FEATURE ARTICLE

Ab Initio Calculation of Nonbonded Interactions: Are We There Yet?

A. K. Rappé* and E. R. Bernstein*

Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523-1872

Received: March 7, 2000; In Final Form: April 18, 2000

We present calculations for the nonbonded interactions in the dimeric complexes: methane dimer, ammonia dimer, water dimer, $H_2O\cdot(NH_3)$, $CH_4\cdot(NH_3)$, and $(FHF)^-$ as a function of theory level (HF, DFT(B3LYP), MP2, LMP2, MP3, MP4, CCSD(T), and others) and basis set (6-31G**, cc-pVXZ, X = D, T, Q, 5). Dimer minimum energy structures are determined at the MP2 theory level for the cc-pVTZ basis set employing analytical second derivatives. For HF and DFT levels of theory, methane dimer and one structure of $CH_4\cdot(NH_3)$ are not bound. The basis set superposition error (BSSE) begins to converge (becomes systematically small) for basis sets larger than cc-pVTZ. For hydrogen-bonded systems, most levels of theory seem to give reasonable estimates of the experimentally known binding energies, but here, too, the BSSE overwhelms the reliability of the binding energies for the smaller basis sets. The $CH_4\cdot(NH_3)$ dimer has two minimum energy conformations with similar binding energies, but very different BSSE values especially for small basis sets (cc-pVXZ, X ≤ T). On the basis of these calculations, we present a discussion of ab initio calculations of nonbonded interactions for molecules, such as phenethylamine, that have different conformations. Suggestions for possible next steps in the calculation of nonbonded interactions are presented.

I. Introduction

Nonbonded interactions play a central role in chemistry and biology.[1−10] They are responsible for polymer properties, tertiary structure in macromolecules, molecular fluid and molecular solid properties, conformational structures and preferences of molecules, and energy transfer between molecules and between molecular moieties, and, in a broad sense, they are even responsible for chemical reactions.

Empirical and semiempirical potential energy functions (e.g., Lennard-Jones−Coulomb, exponential-six, and others) have long been quite successful at predicting and modeling nonbonded interactions between molecules and between molecular moieties.[11] Nonetheless, one can suggest that a fundamental understanding of these nonbonded interaction only arises from an accurate ab initio evaluation of their size, radial and angular dependence, and importance.

The issues we address in this Feature Article are how close are we to this goal of ab initio calculation of nonbonded interactions, and what is or are the approaches one must pursue to get reliable and believable, chemically interesting (ca. 1 kcal/mol) nonbonded interaction energies and properties from ab initio calculations. While a number of benchmark calculations for various systems and at various levels of theory have previously appeared,[12,13] we present in this report a series of systems and theory levels that systematically develop the adequacy of ab initio approaches to nonbonded interactions.

The main systems we discuss are various small dimers that can be treated at a very high level of theory. These include methane dimer, ammonia dimer, water dimer, $H_2O\cdot(NH_3)$, $CH_4\cdot(NH_3)$, and $(FHF)^-$. Additionally, we address conformational calculations with a model study of phenethylamine from which to estimate the nonbonded interactions that control tertiary structure for this molecule. Basis sets employed in these calculations include Pople's[14] 6-31g** and Dunning's[15] cc-pVXZ (X = D, T, Q, 5). The theoretical models employed for the benchmark calculations are Hartree−Fock (HF), density functional theory (DFT-B3LYP),[16] Møller−Plesset perturbation methods[17] (MP2, LMP2, MP3, MP4D, MP4DQ, and MP4SDQ), coupled-cluster singles and doubles (CCSD), and CCSD plus triple excitations included perturbatively (CCSD(T)).[18] With these basis sets and theory algorithms, estimates of the basis set superposition error (BSSE) by the counterpoise method[19] are made for the equilibrium dimer configurations (calculated at the MP2/cc-pVTZ level of theory).

These benchmark studies are informative because the dimers chosen represent the range from dispersion interactions, through hydrogen-bonding interactions, to near-bonding interactions present in low-barrier hydrogen-bonding (LBHB) systems and various intermediate interaction situations. Binding energies vary from ∼150 to 18 000 cm^{-1} (less than 0.5 to ∼50 kcal/mol), with the counterpoise approximation to the BSSE energy varying from ∼0.1 to ∼15 kcal/mol. Model calculations of the counterpoise BSSE correction for conformational energy differences show that the size of the BSSE correction is strongly dependent upon both cluster and molecular structure.

The systematic approach taken in this work for the simple dimer systems suggests that a high level of theory is required to generate reliable, chemically significant results for nonbonded interactions. In particular, at least an MP2 algorithm (i.e., MP2, MP3, MP4, CCSD(T), CASSCF/MP2, ...) with at least a cc-pVTZ basis set is required to begin to observe consistent and convergent behavior for both dimer binding energy and the

* To whom correspondence should be addressed.

10.1021/jp0008997 CCC: $19.00
Published on Web 06/07/2000

TABLE 1: Dimer Metric Parameters (Distances in Å, Angles in deg)

molecule	value	molecule	value
$CH_4 \cdot (CH_4)$		$H_2O \cdot (NH_3)$	
R(C−H)	1.085	R(O−H)	0.957
R(C−H′)	1.086	R(O−H′)	0.971
θ(H−C−H)	109.41	R(N−H)	1.012
θ(H−C−H′)	109.59	R(N−H′)	1.012
ϕ(H−C_a−C_b−H)	60.0	θ(H−O−H′)	103.72
first $CH_4 \cdot (NH_3)$		θ(H−N−H)	106.75
R(C−H)	1.084	θ(H−N−H′)	106.50
R(C−H′)	1.086	ϕ(H−O−N−H′)	180.0
R(C−H″)	1.086	ϕ(H−O−N−H)	64.06
R(N−H)	1.014	$H_2O \cdot (H_2O)$	
R(N−H′)	1.012	R(O_a−H)	0.959
θ(H−C−H)	109.41	R(O_a−H′)	0.966
θ(H′-C−H″)	109.41	R(O_b−H)	0.960
θ(H−N−H)	109.41	θ(H−O_a−H′)	103.62
θ(H−N−H′)	109.41	θ(H−O_b−H)	104.04
second $CH_4 \cdot (NH_3)$		ϕ(H−O_a−O_b−H)	124.92
R(C−H)	1.086	$(FHF)^-$	
R(C−H′)	1.086	R(F−H)	1.138
R(N−H)	1.011		
θ(H′−N−H)	112.50		
ϕ(H−C−N−H)	167.41		
$NH_3 \cdot (NH_3)$			
R(N−H)	1.011		
R(N−H′)	1.013		
θ(H−N−H)	105.99		
θ(H−N−H)	106.48		
θ(N_a−N_b−H′)	41.92		
ϕ(H−N_a−N_b−H′)	0.0		

counterpoise approximation to the BSSE correction. At the MP2/cc-pVTZ theory level, the BSSE correction is roughly 30−50% of the binding energy and at the MP2/cc-pV5Z theory level, the BSSE correction is roughly 10% of the dimer binding energy. For the dimers mentioned above, the binding energy is in essential agreement with both the experimental results and the Lennard-Jones−Coulomb potential energy surface model.[20]

For a conformational system such as phenethylamine, the counterpoise model BSSE correction is 150% of the magnitude of the conformational energy difference between the trans and gauche conformations at the MP2/cc-pVTZ level of theory.

Two conclusions are clear from these studies: (1) the major difficulty with ab initio calculation of nonbonded interactions at the theory levels explored herein lies with inadequate basis sets; and (2) systems with more than 3 or 4 heavy atoms (C, N, O, ...) are, at present, too expensive to treat accurately for both hardware and software reasons.

Given the above, we discuss at the end of this report some possible basis set improvements for the calculation of nonbonded interactions based on floating spherical Gaussian orbitals.

II. Methods and Results

There are several families of Gaussian basis sets that have been used to compute intermolecular interactions. The two most popular are the Pople supplemented 6-31g[14] and Dunning cc-pVXZ sets.[15] We have focused on the cc-pVXZ sets because of the asymptotic convergence they display and the systematic augmentation strategies that have been reported.

Metric parameters for methane dimer, methane-ammonia (two conformations), ammonia dimer, ammonia−water, water dimer, and hydrogen fluoride plus fluoride ion are collected in Table 1 and Figure 1. The geometric results are obtained using a polarized valence triple-ζ basis set (cc-pVTZ)[15] and a second-order Møller−Plesset wave function (MP2).[17] Analytic Hessian matrices are computed at each geometric point to facilitate the optimization of the low-frequency intermolecular degrees of freedom. Cartesian coordinates are provided as Supporting Information.

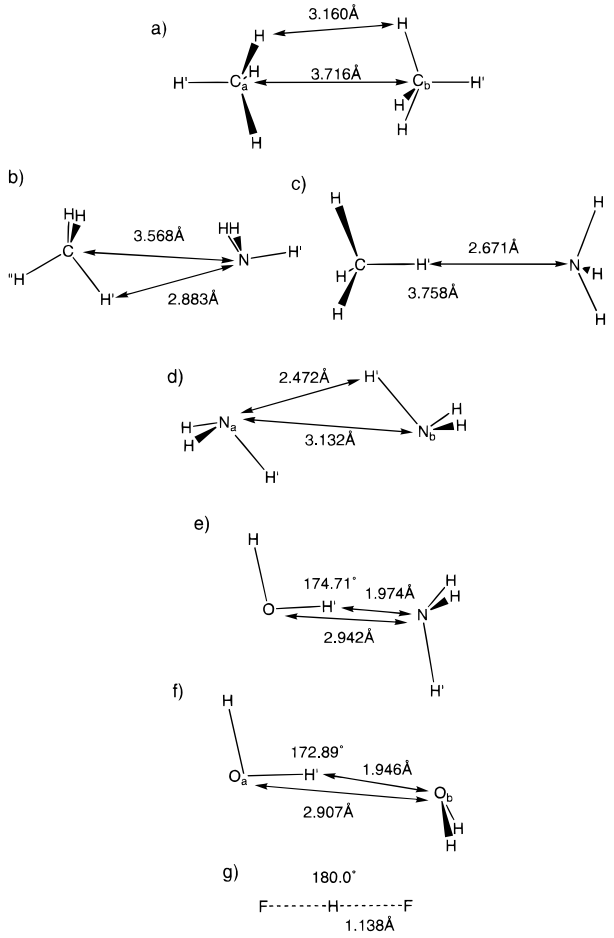

Figure 1. Geometries for the dimers of interest in this report, as indicated in the figure.

Intermolecular binding energies for each complex are computed with Hartree−Fock (HF) and MP2 wave function approaches and the B3LYP gradient corrected density functional method.[16] The Boys−Bernardi counterpoise method,[19] applied at the dimer geometry, is used to account for BSSE. The reported binding energies are with respect to computations at the cc-pVTZ MP2 monomer geometries. Binding energies and BSSE as a function of basis set and method are collected in Figures 2−8 and Table 2. Total energies are provided as Supporting Information.

The convergence characteristics of the Møller−Plesset perturbation series as a function of basis set is examined for the methane−ammonia cluster, second conformation (Figure 1c). Results for the coupled cluster singles and doubles approach (CCSD, CCSD(T))[18] are also provided in Figure 9 and Table 3.

The trans and gauche conformations of phenethylamine are computed with a cc-pVTZ(-f) basis and an LMP2 wave function. Molecular conformations place portions of a molecule that are distant in one conformation close together in other conformations. This also causes basis functions that are distant in one conformation to be close together in others. This geometry variation leads to a conformationally differential BSSE that should be of comparable magnitude to that in intermolecular complexes. Computation of the conformationally differential BSSE for a discrete molecule is problematic. How does one

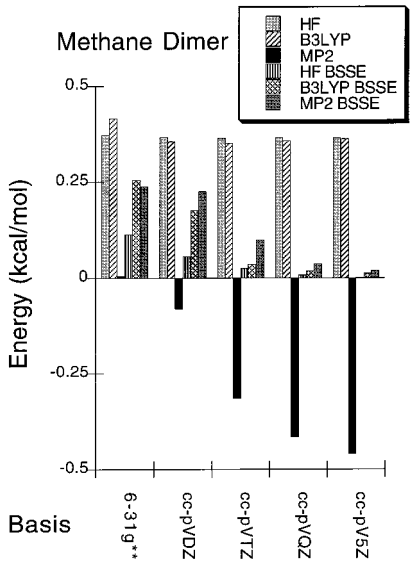

Figure 2. Binding energies (BE) and BSSE for methane dimer in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

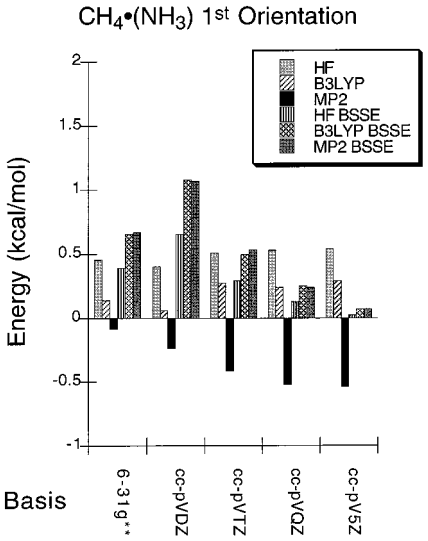

Figure 3. Binding energies (BE) and BSSE for $CH_4 \cdot (NH_3)$, first orientation (see Figures 1b,c) in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

define the fragments? In order to avoid spin polarization issues present in radicals, we have chosen to estimate the BSSE in phenethylamine by breaking the phen−ethyl bond and capping the radical centers with hydrogen atoms, placed in idealized positions. Figure 10 shows these different structures, and Figure 11 presents the net trans and gauche conformer energy differences.

III. Discussion

There is an extensive literature on the ab initio computation of intermolecular interactions. This work is reviewed in section IIIB. In order to provide a contextual framework for this literature, the results of a systematic series of computations on intermolecular complexes that range from strictly van der Waals to a strongly bound LBHB system are discussed in section IIIA. In section IIIC we discuss the comparison between theory and experiment for dimer binding energies.

A. Framework Computations. While Figures 2−8 and Tables 2 and 3 speak for themselves, we present some comments below concerning the calculated data.

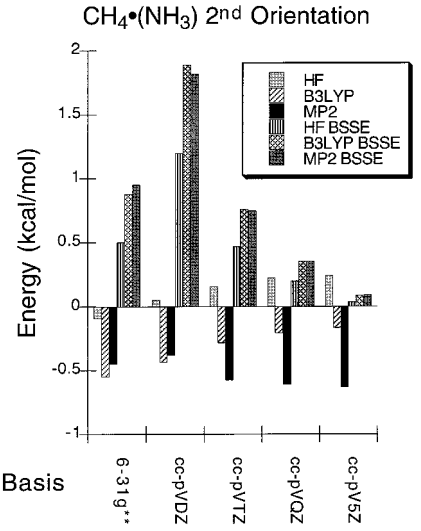

Figure 4. Binding energies (BE) and BSSE for $CH_4 \cdot (NH_3)$, second orientation (see Figures 1b,c) in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

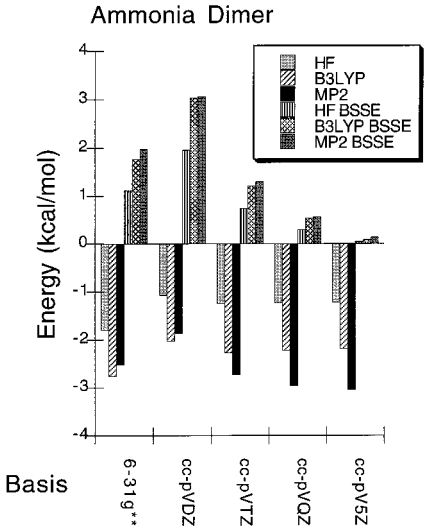

Figure 5. Binding energies (BE) and BSSE for ammonia in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

1. Methane dimer is bound for the MP2 level of theory, but not for the HF and B3LYP levels.

2. Dispersion interactions are not given by the DFT B3LYP level of theory.

3. BSSE is comparable for MP2 and B3LYP for all dimers considered.

4. For the cc-pV5Z basis set, the BSSE is less than 10% of the dimer binding energy.

5. Binding energies for the dimers are convergent for the MP2/cc-pVXZ, X = T, Q, 5, theory level.

6. BSSE is convergent for the MP2/cc-pVXZ, X = T, Q, 5, theory level.

7. For all dimers addressed herein, the BSSE at the polarized VDZ basis set level of theory is nearly equal to or greater than the binding energy and will certainly affect the calculated geometry.

8. The contrast between the BSSE and the binding energy for water dimer and methane dimer is quite significant. For water dimer (an electrostatic or hydrogen-bonding dominated system), the binding energy is remarkably independent of both the basis set and the theory level. For methane dimer (a dispersion-

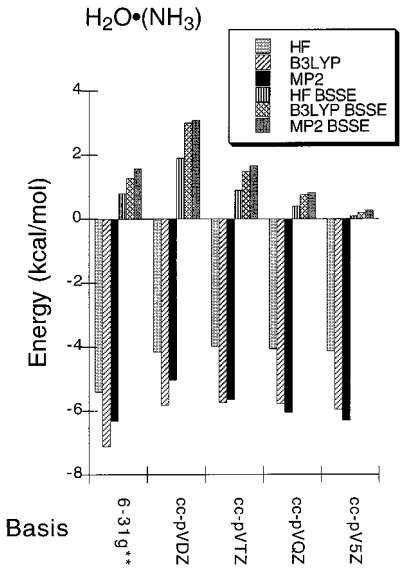

Figure 6. Binding energies (BE) and BSSE for $H_2O \cdot (NH_3)$ in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

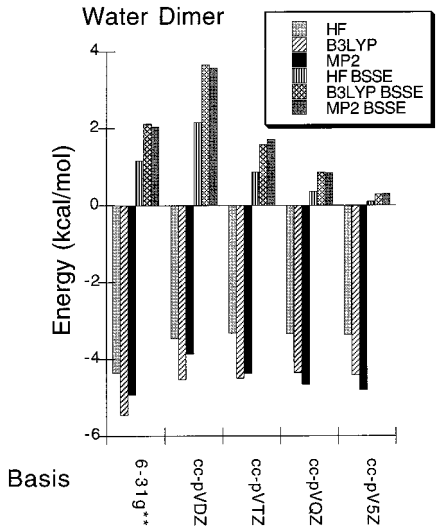

Figure 7. Binding energies (BE) and BSSE for water dimer in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

dominated system), binding is only observed for the MP2 theory level and strongly dependent upon basis set. The BSSE in both instances at the cc-pVTZ basis set is 33% for methane dimer and 40% for water dimer of the respective binding energies.

9. In more detail from the figures and tables, one can conclude the following:

a. The only attractive interaction for the methane dimer is dispersion;

b. For $CH_4 \cdot (NH_3)$ orientation 1 (Figure 1b), dispersion is the dominant interaction;

c. For $CH_4 \cdot (NH_3)$ orientation 2 (Figure 1c), electrostatic (hydrogen bonding) interactions are beginning to become important;

d. For water dimer, $NH_3 \cdot (H_2O)$, ammonia dimer, and $(FHF)^-$, the major component of the nonbonded interaction is electrostatic or hydrogen bonding; and

e. For $CH_4 \cdot (NH_3)$, the conformational preference, binding energy and BSSE are all method and basis set dependent.

10. For $CH_4 \cdot (NH_3)$ orientation 2, the observations made at the MP2 level of theory continue as the perturbation series and

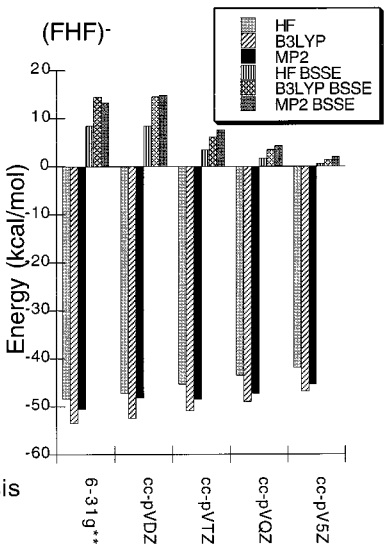

Figure 8. Binding energies (BE) and BSSE for $(FHF)^-$ in the order BE[HF, B3LYP, MP2] and BSSE[HF, B3LYP, MP2] for the indicated basis sets.

basis set are extended. Note that at the cc-pVDZ basis set the BSSE is consistently a factor of 4 larger than the binding energy and at the cc-pVQZ the BSSE is about 33% of the binding energy (Figure 9).

11. In all intermolecular interactions, Coulomb, dispersion, and induction effects are present. Since dispersion interactions are not accounted for in DFT at present, but are given at an MP2 or higher theory level, one should expect that the DFT binding energy should always be less than the MP2 binding energy. If we consider intermolecular interactions to be atom pair wise additive, this difference should be roughly 0.5 kcal/mol for each first-row interacting atomic pair.

Large biologically interesting molecules can be conceptually assembled from the small dimer structures discussed above. One pharmacologically important example is phenethylamine (see Figure 10 for the molecular structure). There are three conformations of the ethyl portion of phenethylamine, trans, and gauche(+) and gauche(−). In the extended or trans conformation of phenethylamine the amine and aromatic functional groups are remote. In the two equivalent gauche conformations the amine and aromatic functional groups are in van der Waals contact (see Figure 10). The conformational preferences of phenethylamine should follow the trends established above for $CH_4 \cdot (NH_3)$. Computed conformational trans−gauche energy differences and differential BSSE data for phenethylamine are collected in Figure 11. As for $CH_4 \cdot (NH_3)$, HF and B3LYP underestimate the conformational energy difference due to a lack of dispersion. The trend with basis set for the conformational energy difference parallels that found for $CH_4 \cdot (NH_3)$ and suggests that a vastly larger basis (cc-pV5Z) would be needed to correctly estimate the trans−gauche energy difference of phenethylamine.

B. Literature Survey. The computation of intermolecular interactions is an active area of theoretical research. In this survey, we classify systems as being van der Waals or hydrogen bonding. For the discussion below, hydrogen-bonding systems are separated into weak hydrogen bonding, classical hydrogen bonding, and LBHB categories.

i. Hydrocarbon/van der Waals. Several theoretical studies of the methane dimer,[21−23] numerous studies of the benzene dimer,[24−27] and a few studies of other hydrocarbon pairs,[28,29] have appeared. For methane dimer, for which basis and

TABLE 2: HF, B3LYP, and MP2 Dimer Binding Energies and BSSE (in kcal/mol)

molecule/method	binding energy	BSSE	molecule/method	binding energy	BSSE
CH_4·(CH_4) 6-31g/HF	0.373	0.113	NH_3·(NH_3) cc-pVTZ/MP2	−2.726	1.299
CH_4·(CH_4) 6-31g/B3LYP	0.417	0.255	NH_3·(NH_3) cc-pVQZ/HF	−1.230	0.295
CH_4·(CH_4) 6-31g/MP2	0.005	0.239	NH_3·(NH_3) cc-pVQZ/B3LYP	−2.223	0.534
CH_4·(CH_4) cc-pVDZ/HF	0.368	0.056	NH_3·(NH_3) cc-pVQZ/MP2	−2.957	0.562
CH_4·(CH_4) cc-pVDZ/B3LYP	0.356	0.176	NH_3·(NH_3) cc-pV5Z/HF	−1.222	0.042
CH_4·(CH_4) cc-pVDZ/MP2	−0.080	0.225	NH_3·(NH_3) cc-pV5Z/B3LYP	−2.177	0.078
CH_4·(CH_4) cc-pVTZ/HF	0.364	0.025	NH_3·(NH_3) cc-pV5Z/MP2	−3.048	0.133
CH_4·(CH_4) cc-pVTZ/B3LYP	0.351	0.035	H_2O·(NH_3) 6-31g/HF	−5.400	0.791
CH_4·(CH_4) cc-pVTZ/MP2	−0.315	0.099	H_2O·(NH_3) 6-31g/B3LYP	−7.114	1.269
CH_4·(CH_4) cc-pVQZ/HF	0.366	0.008	H_2O·(NH_3) 6-31g/MP2	−6.326	1.576
CH_4·(CH_4) cc-pVQZ/B3LYP	0.357	0.019	H_2O·(NH_3) cc-pVDZ/HF	−4.164	1.903
CH_4·(CH_4) cc-pVQZ/MP2	−0.416	0.036	H_2O·(NH_3) cc-pVDZ/B3LYP	−5.808	2.997
CH_4·(CH_4) cc-pV5Z/HF	0.365	0.001	H_2O·(NH_3) cc-pVDZ/MP2	−5.028	3.079
CH_4·(CH_4) cc-pV5Z/B3LYP	0.362	0.012	H_2O·(NH_3) cc-pVTZ/HF	−3.994	0.891
CH_4·(CH_4) cc-pV5Z/MP2	−0.459	0.019	H_2O·(NH_3) cc-pVTZ/B3LYP	−5.739	1.484
first CH_4·(NH_3) 6-31g/HF	0.454	0.394	H_2O·(NH_3) cc-pVTZ/MP2	−5.641	1.666
first CH_4·(NH_3) 6-31g/B3LYP	0.141	0.658	H_2O·(NH_3) cc-pVQZ/HF	−4.060	0.389
first CH_4·(NH_3) 6-31g/MP2	−0.084	0.667	H_2O·(NH_3) cc-pVQZ/B3LYP	−5.773	0.746
first CH_4·(NH_3) cc-pVDZ/HF	0.403	0.654	H_2O·(NH_3) cc-pVQZ/MP2	−6.047	0.815
first CH_4·(NH_3) cc-pVDZ/B3LYP	0.063	1.081	H_2O·(NH_3) cc-pV5Z/HF	−4.126	0.083
first CH_4·(NH_3) cc-pVDZ/MP2	−0.235	1.069	H_2O·(NH_3) cc-pV5Z/B3LYP	−5.950	0.194
first CH_4·(NH_3) cc-pVTZ/HF	0.510	0.293	H_2O·(NH_3) cc-pV5Z/MP2	−6.299	0.261
first CH_4·(NH_3) cc-pVTZ/B3LYP	0.274	0.497	H_2O·(H_2O) 6-31g/HF	−4.358	1.168
first CH_4·(NH_3) cc-pVTZ/MP2	−0.414	0.532	H_2O·(H_2O) 6-31g/B3LYP	−5.449	2.114
first CH_4·(NH_3) cc-pVQZ/HF	0.531	0.130	H_2O·(H_2O) 6-31g/MP2	−4.928	2.041
first CH_4·(NH_3) cc-pVQZ/B3LYP	0.241	0.252	H_2O·(H_2O) cc-pVDZ/HF	−3.447	2.162
first CH_4·(NH_3) cc-pVQZ/MP2	−0.521	0.241	H_2O·(H_2O) cc-pVDZ/B3LYP	−4.528	3.660
first CH_4·(NH_3) cc-pV5Z/HF	0.542	0.025	H_2O·(H_2O) cc-pVDZ/MP2	−3.851	3.566
first CH_4·(NH_3) cc-pV5Z/B3LYP	0.290	0.074	H_2O·(H_2O) cc-pVTZ/HF	−3.316	0.863
first CH_4·(NH_3) cc-pV5Z/MP2	−0.538	0.071	H_2O·(H_2O) cc-pVTZ/B3LYP	−4.491	1.581
second CH_4·(NH_3) 6-31g/HF	−0.089	0.501	H_2O·(H_2O) cc-pVTZ/MP2	−4.365	1.719
second CH_4·(NH_3) 6-31g/B3LYP	−0.550	0.878	H_2O·(H_2O) cc-pVQZ/HF	−3.332	0.363
second CH_4·(NH_3) 6-31g/MP2	−0.450	0.949	H_2O·(H_2O) cc-pVQZ/B3LYP	−4.354	0.871
second CH_4·(NH_3) cc-pVDZ/HF	0.054	1.197	H_2O·(H_2O) cc-pVQZ/MP2	−4.651	0.837
second CH_4·(NH_3) cc-pVDZ/B3LYP	−0.435	1.889	H_2O·(H_2O) cc-pV5Z/HF	−3.352	0.097
second CH_4·(NH_3) cc-pVDZ/MP2	−0.379	1.820	H_2O·(H_2O) cc-pV5Z/B3LYP	−4.417	0.286
second CH_4·(NH_3) cc-pVTZ/HF	0.154	0.467	H_2O·(H_2O) cc-pV5Z/MP2	−4.798	0.303
second CH_4·(NH_3) cc-pVTZ/B3LYP	−0.284	0.759	$(FHF)^-$ 6-31g/HF	−48.40	8.50
second CH_4·(NH_3) cc-pVTZ/MP2	−0.574	0.749	$(FHF)^-$ 6-31g/B3LYP	−53.45	14.47
second CH_4·(NH_3) cc-pVQZ/HF	0.223	0.199	$(FHF)^-$ 6-31g/MP2	−50.47	13.25
second CH_4·(NH_3) cc-pVQZ/B3LYP	−0.206	0.352	$(FHF)^-$ cc-pVDZ/HF	−47.18	8.50
second CH_4·(NH_3) cc-pVQZ/MP2	−0.608	0.352	$(FHF)^-$ cc-pVDZ/B3LYP	−52.48	14.56
second CH_4·(NH_3) cc-pV5Z/HF	0.243	0.035	$(FHF)^-$ cc-pVDZ/MP2	−48.17	14.86
second CH_4·(NH_3) cc-pV5Z/B3LYP	−0.164	0.086	$(FHF)^-$ cc-pVTZ/HF	−45.34	3.42
second CH_4·(NH_3) cc-pV5Z/MP2	−0.631	0.090	$(FHF)^-$ cc-pVTZ/B3LYP	−50.86	6.15
NH_3·(NH_3) 6-31g/HF	−1.793	1.119	$(FHF)^-$ cc-pVTZ/MP2	−48.50	7.59
NH_3·(NH_3) 6-31g/B3LYP	−2.759	1.768	$(FHF)^-$ cc-pVQZ/HF	−43.53	1.69
NH_3·(NH_3) 6-31g/MP2	−2.517	1.980	$(FHF)^-$ cc-pVQZ/B3LYP	−48.99	3.50
NH_3·(NH_3) cc-pVDZ/HF	−1.064	1.961	$(FHF)^-$ cc-pVQZ/MP2	−47.26	4.28
NH_3·(NH_3) cc-pVDZ/B3LYP	−2.029	3.041	$(FHF)^-$ cc-pV5Z/HF	−41.90	0.57
NH_3·(NH_3) cc-pVDZ/MP2	−1.871	3.064	$(FHF)^-$ cc-pV5Z/B3LYP	−46.87	1.36
NH_3·(NH_3) cc-pVTZ/HF	−1.242	0.748	$(FHF)^-$ cc-pV5Z/MP2	−45.38	2.00
NH3·(NH_3) cc-pVTZ/B3LYP	−2.263	1.216			

TABLE 3: MPn and CCSD CH_4·(NH_3) Second Orientation Binding Energies and BSSE (in kcal/mol)

method	binding energy	BSSE	method	binding energy	BSSE
cc-pVDZ/HF	0.049	1.197	cc-pVTZ/MP4DQ	−0.486	0.646
cc-pVDZ/MP2	−0.356	1.820	cc-pVTZ/MP4SDQ	−0.466	0.690
cc-pVDZ/MP3	−0.327	1.672	cc-pVTZ/CCSD	−0.480	0.657
cc-pVDZ/MP4D	−0.308	1.695	cc-pVTZ/CCSD(T)	−0.594	0.701
cc-pVDZ/MP4DQ	−0.264	1.682	cc-pVQZ/HF	0.200	0.199
cc-pVDZ/MP4SDQ	−0.276	1.713	cc-pVQZ/MP2	−0.620	0.350
cc-pVDZ/CCSD	−0.254	1.712	cc-pVQZ/MP3	−0.627	0.283
cc-pVDZ/CCSD(T)	−0.317	1.794	cc-pVQZ/MP4D	−0.612	0.292
cc-pVTZ/HF	0.132	0.467	cc-pVQZ/MP4DQ	−0.539	0.283
cc-pVTZ/MP2	−0.585	0.749	cc-pVQZ/MP4SDQ	−0.536	0.290
cc-pVTZ/MP3	−0.573	0.646	cc-pVQZ/CCSD	−0.523	0.290
cc-pVTZ/MP4D	−0.556	0.660	cc-pVQZ/CCSD(T)	−0.660	0.314

correlation limits can be reached, the intermolecular attraction is dominated by dispersion and modern density functional approaches lack this term.[21−23,30−32] BSSE is observed to diminish as the basis set is enhanced, and it will disappear for an infinite basis set. The Møller−Plesset perturbation series is stable and the MP4SDQ binding energy is within 0.005 kcal/

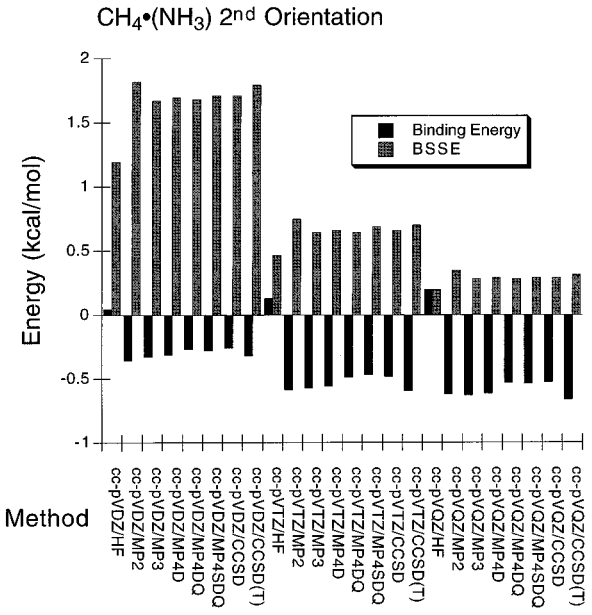

Figure 9. Binding energy and BSSE as a function of method and basis set for the $CH_4 \cdot (NH_3)$, second orientation (see Figures 1c and 4).

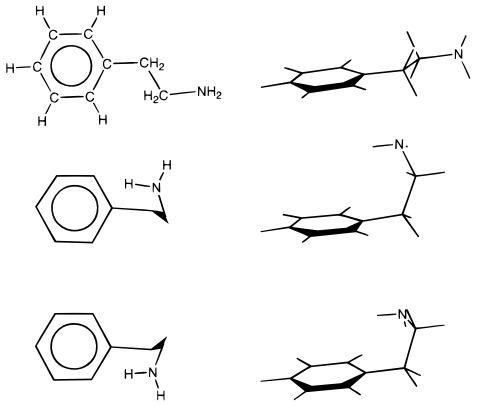

Figure 10. Phenethylamine conformations.

mol of the CCSD(T) binding energy.[21] The present cc-pV5Z/MP2 binding energy of 0.46 kcal/mol is nearly the same as experimental estimates[33−35] of 0.33−0.46 kcal/mol based on spherically averaged potentials.

The preferred orientation of methane•ethylene[28] places one of methane's hydrogens pointed toward the ethylene π-cloud. This orientation is only 0.11 kcal/mol more stable than other non-hydrogen-bonding orientations. The preference for this orientation is ascribed[28] to a 0.15 kcal/mol attractive electrostatic effect. The binding energy in methane•ethane is computed[28] to be 0.16 kcal/mol, stronger than in methane•ethylene. Here electrostatics are found[28] to be insignificant. Thus, the effect of the π-cloud on the intermolecular binding is not clear.

Aromatic−aromatic interactions are found in virtually all areas of chemistry.[1−10] This scope of application has generated a breadth of interest that has caused benzene dimer, despite its relatively large size, to be the most heavily studied hydrocarbon dimer.[25−27] Because its size prevents a definitive study and because experimental measurements are challenging, benzene dimer is somewhat controversial. The controversy surrounding benzene dimer is 2-fold. The first area of disagreement surrounds the binding energy. Experimental estimates place the binding energy at 1.6 kcal/mol.[36] Current theoretical estimates range from 2.3 kcal/mol,[25] through 3.2 kcal/mol,[26] to 4.3 kcal/mol.[27] In view of the difficulty in calculating the binding energy of

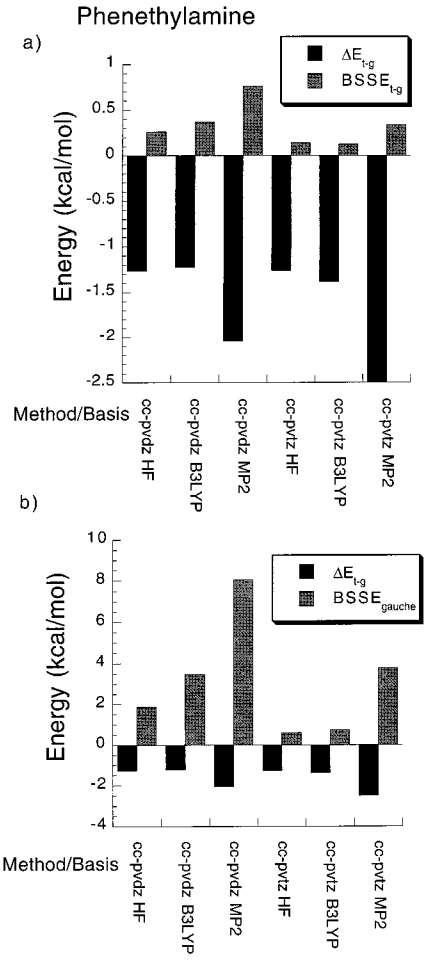

Figure 11. (a) Phenethylamine conformational energy differences between the lowest energy trans conformation and the lowest gauche conformation (ΔE_{t-g}), and the BSSE difference between these conformers ($BSSE_{t-g}$) as a formation of calculational method and basis set. (b) The same conformational energy difference and the gauche conformer BSSE (BSSE gauche), again plotted as a function of calculation method and basis set.

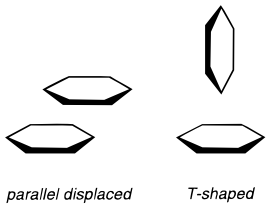

Figure 12. Two possible structures for the benzene dimer: parallel displaced and T-shaped.

methane dimer, it is not surprising that theory is not able to resolve this issue. The cc-pVTZ binding energy of methane dimer is only 69% of the cc-pV5Z binding energy. The magnitude of the cc-pVTZ BSSE is 22% of the final cc-pV5Z binding energy. Surprisingly, the computed binding energy of benzene dimer is larger than the experimental estimate.

The second area of contention in benzene dimer is the intermolecular structure. The most popular structures for benzene dimer are the T-shaped and the parallel-displaced orientations, as given in Figure 12. On the basis of the results for methane•ethylene discussed above and general literature consensus, electrostatics play a differential role between the T-shaped and parallel-displaced structures. Electrostatics are attractive in the T-shaped form and repulsive in the parallel-displaced orientation. Dispersive interactions paint precisely the opposite picture. The

parallel-displaced orientation has more close carbon−carbon contacts than the T-shaped structure. These close contacts lead to enhanced dispersion interactions. From the above discussion of methane dimer we know that the dispersion energy does not converge quickly with basis set. The energy difference between T-shaped and parallel-displaced benzene dimers should be strongly basis set dependent and the true energy difference only obtained with at least a cc-pV6Z basis. The conformational energy difference has been observed to be dependent upon the basis.[25−27] As expected, the parallel-displaced structure is differentially stabilized by an increase in basis set size. Unfortunately, the largest basis set used to date for benzene dimer is only of cc-pVTZ in quality. For methane dimer the MP2/cc-pVTZ BSSE is 31% of the cc-pVTZ binding energy. The final answer on benzene dimer awaits further work.

ii. Hydrogen-Bonding Systems. Hydrogen-bonded systems range from weakly hydrogen-bonded complexes such as CH_4·(NH_3), through classic hydrogen-bonded systems such as water dimer, to very strongly bonded LBHB systems such as H_2O·(H_3O^+). Each class of system has received theoretical study and each is discussed here.

iia. Weak H−Bonded Systems. Intermolecular interactions in weakly hydrogen-bonded systems are an admixture of van der Waals and electrostatic effects. The contribution of each can be conformationally dependent and in fact basis set dependent.

For CH_4·(NH_3) we find two conformations. In the first conformation, Figure 1b, methane is oriented with three hydrogens pointed toward the ammonia; this maximizes van der Waals interactions. In the second conformation, Figure 1c, a single hydrogen is pointed toward the nitrogen of ammonia. As would be expected from the discussion of hydrocarbon van der Waals complexes, the energy difference between the two CH_4·(NH_3) conformations is strongly method and basis set dependent, ranging from 0.69 kcal/mol B3LYP/6-31g** to 0.09 kcal/mol for MP2/cc-pV5Z. The present cc-pV5Z/MP2 binding energy of 0.63 kcal/mol is only marginally larger than our computed binding energy of 0.46 kcal/mol for methane dimer. The magnitude of this binding energy and the basis set dependence is consistent with a van der Waals dominated interaction.

There have been a few studies of CH_4·(H_2O). Scheiner and co-workers[37] report BSSE-corrected MP2/aug-cc-pVDZ and MP2/6-311+g**binding energies of 0.43 and 0.35 kcal/mol. They also report a B3LYP/6-311+g** binding energy of 0.26 kcal/mol. Novoa, Planas, and Rovira report[38] a BSSE-corrected MP2/aug-cc-pVQZ(-g,-f) binding energy of 0.42 kcal/mol. In contrast to these reports of a C−H···O interaction, Szczesniak et al.[39] find the oxygen to be pointed toward a C_3 face of methane. They report a 0.83 kcal/mol binding energy without accounting for BSSE. The magnitude of this binding energy and this preferred orientation are consistent with a van der Waals dominated interaction. Other C−H···O interactions have been studied;[37,40,41] a general trend is that as electronegative substituents are added to the hydrocarbon the binding energy increases through an increase in electrostatics.

A number of studies on C_6H_6·(H_2O) have been reported; the most extensive study is by Feller.[42] A portion of this data is presented in Figure 13. The binding energy and BSSE curves are analogous to those reported here for CH_4·(NH_3) as a function of the MP2/basis set level of theory.

C_6F_6·(H_2O) has been studied using a 6-311g** basis. A BSSE-corrected binding energy of 2.05 kcal/mol binding energy is reported[43] with a BSSE of 1.32 kcal/mol.

The CH_4·(HF) complex completes a methane hydrogen-bonding series. We found CH_4·(NH_3) to have two nearly

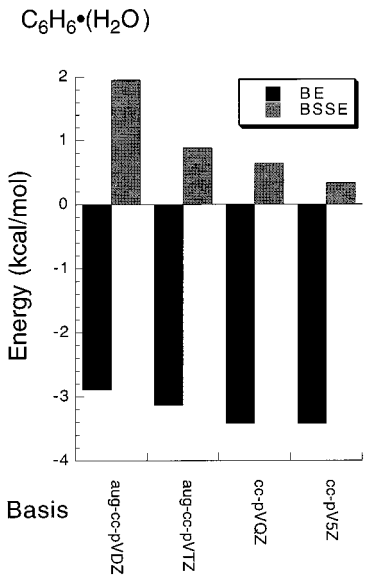

Figure 13. Binding energy and BSSE for C_6H_6·(H_2O) as a function of basis set for the MP2 level of theory.

isoenergetic structures, the hydrogen-bonding form being slightly favored over a more van der Waals orientation. The CH_4·(H_2O) story is still somewhat clouded; both van der Waals and hydrogen-bonding forms have been suggested. For CH_4·(HF) Chandra and Nguyen[44] study hydrogen-bonding and van der Waals orientations and find the van der Waals form to be preferred by 1.0 kcal/mol at the BSSE-corrected MP2/6-31++g-(3d3f, 2pd) level. The computed binding energy is 1.31 kcal/mol. For this basis the BSSE is still 0.8 kcal/mol.

Conformational energy differences for large biologically relevant molecules, such as phenethylamine, have been previously calculated.[45] MP2/6-311g** calculations give a trans−gauche conformational energy difference of −1.1 kcal/mol, whereas the MP2/cc-pVTZ BSSE-corrected value is −2.5 kcal/mol. Both computations favor the gauche conformation. The MP2/cc-pVTZ gauche conformer BSSE is about 3.5 kcal/mol, which is larger than the calculated trans−gauche conformational energy difference. The MP2/cc-pVTZ BSSE for the trans conformer is slightly less than 3 kcal/mol. These BSSE approximate values are obtained from the fragment analysis discussed above (see Figure 11).

iib. Classic H−Bonded Systems. There have been a large number of theoretical studies on conventional hydrogen-bonded systems such as water dimer.[46−50] Other simple systems include ammonia dimer and hydrogen fluoride dimer. Mixed systems include H_2O·(HF), NH_3·(HF), and NH_3·(H_2O).

Water Dimer. Water plays a central role in condensed state chemistry, and as such water dimer is of singular importance. The definitive theoretical study of water dimer was reported in 1992 by Feller.[46] The data reported in Figure 7 are in accord with this study. More recent studies have focused on water trimer,[48] the viability of LMP2 for hydrogen-bonding studies,[49] and the impact of BSSE on the geometry of water dimer.[50]

The cyclic trimer of water shows[48] the same basis set trends as water dimer (see Figure 14). Augmenting the cc-pVXZ basis sets provides an initial decrease in BSSE but then proceeds to follow the same slow asymptotic approach to zero BSSE.

The local MP2 (LMP2) method[51] has been suggested to obviate the need for a BSSE correction. Werner and co-workers have recently demonstrated[49] that for water clusters the validity of this suggestion depends on the basis set. The water dimer LMP2 basis set convergence characteristics are summarized in

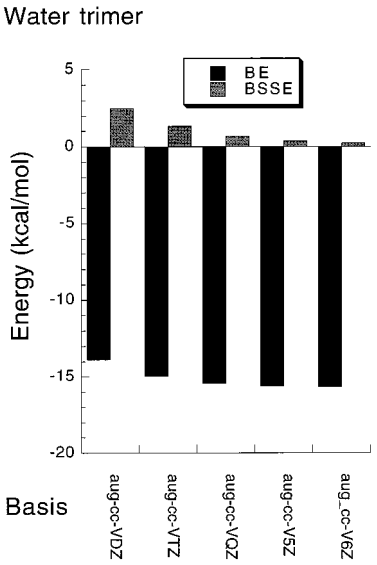

Figure 14. Binding energy and BSSE for the water trimer for the MP2/basis set level of theory.

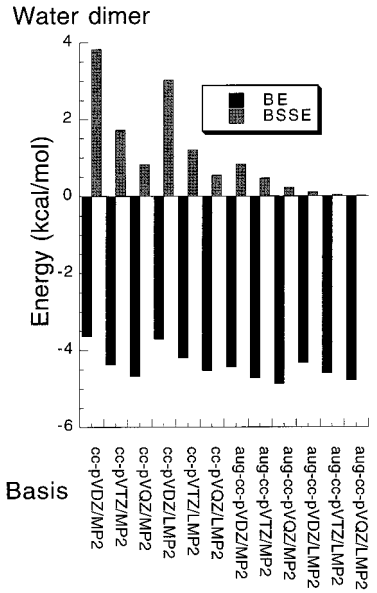

Figure 15. Binding energy and BSSE for the water dimer at the LMP2/basis set level of theory.

Figure 15. For the unaugmented cc-pVXZ basis sets, LMP2 drops the BSSE by ~20%. For augmented cc-pVXZ basis sets, changing from MP2 to LMP2 does cause an order of magnitude drop in BSSE. The BSSE is substantially diminished but not eliminated for augmented basis sets and the LMP2 algorithm. (Compare Figures 7 and 15.)

Since the BSSE is a significant fraction of the binding energy for water dimer for all but the largest of basis sets, the intermolecular geometry should be impacted by BSSE. In 1999 Simon, Duran, and Dannenberg[50] carried out a geometry optimization of water dimer with an explicit inclusion of a BSSE correction during the geometry optimization. The O−O separation increases when BSSE is removed by ~0.1 Å for a D95++(d,p) basis. Correcting for BSSE by using LMP2 also causes an O−O separation increase: by 0.04 Å for the cc-pVDZ basis set, and by 0.02 Å for the aug-cc-pVQZ basis set.

There have been scattered reports of other OH···O bonded systems. In conjunction with a microwave structure determination of ethylene oxide−water, Caminati and co-workers[52] computed the dimer binding energy at the MP2, QCISD(T), B3LYP, and BLYP theory levels with a 6-311++g(d,p) basis set. Comparable binding energies are obtained but BSSE is not corrected for. The most extensive study to date for other hydrogen-bonding small molecules is found in a recent report by Tsuzuki and co-workers.[53] The computed trends in MP2 BSSE and binding energy for the cc-pVXZ (X = D, T, Q, and 5) series for $H_2O\cdot(CH_3OH)$, $H_2O\cdot(CH_3OCH_3)$, $H_2O\cdot(H_2CO)$, methanol dimer, and formic acid dimer are as presented in Figures 6 and 7 above for classic hydrogen-bonded systems. The BSSE is larger for MP2 than for HF and the BSSE drops systematically with basis set improvement for each method.

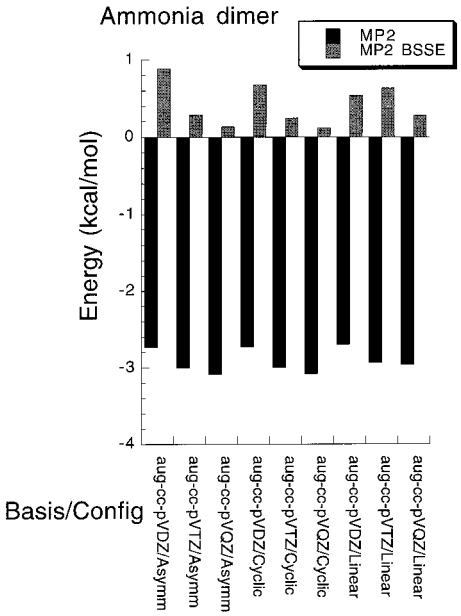

Figure 16. Binding energy and BSSE for ammonia dimer with the MP2/basis set level of theory. Note that the structure of the dimer is sensitive to the BSSE and the basis set.

Ammonia Dimer. The potential surface of ammonia dimer is configurationally complex. In addition to possessing a rather flat intermolecular surface, inversion at nitrogen plays a role in the spectroscopy that is used to determine the experimental structure.[54] The classic microwave study by Nelson, Fraser, and Klemperer,[55] wherein ammonia dimer is found to have an asymmetric structure, is inconsistent with ab initio reports from the 1980s. As shown in Figure 1, we find a symmetric-ring structure to be the global minimum. This geometry optimization is carried out with an analytic Hessian, and thus it establishes a true minimum: the lowest computed vibrational mode lies at 67 cm^{-1}. Subsequent to the completion of this work, Shin and Parks reported[56] a MP2/cc-pVTZ geometry optimization of ammonia dimer, finding essentially the same result. On the basis of the previous observations of Hassett, Marsden, and Smith[57] that diffuse basis functions are important for ammonia dimer, Lee and Parks determine the geometry with an aug-cc-pVTZ basis; they find that, with the augmented basis set, the symmetric-ring structure is no longer a minimum but is a saddle point that connects asymmetric structures. The energy difference between the asymmetric and symmetric structures is basis set dependent, see Figure 16. The energy difference between the cyclic symmetric and asymmetric structures drops from 0.22 kcal/mol for the aug-cc-pVDZ basis (without BSSE correction) to 0.007 kcal/mol for an aug-cc-pVQZ basis. For the aug-cc-pVQZ basis set, the BSSE is still >0.1 kcal/mol suggesting the final answer on NH_3 dimer awaits further work.

In a 1996 summary, Sosa, Carpenter, and Novoa[58,59] report MP2 and density functional studies on the hetero and homo

TABLE 4: MP2/cc-pV5Z and Experimental Binding Energies (in kcal/mol)

method	MP2/ cc-pV5Z	harmonic zero point	net BE	experiment
methane dimer	0.459	0.384	0.075	0.33−0.46[a]
first CH_4·(NH_3)	0.538			
second CH_4·(NH_3)	0.631	0.688	−0.057	
ammonia dimer	3.048	1.530	1.518	<2.80[b]
H_2O·(NH_3)	6.299	2.167	4.132	
water dimer	4.798	2.098	2.700	3.59 (373K)[c]
$(FHF)^-$	45.38			

[a] References 33−35. [b] Fraser, G. T.; Nelson, Jr. D. D.; Gerfen, G. J.; Klemperer, W. *J. Chem. Phys.* **1985**, *82*, 2535. [c] Curtiss, L. A.; Frurip, D. J.; Blander, M. *J. Chem. Phys.* **1977**, *71*, 2703.

dimers of NH_3, H_2O, and HF. The data in Figures 6 and 7 are consistent with their findings; correcting for BSSE is as important for DFT approaches as it is for MP2 and, in the limit of a large basis sets the results for hydrogen bonding are comparable to MP2 results.

iic. LBHB Systems. A number of hydrogen-bonding systems possess unusually large interaction energies. Most are observed/computed to have symmetric structures, while others are observed/computed to possess double-welled potentials with low barriers connecting the two minima. The H-bond donor and acceptor distances to the bridging hydrogen are computed to be equivalent or nearly equivalent. Because of its unusual strength, the hydrogen bond is said to be charge assisted or resonance assisted. Such systems cannot be described in terms of a single valence bond representation or a simple hydrogen bond potential. Two valence bond structures that are connected through resonance are needed to achieve a symmetric structure, as indicated in eq 1:

$$D{-}H\ A \leftrightarrow D^-\ H{-}A^+ \qquad (1)$$

This resonance stabilization model of the hydrogen bond is equivalent to a model in which the hydrogen bond possesses considerable covalency.

Simple examples of LBHB systems include NH_3·(NH_4+), H_2O·(H_3O^+), HF·(H_2F^+), and HF·(F^-). Of these systems, all are computed[60−63] to possess symmetric hydrogen bonds except NH_3·(NH_4^+),[60−62] which is computed to have a double well potential with a barrier of ∼1 kcal/mol between two equivalent asymmetric structures. By today's standards, the basis sets used for NH_3·(NH_4^+) are rather modest and BSSE is not accounted for in assessing the barrier height.

For H_2O·(H_3O^+),[63] the preference for a symmetric or asymmetric structure is found to be dependent upon basis set and level of electron correlation. Electron correlation at the CCSD-(T) level and a basis of cc-pVTZ quality is found to yield a symmetric structure, although BSSE is not estimated in this study.

C. Comparison with Experiment. Direct comparison of computed binding energies with experiment is not viable; zero-point effects as well as the temperature dependence of enthalpy cloud the comparison. Conventional harmonic analyses, useful for isolated molecular heat of formation comparisons, will not work here due to extreme anharmonicities of the intermolecular modes. Computed (MP2/cc-pV5Z) binding energies, computed harmonic zero-point energies, and net binding energies for the systems studied in the present work are collected in Table 4 along with experimental data. The inadequacy of a harmonic analysis is particularly evident for the van der Waals dominated systems.

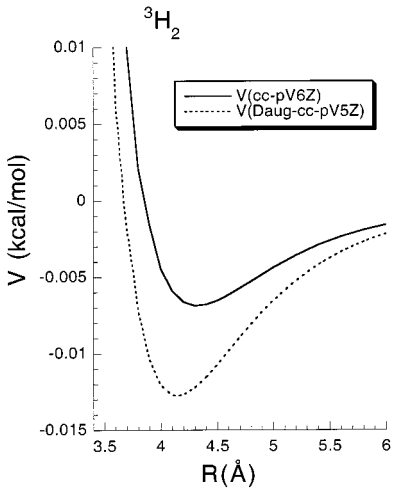

Figure 17. Triplet H_2 potential curves. The solid line is for the cc-pV6Z basis, the dashed line is obtained with the present Daug-cc-pV5Z basis set given in Table 5. The Daug-cc-pV5Z curve is very close to the accepted one (see ref 67). The calculated binding energy with the Daug-cc-pV5Z basis set is 0.0128 kcal/mol.

IV. Future Directions

The clear challenge ahead for the ab initio calculation of intermolecular interactions is the basis set. Møller−Plesset perturbation theory, even at low level, can account for the full range of intermolecular interactions. Density functional theory, with a large basis set, can properly describe induction and other electrostatic effects. Dispersive interactions are the most difficult to determine. They can be described with low-order perturbation theory but require a very large basis set for asymptotic convergence. Until asymptotic convergence is achieved, BSSE is a significant concern with likely structural consequences.

Dispersion Functions. As discussed previously,[38] at least two distinct spatial regions need to be described by a basis set. Functions are needed to describe the valence space and the electron correlation events between pairs of electrons that are geometrically close. Functions are also needed to describe the electron correlation events, predominantly dispersive in character, of the near-valence space. A number of reports have suggested that diffuse basis functions are needed to account for dispersion.[38] Unfortunately, these studies were not carried out with saturated valence basis sets. The diffuse basis was added at the expense of the valence space leading to a large BSSE and to the possibility of apparent binding due to an increase in energy of the monomers rather than a decrease in energy of the dimer.

In the cc-pVXZ series of basis sets the exponents are determined to minimize a total energy. The valence space correlation functions are optimized for the ground state of the atom. Diffuse functions are optimized for the negative ion of the atom. This energy-minimization prescription minimizes the possibility of unintentionally introducing an intermolecular bias. The question remains, however, does this prescription lead to a good description of dispersive interactions?

To test the "off the shelf" viability of the cc-pVXZ basis set family, the cc-pV6Z full-CI potential curve for $^3H_2$ is presented in Figure 17. The computed binding energy is only 57% of the near-exact value of 0.013 kcal/mol and the equilibrium distance is 0.2 Å long! The cc-pV6Z basis simply does not account for dispersion.

This deficiency of the cc-pVXZ basis sets was previously observed by Woon for rare gas dimers.[64] In an attempt to improve the description of dispersion, Woon and Dunning[65]

TABLE 5: cc-pVXZ Hydrogen Polarization Function Exponents, Optimized for $^3H_2$[a]

basis	angular momentum	1st	2nd	3rd	4th
Daug-cc-pVDZ	P	0.158			
[−1.000006736]					
(0.0160)					
Daug-cc-pVTZ	P	0.520	0.127		
[−1.000016201]	D	0.140			
(0.0220)					
Daug-cc-pVQZ	P	0.832	0.3025	0.11	
[−1.0000184984]	D	0.420	0.122		
(0.0234)	F	0.148			
Daug-cc-pV5Z	P	0.869	0.4356	0.219	0.11
[−1.0000188754]	D	0.8624	0.308	0.12	
(0.0237)	F	0.364	0.1425		
	G	0.17			

[a] The total energy is given in brackets in hartrees and the correlation energy is given in parentheses in kcal/mol

developed a layering structure to the augmented cc-pVXZ basis sets. Instead of adding a single set of diffuse functions, two or three sets of diffuse functions were added, exponent ratios optimized for negative ion, and noble gas atoms obtained by extrapolation; if the MP2 method is used, the triply augmented-cc-pV5Z basis recovers only 62% of the neon dimer binding energy; if the CCSD(T) method is used the triply augmented-cc-pV5Z basis recovers 93% of the binding energy.

In order to assess the range of exponents appropriate for computing dispersive forces we have computed the potential curve for $^3H_2$, the simplest van der Waals system. Here BSSE need not be a concern; the 10s Huzinaga set[66] has a total energy of 0.499 9993 hartree, suggesting a 0.000 88 kcal/mol maximal BSSE. To determine the proper range of basis function exponents for dispersion, we have hand-optimized sets of p, d, f, and g functions in the spirit of the cc-pVXZ basis sets. The exponents are optimized to minimize the total energy for $^3H_2$ at a bond distance of 4.15 Å. The new exponents and total energies are collected in Table 5 and the potential curve computed with this D-augmented basis set (dashed curve) is displayed in Figure 17. Using the Daug-cc-pV5Z functions the calculated binding energy for $^3H_2$ is 0.0128 kcal/mol and the bond distance is ∼4.15 Å. The accepted binding energy for $^3H_2$ is 0.0124 kcal/mol.[67] The D-augmented exponents for hydrogen are more diffuse that the exponents of the cc-VXZ sets, reproduced in Table 6, but are not as diffuse as Rydberg functions. For example, the Huzinaga[66] 2p Rydberg function for hydrogen has an exponent of 0.045. Sets of augmenting functions for hydrogen, created using the Woon−Dunning augmentation prescription, are collected in Table 7. Comparison of the data in Tables 6 and 7 confirms the more diffuse nature of the Woon−Dunning augmentation functions. The augmentation sets are not in the same range as the D-augmented exponents of Table 6, however. Interestingly, the most diffuse p, d, f, and g D-augmented exponents all hover in the range of 0.10−0.15. These results suggest that dispersion uses a different class of exponent than described by the cc-pVXZ and aug-cc-pVXZ basis sets.

To test the impact of a D-augmented set of exponents, the methane dimer van der Waals complex is computed with a "dispersion" augmentation. The cc-pVQZ basis hydrogen is augmented as follows: (1) the most diffuse D-augmented p exponent is added, the inner p exponents are adequately represented in the cc-pVQZ basis set; (2) both D-augmented d exponents are added; and (3) the D-augmented f exponent is added. Not surprisingly, the total energies of methane monomer and dimer computed with this Daug-cc-pVQZ basis set are lower than those computed with the cc-pVQZ basis set but not as low as those obtained with the cc-pV5Z basis set. The Daug-cc-pVQZ binding energy of 0.473 kcal/mol is 12% larger than that computed with the cc-pVQZ basis set (0.416 kcal/mol) and even 3% larger than that computed with the cc-pV5Z basis set (0.459 kcal/mol). The Daug-cc-pVQZ methane dimer binding energy is still 3.9% less than the cc-pVXZ exponential extrapolated value of 0.492 kcal/mol (BE = −0.491 95 + 2.2267exp(−0.844 01X)). The Daug-cc-pVQZ BSSE of 0.079 kcal/mol is 54% larger than that computed with the cc-pVQZ basis set (0.036 kcal/mol).

TABLE 6: cc-pVXZ Hydrogen Polarization Function Exponents

basis	angular momentum	1st	2nd	3rd	4th	
cc-pVDZ[a]	P	0.727				
cc-pVTZ[a]	P	1.407	0.388			
	D	1.057				
cc-pVQZ[a]	P	2.292	0.838	0.292		
	D	2.062	0.662			
	F	1.397				
cc-pV5Z[b]	P	4.516	1.712	0.649	0.246	
	D	2.95	1.206	0.493		
	F	2.506	0.875			
	G	2.358				
cc-pV6Z[b]	P	8.649	3.43	1.30	0.539	0.214
	D	4.453	1.958	0.861	0.378	
	F	4.10	1.78	0.773		
	G	3.199	1.326			
	H	2.653				

[a] Reference 15. [b] Reference 15 and K. A. Peterson, to be published.

TABLE 7: Aug-cc-pVXZ Hydrogen Augmentation Function Exponents

basis	angular momentum	1st	2nd	3rd
cc-pVDZ	S	0.029 74	0.007 25	0.001 77
	P	0.727	0.141	0.0273
cc-pVTZ	S	0.025 26	0.006 21	0.001 53
	P	0.102	0.0268	0.007 05
	D	0.247	0.055 25	0.012 35
cc-pVQZ	S	0.023 63	0.006 21	0.001 63
	P	0.0848	0.024 63	0.007 15
	D	0.190	0.054 53	0.015 65
	F	0.360	0.092 77	0.023 91

Tessellated Gaussians. The Dunning and co-workers cc-pVXZ and aug-cc-pVXZ series of basis sets provide a systematic way to approach the asymptotic convergence of the valence space, but do so at great expense. The number and complexity of the basis functions increase significantly as shells of functions are added. The size scaling of the cc-pVXZ basis sets is outlined in Table 7. The n^5 dependence of MP2/cc-pVXZ is illustrated for benzene dimer in Table 8 as well. The real Gaussian functions of Table 8 are calculated from Cartesian Gaussian functions. For example, the 7 real f functions are constructed from 10 Cartesian components (x^3, y^3, z^3, x^2y, x^2z, xy^2, xz^2, y^2z, yz^2, and xyz). The increased number of Cartesian functions can be hidden from the MP2 calculation but the integrals still must be evaluated. Technical advances such as LMP2[51] and resolved identity MP2 (RIMP2)[68] can dramatically decrease the n^5 power law but converged basis set calculations are still prohibitively expensive for biologically interesting molecules due both to the number of functions and to the complexity of the functions.

In order to construct a computationally more attractive family of basis sets, we need to understand what physics the shells of

TABLE 8: Number of Basis Functions Scaling for the cc-pVXZ Basis Sets, Timing Assumes 8.64×10^{13} Floating Point Operations per day and 10 Floating Point Operations per MP2 Term

basis	no. of hydrogens	no. of first-row elements	MP2 terms benzene dimer
cc-pVDZ	$5 = 5$	$6 + 9 = 15$	$(6 \times 21)^5 = 3.2 \times 10^{10}$ (5.3 min)
cc-pVTZ	$5 + 9 = 14$	$6 + 9 + 16 = 31$	$(6 \times 51)^5 = 2.7 \times 10^{12}$ (0.3 days)
cc-pVQZ	$5 + 9 + 16 = 30$	$6 + 9 + 16 + 25 = 46$	$(6 \times 76)^5 = 2.0 \times 10^{13}$ (2.3 days)
cc-pV5Z	$5 + 9 + 16 + 25 = 45$	$6 + 9 + 16 + 25 + 36 = 82$	$(6 \times 127)^5 = 2.6 \times 10^{14}$ (30 days)
cc-pV6Z	$5 + 9 + 16 + 25 + 36 = 81$	$6 + 9 + 16 + 25 + 36 + 49 = 131$	$(6 \times 212)^5 = 3.3 \times 10^{15}$ (380 days)

TABLE 9: Tessellation Level Scaling for Tessellated Spherical Gaussians

basis	no. of hydrogen	no. of first-row elements
1	$2 + 6 = 8$	$6 + 6 = 12$
2	$2 + 6 + 8 = 16$	$6 + 6 + 8 = 20$
3	$2 + 6 + 8 + 24 = 40$	$6 + 6 + 8 + 24 = 44$
4	$2 + 6 + 8 + 24 + 96 = 136$	$6 + 6 + 8 + 24 + 96 = 140$
5	$2 + 6 + 8 + 24 + 96 + 384 = 520$	$6 + 6 + 8 + 24 + 96 + 384 = 524$

functions are adding. One hypothesis is that the shells of functions are adding higher quantum number atomic states to the perturbation series. Another hypothesis is that the shells are merely adding nodes to provide an increasingly refined description of the valence, near-valence, and dispersion electron correlation events. If the first hypothesis is correct, one can only hope for faster computers and better integral evaluation algorithms. If the second hypothesis is correct, computationally expedient alternative approaches to increasing the number of nodes can be developed.

The numerical character of the shells of the hydrogen basis sets of Tables 6−8 provide support for the second hypothesis and a hint to an alternative asymptotic convergence scheme. If the shells of functions are added to describe increasingly higher quantum number atomic states, then the functions should become significantly more diffuse as a function of increasing shell. As shells are added, the smallest exponents (most diffuse components) are getting smaller but not dramatically so. Even for the cc-pV6Z basis the most diffuse p function only has an exponent of 0.214. The Dunning augmentation strategy adds more diffuse functions, but again generally not as diffuse as higher angular momentum atomic states would suggest. For the higher angular momenta, the "dispersion" exponents are more diffuse than the Dunning's augmentation exponents but again not as diffuse as atomic states would suggest.

A vastly more computationally expedient method for increasing the number of nodes can be constructed from Whitten's lobe Gaussian functions[69] or Frost's floating spherical Gaussians.[70] These function sets were developed in the 1960s as inexpensive alternatives to Cartesian Gaussian functions. Rather than use functions with explicit angular character, higher angular momentum functions were constructed as off-center linear combinations of spherically symmetric s functions. The positions of the s functions were included as variational parameters in the floating spherical Gaussian SCF procedure. Increasingly more efficient Cartesian Gaussian algorithms and the large number of variational functions needed to obtain accurate representations of valence orbitals led to the near abandonment of off-center basis functions. The remaining use of off-center basis functions is as "bond functions"; sets of Cartesian Gaussian basis functions are added at bond midpoints as an inexpensive alternative to the use of a more extensive atom-centered basis. Geometry optimization and molecular dissociation are problematic with bond functions because location of the bond functions is indeterminate as a bond breaks and as a geometry optimization occurs.

s basis functions are still simpler to evaluate than higher angular momentum functions. If standard valence basis sets are adopted, the challenge is providing the increasingly complex nodal structure. This can be accomplished by placing the s basis functions on a tessellated surface leading to tessellated spherical Gaussians (TSG). The first level of tessellation is provided by placing s functions at the vertexes of an octahedron. The radial displacement and exponent of the set of functions can be variationally determined. This set of functions formally provides an "s" function, three "p" combinations, and two "d" functions. The second level of tessellation places s functions at the centers of the eight triangular faces of the octahedron. This set of eight functions, when combined with the previous set of six functions, provides higher angular momenta components. Higher levels of tessellation can be provided by recursively placing functions in the centers of the previous level's triangles. The functional growth with tessellation of this TSG approach is provided in Table 9. The viability of the method has been tested on the van der Waals surface of $^3H_2$. The preliminary results are collected in Table 10. A single shell of TGS's provides 77% of the possible correlation energy, while two shells of TSG's recover 96% of the correlation energy. For comparison, the energy optimized set of p functions in Table 5 recover 68% of the correlation energy, while two shells (2p1d) provide 93% of the correlation energy. The TGS sets recover more correlation energy per shell because the shells contain higher angular momentum components. For example, the octahedral set of six s functions can be combined to make 3 p functions, 2 d functions (the e_g set), and an s function. Further work is necessary to establish the viability of tessellated spherical Gaussians, but they do provide hope for an inexpensive method of including dispersion.

TABLE 10: cc-pVXZ Hydrogen Tessellated Spherical Gaussian Functions, Optimized for $^3H_2$[a]

basis	tessellation level	1st (ζ/R)	2nd (ζ/R)
TSG1	0	0.18/0.14	
[−1.0000101706]			
(0.0182)			
TSG2	0	0.63/0.25	0.14/0.25
[−1.0000172964]	1	0.16/0.7	
(0.0227)			

[a] The total energy is given in brackets in hartrees and the correlation energy is given in parentheses in kcal/mol. The tessellation distances, R, are in Å.

V. Conclusions

A number of generalizations can be made regarding the computation of intermolecular interactions.

1. The cc-pVXZ family of basis sets can be used to establish the basis set convergence of modern electronic structure methods.

2. The B3LYP density functional approach and the MP2 wave function approach are observed to have similar BSSE characteristics.

3. When BSSE is accounted for, the binding energies of hydrogen-bonded systems are remarkably insensitive to basis set.

4. When van der Waals interactions are important, an extremely large basis set is needed to remove the BSSE and obtain a converged binding energy.

5. Density functional approaches cannot be used to study systems whose intermolecular interactions are dominated by dispersion.

6. Finally, we note in passing that the Lennard-Jones−Coulomb potential energy surface calculations for all these dimers, if properly calibrated,[20] agree with these converged ab initio binding energies to within ±20% (±10% for each calculational algorithm). This is roughly ±0.1 to ±0.2 kcal/mol.

Supporting Information Available: Tables of total energies and Cartesian coordinates are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Abraham, R. J.; Eivazi, F.; Pearson, H.; Smith, K. M. *J. Chem. Soc., Chem. Commun.* **1976**, 698.

(2) Cozzi, F.; Cinquini, M.; Annuziata, R.; Siegel, J. S. *J. Am. Chem. Soc.* **1993**, *115*, 5330.

(3) Shetty, A. S.; Zhang, J.; Moore, J. S. *J. Am. Chem. Soc.* **1996**, *118*, 1019.

(4) Tanner, D.; Fitzgerald, J. A.; Phillips, B. R. *Adv. Mater.* **1989**, *28*, 649.

(5) Castonguay, L. A.; Rappé, A. K.; Casewit, C. J. *J. Am. Chem. Soc.* **1991**, *113*, 7177.

(6) Pietsch, M. A.; Rappé, A. K. *J. Am. Chem. Soc.* **1996**, *118*, 10908.

(7) Kolb, H. C.; Andersson, P. G.; Sharpless, K. B. *J. Am. Chem. Soc.* **1994**, *116*, 1278.

(8) Gilman, A. G.; Rall, T. W.; Mies, A. S.; Taylor, P. *The Pharmaceutical Basis of Therapeutics*, 8th ed.; McGraw-Hill, Inc.: New York, 1993.

(9) Ishida, T.; Doi, M.; Ueda, H.; Inoue, M.; Scheldrick, B. M. *J. Am. Chem. Soc.*, **1988**, *110*, 2286.

(10) Kamiichi, K.; Danshita, M.; Minamino, N.; Doi, M.; Ishidia, T.; Inoue, M. *FEBS Lett.* **1986**, *195*, 57.

(11) Rappé, A. K.; Casewit, C. J. *Molecular Mechanics Across Chemistry*; University Science Books: Sausalito, CA, 1997.

(12) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247.

(13) Müller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143.

(14) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257. Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

(15) Dunning, T. H. Jr. *J. Chem. Phys.* **1989**, *90*, 1007. Kendall, R. A.; Dunning, T. H. Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. Woon, D. E.; Dunning, T. H. Jr. *J. Chem. Phys.* **1993**, *98*, 1358. Dunning, T. H., Jr.; Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975.

(16) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(17) Møller, C.; Plesset, M. S. *Phys Rev.* **1934**, *46* 618. Head-Gordon, M.; Pople, J. A.; Frish, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503. Frish, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275. Frish, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281. Pople, J. A.; Seeger, R.; Krishnan, R. *Int. J. Quantum Chem. Symp.* **1977**, *11*, 149. Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91. Krishnan, R.; Frish, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244.

(18) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910. Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F.; III *J. Chem. Phys.* **1988**, *89*, 7382. Scuseria, G. E.; Schaefer, H. F.; III *J. Chem. Phys.* **1989**, *90*, 3700. Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

(19) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(20) Yao, J.; Im, H. S.; Foltin, M.; Bernstein, E. R. *J. Phys. Chem.*, submitted for publication.

(21) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.

(22) Nagy, J.; Weaver, D. F.; Smith, V. H., Jr. *Mol. Phys.* **1995**, *85*, 1179.

(23) Rowley, R. L.; Pakkanen, T. *J. Chem. Phys.* **1999**, *100*, 3368.

(24) Hobza, P.; Selzle, H. L.; Schlag, E. W. *Chem. Rev.* **1994**, *94*, 1767.

(25) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.

(26) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780.

(27) Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **1996**, *252*, 206.

(28) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Phys. Chem. A* **1999**, *103*, 8265.

(29) Metzger, T. G.; Ferguson, D. M.; Glauser, W. A. *J. Comput. Chem.* **1997**, *18*, 70.

(30) Meijer, E. J.; Sprik, M. *J. Chem. Phys.* **1996**, *105*, 8684.

(31) Kohn, W.; Meir, Y.; Makarov, D. E. *Phys. Rev. Lett.* **1998**, *80*, 4153.

(32) Andersson, Y.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **1996**, *76*, 102.

(33) Reid, B. P.; O'Loughlin, M. J.; Sparks, R. K. *J. Chem. Phys.* **1985**, *83*, 5656.

(34) Boughton, C. V.; Miller, R. E.; Watts, R. O. *Mol. Phys.* **1985**, *56*, 363.

(35) Hoinkis, J.; Ahlrichs, R.; Böhm, H. J. *Int. J. Quantum Chem.* **1983**, *23*, 821.

(36) Neusser, H. J.; Krause, H. *Chem. Rev.* **1994**, *84*, 1829.

(37) Gu, Y.; Kar, T.; Scheiner, S. *J. Am. Chem. Soc.* **1999**, *121*, 9411.

(38) Novoa, J. J.; Planas, M.; Rovira, M. C. *Chem. Phys. Lett.* **1996**, *251*, 33.

(39) Szczesniak, M. M.; Chalasinski, G.; Cybulski, S. M.; Cieplak, P. *J. Chem. Phys.* **1993**, *98*, 3078.

(40) Novoa, J. J.; Lafuente, P.; Mota, F. *Chem. Phys. Lett.* **1998**, *290*, 519.

(41) Caminati, W.; Melandri, S.; Rossi, I.; Favero, P. G. *J. Am. Chem. Soc.* **1999**, *121*, 10098.

(42) Feller, D. *J. Phys. Chem. A* **1999**, *103*, 7558.

(43) Danten, Y.; Tassaing, T.; Besnard, M. *J. Phys. Chem. A* **1999**, *103*, 3530.

(44) Chandra, A. K.; Nguyen, M. T. *J. Phys. Chem. A* **1998**, *102*, 6865.

(45) Dickinson, J. A.; Hockridge, M. R.; Kroemer, R. T.; Robertson, E. G.; Simons, J. P.; McCombie, J.; Waler, M. *J. Am. Chem. Soc.* **1998**, *120*, 2622.

(46) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.

(47) Feyereisen, M. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1992**, *100*, 2993.

(48) Nielsen, I. M. B.; Seidl, E. T.; Janssen, C. L. *J. Chem. Phys.* **1999**, *110*, 9435.

(49) Schütz, M.; Rauhut, G.; Werner, H.-J. *J. Phys. Chem. A* **1998**, *102*, 5997.

(50) Simon, S.; Duran, M.; Dannenberg, J. J. *J. Phys. Chem. A* **1999**, *103*, 1640.

(51) Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151. Saebø, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, *113*, 13. Azhary, A. E.; Rauhut, G.; Pulay, P.; Werner, H.-J. *J. Chem. Phys.* **1998**, *108*, 5185.

(52) Caminati, W.; Moreschini, P.; Rossi, I.; Favero, P. G. *J. Am. Chem. Soc.* **1998**, *120*, 11144.

(53) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *J. Chem. Phys.* **1999**, *110*, 11906.

(54) Karyakin, E. N.; Fraser, G. T.; Loeser, J. G.; Saykally, R. J. *J. Chem. Phys.* **1999**, *110*, 9555.

(55) Nelson, D. D., Jr.; Fraser, G. T.; Klemperer, W. *J. Chem. Phys.* **1985**, *83*, 6201.

(56) Lee, J. S.; Park, S. Y. *J. Chem. Phys.* **2000**, *112*, 230.

(57) Hassett, D. M.; Marsden, C. J.; Smith, B. J. *Chem. Phys. Lett.* **1991**, *183*, 449.

(58) Sosa, C. P.; Carpenter, J. E.; Novoa, J. J. *ACS Symp. Ser.* **1996**, *629*, 131.

(59) Novoa, J. J.; Sosa, C. *J. Phys. Chem.* **1995**, *99*, 15837.

(60) Del Bene, J. E.; Frisch, M. J.; Pople, J. A. *J. Phys. Chem.* **1985**, *89*, 3669.

(61) Platts, J. A.; Laidig, K. E. *J. Phys. Chem.* **1995**, *99*, 6487.

(62) Jaroszewski, L.; Lesyng, B.; McCammon, J. A. *J. Mol. Struct.* **1993**, *283*, 57.

(63) Xie, Y.; Remington, R. B.; Schaefer, H. F.; III *J. Chem. Phys.* **1994**, *101*, 4878.

(64) Woon, D. E. *J. Chem. Phys.* **1994**, *100*, 2838.

(65) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994**, *100*, 2975.

(66) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.

(67) Kolos, W.; Wolnewicz, L. *J. Chem. Phys.* **1965**, *43*, 2429.

(68) Almlöf, J. *Chem. Phys. Lett.* **1991**, *181*, 319. Häser, M.; Almlöf, J. *J. Chem. Phys.* **1992**, *96*, 489. Vahtras, P.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514. Bernholdt, D. E.; Harrison, R. J. *Chem. Phys. Lett.* **1996**, *250*, 477.

(69) Whitten, J. L. *J. Chem. Phys.* **1966**, *44*, 359.

(70) Frost, A. A. *J. Chem. Phys.* **1967**, *47*, 3707.