Accurate Prediction of Noncovalent Interaction Energies with the Effective Fragment Potential Method: Comparison of Energy Components to Symmetry-Adapted Perturbation Theory for the S22 Test Set

Joanna C. Flick,[†] Dmytro Kosenkov,[†] Edward G. Hohenstein,[‡] C. David Sherrill,[‡] and Lyudmila V. Slipchenko*,[†]

[†]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States
[‡]Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States

S *Supporting Information*

ABSTRACT: Noncovalent interactions play an important role in the stabilization of biological molecules. The effective fragment potential (EFP) is a computationally inexpensive ab initio-based method for modeling intermolecular interactions in noncovalently bound systems. The accuracy of EFP is benchmarked against the S22 and S66 data sets for noncovalent interactions [Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985; Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427]. The mean unsigned error (MUE) of EFP interaction energies with respect to coupled-cluster singles, doubles, and perturbative triples in the complete basis set limit [CCSD(T)/CBS] is 0.9 and 0.6 kcal/mol for S22 and S66, respectively, which is similar to the MUE of MP2 and SCS-MP2 for the same data sets, but with a greatly reduced computational expense. Moreover, EFP outperforms classical force fields and popular DFT functionals such as B3LYP and PBE, while newer dispersion-corrected functionals provide a more accurate description of noncovalent interactions. Comparison of EFP energy components with the symmetry-adapted perturbation theory (SAPT) energies for the S22 data set shows that the main source of errors in EFP comes from Coulomb and polarization terms and provides a valuable benchmark for further improvements in the accuracy of EFP and force fields in general.

INTRODUCTION

Noncovalent interactions play an important role in the stabilization of biological macromolecules. The secondary structure of nucleic acids, protein folding, enzyme—substrate recognition and binding, and enzymatic catalysis are controlled by these interactions.[1−5] Noncovalent interactions are dominated by electrostatic and dispersion forces. Electrostatic interactions have pronounced directional character; they provide the main energetic contribution in hydrogen bonding (HB). Being widespread in many biological systems, hydrogen bonding governs stability of Watson—Crick pairs of DNA. Electrostatic interactions are also strong in charged (e.g., anionic carboxylate, cationic ammonium) and polarized (e.g., hydroxo-, carboxy-, and amino-) groups of DNA and proteins. Dispersion interactions (that is, the London forces) are typically weaker and have less directional character than hydrogen bonding. Dispersion makes a major contribution to stacking interactions. Stacking is substantial in aromatic and other systems with delocalized π-electrons; stacking between DNA bases in the double helix provides an important stabilizing factor of the secondary structure of DNA.[6]

The proper characterization of noncovalent interactions is a challenging task for computational chemistry. For example, although hydrogen bonding is predominantly an electrostatic interaction, a more detailed study reveals that it is a result of the interplay of electrostatic, polarization, exchange-repulsion,

charge transfer, and even dispersion interactions.[7] The dispersion interaction, predominant in π-stacking, can be properly described only if electron correlation effects are accounted for. Thus, computationally intensive high-level ab initio techniques such as coupled cluster methods or methods that incorporate special corrections accounting for dispersion such as in dispersion-corrected density functionals are required for accurate prediction of the noncovalent interactions. Fragmentation approaches[8−17] can often provide a computationally inexpensive alternative to pure ab initio methods without compromising the accuracy of predictions. An even more computationally affordable approach is to use empirically parametrized molecular mechanics force fields (FF). However, they require empirical information, heavily depend on fitted parameters, and often have to be reparameterized to be applicable to new molecules.

In the present inquiry, the general effective fragment potential (EFP) method[18,19] is employed for the description of intermolecular interactions. The EFP method is an ab initio-based potential and is a more sophisticated approach than parametrized molecular mechanics models. EFP is designed as a computationally inexpensive way of modeling intermolecular interactions in noncovalently bound systems. The absence of

fitted parameters and a natural partitioning of the interaction energy into electrostatic (Coulomb), polarization (induction), dispersion, and exchange-repulsion terms make it an attractive choice for analysis and interpretation of the intermolecular forces. The EFP Hamiltonian is pairwise; however, the leading many-body effects are included through a self-consistent treatment of polarization.

Previously, EFP has been successfully applied for investigation of the noncovalent interactions in dimers of benzene[20] and benzene derivatives,[21,22] styrene dimers,[23] and DNA base pairs,[19,24] water–benzene,[25] water–methanol,[26] and water–alanine complexes.[27,28] In these studies, EFP has been shown to provide accurate results as compared to correlated ab initio methods. However, to date, no systematic investigation of the accuracy of EFP has been performed. This work fills this gap by presenting a benchmark of the EFP performance for the S22 and S66 data sets for noncovalent interactions. The S22 data set proposed by Hobza and co-workers[29] consists of 22 dimers of small representative fragments of biomacromolecules containing only C, N, O, and H atoms with single, double, triple bonds, and conjugated cycles. The complexes are divided into three groups based on the interaction types: hydrogen bonded (HB), dispersion-dominated, and mixed. Recently, more accurate complete basis set (CBS) extrapolations of the coupled-cluster with single, double, and perturbative triple excitations (CCSD-(T))[30] interaction energies for these dimers have been published.[31,32] S66[33] is a newer and more balanced data set that includes several variations of each interaction type, for example, complexes with both single and double (cyclic) hydrogen bonds, aromatic–aromatic (stacking), aromatic–aliphatic, and aliphatic–aliphatic interactions.

For complexes from the S22 data set, components of the EFP interaction energies are also compared to the energy decomposition provided by symmetry-adapted perturbation theory (SAPT).[34,35] This comparison provides a fair assessment of the accuracy of different EFP terms. This analysis will be an important benchmark for future developments of EFP and force fields in general.

■ THEORETICAL METHODS AND COMPUTATIONAL DETAILS

There are four interaction terms in the general EFP model potential[18,19] (the general EFP potential has been originally called EFP2 to be distinguished from the water potential EFP1[36–38]), each of which may be thought of as a truncated expansion, Coulomb (electrostatic), induction (polarization), exchange repulsion, and dispersion:

$$E^{\text{EFP–EFP}} = E_{\text{coul}} + E_{\text{pol}} + E_{\text{disp}} + E_{\text{exrep}} \qquad (1)$$

The terms in the EFP potential may be grouped into long-range, $(1/R)^n$ distance dependent, and short-range interactions that decay exponentially. The Coulomb, induction, and dispersion are long-range interactions, whereas the exchange repulsion and damping terms are short-range. The EFP method has been described in detail elsewhere.[18,19,37–39]

The EF potential includes: (i) multipoles (produced by Stone's distributed multipolar analysis[40,41]) for Coulomb and polarization terms; (ii) static polarizability tensors centered at localized molecular orbital (LMO) centroids (obtained from coupled-perturbed Hartree–Fock calculations), which are used for calculations of polarization; (iii) dynamic polarizability tensors centered on the LMOs that are generated by time-

dependent HF calculations and used for calculations of dispersion; and (iv) the Fock matrix, basis set, and localized orbitals needed for the exchange-repulsion term. In sum, all of the EFP parameters are obtained from ab initio calculations on an isolated fragment and contain no empirically fitted parameters. The Coulomb part of these potentials, that is, electrostatic multipoles obtained with analytic Stone DMA, was generated using HF/6-31+G(d)[42–44] and HF/6-31G(d) for nonaromatic and aromatic molecules, respectively. The rest of the potential, that is, static and dynamic polarizability tensors, wave function, Fock matrix, etc., were obtained with the 6-311++G(3df,2p) basis set.[44–46] To account for the short-range charge-penetration effects, overlap-based electrostatic and dispersion screenings as well as Gaussian-like polarization screening were employed.[47]

EFP energy components are compared against those computed by symmetry-adapted perturbation theory (SAPT).[34,35] SAPT partitions the Hamiltonian as

$$H = F_A + F_B + W_A + W_B + V \qquad (2)$$

where F_A is the Fock operator for monomer A, W_A is the fluctuation potential for monomer A (the difference between the exact Coulomb operator and the Fock operator within monomer A), F_B and W_B are corresponding terms for monomer B, and V contains all intermolecular terms. SAPT then evaluates energy components through various orders of the three perturbations, W_A, W_B, and V. Here, we use high-order SAPT theory (including second- and third-order terms) designated SAPT2+(3)[48] along with an aug-cc-pVTZ basis set to obtain reliable energy components against which we can test the corresponding EFP energy components. We have grouped the terms included as follows:

$$E_{\text{electrostatic}} = E_{\text{elst,resp}}^{(10)} + E_{\text{elst,resp}}^{(12)} + E_{\text{elst,resp}}^{(13)}$$

$$E_{\text{exchange}} = E_{\text{exch}}^{(10)} + E_{\text{exch}}^{(11)} + E_{\text{exch}}^{(12)}$$

$$E_{\text{induction}} = E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + {}^t E_{\text{ind}}^{(22)} + {}^t E_{\text{exch-ind}}^{(22)} \\ + \delta E_{\text{HF}}^{(2)}$$

$$E_{\text{dispersion}} = E_{\text{disp}}^{(20)} + E_{\text{disp}}^{(30)} + E_{\text{disp}}^{(21)} + E_{\text{disp}}^{(22)} + E_{\text{exch-disp}}^{(20)} \qquad (3)$$

where the two numbers in parentheses are the perturbation orders in V and in $W = W_A + W_B$, respectively. The meaning of the individual terms is discussed in refs 34,35,48. SAPT computations were carried out using a developers' version of the PSI4 program.[49]

On the basis of the energy decompositions at the SAPT2+(3)/aug-cc-pVTZ level of theory (see the Supporting Information), we find that the original, intuitive assignment[29] of the S22 test set dimers into hydrogen bonding, dispersion-dominated, and mixed influence categories should be reassigned in three cases. Here, we use the rules proposed by Hobza and co-workers,[50] that a complex should be considered dispersion-dominated if the dispersion component is at least twice as large as the electrostatic component, and vice versa for electrostatic dominated complexes. The remainder are categorized as of mixed type. Applying this rule to the SAPT2+(3)/aug-cc-pVTZ energy components, the T-shaped benzene dimer is dispersion-dominated rather than mixed influence, while the stacked uracil dimer and adenine–thymine dimer are more mixed influence than dispersion-dominated. Fortunately, these reassignments agree with those of Hobza and

co-workers for the S22 molecules at their equilibrium geometries based on SAPT-DFT computations[50] except for the benzene–HCN complex, which they categorize as electrostatics dominated, but which we categorize as mixed influence. For the purposes of this study, we have retained the original categories, but it may be useful in future work to begin using the revised assignments.

Geometries from the S22 data set were used in all SAPT calculations. Geometries for the EFP fragments were prepared as follows. Geometries of water, methane, formamide, ammonia, benzene, and phenol were optimized at the MP2[51]/cc-pVTZ[52,53] level of theory. S22 data set geometries[29] were used for other fragments, that is, MP2/cc-pVTZ geometries for adenine, aminopyridine, hydrogen cyanide, indole, pyrazine, pyridoxine, thymine, and uracil; CCSD(T)/cc-pVTZ geometry for formic acid; and CCSD(T)/cc-pVQZ geometries for ethane and ethyne. Two versions of monomer geometries were used for thymine and adenine: one corresponding to hydrogen bonded and another to stacking dimer structure. For the S66 data set, all MP2/cc-pVTZ gas-phase-optimized geometries were used except ethane and ethyne that were taken from the S22 data set. Two versions of acetic acid (AcOH) and acetamide (AcNH$_2$) monomers were used, one corresponding to gas-phase monomer and another to S66 H-bonded dimer. Geometries of the monomers are kept frozen in EFP.

EFP potentials for molecules from S22 and S66 data sets are included in EFP libraries of the GAMESS[54,55] and Q-CHEM[56,57] packages. All EFP calculations were performed in GAMESS.

■ RESULTS AND DISCUSSION

Geometries of S22 Complexes. While the monomer molecules (i.e., EFP fragments) are kept frozen, the intermolecular distances between monomers in S22 and S66 complexes were reoptimized with the EFP potential. The accuracy of EFP in predicting geometries of S22 complexes is analyzed below. The same trends are observed for geometries of the S66 complexes.

The differences between the distances in ab initio geometries and those optimized with EFP can be used as a measure of quality of EFP geometry prediction (bearing in mind that some of the S22 geometries were determined at more robust levels of ab initio theory than others).[29] The geometry data are summarized in Table 1. Geometries of the complexes are shown in the Supporting Information.

In HB complexes, distances between atoms involved in hydrogen bonding are measured. EFP overestimates intermolecular separations by 0.01–0.17 Å. The longer distances predicted by EFP imply that the intermolecular attractive forces in HB complexes are underestimated. Complexes with stronger intermolecular interactions, for example, uracil dimer and formic acid dimer, tend to provide larger discrepancies in distances. This may suggest that the accuracy of EFP deteriorates when the short-range quantum effects become important.

Geometries in nonaromatic dispersion-dominated complexes (methane and ethene dimers) are well predicted by EFP, with slightly shorter distances of −0.06 and −0.07 Å, respectively. For aromatic complexes, two parameters are calculated to describe the change in the geometry: a shift of the parallel-displaced aromatic rings R_2 and a separation between the ring planes R_1. In all cases, the planar separation between the rings is

Table 1. Reference ab Initio and EFP-Optimized Intermolecular Distances in S22 Dimers

	distance[a]	ab initio, Å[b]	EFP, Å	difference, Å
Hydrogen-Bonded Complexes				
ammonia dimer	N1···N5	3.16	3.17	0.01
water dimer	O1···O4	2.91	2.95	0.04
formic acid dimer	O2···O8	2.67	2.83	0.17
formamide dimer	O2···N9	2.86	2.97	0.11
uracil H-bonded dimer	N1···O23	2.80	2.96	0.16
2-pyridoxine 2-aminopyridine	N1···N15	2.90	2.98	0.09
adenine–thymine WC[e]	N1···N20	2.86	2.97	0.10
MUE[f]				0.10
Dispersion-Dominated Complexes				
methane dimer	C1···C6	3.72	3.66	−0.06
ethene dimer	C1···C7	3.84	3.77	−0.07
benzene–methane	C1···RD[c]	3.72	3.95	0.23
benzene stack	R_1/R_2[d]	3.36/1.70	3.80/1.45	0.44/−0.25
pyrazine dimer	R_1/R_2[d]	3.30/1.22	3.61/0.97	0.31/−0.24
uracil stack	R_1/R_2[d]	3.12/0.54	3.27/0.62	0.15/0.08
indole–benzene stack	R_1/R_2[d]	3.25/1.28	3.60/1.25	0.35/−0.03
adenine–thymine stack	R_1/R_2[d]	3.15/0.34	3.41/0.40	0.25/0.05
MUE				0.20[g]
Mixed Complexes				
ethene–ethyne	C8···C2	3.88	3.97	0.10
benzene–water	O1···RD[c]	3.41	3.49	0.08
benzene–ammonia	N···RD[c]	3.57	3.76	0.18
benzene–HCN	C14···RD[c]	3.39	3.57	0.18
benzene dimer T-shaped	C1··· RD[c]	3.51	3.81	0.30
indole–benzene T-shaped	N21···RD[c]	3.24	3.43	0.19
phenol dimer	O7···O20	2.89	3.01	0.12
MUE				0.17
overall MUE				0.17

[a]Atom numbering in accordance with Figure S1 (see the Supporting Information). [b]Reference S22 geometries obtained with MP2/cc-pVTZ for adenine, aminopyridine, benzene, formamide, hydrogen cyanide, indole, phenol, pyrazine, pyridoxine, thymine, and uracil; CCSD(T)/cc-pVTZ for formic acid and methane; and CCSD(T)/cc-pVQZ for ammonia, ethene, ethyne, and water.[29] [c]RD: The distance between the corresponding atom and the plane of the benzene ring. [d]Dimer geometry is characterized by two parameters: distance between planes of the rings R_1 and displacement of the ring centers R_2. [e]Watson–Crick dimer. [f]Mean unsigned error (MUE). [g]MUE for distances between centers of masses of rings.

overestimated, with differences ranging from 0.15 to 0.44 Å for the benzene dimer. The displacement distance R_2 is well reproduced in uracil dimer, indole–benzene, and adenine–thymine complexes, while R_2 is underestimated by ∼0.25 Å for benzene and pyrazine dimers.

In the mixed complexes, the differences in the geometries are about the same as in the HB complexes and do not exceed 0.18 Å with the exception of the T-shaped benzene dimer where it is 0.3 Å. Note that the geometry of the T-shaped benzene dimer used in the S22 set is constrained to the C_{2v} symmetry and is not a true minimum; the real minimum possesses lower C_s

Table 2. Mean Unsigned Errors (MUE) and Root Mean Square Errors (RMSE) of the Total Interaction Energies (kcal/mol) for Hydrogen-Bonded (HB), Dispersion-Dominated (DISP), and Mixed (MIXED) Complexes of the S22 Data Set by EFP, Molecular Mechanics Force Fields, Semiempirical, HF, DFT, and ab Initio Methods

	MUE(HB)	MUE(DISP)	MUE(MIXED)	MUE(overall)[a]	MMUE[b]	RMSE
EFP	1.97	0.48	0.34	0.91	0.93	2.06
Force Fields[c]						
Amber	4.79	0.98	0.98	2.16	2.22	
OPLSAA	4.59	1.04	0.57	2.02	2.07	
MMFF94	3.75	0.88	0.59	1.70	1.74	
Semiempirical Methods[d]						
PM6				3.34		4.16
PM6-DH+				0.60		0.80
AM1-D3H4				1.25		1.76
DFTB-D3H4				0.79		0.97
HF and DFT[e]						
HF	3.29	7.24	3.15		4.56	
B3LYP	1.77	6.22	2.64		3.54	
PBE	1.13	4.53	1.66		2.44	
M05	1.26	3.16	1.09		1.84	
M06	0.89	0.99	0.67		0.85	
M06-2X	0.73	0.36	0.32		0.47	
BLYP-D3				0.23[g]		
ωB97X-D				0.22[h]		
Correlated Methods[f]						
MP2	0.24	1.69	0.61	0.88	0.86	1.05
SCS-MP2	1.54	0.55	0.37	0.80	0.82	0.96
SCS-CCSD	0.40	0.23	0.08	0.24	0.24	0.54
10%[i]	1.38	0.48	0.39	0.74	0.74	

[a]MUE calculated for all complexes of the S22 data set. [b]MMUE: average of MUEs for HB, DISP, and MIXED dimers. [c]Reference 70. [d]Reference 65. [e]Reference 66. [f]Reference 31. [g]Reference 59. [h]Reference 60. [i]10% values of the average interaction energies.

Table 3. Mean Unsigned Errors (MUE) and Root Mean Square Errors (RMSE) of the Total Interaction Energies (kcal/mol) for Hydrogen-Bonded (HB), Dispersion-Dominated (DISP), and Mixed (MIXED) Complexes of the S66 Data Set by EFP, Semiempirical, DFT, and ab Initio Methods

	MUE(HB)	MUE(DISP)	MUE(MIXED)	MUE(overall)[a]	RMSE (overall)[a]
EFP	0.79	0.65	0.35	0.61	0.78
Semiempirical Methods[b]					
PM6				2.67	3.02
PM6-DH+				0.62	0.82
AM1-D3H4				0.96	1.35
DFTB-D3H4				0.52	0.67
DFT[c]					
B3LYP	1.47	5.17	3.00	3.22	
PBE	0.74	3.63	1.94	2.11	
M06-2X	0.24	0.35	0.25	0.28	
BLYP-D3	0.40	0.46	0.31	0.39	
B97-D3	0.26	0.32	0.19	0.26	
Correlated Methods[d]					
MP2				0.45	0.69
SCS-MP2				0.74	0.87
SCS-CCSD				0.15	0.25
10%[e]	0.89	0.36	0.36	0.55	

[a]MUE and RMSE calculated for all complexes of the S66 data set. [b]Reference 65. [c]def2-QZVP basis, from ref 62. [d]CBS basis, from ref 33. [e]10% values of the average interaction energies.

symmetry. For comparison purposes, EFP optimization of the T-shaped benzene dimer was also constrained to the C_{2v} symmetry.[58]

The overall tendency of EFP is to overestimate the distances in the noncovalent complexes. This suggests underestimating attractive forces (Coulomb, polarization, dispersion) or over-estimating repulsion. The following section will shed more light on the origin of the EFP inaccuracies. Despite observed discrepancies, the overall accuracy of EFP is very reasonable; mean unsigned error (MUE) of geometries of HB complexes is 0.10 Å. MUEs for dispersion-dominated and mixed complexes are 0.20 and 0.17 Å, respectively.

Prediction of Total Interaction Energies. The MUE and root-mean-square error (RMSE) of the EFP interaction energies with respect to recently revised CCSD(T)/CBS binding energies for the S22 test set[31] are shown in Table 2 for each subset (hydrogen-bonded, dispersion-dominated, and mixed) and for the test set as a whole. The MUE and RMSE of EFP for the S66 data set are provided in Table 3. The EFP energies used in these comparisons are obtained at the EFP-optimized geometries. The complete list of the EFP interaction energies for the S22 and S66 data sets is available in the Supporting Information. For reference, MUEs and RMSEs of molecular mechanics force fields (Amber, OPLSAA, MMFF94), semiempirical, Hartree—Fock (HF), and popular ab initio methods, as well as several density functional (DFT) methods are also shown in Tables 2 and 3. Because the average interaction energies of HB, dispersion-dominated, and mixed complexes are significantly different (e.g., hydrogen-bonded complexes tend to be much more tightly bound than the other types of complexes), we also list 10% of the value of the average interaction energies that provides an easy estimate of the relative accuracy of the method. Additionally, comparison of the accuracy of EFP and FF methods for S22 data set is shown in Figure 1.
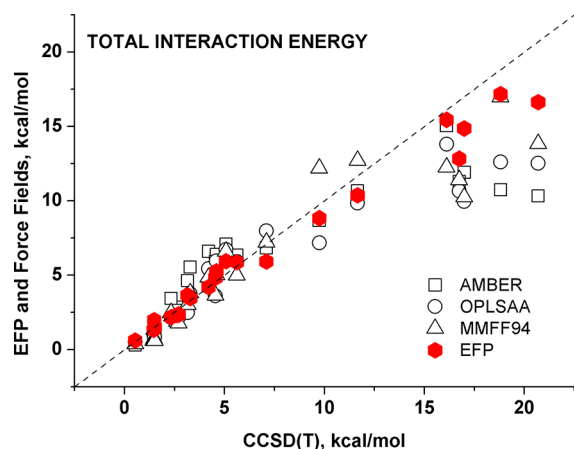

Figure 1. Total interaction energies for S22 data set dimers calculated using EFP and molecular mechanics force fields (AMBER, OPLSAA, MMFF94) from ref 70 as compared to CCSD(T)/CBS data from ref 31. (The detailed information is provided in Table S2 in the Supporting Information.)

Methods that do not account for long-range correlation, such as HF and standard semiempirical or density functional techniques, provide a very poor description of dispersion-dominated complexes, with relative errors often exceeding 100%. On the other hand, wave function correlated methods, such as MP2 and CCSD, can give poor quantitative results for the dispersion energy, which can result in an imbalanced treatment of dispersion-dominated and mixed complexes. As has been demonstrated on multiple occasions and shown in Tables 2 and 3, the accuracy of the uncorrelated (HF or semiempirical) or locally correlated (Kohn—Sham DFT with B3LYP, PBE, etc., functionals) methods can be significantly improved by introducing dispersion corrections. For example, the overall S22 MUE values for ωB97X-D and BLYP-D3 dispersion-corrected functionals is 0.22—0.23 kcal/mol.[59,60] B3LYP-D3 and B97-D3 functionals also show an excellent performance with S22 and S66 MUEs below 0.4 kcal/mol.[61,62]

Semiempirical methods or density functional tight binding (DFTB) models have been also extended to provide accurate description of intermolecular interactions.[63−65] For example, parametrization of H-bonded and dispersion interactions by Rezac and Hobza[65] (so-called D3H4 correction) provided a significant improvement of the accuracy of semiempirical methods and resulted in MUE in the range of 0.4−1.4 kcal/mol. The same correction applied to DFTB lowered its MUE values to 0.5 and 0.8 kcal/mol for the S66 and S22 data sets, respectively. A more balanced description of different kinds of noncovalent bonding can be also achieved through extensive parametrization as is done in the M06-2X functional[66] (although it should be noted that M06-2X does not perform as well for complexes with larger intermolecular separations).[67]

Spin-component-scaled (SCS) corrections have been advocated as a way to improve the treatment of dispersion in wave function methods without incurring additional computational cost. However, while SCS-CCSD[68] provides overall accurate and balanced description of noncovalent interactions,[31,33] on average the SCS version of MP2[69] is only slightly better than the original MP2 for S22[31] (0.80 kcal/mol versus 0.88 kcal/mol) and performs worse than MP2 for S66 (0.74 kcal/mol versus 0.45 kcal/mol).[33]

Performance of classical force fields in description of noncovalent interactions in the S22 data set is surprisingly poor, with an overall MUE of ∼2 kcal/mol (see Table 2). In contrast to the wave function and DFT methods, classical force fields have larger relative errors in the HB than dispersion-dominated or mixed complexes. The analysis of this behavior is provided in the following paragraphs.

The overall MUEs of EFP for the S22 and S66 data sets are 0.91 and 0.61 kcal/mol, respectively, which corresponds to 11−12% relative error in interaction energies. For the S22 data set, the relative accuracy of EFP energies is 9−10% for the mixed and dispersion-dominated complexes and is slightly worse (14%) for the HB dimers, showing generally a balanced description of different kinds of interactions. For the S66 data set, the relative accuracy is 9−10% for HB and mixed dimers and 18% for dispersion-dominated complexes. In S22, the main problem for EFP originates from strongly interacting HB complexes such as DNA base pairs. HB complexes in S66 have on average weaker interactions than in S22 (10% values of the average interaction energies are 1.38 and 0.89 kcal/mol, respectively) and, as a result, are described by EFP more accurately. A larger relative error in dispersion-dominated dimers in S66 is associated with overestimation of dispersion interactions in complexes with pentane.

In terms of overall accuracy, EFP is comparable to MP2 and SCS-MP2 wave function methods, M06 density functional, and semiempirical and DFTB methods with corrections for hydrogen bonding and dispersion. EFP significantly outperforms functionals that do not include dispersion corrections, such as B3LYP or PBE, but is less accurate than dispersion-corrected functionals or new functionals by Truhlar.[66] EFP outperforms classical force fields whose errors are about twice larger.[70] Explicit comparison of the EFP and FF interaction energies in Figure 1 shows that EFP is consistently more accurate for the S22 data set.

The computational cost of EFP allows one to routinely perform molecular dynamics or Monte Carlo simulations of bulk systems for tenths of picoseconds.[71] The scaling of EFP is formally quadratic with the number of fragments in the system; the majority of computational cost for calculation of the
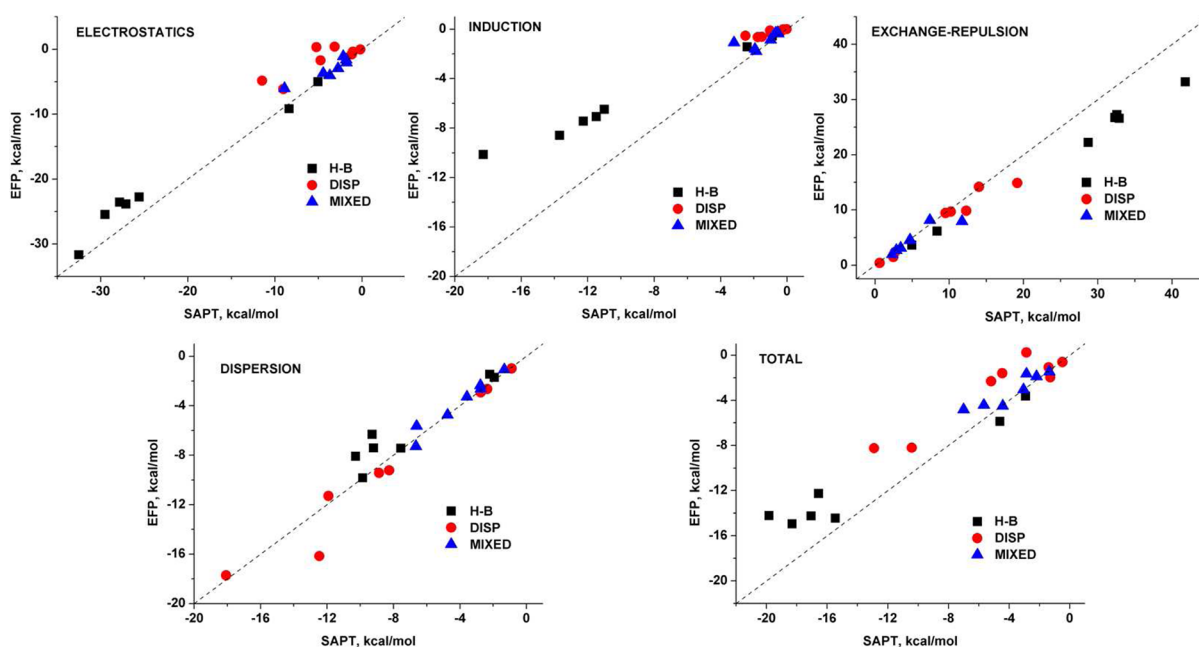
Figure 2. Comparison of EFP and SAPT energy components for the S22 data set dimers: hydrogen-bonded (H-B), dispersion-dominated (DISP), and mixed (MIXED) dimers.

interaction energy or gradient for each pair of fragments (i.e, a prefactor for N^2 scaling) comes from the exchange-repulsion term and is higher than the corresponding cost (prefactor) in classical force fields. Thus, EFP has a lower scaling and is several orders of magnitude computationally cheaper than the wave function or DFT methods. The computational cost of semiempirical and DFTB methods may be comparable with that of EFP.

EFP Energy Components. Evaluation of the accuracy of the EFP energy components, that is, electrostatics, polarization (induction), exchange-repulsion, and dispersion, provides a means to characterize main sources of error in the interaction energies and equilibrium geometries and to target the problematic terms in future developments. Comparison of the EFP and SAPT energy components for the S22 data set is shown in Figure 2. Both SAPT and EFP calculations were performed at the S22 geometries of the dimers.[29] The electrostatic components in the S22 dimers range from ~0 to 35 kcal/mol. EFP underestimates (gives less negative) electrostatic energies in almost all strongly bound HB complexes by 3−4 kcal/mol (i.e., 10−15% of the total electrostatic energies). This is probably due to insufficient capture of the charge-penetration. At short interfragment separations, the classical multipole approximation fails because it does not recover charge delocalization provided by the electronic wave function. The electrostatic term in EFP is augmented by the screening function that is based on the interfragment overlap integrals and accounts for some penetration effects.[47] However, it apparently underestimates charge-penetration in cases of strong electrostatic interactions. Electrostatics is also underestimated in the dispersion-dominated aromatic complexes; the multipole approximation provides repulsive Coulomb energies, but as evidenced by the SAPT computations they are actually attractive due to charge-penetration effects. The electrostatic damping employed in EFP recovers only part of the charge-penetration effect, leading to errors in the electrostatic term of a few kcal/mol in some cases;

because of the small magnitude of the electrostatic terms in the dispersion-dominated complexes, this leads to rather large relative errors. In mixed complexes, the electrostatic energies are described relatively accurately because they are neither large in magnitude nor dominated by charge-penetration. Because EFP electrostatic energies are generally underestimated, the complexes prefer larger intermolecular separations. At larger separations, the accuracy of the EFP electrostatic term improves as the importance of the short-range, overlap-dependent charge-penetration contribution diminishes. To summarize, the complicated nature of the short-range electrostatic interactions is challenging for description by classical or semiclassical approximations. It is not surprising then that the HB complexes dominated by electrostatics are poorly described by force fields. On the other hand, electrostatics is easily and accurately captured by almost any ab initio method.

Similarly to Coulomb interactions, the induction (polarization) energies in EFP are also underestimated. Discrepancies in relative energies are the largest for the dispersion-dominated dimers, but due to small magnitudes of the polarization energies, the absolute errors are not large (the largest error for the adenine−thymine stacked dimer is ~2.0 kcal/mol). The main reason of these errors is probably in the inaccuracy of the field created by the electrostatic multipoles. As discussed in previous paragraphs, the multipole approximation is qualitatively incorrect for the dispersion-dominated systems. Thus, the field due to these multipoles creates induced dipoles that are too weak, leading to an underestimation of the polarization energy. EFP polarization energies are also significantly underestimated in the strongly bound HB complexes, with typical errors of 4−5 kcal/mol and even 8 kcal/mol for the formic acid dimer. In addition to the previously discussed deficiencies of the multipole approximation, one more reason for this troublesome behavior is the neglect of the charge-transfer term in the EFP potentials employed in this work. It should be noted that the induction term in SAPT effectively includes charge-transfer (as long as a dimer-centered basis is

used).[72] The charge-transfer terms are expected to be negligible in the dispersion-dominated complexes but may be of a significant value in the HB complexes. Thus, adding the charge-transfer term may improve the overall accuracy of EFP. Similarly to the Coulomb term, polarization is reasonably well described in the mixed complexes.

The exchange-repulsion energies are generally underestimated by EFP (i.e., they are not repulsive enough), and the errors are roughly proportional to the energy. Thus, the typical relative errors are 10−20%. The absolute errors are the largest for the HB dimers and somewhat cancel the errors in the electrostatic and polarization terms. The underestimation of the exchange-repulsion energies may originate in neglect of the correlation effects in EFP that, based on the SAPT decomposition, provide about 20% of the additional repulsion energy.

The EFP dispersion energy is in very good agreement with the SAPT dispersion. The largest error of 3.7 kcal/mol occurs for the indole−benzene stacked dimer. In 18 dimers, the errors are within 1 kcal/mol. Excellent accuracy of EFP in the prediction of the dispersion interactions is in striking contrast to the nuisances that dispersion causes to many ab initio and DFT methods.

Comparison of the total EFP and SAPT energies reflects the EFP inaccuracies in the polarization and Coulomb terms. The interactions in the strongly bound HB complexes are underestimated, as are the energies in the dispersion-dominated dimers. The accuracy of the interaction energies in the mixed dimers is good. The underestimation of the attractive electrostatic and polarization terms leads to overestimation of the optimal intermolecular separations in EFP.

CONCLUSIONS

The accuracy of the general effective fragment potential method was tested on the S22 and S66 data sets for noncovalent interactions. For S22 complexes, EFP predicts slightly longer intermolecular separations (with a MUE of 0.20 Å) and the interaction energies with a MUE of 0.9 kcal/mol. For the larger S66 set, EFP interaction energies are predicted with MUE of 0.6 kcal/mol. On the basis of the mean errors, the accuracy of EFP is similar to the accuracy of MP2 and SCS-MP2, while it is superior to the accuracy of classical force fields. The analysis of the energy components suggests that the main sources of error in EFP originate from the underestimation of the polarization and Coulomb terms where short-range charge-penetration effects are not fully captured. Somewhat unexpectedly, the accurate treatment of these terms is challenging and crucial not only for the HB dimers but for the dispersion-dominated complexes as well. Underestimation of the Coulomb and polarization energies in EFP is partially compensated by underestimation of exchange-repulsion term, resulting in accurate total interaction energies.

The importance of the short-range charge-penetration effects was recently highlighted in a study of substituted benzene dimers, which singled them out as the only way to understand otherwise counterintuitive behavior of the electrostatic component of substituent effects.[73] The provided analysis suggests that a more accurate and efficient implementation of the short-range electrostatic and polarization damping terms, as well as inclusion of a charge-transfer term, may improve the accuracy of the EFP method.

ASSOCIATED CONTENT

S **Supporting Information**

SAPT interaction energy decomposition (Figure S1), EFP-optimized structures of S22 complexes (Figure S2), detailed comparison of SAPT and EFP energy components (Table S1), and EFP interaction energies of S22 and S66 complexes (Tables S2 and S3). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: lslipchenko@purdue.edu.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) 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.

(2) Privalov, P. L.; Makhatadze, G. I. Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. *J. Mol. Biol.* **1992**, *224*, 715−723.

(3) Hunter, C. A. Meldola lecture. The role of aromatic interactions in molecular recognition. *Chem. Soc. Rev.* **1994**, *23*, 101−109.

(4) Anbarasu, A.; Anand, S.; Babu, M. M.; Sethumadhavan, R. Investigations on C-H...pi interactions in RNA binding proteins. *Int. J. Biol. Macromol.* **2007**, *41*, 251−259.

(5) Gazit, E. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* **2007**, *36*, 1263−1269.

(6) Černý, J.i.; Kabeláč, M.; Hobza, P. Double-helical → ladder structural transition in the B-DNA is induced by a loss of dispersion energy. *J. Am. Chem. Soc.* **2008**, *130*, 16055−16059.

(7) Thanthiriwatte, K. S.; Hohenstein, E. G.; Burns, L. A.; Sherrill, C. D. Assessment of the performance of DFT and DFT-D methods for describing distance dependence of hydrogen-bonded interactions. *J. Chem. Theory Comput.* **2010**, *7*, 88−96.

(8) Fedorov, D. G.; Kitaura, K. Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *J. Phys. Chem. A* **2007**, *111*, 6904−6914.

(9) Fedorov, D. G.; Kitaura, K. *The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems*; CRC Press: Boca Raton, FL, 2009.

(10) Deev, V.; Collins, M. A. Approximate ab initio energies by systematic molecular fragmentation. *J. Chem. Phys.* **2005**, *122*, 154102−154112.

(11) Yang, W. Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett.* **1991**, *66*, 1438.

(12) Zhang, D. W.; Xiang, Y.; Zhang, J. Z. H. New advance in computational chemistry: Full quantum mechanical ab initio computation of streptavidin−biotin interaction energy. *J. Phys. Chem. B* **2003**, *107*, 12039−12041.

(13) Tschumper, G. S. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Cundari, T. A., Eds.; Jon Wiley & Sons: NJ, 2009; p 39.

(14) Xie, W.; Gao, J. Design of a next generation force field: The X-POL potential. *J. Chem. Theory Comput.* **2007**, *3*, 1890−1900.

(15) Gao, J. A molecular-orbital derived polarization potential for liquid water. *J. Chem. Phys.* **1998**, *109*, 2346−2354.

(16) Gao, J. Toward a molecular orbital derived empirical potential for liquid simulations. *J. Phys. Chem. B* **1997**, *101*, 657–663.

(17) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation methods: A route to accurate calculations on large systems. *Chem. Rev.* **2012**, *112*, 632–672.

(18) Gordon, M. S.; Slipchenko, L. V.; Li, H.; Jensen, J. H. The effective fragment potential: A general method for predicting intermolecular forces. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177–193.

(19) Ghosh, D.; Kosenkov, D.; Vanovschi, V.; Williams, C. F.; Herbert, J. M.; Gordon, M. S.; Schmidt, M. W.; Slipchenko, L. V.; Krylov, A. I. Noncovalent interactions in extended systems described by the effective fragment potential method: Theory and application to nucleobase oligomers. *J. Phys. Chem. A* **2010**, *114*, 12739–12754.

(20) Slipchenko, L. V.; Gordon, M. S. Electrostatic energy in the effective fragment potential method: Theory and application to benzene dimer. *J. Comput. Chem.* **2007**, *28*, 276–291.

(21) Smith, T.; Slipchenko, L. V.; Gordon, M. S. Modeling π–π interactions with the effective fragment potential method: The benzene dimer and substituents. *J. Phys. Chem. A* **2008**, *112*, 5286–5294.

(22) Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. Benzene–pyridine interactions predicted by the effective fragment potential method. *J. Phys. Chem. A* **2011**, *115*, 4598–4609.

(23) Adamovic, I.; Li, H.; Lamm, M. H.; Gordon, M. S. Modeling styrene–styrene interactions. *J. Phys. Chem. A* **2006**, *110*, 519–525.

(24) Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. Effective fragment potential study of the interaction of DNA bases. *J. Phys. Chem. A* **2011**, *115*, 11269–11276.

(25) Slipchenko, L. V.; Gordon, M. S. Water–benzene interactions: An effective fragment potential and correlated quantum chemistry study. *J. Phys. Chem. A* **2009**, *113*, 2092–2102.

(26) Adamovic, I.; Gordon, M. S. Methanol–water mixtures: A microsolvation study using the effective fragment potential method. *J. Phys. Chem. A* **2006**, *110*, 10267–10273.

(27) Mullin, J. M.; Gordon, M. S. Alanine: Then there was water. *J. Phys. Chem. A* **2009**, *113*, 8657–8669.

(28) Mullin, J. M.; Gordon, M. S. Water and alanine: From puddles(32) to ponds(49). *J. Phys. Chem. A* **2009**, *113*, 14413–14420.

(29) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(30) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(31) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. Basis set consistent revision of the S22 test set of noncovalent interaction energies. *J. Chem. Phys.* **2010**, *132*, 144104.

(32) Podeszwa, R.; Patkowski, K.; Szalewicz, K. Improved interaction energy benchmarks for dimers of biological relevance. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5974–5979.

(33) Řezáč, 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.

(34) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.

(35) Hohenstein, E. G.; Sherrill, C. D. Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010**, *133*, 014101.

(36) Adamovic, I.; Freitag, M. A.; Gordon, M. S. Density functional theory based effective fragment potential method. *J. Chem. Phys.* **2003**, *118*, 6725–6732.

(37) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The effective fragment potential method: A QM-based MM approach to modeling environmental effects in chemistry. *J. Phys. Chem. A* **2001**, *105*, 293–307.

(38) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. An effective fragment method for modeling solvent effects in quantum mechanical calculations. *J. Chem. Phys.* **1996**, *105*, 1968–1986.

(39) Gordon, M. S.; Mullin, J. M.; Pruitt, S. R.; Roskop, L. B.; Slipchenko, L. V.; Boatz, J. A. Accurate methods for large molecular systems. *J. Phys. Chem. B* **2009**, *113*, 9646–9663.

(40) Stone, A. J. Distributed multipole analysis, or how to describe a molecular charge-distribution. *Chem. Phys. Lett.* **1981**, *83*, 233–239.

(41) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, 1996.

(42) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-consistent molecular-orbital methods. 12. Further extensions of Gaussian-type basis sets for use in molecular-orbital studies of organic-molecules. *J. Chem. Phys.* **1972**, *56*, 2257.

(43) Hariharan, P. C.; Pople, J. A. Influence of polarization functions on molecular-orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(44) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. Efficient diffuse function-augmented basis-sets for anion calculations. 3. The 3-21+G basis set for 1st-row elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294–301.

(45) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular-orbital methods. 20. Basis set for correlated wave-functions. *J. Chem. Phys.* **1980**, *72*, 650–654.

(46) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular-orbital methods. 25. Supplementary functions for Gaussian-basis sets. *J. Chem. Phys.* **1984**, *80*, 3265–3269.

(47) Slipchenko, L. V.; Gordon, M. S. Damping functions in the effective fragment potential method. *Mol. Phys.* **2009**, *107*, 999–1016.

(48) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.

(49) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F.; 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. PSI4: An open-source ab initio electronic structure program. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556–565.

(50) Gráfová, L.; Pitoňák, M.; Řezáč, J.; Hobza, P. Comparative study of selected wave function and density functional methods for noncovalent interaction energy calculations using the extended S22 data set. *J. Chem. Theory Comput.* **2010**, *6*, 2365–2376.

(51) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(52) Dunning, T. H. Gaussian-basis sets for use in correlated molecular calculations. 1. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(53) Woon, D. E.; Dunning, T. H. Gaussian-basis sets for use in correlated molecular calculations. 5. Core-valence basis-sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(54) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic-structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(55) Gordon, M. S.; Schmidt, M. W. Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: New York, 2005; Chapter 41.

(56) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y. H.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W. M.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Voorhis, T.; Oumi, M.; Hirata, S.; Hsu, C. P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. Q-chem 2.0: A high-performance ab initio electronic structure program package. *J. Comput. Chem.* **2000**, *21*, 1532–1548.

(57) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.;

O'Neill, D. P.; DiStasio, R. A.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; 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.; Doerksen, 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.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.

(58) DiStasio, R. A., Jr.; von Helden, G.; Steele, R. P.; Head-Gordon, M. On the T-shaped structures of the benzene dimer. *Chem. Phys. Lett.* **2007**, *437*, 277−283.

(59) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(60) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.

(61) Burns, L. A.; Vazquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *J. Chem. Phys.* **2011**, *134*, 084107.

(62) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking density functional methods against the S66 and S66 × 8 datasets for noncovalent interactions. *ChemPhysChem* **2011**, *12*, 3421−3433.

(63) Yang, Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the self-consistent-charge density-functional tight-binding method: Third-order expansion of the density functional theory total energy and introduction of a modified effective Coulomb interaction. *J. Phys. Chem. A* **2007**, *111*, 10861−10873.

(64) Korth, M. Third-generation hydrogen-bonding corrections for Ssemiempirical QM methods and force fields. *J. Chem. Theory Comput.* **2010**, *6*, 3808−3816.

(65) Řezáč, J.; Hobza, P. Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods. *J. Chem. Theory Comput.* **2011**, *8*, 141−151.

(66) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06 functionals and twelve other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(67) Sherrill, C. D.; Hohenstein, E. G.; Chill, S. T. Assessment of the performance of the M05-2X and M06-2X exchange-correlation functionals for noncovalent interactions in biomolecules. *J. Chem. Theory Comput.* **2008**, *4*, 1996−2000.

(68) Takatani, T.; Hohenstein, E. G.; Sherrill, C. D. Improvement of the coupled-cluster singles and doubles method via scaling same- and opposite-spin components of the double excitation correlation energy. *J. Chem. Phys.* **2008**, *128*, 124111.

(69) Grimme, S. Improved second-order Moller−Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095−9102.

(70) Paton, R. S.; Goodman, J. M. Hydrogen bonding and π-stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions. *J. Chem. Inf. Model.* **2009**, *49*, 944−955.

(71) Hands, M. D.; Slipchenko, L. V. Intermolecular interactions in complex liquids: Effective fragment potential investigation of water−tert-butanol mixtures. *J. Phys. Chem. B* **2012**, *116*, 2775−2786.

(72) Stone, A. J.; Misquitta, A. J. Charge-transfer in symmetry-adapted perturbation theory. *Chem. Phys. Lett.* **2009**, *473*, 201−205.

(73) Hohenstein, E. G.; Duan, J.; Sherrill, C. D. Origin of the surprising enhancement of electrostatic energies by electron-donating substituents in substituted sandwich benzene dimers. *J. Am. Chem. Soc.* **2011**, *133*, 13244−13247.