Accurate Energies and Structures for Large Water Clusters Using the X3LYP Hybrid Density Functional

Julius T. Su, Xin Xu, and William A. Goddard III*

Materials and Process Simulation Center (139-74), Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

We predict structures and energies of water clusters containing up to 19 waters with X3LYP, an extended hybrid density functional designed to describe noncovalently bound systems as accurately as covalent systems. Our work establishes X3LYP as the most practical ab initio method today for calculating accurate water cluster structures and energies. We compare X3LYP/aug-cc-pVTZ energies to the most accurate theoretical values available ($n = 2$−6, 8), MP2 with basis set superposition error (BSSE) corrections extrapolated to the complete basis set limit. Our energies match these reference energies remarkably well, with a root-mean-square difference of 0.1 kcal/mol/water. X3LYP also has *ten times less BSSE* than MP2 with similar basis sets, allowing one to neglect BSSE at moderate basis sizes. The net result is that X3LYP is ∼*100 times faster* than canonical MP2 for moderately sized water clusters.

1. Introduction

We predict structures and energies of water clusters containing up to 19 waters with X3LYP,[1,2] an extended hybrid density functional designed to describe noncovalently bound systems well. Our work establishes X3LYP as the most practical ab initio method today for calculating accurate water cluster structures and energies.

We compare our X3LYP results to the most accurate theory available[3−8] for modest-sized water clusters, MP2 calculations using triple-ζ-plus basis sets with basis set superposition error corrections extrapolated to the complete basis set limit. Our energies match these reference energies to a root-mean-square (rms) deviation of 0.1 kcal/mol of water.

This agreement is remarkable, especially since the noncovalent bonding in water clusters (polar, hydrogen bonded) differs greatly from the bonding in the rare neutral gas dimers used to train X3LYP. In contrast, the popular hybrid functional B3LYP[9−11] provides acceptable geometries and thermochemistry for covalent molecules, but its poor description of London dispersion (van der Waals attraction) leads to poor binding energies[4,12−15] (Table 1) for water clusters.

Two consequences follow:

First, the result establishes the generality of the X3LYP functional, supporting its application to more diverse van der Waals and hydrogen bonded complexes. This validation sets the stage for first principles predictions of noncovalent interactions of ligands to proteins and DNA, with implications for the emerging field of genome-wide structure based drug design.

Second, X3LYP now represents the state of the art for practical ab initio calculations on water clusters, since

(1) *We can use smaller basis sets while preserving accuracy.* Post-Hartree−Fock methods such as MP2 require higher angular momentum basis functions to properly describe the correlation cusp[16] and suffer from slow and unsystematic convergence to the complete basis set limit.[17]

We expect the basis set requirements for DFT methods to be greatly reduced, and our results bear this out: X3LYP/aug-cc-pVTZ agrees with MP2/aug-cc-pV5Z extrapolated to the complete basis set limit to within 0.1 kcal/mol/water, a difference well within the uncertainty of both methods.

(2) *We can neglect BSSE at moderate basis sizes.* Basis set superposition error has long plagued canonical MP2 calculations, with a correction of ∼1.1 kcal/mol for water hexamer even with the aug-cc-pV5Z basis set.[3] This is larger than the energy difference between water hexamer isomers (<0.5 kcal/mol). X3LYP has *ten times less* basis set superposition error than MP2 with comparable basis sets, allowing smaller basis sets to be used. Non-BSSE and BSSE energies converge quickly to the same value with increasing basis set size, so that for moderate sized bases (aug-cc-pVTZ), we can neglect BSSE.

Not including BSSE in X3LYP calculations speeds up our calculations significantly, since a BSSE calculation requires N single point energies with the full system basis, where N is the number of water monomers in the complex.

(3) *Density functional methods are faster than MP2.* For larger clusters, X3LYP is at least 100 times faster than canonical MP2 at the same basis set level, where BSSE is neglected for both calculations. The speed advantage becomes even bigger for larger clusters, since density functional methods scale as a factor of N better than canonical MP2 (formally N^4 vs N^5, with improvements possible for both).

With this superior combination of speed and accuracy, we expect X3LYP to displace MP2-corrected Hartree−Fock (HF) as the preferred method for performing ab initio calculations on water clusters.

2 Computational Details

2.1. X3LYP Functional. The details of the X3LYP hybrid density functional are described elsewhere.[1,2] The X3LYP hybrid functional was developed to describe accurately the thermochemistry of molecules while reproducing the properties (equilibrium distance, binding energy, and Pauli repulsion) of helium

---

* To whom correspondence should be addressed. E-mail: wag@wag.caltech.edu.

Accurate Energies and Structures for Large Water Clusters

J. Phys. Chem. A, Vol. 108, No. 47, 2004 **10519**

TABLE 1: Binding Energies of Presumed Global Minimum $(H_2O)_n$ Clusters ($-\Delta E$, kcal/mol)[a]

n	structure	6-31g**			6-311g**++			aug-cc-pVDZ		aug-cc-pVTZ(-f)			MP2/CBS[3]	MP2/TZ2P++[5]
		LMP2	B3LYP	X3LYP	LMP2	B3LYP	X3LYP	B3LYP	X3LYP	LMP2	B3LYP	X3LYP	Xantheas	Lee
2	linear	6.55	7.56	7.96	5.06	5.82	6.23	4.71	5.11	4.43	4.61	5.00	4.97	4.88
3	cyclic	20.94	25.03	26.24	14.64	17.30	18.45	14.76	15.91	12.31	14.42	15.52	15.82	15.11
4	cyclic	34.94	41.73	43.57	24.48	30.73	32.49	26.79	28.55	17.12	26.03	27.73	27.63	26.72
5	cyclic	44.68	53.34	55.71	32.33	40.78	43.05	35.53	37.83	28.65	34.37	36.57	36.28	35.17
6	cage	58.34	70.73	74.27	39.86	48.83	52.06	42.91	46.14	34.75	41.70	44.78	45.79	44.04
7	prism'	73.04	87.02	91.41	49.12	60.12	64.07	52.74	56.67	42.71	51.45	55.27		54.81
8	D2d	92.61	110.73	116.14	64.14	77.01	81.88	68.35	73.27	59.69	66.60	71.35	72.57	70.06
9	D2dDD	99.37	123.08	129.03	71.38	87.94	93.41	77.36	82.93		75.07	80.36		79.14
10	prism	117.03	139.83	146.80	82.40	99.75	106.00	87.84	93.98		85.81	91.82		90.07
11	Pr443											97.79		96.69
12	Pr444											112.14		112.59
13	Pr454											122.41		
14	Pr2444											133.82		
15	Pr555											142.34		
16	Pr4444											153.08		
17	Pr454(4)											163.20		
18	Pr44244											175.65		
19	globular											184.13		

[a] The LMP2, B3LYP, and X3LYP results have not been corrected for BSSE; the MP2/CBS results have been extrapolated to a complete basis set; and the MP2/TZ2P++ results include 50% of the BSSE correction. The binding energy is given relative to fully separated and relaxed water monomers. Geometry labeling follows the convention of Lee et al.[5−7]

and neon dimers, whose binding is wholly due to dispersion. For these rare gas dimers, the repulsive energy component of X3LYP (total energy minus correlation) is fit to match Hartree−Fock energies. Thus, near equilibrium distances, X3LYP is expected to give correct contributions of dispersion to bonding.

X3LYP extends B3LYP[18] by writing the nonlocal gradient correction in terms of the F^X extended exchange functional, which is written as a linear combination of B88 and PW91 exchange functionals:

$$E_{XC}^{X3LYP} = a_0 E_x^{exact} + (1 - a_0)E_x^{Slater} + a_x \Delta E_x^X + a_c E_c^{VWN} + (1 - a_c)E_c^{LYP} \quad (1)$$

$$\Delta E_x^X = E_X^{LDA} - \int F^X(s)\rho^{4/3}dr \quad (2)$$

$$s = \frac{|\nabla\rho|}{(24\pi^2)^{1/3}\rho^{4/3}} \quad (3)$$

$$F^X(s) = 1 + a_{x1}(F^{B88}(s) - 1) + (1 - a_{x1})(F^{PW91}(s) - 1) \quad (4)$$

$$\{a_0, a_x, a_{x1}, a_c\} = \{0.218, 0.709, 0.765, 0.129\} \quad (5)$$

The four mixing parameters were determined through a least-squares fit to the total energies of 10 atoms, the ionization potentials of 16 atoms, the electron affinities of 10 atoms, and the atomization energies for 33 diatomic and five triatomic molecules selected to represent the important chemistry of first- and second-row elements (including open- and closed-shell molecules; molecules with single, double, and triple bonds; ionic systems; and systems requiring multiple determinants for proper descriptions). Helium and neon rare gas dimers were included as representative van der Waals systems, but no data on water dimer or higher clusters were included.

The accuracy of X3LYP for the thermochemistry (cohesive energies, ionization potentials, electron affinities, proton affinities) of the G2 set of 148 molecules is better than all other DFT methods considered (seven GGA methods and seven hybrid methods) as is the s→d excitation energies for transition-metal atoms. An earlier test for water dimer[2] led to a binding energy (D_e) within 0.05 kcal/mol of the exact value and a O−O distance (R_e) within 0.004 Å of the exact value.

2.2. Quantum Mechanics Calculations. All calculations were performed using the Jaguar 5.0[19] software package, with default options unless indicated otherwise.

In the LMP2 method,[20,21] occupied orbitals are only allowed to correlate with virtual orbitals localized on the atoms of the local occupied Hartree−Fock orbital, with an initial wave function obtained from Pipek−Mezey localization[22] of the HF reference wave function. Only valence electrons were included in the LMP2 correlation. In all cases, SCF convergence under the DIIS scheme was achieved to 50 μhartree.

For LMP2, B3LYP, and X3LYP, the default pseudospectral implementation of Jaguar was used to accelerate evaluation of two-electron integrals. In previous X3LYP calculations,[12,13] the pseudospectral capabilities were turned off to simplify comparison with previous results obtained using other methods.

All geometries were converged to a maximum gradient of 4.5×10^{-4} hartree/bohr, an rms gradient of 3.0×10^{-4} hartree/bohr, a maximum nuclear displacement of 1.8×10^{-3} bohr, and an rms nuclear displacement of 1.2×10^{-3} bohr.

We used the following basis sets: 6-31g**[23] (25 basis functions/water), 6-311++g** (36 basis functions/water), aug-cc-pVDZ[24] (41 basis functions/water), and aug-cc-VTZ(-f)[24] (58 basis functions/water, without f functions). BSSE corrections were carried out where stated explicitly, using the function counterpoise method[25] and taking into account fragment relaxation energy terms[26,27]

$$\Delta E(BSSE) = E^{full}(full) - \sum_{fragments} E^{full}(fragment) + \sum_{fragments} \Delta E_{relax}^{fragment}(fragment) \quad (6)$$

where E^{full} and $E^{fragment}$ indicate the energy calculated with the full- and fragment-only basis sets, respectively.

2.3. Multibody Decomposition. A multibody decomposition of total binding energy for water hexamers was computed by taking into account $2^6 - 1 = 63$ possible present/absent combinations of water fragments and computing their energies using both a fragment-only basis and a full-system basis (to estimate the magnitude of BSSE). Although the complex is symmetric, symmetry was not used. The final multibody contributions Δ^n can be written in terms of linear combinations
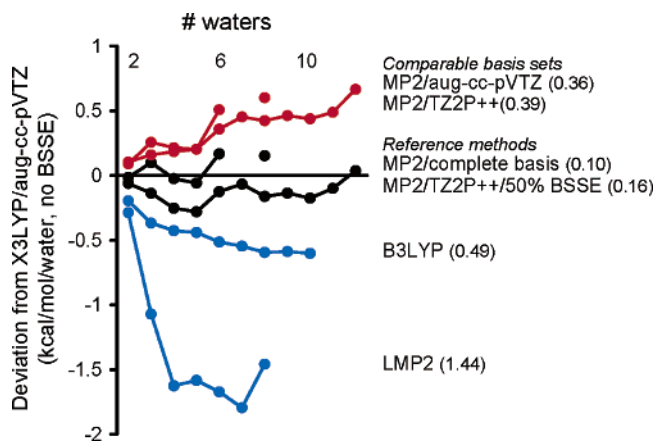
Figure 1. Deviation of global minimum water cluster energies for different levels of theory. Here the reference is X3LYP/aug-cc-pVTZ. We compare MP2 energies from Xantheas[3] and Lee[7] with comparable basis sets; MP2 reference energies obtained from extrapolation to a complete basis[3] (BSSE and no BSSE converge to same energies) and from inclusion of 50% BSSE;[7] and B3LYP and LMP2 energies using the aug-cc-pVTZ(-f) basis set (no BSSE). The total root-mean-squared errors (kcal/mol/water) are indicated in parentheses.

of Σ^n; the sum of energies of all species with n fragments included

$$\begin{pmatrix} \Delta^6 \\ \Delta^5 \\ \Delta^4 \\ \Delta^3 \\ \Delta^2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ & 1 & -2 & 3 & -4 & 5 \\ & & 1 & -3 & 6 & -10 \\ & & & 1 & -4 & 10 \\ & & & & 1 & -5 \end{pmatrix} \begin{pmatrix} \Sigma^6 \\ \Sigma^5 \\ \Sigma^4 \\ \Sigma^3 \\ \Sigma^2 \\ \Sigma^1 \end{pmatrix} \quad (7)$$

and $\Delta^1 = \Sigma^1 - 6\,E(\text{ref water})$. The sum of all Δ^n gives the total binding energy.

3. Results and Discussion

3.1. Water Cluster Global Minima. To compare the overall energetics of clusters up to 19 waters, we started with globally minimized water clusters from Wales et al.[28] (optimized with the TIP4P force field) and carried out a full optimization for each level of theory and basis set presented in Table 1. Figure 1 shows that even *without* using BSSE corrections, the X3LYP/aug-cc-pVTZ(-f) energies are in excellent agreement with the best theoretical estimates available, deviating by an rms of 0.10 kcal/mol/water from the results of Xantheas et al.[3,8] (MP2/CBS extrapolation with a polynomial function from an aug-cc-V5Z basis), who considered up to eight waters. Lee et al.[7] carried out a less complete MP2 than Xantheas (MP2 using the triple-ζ TZ2P^{2+} basis with 50% BSSE correction) but considered up to 12 waters; their cluster binding energies are systematically higher than Xantheas' energies (presumed to be more accurate) by ~0.2–0.3 kcal/mol/water, differing from our energies by an RMS of 0.16 kcal/mol/water. Our results agree well with Xantheas, with little evidence of systematic error (Figure 1).

Figure 1 and Table 1 compare non-BSSE-corrected energies calculated at different levels of theory. Like X3LYP, B3LYP converges quickly to a limiting energy with increasing basis set size, and B3LYP-optimized geometries are similar to X3LYP-optimized geometries (six-cage $C_{rms} = 0.02$ Å, 8-D_{2d} $C_{rms} = 0.01$ Å). However, B3LYP systematically underestimates water cluster binding energies (rms of 0.51 kcal/mol/water vs MP2/CBS).
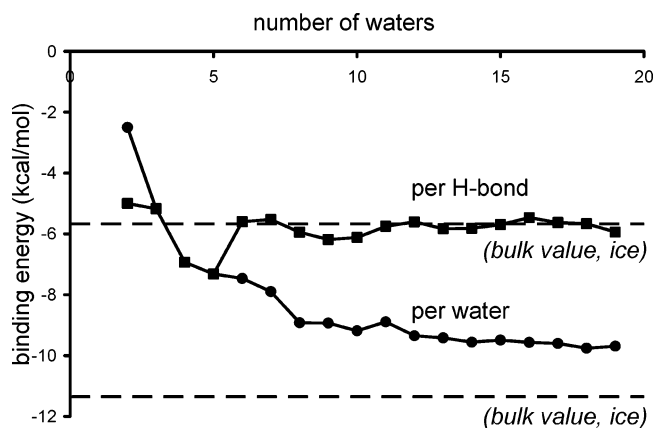
Figure 2. Binding energy (kcal/mol) per hydrogen bond and per water molecule for global minimum water clusters at the level of X3LYP/aug-cc-pVTZ(-f). The energy per hydrogen bond converges quickly to the experimental binding energy of ice at 0 K, $\Delta H/2 = -5.68$ kcal/mol, but the energy per water does not due to the five to seven "dangling" hydrogen bonds present in the larger clusters ($n = 6-19$).

LMP2 performs even more poorly than B3LYP, converging more slowly to a limiting energy with increasing basis set size and more significantly underestimating water cluster binding energies (rms of 1.43 kcal/mol/water vs MP2/CBS). LMP2-optimized geometries are distorted relative to X3LYP-optimized geometries (six-cage $C_{rms} = 0.70$ Å, 8-D_{2d} $C_{rms} = 0.33$ Å) and are characterized by longer hydrogen bonds and larger out of plane distortions for the "cyclic" complexes. Thus, our results suggest LMP2 is unsuitable for describing water clusters, contrary to the conclusion of previous studies,[29] which considered single-point LMP2 energies at MP2-optimized geometries.

Canonical MP2 (non-BSSE corrected) calculations with aug-cc-pVTZ and TZ2P++ basis sets[3,7] perform better, slightly overestimating water cluster binding energies (rms of 0.28 and 0.20 kcal/mol/water, respectively, vs MP2/CBS). Addition of full BSSE tends to overcorrect this overbinding by a factor of ~2—adding 50% BSSE to provide a better estimate of the true binding[30] leads to an rms of 0.04 and 0.24 kcal/mol/water, respectively, vs MP2/CBS.

We emphasize that BSSE calculations are expensive, requiring the calculation of N single-point energies with the full system basis, where N is the number of water monomers in the complex. For canonical MP2 with large basis sets, BSSE is still a large fraction of the total binding energy (9% for aug-cc-pVTZ, 8-D_{2d} geometry). In contrast, with X3LYP we find that BSSE is only 0.9% of the total binding energy (aug-cc-pVTZ(-f), 8-D_{2d} geometry), and we observe good correspondence with MP2/CBS energies despite neglecting BSSE.

We could not find any published MP2 calculations on $(H_2O)_n$ clusters with $n = 13-19$ and, hence, cannot compare our fully optimized X3LYP binding energies for these systems. However, Figure 2 shows that the X3LYP binding energy *per hydrogen bond* for the "three-dimensional" ($n > 5$) water clusters oscillates near the experimentally determined binding energy of ice at 0 K ($\Delta E/2 = -5.68$ kcal/mol).[24] On the other hand, the binding energy *per water* is lower than the bulk value by the five to seven "dangling" hydrogen bonds present in the three-dimensional clusters.

In developing X3LYP, a criterion was that turning off correlation for noble gas dimers should lead to a repulsive curve much like in HF theory. Thus, the correlation functional in X3LYP represents the dispersive contributions to binding. This allows us to separate the correlation component of the binding energy from the electrostatic and hydrogen bonding terms. We

Accurate Energies and Structures for Large Water Clusters

J. Phys. Chem. A, Vol. 108, No. 47, 2004 **10521**

TABLE 2: Theoretical and Experimental Results for the Structure of Water Hexamer

	group	year	method	most stable structure (theory) or obsd (expt)
theory	Tsai and Jordan[31]	1993	MP2/aug-cc-pVDZ′	prism
	Laasonen et al.[14]	1993	GGA/plane wave	cyclic
	Kim et al.[32]	1994	MP2/6-31+G(2d,p) vib freq	cage
	Lee et al.[15]	1994	BLYP/TZVP	cyclic
	Estrin et al.[12]	1996	GGA(PW/P)/"moderate" basis	prism
	Liu et al.[33]	1996	model potential/DQMC(nuclei)	cage
	Kim and Kim[34]	1998	MP2/9s6p4d2f1g/6s4p2d + diffuse	cage
	Kryachko[35]	1999	MP2/aug-cc-pVDZ	prism
	Lee et al.[7]	2000	MP2/TZ2P++	book
	Xantheas et al.[3]	2002	MP2/CBS extrapolation	prism
	Losada and Leutwyler[36]	2003	MP2/aug-cc-pVTZ	cyclic
	Present work	2004	X3LYP/aug-cc-pVTZ(-f)	cyclic
expt	Liu et al.[33,37]	1996	terahertz laser vib−rot. tunnel spec	cage
	Nauta and Miller[38]	2000	IR/liquid He droplets	cyclic + book
	Fajardo and Tam[39]	2001	IR/para-hydrogen matrix	cyclic + cage/book

TABLE 3: Comparison of $(H_2O)_n$ Water Cluster Minima (kcal/mol)[a]

		X3LYP/aug-cc-pVTZ(-f)				ΔE (others)		
n	structure	$-\Delta E$	$-\Delta E_{BSSE}$	$-\Delta E_0$	$-\Delta G_{50}$	Xantheas (MP2)[3]	Lee (MP2)[7]	B3LYP
6	prism	44.69	44.24	30.66	7.78	**45.86**	43.97	41.49
	cage	44.78	44.35	30.87	8.02	**45.79**	**44.04**	41.68
	book	**45.17**	**44.88**	**31.68**	9.34	45.61	**44.06**	**42.26**
	bag	44.39	44.08	31.05	8.61		43.37	41.44
	cyclic	**45.04**	**45.02**	**32.23**	**10.35**	44.86	43.48	**42.35**
	cyclic′	44.10	43.99	**31.64**	**10.00**			41.40
8	D_{2d}	71.05	70.43	50.32	16.94	72.57	70.06	66.31
	S_4	71.35	70.58	50.56	17.20	72.56	70.03	66.53
10	prism	91.17	90.26	65.35	22.08		90.07	85.01
	prism′	91.82	91.06	65.91	22.64		89.98	85.84
	butterfly	84.12	83.43	59.68	16.86		87.93	78.29

[a] ΔE and ΔE_{BSSE} correspond to the non-BSSE and BSSE-corrected binding energies, respectively. ΔE_0 is the non-BSSE binding energy with zero-point energy added; ΔG_{50} is evaluated from $\Delta H + T\Delta S$, $T = 50$ K, based on the non-BSSE binding energy and with zero-point energy added. The most stable hexamer structures are indicated in boldface type.

find that the correlation fraction is remarkably consistent, 45−54% of the total binding energy for all water clusters studied (see the Supporting Information for more details).

3.2. Water Cluster Local Minima. *3.2.1. General Discussion.* It is well-established that water trimers through pentamers have cyclic structures, while water clusters larger than hexamer have three-dimensional structures.[40] Among these three-dimensional structures there is some disagreement on the detailed structure of the decamer but not for the octamer, which has a cubic structure.[41,42] As expected, water octamer isomers (D_{2d} and S_4) have similar energies in both X3LYP and MP2 calculations[42,43] (Table 3). Water decamers appear in both X3LYP and MP2 calculations to prefer a pentagonal prism structure over a less symmetric "butterfly" form derived from the cubic octamer. This contrasts with the interpretation of experimental studies that suggest the butterfly form to be the more stable structure.[43]

However, as indicated in Table 2, the structure of water hexamer, intermediate between the two regimes, has been a subject of active debate. We discuss this case in more detail below.

3.2.2. Water Hexamer. The most commonly considered structures are shown in Figure 3, differing in the balance of ring strain against number of hydrogen bonds. Recent theoretical predictions have been ambiguous, with the energy ordering of isomers highly sensitive to basis set size[32] and BSSE inclusion.[31] In addition, methods using a nuclear QMC scheme to calculate zero-point effects have used different model potentials.[44,45]

Experiments have also been ambiguous, with cage structures observed in water clusters formed from supersonic jets[33] and
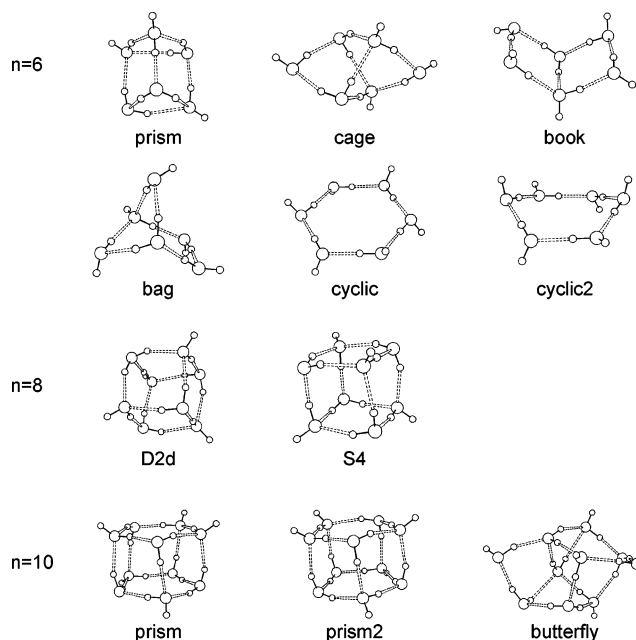

Figure 3. Optimized water cluster minima $(H_2O)_n$; $n = 6, 8, 10$ (X3LYP/aug-cc-VTZ(-f)).

cyclic structures observed in clusters formed in liquid helium droplets[38,46] or solid para-hydrogen matrices.[39]

Our results using X3LYP/aug-cc-pVTZ(-f) indicate that the book and cyclic (chair) structures are the most stable (Table 3, Figure 4). The structures are nearly degenerate (−45.17 and −45.04 kcal/mol, respectively), with an energy ordering that
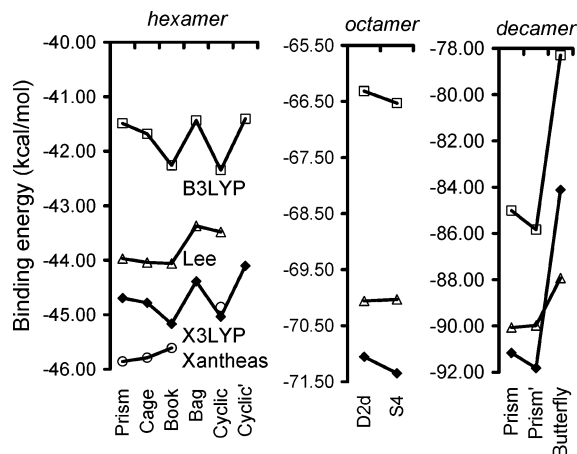
Figure 4. Comparison[3,7] of water cluster minima binding energies (kcal/mol) without BSSE. Negative binding energies are plotted so that the energies of more strongly bound clusters lie at the bottom of the graph.

reverses when BSSE (cyclic now 0.14 kcal/mol more stable), zero-point energy effects (cyclic now 0.55 kcal/mol more stable), or entropic effects (cyclic now 1.01 kcal/mol more stable) are included. We should caution that these zero-point energies and entropic effects are derived using a harmonic normal-mode analysis which may not account for certain "flipping" vibrations in the water hexamer.[36] We find that the cage structure is always less stable and is generally close in energy to the prism.

Our most stable structures (book/cyclic) are different from the most stable structure (prism) predicted with MP2/CBS but are consistent with those observed in the most recent IR/para-hydrogen matrix experiments (book/cyclic). In rationalizing the difference between these experiments and the MP2/CBS results, it has been suggested that the hexamers isolated in para-H2 matrices may represent kinetic and not thermodynamically

favored structures.[39,46] We do not find such an interpretation to be necessary since X3LYP predicts that the book/cyclic structures are the thermodynamically favored structures.

Figure 5 compares the X3LYP results with recent MP2 calculations. With aug-cc-pVTZ(-f), the BSSE error for X3LYP is more than *ten times smaller* than for MP2 methods. X3LYP energies converge quickly to a limiting value with increasing basis set size (Figure 5 and Table 1). For the cyclic and book structures, the X3LYP energies also converge to the MP2 energies in the complete basis set limit; however, for the cage and prism structures, the two methods appear to converge to different energies.

This systematic difference may arise from the fundamental difference in the treatment of electron correlation in MP2 vs X3LYP. Nonetheless, we observe (1) that for the practical triple-ζ basis set the X3LYP energies are well within the uncertainties of similar MP2 calculations and (2) that the B3LYP energies clearly disagree with the MP2 energies, although they follow the same trend as the X3LYP energies.

In the finite basis set description of the hexamer isomers, the X3LYP description of electron correlation is *as consistently valid* as the MP2 perturbative description of electron correlation. Thus the X3LYP cyclic/book geometries are as much "reference" hexamer structures as the MP2 cage geometry currently is considered to be.

3.3. Decomposition of the Total Binding Energy into Multibody Components. It has been estimated that pairwise interactions contribute ∼70% to the total binding energy of water clusters.[47,48] These pairwise interactions are expected to be the ones most sensitive to electron correlation and basis set effects.[47−50] This suggests that one could minimize the computational effort required for high accuracy by using a smaller basis set and lower level of theory to calculate three-body and higher terms and focusing the computation on the two-body terms.[29] To test this idea, Table 4 partitions the binding energy
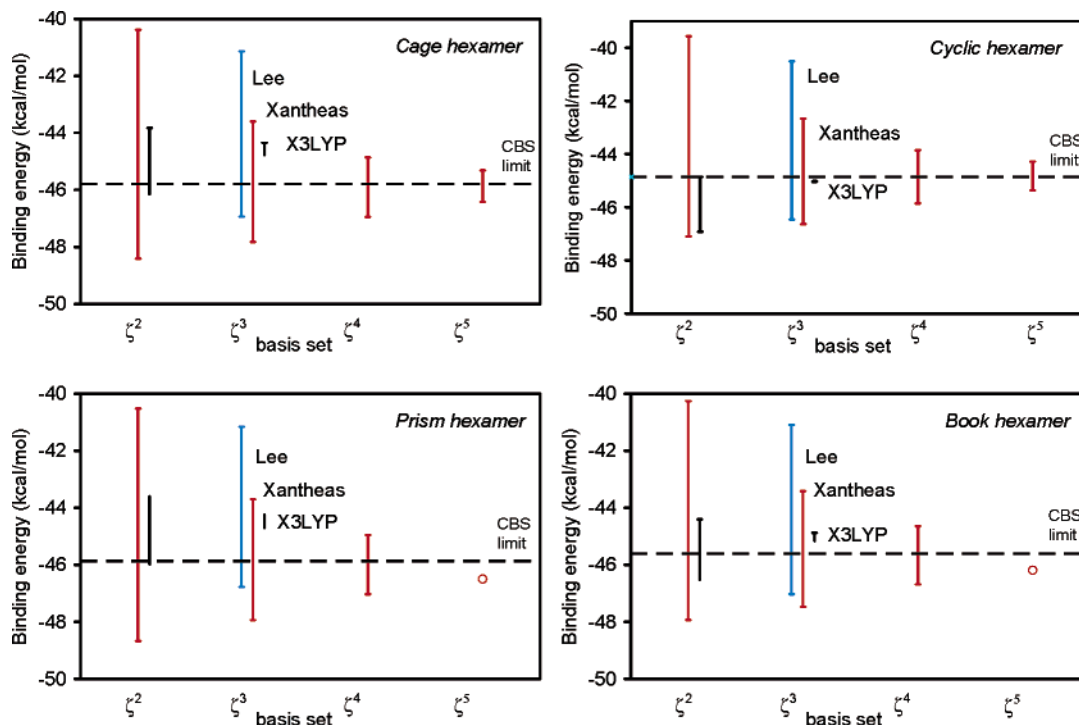
Figure 5. Negative total binding energy as a function of basis set for selected hexamer geometries, with comparison results from Lee[7] and Xantheas.[3] Here the largest basis set on the right and the estimate of the complete basis set (CBS) limit for MP2 is shown with dashes. Lower limits represent non-BSSE energies; upper limits represent BSSE energies. Generally, this lies midway between the BSSE and non-BSSE limits for the finite basis sets. The impact of BSSE for X3LYP is ∼1/10th that for MP2.

Accurate Energies and Structures for Large Water Clusters

J. Phys. Chem. A, Vol. 108, No. 47, 2004 **10523**

TABLE 4: Decomposition of Interaction Energies (kcal/mol) for the Cyclic (S_6) Water Hexamer into Multibody Components[a]

interaction	LMP2			B3LYP			X3LYP			MP2		
	no BSSE	BSSE	50% BSSE	no BSSE	BSSE	50% BSSE	no BSSE	BSSE	50% BSSE	no BSSE	BSSE	50% BSSE
1-body	2.89	4.12	**3.51**	1.89	1.89	**1.89**	1.98	1.96	**1.97**	2.59	1.97	**2.28**
2-body	−30.98	−31.63	**−31.31**	−30.01	−29.98	**−29.99**	−32.79	−32.74	**−32.76**	−34.4	−29.46	**−31.93**
3-body	−6.98	−5.76	**−6.37**	−12.27	−12.08	**−12.17**	−12.19	−12.02	**−12.11**	−11.33	−11.61	**−11.47**
4-body	−0.80	−4.18	**−2.49**	−1.52	−2.01	**−1.76**	−1.60	−2.06	**−1.83**	−1.62	−1.51	**−1.57**
5-body	−0.60	1.35	**0.38**	−0.41	−0.13	**−0.27**	−0.40	−0.13	**−0.27**	−0.62		
6-body	0.11	−0.27	**−0.08**	−0.01	−0.03	**−0.02**	−0.01	−0.03	**−0.02**			
total	−36.36	−36.36	**−36.36**	−42.33	−42.33	**−42.33**	−45.01	−45.01	**−45.01**	−45.38	−40.61	

[a] All geometries were optimized at the level of theory indicated. For LMP2, B3LYP, and X3LYP, aug-cc-pVTZ(-f) single-point energies were calculated from an aug-cc-pVTZ(-f) optimized geometry. For MP2 (results taken from Jordan et al.[29]), aug-cc-pVTZ(-f) energies were calculated from a 6-31+G[2d,p]-optimized geometry. The average (boldfaced) of non-BSSE and BSSE energies is taken to estimate the CBS limit.

into multibody terms, allowing a comparison of the MP2 energy components directly with X3LYP energy components (here we average the non-BSSE and BSSE energies to estimate the CBS limit).

The one-body "monomer relaxation" terms in B3LYP and X3LYP deviate from MP2 by similar amounts (0.39 vs 0.31 kcal/mol, respectively) as do the three-body terms (0.71 vs 0.64 kcal/mol, respectively) and higher. However, B3LYP and X3LYP differ significantly from each other in their two-body terms (1.93 kcal/mol vs 0.83 kcal/mol difference) with X3LYP much closer to MP2. This better description of two-body interactions by X3LYP over B3LYP is expected, since X3LYP also describes water dimer and rare gas dimers much more accurately.

We find that LMP2 has the best description of two-body energies (difference of 0.62 kcal/mol from MP2), but that it fails to reproduce the higher body terms (the three body term is only half the correct value). This probably arises from assumptions in LMP2 about localization of electron correlation that are most valid for pairwise interactions. Hydrogen bonds in the LMP2-optimized cyclic water hexamer are also longer than in the corresponding B3LYP and X3LYP-optimized geometries (1.826 Å vs 1.749 and 1.739 Å, respectively). Thus, LMP2 fails to properly describe water clusters.

It has been reported that B3LYP energies approach MP2/CBS values[29] by a "fortuitous cancellation of terms". However, we find no evidence of this trend. Indeed our results suggest that B3LYP is deficient only in its treatment of two-body interactions. Once this is corrected, as in X3LYP, B3LYP leads to a proper description of larger water clusters.

3.4. Vibrational Frequencies: Theory and Experiment. Vibrational frequencies from theory correspond to force constants at the geometric minimum, while vibrational frequencies from experiment correspond to force constants averaged over the zero-point motions, which are quite large in water clusters. With sufficient experimental data on the vibrational overtones, one can correct for anharmonicity to obtain the harmonic normal-mode vibrational frequencies. However this has been determined only for water monomer[51,52] and water dimer.[53−56] To compare theory and experiment we used the corrections for the monomer and dimer to derive the empirical relation between anharmonic and harmonic vibrational frequencies shown in Figure 6. With this relation, we extrapolated the experimentally determined OH stretching vibrations of larger water clusters to corresponding *harmonic* frequencies. Figure 7 shows a comparison of these harmonic frequencies with our theoretical vibrational frequencies, left unscaled.

For cyclic complexes (dimer to pentamer), the waters are arranged symmetrically leading to a clear distinction between bonded and nonbonded O−H stretches. As the number of waters increases, the bonded OH stretch becomes lower in frequency


$$\omega_{harmonic} = 1.0945\,\omega_{anharmonic} - 168.12\ cm^{-1}$$
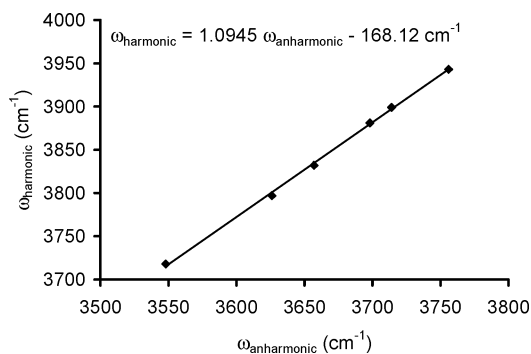
Figure 6. Comparison of experimental harmonic (derived) and anharmonic (measured) O−H stretching frequencies for water monomer and dimer (cm^{-1}). This is used to derive an empirical correction factor to experimental (anharmonic) frequencies for comparison with theoretical (harmonic) frequencies.
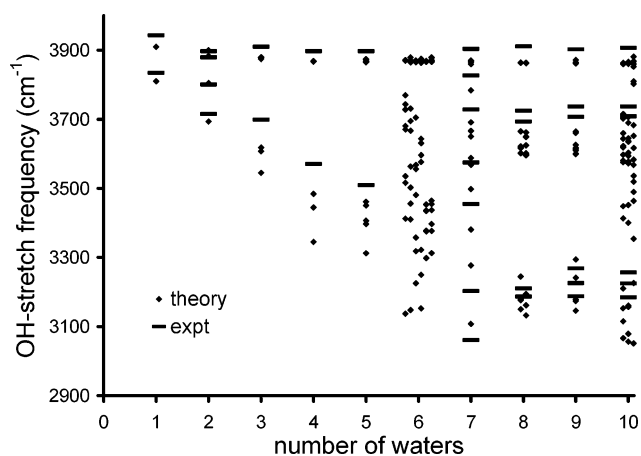
Figure 7. Comparison of O−H stretching frequencies (cm^{-1}), theory (unscaled), and experiment (scaled to obtain the harmonic frequencies). Stretching frequencies for multiple configurations are shown where available: for $n = 6$, prism, cage, book, bag, cyclic, and cyclic′; for $n = 8$, D_{2d} and S_4; and for $n = 10$, prism, prism′, and butterfly.

while the nonbonded OH stretching frequency remains nearly constant. X3LYP clearly reproduces this trend although the overall frequencies are systematically underestimated.

The agreement between theory and experiment for the dimer is good but the monomer agreement is not as close as previously reported.[13] Complexes larger than hexamers are three-dimensional, leading to IR spectra that show a characteristic band structure with a gap between bonded and nonbonded O−H stretches. This band structure and the gap between bands are reproduced well by X3LYP. The OH vibrations from theory and experiment are comparable for all clusters except $n = 6$, consistent with the assignment of cyclic structures to $n \leq 5$ and three-dimensional structures for $n \geq 7$. For $n = 6$ it would
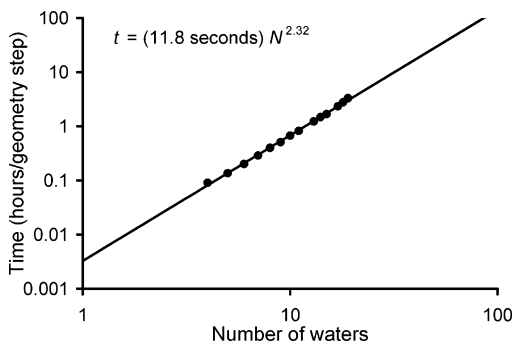
Figure 8. Comparison of computation times (seconds/geometry optimization step) for optimizing the geometry of global minimum water clusters ($n = 4-19$) using Jaguar 5.0 running on one processor of a Pentium 4 Xeon machine (512 K cache), 2.2 GHz, 2 GB memory. The overall computation time scales as $\sim N^{2.3}$.

TABLE 5: Comparison of Computation Times for Optimizing the Geometry of Water Octamer[a]

basis set (basis functions/water)	geometry iteration/s		
	LMP2	B3LYP	X3LYP
6-31g** (25)	1020	154	146
6-311g**++ (36)	4028	475	461
aug-cc-pVTZ(-f) (58)	16035	1572	1380

[a] Computations carried out with Jaguar using one processor of a Pentium 4 Xeon machine (512 K cache), 2.2 GHz, 2 GB memory. These computations all used Jaguar 5.0, which implements pseudo-spectral acceleration.

be valuable to obtain additional vibrational frequencies to check the assignments.

3.5. Benchmark Results and Timing. The cost of carrying out X3LYP calculations is essentially the same as for B3LYP and other hybrid DFT methods, making it quite practical for systems with hundreds of atoms. Figure 8 shows the timings for water cluster calculations for up to 19 waters indicating that the scaling is as $N^{2.3}$. For larger clusters, the scaling may slow to N^3, as initially faster matrix diagonalization and multiply steps become slower and dominate the computation time. Even with this conservative assumption, using 16 processors with a well parallelized DFT code it should be possible to do comparable calculations on clusters up to 50 waters, at an estimated cost of $30-60$ h per geometry step/processor.

In contrast, MP2 calculations are $\sim$100 times slower for the octamer and scale conventionally as $\sim N.^5$ This severe scaling makes canonical MP2 calculations impractical above $8-12$ waters even at national computer centers. Local orbital approximations can accelerate MP2 but as mentioned in section 3.3 may be inaccurate for our application.

Table 5 shows that with the aug-cc-pVTZ(-f) basis set geometry optimization of water octamer with X3LYP/B3LYP is 10 times faster than with LMP2. Canonical MP2 is not implemented in Jaguar, but previous benchmarking studies[57] on systems with a similar number of basis functions indicate that LMP2 (using Jaguar software) is more than 10 times faster than canonical MP2 (using Gaussian software). Thus, X3LYP is expected to be more than 100 times faster than canonical MP2 for geometry optimizations with our given basis set for moderately sized water clusters ($n > 8$).

4. Conclusions

The X3LYP hybrid density functional was designed from first principles to accurately account for the dispersion interactions

of bound clusters while maintaining or improving the accuracy of B3LYP for thermochemistry and other properties. Although water dimer and other water cluster systems were not used in determining the parameters of X3LYP, we find that X3LYP leads to binding energies for water clusters up to 12 waters in excellent agreement (average error in binding energy per water of $\sim$0.1 kcal/mol) with the best theoretical results currently available (MP2/CBS, MP2/TZ2P++).

The accuracy of X3LYP indicates that the DFT description is capable of describing the binding of weakly bound complexes for which dispersion plays an important role.

For the same basis set X3LYP is $\sim$100 times faster than MP2, and these costs scale much more slowly with system size. In addition, the BSSE corrections for X3LYP are $\sim$1/10 that of MP2, allowing BSSE corrections to be neglected even for modest basis sets. This leads to an additional saving in computational cost for high accuracy studies. We tested X3LYP for water clusters here because of the widespread interest in their optimum structures and the availability of high accuracy MP2 calculations for comparison. With X3LYP, we can now extract accurate interaction energies from hydrocarbon clusters and other weakly bound systems, and use those data to create purely ab initio based force fields capable of describing protein–ligand binding, DNA–ligand binding, and macromolecule self-assembly.

The one water cluster for which there remains considerable uncertainty is the water hexamer, which is at the crossover point between small clusters which are cyclic and large clusters which have a cage-like three-dimensional structure. With X3LYP we find that the cyclic (chair) and book forms are particularly stable, which agrees with some recent theoretical and experimental studies, but not with others. We have predicted the vibrational spectrum which may provide a target for experiments to test the predicted structure.

Supporting Information Available: Tables of absolute energies, zero point energies, enthalpies, entropies, and dispersion energies of water clusters; ball-and-stick drawings and coordinates of all water clusters considered; anharmonic and harmonic vibrational frequencies of water monomer and dimer; calculated normal-mode frequencies corresponding to water cluster O–H stretching modes; and experimentally determined vibrational frequencies of water clusters and corresponding harmonic vibrational frequencies estimated through an empirical scaling relation. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Xu, X.; Goddard, W. A., III. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (9), 2673–2677.

(2) Xu, X.; Goddard, W. A., III. Bonding Properties of the Water Dimer: A Comparative Study of Density Functional Theories. *J. Phys. Chem. A* **2004**, *108* (12), 2305−2313.

(3) Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. Development of transferable interaction models for water. II. Accurate energetics of the first few water clusters from first principles. *J. Chem. Phys.* **2002**, *116* (4), 1493−1499.

(4) Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. Structure and Stability of Water Clusters $(H_2O)_n$, $n = 8−20$: An Ab Initio Investigation. *J. Phys. Chem. A* **2001**, *105* (46), 10525−10537.

(5) Lee, H. M.; Suh, S. B.; Kim, K. S. Structures, energies, and vibrational spectra of water undecamer and dodecamer: An ab initio study. *J. Chem. Phys.* **2001**, *114* (24), 10749−10756.

(6) Lee, H. M.; Suh, S. B.; Lee, J. Y.; Tarakeshwar, P.; Kim, K. S. Structures, energies, vibrational spectra, and electronic properties of water monomer to decamer. *J. Chem. Phys.* **2000**, *112*, 9759. Erratum: *J. Chem. Phys.* **2001**, *114* (7), 3343.

(7) Lee, H. M.; Suh, S. B.; Lee, J. Y.; Tarakeshwar, P.; Kim, K. S. Structures, energies, vibrational spectra, and electronic properties of water monomer to decamer. *J. Chem. Phys.* **2000**, *112* (22), 9759−9772.

(8) Xantheas, S. S.; Apra, E. The binding energies of the D_{2d} and S_4 water octamer isomers: High-level electronic structure and empirical potential results. *J. Chem. Phys.* **2004**, *120* (2), 823−828.

(9) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37* (2), 785−9.

(10) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58* (8), 1200−11.

(11) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623−7.

(12) Estrin, D. A.; Paglieri, L.; Corongiu, G.; Clementi, E. Small Clusters of Water Molecules Using Density Functional Theory. *J. Phys. Chem.* **1996**, *100* (21), 8701−11.

(13) Kim, K.; Jordan, K. D. Comparison of Density Functional and MP2 Calculations on the Water Monomer and Dimer. *J. Phys. Chem.* **1994**, *98* (40), 10089−94.

(14) Laasonen, K.; Parrinello, M.; Car, R.; Lee, C.; Vanderbilt, D. Structures of small water clusters using gradient-corrected density functional theory. *Chem. Phys. Lett.* **1993**, *207* (2−3), 208−13.

(15) Lee, C.; Chen, H.; Fitzgerald, G. Structures of the water hexamer using density functional methods. *J. Chem. Phys.* **1994**, *101* (5), 4472−3.

(16) Kutzelnigg, W.; Klopper, W. Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory. *J. Chem. Phys.* **1991**, *94* (3), 1985−2001.

(17) Halkier, A.; Klopper, W.; Helgaker, T.; Jorgensen, P.; Taylor, P. R. Basis set convergence of the interaction energy of hydrogen-bonded complexes. *J. Chem. Phys.* **1999**, *111* (20), 9157−9167.

(18) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−52.

(19) *Jaguar 5.0*; Schrodinger, L.L.C: Portland, OR, 1991−2003.

(20) Saebo, S.; Tong, W.; Pulay, P. Efficient elimination of basis-set-superposition errors by the local correlation method: accurate ab initio studies of the water dimer. *J. Chem. Phys.* **1993**, *98* (3), 2170−5.

(21) Saebo, S.; Pulay, P. Local treatment of electron correlation. *Ann. Rev. Phys. Chem.* **1993**, *44*, 213−36.

(22) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90* (9), 4916−26.

(23) Hariharan, P. C.; Pople, J. A. Effect of d-functions on molecular orbital energies for hydrocarbons. *Chem. Phys. Lett.* **1972**, *16* (2), 217−19.

(24) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96* (9), 6796−806.

(25) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19* (4), 553−566.

(26) Xantheas, S. S. On the importance of the fragment relaxation energy terms in the estimation of the basis set superposition error correction to the intermolecular interaction energy. *J. Chem. Phys.* **1996**, *104* (21), 8821−8824.

(27) Szalewicz, K.; Jeziorski, B. Comment on "On the importance of the fragment relaxation energy terms in the estimation of the basis set superposition error correction to the intermolecular interaction energy" [*J. Chem. Phys.* **1996**, *104*, 8821.]. *J. Chem. Phys.* **1998**, *109* (3), 1198−1200.

(28) Wales, D. J.; Hodges, M. P. Global minima of water clusters $(H_2O)_n$, $n < 21$, described by an empirical potential. *Chem. Phys. Lett.* **1998**, *286* (1, 2), 65−72.

(29) Pedulla, J. M.; Vila, F.; Jordan, K. D., Binding energy of the ring form of $(H_2O)_6$: comparison of the predictions of conventional and localized-orbital MP2 calculations. *J. Chem. Phys.* **1996**, *105* (24), 11091−11099.

(30) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. Molecular Clusters of Pi-Systems: Theoretical Studies of Structures, Spectra, and Origin of Interaction Energies. *Chem. Rev.* **2000**, *100* (11), 4145−4185.

(31) Tsai, C. J.; Jordan, K. D., Theoretical study of the $(H_2O)_6$ water cluster. *Chem. Phys. Lett.* **1993**, *213* (1−2), 181−8.

(32) Kim, K.; Jordan, K. D.; Zwier, T. S., Low-Energy Structures and Vibrational Frequencies of the Water Hexamer: Comparison with Benzene-$(H_2O)_6$. *J. Am. Chem. Soc.* **1994**, *116* (25), 11568−9.

(33) Liu, K.; Brown, M. G.; Carter, C.; Saykally, R. J.; Gregory, J. K.; Clary, D. C. Characterization of a cage form of the water hexamer. *Nature (London)* **1996**, *381* (6582), 501−503.

(34) Kim, J.; Kim, K. S. Structures, binding energies, and spectra of isoenergetic water hexamer clusters: Extensive ab initio studies. *J. Chem. Phys.* **1998**, *109* (14), 5886−5895.

(35) Kryachko, E. S. Ab initio studies of the conformations of water hexamer: modeling the pentacoordinated hydrogen-bonded pattern in liquid water. *Chem. Phys. Lett.* **1999**, *314* (3, 4), 353−363.

(36) Losada, M.; Leutwyler, S. Water hexamer clusters: Structures, energies, and predicted mid-infrared spectra. *J. Chem. Phys.* **2002**, *117* (5), 2003−2016.

(37) Liu, K.; Brown, M. G.; Saykally, R. J. Terahertz Laser Vibration−Rotation Tunneling Spectroscopy and Dipole Moment of a Cage Form of the Water Hexamer. *J. Phys. Chem. A* **1997**, *101* (48), 8995−9010.

(38) Nauta, K.; Miller, R. E. Formation of cyclic water hexamer in liquid helium: the smallest piece of ice. *Science (Washington, D.C.)* **2000**, *287* (5451), 293−295.

(39) Fajardo, M. E.; Tam, S. Observation of the cyclic water hexamer in solid parahydrogen. *J. Chem. Phys.* **2001**, *115* (15), 6807−6810.

(40) Keutsch, F. N.; Saykally, R. J. Water clusters: untangling the mysteries of the liquid, one molecule at a time. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, (19), 10533−10540.

(41) Sadlej, J.; Buch, V.; Kazimirski, J. K.; Buck, U. Theoretical study of structure and spectra of cage clusters $(H_2O)_n$, $n = 7−10$. *J. Phys. Chem. A* **1999**, *103* (25), 4933−4947.

(42) Gruenloh, C. J.; Carney, J. R.; Arrington, C. A.; Zwier, T. S.; Fredericks, S. Y.; Jordan, K. D. Infrared spectrum of a molecular ice cube: the S_4 and D_{2d} water octamers in benzene-(water)8. *Science (Washington, D.C.)* **1997**, *276* (5319), 1678−1681.

(43) Buck, U.; Ettischer, I.; Melzer, M.; Buch, V.; Sadlej, J. Structure and spectra of three-dimensional $(H_2O)_n$ clusters, n=8, 9, 10. *Phys. Rev. Lett.* **1998**, *80*, (12), 2578−2581.

(44) Gregory, J. K.; Clary, D. C. A Theoretical Study of the Cage Water Hexamer Structure. *J. Phys. Chem. A* **1997**, *101* (36), 6813−6819.

(45) Severson, M. W.; Buch, V. Quantum Monte Carlo simulation of intermolecular excited vibrational states in the cage water hexamer. *J. Chem. Phys.* **1999**, *111* (24), 10866−10875.

(46) Burnham, C. J.; Xantheas, S. S.; Miller, M. A.; Applegate, B. E.; Miller, R. E. The formation of cyclic water complexes by sequential ring insertion: Experiment and theory. *J. Chem. Phys.* **2002**, *117* (3), 1109−1122.

(47) Masella, M.; Flament, J. P. A pairwise and two many-body models for water: influence of nonpairwise effects upon the stability and geometry of $(H_2O)_n$ cyclic ($n = 3−6$) and cagelike ($n = 6−20$) clusters. *J. Chem. Phys.* **1997**, *107* (21), 9105−9116.

(48) Hodges, M. P.; Stone, A. J.; Xantheas, S. S. Contribution of Many-Body Terms to the Energy for Small Water Clusters: A Comparison of ab Initio Calculations and Accurate Model Potentials. *J. Phys. Chem. A* **1997**, *101* (48), 9163−9168.

(49) Pedulla, J. M.; Kim, K.; Jordan, K. D. Theoretical study of the n-body interaction energies of the ring, cage and prism forms of $(H_2O)_6$. *Chem. Phys. Lett.* **1998**, *291* (1, 2), 78−84.

(50) Xantheas, S. S. Ab initio studies of cyclic water clusters $(H_2O)_n$, n=1−6. II. Analysis of many-body interactions. *J. Chem. Phys.* **1994**, *100* (10), 7523−34.

(51) Benedict, W. S.; Gailar, N.; Plyler, E. K. Rotation−vibration spectra of deuterated water vapor. *J. Chem. Phys.* **1956**, *24*, 1139−65.

(52) Kuchitsu, K.; Morino, Y. Estimation of anharmonic potential constants. II. Bent XY2 molecules. *Bull. Chem. Soc. Jpn.* **1965**, *38* (5), 814−24.

(53) Fredin, L.; Nelander, B.; Ribbegard, G. Infrared spectrum of the water dimer in solid nitrogen. I. Assignment and force constant calculations. *J. Chem. Phys.* **1977**, *66* (9), 4065−72.

(54) Nelander, B. Infrared spectrum of the water-hydrogen sulfide complex. *J. Chem. Phys.* **1978**, *69*, (8), 3670−1.

(55) Bentwood, R. M.; Barnes, A. J.; Orville-Thomas, W. J. Studies of intermolecular interactions by matrix isolation vibrational spectroscopy. Self-association of water. *J. Mol. Spectrosc.* **1980**, *84* (2), 391−404.

(56) Tursi, A. J.; Nixon, E. R. Matrix-isolation study of the water dimer in solid nitrogen. *J. Chem. Phys.* **1970**, *52* (3), 1521−8.

(57) Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. Pseudospectral localized Moeller−Plesset methods: theory and calculation of conformational energies. *J. Chem. Phys.* **1995**, *103* (4), 1481−90.