

The water-benzene interaction: Insight from electronic structure theories

Jie Ma,^{1,2,3} Dario Alfè,^{2,4,5} Angelos Michaelides,^{2,3,a)} and Enge Wang¹

¹*Institute of Physics, Chinese Academy of Sciences, P.O. Box 603, Beijing 100190, China*

²*London Centre for Nanotechnology, University College London, London WC1E 6BT, United Kingdom*

³*Department of Chemistry, University College London, London WC1E 6BT, United Kingdom*

⁴*Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom*

⁵*Department of Earth Sciences, University College London, London WC1E 6BT, United Kingdom*

(Received 31 January 2009; accepted 10 March 2009; published online 16 April 2009)

Weak noncovalent interactions such as van der Waals and hydrogen bonding are ubiquitous in nature, yet their accurate description with electronic structure theories is challenging. Here we assess the ability of a variety of theories to describe a water-benzene binding energy curve. Specifically, we test Hartree–Fock, second-order Møller–Plesset perturbation theory, coupled cluster, density functional theory with several exchange–correlation functionals with and without empirical vdW corrections, and quantum Monte Carlo (QMC). Given the relative paucity of QMC reports for noncovalent interactions, it is interesting to see that QMC and coupled cluster with single, double, and perturbative triple excitations [CCSD(T)] are in very good agreement for most of the binding energy curve, although at short distances there are small deviations on the order of 20 meV. © 2009 American Institute of Physics. [DOI: [10.1063/1.3111035](https://doi.org/10.1063/1.3111035)]

Weak noncovalent interactions such as van der Waals (vdW) dispersion forces and hydrogen bonds play a central role in holding soft organic and inorganic matter together such as DNA and liquid water. However, the accurate description of these interactions remains a major challenge for *ab initio* electronic structure theories. Density functional theory (DFT), which is widely used in electronic structure calculations, is often inadequate in describing such interactions because of the local or semilocal nature of common exchange–correlation (xc) functionals in widespread use. Some mature quantum chemistry methods, such as second-order Møller–Plesset perturbation theory (MP2) and coupled cluster, can often give the desired accuracy for noncovalent interactions. However, the canonical versions of these theories scale as N^5 or worse, where N is the number of basis functions, so that their application to large systems rapidly becomes computationally prohibitive.¹ Furthermore, their application to periodic systems with large basis sets remains a major challenge and MP2 is not suitable for metallic systems. Quantum Monte Carlo (QMC),² which is available for both periodic and isolated systems, is emerging as a promising approach that can achieve high accuracy for weak interactions.^{3–8} Even the simplest fixed node diffusion Monte Carlo (DMC) method with a single Slater–Jastrow determinant has been shown to give precise results for a wide variety of systems.^{9–14} For example, it gives the correct energy ordering of the low energy isomers of the water hexamer¹³ and the true asymptotic form between metallic wires and layers¹⁴ and high precision (on average less than 1 kcal/mol error) for the S22 benchmark set,¹⁵ the latter emanating from the very promising QMC@Home project.¹⁰ However, considerably more work is needed to establish the widespread validity of

DMC for noncovalent interactions, and especially for energy–distance curves.

In this paper, we briefly compare the performance of several methods in describing the binding energy curve between water and benzene. Water benzene is an interesting model system because it is a reasonably small cluster to which explicitly correlated quantum chemistry methods with large basis sets can be applied. In addition, it is a weakly interacting system involving a complex mixture of hydrogen bonding and vdW bonding, thus providing a stern test for any electronic structure theory. Further, benzene—the smallest aromatic molecule—is a useful prototype for studying sp^2 -bonded systems and its interaction with water is used as the simplest model for the water–graphite interaction. Indeed, for these and other reasons, the water–benzene system has been widely examined before with explicitly correlated and DFT approaches.^{10,16–23} Here, we compare the performance of Hartree–Fock (HF), MP2, coupled cluster with single and double excitations (CCSD), coupled cluster with single, double, and perturbative triple excitations [CCSD(T)], DMC, and several DFT xc functionals with and without empirical vdW corrections. The performance of the traditional quantum chemistry and DFT approaches is very much as expected and consistent with previous studies on this and related systems.^{17,18} DMC and the best of the traditional quantum chemical approaches, namely, CCSD(T), are in excellent agreement at the binding energy minimum and for the long range decay. However, there is a small deviation between DMC and CCSD(T) at short water–benzene distances, leading to a repulsive wall that is too steep compared to CCSD(T). Given the central importance of the shape of the repulsive wall to the dynamics of systems, this small difference between CCSD(T) and DMC warrants further investigation and may be relevant to future explicitly correlated

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

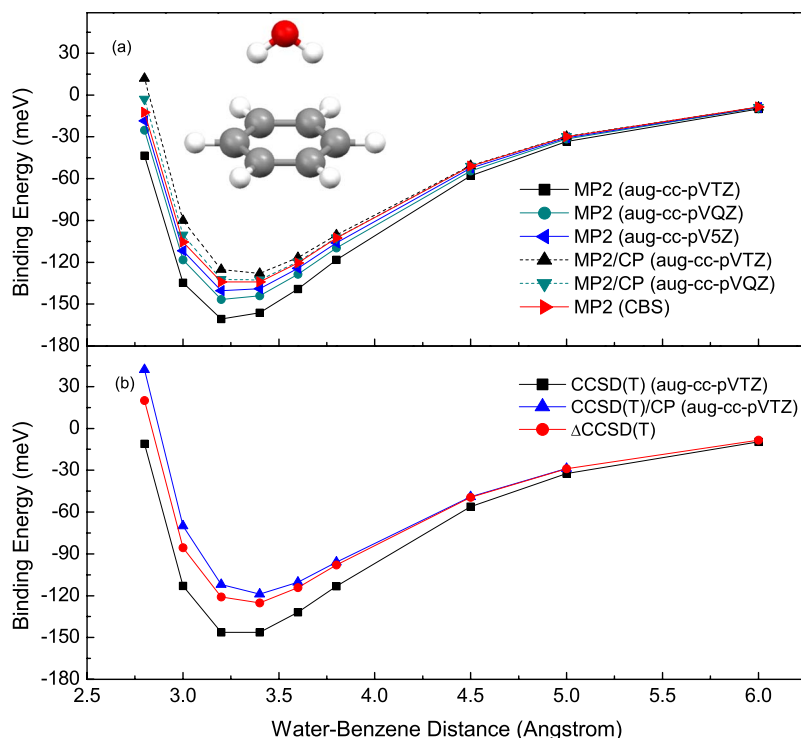


FIG. 1. (Color online) (a) Binding energy curves for uncorrected and CP-corrected MP2 with a variety of basis sets and at the CBS-extrapolated limit. (b) Binding energy curves for uncorrected and CP-corrected CCSD(T) with an aug-cc-pVTZ basis set. The Δ CCSD(T) curve [obtained from Eq. (1)] is also shown. The inset in (a) shows the orientation of the water-benzene complex considered, with the distance between the water and benzene taken as the vertical distance between the O atom of the water molecule and the center of the benzene ring. The lines connecting the points are merely a guide to the eye.

simulations on more complex systems such as aqueous water-benzene or water-graphite interfaces.

All the HF, MP2, and coupled cluster calculations have been performed with the GAUSSIAN03 code.²⁴ Since our purpose here is to benchmark different energy methods rather than to obtain the most reliable binding energy curve for the water-benzene system, we have chosen not to optimize the structure of the benzene-water complex but instead to use the same single set of structures throughout for all electronic structure theories. Specifically, we have used the equilibrium water and benzene structures²⁵ obtained from density functional calculations with the PBE xc functional.²⁶ The oxygen atom of the water molecule is placed above the center of the benzene molecule with two hydrogen atoms of the water pointing symmetrically toward two carbon atoms [inset of Fig. 1(a)]. So as to obtain high quality binding energy curves, all these calculations employed Dunning's augmented correlation consistent basis sets (aug-cc-pVxZ, $x=T, Q$, and 5) (Ref. 27) within the frozen core approximation. Results with and without the counterpoise (CP) correction for basis set superposition error (BSSE) are reported.

The DFT calculations have been performed with the plane-wave (PW) basis set code CPMD.²⁸ Troullier–Martins norm-conserving pseudopotentials²⁹ (PPs) are employed with an energy cutoff of 100 Ry (1360 eV). A large 20 Å³ cell is used to ensure negligible interactions between periodic images. Two popular xc functionals PBE and BLYP (Refs. 30 and 31) and hybrid versions of these, PBE0 and B3LYP,^{32–35} have been considered. All-electron test calculations with GAUSSIAN03 with an aug-cc-pVQZ basis set reproduced the entire PBE and PBE0 binding energy curves to within 2 meV of the CPMD results, demonstrating the accuracy of our CPMD computational setup (basis set and pseudopotentials).³⁶

DMC calculations are performed with the CASINO

code,³⁷ employing trial wave functions (TWs) of the Slater–Jastrow type: $\Psi_T(\mathbf{R})=D^\dagger D^l e^J$, where D^\dagger and D^l are Slater determinants of up- and down-spin single-electron orbitals. The Jastrow factor, e^J , is the exponential of a sum of one-body (electron-nucleus), two-body (electron-electron), and three body (electron-electron-nucleus) terms, which are parametrized functions of electron-nucleus, electron-electron, and electron-electron-nucleus separations, and are designed to satisfy the cusp conditions. The parameters in the Jastrow factor are varied to minimize the variance of the local energy $E_L(\mathbf{R})\equiv\Psi_T^{-1}(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})$. Imaginary time evolution of the Schrödinger equation has been performed with the usual short time approximation, and the locality approximation.³⁸ We used a time step of 0.0125 a.u. Tests using a smaller time step of 0.005 a.u. showed no differences in binding energies within a statistical error of 5 meV. Dirac–Fock PPs for C, O, and H were used.³⁹ The C and O PPs have a frozen He core and core radius of 0.58 and 0.4 Å, respectively. The H PP has a core radius of 0.26 Å. The single particle orbitals have been obtained by DFT PW calculations with a PW cutoff of 300 Ry (4082 eV) using the PWSCF package,⁴⁰ and re-expanded in terms of B-splines^{41,42} using the natural B-spline grid spacing given by $a=\pi/G_{\max}$, where G_{\max} is the length of the largest PW. Most DMC calculations were performed using as TW the wave function built from the Kohn–Sham orbitals obtained from the LDA xc functional, but we also tested TWs obtained with other xc functionals (see below). The PW calculations from which the TWs were obtained were performed using a cell of at least 13 Å³, which is large enough to ensure that the tails of the single particle orbitals decay essentially to zero at the edges of the cell. DMC calculations were performed without periodic boundary conditions and therefore the Ewald technique was used to model electron-electron interactions. The number of walkers was

10240 and the simulations were carried out for 60 000 to 90 000 steps, resulting in statistical errors on each binding energy point between 3 and 4 meV (one standard deviation). In order to maximize DMC cancellation of errors, we have constructed our binding energy curve taking as a reference for the separated water-benzene system the DMC energy of water and benzene at a separation of 7 Å. A calculation at a separation of 9 Å showed no difference within a statistical error of 3 meV. For consistency, this 7 Å reference was also used for calculations with the other theories.

To get the best possible quantum chemistry reference data (that our computational resources allow), we follow the so-called Δ CCSD(T) approach.⁴³ This involves a determination of the energy difference between CCSD(T) and MP2 at a given finite basis set (aug-cc-pVTZ in this study). The difference is then added to the MP2 complete basis set (CBS) extrapolated energy, obtained in a separate series of calculations. Thus the Δ CCSD(T) energy is given by

$$E(\Delta\text{CCSD(T)}) = E(\text{CCSD(T)}/\text{aug-cc-pVTZ}) - E(\text{MP2}/\text{aug-cc-pVTZ}) + E(\text{MP2}/\text{CBS}), \quad (1)$$

with the meaning of each energy term being obvious. All energies in Eq. (1) are CP corrected for BSSE. The MP2/CBS energies were obtained with standard heuristic procedures for extrapolating the energy to the CBS limit, along with aug-cc-pVTZ, aug-cc-pVQZ, and aug-cc-pV5Z basis sets. Since the convergence behavior of the HF and correlation energies is different, we extrapolate them separately. For the extrapolation of the HF part, we use Feller's exponential fit:⁴⁴

$$E_x^{\text{HF}} = E_{\text{CBS}}^{\text{HF}} + Ae^{-Bx}, \quad (2)$$

where $x=3,4,5$ for aug-cc-pVTZ, aug-cc-pVQZ, and aug-cc-pV5Z basis sets, respectively. E_x^{HF} is the HF energy with aug-cc-pV x Z basis sets, and $E_{\text{CBS}}^{\text{HF}}$ is the extrapolated HF energy at the CBS limit. For the correlation part, we use the function⁴⁵⁻⁴⁷

$$E_x^{\text{corr}} = E_{\text{CBS}}^{\text{corr}} + Cx^{-3} + Dx^{-5}, \quad (3)$$

where E_x^{corr} is the MP2 correlation energy with aug-cc-pV x Z basis sets and $E_{\text{CBS}}^{\text{corr}}$ is the extrapolated MP2 correlation energy at the CBS limit. In the equations above, A , B , C , and D are fitting parameters. The MP2 energy curve at the CBS limit is displayed in Fig. 1(a). This is our best estimate of the binding energy curve at the MP2 level. It can be seen that the MP2/CBS curve falls between the raw and CP-corrected MP2 curves with the CP-corrected curves approaching the CBS limit more rapidly with increasing basis set size. This is in agreement with previous observations.^{1,16,48,49} Also the basis set incompleteness errors become smaller as the water-benzene separation increases since the correlation becomes smaller.

The CCSD(T) binding energy curves are displayed in Fig. 1(b). These have been obtained with an aug-cc-pVTZ basis set with and without the CP correction. The energy differences between CCSD(T) and MP2 with an aug-cc-pVTZ basis set with and without CP correction are almost 9

meV at the energy minimum. This difference increases as the separation decreases and decreases as the separation increases. Thus, correlations beyond the MP2 level are important to the binding in this system when the water-benzene separation is small, but less so when it is large. This is physically reasonable since the importance of electron correlation is expected to be greater when water is closer to the benzene molecule. The Δ CCSD(T) binding energy curve is also shown in Fig. 1(b). This is the best estimate of the binding curve obtained from the quantum chemistry methods. The binding energy at the minimum is ~ 125 meV and the equilibrium water-benzene separation is ~ 3.4 Å. Our computed binding energy is about 20–40 meV smaller than previous CCSD(T)/CBS-extrapolated values^{16,19,20} mainly because we have not optimized the structures of the water and benzene molecules. As we have said, we have not done so because our purpose here is not to determine the most accurate binding energy for this system, but rather to obtain an accurate reference with which to systematically investigate the performance of a range of electronic structure techniques absent of the “complications” brought about by slightly different structures.

Let us now consider how the other electronic structure theories tested perform compared to Δ CCSD(T)/CBS. To this end we plot in Fig. 2 binding energy curves for DMC, CCSD, MP2, HF, and DFT with several xc functionals. The bottom panel [Fig. 2(b)] reports the DFT results and the top one [Fig. 2(a)] all other methods. The first thing that is clear is that HF provides a very poor description of the binding energy curve, not even predicting a qualitatively correct curve. Given the lack of electron correlation in the HF description this poor performance is not surprising. Moving to CCSD, specifically Δ CCSD results obtained in a similar manner to Δ CCSD(T), we find that the CCSD binding is consistently smaller than CCSD(T). At the minimum of the binding curve the underbinding compared to CCSD(T) is 18 meV and at the shortest distance considered (2.8 Å) this is as large as 42 meV. This indicates that triple excitations are important to the binding in this system, particularly at short distance. Comparing CCSD and MP2 we find that MP2, which always overestimates the binding in this system, actually outperforms CCSD slightly. At the binding energy minimum MP2 differs from CCSD(T) by only 9 meV and at short separation (2.8 Å) this difference becomes 32 meV. The slightly superior performance of MP2 over CCSD, which obviously arises from a favorable error cancellation in MP2, has been seen before in hydrogen bonded and other systems.¹

Turning our attention to DMC, we find that both the depth (122 meV) and location (3.4 Å) of the binding energy minimum agree very well with CCSD(T). In the equilibrium region and at larger water-benzene separations the difference between CCSD(T) and DMC is never more than 3 meV, i.e., of the same size of the DMC statistical errors. This is rather encouraging and another strong indication that DMC, even at the simplest level [PPs, fixed node DMC with a single Slater–Jastrow trial wave function (TW)], can achieve high accuracy for noncovalent interactions. We note that the agreement obtained in the equilibrium region between DMC and CCSD(T)/CBS is slightly better than the ~ 18 meV dif-

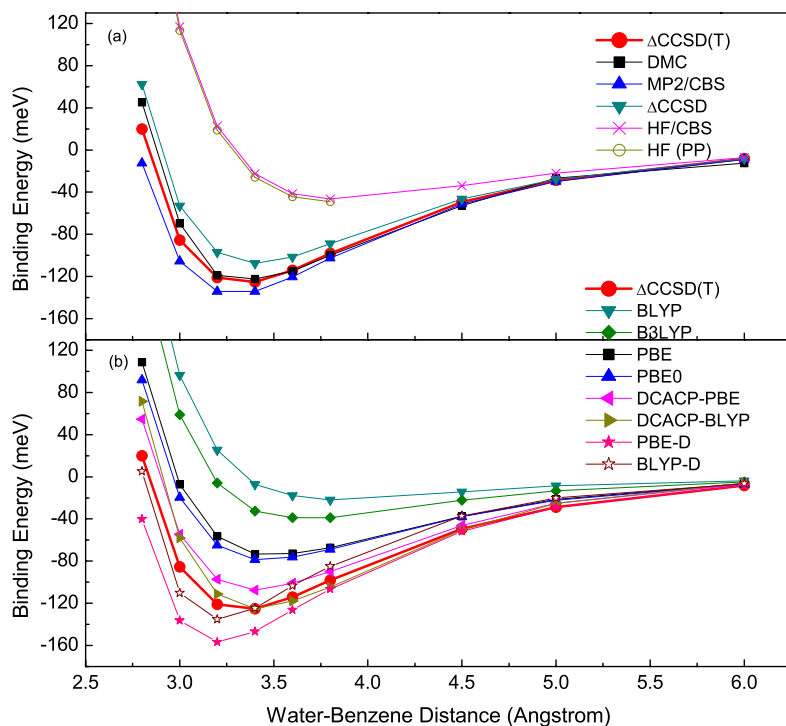


FIG. 2. (Color online) (a) Binding energy curves obtained with DMC, coupled cluster, MP2, and HF. The error bars of DMC are between 3 and 4 meV (one standard deviation), which are similar to the size of the square points and are not shown here. The results from both all-electron and pseudopotential HF (HF PP) calculations, with the same pseudopotentials that were used in the DMC simulations, are reported. (b) Binding energy curves obtained with various exchange-correlation functionals and vdW correction schemes. In (a) and (b) the Δ CCSD(T) binding energy curve [same as Fig. 1(b)] is also reported. The lines connecting the points are merely a guide to the eye.

ference observed recently in Ref. 10. However, as the separation is reduced below the equilibrium distance into the repulsive wall of the binding energy curve the difference between DMC and CCSD(T) grows to around 20 meV; with DMC predicting a steeper repulsive wall. The precise origin of this difference is not clear at this stage. First and foremost this is a small energy difference and it could readily arise from issues with either the coupled cluster or the DMC simulations reported here. With regard to the coupled cluster data, we do not, at present, have the computational resources that would allow us to utilize larger basis sets in the extrapolations or to go beyond the current level of treatment for electron correlation. As far as our DMC calculations are concerned, there are two obvious sources of error, namely, the PPs or the fixed node approximation. At the HF level we have checked the PPs used in the DMC simulations with a comparison to all-electron HF calculations. As reported in Fig. 2(a), for the entire binding energy curve the HF all electron and PP results are indistinguishable. Of course these tests do not tell us about the performance of the PPs in the correlated DMC calculations, and, in particular, about possible errors due to the locality approximation. However, they do show that at least at the HF level the PPs are reliable. A more likely source of the difference is the fixed node approximation. To investigate this issue to some extent, we have carried out a series of calculations using TWs built from the Kohn–Sham orbitals obtained with different DFT xc functionals (PBE, B3LYP, and PBE0) and HF. Calculations with these TWs were only performed with water-benzene separations of 2.8, 3.4, and 9.0 Å, representative of the short, equilibrium, and reference distances, respectively. The results are displayed in Table I. The DMC energies obtained with TWs obtained from HF orbitals are ~ 0.15 eV higher at all points, suggesting that, for this system, TWs built from HF orbitals are not the best choice. The energies

obtained from the TWs built from the Kohn–Sham orbitals obtained from the three DFT xc functionals are all within 20–30 meV of each other, with the TW built from LDA orbitals yielding the lowest total energies in this particular case. These different DMC energies do not directly provide a measure of nodal errors, however, they highlight differences in nodal errors which are of the same order as the energy differences between our best DMC and our CCSD(T) calculations. Therefore, it is reasonable to assume that nodal errors may be one source of the small difference between DMC and CCSD(T) at short water-benzene separations.

Let us now briefly consider how the various DFT xc functionals perform for this system. Only a very small selection of the essentially endless list of xc functionals now in use have been considered. Specifically, PBE, PBE0, BLYP, and B3LYP were examined. As can be seen from Fig. 2(b) all four functionals show poor performance, strongly underestimating the interaction. PBE0 predicts the largest en-

TABLE I. The energies of DMC calculations using TWs built from Kohn–Sham orbitals obtained from various DFT xc functionals (LDA, PBE, PBE0, and B3LYP) and HF. The energy with TWs built from LDA orbitals at 9 Å has been set as the energy zero reference. The DMC calculations with TWs built from LDA orbitals give the lowest total energy at all points. The DMC energies with TWs built from HF orbitals are about 120–160 meV higher than the results obtained from TWs built from LDA orbitals at all points.

TW	Energy (meV)		
	9.0 Å	3.4 Å	2.8 Å
LDA	0	-124(3)	44(3)
PBE	25(3)	-111(3)	62(3)
PBE0	7(3)	-121(3)	54(3)
B3LYP	18(3)	-108(3)	68(3)
HF	158(3)	4(3)	182(3)

ergy of 78 meV, PBE is marginally less (73 meV), and BLYP (22 meV) and B3LYP (39 meV) considerably less. This behavior is quite typical for how these functionals perform in weak interaction systems and is due to the lack of a proper account of vdW forces.⁵⁰ There are various strategies to “fix” this deficiency such as xc functionals which explicitly account for nonlocal correlation, using maximally localized Wannier functions, dispersion-corrected atom-centered potentials (DCACPs), or damped empirical C_6R^{-6} corrections (DFT-D),^{51–55} and indeed, some of these methods have already been applied to the water-benzene system.^{17,18} Having obtained a $\Delta\text{CCSD(T)}/\text{CBS}$ binding energy curve we take the opportunity here to assess the performance of two of the simplest vdW correction schemes, namely, DCACPs and DFT-D. Each method was tested for the PBE and BLYP xc functionals. As can be seen from Fig. 2(b), results in much better agreement within CCSD(T) are obtained. In particular, the combination of BLYP plus DCACPs yields a binding energy of around 125 meV.

In conclusion, a variety of electronic structure techniques have been applied to determine binding energy curves for a water-benzene complex. Specifically, we have compared the performance of HF, MP2, CCSD, CCSD(T), DMC, and several DFT xc functionals with and without empirical vdW corrections. The performance of the traditional quantum chemistry and DFT approaches was very much as expected and consistent with previous studies on this and related systems. On the other hand, a DMC binding curve had not been obtained for this system before and, indeed, DMC binding energy curves remain, in general, very scarce. It was found that DMC and the best quantum chemistry approach considered [i.e., CCSD(T)] agreed with each other very well at the binding energy minimum and beyond. At small water-benzene distances small differences (~ 20 meV) between the DMC and CCSD(T) energies, which yields to a steeper DMC repulsive wall compared to CCSD(T), have been obtained. The origin of the small difference between DMC and CCSD(T) has not been explained, however, explorations with a number of different TWs indicate that it may be related to the fixed node approximation used in DMC simulations.

J.M. and E.W. were supported by the NSFC. E.W. was also supported by an Alexander von Humboldt Research Award. A.M. and D.A. were very grateful to the EURYI scheme (see www.esf.org/euryi) and the EPSRC-GB for support. Computational resources from the London Centre for Nanotechnology are warmly acknowledged. We also acknowledge the UCL Research Computing for access to Legion, where the DMC simulations were run.

¹T. H. Dunning, Jr., *J. Phys. Chem. A* **104**, 9062 (2000).

²W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).

³T. D. Beaudet, M. Casula, J. Kim, S. Sorella, and R. M. Martin, *J. Chem. Phys.* **129**, 164711 (2008).

⁴S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).

⁵J. W. Lawson, C. W. Bauschlicher, Jr., J. Toulouse, C. Filippi, and C. J. Umrigar, *Chem. Phys. Lett.* **466**, 170 (2008).

⁶W. A. Al-Saidi and C. J. Umrigar, *J. Chem. Phys.* **128**, 154324 (2008).

⁷E. San Sebastian, J. M. Matxain, L. A. Eriksson, R. H. Stote, A. Dejaegere, F. P. Cossio, and X. Lopez, *J. Phys. Chem. B* **111**, 9099 (2007).

⁸J. M. Matxain, J. M. Mercero, A. Irigoras, and J. M. Ugalde, *Mol. Phys.* **102**, 2635 (2004).

⁹C. Diedrich, A. Lüchow, and S. Grimme, *J. Chem. Phys.* **123**, 184106 (2005).

¹⁰M. Korth, A. Lüchow, and S. Grimme, *J. Phys. Chem. A* **112**, 2104 (2008).

¹¹J. C. Grossman, *J. Chem. Phys.* **117**, 1434 (2002).

¹²I. G. Gurtubay and R. J. Needs, *J. Chem. Phys.* **127**, 124306 (2007).

¹³B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, *J. Chem. Phys.* **129**, 194111 (2008).

¹⁴N. D. Drummond and R. J. Needs, *Phys. Rev. Lett.* **99**, 166401 (2007).

¹⁵P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006).

¹⁶D. Feller, *J. Phys. Chem. A* **103**, 7558 (1999).

¹⁷U. Zimmerli, M. Parrinello, and P. Koumoutsakos, *J. Chem. Phys.* **120**, 2693 (2004).

¹⁸S. Li, V. R. Cooper, T. Thonhauser, A. Puzder, and D. C. Langreth, *J. Phys. Chem. A* **112**, 9031 (2008).

¹⁹Y. Zhao, O. Tishchenko, and D. G. Truhlar, *J. Phys. Chem. B* **109**, 19046 (2005).

²⁰S. Y. Min, E. C. Lee, D. Y. Kim, D. Kim, and K. S. Kim, *J. Comput. Chem.* **29**, 1208 (2008).

²¹J. K. Gregory and D. C. Clary, *Mol. Phys.* **88**, 33 (1996).

²²B. K. Mishra and N. Sathyamurthy, *J. Phys. Chem. A* **111**, 2139 (2007).

²³M. Nait Achour, M. R. Belmecheri, G. Berthier, and R. Savinelli, *Int. J. Quantum Chem.* **108**, 423 (2008).

²⁴M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., 2004.

²⁵The choice to use PBE structures is a somewhat arbitrary one. However, since full geometry optimizations at the CBS limit are not yet feasible for all the theories considered here, some arbitrary decision of what structures to use is required. See the supporting information for more details (Ref. 56).

²⁶J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).

²⁷R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).

²⁸CPMD V3.13, Copyright IBM Corp. 1990–2008, Copyright MPI für Festkörperforschung Stuttgart, Germany, 1997–2001, <http://www.cpmd.org/>.

²⁹N. Troullier and J. L. Martins, *Phys. Rev. B* **43**, 1993 (1991).

³⁰A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

³¹W. Lee, C. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).

³²C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).

³³A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).

³⁴S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).

³⁵P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).

³⁶DFT with local or semilocal xc functionals such as those used here is well known to be much less sensitive to basis set incompleteness errors than the explicitly correlated quantum chemistry methods considered here. When sufficiently large plane-wave and Gaussian basis sets are used, binding energies can be reproduced with meV precision as found here and elsewhere (see, e.g., Ref. 13).

³⁷R. J. Needs, M. D. Towler, N. D. Drummond, and P. Lopez Rios, *CASINO version 2.1 User Manual* (University of Cambridge, Cambridge, 2007).

³⁸L. Mitáš, E. L. Shirley, and D. M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991).

³⁹J. R. Trail and R. J. Needs, *J. Chem. Phys.* **122**, 014112 (2005); **122**, 174109 (2005); see also www.tcm.phys.ac.uk/~mdt26/casino2_pseudopotentials.html.

⁴⁰S. Baroni, A. Dal Corso, S. de Gironcoli, and P. Giannozzi, <http://www.pwscf.org>.

⁴¹E. Hernández, M. J. Gillan, and C. M. Goringe, *Phys. Rev. B* **55**, 13485 (1997).

⁴²D. Alfè and M. J. Gillan, *Phys. Rev. B* **70**, 161101 (2004).

⁴³P. Jurečka and P. Hobza, *Chem. Phys. Lett.* **365**, 89 (2002).

⁴⁴D. Feller, *J. Chem. Phys.* **98**, 7059 (1993).

⁴⁵C. Schwartz, *Phys. Rev.* **126**, 1015 (1962).

⁴⁶W. Kutzelnigg and J. D. Morgan III, *J. Chem. Phys.* **96**, 4484 (1992).

⁴⁷A. K. Wilson and T. H. Dunning, Jr., *J. Chem. Phys.* **106**, 8718 (1997).

⁴⁸D. Feller and K. D. Jordan, *J. Phys. Chem. A* **104**, 9971 (2000).

- ⁴⁹ B. Santra, A. Michaelides, and M. Scheffler, *J. Chem. Phys.* **127**, 184104 (2007).
- ⁵⁰ Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A* **110**, 5121 (2006).
- ⁵¹ M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, *Phys. Rev. Lett.* **92**, 246401 (2004).
- ⁵² P. L. Silvestrelli, *Phys. Rev. Lett.* **100**, 053002 (2008).
- ⁵³ O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, *Phys. Rev. Lett.* **93**, 153004 (2004).
- ⁵⁴ I.-C. Lin, M. D. Coutinho-Neto, C. Felsenheimer, O. A. von Lilienfeld, I. Tavernelli, and U. Rothlisberger, *Phys. Rev. B* **75**, 205131 (2007).
- ⁵⁵ S. Grimme, *J. Comput. Chem.* **27**, 1787 (2006).
- ⁵⁶ See EPAPS Document No. E-JCPA6-130-047914 for the coordinates of water and benzene molecules (optimized with the G03 code with the PBE xc functional at an aug-cc-pVTZ basis set). For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.