1 / VOL. 3, NO. 5
THE CHEMICAL EDUCATOR

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9

In the Classroom

Linear Water: Would We be Swimming in it?

WILLIAM A. SHIRLEY
Department of Molecular Biology
The Scripps Research Institute
La Jolla, CA 92037
wshirley@scripps.edu

This report records a light-hearted (yet rigorous) study to determine the importance of the bend in the water molecule. We present the answer to what could be a rather open-ended question posed in a computational chemistry course to an advanced undergraduate student with access to a personal computer with one of several quantum chemical programs. We use commonly available computational chemical techniques as well as qualitative molecular orbital theory to probe the hypothetical question: If water were linear, what would be its nature?

…to probe the hypothetical question: If water were linear, what would be its nature?

Introduction

In the spirit of *Flatland* [1], where the three-dimensional world is reduced to two dimensions, we consider the implications of linear water as determined by HOMO–LUMO arguments and high quality ab initio calculations. Our purpose is twofold: to examine an intellectually interesting question (and possibly gain knowledge of the physical forces that bind “real” water), and to introduce people to a few tricks and features of computational chemistry. Those with a knowledge of computational chemistry should skip the next section. Those who find their curiosity piqued by the computational tools should consult the excellent introductory book by Lowe [2].

“Were water linear … it would not be a liquid at ambient temperatures at the earth’s surface” [3]. This would seem to be an accurate statement if we subscribe to conventional theories about the association of water molecules. While a true understanding of the many-body interactions that could lead to the formation of a linear, liquid water is too complex for us, we can study the intermolecular binding of individual water monomers by examination of the binding in $(H_2O)_2$ using conventional HOMO–LUMO arguments [4] as well as high quality ab initio calculations [5]. From this pursuit, we can draw conclusions about the existence of a waterlike liquid in our hypothetical world.

Methods, Basis Sets, and Units

The mysterious world of computational chemistry can be entered fearlessly with a little knowledge. The objective is to solve the (analytically) “unsolvable” Schrödinger equation,

$$\hat{H}\Psi = E\Psi \tag{1}$$

for multielectron systems. The assumption that Hartree made is that the state wavefunction of a molecule can be described by the product of orbital wavefunctions. The wavefunction in a more accurate Hartree–Fock method is an antisymmetrized determinant rather than a simple product. In the Hartree–Fock method the Hamiltonian, $\hat{H}^{\mathrm{HF}}$, that describes the system is related to the wavefunction; the relevant equations are solved to self-consistency by an iterative process.

All molecular properties can be computed from the wavefunction. Basis sets are functions that allow the description of the wavefunction of a molecule. We assume that atoms have hydrogen-like orbitals, and molecules formed by the combination of these atoms have molecular orbitals which are distorted atomic orbitals. Thus we use linear combinations of atomic orbitals, each of which is a Slater function, $e^{-\zeta x}$. For example, a hydrogen 1s atomic orbital is represented by a single Slater function, $\pi^{-\frac{1}{2}}\zeta^{\frac{3}{2}}e^{-\zeta x}$, where the orbital exponent, ζ, in atomic units is equal to one. These Slater functions,

© 1998 SPRINGER-VERLAG NEW YORK, INC.
ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9

for computational expediency, are constructed from several Gaussian functions with fixed coefficients,

$$e^{-\zeta x} \cong \left(c_1 e^{-\alpha_1 x^2} + c_2 e^{-\alpha_2 x^2} + c_3 e^{-\alpha_3 x^2} \right) \tag{2}$$

The use of one basis function to describe each occupied atomic orbital is referred to as single zeta (where zeta is the orbital exponent). The simplest basis set commonly employed is a single zeta basis named STO-3G, where STO stands for Slater Type Orbital and 3G indicates three Gaussian functions grouped to approximate the one STO. As orbitals in molecules do differ from those in atoms, it is useful to include more than one basis function per orbital, to use multiple zetas. It is also useful to increase flexibility by using functions of greater complexity. Because orbitals are described by spherical harmonics, an obvious function to use would be a harmonic of the next orbital (azimuthal) quantum number. Thus, for first-row (heavy) atoms (Li–Ne) this would be a d function, while for hydrogen this would be a p function. These functions are called *polarization functions,* as they allow orbitals to polarize and bend. The most often employed basis set is a double zeta plus polarization function basis set, the 6–31G** basis of Pople, et al. The name of this basis set corresponds to the convention of the Pople group; each digit refers to the number of Gaussian functions (primitives) contracted to one function (Table 1). The numbers before the dash indicate the number of contracted Gaussian functions per core orbital, while those after the dash are for the valence orbitals. The first asterisk (star) indicates polarization on heavy atoms; the second indicates polarization on hydrogen atoms. Other symbols, such as +, *f*, and †, may be appended; these refer to additional special functions.

While the Hartree–Fock method is fast and accurate, an even more accurate method would allow electrons to respond to the instantaneous (rather than the average) potentials of the other electrons. The correlation of electrons and the lowering of energy that this permits can be computed in many ways. A systematic method used for including the affect of correlation corrects the initial Hamiltonian, in this case the Hartree–Fock Hamiltonian, $\hat{H}^{HF}$, by the addition of a perturbing term,

$$\hat{H} = \hat{H}^{\mathrm{HF}} + \lambda \hat{H}' \tag{3}$$

The λ term determines the extent of the perturbation and ranges from 0 (where there is no perturbation) to 1 (with the full perturbation applied). Møller and Plesset devised a

TABLE 1. The 6-31G** Basis Set Composition For Typical First- And Second-Row Atoms.

	H			C		
Atomic Orbital	Gaussians/ Contracted Basis Functions	zeta	Polarization (Second Star)	Gaussians/ Contracted Basis Functions	zeta	Polarization (First Star)
1s	1/1	<= ζ_1		6/1		
	1/1	<= ζ_2				
2s				3/1	<= ζ_1	
				1/1	<= ζ_2	
$2p_x$, , $2p_y$, $2p_z$	1/1		<= *	3/1	<= ζ_1	
				1/1	<= ζ_2	
$2d_{xz}$, $2d_{xy}$, $2d_{yz}$, $2d_{x2-y2}$, $2d_{z2}$ or $2d_{-3}$, $2d_{-2}$, $2d_{-1}$, $2d_1$,$2d_2$, $2d_3$				1/1		<= *

commonly used scheme, which adds back the missing interelectron repulsion in the perturbing Hamiltonian,

$$\hat{H}' = \sum_{i<j} \frac{1}{r_{ij}} - \sum_i v^{\mathrm{HF}}(i) \tag{4}$$

The second-order Møller–Plesset (MP2) energy correction to the Hartree–Fock (HF) energy, which comes from the interaction of each occupied (i) and each virtual (j) orbital from the Hartree–Fock calculation with this Hamiltonian is,

$$\Delta E = \frac{\left|H'_{ij}\right|^2}{E_i^0 - E_j^0}, \quad H'_{ij} \approx -S_{ij} \tag{5}$$

The summation of this term over all pairs of occupied and virtual orbitals accounts for roughly 75% of the correlation energy.

With the Hartree–Fock or MP2-corrected wavefunction in hand, it is possible to calculate the charge on an atom. A common method, not without flaws, is the so-called

Mulliken charge technique. As molecular orbitals are built from atomic orbitals (represented by the basis functions), the molecular orbital coefficients, $c_{\text{ao,mo}}$, can be used to determine the number of electrons (the population) in each atomic orbital. The Mulliken (population analysis) method assigns electron density in an atomic orbital to that atom and evenly partitions the electron density of overlapping atomic orbitals. The resulting expression for the Mulliken charge is

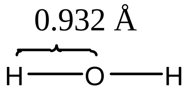
$$\underset{\text{net charge}}{q_{\text{atomic}}} = \underset{\text{charge}}{q_{\text{nuclear}}} - \underset{\substack{\text{electronic}\\ \text{charge}}}{\sum_{i}^{AOs} q_i} = \underset{\text{charge}}{q_{\text{nuclear}}} - \Big(\underset{\text{AO population}}{\sum_{i}^{AOs} \sum_{i}^{MOs} n_i c_{ii}^2} + \tfrac{1}{2} \underset{\substack{\text{overlap population}\\ \text{(shared electrons)}}}{\sum_{i}^{AOs} \sum_{i<k}^{AOs} 2 \sum_{i}^{MOs} n_i c_{ii} c_{ki} S_{ik}} \Big) \tag{6}$$

The 3*N*–6 natural vibrational frequencies (or the 3*N*–5 frequencies of a linear molecule such as a linear water molecule), the normal modes, can be calculated analytically from the Hartree–Fock and Møller–Plesset second-order wavefunctions. An examination of the resultant harmonic frequencies can give an estimate of the force constants for bonds and angles.

The "natural" units used in computational chemistry are the atomic units, hartrees (E_h) and often millihartrees (mE_h). Conversion factors to other commonly used units are provided in Table 2. Frequently used software includes the Gaussian, Spartan, and Jaguar programs. Information about these packages can be found on the WWW at www.gaussian.com, www.wavefun.com, and www.schrodinger.com, respectively.

The Rigorous Calculations

As they are accurate and widely used, we perform our ab initio calculations using the restricted Hartree–Fock/6–31G** and Møller–Plesset second-order/6–31G** models. An MP2/6–31G** optimized linear water molecule has a total energy of –76.1658 E_h (HF = –75.9686 E_h) and the geometry depicted in **1**.

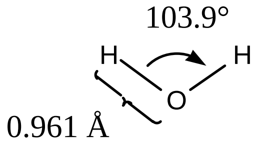


1

TABLE 2. The Conversion Between Units Employed in Computational Chemistry and Other Common Units.

	1 hartree (E_h)	1 millihartree (mE_h)	cm^{-1}
kilocalorie	627.503	0.627	2.859×10^{-3}
eV	27.2107	0.027	1.24×10^{-4}
kilojoule	149.977	0.150	6.83×10^{-4}
hartree	1	0.001	4×10^{-6}
millihartree	1000	1	4.56×10^{-4}

The corresponding fully optimized structure for H_2O at this level of ab initio theory has a total energy of –76.2224 E_h(HF = –76.0229 E_h). The structure, **2**,



2

closely agrees with the experimental result (Table 3). The optimized linear water molecule thus is 56.7 mE_h (35.6 kcal mol^{-1}) above the bent molecule in the MP2/6–31G** model. The linear molecule in the Hartree–Fock model, without the inclusion of correlation energy, is higher in energy than the bent form by 54.2 mE_h (34.0 kcal mol^{-1}). In this system, the inclusion of correlation energy yields only a small correction to the Hartree–Fock energy.

The strength of an interaction between molecules can be determined from the difference between the energies of the individual (lone) molecules and the interacting molecules. The problem with approximating association energies in this way is that there is a computational artifact that lowers the energy of the interacting molecules. A reasonably sized set of functions cannot describe orbitals completely, thus a molecule brought into proximity can and will lend functions, and thus lead to an artificial lowering of energy. This counterintuitive orbital-mixing or basis-set-superposition

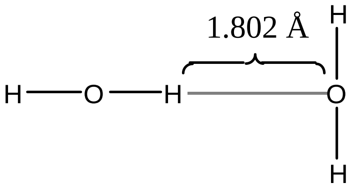
ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9

TABLE 3. The *ab Initio* Optimized and Accepted Experimental Structures of The Water Molecule.

Model	^{r}OH$^{(\text{Å})}$	∠HOH
HF/STO-3G	0.9894	100.02°
HF/6-31G**	0.9430	105.97°
MP2/6-31G**	0.9607	104.36°
Experimental	0.9572	104.54°

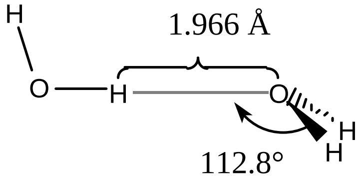
error can be corrected by the computational trick of calculating the energy of a molecule produced by presence of the functions of another molecule, but without the other molecule's nuclei or electrons. The basis-set-superposition error was examined for these calculations, and it was determined not to be significant.

The strength of the hydrogen bond in the optimized structure of the linear water dimer is 14.1 mE_h (8.8 kcal mol^{-1}). The corresponding Hartree–Fock value is 10.2 mE_h(6.4 kcal mol^{-1}) with the geometry of **3**.



3

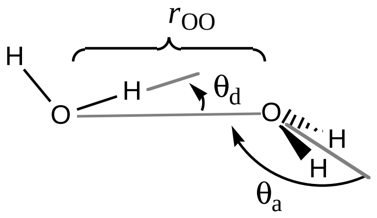
This value is greater than that for the binding of bent water monomers, which is 11.1 mE_h (7.0 kcal mol^{-1}) in the MP2/6–31G** model, and 8.7 mE_h (5.4 kcal mol^{-1}) without correlation energy. It possesses the geometry shown in **4,** which agrees with experimental results as well (Table 4).



4

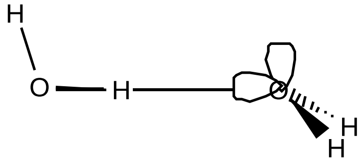
ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9

TABLE 4. The *ab initio* Optimized and Accepted Experimental Structures of Water Dimer From [8].



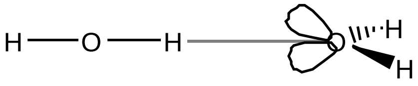
Model	^{r}OO$^{(Å)}$	θ_d	θ_a
HF/STO-3G	2.740	0.2°	124.0°
HF/6-31G**	2.980	5.3°	117.6°
MP2/6-31G**	2.910	9.1°	100.2°
Experimental	2.976 ± 0.001	–1 ± 1°	122.4 ± 0.9

After studying a dimer of two linear water molecules, we also examine a dimer formed from one linear and one normal water molecule. Looking at the dimer formed from two bent water monomers, it appears that the geometry is determined by the orientation of the lone pairs, **5**.



5

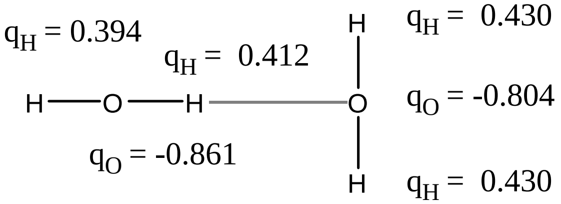
If this were the most important feature of bonding, a linear hydrogen donor and a bent normal monomer acceptor would align one lone pair. Consequently the hydrogens of the acceptor would be bent down or up from the axis of the linear water. The energy-optimized structure is actually planar, **6**.



6

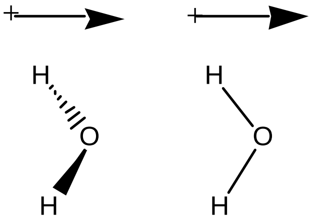
The bending down of the hydrogens in the normal water dimer appears to be a result not of the orientation of the lone pairs, but rather of the Coulombic repulsion between the positive hydrogens of the acceptor and the positive hydrogen of the donor bending up. Further support for this view comes from the electrostatic surface of H_2O. The diffuse lone-pair lobes are only distinct at very small distances, and thus it appears that the positive hydrogens must dictate the structure. Flattening the hydrogen of the donor (bending it down) allows the two hydrogens of the acceptor to lift up.

The Mulliken charges show how the alignment of the quadrupoles determines the binding of the linear water dimer, **7**.



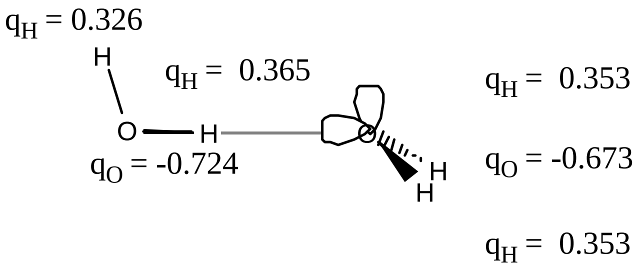
7

In normal water dimer the dipole–dipole interactions are not wholly responsible for the structure because a simple dipole–dipole model indicates that the dipoles of the monomers should be parallel, **8**.



8

The dominant interaction in water dimer appears to be a combination of dipole, quadrupole, and higher multipole interactions, rather than a HOMO–LUMO interaction, **9**.



9

The larger charges in the linear complex can explain the enhancement of binding over the bent complex. Looking at the change in the charge on the atoms involved in the hydrogen bond, it would appear that the binding would be greater by a factor of 1.35 in the linear complex than in a normal water complex,

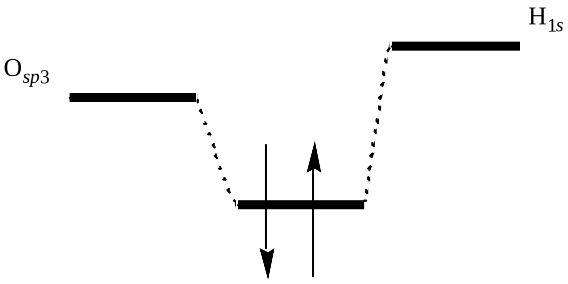
$$\frac{\dfrac{q_O^L q_H^L}{\left(r_{OH}^L\right)^2}}{\dfrac{q_O^B q_H^B}{\left(r_{OH}^B\right)^2}} = \frac{q_O^L q_H^L}{q_O^B q_H^B} = \frac{(0.804)(0.412)}{(0.673)(0.365)} = 1.35; \quad r_{OH}^L \approx r_{OH}^B \tag{7}$$

In fact, the MP2/6–31G** model yields a harmonious result. The calculated energies differ by a factor of 1.28,

$$\frac{14.1}{11.1} = 1.28 \tag{8}$$

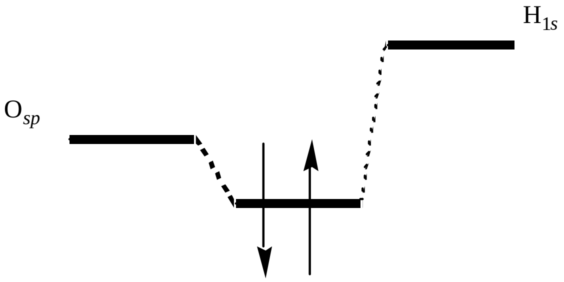
An understanding of the ordering of molecular orbitals allows us to construct a simple argument. The increase in charge on the oxygen and consequently the hydrogens of the linear water can be explained by the increased s character of the O–H bond. The sp^3 hybridization of the normal water molecule is converted to sp when the molecule is straightened [6]. The decrease in s character of the lone pairs is responsible for the higher total energy of the linear monomer, and the increase in s character of the orbitals involved in the O–H bonds increases the charge on the oxygen. Unlike the hydrogenic orbitals, the 2s orbital on oxygen is lower in energy than the 2p orbitals, **10**.

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9



10

Hence, the sp hybrid is lower in energy than the sp^3 hybrid and gives a more polar O–H bond, **11**.



11

While the identification of specific molecular orbitals is difficult with sophisticated ab initio methods such as those employed here, we know that the HOMO of the hydrogen acceptor would describe a lone pair, and the LUMO of the donor would be directed toward this pair. If HOMO–LUMO interactions governed the binding, we would expect torsion about the hydrogen bond axis in normal water dimer to require less energy than bending the hydrogen bond, which would disturb the orientation of the HOMO lone pair. The twelve normal modes of water dimer include six modes of the two monomers, which are only slightly perturbed, and six new "dimer modes" (Table 5 and Table 6). These normal modes of the fully optimized water dimer in the MP2/6-31G† model [7] include a torsion frequency of 229.0 cm^{-1}. The modes that involve bending around the hydrogen bond fall in the range 181.5–405.1 cm^{-1}. In both torsion and bending, the hydrogens of the acceptor and the hydrogen not involved in the hydrogen bond of the donor will travel closer to each other. Thus the comparable frequencies of torsion and bending imply that the same coulombic repulsion between the positive hydrogens determines the frequencies in both types of motion.

TABLE 5. The *ab initio* Calculated and Accepted Experimental Normal Modes of Water Molecule.

	Frequency (cm^{-1})	
Mode	MP2/6-31G**	Experimental [9]
1	1684.3	1595
2	3913.4	3652
3	4040.2	3756

TABLE 6. The *ab initio* Calculated Normal Nodes for Water Dimer.

Mode		Frequence (cm^{-1})	
1		113.4	
2		181.5	
3		201.5	
4		228.9	
5		405.1	
6		617.5	
7	8	1686.0	1718.0
9	10	3873.8	3900.0
11	12	3995.2	4017.0

Conclusion

We have attempted to combine rigorous molecular orbital calculations with a qualitative HOMO–LUMO description of bonding. The use of the dimers ignores the wealth of higher-order multibody interactions, thus giving us only a gross estimate of the properties of liquid water. It has been shown that two-body interactions in water are the dominant force, so our conclusions from this study are valid. We found inconsistency between simple HOMO–LUMO arguments and the results of our more rigorous ab initio calculations. Two conclusions appear to follow from this inquiry:

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 05248-9

1. The forces holding water molecules together (i.e., hydrogen bonds) are primarily electrostatic in origin.

2. Starting from the assumption that linear water monomers exist, linear liquid water would exist, bound by the same electrostatic forces responsible for normal water. At earth's ambient temperature it would not only be a liquid, but it would most likely have a boiling point higher than water! As to the other colligative properties, I am still pondering whether we could skate on the substance; molecular dynamics calculations of putative structures might be revealing.

ACKNOWLEDGEMENT

This work was performed at Wesleyan University under the auspices of George A. Petersson. I would like to thank G. A. P., David Beveridge, and Roald Hoffmann for their aid in this study.

REFERENCES

1. Abbott, E. A. *Flatland: A Romance of Many Dimensions*; Signet Classic, New American Library: New York, 1984; p 160.
2. Lowe, J. P. *Quantum Chemistry*; Academic Press: New York, 1993; p 711.
3. Hoffmann, R.; Torrence, V. *Chemistry Imagined: Reflections on Science*; Smithsonian Institution Press: Washington and London, 1993; pp 57–60, 168.
4. Albright, T. A.; Burdett, J. K.; Whangbo, M.-H. *Orbital Interactions in Chemistry*; John Wiley & Sons: New York, 1985; p 447.
5. Shirley, W. A. Ph.D. thesis, Wesleyan University, Middletown, CT, 1991.
6. Shirley, W. A.; Hoffmann, R. *J. Phys. Chem.* **1995**, *99,* 4025–4033.
7. Petersson, G. A.; Bennett, A.; Al-Laham, M.; Shirley, W. A.; Mantzaris, J. A. *J. Chem. Phys.* **1988**, *89,* 2193–2218.
8. *Cited in* Smith, B. J.; Swanton, D. J.; Pople, J. A.; Schaefer, H. F., III; Radom, L. J. *J. Chem. Phys.* **1990**, *90,* 1240–1247.
9. Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980**, *72,* 5062–5070.
10. Herzberg, G. Infrared *and Raman Spectra of Polyatomic Molecules*; Van Nostrand: Princeton, NJ, 1945; p 632.