Assessment of Density Functional Theory in Predicting Interaction Energies Between Water and Polycyclic Aromatic Hydrocarbons: From Water on Benzene to Water on Graphene

Adeayo O. Ajala,*† Vamsee Voora,‡ Narbe Mardirossian,¶ Filipp Furche,† and Francesco Paesani*†§∥

†Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States
‡Department of Chemistry, University of California Irvine, Irvine, California 92697, United States
¶Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 E. California Boulevard, Pasadena, California 91125, United States
§Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, United States
∥San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, United States

E-mail: aajala@ucsd.edu fpaesani@ucsd.edu
Abstract

The interaction of water with polycyclic aromatic hydrocarbons, from benzene to graphene, is investigated using various exchange-correlation functionals selected across generalized gradient approximation (GGA), meta-GGA, and hybrid families within the density functional theory (DFT) hierarchy. The accuracy of the different functionals is assessed through comparisons with high-level electronic structure methods, including random phase approximation (RPA), diffusion Monte Carlo (DMC), and coupled-cluster with single, double, and perturbative triple excitations (CCSD(T)). Relatively large variations are found in the interaction energies predicted by different DFT models, with GGA functionals underestimating the interaction strength for configurations with the water oxygen pointing toward the aromatic molecules, and the meta-GGA B97M-rV and hybrid ωB97M-V functionals providing nearly quantitative agreement with CCSD(T) values available for the water–benzene, water-coronene, and water–circumcoronene dimers, which, in turn, are within ~1 kcal/mol of the corresponding RPA and DMC results. Similar trends among GGA, meta-GGA, and hybrid functionals are observed for the larger polycyclic aromatic hydrocarbon molecules considered in this analysis (up to C_{216}H_{36}). By performing absolutely localized molecular orbital energy decomposition analyses (ALMO-EDA) of the DFT results, it is found that, independently of the number of carbon atoms and exchange-correlation functional, the dominant contributions to the interaction energies between water and polycyclic aromatic hydrocarbon molecules are the electrostatic and dispersion terms while polarization and charge transfer effects are negligibly small. Calculations carried out with GGA and meta-GGA functionals indicate that, as the number of carbon atoms increases, the interaction energies slowly converge to the corresponding values obtained for an infinite graphene sheet. Importantly, water-graphene interaction energies calculated with the B97M-rV functional appear to deviate by more than 1 kcal/mol from the available RPA and DMC values.
1 INTRODUCTION

Motivated by potential technological applications, including water desalination, electricity generation, and biochemical sensing, there has recently been significant effort in investigating the properties of water interacting with graphene. From a computational modeling perspective, realistic simulations of water at the interface with graphene sheets require an accurate representation of the underlying molecular interactions, at both short and long ranges. Several molecular dynamics (MD) studies, employing either force fields or \textit{ab initio} methods, have been reported to characterize the behavior of water adsorbed on graphene. For example, MD simulations were carried out to investigate the dependence of the contact angle on the strength of carbon–water interactions as well as to determine the importance of polarization effects in the representation of water–graphene interactions. Other studies focused on characterizing the hydrophobic effect on electrically doped graphene layers. MD simulations with the TIP4P water model were used in Ref. to examine the desalination performance of graphene, while water and ion transport through graphene pores was investigated in Refs. and .

On the other hand, most \textit{ab initio} studies of water interacting with graphene rely on density functional theory (DFT). However, it is known that standard exchange-correlation functionals lack the ability to capture long-range dispersion interactions and thus often underestimate the strength of molecular interactions. More reliable results can be obtained by combining standard exchange-correlation functionals with empirically parameterized dispersion corrections, or by constructing nonlocal exchange-correlation functionals. In this context, it was shown that the performance of different DFT models in describing water–graphene interactions depends not only on the specific exchange-correlation functional used in the calculations but also on how dispersion forces are accounted for. In particular, it was found that the application of van der Waals DFT models to investigate the interaction of a single water molecule with a graphene sheet gives results that are in good agreement with those obtained with high-level electronic structure methods.
Diffusion Monte Carlo (DMC)\(^{35-37}\) and coupled-cluster (CC) with single, double, and perturbative triple excitations (CCSD(T))\(^{38}\) have become the reference methods for chemical accuracy. Unfortunately, both DMC and CCSD(T) calculations are computationally expensive, which precludes their routine application to condensed phase systems. The question that arises is if one can turn to a more affordable method that works equally well for non-bonded monomers, while still retaining the ability to be applicable to larger systems. One can address this challenge by contemplating another versatile electronic structure method such as the random phase approximation (RPA)\(^{39-44}\) that has been shown to provide an accurate description of intermolecular interactions at a fraction of the computational cost associated with DMC and CCSD(T) calculations. The RPA method combines a number of attractive features, such as the inclusion of long-range many-body dispersion-type correlation effects and a low computational cost of \(\mathcal{O}(N^4 \log N)\), which makes it appealing for modeling molecular systems with tens of atoms.\(^{45-52}\)

In this study, we report the application of various GGA, meta-GGA, and hybrid exchange-correlation functionals to determine interaction energies between a single water molecule and a series of polycyclic aromatic hydrocarbons (PAHs), from benzene to graphene. The accuracy of the different functionals is established through comparisons with RPA results as well as with DMC and CCSD(T) data available in the literature. Besides enabling a systematic assessment of the performance of different DFT approximations, this study also provides fundamental insights into the nature of the interactions between water and PAHs through the application of the absolutely localized molecular orbital energy decomposition analysis (ALMO-EDA) method.\(^{53,54}\) The analysis reported in this study could thus serve as a guidance for the development of \textit{ab initio}-based force fields MD simulations aimed at determining the structure and dynamics of water at graphene interfaces.
2 COMPUTATIONAL DETAILS

2.1 Molecular systems

In this study, we consider a graphene sheet with C-C bonds and C-C-C angles set to the experimental values of 1.42 Å and 120°, respectively.\(^{55}\) In order to investigate the dependence of the interaction energies as a function of the number of carbon atoms, a series of PAH molecules are extracted from the graphene sheet and the dangling bonds saturated with hydrogen atoms, fixing the C-H bonds and C-C-H angles to the values of 1.089 Å and 120° as defined in Ref. \(^{56}\). The resulting PAHs include: C\(_6\)H\(_6\) (benzene), C\(_{24}\)H\(_{12}\) (coronene), C\(_{54}\)H\(_{18}\) (circumcoronene), C\(_{96}\)H\(_{24}\) (dicircumcoronene), C\(_{150}\)H\(_{30}\), and C\(_{216}\)H\(_{36}\). As shown in Fig. 1, three orientations of the water molecule with respect to the plane of each PAH molecule are considered, which are defined as: 0-leg configurations, with the water oxygen pointing towards the central aromatic ring of the PAH molecule and the bisector of the H-O-H angle perpendicular to the plane of the PAH molecule, 1-leg configurations, with one OH bond of the water molecule perpendicular to the plane of the PAH molecule and pointing to one of the carbon atoms of the central aromatic ring, and 2-leg configurations, corresponding to inverted 0-leg configurations, with the two hydrogen atoms pointing towards the central aromatic ring of the PAH molecule. In all calculations, the O-H bonds and H-O-H angle of the water molecule were kept fixed at the gas-phase experimental values of 0.957 Å and 104.5°, respectively.

2.2 Density Functional Theory Calculations

We analyze the performance of various exchange-correlation functionals belonging to the generalized gradient approximation (GGA), meta-GGA, and hybrid families. Among the GGA functionals considered in this study are BLYP, which combines Becke exchange functional (B88)\(^{57}\) with the semi-local Lee-Yang-Parr correlation functional (LYP),\(^{58}\) the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional\(^{59}\) as well as its revised version (revPBE),\(^{60}\)
and the revised Vydrov and van Voorhis functional (rVV10). As representative meta-GGA functionals, we consider the strongly constrained and appropriately normed SCAN functional and the semi-empirical meta-GGA B97M-V functional, which was recently paired with the rVV10 nonlocal correlation functional. Within the family of hybrid exchange-correlation functionals, we consider B3LYP, PBE0, and revPBE0, which are the hybrid analogues of PBE and revPBE, respectively, and ωB97M-V, which is the range-separated hybrid analogue of B97M-V. To approximate long-range electron correlations associated with dispersion interactions, we adopt the D3(0) semi-empirical approach proposed by Grimme and co-workers. In the remainder of this article, we thus designate the corresponding exchange-correlation functionals with dispersion corrections as BLYP-D3, PBE-D3, revPBE-D3, SCAN-D3, B3LYP-D3, PBE0-D3, and revPBE0-D3. In the case of SCAN, we also consider the SCAN+rVV10 variant, which is derived by pairing SCAN with the nonlocal correlation part of rVV10.

In the analyses presented in Section the interaction energies, $E_{int}$, between water and
the PAH molecules are calculated as

$$E_{\text{int}} = E_{\text{H}_2\text{O} - \text{PAH}} - E^{(\text{H}_2\text{O} - \text{PAH})}_{\text{H}_2\text{O}} - E^{(\text{H}_2\text{O} - \text{PAH})}_{\text{PAH}}$$

(1)

where $E_{\text{H}_2\text{O} - \text{PAH}}$ is the total energy of the $\text{H}_2\text{O} - \text{PAH}$ dimer, and $E^{(\text{H}_2\text{O} - \text{PAH})}_{\text{H}_2\text{O}}$ and $E^{(\text{H}_2\text{O} - \text{PAH})}_{\text{PAH}}$ are the energies of the isolated fragments at the same geometries as in the $\text{H}_2\text{O} - \text{PAH}$ dimer. All energies are corrected for the basis set superposition error (BSSE) according to the counterpoise scheme of Boys and Bernardi,\textsuperscript{71} with the superscripts $(\text{H}_2\text{O} - \text{PAH})$ in Eq. 1 thus indicating that the monomer energies are computed in the dimer basis set. All DFT calculations for the PAH molecules are carried out within the Kohn-Sham formalism as implemented in Q-Chem, version 5.0.\textsuperscript{72} We employ the def2-QZVPPD\textsuperscript{73} basis set for O and H in water and the surrounding C atoms in the central six-membered ring of the PAH, and the 6-31+G\textsuperscript{74,75} basis set for the remaining atoms in the system. This basis set combination guarantees computational efficiency without compromising accuracy in predicting interaction energies.

The DFT calculations of the water–graphene interaction energies are carried out using the QUICKSTEP algorithm\textsuperscript{76} as implemented in the CP2K software.\textsuperscript{77} For these calculations, we employ augmented quadruple-zeta valence basis sets with three polarization functions (aug-QZV3P) for both water and graphene, while the core electrons are described by the Goedecker-Teter-Hutter pseudopotentials.\textsuperscript{78} The calculations are carried out for a $5 \times 5$ supercell with dimensions of 12.325 Å$\times$12.325 Å in the $x - y$ plane, and a vacuum region of 25.0 Å in the $z$-direction.

2.3 Random Phase Approximation

RPA calculations are carried out for water–C$_6$H$_6$, water–C$_{24}$H$_{12}$, and water–C$_{54}$H$_{18}$ using TURBOMOLE v7.3.\textsuperscript{79} The energy calculations use self-consistent PBE\textsuperscript{59} orbitals which are computed using m5 grids\textsuperscript{80} and an energy-convergence criterion of $10^{-7}$ a.u. All interaction
energies are corrected for the BSSE as discussed above. While standard Dunning basis sets can be straightforwardly used for C\textsubscript{6}H\textsubscript{6} and C\textsubscript{24}H\textsubscript{12}, their use for larger PAHs leads to linear-
dependency issues. We therefore use the def2-QZVP\textsuperscript{81} basis sets, which were shown to offer good balance between accuracy and efficiency for interaction energies.\textsuperscript{82} Since for benzene and coronene, the use of def2-QZVP leads to a basis set incompleteness error of 0.2 kcal/mol only, def2-QZVP is then used in all RPA calculations.

### 2.4 Coupled Cluster Theory

All CC calculations presented in this study are based on restricted Hartree–Fock reference functions. Interaction energies are calculated for selected H\textsubscript{2}O–C\textsubscript{6}H\textsubscript{6}, H\textsubscript{2}O–C\textsubscript{24}H\textsubscript{12}, and H\textsubscript{2}O–C\textsubscript{54}H\textsubscript{18} configurations using the linear scaling domain-based pair natural orbital CCSD(T) method, hereafter referred to as L-CCSD(T),\textsuperscript{83} available in the ORCA package.\textsuperscript{84} Dunning basis sets of cc-pVXZ (where X = D,T,Q,5) quality\textsuperscript{85,86} are used in the correlation treatment, while the chemical core is frozen. As a result of linear-dependency problems mentioned in Section 2.3, the cc-pVXZ (X = D,T,Q,5) basis sets are used for H\textsubscript{2}O–C\textsubscript{6}H\textsubscript{6}, the cc-pVXZ (X = D,T,Q) basis sets for H\textsubscript{2}O–C\textsubscript{24}H\textsubscript{12}, and the cc-pVXZ (X = D,T) basis sets for H\textsubscript{2}O–C\textsubscript{54}H\textsubscript{18}. All interaction energies are corrected for the BSSE using Eq. 1. Following Refs. \textsuperscript{87} and \textsuperscript{88} the complete basis set (CBS) limit of the L-CCSD(T) interaction energies is achieved by applying the following two-point extrapolation formula:

\[
E_0^{(CBS)} = \frac{n^3 E_0^{(ccnZ)} - (n - 1)^3 E_0^{(cc(n-1)Z)}}{n^3 - (n - 1)^3},
\]

where \(n\) and \((n - 1)\) are the cardinal numbers of the cc\textsubscript{n}Z and cc\textsubscript{\((n-1)\)Z} basis sets, respectively. For H\textsubscript{2}O–C\textsubscript{6}H\textsubscript{6}, we extrapolate the results obtained with the largest basis sets by setting \(n = 5\) in Eq. 2. Due to the relatively smaller size of the basis sets used for H\textsubscript{2}O–C\textsubscript{24}H\textsubscript{12}, errors that may originate from basis set incompleteness are also taken into account in the (TZ,QZ) extrapolation. These errors are estimated from the difference between
(TZ,QZ) and (QZ,5Z) extrapolations carried out for H$_2$O–C$_6$H$_6$. Following Refs. 89 and 90, the ground-state electronic energies of H$_2$O–C$_{54}$H$_{18}$ dimer are calculated as

$$E_0 = E_0^{(\text{RHF/QZ})} + \Delta E_0^{(\text{CBS/(TZ,DZ)})},$$

where the first term on the right-hand side designates the RHF reference energy obtained using the cc-pVQZ basis set. Due to the fast convergence of the Hartree–Fock energies with respect to the basis set, we can treat the RHF/QZ energies as equivalent to the CBS value. The second term on the right-hand side of Eq. 3 is the CBS limit of the correlation energy and the triples correction due to the (TZ,DZ) extrapolation. We estimate errors due to basis set incompleteness by forming the difference between 5Z and QZ calculations for H$_2$O–C$_6$H$_6$. To assess the accuracy of the L-CCSD(T) results, the differences between interaction energies calculated for H$_2$O–C$_6$H$_6$ using both L-CCSD(T) and standard CCSD(T) in combination with the cc-pVTZ are analyzed. The differences are found to be within 0.2 kcal/mol for 0-leg configuration and 0.1 kcal/mol for both 1- and 2-leg configurations.

3 RESULTS

3.1 Interaction energies for water on C$_6$H$_6$, C$_{24}$H$_{12}$, and C$_{54}$H$_{18}$

Figure 2 shows comparisons between potential energy curves calculated using GGA (top row), meta-GGA (middle row), and hybrid (bottom row) functionals for 0-leg (panels a-c), 1-leg (panels d-f), and 2-leg (panels g-i) configurations of the H$_2$O–C$_6$H$_6$ dimer. Also shown are the corresponding RPA and L-CCSD(T) values, along with DMC data from Ref. 91.

Independently of the level of theoretical treatment, water in the 0-leg configuration is predicted not to be bound to C$_6$H$_6$, in agreement with previous studies. Most GGA and meta-GGA functionals, with the exception of BLYP-D3 and revPBE-D3, predict interaction energies that are in good agreement (with deviations within $\sim$0.1 kcal/mol) with the co-
Figure 2: $\text{H}_2\text{O}–\text{C}_6\text{H}_6$ interaction energies (in kcal/mol) calculated for 0-leg, 1-leg, and 2-leg configurations as a function of the corresponding intermolecular distances. Top panels: GGA functionals (red), middle panels: meta-GGA functionals (green), bottom panels: hybrid functionals (cyan). Also shown for comparison are the RPA, DMC, and L-CCSD(T) values. The intermolecular distances are defined as the vertical distances between the O atom of the water molecule and the plane of the $\text{C}_6\text{H}_6$ molecule. DMC values are taken from Ref. 91.
responding L-CCSD(T) values for 0-leg intermolecular distances between 2.5 Å and 4.0 Å. In contrast, both BLYP-D3 and revPBE-D3 predict interaction energies similar to those obtained at the RPA and DMC levels of theory, which, in turn, deviate by \(\sim 0.5 \text{ kcal/mol}\) from the corresponding L-CCSD(T) values for intermolecular distances between 2.5 Å and 4.0 Å. All hybrid functionals considered in this study predict interaction energies for 0-leg configurations of \(\text{H}_2\text{O–C}_6\text{H}_6\) similar to those obtained with RPA and DMC.

Both 1-leg and 2-leg configurations of water bind to \(\text{C}_6\text{H}_6\), with RPA, DMC, and L-CCSD(T) predicting minimum interaction energies between 2.50 and 3.0 kcal/mol (1-leg) and between 2.80 and 3.20 kcal/mol (2-leg). Among the three methods, RPA and L-CCSD(T) consistently predict the weakest and strongest interactions, with differences on the order of \(\sim 0.6 \text{ kcal/mol}\) in the minimum energy region between 2.5 Å and 4.0 Å. While the DMC curves lie in between the RPA and L-CCSD(T) results, they are generally closer to the L-CCSD(T) values, with deviations always within 0.2 kcal/mol. Among the GGA functionals considered in this study, BLYP-D3, revPBE-D3, and rVV10 systematically overestimate the interaction strength between 2.5 Å and 4.0 Å by 0.2–0.3 kcal/mol. Somewhat larger deviations (up to -0.70 kcal/mol) from the L-CCSD(T) values are associated with meta-GGA and hybrid functionals in the 1-leg and 2-leg energy minimum regions, with the exception of B97M-rV and \(\omega\)B97M-V that closely reproduce the L-CCSD(T) curves at all \(\text{H}_2\text{O–C}_6\text{H}_6\) separations.

As shown in Figure 3, by increasing the size of the PAH molecule from \(\text{C}_6\text{H}_6\) to \(\text{C}_{24}\text{H}_{12}\), more defined trends in the interaction energies begin to emerge. First, the 0-leg configuration of water is bound to coronene (\(\text{C}_{24}\text{H}_{12}\)) contrary to what was observed in the case of benzene (\(\text{C}_6\text{H}_6\)). Second, interaction energies calculated at different levels of theory become more distinct. For instance, the differences between L-CCSD(T) and DMC and RPA results in the minimum energy region of the 0-leg configuration are 0.2 kcal/mol and 0.8 kcal/mol, respectively. The variations in interaction energies are not limited to the 0-leg configuration and are even more pronounced for 1-leg and 2-leg configurations, with the differences in the energy minimum regions between L-CCSD(T) and DMC increasing to 0.5 kcal/mol for both
Figure 3: H$_2$O–C$_{24}$H$_{12}$ interaction energies (in kcal/mol) calculated for 0-leg, 1-leg, and 2-leg configurations as a function of the corresponding intermolecular distances. Top panels: GGA functionals (red), middle panels: meta-GGA functionals (green), bottom panels: hybrid functionals (cyan). Also shown for comparison are the RPA, DMC, and L-CCSD(T) values. The intermolecular distances are defined as the vertical distances between the O atom of the water molecule and the plane of the C$_{24}$H$_{12}$ molecule. DMC values are taken from Ref. 91.
1- and 2-leg configurations, and those between L-CCSD(T) and RPA becoming 1.0 kcal/mol (1-leg) and 0.9 kcal/mol (2-leg). In all cases, the DMC interaction energy curves remain in between the corresponding RPA and L-CCSD(T) curves, in line with the results shown in Figure 2 for water interacting with benzene.

Having established the performance of L-CCSD(T), DMC, and RPA in describing the H$_2$O–C$_{24}$H$_{12}$ dimer, we now turn our attention to assessing the relative accuracy of the different exchange-correlation functionals. Within the GGA family, rVV10 predicts interaction energies that are somewhat in good agreement with the corresponding L-CCSD(T) values for the 0-leg configuration but slightly underbinds and overbinds by $\sim$0.1 kcal/mol for both 1-leg and 2-leg configurations, respectively. Among the meta-GGA functionals, B97M-rV and SCAN-D3 tend to correctly predict water to coronene interaction relative to the L-CCSD(T) data in the 0-leg configuration, whereas SCAN+rVV10 underestimate the interaction energy by 0.3 kcal/mol. However, in the 1- and 2-leg structures, SCAN-D3 overestimates the interaction energy by $\sim$0.6 kcal/mol in the minimum region where both B97M-rV and SCAN+rVV10 show good agreement with L-CCSD(T) predictions. The hybrid functionals, on the other hand, predict interaction energies that tend to underbind water in 0-leg but overbind in both 1-leg and 2-leg configurations. The only exception is $\omega$B97M-V which, independently of the water orientation, consistently display good agreement with L-CCSD(T).

Due to the associated computational cost, L-CCSD(T) calculations for the H$_2$O–C$_{54}$H$_{18}$ dimer are only carried out at the cc-pVDZ and cc-pVTZ basis sets for configurations in the minimum energy region. Since no DMC results are available for this large PAH, the performance of the different exchange-correlation functionals considered in this study is then assessed through comparisons with the available RPA/def2-QZVP and L-CCSD(T)/(DZ,TZ) data. As shown, in Fig. 4, the differences between RPA and L-CCSD(T) interaction energies are more pronounced than for H$_2$O–C$_6$H$_6$ and H$_2$O–C$_{24}$H$_{12}$, being in the range of 0.8–1.0 kcal/mol for 0-leg configurations, 0.9–1.1 kcal/mol for 1-leg configurations, and 1.0–1.3 kcal/mol for 2-leg configurations.
Figure 4: H$_2$O–C$_{54}$H$_{18}$ interaction energies (in kcal/mol) calculated for 0-leg, 1-leg, and 2-leg configurations as a function of the corresponding intermolecular distances. Top panels: GGA functionals (red), middle panels: meta-GGA functionals (green), bottom panels: hybrid functionals (cyan). Also shown for comparison are the RPA and L-CCSD(T) values. The intermolecular distances are defined as the vertical distances between the O atom of the water molecule and the plane of the C$_{54}$H$_{18}$ molecule.
In general, interaction energies calculated with PBE-D3, rVV10, SCAN+rVV10, B97M-rV, and ωB97M-V display better agreement with the available L-CCSD(T) results. In order to facilitate the comparison between various functionals and identify general trends, we limit the discussion to rVV10, B97M-rV, and ωB97M-V as representatives for the GGA, meta-GGA, and hybrid and hybrid/meta-GGA functionals, respectively. As shown in Fig. 4, independently of the water orientation, the differences between rVV10 and L-CCSD(T) interaction energies in the minimum energy regions do not exceed 0.2 kcal/mol. Both B97M-rV and ωB97M-V predict interaction energies that lie within the range of L-CCSD(T) values.

Fig. 4 indicates that going from H$_2$O–C$_{24}$H$_{12}$ to H$_2$O–C$_{54}$H$_{18}$, the differences between rVV10 and RPA interaction energies increases slightly to 0.9 kcal/mol (for 0-leg configurations) and 1.2 kcal/mol (1-leg configurations), while they vary significantly (by ∼1.2 kcal/mol) for 2-leg configurations. A similar analysis carried out for the meta-GGA B97M-rV functional indicates that the deviations from the corresponding RPA values for 0-leg minimum energy configurations are 0.5 kcal/mol, 0.9 kcal/mol, and 1.0 kcal/mol for the H$_2$O–C$_6$H$_6$, H$_2$O–C$_{24}$H$_{12}$, and H$_2$O–C$_{54}$H$_{18}$ dimers, respectively. Similar trends are also observed for both 1-leg and 2-leg configurations, with the differences between B97M-rV and RPA results being 0.55 kcal/mol, 0.85 kcal/mol, and 1.0 kcal/mol for water interacting with benzene, coronene, and cicumcoronene, respectively. While the ωB97M-V functional predicts interaction energies that deviate by the same amount as B97M-rV from the RPA values for 1-leg and 2-leg configurations in the minimum energy regions, it predicts slightly smaller differences (by ∼0.2 kcal/mol) than B97M-rV for 0-leg configurations.

Overall, after taking into account the missing 0.2 kcal/mol due to basis-set incompleteness in the RPA calculations as discussed in Section 2.3 and the differences between RPA, DMC, and L-CCSD(T) results available for H$_2$O–C$_6$H$_6$ and H$_2$O–C$_{24}$H$_{12}$, it is possible to estimate that all exchange-correlation functionals examined in this study should provide interaction energies for the H$_2$O–C$_{54}$H$_{18}$ dimer in the minimum energy region which are within ∼1 kcal/mol of the corresponding values obtained from higher-level electronic structure meth-
ods. In general terms, we can conclude that most functionals examined in this study tend to overestimate the interaction strength in 1-leg and 2-leg configurations, while underestimating the interaction strength in 0-leg configurations, of the H$_2$O–C$_6$H$_6$, H$_2$O–C$_{24}$H$_{12}$, and H$_2$O–C$_{54}$H$_{18}$ dimers. Within this general trend, GGA functionals overestimate the differences between 0-leg and 2-leg configurations, while B97M-rV and ωB97M-V provide closer agreement with higher-level electronic structure methods. In particular, independently of the water orientation, ωB97M-V is found to reproduce, nearly quantitatively, the available L-CCSD(T) data.

In order to provide additional insights into water–PAHs interactions and assess the performance of different exchange-correlation functionals in representing the underlying physics, we also carried out the absolutely localized molecular orbital energy decomposition analysis (ALMO-EDA) to characterize individual physical contributions. The ALMO-EDA separates the intermolecular interaction energy into Pauli repulsion (PAULI), permanent electrostatics (ELEC), polarization (POL), dispersion energy (DISP), and charge transfer (CT) contributions.

Figure 5 shows that there is no qualitative difference in the way different functionals considered in this study represents the individual contributions to the interaction energies. In addition, no qualitative differences exist as a function of water orientation and number of carbon atoms in the PAH molecules. We can then conclude that the purely repulsive nature of the H$_2$O–C$_6$H$_6$ interaction in the 0-leg configuration is a consequence of the dominant role played by the Pauli repulsion term, outweighing the combined attractive contributions from all other energy components. It should be noted that the Pauli repulsion term associated with the 0-leg configuration is not as repulsive as in the 1-leg and 2-leg configurations (top, middle, and bottom panels of Figure 5) for which the closer approach of the water hydrogen atom(s) initiates overlap of atomic orbitals with the benzene $\pi$ system. In spite of the increase in Pauli repulsion for 1-leg and 2-leg configurations, the corresponding interaction energies are negative due to the counterbalance effects associated with the ELEC, POL,
DISP, and CT terms, with ELEC representing the dominant contribution.

Figure 5: ALMO-EDA results for H$_2$O–C$_6$H$_6$ (top), H$_2$O–C$_{24}$H$_{12}$ (middle), and H$_2$O–C$_{54}$H$_{18}$ (bottom) dimers. On the vertical axes are the energies, in kcal/mol, of each component of the interaction energy. Each group of bars from left to right corresponds to 0-leg, 1-leg, and 2-leg configurations. Color scheme: PAULI in red, CT in yellow, POL in green, DISP in blue, ELECT in gray. See main text for details.
Going from $\text{H}_2\text{O}\text{–C}_6\text{H}_6$ to $\text{H}_2\text{O}\text{–C}_{24}\text{H}_{12}$, the PAULI contribution to the 0-leg interaction energy becomes significantly less repulsive, while the POL, DISP, and ELEC terms become relatively more attractive, and the CT term remains effectively unchanged. Compared to the ALMO-EDA results for 1-leg and 2-leg configurations of the $\text{H}_2\text{O}\text{–C}_6\text{H}_6$ dimer, the DISP term becomes more attractive, while both CT and ELEC contributions to the interaction energies become less attractive, with ELEC still representing the dominant contribution. Finally, no appreciable differences in the relative contributions to the interaction energies are found between the $\text{H}_2\text{O}\text{–C}_{24}\text{H}_{12}$ and $\text{H}_2\text{O}\text{–C}_{54}\text{H}_{18}$ dimers.

### 3.2 $\text{H}_2\text{O}$ interactions with larger PAH molecules, from $\text{C}_{96}\text{H}_{24}$ to $\text{C}_{216}\text{H}_{36}$

Owing to the good performance of B97M-rV and $\omega$B97M-V in describing the $\text{H}_2\text{O}\text{–C}_6\text{H}_6$, $\text{H}_2\text{O}\text{–C}_{24}\text{H}_{12}$, and $\text{H}_2\text{O}\text{–C}_{54}\text{H}_{18}$ dimers discussed in Section 3.1, the B97M-rV and $\omega$B97M-V functionals are then used to investigate the dependence of the water–PAH interaction energies on the number of carbon atoms in larger PAH molecules for which RPA, DMC, and L-CCSD(T) calculations are prohibitively expensive. Specifically, continuing on the series of symmetric PAH molecules, interaction energy curves were calculated for 0-leg, 1-leg, and 2-leg configurations of the $\text{H}_2\text{O}\text{–C}_{96}\text{H}_{24}$, $\text{H}_2\text{O}\text{–C}_{150}\text{H}_{30}$, and $\text{H}_2\text{O}\text{–C}_{216}\text{H}_{36}$ dimers.

The evolution of the B97M-rV and $\omega$B97M-V interaction energies from the $\text{H}_2\text{O}\text{–C}_6\text{H}_6$ to the $\text{H}_2\text{O}\text{–C}_{216}\text{H}_{36}$ dimer are shown in the top and bottom panels of Figure 6, respectively. Both functionals predict that the strength of water–PAH interactions for 0-leg configuration increases monotonically with the number of carbon atoms. In the case of 1-leg and 2-leg configurations, the trend is reversed, with the $\text{H}_2\text{O}\text{–C}_6\text{H}_6$ dimer displaying the largest interaction energy among the PAH series. Importantly, while water in the 2-leg configuration is predicted to always interact more strongly with the PAH molecules, independently of the number of carbon atoms, both functionals predict a crossover in the interaction energies associated with 0-leg and 1-leg configurations which occurs between $\text{H}_2\text{O}\text{–C}_{150}\text{H}_{30}$ and $\text{H}_2\text{O}$–
C$_{216}$H$_{36}$, with the 0-leg configuration becoming more attractive than the 1-leg configuration by $\sim 0.02$ kcal/mol.

Building on the analysis presented in Section 3.1, ALMO-EDA calculations carried out for 0-leg, 1-leg, and 2-leg configurations in the minimum energy regions (Figure 7) indicate that

![Figure 6: B97M-rV (top panels) and $\omega$B97M-V (bottom panels) interaction energy curves calculated for 0-leg (a and b), 1-leg (c and d), and 2-leg (e and f) configurations of all H$_2$O–PAH dimers examined in this study. See main text for details.](image-url)
there is no qualitative difference in the relative magnitude of the different terms predicted by
the B97M-rV and ωB97M-V functionals. According to the ALMO-EDA results, polarization
and charge transfer terms do not contribute substantially to the total interaction energy and
remain approximately constant, independently of the size of the PAH molecule. On the
other hand, by counteracting the positive contribution due Pauli repulsion, the dispersion
energy and electrostatic terms are effectively responsible for the attractive nature of the
water–PAH interactions. Importantly, both these terms also exhibit the slowest convergence
with the number of carbon atoms. In particular, while the dispersion energy term becomes
more negative from C₆H₆ to C₂₁₆H₃₆ for all three water orientation, the variation of the
electrostatic contributions depends on the water orientation, increasing in magnitude for
the 0-leg configuration while decreasing for both 1-leg and 2-leg configurations as the size
of the PAH molecules increases. The different dependence of the electrostatic term on the
water orientation thus appears to be responsible for the crossover between 0-leg and 1-leg
interaction energies occurring in Fig. 6 between H₂O–C₁₅₀H₃₀ and H₂O–C₂₁₆H₃₆.

3.3 H₂O interactions with graphene

Having characterized the magnitude and nature of the interactions between water and
PAH molecules from C₆H₆ to C₂₁₆H₃₆, and assessed the performance of various exchange-
correlation functionals, in this section we examine the interaction of water with graphene.
Due to the computational cost associated with hybrid functionals, the analysis will be limited
to GGA and meta-GGA functionals. Based on the good performance exhibited by B97M-rV
in describing the interactions between water and PAH molecules, we first compare in Fig-
ure 8 the 0-leg, 1-leg, and 2-leg interaction energy curves calculated with B97M-rV for the
H₂O–graphene system with the available RPA and DMC data taken from Ref. 91.

Focusing on the minimum energy regions for the three different configurations, it is pos-
sible to see that the RPA and DMC results are in close agreement for the 0-leg configuration
while differences of ~0.5 kcal/mol are found for both 1-leg and 2-leg configurations. These
differences are similar to those found for the H$_2$O–C$_6$H$_6$ and H$_2$O–C$_{24}$H$_{12}$ dimers in Figs. 2 and 3. On the other hand, the differences between the B97M-rV and RPA values are significantly larger (up to $\sim$1.8 kcal/mol) for all water orientations. This is in line with the trend

Figure 7: Decomposition of the total intermolecular interaction energy of water–PAH system as predicted by the B97M-rV (top) and $\omega$B97M-V (bottom) functionals. On the vertical axis are the energies, in kcal/mol, of the energy components and the horizontal axis represents the system size as defined by the 0-leg, 1-leg, and 2-leg configurations. Color scheme: PAULI in red, CT in yellow, POL in green, DISP in blue, ELECT in gray. See main text for details.
observed in the variation of the B97M-rV and RPA interaction energies going from H$_2$O–C$_6$H$_6$ to H$_2$O–C$_{54}$H$_{18}$. Relatively smaller differences which, however, are still on the order of 1 kcal/mol are found between B97M-rV and DMC results. Also in this case the differences between the two sets of results appear to be in line with the trends found for water interacting with smaller PAH molecules. In this context, it should be noted that B97M-rV closely reproduces the L-CCSD(T) interaction energies for both H$_2$O–C$_6$H$_6$ and H$_2$O–C$_{24}$H$_{12}$ dimers which, on the other hand, deviate from the corresponding DMC values by $\sim$0.2 kcal/mol and $\sim$0.5 kcal/mol, respectively. Importantly, Fig. 4 shows that, independently of the water orientation, B97M-rV also predicts interaction energies that lie within the estimated range of L-CCSD(T) values for H$_2$O–C$_{54}$H$_{18}$ in the minimum energy region.

Additional insights into the variation of the strength and anisotropy of water–PAH interactions from H$_2$O–C$_6$H$_6$ to H$_2$O–graphene are gained from the comparisons shown in Figure 9 between interaction energies calculated using the same set of exchange-correlation functionals examined in the previous sections. To facilitate the analysis of water–PAH interaction

Figure 8: Interaction energy curves for the 0-leg (left), 1-leg (middle), and 2-leg (right) of all H$_2$O–graphene dimer resulting from the B97M-rV calculations. Also shown are the corresponding DMC and RPA data taken from Ref. 91.
energies as a function of the number of carbon atoms in the PAH molecules, the comparisons are made for intermolecular distances of 3.075 Å, 3.289 Å, and 3.155 Å, corresponding to the minimum energy regions for 0-leg, 1-leg, and 2-leg configurations, respectively.

General trends can be established for 0-leg interaction energies calculated with different exchange-correlation functionals. In particular, as the size of the PAH molecules increases up to C\textsubscript{216}H\textsubscript{36}, the interaction energies predicted by rVV10, SCAN-D3, B97M-rV, and ωB97M-V effectively converge to the same value of approximately -2.6 kcal/mol. Similarly, PBE-D3, SCAN-rVV10, B3LYP-D3, and PBE0-D3 interaction energies converge to approximately -2.2 kcal/mol. Among all functionals considered in this study, revPBE-D3 predicts the weakest water–PAH interactions for 0-leg configurations.

![Graph showing interaction energies as a function of the number of carbon atoms in the PAH molecules.](image)

Figure 9: Water-PAH interaction energies calculated as a function of the number of carbon atoms in the PAH molecules. The intermolecular separation between the O atom of H\textsubscript{2}O and the PAH molecules is fixed at 3.075 Å, 3.289 Å, and 3.155 Å defining the 0-leg, 1-leg, and 2-leg configurations. RPA and DMC water–graphene values taken from Ref. \[91\] are interpolated. The CCSD(T) data for water–graphene are taken from Ref. \[92\] and were obtained using a different setup from the present L-CCSD(T) calculations. Specifically, the CCSD(T) calculations were carried out for a 3×3 graphene supercells, using the cc-pVDZ and aug-cc-pVTZ basis sets for the C and H atoms of graphene, and the O and H atoms of the water molecule, respectively. In all panels, the dashed lines do not correspond to actual data but are only used as guides to the eye.
In contrast, no common convergence behavior is predicted by the different functionals for both 1-leg and 2-leg configurations, with spreads of up to ~2 kcal/mol. As discussed in the previous section, B97M-rV and ωB97M-V provide nearly identical results from C₆H₆ to C₂₁₆H₃₆, closely reproducing the L-CCSD(T) results for water interacting with the smallest PAH molecules. Importantly, both revPBE-D3 and B3LYP-D3 predict an increase in the interaction strength for 1-leg configurations from water–benzene to water–coronene, which is opposite of the trend predicted by the available RPA, DMC, and L-CCSD(T) data. For 2-leg configurations, it should be noted that RPA and DMC predict similar interaction strengths for C₆H₆ and C₂₄H₁₂, while the available L-CCSD(T) data indicate a decrease of ~0.5 kcal/mol in the interaction energy between the two dimers. Most exchange-correlation functionals considered in this study exhibit, at least qualitatively, the same trend as L-CCSD(T), predicting a decrease in the interaction energy from benzene to coronene, which is followed by a steady increase as the number of carbon atoms increases up to C₂₁₆H₃₆. Notable exception is SCAN+rVV10 that predicts a monotonic increase of the interaction energies from C₆H₆ to C₂₁₆H₃₆. Among all functionals considered in this study, B97M-rV and ωB97M-V tend to provide the strongest interactions for 0-leg configurations, while predicting the weakest interactions for 1-leg and 2-leg configurations. Opposite trend is predicted by BLYP-D3, revPBE-D3, B3LYP-D3, and revPEB0-D3, while SCAN-D3 overall predicts the strongest interactions, independently of the water orientations.

Fig. 9 also shows comparisons of water–graphene interaction energies calculated with GGA and meta-GGA functionals with available RPA and DMC data reported in the literature. The corresponding CCSD(T) data places the water–graphene interaction at -2.49 (0-leg) and -2.84 kcal/mol (2-leg). For 0-leg configurations in the minimum energy region, revPBE-D3 reproduces the results of RPA and DMC, estimated to be around -1.7 kcal/mol, while SCAN+rVV10 data is in good agreement with the CCSD(T) value, which lie ~0.8 kcal/mol below. Compared to CCSD(T), GGA functionals, with the exception of rVV10, tend to underestimate the interaction energy for 0-leg configurations, while meta-GGA func-
tionals, with the exception of B97M-rV, provide closer agreement with the CCSD(T) value. Unfortunately, due to the associated computational cost, no data are available for water–graphene interactions calculated with hybrid functionals. However, the trends established for PAH molecules as a function of carbon atoms can be used to predict that PBE0-D3, B3LYP-D3, and revPBE0-D3, on one hand, and ωB97M-V on the other hand, respectively underestimate and overestimate the interaction energy for 0-leg configurations by ∼0.5 kcal/mol.

Before we comment further on the performance of GGA and meta-GGA functionals in describing the other two orientations of water on graphene, it should be noted that most exchange-correlation functionals predict significant variations in the interaction energy for 1-leg and 2-leg configurations relative to the largest PAH molecule (C$_{216}$H$_{36}$) considered in this study. Specifically, BLYP-D3, revPBE-D3, and SCAN-D3 predict interaction energies for 1-leg water–graphene configurations that are ∼0.1-0.3 kcal/mol smaller than for the corresponding H$_2$O–C$_{216}$H$_{36}$ configurations, while opposite trends are observed for rVV10, SCAN+rVV10, and B97M-rV, which all predict stronger interactions for water on graphene. On the other hand, with the exception of B97M-rV, all GGA and meta-GGA functionals examined in this study predict a decrease in the interaction strength for 2-leg configurations going from H$_2$O–C$_{216}$H$_{36}$ to water–graphene, although the extent of this decrease varies significantly among the different exchange-correlation functionals. These large variations in interaction energies from PAH molecules to graphene can be likely ascribed to basis set incompleteness errors in the water–graphene calculations which, directly modulating the extent of charge transfer, polarization, and electrostatic effects (see Figs. 5 and 7), affect differently the interaction strength of different water orientations. Importantly, even more pronounced differences are found between water-PAH and water-graphene interaction energies calculated at the RPA and DMC levels of theory, which may be related to (and emphasized by) similar basis set incompleteness errors.

In this context, it should be noted that, contrary to the results obtained for 0-leg configurations, noticeable differences (on the order of ∼0.5 kcal/mol) exist between RPA and DMC
interaction energies calculated for both 1-leg and 2-leg configurations of water on graphene, which are in line with the trend observed for water interacting with smaller PAH molecules. Similarly, CCSD(T) interaction energies available for 2-leg water–graphene configurations are consistently lower (by $\sim 1.0$ kcal/mol) than the corresponding RPA and DMC values, with the differences being comparable with those found for $\text{H}_2\text{O}–\text{C}_6\text{H}_6$ and $\text{H}_2\text{O}–\text{C}_{24}\text{H}_{12}$. In general, the comparisons shown in Fig. 8 indicate that the interaction energies calculated with GGA and meta-GGA functionals, which are spread over a range of $\sim 0.3$ kcal/mol (between -3.15 and -2.85 kcal/mol) for 1-leg configurations and $\sim 0.7$ kcal/mol (between -3.70 and -3.00 kcal/mol) for 2-leg configurations, seem to follow more closely the CCSD(T) results than the RPA and DMC predictions. Among the different exchange-correlation functionals examined in this study, the best interaction energy estimate of water–graphene dimer relative to CCSD(T) result is obtained by the calculation of SCAN+rVV10 functional, which is within a marginal error of 0.2 kcal/mol. Other functionals also exhibit a relatively small error within the range 0.4–0.9 kcal/mol.

4 CONCLUSIONS

We have reported a systematic analysis of the interaction energies between a single water molecule and a series of PAH molecules, from benzene to graphene, using different exchange-correlation functionals representative of the generalized gradient approximation (GGA), meta-GGA, and hybrid families. The accuracy of the different functionals has been assessed through direct comparisons with corresponding data calculated at the RPA, DMC, and L-CCSD(T) levels of theory. In this analysis, three orientations of the water molecule relative to the plane of the PAH molecules are considered, corresponding to configurations with the oxygen atom of the water molecule pointing perpendicularly to the center of the PAH molecules (0-leg configurations), one hydrogen of the water molecule pointing perpendicularly to one carbon atom of the central ring of the PAH molecules (1-leg configurations),
and both hydrogen atoms pointing to the plane of the PAH molecules (2-leg configurations).

Relatively large variability is found in the ability of different exchange-correlation functionals to describe water-PAH interactions, independently of the water orientation. In particular, it is found that GGA functionals tend to underestimate the interaction strength for 0-leg configurations, while overestimating the interaction strength for both 1-leg and 2-leg configurations. Overall, the meta-GGA B97M-rV and hybrid, meta-GGA, and range-separated $\omega$B97M-V functionals provide nearly quantitative agreement with L-CCSD(T) values available for H$_2$O–C$_6$H$_6$, H$_2$O–C$_{24}$H$_{12}$, and H$_2$O–C$_{54}$H$_{18}$. Similar trends among GGA, meta-GGA, and hybrid functionals are observed for the larger PAH molecules (up to C$_{216}$H$_{36}$) for which no RPA, DMC, and L-CCSD(T) reference data are currently available.

Further insights into the nature of water–PAH interactions are gained from absolutely localized molecular orbital energy decomposition analyses (ALMO-EDA) of the DFT results, which show that, independently of the number of carbon atoms and exchange-correlation functional, electrostatic and dispersion interactions represent the largest contributions, while polarization and charge transfer effects are negligibly small.

Finally, calculations carried out with GGA and meta-GGA functionals indicate that, as the number of carbon atoms increases, the interaction energies slowly converge to the corresponding values obtained for an infinite graphene sheet. In this context, it should be noted that relatively large deviations (up to $\sim$1.5 kcal/mol) are found among RPA, DMC, and CCSD(T) values reported in the literature for water–graphene interactions which, while highlighting the difficulties associated with converging high-level electronic structure calculations for such larger systems, makes it difficult to quantitatively and unambiguously assess the accuracy of different exchange-correlation functionals.

**Acknowledgement**

This research was supported by the National Science Foundation through grant no. CHE-
1453204 and the Air Force Office of Scientific Research through grant no. FA9550-16-1-0327 awarded to F.P. (DFT and L-CCSD(T) calculations), and the National Science Foundation through grant no. CHE-1800431 awarded to F.F. (RPA calculations). We used computational resources of the Extreme Science and Engineering Discovery Environment (XSEDE) which is supported by the National Science Foundation through grant no. ACI-1053575 under allocation TG-CHE110009 as well as of the Department of Defense.

References


37


TOC Figure