THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

Cite This: J. Phys. Chem. Lett. 2019, 10, 2706−2714

pubs.acs.org/JPCL

Accurate and Efficient *ab Initio* Calculations for Supramolecular Complexes: Symmetry-Adapted Perturbation Theory with Many-Body Dispersion

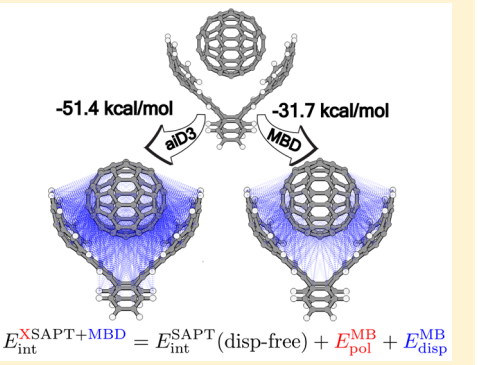
Kevin Carter-Fenk,[†] Ka Un Lao,[‡] Kuan-Yu Liu,[†] and John M. Herbert*[,†]

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

[‡]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, United States

Supporting Information

ABSTRACT: Symmetry-adapted perturbation theory (SAPT) provides a chemically meaningful energy decomposition scheme for nonbonded interactions that is useful for interpretive purposes. Although formally a dimer theory, we have previously introduced an "extended" version (XSAPT) that incorporates many-body polarization via self-consistent charge embedding. Here, we extend the XSAPT methodology to include nonadditive dispersion, using a modified form of the many-body dispersion (MBD) method of Tkatchenko and co-workers. Dispersion interactions beyond the pairwise atom−atom approximation improve total interaction energies even in small systems, and for large π-stacked complexes these corrections can amount to several kilocalories per mole. The XSAPT+MBD method introduced here achieves errors of $\lesssim$1 kcal/mol (as compared to high-level *ab initio* benchmarks) for the L7 data set of large dispersion-bound complexes and $\lesssim$4 kcal/mol (as compared to experiment) for the S30L data set of host−guest complexes. This is superior to the best contemporary density functional methods for noncovalent interactions, at comparable or lower cost. XSAPT+MBD represents a promising method for application to supramolecular assemblies, including protein−ligand binding.



Downloaded via OHIO STATE UNIV on May 10, 2019 at 21:15:15 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Noncovalent interactions are ubiquitous in chemistry, playing a major role in drug design,[1−3] crystal structure prediction,[4−6] synthesis of supramolecular host−guest complexes,[7] and molecular organization at interfaces.[8] Whereas in small systems composed of nonpolar monomers it may be possible to approximate the total intermolecular interaction energy in a pairwise-additive way, this approximation quickly degrades for strongly interacting molecules. In polar systems as small as water trimer, three-body (trimolecular) effects contribute 18% of the total interaction energy,[9,10] and three-body polarization effects can be as large as 13 kcal/mol in magnitude for stable isomers of $(H_2O)_6$.[11] These nonadditive polarization effects have long been a focal point in monomer-based approaches to supramolecular interaction energies,[12−16] while nonadditive dispersion has been largely ignored. However, in systems such as π-stacked uracil dimer, the three-body (triatomic) van der Waals energy is estimated to contribute ∼30% of the total binding energy, with four-body (tetra-atomic) contributions of ∼10%.[17,18] Here, we report a method that captures both nonadditive polarization and nonadditive dispersion, by combining the many-body dispersion (MBD) approach of Tkatchenko and co-workers[19,20] with our "extended" version of symmetry-adapted perturbation theory (XSAPT).[21−26]

The SAPT approach affords a natural partition of intermolecular interaction energies into physically interpretable contributions that include electrostatics, exchange repulsion, induction, and dispersion:[27−29]

$$E_{\rm int}^{\rm SAPT} = E_{\rm elst}^{(1)} + E_{\rm exch}^{(1)} + E_{\rm ind}^{(2)} + E_{\rm exch\text{-}ind}^{(2)} + E_{\rm disp}^{(2)} + E_{\rm exch\text{-}disp}^{(2)} + \cdots \quad (1)$$

(All terms through second order in intermolecular perturbation theory have been written out explicitly.) SAPT is normally used to investigate pairwise interactions between dimers, as the calculation of nonadditive three-body interactions in trimers requires the use of triple excitations,[30−33] which is computationally expensive. However, using monomer wave functions that include self-consistent charge embedding (the "XPol" procedure[13,16,34]), the leading-order many-body polarization effects are incorporated into the zeroth-order SAPT wave functions. This is the basis of XSAPT.[21−26] To date, nonadditive dispersion has been ignored.

The SAPT framework provides varying levels of sophistication for the treatment of electron correlation and thus dispersion.[28,35] Benchmark-quality description of dispersion interactions is *not* obtained at second order (eq 1) and requires either higher-order perturbation theory (e.g., SAPT2+),[35]

Received: April 23, 2019
Accepted: May 7, 2019
Published: May 7, 2019

© XXXX American Chemical Society

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

which incurs $O(N^7)$ scaling with respect to system size, or else SAPT(DFT),[29,36−39] which computes the dispersion interaction based on frequency-dependent density susceptibilities computed for each monomer. The cost of SAPT(DFT) can be reduced from $O(N^6)$ to $O(N^5)$ using density fitting,[40,41] but even this remains prohibitively expensive for large systems.

Atom−atom dispersion potentials are a cost-effective alternative,[23−25,42] and we have developed *ab initio* dispersion corrections for use in SAPT.[24,25] These "+*ai*D" corrections[26] are in some ways similar in spirit to empirical dispersion corrections used in density functional theory (DFT-D),[43,44] but they are fit directly to high-quality SAPT2+(3)(CCD)[28] dispersion energies and thus represent pure dispersion at all length scales. In contrast, empirical dispersion in DFT must be attenuated as intermolecular distance $R \to 0$, in order to avoid double-counting of correlation effects in regions of space where the DFT exchange-correlation potential is significant.[43]

Our third-generation method, XSAPT+*ai*D3,[25] affords accurate interaction energies for small clusters, with a mean absolute deviation (MAD) of just 0.3 kcal/mol versus complete-basis CCSD(T) benchmarks for the S66 data set of small dimers.[45] The assumption of (atomic) pairwise additivity in the dispersion energy necessarily deteriorates in large systems,[17] however, and this manifests as larger errors for supramolecular complexes involving large monomers.[26] In such cases, one must account for screening of the atomic C_6 coefficients in the presence of other, polarizable centers.[46] An *ab initio* method to account for this effect within XSAPT+*ai*D3 has been introduced,[47] but engenders $O(N^4)$ cost. Alternatively, three-body dispersion corrections of the Axilrod−Teller−Muto (ATM) variety can be introduced:[26,48]

$$E_{\mathrm{ATM}} = \sum_{A<B<C}^{\mathrm{atoms}} C_9^{ABC}\left(\frac{3\cos\theta_a\cos\theta_b\cos\theta_c + 1}{\bar{R}_{\mathrm{vdW},ABC}^{9} + R_{AB}^3 R_{AC}^3 R_{BC}^3}\right) \tag{2}$$

These corrections were found to be crucial in XSAPT+*ai*D3 calculations involving large monomers.[26] In the Tkatchenko−Scheffler (TS) version of this correction,[49] $E_{\mathrm{ATM}}^{(\mathrm{TS})}$, the triatomic C_9 coefficients are derived from the electron density, whereas Grimme's version ($E_{\mathrm{ATM}}^{(\mathrm{Grimme})}$) determines C_9 from the geometric mean of the atomic C_6 coefficients.[50]

The present work aims to capture nonadditive dispersion beyond the leading-order ATM correction. For this purpose we turn to the MBD procedure, which incorporates higher-order corrections by accounting for screening of the atomic dispersion interactions by the chemical environment. This is accomplished via a range-separated, self-consistent screening approach or "MBD@rsSCS".[19,20,51] This method, which we will simply call "MBD" hereafter, has been reviewed elsewhere,[52] so we discuss only the necessary modifications to incorporate this theory into the XSAPT formalism.

The original MBD correction was designed to capture only the long-range dipole−dipole component of the dispersion interaction, but SAPT dispersion must be well-defined even at short intermolecular distances. As such, the original MBD method cannot simply be reparameterized using *ab initio* dispersion data, and our initial attempts to do so produced unsatisfying results. To extend the MBD model into the short-range regime, we include dipole−quadrupole response terms via "effectively-screened" C_8 coefficients, $C_{8,\mathrm{es}}$. The atomic coefficient $C_{8,\mathrm{es}}^{AA}$ for atom A is derived from a modified recursion relation[50]

$$C_{8,\mathrm{es}}^{AA} = 3C_{6,\mathrm{SCS}}^{AA}Q_A \tag{3}$$

where $C_{6,\mathrm{SCS}}^{AA}$ is the atomic dipole−dipole dispersion coefficient derived from the self-consistently screened (SCS) atomic polarizabilities,[52] and

$$Q_A = \frac{\gamma Z_A^{1/2}\langle r^4\rangle_A}{\langle r^2\rangle_A} \tag{4}$$

The scaling factor

$$\gamma = s_{42} + \exp(-Z_A^{1/2}/2) \tag{5}$$

was chosen (with $s_{42} = 0.29$) such that eq 3 reproduces quadrupole polarizabilities for rare-gas atoms (α_2, from ref 53), with average deviations of ∼1.5%. The quadrupole polarizabilities are derived from the coefficients $C_{8,\mathrm{es}}$ via the quantum Drude oscillator invariant[53]

$$C_{8,\mathrm{es}}^{AA} = 5\alpha_1^A\alpha_2^A\bar{\omega}_A \tag{6}$$

This approach differs from the scaling proposed by Grimme et al.[50] and shows improved results for atomic response properties of light elements.

Rather than relying on eq 3 alone to transfer interatomic response properties to the molecular environment, we use the quantum Drude oscillator expression for the pairwise C_8 coefficients:[53]

$$C_8^{AB} = \frac{15}{2}\bar{\omega}_A\bar{\omega}_B\left(\frac{\alpha_1^A\alpha_2^B}{\bar{\omega}_A + 2\bar{\omega}_B} + \frac{\alpha_2^A\alpha_1^B}{2\bar{\omega}_A + \bar{\omega}_B}\right) \tag{7}$$

Using eq 6, we then express the effectively screened dipole−quadrupole coefficients in terms of atomic properties, in a manner similar to the original TS scheme:[49]

$$C_{8,\mathrm{es}}^{AB} = \frac{3}{2}\left[\frac{C_{8,\mathrm{es}}^{AA}C_{8,\mathrm{es}}^{BB}}{(\alpha_1^B/\alpha_1^A)C_{8,\mathrm{es}}^{AA} + 2(\alpha_2^A/\alpha_2^B)C_{8,\mathrm{es}}^{BB}} + \frac{C_{8,\mathrm{es}}^{AA}C_{8,\mathrm{es}}^{BB}}{(\alpha_1^A/\alpha_1^B)C_{8,\mathrm{es}}^{BB} + 2(\alpha_2^B/\alpha_2^A)C_{8,\mathrm{es}}^{AA}}\right] \tag{8}$$

Finally, effectively screened dipole−quadrupole (esDQ) dispersion can be added as a short-range correction to MBD dispersion:

$$E_{\mathrm{disp}}^{\mathrm{MBD+esDQ}} = E_{\mathrm{disp}}^{\mathrm{MBD}} - s_8\sum_{A>B}^{\mathrm{atoms}} f_8(R_{AB})\frac{C_{8,\mathrm{es}}^{AB}}{R_{AB}^8} \tag{9}$$

Here, s_8 is a linear scaling parameter and f_8 is the Tang−Toennies damping function,

$$f_8(R_{AB}) = 1 - \mathrm{e}^{-bR_{AB}}\sum_{k=0}^{8}\frac{(bR_{AB})^k}{k!} \tag{10}$$

with

$$b = a_1(R_{\mathrm{vdW},A} + R_{\mathrm{vdW},B}) + a_2 \tag{11}$$

As in the original formulation of MBD,[20] a Fermi-type damping function is used to calculate $E_{\mathrm{disp}}^{\mathrm{MBD}}$, but the range-separation parameter β in that damping function is refit for use with SAPT. Both the effective atomic radii (R_{vdW}) and the effective dipole polarizabilities (α_1) that are inputs to the MBD procedure are obtained from volume ratios computed using

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

Hirshfeld partitioning of the density, as in ref 49, but using XPol densities in the present case.

Parameters β, s_8, a_1, and a_2 were optimized by fitting to high-quality dispersion energies, minimizing mean absolute percentage deviations (MAPDs), with results reported in Table 1. The training set for this fitting consists of the S66 × 8 data set of biologically relevant dimers,[45] which are each evaluated at 8 intermolecular distances; the X40 data set of halogenated dimers;[54] and a new data set that we call "π11 × 8", as it consists of 11 π-stacked complexes at each of 8 intermolecular separations. The π11 × 8 complexes include four substituted benzene dimers from ref 55 and seven complexes involving benzene or naphthalene with a heterocycle. The combined training set (S66 × 8 + X40 + π11 × 8, or "SXπ") contains a total of 656 dimer geometries. For the S66 × 8 and X40 data sets, we use SAPT(DFT) dispersion energies from refs 56 and 57 that were computed at the SAPT(DFT)/aug-cc-pVTZ level and scaled by a factor of 1.051 to account for basis-set incompleteness.[58] For π11 × 8, we report SAPT2+3(CCD)/ and SAPT2+(3)(CCD)/aug-cc-pVTZ benchmarks for the first time. Geometries and benchmark energies for π11 × 8 complexes are provided in the Supporting Information.

Table 1. Results of the MBD Parameterization

data set	parameters				error (kcal/mol)[a]		
	β	a_1	a_2	s_8	MAD	(%)	RMSD
SXπ	0.511	0.221	0.201	1.160	0.193	(4.9%)	0.306
A55	0.503	0.171	0.271	1.756	0.192	(7.5%)	0.263

[a]Mean absolute deviation (MAD) and root-mean square deviation (RMSD) are in kcal/mol. Mean absolute percent deviation is given in parentheses.

A validation data set that we call A55 was constructed by combining the A24 set of small dimers[59] with a set of 31 dimers used in ref 58 to validate SAPT(DFT). (The latter contain up to nine non-hydrogen atoms.) Very similar parameters are obtained by fitting to A55 instead of SXπ (see Table 1). When the A55-optimized parameters are applied to SXπ, errors increase only very slightly as compared to those obtained using SXπ-optimized parameters. Specifically, the MAD increases to 0.258 kcal/mol (ΔMAD = 0.065 kcal/mol) and the RMSD increases to 0.394 kcal/mol (ΔRMSD = 0.088 kcal/mol). This suggests that the parametrization is robust. It is also worth noting that setting $s_8 = 1$ in eq 9 results in only a small increase in the errors (RMSD = 0.48 kcal/mol, MAD = 0.28 kcal/mol, and MAPD = 6.2%), which suggests that our "effective screening" ansatz for C_8 is reasonably accurate. Because the SXπ data set is more diverse than A55, we will use the SXπ-optimized parameters in what follows.

Having settled on a parametrization of MBD, we proceed to assess the combined XSAPT+MBD method against the best available noncovalent benchmarks, starting with CCSD(T) data at the complete basis set (CBS) limit for the 3B-69 data set.[60] XPol embedding charges were obtained from the molecular electrostatic potential,[22] as in previous work.[22−26] Previous studies suggest that the individual energy components in XSAPT are converged in triple-ζ basis sets,[24] and we use def2-TZVPPD throughout.[61] (A detailed comparison of basis sets can be found in the Supporting Information.) All XSAPT(KS) calculations were done with the long-range corrected (LRC-)ωPBE functional,[62] using the global density-dependent (GDD) tuning procedure to determine the range-separation parameter, ω.[25,26]

Error statistics for three-body contributions to the interaction energies in 3B-69 are presented in Table 2. For the XSAPT+*ai*D3 method, we present results both with and without an explicit three-body dispersion correction, $E_{\mathrm{ATM}}^{(\mathrm{TS})}$. The three-body contribution to the XSAPT interaction energy for trimer *ABC* can be isolated by subtracting the pairwise XSAPT interactions,

$$E_{\mathrm{int},3\mathrm{B}}^{\mathrm{XSAPT}} = \sum_{A<B<C} (E_{\mathrm{int},ABC}^{\mathrm{XSAPT}} - E_{\mathrm{int},AB}^{\mathrm{XSAPT}}) \tag{12}$$

Three-body interaction energies are rather small in the 3B-69 systems; nevertheless, XSAPT+MBD provides an impressive 31% improvement in the accuracy of the three-body interactions, relative to XSAPT+*ai*D3. The latter method includes many-body polarization but neglects many-body dispersion. Addition of $E_{\mathrm{ATM}}^{(\mathrm{TS})}$ to XSAPT+*ai*D3 improves the results by 26%. We conclude that nonadditive dispersion is essential for an accurate description of three-body interactions within the XSAPT formalism. Nonadditive dispersion is likely even more important in this context than it is in DFT-D, because the latter method may already include some nonadditive, short-range dispersion within the density functional itself.

Following development of our second-generation +*ai*D2 potential,[24] it became clear that this approach systematically overestimates the dispersion energy in π-stacked systems, prompting us to develop +*ai*D3.[25] To assess the accuracy of the MBD correction for π-stacked systems, we examine a set of 10 nucleic acid tetramers for which SAPT(DFT)/aug-cc-pVTZ dispersion energies are available.[63] These dispersion energies are scaled to account for basis set incompleteness error, but different scaling factors have been suggested.[42,58] In Figure 1, the difference between scaling by 1.051 (as suggested in ref 58) versus scaling by 1.100 (suggested ref 42) is taken as an error bar on the benchmark interaction energies for the nucleobase tetramers.

That said, the scaling parameter of 1.051 is obtained by averaging SAPT(DFT) results for small dimers obtained using basis sets ranging up to aug-cc-pV5Z,[58] whereas the scaling

Table 2. Errors in Three-Body Interaction Energies[a]

method	MAD		RMSD		MSD[c]
	error (kcal/mol)	improvement[b] (%)	error (kcal/mol)	improvement[b] (%)	(kcal/mol)
XSAPT+MBD	0.067	(30.6%)	0.096	(25.8%)	−0.012
XSAPT+*ai*D3	0.099		0.136		−0.059
XSAPT+*ai*D3+$E_{\mathrm{ATM}}^{(\mathrm{TS})}$	0.072	(25.7%)	0.107	(19.9%)	−0.028

[a]Versus CCSD(T)/CBS benchmarks for the 3B-69 data set of ref 60. [b]Improvement relative to XSAPT+*ai*D3. [c]Mean signed deviation.

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

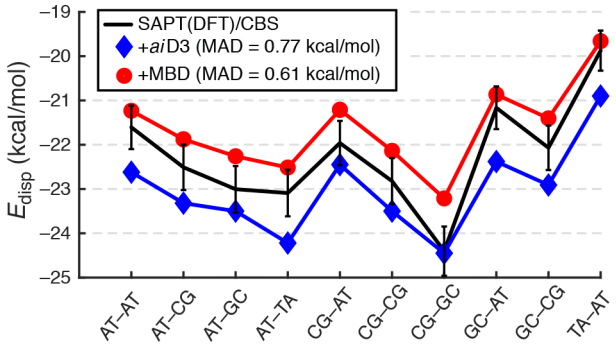


Figure 1. MBD and *ai*D3 dispersion energies for nucleic acid tetramers. SAPT(DFT)/aug-cc-pVTZ data have been scaled to approximate the CBS limit, and the benchmark SAPT(DFT) data (in black) are presented with error bars that indicate scaling by a factor of 1.051 (as suggested in ref 58, upper delimiter) as compared to a factor of 1.100 (as suggested in ref 42, lower delimiter). MADs are reported relative to the average of the two scaled SAPT(DFT) values.

parameter of 1.100 is more *ad hoc*.[42] As a test, we considered the parallel-displaced isomer of $(C_6H_6)_2$, scaling the SAPT(DFT)/aug-cc-pVTZ dispersion energy in an attempt to match the SAPT2+(3)(CCD)/CBS value obtained by two-point extrapolation. Scaling by 1.051 results in a discrepancy of just 0.02 kcal/mol between these two values, whereas scaling by 1.100 affords a discrepancy of 0.4 kcal/mol. As such, we consider the factor of 1.051 to be more reliable in the present context.

Applying this factor to SAPT(DFT)/aug-cc-pVTZ dispersion energies to estimate the SAPT(DFT)/CBS benchmarks for the nucleobase tetramers, we find that the MBD procedure eliminates the systematic overestimation of π-stacking energies that we have previously documented for the + *ai*D family of methods, reducing the MAD from 0.8 kcal/mol (*ai*D3) to 0.2 kcal/mol (MBD). Because the leading nonadditive term in the dispersion energy is repulsive (whereas pairwise dispersion is always attractive), these results might have been anticipated and suggest that the overbinding exhibited by SAPT+*ai*D for π-stacked systems is at least partly attributable to the pairwise approximation. The extremely small errors obtained using XSAPT+MBD makes this method very promising for future applications to large biological systems.

Recent NMR measurements of the binding affinities between noble gas atoms and cucurbit[5]uril (CB5) macrocyclic receptors provide a point of contact with experiment for purely dispersion-bound systems.[64] Free energies of association at the XSAPT+MBD level were computed here, using thermodynamic quantities provided in ref 64. Results for guest atoms He−Kr are reasonably good (Table 3), but for reasons that are not entirely clear, the Xe@CB5 complex is significantly overbound, as compared to either experiment or CCSD(T). Note, however, that we are unable to obtain the corresponding values at the XSAPT+*ai*D3 level because the noble gas atoms were not included in the parametrization. Similarly, a ferrocene complex that is part of the S12L data set[65] was excluded in our previous evaluation of XSAPT+*ai*D3 for large supramolecular complexes[26] because of lack of *ai*D3 parameters for Fe. The more universal parametrization of MBD is a distinct advantage of the method introduced here.

We next consider the L7 set of large dispersion-bound dimers,[66] as well as the S30L data set of large host−guest complexes.[67] The L7 set was originally published in ref 66 with interaction energies computed at the QCISD(T)/CBS level. Interaction energies were later updated to the CCSD(T)/CBS level,[68] using the domain-based local pair natural orbital (DLPNO) implementation of CCSD(T) by Neese and co-workers.[69−71] However, for noncovalent interactions the DLPNO approximation is sensitive to user-defined thresholds that control truncation of the PNO basis and domain size. These thresholds have recently been standardized, and both TightPNO and NormalPNO thresholds (as defined in ref 71) are reported here. TightPNO benchmarks reported in our previous work,[26] obtained via private correspondence from Grimme and co-workers,[72] have been recomputed here, using Orca v. 4.1.1.[73]

As compared to previous values,[26,72] our new DLPNO-CCSD(T)/CBS interaction energies are in better agreement with DLPNO-CCSD(T)-F12/cc-pVDZ (VeryTightPNO) results reported elsewhere[74] (see Table 4). Results for the smallest L7 dimers are insensitive to the inclusion of diffuse functions, so for the sake of cost we estimate the CCSD(T)/CBS interaction energy via two-point extrapolation of cc-pVDZ and cc-pVTZ results, using the extrapolation scheme suggested by Neese and Valeev.[75] For the two L7 complexes involving circumcoronene, we were unable to complete the DLPNO-CCSD(T) calculations using TightPNO thresholds on the computational resources available to us, so for these two complexes the values from ref 72 are used instead.

Results for the L7 complexes (Table 4) demonstrate that XSAPT+MBD affords excellent agreement with the TightPNO benchmarks, with a MAD of 1.0 kcal/mol. This is a small but noticeable improvement upon XSAPT+*ai*D3+$E_{ATM}^{(TS)}$, which exhibits a MAD of 1.4 kcal/mol. Omitting the three-body

Table 3. Binding Energies for Noble Gas@Cucurbit[5]uril Complexes

	ΔE^a (kcal/mol)				ΔG(kcal/mol)	
guest	TPSS -D3[b]	CCSD(T)[b]	MP2[b]	XSAPT + MBD	exptl[c]	XSAPT + MBD
He	−1.2	−1.4	−0.6	−1.0	0.1	2.0
Ne	−2.5	−1.7	−1.5	−2.0	0.1	1.0
Ar	−6.4	−6.1	−6.5	−6.7	−1.5	−1.7
Kr	−8.4	−7.9	−9.0	−8.9	−3.0	−3.7
Xe	−10.1	−9.2	−12.5	−13.9	−4.1	−8.7
MAD[d] (all)	0.5		1.5	1.4		1.7
MAD[d] (He−Kr)	0.5		0.6	0.6		0.9

[a]Including MP2/def2-TZVP deformation energies from ref 64. [b]From ref 64, def2-TZVP basis set. [c]Back-corrected "gas phase" ΔG, from a solution-phase NMR measurement.[64] [d]Relative to CCSD(T) for ΔE and relative to experiment for ΔG.

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

Table 4. Interaction Energies (kcal/mol) for the L7 Dimers

		DLPNO-CCSD(T)/CBS			
system	QCISD(T)/CBS[a]	NormalPNO[b]	TightPNO[c]	CCSD(T)-F12/cc-pVDZ[d]	XSAPT+MBD
C2C2PD	−24.81	−24.81	−23.32	−19.14	−21.17
C3A	−18.19	−17.98	−15.80[e]		−16.99
C3GC	−31.25	−29.86	−26.70[e]		−27.89
CBH	−11.06	−11.64	−11.48	−11.13	−12.82
GCGC	−14.37	−13.21	−13.80	−13.69	−14.55
GGG	−2.40	−1.68	−2.22	−2.36	−2.13
PHE	−25.76	−22.81	−25.01	−25.09	−25.64

[a]From ref 66. [b]From ref 68. [c]This work, except as noted. [d]DLPNO approximation with VeryTightPNO threshold; from ref 74. [e]Private communication (ref 72).

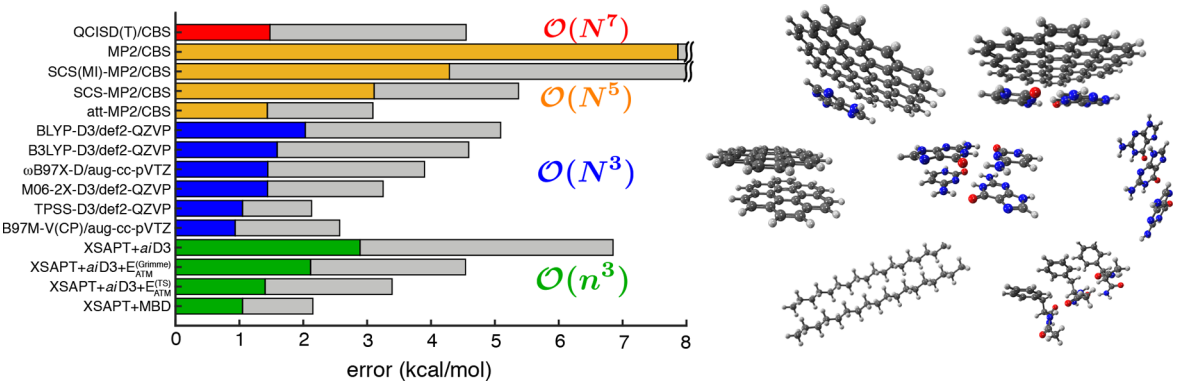


Figure 2. Errors in interaction energies for the L7 set of dimers (depicted on the right), as predicted by a variety of quantum-chemical methods in comparison to DLPNO-CCSD(T)/CBS benchmarks. Gray bars indicate maximum errors (truncated in two cases), whereas colored bars indicate MAEs color-coded according to cost. The computational scaling is indicated for each method: $O(N^p)$ means that the method scales asymptotically as N^p with respect to dimer size, N, whereas $O(n^3)$ means that the cost grows cubically with respect to monomer size, n. The att-MP2 data are from ref 77; the ωB97X-D data are from ref 78; the B97M-V and XSAPT+*ai*D data are from ref 26. All other calculations are from ref 66, except for the XSAPT+MBD values, which are reported here for the first time. Note that the DFT-D3 methods contain three-body ATM-style dispersion corrections.[50]

dispersion correction (i.e., XSAPT+*ai*D3) results in dramatic overestimation of the dispersion energies and a MAD of 2.9 kcal/mol. This overestimation is not surprising, because three-body dispersion is usually repulsive, but note that XSAPT+*ai*D3+$E_{ATM}^{(TS)}$ systematically *underestimates* the L7 interaction energies, with a mean signed error of +1.1 kcal/mol. This suggests that the ATM-style three-body dispersion over-compensates, and indeed this correction is known to overestimate the repulsiveness of the many-body dispersion,[76] because it neglects higher-order effects that are again attractive. In contrast, the mean signed error for XSAPT+MBD is −0.4 kcal/mol. Higher-order effects included by the MBD model do appear to have the anticipated effect.

Figure 2 summarizes errors in L7 interaction energies for a slate of wave function and DFT methods. MP2/CBS errors are large (up to a maximum of 17.7 kcal/mol),[26,66] as expected for these π-stacked complexes where MP2 severely overestimates the dispersion energy. Spin-component scaled (SCS) versions of MP2 only partially mitigate these errors. The att-MP2 method[79] attenuates the MP2 interaction operator and exhibits a maximum error of only 3.2 kcal/mol[26,77] while achieving high-quality results with a triple-ζ basis set.[80] However, each of these MP2-based methods exhibits an asymptotic scaling of $O(N^5)$ with respect to the size of the supramolecular complex.

In contrast, XSAPT+MBD outperforms even att-MP2 and rivals the accuracy of the best available methods, at significantly lower cost. The cost of XSAPT+MBD scales as $O(n^3)$ with respect to *monomer* size, n, whereas even DFT calculations exhibit a cost that scales as $O(N^3)$ with respect to *supersystem* size, N. In addition, XSAPT calculations are inherently free of basis-set superposition error and therefore do not require counterpoise (CP) correction, and for this reason there is no compelling reason to use quadruple-ζ basis sets with XSAPT+MBD. Furthermore, XSAPT+MBD provides an energy decomposition analysis scheme,[81] though we defer a detailed analysis of this aspect to a future publication.

The L7 data set consists mostly of complexes involving π-stacked nucleobases, coronene, and/or circumcoronene and is therefore dispersion-heavy. A more diverse set of interactions is found within the S30L data set,[67] which extends the earlier S12L data set[65] and consists of host−guest complexes involving some rather large monomers such as cucurbit[n]urils, resorcin[4]arenes, and cyclodextrins, as shown in Figure 3a−c. Benchmark interaction energies for these complexes are obtained from solution-phase binding affinity measurements that are then back-corrected to obtain effective "gas phase" experimental benchmarks.[67] Many of the guest molecules are flexible, and it is therefore crucial to include monomer deformation energies. These are included in the benchmarks of ref 67, but the deformation energies were not provided; therefore, we have recomputed them at the SCS-MP2/CBS level and included them in the Supporting Information.

When deformation energies are included at this level, the XSAPT+MBD method predicts the experimental interaction

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

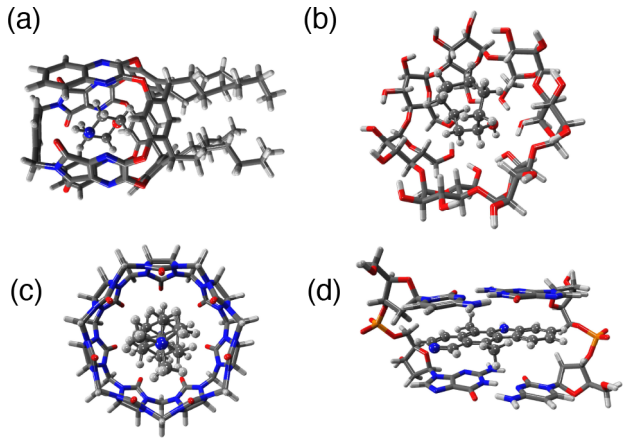


Figure 3. (a) Morpholine@resorcin[4]arene (250 atoms), (b) cyclooctanol@β-cyclodextrin (172 atoms), and (c) diamantane diammonium@cucurbit[7]uril (184 atoms), each of which is taken from the S30L data set. (d) Ellipticine/DNA intercalation complex (157 atoms).

Table 5. Error Statistics for the S30L∩S12L Set

	error (kcal/mol)		
method	MAD	MSD	Max
MP2/CBS[a]	15.67	−15.67	−53.30
SCS-MP2/CBS[a]	6.84	−3.15	−26.25
B97M-V(CP)/aug-cc-pVTZ[b]	5.93	−0.04	11.25
XSAPT+*ai*D3[b]	9.15	−8.47	−21.85
XSAPT+*ai*D3+$E_{ATM}^{(TS)}$[b]	5.73	−2.83	−10.59
XSAPT+*ai*D3+$E_{ATM}^{(Grimme)}$[b]	7.65	−6.51	−18.65
XSAPT+MBD	4.16	−3.58	−10.36

[a]From ref 82. [b]From ref 26.

energies with a MAD of 5.9 kcal/mol and a MSD of −3.6 kcal/mol that suggests systematic overbinding. However, when considering the structures from S30L that are also present in S12L (S30L∩S12L data set), for which we also have XSAPT+*ai*D3 data from ref 26, it becomes clear that XSAPT+MBD provides a significant improvement relative to XSAPT+*ai*D3 and even XSAPT+*ai*D3+$E_{ATM}^{(TS)}$ (see Table 5). XSAPT+MBD is also slightly more accurate than the B97M-V density functional, which is notable because the latter is one of the most accurate DFT methods for noncovalent interactions,[83] especially if one considers only functionals that omit Hartree−Fock exchange and can thus be applied more easily to very large systems.

As a final application, we present interaction energies and energy decomposition analysis for the ellipticine/DNA intercalation complex depicted in Figure 3d.[25,84] Dividing this system into three fragments, and using the latest version of our XSAPT code,[26,85] the XSAPT/def2-TZVPPD calculations require a total of 40 h on a single compute node with 40 processors, while CP-corrected ωB97M-V/def2-TZVPPD calculations require 38 h on the same hardware. While the formal scaling of XSAPT+MBD is superior to supramolecular DFT, a total of seven calculations (averaging 6 h each) is required to obtain the many-body induction energy in XSAPT, whereas CP-corrected DFT requires three separate calculations averaging 13 h each. In each case these separate calculations are trivially parallelizable, so that by running on seven nodes instead of one, the wall time for the XSAPT calculations would be reduced to a mere 6 h. The three DFT calculations could similarly be run on separate nodes, but this would only reduce the wall time to 13 h.

Table 6. Interaction Energies for the DNA Intercalation Complex of Figure 3d

method	E_{int} (kcal/mol)
QMC[a]	−33.6 ± 0.9
B97M-V(CP)[b]	−41.3
ωB97M-V(CP)[b]	−43.7
XSAPT+*ai*D3+$E_{ATM}^{(TS)}$[c]	−35.7
XSAPT+MBD[c]	−40.4
XSAPT[c] Components	
E_{elst}	−22.2
E_{exch}	59.2
E_{ind}	−7.0
$E_{disp}^{aiD3}+E_{ATM}^{(TS)}$	−65.7
$E_{disp}^{MBD+esDQ}$	−70.4

[a]From ref 84. [b]CP-corrected def2-TZVPPD. [c]Heavy-augmented def2-hpTZVPP basis set defined in ref 25.

Results for the DNA intercalation complex (Table 6) show that the XSAPT+MBD and B97M-V calculations afford interaction energies within 1 kcal/mol of one another, while ωB97M-V is somewhat further from the quantum Monte Carlo (QMC) result. In this case, the XSAPT+*ai*D3+$E_{ATM}^{(TS)}$ prediction is in best agreement with the QMC result. Importantly, the energy decomposition reveals that the intercalating drug molecule would not bind to DNA were it not for the dispersion contribution, demonstrating the profound importance of dispersion in this system.

In this work, we have introduced a novel methodology that allows for efficient and accurate calculation and analysis of noncovalent interaction energies for large systems, including both nonadditive polarization and nonadditive dispersion effects, at $O(n^3)$ cost. The method shows appreciable improvement upon the previously reported XSAPT+*ai*D3 method, even when the latter is corrected for the leading-order Axilrod−Teller−Muto three-body dispersion interaction. Dispersion corrections beyond the pairwise-additive approximation offer useful improvements even for small systems, and for larger systems (especially where π-stacking interactions are involved) they often improve the results by several kilocalories per mole as compared to XSAPT+*ai*D3. Furthermore, the present approach puts the dispersion correction to XSAPT on a more solid theoretical basis as compared to the use of atom−atom dispersion potentials in the + *ai*D3 correction. At the same time, we are working to develop a version of MBD that is valid at all length scales, which would obviate the need for explicit C_8 potentials (and corresponding damping functions) in eq 9.

For the L7 and S30L data sets, where the monomers range in size up to almost 200 atoms, XSAPT+MBD affords interaction energies that are as accurate as the best-available DFT methods that include nonlocal correlation, at comparable or lower computational cost. As this method is pushed to even larger supramolecular complexes, or those involving numerous monomers, the superior formal scaling of XSAPT, combined with its ease of parallelization across fragments, will ensure that XSAPT+MBD is considerably more affordable than traditional supramolecular DFT. That said, XSAPT is primarily a tool for analysis of intermolecular interactions, as forces are not available.

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

The XSAPT+MBD method has been implemented in a locally modified version of the Q-Chem software[86] and will be released with v. 5.3. Having established its accuracy and robustness in the present work, we look forward to applying this method to analyze noncovalent interactions in large systems, taking advantage of the natural way in which XSAPT decomposes intermolecular interaction energies into chemically meaningful components.[81] An interesting question for future work is the extent to which empirical "-D" corrections to DFT reproduce true dispersion energies. The present methodology is poised to address this issue and to offer incisive analysis of the role of dispersion in intermolecular interactions in complex systems.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.9b01156.

Benchmark interaction energies, energy components, and tuned ω_{GDD} parameters for various data sets; definition and benchmarks for the $\pi 11 \times 8$ data set; tests of the basis-set dependence of the XSAPT+*ai*D3 and XSAPT+MBD methods (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

ORCID

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

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

Notes

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

■ ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award No. DE-SC0008550. Calculations were performed at the Ohio Supercomputer Center under project no. PAA-0003.[87]

■ REFERENCES

(1) Zhou, P.; Huang, J.; Tian, F. Specific noncovalent interactions at protein−ligand interface: Implications for rational drug design. *Curr. Med. Chem.* **2012**, *19*, 226−238.

(2) Li, S.; Xu, Y.; Shen, Q.; Liu, X.; Lu, J.; Chen, Y.; Lu, T.; Luo, C.; Luo, X.; Zheng, M.; Jiang, H. Non-covalent interactions with aromatic rings: Current understanding and implications for rational drug design. *Curr. Pharm. Des.* **2013**, *19*, 6522−6533.

(3) Ulbrich, K.; Holá, K.; Šubr, V.; Bakandritsos, A.; Tuček, J.; Zbořil, R. Targeted drug delivery with polymers and magnetic nanoparticles: Covalent and noncovalent approaches, release control, and clinical studies. *Chem. Rev.* **2016**, *116*, 5338−5431.

(4) Woodley, S. M.; Catlow, R. Crystal structure prediction from first principles. *Nat. Mater.* **2008**, *7*, 937−946.

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

(6) Price, S. L.; Brandenburg, J. G. Molecular crystal structure prediction. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; de la Roza, A. O., DiLabio, G. A., Eds.; Elsevier, 2017.

(7) Wei, P.; Yan, X.; Huang, F. Supramolecular polymers constructed by orthogonal self-assembly based on host−guest and metal−ligand interactions. *Chem. Soc. Rev.* **2015**, *44*, 815−832.

(8) Auwärter, W.; Écija, D.; Klappenberger, F.; Barth, J. V. Porphyrins at interfaces. *Nat. Chem.* **2015**, *7*, 105−120.

(9) Xantheas, S. S. *Ab initio* studies of cyclic water clusters $(H_2O)_n$, n = 1−6. II. Analysis of many-body interactions. *J. Chem. Phys.* **1994**, *100*, 7523−7534.

(10) Cisneros, G. A.; Wikfeldt, K. T.; Ojamäe, L.; Lu, J.; Xu, Y.; Torabifard, H.; Bartók, A. P.; Csányi, G.; Molinero, V.; Paesani, F. Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.* **2016**, *116*, 7501−7528.

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

(12) Dahlke, E. E.; Truhlar, D. G. Electrostatically embedded many-body expansion for large systems, with applications to water clusters. *J. Chem. Theory Comput.* **2007**, *3*, 46−53.

(13) 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.

(14) Jacobson, L. D.; Richard, R. M.; Lao, K. U.; Herbert, J. M. Efficient monomer-based quantum chemistry methods for molecular and ionic clusters. *Annu. Rep. Comput. Chem.* **2013**, *9*, 25−56.

(15) Richard, R. M.; Lao, K. U.; Herbert, J. M. Aiming for benchmark accuracy with the many-body expansion. *Acc. Chem. Res.* **2014**, *47*, 2828−2836.

(16) 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.

(17) 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.

(18) 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.

(19) 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.

(20) 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.

(21) 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.

(22) 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.

(23) 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.

(24) Lao, K. U.; Herbert, J. M. An improved treatment of empirical dispersion and a many-body energy decomposition scheme for the explicit polarization plus symmetry-adapted perturbation theory (XSAPT) method. *J. Chem. Phys.* **2013**, *139*, 034107 Erratum: *ibid.* 140, 119901 (2014) .

(25) Lao, K. U.; Herbert, J. M. Accurate and efficient quantum chemistry calculations of noncovalent interactions in many-body systems: The XSAPT family of methods. *J. Phys. Chem. A* **2015**, *119*, 235−253.

(26) Lao, K. U.; Herbert, J. M. Atomic orbital implementation of extended symmetry-adapted perturbation theory (XSAPT) and benchmark calculations for large supramolecular complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955−2978.

(27) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254−272.

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

(28) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304−326.
(29) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 127−144.
(30) Moszynski, R.; Wormer, P. E. S.; Jeziorski, B.; van der Avoird, A. Symmetry-adapted perturbation theory of nonadditive three-body interactions in van der Waals molecules. I. General theory. *J. Chem. Phys.* **1995**, *103*, 8058−8074 Erratum: ibid. 107, 672−673 (1997) .
(31) Lotrich, V. F.; Szalewicz, K. Symmetry-adapted perturbation theory of three-body nonadditivity of intermolecular interaction energy. *J. Chem. Phys.* **1997**, *106*, 9668−9687.
(32) Lotrich, V. F.; Szalewicz, K. Symmetry-adapted perturbation theory of three-body nonadditivity in Ar trimer. *J. Chem. Phys.* **1997**, *106*, 9688−9702.
(33) Lotrich, V. F.; Szalewicz, K. Perturbation theory of three-body exchange nonadditivity and application to helium trimer. *J. Chem. Phys.* **2000**, *112*, 112−121.
(34) 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.
(35) 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.
(36) Heßelmann, A.; Jansen, G. Intermolecular dispersion energies from time-dependent density functional theory. *Chem. Phys. Lett.* **2003**, *367*, 778−784.
(37) 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.
(38) Misquitta, A. J.; Szalewicz, K. Symmetry-adapted perturbation-theory calculations of intermolecular forces employing density-functional description of monomers. *J. Chem. Phys.* **2005**, *122*, 214109.
(39) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J. Chem. Phys.* **2005**, *123*, 214103.
(40) Bukowski, R.; Podeszwa, R.; Szalewicz, K. Efficient calculation of coupled Kohn−Sham dynamic susceptibility functions and dispersion energies with density fitting. *Chem. Phys. Lett.* **2005**, *414*, 111−116.
(41) 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.
(42) Heßelmann, A. Comparison of intermolecular interaction energies from SAPT and DFT including empirical dispersion contributions. *J. Phys. Chem. A* **2011**, *115*, 11321−11330.
(43) Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211−228.
(44) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105−5154.
(45) 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 Erratum: *ibid.* **10**, 1359−1360 (2014) .
(46) Dobson, J. F. Beyond pairwise additivity in London dispersion interactions. *Int. J. Quantum Chem.* **2014**, *114*, 1157−1161.
(47) Lao, K. U.; Herbert, J. M. A simple correction for nonadditive dispersion within extended symmetry-adapted perturbation theory (XSAPT). *J. Chem. Theory Comput.* **2018**, *14*, 5128−5142.
(48) 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.
(49) 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.
(50) 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.
(51) Donchev, A. G. Many-body effects of dispersion interaction. *J. Chem. Phys.* **2006**, *125*, 074713.
(52) 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.
(53) Jones, A. P.; Crain, J.; Sokhan, V. P.; Whitfield, T. W.; Martyna, G. J. Quantum Drude oscillator model of atoms and molecules: Many-body polarization and dispersion interactions for atomistic simulation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 144103.
(54) Řezác, J.; Riley, K. E.; Hobza, P. Benchmark calculations of noncovalent interactions of halogenated molecules. *J. Chem. Theory Comput.* **2012**, *8*, 4285−4292.
(55) Sinnokrot, M. O.; Sherrill, C. D. Substituent effects in π-π interactions: Sandwich and T-shaped configurations. *J. Am. Chem. Soc.* **2004**, *126*, 7690−7697.
(56) Sedlak, R.; Řezác, J. Empirical D3 dispersion as a replacement for *ab initio* dispersion terms in density functional theory-based symmetry-adapted perturbation theory. *J. Chem. Theory Comput.* **2017**, *13*, 1638−1646.
(57) Řezáč, J.; de la Lande, A. On the role of charge transfer in halogen bonding. *Phys. Chem. Chem. Phys.* **2017**, *19*, 791−803.
(58) 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.
(59) 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.
(60) 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.
(61) Papajak, E.; Zheng, E. P. J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.
(62) 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.
(63) Fiethen, A.; Jansen, G.; Hesselmann, A.; Schütz, M. Stacking energies for average B-DNA structures from the combined density functional theory and symmetry-adapted perturbation theory approach. *J. Am. Chem. Soc.* **2008**, *130*, 1802−1803.
(64) He, S.; Biedermann, F.; Vankova, N.; Zhechkov, L.; Heine, T.; Hoffman, R. E.; De Simone, A.; Duignan, T. T.; Nau, W. M. Cavitation energies can outperform dispersion interactions. *Nat. Chem.* **2018**, *10*, 1252−1257.
(65) Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. - Eur. J.* **2012**, *18*, 9955−9964.
(66) 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.
(67) Sure, R.; Grimme, S. Comprehensive benchmark of association (free) energies of realistic host−guest complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785−3801 Erratum: ibid. 11, 5990 (2015) .

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714

(68) Calbo, J.; Sancho-García, J. C.; Ortí, E.; Aragó, J. DLPNO-CCSD(T) scaled methods for the accurate treatment of large supramolecular complexes. *J. Comput. Chem.* **2017**, *38*, 1869−1878.
(69) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.
(70) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.
(71) 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.
(72) Hansen, A.; Riplinger, C.; Neese, F.; Grimme, S. Unpublished data.
(73) Neese, F. Software update: The ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.
(74) 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.
(75) Neese, F.; Valeev, E. F. Revisiting the atomic natural orbital approach for basis sets: Robust systematic basis sets for explicitly correlated and conventional correlated *ab initio* methods. *J. Chem. Theory Comput.* **2011**, *7*, 33−43.
(76) 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.
(77) 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.
(78) 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.
(79) 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.
(80) Goldey, M.; Head-Gordon, M. Convergence of attenuated second order Møller−plesset perturbation theory towards the complete basis set limit. *Chem. Phys. Lett.* **2014**, *608*, 249−254.
(81) 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.
(82) Heßelmann, A.; Korona, T. Intermolecular symmetry-adapted perturbation theory study of large organic complexes. *J. Chem. Phys.* **2014**, *141*, 094107.
(83) 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.
(84) Benali, A.; Shulenburger, L.; Romero, N. A.; Kim, J.; von Lilienfeld, O. A. Application of diffusion Monte Carlo to materials dominated by van der Waals interactions. *J. Chem. Theory Comput.* **2014**, *10*, 3417−3422.
(85) Liu, K.-Y.; Carter-Fenk, K.; Herbert, J. M. Intermolecular energy decomposition analysis in large supramolecular complexes using symmetry-adapted perturbation theory. Manuscript in preparation.
(86) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kús, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A., Jr.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyaev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Pieniazek, P. A.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sergueev, N.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Vanovschi, V.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W. Head-Gordon. *Mol. Phys.* **2015**, *113*, 184−215.
(87) Ohio Supercomputer Center. http://osc.edu/ark:/19495/f5s1ph73.

DOI: 10.1021/acs.jpclett.9b01156
J. Phys. Chem. Lett. 2019, 10, 2706−2714