JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2018, 14, 2955−2978

pubs.acs.org/JCTC

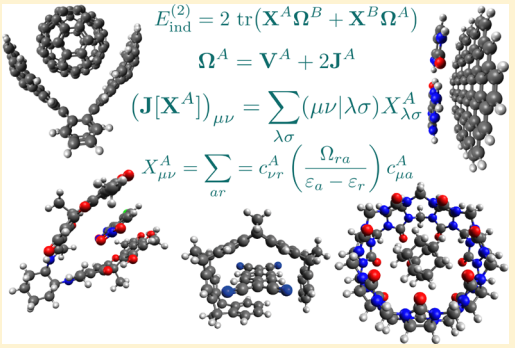
Atomic Orbital Implementation of Extended Symmetry-Adapted Perturbation Theory (XSAPT) and Benchmark Calculations for Large Supramolecular Complexes

Ka Un Lao† and John M. Herbert*

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

Supporting Information

ABSTRACT: We report an implementation of extended symmetry-adapted perturbation theory (XSAPT) in the atomic orbital basis, extending this method to systems where the monomers are large. In our "XSAPT(KS)" approach, monomers are described using range-separated Kohn−Sham (KS) density functional theory (DFT), with correct asymptotic behavior set by tuning the range-separation parameter ω in a monomer-specific way. This is accomplished either by conventional ionization potential (IP)-based tuning, in which ω is adjusted to satisfy the condition $\varepsilon_{\mathrm{HOMO}}(\omega) = -\mathrm{IP}(\omega)$, or else using a "global density-dependent" (GDD) condition, in which ω is fixed based on the size of the exchange hole. The latter procedure affords better results for both total interaction energies and energy components, when used in conjunction with our third-generation pairwise atom−atom dispersion potential (+*ai*D3). Three-body (triatomic) dispersion terms are found to be important when the monomers are large, and we incorporate these by means of an Axilrod−Teller−Muto term, $E_{\mathrm{disp,3B}}^{\mathrm{ATM}}$, which reduces errors in supramolecular interaction energies by about a factor of 2. The XSAPT(KS) + *ai*D3 + $E_{\mathrm{disp,3B}}^{\mathrm{ATM}}(\omega_{\mathrm{GDD}})$ approach affords mean absolute errors as low as 1.2 and 4.2 kcal/mol, respectively, for the L7 and S12L benchmark test sets of large dimers. Such errors are comparable to those afforded by far more expensive methods, such as DFT-SAPT, and the closely related second-order perturbation theory with coupled dispersion (MP2C). We also survey the performance of various other quantum-chemical methods for these data sets and identify several semiempirical and DFT-based approaches whose accuracy approaches that of MP2C, at dramatically reduced cost.



1. INTRODUCTION

Efficient and accurate modeling of noncovalent interactions and, in particular, the dispersion or van der Waals interaction, is an active topic of research in computational chemistry because of the importance of such interactions in the supramolecular chemistry of host/guest complexes, crystal packing and polymorphism, protein folding, and conformational energies of large molecules in general.[1−10] Density functional theory (DFT) calculations based on popular semilocal functions are fundamentally incapable of describing the long-range electron correlation effects that give rise to dispersion, as these have inherently nonlocal, many-body properties. To do better requires either a nonlocal correlation functional,[11−14] or else the addition of *a posteriori* corrections.[15−19] Examples of the latter approach include empirical atom−atom dispersion (+D) potentials,[15,18,19] as well as similar potentials wherein the C_6 (and perhaps higher C_n) coefficients are determined on-the-fly based on the electron density.[20−30] Three-body (triatomic) dispersion correction may be necessary as well,[31−34] especially for larger molecules, leading ultimately to many-body approaches that include the electrodynamic response of the surrounding polarizable atoms, by solving the self-consistent screening equation of classical electrodynamics.[10,28,35−38]

From a wave function perspective, second-order Møller−Plesset perturbation theory (MP2) is effective at describing intermolecular interactions in systems dominated by electrostatics and polarization[39,40] but usually overestimates the dispersion energy,[39] especially for systems exhibiting π-stacking interactions.[41−43] The reason is that dispersion at the MP2 level in effect arises from an "uncoupled Hartree-Fock" expression for C_6.[44] That is, the dispersion energy has the asymptotic form $-C_6/R^6$ with

$$C_6 = \frac{3}{\pi}\int_0^\infty \alpha_A(i\omega)\alpha_B(i\omega)\,\mathrm{d}\omega \tag{1}$$

where $\alpha_A(i\omega)$ and $\alpha_B(i\omega)$ are dynamic polarizabilities (evaluated at imaginary frequency $i\omega$), computed at the Hartree−Fock (HF) level not the correlated level.[44] Various strategies have been devised to address this deficiency in MP2's description of dispersion, including spin-component-scaled MP2 (SCS-MP2),[45] long-range attenuated MP2 (att-MP2),[46] and a "coupled MP2" approach (MP2C).[47,48] At present, however, the $O(N^7)$ coupled-cluster method with single,

Received: January 19, 2018
Published: April 16, 2018

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

double, and perturbative triple excitations [CCSD(T)], extrapolated to the complete basis-set (CBS) limit, remains the gold standard for noncovalent interactions. Higher-order corrections to interactions energies are typically <0.1 kcal/mol.[49,50]

A perturbative view of intermolecular interactions results in an alternative formalism called symmetry-adapted perturbation theory (SAPT),[3,39,43,51−55] which provides a solid theoretical basis to resolve one of the principle shortcomings of the supramolecular approach, namely, basis set superposition error (BSSE).[56] This is bypassed because SAPT is a perturbative expansion of the interaction energy itself, and energy subtraction ($\Delta E_{AB} = E_A - E_B$), from which arises BSSE, is not required. Moreover, the intermolecular energy in SAPT is naturally decomposable into physically meaningful components.[3,43,57]

The most common form of SAPT, usually called SAPT0,[3] treats intermolecular terms in the Hamiltonian at second order in perturbation theory but describes the monomers at the HF level. This approach is efficient enough for application to large systems, for example, an intercalation complex in DNA,[58,59] protein/ligand complexes,[60] and π-stacked complexes involving graphene.[61,62] However, SAPT0 affords accurate interaction energies only in small basis sets, where it benefits from error cancellation in the dispersion energy, or upon empirical scaling of certain terms.[63] Higher-level methods such as SAPT2+3 or SAPT(CCSD),[3] which incorporate additional electron correlation effects in a proper way, are generally quite accurate[3,63,64] but exhibit the same $O(N^7)$ scaling as CCSD(T). A more affordable alternative, and one that is pursued here, is to replace the MP2-style dispersion and exchange-dispersion terms in SAPT0 with analytic atom−atom dispersion potentials,[65,66] similar to dispersion-corrected DFT. This affords an $O(N^3)$ method, albeit one that treats the monomers at the HF level of theory.

A correlated description of the monomers can sometimes be important, as in the case of strong hydrogen bonds.[39] With this in mind, we have pursued SAPT calculations based on Kohn−Sham (KS) molecular orbitals as a low-cost way to incorporate intramolecular electron correlation,[43,67,68] with the remaining long-range intermolecular correlation handled via the perturbative expansion. This approach is known as SAPT(KS).[69−71] Proper asymptotic behavior of the monomer exchange-correlation potentials is crucial to its success,[69−71] and we have demonstrated that this can be achieved using long-range corrected (LRC) density functionals,[67] otherwise known as range-separated hybrid functionals.[72]

An alternative DFT-based SAPT approach is DFT-SAPT,[54] also known as SAPT(DFT),[55] in which DFT response theory is used to compute frequency-dependent density susceptibilities $\alpha(\mathbf{r}_1,\mathbf{r}_2|i\omega)$ for the monomers. These susceptibilities, evaluated at imaginary frequencies $i\omega$, are then used to compute the "coupled Kohn−Sham" (CKS) dispersion energy:[73,74]

$$E_{\mathrm{disp}}^{\mathrm{CKS}} = -\frac{1}{2\pi}\int_0^\infty \mathrm{d}\omega \int \frac{\alpha_A(\mathbf{r}_1,\mathbf{r}'_1|i\omega)\alpha_B(\mathbf{r}_2,\mathbf{r}'_2|i\omega)}{|\mathbf{r}_1-\mathbf{r}_2||\mathbf{r}'_1-\mathbf{r}'_2|} \times \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}'_1\,\mathrm{d}\mathbf{r}'_2 \quad (2)$$

(This expression should be compared to the uncoupled C_6 expression in eq 1.) As with SAPT(KS) calculations, proper asymptotic behavior of the exchange-correlation potential is important for the accuracy of DFT-SAPT, and has been accomplished using various asymptotic correction schemes.[75−77]

Traditional SAPT is a well-defined method for dimers, and while three-body nonadditivity corrections have been derived,[78−80] they are expensive to evaluate. For this reason, our group has developed an "extended" version of SAPT (XSAPT),[39,43,66,81] in which the monomer wave functions incorporate many-body polarization effects via the variational "explicit polarization" (XPol) method.[82−84] These monomer wave functions are then used in subsequent SAPT calculations, thereby extending traditional SAPT to an arbitrary number of monomers. The computational cost of XSAPT is pairwise-additive with respect to the number of monomers,[66] while its computational scaling with respect to monomer size depends upon the treatment of the second-order dispersion and exchange-dispersion terms.[43] The limitations of second-order (SAPT0) dispersion were discussed above; an alternative is to scale the direct dispersion term by an empirical factor while omitting the exchange-dispersion term, which yields improved accuracy at $O(N^4)$ cost.[43,60] The method becomes even more affordable if the direct dispersion term is replaced by atom−atom potentials, as suggested originally by Heßelmann.[65] This "XSAPT(KS)+D" approach[43,66,68] exhibits only cubic scaling with respect to *monomer* (not supersystem) size, as the rate-limiting step is a DFT-based XPol calculation on each monomer. Monomer DFT calculations and subsequent pairwise SAPT calculations are trivially parallelizable.

Unlike dispersion-corrected DFT, which has something of a double-counting problem because some contribution to the dispersion interaction is contained already in the uncorrected functional,[85−89] the XSAPT(KS)+D approach is based on a well-defined partition of the interaction energy, with the "+D" substituting for the (expensive and inaccurate) second-order dispersion terms. Double-counting is thereby avoided. After two initial attempts,[66,68] we ultimately introduced a third-generation "+*ai*D3" *ab initio* dispersion potential,[43,90] in which atom−atom potentials of the form[43,91]

$$E_{\mathrm{disp}}^{ai\mathrm{D3}} = -\sum_{i\in A}\sum_{\substack{j\in B\\(B\neq A)}}\left[f_6(R_{ij})\left(\frac{C_{6,ij}}{R_{ij}^6}\right) + f_8(R_{ij})\left(\frac{C_{8,ij}}{R_{ij}^8}\right)\right] \quad (3)$$

are fit directly to *ab initio* dispersion energies from benchmark SAPT2+(3) calculations. This XSAPT(KS)+*ai*D3 approach affords accurate results for both total interaction energies as well as individual energy components.[43]

For clusters of small molecules, the performance of XSAPT(KS)+*ai*D3 is typically superior to supramolecular methods with similar scaling.[43] However, only limited results are available for systems where the monomers are large, and there is reason to question the validity of atomic-pairwise dispersion potentials in systems composed of a large number of highly polarizable centers,[92] such as molecular crystals or supramolecular complexes assembled from highly conjugated monomers.[37,93,94] The present work focuses on examples from the latter category. Such examples often involve monomers that are rather large, so to facilitate XSAPT(KS) + *ai*D3 calculations in these cases, we have reformulated XSAPT in the atomic orbital (AO) basis. This formulation avoids the four-index integral transformation that is required in the original, molecular orbital (MO) version of the method. This work also presents the spin-unrestricted implementation of XSAPT(KS), which was already used in ref 57.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

2. THEORY

2.1. Atomic Orbital Implementation of XSAPT. In XSAPT(KS), we consider intermolecular correlation through second order and incorporate intramolecular correlation via monomer KS calculations. This corresponds to the SAPT0 version of SAPT(KS), and the interaction energy at this level can be expressed as

$$E_{\mathrm{int}}^{\mathrm{SAPT0}} = E_{\mathrm{elst}}^{(1)} + E_{\mathrm{exch}}^{(1)} + E_{\mathrm{ind}}^{(2)} + E_{\mathrm{exch\text{-}ind}}^{(2)} + E_{\mathrm{disp}}^{(2)} + E_{\mathrm{exch\text{-}disp}}^{(2)} + \delta E_{\mathrm{int}}^{\mathrm{HF}} \tag{4}$$

The various terms in eq 4 are discussed below, but we note here that the final term

$$\delta E_{\mathrm{int}}^{\mathrm{HF}} = E_{\mathrm{int}}^{\mathrm{HF}} - [E_{\mathrm{elst}}^{(10)}(\mathrm{HF}) + E_{\mathrm{exch}}^{(10)}(\mathrm{HF}) + E_{\mathrm{ind,resp}}^{(20)}(\mathrm{HF}) + E_{\mathrm{exch\text{-}ind,resp}}^{(20)}(\mathrm{HF})] \tag{5}$$

is incorporated to capture polarization effects beyond second order, by means of a counterpoise-corrected HF calculation of the interaction energy, $E_{\mathrm{int}}^{\mathrm{HF}}$. The monomers are described at the HF level in both eqs 4 and 5, but a response ("resp") correction, as described below, is included in the second-order induction and exchange-induction terms, as indicated in eq 5.

The nondispersion part of $E_{\mathrm{int}}^{\mathrm{SAPT0}}$ is

$$E_{\mathrm{non\text{-}disp}}^{\mathrm{SAPT0}} = E_{\mathrm{elst}}^{(1)} + E_{\mathrm{exch}}^{(1)} + E_{\mathrm{ind}}^{(2)} + E_{\mathrm{exch\text{-}ind}}^{(2)} \tag{6}$$

Evaluation of these terms in the AO basis sidesteps the need for a costly four-index integral transformation, resulting in an $O(N^3)$ rather than $O(N^5)$ implementation, where N measures monomer size. We also find that the AO formulation circumvents problems with linear dependencies that we have encountered using our resolution-of-identity implementation of XSAPT,[81] specifically when the monomers and basis sets are large. Finally, the AO-based approach naturally facilitates the use of linear-scaling algorithms for construction of the Coulomb ($\mathbf{J}$) and exchange ($\mathbf{K}$) matrices. The SAPT0 dispersion energy

$$E_{\mathrm{disp}}^{\mathrm{SAPT0}} = E_{\mathrm{disp}}^{(2)} + E_{\mathrm{exch\text{-}disp}}^{(2)} \tag{7}$$

is not considered here, as it is replaced by an *ab initio* dispersion potential; see section 2.3.

Closed-shell SAPT0 formulas in the AO basis were first provided by Heßelmann et al.,[95,96] and a more efficient formulation of the $E_{\mathrm{exch}}^{(1)}(S^2)$ and $E_{\mathrm{exch\text{-}ind}}^{(2)}$ terms was introduced later by Beer.[97] Spin-unrestricted formulas have been presented by Hapka et al.[98] These formulas are consolidated in the present work, using Beer's formulation since it requires construction of one fewer exchange matrix for $E_{\mathrm{exch}}^{(1)}(S^2)$ and one fewer Coulomb matrix for $E_{\mathrm{exch\text{-}ind}}^{(2)}$, as compared to Heßelmann's formulation. Our closed-shell AO expression for $E_{\mathrm{exch}}^{(1)}$ is based on the open-shell formulas in ref 98, and two typographical errors in refs 97 and 98 are corrected here. Closed-shell expressions are presented in this section and the analogous spin-unrestricted formulas can be found in Appendix A.

For a closed-shell (CS) system, the first-order electrostatic energy can be expressed as

$$[E_{\mathrm{elst}}^{(1)}]_{\mathrm{CS}} = \mathrm{tr}(2\mathbf{P}^A\mathbf{V}^B + 2\mathbf{P}^B\mathbf{V}^A + 4\mathbf{P}^B\mathbf{J}^A) + V_0 \tag{8}$$

Here, $\mathbf{P}^{A,B}$ are the one-electron density matrices for monomers A and B, $\mathbf{V}^{A,B}$ are the corresponding electrostatic potentials, $\mathbf{J}^{A,B} \equiv \mathbf{J}[\mathbf{P}^{A,B}]$ are the Coulomb matrices, and V_0 is the internuclear repulsion energy. Note that $\mathrm{tr}(\mathbf{P}^B\mathbf{J}^A) = \mathrm{tr}(\mathbf{P}^A\mathbf{J}^B)$, which is why eq 8 does not at first appear to be symmetric with respect to interchange of A and B.

Defining

$$\mathbf{\Omega}^{A,B} = \mathbf{V}^{A,B} + 2\mathbf{J}^{A,B} \tag{9}$$

and

$$\mathbf{h}^{A,B} = \mathbf{\Omega}^{A,B} - \mathbf{K}^{A,B} \tag{10}$$

the first-order exchange energy can be written[99]

$$\begin{aligned}[E_{\mathrm{exch}}^{(1)}]_{\mathrm{CS}} = 2\mathrm{tr}\{&-\mathbf{P}^A\mathbf{K}^B + \mathbf{T}^A\mathbf{h}^B + \mathbf{T}^B\mathbf{h}^A + \mathbf{T}^{AB}\mathbf{h}^A \\ &+ \mathbf{T}^{AB}\mathbf{h}^B + \mathbf{T}^A(2\mathbf{J}[\mathbf{T}^B] - \mathbf{K}[\mathbf{T}^B]) \\ &+ \mathbf{T}^{AB}(2\mathbf{J}[\mathbf{T}^A] - \mathbf{K}[\mathbf{T}^A]) + \mathbf{T}^{AB}(2\mathbf{J}[\mathbf{T}^B] - \mathbf{K}[\mathbf{T}^B]) \\ &+ \mathbf{T}^{AB}(2\mathbf{J}[\mathbf{T}^{BA}] - \mathbf{K}[\mathbf{T}^{BA}])\}\end{aligned} \tag{11}$$

The matrices $\mathbf{K}^{A,B} \equiv \mathbf{K}[\mathbf{P}^{A,B}]$ in this equation are the usual HF exchange matrices for the monomers, whereas $\mathbf{J}[\mathbf{X}]$ and $\mathbf{K}[\mathbf{X}]$ are Coulomb and exchange matrices constructed from a generalized density matrix, $\mathbf{X}$:

$$(\mathbf{J}[\mathbf{X}])_{\mu\nu} = \sum_{\lambda\sigma} (\mu\nu|\lambda\sigma)X_{\lambda\sigma} \tag{12a}$$

$$(\mathbf{K}[\mathbf{X}])_{\mu\nu} = \sum_{\lambda\sigma} (\mu\lambda|\sigma\nu)X_{\lambda\sigma} \tag{12b}$$

The generalized density matrices in eq 11,

$$\mathbf{T}^A = \mathbf{C}^A\mathbf{D}_{aa}(\mathbf{C}^A)^\dagger \tag{13a}$$

$$\mathbf{T}^B = \mathbf{C}^B\mathbf{D}_{bb}(\mathbf{C}^B)^\dagger \tag{13b}$$

$$\mathbf{T}^{AB} = \mathbf{C}^A\mathbf{D}_{ab}(\mathbf{C}^B)^\dagger \tag{13c}$$

$$\mathbf{T}^{BA} = \mathbf{C}^B\mathbf{D}_{ba}(\mathbf{C}^A)^\dagger \tag{13d}$$

are obtained by transforming the blocks of a matrix

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{aa} & \mathbf{D}_{ab} \\ \mathbf{D}_{ba} & \mathbf{D}_{bb} \end{pmatrix} \tag{14}$$

from the MO to the AO basis, using the MO coefficient matrices $\mathbf{C}^A$ and $\mathbf{C}^B$. The matrix $\mathbf{D}$ is defined as

$$\begin{aligned}\mathbf{D} &= (\mathbf{1} + \mathbf{S}^{AB})^{-1} - \mathbf{1} \\ &= \sum_{n=1}^{\infty} (-\mathbf{S}^{AB})^n\end{aligned} \tag{15}$$

where $\mathbf{S}^{AB}$ is the overlap matrix between occupied MOs of the dimer. This matrix can be written in a blocked form analogous to $\mathbf{D}$:

$$\mathbf{S}^{AB} = \begin{pmatrix} \mathbf{0} & \mathbf{S}_{ab} \\ \mathbf{S}_{ba} & \mathbf{0} \end{pmatrix} \tag{16}$$

Here $S_{ab} = \langle\phi_a|\phi_b\rangle$, with $a \in A$ and $b \in B$ used to label occupied MOs on either monomer.

Although the exact expression for $E_{\mathrm{exch}}^{(1)}$ was written down almost 40 years ago,[100] exact expressions for $E_{\mathrm{exch\text{-}ind}}^{(2)}$ and $E_{\mathrm{exch\text{-}disp}}^{(2)}$ were only published recently.[101,102] As such, second-order exchange energies have historically been evaluated using the "single exchange" approximation,[52,53] also known as the

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

"S^2" approximation. This approximation is accurate beyond the van der Waals (vdW) contact distance,[52,101,102] except when one of the monomers is an anion, in which case the $E^{(1)}_{\mathrm{exch}}(S^2)$ and $E^{(2)}_{\mathrm{exch\text{-}ind}}(S^2)$ terms must be scaled to obtain accurate results.[40,103] The scaling that is employed is based on the ratio of the exact and S^2 results for first-order exchange.

The S^2 approximation to $E^{(1)}_{\mathrm{exch}}$ will be used for all calculations reported in this work. This avoids the need to invert the dimer overlap matrix, and errors introduced by this approximation may be partly canceled by errors arising when the S^2 approximation is used in the $\delta E^{\mathrm{HF}}_{\mathrm{int}}$ term, eq 5. Within the S^2 approximation, the first-order exchange energy is

$$[E^{(1)}_{\mathrm{exch}}(S^2)]_{\mathrm{CS}} = -2\mathrm{tr}(\mathbf{P}^A\mathbf{K}^B + \mathbf{O}^\dagger\mathbf{h}^A + \mathbf{O}\mathbf{h}^B - \mathbf{P}^B\mathbf{S}\mathbf{O}\boldsymbol{\Omega}^A - \mathbf{O}\mathbf{S}\mathbf{P}^A\boldsymbol{\Omega}^B + \mathbf{O}\mathbf{K}^\dagger[\mathbf{O}]) \tag{17}$$

where

$$\mathbf{O} = \mathbf{P}^A\mathbf{S}\mathbf{P}^B \tag{18}$$

is a generalized density matrix, with $\mathbf{S}$ and $\mathbf{P}$ denoting the (monomer) overlap and density matrices, respectively.

The second-order induction energy consists of two terms,

$$E^{(2)}_{\mathrm{ind}} = E^{(2)}_{\mathrm{ind}}(A \leftarrow B) + E^{(2)}_{\mathrm{ind}}(B \leftarrow A) \tag{19}$$

where $E^{(2)}_{\mathrm{ind}}(A \leftarrow B)$ indicates the induction energy for monomer A due to the perturbing field of a frozen charge density from B, and vice versa for $E^{(2)}_{\mathrm{ind}}(B \leftarrow A)$. In the AO basis

$$[E^{(2)}_{\mathrm{ind}}]_{\mathrm{CS}} = 2\mathrm{tr}(\mathbf{X}^A\boldsymbol{\Omega}^B + \mathbf{X}^B\boldsymbol{\Omega}^A) \tag{20}$$

where

$$X^A_{\mu\nu} = \sum_{ar} c^A_{\nu r}U_{ra}(c^A_{\mu a})^\dagger \tag{21a}$$

$$X^B_{\mu\nu} = \sum_{bs} c^B_{\nu s}U_{sb}(c^B_{\mu b})^\dagger \tag{21b}$$

and

$$U_{ra} = \frac{\Omega_{ra}}{\varepsilon_a - \varepsilon_r} \tag{22}$$

Indices $r \in A$ and $s \in B$ are used to label virtual MOs belonging to monomers A and B, respectively, so that Ω_{ra} is a matrix element of $\boldsymbol{\Omega}^A$, expressed in the MO basis for monomer A. The quantity U_{sb} is defined similarly.

The second-order exchange-induction term can be separated in analogy to eq 19, with the $A \leftarrow B$ contribution expressed as[104]

$$[E^{(2)}_{\mathrm{exch-ind}}(S^2)(A \leftarrow B)]_{\mathrm{CS}} = -2\mathrm{tr}(\mathbf{X}^A\mathbf{K}^B + \mathbf{X}^A\mathbf{S}\mathbf{P}^B\mathbf{h}^A + \mathbf{P}^B\mathbf{S}\mathbf{X}^A\mathbf{h}^B - \mathbf{P}^B\mathbf{S}\mathbf{X}^A\mathbf{S}\mathbf{P}^B\boldsymbol{\Omega}^A - \mathbf{X}^A\mathbf{S}\mathbf{O}^\dagger\boldsymbol{\Omega}^B - \mathbf{O}\mathbf{S}\mathbf{X}^A\boldsymbol{\Omega}^B + 2\mathbf{O}^\dagger\mathbf{J}[\mathbf{X}^A] - 2\mathbf{P}^B\mathbf{S}\mathbf{O}\mathbf{J}[\mathbf{X}^A] - \mathbf{X}^A\mathbf{K}[\mathbf{O}] + \mathbf{X}^A\mathbf{S}\mathbf{P}^B\mathbf{K}^\dagger[\mathbf{O}] + \mathbf{P}^B\mathbf{S}\mathbf{X}^A\mathbf{K}[\mathbf{O}]) \tag{23}$$

within the S^2 approximation. The $B \leftarrow A$ contribution is evaluated similarly.

The $E^{(2)}_{\mathrm{ind}}$ and $E^{(2)}_{\mathrm{exch\text{-}ind}}$ terms are often replaced by their "response" analogues, $E^{(2)}_{\mathrm{ind,resp}}$ and $E^{(2)}_{\mathrm{exch\text{-}ind,resp}}$, in which coupled-perturbed HF (CPHF) equations are solved in order to compute the infinite-order correction for induction arising from a frozen partner density.[105,106] To obtain the response-corrected energies, the CPHF coefficients should be transformed into the AO basis and used as the matrix $\mathbf{U}$ in eq 21, in place of the definition in eq 22. Using the modified matrices $\mathbf{X}^A$ and $\mathbf{X}^B$ from eq 21, the formulas given above for $E^{(2)}_{\mathrm{ind}}$ and $E^{(2)}_{\mathrm{exch\text{-}ind}}$ then provide the response-corrected energies $E^{(2)}_{\mathrm{ind,resp}}$ and $E^{(2)}_{\mathrm{exch\text{-}ind,resp}}$.

In the absence of the SAPT0 dispersion or exchange-dispersion terms, the rate-limiting step in this formalism is construction of the nonsymmetric matrix $\mathbf{K}[\mathbf{T}^{BA}]$, in the case of exact first-order exchange (eq 11), or the nonsymmetric $\mathbf{K}^\dagger[\mathbf{O}]$ matrix in the S^2 approximation (eq 17). In the open-shell case (Appendix A), both matrices must be constructed from both α- and β-spin densities, so that the cost is essentially twice that of the closed-shell formalism.

2.2. Intramolecular Correlation Using Tuned LRC-DFT.

The asymptotic (large-r) behavior of the exchange-correlation potential is[76,107]

$$v_{\mathrm{xc}}(r) \sim -\frac{1}{r} + \Delta_\infty \tag{24}$$

with

$$\lim_{r\to\infty} \Delta_\infty = \mathrm{IP} + \varepsilon_{\mathrm{HOMO}} \tag{25}$$

Here, "IP" denotes the lowest ionization potential and $\varepsilon_{\mathrm{HOMO}}$ is the KS eigenvalue for the highest occupied MO (HOMO). To achieve proper asymptotic behavior, $v_{\mathrm{xc}}(r) \sim -1/r$, Baer and co-workers[72,108] propose to tune the range-separation parameter ω in LRC-DFT in order to satisfy the condition

$$\varepsilon_{\mathrm{HOMO}}(\omega) = -\mathrm{IP}(\omega) \tag{26}$$

That is, ω is tuned such that $\Delta_\infty = 0$.

Correct asymptotic behavior is crucial for obtaining accurate energy components in DFT-based SAPT,[69−71] and the nonempirical tuning procedure of Baer and co-workers, applied separately to each monomer, affords such behavior and provides accurate energy components.[43,67,68] At the same time, tuned LRC-DFT retains the relationship $v_{\mathrm{xc}}(\mathbf{r}) = \delta E_{\mathrm{xc}}/\delta\rho(\mathbf{r})$ that is sacrificed when using the asymptotic "splicing" schemes[75−77] that have traditionally been employed in DFT-based SAPT, and which amount to *ad hoc* correction of $v_{\mathrm{xc}}(\mathbf{r})$ without modification of $E_{\mathrm{xc}}[\rho]$. Tuned LRC-DFT functionals also provide a better description of properties such as polarizabilities and isotropic C_6 coefficients, and afford smaller delocalization errors, as compared to those obtained using splicing methods,[109,110] and are generally superior to splicing approaches in SAPT(KS) calculations.[110] We will denote by "ω_{IP}" the value of ω that satisfies the condition in eq 26. Tuned values of ω_{IP} are available in the Supporting Information, for all of the monomers considered in this work.

One drawback of the IP-tuning approach is that ω_{IP} exhibits a troublesome dependence on system size.[111−115] An alternative means to set the asymptotic decay of v_{xc} is the so-called "global density-dependent" (GDD) tuning procedure,[116] in which the optimal value

$$\omega_{\mathrm{GDD}} = C\langle d_x^2\rangle^{-1/2} \tag{27}$$

is related to the average distance d_x between an electron in the outer regions of a molecule and the exchange hole in the region of valence MOs. The quantity C in eq 27 is an empirical constant for a given LRC functional, and following the procedure in ref 116, we determined $C = 0.885$ for the LRC-ωPBE functional.[117,118] (Details can be found in the Supporting Information.) The basis set in these calculations is def2-TZVPP

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

augmented with diffuse functions on non-hydrogen atoms that are taken from Dunning's aug-cc-pVTZ basis set. (Henceforth, we refer to this composite basis as "haTZVPP", and we will abbreviate aug-cc-pVTZ as "aTZ".) Our optimized parameter C is almost the same as that determined in ref 116 (C = 0.90) for LRC-ωPBE with ω = 0.4 a_0^{-1} and the def2-TZVPP basis set. Since LRC-ωPBE(ω_{GDD}) provides a better description of polarizabilities in polyacetylene as compared to ω_{IP},[110] it is anticipated that using ω_{GDD} in place of ω_{IP} may afford more accurate energy components, especially in conjugated systems. Many of the supramolecular complexes considered here fall into this category.

2.3. Dispersion Corrections. In the XSAPT(KS) + *ai*D3 approach that is employed here, the second-order dispersion energy (eq 7) is replaced by an analytic atom−atom potential (eq 3) containing C_6 and C_8 terms.[43,90] Although the form of this potential is the same as in dispersion-corrected DFT methods, it is fit to dispersion energies from high-level SAPT2+(3)/aTZ and SAPT2+3(CCD)/aTZ calculations; see refs 3 or 119 for an explanation of the SAPT terminology. In that sense, E_{disp}^{aiD3} can be interpreted as a true dispersion energy, unlike early versions of dispersion-corrected DFT such as Grimme's DFT+D1[86] and DFT+D2.[87] (Grimme's DFT+D3,[31] on the other hand, uses atomic dispersion coefficients computed from first principles and is thus more easily interpretable as a dispersion energy.)

Although the *ai*D3 potential exhibits errors of $\lesssim$0.2 kcal/mol with respect to the benchmark dispersion energies on which it was parametrized,[43] its training set consists of monomers with no more than 20 non-hydrogen atoms whereas the monomers considered here are much larger. Pairwise dispersion corrections are known to overestimate the dispersion energy in certain supramolecular complexes and molecular crystals,[37,93,94] essentially because the presence of numerous polarizable centers can screen the interactions between any two centers, leading to effective atom-in-molecule C_n coefficients that differ from those derived for small molecules. Dobson[92] classifies this as "type B" nonadditivity, and it is missing not only from *ai*D3 but also (perhaps surprisingly) from MP2 and SAPT0 as well.[92] Less surprisingly, type B effects are absent in most DFT+D approaches, with the notable exception of DFT+D3, since it includes three-body dispersion corrections of the type described below. Other methods that can capture type B effects include the many-body dispersion (MBD) method of Tkatchenko et al.,[28,35,36,38] the Becke−Johnson exchange-dipole model (XDM),[34,120] and Axilrod−Teller−Muto (ATM) triple-dipole corrections.[31,32,121]

We will consider two different ATM-style corrections in this work. The first is an empirical one used by Grimme et al.[31] in the DFT+D3 dispersion correction. It has the form

$$E_{disp,3B}^{ATM(Grimme)} = \sum_{A<B<C}^{atoms} C_9^{ABC} f(\bar{R}_{ABC}) \left(\frac{3\cos\theta_a \cos\theta_b \cos\theta_c + 1}{R_{AB}^3 R_{AC}^3 R_{BC}^3} \right) \tag{28}$$

where the quantities θ_x are the internal angles of the atomic ABC triangle, $R_{XY} = |\mathbf{R}_X - \mathbf{R}_Y|$, and the C_9 coefficient for triatomic unit ABC is estimated from pairwise C_6 coefficients:

$$C_9^{ABC} = (C_6^{AB} C_6^{AC} C_6^{BC})^{1/2} \tag{29}$$

The function $f(\bar{R}_{ABC})$ in eq 28 damps the correction to zero as the averaged trimer distance $\bar{R}_{ABC} \to 0$.

A less empirical ATM-style correction takes the form of eq 28 but determines C_9^{ABC} from the electron density as suggested by Tkatchenko and Scheffler (TS).[26] Because there has been some debate about the best form for the damping function f in eq 28,[32,34,120,122] we use the modified three-body correction

$$E_{disp,3B}^{ATM(TS)} = \sum_{A<B<C}^{atoms} C_9^{ABC} \left(\frac{3\cos\theta_a \cos\theta_b \cos\theta_c + 1}{\bar{R}_{vdW,ABC}^9 + R_{AB}^3 R_{AC}^3 R_{BC}^3} \right) \tag{30}$$

that was suggested in ref 34. The $\bar{R}_{vdW,ABC}^9$ term in the denominator ensures that the correction remains finite as $\bar{R}_{ABC} \to 0$. The quantity

$$\bar{R}_{vdW,ABC} = (R_{vdW}^{AB} R_{vdW}^{AC} R_{vdW}^{BC})^{1/3} \tag{31}$$

measures the size of the triatomic unit ABC, with

$$R_{vdW}^{AB} = a(R_{vdW}^{A} + R_{vdW}^{B}) + b \tag{32}$$

Parameters a = 0.2525480 and b = 2.9273315 Å were determined in ref 34. As suggested in ref 32, we take the C_9 coefficients in eq 30 to be

$$C_9^{ABC} = \frac{8 P_A^{BC} P_B^{AC} P_C^{AB} (P_A^{BC} + P_B^{AC} + P_C^{AB})}{3(P_A^{BC} + P_B^{AC})(P_A^{BC} + P_C^{AB})(P_B^{AC} + P_C^{AB})} \tag{33}$$

where

$$P_X^{YZ} = C_9^X \left(\frac{\alpha_0^Y \alpha_0^Z}{\alpha_0^X} \right) \tag{34}$$

Atomic vdW radii (in eq 32), along with atomic C_9 coefficients and static polarizabilities (in eq 34) are computed using an atoms-in-molecules approach that we next describe.

The static polarizability for atom X, α_0^X, is computed using the electron density of the entire molecule by scaling the free-atom polarizability according to

$$\alpha_0^X = \left(\frac{V^X}{V_{free}^X} \right) \alpha_{0,free}^X \tag{35}$$

The quantity V^X/V_{free}^X is the ratio of the volume that atom X occupies in the molecule relative to the volume of the free atom. This ratio can be computed using Hirshfeld volume partitioning,[21] in which the Hirshfeld (or "stockholder") partitioning of the electron density[123] is used to partition three-dimensional space into atomic basins. In this work, we use Hirshfeld partitioning to evaluate V^X but take the free-atom volumes V_{free}^X from ref 124. The electron density used to evaluate V^X is the superposition of self-consistent XPol densities for the monomers, which avoids the needs for a supersystem calculation.

Atom-in-molecule C_9 coefficients are computed in a similar manner:

$$C_9^X = \left(\frac{V^X}{V_{free}^X} \right)^3 C_{9,free}^X \tag{36}$$

Free-atom coefficients for the elements H, He, C, N, O, F, Ne, Si, P, S, Cl, Ar, Br, and Kr are taken from ref 32, where they were computed from free-atom polarizabilities. For the remaining elements, $C_{9,free}^X$ is estimated using the relation[125]

$$C_{9,free}^X \approx \frac{3}{4} C_{6,free}^X \alpha_{0,free}^X \tag{37}$$

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

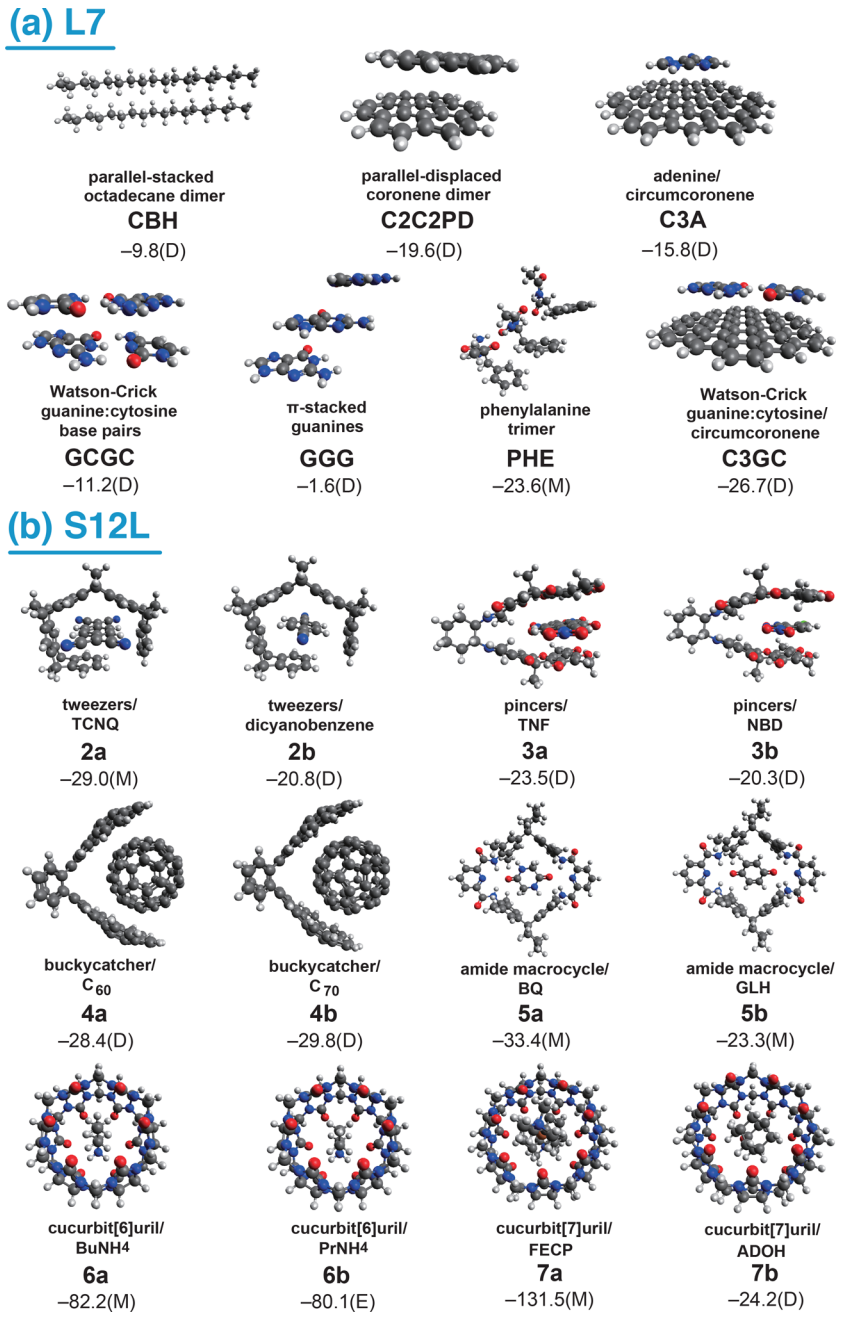


Figure 1. Structures of the complexes in the (a) L7 and (b) S12L data sets.[131,132] Interaction energies (in kcal/mol) for L7 are taken from ref 133, and binding energies for S12L are taken from ref 134, except that the value for **7a** is from ref 135. The coloring system is as follows: white spheres (hydrogen), gray (carbon), dark blue (nitrogen), red (oxygen), green (chlorine), and orange (iron). The designations "D", "E", and "M" indicate cases where the intermolecular interactions are dispersion-dominated, electrostatics-dominated, or of mixed character, respectively. Specifically, case D means $|E_{disp}/E_{elst}| > 2$, case E means $|E_{disp}/E_{elst}| < 0.5$, and case M lies in the intermediate regime, according to the classification scheme used in ref 136. These classifications are based on energetics computed at the XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}(\omega_{GDD})$/hpTZVPP level.

with the relevant data taken from ref 26.

Atomic vdW radii are also determined via Hirshfeld volume partitioning,

$$R_{vdW}^{X} = \left(\frac{V^{X}}{V_{free}^{X}}\right)^{1/3} R_{vdW,free}^{X} \tag{38}$$

with numerical values as reported in ref 26. These values are used to evaluate $\overline{R}^{9}_{vdW,ABC}$, the limiting value of the denominator in eq 30 as the interatomic *ABC* distances all tend to zero.

It has been argued[126] that inclusion of a three-body dispersion term may overestimate the extent to which the two-body dispersion should be reduced, and that contributions up to sixth order are necessary for a converged dispersion interaction. However, a C_8/R^8 term is included in our *ai*D3 dispersion potential (eq 3), whereas this term is omitted in the "MBD*" many-body dispersion approach of ref 126. The magnitude of the C_8/R^8 term is roughly 1/3 that of the C_6/R^6 term,[127] so we expect the former to reduce any over-compensation caused by the three-body correction $E_{disp,3B}^{ATM(TS)}$.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955–2978

To examine this possible overcompensation in a quantitative way, we tested the $E_{disp,3B}^{ATM(TS)}$ correction in conjunction with XSAPT(KS) + *ai*D3, as applied to the S22[128] and S66[129] benchmark data sets of small molecular dimers. (See section 3.2 for a discussion of functionals and basis sets used in this work.) As compared to plain XSAPT(KS) + *ai*D3, the three-body dispersion correction slightly reduces the mean absolute error (MAE) for S22 interaction energies, from 0.39 to 0.32 kcal/mol, and leaves the MAE for S66 unchanged at 0.28 kcal/mol. Although three-body dispersion effects are not anticipated to be large when the monomers are small, as they are in S22 and S66, this test nevertheless demonstrates that addition of $E_{disp,3B}^{ATM(TS)}$ is not deleterious to the performance of the method, and thus suggests that dispersion is not significantly overestimated when this nonadditive term is combined with the pairwise-additive *ai*D3 potential.

We furthermore evaluated the behavior of $E_{disp,3B}^{ATM(TS)}$ for several isomers of the benzene trimer. The first is a trimer extracted from the benzene crystal structure, for which SAPT(DFT) calculations afford a three-body dispersion energy of 0.18 kcal/mol.[80] In the context of SAPT(DFT), "three-body" dispersion means three distinct *molecules*, whereas in the present context it means three different *atoms*, hence the appropriate comparison is to restrict the sum in eq 30 to triples *ABC* involving one atom per monomer, which affords a three-body dispersion energy of 0.14 kcal/mol, in excellent agreement with the SAPT(DFT) result. In a second test, we considered three isomers of $(C_6H_6)_3$ frpm, ref 130, where benchmark three-body dispersion energies of 0.10, 0.07, and 0.03 kcal/mol were reported, based on the difference between CCSD(T) and MP2 results since the latter does not capture type-B nonadditivity.[92] For the same three isomers, the $E_{disp,3B}^{ATM(TS)}$ corrections are in excellent agreement at 0.09, 0.08, and 0.05 kcal/mol, respectively.

3. COMPUTATIONAL DETAILS

3.1. Data Sets. We will use the L7[131] and S12L[132] data sets to assess the performance of various methods on large supramolecular complexes. Structures and nomenclature for these complexes are shown in Figure 1.

For L7, the original QCISD(T)/CBS interaction energies of ref 131 were upgraded to the CCSD(T)/CBS level by Calbo et al.[137] and subsequently by Hansen et al.,[133] using a new CBS extrapolation.[139] In either case, the CCSD(T) calculations use the domain-based local-pair natural orbital (DLPNO) approximation,[140−143] albeit with tighter thresholds in ref 133, that is, `TightPNO` instead of `NormalPNO` in the ORCA program.[143] The tighter threshold affords differences as large as 5.2 kcal/mol in the interaction energy; see Table 1. As noted in ref 138, π-stacking interactions are especially sensitive to how many DLPNO pairs are included at the CCSD level, so it is no surprise that the coronene dimer (C2C2PD) is most sensitive to the threshold.

Explicitly correlated CCSD(T)-F12/cc-pVDZ calculations, using the `VeryTightPNO` threshold,[138] seem to resolve the aforementioned discrepancy for C2C2PD in favor of the calculation with the tighter threshold. For GCGC, however, the explicitly correlated result is closer to the DLPNO-CCSD(T) calculation based on the *looser* threshold, although the difference between tight and loose thresholds is smaller, at 2.0 kcal/mol. In view of this, plus the fact that CCSD(T)-F12 results are unavailable for the two L7 dimers containing circumcoronene, we will use the `TightPNO` CCSD(T)/CBS interaction energies from ref 133 as our benchmarks. (These values have also been used in other recent benchmarking studies.[144,145]) We take the variation in Table 1 as an indication of the accuracy to which these values should be trusted.

Table 1. Benchmark Interaction Energies for L7 Complexes, in kcal/mol

	CCSD(T)/CBS		
complex	normal[a]	tight[b]	CCSD(T)-F12/cc-pVDZ[c]
CBH	11.64	9.80	11.13
C2C2PD	24.81	19.60	19.14
C3A	17.98	15.80	–
GCGC	13.21	11.20	13.69
GGG	1.68	1.60	2.36
PHE	22.81	23.60	25.09
C3GC	29.86	26.70	–

[a]"Normal" DLPNO threshold, from ref 137. [b]"Tight" DLPNO threshold, from ref 133. [c]From ref 138.

The S12L data set consists of 12 host/guest complexes assembled from among six host molecules and two guests. Whereas six of seven L7 complexes are dispersion-dominated, the S12L complexes present a wider variety of noncovalent motifs including hydrogen bonding, dispersion and π-stacking, and cation-dipole interactions. Benchmark "gas-phase" binding energies for S12L have been reported,[132,135] based on back-correction of experimental free energies of association using semiempirical harmonic vibrational frequencies and COSMO-RS[146] solvation energies. The resulting binding energies range from 20−132 kcal/mol, with estimated uncertainties of ∼2 kcal/mol[132,135] that are thought mainly to arise from the solvation correction, which amounts to 1.4 kcal/mol on average.[134] The newer benchmarks from ref 134 are used for all S12L complexes except **7a**, for which only the older benchmark from ref 135 is available.

Quantum Monte Carlo (QMC) interaction energies with narrow statistical error bars have also been determined for six of the S12L systems (**2a**, **2b**, **4a**, **5a**, **6a**, and **7b**).[126] The MAE between these six QMC benchmarks and the latest back-corrected experimental benchmarks is 1.6 kcal/mol,[134] versus 2.4 kcal/mol when compared to older back-corrected benchmarks.[132,135] This comparison provides another suggestion as to the level of accuracy that can be anticipated from the benchmarks.

3.2. Functionals and Basis Sets. We tested the basis-set dependence of SAPT(KS) calculations using the LRC-ωPBE functional,[117] with either ω_{GDD} or ω_{IP} tuning, for a few representative systems including $F^-(H_2O)$, $(H_2O)_2$, and two isomers of $(C_6H_6)_2$. For this exercise, we examined the energy components $E_{elst}^{(1)}$ and $E_{exch}^{(1)}$ in comparison to benchmarks computed at the SAPT2+(3)/aTZ level, with results presented in Table S3. Results using a pseudocanonicalized[147,148] hpTZVPP basis differ from the benchmarks by only 5%. Unlike the traditional two-body SAPT approach based on a dimer-centered basis set,[149] which is ill-defined in many-body systems, use of the pseudocanonicalized dimer basis (which we originally called the "projected" basis[81]) requires self-consistent field calculations in monomer basis sets only, and furthermore captures some intermolecular charge transfer.[81] Since the pseudocanonicalized hpTZVPP basis also works well in small-molecule XSAPT calculations,[43,68] we use this basis for all of the XSAPT(KS) calculations reported here.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

For the δE_{int}^{HF} correction in eq 5, the 6-31+G(3d,3pd) basis set is used for the L7 complexes, as in previous work,[68] whereas for S12L we use cc-pVTZ (hereafter abbreviated "TZ") values of δE_{int}^{HF}, as reported in ref 150. For the two smallest systems in S12L (**2a** and **2b**), we have verified that the δE_{int}^{HF} correction computed using 6-31+G(3d,3pd) differs by at most 0.1 kcal/mol from the TZ value in ref 150.

The LRC-ωPBE functional,[117] with no short-range HF exchange, is used for the XSAPT(KS) calculations, although monomer-specific tuning parameters are provided also the LRC-ωPBEh functional,[117,151] which contains 20% short-range HF exchange. (See Table S1.) Atom-centered ChElPG charges are used for the XPol embedding,[39,81,152] computed using an Euler−Maclaurin−Lebedev quadrature grid with a "head space" of 3.0 Å (see ref 152) and radial shells spaced 0.25 Å apart for L7, and 0.50 Å apart for S12L. Weak interactions require the use of unusually dense integration grids for DFT,[153,154] and we use Lebedev grids with (N_{radial}, $N_{angular}$) = (110, 590) for the L7 data set and (86, 434) for S12L, which exceeds recommendations.[154] Neither the *ai*D2[68] nor the *ai*D3[43] dispersion potential is parametrized for iron, so in the ferrocene complex **7a** the pairwise dispersion potentials involving Fe are set to zero.

Whereas the benchmark results for L7 are *interaction* energies based on unrelaxed monomers, the benchmark *binding* energies for S12L include monomer relaxation energies. To compute these with the LRC-ωPBE function that is used in the XSAPT(KS) calculations is somewhat dubious since the monomers are large and this functional is not dispersion-corrected. Instead, we turn to results from ref 150, where S12L deformation energies were computed using a nonlocal DFT functional (NLDFT),[155] and also using MP2 and spin-component scaled (SCS) MP2. These methods deviate from one another, and from B97M-V/aTZ results reported here, by as much as 11 kcal/mol; see Table S6. For reasons of cost, the B97M-V calculations for complexes **4a**, **4b**, **7a**, and **7b** use the "heavy augmented" haTZ basis set, which differs from aTZ by removal of the diffuse functions on hydrogen. MAEs for S12L binding energies, with respect to the back-corrected experimental benchmarks of ref 134, are 2.4 kcal/mol (NLDFT), 6.4 kcal/mol (SCS-MP2), and 15.9 kcal/mol (MP2). As such, we select NLDFT to compute the monomer relaxation energies, and note that the B97M-V/aTZ relaxation energies reported here differ from the NLDFT results in ref 150 by no more than 2.8 kcal/mol.

Some results from traditional SAPT calculations will also be presented for comparison, including complete-basis DFT-SAPT results from ref 150. We also consider a version of SAPT0 called *s*SAPT0,[63] in which the second-order exchange-induction and exchange-dispersion terms are scaled in an attempt to mitigate the effects of the S^2 approximation. The *s*SAPT0 approach has been called the "bronze standard" of SAPT,[63] when used with the jun-cc-pVDZ (jaDZ) basis set that removes some diffuse functions relative to aug-cc-pVDZ. This basis is recommended for SAPT0 because it often affords favorable error cancellation,[63,156,157] since SAPT0 greatly overestimates dispersion energies in the basis-set limit yet the dispersion energy converges very slowly to that limit. Finally, we test a scaled-dispersion version of SAPT0 ("sd-SAPT") that neglects $E_{exch\text{-}disp}^{(2)}$ and empirically scales $E_{disp}^{(2)}$ to compensate.[60] This also reduces the computational effort from $O(N^5)$ to $O(N^4)$. The sd-SAPT calculations reported here use KS orbitals for the monomers, which does not significantly affect the accuracy (as compared to HF orbitals) in small dimers.[67] We previously optimized a scaling factor c_{disp} = 0.657 for sd-XSAPT(KS)/6-31G(d,2p) calculations using the LRC-ωPBE(ω_{IP}) functional,[43] and essentially the same scaling factor (c_{disp} = 0.661) is obtained in conjunction with ω_{GDD} tuning, for the S22 data set. The root-mean-square deviation (RMSD) in S22 interaction energies is 0.37 kcal/mol (ω_{IP}) versus 0.34 kcal/mol (ω_{GDD}).

Lastly, results from several DFT and MP2-based methods will also be discussed. The B97M-V functional is one of the best-performing DFT methods for noncovalent interactions,[43,153,158−160] and we will report B97M-V/aTZ results with counterpoise (CP) correction. A (75, 302) quadrature grid is used to evaluate the semilocal part of these functionals and the SG-1 grid[161] is used for the VV10 nonlocal correlation part. We also examine the semiempirical HF-3c and PBEh-3c methods,[162,163] which consist of a small-basis HF or DFT calculation combined with three empirical corrections: for dispersion, for BSSE, and for basis-set incompleteness. As introduced in ref 163, the PBEh-3c method includes the three-body dispersion correction $E_{disp,3B}^{ATM(Grimme)}$ by default, whereas HF-3c (ref 162) does not. In the present work, however, we separate this correction so that PBEh-3c can be compared directly to HF-3c, and $E_{disp,3B}^{ATM(Grimme)}$ can subsequently be added to either.

SAPT0 calculations were performed using the BETA5 version of the PSI4 program,[164] and HF-3c calculations were performed using version 3.0.3 of ORCA.[165] All other calculations were performed using Q-CHEM.[166] Our AO-based XSAPT algorithm was released with v. 4.4 of Q-CHEM. The self-consistent field convergence criterion was set to 10^{-7} a.u. because we have observed that interaction energies of the buckycatcher complexes **4a** and **4b** change by several kcal/mol as the threshold is tightened from the Q-CHEM default value of 10^{-5} a.u.

4. RESULTS AND DISCUSSION

In what follows, we evaluate not only the performance of various (X)SAPT-based methods, but also selected wave function and DFT approaches as well, for the L7[131] and S12L[132] sets of dimers. Many of the results using traditional approaches are taken from the literature (though collected together and compared for the first time here); the HF-3c, PBEh-3c, and B97M-V calculations are new. Regarding DFT approaches, we make no attempt at a comprehensive survey of functionals but rather evaluate only a selected set that have shown promise for prediction of noncovalent interaction energies.

4.1. L7 Data Set. *4.1.1. Wave Function Methods.* MAEs, mean deviations (MDs), and maximum errors in interaction energies for the L7 data set, using a variety of supersystem methods, are listed in Table 2. As expected, MP2/CBS is unreliable for these dispersion-bound complexes, with a MAE of 8.1 kcal/mol and a MD of −8.1 kcal/mol, symptomatic of a gross overestimation of dispersion energies. This is especially true for the π-stacked parallel-displaced coronene dimer (C2C2PD in Figure 1a), whose interaction energy is overestimated by almost a factor of 2, an error of 17.7 kcal/mol. This problem is greatly mitigated by including half of the MP3 correlation energy: the MP2.5/CBS method reduces the MAE to 1.1 kcal/mol with a maximum error of 2.5 kcal/mol, albeit at $O(N^6)$ cost.

Alternatively, spin-component scaling tends to improve the performance of MP2 for noncovalent interactions. In the

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

Table 2. Error Statistics in L7 Interaction Energies[a] Computed Using Supersystem Methods

method	error (kcal/mol) MAE	MD	max
wave function methods			
MP2/CBS[b]	8.10	−8.10	17.68
MP2.5/CBS[b]	1.13	−0.92	2.46
SCS-MP2/CBS[b]	2.43	−0.95	6.23
SCS(MI)-MP2/CBS[b]	4.55	−3.99	10.41
MP2C/CBS[b]	0.49	−0.14	1.82
att-MP2/aTZ[c]	1.14	−0.15	3.21
att-SCS-MP2/aTZ[c]	1.82	0.97	3.66
HF-3c	0.92	−0.26	1.67
HF-3c+$E_{disp,3B}^{ATM(Grimme)}$	1.21	0.82	3.48
PBEh-3c	1.39	0.02	3.32
PBEh-3c+$E_{disp,3B}^{ATM(Grimme)}$	1.59	1.12	5.52
DFT methods			
PBE-TS[d,e,f]	2.95	−2.95	4.85
PBE-TS(HI)[d,e,g]	2.43	−2.43	6.73
TPSS+D3(BJ)/def2-QZVP[b]	0.50	−0.23	1.23
B3LYP+D3(BJ)/def2-QZVP[b]	1.76	−1.76	2.99
M06-2X/def2-QZVP[b]	3.39	3.28	6.89
M06-2X+D3(0)/def2-QZVP[b]	1.37	0.19	3.37
PW6B95+D3(BJ)/def2-QZVP[b]	0.79	0.31	1.59
B2PLYP+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[h]	0.95	0.56	2.01
B2PLYP+NL+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[h]	1.07	−0.58	1.72
PBE-XDM/pc-2-spd[i]	2.83	2.18	7.19
LC-ωPBE-XDM/pc-2-spd[i]	0.69	−0.56	2.07
ωB97X-D/aTZ (CP)[j]	1.64	−1.64	2.41
B97M-V/aTZ	1.46	−1.46	2.87
B97M-V/aTZ (CP)	0.71	−0.58	2.21

[a]With respect to CCSD(T)/CBS benchmarks.[4,133] [b]From ref 131. [c]From ref 167. [d]From ref 168. [e]Using a plane-wave basis set. [f]With TS dispersion based on Hirshfeld partitioning of the density.[123] [g]With TS dispersion based on iterative Hirshfeld partitioning of the density.[169] [h]From ref 137. [i]From ref 144. [j]From ref 170.

original SCS-MP2 method,[45] the same- and opposite-spin scaling parameters were optimized using high-quality reaction energies, although they have also been optimized for molecular interaction (MI) using S22 interaction energies, in a method called SCS(MI)-MP2.[171] For the L7 data set, SCS-MP2 actually outperforms SCS(MI)-MP2, with an MAE of 2.4 kcal/mol for the former and 4.6 kcal/mol for the latter. This suggests that SCS-MP2 is more suitable for general applications as compared to SCS(MI)-MP2, or perhaps that the SCS(MI)-MP2 scaling parameters should be reoptimized using larger molecules. Nevertheless, the SCS-MP2/CBS method still overestimates the interaction energy of C2C2PD by 6.2 kcal/mol.

MP2's deficiencies here stem primarily from that method's description of dispersion,[42] and the MP2C method replaces the MP2 dispersion energy by a more accurate one computed using DFT response theory.[47,48] The cost of evaluating the dispersion energy scales as $O(N^4)$ if resolution-of-identity techniques are used,[48,172] although the overall cost remains $O(N^5)$. MP2C exhibits a MAE of 0.5 kcal/mol for L7, with a maximum error of 1.8 kcal/mol.

We have focused on CBS results in this discussion owing to the slow convergence of MP2 interaction energies. This convergence problem is ameliorated by the attenuated MP2 (att-MP2) method,[46] which attenuates the Coulomb operator used in the correlation energy calculation. This method not only reduces MP2's tendency to grossly overestimate π-stacking energies, but also removes significant BSSE so that the method is accurate in modest basis sets. Promising results are obtained using the aTZ basis set,[173] and the att-MP2/aTZ method affords a MAE of 1.1 kcal/mol for L7, comparable to MP2.5/CBS but without the $O(N^6)$ scaling or the need for quadruple-ζ results for use in CBS extrapolation. The SCS concept can be incorporated into att-MP2,[167] and for L7 the att-SCS-MP2 method affords a MAE of 1.8 kcal/mol with a maximum error of 3.7 kcal/mol.

At the other end of the spectrum of computational cost, we find that the semiempirical HF-3c and PBEh-3c methods perform quite well, with MAEs of only 0.9 and 1.4 kcal/mol, respectively. For reasons that are unclear, the correlated PBEh-3c results are slightly worse than the HF-3c ones, and furthermore both sets of results become slightly worse when the $E_{disp,3B}^{ATM(Grimme)}$ correction is added. Neither of these apparent anomalies is reflected in the S12L data discussed below, however. In any case, these results make HF-3c and PBEh-3c appear very promising for screening of large structures,[163] in computational drug computational drug design for example.

In summary, the best MP2-based method for the L7 data set is MP2C/CBS, although the attenuated MP2/aTZ methods are only marginally worse yet considerably more efficient. Performance of the low-cost HF-3c and PBEh-3c methods is comparable to that of the best MP2-based methods, at $O(N^3)$ cost.

4.1.2. DFT Methods. Traditional semilocal density functionals cannot properly describe nonlocal dispersion interactions; the PBE functional, for example, predicts that five of the seven L7 complexes are unbound. One approach to correct this behavior is to incorporate the pairwise TS dispersion model,[26] which uses atom-in-molecule C_6 coefficients determined from DFT densities to apply an *a posteriori* (but density-dependent) dispersion correction. The combined PBE-TS approach binds each of the L7 dimers, with a MAE of 3.0 kcal/mol in the interaction energies. (See Table 2.) The errors, which tend toward overbinding, are however considerably larger than those reported for small dimers using the same approach,[26] suggesting that corrections beyond the atomic-pairwise approach are necessary.

We next consider ATM-style three-body dispersion corrections, eq 28 and eq 30, which are listed in Table 3 for each of the L7 complexes. The larger π-stacked systems (C3A, C2C2PD, and C3GC) afford large, repulsive values of $E_{disp,3B}^{ATM(TS)}$, up to 6.6 kcal/mol for C3GC, whereas $E_{disp,3B}^{ATM(Grimme)}$ is considerably smaller. In what follows, we will show that the TS three-body correction is more suitable for use in conjunction with our

Table 3. Three-Body Dispersion Energies (in kcal/mol) for L7 Complexes[a]

system	$E_{disp,3B}^{ATM(TS)}$	$E_{disp,3B}^{ATM(Grimme)}$
CBH	1.45	0.73
GGG	0.76	0.26
C3A	3.85	1.23
C3GC	6.59	2.31
C2C2PD	6.18	1.72
GCGC	2.32	1.02
PHE	0.73	0.39

[a]Computed using XPol densities at the LRC-ωPBE(ω_{GDD})/hpTZVPP level.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

pairwise *ai*D3 potential to reproduce dispersion energies calculated using DFT response theory. Adding the $E_{disp,3B}^{ATM(TS)}$ correction to PBE-TS slightly reduces the MAEs for L7, to 2.2 kcal/mol.

Among DFT methods that incorporate Grimme's D3 dispersion correction,[31] the TPSS meta-GGA functional performs the best, with a MAE of 0.5 kcal/mol, and has been recommended for geometry optimization in large systems.[132] The best DFT method is TPSS+D3(BJ), where "D3(BJ)" indicates use of the Becke−Johnson (BJ) damping function for the D3 dispersion correction,[174] as compared to the "zero-damping" D3(0) approach.[31] Error statistics for this approach are comparable to those of the best-performing wave function approach, MP2C. The PW6B95 hybrid meta-GGA functional,[175] combined with the D3(BJ) dispersion correction, also performs well for L7, with a MAE of 0.8 kcal/mol. The PW6B95+D3(BJ) approach has also been recommended for use in calculations involving water clusters,[176] biomolecules,[177] transition metal catalysts,[178] large host/guest complexes,[132,134,179] and geometry optimizations.[180] We conclude that PW6B95+D3(BJ) appears to be reasonably accurate for a wide range of applications.

With the exception of the phenylalanine trimer PHE, each of the L7 systems is unbound at the B3LYP level,[170] and while Grimme's D3 correction fixes this, it overcorrects and the B3LYP+D3(BJ) method overestimates the L7 interaction energies with a MAE of 1.8 kcal/mol. The highly parametrized M06-2X functional, which only captures middle-range (not long-range) nonlocal correlation, affords a MAE of 3.4 kcal/mol. Long-range correlation is important in complexes of this size and must be added explicitly to M06-2X, via an empirical dispersion correction. The M06-2X+D3(0) method affords reasonable results for L7, with a MAE of 1.4 kcal/mol. The Becke-Johnson XDM dispersion correction, with two different density functionals, affords MAEs of 0.7 and 2.8 kcal/mol that are comparable to—but not significantly better than—the performance of the best DFT+D3 methods.

Considering a more recent functional containing proper nonlocal correlation, the B97M-V/aTZ approach[158] overestimates the L7 interaction energies with a MAE of 1.5 kcal/mol, despite affording more accurate interaction energies for many other noncovalent complexes.[43,158,160] This is partly a consequence of BSSE, and counterpoise correction of the B97M-V/aTZ results reduces the MAE to 0.7 kcal/mol. As an example, the B97M-V/aTZ interaction energies for C2C2PD are −21.5 and −22.3 kcal/mol with and without counterpoise correction, respectively.

4.1.3. (X)SAPT Methods. Next, we turn from supersystem methods to SAPT-based methods, for which L7 error statistics are listed in Table 4. By design, SAPT and XSAPT afford similar results for dimers,[81] which proves to be true for the systems considered here. As such, we mainly focus on XSAPT in the following discussion, except to note the poor performance of SAPT0/jaDZ and its scaled-dispersion counterpart, *s*SAPT0/jaDZ, which exhibit errors as large as 10 kcal/mol. Even the MAEs (4.8 kcal/mol for both methods) are larger than those documented in previous SAPT0/jaDZ studies of smaller molecules,[63] despite the favorable error cancellation that often accompanies use of the jaDZ basis set for SAPT0 calculations.[156,157] This error cancellation arises from the fact that SAPT0 dispersion energies are much too large in the CBS limit, but that dispersion converges much more slowly to that limit as compared to other energy components.[39] Failure of this cancellation in the present cases, as compared to the much smaller molecules considered in previous studies,[63,156,157] is likely attributable to the fact that the dispersion energies are simply much larger for the L7 complexes, and suggests that SAPT0/jaDZ should be used with caution in large systems.

Table 4. Error Statistics for L7 Interaction Energies[a] Computing Using (X)SAPT-Based Methods[b,c]

method	error (kcal/mol) MAE	MD	max
SAPT-based methods			
SAPT0/jun-cc-pVDZ	4.83	−3.67	10.28
*s*SAPT0/jun-cc-pVDZ	4.83	−3.66	10.34
XSAPT-based methods			
XSAPT(KS)+*ai*D3(ω_{IP})	2.54	−1.93	4.76
XSAPT(KS)+*ai*D3(ω_{GDD})	2.85	−2.27	5.09
sd-XSAPT(KS)(ω_{IP})	1.83	−1.00	2.77
sd-XSAPT(KS)(ω_{GDD})	1.32	−0.39	2.69
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{IP})	1.36	1.20	3.61
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD})	1.19	0.86	3.51
sd-XSAPT(KS)+$E_{disp,3B}^{ATM(TS)}$(ω_{IP})	2.34	2.13	4.52
sd-XSAPT(KS)+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD})	2.94	2.74	5.93
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(Grimme)}$($\omega_{IP}$)	1.66	−0.83	3.04
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(Grimme)}$($\omega_{GDD}$)	1.97	−1.17	3.37
sd-XSAPT(KS)+$E_{disp,3B}^{ATM(Grimme)}$($\omega_{IP}$)	1.09	0.10	3.21
sd-XSAPT(KS)+$E_{disp,3B}^{ATM(Grimme)}$($\omega_{GDD}$)	1.23	0.70	3.42

[a]With respect to CCSD(T)/CBS benchmarks.[4,133] [b]All complexes treated as dimers. [c]The jun-cc-pVDZ basis set is used for (*s*)SAPT0 calculations, hpTZVPP for (X)SAPT(KS+*ai*D3, and 6-31G(d,2p) for sd-XSAPT(KS).

Considering the XSAPT results, we note that both XSAPT(KS)+*ai*D3 and sd-XSAPT(KS) interaction energies are improved, in nearly all cases, by inclusion of either $E_{disp,3B}^{ATM(TS)}$ or $E_{disp,3B}^{ATM(Grimme)}$. The TS version performs better in combination with XSAPT(KS)+*ai*D3 whereas the Grimme version is superior when combined with sd-XSAPT(KS). (However, we will demonstrate below that for S12L complexes the TS version affords better agreement with dispersion energies computed using DFT response theory.) Although we also tested XSAPT(KS)+*ai*D2, the resulting MAEs were found to be at least 0.3 kcal/mol larger than the corresponding +*ai*D3 results, consistent with previous observations that the latter, third-generation approach is generally more accurate for both dispersion energies and total interaction energies.[43]

The best-performing XSAPT methods for L7 are XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD}) and sd-XSAPT(KS)+$E_{disp,3B}^{ATM(Grimme)}$($\omega_{IP}$), with MAEs of 1.2 and 1.1 kcal/mol, respectively. This is comparable to the best supersystem methods tested here, and Figure 2 compares error statistics for some of the better-performing methods (including supersystem methods), as applied to L7. Among XSAPT methods, we favor XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD}) because of the good agreement between $E_{disp,3B}^{ATM(TS)}$ and dispersion energies from DFT response theory; see section 4.2.1. Furthermore, we prefer ω_{GDD}-tuning to ω_{IP}-tuning because the former affords more accurate polarizabilities and thus, presumably, more accurate dispersion energies.[110] In conjunction with XSAPT(KS)+*ai*D3, the largest difference between the two tuning approaches occurs for the largest L7 monomer, circumcoronene, which is interesting in light of the documented dependence of ω_{IP} on system size.[111−115]

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

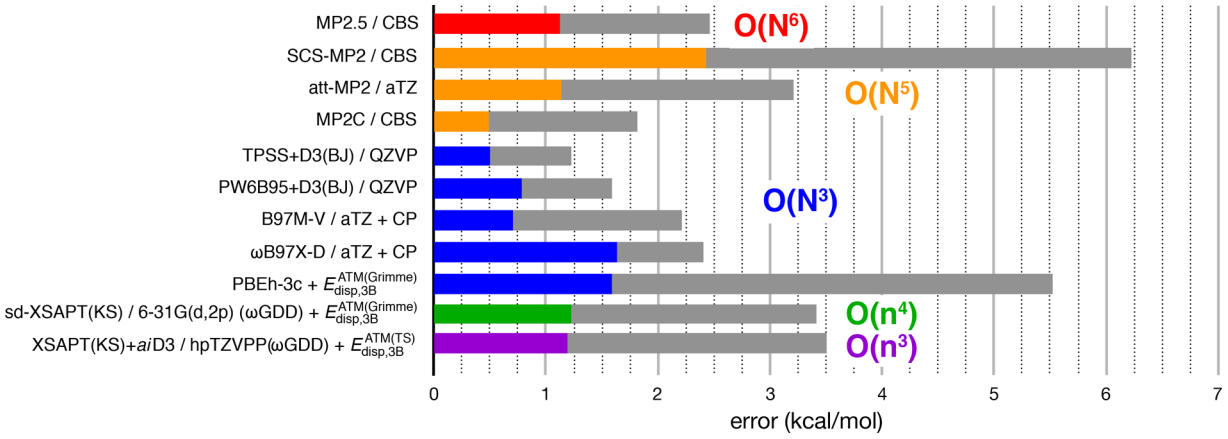


Figure 2. Error statistics for L7 interaction energies, with respect to CCSD(T)/CBS benchmarks, for some of the better-performing methods examined in this work. (Numerical data can be found in Tables 2 and 4.) Gray bars represent maximum errors and colored bars are MAEs, with the latter color-coded according to how the cost of the method scales with system size, with $O(N^3)$ (for example) representing cubic-scaling with respect to supersystem size whereas $O(n^3)$ means cubic scaling with respect to monomer size.

Ratios E_{comp}/E_{int} of various XSAPT energy components (E_{comp}) to the total interaction energy (E_{int}) can be used to categorize the nature of the association interaction.[68] Because the sd-XSAPT(KS)+$E^{ATM(Grimme)}_{disp,3B}(\omega_{IP})$ approach likely benefits from error cancellation, and cannot therefore be recommended for energy decomposition analysis, we compute the energy components at the XSAPT(KS) + *ai*D3+$E^{ATM(TS)}_{disp,3B}(\omega_{GDD})$ level and plot the ratios $-E_{comp}/E_{int}$ in Figure 3. (Numerical values

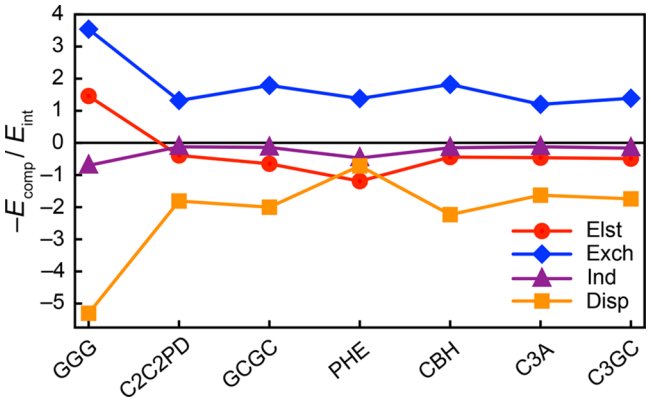


Figure 3. Ratio of various energy components to the total interaction energy, computed at the XSAPT(KS)+*ai*D3+$E^{ATM(TS)}_{disp,3B}(\omega_{GDD})$/hpTZVPP level for the L7 data set.

can be found in the Supporting Information.) According to this analysis, all of the L7 complexes are dispersion-dominated except for the hydrogen-bonded phenylalanine trimer, where the electrostatic interaction is slightly larger than the dispersion interaction. This is consistent with Hobza's classification;[136] see Figure 1a.

In the (X)SAPT calculations discussed thus far, each system is treated as a dimer even though the GGG, C3GC, and PHE complexes contain more than two monomers. We used the π-stacked guanine dimer as one monomer in GGG, the hydrogen-bonded GC base pair as one monomer in both GCGC and C3GC, and phenylalanine dimer as one monomer in PHE. The XSAPT approach affords the possibility to treat each monomer as a separate fragment, and in Table 5 we recompute the XSAPT(KS) + *ai*D3 + $E^{ATM(TS)}_{disp,3B}(\omega_{GDD})$ interaction energies

Table 5. Interaction Energies (in kcal/mol) for Many-Body Systems, Computed at the XSAPT(KS)+*ai*D3+$E^{ATM(TS)}_{disp,3B}(\omega_{GDD})$ Level[a]

	fragmentation scheme	
system	two-body[b]	many-body[c]
GGG	−1.84	−1.92
C3GC	−26.96	−28.32
PHE	−23.22	−23.22
GCGC	−13.74	−14.63

[a]hpTZVPP basis set. [b]System treated as a dimer, as in Table 4. [c]GGG, C3GC, and PHE systems treated as trimers and GCGC as a tetramer.

using this "many-body" fragmentation scheme, comparing it to the dimer fragmentation scheme that was used for the calculations reported in Table 4.

Interaction energies for GGG and PHE are essentially the same in either fragmentation scheme. This makes sense in view of the fact that these two systems are characterized by monomers that are arranged into layers, and the interaction between the second and third layers is only slightly disturbed by the presence of the first layer. The difference between the two- and many-body approaches is larger for GCGC (0.9 kcal/mol) and C3GC (1.4 kcal/mol), which may stem from the assumption that the δE^{HF}_{int} correction is pairwise-additive in many-body systems, and the related fact that the δE^{HF}_{int} correction used in XSAPT is obtained from two-body SAPT without electrostatic embedding, using a dimer-centered basis.[68] Furthermore, the infinite-order response correction for induction is assumed to be included implicitly by the XPol step with negligible double-counting of higher-order corrections for induction.[68] These assumptions appear to be robust, since XSAPT affords accurate interaction energies for systems such as water clusters and halide−water clusters,[43] and moreover the many-body XSAPT(KS)+*ai*D3+$E^{ATM(TS)}_{disp,3B}(\omega_{GDD})$ approach affords about the same MAE for L7 as does the strictly pairwise treatment.

4.2. S12L Data Set. Error statistics for S12L, using a variety of supersystem and (X)SAPT-based methods, are compiled in Table 6 and will be discussed below. Note that the mean deviations in Table 6 are almost uniformly negative, indicating that these methods consistently overbind the S12L dimers. The back-corrected experimental binding energies for S12L are

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

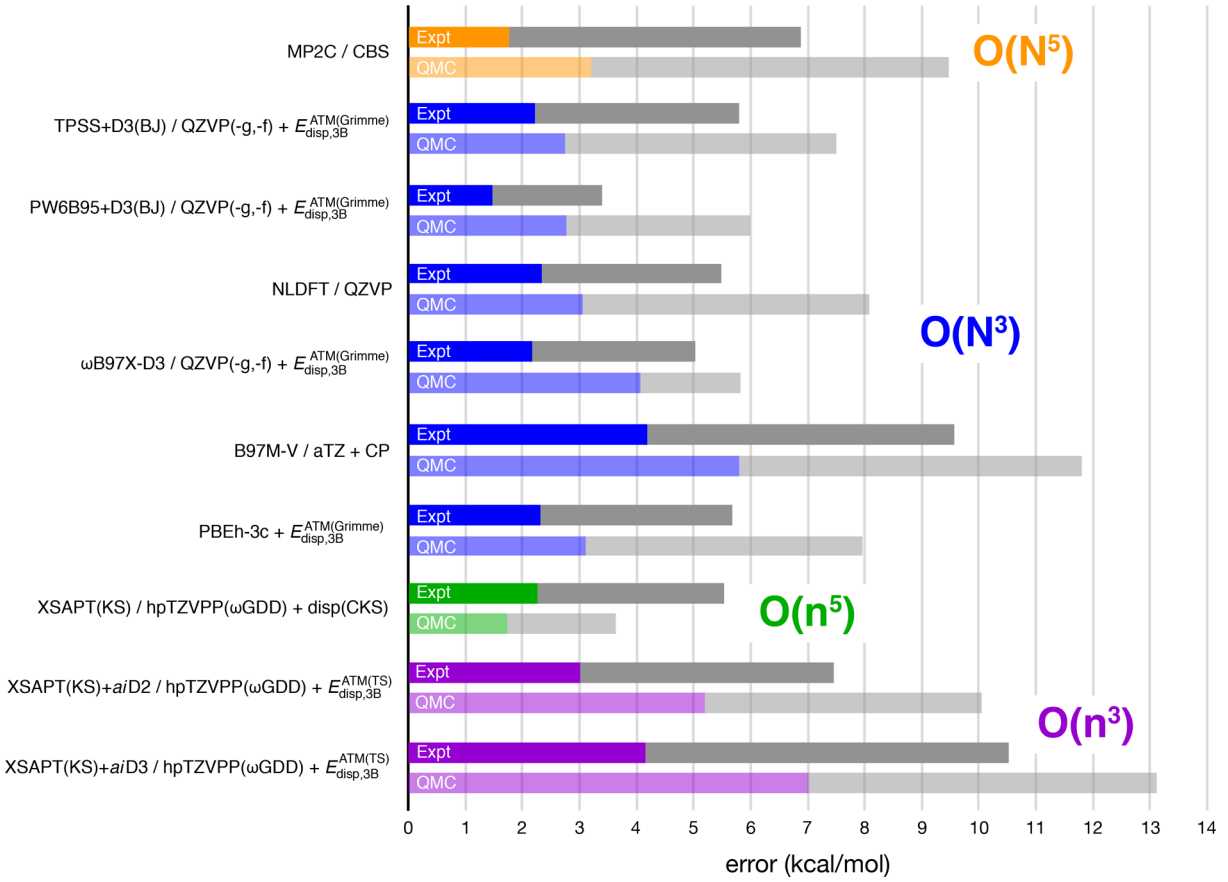


Figure 4. Error statistics for S12L interaction energies for some of the better-performing methods examined in this work. (Numerical data are taken from Table 6.) Gray bars represent maximum errors and colored bars are MAEs, with the latter color-coded according to how the cost of the method scales with system size. Two sets of error statistics are plotted for each method, with the upper (solid color) bar representing errors relative to back-corrected experimental benchmarks and the lower (translucent) bar indicating errors relative to QMC benchmarks.

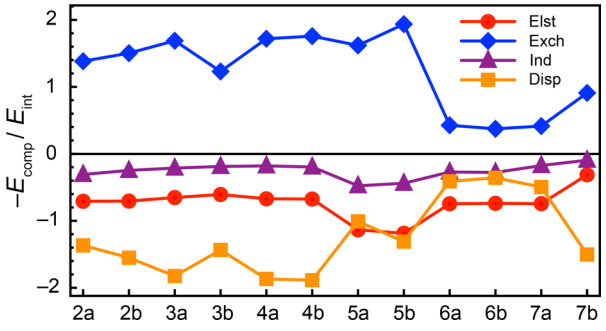


Figure 5. Ratio of each energy contribution relative to the total interaction energy for the S12L data set, computed at the XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}(\omega_{GDD})$/hpTZVPP level.

consistently larger than the QMC ones, by an average of 1.6 kcal/mol, which explains why the errors in Table 6 are slightly larger when compared to QMC results.

Although ∼1 kcal/mol is often cited as the standard for "chemical accuracy", this is unrealistically stringent for interaction energies in large supramolecular complexes, where deviations of 2−3 kcal/mol are to be expected.[181] Estimated uncertainties for the back-corrected experimental binding energies are 2 kcal/mol for the S12L data set,[132,135] while the average statistical error in the QMC benchmarks is 1.2 kcal/mol.[126] S12L error statistics for a variety of quantum chemistry methods can be found in Figure 4. In view of the uncertainties in the benchmarks, the following methods each exhibit errors that are comparable to the anticipated accuracy of the benchmarks: Heßelmann's NLDFT functional,[155] the DFT methods TPSS+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$ and PW6B95+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$, XSAPT(KS)+*ai*D2+$E_{disp,3B}^{ATM(TS)}(\omega_{GDD})$, and XSAPT(KS)+disp(CKS). In the latter approach, "disp(CKS)" means that the second-order SAPT0-style dispersion terms are replaced by dispersion interactions computed using DFT response theory; in other words, this is the XSAPT version of DFT-SAPT.[54,55]

Below, we present a detailed analysis of these results.

4.2.1. Atomic-Pairwise Dispersion Potentials. To address the validity of atomic-pairwise dispersion potentials for S12L complexes, we compare *ai*D2 and *ai*D3 dispersion energies to those computed using DFT response theory (as in DFT-SAPT), with results shown in Table 7. In general, both the TS and Grimme variants of $E_{disp,3B}^{ATM}$ significantly reduce the errors, as compared to the pairwise approach, but the TS correction performs better. Surprisingly, *ai*D2-based potentials give smaller errors than the corresponding *ai*D3 methods, consistent with results for S12L in Figure 4 but in contrast to previous results for S22, S66, and π-stacked nucleobase tetramers.[43] We have noted that the *ai*D2 potential slightly overestimates dispersion energies, especially for π-stacked systems, and that *ai*D3 corrects this artifact,[43] but since $E_{disp,3B}^{ATM(TS)}$ is repulsive it is possible that there is some error cancellation when this three-body term is combined with *ai*D2.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

Table 6. Error Statistics for S12L Interaction Energies, versus Back-Corrected Experimental Benchmarks,[132,134] or QMC Benchmarks.[126]

	error (kcal/mol)					
	experiment			QMC		
method	MAE	MD	max	MAE	MD	max
			supersystem methods			
MP2/CBS[a]	15.90	−15.90	53.28	17.87	−17.87	52.98
SCS-MP2/CBS[a]	6.36	−2.96	26.23	7.19	−5.05	27.65
MP2C/CBS [(DZ,TZ) extrapolation][a,b]	1.77	−0.99	6.87	3.21	−3.21	9.47
MP2C/CBS [1.08×(DZ,TZ)][a,b]	4.37	−4.37	12.61	6.40	−6.40	15.21
NLDFT/def2-QZVP[a]	2.35	−1.08	5.48	3.05	−2.93	8.08
TPSS+D3(BJ)/def2-QZVP(-g,-f)[c,d]	2.62	−2.02	8.30	4.50	−4.33	10.70
TPSS+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-QZVP(-g,-f)[c,d]	2.23	0.16	5.80	2.75	−2.22	7.50
PW6B95+D3(BJ)/def2-QZVP(-g,-f)[c,d]	2.75	−2.35	6.60	4.57	−4.40	9.20
PW6B95+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-QZVP(-g,-f)[c,d]	1.48	−0.18	3.40	2.78	−2.28	6.00
B2PLYP+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[e]	3.62	−0.92	12.33	4.00	−3.89	14.78
B2PLYP+NL+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[e]	5.22	−5.22	17.91	6.74	−6.74	20.00
PBE-XDM/pc-2-spd[f]	2.28	2.00	4.40	1.22	−0.05	2.70
LC-ωPBE-XDM/pc-2-spd[f]	5.61	−5.61	11.70	7.65	−7.65	14.20
ωB97X-D3+$E_{disp,3B}^{ATM(Grimme)}$/def2-QZVP(-g,-f)[c,g]	2.16	−1.29	5.04	4.08	−4.08	5.82
B97M-V/aTZ	6.07	−6.07	12.78	7.51	−7.51	14.74
B97M-V/aTZ (CP)	4.20	−4.17	9.58	5.79	−5.79	11.80
HF-3c	5.39	−3.92	11.38	6.51	−4.90	11.57
HF-3c+$E_{disp,3B}^{ATM(Grimme)}$	3.86	−1.74	8.01	4.73	−2.78	9.21
PBEh-3c	3.38	−2.77	9.18	4.35	−4.35	11.17
PBEh-3c+$E_{disp,3B}^{ATM(Grimme)}$	2.31	−0.60	5.68	3.12	−2.23	7.97
			(X)SAPT-based methods			
DFT-SAPT [(DZ,TZ) extrapolation][a,b,h,i]	2.86	2.13	6.29	1.99	−0.43	4.89
DFT-SAPT [1.08 × (DZ,TZ)][a,b,h,i]	2.10	−1.56	7.64	3.87	−3.87	10.13
XSAPT(KS)+*ai*D2(ω_{IP})[b,j]	9.63	−9.63	18.50	12.19	−12.19	21.10
XSAPT(KS)+*ai*D2(ω_{GDD})[b,j]	8.38	−8.38	18.72	10.70	−10.70	21.10
XSAPT(KS)+*ai*D2+$E_{disp,3B}^{ATM(TS)}$(ω_{IP})[b,j]	3.69	−3.63	10.78	6.66	−6.66	10.78
XSAPT(KS)+*ai*D2+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD}))[b,j]	3.02	−2.37	7.46	5.19	−5.19	10.06
XSAPT(KS)+*ai*D3(ω_{IP}))[b,j]	10.93	−10.93	21.56	14.02	−14.02	24.16
XSAPT(KS)+*ai*D3(ω_{GDD})[b,j]	9.67	−9.67	21.86	12.57	−12.57	24.38
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{IP})[b,j]	5.80	−4.92	10.30	8.49	−8.49	12.90
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD})[b,j]	4.17	−3.66	10.53	7.02	−7.02	13.13
XSAPT(KS)+disp[CKS, (DZ,TZ)](ω_{IP})[b,h,j]	2.24	0.52	7.08	2.50	−2.45	7.08
XSAPT(KS)+disp[CKS, (DZ,TZ)](ω_{GDD})[b,h,j]	2.26	1.77	5.53	1.74	−0.99	3.64
XSAPT(KS)+disp[CKS, 1.08×(DZ,TZ)](ω_{IP})[b,h,j]	3.64	−3.17	10.34	5.89	−5.89	10.34
XSAPT(KS)+disp[CKS, 1.08×(DZ,TZ)](ω_{GDD})[b,h,j]	2.28	−1.92	6.29	4.43	−4.43	8.89

[a]From ref 150. [b]Using the NLDFT deformation energies from ref 150. [c]The "(-g,-f)" notation indicates removal of *f* function on hydrogen and *g* functions on other atoms. [d]From ref 132. [e]From ref 137. [f]From ref 144. [g]From ref 134. [h]Including DFT-SAPT dispersion energies extrapolated to the CBS limit using a (DZ,TZ) scheme.[150] [i]Using cc-pVTZ for the nondispersion terms. [j]Using the hpTZVPP basis set.

Table 7. Errors in the *ai*D2 and *ai*D3 Dispersion Potentials[a] for S12L Complexes

	error (kcal/mol)	
dispersion potential	MAE	max
	*ai*D2 potential	
*ai*D2	6.46	12.43
*ai*D2+$E_{disp,3B}^{ATM(TS)}$	1.18	2.50
*ai*D2+$E_{disp,3B}^{ATM(Grimme)}$	4.39	9.23
	*ai*D3 potential	
*ai*D3	7.75	15.69
*ai*D3+$E_{disp,3B}^{ATM(TS)}$	2.62	4.72
*ai*D3+$E_{disp,3B}^{ATM(Grimme)}$	5.66	12.30

[a]With respect to DFT-SAPT/CBS dispersion energies extrapolated using the 1.08 × (DZ,TZ) scheme of ref 150.

Alternatively, it may be that the DFT-SAPT dispersion benchmarks in Table 7 are not completely converged. These benchmarks use a 1.08 × (DZ,TZ) extrapolation[150] to correct for the absence of diffuse functions, where the scaling factor of 1.08 attempts to compensate for basis-set incompleteness, and was determined using S22 benchmarks.[182] Finally, the quality of DFT-SAPT dispersion benchmarks is also influenced by the fact that the coupled value of $E_{exch\text{-}disp}^{(2)}$ is *not* obtained from DFT response theory but rather via empirical scaling of the uncoupled value.[150] A scaling factor of 0.686 was determined in ref 150 using the S22 data set. Although the fit was quite good (R^2 = 0.9993), it is possible that the relationship between the coupled and uncoupled values of $E_{exch\text{-}disp}^{(2)}$ is different in large molecules where dispersion plays a more prominent role. Thus, the superior performance of *ai*D2 dispersion energies, in comparison to *ai*D3, may partly reflect the quality of the

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

benchmarks. The performance of both *ai*D2 and *ai*D3 combined with XSAPT(KS) will be further examined below.

4.2.2. Supersystem Approaches. Let us now discuss S12L errors for supersystem methods, which are listed in Table 6 and summarized in Figure 4. The MP2/CBS method significantly overestimates the binding energies, as expected, with errors as large as 53 kcal/mol, whereas SCS-MP2/CBS affords a dramatic reduction in the MAE, to 6−7 kcal/mol. (Throughout this discussion, we will quote MAEs as a range that reflects the fact that the putative error depends on which set of benchmarks is selected.) MP2C/CBS achieves MAEs as low as 2−3 kcal/mol, using a (DZ,TZ) extrapolation for the correlation contributions,[150] but when these extrapolated results are scaled by 1.08 as suggested in ref 150, MAEs increase to 4−6 kcal/mol. Note that the S12L benchmarks have been revised since publication of ref 150, and the new benchmarks, as well as results for L7 discussed above, call into question this factor of 1.08.

The semiempirical PBEh-3c method affords MAEs of 3−4 kcal/mol, albeit with somewhat larger outliers as compared to the (vastly more expensive) MP2C/CBS approach. In contrast to results for L7, here the DFT-based PBEh-3c is clearly superior to HF-3c. More importantly, and again in contrast to L7 results, addition of the $E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ improves the PBEh-3c results, reducing the MAE to 2−3 kcal/mol and also reducing the maximum errors, such that the error statistics for the PBEh-3c+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ approach are comparable to, or slightly better than, those for MP2C/CBS, and in view of cost the former therefore seems far preferable. To reiterate: although broken out here, the $E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ correction is included by default in the PBEh-3c method introduced in ref 163 and implemented in the ORCA, TURBOMOLE, and Q-CHEM programs.

The NLDFT functional[155] and the dispersion-corrected TPSS+D3(BJ) and PW6B95+D3(BJ) functionals perform well for S12L, with MAEs of 2−3 kcal/mol with respect to the back-corrected experimental benchmarks. For the two DFT+D approaches, the MAEs are reduced slightly if $E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ is included, and in particular PW6B95+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ appears to be a promising supersystem approach for use in large complexes, with a MAE of only 1.5 kcal/mol. This is significantly better than even the B2PLYP-based double-hybrid results in Table 6.

With the exception of NLDFT, nonlocal functionals generally do not perform as well as those outfitted with empirical dispersion corrections. Counterpoise-corrected B97M-V/aTZ calculations, for example, overbind most of the S12L complexes, with MAEs of 4.2−5.8 kcal/mol that are larger than those observed for L7 complexes. The maximum error is also larger: 9.6 kcal/mol for complex **4b**. Despite the use of a triple-ζ basis set, the counterpoise correction for complex **4b** is 3.2 kcal/mol for B97M-V, and in the absence of counterpoise correction the MAE for B97M-V/aTZ calculations increases to 6.1−7.5 kcal/mol, with a maximum error of 12.8 kcal/mol.

Considering XDM-based methods we find that the favorable performance of LC-ωPBE-XDM for the L7 complexes does not carry over to S12L, and in the latter case this approach exhibits a MAE of 5.6−7.7 kcal/mol. PBE-XDM performs reasonable well for both L7 and S12L, with MAEs of 2−3 kcal/mol in both, and in particular exhibits the smallest MAE with respect to the QMC benchmarks for S12L. However, the strong dependence of the XDM correction on the underlying functional, which has been noted previously,[144] is somewhat bothersome.

4.2.3. (X)SAPT-Based Methods. Turning to SAPT-based results, we note first that the maximum difference between the GDD- and IP-tuned values of ω, namely, $\Delta\omega$ = 0.218 a_0^{-1} for complex **5a**, is much larger than for the L7 data set. Error statistics are again summarized in Table 6 and in Figure 4. Some of these results are labeled "XSAPT(KS)+disp(CKS)", by which we mean that dispersion energies are computed at the coupled Kohn−Sham (CKS) level based on DFT response theory, as in DFT-SAPT, but are then combined with an XSAPT(KS) description of the remaining energy components.

Ratios $E_{\mathrm{comp}}/E_{\mathrm{int}}$ for each energy component are plotted in Figure 5, as computed at the XSAPT(KS)+*ai*-D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}(\omega_{\mathrm{GDD}})$ level. (The raw data can be found in the Supporting Information.) Almost all of the systems are dominated by dispersion except complexes **5a**, **6a**, **6b**, and **7a**, where the dispersion component is large but not dominant. The electrostatic and induction interactions have about the same pattern across all complexes although the mixed-type dimers **5a** and **5b** have slightly larger electrostatic and induction contributions. On the other hand, most complexes have large exchange-repulsion and attractive dispersion, with no clear pattern. The nondispersion part of the interaction energy is positive for all complexes except **6a**, **6b**, and **7b**, whose nondispersion energies are −54.0, −57.2, and −70.5 kcal/mol, respectively. Clearly, accurate description of dispersion interactions is crucial in these complexes.

DFT-SAPT affords a good description of binding energies for S12L (see Table 6), with a MAEs of 2−4 kcal/mol. For reasons of cost, the DFT-SAPT calculations use a (DZ,TZ) extrapolation for the dispersion energy,[150] and in ref 150, it is suggested to multiply the extrapolated energies by an empirical factor of 1.08 to account for basis-set incompleteness. This correction slightly reduces the MAE (with respect to the back-corrected experimental benchmarks) by 0.8 kcal/mol, although the maximum error increases by 1.4 kcal/mol. With respect to the QMC benchmarks, however, empirical scaling seriously degrades the quality of the DFT-SAPT/CBS results, increasing the MAE from 2.0 to 3.9 kcal/mol and increasing the maximum error from 4.9 to 10.1 kcal/mol. In view of this, the empirical scaling factor appears to be ill-advised.

For reasons of cost, we have performed SAPT0/jaDZ calculations on only the two smallest S12L complexes, **2a** and **2b**, obtaining binding energies of −43.3 and −27.2 kcal/mol, respectively. These are too large by 14.3 and 6.4 kcal/mol, respectively, as compared to the back-corrected experimental benchmarks, which is consistent with the overstabilization of the L7 complexes at the SAPT0/jaDZ level (Table 2). Also consistent with L7 results is the fact that the "bronze standard" (empirically scaled) *s*SAPT0/jaDZ approach affords essentially the same binding energies (−43.2 and −27.1 kcal/mol) as *unscaled* SAPT0/jaDZ. Favorable error cancellation between the SAPT0 method and the jaDZ basis set, observed consistently in small dimers,[3,156,157] seems not to extend to these larger systems.

For the L7 dimers, the three-body dispersion correction $E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$ was as large as 6.6 kcal/mol (for C3GC, Table 3) and this correction is even larger for S12L dimers, up to 12.2 kcal/mol for **4b** (Table 8). Incorporating $E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$ into XSAPT(KS)+*ai*D3(ω_{GDD}) reduces the MAE from 9.7 to 4.2 kcal/mol. As for the L7 complexes, XSAPT(KS) results for S12L are superior when using ω_{GDD} tuning as compared to ω_{IP} tuning. To investigate the latter point in more detail, Table 9 compares how the first-order electrostatic and exchange energy

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

Table 8. Three-Body Dispersion Corrections for S12L Complexes

system	energy (kcal/mol)	
	$E_{disp,3B}^{ATM(TS)}$[a]	$E_{disp,3B}^{ATM(Grimme)}$[b]
2a	5.55	1.80
2b	4.04	1.20
3a	7.14	1.80
3b	4.40	0.70
4a	11.26	3.20
4b	12.22	3.50
5a	2.03	1.00
5b	2.11	1.00
6a	3.76	2.20
6b	2.95	1.80
7a	10.06	4.60
7b	6.62	3.30

[a]Computed using XPol at the level of LRC-ωPBE/hpTZVPP with ω_{GDD} tuning. [b]From ref 135.

Table 9. XSAPT(KS) Energy Components for S12L[a] Compared to DFT-SAPT Benchmarks

component	deviation (kcal/mol)			
	ω_{IP}		ω_{GDD}	
	MAE	max	MAE	max
$E_{elst}^{(1)}$	0.70	1.87	0.67	2.07
$E_{exch}^{(1)}$	2.23	6.13	0.77	1.82

[a]LRC-ωPBE functional with two different tuning schemes.

components in XSAPT(KS) deviate from the corresponding DFT-SAPT values.[150] Although $E_{elst}^{(1)}$ is about the same for either tuning scheme, and lies within ~1 kcal/mol of the DFT-SAPT result, ω_{GDD} tuning affords first-order exchange energies that are much more consistent with DFT-SAPT as compared to ω_{IP} tuning. We therefore recommend the ω_{GDD} approach for (X)SAPT calculations that employ LRC functionals.

The XSAPT(KS)+disp(CKS) results in Table 6 use CKS dispersion energies computed using asymptotically spliced functionals,[150] as in DFT-SAPT, and extrapolated to the CBS limit. Using ω_{GDD} tuning for the density functional, we obtain a MAE of 2.3 kcal/mol as compared to the back-corrected experimental benchmarks, which makes the XSAPT(KS)+disp(CKS) superior to XSAPT(KS)+*ai*D3, even when three-body dispersion terms are included in the latter. The accuracy of XSAPT(KS)+disp(CKS) is similar to that of DFT-SAPT itself, which supports the notion that CKS dispersion energies should be more accurate than atom−atom dispersion potentials.

Regarding the empirical scaling factor of 1.08 used to extrapolate the DFT-SAPT dispersion energies,[150] which has been mentioned several times already, we find that XSAPT(KS)+disp(CKS) performs better *without* this factor (MAE of 1.7 kcal/mol as compared to 4.4 kcal/mol, versus QMC benchmarks). For both this method and the closely related DFT-SAPT approach, the mean deviations are negative when the scaling factor is included and positive when it is omitted, suggesting that this scaling factor overcorrects for large systems. The optimal scaling factor for S12L probably lies between 1.00 and 1.08, but the dispersion energies are so large for these systems that this relatively small change can alter the error statistics by several kcal/mol.

Interestingly, all XSAPT(KS)+*ai*D2 methods perform slightly better than the corresponding +*ai*D3 approaches. The best version of the former is XSAPT(KS)+*ai*D2+$E_{disp,3B}^{ATM(TS)}(\omega_{GDD})$, with a MAE of 3.0 kcal/mol.

4.2.4. A Closer Look at a Few Examples. The largest $E_{disp,3B}^{ATM(TS)}$ corrections among the S12L complexes arise from dimers **4a** (11.3 kcal/mol) and **4b** (12.2 kcal/mol), which are the buckycatcher complexes with C_{60} and C_{70}; the smallest corrections are for the amide macrocycle complexes **5a** (2.0 kcal/mol) and **5b** (2.1 kcal/mol). Whether these nonpairwise dispersion corrections are important in a given complex depends upon the shape, topology, and conjugation, and below we analyze two extreme cases: the $C_{60}@C_{60}H_{28}$ buckycatcher/C_{60} complex **4a**, which has been widely used as a stringent test of theoretical methods for describing dispersion,[31,37,126,132,135,137,150,183−185] and the amide macrocycle/benzoquinone complex **5a**, whose intermolecular interactions are classified as being of mixed type.

Binding energies for these two complexes are provided in Table 10, and the methods listed overestimate the binding

Table 10. Binding Energies for S12L Complexes 4a and 5a

method (kcal/mol)	binding energy	
	4a	**5a**
experiment[a]	−28.4	−33.4
QMC[b]	−25.8 ± 1.5	−33.4 ± 1.0
HF-3c+$E_{disp,3B}^{ATM(Grimme)}$	−34.2	−27.6
MP2/CBS[c]	−78.8	−39.5
SCS-MP2/CBS[c]	−53.5	−30.1
MP2C/CBS [(DZ,TZ)][c,d]	−35.3	−35.4
MP2C/CBS [1.08×(DZ,TZ)][c,d]	−41.0	−37.9
B2PLYP+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[e]	−40.6	−33.1
B2PLYP+NL+$E_{disp,3B}^{ATM(Grimme)}$/def2-TZVP[e]	−45.8	−35.2
NLDFT/def2-QZVP[c]	−33.9	−33.0
TPSS+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-QZVP(-g,-f)[a]	−33.3	−31.9
PW6B95+D3(BJ)+$E_{disp,3B}^{ATM(Grimme)}$/def2-QZVP(-g,-f)[a]	−31.8	−31.9
PBE+MBD*[b]	−28.3	−33.8
PBEh-3c + $E_{disp,3B}^{ATM(Grimme)}$	−33.8	−32.4
DFT-SAPT[c,d,f,g]	−35.9	−34.5
XSAPT(KS)+*ai*D3(ω_{GDD})[d,h]	−50.2	−38.2
XSAPT(KS)+*ai*D3+$E_{disp,3B}^{ATM(TS)}$(ω_{GDD})[d,h]	−38.9	−36.2
XSAPT(KS)+disp(CKS)(ω_{GDD})[d,f,h]	−34.7	−37.1
PBE-XDM/pc-2-spd[i]	−27.5	−30.7
B97M-V/aTZ	−40.5	−36.7
B97M-V/aTZ (CP)	−37.6	−35.2

[a]From ref 134. [b]From ref 126. [c]From ref 150. [d]Using NLDFT deformation energies from ref 150. [e]From ref 137. [f]DFT-SAPT dispersion energies [+disp(CKS)] were extrapolated using the 1.08 × (DZ,TZ) scheme of ref 150. [g]Using cc-pVTZ for the nondispersion terms. [h]Using the hpTZVPP basis set. [i]From ref 144.

energy of $C_{60}@C_{60}H_{28}$ by anywhere from 2−53 kcal/mol, as compared to QMC benchmarks. MP2 performs exceptionally poorly, as expected, and the MP2/CBS binding energy is 53 kcal/mol too large, whereas the SCS-MP2/CBS and MP2C/CBS approaches afford binding energies that are overestimated by "only" 28 and 10 kcal/mol, respectively. These errors may be due in part to residual BSSE arising from the (DZ,TZ) extrapolation that is used. MP2 results for the c2c2pd complex show that (aTZ,aQZ) extrapolation affords results much closer

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

to the CBS limit as compared to (DZ,TZ) extrapolation.[158] The 1.08 × (DZ,TZ) extrapolation scheme degrades the performance of MP2C and increases the overestimation to 15 kcal/mol.

The B2PLYP+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ and B2PLYP+NL+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ methods[137] overestimate the binding energy of **4a** by 15 and 20 kcal/mol, respectively. These double-hybrid functionals inherit MP2's uncoupled description of dispersion, which is especially problematic for π-stacked systems. In smaller systems this can be remedied using attenuated MP2 correlation in the double hybrids,[167] but we have not attempted any such calculations here.

As is true generally for S12L, the NLDFT functional, as well as the TPSS+D3(BJ) + $E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ and PW6B95+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ approaches, each performs reasonably well, overestimating the binding energy of **4a** by 6−8 kcal/mol as compared to the QMC benchmark. B97M-V/aTZ, with and without including counterpoise correction, overestimates the binding energy of **4a** by 12 and 15 kcal/mol, respectively. The best methods for **4a** are PBE+MBD* and PBE-XDM, which only overestimate the binding energy by 2.5 and 1.7 kcal/mol, respectively. The DFT-SAPT and XSAPT(KS)+ disp(CKS)-(ω_{GDD}) methods overestimate the binding energy by 10 and 9 kcal/mol, respectively, which is probably at least partially due to uncertainty in the DFT-SAPT dispersion energies at the CBS limit, owing to use of the aforementioned 1.08 × (DZ,TZ) extrapolation procedure.

The $E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$ correction is very important in **4a**, reducing the error in the XSAPT(KS)+*ai*D3 binding energy from 24 to 13 kcal/mol. The remaining 13 kcal/mol may come from the lack of dynamical dielectric screening effects in the dispersion coefficients of the pairwise *ai*D3 dispersion potential.[37] These effects are important in large supramolecular and solid-state calculations, with C_6 coefficients reduced by a factor of 1.6−1.8 in diamond and silicon relative to free atoms, for example.[37,186] The dynamical screening effect in complex **4a**, as computed using the PBE+TS method, is +9 kcal/mol,[37] which comes from the reduction of the atomic C_6 coefficients for C_{60} by a factor of 1.1−1.3 upon formation of the complex.[37] Combining the 9 kcal/mol dynamical screening effect in **4a** with the XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$($\omega_{\mathrm{GDD}}$) binding energy affords −29.9 kcal/mol, which is close to both the QMC and the experimental benchmark.

The mixed-type complex **5a** contains hydrogen bonds between host and guest, although dispersion still plays a prominent role. All of the methods in Table 10 are in good agreement with the QMC benchmark, with the maximum deviation being only 6 kcal/mol, for MP2/CBS. The aforementioned dynamical screening effect for **5a**, computed using the PBE+TS method, is only 0.1 kcal/mol,[37] which is probably the reason for the good performance of XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$($\omega_{\mathrm{GDD}}$) in this case, as compared to the buckycatcher complex **4a**.

Prediction of relative interaction energies just as important as prediction of absolute association energies, and both the QMC and experimental benchmarks suggest that the binding energy of **5a** is at least 5 kcal/mol larger than that of **4a**. Many of the methods listed in Table 10, however, get this ordering wrong (including B97M-V/aTZ with or without counterpoise correction), with MP2-based approaches grossly exaggerating the stability of **4a**. Only the PBE+MBD*, PBE-XDM, and XSAPT(KS)+disp(CKS)(ω_{GDD}) methods predict a significantly larger binding energy for **5a**.

Finally, let us discuss complexes **6a** and **6b** whose binding energies are consistently overestimated by the methods considered here. The guest molecule is a cation in these complexes, and the cation-dipole interaction leads to large errors in continuum solvation corrections, up to 6 kcal/mol according to the estimates in ref 134. This solvation correct is used obtain a "gas phase" binding energy from experimental data, and is likely the reason why **6a** has the largest deviation (3.6 kcal/mol) between the QMC and experimental binding energies. Use of counterions, leading to a system with overall neutral charge, has been shown to improve results as compared to experimental free energies,[179,187−189] and COSMO-RS solvation energies for multiply charged species are also improved by including counterions.[188] The S12L benchmarks used here did not include counterions in the solvation calculations, however.[134]

The XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$($\omega_{\mathrm{GDD}}$) method overestimates the binding energies of **6a** and **6b** by 4.4 and 3.4 kcal/mol, respectively. We have recently observed some problems with XPol charge embedding for systems involving ions, and preliminary results suggest that significantly improved results for ion−molecule binding energies are obtained using Gaussian blurring of the embedding point charges. (This has been seen in other contexts as well.[190]) Here, however, XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$($\omega_{\mathrm{GDD}}$) with Gaussian blurring slightly increases the overestimation of the **6a** and **6b** binding energies, so the point-charge embedding is not the culprit. As compared to the more accurate QMC benchmark, XSAPT(KS)+*ai*-D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}$($\omega_{\mathrm{GDD}}$) overestimates the binding energy for **6a** by 5.6 kcal/mol, and replacing its dispersion potential by disp(CKS) reduces the error to 1.2 kcal/mol. Even in this electrostatically dominated system, an accurate description of dispersion interactions is important if quantitative results are required.

One final point bears mentioning. In XSAPT calculations, the polarized wave function for each fragment is used for the subsequent SAPT calculations, and typically the dipole moment of this polarized wave function is larger than that of the unpolarized one, often significantly. For example, upon polarization by the neighboring charges, the dipole moment of circumcoronene monomer in **c3gc** increases from 0.09 to 2.48 D, and the dipole moment of C_{70} in $C_{70}@C_{60}H_{28}$ increases from 0.01 to 1.39 D. For complexes **6a** and **6b**, the dipole moment for the host molecule increases from approximately zero to almost 7 D, because the guest is a cation. (Dipole moments of the polarized and unpolarized wave functions can be found in the Supporting Information.) The polarized wave functions generated in the XPol step should capture the environmental effects of the surrounding monomers and in principle can be used density-based schemes for describing dispersion, such as the TS vdW model.[26] This should capture some many-body dispersion effects without the need for supersystem calculations. We are currently investigating such an approach.

4.3. Computational Cost. As an example to illustrate the cost of AO-XSAPT calculations, we consider the $C_{60}@C_{60}H_{28}$ complex **4a** (148 atoms) in the hpTZVPP basis set (4592 basis functions). Timing data in are presented in Figure 6 for the three steps in an XSAPT calculation: the self-consistent XPol iterations (46% of total job time, for this example), pseudocanonicalization of the dimer basis (33% of total time), and the AO-based SAPT calculation (20% of total time). In a many-body system, the latter two steps must be

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

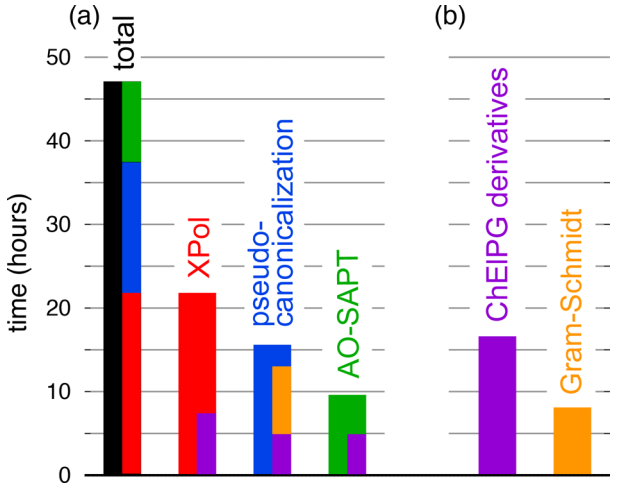


Figure 6. Wall time for specific parts of an XSAPT(KS)+D/hpTZVPP calculation for the $C_{60}@C_{60}H_{28}$ complex **4a**, threaded across 28 cores. (a) Total time (in black), broken down into the required steps for an XSAPT calculation: self-consistent XPol iterations (in red), formation of the dimer basis via pseudocanonicalization (in blue), and finally SAPT calculations (in green). (b) Two significant sources of time, namely, evaluation of the ChElPG charge derivatives $\partial Q_A/\partial P_{\mu\nu}$ (in purple), which is required in order to form the Fock matrix, and Gram−Schmidt orthogonalization of the occupied and virtual orbitals during the pseudocanonicalization step (in orange), where the cost is primarily due to a very large number of matrix multiplications. The cost of each of these operations within the XPol, pseudocanonicalization, and SAPT steps is shown in (a).

performed for each pair of dimers, but these calculations are independent of one another and can be trivially distributed across processors. It is worth noting that the AO-SAPT step is the least time-consuming step, and also does not require significant memory or disk space as compared to an MO implementation of SAPT. The majority (52%) of the pseudocanonicalization time is spent in Gram-Schmidt orthogonalization, which in our implementation is cast in terms of multithreaded matrix multiplications so any further speedup in matrix multiplication—by porting this operation to GPUs, for example—will engender an immediate performance improvement for XSAPT. We mention in this capacity that a GPU implementation of the nondispersion terms in SAPT0 has recently been reported.[59]

Formation of the derivatives $\partial Q_A/\partial P_{\mu\nu}$ of the atomic ChElPG charges Q_A, which are required to form the XPol Fock matrix,[39,81] is clearly a major bottleneck, accounting for 35% of the total job time and contributing to each of the three steps in an AO-XSAPT calculations. ChElPG charges and especially their derivatives are significantly more expensive as compared to Mulliken or Löwdin charges and derivatives, which simply only on the AO overlap matrix, but unfortunately the XPol iterations become unstable in large basis sets (or basis sets containing diffuse functions) when those kind of charges are used.[152] The charge-derivative step represents a clear target for efficiency improvements, which would also enhance the performance of a ChElPG-based Ewald summation technique for QM/MM calculations.[152] Alternative point-charge representations of electrostatics that are tied to the basis functions themselves[191,192] and which do not require evaluation of the electrostatic potential on a real-space grid, represent possible alternatives to ChElPG charges.

5. SUMMARY

5.1. Conclusions. An accurate description of dispersion is important for modeling large supramolecular assemblies, and here, we have examined sizable heterodimers where dispersion interactions make a major contribution to the association energy. To facilitate XSAPT(KS)+D calculations when the monomers are large, we have developed an AO-based implementation of this method that eliminates memory bottlenecks and provides better integral screening as compared to our original MO-based implementation. As found in previous work,[43,67,68] an essential requirement for the SAPT(KS) part of the calculation is monomer-specific tuning of a range-separated hybrid functional. Although we have previously relied on Baer's "optimal tuning" scheme (eq 26),[72] here we find that the GDD tuning procedure (eq 27)[116] is superior. The difference between the two can be significant and we therefore recommend GDD tuning for SAPT(KS) calculations.

For the L7 and S12L data sets, three-body (triatomic) dispersion corrections significantly improve interaction energies computed using XSAPT(KS)+*ai*D3, reducing MAEs from 2.8 to 1.2 kcal/mol for L7 while reducing the maximum error from 5.1 to 3.5 kcal/mol. For S12L, where two sets of benchmarks are available (experimental binding affinities back-corrected to gas-phase interaction energies, and also QMC interaction energies), addition of three-body dispersion corrections to XSAPT(KS)+*ai*D3 reduces the MAE from 9.7 (or 12.6) kcal/mol to 4.2 (or 7.0) kcal/mol, depending on the choice of benchmarks. Although these three-body corrections significantly reduce the outliers for S12L, maximum errors remain large: 10.5 (or 13.1) kcal/mol.

That said, the performance of the XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}(\omega_{\mathrm{GDD}})$ approach is comparable, both in mean and maximum errors, to that of the MP2C/CBS method, at a cost that is rate-limited by the monomer DFT calculations, whereas MP2C scales as $O(N^5)$ with respect to the size of the supersystem. The XSAPT(KS)+disp(CKS) method introduced here exhibits fifth-order scaling with respect to *monomer* size rather than supersystem size. This approach solves CKS equations for dispersion and is the XSAPT analogue of DFT-SAPT;[54,55] it can be viewed, in some sense, as MP2C with monomer correlation effects. This method performs slightly better even than MP2C itself.

Remaining errors in the empirically corrected XSAPT(KS) methods may arise from the absence of dynamical screening effects in the *ai*D3 pairwise dispersion coefficients, which become important when the monomers are large.[92] For example, the interaction energy of the buckycatcher/C_{60} complex computed at the XSAPT(KS)+*ai*D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}(\omega_{\mathrm{GDD}})$ level is −38.9 kcal/mol as compared to a QMC benchmark of −25.8 ± 1.5 kcal/mol.[126] Using the Tkatchenko−Scheffler many-body dispersion approach,[37] we estimate that the dynamical screening affect in this system amounts to a correction of +9 kcal/mol, which would reduce the aforementioned XSAPT binding energy to −29.9 kcal/mol and bring it into good agreement with the QMC benchmark.

In addition to these new XSAPT calculations, this work presents a survey of supersystem results for the L7 and S12L data sets. The recently developed B97M-V functional[158] is a meta-GGA that has shown great promise for noncovalent interactions,[160] and since it does not contain Hartree−Fock exchange it is an attractive option for large systems. Counterpoise-corrected B97M-V/aTZ calculations afford a

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

MAE of 0.7 kcal/mol for L7 and 4.2 (or 5.8) kcal/mol for S12L, albeit with outliers as large as 9.6 (or 11.8) kcal/mol for S12L. The MAEs for S12L are significantly larger than the typical errors reported for B97M-V interaction energies in smaller complexes.[40,43,153,160] Modeling of large supramolecular complexes therefore remains a challenge, even when armed with the latest density functionals.

For L7 and S12L, the two best-performing DFT methods that we have identified are TPSS+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$, with MAEs of 1.1 kcal/mol for L7 and 2.2 (or 2.8) kcal/mol for S12L, and PW6B95+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$, with MAEs of 1.6 kcal/mol for L7 and 1.5 (or 2.8) kcal/mol for S12L. These are unusual functional choices, but ones that were considered already in the original work that introduced S12L.[132] Reasonable performance at significantly lower cost is obtained using the semiempirical PBEh-3c method. (For consistency of notation, we called this approach PBEh-3c+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ in the present work, but the three-body correction is included by default in computer implementations of PBE-3c.[163]) This method affords MAEs of 1.6 kcal/mol for L7 and 2.3 (or 3.1) kcal/mol for S12L. However, the PBEh-3c maximum errors are somewhat larger than those for the other best-performing methods examined here, at 5.5 kcal/mol for L7 and 5.7 (or 8.0) kcal/mol for S12L.

Lastly, the comparison of theoretical methods presented here affords several observations seemingly at odds with the conventional wisdom regarding noncovalent quantum chemistry, which is largely based upon studies of smaller complexes and may therefore need to be revisited. The performance of PBEh-3c is comparable to or better than many of the best wave function and DFT methods examined here, which is something of a pleasant surprise given that noncovalent interactions have historically been challenging for semiempirical quantum chemistry.[4,9] Also notable is the fact that the best-performing DFT methods for small noncovalent complexes are usurped here by less traditional functionals, such as TPSS and PW6B95, especially when these functionals are combined with three-body dispersion corrections. That said, the ωB97X-D3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ approach, where the functional is a reparameterized version[193] of the widely used ωB97X-D functional,[194] also performs well. It is debatable whether it is appropriate to add three-body dispersion corrections to nonlocal correlation functionals such as NLDFT, B97M-V, or double-hybrid functionals,[195] although there have been attempts to do so.[137]

Finally, it is worth noting that discrepancies of ∼2 kcal/mol exist between two sets of "benchmark" S12L interaction energies[126,134,135] and also between several sets of CCSD(T)/CBS benchmarks for L7.[133,137,138] This observation provides some perspective on the precision to which other methods can be assessed based on these data.

5.2. Recommendations. Our recommendations for calculation of supramolecular interaction energies, based on our experience with L7 and S12L at various levels of theory, are summarized in this section. With the exception of Heßelmann's NLDFT,[155] nonlocal density functionals generally do not perform as well as those outfitted with empirical dispersion corrections. Two dispersion-corrected DFT methods, TPSS+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ and PW6B95+D3(BJ)+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$, appear to be reasonably accurate for supramolecular interaction energies, when combined with quadruple-ζ basis sets, in which case explicit counterpoise correction is unnecessary. TPSS is a meta-GGA that does not include Hartree−Fock exchange, so is the more attractive option for large systems, as compared to PW6B95. The cost of these methods is $O(N^3)$.

The only acceptable wave function approach, among those examined here, is MP2C/CBS. This approach is much more expensive than DFT, scaling formally as $O(N^5)$, and its error statistics are actually not better that the two DFT approaches mentioned above. In contrast, the semiempirical HF-3c+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ and PBEh-3c+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(Grimme)}}$ methods are rather attractive options in terms of the price-to-performance ratio. Error statistics for these methods are comparable to the best supersystem approaches examined here, yet at a cost that scales formally as only $O(N^3)$. Moreover, HF-3c requires only a minimal basis set and PBEh-3c a double-ζ basis set, so the actual cost is significantly smaller than traditional DFT, especially in the case of HF-3c.

Regarding (X)SAPT-based methods, the cubic-scaling XSAPT(KS)+aiD2+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}(\omega_{\mathrm{GDD}})$ and XSAPT(KS)+aiD3+$E_{\mathrm{disp,3B}}^{\mathrm{ATM(TS)}}(\omega_{\mathrm{GDD}})$ methods, and the fifth-order scaling XSAPT(KS)+disp(CKS) approach (or DFT-SAPT, for dimers) afford reasonable interaction energies. While certainly more expensive than semiempirical approaches, these methods can actually be less expensive than traditional DFT due to the monomer-based nature of the SCF calculations.[66] As such, they offer an affordable means to understand supramolecular interaction energies in terms of their various energy components.

APPENDIX A: OPEN-SHELL SAPT EQUATIONS

Here, we present open-shell (OS, which is to say spin-unrestricted) analogues of the AO-based SAPT0 electrostatic, exchange, and induction formulas.[97,98,197,198] For the first-order electrostatic energy, the OS version of eq 8 is

$$[E_{\mathrm{elst}}^{(1)}]_{\mathrm{OS}} = \mathrm{tr}[(\mathbf{P}_\alpha^A + \mathbf{P}_\beta^A)\mathbf{V}^B + (\mathbf{P}_\alpha^B + \mathbf{P}_\beta^B)(\mathbf{V}^A + \mathbf{J}_\alpha^A + \mathbf{J}_\beta^A)] + V_0 \tag{A1}$$

which involves spin-density matrices for the monomers, with $\mathbf{P}^{A,B} = \mathbf{P}_\alpha^{A,B} + \mathbf{P}_\beta^{A,B}$. The OS version of the first-order exchange energy in eq 11 is

$$\begin{aligned}[E_{\mathrm{exch}}^{(1)}]_{\mathrm{OS}} = \sum_{\sigma=\alpha,\beta} \mathrm{tr}\{&-\mathbf{P}_\sigma^A\mathbf{K}_\sigma^B + \mathbf{T}_\sigma^A\mathbf{h}_\sigma^B + \mathbf{T}_\sigma^B\mathbf{h}_\sigma^A + \mathbf{T}_\sigma^{AB}\mathbf{h}_\sigma^A + \mathbf{T}_\sigma^{AB}\mathbf{h}_\sigma^B \\ &+ \mathbf{T}_\sigma^A(\mathbf{J}[\mathbf{T}_\sigma^B] - \mathbf{K}[\mathbf{T}_\sigma^B]) + \mathbf{T}_\sigma^{AB}(\mathbf{J}[\mathbf{T}_\sigma^A] - \mathbf{K}[\mathbf{T}_\sigma^A]) \\ &+ \mathbf{T}_\sigma^{AB}(\mathbf{J}[\mathbf{T}_\sigma^B] - \mathbf{K}[\mathbf{T}_\sigma^B]) + \mathbf{T}_\sigma^{AB}(\mathbf{J}[\mathbf{T}_\sigma^{BA}] - \mathbf{K}[\mathbf{T}_\sigma^{BA}])\} \\ &+ \mathbf{T}_\alpha^A\mathbf{J}[\mathbf{T}_\beta^B] + \mathbf{T}_\beta^A\mathbf{J}[\mathbf{T}_\alpha^B] + \mathbf{T}_\alpha^{AB}\mathbf{J}[\mathbf{T}_\beta^A] + \mathbf{T}_\beta^{AB}\mathbf{J}[\mathbf{T}_\alpha^A] \\ &+ \mathbf{T}_\alpha^{AB}\mathbf{J}[\mathbf{T}_\beta^B] + \mathbf{T}_\beta^{AB}\mathbf{J}[\mathbf{T}_\alpha^B] + \mathbf{T}_\alpha^{AB}\mathbf{J}[\mathbf{T}_\beta^{BA}] + \mathbf{T}_\beta^{AB}\mathbf{J}[\mathbf{T}_\alpha^{BA}]\end{aligned} \tag{A2}$$

Relative to the CS case, the matrices $\mathbf{T}^A$, $\mathbf{T}^{AB}$, etc., acquire a spin index, $\sigma = \alpha$ or β. This means that, for example, the matrix $\mathbf{T}^A$ defined in eq 13a should be replaced with

$$\mathbf{T}_\sigma^A = \mathbf{C}_\sigma^A\mathbf{D}_{aa}(\mathbf{C}_\sigma^A)^\dagger \tag{A3}$$

For the first-order exchange energy within the S^2 approximation, we have (cf. eq 17)

$$[E_{\mathrm{exch}}^{(1)}(S^2)]_{\mathrm{OS}} = -\sum_{\sigma=\alpha,\beta} \mathrm{tr}(\mathbf{P}_\sigma^A\mathbf{K}_\sigma^B + \mathbf{O}_\sigma^\dagger\mathbf{h}_\sigma^A + \mathbf{O}_\sigma\mathbf{h}_\sigma^B - \mathbf{P}_\sigma^B\mathbf{S}\mathbf{O}_\sigma\mathbf{\Omega}^A - \mathbf{O}_\sigma\mathbf{S}\mathbf{P}_\sigma^A\mathbf{\Omega}^B + \mathbf{O}_\sigma\mathbf{K}^\dagger[\mathbf{O}_\sigma]) \tag{A4}$$

For second-order induction, the OS version of eq 20 is

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

$$[E^{(2)}_{\mathrm{ind}}]_{\mathrm{OS}} = \mathrm{tr}(\mathbf{X}^A_\alpha \mathbf{\Omega}^B + \mathbf{X}^A_\beta \mathbf{\Omega}^B + \mathbf{X}^B_\alpha \mathbf{\Omega}^A + \mathbf{X}^B_\beta \mathbf{\Omega}^A) \quad \text{(A5)}$$

and finally for exchange-induction the analogue of eq 23 is

$$\begin{aligned}[E^{(2)}_{\text{exch-ind}}(S^2)(A \leftarrow B)]_{\mathrm{OS}} &= -\sum_{\sigma=\alpha,\beta} \mathrm{tr}(\mathbf{X}^A_\sigma \mathbf{K}^B_\sigma + \mathbf{X}^A_\sigma \mathbf{S} \mathbf{P}^B_\sigma \mathbf{h}^A_\sigma + \mathbf{P}^B_\sigma \mathbf{S} \mathbf{X}^A_\sigma \mathbf{h}^B_\sigma - \mathbf{P}^B_\sigma \mathbf{S} \mathbf{X}^A_\sigma \mathbf{S} \mathbf{P}^B_\sigma \mathbf{\Omega}^A \\ &\quad - \mathbf{X}^A_\sigma \mathbf{S} \mathbf{O}^\dagger_\sigma \mathbf{\Omega}^B - \mathbf{O}_\sigma \mathbf{S} \mathbf{X}^A_\sigma \mathbf{\Omega}^B + \mathbf{O}^\dagger_\sigma \mathbf{J}[\mathbf{X}^A_\sigma] - \mathbf{P}^B_\sigma \mathbf{S} \mathbf{O}_\sigma \mathbf{J}[\mathbf{X}^A_\sigma] - \mathbf{X}^A_\sigma \mathbf{K}[\mathbf{O}_\sigma] \\ &\quad + \mathbf{X}^A_\sigma \mathbf{S} \mathbf{P}^B_\sigma \mathbf{K}^\dagger[\mathbf{O}_\sigma] + \mathbf{P}^B_\sigma \mathbf{S} \mathbf{X}^A_\sigma \mathbf{K}[\mathbf{O}_\sigma]) + \mathrm{tr}(\mathbf{O}^\dagger_\alpha \mathbf{J}[\mathbf{X}^A_\beta] + \mathbf{O}^\dagger_\beta \mathbf{J}[\mathbf{X}^A_\alpha] \\ &\quad - \mathbf{P}^B_\alpha \mathbf{S} \mathbf{O}_\alpha \mathbf{J}[\mathbf{X}^A_\beta] - \mathbf{P}^B_\beta \mathbf{S} \mathbf{O}_\beta \mathbf{J}[\mathbf{X}^A_\alpha]) \end{aligned} \quad \text{(A6)}$$

As in the CS case (eq 23), in the context of eq A6 the $B \leftarrow A$ term that completes the exchange-induction energy is defined analogously. "Response" analogues of $E^{(2)}_{\text{exch-ind}}(S^2)$ and $E^{(2)}_{\text{exch-ind}}(S^2)$ are obtained just as in the CS case, by modifying the definition of $\mathbf{U}$ in eq 22 using CPHF amplitudes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b00058.

Tuned ω_{IP} and ω_{GDD} values, total binding energies, interaction energy components and their dependence on basis set, deformation energies, and dipole moments for the L7 and S12L data sets. (PDF)

L7 and S12L data sets (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

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

ORCID

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

Present Address

†Department of Chemistry & Chemical Biology, Cornell University, Ithaca, New York 14853, United States.

Notes

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

■ ACKNOWLEDGMENTS

The authors thank Prof. Stefan Grimme and Dr. Andreas Hansen for providing the L7 binding energies from ref 133, and Prof. Alexandre Tkatchenko for providing a full list of $C^X_{6,\mathrm{free}}$, $\alpha^X_{0,\mathrm{free}}$, and $R^X_{\mathrm{vdW,free}}$ values from ref 26. 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, and calculations were performed at the Ohio Supercomputer Center under project PAA-0003.[196] 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) Kaplan, I. G. *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials*; John Wiley and Sons, Ltd: Chichester, UK, 2006.

(2) Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; Oxford University Press: Oxford, UK, 2013.

(3) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304−326.

(4) Brandenburg, J. G.; Hochheim, M.; Bredow, T.; Grimme, S. Low-cost quantum chemical methods for noncovalent interactions. *J. Phys. Chem. Lett.* **2014**, *5*, 4275−4284.

(5) Brandenburg, J. G.; Gerit, J.; Grimme, S. Dispersion corrected Hartree-Fock and density functional theory for organic crystal structure prediction. In *Prediction and Calculation of Crystal Structures*, Vol. *345*; Atahan-Evrenk, Ş., Aspuru-Guzik, A., Eds.; Springer: Heidelberg, Germany, 2014.

(6) Grimme, S. Dispersion interaction and chemical bonding. In *The Chemical Bond: Chemical Bonding across the Periodic Table*; Frenking, G., Shaik, S., Eds.; Wiley-VCH, 2014.

(7) Beran, G. J. O. A new era for *ab initio* molecular crystal lattice energy prediction. *Angew. Chem., Int. Ed.* **2015**, *54*, 396−398.

(8) Beran, G. J. O. Modeling polymorphic molecular crystals with electronic structure theory. *Chem. Rev.* **2016**, *116*, 5567−5613.

(9) Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical quantum mechanical methods for noncovalent interactions for chemical and biochemical applications. *Chem. Rev.* **2016**, *116*, 5301−5337.

(10) 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.

(11) 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.

(12) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2010**, *82*, 081101.

(13) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

(14) Calbo, J.; Ortí, E.; Sancho-García, J. C.; Aragó, J. The nonlocal correlation density function VV10: A sucessful attempt to accurately capture noncovalent interactions. *Annu. Rep. Comput. Chem.* **2015**, *11*, 37−102.

(15) Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211−228.

(16) Kronik, L.; Tkatchenko, A. Understanding molecular crystals with dispersion-inclusive density functional theory: Pairwise corrections and beyond. *Acc. Chem. Res.* **2014**, *47*, 3208−3216.

(17) Reilly, A. M.; Tkatchenko, A. Van der Waals dispersion interactions in molecular materials: Beyond pairwise additivity. *Chem. Sci.* **2015**, *6*, 3289−3301.

(18) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105−5154.

(19) Goerigk, L. A comprehensive overview of the DFT-D3 London-Dispersion correction. In *Non-Covalent Interactions in Quantum Chemistry and Physics*, 1st ed.; de la Roza, A. O.; DiLabio, G. A., Eds.; Elsevier: Oxford, UK, 2017.

(20) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction. *J. Chem. Phys.* **2005**, *122*, 154104.

(21) Johnson, E. R.; Becke, A. D. A post-Hartree−Fock model of intermolecular interactions. *J. Chem. Phys.* **2005**, *123*, 024101.

(22) Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *J. Chem. Phys.* **2005**, *123*, 154101.

(23) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction: High-order dispersion coefficients. *J. Chem. Phys.* **2006**, *124*, 014104.

(24) Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction revisited. *J. Chem. Phys.* **2007**, *127*, 154108.

(25) Otero-de-la-Roza, A.; Johnson, E. R. Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

(26) 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.

(27) Marom, N.; Tkatchenko, A.; Rossi, M.; Gobre, V. V.; Hod, O.; Scheffler, M.; Kronik, L. Dispersion interactions with density-functional theory: Benchmarking semiempirical and interatomic

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

pairwise corrected density functionals. *J. Chem. Theory Comput.* **2011**, *7*, 3944−3951.
(28) Tkatchenko, A.; Ambrosetti, A.; DiStasio, R. A., Jr. Interatomic methods for the dispersion energy derived from the adiabatic connection fluctuation-dissipation theorem. *J. Chem. Phys.* **2013**, *138*, 074106.
(29) Tao, J.; Perdew, J. P.; Ruzsinszky, A. Accurate van der Waals coefficients from density functional theory. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 18−21.
(30) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.
(31) 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.
(32) von Lilienfeld, O. A.; Tkatchenko, A. Two- and three-body interatomic dispersion energy contributions to binding in molecules and solids. *J. Chem. Phys.* **2010**, *132*, 234109.
(33) Agrawal, P.; Tkatchenko, A.; Kronik, L. Pair-wise and many-body dispersive interactions coupled to an optimally tuned range-separated hybrid functional. *J. Chem. Theory Comput.* **2013**, *9*, 3473−3478.
(34) 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.
(35) Tkatchenko, A.; DiStasio, R. A., Jr.; Car, R.; Scheffler, M. Accurate and efficient method for many-body van der Waals interactions. *Phys. Rev. Lett.* **2012**, *108*, 236402.
(36) DiStasio, R. A., Jr.; von Lilienfeld, O. A.; Tkatchenko, A. Collective many-body van der Waals interactions in molecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14791−14795.
(37) Tkatchenko, A.; Alfe, D.; Kim, K. S. First-principles modeling of non-covalent interactions in supramolecular systems: The role of many-body effects. *J. Chem. Theory Comput.* **2012**, *8*, 4317−4322.
(38) Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A., Jr.; Tkatchenko, A. Long-range correlation energy calculated from coupled atomic response functions. *J. Chem. Phys.* **2014**, *140*, 18A508.
(39) 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.
(40) 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.
(41) Sinnokrot, M. O.; Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *J. Phys. Chem. A* **2004**, *108*, 10200−10207.
(42) Cybulski, S. M.; Lytle, M. L. The origin of the deficiency of the supermolecule second-order Møller−Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*, 141102.
(43) 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.
(44) Szabo, A.; Ostlund, N. S. The correlation energy in the random phase approximation: Intermolecular forces between closed-shell systems. *J. Chem. Phys.* **1977**, *67*, 4351−4360.
(45) Grimme, S. Improved second-order Møller−Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095−9102.
(46) Goldey, M.; Head-Gordon, M. Attenuating away the errors in inter- and intramolecular interactions from second-order Møller−Plesset calculations in the small aug-cc-pVDZ basis set. *J. Phys. Chem. Lett.* **2012**, *3*, 3592−3598.
(47) Heßelmann, A. Improved supermolecular second order Møller−Plesset intermolecular interaction energies using time-dependent density functional response theory. *J. Chem. Phys.* **2008**, *128*, 144112.
(48) Pitoňák, M.; Heßelmann, A. Accurate intermolecular interaction energies from a combination of MP2 and TDDFT response theory. *J. Chem. Theory Comput.* **2010**, *6*, 168−178.
(49) Řezáč, J.; Hobza, P. Describing noncovalent interactions beyond the common approximations: How accurate is the "gold standard", CCSD(T) at the complete basis set limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151−2155.
(50) Šimová, L.; Řezáč, J.; Hobza, P. Convergence of the interaction energies in noncovalent complexes in the coupled-cluster methods up to full configuration interaction. *J. Chem. Theory Comput.* **2013**, *9*, 3420−3428.
(51) Jeziorski, B.; Moszynski, R.; Ratkiewicz, A.; Rybak, S.; Szalewicz, K.; Williams, H. L. SAPT: A program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies. In *Methods and Techniques in Computational Chemistry: METECC-94*, Vol. *B*; Clementi, E., Ed.; STEF: Cagliari, 1993.
(52) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.
(53) Szalewicz, K.; Patkowski, K.; Jeziorski, B. Intermolecular interactions via perturbation theory: From diatoms to biomolecules. In *Intermolecular Forces and Clusters II*, Vol. *116*; Wales, D. J., Eds.; Springer-Verlag: Berlin, 2005.
(54) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254−272.
(55) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2014**, *4*, 127−144.
(56) 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.
(57) 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.
(58) 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.
(59) Parrish, R. M.; Thompson, K. C.; Martínez, T. J. Large-scale functional group symmetry-adapted perturbation theory of graphical processing units. *J. Chem. Theory Comput.* **2018**, *14*, 1737−1753.
(60) Maurer, S. A.; Beer, M.; Lambrecht, D. S.; Ochsenfeld, C. Linear-scaling symmetry-adapted perturbation theory with scaled dispersion. *J. Chem. Phys.* **2013**, *139*, 184104.
(61) Wang, W.; Zhang, Y.; Wang, Y.-B. Noncovalent π-π interaction between graphene and aromatic molecule: Structure, energy, and nature. *J. Chem. Phys.* **2014**, *140*, 094302.
(62) Wang, W.; Sun, T.; Zhang, Y.; Wang, Y.-B. Substituent effects in the π-π interaction between graphene and benzene: An indication for the noncovalent functionalization of graphene. *Comput. Theor. Chem.* **2014**, *1046*, 64−69.
(63) 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.
(64) Korona, T. A coupled cluster treatment of intramonomer electron correlation within symmetry-adapted perturbation theory: Benchmark calculations and a comparison with a density-functional theory description. *Mol. Phys.* **2013**, *111*, 3705−3715.
(65) Heßelmann, A. Comparison of intermolecular interaction energies from SAPT and DFT including empirical dispersion contributions. *J. Phys. Chem. A* **2011**, *115*, 11321−11330.
(66) 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.
(67) 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.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

(68) 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; **2014**, *140*, 119901.
(69) Jansen, G.; Heßelmann, A. Comment on "Using Kohn-Sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interactions. *J. Phys. Chem. A* **2001**, *105*, 11156−11157.
(70) Williams, H. L.; Chabalowski, C. F. Using Kohn-Sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interactions. *J. Phys. Chem. A* **2001**, *105*, 646−659.
(71) Misquitta, A. J.; Szalewicz, K. Intermolecular forces from asymptotically corrected density functional description of monomers. *Chem. Phys. Lett.* **2002**, *357*, 301−306.
(72) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85−109.
(73) Longuet-Higgins, H. C. Intermolecular forces. *Discuss. Faraday Soc.* **1965**, *40*, 7−18.
(74) Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. Dispersion energy from density-function theory description of monomers. *Phys. Rev. Lett.* **2003**, *91*, 033201.
(75) van Leeuwen, R.; Baerends, E. J. Exchange-correlation potential with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *49*, 2421−2431.
(76) Tozer, D. J.; Handy, N. C. Improving virtual Kohn-Sham orbitals and eigenvalues: Application to excitation energies and static polarizabilities. *J. Chem. Phys.* **1998**, *109*, 10180−10189.
(77) Grüning, M.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. Shape corrections to exchange-correlation potentials by gradient-regulated seamless connection of model potentials for inner and outer region. *J. Chem. Phys.* **2001**, *114*, 652−660.
(78) Lotrich, V. F.; Szalewicz, K. Symmetry-adapted perturbation theory of three-body nonadditivity of intermolecular interaction energy. *J. Chem. Phys.* **1997**, *106*, 9668−9687.
(79) Lotrich, V. F.; Szalewicz, K. Perturbation theory of three-body exchange nonadditivity and application to helium trimer. *J. Chem. Phys.* **2000**, *112*, 112−121.
(80) Podeszwa, R.; Szalewicz, K. Three-body symmetry-adapted perturbation theory based on Kohn-Sham description of the monomers. *J. Chem. Phys.* **2007**, *126*, 194101.
(81) 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.
(82) 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.
(83) 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.
(84) 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.; Pan Stanford: Boca Raton, FL, 2016.
(85) Wu, Q.; Yang, W. Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.* **2002**, *116*, 515−524.
(86) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(87) Grimme, S. Semiempirical GGA-type density functional constructed with long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(88) Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. Dispersionless density functional theory. *Phys. Rev. Lett.* **2009**, *103*, 263201.
(89) Podeszwa, R.; Szalewicz, K. Communication: Density functional theory overcomes the failure of predicting intermolecular interaction energies. *J. Chem. Phys.* **2012**, *136*, 161102.
(90) In earlier work,[43,57] we referred to the "D3" potential simply as "*ai*D3". We subsequently spent a considerable amount of time explaining to others that it was not Grimme's D3 (ref 31), hence the name change going forward.
(91) 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.
(92) Dobson, J. F. Beyond pairwise additivity in London dispersion interactions. *Int. J. Quantum Chem.* **2014**, *114*, 1157−1161.
(93) Otero-de-la-Roza, A.; Johnson, E. R. A benchmark for non-covalent interactions in solids. *J. Chem. Phys.* **2012**, *137*, 054103.
(94) Reilly, A. M.; Tkatchenko, A. Seamless and accurate modeling of organic molecular materials. *J. Phys. Chem. Lett.* **2013**, *4*, 1028−1033.
(95) Heßelmann, A. PhD Thesis, Universität Duisburg-Essen, Germany, 2003.
(96) Heßelmann, A.; Jansen, G.; Schütz, M. Density-functional theory symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.
(97) Beer, M. Effiziente "ab-initio" Berechnung molekularer Eigenschaften großer Systeme. Thesis, Ludwig-Maximilians-Universität (LMU), Munich, Germany, 2011.
(98) Hapka, M.; Zuchowski, P. S.; Szczęśniak, M. M.; Chałasiński, G. Symmetry-adapted perturbation theory based on unrestricted Kohn-Sham orbitals for high-spin open-shell van der Waals complexes. *J. Chem. Phys.* **2012**, *137*, 164104.
(99) The $\mathbf{T}^{AB}(2\,\mathbf{J}[\mathbf{T}^{BA}] - \mathbf{K}[\mathbf{T}^{BA}])$ term in eq 11 was mistyped as $\mathbf{T}^{AB}(2\,\mathbf{J}[\mathbf{T}^{AB}] - \mathbf{K}[\mathbf{T}^{AB}])$ in ref 98.
(100) Jeziorski, B.; Bulski, M.; Piela, L. First-order perturbation treatment of the short-range repulsion in a system of many closed-shell atoms or molecules. *Int. J. Quantum Chem.* **1976**, *10*, 281−297.
(101) Schäffer, R.; Jansen, G. Intermolecular exchange-induction energies without overlap expansion. *Theor. Chem. Acc.* **2012**, *131*, 1235.
(102) Schäffer, R.; Jansen, G. Single-determinant-based symmetry-adapted perturbation theory without single-exchange approximation. *Mol. Phys.* **2013**, *111*, 2570−2584.
(103) 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.
(104) The $\mathbf{X}^{A}\mathbf{K}[\mathbf{O}]$ term in eq 23 was mistyped as $\mathbf{X}^{A}\mathbf{K}[\mathbf{O}^{\dagger}]$ in ref 97.
(105) Moszynski, R.; Jeziorski, B.; Ratkiewicz, A.; Rybak, S. Many-body perturbation theory of electrostatic interactions between molecules: Comparison with full configuration interaction for fourelectron dimers. *J. Chem. Phys.* **1993**, *99*, 8856−8869.
(106) Moszyński, R.; Cybulski, S. M.; Chałasiński, G. Many-body theory of intermolecular induction interactions. *J. Chem. Phys.* **1994**, *100*, 4998−5010.
(107) Perdew, J. P.; Burke, K. Comparison shopping for a gradient-corrected density functional. *Int. J. Quantum Chem.* **1996**, *57*, 309−319.
(108) Stein, T.; Kronik, L.; Baer, R. Prediction of charge-transfer excitations in coumarin-based dyes using a range-separated functional tuned from first principles. *J. Chem. Phys.* **2009**, *131*, 244119.
(109) Tsai, C.-W.; Su, Y.-C.; Li, G.-D.; Chai, J.-D. Assessment of density functional methods with correct asymptotic behavior. *Phys. Chem. Chem. Phys.* **2013**, *15*, 8352−8361.
(110) Hapka, M.; Rajchel, L.; Modrzejewski, M.; Chałasiński, G.; Szczęśniak, M. M. Tuned range-separated hybrid functionals in the symmetry-adapted perturbation theory. *J. Chem. Phys.* **2014**, *141*, 134120.
(111) Körzdörfer, T.; Sears, J. S.; Sutton, C.; Brédas, J.-L. Long-range corrected hybrid functionals for π-conjugated systems: Dependence of the range-separation parameter on conjugation length. *J. Chem. Phys.* **2011**, *135*, 204107.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

(112) Karolewski, A.; Kronik, L.; Kümmel, S. Using optimally tuned range separated hybrid functionals in ground-state calculations: Consequences and caveats. *J. Chem. Phys.* **2013**, *138*, 204115.
(113) Körzdörfer, T.; Brédas, J.-L. Organic electronic materials: Recent advances in the DFT description of the ground and excited states using tuned range-separated hybrid functionals. *Acc. Chem. Res.* **2014**, *47*, 3284−3291.
(114) Uhlig, F.; Herbert, J. M.; Coons, M. P.; Jungwirth, P. Optical spectroscopy of the bulk and interfacial hydrated electron from ab initio calculations. *J. Phys. Chem. A* **2014**, *118*, 7507−7515.
(115) Garrett, K.; Sosa Vazquez, X. A.; Egri, S. B.; Wilmer, J.; Johnson, L. E.; Robinson, B. H.; Isborn, C. M. Optimum exchange for calculation of excitation energies and hyperpolarizabilities of organic electro-optic chromophores. *J. Chem. Theory Comput.* **2014**, *10*, 3821−3831.
(116) 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.
(117) 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.
(118) The value $\omega = 0.3\ a_0^{-1}$ is used in the LRC-ωPBE functional to determine the self-consistent density. This density is then used to compute the value of ω_{GDD} according to eq 27, and then ω_{GDD} is used in the LRC-ωPBE functional for the XSAPT(KS) calculations. We find that the value of ω_{GDD} is not very sensitive (within the range 0.2−0.3 a_0^{-1}, for example) to the value of ω that is used to determined the self-consistent density.
(119) Parrish, R. M.; Hohenstein, E. G.; Sherrill, C. D. Tractability gains in symmetry-adapted perturbation theory including coupled double excitations: CCD + ST(CCD) dispersion with natural orbital truncations. *J. Chem. Phys.* **2013**, *139*, 174102.
(120) Otero-de-la-Roza, A.; Johnson, E. R. Many-body dispersion interactions from the exchange-hole dipole moment model. *J. Chem. Phys.* **2013**, *138*, 054103.
(121) Axilrod, B. M.; Teller, E. Interaction of the van der Waals type between three atoms. *J. Chem. Phys.* **1943**, *11*, 299−301.
(122) Kennedy, M. R.; McDonald, A. R.; DePrince, A. E., III; Marshall, M. S.; Podeszwa, R.; Sherrill, C. D. Communication: Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory. *J. Chem. Phys.* **2014**, *140*, 121104.
(123) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chem. Acc.* **1977**, *44*, 129−138.
(124) Kannemann, F. O.; Becke, A. D. Atomic volumes and polarizabilities in density-functional theory. *J. Chem. Phys.* **2012**, *136*, 034109.
(125) Midzuno, Y.; Kihara, T. Non-additive intermolecular potential in gases I. van der waals interactions. *J. Phys. Soc. Jpn.* **1956**, *11*, 1045−1049.
(126) Ambrosetti, A.; Alfè, D.; DiStasio, R. A., Jr.; Tkatchenko, A. Hard numbers for large molecules: Toward exact energetics for supramolecular systems. *J. Phys. Chem. Lett.* **2014**, *5*, 849−855.
(127) Adamovic, I.; Gordon, M. S. Dynamic polarizability, dispersion coefficient C_6 and dispersion energy in the effective fragment potential method. *Mol. Phys.* **2005**, *103*, 379−387.
(128) 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.
(129) Řezác, 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; **2014**, *10*, 1359−1360.
(130) Řezáč, J.; Huang, Y.; Hobza, P.; Beran, G. J. O. Benchmark calculations of three-body intermolecular interactions and the performance of low-cost electronic structure methods. *J. Chem. Theory Comput.* **2015**, *11*, 3065−3079.
(131) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364−3374.
(132) Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. - Eur. J.* **2012**, *18*, 9955−9964.
(133) Hansen, A.; Riplinger, C.; Neese, F.; Grimme, S. Private communication.
(134) Sure, R.; Grimme, S. Comprehensive benchmark of association (free) energies of realistic host-guest complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785−3801.
(135) Risthaus, T.; Grimme, S. Benchmarking of London dispersion-accounting density functional theory methods on very large molecular complexes. *J. Chem. Theory Comput.* **2013**, *9*, 1580−1591.
(136) Grafova, L.; Pitonak, M.; Řezác, J.; Hobza, P. Comparative study of selected wave function and density functional methods for noncovalent interaction energy calculations using the extended S22 data set. *J. Chem. Theory Comput.* **2010**, *6*, 2365−2376.
(137) Calbo, J.; Ortí, E.; Sancho-García, J. C.; Aragó, J. Accurate treatment of large supramolecular complexes by double-hybrid density functionals coupled with nonlocal van der Waals corrections. *J. Chem. Theory Comput.* **2015**, *11*, 932−939.
(138) Pavošević, F.; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps−A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. *J. Chem. Phys.* **2017**, *146*, 174108.
(139) Kruse, H.; Mladek, A.; Gkionis, K.; Hansen, A.; Grimme, S.; Sponer, J. Quantum chemical benchmark study on 46 RNA backbone families using a dinucleotide unit. *J. Chem. Theory Comput.* **2015**, *11*, 4972−4991.
(140) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.
(141) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.
(142) Sparta, M.; Neese, F. Chemical applications carried out by local pair natural orbital based coupled-cluster methods. *Chem. Soc. Rev.* **2014**, *43*, 5032−5041.
(143) 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.
(144) Otero-de-la-Roza, A.; Johnson, E. R. Predicting energetics of supramolecular systems using the XDM dispersion model. *J. Chem. Theory Comput.* **2015**, *11*, 4033−4040.
(145) Christensen, A. S.; Elstner, M.; Cui, Q. Improving intermolecular interactions in DFTB3 using extended polarization from chemical-potential equalization. *J. Chem. Phys.* **2015**, *143*, 084123.
(146) Klamt, A. The COSMO and COSMO-RS solvation models. *WIREs Comput. Mol. Sci.* **2011**, *1*, 699−709.
(147) Wolinski, K.; Pulay, P. Second-order Møller−Plesset calculations with dual basis sets. *J. Chem. Phys.* **2003**, *118*, 9497−9503.
(148) Steele, R. P.; DiStasio, R. A., Jr.; Shao, Y.; Kong, J.; Head-Gordon, M. Dual-basis second-order Møller−Plesset perturbation theory: A reduced-cost reference for correlation calculations. *J. Chem. Phys.* **2006**, *125*, 074108.
(149) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, 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.
(150) Heßelmann, A.; Korona, T. Intermolecular symmetry-adapted perturbation theory study of large organic complexes. *J. Chem. Phys.* **2014**, *141*, 094107.
(151) 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.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

(152) Holden, Z. C.; Richard, R. M.; Herbert, J. M. Periodic boundary conditions for QM/MM calculations: Ewald summation for extended Gaussian basis sets. *J. Chem. Phys.* **2013**, *139*, 244108; erratum: **2015**, *142*, 059901.
(153) Mardirossian, N.; Head-Gordon, M. ωB97X-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.
(154) Dasgupta, S.; Herbert, J. M. Standard grids for high-precision integration of modern density functionals: SG-2 and SG-3. *J. Comput. Chem.* **2017**, *38*, 869−882.
(155) Heßelmann, A. Assessment of a nonlocal correction scheme to semilocal density functional theory methods. *J. Chem. Theory Comput.* **2013**, *9*, 273−283.
(156) 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.
(157) Hohenstein, E. G.; Sherrill, C. D. Density fitting of intramonomer correlation effects in symmetry adapted perturbation theory. *J. Chem. Phys.* **2010**, *133*, 014101.
(158) 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.
(159) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
(160) 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.
(161) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. A standard grid for density-functional calculations. *Chem. Phys. Lett.* **1993**, *209*, 506−512.
(162) Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672−1685.
(163) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **2015**, *143*, 054107.
(164) Turney, J. M.; et al. Psi4: An open-source *ab initio* electronic structure program. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556−565.
(165) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73−78.
(166) Shao, Y.; et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184−215.
(167) Goldey, M.; Head-Gordon, M. Separate electronic attenuation allowing a spin-component-scaled second-order Møller−Plesset theory to be effective for both thermochemistry and noncovalent interactions. *J. Phys. Chem. B* **2014**, *118*, 6519−6525.
(168) Bucko, T.; Lebegue, S.; Angyan, J. G.; Hafner, J. Extending the applicability of the Tkatchenko-Scheffler dispersion correction via iterative Hirshfeld partitioning. *J. Chem. Phys.* **2014**, *141*, 034114.
(169) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. Critical analysis and extension of the Hirshfeld atoms in molecules. *J. Chem. Phys.* **2007**, *126*, 144111.
(170) Li, A.; Muddana, H. S.; Gilson, M. K. Quantum mechanical calculation of noncovalent interactions: A large-scale evaluation of PMx, DFT, and SAPT approaches. *J. Chem. Theory Comput.* **2014**, *10*, 1563−1575.
(171) DiStasio, R. A., Jr.; Head-Gordon, M. Optimized spin-component scaled second-order Møller−Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007**, *105*, 1073−1083.
(172) Huang, Y.; Shao, Y.; Beran, G. J. O. Accelerating MP2C dispersion corrections for dimers and molecular crystals. *J. Chem. Phys.* **2013**, *138*, 224112.
(173) Goldey, M.; Dutoi, A.; Head-Gordon, M. Attenuated second-order Møller−Plesset perturbation theory: Performance within the aug-cc-pVTZ basis. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15869−15875.
(174) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(175) 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.
(176) Anacker, T.; Friedrich, J. New accurate benchmark energies for large water clusters: DFT is better than expected. *J. Comput. Chem.* **2014**, *35*, 634−643.
(177) Mladek, A.; Krepl, M.; Svozil, D.; Cech, P.; Otyepka, M.; Banas, P.; Zgarbova, M.; Jurecka, P.; Sponer, J. Benchmark quantum-chemical calculations on a complete set of rotameric families of the DNA sugar-phosphate backbone and their comparison with modern density functional theory. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7295−7310.
(178) Steinmetz, M.; Grimme, S. Benchmark study of the performance of density functional theory for bond activations with (Ni,Pd)-based transition-metal catalysts. *ChemistryOpen* **2013**, *2*, 115−124.
(179) Antony, J.; Sure, R.; Grimme, S. Using dispersion-corrected density functional theory to understand supramolecular binding thermodynamics. *Chem. Commun.* **2015**, *51*, 1764−1774.
(180) Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031−16042.
(181) Steinmetz, M.; Hansen, A.; Ehrlich, S.; Risthaus, T.; Grimme, S. Accurate thermochemistry for large molecules with modern density functionals. *Top. Curr. Chem.* **2014**, *365*, 1−24.
(182) Řezác, 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.
(183) Zhao, Y.; Truhlar, D. G. Computational characterization and modeling of buckyball tweezers: Density functional study of concave-convex π-π interactions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2813−2818.
(184) Muck-Lichtenfeld, C.; Grimme, S.; Kobryn, L.; Sygula, A. Inclusion complexes of buckycatcher with C_{60} and C_{70}. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7091−7097.
(185) Podeszwa, R.; Cencek, W.; Szalewicz, K. Efficient calculations of dispersion energies for nanoscale systems from coupled density response functions. *J. Chem. Theory Comput.* **2012**, *8*, 1963−1969.
(186) Zhang, G.-X.; Tkatchenko, A.; Paier, J.; Appel, H.; Scheffler, M. van der Waals interactions in ionic and semiconductor solids. *Phys. Rev. Lett.* **2011**, *107*, 245501.
(187) Iali, W.; Petrovic, P.; Pfeffer, M.; Grimme, S.; Djukic, J.-P. The inhibition of iridium-promoted water oxidation catalysis (WOC) by cucurbit[n]urils. *Dalton Trans.* **2012**, *41*, 12233−12243.
(188) Ehrlich, S.; Moellmann, J.; Grimme, S. Dispersion-corrected density functional theory for aromatic interactions in complex systems. *Acc. Chem. Res.* **2013**, *46*, 916−926.
(189) 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.
(190) Tempkin, J. O. B.; Leverentz, H. R.; Wang, B.; Truhlar, D. G. Screened electrostatically embedded many-body method. *J. Phys. Chem. Lett.* **2011**, *2*, 2141−2144.
(191) Stone, A. J. Distributed multipole analysis: Stability for large basis sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.
(192) Gilbert, A. T. B.; Gill, P. M. W. A point-charge model for electrostatic potentials based on a local projection of multipole moments. *Mol. Simul.* **2006**, *32*, 1249−1253.
(193) 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.
(194) 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.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978

(195) The derivation of the VV10 nonlocal correlation functional is based on a two-body *ansatz*,[13] and it is somewhat unclear what this means when VV10 is used in an empirically parametrized density functional such as B97M-V. Calbo et al.[14,137] suggest that since VV10 fails to capture many-body dispersion, addition of explicit three-body $E_{\mathrm{disp,3B}}^{\mathrm{ATM}}$ terms is not inconsistent even in such cases.
(196) Ohio Supercomputer Center. http://osc.edu/ark:/19495/f5s1ph73.
(197) Zuchowski, P. S.; Podeszwa, R.; Moszynski, R.; Jeziorski, B.; Szalewicz, K. Symmetry-adapted perturbation theory utilizing density functional description of monomers for high-spin open-shell complexes. *J. Chem. Phys.* **2008**, *129*, 084101.
(198) Gonthier, J. F.; Sherrill, C. D. Density-fitted open-shell symmetry-adapted perturbation theory and its application to π-stacking in benzene dimer cation and ionized DNA base pair steps. *J. Chem. Phys.* **2016**, *145*, 134106.

DOI: 10.1021/acs.jctc.8b00058
J. Chem. Theory Comput. 2018, 14, 2955−2978