

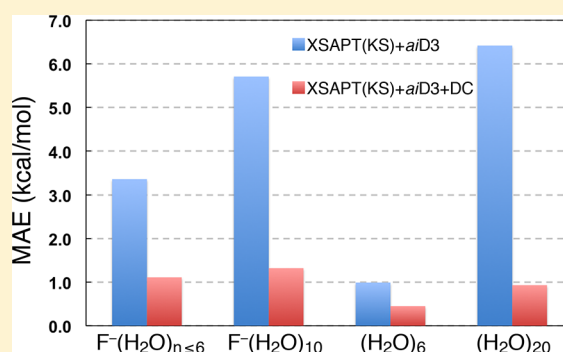
A Simple Correction for Nonadditive Dispersion within Extended Symmetry-Adapted Perturbation Theory (XSAPT)

Ka Un Lao[†] and John M. Herbert^{*‡}

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

Supporting Information

ABSTRACT: Extended symmetry-adapted perturbation theory (XSAPT), in conjunction with empirical “+aiD” potentials fit to *ab initio* dispersion data, is a low-scaling approach to compute intermolecular interaction energies in noncovalent clusters. One shortcoming is that the aiD atom–atom dispersion potentials are independent of the chemical environment of the atoms in question and therefore neglect nonadditive dispersion effects. These can be significant in large systems, so to account for them we test a simple correction to XSAPT(KS)+aiD, where “KS” indicates the use of Kohn–Sham orbitals. This correction, which can be evaluated at fourth-order cost using double- ζ basis sets, is based on comparing second-order SAPT dispersion with and without a self-consistent charge embedding for the monomer wave functions. The correction amounts to ~ 1.4 kcal/mol in $(\text{H}_2\text{O})_6$ but ~ 5.5 kcal/mol in $(\text{H}_2\text{O})_{20}$. With the nonadditive dispersion correction, XSAPT(KS)+aiD affords errors of ~ 1 kcal/mol for isomers of $\text{F}^-(\text{H}_2\text{O})_{10}$ and $(\text{H}_2\text{O})_{20}$, where the benchmarks are complete-basis CCSD(T) energies, as well as for ion–water clusters $\text{X}(\text{H}_2\text{O})_n$ where $n \leq 6$ and $\text{X} = \text{F}^-, \text{Cl}^-, \text{SO}_4^{2-}, \text{Li}^+, \text{Na}^+, \text{or } \text{K}^+$. We also test the MP2 method and a variety of density-functional methods that have been specifically recommended for noncovalent interactions. Among the latter, only $\omega\text{B97X-V}$ and $\omega\text{B97M-V}$ can be recommended for ion–water clusters, as mean errors for other popular approaches (including $\omega\text{B97X-D3}$ and several Minnesota functionals) exceed 1 kcal/mol. Lastly, we examine clathrate-hydrate host/guest complexes whose mixture of hydrogen bonding and dispersion make them challenging tests for noncovalent quantum chemistry. Although the B97-D2 functional performs best for clathrate hydrates and has been previously recommended in other studies of these inclusion complexes, its performance for other systems examined here is quite poor. We are unable to find a functional whose accuracy is $\lesssim 1$ kcal/mol accuracy for both clathrate hydrates and ion–water clusters. However, the XSAPT(KS)+aiD method with the nonadditive dispersion correction can achieve this, with a mean error for the clathrate hydrates of 0.3 kcal/mol.



I. INTRODUCTION

The failure of semilocal and hybrid density functionals to describe dispersion interactions^{1,2} has motivated the development of empirical dispersion corrections for density functional theory (“DFT+D”).^{3–6} The simplest way to incorporate dispersion is via pairwise-additive contributions from atomic multipoles:

$$E_{\text{disp}} = - \sum_{i,j < i}^{\text{atoms}} \left(\frac{C_{6,ij}}{R_{ij}^6} + \frac{C_{8,ij}}{R_{ij}^8} + \frac{C_{10,ij}}{R_{ij}^{10}} + \dots \right) \quad (1)$$

The coefficients $C_{n,ij}$ are obtained from atomic coefficients $C_{n,i}$ according to

$$C_{n,ij} = (C_{n,i}C_{n,j})^{1/2} \quad (2)$$

For each pair of atoms ij , $C_{6,ij}$ is the dipole–dipole dispersion coefficient, $C_{8,ij}$ represents dipole–quadrupole dispersion, $C_{10,ij}$ expresses quadrupole–quadrupole and dipole–octupole dispersion, and so on. The dispersion coefficients can be determined from *ab initio* calculations^{7–11} or by fitting to a

benchmark training set.^{12–20} Because eq 1 diverges for small interatomic separations R_{ij} , however, a damping function must be used at short range. This also helps to minimize double-counting of electron correlation effects that are already included at short range via the exchange–correlation functional. These damping functions often include several functional-dependent parameters.³

Nonadditive environment-dependent effects cause the dispersion coefficients to deviate from their gas-phase values but early versions of the “+D” correction in eq 1 did not account for this. Use of environment-dependent $C_{n,ij}$ coefficients is a first step to account for what Dobson has termed “type A” nonadditive dispersion,²¹ in which crowding or squeezing by neighboring atoms reduces the atomic polarizability and therefore the dispersion energy. (This leads to a many-body contribution to dispersion that is repulsive,²⁰ whereas the pairwise-additive dispersion in eq 1 is strictly

Received: May 30, 2018

Published: September 10, 2018



attractive.) Type A effects are incorporated in methods such as the Tkatchenko–Scheffler dispersion correction^{10,22} and the Becke–Johnson exchange-hole dipole model,^{8,23} because the coefficients $C_{n,i}$ used in these methods are computed directly from the local electron density. These effects are also included, to some extent, by Grimme’s D3 dispersion correction,¹¹ which still uses the pairwise-additive form of eq 1 but with atomic dispersion parameters that are dependent on the hybridization state of the atom.

An alternative, *ab initio* way to compute intermolecular interactions is based on symmetry-adapted perturbation theory (SAPT),^{24–26} where the dispersion energy can be separated from the remaining components of the interaction energy in a manner that is well-defined at all distances. Electron correlation for the monomers can be included in a low-cost way using Kohn–Sham (KS) orbitals, provided that the exchange–correlation potential is asymptotically corrected for each monomer.²⁷ This can be accomplished using range-separated hybrid functionals.^{27,28} At second order in perturbation theory, this “SAPT(KS)” approach affords accurate results for all of the interaction energy components *except* dispersion.^{20,27,29,30} The dispersion energy is cleanly separable, however, so as compared to DFT, the SAPT approach actually lends itself much more naturally to an empirical dispersion potential of the form in eq 1,^{19,20,30,31} because there is no double-counting of electron correlation.

A damped version of the pairwise dispersion potential in eq 1 can be fit to dispersion energies computed using higher-level SAPT methods that afford accurate dispersion energies, e.g., SAPT2+(3).²⁴ We call these “*ab initio* dispersion potentials” (“+*aiD*”),³² since they represent true dispersion energies in a manner that is much closer to a first-principles approach as compared to DFT+D.^{17,18,20,30} In particular, our second-generation dispersion potentials (+*aiD2*)³⁰ and third-generation dispersion potentials (+*aiD3*)²⁰ afford an accuracy of $\lesssim 1$ kcal/mol for small-molecule dimers, for both total intermolecular interaction energies as well as individual energy components.³³ The SAPT(KS)+*aiD* method exhibits cubic scaling with respect to fragment (monomer) size, $O(N_f^3)$, and in conjunction with our third-generation *aiD3* potential exhibits very high accuracy for a variety of noncovalent datasets.²⁰

We have also generalized traditional two-body SAPT to obtain an “extended” version that we call XSAPT,^{20,29,34,35} in which many-body polarization effects are incorporated self-consistently into the monomer wave functions by means of the variational explicit polarization (XPol) method.³⁶ The XSAPT(KS)+*aiD* method^{20,32} affords accurate interaction energies and energy components for standard databases of dimer interaction energies, as well as other small clusters where accurate benchmarks are available,^{20,30,31} and for large collections of monomers it is more efficient than DFT.^{20,31} On the other hand, it is precisely in these larger systems where one starts to worry about not only type-A but also Dobson’s type-B nonadditivity effects.²¹ The latter occur when a third polarizable center screens the interaction between two other centers; the Axilrod–Teller–Muto triple-dipole interaction^{37,38} is the lowest-order example. At present, available SAPT-based dispersion potentials are fixed and independent of molecular environment, similar to early-stage DFT+D methods.^{12,13,15} In the present work, we propose a correction to XSAPT(KS)+*aiD* that accounts for type-A nonadditive dispersion by examining how the second-order SAPT

dispersion energy changes in the presence of the XPol embedding environment, modifying the atomic-pairwise *aiD* dispersion to account for the squeezing and crowding effects of the molecular environment.

II. THEORY

We seek an approach that can respond to changes in the local electron density as an alternative to reparameterizing the dispersion coefficients for different hybridization environments. The goal is to incorporate many-body dispersion, but we note that the only many-body part of the XSAPT procedure is the self-consistent XPol charge embedding.^{20,29,34} This is built upon a double self-consistent field approach,^{34,39,40} wherein each monomer wave function is iteratively converged in the presence of atomic point charges for the other monomers. These point charges are determined in order to reproduce the molecular electrostatic potential of the environment,²⁹ and are themselves iteratively updated until convergence. SAPT corrections through second order are then applied in a pairwise-additive way.²⁰ In view of this, we propose to examine how the second-order dispersion energy,

$$E_{\text{disp}} = E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)} \quad (3)$$

changes when XPol wave functions are substituted for the gas-phase monomer wave functions used in traditional SAPT. We propose the following dispersion correction (DC):

$$E_{\text{DC(full)}} = \sum_A \sum_{B>A} (E_{\text{disp,AB}}^{(2),\text{XSAPT}} + E_{\text{exch-disp,AB}}^{(2),\text{XSAPT}} - E_{\text{disp,AB}}^{(2),\text{SAPT}} - E_{\text{exch-disp,AB}}^{(2),\text{SAPT}}) \quad (4)$$

The double sum runs over all *pairs* of monomers and evaluates the difference between the SAPT and XSAPT versions of the second-order dispersion energy in eq 3. For the XSAPT terms, the XPol wave functions are computed in the environment of *all* monomers, so eq 4 is a correction for many-body dispersion or, equivalently, an adjustment of the pairwise dispersion to account for the many-body polarization effects of the environment.

Within the context of noncovalent quantum chemistry, the term “many-body” (or more specifically, “body”) is commonly used in two different ways by different practitioners. On the one hand there is the SAPT community, to whom the “bodies” are monomers and therefore nonadditive effects do not appear until trimers, by definition. It is in this sense of the term that XSAPT implicitly incorporates many-body induction, by means of the XPol procedure that polarizes the monomers so that their charge distributions are representative of the many-body environment. At the same time, however, there is a sizable community of developers working on DFT+D methods, a category that includes both Grimme-style empirical dispersion potentials^{3,6} but also Tkatchenko–Scheffler-type models.²² Within this latter community the “bodies” are implicitly *atoms*, and dispersion effects not included in pairwise-atomic C_n/R^n potentials are considered nonadditive. This is Dobson’s point of view,²¹ in defining nonadditivity of types A and B. XSAPT+*aiD* contains elements from both of these communities and as such we are forced into a bifurcated usage of the term “many-body”. The “many-body dispersion” that is captured by the correction in eq 4 corresponds to Dobson’s type A, wherein squeezing and crowding due to

many-body polarization effects (arising from the other monomers) modifies the interatomic dispersion.

The cost to evaluate $E_{\text{exch-disp}}^{(2)}$ grows as $O(N_f^5)$, whereas $E_{\text{disp}}^{(2)}$ is only $O(N_f^4)$. At the same time, the former is typically several times smaller than the latter and decays exponentially with increasing monomer separation near the equilibrium separation. As such, the SAPT and XSAPT exchange-dispersion terms may cancel to a much greater extent, compared to the “direct” dispersion terms $E_{\text{disp}}^{(2)}$. In practice, we will therefore approximate eq 4 by omitting the exchange-dispersion terms. This defines

$$E_{\text{DC(partial)}} = \sum_A \sum_{B>A} (E_{\text{disp,AB}}^{(2),\text{XSAPT}} - E_{\text{disp,AB}}^{(2),\text{SAPT}}) \quad (5)$$

The extent to which $E_{\text{DC(full)}} \approx E_{\text{DC(partial)}}$ will be tested herein.

Among the components of the SAPT interaction energy, the dispersion energy converges most slowly with respect to the basis set.^{41,42} However, we anticipate that the difference in eq 5 may converge faster, which is borne out by the tests presented below. (Similarly, the difference between coupled Kohn–Sham and uncoupled Hartree–Fock dispersion energies exhibits a weaker dependence on the basis set than the two dispersion energies themselves.⁴³) Previous work has shown that triple- ζ basis sets are needed to converge the electrostatic and induction energies in (X)SAPT,^{29,30} so in the end we arrive at a composite method that employs a triple- ζ basis set to perform the XSAPT(KS)+aiD3 calculation, whose cost scales as $O(N_f^3)$, but a double- ζ basis set to evaluate the $O(N_f^4)$ correction in eq 5.

III. PRELIMINARY TESTS

III.A. Computational Details. Ions have proven to be challenging for SAPT,^{44,45} and we will examine F^- , Cl^- , SO_4^{2-} , Li^+ , Na^+ , and K^+ in $(\text{H}_2\text{O})_n$ clusters ($n = 1-6$). To accomplish this, our aiD3 dispersion potential²⁰ was extended to include parameters for Li, Na, and K, as described in the Supporting Information. Some larger clusters including $(\text{H}_2\text{O})_{20}$, $\text{F}^-(\text{H}_2\text{O})_{10}$, and clathrate hydrates will be examined as well. All cluster structures and the details of how they were obtained can be found in the Supporting Information.

Energies reported in this work are intermolecular interaction energies, not including relaxation energies for the monomers. Benchmark interaction energies at the CCSD(T) level, extrapolated to the complete basis set (CBS) limit, are provided in the Supporting Information, along with details of the procedures used to obtain them. As in previous work,³⁰ we benchmark individual energy components (electrostatics, exchange repulsion, induction, and dispersion) at the SAPT2+(3)/aug-cc-pVTZ level. (Hereafter, we will abbreviate aug-cc-pVXZ as aXZ.) Benchmarks at this level of theory were taken from ref 27 for the S22 dataset,⁴⁶ and reported here for the S66 dataset⁴⁷ (see the Supporting Information).

The success of SAPT(KS) calculations is crucially dependent on achieving correct asymptotic behavior of the exchange-correlation functional for each monomer.²⁷ XSAPT(KS) calculations reported here use the LRC- ω PBE functional,⁴⁸ for which the range-separation parameter (ω) is tuned in a monomer-specific fashion in order to obtain proper asymptotic behavior. A well-established way to do this is the “IP tuning” procedure of Baer and co-workers,⁴⁹ wherein ω is adjusted so that

$$-\epsilon_{\text{HOMO}}(\omega) = \text{IP}(\omega) \quad (6)$$

where ϵ_{HOMO} is the highest occupied energy level obtained using LRC- ω PBE, and “IP” denotes the ionization potential computed using the same functional. The value of ω is adjusted until eq 6 is satisfied, and we call this optimized value ω_{IP} . An alternative “global density-dependent” (GDD) tuning procedure has also been suggested⁵⁰ and used in our most recent XSAPT work.³² In this approach, an optimal value of ω is obtained from the average distance between an electron in the outer regions of the molecule and the exchange hole in the region of compact orbitals. We examine both procedures, and tuned values of ω for each of the monomers considered here are available in the Supporting Information. All XSAPT(KS)+aiD3 calculations use the hpTZVPP basis set, as recommended in our previous work.^{20,30} This basis consists of def2-TZVPP augmented with diffuse functions on non-hydrogen atoms that are taken from Pople’s 6-311+G basis set.

As compared to the asymptotic correction schemes that are most often employed in DFT-SAPT calculations,^{25,26} which involve splicing together a short-range exchange-correlation potential with an asymptotically correct one, the LRC- ω PBE tuning procedure has the advantage that it preserves the relationship $v_{\text{xc}} = \delta E_{\text{xc}}/\delta \rho$. This relationship is sacrificed when v_{xc} is spliced together from different sources. Alternatively, Chai and co-workers⁵¹ have developed an asymptotic correction called the strictly localized Fermi-Amaldi (LFAs) scheme that can be added to any semilocal functional, which achieves correct asymptotic behavior but also ensures that v_{xc} is the functional derivative of a well-defined energy functional. We will test the LFAs correction for use with XSAPT(KS).

Finally, we will test some density functionals that have emerged as reasonable choices for describing noncovalent interactions.^{20,52-57} These include ω B97X-D3;⁵⁵ B97-D2;^{15,58,59} PW6B95-D3(0), which uses the “zero-damping” D3(0) potential;^{11,60} the Minnesota functionals M06-2X,⁶¹ M11,⁶² and MN15;⁶³ and finally the functionals B97M-V,⁵⁷ ω B97X-V,⁵⁶ and ω B97M-V,⁶⁴ each of which includes nonlocal VV10 correlation.⁶⁵ These were selected for their performance on noncovalent interactions, but we will also examine the recently developed SCAN functional,⁶⁶ which satisfies all known exact constraints on the Hohenberg–Kohn functional, along with its hybrid variant SCAN0.⁶⁷ We include these because the SCAN functional has recently been used in *ab initio* molecular dynamics simulations of both liquid water⁶⁸⁻⁷⁰ and aqueous NaCl,⁷¹ but to the best of our knowledge there has been only one previous study of SCAN for the analogous cluster systems.⁷² Finally, we test a variety of DFT methods that have been recommended for clathrate hydrates,⁷³ as detailed in Section IV.C. All DFT calculations reported here use the def2-TZVPPD basis set,^{74,75} and all are counterpoise-corrected.

The ORCA program⁷⁶ was used for the alkali–water cluster optimizations with TPSS+D3, the CCSD(T) calculations using the domain-based local pair natural orbital (DPLNO) approximation,⁷⁷ and the MP2 calculations for clathrate hydrates. The Psi4 program⁷⁸ was used for all SAPT2+(3) calculations, and for single-point MP2 and CCSD(T) calculations for $\text{SO}_4^{2-}(\text{H}_2\text{O})_n$ and alkali–water clusters. All remaining calculations were performed using a locally modified version of Q-Chem.⁷⁹

III.B. Density Functionals for XSAPT(KS). The SAPT energy decomposition has been generalized for use with

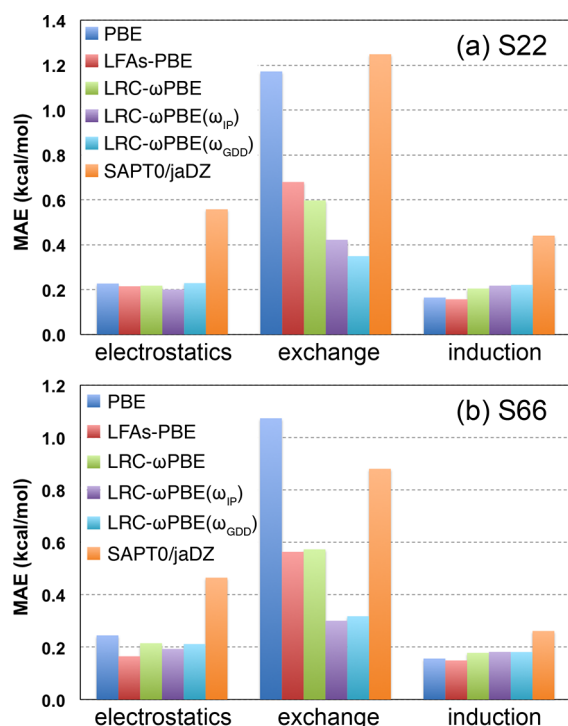


Figure 1. Mean absolute errors (MAEs) for individual energy components in XSAPT(KS)/hpTZVPP calculations, using various KS density functionals, for (a) the S22 dataset and (b) the S66 dataset. SAPT0/jun-cc-pVDZ results are also shown as a point of comparison, and the benchmarks are SAPT2+(3)/aug-cc-pVTZ calculations. The LRC- ω PBE calculations used $\omega = 0.3$ bohr⁻¹ (see ref 80), whereas the “ ω_{IP} ” and “ ω_{GDD} ” values are tuned.

XSAPT,³⁰ and in Figure 1 we evaluate the performance of XSAPT(KS) for reproducing energy components for the S22 and S66 sets of dimers.^{46,47} Five different density functionals are tested, where four of them are asymptotically corrected, and the benchmark is SAPT2+(3)/aug-cc-pVTZ. (Hereafter, we abbreviate aug-cc-pVXZ as “aXZ”, for X = D, T, Q.) We do not consider the dispersion contribution to the energy in this analysis, because we know it to be inaccurate,^{29,30} hence the motivation for XSAPT(KS)+*aiD*.

Mean absolute errors (MAEs) for the electrostatic and induction (polarization) energies are similar for all of the functionals tested here. This is consistent with the generally well-behaved basis-set convergence of the electrostatic energy in SAPT, which has a tendency to converge in triple- ζ basis sets.²⁹ For the exchange (Pauli repulsion) energy, MAEs decrease in the following order:

$$\text{PBE} > \text{LFA-PBE} \approx \text{LRC-}\omega\text{PBE} > \text{LRC-}\omega\text{PBE}(\omega_{IP}) \\ \approx \text{LRC-}\omega\text{PBE}(\omega_{GDD})$$

The order suggests that the improvement is due to better asymptotic behavior of the exchange-correlation potential. Consistent with previous work,²⁷ this demonstrates the necessity of tuning the range separation parameter for SAPT(KS) and XSAPT(KS) calculations. Both the ω_{IP} and ω_{GDD} tuning schemes afford MAEs of <0.5 kcal/mol in each of the energy components. Whereas LFA-PBE is more computationally efficient, because it does not require Hartree–Fock exchange, the MAE for the exchange energy is doubled, compared to LRC-DFT with ω tuning.

The SAPT0 method,⁸¹ when used in conjunction with the jun-cc-pVDZ (jaDZ) basis set, has been suggested for use in large systems.⁸² SAPT0 is essentially Hartree–Fock theory for the monomers combined with MP2 theory for the intermolecular interactions, and thus suffers from the usual MP2-like overestimation of dispersion. However, dispersion converges more slowly to the basis-set limit, compared to the other energy components,^{41,42} and use of the partially augmented jaDZ basis exploits error cancellation to obtain reasonable interaction energies.⁸³ (For large monomers, however, this cancellation appears to work less well.³²) SAPT0/jaDZ calculations may not be fully converged for electrostatics and polarization, and this is borne out in Figure 1, where it is shown that SAPT0/jaDZ exhibits larger errors for the electrostatic, induction, and exchange energies, compared to any of the SAPT(KS) methods except that based on PBE.

Although SAPT0/jaDZ can be used to estimate energy components in reasonably large systems,^{81,84} it exhibits fifth-order scaling, compared to the cubic-scaling XSAPT(KS)+*aiD* approach, which actually affords smaller errors. Upon adding the empirical dispersion potential and comparing to CCSD(T)/CBS interaction energies (Figure 2), we find that XSAPT-

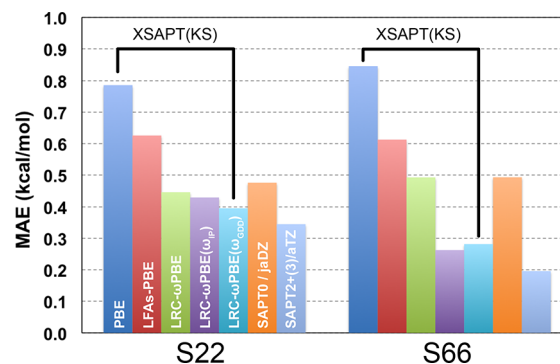


Figure 2. MAEs for S22 and S66 with respect to CCSD(T)/CBS benchmarks, for XSAPT(KS)+*aiD*3/hpTZVPP with five different density functionals. Also shown are SAPT0/jaDZ and SAPT2+(3)/aTZ results.

(KS)+*aiD*3 is indeed more accurate than SAPT0/jaDZ, predicting each energy component to within 0.1 kcal/mol of SAPT2+(3). Note that both SAPT2+(3) and CCSD(T) are $O(N^7)$ methods but the former is inherently free of basis set superposition error (BSSE) and thus can be trusted in triple- ζ basis sets, without counterpoise correction.

In view of these results, all XSAPT(KS)+*aiD*3 calculations for the remainder of this article will be based on tuned LRC- ω PBE.

III.C. Basis Sets for Nonadditive Dispersion. We expect error cancellation to reduce the basis-set dependence of the dispersion corrections in eqs 4 and 5, so this dependence is examined here using three small systems for which the nonadditive correction is relatively large: F[−](H₂O)₆, Cl[−](H₂O)₆, and SO₄^{2−}(H₂O)₆. (Results are similar when LRC- ω PBE is used with either ω_{IP} or ω_{GDD} tuning, so only the ω_{IP} data are shown here.) We take XSAPT(KS)/hpTZVPP results using the “full” correction (eq 4) as the benchmark, and present error statistics for these three systems in Figure 3, using a variety of basis sets. Basis sets are ordered left to right in Figure 3, in order of increasing basis size, as a guide to the eye in gauging convergence (or the lack thereof).

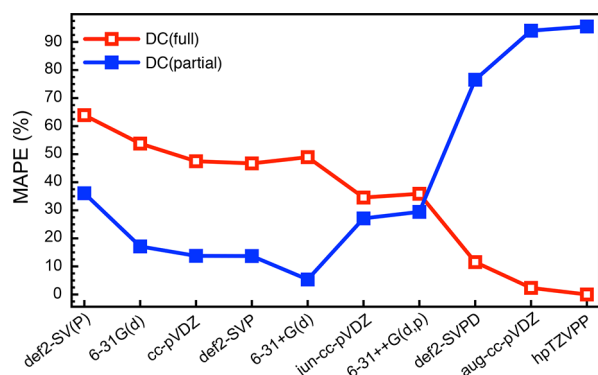


Figure 3. Mean absolute percentage errors (MAPEs) in the “full” (eq 4) and “partial” (eq 5) nonadditive dispersion corrections, in a variety of double- ζ basis sets. The test set consists of $\text{F}^-(\text{H}_2\text{O})_6$, $\text{Cl}^-(\text{H}_2\text{O})_6$, and $\text{SO}_4^{2-}(\text{H}_2\text{O})_6$. Errors are computed relative to the full nonadditive correction computed using hpTZVPP. All calculations including the benchmarks use LRC- ω PBE(ω_{IP}). Basis size increases from left to right.

For the basis sets examined in Figure 3, the full nonadditive correction converges smoothly to the hpTZVPP result as the size of the double- ζ basis set is increased, with aDZ converged to within 2%. Good performance in this scheme requires converging both the direct- and exchange-dispersion energy corrections. In contrast, the partial correction (eq 5) omits the exchange-dispersion component and is therefore more reliant on error cancellation, and convergence is found to be correspondingly more erratic. The minimum error is 5%, for 6-31+G(d), but increases for smaller basis sets. Because fourth-order scaling is quite affordable with 6-31+G(d), we will use this basis to evaluate $E_{\text{DC}(\text{partial})}$ in the remainder of this work, except clathrate hydrate host/guest complexes, where $E_{\text{DC}(\text{full})}$ calculated with hpTZVPP will be used in order to obtain highly accurate results in these systems.

IV. RESULTS FOR CLUSTERS

IV.A. Small Ion–Water Clusters. Ionic systems prove to be challenging test cases for SAPT and other methods.^{20,45} Table 1 shows MAEs for interaction energies in small ion–water clusters, $\text{X}(\text{H}_2\text{O})_{n=1-6}$ with $\text{X} = \text{F}^-$, Cl^- , SO_4^{2-} , Li^+ , Na^+ , and K^+ , compared to CCSD(T)/CBS benchmarks. The XSAPT(KS)+aiD3 method based on ω_{GDD} -tuned LRC- ω PBE works slightly better (0.9 kcal/mol on average) than ω_{IP} tuning for the anionic clusters. On the other hand, ω_{IP} tuning works slightly better (0.3 kcal/mol on average) than ω_{GDD} tuning for the cations. For the entire dataset, the MAEs are 2.0 kcal/mol (ω_{IP}) and 1.6 kcal/mol (ω_{GDD}).

The performance of XSAPT(KS)+aiD3 for the anions is largely improved by including the DC term. For example, the MAEs for the fluoride–water clusters are reduced from 3.36 kcal/mol (ω_{IP}) and 2.39 kcal/mol (ω_{GDD}) to 1.11 and 0.55 kcal/mol, respectively. Although errors increase slightly when the DC term is added to the cation calculations, the correction is small in those cases; its maximum value is only 0.65 kcal/mol, for $\text{K}^+(\text{H}_2\text{O})_6$.

In ref 20, we speculated that relatively large errors in XSAPT(KS)+aiD3 calculations for halide–water clusters might arise from missing three-body induction couplings, the formalism for which has been worked out and implemented,²⁹ but which are somewhat expensive to evaluate. We further speculated that a “ δ MP2” correction might help, as it accounts

Table 1. Mean Absolute Errors^a in Interaction Energies for Ion–Water Clusters^b

method	Mean Absolute Error, MAE ^c (kcal/mol)		
	anions	cations	all
Supersystem Methods			
MP2/CBS	0.52	0.34	0.43
B97-D2 ^d	4.16	3.74	3.95
ω B97X-D3 ^d	1.30	3.17	2.24
B97M-V ^d	0.37	0.96	0.66
ω B97X-V ^d	0.42	0.50	0.46
ω B97M-V ^d	0.45	0.51	0.48
M06-2X ^d	3.26	1.50	2.38
M11 ^d	1.55	1.19	1.37
MN15 ^d	1.64	1.14	1.39
PW6B95-D3(0) ^d	0.41	0.85	0.63
SCAN ^d	4.73	1.10	2.92
SCAN0 ^d	4.23	1.46	2.85
XSAPT Methods			
XSAPT(KS)+aiD3 (ω_{IP}) ^e	3.08	0.82	1.95
XSAPT(KS)+aiD3+DC (ω_{IP}) ^{e,f}	1.37	0.87	1.12
XSAPT(KS)+aiD3 (ω_{GDD}) ^e	2.20	1.07	1.63
XSAPT(KS)+aiD3+DC (ω_{GDD}) ^{e,f}	0.71	1.14	0.92

^aError is defined with respect to CCSD(T)/CBS benchmarks.

^b $\text{X}(\text{H}_2\text{O})_{n=1-6}$ with $\text{X} = \text{F}^-$, Cl^- , SO_4^{2-} , Li^+ , Na^+ , and K^+ . ^cThe lowest three MAEs in each column are highlighted in boldface italic type. ^dUsing def2-TZVPPD with counterpoise correction. ^eUsing hpTZVPP for XSAPT(KS). ^fUsing 6-31+G(d) for the nonadditive dispersion correction.

for higher-order induction–dispersion coupling,⁸² and has been used successfully for other SAPT calculations involving anions.⁴⁵ (Unfortunately, the δ MP2 correction requires a supersystem MP2 calculation.) The present results demonstrate that increased accuracy (relative to the calculations reported in ref 20) can be achieved without turning to these more-expensive corrections, simply by correcting for type-A squeezing effects on the dispersion coefficients. The non-additive dispersion correction ranges up to 3 kcal/mol in magnitude, with the largest value for $\text{SO}_4^{2-}(\text{H}_2\text{O})_6$. When it is included, the composite $\mathcal{O}(N_f^4)$ method, XSAPT(KS)+aiD3+DC, affords MAEs of 1.1 kcal/mol (ω_{IP}) and 0.9 kcal/mol (ω_{GDD}), for the entire ion–water dataset.

Considering some supersystem methods applied to the same dataset (see Table 1), the MP2/CBS method affords mean errors of ≤ 1.0 kcal/mol in all cases, with a MAE of 0.4 kcal/mol for the entire dataset and slightly better performance for cations than for anions. However, quadruple- ζ basis sets are required for the extrapolation. A correction defined as

$$\delta_{\text{CCSD(T)}} = E_{\text{CCSD(T)}} - E_{\text{MP2}} \quad (7)$$

for higher-order correlation effects contributes between +0.2% and +1.6% for $\text{X} = \text{Cl}^-$, SO_4^{2-} , and K^+ . For $\text{X} = \text{F}^-$, Li^+ , and Na^+ the percentage magnitude of this correction is comparable but the sign is negative. Details regarding the $\delta_{\text{CCSD(T)}}$ correction can be found in the Supporting Information.

Of the 11 DFT functionals examined here, which were chosen specifically for their performance on noncovalent interactions, only B97M-V, ω B97X-V, ω B97M-V, and PW6B95-D3(0) afford MAEs of <1 kcal/mol for both cations and anions. MAEs for these functionals are 0.5 kcal/mol for both ω B97X-V and ω B97M-V, 0.6 kcal/mol for PW6B95-

Table 2. Mean Absolute Errors for Interaction Energies and Relative Energies of Various Clusters^a

method	(H ₂ O) ₆			(H ₂ O) ₂₀			F [−] (H ₂ O) ₁₀		
	<i>E</i> _{int}		<i>E</i> _{rel} ^{<i>b</i>}	<i>E</i> _{int}		<i>E</i> _{rel} ^{<i>c</i>}	<i>E</i> _{int}		<i>E</i> _{rel} ^{<i>d</i>}
	MAE (kcal/mol)	(%)		MAE (kcal/mol)	(%)		MAE (kcal/mol)	(%)	
Supersystem Methods									
MP2/CBS	0.33	(0.7)	0.46	2.69	(1.4)	0.18	2.02	(1.2)	0.31
B97-D2 ^{<i>e</i>}	4.50	(9.6)	0.51	23.04	(11.6)	0.49	11.34	(6.7)	0.63
<i>ω</i> B97X-D3 ^{<i>e</i>}	1.54	(3.3)	0.22	6.12	(3.1)	0.33	4.96	(2.9)	0.46
B97M-V ^{<i>e</i>}	0.38	(0.8)	0.70	2.86	(1.4)	0.33	0.15	(0.1)	0.31
<i>ω</i> B97X-V ^{<i>e</i>}	0.27	(0.6)	0.06	1.68	(0.8)	0.07	0.24	(0.1)	0.32
<i>ω</i> B97M-V ^{<i>e</i>}	0.26	(0.6)	0.19	0.93	(0.5)	0.13	0.48	(0.3)	0.11
M06-2X ^{<i>e</i>}	1.31	(2.8)	1.70	3.03	(1.5)	1.18	8.33	(5.0)	0.53
M11 ^{<i>e</i>}	0.88	(1.9)	1.39	10.20	(5.2)	1.06	1.65	(1.0)	0.77
MN15 ^{<i>e</i>}	0.72	(1.5)	1.26	8.91	(4.5)	0.86	2.11	(1.2)	0.73
PW6B95-D3 ^{<i>e</i>}	0.27	(0.6)	0.20	6.31	(3.2)	0.29	1.00	(0.6)	0.32
SCAN ^{<i>e</i>}	4.75	(10.1)	0.53	17.74	(9.0)	0.45	14.30	(8.5)	2.27
SCAN0 ^{<i>e</i>}	4.47	(9.5)	0.04	14.91	(7.5)	0.27	12.47	(7.4)	0.41
XSAPT Methods									
XSAPT(KS)+ <i>ai</i> D3 (<i>ω</i> _{IP}) ^{<i>f</i>}	0.99	(2.1)	0.43	6.42	(3.2)	1.02	5.71	(3.4)	0.56
XSAPT(KS)+ <i>ai</i> D3+DC (<i>ω</i> _{GDD}) ^{<i>f,g</i>}	0.45	(0.9)	0.20	0.93	(0.5)	0.72	1.32	(0.8)	0.32
XSAPT(KS)+ <i>ai</i> D3 (<i>ω</i> _{GDD}) ^{<i>f</i>}	0.38	(0.8)	0.45	3.48	(1.8)	1.02	3.48	(92.1)	0.53
XSAPT(KS)+ <i>ai</i> D3+DC (<i>ω</i> _{IP}) ^{<i>f,g</i>}	1.10	(2.3)	0.21	2.23	(1.1)	0.72	0.98	(0.6)	0.33

^aError is defined with respect to CCSD(T)/CBS benchmarks, and the lowest three MAEs among supersystem methods are indicated in boldface italic type. ^bRelative to the prism isomer. ^cRelative to isomer 1. ^dRelative to isomer 6. ^eUsing def2-TZVPPD with counterpoise correction. ^fUsing hpTZVPP for XSAPT(KS). ^gUsing 6-31+G(d) for the nonadditive dispersion correction.

D3(0), and 0.7 kcal/mol for B97M-V, and these are comparable to MP2/CBS results. These results are consistent with those in other recent studies. For example, in a recent DFT study of interaction energies and harmonic vibrational frequencies for M⁺(H₂O) systems,⁸⁵ with M = Li, Na, K, Rb, and Cs, it was found that *ω*B97M-V performed best among a variety of functionals. More generally, *ω*B97X-V and *ω*B97M-V afford the best performance for noncovalent interactions in a comprehensive study of 200 modern density functionals.⁸⁶ For the ion–water data set assembled here, the lowest overall MAE is obtained using *ω*B97X-V, which furthermore affords a similar MAE for both the anion–water and the cation–water clusters, and is therefore recommended for use in ion–water clusters.

The B97-D2 functional exhibits the worst overall performance of any of the methods considered in Table 1. For the cations, its MAE is the largest of the functionals tested here; for the anions, only the SCAN and SCAN0 functionals exhibit larger MAEs. The performance of *ω*B97X-D3 is also quite poor, especially for the cations where the errors range up to 5.7 kcal/mol, for lithium–water clusters. The Minnesota functionals M06-2X, M11, and MN15 exhibit similarly large errors for the anions, although they are superior to *ω*B97X-D3 for the cations. The SCAN functional overestimates interaction energies for all anionic clusters and some cationic clusters and affords a MAE of 2.9 kcal/mol for the entire dataset.

It is worth remarking that despite satisfying all known exact constraints on the Hohenberg–Kohn functional,⁶⁶ the SCAN functional nevertheless requires an empirical dispersion correction for best performance.^{72,87,88} For ion–water clusters, however, SCAN already overestimates the interaction energies and this will only be exacerbated by the addition of an attractive dispersion correction. SCAN, with no dispersion correction, has been used in *ab initio* molecular dynamics simulations to simulate the potential mean force of a Na⁺⋯Cl[−] ion pair in liquid water, as has *ω*B97X-V, and these two

functionals predict different relative stabilities for the two free-energy minima.⁷¹ The *ω*B97X-V prediction is consistent with that inferred from neutron diffraction experiments,⁷¹ and our results support the conclusion that *ω*B97X-V is preferable for simulations of aqueous ions.

The hybrid SCAN0 functional offers only a marginal improvement upon SCAN, overestimating the interaction energies for all of the ionic clusters except K⁺(H₂O) and K⁺(H₂O)₂. As with SCAN itself, the addition of a van der Waals model⁸⁹ to SCAN0 will only further exaggerate the already too-large interaction energies for ion–water clusters. Although SCAN0 predicts accurate interaction energies for ion pairs containing halide anions,⁹⁰ it does not provide accurate interaction energies for the ion–water cluster considered here. Note that, along with SCAN0, each of the functionals such as M06-2X, M11, MN15, *ω*B97X-D3, and *ω*B97X-V contains Hartree–Fock exchange, whereas B97M-V does not. This makes the latter functional an attractive option for large systems, especially given that its MAE is only 0.7 kcal/mol. In cases where Hartree–Fock exchange is not a limiting factor, the use of *ω*B97X-V is clearly preferable.

Finally, there is PW6B95-D3(0), which is a seemingly odd combination but one that has been used previously by Grimme and co-workers,^{91–96} and subsequently by us.³² It has shown surprisingly good performance in noncovalent applications involving large monomers, and its MAE for these ion–water clusters is only 0.6 kcal/mol. However, PW6B95-D3(0) performs quite poorly for (H₂O)₂₀ clusters, as demonstrated below. Nevertheless, it has previously been recommended for use in water clusters.⁹⁷

IV.B. Larger Clusters. For the ion–water clusters, the largest values of the nonadditive dispersion correction are obtained for the largest clusters we examined. (Large ion polarizabilities also increase the magnitude of this correction, as in the sulfate clusters.) In this section, we will therefore

examine results for some larger clusters, including $(\text{H}_2\text{O})_{20}$ and $\text{F}^-(\text{H}_2\text{O})_{10}$.

First, however, we consider eight low-lying structures of $(\text{H}_2\text{O})_6$, from ref 98. MAEs for both total interaction energies and relative isomer energies, computed using various methods and evaluated with respect to CCSD(T)/CBS benchmarks, are shown in Table 2. For interaction energies, XSAPT(KS)+*ai*D3 affords MAEs of 1.0 kcal/mol (ω_{IP}) and 0.4 kcal/mol (ω_{GDD}). Adding the DC term, which averages 1.4 kcal/mol for these eight clusters and is repulsive in nature, changes the MAEs to 0.5 kcal/mol (ω_{IP}) and 1.1 kcal/mol (ω_{GDD}).

All of the supersystem methods examined here afford MAEs below 1 kcal/mol for the interaction energies of $(\text{H}_2\text{O})_6$ isomers, except for B97-D2, ω B97X-D3, M06-2X, SCAN, and SCAN0. B97-D2 significantly underestimates the interaction energies whereas SCAN and SCAN0 overestimate them. As with the ion–water clusters, the best methods are ω B97X-V, ω B97M-V, and PW6B95-D3(0), each with a MAE of 0.3 kcal/mol for this dataset. However, PW6B95-D3(0) predicts the wrong relative energies for the “bag” and “cyclic chair” isomers, a subtlety that is also problematic for the XSAPT methods. Surprisingly, SCAN0 gives the best relative energies, compared with CCSD(T)/CBS results, even while it significantly overestimates the interaction energies. ω B97X-V and ω B97M-V give the correct energetic ordering for all of the isomers and simultaneously are the best two methods for the total interaction energies themselves. These appear to be good functionals for noncovalent interactions in water.

To examine what happens in larger clusters, we consider 10 low-energy isomers of $(\text{H}_2\text{O})_{20}$, from refs 99 and 100. Error statistics are again listed in Table 2. All supersystem methods underbind all isomers, with the exception of the ω B97X-D3, M06-2X, SCAN, and SCAN0 functionals. These four functionals overestimate the interaction energies, with SCAN (MAE = 18 kcal/mol) and SCAN0 (MAE = 15 kcal/mol) exhibiting particularly large errors. (A recent study of water clusters using the SCAN functional also found that interaction energies were overestimated;⁷² nevertheless, this functional has been used recently to model liquid water.^{68–70}) The B97-D2 functional, on the other hand, exhibits very large errors in the opposite direction, underestimating the interaction energies by 23 kcal/mol, on average. This inconsistency between functionals mirrors the situation for polymorphs of ice,^{101,102} for which DFT fails to provide consistent performance and the inclusion of dispersion corrections is not always helpful.

The MP2/CBS method (MAE = 2.7 kcal/mol) performs reasonably well for $(\text{H}_2\text{O})_{20}$, as do the functionals B97M-V (MAE = 2.9 kcal/mol), ω B97X-V (1.7 kcal/mol), and ω B97M-V (0.9 kcal/mol). All other functionals considered here exhibit MAEs of >3 kcal/mol. This includes PW6B95-D3(0), whose performance is significantly worse for $(\text{H}_2\text{O})_{20}$ (MAE = 6.3 kcal/mol) than it is for $(\text{H}_2\text{O})_6$ (MAE = 0.3 kcal/mol). Although M11 and MN15 are from the new generations of Minnesota functionals, their performance for $(\text{H}_2\text{O})_{20}$ is not competitive with the best-performing functionals and is considerably worse than that of the older M06-2X functional.

Considering the XSAPT approaches, XSAPT(KS)+*ai*D3 performs better with ω_{GDD} tuning than with ω_{IP} tuning, as was also the case for $(\text{H}_2\text{O})_6$. In either case, the errors are significantly reduced by the DC term, affording MAEs of 0.9 kcal/mol (ω_{IP}) and 2.2 kcal/mol (ω_{GDD}). Of the methods examined here, the only ones that afford MAEs of <1 kcal/mol for $(\text{H}_2\text{O})_{20}$ are XSAPT(KS)+*ai*D3+DC (ω_{IP}) and ω B97M-V.

Comparing absolute errors in interaction energies between $(\text{H}_2\text{O})_6$ and $(\text{H}_2\text{O})_{20}$ is potentially deceptive, as a roughly constant error per monomer (or per hydrogen bond) might pass unnoticed in such an analysis. Therefore, we have also listed the MAEs in Table 2 in percentage terms. For many of the methods examined here, percentage errors are larger in the larger system, with M11, MN15, and PW6B95-D3(0) being egregious examples of this trend. However, the same is not true for XSAPT(KS)+*ai*D3+DC, nor for the functionals ω B97X-D3, ω B97M-V, M06-2X, SCAN, and SCAN0. (Percentage errors for ω B97X-V are also only slightly larger for the 20-mer.) Particularly promising are the ω B97X-V and ω B97M-V functionals and the XSAPT(KS)+*ai*D3+DC (ω_{IP}) method, each of which exhibits a percentage error of $\leq 1.0\%$ in both cluster sizes. These three are the only methods examined here for which this is true.

Compared to total interaction energies, predicting the relative energies of different cluster isomers appears to be an easier problem and all supersystem methods afford MAEs of ≤ 0.5 kcal/mol in $(\text{H}_2\text{O})_{20}$, except for the three Minnesota functionals, where the MAEs are ~ 1 kcal/mol. The two XSAPT(KS)+*ai*D3 methods also afford MAEs of 1.0 kcal/mol, decreasing to 0.7 kcal/mol when the DC term is added. The DC term also reduces the largest relative energy error for the XSAPT methods (in isomer 6) from 2.7 kcal/mol to 1.9 kcal/mol.

Isomer 6 is the most highly coordinated of the 10 $(\text{H}_2\text{O})_{20}$ isomers, and it has been noted that it affords the largest relative energy error for some DFT methods as well.¹⁰⁰ None of the supersystem methods tested in this study give the correct energetic ordering for all of the $(\text{H}_2\text{O})_{20}$ isomers but the best performance for relative energies comes from the three functionals B97M-V, ω B97X-V, and ω B97M-V. The ω B97X-V functional predicts the wrong relative energy ordering only for isomers 5 and 6, and the functionals B97M-V and ω B97M-V only predict the wrong order for isomers 4 and 7. The XSAPT methods also do well. For example, XSAPT(KS)+*ai*D3+DC (ω_{GDD}) exhibits a large error for just one isomer (6, as shown in the Supporting Information). Further improvement may require an even more-sophisticated treatment of dispersion, such as the coupled Kohn–Sham dispersion that is used in SAPT(DFT).^{25,26}

Lastly in this section, we consider 10 isomers of $\text{F}^-(\text{H}_2\text{O})_{10}$, where the presence of the anion increases the dispersion energy. Error statistics for total interaction energies and relative energies are listed in Table 2. Among supersystem methods, B97M-V, ω B97X-V, and ω B97M-V perform the best; these are the only supersystem methods whose MAEs are <0.5 kcal/mol. All other supersystem methods exhibit MAEs of >1 kcal/mol, and for B97-D2, M06-2X, SCAN, and SCAN0, the MAEs are >8 kcal/mol. For relative energies, most supersystem methods exhibit MAEs of 0.1–0.5 kcal/mol, the exceptions being B97-D2 (0.6 kcal/mol), MN15 (0.7 kcal/mol), and SCAN (2.3 kcal/mol).

As also seen for the $(\text{H}_2\text{O})_{20}$ clusters, the addition of the DC term to XSAPT(KS)+*ai*D3 affords a modest increase in accuracy for relative energies, bringing the results into alignment with the best supersystem approaches in this case, but affects a significant increase in accuracy for total interaction energies. Errors of 3.5 kcal/mol (ω_{GDD}) and 5.7 kcal/mol (ω_{IP}) in the latter are reduced to 1.0 and 1.3 kcal/mol, respectively, when the DC term is added. As such, our

conclusion is that the DC term is essential in XSAPT(KS)+D calculations involving either large clusters or clusters with ions.

IV.C. Clathrate Hydrate Host/Guest Complexes.

Clathrate hydrates are host/guest complexes consisting of polyhedral water frameworks (some building blocks of which are depicted in Figure 4) that encapsulate small gas molecules.

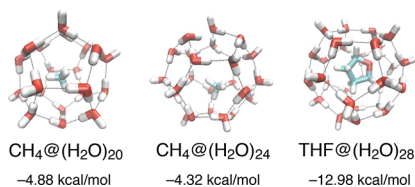


Figure 4. Lowest energy structure for each of three classes of clathrate hydrate building blocks, each containing a small guest molecule. Benchmark interaction energies were computed at the CCSD(T)/CBS level.

These frameworks self-assemble around methane at high pressure, causing problems for the natural gas industry,^{103,104} but they have also been presented as possible matrices for gaseous fuel storage.^{105–108} Indeed, they already serve as such (for the time being) in arctic permafrost and in oceanic sediments,^{103,104,108,109} and on Saturn's largest moon, Titan.¹¹⁰

The most common structures are based on cubic lattices known as type I and type II clathrates; a hexagonal cubic structure known as type H is rare and is not considered here. The type I clathrate hydrates involve two types of water cages: (H₂O)₂₀ with 12 pentagonal faces (5¹²) and (H₂O)₂₄ with 12 pentagonal and two hexagonal faces (5¹²6²). Type II clathrate hydrates are formed from either 5¹² cages or else from (H₂O)₂₈ in a 5¹²6⁴ configuration.^{103,106} The type of guest molecules inside the host cages of clathrate hydrates is dependent on the cavity size and shape. Furthermore, the lattice of the host cage in clathrate hydrates is maintained by the interactions between the host cages and guest molecules under suitable thermodynamic conditions.¹⁰³

There have been numerous theoretical studies of the interactions between host water cages and small guest molecules,^{73,111–128} using functionals including revPBE,¹²⁹ B3LYP,^{130,131} B3LYP-D2,^{15,59,130,131} M05-2X,¹³² M06-L,¹³² M06-HF,¹³² M06-2X,¹³² B97-D2,^{15,58,59} ωB97X-D,¹³³ and vdW-DF.¹³⁴ Small-basis Hartree–Fock and MP2 calculations have also been reported for these systems.^{119,120,124,135–138}

Despite this considerable computational effort, there has been no attempt to benchmark the computed interaction energies with respect to high-quality *ab initio* calculations. Instead, interactions in CH₄@(H₂O)₂₀ have been benchmarked with respect to MP2 calculations using double- and triple-ζ basis sets (only) to extrapolate the CBS limit,^{73,139} which is a procedure that is known to be inadequate.¹⁴⁰ These studies resulted in recommendations^{73,120,124,141} to use B97-D2, ωB97X-D, M06-2X, PBE-D2, and CAM-B3LYP, and these functionals were subsequently used to study various combinations of water cages and guest molecules.^{127,141–144} BLYP-D3 was also suggested for use in clathrate hydrates,¹⁴⁵ based on comparison to CCSD/CBS benchmarks for a variety of guest molecules in a 5¹² cage. Finally, quantum Monte Carlo (QMC) benchmarks are available for CH₄ in the 5¹² cage,¹⁴⁶ but not for other species. There is a need for high-level benchmarks for a wider variety of clathrate hydrates.

The water cages 5¹² and 5¹²6² are used here as methane hosts in various MP2, DFT, and XSAPT calculations, where they are compared to CCSD(T)/CBS benchmarks reported here for the first time. We also consider tetrahydrofuran (THF) in 5¹²6⁴, since hydrogen storage has been achieved under industrially attainable conditions using this framework, where THF stabilizes the type II clathrate hydrate.¹⁴⁷ The water cage 5¹²6⁴ is used here as THF host to study the interactions between them. Geometries of CH₄@(H₂O)₂₀, CH₄@(H₂O)₂₄, and THF@(H₂O)₂₈, optimized at the B97M-V/6-31+G(d) level, are shown in Figure 4.

Error statistics for total interaction energies using various methods are shown in Table 3. The CCSD(T)/CBS

Table 3. Mean Absolute Errors in Interaction Energies for Three Clathrate Hydrate Host/Guest Complexes^a

method	MAE	
	(kcal/mol)	(%)
Supersystem Methods		
MP2/CBS	0.19	2.6
ωB97X-V ^b	1.97	29.0
ωB97M-V ^b	1.32	19.2
B97M-V ^b	0.92	13.3
ωB97X-D3 ^b	2.82	44.6
M11 ^b	1.85	24.5
MN15 ^b	3.59	61.4
PW6B95-D3(0) ^b	1.57	25.2
SCAN ^b	1.14	14.0
SCAN0 ^b	2.35	31.6
Functionals Recommended Previously		
revPBE ^b	12.79	187.8
B3LYP ^b	9.79	143.7
B3LYP-D2 ^b	0.72	9.9
PBE-D2 ^b	0.77	12.0
B97-D2 ^b	0.10	2.0
BLYP-D3 ^b	2.15	32.2
M05-2X ^b	0.64	11.7
M06-2X ^b	1.12	15.7
ωB97X-D ^b	1.33	21.3
XSAPT(KS)+aiD2 Methods		
XSAPT(KS)+aiD2(ω _{IP}) ^c	0.55	8.1
XSAPT(KS)+aiD2+DC(ω _{IP}) ^{c,d}	0.26	4.0
XSAPT(KS)+aiD2(ω _{GDD}) ^c	0.49	7.0
XSAPT(KS)+aiD2+DC(ω _{GDD}) ^{c,d}	0.24	3.7
XSAPT(KS)+aiD3 Methods		
XSAPT(KS)+aiD3(ω _{IP}) ^c	0.78	14.6
XSAPT(KS)+aiD3+DC(ω _{IP}) ^{c,d}	1.07	18.7
XSAPT(KS)+aiD3(ω _{GDD}) ^c	0.84	15.7
XSAPT(KS)+aiD3+DC(ω _{GDD}) ^{c,d}	1.14	19.8

^aError is defined with respect to CCSD(T)/CBS benchmarks. ^bUsing def2-TZVPPD with counterpoise correction. ^cUsing hpTZVPP for XSAPT(KS). ^dUsing hpTZVPP for the nonadditive dispersion correction, $E_{\text{DC}}(\text{full})$.

benchmarks, reported here for the first time, are computed by adding a correction $\delta_{\text{CCSD(T)}}$ (eq 7) to the MP2/CBS interaction energy. The $\delta_{\text{CCSD(T)}}$ correction is evaluated at the DPLNO–CCSD(T)/def2-TZVP level, using a tight PNO truncation threshold.¹⁴⁸ This approximation has been shown to approach the canonical CCSD(T) result within 1 kJ/mol.¹⁴⁹ We find that $\delta_{\text{CCSD(T)}} = 0.1\text{--}0.3$ kcal/mol for these species,

and is destabilizing. This correction is small enough that MP2/CBS benchmarks are probably adequate for these systems.

High-level *ab initio* studies of the $\text{CH}_4@(\text{H}_2\text{O})_{20}$ isomer considered here have also been reported by Jordan and co-workers,¹⁴⁶ and although the geometry in that work was obtained at a different level of theory, it is probably still safe to make some comparisons. In ref 146, a MP2/CBS interaction energy of -5.04 kcal/mol (almost identical to the value determined here) was obtained from extrapolation of aDZ, aTZ, and aQZ results. The $\delta_{\text{CCSD(T)}}$ correction reported in ref 146 was estimated using a many-body expansion truncated at the two- or three-body level, and taking the difference between CCSD(T)-F12b¹⁵⁰ and MP2-F12¹⁵¹ interaction energies computed using the VTZ-F12¹⁵² basis set to define $\delta_{\text{CCSD(T)}}$. The two-body and three-body estimates are $\delta_{\text{CCSD(T)}} \approx -0.90$ kcal/mol and -0.14 kcal/mol, respectively. These corrections are slightly attractive, but neglect four-body and higher-order terms that are likely *not* completely negligible.^{153–155} In contrast, our $\delta_{\text{CCSD(T)}}$ corrections are slightly repulsive and were computed via a CCSD(T) calculation on the full system, albeit within the DLPNO approximation. Our CCSD(T)/CBS interaction energy of -4.88 kcal/mol agrees with the QMC result (-5.3 ± 0.5 kcal/mol) to within the statistical uncertainty of the latter, suggesting that $\delta_{\text{CCSD(T)}}$ is indeed slightly repulsive or at least not attractive. This agreement also suggests that our benchmarks for these three host/guest complexes are reliable. To the best of our knowledge, these are the first CCSD(T)/CBS benchmarks for host/guest complexes involving water clusters of this size. That being said, the fact that $|\delta_{\text{CCSD(T)}}| \lesssim 0.3$ kcal/mol, according to our calculations, implies that MP2/CBS is a reliable method to study interactions between water cages and guest molecules, and for the three clathrate hydrates considered here, the MP2/CBS interaction energies differ from CCSD(T)/CBS results by an average of only 0.2 kcal/mol, or 2.6%.

Considering some of the previously recommended density functionals, we find (not surprisingly) that the guest molecules are unbound when revPBE and B3LYP are applied to the clathrate hydrates, although they become bound upon the addition of Grimme's D2 correction.¹⁵ The MAEs for B3LYP-D2 and PBE-D2 are 0.7 and 0.8 kcal/mol, respectively. However, if the D2 correction is replaced by D3, which is generally superior to D2,¹¹ the mean errors actually increase, to 2.2 kcal/mol (BLYP-D3), 2.8 kcal/mol (ω B97X-D3), and 1.6 kcal/mol [PW6B95-D3(0)]. The three functionals recommended by Liu et al.,⁷³ namely, ω B97X-D, M06-2X, and B97-D2, exhibit MAEs of 1.3, 1.1, and 0.1 kcal/mol, respectively. Of all of these functionals, therefore, only B97-D2 can be considered reliable, with errors averaging just 2%, and indeed this functional has been used in several previous studies of clathrate hydrates.^{73,120,124,127}

On the other hand, the recommendation from ref 73 is to use B97-D2 in conjunction with the 6-311++G(2d,2p) basis set and *without* counterpoise correction, whereas all of the DFT calculations reported in Table 3 use def2-TZVPPD *with* counterpoise correction. Therefore, in Table 4, we compare B97-D2 results using either basis set, both with and without counterpoise correction. The recommended procedure from ref 73 affords a MAE of 1.7 kcal/mol, or 21.7%. Use of def2-TZVPPD reduces the error to 0.4 kcal/mol (5.2%), even without counterpoise correction, and with counterpoise correction the MAE is reduced to <0.2 kcal/mol for either basis set. In view of these data, the recommendation from ref

Table 4. Mean Absolute Errors in B97-D2 Interaction Energies for the Clathrate Hydrates in Figure 4^a

basis set	CP? ^b	MAE	
		(kcal/mol)	(%)
6-311++G(2d,2p)	no	1.69	21.7
6-311++G(2d,2p)	yes	0.17	3.3
def2-TZVPPD	no	0.42	5.2
def2-TZVPPD	yes	0.10	2.0

^aError is defined with respect to CCSD(T)/CBS benchmarks. ^bCP = counterpoise correction.

73 cannot be taken seriously and indeed one should be skeptical, in our opinion, of any recommendation to avoid counterpoise correction, as this implies that good results are balanced on an error cancellation whose robustness may be questionable. Concerns have occasionally been raised that the counterpoise procedure may overcorrect for BSSE but these concerns have been shown to be largely unfounded.¹⁵⁶ Cases where the counterpoise-corrected result lies further from the true intermolecular interaction energy as compared to the uncorrected result are usually a sign that the basis set employed is not of benchmark quality.^{156–158} Basis-set incompleteness should not be conflated with BSSE. In this context, it is apropos to note that XSAPT calculations are inherently free of BSSE.

It is interesting that MAEs for the clathrate hydrates are larger for the newer generation of Minnesota functionals, as compared to many of the other functionals considered here. Among M05-2X, M06-2X, M11, and MN15, the smallest MAE (0.6 kcal/mol) is obtained using M05-2X and the largest (3.6 kcal/mol) is obtained using MN15. (Similarly, it was noted above that the older ω B97X-D functional performs better than the newer ω B97X-D3 that was reparameterized in an effort to obtain a better dispersion correction.) The recently developed SCAN and SCAN0 functionals give MAEs of 1.1 and 2.4 kcal/mol, respectively. In principle, the “right way” to include dispersion in DFT is using a nonlocal correlation functional, but the three functionals that contain nonlocal correlation (ω B97M-V, ω B97X-V, and B97M-V) exhibit MAEs ranging from 0.9–2.0 kcal/mol and do not perform better than some of the older dispersion-corrected functionals including M05-2X, B97-D2, and B3LYP-D2. The performance of ω B97M-V and ω B97X-V for these clathrate hydrates is especially disappointing, in view of the favorable performances of these functionals for the other systems considered here, and for noncovalent interactions more generally.^{56,86} These results perhaps provide a target for further improvement of these functionals. Inversely, B97-D2 performs well for these clathrate hydrates but very poorly for the other systems examined in this work, suggesting significant error cancellation in the case of the clathrate hydrates.

Turning to the XSAPT methods, we first of all note that, for these clathrate hydrates, the computational cost of XSAPT is considerably lower than that of supersystem DFT, since self-consistent field iterations are required only on the (very small) monomers, not the entire cluster, and subsequent SAPT calculations are required only on dimers (H_2O)₂, $\text{CH}_4(\text{H}_2\text{O})$, and $\text{THF}(\text{H}_2\text{O})$. XSAPT(KS)+*ai*D3 affords MAEs of 0.8 kcal/mol that are actually slightly degraded (to 1.1 kcal/mol) via addition of the DC term. The $E_{\text{DC}}^{\text{(partial)}}$ correction, using ω_{IP} tuning, contributes 0.07, 0.04, and 0.18 kcal/mol for $\text{CH}_4@(\text{H}_2\text{O})_{20}$, $\text{CH}_4@(\text{H}_2\text{O})_{24}$, and $\text{THF}@(\text{H}_2\text{O})_{28}$, respec-

Table 5. Energy Decomposition Analysis at the XSAPT(KS)+*ai*D2+DC Level^a

species	Energy Component (kcal/mol)					
	electrostatic	exchange	Induction			dispersion
			total	MB	CT ^b	
CH ₄ @(H ₂ O) ₂₀	−2.87	8.92	−1.09	0.02	−0.10	−9.86
CH ₄ @(H ₂ O) ₂₄	−1.13	3.69	−0.37	0.32	−0.06	−6.87
THF@(H ₂ O) ₂₀	−5.65	11.52	−2.04	0.35	−0.19	−17.20
						−13.38

^a ω_{IP} tuning, hpTZVPP basis set. ^bUsing LRC- ω PBE with ω_{IP} tuning, hpTZVPP basis set.

tively, and the corresponding values of $E_{\text{DC}(\text{full})}$ are 0.23, 0.17, and 0.47 kcal/mol. Given that $E_{\text{DC}(\text{full})}$ is ~ 3 times larger than $E_{\text{DC}(\text{partial})}$, we will use the former for calculations on the clathrate hydrates. XSAPT(KS)+*ai*D2 works surprisingly well for these systems, with a MAE of 0.5 kcal/mol, which improves to 0.3 kcal/mol (4% error) upon the addition of $E_{\text{DC}(\text{full})}$. The choice of ω_{GDD} versus ω_{IP} matters little. We note in this case that the second-generation *ai*D2 potential performs slightly better as compared to *ai*D3.

Most of the supersystem methods predict that encapsulation of methane is more favorable in the (H₂O)₂₀ cage than it is in (H₂O)₂₄, in agreement with the CCSD(T)/CBS prediction. (The revPBE and B3LYP functionals are exceptions, but these have no dispersion corrections and therefore should not be trusted.) The XSAPT(KS)+*ai*D2 method predicts the same, as do the M05-2X and B97-D2 functionals,¹²³ but XSAPT(KS)+*ai*D3 predicts the opposite. As such, only XSAPT(KS)+*ai*D2+DC (with either ω_{IP} or ω_{GDD} tuning), M05-2X, and B97-D2 afford acceptable results for the clathrate hydrates, although the very poor performance of these two functionals for the other systems considered here makes it difficult to recommend them, despite their (likely coincidentally) good performance for clathrate hydrates.

Finally, we report results from a many-body energy decomposition analysis³⁰ of the clathrate hydrates, evaluated at the XSAPT(KS)+*ai*D2+DC (ω_{IP}) level and using constrained DFT (cDFT) calculations on dimers to extract a well-defined charge-transfer (CT) interaction energy from the SAPT induction energy, as described in our previous work.¹⁵⁹ The XPol procedure can also be exploited to obtain the many-body (MB) contribution to the induction energy.³⁰ Results are presented in Table 5. Previous calculations on clathrate hydrates revealed that the electrostatic potential inside the water cage is relatively uniform and small, leading to a suggestion that van der Waals (i.e., dispersion) interactions should dominate the host/guest interaction,¹⁶⁰ but our calculations suggest that the picture is not quite so simple. The CT and MB-polarization contributions are rather small for all three clathrate hydrates, but although dispersion is the single largest energy component in all three cases, the exchange energy is certainly not negligible in comparison, and the electrostatic energy ranges from −1.1 kcal/mol to −5.6 kcal/mol. For the two methane hydrates, electrostatics and induction are similar but dispersion and exchange are larger in CH₄@(H₂O)₂₀ than in CH₄@(H₂O)₂₄, which may reflect the fact that CH₄ inclines to one side in (H₂O)₂₄ but is centered in (H₂O)₂₀. This is somewhat evident from the structures shown in Figure 4, and the average C–O distance between CH₄ and H₂O is 3.87 Å in CH₄@(H₂O)₂₀ and 4.29 Å in CH₄@(H₂O)₂₄. The significantly smaller C–O distance in CH₄@(H₂O)₂₀ explains the larger value of the Pauli repulsion (exchange) energy and likely also explains the larger dispersion

energy, despite the larger number of electrons in CH₄@(H₂O)₂₄.

The host/guest interaction is larger for the polar THF molecule, and all energy components are larger in THF@(H₂O)₂₈ as compared to the two methane hydrates. Larger induction and electrostatic interactions likely reflect hydrogen bonding, while exchange and dispersion reflect the larger size of the guest molecule. In the end, the increase in attractive dispersion and electrostatic interactions more than compensates for the larger exchange energy and the interaction energy of THF in (H₂O)₂₈ is significantly larger than that of CH₄ in either (H₂O)₂₀ or (H₂O)₂₄. The large interaction of THF with water cages is known and has been used to stabilize type II clathrate hydrates.¹⁴⁷ It is possible that a theoretical exploration of possible guest molecules, combined with energy decomposition analysis, could aid in the search for hydrates that can achieve gas storage under industrially useful conditions.

V. SUMMARY

A dispersion correction (DC) has been introduced to approximate nonpairwise-additive dispersion effects in XSAPT(KS)+*ai*D calculations in a first-principles way. These effects arise from the “squeezing” of atoms by the molecular environment, which modifies atomic contributions to dispersion relative to those arising from isolated atoms. Unlike empirical three-body (i.e., triatomic) dispersion corrections that have been employed in dispersion-corrected DFT^{11,161} and dispersion-corrected XSAPT,³² the correction introduced here contains no empirical parameters. It simply considers the difference in second-order SAPT dispersion energies computed using XSAPT with charge embedding versus a traditional SAPT calculation without embedding. This difference can be approximated accurately using basis sets as small as 6-31+G(d), even if the exchange-dispersion term is neglected, leading to an affordable correction whose cost scales as the fourth power of fragment size, $O(N_f^4)$, which represents an increase over the $O(N_f^3)$ scaling of an ordinary XSAPT(KS)+*ai*D, but whose cost is mitigated by the ability to use a small basis set for the fourth-order part of the calculation.

The DC term introduced here encodes changes in dispersion that are brought about by polarization and affords no nonadditive dispersion correction in the absence of polarization. As such, the new approach will not afford any nonadditive correction in, e.g., clusters of noble gas atoms, since the constituent fragments will have neither a charge nor a dipole moment with which to polarize one another. That being said, there are numerous important applications involving large numbers of polar monomers, and we have focused on ion–water clusters and clathrate hydrate host/guest complexes in the present work.

The DC term proves to be important in clusters of polar molecules and amounts to ~ 3 kcal/mol for isomers of $\text{SO}_4^{2-}(\text{H}_2\text{O})_6$, ~ 4 kcal/mol for $\text{F}^-(\text{H}_2\text{O})_{10}$, and ~ 5 kcal/mol for $(\text{H}_2\text{O})_{20}$. For systems such as these, the nonpairwise dispersion correction to XSAPT(KS)+*aiD* is essential if high accuracy is desired. As compared to CCSD(T)/CBS benchmarks, the XSAPT(KS)+*aiD*3+DC method with ω_{IP} tuning of the monomer Kohn–Sham potentials affords MAEs of ~ 1 kcal/mol for $\text{X}(\text{H}_2\text{O})_n$ clusters with $n \leq 6$, for $\text{X} = \text{F}^-$, Cl^- , SO_4^{2-} , Li^+ , Na^+ , and K^+ .

We have also tested a selection of density functionals for these systems, focusing on functionals that have been specifically recommended to describe noncovalent interactions. Of the functionals tested, only $\omega\text{B97X-V}$ and $\omega\text{B97M-V}$ can be recommended for use with ion–water clusters. MAEs for our $\text{X}(\text{H}_2\text{O})_n$, $\text{F}^-(\text{H}_2\text{O})_{10}$, and $(\text{H}_2\text{O})_{20}$ datasets are 0.5, 0.2, and 1.7 kcal/mol, respectively, for $\omega\text{B97X-V}$, and 0.5, 0.5, and 0.9 kcal/mol, respectively, for $\omega\text{B97M-V}$. The outstanding performance of these two functionals is consistent with their generally good performance for noncovalent interactions,⁸⁶ but the lackluster performance of the Minnesota functionals M06-2X, M11, and MN15 is somewhat surprising. For the latter functionals, MAEs for $(\text{H}_2\text{O})_{20}$ isomers range from 3 kcal/mol to 10 kcal/mol and those for $\text{F}^-(\text{H}_2\text{O})_{10}$ isomers range from 2 kcal/mol to 8 kcal/mol.

The $\omega\text{B97X-V}$ and $\omega\text{B97M-V}$ functionals exhibit somewhat larger errors of ~ 2 kcal/mol and ~ 1.3 kcal/mol, respectively, for clathrate hydrate complexes such as $\text{CH}_4@(\text{H}_2\text{O})_{20}$, $\text{CH}_4@(\text{H}_2\text{O})_{24}$, and $\text{THF}@(\text{H}_2\text{O})_{28}$. Nevertheless, these are the functionals that we still recommend for these systems because the few functionals that afford better error statistics for clathrate hydrates (e.g., B97-D2, with a MAE of just 0.1 kcal/mol) exhibit unacceptably poor performance for water clusters and ion–water clusters, which calls into question just how robust the MAE for clathrate hydrates really is. This work suggests that clathrate hydrates, with a myriad of hydrogen bonds but where dispersion still plays an important role in the binding of small molecules, may provide stringent tests of methods for predicting noncovalent interaction energies. The XSAPT(KS)+*aiD*2+DC and XSAPT(KS)+*aiD*3+DC approaches with ω_{IP} tuning afford MAEs of 0.3 and 1.1 kcal/mol, respectively, for the clathrate hydrates.

Finally, not just total interaction energies but also energy components for SAPT(KS) calculations have been tested against SAPT2+(3)/aTZ benchmarks, and we used these tests to evaluate the performance of various density functionals for use in SAPT(KS) calculations. For the S22 and S66 datasets of noncovalent dimers, all of the functionals tested afford similar electrostatic and induction energies in SAPT(KS) calculations, but tuned LRC- ωPBE affords the most accurate exchange energies. This is the functional that we have used in SAPT(KS) calculations for some time,^{20,27,30,32} and we have previously demonstrated that tuning the range-separation parameter is essential to its success.²⁷

XSAPT(KS)+*aiD*3 with tuned LRC- ωPBE also provides excellent accuracy for the S22 and S66 datasets, where its accuracy is comparable to CCSD(T)/CBS benchmarks for total interaction energies and comparable to SAPT2+(3)/aTZ benchmarks for energy components. Notably, XSAPT(KS)+*aiD*3 calculations can be performed at a fraction of the cost of supersystem DFT calculations, especially for systems such as $\text{THF}@(\text{H}_2\text{O})_{28}$ where the monomers are small. The cost reduction is even more dramatic considering that the DFT

calculations must be counterpoise-corrected, which for $\text{THF}@(\text{H}_2\text{O})_{28}$, means 30 separate DFT calculations in the full system basis set. Traditionally, the SAPT0/jun-cc-pVDZ procedure has been considered to be the most cost-effective form of SAPT, where the basis set is a partially augmented double- ζ one that is recommended, because it affords good error cancellation in the dispersion term,⁸² although this cancellation appears to work less well in large dispersion-dominated systems.³² Here, we find that the error cancellation in total interaction energies does not percolate into the energy components, which exhibit large errors as compared to SAPT2+(3)/aTZ.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b00527.

Additional details of the calculations (PDF)

Coordinates of structures (TXT)

Raw data (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: herbert@chemistry.ohio-state.edu.

ORCID

Ka Un Lao: 0000-0002-3993-536X

John M. Herbert: 0000-0002-1663-2278

Present Address

[†]Department of Chemistry & Chemical Biology, Cornell University, Ithaca, NY 14853, USA.

Notes

The authors declare the following competing financial interest(s): J.M.H. serves on the Board of Directors of Q-Chem, Inc.

■ ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, under Award No. DE-SC0008550. Calculations were performed at the Ohio Supercomputer Center, under Project No. PAA-0003.¹⁶² K.U.L. acknowledges support from a Presidential Fellowship awarded by The Ohio State University. J.M.H. is a fellow of the Alexander von Humboldt Foundation.

■ REFERENCES

- (1) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Challenges for density functional theory. *Chem. Rev.* **2012**, *112*, 289–320.
- (2) Pribram-Jones, A.; Gross, D. A.; Burke, K. DFT: A theory full of holes? *Annu. Rev. Phys. Chem.* **2015**, *66*, 283–304.
- (3) Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.
- (4) Klimeš, J.; Michaelides, A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *J. Chem. Phys.* **2012**, *137*, 120901.
- (5) Reilly, A. M.; Tkatchenko, A. Van der Waals dispersion interactions in molecular materials: Beyond pairwise additivity. *Chem. Sci.* **2015**, *6*, 3289–3301.
- (6) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

- (7) Chu, X.; Dalgarno, A. Linear response time-dependent density functional theory for van der Waals coefficients. *J. Chem. Phys.* **2004**, *121*, 4083–4088.
- (8) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction revisited. *J. Chem. Phys.* **2007**, *127*, 154108.
- (9) Sato, T.; Nakai, H. Density functional method including weak interactions: Dispersion coefficients based on the local response approximation. *J. Chem. Phys.* **2009**, *131*, 224104.
- (10) 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.
- (11) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *ab initio* parameterization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (12) Wu, Q.; Yang, W. Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.* **2002**, *116*, 515–524.
- (13) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
- (14) Neumann, M. A.; Perrin, M.-A. Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction. *J. Phys. Chem. B* **2005**, *109*, 15531–15541.
- (15) Grimme, S. Semiempirical GGA-type density functional constructed with long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (16) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with *ab initio* quantum mechanics calculations. *J. Comput. Chem.* **2007**, *28*, 555–569.
- (17) Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. Dispersionless density functional theory. *Phys. Rev. Lett.* **2009**, *103*, 263201.
- (18) Podeszwa, R.; Pernal, K.; Patkowski, K.; Szalewicz, K. Extension of the Hartree-Fock plus dispersion method by first-order correlation effects. *J. Phys. Chem. Lett.* **2010**, *1*, 550–555.
- (19) Heßelmann, A. Comparison of intermolecular interaction energies from SAPT and DFT including empirical dispersion contributions. *J. Phys. Chem. A* **2011**, *115*, 11321–11330.
- (20) Lao, K. U.; Herbert, J. M. Accurate and efficient quantum chemistry calculations of non-covalent interactions in many-body systems: The XSAPT family of methods. *J. Phys. Chem. A* **2015**, *119*, 235–253.
- (21) Dobson, J. F. Beyond pairwise additivity in London dispersion interactions. *Int. J. Quantum Chem.* **2014**, *114*, 1157–1161.
- (22) Hermann, J.; DiStasio, R. A., Jr.; Tkatchenko, A. First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications. *Chem. Rev.* **2017**, *117*, 4714–4758.
- (23) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction. *J. Chem. Phys.* **2005**, *122*, 154104.
- (24) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.
- (25) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254–272.
- (26) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2014**, *4*, 127–144.
- (27) Lao, K. U.; Herbert, J. M. Symmetry-adapted perturbation theory with Kohn-Sham orbitals using non-empirically tuned, long-range-corrected density functionals. *J. Chem. Phys.* **2014**, *140*, 044108.
- (28) Hapka, M.; Rajchel, L.; Modrzejewski, M.; Chalaśiński, G.; Szczęśniak, M. M. Tuned range-separated hybrid functionals in the symmetry-adapted perturbation theory. *J. Chem. Phys.* **2014**, *141*, 134120.
- (29) Herbert, J. M.; Jacobson, L. D.; Lao, K. U.; Rohrdanz, M. A. Rapid computation of intermolecular interactions in molecular and ionic clusters: Self-consistent polarization plus symmetry-adapted perturbation theory. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7679–7699.
- (30) Lao, K. U.; Herbert, J. M. An improved treatment of empirical dispersion and a many-body energy decomposition scheme for the explicit polarization plus symmetry-adapted perturbation theory (XSAPT) method. *J. Chem. Phys.* **2013**, *139*, 034107; Erratum. *J. Chem. Phys.* **2014**, *140*, 119901.
- (31) Lao, K. U.; Herbert, J. M. Accurate intermolecular interactions at dramatically reduced cost: XPol+SAPT with empirical dispersion. *J. Phys. Chem. Lett.* **2012**, *3*, 3241–3248.
- (32) Lao, K. U.; Herbert, J. M. Atomic orbital implementation of extended symmetry-adapted perturbation theory (XSAPT) and benchmark calculations for large supramolecular complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955–2978.
- (33) Our first-generation *aiD* model (ref 31) was fit to reproduce total interaction energies rather than dispersion energies. While it does afford accurate interaction energies for small-molecule dimers (e.g., for the S22 dataset), it benefits from a cancellation of errors and this model's individual energy components are less reliable. The first generation version is therefore not recommended.
- (34) Jacobson, L. D.; Herbert, J. M. An efficient, fragment-based electronic structure method for molecular systems: Self-consistent polarization with perturbative two-body exchange and dispersion. *J. Chem. Phys.* **2011**, *134*, 094118.
- (35) Jacobson, L. D.; Richard, R. M.; Lao, K. U.; Herbert, J. M. Efficient monomer-based quantum chemistry methods for molecular and ionic clusters. *Annu. Rep. Comput. Chem.* **2013**, *9*, 25–56.
- (36) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. The variational explicit polarization potential and analytical first derivative of energy: Towards a next generation force field. *J. Chem. Phys.* **2008**, *128*, 234108.
- (37) Axilrod, B. M.; Teller, E. Interaction of the van der Waals type between three atoms. *J. Chem. Phys.* **1943**, *11*, 299–301.
- (38) Muto, Y. Forces between nonpolar molecules. *J. Phys.-Math. Soc. Jpn.* **1943**, *17*, 629.
- (39) Gao, J.; Truhlar, D. G.; Wang, Y.; Mazack, M. J. M.; Löffler, P.; Provorse, M. R.; Rehak, P. Explicit polarization: A quantum mechanical framework for developing next generation force fields. *Acc. Chem. Res.* **2014**, *47*, 2837–2845.
- (40) Wang, Y.; Mazack, M. J. M.; Truhlar, D. G.; Gao, J. Explicit polarization theory. In *Many-Body Effects and Electrostatics in Biomolecules*; Cui, Q., Meuwly, M., Ren, P., Eds.; CRC Press/Taylor and Francis: Boca Raton, FL, 2016; pp 33–64.
- (41) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jezierski, B. On the effectiveness of monomer-, dimer-, and bond-centered basis functions in calculations of intermolecular interaction energies. *J. Chem. Phys.* **1995**, *103*, 7374–7391.
- (42) Řezáč, J.; Hobza, P. Extrapolation and scaling of the DFT-SAPT interaction energies toward the basis set limit. *J. Chem. Theory Comput.* **2011**, *7*, 685–689.
- (43) Huang, Y.; Shao, Y.; Beran, G. J. O. Accelerating MP2C dispersion corrections for dimers and molecular crystals. *J. Chem. Phys.* **2013**, *138*, 224112.
- (44) Lao, K. U.; Herbert, J. M. Breakdown of the single-exchange approximation in third-order symmetry-adapted perturbation theory. *J. Phys. Chem. A* **2012**, *116*, 3042–3047.
- (45) Lao, K. U.; Schäffer, 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.
- (46) Jurečka, P.; Šponer, J.; Černý, 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.
- (47) Řezáč, 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; Erratum. *J. Chem. Theory Comput.* **2014**, *10*, 1359–1360.

- (48) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.* **2009**, *130*, 054112.
- (49) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.
- (50) Modrzejewski, M.; Rajchel, L.; Chalasinski, G.; Szczesniak, M. M. Density-dependent onset of the long-range exchange: A key to donor-acceptor properties. *J. Phys. Chem. A* **2013**, *117*, 11580–11586.
- (51) Pan, C.-R.; Fang, P.-T.; Chai, J.-D. Asymptotic correction schemes for semilocal exchange-correlation functionals. *Phys. Rev. A: At., Mol., Opt. Phys.* **2013**, *87*, 052510.
- (52) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking density functional methods against the S66 and S66 × 8 datasets for non-covalent interactions. *ChemPhysChem* **2011**, *12*, 3421–3433.
- (53) Goerigk, L.; Grimme, S. Efficient and accurate double-hybrid-meta-GGA density functionals—evaluation with the extended GMTKN30 database for general main group thermochemistry, kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
- (54) Mardirossian, N.; Lambrecht, D. S.; McCaslin, L.; Xantheas, S. S.; Head-Gordon, M. The performance of density functionals for sulfate-water clusters. *J. Chem. Theory Comput.* **2013**, *9*, 1368–1380.
- (55) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272.
- (56) Mardirossian, N.; Head-Gordon, M. ω B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.
- (57) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.
- (58) Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (59) It seems to have become customary in the literature to use simply “-D” to mean the correction formerly known as “-D2”, i.e., the empirical dispersion correction from ref 15. We find this convention to be ambiguous so we continue to use “-D2” in this work.
- (60) Zhao, Y.; Truhlar, D. G. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
- (61) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent 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; Erratum. *Theor. Chem. Acc.* **2008**, *119*, 525.
- (62) Peverati, R.; Truhlar, D. G. Improving the accuracy of hybrid meta-GGA density functionals by range separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.
- (63) Yu, H. S.; 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.
- (64) Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
- (65) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.
- (66) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (67) Hui, K.; Chai, J.-D. SCAN-based hybrid and double-hybrid density functionals from models without fitted parameters. *J. Chem. Phys.* **2016**, *144*, 044114.
- (68) Chen, M.; Ko, H.-Y.; Remsing, R. C.; Calegari Andrade, M. F.; Santra, B.; Sun, Z.; Selloni, A.; Car, R.; Klein, M. L.; Perdew, J. P.; Wu, X. Ab initio theory and modeling of water. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 10846–10851.
- (69) Wiktor, J.; Ambrosio, F.; Pasquarello, A. Note: Assessment of the SCAN+rVV10 functional for the structure of liquid water. *J. Chem. Phys.* **2017**, *147*, 216101.
- (70) Zheng, L.; Chen, M.; Sun, Z.; Ko, H.-Y.; Santra, B.; Dhruv, P.; Wu, X. Structural, electronic, and dynamical properties of liquid water by *ab initio* molecular dynamics based on SCAN functional within the canonical ensemble. *J. Chem. Phys.* **2018**, *148*, 164505.
- (71) Yao, Y.; Kanai, Y. Free energy profile of NaCl in water: First-principles molecular dynamics with SCAN and ω B97X-V exchange-correlation functionals. *J. Chem. Theory Comput.* **2018**, *14*, 884–893.
- (72) Brandenburg, J. G.; Bates, J. E.; Sun, J.; Perdew, J. P. Benchmark tests of a strongly constrained semilocal functional with a long-range dispersion correction. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2016**, *94*, 115144.
- (73) Liu, Y.; Zhao, J.; Li, F.; Chen, Z. Appropriate description of intermolecular interactions in the methane hydrates: An assessment of DFT methods. *J. Comput. Chem.* **2013**, *34*, 121–131.
- (74) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (75) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*, 134105.
- (76) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (77) Sparta, M.; Neese, F. Chemical applications carried out by local pair natural orbital based coupled-cluster methods. *Chem. Soc. Rev.* **2014**, *43*, 5032–5041.
- (78) Turney, J. M.; et al. Psi4: An open-source *ab initio* electronic structure program. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556–565.
- (79) Shao, Y.; et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.
- (80) Lange, A. W.; Herbert, J. M. Both intra- and interstrand charge-transfer excited states in B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states. *J. Am. Chem. Soc.* **2009**, *131*, 3913–3922.
- (81) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F., III Large-scale symmetry-adapted perturbation theory computations via density fitting and Laplace transformation techniques: Investigating the fundamental forces of DNA-intercalator interactions. *J. Chem. Phys.* **2011**, *135*, 174107.
- (82) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, *140*, 094106.
- (83) Hohenstein, E. G.; Sherrill, C. D. Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of π - π interactions in linear acenes. *J. Chem. Phys.* **2010**, *132*, 184111.
- (84) Sutton, C.; Marshall, M. S.; Sherrill, C. D.; Risko, C.; Brédas, J.-L. Rubrene: The interplay between intramolecular and intermolecular interactions determines the planarization of its tetracene core in the solid state. *J. Am. Chem. Soc.* **2015**, *137*, 8775–8782.
- (85) 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.
- (86) 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.

(87) 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.

(88) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(89) Hermann, J.; Tkatchenko, A. Electronic exchange and correlation in van der Waals systems: Balancing semilocal and nonlocal energy contributions. *J. Chem. Theory Comput.* **2018**, *14*, 1361–1369.

(90) Karu, K.; Mišin, M.; Ers, H.; Sun, J.; Ivaniššev, V. Performance of SCAN density functional for a set of ionic liquid ion pairs. *Int. J. Quantum Chem.* **2018**, *118*, e25582.

(91) Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. - Eur. J.* **2012**, *18*, 9955–9964.

(92) Sure, R.; Antony, J.; Grimme, S. Blind prediction of binding affinities for charged supramolecular host-guest systems: Achievements and shortcomings of DFT-D3. *J. Phys. Chem. B* **2014**, *118*, 3431–3440.

(93) Sure, R.; Grimme, S. Comprehensive benchmark of association (free) energies of realistic host-guest complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785–3801; Erratum. *J. Chem. Theory Comput.* **2015**, *11*, 5990.

(94) Antony, J.; Sure, R.; Grimme, S. Using dispersion-corrected density functional theory to understand supramolecular binding thermodynamics. *Chem. Commun.* **2015**, *51*, 1764–1774.

(95) Sure, R.; Grimme, S. Halogen bonded supramolecular capsules: A challenging test case for quantum chemical methods. *Chem. Commun.* **2016**, *52*, 9893–9896.

(96) Sure, R.; Hansen, A.; Schwerdtfeger, P.; Grimme, S. Comprehensive theoretical study of all 1812 C₆₀ isomers. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14296–14305.

(97) Anacker, T.; Friedrich, J. New accurate benchmark energies for large water clusters: DFT is better than expected. *J. Comput. Chem.* **2014**, *35*, 634–643.

(98) Chen, Y.; Li, H. Intermolecular interaction in water hexamer. *J. Phys. Chem. A* **2010**, *114*, 11719–11724.

(99) Kazimirski, J. K.; Buch, V. Search for low energy structures of water clusters (H₂O)_n, *n* = 20–22, 48, 123, and 293. *J. Phys. Chem. A* **2003**, *107*, 9762–9775.

(100) Wang, K.; Li, W.; Li, S. Generalized energy-based fragmentation CCSD(T)-F12a method and application to the relative energies of water clusters (H₂O)₂₀. *J. Chem. Theory Comput.* **2014**, *10*, 1546–1553.

(101) Santra, B.; Klimeš, J.; Alfè, D.; Tkatchenko, A.; Slater, B.; Michaelides, A.; Car, R.; Scheffler, M. Hydrogen bonds and van der Waals forces in ice at ambient and high pressures. *Phys. Rev. Lett.* **2011**, *107*, 185701.

(102) Brandenburg, J. G.; Maas, T.; Grimme, S. Benchmarking DFT and semiempirical methods on structures and lattice energies for ten ice polymorphs. *J. Chem. Phys.* **2015**, *142*, 124104.

(103) Buch, V.; Devlin, J. P.; Monreal, I. A.; Jagoda-Cwiklik, B.; Uras-Aytemiz, N.; Cwiklik, L. Clathrate hydrates with hydrogen-bonding guests. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10245–10265.

(104) Koh, C. A.; Sloan, E. D.; Sum, A. K.; Wu, D. T. Fundamentals and applications of gas hydrates. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 237–257.

(105) Koh, C. A. Towards a fundamental understanding of natural gas hydrates. *Chem. Soc. Rev.* **2002**, *31*, 157–167.

(106) Struzhkin, V. V.; Militzer, B.; Mao, W. L.; Mao, H.; Hemley, R. J. Hydrogen storage in molecular clathrates. *Chem. Rev.* **2007**, *107*, 4133–4151.

(107) Kim, D.-Y.; Park, Y.; Lee, H. Tuning clathrate hydrates: Application to hydrogen storage. *Catal. Today* **2007**, *120*, 257–261.

(108) Roshandell, M.; Santacana-Vall, J.; Karnani, S.; Botimer, J.; Taborek, P.; Dunn-Rankin, D. Burning ice—direct combustion of methane clathrates. *Combust. Sci. Technol.* **2016**, *188*, 2137–2148.

(109) Kvenvolden, K. A. Potential effects of gas hydrate on human welfare. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3420–3426.

(110) Mousis, O.; Choukroun, M.; Lunine, J. I.; Sotin, C. Equilibrium composition between liquid and clathrate reservoirs on Titan. *Icarus* **2014**, *239*, 39–45.

(111) Khan, A. Theoretical studies of CO₂(H₂O)_{20,24,28} clusters: Stabilization of cages in hydrates by CO₂ guest molecules. *J. Mol. Struct.: THEOCHEM* **2003**, *664–665*, 237–245.

(112) Patchkovskii, S.; Tse, J. S. Thermodynamic stability of hydrogen clathrates. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14645–14650.

(113) Grimme, S.; Antony, J.; Schwabe, T.; Muck-Lichtenfeld, C. Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio)organic molecules. *Org. Biomol. Chem.* **2007**, *5*, 741–758.

(114) Alavi, S.; Ripmeester, J. A. Hydrogen-gas migration through clathrate hydrate cages. *Angew. Chem.* **2007**, *119*, 6214–6217.

(115) Geng, C.-Y.; Wen, H.; Zhou, H. Molecular simulation of the potential of methane reoccupation during the replacement of methane hydrate by CO₂. *J. Phys. Chem. A* **2009**, *113*, 5463–5469.

(116) Loboda, O.; Goncharuk, V. Theoretical study on icosahedral water clusters. *Chem. Phys. Lett.* **2010**, *484*, 144–147.

(117) Roman-Perez, G.; Moaied, M.; Soler, J. M.; Yndurain, F. Stability, adsorption, and diffusion of CH₄, CO₂, and H₂ in clathrate hydrates. *Phys. Rev. Lett.* **2010**, *105*, 145901.

(118) Li, Q.; Kolb, B.; Roman-Perez, G.; Soler, J. M.; Yndurain, F.; Kong, L.; Langreth, D. C.; Thonhauser, T. Ab initio energetics and kinetics study of H₂ and CH₄ in the SI clathrate hydrate. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *84*, 153103.

(119) Chattaraj, P. K.; Bandaru, S.; Mondal, S. Hydrogen storage in clathrate hydrates. *J. Phys. Chem. A* **2011**, *115*, 187–193.

(120) Xu, L.; Wang, X.; Liu, L.; Yang, M. First-principles investigation on the structural stability of methane and ethane clathrate hydrates. *Comput. Theor. Chem.* **2011**, *977*, 209–212.

(121) Srivastava, H. K.; Sastry, G. N. Viability of clathrate hydrates as CO₂ capturing agents: A theoretical study. *J. Phys. Chem. A* **2011**, *115*, 7633–7637.

(122) Bravo-Perez, G.; Saint-Martin, H. A theoretical study of the confinement of methane in water clusters. *Int. J. Quantum Chem.* **2012**, *112*, 3655–3660.

(123) Ramya, K. R.; Venkatnathan, A. Stability and reactivity of methane clathrate hydrates: Insights from density functional theory. *J. Phys. Chem. A* **2012**, *116*, 7742–7745.

(124) Fan, X.; Xu, L.; Liu, L.; Yang, M.; Zeng, Q.; Yang, M. Polarization response of methane encapsulated in water cages. *Comput. Theor. Chem.* **2013**, *1013*, 52–56.

(125) Park, S. S.; Lee, S.; Won, Y. S.; Ahn, Y. J. Comparative investigation of polyhedral water cages of (H₂O)_n (*n* = 20, 24, and 28) encaging CH₄ and SF₆ as guest molecules. *Chem. Phys.* **2014**, *441*, 128–136.

(126) Teeratchanan, P.; Hermann, A. Computational phase diagrams of noble gas hydrates under pressure. *J. Chem. Phys.* **2015**, *143*, 154507.

(127) Zeng, Q.; Li, J.; Huang, H.; Wang, X.; Yang, M. Polarization response of clathrate hydrates capsulated with guest molecules. *J. Chem. Phys.* **2016**, *144*, 204308.

(128) Sun, N.; Li, Z.; Qiu, N.; Yu, X.; Zhang, X.; Li, Y.; Yang, L.; Luo, K.; Huang, Q.; Du, S. Ab initio studies on the clathrate hydrates of some nitrogen- and sulfur-containing gases. *J. Phys. Chem. A* **2017**, *121*, 2620–2626.

(129) Zhang, Y.; Yang, W. Comment on “Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1998**, *80*, 890–893.

(130) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(131) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism

- spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (132) Zhao, Y.; Truhlar, D. G. Applications and validations of the Minnesota density functionals. *Chem. Phys. Lett.* **2011**, *502*, 1–13.
- (133) 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.
- (134) 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.
- (135) Khan, A. Theoretical studies of $\text{CH}_4(\text{H}_2\text{O})_{20}$, $(\text{H}_2\text{O})_{21}$, $(\text{H}_2\text{O})_{20}$, and fused dodecahedral and tetrakaidecahedral structures: How do natural gas hydrates form? *J. Chem. Phys.* **1999**, *110*, 11884–11889.
- (136) Khan, A. Stabilization of hydrate structure H by N_2 and CH_4 molecules in $4^35^66^3$ and 5^{12} cavities, and fused structure formation with $5^{12}6^8$ cage: A theoretical study. *J. Phys. Chem. A* **2001**, *105*, 7429–7434.
- (137) Khan, A. *Ab initio* studies of $(\text{H}_2\text{O})_{28}$ hexakaidecahedral cluster with Ne , N_2 , CH_4 , and C_2H_6 guest molecules in the cavity. *J. Chem. Phys.* **2002**, *116*, 6628–6633.
- (138) Lebsir, F.; Bouyacoub, A.; Bormann, D.; Krallafa, A. Theoretical investigations of CH_4 , C_2H_6 , CO_2 and N_2 guest molecules into a dodecahedron water cluster cavities. *J. Mol. Struct.: THEOCHEM* **2008**, *864*, 42–47.
- (139) Kumar, P.; Sathyamurthy, M. Theoretical studies of host-guest interaction in gas hydrates. *J. Phys. Chem. A* **2011**, *115*, 14276–14281.
- (140) Bakowies, D. Accurate extrapolation of electron correlation energies from small basis sets. *J. Chem. Phys.* **2007**, *127*, 164109.
- (141) Cao, X.; Su, Y.; Zhao, J. Stability and vibrations of guest molecules in the type II clathrate hydrate: A first-principles study of solid phase. *J. Phys. Chem. A* **2015**, *119*, 7063–7069.
- (142) Cao, X.; Su, Y.; Liu, Y.; Zhao, J.; Liu, C. Storage capacity and vibration frequencies of guest molecules in CH_4 and CO_2 hydrates by first-principles calculations. *J. Phys. Chem. A* **2014**, *118*, 215–222.
- (143) Liu, Y.; Ojamäe, L. C–C stretching raman spectra and stabilities of hydrocarbon molecules in natural gas hydrates: A quantum chemical study. *J. Phys. Chem. A* **2014**, *118*, 11641–11651.
- (144) Ramya, K. R.; Kumar, R.; Venkatnathan, A. Energy and spectral characteristics of hydrogen occupied pure and tetrahydrofuran doped water cages. *Comput. Theor. Chem.* **2014**, *1039*, 28–32.
- (145) Kumar, P.; Mishra, B. K.; Sathyamurthy, N. Density functional theoretic studies of host-guest interaction in gas hydrates. *Comput. Theor. Chem.* **2014**, *1029*, 26–32.
- (146) Deible, M. J.; Tuguldur, O.; Jordan, K. D. Theoretical study of the binding energy of a methane molecule in a $(\text{H}_2\text{O})_{20}$ dodecahedral cage. *J. Phys. Chem. B* **2014**, *118*, 8257–8263.
- (147) Florusse, L.; Peters, C.; Schoonman, J.; Hester, K.; Koh, C.; Dec, S.; Marsh, K.; Sloan, E. Stable low-pressure hydrogen clusters stored in a binary clathrate hydrate. *Science* **2004**, *306*, 469–471.
- (148) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.
- (149) Liakos, D. G.; Neese, F. Is it possible to obtain coupled cluster quality energies at near density functional theory cost? Domain-based local pair natural orbital coupled cluster vs. modern density functional theory. *J. Chem. Theory Comput.* **2015**, *11*, 4054–4063.
- (150) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.
- (151) Werner, H.-J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, 164102.
- (152) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B–Ne, and Al–Ar. *J. Chem. Phys.* **2008**, *128*, 084102.
- (153) Richard, R. M.; Lao, K. U.; Herbert, J. M. Aiming for benchmark accuracy with the many-body expansion. *Acc. Chem. Res.* **2014**, *47*, 2828–2836.
- (154) Lao, K. U.; Liu, K.-Y.; Richard, R. M.; Herbert, J. M. Understanding the many-body expansion for large systems. II. Accuracy considerations. *J. Chem. Phys.* **2016**, *144*, 164105.
- (155) Liu, K.-Y.; Herbert, J. M. Understanding the many-body expansion for large systems. III. Critical role of four-body terms, counterpoise corrections, and cutoffs. *J. Chem. Phys.* **2017**, *147*, 161729.
- (156) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. State of the art in counterpoise theory. *Chem. Rev.* **1994**, *94*, 1873–1885.
- (157) Richard, R. M.; Lao, K. U.; Herbert, J. M. Achieving the CCSD(T) basis-set limit in sizable molecular clusters: Counterpoise corrections for the many-body expansion. *J. Phys. Chem. Lett.* **2013**, *4*, 2674–2680.
- (158) Richard, R. M.; Lao, K. U.; Herbert, J. M. Approaching the complete-basis limit with a truncated many-body expansion. *J. Chem. Phys.* **2013**, *139*, 224102.
- (159) Lao, K. U.; Herbert, J. M. Energy decomposition analysis with a stable charge-transfer term for interpreting intermolecular interactions. *J. Chem. Theory Comput.* **2016**, *12*, 2569–2582.
- (160) Fleischer, E. B.; Janda, K. C. Prediction of clathrate structure type and guest position by molecular mechanics. *J. Phys. Chem. A* **2013**, *117*, 4001–4010.
- (161) Proynov, E.; Liu, F.; Gan, Z.; Wang, M.; Kong, J. Density-functional approach to the three-body dispersion interaction based on the exchange dipole moment. *J. Chem. Phys.* **2015**, *143*, 084125.
- (162) Ohio Supercomputer Center; <http://osc.edu/ark:/19495/f5s1ph73>.