Assessing DFT-D3 damping functions across widely-used density functionals: can we do better?

Journal:	*Journal of Chemical Theory and Computation*
Manuscript ID	ct-2017-00176x.R1
Manuscript Type:	Article
Date Submitted by the Author:	06-Apr-2017
Complete List of Authors:	Witte, Jonathon; University of California, Berkeley, Chemistry Mardirossian, Narbe; UC Berkeley, Neaton, Jeffrey; Lawrence Berkeley National Laboratory, Molecular Foundry Head-Gordon, Martin; University of California, Berkeley, Chemistry

Assessing DFT-D3 damping functions across widely-used density functionals: can we do better?

Jonathon Witte,[†,‡,⊥] Narbe Mardirossian,[†,⊥] Jeffrey B. Neaton,[‡,¶,§] and Martin Head-Gordon[*,†,‖]

†*Department of Chemistry, University of California, Berkeley, California 94720, United States*

‡*Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

¶*Department of Physics, University of California, Berkeley, California 94720*

§*Kavli Energy Nanosciences Institute at Berkeley, Berkeley, California 94720, United States*

‖*Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

⊥*These authors contributed equally to this work*

E-mail: mhg@cchem.berkeley.edu

Abstract

With the aim of improving the utility of the DFT-D3 empirical dispersion correction, we herein generalize the DFT-D3 damping function by optimizing an additional parameter, an exponent, which controls the rate at which the dispersion tail is activated. This method – DFT-D3(op), shorthand for "optimized power," where power

refers to the newly introduced exponent – is then parameterized for use with ten popular density functional approximations across a small set of non-covalent interactions and isomerization energies; the resulting methods are then evaluated across a large independent test set of 2475 non-covalent binding energies and isomerization energies. We find that the DFT-D3(op) tail represents a substantial improvement over existing damping functions, as it affords significant reductions in errors associated with non-covalent interaction energies and geometries. The revPBE0-D3(op) and MS2-D3(op) methods in particular stand out, and our extensive testing indicates they are competitive with other modern density functionals.

1 Introduction

Kohn-Sham density functional theory (DFT)[1] is the most widely used formalism in electronic structure theory, a consequence of its relative simplicity and the nice balance it strikes between computational expense and accuracy. Nevertheless, DFT has its drawbacks. Although there is a degree of hierarchy within DFT, as exemplified by the proverbial Jacob's ladder of DFT,[2] there is no prescription for systematically improving results. Moreover, standard approximations within the formalism are inherently semi-local, and hence incapable of correctly describing long-range electron correlation, i.e. strong correlations and dispersion forces.[3] This latter deficiency is particularly troubling, since dispersion is integral to the correct description of non-covalent interactions. To address this issue, a "stairway" of dispersion corrections has been constructed over the years.[4]

The simplest such corrections can be traced back to a Hartree–Fock+D approach,[5,6] which, channeling second-order Rayleigh-Schrödinger perturbation theory, adds in a pairwise atomic correction involving empirical isotropic dispersion coefficients with the correct r^{-6} asymptote. This scheme was later adapted to DFT,[7] which introduced an additional complication: since DFT already describes local electron correlation, the added +D component needs to be damped at small separations in order to avoid double counting. Grimme

systematized this approach, first with his introduction of the DFT-D method,[8] then subsequently DFT-D2[9] and DFT-D3.[10] These methods are widely used for the same reason that DFT is so prolific within the electronic structure community: they are simple, incredibly efficient, and quite accurate for a variety of interesting systems.

In recent years, a number of additional approaches to dispersion have been developed. Von Lilienfeld *et al.*[11] proposed adding in dispersion-corrected atom-centered potentials (DCACPs) within the effective core-potential approximation; this approach was later adapted to atom-centered basis sets in what is now known as the DCP approach.[12] Several approaches for self-consistently calculating dispersion coefficients have been introduced, most notably the exchange-dipole moment (XDM) model,[13–16] and the TS-VdW method.[17] The past decade has even seen the proliferation of methods that attempt to explicitly incorporate non-local correlation, such as the vdW-DF[18] and vdW-DF2[19] approaches, which are popular within the solid-state community, and the VV09[20] and VV10[21] methods. For a more thorough description of these various methods, the reader is referred to a recent review by Klimeš and Michaelides[4].

Within this study, we focus on the most computationally inexpensive brand of dispersion corrections, the aforementioned DFT-D2 and DFT-D3 approaches. Specifically, we explore the effect of including an additional degree of freedom within the -D3 damping function, thereby introducing a new, more general damping function. This new damping function is then optimized for several popular density functionals, and the resulting methods are compared to those obtained with existing -D3 damping functions. We find that this new iteration, which we term DFT-D3(op), shorthand for optimized power, substantially improves the description of non-covalent interactions – particularly those involving molecular clusters – and isomerization energies.

2 Theory

Within the DFT-D family of methods, the two-body component of the empirical dispersion energy is given by

$$E^{(2)} = -\sum_{i<j} \sum_{n=6,8,10,\ldots} s_n \frac{C_{n,ij}}{r_{ij}^n} f_{\mathrm{damp},n}(r_{ij}). \tag{1}$$

The first sum in eq. (1) runs over all unique pairs of atoms i and j; $C_{n,ij}$ are isotropic nth-order dispersion coefficients for atom pair ij; r_{ij} is the internuclear distance between atoms i and j; s_n are global, density functional-dependent scaling parameters; and $f_{\mathrm{damp},n}(r_{ij})$ are damping functions intended to address small-r_{ij} singularities, as well as double-counting of correlation effects. Early iterations of these dispersion models – namely the original DFT-D,[7,8] as well as DFT-D2[9] – truncated the sum at $n = 6$, employed chemically-insensitive, pre-tabulated dispersion coefficients $C_{6,ij}$ and van der Waals radii $r_{0,ij}$, and utilized Fermi-type damping functions of the form

$$f_{\mathrm{damp},6}^{\mathrm{D2}}(r_{ij}) = \left[1 + \exp\left(-\alpha\left(\frac{r_{ij}}{r_{0,ij}} - 1\right)\right)\right]^{-1}, \tag{2}$$

with α generally fixed to 20.

In 2010, Grimme et al.[10] introduced the now widely used DFT-D3 scheme. Although many aspects of -D3 are similar to -D2, there are some key fundamental differences: a counting function is introduced to allow the $C_{6,ij}$ coefficients to be weakly environmentally-dependent; the sum in eq. (1) is extended to include the $n = 8$ term, and a CHG-style[22] damping function – given in eq. (3) – is used. For the damping function, Grimme et al.[10] chose to fix $s_{r,8} = 1$, $\alpha_6 = 12$, and $\alpha_8 = 14$, thereby optimizing only one nonlinear parameter – $s_{r,6}$ – for each density functional. Moreover, for almost all density functionals, s_6 is fixed to unity, leaving only one linear parameter, s_8. This version of DFT-D3 is now known as zero-damping, or DFT-D3(0).

$$f_{\mathrm{damp},n}^{\mathrm{D3(0)}}(r_{ij}) = \left[1 + 6\left(\frac{r_{ij}}{s_{r,n} r_{0,ij}}\right)^{-\alpha_n}\right]^{-1} \tag{3}$$

One year later, Grimme et al.[23] combined the basic principles of DFT-D3 with the finite-damping scheme Johnson and Becke[15] had utilized in their XDM approach to dispersion; this is now the generally preferred style of -D3, termed DFT-D3(BJ). The damping function is of the form

$$f_{\mathrm{damp},n}^{\mathrm{D3(BJ)}}(r_{ij}) = \frac{r_{ij}^n}{r_{ij}^n + \left(\alpha_1 r_{0,ij} + \alpha_2\right)^n}, \tag{4}$$

where α_1 and α_2 are adjustable nonlinear parameters. At short internuclear distances r_{ij}, the dispersion energy $E^{(2)}$ in the zero-damping approach vanishes, since $f_{\mathrm{damp,n}}(r_{ij})$ decays more quickly than $C_{n,ij} r_{ij}^{-n}$. In the BJ-damping scheme, however, these two terms decay at the same rate, and hence $E^{(2)}$ asymptotes to a finite value. This is the key difference between the -D3(0) and -D3(BJ) approaches. Although van der Waals radii $r_{0,ij}$ in -D3(BJ) are given by $\sqrt{\frac{C_{8,ij}}{C_{6,ij}}}$ instead of their -D3(0) values, the tabulated coefficients $C_{n,ij}$ are the same, and – as with -D3(0) – s_6 is generally fixed to unity in -D3(BJ), leaving s_8 as the sole linear parameter.

This new version of -D3 with BJ-damping has become the preferred version of -D3 due to the fact that it consistently outperforms -D3(0).[23] Recently, Schröder et al.[24] have attempted to simplify the model with their C-Six-Only (CSO) approach, wherein they introduce a sigmoidal interpolation function to approximate the eighth-order term. In so doing, they eliminate one linear parameter and one nonlinear parameter without significantly impacting performance across GMTKN30 or S66.[24] The damping function for this approach, -D3(CSO), is given by

$$f_{\mathrm{damp},6}^{\mathrm{D3(CSO)}}(r_{ij}) = \frac{r_{ij}^6}{r_{ij}^6 + \left(\alpha_3 r_{0,ij} + \alpha_4\right)^6}\left[1 + \frac{\alpha_1}{s_6\left[1 + \exp\left(r_{ij} - \alpha_2 r_{0,ij}\right)\right]}\right]. \tag{5}$$

Note the similarities between eqs. (4) and (5): α_3 and α_4 in the CSO scheme correspond to

α_1 and α_2 in BJ-damping, respectively, and the bracketed term in eq. (5) is the interpolation function. For the density functionals they examined, Schröder et al.[24] found $\alpha_3 \approx 0$, $\alpha_4 \approx 2.5$, and $\alpha_2 \approx 2.5$, leaving α_1 as the sole functional-dependent parameter in -D3(CSO).

In addition to the dispersion corrections mentioned thus far, we consider in this work the modified version of -D3(BJ) recently proposed by Smith et al.[25]. The damping function in this approach, -D3M(BJ), is identical to that in eq. (4); the method constitutes a refitting of the BJ-damping parameters to a much broader set of data. Grimme et al.[23] originally fit s_8, α_1, and α_2 on 130 data points; Smith et al.[25] utilize a training set of 1526 energies, with an emphasis on non-equilibrium – particularly compressed – geometries.

It is illustrative to rewrite the damping function for -D3(0) from eq. (3). By doing so, we obtain

$$f_{\mathrm{damp},n}^{\mathrm{D3(0)}}(r_{ij}) = \frac{r_{ij}^{\alpha_n}}{r_{ij}^{\alpha_n} + \left(\sqrt[\alpha_n]{6}s_{r,n}r_{0,ij}\right)^{\alpha_n}}. \tag{6}$$

The similarities between the functional forms of -D3(0) and -D3(BJ) are striking (cf. eqs. (4) and (6)). There is a direct correspondence between α_1 in BJ-damping and $s_{r,n}$. Whereas in -D3(BJ) α_1 is the same for the sixth- and eighth-order terms, in -D3(0) this is no longer the case, as $s_{r,6} \neq s_{r,8}$. The original zero-damping scheme thus resembles a slightly constrained version of BJ-damping, wherein α_2 (the constant added to the van der Waals radii) is zero, and the sixth-order power is 14 instead of 6.

In this work, we generalize the BJ-damping function, adding a parameter to control how quickly the damping occurs, i.e. the power. The damping function we employ is given by

$$f_{\mathrm{damp},n}^{\mathrm{D3(op)}}(r_{ij}) = \frac{r_{ij}^{\beta_n}}{r_{ij}^{\beta_n} + (\alpha_1 r_{0,ij} + \alpha_2)^{\beta_n}} \tag{7}$$

This new scheme – optimized-power-damping, or -D3(op) – is mathematically similar to both BJ-damping and zero-damping. As is the case with both BJ- and zero-damping, we constrain $\beta_8 = \beta_6 + 2$. We utilize the same isotropic dispersion coefficients as both -D3(0)

and -D3(BJ), the same van der Waals radii as -D3(BJ) – i.e. $r_{0,ij} = \frac{C_{8,ij}}{C_{6,ij}}$ – and optimize three nonlinear parameters ($\alpha_1, \alpha_2, \beta = \beta_6$) and one linear parameter (s_8). The effects of varying the nonlinear parameters in this model are visualized in Figure 1. The holdovers from BJ-damping – α_1 and α_2 – primarily control the distance r_{ij} at which the damping function switches off the dispersion correction: α_1 and α_2 are just multiplicative and additive terms, respectively, for the sum of van der Waals radii $r_{0,ij}$. The newly introduced β, on the other hand, controls the rate at which the dispersion correction is switched off; in the limit $\beta \to \infty$, $f_{\text{damp}}(r_{ij}) \to \theta(r_{ij})$, i.e. the damping function just becomes a step function. Most small changes in β correspond to subtle changes in the dispersion energy, $E^{(2)}$; the exception is the transition from $\beta = 6$ to $\beta = 6 + \epsilon$, which is a fundamentally different change than that from, say, $\beta = 10$ to $\beta = 10 + \epsilon$. For all $\beta > 6$, the contribution of atom pair ij to the dispersion energy at $r_{ij} = 0$ is zero; for $\beta = 6$, this contribution is nonzero. In practice, however, this difference is not so significant; due to Pauli repulsion, the limit $r_{ij} = 0$ is not particularly meaningful.

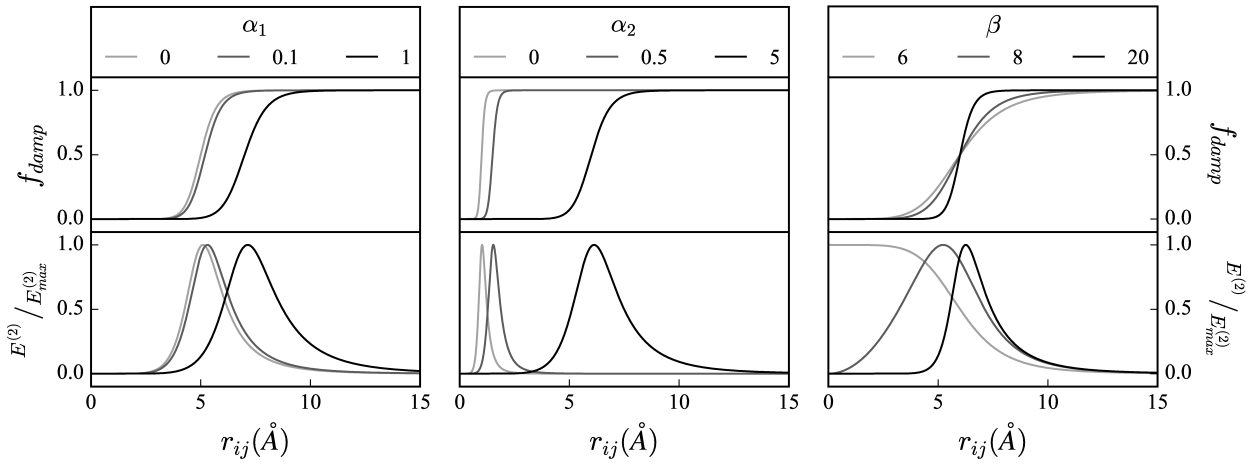


Figure 1: Visualization of the effects of varying parameters α_1, α_2, and β from eq. (7). The leftmost plots are generated by varying α_1 with fixed α_2 and β; the center plots are generated by varying α_2 with fixed α_1 and β; and the rightmost plots are generated by varying β with fixed α_1 and α_2. When not being varied, the parameters are fixed to $\alpha_1 = 0.5$, $\alpha_2 = 5$, and $\beta = 14$, with $r_{0,ij} = 2$ Å. The top plot in each section shows the damping function for the sixth-order term, $f^{\text{D3(op)}}_{\text{damp},6}(r_{ij})$ from eq. (7). The bottom plot in each section shows the sixth-order contribution to the two-body dispersion energy, $E^{(2)}$, which has been normalized to span the range [0,1].

An overview of the six forms of dispersion corrections considered in this study – -D2, -D3(0), -D3(BJ), -D3M(BJ), -D3(CSO), and -D3(op) – can be found in Table 1. Note we have not considered the modified version of zero-damping, -D3M(0), as it was found to be inferior to -D3M(BJ) by the original authors.[25]

Table 1: Summary of empirical dispersion corrections considered in this study.

	Fit Parameters			
Type	Linear	Nonlinear	C_8?	Reference
-D2	s_6	None	No	9
-D3(0)	s_8	$s_{r,6}$	Yes	10
-D3(BJ)	s_8	α_1, α_2	Yes	23
-D3M(BJ)	s_8	α_1, α_2	Yes	25
-D3(CSO)	None	α_1	No	24
-D3(op)	s_6 or s_8	α_1, α_2, β	Maybe	This work

3 Computational Details

We have optimized the DFT-D3(op) damping function given in eq. (7) for several density functional approximations and compared its performance to that of existing damping functions. Specifically, we have considered the ten density functionals outlined in Table 2.

These ten representative density functionals were carefully chosen. There are five natural pairs of pure/hybrid functionals: BLYP/B3LYP, B97/B97h, revPBE/revPBE0, TPSS/TPSSh, and MS2/MS2h. The first three of these pairs are generalized gradient approximations (GGAs), and the last two are meta-GGAs. Moreover, each of these ten density functionals exhibits positive mean signed errors across every dataset of non-covalent interactions we considered; that is to say, they consistently underbind every type of system at which we have looked, and hence can all profit greatly from the addition of a dispersion correction. One popular functional we excluded from this study is PBE. Since PBE is known to overbind water clusters,[37] we believe it is not a good candidate for a blanket dispersion correction. That being said, although PBE (and its complement, PBE0) are not found within the main study, parameterizations for both may be found in the Supporting Information.

Table 2: Summary of density functionals. The names in the first column are standard, with two exceptions: B97 corresponds to Grimme's pure functional B97-D,[9] which has had the dispersion tail stripped away; and B97h corresponds to the original hybrid functional B97, as parameterized by Becke.[26] The third column details the existing parameterized DFT-D-style dispersion corrections we consider in this study, and the fourth column lists the references for the method.

Functional	Class	Empirical Dispersion	Ref
BLYP	GGA	D2,D3(0),D3(BJ), D3M(BJ),D3(CSO)	27,28
B3LYP	hybrid GGA	D2,D3(0),D3(BJ), D3M(BJ),D3(CSO)	27–30
B97	GGA	D2,D3(0),D3(BJ), D3M(BJ)	9
B97h	hybrid GGA	D2	26
revPBE	GGA	D3(0),D3(BJ)	31,32
revPBE0	hybrid GGA	D3(0),D3(BJ)	31–33
TPSS	meta-GGA	D2,D3(0),D3(BJ), D3(CSO)	34
TPSSh	hybrid meta-GGA	D3(0),D3(BJ)	35
MS2	meta-GGA	D3(0)	36
MS2h	hybrid meta-GGA	D3(0)	36

All density functional calculations were performed in the def2-QZVPPD basis,[38,39] near the basis set limit for standard non-covalent interactions.[40] A fine Lebedev integration grid consisting of 99 radial shells – each with 590 angular points – was utilized in the computation of all semi-local components of exchange and correlation; non-local correlation in the VV10-containing functionals was calculated with the coarser SG-1 grid.[41] All calculations were performed within a development version of Q-Chem 4.4.[42]

For each density functional, we performed an exhaustive determination of the optimal parameters for the DFT-D3(op) method. Specifically, we scanned α_1 from 0 to 1 in steps of 0.025, α_2 from 0 to 10 in increments of 0.25, and β from 6 to 18 in increments of 2; this resulted in 11767 possible forms of the -D3(op) tail for each density functional.

To identify the best of these many candidate fits, we utilized the comprehensive database assembled by Mardirossian and Head-Gordon.[43] This database contains 4419 data points which are spread out among 82 smaller datasets. These smaller constituent datasets are classified according to eight distinct datatypes: NCED (easy non-covalent interactions of dimers), NCEC (easy non-covalent interactions of clusters), NCD (difficult non-covalent interactions of dimers), IE (easy isomerization energies), ID (difficult isomerization energies),

TCE (easy thermochemistry), TCD (difficult thermochemistry), and BH (barrier heights). "Difficult" interactions involve either strong correlation or self-interaction error, whereas "easy" interactions are not heavily characterized by either. In order to facilitate the testing of -D3(op) candidates, the datasets were divided into two categories. A training set was used to identify the best set of parameters, and a test set was used to assess the performance of the resulting method relative to existing dispersion corrections. A summary of the datasets can be found in Table 3.

Table 3: Summary of datasets that comprise the training and test sets. For more details, see Ref. 43.

Set	Datatype	#	Constituent Datasets	References
Train	NCED	127	S66, HB49, AlkBind12	44–49
	NCEC	18	H2O6Bind8, H2O20Bind4, HW6Cl	50–55
	IE	122	Butanediol65, Melatonin52, H2O16Rel5	56–58
	BL	20	Interpolated equilibrium binding lengths from BzDC215 and NBC10	59–63
Test	NCED	1617	A24, DS14, HB15, HSG, NBC10, S22, X40, A21x12, BzDC215, HW30, NC15, S66x8, 3B-69-DIM, CO2Nitrogen16, Ionic43	45,59–75
	NCEC	225	HW6F, FmH2O10, Shields38, SW49Bind345, SW49Bind6, WATER27, 3B-69-TRIM, CE20, H2O20Bind10	50–53,73,76–79
	NCD	91	TA13, XB18, Bauza30, CT20, XB51	80–84
	IE	633	AlkIsomer11, ACONF, CYCONF, Pentane14, SW49Rel345, SW49Rel6, H2O20Rel10, H2O20Rel4, YMPJ519	51–53,53–55,77,85–89
	ID	155	EIE22, Styrene45, DIE60, ISOMERIZATION20, C20C24	78,90,90–92
	TCE	947	AlkAtom19, BDE99nonMR, G21EA, G21IP, TAE140nonMR, AlkIsod14, BH76RC, EA13, HAT707nonMR, IP13, NBPRC, SN13, BSR36, HNBrBDE18, WCPT6	53,85,91,93–101
	TCD	258	BDE99MR, HAT707MR, TAE140MR, PlatonicHD6, PlatonicID6, PlatonicIG6, PlatonicTAE6	91,102
	BH	206	BHPERI26, CRBH20, DBH24, CR20, HTBH38, NHTBH38, PX13, WCPT27	53,78,79,94,95,101,103–107

Once all 11767 possible fits were generated for a given functional, we set s_6 to unity and performed a least-squares fit of s_8 to a small subset of NCEDTrain and IETrain, namely S66 and Butanediol65. The resulting methods were then sorted according to a simple product of root-mean-square errors (RMSEs) across eleven training datasets – S66, HB49, AlkBind12, H2O6Bind8, H2O20Bind4, HW6Cl, Butanediol65, Melatonin52, H2O16Rel5 and two geometric datasets: interpolated equilibrium binding lengths of BzDC215 and NBC10. To prevent the appearance of unphysical parameters (negative values of s_8), when s_8 optimized to a value less than 0.1, we set $s_8 = 0$ and performed a least-squares fit of s_6 instead. The only functional within this study for which this occurred is B97h.

4 Results and Discussion

In the course of this study, we have introduced a new damping function for use in the DFT-D3 empirical dispersion correction – DFT-D3(op) – which encompasses the space spanned by DFT-D3(BJ) and an unconstrained form of DFT-D3(0). We have optimized this new damping function for ten distinct density functionals across a small yet diverse training set. The resulting optimized fit parameters are listed in Table 4; parameterizations for additional functionals may be found in the Supporting Information.

Table 4: Optimized values of -D3(op) fit parameters for each density functional.

Functional	s_6	s_8	α_1	α_2	β
BLYP	1.00000	1.31867	0.425	3.50	8
B3LYP	1.00000	0.78311	0.300	4.25	10
B97	1.00000	1.46861	0.600	2.50	6
B97h	0.97388	0.00000	0.150	4.25	12
revPBE	1.00000	1.44765	0.600	2.50	6
revPBE0	1.00000	1.25684	0.725	2.25	6
TPSS	1.00000	0.51581	0.575	3.00	14
TPSSh	1.00000	0.43185	0.575	3.00	14
MS2	1.00000	0.90743	0.700	4.00	8
MS2h	1.00000	1.69464	0.650	4.75	6

This new variant of DFT-D3 was then evaluated in tandem with existing versions across a large independent test set. The results are given in Figure 2, within which we show for each method the root-mean-square errors (RMSEs) across the three energetic categories for which empirical dispersion methods are best suited – namely NCEDTest, NCECTest, and IETest – as well as one geometric category, S66x8 interpolated equilibrium binding lengths (and corresponding interpolated equilibrium binding energies). For each of NCEDTest, NCECTest, and IETest, both the aggregate results as well as results for two representative constituent datasets are provided.

From Figure 2, it is evident that the newly proposed -D3(op) dispersion correction represents an improvement over existing corrections for a diverse set of density functionals and systems. Even in cases where the power optimizes to $\beta = 6$, which corresponds to the DFT-D3(BJ) scheme, we see large improvements in some categories. This is the case, for instance, for revPBE. Relative to revPBE-D3(BJ), revPBE-D3(op) exhibits significantly

	Non-Covalent: Dimers			Non-Covalent: Clusters			Isomerization			Equilibrium		Overall
Functional	**NCED***	**HSG**	**S22**	**NCEC***	**Shields**	**3B-69**	**IE***	**Pent14**	**YMPJ**	**BL (Å)**	**BE**	**Product**
BLYP	3.46	3.39	6.00	14.20	7.82	8.39	1.58	0.55	1.51	0.225	3.80	
-D2	0.54	0.48	0.25	2.13	3.23	0.73	1.17	0.83	1.18	0.095	0.58	12.85
-D3(0)	0.36	0.39	0.28	**1.67**	2.52	0.52	0.83	0.20	0.80	0.028	0.42	1.41
-D3(BJ)	0.33	0.40	0.32	2.07	1.59	0.71	0.71	0.48	0.70	**0.027**	0.23	1.32
-D3M(BJ)	0.33	0.29	0.34	2.57	3.81	0.69	0.72	0.45	0.71	0.038	0.31	2.33
-D3(CSO)	0.40	0.48	0.38	3.21	0.61	0.80	**0.68**	0.49	0.65	0.061	0.29	5.29
-D3(op)	**0.31**	0.40	0.20	1.77	2.24	0.75	0.73	0.41	0.74	0.029	**0.21**	**1.14**
B3LYP	2.82	2.68	4.89	8.67	4.54	6.61	1.34	0.47	1.28	0.215	3.12	
-D2	0.58	0.45	0.61	5.19	5.14	0.80	0.89	0.54	0.88	0.100	0.81	26.63
-D3(0)	0.35	0.27	0.42	3.69	4.00	0.79	0.50	0.13	0.45	0.027	0.44	1.71
-D3(BJ)	0.30	0.21	0.43	2.67	3.21	0.74	0.51	0.20	0.49	**0.022**	0.38	0.89
-D3M(BJ)	0.34	0.19	0.52	3.32	3.79	0.80	0.50	0.19	0.49	0.023	0.44	1.32
-D3(CSO)	0.30	0.25	0.30	1.51	2.15	0.63	**0.48**	0.20	0.44	0.047	0.28	1.01
-D3(op)	**0.26**	0.30	0.25	**1.28**	1.88	0.56	0.55	0.20	0.54	0.027	**0.24**	**0.49**
B97	3.66	3.70	6.49	21.42	15.11	9.30	1.87	0.64	1.77	0.239	3.86	
-D2	0.57	0.72	0.62	5.63	3.66	1.42	1.28	0.66	1.33	0.068	0.53	27.50
-D3(0)	**0.39**	0.50	0.54	4.12	2.22	1.17	0.76	0.08	0.77	0.034	0.35	4.16
-D3(BJ)	0.46	0.55	0.49	4.75	2.75	1.11	0.82	0.28	0.82	0.047	0.38	8.40
-D3M(BJ)	0.40	0.40	0.38	2.75	0.94	1.17	**0.61**	0.19	0.60	0.032	0.30	2.14
-D3(op)	0.41	0.39	0.38	**1.83**	0.42	1.08	0.72	0.30	0.70	**0.026**	**0.29**	**1.38**
B97h	2.41	2.26	4.39	10.42	6.50	5.77	1.27	0.41	1.19	0.218	2.56	
-D2	0.49	0.39	0.71	**1.14**	0.44	0.93	0.72	0.42	0.73	0.045	0.40	1.80
-D3(op)	**0.29**	0.33	0.44	1.29	0.30	0.55	**0.56**	0.10	0.59	**0.034**	**0.26**	**0.71**
revPBE	3.61	3.68	6.45	21.85	14.82	9.25	1.76	0.60	1.66	0.238	3.88	
-D3(0)	**0.41**	0.56	0.61	5.50	2.71	1.12	0.81	0.08	0.84	0.037	0.42	6.67
-D3(BJ)	0.48	0.60	0.62	6.56	3.51	1.25	0.81	0.39	0.81	0.056	0.43	14.08
-D3(op)	0.42	0.39	0.40	**2.44**	0.42	1.06	**0.75**	0.41	0.74	**0.031**	**0.30**	**2.36**
revPBE0	3.08	3.14	5.45	16.52	12.07	7.78	1.51	0.58	1.41	0.229	3.38	
-D3(0)	0.32	0.34	0.36	2.14	1.56	0.53	0.59	0.26	0.62	0.032	0.40	1.32
-D3(BJ)	**0.32**	0.37	0.32	3.44	2.53	0.77	**0.56**	0.09	0.56	0.037	**0.23**	2.26
-D3(op)	0.33	0.26	0.39	**0.83**	0.33	0.80	0.56	0.09	0.55	**0.024**	0.29	**0.37**
TPSS	2.53	2.49	4.61	8.81	3.64	6.48	1.22	0.47	1.08	0.217	2.90	
-D2	0.63	0.48	0.64	4.74	5.67	0.89	1.07	0.89	1.03	0.118	0.77	37.47
-D3(0)	0.34	0.25	0.44	2.75	3.90	0.74	**0.67**	0.38	0.53	0.059	0.35	3.73
-D3(BJ)	0.36	0.26	0.47	1.90	2.85	0.96	0.70	0.55	0.57	0.074	0.34	3.51
-D3(CSO)	0.38	0.25	0.43	1.85	2.69	0.90	0.69	0.56	0.59	0.075	0.33	3.62
-D3(op)	**0.32**	0.31	0.36	**1.76**	2.54	0.87	0.78	0.45	0.67	**0.053**	**0.28**	**2.28**
TPSSh	2.46	2.42	4.43	8.33	3.97	6.24	1.19	0.40	1.06	0.214	2.82	
-D3(0)	0.34	0.20	0.49	2.59	3.35	0.74	**0.56**	0.24	0.46	0.056	0.36	2.77
-D3(BJ)	0.35	0.23	0.44	1.49	2.13	0.86	0.63	0.42	0.53	0.072	0.33	2.37
-D3(op)	**0.30**	0.27	0.35	**1.41**	2.02	0.80	0.69	0.31	0.59	**0.048**	**0.28**	**1.41**
MS2	1.24	1.28	2.36	4.83	2.77	3.48	0.68	0.29	0.67	0.126	1.70	
-D3(0)	0.38	0.38	0.41	1.38	1.24	0.62	0.49	0.09	0.41	**0.034**	0.43	0.86
-D3(op)	**0.29**	0.19	0.43	**0.91**	0.29	0.50	**0.35**	0.11	0.33	0.039	**0.23**	**0.36**
MS2h	1.27	1.31	2.36	4.49	2.84	3.50	0.68	0.34	0.67	0.130	1.72	
-D3(0)	0.34	0.33	0.34	1.68	1.14	0.59	0.43	0.12	0.34	0.033	0.39	0.80
-D3(op)	**0.26**	0.13	0.32	**0.56**	0.21	0.43	**0.27**	0.14	0.24	**0.027**	**0.20**	**0.11**

Figure 2: Root-mean-square errors across various datasets for all combinations of functionals and dispersion corrections examined. Abbreviations of the dataset names are used; the full names are, in order, NCEDTest, HSG, S22, NCECTest, Shields38, 3B-69-TRIM, IETest, Pentane14, YMPJ519, S66x8 BL, and S66x8 BE. The Product column corresponds to a simple product across the aggregate dataset RMSEs (times 100). All RMSEs are in units of kcal/mol, with the exception of the equilibrium binding lengths (BL), which are in units of angstroms. Each column within every functional block is color-coded for ease of reading, with darker cells corresponding to lower RMSEs; note, however, that the color-gradient is not uniform, but rather skewed to emphasize significant differences among the top fits for each functional. Additionally, the lowest RMSE achieved by any functional-dispersion combination on each aggregate dataset is indicated in bold.

lower RMSEs across all metrics related to non-covalent interactions; the reductions in errors on molecular clusters (NCEC, 63%) and geometries (BL, 45%) are particularly striking. This sort of improvement highlights the benefits a simple re-optimization incorporating the plethora of new high quality data that have been published in the past six years can bring; after all, revPBE-D3(op) is effectively just a reparameterization of revPBE-D3(BJ). Note that for this particular functional, there is no DFT-D3M(BJ) version with which to compare our reparameterization; Smith et al.[25] chose to parameterize PBE instead of revPBE.

Another manifestation of the importance of having a well-balanced training set is the performance of DFT-D3(CSO). In a recent study, it was found that DFT-D3(CSO) reproduces bond lengths and rotational constants quite well.[108] Here, we find that this satisfactory intramolecular performance does not transfer to intermolecular metrics, despite the fact that DFT-D3(CSO) is decent at reproducing accurate energies (particularly isomerization energies, at which it seems to excel). Regardless of density functional, -D3(CSO) exhibits significantly larger errors in equilibrium binding lengths than any of the other iterations of -D3; for B3LYP-D3(CSO), for instance, the RMSE across BL is nearly double the next-highest B3LYP-D3 BL RMSE. This is likely a consequence of the fact that DFT-D3(CSO) was trained exclusively on equilibrium systems, specifically S66. This is in stark contrast to other DFT-D3 variants, which were trained on geometries either implicitly through the inclusion of nonequilibrium systems – as was the case for -D3(0), -D3(BJ), and -D3M(BJ) – or by explicitly including geometries in the training metric, as is the case here for -D3(op).

Similarly, deviations among performances across single datasets highlight the need for diverse aggregate datasets. From the exemplary performance of BLYP-D3(CSO) on the Shields38 set, one might infer that this method should be particularly well-suited to water clusters. This is not the case, however; BLYP-D3(CSO) has the highest RMSE of all BLYP-D3 variants across the H2O20Bind10 set of large water clusters by nearly a factor of two. Oftentimes, outstanding performance of a particular method on a particular set of systems will transfer to similar systems, but that sort of transferability is not guaranteed.

The larger and more diverse the independent test set, the more likely the results will be transferable.

This point is further driven home by the relative performances of the DFT-D3(BJ) and -D3M(BJ) methods. For instance, compare B3LYP-D3(BJ) and B3LYP-D3M(BJ) in Figure 2. The -D3M(BJ) method is simply a reparameterization of -D3(BJ) across a significantly larger training set. In the case of B3LYP-D3M(BJ), however, this reparameterization has resulted in significant loss of performance as measured by RMSEs across equilibrium non-covalent interactions. This is likely a consequence of the large emphasis placed on compressed geometries in the parameterization of -D3M(BJ) by both the composition of the datasets and the unique error metric employed.

A consistent feature of functionals employing the DFT-D3(op) tail is enhanced performance on molecular clusters and geometries, as illustrated by low RMSEs across the NCECTest and S66x8 BL sets. For instance, B97-D3(op) performs similarly to B97-D3(BJ) and B97-D3(0) across NCEDTest and IETest, but with dramatic reduction in error across NCECTest and interpolated S66x8 binding lengths. Even in cases where performance on molecular dimers is significantly improved, e.g. B3LYP-D3(op) and MS2h-D3(op), we still see solid performances on clusters and geometries. Whereas with other dispersion tails, performance on molecular clusters is very dependent on the choice of base functional, the -D3(op) approach consistently captures cluster binding energies quite well. For instance, from Figure 2, it is clear that BLYP-D3(0) is the best BLYP-based approach examined on NCECTest, even beating out BLYP-D3(op) by 0.10 kcal/mol; on the other hand, B97-D3(0) has a NCECTest RMSE more than double that of B97-D3(op). This high degree of consistency in the new correction, -D3(op), is likely attributable to the incorporation of molecular clusters in the training set. It is worth noting that a three-body correction of the form suggested by Grimme et al.[10] was also considered in this study; its inclusion, however, did not impact results in any meaningful way. It is quite possible that for larger systems, a three-body term would become relevant.

The one category for which we consistently see the least improvement with the new DFT-D3(op) correction is IETest, i.e. isomerization energies. For certain functionals – most notably MS2 and MS2h – the -D3(op) tail actually improves performance on IETest dramatically. For others, such as TPSS and B3LYP, the new approach is slightly worse than the existing tails – though still significantly better than the base functionals – for isomerization energies. There is inevitably some degree of trade-off between performances across these various properties; of the 11767 combinations of parameters we examined for any given functional, there were certain fits which excelled at one particular category. Although in theory it would be possible to recommend specialized fits for each functional which are best suited for a particular property – B3LYP-D3(op,NCED), B3LYP-D3(op,NCEC), B3LYP-D3(op,IE), etc. – that sort of approach has limited utility. It is far simpler to recommend a parameterization that provides a balanced description of all relevant properties. Comparing B3LYP-D3(op) and B3LYP-D3(BJ), for instance, we see the former has 8% higher RMSE across IETest; however, this slight reduction in performance for isomerization energies is compensated by a 13% reduction in RMSE across NCEDTest and a 52% reduction in RMSE across NCECTest.

Based on the data in Figure 2, we can make recommendations regarding which fit should be used with each density functional. As a metric of optimality, we use a simple product of the RMSEs across each of the aggregate test sets in Figure 2, i.e. the Product column. It is worth noting that -D2 dramatically underperforms relative to all variants of -D3 for all functionals examined. That being said, even the -D2 corrections represent a major improvement over the base functional. For BLYP, B3LYP, TPSS, and TPSSh, the DFT-D3(op) correction is best, followed by -D3(BJ); for B97, the DFT-D3(op) correction is best, followed by -D3M(BJ); and for revPBE, revPBE0, MS2, and MS2h, DFT-D3(op) is best, followed by -D3(0).

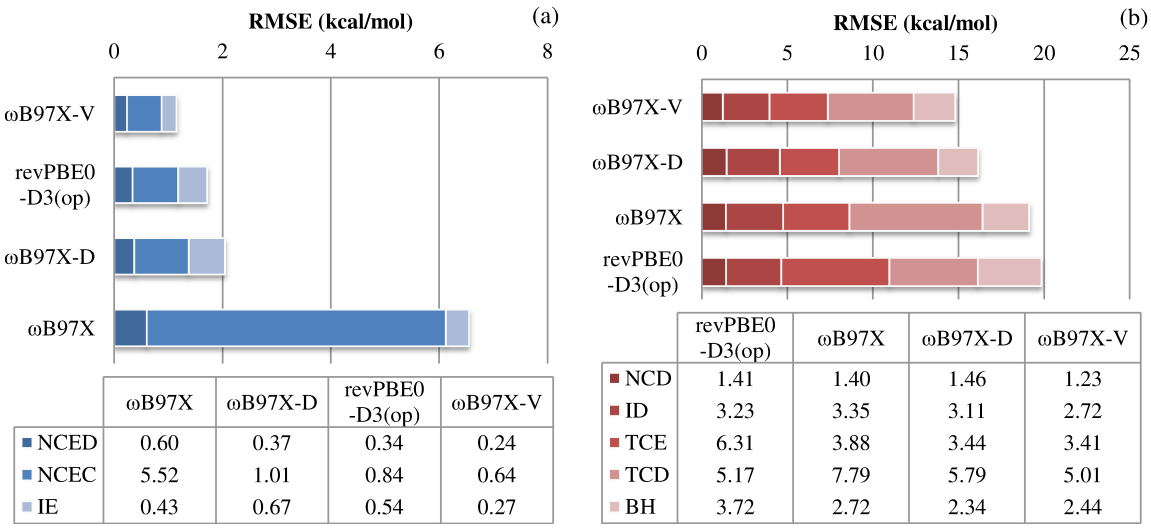
In a similar vein, we can use the data from Figure 2 to make recommendations for the "best" functional/tail combination within each rung of Jacob's ladder. The top performer among pure GGAs is BLYP-D3(op); at the hybrid GGA level, revPBE0-D3(op); from the pure meta-GGAs, MS2-D3(op); and at the hybrid meta-GGA rung, MS2h-D3(op). Based

on the data of Figure 2, these four functional/tail groupings provide the best, most balanced descriptions of non-covalent interactions and isomerization energies, the two datatypes which empirical dispersion corrections should be capable of improving.

Thus far, we have established two major points: first, when a functional consistently underpredicts intermolecular interactions, the addition of an empirical dispersion correction can greatly improve results; and second, among the dispersion corrections examined, for the systems and density functionals we considered, DFT-D3(op) is best, regardless of the choice of base functional.

To put these results in broader context, we can compare these "best-in-class" DFT-D3 functionals with state-of-the-art density functionals corresponding to the same rungs of Jacob's ladder. In Figure 3, the RMSEs across all eight categories of Table 3 are plotted and tabulated for revPBE0-D3(op) and three other hybrid GGAs: ωB97X,[109] ωB97X-D,[22] and ωB97X-V.[110] The left side of Figure 3, (a), contains NCED, NCEC, and IE – the three categories a D3 correction is capable of improving – while the right side, (b), encompasses the remaining datatypes: difficult cases characterized by strong correlation or self-interaction error, as well as thermochemistry and barrier heights. Although revPBE0-D3(op) represents a substantial improvement over ωB97X as far as non-covalent interactions are concerned, the method is somewhat lacking when it comes to the other datatypes: its performance on standard thermochemical problems (TCE) is particularly lackluster. That being said, it represents a decent alternative to ωB97X-D, and even rivals ωB97X-V for non-covalent interactions.

The same sort of comparison is made at the pure meta-GGA level in Figure 4. Therein, the top performer from this study, MS2-D3(op), is compared to M06-L,[111] TM,[112] and B97M-rV.[43,113] Here, it seems MS2-D3(op) is actually a viable alternative to standard meta-GGAs; the method generally outperforms both M06-L and TM across the 4419 systems examined. We see a 47% reduction in RMSE across NCED, a 60% reduction in error across NCEC, and a 51% reduction in RMSE across IE for MS2-D3(op) relative to M06-L.



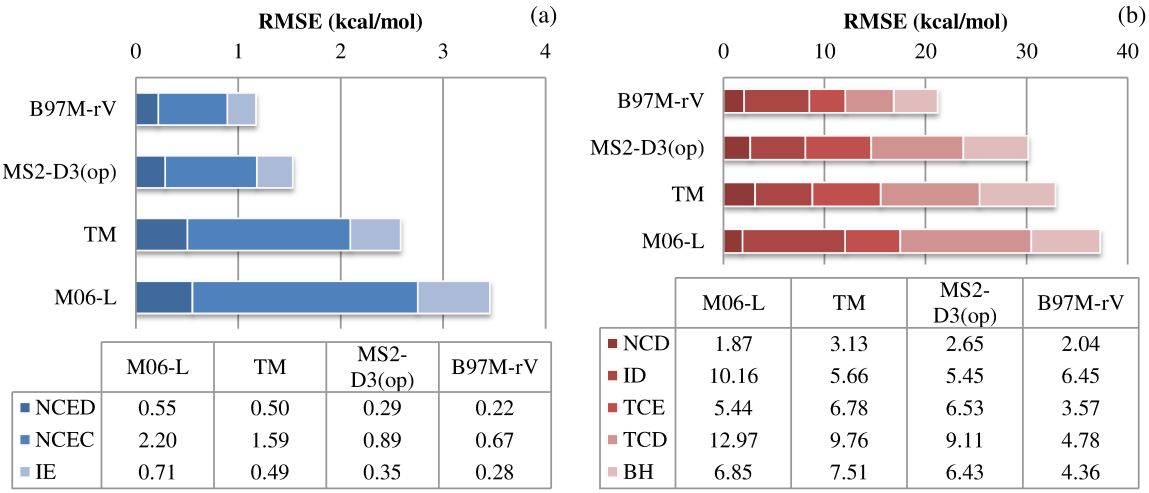
	ωB97X	ωB97X-D	revPBE0 -D3(op)	ωB97X-V
NCED	0.60	0.37	0.34	0.24
NCEC	5.52	1.01	0.84	0.64
IE	0.43	0.67	0.54	0.27

	revPBE0 -D3(op)	ωB97X	ωB97X-D	ωB97X-V
NCD	1.41	1.40	1.46	1.23
ID	3.23	3.35	3.11	2.72
TCE	6.31	3.88	3.44	3.41
TCD	5.17	7.79	5.79	5.01
BH	3.72	2.72	2.34	2.44

Figure 3: Root-mean-square errors (RMSEs) across various datatypes (see Table 3) for the top-performing D3-corrected hybrid GGA functional examined in this work, as well as three other state-of-the-art hybrid GGA functionals. The left plot – (a) – contains datasets pertaining to standard non-covalent interactions and isomerization energies, whereas the right plot – (b) – encompasses other data, such as thermochemistry, which is beyond the realm of a D3 correction. Within each plot, the methods are ordered from best performance across the constituent datasets at the top, to worst performance at the bottom. Tables of RMSEs are provided below the bar graphs to facilitate quantitative comparison.

This strikes a stark contrast to the performance of the uncorrected MS2 functional, which, relative to M06-L, has a NCED RMSE of 1.27 kcal/mol, a NCEC RMSE of 4.99 kcal/mol, and an IE RMSE of 0.76 kcal/mol; without the DFT-D3(op) correction, MS2 is significantly worse than M06-L across all three categories. This is certainly a testament to the utility of dispersion corrections. That being said, the exemplary performance of MS2-D3(op) – particularly on non-covalent interactions and isomerization energies – is eclipsed by that of B97M-rV, which benefits from the inclusion of non-local van der Waals correlation.

As a final example, three popular hybrid meta-GGA functionals – M06-2X,[114] MN15,[115] and ωB97M-V[43] – are compared head-to-head with MS2h-D3(op) in Figure 5. Although the DFT-D3(op) dispersion correction allows MS2h to significantly outperform M06-2X and MN15 for standard intermolecular binding energies and isomerization energies – and even come close to the stellar performance of ωB97M-V – it can do nothing to rectify the other deficiencies of the method, which manifest themselves in poor performance across the TCE



	M06-L	TM	MS2-D3(op)	B97M-rV
NCED	0.55	0.50	0.29	0.22
NCEC	2.20	1.59	0.89	0.67
IE	0.71	0.49	0.35	0.28

	M06-L	TM	MS2-D3(op)	B97M-rV
NCD	1.87	3.13	2.65	2.04
ID	10.16	5.66	5.45	6.45
TCE	5.44	6.78	6.53	3.57
TCD	12.97	9.76	9.11	4.78
BH	6.85	7.51	6.43	4.36

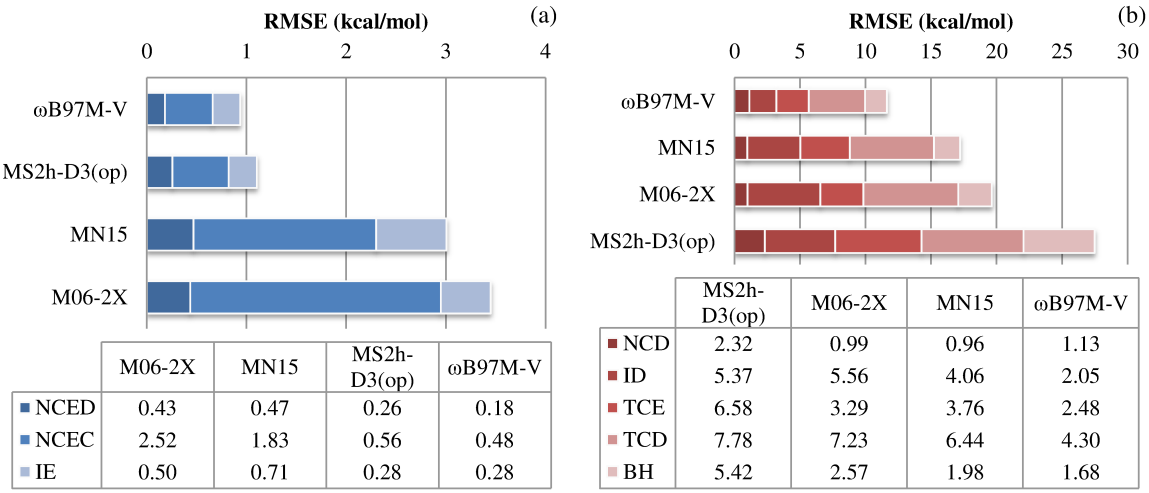
Figure 4: Root-mean-square errors (RMSEs) across various datatypes (see Table 3) for the top-performing D3-corrected pure meta-GGA functional examined in this work, as well as three other state-of-the-art pure meta-GGA functionals. For further details, refer to the caption of Figure 3.

and BH categories in particular. At the hybrid meta-GGA level, the top performer is unambiguously ωB97M-V.

5 Discussion and Conclusion

In this study, we have introduced a new damping function for the DFT-D3 brand of empirical dispersion corrections. This correction is effectively a generalization of DFT-D3(BJ) wherein the power is treated as an additional parameter, and is accordingly named DFT-D3(op), for "optimized-power"-damping. We have parameterized this method for ten distinct density functionals and compared its performance across an external test set to that of existing forms of DFT-D3.

This new approach, -D3(op), consistently yields substantial improvements in the descriptions of molecular clusters, regardless of the base functional with which it is paired. Moreover, it provides a well-balanced description of intermolecular binding energies, equilibrium geometries, and isomerization energies – the three broad classes of data that an empirical dispersion correction can reasonably be expected to improve. Unfortunately, the DFT-D3



	M06-2X	MN15	MS2h-D3(op)	ωB97M-V
NCED	0.43	0.47	0.26	0.18
NCEC	2.52	1.83	0.56	0.48
IE	0.50	0.71	0.28	0.28

	MS2h-D3(op)	M06-2X	MN15	ωB97M-V
NCD	2.32	0.99	0.96	1.13
ID	5.37	5.56	4.06	2.05
TCE	6.58	3.29	3.76	2.48
TCD	7.78	7.23	6.44	4.30
BH	5.42	2.57	1.98	1.68

Figure 5: Root-mean-square errors (RMSEs) across various datatypes (see Table 3) for the top-performing D3-corrected hybrid meta-GGA functional examined in this work, as well as three other state-of-the-art pure hybrid meta-GGA functionals. For further details, refer to the caption of Figure 3.

correction is not intended to address deficiencies in other aspects – for instance, description of thermochemical properties – of the base functionals to which it is applied. From Figures 3, 4, and 5, it is apparent that there exists at each level of Jacob's ladder beyond pure GGA at least one functional which will outperform the best comparable DFT-D3 method. This is not particularly surprising; after all, this isn't a very fair comparison. Those better methods were built from the ground up; all their components were optimized together, self-consistently. If anything, it is remarkable just how well this sort of simple post-SCF dispersion correction performs.

The most impressive methods – namely ωB97X-V, B97M-rV, and ωB97M-V – explicitly include non-local correlation, which is a more robust, more powerful treatment of dispersion.[116,117] Similarly, it might be the case that other self-consistent approaches, such as dispersion-corrected potentials[11] and TS-vdW,[118] will outperform the methods introduced here. Such comparisons are beyond the scope of this particular study; the aim here has been simply to introduce an updated, more accurate version of the most computationally economical form of dispersion correction, DFT-D3. In this particular endeavor, we have been successful: for all density functionals examined, DFT-D3(op) yields the best, most balanced

performance across the available DFT-D3 dispersion tails.

Specifically, among the pure GGAs examined, BLYP-D3(op) is the stand-out candidate, offering an unparalleled account of non-covalent interactions. At the hybrid GGA level, revPBE0-D3(op) shines. As far as pure and hybrid meta-GGA functionals go, MS2-D3(op) and MS2h-D3(op) are both quite impressive. For these latter two functionals in particular, the new DFT-D3(op) tail represents a major improvement across all energetic categories over the existing -D3(0) versions. These DFT-D3(op) methods have a significant cost advantage over those incorporating VV10 non-local correlation. Moreover, they make relatively small sacrifices on performance across non-covalent interactions and isomerization energies, which makes them well suited for certain applications, e.g. calculations on large or condensed-phase systems, and molecular dynamics.

Acknowledgement

This research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-12ER16362. This work was also supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, and a subcontract from MURI Grant W911NF-14-1-0359.

References

(1) Kohn, W.; Sham, L. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(2) Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. AIP Conf. Proc. Melville, NY, 2001; pp 1–20.

(3) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(4) Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.

(5) Hepburn, J.; Scoles, G.; Penco, R. *Chem. Phys. Lett.* **1975**, *36*, 451–456.

(6) Ahlrichs, R.; Penco, R.; Scoles, G. *Chem. Phys.* **1977**, *19*, 119–130.

(7) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515.

(8) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

(9) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(10) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

(11) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.

(12) DiLabio, G. A. *Chem. Phys. Lett.* **2008**, *455*, 348–353.

(13) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, 154104.

(14) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.

(15) Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* **2005**, *123*, 24101.

(16) Johnson, E. R.; DiLabio, G. A. *Chem. Phys. Lett.* **2006**, *419*, 333–339.

(17) Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(18) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(19) Lee, K.; Murray, D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.

(20) Vydrov, O.; Van Voorhis, T. *Phys. Rev. Lett.* **2009**, *103*, 063004.

(21) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.

(22) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–20.

(23) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(24) Schröder, H.; Creon, A.; Schwabe, T. *J. Chem. Theory Comput.* **2015**, *11*, 3163–3170.

(25) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *J. Phys. Chem. Lett.* **2016**, *7*, 2197–2203.

(26) Becke, A. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(27) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(28) Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*, 785–789.

(29) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(30) Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(31) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(32) Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890–890.

(33) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889.

(34) Tao, J.; Perdew, J.; Staroverov, V.; Scuseria, G. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(35) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129.

(36) Sun, J.; Haunschild, R.; Xiao, B.; Bulik, I. W.; Scuseria, G. E.; Perdew, J. P. *J. Chem. Phys.* **2013**, *138*, 1–9.

(37) Gillan, M. J.; Alfè, D.; Michaelides, A. *J. Chem. Phys.* **2016**, *144*, 130901.

(38) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(39) Rappoport, D.; Furche, F. *J. Chem. Phys.* **2010**, *133*, 134105.

(40) Witte, J.; Neaton, J. B.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 194306.

(41) Gill, P. M.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(42) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W.; Harbach, P. H.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. a.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.;

Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M. *Mol. Phys.* **2015**, *113*, 184–215.

(43) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2016**, *144*, 214110.

(44) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

(45) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 3466–3470.

(46) Boese, A. D. *J. Chem. Theory Comput.* **2013**, *9*, 4403–4413.

(47) Boese, A. D. *Mol. Phys.* **2015**, *113*, 1–12.

(48) Boese, A. D. *ChemPhysChem* **2015**, *16*, 978–985.

(49) Granatier, J.; Pitoňák, M.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 2282–2292.

(50) Lao, K. U.; Herbert, J. M. *J. Chem. Phys.* **2013**, *139*, 034107.

(51) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.

(52) Bryantsev, V. S.; Diallo, M. S.; Van Duin, A. C. T.; Goddard, W. A. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.

(53) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.

(54) Fanourgakis, G. S.; Aprà, E.; Xantheas, S. S. *J. Chem. Phys.* **2004**, *121*, 2655–2663.

(55) Anacker, T.; Friedrich, J. *J. Comput. Chem.* **2014**, *35*, 634–643.

(56) Kozuch, S.; Bachrach, S. M.; Martin, J. M. L. *J. Phys. Chem. A* **2014**, *118*, 293–303.

(57) Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. *J. Phys. Chem. A* **2013**, *117*, 2269–2277.

(58) Yoo, S.; Aprà, E.; Zeng, X. C.; Xantheas, S. S. *J. Phys. Chem. Lett.* **2010**, *1*, 3122–3127.

(59) Crittenden, D. L. *J. Phys. Chem. A* **2009**, *113*, 1663–1669.

(60) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *135*, 194102.

(61) Hohenstein, E. G.; Sherrill, C. D. *J. Phys. Chem. A* **2009**, *113*, 878–886.

(62) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146–10159.

(63) Takatani, T.; David Sherrill, C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6106–6114.

(64) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.

(65) Mintz, B. J.; Parks, J. M. *J. Phys. Chem. A* **2012**, *116*, 1086–1092.

(66) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.

(67) Faver, J. C.; Benson, M. L.; He, X.; Roberts, B. P.; Wang, B.; Marshall, M. S.; Kennedy, M. R.; Sherrill, C. D.; Merz, K. M. *J. Chem. Theory Comput.* **2011**, *7*, 790–797.

(68) Jurecka, P.; Sponer, J.; Cerný, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(69) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141–151.

(70) Witte, J.; Goldey, M.; Neaton, J. B.; Head-Gordon, M. *J. Chem. Theory Comput.* **2015**, *11*, 1481–1492.

(71) Copeland, K. L.; Tschumper, G. S. *J. Chem. Theory Comput.* **2012**, *8*, 1646–1656.

(72) Smith, D. G. A.; Jankowski, P.; Slawik, M.; Witek, H. A.; Patkowski, K. *J. Chem. Theory Comput.* **2014**, *10*, 3140–3150.

(73) Řezáč, J.; Huang, Y.; Hobza, P.; Beran, G. J. O. *J. Chem. Theory Comput.* **2015**, *11*, 3065–3079.

(74) Li, S.; Smith, D. G. A.; Patkowski, K. *Phys. Chem. Chem. Phys.* **2015**, *17*, 16560–16574.

(75) Lao, K. U.; Herbert, J. M. *J. Phys. Chem. A* **2015**, *119*, 235–252.

(76) Temelso, B.; Archer, K. A.; Shields, G. C. *J. Phys. Chem. A* **2011**, *115*, 12034–12046.

(77) Mardirossian, N.; Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Head-Gordon, M. *J. Chem. Theory Comput.* **2013**, *9*, 1368–1380.

(78) Karton, A.; O'Reilly, R. J.; Chan, B.; Radom, L. *J. Chem. Theory Comput.* **2012**, *8*, 3128–3136.

(79) Chan, B.; Gilbert, A. T. B.; Gill, P. M. W.; Radom, L. *J. Chem. Theory Comput.* **2014**, *10*, 3777–3783.

(80) Tentscher, P. R.; Arey, J. S. *J. Chem. Theory Comput.* **2013**, *9*, 1568–1579.

(81) Kozuch, S.; Martin, J. M. L. *J. Chem. Theory Comput.* **2013**, *9*, 1918–1931.

(82) Bauzá, A.; Alkorta, I.; Frontera, A.; Elguero, J. *J. Chem. Theory Comput.* **2013**, *9*, 5201–5210.

(83) Otero-de-la Roza, A.; Johnson, E. R.; DiLabio, G. A. *J. Chem. Theory Comput.* **2014**, *10*, 5436–5447.

(84) Steinmann, S. N.; Piemontesi, C.; Delachat, A.; Corminboeuf, C. *J. Chem. Theory Comput.* **2012**, *8*, 1629–1640.

(85) Karton, A.; Gruzman, D.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 8434–8447.

(86) Gruzman, D.; Karton, A.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 11974–11983.

(87) Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Csaszar, A. G.; Allen, W. D. *J. Chem. Theory Comput.* **2009**, *5*, 1511–1523.

(88) Martin, J. M. L. *J. Phys. Chem. A* **2013**, *117*, 3118–3132.

(89) Kesharwani, M. K.; Karton, A.; Martin, J. M. L. *J. Chem. Theory Comput.* **2016**, *12*, 444–454.

(90) Yu, L.-J.; Karton, A. *Chem. Phys.* **2014**, *441*, 166–177.

(91) Karton, A.; Daon, S.; Martin, J. M. L. *Chem. Phys. Lett.* **2011**, *510*, 165–178.

(92) Manna, D.; Martin, J. M. L. *J. Phys. Chem. A* **2016**, *120*, 153–160.

(93) Curtiss, L. A.; Raghavachari, K. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(94) Zhao, Y.; González-Garda, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.

(95) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.

(96) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384–1388.

(97) Grimme, S.; Kruse, H.; Goerigk, L.; Erker, G. *Angew. Chemie - Int. Ed.* **2010**, *49*, 1402–1405.

(98) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

(99) Krieg, H.; Grimme, S. *Mol. Phys.* **2010**, *108*, 2655–2666.

(100) O'Reilly, R. J.; Karton, A. *Int. J. Quantum Chem.* **2016**, *116*, 52–60.

(101) Karton, A.; Martin, J. M. L. *Mol. Phys.* **2012**, *110*, 19–20.

(102) Karton, A.; Schreiner, P. R.; Martin, J. M. L. *J. Comput. Chem.* **2016**, *37*, 49–58.

(103) Karton, A.; Goerigk, L. *J. Comput. Chem.* **2015**, *36*, 622–632.

(104) Yu, L.-J.; Sarrami, F.; O'Reilly, R. J.; Karton, A. *Chem. Phys.* **2015**, *458*, 1–8.

(105) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 569–582.

(106) Karton, A.; Tarnopolsky, A.; Lamère, J. F.; Schatz, G. C.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.

(107) Yu, L.-J.; Sarrami, F.; O'Reilly, R. J.; Karton, A. *Mol. Phys.* **2016**, *114*, 21–33.

(108) Schröder, H.; Hühnert, J.; Schwabe, T. *J. Chem. Phys.* **2017**, *146*, 044115.

(109) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

(110) Mardirossian, N.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

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

(112) Tao, J.; Mo, Y. *Phys. Rev. Lett.* **2016**, *117*, 73001.

(113) Mardirossian, N.; Head-Gordon, M. *J. Chem. Phys.* **2015**, *142*, 074111.

(114) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(115) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. *Chem. Sci.* **2016**, *7*, 5032–5051.

(116) Goerigk, L. *J. Chem. Theory Comput.* **2014**, *10*, 968–980.

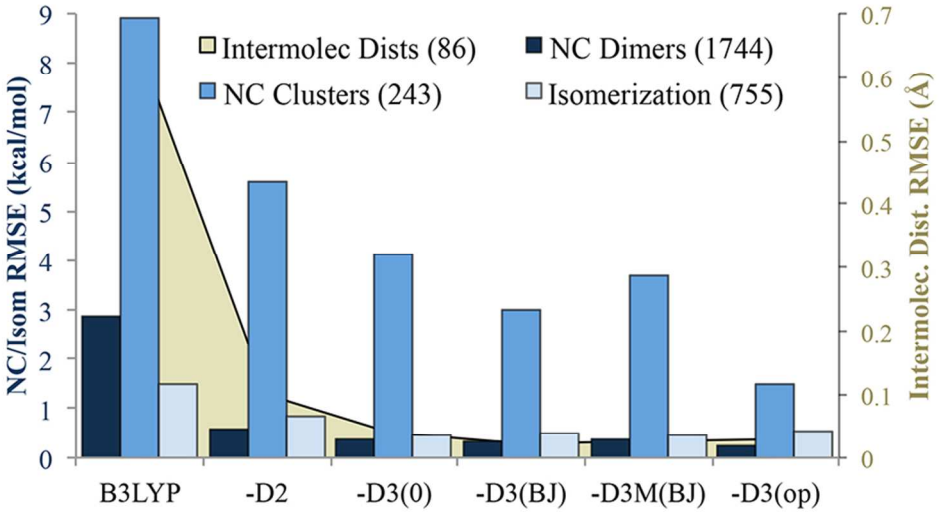
(117) Hujo, W.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 3866–3871.

(118) Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012**, *108*, 236402.

Supporting Information Available

DFT-D3(op) parameterizations for PBE, PBE0, revTPSS, and revTPSSh may be found online, as may data for rare gas dimers and maximum errors for each method across the various datasets.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.



For Table of Contents Only

82x44mm (300 x 300 DPI)