ACCOUNTS of chemical research

pubs.acs.org/accounts Article

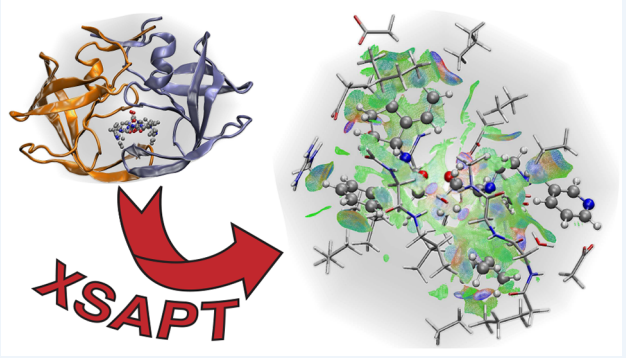
Predicting and Understanding Non-Covalent Interactions Using Novel Forms of Symmetry-Adapted Perturbation Theory

Kevin Carter-Fenk, Ka Un Lao, and John M. Herbert*

Cite This: *Acc. Chem. Res.* 2021, 54, 3679–3690 Read Online

ACCESS | Metrics & More | Article Recommendations

CONSPECTUS: Although sometimes derided as "weak" interactions, non-covalent forces play a critical role in ligand binding and crystal packing and in determining the conformational landscape of flexible molecules. Symmetry-adapted perturbation theory (SAPT) provides a framework for accurate *ab initio* calculation of intermolecular interactions and furnishes a natural decomposition of the interaction energy into physically meaningful components: semiclassical electrostatics (rigorously obtained from monomer charge densities), Pauli or steric repulsion, induction (including both polarization and charge transfer), and dispersion. This decomposition helps to foster deeper understanding of non-covalent interactions and can be used to construct transferable, physics-based force fields. Separability of the SAPT interaction energy also provides the flexibility to construct composite methods, a feature that we exploit to improve the description of dispersion interactions. These are challenging to describe accurately because they arise from nonlocal electron correlation effects that appear for the first time at second order in perturbation theory but are not quantitatively described at that level.



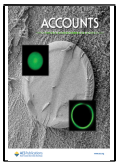
As with all quantum-chemical methods, a major limitation of SAPT is nonlinear scaling of the computational cost with respect to system size. This cost can be significantly mitigated using "SAPT0(KS)", which incorporates monomer electron correlation by means of Kohn–Sham (KS) molecular orbitals from density functional theory (DFT), as well as by an "extended" theory called XSAPT, developed by the authors. XSAPT generalizes traditional dimer SAPT to many-body systems, so that a ligand–protein interaction (for example) can be separated into contributions from individual amino acids, reducing the cost of the calculation below that of even supramolecular DFT while retaining the accuracy of high-level *ab initio* quantum chemistry.

This Account provides an overview of the SAPT0(KS) approach and the XSAPT family of methods. Several low-cost variants are described that provide accuracy approaching that of the best *ab initio* benchmarks yet are affordable enough to tackle ligand–protein binding and sizable host–guest complexes. These variants include SAPT+*ai*D, which uses *ab initio* atom–atom dispersion potentials ("+*ai*D") in place of second-order SAPT dispersion, and also SAPT+MBD, which incorporates many-body dispersion (MBD) effects that are important in the description of nanoscale materials. Applications to drug binding highlight the size-extensive nature of dispersion, which is not a weak interaction in large systems. Other applications highlight how a physics-based analysis can sometimes upend conventional wisdom regarding intermolecular forces. In particular, careful reconsideration of π–π interactions makes clear that the quadrupolar electrostatics (or "Hunter–Sanders") model of π–π stacking should be replaced by a "van der Waals model" in which conformational preferences arise from a competition between dispersion and Pauli repulsion. Our analysis also suggests that molecular shape, rather than aromaticity *per se*, is the key factor driving strong stacking interactions. Looking forward, we anticipate that XSAPT-based methods can play a role in screening of drug candidates and in materials design.

KEY REFERENCES

- Lao, K. U.; Herbert, J. M. Atomic orbital implementation of extended symmetry-adapted perturbation theory (XSAPT) and benchmark calculations for large supramolecular complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955–2978.[1] *This paper describes an efficient implementation of XSAPT and documents the breakdown of pairwise-additive dispersion for large molecules.*
- Carter-Fenk, K.; Lao, K. U.; Liu, K.-Y.; Herbert, J. M. Accurate and efficient *ab initio* calculations for supra-

Received: June 28, 2021
Published: September 22, 2021



https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679–3690

molecular complexes: Symmetry-adapted perturbation theory with many-body dispersion. *J. Phys. Chem. Lett.* **2019**, *10*, 2706−2714.[2] *This paper describes XSAPT +MBD, which is presently the most accurate variant of XSAPT.*

- Carter-Fenk, K.; Herbert, J. M. Electrostatics does not dictate the slip-stacked arrangement of aromatic $\pi-\pi$ interactions. *Chem. Sci.* **2020**, *11*, 6758−6765.[3] *This paper highlights an application of XSAPT to debunk the Hunter−Sanders model of π stacking and replace it with a van der Waals model.*
- Carter-Fenk, K.; Herbert, J. M. Reinterpreting π-stacking. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24870−24886.[4] *This work used XSAPT to address the question "does π stacking constitute a unique form of dispersion?". (The answer is yes, but the reason is molecular shape rather than aromaticity.)*

1. INTRODUCTION

Although sometimes regarded as "weak" forces, non-covalent interactions are pervasive in nature, and their cumulative influence has macroscopic manifestations. Johannes van der Waals was the first to connect these forces to the condensation of gases,[5] but we now understand that the role of non-covalent forces is not limited to causing deviations from ideal-gas behavior. In addition to governing condensed-phase properties ranging from water's density anomaly[6] to the structure of molecular crystals,[7] these forces play a critical role in organic chemistry[8] and in the three-dimensional structure of macromolecules.[9−12]

Non-covalent interaction energies are often computed *via* the supramolecular approach, by energy difference:

$$\Delta E_{\text{int}} = E_{\text{supra}} - \sum_{A=1}^{N} E_A \tag{1}$$

Here, E_{supra} is the energy of the supramolecular system and E_A is the energy of monomer A. Given the poor scaling of computational quantum chemistry with respect to system size, this approach is already expensive in systems of moderate size. When atom-centered basis sets are used, the supramolecular approach also introduces basis-set superposition error (BSSE),[13] leading to a dramatic overestimation of ΔE_{int} in small basis sets.[14]

Symmetry-adapted perturbation theory (SAPT)[15−17] is an alternative approach in which E_{int} for a dimer is computed perturbatively starting from isolated-monomer wave functions, thus avoiding BSSE. The terms in the perturbation expansion have straightforward physical interpretations,[18] furnishing an energy decomposition analysis.[16] The present work describes novel variants of SAPT that the authors have developed over the past decade.[1,2,19−30] These include an extension of traditional two-body SAPT that we call "XSAPT", which is intended to describe many-body polarization effects in clusters.[1,19−24] This is described in section 2 following a brief outline of traditional SAPT.

A key advantage of SAPT over energy decomposition analyses based on supramolecular density functional theory (DFT)[31] is a well-defined and predictive decomposition. (Dispersion interactions are especially difficult to isolate in DFT,[29,32,33] even for functionals with empirical dispersion corrections.[34]) Separability in SAPT can be exploited to replace low-order perturbation theory for dispersion, which is ill-behaved for large molecules,[1,35] with accurate and cost-effective alternatives.[1,2,23−26] These are discussed in section 3.

Section 4 provides illustrative applications highlighting the role that SAPT and its extensions can play in obtaining a qualitative understanding of non-covalent interactions, yet one that is firmly grounded in quantitative energetics. Additional examples of cases where textbook explanations for non-covalent phenomena fail under close scrutiny can be found in a recent exposé.[36]

2. SYMMETRY-ADAPTED PERTURBATION THEORY

2.1. SAPT0

SAPT partitions the supramolecular Hamiltonian into a zeroth-order part consisting of monomer Fock operators, plus Møller−Plesset fluctuation potentials (representing intramolecular electron correlation), and finally intermolecular Coulomb operators.[15−17] Taking the latter perturbation to second order, with Hartree−Fock (HF) wave functions used to describe the isolated monomers, results in a method called "SAPT0".[15,37] The corresponding partition of E_{int} is

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

Each term represents a physically meaningful contribution:[18,36]

- **Electrostatics:** $E_{\text{elst}}^{(1)}$ is the Coulomb interaction between the isolated-monomer charge densities.
- **Exchange:** $E_{\text{exch}}^{(1)}$ is the penalty for enforcing antisymmetry between monomer wave functions, which is known as *Pauli repulsion*.
- **Induction:** Polarization of the monomer charge densities is captured in

$$E_{\text{ind}} = E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)} + \delta E_{\text{HF}} \tag{3}$$

  where δE_{HF} represents a correction for higher-order induction, as described below. (For separation of induction into polarization and charge-transfer contributions, see refs 29 and 30.)
- **Dispersion:** As originally described by London,[38] dispersion arises at second order in perturbation theory,

$$E_{\text{disp}} = E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)} \tag{4}$$

  due to correlation-induced fluctuations in the mean-field electron densities of the monomers.

The terms $E_{\text{exch}}^{(1)}$, $E_{\text{exch-ind}}^{(2)}$, and $E_{\text{exch-disp}}^{(2)}$ arise from antisymmetrization of the monomer wave functions and serve to remove Pauli-forbidden contributions to $E_{\text{elst}}^{(1)}$, $E_{\text{ind}}^{(2)}$, and $E_{\text{disp}}^{(2)}$, respectively.

Polar molecules generally require a higher-order treatment of induction. Recognizing that a supramolecular HF calculation includes induction to all orders, the "δHF" correction in eqs 2 and 3 is defined as[15]

$$\delta E_{\text{HF}} = \Delta E_{\text{int}}^{\text{HF}} - (E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind,resp}}^{(2)} + E_{\text{exch-ind,resp}}^{(2)}) \tag{5}$$

where $\Delta E_{\text{int}}^{\text{HF}}$ is the HF interaction energy computed by difference (eq 1) and $E_{\text{ind,resp}}^{(2)}$ and $E_{\text{exch-ind,resp}}^{(2)}$ are the "response" (or "relaxed") analogues to the second-order induction and exchange−induction energies, computed *via* coupled-perturbed self-consistent field (SCF) equations.[39] The SAPT0 interaction

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

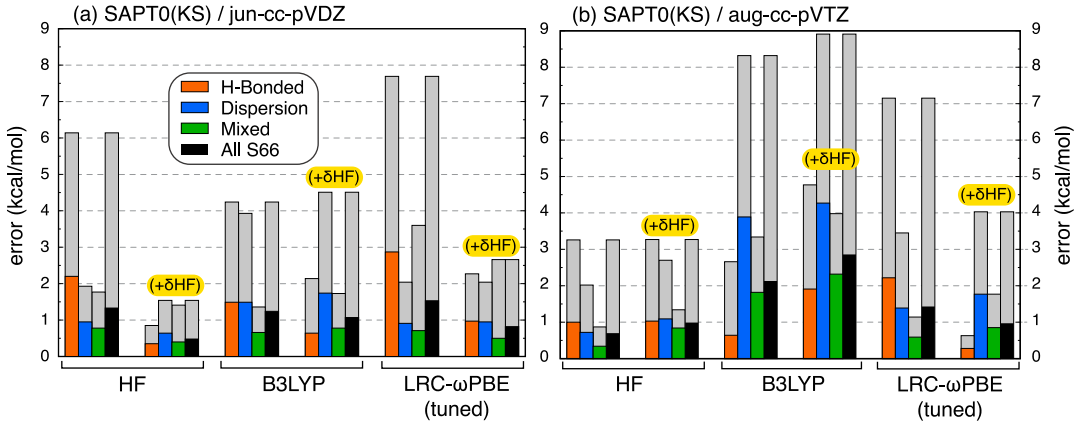


Figure 1. Performance of SAPT0(KS) versus CCSD(T)/CBS benchmarks for the S66 data set and subsets thereof using either (a) the jun-cc-pVDZ basis set or (b) the aug-cc-pVTZ basis set. Results from three different SCF methods are shown, with and without the δHF correction. Colored bars indicate mean absolute errors and gray bars indicate maximum errors. Data are from ref 28. The SAPT0(HF) approach is the method that is normally called simply "SAPT0".

energy in eq 2 can thus be recognized as essentially HF plus dispersion:

$$E_{\text{int}}^{\text{SAPT0}} \approx \Delta E_{\text{int}}^{\text{HF}} + E_{\text{disp}} \tag{6}$$

To avoid introducing BSSE, $\Delta E_{\text{int}}^{\text{HF}}$ should be computed in the dimer basis.

2.2. SAPT0(KS)

The use of Kohn–Sham (KS) molecular orbitals within the SAPT0 formalism affords a method that we have called SAPT0(KS),[28] which provides a low-cost means to incorporate intramolecular electron correlation. The results are somewhat disappointing, however, as demonstrated in Figure 1 using the S66 database,[40] where the benchmarks are coupled-cluster interaction energies [including single, double, and perturbative triple excitations, or CCSD(T)] extrapolated to the complete basis set (CBS) limit. Anomalously small KS gaps exacerbate problems with second-order dispersion, such that dispersion energies are significantly worse than HF-based results unless asymptotically correct density functionals are employed.[20,27,28] Correct asymptotic behavior can be enforced using long-range-corrected functionals with "tuned" range-separation parameters.[27,28] However, one should not simply insert orbitals from an arbitrary density functional into the SAPT0 formalism and expect good results!

In contrast to SAPT0(KS), DFT-SAPT[41,42] is a benchmark-quality method that exhibits $O(n^5)$ scaling with respect to system size, albeit with a much larger prefactor compared to SAPT0. In DFT-SAPT, the dispersion energy is computed from DFT density susceptibilities and combined with other energy components computed using traditional SAPT. Although the present work is focused on more cost-effective approaches, DFT-SAPT does represent a type of hybrid method similar to those described in section 3.

2.3. XSAPT

Polarization exhibits significant nonadditivity, whereas other energy components decay much faster with intermolecular separation and are pairwise-additive to a good approximation.[43] To extend pairwise SAPT to a many-body approach, it is therefore necessary to incorporatate nonadditive polarization, but it may be sufficient to describe other components of E_{int} in a pairwise-additive way. This is the assumption that underlies the extended SAPT (XSAPT) approach,[1,19−24] which combines pairwise-additive SAPT with monomer wave functions obtained from the "explicit polarization" (XPol) procedure.[44] XPol is a self-consistent charge-embedded SCF method that is used to compute monomer wave functions in the presence of classical atomic charges derived from the monomer wave functions themselves.[19,20] Charges derived from the molecular electrostatic potential,[20] or alternatively from modified Hirshfeld population analysis (so-called "CM5" charges),[26] have been used in this capacity. The monomer wave functions used in the subsequent pairwise SAPT calculation are polarized for the cluster environment, and many-body polarization effects are included implicitly.

The XSAPT *ansatz* for the interaction energy is[20,25]

$$E_{\text{int}}^{\text{XSAPT}} = \sum_{A,B>A} (E_{\text{elst}}^{AB} + E_{\text{exch}}^{AB} + E_{\text{disp}}^{AB} + E_{\text{exch-disp}}^{AB} + E_{\text{ind}}^{AB} + E_{\text{exch-ind}}^{AB} + \delta E_{\text{HF}}^{AB}) + E_{\text{pol}}^{\text{PW}} + E_{\text{pol}}^{\text{MB}} \tag{7}$$

The terms E_{elst}^{AB}, E_{exch}^{AB}, ... are the gas-phase SAPT0(KS) energy components for dimer AB, and the final two terms ($E_{\text{pol}}^{\text{PW}}$ and $E_{\text{pol}}^{\text{MB}}$) are pairwise and many-body contributions to the polarization energy. The pairwise contribution,

$$E_{\text{pol}}^{\text{PW}} = \sum_{A,B>A}^{N} [E_{AB}^{\text{XSAPT}}(AB) - E_{AB}^{\text{SAPT}}] \tag{8}$$

is computed by summing the differences between pairwise XSAPT (with charge embedding) and SAPT (isolated dimer) interaction energies. The many-body contribution is

$$E_{\text{pol}}^{\text{MB}} = \sum_{A,B>A}^{N} [E_{AB}^{\text{XSAPT}}(AB\cdots N) - E_{AB}^{\text{XSAPT}}(\text{AB})] \tag{9}$$

where the notation $E_{XY}^{\text{XSAPT}}(AB\cdots N)$ indicates that dimer XY is embedded in a charge field due to $AB\cdots N$.

3. ALTERNATIVE DESCRIPTIONS OF DISPERSION

Second-order perturbation theory, including SAPT0 but also Møller–Plesset theory (MP2), does not afford a quantitative treatment of dispersion. For example, MP2 overestimates the interaction energy of π-stacked $(C_6H_6)_2$ by a factor of 2, versus a 30% error for the T-shaped isomer whose dispersion energy is smaller.[45] For larger systems, MP2 errors are often considerably

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

larger.[1,35] For this reason, SAPT0 calculations are often performed in limited basis sets (typically jun-cc-pVDZ, as in Figure 1a),[37] relying on the slow convergence of the dispersion energy to mitigate the errors. A quantitative treatment of dispersion that does not rely on error cancellation is possible by including third-order terms in SAPT,[37] but at $O(n^7)$ cost.[15]

The most expensive terms in a SAPT0 calculation are $E^{(2)}_{\mathrm{disp}}$ and $E^{(2)}_{\mathrm{exch-disp}}$, which scale as $O(n^4)$ and $O(n^5)$, respectively. At least when high-quality basis sets are used, this leads to a peculiar situation in which the most expensive parts of a SAPT0 calculation are also the least accurate. Our strategy is to exploit the separability of the SAPT decomposition to replace second-order dispersion (eq 4) with alternative models that are both more accurate and less expensive. This leads to several $O(n^3)$ methods that are capable of quantitative (~1 kcal/mol) accuracy in many cases. These alternatives are described next.

3.1. *Ab Initio* Dispersion Potentials

Use of atom−atom potentials of the form $-C_6/R^6$ is well-established in DFT,[34,46] but these are *ad hoc* corrections that are intended to improve the description of nonbonded interactions, not quantities that should be interpreted as dispersion *per se*.[29,32] This complexity results from double-counting in the van der Waals region,[33,34] where both the semilocal exchange−correlation potential and the empirical correction are nonzero.

This ambiguity can be sidestepped by exploiting the inherent separability of the SAPT interaction energy to define *ab initio* dispersion (*ai*D) potentials that represent true dispersion.[1,23−25] A typical functional form, including not only dipole−dipole (C_6) but also dipole−quadrupole (C_8) dispersion, is

$$E^{aiD}_{\mathrm{disp}} = -\sum_{a\in A}\sum_{\substack{b\in B\\(B>A)}}\left[f_6(R_{ab})\frac{C_{6,ab}}{R^6_{ab}} + f_8(R_{ab})\frac{C_{8,ab}}{R^8_{ab}}\right] \tag{10}$$

where $a \in A$ and $b \in B$ are atoms on monomers A and B, respectively. The R^{-n}_{ab} terms are damped at short range (using functions f_n) to avoid singularities, and

$$C_{n,ab} = (C_{n,a}C_{n,b})^{1/2} \tag{11}$$

The parameters in eq 10 are the atomic C_6 and C_8 coefficients, $\{C_{6,a}\}$ and $\{C_{8,a}\}$.

Although eq 10 has the same functional form as that used in dispersion-corrected DFT,[34] separability of the SAPT interaction energy ensures that there is no double-counting. In the second-generation (*ai*D2)[25] and third-generation (*ai*D3)[24] versions of this approach, parameters were obtained by fitting to pure dispersion energies obtained from DFT-SAPT and from partial third-order SAPT.[24,25] (The first-generation method,[23] which was fit to reproduce total interaction energies, benefited from some error cancellation and is no longer recommended.[25]) The performance of SAPT+*ai*D2 and SAPT+*ai*D3 is quite similar, but the training set for the latter was somewhat expanded in order to reduce errors for $\pi-\pi$ stacking.[24]

Although XSAPT+*ai*D was designed as an affordable method for large systems, the atomic-pairwise dispersion approximation (eq 10) breaks down for large monomers.[1,22] The culprit is nonadditive dispersion arising from the simultaneous coupling of three atomic dipoles, or in other words, C_9 contributions of the Axilrod−Teller−Muto (ATM) "triple-dipole" type.[46] Three-body ATM-style dispersion corrections are included in Grimme's empirical D3 dispersion potential for DFT,[46] and similar corrections have been included in XSAPT+*ai*D.[1] The need to resort to such corrections, however, motivated the development of a method that more naturally captures dispersion nonadditivity.

3.2. Many-Body Dispersion

Nonadditive dispersion emerges from several sources:[47] from changes in polarizability due to the molecular environment, from electrodynamic screening between atoms at short range that attenuates dispersion interactions between atoms at longer range, and finally from simultaneous coupling of more than two dipoles. These effects modify the effective C_n coefficients in an atomic-pairwise dispersion expression, meaning that the parameters in eq 10 ought to be density-dependent. In contrast, XSAPT+*ai*D treats eq 10 as an unscreened sum involving parameters $C_{n,a}$ that are set by the atomic number of atom a. This leads to overestimation of dispersion energies for large monomers, especially conjugated π systems with numerous polarizable centers. To account for nonadditive dispersion, a range-separated and self-consistently screened many-body dispersion (MBD) protocol[48,49] has been modified for use with XSAPT.[2]

The MBD approach treats electrodynamic screening *via* range separation of the dipole−dipole interaction tensor $\mathbf{T}_{pq}$ between atomic centers p and q. The short-range component is used to screen the atomic polarizabilities, obtained by scaling of the free-atom data based on a Hirshfeld partition of the density that accounts for the size of an atom in the molecule.[50] The long-range part of $\mathbf{T}_{pq}$ is used to compute the dispersion energy. This long-range interaction between self-consistently screened, isotropic dipole polarizabilities $\bar{\alpha}_p$ and $\bar{\alpha}_q$ yields the leading-order term in the dispersion energy:

$$E_{\mathrm{disp}} = -\frac{1}{2\pi}\int_0^\infty \bar{\alpha}_p(i\omega)\bar{\alpha}_q(i\omega)\,\mathrm{Tr}(\mathbf{T}^2_{pq}) = -\frac{C_{6,pq}}{R_{pq}{}^6} \tag{12}$$

These $C_{6,pq}$ coefficients include self-consistent screening among all of the atoms (hence, "many-body"), reflecting nonadditive dispersion interactions beyond the leading-order ATM correction.

Unlike the situation in dispersion-corrected DFT, in which the dispersion correction is damped away at short range, true dispersion is defined at all length scales. In the context of SAPT, reparametrization of the switching function that defines the range separation in $\mathbf{T}_{pq}$ results in dispersion that is insufficiently attractive at short range, and we found it necessary to augment the MBD potential with pairwise "effectively-screened" dipole−quadrupole (esDQ) terms.[2] Effective C_8 coefficients can be obtained from the self-consistent C_6 coefficients *via* mathematical relationships between the quantum harmonic oscillators that define the MBD model. A C_8 dispersion potential is then added to the MBD dispersion energy:

$$E^{\mathrm{MBD+esDQ}}_{\mathrm{disp}} = E^{\mathrm{MBD}}_{\mathrm{disp}} - s_8\sum_{p\in A}\sum_{\substack{q\in B\\(B>A)}} f_8(R_{pq})\frac{C_{8,pq}}{R^8_{pq}} \tag{13}$$

The complete parametrization of XSAPT+MBD is provided in ref 26.

3.3. Performance

Results for the S66 data set (Figure 2) demonstrate that both XSAPT+*ai*D3 and XSAPT+MBD achieve ~1 kcal/mol accuracy for small dimers, although for hydrogen-bonded systems this does require the δHF correction. The accuracy for dispersion

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

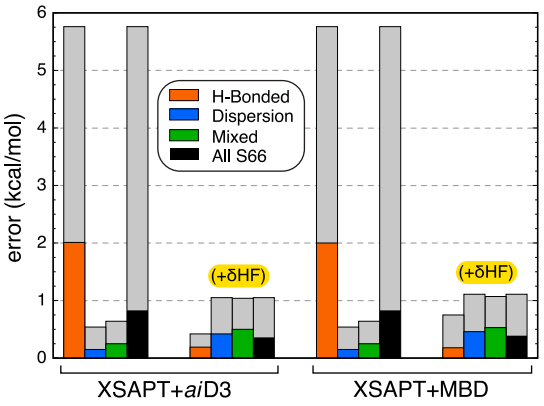


Figure 2. Performance of XSAPT/def2-TZVPPD versus CCSD(T)/CBS benchmarks for the S66 data set, using CM5 embedding charges both with and without the δHF correction. Colored bars indicate mean absolute errors and gray bars indicate maximum errors. Data are from ref 28.

interactions does not rely on a choice of basis set designed for error cancellation.

For larger supramolecular complexes, XSAPT+MBD reproduces benchmark interaction energies to an accuracy that is achieved by few other methods that are cost-effective enough to tackle large systems. Results for the L7 data set[51] (Figure 3) underscore the sizable errors that can be expected in MP2-based interaction energies for large π-stacked systems. Errors are much smaller for the XSAPT methods, but nonadditive dispersion is essential; XSAPT+*ai*D3 must be augmented with three-body dispersion corrections ($E^{ATM}_{disp,3B}$) for reasonable results. Even so, it is less accurate than XSAPT+MBD.

Recently, discrepancies of 1−2 kcal/mol between CCSD(T)/CBS and quantum Monte Carlo (QMC) benchmarks have been noted for two of the L7 complexes.[52,53] Our comparisons employ recent CCSD(T)/CBS benchmarks,[54] where sensitivity to numerical thresholds was considered carefully. For the cases in question, QMC interaction energies are smaller than the CCSD(T)/CBS values,[53] whereas MP2-based methods consistently overestimate E_{int}. As such, this discrepancy does not affect our conclusion regarding the poor performance of MP2-based methods.

For complexes with mixed electrostatic and dispersion interactions, as exemplified by the S12L∩S30L data set of host/guest complexes that is shown in Figure 4a,[2] MP2 errors are smaller, and the difference among the XSAPT approaches is less pronounced (Figure 4b). Notably, the S12L∩S30L benchmarks are experimental binding affinities back-corrected to gas-phase interaction energies, with estimated uncertainties of ±3 kcal/mol.[55]

L7 results for a selection of DFT methods that are known to perform well for non-covalent interactions are presented in Figure 3. To avoid BSSE, quadruple-ζ basis sets, or else triple-ζ basis sets with counterpoise correction, were used for all of the DFT calculations. While it is possible to find functionals that outperform XSAPT+MBD, the energy decomposition is much cleaner for the SAPT-based approach. It is also worth noting that the asymptotic cost of XSAPT+MBD is considerably less than that of supramolecular DFT.[2,23,26] Whereas the former exhibits $O(n^3)$ cost with respect to *monomer* size (n), supramolecular DFT scales as $O(N^3)$ with respect to *supersystem* size (N).

4. APPLICATIONS

4.1. Perspectives on Drug Design

The efficiency of pharmaceutical research and development, as measured by the number of new drugs put to market per dollar of research investment, has been in decline for decades, a trend that has been attributed to over-reliance on brute-force screening methods.[56] Given that about 70% of drugs function *via* non-covalent ligand−protein binding,[57] it is natural to wonder whether efficient but accurate quantum-chemical methods could be used to improve "knowledge-based" (but largely physics-free) scoring functions. With that application on the long-term horizon, we illustrate the application of XSAPT+MBD to a pair of large complexes relevant to drug binding.

The first of these is a DNA intercalation complex involving the antitumor agent ellipticine (Figure 5a), for which a CCSD(T_0)/

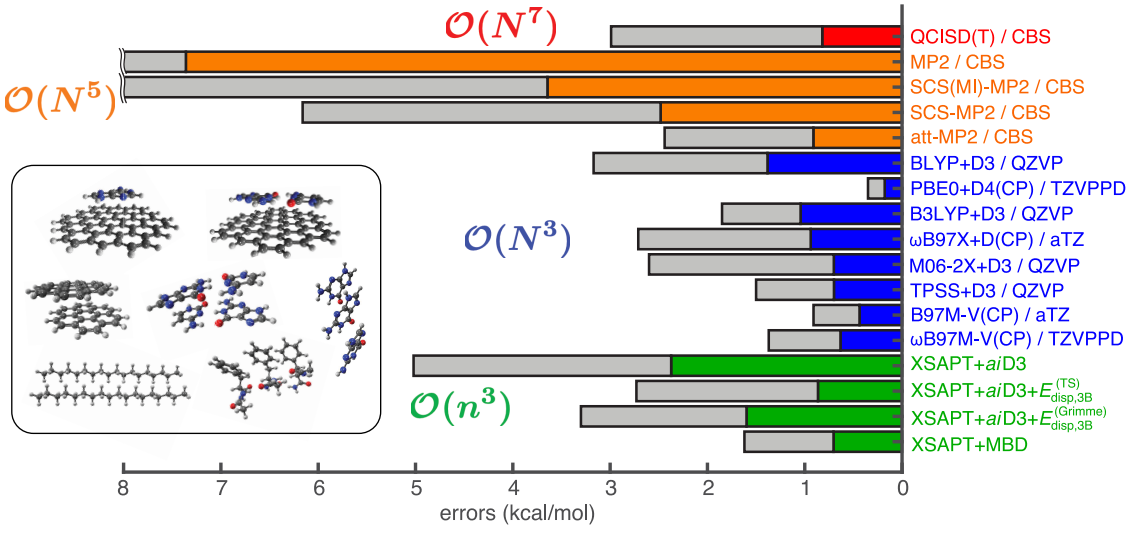


Figure 3. Mean absolute errors (colored bars) and maximum errors (gray bars) versus CCSD(T)/CBS benchmarks for the L7 data set (inset). The DFT calculations used def2-TZVPPD, def2-QZVP, or aug-cc-pVTZ, whereas all of the XSAPT calculations used def2-TZVPPD. Triple-ζ DFT results are counterpoise (CP)-corrected. Each method is color-coded by cost, where n is the number of monomer basis functions and N is the number of functions for the entire supramolecular complex.

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

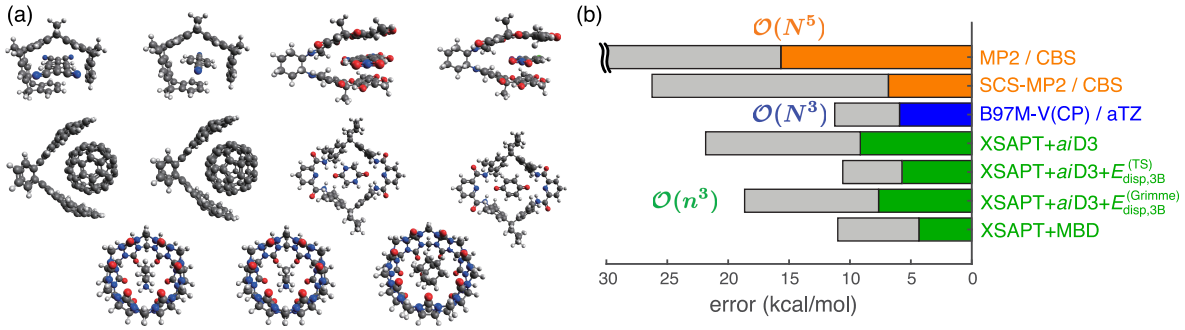


Figure 4. (a) The S12L∩S30L data set of host/guest complexes. (b) Mean absolute errors (colored bars) and maximum errors (gray bars) for various methods versus back-corrected experimental benchmarks with estimated uncertainties of ±3 kcal/mol. XSAPT calculations used the def2-TZVPPD basis set.

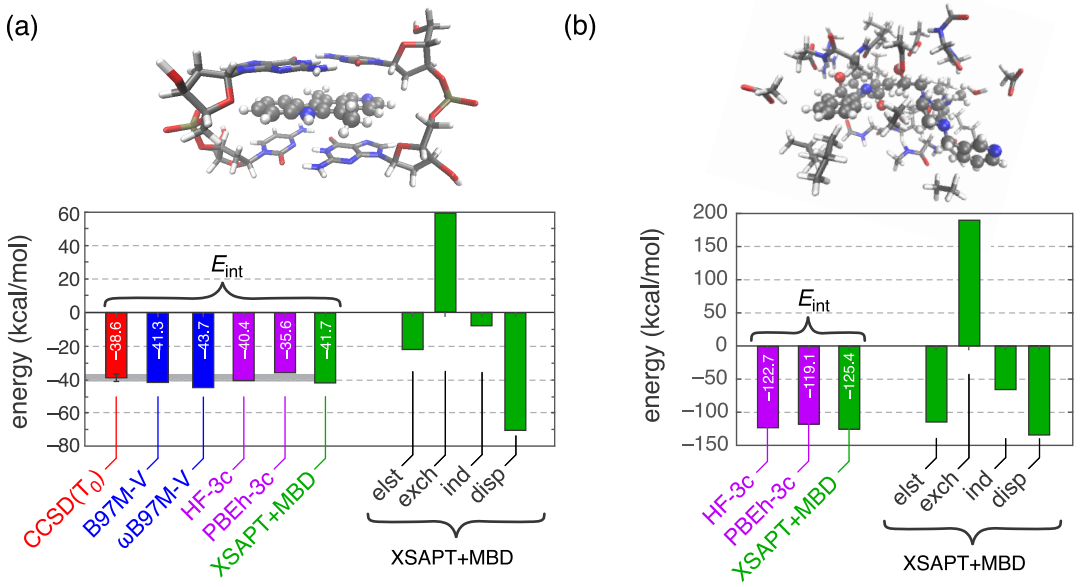


Figure 5. Interaction energies and XSAPT+MBD decompositions for two pharmacologically relevant complexes: (a) ellipticine intercalated into DNA (157 atoms) and (b) indinavir in the binding pocket of HIV-2 protease (323 atoms). The B97M-V and ωB97M-V calculations used the def2-TZVPPD basis set and are counterpoise-corrected, whereas XSAPT+MBD calculations used the def2-hpTZVPP basis set described in ref 24. In (b), only XSAPT and the semiempirical HF-3c and PBEh-3c methods are affordable. DFT and XSAPT data are from ref 26 and the CCSD(T$_0$) value is from ref 54. HF-3c and PBEh-3c values differ somewhat from those in ref 26, where the the def2-SV(P) basis set was used. Here, we use the preferred basis sets for these methods, which are MINIX for HF-3c and def2-mSVP for PBEh-3c.

CBS benchmark of −38.6 ± 2.2 kcal/mol is available.[54] The XSAPT+MBD interaction energy is within 1 kcal/mol of this result. A counterpoise-corrected ωB97M-V/def2-TZVPPD calculation (4651 basis functions) is similarly accurate but several times more expensive.

The XSAPT+MBD energy components are also shown in Figure 5a and typify the results for equilibrium geometries of π-stacked systems in that the ligand would be unbound were it not for dispersion. The dispersion energy is mostly offset by a sizable Pauli repulsion term, leading to a net "van der Waals" (vdW) energy,

$$E_{vdW} = E_{disp} + E_{exch} \quad (14)$$

of −12 kcal/mol. Near-cancellation of exchange and dispersion for π-stacked complexes has been noted before.[58] Given the important role of dispersion in stabilizing π−π interactions, a π-stacked system for which E_{exch} does not approach $|E_{disp}|$ in magnitude is probably not one that is very close to its equilibrium geometry.

A second pharmacological example (Figure 5b) features the antiretroviral drug indinavir bound to the active site of HIV-2 protease. A realistic model containing all nearby amino acids and some crystallographic waters is twice as large as the DNA−ellipticine complex, and the ωB97M-V/def2-TZVPPD calculation (8346 basis functions) would pose a significant computational challenge. Semiempirical HF-3c and PBEh-3c calculations can instead be used to provide a sanity check on the XSAPT+MBD result.[26] Notably, the dispersion energy obtained for the HIV−indinavir system is significantly larger than that computed for DNA−ellipticine, despite the absence of obvious π−π interactions in the former. This is a compelling demonstration of the ubiquity and size extensivity of dispersion.

4.2. Insights into π−π Interactions: Slip Stacking

Because of their rich substitution chemistry, π−π architectures offer flexible platforms on which to build pharmaceuticals with exceptionally specific binding sites. Before considering substituent effects, however, one should understand the basic molecular physics of π−π stacking, which has been a topic of

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

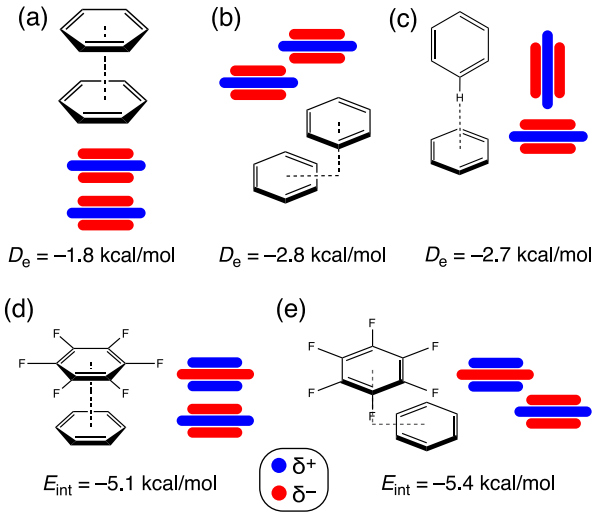


Figure 6. Isomers of $(C_6H_6)_2$ and $C_6H_6\cdots C_6F_6$, alongside representations of their quadrupolar charge distributions: (a) eclipsed-cofacial (or "sandwich") isomer of $(C_6H_6)_2$, (b) parallel-offset isomer of $(C_6H_6)_2$, (c) T-shaped isomer of $(C_6H_6)_2$, (d) cofacial isomer of $C_6H_6 \cdots C_6F_6$, and (e) parallel-displaced isomer of $C_6H_6\cdots C_6F_6$. CCSD(T)/CBS interaction energies are shown (data are from refs 45 and 61).

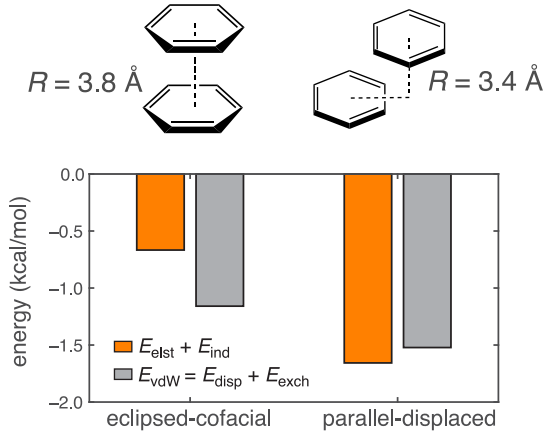


Figure 7. XSAPT+MBD energy decomposition for $(C_6H_6)_2$ at two stationary-point geometries.

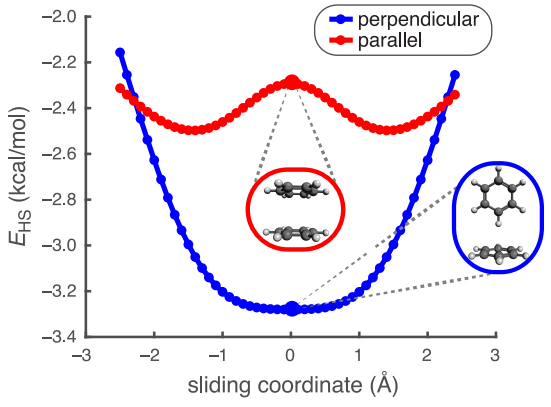


Figure 8. Scans of the Hunter–Sanders model potential (eq 17) for parallel and perpendicular arrangements of $(C_6H_6)_2$.

some debate.[4] The accuracy of XSAPT+MBD allows us to interrogate the competing forces that drive π–π conformational preferences,[3] free of the vagaries of this or that density

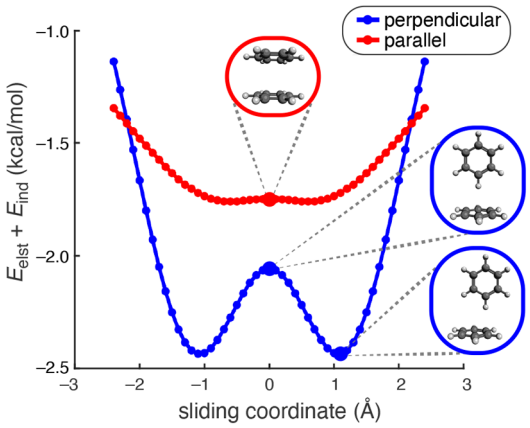


Figure 9. Scans of $E_{elst} + E_{ind}$ from XSAPT+MBD calculations for $(C_6H_6)_2$ in both parallel and perpendicular orientations. From ref 3. CC BY 3.0.

functional. Doing so led us to question the textbook "Hunter–Sanders" model of π–π interactions,[59,60] which is based on quadrupolar electrostatics.

Figure 6a–c depicts three canonical stereoisomers of $(C_6H_6)_2$ along with cartoon depictions of the charge distribution that gives the benzene monomer a significant quadrupole moment. The lowest-energy structure is the slip-stacked (or parallel-displaced) isomer. A survey of phenylalanine close contacts in protein crystal structures suggests that slip-stacked and T-shaped configurations are common but that eclipsed-cofacial π stacking (as in Figure 6a) is essentially absent.[62]

According to the Hunter–Sanders model,[59] slip stacking emerges from a competition between dispersion (favoring cofacial π stacking) and quadrupolar electrostatics that is repulsive in cofacial arrangements but attractive in T-shaped geometries. This has become the dominant paradigm for understanding π–π conformational preferences,[60] despite its failure to explain why the $C_6H_6\cdots C_6F_6$ heterodimer (Figures 6d,e) *also* exhibits slip stacking.[61] In that system, the polarity of the C–F bonds reverses the sign of the quadrupolar electrostatics term, which is therefore *attractive* in the eclipsed-cofacial geometry. A more general criticism is that the very concept of quadrupolar electrostatics is based on an asymptotic expansion that may be qualitatively incorrect at vdW contact distances, which is indeed the case for $(C_6H_6)_2$.[36]

Given how small the energy differences are between isomers of $(C_6H_6)_2$, a reliable understanding of the intermolecular forces that drive π–π conformational preferences requires a method that affords accurate energetics, free of the vagaries of DFT functional selection. Our work has also emphasized the need to examine potential energy *surfaces*,[3] not just stationary points, because the cancellation of attractive and repulsive forces at such points can obscure the driving forces that cause a particular geometry to be stationary.

Consider the energy decomposition at the cofacial and parallel-displaced stationary points of $(C_6H_6)_2$ (Figure 7). At first, these results appear to corroborate the Hunter–Sanders picture insofar as the "polarized electrostatics",

$$E_{elst+ind} = E_{elst} + E_{ind} \quad (15)$$

increases by 150% upon parallel displacement whereas E_{vdW} (eq 14) increases by only 30%. This result is sometimes taken as evidence that electrostatic forces guide the parallel offset, despite the fact that E_{elst} is attractive in both geometries, in contradiction

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679–3690

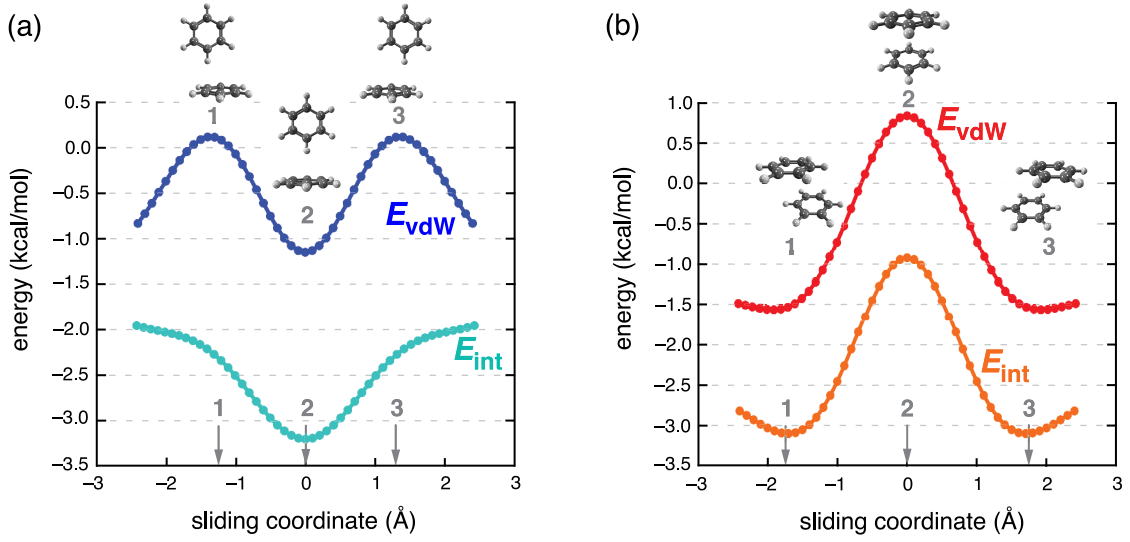


Figure 10. XSAPT+MBD interaction potentials (E_{int}) versus vdW potentials (E_{vdW}, eq 14) along the sliding coordinate of $(C_6H_6)_2$ in either (a) the perpendicular orientation or (b) the cofacial orientation. Arrows at the bottom indicate the position coordinates of the structures shown at the top. From ref 3. CC BY 3.0.

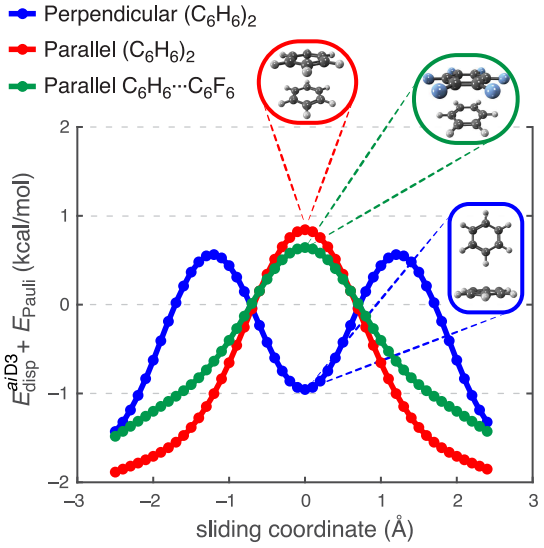


Figure 11. Potential energy surfaces along the sliding coordinates of $(C_6H_6)_2$ and $C_6H_6\cdots C_6F_6$ as predicted by a vdW force field model, $E^{aiD3}_{disp} + E_{Pauli}$. From ref 3. CC BY 3.0.

to the underlying hypothesis of the Hunter−Sanders model. Figure 7 also highlights the fact that the parallel-displaced geometry is characterized by 0.4 Å smaller separation between the benzene rings compared with the eclipsed-cofacial geometry. This difference arises from a reduction in Pauli repulsion upon parallel displacement,[3] which becomes clear only upon performing potential energy scans.

To rationalize offset-stacking, Hunter and Sanders[59] introduced a model potential consisting of an arrangement of point charges that mimics the charge distributions shown in Figure 6a−c and reproduces the experimental quadrupole moment for C_6H_6. The model electrostatic energy is then

$$E^{Q}_{elst} = \sum_{i\in A}\sum_{j\in B}\frac{q_i q_j}{r_{ij}} \tag{16}$$

to which Hunter and Sanders add an atom−atom potential of "exponential-6" type, to obtain the model potential

$$E_{HS} = E^{Q}_{elst} + \sum_{i\in A}\sum_{i\in B}\left(A_{ij}e^{-\alpha_{ij}R_{ij}} - \frac{C_{6,ij}}{R^6_{ij}}\right) \tag{17}$$

One-dimensional scans of this potential surface for $(C_6H_6)_2$ in both parallel and perpendicular arrangements are plotted in Figure 8. The model correctly predicts that the eclipsed-cofacial geometry is a saddle point connecting symmetric parallel-displaced minima, but in the perpendicular orientation it fails to predict a preference for the T-shaped (C_{2v}-symmetric) geometry. Instead, distortions of ±1 Å are possible without penalty, leading to L-shaped geometries that are inconsistent with accurate calculations. In three decades of discussing the Hunter−Sanders model, it is unclear that anyone ever bothered to check its behavior for the T-shaped isomer.

The picture is considerably more nuanced when XSAPT+MBD is used to compute the energy components.[3] Considering polarized electrostatics first, one-dimensional scans of $E_{elst+ind}$ are shown in Figure 9. As a result of charge penetration,[4,63] leading to a breakdown of the multipole expansion at vdW contact distances,[36] electrostatic interactions for $(C_6H_6)_2$ are attractive even in the eclipsed-cofacial geometry. In the parallel arrangement, $E_{elst+ind}$ hardly changes for parallel displacements of ±1 Å (similar to the total interaction potential in the Hunter−Sanders model), suggesting that something other than electrostatics is responsible for the emergence of slip stacking. In perpendicular arrangements, polarized electrostatics alone favors L-shaped geometries that are not present on the full XSAPT+MBD potential energy surface. These results underscore the importance of considering the potential energy *surface*, as opposed to single-point energies at a few geometries, when attempting to ascribe causal links between energy components and geometric isomers.

Whereas polarized electrostatics fails to predict offset stacking, the vdW potential (eq 14) predicts the same geometries as the full interaction potential, as shown using XSAPT+MBD in Figure 10. The origin of offset stacking proves to be a reduction in Pauli repulsion (which competes with

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

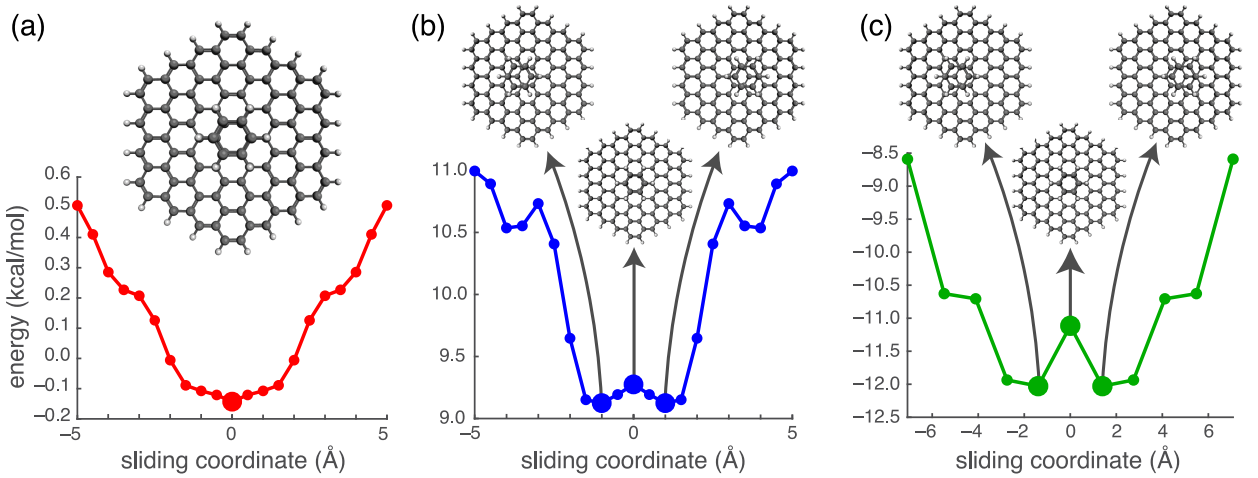


Figure 12. Potential energy surfaces for sliding benzene along the surface of $C_{96}H_{24}$ in a parallel orientation, as predicted by (a) the Hunter–Sanders model potential (eq 17), (b) the vdW model potential ($E_{\text{disp}}^{aiD3} + E_{\text{Pauli}}$), and (c) XSAPT+MBD.

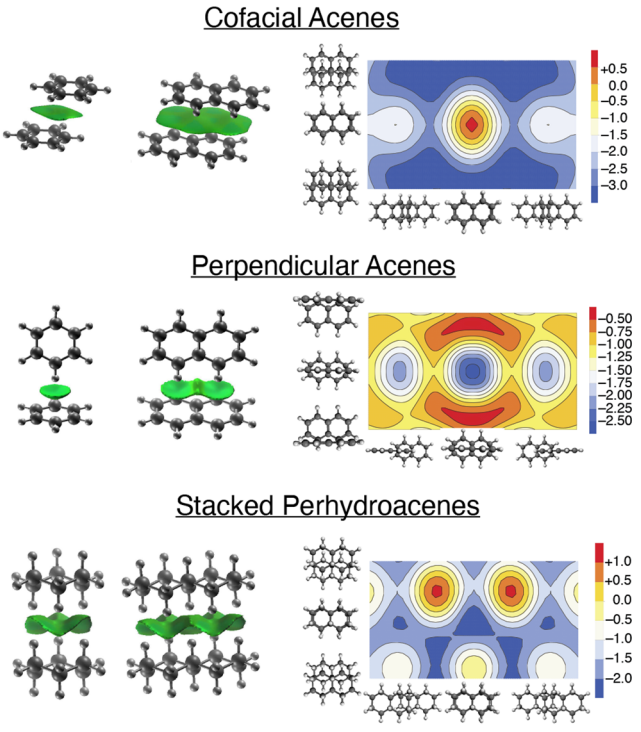


Figure 13. (left) Isosurfaces of the reduced density gradient (eq 19). (right) vdW potential surfaces (E_{vdW}) from XSAPT+MBD calculations, in units of kcal/mol.

dispersion), with electrostatics as an ambivalent spectator.[3,4] These results inspired us to develop an alternative semiclassical model (to replace eq 17)[3] in which we dispense with electrostatics altogether, replace the dispersion model with *ai*D3,[24] and replace the exponential repulsion model with one based on the overlap of atom-centered spherical Gaussian functions. The form of the latter is

$$E_{\text{Pauli}} = \sum_{i\in A}\sum_{j\in B} \frac{Z_i Z_j S_{ij}^{\ 2}}{R_{ij}} \tag{18}$$

The resulting "vdW model" ($E_{\text{disp}}^{aiD3} + E_{\text{Pauli}}$) correctly captures the topography of both parallel and perpendicular configurations of $(C_6H_6)_2$, as shown in Figure 11. In addition, the vdW model correctly predicts a saddle point at the eclipsed-cofacial geometry of $C_6H_6\cdots C_6F_6$, leading to offset stacking in that system as well. These data suggest a useful starting point for constructing classical force fields that describe π stacking for correct physical reasons.

To test whether this force field model is generalizable to $\pi-\pi$ interactions in larger systems, we computed the potential energy surface for dragging a benzene molecule across the surface of a $C_{96}H_{24}$ graphene nanoflake (Figure 12). Here, the accuracy and scalability of XSAPT+MBD are crucial for obtaining high-quality energetics against which simple models can be tested. The Hunter–Sanders model qualitatively fails to describe the $C_6H_6\cdots C_{96}H_{24}$ potential surface (Figure 12a), incorrectly predicting that the benzene molecule should be perfectly centered atop the nanoflake, aligned with the lattice of carbon atoms below. In contrast, the vdW model predicts an offset (Figure 12b), consistent with the XSAPT+MBD results (Figure 12c). Notably, the vdW model was parametrized for $(C_6H_6)_2$, and we did not modify the parameters for this larger example.

4.3. Origins of $\pi-\pi$ Stacking

Given the abject failure of the conventional Hunter–Sanders model, we thought it pertinent to revisit an old question: does $\pi-\pi$ stacking constitute a unique form of dispersion?[4] The answer can be summarized in pictorial form by examining the reduced density gradient,

$$s(\mathbf{r}) = \frac{\|\nabla\rho(\mathbf{r})\|}{2(3\pi^2)^{1/3}\rho(\mathbf{r})^{4/3}} \tag{19}$$

This function can indicate the presence of non-covalent interactions, which manifest as oscillations in the density in regions where $\rho(\mathbf{r})$ itself is small. Isosurfaces of $s(\mathbf{r})$ are plotted in Figure 13 for the benzene and naphthalene dimers and also for their saturated analogues, the cyclohexane and perhydronaphthalene dimers. The fact that these isosurfaces resemble molecular surfaces is no accident and underscores the primacy of the vdW potential energy surface. The latter surfaces are also presented in Figure 13 for the cofacial and perpendicular arrangements of $(\text{naphthalene})_2$ as well as the stacked arrangement of $(\text{perhydronaphthalene})_2$. Additional examples can be found in ref 4.

The planarity of the acene molecules facilitates offset stacking in the cofacial arrangement because E_{vdW} is attractive over a wide

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679–3690

range of lateral configurations. This contrasts with both the perpendicular acenes and the stacked perhydroacenes, for which E_{vdW} exhibits a sawtooth topography and favors just one geometry, characterized by C−H moieties of one monomer that are locked within the ring structure of its partner, presenting a significant barrier to lateral motion. The perpendicular acenes prove to be much more similar to the saturated perhydroacenes (models of graphane) than they are to the cofacial acenes (models of graphene), and this is true for the total interaction energies as well, where the cofacial acenes are systematically more strongly bound and the perpendicular isomers are surprisingly similar to the stacked perhydroacene dimers.[4] Benzene dimer is an anomaly in this respect, insofar as the parallel and perpendicular orientations have nearly the same interaction energy, which is also nearly identical to that of stacked cyclohexane dimer.[4] For that reason, $(C_6H_6)_2$ is not the archetypal example of $\pi-\pi$ stacking that it is often assumed to be.

The explanation for these observations reveals that there is indeed a unique aspect to $\pi-\pi$ stacking, but it is only indirectly related to dispersion. Compared with the perpendicular acenes or the stacked perhydroacenes, the cofacial acene dimers do not experience significant Pauli repulsion until much shorter intermolecular separation. As a result, the cofacial isomers are able to access electrostatic interactions that are *attractive* at short range, as a result of charge penetration effects and in direct contradiction to the Hunter−Sanders hypothesis. In the end, $\pi-\pi$ stacking turns out to be a largely geometric phenomenon, wherein flat molecular architectures "serve up" interactions on a vdW surface that is much less perturbed by intermolecular exchange compared to the more complex topography of a flexible hydrocarbon. This "pizza-π" model of stacking[4] explains the existence of strong stacking interactions in other molecules that are planar but not aromatic.[64]

5. SUMMARY

Key features of the XSAPT methodology are a balance of accuracy and affordability. By managing to strike this balance with minimal parameter fitting, XSAPT fills a unique niche in that it can be applied to relatively large systems of biochemical interest yet comes with an energy partition analysis that is reliably founded in rigorous quantum mechanics. Apparently, there is still basic molecular physics to be learned from systems as simple as benzene dimer,[3] acene dimers,[4] and halide−water complexes,[30] and in each of these examples we find that XSAPT-based calculations challenge textbook explanations for non-covalent interactions.[36] Given the computational extensibility of XSAPT, "chemical space" is wide open to bring this insight into much larger systems of pharmacological or materials-science interest.

AUTHOR INFORMATION

Corresponding Author

John M. Herbert − *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0002-1663-2278; Email: herbert@chemistry.ohio-state.edu

Authors

Kevin Carter-Fenk − *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0001-8302-4750

Ka Un Lao − *Department of Chemistry, Virginia Commonwealth University, Richmond, Virginia 23284, United States;* orcid.org/0000-0002-3993-536X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.accounts.1c00387

Notes

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

Biographies

Kevin Carter-Fenk was born on December 22, 1993, in Alliance, Ohio. He graduated from The Ohio State University with a B.Sc. in Chemistry in 2015 and then again with a Ph.D. in 2021 under the supervision of Prof. John Herbert. He is currently a postdoctoral scholar at the University of California, Berkeley under the direction of Prof. Martin Head-Gordon.

Ka Un Lao was born in Macau and received B.Sc. and M.Sc. degrees in chemistry from National Tsing Hua University in Hsinchu, Taiwan. He obtained his Ph.D. from The Ohio State University, where he studied under Prof. John Herbert. Following postdoctoral work at Cornell University with Prof. Robert DiStasio, he joined Virginia Commonwealth University as an Assistant Professor in 2019. His research interests include intermolecular interactions, fragment-based quantum chemistry, and machine learning.

John M. Herbert is a quantum chemist with wide-ranging interests in physical chemistry. In addition to non-covalent interactions, his group works on computational methods for condensed-phase spectroscopy, photochemistry, solvation modeling, and the electron correlation problem, with particular focus on the development of new methods and algorithms for large molecules and condensed-phase systems. His research group is a major contributor to the Q-Chem electronic structure program. He is a previous Fellow of the Humboldt, Dreyfus, and Sloan Foundations and currently holds the rank of Professor at The Ohio State University.

ACKNOWLEDGMENTS

The development of XSAPT has been funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, under Award DE-SC0008550 (to J.M.H.).

REFERENCES

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

(2) Carter-Fenk, K.; Lao, K. U.; Liu, K.-Y.; Herbert, J. M. Accurate and efficient *ab initio* calculations for supramolecular complexes: Symmetry-adapted perturbation theory with many-body dispersion. *J. Phys. Chem. Lett.* **2019**, *10*, 2706−2714.

(3) Carter-Fenk, K.; Herbert, J. M. Electrostatics does not dictate the slip-stacked arrangement of aromatic $\pi-\pi$ interactions. *Chem. Sci.* **2020**, *11*, 6758−6765.

(4) Carter-Fenk, K.; Herbert, J. M. Reinterpreting π-stacking. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24870−24886.

(5) Clerk-Maxwell, J. Van der Waals on the continuity of the gaseous and liquid states. *Nature* **1874**, *10*, 477−480.

(6) Tröster, P.; Tavan, P. The microscopic physical cause for the density maximum of liquid water. *J. Phys. Chem. Lett.* **2014**, *5*, 138−142.

(7) Beran, G. J. O.; Heit, Y. N.; Hartman, J. D. Noncovalent interactions in molecular crystals. In *Non-Covalent Interactions in*

Quantum Chemistry and Physics; de la Roza, A. O., DiLabio, G. A., Eds.; Elsevier: Amsterdam, 2017; Chapter 10, pp 303−331.
(8) Wagner, J. P.; Schreiner, P. R. London dispersion in molecular chemistry−Reconsidering steric effects. *Angew. Chem., Int. Ed.* **2015**, *54*, 12274−1229.
(9) McGaughey, G. B.; Gagné, M.; Rappé, A. K. π-stacking interactions: Alive and well in proteins. *J. Biol. Chem.* **1998**, *273*, 15458−15463.
(10) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. C-H···π interactions in proteins. *J. Mol. Biol.* **2001**, *307*, 357−377.
(11) He, X.; Fusti-Molnar, L.; Cui, G.; Merz, K. M., Jr. The importance of dispersion and electron correlation in ab initio protein folding. *J. Phys. Chem. B* **2009**, *113*, 5290−5300.
(12) Xu, Z.; Zhang, Q.; Shi, J.; Zhu, W. Underestimated noncovalent interactions in protein data bank. *J. Chem. Inf. Model.* **2019**, *59*, 3389−3399.
(13) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. State of the art in counterpoise theory. *Chem. Rev.* **1994**, *94*, 1873−1885.
(14) Richard, R. M.; Lao, K. U.; Herbert, J. M. Achieving the CCSD(T) basis-set limit in sizable molecular clusters: Counterpoise corrections for the many-body expansion. *J. Phys. Chem. Lett.* **2013**, *4*, 2674−2680.
(15) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 304−326.
(16) Francisco, E.; Pendás, A. M. Energy partition analyses: Symmetry-adapted perturbation theory and other techniques. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; de la Roza, A. O., DiLabio, G. A., Eds.; Elsevier: Amsterdam, 2017; Chapter 2, pp 27−64.
(17) Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, e1452.
(18) Stone, A. J. Physical basis of intermolecular interactions. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; de la Roza, A. O., DiLabio, G. A., Eds.; Elsevier: Amsterdam, 2017; Chapter 1, pp 3−26.
(19) Jacobson, L. D.; Herbert, J. M. An efficient, fragment-based electronic structure method for molecular systems: Self-consistent polarization with perturbative two-body exchange and dispersion. *J. Chem. Phys.* **2011**, *134*, 094118.
(20) Herbert, J. M.; Jacobson, L. D.; Lao, K. U.; Rohrdanz, M. A. Rapid computation of intermolecular interactions in molecular and ionic clusters: Self-consistent polarization plus symmetry-adapted perturbation theory. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7679−7699.
(21) Jacobson, L. D.; Richard, R. M.; Lao, K. U.; Herbert, J. M. Efficient monomer-based quantum chemistry methods for molecular and ionic clusters. *Annu. Rep. Comput. Chem.* **2013**, *9*, 25−56.
(22) Lao, K. U.; Herbert, J. M. A simple correction for nonadditive dispersion within extended symmetry-adapted perturbation theory (XSAPT). *J. Chem. Theory Comput.* **2018**, *14*, 5128−5142.
(23) Lao, K. U.; Herbert, J. M. Accurate intermolecular interactions at dramatically reduced cost: XPol+SAPT with empirical dispersion. *J. Phys. Chem. Lett.* **2012**, *3*, 3241−3248.
(24) Lao, K. U.; Herbert, J. M. Accurate and efficient quantum chemistry calculations of noncovalent interactions in many-body systems: The XSAPT family of methods. *J. Phys. Chem. A* **2015**, *119*, 235−253.
(25) Lao, K. U.; Herbert, J. M. An improved treatment of empirical dispersion and a many-body energy decomposition scheme for the explicit polarization plus symmetry-adapted perturbation theory (XSAPT) method. *J. Chem. Phys.* **2013**, *139*, 034107; Erratum: *J. Chem. Phys.* **2014**, *140*, 119901.
(26) Liu, K.-Y.; Carter-Fenk, K.; Herbert, J. M. Self-consistent charge embedding at very low cost, with application to symmetry-adapted perturbation theory. *J. Chem. Phys.* **2019**, *151*, 031102.
(27) Lao, K. U.; Herbert, J. M. Symmetry-adapted perturbation theory with Kohn-Sham orbitals using non-empirically tuned, long-range-corrected density functionals. *J. Chem. Phys.* **2014**, *140*, 044108.
(28) Gray, M.; Herbert, J. M. Simplified tuning of long-range corrected density functionals for symmetry-adapted perturbation theory. *J. Chem. Phys.* **2021**, *155*, 034103.
(29) Lao, K. U.; Herbert, J. M. Energy decomposition analysis with a stable charge-transfer term for interpreting intermolecular interactions. *J. Chem. Theory Comput.* **2016**, *12*, 2569−2582.
(30) Herbert, J. M.; Carter-Fenk, K. Electrostatics, charge transfer, and the nature of the halide-water hydrogen bond. *J. Phys. Chem. A* **2021**, *125*, 1243−1256.
(31) Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy decomposition analysis approaches and their evaluation on prototypical protein-drug interaction patterns. *Chem. Soc. Rev.* **2015**, *44*, 3177−3211.
(32) Shahbaz, M.; Szalewicz, K. Do semilocal density-functional approximations recover dispersion energies at small intermonomer separations? *Phys. Rev. Lett.* **2018**, *121*, 113402.
(33) Price, A. J. A.; Bryenton, K. R.; Johnson, E. R. Requirements for an accurate dispersion-corrected density functional. *J. Chem. Phys.* **2021**, *154*, 230902.
(34) Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211−228.
(35) Nguyen, B.; Chen, G. P.; Agee, M. M.; Burow, A. M.; Tang, M.; Furche, F. Divergence of many-body perturbation theory for non-covalent interactions of large molecules. *J. Chem. Theory Comput.* **2020**, *16*, 2258−2273.
(36) Herbert, J. M. Neat, simple, and wrong: Debunking electrostatic fallacies regarding noncovalent interactions. *J. Phys. Chem. A* **2021**, *125*, 7125−7137.
(37) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, *140*, 094106.
(38) London, F. The general theory of intermolecular forces. *Trans. Faraday Soc.* **1937**, *33*, 8−26.
(39) Moszynski, R.; Jeziorski, B.; Ratkiewicz, A.; Rybak, S. Many-body perturbation theory of electrostatic interactions between molecules: Comparison with full configuration interaction for four-electron dimers. *J. Chem. Phys.* **1993**, *99*, 8856−8869.
(40) Řezác, J.; Riley, K. E.; Hobza, P. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438; Erratum: *J. Chem. Theory Comput.* **2014**, *10*, 1359−1360.
(41) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 254−272.
(42) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 127−144.
(43) Herbert, J. M. Fantasy versus reality in fragment-based quantum chemistry. *J. Chem. Phys.* **2019**, *151*, 170901.
(44) Gao, J.; Truhlar, D. G.; Wang, Y.; Mazack, M. J. M.; Löffler, P.; Provorse, M. R.; Rehak, P. Explicit polarization: A quantum mechanical framework for developing next generation force fields. *Acc. Chem. Res.* **2014**, *47*, 2837−2845.
(45) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the ab initio limit for π-π interactions: The benzene dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.
(46) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-corrected mean-field electronic structure methods. *Chem. Rev.* **2016**, *116*, 5105−5154.
(47) Dobson, J. F. Beyond pairwise additivity in London dispersion interactions. *Int. J. Quantum Chem.* **2014**, *114*, 1157−1161.
(48) Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A., Jr.; Tkatchenko, A. Long-range correlation energy calculated from coupled atomic response functions. *J. Chem. Phys.* **2014**, *140*, 18A508.

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

(49) Hermann, J.; DiStasio, R. A., Jr.; Tkatchenko, A. First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications. *Chem. Rev.* **2017**, *117*, 4714−4758.
(50) Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102*, 073005.
(51) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364−3374.
(52) Benali, A.; Shin, H.; Heinonen, O. Quantum Monte Carlo benchmarking of large noncovalent complexes in the L7 benchmark set. *J. Chem. Phys.* **2020**, *153*, 194113.
(53) Al-Hamdani, Y. S.; Nagy, P. R.; Zen, A.; Barton, D.; Kállay, M.; Brandenburg, J. G.; Tkatchenko, A. Interactions between large molecules pose a puzzle for reference quantum mechanical methods. *Nat. Commun.* **2021**, *12*, 3927.
(54) Ballesteros, F.; Dunivan, S.; Lao, K. U. Coupled cluster benchmarks of large noncovalent complexes: The L7 dataset as well as DNA-ellipticine and buckycatcher-fullerene. *J. Chem. Phys.* **2021**, *154*, 154104.
(55) Sure, R.; Grimme, S. Comprehensive benchmark of association (free) energies of realistic host-guest complexes. *J. Chem. Theory Comput.* **2015**, *11*, 3785−3801; Erratum: *J. Chem. Theory Comput.* **2015**, *11*, 5990.
(56) Scannell, J. W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the decline in pharamceutical R&D efficiency. *Nat. Rev. Drug Discovery* **2012**, *11*, 191−200.
(57) Kumalo, H. M.; Bhakat, S.; Soliman, M. E. S. Theory and applications of covalent docking in drug discovery: Merits and pitfalls. *Molecules* **2015**, *20*, 1984−2000.
(58) Langner, K. M.; Sokalski, W. A.; Leszczynski, J. Intriguing relations of interaction energy components in stacked nucleic acids. *J. Chem. Phys.* **2007**, *127*, 111102.
(59) Hunter, C. A.; Sanders, J. K. M. The nature of $\pi-\pi$ interactions. *J. Am. Chem. Soc.* **1990**, *112*, 5525−5534.
(60) Fagnani, D. E.; Sotuyo, A.; Castellano, R. K. $\pi-\pi$ interactions. In *Comprehensive Supramolecular Chemistry II*; Elsevier: Oxford, U.K., 2017; Vol. *1*, Chapter 6, pp 121−148.
(61) Tsuzuki, S.; Uchimaru, T.; Mikami, M. Intermolecular interaction between hexafluorobenzene and benzene: Ab initio calculations including CCSD(T) level electron correlation correction. *J. Phys. Chem. A* **2006**, *110*, 2027−2033.
(62) Hunter, C. A.; Singh, J.; Thornton, J. M. π-π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol.* **1991**, *218*, 837−846.
(63) Sherrill, C. D. Energy component analysis of π interactions. *Acc. Chem. Res.* **2013**, *46*, 1020−1028.
(64) Bloom, J. W. G.; Wheeler, S. E. Taking the aromaticity out of aromatic interactions. *Angew. Chem., Int. Ed.* **2011**, *50*, 7847−7849.

https://doi.org/10.1021/acs.accounts.1c00387
Acc. Chem. Res. 2021, 54, 3679−3690

ACCOUNTS of chemical research

pubs.acs.org/accounts

Addition/Correction

Correction to "Predicting and Understanding Noncovalent Interactions Using Novel Forms of Symmetry-Adapted Perturbation Theory"

Kevin Carter-Fenk, Ka Un Lao, and John M. Herbert*

Acc. Chem. Res. **2021**, *54* (19), 3679−3690. DOI: 10.1021/acs.accounts.1c00387

Cite This: *Acc. Chem. Res.* 2025, 58, 1051−1053

Read Online

ACCESS | Metrics & More | Article Recommendations

In a previous Account,[1] we surveyed the use of extended symmetry-adapted perturbation theory (XSAPT), a family of methods for computing accurate intermolecular interaction energies and components thereof. In considering π-stacking interactions, we made comparisons between the XSAPT + many-body dispersion (MBD) method and a model potential introduced in a seminal paper on $\pi-\pi$ interactions by Hunter and Sanders (HS).[2] Unfortunately, our implementation of the HS model contained an error in the van der Waals (vdW) term, which is corrected here alongside some additional clarifications. Because there are subtleties in how the vdW parameters were originally reported,[2] as well as ambiguity regarding which point charges constitute the HS model,[2,3] additional details are provided here.

The HS model consists of a point-charge electrostatic term ($E_{\rm elst}^{\rm Q}$) and a vdW term ($E_{\rm vdW}$),

$$E_{\rm HS} = E_{\rm elst}^{\rm Q} + E_{\rm vdW} \tag{1}$$

The latter has a typical form,

$$E_{\rm vdW} = \sum_{i\in I}\sum_{j\in J}\left(A_{ij}e^{-\alpha_{ij}r_{ij}} - \frac{C_{ij}}{r_{ij}^6}\right) \tag{2}$$

In the original HS paper,[2] parameters for eq 2 were taken from ref 4 but with notation that is reversed relative to that work. (Our notation follows that in ref 2.) Although HS purport to have used parameters A_{ij}, C_{ij}, and α_{ij} from ref 4, they instead use

$$\alpha_{ij} = \alpha/R_{ij}^0 \tag{3a}$$

$$C_{ij} = C(R_{ij}^0)^6 \tag{3b}$$

$$A_{ij} = A \tag{3c}$$

where α, C, and A are the parameters from ref 4, with A and C swapped. The parameter R_{ij}^0 is the geometric mean of the vdW radii of atoms i and j,

$$R_{ij}^0 = 2(R_iR_j)^{1/2} \tag{4}$$

with $R_{\rm C}$ = 1.77 Å and $R_{\rm H}$ = 1.2 Å used for benzene. Numerical values for the other parameters are

$$\alpha = 12.35\ \text{Å} \tag{5a}$$

$$A = 47 \times 10^3\ \text{kcal/mol} \tag{5b}$$

$$C = 0.214\ \text{kcal/mol} \tag{5c}$$

These values were incorrect in our original work,[1] with the result that $E_{\rm vdW}$ was insufficiently repulsive.

There is some ambiguity regarding the point charges to be used in $E_{\rm elst}^{\rm Q}$. What is clear is that the HS model contains atom-centered point charges for carbon atoms within the π-system ($q_{\rm C}$) along with out-of-plane displaced charges (q_π) to represent the π-electrons. In their original 1990 paper, HS first discuss "unpolarized" or "idealized" charges, in which carbon atoms within the π-system are described by charges $q_{\rm C}$ = +1.0 and q_π = −0.5 (in atomic units).[2] The π charges are displaced from the nuclei by δ = 0.47 Å, both above and below the arene plane, a value that is determined in order to reproduce the experimental quadrupole moment of C_6H_6.[5] Although the HS paper includes a discussion of polarizing this idealized framework, no actual values for hydrogen-atom charges are provided in ref 2. Moreover, Figure 3 of ref 2 depicts only $q_{\rm C}$ = +1.0 and q_π = −0.5, with no indication that there are charges on the hydrogen atoms.

In 1991, Hunter et al.[3] suggested a model in which the charge on carbon is reduced to $q_{\rm C}$ = +0.95 and a charge $q_{\rm H}$ = +0.05 is placed on hydrogen, retaining q_π = −0.5. This scheme (in Figure 3 of ref 3) is attributed to the original HS model even though the value of $q_{\rm H}$ was not provided in the original. In other work by Hunter and co-workers, only $q_{\rm C}$ and q_π are discussed, e.g., in Figure 3 of ref 6. These ambiguities are consistent with widespread confusion in the literature regarding what the HS model actually is, as discussed

Published: March 3, 2025

elsewhere.[7] For this Correction, we implemented E^{Q}_{elst} according to ref 3 using q_C = +0.95, q_H = +0.05, and q_π = −0.5. For $(C_6H_6)_2$, the presence or absence of q_H makes only a minor difference.

Using the corrected parameters in E_{vdW} and the updated parameters in E^{Q}_{elst}, we recomputed HS potentials for the lateral displacement of parallel and perpendicular arrangements of the benzene dimer. Figure 1 of this Correction should replace Figure 8 in ref 1, illustrating how E_{HS} and E^{Q}_{elst} vary with lateral displacement. Note that the electrostatic component (Figure 1b) remains qualitatively incorrect in comparison to a calculation based on full monomer charge densities, the latter of which can be found in Figure 9 of ref 1. The crux of our argument is unchanged, namely, that electrostatics does not explain parallel-displaced π-stacking.

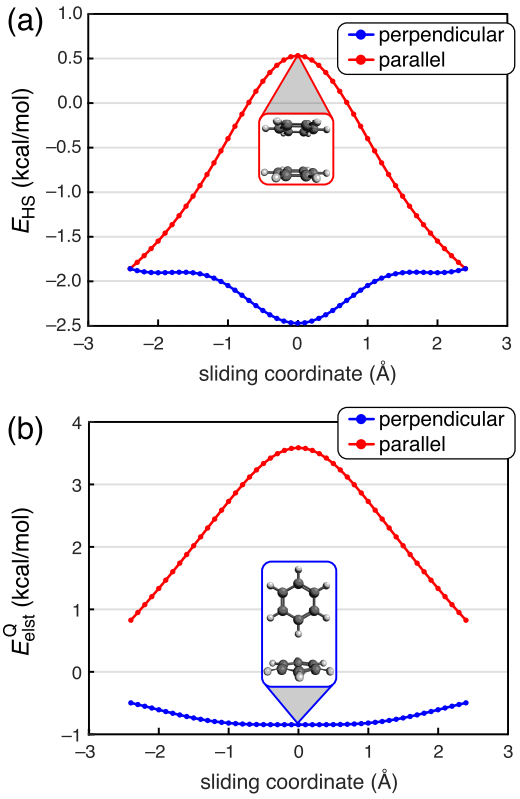


Figure 1. (a) Total HS model potential and (b) its electrostatic component, for lateral displacement of $(C_6H_6)_2$ in either a coplanar configuration at 3.4 Å separation (consistent with the parallel-displaced minimum-energy geometry) or else a perpendicular edge-to-face arrangement with a 5.0 Å center-to-center distance, consistent with the T-shaped saddle point of $(C_6H_6)_2$. This figure should replace Figure 8 of ref 1.

Figure 2 of this Correction plots E_{HS} and E^{Q}_{elst} for lateral displacement of benzene atop a $C_{96}H_{24}$ graphene nanoflake; this figure should replace Figure 12a of ref 1. As compared to XSAPT + MBD calculations (in Figure 12c of ref 1), the corrected version of the HS model potential is qualitatively correct insofar as the zero-displacement structure is a saddle point between symmetry-equivalent minima corresponding to parallel-displaced π-stacking. Nevertheless, the electrostatic component E^{Q}_{elst} (Figure 2b of this Correction) is repulsive at all values of the lateral displacement coordinate, which is not consistent with exact electrostatics.[8] Indeed, a central aspect of our "pizza-π" model of $\pi-\pi$ interactions,[8] which explains how π-stacking is different from ordinary dispersion, is that electrostatics is attractive for two coplanar arenes separated by typical π-stacking distances of 3.4−3.8 Å. This conclusion is borne out in a wide variety of $\pi-\pi$ interactions.[7−11]

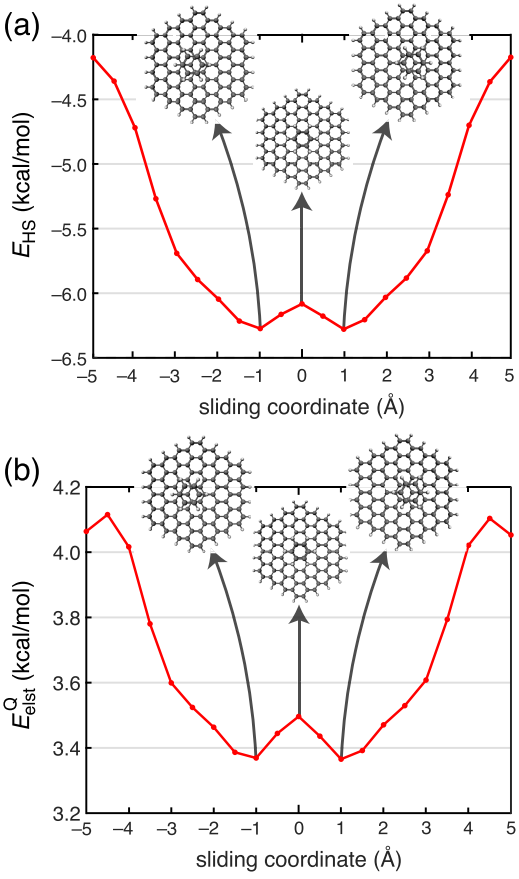


Figure 2. (a) Total HS model potential and (b) its electrostatic component, for lateral displacement of C_6H_6 atop $C_{96}H_{24}$ in a cofacial configuration at 3.4 Å separation. This figure should replace Figure 12a of ref 1.

A final topic of discussion is the use of $\delta^{\pm}$ to label the cartoons in Figure 6 of ref 1. This is potentially misleading, as it may suggest that C_6F_6 has a positive charge density on its π faces, which is absurd. Rather, these diagrams are intended to convey the change in *quadrupolar* electrostatics that occurs when one C_6H_6 monomer is replaced by C_6F_6, since the quadrupole−quadrupole interactions in $(C_6H_6)\cdots(C_6F_6)$ are attractive in the face-to-face orientation. A similar diagram, coloring C_6F_6 opposite to C_6H_6, can be found in Hunter et al.[6] Although these cartoon charge distributions are widely used in discussing π-stacking,[6,12,13] we have elsewhere suggested that they are misleading and that their use ought to be discontinued.[7]

We thank Prof. Steven Wheeler (University of Georgia) for bringing these issues to our attention based on his independent implementation of the HS model.

https://doi.org/10.1021/acs.accounts.5c00107
Acc. Chem. Res. 2025, 58, 1051−1053

AUTHOR INFORMATION

Corresponding Author

John M. Herbert − *Department of Chemistry & Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0002-1663-2278; Email: herbert@chemistry.ohio-state.edu

Authors

Kevin Carter-Fenk − *Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States;* orcid.org/0000-0001-8302-4750

Ka Un Lao − *Department of Chemistry, Virginia Commonwealth University, Richmond, Virginia 23284, United States;* orcid.org/0000-0002-3993-536X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.accounts.5c00107

REFERENCES

(1) Carter-Fenk, K.; Lao, K. U.; Herbert, J. M. Predicting and understanding non-covalent interactions using novel forms of symmetry-adapted perturbation theory. *Acc. Chem. Res.* **2021**, *54*, 3679−3690.

(2) Hunter, C. A.; Sanders, J. K. M. The nature of $\pi-\pi$ interactions. *J. Am. Chem. Soc.* **1990**, *112*, 5525−5534.

(3) Hunter, C. A.; Singh, J.; Thornton, J. M. $\pi-\pi$ interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. *J. Mol. Biol.* **1991**, *218*, 837−846.

(4) Caillet, J.; Claverie, P. Theoretical evaluation of the intermolecular interaction energy of a crystal: Application of the analysis of crystal geometry. *Acta Cryst. A* **1975**, *31*, 448−461.

(5) Battaglia, M. R.; Buckingham, A. D.; Williams, J. H. The electric quadrupole moments of benzene and hexafluorobenzene. *Chem. Phys. Lett.* **1981**, *78*, 421−423.

(6) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. Aromatic interactions. *J. Chem. Soc., Perkin Trans.* **2001**, *2*, 651−669.

(7) Schramm, B.; Gray, M.; Herbert, J. M. Substituent and heteroatom effects on $\pi-\pi$ interactions: Evidence that parallel-displaced π-stacking is not driven by quadrupolar electrostatics. *J. Am. Chem. Soc.* **2025**, *147*, 3243−3260.

(8) Carter-Fenk, K.; Herbert, J. M. Reinterpreting π-stacking. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24870−24886.

(9) Carter-Fenk, K.; Herbert, J. M. Electrostatics does not dictate the slip-stacked arrangement of aromatic $\pi-\pi$ interactions. *Chem. Sci.* **2020**, *11*, 6758−6765.

(10) Herbert, J. M. Neat, simple, and wrong: Debunking electrostatic fallacies regarding noncovalent interactions. *J. Phys. Chem. A* **2021**, *125*, 7125−7137.

(11) Gray, M.; Herbert, J. M. Origins of offset-stacking in porous frameworks. *J. Phys. Chem. C* **2023**, *127*, 2675−2686.

(12) Martinez, C. R.; Iverson, B. L. Rethinking the term "pi-stacking. *Chem. Sci.* **2012**, *3*, 2191−2201.

(13) Fagnani, D. E.; Sotuyo, A.; Castellano, R. K. $\pi-\pi$ interactions. In *Comprehensive Supramolecular Chemistry II*, Vol. *1*; Elsevier: Oxford, UK, 2017; pp 121−148.

https://doi.org/10.1021/acs.accounts.5c00107
Acc. Chem. Res. 2025, 58, 1051−1053