

Interaction between water and carbon nanostructures: How good are current density functional approximations?

Cite as: J. Chem. Phys. **151**, 164702 (2019); <https://doi.org/10.1063/1.5121370>

Submitted: 29 July 2019 . Accepted: 16 September 2019 . Published Online: 22 October 2019

Jan Gerit Brandenburg , Andrea Zen , Dario Alfè , and Angelos Michaelides 

COLLECTIONS

Paper published as part of the special topic on [JCP Emerging Investigators Special Collection](#)

Note: This paper is part of the JCP Emerging Investigators Special Collection.

 This paper was selected as an Editor's Pick



View Online



Export Citation



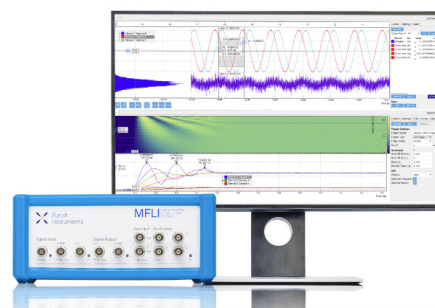
CrossMark

Challenge us.

What are your needs for periodic signal detection?



Zurich
Instruments



Interaction between water and carbon nanostructures: How good are current density functional approximations?

Cite as: J. Chem. Phys. 151, 164702 (2019); doi: 10.1063/1.5121370

Submitted: 29 July 2019 • Accepted: 16 September 2019 •

Published Online: 22 October 2019



View Online



Export Citation



CrossMark

Jan Gerit Brandenburg,^{1,a),b)}  Andrea Zen,^{2,3,a)}  Dario Alfè,^{2,3,4}  and Angelos Michaelides^{5,3,c)} 

AFFILIATIONS

¹Interdisciplinary Center for Scientific Computing, University of Heidelberg, Im Neuenheimer Feld 205A, 69120 Heidelberg, Germany

²Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom

³Thomas Young Centre and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom

⁴Dipartimento di Fisica Ettore Pancini, Università di Napoli Federico II, Monte S. Angelo, I-80126 Napoli, Italy

⁵Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

Note: This paper is part of the JCP Emerging Investigators Special Collection.

^{a)}**Contributions:** J. G. Brandenburg and A. Zen contributed equally to this work.

^{b)}**Electronic mail:** j.g.brandenburg@gmx.de

^{c)}**Electronic mail:** angelos.michaelides@ucl.ac.uk

ABSTRACT

Due to their current and future technological applications, including realization of water filters and desalination membranes, water adsorption on graphitic sp^2 -bonded carbon is of overwhelming interest. However, these systems are notoriously challenging to model, even for electronic structure methods such as density functional theory (DFT), because of the crucial role played by London dispersion forces and noncovalent interactions, in general. Recent efforts have established reference quality interactions of several carbon nanostructures interacting with water. Here, we compile a new benchmark set (dubbed **WaC18**), which includes a single water molecule interacting with a broad range of carbon structures and various bulk (3D) and two-dimensional (2D) ice polymorphs. The performance of 28 approaches, including semilocal exchange-correlation functionals, nonlocal (Fock) exchange contributions, and long-range van der Waals (vdW) treatments, is tested by computing the deviations from the reference interaction energies. The calculated mean absolute deviations on the WaC18 set depend crucially on the DFT approach, ranging from 135 meV for local density approximation (LDA) to 12 meV for PBE0-D4. We find that modern vdW corrections to DFT significantly improve over their precursors. Within the 28 tested approaches, we identify the best performing within the functional classes of generalized gradient approximated (GGA), meta-GGA, vdW-DF, and hybrid DF, which are BLYP-D4, TPSS-D4, rev-vdW-DF2, and PBE0-D4, respectively.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5121370>

I. INTRODUCTION

The interaction of molecules or liquids with nanostructured surfaces is central to many real-life applications, including catalysis, gas storage, desalination, and more. Interfaces involving water and carbon show unique and fascinating behavior, which can be employed in important applications, for instance, in water purification devices.^{1–10} Topologically similar materials can have

substantially different properties^{11–15} emphasizing that the understanding of the nature of the interaction has to be sought at a quantum mechanical electronic structure level.

Density functional theory (DFT) is the simulation method of choice for many materials applications due to its favorable accuracy to computational cost ratio.^{16–19} Modern density functional approximations (DFAs) combine semilocal expansions of the exchange-correlation with long-range corrections for missing

London dispersion interactions, i.e., the attractive part of the van der Waals (vdW) forces. The approximations are physically motivated but additionally require adjustment of a small number of parameters, which are either based on exact constraints or on empirical data. Adjusting the parameters to optimally describe short, long, and middle-ranges of interactions is challenging,^{20,21} and a solution that is good for interactions between small molecules is not necessarily good for the interaction of a molecule with extended surfaces or within the bulk (e.g., molecules in solutions and molecular crystals).^{19,22} In either case, it is mandatory to carefully benchmark the DFT methods, especially as the “zoo” of methodologies is growing and it is often unclear what is the expected reliability of a possible DFT setup or how to pick the best DFT flavor for a specific application. In the past decade, these DFA benchmarks mainly focused on molecular properties with recent studies testing more than 200 DFAs on thousands of references including thermochemistry, kinetics, and noncovalent interactions.^{23–27} Recently, some focus of DFT benchmarking moved to the description of equilibrium geometries.^{28–31} Similar large-scale benchmarks for condensed phase properties are much more rare, which is mainly due to the lack of theoretical reference data. While for bulk solids, experimental lattice constants and cohesive energies have been used successfully for DFT benchmarking,^{32–35} similar data for more weakly bound molecular crystals have substantially higher uncertainty.^{36–38} This is due to the experimental measurement uncertainty,³⁹ indirect measurements that cannot directly be compared to simple equilibrium geometries and energies, or the challenge to do the measurement itself. The latter point holds for a single water molecule adsorbed at surfaces as water readily forms clusters.⁴⁰

Concerning theoretical reference calculations, exciting progress has been made in the field of high-level wavefunction methods. On the one hand, embedding techniques^{41–43} and local approaches of coupled cluster theories^{44–49} have made the gold standard of quantum chemistry applicable to molecular systems with a few hundred atoms and molecular crystals of small molecules.^{50–52} On the other hand, new algorithmic developments in the field of diffusion Monte Carlo [DMC, a quantum Monte Carlo (QMC) technique] have made the computation of chemically accurate lattice energies of small molecular crystals feasible within reasonable computational effort.^{53,54} These developments together with an increased capacity of available computational resources make the interaction energy determination of extended systems feasible. This is an important step toward the better understanding of large noncovalently interacting systems as recently highlighted.^{19,55,56}

Here, we capitalize on this by gathering the benchmark quality interaction energies of water with carbon nanomaterials that we have studied in the past few years. This involves the adsorption of water on graphene,⁵⁷ water on benzene and coronene as two representative aromatic hydrocarbons (abbreviated as AH),⁵⁷ and water on a carbon nanotube (CNT).⁵⁸ In practical applications, it is important to describe correctly also the interaction between water molecules; thus, we additionally analyze different phases of two-dimensional (2D) ice⁵⁹ as well as bulk (3D) ice polymorphs.⁵⁴ Therefore, we obtain a dataset of 18 configurations and associated reference interaction energies, dubbed **WaC18** set. We test as many as 28 DFAs on the WaC18 set, including several recently developed vdW corrections to DFAs. Some DFT benchmarks already exist on these

or related system types,^{58–76} but they typically include a limited set of DFAs, the recent vdW developments are not included, and the used reference data are not equally well converged. Here, we will address all of these issues.

The WaC18 benchmark test will help in understanding the essential ingredients needed to describe seamlessly both the strong hydrogen bonds and the weak interaction with surfaces. Furthermore, the identification of the most accurate DFAs can be used by researchers aiming to describe these widely spread system types, complementing and updating the perspective in Ref. 77 that focused on DFT recommendations for water.

In Sec. II, we describe the benchmark systems considered, discuss the best estimates of the interaction energies including additional DMC calculations to have equally well converged reference data, and give the computational details of the DFT calculations. Following this, we report the results of a variety of DFAs and vdW corrections from several functional classes and analyze the critical aspects determining the DFA-vdW performance (Sec. III). Conclusions and a future perspective are given in Sec. IV.

II. BENCHMARK SETUP

A. Systems under consideration

In our current benchmark, we will focus on water interacting with carbon nanostructures and ice. We separate the analysis into a single water molecule interacting with graphene, a carbon nanotube (CNT), aromatic hydrocarbons (AHs), and the interaction of solid water in three-dimensional and two-dimensional ice (see Fig. 1). Water adsorption on graphene and on the AHs is tested with three different water orientations, dubbed 0-leg, 1-leg, and 2-leg. Adsorption on the CNT is considered outside the CNT in a 2-leg configuration (external) and inside the CNT in a 2-leg configuration (internal). The two-dimensional ice polymorphs have been constructed in Ref. 59 by confining a single water layer resulting in four stable polymorphs of hexagonal, pentagonal, square, and rhombic ice structures. The three bulk ice phases cover the subtle balance of the competing polymorphs Ih, II, and the high-pressure form VIII. Overall, this benchmark dataset has been designed to investigate both the water surface interaction with different adsorption motifs, different surface sizes and curvature, as well as the transferability to many-water systems at different pressures and confinements.

For all systems under consideration in our study, we use the interaction energy at fixed equilibrium geometry, where we have well converged theoretical reference energies available. We provide the reference geometries and energies as the [supplementary material](#) to make our compiled benchmark easily accessible to other researchers.

B. Reference interaction energies

The interaction energies considered are defined as the difference between a bound and an unbound configuration. Adsorption on nanostructure is computed as

$$E_{\text{int}} = E_{\text{w@nano}} - (E_{\text{w}} + E_{\text{nano}}), \quad (1)$$

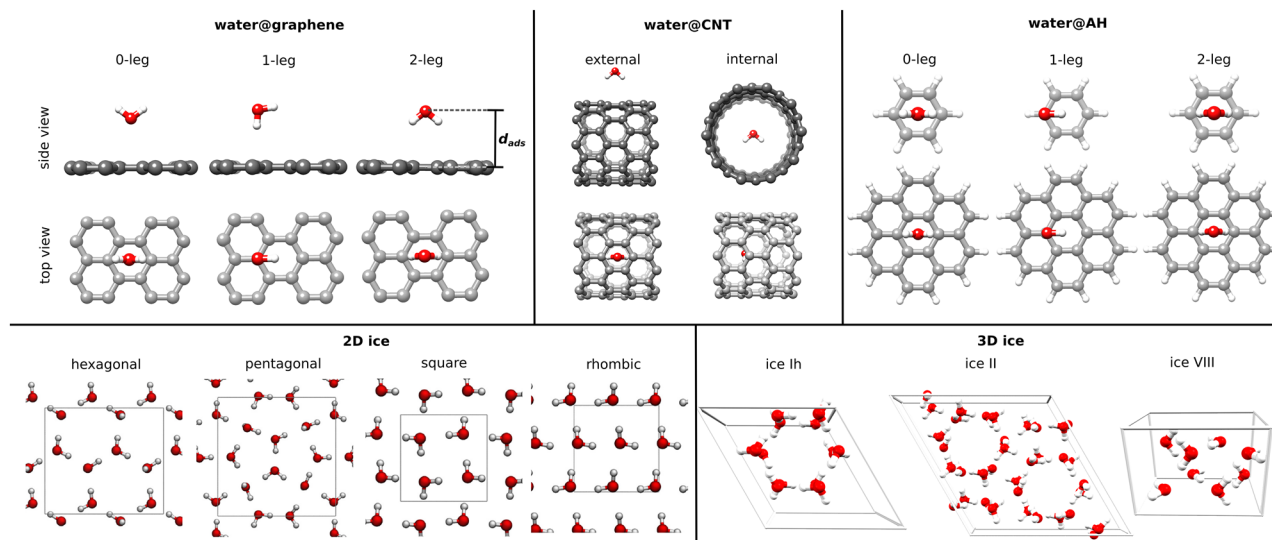


FIG. 1. WaC18 benchmark set: Structures are shown as provided in the [supplementary material](#). The top panel shows the water adsorption structures: 0-leg, 1-leg, and 2-leg motif on graphene, benzene, and coronene, as well as the adsorption inside and outside of a CNT. The lower panel shows the primitive unit cells of 2D and 3D ice structures.

where $E_{w@nano}$ is in the bound configuration and the geometries of the individual fragments E_w and E_{nano} for the unbound configuration are kept frozen.⁷⁸ The interaction energies for 2D and 3D ice are given per molecule as

$$E_{int} = E_{solid}/N_{mol} - E_w. \quad (2)$$

The reference interaction energies used in the following analysis are obtained from studies published in the past four years. All the reference values are reported in [Table I](#). The reference values for the two-dimensional ice crystals are taken from Ref. 59, while the

TABLE I. WaC18 reference interaction energies given per water molecule in meV. The reported error represents the uncertainty on the evaluation.

w@graphene ⁵⁷		w@benzene ⁵⁷		w@coronene ⁵⁷	
0-leg	-90 ± 6	0-leg	+43 ± 1	0-leg	-61 ± 3
1-leg	-92 ± 6	1-leg ^a	-124 ± 3	1-leg	-118 ± 5
2-leg	-99 ± 6	2-leg ^a	-136 ± 2	2-leg	-143 ± 4
w@CNT ^b					
External	-85 ± 18				
Internal	-287 ± 16				
3D ice ⁵⁴		2D ice ⁵⁹			
Ih	-615 ± 5	Hex.	-423 ± 3	Pent.	-419 ± 3
II	-613 ± 6	Square	-404 ± 3	Rhombic	-389 ± 3
VIII	-594 ± 6				

^aSome earlier studies reported the 1-leg structure as most stable, which might be due to small differences in numerical and geometrical setups.^{80,81}

^bEvaluation based on Ref. 58, but recomputed in this work. See details in Sec. II B.

reference values for the three bulk-ice crystals are given in Ref. 54. The reference values for water on benzene, coronene, and graphene with different orientations are taken from Ref. 57. The reference interaction energy between water and CNT where previously investigated in Ref. 58, but here we report new values that are as tightly converged as the other references.

The computational approaches used to obtain the reference values are coupled cluster theory including singles, doubles, and perturbative triples [CCSD(T)] and fixed-node DMC. In particular, the values for water on benzene and on coronene are from CCSD(T) (where DMC yields identical results within the stochastic error), and all the other values are from fixed-node DMC. Indeed, most of these systems are very challenging or even out-of-reach for CCSD(T) due to their large size and the presence of periodic boundary conditions. With DMC, it is easier to assess the binding energy in large complexes because it is straightforward to simulate periodic systems, and DMC has favorable scaling with system size. In terms of accuracy, there is generally very good agreement between CCSD(T) and DMC in the evaluation of noncovalent interactions, provided that care is taken to ensure that an accurate setup is used for each method.^{53,56,79} Agreement between methods is not expected beyond a given precision. To this aim, in [Table I](#), we report the estimated uncertainty associated with any evaluation. For the water on benzene and coronene systems, where both CCSD(T) and DMC are affordable, there is excellent agreement among them.⁵⁷

Although reported DMC results are coming from different studies, the setup is consistent among them. DMC simulations were carried out with the CASINO code.⁸² Dirac-Fock pseudopotentials^{83,84} with the localization approximation⁸⁵ (LA) are used. The trial wavefunctions were of the Slater-Jastrow type with single Slater determinants and the single particle orbitals obtained from DFT-local density approximation (LDA) calculations performed with the plane-wave self consistent field program (PWscf)⁸⁶ and

re-expanded in terms of B-splines.⁸⁷ The Jastrow factor included electron-electron, electron-nucleus, and electron-electron-nucleus terms. The parameters of the Jastrow factor were carefully optimized by minimizing the variance, within a variational QMC scheme. The recently developed size-consistent DMC algorithm (ZSGMA)⁵³ was used. Finite time-step errors are carefully minimized by performing simulations with different values of the time step, until the bias appears safely smaller than the stochastic error. In periodic calculations, finite-size corrections are applied either using the model periodic Coulomb interaction^{88–90} or the Kwee, Zhang, and Krakauer⁹¹ approach.

We reevaluate here the binding energy of water on the CNT because one specific aspect of the setup used in Ref. 58 is now known to be a possible issue: the optimization of the Jastrow factor for the configuration of water inside the CNT was not optimal. Since in the standard LA approach the Jastrow factor plays a major role in the pseudopotential error, we have developed a new method (to be reported elsewhere⁹²) that removes the LA bias. In repeating the evaluation, we also tuned some other aspect of the setup to be in line with the actual state-of-the-art. In particular, we used the recently developed eCEPP pseudopotentials⁹³ and the ZSGMA algorithm.⁵³ The reported binding values are obtained with time step $\tau = 0.01$ a.u., which gives errors less than the stochastic uncertainty (standard deviation of 18 meV).

C. DFT computational details

DFT test calculations based on the Perdew-Burke-Ernzerhof (PBE) functional⁹⁴ were done on selected systems to ensure convergence of all relevant numerical settings. Several different electronic structure codes and orbital basis expansions have been employed: ORCA 4⁹⁵ with large aug-cc-pVQZ, aug-cc-pV5Z,^{96,97} and def2-QZVPPD⁹⁸ basis sets; CRYSTAL17^{99,100} with a def2-QZVPPD(-f) basis; and Vienna *Ab initio* Simulation Package (VASP 5.4)^{101,102} with projector-augmented plane waves (hard PAWs^{103,104}) with an energy cutoff of 1000 eV. In all codes, tight self-consistent field settings and large integration (and fine FFT) grids are used. The Brillouin zone sampling has been increased to converge the interaction energy to 1 meV. For the adsorption on graphene (5×5 supercell with 53 atoms in the unit cell) and the CNT [nonmetallic (10,0) configuration with 83 atoms in the unit cell], this reduces to a Γ -point calculation. For all PAW calculations, the nonperiodic directions use a vacuum spacing of 20 Å for the adsorbed geometries and the same unit cell is used for the individual fragments, which compensates possible remaining image interactions for the binding energies. CRYSTAL and ORCA use open conditions in the nonperiodic directions consistent with the reference calculations.

For the water at the AH system, we established that the PBE interaction energy is converged within 2 meV using the def2-QZVPPD basis and the codes ORCA and CRYSTAL yield results within 1 meV. We additionally compared the PBE/def2-QZVPPD interaction energy with the PBE/hard PAW/1000 eV ones for water at graphene, water at AH, and ice Ih with a maximum deviation of 2 meV. To reach the DFT convergence for these systems, unusually tight thresholds are needed. In Table II, we list the convergence of the PBE ice Ih lattice energy for three complementary basis set expansions and software codes. For the PAW

TABLE II. Lattice energy convergence in meV of ice Ih using various numerical settings based on PBE and a Γ -centered $3 \times 3 \times 3$ k -grid.^{105,106,161–163} The best estimates are highlighted in bold.

VASP PW cutoff (eV)	PAW		
	Soft	Normal	Hard
500	−798.4	−665.3	−646.2
700	−798.1	−664.6	−636.8
1000	−798.0	−664.7	− 637.1
Crystal		w.o.c. ^a	CP-corrected
def2-mSVP		−1104.2	−858.8
def2-TZVPP		−721.2	−646.7
def2-QZVPP		−665.6	−638.2
CBS(TZ,QZ) ^b		−657.9	− 637.0

^aSupramolecular approach without counterpoise (CP) correction.

^bBasis set extrapolation using optimized exponents.¹⁰⁷

code, VASP, hard PAWs (i.e., small potential core) are needed as well as a minimum PW cutoff of 700 eV. CRYSTAL employs an all-electron basis set with atom-centered functions. Here, the interaction energy is converged employing neither a counterpoise-corrected def2-TZVPP nor a def2-QZVPP calculation. For fully converged values, a counterpoise-corrected def2-QZVPP calculation is needed and an extrapolation to the basis set limit reduced the binding by only 1.2 meV. Production level calculations for all reported DFT interactions on the full WaC18 benchmark are performed with VASP 5.4 using hard PAWs and a PW energy cutoff of 1000 eV.

DFAs from several rungs are tested: local density approximation (LDA), generalized gradient approximation (GGA), meta-GGA, and hybrid functionals (incorporation of nonlocal Fock exchange). The semilocal DFAs are corrected for missing long-range correlation effects by means of a variety of different semiclassical and nonlocal density based corrections (D2,¹⁰⁸ D3,^{109,110} D4,^{111,112} TS,¹¹³ MBD,¹¹⁴ VV10,¹¹⁵ dDSC,¹¹⁶ and vdW-DF¹¹⁷). TS and MBD are used with the noniterative Hirshfeld partitioning, D3 is used in the rational damping variant including Axilrod-Teller-Muto type three-body contributions, for D4, partitioned charges are generated by the electronegativity equilibration procedure (EEQ), and many-body contributions are covered by a standard coupled fluctuating dipole expression retaining an RPA-like expression. See Refs. 118–121 for further overview on vdW interactions in the DFT framework.

III. RESULTS AND DISCUSSION

Using the DMC and CCSD(T) high-quality interaction energies, we can now test the capability of standard and new DFAs and a few simplified electronic structure methods for water adsorption on nanostructured surfaces and within ice polymorphs. While discussing the performances of each method, it is important to keep in mind the numerical DFT uncertainty of about 2 meV as well as the

reference uncertainty, ranging from 1 meV (w@benzene, 0-leg) to 18 meV (w@CNT, external).

Table III summarizes the individual interaction energies from several electronic structure approaches and gives a root-mean-square (rms) error over the full WaC18 set. Additional statistical information on the performance of the methods is given in Table IV.

A. Impact of vdW interactions

From looking at Tables III and IV, it is clear that Hartree-Fock (HF) misses all (Coulomb) correlation effects and cannot describe any of these noncovalently bound systems appropriately.^{139,140} All systems are systematically underbound by up to 342 meV. While there is some weak binding for water on benzene, this diminishes for larger substrates and becomes repulsive for the adsorption on graphene and on the CNT. Thus, exchange repulsion, induction, and electrostatics are not sufficient to lead to a net binding of water on the carbon nanostructures considered. This is consistent with our symmetry adapted perturbation theory analysis in Ref. 57 as well as many studies on molecular dimers.¹⁴¹ That the local density approximation (LDA) yields an inconsistent description of small vdW complexes has been known since the mid-1990s.¹⁴² This is confirmed here, where LDA systematically overbinds all systems, in particular, the ice polymorphs. The LDA results for water adsorption, on the other hand, seem reasonably good. However, this is a fortuitous event due to the fixed geometries. While all other DFAs give equilibrium adsorption distances within 0.1–0.2 Å (see Refs. 22, 58, and 60), the LDA minimum is at substantially smaller distance and cannot be recommended for either geometry or stability estimates of vdW bound systems.

Including semilocal exchange-correlation effects as in the popular PBE GGA functional improves the behavior although most systems are bound a bit too weakly. Clearly, the long-range correlation effects leading to vdW attraction are missing. In the past decade, several methods have been developed for including these missing interactions (see, e.g., Refs. 118, 119, and 121). We test a broad range of these vdW corrections coupled with PBE (see Tables III and IV). The error spread is still substantial, and, in particular, the older effective pair-wise schemes (PBE-VV10, PBE-dDsC, PBE-TS, and PBE-D2) do not perform well. Recent vdW developments pay off, and we can see a clear improvement of PBE-MBD¹¹⁴ and PBE-D4¹¹² over their predecessors. The many-body vdW contributions decrease the binding yielding better agreement with the references. Most of this effect is already covered at the Axilrod-Teller-Muto type three-body level^{143,144} as included in the D3 method.¹⁰⁹ At the PBE-vdW level, only D4 and VV10 are able to reproduce the relative stability of the water adsorption, i.e., 0-leg vs 2-leg and 1-leg vs 2-leg stability, to good accuracy (coming within the error of the reference energy). The best PBE based method (PBE-D4) yields an excellent description of the water adsorption with mean absolute deviation (MAD) from the reference of 13 meV. The description of the ice polymorphs is less satisfactory, which can be traced back to the intrinsic overpolarization and thus overestimated induction interaction of the PBE functional.^{145–149} For instance, uncorrected PBE already overbinds hexagonal ice Ih by 24 meV, which clearly is not corrected by a (mostly) attractive vdW interaction. Overall, PBE-vdW does not perform well for strong hydrogen bonded systems.⁷⁷

DFAs with a nonlocal kernel to describe vdW interactions have been pioneered by Dion *et al.* (vdW-DF1).^{117,150} This first-generation nonlocal functional gives unsatisfactory results on our benchmarks, the adsorption strength is overestimated, and the ice lattice energy is underestimated. The revised version with optimized semilocal exchange-correlation optB86b-vdW improves upon this, but the overall MAD is at 44 meV still rather high. Binding to graphene and the CNT seems to be extremely challenging for the nonlocal functionals with maximum deviation of 168 meV for water inside the CNT, as already noted in Ref. 58. Consistent with previous studies,^{130,131} the second generation of nonlocal functionals significantly improves the results on all benchmark systems and the revised variant rev-vdW-DF2 more than halves the overall MAD to 20 meV. Only water inside the CNT remains challenging being overestimated by 72 meV, which is worse than all other vdW corrected semilocal DFAs (with the exception of PBE-TS).

B. Performance by Jacob's ladder classification

For a better visual comparison of the performance of the DFAs for the different WaC18 subsets, we show a graphical representation of the individual rms errors separated into the different functional classes in Fig. 2. LDA is not reported as it yields very bad results. In the GGA panel, we show the results for the three most accurate GGA functionals, after inclusion of D4 for long-range vdW interactions. PBE-D4 is not included as some of the various modifications of the PBE exchange enhancement factors prove better, most notably RPBE and revPBE that are both known to give more reasonable hydrogen bond strengths.^{38,148} While the water adsorption computed by RPBE-D4 and revPBE-D4 is very similar to PBE-D4, we see a clear improvement for the 2D/3D ice polymorphs. However, the MADs at 29 and 37 meV are still unexpectedly high. Especially, the denser ice structures (rhombohedral 2D-ice and high-pressure ice VIII) are systematically underbound. Note that the use of normal PAWs or smaller basis set expansions results in a systematic shift toward more strongly bound systems (see Table II), removing most of the bias for RPBE-D4 and revPBE-D4 and giving artificially better results.^{151,164–166} The most successful GGA tested by us is BLYP-D4 giving a very consistent performance with MADs below 20 meV for all considered benchmark sets.

Including higher derivatives (like the kinetic energy density τ) in the exchange-correlation enhancement factors give rise to the meta-GGA class. Formally, their computational cost scales with system size as the GGAs, but a stable self-consistent field convergence can be numerically more involved and typically requires larger integration grids.^{20,152} TPSS is based on the PBE GGA and has a rather weak τ -dependency but still improves over PBE for many physical properties.²³ This also holds for our benchmark systems, TPSS-D4 has a rather balanced description of very accurate adsorption energies and decently good ice lattice energies, and its overall MAD of 18 meV is identical to that of BLYP-D4. SCAN and M06L are modern meta-GGAs that can cover part of the medium-range correlation and have been shown to yield good structures and energies for diversely bonded systems.^{126,153} Still, both underbind all water adsorption systems, which can be partially compensated by correction schemes (see SCAN-D4). However, since SCAN already includes some vdW forces, it is nontrivial to

TABLE IV. Mean deviation (MD) and mean absolute deviation (MAD) in meV of the interaction energies from the references; see Table I and Fig. 1. Best estimates are highlighted in bold.

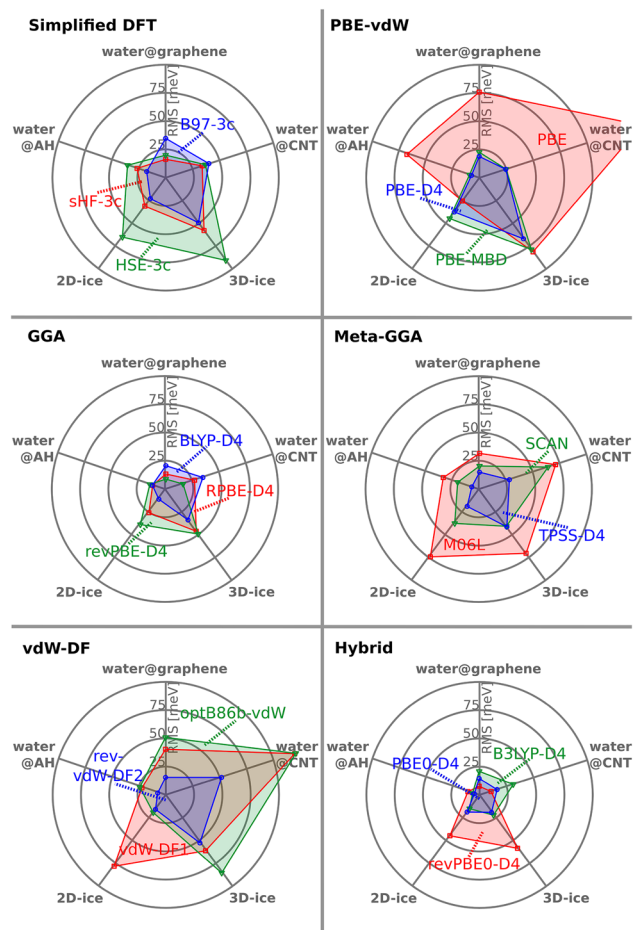
Method	w@nano ^a		2D/3D-ice		All	
	MD ^b	MAD	MD	MAD	MD	MAD
Local density approximation						
LDA	-20	29	-302	302	-130	135
PBE with various vdW corrections						
PBE	79	79	30	40	60	64
PBE-VV10	-30	30	-83	83	-51	51
PBE-dDsC	-32	32	-77	77	-49	49
PBE-TS	-40	40	-49	49	-43	43
PBE-MBD	-12	14	-55	55	-29	30
PBE-D2	-18	20	-65	65	-37	38
PBE-D3	-10	13	-46	46	-24	26
PBE-D4	-12	13	-42	43	-24	25
GGAs with D4 vdW correction						
RPBE-D4	-4	14	24	29	7	20
revPBE-D4	2	12	35	37	15	21
BLYP-D4	-10	18	-12	18	-10	18
meta-GGA (with D4 vdW correction)						
M06L	19	37	68	68	38	49
SCAN	22	23	-38	38	-1	29
SCAN-D4	-16	16	-72	72	-38	38
TPSS-D4	-6	12	-5	27	-6	18
1st and 2nd generation vdW-DFs						
vdW-DF1	-32	39	70	70	7	51
optB86b-vdW	-44	44	-45	45	-44	44
vdW-DF2	-26	28	4	11	-14	21
Rev-vdW-DF2	-11	16	-26	26	-17	20
Hybrids (with D4 vdW correction)						
HF	142	142	247	247	182	182
HF-D4	1	12	79	79	32	39
revPBE0-D4	2	10	48	48	20	25
B3LYP-D4	-11	14	-14	16	-12	15
PBE0-D4	-7	9	-13	16	-9	12
Simplified density functional approximations						
sHF-3c	8	20	-32	40	-8	28
HSE-3c	-16	29	-70	70	-37	45
B97-3c	-25	25	-26	28	-26	26

^aCombination of the three adsorption sets water@graphene, water@CNT, and water@AH.

^bNegative MD means too strongly bounded system.

combine SCAN with correction schemes.^{20,21} This is particularly relevant for 2D/3D-ice, where plain SCAN already overbinds all systems.

The next DFA rung requires the inclusion of nonlocal (Fock) exchange resulting in hybrid functionals. The most widely used hybrid DFAs are PBE0 and B3LYP, and while they are typically only of medium quality for many chemical properties,²³ PBE0-D4 and

**FIG. 2.** Root-mean-square (rms) errors in meV of the computed interaction energies for the WaC18 benchmark, separated into five subsets with various theoretical methods. The systems are defined in Fig. 1 and Table I.

B3LYP-D4 consistently improve over their GGA parents. In particular, the improvement for the 2D/3D ice polymorphs is significant. Of all tested methods, PBE0-D4 has the closest agreement with the reference interaction energies for all considered systems with an overall smallest MAD of 12 meV. Importantly, all tested relative stability sequences (ice polymorphs and the adsorption motifs) of PBE0-D4 are correct. The relative adsorption energies of the different binding motifs on graphene and the CNT are actually within the stochastic uncertainty of the DMC references.

The simplified DFAs (sHF-3c,^{135,136} HSE-3c,¹³⁷ B97-3c;¹³⁸ see Ref. 154 for an overview) give mixed results. Overall, their accuracy is similar to the average vdW corrected DFA. Especially, HSE-3c has problems at describing the strong hydrogen bonds in 2D/3D-ice, most likely due to remaining basis set superposition errors that cannot be fully compensated. B97-3c, on the other hand, employs a slightly larger basis set expansion and gives reasonably balanced results. As those methods have been designed for increased computational speed (speedup of up to 2 orders of magnitude compared to

converged basis set DFT¹⁵⁴), they might still be useful for screening applications.

C. Essential ingredients for well-balanced DFA

We find the following points essential for a well-balanced description of both water adsorbed on nanostructures as well as within ice polymorphs:

- Correlation effects beyond the local density approximation (avoid HF and LDA).
- Use of a modern vdW correction (D3/D4, MBD, or vdW-DF2).
- Converged numerical settings with hard cores for PAWs or expansions beyond triple- ζ quality for atom-centered basis sets.
- GGA, meta-GGA, and hybrid DFAs are similarly good for adsorption.
- Fock exchange improves strong H-bonds in ice polymorphs.

As shown in Fig. 2, the best GGA, meta-GGA, vdW-DF, and simplified DFT methods (BLYP-D4, TPSS-D4, rev-vdW-DF2, and B97-3c, respectively) fulfill the first three points and perform rather similarly. Significantly more accurate are the best hybrid functionals mainly due to their improved description of 2D/3D ice. In terms of computational cost, GGAs seem to have the best accuracy vs effort ratio, while hybrids should be used when aiming for the highest accuracy. The most successful DFAs

have errors consistently well below chemical accuracy (1 kcal/mol = 43 meV), challenging experimental errors of sublimation enthalpies.^{39,155}

D. Comparison to water scoring scheme

In a previous effort to judge “How good is DFT for water?”⁷⁷ a scoring scheme has been devised to judge the performance of approximated methods for the properties of the water monomer, the dimer, the hexamer, and ice structures. Physical quantities scored are monomer symmetric stretch frequency f_{ss}^{mono} , dimer binding energy E_b^{dim} , ring-hexamer binding energy per monomer E_b^{ring} , ice Ih lattice energy E_{Ih} , difference $\Delta E_b^{\text{pr-ring}}$ of binding energies per monomer of prism and ring isomers of the hexamer, difference $\Delta E_{\text{Ih-VIII}}$ of lattice energies of ice Ih and VIII, equilibrium O-O distance R_{OO} in dimer, and equilibrium volumes per monomer $V_{\text{eq}}^{\text{Ih}}$, $V_{\text{eq}}^{\text{VIII}}$ of ice Ih and VIII.¹⁵⁶ For more details on the scoring system, see Ref. 77. For some of the DFAs examined in our benchmark study, we report their performance for this scoring system in Table V. As expected, LDA and HF are not reliable to describe water though some individual scores like the relative stability of the prism and ring hexamer are fortuitously good. The same holds for uncorrected as well as dispersion corrected PBE as already recognized in the original study.⁷⁷ On the other hand, BLYP-D4, TPSS-D4, optB88-vdW are performing reasonably well. Especially energetic and geometric properties are well reproduced although the lattice energy of ice Ih seems to be problematic. Symmetric stretch frequencies of the water monomer

TABLE V. Percentage scores of selected DFAs using the scoring scheme from Ref. 77. Physical quantities scored are monomer symmetric stretch frequency f_{ss}^{mono} , dimer binding energy E_b^{dim} , ring-hexamer binding energy per monomer E_b^{ring} , ice Ih lattice energy E_{Ih} , difference $\Delta E_b^{\text{pr-ring}}$ of binding energies per monomer of prism and ring isomers of the hexamer, difference $\Delta E_{\text{Ih-VIII}}$ of lattice energies of ice Ih and VIII, equilibrium O-O distance R_{OO} in dimer, and equilibrium volumes per monomer $V_{\text{eq}}^{\text{Ih}}$, $V_{\text{eq}}^{\text{VIII}}$ of ice Ih and VIII. The total score in the final column is the average of the individual scores, and unmarked values have been computed in this study.

	f_{ss}^{mono}	E_b^{dim}	E_b^{ring}	E_{Ih}	$\Delta E_b^{\text{pr-ring}}$	$\Delta E_{\text{Ih-VIII}}$	R_{OO}	$V_{\text{eq}}^{\text{Ih}}$	$V_{\text{eq}}^{\text{VIII}}$	Total
Reference	3812 cm ^{-1a}	-216 meV ^a	-319 meV ^a	-615 meV ^b	13 meV ^a	21 meV ^b	2.91 Å ^a	30.9 Å ^{3c}	19.1 Å ^{3c}	100
Tolerance	20 cm ⁻¹	10 meV	10 meV	10 meV	10 meV	10 meV	0.01 Å	1%	1%	
LDA	60	0	0	0	100	0	0	... ^d	... ^d	23
HF	0	30	0	0	80	90	0	... ^d	... ^d	29
PBE	50	100	90	80	0	0	90	100 ^e	20 ^e	59
PBE-D4	50	90	60	0	100	10	80	50 ^f	100 ^f	60
BLYP-D4	30	100	90	60	100	40	100	90 ^f	80 ^f	77
TPSS-D4	50	100	80	50	90	10	80	60 ^f	70 ^f	66
optB88-vdW	60 ^e	100 ^e	90 ^e	20 ^e	100 ^e	100 ^e	50 ^e	80 ^e	100 ^e	78
PBE0-D4	80	100	80	80	90	70	100	70 ^g	70 ^g	82
B97-3c	70	90	70	30	100	30	80	50	70	66

^aReferences taken as gathered in Ref. 77.

^bFor consistency within this article, references taken from Table I.

^cValues taken from Ref. 38 to consistently exclude zero-point and thermal effects.

^dValues not computed as expected to be unreliable.

^eValues taken from Ref. 77.

^fValues taken from Ref. 38 using the D3 dispersion correction.

^gValues taken from Ref. 77 using the TS dispersion correction.

are as usual underestimated by all (meta-)GGA functionals. Here, it is worth pointing out that out of the nonhybrid functionals, the low-cost method B97-3c yields the best frequencies and overall performs competitively to the dispersion corrected DFAs. In agreement with our current benchmark, PBE0-D4 is the best performing method yielding an overall score of 82, which is indeed higher than any other DFT method tested on this set so far.

IV. CONCLUSIONS

We have gathered and computed well converged reference interaction energies of water with carbon nanostructures and of water within two- and three-dimensional ice polymorphs compiled in the new WaC18 benchmark set. Combined, this gives a challenging set of large and complex systems, ideally suited to benchmark approximated methods. Importantly, those systems are larger than standard noncovalent interaction benchmark sets based on molecular dimers of small to medium sized molecules like S22¹⁴¹ and S66.¹⁵⁷ The 0D, 1D, 2D, and 3D periodicity covered here also includes new aspects compared to benchmark sets of large supramolecular complexes like L7¹⁵⁸ and S12l.¹⁵⁹ In contrast to benchmarks using back-corrected experimental references, our high-level theoretical interaction energies make the comparison much more straightforward without the complication of thermal and zero-point energy contributions. We span a broad range of interaction energies from nonbinding (water@benzene, 0-leg motif) to moderately strong binding (ice Ih with lattice energy of by -615 meV). Figure 3 shows an overview of the different binding strengths by correlating the reference energies with a few considered DFAs. The correlation highlights that we roughly follow Jacob's ladder classification of DFAs with HF and LDA being unreliable, and PBE, PBE-D4, and PBE0-D4 step by step increasing the accuracy. Of all methods considered, BLYP-D4, TPSS-D4, rev-vdW-DF2, and PBE0-D4 are the most accurate within their respective functional

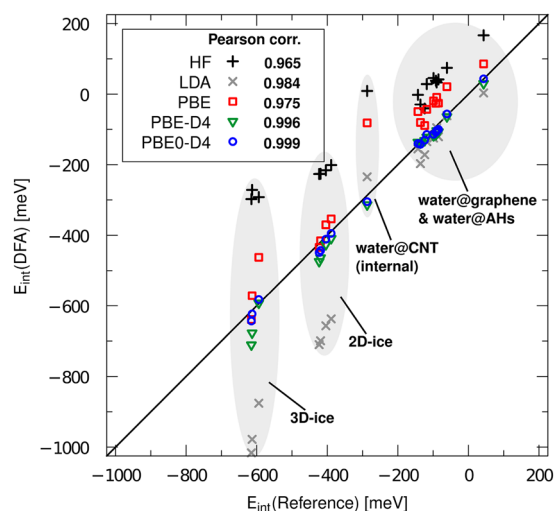


FIG. 3. Correlation between reference energies and interaction energies computed with a selection of density functional approximations and HF. Explicit data are given in Table III.

class. Replacing D4 with the older D3 or the MBD vdW correction leads to minor deterioration and can be used when D4 is not available. Our present benchmark focuses on specific noncovalent interactions only; from previous studies, it is known that TPSS-D3 and PBE0-D3 yield very good equilibrium geometries²⁸ and organometallic reaction energies.¹⁶⁰ In the large GMTKN55 benchmark,²³ BLYP-D3 and TPSS-D3 are among the best performing GGAs and meta-GGAs, respectively, consistent with our findings.

We see our benchmark results as a guideline for future simulation studies of water in the condensed phase (liquid or solid) and water-carbon nanostructure interfaces. Additionally, the provided WaC18 dataset can help as a challenging cross check for other DFAs, classical force fields, machine learning potentials, tight-binding Hamiltonians, and to test other many body electronic structure methods such as the Random Phase Approximation (RPA) or Møller-Plesset (MP) perturbation theory.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for geometries and reference energies of all considered structures of the WaC18 benchmark set, including water@graphene, water@CNT, water@AH, 2D-ice, and 3D-ice. Geometries are provided in VASP POSCAR and XYZ format.

ACKNOWLEDGMENTS

We thank Stefan Grimme and Eike Caldeweyher for providing access to the *dftd4* code. J.G.B. acknowledges support by the Alexander von Humboldt Foundation. A.Z. and D.A.'s work is sponsored by the Air Force Office of Scientific Research, Air Force Material Command, U.S. Air Force, under Grant No. FA9550-19-1-7007. A.Z. and A.M. were supported by the European Research Council (ERC) under the European Union's Seventh Framework Program (No. FP/2007-2013)/ERC Grant Agreement No. 616121 (Heterolce project). This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract No. DE-AC05-00OR22725. We are also grateful, for computational resources, to the ARCHER UK National Supercomputing Service, United Kingdom Car-Parrinello (UKCP) consortium (No. EP/F036884/1), the London Centre for Nanotechnology, University College London (UCL) Research Computing, and the UK Materials and Molecular Modelling Hub, which is partially funded by EPSRC (No. EP/P020194/1).

REFERENCES

- ¹L. Fumagalli, A. Esfandiari, R. Fabregas, S. Hu, P. Ares, A. Janardanan, Q. Yang, B. Radha, T. Taniguchi, K. Watanabe, G. Gomila, K. S. Novoselov, and A. K. Geim, *Science* **360**, 1339 (2018).
- ²J. Abraham, K. S. Vasu, C. D. Williams, K. Gopinadhan, Y. Su, C. T. Cherian, J. Dix, E. Prestat, S. J. Haigh, I. V. Grigorieva, P. Carbone, A. K. Geim, and R. R. Nair, *Nat. Nanotechnol.* **12**, 546 (2017).
- ³R. K. Joshi, P. Carbone, F. C. Wang, V. G. Kravets, Y. Su, I. V. Grigorieva, H. A. Wu, A. K. Geim, and R. R. Nair, *Science* **343**, 752 (2014).
- ⁴R. R. Nair, H. A. Wu, P. N. Jayaram, I. V. Grigorieva, and A. K. Geim, *Science* **335**, 442 (2012).
- ⁵G. Algara-Siller, O. Lehtinen, F. C. Wang, R. R. Nair, U. Kaiser, H. A. Wu, A. K. Geim, and I. V. Grigorieva, *Nature* **519**, 443 (2015).

occur due to updated references and slightly different numerical settings. Here, all molecular properties are computed with Orca in a larger def2-QZVPDD basis set, large integration grid, and tight convergence thresholds.

- ¹⁵⁷J. Řezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011).
- ¹⁵⁸R. Sedlak, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay, and P. Hobza, *J. Chem. Theory Comput.* **9**, 3364 (2013).
- ¹⁵⁹S. Grimme, *Chem. Eur. J.* **18**, 9955 (2012).
- ¹⁶⁰S. Dohm, A. Hansen, M. Steinmetz, S. Grimme, and M. P. Checinski, *J. Chem. Theory Comput.* **14**, 2596 (2018).
- ¹⁶¹Y. Yao and Y. Kanai, *J. Chem. Phys.* **146**, 224105 (2017).
- ¹⁶²J. Yang, L. Z. Tan, and A. M. Rappe, *Phys. Rev. B* **97**, 085130 (2018).
- ¹⁶³I. Hamada and S. Yanagisawa, *Phys. Rev. B* **84**, 153104 (2011).
- ¹⁶⁴M. Dodia, T. Ohto, S. Imoto, and Y. Nagata, *J. Chem. Theory Comput.* **15**, 3836 (2019).
- ¹⁶⁵L. Ruiz Pestana, O. Marsalek, T. E. Markland, and T. Head-Gordon, *J. Phys. Chem. Lett.* **9**, 5009 (2018).
- ¹⁶⁶L. Ruiz Pestana, N. Mardirossian, M. Head-Gordon, and T. Head-Gordon, *Chem. Sci.* **8**, 3554 (2017).