Thermodynamics of Hydrogen Bonding from Molecular Orbital Theory: 1. Water

Jeffrey P. Wolbach and Stanley I. Sandler
Center for Molecular and Engineering Thermodynamics, Dept. of Chemical Engineering,
University of Delaware, Newark, DE 19716

To reduce the number of adjustable parameters that appear in the thermodynamic modeling of hydrogen-bonding fluids, the use of ab-initio molecular orbital calculations in the form of Hartree–Fock and density functional theories to compute thermodynamic properties are considered. In particular, the enthalpy and entropy changes on hydrogen bonding for water are computed. The theory, the computational methods and the basis sets used, as well as the results for water, are discussed. The results obtained are compared with those of others and with experimental estimates from the literature to determine the validity of the methods that are then used in this extended study. The results obtained here are also used to examine the validity of mixture group contribution methods for hydrogen-bonding mixtures, to develop simple estimation procedures for model parameters, and to reduce the number of calculations such as those described that need to be done.

Introduction

A challenging problem in thermodynamics is the accurate modeling of the phase behavior of hydrogen-bonding compounds. This is often done imperfectly with equations of state that have a number of adjustable parameters with either no well-defined physical meaning, as in equations of state derived from chemical theories of hydrogen bonding, or that cannot be determined independently, as in equations of state derived from physical theories of hydrogen bonding. The thermodynamic description of hydrogen bonding would be improved if one were able to independently determine the thermodynamic parameters of hydrogen-bonding "reactions." Knowledge of the types of association and their reaction parameters would allow one to assign association schemes that better represent the physics of a particular system, without adding adjustable parameters to an equation of state.

In principle, one could attempt to determine the association parameters by measuring them experimentally. However, these parameters cannot be measured directly so that reported experimental results depend on the model used to interpret the data. An alternate way of determining the thermodynamic parameters is through the use of *ab initio* molecular orbital calculations (Frisch et al., 1986; Del Bene, 1987; Zheng and Merz, 1992). This method permits the direct calculation of the quantities of interest, and also allows one to investigate particular reactions (e.g., monomer + linear dimer → cyclic trimer). Unfortunately, there is no universally accepted "correct" method for performing such calculations.

Here and in the companion article we use *ab initio* molecular orbital calculations to perform an extensive study of hydrogen-bonded dimers ranging in size from $(H_2O)_2$ to $(C_2H_5COOH)_2$, and including both pure and cross dimers. We wanted to perform all of the calculations using a single method and basis set in order to make comparisons of the results meaningful. This required the use of a method and basis set that were computationally feasible not only for the water dimer, a 20-electron system, but also for the propionic acid dimer, an 80-electron system. This requirement eliminated from consideration the most rigorous *ab-initio* methods available.

We investigated two methods that fulfilled this requirement. The first, less computationally intensive method was that of Hartree (1928a,b) and Fock (1930), and the second was density functional theory in the form of Becke's three-parameter exchange functional (1992a,b, 1993) with the nonlocal correlation functional of Lee et al. (1988). We first pro-

---

Correspondence concerning this article should be addressed to S. I. Sandler.

vide a brief explanation of each calculational method and procedure, and then present the results obtained by applying these methods to the water monomer and linear water dimer, the systems most extensively studied in the literature. These results are compared to experimental values, as well as to the reported results of higher-level molecular orbital calculations. We will show that the relatively low-level methods we use describe these systems as accurately as the more computationally intensive methods, providing confidence in our methods for the study of molecules larger than water. The results of our calculations using these methods for a variety of pure and cross dimers are presented in Part 2.

Theory

The starting point for *ab initio* molecular orbital calculations is the time-independent Schrödinger equation. This may be simplified by using the Born–Oppenheimer (1927) approximation, which states that since the nuclei move much more slowly than do the electrons, one needs to solve only for the electron motion around nuclei at fixed positions. This leads to the electronic Schrödinger equation

$$\hat{H}_e \psi_e(\boldsymbol{r};\boldsymbol{R}) = E^{\text{elec}}(\boldsymbol{R})\psi_e(\boldsymbol{r};\boldsymbol{R}). \tag{1}$$

In Eq. 1, E^{elec} is the electronic energy of the molecule, $\hat{H}_e$ is the electronic Hamiltonian, and ψ is the electronic wave function. The vector $\boldsymbol{r}$ refers to the coordinates of the electrons, and $\boldsymbol{R}$ refers to the coordinates of the nuclei. From this point forward, the electronic wave function, $\psi_e(\boldsymbol{r};\boldsymbol{R})$, will be abbreviated as $\psi_e(\boldsymbol{r})$ due to the assumption of fixed nuclear positions. The electronic Hamiltonian is given by

$$\hat{H}_e = \frac{-\hbar^2}{2}\sum_i^{\text{electrons}} \frac{1}{m_i}\nabla_i^2 - \sum_i^{\text{electrons}}\sum_I^{\text{nuclei}}\left(\frac{Z_I e'^2}{|\boldsymbol{R}_I - \boldsymbol{r}_i|}\right) + \sum_i^{\text{electrons}}\sum_{j<i}\left(\frac{e'^2}{|\boldsymbol{r}_j - \boldsymbol{r}_i|}\right) + \sum_I^{\text{nuclei}}\sum_{J<I}\left(\frac{Z_I Z_J e'^2}{|\boldsymbol{R}_J - \boldsymbol{R}_I|}\right), \tag{2}$$

where $\hbar = h/2\pi$, and $e'^2 = e^2/4\pi\epsilon_0$, with e being the fundamental electron charge (1.6202×10^{-19} C) and $\hbar$ being Planck's constant (6.626×10^{-34} J·s). The first term in Eq. 2 is the kinetic energy of electronic motion, the second is the attractive potential energy between the electrons and the fixed nuclei, and the fourth term is the repulsive potential between the nuclei. The third term in Eq. 2 is the repulsive potential between electrons, and results in correlation of the electron motion. It is due to this correlation of motion that the electronic Schrödinger equation cannot be solved exactly except in trivial cases. Thus some approximate method is required to solve Eq. 2. The first approximate method used in this work was the Hartree–Fock theory.

Hartree – Fock theory

In the Hartree–Fock (H–F) theory, one assumes that the n-electron wave function can be approximated as a product of n one-electron functions or orbitals. To obtain a physically correct wave function, one must use the determinant of a matrix of one-electron orbitals that are themselves products of spatial orbitals (ϕ_i) and spin functions (α or β). This form of the wave function is termed a Slater determinant (Slater, 1960, 1963).

$$\tilde{\psi}_e(\boldsymbol{r}) = \frac{1}{\sqrt{n!}}\begin{vmatrix} \phi_1(\boldsymbol{r}_1)\alpha(1) & \phi_1(\boldsymbol{r}_1)\beta(1) & \cdots & \phi_{n/2}(\boldsymbol{r}_1)\alpha(1) & \phi_{n/2}(\boldsymbol{r}_1)\beta(1) \\ \phi_1(\boldsymbol{r}_2)\alpha(2) & \phi_1(\boldsymbol{r}_2)\beta(2) & \cdots & \phi_{n/2}(\boldsymbol{r}_2)\alpha(2) & \phi_{n/2}(\boldsymbol{r}_2)\beta(2) \\ \vdots & \vdots & & \vdots & \vdots \\ \phi_1(\boldsymbol{r}_n)\alpha(n) & \phi_1(\boldsymbol{r}_n)\beta(n) & \cdots & \phi_{n/2}(\boldsymbol{r}_n)\alpha(n) & \phi_{n/2}(\boldsymbol{r}_n)\beta(n) \end{vmatrix}. \tag{3}$$

Each electron will possess a spin of either $+1/2$ or $-1/2$, and the spin functions α and β are defined as delta functions for each spin:

$$\text{if spin electron } j = +\frac{1}{2} \rightarrow \begin{cases} \alpha(j) = 1 \\ \beta(j) = 0 \end{cases}$$

$$\text{if spin electron } j = -\frac{1}{2} \rightarrow \begin{cases} \alpha(j) = 0 \\ \beta(j) = 1. \end{cases} \tag{4}$$

The orbitals are then varied so as to minimize the expectation value of the electronic Hamiltonian, or equivalently, the electronic energy:

$$E^{\text{elec}} = \langle \tilde{\psi}_e | \hat{H}_e | \tilde{\psi}_e \rangle = \frac{\int \tilde{\psi}_e^* \hat{H}_e \tilde{\psi}_e d\tau}{\int \tilde{\psi}_e^* \tilde{\psi}_e d\tau}, \tag{5}$$

where the * indicates complex conjugation. Since the spin on an electron cannot change, only the spatial orbitals are variable when optimizing the wave function.

The procedure for determining the functions that minimize Eq. 5 was introduced by Hartree (1928a,b) and is called the Hartree self-consistent field method (SCF). The method begins by focusing on electron 1 and considering electrons 2 through n to be smeared out to form a static distribution of charge through which electron 1 moves. This corresponds to an averaging of the instantaneous interactions, and allows one to approximate the third term in the electronic Hamiltonian (Eq. 2), the interelectron potential.

The averaging of the interactions removes the correlation of the electron motion, and enables one to write a one-electron eigenvalue problem for the spatial orbitals ϕ_i

$$\hat{F}\phi_i(1) = \epsilon_i \phi_i(1). \tag{6}$$

In Eq. 6, $\hat{F}$ is the Fock operator, and is a function only of the position of electron 1. Solution of Eq. 6 provides the optimized orbitals needed to approximate the wave function in Eq. 3. Further details on both the H–F theory and the solution procedure may be found in Levine (1991) and Parr (1963).

While H–F calculations are the least computationally intensive *ab initio* calculations available, the neglect of electron correlation introduces a potentially large error, as it is believed that electron correlation is an important contributor to the interaction energies of nonbonded complexes. Due to our concern over this possible error, we next considered a calculational method that included some degree of electron correlation without requiring a great increase in computational expense. Previously published studies of the water dimer (Kim and Jordan, 1994; Combariza and Kestner, 1995) using density functional theory with Becke's family of three-parameter functionals, produced reasonable results for geometries, harmonic vibrational frequencies, and energetics when compared to the measured quantities. We next present a brief overview of density functional theory.

Density functional theory

The electronic wave function of an n-electron molecule depends on $3n$ spatial and n spin coordinates. However, the electronic Hamiltonian (Eq. 2) contains only one- and two-electron spatial terms, thus the n-electron wave function contains much more information than is needed for the calculation of thermodynamic properties. In theory, the energy of the molecule could be determined using integrals involving only six spatial coordinates. This led to a search for functions involving fewer variables than in the wave function that could be used to calculate the energy and other properties of molecules.

Hohenberg and Kohn (1964) showed that the ground-state energy, the wave function, and many other properties of a molecule could be determined from a knowledge of the electron probability density, $\rho(x, y, z)$, which is a function of only three variables. Kohn and Sham (1965) later provided the equation for the ground-state electronic energy, E_0^{elec}, of an n-electron molecule

$$E_0^{\text{elec}} = -\frac{\hbar^2}{2}\sum_{i=1}^{n}\langle \xi_i(1)|\nabla_1^2|\xi_i(1)\rangle - \sum_I \int \frac{Z_I e' \rho(1)}{|\boldsymbol{R}_I - \boldsymbol{r}_1|} d\boldsymbol{r}_1 + \frac{1}{2}\iint \frac{\rho(1)\rho(2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} d\boldsymbol{r}_1 d\boldsymbol{r}_2 + E_{xc}[\rho], \quad (7)$$

where the ξ_i are the one-electron Kohn–Sham orbitals, and the exact ground-state electron density is found from the one-electron orbitals using

$$\rho(k) = \sum_{i=1}^{n} |\xi_i(k)|^2. \quad (8)$$

The first term in Eq. 7 is the electronic kinetic energy, the second is the attractive nuclear-electron potential energy, and the third is the noncorrelated repulsive electron–electron potential energy. The last term, $E_{xc}[\rho]$, is the exchange-correlation energy in which $[\rho]$ indicates a functional of the electron density. Similar to H–F theory, density functional theory replaces an n-electron problem with n one-electron problems. For specificity, the notation $\xi_i(1)$ and $\rho(1)$ indicate that ξ_i and ρ are to be taken as functions of the spatial coordinates of electron 1.

Similarly to the H–F procedure, the Kohn–Sham orbitals needed to solve Eqs. 7 and 8 can be found by solving a series of one-electron eigenvalue problems

$$\hat{F}_{ks}(1)\xi_i(1) = \epsilon_{i,ks}\xi_i(1), \quad (9)$$

where $\hat{F}_{ks}$ is the one-electron Kohn–Sham operator. Further details are found in Levine (1991).

Density functional theory, using the one-electron orbitals defined in Eqs. 7–9, would give the exact ground-state energy if $E_{xc}[\rho]$ were known. However, this is not the case, and all DFT calculations involve approximating $E_{xc}[\rho]$. The approximation for $E_{xc}[\rho]$ we have chosen is based on the three-parameter functional of Becke (1993). The functional contains three semiempirical constants, and has been shown to provide accurate molecular geometries, energies, and vibrational frequencies, as well as being able to accurately describe nonelectron conserving processes, such as ionization. The semiempirical constants were determined by Becke (1993) using a linear least-squares fit to 56 atomization energies, 42 ionization potentials, eight proton affinities, and ten first-row total atomic energies. The resulting parameters yielded an average absolute error in atomization energies of 2.6%, and good predictions for ionization potentials and proton affinities.

We do not use the exact functional proposed in Becke (1993), replacing one component of the functional, the gradient correction to the correlation energy of Perdew and Wang (Perdew, 1991; Perdew and Wang, 1992; Perdew et al., 1992) with that of Lee et al. (1988). This is because the exact functional proposed by Becke is not available in Gaussian92/DFT (Gaussian, Inc., 1993), the commercial program used for most of the calculations to be described in this and the companion article. From this point forward, the functional we used will be abbreviated as B3LYP.

Procedure

In most commercially available computational chemistry programs, spatial orbitals in Eqs. 6 and 9 are taken to be linear combinations of a predefined set of one-electron functions, the basis functions. This is done to simplify the minimization procedure. An individual spatial orbital is defined as

$$\phi_i(\boldsymbol{r}) = \sum_{\mu=1}^{M} c_{\mu i}\chi_\mu(\boldsymbol{r}), \quad (10)$$

where the variable coefficients $c_{\mu i}$ are referred to as the spatial orbital expansion coefficients, and M is the number of basis functions in each spatial orbital. The basis functions $\chi_1 \cdots \chi_M$ are normalized and fixed. The collection of basis functions is known as the basis set. The use of basis functions changes the problem from one of optimizing the arbitrary normalized functions ϕ_i to the simpler one of optimizing the coefficients $c_{\mu i}$. The more basis functions assigned to a particular molecule, the more flexible the spatial orbitals will become, and the more accurate will be the resulting solution.

The basis set used for the H–F calculations in this work is known in the Gaussian 9X series of programs (Gaussian, Inc., 1993, 1995) as 6-31*g*(*d,p*) (Ditchfield et al., 1971; Hehre et al., 1972; Hariharan and Pople, 1973, 1974; Gordon, 1980). This basis set assigns 5 basis functions to each hydrogen and 15 basis functions to each first-row atom such as oxygen or carbon. This is considered to be a relatively small basis set. For the B3LYP calculations, we used the basis set 6-31++*g*(2*d,p*) (Ditchfield et al., 1971; Hehre et al., 1972; Hariharan and Pople, 1973, 1974; Gordon, 1980; Clark et al., 1983; Frisch et al., 1984), which assigns 6 basis functions to each hydrogen and 25 to each first-row atom, and is considered to be a medium-sized basis set. Explanation of the basis-set nomenclature may be found in Forseman and Frisch (1992).

The calculational procedure below that we used in the H–F calculations is known to give reasonable results for hydrogen-bonding systems (Zheng and Merz, 1992; Ferguson and Kollman, 1991; Dill et al., 1975; Del Bene, 1987; Frisch et al., 1986):

1. The geometry of monomeric species was optimized to the default convergence criterion of the Gaussian program.
2. The output from step 1 served as the input to a second geometric optimization utilizing the "tight" convergence criterion in which each convergence criterion was lowered by a factor of 30 from its default value.
3. A harmonic frequency calculation was performed using the optimized geometry of step 2. This step was also used to calculate the rotational constants and the energy of the molecule. This energy is designated $E_1^1(\text{mon})$.
4. A possible dimer was chosen, and its geometry optimized as in steps 1 and 2. Since we were interested in calculating the vibrational frequencies of the dimer, we allowed the monomers to deform.
5. The frequencies, rotational constants, and energy of the dimer were calculated as in step 3. This energy is designated $E_2^2(\text{dim})$.
6. The change in self-consistent field energy (ΔE^{SCF}) was then calculated as

$$\Delta E^{SCF} = E_2^2(\text{dim}) - \Sigma E_1^1(\text{mon}). \tag{11}$$

This is the energy change upon dimerization at 0 K, and is uncorrected for zero-point vibration.

7. Vibrational frequencies calculated in steps 2 and 4 were then used to correct the value of ΔE^{SCF} for the zero-point vibrational energy (ZPE). However, before doing so, all frequencies were scaled by a factor of 0.8929. This factor was found to improve the agreement between the H–F calculated and experimentally observed frequencies (Pople et al., 1981).
8. The calculated rotational constants and the scaled frequencies were then used to calculate the entropies of the monomers and dimer, and also to provide thermal corrections to the energy. This was done by the use of standard statistical mechanical formulas, assuming the monomers and dimer were ideal polyatomic gases (McQuarrie, 1976). The formulas for the calculation of the thermodynamic quantities of interest here as

$$\Delta S(T) = S(T,\text{dimer}) - \Sigma S(T,\text{mon}) \tag{12}$$

$$\begin{aligned}\Delta U(T) = \Delta E^{SCF} &+ E^{rot}(T,\text{dimer}) + E^{trans}(T,\text{dimer})\\ &+ E^{vib}(T,\text{dimer}) + ZPE(\text{dimer}) - \Sigma ZPE(\text{mon})\\ &- \Sigma[E^{rot}(T,\text{mon}) + E^{trans}(T,\text{mon}) + E^{vib}(T,\text{mon})]\end{aligned} \tag{13}$$

and

$$\begin{aligned}\Delta H(T) &= \Delta U(T) + \Delta(PV) = \Delta U(T) + \Delta(nRT)\\ &= \Delta U(T) - RT.\end{aligned} \tag{14}$$

While this method is not computationally intensive and has been shown to yield quite reasonable results, it contains several important errors:

1. Neglect of electron correlation. The H–F method does not account for electron correlation, and this leads to an underestimation of the magnitude of ΔE^{SCF}.
2. Basis set superposition error (BSSE). This error occurs when performing "supermolecular" calculations on nonbonded complexes, as we have done here, and results in an artificial increase in the magnitude of the interaction energy, ΔE^{SCF}. The reason for this error is an artificial lowering of the energy for each monomer fragment in the dimer due to the presence of the basis functions of the other monomer; that is, a larger basis set is used for the dimer than for the monomers, and in H–F theory, the more basis functions assigned to a molecule, the lower will be its energy. Therefore direct use of Eq. 11 compares energies calculated in different basis sets. A better procedure is to compare energies calculated using the same basis set for both the monomer and dimer. We chose not to correct for the BSSE at the H–F level, since doing so in conjunction with the neglect of electron correlation could lead to a severe underestimation of the interaction energies, and also because we wanted to keep the initial calculational procedure as simple as possible.
3. Distortion energy error. Since the calculations here require the intramolecular vibrational frequencies of the dimer, we chose not to use the "frozen-monomer" technique for dimer optimization. This introduces an error related to the energy difference between the isolated monomer and the monomer in the geometry of the dimer. This error is usually small compared to the basis-set superposition error, and results in an overestimation of the magnitude of ΔE^{SCF}.
4. The scaling factor for the vibrational frequencies we used was reported elsewhere based on the results of a study of selected monomers using the smaller 3-21*g*(*d*) basis set (Pople et al., 1981). This scaling factor may not be appropriate for the larger 6-31*g*(*d,p*) basis set; however, it is used throughout the computational chemistry literature (Forseman and Frisch, 1993; Pople et al., 1989; Feller and Feyereisen, 1993). It is not possible to quantify the effect of this error on the calculated quantities.

We attempted to correct for these errors when performing the B3LYP calculations. Our procedure was identical to the H–F procedure with two exceptions. First, we did not scale the vibrational frequencies of the monomers or the dimer. Second, to account for the basis-set superposition error, a counterpoise correction was applied (Boys and Bernardi, 1970) by calculating the energy of the monomers in the dimer basis set using the geometry they assume in the dimer. This required taking the optimized dimer geometry and removing the atoms of one monomer, while leaving behind their basis functions. This procedure was applied to both monomers in the dimer, and these energies are designated as $E_2^2(\text{mon})$. To calculate the distortion energy, we took the optimized geometry of the dimer and eliminated both the atoms and basis functions of one of the monomers. The molecular energy was then calculated. This procedure was applied to both monomers in the dimer. These energies are designated as $E_2^1(\text{mon})$. This led to the following expression for the basis-set superposition error corrected self-consistent field energy:

$$\Delta E^{BSSE} = E_2^2(\text{dim}) - \Sigma E_2^2(\text{mon}) + \Sigma[E_2^1(\text{mon}) - E_1^1(\text{mon})] \quad (15)$$

The thermodynamic quantities of interest were then calculated as in step 8 of the H–F procedure, with ΔE^{BSSE} replacing ΔE^{SCF}. No attempt was made to correct the vibrational frequencies for the basis-set superposition error.

All of the B3LYP calculations were performed with the int=finegrid option in the Gaussian 9X package, which is used to select a larger number of points per atom over which to perform the numerical integration of the E_{xc} term, and the counterpoise and distortion energy correction calculations used the option SCF=tight, which lowers the convergence criterion for the SCF energy to that used in the frequency calculations.

Even with the corrections in the B3LYP procedure there still exist potential errors.

1. Basis-set superposition error (BSSE). Even though we used a counterpoise correction, BSSE can still result in errors in the calculation of the interaction energy. This is because one effect of BSSE is to change the shape of the potential surface from that used in the minimization. The result of this is that the "optimized" geometry may not be at the minimum of the true potential surface, but merely close to it. As such, the reported ΔE^{BSSE} could be viewed as a lower bound on the magnitude of the true ΔE.

We do not expect the effect of this error to be large, as the magnitude of the BSSE is not large. The sum of the distortion energy and the BSSE is less than 10% of the interaction energy for the water dimer at the B3LYP level of calculation. We also found (but do not report here) that these BSSE nearly equal the errors incurred in the H–F calculations using the same basis set. This is a useful result, as it is known that the BSSE for H–F calculations tends to zero with much smaller basis sets than in MP2 (second-order many-body perturbation theory) calculations (Hobza and Zahradnik, 1988), the method that has been found to be most nearly equivalent to the B3LYP calculations (Kim and Jordan, 1994). [For comparison, if these corrections were applied to our H–F/6-31$g(d,p)$ results, the errors would be close to 30% of the interaction energy, due primarily to the smaller basis set used.]

2. Errors resulting from having to numerically integrate the exchange-correlation energy in the B3LYP functional. It is unknown what effect this would have on the results, but since the int=finegrid option was used throughout this work, it is believed that any effect would be minimal.

3. Utilization of the "wrong" constants in the B3LYP functional. We used the constants published by Becke (1993) for his three-parameter functional that were determined using a computational procedure to avoid basis-set incompleteness. While these constants should have a basis-set dependence, no effort was made to reoptimize the coefficients. A minimal error may also occur due to using the Lee–Yang–Parr gradient-corrected correlation functional rather than the Perdew–Wang functional.

Results

Our goal was to determine whether our procedures and methods are sufficient to accurately describe the water monomer and linear dimer. We first compare the results of our calculations for the water monomer to experimentally determined quantities, and then to the results of previously published higher-level *ab-initio* calculations. The higher-level calculations were performed by others using larger basis sets and/or more theoretically rigorous calculational methods. The higher-level method used was MP2, which accounts for electron correlation by a perturbation of the H–F wave function. This is a purely *ab-initio* method, in contrast to DFT implementations that contain semiempirical terms.

Table 1 compares the calculated geometries, dipole moments, and harmonic frequencies of both our calculational methods with measured results, as well as with higher-level calculations. The headings of each column are in the form {calculational method/basis set}. The notation aug-cc-pVxZ refers to members of a different class of basis sets. The MP2 calculations using the basis set aug-cc-pVDZ* are useful for comparison to our B3LYP calculations, as aug-cc-pVDZ* and 6-31 + + $g(2d, p)$ are nearly identical basis sets. The results of the MP2/6-311 + + $g(2d, 2p)$ calculation are included, as these are the most rigorous calculations in the literature for which a fully optimized dimer geometry and harmonic frequencies are reported. The two sets of calculations using the large basis set aug-cc-pVQZ′ are included to examine the effect enlarging the basis set has on both MP2 and B3LYP calculations.

Our B3LYP/6-31 + + $g(2d,p)$ calculation produces excellent agreement for the monomeric geometry and harmonic frequencies, but gives only a fair result for the dipole moment. Also, these results are nearly identical to those for the MP2/aug-cc-pVDZ* calculation. Our B3LYP results also compare favorably to those of the calculations performed using larger basis sets. The primary effect of increasing the basis-set size is to improve the calculated dipole moment of the monomer, independent of the calculational method, but it does not significantly improve the predictions of the geometry or the vibrational frequencies. The H–F calculation produces fair results for all the quantities, and is the least computationally intensive method considered. The reported H–F frequencies are after scaling, and underestimate the experimental frequencies by about 5%. Prior to scaling, the H–F frequencies overestimate the experimental frequencies by a similar amount.

We next performed calculations on the linear dimer of water, in which the two molecules form a single hydrogen bond. Table 2 gives a comparison of the calculated geometry and intramolecular harmonic frequencies from both our methods to experimentally measured values and two higher-level MP2 calculations. The CPU times and disk usage are reported for the harmonic frequency calculations at the various levels. The frequency calculation was chosen as the basis for comparison, as it is the most computationally intensive step in the proce-

Table 1. Comparison of Calculated and Experimental Results: Water Monomer

	Expt.	H–F/ 6–31$g(d,p)$	B3LYP/ 6–31++ $g(2d,p)$	MP2/ aug-cc-pVDZ*[†]	MP2/ 6–311++ $g(2d,2p)$[‡]	B3LYP/ aug-cc-pVQZ[†]	MP2/ aug-cc-pVQZ[†]
O–H bond distance (Å)*	0.959	0.943	0.964	0.965	0.957	0.961	0.958
∠HOH (deg)*	105.0	105.97	105.27	104.2	104.3	105.1	104.2
Dipole moment (debye)**	1.854	2.148	2.001	1.986	2.064	1.845	1.857
Vibrational frequencies (cm^{-1})*	1,648	1,580	1,633	1,630	1,661	1,629	1,638
	3,832	3,704	3,807	3,818	3,869	3,804	3,843
	3,943	3,808	3,919	3,951	3,989	3,905	3,968
Basis functions	—	25	37	35	47	126	126

*Experimental data from Benedict et al. (1956).

**Experimental data from Dyke and Muenter (1973).

[†] Results from Kim and Jordan (1994). The [$4s3p2d/3s1p$] aug-cc-pVDZ* basis set was formed by Kendall et al. (1992) by adding diffuse functions to the [$3s2p1d/2s1p$] basis set of Dunning (1989).

[‡] Results from Frisch et al. (1986). The [$6s5p4d3f/5s4p3d$] aug-cc-pVQZ′ basis set was formed by Kim and Jordan (1994) by deleting functions from the [$6s5p4d3f2g/5s4p3d2f$] aug-cc-pVQZ basis set of Kendall et al. (1992).

dure. All the frequency calculations were performed using Gaussian 94 on an IBM RS/6000 computer, and utilized the option SCF = direct.

The B3LYP calculation produces very good results for the dimer geometry and agreement within 1% with the experimental harmonic frequencies. The MP2/aug-cc-pVDZ* results again show that the B3LYP and MP2 methods produce nearly identical results when using similar basis sets. Table 2 also shows that increasing the basis-set does not improve the quality of the geometric parameters or the harmonic frequencies at the MP2 level. Our H–F calculations produce reasonable results for the geometric parameters, and fair agreement with the harmonic frequencies. This agreement is pleasing, since the H–F procedure is by far the least computationally intensive procedure considered in this comparison.

Table 3 presents the results for the changes in the thermodynamic properties upon dimerization for the linear water dimer. We first consider the thermodynamic changes at ambient conditions, 298 K, and 1 atm. At these conditions, and recalling that our calculations correspond to vapor-phase association, the expected value of ΔG for this "reaction" is small and positive. If ΔG were negative, and if the linear dimer was a good approximation for all higher-order association "reactions," a high degree of polymerization would exist at gas-phase conditions. If ΔG were large and positive, there would be no appreciable association, even at liquid densities. All of the computational methods produce qualitatively correct results, with the methods incorporating electron correlation (B3LYP and MP2) being in reasonable agreement.

We next compare the results of these calculational methods with experimental estimates at two different temperatures. The "experimental" result at 0 K is taken from a very high-level molecular orbital calculation (Szalewicz et al., 1988). Except for our H–F result, all of the calculated values of ΔE (0 K) include the counterpoise correction to account for BSSE. As described earlier, we did not correct for BSSE at the H–F level. Our B3LYP result is in excellent agreement with the "experimental" value, with the two MP2 calculations also falling within the error bars. The H–F result overestimates the magnitude of ΔE, as might be expected due to the lack of a correction for the BSSE.

The experimental estimates at 373 K were derived from gas-phase thermal conductivity data (Blander et al., 1979). In this case, the H–F results are the only ones that fall within the error bars for either ΔH or ΔS. It is surprising that the least computationally intensive, and the least theoretically rigorous, procedure produces the best results. All of the calculations that included electron correlation are in good agreement with each other; however, they fall outside the error bars of the experimental estimates for the thermodynamic properties of the water dimer.

Conclusions

After examining all of the results, and recognizing that

Table 2. Comparison of Calculated and Experimental Results: Linear Water Dimer

	Expt.	H–F/ 6–31$g(d,p)$	B3LYP/ 6–31++ $g(2d,p)$	MP2/ aug-cc-pVDZ*[‡]	MP2/ 6–311++ $g(2d,2p)$[§]
O–O distance (Å)*	2.946	2.980	2.914	2.922	2.911
Dipole Moment (debye)**	2.600	2.643	2.533	2.728	2.814
Intramolecular vibrational frequencies (cm^{-1})[†]	1,653	1,579	1,637	1,634	1,661
	1,669	1,605	1,655	1,653	1,687
	3,718	3,660	3,682	3,716	3,774
	3,797	3,698	3,800	3,811	3,858
	3,881	3,784	3,888	3,917	3,958
	3,899	3,800	3,908	3,940	3,975
Basis functions in dimer	—	50	74	70	94
Frequency CPU time (s)	—	59	2,334	1,187	3,550
Frequency disk usage (Mbytes)	—	20	13	252	393

*Experimental data from Odutola and Dyke (1980).

**Experimental data from Dyke et al. (1977).

[†] Experimental data from Fredin et al. (1977) and Nelander (1978).

[‡] Results from Kim and Jordan (1994).

[§] Results from Frisch et al. (1986).

Table 3. Comparison of Association Parameters for Linear Water Dimerization

	Expl.	H–F/ 6-31$g(d,p)$	B3LYP/ 6-31++ $g(2d,p)$	MP2/ aug-cc-pVDZ*‡	MP2/ 6-311++ $g(2d,2p)$§
ΔH (298 K, kcal/mol)	—	−3.929	−2.993	−2.764	−2.707
ΔS (298 K, cal/mol·K)	—	−18.324	−21.087	−20.766	−20.864
ΔG (298 K, kcal/mol)	—	1.536	3.294	3.427	3.513
K_{eq} (298 K)	—	7.480×10^{-2}	3.847×10^{-3}	3.07×10^{-3}	2.66×10^{-3}
$\Delta E^{\dagger}$ (0 K, kcal/mol)	−4.7±0.35	−5.541	−4.714	−4.47	−4.45
BSSE (kcal/mol)	—	not calc.	0.388	0.77	0.99
ΔH (373 K, kcal/mol)**	−3.59±0.5	−3.72	−2.82	−2.58	−2.53
ΔS (373 K, cal/mol·K)**	−18.59±1.3	−17.71	−20.49	−20.33	−20.35

*Results from Szalewicz et al. (1988).
**Results from Blander et al. (1979).
†The calculated ΔE includes the BSSE correction.
‡Results from Kim and Jordan (1994).
§Results from Frisch et al. (1986).

neither the dimer nor monomer results improve significantly with increasing basis-set size, we conclude that the 6-31++ $g(2d,p)$ basis set is sufficient for use in extending our study to other molecules. This conclusion is in agreement with earlier studies (Kim and Jordan, 1994; Feller, 1992). Also the similarity of the B3LYP and MP2 results when using similar basis sets allows us to choose between the methods based solely on computational requirements.

The first measure of the computational requirements is the CPU time required for a frequency calculation. As seen in Table 2, for small systems MP2 calculations can be faster than DFT calculations, primarily due to the need to perform a numerical integration in the DFT calculations. However, Table 2 also shows the steep scaling of CPU time with the number of basis functions for MP2 calculations; MP2 calculations scale as ON^4 (Forseman and Frisch, 1993) where N is the number of basis functions, and O is the number of occupied orbitals (a function of the molecule studied, independent of the basis set). When comparing results obtained using different basis sets on the same compound, the scaling is simply N^4. This scaling agrees very well with the results of the two MP2 calculations presented in Table 2. We have performed frequency calculations using the B3LYP method on a number of compounds, and have concluded that the scaling of required CPU time with the number of basis functions is $N^{3.5}$. The difference in the scaling of the two methods indicates that above some number of basis functions, MP2 calculations require greater CPU time than B3LYP calculations. Based on the timing results in Table 2, this crossover (on an IBM RS/6000) occurs at approximately 150 basis functions. For reference, if one uses the 6-31++ $g(2d,p)$ basis set, there are 148 basis functions for the methanol dimer and 322 basis functions for the propionic acid dimer.

A second measure of the computational requirement is the amount of disk storage needed to complete the frequency calculation. The minimum amount of disk storage needed to perform an MP2 frequency calculation is $\sim 4N^4$ bytes (Forseman and Frisch, 1993). The results in Table 2 agree with this formula. The frequency calculations we performed using B3LYP indicate that the disk storage requirement of this method is $\sim 2{,}250N^2$ bytes. The difference in scaling between the methods allows larger systems to be studied using B3LYP before available disk storage is exhausted.

Because the B3LYP method yields reasonable results and requires less computational resources than the MP2 method, we have extended our study to other molecules using B3LYP/6-31++ $g(2d,p)$. Since we are primarily interested in estimating the thermodynamic parameters of hydrogen-bonding, and since the H–F/6-31$g(d,p)$ calculation produced the best agreement for these parameters with experimental data, we also perform calculations using the H–F method. The H–F method has the advantage of requiring almost an order of magnitude less computational resources than does the B3LYP method, and thus allows even larger clusters to be studied.

Acknowledgments

The authors thank the National Science Foundation (Grant CTS-9521406) and the Department of Energy (Grant DE-FG02-93ER13436) for financial support of this research. One of the authors (J.W.) thanks the National Science Foundation for a Pre-Doctoral Fellowship.

Notation

K_{eq} = equilibrium constant for dimerization "reaction"
m = particle mass
Z = charge on a particle (nucleus or electron)
ϵ_0 = permittivity of a vacuum = 8.8541878×10^{-12} C^2/N·m^2
ϵ_i = energy of one-electron orbital i, eigenvalue solution of Eq. 6
μ = dipole moment, debye
ν = vibrational frequency, cm^{-1}
∇^2 = Laplacian = $\nabla\cdot\nabla$
∇ = gradient

Subscripts

i, j = for electrons i, j
I, J = for nuclei I, J

Literature Cited

Becke, A. D., "Density-functional Thermochemistry: I. The Effect of the Exchange-only Gradient Correction," *J. Chem. Phys.*, **96**(3), 2155 (1992a).

Becke, A. D., "Density-functional Thermochemistry: II. The Effect of the Perdew-Wang Generalized-gradient Correlation Correction," *J. Chem. Phys.*, **97**(12), 9173 (1992b).

Becke, A. D., "Density-functional Thermochemistry: III. The Role of Exact Exchange," *J. Chem. Phys.*, **98**(7), 5648 (1993).

Benedict, W. S., N. Gailar, and E. K. Plyer, "Rotation-Vibration Spectra of Deuterated Water Vapor," *J. Chem. Phys.*, **24**, 1139 (1956).

Blander, M., D. J. Frurip, and L. A. Curtiss, "Studies of Molecular Association in H_2O and D_2O Vapors by Measurement of Thermal Conductivity," *J. Chem. Phys.*, **71**(6), 2703 (1979).

Born, M., and J. R. Oppenheimer, *Ann. Phys.*, **84**, 457 (1927).

Boys, S. F., and F. Bernardi, "The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors," *Mol. Phys.*, **19**, 553 (1970).

Clark, T., J. Chandrasekar, G. W. Spitznagel, and P.v.R. Schleyer, "Efficient Diffuse Function-augmented Basis Sets for Anion Calculations: III. The 3-21+G Basis Set for First-row Elements, Li-F," *J. Comp. Chem.*, **4**, 294 (1983).

Combariza, J. E., and N. R. Kestner, "Density Functional Study of Short-Range Interaction Forces between Ions and Water Molecules," *J. Phys. Chem.*, **99**, 2717 (1995).

Del Bene, J. E., "Bais Set and Correlation Effects on Computed Hydrogen Bond Energies of the Dimers $(AH_n)_2$: $AH_n = NH_3$, OH_2, and FH," *J. Chem. Phys.*, **86**(4), 2110 (1987).

Dill, J. D., L. C. Allen, W. C. Topp, and J. A. Pople, *J. Amer. Chem. Soc.*, **97**, 7220 (1975).

Ditchfield, R., W. J. Hehre, and J. A. Pople, "Self-consistent Molecular Orbital Methods: IX. An Extended Gaussian-type Basis for Molecular-Orbital Studies of Organic Molecules," *J. Chem. Phys.*, **54**, 724 (1971).

Dunning, T. H., Jr., "Gaussian Basis Sets for Use in Correlated Molecular Calculations: I. The Atoms Boron through Neon and Hydrogen," *J. Chem. Phys.*, **90**(2), 1007 (1989).

Dyke, T. R., and J. S. Muenter, "Electric Dipole Moments of Low J States of H_2O and D_2O," *J. Chem. Phys.*, **59**, 3125 (1973).

Dyke, T. R., K. M. Mack, and J. S. Muenter, "The Structure of Water Dimer from Molecular Beam Electric Resonance Spectroscopy," *J. Chem. Phys.*, **66**, 498 (1977).

Feller, D., "Application of Systematic Sequences of Wave Functions to the Water Dimer," *J. Chem. Phys.*, **96**(8), 6104 (1992).

Feller, D., and M. W. Feyereisen, "*Ab Initio* Study of Hydrogen Bonding in the Phenol-Water System," *J. Comp. Chem.*, **14**(9), 1027 (1993).

Ferguson, D. M., and P. A. Kollman, "Can the Lennard-Jones 6-12 Function Replace the 10-12 Form in Molecular Mechanics Calculations," *J. Comp. Chem.*, **12**, 620 (1991).

Fock, V., "Näherungsmethode zur Lösung des quantenmechanischen Mehrkör-perproblems," *Z. Phys.*, **61**, 126 (1930).

Forseman, J. B., and A. Frisch, *Exploring Chemistry with Electronic Structure Methods: A Guide to Using Gaussian*, Gaussian, Inc., Pittsburgh (1993).

Fredin, L., B. Nelander, and G. Ribbegard, "Infrared Spectrum of the Water Dimer in Solid Nitrogen. I. Assignment and Force Constant Calculations," *J. Chem. Phys.*, **66**, 4065 (1977).

Frisch, M. J., J. A. Pople, and J. S. Binkley, "Self-Consistent Molecular Orbital Methods 25. Supplementary Functions for Gaussian Basis Sets," *J. Chem. Phys.*, **80**, 3265 (1984).

Frisch, M. J., J. E. Del Bene, J. S. Binkley, and H. F. Schaefer, "Extensive Theoretical Studies of the Hydrogen-Bonded Complexes $(H_2O)_2$, $(H_2O)_2H^+$, $(HF)_2$, $(HF)_2H^+$, F_2H^-, and $(NH_3)_2$," *J. Chem. Phys.*, **84**(4), 2279 (1986).

Gaussian 92/DFT, Revision G.2: M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. W. Wong, J. B. Foresman, M. A. Robb, M. Head-Gordon, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian, Inc., Pittsburgh (1993).

Gaussian 94, Revision B.2: M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzales, and J. A. Pople, Gaussian, Inc., Pittsburgh (1995).

Gordon, M. S., "The Isomers of Silacyclopropane," *Chem. Phys. Lett.*, **76**, 163 (1980).

Hariharan, P. C., and J. A. Pople, "The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies," *Theor. Chim. Acta.*, **28**, 213 (1973).

Hariharan, P. C., and J. A. Pople, "Accuracy of AH_n Equilibrium Geometries by Single Determinant Molecular Orbital Theory," *Mol. Phys.*, **27**, 209 (1974).

Hartree, D. R., "The Wave Mechanics of an Atom with a Non-Coulomb Central Field: I. Theory and Methods," *Proc. Cambridge Philos. Soc.*, **24**, 89 (1928a).

Hartree, D. R., "The Wave Mechanics of an Atom with a Non-Coulomb Central Field: II. Some Results and Discussions," *Proc. Cambridge Philos. Soc.*, **24**, 111 (1928b).

Hehre, W. J., R. Ditchfield, and J. A. Pople, "Self-consistent Molecular Orbital Methods: XIII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules," *J. Chem. Phys.*, **56**, 2257 (1972).

Hobza, P., and R. Zahradnik, "Intermolecular Interactions between Medium-Sized Systems. Nonempirical and Empirical Calculations of Interaction Energies: Successes and Failures," *Chem. Rev.*, **88**, 871 (1988).

Hohenberg, P., and W. Kohn, "Inhomogeneous Electron Gas," *Phys. Rev.*, **136**, B864 (1964).

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

Kendall, R. A., T. H. Dunning, Jr., and R. J. Harrison, "Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions," *J. Chem. Phys.*, **96**(9), 6796 (1992).

Kohn, W., and L. J. Sham, "Self-consistent Equations Including Exchange and Correlation Effects," *Phys. Rev.*, **140**, A1133 (1965).

Lee, C., W. Yang, and R. G. Parr, "Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of Electron Density," *Phys. Rev. B*, **37**, 785 (1988).

Levine, I. N., *Quantum Chemistry*, 4th ed., Prentice Hall, Englewood Cliffs, NJ (1991).

McQuarrie, D. A., *Statistical Mechanics*, HarperCollins, New York (1976).

Nelander, B., "Infrared Spectrum of the Water-Hydrogen Sulfide Complex," *J. Chem. Phys.*, **69**, 3870 (1978).

Odutola, J. A., and T. R. Dyke, "Partially Deuterated Water Dimers: Microwave Spectra and Structure," *J. Chem. Phys.*, **72**, 5062 (1980).

Parr, R. G., *Quantum Theory of Molecular Electronic Structure*, Benjamin, New York (1963).

Perdew, J. P., *Electronic Structure of Solids*, P. Ziesche and H. Eschrig, eds., Akademie Verlag, Berlin (1991).

Perdew, J. P., and Y. Wang, "Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy," *Phys. Rev. B*, **45**(13), 244 (1992).

Perdew, J. P., J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation," *Phys. Rev. B*, **46**, 6671 (1992).

Pople, J. A., R. Krishnan, H. B. Schlegel, D. DeFrees, J. S. Binkley, M. J. Frisch, R. A. Whiteside, R. F. Hout, and W. J. Hehre, "Molecular Orbital Studies of Vibrational Frequencies," *Int. J. Quantum Chem., Symp.*, **15**, 269 (1981).

Pople, J. A., M. Head-Gordon, D. J. Fox, K. Raghavachri, and L. A. Curtiss, "Gaussian-1 Theory: A General Procedure for Prediction of Molecular Energies," *J. Chem. Phys.*, **90**(10), 5622 (1989).

Slater, J. C., *Quantum Theory of Atomic Structure*, Vols. I and II, McGraw-Hill, New York (1960).

Slater, J. C., *Quantum Theory of Molecules and Solids*, Vol. I, *Electronic Structure of Molecules*, McGraw-Hill, New York (1963).

Szalewicz, K., S. J. Cole, W. Kolos, and R. J. Bartlett, "A Theoretical Study of the Water Dimer Interaction," *J. Chem. Phys.*, **89**(6), 3662 (1988).

Zheng, Y.-J., and K. M. Merz, "Study of Hydrogen Bonding Interactions Relevant to Biomolecular Structure and Function," *J. Comp. Chem.*, **13**(9), 1151 (1992).

Manuscript received Apr. 22, 1996, and revision received July 22, 1996.