Balance Between Physical Interpretability and Energetic Predictability in Widely Used Dispersion-Corrected Density Functionals

Saswata Dasgupta,[*,†] Etienne Palos,[†,§] Yuanhui Pan,[†,§] and Francesco Paesani[*,†,‡,¶]

†Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States

‡Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, United States

¶San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, United States

§Contributed equally to this work

E-mail: s1dasgupta@ucsd.edu; fpaesani@ucsd.edu

1

Abstract

We assess the performance of different dispersion models for several popular density functionals across a diverse set of non-covalent systems, ranging from the benzene dimer to molecular crystals. By analyzing the interaction energies and their individual components, we demonstrate that there exists variability across different systems for empirical dispersion models, which are calibrated for reproducing the interaction energies of specific systems. Thus, parameter fitting may undermine the underlying physics, as dispersion models rely on error compensation among the different components of the interaction energy. Energy decomposition analyses reveal that, the accuracy of revPBE-D3 for some aqueous systems originates from significant compensation between dispersion and charge transfer energies. However, revPBE-D3 is less accurate in describing systems where error compensation is incomplete, such as the benzene dimer. Such cases highlight the propensity for unpredictable behavior in various dispersion-corrected density functionals across a wide range of molecular systems, akin to the behavior of force fields. On the other hand, we find that SCAN-rVV10, a targeted-dispersion approach, affords significant reductions in errors associated with the lattice energies of molecular crystals, whilst it has limited accuracy in reproducing structural properties. Given the ubiquitous nature of non-covalent interactions and the key role of density functional theory in computational sciences, the future development of dispersion models should prioritize the faithful description of the dispersion energy, a shift that promises greater accuracy in capturing the underlying physics across diverse molecular and extended systems.

INTRODUCTION

Non-covalent interactions play a central role in a broad range of scientific fields, including but not limited to chemistry, condensed matter physics, and biology.[1–6] In recent years, there has been significant interest in developing efficient and accurate methods to study non-covalent interactions theoretically. Although coupled cluster theory including single, double,

2

and perturbative triple excitations in the complete basis set limit, i.e., CCSD(T)/CBS, currently represents the "gold standard" for modeling non-covalent interactions,[7] the associated computational cost limits its applicability to small systems. In this context, the recent development of the domain-based local pair natural orbital CCSD(T), i.e., DLPNO-CCSD(T), represents significant progress toward achieving high accuracy for large non-covalent systems.[8] Comparatively inexpensive, second-order Møller-Plesset perturbation theory (MP2),[9] has been widely used for systems consisting of a few tens of atoms. It has, however, been found that MP2 displays poor convergence behavior for systems dominated by dispersion interactions, such as $\pi - \pi$ stacking.[10–13] On the other hand, Kohn-Sham density functional theory (DFT) has been successful in handling larger systems albeit with limited ability in describing non-covalent interactions,[14] especially the London dispersion interaction (hereafter simply referred to as dispersion interaction), which is an electron correlation effect. In the limit of two atoms, the dispersion interaction decays as R^{-6}, where R is the interatomic distance.[15]

By definition, the dispersion interaction is an intrinsic component of the total interaction energy and is always attractive. The principal driving force behind the dispersion interaction is the correlated motion of electrons, particularly over extended distances. Using semi-local and hybrid density functionals (DFs), the dependence of the dispersion interactions on the intermolecular distances is challenging to capture with the exponentially decaying electron density.[16,17] None of the conventional components of a DF, such as the local electron density, its gradient, or its kinetic energy, fully account for the correlated motion of electrons over the appropriate range of distances necessary for a correct description of dispersion interactions. Although hybrid DFs incorporate long-range effects (nonlocality) through Hartree-Fock exchange, they remain local in correlation and are, consequently, insufficient for accurately describing the R^{-6} asymptotic behavior of the dispersion energy.

The limitations of conventional DFs become clear when examining the underlying wavefunction and its role in the dispersion energy. Within MP2 theory,[9] the dispersion energy

3

arises from Coulomb and exchange interactions involving single-electron transition densities associated with two interacting fragments:

$$E_{\text{disp}}^{(2)} = -\sum_{ia} \sum_{jb} \frac{(ia \mid jb)\big[(ia \mid jb) - (ib \mid ja)\big]}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{1}$$

As shown in eq 1, $E_{\text{disp}}^{(2)}$ involves a sum over all possible single-particle hole excitations from orbital i to orbital a (localized on the first fragment) and from orbital j to orbital b (localized on the second fragment), with the term $(ia \mid jb)$ representing a two-electron integral, and ε_i, ε_a, ε_j and ε_b corresponding to the energies of orbitals i, a, j, and b, respectively. It follows that the term $(ia \mid jb)$ involves instantaneous electron correlations arising from electromagnetic zero-point energy fluctuations, leading to virtual excitations and electrostatic interactions. Conventional DFs, which only consider electron exchange without using virtual orbitals, fall short in adequately representing these instantaneous electron correlations. While DFT-related virtual-orbital dependent methods, such as the random phase approximation (RPA),[18–20] are able to accurately capture nonlocal correlations, they are significantly more computationally expensive than conventional DFs.[21]

An accurate treatment of the dispersion energy is essential for understanding the stability and properties of diverse molecular systems. To overcome the limitations of conventional DFs to describe dispersion interactions, the dispersion energy in Kohn-Sham DFT is typically incorporated as an additive correction.[22–36] Dispersion correction schemes commonly employed in DFT calculations can be categorized into five main approaches:

a) **Nonlocal density-based schemes that introduce corrections to the electronic potential**: These nonlocal schemes, which involve a supermolecular calculation of both the total system's and fragments' energies to determine the interaction energy, are widely used to compute the dispersion energy in various systems based on their electron density.[37] All van der Waals density functional (vdW-DF) schemes employ an approximation for the exchange-correlation energy using standard exchange and correlation components for short-

4

range contributions and a nonlocal term for the dispersion energy. Earlier versions of vdW-DF required empirical damping functions,[38] but modern schemes provide improved results by going beyond local approximations in the nonlocal correlation kernel.[33–35,39,40] Several variants of the vdW-DF scheme exist that use different nonlocal correlation kernels based on approximations of the dipole polarizability.[41] The dispersion energy is then determined by integrating the polarizability using the Casimir-Polder relationship,[42] which enables the calculation of the C_6 dispersion coefficients for the interacting fragments. Dielectric functions (e.g., the Drude model) and local plasma frequencies are used to relate the local polarizability to the electron density.[43,44] Various versions of the vdW-DF scheme have been reported, including vdW-DF,[45,46] vdW-DF2,[39] VV09,[33] and VV10.[35] One advantage of the vdW-DF schemes is the ability to naturally include dispersion effects based on the charge density, accounting for the dependence on charge transfer.[37]

b) **Semiclassical schemes based on C_6 parameters that primarily modify the total energy**: The DFT-D approach[23–26,36,47–50] involves augmenting DFT calculations with a damped dispersion energy that takes into account the interactions between pairs of atoms. Over the years, several variants of the DFT-D scheme have been proposed, each with its own damping functions and refinement strategies.[24–26,47–49,51–53] Among these approaches, the DFT-D3 scheme[26,51–53] has emerged as one of the most widely used, enhancing the accuracy of the DFT-D scheme, providing broader applicability, and reducing empiricism. DFT-D3 incorporates atom-specific pairwise-additive dispersion coefficients, refined cutoff radii computed from "first principles", and system-dependent information through fractional coordination numbers.[26] DFT-D3 requires adjustment of three global parameters for each DF to provide the dispersion energy for molecules and solids. The next-generation, DFT-D4, exhibits improved performance by incorporating atomic charge information to enhance the "geometry-only" model.[36] In the DFT-D4 scheme, computed atom-in-molecule dynamic polarizabilities are scaled using element-specific functions that are derived from Mulliken-type atomic charges.[36] The DFT-D4 scheme retains the strengths of DFT-D3 while introducing

5

charge dependence and technical refinements, such as less empirical functions and classical electronegativity equilibration (EEQ) partial charges that are adopted instead of Mulliken partial charges.[36]

c) **Many-body dispersion model:** The many-body dispersion (MBD) model developed by Tkatchenko and co-workers has gained popularity in recent times.[54] It originates from the Tkatchenko-Scheffler model in which C_6 coefficients and vdW radii are determined from the mean-field electron density.[55] Tkatchenko and Lilienfeld estimated 2-body and 3-body dispersion contributions using C_6 and C_9 coefficients within a many-body expansion formalism.[56] The inclusion of 3-body dispersion contributions was found to be necessary for many non-covalent interactions, leading to the development of the MBD model. The MBD approach incorporates long-range screening effects (SCS) and nonadditive dispersion energy using the coupled fluctuating dipole model (CFDM).[57] The MBD model is based on the adiabatic-connection fluctuation-dissipation theorem, providing an exact expression for the total electron correlation energy.[19] The original MBD model was shown to suffer from underbinding issues due to short-range correlation effects, which led to the development of the MBD@rsSCS model that adopts range separation.[58] In the following, we will refer to MBD@rsSCS as MBD for simplicity.

d) **Exchange-hole dipole moment (XDM) model**: In a less empirical fashion, the dispersion energy can be incorporated into Kohn-Sham DFT by means of the XDM model developed by Becke and Johnson.[27,59,60] The XDM model relies on the observation that in a non-overlapping system the nonzero dipole moment of an exchange-hole can induce an instantaneous dipole moment that leads to dispersion interactions.

e) **Effective one-electron potentials**: The London dispersion interaction resulting from correlated electron motion can be empirically described via effective one-electron potentials. Two common approaches include atom-centered external potentials[22] and semi-local DFs, i.e., functionals derived within the generalized gradient approximation (GGA) as well as meta-GGA functionals, describing dispersion interactions.[61] The concept of dispersion-

6

corrected atom-centered potentials interprets London dispersion forces as arising from distorted charge distributions, inducing a dipole moment in each atom.[22,62–65] Alternatively, semi-local DFs[61,66,67] can glean information about the dispersion interaction from the total density and its deformation caused by the overlapping of wavefunctions.

Since most dispersion models currently used in studies involve empirical elements in various ways, it is important to conduct rigorous benchmarking against reliable experimental or high-level theoretical reference data to ensure the accuracy and validity of the results obtained with different dispersion models. The conventional parameterization of dispersion coefficients relies on evaluating the error in interaction energies of datasets involving non-covalently interacting dimers, such as the S22,[68] S22x5,[69] S66,[70] and S66×8[71] datasets. By design, such an approach does not necessarily yield optimized parameters that accurately describe the dispersion energy. Instead, it tends to determine an optimal set of parameters that reproduce the total interaction energy, which, however, may be achieved through error compensation among different energy components. Although the incorporation of dispersion models does apparently enhance the overall accuracy of conventional DFs, there remains uncertainty regarding whether empirically fitted dispersion models can also lead to physical interpretability as they might not be able to correctly reproduce the dispersion energy contribution to the interaction energy. Furthermore, it has been observed that globally optimized dispersion parameters may not perform as well as dispersion parameters optimized specifically for a certain class of systems.[72,73] In the context of *ab initio* molecular dynamics (AIMD) simulations, the error resulting from inaccurate dispersion coefficients can be amplified due to the size of the system, leading to potentially incorrect predictions of physical behavior and, consequently, misleading interpretations of the results.

To address these challenges, this study aims to investigate the role of different models in representing the dispersion energy in various non-covalently bound systems and explore their impact on both physical interpretability and energetic predictability. Through thorough analyses and benchmarking against high-quality reference data, we aim to enhance our

7

understanding of dispersion interactions in molecular systems.

THEORY AND METHODS

Energy Calculations

The interaction energies for non-covalently interacting systems were calculated as follows:

$$E_{\text{Int}} = E_{\text{N}} - \sum_i^N E_i \qquad (2)$$

where E_{Int}, E_{N}, and E_i represent the interaction energies, total energy of the N-monomer system, and monomer energies, respectively. The energy calculations for all datasets were performed using Q-Chem 6.0[74] with the PBE,[75] revPBE,[76] and PBE0[77] functionals combined with various dispersion models, including the -D3,[26,51] -D4,[36] -VV10,[35] and -MBD models.[58] The calculations were carried out with the def2-QZVPPD basis set[78] unless stated otherwise using the SG-3 integration grid.[79] For the S66×8 dataset, we used both the zero-damping, -D3(0),[26] and the Becke-Johnson damping, -D3(BJ),[51] schemes. The -D4 dispersion correction was calculated for DFs that include the 3-body dispersion term by default. The -VV10 dispersion model[35] was incorporated with all DFs. The -VV10 b parameters were set to 6.5 for PBE, 3.6 for revPBE, and 6.6 for PBE0, with C fixed at its original value ($C = 0.0093$) as determined in ref 80. The errors for the interaction energies of water clusters, $(H_2O)_n$, were calculated relative the reference values obtained at the CCSD(T)/CBS level of theory in ref 81. The errors for the interaction energies of the hydrate sodium ion, $Na^+(H_2O)_n$, and hydrated chloride ion, $Cl^-(H_2O)_n$, clusters were calculated relative to DLPNO-CCSD(T)[82] calculations performed using the def2-QZVPPD basis set in combination with the def2-QZVPPD/C auxiliary basis set,[83] as implemented in ORCA 4.2.1.[84]

8

Energy Decomposition Analysis (EDA)

The second-generation absolutely localized molecular orbital energy decomposition analysis (ALMO-EDA) method was used to decompose the interaction energy of a given N-monomer system, E_{Int}, into a sum of physical contributions.[85–87] Specifically, the interaction energy was decomposed as

$$E_{\text{Int}} = E_{\text{Pol}} + E_{\text{Frz}} + E_{\text{CT}} + E_{\text{Disp}} \tag{3}$$

where E_{Pol}, E_{Frz}, E_{CT}, and E_{Disp} are the polarization, frozen, charge transfer, and dispersion energies of the N-monomer system, respectively. Based on this decomposition, E_{Frz} is the sum of energy contributions associated with permanent electrostatics and Pauli repulsion.[87] The ALMO-EDA calculations were performed using the PBE, revPBE, and PBE0 functionals without any dispersion correction as well as with the -D3, -D4, and -VV10 dispersion models (see Supporting Information for details). For the dispersion-corrected PBE and revPBE functionals, revPBE was used as the dispersion-free functional in the ALMO-EDA calculations. The b and C parameters of the -VV10 model for PBE, revPBE, and PBE0 were obtained from ref 80. All ALMO-EDA calculations used the def2-QZVPPD basis set[88] except those for the aqueous clusters. The ALMO-EDA calculations for the water clusters were performed using the def2-QZVPP basis set, while the def2-TZVPPD basis set[89] was used in the ALMO-EDA calculations for the hydrated sodium and chloride clusters. The reference ALMO-EDA energies were calculated with the ωB97M-V density functional[90] using the corresponding basis sets. The selection of ωB97M-V as the reference DF is based on its accuracy in describing non-covalent interactions as demonstrated in previous studies.[91–93]

Cell Optimizations and Calculations of Lattice Energies

The geometry optimizations of the X23 dataset of molecular crystals[94] and corresponding isolated molecules, followed by the calculations of the lattice energies, were performed using the Vienna Ab initio Simulation Package (VASP).[95,96] The core electrons were treated by

9

the projector augmented wave (PAW) pseudopotential.[97] For these calculations, we used PBE and revPBE as representative of GGA functionals[98,99] with the -D3(BJ) and -MBD dispersion models. The range-separation parameter $\beta = 0.54$ for revPBE-MBD was obtained from ref 100. VASP supports the revised version of the -VV10 model, known as -rVV10.[101] We used the strongly constrained and appropriately normed (SCAN)[67] as a representative meta-GGA functional with and without -rVV10,[101] as both SCAN and SCAN-rVV10[102] are becoming increasingly popular for modeling diversely bonded systems and materials.[103–106] In this context, it is important to mention that in the version of VASP available to us (VASP 5.4.4), DFT calculations with -rVV10 are only supported for meta-GGA functionals such as SCAN. The $b = 15.7$ and $C = 0.0093$ parameters of the -rVV10 model for SCAN were obtained from ref 102. In the optimizations of the molecular crystals, the unit cell parameters as well as the atomic coordinates were allowed to relax. For the optimizations and energy calculations involving the isolated molecules, we used large cubic unit cells with each side measuring 25 Å to minimize finite-size effects. A plane-wave basis set with an energy cutoff of 1000 eV was used in all calculations. The convergence criteria for the energies and forces were set to 1×10^{-5} eV and 0.005 eV Å^{-1}, respectively. Given the different unit cell dimensions along different directions of the crystals in the X23 dataset, Γ-centered k-point grids with a k-point resolution of $2\pi \times 0.03$ Å^{-1} were used for the structural optimizations, which were generated using VASPKIT.[107] For the isolated molecules, only the Γ point was used. The lattice energy of each molecular crystal was calculated as

$$E_{\text{latt}} = \frac{E_{\text{crys}}}{n_{\text{mol}}} - E_{\text{iso}}, \tag{4}$$

where E_{crys} is the energy of the molecular crystal unit cell, n_{mol} is the number of molecules in the crystal unit cell, and E_{iso} is the energy of the corresponding isolated molecule.

RESULTS AND DISCUSSION

Interaction Energies and Energy Decomposition Analyses of the S66×8 Dataset

S66×8[71] is an exemplary dataset for studying non-covalent interactions, encompassing a diverse range of interactions such as hydrogen bonding, $\pi - \pi$ stacking, dispersion, and mixed interactions involving both electrostatics and dispersion. In addition to dimer configurations in their equilibrium geometries, it also includes stretched and compressed dimer configurations, thereby assessing the influence of dispersion interactions at both short and long ranges. The interaction energies within the S66×8 dataset were calculated using PBE, revPBE, and PBE0 in combination with the -D3, -D4, -VV10, and -MBD models. The accuracy of these calculations was assessed relative to the CCSD(T)/CBS reference values. In the case of the -D3 model, we considered both -D3(0) and -D3(BJ) damping schemes.

Figure 1a reports the error statistics for the complete S66×8 calculated with PBE combined with the different dispersion models. The dispersion-corrected PBE functionals exhibit remarkably similar error statistics across the entire dataset, with mean unsigned errors (MUEs) ranging between 0.29 and 0.42 kcal/mol. Among the dispersion-corrected PBE functionals, PBE-D4 and PBE-D3(0) demonstrate the best and poorest performance, respectively. Interestingly, when -D3(0) is replaced by -D3(BJ), the accuracy of the corresponding dispersion-corrected PBE functional, i.e., PBE-D3(BJ) exhibits significant better performance (MUE = 0.33 kcal/mol). It should be noted that the dispersion-corrected PBE functionals also exhibit comparable accuracy for the different subsets of interactions included in the S66×8 dataset as shown in Figure S1. Similar trends in accuracy are also observed for the dispersion-corrected PBE0 models (Figure 1c), with only a negligible uptick in error statistics found for the PBE0-D3(0) functional. In contrast, the accuracy of the dispersion-corrected revPBE functionals displays notable variations, with the empirical dispersion models (i.e., -D3 and -D4) performing surprisingly better than the more sophisticated
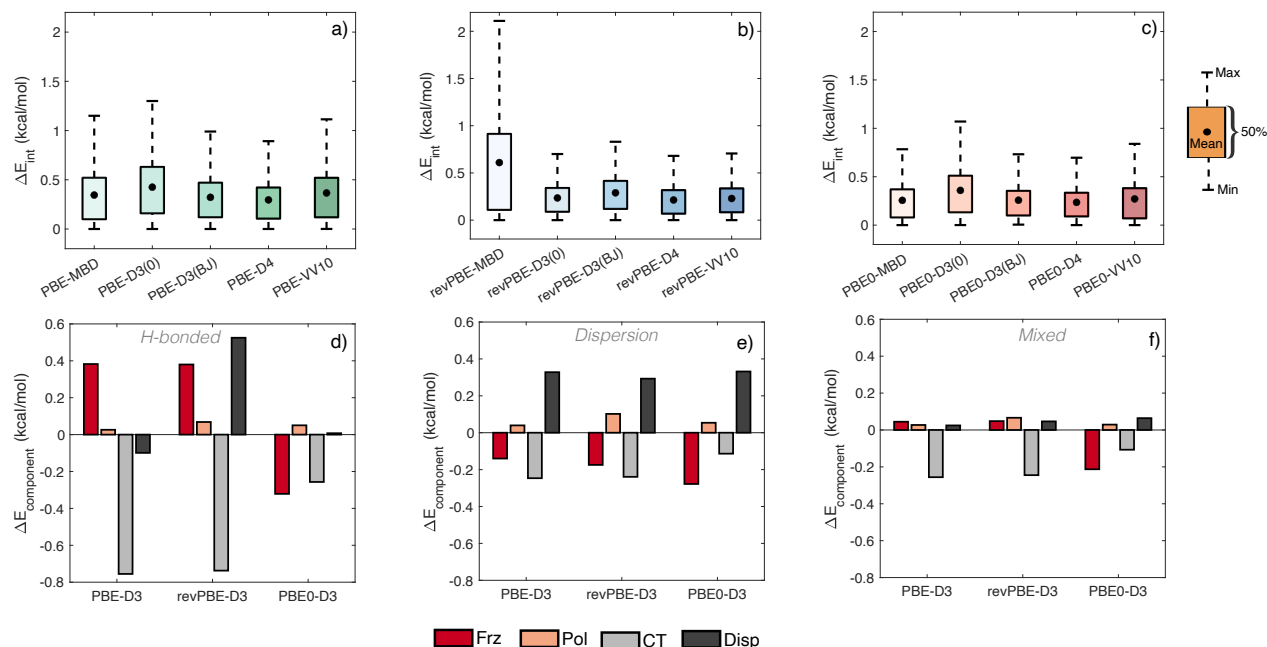
11

Figure 1: Box plots of absolute errors in total interaction energies relative to the CCSD(T)/CBS reference values calculated for the S66×8 dataset using: (a) PBE, (b) revPBE, and (c) PBE0 combined with the -MBD, -D3(0), -D3(BJ), -D4, and -VV10 models. Each box represents the interquartile range (IQR), the whiskers indicate the full range of the data, and the mean error is indicated by the dot. Panels (d-f) display the ALMO-EDA mean errors relative to ωB97M-V calculated for specific subsets of the S66×8 dataset: (d) hydrogen-bonding interactions, (e) dispersion-dominated interactions, and (f) mixed interactions. See main text for details.

-MBD model (Figure 1b). This discrepancy becomes particularly evident when analyzing dispersion-dominated and mixed interactions for which revPBE-MBD exhibits the highest error statistics among all dispersion models considered in this study (Figure S1). It is worth noting that the inclusion of the Axilrod-Teller-Muto (ATM) term to account for 3-body interactions[108] does not significantly alter the performance of PBE-D3(BJ), revPBE-D3(BJ), and PBE0-D3(BJ) for S66×8 (Figure S2).

Although the incorporation of dispersion models leads to a notable enhancement in the accuracy of density functionals in modeling non-covalent interactions within the S66×8 dataset,[109] it is important to consider that the choice of a specific dispersion model can impact the level of accuracy attained, as in the case of revPBE where the more physically

https://doi.org/10.26434/chemrxiv-2023-nc64t ORCID: https://orcid.org/0000-0002-4451-1203 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

motivated -MBD model exhibits larger errors across all individual subsets. For hydrogen-bonding interactions, the combination of -D3 or -D4 with revPBE results in notably better performance than when the same two dispersion models are combined with PBE and PBE0 (Figures S1). A previous study by Boese[110] also reported a significant error reduction with revPBE-D3. Because of this apparent good performance, revPBE-D3 has become a common density functional in AIMD simulations of hydrogen-bonded systems.[111–115]

When PBE and PBE0 are combined with the -D3, -D4, and -VV10 models, the dispersion errors for both stretched and compressed configurations show similar trends as when they are combined with the -MBD model (Figure S3). This seems to suggest that the -D3, -D4, and -VV10 models are able to effectively capture both short-range and long-range dispersion contributions. However, revPBE-D3, revPBE-D4, and revPBE-VV10 deviate significantly from the dispersion energy provided by revPBE-MBD, resulting in significant underbinding (Figure S3). Nevertheless, as the geometries become highly stretched and, consequently, dispersion contributions become nearly negligible, the differences between revPBE-D3, revPBE-D4, revPBE-VV10, and revPBE-MBD become smaller.

To investigate the surprising accuracy exhibited by revPBE-D3, ALMO-EDA calculations were conducted on the S66×8 dataset with the D3(0)-corrected PBE, revPBE, and PBE0 functionals. As shown by the analysis of hydrogen-bonded systems reported in Figure 1d, all three functionals display significant errors in the frozen energy, which, as mentioned above, encompasses permanent electrostatics and Pauli repulsion. Specifically, PBE-D3(0) and revPBE-D3(0) tend to underestimate the frozen energy, with mean errors of 0.39 and 0.38 kcal/mol, respectively, whereas PBE0-D3(0) tends to overestimate it, with a mean error of -0.32 kcal/mol. In terms of polarization contributions (denoted as Pol in Figure 1), all D3(0)-corrected functionals exhibit negligible errors. The GGA functionals (i.e., PBE and revPBE) exhibit substantial errors in charge transfer energies (denoted as CT in Figure 1) due to the presence of larger delocalization errors. Notably, the inclusion of 25% Hartree-Fock exchange, a known approach to mitigate delocalization errors, reduces the charge transfer

13

error as demonstrated by the PBE0 results. While PBE-D3 and PBE0-D3 exhibit small mean errors for the dispersion energy (denoted as Disp in Figure 1), revPBE-D3 exhibits a much larger mean error of 0.53 kcal/mol. The positive errors in the dispersion and frozen energies displayed by revPBE-D3 significantly counterbalance the negative error introduced by the charge transfer energy. The more pronounced error compensation among the different energy contributions thus explains why revPBE-D3 overall performs better than PBE-D3 for hydrogen-bonded systems.

Error compensation among different energy components as predicted by the various dispersion-corrected functionals is especially notable in systems primarily influenced by dispersion interactions. The positive error (i.e., underbinding) associated with the dispersion energy effectively offsets the negative errors (i.e., overbinding) associated with the frozen and charge transfer energies, resulting in an apparent better description of the interaction energy. In the case of mixed interactions, the errors for all energy components predicted by the various dispersion-corrected DFs is smaller, which also results in a more accurate representation of the dispersion energy.

Based on these analyses, it is natural to inquire about the underlying factors contributing to these significant discrepancies in the representation of the dispersion energy. The widely used -D3 and -D4 dispersion models require specific parameterization for any given DF. These parameters are optimized using extensive datasets, such as S66,[70] S22×5,[69] and S66×8,[71] which encompass various non-covalent interactions. The objective of this parameter fitting is to reproduce the reference interaction energies and not the actual dispersion contributions. However, different DFs suffer, to varying degrees, from inherent errors, such as self-interaction and delocalization errors,[116–118] which lead to inaccuracies in the representation of charge transfer energies. As shown by the ALMO-EDA results reported in Figure 1, significant negative charge transfer errors are often counterbalanced by positive errors in the representation of the dispersion energy by empirical dispersion models, which are particularly pronounced for revPBE-D3. This explains why empirical dispersion models, such as

14

-D3 and -D4, can sometimes outperform the more physically motivated -MBD model when combined with the revPBE functional.

These analyses thus demonstrate that balancing the errors associated with the different energy contributions to reproduce reference interaction energies through parameter fitting can result in an incorrect and crumbling representation of the dispersion energy, compromising the underlying physics while only apparently preserving the overall predictive accuracy in energetic calculations.

Role of the Dispersion Energy in the Benzene Dimer

The accurate prediction of the interaction energy of the benzene dimer serves as a rigorous test for the capability of a particular electronic structure method to adequately describe non-covalent interactions, as the attraction between two benzene molecules primarily arises from a delicate balance between dispersion and Pauli-repulsion, depending on the specific binding arrangement of the two molecules.[119] As a result, the interaction energy of the benzene dimer is relatively small, typically 2-3 kcal/mol, leading to a remarkably flat potential energy surface.[120]

Studies of the benzene dimer generally focus on three distinct configurations: "parallel-stacked", "parallel-displaced", and "T-shape" configurations.[120] Both theoretical and experimental investigations face significant challenges when attempting to determine the stationary points on the benzene dimer potential energy surface. Wavefunction methods, such as second-order perturbation theory encompassing SAPT0 and MP2, fail to provide a quantitative description of the dispersion energy.[121] MP2 overestimates the interaction energy of the "parallel-stacked" benzene dimer by a factor of 2, whereas it predicts the interaction energy for the T-shaped configuration with a 30% error due to this configuration's smaller dispersion energy.[122] The benzene dimer thus represents a prototypical system to assess the performance of various dispersion-corrected DFs.

Podeszwa *et al.*[120] used DFT-based symmetry adapted perturbation theory (DFT-SAPT)[124]

15

to scan the potential energy surface of the benzene dimer, identifying multiple station-ary points, including minima (M) and saddle (S) points. Bludský *et al.*[123] performed counterpoise-corrected optimizations at the PBE/CCSD(T) level of theory for these con-figurations using the accurate fixed benzene monomer geometry provided by Gauss and Stanton.[125]

We calculated the interaction energies between two minima (M1 and M2) and eight saddle points (S1-S8) using PBE, revPBE, and PBE0 combined with the -D3 and -VV10 dispersion models. In addition, we also calculated the interaction energies using ωB97M-V, as it exhibits high accuracy in describing non-covalent interactions among existing DFs.[92] All calculations were performed using the def2-QZVPPD basis set. As shown in Figure 2, ωB97M-V exhibits close agreement with the CCSD(T)/CBS reference values of ref 123, with an MUE of 0.12 kcal/mol and a maximum unsigned error of 0.25 kcal/mol. For comparison, the MUE for the PBE-D3 functional is 0.18 kcal/mol, with a maximum unsigned error of 0.35 kcal/mol. However, when PBE is combined with -VV10, the agreement improves significantly, with PBE-VV10 exhibiting a MUE of 0.08 kcal/mol and maximum unsigned error of 0.28 kcal/mol. In contrast, despite their accuracy in modeling the hydrogen-bonded
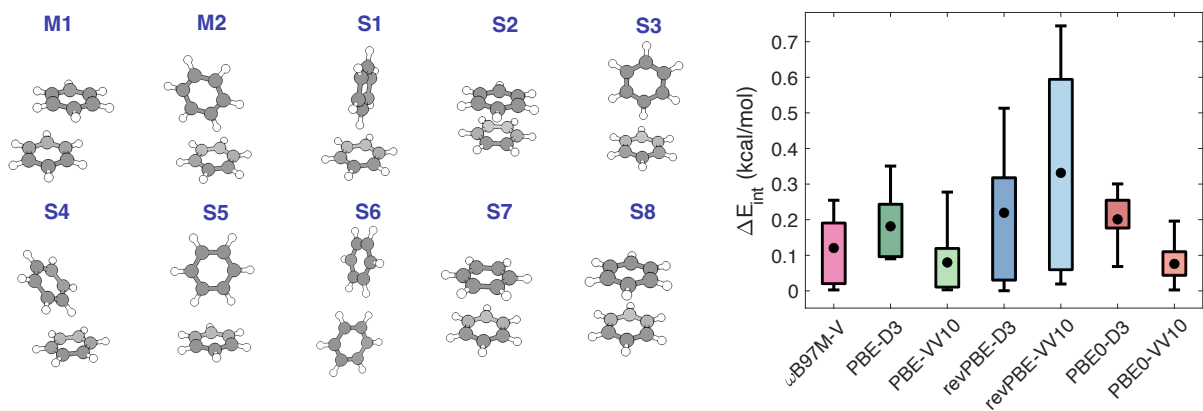

Figure 2: Box plot of absolute errors in interaction energies relative to the CCSD(T)/CBS reference values[123] calculated for 10 stationary points on the benzene dimer potential energy surface using ωB97M-V, PBE-D3, PBE-VV10, revPBE-D3, revPBE-VV10, PBE0-D3, and PBE0-VV10. See main text for details.

16

systems of the S66×8 dataset, both revPBE-D3 and revPBE-VV10 perform poorly on the benzene dimer. In particular, revPBE-D3 exhibits an MUE of 0.21 kcal/mol and a maximum unsigned error of 0.51 kcal/mol, with revPBE-VV10 performing even more poorly. Both PBE0-D3 and PBE0-VV10 demonstrate high accuracy, comparable to that of ωB97M-V.

The interaction energies of T-shaped benzene dimers (S3 saddle point) were calculated at intermolecular distances ranging from -2.0 to 2.0 Å with an increment of 0.2 Å (Figure 3a), and were further decomposed using ALMO-EDA to characterize the different energy contributions. In Figure 3a, we compare the errors in interaction energies predicted by PBE-D3, PBE-VV10, revPBE-D3, revPBE-VV10, PBE0-D3, and PBE0-VV10 with respect to the reference ωB97M-V values. This comparison reveals notable trends in the errors exhibited by each DF. First, PBE-D3 exhibit errors ranging from 0.07 to 0.43 kcal/mol relative to ωB97M-V, with the errors becoming smaller near the equilibrium distance, i.e., at zero displacement from the equilibrium distance between the centers of mass (COM) of the two benzene rings. Interestingly, PBE-VV10 exhibits slightly better performance for dimer configurations far from the equilibrium geometry. In contrast, revPBE-D3 exhibits large deviation from ωB97M-V, with errors ranging from 0.30 to 0.54 kcal/mol. Also in this case, the combination with the -VV10 model improves the overall performance of revPBE. Regarding the hybrid DFs, PBE0-D3 exhibits errors ranging from 0.15 to 0.35 kcal/mol, which consistently decrease as the benzene-benzene separation within the dimer approaches the equilibrium distance. Notably, PBE0-VV10 provides the best agreement with ωB97M-V as the errors range from -0.01 to 0.16 kcal/mol, indicating a significant improvement over the other combinations of DFs and dispersion models considered in this study. The results for revPBE-D3 are somewhat unexpected, considering that revPBE-D3 shows the best error statistics for the hydrogen-bonded systems in the S66×8 dataset (Figure 1). It thus becomes apparent that highly parameterized dispersion models cannot serve as a proxy for a universal solution to all problems concerning the shortcomings of any given DF.

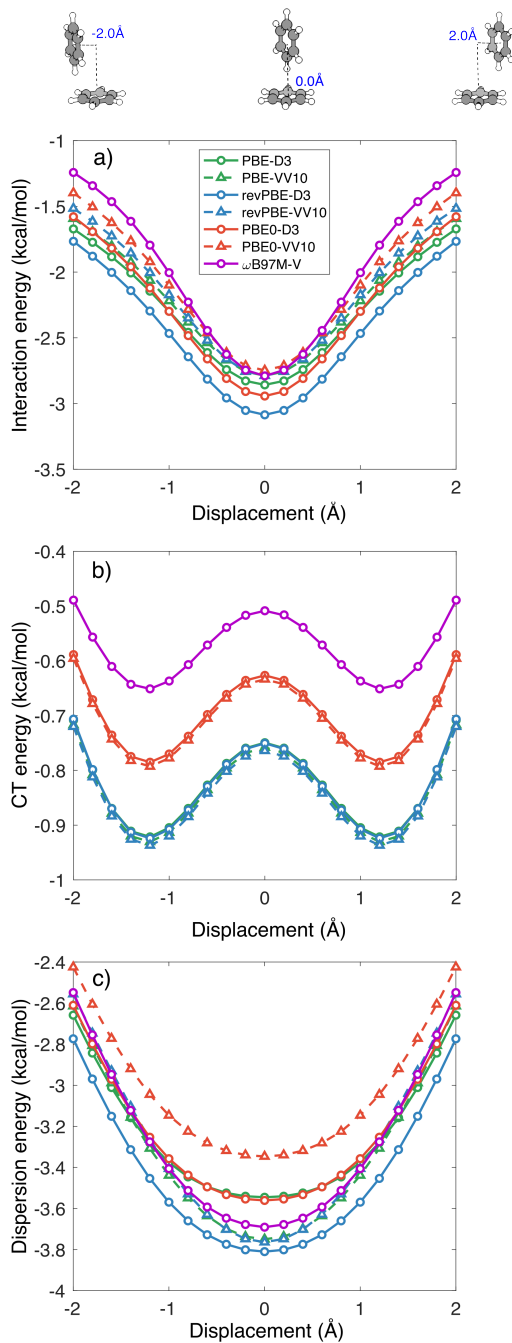After decomposing the interaction energies through ALMO-EDA calculations, ωB97M-V

17

Figure 3: (a) Interaction energies of the T-shaped benzene dimers (S3) calculated at intermolecular displacements from -2.0 to 2.0 Å relative to the equilibrium distance using PBE, revPBE, and PBE0 functionals combined with the -D3 and -VV10 models, as well as ωB97M-V. (b) Charge-transfer contributions to the interaction energies reported in (a). (c) Dispersion contributions to the interaction energies reported in (a).

provides charge transfer energies of -0.49 and -0.51 kcal/mol for the terminal and equilibrium dimer configurations along the intermolecular distance scan, respectively (Figure 3). In comparison, due to the delocalization error, both PBE and revPBE display larger charge transfer stabilization of -0.71 and -0.75 kcal/mol for the same structures, while PBE0 predicts charge transfer energies of -0.59 and -0.63 kcal/mol, respectively. Interestingly, all DFs examined in this study show optimal charge transfer for dimer configurations displaced by $\sim$1.2 Å relative to the equilibrium distance.

Figure 3c illustrates the impact of the different dispersion models on the performance of PBE, revPBE, and PBE0. For the terminal structures along the intermolecular distance scan (i.e., structures with a displacement of 2 Å relative to the equilibrium distance), PBE0-VV10 predicts the smallest dispersion energy (-2.43 kcal/mol). Both ωB97M-V and revPBE-VV10 predict a dispersion energy of -2.56 kcal/mol for the same structures, while the other dispersion-corrected DFs overestimate the stabilization due to the dispersion energy. Specifically, PBE0-D3 and PBE-VV10 predict a dispersion energy of -2.61 kcal/mol, while the dispersion energy predicted by PBE-D3 is -2.66 kcal/mol. Among all functionals examined in this study, revPBE-D3 predicts the largest overstabilization due to the dispersion energy (-2.78 kcal/mol).

In the case of the equilibrium geometry, where the dispersion energy is most prominent, the differences in dispersion energies predicted by the different dispersion-corrected DFs become more apparent. PBE0-VV10 exhibits significant understabilization (-3.34 kcal/mol) compared to ωB97M-V (-3.69 kcal/mol). Both PBE-D3 and PBE0-D3, which show slight overbinding compared to ωB97M-V for the displaced structure, exhibit underbinding for the equilibrium geometry, with dispersion energies of -3.54 and -3.55 kcal/mol, respectively. On the other hand, PBE-VV10 and revPBE-VV10 display overstabilization, predicting a dispersion energy of -3.75 kcal/mol. revPBE-D3 deviates the most from ωB97M-V also in the case of the equilibrium geometry, predicting a dispersion contribution of -3.80 kcal/mol.

The ALMO-EDA results demonstrate that, while ωB97M-V and PBE0-VV10 exhibit

similar interaction energy profiles (Figure 3a), this similarity is purely accidental, arising from substantial error compensation in PBE0-VV10, particularly near the equilibrium geometry. In particular, PBE0-VV10 displays an overbinding tendency in terms of charge transfer (Figure 3b) and frozen energies (Figure S4a), but it significantly underbinds the benzene dimer in terms of dispersion energy (Figure 3c) and slightly underestimates the polarization energy (Figure S4b). The same trend is observed for the other dispersion-corrected DFs where various levels of error compensation occur among the different energy components, which overall tend to improve the agreement with ωB97M-V. Although several dispersion-corrected DFs demonstrate similar accuracy to ωB97M-V for the benzene dimer (Figure 2), the ALMO-EDA results demonstrate that this apparent accuracy stems from the delicate balance of error compensation among the different energy components and not from a correct representation of the corresponding physical interactions. In particular, the DFs analyzed in this study tend to overestimate charge transfer and undermine polarization effects, while the fitted dispersion models attempt to counterbalance these errors by either overemphasizing or underemphasizing the dispersion contributions.

Non-Covalent Interaction in Charged Systems

The interaction energies in charged systems are influenced by the presence of the charges on the atoms, which affect the polarizability and alter the dispersion contributions. Charged systems interacting through non-covalent forces are included in the AHB21, IL16, and CHB6 datasets.[126] Each dataset focuses on specific types of interactions: anionic hydrogen bonding (AHB21), cation-neutral interactions (CHB6), and cation-anion pairs in ionic liquids (IL16).[126] In our analyses reported in Figure 4, we compare the results obtained with the different dispersion-corrected DFs with the CCSD(T)/CBS reference values reported in ref 126.

In the case of the AHB21 dataset, both PBE-D3 and PBE-D4 exhibit similar error statistics, while the magnitude of the errors increases when PBE is combined with -VV10.
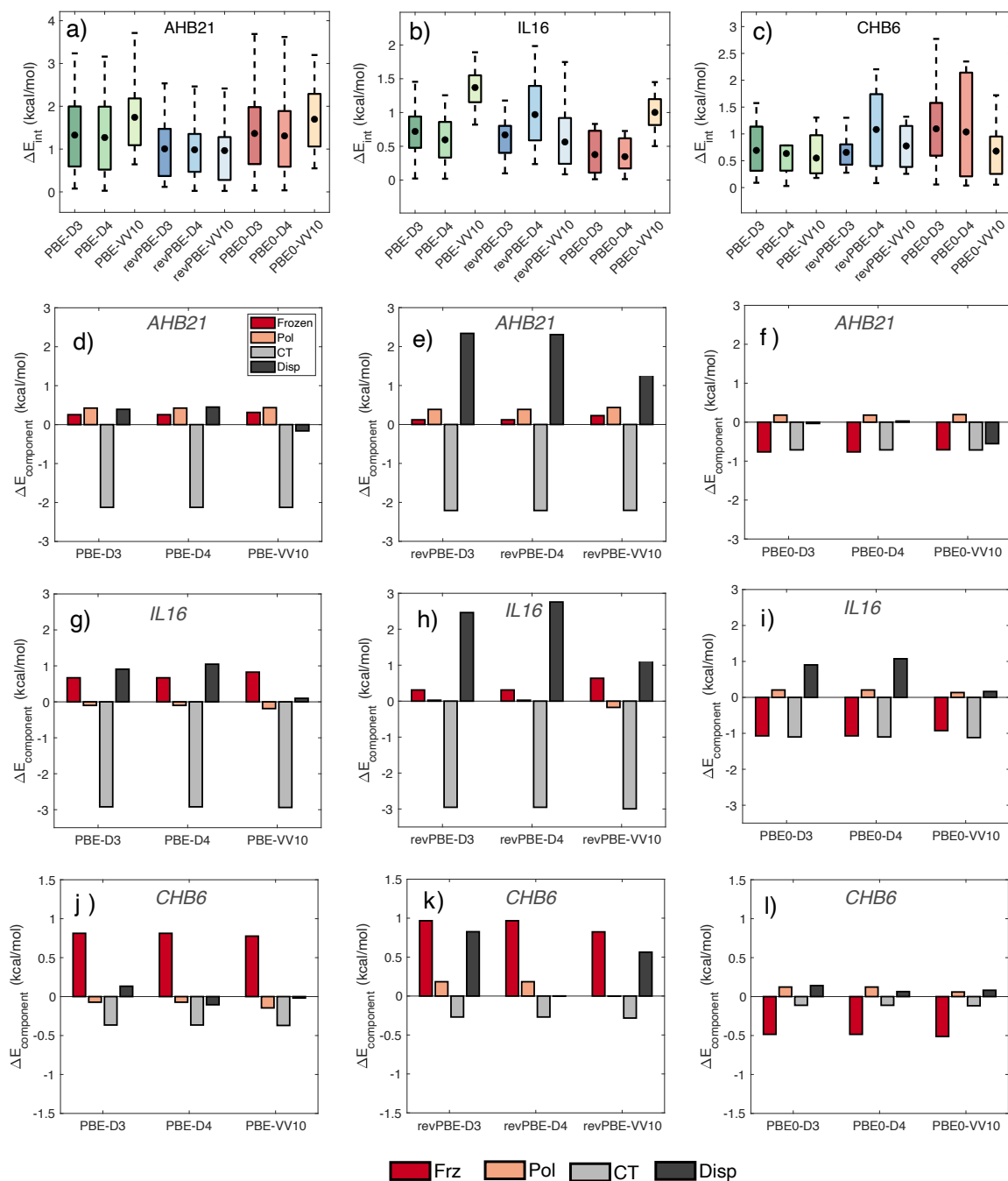
20

Figure 4: Box plots for the absolute errors in interaction energies relative to the CCSD(T)/CBS reference values[126] calculated for the AHB21 (a), IL16 (b), and CHB6 (c) datasets using PBE, revPBE, and PBE0 combined with the -D3, -D4, and -VV10 models. Panels (d-l) display the mean errors in individual energy components relative to the ωB97M-V reference values obtained from ALMO-EDA calculations carried out for specific subsets of the AHB21 (panels (d-f)), IL16 (panels (g-i)), and CHB6 (panels j-l) datasets.

21

All dispersion-corrected variants of revPBE (i.e., revPBE-D3, revPBE-D4, and revPBE-VV10) exhibit comparable error statistics. Upon application of the dispersion correction to the PBE0 functional, we find similar error statistics to those exhibited by the dispersion-corrected PBE functional, with PBE0-VV10 exhibiting larger errors than PBE0-D3 and PBE0-D4. As for the AHB21 dataset, the comparisons for the IL16 dataset indicate that the combination of the -VV10 model with PBE and PBE0 results in worse agreement with the reference values compared to when the same DFs are combined with the -D3 or -D4 models. In contrast, revPBE-D4 exhibit a larger MUE and a wider range of errors compared to revPBE-D3 and revPBE-VV10. Finally, for the CHB6 dataset (Figure 4), the combination of PBE and PBE0 with the -VV10 model results in slightly smaller errors compared to those exhibited by the two DFs when they are combined with the -D3 and -D4 models. Interestingly, combining revPBE with the -D4 model significantly increases the error and widens the range of errors compared to revPBE-D3 and revPBE-VV10. Notably, PBE0-D4 exhibits a wide range of errors similar to revPBE-D4, while PBE0-D3 and PBE0-D4 exhibit comparable MUEs.

Overall, the most consistent DF in producing the lowest MUEs across all three datasets is the dispersion-corrected revPBE functional. However, its performance varies with the type of dispersion model used. For instance, all dispersion-corrected revPBE functionals perform similarly in the case of the AHB21 dataset, while revPBE-VV10 and revPBE-D3 perform better for the IL16 and CHB6 datasets, respectively. Although all dispersion-corrected revPBE functionals exhibit reasonably good performance across the three datasets, no single dispersion-corrected DF consistently yield the lowest MUEs across all three datasets. Based on the analyses for PBE, revPBE, and PBE0, it thus follows that there is no definitive "best" combination of DF and dispersion model that consistently applies to all types of non-covalent interactions in charged systems.

To understand the trends in interaction energies in terms of the underlying physical contributions, ALMO-EDA calculations were performed on the three datasets using all

dispersion-corrected DFs examined in this study. The ALMO-EDA results are shown Figure 4d-l relative to the ωB97M-V references values. For the AHB21 dataset, the PBE functional predicts a significant charge transfer stabilization of approximately -2 kcal/mol. Both PBE-D3 and PBE-D4 predict slightly underbinding dispersion energies relative to the ωB97M-V reference values, while PBE-VV10 exhibits opposite behavior. As shown in Figure 4d, this delicate balance of errors improves the performance of PBE-D3 and PBE-D4 in predicting the interaction energies for systems included the AHB21 dataset. Conversely, all dispersion-corrected revPBE functionals consistently exhibit large (positive) errors for the dispersion energy, independently of the dispersion model employed. This underestimation of the dispersion energy is counteracted by the delocalization error, which is reflected in an artificial stabilization provided by the charge transfer energy. This pronounced error compensation leads to an apparent good agreement with the reference ωB97M-V values for the interaction energies (Figure 4e). On the other hand, the PBE0 functional yields smaller errors for all energy components. Being a hybrid functional, PBE0 mitigates the error in charge transfer energy by construction, while keeping the error in the dispersion energy consistently low. However, since the errors for all energy components of the dispersion-corrected PBE0 functionals display the same signs, their accumulation decreases the agreement with the ωB97M-V interaction energies compared to the corresponding dispersion-corrected PBE and revPBE functionals (Figure 4a).

The ALMO-EDA calculations for the IL16 dataset indicate that large dispersion errors negate the effects of over-stabilization caused by the delocalization error, which is particularly notable in the case of revPBE (Figure 4g-i). This error compensation is more prominent when revPBE is combined with the -D3 and -D4 models, while it is less marked in revPBE-VV10 due to a smaller dispersion error (Figure 4h). Similar error compensation improves the performance of PBE0-D3 and PBE0-D4 but worsens the performance of PBE0-VV10 for which the errors in charge transfer and dispersion energy have the same sign (Figure 4i).

Figure 4j-l shows that the error compensation is less pronounced in the case of the CHB6

23

dataset. This is primarily due to smaller errors in the dispersion energy displayed by all dispersion-corrected DFs, which can be traced back to the intrinsic electronic properties of cations. Specifically, due to their compact electron density, cations exhibit smaller polarizabilities than neutral and anionic species. A smaller polarizability implies that the electron density of a cation is less easily distorted by an external electric field. Since dispersion interactions depend on these distortions of the electron densities, they are relatively weaker for cationic species. This explains the smaller errors in the dispersion energies exhibited by the different dispersion-corrected DFs for systems in the CHB6 dataset. On the other hand, electrostatic interactions, which arise from the forces that electric charges exert on each other, play a key role for cationic systems. Since charges get artificially distributed over larger volumes due to the delocalization error, Figure 4j-l shows that the errors associated with electrostatic contributions predicted by the different dispersion-corrected DFs are consistently larger for systems in the CHB6 dataset.

Energy Decomposition Analysis of Aqueous Clusters

To elucidate the effects of various dispersion models on the DFT description of molecular interactions in aqueous environments, we focus in this section on water $[(H_2O)_{n=2-10}]$, hydrated sodium cation $[Na^+(H_2O)_{n=2-10}]$, and hydrated chloride anion $[Cl^-(H_2O)_{n=2-10}]$ clusters. As in the previous sections, our analyses encompass the -D3, -D4, and -VV10 models combined with the PBE, revPBE, and PBE0 functionals that are commonly employed in AIMD simulations of aqueous systems. Unless otherwise stated, the analyses with the -D3 model presented in this section are discussed in the context of the -D3(0) damping scheme. The structures of the $(H_2O)_{n=2-10}$ clusters were obtained from the BEGDB water clusters dataset,[127] which were originally optimized at the RI-MP2/aug-cc-pVDZ level of theory.[81] All analyses reported for the BEGDB clusters are averaged over the number of isomers in the dataset for each cluster size n.

The combination of the -D3, -D4, and -VV10 models with the PBE and PBE0 function-

Table 1: Errors (kcal/mol) in interaction energies per water molecule relative to the CCSD(T)/CBS[128] and DLPNO-CCSD(T)/def2-QZVPPD reference values for water and ion-water clusters, respectively, calculated with PBE, revPBE, and PBE0 combined with the -D3, -D4, and -VV10 models as well as ωB97M-V.

Method	$(H_2O)_2$	$(H_2O)_{10}$	$Na^+(H_2O)_2$	$Na^+(H_2O)_{10}$	$Cl^-(H_2O)_2$	$Cl^-(H_2O)_{10}$
PBE-D3	-0.37	-1.52	0.82	0.12	-1.18	-0.76
PBE-D4	-0.34	-1.33	0.65	0.23	-1.33	-0.63
PBE-VV10	-0.46	-1.87	0.60	-0.04	-1.84	-1.06
revPBE-D3	0.05	0.19	1.70	1.10	0.51	0.30
revPBE-D4	0.07	0.18	1.21	1.10	0.08	0.41
revPBE-VV10	0.30	1.05	1.44	0.91	-0.75	-0.21
PBE0-D3	-0.34	-1.40	0.65	0.05	-1.01	-0.67
PBE0-D4	-0.27	-1.12	0.58	0.28	-1.06	-0.46
PBE0-VV10	-0.39	-1.62	0.51	0.03	-1.55	-0.85
ωB97M-V	-0.13	-0.50	0.79	0.50	-0.13	-0.18

als results in significant errors in the interaction energies with respect to CCSD(T)/CBS reference values,[128] which tend to increase with cluster's size as summarized in Table 1 for $(H_2O)_2$, $Na^+(H_2O)_n$ and $Cl^-(H_2O)_n$, with $n = 2, 10$. The errors per water molecule for the interaction energies of the full datasets are reported in Figures S5-S13.

The ALMO-EDA errors relative to the ωB97M-V reference values for each individual energy component of the corresponding interaction energy for the $(H_2O)_{n=2-10}$ are shown in Figure 5a as mean errors (ME) per water molecule, $\Delta e_{\text{component}} = (E_{\text{component}}^{\text{model}} - E_{\text{component}}^{\text{ref}})/n$. The error per water molecule in the frozen component (i.e., electrostatics + Pauli repulsion) of the interaction energies exhibits an increasing trend with cluster size. This trend is more pronounced for PBE and revPBE than the hybrid PBE0 functional. Overall, all three DFs perform similarly in describing the polarization contribution, with the associated errors contributing minimally to the total error for all clusters.

As discussed in the previous sections, GGA functionals, such as PBE and revPBE, suffer from a non-negligble delocalization error, which, in turn, results in a systematic overestimation of charge transfer. This is particularly pronounced in aqueous systems, as observed

in our previous studies.[129–132] In the case of PBE-D3, the error in charge transfer energies increases from -0.35 kcal/mol to -1.65 kcal/mol as the cluster size increases from the dimer to the decamer. Similar trend is displayed by revPBE-D3. As expected, by incorporating 25% Hartree-Fock exchange, this artificial charge transfer error is partially mitigated in PBE0-D3, which exhibits errors of -0.12 kcal/mol and -0.57 kcal/mol for the dimer and decamer, respectively.

For the dispersion energy, revPBE-D3 displays the most notable deviation from the ωB97M-V reference values, with an error of 0.36 kcal/mol for the dimer that progressively increases to 1.60 kcal/mol for the decamer. In contrast, PBE-D3 and PBE0-D3 exhibit minimal errors for all clusters from the dimer to the decamer. While PBE-D3 and PBE0-D3 show improved agreement with ωB97M-V for the dispersion energies, their overall performance in reproducing the CCSD(T)/CBS interaction energies is inferior to revPBE-D3. This poor performance is rationalized by considering the large dispersion errors associated with revPBE-D3. Paradoxically, the smaller errors in dispersion energies exhibited by PBE-D3 and PBE0-D3 are insufficient to compensate the charge transfer errors, which thus results in higher inaccuracy in the description of the interaction energies as reported in Table 1. Figure S14 and S15 show the same analyses for PBE, revPBE, and PBE0 combined with the -D4 and -VV10 models, respectively. Interestingly, although the trends in dispersion energies predicted by PBE and PBE0 combined with the -D4 and -VV10 models are similar to those provided by the -D3 model, the performance of revPBE-VV10 is significantly worse than that of revPBE-D3 and revPBE-D4.

Given the apparent agreement with the CCSD(T)/CBS reference energies for the water clusters, which, as demonstrated above, effectively derives from fortuitous error compensation between the representation of charge transfer and dispersion energy, revPBE-D3(0) has gained widespread popularity as a suitable DF for AIMD simulations of liquid water.[111–113] This popularity has made revPBE-D3(0) one of the preferred DFs for training "DFT-level" machine-learned potentials for water.[115] In this context, it should be noted

26

that a recent study[131] demonstrated that the -D3(BJ) model and its generalized form -D3(op)[53] introduces smaller dispersion errors compared to the -D3(0) model in the case of revPBE. However, due a less-than-optimal error compensation between charge transfer
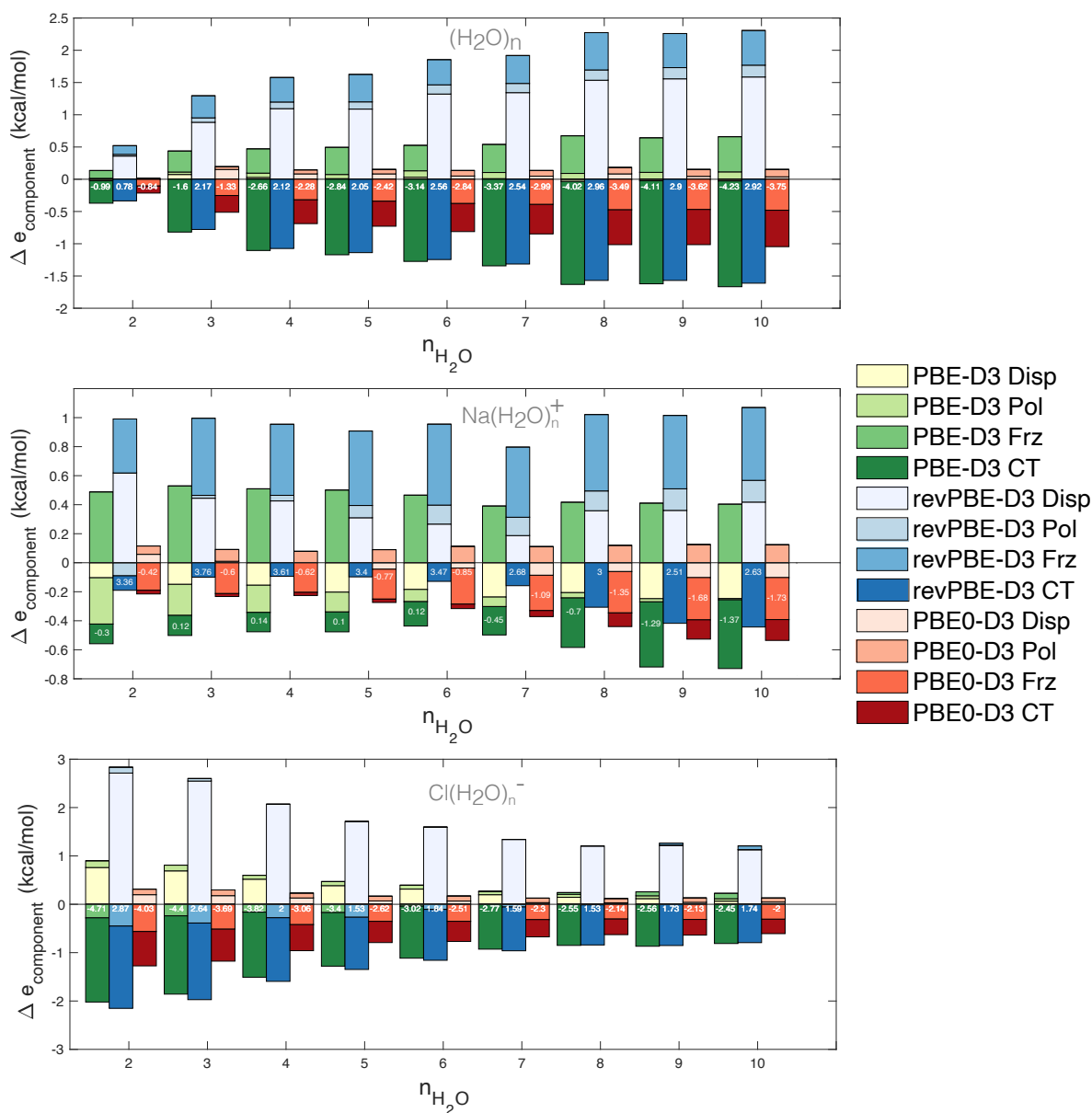

Figure 5: Errors (kcal/mol) relative to the ωB97M-V reference values for the energy components of the interaction energies obtained from ALMO-EDA calculations carried out for $(H_2O)_n$ (a), $Na^+(H_2O)_n$ (b), and $Cl^-(H_2O)_n$ (c) clusters, with $n = 2 - 10$. All errors are divided by the number of water molecules in each cluster and the total error is shown on each bar. The energy components of PBE-D3, revPBE-D3, and PBE0-D3 are shown in green, blue, and red shades, respectively. See main text for details.

and dispersion energy, both revPBE-D3(BJ) and revPBE-D3(op) provide worse agreement with the CCSD(T)/CBS reference energies than revPBE-D3(0). This contradicting trends in reproducing dispersion and interaction energies manifest, for example, in revPBE-D3(0) better reproducing the sum-frequency generation spectrum of the air/water interface than revPBE-D3(BJ) and PBE-D3.[133] The underbinding dispersion interaction in revPBE-D3(0) effectively prevents hydrogen bonds from becoming overly attractive, which counteracts the relative large delocalization error. On the other hand, a better description of the dispersion energy combined with GGA functionals, such as PBE and revPBE, lead to overly stronger hydrogen bonds, causing the vibrational spectra to redshift as shown in simulations with PBE-D3(0), PBE-D3(BJ), and revPBE-D3(BJ).[133]

Simulations carried out based on revPBE-D3 were shown to reproduce some properties of liquid water.[134,135] However, employing a data-driven many-body potential based on revPBE-D3, which includes a pairwise dispersion term resulted in a more structured liquid. This seemingly counterintuitive result is again explained by the lack of error compensation in dispersion-corrected revPBE variants that adopt an accurate description of the dispersion energy.[131] Head-Gordon and coworkers also suggested the presence of error-compensation in revPBE-D3.[136] In particular, they demonstrated that the inherent inaccuracies in revPBE-D3 seem to fortuitously coincide with the neglect of nuclear quantum effect in classical simulations, as the agreement with the experimental data deteriorates when these effects are explicitly included in path-integral molecular dynamics simulations.

In addition to pure water clusters, we also investigated the impact of dispersion interactions on ionic aqueous clusters. In the following, we focus on $Na^+(H_2O)_n$ and $Cl^-(H_2O)_n$ clusters, obtained from molecular dynamics simulations carried out in the isothermal-isobaric (NPT = constant number of particles, pressure, and temperature) ensemble at $T = 298.15$ K and $P = 1$ atm.[137,138] These simulations employed MB-nrg many-body potentials that were trained on CCSD(T)/CBS 2-body and 3-body energies for ion-water dimers and trimers[137–140] The cluster configurations extracted from the NPT simulations were further optimized using

28

the MB-nrg framework implemented in the MBX software.[141]

Contrary to the water clusters, dispersion-corrected revPBE exhibits significantly larger error in interaction energies compared to the DLPNO-CCSD(T)/def2-QZVPPD reference values, especially when compared to the results obtained with the dispersion-corrected PBE and PBE0 functionals as shown in Table 1 and Figure S6. The analysis of the dispersion energies predicted by PBE-D3, PBE0-D3, and revPBE-D3 demonstrates that, while both PBE-D3 and PBE0-D3 exhibit relatively small errors, revPBE-D3 consistently exhibits larger errors as shown in Figure 5b. For example, the errors associated with the PBE-D3 values for the dispersion energy range from -0.15 kcal/mol for $n = 2$ to -0.27 kcal/mol for $n = 10$. Similar trend is observed for PBE0-D3 with errors of 0.01 kcal/mol and -0.11 kcal/mol for the corresponding clusters. In contrast, the errors for revPBE-D3 are 0.43 kcal/mol for $Na^+(H_2O)_2$ and 0.47 kcal/mol for $Na^+(H_2O)_{10}$. The corresponding ALMO-EDA results obtained with the -D4 and -VV10 models are reported in Figures S16 and S17. For the -D4 model, the errors associated with the dispersion energy decrease as a function of cluster size for PBE-D4 but increase for revPBE-D4. In the case of PBE0-D3, the errors remain relatively small and only slightly increase as the cluster size increases. It is noteworthy that all D4-corrected DFs exhibit errors in dispersion energies comparable in magnitude to those of their D3-corrected counterparts. Both PBE-VV10 and revPBE-VV10 exhibit substantial errors in dispersion energies, which persist independently of cluster size. Conversely, PBE0-VV10 displays a consistent trend, with comparatively smaller errors in dispersion energies for all clusters.

It is noteworthy that the dispersion errors in $Na^+(H_2O)_n$ clusters are smaller compared to those observed in $(H_2O)_n$ clusters, primarily because the dispersion contributions to the interaction energies is much smaller in $Na^+(H_2O)_n$ clusters as cations have smaller polarizabilities. As mentioned above, $Na^+(H_2O)_n$ clusters are governed by electrostatic interactions that are poorly described by GGA functionals, such as PBE and revPBE, due to relatively large delocalization errors. This is evident in Figure 5b where both PBE and revPBE exhibit

29

large errors in the description of the electrostatic energy.

Mundy and coworkers carried out AIMD simulations of alkali-metal ions in water and compared the performance of revPBE-D3, a GGA functional, and SCAN, a meta-GGA functional.[142] Their results indicate that SCAN provides a more reliable description of the hydration structure for Na^+, although revPBE-D3 outperforms SCAN in describing the hydration structure of liquid water. The observed differences between the behavior of SCAN and revPBE-D3 in the context of liquid water were attributed to SCAN's known tendency to overbind.[142] In this context, it is important to note that both revPBE-D3 and SCAN are susceptible to delocalization errors, which are more prominent for GGA functionals, such as revPBE-D3, in case of ion-water clusters.[131,132] As discussed in ref 131 for revPBE-D3(0), revPBE-D3(op), and SCAN, delocalization errors lead to an overstructured representation of liquid water. In the case of neutral water, the errors in charge transfer and dispersion energy in revPBE-D3 accidentally compensate for each other, leading to a fortuitous improvement in the performance of revPBE-D3 for liquid water. In contrast, SCAN outperforms revPBE-D3 in predicting the energetics of $Na^+(H_2O)_n$ clusters as well as the hydration structure of Na^+ in liquid water.[142] The reason for the different performance of revPBE-D3 in describing pure water and hydrated Na^+ thus lies in the different nature of the underlying physical interactions. Specifically, permanent electrostatics become the dominant contribution in Na^+-water clusters, which prevents the fortuitous error compensation observed in water clusters and consequently prevents the -D3 model, which shows remarkable accuracy for neutral water simulations, from accurately representing Na^+-water interactions. Recently, Jungwirth and coworkers reported significant improvement for hydrated alkali-metal ion simulations by selectively deactivating the dispersion interaction between alkali-metal ions and water molecules while still preserving the dispersion interaction among water molecules.[143] It should be emphasized that this improved agreement does not imply that alkali-metal ions are not polarizable or that their dispersion interactions can be neglected. Our analyses, in fact, demonstrate that removing Na^+-metal water dispersion interactions can, at times,

30

result in an improved performance of revPBE-D3. However, this improvement is primarily due to error compensation rather than a more accurate representation of the actual physical interactions.

Unlike the $Na^+(H_2O)_n$ clusters, Table 1 and Figure S7 show that the errors in the interaction energies of the $Cl^-(H_2O)_n$ clusters calculated with all dispersion-corrected DFs decrease as the cluster size increases. It is, however, important to note that these errors are significantly larger for $Cl^-(H_2O)_n$ than for $Na^+(H_2O)_n$ and $(H_2O)_n$ clusters. This behavior is reminiscent of the delocalization error that affects any DFT representation of molecular interactions and is particularly pronounced in the case of anionic systems due to the more diffuse delocalization of the excess electrons. As discussed above, the delocalization error, in turn, has a significant effect on the charge transfer energies.[132,144] Interestingly, the errors in interaction energies for the larger clusters remain quite comparable, regardless of the dispersion model used (Figures S10 and S13). As in the case of water clusters, revPBE combined with any of the dispersion models provides better agreement with the reference values than the dispersion-corrected PBE and PBE0 functionals.

Figure 5c shows the ALMO-EDA results for the $Cl^-(H_2O)_{n=1-10}$ clusters. Focusing on the dispersion energy, the error associated with revPBE-D3 still constitutes a substantial fraction of the total error, being 2.56 kcal/mol at $n = 2$ and 1.18 kcal/mol at $n = 10$. The error in the dispersion energy varies from 0.70 kcal/mol to 0.07 kcal/mol in the case of PBE-D3 and from 0.18 kcal/mol to 0.04 kcal/mol in the case of PBE0-D3 as the cluster size increases from $n = 2$ to $n = 10$.

As shown in Figures S18 and S19, the errors in dispersion energies for the $Cl^-(H_2O)_2$ cluster calculated with the -D4 and -VV10 models present a significant spread in values. Specifically, in the case of the PBE functional, the error varies from 0.55 kcal/mol for PBE-D4 to -0.04 kcal/mol for PBE-VV10. Similarly, the error associated with revPBE-D4 is 2.14 kcal/mol and becomes 1.13 kcal/mol for revPBE-VV10, while PBE0-D4 and PBE0-VV10 exhibit errors of 0.13 kcal/mol and -0.43 kcal/mol, respectively. The same trends hold for

31

the larger $Cl^-(H_2O)_{10}$ cluster. In this case, the error varies from 0.20 kcal/mol for PBE-D4 to -0.25 kcal/mol for PBE-VV10, from 1.28 kcal/mol for revPBE-D4 to 0.60 kcal/mol for revPBE-VV10, and from 0.25 kcal/mol for PBE0-D4 to -0.16 kcal/mol for PBE0-VV10.

These analyses highlight the significant impact of the selected dispersion models on the representation of dispersion interactions in water and ionic aqueous systems within DFT. Notably, they shed light on the unexpected accuracy of revPBE-D3 in AIMD simulations of liquid water and hydrated monovalent anions (e.g., Cl^-),[145] attributing it to error compensation among various physical contributions, especially charge transfer and dispersion energy. Moreover, these findings underscore the need for caution in AIMD simulations of aqueous systems with dispersion-corrected DFs since, not being able to correctly represent the underlying physical interactions, these DFs may lead to misinterpretations of the experimental data.

Impact of Dispersion Energy on Molecular Crystals

DFT stands as the *de facto* electronic structure theory for modeling extended systems due to its reasonable accuracy and efficiency relative to post-Hartree-Fock wavefunction methods. To assess how a given dispersion model affects the DFT description of structural properties and energetics of molecular crystals, we analyzed the performance of dispersion-corrected DFs on the X23 dataset that consists of a diverse collection of 23 molecular crystals with well-understood rigid molecular structures (Figure S20).[94] The reference values in the X23 dataset[94] includes experimentally measured sublimation enthalpies, adjusted to account for vibrational contributions. This adjustment enables a direct comparison of static lattice energies. Since the interactions in the molecular crystals included in the X23 dataset are dominated by long-range effects, the treatment of the dispersion energy within DFT is expected to significantly affect the lattice structures.

Full unit cell optimizations were carried out for all 23 molecular crystals using PBE-D3, PBE-MBD, revPBE-D3, SCAN, and SCAN-rVV10. The analysis of the errors in the unit cell

32

volumes, represented as percentages of the corresponding reference experimental volumes,[146] provides valuable insights into the accuracy of dispersion-corrected DFs in predicting the structural properties of the molecular crystals in the X23 dataset. Evaluating the entire dataset, the mean absolute errors (MAEs) for the PBE-D3 and PBE-MBD functionals are 1.83% and 1.77%, respectively. Poorer performance is displayed by revPBE-D3 and revPBE-MBD, with MAEs of 3.41% and 6.99%, respectively.

SCAN represents a particular case among the DFs examined in this study. By construction, SCAN includes medium-range dispersion contributions in its functional form,[67] but, as any semi-local functional, also suffers from a nonnegligble delocalization error that leads to over-attractive interactions.[147] The effect of the delocalization error is particularly evident in the case of the hydrogen-bonded molecular crystals of the X23 dataset for which SCAN predicts more compact unit cells as shown in Figure S20. On the other hand, the absence of long-range dispersion contributions partially counteracts the artificial unit cell shrinkage due to the delocalization error in dispersion-bound molecular crystals. The MAE of SCAN over the entire X23 dataset is 3.47%. Combining the -rVV10 model with SCAN[102] leads to a significant reduction in unit cell size, resulting in an error of 11.13%. The unit cell shrinkage provided by SCAN-rVV10 is more pronounced in the hydrogen-bonded molecular crystals as shown in Figure S20. These results are in line with the observation that AIMD simulations of liquid water reported in the literature are mostly carried out with SCAN because the incorporation of the dispersion correction through the -rVV10 model leads to a significant overestimation of the liquid density.[148]

Figure 6a shows the normal distributions of the volume errors associated with each dispersion-corrected DF examined in this study. Both PBE-D3 and PBE-MBD distributions display small mean errors and standard deviations. Although revPBE-D3 and revPBE-MBD, on average, predict more contracted and expanded unit cells, respectively, the corresponding distributions are relatively wide, spanning both negative (i.e., smaller unit cell volumes) and positive errors (i.e., larger unit cell volumes). As expected from the discussion above, SCAN
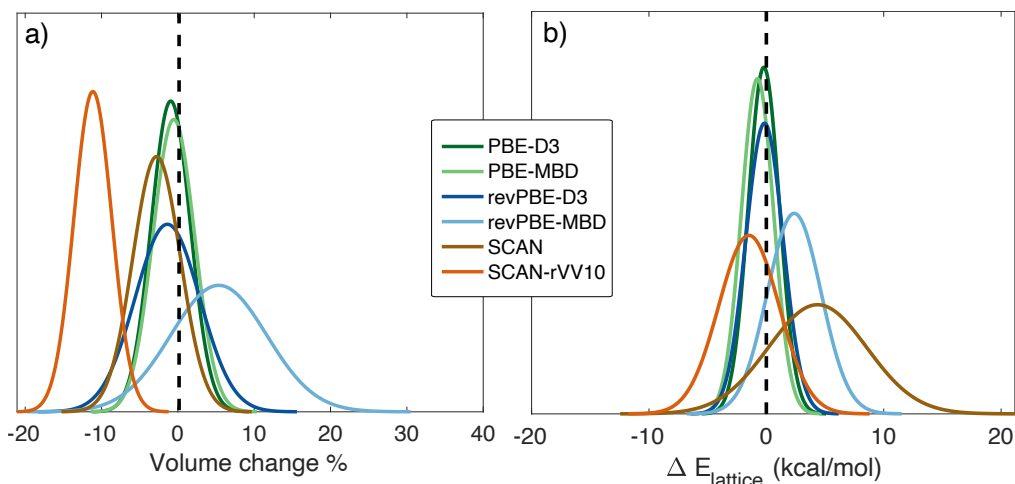
33

Figure 6: Normal distribution of errors relative to the experimental values[94,146] for cell volumes (a) and lattice energies (b) of the molecular crystals in the X23 dataset calculated with PBE-D3, PBE-MBD, revPBE-D3, revPBE-MBD, SCAN, and SCAN-rVV10. See main text for details.

and SCAN-rVV10, on average, predict more contracted unit cells. However, while the error distribution of SCAN also extends into positive errors, SCAN-rVV10 underestimates the unit cell volumes for every molecular crystal in the X23 dataset.

The normal distributions of the lattice energy errors reported in Figure 6b demonstrate that both PBE-D3 and PBE-MBD exhibit a similar level of agreement with the experimental values, resulting in MUEs of 1.02 kcal/mol and 1.36 kcal/mol, respectively. Interestingly, revPBE-D3 predicts X23 lattice energies with a comparable MUE of 1.26 kcal/mol. However, the combination of revPBE with the -MBD model compromises the overall accuracy, leading to a MUE of 2.66 kcal/mol. Although predicting significantly more contracted unit cells, SCAN-rVV10 performs reasonably well in reproducing the lattice energies. Specifically, SCAN yields an MUE of 4.60 kcal/mol, which reduces to 2.26 kcal/mol for SCAN-rVV10. The performance of SCAN-rVV10 in predicting unit cell volumes and lattice energies is particularly revealing. It underscores the capability of a dispersion-corrected DF to reliably estimate the energetics of molecular crystals, even amidst clear limitations in accurately reproducing structural properties.

34

CONCLUSIONS

Maintaining a balance between physical interpretability and energy predictability is paramount when employing dispersion models within DFT. Although the incorporation of empirical dispersion can sometimes enhance the accuracy of DFT calculations, it concurrently brings the potential for introducing physical inconsistencies, thereby adding misleading complexity to the interpretation of the results.

A key challenge arises when tuning parameters to predict the interaction energy of a given molecular system rather than only the component associated with the dispersion energy. This study demonstrates that parameter fitting can compromise the underlying physics and hinder transferability, as it leads to significant error compensation among the different energy components in order to reproduce the reference data. Consequently, the performance of dispersion models varies across different systems. For example, while revPBE-D3 proves highly effective in hydrogen-bonded systems, it struggles to faithfully reproduce the potential energy surface of the benzene dimer. This discrepancy highlights the potential for inconsistent - and thus unpredictable - behavior of various dispersion-corrected DFs across a broad range of molecular systems, a trait reminiscent of the behavior displayed by force fields.

Further complicating matters is a proposed solution to optimally tune the dispersion correction in accordance to a tailored training set. While Perdew *et. al* recommended distinct sets of dispersion coefficients for different systems to optimize performance for a certain system class,[72] this method increases the complexity of the resulting optimally-tuned dispersion-corrected DF. Additionally, the variability of dispersion coefficients in optimally-tuned dispersion models for molecular systems and layered materials may, for example, limit the applicability of the optimally-tuned dispersion-corrected DF to surface science, thus complicating the description of processes like molecular adsorption.

In pursuit of obtaining physically-meaningful results rooted in accurate physics, future development of dispersion models should prioritize reproducing the dispersion energy, rather than solely aiming to reproduce the interaction energy through parameter fitting. This

refined approach offers the potential to more accurately capture the underlying physics, ensuring broad applicability across a diverse range of molecular systems.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

(1) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, CA, 2004.

(2) Schalley, C. A. *Analytical Methods in Supramolecular Chemistry*, 2nd ed.; Wiley, 2012.

(3) Cockroft, S. L.; Hunter, C. A. Chemical Double-Mutant Cycles: Dissecting Non-Covalent Interactions. *Chem. Soc. Rev.* **2007**, *36*, 172–188.

(4) Brown, T. L.; Bursten, B. E.; Eugene, H.; LeMay, H. *Chemistry: The Central Science*, 11th ed.; Pearson Prentice Hall: Upper Saddle River, NJ, 2009.

(5) Eisler, M. In *Encyclopedia of Nanoscience and Society*; Guston, D. H., Ed.; SAGE Publications: Thousand Oaks, CA, 2010.

(6) Biedermann, F.; Schneider, H.-J. Experimental Binding Energies in Supramolecular Complexes. *Chem. Rev.* **2016**, *116*, 5216–5300.

(7) Rezac, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.

(8) Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An Improved Linear Scaling Perturbative Triples Correction for the Domain Based Local Pair-Natural Orbital Based Singles and Doubles Coupled Cluster Method [DLPNO-CCSD (T)]. *J. Chem. Phys.* **2018**, *148*.

(9) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618.

(10) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 104–112.

(11) Grimme, S. Do Special Noncovalent π–π Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008**, *47*, 3430–3434.

(12) Tkatchenko, A.; DiStasio, R. A.; Head-Gordon, M.; Scheffler, M. Dispersion-Corrected Møller-Plesset Second-Order Perturbation Theory. *J. Chem. Phys.* **2009**, *131*.

(13) Hohenstein, E. G.; Sherrill, C. D. Wavefunction Methods for Noncovalent Interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.

(14) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(15) Ullrich, C. A. Time-Dependent Density-Functional Theory: Concepts and Applications. **2011**,

(16) Pérez-Jordá, J.; Becke, A. D. A Density-Functional Study of Van Der Waals Forces: Rare Gas Diatomics. *Chem. Phys. Lett.* **1995**, *233*, 134–137.

(17) Kristyán, S.; Pulay, P. Can (Semi) Local Density Functional Theory Account for the London Dispersion Forces? *Chem. Phys. Lett.* **1994**, *229*, 175–180.

(18) Langreth, D. C.; Perdew, J. P. The Exchange-Correlation Energy of a Metallic Surface. *Solid State Communications* **1975**, *17*, 1425–1429.

(19) Langreth, D. C.; Perdew, J. P. Exchange-Correlation Energy of a Metallic Surface: Wave-Vector Analysis. *Phys. Rev. B* **1977**, *15*, 2884.

(20) Adamo, C.; Barone, V. Toward Reliable Adiabatic Connection Models Free From Adjustable Parameters. *Chem. Phys. Lett.* **1997**, *274*, 242–250.

(21) Kaltak, M.; Klimeš, J.; Kresse, G. Cubic Scaling Algorithm for the Random Phase Approximation: Self-Interstitials and Vacancies in Si. *Phys. Rev. B* **2014**, *90*, 054115.

(22) Von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory. *Phys. Rev. Lett.* **2004**, *93*, 153004.

(23) Wu, Q.; Yang, W. Empirical Correction to Density Functional Theory for Van Der Waals Interactions. *J. Chem. Phys.* **2002**, *116*, 515–524.

(24) Grimme, S. Accurate Description of Van Der Waals Complexes by Density Functional Theory Including Empirical Corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

38

(25) Grimme, S. Semiempirical Gga-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(26) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(27) Becke, A. D.; Johnson, E. R. A Density-Functional Model of the Dispersion Interaction. *J. Chem. Phys.* **2005**, *123*, 154101.

(28) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction: High-Order Dispersion Coefficients. *J. Chem. Phys.* **2006**, *124*, 014104.

(29) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction. *J. Chem. Phys.* **2005**, *122*, 154104.

(30) Ángyán, J. G. On the Exchange-Hole Model of London Dispersion Forces. *J. Chem. Phys.* **2007**, *127*, 024108.

(31) Heßelmann, A. Derivation of the Dispersion Energy as an Explicit Density-and Exchange-Hole Functional. *J. Chem. Phys.* **2009**, *130*, 084104.

(32) Steinmann, S. N.; Corminboeuf, C. A System-Dependent Density-Based Dispersion Correction. *J. Chem. Theory Comput.* **2010**, *6*, 1990–2001.

(33) Vydrov, O. A.; Van Voorhis, T. Nonlocal Van der Waals Density Functional Made Simple. *Phys. Rev. Lett.* **2009**, *103*, 063004.

(34) Vydrov, O. A.; Van Voorhis, T. Implementation and Assessment of a Simple Nonlocal Van Der Waals Density Functional. *J. Chem. Phys.* **2010**, *132*, 164113.

(35) Vydrov, O. A.; Van Voorhis, T. Nonlocal Van der Waals Density Functional: The Simpler the Better. *J. Chem. Phys.* **2010**, *133*, 244103.

(36) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.

(37) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.

(38) Andersson, Y.; Hult, E.; Rydberg, H.; Apell, P.; Lundqvist, B. I.; Langreth, D. C. Van der Waals Interactions in Density Functional Theory. *Electronic Density Functional Theory: Recent Progress and New Directions* **1998**, 243–260.

(39) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-Accuracy Van der Waals Density Functional. *Phys. Rev. B* **2010**, *82*, 081101.

(40) Berland, K.; Borck, Ø.; Hyldgaard, P. Van der Waals Density Functional Calculations of Binding in Molecular Crystals. *Comput. Phys. Commun.* **2011**, *182*, 1800–1804.

(41) Vydrov, O. A.; Van Voorhis, T. Dispersion Interactions From a Local Polarizability Model. *Phys. Rev. A* **2010**, *81*, 062708.

(42) Casimir, H. B.; Polder, D. The Influence of Retardation on the London-Van der Waals Forces. *Phys. Rev.* **1948**, *73*, 360.

(43) Bade, W. Drude-Model Calculation of Dispersion Forces. I. General Theory. *J. Chem. Phys.* **1957**, *27*, 1280–1284.

(44) Jones, A. P.; Crain, J.; Sokhan, V. P.; Whitfield, T. W.; Martyna, G. J. Quantum Drude Oscillator Model of Atoms and Molecules: Many-Body Polarization and Dispersion Interactions for Atomistic Simulation. *Phys. Rev. B* **2013**, *87*, 144103.

(45) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. Van der Waals Density Functional for General Geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(46) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D.; Lundqvist, B. Erratum: Van der Waals Density Functional for General Geometries [Phys. Rev. Lett. 92 246401 (2004)]. *Phys. Rev. Lett.* **2005**, *95*, 109902.

(47) Gianturco, F.; Paesani, F. In *Conceptual Perspectives in Quantum Chemistry*; Calais, J.-L., Kryachko, E., Eds.; Springer, 1997; pp 337–382.

(48) Gianturco, F.; Paesani, F.; Laranjeira, M.; Vassilenko, V.; Cunha, M. Intermolecular Forces From Density Functional Theory. III. A Multiproperty Analysis for the Ar (1 S)-Co (1 σ) Interaction. *J. Chem. Phys.* **1999**, *110*, 7832–7845.

(49) Gianturco, F.; Paesani, F. The He–OCS Van Der Waals Potential From Model Calculations: Bound States Stable Structures and Vibrational Couplings. *J. Chem. Phys.* **2000**, *113*, 3011–3019.

(50) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. Hydrogen Bonding and Stacking Interactions of Nucleic Acid Base Pairs: A Density-Functional-Theory Based Treatment. *J. Chem. Phys.* **2001**, *114*, 5149–5155.

(51) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(52) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

(53) Witte, J.; Mardirossian, N.; Neaton, J. B.; Head-Gordon, M. Assessing DFT-D3 Damping Functions Across Widely Used Density Functionals: Can We Do Better? *J. Chem. Theory Comput.* **2017**, *13*, 2043–2052.

(54) Tkatchenko, A.; DiStasio Jr, R. A.; Car, R.; Scheffler, M. Accurate and Efficient

Method for Many-Body Van der Waals Interactions. *Phys. Rev. Lett.* **2012**, *108*, 236402.

(55) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van der Waals Interactions From Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102*, 073005.

(56) Anatole von Lilienfeld, O.; Tkatchenko, A. Two-and Three-Body Interatomic Dispersion Energy Contributions to Binding in Molecules and Solids. *J. Chem. Phys.* **2010**, *132*.

(57) Donchev, A. Many-Body Effects of Dispersion Interaction. *J. Chem. Phys.* **2006**, *125*.

(58) Tkatchenko, A.; Ambrosetti, A.; DiStasio, R. A. Interatomic Methods for the Dispersion Energy Derived From the Adiabatic Connection Fluctuation-Dissipation Theorem. *J. Chem. Phys.* **2013**, *138*.

(59) Johnson, E. R.; Becke, A. D. A Post-Hartree–Fock Model of Intermolecular Interactions. *J. Chem. Phys.* **2005**, *123*.

(60) Johnson, E. R.; Becke, A. D. A Post-Hartree-Fock Model of Intermolecular Interactions: Inclusion of Higher-Order Corrections. *J. Chem. Phys.* **2006**, *124*, 174104.

(61) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry Thermochemical Kinetics Noncovalent Interactions Excited States and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(62) Tkatchenko, A.; von Lilienfeld, O. A. Adsorption of Ar on Graphite Using London Dispersion Forces Corrected Kohn-Sham Density Functional Theory. *Phys. Rev. B* **2006**, *73*, 153406.

(63) Arey, J. S.; Aeberhard, P. C.; Lin, I.-C.; Rothlisberger, U. Hydrogen Bonding Described Using Dispersion-Corrected Density Functional Theory. *J. Phys. Chem. B* **2009**, *113*, 4726–4732.

(64) Mackie, I. D.; DiLabio, G. A. Accurate Dispersion Interactions from Standard Density-Functional Theory Methods with Small Basis Sets. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6092–6098.

(65) van Santen, J. A.; DiLabio, G. A. Dispersion Corrections Improve the Accuracy of Both Noncovalent and Covalent Interactions Energies Predicted by a Density-Functional Theory Approximation. *J. Phys. Chem. A* **2015**, *119*, 6703–6713.

(66) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham Global-Hybrid Exchange–Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7*, 5032–5051.

(67) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

(68) Jurečka, P.; Šponer, J.; Černỳ, J.; Hobza, P. Benchmark Database of Accurate (MP2 and CCSD(T) Complete Basis Set Limit) Interaction Energies of Small Model Complexes DNA Base Pairs and Amino Acid Pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(69) Gráfová, L.; Pitonak, M.; Rezac, J.; Hobza, P. Comparative Study of Selected Wave Function and Density Functional Methods for Noncovalent Interaction Energy Calculations Using the Extended S22 Data Set. *J. Chem. Theory Comput.* **2010**, *6*, 2365–2376.

(70) Rezác, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark

Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.

(71) Rezáč, J.; Riley, K. E.; Hobza, P. Extensions of the S66 Data Set: More Accurate Interaction Energies and Angular-Displaced Nonequilibrium Geometries. *J. Chem. Theory Comput.* **2011**, *7*, 3466–3470.

(72) Peng, H.; Perdew, J. P. Rehabilitation of the Perdew-Burke-Ernzerhof Generalized Gradient Approximation for Layered Materials. *Phys. Rev. B* **2017**, *95*, 081105.

(73) Ferretti, A.; Canal, L.; Sorodoc, R. A.; Sinha, S.; Brancato, G. Fine Tuning the Intermolecular Interactions of Water Clusters Using the Dispersion-Corrected Density Functional Theory. *Molecules* **2023**, *28*, 3834.

(74) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergifosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, Á.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.;

Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Lieben-thal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, Á.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sund-strom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Co-riani, S.; Cramer, C. J.; Cserey, G.; DePrince, A. E.; DiStasio, R. A.; Dreuw, A.; Duni-etz, B. D.; Furlani, T. R.; Goddard, W. A.; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lam-brecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsen-feld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truh-lar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, H. L.; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I. Software for the Frontiers of Quantum Chemistry: An Overview of De-velopments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155*, 084801.

(75) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(76) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Sim-

ple". *Phys. Rev. Lett.* **1998**, *80*, 890.

(77) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(78) Rappoport, D.; Furche, F. Property-Optimized Gaussian Basis Sets for Molecular Response Calculations. *J. Chem. Phys.* **2010**, *133*.

(79) Dasgupta, S.; Herbert, J. M. Standard Grids for High-Precision Integration of Modern Density Functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017**, *38*, 869–882.

(80) Arago, J.; Orti, E.; Sancho-Garcia, J. C. Nonlocal Van der Waals Approach Merged with Double-Hybrid Density Functionals: Toward the Accurate Treatment of Noncovalent Interactions. *J. Chem. Theory Comput.* **2013**, *9*, 3437–3443.

(81) Temelso, B.; Archer, K. A.; Shields, G. C. Benchmark Structures and Binding Energies of Small Water Clusters with Anharmonicity Corrections. *J. Phys. Chem. A* **2011**, *115*, 12034–12046.

(82) Riplinger, C.; Neese, F. An Efficient and Near Linear Scaling Pair Natural Orbital Based Local Coupled Cluster Method. *J. Chem. Phys.* **2013**, *138*.

(83) Hellweg, A.; Rappoport, D. Development of New Auxiliary Basis Functions of the Karlsruhe Segmented Contracted Basis Sets Including Diffuse Basis Functions (Def2-SVPD Def2-TZVPPD and Def2-Qvppd) for Ri-Mp2 and Ri-Cc Calculations. *Phys. Chem. Chem. Phys.* **2015**, *17*, 1010–1017.

(84) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152*.

(85) Horn, P. R.; Head-Gordon, M. Polarization Contributions to Intermolecular Interactions Revisited with Fragment Electric-Field Response Functions. *J. Chem. Phys.* **2015**, *143*, 114111.

(86) Horn, P. R.; Mao, Y.; Head-Gordon, M. Defining the Contributions of Permanent Electrostatics Pauli Repulsion and Dispersion in Density Functional Theory Calculations of Intermolecular Interaction Energies. *J. Chem. Phys.* **2016**, *144*, 114107.

(87) Horn, P. R.; Mao, Y.; Head-Gordon, M. Probing Non-Covalent Interactions with a Second Generation Energy Decomposition Analysis Using Absolutely Localized Molecular Orbitals. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079.

(88) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New Basis Set Exchange: An Open, Up-to-date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.

(89) Weigend, F.; Furche, F.; Ahlrichs, R. Gaussian Basis Sets of Quadruple Zeta Valence Quality for Atoms H–Kr. *J. Chem. Phys.* **2003**, *119*, 12753–12762.

(90) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A Combinatorially Optimized Range-Separated Hybrid meta-GGA Density Functional with VV10 Nonlocal Correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(91) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(92) Najibi, A.; Goerigk, L. The Nonlocal Kernel in Van der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ωB97M-V Approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738.

(93) Santra, G.; Martin, J. M. Some Observations on the Performance of the Most Recent Exchange-Correlation Functionals for the Large and Chemically Diverse GMTKN55 Benchmark. AIP Conference Proceedings. 2019.

(94) Reilly, A. M.; Tkatchenko, A. Understanding the Role of Vibrations Exact Exchange and Many-Body Van der Waals Interactions in the Cohesive Properties of Molecular Crystals. *J. Chem. Phys.* **2013**, *139*.

(95) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169.

(96) Kresse, G.; Furthmüller, J. Efficiency of *Ab-Initio* Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.

(97) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953.

(98) Langreth, D. C.; Mehl, M. Beyond the Local-Density Approximation in Calculations of Ground-State Electronic Properties. *Phys. Rev. B* **1983**, *28*, 1809.

(99) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(100) Markovich, T.; Blood-Forsythe, M. A.; Rappoport, D.; Kim, D.; Aspuru-Guzik, A. Calibration of the Many-Body Dispersion Range-Separation Parameter. *arXiv preprint arXiv:1605.04987* **2016**,

(101) Sabatini, R.; Gorni, T.; De Gironcoli, S. Nonlocal Van Der Waals Density Functional Made Simple and Efficient. *Phys. Rev. B* **2013**, *87*, 041108.

(102) Peng, H.; Yang, Z.-H.; Perdew, J. P.; Sun, J. Versatile Van Der Waals Density Functional Based on a Meta-Generalized Gradient Approximation. *Phys. Rev. X* **2016**, *6*, 041005.

(103) Sun, J.; Remsing, R. C.; Zhang, Y.; Sun, Z.; Ruzsinszky, A.; Peng, H.; Yang, Z.; Paul, A.; Waghmare, U.; Wu, X., et al. Accurate First-Principles Structures and

Energies of Diversely Bonded Systems From an Efficient Density Functional. *Nat. Chem.* **2016**, *8*, 831–836.

(104) Yang, J. H.; Kitchaev, D. A.; Ceder, G. Rationalizing Accurate Structure Prediction in the meta-GGA SCAN Functional. *Phys. Rev. B* **2019**, *100*, 035132.

(105) Shepard, S.; Smeu, M. First Principles Study of Graphene on Metals with the SCAN and SCAN+rVV10 Functionals. *J. Chem. Phys.* **2019**, *150*.

(106) Varadwaj, A.; Miyake, T. Geometrical- Electronic-and Optical Properties of Vanadium Dioxide: A Theoretical Perspective From meta-GGA SCAN. *ChemistrySelect* **2022**, *7*, e202200171.

(107) Wang, V.; Xu, N.; Liu, J.-C.; Tang, G.; Geng, W.-T. Vaspkit: A User-Friendly Interface Facilitating High-Throughput Computing and Analysis Using Vasp Code. *Comput. Phys. Commun.* **2021**, *267*, 108033.

(108) Moellmann, J.; Grimme, S. DFT-D3 Study of Some Molecular Crystals. *J. Phys. Chem. C* **2014**, *118*, 7615–7621.

(109) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking Density Functional Methods Against the S66 and S66x8 Datasets for Non-Covalent Interactions. *ChemPhysChem* **2011**, *12*, 3421–3433.

(110) Boese, A. D. Density Functional Theory and Hydrogen Bonds: Are We There Yet? *ChemPhysChem* **2015**, *16*, 978–985.

(111) Lin, I.-C.; Seitsonen, A. P.; Tavernelli, I.; Rothlisberger, U. Structure and Dynamics of Liquid Water From *Ab Initio* Molecular Dynamics Comparison of BLYP PBE and revPBE Density Functionals with and Without Van der Waals Corrections. *J. Chem. Theory Comput.* **2012**, *8*, 3902–3910.

(112) Skinner, L.; Galib, M.; Fulton, J.; Mundy, C.; Parise, J.; Pham, V.-T.; Schenter, G.; Benmore, C. The Structure of Liquid Water Up to 360 Mpa From X-Ray Diffraction Measurements Using a High Q-Range and From Molecular Simulation. *J. Chem. Phys.* **2016**, *144*.

(113) Bankura, A.; Karmakar, A.; Carnevale, V.; Chandra, A.; Klein, M. L. Structure Dynamics and Spectral Diffusion of Water From First-Principles Molecular Dynamics. *J. Phys. Chem. C* **2014**, *118*, 29401–29411.

(114) Galib, M.; Limmer, D. T. Reactive Uptake of N_2O_5 by Atmospheric Aerosol is Dominated by Interfacial Processes. *Science* **2021**, *371*, 921–925.

(115) Niblett, S. P.; Galib, M.; Limmer, D. T. Learning Intermolecular Forces at Liquid–Vapor Interfaces. *J. Chem. Phys.* **2021**, *155*.

(116) Perdew, J. P.; Zunger, A. Self-Interaction Correction to Density-Functional Approximations for Many-Electron Systems. *Phys. Rev. B* **1981**, *23*, 5048.

(117) Johnson, E. R.; Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Delocalization Errors in Density Functionals and Implications for Main-Group Thermochemistry. *J. Chem. Phys.* **2008**, *129*.

(118) Hait, D.; Head-Gordon, M. Delocalization Errors in Density Functional Theory Are Essentially Quadratic in Fractional Occupation Number. *J. Phys. Chem. Lett.* **2018**, *9*, 6280–6288.

(119) Carter-Fenk, K.; Herbert, J. M. Electrostatics Does Not Dictate the Slip-Stacked Arrangement of Aromatic π–π Interactions. *Chem. Sci.* **2020**, *11*, 6758–6765.

(120) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential Energy Surface for the Benzene Dimer and Perturbational Analysis of π- π Interactions. *J. Phys. Chem. A* **2006**, *110*, 10345–10354.

(121) Nguyen, B. D.; Chen, G. P.; Agee, M. M.; Burow, A. M.; Tang, M. P.; Furche, F. Divergence of Many-Body Perturbation Theory for Noncovalent Interactions of Large Molecules. *J. Chem. Theory Comput.* **2020**, *16*, 2258–2273.

(122) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the *Ab Initio* Limit for π- π Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893.

(123) Bludskỳ, O.; Rubeš, M.; Soldán, P.; Nachtigall, P. Investigation of the Benzene-Dimer Potential Energy Surface: DFT/CCSD (T) Correction Scheme. *J. Chem. Phys.* **2008**, *128*, 114102.

(124) Williams, H. L.; Chabalowski, C. F. Using Kohn- Sham Orbitals in Symmetry-Adapted Perturbation Theory to Investigate Intermolecular Interactions. *J. Phys. Chem. A* **2001**, *105*, 646–659.

(125) Gauss, J.; Stanton, J. F. The Equilibrium Structure of Benzene. *J. Phys. Chem. A* **2000**, *104*, 2865–2868.

(126) Lao, K. U.; Schaffer, R.; Jansen, G.; Herbert, J. M. Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.

(127) Řezáč, J.; Jurečka, P.; Riley, K. E.; Černỳ, J.; Valdes, H.; Pluháčková, K.; Berka, K.; Řezáč, T.; Pitoňák, M.; Vondrášek, J.; Hobza, P. Quantum Chemical Benchmark Energy and Geometry Database for Molecular Clusters and Complex Molecular Systems (Www.Begdb.Com): A Users Manual and Examples. *Collect. Czechoslov. Chem. Commun.* **2008**, *73*, 1261–1270.

(128) Manna, D.; Kesharwani, M. K.; Sylvetsky, N.; Martin, J. M. Conventional and Explicitly Correlated *Ab Initio* Benchmark Study on Water Clusters: Revision of the Begdb and Water27 Data Sets. *J. Chem. Theory Comput.* **2017**, *13*, 3136–3152.

(129) Lambros, E.; Dasgupta, S.; Palos, E.; Swee, S.; Hu, J.; Paesani, F. General Many-Body Framework for Data-Driven Potentials with Arbitrary Quantum Mechanical Accuracy: Water as a Case Study. *J. Chem. Theory Comput.* **2021**, *17*, 5635–5650.

(130) Dasgupta, S.; Lambros, E.; Perdew, J.; Paesani, F. Elevating Density Functional Theory to Chemical Accuracy for Water Simulations Through a Density-Corrected Many-Body Formalism. *Nat. Commun.* **2021**, *12*, 1–12.

(131) Palos, E.; Lambros, E.; Swee, S.; Hu, J.; Dasgupta, S.; Paesani, F. Assessing the Interplay Between Functional-Driven and Density-Driven Errors in DFT Models of Water. *J. Chem. Theory Comput.* **2022**, *18*, 3410–3426.

(132) Dasgupta, S.; Shahi, C.; Bhetwal, P.; Perdew, J. P.; Paesani, F. How Good Is the Density-Corrected SCAN Functional for Neutral and Ionic Aqueous Systems and What Is So Right About the Hartree-Fock Density? *J. Chem. Theory Comput.* **2022**, *18*, 4745–4761.

(133) Zhong, K.; Yu, C.-C.; Dodia, M.; Bonn, M.; Nagata, Y.; Ohto, T. Vibrational Mode Frequency Correction of Liquid Water in Density Functional Theory Molecular Dynamics Simulations with Van der Waals Correction. *Phys. Chem. Chem. Phys.* **2020**, *22*, 12785–12793.

(134) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How Van der Waals Interactions Determine the Unique Properties of Water. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 8368–8373.

(135) Galib, M.; Duignan, T. T.; Misteli, Y.; Baer, M. D.; Schenter, G. K.; Hutter, J.; Mundy, C. J. Mass Density Fluctuations in Quantum and Classical Descriptions of Liquid Water. *J. Chem. Phys.* **2017**, *146*.

(136) Pestana, L. R.; Mardirossian, N.; Head-Gordon, M.; Head-Gordon, T. *Ab Initio* Molec-

ular Dynamics Simulations of Liquid Water Using High Quality Meta-Gga Functionals. *Chem. Sci.* **2017**, *8*, 3554–3565.

(137) Zhuang, D.; Riera, M.; Zhou, R.; Deary, A.; Paesani, F. Hydration Structure of Na^+ and K^+ Ions in Solution Predicted by Data-Driven Many-Body Potentials. *J. Phys. Chem. B* **2022**, *126*, 9349–9360.

(138) Caruso, A.; Paesani, F. Data-Driven Many-Body Models Enable a Quantitative Description of Chloride Hydration From Clusters to Bulk. *J. Chem. Phys.* **2021**, *155*.

(139) Riera, M.; Mardirossian, N.; Bajaj, P.; Götz, A. W.; Paesani, F. Toward Chemical Accuracy in the Description of Ion–Water Interactions Through Many-Body Representations. Alkali-Water Dimer Potential Energy Surfaces. *J. Chem. Phys.* **2017**, *147*, 161715.

(140) Bajaj, P.; Götz, A. W.; Paesani, F. Toward Chemical Accuracy in the Description of Ion–Water Interactions Through Many-Body Representations. I. Halide–Water Dimer Potential Energy Surfaces. *J. Chem. Theory Comput.* **2016**, *12*, 2698–2705.

(141) Riera, M.; Knight, C.; Bull-Vulpe, E. F.; Zhu, X.; Agnew, H.; Smith, D. G.; Simmonett, A. C.; Paesani, F. MBX: A Many-Body Energy and Force Calculator for Data-Driven Many-Body Simulations. *J. Chem. Phys.* **2023**, *159*.

(142) Duignan, T. T.; Schenter, G. K.; Fulton, J. L.; Huthwelker, T.; Balasubramanian, M.; Galib, M.; Baer, M. D.; Wilhelm, J.; Hutter, J.; Del Ben, M., et al. Quantifying the Hydration Structure of Sodium and Potassium Ions: Taking Additional Steps on Jacob'S Ladder. *Phys. Chem. Chem. Phys.* **2020**, *22*, 10641–10652.

(143) Kostal, V.; Mason, P. E.; Martinez-Seara, H.; Jungwirth, P. Common Cations Are Not Polarizable: Effects of Dispersion Correction on Hydration Structures From *Ab Initio* Molecular Dynamics. *J. Phys. Chem. Lett.* **2023**, *14*, 4403–4408.

(144) Wagle, K.; Santra, B.; Bhattarai, P.; Shahi, C.; Pederson, M. R.; Jackson, K. A.; Perdew, J. P. Self-Interaction Correction in Water–Ion Clusters. *J. Chem. Phys.* **2021**, *154*.

(145) Zhou, K.; Qian, C.; Liu, Y. Quantifying the Structure of Water and Hydrated Monovalent Ions by Density Functional Theory-Based Molecular Dynamics. *J. Phys. Chem. B* **2022**, *126*, 10471–10480.

(146) Mortazavi, M.; Brandenburg, J. G.; Maurer, R. J.; Tkatchenko, A. Structure and Stability of Molecular Crystals With Many-Body Dispersion-Inclusive Density Functional Tight Binding. *J. Phys. Chem. Lett.* **2018**, *9*, 399–405.

(147) Brandenburg, J.; Bates, J.; Sun, J.; Perdew, J. Benchmark Tests of a Strongly Constrained Semilocal Functional with a Long-Range Dispersion Correction. *Phys. Rev. B* **2016**, *94*, 115144.

(148) Wiktor, J.; Ambrosio, F.; Pasquarello, A. Note: Assessment of the SCAN+rVV10 Functional for the Structure of Liquid Water. *J. Chem. Phys.* **2017**, *147*.

TOC graphic