On the accuracy of density-functional theory exchange-correlation functionals for H bonds in small water clusters. II. The water hexamer and van der Waals interactions

Biswajit Santra,[1] Angelos Michaelides,[1,2,a)] Martin Fuchs,[1] Alexandre Tkatchenko,[1] Claudia Filippi,[3] and Matthias Scheffler[1]
[1]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*
[2]*Materials Simulation Laboratory, London Centre for Nanotechnology and Department of Chemistry, University College London, London WC1E 6BT, United Kingdom*
[3]*Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, NL-2300 RA Leiden, The Netherlands and Faculty of Science and Technology and MESA+Research Institute, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*

(Received 13 August 2008; accepted 13 October 2008; published online 20 November 2008)

Second order Møller–Plesset perturbation theory at the complete basis set limit and diffusion quantum Monte Carlo are used to examine several low energy isomers of the water hexamer. Both approaches predict the so-called prism to be the lowest energy isomer, followed by cage, book, and cyclic isomers. The energies of the four isomers are very similar, all being within 10–15 meV/H_2O. These reference data are then used to evaluate the performance of several density-functional theory exchange-correlation (xc) functionals. A subset of the xc functionals tested for smaller water clusters [I. Santra *et al.*, J. Chem. Phys. **127**, 184104 (2007)] has been considered. While certain functionals do a reasonable job at predicting the absolute dissociation energies of the various isomers (coming within 10–20 meV/H_2O), *none* predict the correct energetic ordering of the four isomers nor does any predict the correct low total energy isomer. All xc functionals tested either predict the book or cyclic isomers to have the largest dissociation energies. A many-body decomposition of the total interaction energies within the hexamers leads to the conclusion that the failure lies in the poor description of van der Waals (dispersion) forces in the xc functionals considered. It is shown that the addition of an empirical pairwise (attractive) C_6R^{-6} correction to certain functionals allows for an improved energetic ordering of the hexamers. The relevance of these results to density-functional simulations of liquid water is also briefly discussed. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3012573]

I. INTRODUCTION

How good is density-functional theory (DFT) for hydrogen (H) bonds? What is the best exchange-correlation (xc) functional for treating H bonds? Questions like these are far from uncommon for developers and practitioners of Kohn–Sham DFT, particularly those interested in simulating collections of atoms held together with H bonds. Clearly imprecise and vague questions, it is nonetheless important to answer them once, of course, terms such as "good" and "best" have been defined and consideration made to the properties of interest (energetic, structural, dynamical, electronic). Indeed considerable effort has been expended in an attempt to answer questions like these,[1–8] and with new xc functionals regularly appearing, there appears to be no end in sight for such studies.

One particularly important class of H bonded systems, arguably the most important, are the H bonds that hold water molecules together, either as gas phase molecular clusters or condensed phase solid (ice) and liquid water. Kohn–Sham DFT has been widely used to examine water under various conditions and environments.[1,5–7,9–24] Along with this widespread application there have also been various benchmark studies specifically aimed at accessing the performance of various xc functionals in treating gas phase water clusters,[1,5–7] adsorbed clusters,[19–26] and liquid water.[9–18] In particular, the question of the performance of DFT xc functionals in describing the structure and dynamics of liquid water has become a particularly hot and contentious issue due to apparent discrepancies between experiment and DFT.[9–18] Reconciling these differences, which are mainly concerned with the radial distribution functions (RDFs) and diffusion coefficient of liquid water, remains an immensely important open question and is one that is actively being addressed by many. However, simultaneously addressing all the possible factors which could account for the difference between the experimental and theoretical RDFs and diffusion coefficients (e.g., quantum nuclear effects, xc functional, density, and basis set) is far from straightforward and not particularly practicable. Instead the course we and others have chosen to follow to shed light on the performance of DFT xc functionals for treating water is to investigate well-defined gas phase water clusters for which precise comparison can be made to high level quantum chemistry calculations. This approach allows the precise performance

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

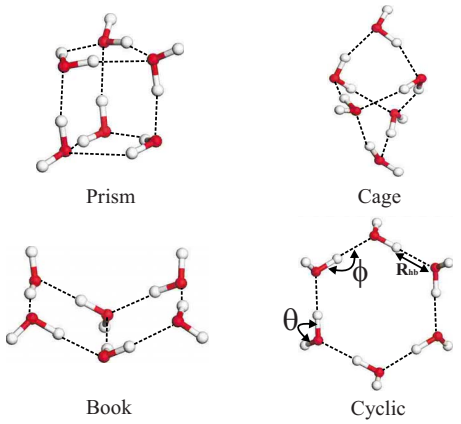

FIG. 1. (Color online) Structures of the four isomers of the water hexamer considered here (obtained with MP2 and an aug-cc-pVTZ basis). The dashed lines indicate H bonds, with the conventional number of H bonds each cluster is assumed to have (prism=9; cage=8; book=7; and cyclic =6) (Ref. 92). Some of the structural parameters discussed in the text are included alongside the cyclic structure.

limitations for a range of xc functionals to be obtained, information that is likely to be of relevance to liquid water.

Previously we tested the performance of 16 xc functionals for the equilibrium structures of the water dimer to pentamer, making reference to complete basis set (CBS) extrapolated second order Møller–Plesset perturbation theory (MP2) data.[1] That study revealed that of the functionals tested the hybrid X3LYP (Ref. 27) and PBE0 (Ref. 28) functionals were the most accurate, both coming within 10 meV/H bond of MP2 for each cluster. Among the nonhybrid functionals mPWLYP (Refs. 29 and 30) and PBE1W (Ref. 5) offered the best performance.[1] Here, we extend this work to the water hexamer. The water hexamer is interesting and warrants particular attention, not least because it provides a critical test for DFT xc functionals since there are four distinct isomers which lie within 10–20 meV/H_2O of each other. The isomers are known most commonly as the "prism," "cage," "book," and "cyclic" isomers (Fig. 1). Which one is the lowest energy on the Born–Oppenheimer potential energy surface with or without corrections for zero point vibrations or the experimental ground state structure at finite temperatures has been a matter of debate for some time.[6,31–40] For this paper we focus exclusively on the question of the lowest total energy isomer without zero point corrections, for which a consensus from wave function based methods appears to have emerged recently in favor of the prism isomer as being the lowest energy structure.[6,36,39,40] How many of the widely used xc functionals such as PBE, BLYP, and B3LYP perform for the relative energies of these isomers remains unclear, although there are indications that these and other DFT xc functionals are likely to encounter problems for the hexamer.[7,41–43] Other often cited reasons for being interested particularly in water hexamers are that they represent a transition from cyclic structures favored by smaller water clusters to three-dimensional structures favored by larger water clusters and that water hexamers are believed to be important constituents of liquid water and known to be building blocks of various phases of ice.[44]

In the following, we report a study in which the ability of several popular xc functionals to describe the energies and structures of the four water hexamers mentioned above is addressed. Comparisons are made with reference data generated by ourselves with MP2 at the CBS limit and diffusion Monte Carlo (DMC). The total energy ordering (i.e., neglecting zero point energies and finite temperature effects) predicted by MP2 and DMC is the same and in the order prism < cage < book < cyclic. However, all popular and widely used xc functionals tested fail to predict the correct ordering of the isomers; instead, they opt for either the book or cyclic isomers as the lowest energy ones. This discrepancy is largely attributed to the inability of DFT to correctly capture the van der Waals (vdW) interaction between widely separated molecules in the clusters. By including an empirical C_6R^{-6} correction we are able to explain the origin of the failure of the tested xc functionals and recover the correct energetic ordering between the different conformers.

II. METHODS AND REFERENCE DATA

This paper involves the application of a variety of theoretical approaches, which we now briefly describe. Specifically, we discuss how the MP2 and DMC reference data are acquired and then the set up for the DFT calculations.

A. MP2

MP2 has been used to compute structures and binding energies for each of the four isomers. All MP2 calculations have been performed with the GAUSSIAN 03 (Ref. 45) and NWCHEM (Ref. 46) codes and all geometries were optimized with an aug-cc-pVTZ basis set within the "frozen core" approximation, i.e., correlations of the oxygen 1*s* orbital were not considered.[47] Although the aug-cc-pVTZ basis set is moderately large (92 basis functions/H_2O), this finite basis set will introduce errors in the predicted MP2 structures. However, a test with the H_2O dimer reveals that the aug-cc-pVTZ and aug-cc-pVQZ MP2 structures differ by only 0.004 Å in the O–O bond length and 0.16° in the H bond angle (ϕ, Fig. 1). Likewise, Nielsen *et al.*[48] showed that the MP2 O–O distances in the cyclic trimer differ by 0.006 Å between the aug-cc-pVTZ and aug-cc-pVQZ basis sets with all other bonds differing by <0.003 Å. For our present purposes these basis set incompleteness errors on the structures are acceptable and it seems reasonable to assume that the MP2 aug-cc-pVTZ structures reported here come with error bars compared to the MP2/CBS limit of ±0.01 Å for bond lengths and ±0.5° for bond angles.

Total energies and dissociation energies are known to be more sensitive to basis set incompleteness effects than the geometries. To obtain reliable MP2 total energies and dissociation energies we employ the aug-cc-pVTZ, aug-cc-pVQZ (172 basis functions/H_2O), and aug-cc-pV5Z (287 basis functions/H_2O) basis sets in conjunction with the well-established methods for extrapolating to the CBS limit. Usually the extrapolation schemes rely on extrapolating separately the Hartree–Fock (HF) and correlation contributions to the MP2 total energy. For extrapolation of the HF part we use Feller's[49] exponential fit:

$$E_X^{\mathrm{HF}} = E_{\mathrm{CBS}}^{\mathrm{HF}} + Ae^{-BX}, \quad (1)$$

where X is the cardinal number corresponding to the basis set ($X=3$, 4, and 5 for the aug-cc-pVTZ, aug-cc-pVQZ, and aug-cc-pV5Z basis sets, respectively). E_X^{HF} is the corresponding HF energy, $E_{\mathrm{CBS}}^{\mathrm{HF}}$ is the extrapolated HF energy at the CBS limit, and A and B are fitting parameters. For the correlation part of the MP2 total energy we follow an inverse power of highest angular momentum equation:[50–52]

$$E_X^{\mathrm{corr}} = E_{\mathrm{CBS}}^{\mathrm{corr}} + CX^{-3} + DX^{-5}, \quad (2)$$

where E_X^{corr} is the correlation energy corresponding to X, $E_{\mathrm{CBS}}^{\mathrm{corr}}$ is the extrapolated CBS correlation energy, and C and D are fitting parameters.[53]

B. Quantum Monte Carlo

In order to assess the importance of correlation effects beyond the MP2 level, we evaluated the binding energies of the water clusters using quantum Monte Carlo (QMC). QMC is a stochastic approach to solve the many-electron Schrödinger equation.[54] The central quantity which determines the accuracy of a QMC calculation is the trial wave function, i.e., a correlated ansatz for the many-electron wave function. In variational Monte Carlo the expectation value of the many-electron Hamiltonian is computed as a statistical average over a large number of electronic configurations which are sampled from the square of the trial wave function using the Metropolis algorithm. An optimized trial wave function may be obtained within variational Monte Carlo based on variational principles for the variance of the local energy or the energy.[55] This trial wave function is then used in DMC which yields the best energy within the fixed-node approximation (i.e., projecting out the lowest energy state with the same nodes as the trial wave function).

DMC calculations yield highly accurate results for a wide variety of chemical systems (molecules and solids) and properties (binding energies, reaction energetics) as shown, for example, in Refs. 54 and 56–58. Recent studies of H bonded (and stacked aromatic) molecular dimers[59–63] demonstrate that DMC describes the interaction energies of such noncovalently bonded systems in very close agreement with the best available CCSD(T)/CBS estimates, in particular, where corrections beyond MP2/CBS are significant.[64]

In the present work, the trial wave functions are chosen of the Slater–Jastrow form with $\Psi = D_\uparrow D_\downarrow e^J$, i.e., as the product of Slater determinants D_σ of one-particle orbitals for the spin-up and spin-down electrons and a Jastrow correlation factor e^J depending on the electron-electron and electron-nucleus distances.[65] The one-electron orbitals are represented in an atomic Gaussian basis and generated from DFT-B3LYP calculations using the GAMESS code.[66] The parameters of the Jastrow correlation factor are optimized using the variance minimization method.[67] The atomic cores are represented using the nonlocal pseudopotentials of Ref. 68 and included in DMC within the usual localization approximation (i.e., nonlocal potentials are transformed into local operators by projection onto the trial wave function). The QMC calculations are performed using the CHAMP package.[69,70] All QMC results reported below are from DMC. The results for the different isomers are given in Table I and calculated for MP2/aug-cc-VTZ geometries (see Sec. II A) using a time step of $\tau = \frac{1}{80}E_h^{-1}$ and a target population of 800 walkers.

The main approximations in our DMC calculations are the fixed-node and pseudopotential localization approximations, and the quality of both is determined by the choice of the trial wave function. To estimate their effect on the H bond energies we have carefully analyzed the results of DMC calculations for the water dimer (in terms of the choice of pseudopotentials, basis set, terms in the Jastrow factor, and the time step in DMC) using the same form of the trial wave function as for the water hexamers. For the dissociation energy of the water dimer we obtain $D_e = 218 \pm 3$ meV, in agreement with the recent result of Gurturbay and Needs,[59] 218 ± 3 meV, obtained by a DMC calculation that employed a different set of pseudopotentials but appears otherwise essentially analogous to ours. These pseudopotential DMC results are consistent with the CCSD(T)/CBS result of Ref. 71 (217.7 meV) and the all-electron DMC result of Gurturbay and Needs[59] (224 ± 4 meV, for a Slater–Jastrow wave function using DFT-B3LYP orbitals). Gurturbay and Needs[59] furthermore showed that going beyond the localization approximation for nonlocal pseudopotentials produces equivalent results for the water dimer dissociation energy to within 5 meV. As the nodes of the trial wave function are given by its determinantal part, i.e., by D_σ, we also use orbitals from HF instead of DFT-B3LYP calculations to build the Slater determinants and thus provide a test of the sensitivity of the DMC results to changes in the nodes. While HF orbitals noticeably increase the total energies of the monomer and dimer compared to DFT-B3LYP orbitals, we find that these changes cancel in the DMC dissociation energy. Our DMC-HF result is 214 ± 6 meV, compared to 218 ± 8 meV for all-electron DMC-HF.[72] On the other hand, Gurturbay and Needs[59] showed that the inclusion of so-called backflow correlations to alter the nodes in a Slater–Jastrow wave function produces a slightly stronger H bond, changes being <20 meV in their pseudopotential DMC and somewhat smaller in their all-electron DMC calculation. From the above comparison of our DMC results for the water dimer with the best available theoretical reference data, DMC and CCSD(T), errors in the H bond strength due to the fixed-node and pseudopotential (localization) approximations appear small. We therefore expect that our DMC calculations provide an accurate account of the interactions between water molecules also in the hexamers, i.e., within ≈ 10 meV/H_2O. To further corroborate this estimate requires additional investigation, in particular, of the accuracy of the available, different pseudopotentials as well as of refinement of the trial wave functions. This is beyond the scope of the present study, but we note that previous DMC-HF studies using different pseudopotentials (and slightly different geometries) than employed here found somewhat larger dissociation energies of the water dimer [245 ± 9 meV (Ref. 73) and 232 ± 4 meV (Ref. 60)] than in the present work and in Ref. 59. The pseudopotentials used in these and the present study are both based on atomic HF calculations, yet their functional form is different. The accu-

TABLE I. Dissociation energies of the four water hexamers obtained from various electronic structure approaches: MP2/CBS; DMC; CCSD(T) with a triple-ζ basis set (Ref. 40); 12 different DFT xc functionals computed, unless indicated otherwise, with an aug-cc-pV5Z basis set; and HF at the CBS limit. The most stable isomer from each method is indicated in bold and the relative energies of the other isomers are given in parenthesis. MEs and MAEs in dissociation energies, averaged over the four hexamers in comparison with MP2 and DMC, are also given. All structures were optimized consistently with MP2, HF, and each DFT functional with an aug-cc-pVTZ basis set except for the DMC calculations which used the MP2 structures. DFT xc functionals are arranged here with increasing value of MAE from MP2. All values are in meV/H_2O (1 kcal/mol=43.3641 meV).

					MP2		DMC	
Method	Prism	Cage	Book	Cyclic	MAE	ME	MAE	ME
MP2	**332.3**	331.9(0.4)	330.2(2.1)	324.1(8.2)	···	···	···	···
DMC[a]	**331.9**	329.5(2.4)	327.8(4.1)	320.8(11.1)	···	···	···	···
CCSD(T)[b]	**347.6**	345.5(2.1)	338.9(8.7)	332.5(15.1)	···	···	···	···
PBE0	322.9(8.0)	325.3(5.7)	**330.9**	330.8(0.1)	5.9	−2.1	6.6	0.0
mPWLYP	323.2(10.4)	325.9(7.7)	**333.6**	333.3(0.3)	6.9	−0.6	7.7	1.5
X3LYP	317.2(8.8)	319.2(6.8)	325.8(0.2)	**326.0**	8.5	−7.6	7.1	−5.5
PBE1W	315.2(6.9)	314.8(7.3)	**322.1**	321.5(0.6)	11.3	−11.3	9.5	−9.1
PBE	336.1(9.5)	339.4(6.2)	**345.6**	344.1(1.5)	11.7	11.7	13.8	13.8
B98	305.3(7.3)	306.8(5.8)	**312.6**	312.5(0.1)	20.4	−20.4	18.2	−18.2
TPSS	303.9(12.8)	302.8(13.9)	313.6(3.1)	**316.7**	20.4	−20.4	18.3	−18.3
PW91	351.4(10.2)	354.7(6.9)	**361.6**	360.3(1.3)	27.3	27.3	29.5	29.5
BP86	294.9(13.6)	297.4(11.1)	**308.5**	306.6(1.9)	27.8	−27.8	25.7	−25.7
B3LYP	294.4(12.3)	297.1(9.6)	305.1(1.6)	**306.7**	28.8	−28.8	26.7	−26.7
XLYP	287.9(10.0)	286.9(11.0)	296.3(1.6)	**297.9**	37.4	−37.4	35.3	−35.3
BLYP	273.6(16.2)	277.4(12.4)	287.5(2.3)	**289.8**	47.6	−47.6	45.4	−45.4
BLYP[c]	273.6(16.2)	277.3(12.5)	287.4(2.4)	**289.8**	47.6	−47.6	45.5	−45.5
BLYP[d]	272.1(17.6)	276.0(13.7)	286.7(3.0)	**289.7**	48.5	−48.5	46.4	−46.4
HF	222.9(12.2)	224.4(10.7)	230.6(4.5)	**235.1**	101.4	−101.4	99.3	−99.3

[a]The statistical errors on the dissociation energies of prism, cage, book, and cyclic are ±1.0, ±0.9, ±1.0, and ±1.0 meV/H_2O, respectively. For the relative energies of the cage, book, and cyclic with respect to the prism (calculated as differences of total energies of the isomers rather than their dissociation energies) the statistical errors are ±1.0, ±1.1, and ±1.1 meV/H_2O, respectively.
[b]Reference 40.
[c]Full geometry optimization with the FHI-AIMS code.
[d]Full geometry optimization with the CPMD code using a plane-wave basis set and a 250 Ry cutoff. Also see Ref. 87 for results with a lower (70 Ry) cutoff.

racy of the pseudopotentials used here has been explicitly demonstrated for molecular properties of diatomic molecules at the CCSD(T) level.[68]

C. Kohn–Sham DFT

Of course, in a study such as this, there is an essentially endless list of functionals that we could consider evaluating. We have performed DFT calculations with 12 different xc functionals, chosen because they are either popular or have previously been shown to perform well for the strengths of H bonds between water molecules. Specifically, we have examined the following generalized gradient approximation (GGA) functionals: PW91,[74] PBE,[75] PBE1W,[5] mPWLYP,[29,30] BP86,[76,77] BLYP,[30,76] and XLYP (Ref. 27). The meta-GGA TPSS (Ref. 78) has also been considered as well as the following hybrid functionals: PBE0,[28] X3LYP,[27] B3LYP,[30,79–81] and B98.[82] The local-density approximation (LDA) has not been considered since it is known to overestimate the dissociation energy of water clusters by >50%.[42]

Most DFT calculations have been performed with the GAUSSIAN 03 (Ref. 45) and NWCHEM (Ref. 46) codes. Such calculations are all electron and employ Gaussian-type orbital basis sets. Geometries were optimized with an aug-cc-pVTZ basis set and energies with an aug-cc-pV5Z basis set. We have shown before that such large basis sets are for DFT sufficiently large to reflect the true performance of each xc functional at a level of accuracy that is reasonably expected to approach the basis set limit to within about 0.5 meV/H bond or better.[1]

While standard quantum chemistry software packages, such as the ones mentioned above, can be conveniently used for the simulation of small water clusters, one of our longer term goals is the accurate simulation of condensed phases of water such as ice or liquid water. Therefore, we have also performed selected DFT calculations with codes suitable for condensed phase simulations, such as the plane-wave pseudopotential code CPMD (Ref. 83) and the all-electron numeric atom-centered orbital (NAO) code FHI-AIMS, which originates from our laboratory.[84] A by-product of such effort is the interesting comparison of three different methodologies for DFT calculations (Gaussians, plane waves, and NAOs) of the energetics of H bonded systems. For the pseudopotential plane-wave DFT calculations in CPMD we have used hard pseudopotentials of Goedecker *et al.*[85,86] along with an energy cutoff of at least 200 Ry for the plane-wave kinetic energy.[87] For each hexamer an appropriate cell

size was chosen to leave at least 10 Å of vacuum on each side of the cluster. This cell size was found to be converged by performing selected simulations with a larger vacuum of 15 Å and with or without the Hockney Poisson solver[88] for electrostatic decoupling between neighboring cells. In the case of the all-electron NAO calculations with the FHI-AIMS code we have employed hydrogenic basis functions and carefully benchmarked all calculation parameters (basis set, grid size, cutoff potential) to achieve extreme convergence equivalent to or better than an aug-cc-pV5Z Gaussian basis set. As we will discuss below, a comparison between these three methods on the energetics of small water clusters (dimer-pentamer) reveals that the differences in the binding energy are on the order of 0.1 meV between GAUSSIAN and FHI-AIMS and of 1 meV between GAUSSIAN and CPMD, a value which is negligible for all our conclusions. Further, we have implemented a C_6R^{-6} empirical correction for vdW interactions both in CPMD and FHI-AIMS, so that we could perform full geometry optimizations with and without this correction.

D. Dissociation energy

For MP2, DFT, and DMC we have calculated dissociation energies per H_2O (D_e^n) which are given by

$$D_e^n = (E^{nH_2O} - nE^{H_2O})/n_{H_2O}, \quad (3)$$

where E^{nH_2O} is the total energy of each cluster with n H_2O molecules, E^{H_2O} is the total energy of a H_2O monomer, and n_{H_2O} is the number of water molecules in the cluster.

III. RESULTS

Now we present and discuss our MP2 and DMC reference data. Following this we evaluate the accuracy of the 12 xc functionals considered and then present a many-body decomposition of the total dissociation energies as well as a detailed discussion of the value of accounting for vdW dispersion forces in these clusters.

A. Reference dissociation energies

Following the procedure outlined above, we obtain MP2 dissociation energies at the CBS limit for the prism, cage, book, and cyclic hexamers of 332.3, 331.9, 330.2, and 324.1 meV/H_2O, respectively (see Table I).[89] Thus with MP2 the prism is the most stable structure and the energetic ordering of the isomers is prism < cage < book < cyclic. We note that this is consistent with the previous MP2/CBS study of the water hexamer reported by Xantheas *et al.*[39,90] The DMC calculations also find the prism to be the most stable isomer and predict the same energetic ordering as MP2. Clearly, the cyclic is the least stable isomer while the prism and the cage isomers appear energetically very close as they only differ by about two standard errors. Moreover, the absolute dissociation energies obtained with DMC and MP2 are within 4 meV/H_2O of each other for all four clusters (Table I). The sequence prism < cage < book < cyclic is also consistent with recent CCSD(T) calculations,[6,40] although the absolute binding energies from CCSD(T) when reported[40] are some 10–15 meV/H_2O larger than our MP2/CBS dissociation energies. Most of this difference can, however, be attributed to the smaller (aug-cc-pVTZ) basis set used in the CCSD(T) study.[91] Therefore, it is clear that all the explicitly correlated wave function based methods [MP2, DMC, CCSD(T)] predict the same low energy structure—prism—and the same energetic ordering— prism < cage < book < cyclic. With this consensus from different methods it now seems that the question of which isomer is the lowest energy on the Born–Oppenheimer potential energy surface (in the absence of contributions from zero point vibrations) is resolved in favor of the prism, and that suggestions to the contrary are not correct.[7] There remain, of course, minor differences in the relative energetic ordering of some structures on the order of 5 meV/H_2O [notably CCSD(T) predicts particularly unstable book and cyclic structures compared to MP2, with DMC being in between]. Resolving such small remaining differences is beyond the scope of the current paper, which instead now focuses on how the various DFT functionals do in describing the energies and structures of these clusters.

B. DFT dissociation energies

We turn now to the results obtained with the various DFT xc functionals and first consider (i) if the DFT xc functionals tested are able to predict the correct energetic ordering of the four hexamer isomers and (ii) what are the absolute errors in the total dissociation energies for each of the isomers. The answer to the first question is simple. All popular and widely used functionals tested fail to predict the correct minimum energy isomer. Instead of identifying the prism as the minimum energy conformer, all xc functionals tested either opt for the cyclic or book conformers (Table I). This includes the X3LYP and PBE0 functionals, which, in our previous study,[1] were identified as the most accurate xc functionals of those tested on the global minimum structures of small water clusters. It is somewhat discouraging that most of the xc functionals tested, despite being immensely popular for liquid water simulations, fail to predict the correct low energy structure for a system as seemingly simple as six water molecules. However, the failure is not entirely unexpected given that according to the wave function methods all four structures are so close in energy (within 10–15 meV/H_2O).

With regard to the second issue of how well the functionals perform at predicting the absolute binding energies of the clusters, the best functionals are PBE0, mPWLYP, and X3LYP, producing mean absolute errors (MAEs) averaged over the four clusters of 6, 7, and 9 meV/H_2O. PBE and PW91 produce errors of 12 and 28 meV/H_2O, respectively. B98 and TPSS both have a MAE of 20 meV/H_2O. B3LYP and BLYP underbind by ~29 and ~48 meV/H_2O, respectively. All of these conclusions are largely consistent with our previous study on smaller water clusters.[1] We note that the MAEs discussed are those obtained with respect to the MP2/CBS reference data. If instead we use the DMC results as the reference, the conclusions all remain essentially the

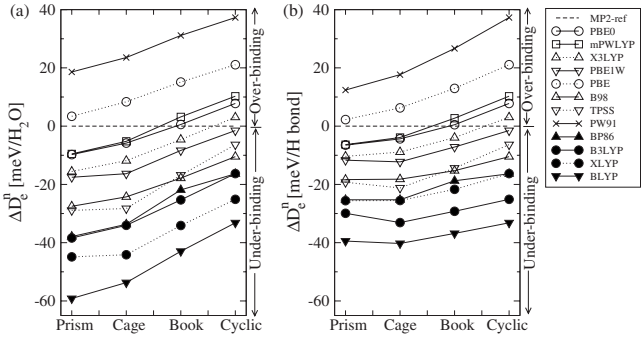

FIG. 2. Difference in the dissociation energy (ΔD_e^n) in (a) meV/H_2O and (b) meV/H bond between the various DFT xc functionals and MP2. In (b) the generally accepted number of H bonds in the prism, cage, book, and cyclic isomers of nine, eight, seven, and six, respectively, have been used (Ref. 92). Positive values correspond to an overestimation of the dissociation energy by a given DFT xc functional. We note that the reference MP2 dissociation energies are at the CBS limit, whereas for the DFT xc functionals an aug-cc-pV5Z basis set has been employed. Lines are drawn to guide the eye only.

same. This can be seen from Table I, and is, of course, due to the fact that the DMC and MP2 reference data are so similar (always within 4 meV/H_2O).

Looking more closely at how the functionals perform for specific clusters, we have plotted in Figs. 2(a) and 2(b) the difference between each functional and MP2/CBS (ΔD_e^n) for all four isomers. Since each cluster nominally has a different number of H bonds[92] and we are interested also in the description of H bonds, in Fig. 2(b) we also plot the error per H bond for each of the clusters. Figure 2 proves to be very illuminating and from it we extract the following key conclusions: (i) Upon moving from the prism to the cyclic isomer (as plotted in Fig. 2), all xc functionals display a trend toward increased binding; (ii) most functionals underbind the prism, with PBE and PW91 being the only exceptions; (iii) as we saw before for the dimer to pentamer,[1] here also BLYP performs consistently when we consider the error per H bond, coming around ~35 meV/H bond off MP2. Likewise XLYP yields very similar errors for all four isomers when considered on a per H bond basis. We will draw upon these conclusions later.

Another interesting finding is that the calculations on different water hexamers agree within 0.1 meV/H_2O between the all-electron GAUSSIAN 03 and FHI-AIMS codes and within 1.5 meV/H_2O between GAUSSIAN 03 and the pseudopotential plane-wave CPMD code (Table I). The latter value is most probably due to the difference in treatment of core electrons; however, this difference is still very small for all practical purposes. This level of agreement is also achieved for the smaller clusters—dimer to pentamer—in their equilibrium geometries.[93] This again reinforces that the basis sets employed here are sufficiently large to reflect the true performance of a given xc functional, absent of basis set incompleteness errors.

C. Geometry

Let us now consider the quality of the geometrical predictions made by the various xc functionals. The five key structural parameters of the H_2O clusters (some of them are shown in Fig. 1) that we evaluate are (i) the distance between adjacent oxygen atoms involved in a H bond, $R_{O–O}$; (ii) the length of a H bond, given by the distance between the donor H and the acceptor O, $R_{O\cdots H}=R_{hb}$ (Fig. 1); (iii) the H bond angle, $\angle(O\cdots H–O)=\phi$ (Fig. 1); (iv) the internal O–H bond lengths of each water, $R_{O–H}$; and (v) the internal H–O–H angle of each water, $\angle(H–O–H)=\theta$ (Fig. 1).

In Table II, the MAE and mean error (ME) of each xc functional compared to MP2 and averaged over all four clusters are reported. This provides a broad overview of how each functional performs, revealing that for structural predictions X3LYP is the most accurate functional. X3LYP outper-

TABLE II. MAE of the various DFT functionals from MP2 for five different structural parameters, averaged over the four water hexamers examined here. The numbers in bold all have MAE ≤0.010 Å for bond lengths and ≤0.50° for bond angles. MEs are given in parenthesis. MP2 and DFT (and HF) structures were optimized consistently with MP2 and with each DFT functional (and HF) with an aug-cc-pVTZ basis set. The DFT +vdW structures were optimized with a numerical atom-centered basis set (FHI-AIMS code). The order of the DFT xc functionals is the same as in Table I.

	$\Delta R_{O–O}$ (Å)	ΔR_{hb} (Å)	ΔR_{O-H} (Å)	$\Delta\phi$ (deg)	$\Delta\theta$ (deg)
PBE0	0.023(−0.017)	0.028(−0.018)	**0.002**(0.000)	0.96(−0.01)	0.69(+0.69)
mPWLYP	0.021(+0.021)	0.019(+0.008)	0.013(+0.013)	0.95(−0.25)	**0.49**(+0.49)
X3LYP	**0.009**(+0.008)	0.012(+0.009)	**0.000**(0.000)	**0.48**(−0.29)	0.98(+0.98)
PBE1W	0.062(+0.045)	0.096(+0.051)	0.011(+0.011)	3.98(−0.64)	**0.33**(+0.31)
PBE	0.032(−0.019)	0.055(−0.036)	0.014(+0.014)	1.87(+0.12)	**0.24**(+0.18)
PBE+vdW	0.026(−0.022)	0.044(−0.039)	0.012(+0.012)	1.11(+0.26)	**0.21**(+0.03)
B98	0.025(+0.025)	0.028(+0.028)	**0.001**(−0.001)	1.07(−0.20)	0.66(+0.66)
TPSS	0.094(+0.040)	0.155(+0.058)	0.011(+0.011)	6.03(−0.87)	0.58(+0.53)
PW91	0.039(−0.034)	0.060(−0.051)	0.014(+0.014)	1.59(+0.15)	**0.36**(+0.33)
BP86	0.032(−0.026)	0.055(−0.046)	0.016(+0.016)	1.65(+0.27)	**0.28**(+0.16)
B3LYP	0.019(+0.019)	0.020(+0.020)	**0.000**(+0.000)	0.61(−0.28)	0.89(+0.89)
XLYP	0.092(+0.082)	0.113(+0.091)	0.011(+0.011)	3.73(−0.99)	0.52(+0.52)
BLYP	0.039(+0.039)	0.029(+0.028)	0.012(+0.012)	1.29(−0.20)	**0.39**(+0.39)
BLYP+vdW	0.030(−0.026)	0.052(−0.044)	0.013(+0.013)	1.94(+0.59)	0.63(+0.63)
HF	0.165(+0.165)	0.200(+0.200)	0.026(−0.026)	1.66(−1.39)	1.62(+1.62)

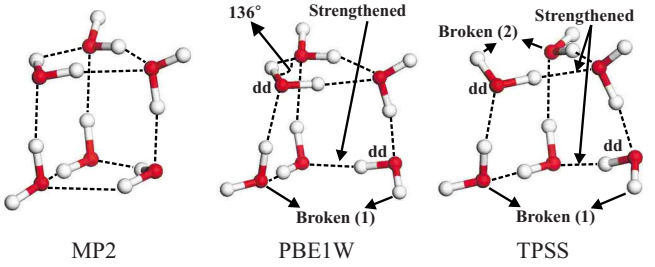

FIG. 3. (Color online) Structures of the prism isomer optimized with MP2 and the PBE1W and TPSS xc functionals. Dashed lines indicate H bonds. For PBE1W one H bond is broken and for TPSS two H bonds are broken, each broken H bond being associated with a double donor (dd) water molecule. The other H bonds which get stronger as a result of the bond breaking are also indicated. A very bent H bond angle of 136° is also shown in the upper triangle of the PBE1W structure.

forms all other functionals for almost all structural parameters considered with an average error of only 0.02 Å for the bond lengths and 0.5° for the bond angles. Considering the predicted O–O distances, on average, X3LYP, mPWLYP, PBE1W, TPSS, B98, B3LYP, BLYP, and XLYP predict slightly longer (0.008–0.082 Å) distances, whereas PBE0, PBE, BP86, and PW91 produce slightly shorter O–O distances (0.017–0.034 Å). This conclusion also holds for the related quantity R_{hb}. For the O–H bond length, R_{O-H}, on average all functionals perform reasonably well, coming within 0.02 Å. In particular, the results for X3LYP, PBE0, B98, and B3LYP are nearly identical to MP2. For the internal H–O–H angle θ, the MAE from all the functionals is within ~1.0°. Finally, for the H bond angle ϕ, X3LYP, B3LYP, PBE0, and mPWLYP perform the best, all coming within 1.0°. For this quantity, however, several functionals exhibit quite large discrepancies. Specifically, XLYP, PBE1W, and TPSS yield average MAEs of 3.7°, 3.9°, and 6.0°, respectively. As we go from cyclic to book to cage to prism, the H bond angles in the clusters become increasingly nonlinear (179° for cyclic, ~160°–170° for book, ~152°–166° for cage, and ~135°–168° for prism) and it appears that certain xc functionals struggle to reliably describe such nonlinear H bonds. Indeed closer inspection reveals that the largest errors in ϕ are encountered for the prism isomer. In this isomer there are two water molecules that are each involved in donating two H bonds (the molecules labeled dd for double donor in Fig. 3), and according to MP2 the H bonds these molecules donate are very bent (i.e., values of $\phi \sim 135°$). Several of the xc functionals fail to describe these very nonlinear essentially putative H bonds, and for one or both of the waters in the prism sacrifice a single very nonlinear H bond to enable the other to become more linear and hence stronger (Fig. 3). TPSS fails for both double donor water molecules and PBE1W and XLYP fail to describe one of them. The limitations of functionals such as those considered here in describing nonlinear putative H bonds in water clusters have also recently been pointed out by Shields and Kirschner.[94] There it was argued that vdW dispersion forces are critical to the binding of such weak H bond structures. We tend to agree with this conclusion and will show more evidence in support of it below.

D. Many-body decomposition of the dissociation energies

To identify precisely where the problem with the DFT xc functionals lies in correctly describing the energetic ordering of the various isomers, we have performed a many-body decomposition of the total dissociation energies of the hexamers. This has involved decomposing the total interaction energy within the clusters into one-body up to six-body contributions. Such many-body expansions have before proved useful in understanding the binding in H bonded clusters (including water clusters). A full description of the procedure involved can be found in Refs. 95–98. Very briefly, the total one-body energy is the energy cost incurred upon deforming all six monomers from the equilibrium isolated monomer structure to the structures they assume in a given hexamer. The total two-body interaction energy is the sum of all possible dimer interactions within the hexamer, i.e., the total energy (gain) to form all possible water dimers within a given hexamer from each of its (deformed) monomers. The total three-body interaction corresponds to the energy (gain) to form all possible trimer combinations (excluding dimer interactions inside the trimers) and so on for the four-, five-, and six-body interactions. We have performed such a many-body decomposition for the prism and cyclic conformers since the prism conformer is favored by the wave function approaches and the cyclic conformer is favored by many of the DFT xc functionals. The decomposition, the results of which are reported in Table III, has been performed with MP2 (with an aug-cc-pV5Z basis set) and with the X3LYP, PBE0, and BLYP xc functionals. To enable an exact comparison between MP2 and the various XC functionals, absent of any contributions arising from the slightly different structures obtained with the different approaches, we have used the MP2 geometries for all decompositions.

Let us first consider the MP2 reference data. For each cluster a small positive one-body energy of ~17 meV/H_2O is observed. The two-body interaction is attractive (negative) and at −244 and −283 meV/H_2O for the cyclic and prism isomers, respectively, comprises by far the largest contribution to the many-body expansion. The three-body interaction is also large and overall attractive: −84 and −64 meV/H_2O for the cyclic and prism structures, respectively. Indeed because of their magnitude the two- and three-body interactions almost decide what the total dissociation energies are. The four-, five-, and six-body terms are all considerably smaller. These results are consistent with those reported by Xantheas[96] with a smaller basis set.

Turning our attention now to how the DFT xc functionals perform, we first consider the two more accurate xc functionals for which the many-body decomposition has been performed (PBE0 and X3LYP). For the one-, four-, five-, and six-body contributions, we find reasonably good agreement with MP2. As we have said, these terms are small and the difference between MP2 and the two xc functionals is typically $\ll 8$ meV/H_2O. For the (larger) thee-body terms we observe variable performance with overbinding (8–9 meV) for the cyclic isomer and underbinding (3–5 meV) for the prism. It is for the two-body terms that we observe the

TABLE III. Many-body contributions to the total dissociation energies of the cyclic and prism isomers as obtained from MP2, X3LYP, PBE0, BLYP, and BLYP+vdW. For the MP2 many-body decomposition an aug-cc-pV5Z basis set is employed and so the total MP2 dissociation energies differ slightly from the MP2/CBS values given in Table I. Likewise, to avoid complications from the slightly different optimized structures obtained from MP2 and the DFT xc functionals, the DFT many-body decompositions are performed on the optimized MP2 structures (with an aug-cc-pV5Z basis set for the DFT energies). Values in the parenthesis are the difference between each functional and the MP2 results. Negative values indicate a gain in energy, i.e., a net attraction when all the n-body interactions of a given class are summed up, and positive values a net repulsion. All values are in meV/H_2O.

	MP2	X3LYP	PBE0	BLYP	BLYP+vdW
			Cyclic		
One body	+16.6	+12.9(−3.7)	+16.5(−0.1)	+2.4(−14.2)	+2.0(−14.6)
Two body	−244.2	−231.2(+13.0)	−240.8(+3.4)	−175.8(+68.4)	−227.8(+16.4)
Three body	−83.6	−92.1(−8.5)	−92.8(−9.2)	−97.7(−14.1)	−97.7(−14.1)
Four body	−16.0	−13.9(+2.1)	−8.1(+7.9)	−14.8(+1.2)	−14.8(+1.2)
Five body	+0.5	−1.7(−2.2)	−6.4(−6.9)	−1.9(−2.4)	−1.9(−2.4)
Six body	−0.9	+0.0(+0.9)	+1.2(+2.1)	+0.0(+0.9)	+0.0(+0.9)
Total	−327.6	−326.0(+1.6)	−330.4(−2.8)	−287.8(+39.8)	−340.2(−12.6)
			Prism		
One body	+16.7	+14.4(−2.3)	+17.3(+0.6)	+3.4(−13.3)	+3.2(−13.5)
Two body	−283.4	−263.6(+19.8)	−274.4(+9.0)	−191.8(+91.6)	−278.0(+5.4)
Three body	−63.8	−61.3(+2.5)	−59.3(+4.5)	−79.3(−15.5)	−79.3(−15.5)
Four body	−5.2	−7.6(−2.4)	−5.2 (0.0)	−2.8(+2.4)	−2.8(+2.4)
Five body	−2.6	+1.4(+4.0)	−3.7(−1.1)	+0.1(+2.7)	+0.1(+2.7)
Six body	+2.2	−0.1(−2.3)	+2.5(+0.3)	+0.1(−2.1)	+0.1(−2.1)
Total	−336.1	−316.8(+19.3)	−322.8(+13.3)	−270.3(+65.8)	−356.7(−20.6)

largest deviations from MP2 with a consistent underbinding for each functional and cluster. Both PBE0 and X3LYP underestimate the two-body contribution in the prism isomer by 9 and 20 meV/H_2O, respectively, and for the cyclic isomer PBE0 and X3LYP underestimate the two-body contribution by 4 and 13 meV/H_2O, respectively. It is interesting that these errors are noticeably larger than the 1–2 meV/H_2O errors obtained with these functionals for the equilibrium water dimer.[1] Thus we observe from the many-body analysis that *these xc functionals yield larger errors when describing the nonequilibrium dimer configurations present in the various water hexamers, compared to the equilibrium water dimer.* Upon inspection of the errors associated with the individual dimer configurations within the hexamers we find that there is a systematic underbinding for dimers at intermediate separations (O–O distances of ~3.0–5.0 Å) typical of vdW bonded complexes and also for certain orientations of water molecules held together with very nonlinear H bonds. There are not enough distinct dimer configurations within the hexamers to allow us to understand the precise dependence of the two-body error on orientation and H bond angle. However, the distance dependence of the underbinding is more clear and is something that we now address with a distance dependent vdW correction. Before moving on we note that the BLYP errors from the many-body analysis are consistently larger compared to PBE0 and X3LYP, consistent with the generally inferior performance of this functional. However, the main conclusion from the many-body analysis that the two-body terms are underbound (and are more poorly described than the equilibrium dimer) still holds.

E. DFT+vdW dissociation energy

Nowadays it is well known that most popular xc functionals generally show unsatisfactory performance for vdW forces, which inherently arise due to nonlocal correlations.[4,99,100] In order to test if the lack of vdW forces is indeed responsible for the underestimation in the two-body interactions, we use a simple C_6R^{-6} correction for the DFT total energies. The C_6R^{-6} correction method was early proposed for correcting HF calculations[101] and specifically applied to DFT by Wu and Yang,[102] Grimme,[103] and Jurečka *et al.*[104] Certainly the C_6R^{-6} pairwise scheme is a simple one for incorporating dispersion interactions into DFT calculations in contrast to other approaches (e.g., DFT xc functionals explicitly accounting for nonlocal correlation,[105] interaction of the instantaneous dipole moment of the exchange hole,[106] using maximally localized Wannier functions[107] or modified pseudopotentials[108]). However, accurate results have been obtained with the C_6R^{-6} correction and it has a well-established physical basis. With this approach the pairwise vdW interaction (E_{disp}) is calculated by

$$E_{\text{disp}} = -\sum_{j>i} f_{\text{damp}}(R_{ij},R_{ij}^0)C_{6ij}R_{ij}^{-6}, \tag{4}$$

where C_{6ij} are the dispersion coefficients for an atom pair ij (here taken from the work of Wu and Yang[102]), R_{ij} is the interatomic distance, R_{ij}^0 is the sum of equilibrium vdW distances for the pair (derived from atomic vdW radii[109]), and f_{damp} is a damping function. The damping function is needed to avoid the divergence of the R^{-6} term at short distances and reduces the effect of the correction on covalent bonds. We use a Fermi-type function f_{damp},

TABLE IV. Absolute values of vdW interaction energies and vdW corrected total dissociation energies for the four water hexamers for three different xc functionals. The DFT structures employed are fully relaxed geometries calculated with the FHI-AIMS code [the CPMD code gives very similar numbers (Ref. 120)]. For comparison the MP2/CBS results are also displayed. The energies of the most stable isomers are indicated in bold and the relative energies of the other structures with respect to the prism are given in parenthesis. MAEs in total dissociation energies are calculated from the MP2/CBS values averaging over the four hexamers. All numbers are in meV/H_2O.

	vdW interaction energy			
Method	Prism	Cage	Book	Cyclic
BLYP+vdW	93.8	90.5	75.8	60.7
PBE+vdW	40.9	40.5	31.6	22.9
PBE+vdW(TS)[a]	31.9	32.2	23.7	15.8
PBE0+vdW	35.2	35.4	27.4	19.4

	Total dissociation energy				
Method	Prism	Cage	Book	Cyclic	MAE
MP2	**332.3**	331.9 (0.4)	330.2 (2.1)	324.1 (8.2)	···
BLYP+vdW	**359.9**	359.7 (0.2)	356.3 (3.6)	344.8 (15.1)	25.5
PBE+vdW	377.8	**380.1** (−2.3)	377.8 (0.0)	367.3 (10.5)	46.1
PBE+vdW(TS)[a]	369.6	**372.6** (−3.0)	370.6 (−1.0)	360.7 (8.9)	38.8
PBE0+vdW	360.6	**361.9** (−1.3)	359.2 (1.4)	351.4 (9.2)	28.6

[a]Nonempirical vdW method of Tkatchenko and Scheffler (Ref. 113).

$$f_{\mathrm{damp}}(R_{ij},R_{ij}^0)=\left(1+\exp\left(-d\left(\frac{R_{ij}}{s_R R_{ij}^0}-1\right)\right)\right)^{-1}, \tag{5}$$

where d determines the steepness of the damping function (the higher the value of d, the closer it is to a step function) and s_R reflects the range of interaction covered by the chosen DFT xc functional.[104] The value of d was set to 20 and s_R is 0.80 for BLYP, 1.00 for PBE, and 1.03 for PBE0. These values of d and s_R were obtained by fitting to the intermolecular binding energies of the S22 database[104] at the CBS limit for all DFT xc functionals.[110,111]

The results for the PBE, PBE0, and BLYP functionals after applying the correction to the four hexamers are shown in Table IV. Also the total vdW interaction within each hexamer is reported. One can see that the vdW correction is largest for the prism and cage structures and noticeably less for book and cyclic structures, favoring the prism or cage over the cyclic or the book structure. The new energetic orderings of the hexamers are thus in contrast to all pure DFT functionals, which predict the book or cyclic structures to have the lowest energy (Table I), and in better agreement with the wave function based methods. The energy difference between the most stable and the least stable hexamers is also in reasonably good agreement with MP2 and DMC results (around 10–15 meV). Of the three functionals to which the correction has been applied, the BLYP+vdW method gives the best agreement with MP2. The MAE in the total dissociation energies for all four hexamers is reduced from 15% to 8%. Moreover, the correct energetic ordering of the four isomers is recovered, i.e., BLYP+vdW predicts the sequence prism<cage<book<cyclic. The results for BLYP+vdW are encouraging; however, it is important to note that there remains an 8% error (a significant overbinding). In addition, the "success" of BLYP+vdW is achieved at the expense of a smaller s_R parameter which shifts the vdW minima to quite short distances (see below). Also, the three-body contribution of BLYP, unaffected by the pairwise vdW correction, shows substantial error. Recently it was pointed out that many DFT functionals grossly overestimate many-body interactions in vdW systems.[112] However, in the present case the combination of electrostatic and vdW contributions does not allow one to clearly discern which part is responsible for the overbinding. Thus, further investigation is required to rule out fortuitous error cancellation for BLYP+vdW. Nonetheless these findings for at least three different functionals support the suggestion that the origin of the incorrect prediction of the energetic ordering of the water hexamers lies in the absence of vdW dispersion forces in the functionals considered.

At the final stages of the present work, we completed the development of a novel scheme to determine the vdW C_6 coefficients in Eq. (4) and the vdW radii R_0 in Eq. (5) in a nonempirical fashion. The scheme, which will be presented in a forthcoming publication,[113] involves three key elements: (i) Hirshfeld partitioning of the electron density to calculate the *relative* polarizability of an atom inside a molecule; (ii) the use of very accurate reference free-atom static dipole polarizabilities and C_6 coefficients, calculated with converged wave function based methods; and (iii) accurate combination rules to derive heteronuclear C_6 coefficients from static dipole polarizabilities and C_6 coefficients of homonuclear atoms. The scheme, which turns out to be very accurate (5.6% mean absolute relative error on a database of 148 experimental C_6 coefficients), is implemented in the FHI-AIMS code.[84] It allows geometry optimizations in which the C_6 coefficients and vdW radii of individual atoms can change throughout the simulation according to the DFT charge density. We find that this is particularly important for water clusters where H atoms participating in a H bond can

yield different values of C_6 and R_0 from those not involved in a H bond (and from the values used in empirical schemes). Preliminary results for the water hexamers are shown in Table IV for the PBE functional. As with the empirical correction scheme, the prism and cage structures are favored over book and cyclic structures and it is encouraging that the overestimation in binding energies with PBE+vdW is reduced by 7–8 meV/H_2O for all isomers. This reduction is due to the larger effective vdW radius of the atoms participating in the H bonds.

IV. DISCUSSION AND CONCLUSIONS

Having presented a lot of data obtained with various approaches, let us now recap the main results and discuss them in a somewhat broader context. To begin, there are the reference data themselves, which have been acquired with MP2 and DMC. From these we conclude that the prism is the lowest total energy isomer for six water molecules in the absence of contributions from zero point vibrations. This conclusion agrees with the general consensus that has emerged, being consistent with the very recent triple-ζ CCSD(T) results.[6,40] There remain, of course, minor differences in the relative energetic ordering of some structures on the order of 5 meV/H_2O [notably CCSD(T) predicts particularly unstable book and cyclic structures compared to MP2, with DMC being in between]. Resolving such small remaining differences will provide interesting work for the future. In this regard CCSD(T) calculations at the CBS limit would be welcome. We stress that the ordering arrived at here, prism < cage < book < cyclic, is the ordering obtained in the absence of corrections for zero point contributions. It is known that zero point energies will alter the relative energy spacings with indications that the cage becomes the most stable isomer.[34–36]

It is interesting to see that DMC and MP2 dissociation energies of the different isomers are so similar to each other, within 4 meV/H_2O. This may indicate that correlation effects beyond MP2 have little effect on the H bond energetics in these water clusters or it may indicate a favorable cancellation of errors in the MP2 and/or DMC calculations. Nonetheless, it demonstrates that DMC can achieve high accuracy in describing the energetics of H bonds between water molecules, already at the simplest DMC level, i.e., pseudopotential fixed-node DMC with a single-determinant Slater–Jastrow trial wave function, as has been found for a number of other H bonded systems (including DNA base pairs).[60,61] For the water hexamers studied here, the fixed-node and pseudopotential approximations in DMC incur no significant errors on the calculated H bond energies. We stress, however, that in general such errors depend on the system considered and still need to be carefully assessed by comparing to standard quantum chemistry approaches such as CCSD(T)/CBS and monitoring the quality of the trial wave function and, when used, also the pseudopotentials.

The main part of this paper was concerned with using the reference data from the wave function based methods to evaluate the performance of several DFT xc functionals. A subset of the xc functionals previously tested for small water clusters[1] was considered. It was found that while certain functionals did a reasonable job at predicting the absolute dissociation energies of the various isomers (coming within 10–20 meV/H_2O), *none* of the functionals tested predict the correct energetic ordering of the four isomers nor does any predict the correct lowest energy isomer. All xc functionals either predict the book or cyclic isomers to have the largest dissociation energies. There have been indications before that certain DFT xc functionals may not predict the correct lowest energy structure for the water hexamer. BLYP, for example, was long ago shown to favor the cyclic isomer.[42] Likewise X3LYP, B3LYP, and PBE1W have been shown to favor the cyclic structure.[6,7] Here, we have shown that several other popular xc functionals fail to predict the correct lowest energy structure too, yielding results for relative energies that are unreliable and misleading. Furthermore, by attributing the failure to an improper treatment of vdW forces it seems likely that many other semilocal and hybrid xc functionals which do not account for vdW in some way will also fail in this regard. We have shown that by augmenting the BLYP functional with an empirical pairwise C_6R^{-6} correction the correct energetic ordering of the four hexamers is recovered. Equivalent empirical corrections to other functionals (PBE, PBE0) also improves the ordering somewhat, favoring the prism and cage isomers over the book and cyclic ones. Of course there are other means of incorporating vdW dispersion forces implicitly into DFT xc functionals such as the approaches pioneered in Refs. 105–108. It will be interesting to see if these functionals can predict the correct lowest energy structure for the water hexamer and, at the same time, yield accurate total dissociation energies. Indeed on the general point of benchmarking and assessing the performance of existing and new xc functionals for the treatment of H bonded systems, it seems that the water hexamer would be an appropriate test case to add to existing H bond test sets since it presents a stern challenge for any xc functional. We reiterate that we are not suggesting that all xc functionals which do not account for vdW forces in one way or another are likely to fail to predict the correct energy ordering for the water hexamer. Indeed Dahlke *et al.*[6] very recently reported that a few empirical hybrid meta-GGA functionals achieve the correct energetic ordering for the hexamers,[6] and, in agreement with Ref. 6 our calculations with the M05-2X (Ref. 114) functional with an aug-cc-pVQZ basis set also find the prism to be the lowest energy structure.[115] This looks like an exciting development but what the precise reason for the success of the functionals tested is remains unclear to us at present.

Having identified a lack of vdW dispersion forces as being at the heart of the incorrect energy ordering of the various water hexamers, we now consider why the C_6R^{-6} correction scheme applied here works to alter the relative energies of the four isomers. Since the empirical BLYP+vdW scheme recovers the correct energetic ordering for the four hexamers we focus on analyzing the details of this correction. First we consider the functional form of the specific empirical dispersion corrections applied in these systems. These are displayed in Fig. 4(a) for the three individual types of atom-atom interaction: O–O, O–H, and H–H. Dispersion

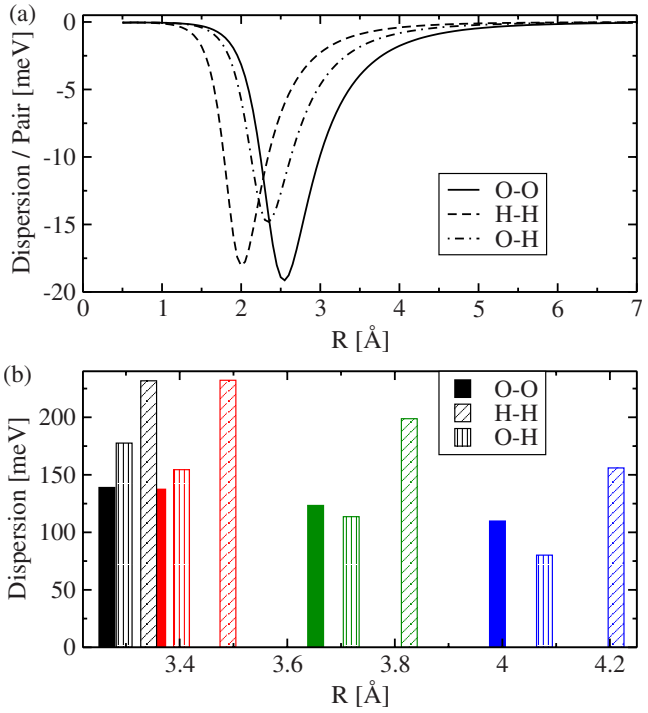

FIG. 4. (Color online) (a) Variation in the dispersion contribution with distance from different atom pairs with parameters for BLYP. (b) Intermolecular dispersion interaction for the four isomers as a function of the average interatomic distances of different atom pairs (on BLYP+vdW optimized structures). Here black, red, green, and blue refer to prism, cage, book, and cyclic isomers, respectively.

forces are generally considered to be long range and indeed the tails of all three vdW curves extend to beyond 4 Å. However, the minima of the vdW curves with the specific parameters employed here are located at considerably shorter distances: ~2.80, ~2.20, and ~2.55 Å for the O–O, H–H, and O–H curves, respectively. It is the location of these vdW minima relative to the structures of the various isomers that leads to the revised energetic ordering of the four isomers. In simplest terms the mean intermolecular distances of the four clusters decreases upon going from cyclic to book to cage to prism and so the magnitude of the dispersion correction decreases in the order of prism to cage to book to cyclic, which ultimately leads to the correct stability sequence of prism to cage to book to cyclic. Considering this in more detail we show in Fig. 4(b) the contributions to the total intermolecular dispersion interaction in each cluster for each type of atomic pair interaction (O–O, H–H, and O–H), plotted as a function of distance.[116] It can be seen from the histogram that the average intermolecular O–O, O–H, and H–H distances steadily increase along the sequence of prism-cage-book-cyclic and that likewise the dispersion contribution decreases. Further, we note that by simply summing up the contributions from each type of interaction in the hexamers we find that the majority of the vdW correction comes from H–H interactions (~44%–48%), followed by the O–H (~22%–32%) and then the O–O (~25%–30%) interactions. The H–H interaction dominates simply because there are more of them. For brevity we do not show the results of similar analysis performed for the PBE and PBE0 vdW corrections. However, the general conclusion that the vdW dispersion contribution favors the more compact prism and cage isomers over the less compact book and cyclic isomers because the former are closer to the minima of the vdW curves than the latter also holds for the PBE and PBE0 vdW corrections.

Finally, this paper has focused on water clusters. Clusters often exhibit quite different properties from the corresponding bulk substance. However, it does not seem unreasonable to make some speculations about the relevance of the results presented here to DFT simulations of liquid water. As indicated in Sec. I, the simulation of liquid water with DFT is by no means free from controversy.[9–12,15–18] Of the many functionals tested for liquid water, BLYP appears to provide comparatively good agreement with experiment in terms of, e.g., the O–O RDF and diffusion coefficient.[16] However, precise quantitative agreement with experiment for BLYP or, indeed, any xc functional remains beyond reach. It seems likely that if an xc functional fails to predict the correct energetic ordering of the low energy isomers of the water hexamer, then similar errors will exist in describing the many more competing configurations of water clusters present transiently or otherwise in the liquid. Given that the hybrid xc functionals PBE0 and X3LYP also fail for the hexamer despite otherwise predicting equilibrium H bond strengths and structures for smaller water clusters in excellent agreement with MP2, it seems likely that these functionals may not offer the promise anticipated for liquid water.[1] Indeed a very recent PBE0 simulation for liquid water, which ran for a reasonably respectable 10 ps, found that the PBE and PBE0 RDFs were essentially indistinguishable.[117] Based on the foregoing results and discussion with unreliable results obtained for the hexamer the lack of a significant improvement in describing the liquid is not entirely unexpected. We suggest instead that density-functional methodologies which account for vdW dispersion forces are likely to offer more promise in the quest to improve the description of liquid water. Again very recent MD simulations of liquid water are consistent with this suggestion. Lin *et al.*[118] reported BLYP simulations for liquid water corrected with a similar C_6R^{-6} correction scheme to the one employed here (but with a different damping function) as well as a separate account for vdW through the use of modified pseudopotentials. These simulations indicate that (at the experimental density and temperatures tested) accounting for vdW forces lowers the peak maximum in the O–O RDF and in so doing brings the experimental and theoretical RDFs into better agreement. However, others have suggested that dispersion interactions are not very important for liquid water under ambient conditions[119] and so it appears that considerably more work is needed to address this issue. In current work we are focusing on establishing precisely how vdW dispersion forces impact on the properties of liquid water in addition to understanding how reliable gradients (forces) of structures displaced from minimum energy positions are with and without vdW corrections.

ACKNOWLEDGMENTS

This work is supported by the European Commission through the Early Stage Researcher Training Network MONET, Grant No. MEST-CT-2005-020908 (www.sljus.lu.se/monet). A.T. thanks the Alexander von Humboldt (AvH) Foundation for funding. A.M.'s work is supported by a EURYI award (www.esf.org/euryi) and by the EPSRC. Some of the calculations reported here were performed at the London Centre for Nanotechnology, to which we are very grateful for support.

[1] B. Santra, A. Michaelides, and M. Scheffler, J. Chem. Phys. **127**, 184104 (2007).
[2] S. Tsuzuki and H. P. Lüthi, J. Chem. Phys. **114**, 3949 (2001).
[3] J. J. Novoa and C. Sosa, J. Phys. Chem. **99**, 15837 (1995).
[4] Y. Zhao and D. G. Truhlar, J. Chem. Theory Comput. **1**, 415 (2005).
[5] E. E. Dahlke and D. G. Truhlar, J. Phys. Chem. B **109**, 15677 (2005).
[6] E. E. Dahlke, R. M. Olson, H. R. Leverentz, and D. G. Truhlar, J. Phys. Chem. A **112**, 3976 (2008).
[7] J. T. Su, X. Xu, and W. A. Goddard III, J. Phys. Chem. A **108**, 10518 (2004).
[8] J. Ireta, J. Neugebauer, and M. Scheffler, J. Phys. Chem. A **108**, 5692 (2004).
[9] T. Todorova, A. P. Seitsonen, J. Hutter, I.-F. W. Kuo, and C. J. Mundy, J. Phys. Chem. B **110**, 3685 (2006).
[10] M. V. Fernández-Serra and E. Artacho, J. Chem. Phys. **121**, 11136 (2004).
[11] H.-S. Lee and M. E. Tuckerman, J. Phys. Chem. A **110**, 5549 (2006).
[12] J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik, and M. Parrinello, J. Chem. Phys. **122**, 014515 (2005).
[13] Ph. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. Å. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, and A. Nilsson, Science **304**, 995 (2004).
[14] J. D. Smith, C. D. Cappa, K. R. Wilson, B. M. Messer, R. C. Cohen, and R. J. Saykally, Science **306**, 851 (2004).
[15] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).
[16] H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **125**, 154507 (2006).
[17] H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **126**, 164501 (2007).
[18] P. H.-L. Sit and N. Marzari, J. Chem. Phys. **122**, 204510 (2005).
[19] X. L. Hu and A. Michaelides, Surf. Sci. **601**, 5378 (2007).
[20] A. Michaelides, Appl. Phys. A: Mater. Sci. Process. **85**, 415 (2006).
[21] P. J. Feibelman, Science **295**, 99 (2002).
[22] V. A. Ranea, A. Michaelides, R. Ramírez, P. L. de Andres, J. A. Vergés, and D. A. King, Phys. Rev. Lett. **92**, 136104 (2004).
[23] A. Michaelides and P. Hu, J. Chem. Phys. **114**, 513 (2001).
[24] S. Meng, E. G. Wang, and S. Gao, Phys. Rev. B **69**, 195404 (2004).
[25] A. Michaelides and K. Morgenstern, Nature Mater. **6**, 597 (2007).
[26] A. Michaelides, Faraday Discuss. **136**, 287 (2007).
[27] X. Xu and W. A. Goddard III, Proc. Natl. Acad. Sci. U.S.A. **101**, 2673 (2004).
[28] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[29] C. Adamo and V. Barone, J. Chem. Phys. **108**, 664 (1998).
[30] W. Lee, C. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[31] C. J. Tsai and K. D. Jordan, Chem. Phys. Lett. **213**, 181 (1993).
[32] K. Kim, K. D. Jordan, and T. S. Zwier, J. Am. Chem. Soc. **116**, 11568 (1994).
[33] K. Liu, M. G. Brown, R. J. Saykally, J. K. Gregory, and D. C. Clary, Nature (London) **381**, 501 (1996).
[34] J. K. Gregory and D. C. Clary, J. Phys. Chem. **100**, 18014 (1996).
[35] J. K. Gregory and D. C. Clary, J. Phys. Chem. A **101**, 6813 (1997).
[36] J. Kim and K. S. Kim, J. Chem. Phys. **109**, 5886 (1998).
[37] M. Losada and S. Leutwyler, J. Chem. Phys. **117**, 2003 (2002).
[38] K. Nauta and R. E. Miller, Science **287**, 293 (2000).
[39] S. S. Xantheas, C. J. Burnham, and R. J. Harrison, J. Chem. Phys. **116**, 1493 (2002).
[40] R. M. Olson, J. L. Bentz, R. A. Kendall, M. W. Schmidt, and M. S. Gordon, J. Chem. Theory Comput. **3**, 1312 (2007).
[41] K. Laasonen, M. Parrinello, R. Car, C. Lee, and D. Vanderbilt, Chem. Phys. Lett. **207**, 208 (1993).
[42] C. Lee, H. Chen, and G. Fitzgerald, J. Chem. Phys. **101**, 4472 (1994).
[43] D. A. Estrin, L. Paglieri, G. Corongiu, and E. Clementi, J. Phys. Chem. **100**, 8701 (1996).
[44] V. F. Petrenko and R. W. Whitworth, *Physics of Ice* (Oxford University Press, New York, 2003).
[45] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
[46] E. J. Bylaska, W. A. de Jong, K. Kowalski, T. P. Straatsma, M. Valiev, D. Wang, E. Aprà, T. L. Windus, S. Hirata, M. T. Hackler, Y. Zhao, P.-D. Fan, R. J. Harrison, M. Dupuis, D. M. A. Smith, J. Nieplocha, V. Tipparaju, M. Krishnan, A. A. Auer, M. Nooijen, E. Brown, G. Cisneros, G. I. Fann, H. Fruchtl, J. Garza, K. Hirao, R. Kendall, J. A. Nichols, K. Tsemekhman, K. Wolinski, J. Anchell, D. Bernholdt, P. Borowski, T. Clark, D. Clerc, H. Dachsel, M. Deegan, K. Dyall, D. Elwood, E. Glendening, M. Gutowski, A. Hess, J. Jaffe, B. Johnson, J. Ju, R. Kobayashi, R. Kutteh, Z. Lin, R. Littlefield, X. Long, B. Meng, T. Nakajima, S. Niu, L. Pollack, M. Rosing, G. Sandrone, M. Stave, H. Taylor, G. Thomas, J. van Lenthe, A. Wong, and Z. Zhang, NWCHEM, a computational chemistry package for parallel computers, Version 5.0, 2006, Pacific Northwest National Laboratory, Richland, Washington 99352-0999.
[47] We have used GAUSSIAN 03 (Ref. 45) and NWCHEM (Ref. 46) interchangeably, since total energies of the water clusters obtained from the two codes differ by no more than 0.4 meV per water molecule.
[48] I. M. B. Nielsen, E. T. Seidl, and C. L. Janssen, J. Chem. Phys. **110**, 9435 (1999).
[49] D. Feller, J. Chem. Phys. **98**, 7059 (1993).
[50] C. Schwartz, Phys. Rev. **126**, 1015 (1962).
[51] W. Kutzelnigg and J. D. Morgan III, J. Chem. Phys. **96**, 4484 (1992).
[52] A. K. Wilson and T. H. Dunning, Jr., J. Chem. Phys. **106**, 8718 (1997).
[53] As we have said before (Ref. 1), we opted for the scheme provided by Eqs. (1) and (2) because we found that with input from triple-, quadruple-, and pentuple-ζ basis sets this method was best able to predict the total energy of a water monomer and dimer explicitly calculated with an aug-cc-pV6Z basis set (443 basis functions/H_2O).
[54] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[55] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).
[56] S. Manten and A. Lüchow, J. Chem. Phys. **115**, 5362 (2001).
[57] J. C. Grossman, J. Chem. Phys. **117**, 1434 (2002).
[58] C. Filippi, S. B. Healy, P. Kratzer, E. Pehlke, and M. Scheffler, Phys. Rev. Lett. **89**, 166102 (2002).
[59] I. G. Gurtubay and R. J. Needs, J. Chem. Phys. **127**, 124306 (2007).
[60] M. Korth, A. Lüchow, and S. Grimme, J. Phys. Chem. A **112**, 2104 (2008).
[61] M. Fuchs, C. Filippi, J. Ireta, and M. Scheffler (unpublished), http://w3.rz-berlin.mpg.de/%7Efuchs/member/vdwdft2005.pdf
[62] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).
[63] M. Zaccheddu, C. Filippi, and F. Buda, J. Phys. Chem. A **112**, 1627 (2008).
[64] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[65] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).
[66] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., J. Comput. Chem. **14**, 1347 (1993).
[67] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).
[68] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007); **27**, 109901 (2007).
[69] C. J. Umrigar and C. Filippi, CHAMP, Cornell Holland *ab Initio* materials package.
[70] The Gaussian basis sets are $(12s11p1d)/[3s3p1d]$ for oxygen and $(12s10p)/[2s2p]$ for hydrogen based on the VDZ basis sets of Ref. 68 with additional s (O) and p (O, H) functions. The Jastrow correlation factor is the product of exponentials of fifth order polynomials in the (i) electron-nucleus (e-n), (ii) electron-electron (e-e), and (iii) mixed three-body e-n and e-e distances of Ref. 65, adapted to the use of pseudopotentials. Our DMC results for the binding energy of the water dimer calculated at different DMC time steps $\tau=1/20,\ldots,1/10^3$ E_h^{-1} showed no significant differences. In our production runs we chose $\tau=1/80$ E_h^{-1}.
[71] W. Klopper, J. G. C. M. van Duijneveldt-van de Rijdt, and F. B. van Duijneveldt, Phys. Chem. Chem. Phys. **2**, 2227 (2000), in our DMC

calculations of the water dimer we use the same geometry as reported in this paper.
[72] N. A. Benedek, I. K. Snook, M. D. Towler, and R. J. Needs, J. Chem. Phys. **125**, 104302 (2006).
[73] C. Diedrich, A. Lüchow, and S. Grimme, J. Chem. Phys. **123**, 184106 (2005).
[74] J. P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie, Berlin, 1991), p. 11.
[75] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[76] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[77] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).
[78] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
[79] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[80] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[81] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[82] H. L. Schmider and A. D. Becke, J. Chem. Phys. **108**, 9624 (1998).
[83] J. Hutter *et al.*, computer code CPMD, version 3.11; Copyright IBM Corp. 1990–2006, Copyright für Festkörperforschung Stuttgart, Germany, 1997–2001 (http://www.cpmd.org/).
[84] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, "*Ab initio* molecular simulations with numeric atom-centered orbitals: FHI-aims," Comput. Phys. Commun. (in press).
[85] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).
[86] M. Krack, Theor. Chem. Acc. **114**, 145 (2005).
[87] Since 200–250 Ry plane-wave cutoff energies are large and far from common, we have also calculated dissociation energies for BLYP using a 70 Ry cutoff with CPMD. The dissociation energies of the prism, cage, book, and cyclic structures from CPMD with the 70 Ry (250 Ry) cut-off are 270.3 (272.1), 272.5 (276.0), 281.6 (286.7), and 279.7 (289.7) meV/H_2O, respectively. Thus the absolute energies at the smaller cutoff differ from those at the larger cutoff by 2–10 meV/H_2O. This error in the basis set at 70 Ry is sufficient to change the energetic ordering of the clusters, incorrectly making the book the most stable conformer. These results clearly illustrate the importance of fully converged plane-wave cutoffs when high accuracy is desired for weakly bonded systems such as those under consideration here.
[88] R. W. Hockney, Methods Comput. Phys. **9**, 136 (1970).
[89] It is somewhat difficult to determine the precise "error bar" associated with our MP2/CBS estimates. However, as we noted before through comparisons of MP2/CBS and CCSD(T)/CBS, the error [due to MP2 (valence only) treatment of correlation, aug-cc-pVTZ structures, extrapolation to reach the CBS, etc.] is not likely to be more than 5 meV/H bond.
[90] All our dissociation energies come within ~1.3 meV/H_2O of Ref. 39. These differences are small. Most of it can be attributed to the use of different extrapolation methods. Applying the extrapolation scheme used by Xantheas *et al.* (Ref. 39) (extrapolating the dissociation energies with a 4, 5 polynomial and taking double-, triple-, quadruple, and pentuple-ζ values) the dissociation energies obtained by us (Xantheas *et al.*) are 331.54 (331.45), 331.19 (330.94), 329.59 (329.63), and 324.19 (324.22) meV/H_2O for the prism, cage, book, and cyclic, respectively. Thus when the same extrapolation scheme is used ours and the MP2/CBS results of Xantheas *et al.* agree to within 0.3 meV/H_2O.
[91] We believe that the differences between CCSD(T) and MP2 (10–15 meV/H_2O) come mostly from the use of an aug-cc-pVTZ basis set for the CCSD(T) energy calculations. Indeed with this basis set the difference between MP2 and CCSD(T) is reduced to only 4 meV/H_2O. Specifically, our MP2/aug-cc-pVTZ dissociation energies are 346.5, 345.9, 343.0, and 335.1 meV/H_2O, respectively, for prism, cage, book, and cyclic isomers. In addition, the different basis sets used to obtain the structures is likely to account for some of the remaining difference. Here we use MP2 structures optimized with an aug-cc-pVTZ basis set where as in Ref. 40, MP2 structures optimized with a polarized double-ζ basis set of Dunning, Jr. and Hay (Ref. 121) [DH(d,p)] are used.
[92] For these hexamers it is invariably assumed that the prism, cage, book, and cyclic isomers have nine, eight, seven, and six H bonds, respectively (Refs. 36, 37, and 39). In Fig. 2 we have used these conventional H bond numbers. However, it is interesting to note that upon inspection of our optimized MP2 and DFT structures of the four hexamers and employing several standard geometric definitions of H bonds between water molecules, we find that the number of H bonds counted depends sensitively on which definition is used. See the supporting information for more details (Ref. 122).
[93] Dissociation energies calculated (using BLYP), respectively, with G03/aug-cc-pV5Z, CPMD, and FHI-AIMS codes are the following: dimer (180.7, 179.3, 180.6), trimer (191.7, 190.8, 191.9), tetramer (264.9, 263.9, 265.1), and pentamer (281.2, 280.7, 281.3). Likewise in comparison to G03/aug-cc-pVTZ structures, CPMD and FHI-AIMS, respectively, concede maximum errors of 0.031 and 0.009 Å for bond lengths and 0.57° and 0.03° for bond angles (parameters considered are shown in Table II).
[94] G. C. Shields and K. N. Kirschner, Synth. React. Inorg., Met.-Org., Nano-Met. Chem. **38**, 32 (2008).
[95] S. S. Xantheas, J. Chem. Phys. **100**, 7523 (1994).
[96] S. S. Xantheas, Chem. Phys. **258**, 225 (2000).
[97] J. M. Pedulla, F. Vila, and K. D. Jordan, J. Chem. Phys. **105**, 11091 (1996).
[98] M. Quack, J. Stohner, and M. A. Suhm, J. Mol. Struct. **599**, 381 (2001).
[99] S. Kristyán and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).
[100] J. M. Pérez-Jordá and A. D. Becke, Chem. Phys. Lett. **233**, 134 (1995).
[101] R. Ahlrichs, R. Penco, and G. Scoles, Chem. Phys. **19**, 119 (1977).
[102] Q. Wu and W. Yang, J. Chem. Phys. **116**, 515 (2002).
[103] S. Grimme, J. Comput. Chem. **25**, 1463 (2004).
[104] P. Jurečka, J. Černý, P. Hobza, and D. R. Salahub, J. Comput. Chem. **28**, 555 (2007).
[105] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **92**, 246401 (2004).
[106] A. D. Becke and E. R. Johnson, J. Chem. Phys. **122**, 154104 (2005).
[107] P. L. Silvestrelli, Phys. Rev. Lett. **100**, 053002 (2008).
[108] O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, Phys. Rev. Lett. **93**, 153004 (2004).
[109] We have used Bondi's (Ref. 123) vdW radii for different atoms. More precisely for the atom pairs we used the following radii and vdW coefficients (in a.u.): R^0_{H-H}=4.54, R^0_{O-O}=5.74, R^0_{O-H}=5.28, C_{6H-H}=2.78, C_{6O-O}=12.14, and C_{6O-H}=5.67. For more details on the averaging of vdW radii and dispersion coefficients for pairs of atoms see Ref. 104.
[110] A. Tkatchenko, V. Blum, and M. Scheffler (unpublished).
[111] Also we made fits on various water dimers to reproduce the MP2 two-body interaction by optimizing the correction parameters using BLYP. A large set of parameters was considered by a careful scan over the range $0.8 \leq s_R \leq 1.5$ and $20 \leq d \leq 40$. All possible combinations of water dimers from all four hexamers (a total of 48 dimers) were included in the fitting and interestingly the set of parameters which yielded the smallest MAE for the 48 dimers are s_R=0.82 and d=20, which are nearly the same as those found from the S22 database.
[112] A. Tkatchenko and O. A. von Lilienfeld, Phys. Rev. B **78**, 045116 (2008).
[113] A. Tkatchenko and M. Scheffler, "Accurate van-der-Waals interactions from (semi-)local density functional theory," Phys. Rev. Lett. (submitted).
[114] Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).
[115] For the M05-2X (Ref. 114) functional with an aug-cc-pVQZ basis set the dissociation energies obtained are 353.5, 347.9, 339.9, and 333.1 meV/H_2O for prism, cage, book, and cyclic, respectively.
[116] The total intramolecular O–H and H–H dispersion contributions are 0.6 and ~13–17 meV, respectively. These are small compared to intermolecular contributions and remain almost constant for all the hexamers and for brevity are not shown in Eq. 4.
[117] M. Guidon, F. Schiffmann, J. Hutter, and J. VandeVondele, J. Chem. Phys. **128**, 214104 (2008).
[118] I.-C. Lin, A. P. Seitsonen, M. D. Coutinho-Neto, I. Tavernelli, and U. Rothlisberger, "Importance of van der Waals interactions in liquid water," J. Phys. Chem. B (in press).
[119] O. Akin-Ojo, Y. Song, and F. Wang, J. Chem. Phys. **129**, 064108 (2008).
[120] We have also performed full geometry optimizations with vdW corrected BLYP and PBE using the CPMD code and the total dissociation energies are within 1 meV from the numbers presented in Table III. We also note that with the FHI-AIMS code PBE0+vdW calculations were performed on PBE+vdW optimized geometries.
[121] T. H. Dunning, Jr. and P. J. Hay, *Methods of Electronic Structure Theory* (Plenum, New York, 1997), p. 1.

[122] See EPAPS Document No. E-JCPSA6-129-042842 for a database of the coordinates (optimized consistently with MP2, 12 DFT xc functionals, and HF with an aug-cc-pVTZ basis set and vdW corrected geometries optimized with all-electron NAOs) and the total energies (obtained with an aug-cc-pV5Z basis set) of each isomer studied here. A table with the number of H bonds in the prism isomer, obtained with various H bond definitions, is also provided as well as information on the typical computational cost of the various methods considered here. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[123] A. J. Bondi, J. Phys. Chem. **68**, 441 (1964).