The influence of a solvent environment on direct non-covalent interactions between two molecules: A symmetry-adapted perturbation theory study of polarization tuning of π–π interactions by water

Dominic A. Sirianni, Xiao Zhu, Doree F. Sitkoff, et al.

ARTICLES YOU MAY BE INTERESTED IN

The influence of a solvent environment on direct non-covalent interactions between two molecules: A symmetry-adapted perturbation theory study of polarization tuning of π–π interactions by water

Dominic A. Sirianni,[1]   Xiao Zhu,[2] Doree F. Sitkoff,[2] Daniel L. Cheney,[2] and C. David Sherrill[1,a]

AFFILIATIONS

[1] Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA
[2] Molecular Structure and Design, Bristol Myers Squibb Company, P.O. Box 5400, Princeton, New Jersey 08543, USA

Note: This paper is part of the JCP Special Topic on Nature of the Chemical Bond.
[a] Author to whom correspondence should be addressed: sherrill@gatech.edu

ABSTRACT

High-level quantum chemical computations have provided significant insight into the fundamental physical nature of non-covalent interactions. These studies have focused primarily on gas-phase computations of small van der Waals dimers; however, these interactions frequently take place in complex chemical environments, such as proteins, solutions, or solids. To better understand how the chemical environment affects non-covalent interactions, we have undertaken a quantum chemical study of π–π interactions in an aqueous solution, as exemplified by T-shaped benzene dimers surrounded by 28 or 50 explicit water molecules. We report interaction energies (IEs) using second-order Møller–Plesset perturbation theory, and we apply the intramolecular and functional-group partitioning extensions of symmetry-adapted perturbation theory (ISAPT and F-SAPT, respectively) to analyze how the solvent molecules tune the π–π interactions of the solute. For complexes containing neutral monomers, even 50 explicit waters (constituting a first and partial second solvation shell) change total SAPT IEs between the two solute molecules by only tenths of a kcal mol^{-1}, while significant changes of up to 3 kcal mol^{-1} of the electrostatic component are seen for the cationic pyridinium–benzene dimer. This difference between charged and neutral solutes is attributed to large non-additive three-body interactions within solvated ion-containing complexes. Overall, except for charged solutes, our quantum computations indicate that nearby solvent molecules cause very little "tuning" of the *direct* solute–solute interactions. This indicates that differences in binding energies between the gas phase and solution phase are primarily *indirect* effects of the competition between solute–solute and solute–solvent interactions.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0087302

I. INTRODUCTION

Quantum mechanics is fundamental to understanding the nature of the chemical bond, whether from a molecular-orbital, valence-bond, or other point of view. It can be equally important in understanding the nature of weak interactions, whether hydrogen bonds, "halogen bonds," CH/π "bonds," and π–π interactions.[1] From ranking the relative stability of polymorphs of organic molecular crystals to probing differential host–guest binding in

pharmaceutical design, understanding such non-covalent interactions (NCIs) is crucial. For the past couple of decades, it has been possible to perform very high-level quantum mechanical studies of NCI in small, gas-phase model systems. The binding energy between two molecules $\mathcal{A}$ and $\mathcal{B}$ is the energy change for the reaction,

$$\mathcal{A}(gas) + \mathcal{B}(gas) \rightarrow \mathcal{A}\mathcal{B}(gas), \tag{1}$$
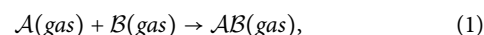
and is computable as

$$\Delta E_{\mathcal{A}\mathcal{B}}^{\text{bind}} = E_{\mathcal{A}\mathcal{B}} - E_{\mathcal{A}} - E_{\mathcal{B}}, \qquad (2)$$
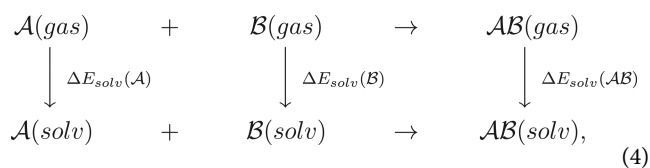
where $E_{\mathcal{A}}$ and $E_{\mathcal{B}}$ are the energies of the monomers and $E_{\mathcal{A}\mathcal{B}}$ is the energy of the dimer. For computational simplicity, studies often use the geometries of $\mathcal{A}$ and $\mathcal{B}$ as they appear in the dimer $E_{\mathcal{A}\mathcal{B}}$, rather than the equilibrium geometries they have in isolation. That energy difference is referred to as an "interaction energy (IE)," $\Delta E_{\mathcal{A}\mathcal{B}}^{\text{int}}$, and it ignores the energy penalties incurred by deforming the equilibrium geometries of $\mathcal{A}$ and $\mathcal{B}$ to the geometries they attain in the dimer. If these deformation energies are added to the interation energy, then one recovers the binding energy. For rigid molecules, the deformation energy is essentially zero, and the interaction energy and binding energy are equivalent. Again for reasons of computational simplicity, most quantum chemical studies use total electronic energies in computing the interaction energy or binding energy, although at increasing levels of sophistication one may also consider estimates of the enthalpies or Gibbs free energies.

Although there are now numerous high-level studies of gas-phase van der Waals dimers in the literature, it remains unclear to what extent these studies are relevant for understanding NCI in the context of more complex systems, e.g., protein–ligand binding or solvation. In particular, it is not understood how or to what degree individual NCI between chemical groups is modified by surrounding chemical environments. Our goal here is to explore aspects of this question, with the hope of better understanding how NCI should be thought about and computed in complex systems. In this study, we choose an aqueous solvent as the chemical environment.

Analogous to Eq. (1), we may consider the formation of a dimeric complex $\mathcal{A}\mathcal{B}$ now in solvent,

$$\mathcal{A}(solv) + \mathcal{B}(solv) \rightarrow \mathcal{A}\mathcal{B}(solv). \qquad (3)$$

The energy change for the formation of this complex in solvent may be related to the energy change for the formation of this complex in the gas phase using a thermodynamic cycle,

$$
\begin{array}{ccccc}
\mathcal{A}(gas) & + & \mathcal{B}(gas) & \rightarrow & \mathcal{A}\mathcal{B}(gas) \\
\downarrow{\scriptstyle \Delta E_{solv}(\mathcal{A})} & & \downarrow{\scriptstyle \Delta E_{solv}(\mathcal{B})} & & \downarrow{\scriptstyle \Delta E_{solv}(\mathcal{A}\mathcal{B})} \\
\mathcal{A}(solv) & + & \mathcal{B}(solv) & \rightarrow & \mathcal{A}\mathcal{B}(solv),
\end{array}
$$
$$ (4) $$

so that

$$\Delta E_{solv}^{\text{bind}}(\mathcal{A}\mathcal{B}) = \Delta E_{gas}^{\text{bind}}(\mathcal{A}\mathcal{B}) + \Delta E_{solv}(\mathcal{A}\mathcal{B})$$
$$- [\Delta E_{solv}(\mathcal{A}) + \Delta E_{solv}(\mathcal{B})]. \qquad (5)$$

Thus, the gas-phase binding energy is modified by the *difference* between the solvation energy of the dimer and the sum of the solvation energies of the monomers. If more energy is released by solvating the dimer than by solvating the separated monomers, then the dimer will exhibit greater binding in the solution. Typically, the reverse is true: more energy is released by solvating separated monomers, meaning that the dimer exhibits reduced

binding in solution. One might consider this an *indirect*, "effective" modification of the interaction due to the chemical environment.

Equation (5) is simple but some subtle effects lurk within it. One expects the solvation energy of the dimer, $\Delta E_{solv}(\mathcal{A}\mathcal{B})$, to arise primarily from solute–solvent interactions. However, this term contains other effects as well. The presence of the solvent may affect the *direct* electronic interaction between $\mathcal{A}$ and $\mathcal{B}$. The solvent molecules will polarize the electrons in the solute monomers, thus modifying the $\mathcal{A}$–$\mathcal{B}$ interaction. We might describe this modified interaction energy as

$$\Delta E_{\mathcal{A}\mathcal{B}}^{\text{direct-int}}(solv) = E_{\mathcal{A}\mathcal{B}}^{\text{elec}}(solv) - E_{\mathcal{A}}^{\text{elec}}(solv) - E_{\mathcal{B}}^{\text{elec}}(solv), \qquad (6)$$

where $E_{\mathcal{X}}^{\text{elec}}(solv)$ represents the total electronic energy of monomer or dimer $\mathcal{X}$ after it has been polarized by the solvent environment.

Furthermore, the solvent-induced polarization of the monomers depends on the particular locations of the solvent molecules, and thus, one needs to consider averages over solvent configurations. In addition, the presence of solvent will affect the favorability of different geometric orientations between $\mathcal{A}$ and $\mathcal{B}$, thus affecting geometric averages of the energetics of their interaction. Of course, the presence of solvent will also have a significant entropic effect, which will contribute to the Gibbs free energy difference $\Delta G_{solv}(\mathcal{A}\mathcal{B})$ $- [\Delta G_{solv}(\mathcal{A}) + \Delta G_{solv}(\mathcal{B})]$ and will affect the sampling of different $\mathcal{A}$–$\mathcal{B}$ and solute–solvent geometric configurations.

In this work, we will explore the solvent-induced changes in the direct electronic interaction between $\mathcal{A}$ and $\mathcal{B}$ [Eq. (6)], which to our knowledge has received little detailed scrutiny in large chemical systems. We will refer to this effect as "polarization tuning" of the direct solute–solute interaction. We do this here by investigating the manner in which the chemical environment modulates π–π interactions within T-shaped configurations of eight mono-functionalized aromatic molecules (ArX; Ar = benzene and pyridine; X = H, NH_2, NO_2, OCH_3, and CH_3) interacting with benzene (Bz), solvated by statistically significant configurations of up to one or two hydration shells (with 28 and 50 water molecules, respectively). For each solvent configuration, the tuning of the PhX–Bz interaction will be assessed by computing the direct ArX–Bz interaction within the solvent environment using theoretical tools particularly suited to this task: functional-group and intramolecular versions of symmetry-adapted perturbation theory (SAPT). Our study complements earlier works using molecular dynamics and Monte Carlo simulations on benzene dimer in water,[2,3] which accounted for configurational sampling, solute–solvent interactions, and entropic effects entering into $\Delta G_{solv}(\mathcal{A}\mathcal{B})$ $- [\Delta G_{solv}(\mathcal{A}) + \Delta G_{solv}(\mathcal{B})]$ using standard force fields. However, those studies were unable to examine how polarization of the solutes by the solvent would affect the direct electronic interaction between the solutes because the standard force fields used lacked polarization terms.

II. THEORETICAL AND COMPUTATIONAL APPROACH

A. Symmetry-adapted perturbation theory (SAPT)

Despite its utility for quantifying the strength of NCI, the supermolecular approach of Eq. (2) offers only a single scalar quantity

by which to do so. Thus, it provides no direct insight into the *reasons* for the different behavior of, e.g., hydrogen bonds vs π–π stacking, in terms of underlying physical forces, such as electrostatics, induction/polarization, London dispersion interactions, or exchange (steric) repulsion. The contributions of these forces to an interaction energy (IE) may be quantified either by the post-hoc partitioning of the total IE (an energy decomposition analysis, EDA) or by computing them directly via symmetry-adapted perturbation theory (SAPT).[4,5] SAPT has been particularly successful at providing a detailed description of the physics governing NCI by directly computing each of the components contributing to the IE of a bimolecular complex and has been applied to analyze and classify interaction motifs in a wide range of chemical systems.[6–8] Thanks to the formulation of the lowest-order truncation of SAPT (SAPT0) to leverage density-fitted two-electron integrals (DF-SAPT0), SAPT has become routinely applicable to systems as large as ~300 atoms,[9–11] or more if empirical dispersion estimates are used.[12] Furthermore, IEs computed with SAPT0 in the jun-cc-pVDZ basis set—where the diffuse space is truncated by neglecting diffuse functions on H atoms and diffuse d functions on heavy atoms—are within "chemical accuracy" (mean absolute errors of less than 1 kcal mol^{-1}) vs high-level coupled-cluster singles, doubles, and perturbative triples [CCSD(T)][13] benchmarks.[14]

The functional-group partition of SAPT (F-SAPT)[11] extends the functionality of SAPT by providing an additional partitioning of the SAPT0 interaction energy (and its components) into contributions from between pairs of functional groups on opposite monomers. However, it retains a limitation of its parent theory, two-body SAPT: all species must be accounted for as belonging to either monomer $\mathcal{A}$ or monomer $\mathcal{B}$.

The intramolecular formulation of SAPT (ISAPT)[15] provides for the computation of the interaction energy between functional groups of the same molecule, rather than the traditional two-monomer formulation of SAPT. ISAPT does this by first partitioning a single molecule $\mathcal{X}$ into interacting fragments $\mathcal{A}$ and $\mathcal{B}$, separated (but linked to one another) by a third fragment, $\mathcal{C}$. The zeroth-order wavefunctions for $\mathcal{A}$ and $\mathcal{B}$ are then prepared via a Hartree–Fock-in-Hartree–Fock embedding approach inspired by the procedure of Manby et al.,[16] in which the orbitals of $\mathcal{A}$ and $\mathcal{B}$ are electronically deformed by the presence of $\mathcal{C}$ before a standard SAPT0 computation is performed. The effect of the linker $\mathcal{C}$ is, therefore, *effectively* captured since the resulting ISAPT0 interaction energy and components are computed between the pre-polarized electron densities of fragments $\mathcal{A}$ and $\mathcal{B}$. Although the original test cases reported with ISAPT utilized a covalent linker $\mathcal{C}$, there is nothing in the theory that demands a covalent attachment, and, in this work, we identify group $\mathcal{C}$ with the solvent molecules.

B. Assessing polarization tuning of NCI by chemical environment via F-/ISAPT

In order to assess the extent to which NCI is modified by the electronic deformation of the interacting species by their chemical environment, we will leverage each of the F-SAPT and ISAPT (collectively, F-/ISAPT) approaches. To incorporate the chemical environment into our F-SAPT computations, we explicitly include solvent molecules in "monomer" $\mathcal{A}$ or $\mathcal{B}$; solute–solute interaction may then be extracted from the F-SAPT procedure by treating the solvent and solute molecules as different "functional groups" within

the monomer. When the solvent molecules are grouped with one of the solute molecules (here, benzene or substituted benzene), that molecule is "pre-polarized" by the solvent molecules when solving for the Hartree–Fock wavefunction of that overall solute + solvent grouping. This "pre-polarization" will affect a solute's interaction with the other solute molecule. Regrettably, this F-SAPT strategy allows the solvent molecules to pre-polarize only *one* of the solute molecules, and so the effect of solvent-induced polarization of the other molecule is neglected. Moreover, the computed solute–solute interaction energy will be different, depending on whether we chose to add the solvent molecules to "monomer" $\mathcal{A}$ or $\mathcal{B}$. Nevertheless, by performing the F-SAPT computations with both groupings, and comparing to gas-phase results, we can gain some insight into how polarization of a solute molecule by the solvent affects the solute–solute interaction.

From a theoretical perspective, ISAPT offers a more appealing approach to the problem. We collect all solvent molecules into the third "monomer," i.e., $\mathcal{C}$. In the ISAPT procedure, we solve for the Hartree–Fock wavefunction of the entire system (both solute molecules, $\mathcal{A}$ and $\mathcal{B}$, plus the collection of all solvent molecules, $\mathcal{C}$). We then localize the orbitals, freeze the density corresponding to monomer $\mathcal{C}$, and solve for the zeroth-order Hartree–Fock orbitals for $\mathcal{A}$ and $\mathcal{B}$ in the absence of each other, but in the continued presence of the Hartree–Fock potential of monomer $\mathcal{C}$. This approach allows the solvent molecules in $\mathcal{C}$ to pre-polarize *both* solute molecules before they interact with each other in the SAPT procedure.

We will differentiate these three schemes for handling the solvent by denoting them as "EnvX" (X = A, B, and C), where X refers to the F-/ISAPT "monomer" in which explicit solvent molecules are grouped. For clarity, we have illustrated these three schemes in Fig. 1 for the solvated aniline–benzene complex (HYD8-1), and we summarize these three approaches as follows:

(a) "EnvA": F-SAPT "monomer $\mathcal{A}$" consists of aniline plus all water molecules, and monomer $\mathcal{B}$ is the unsubstituted benzene molecule. The aniline–benzene interaction energy and its SAPT components are extracted via F-SAPT partitioning, with aniline defined as a separate "functional group" within "monomer $\mathcal{A}$." Aniline is "pre-polarized" by water molecules (but benzene is not).

(b) "EnvB": F-SAPT "monomer $\mathcal{B}$" consists of benzene plus all water molecules, and monomer $\mathcal{A}$ is aniline. The aniline–benzene interaction energy and its SAPT components are extracted via F-SAPT partitioning, with benzene defined as a separate "functional group" within "monomer $\mathcal{B}$." Benzene is "pre-polarized" by water molecules (but aniline is not).

(c) "EnvC": All water molecules grouped as ISAPT "monomer $\mathcal{C}$;" both aniline and benzene are "pre-polarized" by the Hartree–Fock embedding potential of monomer $\mathcal{C}$, which is obtained by localizing the Hartree–Fock orbitals of the entire system and retaining those orbitals localized on the solvent molecules. The $\mathcal{A}$–$\mathcal{B}$ interaction is computed using normal two-body SAPT theory using the aniline and benzene densities "pre-polarized" by the water molecules.

These three approaches (EnvA/B leveraging F-SAPT and EnvC leveraging ISAPT) offer complementary perspectives against which they may be mutually validated.
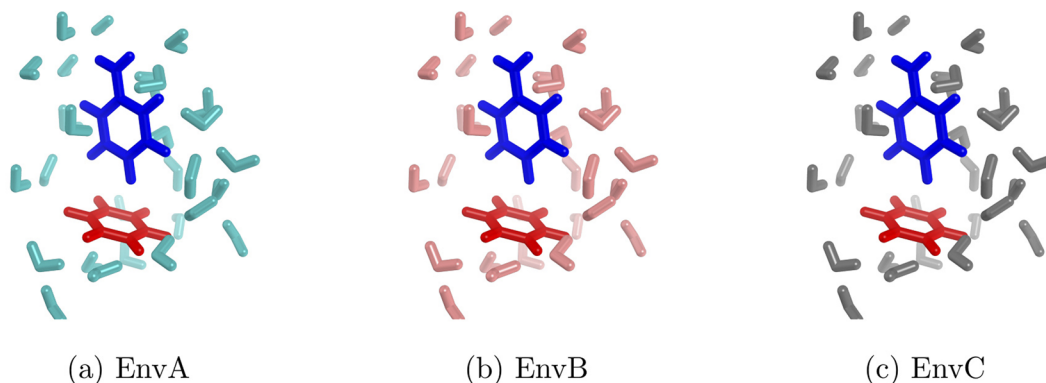
(a) EnvA (b) EnvB (c) EnvC

FIG. 1. Environment binning schemes employed in this work, illustrated for the HYD8-1 (solvated aniline–benzene) complex; here, aniline is designated as monomer A and benzene as monomer B. In "EnvA," the solvent molecules (cyan) are grouped with aniline (blue). In "EnvB," solvent molecules (light red) are grouped with benzene (red). In these cases, the direct electronic solute–solute interactions are computed through the F-SAPT approach. In "EnvC," the solvent molecules are treated as a separate "monomer C" in an ISAPT computation.

For additional insight into the basic physics of NCI embedded in a chemical environment, we have supplemented our F-/ISAPT computations with a traditional many-body cluster expansion approach, wherein the embedded NCI is described according to a "three body" decomposition. In this picture, "monomer $\mathcal{A}$" and "monomer $\mathcal{B}$" are the interacting solute molecules and group "$\mathcal{C}$" collects together the entire chemical environment (i.e., the water molecules). The overall interaction energy between monomers/groups $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ may be defined as

$$\Delta E_{\mathcal{ABC}}^{\text{IE}} = E_{\mathcal{ABC}}(\mathcal{ABC}) - E_{\mathcal{A}}(\mathcal{ABC}) - E_{\mathcal{B}}(\mathcal{ABC}) - E_{\mathcal{C}}(\mathcal{ABC}), \quad (7)$$

where the subscripts denote the identity of the species and the parenthetical ($\mathcal{ABC}$) denotes that each of the total energies in the expression have been computed in the trimer basis set according to the counterpoise correction scheme of Boys and Bernardi[17] to mitigate the basis set superposition error. The counterpoise correction entails computing all required energies using the union of all basis functions in the entire cluster (all three monomers/groups), even when some of the atoms are not required in the computation. Note again that we have grouped the entire environment together as a single group "$\mathcal{C}$," meaning that $\Delta E_{\mathcal{ABC}}^{\text{IE}}$ computed as above will be smaller in magnitude than if we computed the interaction relative to the limit in which all molecules (including those collected in the environment molecules) are infinitely separated; this leads to a simpler analysis because our primary concern here is the interaction between the two solute molecules, $\Delta E_{\mathcal{AB}}^{(2)}$, and how it is affected by the environment.

The three-body interaction energy, $\Delta E_{\mathcal{ABC}}^{\text{IE}}$, can also be computed according to the many-body expansion[18] as

$$\Delta E_{\mathcal{ABC}}^{\text{IE}}(\mathcal{ABC}) = \sum_{\mathcal{I} < \mathcal{J}} \Delta E_{\mathcal{IJ}}^{(2)}(\mathcal{ABC}) + \Delta E_{\mathcal{ABC}}^{(3)}(\mathcal{ABC}), \quad (8)$$

where each of the $\Delta E_{\mathcal{IJ}}^{(2)}(\mathcal{ABC})$ are the standard two-body interaction energies between monomers $\mathcal{I}$ and $\mathcal{J}$, and $\Delta E_{\mathcal{ABC}}^{(3)}(\mathcal{ABC})$ is the non-additive three-body contribution to the interaction energy. This non-additive contribution can be written as the difference between the overall interaction energy and the sum of the interactions between all pairs,

$$\Delta E_{\mathcal{ABC}}^{(3)} = \Delta E_{\mathcal{ABC}}^{\text{IE}} - \Delta E_{\mathcal{AB}}^{(2)}(\mathcal{ABC}) - \Delta E_{\mathcal{BC}}^{(2)}(\mathcal{ABC}) - \Delta E_{\mathcal{AC}}^{(2)}(\mathcal{ABC}). \quad (9)$$

By computing this quantity, we will further investigate the extent to which a mutual three-body interaction is relevant for each system. F-/ISAPT partially includes mutual polarization effects (as described above, the waters polarize one aromatic monomer in F-SAPT, and both aromatic monomers in ISAPT, and then the monomers interact), but it does not directly provide the non-additive three-body interaction energy, $\Delta E_{\mathcal{ABC}}^{(3)}$.

C. Preparation of geometries for functionalized complexes

Eight bimolecular complexes, each consisting of one benzene molecule and one benzene derivative, were prepared via functionalization of the tilted T-shaped pyridine–benzene complex from the S66 dataset of Řezáč *et al.*[19,20] Each structure was then optimized within enforced C_s symmetry using a development version of the PSI4 electronic structure package,[21,22] using the dispersion-corrected B3LYP density functional and aug-cc-pVDZ basis set.[23–25] Optimizations were performed with default convergence thresholds, as recommended previously,[26] and employed the recently modified[27] parameters for the "-D3" dispersion correction of Grimme,[28] together with Becke–Johnson damping.[29,30] For clarity, we will denote this combination of density functional, dispersion correction, and damping scheme here as "B3LYP-D3M(BJ)." This test set of functionalized benzene dimer complexes was constructed to provide both structural diversity and differentiation with respect to (i) local substituent dipole (e.g., toluene vs nitrobenzene), (ii) ability to form hydrogen bonds with the water solvent environment (e.g., phenol vs benzene), (iii) molecular polarizability (e.g., anisole vs benzene), and (iv) polarizing effect (e.g., pyridinium vs benzene).

In this way, we hope that conclusions drawn can be extended to the broader chemical space spanned by biologically relevant bimolecular complexes.

The relative orientations of the benzene and substituted benzenes were fixed at the gas-phase optimal geometries. Locations of solvating water molecules were determined by taking snapshots from molecular dynamics simulations, followed by a clustering analysis to help us select ten representative snapshots of water configurations around each of our eight test systems (a total of 80 snapshots). To eliminate any energetically spurious solvent configurations from the gathered snapshots, we performed partial geometry optimizations (with solute atoms fixed) to obtain relaxed or minimized structures. For comparison purposes, we also considered ten raw, unminimized snapshots each for benzene dimer, aniline–benzene, pyridine–benzene, and pyridinium–benzene (an additional 40 snapshots). For each of these 120 snapshots, we generated two chemical model systems, one including 28 waters, corresponding to an approximate first "solvation shell" (although some non-polar portions of the solute molecules remain exposed to vacuum), and one including 50 waters, corresponding to a second solvation shell for polar solute atoms and a complete first solvation shell for nonpolar solute atoms. This yields a total of 240 structures, which we label according to the format HYD8-IxJ-wN, where I = 1–8 identifies the complex (see numbering in Fig. 2), x = m, u indicates the solvent molecules were relaxed/minimized (m) or unrelaxed (u), J = 1–10 identifies the particular snapshot, and N = 28, 50 indicates the number of water molecules. Additional technical details on the construction of these 240 structures are provided in the supplementary material. No significant differences were found in the analysis of the minimized/relaxed snapshots vs the unminimized snapshots, so we focus on the former in our analysis (see the supplementary material).

D. Additional computational details

Since both F-SAPT and ISAPT are formulated at the SAPT0 level, all F-/ISAPT computations will employ the truncated jun-cc-pVDZ[25,31] basis set, which was previously recommended[14] for pairing with SAPT0. Throughout this work, we will adopt the convention that for all HYD8 complexes, the substituted aromatic monomer (ArX) will be denoted monomer $\mathcal{A}$, while the unsubstituted benzene (Bz) will be denoted as monomer $\mathcal{B}$. In the case of HYD8-3 (the solvated benzene dimer), the benzene with C–H bond pointing toward the π cloud of its partner will be monomer $\mathcal{A}$, for consistency with the other complexes. In addition to describing substituted benzene–benzene interactions in solution, we performed a conventional two-body F-SAPT analysis of the interactions between monomers in the gas phase. In this way, we may investigate directly the effect of solvation on the interactions of interest.

To construct the non-additive three-body energy correction via Eq. (9) for each of the 240 complexes in the HYD8 test set, we must compute seven individual computations (dimer energies for $\mathcal{AB}$, $\mathcal{BC}$, $\mathcal{AC}$; monomer energies for $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$; trimer energy $\mathcal{ABC}$, all in the trimer basis set) for a total of 1680 individual single-point computations. Considering that these complexes are comprised of up to 64 heavy atoms (for 178 atoms total) and each computation must be performed in the trimer basis set, choosing a level of theory (combination of method and basis set) that can be afforded is of critical concern. Since interaction energies are surprisingly sensitive to
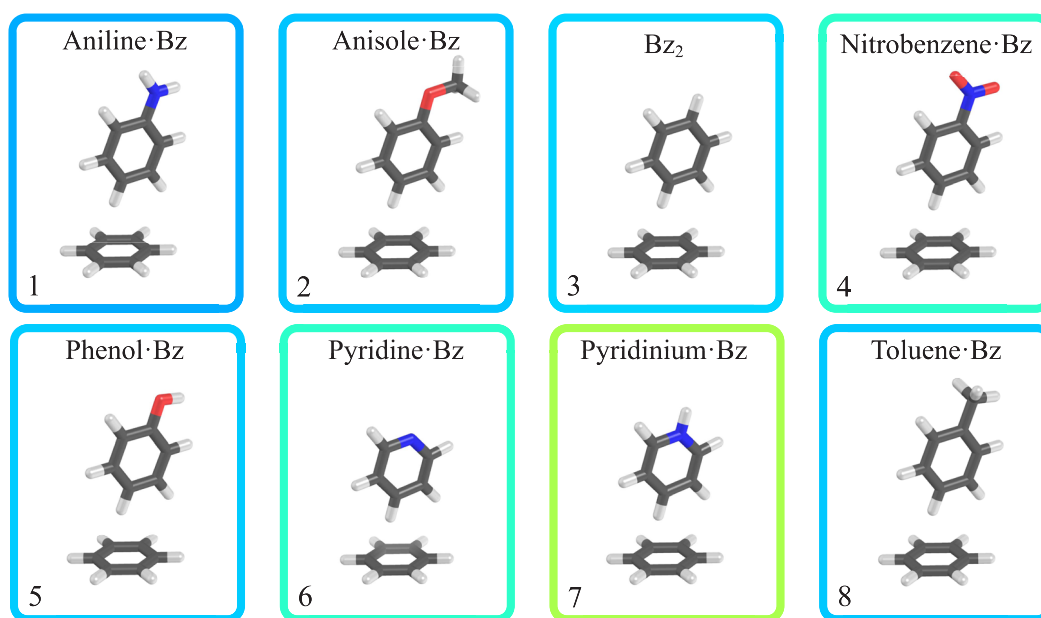

FIG. 2. Bimolecular complexes from which the HYD8 test set is constructed. Box coloring is based on SAPT0/jun-cc-pVDZ results computed in the gas phase and indicates the interaction type: blue for dispersion-dominated interactions and yellow-green for mixed electrostatics and dispersion contributions.

the choice of theoretical method,[14,27,32–36] we choose second-order Møller–Plesset perturbation theory (MP2) computations using the jun-cc-pVDZ basis set.[31,37] This level of theory represents a compromise between the computational accuracy and speed,[38–43] and it is also expected to yield the most similar interaction energies to the above discussed SAPT/jun-cc-pVDZ results that are the primary focus of this work. The largest of the interaction energy computations comprised 178 atoms (HYD8-2), with 1786 orbital basis functions and 8674 auxiliary basis functions for the density fitting procedure. Computations were performed on workstations with 6-core Intel i7 processors and 64 GB of RAM. All quantum computations were performed using the open-source PSI4 electronic structure package.[21,22]

III. RESULTS AND DISCUSSION

A. π–π interactions in the gas phase

Before considering the tuning of direct ArX–Bz interactions by the solvent environment, it is important to first understand the nature of the interactions of the HYD8 dimers in the gas phase; SAPT0 components are provided in Table I. In general, these eight ArX–Bz complexes are electrostatically attractive, but with an even larger dispersion term (which is sometimes up to twice as large), and a small stabilizing induction/polarization interaction. This is expected for T-shaped π–π interactions, and in an idealized T-shaped benzene dimer, SAPT2/jun-cc-pDVZ computes electrostatics, exchange, induction, and dispersion components to be -2.2, 4.9, -0.7, and -4.4 kcal mol^{-1}, respectively, yielding a total IE of -2.4 kcal mol^{-1}.[42] These results for the idealized T-shaped dimer are very similar to those computed for the tilted T-shaped benzene dimer in this study with SAPT0. Of course, differences in geometry and substituents influence these values somewhat, but they remain similar for each neutral HYD8 complex. For the cationic HYD8-7 (pyridinium–benzene), however, both the total SAPT0 IE and components are enhanced relative to the values for the neutral dimers, with the electrostatic component overshadowing dispersion as the dominant contributor to this increase in total attraction. Induction also becomes much larger for this charged complex and is now only slightly less attractive than dispersion.

Significant electrostatic stabilization is to be expected. In these T-shaped complexes, the *para*-hydrogen of monomer $\mathcal{A}$, bearing a partial positive charge, is the atom closest to the electron-rich π face of monomer $\mathcal{B}$. Moreover, among the HYD8 complexes involving an exocyclic substituent on the functionalized benzene (PhX), the magnitude of the electrostatic attraction correlates with the electron donating or electron withdrawing character of the substituent, as quantified by the Hammett σ_p parameter. The electrostatic attraction is the weakest for the dimer with the electron-donating NH$_2$ substituent and is the strongest for the electron-withdrawing NO$_2$ substituent. This is consistent with the electron-donating substituents reducing the electron-deficient character near the *para*-hydrogen, and thus reducing the strength of the electrostatic interaction, and vice versa for electron-withdrawing substituents. This is most pronounced for the nitrobenzene–benzene complex (HYD8-4), where the electrostatic interaction is a full kcal mol^{-1} more attractive than for the next dimer in the series, the benzene dimer (HYD8-3). Finally, both the pyridine–benzene and

TABLE I. Snapshot-averaged F-/ISAPT0/jun-cc-pVDZ solute–solute interaction energies and SAPT components (kcal mol^{-1}) for the HYD8 complexes, hydrated by 50 explicit water molecules (w50 subset), with each environment partitioning scheme. Each column corresponds to one of the dimers in Fig. 2. "X" row labels (A, B, and C) indicate that the explicit solvent molecules are contained within monomer "X" during the SAPT computation (see the text). Also included are the total IE and components from a conventional two-body SAPT computation in the gas phase, i.e., in the absence of explicit solvent molecules. Note that values are *direct electronic solute–solute interaction energies*, Eq. (6), not the solution-phase binding energies of Eq. (5).

	1	2	3	4	5	6	7	8
				Electrostatics				
A	-1.82	-2.04	-2.26	-3.36	-1.97	-4.50	-7.20	-2.16
B	-2.05	-2.24	-2.46	-3.14	-2.38	-4.42	-5.11	-2.38
C	-1.86	-2.04	-2.31	-3.09	-2.02	-4.50	-4.47	-2.24
Gas	-1.97	-2.22	-2.41	-3.38	-2.27	-4.39	-7.53	-2.25
				Exchange				
A	4.64	4.78	4.85	5.27	4.83	9.15	8.62	4.78
B	4.76	4.83	4.93	5.24	4.91	9.32	8.11	4.95
C	4.75	4.84	4.92	5.23	4.94	9.22	8.30	4.92
Gas	4.83	4.94	5.05	5.46	4.89	9.50	8.67	4.93
				Induction				
A	-0.40	-0.38	-0.47	-0.59	-0.44	-0.99	-2.11	-0.44
B	-0.36	-0.27	-0.33	-0.66	-0.23	-1.03	-3.34	-0.36
C	-0.67	-0.68	-0.70	-0.98	-0.69	-1.47	-3.36	-0.72
Gas	-0.60	-0.64	-0.68	-0.96	-0.64	-1.33	-4.14	-0.64
				Dispersion				
A	-4.92	-4.99	-5.00	-5.21	-4.98	-6.64	-6.34	-4.96
B	-4.99	-5.01	-5.03	-5.17	-5.00	-6.67	-6.06	-5.04
C	-4.91	-4.96	-4.97	-5.13	-4.98	-6.59	-6.16	-4.97
Gas	-4.73	-4.78	-4.79	-5.00	-4.73	-6.44	-5.97	-4.77
				Total IE				
A	-2.50	-2.64	-2.88	-3.89	-2.56	-2.98	-7.03	-2.79
B	-2.63	-2.69	-2.90	-3.72	-2.70	-2.80	-6.39	-2.83
C	-2.68	-2.83	-3.06	-3.97	-2.75	-3.34	-5.68	-3.00
Gas	-2.47	-2.70	-2.84	-3.89	-2.76	-2.66	-8.97	-2.74

pyridinium–benzene complexes exhibit larger magnitude total interaction energies and components relative to complexes involving PhX, consistent with previous observations that heteroatoms can enhance the strength of π–π interactions.[44]

B. Quantifying polarization tuning of ArX–Bz interactions by solvent using F-/ISAPT

As mentioned above, the focus of this study is to ascertain to what extent a chemical environment may modulate NCI by electronic deformation/polarization of the interacting species, which we call polarization tuning. We have examined this question through the use of F-/ISAPT methods. Because this polarization tuning

depends on the solvent configuration, we have obtained averages over ten diversified snapshots of each configuration.

1. Environment binning scheme

Visualized in Fig. 3 are total F-/ISAPT0 IEs and components for each environment grouping (EnvX; X = A, B, and C) of the hydrated benzene dimer (HYD8-3; top panel) and the hydrated pyridinium–benzene complex (HYD8-7; bottom panel). Most strikingly, the hydrated benzene dimer [Fig. 3(a)] exhibits very little variation between the average IE or components depending on the environment binning scheme; furthermore, F-/ISAPT IEs and components for each binning scheme are quite similar to those computed for the gas-phase benzene dimer with conventional SAPT0. This behavior is also exhibited by each of the other neutral HYD8 systems, with all differences in average total F-/ISAPT0 IE or components computed between EnvX binning schemes being less than 0.2 kcal mol^{-1} in nearly all cases except for induction, where
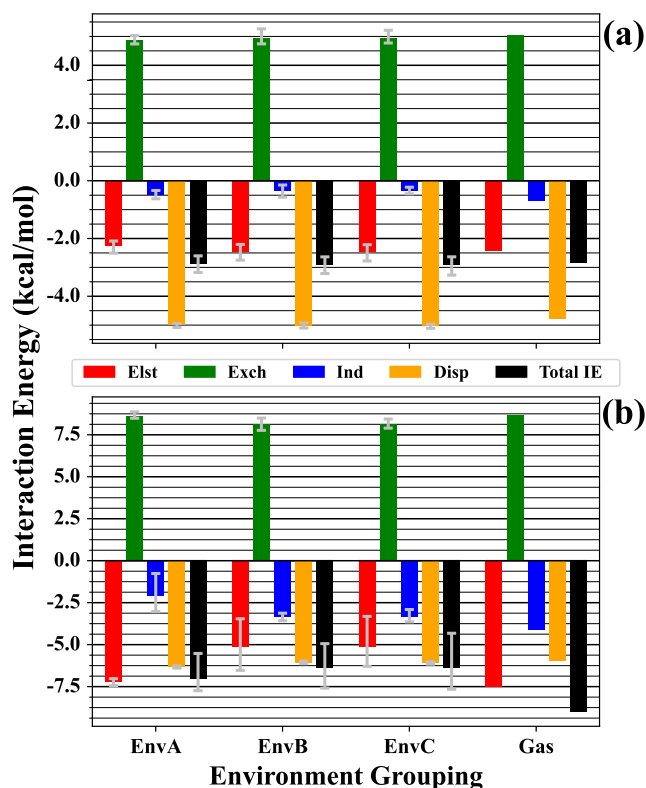

FIG. 3. Total direct electronic solute–solute interaction energies, Eq. (6), and SAPT components for (a) HYD8-3m-w50 (benzene dimer) and (b) HYD8-7m-w50 (pyridinium–benzene) complexes solvated by 50 explicit solvent molecules, computed at the F-/ISAPT0/jun-cc-pVDZ level of theory and averaged over all ten relaxed solvent configurations. "EnvX" labels indicate that explicit solvent molecules are contained within "monomer X" during the SAPT computation (see the text). Error bars encompassing the full range of values across all snapshots are also provided for SAPT terms and total IEs. On the right are results for conventional SAPT computations on gas-phase dimers in the absence of solvent molecules.

variations of 0.3–0.5 kcal mol^{-1} are observed. The solution-phase F-/ISAPT IEs and components also remain close to their conventional gas-phase SAPT0 counterparts across all the neutral dimers, with differences generally ~0.2 kcal mol^{-1} except for induction, where differences of 0.2–0.4 kcal mol^{-1} are observed. This similarity of F-/ISAPT0 IEs and components between environment binning schemes for neutral HYD8 complex indicates that the presence of explicit solvent molecules does not significantly tune the direct ArX–Bz π–π interactions in neutral systems.

This result demonstrates that any understanding of the fundamental physics of direct interaction between two substituted neutral benzenes is not invalidated if the benzenes are placed in water. Indeed, even on a quantitative basis, the electrostatics, exchange-repulsion, induction/polarization, and London dispersion interactions between the two benzenes generally change very little on average due to the presence of surrounding solvent molecules (more significant changes are seen in some of the individual snapshots, as discussed in Subsection III B 2). Thus, it does not appear that the solvent reduces (or "screens") the *direct* interaction between the aromatic molecules, what we called polarization tuning in the Introduction. Of course, as also mentioned in the Introduction, it remains the case that solute–solvent interactions may make large contributions to $\Delta E_{solv}(\mathcal{AB}) - [\Delta E_{solv}(\mathcal{A}) + \Delta E_{solv}(\mathcal{B})]$ from Eq. (5), and thus provide an *indirect* or *effective* "screening" of the interaction.

The cationic HYD8-7 system [hydrated pyridinium–benzene complex, visualized in Fig. 3(b)], on the other hand, exhibits notable variations between binning schemes for total IEs and components. In addition, the overall strength of the interaction is significantly reduced in solution vs the gas phase (by ~2–3 kcal mol^{-1} compared to 9 kcal mol^{-1} in the gas phase), indicating significant polarization tuning of the π–π interaction by the solvent. For this complex, large differences from the gas phase are found for the electrostatic and induction components. The electrostatic contribution is decreased in magnitude for all environment binning schemes, either by a modest amount (0.3 kcal mol^{-1}, EnvA) or by a large amount (2.4 and 3.1 kcal mol^{-1}, for EnvB and EnvC, respectively). The induction/polarization interaction between the solute molecules is reduced by 2.0, 2.8, or 0.8 kcal mol^{-1} in EnvA, EnvB, or EnvC, respectively. Exchange-repulsion is reduced somewhat, and London dispersion interactions become somewhat more attractive, depending on the binning scheme.

2. Ranges due to solvent configuration

Although the *average* interaction energies or their components are quite similar between envA, envB, envC, and the gas phase for the neutral dimers, there can be some significant variations between individual snapshots of the solvent configurations. Thus, even for the neutral dimers, there can be significant polarization tuning of the solute–solute interactions by the solvent, but this effect tends to be washed out when taking averages over solvent configurations.

Tables S4–S15 of the supplementary material present the detailed breakdowns of the SAPT0 components for every dimer, environment binning, and solvent snapshot. The snapshot averages, and their ranges, are visualized in Figs. S1–S12 of the supplementary material. For each F-/ISAPT component, differences between solvent configurations typically range from ~0.2 to 0.5 kcal mol^{-1} for exchange or induction, 0.5 to 1.0 kcal mol^{-1} for electrostatics,

0.1 kcal mol^{-1} for dispersion, and 0.5 to 1.5 kcal mol^{-1} for total interaction energies. Larger ranges are seen in some cases, including the pyridinium–benzene system in panel (b) of Fig. 3, which demonstrates spreads of 3 kcal mol^{-1} among the electrostatic energies of different solvent configurations for EnvB/C and spreads of 2–3 kcal mol^{-1} in the total interaction energies. The range in electrostatics is also larger than normal for some environmental binnings of HYD8-4 (nitrobenzene–benzene) and the range in induction values across snapshots increases to 1 kcal mol^{-1} for some environment binnings for HYD8-4 and HYD8-6 (pyridine–benzene) and can grow to more than 2 kcal mol^{-1} for HYD8-7 (pyridinium–benzene).

The increased susceptibility of solute–solute interactions to solvent configuration for nitrobenzene–benzene, pyridine–benzene, and pyridinium–benzene may be related to the fact that these dimers contain the three substituted benzenes with the largest dipole moments (5.52, 2.27, and 1.90 D, respectively, at the same Hartree–Fock/jun-cc-pVDZ level of theory the SAPT computations are based upon). The dipole moments for the other substituted benzenes are smaller (aniline: 1.60, anisole: 1.44, benzene: 0.00, phenol: 1.47, and toluene: 0.41 D). Because solvation can enhance a molecule's dipole moment by ~30%–40% relative to its gas phase value[45] (a polarization effect), the variation of ArX–Bz interactions between different solvent configurations may be related to magnifications in the ArX dipole, which one might expect to depend sensitively on the solvent configuration. One might also expect this solvent-induced dipole enhancement to be larger in monomers with larger initial dipole moments, thus explaining why the other complexes exhibit reduced variations among the solvent configurations. This could account for larger variations for these molecules seen in the envA or envC binning schemes, where the explicit water molecules can polarize the substituted benzene ("monomer $\mathcal{A}$"). On the other hand, it cannot account for the larger variations in electrostatics for envB for these three molecules because that binning scheme separates the waters from the substituted benzene when obtaining the electron distribution around the substituted benzene. In these cases, one may expect the charged molecule (pyridinium) and the two neutral molecules with the largest dipole moments (nitrobenzene and pyridine) will have electrostatic interactions that are more sensitive to changes in the π electron distributions of the unsubstituted benzene, which is polarized differently by different water snapshots in the envB binning scheme. These three dimers (involving nitrobenzene, pyridine, and pyridinium) have significantly larger gas-phase electrostatic interaction energies (−3.4 to −7.5 kcal mol^{-1}) than the others (strongest interaction −2.4 kcal mol^{-1}).

C. Many-body analysis of solvated interactions

To better understand why the environment has little effect on the π–π interactions between neutral monomers (at least when averaged over solvent snapshots), while it has a larger effect for systems involving a cationic monomer, we have computed the non-additive three-body component of the total trimer energies according to Eq. (9) at the Hartree–Fock/jun-cc-pVDZ and MP2/jun-cc-pVDZ levels of theory, as these combinations of methods and basis set provide the most direct comparison to the F-/ISAPT0 results. A large $\Delta E_{\mathcal{ABC}}^{(3)}$ indicates that the mutual interactions between the benzene

(Bz), the substituted benzene (PhX), and the waters are not accurately modeled as a simple sum of isolated Bz-PhX, Bz–water, and PhX–water interactions. Conversely, if $\Delta E_{\mathcal{ABC}}^{(3)}$ is essentially zero, it means that the interactions between these groups are all pairwise-additive, and thus, we would expect no tuning of the direct electronic solute–solute interaction by the presence of the solvent.

Generally, the largest source of non-additive interactions is mutual polarization of electron densities, and we expect this is the case in the systems studied here. This effect can be captured well even with simple Hartree–Fock theory. ISAPT also incorporates such mutual polarization effects. In the ISAPT procedure, one obtains Hartree–Fock orbitals for the entire system (here, both solutes and all waters); this fully includes all mutual polarization effects. Then, the orbitals of "monomer $\mathcal{C}$" (here, the waters) are frozen, and the local orbitals of monomers $\mathcal{A}$ and $\mathcal{B}$ are obtained in the embedding field of the density of $\mathcal{C}$ but in the absence of each other (so that a normal SAPT procedure may be applied after that). Because the presence of $\mathcal{C}$ is felt throughout by both $\mathcal{A}$ and $\mathcal{B}$, and because $\mathcal{C}$ is obtained from a supermolecular Hartree–Fock computation, many-body polarization is effectively folded into the procedure. Our F-SAPT computations, on the other hand, which group the water molecules with monomers $\mathcal{A}$ or $\mathcal{B}$, only partially account for mutual polarization effects because only one of the benzene solute molecules can polarize in response to the water. In this sense, the ISAPT model is more appropriate for understanding these systems than the F-SAPT models.

Hartree–Fock will also model non-additive three-body exchange repulsion, which should be small in these systems because they would require very close three-body contacts involving a water molecule and the two solute molecules simultaneously. Unfortunately, neither Hartree–Fock nor MP2 (nor our F-/ISAPT computations) will capture three-body dispersion interactions, which would be modeled by the correlated motion of three electrons (one from each monomer). However, due to the limited polarizability of a water molecule and the reasonable separations between molecules, we expect this effect to be modest. To confirm this expectation, we performed a series of computations on trimers consisting of the two benzenes plus a single water; we did this for all 28 waters in structure HYD8-3m3-w28. For each trimer, we estimated the three-body dispersion as the difference between the counterpoise-corrected three-body non-additive contribution to the interaction energy, $\Delta E_{\mathcal{ABC}}^{(3)}$ [Eq. (9)], computed using MP2.5,[46] and the corresponding quantity computed using MP2. MP2 will provide a reasonable description of three-body (non-additive) induction and exchange but lacks any three-body (non-additive) dispersion. MP2.5 improves upon MP2 by adding half of the third-order correlation in MP3 and has been found to capture three-body dispersion fairly accurately.[47] Thus, $\Delta E_{\mathcal{ABC}}^{(3)}$ (MP2.5) − $\Delta E_{\mathcal{ABC}}^{(3)}$ (MP2) should provide a good estimate of three-body dispersion. The three-body dispersion estimates had magnitudes under 0.02 kcal mol^{-1} for the 28 trimers and summed to 0.08 kcal mol^{-1} (see Table S-32 of the supplementary material), demonstrating the small size of three-body dispersion in these systems.

Presented in Fig. 4 are box-and-whisker plots visualizing the distribution of non-additive three-body energy contribution, $\Delta E_{\mathcal{ABC}}^{(3)}$, computed with Hartree–Fock over different solvent configurations of each HYD8 complex. Immediately, it is apparent that complex
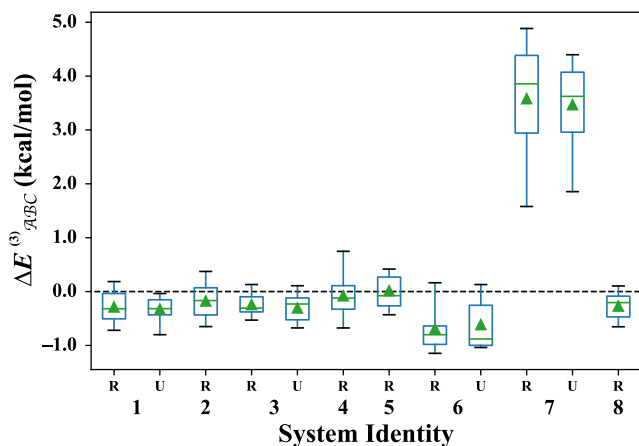
FIG. 4. Box-and-whisker plots representing the non-additive three-body correction ($\Delta E_{ABC}^{(3)}$; kcal mol^{-1}) to the total "trimer" energy, for both relaxed (R) and unrelaxed (U) solvent configurations of each HYD8 complex (with 50 waters), computed at the HF/jun-cc-pVDZ level of theory. Boxes encompass the first (Q1) through third (Q3) quartiles of $\Delta E_{ABC}^{(3)}$ across all solvent snapshots, with values corresponding to the median (Q2) and mean $\Delta E_{ABC}^{(3)}$ indicated as a solid green bar and green triangle, respectively. Whiskers encompass the full range of $\Delta E_{ABC}^{(3)}$ values.

FIG. 5. Box-and-whisker plots representing the total "trimer" interaction energy (ΔE_{ABC}^{IE}; kcal mol^{-1}) for both relaxed (R) and unrelaxed (U) solvent configurations of benzene dimer (HYD8-3) and pyridinium–benzene (HYD8-7) complexes hydrated by 50 explicit water molecules, computed at the HF/jun-cc-pVDZ (orange boxes) and MP2/jun-cc-pVDZ (blue boxes) levels of theory. Boxes encompass the first (Q1) through third (Q3) quartiles of ΔE_{ABC}^{IE}, with values corresponding to the median (Q2) and mean ΔE_{ABC}^{IE} indicated as a solid green bar and green triangle, respectively. In addition, whiskers encompass the full range of ΔE_{ABC}^{IE} values for all solvent configurations.

HYD8-7, the cationic pyridinium–benzene complex, exhibits a much larger value for $\Delta E_{ABC}^{(3)}$ (typically ~3–4 kcal mol^{-1}) than the other dimers (typically ~−1 to 0 kcal mol^{-1}). The much larger size of $\Delta E_{ABC}^{(3)}$ for pyridinium–benzene is consistent with the much larger difference observed between solvated and gas-phase solute–solute SAPT0 energies, and the larger variations between binning schemes (EnvA/B/C), for this dimer compared to the other dimers considered. $\Delta E_{ABC}^{(3)}$ also exhibits a much wider range (~ 3 kcal mol^{-1}) for this complex than for the other complexes, which is consistent with the wider variation between snapshots observed for this complex in the F-/ISAPT results.

Above, we noted that after the pyridinium–benzene complex, the next-biggest variations in SAPT0 interactions with respect to snapshot configuration occurred for HYD8-4 (nitrobenzene–benzene) and HYD8-6 (pyridine–benzene). Correspondingly, in Fig. 4, we see that these systems have the next-largest variations in $\Delta E_{ABC}^{(3)}$.

While the preceding analysis has been performed for three-body interactions at the HF/jun-cc-pVDZ level of theory, it is worth noting that the two-body ArX–Bz interaction energy at this level is in many cases repulsive (see, e.g., Table S-16–S-23 in the supplementary material). This is to be expected because London dispersion interactions are very important to the binding of these systems, but they are absent at the Hartree–Fock level because they are inherently electron correlation effects. Therefore, to provide a more complete picture of the interactions in these systems, we have also performed a similar analysis at the MP2/jun-cc-pVDZ level of theory, which is the closest supermolecular wavefunction method to SAPT0/jun-cc-pVDZ. Provided in Fig. 5 is a comparison of the full trimer IE computed with HF/jun-cc-pVDZ and
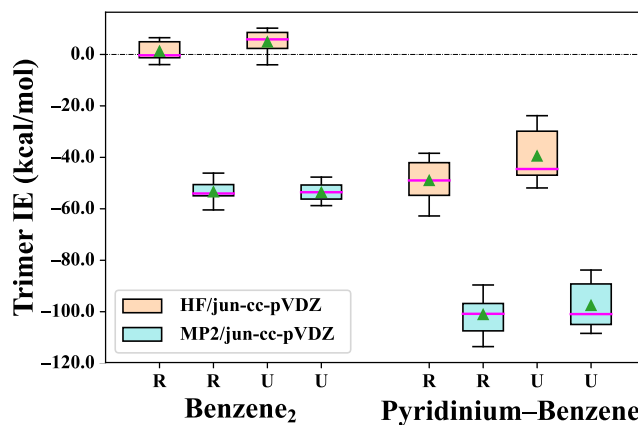
MP2/jun-cc-pVDZ for the hydrated benzene dimer (HYD8-3) and hydrated pyridinium–benzene complex (HYD8-7) (all components of the three-body MBE are given in Tables S-23–S-30 in the supplementary material). When using MP2/jun-cc-pVDZ, all two-body ArX–Bz IEs now become attractive. All trimer IEs are also now attractive, and in these large model systems with 50 waters, the shift between Hartree–Fock and MP2 trimer interaction energies is ~50–60 kcal mol^{-1}. Nevertheless, both the ranges and the values of the three-body interaction ΔE_{ABC}^{IE} are similar between Hartree–Fock and MP2.

D. Effect of multiple hydration shells

We expect the major effects of polarization tuning of π–π interactions to come from the closest water molecules. To confirm this expectation, and to probe whether a partial second solvation shell is sufficient to capture most of the effects of polarization tuning by bulk solvent, we have also performed F-/ISAPT0 computations using an approximate first hydration shell (with 28 explicit water molecules), and we compare those results to the computations previously discussed with including a partial second solvation shell (with 50 explicit water molecules).

We depict the difference in F-/ISAPT0 IEs and components upon the addition of the first solvation shell to the gas-phase complex (ΔE_{int}^{1-G}) and upon addition of the second solvation shell (ΔE_{int}^{2-1}) in Fig. 6 for each environment binning of the hydrated benzene dimer [HYD8-3; top panel (a)] and hydrated pyridinium–benzene complex [HYD8-7; bottom panel (b)]. Results for the other dimers are presented in the supplementary material.

For neutral systems, where the non-additive three-body interaction is small [e.g., benzene dimer, Fig. 6(a)], we find only small

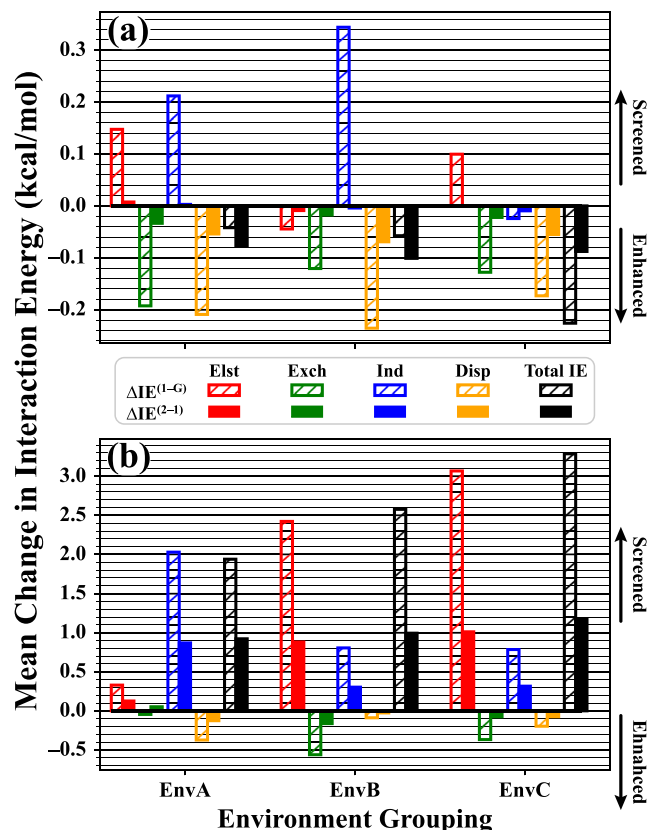**FIG. 6.** Mean change in total direct electronic solute–solute interaction energies, Eq. (6), and SAPT components upon first-shell solvation (ΔIE^{1-G}; striped bars) and second-shell solvation (ΔIE^{2-1}; solid bars) for the (a) HYD8-3mX (benzene dimer) and (b) HYD8-7mX (pyridinium–benzene) complexes, averaged over values computed at the F-/ISAPT0/jun-cc-pVDZ level of theory for all ten relaxed solvent configurations. "EnvX" labels indicate that explicit solvent molecules are contained within monomer "X" during the SAPT computation (see the text). See Sec. II C for additional details regarding our nomenclature and details of the F-/ISAPT computations.

components seem largely converged by the time an approximate single solvation shell has been added.

For the cationic pyridinium–benzene complex, where the non-additive three-body interaction is significant (see above), the effects of adding both the first and second solvation shells are more notable, as shown in Fig. 6(b). As for the neutral complexes, dispersion interactions are enhanced slightly upon the addition of solvent, and essentially all of the effect is captured by the addition of the first solvation shell. Similarly, consistent with the benzene dimer and many of the neutral complexes, exchange-repulsion between the solute molecules is slightly less repulsive with the addition of the first solvation shell, at least for EnvB and EnvC, and the addition of the second shell produces a little additional effect. Significantly larger changes are seen for the induction, electrostatics, and total interaction energies, with most ΔE_{int}^{1-G} values in the range of 0.8–3.3 kcal mol^{-1} and changes to total interaction energies of 1.9–3.3 kcal mol^{-1}. All of these shifts are in the direction of a weaker interaction between the pyridinium and the benzene (the interactions are directly "screened" by the solvent-induced solute polarization). Upon the addition of the second solvation shell, similar behavior is observed for each of these three energies (electrostatics, induction, and total IE); however, the difference is only about half or a third as large as was observed for the addition of the first solvation shell. This suggests that the computed solvent-induced polarization tuning of the pyridinium–benzene interaction is not fully converged with our 50-water model systems (including a first and partial second solvation shell). Nevertheless, we expect that these 50-water computations have captured the majority of the relevant physics.

IV. SUMMARY AND CONCLUSIONS

The extent to which the chemical environment "tunes" non-covalent interactions, as well as the best manner in which to account for this effect, are open questions in the computational molecular sciences. Here, we considered eight functionalized, T-shaped arene–benzene complexes in the gas phase and hydrated by a statistically diverse set of solvent configurations consisting of either an approximate single solvent shell of 28 explicit water molecules or a complete single shell and partial double shell consisting of 50 explicit water molecules. We have used a traditional many-body cluster expansion, at the Hartree–Fock/jun-cc-pVDZ and MP2/jun-cc-pVDZ levels of theory, to quantify the two-body and non-additive three-body effects in these systems, where the "bodies" are a functionalized benzene ($\mathcal{A}$), the benzene molecule ($\mathcal{B}$), and the collection of all the explicit water molecules ($\mathcal{C}$). While the size of the non-additive three-body contribution reflects mutual interactions involving all three of these groups and is therefore relevant in understanding the effect of the chemical environment, it is not formulated to directly answer the original question of "how does a chemical environment tune a non-covalent interaction?"

On the other hand, functional-group partitioning of symmetry-adapted perturbation theory (F-SAPT) and its "intramolecular" formulation (ISAPT) are able to directly answer the question as posed, if we seek to understand the tuning of the direct electronic solute–solute interaction arising from solvent-induced polarization of the solute molecules ("polarization tuning"). To the extent to which they can be compared, findings from the three-body

differences in components upon addition of the first solvation shell, ΔE_{int}^{1-G} (generally less than 0.2 kcal mol^{-1}, or less than 10% of the gas-phase IE). For this system, polarization by solvent enhances dispersion slightly, and reduces unfavorable exchange-repulsion slightly, on average compared to the gas phase. Changes in electrostatics and induction depend on the solvent binning scheme, but the overall average interaction energies are consistently more attractive in the solvent than the gas phase. Further differences upon addition of the second solvation shell, ΔE_{int}^{2-1}, are much smaller (only a few hundredths of 1 kcal mol^{-1}, up to 0.1 kcal mol^{-1} for the shift in the overall interaction energy). Except in cases where the effect is essentially zero, the additional shifts due to the addition of a second solvation shell, ΔE_{int}^{2-1}, tend to have the same sign as the shifts due to the addition of the first solvation shell, ΔE_{int}^{1-G}. As illustrated in Figs. S19–S27 of the supplementary material, these trends hold across the other neutral dimers considered. Hence, for the neutral systems, solvent-induced shifts in interaction energy and its

Hartree–Fock and MP2 computations are consistent with the findings from the F-/ISAPT computations.

Our test molecules included seven neutral dimer complexes and one charged complex (pyridinium–benzene). For the neutral complexes, we find that the non-additive three-body interaction energies can be substantial if one looks across all the snapshot solvent configurations considered. The SAPT0/jun-cc-pVDZ interaction energies for the neutral complexes range from −2.5 to −3.9 kcal mol^{-1}, whereas the Hartree–Fock/jun-cc-pVDZ non-additive three-body interaction energies (sufficient to capture three-body induction/polarization effects) have a range of ~ ±1 kcal mol^{-1} (MP2 results are not much different). Thus, solvent molecules certainly can "tune" direct solute–solute interactions by up to a significant fraction of the total interaction energy. Correspondingly, across the neutral dimers, we often see significant variations in the electrostatics, induction, and exchange-repulsion components of the energy as the solvent configuration is varied. London dispersion interactions appear to be mostly immune to these variations and are hardly affected by the environment. On the other hand, for the solute–solute interaction, the *average* values of the energy components, or the total interaction energy, are all very similar across any of our "environment binning" schemes (F-SAPT vs ISAPT), and are also very similar to the gas-phase values. This suggests that dynamical averaging of solvent configurations tends to wash out any solvent-induced tuning of the solute–solute interactions.

The situation is different for the charged pyridinium–benzene complex. Non-additive three-body interaction energies are substantially larger for this complex (typically ~3–4 kcal mol^{-1} for Hartree–Fock/jun-cc-pVDZ), and their variation is much larger across solvent snapshots. Correspondingly, we see much larger differences between the gas-phase and solvent-including SAPT computations for the solute–solute interactions. Again, London dispersion forces are hardly affected by the solvent environment, but there can be large changes in the electrostatic and induction terms. There are also larger variations in the SAPT components across solvent snapshots. Overall, polarization of the solute molecules by the environment tends to "screen" the direct solute–solute interactions and reduce their attraction to each other, by ~1.5–2.5 kcal mol^{-1} on average out of a gas-phase value of −9.0 kcal mol^{-1} for pyridinium–benzene.

For the neutral dimers, a single hydration shell of 28 explicit water molecules within ~3–4 Å is sufficient to converge the solvent-induced changes to the solute–solute interaction energy or its components. For the charged pyridinium–benzene complex, even 50 explicit water molecules within 7 Å of the solute complex (constituting a full single solvent shell and a partial second solvent shell around polar atoms) may not be sufficient to ensure convergence of the solute–solute interactions toward the continuum limit.

Although significant changes are observed in the solute–solute total interaction energy and its components for the charged pyridinium–benzene complex, and for certain snapshots of the neutral systems, in no case did the polarization by solvent lead to a change in the qualitative nature of the solute–solvent interaction (its rough strength and breakdown into electrostatic, exchange-repulsion, induction, and London dispersion components). Moreover, variations in solute–solute interactions due to solvent configuration were generally modest for the neutral dimers, and on average, the SAPT component and total interaction energies were only slightly changed from their gas-phase values. This suggests that insights into the fundamental physics of non-covalent interactions, gained through gas-phase theoretical studies by many research groups on many prototype chemical systems, should remain valid when considering similar direct interactions in condensed phases, certainly at a qualitative level, and also probably at a quantitative level, at least for neutral–neutral interactions. On the other hand, the significant solvent-induced polarization tuning of solute–solute interactions in the charged pyridinium–benzene complex reinforces the idea that polarizable force-fields models, which can capture this kind of polarization tuning, will be more appropriate for studies of charged solutes than standard non-polarizable force-fields.

SUPPLEMENTARY MATERIAL

See the supplementary material for complete tables of all energetic quantities and figures providing additional analysis of results, including snapshot-averaged interaction energy components for all eight dimers considered, variations in components among snapshots, the effect of the first and second solvation shells upon average energy components for each dimer considered, comparison of relaxed vs unrelaxed solvent geometries, and detailed two-body and three-body interaction energies for Hartree–Fock and MP2 for all snapshots considered. Also provided are Cartesian coordinates for geometries considered in this work.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

REFERENCES

[1] E. G. Hohenstein and C. D. Sherrill, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 304 (2012).

[2] W. L. Jorgensen and D. L. Severance, J. Am. Chem. Soc. **112**, 4768 (1990).

[3] C. Chipot, R. Jaffe, B. Maigret, D. A. Pearlman, and P. A. Kollman, J. Am. Chem. Soc. **118**, 11217 (1996).

[4] B. Jeziorski, R. Moszynski, and K. Szalewicz, Chem. Rev. **94**, 1887 (1994).

[5] K. Szalewicz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 254 (2012).

[6] K. Patkowski, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **10**, e1452 (2019).

[7] E. Francisco and A. Martín Pendás, in *Non-Covalent Interactions in Quantum Chemistry and Physics*, edited by A. O. de la Roza and G. A. DiLabio (Elsevier, 2017), pp. 27–64.

[8] C. D. Sherrill, Acc. Chem. Res. **46**, 1020 (2013).

[9] E. G. Hohenstein and C. D. Sherrill, J. Chem. Phys. **132**, 184111 (2010).

[10] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, J. M. Turney, and H. F. Schaefer, J. Chem. Phys. **135**, 174107 (2011).

[11] R. M. Parrish, T. M. Parker, and C. D. Sherrill, J. Chem. Theory Comput. **10**, 4417 (2014).

[12] J. B. Schriber, D. A. Sirianni, D. G. A. Smith, L. A. Burns, D. Sitkoff, D. L. Cheney, and C. D. Sherrill, J. Chem. Phys. **154**, 234107 (2021).

[13] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[14] T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherrill, J. Chem. Phys. **140**, 094106 (2014).

[15] R. M. Parrish, J. F. Gonthier, C. Corminbœuf, and C. D. Sherrill, J. Chem. Phys. **143**, 051103 (2015).

[16] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, J. Chem. Theory Comput. **8**, 2564 (2012).

[17] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[18] D. Hankins, J. W. Moskowitz, and F. H. Stillinger, J. Chem. Phys. **53**, 4544 (1970).

[19] J. Řezáč, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **10**, 1359 (2014).

[20] J. Řezáč, K. E. Riley, and P. Hobza, J. Chem. Theory Comput. **7**, 2427 (2011).

[21] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 3185 (2017).

[22] D. G. A. 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. **152**, 184108 (2020).

[23] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[24] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).

[25] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[26] D. A. Sirianni, A. Alenaizan, D. L. Cheney, and C. D. Sherrill, J. Chem. Theory Comput. **14**, 3004 (2018).

[27] D. G. A. Smith, L. A. Burns, K. Patkowski, and C. D. Sherrill, J. Phys. Chem. Lett. **7**, 2197 (2016).

[28] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[29] S. Grimme, S. Ehrlich, and L. Goerigk, J. Comput. Chem. **32**, 1456 (2011).

[30] E. R. Johnson and A. D. Becke, J. Chem. Phys. **123**, 024101 (2005).

[31] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, and D. G. Truhlar, J. Chem. Theory Comput. **7**, 3027 (2011).

[32] D. A. Sirianni, L. A. Burns, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 86 (2017).

[33] L. A. Burns, M. S. Marshall, and C. D. Sherrill, J. Chem. Theory Comput. **10**, 49 (2014).

[34] L. A. Burns, M. S. Marshall, and C. D. Sherrill, J. Chem. Phys. **141**, 234111 (2014).

[35] L. A. Burns, Á. Vázquez-Mayagoitia, B. G. Sumpter, and C. D. Sherrill, J. Chem. Phys. **134**, 084107 (2011).

[36] M. S. Marshall, L. A. Burns, and C. D. Sherrill, J. Chem. Phys. **135**, 194102 (2011).

[37] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[38] M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, J. Am. Chem. Soc. **124**, 10887 (2002).

[39] M. O. Sinnokrot and C. D. Sherrill, J. Am. Chem. Soc. **126**, 7690 (2004).

[40] M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **108**, 10200 (2004).

[41] T. P. Tauer and C. D. Sherrill, J. Phys. Chem. A **109**, 10475 (2005).

[42] A. L. Ringer, M. O. Sinnokrot, R. P. Lively, and C. D. Sherrill, Chem. - Eur. J. **12**, 3821 (2006).

[43] M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **110**, 10656 (2006).

[44] E. G. Hohenstein and C. D. Sherrill, J. Phys. Chem. A **113**, 878 (2009).

[45] D. J. Tannor, B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, B. Honig, M. Ringnalda, and W. A. Goddard, J. Am. Chem. Soc. **116**, 11875 (1994).

[46] M. Pitoňák, P. Neogrády, J. Černý, S. Grimme, and P. Hobza, ChemPhysChem **10**, 282 (2009).

[47] J. Řezáč, Y. Huang, P. Hobza, and G. J. O. Beran, J. Chem. Theory Comput. **11**, 3065 (2015).