

# The Water Dimer: Correlation Energy Calculations

Subhas J. Chakravorty and Ernest R. Davidson\*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

Received: December 28, 1992; In Final Form: March 29, 1993

The binding energy of the water dimer at the experimental geometry has been calculated employing the Hartree–Fock (HF) method, Møller–Plesset theories MP2, MP3, and MP4, and the coupled cluster method with double excitations. We have also used configuration interaction wavefunctions both at the HFSDCI (single and double excitations from the HF configuration) and MRSDCI (single and double excitations from multiple reference configurations) levels to calculate the interaction energy. Binding energies within quasi-degenerate variation perturbation theory, linearized coupled cluster method, and the average pair correction model are also calculated. The basis set developed and employed in the study yields  $-76.0674$  au for the HF energy of the water monomer at the experimental geometry. The basis set superposition error for the HF and MP2 models is 0.003 and 0.251 kcal/mol. The MP2 interaction energy has been partitioned into pair energies to give a better understanding of the hydrogen bond. The present study is unable to yield results close to  $-5.4$  kcal/mol, the experimental estimate for the interaction energy. Instead, values ranging from  $-4.2$  to  $-4.7$  kcal/mol have been obtained. On the other hand, very large scale MRSDCI wavefunctions which are almost size consistent yield values as poor as  $-3.3$  kcal/mol. Correlation energy density functional models yield interaction energies in the range of  $-5.2$  to  $-5.8$  kcal/mol. However, when the Hartree–Fock exchange term is replaced with exchange energy density functional models, values from  $-2.8$  to  $-4.5$  kcal/mol are obtained.

## 1. Introduction

The experimental estimate of the binding energy,  $D_e$ , of the water dimer at the equilibrium geometry with respect to the individual water monomers is  $-5.4 \pm 0.7$  kcal/mol,<sup>1,2</sup> and theoretical studies<sup>3–23</sup> to accurately calculate this quantity continue unabated. Recent studies on the water dimer by Feller,<sup>16</sup> van Duijneveldt et al.,<sup>17</sup> and Kim et al.<sup>18</sup> revisit this difficult problem of calculating the binding energy of the water dimer with ab initio methods.

Feller<sup>16</sup> estimated the binding energy to be about  $-4.7$  kcal/mol using the second-order Møller–Plesset perturbation method, MP2. He presented empirical evidence similar to that presented by Del Bene<sup>19</sup> that the counterpoise corrected MP2 binding energy is less constant as the basis set is augmented than not making any correction at all. The MP2 binding energy increases as the basis set is augmented, while the BSSE is expected to decrease. Hence, these two sources of error may tend to cancel. Feller and Del Bene seem to have found a sequence of basis sets for which the sum of the incompleteness error plus the BSSE error remains approximately constant so that the uncorrected MP2 result varies slowly with the basis set. On the basis of an extrapolation of his computed results, Feller estimated that the binding energy could be about  $-5.1$  kcal/mol. The largest calculation actually performed including the counterpoise correction<sup>24</sup> yielded about  $-4.67$  kcal/mol.

Van Duijneveldt et al.<sup>17</sup> used MP2 and the coupled electron pair approximation, CEPA, to compute the binding energy to be about  $-4.73$  kcal/mol and advocated further studies using multireference configuration interaction, MRCI, and correlated methods beyond MP4 to further reduce the uncertainty in the binding energy obtainable from theoretical calculations.

In the work of Kim et al.,<sup>18</sup> the binding energy was determined at the MP2 level with a 13s8p4d2f18s4p2d basis set. They have used optimized Hartree–Fock, HF, and optimized MP2 structures to compute the total energies and found that the interaction energy is about  $-4.99$  kcal/mol without, and  $-4.66$  kcal/mol with, the counterpoise correction.

TABLE I: Hartree–Fock Energies, ( $-E_{\text{HF}}$ ) for Oxygen and Hydrogen Atoms<sup>a</sup>

atom	state	contracted	uncontracted	$E_{\text{HF}}$
O <sup>+</sup>	4S	-74.372 559 31	-74.372 602 72	-74.372 606
	2P	-74.142 051 80	-74.142 095 31	-74.142 098
	2D	-74.233 296 55	-74.233 339 98	-74.233 343
O	3P	-74.809 392 07	-74.809 396 15	-74.809 398
	1D	-74.729 257 80	-74.729 261 98	-74.729 264
	1S	-74.611 013 70	-74.611 018 34	-74.611 021
O <sup>-</sup>	2P	-74.789 654 62	-74.789 743 44	-74.789 746
H	2S	-0.499 913 45	-0.499 999 24	-0.5
H <sup>-</sup>	1S	-0.477 134 63	-0.487 910 28	-0.487 929

<sup>a</sup> All values are in hartrees.

In this paper, we present some results of our investigations of the hydrogen bonding in the water dimer that emphasize aspects of the water dimer interaction not discussed in these other recent papers. We have sought, first, to develop a useful basis set for the binding energy of the water dimer. The emphasis of our work was primarily to reduce the total basis set superposition error (BSSE) at the HF and MP2 levels. Then, we have estimated the binding energy with a variety of size-extensive methods such as Møller–Plesset, MP $n$ , and coupled cluster. The MP2 binding energy has been partitioned into pair correlation interaction energies for a more detailed analysis of the individual magnitude of various contributions to the total binding energy. The effect of stretching of the hydrogen–oxygen bonds on the binding energy has been examined. We have examined the difficulties in using large-scale configuration interaction (CI), quasi-degenerate variation perturbation theory (QDVART),<sup>25</sup> multireference linearized coupled cluster (MRLCC),<sup>26</sup> and average pair correction models<sup>27,28</sup> for estimating the binding energy. Finally, we have performed an analysis of the hydrogen bonding at the SCF level using deformation density maps and frozen orbital SCF calculations.

## 2. Basis Set, Geometry, and SCF Results

The beginning basis set was the Partridge extended basis (18s13p) for the oxygen atom<sup>29</sup> and the van Duijneveldt basis (10s) for the hydrogen atom.<sup>30</sup> The oxygen basis was augmented

\* Author to whom correspondence should be addressed.

TABLE II: Cartesian Coordinates (in bohr) of the Water Dimer Used in the Present Study<sup>a</sup>

	atom	X	Y	Z
monomer B	Ob	0.0	0.0	0.0
	Hb1	0.0	-1.107 156 942	1.430 428 676
	Hb2	0.0	-1.107 156 942	-1.430 428 676
$R_{OH} = 0.9572 \text{ \AA}$ angle HOH = $104.52^\circ$ XO is the bisector of angle HOH				
monomer A Dyke et al.	Oa	4.775 683 911	2.984 178 510	0.0
	Ha1	3.241 695 777	2.025 636 339	0.0
	Ha2	4.232 355 741	4.709 495 058	0.0
$R_{OO} = 2.98 \text{ \AA}$ angle OHO = $180^\circ$ angle XOO = $122^\circ$				
monomer A Popkie et al.	Oa	2.834 588 924	4.909 652 035	0.0
	Ha1	1.930 166 084	3.343 145 725	0.0
	Ha2	1.544 871 568	6.177 939 003	0.0
$R_{OO} = 3.00 \text{ \AA}$ angle XOO = $150^\circ$				
monomer A supermolecule	Oa	852.823 780 067	532.903 442 744	0.0
	Ha1	851.289 791 933	531.944 900 572	0.0
	Ha2	852.280 451 897	534.628 759 291	0.0
stretched	Oa	4.775 683 911	2.984 178 510	0.0
	Ha1	3.209 644 207	2.005 608 295	0.0
	Ha2	4.221 003 293	4.745 544 296	0.0
	Ob	0.0	0.0	0.0
	Hb1	0.0	-1.130 290 183	1.460 316 446
	Hb2	0.0	-1.130 290 183	-1.460 316 446
$R_{OH} = 0.9772 \text{ \AA}$				

<sup>a</sup> The Popkie et al.<sup>4</sup> and the Dyke et al.<sup>25</sup> geometries have been adjusted to a convenient orientation so that the free monomers give identical energies.

TABLE III: Hartree-Fock Energies and MP2 Energies for Water<sup>a</sup>

GTO basis	$E_{HF}$	$E_c(MP2)$		ref
		frozen core		
19s14p7d2f/11s7p2d	-76.067 400	-0.280 87	-0.335 20	
7s14p7d2f/11s7p2d	-76.067 400	-0.280 67	-0.333 36	
7s8p7d2f/11s7p2d	-76.067 400	-0.280 45	-0.323 10	
7s8p6d2f/11s7p2d	-76.067 395	-0.280 28	-0.32171	
7s8p6d1f/11s7p2d	-76.067 385	-0.277 47	-0.318 71	
7s8p6d1f/4s7p2d	-76.067 384	-0.277 02	-0.318 04	
7s8p6d1f/4s2p2d	-76.067 375	-0.275 81	-0.316 68	
final basis set				
7s8p6d1f/4s2p1d	-76.067 361	-0.274 84	-0.315 63	
[12s10p4d2f/10s4p2d]	-76.067 3		-0.329 (MRSDCI)	Feller et al. <sup>32</sup>
[14s10p4d1f/6s4p1d]	-76.067 2		-0.3189 (MRSDCI)	
			-0.3112 (HFSDCI)	
cc-pV5Z		-76.067 1	-0.2915	Feller <sup>16</sup>
cc-pVQZ		-76.064 8	-0.2828	
aug-cc-pVTZ	-76.060 6	-0.268 4		
(13s8p4d2f/8s4p2d)	-76.067 56		-76.391 18	Kim et al. <sup>18</sup>
(28s18p8d3f/11s7p2d)	-76.067 42		-0.336 62	Clementi et al. <sup>33</sup>

<sup>a</sup> All values at experimental geometry of monomer except those of Kim et al.,<sup>18</sup> who optimized the geometry with each method. They did not report the HF energy at the MP2 geometry.

to the 19s14p set using the supplementary set of 1s1p provided by Partridge<sup>29</sup> to yield lower HF values for the negative ion O<sup>-</sup>. For the hydrogen negative ion, H<sup>-</sup>, we added one 1s function with a low exponent value and optimized the HF energy. The HF energies for the uncontracted isotropic part of the basis set for the oxygen atom and hydrogen are reported in Table I along with their numerical HF counterparts.

The d and the f polarization functions for oxygen and the p and d polarization functions for hydrogen have been determined by partially optimizing the HF energy for the water molecule at the experimental geometry (see Table II). Before the optimization of exponents was done, we selected seven even-tempered d functions and two f functions on oxygen. The d set had exponents ranging from 0.08 to 5.12, and the f functions were from suggested values of Feller and Davidson.<sup>31</sup> The seven p functions for hydrogen again were even tempered with exponents ranging from 0.08 to 5.12. The d functions were taken from the work of Feller et al.<sup>32</sup> Only the largest and smallest exponents of the d set for oxygen and the p set of hydrogen were optimized followed by a

final rough optimization of the f set on oxygen. The HF energy for the uncontracted basis set (19s14p7d2f/11s7p2d) of the water molecule was -76.067 400 hartrees, which seemed to be satisfactory enough for the purposes of this work. The lowest reported value for the HF energy for the water molecule at this experimental geometry is -76.067 417 hartrees obtained by Clementi et al.<sup>33</sup> with a (28s18p8d3f/18s8p3d) basis set. They also estimate the basis set limit HF energy to be -76.067 55 hartrees, in good agreement with the result of Bawagan et al.<sup>34</sup>

Contraction of this primitive set was necessary in order to carry out a multireference CI calculation for the water dimer with our computational resources. We concluded that the largest basis set we could use was 7s8p6d1f on oxygen and the 4s3p1d set on hydrogen. This choice gives 98 functions on the water monomer and 196 basis functions on the water dimer. The next step of actually constructing suitable contractions was done using the natural orbitals of a frozen core HFSDCI (configuration interaction of single and double excitations with respect to the Hartree-Fock ground-state configuration) calculation on the

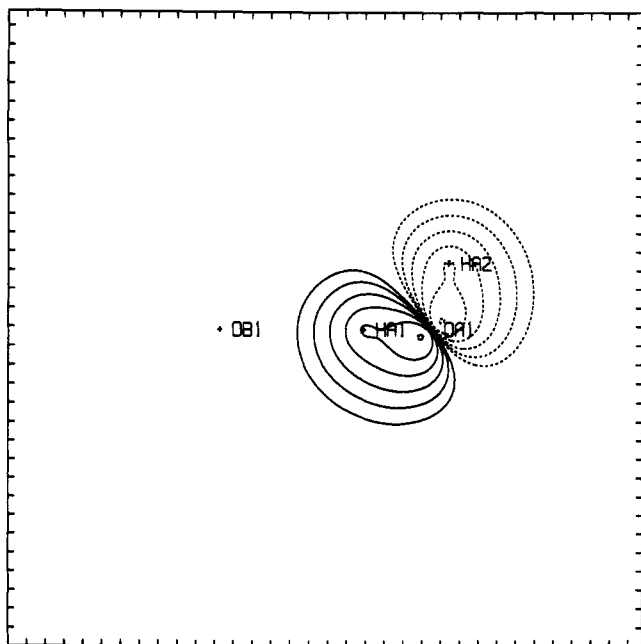


Figure 1.  $1b_2$  orbital of monomer A. Tic marks at  $0.5a_0$  spacing.

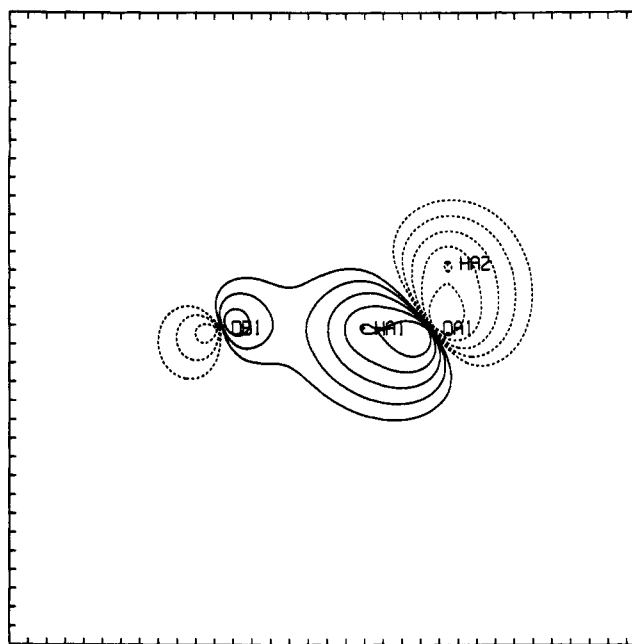


Figure 3.  $5a'$  orbital of dimer.

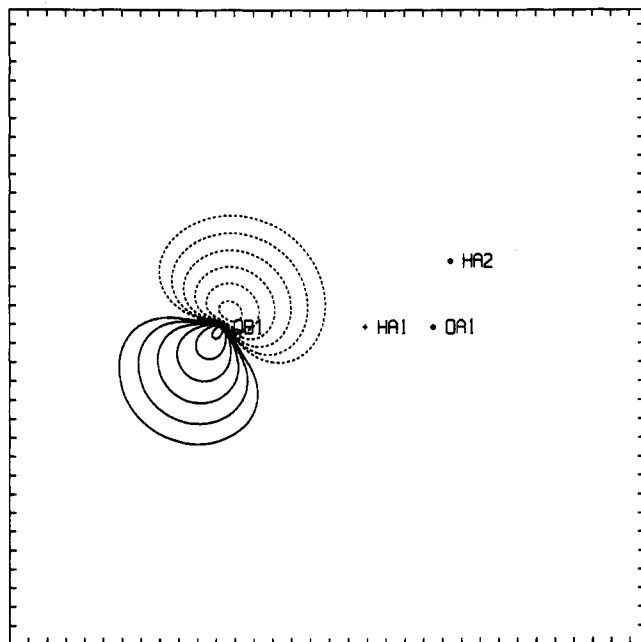


Figure 2.  $3a_1$  orbital of monomer B.

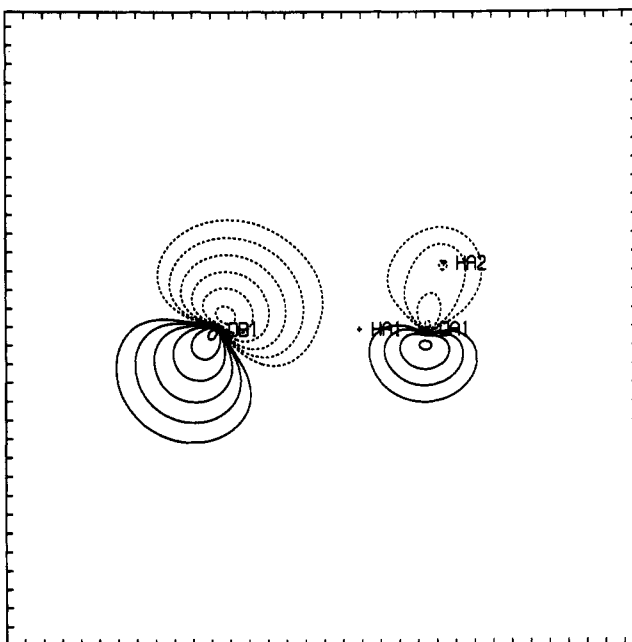


Figure 4.  $6a'$  orbital of dimer.

water molecule. For the seven s oxygen contractions, we extracted the oxygen s part of the  $1-7a_1$  frozen natural orbitals (three of which are HF orbitals) and performed a linear transformation to make shorter contractions, yielding the 7s basis set for oxygen. For the p set of oxygen, we considered the frozen natural orbitals  $2a_1$ ,  $3a_1$ ,  $1b_1$ ,  $1b_2$ ,  $4a_1$ ,  $2b_1$ , and  $2b_2$  to synthesize an 8p contraction. For the d set, we concluded that we could contract only the orbitals with the largest exponents, and the 2f functions were contracted to an optimum 1f function. The corresponding energy after contraction both at the Hartree-Fock and MP2 levels has been reported in Table III. For the hydrogen atom, the  $2a_1$ ,  $3a_1$ , and  $4a_1$  and the  $1b_2$  frozen natural orbitals were used to determine the s contraction coefficients. For the p contractions, we considered the  $1b_1$  and  $1b_2$  orbitals. An optimum contraction for the d functions of hydrogen was then determined. The contracted basis set 7s8p6d1f/4s2p1d used for water and water dimer calculations is available from the authors upon request. In Table III, a summary of the SCF and MP2 energies of the water molecule at the experimental geometry for various stages of contraction

of the basis set is reported. In Table III, we also compare the HF and MP2 energies with our basis set with a few selected basis sets recently used for the water monomer and water dimer. The HF energy is believed to be within 0.2 mhartrees of the HF limiting value at the experimental geometry. In later tables, we will show the MP2 energy will be within 1% of the estimated full CI value for these basis sets. On the other hand, with our final basis set, it is 20% from the true correlation energy,  $-0.372$  of water. The frozen core values are similarly 15% from the estimated frozen core limit of  $-0.32$ . Compared with Feller,<sup>16</sup> our contracted basis is better than his correlation consistent triple  $\zeta$ , but worse than his quadrupole  $\zeta$  basis.

The geometry we selected to perform the computations is presented in Table II. The water monomers were kept identical and left in their experimental geometry. However, for the dimer structure, there are many different geometries<sup>4,35,36</sup> used for the binding energy calculations. The work of Popkie et al.<sup>4</sup> and Szalewicz et al.<sup>12</sup> used the geometry where the water monomers were left in their experimental geometry and the interoxygen



TABLE IV: SCF Orbital Energies for the Water Dimer<sup>a,b</sup>

dimer orbital	$\epsilon_{\text{dimer}}$	orbital	$\epsilon_{\text{monomer}}$	$\epsilon_{\text{frozen AB}}^c$	$\epsilon_{\text{frozen B}}^d$	$\epsilon_{\text{frozen A}}^d$
1a'	-20.5883	1a <sub>1</sub> (B)	-20.5640	-20.5838	-20.5888	-20.5850
2a'	-20.5345	1a <sub>1</sub> (A)	-20.5640	-20.5387	-20.5392	-20.5321
3a'	-1.3793	2a <sub>1</sub> (B)	-1.3527	-1.3751	-1.3784	-1.3745
4a'	-1.3257	2a <sub>1</sub> (A)	-1.3527	-1.3297	-1.3291	-1.3248
5a'	-0.6965	1b <sub>2</sub> (A)	-0.7179	-0.6989	-0.6941	-0.6911
6a'	-0.6144	3a <sub>1</sub> (B)	-0.5850	-0.6119	-0.6162	-0.6097
7a'	-0.5622	3a <sub>1</sub> (A)	-0.5850	-0.5647	-0.5634	-0.5606
8a'	-0.5317	1b <sub>1</sub> (B)	-0.5105	-0.5298	-0.5433	-0.5350
1a''	-0.7427	1b <sub>2</sub> (B)	-0.7179	-0.7374	-0.7408	-0.7395
2a''	-0.4840	1b <sub>1</sub> (A)	-0.5105	-0.4900	-0.4884	-0.4852

<sup>a</sup> The dimer geometry of Dyke et al. <sup>b</sup> All values are in hartrees.

<sup>c</sup> Orbital energies of dimer using only occupied monomer orbitals as a basis. <sup>d</sup> Orbital energies when the orbitals of one monomer are kept frozen.

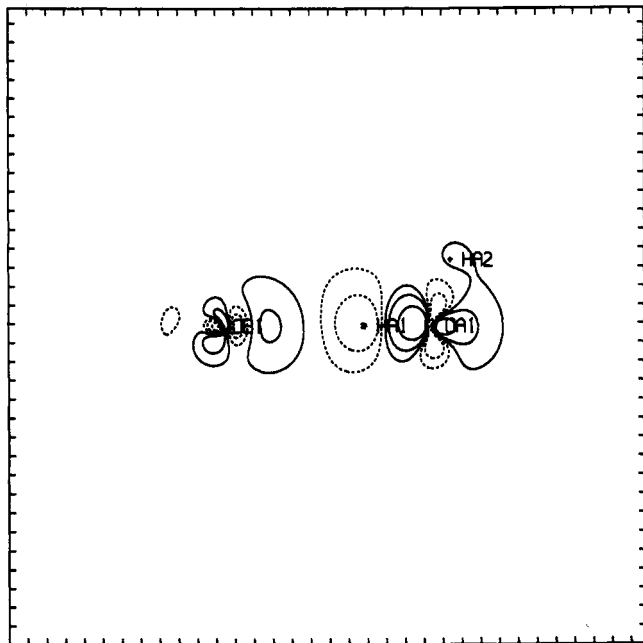


Figure 9. Density difference map, dimer minus sum of monomers.

To verify that this is only due to polarization within each monomer and not due to charge transfer, we performed frozen orbital HF calculations. These dimer Hartree-Fock calculations were performed by constraining one of the sets of monomer orbitals to remain unchanged and minimizing the energy with respect to the orbitals of the other monomer. In Table IV, we have presented the orbital energies of such calculations alongside the dimer and monomer orbital energies. Figures 10 and 11 show the polarization of each monomer in the field of the other. Figure 12 shows the sum of these two deformation densities. Figure 9 and Figure 12 are almost identical, which confirms that the primary distortion in the water dimer is due to polarization of each monomer. These conclusions and figures are nearly identical to those obtained by Krijn and Feil<sup>21</sup> using density functional methods.

In continuation of our analyses of the SCF interaction energy of the water dimer, a Morokuma partitioning<sup>37</sup> of the interaction energy was done. In this method, the total interaction energy  $\Delta E$  is partitioned into a sum of electrostatic (ES), polarization (PL), exchange (EX), and charge transfer and other higher order terms (CT). Thus,

$$\Delta E = \text{CT} + \text{ES} + \text{EX} + \text{PL} \quad (1)$$

The ES term is calculated as the Coulombic interaction of the free monomer charge distributions. The sum, ES + EX, is defined as the average energy of a Slater determinant formed from the monomer MOs. PL is the sum of polarization energy of each monomer using only its own basis set in the electrostatic field of

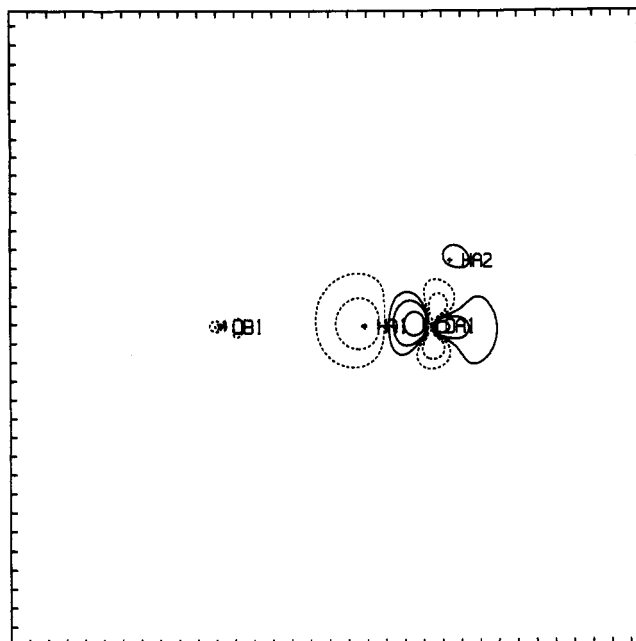


Figure 10. Polarization density for monomer A.

the other (without regarding orthogonality of orbitals). The sum, PL + CT, is defined here as the total orbital relaxation, -1.809 kcal/mol, and gives the difference between the ES + EX energy and the final HF energy. Umeyama and Morokuma evaluated these quantities with a 6-31G\*\* basis set and obtained  $\Delta E$  to be -5.6 kcal/mol, with -7.5 kcal/mol for the ES term, +4.3 kcal/mol for the EX term, -0.5 kcal/mol for the PL term, and -1.8 kcal/mol for the CT term, respectively. These 6-31G\*\* results were pleasing because the Hartree-Fock interaction energy agreed well with experiment. Table V summarizes some of the calculations for the binding energy of the dimer for improved basis sets. Unfortunately, as this table shows, as the basis set is improved, the HF interaction energy converges to -3.6 kcal/mol at the Dyke geometry. The electrostatic energy for the water dimer for our basis set was -7.301 kcal/mol, with 5.497 kcal/mol for exchange, -1.431 kcal/mol for polarization, and -0.378 kcal/mol for charge transfer. These values are in closer conformity with the results White and Davidson<sup>15</sup> obtained for their SVP basis ( $\Delta E = -4.2$ , ES = -7.6, EX = 4.7, PL = -0.7, and CT = -0.7 kcal/mol). The SVP basis set [3s2p1d/2s1p] was developed to study larger water clusters in ice. The ES energy term is the largest attractive term in the dissected interaction energy and is quite similar for these three basis sets. However, the exchange term almost cancels it and becomes larger with improved basis sets. Other work<sup>34</sup> has shown improved basis sets also result in a more diffuse b<sub>1</sub> orbital. Presumably, this increase in the electron density in the tail of the wavefunction results in larger overlap with the neighboring monomer and, hence, a larger EX energy.

Another way to analyze the total interaction energy is to consider a different partitioning of the relaxation energy of orbitals of individual monomers. Define

$$\Delta E = \Delta E(\text{frozen A, B}) + \Delta E(\text{frozen A}) + \Delta E(\text{frozen B}) + \text{mixing} \quad (2)$$

where

$$\Delta E(\text{frozen A, B}) \equiv \text{ES} + \text{EX} \quad (3)$$

and  $\Delta E(\text{frozen A})$  is the improvement in energy when the orbitals of monomer B are allowed to relax in the full basis of the dimer keeping the orbitals of A unchanged. The values for the three terms are -1.32 kcal/mol for the relaxation in the proton donor (monomer A), -0.45 kcal/mol for the proton acceptor (monomer B), and -1.804 kcal/mol for the dimerization energy with frozen

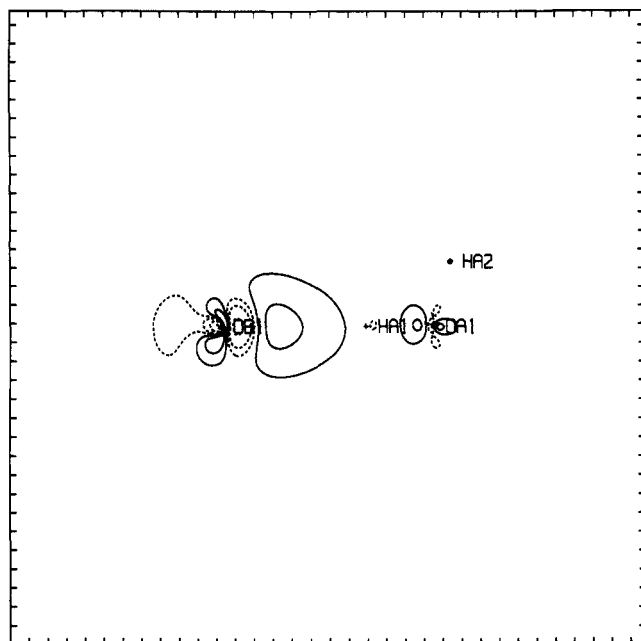


Figure 11. Polarization density for monomer B.

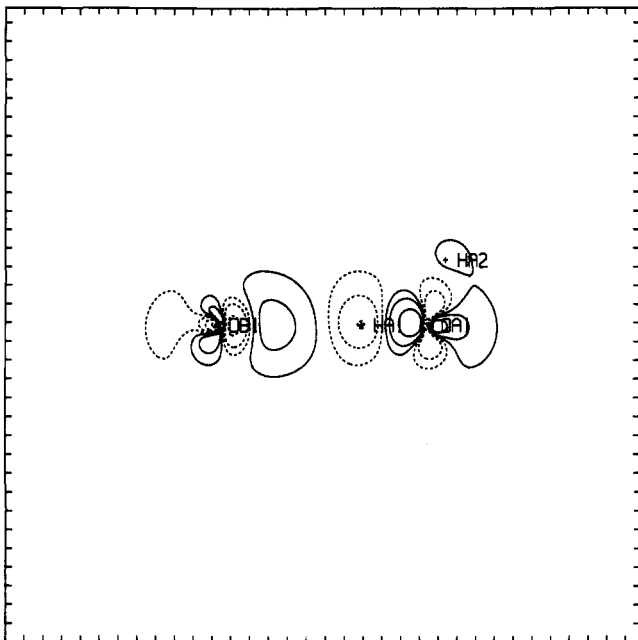


Figure 12. Sum of polarization densities of A and B.

orbitals of both monomers A and B. The mixing term is only  $-0.04$  kcal/mol, so the polarized monomers with orthogonality constraints almost account for the total HF interaction energy.

The distinction between polarization and charge transfer in the Morokuma analysis is artificial and very basis set dependent. With a complete basis on each monomer, the orbital relaxation can be completely represented either as polarization or as charge transfer.<sup>38</sup> With a very incomplete basis, much of what has been called charge transfer is really BSSE and must be regarded with suspicion. Also, calculation of polarization without regard to orthogonality is dangerous since with a complete basis, this leads to collapse of the results.<sup>38</sup> Hence, use of free monomer polarizabilities and monomer fields to estimate interaction energies will lead to incorrect results if carried to high order in the multipole expansion.

### 3. Size-Extensive Methods and Pair Correlation Interaction Energies

Calculations using the Møller–Plesset perturbation theory at various orders MP2, MP3, and MP4 were computationally

amenable with the basis developed in this work. The MP3 and the various kinds of MP4 energies were determined using the Gaussian92 package of programs.<sup>39</sup> We could also calculate the interaction energy within the coupled cluster doubles (CCD) approximation with the Gaussian92 package. In Table VI, we have summarized the results for the interaction energy. The MP2 interaction energy of  $-4.53$  kcal/mol seems to be more bound by  $0.2$  kcal/mol than the best methods used (CCD or MP4(SDQ)). In MP2, all virtual orbital contributions are negative so the BSSE is also stabilizing. Higher order contributions can have either sign, and higher order BSSE is actually destabilizing. The best result with this basis set and geometry is probably  $-4.4$  kcal/mol.

To appreciate the delicate effect of electron correlation, the reader should recall that the total dimer valence correlation energy is about  $400$  kcal/mol and this basis set gives an error of about  $60$  kcal/mol in this number. The BSSE of  $0.2$  kcal/mol is very small compared with this possible energy improvement of  $60$  kcal/mol which could be accomplished by completing the basis set. Our objective is to compute the change of  $1$ – $2$  kcal/mol in the total correlation energy when the dimer is formed. This change is tiny ( $0.3\%$ ) compared with the actual error of  $15\%$  in the computed correlation energy. While our basis is not the best that has been used recently, the others are qualitatively similar in the magnitude of the errors.

Density functional models are being advocated for use in molecular structure studies. It is pertinent to measure the performance of these models for this problem. We have compared the MP2 interaction energy (including the  $1s$  core) with some of the recent density functional models. The density functional estimates of the interaction energy were obtained in two ways. In the first, the HF density and the HF exchange energy have been used, and only the correlation energy is estimated with a density functional model. In particular, we have used the correlation energy density functionals of Lee et al.<sup>40</sup> as simplified by Miehlich et al.<sup>41</sup> (LYPc), the generalized gradient approximation of Perdew et al.<sup>42</sup> (PWC) and the reparametrized Colle–Salvetti functional<sup>43</sup> by Caravetta and Clementi.<sup>6</sup> In the second way, we have replaced the HF exchange energy with a density functional model for exchange energy evaluated with the HF density. We have used Becke's exchange energy functional<sup>44</sup> and the generalized gradient approximation to the exchange energy by Perdew et al.<sup>42</sup> The results are summarized in Table VII. The density functional estimates for the binding energy obtained by retaining the HF exchange energy yield values in the range  $-5.2$  to  $-5.9$  kcal/mol. However, when the HF exchange energy term is replaced with the density functional models, values ranging from  $-2.8$  to  $-4.5$  kcal/mol are obtained. Thus, it may be preferable in this case to resort to the first approach. We also note that the monomer correlation energies in the DF model are closer to the true monomer correlation energy, estimated to be  $-0.372$ , than to the MP2 monomer correlation energy. The BSSE with the density functional models comes from the HF part of the energy and is small.

In Table VIII, the MP2 energies and interaction energies with stretched OH bonds ( $+0.02$  Å) and the SCF dimer geometry of Popkie et al.<sup>4</sup> using the same basis set are tabulated. The average OH bond length in the monomer is  $0.02$  Å longer than the equilibrium bond length due to anharmonicity in the zero-point wavefunction for OH stretch. Consequently, this geometry may be more representative of the OH bond length in the dimer after averaging over the high-frequency OH vibrations. There is less than a  $0.1$  kcal/mol increase in the dissociation energy due to the effect of this OH stretching. Therefore, this monomer vibration effect is not responsible for the discrepancy between the calculated and experimental interaction energy.

TABLE V: Previous Calculations for the Binding Energies for the Dimer<sup>a</sup>

basis set	$E_A(\text{HF})$	$E_{cA}(\text{MP2})$	$\Delta E_{ab}(\text{HF})$	$\Delta E_{AB}(\text{HF})$	$\Delta E_{ab}(\text{MP2})$	$\Delta E_{AB}(\text{MP2})$	$\Delta E(\text{CPC})$		ref
6-31G*	-76.0105		-5.43	-4.76					
6-31G**+	-76.0307		-5.08	-4.47					
[9,6,3,2/6,3,2]	-76.0658		-3.78	-3.73					
	-76.0665		-3.75	-3.73					
cc-pVDZ	-76.0268	-0.2016	-5.75	-3.70	-7.30	-3.93	-7.10(-3.73) +4.67(+7.67)	MP4 SDCI	Feller <sup>16</sup>
							-6.84(-3.64) -7.02(-3.71)	QCISD QCISD(T)	
aug-cc-pVDZ	-76.0414	-0.2194	-3.79	-3.55	-5.17	-4.32	-5.20(-4.27) -4.93(-4.12)	MP4 QCISD	Feller <sup>16</sup>
							-5.21(-4.29)	QCISD(T)	
cc-pVTZ	-76.0572	-0.2614	-4.37	-3.55	-6.04	-4.40	-5.95(-4.32) -5.60(-4.15)	MP4 QCISD	Feller <sup>16</sup>
							-5.87(-4.32)	QCISD(T)	
aug-cc-pVTZ	-76.0606	-0.2684	-3.61	-3.54	-5.10	-4.64	-5.14(-4.66) -4.89(-4.46)	MP4 QCISD	Feller <sup>16</sup>
							-5.16(-4.69)	QCISD(T)	
cc-pVQZ	-76.0648	-0.2828	-3.90	-3.57	-5.45	-4.67			Feller <sup>16</sup>
aug-cc-pVQZ	-76.0660	-0.2859	-3.60	-3.56	-5.05	-4.81			Feller <sup>16</sup>
cc-pV5Z	-76.0671	-0.2915	-3.66						Feller <sup>16</sup>
[3s2p1d/2s1p]	-76.0630	-0.200	-4.25	-4.13	-1.63	-0.50	+11.51(+12.25)	SDCI	White and Davidson <sup>15</sup>
							+0.06(+0.78)	MRVARPT	
[4s3p2d/2s1p]			-3.79	-3.64	-5.48	-4.22	-5.57(-5.15)	MP4	Niesar et al. <sup>14</sup>
[4s3p2d/2s1p]									Szczesniak and Scheiner <sup>10</sup>
[5s4p1d/3s1p]	-76.0520		-4.99	-4.11			-6.00(-4.37)	SDCI+Q	Kolos <sup>5</sup>
6-311++G(2df,2p)	-76.0494		-4.41	-3.79	-5.90	-4.25	-5.90(-4.30)	MBPT(4)	Szalewicz et al. <sup>12</sup>
							-5.78(-4.34)	CCSD-T	
6-311++G(3df,3pd)	-76.0587		-4.01	-3.72	-5.08	-4.42		MBPT	Szalewicz et al. <sup>12</sup>
6-311++G(3d2f3p2d)	-76.0658		-3.78	-3.73					Szalewicz et al. <sup>12</sup>
6-311++G(3df,3pd)	-76.0587	-0.2656	-3.8	-3.4	-5.4	-4.5	-5.3(-4.6)	MP4	Frisch et al. <sup>11</sup>
DZP	-76.04680		-5.03	-4.73	-6.36	-5.03	-6.14(-4.81)	MP4	Kim et al. <sup>18</sup>
6-311++G(2d,2p)	-76.05683		-4.08	-3.7	-5.45	-4.45			Kim et al. <sup>18</sup>
[6s3p3d/3s3p]	-76.06328		-3.73	-3.64	-5.24	-4.30	-5.24(-4.30)	MP4	Kim et al. <sup>18</sup>
[7s4p2d1f/4s2p1d]	-76.06439		-4.08	-3.64	-5.24	-4.30			Kim et al. <sup>18</sup>
(13s8p4d2f/8s4p2d)	-76.06756		-3.71	-3.69	-4.99	-4.66			Kim et al. <sup>18</sup>
[6s4p3d1f/4s4p2d]	-76.05207		-3.95						Kim et al. <sup>18</sup>
EZPPBFD			-4.06	-3.76	-8.21	-4.68	-7.82(-4.63)	CEPA	van Duijneveldt et al. <sup>17</sup>
(28s18p8d3f/11s7p2d)	-76.06742		-3.75	-3.73	-5.79	-5.09	-6.01(-5.76)	DFT	Clementi et al. <sup>33</sup>
			-3.67	-3.42					Caravetta and Clementi <sup>6</sup>
cc-pVTZ+							-5.1(-4.4)	MP4	DelBene <sup>19</sup>
cc-pVQZ'							-5.0(-4.7)	MP4	DelBene <sup>19</sup>
Basis-D	-76.06243		-3.72	-3.58	-5.19	-4.57	-4.94(4.40)	MP4	Saebo et al. <sup>23</sup>

<sup>a</sup>  $E_A(\text{HF})$  is the monomer Hartree-Fock energy in au;  $E_{cA}(\text{MP2})$  is the monomer MP2 estimate of the frozen core monomer correlation energy;  $\Delta E_{ab}$  and  $\Delta E_{AB}$  are the binding energies before and after counterpoise correction.  $\Delta E$  is the binding energy by higher level methods, and the number in parentheses includes the counterpoise correction.

TABLE VI: Interaction Energies<sup>a</sup>  $\Delta E$  for the Water Dimer<sup>b</sup>

	HF	MP2	MP3	MP4(D)	MP4(DQ)	MP4(SDQ)	CCD
1 dimer	-152.140 479 57	-0.551 55	-0.555 59	-0.565 09	-0.556 42	-0.560 40	-0.556 52
2 monomer A (Ab) <sup>c</sup>	-76.067 364 79	-0.275 04	-0.277 06	-0.281 87	-0.277 63	-0.279 60	-0.277 68
3 monomer A (A)	-76.067 361 48	-0.274 84	-0.276 91	-0.281 71	-0.277 48	-0.279 45	-0.277 53
4 monomer B (aB)	-76.067 364 09	-0.275 04	-0.277 06	-0.281 86	-0.277 63	-0.279 60	-0.277 68
5 monomer B (B)	-76.067 361 48	-0.274 84	-0.276 91	-0.281 71	-0.277 48	-0.279 45	-0.277 53
6 $\Delta E_{AB}$ (1)-(2)-(4)	-3.609	-4.529	-4.533	-4.465	-4.343	-4.362	-4.335
7 $\Delta E_{ab}$ (1)-(3)-(5)	-3.612	-4.781	-4.728	-4.665	-4.532	-4.555	-4.526
8 BSSE (7)-(6)	-0.003	-0.252	-0.195	-0.200	-0.189	-0.193	-0.191

<sup>a</sup> Values in rows 1-5 are in hartrees, and values in 6-8 are in kcal/mol. <sup>b</sup> The Dyke geometry from Table II was used with the basis set from Table III. <sup>c</sup> Ab indicates basis of A plus ghost basis on b, etc.

In Table IX, the MP2 correlation energy has been partitioned into orbital correlation energies

$$E_c(\text{MP2}) = \sum_{i \leq j}^N \epsilon_{ij} \quad (4)$$

Notice that  $\epsilon_{ij}$  ( $i \neq j$ ) includes all four spin-orbital pair contributions. These pair energies are easily amenable with the MELD program package<sup>45</sup> by selective excitations. In Table

IX, the pair energy and contribution to the interaction energy for each individual pair are presented. All the diagonal pair energies, i.e., the intraorbital pair correlations for the dimer, are smaller in magnitude than the monomer pairs, leading to a repulsive contribution to the total interaction energy. However, the nondiagonal terms which represent the interorbital pair energies are mostly attractive and finally lead to a net total attractive contribution of -0.92 kcal/mol toward the binding energy within the MP2 theory. The largest terms are the interpairs corre-

TABLE VII: Total Correlation Energies and Interaction Energies of the Water Dimer with Density Functional Models<sup>a</sup>

	MP2 <sup>b</sup>	HFxLYPc	HFxPWc	HFxCC-CS'	PWxPWc	BxPWc	PxLYPc <sup>d</sup>
1 dimer	-0.633 35	-0.684 83	-0.707 52	-0.681 65	-0.739 08	-0.776 33	-0.753 64
2 monomer A (Ab)	-0.315 93	-0.340 63	-0.352 48	-0.339 24	-0.368 81	-0.388 76	-0.376 90
3 monomer A (A)	-0.315 63	-0.340 63	-0.352 49	-0.339 24	-0.368 81	-0.388 76	-0.376 90
4 monomer B (aB)	-0.315 91	-0.340 62	-0.352 49	-0.339 21	-0.368 90	-0.388 86	-0.376 99
5 monomer B (B)	-0.315 63	-0.340 62	-0.352 49	-0.339 21	-0.368 90	-0.388 86	-0.376 99
6 $\Delta E_{AB}$	-4.557	-5.86	-5.21	-5.61	-4.47	-2.80	-3.45
7 $\Delta E_{ab}$	-4.923	-5.86	-5.21	-5.62	-4.47	-2.80	-3.46
8 BSSE	-0.366	0.00	0.00	0.01	0.00	0.00	0.01

<sup>a</sup> All values in rows 1–5 are in hartrees, and values in 6–8 are in kcal/mol. <sup>b</sup> All electron. <sup>c</sup> HFxLYPc means the total energy is partitioned as  $E = E(\text{HF}) + E_c[\rho_{\text{HF}}](\text{LYP})$ , only  $E_c$  is tabulated. <sup>d</sup> BxLYPc means the total energy is partitioned as  $E = E(\text{HF}) + [E_x[\rho_{\text{HF}}] - E_x(\text{HF}) + E_c[\rho_{\text{HF}}](\text{LYP})]$ , only the change from the HF energy is tabulated in rows 1–5. <sup>e</sup> Bx: Becke's exchange functional (1988) [ref 44]. PWx: Perdew et al. generalized gradient exchange functional (1992) [ref 42]. PWc: Perdew et al. generalized gradient correlation functional (1992) [ref 42]. LYP: Lee, Yang, and Parr's functional simplified by Miehlich et al. [ref 41]. CC-CS': Caravetta and Clementi's reparametrized Colle and Salvetti functional [ref 6].

TABLE VIII: HF Total Energies, MP2 Correlation Energies,<sup>a</sup> and Interaction Energies for the Water Dimer at Other Geometries

	Popkie <sup>b</sup>			stretched <sup>b</sup>		
	HF	MP2(FC)	MP2(all)	HF	MP2(FC)	MP2(all)
1 dimer	-152.140 638 62	-0.551 31	-0.633 08	-152.136 135 68	-0.554 64	-0.636 31
2 monomer A (Ab) <sup>c</sup>	-76.067 364 50	-0.275 03	-0.315 91	-76.065 192 55	-0.276 55	-0.317 37
3 monomer A (A)	-76.067 361 48	-0.274 84	-0.315 63	-76.065 188 92	-0.276 35	-0.317 07
4 monomer B (aB)	-76.067 362 67	-0.275 03	-0.315 87	-76.065 191 82	-0.276 54	-0.317 35
5 monomer B (B)	-76.067 361 48	-0.274 84	-0.315 63	-76.065 188 92	-0.276 35	-0.317 07
6 $\Delta E_{AB}$	-3.709	-4.493	-4.525	-3.609	-4.582	-4.607
7 $\Delta E_{ab}$	-3.712	-4.735	-4.854	-3.613	-4.830	-4.975
8 BSSE	-0.004	-0.242	-0.330	-0.004	-0.249	-0.368

<sup>a</sup> Values in rows 1–5 are in hartrees, and values in 6–8 are in kcal/mol. <sup>b</sup> See Table II for coordinates. <sup>c</sup> Ab indicates basis of A plus ghost basis on b, etc.

sponding to the bonding orbitals, 5a', 6a', 7a', and 8a'. Interpretation of this table is clouded by orbital mixing that leads to delocalization. For the simple case of two noninteracting monomers, the mixing of doubly occupied monomer orbitals  $\phi_A$  and  $\phi_B$  to give

$$\phi_1 = c_A \phi_A + c_B \phi_B \quad (5)$$

and

$$\phi_2 = -c_B \phi_A + c_A \phi_B \quad (6)$$

leads to

$$\begin{aligned} \epsilon_{11} &= c_A^4 \epsilon_{AA} + c_B^4 \epsilon_{BB} \\ \epsilon_{22} &= c_B^4 \epsilon_{AA} + c_A^4 \epsilon_{BB} \\ \epsilon_{12} &= 2c_A^2 c_B^2 (\epsilon_{AA} + \epsilon_{BB}) \end{aligned} \quad (7)$$

Notice  $\epsilon_{11} + \epsilon_{22} + \epsilon_{12}$  equals  $\epsilon_{AA} + \epsilon_{BB}$ . But if  $\epsilon_{AA}$  and  $\epsilon_{BB}$  are similar in magnitude,  $\epsilon_{11} - \epsilon_{AA}$  and  $\epsilon_{22} - \epsilon_{BB}$  are positive and  $\epsilon_{12}$  is negative.

This effect from orbital mixing makes it difficult to see in Table IX the expected effect on the monomer correlation from the Pauli exclusion effect. In the free monomer A with a complete basis set, the correlation energy would be more negative than for the correlation energy of A with a virtual space orthogonal to the occupied orbitals of B. In a previous paper,<sup>15</sup> this effect was found to contribute a destabilizing correlation effect of about 1 kcal/mol to the total energy. With delocalized canonical HF dimer orbitals, it is difficult to identify this effect unambiguously. This has been confirmed by Vos et al.,<sup>20</sup> who partitioned the MP2 interaction energy into intra- and intermolecular (dispersion) contributions and found these terms are of opposite sign and large. They obtained a 1.3 kcal/mol destabilizing intramolecular correlation effect and a -1.6 kcal/mol stabilizing effect from the intermolecular correlation effect.

#### 4. CI Calculations

In recent studies, a careful CI study has been avoided due to the obvious difficulties in realizing meaningful size-consistent

calculations. It is one of the purposes of this work to attempt such a study. Moreover, we have also assessed the performance of nearly size-consistent methods such as QDVARPT (quasi-degenerate variation perturbation theory<sup>25</sup>), LCC (linearized coupled cluster<sup>26</sup>), and the average pair correction models.<sup>27</sup> It may be recalled that some of the features commonly associated with CI-based methods are the choice of orbitals for the dimer and monomers, the choice of reference space, and finally the selection of a suitable threshold for the perturbative selection of configurations to be actually considered in the correlation energy calculation. These features are both an advantage because of the flexibility introduced by these choices and a disadvantage because careful thought is required to obtain meaningful results.

Because the mixing of monomer orbitals for the dimer is so great, it is difficult to identify a single-monomer orbital with a single-dimer molecular orbital. A solution to this problem in the water dimer has been given by White and Davidson.<sup>15</sup> They localized the occupied HF dimer orbitals using the method of corresponding orbitals. In this method, the overlap between the localized dimer orbitals and the individual monomer occupied HF orbitals is maximized. Racine and Davidson<sup>46</sup> have generalized this procedure to produce dimer virtual orbitals with maximum overlap with monomer (virtual) natural orbitals.

The HFSDCI calculation, a single and doubles excitation CI calculation from the HF ground-state configuration, was done for monomers. From this calculation, frozen natural orbitals for the individual monomers were determined. The frozen natural orbitals retain the original occupied HF orbitals. In the next step, the dimer orbitals were localized using the method of corresponding orbitals. The SDCI calculation was then performed for the monomers and dimer subsequent to a perturbation theory selection of configurations. In the next step, nearly size-extensive methods such as LCC and APC were employed. The LCC and QDVARPT are identical models for the single-reference function. The APC model requires the knowledge of the number of direct electron pairs,  $N/2$ , where  $N$  is the number of correlated electrons for the closed-shell molecule which we chose as 8 for  $\text{H}_2\text{O}$  and 16 for  $(\text{H}_2\text{O})_2$ .



TABLE IX: MP2 Pair Correlation Energy Analysis of the Water Dimer<sup>a</sup>

		1a'	2a'	3a'	4a'	5a'	6a'	7a'	8a'	1a''	2a''
		1a <sub>1</sub> (B)	1a <sub>1</sub> (A)	2a <sub>1</sub> (B)	2a <sub>1</sub> (A)	1b <sub>2</sub> (A)	3a <sub>1</sub> (B)	3a <sub>1</sub> (A)	1b <sub>1</sub> (B)	1b <sub>2</sub> (B)	1b <sub>1</sub> (A)
dimer	1a'	-0.025 05									
monomer		-0.025 05									
dimer-monomer		0.000 00									
dimer	2a'	0.000 00	-0.025 06								
monomer		0.000 00	-0.025 06								
dimer-monomer		0.000 00	0.000 00								
dimer	3a'	-0.003 37	-0.000 02	-0.011 64							
monomer		-0.003 39	0.000 00	-0.011 87							
dimer-monomer		0.000 02	-0.000 02	0.000 23							
dimer	4a'	-0.000 02	-0.003 37	-0.000 29	-0.011 66						
monomer		0.000 00	-0.003 39	0.000 00	-0.011 87						
dimer-monomer		-0.000 02	0.000 02	-0.000 29	0.000 21						
dimer	5a'	-0.000 18	-0.003 15	-0.001 22	-0.024 45	-0.021 01					
monomer		0.000 00	-0.003 33	0.000 00	-0.026 28	-0.023 64					
dimer-monomer		-0.000 18	-0.001 22	0.001 83	0.002 63	0.002 63					
dimer	6a'	-0.003 89	-0.000 32	-0.021 20	-0.002 20	-0.005 36	-0.019 67				
monomer		-0.004 25	0.000 00	-0.023 29	0.000 00	0.000 00	-0.023 33				
dimer-monomer		0.000 36	-0.000 32	0.002 09	-0.002 20	-0.005 36	0.003 66				
dimer	7a'	-0.001 00	-0.003 33	-0.005 22	-0.018 06	-0.029 27	-0.011 35	-0.015 37			
monomer		0.000 00	-0.004 25	0.000 00	-0.023 30	-0.039 52	0.000 00	-0.023 33			
dimer-monomer		-0.001 00	0.000 92	-0.005 22	0.005 24	0.010 25	-0.011 35	0.007 96			
dimer	8a'	-0.004 02	-0.000 86	-0.020 53	-0.005 01	-0.009 28	-0.031 58	-0.015 15	-0.017 10		
monomer		-0.004 84	0.000 00	-0.024 94	0.000 00	0.000 00	-0.040 77	0.000 00	-0.023 57		
dimer-monomer		0.000 82	-0.000 86	0.004 41	-0.005 01	-0.009 28	0.009 19	-0.015 15	0.006 47		
dimer	1a''	-0.003 36	0.000 00	-0.025 98	-0.000 21	-0.001 81	-0.036 08	-0.008 16	-0.031 42	-0.023 58	
monomer		-0.003 33	0.000 00	-0.026 26	0.000 00	0.000 00	-0.039 51	0.000 00	-0.037 85	-0.023 63	
dimer-monomer		-0.000 03	0.000 00	0.000 28	-0.000 21	-0.001 81	0.003 43	-0.008 16	0.006 43	0.000 05	
dimer	2a''	0.000 00	-0.004 80	-0.000 15	-0.024 74	-0.035 44	-0.003 46	-0.031 83	-0.008 42	-0.000 10	-0.023 56
monomer		0.000 00	-0.004 84	0.000 00	-0.024 94	-0.037 85	0.000 00	-0.040 75	0.000 00	-0.023 56	-0.023 56
dimer-monomer		0.000 00	0.000 04	-0.000 15	0.000 20	0.002 41	-0.003 46	0.008 92	-0.008 42	-0.000 10	0.000 00

<sup>a</sup> All values are in hartrees. Monomer calculations are done in the full dimer basis with *C<sub>v</sub>* symmetry.

TABLE X: Single-Reference Single- and Double-Excitations Correlation Energies<sup>a</sup>

method	dimer	supermolecule	monomer B (B)	monomer B (aB)	monomer A (A)	monomer A (Ab) <sup>e</sup>	$\Delta E_{\text{super}}$	$\Delta E_{\text{ab}}$ <sup>f</sup>	$\Delta E_{\text{AB}}$ <sup>f</sup>	BSSE	size consistency
SAC <sup>b</sup>	565 945	565 945	35 349	149 769	36 091	149 679					
PT energy threshold	$1 \times 10^{-13}$	$1 \times 10^{-13}$	0.0	$1 \times 10^{-13}$	0.0	$1 \times 10^{-13}$					
SACs kept	79 222	35 538	35 349	38 159	36 091	39 175					
energy(kept)	-0.807 90	-0.813 10		-0.406 35		-0.406 13					
energy(discarded)	-0.000 25	0.0		0.0		0.0					
SDCI energy	-0.508 77	-0.507 90	-0.267 47	-0.267 59	-0.267 47	-0.267 59	-4.16	12.81	12.96	-0.15	16.98
extrapolated	-0.508 93	-0.507 90	-0.267 47	-0.267 59	-0.267 47	-0.267 59	-4.26	12.71	12.86	-0.15	16.98
Quadruples Correction											
$(1 - C_0)^2$ <sup>c</sup>	-0.566 92	-0.555 99	-0.282 72	-0.282 85	-0.282 72	-0.282 84	-4.20	1.73	1.89	-0.16	5.93
$(1 - C_0)/C_0^2$ <sup>c</sup>	-0.561 92	-0.561 02	-0.283 64	-0.283 77	-0.283 64	-0.283 76	-4.17	-0.25	-0.09	-0.16	3.93
$(1 - C_0^2)/(2C_0^2 - 1)$ <sup>c</sup>	-0.568 08	-0.567 23	-0.284 68	-0.284 81	-0.284 68	-0.284 80	-4.14	-2.81	-2.65	-0.16	1.34
Pople correction	-0.567 33	-0.566 48	-0.284 45	-0.284 58	-0.284 45	-0.284 57	-4.15	-2.62	-2.47	-0.16	1.52
LCC extrapolated	-0.571 31	-0.570 51	-0.285 29	-0.285 42	-0.285 29	-0.285 41	-4.12	-4.07	-3.91	-0.16	0.05
APC extrapolated	-0.561 32	-0.560 45	-0.280 26	-0.280 39	-0.280 26	-0.280 38	-4.16	-4.11	-3.96	-0.16	0.04

<sup>a</sup> The correlation energies are in hartrees, and the interaction energies, BSSE, and size consistency errors are in kcal/mol. The  $\Delta E_{\text{HF}}$  values used are from Table V. All values computed are at the Dyke et al. geometry. <sup>b</sup> SOP are space orbital products, and SAC are spin-adapted configurations of desired spatial symmetry, *C<sub>v</sub>*. <sup>c</sup> Davidson quadruples correction formulae for SDCI wavefunctions. *C<sub>0</sub>* is the coefficient of the Hartree-Fock configuration in the CI wavefunction. <sup>d</sup> All interaction energies are in kcal/mol. All correlation energies are in hartrees. <sup>e</sup> Ab indicates basis of A plus ghost basis on b, etc. <sup>f</sup>  $\Delta E_{\text{AB}}$  and  $\Delta E_{\text{ab}}$  are the binding energies with and without counterpoise correction.

The results of this calculation are summarized in Table X. In order to show how this calculation was really done, more information than usual is included. The total number of single and double excitations is shown. A subset of these was selected based on their second-order Rayleigh-Schrödinger Epstein-Nesbet perturbation contribution

$$\Delta E_1 = - \frac{\langle \phi_1 | H | \phi_{\text{HF}} \rangle^2}{H_{11} - E_{\text{HF}}} \quad (8)$$

using the threshold indicated in Table X for the magnitude of  $\Delta E_1$ . All single excitations were included even though they give  $\Delta E_1 = 0$ . Table X also indicates the number of configurations selected, the total second-order energy contribution from those kept, and the total from those discarded. The actual CI correlation energy is given along with an extrapolated estimate of what could

have been obtained had all configurations been included. This was obtained from

$$\Delta E_{\text{extrap}} = \Delta E_{\text{CI}}(\Delta E_{\text{kept}} + \Delta E_{\text{discarded}})/\Delta E_{\text{kept}} \quad (9)$$

From the table, we can immediately see the huge size consistency error due to SDCI of about 16.98 kcal/mol. The quadruples corrections such as the Davidson<sup>47</sup> and the Pople<sup>48</sup> algorithms provide a substantial improvement of about 15 kcal/mol for the size consistency error; however, for the problem at hand in which the correlation contribution to the interaction energy is about 1–2 kcal/mol, these methods are inadequate.

The LCC and APC methods reduce the size consistency error to about 0.04 kcal/mol. This BSSE would have been zero if all SDCI configurations were included. The total binding energy including the BSSE and size consistency correction effects is

**TABLE XI: List of Monomer A (Proton Donor) Configurations Used To Form the Direct Product Reference Space for the Water Dimer<sup>a</sup>**

1	2a' 23a' 24a' 2	1a'' 2
2	2a' 23a' 24a' 15a' 1	1a'' 2
3	2a' 23a' 14a' 26a' 1	1a'' 2
4	2a' 23a' 24a' 2	1a'' 12a'' 1
5	2a' 23a' 24a' 15a' 1	1a'' 12a'' 1
6	2a' 23a' 14a' 26a' 1	1a'' 22a'' 1
7	2a' 23a' 14a' 15a' 16a' 1	1a'' 2
8	2a' 24a' 26a' 2	1a'' 2
9	2a' 23a' 25a' 2	1a'' 2
10	2a' 23a' 24a' 2	2a'' 2

<sup>a</sup> The configurations are listed in  $C_s$  symmetry. For the correlation with the water monomers in  $C_{2v}$  symmetry, refer to Table V.

about -4.0 kcal/mol for the LCC and APC models. These values are clearly slight underestimates when compared to the MP $n$  and CCD methods that give -4.3 kcal/mol. The APC correlation contribution to binding is only -0.4 kcal/mol compared to -0.7 kcal/mol for MP4 (SDQ) or CCD.

For the multireference single- and double-excitations CI calculations, there are a number of obstacles. The first one is the selection of a size-consistent reference space, and the second one is the suitable choice of orbitals for the reference space and virtual space. Again this problem was solved in various stages. For each of the monomers A and B, we selected the seven most important configurations and their corresponding single-excitation configurations to be the reference. This results in 10 reference configurations for each monomer, listed in Table XI. The direct product of these reference configurations results in a reference space of 100 configurations (322 spin-adapted configurations). However, it is yet necessary to specify the orbitals for the dimer. In the next step, the localized orbitals of the dimer from the previous SDCI calculation were used to perform MR(sd)(S)CI calculation. This is defined here as a CI calculation with the above-defined reference configurations with all single and double excitations within the reference space, plus all single excitations to the virtual space of orbitals. From this calculation, a new set of frozen natural orbitals was derived where the HF space of

orbitals was frozen. With this calculation, we hoped to further improve the reference space of orbitals. The monomer and supermolecule orbitals were constructed in the same way. In the next step, the localization was performed separately for three sets of orbitals: the HF orbitals, the correlating orbitals of the reference set, and the virtual orbitals.

In the next stage, the dimer CI calculation was performed first to derive an energy threshold for perturbative selection of configurations for the dimer. Monomers and the supermolecule calculations were then done with the same threshold. A total of about 154 million configurations were scanned by perturbation theory for the dimer MRSDCI calculation.<sup>45</sup> Present limitations of the program only permit up to eight open shells in a particular configuration. However, with the current choice of reference set, configurations with a maximum of 12 open shells would be required for a complete MRSDCI calculation. The results of our MRSDCI calculations have been summarized in Table XII. From the tabulation, the inadequacies of the multireference CI methods become transparent. The first point to note is that though the BSSE and size consistency errors have been reduced, the choice of the reference set of orbitals for the dimer seems to be an inadequate one. This can be seen by comparing the dimer and supermolecule MRSDCI correlation energies. Furthermore, the subsequent attempts by various models such as QDVARPT, MRLCC, and MR-APC to correct for this inadequacy are unsuccessful. The best energy estimate including BSSE and size consistency correction from this multireference approach is -3.1 kcal/mol. The correlation contribution to the result is actually +0.51 kcal/mol repulsive. None of the multireference results are more bound than the HF energy. All of the unlinked cluster correction methods used in this table are derived for a Hartree-Fock reference and do not necessarily apply to a multireference wavefunction, especially when no other large coefficients are involved. On the other hand, the extrapolated MRSDCI result with BSSE correction ( $\Delta E_{\text{super}}$  and BSSE) is already -3.1 kcal/mol and is unchanged by the attempts to correct for higher excitations. Clearly, even though the actual MRSDCI correlation energy is improved by 20 kcal/mol compared to SDCI, the delicate

**TABLE XII: Multireference Single- and Double-Excitations Correlation Energies with CI, QDVARPT, LCC, and APC Methods for the Water Dimer<sup>a</sup>**

method	dimer	supermolecule	monomer B (B)	monomer B (aB)	monomer A (A)	monomer A (Ab)	$\Delta E_{\text{super}}$	$\Delta E_{\text{ab}}$	$\Delta E_{\text{AB}}$	BSSE	size consistency
MR(sd)(S)CI energy <sup>b</sup>	-0.200 16	-0.200 92	-0.103 89	-0.103 89	-0.103 89	-0.103 90	-3.13	1.17	1.18	-0.01	4.30
reference energy	-0.156 04	-0.158 39	-0.078 81	-0.078 81	-0.078 81	-0.078 85	-2.14	-2.62	-2.60	-0.02	-0.48
generated SAC <sup>c</sup>	154 750 190	154 750 190	555 595	2410 937	569 371	2410 937					
PT energy threshold	$2.5 \times 10^{-7}$	$2.5 \times 10^{-7}$	$2.5 \times 10^{-7}$	$2.5 \times 10^{-7}$	$2.5 \times 10^{-7}$	$2.5 \times 10^{-7}$					
SACs(kept)	92 987	89 416	26 066	30 050	26 109	29 181					
energy(kept)	-0.507 16	-0.507 36	-0.251 36	-0.251 91	-0.252 04	-0.251 13					
energy(discarded)	-0.005 19	-0.003 88	-0.001 22	-0.001 79	-0.001 23	-0.001 77					
MRSDCI	-0.539 23	-0.540 63	-0.277 81	-0.277 61	-0.277 86	-0.277 59	-2.74	6.71	6.41	0.30	9.45
extrapolated	-0.543 15	-0.543 55	-0.278 78	-0.279 02	-0.278 84	-0.278 99	-3.36	5.47	5.72	-0.25	8.83
Quadruples Correction											
(1 - $C_0^2$ ) <sup>d</sup>	-0.570 05	-0.570 43	-0.286 82	-0.287 07	-0.286 88	-0.287 01	-3.37	-1.33	-1.08	-0.25	2.05
(1 - $C_0^2$ )/ $C_0^2$ <sup>d</sup>	-0.572 06	-0.572 45	-0.287 16	-0.287 41	-0.287 21	-0.287 35	-3.37	-2.16	-1.92	-0.25	1.20
(1 - $C_0^2$ )/ $(2C_0^2 - 1)$ <sup>d</sup>	-0.574 40	-0.574 79	-0.287 52	-0.287 78	-0.287 58	-0.287 71	-3.36	-3.17	-2.92	-0.25	0.19
Pople correction	-0.574 13	-0.574 53	-0.287 44	-0.287 70	-0.287 50	-0.287 63	-3.36	-3.10	-2.85	-0.25	0.26
QDVARPT	-0.575 40	-0.575 84	-0.287 71	-0.287 95	-0.287 77	-0.287 90	-3.33	-3.56	-3.33	-0.24	-0.23
extrap.											
MRLCC	-0.575 06	-0.575 51	-0.287 52	-0.287 77	-0.287 58	-0.287 71	-3.33	-3.59	-3.35	-0.24	-0.25
extrap.											
APC extrap.	-0.570 65	-0.571 07	-0.285 29	-0.285 54	-0.285 35	-0.285 49	-3.34	-3.62	-3.37	-0.24	-0.27

<sup>a</sup> The correlation energies are in hartrees, and the interaction energies, BSSE, and size consistency errors are in kcal/mol. The  $\Delta E_{\text{HF}}$  values used are from Table V. All values computed are at the Dyke et al. geometry. <sup>b</sup> MR(sd)(S)CI is a multireference CI calculation, including all single and double excitations within the reference space and all single excitations from the reference space into the virtual space. <sup>c</sup> SAC are spin-adapted configurations of desired spatial symmetry,  $C_s$ . <sup>d</sup> Davidson quadruples correction formulae for SDCI wavefunctions.  $C_0^2$  is the sum of the coefficients from the reference space. <sup>e</sup> All interaction energies are in kcal/mol. All correlation energies are in hartrees. <sup>f</sup> Ab indicates basis of A plus ghost basis on b, etc. <sup>g</sup>  $\Delta E_{\text{AB}}$  and  $\Delta E_{\text{ab}}$  are binding energies with and without counterpoise correction.

cancellation of huge errors required to get a reasonable binding energy has been lost. From Table VI, it appears that the actual total correlation energy with this basis is  $-0.560$  for the dimer and  $-0.279$  for the monomer.

The single-reference APC result is very close to this, but the SDCI result is low (as expected). The quadruples correction actually overestimates the quadruples effect by 10–20%. The multireference MRSDCI is much closer for the monomer but is not as good for the dimer (again as expected). All of the quadruples methods including QDVARPT, MRLCC, and APC overestimate the correlation energy in both the monomer and dimer.

## 5. Summary

We have performed a sequence of computations of the binding energy of the water dimer employing a variety of ab initio methods with one fixed basis set. Our studies give a dimerization energy near the values obtained in recent studies of Kim et al.<sup>18</sup> and that of van Duijneveldt et al.<sup>17</sup> This basis was unable to obtain the larger values obtained by Feller<sup>16</sup> and Clementi et al.<sup>33</sup> Further accurate treatments will require size-consistent approaches such as many body perturbation methods and multireference coupled cluster involving higher levels of excitations. Larger basis sets that would recover 99% of the correlation energy rather than just 85% are also needed.

We have demonstrated that it seems inevitable that CI-based methods have to be abandoned in preference to size-consistent methods such as MP $n$  and coupled cluster for problems such as the binding energy of water dimer. The necessary precision of 0.1% in the calculated CI correlation energy for dimers and monomers cannot be achieved while using large basis sets that generate such large numbers of configurations.

**Acknowledgment.** This work was supported by Grant No. PHS R01 GM 34081 from the National Institutes of Health.

## References and Notes

- (1) Curtiss, L. A.; Frurip, D. J.; Blander, M. J. *Chem. Phys.* **1979**, *71*, 2703.
- (2) Reimers, J.; Watts, R.; Klein, M. *Chem. Phys.* **1982**, *64*, 95.
- (3) Clementi, E.; Popkie, H. J. *Chem. Phys.* **1972**, *57*, 1077.
- (4) Popkie, H.; Kistenmacher, H.; Clementi, E. J. *Chem. Phys.* **1973**, *59*, 1325.
- (5) Kolos, W. *Theoret. Chim. Acta* **1979**, *51*, 219.
- (6) Caravetta, V.; Clementi, E. J. *Chem. Phys.* **1984**, *81*, 2646.
- (7) Yoon, B. J.; Morokuma, K.; Davidson, E. R. *J. Chem. Phys.* **1985**, *83*, 1223.
- (8) Baum, J. O.; Finney, J. L. *Molec. Phys.* **1985**, *55*, 1097.
- (9) Frisch, M. J.; Pople, J. A.; Del Bene, J. E. *J. Phys. Chem.* **1985**, *89*, 3664.
- (10) Szczesniak, M. M.; Scheiner, S. J. *Chem. Phys.* **1986**, *84*, 6328.
- (11) Frisch, M. J.; Del Bene, J. E.; Binkley, J. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 2279.
- (12) Szalewicz, K.; Cole, S. J.; Kolos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662.
- (13) Del Bene, J. E. *J. Chem. Phys.* **1987**, *86*, 2110.
- (14) Niesar, U.; Corongiu, G.; Huang, M.-J.; Dupuis, M.; Clementi, E. *Int. J. Quant. Chem. Symp.* **1989**, *23*, 421.
- (15) White, J. C.; Davidson, E. R. *J. Chem. Phys.* **1990**, *93*, 8029.
- (16) Feller, D. J. *Chem. Phys.* **1992**, *96*, 6104.
- (17) van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. J. *Chem. Phys.* **1992**, *97*, 5019.
- (18) Kim, K. S.; Mhin, B. J.; Choi, U.-C.; Lee, K. J. *Chem. Phys.* **1992**, *97*, 6649.
- (19) Del Bene, J. E. *Int. J. Quant. Chem. Symp.* **1992**, *26*, 527.
- (20) Vos, R. J.; Hendriks, R.; van Duijneveldt, F. B. *J. Comput. Chem.* **1990**, *11*, 1.
- (21) Krijn, M. P. C.; Feil, D. J. *Chem. Phys.* **1988**, *89*, 5787.
- (22) Rybak, S.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **1991**, *95*, 6576.
- (23) Saebo, S.; Tong, W.; Pulay, P. J. *Chem. Phys.* **1993**, *98*, 2170.
- (24) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (25) Cave, R. J.; Davidson, E. R. *J. Chem. Phys.* **1988**, *89*, 6798. Racine, S. Ph.D. Dissertation, Indiana University, Bloomington, 1992 (unpublished).
- (26) Laidig, W. D.; Bartlett, R. J. *Chem. Phys. Lett.* **1984**, *104*, 424.
- (27) Murray, C.; Racine, S. C.; Davidson, E. R. *Int. J. Quant. Chem.* **1992**, *42*, 273.
- (28) Gdanitz, J.; Ahlrichs, R. *Chem. Phys. Lett.* **1988**, *143*, 413.
- (29) Partridge, H. NASA Technical Memorandum 101044, 1989.
- (30) van Duijneveldt, F. B. IBM Research Report No. RJ-945, 1971.
- (31) Feller, D.; Davidson, E. R. *J. Chem. Phys.* **1985**, *82*, 4135.
- (32) Feller, D.; Boyle, C. M.; Davidson, E. R. *J. Chem. Phys.* **1987**, *86*, 3424.
- (33) Clementi, E.; Corongiu, G.; Stradella, O. G. *Modern Techniques in Computational Chemistry: MOTECC-91*; ESCOM Science Publishers: Leiden, 1991; pp 295–380.
- (34) Bawagan, A. O.; Brion, C. E.; Davidson, E. R.; Feller, D. *Chem. Phys.* **1987**, *113*, 19.
- (35) Dyke, T. R.; Mack, K. M.; Muentner, J. S. *J. Chem. Phys.* **1977**, *66*, 498.
- (36) Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980**, *72*, 5062.
- (37) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc.* **1977**, *99*, 1316.
- (38) Gutowski, M.; Piela, L. *Molec. Phys.* **1988**, *64*, 337. Chalasinski, G.; Gutowski, M. *Chem. Rev.* **1988**, *88*, 943. Frey, R. F.; Davidson, E. R. *J. Chem. Phys.* **1989**, *90*, 5555. Kunze, K. L.; Davidson, E. R. *J. Phys. Chem.* **1992**, *96*, 2129.
- (39) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92, Revision C*; Gaussian, Inc.: Pittsburgh PA, 1992.
- (40) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (41) Miehlisch, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.
- (42) Perdew, J. P.; Chevary, J.; Vosko, S. H.; Jackson, K.; Pederson, M. R.; Singh, D. J.; Fiolhas, C. *Phys. Rev. B* **1992**, *46*, 6671. Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
- (43) Colle, R.; Salvetti, O. *Theoret. Chim. Acta* **1975**, *37*, 329; **1979**, *53*, 55; *J. Chem. Phys.* **1983**, *79*, 1404.
- (44) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.
- (45) The MELD series of electronic structure codes were developed by L. E. McMurchie, S. T. Elbert, S. R. Langhoff, and E. R. Davidson and was extensively modified by D. Feller and D. C. Rawlings.
- (46) Racine, S.; Davidson, E. R. *J. Phys. Chem.*, preceding paper in this issue.
- (47) Davidson, E. R.; Silver, D. W. *Chem. Phys. Lett.* **1977**, *53*, 403.
- (48) Pople, J. A.; Seeger, R.; Krishnan, R. *Int. J. Quant. Chem. Symp.* **1977**, *11*, 149.