PCCP

Spin-component-scaled and dispersion-corrected second-order Møller-Plesset perturbation theory: A path toward chemical accuracy

SCHOLARONE™
Manuscripts

Journal Name

Spin-component-scaled and dispersion-corrected second-order Møller-Plesset perturbation theory: A path toward chemical accuracy[†]

Chandler Greenwell,[a] Jan Řezáč,[b‡] and Gregory J. O. Beran[*a]

Second-order Møller-Plesset perturbation theory (MP2) provides a valuable alternative to density functional theory for modeing problems in organic and biological chemistry. However, MP2 suffers from known limitations in the description of van der Waals dispersion interactions and reaction thermochemistry. Here, a spin-component-scaled, dispersion-corrected MP2 model (SCS-MP2D) is proposed that addresses these weaknesses. The dispersion correction, which is based on Grimme's D3 formalism, replaces the uncoupled Hartree-Fock dispersion inherent in MP2 with a more robust coupled Kohn-Sham treatment. The spin-component scaling of the residual MP2 correlation energy then reduces the remaining errors in the model. This two-part correction strategy solves the problem found in earlier spin-component-scaled MP2 models where completely different spin-scaling parameters were needed for describing reaction energies versus intermolecular interactions. Results on 18 benchmark data sets and two challenging potential energy curves demonstrate that SCS-MP2D considerably improves upon the accuracy of MP2 for intermolecular interactions, conformational energies, and reaction energies. Its accuracy and computational cost are competitive with state-of-the-art density functionals such as DSD-BLYP-D3(BJ), revDSD-PBEP86-D3(BJ), ωB97X-V, and ωB97M-V for systems with $\sim$100 atoms.

1 Introduction

Accurately modeling many chemically-interesting systems with electronic structure theory requires models capable of describing diverse mixtures of covalent and non-covalent interactions. Chemical reactions occurring in enzyme active sites demand models that can treat the thermochemistry associated with changes in the substrate chemical bonding together with the hydrogen bonding, electrostatic, and dispersion interactions that govern the substrate-protein interaction.[1] Furthermore, the stabilities of molecular crystal conformational polymorphs are governed by the competition between intramolecular conformation and intermolecular packing.[2] In principle, high-accuracy methods like coupled cluster singles, doubles, and perturbative triples (CCSD(T)) can provide the requisite accuracy for modeling systems like these, but the steep $O(N^7)$ computational cost with system size N frequently makes it cost-prohibitive in practice.

Instead, Kohn-Sham density functional theory (DFT) has become the standard tool of choice for modeling such systems. Many successful density functionals have been developed over the years, some of which can approach CCSD(T) accuracy. Large benchmark studies[3–5] have identified some of the best current functionals. These include, for example, the hierarchy of functionals developed by Mardirossian and Head-Gordon: the range-separated hybrid functional ωB97X-V,[6] its hybrid meta-GGA variant ωB97M-V,[7] and the double-hybrid meta-GGA functional ωB97M(2).[8] The family of dispersion-corrected, spin-component-scaled double hybrid (DSD) density functionals developed in the Martin group are also highly competitive, both in their original[9] and recently-revised forms.[10]

At the same time, there have been many efforts to achieve near-coupled cluster accuracy using wave function methods that are less computationally demanding than CCSD(T). The domain-based local pair natural orbital variant (DLPNO-CCSD(T))[11] achieves most of the accuracy of CCSD(T) at far lower computational cost, for example. At the other extreme, inexpensive machine learning models that target CCSD(T) accuracy continue to improve.[12,13] Models based on second-order Møller-Plesset perturbation theory (MP2), the least-expensive correlated wavfunction method, have also garnered considerable attention over the years. Although MP2 scales $O(N^5)$ with system size, efficient

[a] Department of Chemistry, University of California Riverside, Riverside CA 92521, USA. E-mail: gregory.beran@ucr.edu

[b] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 166 10 Prague, Czech Republic.

density-fitting algorithms mean that the computational cost of evaluating the non-iterative MP2 correlation energy is small compared to that associated with the underlying iterative Hartree-Fock (HF) self-consistent field energy for many systems. This means that MP2 can be computationally competitive with modern state-of-the-art density functionals for systems with up to ∼100 atoms.

Furthermore, MP2 inherently includes exact exchange and does not suffer from the issues of self-interaction error/delocalization error that plagues approximate density functionals. [14,15] Delocalization error leads to the underestimation of dissociation energies [16–18] and barrier heights, [19–22] for example. In the context of organic molecular crystals, delocalization error can cause spurious proton transfer in multi-component crystals [23] and problematic conformational energies that produce highly incorrect polymorph stabilities. [24–27]

Despite its advantages, MP2 has its own significant deficiencies which manifest in systems with strong static correlation (e.g. stretched covalent bonds) or in systems where van der Waals dispersion interactions are important (e.g. benzene dimer). The second-order correlation energy can be partitioned into contributions arising from same-spin ($\alpha\alpha$ and $\beta\beta$) and opposite-spin ($\alpha\beta$ and $\beta\alpha$) electron pairs. The same-spin contributions are more connected with long-range static correlation, while the opposite-spin ones are more important for the dynamic correlation that is associated with dispersion. The MP2 perturbation series is biased toward the same-spin correlation and frequently overestimates its contribution. [28] In 2003, Grimme's spin-component-scaled MP2 (SCS-MP2) model [29] demonstrated how scaling the same-spin and opposite-spin energy components of the MP2 correlation energy with constant coefficients improves the accuracy of MP2 on systems that would otherwise be poorly described, without any increase in the computational effort required. The SCS-MP2 prescription greatly improves upon canonical MP2 for predicting reaction thermochemistry.

However, it soon became apparent that the spin-component scaling coefficients appropriate for one type of chemical problem do not always transfer well to other chemical problems. For example, the original SCS-MP2 model scales the same-spin correlation energy by $c_{ss} = 1/3$, while the opposite-spin correlation energy is scaled up by $c_{os} = 6/5$. [29] These values were initially determined from studying reaction energies, though subsequent work established a theoretical basis for these scaling parameter values. [28,30,31] Studying the S22 benchmark set a few years later, Distasio and Head-Gordon found optimal scaling coefficients $c_{ss} = 1.29$ and $c_{os} = 0.40$ for molecular interactions (MI), denoting the resulting model as SCS(MI)-MP2. [32] These SCS(MI)-MP2 c_{ss} and c_{os} scaling coefficients are nearly reversed compared to those found in the original SCS-MP2.

Given that no single set of spin-scaling coefficients can fully address the MP2 problems, a few strategies for SCS-MP2 models have emerged over the years. One approach tailors the SCS coefficients for specific chemical systems, such as for nucleic acid base pair interactions, [33] ethylene dimers, [34] or ionic liquids [35]. Such models can potentially work well, though this parameterization strategy inherently limits transferability of the model. Another

approach adapts the spin-component scaling coefficients to each given system on the fly. This has been done via spin-ratio scaled spin components (SRS-MP2) [36] or by machine learning the optimal scaling parameters as in SNS-MP2. [13] Adaptive spin-scaling approaches can be more universal, though care must be taken to ensure that the coefficient adaptations retain smooth and continuous potential energy surfaces. Moreover, adaptive schemes can still have limited applicability: the design of the neural network-based SNS-MP2 model limits its application to dimer intermolecular interactions, for instance. Other SCS models seek to exploit the greater computational efficiency associated with the opposite-spin correlation, [37] to improve the long-range behavior of SCS methods, [38] or to apply these ideas to higher-levels of theory such as MP3, [39] coupled cluster models, [40,41] and excited state approaches. [42–45]

The fundamental challenge for existing spin-component-scaled MP2 methods is that they attempt to use spin component scaling to address multiple, physically distinct weaknesses inherent in MP2 simultaneously. These limitations can generally be partitioned into the suitability of MP2 pair correlations for covalent bond chemistry (termed "thermochemistry" here for simplicity) versus the problems associated with describing van der Waals dispersion in non-covalent interactions. The former typically include more significant amounts of static correlation energy, while the dispersion interaction arises from dynamical correlation. The differences between the optimal scaling coefficients in SCS-MP2 and SCS(MI)-MP2 highlight the challenge associated with addressing both problems simultaneously with spin-component scaling.

In the language of intermolecular perturbation theory, the dispersion problem arises from the uncoupled Hartree-Fock (UCHF) description of dispersion that is inherent in MP2. [46] Hesselmann's corrected MP2 model (MP2C), [47,48] addresses this by subtracting out the UCHF dispersion energy and replacing it with a better treatment computed at the coupled Kohn-Sham (CKS) level of theory. MP2C has proved very successful, [49] though its intermolecular perturbation theory formulation limits its application to dimer intermolecular interactions. We recently developed a new version of MP2C, called MP2D, [50] which recasts the MP2C dispersion correction in terms of Grimme's D3 dispersion correction. [51] MP2D is similar to the MP2 plus van der Waals approach proposed by Tkatchenko et al; [52] however, use of the D3 model in MP2D makes the dispersion correction more straightforward to compute.

The MP2D dispersion correction computes the CKS and UCHF dispersion contributions using atom-centered dispersion coefficients which are inexpensively interpolated from a small set of *ab initio* dispersion coefficients computed for simple hydrides of the elements. Because MP2D employs atomic dispersion coefficients, the dispersion correction is applicable to both intra- and intermolecular interactions. Its performance for intermolecular interactions is similar to that of MP2C, and it has proved very useful in describing intramolecular interactions that prove difficult for many widely used density functionals, such as in conformational polymorphs of molecular crystals. [25,26]

The present study introduces spin-component-scaled MP2D (SCS-MP2D). Because the MP2D dispersion correction already

addresses the MP2 dispersion problems for non-covalent interactions well, the spin-component scaling coefficients can focus solely on correcting the residual problems beyond dispersion that impact MP2 performance for thermochemistry. The proposed SCS-MP2D model employs seven global empirical parameters, which is a modest number of parameters compared to many models derived from big data and machine learning approaches. SCS-MP2D represents the first spin-component-scaled MP2 model where a single set of spin-scaling coefficients can describe a wide variety of intra- and intermolecular organic chemistry. It appears highly transferable to new systems, despite only a modest amount of benchmark data being used to fit the empirical parameters.

The addition of spin component scaling to MP2D makes it similar in many ways to the family of DSD double-hybrid density functionals.[9,10] Those density functionals start from a hybrid functional DFT treatment with a modest fraction of exact exchange, mix in some amount of spin-component-scaled MP2-like correlation (evaluated in terms of the Kohn-Sham orbitals), and include long-range Grimme dispersion. SCS-MP2D has its foundation in HF (instead of DFT), includes spin-component-scaled MP2 correlation from which the long-range dispersion has been removed, and Grimme D3 dispersion. Starting from the exact exchange treatment in HF circumvents the problems of DFT delocalization error. By subtracting the UCHF dispersion from the MP2 correlation energy before adding the Grimme dispersion correction, SCS-MP2D also avoids any issues of double-counting dispersion energy contribtutions that can hinder dispersion-corrected DFT models.

As will be demonstrated below, the accuracy of SCS-MP2D is competitive with some of the very best density functionals on a large set of benchmark data sets of organic species that span intermolecular interactions, conformational energies, and reaction energies. Furthermore, SCS-MP2D outperforms several state-of-the-art density functionals for two particularly challenging potential systems for which delocalization error is known to be an issue: organic crystal polymorphs of the ROY molecule and the photodimerization reaction of anthracene. The cost of SCS-MP2D is effectively identical to that of MP2, and it is ∼2–2.5 times faster than top-tier functionals such as ωB97X-V or ωB97M-V for systems approaching ∼100 atoms. Overall, the results presented below highlight how MP2-based wave function methods offer a viable route toward high-accuracy quantum chemistry in organic systems.

2 Theory

2.1 SCS-MP2D Energy

The canonical MP2 energy can be decomposed into the HF energy plus the same-spin (ss) and opposite-spin (os) correlation energies,

$$E_{MP2} = E_{HF} + c_{os}E_{corr}^{os} + c_{ss}E_{corr}^{ss} \tag{1}$$

In canonical MP2, the spin-scaling coefficients c_{os} and c_{ss} both equal one. Spin-component-scaled MP2 methods change those spin-scaling coefficients to improve the performance of the model. Grimme's original SCS-MP2 model employed $c_{os} = 6/5$ and $c_{ss} = 1/3$. As described earlier, however, these coefficients can vary con-

siderably depending on the nature of the chemical system being studied.

One of the key problems in canonical MP2 lies in its treatment of van der Waals dispersion, such as its well-known overestimation of the benzene dimer interaction energy and many other π-π interactions. The successful MP2C model addresses this for intermolecular interactions by subtracting out the UCHF dispersion that is inherent in MP2 and replacing it with a more reliable CKS description,[47,48]

$$E_{MP2C} = E_{MP2} - E_{disp}^{UCHF} + E_{disp}^{CKS} \tag{2}$$

However, the reliance on intermolecular perturbation theory for the dispersion correction limits MP2C to dimer intermolecular interactions, and the analytic nuclear gradients that would facilitate geometry optimizations are complicated and have not yet been implemented.

Our recently proposed dispersion-corrected MP2D model[50] addresses both limitations. MP2D adopts the same basic formalism as MP2C, but it computes the UCHF and CKS dispersion contributions according to Grimme's D3 strategy,[51] which estimates the atom-centered dispersion coefficients via interpolation among pre-tabulated reference values for each element in different coordination environments. The D3 dispersion correction can be computed with trivial force-field like cost, is readily differentiated for analytic nuclear gradients, and is applicable to both intra- and intermolecular interactions.

SCS-MP2D combines spin-scaling of the correlation energy with a CKS dispersion correction. The SCS-MP2D energy is given by,

$$E_{SCS\text{-}MP2D} = E_{HF} + c_{os}E_{MP2\text{-}corr}^{os} + c_{ss}E_{MP2\text{-}corr}^{ss}$$
$$- c_{os}E_{UCHF}^{disp,os} - c_{ss}E_{UCHF}^{disp,ss} + E_{CKS}^{disp,tot} \tag{3}$$

Because the MP2 correlation energies are scaled by the spin-scaling coefficients, the UCHF dispersion energy being removed from the correlation energy must also be scaled accordingly to obtain a "dispersion-free" SCS-MP2 energy. One is then free to add an appropriate dispersion treatment onto it—the D3 CKS dispersion energy in this case, without any spin-component scaling. Since the CKS dispersion contribution is unchanged compared to MP2D, the following discussion focuses on the UCHF contribution.

Spin-component scaling of the UCHF dispersion energy begins with partitioning the molecular frequency-dependent dipole-dipole polarizabilities into their α (↑) and β (↓) spin contributions,

$$\alpha_{\lambda\sigma}(i\omega) = \alpha_{\lambda\sigma}^{\uparrow}(i\omega) + \alpha_{\lambda\sigma}^{\downarrow}(i\omega)$$
$$= \sum_{ia} \frac{2\varepsilon_{ia}\langle i|\hat{\mu}_{\lambda}|a\rangle\langle a|\hat{\mu}_{\sigma}|i\rangle}{\hbar(\varepsilon_{ia}^2 + \omega^2)} + \sum_{\bar{i}\bar{a}} \frac{2\varepsilon_{\bar{i}\bar{a}}\langle \bar{i}|\hat{\mu}_{\lambda}|\bar{a}\rangle\langle \bar{a}|\hat{\mu}_{\sigma}|\bar{i}\rangle}{\hbar(\varepsilon_{\bar{i}\bar{a}}^2 + \omega^2)} \tag{4}$$

where i and a are α spin occupied and virtual orbitals, $\bar{i}$ and $\bar{a}$ are the analogous β spin orbitals, ε_{ia} is the HF energy difference between orbitals i and a, $\hat{\mu}_{\lambda}$ is the λ-th component of the dipole moment operator, and $i\omega$ is the imaginary frequency at which the polarizability is being evaluated. The isotropic frequency-dependent

polarizability α_{iso} is computed as the trace of $\alpha_{\lambda\sigma}$ divided by 3.

In the D3 approach, the isotropic molecular frequency-dependent polarizabilities are computed and tabulated for a series of elemental hydrides A_mH_n (e.g. C_2H_6, C_2H_4, C_2H_2, CH, and C). The atomic frequency-dependent polarizability for atom A is then determined by subtracting out the hydrogen contributions and distributing the remaining polarizability evenly across the heavy atoms. For the spin up contributions, this takes the form:

$$\alpha_{iso}^{A\uparrow}(i\omega) = \frac{1}{m}[\alpha_{iso}^{A_mH_n\uparrow}(i\omega) - \frac{n}{2}\alpha_{iso}^{H_2\uparrow}(i\omega)] \tag{5}$$

An analogous expression can be written for the β ($\downarrow$) polarizabilities. The total UCHF C_6 dispersion coefficients for the interaction of atoms A and B can be computed from the spin-partitioned isotropic atomic frequency-dependent polarizabilities as,

$$C_6^{AB} = \frac{3}{\pi}\int_0^\infty d\omega \left(\alpha_{iso}^{A\uparrow}(i\omega) + \alpha_{iso}^{A\downarrow}(i\omega)\right)\left(\alpha_{iso}^{B\uparrow}(i\omega) + \alpha_{iso}^{B\downarrow}(i\omega)\right) \tag{6}$$

By multiplying out integrand and regrouping terms, one can partition the total UCHF C_6 coefficient into separate same-spin and opposite spin contributions,

$$\begin{aligned}C_6^{AB} &= \frac{3}{\pi}\int_0^\infty d\omega \left(\alpha_{iso}^{A\uparrow}(i\omega)\alpha_{iso}^{B\uparrow}(i\omega) + \alpha_{iso}^{A\downarrow}(i\omega)\alpha_{iso}^{B\downarrow}(i\omega)\right)\\ &+ \frac{3}{\pi}\int_0^\infty d\omega \left(\alpha_{iso}^{A\uparrow}(i\omega)\alpha_{iso}^{B\downarrow}(i\omega) + \alpha_{iso}^{A\downarrow}(i\omega)\alpha_{iso}^{B\uparrow}(i\omega)\right) \end{aligned} \tag{7}$$

$$= C_6^{AB,ss} + C_6^{AB,os} \tag{8}$$

Once the same-spin UCHF, opposite-spin UCHF, and total CKS C_6 coefficients have been obtained for each atom type in each coordination number environment, the dispersion energies are computed according to the D3 scheme as described previously.[50,51] Specifically, the final C_6 coefficients for a given atom in a particular chemical environment are interpolated using a slightly modified version[50] of the original D3 coordination number scheme, the C_8 coefficients are estimated as proscribed by the D3 model, and then the dispersion energy is obtained as,

$$E_{disp} = s_6 \sum_{AB} f_6(R_{AB})\frac{C_{6,AB}}{R_{AB}^6} + s_8 \sum_{AB} f_8(R_{AB})\frac{C_{8,AB}}{R_{AB}^8} \tag{9}$$

where s_6 and s_8 are scaling coefficients and R_{AB} is the distance between atoms A and B. The Tang-Toennies damping function $f_N(R_{AB})$ is given by,

$$f_N(R_{AB}) = 1 - \exp(s_R R_{AB}) \sum_{k=0}^{N} \frac{(s_R R_{AB})^k}{k!} \tag{10}$$

where N is the order of the dispersion term (6 or 8), and s_R is a distance scaling factor calculated from the cutoff radius $R_{0,AB}$ (taken from the D3 dispersion correction[51]) using two empirical parameters a_1 and a_2:

$$s_R = a_1 R_{0,AB} + a_2. \tag{11}$$

Finally, during the development of the original MP2D model, it was found that the Tang-Toennies damping function decays too slowly at covalent-bond distances.[50] To address this, the inter-

atomic distance R_{AB} was modified via a secondary damping at very short distances according to,

$$R_{AB}' = \begin{cases} r_{cut}R_{0,AB} & \text{if } R_{AB} <= R_{0,AB}(r_{cut} - w/2) \\ R_{AB} & \text{if } R_{AB} >= R_{0,AB}(r_{cut} + w/2) \\ r_{cut}R_{0,AB} + g(R_{AB}, r_{cut}', w') & \text{otherwise;} \end{cases} \tag{12}$$

$$g(R_{AB}, r_{cut}', w') = (-2.5x^8 + 10x^7 - 14x^6 + 7x^5) * w'; \tag{13}$$

$$x = \frac{R_{AB} - (r_{cut}' - w'/2)}{w'} \tag{14}$$

where $r_{cut}' = R_{0,AB}r_{cut}$ and $w' = R_{0,AB}w$. The empirical parameters r_{cut} and w define the distance and width over which the damping occurs. This short-range damping leaves R_{AB} in Eq 9 unchanged at longer distances, while fixing it at a constant fraction of $R_{0,AB}$ value for very short distances. The polynomial $g(R_{AB}, r_{cut}', w')$ smoothly interpolates between the two regimes. See ref 50 for more details.

For restricted wave functions, the spin up and spin down frequency-dependent polarizabilities in Eq 4 are identical and each equal to half the total polarizability. As a result, the same-spin and opposite-spin UCHF C_6 coefficients in Eqs 7 and 8 are each equal to one half to the total C_6 coefficient, and the spin components each contribute half of the UCHF dispersion energy E_{disp} (Eq 9),

$$E_{UCHF}^{disp,ss} = E_{UCHF}^{disp,os} = \frac{1}{2}E_{UCHF}^{disp,tot} \tag{15}$$

In this scenario, the final SCS-MP2D energy (Eq 3) can be expressed as,

$$\begin{aligned} E_{SCS-MP2D} =& E_{HF} + c_{os}E_{MP2-corr}^{os} + c_{ss}E_{MP2-corr}^{ss} \\ & - \frac{1}{2}(c_{os} + c_{ss})E_{UCHF}^{disp,tot} + E_{CKS}^{disp,tot} \end{aligned} \tag{16}$$

Because the present study focuses only on closed-shell species with restricted wave functions for which the open-shell reference hydrides coordination environments contribute negligibly, Eq 16 represents the final equation implemented here.

The SCS-MP2D model contains seven empirical parameters in total: c_{os}, c_{ss}, s_8, a_1, a_2, r_{cut}, and w. The s_6 parameter scaling the C_6 dispersion energy contribution in Eq 9 is set to unity unless otherwise noted. The atom-pairwise dispersion coefficients C_6^{AB} coefficients from the reference hydrides are identical to those used in ref 50. More specifically, the CKS coefficients are taken directly from the original D3 model, while the UCHF ones were computed for H, B, C, N, O, F, Ne, P, S, Cl, Ar, and Br atoms according to the same Grimme D3 scheme. Accordingly, SCS-MP2D is presently applicable to typical organic and biological systems.

Finally, one could conceivably design a similar SCS-MP2D-like approach using the more recent D4 dispersion correction[53] or even the exchange-hole dipole moment (XDM)[54] or many-body dispersion (MBD) models.[55,56] Computing both the UCHF and CKS dispersion contributions in the context of XDM or MBD would require both Hartree-Fock and DFT calculations, making it considerably more computationally demanding. Adapting D4

would be more straightforward computationally and might improve the performance of SCS-MP2D modestly, as evidenced by its performance with DFT functionals.[5,10,53] However, we opt to use D3 here for consistency with our earlier MP2D model, thereby enabling more direct assessment of the impact of spin-component scaling.

2.2 Empirical Parameter Fitting Procedure

The seven empirical parameters in MP2D were fitted to a total of 559 benchmark data points taken from the S66x8 set of dimer intermolecular interactions,[57] the Diels-Alder reaction energy (DARC) subset of the GMTKN55 data set,[3,58] and the sugar conformational energy (SCONF) subset of GMTKN55.[3] Because each of the three data sets differ in the number of data points contained and the magnitudes of the benchmark energy values, an objective function F was constructed from the weighted sum of the relative root-mean-square error (RMSE) for each data set. Relative RMSE values were obtained by dividing the RMSE of each set by the mean absolute value of the benchmark energies in the set, $\langle |E| \rangle$,

$$F = w_{S66x8} \frac{RMSE_{S66x8}}{\langle |E| \rangle_{S66x8}} + w_{DARC} \frac{RMSE_{DARC}}{\langle |E| \rangle_{DARC}} + w_{SCONF} \frac{RMSE_{SCONF}}{\langle |E| \rangle_{SCONF}}$$
(17)

Using relative RMSE values compensates for the fact that the Diels-Alder reaction energies are many-fold larger than the typical intermolecular or conformational energies. Different weights w_i for the three relative RMSEs in the objective function were tested during the fitting. By trial and error, it was determined that increasing the weight of the DARC data set in the objective function led to particularly good, transferable parameters. The final SCS-MP2D parameters were obtained with weight $w_{DARC} = 2.0$ and $w_{S66x8} = w_{SCONF} = 1.0$. The enhanced weight on the DARC set is consistent with the desire to use spin-component scaling to improve reaction energies and the importance of including short-range interactions in fitting the damping function parameters.

Initial exploratory optimizations of the empirical parameters revealed a rugged landscape containing many local minima. Therefore, an evolutionary algorithm was implemented to seek out (nearly) globally optimal parameters that minimize the objective function. Initial values of the parameters were generated randomly within a predefined range of plausibly physical values. Initial spin-component scaling coefficients and most other parameters were restricted to lie between 0 and 2, though a_2 was given the range -1 to 2. Each generation of the search algorithm was populated with 30–40 distinct parameter sets. Each parameter set was optimized using a quasi-Newton algorithm in Cuby4.[59] This gradient-based optimization converged slowly, but it was observed that the parameters varied little after the first ten optimization cycles; therefore, 10 cycles were used for each optimization during the evolutionary search.

After randomly seeding the initial generation, subsequent generations were created as a mixture of, for example: the five best-performing parameter sets from the previous generation, five parameter sets obtained by randomly combining parameters from the top five performers (inheritance), five sets where new damp-

ing function parameters were generated for the top 5 performers (mutation), 5 populations where the spin-component constants were replaced by new randomly generated constants (mutation), and 10 entirely new randomly generated parameter sets to add diversity to the population. Optimization runs which varied the partitioning among inheritance, mutation, and random generation and the total population size were explored. The genetic algorithm was allowed to run for between 5 and 20 generations. To avoid biasing the search in favor of the top performing parameters against newly generated parameter sets, each generation passed initial starting parameters to the next generation rather than the optimized parameters. Once the optimal parameters were found after many searches, they were fully-optimized to ensure a minimum in the parameter landscape had been reached (though their values changed only slightly). Convergence of the evolutionary optimization algorithm was tested both by (1) performing dozens of independent runs of the evolutionary optimizer, and (2) by seeding a Bayesian search algorithm with good parameter sets from the genetic algorithm (see Supporting Information Section S1). The final parameter set discussed in Section 4.1 was discovered relatively early and repeatedly in the search process, and subsequent searching did not reveal any better-performing parameter sets.

3 Computational Methods

The MP2D and SCS-MP2D dispersion corrections were calculated using a developmental version of the freely available Cuby4 software.[59] Energies were computed at the complete basis set limit by combining HF/aug-cc-pVQZ with correlation energies extrapolated[60] from aug-cc-pVTZ and aug-cc-pVQZ results. Counterpoise corrections were employed for all benchmark sets involving purely intermolecular interactions: S66x8, 3B-69, SSI, HBC6, NBC10, Charge Transfer, HB375, and IHB100. MP2 data for S66x8, 3B-69, SSI, and IDISP was taken from the original sources; data for all other sets was computed here using PSI4 version 1.3.[61] For ISOL24, the 24th isomerization reaction was omitted because MP2D and SCS-MP2D dispersion coefficients have not been computed for silicon. For consistency, reaction 24 was excluded for all other tested methods as well.

DFT calculations were performed using DSD-BLYP-D3(BJ), ωB97X-V, and ωB97M-V functionals in PSI4 and the revDSD-PBEP86-D3(BJ) functional in Orca version 4.2.[62] The PSI4 DSD-BLYP-D3(BJ) calculations employ the parameters reported in ref 9, rather than the earlier parameters[63,64] which were used in the GMTKN55 benchmarks.[3] Counterpoise corrections were employed for all intermolecular benchmark sets, except for the revDSD-PBEP86-D3(BJ) functional, for which the counterpoise-corrected results were substantially worse than the uncorrected ones. In general, DSD-BLYP-D3(BJ), revDSD-PBEP86-D3(BJ), and ωB97X-V calculations were performed using the def2-QZVP basis set, while the ωB97M-V ones were performed in aug-cc-pVQZ since the latter basis set is one of the recommended ones from Ref 7. There are a few exceptions, however: For the SSI data set, aug-cc-pVTZ results were taken from Burns et al[65] for ωB97X-V and ωB97M-V. The aug-cc-pVQZ basis set was used with DSD-BLYP-D3(BJ) and revDSD-PBEP86-D3(BJ) for the SSI

Table 1 Comparison of the MP2D parameters from ref 50 and the SCS-MP2D ones determined here. All parameters are dimensionless except for a_2.

	MP2D	SCS-MP2D
c_{os}	1	0.8263
c_{ss}	1	0.9004
a_1	0.9436	1.5359
a_2 (Å)	0.4802	-0.7595
s_8	1.1873	1.2092
r_{cut}	0.72	0.8254
w	0.20	0.1198

set due to the presence of anionic species. The aug-cc-pVQZ basis set was similarly employed for all functionals on the IHB100 set of ionic species. Finally, ωB97X-V results for SCONF, ACONF, Amino20x4, MCONF, PCONF21, DARC, ISOL24, ISO34, and IDISP were taken directly from the GMTKN55 database.[3]

The genetic optimization algorithm was implemented by the authors. The searches employing Bayesian optimization with Gaussian processes utilized the python scikit-learn library `skopt.gp_minimize`.[66]

4   Results and Discussion

4.1   Parameter optimization

The seven empirical parameters were optimized using the evolutionary algorithm discussed in Section 2.2, and the final SCS-MP2D model parameters are listed in Table 1. Figure 1 shows the progress over a single run of the evolutionary optimizer. As the algorithm proceeds through the generations, the population is enriched with low-error parameter sets. By the ninth generation, quite a few parameter sets have been found that perform well, and these best-performing models change little over the next five generations. As shown in the radar plot in Figure 1, the five best-performing members of the population in this optimization run in blue exhibit parameters that are quite similar to those in the final SCS-MP2D model in red. The only appreciable variations occur for the a_2 Tang-Toennies damping parameter. The next five best-performing parameter sets in green differ a little more from the SCS-MP2D ones, most notably in s_8 and a_2, but they are again fairly similar. Moving beyond the ten best parameter sets from this search, one finds greater parameter diversity, indicating that algorithm is searching widely. Overall, many independent runs of the optimizer generated parameter sets that are similar to the final SCS-MP2D ones, and they were always among the very best performing models.

Some searches revealed a few alternative parameter sets that also performed very well, though they had unphysical parameters such as a negative s_8 value. In those models, the MP2D dispersion correction had the wrong sign, with the CKS term effectively increasing the van der Waals binding energy compared to UCHF for systems such as the π-stacked benzene dimer. This behavior is contrary to the well-known behaviors from intermolecular perturbation theory where UCHF over-binds such systems.[46,48] Other tests that allowed s_6 to deviate from unity (increasing the number of parameters from seven to eight) produced good-performing

parameter sets with $s_6 > 1$ and $s_8 \approx 2$. From intermolecular perturbation theory, s_6 and s_8 should both equal one, though the D3 model typically allows s_8 to deviate from unity to compensate for the neglect of higher-order dispersion terms. Given the rapid decay of those higher-order terms with distance, however, it seems unlikely that those neglected contributions should effectively double the s_8 contribution. When tested for transferability to other data sets not employed in the parameter fitting, the final chosen parameter set in Table 1 performed as well as or better than any of these alternative parameter sets. Accordingly, the final chosen parameter set in Table 1 was selected on the basis of its performance on the training set, physically reasonable parameter values, and its transferability to other benchmark sets (Section 4.2).

Table 1 compares the final SCS-MP2D parameters against those published previously for MP2D, which does not scale the spin components of the correlation energy. The SCS-MP2D dispersion correction parameters are fairly similar to those in MP2D. For example, s_8 differs by only 2% between the two models. Values of s_8 near 1.2 are intermediate relative to the range of $s_8 \sim$ 0.8–1.7 typically found for D3 with various density functionals.[51] The parameters r_{cut} and w that govern damping at very short (covalent) distances differ by $\sim$0.1 between MP2D and SCS-MP2D. However, as discussed in the original MP2D study,[50] a relatively broad range of parameters r_{cut} and w performs well, and the differences in these parameters between the two models has a small impact on the overall performance.

In contrast, the SCS-MP2D Tang-Toennies dispersion damping parameters differ noticeably from the earlier MP2D ones. Compared to MP2D, the new parameters enhance the SCS-MP2D dispersion correction contribution at shorter distances (SI Section S2). This increased contribution from the dispersion correction offsets the diminished contribution of the MP2 correlation energy that results from having spin-component scaling coefficients less than one. Interestingly, the spin-scaling coefficients $c_{ss} = 0.8263$ and $c_{os} = 0.9004$ are similar to each other, in contrast to many previous SCS-type MP2 models. Furthermore, the SCS-MP2D spin-component scaling enhances the same-spin contribution relative to the opposite-spin one, which is contrary to theoretical arguments that MP2 typically overestimates the same-spin correlation more than the opposite-spin contributions.[28] On the other hand, such arguments may no longer apply when such a sizable fraction of the final SCS-MP2D correlation energy originates from the CKS dispersion correction.

The importance of the CKS dispersion can be seen from a few example systems. Along the eight points of the S66x8 π-stacked benzene dimer potential energy curve, for instance, an average 93% of the SCS-MP2D correlation energy arises from the CKS dispersion energy. Only $\sim$7% stems from what remains of the spin-component-scaled MP2 correlation energy after subtracting out the UCHF contribution. Of course, dispersion is expected to be important for the benzene π dimer. However, even in the DARC data set chemical reaction energies for which non-dispersion components of the correlation energy are more important, the CKS dispersion still contributes an average 24 kcal/mol, compared to only 10 kcal/mol from the residual spin-component-scaled MP2 correlation energy that remains after removing the UCHF disper-
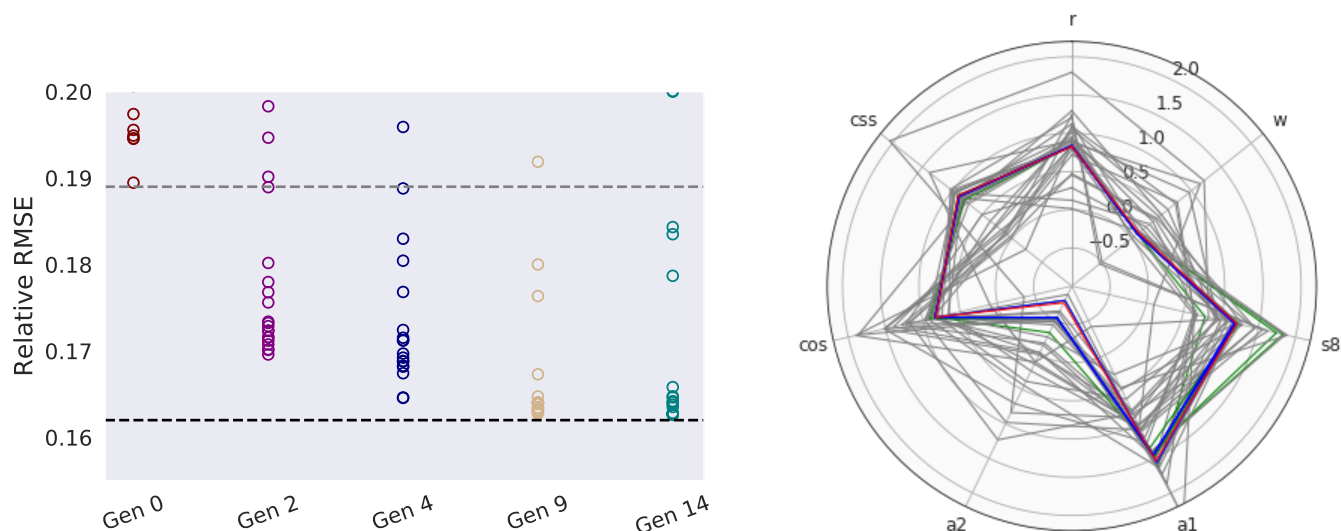
Fig. 1 Left: Performance of one run of the evolutionary SCS-MP2D parameter optimization, plotting the relative root-mean-square error (the objective function) versus the generation number. The figure focuses only on the low-error region, and the horizontal lines indicate the best-performing models in the initial (Gen 0) and final (Gen 14) generations. Right: Radar plot comparing the final SCS-MP2D parameter values (red) to those of the five lowest-error parameter sets (blue), the next five lowest-error models (green), and all other parameter sets (gray) from this particular optimization.

sion component.

That said, a few additional considerations should be noted. First, there are multiple ways to decompose and group the SCS-MP2D energy components. If one partitions it into the SCS-MP2 correlation energy plus a dispersion correction (computed as the difference between the UCHF and CKS dispersion energies), the dispersion correction amounts to only 0.5 kcal/mol for benzene dimer at its equilibrium geometry and an average of 1 kcal/mol (a few percent) for the reaction energies in the DARC set. In other words, the individual UCHF and CKS dispersion terms are very large individually, but the difference between the two contributions is much smaller. Finally, note that the parameters were fitted to the total interaction, conformational, or reaction energies, rather than to individual components of the correlation energy. As discussed for MP2D,[50] this choice leads to good overall performance with fewer empirical parameters, but the individual UCHF and CKS components in SCS-MP2D do not quantitatively reproduce those from intermolecular perturbation theory and/or MP2C.

4.2 Performance on Benchmark Data Sets

As discussed above, SCS-MP2D was trained on three datasets consisting of noncovalent interactions (S66x8), reaction energies (DARC), and sugar conformational energies (SCONF). MP2D was fitted against S66x8.[50] To assess overall performance and transferability, SCS-MP2D was tested on 14 additional benchmark data sets for intermolecular interactions (3B-69 dimers,[67] NBC10,[68] HBC6,[68] HB375,[69] IHB100,[69] SSI,[65] & charge transfer reactions[70]), conformational energies (Amino20x4,[3] ACONF,[71] MCONF,[72] and PCONF21[3]), and thermochemical reaction energies (ISO34,[73] ISOL24,[3] IDISP[3]). Note that IDISP contains a mixture of interaction types that all involve substantial changes in intramolecular dispersion energy, but it is grouped with the

reaction energy data sets here because four of the six examples involve chemical reactions and/or isomerizations. Table 2 summarizes root-mean-square errors for each data set as computed with several different quantum chemistry models. Figure 2 plots the relative RMSEs, where the RMSE for each model is divided by the mean absolute value of the target reference energies for that data set.

4.2.1 Performance of wave function methods

Consider first the performance of MP2 and MP2D for the intermolecular interaction data sets. The MP2D dispersion correction seeks to address systems like the benzene dimer where dispersion is important and which are often considerably over-bound by MP2. The dispersion correction has minimal impact on hydrogen-bonded systems like the water dimer, for which MP2 already performs fairly well. Overall, MP2D reduces the MP2 RMSE on S66x8 four-fold, from 0.67 to 0.16 kcal/mol. That improvement partly reflects that MP2D was trained against S66x8 benchmark data, but the MP2D parameters also prove highly transferable to other benchmark sets. The MP2D dispersion correction provides several-fold error reductions in the non-bonded potential energy curves of NBC10, the protein side-chain side-chain interactions of SSI, the large database of hydrogen bonds in HCNO-containing species (HB375), and a set of charge transfer reactions. Smaller MP2D improvements occur in 3B-69 and the hydrogen-bonded dimer curves of HBC6, and no appreciable improvement is found for the ionic H-bonds of IHB100. The smaller improvements seen in those latter sets largely reflects the lesser importance of dispersion interactions in those dimers rather than any weaknesses in MP2D.

Incorporating spin-component-scaling into MP2D leads to further modest improvements for the intermolecular interaction data sets. In S66x8 (a training set), the RMSE reduces from 0.16 to 0.13 kcal/mol. Improvements are observed for many of the test-

Table 2 Root mean square errors calculated relative to the benchmark reference values (kcal/mol). The asterix (*) indicates data sets that were used to fit SCS-MP2D. Cell color indicates the relative RMSE: dark blue ≤ 5%, light blue = 5–10%, orange = 10–25%, light red = 25–50%, and dark red ≥ 50%.

Data Set	MP2 CBS	MP2D CBS	SCS-MP2D CBS	DSD-BLYP-D3(BJ) def2-QZVP	revDSD-PBEP86-D3(BJ) def2-QZVP	ωB97X-V def2-QZVP	ωB97M-V aQZ
Intermolecular Interactions							
S66x8	0.67	0.16	0.13*	0.18	0.16	0.21	0.11
3B-69 Dimers	0.33	0.21	0.18	0.19	0.19	0.20	0.17
SSI	0.36	0.16	0.17	0.15^a	0.12	0.16^b	0.15^b
HBC6	0.32	0.26	0.25	0.37	0.17	0.32	0.24
NBC10	1.55	0.29	0.14	0.33	0.07	0.34	0.17
Charge Transfer	2.72	0.56	0.34	0.77	0.62	0.57	0.45
HB375	0.43	0.16	0.13	0.14	0.13	0.17	0.19
IHB100	0.45	0.46	0.34	0.41^a	0.26^a	0.37^a	0.35
Conformational Energies							
SCONF	0.31	0.35	0.18*	0.26	0.13	0.21^c	0.24
ACONF	0.11	0.07	0.12	0.08	0.24	0.06^c	0.08
Amino20x4	0.26	0.17	0.18	0.16	0.17	0.24^c	0.24
MCONF	1.02	0.40	0.33	0.55	0.19	0.27^c	0.39
PCONF21	1.11	0.42	0.31	0.48	0.23	0.35^c	0.69
Reaction Energies							
DARC	3.97	1.90	1.41*	1.10	0.64	4.38^c	0.98
ISO34	1.68	1.42	0.96	1.06	0.49	1.56^c	0.82
ISOL24	3.72	2.81	2.24	2.71	1.73	4.20^c	2.41
IDISP	7.03	1.42	1.29	1.60	0.67	3.88^c	2.83
Overall Relative RMSE Statistics (%)							
Mean	21.3	7.7	5.9	8.1	4.9	9.5	8.2
Median	11.5	5.8	4.4	5.5	4.1	5.3	4.5

^{a} aug-cc-pVQZ basis ^{b} Ref 65, aug-cc-pVTZ basis. ^{c} Ref 3

ing sets as well. For example, spin-component-scaling reduces the NBC10 RMSE by a factor of two, from 0.29 kcal/mol for MP2D to 0.14 kcal/mol for SCS-MP2D. More typically, SCS-MP2D reduces the MP2D errors by around a third or less for many of the intermolecular interaction sets. Spin-component scaling tends to improve the MP2D performance on ionic and hydrogen-bonded systems, even though those generally are reasonable even with MP2 and MP2D already. Interestingly, SCS-MP2D also performs better than MP2D for systems like the benzene dimer, which explains much of the improvement observed for NBC10. As discussed previously,[50] the D3 correction has the weakness that its highly local interpolation scheme for the C_6 dispersion coefficients distinguishes poorly between benzene and ethene, for example. So while MP2D performs well for many systems where dispersion is important, its performance for the benzene dimer is actually somewhat worse than the fully *ab initio* treatment in MP2C (though MP2D still improves dramatically upon MP2). SCS-MP2D suffers from the same limitations of the dispersion coefficients, but apparently the spin-component-scaling compensates somewhat, perhaps through an improved description of the sizable non-dispersion components of the benzene-benzene interaction. For instance, the RMSE for the π-stacked benzene dimer decreases from from 0.33 kcal/mol with MP2D to 0.13 kcal/mol with SCS-MP2D. SCS-MP2D performs marginally worse

than MP2D on the SSI data set; this issue will be explored in detail in Section 4.2.3.

One of the key strengths of MP2D and SCS-MP2D over MP2C is that the atom-pairwise definition of the dispersion correction allows for correcting both intra- and intermolecular dispersion. Because it is based on intermolecular perturbation theory, the MP2C dispersion correction has no effect on intramolecular conformational energies. Intramolecular dispersion corrections can be essential in systems such as the conformational polymorphs of organic crystals.[25,26]

Looking at the six conformational energy data sets (1 training and 5 testing), the MP2D and SCS-MP2D performance trends are similar to what was observed for the intermolecular interaction data sets. The MP2D dispersion correction already improves upon MP2 nicely. Adding the dispersion correction reduces the MP2 RMS errors from 1.11 to 0.41 kcal/mol in peptide conformers (PCONF21) and from 1.02 to 0.42 kcal/mol in melatonin conformers (MCONF). Including spin-component-scaling reduces those errors ~20–25% further. Both MP2D and SCS-MP2D perform about one third better than MP2 for the amino acid conformations in Amino20x4. The alkane conformations in (ACONF) are the only data set here where SCS-MP2D (RMSE 0.12 kcal/mol) does not improve upon MP2 (0.11 kcal/mol) and is somewhat worse than MP2D (0.07 kcal/mol). Fortunately,
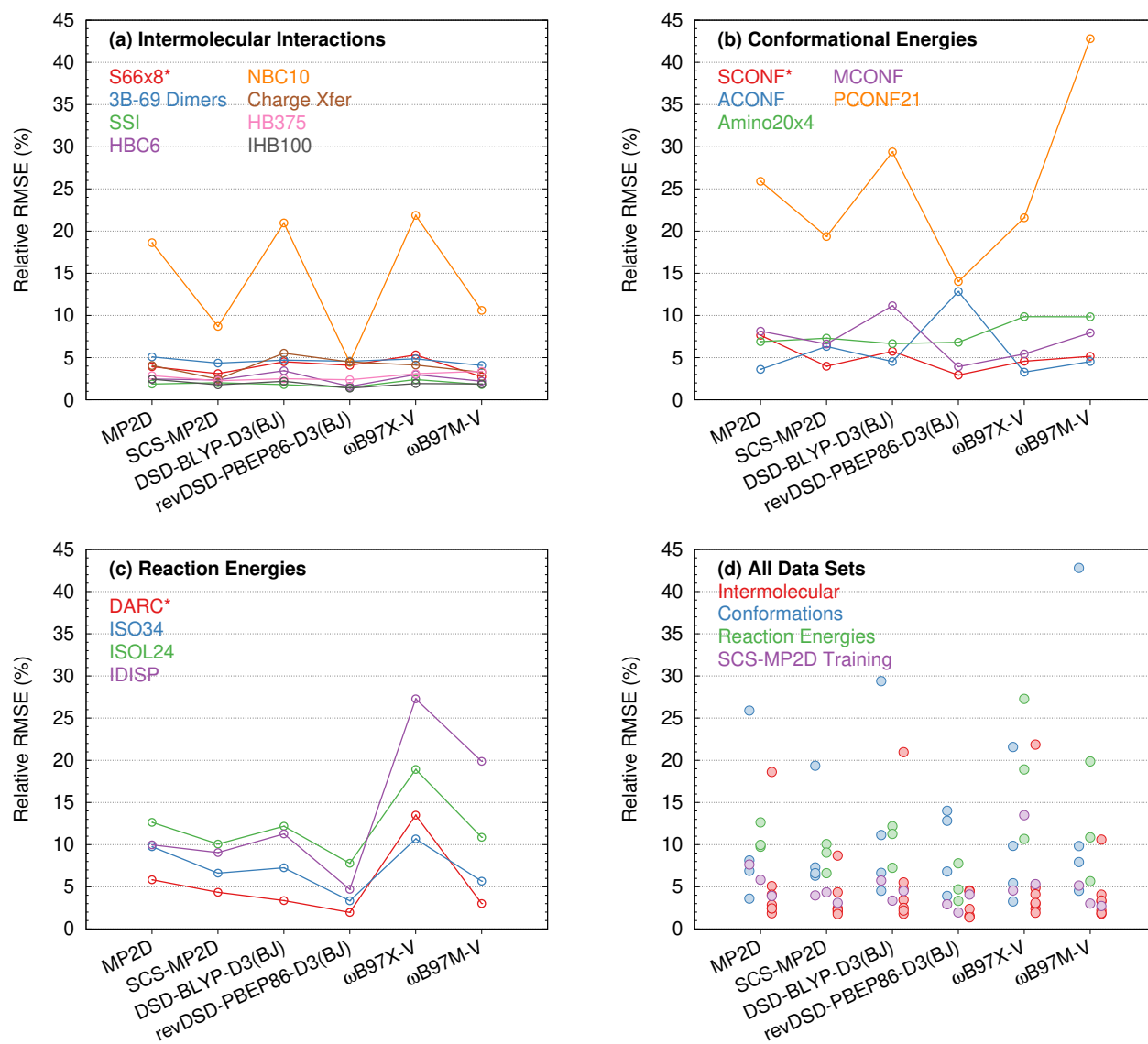
Fig. 2 Comparison of the relative RMSEs for several models on benchmark data sets focusing on (a) intermolecular interactions, (b) conformational energies, (c) reaction energies, and (d) the union of all training and testing sets. The asterix indicates that S66x8, SCONF, and DARC were involved in fitting the SCS-MP2D parameters. Note that the large relative RMSE values for NBC10 and PCONF21 partly reflect the smaller magnitudes of the benchmark energies in those two sets. The absolute errors for the best-performing models in NBC10 and PCONF21 are similar to those in HB375 and MCONF, respectively (see Table 2).

these errors are small in both absolute and relative terms (e.g. 6.3% for SCS-MP2D).

Finally, we examine the reaction energies associated with Diels-Alder reactions (DARC, training set), the smaller- and larger-molecule isomerizations (ISO34, ISOL24), and the IDISP set, which contains several dimerization and isomerization reactions for which intramolecular dispersion matters. Non-dispersion contributions to the correlation energy are expected to be sizable for reaction energies, so spin-component scaling might be expected to have a significant impact in these data sets. Indeed, while MP2D did improve upon MP2 for all four data sets, SCS-MP2D performs even better. For example, MP2D (1.42 kcal/mol) only improved upon MP2 (1.68 kcal/mol) by about 15% for the ISO34 small-molecule isomerizations. The MP2D dispersion correction

has a slightly larger impact on the larger-molecule isomerizations of ISOL24, reducing the MP2 error by $\sim 25\%$ (3.72 to 2.81 kcal/mol). In both cases, however, SCS-MP2D reduces the errors by $\sim40\%$ compared to MP2, with RMSE values of 0.96 and 2.24 kcal/mol, respectively. For DARC, which was included in the SCS-MP2D fitting, SCS-MP2D reduces the MP2 error by 65%, and it reduces the MP2D error by $\sim25\%$. The IDISP set tends to exhibit considerable variability in the error statistics achieved by different models, due to the diverse chemistry and the disparate energy scales for the different reactions. Regardless, both MP2D and SCS-MP2D perform very well for this set, with RMS errors of 1.42 and 1.29 kcal/mol, respectively.

As a whole, these benchmark results demonstrate that the SCS-MP2D model is highly transferable to a wide variety of organic

chemistry, despite being fitted to a modest amount of training data. In most of the benchmarks performed here, the MP2 dispersion contributions account for the largest share of the improvement, but the spin-component scaling almost always improves the quality of the predicted energies further. Given the error statistics presented here, SCS-MP2D is arguably one of the best-performing $O(N^5)$ correlated wave function methods available today for describing intra- and intermolecular interactions in organic chemistry. However, a better understanding of its overall performance requires comparing it to state-of-the-art density functionals.

4.2.2 Comparisons to selected density functional models

Four top-performing density functionals were chosen for comparison against the dispersion-corrected MP2 models: the range-separated hybrid functional ωB97X-V, the meta-GGA variant ωB97M-V, and the double-hybrid spin-component-scaled functionals DSD-BLYP-D3(BJ) and revDSD-PBEP86-D3(BJ). The ωB97X-V and DSD-BLYP-D3(BJ) functionals were selected based on their excellent performance on the GMTKN55 test suite,[3] while ωB97M-V was selected because it represents the meta-GGA rung on Jacob's ladder of density functionals and generally performs even better than ωB97X-V.[4,7]

The recently revised DSD functionals (revDSD) perform even better than the original DSD functionals.[10] The improvements stem primarily from replacing the D3 dispersion correction with the newer D4 one[53] in some of the functionals and from fitting the empirical parameters to a much larger set of training data. The revDSD-PBEP86-D3(BJ) functional was selected as a representative example of these new functionals. Because SCS-MP2D could plausibly be developed based on the D4 correction instead of the D3 one as well, we opted to compare against the D3 version of revDSD-PBEP86 for the sake of consistency. On the GMTKN55 data set, the revDSD-PBEP86-D3(BJ) functional performs about 0.1 kcal/mol worse than the D4 version in the weighted mean absolute deviation.[10]

The double-hybrid ωB97M(2) functional[8] would be another interesting potential comparison, since it performs noticeably better than the ωB97X-V and ωB97M-V functionals tested here and even slightly better than the revDSD functionals (e.g. $\sim$0.1 kcal/mol better than revDSD-PBEP86-D4 for GMTKN55).[4] However ωB97M(2) is not presently implemented in any of the software packages used here. Overall, the four functionals selected here are representative of top-performing functionals in their respective categories.

To facilitate comparisons between the MP2-based methods and the DFT functionals, Table 2 employs color-coding based on the relative RMS errors (i.e. RMSE divided by the average magnitude of the benchmark energy in each set). Dark blue corresponds to relative RMSEs of 5% or less, light blue to relative RMSEs in the range 5–10%, orange to those in the range 10–25%, light red for the range 25–50%, and dark red for larger relative RMSEs. Figure 2 plots the relative RMSEs for all models except MP2. MP2 is omitted from Figure 2 because its large errors would obscure the comparison among the better-performing methods.

Considering first the intermolecular interactions, Figure 2a highlights how most of these dispersion-corrected MP2 and DFT

models perform well, but revDSD-PBEP86-D3(BJ), SCS-MP2D and ωB97M-V are the clearly the top performers. The most noticeable differences among the models occur for NBC10 and the charge transfer set. In NBC10, revDSD-PBEP86-D3(BJ) (RMSE 0.07 kcal/mol), SCS-MP2D (RMSE 0.14 kcal/mol) and ωB97M-V (0.16 kcal/mol) exhibit errors that are a factor of 2–3 times smaller than those for MP2D and the other two functionals. For the charge transfer set, all four functionals perform noticeably worse than SCS-MP2D, especially DSD-BLYP-D3(BJ). Delocalization error in approximate functionals hinders the description of such systems. The inclusion of exact exchange and/or range-separation in these functionals reduces delocalization error considerably, but the RMSEs for this set remain appreciably larger than for any of the other intermolecular interaction data sets.

The performance of SCS-MP2D is also competitive with these density functionals for the conformational energy data sets. No single model performs uniformly well across all five test sets, but SCS-MP2D and revDSD-PBEP86-D3(BJ) exhibit the most consistent performance. The revDSD-PBEP86-D3(BJ) functional performs noticeably better for SCONF, MCONF, and PCONF21, about the same for Amino20x4, and appreciably worse for ACONF. However, SCS-MP2D either performs better than or is on par with the other three functionals. ACONF provides the most notable exception, with the other three functionals performing quite a bit better than SCS-MP2D and revDSD-PBEP86-D3(BJ). The peptide conformations in PCONF21 are interesting for two reasons. First, due to the small average conformational energies, the relative RMSE values for most models are more than double those of the other sets, even if the absolute RMSE values are similar to those in MCONF. Second, while the ωB97M-V functional generally performs very well for conformational energies, its 0.69 kcal/mol RMSE for PCONF21 is roughly double that of SCS-MP2D (0.32 kcal/mol) and noticeably worse than the other three functionals (0.23–0.48 kcal/mol).

The behavior of the various models for the reaction energy data sets follows similar patterns. SCS-MP2D exhibits root-mean-square errors that are typically $\sim$20–70% smaller than those for DSD-BLYP-D3(BJ) and ωB97X-V for DARC, ISO34, ISOL24, and IDISP. The only exception is that SCS-MP2D and DSD-BLYP-D3(BJ) perform about the same on IDISP (1.29 and 1.31 kcal/mol RMSE, respectively). The comparison between SCS-MP2D and ωB97M-V is more mixed, with SCS-MP2D performing better on ISOL24 and IDISP, and ωB97M-V giving smaller errors for DARC and ISO34. As shown in Figure 2c, SCS-MP2D does notably provide much more consistent relative errors than ωB97M-V, however (Figure 2). Finally, revDSD-PBEP86-D3(BJ) is the clear winner for these reaction energies, with errors that are often only half those of SCS-MP2D.

Figure 2d aggregates the relative errors for all methods except MP2 on all the data sets. It highlights how SCS-MP2D, MP2D, and all of the density functionals examined here generally perform well. The best-performing model is revDSD-PBEP86-D3(BJ), but SCS-MP2D is only moderately worse. The general consistency of both models across the different data sets is particularly notable. This can also be seen from the mean and median statistics of the relative RMSEs for all data sets in Table 2: revDSD-

PBEP86-D3(BJ) exhibits the smallest mean and median errors of 4.9% and 4.1%, respectively. SCS-MP2D performs a little worse at 5.9% (mean) and 4.4% (median), and ωB87M-V is close behind with a median error of 4.5%, though its mean error of 8.2% is much larger due to its poor performance on PCONF21. The statistics also highlight that MP2D is fairly competitive with DSD-BLYP-D3(BJ) and ωB87X-V, but those three models represent a noticeable decrease in accuracy compared to the best three.

In other words, SCS-MP2D is highly competitive with some of the best density functionals on these benchmark sets. Based on earlier benchmarks, one anticipates that revDSD-PBEP86-D4 and ωB87M(2) would perform even a little better than any of the models here. On the other hand, there are some very encouraging features of SCS-MP2D here. First, while SCS-MP2D has 7 empirical parameters and the DSD functionals have six, ωB97X-V has 10 and ωB97M-V has 12. The fact that SCS-MP2D exhibits good performance and transferability with a modest number of empirical parameters speaks well to the physical partitioning of the dispersion and residual correlation energies in the model. Second, substantial error reduction was obtained with the revDSD functionals compared to the original DSD versions by optimizing the empirical paramaters against a much larger data set.[10] This raises the prospect that a similar strategy might lead to further improvements for SCS-MP2D as well.

4.2.3 Anion-Anion interactions in the SSI data set

Despite overall good performance on the SSI data set[65] (RMSE of 0.17 kcal/mol), SCS-MP2D actually performs slightly worse than the original MP2D method (0.16 kcal/mol). The subset breakdown in Figure 3 makes clear that the interactions involving anions, especially the anion-anion subset, are the primary driver of this larger RMSE. In fact, MP2 actually performs better than SCS-MP2, MP2D, and MP2C on this anion-anion subset, which indicates there may be a general error in the dispersion correction scheme used in these methods for ionic species (Table 3). This behavior contrasts the results of the IHB100 data set of 100 hydrogen-bonded ion pairs, where SCS-MP2D performs somewhat better than either MP2 or MP2D (Table 2).

The SSI reference data uses Sherrill's silver standard DW-CCSD(T**)-F12 approach. For comparison, we also computed the energies using conventional CCSD(T)/CBS as computed from MP2/aug-cc-pV[TQ]Z and CCSD(T)/aug-cc-pVDZ. Changing the reference data reduces the SCS-MP2D RMSE modestly from 0.77 to 0.65 kcal/mol, but it does not alter the fundamental story that SCS-MP2D is performing worse than MP2D for the anion-anion interactions.

For comparison, Table 3 also compares the performance of several density functionals on the anion-anion subset. With an RMSE of 0.73 kcal/mol, ωB97X-V performs comparably to SCS-MP2D. revDSD-PBEP86-D3(BJ) and DSD-BLYP-D3(BJ) perform only moderately better at 0.62 and 0.50 kcal/mol, respectively. In contrast, ωB97M-V performs outstandingly with an RMSE of only 0.29 kcal/mol. These anion-anion interactions represent only a small fraction of the full set, however, and the different models exhibit much smaller variations in RMSE across the full SSI data set.

One possible source of the SCS-MP2D behavior (and the two DSD functionals, to a lesser extent) could lie in the D3 dispersion correction. The D3 dispersion coefficients are interpolated from neutral hydrides based on the geometry-dependent coordination number. They do not, however, directly differentiate between neutral and ionic environments. Anions tend to be more polarizable and likely exhibit stronger dispersion interactions that are perhaps not handled ideally in the MP2D and SCS-MP2D models. The newer D4 dispersion correction[53] accounts for the effect of atomic charge in the coordination numbers used to calculate the dispersion coefficients. Indeed, the performance of the DSD-family of double-hybrid functionals improves noticeably over a variety of benchmark sets when the D4 correction is used in place of D3.[10] Perhaps a version of the MP2D-type methods based on D4 would perform better for these anion-anion interactions. On the other hand, it is also worth noting that MP2C, which computes the UCHF and CKS dispersion contributions from first principles, still performs worse than canonical MP2 and ωB97M-V. In other words, the worse performance of the dispersion-corrected MP2 models is probably not entirely due to the D3 dispersion correction. Exchange-dispersion and induction-dispersion couplings are also important in ionic interactions,[74] and it is possible that the CKS dispersion correction and/or spin-component scaling in the MP2D-type methods disrupt some favorable error cancellation between the UCHF dispersion and those other terms.

4.3 Two Challenging Examples

The data sets examined above provide a broad perspective for the performance of SCS-MP2D relative to other models, but it can also be instructive to look at specific, challenging systems. Here we focus on two: the torsional scan about the key dihedral angle in the ROY molecule[24,25,75] and the dissociation of the anthracene photodimer.[50,76,77] Both have proved challenging for common GGA and hybrid density functionals due to delocalization error stemming from changes in the extent of π conjugation in these systems. Comparing how several top-performing models behave on these potential energy curves provides further insights into their capabilities.

4.3.1 ROY dihedral angle scan

The ROY molecule (Figure 4a) holds the current record for the largest number of fully characterized crystal polymorphs.[78–83] These polymorphs exhibit vibrant red, orange, or yellow crystals, depending on the degree of conjugation between the two aromatic rings as governed by the dihedral angle θ_{thio} (Figure 4). The 12 characterized polymorphs lie within a narrow $\sim$1–2 kcal/mol energy window, and even modest failures to predict these conformational energies correctly inhibits accurate energy ranking of the different crystal polymorphs.[24–26,75,79] GGA density functionals such as B86bPBE-XDM typically predict the thermodynamically preferred Y polymorph to be one of the least stable crystal forms (Figure 5), while red/orange forms are artificially stabilized. Compared to CCSD(T) benchmarks, MP2D predicts the conformational energies much more reliably than conventional GGA and hybrid functionals, and it is one of the relatively few methods that has been shown to predict the polymorph stabilities
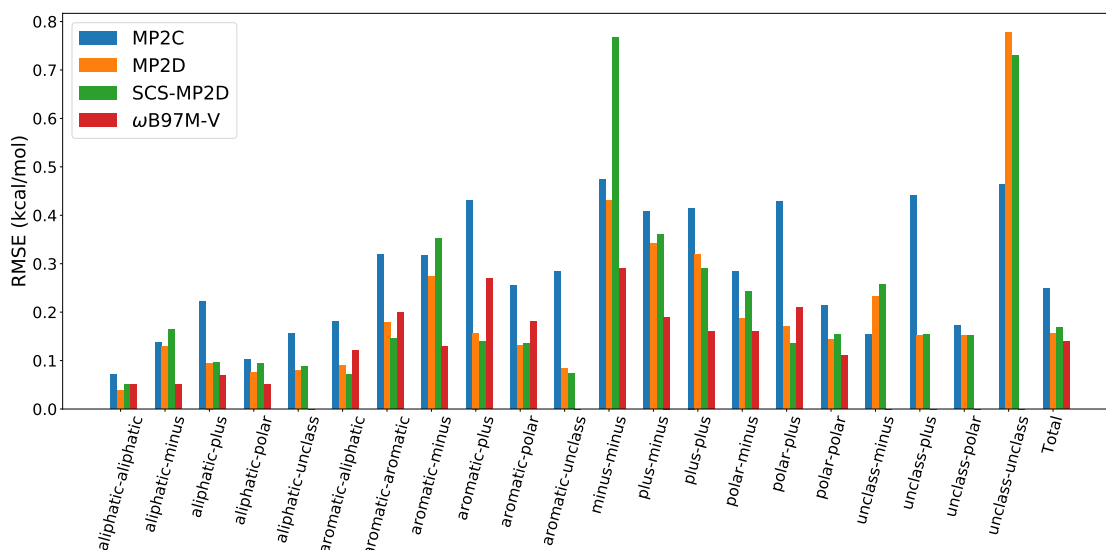
Fig. 3 Root mean square error of MP2C, MP2D, SCS-MP2D, and ωB97M-V for the different interaction categories in SSI relative to the DW-CCSD(T)-F12 reference values of ref 65.

Table 3 Root-mean-square errors for various models in the anion-anion interaction subset of SSI compared to those for the full SSI data set, using the benchmark DW-CCSD(T**)-F12 reference values of ref 65.

Method	Anion-Anion Subset	Full SSI Set
MP2/CBS (Ref 50)	0.36	0.36
MP2C/CBS (Ref 50)	0.43	0.12
MP2D/CBS (Ref 50)	0.43	0.16
SCS-MP2D/CBS (this work)	0.77	0.17
DSD-BLYP-D3(BJ)/aug-cc-pVQZ (this work)	0.50	0.15
revDSD-PBEP86-D3(BJ)/aug-cc-pVQZ (this work)	0.62	0.12
ωB97X-V/aug-cc-pVTZ (Ref 65)	0.73	0.16
ωB97M-V/aug-cc-pVTZ (Ref 65)	0.29	0.15

largely correctly.[25,26]

Figure 4 plots the one-dimensional conformational energy scan for θ_{thio}. The geometries were taken from ref 25, where they were obtained by constraining the dihedral angle at different angles 0–150° (in 10° steps) and relaxing all other degrees of freedom using B3LYP-D3(BJ)/def2-TZVP. Single-point energies were then computed on these geometries with the various methods considered in Figure 4. CCSD(T) benchmarks[25] predict a global minimum around 120°. A secondary, more shallow minimum occurs around 50°, and it is separated from the global minimum by a small barrier near 70°. The biggest challenge along this potential energy coordinate occurs in the ~0–80° region and above 120°—i.e. for the more planar structures.

Typical GGA functionals like B86bPBE-XDM dramatically over-stabilize the more planar structures (Figure 4a), and this leads to over-stabilization of the polymorphs with red and orange colors (R, OP, ON, & ORP; $\theta_{thio} \sim 20$–60°) relative to the yellow ones (Y, YN, & YT04; $\theta_{thio} \sim 100$–120°).[24,25,75] This behavior stems from delocalization error in the functionals artificially stabilizing the more planar conformations that allow greater conjugation between the two rings.[26] B86bPBE-XDM also shifts the global minimum of the scan closer to 130° than the true 120° value, and it incorrectly predicts both the position and magnitude of the bar-

rier between the two wells. The hybrid B3LYP-D3(BJ) partially corrects the energies and positions of the two minima and the barrier, but not enough.

MP2D performs considerably better than these traditional GGA and hybrid functionals. It predicts the correct position of the global minimum, and overestimates the barrier height and secondary minimum stability by < 0.25 kcal/mol. While the position of the secondary minimum is ~10° degrees too high with MP2D, it does position the barrier maximum correctly. SCS-MP2D improves upon MP2D modestly throughout the low-angle range, such that the minima and barrier maximum are all positioned correctly. The overall RMSE relative to CCSD(T) is about a third smaller than that of MP2D (0.12 kcal/mol vs 0.17 kcal/mol).

Consider next the double hybrid DSD-BLYP-D3(BJ) functional (Figure 4b). It performs fairly well near 120° and below 30°, but it overestimates the barrier near 80° and shifts the angle at which the minima occur by about 10° in opposite directions. This leads to a distorted potential energy curve. In contrast, the range-separated hybrid ωB97X-V and hybrid meta-GGA ωB97M-V functionals under-stabilize the low-angle conformations, such that no minimum occurs near 50° at all. So while the magnitude of the errors at any individual point along the energy surface is not especially large for those three functionals, the shapes of the po-
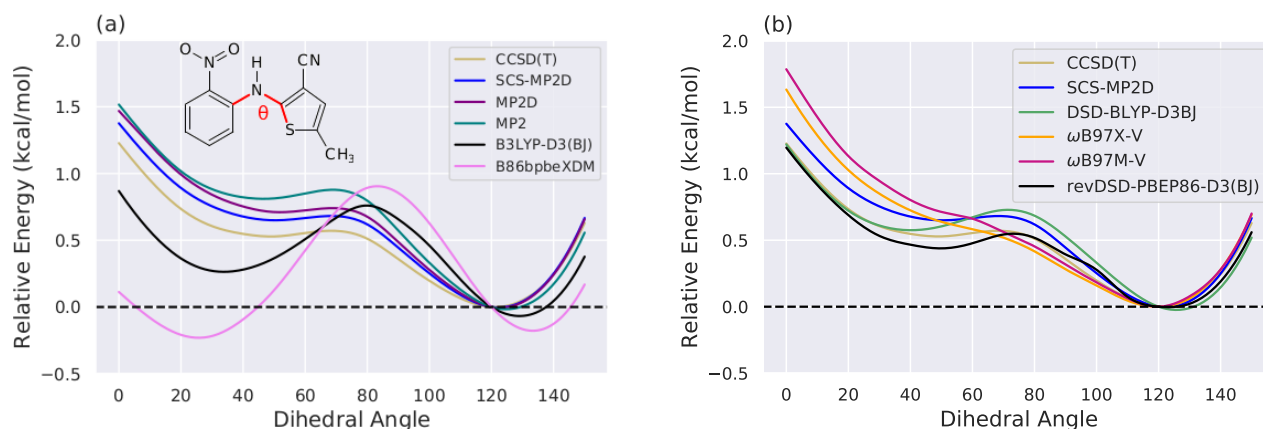
Fig. 4 Relaxed potential energy scan curves for the θ_{thio} dihedral angle of ROY, as computed with several different density functionals and correlated wave function models. The potential energy curves are separated into two panels for easier viewing. Yellow polymorphs exhibit dihedral angles of approximately $\sim 90 \pm 40°$, while red and orange forms have more planar dihedral angles.

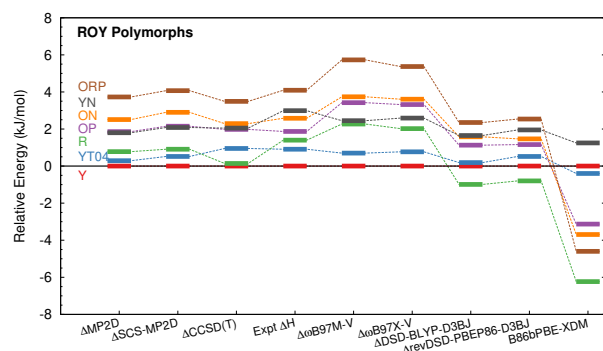**Fig. 5** Comparison of relative lattice energies for seven ROY polymorphs against experimentally measured enthalpies. The methods listed with a Δ were used to compute the intramolecular conformational energy correction; the intermolecular part was calculated with periodic B86bPBE-XDM throughout.

tential energy curves are qualitatively incorrect, particularly for the ωB97-based functionals. The revDSD-PBEP86-D3(BJ) functional performs the best among the DFT models for this curve, though whether it or SCS-MP2D performs better is debatable and depends on which regions of the curve one focuses on.

Figure 5 examines the impact of the conformational energy differences on the crystal polymorph stabilities by comparing the relative lattice energies for the seven polymorphs with experimentally reported enthalpies.[84–87] The lattice energies were computed via the monomer-correction approach,[26] which models the crystal energy as a combination of periodic DFT for the intermolecular interactions and a higher-level of theory for the intramolecular conformational energy. Here, the intermolecular part is computed with the B86bPBE-XDM functional (results and fixed-cell optimized geometries taken from ref 25), while the intramolecular conformational energy correction is computed with SCS-MP2D and the other methods listed in Figure 5.

Figure 5 highlights how the GGA B86bPBE-XDM overstabilizes the red and orange polymorphs (R, ON, OP, ORP) relative to the yellow ones (YN, YT04, Y). Correcting the conformational

energies with MP2D or SCS-MP2D gives results in much better agreement with experiment, and only the position of the YN polymorph differs appreciably from experiment. In contrast, correcting the conformational energies with DSD-BLYP-D3(BJ) or revDSD-PBEP86-D3(BJ) only partially resolves the B86bPBE-XDM problems, with the R polymorph still predicted to be more stable than form Y. ωB97M-V performs somewhat better, though it seemingly over-estimates the destabilization of R, OP, ON, and ORP, which is consistent with the errors seen in the low-angle region of conformational energy scan (Figure 5). Overall, Figure 5 highlights how the conformational energy errors seen in Figure 4 impact polymorph energy differences.

The comparison against experiment assumes that B86bPBE-XDM is adequate for describing the intermolecular interactions in the crystal, and it neglects phonon contributions to the enthalpies.[25] Performing the conformational energy correction with CCSD(T) leads to polymorph stabilities that are seemingly slightly worse than the MP2D or SCS-MP2D ones compared to experiment, which is presumably an artifact of those other approximations and differing degrees of fortuitous error cancellation. In any case, the SCS-MP2D-corrected polymorph energies exhibit better agreement with the experimental and CCSD(T)-corrected energies than do any of the density functionals.

4.3.2 Anthracene photodimerization

The anthracene photodimerization potential energy surface is another system that challenges wave function and DFT methods alike. This photochemical reaction converts two non-covalent π-stacked anthracene molecules to form the covalently linked "butterfly" photodimer product (Figure 6). The difficulty of modeling the energy difference between reactants and products with electronic structure methods was first highlighted by Grimme.[76] The largest difficulty lies in the photodimer product, which exhibits atypically long C-C single bonds between the rings, highly distorted "anthracene" moieties with diminished π conjugation, and strong close-range van der Waals dispersion interactions between rings. This reaction is also of practical importance, as there has been considerable interest in anthracene-based photomechanical
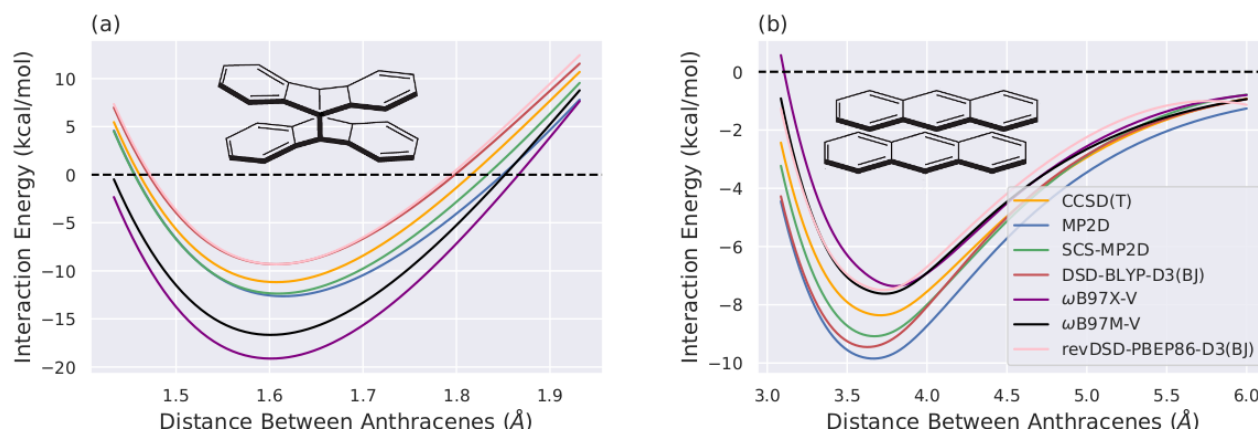
Fig. 6 Potential curves along the anthracene photodimerization curve for (a) the photodimer and (b) the π-stacked non-covalent dimer.

materials,[88–91] Problematic energy predictions for this reaction directly impact the ability for modeling to help understand those systems[77] and design anthracene derivatives with improved photomechanical switching properties.

To study this system, a one-dimensional reaction coordinate as a function of the separation between the two anthracenes has previously been constructed.[50] At each constrained separation of the four carbons involved in forming the new bonds, all other degrees of freedom were relaxed. Two caveats regarding this energy curve here should be noted: (1) The constraints applied to generate this curve enforce a symmetrically stacked sandwich π dimer, while the true system would offset the anthracene molecules laterally at longer interdimer separations. (2) No effort was made to model the excited-state chemistry or the multi-reference character associated with the forming/breaking of two covalent C-C bonds simultaneously in the intermediate region between $\sim$2–3 Å. The spin-restricted wave functions used here will perform poorly in this region; our focus lies on the reactant and product basins. A full potential energy curve is provided in SI Section S4 for completeness.

Figure 6 shows how several different methods perform in the non-covalent π stacked region (near 3.6–3.8 Å) and the covalent photodimer region (near 1.6 Å). Though not shown here, traditional GGA and hybrids like B86bPBE-XDM and B3LYP-D3(BJ) perform poorly for this system, exhibiting errors of up to tens of kcal/mol and predicting the photodimer product to be less stable than two non-interacting anthracene molecules.[50,76,77] In contrast, MP2D performs quite well relative to CCSD(T).[50,77] It binds both the π-stacked dimer and the photodimer a few kcal/mol too tightly, but the systematic nature of the error between the two energy wells leads to a photodimerization reaction energy of -2.8 kcal/mol that agrees almost perfectly with the CCSD(T) value of -2.9 kcal/mol (Table 4). It also improves considerably over MP2 (not shown here).[50] SCS-MP2D performs a little better than MP2D—it slightly reduces the errors relative to CCSD(T) in the two minima, and it actually mirrors CCSD(T) almost perfectly in the bond-breaking region near 2.5 Å, unlike MP2D. The SCS-MP2D reaction energy of -3.3 kcal/mol is also in excellent agreement with CCSD(T).

The density functionals examined here perform better than traditional GGAs and hybrids for this reaction, but flaws remain. In particular, ωB97X-V and ωB97M-V both overbind the photodimer and underbind the π dimer. This means that the errors compound when computing the reaction energy, and the resulting ΔE is much too exothermic (-9.2 and -12.3 kcal/mol, respectively, Table 4). DSD-BLYP-D3(BJ) reverses the binding trends compared to those two functionals, and it incorrectly predicts almost zero energy difference between the two species. revDSD-PBEP86-D3(BJ) gives the best DFT reaction energy of -2.0 kcal/mol, though it still performs a little worse that SCS-MP2D.

Table 4 also lists RMSE values relative to the CCSD(T) benchmarks, as computed across all data points used in each of the two basins. In the non-covalent π dimer basin (3.2–6 Å), MP2D, SCS-MP2D, and all three functionals perform very well, with errors of 0.5 kcal/mol for SCS-MP2D and 0.7–1.2 kcal/mol for the four functionals. In contrast, the errors in the photodimer basin (1.4–2.0 Å) are somewhat larger for ωB97X-V (6.4 kcal/mol) and ωB97M-V (4.6 kcal/mol), compared to only 1.2 kcal/mol for SCS-MP2D and 1.8 kcal/mol for revDSD-PBEP86-D3(BJ). Overall, spin-component scaling reduces the MP2D errors by about a factor of two across these two basins, and SCS-MP2D reproduces this challenging CCSD(T) potential energy curve more faithfully than any of the four density functionals.

Table 4 Reaction energy ΔE_{rxn} for anthracene photodimerization, 2 $C_{14}H_{10} \longrightarrow (C_{14}H_{10})_2$,[a] in kcal/mol. Root-mean-square errors relative to the CCSD(T) benchmarks are also presented for the photodimer (1.4–2.0 Å) and π dimer (3.2–6.0 Å) basins from Figure 6.

Method	ΔE_{rxn}[a]	Root-Mean-Square Error	
		Photodimer	π Dimer
CCSD(T)	-2.9		
MP2	-5.7	10.0	5.1
MP2D	-2.8	2.0	1.3
SCS-MP2D	-3.3	1.2	0.5
DSD-BLYP-D3(BJ)	0.1	1.5	0.9
revDSD-PBEP86-D3(BJ)	-2.0	1.8	0.8
ωB97X-V	-12.3	6.4	1.2
ωB97M-V	-9.2	4.6	0.7

[a] $\Delta E_{rxn} = E(1.6\text{ Å}) - E(3.6\text{ Å})$

5 Timings

The computational cost of SCS-MP2D is functionally equivalent to that of MP2, since the dispersion correction incurs only force-field-like cost. With modern density-fitting approximations, the cost of MP2 is comparable to or faster than several of the other functionals considered here. Table 5 compares the relative timings of several density functionals against SCS-MP2D. In addition to DSD-BLYP-D3(BJ), ωB97X-V, and ωB97M-V, Table 5 also includes timings for the widely-used global hybrid B3LYP-D3(BJ) functional and the meta-GGA B97M-V (for comparison against the range-separated version). For consistency, all calculations were performed in PSI4 and with the aug-cc-pVQZ basis set. Density fitting was used for both the self-consistent field (SCF) and perturbative correction portions of the calculations, and core electrons were frozen for the perturbation corrections. All other parameters, including integral thresholds and exchange correlation integration grids, were left at their default values. Timing data is presented the 33-atom amino acid tyrosine (Amino20x4 geometry TYR_xab) and for the 81-atom lanosterol (ISOL24 geometry i4e). The timings are relative to SCS-MP2D, which took 27 minutes and 9.7 hours of wall time, respectively. As can be seen from the table, SCS-MP2D, B3LYP-D3(BJ), and DSD-BLYP-D3(BJ) exhibit comparable computational cost. The DSD-BLYP-D3(BJ) calculations are slightly more expensive per SCF iteration, but SCS-MP2D requires 2-3 additional SCF cycles to converge in these examples due to the tighter default PSI4 convergence criterion (10^{-8} hartrees for MP2 calculations vs. 10^{-6} hartrees for the DFT functionals). On the other hand, SCS-MP2D is $\sim$2–2.5 times faster than the ωB97X-V and ωB97M-V functionals, despite the stricter SCF convergence criterion in SCS-MP2D. In other words, for a given choice of basis set, SCS-MP2D requires similar or less time than these state-of-the-art functionals. A significant fraction of that cost is associated with the range-separated hybrid nature of the function; the meta-GGA B97M-V functional is considerably faster than its range-separated hybrid analog, ωB97M-V.

Timing comparisons in practical applications will depend on basis set, of course. For example, including an SCS-MP2D/aug-cc-pVTZ calculation to enable extrapolation to the CBS limit would increase the SCS-MP2D cost by 10–20% in these two species, but it would still be roughly twice as fast as the ωB97X-V and ωB97M-V functionals. On the other hand, changing the basis set to def2-QZVP or def2-QZVPPD would reduce the DFT costs appreciably while probably maintaining good accuracy. DFT functionals often perform better in smaller basis sets than do MP2 methods, though large basis sets such as def2-QZVPPD, pc-3, aug-pc3, or aug-cc-pVQZ are recommended for use with the ωB97M-V functional, for example.[7]

6 Conclusions

The last decade has witnessed substantial performance improvements in lower-cost models based on DFT and MP2. This study presented a new spin-component-scaled, dispersion-corrected MP2 model that provides accuracy that is competitive with some of the best density functional models for intermolecular interactions, conformational energies, and thermochemistry in organic

Table 5 Relative PSI4 wall times for several of the density functionals considered here compared to SCS-MP2D, all employing the aug-cc-pVQZ basis set and density fitting. The benchmarks employ 4 (tyrosine) or 6 (lanosterol) cores of an AMD EPYC 7282 processor with 4 gigabytes of DDR4-3200 MHz random access memory per core and Micron 5300 PRO solid-state disk storage.

| | Tyrosine $C_9H_{11}NO_3$ | Lanosterol $C_{30}H_{50}O$ |
# of basis functions:	2096	4780
SCS-MP2D	1.0	1.0
B3LYP-D3(BJ)	1.1	1.2
DSD-BLYP-D3(BJ)	1.1	0.9
B97M-V	1.6	1.0
ωB97X-V	2.1	2.4
ωB97M-V	2.5	2.6

systems. Indeed, SCS-MP2D out-performs the several state-of-the-art density functionals in the particularly challenging potential energy curves examined here. Among the benchmark test sets examined, the largest SCS-MP2D performance improvements over MP2D occur for the non-bonded complexes in NBC10, charge transfer reactions, the ionic hydrogen bonds of IHB100, and many of the conformational and reaction energy data sets. For systems with $\sim$100 atoms, the computational cost of SCS-MP2D is also similar to or lower than that of top-tier range-separated and double-hybrid density functionals.

SCS-MP2D is the first spin-component-scaled model for which a single set of spin scaling coefficients can describe both inter- and intramolecular interactions well. The fact that the seven empirical parameters in SCS-MP2D could be trained using a relatively small amount of training data while maintaining excellent transferability to new systems suggest that the partitioning of the MP2 corrections into an improved treatment of dispersion plus a re-weighting of the residual correlation energy components is theoretically sound. Overall, this research shows that correlated wavefunction methods continue to provide a viable alternative to DFT for even reasonably large molecules.

Author Contributions

Beran planned and supervised the research with input from Řezáč; Greenwell carried out the software implementation, model fitting, and data collection; Greenwell and Beran analyzed the data, and all authors contributed to the writing the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 *Multi-scale Quantum Models for Biocatalysis*, ed. D. M. York and T.-S. Lee, Springer Netherlands, Dordrecht, 2009, vol. 7.

2 A. J. Cruz-Cabeza and J. Bernstein, *Chem. Rev.*, 2014, **114**, 2170–2191.

3 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.

4 G. Santra and J. M. L. Martin, *AIP Conf. Proc.*, 2019, **2186**, 030004.

5 J. M. L. Martin and G. Santra, *Israel J. Chem.*, 2020, **60**, 787–804.

6 N. Mardirossian and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2014, **16**, 9904.

7 N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.*, 2016, **144**, 214110.

8 N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.*, 2018, **148**, 241736.

9 S. Kozuch and J. M. L. Martin, *J. Comput. Chem.*, 2013, **34**, 2327–2344.

10 G. Santra, N. Sylvetsky and J. M. L. Martin, *J. Phys. Chem. A*, 2019, **123**, 5129–5143.

11 C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *J. Chem. Phys.*, 2013, **139**, 134101.

12 J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev and A. E. Roitberg, *Nature Commun.*, 2019, **10**, 2903.

13 R. T. McGibbon, A. G. Taube, A. G. Donchev, K. Siva, F. Hernández, C. Hargus, K.-H. Law, J. L. Klepeis and D. E. Shaw, *J. Chem. Phys.*, 2017, **147**, 161725.

14 A. J. Cohen, P. Mori-Sanchez and W. Yang, *Science*, 2008, **321**, 792–794.

15 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.*, 2012, **112**, 289–320.

16 T. Bally and G. N. Sastry, *J. Phys. Chem. A*, 1997, **101**, 7923–7925.

17 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 194112.

18 A. D. Dutoi and M. Head-Gordon, *Chem. Phys. Lett.*, 2006, **422**, 230–233.

19 B. G. Johnson, C. A. Gonzales, P. M. Gill and J. A. Pople, *Chem. Phys. Lett.*, 1994, **221**, 100–108.

20 O. V. Gritsenko, B. Ensing, P. R. T. Schipper and E. J. Baerends, *J. Phys. Chem. A*, 2000, **104**, 8558–8565.

21 S. Patchkovskii and T. Ziegler, *J. Chem. Phys.*, 2002, **116**, 7806–7813.

22 S. Andersson and M. Grüning, *J. Phys. Chem. A*, 2004, **108**, 7621–7636.

23 L. M. LeBlanc, S. G. Dale, C. R. Taylor, A. D. Becke, G. M. Day and E. R. Johnson, *Angew. Chem. Int. Ed.*, 2018, **57**, 14906–14910.

24 J. Nyman, L. Yu and S. M. Reutzel-Edens, *CrystEngComm*, 2019, **21**, 2080–2088.

25 C. Greenwell, J. L. McKinley, P. Zhang, Q. Zeng, G. Sun, B. Li, S. Wen and G. J. O. Beran, *Chem. Sci.*, 2020, **11**, 2200–2214.

26 C. Greenwell and G. J. O. Beran, *Cryst. Growth Des.*, 2020, **20**, 4875–4881.

27 C. Greenwell and G. J. O. Beran, *J. Mater. Chem. C*, 2021, **9**, 2848–2857.

28 S. Grimme, L. Goerigk and R. F. Fink, *WIRES: Comput. Mol. Sci.*, 2012, **2**, 886–906.

29 S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095.

30 Á. Szabados, *J. Chem. Phys.*, 2006, **125**, 214105.

31 R. F. Fink, *J. Chem. Phys.*, 2010, **133**, 174113.

32 R. A. Distasio and M. Head-Gordon, *Mol. Phys.*, 2007, **105**, 1073–1083.

33 J. G. Hill and J. A. Platts, *J. Chem. Theory Comput.*, 2007, **3**, 80–85.

34 R. A. King, *Molec. Phys.*, 2009, **107**, 789–795.

35 J. Rigby and E. I. Izgorodina, *J. Chem. Theory Comput.*, 2014, **10**, 3111–3122.

36 S. Tan, S. Barrera Acevedo and E. I. Izgorodina, *J. Chem. Phys.*, 2017, **146**, 064108.

37 Y. Jung, R. C. Lochan, A. D. Dutoi and M. Head-Gordon, *J. Chem. Phys.*, 2004, **121**, 9793–802.

38 R. C. Lochan, Y. Jung and M. Head-Gordon, *J. Phys. Chem. A*, 2005, **109**, 7598–7605.

39 S. Grimme, *J. Comp. Chem.*, 2003, **24**, 1529–1537.

40 T. Takatani, E. G. Hohenstein and C. D. Sherrill, *J. Chem. Phys.*, 2008, **128**, 124111.

41 M. Pitonak, J. Rezac and P. Hobza, *Phys. Chem. Chem. Phys.*, 2010, **12**, 9611–9614.

42 S. Grimme and E. I. Izgorodina, *Chem. Phys.*, 2004, **305**, 223–230.

43 Y. M. Rhee and M. Head-Gordon, *J. Phys. Chem. A*, 2007, **111**, 5314–26.

44 D. Casanova, Y. M. Rhee and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 164106.

45 A. Hellweg, S. A. Grün and C. Hättig, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4119.

46 S. M. Cybulski and M. L. Lytle, *J. Chem. Phys.*, 2007, **127**, 141102.

47 A. Hesselmann, *J. Chem. Phys.*, 2008, **128**, 144112.

48 M. Pitonak and A. Hesselmann, *J. Chem. Theory Comput.*, 2010, **6**, 168–178.

49 L. A. Burns, M. S. Marshall and C. D. Sherrill, *J. Chem. Phys.*, 2014, **141**, 234111.

50 J. Řezáč, C. Greenwell and G. J. O. Beran, *J. Chem. Theory Comput.*, 2018, **14**, 4711–4721.

51 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

52 A. Tkatchenko, R. A. DiStasio, M. Head-Gordon and M. Scheffler, *J. Chem. Phys.*, 2009, **131**, 094106.

53 E. Caldeweyher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2017, **147**, 034112.

54 A. D. Becke and E. R. Johnson, *J. Chem. Phys.*, 2007, **127**, 154108.

55 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.

56 A. Ambrosetti, A. M. Reilly, R. A. DiStasio and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.

57 J. Řezáč, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2011, **7**, 2427–2438.

58 E. R. Johnson, P. Mori-Sánchez, A. J. Cohen and W. Yang, *J. Chem. Phys.*, 2008, **129**, 204112.

59 J. Řezáč, *J. Comp. Chem.*, 2016, **37**, 1230–1237.

60 T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639–9646.

61 R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Theory Comput.*, 2017, **13**, 3185–3197.

62 F. Neese, *WIREs Comput. Molec. Sci.*, 2012, **2**, 73–78.

63 S. Kozuch, D. Gruzman and J. M. L. Martin, *J. Phys. Chem. C*, 2010, **114**, 20801–20808.

64 L. Goerigk and S. Grimme, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6670.

65 L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. Smith, K. Vanommeslaeghe, A. D. MacKerell, K. M. Merz and C. D. Sherrill, *J. Chem. Phys.*, 2017, **147**, 161727.

66 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *Journal of Machine Learning Research*, 2011, **12**, 2825–2830.

67 J. Řezáč, Y. Huang, P. Hobza and G. J. O. Beran, *J. Chem. Theory Comput.*, 2015, **11**, 2065–3079.

68 M. S. Marshall, L. A. Burns and C. D. Sherrill, *J. Chem. Phys.*, 2011, **135**, 194102.

69 J. Řezáč, *J. Chem. Theory Comput.*, 2020, **16**, 2355–2368.

70 J. Řezáč and A. de la Lande, *J. Chem. Theory Comput.*, 2015, **11**, 528–537.

71 D. Gruzman, A. Karton and J. M. L. Martin, *J. Phys. Chem. A*, 2009, **113**, 11974–11983.

72 U. R. Fogueri, S. Kozuch, A. Karton and J. M. Martin, *J. Phys. Chem. A*, 2013, **117**, 2269–2277.

73 S. Grimme, M. Steinmetz and M. Korth, *J. Org. Chem.*, 2007, **72**, 2118–2126.

74 K. U. Lao, R. Schäffer, G. Jansen and J. M. Herbert, *J. Chem. Theory Comput.*, 2015, **11**, 2473–2486.

75 S. P. Thomas and M. A. Spackman, *Austral. J. Chem.*, 2018, **71**, 279.

76 S. Grimme, C. Diedrich and M. Korth, *Angew. Chem. Int. Ed.*, 2006, **45**, 625–629.

77 G. J. O. Beran, *CrystEngComm*, 2019, **21**, 758–764.

78 L. Yu, *Acc. Chem. Res.*, 2010, **43**, 1257–1266.

79 M. Tan, A. G. Shtukenberg, S. Zhu, W. Xu, E. Dooryhee, S. M. Nichols, M. D. Ward, B. Kahr and Q. Zhu, *Faraday Disc.*, 2018, **211**, 477–491.

80 K. S. Gushurst, J. Nyman and S. X. M. Boerrigter, *CrystEngComm*, 2019, **21**, 1363–1368.

81 A. R. Tyler, R. Ragbirsingh, C. J. McMonagle, P. G. Waddell, S. E. Heaps, J. W. Steed, P. Thaw, M. J. Hall and M. R. Probert, *Chem*, 2020, **6**, 1755–1765.

82 A. Lévesque, T. Maris and J. D. Wuest, *J. Am. Chem. Soc.*, 2020, **142**, 11873–11883.

83 X. Li, X. Ou, H. Rong, S. Huang, J. Nyman, L. Yu and M. Lu, *Cryst. Growth Des.*, 2020, **20**, 7093–7097.

84 L. Yu, G. A. Stephenson, C. A. Mitchell, C. A. Bunnell, S. V. Snorek, J. J. Bowyer, T. B. Borchardt, J. G. Stowell and S. R. Byrn, *J. Am. Chem. Soc.*, 2000, **122**, 585–591.

85 S. Chen, I. A. Guzei and L. Yu, *J. Am. Chem. Soc.*, 2005, **127**, 9881–9885.

86 S. Chen, H. Xi and L. Yu, *J. Am. Chem. Soc.*, 2005, **127**, 17439–17444.

87 L. Yu, J. Huang and K. J. Jones, *J. Phys. Chem. B*, 2005, **109**, 19915–19922.

88 R. O. Al-Kaysi, A. M. Müller and C. J. Bardeen, *J. Am. Chem. Soc.*, 2006, **128**, 15938–15939.

89 L. Zhu, R. O. Al-Kaysi and C. J. Bardeen, *J. Am. Chem. Soc.*, 2011, **133**, 12569–12575.

90 S. R. Jezowski, L. Zhu, Y. Wang, A. P. Rice, G. W. Scott, C. J. Bardeen and E. L. Chronister, *J. Am. Chem. Soc.*, 2012, **134**, 7459–7466.

91 L. Zhu, F. Tong, C. Salinas, M. K. Al-Muhanna, F. S. Tham, D. Kisailus, R. O. Al-Kaysi and C. J. Bardeen, *Chem. Mater.*, 2014, **26**, 6007–6015.