_c=0.27 is a bit high due to compensation for basis set incompleteness error), but that the actual surface exhibits a narrow "canal of best performance" (spanning roughly the c_x=45-75% range) rather than a sharply defined minimum. This leaves a "degree of freedom" to optimize performance for other properties such as reaction barrier heights from the DBH24 set:[36] thus, at c_x=0.65 and c_c=0.36, the B2GP-PLYP functional (GP standing for "general purpose") was obtained, which yields[3,19,24] near-G3 accuracy for both atomization energies and barrier heights.

Both Grimme and Martin considered numerous options for the exchange and correlation parts of the double hybrid. In both cases, it was found that while the choice of the exchange part is less critical (note, e.g., the mPW2PLYP functional[16]) the LYP correlation functional outperformed all others. We shall return to this point later.

The high percentage of exact exchange in B2-PLYP and B2GP-PLYP leads to much reduced self-interaction errors: we note that it was seen earlier (see ref. 37 and references therein) on the fourth-rung that "kinetics-friendly" functionals require much higher c_x than those optimized for thermochemistry only.

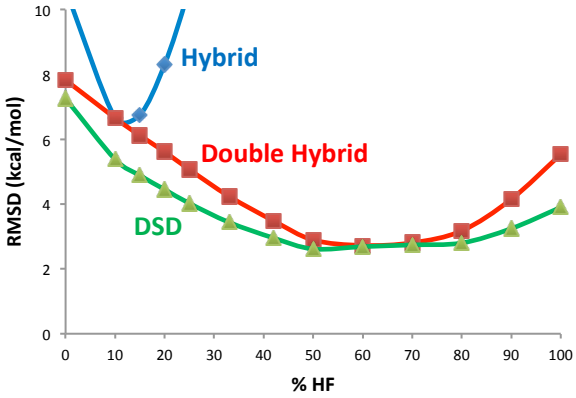

Fig. 1 BxLYP, B2x-PLYP, and DSDx-BLYP performance (RMSD, kcal/mol) for the W4-11 thermochemistry benchmark as a function of the percentage of HF exchange. Note that the seemingly poor performance of BxLYP is degraded by a number of very large errors for pseudohypervalent compounds.

Figure 1 presents the RMSD for the W4-11 thermochemical benchmark[38] of 140 accurate total atomization energies — which span a wide range of degrees of nondynamical correlation — for BxLYP hybrids and B2xPLYP double hybrids as a function of the percentage of Hartree-Fock exchange. (For each double hybrid, c_c was optimized.) As can be seen there, while BxLYP exhibits a sharp minimum, the performance profile of B2x-PLYP is much flatter, and that for the spin-component-scaled DSDx-BLYP[24]

variant — in which separate scaling factors are permitted for same-spin and opposite-spin correlation (see below for more detail) — flatter still. In a very recent paper[39] on diagnostics for nondynamical correlation, it was shown that the slope of the total atomization energy with respect to the fraction of HF exchange is a good measure for static correlation: this explains why it was found in Ref. 24 that, compared to B2GP-PLYP, DSD-BLYP exhibited greatly improved resilience toward nondynamical correlation.

Dispersion Correction

Treatment of non-bonding interactions inevitably requires treatment of dispersion (London) attraction, which varies in importance[40] from about 20% for hydrogen bonds to essentially the whole interaction for rare gas dimers. Such interactions, which are long-range by definition (on the electronic structure scale), appear to be intrinsically beyond the reach of semilocal functionals, although attempts have been made to implicitly treat them by reparametrization of flexible functional forms.[37,41,42] Double-hybrids recover at least part of the dispersion effects through the MP2-like term, but still performance can be improved by introducing an empirical dispersion correction,[43] such as the **D2**[44] and **D3**[45] methods. Most regular DFT of lower rungs in the ladder (i.e. LDA, GGA, meta-GGA and hybrids) have almost null long distance correlation energy, in such a way that van der Waals forces are severely underestimated. MP2 correlation can deal with dispersion forces (and even overestimate them), so double hybrids tend to provide some bonding where other DFT methods fail. However, since the amount of perturbation correction is relatively small, adding a well parametrized dispersion method still benefits the DHs.

By analogy with the 'Jacob's Ladder', Klimeš and Michaelides[46] introduced the "stairway to Heaven" as a hierarchy for dispersion corrections. On step zero, aside from pseudopotential approaches,[46] one finds implicit treatment by parametrization of the semilocal functional (such as is done in the BMK[37], PW6B95[47], and M06[41] functionals): in the present work, this step is represented by the M06 family of functionals.

On step one are simple "molecular mechanics" or "London-type" empirical dispersion corrections of the form:

$$E_D = -\sum_{A>B} s_6 \frac{C_6^{AB}}{R_{AB}^6} f_{damp}\left(R_{AB}\right) \qquad (2)$$

where typically an approximation like $C_6^{AB} \approx \sqrt{C_6^A C_6^B}$ is made, and the C_6^{A} are Lennard-Jones constants independent of the molecular environment and dependent only on the atomic number. The damping function $f_{\mathrm{damp}}(R_{\mathrm{AB}})$, described in more detail below, should reach unity in the large-R_{AB} asymptote and taper off to zero as R_{AB} approaches the covalent region. Grimme's original D2 correction[48] is an example of a "step one" correction, as is Jurečka's similar variant.[49]

More sophisticated variants also may include higher-order r^{-8} or even r^{-10} terms:[43,50]

$$E_D = -\frac{1}{2}\sum_{A\neq B}\sum_{n=6,8,10,\ldots} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{damp}\left(R_{AB}\right) \qquad (3)$$

On step two we find methods in which the $C_{\mathrm{n}}^{\mathrm{AB}}$ can have some dependence on the chemical environment.

Grimme's D3 and D3BJ corrections, in which the C_n^{AB} depend on the valence state, are one example (with two different damping functions); other examples include the Tkatchenko-Scheffler[51] and Becke-Johnson[52] models.

On step three, one finds long-range or "nonlocal" functionals such as VV10,[53] and the vdwDF[54] and vdw-DF2 models;[55] higher steps would consider three-body terms and/or explicit long-range correlation by perturbation theory (such as the double hybrids considered in the present work) or the random phase approximation (RPA).

Both Grimme's D2[48] and Jurečka[49] use the Fermi-type damping function:

$$f_{\mathrm{damp}}(R_{AB}) = \frac{1}{1+\exp(-\gamma(R_{AB}/s_R R_{0,AB}-1))} \tag{4}$$

where $R_{0,AB}$ is the van der Waals distance for the AB pair and s_R is a length scaling parameter. While Grimme uses s_6 as an adjustable parameter, Jurečka et al. set s_6=1 to ensure proper asymptotic behavior, and adjust s_R instead.

Both Grimme's D3 and the Chai–Head-Gordon correction[56] employ a different damping function:

$$f_{\mathrm{damp}}(R_{AB}) = \frac{1}{1+a\left(R_{0,AB}/R_{AB}\right)^{-\alpha}} \tag{5}$$

where a=6, and α=12 (Chai–Head-Gordon) or α=14 (Grimme D3).

Grimme's D3BJ correction[50] employs a different damping function due to Becke and Johnson,[52] which tapers off to a small but finite energy at small distances:

$$E_D^{BJ} = -\frac{s_n}{2}\sum_{A\neq B}\sum_{n=6,8,10,\ldots}\frac{C_n^{AB}}{R_{AB}^n+\left(a_1R_{AB}^0+a_2\right)^n} = -\frac{s_n}{2}\sum_{A\neq B}\sum_{n=6,8,10,\ldots}\frac{\frac{C_n^{AB}}{R_{AB}^n}}{1+\left(\frac{a_1R_{AB}^0+a_2}{R_{AB}}\right)^n} \tag{6}$$

Formally, s_6 should be set equal to one for any method that lacks long-range interactions (most of DFT functionals), but in DHs it must be optimized to complement the same spin MP2 fraction. D2 includes only the r^{-6} term with the "zero" damping,[44] while D3 (a.k.a. D3zero) and D3BJ also include an r^{-8} term to model the medium range region.

We primarily considered the D2 and D3BJ corrections, since the latter is arguably better motivated physically than D3zero, although in practice we found that the two damping functions yielded very similar accuracies.[30] At least with the D3BJ damping, fitted parameters[30,50,57] for double hybrids typically indicate very small s_8 coefficients: presumably the medium-range dispersion is reasonably well covered by the DFT-MP2 tandem. Consequently, we decided to omit the r^{-8} term for D3BJ, thus reducing the number of parameters to be fitted by one. At the same time, it was found that a_1 is typically quite small for double hybrids, and it was concluded that this parameter could safely be set to zero. This leaves us then with just two parameters in D3BJ, namely s_6 and a_2.

Spin-Component Scaled MP2 (SCS-MP2) and Dispersion Corrected, Spin-Component Scaled Double Hybrid DFT (DSD-DFT)

Spin component scaled MP2 was originally proposed on intuitive grounds in an attempt to improve on MP2 without additional cost. Grimme argued[58] that on the one hand, some same-spin correlation is already recovered at first order in a Møller-Plesset series (namely, as Fermi exchange repulsion), while opposite-spin correlation only enters at second order. Considering then that (a) MP2 underestimates the correlation energy of prototype 2-electron systems such as H_2 and He by about 20%; (b) the E_o/E_s (opposite and same spin perturbation energies) ratio is typically in the 3–4 range; (c) imposing the constraint $c_sE_s+c_oE_o=E_s+E_o$, i.e., $c_s=1-(E_o/E_s)(c_o-1)$, Grimme proposed the parameters $c_o=6/5$ and $c_s=1/3$, i.e. $E_{SCS\text{-}MP2}=(6/5)E_o+(1/3)E_s$.

For the W4-11 thermochemical benchmark[38] SCS-MP2 reaches RMSD=4.8 kcal/mol, cutting the original MP2 value (9.5 kcal/mol) basically in half. Minimization of the RMSD with respect to the parameters brings it down to 4.0 kcal/mol, for $c_o=6/5$ and $c_s=1/4$.

Later, reparametrizations for weak interactions such as SCS(MI)MP2[59] ($c_s=1.29$,$c_o=0.40$) and SCS-S66-MP2[60] ($c_s=1.46$,$c_o=0.33$ for the aVQZ basis set) have been proposed. Noteworthy is that for weak interactions, $c_o<c_s$ is preferred, while the opposite holds for covalent bonding: this illustrates that same-spin correlation is the more crucial term for long-range correlation, while opposite-spin correlation is the more important term at short range. This also implies that a universal SCS-MP2 scheme that deals with both covalent and dispersion bonds is not possible,[24] since the former requires $c_o>c_s$, while the latter uses $c_o<c_s$.

Inclusion of spin component scaling into double hybrids yields improved results, especially in conjunction with an empirical dispersion correction.[24] We denoted these functionals **DSD-DFT** (dispersion corrected, spin-component scaled double hybrid DFT). The corresponding equation is:

$$E_{xc}=(1-c_x)E_{x,DFT}+c_xE_{x,HF}+c_cE_{c,DFT}+c_oE_o+c_sE_s+c_DE_D \qquad (7)$$

with E_D being the dispersion correction. It is of interest to note here that in Ref. 24 we found that c_s and c_D were strongly coupled in the optimization and only fairly weakly coupled to the other parameters. Moreover, in a recent study on the pentane conformer surface[61] one of us found that contour plots of the effect of dispersion and of same-spin correlation on the conformer surface look very similar, while both are qualitatively different from opposite-spin correlation. Once again the long-range role of same-spin correlation is illustrated.

A downside of DSD-DFT compared to the simpler double hybrids is the introduction of two additional empirical parameters (compared to two in eqn 1 without dispersion correction, or at least three if a dispersion correction is added). One parameter could be removed by constraining the sum of the coefficients for the short-range correlation contributions to unity, $c_o+c_c\approx1$ — this constraint was found to be approximately, but not exactly, satisfied by the optimum parameters.[24]

Another dimension in the quest for the best DH is the selection of the exchange and correlation functionals. Initially our search was limited to B88 exchange and LYP correlation in light of the earlier findings[3,19] for

"ordinary" double hybrids, although some might argue that the great popularity of B88 and LYP has been a factor. However, in our preliminary communication on the present subject,[30] and in more detail in the present work, we shall show that the "bias" toward BxPLYP was an artifact of not allowing for spin component scaling, that in fact a sizable cohort of DFT exchange and correlation functionals yield good performance, and that several choices surpass the performance of DSD-BLYP by comfortable margins.

In light of the c_s/c_o=5/18 ratio of SCS-MP2, one might wonder what would be the effect of setting c_s=0. Head-Gordon et al.[62] proved that using a Laplace transform, the calculation of E_o by itself can be recast in a form with $O(N^4)$ computational cost scaling, which is an order less steep than the $O(N^5)$ of full MP2. This approach was called **SOS-MP2** (scaled opposite spin MP2), and is almost as accurate as SCS-MP2, as long as van der Waals interactions are not involved. In the present perspective in light of the small c_s/c_o rations obtained for certain exchange-correlation combinations, we will also consider double hybrids with only E_o, which we will term **DOD-DFT**.[30] This DOD-DFT approach was previously applied in the B2OS3LYP,[18] XYGJ-OS[29] and PWPB95[13] double-hybrids.

Exchange and correlation functionals considered for DSD-DFT and DOD-DFT

Following the original B-LYP based DHs,[3,18,19,21,22,24,29] several combinations of functionals have been proposed for double hybrids, including mPW-LYP,[16] B-PW91,[31] PBE-PBE,[26] PW-B95,[13] TPSS,[13] and B97.[23]

In our preliminary communication[30] we considered primarily the S(later), B(ecke88), PBE, mPW, and HSE exchange functionals in combination with the VWN5, LYP, PBE, P86, and PW91 correlation functionals, plus some other individual combinations such as TPSS. Out of that set, the best three performers were very close in performance, and all involved the P86 correlation functional in conjunction with PBE, mPW, and HSE exchange functionals: DSD-PBEP86 was then selected on availability grounds, as PBE is the only one among the three exchange functionals that is available in all major quantum chemistry packages.

For most systems in our training set, the range-separated HSE exchange functional is de facto equivalent to its short-range component, which is PBEh (where "h" stands for "exchange hole model"). As DSD-HSEP86 (which for small molecules effectively corresponds to DSD-PBEhP86) yielded marginally better performance than DSD-PBEP86, we have added PBEh to the list of exchange functionals: in addition we are also considering the B95 correlation functional in the present perspective.

Theoretical Methods

Training Sets

Although there have been some efforts at constructing a truly non-empirical DH,[26,28,63] the accuracy expected for this methods requires fine-tuning the parameters. All the ingredients of double hybrids (DFT, HF and MP2) are in practice approximations, and therefore there is a limit on what theory can predict over the optimal coefficients in the mixture. Because of this, a "well-tempered" training set is still the best way to achieve accuracy for these methods.

We are presently using the same combination of six training sets as in the DSD-BLYP paper and in the preliminary communication.[24,30] The sets, which strive to achieve a good balance of properties, are:

• the "W4-08" of atomization energies of Karton et al.[19] with 101 molecules calculated with the high accuracy W4 method,[64] including 14 high multireference systems;

• the "DBH24" of Zheng et al.[36] of reaction barriers;

• the "Grubbs" set by Zhao et al.[65] for metathesis reactions on a Ru catalyst;

• the "mindless benchmark" (MB) set of Korth et al.[66] consisting on 167 eight-atom random systems and their energy of conversion to diatomic and hydride molecules;

• the "S22" set of van der Waals forces and hydrogen bonds of Jurečka et al.,[67] with revised benchmark energies by Takatani et al.;[68]

• the "Pd" set of oxidative addition reactions at a Pd atom of Quintal et al.[69]

These sets cover main group chemistry (W4-08, DBH24, MB and S22) as well as transition metals (Pd and Grubbs), and thermochemistry (W4-08, Grubbs, MB and Pd) and kinetics (DBH24, Grubbs and Pd), as well as long-range interactions (S22 and Grubbs). The metric used for optimization and evaluation is the arithmetic mean of the six individual RMSD values for the respective training sets. Considering the diversity of chemical systems covered, we have reason to expect that the final functionals will not only be accurate, but robust for a broad variety of chemical situations.

Unless specifically noted otherwise, all reference data for the training and validation sets are obtained from benchmark ab initio calculations rather than experiment.

Basis Sets and Frozen Core MP2

Ab-initio post-HF methods require big basis sets for accuracy. In other words, convergence for basis set completeness is slower for MP2 than for HF or regular DFT. When dealing with double hybrids, the convergence is intermediate between DFT and MP2, thus in principle the size of the basis set should be less critical than with pure MP2. However, the expected accuracy of DHs is also much better than that of MP2, and therefore this issue must be carefully addressed.

As in our previous studies,[24,30] we used quadruple-zeta basis sets where computationally feasible, and triple-zeta basis sets otherwise. Specifically, we employed:

• Aug'-pc3+d[70–74] for W4-08 (with diffuse functions for non-hydrogen atoms and a high exponent d function).

• Aug'-pc3 for DBH24.

• Def2-QZVP[75,76] for MB.

• Def2-QZVPP for Pd.

• Def2-TZVP for Grubbs.

• Def2-TZVP with half counterpoise correction[24,77] for S22.

The size of the basis set affects the optimal parameters of the DH. For instance, the DuT-D3 and PoT-D3 methods[27] were optimized for Dunning's aug′-cc-pVTZ+d and Pople's 6-311+G(3df,2p)+d. In a presumably

futile attempt to obtain a fast but still reliable DH properly parametrized for a small basis set, we also checked Def2-SV(P) as a possible low-cost alternative.

Once having the optimized parameters for each DSD-DFT based on the training set, we tested some functional combinations with several validation sets (see below). The basis sets used in these tests were aug'-pc3+d for HTBH38, NHTBH38, the singlet-triplet excitations, the H-bonds and the van der Waals complexes. For the pericyclic reactions we used aug'-pc2+2d, Def2-TZVP for S66x8, Def2-TZVPP for monoligated Zn(II), and Def2-QZVP for gold complexes. The anthracene dimerization and the alkane conformers were studied with Def2-TZVPP for MP2 and the double hybrids, and pc-2 for the hybrid functionals. For the H-bonds, van der Waals and S66x8 sets, half the counterpoise correction was added as per established procedure[24,77], averaging between overestimated raw and underestimated counterpoise interaction energies. For the frequency calculations, besides the aforementioned def2 basis sets, aug-pc2+2d and aug-pc3+d were applied.

The effect of correlating all electrons in the MP2-like step. instead of just the valence electrons (frozen core), was found to be insignificant, as long as the coefficients are optimized accordingly.[24] Therefore, in this work we continue to use the frozen core option, consisting of all valence electrons, augmented by outer-core s and p core orbitals in the cases of transition metals as well as alkali and alkaline earth atoms heavier than Ar.

Computational Details

All the calculations were done with Gaussian09 RevB.01 or RevC.01.[78] Sample inputs for Gaussian, Molpro,[79] Orca[80] and Q-Chem[81] can be found in the electronic supporting information.

The standard version of Gaussian 09 does not allow for scaling the correlation energy of the B95c correlation functional: this capability was added to the locally modified version of Gaussian 09 by one of us (JMLM). In addition, to resolve a numerical underflow problem in the same-spin correlation term of B95c, the patch detailed in Ref. 82 was applied.

D2 and D3BJ dispersion corrections were obtained with the DFT-D3 standalone program.[45,50,57]

Macroiterations of the optimization were carried out as a 2D grid search in c_x and c_c of Eq. (7): for each (c_x,c_c) pair all electronic structure calculations for the training set were redone. In the D3BJ case, microiterations in the additional nonlinear parameter a_2 were carried out as a one-dimensional scan: for each a_2 value, only the dispersion corrections need to be recalculated. At each (c_x,c_c) or (c_x,c_c,a_2) point, a multivariate linear optimization problem needs to be solved for (c_s,c_o,s_6): this was done using the Solver function of commonly used business spreadsheet software.

In all cases, the optimization yielded a single well-defined minimum.

Results and Discussion: Training Sets

In Table 2 we present the average RMSD of our six training sets for a large number of exchange and correlation functional combinations. The top pane (Table 2a) presents statistics in the absence of any dispersion correction, while results including D2 and D3-BJ dispersion corrections are presented in the following panes (b and c). Table 2d shows the DOD-DFT-D3BJ results, while at the bottom of the table, error statistics for representative non-DSD methods are shown for comparative purposes. In each DH case, all parameters were 'self-consistently' optimized.

From a theoretical perspective, perhaps the most stunning finding is the surprisingly good performance of DSD-LDA – without dispersion it actually edges out DSD-BLYP, while it performs marginally less well than DSD-BLYP if a dispersion correction is included.

In fact, all double hybrids in the table — even clearly mismatched combos like DSD-BVWN5 — do better than SCS-MP2, let alone straight MP2 or "single" hybrids.

In the absence of a dispersion correction, the rule for RMSD appears to be B88>mPW>PBE>≈ PBEh for exchange and PBE>PW91>LYP>P86>B95 for correlation. When a dispersion correction is added, the differences between the exchange functionals become greatly reduced, except that B88 performs slightly worse than the other three. By and large, the ordering for correlation functionals is preserved, however.

In the preliminary communication,[30] we did not consider the B95 correlation functional. Here, we actually find that it— at least for this combination and weighting of training sets — significantly improves RMSD over DSD-PBEP86. The best performer overall is found to be DSD-PBEhB95, followed by DSD-PBEB95 and then DSD-mPWB95.

How does DSD-PBEhB95 compare with DSD-PBEP86, the functional reported as "best" in our preliminary communication? On average, the former DSD-PBEhB95 well outperforms DSD-PBEP86, especially if a dispersion correction is included. Considering the six training sets in detail, we find that DSD-PBEhB95 outperforms DSD-PBEP86 for the Grubbs set (with D3BJ only), for the Pd prototype reactions, and for the "mindless benchmark" (main group artificial structures), while DSD-PBEP86 has the upper hand for S22 and W4-08. While the training sets for Grubbs and Pd are probably too small and idiosyncratic for fine distinctions (note that the "lesser" performance is still excellent by DFT standards), the "mindless benchmark" is large and diverse enough for the difference to be significant.

The B95 correlation functional, which was not studied in our previous communication, is the clear winner. This functional exceeds the accuracy of P86 parametrization (our previous "most favored correlation"[30]), except in the long range interaction and possibly on heavy multireference systems (*vide infra*). The best combination comes with the PBEh exchange (or HSE, its screened-hybrid sibling that brings virtually the same results), but using the standard PBE yields similar performance.

Table 2 Mean error in kcal/mol for the RMSD of all the six training sets. The parametrizations shown here include DSD-DFT (a) without dispersion correction, (b) with D2, (c) with D3-BJ, and (d) some selected combinations of functionals optimized for DOD-DFT with D3-BJ. (e) As a reference, some non-DH methods.[a]

a) DSD-DFT Without Dispersion:

x\c	VWN5	LYP	PBE	P86	PW91	B95	TPSS	Stand Alone:	
S	2.12	2.30	2.61	2.43				B98	2.04
B	3.02	2.18	2.36	2.06	2.33	1.99		BMK	3.07
PBE	2.79	2.04	2.17	1.89	2.13	1.71		τHCTH	2.57
mPW		2.10	2.25	1.96	2.20	1.80		HCTH	2.76
PBEh		2.03	2.12	1.85	2.08	1.65			
O		2.55							
X		2.14				1.90			
TPSS						1.99	2.25		

b) DSD-DFT-D2:

x\c	VWN5	LYP	PBE	P86	PW91	B95	TPSS	Stand Alone:	
S	1.83	2.21	2.51	2.36				B98	2.04
B	2.60	1.80	1.83	1.68	1.78			BMK	3.03
PBE	2.42	1.82	1.75	1.62	1.71	1.43		τHCTH	2.50
mPW		1.81	1.76	1.63	1.72	1.46		HCTH	2.55
PBEh		1.83	1.71	1.60	1.67	1.38			
O		1.94							
X		1.80							
TPSS						1.94	2.09		

c) DSD-DFT-D3BJ:

x\c	VWN5	LYP	PBE	P86	PW91	B95	TPSS	Stand Alone:	
S	1.72	2.12	2.41	2.27				B98	1.99
B	2.49	1.69	1.69	1.55	1.65	1.56		BMK	2.94
PBE	2.29	1.70	1.62	1.50	1.59	1.36		τHCTH	2.36
mPW		1.70	1.63	1.50	1.59	1.40		HCTH	2.40
PBEh		1.73	1.60	1.49	1.57	1.33			
O		1.81							
X		1.69				1.48			
TPSS						1.84	1.94		

d) DOD-DFT-D3BJ

x	S	B	PBE	PBE	PBEh	PBE	PBEh
c	VWN5	LYP	PBE	P86	P86	B95	B95
	1.77	2.55	1.72	1.82	1.84	1.43	1.41

e) Non-DSD methods:

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0	PBE0 +D3BJ
4.16	6.60	4.35	4.65	4.96	5.79	4.81	4.16
TPSSh	M06	M06-2X	M11	ωB97XD	MP2	SCS-MP2	
6.36	4.14	4.59	4.10	3.99	5.58	4.86	

[a] At a referee's request, these are the errors for PWPB95:[13] (a) without D3BJ, 2.45; (b) with D3BJ, 1.90.

Jacob's Sub-Ladder in Double Hybrids

Although the DHs are formally on the fifth rung of Jacob's ladder,[1] there is the possibility of climbing a "sub-ladder" when applying LDA, GGA or meta-GGA functionals into the DH formalism. The effect of climbing this second ladder is far from obvious. The LDA S-VWN5 functional has a very good performance, even better than several GGA options. However, mixing LDA exchange with GGA correlation (or conversely) is found to be counterproductive, owing to error compensation between underestimated exchange and excess correlation in pure LDA. Using a meta-GGA does not bring any improvement in DHs, as can be seen for example by comparison of PBE and TPSS. The notable exception to this rule, of course, is B95, which contains a very simple meta-GGA ingredient (a multiplicative factor in the same-spin contribution involving the difference between true and von Weiszäcker kinetic energy densities) that ensures that no unphysical self-correlation energy will be found for a single electron.

Dispersion Correction and SCS-MP2 Coefficients

Turning now to the effect of the dispersion correction, there is a clear improvement when passing from 'no correction' to D2 and finally to D3BJ. For instance, in the DSD-PBEhB95 method, the results with D2 is (on average) 16% more accurate than without dispersion, with a further 4% improvement seen for D3BJ. The latter has better atomic parameters, but more importantly, we optimized both a_2 and s_6, while for the D2 correction only s_6 was changed. The enhancement coming from the upgrade of the dispersion correction is not large, but it comes at effectively zero computational cost. We do not recommend using the variant without dispersion corrections, unless there is a specific need or one has principled/purist objections to their inclusion. DOD-DFT will by construction entail deterioration in error statistics (see bottom of Table 2) since it lacks same-spin MP2-like correlation. However, for some DSD-DFT functionals, the optimum c_s is so small (especially in the presence of a dispersion correction) that constraining it to zero does not appreciably impact the error statistics (see Table 3): this is, for instance, the case for DOD-LDA and DOD-PBE(h)B95. In most other cases, c_s is significantly different from zero but the c_s/c_o ratio is still tolerably small and hence the deterioration in DOD-DFT is also tolerable: this is found, for instance, for DOD-PBEPBE (with D2: c_s/c_o=0.23, compared to 0.28 for SCS-MP2). The case of DSD-BLYP stands out in that it clearly prefers nearly equal values of the same-spin and opposite-spin coefficients (with D2: c_s/c_o=0.93). The anomalous behavior of DSD-BLYP stems from the imbalanced treatment of same-spin and opposite-spin correlation in LYP:[83] as easily seen from eq. (2) in Ref. 84 the LYP correlation energy for a fully spin-polarized electron gas is actually *zero*. As we pointed out in the communication,[30] our earlier assertion that LYP correlation is essential for good double hybrids was an artifact of imposing c_s=c_o. As a result, a DSD treatment has the potential to improve the accuracy of most double hybrids, but LYP based parametrizations will benefit the least.

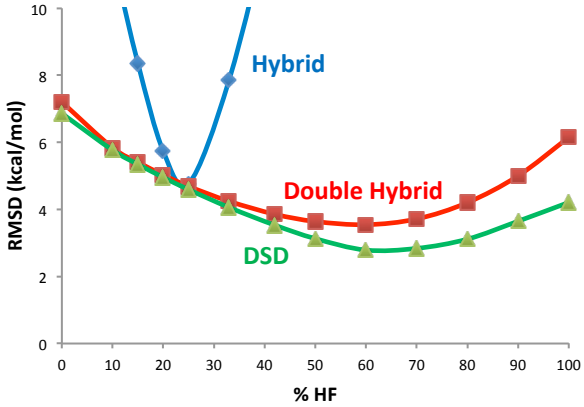

Fig. 2 PBExPBE, PBE2x-PBE, and DSD-PBEPBE performance (RMSD, kcal/mol) for the W4-11 thermochemistry benchmark as a function of the percentage of HF exchange

Coming back to the W4-11 performance profile for B2x-PLYP and DSDx-BLYP in Figure 1, we see that DSDx-BLYP offers no significant advantage over B2x-PLYP in the significant 50-80% exact exchange region. In contrast, the corresponding profile (Fig. 2) for PBE2x-PBE and DSDx-PBE shows a marked advantage of the DSD form basically anywhere from 50% HF exchange on up.

DOD-DFT

Moving from D**S**D-PBEhB95-D3BJ to D**O**D-PBEhB95-D3BJ only worsens the mean error by 6%, still maintaining the leadership as one of the best DHs. The resulting functional bears some similarity to Grimme's PWPB95 method,[13] which t differ as follows: while in PWPB95 the fractions of HF exchange and opposite-spin MP2-like correlation were fixed at 50% and 25%, respectively, and the parameters in the mPW91x and B95c enhancement factors were optimized simultaneously with the other functional parameters, in DSD-PBEh-B95 the two percentages were varied (Eq. 7) while the PBEh exchange and B95 correlation functionals were left unmodified. Comparing the performance of DOD-PBEhB95-D3BJ and PWPB95-D3BJ (Table 3), we find that PWPB95 has a clear advantage for S22 but DOD-PBEhB95-D3BJ (and the functionally nearly identical DSD-PBEhB95-D3BJ) perform markedly better on the DBH24 barrier heights and on MB, leading to a significantly lower mean RMSD for the training set. In view of the reduced CPU time scaling if a Laplace transform algorithm is used,[62] DOD-PBEhB95-D3BJ may be attractive for larger molecules, although the computational cost difference between DSD-DFT and DOD-DFT may become much less of an issue if density fitting/resolution of the identity is employed.[85–88]

When working with DOD-DFT, using a dispersion correction cannot be avoided. As pointed out before, same-spin MP2 is the more crucial term for long-range correlation, and a functional lacking them will not be able to cope with van der Waals forces. It can be seen that when the dispersion correction is omitted from the D**S**D form, c_s significantly grows[24] (from 0.09 to 0.21 for PBEh-B95, see table 3). This negatively affects even the small systems where the dispersion correction is almost negligible (e.g., the W4-08 set) since, as in the SCS-MP2 case, high same-spin MP2 lowers the accuracy in reactions involving covalent bonds.

Table 3 Optimized coefficients and detailed errors (in kcal/mol) for selected functional combinations (see SI for the complete table).[a,b]

		DSD-DFT Without Dispersion					DSD-DFT-D2					DSD-DFT-D3BJ					DOD-DFT-D3BJ				
DFTx		S	B	PBE	PBEh	PBEh	S	B	PBE	PBEh	PBEh	S	B	PBE	PBEh	PBEh	S	B	PBE	PBEh	PBEh
DFTc		VWN5	LYP	PBE	P86	B95	VWN5	LYP	PBE	P86	B95	VWN5	LYP	PBE	P86	B95	VWN5	LYP	PBE	P86	B95
Coeff.	c_c	0.35	0.53	0.48	0.45	0.54	0.34	0.55	0.51	0.46	0.56	0.33	0.54	0.49	0.44	0.55	0.34	0.58	0.54	0.48	0.59
	c_x	0.74	0.75	0.72	0.71	0.69	0.71	0.71	0.66	0.68	0.65	0.72	0.71	0.68	0.69	0.66	0.69	0.65	0.64	0.65	0.63
	c_s	0.31	0.60	0.31	0.36	0.21	0.11	0.43	0.12	0.23	0.09	0.12	0.40	0.13	0.23	0.09					
	c_o	0.55	0.46	0.54	0.50	0.47	0.58	0.46	0.53	0.50	0.46	0.59	0.47	0.55	0.52	0.47	0.58	0.53	0.52	0.54	0.45
	s_6						0.28	0.35	0.42	0.27	0.29	0.46	0.57	0.78	0.46	0.58	0.57	0.96	0.91	0.69	0.67
	a_2											5.6	5.4	6.1	5.6	6.2	5.6	5.1	5.9	5.4	6.0
RMSD	Grubbs	1.26	1.39	1.50	1.11	1.17	0.96	1.10	1.18	0.91	1.00	1.10	1.75	1.22	1.12	0.76	1.44	4.35	1.67	2.14	1.03
	DBH24	1.65	1.41	1.87	1.31	1.28	1.38	1.09	1.33	1.10	1.01	1.46	1.11	1.47	1.15	1.05	1.24	1.16	1.18	0.97	0.88
	S22	0.89	0.81	1.42	0.85	1.05	0.33	0.44	0.48	0.36	0.72	0.34	0.45	0.56	0.42	0.78	0.38	0.66	0.56	0.56	0.82
	Pd	1.25	1.23	1.25	1.37	0.58	1.27	0.95	1.37	1.28	0.45	1.30	1.04	1.17	1.42	0.33	1.37	1.18	1.04	1.17	0.41
	MB	4.58	4.95	4.05	3.98	3.01	4.47	4.51	3.49	3.61	2.59	3.63	3.05	2.53	2.51	2.23	3.85	3.25	2.78	2.77	2.20
	W4-08	3.08	3.27	2.92	2.48	2.81	2.58	2.69	2.64	2.34	2.51	2.48	2.75	2.77	2.35	2.81	2.32	4.67	3.11	3.46	3.12
	non MR	2.93	3.02	2.82	2.21	2.57	2.30	2.25	2.56	2.05	2.20	2.14	2.24	2.63	2.04	2.49	1.90	4.34	2.95	3.21	2.81
	MR	3.86	4.54	3.48	3.73	4.01	3.93	4.57	3.09	3.68	3.94	3.98	4.81	3.54	3.74	4.28	4.05	6.35	3.97	4.74	4.59
Mean		2.12	2.18	2.17	1.85	1.65	1.83	1.80	1.75	1.60	1.38	1.72	1.69	1.62	1.49	1.33	1.77	2.55	1.72	1.84	1.41

[a] $E_{xc}=(1-c_x)E_{x,DFT}+c_xE_{x,HF}+c_cE_{c,DFT}+c_oE_{c,MP2o}+c_sE_{c,MP2s}+E_D$

[b] At a referee's request, here are the errors for PWPB95:[13] (a) without D3BJ: Grubbs=3.15, DBH24=2.24, S22=1.50, Pd=0.93, MB=3.82, W4-08=3.06, Mean=2.45; (b) with D3BJ: Grubbs=1.44, DBH24=2.48, S22=0.32, Pd=0.75, MB=3.36, W4-08=3.04, Mean=1.90.

Table 4 RMSDs in kcal/mol for the training sets for (SCS)MP2 and a number of representative non-DSD functionals

	MP2	SCS-MP2	B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0	PBE0 +D3BJ	TPSSh	M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2
Grubbs	5.46	3.15	5.00	12.22	4.26	6.49	6.47	10.30	4.21	1.84	4.84	2.39	5.80	4.22	3.90	1.39
DBH24	7.26	8.69	3.03	4.80	5.24	5.40	1.68	3.31	4.41	4.70	7.19	2.95	1.13	1.42	2.05	0.98
S22	1.14	1.53	3.77	4.77	0.81	0.55	0.61	5.01	3.32	0.64	4.36	1.17	0.62	0.72	0.33	0.32
Pd	4.20	1.71	1.20	2.10	1.76	1.39	7.87	2.04	0.77	0.85	3.70	5.94	8.34	3.88	2.68	0.51
MB	6.11	8.44	7.61	10.34	8.61	9.09	7.60	9.52	11.05	11.82	11.81	7.84	6.32	7.90	9.90	5.38
W4-08	9.34	5.66	4.34	5.38	5.40	4.98	5.52	4.52	5.12	5.11	6.25	4.56	5.35	6.42	5.08	3.17
non MR	9.40	5.59	3.18	4.59	4.61	4.20	3.27	3.73	4.23	4.25	5.70	3.81	3.22	3.65	2.91	2.12
only MR	8.97	6.03	8.55	8.86	8.83	8.35	12.38	7.81	8.84	8.74	8.95	7.76	11.90	13.70	11.58	6.68
Mean	5.58	4.86	4.16	6.60	4.35	4.65	4.96	5.79	4.81	4.16	6.36	4.14	4.59	4.09	3.99	1.96

Analysis of DSD-PBEh-B95 and DSD-PBEh-P86

From the detailed decomposition of the energy accuracy in table 3, it can be seen that, even though PBEh-B95 is the best overall combination, it is not necessarily the best answer for individual properties. Notable is the reduced accuracy for long-range interactions (set S22) compared to the other DSD-DFT. However, it still outperforms MP2 and rivals the accuracy of simple hybrid functionals with dispersion corrections (without this correction not even the M06 functional is close). Therefore, if long-range interactions are important for

the system to study, DSD-PBEhP86 is a better choice.[30] One might conclude that the B95 functional does not `mesh' with the dispersion correction as well as P86 does.

The W4-08 set of atomization energies is also somewhat less well handled by DSD-PBEh-B95 than by other DSD-DFT, although again (see tables 2 and 3), DSD-PBEhB95 still does much better than other DFT methods or (SCS)MP2 methods. Once more, for atomization energies and especially for heavy multireference systems, DSD-PBEhP86 is "the way to go", a method that approaches the accuracy of much more demanding composite ab-initio schemes.[30] Surprisingly, in conjunction with $B95_c$, the D3BJ correction actually degrades the W4-08 results compared to the D2 method. As before, this may be a problem at the interface of the DFT functional and the dispersion correction. In any case, D3BJ is only slightly better than D2 in average, and although these dispersion methods have been highly beneficial to computational chemistry, there is still more room for improvement.

The strong points of DSD-PBEhB95 are in the transition metals reactions (Grubbs and especially Pd sets) and in the very difficult Mindless Benchmark set for main group reactions. In the DBH24 set for kinetics (i.e. activation energies) the quality is also very high, but all the double hybrids tend to give similar accuracy.

Considering the possibility of having a very fast DH which uses a small basis set,[27] we optimized the coefficients for DSD-PBEhB95 with Def2-SV(P), in the hope that the parametrization may counteract the deficiencies of the basis set. The results were very disappointing. The mean error for all the sets was 5.8 kcal/mol (not better than functionals on lower rungs of Jacob's Ladder).

Summarizing, the DSD-PBEhB95 functional has excellent performance both for main group chemistry and transition metals, including kinetics and thermodynamics. For dispersion forces the results are good, but other combinations (especially PBEh-P86) are more promising. However, all these conclusions are taken from the training sets; we shall turn to validation sets in the next section.

Performance Comparison of DSD-DFT with non-DSD Functionals

Tables 2 and 4 also include training set error statistics for a number of non-DSD functionals. Of all the hybrids considered, only the range-separated hybrid ωB97X-D breaks the 4 kcal/mol barrier (just barely), which is still three times the RMSD of DSD-PBEhB95. Several other functionals have RMSDs near 4 kcal/mol, such as B1B95, B3LYP+D2, PBE0-D3BJ, and M06, as well as the range-separated hybrid M11.[89] (These latter calculations were performed using GAMESS-US.[90]) The edge of DSD double hybrids is most pronounced for the W4-08, MB, and Grubbs subsets: for the DBH24 and Pd subsets, respectively, the M11 and PBE0+D3BJ functionals come close, while ωB97X-D actually barely edges out the DSD functionals for S22.

Compared to the conventional double hybrid B2GP-PLYP-D2 (RMSD=1.96 kcal/mol overall), DSD shaves as much as 1/3 off the overall RMSD. Inspection of the RMSD for the individual subsets reveals the largest improvements are in the "mindless benchmark" (MB) and in the "multireference" subset of W4-08. We have previously noted[24,30] the greater resilience to nondynamical correlation of DSD functionals. (For a more general discussion of DFT and nondynamical correlation, see Fogueri et al.[39] and references therein.)

Results and Discussion: Validation Sets

We tested some of the new combinations of double hybrids on validation sets, which test several chemical properties. Some of these tests were considered in previous studies,[24,30] and others are considered here for the first time with our DHs.

HTBH38 (Hydrogen Transfer Barrier Heights) and NHTBH38 (Non Hydrogen Transfer Barrier Heights) set of Truhlar[91]

These are larger benchmark sets of the two most common types of barrier heights. (The DBH24 set used in the parametrization includes six barriers each from HTBH38 and NHTBH38, plus six nucleophilic substitution barriers and six unimolecular & recombination barriers were added.) It has been observed that dispersion corrections may actually be more of a hindrance than a help in small systems like these,[24] since the parametrization of the damping function is usually done on sets dealing only with long range interactions, rather than at "stretched covalent" distances. In the DSD-DFT case, on the other hand, the dispersion correction was optimized in conjunction with all the other parameters taking into account a mix of covalent and long-range training sets, thus they may be somewhat beneficial. Table 5 shows error statistics for selected double hybrids for heavy atom transfers, while table 6 does the same for hydrogen transfers. There is little to choose between the different DSDs based on hydrogen transfer barriers, as they all perform excellently: perhaps DSD-PBEhB95 has a slight edge over the others in the absence of dispersion corrections, but the difference is really quite small. All handily beat the "kinetics" functional BMK, and have a small but significant edge over B2GP-PLYP, M06-2X and M11. Heavy atoms transfer barrier heights are much more sensitive to the level of theory,[19,91] and here DSD-PBEhB95 is clearly superior over the others, and incidentally is the only one that beats M11, BMK and B2GP-PLYP. DSD-PBEhP86, on the other hand, still puts in a respectable performance if a dispersion correction is added. Somewhat startlingly, the DOD-PBEhB95 fit actually does slightly *better* for NHTBH38 than its DSD congener.

Also somewhat counterintuitively, ωB97X-D actually performs considerably worse for H-transfer barrier heights (RMSD=2.86 kcal/mol) than for the more demanding heavy atom transfer barrier heights (RMSD=1.89 kcal/mol). In the former case a systematic underestimate is seen (MSD=-2.34 kcal/mol) which is essentially absent for the latter case.

Finally, we note that DSD- and DOD-PBEhB95 clearly outperform PWPB95 for both hydrogen and non-hydrogen transfer reactions.

Table 5 RMSD for the NHTBH38 validation set in kcal/mol.

	SVWN5	BLYP	PBEPBE	PBEhP86	PBEhB95
DSD-DFT (no disp.)	2.51	2.39	2.73	2.23	1.84
DSD-DFT-D2	2.21	2.04	2.07	1.97	1.46
DSD-DFT-D3BJ	2.33	2.06	2.29	2.08	1.55
DOD-DFT-D3BJ	2.00	1.82	1.87	1.82	1.28

B1B95	B3LYP	B3LYP+D2	B3LYP+D3BJ	BMK+D2	B972	PBE0	PBE0+D3BJ
3.26	5.54	5.95	6.02	1.79	2.94	4.48	4.71
TPSSh	**M06**	**M06-2X**	**M11**	**ωB97XD**	**B2GP-PLYP+D2**	**PWPB95[a]**	**PWPB95+D3BJ[a]**
7.75	2.83	1.75	1.6	1.89	1.65	2.12	2.32

[a] Extracted from SI of ref. 13.

Table 6 RMSD for the HTBH38 validation set in kcal/mol.

	SVWN5	BLYP	PBEPBE	PBEhP86	PBEhB95
DSD-DFT (no disp.)	1.36	1.28	1.29	1.25	1.05
DSD-DFT-D2	1.26	1.07	1.17	1.18	1.11
DSD-DFT-D3BJ	1.22	1.05	1.06	1.13	1.01
DOD-DFT-D3BJ	1.28	1.32	1.05	1.28	1.09

B1B95	B3LYP	B3LYP+D2	B3LYP+D3BJ	BMK+D2	B972	PBE0	PBE0+D3BJ
3.42	5.09	5.78	5.88	2.14	3.94	4.95	5.41
TPSSh	**M06**	**M06-2X**	**M11**	**ωB97XD**	**B2GP-PLYP+D2**	**PWPB95[a]**	**PWPB95+D3BJ[a]**
7.04	2.41	1.52	1.64	2.86	1.4	1.77	1.96

[a] Extracted from SI of ref. 13.

Hydrogen Bonds[92]

Here we consider a set of 16 hydrogen-bonded dimers, with reference data at the CCSD(T) basis set limit from the work of Boese et al.[92] Error statistics can be found in table 7. As dispersion is a fairly minor component in hydrogen bonds, most DFT functionals yield fairly decent results even without dispersion corrections, and some such as BMK, M06, ωB97XD and M11 achieve RMSDs of as low as 0.4 kcal/mol. DSD-BLYP and DSD-PBEhP86 yield approximately the same performance, while DSD-PBEhB95 and DSD-PBEPBE do even better, with RMSDs just above the 1 kJ/mol (0.24 kcal/mol) mark.

Noteworthy is the similar quality of DSD-DFT on halogen bonds, the "twin sister" of the hydrogen bonds (which also includes charge transfer, electrostatic and polarization interactions, and a small amount of dispersion). In the XB51 benchmark for halogen systems[93] the winner was DSD-PBEhB95 without any dispersion, followed by DSD-PBEP86-D3BJ (the inclusion of dispersion corrections is not necessarily advantageous for all the functionals).

Table 7 RMSD for the H-Bonds validation set in kcal/mol.

	S VWN5	B LYP	PBE PBE	PBEh P86	PBEh B95
DSD-DFT (no disp.)	0.52	0.4	0.24	0.4	0.25
DSD-DFT-D2	0.60	0.43	0.30	0.45	0.25
DSD-DFT-D3BJ	0.57	0.42	0.27	0.42	0.27
DOD-DFT-D3BJ	0.60	0.34	0.28	0.41	0.27

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0
0.95	0.44	0.95	0.84	0.39	0.75	0.89
PBE0 +D3BJ	TPSSh	M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2
1.23	0.61	0.42	0.82	0.49	0.41	0.51

Zhao and Truhlar's set of 28 Nonbonded Complexes[82]

This set consists of 28 dimers bonded by van der Waals forces, being six hydrogen-bonded species, seven charge transfer complexes, six dipole complexes, and nine dispersion-bound ones. Except for M06 and a few other functionals, without a dispersion correction regular DFT yields RMSD comparable to the actual interaction energies themselves. Again we see in table 8 that DSD-DFT does quite well, with the two best performances from DSD-PBEhB95 and DSD-PBEPBE, respectively. Note that even the versions parametrized without dispersion corrections do quite well.

Table 8 RMSD for the Zhao-Truhlar nonbonded validation set in kcal/mol.

	S VWN5	B LYP	PBE PBE	PBEh P86	PBEh B95
DSD-DFT (no disp.)	0.38	0.39	0.35	0.42	0.33
DSD-DFT-D2	0.44	0.40	0.30	0.45	0.26
DSD-DFT-D3BJ	0.51	0.52	0.37	0.53	0.30
DOD-DFT-D3BJ	0.53	0.46	0.38	0.54	0.31

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0
1.20	0.90	0.82	0.95	0.52	1.14	0.68
PBE0 +D3BJ	TPSSh	M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2
1.00	1.04	0.49	0.29	0.56	0.32	0.47

Monoligated Zn(II)[94]

This benchmark (Table 9) consists of 19 neutral and monoanionic ligand binding energies to a bare Zn^{II} atom. This set has proven[94] to be an extremely difficult test for DFT, which systematically overestimates the interaction energies. This trend is usually seen for cationic systems.[24,37] As MP2 does not suffer from this deficiency, [94] it was expected that DH would mitigate the problems of lower-rung DFT functionals,[24] and it

indeed does so even for DSD-LDA! We note that the conventional double hybrid B2GP-PLYP has a rather hard time with this benchmark, and that its error can easily be cut in half by spin-component scaling. Improvements are seen with DSD-PBEPBE and to a lesser extent with DSD-PBEhB95, but we are still some distance from chemical accuracy. In fact, the functionals parametrized without dispersion corrections do better, possibly because they have a larger same-spin correlation coefficient and this problem appears to be sensitive to same-spin correlation. The latter point is illustrated by the good accuracy of straight MP2 compared to the SCS-MP2 version. Accordingly, within the DH methods DOD-DFT version is the worst for each functional combination. Although (especially SCS) double hybrids are usually flexible enough to achieve good compromises, this is one of the specific cases where it is not possible "to have one's cake and eat it too".

Table 9 RMSD for the monoligated Zn(II) validation set in kcal/mol.

	SVWN5	BLYP	PBEPBE	PBEhP86	PBEhB95
DSD-DFT (no disp.)	3.06	3.53	1.61	2.89	2.65
DSD-DFT-D2	3.44	3.89	2.42	3.17	3.27
DSD-DFT-D3BJ	3.87	4.78	2.68	3.66	3.47
DOD-DFT-D3BJ	4.37	5.81	3.11	4.28	3.97

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0
7.49	11.91	13.72	14.22	14.51	3.98	8.63
PBE0 +D3BJ	**TPSSh**	**M06**	**M06-2X**	**M11**	**ωB97XD**	**B2GP-PLYP +D2**
9.75	9.61	8.13	4.47	6.90	4.55	4.18

Pericyclic Reactions[37,95,96]

This is a set of 11 pericyclic reactions proposed by Houk et al. The reference energies for the first eight reactions were also calculated by the W1 method.[19] As the accuracy of DHs is typically higher than of the other DFT methods, we limited ourselves to the systems with W1 benchmark data rather than consider lower-level reference data for additional reactions. Table 10 shows the RMSD for these eight reactions. For this particular problem, DSD-LDA with dispersion actually does surprisingly well, while DSD-PBEhB95 has a clear edge over DSD-PBEhP86 and DSD-PBEPBE. Here, somewhat surprisingly, DOD-PBEhB95 has the best performance of all, although the difference with DSD-PBEhB95 is probably within overlapping uncertainties of the reference values. Likewise, the performance of M11, ωB97X-D and M06-2X can be deemed comparable to the DSDs for this particular criterion.

Table 10 RMSD for the pericyclic reactions validation set in kcal/mol, taking as benchmark the W1 method.

	S VWN5	B LYP	PBE PBE	PBEh P86	PBEh B95
DSD-DFT (no disp.)	2.47	3.62	2.98	3.11	1.91
DSD-DFT-D2	1.87	2.99	2.66	2.63	1.7
DSD-DFT-D3BJ	1.93	3.01	2.48	2.7	1.54
DOD-DFT-D3BJ	1.49	1.34	1.84	1.57	1.22

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	BMK +D2	B972	PBE0
2.70	3.21	2.56	2.43	2.32	2.85	3.43
PBE0 +D3BJ	TPSSh	M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2
4.51	3.41	2.08	1.69	1.99	1.82	2.65

Gold Complexes[97]

Ref. 97 presents a benchmark set of aliphatic η^2 interactions with cationic and neutral gold complexes. As in the original benchmark paper, we found that the dispersion corrections are important, that PBE0 is the most reliable hybrid functional of all the varieties studied, and that double hybrids outperform their "simple hybrid" counterparts. DSD-PBEhB95 is the clear winner here, followed by DSD-PBEPBE and DSD-PBEhP86 (see table 11). Again, we see the somewhat counterintuitive result that DOD actually outperform DSD for this problem. It appears that the DxD-PBEhB95 "combo" works especially well for transition metal complexes.

Table 11 RMSD for the gold complexes validation set in kcal/mol.[(a)]

	S VWN5	B LYP	PBE PBE	PBEh P86	PBEh B95
DSD-DFT (no disp.)	2.57	2.8	2.11	2.5	0.86
DSD-DFT-D2	1.72	1.73	1.21	1.67	0.49
DSD-DFT-D3BJ	2.59	2.79	1.98	2.53	0.68
DOD-DFT-D3BJ	1.93	0.87	1.14	1.35	0.47

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	B972	PBE0
4.17	8.40	4.29	4.29	7.09	2.73
PBE0 +D3BJ	TPSSh	M06	M06-2X	ωB97XD	
1.77	3.87	5.37	8.85	3.47	

[(a)] M11 was omitted due to convergence problems that we were unable to resolve.

Singlet-Triplet Excitations

We compiled experimental and high-level ab initio singlet-triplet excitation energies for a set of 11 small carbenes and silylenes. The detailed values and references are given in the Electronic Supporting Information.[38,98–100] Due to nondynamical correlation effects, even CCSD(T) is not a guaranteed panacea in this case.

It is received wisdom that DFT methods intrinsically have great difficulties with spin splittings, with pure DFT functionals tending to biased towards low spin and hybrid functionals towards high spin,[101,102] thus the results for hybrids tend to be sharply dependent on the percentage of HF exchange. PBE0 performs abysmally, while M06, B3LYP, and especially B3LYP* (which is B3LYP specifically reparametrized for spin splitting[101]) do quite well. DSD-DFTs show an accuracy similar to the well behaved hybrid DFTs, but DSD-PBEhB95 surpasses even the B3LYP* method (see table 12).

Table 12 RMSD for the singlet-triplet excitations validation set in kcal/mol.

	S VWN5	B LYP	PBE PBE	PBEh P86	PBEh B95
DSD-DFT (no disp.)	3.12	3.01	3.46	3.4	1.55
DSD-DFT-D2	2.12	2.16	2.46	2.78	1.24
DSD-DFT-D3BJ	2.14	2.01	2.46	2.66	1.22
DOD-DFT-D3BJ	1.8	2.04	2.1	1.44	1.21

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	B3LYP*[(a)]	B972	BP86	
2.67	2.22	2.22	2.21	1.64	1.85	3.19	
PBE0	PBE0 +D3BJ	TPSSh	M06	M06-2X	M11	ωB97XD	MP2
6.73	6.72	6.70	2.20	3.18	3.10	2.53	3.17

[(a)] Reparametrized B3LYP with 15% exact exchange, optimized for spin-splitting.[101]

S66x8 Set for Dispersion Forces at Different Distances[103,104]

Hobza and coworkers proposed this expanded set of 66 dimers covering a wide variety of interactions of biochemical interest (strong and weak hydrogen bonds, pi-stacking, purely dispersive alkane-alkane interactions,…) which are compressed and stretched to eight multiples of the optimum intermonomer distance (namely 0.9, 0.95, 1.0, 1.05 1.10, 1.20, 1.50, and 2.0). This is a much more stringent test than S22 or S66 by itself, since this requires dispersion corrections to not merely give a fortuitously correct result near equilibrium but to display the correct dissociation behavior. With the D3BJ correction included, B3LYP and PBE0 achieve the accuracy level of MP2, while ωB97X-D surpasses it with its built-in D3zero-like correction, and M11 does so without any correction (see table 13). It was observed that SCS-MI-MP2 can achieve much better accuracy (RMSD=0.32 kcal/mol[103,104]), but at the expense of sharply degraded performance for covalent bonds.

Because of the size of the test set, we only considered our two top choices, namely DSD-PBEhP86 and DSD-PBEhB95. We had observed for the training sets that, although with B95 dispersion forces are well described, P86 provided better accuracy. Herein we see a similar trend, with both correlation functionals having better accuracy than hybrid DFT or MP2, but with P86 as the best option. This functional even outperforms SCS-MI-MP2 (which was specially parametrized for this problem) at a similar computational cost, and without its degradation for covalent bonds. Interestingly, in this case no improvement was observed

on passing from D2 to D3BJ, or the improvement was masked by residual basis set error in the modest basis set. As a probe for the latter, we recalculated the set at the PBE0/QZVP-D3BJ level, finding RMSD=0.47 without counterpoise, 0.42 kcal/mol with, and 0.44 kcal/mol with half-counterpoise, all three RMSD values being noticeably less than 0.54 kcal/mol averaged between counterpoised and non-counterpoised def2-TZVP. The good performance [104] of PWPB95-D3BJ/def2-QZVP (MAD=0.22 kcal/mol, which corresponds to RMSD=0.28 kcal/mol assuming an unbiased normal distribution[38]) is probably within basis set incompleteness error of our DSD-PBEhB95-D3BJ results.

Table 13 RMSD for the S66x8 validation set in kcal/mol, including the RMSD for each distance and the overall error.

	PBEh-P86				PBEh-B95				B3LYP +D3BJ	PBE0 +D3BJ	M06-2X	ωB97XD	MP2
	DSD-DFT (no disp.)	DSD-DFT D2	DSD-DFT D3BJ	DOD-DFT D3BJ	DSD-DFT (no disp.)	DSD-DFT D2	DSD-DFT D3BJ	DOD-DFT D3BJ					
All	0.67	0.25	0.25	0.36	0.68	0.33	0.34	0.39	0.46	0.54	0.41	0.43	0.58
0.90	1.04	0.30	0.34	0.46	0.99	0.56	0.52	0.59	0.58	0.81	0.45	0.66	0.96
0.95	0.91	0.29	0.31	0.45	0.91	0.44	0.45	0.51	0.56	0.68	0.38	0.57	0.80
1.00	0.80	0.28	0.29	0.43	0.82	0.35	0.40	0.46	0.55	0.59	0.39	0.51	0.68
1.05	0.70	0.27	0.28	0.40	0.72	0.30	0.36	0.42	0.53	0.53	0.47	0.44	0.58
1.10	0.61	0.27	0.26	0.37	0.64	0.27	0.32	0.37	0.50	0.50	0.55	0.37	0.50
1.25	0.41	0.23	0.21	0.29	0.43	0.21	0.24	0.28	0.38	0.43	0.53	0.26	0.32
1.50	0.23	0.15	0.15	0.18	0.22	0.13	0.15	0.17	0.22	0.30	0.25	0.20	0.16
2.00	0.07	0.06	0.06	0.07	0.08	0.06	0.07	0.07	0.09	0.12	0.12	0.10	0.07

Harmonic Frequencies

In the original DSD-BLYP paper, we considered the performance of double hybrids for the HFREQ27 set of harmonic frequencies for 27 small molecules (see ref. 24). We found that the optimum B2x-PLYP double hybrid was located at 43% HF-like exchange and 18% MP2-like correlation, yielding RMSD=18.6 cm^{-1}; however, its performance for thermochemistry and especially kinetics is unacceptably poor. B2GP-PLYP and DSD-BLYP yielded much worse RMSDs of 29.8 and 30.9 cm^{-1}, respectively, compared to 33.9 for B3LYP and 44.0 cm^{-1} for MP2. Using scaling factors (given in parentheses), this can be reduced to 17.7 (0.989) for B2GP-PLYP and 22.4 (0.990) for DSD-BLYP.

Arguably, since the DSD double hybrids require at least the computational toil of SCS-MP2, they should at least outperform SCS-MP2 to justify their existence for frequencies. As it happens, even DSD-LDA handily meets this criterion, *without* the assistance of scale factors: its RMSD for the HFREQ27 set with the Def2-QZVP basis set set is 16.9 cm^{-1}, which is half the corresponding SCS-MP2 value (see table 14). Even DSD-BLYP bests SCS-MP2, as does DSD-PBEhB95, while DSD-BP86, DSD-mPWP86 and DSD-PBEPBE all meet or exceed the performance of DSD-LDA. In fact, all of them meet or exceed the performance of CCSD/apc2+2d. Then again, so does SCS-MP2; only when we introduce a constant scaling factor (to correct systematic bias) does CCSD exhibit a clear advantage over SCS-MP2, especially when the problematic F_2 molecule is removed as an outlier.

With such a scaling factor, even B2GP-PLYP puts in a performance similar to CCSD, and several

functionals are able to exceed this, notably DSD-PBEP86 and DSD-PBEPBE. With the Def2-QZVP basis set, these two functionals reach RMSDs for the scaled frequencies of about 13 cm^{-1}, which is in shouting distance of, but not quite reaching, the sub-10 cm^{-1} accuracy of unscaled CCSD(T). (Scaling offers no advantage for CCSD(T), as expected). If in addition F_2 is removed as an outlier, both DSD-PBEPBE and DSD-PBEP86 do indeed skirt the edge of 10 cm^{-1}.

Table 14: Error statistics (cm^{-1}) over the HFREQ27 dataset for a variety of methods.[a]

	MSD unscaled	RMSD unscaled	Scaling factor	RMSD scaled	w/o F_2
DSD-SVWN5 -D2/Def2-QZVP	9.6	16.9	0.9959	14.4	11.50
DSD-BLYP -D2/Def2-QZVP	18.9	27.3	0.9901	16.6	13.8
DSD-BP86 -D2/Def2-QZVP	8.0	16.8	0.9959	14.2	11.5
DSD-PBEPBE -D2/apc2+2d	10.9	19.2	0.9941	14.3	12.6
DSD-PBEPBE -noD/apc2+2d	20.8	30.1	0.9889	17.6	15.1
DOD-PBEPBE -D3BJ/apc2+2d	12.9	18.9	0.9935	12.6	10.7
DSD-PBEPBE -D2/apc3+d	9.0	16.5	0.9954	13.1	10.8
DSD-PBEPBE -D2/Def2-TZVPP	9.6	17.0	0.9952	13.4	11.1
DSD-PBEPBE -D2/Def2-ATZVPP	5.6	15.3	0.9971	13.9	12.0
DSD-PBEPBE -D2/Def2-QZVP	8.9	16.2	0.9955	12.8	10.7
DSD-PBEP86 -D2/apc2+2d	9.5	18.8	0.9950	15.3	13.0
Ditto including D2 derivs	7.6	17.9	0.9959	15.5	13.2
DSD-PBEP86 -D2/apc3+d	7.6	16.1	0.9963	13.9	11.0
DSD-PBEP86 -D2/Def2-QZVP	7.6	15.8	0.9962	13.5	10.8
DSD-PBEhB95 -D2/Def2-QZVP	18.1	24.0	0.9913	14.6	12.2
DSD-PBEhB95 -noD/Def2-QZVP	23.6	30.1	0.9883	15.6	12.5
DOD-PBEhB95 -D3BJ/apc2+2d	20.4	26.8	0.9899	15.2	13.4
DSD-PBEhB95 -D3BJ/apc2+2d	20.3	27.0	0.9898	15.0	12.7
SCS-MP2 /cc-pV(T+d)Z	14.2	31.3	0.9920	26.1	25.8
SCS-MP2 /cc-pV(Q+d)Z	15.2	30.5	0.9915	24.1	23.9
CCSD/cc-pV(T+d)Z	24.9	31.9	0.9897	22.1	20.5
CCSD(T) /cc-pV(T+d)Z	12.7	12.7	1.0005	12.6	12.7
CCSD(T) /cc-pV(Q+d)Z [c]	-0.6	8.4	1.000	8.4	8.4
MP2/apc3+d [b]		44.0	0.988	35.6	
B2GP-PLYP /apc3+d [b]		29.8	0.989	17.7	
B2-PLYP/apc3+d [b]		18.6	0.997	17.4	
B3LYP/apc3+d [b]		33.9	1.004	32.6	
M06/apc3+d [b]		47.5	0.996	46.6	

[a] Except where indicated otherwise, derivatives of the Grimme correction itself were not included. For the D3BJ case, the effect of their inclusion is insignificant at the accuracy scale of interest. In the D2 case, these terms are artificially large due to the shape of the cutoff function near turnover (see fig. 3).

[b] Taken from Ref. 24.

[c] Compiled from published data: C_2H_2;[105] C_2H_4;[106] C_2H_6;[107] CH_3NH_2;[108] CH_4;[109] Cl_2;[110] ClCN (only cc-pVTZ);[106] ClF, HCl PW; CO, CO_2, F_2, H_2CO, H_2, H_2O, HCN, HF, N_2, N_2O, NH_3;[111] CS, OCS, H_2S;[112] HNO;[113] HOCl;[114] PH_3;[115] SO, SO_2.[116]

The scaling factors for DSD-PBEPBE and DSD-PBEP86 hover around 0.995, while that for DSD-PBEhB95 is about 0.99. The latter is definitely the weakest of the three when it comes to frequencies.

Higher same-spin correlation, as present in the "noD" variants, is associated with worse HFREQ27 error statistics, while DOD variants (that omit same-spin MP2-like correlation altogether) actually perform marginally better.

Another remark is due regarding dispersion corrections. In the harmonic frequency calculations using Gaussian, we neglected the second derivatives of the dispersion correction itself. We were however able to consider their effect through independent recalculation using the latest version of the ORCA program system.[80] At the resolution of interest to us, those for the D3BJ correction can indeed be considered negligible, while those for the D2 correction can reach several cm^{-1} especially on hydrogen stretches. Consideration of the distance profile of these corrections reveals a sharp peak for the D2 correction near the van der Waals distance; said peak is absent in the D3BJ correction (see fig. 3).

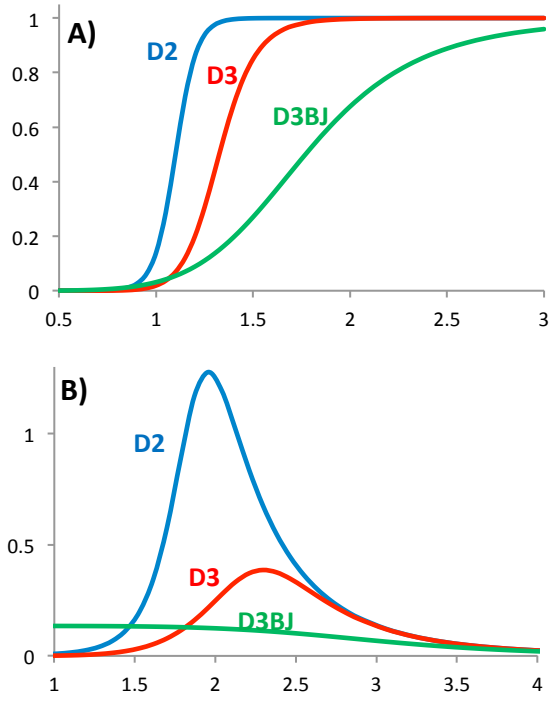

Fig. 3 (A) Damping functions and (B) the r^{-6} term of E_{disp}/C_6^{AB} for different damping functions. (The sum of the van der Waals radii in D3BJ was arbitrarily set to 3.4 Å, twice the value for carbon.) The sharpness of the D2 method worsens the accuracy of interactions occurring close to the van der Waals distances. Units are dimensionless.

Concerning basis set convergence of frequencies with the double hybrids, we have explored this for the DSD-PBEPBE functional. Def2-QZVP and apc3+d yield results of essentially the same quality (the former basis set being smaller), while Def2-TZVPP is acceptably close to convergence. The somewhat worse performance of apc2+2d is actually almost exclusively due to the hydrogen basis set. Detailed reference data and results for additional levels of theory are given in the SI.

As a final remark on the frequencies, the good performance of DSD-PBEP86 and DSD-PBEPBE for HFREQ27, which was not part of their "training set" in any way, pleads in favor of their "universality".

Anthracene Dimerization Energy

Anthracene can undergo [4+4] cycloadditive photodimerization to a D_{2h} "sandwich". Both covalent bonding (between the central rings) and dispersion forces (between the outer rings) exist, and on balance the sandwich dimer is 10.7±1 kcal/mol more stable than two separated monomers. (For a detailed study of the dimerization mechanism see Ref. 117). In the preliminary communication, we reported a CCSD(T)/cc-pVTZ(f/p) dimerization energy of 11.15 kcal/mol; using the same basis set, we obtained 20.59 kcal/mol for MP2 (which is clearly inadequate) but 13.13 kcal/mol at the SCS-MP2 level. Using the cc-pVQZ basis set, we found 20.21 kcal/mol MP2 and 12.55 kcal/mol for CCSD. Applying additivity approximations to obtain an estimated CCSD(T)/cc-pVQZ dimerization energy, we finally obtained 10.77 and 10.57 kcal/mol via MP2 and SCS-MP2 basis set extrapolations, respectively, or a proposed best estimate of 10.7±1 kcal/mol (with a conservative error bar), somewhat higher than an earlier best value[118] in which only a cc-pVDZ basis set was invoked for post-MP2 correlation effects.

Perhaps the most pleasing finding from the table is that all DSD-double hybrids (except perhaps for DSD-BLYP without dispersion) do at least a decent job of reproducing the reference value. This is even true for DSD-SVWN! DSD-PBEP86 and DSD-PBEhB95 both do very well, and discrimination between them is not possible on the basis of this problem (see table 15). ωB97X-D yields an excellent 10.17 kcal/mol, while M11 falls short at 8.19 kcal/mol.

Table 15 Calculated anthracene dimerization energy (kcal/mol).[a]

	S VWN5	B LYP	PBE PBE	PBE P86	PBEh P86	PBEh B95
DSD-DFT (no disp.)	9.03	7.04	9.90	9.06	8.58	8.36
DSD-DFT-D2	12.09	10.05	13.89	11.66	11.02	11.07
DSD-DFT-D3BJ	12.82	10.65	14.91	12.00	11.82	11.85
DOD-DFT-D3BJ	12.43	9.96	14.54	9.50	11.78	11.60

B3LYP	B3LYP +D2	B3LYP +D3BJ	PBE0	PBE0 +D3BJ	TPSSh	B972	M06
-23.19	-1.43	-.2.23	-4.61	8.70	-15.17	-15.44	-3.04
M06-2X	M11	ωB97XD	B2GP-PLYP +D2	MP2	SCS-MP2	CCSD/VTZ (no d on H)	Best estimate (see text)
3.99	8.19	10.17	7.39	20.21	12.55	11.15	10.7±1

[a] The pc-2 and def2-TZVPP basis sets were used for the hybrid and double hybrid functionals, respectively.

Alkane and Melatonin Conformer Energies, plus n-Pentane Torsion Surface

Our test set here is identical to the ACONF subset of GMTKN30,[119] and consists of relative energies of the twelve conformers of n-hexane, the four of n-pentane, and the trans/gauche equilibrium in n-butane. The reference data, as in GMTKN30, were taken from our own earlier paper.[120] It was shown there that dispersion is an essential component of these equilibria[121] (although the role of dispersion appears to have been first suggested as early as a 1970 paper by Schleyer[122]), and, conversely, that conformer equilibria are a sensitive test for the proper description of medium-range dispersion effects.

The results are summarized in table 16. Additional pieces of the puzzle can be found in recent papers on the global torsion surface of pentane[61] and on the conformer manifold of melatonin.[123]

Analysis revealed[120] that the most important interactions for the conformer equilibria of alkanes are right in the region around the sum of the van der Waals radii (i.e., in the middle of the D2 cutoff function hysteresis). This makes the shape of the cutoff function quite important. We found earlier[120] that D2 tends to overestimate the correction, and here we do see that for all four functional pairs in the table, D2 performs worse than no dispersion correction at all. For all four functionals, D3BJ offers an improvement over D2, sometimes quite drastically so, as in DSD-PBEPBE-D3BJ and DSD-PBEP86-D3BJ. DSD-PBEhB95-noD is the exception, as it involves no dispersion correction at all yet yields an astonishingly low RMSD of 0.05 kcal/mol.

The pentane torsion surface (which spans an energy range of 16.66 kcal/mol[61]) is in some ways a more demanding test. It was found there (compared to basis set limit CCSD(T)-F12 data) that PBE0-D3BJ performs best of all conventional DFT functionals considered (RMSD=0.07 kcal/mol over all 343 unique points); both B2GP-PLYP-D2 and the DSD-DFT-D3BJ functionals match or surpass that performance, with RMSD=0.03 kcal/mol for DSD-PBEPBE-D3BJ being almost too good to be true. DSD-PBEhB95-noD has the best performance of a functional without dispersion corrections (RMSD=0.09 kcal/mol).

Conformer equilibria in melatonin,[123] on the other hand, are to some extent driven by internal weak

hydrogen bonds. Here DSD-PBEhB95 has a harder time, while DSD-PBEP86-D3BJ (RMSD=0.15 kcal/mol) actually surpasses the already excellent performance of DSD-PBEPBE-D3BJ (RMSD=0.27 kcal/mol).

Table 16 RMSD errors (kcal/mol) for (A) alkane conformer energies (ACONF benchmark),[120] (B) pentane torsion surface,[61] and (C) melatonin conformers.[123]

A) ACONF

	B LYP	PBE PBE	PBE P86	PBEh B95
DSD-DFT (no disp.)	0.20	0.25	0.16	0.05
DSD-DFT-D2	0.25	0.31	0.24	0.42
DSD-DFT-D3BJ	0.24	0.08	0.09	0.26
DOD-DFT-D3BJ	0.20	0.11	0.17	0.31

B1B95	B3LYP	B3LYP +D2	B3LYP +D3BJ	B972	PBE0	PBE0 +D3BJ	TPSSh
0.43	1.05	0.68	0.08	1.14	0.7	0.07	0.8
M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2	PWPB95[a]	PWPB95 +D3BJ[a]	MP2
0.37	0.27	0.65	0.3	0.22	0.12	0.22	0.21

[a] Extracted from SI of ref. 13.

B) Pentane torsion surface

	B LYP	PBE PBE	PBE P86	PBEh B95
DSD-DFT (no disp.)	0.26	0.23	0.20	0.09
DSD-DFT-D2	0.08	0.12	0.07	0.17
DSD-DFT-D3BJ	0.11	0.03	0.08	0.08

B3LYP	B3LYP +D2	B3LYP +D3BJ	PBE0	PBE0 +D2	PBE0 +D3BJ	
0.43	0.49	0.17	0.30	0.21	0.07	
M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2	PWPB95 +D3BJ	MP2
0.20	0.13	0.49	0.17	0.07	0.14	0.18

C) Melatonin conformers

	B LYP	PBE PBE	PBE P86	PBE B95
DSD-DFT (no disp.)		0.63	0.41	0.56
DSD-DFT-D2		0.25	0.21	0.33
DSD-DFT-D3BJ	0.26	0.27	0.15	0.33

B3LYP	B3LYP +D2	B3LYP +D3BJ	PBE0	PBE0 +D2	PBE0 +D3BJ	
3.10	0.48	0.33	2.16	0.73	0.39	
M06	M06-2X	M11	ωB97XD	B2GP-PLYP +D2	PWPB95 +D3BJ	MP2
0.52	0.29	0.77	0.31	0.17	0.30	0.91

Among conventional functionals, the performance of PBE0-D3BJ is very satisfying for all three problems (particularly the pentane surface and ACONF), but both ωB97X-D and M06-2X perform reasonably well on average. M11, on the other hand, yields comparatively poor RMSDs on all three tests (0.49 kcal/mol for the pentane surface, 0.77 kcal/mol for the 52 conformers of melatonin, and 0.65 kcal/mol for ACONF). Detailed inspection of the pentane results reveals that M11 (like M06, as was shown earlier[120]) inverts the ordering of the trans-gauche and gauche-gauche conformers, as well as yields an unrealistically small trans-gauche difference of just 0.18 kcal/mol (benchmark value[61] 0.58 kcal/mol) due to gauche overstabilization. (Similar behavior is seen for the hexane conformers.)

Isomers of C_8H_8

In ref. 124, Karton and Martin carried out an extensive benchmark study of relative energies of 45 isomers of C_8H_8, spanning a wide range of organic bonding situations. Reference energies were obtained at the W1-F12 level.[125]

The best performance of any DFT method there was put in by the DSD-PBEP86 double hybrid, with an RMSD for relative energies of just 1.2 kcal/mol, which outperforms CCSD and approaches the performance of the CBS-QB3 composite method.

In this case, dispersion is a small player relative to the energy differences being considered: RMSD for PBE0-D2 relative to PBE0, for example, only changes from 3.5 to 3.7 kcal/mol, with a prefactor s_6 over twice the size of that in DSD-PBEP86-D2.

The DSD-PBEhB95 functional was not considered there: in the present work, we find RMSD=2.40 kcal/mol for DSD-PBEhB95-D2, which goes up to 2.61 kcal/mol for DSD-PBEhB95-noD.

Some Further Observations about Double Hybrids' Coefficients

There have been recent efforts by Adamo, Savin and coworkers to try and achieve "nonempirical" double hybrids.[26,28] Based on density scaling arguments they found that (eqn 7 of ref. 28)

$$E^{\lambda} = \langle \phi | \hat{T} + \hat{V}_{ext} | \phi \rangle + E_H[\rho] + \lambda E_x^{HF}[\phi] + (1-\lambda) E_x[\rho] + E_c[\rho] - \lambda^2 E_c\left[\rho_{1/\lambda}\right] \quad (8)$$

If one makes the approximation that $E_c[\rho_{1/\lambda}] \approx E_c[\rho]$ and assumes that $c_{DFTc}E_c[\rho]+c_{MP2}=1$, then this leads to a quadratic dependence of the MP2 prefactor on the percentage of HF exchange, $c_{MP2}=c_{HF}^2$.[28] If one instead makes a slight less crude approximation, $E_c[\rho_{1/\lambda}] \approx \lambda E_c[\rho]$, one instead obtains a cubic dependence, $c_{MP2}=c_{HF}^3$. Somewhat arbitrarily choosing $c_{HF}=0.5$ (i.e., 50% HF exchange) leads to just 12.5% MP2-like correlation and 87.5% PBEc correlation in the PBE0-DH double hybrid.[26]

In Figures 2 and 4 we have taken another tack. For the DSD-PBExPBEc form at different values of c_{HF} we have optimized, for the W4-11 benchmark set of 140 accurate total atomization energies, both the value of c_{MP2} for a conventional PBE2PPBE double hybrid (c_{PBEc} being set to $1-c_{MP2}$) and the values of c_s, c_o, and c_{PBEc} for DSD-PBEPBE.

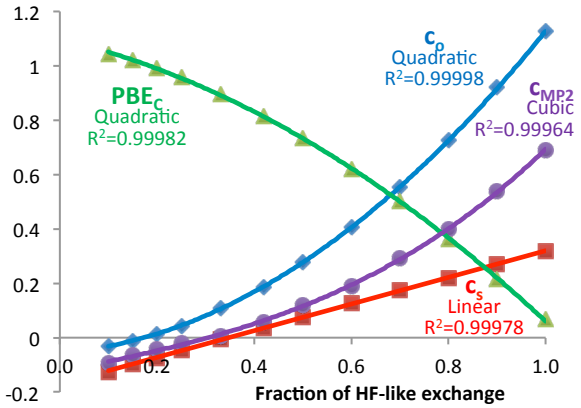

Fig. 4 dependence of the c_{MP2} coefficient in the PBE2xPPBE and c_c, c_s, and c_o in the DSD-PBEPBE form, fitted to the W4-11 benchmark, on the fraction of Hartree-Fock-like exchange. Note the linear dependence for c_s, the quadratic dependences for c_o and c_c, and the cubic dependence for μ if $c_s = c_o = \mu$ is imposed.

It can be seen there that for the PBE2PPBE form, c_{MP2} can be very well fitted (R^2=0.99964) to a cubic polynomial. The quadratic coefficient is very small, leaving the cubic term plus a linear term of the form ac_{HF}-b. This latter term counteracts the almost perfect a_0–$a'c_{HF}$ dependence seen[39] in the exchange component: as c_{HF} is increased and thus c_{PBEx}=1–c_{HF} decreased, the static correlation absorbed in the DFT exchange[126,127] is lost, and the linear ac_{HF}-b term in the double hybrid compensates for this, after a fashion. It should be noted that for small c_{HF} (in the 0-25% range), c_{MP2} is actually negative: in Figure 2 one can see that RMSD for the ordinary hybrid has a sharp minimum near 25% (i.e., PBE0) and steeply rises if one goes away from that point, while the ordinary double hybrid has a much flatter profile. Its behavior to the left of PBE0 reflects, indeed, negative c_{MP2} as a way to compensate for overcorrection for nondynamical correlation.

Turning now to the DSD-PBEPBE form, here we see rather different behavior. While c_s depends almost perfectly linearly (R^2=0.99978) on c_{HF}, c_o and c_{DFTc} exhibit clear parabolic and inverted-parabolic dependence, respectively, on c_{HF}, with respective squared correlation coefficients R^2=0.99998 and 0.99982. As seen in fig. 2, the performance advantage of the DSD form over the conventional DH form increases with increasing c_{HF}, widening to 2 kcal/mol at 100% HF exchange. The DSD curve is particularly flat, showing RMSDs of 4 kcal/mol or better over almost the entire 30-100% range and values at or below 3 kcal/mol over almost the entire 50-80% range. While a minimum of sorts can be identified around 65% HF exchange (not coincidentally, near the optimized c_{HF} for our various double hybrids despite their more diverse training sets), basically anything in the 60-70% range should be regarded as "optimum plains".

As an aside, we note that c_{DFTc}+c_o stays at 1.00±0.01 over the 10-60% HF exchange region, then starts climbing away from unity until it hits 1.17 at 100% HF exchange. In contrast, for the ordinary double hybrid, c_{DFTc}+c_{MP2} goes through a basin as HF exchange is increased from 10% to 100%, dipping down to 0.83 near 70% HF exchange.

How specific is the c_{HF} dependence of DSD parameters to the PBE functional? For the DSD-LDA form,

essentially the same qualitative behavior is obtained (see ESI). The odd man out is the DSD-BLYP form (see ESI), the anomalous behavior of which (resulting from the unbalanced treatment of same and opposite spin in LYP) we have already noted at length earlier in this article (see also our previous communication[30]).

Conclusions

In this full paper following our earlier communication, we analyzed the variables and components that maximize the accuracy of double hybrids, i.e. the fusion of DFT with perturbation theory (a "fifth rung on Jacob's ladder"). This included the selection of the exchange and correlation functionals, the coefficients of each component (DFT, exact exchange and perturbative correlation in both the same spin and opposite spin terms), and the addition of an ad-hoc dispersion correction; we called these parametrizations "DSD-DFT" ("**D**ispersion corrected, **S**pin component scaled, **D**ouble hybrid - DFT"). We have carried out an extensive survey against a diverse training set, and validated for a wide variety of thermochemical (including weak interaction), kinetic, and spectroscopic benchmarks. We are now in a position to make the following observations:

• Such spin-component scaled double hybrids, for most exchange-correlation contributions, exceed the performance of conventional double hybrids and approach performance of composite ab initio methods.

• Somewhat surprisingly, the quality of the results is only mildly sensitive to the choice of the underlying DFT exchange and correlation components: even DSD-LDA yields respectable performance.

• Simple, nonempirical GGAs appear to work best, while meta-GGAs offer no advantage, with the notable exception of the B95c correlation functional.

• Of the various DFT correlation components considered, B95c, P86, PW91 and PBEc (in that order) yield the best performance.

• The choice of the DFT exchange component is less important: PBEx and mPW appear to work well, but PBEh[128] (or, functionally equivalently for smallish molecules, HSE[129]) yields marginally better error statistics.

• The earlier assertion that LYP correlation is essential for double hybrids is an artifact of constraining the same-spin and opposite-spin coefficients to be equal.

• The best functionals of the pool considered are DSD-PBEhB95 and DSD-PBEP86. DSD-PBEhB95 yields the best overall performance statistics (especially for main group and transition metals), while DSD-PBEP86 is available (without reprogramming) in a wider variety of QM codes and has an edge for such properties as weak interaction energies and vibrational frequencies.

• Performance for vibrational frequencies is intermediate between CCSD and CCSD(T) for all functionals considered.

• DSD-PBEhB95 yields pleasing performance for singlet-triplet splitting energies, despite not being parametrized for these.

• The functionals were parametrized for both the D2 and D3BJ methods. However, a version without

any dispersion correction is also considered. (We note that for certain subsets, e.g., for hydrogen-bonded species, dispersion corrections consistently appear to do more harm than good — we observed a similar phenomenon in a very recent study on halogen-bonded species.[93])

• Opposite-spin-only double hybrids are additional options, for program systems that can exploit the speedup from neglecting same-spin correlation.

• Input examples for applying these functionals using unmodified versions of several common quantum chemistry programs are given in the Electronic Supporting Information.

• In difficult systems, double hybrids tend to "inherit" the inaccuracies of regular DFT, but to a lesser extent.

Some caveats are in order, lest we give the mistaken impression that we regard these double hybrids as a panacea. For instance, solid state calculations are still mostly outside their reach (and beyond the scope of our activities), while we also have not yet considered electronic excitations other than lowest singlet-lowest triplet, nor severely multireference transition metal systems.

Yet we believe we have shown throughout this paper that DSD-DFT (especially DSD-PBEhB95 and DSD-PBEP86) offer a level of accuracy and robustness generally associated with composite ab initio schemes.[130–132] Particularly if density fitting/"resolution of the identity" approximations are applied,[85–88] the extra cost over conventional hybrids is small enough that it is more than justified by the accuracy gained. Moreover, this accuracy is achieved using, by the standards of contemporary empirical DFT functionals,[41,56,89] quite a modest number (half a dozen or less) of empirical parameters.

Notes

Supplementary Information available: Input examples of DSD-DFT methods for Gaussian, Orca, Molpro and Q-Chem, reference energies of the singlet-triplet excitation energies validation set, dependence of DSD-DFT coefficients for B-LYP and LDA, and full references 78–81.

Acknowledgements: This research was made possible by a generous grant of computer time from the Faculty of Chemistry at the Weizmann Institute of Science and by a startup grant at the University of North Texas. *** The authors would like to thank Dr. Brina Brauer for assistance with GAMESS and with editing the manuscript, and Prof. Leeor Kronik for helpful discussions. *** SK acknowledges a Koshland Fellowship.

References

[1] J. P. Perdew, K. Schmidt, *AIP Conference Proceedings* **2001**, *577*, 1–20.
[2] S. Kümmel, L. Kronik, *Rev. Mod. Phys.* **2008**, *80*, 3–60.
[3] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.
[4] S. Kurth, J. P. Perdew, *Phys. Rev. B* **1999**, *59*, 10461–10468.
[5] J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, J. G. Ángyán, *Phys. Rev. Lett.* **2008**, *102*, 096404.
[6] B. G. Janesko, T. M. Henderson, G. E. Scuseria, *J. Chem. Phys.* **2009**, *130*, 081105.
[7] H. Eshuis, J. Bates, F. Furche, *Theor. Chem. Acc.* **2012**, *131*, 1–18.
[8] G. E. Scuseria, T. M. Henderson, D. C. Sorensen, *J. Chem. Phys.* **2008**, *129*, 231101.
[9] A. Görling, M. Levy, *Phys. Rev. B* **1993**, *47*, 13105.
[10] A. Görling, M. Levy, *Phys. Rev. A* **1994**, *50*, 196–204.
[11] T. Schwabe, S. Grimme, *Acc. Chem. Res.* **2008**, *41*, 569–579.
[12] I. Y. Zhang, X. Xu, *Int. Rev. Phys. Chem.* **2011**, *30*, 115–160.
[13] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
[14] Y. Zhao, B. J. Lynch, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 4786–4791.
[15] J. G. Ángyán, I. C. Gerber, A. Savin, J. Toulouse, *Phys. Rev. A* **2005**, *72*, 012510.
[16] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
[17] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, J. M. L. Martin, *J. Phys. Chem. A* **2008**, *112*, 3–8.
[18] T. Benighaus, R. A. DiStasio, R. C. Lochan, J.-D. Chai, M. Head-Gordon, *J. Phys. Chem. A* **2008**, *112*, 2702–2712.
[19] A. Karton, A. Tarnopolsky, J.-F. Lamère, G. C. Schatz, J. M. L. Martin, *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
[20] Y. Zhang, X. Xu, W. A. Goddard, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4963–4968.
[21] D. C. Graham, A. S. Menon, L. Goerigk, S. Grimme, L. Radom, *J. Phys. Chem. A* **2009**, *113*, 9861–9873.
[22] J. C. Sancho-García, A. J. Pérez-Jiménez, *J. Chem. Phys.* **2009**, *131*, 084108.
[23] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2009**, *131*, 174105.
[24] S. Kozuch, D. Gruzman, J. M. L. Martin, *J. Phys. Chem. C* **2010**, *114*, 20801–20808.
[25] K. Sharkas, J. Toulouse, A. Savin, *J. Chem. Phys.* **2011**, *134*, 064113.
[26] E. Brémond, C. Adamo, *J. Chem. Phys.* **2011**, *135*, 024106.
[27] B. Chan, L. Radom, *J. Chem. Theory Comput.* **2011**, *7*, 2852–2863.
[28] J. Toulouse, K. Sharkas, E. Brémond, C. Adamo, *J. Chem. Phys.* **2011**, *135*, 101102.
[29] I. Y. Zhang, X. Xu, Y. Jung, W. A. Goddard, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 19896 –19900.
[30] S. Kozuch, J. M. L. Martin, *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.
[31] A. Mohajeri, M. Alipour, *J. Chem. Phys.* **2012**, *136*, 124111.
[32] I. Y. Zhang, N. Q. Su, É. A. G. Brémond, C. Adamo, X. Xu, *J. Chem. Phys.* **2012**, *136*, 174103.
[33] J.-D. Chai, S.-P. Mao, *Chem. Phys. Lett.* **2012**, *538*, 121–125.
[34] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.
[35] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[36] J. Zheng, Y. Zhao, D. G. Truhlar, *J. Chem. Theory Comput.* **2009**, *5*, 808–821.
[37] A. D. Boese, J. M. L. Martin, *J. Chem. Phys.* **2004**, *121*, 3405.
[38] A. Karton, S. Daon, J. M. L. Martin, *Chem. Phys. Lett.* **2011**, *510*, 165–178.
[39] U. R. Fogueri, S. Kozuch, A. Karton, J. M. L. Martin, *Theor. Chem. Acc.* **2013**, *132*, 1291.
[40] A. J. Stone, *The theory of intermolecular forces*, Clarendon Press ; Oxford University Press, Oxford; New York, **1997**.
[41] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2007**, *120*, 215–241.
[42] C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664–675.
[43] S. Grimme, *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.
[44] S. Grimme, *J. Comput. Chem.* **2004**, *25*, 1463–1473.
[45] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
[46] J. Klimeš, A. Michaelides, *J. Chem. Phys.* **2012**, *137*, 120901.
[47] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
[48] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.
[49] P. Jurecka, J. Cerný, P. Hobza, D. R. Salahub, *J. Comput. Chem.* **2007**, *28*, 555–569.
[50] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[51] A. Tkatchenko, M. Scheffler, *Phys. Rev. Lett.* **2009**, *102*, 073005.
[52] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.
[53] O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **2010**, *133*, 244103.
[54] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, B. I. Lundqvist, *Phys. Rev. Lett.* **2004**, *92*, 246401.
[55] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, D. C. Langreth, *Phys. Rev. B* **2010**, *82*, 081101.
[56] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
[57] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *"DFT-D3 - A Dispersion Correction for DFT-functionals"; http://toc.uni-muenster.de/DFTD3*, **n.d.**
[58] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095.
[59] R. A. Distasio, M. Head-Gordon, *Mol. Phys.* **2007**, *105*, 1073–1083.
[60] K. E. Riley, J. A. Platts, J. Řezáč, P. Hobza, J. G. Hill, *J. Phys. Chem. A* **2012**, *116*, 4159–4169.
[61] J. M. L. Martin, *J. Phys. Chem. A* **2013**,*117*, 3118-3132.
[62] Y. Jung, R. C. Lochan, A. D. Dutoi, M. Head-Gordon, *J. Chem. Phys.* **2004**, *121*, 9793.
[63] E. Fromager, *J. Chem. Phys.* **2011**, *135*, 244106.
[64] A. Karton, E. Rabinovich, J. M. L. Martin, B. Ruscic, *J. Chem. Phys.* **2006**, *125*, 144108.
[65] Y. Zhao, D. G. Truhlar, *J. Chem. Theory Comput.* **2009**, *5*, 324–333.
[66] M. Korth, S. Grimme, *J. Chem. Theory Comput.* **2009**, *5*, 993–1003.
[67] P. Jurečka, J. Šponer, J. Černý, P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
[68] T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, C. D. Sherrill, *J. Chem. Phys.* **2010**, *132*, 144104.
[69] M. M. Quintal, A. Karton, M. A. Iron, A. D. Boese, J. M. L. Martin, *J. Phys. Chem. A* **2006**, *110*, 709–716.
[70] F. Jensen, *J. Chem. Phys.* **2001**, *115*, 9113.
[71] F. Jensen, *J. Chem. Phys.* **2002**, *116*, 7372.
[72] F. Jensen, *J. Chem. Phys.* **2002**, *117*, 9234.
[73] F. Jensen, T. Helgaker, *J. Chem. Phys.* **2004**, *121*, 3463.
[74] F. Jensen, *J. Phys. Chem. A* **2007**, *111*, 11198–11204.
[75] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829.
[76] (a) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297. (b) D. Rappoport and F. Furche, *J. Chem. Phys.* **2010**, *133*, 134105.
[77] A. Halkier, H. Koch, P. Jørgensen, O. Christiansen, I. M. B. Nielsen, T. Helgaker, *Theor. Chem. Acc.* **1997**, *97*, 150–157.
[78] *Gaussian 09, Rev. C01;* M. J. Frisch et al.,; *See full ref. in SI.*
[79] *MOLPRO, Version 2010.1;* H. J. Werner, et al*; See full ref. in SI.*
[80] *ORCA,* F. Neese, et al*; See full ref. in SI.*
[81] *Q-CHEM,* Y. Shao, et al*; See full ref. in SI.*
[82] Y. Zhao, D. G. Truhlar, *J. Chem. Theory Comput.* **2005**, *1*, 415–432.
[83] J. P. Perdew, K. Burke, *Int. J. Quantum Chem.* **1996**, *57*, 309–319.
[84] B. Miehlich, A. Savin, H. Stoll, H. Preuss, *Chem. Phys. Lett.* **1989**, *157*, 200–206.
[85] O. Vahtras, J. Almlöf, M. W. Feyereisen, *Chem. Phys. Lett.* **1993**, *213*, 514–518.
[86] R. A. Kendall, H. A. Früchtl, *Theor. Chem. Acc.* **1997**, *97*, 158–163 (and references therein).
[87] D. E. Bernholdt, *Parallel Comput.* **2000**, *26*, 945–963 (and references therein).
[88] F. Weigend, M. Häser, *Theor Chem Acta* **1997**, *97*, 331–340.
[89] R. Peverati, D. G. Truhlar, *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.
[90] *GAMESS US,* M.W.Schmidt, et al*; See full ref. in SI.*
[91] Y. Zhao, N. González-García, D. G. Truhlar, *J. Phys. Chem. A* **2005**, *109*, 2012–2018.
[92] A. D. Boese, J. M. L. Martin, W. Klopper, *J. Phys. Chem. A* **2007**, *111*, 11122–11133.
[93] S. Kozuch, J. M. L. Martin, *J. Chem. Theory Comput.* **2013**, *9*, 1918-1931.
[94] V. M. Rayón, H. Valdés, N. Díaz, D. Suárez, *J. Chem. Theory Comput.* **2008**, *4*, 243–256.
[95] V. Guner, K. S. Khuong, A. G. Leach, P. S. Lee, M. D. Bartberger, K. N. Houk, *J. Phys. Chem. A* **2003**, *107*, 11445–11459.
[96] D. H. Ess, K. N. Houk, *J. Phys. Chem. A* **2005**, *109*, 9542–9553.
[97] R. Kang, H. Chen, S. Shaik, J. Yao, *J. Chem. Theory Comput.* **2011**, *7*, 4002–4011.
[98] V. M. García, O. Castell, M. Reguero, R. Caballol, *Mol. Phys.* **1996**, *87*, 1395–1404.
[99] S. Matzinger, M. P. Fuelscher, *J. Phys. Chem.* **1995**, *99*, 10747–10751.

[100] K. Balasubramanian, A. D. McLean, *J. Chem. Phys.* **1986**, *85*, 5117–5119.
[101] M. Reiher, O. Salomon, B. A. Hess, *Theor. Chem. Acc.* **2001**, *107*, 48–55.
[102] D. H. Ess, T. C. Cook, *J. Phys. Chem. A* **2012**, *116*, 4922–4929.
[103] (a) J. Řezáč, K. E. Riley, P. Hobza, *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438. (b) J. Řezáč, K. E. Riley, P. Hobza, *J. Chem. Theory Comput.* **2011**, *7*, 3466–3470.
[104] Goerigk, L., Kruse, H., Grimme, S., *ChemPhysChem* **2011**, *12*, 3421–3433.
[105] J. M. L. Martin, T. J. Lee, P. R. Taylor, *J. Chem. Phys.* **1998**, *108*, 676–691.
[106] J. M. L. Martin, P. R. Taylor, *Chem. Phys. Lett.* **1996**, *248*, 336–344.
[107] A. Karton, B. Ruscic, J. M. L. Martin, *J. Mol. Struct. THEOCHEM* **2007**, *811*, 345–353.
[108] C. Levi, J. M. L. Martin, I. Bar, *J. Comput. Chem.* **2008**, *29*, 1268–1276.
[109] T. J. Lee, J. M. L. Martin, P. R. Taylor, *J. Chem. Phys.* **1995**, *102*, 254–261.
[110] J. M. L. Martin, O. Uzan, *Chem. Phys. Lett.* **1998**, *282*, 16–24.
[111] J. M. L. Martin, *J. Chem. Phys.* **1994**, *100*, 8186–8193.
[112] J. M. L. Martin, J. P. Francois, R. Gijbels, *J. Mol. Spectrosc.* **1995**, *169*, 445–457.
[113] C. E. Dateo, T. J. Lee, D. W. Schwenke, *J. Chem. Phys.* **1994**, *101*, 5853.
[114] J. Koput, K. A. Peterson, *Chem. Phys. Lett.* **1998**, *283*, 139–146.
[115] D. Wang, Q. Shi, Q.-S. Zhu, *J. Chem. Phys.* **2000**, *112*, 9624.
[116] J. M. L. Martin, *J. Chem. Phys.* **1998**, *108*, 2791–2800.
[117] S. S. Zade, N. Zamoshchik, A. R. Reddy, G. Fridman-Marueli, D. Sheberla, M. Bendikov, *J. Am. Chem. Soc.* **2011**, *133*, 10803–10816.
[118] S. Grimme, C. Diedrich, M. Korth, *Angew. Chem. Int. Ed.* **2006**, *45*, 625–629.
[119] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2010**, *6*, 107–126.
[120] D. Gruzman, A. Karton, J. M. L. Martin, *J. Phys. Chem. A* **2009**, *113*, 11974–11983.
[121] E. R. Johnson, J. Contreras-García, W. Yang, *J. Chem. Theory Comput.* **2012**, *8*, 2676–2681.
[122] P. von R. Schleyer, J. E. Williams, K. R. Blanchard, *J. Am. Chem. Soc.* **1970**, *92*, 2377–2386.
[123] U. R. Fogueri, S. Kozuch, A. Karton, J. M. L. Martin, *J. Phys. Chem. A* **2013**, *117*, 2269-2277.
[124] A. Karton, J. M. L. Martin, *Mol. Phys.* **2012**, *110*, 2477–2491.
[125] A. Karton, J. M. L. Martin, *J. Chem. Phys.* **2012**, *136*, 124114.
[126] N. C. Handy, A. J. Cohen, *Mol. Phys.* **2001**, *99*, 403–412.
[127] A. J. Cohen, N. C. Handy, *Mol. Phys.* **2001**, *99*, 607–615.
[128] M. Ernzerhof, J. P. Perdew, *J. Chem. Phys.* **1998**, *103*, 3313-3320.
[129] J. Heyd, G. Scuseria, M. Ernzerhof, *J. Chem. Phys.* **2003**, *118*, 8207-8215
[130] J. A. Montgomery, M. J. Frisch, J. W. Ochterski, G. A. Petersson, *J. Chem. Phys.* **1999**, *110*, 2822–2827.
[131] L. A. Curtiss, P. C. Redfern, K. Raghavachari, *J. Chem. Phys.* **2007**, *126*, 084108.
[132] N. J. DeYonker, T. R. Cundari, A. K. Wilson, *J. Chem. Phys.* **2006**, *124*, 114104.