PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 23597

Influence of three-body effects on halogen bonding†

Sharon A. Ochieng and Konrad Patkowski *

We construct a new noncovalent benchmark dataset 3BXB that combines halogen-bonded bimolecular complexes from the SH250 dataset [K. Kříž and J. Řezáč, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14794–14804] with a third interacting partner, either H_2O or CH_4. The reference total and three-body interaction energies are computed at the CCSD(T) level. To shed light on the physical origins of binding and cooperativity in complexes of this kind, several symmetry-adapted perturbation theory (SAPT)-based energy decompositions were performed for both pairwise additive and nonadditive terms. We found that the two-body attractions in the 3BXB complexes are dominated by either electrostatics or dispersion, while the three-body effect is dominated by induction and can be either attractive or repulsive. An accurate recovery of reference interaction energies is attained by the wavefunction-based two-body SAPT variants including the δMP2 correction, combined with the SAPT(DFT) estimates of nonadditive induction and first-order exchange and any estimate of nonadditive dispersion. The values for the latter term are sometimes quite inconsistent between different approaches; fortunately, nonadditive dispersion is a relatively minor effect for complexes studied here, and all reasonable estimates lead to total interaction energies of similar accuracy.

Received 23rd July 2025,
Accepted 14th October 2025

DOI: 10.1039/d5cp02809d

rsc.li/pccp

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

1 Introduction

Accurate computational data are vital in advancing chemical understanding in domains such as template-directed synthesis, enzyme-ligand binding mechanisms, and the structure and properties of materials. The significance of computations cannot be overemphasized, because many molecular and electronic properties responsible for a broad range of phenomena cannot be measured directly. Benchmarks, therefore, play a vital role in validation of the accuracy of approximate computational methods and the development of those that rely on empirical parameters. In the realm of noncovalent interactions, where experimental data does not typically offer a straightforward comparison, these datasets are vital for ensuring accuracy and facilitating method development. Many datasets addressing different kinds of noncovalent interactions have been introduced. They range from well established but relatively small sets of complexes such as S22,[1] S66,[2] X40,[3] or S12L,[4] to more extensive sets such as SSI[5] (interactions between amino acid side chains in proteins), NENCI[6] (a wide variety of complexes in radially and orientationally nonequilibrium configurations), or the NCI Atlas by Řezáč *et al.* with several subsets mapping pairwise interactions over broad chemical space one binding type at a time.[7–11]

In molecular clusters, liquids, and solids, interactions between pairs of molecules provide only an approximate description of the total binding energy. Nonadditive contributions from triplets (and sometimes even quadruplets) of molecules are essential for obtaining the correct energetic order of water hexamers[12] and energetically ranking different polymorphs of molecular crystals.[13] However, the existing benchmark datasets for molecular trimers (or larger clusters) are not nearly as broad as for dimers and are dominated by data for water clusters[14] or for clusters involving an ion such as Cl^- solvated by a few water molecules.[15] Only a handful of benchmark three-body datasets contains trimers of different chemical character. The 3B-69 dataset contains 69 $(A)_3$ trimers from 23 different molecular crystals.[16] The S22(3) set[17] includes 24 $(A)_3$ and $(A)_2$ B-type trimers obtained by augmenting the complexes in the popular S22 database.[1] Finally, the 3BHET dataset from our earlier work[18] comprises 20 model $(A)_2B$- and ABC-type trimers exhibiting diverse types of interactions. However, these are relatively small datasets built for specific purposes and only focus on fairly specific types of interactions. Therefore, the field would benefit from the development of benchmark nonadditive datasets that are larger, more general, and as accurate as possible.

In this work, we build on the SH250[10] dataset of halogen bonded complexes featuring Cl, Br, I, and extend it to trimers to study the effects of three-body interactions on the halogen

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA. E-mail: patkowsk@auburn.edu

† Dedicated to Professor Giuseppe Resnati, celebrating a career in fluorine and noncovalent chemistry on the occasion of his 70th birthday.

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

bond (XB). In recent years, noncovalent complexes involving σ-hole interactions, including halogen and chalcogen bonding, have gained significant attention.[19–21] These interactions are characterized by attractive forces between an electrophilic region on a halogen or chalcogen atom and a Lewis base. This is a result of an anisotropic redistribution of the halogen electron density. The positive region that appears at the extension of a σ covalent bond is termed a "σ-hole".[22–24] Hydrogen and halogen bonding act as complementary design features for the design of functional molecules and materials, stabilization of drug–protein complexes, and self-assembly of supramolecular structures.[25–27] Consequently, the development of robust computational methods to describe halogen bonding is widely important. Ideal methods not only aid in accurately modeling and understanding these interactions, but also ensure reliability in predicting interaction energies for new systems. Additionally, when the accurate methods become prohibitively costly, it is common practice to test the interaction energies from approximate theoretical methods against accurate benchmark calculations.

The extension of the halogen bonded part of the SH250 dataset to trimers presented here aims to investigate the interplay between halogen bonding and other interactions. Besides introducing the dataset, this work also focuses on its applications to benchmarking different variants of symmetry-adapted perturbation theory (SAPT). This study was motivated by our previous work on heteromolecular trimers which observed that the XB complexes exhibited a comparatively smaller magnitude of the nonadditive exchange and dispersion contributions from a SAPT(DFT) energy decomposition analysis.[18] With only three systems from a dataset with 20 trimers, that observation was possibly accidental as the data were too small to be useful in making predictions on general trends in specific interactions such as XB. Building on these findings, the focus on trimers featuring XB in this work seeks to attain a broader coverage with a more comprehensive benchmark of trimers in a larger chemical space.

2 Computational methods

2.1 Composition of the data set and benchmark interaction energies

The dataset introduced in this work, which will be denoted 3BXB, follows the same strategy as our earlier 3BHET dataset[18] but with a focus on halogen bonding. We examine the influence of the three-body non-additive interactions on the halogen bond, looking for common trends in the three-body contributions that might characterize those halogen-bonded structures. 3BXB comprises 214 trimers, 107 with water and 107 with methane, obtained by extending the halogen-bonded dimers featuring Cl, Br and I from the SH250 dataset.[10] We picked either water or methane as the third partner because the two molecules have the same number of electrons (10), but exhibit very different kinds of interactions, predominantly electrostatic/hydrogen bonded for water and mostly dispersive for methane. The starting geometries were obtained from dimers in the SH250[10] dataset by adding the third partner and reoptimizing only the variables pertaining to the third molecule in the trimer. This ensures that the halogen-bonded complex is not broken, but simply augmented with an external partner. The structures of some of these trimers are illustrated in Fig. 1, and the Cartesian coordinates for all 214 trimers are provided in the SI. All restricted geometry optimizations were carried out using the MP2 method and the aug-cc-pVDZ basis set. Frequency calculations (in an appropriate subspace of parameters) were also carried out to confirm that a minimum has been reached in that subspace.

To ensure a seamless comparison with the previous 3BHET dataset,[18] the benchmark interaction energy calculations are performed using the same composite CCSD(T)/CBS scheme. The interaction energies of interest are the two-body and nonadditive three-body terms, where the latter is defined as the difference between the trimer interaction energy and the sum of the pair interaction energies.[28] Specifically, the complete interaction energy in the trimer

$$E_{\text{int}} = E_{\text{ABC}} - E_{\text{A}} - E_{\text{B}} - E_{\text{C}} \quad (1)$$

can be decomposed as

$$E_{\text{int}} = E^{2}_{\text{int}} + E^{3}_{\text{int}} \quad (2)$$

where the two-body part is the sum of pairwise interactions:

$$E^{2}_{\text{int}} = (E_{\text{AB}} - E_{\text{A}} - E_{\text{B}}) + (E_{\text{BC}} - E_{\text{B}} - E_{\text{C}}) + (E_{\text{AC}} - E_{\text{A}} - E_{\text{C}}) \quad (3)$$

The monomer and dimer calculations were performed in the full trimer basis set to eliminate basis set superposition error (BSSE). In addition, the interaction energies were computed using fixed monomer geometries ignoring all monomer deformation effects.[29] The benchmark interaction energies at both the two-body and three-body level were calculated using the composite CCSD(T)/CBS approach where correlation consistent basis sets fully augmented with diffuse functions were used; in this text, aug-cc-pVXZ (X = D, T, Q, 5) are abbreviated as aXZ. The benchmark calculations employ the well-kown composite CCSD(T)/CBS scheme that has contributions of the Hartree–Fock method (HF), the MP2 correlation energy, and a small-basis CCSD(T) correction:

$$E^{(\text{CCSD(T)/CBS})} = E^{\text{HF}} + E^{\text{MP2/CBS}}_{\text{corr}} + \delta E^{\text{CCSD(T)}} \quad (4)$$

In this formalism, a single HF calculation in a large basis set, a5Z basis for this purpose, is sufficient as it converges fast with the basis set size. The MP2 correlation energy is extrapolated to the CBS limit from the aQZ and a5Z basis sets using Helgaker's X^{-3} formula.[30] The coupled cluster correction $\delta E^{\text{CCSD(T)}}$ is calculated mostly in the aDZ basis set except for some small complexes where aTZ was used (the specific CCSD(T) basis set for each system is indicated in Tables SI–SIII in the SI). Thus, the benchmark interaction energies produced in this work correspond to the so-called silver standard of theory.[31,32] For the 38 trimers for which we computed the $\delta E^{\text{CCSD(T)}}$ term in the aTZ basis, the mean absolute deviations for this term relative to the aDZ basis amount to 0.064 kcal mol^{-1}

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

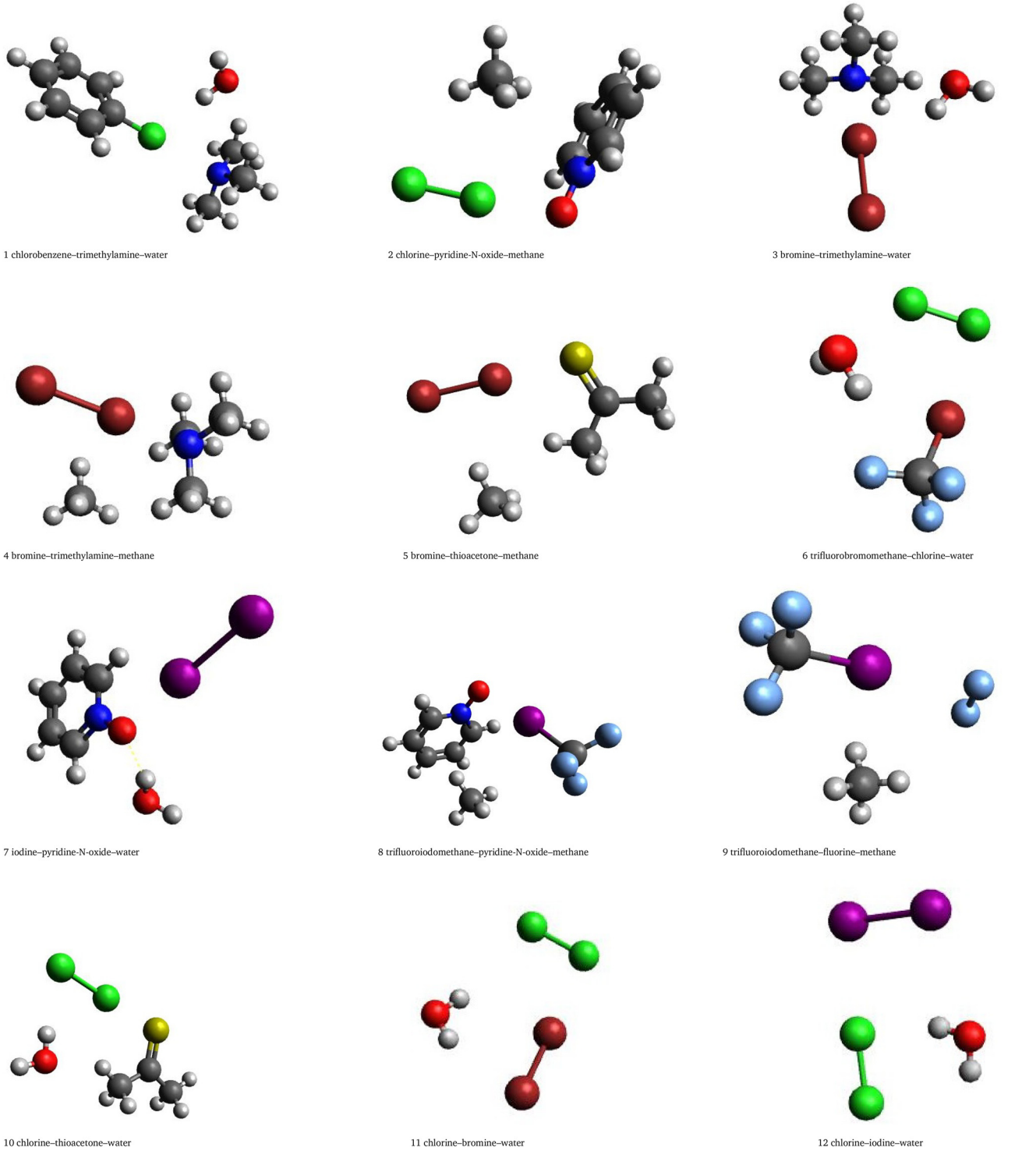

Fig. 1 Representative trimer structures from the 3BXB dataset.

for the complete E_{int} and only 0.004 kcal mol^{-1} for E^{3}_{int}, confirming that our "silver standard" level of theory is adequate for the purpose of the present work.

2.2 Two-body SAPT interaction energy decomposition

For appropriate interpretation of the supermolecular data obtained from the benchmark calculations, we need to know the nature of both the two- and three-body binding energies in each trimer. This is efficiently achieved through energy decomposition analysis such as SAPT. In this study, we employ several SAPT formalisms[28,33,34] to decompose the interaction energy into physically meaningful components. While we utilize various flavors of SAPT, the formalisms described below follow a set of general rules that apply across these methods. These

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

rules provide a consistent framework for interpreting the interaction energies and their components, regardless of the specific SAPT variant employed.

In the two-body and three-body SAPT based on the HF wavefunctions, the Hamiltonian is partitioned into the Fock operator F, the intermolecular interaction operator V, and the intramolecular correlation operator W encompassing all monomers:

$$H = F + V + W \tag{5}$$

where (in the three-body case)

$$F = F_A + F_B + F_C \quad V = V_{AB} + V_{BC} + V_{AC} \quad W = W_A + W_B + W_C \tag{6}$$

The simplest approximation that provides qualitatively reasonable total interaction energies is the SAPT0 one, which totally neglects the intramonomer electron correlation.[28] The two-body SAPT0 interaction energy is calculated as

$$E_{int}^{SAPT0} = E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} + \delta E_{HF}^{(2)} \tag{7}$$

In a SAPT correction $E^{(kl)}$ in eqn (7) and throughout the text, the subscripts k and l denote orders of perturbation theory with respect to V and W, respectively. SAPT0 provides the physical components that represent the electrostatic, first-order exchange, induction, exchange-induction (both with the inclusion of monomer relaxation/response effects as denoted by the additional subscript "resp"), dispersion, exchange-dispersion, and $\delta E_{HF}^{(2)}$ contributions. The $\delta E_{HF}^{(2)}$ term is meant to estimate the higher-order induction and exchange-induction effects from a supermolecular HF calculation:[28,34]

$$\delta E_{int}^{HF} = E_{int}^{HF} - E_{elst}^{(10)} - E_{exch}^{(10)} - E_{ind,resp}^{(20)} - E_{exch-ind,resp}^{(20)} \tag{8}$$

This energy decomposition allows one to classify complexes according to the relative importance of electrostatics, induction, and dispersion for their binding, for example, using a ternary diagram.[35]

To attain quantitative accuracy of total energies and their contributions, a higher-order SAPT level that includes both intra- and intermonomer electron correlation effects is required. In this study, the two-body interaction energy decomposition was performed by wavefunction-based SAPT up to the SAPT2+3 level of theory, simultaneously including the effects of the W operator on most SAPT0-level corrections and some terms that are third-order in V. Beyond SAPT0, the following levels of SAPT have been established:[34]

$$E_{int}^{SAPT2} = E_{elst}^{(10)} + E_{elst,resp}^{(12)} + E_{exch}^{(10)} + E_{exch}^{(11)} + E_{exch}^{(12)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{ind}^{(22)} + E_{exch-ind}^{(22)} + \delta E_{HF}^{(2)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} \tag{9}$$

$$E_{int}^{SAPT2+} = E_{int}^{SAPT2} + E_{disp}^{(21)} + E_{disp}^{(22)} \tag{10}$$

$$E_{int}^{SAPT2+(3)} = E_{int}^{SAPT2+} + E_{elst,resp}^{(13)} + E_{disp}^{(30)} \tag{11}$$

$$E_{int}^{SAPT2+3} = E_{int}^{SAPT2+(3)} + \left[\delta E_{HF}^{(3)} - \delta E_{HF}^{(2)} + E_{ind,resp}^{(30)} + E_{exch-ind,resp}^{(30)}\right] + \left[E_{exch-disp}^{(30)} + E_{ind-disp}^{(30)} + E_{exch-ind-disp}^{(30)}\right] \tag{12}$$

Eqn (9)–(12) show the progressive improvements to the SAPT0 approximation. In addition to those improvements, for levels of SAPT beyond SAPT2, we may include extra terms denoted δE_{int}^{MP2} and (CCD). The $\delta E_{int}^{MP2} = E_{int}^{MP2} - E_{int}^{SAPT2}$ term is based on a supermolecular MP2 interaction energy and may account for higher-order couplings between induction and dispersion as well as for some exchange effects, including those beyond the commonly used single exchange approximation.[34] The (CCD) notation in turn signifies that the second-order dispersion energy has been computed nonperturbatively, using the CCD + ST(CCD) dispersion approach introduced by Williams *et al.*[36] in place of the perturbative sum $E_{disp}^{(20)} + E_{disp}^{(21)} + E_{disp}^{(22)}$.

Alternative variants to the wavefunction-based SAPT approach, SAPT(DFT) and XSAPT, were also used to perform energy decompositions of both two- and three-body interaction energies. Here, we follow the same workflow as employed in our previous work for both SAPT(DFT) and XSAPT.[18] For SAPT(DFT), the theory based on the Kohn–Sham (KS) description of the monomers, the Fock operator is replaced by the Kohn–Sham operator, $K = K_A + K_B$ or $K = K_A + K_B + K_C$ for dimers and trimers, respectively. Then, the two-body SAPT(DFT) interaction energy is computed as

$$E_{int}^{SAPT(DFT)} = E_{elst}^{(1)}(KS) + E_{exch}^{(1)}(KS) + E_{ind}^{(2)}(CKS) + E_{exch-ind}^{(2)}(CKS) + E_{disp}^{(2)}(CKS) + E_{exch-disp}^{(2)}(CKS) + \delta E_{int}^{HF} \tag{13}$$

KS denotes standard uncoupled Kohn–Sham calculations (such SAPT(KS) quantities will also be used as components of the XSAPT approach below), CKS stands for the coupled Kohn–Sham, that is, the linear-response DFT approach to molecular frequency-dependent density susceptibilities (FDDS)[18,37] which are then used to compute SAPT(DFT) induction (static FDDS) and dispersion (imaginary-frequency FDDS). The exchange counterparts to dispersion and induction are also evaluated using the respective CKS amplitudes[37] (that is, no scaling of the uncoupled quantities is performed), and they are computed using the single-exchange approximation prevalent in standard SAPT.[33] The δE_{int}^{HF} term is once again taken from the HF-based SAPT theory, eqn (8). Since DFT fails to account for the correct asymptotic behavior of the exchange–correlation potential, the PBE0 functional used in all SAPT(DFT) computations was augmented with an asymptotic correction.[38]

The extended SAPT (XSAPT) approach allows for calculations for an arbitrary number of monomers with all-body induction interactions taken into account in an iterative manner.[15] This goal is achieved through the explicit polarization (XPol) approach[39] that solves the self-consistent field equations for

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

each fragment in an electrostatic embedding of all the other fragments in a form of point charges. The electron density obtained in this way is used to construct a new set of charges, and the process is iterated for all molecules until self-consistency is achieved. When combining SAPT with XPol, we obtain the total XSAPT interaction energy as

$$E_{\mathrm{int}}^{\mathrm{XSAPT}} = E_{\mathrm{elst}}^{(1)} + E_{\mathrm{exch}}^{(1)} + E_{\mathrm{disp+exch-disp}}^{(2)} + \Big[E_{\mathrm{ind}}^{(2)} + E_{\mathrm{exch-ind}}^{(2)} + \sum_{\mathrm{A}}\sum_{\mathrm{B>A}} \delta E_{\mathrm{AB}}^{\mathrm{HF}} + \sum_{\mathrm{A}}\sum_{\mathrm{B>A}} \left(E_{\mathrm{AB}}^{\mathrm{XSAPT}} - E_{\mathrm{AB}}^{\mathrm{SAPT(KS)}}\right) + E_{\mathrm{int}}^{\mathrm{MB}}\Big] \quad (14)$$

where the $E_{\mathrm{elst}}^{(1)}$ and $E_{\mathrm{exch}}^{(1)}$ terms represent the sum of these energy components over all pairs of dimers, and the $E_{\mathrm{disp+exch-disp}}^{(2)}$ term is modeled by a suitably adjusted many-body dispersion (MBD) approach.[40,41] The term in square brackets represents the total induction energy, which includes a many-body contribution,[42] and the δE^{HF} term is assumed to be pairwise additive. The $E_{\mathrm{AB}}^{\mathrm{XSAPT}}$ and $E_{\mathrm{AB}}^{\mathrm{SAPT(KS)}}$ quantities are total SAPT(KS) energies for dimer AB computed with and without the XPol embedding, respectively, and $E_{\mathrm{int}}^{\mathrm{MB}}$ denotes the pairwise-nonadditive effect of this embedding. We neglected the additional three-body induction coupling terms that are optional in XSAPT,[43] as they are expensive to compute, and our previous study found that the effect of those couplings is small.[18] Thus, XSAPT can take into account some nonadditive three-body (and higher-body) effects, namely induction (through the XPol iterative scheme) and dispersion (through the MBD formalism). Note that the SAPT(KS) calculations underlying the XSAPT scheme are computed without an explicit asymptotic correction; instead, the long-range corrected functional LRC-ωPBEh[44,45] was used, with the range separation parameter ω tuned separately for each molecule using the global density-dependent (GDD) scheme.[46]

2.3 Three-body SAPT energy decomposition and its alternatives

The energy decomposition of nonadditive three-body effects was performed by either XSAPT or SAPT(DFT). While the XSAPT method was described in the previous section including all-body terms, here we focus on the three-body SAPT(DFT) approach.[47] In this case, the electrostatic energy is rigorously pairwise additive and so is the second-order dispersion energy; however, important nonadditivity effects do show up in the third-order dispersion energy. The remaining corrections, most notably $E_{\mathrm{ind}}^{(2)}$, are not pairwise additive and, in the three-body SAPT(DFT) framework, the nonadditive interaction energy is computed as:[47]

$$E_{\mathrm{int,3}}^{\mathrm{SAPT(DFT)}} = \left(E_{\mathrm{exch,3}}^{(1)}(\mathrm{KS})\right)_{\mathrm{exch}} + \left(E_{\mathrm{ind,3}}^{(2)}(\mathrm{CKS}) + \tilde{E}_{\mathrm{exch-ind,3}}^{(2)}(\mathrm{CKS}) + \delta E_{\mathrm{int,3}}^{\mathrm{HF}}\right)_{\mathrm{ind}} + \left(\tilde{E}_{\mathrm{exch-disp,3}}^{(2)}(\mathrm{CKS}) + E_{\mathrm{disp,3}}^{(3)}(\mathrm{CKS})\right)_{\mathrm{disp}} \quad (15)$$

where the grouping in parentheses defines the overall splitting of nonadditive effects into exchange, induction, and dispersion terms as utilized below. As for the two-body energy, the first-order exchange is computed using the (asymptotically corrected) Kohn–Sham orbitals and orbital energies. In the second and third order, the induction and dispersion terms are computed from monomer FDDSs, and their exchange counterparts are approximated by scaling the respective uncoupled quantities as[47]

$$\tilde{E}_{\mathrm{exch-ind,3}}^{(2)}(\mathrm{CKS}) = E_{\mathrm{exch-ind,3}}^{(2)}(\mathrm{KS})\frac{E_{\mathrm{ind,3}}^{(2)}(\mathrm{CKS})}{E_{\mathrm{ind,3}}^{(2)}(\mathrm{KS})} \quad (16)$$

$$\tilde{E}_{\mathrm{exch-disp,3}}^{(2)}(\mathrm{CKS}) = E_{\mathrm{exch-disp,3}}^{(2)}(\mathrm{KS})\frac{E_{\mathrm{disp,3}}^{(3)}(\mathrm{CKS})}{E_{\mathrm{disp,3}}^{(3)}(\mathrm{KS})} \quad (17)$$

For the exchange-induction term, the scaling factor was restricted to the $[-2.0;2.0]$ range to prevent overestimation when the uncoupled induction term is way smaller than the CKS term (which might accidentally happen as these contributions change sign between different trimer configurations). The choice of a value 2.0 as the scaling restriction is completely arbitrary and made to avoid some particularly inappropriate scaling factors that would otherwise occur in rare cases within our dataset. Finally, the nonadditive HF delta term is defined from the supermolecular HF nonadditive energy and its three-body SAPT0 approximation as

$$\delta E_{\mathrm{int,3}}^{\mathrm{HF}} = E_{\mathrm{int,3}}^{\mathrm{HF}} - E_{\mathrm{exch,3}}^{(10)} - E_{\mathrm{ind,3}}^{(20)} - E_{\mathrm{exch-ind,3}}^{(20)} \quad (18)$$

Among the nonadditive components, the three-body dispersion is the most computationally demanding. Its many-body quantum nature makes it inherently difficult to calculate accurately. Previous studies have shown that the subtle three-body dispersion forces are overestimated by popular *ab initio* quantum mechanical approximations such as DFT[48] and they are completely missed by supermolecular MP2. To reduce costs, it is common to adopt a simplified approximation and model three-body dispersion by its leading Axilrod–Teller–Muto (ATM) asymptotic term.[49–51]

The damped ATM asymptotic dispersion model used in this work follows the workflow of ref. 52 and expresses the three-body dispersion energy as a summation over triplets of atoms in the complex, one from each molecule:[52,53]

$$E_{\mathrm{ATM}} = \sum_{\mathrm{a}\in\mathrm{A}}\sum_{\mathrm{b}\in\mathrm{B}}\sum_{\mathrm{c}\in\mathrm{C}} f_9^{\mathrm{abc}}(\beta) E_{\mathrm{ATM}}^{\mathrm{abc}} \quad (19)$$

where

$$E_{\mathrm{ATM}}^{\mathrm{abc}} = C_9^{\mathrm{abc}}\frac{(1+3\cos\theta_{\mathrm{a}}\cos\theta_{\mathrm{b}}\cos\theta_{\mathrm{c}})}{(R_{\mathrm{ab}}R_{\mathrm{ac}}R_{\mathrm{bc}})^3} \quad (20)$$

represents the triple-dipole ATM contribution in terms of interatomic distances R_{ij} and angles θ_i in the triangle connecting the three atoms. The appropriate atom–atom–atom dispersion coefficients C_9^{abc} are modeled as $C_9^{\mathrm{abc}} \approx \sqrt{C_6^{\mathrm{ab}}C_6^{\mathrm{ac}}C_6^{\mathrm{bc}}}$ using standard atom–atom dispersion coefficients C_6^{ij} from the DFT-D4

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

formalism.[54] At short range, it is critical to damp the ATM dispersion terms to prevent unphysical divergence as well as account for overlap effects. The damping function $f_9^{abc}(\beta)$ is chosen as a product of three two-body Tang–Toennies (TT) damping functions.[55,56]

$$f_9^{abc}(\beta) = f_6^{ab}(R_{ab}\beta) f_6^{ac}(R_{ac}\beta) f_6^{bc}(R_{bc}\beta) \quad (21)$$

$$f_6(R\beta) = 1 - \sum_{k=0}^{6} \frac{(R\beta)^k}{k!} e^{-\beta R} \quad (22)$$

with the damping parameter β estimated empirically as in ref. 52, which can be consulted for further details of the ATM algorithm employed here.

We performed the supermolecular benchmark and two-body SAPT(DFT) calculations using the MOLPRO program.[57] For all wavefunction-based SAPT calculations (from SAPT0 to SAPT2+3(CCD)δMP2), the Psi4 code was used.[58] The nonadditive SAPT(DFT) results were obtained using the three-body module of the SAPT package,[59] importing the DFT quantities from DALTON 2.0.[60] The XSAPT computations utilized the Q-Chem code.[61] Finally, the ATM three-body dispersion terms were computed by a code written by Austin Wallace (Georgia Tech): see **https://github.com/Awallace3/dispersion**.

3 Results and discussion

3.1 Benchmark interaction energies

Fig. 2 and 3 provide a summary distribution of the total benchmark interaction energies as well as their nonadditive three-body contributions. The individual benchmark values for each system are provided in Tables SIV–SIX in the SI. The reference total interaction energy ranges from −2.19 to −21.76 kcal mol^{-1} for water complexes and from −1.48 to −17.59 kcal mol^{-1} for methane complexes. The trimers with water show overall stronger binding (average interaction energy of −7.83 kcal mol^{-1}) than the trimers with methane (average interaction energy of −5.37 kcal mol^{-1}), with I showing the largest interactions; the average trimer interaction energy amounts to −8.19 kcal mol^{-1} for iodine–water complexes compared to −7.53 kcal mol^{-1} for bromine–water and −7.69 kcal mol^{-1} for chlorine–water ones. This can be linked to the size and polarizability of the halogen atom involved in the interaction. The benchmark nonadditive contribution can be either attractive or repulsive and ranges between −2.87 and +0.69 kcal mol^{-1} for trimers with water, and only between −0.37 and +0.18 kcal mol^{-1} for trimers with methane. Complexes 5 and 10 in Fig. 1 show the systems that exhibit the largest nonadditive effects for methane and water complexes, respectively, in the 3BXB dataset.

Our CCSD(T)-level interaction energies for the halogen-bonded dimers, listed in the last columns of Tables SIV–SIX, are expected to be close to, but not identical, as the respective benchmark values reported for those dimers in ref. 10 due to several technical differences. We observed that, for systems with chlorine as the XB donor, both sets of values are very close (a mean absolute deviation (MAD) of 0.03 kcal mol^{-1}), with the small differences mostly due to the basis set used for the CCSD(T) calculations (it was heavy-aug-cc-pVTZ in ref. 10 and either aDZ or aTZ here). When the XB donor is iodine, the MAD between both sets of data increases to 0.18 kcal mol^{-1}, with the primary source of this difference being the correlation effects coming from the sub-valence shell of iodine, included in ref. 10 but neglected in this work. Finally, for bromine as the XB donor, the two sets of values exhibit the largest MAD of 0.31 kcal mol^{-1} due to both sub-valence correlation and relativistic effects, which were included (*via* an effective core potential) in ref. 10 but not in this work. This is a moderate difference larger than what is typically expected as benchmark accuracy, however, only a part of it is relevant for our calculations since we never compare relativistic results to nonrelativistic ones (the relativistic effects will be always neglected for bromine and always included for iodine).

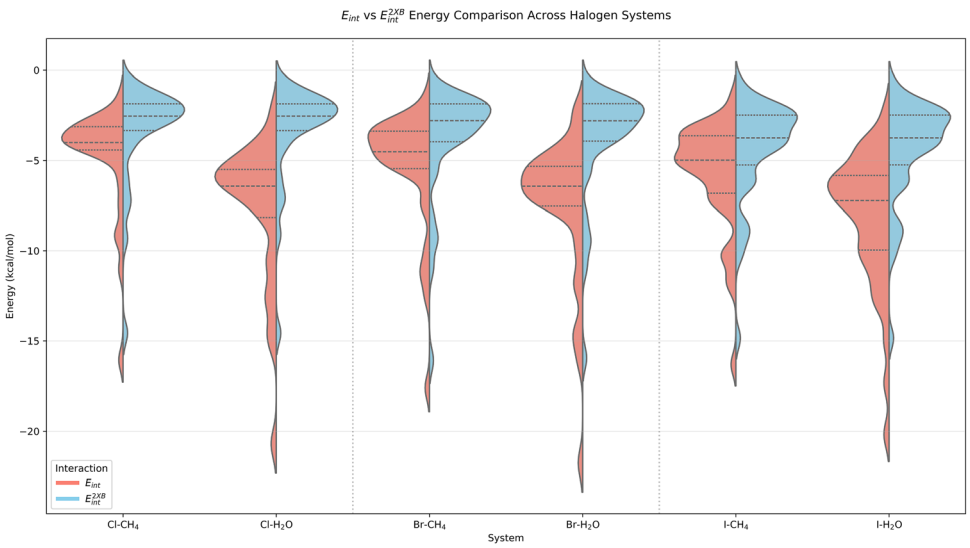

Fig. 2 The distribution of complete interaction energies E_{int} for the 3BXB complexes, partitioned by the halogen and by the third interaction partner (H_2O/CH_4), compared to the respective halogen-bonded value E_{int}^{2XB}, that is, the interaction energy for the unaltered halogen-bonded dimer AB.

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

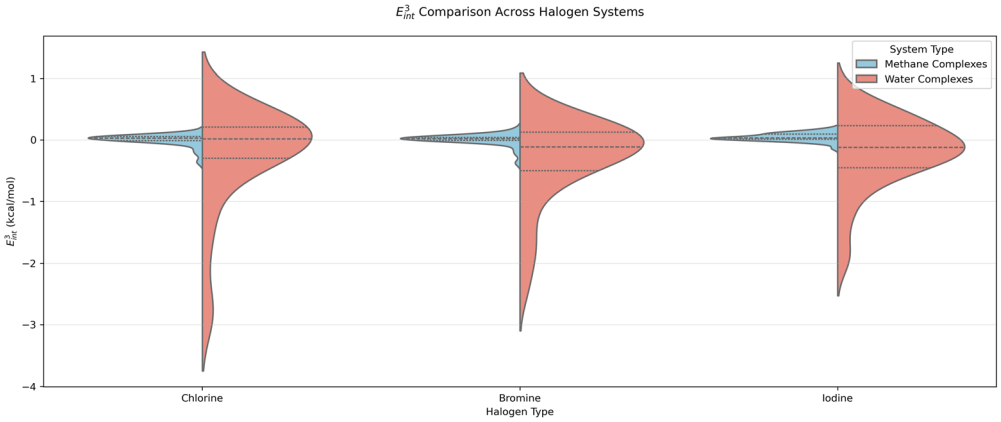

Fig. 3 The distribution of the nonadditive three-body energies E^3_{int} for the 3BXB complexes, partitioned by the halogen and by the third interaction partner (H_2O/CH_4).

In the next subsections, we will analyze the performance of various SAPT flavors in addition to the MP2 method applied to our set of σ-hole noncovalent interactions. The MP2/CBS data includes the correlation energy extrapolated to the CBS limit from aQZ and a5Z. The wavefunction SAPT (SAPT0 to SAPT2+3(CCD)δMP2), two-body SAPT(DFT), and XSAPT calculations were performed in the aTZ basis. The three-body SAPT(DFT) calculations were only feasible for complexes not involving iodine, and were performed in jun-cc-pVDZ for complexes with more than 25 atoms and aDZ otherwise. The XSAPT calculations are also problematic for iodine complexes (because of the need of an explicit-core-potential treatment of inner electrons), but we were able to perform incomplete XSAPT to extract the MBD three-body dispersion data. Further details on auxiliary basis sets, convergence thresholds, and other technical details of our calculations are provided in the SI. We will first investigate the individual interaction energy contributions, at the two-body and three-body levels, before moving on to examining the accuracy of the entire interaction energies using mean unsigned errors (MUE) relative to the CCSD(T)-level benchmark values.

3.2 SAPT two-body interaction energy contributions

Tables SX–SXXV in the SI show the two-body interaction energy components for the 3BXB structures computed using SAPT2+3(CCD)δMP2, SAPT(DFT), and XSAPT. The electrostatic, induction, and dispersion components to the interaction energy are all attractive for those minimum structures. The magnitude of the interaction energy components depends on the halogen type with additional variation influenced by the presence of water or methane in the complex.[3] The trimers constructed by adding water to the pre-existing dimers are predominantly bound by electrostatics with the exception of the aromatic complexes that show larger dispersion magnitudes. Second-order dispersion is the leading term for complexes constructed with methane, with a few exceptions for systems with Cl_2, Br_2, and I_2 as the donor. This behavior is expected because H_2O is a highly polar molecule, therefore significantly contributing to electrostatics whereas a nonpolar methane molecule drives dispersion forces. On the other hand, complexes with Cl_2, Br_2, and I_2 (for example, structures 11 and 12 in Fig. 1) exhibit strong electrostatic interactions because of the increased polarizability of the halogen atom.[3] With an increase in size, and thus polarizability, of the halogen atom, the largest contributors to the interaction, namely electrostatics and dispersion, also get stronger. In addition, the presence of electron withdrawing groups enlarges the σ-hole by pulling electron density away from the halogen, strengthening the interaction, while electron-donating groups act in the opposite direction.[3] A detailed SAPT2+(CCD)δMP2 energy decomposition analysis of the 3BXB dataset shows that the interactions span a relatively small section of the ternary diagram with a varying but fairly equal mix between dispersion and electrostatics (Fig. 4). The two-body energy decomposition highlights that the methane complexes cluster towards the dispersion dominated region while the water complexes towards the electrostatic corner, also with a stronger induction influence. The chlorine and bromine complexes tend to exhibit similar bonding, clustering in the same areas of the ternary diagram.

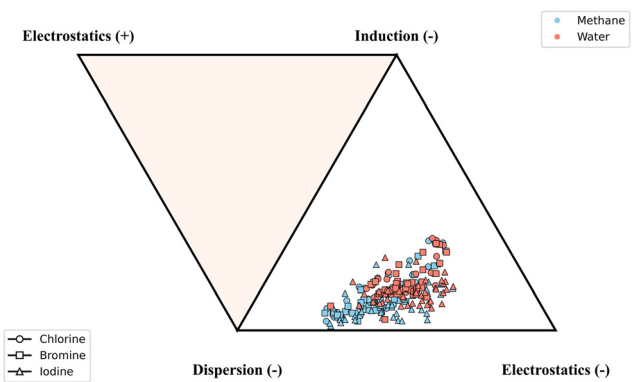

Fig. 4 Two-body SAPT2+(CCD)δMP2 interaction energy distribution for different parts of the 3BXB dataset (separated by the XB halogen as well as water/methane), displayed on a ternary diagram.

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

On the other hand, many but not all iodine complexes cluster lower on the triangle, displaying somewhat diminished induction contributions.

As we will show in Section 3.4, the two-body SAPT2+(CCD)δMP2 level of theory recovers very well, and better than any other SAPT variant, the benchmark CCSD(T) pairwise additive interaction energies. Therefore, we will examine the interaction energy component differences from various SAPT flavors treating the SAPT2+(CCD)δMP2 values as reference. Tables SXXXVII–SXL in the SI show the mean signed and unsigned deviations of each SAPT component between a given variant and the reference one. We see that the simplest SAPT0 variant generally overbinds (the interaction energies are too negative), primarily because the simple $E^{(10)}_{\mathrm{exch}}$ expression underestimates the reference first-order exchange energy (on the average by 1.72 kcal mol^{-1} for water complexes and 1.26 kcal mol^{-1} for methane complexes). In contrast, XSAPT significantly underbinds, and the primary reason is underestimating the magnitude of the dispersion energy, on the average by 2.11 kcal mol^{-1} for either water or methane complexes (this particular value is averaged over non-iodine systems only). It seems like the short-range adjustment of the MBD formalism to adapt to XSAPT[41] might benefit from a reoptimization for halogen-bonded systems. The SAPT(DFT) dispersion energy is also underestimated in magnitude, albeit by much less (by 1.12 and 0.99 kcal mol^{-1} on the average for trimers with water and methane, respectively) and again this underestimation is the leading cause of the SAPT(DFT) underbinding. The correlated wavefunction-based SAPT variants exhibit smaller differences relative to one another, and these differences are harder to interpret. Note that the reference electrostatic energy is formally more approximated than the SAPT2+(3) and SAPT2+3 one because it omits the $E^{(13)}_{\mathrm{elst,resp}}$ term, and the $E^{(30)}_{\mathrm{ind-disp}}$ + $E^{(30)}_{\mathrm{exch-ind-disp}}$ contribution is computed explicitly and classified as dispersion in SAPT2+3δMP2, but forms a part of the δMP2 term, classified as induction in SAPT2+(3)δMP2.[34] The inclusion of the CCD + ST(CCD) dispersion terms beyond $E^{(20)}_{\mathrm{disp}}$ + $E^{(21)}_{\mathrm{disp}}$ + $E^{(22)}_{\mathrm{disp}}$ makes the dispersion energy slightly less negative, by about 0.3 kcal mol^{-1} on the average for either class of trimers. Such a change is similar in magnitude to the differences in other SAPT corrections and may be beneficial or detrimental, however, it does strongly benefit the accuracy at the reference SAPT2+(CCD)δMP2 level.

3.3 SAPT analysis of the three-body nonadditive interaction

The SAPT energy decomposition of the nonadditive interaction in the 3BXB dataset shows that the three-body induction dominates while the three-body exchange and dispersion play smaller roles. Fig. 3, 5 and 6 present an analysis of the nonadditive interactions. Here, we can no longer use the two-body SAPT2+3(CCD)δMP2 approach, but instead shift our attention to SAPT(DFT) and XSAPT. As mentioned before, these nonadditive interaction energy calculations were not feasible for iodine complexes. Consequently, Fig. 5 and 6 include data only from the chlorine complexes, and we present analogous figures for the bromine complexes (and partial nonadditive dispersion data for the iodine ones) in the SI. As explained

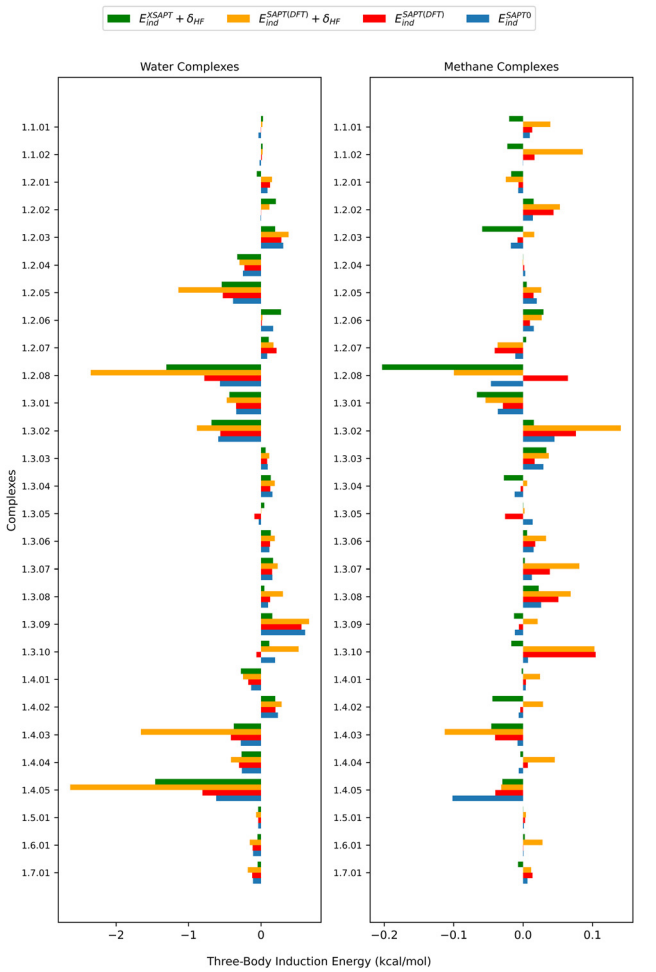

Fig. 5 Comparison of the nonadditive three-body induction energies for the systems with chlorine as the XB donor, predicted by SAPT0, XSAPT, and SAPT(DFT). The system labels are defined in the SI; the first '1.' part denotes chlorine. Note the different horizontal scale in the left and right panels.

above, SAPT(DFT) includes all the leading order nonadditive effects while XSAPT neglects the nonadditive first-order exchange component. XSAPT includes, however, nonadditive induction applicable to any number of molecules (not just three) and can be used with or without the three-body induction couplings.[43] As our earlier study[18] indicates that the latter couplings play a minor role, we neglect them here.

We now examine the SAPT(DFT) energy decomposition for the (non-iodine) 3BXB dataset, with the individual components for each system given in Tables SXXVI–SXXIX in the SI. The nonadditive first-order exchange is predominantly attractive (negative for 112 systems and positive for 12 ones) and ranges from −0.49 to 0.20 kcal mol^{-1} for the water complexes and from −0.31 to 0.003 kcal mol^{-1} for the methane ones. The nonadditive dispersion in this dataset is repulsive, aiming to remedy the overbinding that might come from the two-body dispersion contribution. This component at the SAPT(DFT) level contributes up to 0.33 kcal mol^{-1} for complexes with water and up to 0.31 kcal mol^{-1} for complexes with methane. Conversely, the nonadditive induction can be either attractive

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

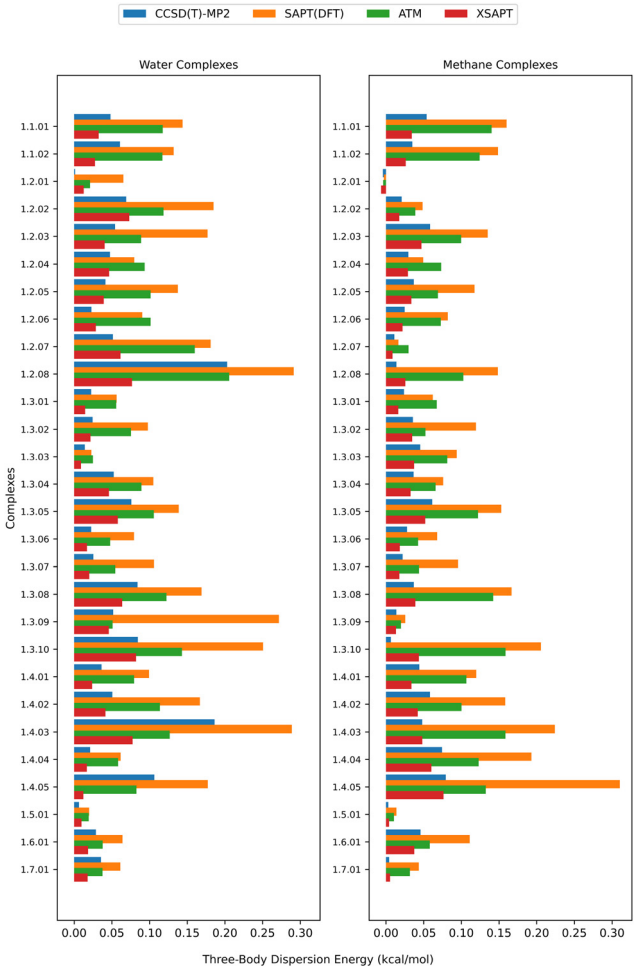

Fig. 6 Nonadditive three-body dispersion energy for the systems with chlorine as the XB donor, estimated as the supermolecular CCSD(T)/CBS–MP2/CBS difference as well as computed with SAPT(DFT), XSAPT (that is, MBD) and estimated through the asymptotic ATM term. The system labels are defined in the SI; the first '1.' part denotes chlorine.

or repulsive, and it ranges from −2.63 to 0.66 kcal mol^{-1} for trimers involving water and from −0.24 to 0.14 kcal mol^{-1} for trimers involving methane. An in-depth analysis of the nonadditive induction is shown in Fig. 5. Here, we compare this three-body component computed by SAPT(DFT), SAPT0, and XSAPT. The SAPT(DFT) variant can include or omit the nonadditive δE^{HF}, while the δE^{HF} contribution in XSAPT is assumed to be pairwise additive and thus has no effect on the nonadditive induction. The water-containing systems in the 3BXB dataset follow trends consistent with our previous work on heteromolecular trimers,[18] where we observed that SAPT0 and SAPT(DFT) without the δE^{HF} correction have similar values of nonadditive induction. The incorporation of the δE^{HF} term into SAPT(DFT) results in an increase in magnitude making the SAPT(DFT) + δE^{HF} values the largest, followed by XSAPT. Overall, we do expect the inclusion of δE^{HF} to improve accuracy as this term covers physical effects of nonadditive higher-order induction.

Fig. 6 shows the different dispersion estimates from the supermolecular CCSD(T)–MP2, XSAPT, SAPT(DFT), and ATM approaches. Note that while supermolecular MP2 entirely misses the three-body dispersion, the CCSD(T)–MP2 difference is not strictly due to dispersion energy as the molecular multipole moments, polarizabilities, and electron densities somewhat change between MP2 and CCSD(T), leading to differences in other contributions. Nevertheless, this difference has been previously taken as a measure of the three-body dispersion.[62] Consistent with previous observations from Huang and Beran,[51] Carter-Fenk *et al.*,[41,63] and our earlier work,[18] the SAPT(DFT) and ATM models yield the largest dispersion estimates. While both of these methods are able to capture the repulsiveness of the many-body dispersion, they tend to overestimate it as they both neglect higher-order effects that are again attractive.[41,51] Indeed, relative to the leading asymptotic term (that is, ATM), SAPT(DFT) amplifies the nonadditive dispersion (through the inclusion of repulsive third-order effects beyond the triple dipole approximation) while XSAPT (that is, MBD) reduces it (through an effective inclusion of some attractive higher-order terms). The MBD estimate, which is obtained from the supermolecular trimer–dimer differences, has the smallest magnitude and is mostly consistent with CCSD(T)–MP2, especially for the trimers involving methane. The reasonable agreement between XSAPT (MBD) and CCSD(T)–MP2 suggests that both methods can reliably capture nonadditive dispersion in our systems. The MBD approach has been previously observed to produce better three-body interaction energies when combined with supermolecular MP2 than on top of supermolecular DFT since, in the latter case, MBD would have to account for deficiencies in the underlying density functional, which is not possible with a single damping/switching parameter.[64] In particular, unlike MP2, supermolecular DFT is not entirely free of three-body dispersion, and it has trouble correctly reproducing the nonadditive first-order exchange.[65] Those issues do not apply to either XSAPT or SAPT(DFT), so MBD has the potential to perform well in our case as long as the damping parameter value is suitable.

3.4 Total SAPT interaction energies

So far, we have investigated the physical contributions to the two-body and three-body interaction energies in the 3BXB trimers as provided by various variants of SAPT. As we want to exploit the full strength of SAPT, that is, the combination of a meaningful energy decomposition and accurate total interaction energies, we now examine the accuracy of complete interaction energies from different SAPT variants relative to our CCSD(T)-level benchmark data. The interaction energy in a cluster is dominated by its pairwise-additive term, so we first assess the accuracy of two-body SAPT interaction energies. Subsequently, we combine the best two-body SAPT performers with various measures of the three-body nonadditive contribution to identify the variants that most closely reproduce the reference total energies in the 3BXB trimers. In this analysis, we do include hybrid SAPT approaches, in which the two-body and three-body contributions are obtained from different

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

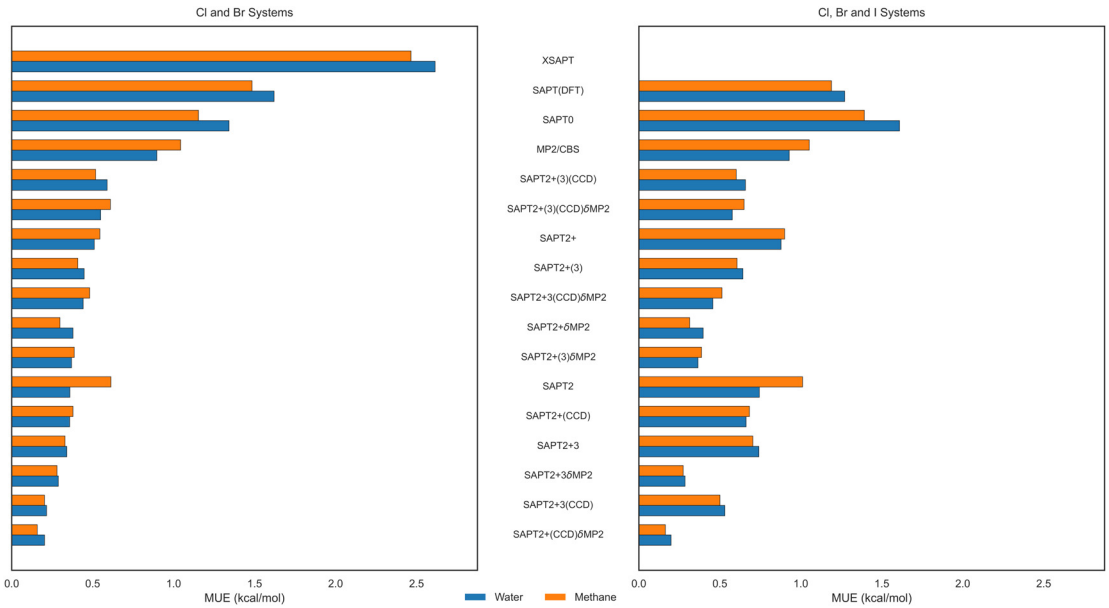

Fig. 7 Mean unsigned errors (MUE, kcal mol^{-1}) of various two-body interaction energy estimates relative to the benchmark CCSD(T)-level E^{2}_{int} values. The values in the left panel are averaged over the 124 complexes containing chlorine and bromine only, while the right panel values are averaged on the entire 214-system dataset.

SAPT flavors. We even allow three-body estimates obtained by combining different sources, for example, the SAPT(DFT) nonadditive induction and first-order exchange with the MBD nonadditive dispersion.

We consider a large variety of SAPT variants with different intramolecular electron correlation treatments. These include SAPT0, where dispersion is introduced through second-order perturbation theory and the monomers are treated at the HF level (no intramolecular correlation), along with its progressive improvements through SAPT2, SAPT2+, SAPT2+(3), and SAPT2+3. In the higher-level wavefunction SAPT variants (SAPT2+ and above), we also explore including additional corrections to dispersion (through the CCD + ST(CCD) level) and induction (the δMP2 term). While these wavefunction-based SAPT treatments are only available for the two-body interaction energy, we also consider the SAPT(DFT) and XSAPT approaches which provide both the two-body and three-body decomposition.

The mean unsigned errors of different approaches relative to the reference CCSD(T)-level data are presented in Fig. 7–9 for the two-body, nonadditive three-body, and total trimer interaction energies, respectively. Since our perturbative calculations of three-body contributions, both SAPT(DFT) and XSAPT, are limited to complexes not involving iodine, the errors in Fig. 8 are averaged over the 124 non-iodine complexes (62 each with water and methane). In the two-body case, the analogous MUE values in Fig. 7 are given both for the 124 non-iodine systems (left panel) and for the full set of 214 complexes (right panel).

The ordering of the two-body errors presented in Fig. 7 is overall quite expected. The lowest SAPT0 level of theory leads to the largest MUE value on the whole dataset, although XSAPT yields the highest errors for the reduced dataset for which it was computed. Interestingly, the SAPT(DFT) variant is overall as inaccurate as SAPT0, and both SAPT0 and SAPT(DFT) exhibit comparable errors for the water and methane trimers, indicating that the error of the sum $E^{\text{AB}}_{\text{int}} + E^{\text{BC}}_{\text{int}} + E^{\text{CA}}_{\text{int}}$ is dominated by the (common to both) halogen-bonded dimer $E^{\text{AB}}_{\text{int}}$. All correlated levels of wavefunction-based SAPT, starting from SAPT2, are more accurate than SAPT(DFT), and they also all surpass the accuracy of supermolecular MP2. Among these correlated variants, those that include the δMP2 correction are clearly superior, indicating that, without this term, the SAPT data are missing some important induction-dispersion couplings or some exchange effects beyond the commonly used approximations. The best performing SAPT variant happens to be SAPT2+(CCD)δMP2, as it clearly combines a high level of SAPT theory with a favorable cancellation of remaining residual errors. SAPT2+(CCD)δMP2 leads to impressively low MUE values of 0.20 (trimers with water) and 0.16 kcal mol^{-1} (trimers with methane) when averaged over either the full 214-complex dataset or the 124-complex reduced non-iodine dataset. It should be noted that many other variants of SAPT perform worse on the complexes containing iodine than for the lighter halogens, which leads to SAPT errors typically becoming smaller when evaluated over the reduced dataset.

We move on to the nonadditive three-body interaction energy contributions displayed in Fig. 8. Here, our assortment of methods has shrunk since the correlated wavefunction-based SAPT approaches are only available in the two-body case. In the context of the energy decomposition presented above, only SAPT(DFT) can account for all leading sources of three-body nonadditivity, while XSAPT is missing first-order nonadditive exchange and supermolecular MP2 is missing

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

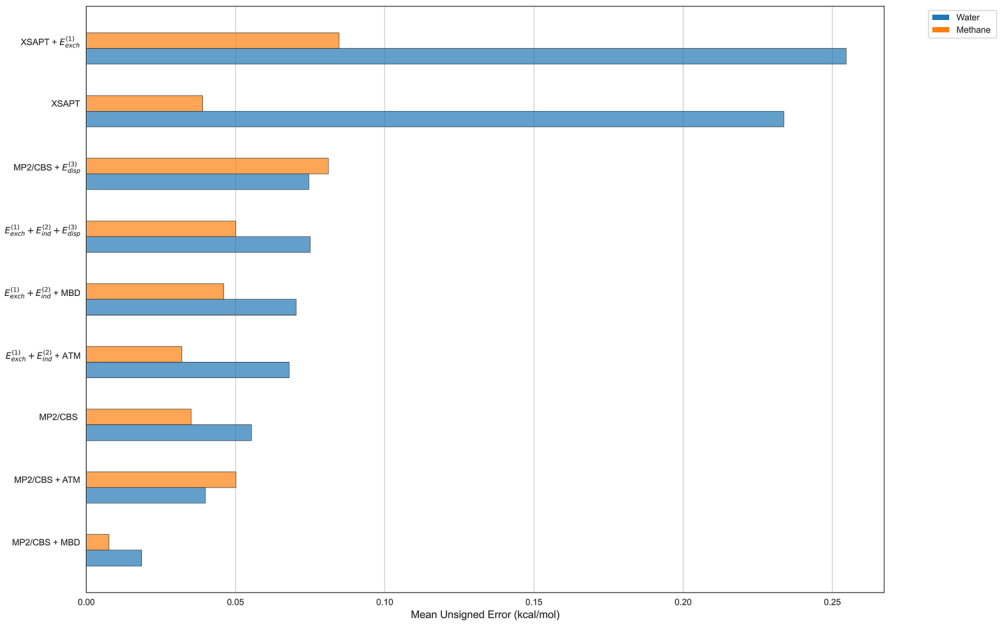

Fig. 8 Mean unsigned errors (MUE, kcal mol^{-1}) of various nonadditive three-body interaction energy estimates relative to the benchmark CCSD(T)-level E^{3}_{int} values. The corrections $E^{(1)}_{exch}$, $E^{(2)}_{ind}$, and $E^{(3)}_{disp}$ are taken from three-body SAPT(DFT), and their sum $E^{(1)}_{exch}$ + $E^{(2)}_{ind}$ + $E^{(3)}_{disp}$ is by definition the SAPT(DFT) three-body interaction energy.

nonadditive dispersion. On the other hand, three-body SAPT(DFT) has been observed, here and earlier,[47] to overestimate nonadditive dispersion, so it is worthwhile to examine the performance of hybrid SAPT(DFT)-based variants replacing its CKS dispersion by either MBD (as included in XSAPT) or the leading asymptotic term, ATM. The same measures of dispersion can be added to supermolecular MP2.

The errors of different nonadditive three-body estimates in Fig. 8 are overall quite similar to each other, with an exception of XSAPT which is less accurate for the trimers with water. At first, one could think that this is due to XSAPT neglecting nonadditive first-order exchange. However, the addition of the SAPT(DFT) $E^{(1)}_{exch}$ estimate to XSAPT makes the errors larger, not smaller, leading us to believe that the issue with XSAPT might be its comparatively inaccurate description of nonadditive induction. The MUE values of all other methods do not exceed 0.1 kcal mol^{-1} for either interaction partner (water or methane), which is significantly lower than the two-body errors, indicating that any of those variants is in principle an acceptable strategy to account for nonadditive effects in our trimers. The complete SAPT(DFT) nonadditive energy has a very reasonable average error of 0.075 and 0.050 kcal mol^{-1} for the trimers with water and methane, respectively. Supermolecular nonadditive MP2 performs slightly better (with MUE of 0.055 and 0.035 kcal mol^{-1}, respectively, for the water and methane trimers), and further improvement to the accuracy can be attained by hybrid nonadditive approaches. The SAPT(DFT) nonadditive exchange and induction performs best when combined with the ATM three-body dispersion estimate (with water and methane MUE values of 0.068 and 0.032 kcal mol^{-1}, respectively) and it can also be combined with the MBD three-body dispersion for a very similar accuracy. The closest recovery of the benchmark nonadditive energies, with the water and methane MUE amounting to a respective 0.019 and 0.008 kcal mol^{-1}, is afforded by combining the supermolecular MP2 method, which accounts for the exchange and induction nonadditivity, with the three-body dispersion estimate from MBD. While such a good agreement has to be to some extent accidental, it does lend even more credibility to MBD as the truthful descriptor of nonadditive dispersion effects in our halogen-bonded clusters.

We now analyze how different combinations of two-body and three-body approaches influence the accuracy of total interaction energies in the 3BXB clusters, employing the MUE values for supermolecular, SAPT, and hybrid approaches displayed in Fig. 9. As expected, the two-body accuracy is the most important factor, and the best two-body performer, SAPT2+(CCD)δMP2, leads to reasonable MUE values of 0.48 and 0.15 kcal mol^{-1} for water and methane trimers, respectively (computed on the 124-member reduced non-iodine dataset), even without any nonadditive three-body estimate. This level of accuracy is not attained by any available method that provides both two-body and three-body interaction energies, including SAPT(DFT), XSAPT, and supermolecular MP2. Only hybrid approaches can further improve the accuracy of the total interaction energies, and it is no surprise that the best hybrid variants take the best two-body performer, SAPT2+(CCD)δMP2, and augment it with an estimate of the nonadditive three-body effects. For the latter estimate, all choices (including a choice not to add anything at all) perform equally well for the methane-containing trimers, with all MUE values in the 0.15–0.16 kcal mol^{-1} range. For the water-containing trimers, there

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

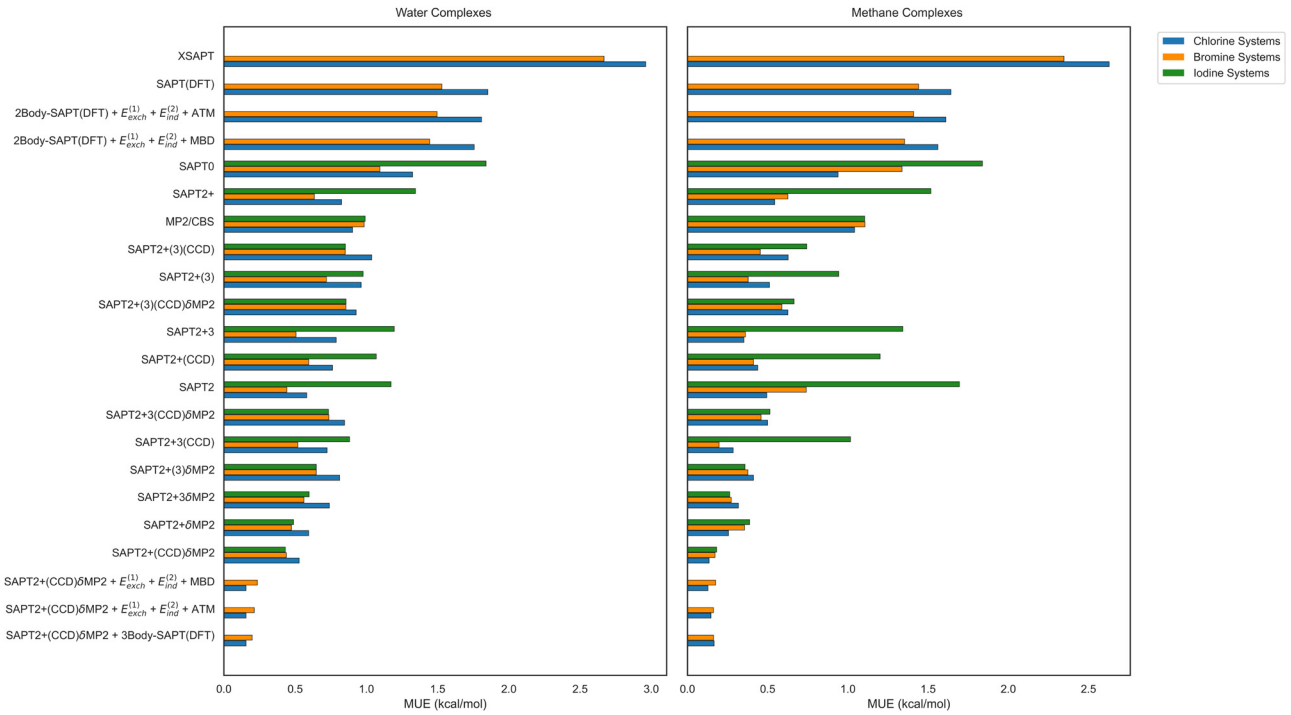

Fig. 9 Mean unsigned errors (MUE, kcal mol^{-1}) of the total (two-body plus three-body) interaction energy estimates relative to the benchmark CCSD(T)-level complete E_{int} values, grouped by the identity of the halogen bond donor. For consistency between different methods, the data with chlorine and bromine as the XB donor do not contain the 6 systems (3 with water and 3 with methane) involving iodine in a different capacity (for which the XSAPT and three-body SAPT(DFT) data are not available).

is a bit more variability, and the lowest MUE of 0.18 kcal mol^{-1} is afforded by the addition of the complete SAPT(DFT) three-body energy to the SAPT2+(CCD)δMP2 value. However, nearly the same accuracy is obtained if one keeps the SAPT(DFT) nonadditive induction and first-order exchange, but replaces its nonadditive dispersion by either MBD or ATM. Overall, the best strategy to compute both accurate trimer interaction energies and their decomposition into physically meaningful terms appears to be combining a correlated wavefunction-based two-body SAPT variant utilizing the δMP2 correction with any SAPT(DFT)-like decomposition of the three-body energy, with dispersion either taken from SAPT(DFT) itself, computed with MBD, or approximated by its leading ATM term.

For the two methods that show the worst total interaction energy accuracy, XSAPT and SAPT(DFT), the biggest outliers are the chlorine–trimethylamine–water/methane trimers for both methods. These systems show respective errors of 8.59 and 6.26 kcal mol^{-1} for XSAPT and 6.80 and 6.17 kcal mol^{-1} for SAPT(DFT). For the two-body interaction energy only, the worst errors pertain to XSAPT with the chlorine–trimethylamine–water/methane complexes (7.09 and 6.08 kcal mol^{-1}, respectively), SAPT(DFT) also with chlorine–trimethylamine–water/methane (6.32 and 5.90 kcal mol^{-1}, respectively), and SAPT0 for the iodine–pyridine-*N*-oxide–water/methane systems (−7.19 and −5.21 kcal mol^{-1}, respectively). Most methods in Fig. 9 perform very similarly for all XB donors, Cl, Br, and I (when the latter is available). One exception is the wavefunction-based SAPT without δMP2 being somewhat less accurate for iodine complexes (this difference is completely alleviated when δMP2 is added). Second, all SAPT variants seem to perform a little worse for chlorine–water systems than for bromine–water ones. However, the differences are not particularly significant.

4 Conclusions

In this work, we constructed a new 3BXB benchmark dataset which contains noncovalently interacting trimers with two halogen-bonded molecules taken from the SH250 dataset[10] and either water or methane as the third partner. The benchmark interaction energies and their nonadditive three-body contributions were generated at the composite CCSD(T) level of theory, with basis sets appropriate for a ''silver-standard'' calculation. The new 214-member dataset helps alleviate the scarcity of benchmark-quality nonadditive interaction data, especially for trimers composed of three different molecules, where our previous 20-member dataset[18] was, to our knowledge, the only one of its kind.

The decomposition of both two- and three-body interaction energies into physically meaningful terms provides key insights into the nature of specific interactions. Therefore, we used several variants of SAPT to analyze the physical origins and cooperativity of the binding in the 3BXB systems. On the two-body level, the high-level wavefunction-based SAPT variants

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

including the so-called δMP2 correction accurately reproduced the benchmark total interaction energies, and indicated that the 3BXB complexes are held together, in various proportions, by electrostatic and dispersion forces. The nonadditive three-body interaction energy can in turn be decomposed into the effects of first-order exchange (mostly attractive), induction (dominant, can be attractive or repulsive), and dispersion (repulsive). While the first of those terms is only provided by SAPT (in this case, the three-body SAPT(DFT) variant of ref. 47), the other two terms can be computed from various sources. As far as nonadditive induction is concerned, the SAPT(DFT) and XSAPT values are reasonably consistent only when the δHF term is not included in the former; however, this term represents important high-order induction effects and should be normally taken into account. The nonadditive dispersion energies, while small, display quite significant inconsistencies. Relative to the leading-order asymptotic ATM contribution, the three-body SAPT(DFT) mostly predicts dispersion energy enhancement while MBD (as included in the XSAPT formalism) mostly suggests a reduction in magnitude for this effect. The supermolecular CCSD(T)–MP2 interaction energy difference, which does include three-body dispersion but also contains other effects, also predicts reduced dispersion energies relative to ATM but exhibits some inconsistencies relative to MBD. It appears that further analysis and benchmarking of the nonadditive dispersion effects is needed to provide a more comprehensive understanding.

Different SAPT variants display a broad range of accuracy relative to the reference CCSD(T)-level interaction energies. Some high-level variants of wavefunction-based SAPT, especially SAPT2+(CCD)δMP2, are so accurate for the dominant two-body interaction energy term that they overperform SAPT(DFT) and XSAPT (variants that can be used for all two-body and three-body terms) even as they are, neglecting the nonadditive contributions entirely. However, the highest interaction energy accuracy is attained by hybrid approaches generated by mixing and matching different two-body and three-body estimates. Overall, the best strategy that yields both accurate interaction energies and their physical decomposition involves augmenting a high-level wavefunction-based two-body SAPT energy such as SAPT2+(CCD)δMP2 with the SAPT(DFT) estimates of nonadditive induction and first-order exchange plus *any* estimate of nonadditive dispersion, even from the simple damped asymptotic ATM expression.

Conflicts of interest

There are no conflicts to declare.

Data availability

The numerical data computed in the course of this project are available in the article and the supplementary information (SI). The latter document contains additional tables and figures as well as the XYZ geometries of all studied complexes. Supplementary information is available. See DOI: **https://doi.org/10.1039/d5cp02809d**.

Acknowledgements

This work was supported by the U.S. National Science Foundation award CHE-1955328. We thank Austin Wallace (Georgia Tech) for the code to compute the ATM three-body contribution.

References

1 P. Jurečka, J. Šponer, J. Černý and P. Hobza, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1985–1993.
2 J. Řezáč, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2011, **7**, 2427–2438.
3 J. Řezáč, K. E. Riley and P. Hobza, *J. Chem. Theory Comput.*, 2012, **8**, 4285–4292.
4 S. Grimme, *Chem. – Eur. J.*, 2012, **18**, 9955–9964.
5 L. A. Burns, J. C. Faver, Z. Zheng, M. S. Marshall, D. G. Smith, K. Vanommeslaeghe, A. D. MacKerell, K. M. Merz and C. D. Sherrill, *J. Chem. Phys.*, 2017, **147**, 161727.
6 Z. M. Sparrow, B. G. Ernst, P. T. Joo, K. U. Lao and R. A. DiStasio, Jr., *J. Chem. Phys.*, 2021, **155**, 184303.
7 J. Řezáč, *J. Chem. Theory Comput.*, 2020, **16**, 2355–2368.
8 J. Řezáč, *J. Chem. Theory Comput.*, 2020, **16**, 6305–6316.
9 K. Kříž, M. Nováček and J. Řezáč, *J. Chem. Theory Comput.*, 2021, **17**, 1548–1561.
10 K. Kříž and J. Řezáč, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14794–14804.
11 J. Řezáč, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14780–14793.
12 U. Góra, R. Podeszwa, W. Cencek and K. Szalewicz, *J. Chem. Phys.*, 2011, **135**, 224102.
13 G. J. Beran, *Chem. Rev.*, 2016, **116**, 5567–5613.
14 B. Temelso, K. A. Archer and G. C. Shields, *J. Phys. Chem. A*, 2011, **115**, 12034–12046.
15 K. U. Lao and J. M. Herbert, *J. Phys. Chem. A*, 2015, **119**, 235–252.
16 J. Řezáč, Y. Huang, P. Hobza and G. J. Beran, *J. Chem. Theory Comput.*, 2015, **11**, 3065–3079.
17 M. Alkan, P. Xu and M. S. Gordon, *J. Phys. Chem. A*, 2019, **123**, 8406–8416.
18 S. A. Ochieng and K. Patkowski, *Phys. Chem. Chem. Phys.*, 2023, **25**, 28621–28637.
19 G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati and G. Terraneo, *Chem. Rev.*, 2016, **116**, 2478–2601.
20 G. Haberhauer and R. Gleiter, *Angew. Chem., Int. Ed.*, 2020, **59**, 21236–21243.
21 K. T. Mahmudov, A. V. Gurbanov, V. A. Aliyeva, M. F. C. Guedes da Silva, G. Resnati and A. J. Pombeiro, *Coord. Chem. Rev.*, 2022, **464**, 214556.
22 A. J. Stone, *J. Am. Chem. Soc.*, 2013, **135**, 7005–7009.
23 H. J. Davis and R. J. Phipps, *Chem. Sci.*, 2017, **8**, 864–877.
24 P. Politzer, J. S. Murray, T. Clark and G. Resnati, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32166–32178.

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

25 P. Metrangolo, H. Neukirch, T. Pilati and G. Resnati, *Acc. Chem. Res.*, 2005, **38**, 386–395.
26 M. Baldrighi, G. Cavallo, M. R. Chierotti, R. Gobetto, P. Metrangolo, T. Pilati, G. Resnati and G. Terraneo, *Mol. Pharmaceutics*, 2013, **10**, 1760–1772.
27 A. Bertolani, L. Pirrie, L. Stefan, N. Houbenov, J. S. Haataja, L. Catalano, G. Terraneo, G. Giancane, L. Valli, R. Milani, O. Ikkala, G. Resnati and P. Metrangolo, *Nat. Commun.*, 2015, **6**, 1–9.
28 K. Patkowski, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2020, **10**, e1452.
29 J. Řezáč and P. Hobza, *Chem. Rev.*, 2016, **116**, 5038–5071.
30 A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen and A. K. Wilson, *Chem. Phys. Lett.*, 1998, **286**, 243–252.
31 L. A. Burns, M. S. Marshall and C. D. Sherrill, *J. Chem. Phys.*, 2014, **141**, 234111.
32 M. Kodrycka and K. Patkowski, *J. Chem. Phys.*, 2019, **151**, 070901.
33 B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887–1930.
34 T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno and C. D. Sherrill, *J. Chem. Phys.*, 2014, **140**, 094106.
35 D. G. Smith, L. A. Burns, K. Patkowski and C. D. Sherrill, *J. Phys. Chem. Lett.*, 2016, **7**, 2197–2203.
36 H. L. Williams, K. Szalewicz, R. Moszynski and B. Jeziorski, *J. Chem. Phys.*, 1995, **103**, 4586–4599.
37 A. Heßelmann, G. Jansen and M. Schütz, *J. Chem. Phys.*, 2005, **122**, 14103.
38 A. J. Misquitta and K. Szalewicz, *J. Chem. Phys.*, 2005, **122**, 214109.
39 W. Xie, L. Song, D. G. Truhlar and J. Gao, *J. Chem. Phys.*, 2008, **128**, 234108.
40 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.
41 K. Carter-Fenk, K. U. Lao, K. Y. Liu and J. M. Herbert, *J. Phys. Chem. Lett.*, 2019, **10**, 2706–2714.
42 K. U. Lao and J. M. Herbert, *J. Chem. Phys.*, 2014, **140**, 119901.
43 J. M. Herbert, L. D. Jacobson, K. Un Lao and M. A. Rohrdanz, *Phys. Chem. Chem. Phys.*, 2012, **14**, 7679–7699.
44 O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 6.
45 M. A. Rohrdanz, K. M. Martins and J. M. Herbert, *J. Chem. Phys.*, 2009, **130**, 21.
46 M. Modrzejewski, Ł. Rajchel, G. Chałasiński and M. M. Szcześniak, *J. Phys. Chem. A*, 2013, **117**, 11580–11586.
47 R. Podeszwa and K. Szalewicz, *J. Chem. Phys.*, 2007, **126**, 194101.
48 A. Tkatchenko and O. A. Von Lilienfeld, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2008, **78**, 045116.
49 B. M. Axilrod and E. Teller, *J. Chem. Phys.*, 1943, **11**, 299–300.
50 Y. Muto, *Proc. Phys.-Math. Soc. Jpn.*, 1943, **17**, 629–631.
51 Y. Huang and G. J. Beran, *J. Chem. Phys.*, 2015, **143**, 044113.
52 Y. Xie, Z. L. Glick and C. D. Sherrill, *J. Chem. Phys.*, 2023, **158**, 94110.
53 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
54 E. Caldeweyher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2017, **147**, 34112.
55 O. A. Von Lilienfeld and A. Tkatchenko, *J. Chem. Phys.*, 2010, **132**, 234109.
56 K. T. Tang and J. P. Toennies, *J. Chem. Phys.*, 1984, **80**, 3726–3741.
57 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 242–253.
58 D. G. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Phys.*, 2020, **152**, 84108.
59 R. Bukowski, W. Cencek, P. Jankowski, M. Jeziorska, B. Jeziorski, S. A. Kucharski, V. F. Lotrich, A. J. Misquitta, R. Moszyński, K. Patkowski, R. Podeszwa, S. Rybak, K. Szalewicz, H. L. Williams, R. J. Wheatley, P. E. S. Wormer and P. S. Żuchowski, *SAPT2012: An Ab Initio Program for Many-Body Symmetry- Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies*, University of Delaware and University of Warsaw, 2012.
60 K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I. M. Høyvik, M. F. Iozzi, B. Jansík, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Sałek, C. C. Samson, A. S. de Merás, T. Saue, S. P. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. Wilson, M. Ziolkowski and H. Ågren, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 269–284.
61 Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock,

 This journal is © the Owner Societies 2025

Published on 14 octobre 2025. Downloaded on 2026-04-18 05:35:58.

P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C. M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. Distasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T. C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S. P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. Oneill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y. C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z. Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J. D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C. P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. Gill and M. Head-Gordon, *Mol. Phys.*, 2015, **113**, 184–215.

62 M. R. Kennedy, A. R. McDonald, A. E. DePrince, M. S. Marshall, R. Podeszwa and C. D. Sherrill, *J. Chem. Phys.*, 2014, **140**, 121104.

63 K. Carter-Fenk, K. U. Lao and J. M. Herbert, *Acc. Chem. Res.*, 2021, **54**, 3679–3690.

64 W. Jankiewicz, R. Podeszwa and H. A. Witek, *J. Chem. Theory Comput.*, 2018, **14**, 5079–5089.

65 M. Hapka, Ł. Rajchel, M. Modrzejewski, R. Schäffer, G. Chałasiński and M. M. Szcześniak, *J. Chem. Phys.*, 2017, **147**, 084106.