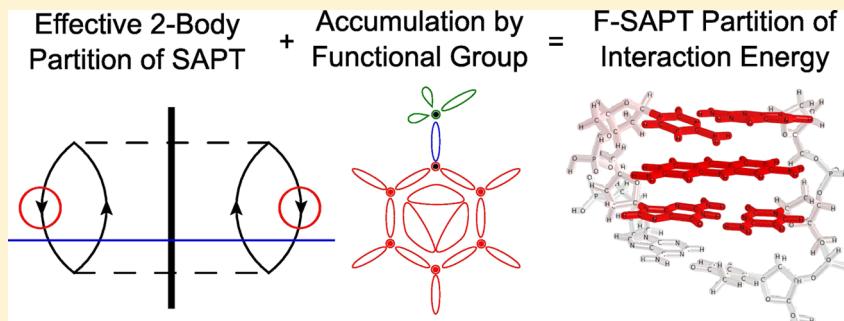


# Chemical Assignment of Symmetry-Adapted Perturbation Theory Interaction Energy Components: The Functional-Group SAPT Partition

Robert M. Parrish, Trent M. Parker, and C. David Sherrill\*

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, United States

## Supporting Information



**ABSTRACT:** Recently, we introduced an effective atom-pairwise partition of the many-body symmetry-adapted perturbation theory (SAPT) interaction energy decomposition, producing a method known as atomic SAPT (A-SAPT) [Parrish, R. M.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *141*, 044115]. A-SAPT provides *ab initio* atom-pair potentials for force field development and also automatic visualizations of the spatial contributions of noncovalent interactions, but often has difficulty producing chemically useful partitions of the electrostatic energy, due to the buildup of oscillating partial charges on adjacent functional groups. In this work, we substitute chemical functional groups in place of atoms as the relevant local quasiparticles in the partition, resulting in a functional-group-pairwise partition denoted as functional-group SAPT (F-SAPT). F-SAPT assigns integral sets of local occupied electronic orbitals and protons to chemical functional groups and linking  $\sigma$  bonds. Link-bond contributions can be further assigned to chemical functional groups to simplify the analysis. This approach yields a SAPT partition between pairs of functional groups with integral charge (usually neutral), preventing oscillations in the electrostatic partition. F-SAPT qualitatively matches chemical intuition and the cut-and-cap fragmentation technique but additionally yields the quantitative many-body SAPT interaction energy. The conceptual simplicity, chemical utility, and computational efficiency of F-SAPT is demonstrated in the context of phenol dimer, proflavine<sup>+</sup>-DNA intercalation, and a cucurbituril host-guest inclusion complex.

## 1. INTRODUCTION

While quantitative determination of noncovalent interactions is considered to be a critical challenge facing modern quantum chemistry, even robust qualitative chemical explanation of such interactions is difficult to attain. Consider, for just one instance, the large volume of work devoted to understanding the origins of interaction energy differences between stacked substituted benzene dimers.<sup>1–10</sup> Even simply producing a high-quality interaction energy for such a system is a formidable task: the extreme sensitivity of the interaction to the electronic structure of the constituent monomers and to the intermolecular orientations and distances present in the dimer mandate the use of such methods as complete basis set coupled cluster with singles, doubles, and perturbative triples [CCSD(T)/CBS].<sup>11</sup> For more qualitative analysis, less demanding electronic structure methods will often suffice, but the qualitative assignment of an interaction energy component to a particular chemical progenitor remains a delicate art. For instance, it

required many years of effort to reach the conclusion that substituent effects in stacked benzenes could be largely explained by local interactions between the substituent and nearby portion of the neighboring benzene.<sup>6,9,10</sup> Moreover, while this conclusion is quite useful to chemists interested in pure face-to-face stacking interactions, it does not seem to immediately transfer to slipped or T-shaped conformations of the benzene dimer, or to more exotic intermolecular motifs.<sup>4,5,12,13</sup> At present, intuitive intermolecular chemistry is lacking the simple and broadly transferrable pictures defining covalent chemistry; i.e., there are few analogs of the guiding concepts of Lewis structures, electronegativities, bond enthalpies, frontier orbitals, etc. in intermolecular chemistry. If we cannot currently provide the applications chemist with such a rulebook with which to design intermolecular interactions, then we

Received: August 8, 2014

Published: September 9, 2014

should at least provide fast and accurate methods to determine why an arbitrary intermolecular motif has a particular interaction energy. To that end, the standard theoretical analysis these days is a high-quality interaction energy and possibly a symmetry adapted perturbation theory (SAPT)<sup>14</sup> or energy decomposition analysis (EDA)<sup>15</sup> partition of the noncovalent interaction into electrostatics, steric repulsion, polarization, and dispersion contributions. While such analysis can provide useful insights for small model systems,<sup>16</sup> this insight becomes limited as one progresses to larger systems where multiple noncovalent contacts can occur. Unfortunately, standard SAPT or EDA methods provide only a single value for each of the interaction energy components, precluding elucidation of how each intermolecular contact contributes to the total interaction energy and its components.

Recently, we have been working to produce a robust and efficient partition of the SAPT decomposition into physically localized progenitors, i.e., to assign the interaction energy components to contributions from such chemically useful quasiparticles as protons and electrons, atoms, or functional groups. In our previous work,<sup>17</sup> we identified an effective two-body partition of the SAPT0 expressions<sup>18</sup> which both matches the chemists' intuition as to what each term represents and also respects the many-body limit of the SAPT interaction energy. In the previous approach, terms involving the response of the monomer wave functions (e.g., dispersion) used local occupied orbitals as the relevant quasiparticles, while terms involving electrostatic potentials (e.g., electrostatics) grouped fractions of the molecular electronic density and nuclei into atoms to prevent enormous oscillations in the partition. The terms involving local occupied orbitals could additionally be assigned to atoms via orbital atomic charges, yielding a complete spatial partition of SAPT in terms of pairs of atoms. We denoted this method atomic SAPT (A-SAPT). By automatically providing an effective atom-pairwise partition of the *ab initio* SAPT interaction energy, A-SAPT has obvious applications in the development of next-generation molecular dynamics force fields. With respect to supporting chemical understanding of noncovalent contacts, A-SAPT can also produce automatic visualizations which clearly show the most important chemical elements in complicated intermolecular interactions. For response terms, A-SAPT can even provide both qualitative and quantitative understanding; for example, it was able to identify the origin of the induction stabilization of side-on cation–benzene and cation–cyclohexane interactions in terms of C–H  $\sigma$ , C–C  $\sigma$ , and  $\pi$  orbital contributions.<sup>17</sup> Unfortunately, as was previously reported<sup>17</sup> for the case of a charged proflavine<sup>+</sup>–DNA intercalation complex, A-SAPT sometimes exhibits oscillating and nearly canceling contributions to the electrostatic energy, particularly in systems with charged or highly polar monomers. This complicates attempts to analyze and understand how different functional groups of atoms contribute to the electrostatic component of the interaction energy. The root of this difficulty has been identified as the very concept of atoms in molecules:  $\sigma$  bonds linking neighboring functional groups will necessarily be assigned preferentially to the more electronegative atom, leading to oscillating charges between the neighboring functional groups. This gives rise to large charge–charge interactions between intermolecular pairs of functional groups, obfuscating the true higher-order multipole–multipole interactions which are intuitively expected. Note that the trouble with the atoms-in-molecules picture only arises when a

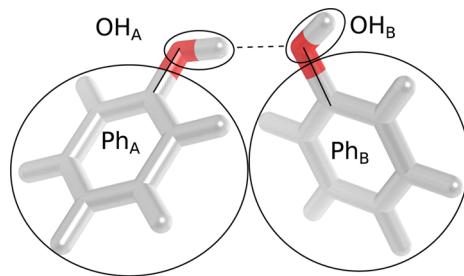
partition into chemically relevant functional group contributions is attempted. For many problems, the atoms-in-molecules picture is the correct *ansatz*, e.g., as in the development of atom-pairwise force fields, for which A-SAPT is a well-posed tool.

In this work, we ameliorate the problems with respect to chemical intuition in the electrostatics term by switching from an atom-based *ansatz* to a functional-group-based *ansatz*. In the new functional-group SAPT (F-SAPT) approach, integral sets of localized electrons and protons are assigned to chemical functional groups and linking  $\sigma$  bonds. All of these fragments are neutral or carry a chemically expected net charge (e.g., the carboxylate ion functional group would carry a net charge of  $-1$  in our analysis), preventing canceling oscillations in the electrostatic partition. To simplify the analysis, an even assignment of the link-bond contributions to their adjacent functional groups is used after the full link-bond and functional group partition is accumulated. The division of an intermolecular interaction energy decomposition into identifiable orbital or functional group contributions is somewhat reminiscent of the natural orbital EDA (NEDA) technique,<sup>19</sup> absolutely localized molecular orbital EDA (ALMO EDA) methodology,<sup>20–22</sup> fragment molecular orbital (FMO)<sup>23</sup> based pairwise interaction EDA (PIEDA),<sup>24</sup> or divide-and-conquer semiempirical techniques.<sup>25</sup> We also note a completely different but complementary approach: the use of density-based heuristics to visually indicate the presence of noncovalent interactions, including those based on electron localization functions (ELF)<sup>26</sup> or localized orbital locators (LOL),<sup>27</sup> the noncovalent interactions (NCI) scalar field,<sup>28,29</sup> or the recent density overlap regions indicator (DORI).<sup>30</sup> Finally, we note that many important developments toward spatially localizing interaction energy components to fragments of monomers have been made by Stone and co-workers, e.g., in distributed multipole analysis (DMA) of electrostatics<sup>31–34</sup> or distributed polarizabilities.<sup>35,36</sup> However, these approaches work almost wholly in terms of atomic contributions and, as such, are much more analogous to our previous A-SAPT approach than the present F-SAPT approach.

Below, we outline the intuitive development of the F-SAPT partition by first identifying desired characteristics of the new approach and weaknesses of other partition methods such as the cut-and-cap fragmentation approach and A-SAPT. We then construct F-SAPT to satisfy all of the desired constraints and briefly detail the straightforward modifications to the A-SAPT program needed to obtain an implementation of F-SAPT at the level of SAPT0. As with any heuristic partition, F-SAPT can be neither derived nor proved, so we rely on practical discussion and chemical examples to justify the method. We focus on the most-difficult electrostatic term to simplify the discussion and use phenol dimer as an illustrative example. In the results and discussion section, we apply cut-and-cap fragmentation, A-SAPT, and F-SAPT to provide functional-group pairwise partitions of phenol dimer and the proflavine<sup>+</sup>–DNA intercalator (whose electrostatic oscillations under A-SAPT motivated this work). These case studies demonstrate that F-SAPT produces chemically useful partitions of the electrostatic energy while also respecting the many-body limit of the interaction energy, placing it clearly above cut-and-cap or A-SAPT methodologies. We close by discussing a complete F-SAPT analysis for an interesting cucurbituril inclusion complex.

## 2. THEORY

**2.1. Motivation: Phenol Dimer.** To frame the discussion, we consider the case of phenol dimer in a hydrogen-bonded arrangement (from the S22 database),<sup>37</sup> depicted in Figure 1.



**Figure 1.** Phenol dimer geometry and functional group partitioning scheme. Ovals denote functional group selection. Link bonds are indicated by solid black lines. Intermolecular hydrogen bonds are indicated by dashed black lines.

We denote the hydrogen-bond-donating member as monomer A and the hydrogen-bond-receiving member as monomer B. Standard SAPT0/jun-cc-pVDZ<sup>18,38,39</sup> analysis indicates a total interaction energy of  $-7.1 \text{ kcal mol}^{-1}$ , with contributions of  $-9.1$ ,  $+9.9$ ,  $-0.5$ ,  $-2.3$ , and  $-5.1 \text{ kcal mol}^{-1}$  from  $E_{\text{elst}}$ ,  $E_{\text{exch}}$ ,  $E_{\text{ind(A-B)}}$ ,  $E_{\text{ind(B-A)}}$ , and  $E_{\text{disp}}$ , respectively. This interaction is somewhat stronger than the  $-5.1 \text{ kcal mol}^{-1}$  of the water dimer (also from the S22 database) and considerably different in both geometry and interaction energy from the standard configurations of the benzene dimer. What, then, are the chemical origins of this interaction? How important is the  $\pi-\pi$  interaction relative to the hydrogen bonding interaction, and are OH- $\pi$  interactions relevant? Below, we will pursue several approaches to answer these questions, focusing on the critical and difficult electrostatics term.

To the chemist, these phenol monomers ( $\text{PhOH}$ ) are intuitively defined as a phenyl group ( $\text{Ph}$ ) attached to a hydroxyl group ( $\text{OH}$ ) via a  $\sigma$  bond, as seen in Figure 1. Therefore, the most useful partition of the  $\text{PhOH}_A \cdots \text{PhOH}_B$  interaction would be into the pair contributions of  $\text{Ph}_A \cdots \text{Ph}_B$ ,  $\text{Ph}_A \cdots \text{OH}_B$ ,  $\text{OH}_A \cdots \text{Ph}_B$ , and  $\text{OH}_A \cdots \text{OH}_B$ . There are two fundamental paradigms for such partitions. The first is the family of fragmentation techniques, which *a priori* divide the monomers into relevant subsystems and then perform isolated pairwise interaction energy computations on the subsystems. The other is the family of many-body assignment techniques, wherein a single interaction energy is computed between the two complete monomers, and the interaction energy components are *a posteriori* divided between the functional groups. A-SAPT is a member of the latter paradigm. The fragmentation approaches are compelling due to their conceptual and computational simplicity. However, the fragment approaches have three key problems that can be remedied by the many-body assignment techniques:

1. The computation of the interaction energy contributions in terms of pairs of isolated fragments ignores possibly very large intramolecular many-body coupling effects between different functional groups within the monomers. This implies that the *ab initio* limit cannot ever be achieved for such methods, even if a convergent *ab initio* methodology such as SAPT or coupled-cluster theory is used for the isolated pairwise interaction energy computations. Note that this caveat directly applies only to explicitly pairwise interaction fragment models:

higher order fragmentation models (e.g., which include trimers and perhaps even larger clusters) can approach the *ab initio* limit of the interaction energy, albeit at the cost of losing a straightforward pairwise partition of the interaction energy.

2. When covalent bonds link two functional groups, myriad problems arise in the division of the monomers into fragments. In particular, it is difficult to provide a scheme which simultaneously respects particle count, charge neutrality of adjacent functional groups, link bond polarization, etc.

3. The best schemes involve cleavage of the link bonds and capping with hydrogen atoms to maintain the neutral closed-shell nature of the fragments, which can involve additional user effort to produce geometries and analyze results.

**2.1.1. Fragmentation: Homolytic Cleavage.** One particularly straightforward but ill-formed fragmentation approach is to simply cleave the link bond homolytically, assigning one link electron to the  $\text{Ph}$  group and the other to the  $\text{OH}$  group in each monomer. This effectively defines the monomers as the neutral radical species  $\text{Ph}^\bullet$  and  ${}^{\bullet}\text{OH}$ . Pairwise interaction energy computations could then be performed between these radical species, e.g.  $(\text{Ph}^\bullet)_A \cdots ({}^{\bullet}\text{OH})_B$ . This approach has the advantages that the particle count (the total number of protons and electrons) is maintained and that the  $\text{Ph}^\bullet$  and  ${}^{\bullet}\text{OH}$  are both neutral. However, this approach is fraught with peril: the electronic structures of the radical fragments are often nothing like the underlying monomer functional groups, which are stabilized by electron pairing in the link  $\sigma$  orbital.

**2.1.2. Fragmentation: Heterolytic Cleavage.** One possible remedy for the homolytic cleavage approach is to recognize that the pairing of electrons is vital and to assign *both* of the electrons in the link  $\sigma$  bond to the more electronegative functional group. This effectively defines the monomers as the charged closed-shell species  $\text{Ph}^+$  and  ${}^-\text{OH}$ . This approach has the advantages that the particle count is maintained and that there are no radical species involved but has the distinct disadvantage that the fragments are charged to the magnitude of a complete electron, in an oscillating  $\pm$  fashion that cancels in the net. For instance, the  ${}^-\text{OH}_A \cdots {}^-\text{OH}_B$  electrostatic interaction is dominated by a repulsive charge-charge interaction on the order of  $(-1)^2/R$ , where  $R$  is the average separation between the hydroxyl groups. The contribution is a sizable fraction of a Hartree and completely masks the expected attractive dipole-dipole interaction that characterizes a hydrogen-bonding interaction.

**2.1.3. Fragmentation: Cut-and-Cap.** A superior approach is to homolytically cut each link bond and then cap the unpaired electrons in each functional group with hydrogens, each of which contains a proton and an electron. This effectively defines the monomers as the neutral closed-shell species benzene ( $\text{Bz}$ ) and  $\text{H}_2\text{O}$ . This approach maintains the pairing of electrons and neutrality of the functional groups and roughly retains the anisotropy of the link bond (though the overall link bond polarity is almost always lost in the new fragments). Unfortunately, the link bonds are essentially counted twice in this approach, as the particle count is increased by two protons and two electrons for each link bond. Often, the particle count effect is assumed to be minor and ignored. Alternatively, one can subtract a hydrogen molecule placed along each link bond to maintain the particle count, e.g., defining the phenol as  $\text{Bz} + \text{H}_2\text{O} - \text{H}_2$ . This correction can significantly clutter the analysis, e.g., due to the ambiguity in assigning the  $\text{H}_2$  contributions to the functional groups. An additional concern is the polarization of the capped link bonds. Link bonds are usually of X-Y

topology where X and Y are heavy atoms and are much less polarized than the X–H and Y–H bonds they will be replaced with, potentially cluttering the electrostatics and induction terms. Finally, one last ambiguity arises in the placement of the protons in the capped bonds; these are often handled by constrained DFT optimizations with all nuclei but those of the capping hydrogens frozen, mandating additional computational effort.

This approach has previously been used to roughly elucidate the relative contributions of chemical functional groups to the interaction energy, e.g., in probing the hypothesis that substituent effects in benzene dimers are primarily direct, through-space effects,<sup>6</sup> or in illuminating the many competing contributors in proflavine<sup>+</sup>–DNA intercalation.<sup>40</sup>

**2.1.4. Many-Body Assignment: A-SAPT.** An alternative to fragmentation is to compute the many-body interaction through a convergent *ab initio* method, while retaining a two-body quasiparticle partition identified from the many-body interaction energy diagrams. One such method is A-SAPT, which extracts the effective two-body partition from the formalism of SAPT (particularly SAPT0 in our previous work) and then uses either local occupied orbitals or iterative stockholder analysis (ISA)<sup>41–43</sup> atoms as the quasiparticle set. The invocation of the concept of atoms is critical to obtain an automatic and sensible partition of the electrostatic energy component and the electrostatic sources in the induction energy component: in this paradigm, fractions of the molecular electronic density are assigned to atoms to provide reasonable cancellation between the extremely large electronic and nuclear contributions to these terms. This atom-based assignment leads to an efficient and black-box method which produces useful visualizations identifying the active portions of complicated electrostatic interactions.

For the full technical details of A-SAPT, we refer the reader to our previous work,<sup>17</sup> but it is appropriate at this point to provide a self-contained description of A-SAPT, to facilitate the discussion below. The development of the A-SAPT partition is a two-stage procedure. In the first stage, the many-body interaction energy diagrams (or corresponding algebraic equations) for a given level of SAPT are examined, and an effective two-body interaction is identified for each SAPT diagram, with one body taken from each monomer. As the SAPT diagrams may contain anywhere from two to an infinite number of occupied or virtual quasiparticles (and include both electronic and nuclear contributions), this choice is not unique. However, we have found that the identification of a chemically sensible effective two-body interaction is entirely straightforward, at least at the level of SAPT0 or DFT-SAPT. In particular, we restrict the partition to occupied quasiparticles only, summing over the virtual quasiparticles. For electrostatics and for the electrostatic sources in induction, this is an entirely rigorous choice, as no corresponding virtual quasiparticles are present. For exchange, dispersion, and the response portion of induction, the diagrams are all characterized by excitations from the occupied quasiparticles to virtual quasiparticles of one monomer, in response to some perturbation provided by the adjacent monomer. For these terms, the intuitive choice for the partition is the occupied quasiparticle from which the excitation arose, not the virtual quasiparticle to which the excitation concluded, as we are concerned with those portions of the monomers from which the interaction energy term originate. Having made the restriction to occupied quasiparticles, the choice of a specific pair of occupied quasiparticles is obvious.

For instance, the dispersion contribution involves the trace of a double excitation  $t_{ab}^s$  with the Coulomb integral  $v_{rs}^{ab}$ , which obviously involves only the two occupied quasiparticles *a* and *b*. The corresponding exchange-dispersion term involves additional occupied quasiparticles, but these encapsulate Pauli-forbidden collisions between the electrons involved in the double excitation and the occupied spaces of the complementary monomers, so we continue to focus on the two occupied quasiparticles *a* and *b*.

In the second stage, an appropriate set of spatially localized quasiparticles must be selected to provide a spatially meaningful partition of the effective two-body interaction identified in the first stage. In the development of A-SAPT, our initial aim was to provide a completely black-box partition (e.g., there was to be no invocation of user-specified subsets of atoms to denote functional groups). Moreover, we also intended to provide an automatic partition of the interaction energy to atom-pairwise contributions, to facilitate the development of *ab initio* force fields. Both of these considerations mandated the selection of atoms as the proper quasiparticles, and we further selected the iterative stockholder analysis (ISA) method of Wheatley and co-workers to provide these atoms.<sup>41–43</sup> ISA is a variant of the Hirshfeld family of partitions to automatically extract a set of atomic density fields which sum to the total molecular density field. Hirshfeld-type partitions assign a spherical proatomic density to each atom and then obtain the true atomic density contribution by weighting the total molecular density by the ratio between the proatomic density and the total promolecular density at each point in space. This results in smooth, overlapping atomic density fields. The difference between the various Hirshfeld-type partitions lies in the particular choice of spherical proatomic densities: Hirshfeld<sup>44</sup> uses the sphericalized densities of the underlying free atoms which are neutral; Hirshfeld-I<sup>45</sup> uses the sphericalized densities of underlying free atoms with charges selected to be equal to those of the true atomic densities, while ISA uses spherical proatomic densities which are wholly unconstrained in the radial coordinate. In the latter two approaches, the radial forms of the proatomic densities are determined iteratively by minimizing the information loss between the true molecular density and the promolecular density. Our selection of ISA was based on the fact that ISA makes no *a priori* constraints on the proatomic densities beyond being spherical, though we note that A-SAPT could also be easily adapted to use Hirshfeld, Hirshfeld-I, or any other atoms-in-molecules scheme. Having selected ISA as the atoms-in-molecules approach, it remains to connect the effective two-body partition of the SAPT diagrams to the ISA atom-pairwise contributions. For electrostatics and for the electrostatic sources in induction, the atomic contributions are directly obtained by finding the electrostatic potential or electrostatic energy contribution of each ISA atom. For the exchange, dispersion, and response partition of induction, the atomic contributions are found by first accumulating in terms of local occupied orbitals and then assigning fractions of each local occupied orbital contribution to atomic contributions via ISA orbital atomic charges. This procedure closes the definition of ISA-based A-SAPT0, yielding an atom-pairwise partition for all SAPT0 terms.

It seems obvious that A-SAPT should be able to provide a quantification of functional-group...functional-group interactions, i.e., by simply summing up the relevant atomic contributions *A* and *B* to the two functional groups  $\mathcal{A}$  and  $\mathcal{B}$ :

$$E_{\text{int-term},A-\text{SAPT}}^{\mathcal{AB}} = \sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} E_{\text{int-term},A-\text{SAPT}}^{AB} \quad (1)$$

Unfortunately, as noted in the context of the proflavine–DNA intercalation complex of our previous work, A-SAPT can produce unexpected results if quantitative chemical partitions of the electrostatic energy are attempted. The root of this problem is the very concept of atoms in molecules: regardless of the definition of the atom type (Hirshfeld,<sup>44</sup> Hirshfeld-I,<sup>45</sup> ISA,<sup>41</sup> GISA,<sup>46</sup> Bader,<sup>47</sup> etc.), every atom in each molecule will, in general, carry a net partial charge, determined by its arrangement in space relative to the atoms around it, and to the electronegativities of the surrounding atoms. In particular, link bonds between adjacent functional groups are almost always assigned preferentially to the functional group with the more electronegative linking atom, leading to very large charge–charge interactions between functional groups on the two monomers. In the case of phenol dimer, the problem is obvious: the greater electronegativity of the oxygen means that the OH will carry a charge of roughly −0.2 electrons, while the Ph will carry the balance of +0.2 electrons (at ISA/jun-cc-pVDZ as in standard A-SAPT0 analysis). This implies, e.g., that the favorable hydrogen-bond interaction expected in the OH<sub>A</sub>⋯OH<sub>B</sub> electrostatic interaction is actually dominated by the unfavorable (−0.2)<sup>2</sup>/R charge–charge interaction. This charge–charge repulsion, plus the corresponding charge–charge repulsion of the Ph<sub>A</sub>⋯Ph<sub>B</sub> interaction, largely cancels the charge–charge attractions from the two Ph⋯OH interactions. Overall, this leads to a muddled analysis which obfuscates the desired dipole–dipole and higher interactions between neutral functional groups. Note that this phenomenon is not unique to A-SAPT: all other atom-based methods (e.g., DMA or atom-pairwise force fields) will necessarily suffer this same artifact.

**2.1.5. Many-Body Assignment: F-SAPT.** Motivated by the failure of A-SAPT to produce a chemically reasonable functional-group partition of the electrostatic energy, we now develop a paradigm which ameliorates this difficulty by directly considering chemically separable functional groups of integral charge (usually neutral) connected by neutral link bonds. This approach closely resembles the cut-and-cap fragmentation technique above but retains the possibility of reaching the *ab initio* limit by not performing pairwise computations in isolation. Additionally, arbitrary division of the monomers into any desired patterns of functional groups can easily be obtained in postprocessing, from the results of a single SAPT-type computation on the original dimer. We refer to this new approach as “functional-group SAPT,” or F-SAPT.

F-SAPT starts from the A-SAPT-like assertion that the interaction energy terms can be partitioned to pairwise effective interactions between nuclei (*A* or *B*) or localized orbitals ( $\bar{a}$  or  $\bar{b}$ , doubly occupied), e.g.,

$$E_{\text{int-term}} = E_{\text{int-term}}^{\bar{a}\bar{b}} + E_{\text{int-term}}^{\bar{a}B} + E_{\text{int-term}}^{A\bar{b}} + E_{\text{int-term}}^{AB} \quad (2)$$

A-SAPT is already nearly in this form. For instance, the exchange and dispersion terms of A-SAPT0 are already explicitly defined in terms of local orbital pairs and involve no partitioned nuclear contributions. Only terms involving electrostatic potentials need modification: these were previously assigned directly to atomic quasiparticle contributions in A-SAPT, to avoid extreme oscillations in the visualization. The new nuclear and orbital partition for F-SAPT0 is quite easily obtained, as the contribution of each charge source to the

electrostatic potential is linear. For the electrostatics term in SAPT0 (closed-shell RHF), the partition is

$$\begin{aligned} E_{\text{elst,r}}^{(10),\bar{a}\bar{b}} &\equiv 4(\bar{a}\bar{a}|\bar{b}\bar{b}) \\ &= 4 \iint_{\mathbb{R}^6} d^3r_1 d^3r_2 \phi_{\bar{a}}(\vec{r}_1) \phi_{\bar{a}}(\vec{r}_2) \frac{1}{r_{12}} \phi_{\bar{b}}(\vec{r}_2) \phi_{\bar{b}}(\vec{r}_2) \end{aligned} \quad (3)$$

where  $\phi_{\bar{a}}(\vec{r})$  is the  $\bar{a}$ th local occupied orbital,

$$E_{\text{elst,r}}^{(10),\bar{a}B} \equiv -2(\bar{a}\bar{a}|Z_B) = -2 \iint_{\mathbb{R}^3} d^3r_1 \phi_{\bar{a}}(\vec{r}_1) \phi_{\bar{a}}(\vec{r}_1) \frac{Z_B}{r_{1B}} \quad (4)$$

where  $Z_B$  is the charge of the *B*th nucleus, and similarly for  $E_{\text{elst,r}}^{(10),A\bar{b}}$  and

$$E_{\text{elst,r}}^{(10),AB} \equiv \frac{Z_A Z_B}{r_{AB}} \quad (5)$$

For the induction term, the new partition produces separate nuclear and local orbital contributions to the total electrostatic potential

$$\begin{aligned} V_{\bar{a}r}^B &\equiv -(\bar{a}r|\rho_B) \\ &= - \iint_{\mathbb{R}^6} d^3r_1 d^3r_2 \phi_{\bar{a}}(\vec{r}_1) \phi_r(\vec{r}_1) \frac{1}{r_{12}} \rho_B(\vec{r}_2) \\ &\equiv V_{\bar{a}r}^{\bar{b}} + V_{\bar{a}r}^B \end{aligned} \quad (6)$$

where  $\rho_B(\vec{r})$  is the total charge density of monomer B. Here,

$$\begin{aligned} V_{\bar{a}r}^{\bar{b}} &\equiv 2(\bar{a}r|\bar{b}\bar{b}) \\ &= 2 \iint_{\mathbb{R}^6} d^3r_1 d^3r_2 \phi_{\bar{a}}(\vec{r}_1) \phi_r(\vec{r}_1) \frac{1}{r_{12}} \phi_{\bar{b}}(\vec{r}_2) \phi_{\bar{b}}(\vec{r}_2) \end{aligned} \quad (7)$$

and

$$V_{\bar{a}r}^B \equiv -(\bar{a}r|Z_B) = - \iint_{\mathbb{R}^6} d^3r_1 d^3r_2 \phi_{\bar{a}}(\vec{r}_1) \phi_r(\vec{r}_1) \frac{Z_B}{r_{1B}} \quad (8)$$

These nuclear and local occupied orbital sources each produce linearly additive induction responses, encapsulated by the induction amplitudes  $x_{\bar{a}r}^{\bar{b}}$  or  $x_{\bar{a}r}^B$ . As in our previous work (see eqs 18 and 19 in ref 17), these amplitudes can be obtained through either the uncoupled approximation (used throughout this work, with scalings as in the previous work) or through coupled-perturbed Hartree–Fock. Once the amplitudes are obtained, the full pairwise partition to the induction energy is obtained as

$$E_{\text{ind}(A \leftarrow B)}^{(20),\bar{a}\bar{b}} \equiv 2x_{\bar{a}r}^{\bar{b}} [V_{\bar{a}r}^B + \tilde{V}_{\bar{a}r}^B] \quad (9)$$

and

$$E_{\text{ind}(A \leftarrow B)}^{(20),\bar{a}B} \equiv 2x_{\bar{a}r}^B [V_{\bar{a}r}^B + \tilde{V}_{\bar{a}r}^B] \quad (10)$$

Here,  $\tilde{V}_{\bar{a}r}^B$  is the exchange-renormalization potential of the exchange-induction term, defined in our previous work (see eq 21 of ref 17). Note that the blocks of  $E_{\text{ind}(A \leftarrow B)}$  corresponding to the nuclei *A* are zero, as the involvement of monomer *A* in this term is an electronic response only. Further, corresponding expressions for  $E_{\text{ind}(B \leftarrow A)}$  can be obtained by permutation of indices.

In practice, the above modifications are easily and efficiently implemented, using the density fitting approximation<sup>48–51</sup> in the -JKFIT auxiliary basis sets.<sup>52</sup> These terms are actually much

easier to implement and evaluate than those in the atom-based A-SAPT partition, as the above F-SAPT expressions are all tractable in a spectral basis set and do not require numerical quadrature to evaluate.

The nuclear and local orbital partition above is easy to attain, but it does not immediately yield interpretable data. One reason for this is that electronic and nuclear sources have been split, leading to enormous and nearly canceling contributions to the partition. To overcome this, we group chemically intuitive electronic and nuclear contributions into fragments  $\mathcal{A}$  and  $\mathcal{B}$ , including chemical functional groups and link-bond subsystems. The mathematical form for this process is given as

$$\begin{aligned} E_{\text{int-term,full}}^{\mathcal{A},\mathcal{B}} \equiv & w_{\bar{a}}^{\mathcal{A}} w_{\bar{b}}^{\mathcal{B}} E_{\text{int-term}}^{\bar{a}\bar{b}} + w_{\bar{a}}^{\mathcal{A}} w_{\bar{B}}^{\mathcal{B}} E_{\text{int-term}}^{\bar{a}B} \\ & + w_A^{\mathcal{A}} w_{\bar{b}}^{\mathcal{B}} E_{\text{int-term}}^{AB} + w_A^{\mathcal{A}} w_B^{\mathcal{B}} E_{\text{int-term}}^{AB} \end{aligned} \quad (11)$$

The quantities  $w_{\bar{a}}^{\mathcal{A}}$  and  $w_A^{\mathcal{A}}$  are the orbital-to-fragment and nucleus-to-fragment weights, respectively. The orbital-to-fragment weights  $w_{\bar{a}}^{\mathcal{A}}$  are always 0 or 1, as each orbital is wholly assigned to a single fragment in this scheme. The nucleus-to-fragment weights are also 0 or 1 for any atom not included in a link bond, in which case the weight will be a fraction between 0 and 1, as described below.

The general procedure for assigning orbital and nuclear weights is as follows. First, the monomer geometry is traversed to identify bridging  $\sigma$  bonds between user-specified functional groups. If linking covalent bonds are found (typically by inspection of the interatomic distances and van der Waals radii, followed by orbital charge analysis), the corresponding link fragment will be assigned a single  $\sigma$  orbital bridging the two constituent atoms (by “assigned,” it is meant that the corresponding weight  $w_{\bar{a}}^{\mathcal{A}}$  is set to 1). Once this is done, the remaining orbitals are assigned to the user-specified functional groups according to the functional group on which each orbital has the largest atomic charge. With the orbital-to-fragment weights assigned, the next task is the nucleus-to-fragment weights. To seed these weights,  $w_A^{\mathcal{A}}$  is set to 1 if the nucleus  $A$  is in the user-defined functional group  $\mathcal{A}$  and 0 otherwise. From this, the fraction  $1/Z_A$  is removed from the weight  $w_A^{\mathcal{A}}$  and assigned to the weight  $w_A^{\mathcal{A}'}$ , for each of the two atoms in the link fragment  $\mathcal{A}'$ . This effectively removes a proton from each functional group involved in the link bond and assigns these protons to the link bond fragment, resulting in a neutral link fragment with two electrons in a  $\sigma$  orbital and two protons. This shuffling of two protons to each linking  $\sigma$  bond is critical to achieve a sensible partition: if the protons were left with the functional groups, the functional groups would usually carry net positive charges while the link bond groups would each carry a charge of  $-2$ , leading to large oscillations in the electrostatic partition. Another way to view this division of the monomer is into a set of neutral functional groups interconnected by a set of link fragments (each of which loosely resembles an extended hydrogen molecule).

In practice, much of the above weight selection can be automated, with excellent results for chemically reasonable choices of functional groups. The user first provides the F-SAPT postprocessing script with the set of atoms defining each chemically relevant functional group (e.g., O and H to define the OH functional group and C<sub>6</sub>H<sub>6</sub> to define the Ph functional group). The distances between pairs of atoms in difference are then checked against the sum of van der Waals radii of the

constituent atoms, and candidates for link bonds are identified automatically (this can be specified manually in special cases). For each link bond between atoms  $A$  and  $A'$ , the local orbital corresponding to the linking  $\sigma$  bond is easily identified as that which maximizes the bond-index-like metric  $M_{\bar{a}} = Q_A^{\bar{a}} Q_{A'}^{\bar{a}}$ , where  $Q_A^{\bar{a}}$  are local orbital atomic charges computed at the same level as the Pipek–Mezey-like orbital localization routine. Once the link orbitals are assigned, the remaining orbitals are assigned to the chemical functional group containing the atom with the maximal orbital atomic charge  $Q_A^{\bar{a}}$ . The nuclear weights are then automatically assigned as above.

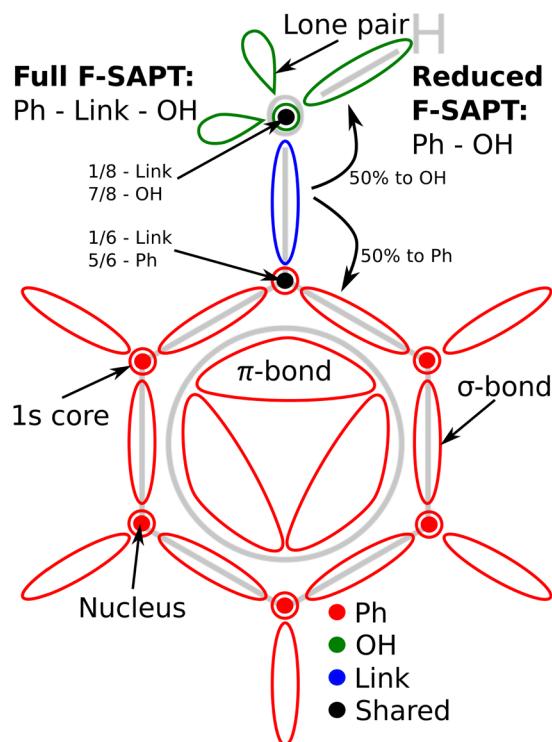
A key assumption in the F-SAPT analysis is that the functional groups are chemically separable entities linked by, at most, simple  $\sigma$  bonds. This is obvious when one considers the weights  $w_{\bar{a}}^{\mathcal{A}}$ , which are integral, meaning that we have chosen the partition in such a way that nonlink electrons are wholly owned by a single functional group. This implies that neighboring functional groups linked by double, aromatic, etc. bonds or featuring significant fractional charge transfer are not appropriate choices within F-SAPT. We do not believe this represents a limitation of F-SAPT but rather a form of maximum resolution of effective quasiparticle separability in the many-body interaction: two user-defined fragments linked by more than simple  $\sigma$  bonds are actually acting as a single chemical entity, and should be treated as such. Moreover, F-SAPT does provide for polarization of the electrons within each functional group, to respond to neighboring functional groups (this flexibility is not present in fragmentation approaches, which are performed on pairs of fragments in isolation). An excellent automatic check of the validity of user-defined functional groups is the completeness of the orbital atomic charges

$$C_{\bar{a}} = 1 - \sum_{A \in \mathcal{A}} Q_A^{\bar{a}} \quad (12)$$

If  $C_{\bar{a}}$  is greater than a few hundredths of an electron pair, this is an indication that a significant piece of this orbital lies across more than one fragment and that the functional group partition should be redefined.

For example, consider the F-SAPT partition scheme for phenol, depicted in Figure 2. The link  $\sigma$  bond is clearly identifiable from the orbital atomic charges and is also assigned one proton from the OH oxygen (a nuclear-to-fragment weight of  $1/8$ ) and one proton from the nearest carbon on the phenyl group (a nuclear-to-fragment weight of  $1/6$ ). The hydroxyl group retains seven of the eight oxygen protons and the hydrogen proton, as well as the oxygen 1s orbital, the OH  $\sigma$  orbital and the two lone-pair orbitals. The phenyl group retains five of the six protons on the linking carbon, the five remaining carbon nuclei, the five hydrogen nuclei, the six carbon 1s orbitals, the six C–C  $\sigma$  orbitals, the five C–H  $\sigma$  orbitals, and the three  $\pi$  orbitals. Note that the  $\pi$  orbitals form a delocalized system over the carbons of the phenyl group and cannot be unambiguously localized to subsets of these atoms. However, the  $\pi$  system can be localized to the level of the phenyl functional group, which functions as a single, separable chemical unit. The largest value of  $C_{\bar{a}}$  for this system is 0.045 electron pairs, indicating mild polarization of one of the oxygen lone pairs toward the  $\pi$  system of the benzene.

To this point, the F-SAPT analysis produces a pairwise partition of the interaction energy terms between chemical functional groups and link bond fragments, e.g., Ph<sub>A</sub>…Ph<sub>B</sub>,



**Figure 2.** Phenol monomer F-SAPT partitioning scheme. In full F-SAPT, one proton from the hydroxyl O and one proton from the nearest phenyl C are assigned to the link  $\sigma$  bond to produce neutral Ph, Link, and OH fragments. In reduced F-SAPT, the full F-SAPT partition for the Link fragment is split equally between the Ph and OH fragments.

$\text{Ph}_A \cdots \text{OH}_B$ ,  $\text{Ph}_A \cdots \text{Link}_B$ , etc. contributions. We refer to this level of detail as a “full F-SAPT” partition. For detailed analysis, the explicit enumeration of the link fragment contributions may be important, particularly if these (often polar) bonds are significantly involved in the interaction. In these cases, the link bonds are chemical entities of their own and should be considered separately from the chemical functional groups. However, if the link bonds do not contribute significantly to the interaction, or if the user is only concerned with semi-quantitative trends, the analysis can be vastly simplified by assigning the link fragment contributions to the two functional groups that they bridge, i.e.,

$$\begin{aligned} E_{\text{int-term,reduced}}^{\mathcal{AB}} &= E_{\text{int-term,full}}^{\mathcal{AB}} + f_{\mathcal{B}'}^{\mathcal{B}} E_{\text{int-term,full}}^{\mathcal{AB}'} \\ &+ f_{\mathcal{A}'}^{\mathcal{A}} E_{\text{int-term,full}}^{\mathcal{A}B} + f_{\mathcal{A}'}^{\mathcal{A}} f_{\mathcal{B}'}^{\mathcal{B}} E_{\text{int-term,full}}^{\mathcal{A}\mathcal{B}'} \end{aligned} \quad (13)$$

Here  $f_{\mathcal{A}'}^{\mathcal{A}}$  is a weight assigning all or part of the link fragment  $\mathcal{A}'$  to the chemical functional group  $\mathcal{A}$ , which is only nonzero if  $\mathcal{A}'$  and  $\mathcal{A}$  share a nucleus. Unfortunately, there is no clearly unambiguous scheme to provide these weights: one could imagine assigning all of the link bond to one of the two functional groups, assigning some fraction of the link fragment to the two functional groups based on the ratio of orbital atomic charges, or even dividing the link fragment between the two functional groups. We have found that, for chemically reasonable choices of functional groups, all of these schemes generally produce qualitatively identical results, with the last two producing nearly quantitatively identical results in most cases. The last two schemes also contain a bit more physical

merit than the first: partitioning based on orbital charges respects the concept of how much each fragment “owns” the link  $\sigma$  orbital, while equal partitioning produces chemical functional groups with integral particle counts in both electrons and protons (e.g., the OH group in phenyl contains nine electrons and nine protons with equal partitioning). Due to the particular elegance of particle count of the last choice, and the lack of a concrete definition of an orbital atomic charge, we have selected the equal partitioning scheme ( $f_{\mathcal{A}'}^{\mathcal{A}} = 0.5$ ) for use in simpler, semiquantitative F-SAPT analysis. We denote this choice of partition as a “reduced F-SAPT” analysis. By default, “F-SAPT” below refers to reduced F-SAPT, while consideration of full F-SAPT will be explicitly enumerated.

**2.2. Interpretations and Limitations.** In the process of review, a reviewer raised the admirable critique that F-SAPT is necessarily not the only possible decomposition of the SAPT interaction energy and, moreover, raised the complaint that our choice of partition ignores the virtual quasiparticles present in the response portions of the SAPT expressions (the exchange, dispersion, and induction response terms). We provide this section to discuss the limitations and choices made in the development of F-SAPT.

One way to view F-SAPT is to assert that the monomers are composed of transferable collections of localized electrons and protons/neutrons (functional groups) which, after selection under some localization metric and prescribed set of functional groups, are allowed to interact naturally with one another. At the level of Hartree–Fock theory (as encountered in SAPT0 here), the protons/neutrons and occupied orbitals completely parametrize the states of the noninteracting monomers, albeit with a mean-field description of the wave functions. If Kohn–Sham theory were invoked (as in DFT-SAPT, which we intend to apply in later efforts), the protons/neutrons and localized occupied Kohn–Sham orbitals could theoretically be used to completely parametrize the *exact* states of the noninteracting monomers, if the exact Kohn–Sham functional were used. To this point, no mention has been made of the concept of the virtual space. We now ask the same question as SAPT itself asks: how do these monomers interact with one another? Further, we will retain the labels of the localized occupied quasiparticles in computing this interaction, yielding F-SAPT. In doing so, we will arrive at the F-SAPT partition of the SAPT equations, which do contain references to the virtual space. The virtual space appears here as a mechanism to provide for the response of one monomer’s wave function in a perturbation (usually from the other monomer). However, here we are interested only in which portion of the monomer the electron was localized in before the interaction occurred (the functional group of origin), not to where it excited to after the interaction was invoked. In this view, the virtual space is not relevant as anything other than a mathematical construction to allow for response interactions, i.e., as a medium through which the occupied quasiparticles interact. Further evidence for this choice is the fact that the local occupied orbitals and nuclei directly outline the intuitive shapes of the chemical functional groups and are localizable and largely invariant to basis set size or type. By contrast, the virtual space is intrinsically nonlocalizable, particularly in the complete-basis or plane-wave limits. All of these considerations point to the intuitive validity of our selection of local occupied quasiparticles in designing F-SAPT.

At this point, it is imperative to remember that (1) any partition of the interaction energy terms into spatially localized contributions will not correspond to any set of observable quantities and (2) any partition of interaction energy terms will necessarily not be unique. Such partitions are intended to aid our chemical understanding of noncovalent contacts but should not be mistaken for absolute physical realism. We have designed F-SAPT to correspond to chemical intuition while using only minimal prescription (it requires only a localization metric and a human-specified set of functional groups, together with the picture of interactions between localized occupied quasiparticles discussed above), hopefully resulting in a robust, reliable partition. See section II.E on pages 10–11 of our A-SAPT paper<sup>17</sup> for an additional discussion of the related ambiguities present in A-SAPT.

### 3. RESULTS

**3.1. Computational Details.** All results presented in this work use A-SAPT0 and F-SAPT0 implementations in a development version of the PSI4 package,<sup>53</sup> largely according to the methodology discussed in our original A-SAPT0 paper. The dimer-centered jun-cc-pVDZ orbital basis set is used for all computations.<sup>38,39</sup> The density fitting approach is used to accelerate all Hartree–Fock<sup>52</sup> and SAPT0<sup>40,54,55</sup> computations, with the jun-cc-pVDZ-JKFIT auxiliary basis set used for the Hartree–Fock wave functions, electrostatics, exchange, induction, and one-particle contributions to the exchange-dispersion terms, while the jun-cc-pVDZ-RI auxiliary basis set is used for the two-body contributions to the dispersion and exchange-dispersion terms. The frozen core treatment is invoked in the dispersion term, freezing 1s electrons on first-row atoms and 1s, 2s, and 2p electrons on second-row atoms. For A-SAPT, atomic assignment of the electronic density is accomplished through the ISA procedure, using large Treutler–Ahlrichs grids with 100 radial points and 302 angular points.<sup>56,57</sup> Scalings for infinite-order exchange, induction coupling, and higher-order induction are applied as recommended in our previous work.

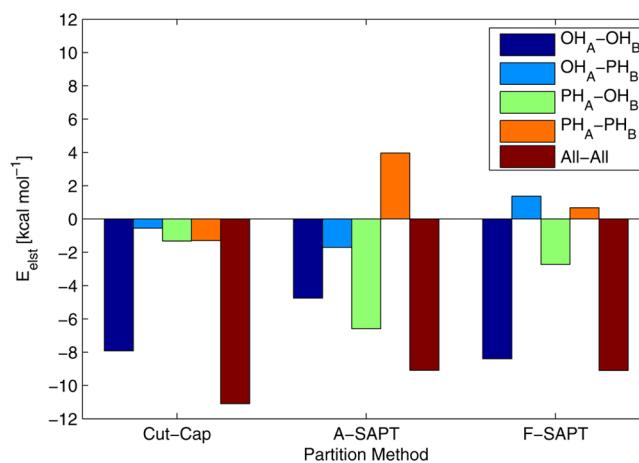
In contrast to our previous paper, we use the intrinsic bond orbital (IBO) methodology developed by Knizia to provide local occupied orbitals and orbital atomic charges.<sup>58</sup> In our previous work, these were provided by Pipek–Mezey local orbitals<sup>59</sup> and a form of ISA orbital charges, respectively. IBO is an extremely elegant Pipek–Mezey-like approach that is essentially invariant to basis set quality (a major problem with Mulliken-charge-based Pipek–Mezey). IBO methodology has very recently been shown<sup>60</sup> to be essentially equivalent to (albeit numerically simpler than) the quasiatomic minimal basis set (QUAMBO) methodology of Ruedenberg and co-workers.<sup>61</sup> As in Knizia’s work, we use the fully contracted cc-pVTZ core and valence functions<sup>62</sup> to define the minimal atomic orbitals and use the fourth-power metric in the localization procedure. Separate localizations are performed for the core and valence orbitals (after the unified set of intrinsic atomic orbitals are generated), to prevent mixing of core and valence spaces in the frozen-core dispersion treatment. Minimal AOs are placed only on the active monomer: ghost atoms do not receive minimal AOs.

For cut-and-cap analyses, constrained optimizations for capping hydrogens are performed at the B970-D2/aug-cc-pVDZ<sup>63,64</sup> level in Q-Chem 4.1,<sup>65</sup> with monomer-centered basis sets. SAPT0/jun-cc-pVDZ is used for the pairwise cut-and-cap electronic structure computations, as is consistent with

the A-SAPT and F-SAPT analyses. The H<sub>2</sub> particle-count corrections are not applied in this work.

Dimer geometries, cut-and-cap fragment geometries, and complete cut-and-cap, A-SAPT, and F-SAPT analyses for all systems and SAPT terms are provided in the Supporting Information.<sup>66</sup>

**3.2. Phenol Dimer Example.** Figure 3 depicts various decompositions of the electrostatic energy for the framing



**Figure 3.** Phenol dimer SAPT0/jun-cc-pVDZ electrostatic energy partitioned by Cut-Cap, A-SAPT, and reduced F-SAPT schemes.

phenol dimer example of the theory section, partitioned by the cut-and-cap fragmentation method, and by the A-SAPT and reduced F-SAPT many-body partitioning techniques. The most immediate feature of this illustration is that A-SAPT provides a significantly weaker hydrogen bonding contribution than one might expect (vs say that predicted by cut-and-cap analysis), due to the obfuscating influence of partial charges on the ISA atom-based fragments. The dipole–dipole terms manage to overcome the  $\delta^- \cdots \delta^-$  charge–charge repulsion to produce a net attraction between the two hydroxyl groups, but the magnitude of this attraction is much weaker than expected. In fact, the A-SAPT Ph<sub>A</sub>···OH<sub>B</sub> interaction is significantly more attractive than that of OH<sub>A</sub>···OH<sub>B</sub>, due to the spurious  $\delta^+ \cdots \delta^-$  charge–charge attraction in the former. The Ph<sub>A</sub>···Ph<sub>B</sub> electrostatic interaction is also poorly described by A-SAPT, with substantial repulsion arising from the  $\delta^+ \cdots \delta^+$  charge–charge term.

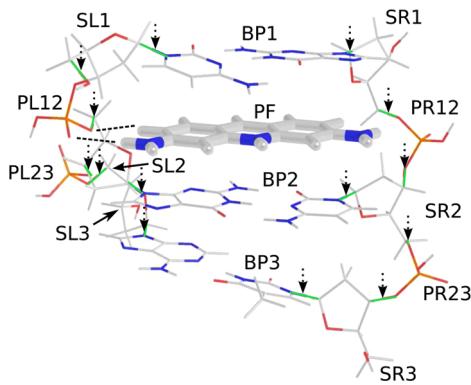
The second key feature of this figure is the strong similarity between the cut-and-cap partition and the reduced F-SAPT partition. Both techniques clearly indicate the OH<sub>A</sub>···OH<sub>B</sub> hydrogen bonding, with electrostatic contributions of  $-7.9$  kcal mol<sup>-1</sup> and  $-8.4$  kcal mol<sup>-1</sup>, respectively. The next strongest contributor is Ph<sub>A</sub>···OH<sub>B</sub> interaction, with electrostatic contributions of  $-1.3$  kcal mol<sup>-1</sup> and  $-2.7$  kcal mol<sup>-1</sup>, respectively. Note that the total SAPT0/jun-cc-pVDZ electrostatic energy for this system is  $-9.1$  kcal mol<sup>-1</sup>, while the cut-and-cap electrostatic energy is  $-11.1$  kcal mol<sup>-1</sup>, indicating that isolated pairwise computations induce an error of  $2.0$  kcal mol<sup>-1</sup> in cut-and-cap analysis. As the differences between the reduced F-SAPT and cut-and-cap electrostatic partitions are at most  $2.0$  kcal mol<sup>-1</sup>, we can conclude the F-SAPT and cut-and-cap analyses agree to within the maximum resolution of the latter method.

The link bond contributions to the full F-SAPT electrostatic partition are not entirely negligible but never exceed a magnitude of  $1.5$  kcal mol<sup>-1</sup>, e.g., much smaller than the total

$\text{OH}_A \cdots \text{OH}_B$  interaction. Moreover, partitioning the link bonds by IBO orbital charges vs equal assignment results in a maximum deviation of only  $0.34 \text{ kcal mol}^{-1}$  in the reduced F-SAPT partition.

The results obtained for this case study are quite encouraging. F-SAPT produces chemically intuitive results for the electrostatic contributions for this system (in contrast to A-SAPT) and also returns the many-body SAPT0 limit (in contrast to cut-and-cap). Additionally, the ambiguities associated with link-bond assignment in reduced F-SAPT are far smaller in magnitude than the key contributions to the electrostatics partition. We have observed similar correspondence in many other chemical systems and in the other (easier) SAPT terms.

**3.3. Proflavine Intercalation Example.** We now turn our attention to a particularly problematic and motivating case encountered in our A-SAPT paper: the intercalation of a proflavine cation into a double-stranded CGA trinucleotide fragment of DNA, including the sugar–phosphate backbone. The geometry and chemical fragmentation pattern for this system are depicted in Figure 4. The geometry<sup>67</sup> is the same as



**Figure 4.** Proflavine<sup>+</sup>–DNA intercalation complex geometry and functional group partitioning scheme. Link bonds are indicated by solid green lines and emphasized by downward-pointing dotted arrows. Intermolecular hydrogen bonds are indicated by dashed black lines.

that used in our 2011 DF-SAPT0 paper<sup>40</sup> and our A-SAPT paper.<sup>17</sup> Note that we have protonated each phosphate group to simulate the presence of a counterion, resulting in a neutral DNA monomer. The proflavine cation carries a +1 net charge. The DNA fragment is labeled according to the nucleotide on the left side of Figure 4, going from 5' to 3' (top to bottom). The proflavine has intercalated between the C and G steps in this ordering, causing significant distortion of the DNA structure. We label the base pairs as BP1 (C), BP2 (G), and BP3 (A), from top to bottom. The sugars are labeled according to their side and neighboring base pair; e.g., SL1 is the left-hand sugar bonded to BP1. The phosphates are labeled by their side and the two base pairs they link; e.g., PL12 is the left-hand phosphate bridging BP1 and BP2, via SL1 and SL2. Sugar–base pair functional groups are separated by cutting glycosidic C–N bonds, while sugar–phosphate functional groups are separated by cutting C–O bonds.

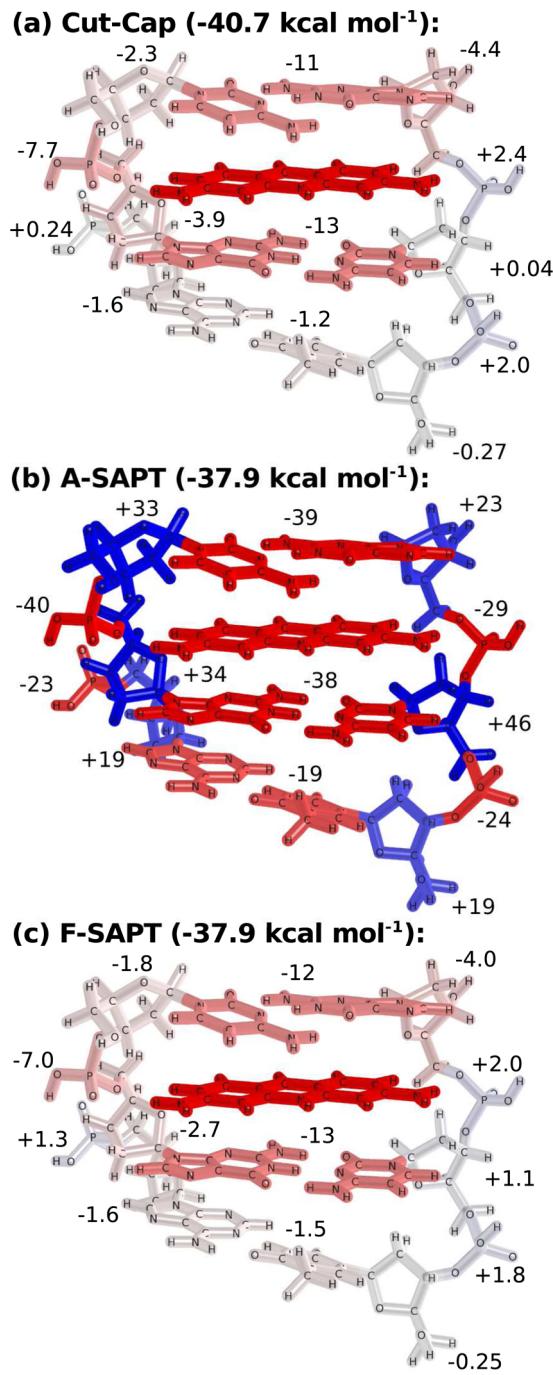
This is an interesting test case due to the complex interplay between various contributions to the intermolecular interaction: in our 2011 paper,<sup>40</sup> we performed a coarse SAPT0/jun-cc-pVDZ cut-and-cap analysis (to the level of base pairs and backbones) and determined that (1) roughly 2/3 of the

interaction strength comes from  $\pi$ -stacking between the proflavine and neighboring base pairs, (2) the next-nearest neighbor base pair contributes negligibly to the interaction, and (3) roughly 1/3 of the interaction strength comes from interactions with the closer left-side backbone, possibly including long hydrogen bonds forming between the proflavine and left upper phosphate. In our A-SAPT paper, the first and second findings were easily visible in order-1 atomic visualizations of the dispersion and exchange terms, which clearly showed the nearest-neighbor  $\pi$ -stacking interactions, and negligible interactions with the next-nearest neighbor base pair. However, the third finding was much harder to elucidate, due to the sizable oscillations in the order-1 atomic electrostatics visualizations induced by the uneven assignment of bonds linking adjacent atoms of the DNA fragments, under the strong driving field of the proflavine cation. By differentially increasing the color scales in the order-1 atomic visualizations, or looking through the extensive order-2 atom-pairwise data, it was apparent that PL12 was the most active phosphate in the electrostatic interaction, but this was, at best, a qualitative conclusion.

Here, we revisit this problem by quantitatively partitioning the electrostatic energy to base pair, sugar, and phosphate contributions via cut-and-cap, A-SAPT, and F-SAPT analyses. Note that the proflavine cation is treated as a single functional group, so the order-1 and order-2 analyses are equivalent for this system. This allows us to use visualization techniques adapted from A-SAPT to quickly depict the relevant interactions in the system: Figure 5 shows the qualitative and quantitative results from these analyses. In each panel, the functional groups are colored according to the sign and magnitude of their electrostatic interaction with the other monomer, with red indicating attraction and blue indicating repulsion. The small numbers in the figure indicate the quantitative electrostatic contributions, in  $\text{kcal mol}^{-1}$ . The color scales saturate at  $\pm 30 \text{ kcal mol}^{-1}$  to provide adequate resolution for the smaller contributions. This means that the proflavine contributions will be saturated in all three analyses, as will some of the DNA contributions in the A-SAPT analysis.

Just as with the phenol dimer, the quantitative A-SAPT electrostatic partition does not provide much insight. The sugars have large repulsive electrostatic interactions, while the base pairs and phosphates have large attractive electrostatic interactions. This is easily explained when one considers that the link bonds involving sugars are either C–N or C–O, with the C on the sugar: the electron pairs in these polar covalent bonds will be assigned more preferentially to the N or O atoms of the base pairs or phosphates, respectively. This means that the sugars will attain net positive charges, while the base pairs and phosphates will attain net negative charges. The charge–charge interactions between these partially charged functional groups and the +1 charged proflavine cation strongly dominate the quantitative A-SAPT partition. Within the base pairs and sugars, it is apparent that the nearest neighbor base pairs and PL12 phosphate have the strongest electrostatic interactions and that the next nearest neighbor base pair is less involved. However, as was the case with order-1 atomic visualizations in A-SAPT, this is, at best, a qualitative conclusion: there are oscillating contributions of  $> \pm 10 \text{ kcal mol}^{-1}$  from all functional groups, and the strongest overall contributor is actually SR2, with a contribution of  $+46 \text{ kcal mol}^{-1}$ .

The invocation of reduced F-SAPT markedly cleans up the electrostatic partition, yielding results which strongly resemble



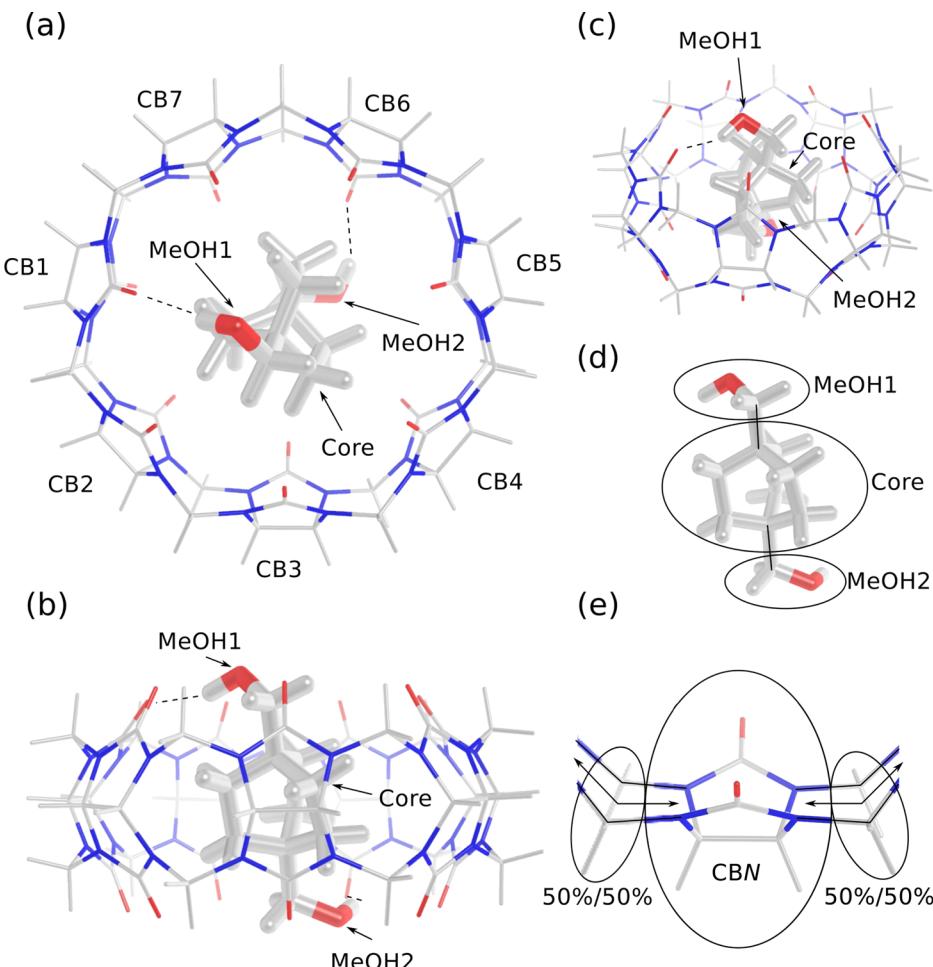
**Figure 5.** Proflavine<sup>+</sup>–DNA SAPTO/jun-cc-pVDZ electrostatic energy partitioned to functional groups via (a) Cut-Cap, (b) A-SAPT, and (c) reduced F-SAPT schemes. Red indicates attraction. Blue indicates repulsion. White indicates no net interaction. The color scales saturate at  $\pm 30$  kcal mol<sup>-1</sup>. Small numbers indicate functional group electrostatic contributions in kcal mol<sup>-1</sup>.

cut-and-cap fragmentation, while reproducing the full SAPTO electrostatic energy. Quoting the F-SAPT partition, the rough conclusions from our DF-SAPTO cut-and-cap analysis are immediately ratified, along with additional detail. The first finding above is apparent in the substantial  $-12$  and  $-13$  kcal mol<sup>-1</sup> electrostatic contributions of BP1 and BP2, respectively, indicating an interesting cation– $\pi$  and  $\pi$ -stacking electrostatic interaction between the proflavine and these two neighboring base pairs. The second finding above is apparent from the small

contributions of PL23, PR23, SL3, SR3, and BP3, which individually contribute no more than  $\pm 1.8$  kcal mol<sup>-1</sup> to the electrostatic energy and collectively contribute a negligible  $-0.25$  kcal mol<sup>-1</sup> to the electrostatic energy. The third finding above is also visible, with some enhancement provided by F-SAPT. With an electrostatic contribution of  $-7.0$  kcal mol<sup>-1</sup>, PL12 is clearly the strongest contributor from the backbone and is the strongest overall contributor after the nearest neighbor base pairs. This supports our intuitive conclusion of long hydrogen bonding between proflavine and PL12 from the 2011 paper. However, it is interesting to note that there are other non-negligible contributors from the two backbones, e.g., SR1 ( $-4.0$  kcal mol<sup>-1</sup>) and SL2 ( $-2.7$  kcal mol<sup>-1</sup>). Both of these sugars are oriented so that the lone pairs of the ring oxygen are facing the proflavine cation, providing favorable electrostatic orientations. Additionally, SL1 is distorted so that the terminal hydroxyl group is oriented favorably with respect to the proflavine cation; while the ring oxygen is not favorably aligned for this sugar, the hydroxyl group is able to provide for a net favorable contribution of  $-1.8$  kcal mol<sup>-1</sup>. The contributions from SL1, PL12, and SL2 all add to the attractive contribution of the left-side backbone, indicating that while long hydrogen bonds with PL12 are the dominant contributors, the nearby sugars of the left-side backbone also contribute to the overall stability of the complex. Moreover, the favorable electrostatic contribution from SR1 is largely canceled by unfavorable contributions from PR12, SR2, etc. This implies that our original conclusion of negligible interactions between proflavine and the right-side backbone only holds in the net: nontrivial interactions do occur with various units of the right-side backbone, but these nearly cancel overall.

Note that the total electrostatic energy at SAPTO/jun-cc-pVDZ is  $-37.9$  kcal mol<sup>-1</sup>, while the pairwise cut-and-cap electrostatic energy is  $-40.7$  kcal mol<sup>-1</sup>, i.e., a many-body error of  $-2.8$  kcal mol<sup>-1</sup>. In particular, the cut-and-cap contributions are generally slightly more attractive than their F-SAPT counterparts, which is possibly due to the greater polarization of capping C–H, O–H, and N–H  $\sigma$  bonds in the cut-and-cap analysis. F-SAPT is able to avoid this error while producing a chemically useful partition.

**3.4. Cucurbituril Inclusion Example.** To demonstrate the insight provided by a complete F-SAPT analysis, we consider a cucurbit[7]uril... bicyclo[2.2.2]octane-1,4-dimethanol inclusion complex (geometry obtained from ref 29, originally proposed in ref 68), depicted in Figure 6. Inclusion complexes involving cucurbituril hosts and polar or cationic guest molecules are of great interest in supramolecular chemistry, due to their enormous binding constants and amazing specificities. Myriad inclusion complexes involving neutral and cationic ferrocene derivatives,<sup>68–71</sup> cationic adamantyl derivatives,<sup>70</sup> and neutral and cationic bicyclo[2.2.2]octane derivatives<sup>68</sup> as guests have been reported in both experimental and theoretical studies. The proposed origins for such strong binding affinities are the complementary shapes of the host and guest, which lead to strong noncovalent enthalpic contributions, and to a remarkably small entropic penalty in the inclusion complex, arising from the rigidity of the constituent monomers. Cationic guests are found to have somewhat higher binding affinities, presumably due to favorable electrostatic interactions between the cationic side substituents of the guest and the rings of carbonyl groups in the cucurbituril (though this favorability is somewhat diminished by the corresponding desolvation penalty). The complex considered in the present study does



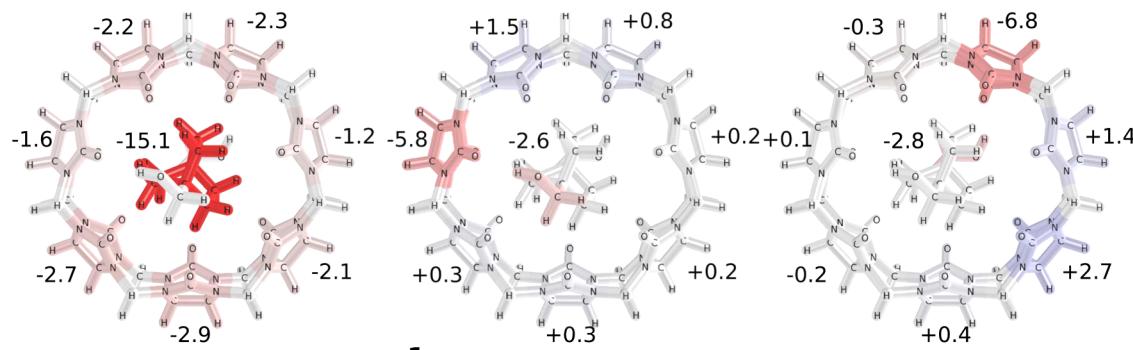
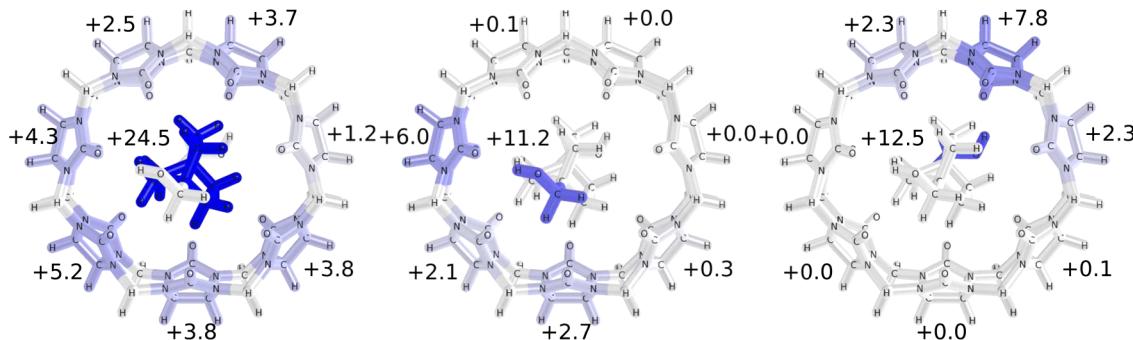
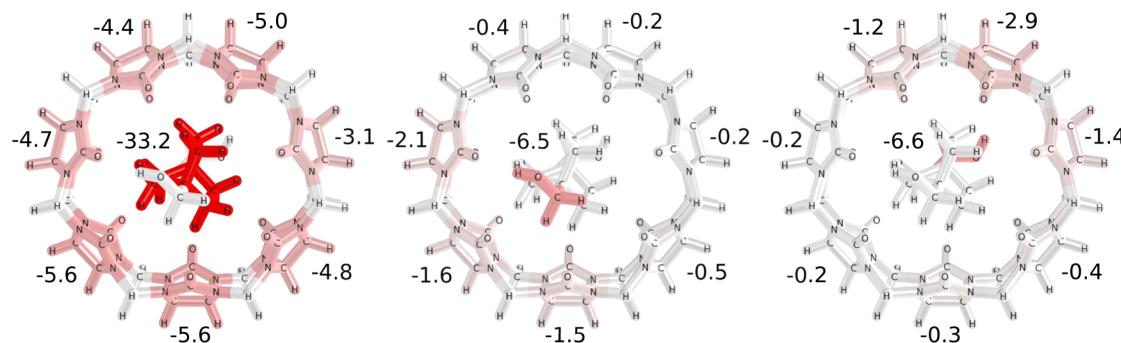
**Figure 6.** Cucurbit[7]uril···bicyclo[2.2.2]octane-1,4-dimethanol inclusion complex geometry and functional group partitioning scheme. (a) Top view. (b) Side view. (c) Birds-eye view. (d) Bicyclo[2.2.2]octane-1,4-dimethanol functional group partition. (e) Cucurbit[7]uril functional group partition. Ovals denote functional group selection. Link bonds are indicated by solid black lines. Intermolecular hydrogen bonds are indicated by dashed black lines.

not have a cationic guest but rather features a pair of polar, neutral methanol groups attached to a bicyclo[2.2.2]octane core. This guest was proposed in the mining minima algorithm (M2) theoretical study of Gilson and co-workers<sup>68</sup> and was further investigated by the noncovalent interactions index (NCI) of Yang and co-workers.<sup>29</sup> The NCI visualization heuristic of the latter was able to detect strong van der Waals interactions throughout the inclusion cavity (presumably indicating dispersion between the host and the core of the guest) and additionally indicated that “the hydroxyl substituents on the guest establish strong hydrogen bonds with two carbonyls of the [cucurbituril] host.”

The functional-group partition selected for this complex is depicted in Figure 6. For the guest, we divide the monomer into a bicyclo[2.2.2]octane core (Core) and two methanol groups (MeOH1 and MeOH2). We alternatively could have partitioned to the level of hydroxyl groups and obtained similar results; the main deviations between these choices of partition are small exchange and dispersion contributions from the CH<sub>2</sub> links in the methanols. For the host, we use a slightly more exotic partition to simplify the analysis. We first perform a reduced F-SAPT analysis with 14 host functional groups: the seven uniquely identifiable cucurbituril units and the seven sets of linking CH<sub>2</sub> pairs. Then, we assign the (minor) contributions from the CH<sub>2</sub> pairs evenly to their adjacent cucurbituril units

(CBN). This “collapsed F-SAPT” scheme allows us to effectively slice the linking CH<sub>2</sub> pairs down the middle. Figures 7 and 8 depict the quantitative results from the collapsed F-SAPT partition, for all SAPT terms. Each row of these figures depicts an individual SAPT term, while each column within the row depicts the contributions to the core, MeOH1, and MeOH2 portions of the guest, from left to right. The colorings of the cucurbituril units correspond to their contributions from the current term and portion of the guest, while the colorings of the guest units correspond to their contributions from the current term and all cucurbituril units. The color scales saturate at  $\pm 20$  kcal mol<sup>-1</sup>, implying that the total core contributions will sometimes be saturated.

The obvious conclusion from the collapsed F-SAPT analysis is the large and favorable interaction between the host and guest core, which contributes  $-28.7$  kcal mol<sup>-1</sup> of the net interaction strength of  $-28.2$  kcal mol<sup>-1</sup>. The major contributors to this interaction are electrostatics ( $-15.1$  kcal mol<sup>-1</sup>), exchange ( $+24.5$  kcal mol<sup>-1</sup>), and especially dispersion ( $-33.2$  kcal mol<sup>-1</sup>). The dispersion contribution is quite isotropic throughout the cucurbituril, while the electrostatic and exchange interactions are much more anisotropic. The contributions of CB5 are the smallest overall, corresponding to CB5 being the most geometrically distant cucurbituril unit from the guest. The always attractive and strongly distance-

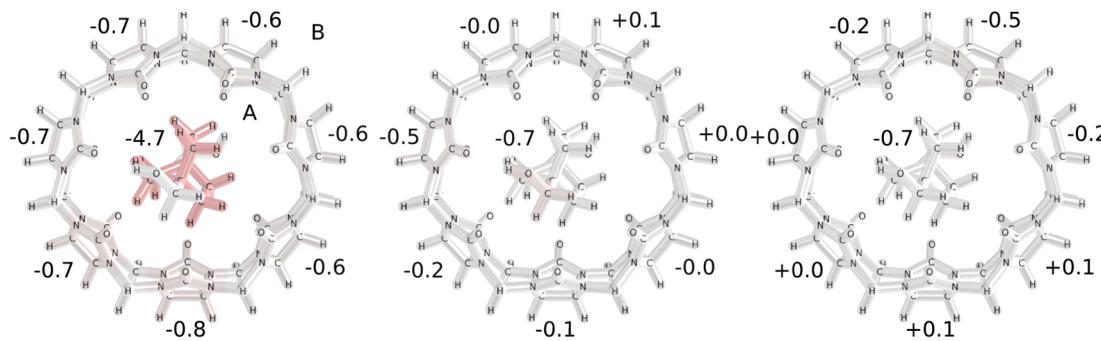
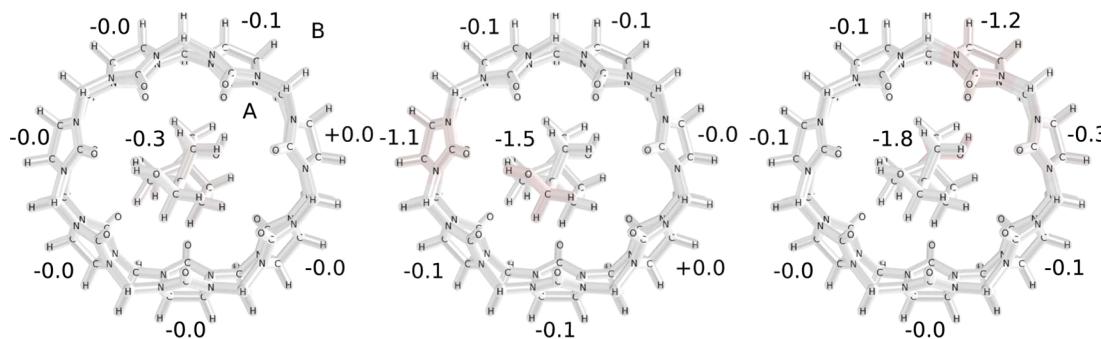
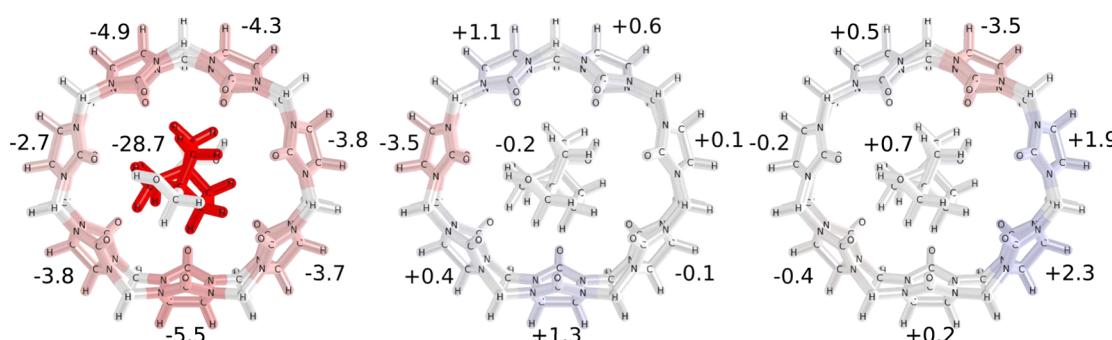
**Electrostatics (-20.4 kcal mol<sup>-1</sup>):****Exchange (+48.2 kcal mol<sup>-1</sup>):****Dispersion (-46.2 kcal mol<sup>-1</sup>):**

**Figure 7.** Cucurbit[7]uril···bicyclo[2.2.2]octane-1,4-dimethanol inclusion complex SAPTO/jun-cc-pVDZ electrostatic, exchange, and dispersion terms partitioned by collapsed F-SAPT scheme discussed in the main text. Left: contributions to Core. Middle: contributions to MeOH1. Right: contributions to MeOH2. Red indicates attraction. Blue indicates repulsion. White indicates no net interaction. The color scales saturate at  $\pm 20$  kcal mol<sup>-1</sup>. Small numbers indicate functional group electrostatic contributions in kcal mol<sup>-1</sup>.

dependent electrostatic interaction between the host and nonpolar guest core is strongly indicative of a favorable charge penetration interaction (though a distributed multipole analysis would be required to quantitatively prove this claim). The induction contribution from the cucurbituril polarizing the core is surprisingly large, with a net contribution of  $-4.7$  kcal mol<sup>-1</sup>. Each polar cucurbituril contributes a small  $-0.6$  to  $-0.7$  kcal mol<sup>-1</sup> fraction, but these add up to a substantial total. The induction term for the nonpolar guest core polarizing the cucurbituril is quantitatively negligible ( $-0.3$  kcal mol<sup>-1</sup>), as expected.

The surprising finding from F-SAPT is the quantitatively negligible net contribution from the methanol groups ( $-0.2$  kcal mol<sup>-1</sup> and  $+0.7$  kcal mol<sup>-1</sup> for MeOH1 and MeOH2, respectively). Evidence of hydrogen bonding with the nearest carbonyl units is present in favorable net contributions in the electrostatics, exchange, induction ( $B \leftarrow A$ ), and dispersion

terms for MeOH1···CB1 and MeOH2···CB6. However, this is essentially wholly canceled by unfavorable interactions in electrostatics and exchange from other cucurbituril units. The unfavorable electrostatic contributions are strongest on the next two cucurbituril units clockwise from the closest contact and probably arise from a clash between the hydroxyl and carbonyl lone pairs. The unfavorable exchange contributions appear counterclockwise from the contact in MeOH1 and on the two adjacent cucurbituril units to the contact in MeOH2, and thus are more dependent on the distance between the host and MeOH unit than was the case for electrostatics. Note that some of the unfavorable exchange contributions can be attributed to the linking CH<sub>2</sub> units in the methanols. However, repeating the collapsed F-SAPT analysis with partitioning at the level of hydroxyl groups and bicyclo[2.2.2]octane-1,4-CH<sub>2</sub> indicates that this is a minor effect. In the latter analysis, the two hydroxyls contribute total interaction strengths of  $-1.8$  kcal<sup>-1</sup>

**Induction ( $\mathbf{A} \leftarrow \mathbf{B}$ ) (-6.1 kcal mol $^{-1}$ ):****Induction ( $\mathbf{B} \leftarrow \mathbf{A}$ ) (-3.6 kcal mol $^{-1}$ ):****Total SAPTO (-28.2 kcal mol $^{-1}$ ):**

**Figure 8.** Cucurbit[7]uril···bicyclo[2.2.2]octane-1,4-dimethanol inclusion complex SAPTO/jun-cc-pVDZ induction and total terms partitioned by collapsed F-SAPT scheme discussed in the main text. Left: contributions to Core. Middle: contributions to MeOH1. Right: contributions to MeOH2. Red indicates attraction. Blue indicates repulsion. White indicates no net interaction. The color scales saturate at  $\pm 20$  kcal mol $^{-1}$ . Small numbers indicate functional group electrostatic contributions in kcal mol $^{-1}$  h.

and -0.8 kcal $^{-1}$  overall, e.g., much less than one hydrogen bond each. Thus, it is apparent that the presence of neutral, polar methanol, or hydroxyl units do not quantitatively stabilize the complex: a favorable dipole–dipole interaction does occur with the carbonyl of the nearest cucurbituril unit, but this is canceled by unfavorable electrostatic and exchange interactions with other cucurbituril units. This may help to explain the enhanced binding affinities generally observed with cationic guests. A guest with cationic side substituents can adopt a geometry in which the cations are in the middle of the ring of carbonyls on both sides of the cucurbituril, interacting favorably with all carbonyls. This is not possible for neutral side substituents possessing a dipole oriented in the plane of the ring of carbonyls, for which a single favorable dipole–dipole interaction will be largely canceled by unfavorable dipole–dipole interactions with the rest of the ring. An additional, more subtle effect may also be relevant: the symmetry-breaking hydrogen-bonded association of a neutral, polar side substituent

to any one cucurbituril unit presents a barrier to free rotation of the guest within the host. This barrier to rotation may present an entropic penalty relative to a guest with a charged side substituent which does not symmetry break.

It is worth noting that the collapsed cut-and-cap, A-SAPT, and F-SAPT analyses for this system (presented in the Supporting Information) all produce quite similar partitions for this system. For A-SAPT, this is serendipitous: the raw A-SAPT partition summed to the level of CB units and linking CH<sub>2</sub> pairs shows significant oscillations in the electrostatics, but this is attenuated by the even assignment of the links to the neighboring CB units in the collapsed analysis. Moreover, the cut-and-cap analysis deviates from the total SAPTO limit by 2.3 kcal mol $^{-1}$ , with larger deviations in individual terms, e.g., 5.6 kcal mol $^{-1}$  in exchange.

## 4. CONCLUSIONS

In this work, we have detailed an “F-SAPT” modification to the A-SAPT partition, which yields an alternative decomposition in terms of pairs of chemical functional groups. A-SAPT provides a partition in the force-field-type picture of atom-pairwise contributions to the interaction energy and is therefore much more suited to the development of *ab initio* force fields than to a chemical understanding of complicated noncovalent interactions. In particular, summing A-SAPT atomic contributions to produce contributions in terms of chemical functional groups can provide nonintuitive results, especially for the electrostatics term, due to the presence of intrinsic charges in the A-SAPT functional groups. By contrast, F-SAPT is formulated entirely in terms of the chemically intuitive picture of fragments of neutral or integral charge via assignment of whole occupied electrons or protons to fragments, thus ameliorating the charge–charge contaminations present in the A-SAPT electrostatics partition. The F-SAPT partition shows a strong resemblance to the chemically intuitive results produced by cut-and-cap fragmentation but additionally respects the many-body limit of the SAPT approach. The F-SAPT partition is shown to be useful and robust over such complicated intermolecular interaction motifs as a phenol dimer, a proflavine<sup>+</sup>–DNA intercalation complex, and a cucurbituril host–guest inclusion complex. Additionally, the computational overhead of F-SAPT is very small, so F-SAPT should be tractable on any system for which SAPT is tractable. The largest F-SAPT0/jun-cc-pVDZ result shown in this work (proflavine<sup>+</sup>–DNA) has 228 atoms and 2845 basis functions. Moreover, we have implemented F-SAPT0 as a minor extension to our A-SAPT0 code, so both analyses may be obtained together.

This work and the preceding one form a powerful toolset for the applications chemist. A-SAPT provides both automatic extraction of atom-pairwise potentials for use in force field development and black box visualizations of the key contributions in noncovalent interactions. F-SAPT provides a more chemically robust assignment of the contributions of individual noncovalent contacts, for use in understanding and designing complex intermolecular interactions. Future improvements to the methodology include sensitivity testing over a broad range of interaction motifs and quasiparticle localization techniques, application to more-accurate SAPT levels, and development of a streamlined workflow for A-SAPT/F-SAPT computation and postprocessing. We should be careful to remember the A-SAPT/F-SAPT are neither unique nor experimentally observable partitions of the interaction energy components, but so far they seem to be providing robust and deep insight into the origins of complicated noncovalent contacts. Even at the current level of completeness, we are already beginning to use A-SAPT/F-SAPT to probe the origins of intermolecular interactions in drug–ligand interactions, organic materials stacking, dispersion-assisted hydrogen bonding, and substituent effects in stacked benzene dimers.

## ■ ASSOCIATED CONTENT

### Supporting Information

Molecular geometries and detailed numerical results for all partitions applied in this work. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: sherrill@gatech.edu.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

R.M.P. is a DOE Computational Science Graduate Fellow (Grant DE-FG02-97ER25308). This work was supported by the National Science Foundation (Grant No. ~CHE-1300497). The Center for Computational Molecular Science and Technology is funded through an NSF CRIF award (Grant No. CHE-0946869) and by Georgia Tech.

## ■ REFERENCES

- (1) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.
- (2) Cozzi, F.; Cinquini, M.; Annuziata, R.; Dwyer, T.; Siegel, J. S. *J. Am. Chem. Soc.* **1992**, *114*, 5729–5733.
- (3) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2003**, *107*, 8377–8379.
- (4) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690–7697.
- (5) Ringer, A. L.; Sinnokrot, M. O.; Lively, R. P.; Sherrill, C. D. *Chem.—Eur. J.* **2006**, *12*, 3821–3828.
- (6) Wheeler, S. E.; Houk, K. N. *J. Am. Chem. Soc.* **2008**, *130*, 10854–10855.
- (7) Ringer, A. L.; Sherrill, C. D. *J. Am. Chem. Soc.* **2009**, *131*, 4574–4575.
- (8) Hohenstein, E. G.; Duan, J.; Sherrill, C. D. *J. Am. Chem. Soc.* **2011**, *133*, 13244–13247.
- (9) Wheeler, S. E. *J. Am. Chem. Soc.* **2011**, *133*, 10262–10274.
- (10) Raju, R. K.; Bloom, J. W. G.; Wheeler, S. E. *J. Chem. Theory Comput.* **2013**, *9*, 3479–3490.
- (11) Dunning, T. H. *J. Phys. Chem. A* **2000**, *104*, 9062–9080.
- (12) Arnstein, S. A.; Sherrill, C. D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2646–2655.
- (13) Wheeler, S. E.; Bloom, J. W. G. *J. Phys. Chem. A* **2014**, *118*, 6133–6147.
- (14) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (15) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325–340.
- (16) Sherrill, C. D. *Acc. Chem. Res.* **2013**, *46*, 1020–1028.
- (17) Parrish, R. M.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *141*, 044115.
- (18) Jeziorski, B.; Moszynski, R.; Ratkiewicz, A.; Rybak, S.; Szalewicz, K.; Williams, H. L. In *Methods and Techniques in Computational Chemistry: METECC94*; Clementi, E., Ed.; STEF: Cagliari, 1993; Vol. B (Medium-Size Systems), p 79.
- (19) Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994**, *100*, 2900–2909.
- (20) Khalilullin, R. Z.; Head-Gordon, M.; Bell, A. T. *J. Chem. Phys.* **2006**, *124*, 204105.
- (21) Khalilullin, R. Z.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 184112.
- (22) Azar, R. J.; Head-Gordon, M. *J. Chem. Phys.* **2012**, *136*, 024103.
- (23) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701–706.
- (24) Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2007**, *28*, 222–237.
- (25) van der Vaart, A.; Merz, K. M. *J. Phys. Chem. A* **1999**, *103*, 3321–3329.
- (26) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397–5403.
- (27) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683–686.
- (28) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

- (29) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.
- (30) de Silva, P.; Corminboeuf, C. *J. Chem. Theory Comput.* **2014**, *10*, 3745–3756.
- (31) Stone, A. *J. Chem. Phys. Lett.* **1981**, *83*, 233–239.
- (32) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047–1064.
- (33) Stone, A. *J. J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.
- (34) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1996.
- (35) Le Sueur, C. R.; Stone, A. J. *Mol. Phys.* **1994**, *83*, 293–307.
- (36) Misquitta, A. J.; Stone, A. J. *J. Chem. Phys.* **2006**, *124*, 024111.
- (37) Jurečka, P.; Sponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- (38) Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *133*, 014101.
- (39) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 3027–3034.
- (40) Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F. *J. Chem. Phys.* **2011**, *135*, 174107.
- (41) Lillestolen, T. C.; Wheatley, R. *J. Chem. Commun.* **2008**, 2008, 5909–5911.
- (42) Lillestolen, T. C.; Wheatley, R. *J. J. Chem. Phys.* **2009**, *131*, 144101.
- (43) Wheatley, R. J.; Gopal, A. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2087–2091.
- (44) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129–138.
- (45) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.
- (46) Verstraelen, T.; Ayers, P.; Speybroeck, V. V.; Waroquier, M. *Chem. Phys. Lett.* **2012**, *545*, 138–143.
- (47) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford: New York, 1990; International Series of Monographs on Chemistry, Vol. 22.
- (48) Whitten, J. L. *J. Chem. Phys.* **1973**, *58*, 4496–4501.
- (49) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *Int. J. Quantum Chem. Symp.* **1977**, *12*, 81.
- (50) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396–3402.
- (51) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
- (52) Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.
- (53) Turney, J. M.; Simonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556–565.
- (54) Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
- (55) Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 184111.
- (56) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.
- (57) Treutler, O.; Ahlrichs, R. *J. Chem. Phys.* **1995**, *102*, 346.
- (58) Knizia, G. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (59) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- (60) Janowski, T. *J. Chem. Theory Comput.* **2014**, *10*, 3085–3091.
- (61) Lu, W. C.; Wang, C. Z.; Schmidt, M. W.; Bytautas, L.; Ho, K. M.; Ruedenberg, K. *J. Chem. Phys.* **2004**, *120*, 2629–2637.
- (62) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (63) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (64) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (65) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J.; Besley, N. A.; Herbert, J. M.; Yeh Lin, C.; Van Voorhis, T.; Hung Chien, S.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerkens, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simonett, A. C.; Subotnik, J. E.; Lee Woodcock, H., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.
- (66) See Supporting Information.
- (67) Maehigashi, T.; Persil, O.; Hud, N. V.; Williams, L. D. Unpublished results, NDB ID: DD0103.
- (68) Moghaddam, S.; Inoue, Y.; Gilson, M. K. *J. Am. Chem. Soc.* **2009**, *131*, 4012–4021.
- (69) Jeon, W. S.; Moon, K.; Park, S. H.; Chun, H.; Ko, Y. H.; Lee, J. Y.; Lee, E. S.; Samal, S.; Selvapalam, N.; Rekharsky, M. V.; Sindelar, V.; Sobransingh, D.; Inoue, Y.; Kaifer, A. E.; Kim, K. *J. Am. Chem. Soc.* **2005**, *127*, 12984–12989.
- (70) Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P. Y.; Isaacs, L. *J. Am. Chem. Soc.* **2005**, *127*, 15959–15967.
- (71) Rekharsky, M. V.; Mori, T.; Yang, C.; Ko, Y. H.; Selvapalam, N.; Kim, H.; Sobransingh, D.; Kaifer, A. E.; Liu, S.; Isaacs, L.; Chen, W.; Moghaddam, S.; Gilson, M. K.; Kim, K.; Inoue, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20737–20742.