Ab Initio Study of Hydrogen-Bonded Complexes of Small Organic Molecules with Water

Paul R. Rablen,*,† Jeffrey W. Lockman,‡ and William L. Jorgensen§

Contribution from the Department of Chemistry, Swarthmore College, 500 College Avenue, Swarthmore, Pennsylvania 19081-1397, and the Department of Chemistry, Yale University, P.O. Box 208107, New Haven, Connecticut 06520-8107

Received: January 15, 1998; In Final Form: March 11, 1998

Hydrogen bonding between water and a series of small organic molecules was examined via electronic structure calculations. Several computational methods were examined, including both a hybrid density functional procedure (Becke3LYP) and second-order Møller−Plesset theory (MP2) coupled with a double-ζ basis set augmented by diffuse polarization functions on heteroatoms. The agreement between Becke3LYP and MP2 energies was generally good, as was the agreement with energies obtained using more sophisticated and costly methods. The energies and structures of 53 hydrogen-bonded complexes of water with various small organic molecules, including alcohols, thiols, ethers, thioethers, carboxylic acids, esters, amines, amides, nitriles, and nitro compounds, were then examined systematically using the Becke3LYP and MP2 procedures. The hydrogen bond geometries were generally linear, and acceptor sites corresponded closely to the positions of lone pairs as predicted by simple hybridization arguments. Structures with sulfur and chlorine atoms showed some deviation from these simple expectations and seemed to be largely determined by molecular dipole−dipole interactions. Categorization of the hydrogen bonds involved in the various complexes led to an ordering of hydrogen bond donor and acceptor abilities for some common functional groups. The strength of association was found to correlate moderately well with experimental gas-phase basicity in those cases where water acted unambiguously as the hydrogen bond donor at a single site. Interestingly, sulfur was found to be close to oxygen in hydrogen bond acceptor strength, and the surprisingly strong acceptor ability of sulfur could not be explained in terms of its enhanced polarizability relative to oxygen. Calculations were also carried out on the AT and GC base pairs and yielded results in very close agreement with the highest levels of calculation previously reported.

Introduction

Hydrogen bonding is of central importance in the molecular sciences for both practical and theoretical reasons.[1−8] It largely determines the physical properties of many common condensed-phase systems including water.[9,10] It represents the strongest force governing the influence of solvents on molecular structure and reactivity, and a quantitative accounting for hydrogen-bonding interactions is a prerequisite for the proper understanding of chemical activity taking place in aqueous solution. Hydrogen bonding has held particular interest in recent years due to the central role it plays with regard to molecular recognition in both biological and artificial systems.[11−14] From a practical point of view, quantitative knowledge of the energies and geometries of hydrogen-bonding interactions is particularly important for the development and validation of the empirical force fields used for conformational analysis and for the statistical mechanical simulation of solution environments.[15,16] Accordingly, bimolecular complexes bound by hydrogen bonds have been the subject of numerous computational[17−37] and experimental[38−42] studies in the past.

Despite the prior research in this area, a systematic study of the interactions of a single water molecule with a variety of small organic molecules having diverse functional groups at a consistent and reliable level of ab initio theory was desired. Such an investigation could establish relative hydrogen-bonding abilities and a database for force field validation. It is the principle aim of this study to provide exactly this sort of database, for a series of compounds of general interest to the biological and organic modeling community.

The HF/6-31G* and HF/6-31G** levels of theory often give fortuitously accurate energies of both conformational change and hydrogen bonding. This happy coincidence permits the study of rather large systems where more sophisticated calculations are not feasible. In fact, substantially more expensive calculations are required before any real improvement beyond HF/6-31G* or HF/6-31G** is observed. Many important studies of hydrogen bonding have been carried out at these economical and frequently sufficient levels of theory.[18,33−37]

However, it is well-known that the reliably accurate description of weak interactions generally requires a treatment of electron correlation. Density functional methods have recently proved quite useful in this regard for hydrogen-bonded complexes.[43−52] The B3YLP functional in particular has proven highly effective, at least as long as basis sets at least as large as 6-31+G** are used.[45,49,50] Density functional theory offers an electron correlation correction frequently comparable to second-order Møller−Plesset theory (MP2), or in certain cases and for certain purposes even superior to MP2,[53] but at considerably lower computational cost. The cost advantage becomes progressively greater as the size of a system increases. We have chosen here to use density functional theory as a means to study hydrogen-bonded complexes with a higher level of reliability

† Swarthmore College; prablen1@swarthmore.edu.
‡ Swarthmore College; jlockma1@swarthmore.edu.
§ Yale University; bill@adrik.chem.yale.edu.

S1089-5639(98)00708-7 CCC: $15.00
Published on Web 05/02/1998

and accuracy than would otherwise be possible. The cost advantage is particularly important for the time-consuming process of geometry optimization. Through the inclusion of a correlation correction and the use of an extended basis set, we believe that the calculations presented here should serve as a highly dependable guide to the relative strengths and preferred geometries of hydrogen bonds.

The computational procedures for which we report association energies have been chosen specifically to reproduce the hydrogen bond strengths derived from experiment and from the highest levels of calculation possible for very small systems. It is important to remember, however, that it is not well understood why density functional methods work as well as they do. Although we believe our calculations are reliable for the hydrogen-bonded complexes examined here, in which electrostatic interactions play a dominant role, there is reason to believe that density functional theory is generally not appropriate for complexes in which dispersion interactions are predominant.[54] In such cases, there is currently no adequate alternative to high-cost traditional correlation treatments. Wherever feasible, we have included in this study energies calculated at MP2 and at other, higher levels of theory for comparison.

Calculations

All ab initio calculations were carried out using the Gaussian 94[55] suite of programs. Density functional calculations employed the Becke3LYP keyword, which invokes Becke's three-parameter hybrid method[56] using the correlation functional of Lee, Yang, and Parr.[57,58] Density = current was specified for all calculations (HF, DFT, and MP2) for which dipole moments are reported in order to ensure use of the correlated charge density distribution. The CBS-4 and CBS-Q composite procedures,[59] which incorporate a complete basis set extrapolation,[60] are included in Gaussian 94 and were used without modification.

Standard Pople basis sets[61] were used for all calculations, although in some cases with customized augmentation. The nonstandard augmentations were adapted from the work of Ochterski.[62] The (d+) and (2d+) descriptors indicate an additional set of diffuse d polarization functions having an exponent one-fourth as large as that for the previous set. The basis sets 6-31+G(d(X+),p) and 6-31++G(2d(X+),p) indicate that the additional diffuse polarization functions are included only on atoms having lone pairs.

Geometry optimizations were carried out without constraints. All stationary points were confirmed as minima via vibrational frequency calculations. In some cases, symmetrical structures that were investigated did yield imaginary frequencies, and in these instances, the symmetry was lowered and optimization carried out anew. Both the geometry optimizations and the vibrational frequency calculations were carried out at B3LYP/6-31+G(d(X+),p), a level found to be sufficient for these purposes, as described in the Results and Discussion section.

The basis set superposition error (BSSE) for the association energy was estimated using the counterpoise method. For each hydrogen-bonded complex, the energy of each monomer was computed with and without the basis functions for the other monomer present. The former was accomplished by using the "massage" keyword. The monomer geometries corresponding to the optimized complexes were used in all cases. The sum of the energy differences from including the basis functions then provided the counterpoise correction for the BSSE. For instance, for formaldehyde−water, the optimized geometry of the complex was used for a series of four calculations: (1) water with all the basis functions, (2) water with only its own basis functions, (3) formaldehyde with all the basis functions, and (4) formaldehyde with only its own basis functions. The sum of the energy differences (1) minus (2) and (3) minus (4) then yielded the counterpoise estimate of the BSSE.

Difference density distributions were computed from the appropriate wave function files obtained from Gaussian using the CASGEN package written at Yale University.[63]

TABLE 1: Comparison of Dipole Moments and Dimerization Energies for Water Calculated at the Hartree−Fock Level Using Different Standard Basis Sets

basis set	μ (H_2O)	$-\Delta E$
HF/6-31G*	2.199	5.62
HF/6-31+G*	2.285	5.38
HF/6-31G**	2.148	5.53
HF/6-31+G**	2.234	5.04
HF/6-311++G(3df,2p)	1.972	3.93
HF/6-311++G(3df,2p)//HF/6-31+G**	1.964	3.90
HF/6-311++G(3df,3pd)	1.967	4.00
HF/6-311++G(3df,3pd)//HF/6-31+G**	1.956	3.98

TABLE 2: Comparison of Dipole Moments and Dimerization Energies for Water Calculated at the MP2 Level Using Different Standard Basis Sets

level of theory	μ (H_2O)	$-\Delta E$
MP2/6-31G*	2.199	7.32
MP2/6-31+G*	2.327	7.00
MP2/6-31G**	2.112	7.04
MP2/6-31+G**	2.237	6.38
MP2/6-311++G(3df,2p)	1.938	5.32
MP2/6-311++G(3df,2p)//MP2/6-31G*	1.952	5.18
MP2/6-311++G(3df,2p)//MP2/6-31+G*	1.931	5.32
MP2/6-311++G(3df,2p)//MP2/6-31G**	1.950	5.18
MP2/6-311++G(3df,2p)//MP2/6-31+G**	1.927	5.29
MP2/6-311++G(3df,3pd)	1.899	5.30
MP2/6-311++G(3df,3pd)//MP2/6-31+G**	1.882	5.30

Results and Discussion

Selection of an Appropriate Level of Theory. Before performing calculations on larger systems, we chose to study the water dimer and a few other very small systems in great detail in order to find a procedure that would yield optimal energies and geometries for hydrogen-bonding interactions at reasonable cost. Two criteria were used to assess the quality of prospective methods: (1) accurate reproduction of the dipole moment of an isolated water molecule, and (2) accurate reproduction of the dimerization energy of water. The first criterion was adopted because hydrogen bonds are largely electrostatic in nature, and so an accurate dipole moment would seem a prerequisite for the proper treatment of hydrogen-bonding interactions. Requiring agreement with experiment and the highest level calculations for both the dimerization energy *and* the dipole moment also reduces the possibility of unwittingly selecting a method that achieves only fortuitous accuracy and that breaks down when applied to other systems.

It has been shown previously that computing the dipole moment of water accurately requires both a correction for electron correlation and the inclusion of diffuse polarization functions in the basis set.[67] The data in Tables 1−6 illustrate these points quite clearly. Table 1 shows that the Hartree−Fock (HF) level consistently overestimates the experimental dipole moment of water (1.854 D). With the popular 6-31G* basis set, the calculated dipole moment exceeds the experimental value by 18.6%. The tendency of HF/6-31G* to exaggerate the degree of charge separation is well-known and, in fact, is often invoked to justify the use of gas-phase calculations to

TABLE 3: Comparison of Dipole Moments and Dimerization Energies for Water Calculated Using Density Functional Theory (Becke3LYP Functional) with Different Standard Basis Sets

level of theory[a]	μ (H_2O)	$-\Delta E$
DFT/6-31G*	2.096	7.69
DFT/6-31+G*	2.249	6.44
DFT/6-31G**	2.044	7.54
DFT/6-31+G**	2.195	6.04
DFT/6-311++G(3df,2p)	1.914	4.83
DFT/6-311++G(3df,2p)//HF/6-31G*	1.902	4.76
DFT/6-311++G(3df,2p)//HF/6-31+G*	1.887	4.77
DFT/6-311++G(3df,2p)//HF/6-31G**	1.893	4.73
DFT/6-311++G(3df,2p)//HF/6-31+G**	1.875	4.69
DFT/6-311++G(3df,2p)//DFT/6-31G*	1.942	4.51
DFT/6-311++G(3df,2p)//DFT/6-31+G*	1.914	4.81
DFT/6-311++G(3df,2p)//DFT/6-31G**	1.939	4.57
DFT/6-311++G(3df,2p)//DFT/6-31+G**	1.908	4.81
DFT/6-311++G(3df,2p)//CD1993[b]	1.922	4.66
DFT/6-311++G(3df,3pd)	1.891	4.83
DFT/6-311++G(3df,3pd)//DFT/6-31+G**	1.880	4.82

[a] DFT = Becke3LYP. [b] CD1993 = geometry from Chakravorty, S. J.; Davidson, E. R. *J. Phys. Chem.* **1993**, *97*, 6373−6383.

TABLE 4: Comparison of Dimerization Energies for Water Calculated Using Various Compound Levels of Theory

level of theory[a]	$-\Delta E$
CBS-4 (w/o ZPE)	4.82
CBS-Lq	4.84
CBS-Q (w/o ZPE)	5.13

[a] ZPE = zero-point vibrational energy.

TABLE 5: Comparison of Dipole Moments and Dimerization Energies for Water Calculated Using Density Functional Theory (Becke3LYP Functional) with Different Nonstandard Basis Sets

level of theory[a]	μ (H_2O)	$-\Delta E$
DFT/6-31+G(d+,p)	1.901	4.91
DFT/6-31+G(2d+,p)	1.860	4.79
DFT/6-31+G(3d,p)	1.885	4.81
DFT/6-31+G(2d+,p)//DFT/6-31+G(d+,p)	1.860	4.79
DFT/6-31+G(3d,p)//DFT/6-31+G(d+,p)	1.883	4.81
DFT/6-31++G(2d+,p)//DFT/6-31+G(d+,p)	1.856	4.76
DFT/6-31++G(3d,p)//DFT/6-31+G(d+,p)	1.872	4.78
DFT/6-311+G(2d+,p)//DFT/6-31+G(2d+,p)	1.866	4.81
DFT/6-311++G(2d+,p)//DFT/6-31+G(2d+,p)	1.864	4.78
DFT/6-311+G(3d,p)//DFT/6-31+G(3d,p)	1.955	5.01

[a] DFT = Becke3LYP.

TABLE 6: Comparison of Dipole Moments and Dimerization Energies for Water Calculated at the MP2 Level Using Different Nonstandard Basis Sets

level of theory[a]	μ (H_2O)	$-\Delta E$
MP2/6-31+G(d+,p)	1.913	5.31
MP2/6-31+G(2d+,p)	1.887	5.23
MP2/6-31+G(3d,p)	1.900	5.21
MP2/6-31+G(2d+,p)//DFT/6-31+G(d+,p)	1.879	5.22
MP2/6-31+G(3d,p)//DFT/6-31+G(d+,p)	1.891	5.20
MP2/6-31++G(2d+,p)//DFT/6-31+G(d+,p)	1.874	5.22
MP2/6-31++G(3d,p)//DFT/6-31+G(d+,p)	1.884	5.22
MP2/6-311+G(2d+,p)//MP2/6-31+G(2d+,p)	1.883	5.22
MP2/6-311++G(2d+,p)//MP2/6-31+G(2d+,p)	1.883	5.15
MP2/6-311+G(3d,p)//MP2/6-31+G(3d,p)	1.978	5.36

[a] DFT = Becke3LYP.

describe the charge distribution of a molecule in solution.[65] Increasing the basis set size reduces the calculated dipole moment substantially, but even with the largest basis set investigated, 6-311++G(3df,3pd), the calculation overestimates the experimental dipole moment by 5.5%.

The treatment of electron correlation through either density functional theory (Becke3LYP) or MP2 brings the dipole moment into accord with experiment. Tables 2 and 3 demonstrate that although DFT and MP2 yield somewhat different dipole moments when used with the smaller basis sets, they approach essentially the same limiting value of 1.89 ± 0.01 when paired with the larger basis sets. This value overestimates the experimental dipole moment by only 1.9%. Furthermore, using the largest basis sets for the geometry optimization is not necessary for achieving the limiting value for the dipole moment. For instance, the later lines in Tables 2 and 3 show that while the 6-31G*-based geometry is not entirely adequate, geometry optimization beyond the 6-31+G** basis yields little improvement.

We next investigated the water dimer to see what was necessary to obtain agreement for the interaction energy within ±0.5 kcal/mol with respect to the best estimates available. The experimental binding enthalpy of the water dimer has been reported as 3.6 ± 0.5 kcal/mol.[66] Del Bene has back-calculated an "experimental" electronic binding energy of 5.5 ± 0.5 kcal/mol using the experimental data and ab initio vibrational frequency calculations.[50] The best computational estimates available, which range from 4.9 to 5.3 kcal/mol, fall within the experimental error bars.[17−19,21,31,48] The reported experimental uncertainty of ±0.5 kcal/mol, which is further compounded by uncertainty in the back-calculation of an electronic binding energy from the experimental measurement, is probably greater than the theoretical uncertainty for this small system. Consequently, agreement with the best levels of theory available was deemed the most appropriate measure of suitability.

Tables 1−3 list the dimerization energies obtained at the HF, DFT, and MP2 levels of theory discussed above. These values are not corrected for BSSE (vide infra). Hartree−Fock theory tends to overestimate the binding energy when used with the smaller basis sets, in accord with its overestimate of the dipole moment. With the larger basis sets, it underestimates the binding energy, presumably due to omission of dispersion forces, which can only be described by correlated levels of theory. When correlation is included, the calculated dimerization energy converges to a limiting value of 4.8 kcal/mol for DFT and of 5.3 kcal/mol for MP2. These values are very close to the best computational estimates available, which range from 4.9 to 5.3 kcal/mol.[17−19,21,31,48] The C_s symmetric linear structure was invariably the lowest in energy at all levels of theory, in agreement with previous studies.[17−19,21,31,48,50] The difference in the limiting binding energies obtained with DFT and MP2 might represent a greater sensitivity of the MP2 calculations to basis set superposition error (BSSE), which was consistently much larger for the MP2 calculations than for the DFT calculations (vide infra). As was the case for the dipole moment, more modest basis sets sufficed for the geometry optimization even when the final calculation of binding energy required a large basis set.

Several composite theoretical procedures are available in the literature for the calculation of accurate relative energies, including Pople's G1 and G2[67] techniques and the CBS (complete basis set) series of methods developed by Ochterski and Petersson.[59,60] In Table 4 the results are tabulated for the dimerization energy of water using three of these procedures. The CBS-4 method is both highly reliable and remarkably economical for many purposes, particularly the computation of bond dissociation energies (BDEs).[68,69] It includes treatment of electron correlation at the MP2 and MP4(SDQ) levels as well as an extrapolation of the HF and MP2 components to the

TABLE 7: Electronic Binding Energies ($-\Delta E$) of Selected Complexes (kcal/mol)

compound	pt. grp.	HF/6-31G*	MP2/6-31G*	CBS-4 (w/o ZPE)	CBS-Lq (w/o ZPE)	CBS-Q (w/o ZPE)	MP2[a]	B3LYP[b]
HF−HF	C_s	26.81	8.01	3.97	3.97	4.54	4.76	4.61
H_2O-H_2O	C_s	5.62	7.32	4.81	4.84	5.13	5.22	4.76
HCN−HF	$C_{\infty v}$	6.53	7.54	7.56	7.34	7.56	7.89	7.71
H_2CO-H_2O	C_1	5.28	6.93	4.07	4.13	5.15	5.42	4.46
CH_3CN−HF	C_{3v}	7.89	8.95	9.54	9.16	9.26	9.48	9.49

[a] MP2 = MP2/6-31++G(2d+,p)//Becke3LYP/6-31+G(d+,p) [b] B3LYP = Becke3LYP/6-31++G(2d+,p)//Becke3LYP/6-31+G(d+,p).

complete basis set limit. In the current application, however, CBS-4 appears less than optimal, as illustrated by the low estimate it yields for the dimerization energy of water. The error quite possibly results from using the HF/3-21G* optimized geometry, which is sufficient for most isolated molecules but is not adequate to describe weakly bound complexes. The slightly more sophisticated CBS-Lq method, which still relies upon a HF/3-21G* geometry, offers no improvement. The more elaborate CBS-Q procedure, on the other hand, yields a more accurate result, albeit at substantially greater computational cost. CBS-Q is built around an MP2/6-31G† optimized geometry and includes electron correlation at the MP2, MP4, and QCISD(T) levels. It generally has an accuracy comparable to, or even somewhat superior to, that of G2. However, the calculations are quite demanding of computational resources and are only feasible for smaller systems. In the present application, we have been able to carry out CBS-Q calculations for the smaller hydrogen bonded complexes, and in these cases CBS-Q serves as a useful check on the accuracy of the DFT and MP2 calculations.

Tables 1−3 thus demonstrate that calculations using DFT or MP2 with the 6-311++G(3df,3pd) basis set should yield reliable hydrogen-bonding interaction energies. However, this basis set is too large to be practical for all but the smallest of molecules. Consequently, we sought to find the smallest basis set with which the limiting values for the dipole moment and dimerization energy of water could be achieved. Tables 5 and 6 summarize the results of this investigation, in which nonstandard diffuse polarization functions of the sort recommended by Ochterski were used.[62] The comparatively modest 6-31++G(2d+,p) basis set achieves the desired level of accuracy in dipole moment and dimerization energy. The even smaller 6-31+G(d+,p) basis set slightly increases the error in the calculated dipole moment, but is sufficient for geometry optimization. These basis sets are very closely related to those recommended by Ochterski[62] and by Del Bene[17] or the economical study of hydrogen-bonded complexes. We have made the additional simplification of including the supplemental diffuse polarization functions only on atoms having lone pairs (presumably the atoms most important in hydrogen-bonding interactions) and designate these basis sets 6-31++G(2d(X+),p) and 6-31+G(d(X+),p).

As a final test of the proposed methodology, a series of five hydrogen-bonded complexes (HF dimer, water dimer, hydrogen cyanide−HF, formaldehyde−water, and acetonitrile−HF) were calculated at a few representative and well-established levels of theory as well as at DFT/6-31++G(2d(X+),p)//DFT/6-31+G(d(X+),p) and MP2/6-31++G(2d(X+),p)//DFT/6-31+G(d(X+),p). The results are given in Table 7. All the methods using the 6-31G* basis set are clearly inadequate, as are CBS-4 and CBS-Lq. The former typically overestimates the strength of interaction, while the latter tends to underestimate it. CBS-Q, DFT/6-31++G(2d(X+),p), and MP2/6-31++G-(2d(X+),p), on the other hand, appear to yield values consistently in close agreement with one another and with the best reported calculations for water dimer (4.9−5.3 kcal/mol).[17−19,21,31,48] The MP2 values are consistently slightly higher than the CBS-Q numbers, while the DFT values show no pattern of systematic deviation. At least superficially, MP2 appears to give somewhat closer agreement with the highest levels of theory, but the corresponding BSSE corrections are also larger, leaving open to question whether the increases in binding energy are "real" (vide infra). We have therefore chosen the DFT energies as our "final" results. As an additional benefit, the DFT calculations can be used for larger systems than are accessible to MP2 calculation. Our final selected level of calculation is quite similar to that developed by Del Bene using MP2 methodology.[17] It is important to point out, however, that the results and conclusions presented below do not depend on our use of the DFT association energies in preference to the MP2 values. With the exception of one or two particular instances that are explicitly noted in the text, all of the analysis that follows could equally well have been carried out using the MP2 binding energies, and the same conclusions would be reached.

Scaling of Vibrational Frequencies. The calculated vibrational frequencies are necessary to make zero-point energy corrections. It is customary to scale the calculated harmonic frequencies in order to improve agreement with experiment, since it is well-known that the calculations consistently overestimate the vibrational frequencies. For HF/6-31G* calculations, the accepted scaling factor is 0.8934.[67] However, we have used B3LYP/6-31+G(d(X+),p) vibrational frequencies so as to correspond properly to the optimized geometries. It has been shown previously that density functional theory generally yields vibrational frequencies superior to those obtained at the HF level, or even the MP2 level.[70−73] An empirical scaling factor appropriate for B3LYP/6-31+G(d(X+),p) was obtained by computing both HF/6-31G* and B3LYP/6-31+G(d(X+),p) frequencies for all the molecules used in this study, but not the bimolecular complexes, and the data are presented in Table 8. Multiplication of the DFT frequencies by 0.97 was found to give the same average result as multiplying the HF/6-31G* frequencies by 0.8934, and so 0.97 was used as the DFT scale factor. Isolated molecules were used to establish the scale factor since the reference HF/6-31G* level of theory should provide an adequate description of individual molecules, even if it is somewhat deficient in the treatment of the bimolecular complexes.

Basis Set Superposition Error. It has been observed in numerous studies that counterpoise corrections to the BSSE are much larger at correlated levels of theory than at the HF level.[77] We have found that the BSSE is much smaller at the DFT level than at the MP2 level for the same basis set. There is a general lack of agreement as to whether BSSE counterpoise corrections are accurate and appropriate,[19,75] and so we have chosen not to include the BSSE correction in our final reported binding energies. We provide the counterpoise corrections in separate columns in Table 9 as an estimate of the error likely to stem from BSSE for both the DFT and MP2 binding energies. At least at the DFT level, however, the basis sets used are

TABLE 8: Zero-Point Energies of Molecules (hartrees)

compound	pt. grp.	HF/[a]3-21G*	HF/[b]6-31G*	DFT[c]
HF	$C_{\infty v}$	0.008 48	0.008 87	0.008 99
H_2O	C_{2v}	0.019 97	0.020 53	0.020 67
HCN	$C_{\infty v}$	0.016 84	0.016 07	0.015 76
H_2CO	C_{2v}	0.026 56	0.026 10	0.025 78
CH_3OH	C_s	0.049 95	0.049 44	0.049 68
CH_3NH_2	C_s	0.062 35	0.061 56	0.062 06
CH_3Cl	C_{3v}	0.037 21	0.036 36	0.036 66
CH_3SH	C_s	0.045 51	0.044 34	0.044 64
HCOOH (Z)	C_s	0.033 04	0.033 10	0.032 66
HCOOH (E)	C_s	0.032 47	0.032 73	0.032 34
CH_3CN	C_{3v}	0.045 27	0.043 70	0.043 85
CH_3CHO	C_s	0.054 81	0.053 54	0.053 71
$HCONH_2$	C_s	0.045 47	0.043 76	0.043 82
CH_3OCH_3	C_{2v}	0.078 44	0.076 96	0.077 22
CH_3SCH_3	C_{2v}	0.074 82	0.072 85	0.073 36
CH_3COOH (Z)	C_s	0.060 47	0.059 74	0.059 76
CH_3COOH (E)	C_s	0.060 03	0.059 46	0.059 56
$HCOOCH_3$ (Z)	C_s	0.061 13	0.060 37	0.060 05
$HCOOCH_3$ (E)	C_s	0.060 26	0.059 61	0.059 17
CH_3COCH_3	C_{2v}	0.082 44	0.080 34	0.080 89
$(CH_3)_3N$	C_{3v}	0.118 32	0.115 70	0.116 51
$HCONHCH_3$ (Z)	C_s	0.073 50	0.071 51	0.071 81
$HCONHCH_3$ (E)	C_s	0.073 30	0.071 48	0.071 73
CH_3NO_2	C_s	0.047 63	0.048 78	0.048 24
CH_3SOCH_3	C_s	0.078 72	0.076 69	0.076 74
CH_3COOCH_3 (Z)	C_s	0.088 29	0.086 72	0.086 77
CH_3COOCH_3 (E)	C_s	0.088 24	0.086 59	0.086 63
$CH_3CONHCH_3$ (Z)	C_s	0.100 62	0.097 89	0.098 70
$CH_3CONHCH_3$ (E)	C_s	0.100 84	0.098 16	0.098 53
imidazole	C_s	0.070 77	0.068 81	0.068 92
pyridine	C_{2v}	0.088 29	0.085 26	0.086 03
benzene	D_{6h}	0.099 75	0.096 20	0.097 43

[a] HF/3-21G* scaled by 0.9167. [b] HF/6-31G* scaled by 0.8934. [c] Becke3LYP/6-31+G(d(X+),p) scaled by 0.97.

sufficiently large that the BSSE appears to be below the intrinsic error limits in the calculations.

Full Set of Complexes. Table 9 lists the energies of association for water with a wide variety of small organic molecules, sometimes in more than one geometry. The energies are reported at the MP2 and DFT levels, using the optimized basis sets described above. Zero-point vibrational energy corrections are also listed, although they have not been explicitly added to the association energies reported. Where feasible, CBS-Q calculations were carried out, and the corresponding association energies are also listed in Table 9, while CBS-4 values are reported for all cases. BSSEs for the DFT and MP2 binding energies are also given, but have not been subtracted from the binding energies, as explained above. The MP2 association energies are consistently larger in magnitude than the corresponding DFT values, but are compensated by larger BSSEs. The DFT numbers are likely the most reliable overall, given the relatively smaller BSSE values, although the MP2 and CBS-Q values serve as an important check. The *ordering* of association energies within any consistent series except CBS-4 (i.e., DFT, MP2, or CBS-Q) is expected to be reliable.

Correlations between Different Levels of Theory. Although the different levels of calculation often yield significantly different absolute binding energies, there is an excellent linear relationship between the energies computed using the different procedures. The closest correlation occurs between energies computed using the same treatment of electron correlation, with only the basis set being different. Figure 2 compares the binding energies calculated at the DFT/6-31+G(d(X+),p) and the DFT/6-31++G(2d(X+),p)//DFT/6-31+G(d(X+),p) levels and shows a superb correlation ($R^2 = 0.999$), despite the small but significant difference in the absolute binding energies obtained using the two procedures. This correlation could be useful for accurately estimating the binding energy that would be calculated with the larger basis set when only calculations using the smaller basis set are feasible. This approach is demonstrated later in deriving the energy of association of the GC and AT DNA base pairs.

There is also a good correlation between the energies obtained at the DFT and MP2 levels ($R^2 = 0.94$) and between the DFT energies and the CBS-Q energies, as shown in Figure 3. The correlation between the MP2 and CBS-Q numbers is even closer ($R^2 = 0.993$). The generally excellent correlations lend confidence to the calculated ordering of association energies.

Nature of Hydrogen-Bonding Interactions. The computed geometries of the complexes provide some insights into the nature of the stabilizing interactions. Key results from the DFT/6-31+G(d(X+),p) optimized geometries are summarized in Table 10. The distances and angles that most simply define the hydrogen-bonding geometry are listed, along with comments that describe the overall structure (e.g., whether water is the donor or acceptor). For those complexes where the short descriptions do not suffice to define the geometry, full three-dimensional structures are shown in Figure 1. Each complex has been assigned an identifier code, listed in Table 9, that serves as a key for Figure 1 and the other tables. Most of the structures follow the expected patterns, with an X−H−Y angle approaching 180° and X−Y distances of about 3 Å. The linear hydrogen bond is preferred on electrostatic grounds and is known to be common. For instance, the water dimer and the complexes of water with methanol and dimethyl ether all conform exactly to expectations for a linear O−H···O arrangement and tetrahedral (sp^3-hybridized) interaction points on the acceptor oxygen (a "linear/tetrahedral" arrangement). Also as expected, the complexes in which water and methanol act as the hydrogen bond donors are similar in energy, differing by only about 0.5 kcal/mol, with methanol favored as the acceptor. The geometries of the methanol complexes very closely resemble what one would obtain by simply substituting a methyl group for one of the hydrogen atoms in the water dimer structure.

For methylamine and water, it was not possible to locate a minimum in which the amine was the hydrogen bond donor and water the hydrogen bond acceptor. This failure is not surprising, given that the amine is the stronger hydrogen bond acceptor, but the weaker donor, whereas water is the stronger donor but the weaker acceptor. The complex with methylamine as the hydrogen bond acceptor and water as the hydrogen bond donor was located without difficulty. It is more strongly bound than either the water dimer or any of the methanol−water complexes. The hydrogen bond is linear for the H−O···N angle and lies along the putative position of the lone pair on nitrogen.

The hydrogen bond between water and trimethylamine likewise conforms to a "linear/tetrahedral" pattern. The interaction is slightly weaker than for methylamine with the DFT calculations, but stronger from the MP2 results by 0.5 kcal/mol with BSSE included. The hydrogen bonding of water with imidazole and pyridine follows the expected "linear/trigonal planar" template at nitrogen when water is the donor, and the "linear/tetrahedral" arrangement at oxygen when water is the acceptor. The DFT results place pyridine with the saturated amines as a hydrogen bond acceptor, while imidazole shows somewhat enhanced basicity.

In aggregate, then, the patterns of hydrogen bonding follow predictions made on the basis of a simple electrostatic and hybridized valence bond view of hydrogen bonding. However, there are some deviations from the expected geometries. For

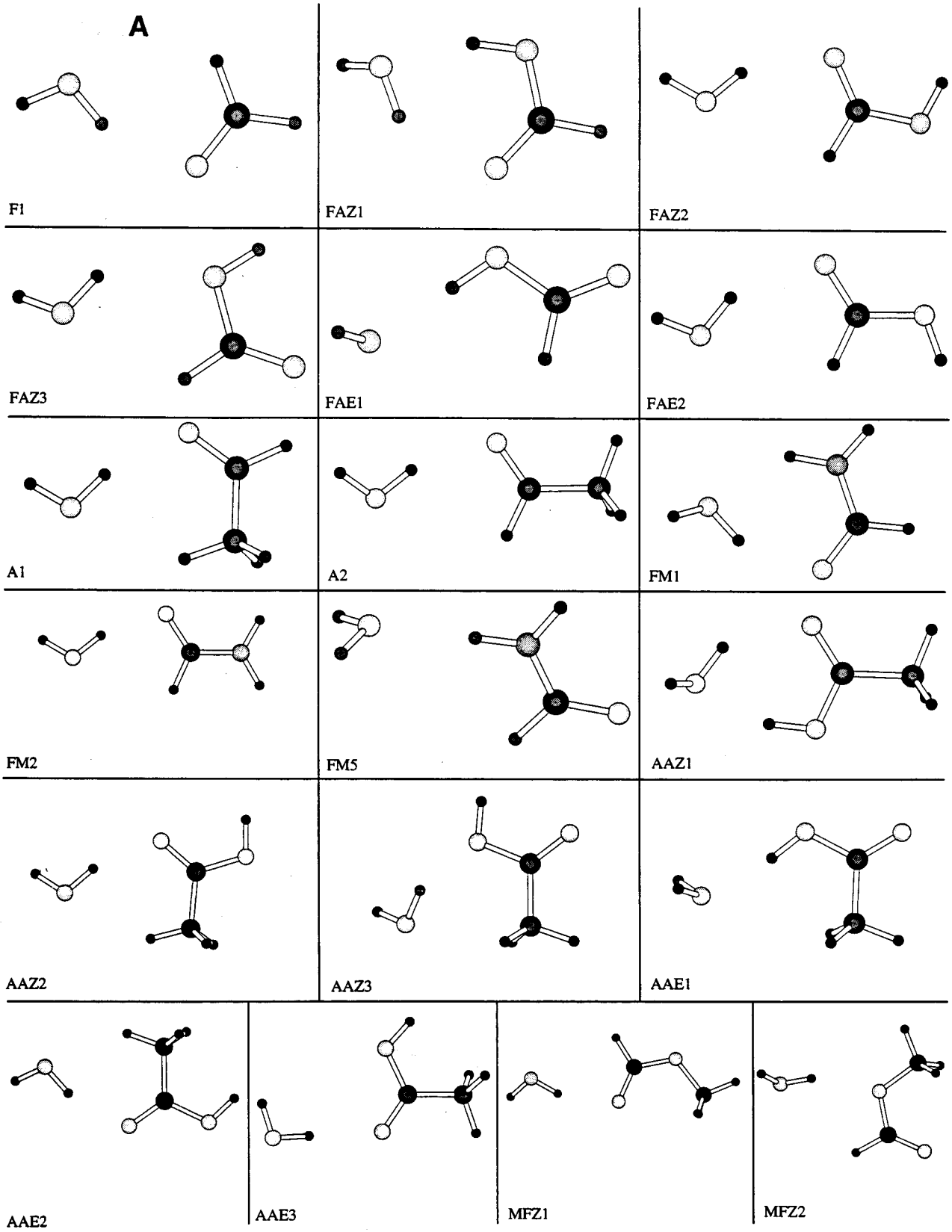
TABLE 9: Binding Energies ($-\Delta E$), BSSE, and Zero-Point Corrections of Complexes (kcal/mol)

compound	structure name	pt. grp.	CBS-4	CBS-Q	DFT[a]	DFT BSSE	MP2[b]	MP2 BSSE	DFTB[c]	ΔZPE[d]
HF−HF	HFD	C_s	3.97	4.54	4.61	0.11	4.76	0.77	4.79	1.79
HCN−HF	HCNHF	$C_{\infty v}$	7.56	7.56	7.71	0.38	7.89	1.14	7.96	2.08
CH_3CN-HF	MNHF	C_{3v}	9.54	9.26	9.49	0.37	9.48	1.27	9.81	2.01
H_2O-H_2O	WD	C_s	4.81	5.13	4.76	0.18	5.22	0.79	4.91	2.17
H_2CO-H_2O	F1	C_1	4.07	5.16	4.46	0.18	5.42	0.80	4.67	1.85
CH_3OH-H_2O	M1	C_1	4.07	5.69	5.09	0.22	6.07	1.02	5.25	1.89
CH_3OH-H_2O	M2	C_s	4.79	5.27	4.59	0.21	5.37	0.97	4.69	1.63
$CH_3NH_2-H_2O$	MM1	C_1	6.01	7.23	6.68	0.18	7.52	1.01	6.86	1.97
CH_3Cl-H_2O	MC1	C_1	1.21	3.31	2.64	0.16	3.80	0.80	2.75	1.23
CH_3SH-H_2O	MT1	C_1	1.68	3.90	3.64	0.18	4.69	1.07	3.72	1.41
CH_3SH-H_2O	MT2	C_s	2.06	2.50	1.84	0.13	2.79	0.66	1.88	0.97
HCOOH(Z)-H_2O	FAZ1	C_1	8.84	10.28	9.51	0.29	10.18	1.46	9.86	2.50
HCOOH(Z)-H_2O	FAZ2	C_1	4.00	5.23	4.43	0.16	5.28	0.82	4.68	1.64
HCOOH(Z)-H_2O	FAZ3	C_1	1.96	3.39	2.44	0.18	3.73	0.77	2.59	1.20
HCOOH(E)-H_2O	FAE1	C_s	7.78	8.51	7.73	0.18	8.44	1.05	7.96	1.73
HCOOH(E)-H_2O	FAE2	C_1	4.48	5.75	4.91	0.16	5.87	0.84	5.21	1.71
CH_3CHO-H_2O	A1	C_1	4.48	5.88	5.27	0.16	6.25	0.95	5.47	1.79
CH_3CHO-H_2O	A2	C_1	4.60	5.78	5.03	0.16	5.90	0.88	5.29	1.65
CH_3CN-H_2O	CN1	C_s	4.30	5.06	4.57	0.29	5.28	0.93	4.84	1.51
$HCONH_2-H_2O$	FM1	C_1	8.32	9.61	8.64	0.28	9.72	1.26	9.00	2.60
$HCONH_2-H_2O$	FM2	C_1	5.43	6.66	5.94	0.16	6.75	0.88	6.27	1.90
$HCONH_2-H_2O$	FM3	C_1	5.05	5.37	4.60	0.12	5.50	0.74	4.69	1.36
$CH_3OCH_3-H_2O$	DME1	C_1	4.19		4.98	0.24	6.60	1.29	5.15	1.78
$CH_3SCH_3-H_2O$	DMS1	C_1	1.66		4.12	0.18	5.55	1.16	4.38	1.51
$CH_3COOH(Z)-H_2O$	AAZ1	C_1	8.70		9.46	0.28	10.16	1.53	9.86	2.37
$CH_3COOH(Z)-H_2O$	AAZ2	C_1	4.87		5.31	0.18	6.28	1.00	5.52	1.72
$CH_3COOH(Z)-H_2O$	AAZ3	C_1	2.50		2.91	0.16	4.34	0.92	3.02	1.31
$CH_3COOH(E)-H_2O$	AAE1	C_1	7.12		6.89	0.23	8.04	1.17	7.12	1.62
$CH_3COOH(E)-H_2O$	AAE2	C_1	5.28		5.79	0.18	6.83	1.02	6.06	1.75
$CH_3COOH(E)-H_2O$	AAE3	C_1	4.62		4.49	0.14	5.27	0.83	4.65	1.45
$HCOOCH_3(Z)-H_2O$	MFZ1	C_s	4.89		4.74	0.16	5.57	0.90	4.97	1.60
$HCOOCH_3(Z)-H_2O$	MFZ2	C_1	2.17		2.51	0.16	3.88	0.92	2.65	1.33
$HCOOCH_3(Z)-H_2O$	MFZ3	C_s	2.96		4.99	0.16	6.30	1.07	5.17	1.68
$HCOOCH_3(E)-H_2O$	MFE1	C_1	4.70		5.35	0.19	6.01	0.85	5.65	1.85
$HCOOCH_3(E)-H_2O$	MFE2	C_1	4.77		4.71	0.21	5.23	0.88	4.81	1.68
$CH_3COCH_3-H_2O$	AN1	C_1	4.96		5.73	0.18	6.70	1.04	5.94	1.83
$(CH_3)_3N-H_2O$	TA1	C_1	6.30		6.36	0.27	8.72	1.74	6.65	1.90
$HCONHCH_3(Z)-H_2O$	NMFZ1	C_1	5.28		6.35	0.17	7.89	1.16	6.58	1.85
$HCONHCH_3(Z)-H_2O$	NMFZ2	C_1	5.25		4.29	0.11	5.80	0.91	4.41	1.10
$HCONHCH_3(E)-H_2O$	NMFE1	C_1	8.84		8.95	0.26	10.28	1.46	9.33	2.42
$HCONHCH_3(E)-H_2O$	NMFE2	C_1	5.84		6.34	0.15	7.15	0.93	6.66	1.84
$CH_3NO_2-H_2O$	NM1	C_1	3.71		4.65	0.21	4.07	1.05	4.83	1.68
$CH_3SOCH_3-H_2O$	DMSO1	C_1	8.45		8.30	0.24	10.49	1.78	8.87	2.15
$CH_3SOCH_3-H_2O$	DMSO2	C_1	6.05		7.01	0.21	8.87	1.49	7.51	1.76
$CH_3COOCH_3(Z)-H_2O$	MAZ1	C_1	3.36		5.28	0.17	6.58	1.12	5.48	1.65
$CH_3COOCH_3(Z)-H_2O$	MAZ2	C_1	5.05		5.55	0.18	6.60	1.12	5.77	1.75
$CH_3COOCH_3(E)-H_2O$	MAE1	C_1	5.67		5.16	0.19	6.03	0.96	5.38	1.63
$CH_3COOCH_3(E)-H_2O$	MAE2	C_1	5.42		6.01	0.20	7.13	1.08	6.34	1.82
$CH_3CONHCH_3(Z)-H_2O$	NMAZ1	C_1	5.35		6.68	0.20	8.35	1.20	6.93	1.82
$CH_3CONHCH_3(Z)-H_2O$	NMAZ2	C_1	6.32		6.82	0.17	8.01	1.15	7.13	1.88
$CH_3CONHCH_3(E)-H_2O$	NMAE1	C_1	9.00		9.18	0.30	10.38	1.51	9.53	2.44
$CH_3CONHCH_3(E)-H_2O$	NMAE2	C_1	6.46		6.97	0.23	7.94	1.12	7.23	1.95
imidazole-H_2O	IM1	C_1	5.84		6.85	0.16	7.79	1.10	6.95	1.77
imidazole-H_2O	IM2	C_s	6.09		5.20	0.16	6.69	0.98	5.33	1.17
pyridine-H_2O	PYR1	C_1	5.74		6.26	0.23	7.32	1.17	6.41	1.77
benzene-H_2O	BZ1	C_1	4.16		1.69	0.28	3.45	0.83	1.78	0.78

[a] DFT = Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p). [b] MP2 = MP2(fc)/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p). [c] DFTB = Becke3LYP/6-31+G(d(X+),p) [d] Change in zero-point vibrational energy upon complexation, calculated at Becke3LYP/6-31+G(d(X+),p) and scaled by 0.97.

instance, the complexes in which chloromethane, dimethyl sulfide, and methanethiol accept a hydrogen bond from water do not conform to the usual pattern. The structures are shown in Figure 4, from which it is apparent that the H−O···X angle departs quite strongly from linearity in each case (see also Table 10). Furthermore, the region of the heteroatom (either sulfur or chlorine) to which the water OH bond "points" does not correspond to the location expected for sp^3-hybridized lone pairs. The latter discrepancy can be understood based on the tendency of second-row atoms such as sulfur and chlorine to adopt nonhybridized geometries; for example, bond angles at sulfur and phosphorus are generally closer to 90° than to 109.5°. However, the overall impression suggested by the structures is of a more complex interaction between the molecules that involves optimization of dipole−dipole and dispersion interactions and an absence of true hydrogen bonding. In each case, the water molecule arranges itself so that its total dipole moment, not just the dipole moment from a single OH bond, is aligned in an antiparallel fashion with respect to the total dipole moment of its complexation partner.

The complexes of water with aldehydes and ketones (formaldeyde, acetaldehyde, and acetone) suggest that in these cases



the total molecular dipole also affects the optimal geometry. The water molecule always distorts away from the idealized geometry in which the O···H−O angle would be 180° and C═O···H angle would be 120° (position of an sp^2-hybridized lone pair). The distortion serves to align the water molecule's dipole moment more closely antiparallel to that of the carbonyl bond. Similar distortions from the idealized "linear/trigonal planar" geometry are observed for the other carbonyl functionalities examined, i.e., carboxylic acids, esters, and amides. The unusually strong DMSO−water complex (DMSO1, Figure 1) also appears to have a geometry decisively influenced by the interaction of the overall molecular dipole moments.

Carboxylic acids and amides give the strongest complexes of all, with some binding energies as large as 8−10 kcal/mol.

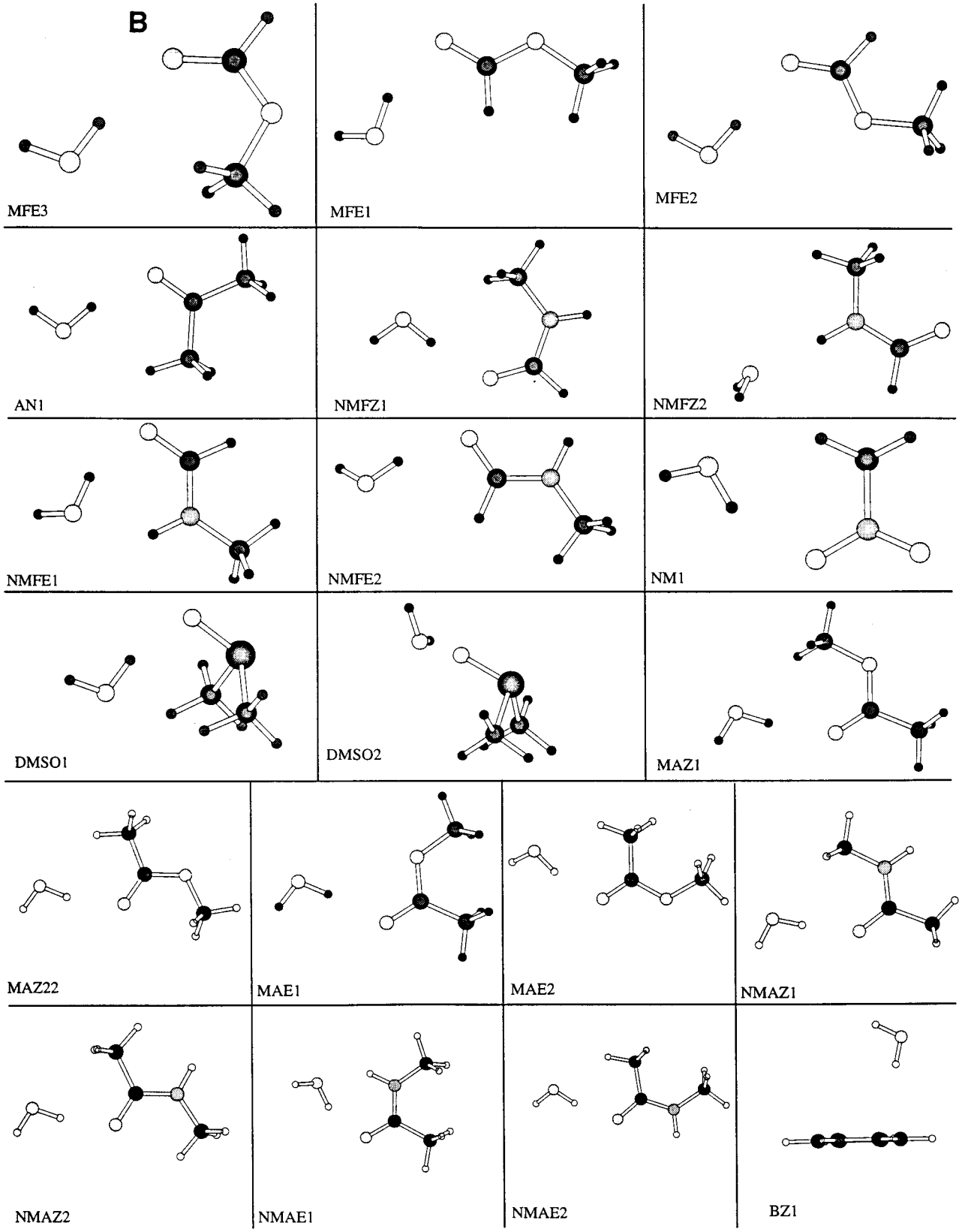


Figure 1. Becke3LYP/6-31+G(d(X+),p) optimized geometries of selected hydrogen-bonded complexes.

The potential energy surfaces generally have multiple minima, not all of which are equally deep. However, the most strongly bound structures are the ones in which there is some component of bidentate binding; that is, more than one hydrogen-bonding interaction is present (Figure 1). The optimized geometries show a cyclic arrangement with donation of a hydrogen bond by the carboxylic acid hydroxy group or amide amino group to water and with a second hydrogen bond in which water donates to the carbonyl oxygen. These complexes, not surprisingly, show quite marked deviations from the idealized 180° XHY angle, as a result of the geometric constraints of the small-ring cyclic nature of the structures. It is notable that for both

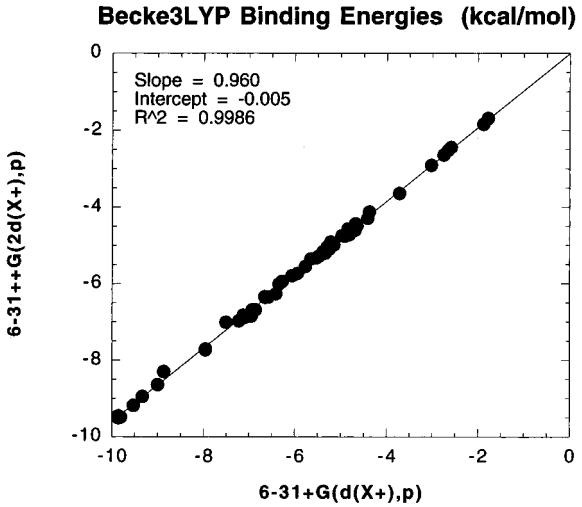


Figure 2. Comparison of association energies calculated using DFT (Becke3LYP functional) and two different basis sets: 6-31+G(d(X+),p) and 6-31++G(2d(X+),p). The Becke3LYP/6-31+G(d(X+),p) optimized geometries were used for both sets of calculations. The best fit line yields the relationship E(bigger basis) $= -0.005 + 0.960 \times E$(smaller basis). The correlation coefficient is 0.9986.

formic and acetic acid the most favorable complex for the Z conformer is considerably more strongly bound than that for the E conformer. Likewise for *N*-methylacetamide and *N*-methylformamide, the E isomer, in which the NH hydrogen is syn to the carbonyl, binds more strongly to water than does the Z isomer.

Benzene forms a weak complex with water of the sort previously reported,[79] in which one of the OH bonds of water points toward the π system of the aromatic ring. Benzene is thus different from all the other acceptor species studied here, in that the site of hydrogen bond acceptance is not associated with a specific atom. Benzene also represents one of the few cases reported here where the DFT and MP2 binding energies are not in good agreement, perhaps as a result of dispersion interactions playing an increased role.

Relative Hydrogen Bond Acceptor Ability. By visual examination of their geometries it is possible to divide the various complexes into three classes according to whether water is the hydrogen bond donor, acceptor, or both. The complexes in which water is clearly the donor can then further be categorized according to the nature of the atom acting as the hydrogen bond acceptor site. Table 11 shows the results of such an analysis, along with the average value and range for the binding energies obtained in each category. The variation of energies within a category turns out to be remarkably small, so that the average binding energy is strongly characteristic of a given type of hydrogen bond. The average energies then allow the assignment of an ordering of hydrogen bond acceptor abilities for typical functional groups, and the entries in Table

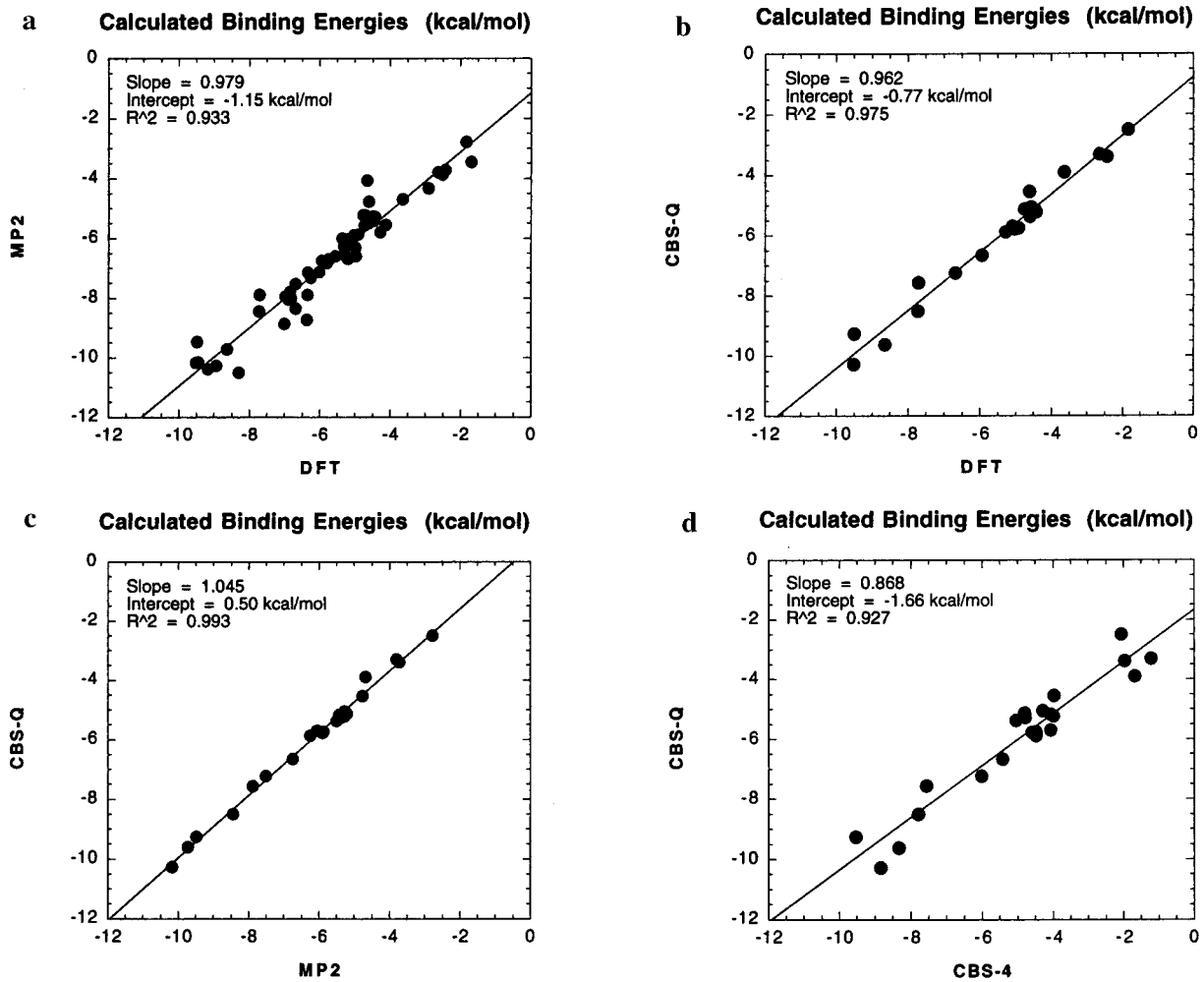


Figure 3. Comparison of association energies computed using different levels of theory: (a) MP2/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) vs Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p); (b) CBS-Q vs Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p); (c) CBS-Q vs MP2/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p). (d) CBS-Q vs CBS-4.

TABLE 10: Selected Geometric Parameters of Complexes from Becke3LYP/6-31+G(d(X+),p) Optimized Structures

complex	structure name	pt. grp.	r(X−H) (Å)	∠(X−H−Y) (deg)	comment[a]
HF−HF	HFD	C_s	1.802 (FH−F)	169.1 (FHF)	planar, bent
HCN−HF	HCNHF	$C_{\infty v}$	1.818 (CN−H)	180.0 (NHF)	linear
CH_3CN-HF	MNHF	C_{3v}	1.766 (CN−H)	180.0 (NHF)	linear
H_2O-H_2O	WD	C_s	1.949 (OH−O)	170.1 (OHO)	symmetric
H_2CO-H_2O	F1	C_1	1.977 (CO−H)	153.4 (OHO)	H_2O as D
CH_3OH-H_2O	M1	C_1	1.911 (CO−H)	170.6 (OHO)	H_2O as D
CH_3OH-H_2O	M2	C_s	1.959 (OH−O)	171.6 (OHO)	H_2O as A
$CH_3NH_2-H_2O$	MM1	C_1	1.920 (CN−H)	169.5 (NHO)	H_2O as D
CH_3Cl-H_2O	MC1	C_1	2.545 (CCl−H)	143.4 (ClHO)	H_2O as neither D nor A
CH_3SH-H_2O	MT1	C_1	2.449 (CS−H)	159.6 (SHO)	H_2O as D
CH_3SH-H_2O	MT2	C_s	2.288 (SH−O)	177.6 (SHO)	H_2O as A
HCOOH(Z)-H_2O	FAZ1	C_1	1.775 (OH−O)	157.1 (OHO)	H_2O as both D and A
			2.016 (C═O−H)	136.5 (OHO)	(D to carbonyl, A to hydroxy)
HCOOH(Z)-H_2O	FAZ2	C_1	1.992 (C═O−H)	151.4 (OHO)	H_2O as D to carbonyl
HCOOH(Z)-H_2O	FAZ3	C_1	2.212 (C−O−H)	132.2 (OHO)	H_2O as D to hydroxy
HCOOH(E)-H_2O	FAE1	C_s	1.814 (OH−O)	178.6 (OHO)	H_2O as A
HCOOH(E)-H_2O	FAE2	C_1	1.991 (C═O−H)	148.7 (OHO)	H_2O as D to carbonyl
CH_3CHO-H_2O	A1	C_1	1.941 (CO−H)	162.6 (OHO)	H_2O as D, syn to methyl
CH_3CHO-H_2O	A2	C_1	1.931 (CO−H)	159.2 (OHO)	H_2O as D, anti to methyl
CH_3CN-H_2O	CN1	C_s	2.063 (CN−H)	178.4 (NHO)	H_2O as D
$HCONH_2-H_2O$	FM1	C_1	1.893 (CO−H)	151.6 (OHO)	H_2O as both D and A
			2.078 (NH−O)	137.3 (NHO)	(D to carbonyl, A to amino)
$HCONH_2-H_2O$	FM2	C_1	1.892 (CO−H)	160.7 (OHO)	H_2O as D to carbonyl
$HCONH_2-H_2O$	FM3	C_1	2.040 (NH−O)	179.8 (NHO)	H_2O as A
$CH_3OCH_3-H_2O$	DME1	C_1	1.899 (CO−H)	173.8 (OHO)	H_2O as D
$CH_3SCH_3-H_2O$	DMS1	C_1	2.391 (CS−H)	159.4 (SHO)	H_2O as D
$CH_3COOH(Z)-H_2O$	AAZ1	C_1	1.793 (OH−O)	157.2 (OHO)	H_2O as both D and A
			1.946 (CO−H)	140.3 (OHO)	(D to carbonyl, A to hydroxy)
$CH_3COOH(Z)-H_2O$	AAZ2	C_1	1.936 (C═O−H)	161.9 (OHO)	H_2O as D to carbonyl
$CH_3COOH(Z)-H_2O$	AAZ3	C_1	2.085 (C−O−H)	155.8 (OHO)	H_2O as D to hydroxy
$CH_3COOH(E)-H_2O$	AAE1	C_1	1.858 (OH−O)	169.8 (OHO)	H_2O as A
$CH_3COOH(E)-H_2O$	AAE2	C_1	1.923 (C═O−H)	160.3 (OHO)	H_2O as D to carbonyl
$CH_3COOH(E)-H_2O$	AAE3	C_1	2.003 (C═O−H)	179.8 (OHO)	H_2O as D to carbonyl
$HCOOCH_3(Z)-H_2O$	MFZ1	C_1	1.955 (C═O−H)	158.0 (OHO)	H_2O as D to carbonyl
$HCOOCH_3(Z)-H_2O$	MFZ2	C_1	2.058 (C−O−H)	157.8 (OHO)	H_2O as D to hydroxy
$HCOOCH_3(Z)-H_2O$	MFZ3	C_s	1.964 (C═O−H)	167.9 (OHO)	H_2O as D to carbonyl
$HCOOCH_3(E)-H_2O$	MFE1	C_1	1.954 (C═O−H)	154.1 (OHO)	H_2O as D to carbonyl
$HCOOCH_3(E)-H_2O$	MFE2	C_1	2.002 (C═O−H)	178.7 (OHO)	H_2O as D to carbonyl
$CH_3COCH_3-H_2O$	AN1	C_1	1.908 (CO−H)	165.6 (OHO)	H_2O as D
$(CH_3)_3N-H_2O$	TA1	C_1	2.102 (CN−H)	170.2 (NHO)	H_2O as D
$HCONHCH_3(Z)-H_2O$	NMFZ1	C_1	1.892 (CO−H)	171.9 (OHO)	H_2O as D to carbonyl
$HCONHCH_3(Z)-H_2O$	NMFZ2	C_1	2.058 (NH−O)	176.9 (NHO)	H_2O as A
$HCONHCH_3(E)-H_2O$	NMFE1	C_1	1.864 (CO−H)	152.8 (OHO)	H_2O as both D and A
			2.074 (NH−O)	140.1 (NHO)	(D to carbonyl, A to amino)
$HCONHCH_3(E)-H_2O$	NMFE2	C_1	1.874 (CO−H)	162.7 (OHO)	H_2O as D to carbonyl
$CH_3NO_2-H_2O$	NM1	C_1	2.032 (NO−H)	152.5 (OHO)	H_2O as D
$CH_3SOCH_3-H_2O$	DMSO1	C_1	1.822 (SO−H)	159.7 (OHO)	H_2O as D
$CH_3SOCH_3-H_2O$	DMSO2	C_1	1.868 (SO−H)	162.2 (OHO)	H_2O as D
$CH_3COOCH_3(Z)-H_2O$	MAZ1	C_s	1.938 (C═O−H)	170.8 (OHO)	H_2O as D to carbonyl
$CH_3COOCH_3(Z)-H_2O$	MAZ2	C_1	1.916 (C═O−H)	164.5 (OHO)	H_2O as D to carbonyl
$CH_3COOCH_3(E)-H_2O$	MAE1	C_1	1.976 (C═O−H)	176.7 (OHO)	H_2O as D to carbonyl
$CH_3COOCH_3(E)-H_2O$	MAE2	C_1	1.901 (C═O−H)	162.7 (OHO)	H_2O as D to carbonyl
$CH_3CONHCH_3(Z)-H_2O$	NMAZ1	C_1	1.875 (CO−H)	174.8 (OHO)	H_2O as D to carbonyl
$CH_3CONHCH_3(Z)-H_2O$	NMAZ2	C_1	1.853 (CO−H)	168.8 (OHO)	H_2O as D to carbonyl
$CH_3CONHCH_3(E)-H_2O$	NMAE1	C_1	1.833 (CO−H)	155.2 (OHO)	H_2O as both D and A
			2.074 (NH−O)	143.1 (NHO)	(D to carbonyl, A to amino)
$CH_3CONHCH_3(E)-H_2O$	NMAE2	C_1	1.850 (CO−H)	167.6 (OHO)	H_2O as D to carbonyl
imidazole-H_2O	IM1	C_1	1.936 (CN−H)	170.8 (NHO)	H_2O as D
imidazole-H_2O	IM2	C_s	1.994 (NH−O)	178.6 (NHO)	H_2O as A
pyridine-H_2O	PYR1	C_1	1.932 (CN−H)	176.3 (NHO)	H_2O as D
benzene-H_2O	BZ1	C_1	2.605 (CC−H)	165.1 (CHO)	H_2O as D

[a] D = hydrogen bond donor; A = hydrogen bond acceptor.

11 are listed in just such an order. Although the DFT association energies are given in Table 11, the same ordering and almost the same average magnitudes result from using the BSSE-corrected MP2 energies. The only exceptions are the nitro group, which is predicted to be 1.6 kcal/mol weaker at the MP2 level, and the benzene ring, which is predicted to be significantly stronger at the MP2 level. Aside from these two cases, the difference between the average energy at the DFT and the BSSE-corrected MP2 levels is generally within ±0.3 kcal/mol and is always within ±0.5 kcal/mol.

The very strongest acceptor type appears to be the oxygen of DMSO, which defines its own category with a binding energy of 7.7 kcal/mol. The next best acceptor type is nitrogen, with an average binding energy to water of 6.5 kcal/mol. The hybridization seems to matter little, with imidazole and pyridine yielding strengths close to those for methylamine and trimeth-

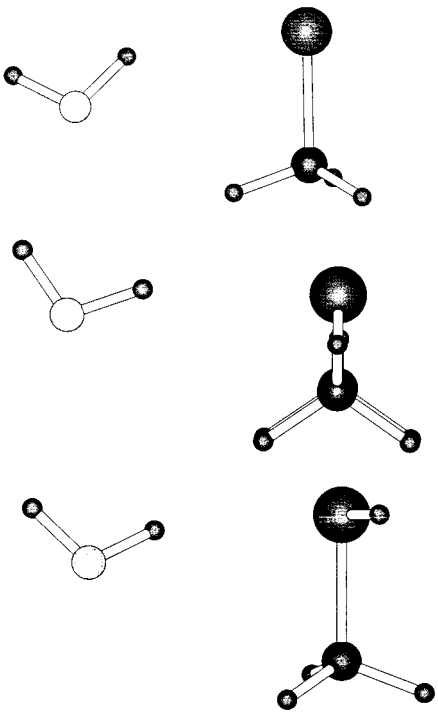


Figure 4. Becke3LYP/6-31+G(d(X+),p) optimized geometries of chloromethane, methanethiol, and dimethyl sulfide as hydrogen bond acceptors with water.

ylamine. The carbonyl oxygen of amides is an equally good acceptor, with the same average binding energy, 6.5 kcal/mol. The other types of carbonyls (esters, acids, ketones, and aldehydes) all give average strengths of 5.1 ± 0.1 kcal/mol, while sp^3-hybridized oxygens are slightly weaker acceptors (4.9 kcal/mol). Cyano and nitro groups also fall in this region, with an average acceptor strength of 4.6 kcal/mol. Sulfur is a still weaker acceptor (3.9 kcal/mol), while the very weakest acceptors of all by far are the saturated oxygen atoms of esters and acids (2.6 kcal/mol).[77] The latter follows from the presence of the adjacent electron-withdrawing carbonyl group.

Relative Hydrogen Bond Donor Ability. Similarly, it is possible to catalog hydrogen bond donor abilities, as shown in the lower part of Table 11. The best donor by far is the hydroxy group of a carboxylic acid (7.3 kcal/mol), as a result of its high acidity. Alcoholic OH and amide or imidazole NH groups are donors of intermediate strength, having average energies of 4.7 kcal/mol. The SH group is clearly a very weak donor, yielding a binding energy of only 1.8 kcal/mol.

The strongest complexes of all are obtained when water participates in two hydrogen bonds, one as a donor and one as an acceptor. The average strength of interaction in this bidentate category is 9.2 kcal/mol. No cases were found in which water unambiguously acted as the donor in two hydrogen bonds, although the geometry of the AAE3 complex is suggestive of such an arrangement.

The bidentate systems have been listed separately from the monodentate donor and acceptor complexes in Table 11. The binding energies are generally strong, but it is difficult to draw further conclusions of a general nature for these systems.

Why Is Sulfur Almost as Good a Hydrogen Bond Acceptor as Oxygen? The structure in which methanethiol is a hydrogen bond donor to water quite closely resembles that in which methanol is a donor to water, although the former interaction is, not surprisingly, only about half as strong (1.84 kcal/mol vs 4.59 kcal/mol). The lesser electronegativity of sulfur relative to that of oxygen results in sulfur having a much smaller negative charge. This, coupled with the larger radius of sulfur, greatly reduces its ability to participate in stabilizing electrostatic interactions. When water acts as the donor, however, the interaction energy of 3.64 kcal/mol is closer to that for the corresponding methanol−water complex (5.09 kcal/mol). The somewhat surprising strength of this interaction might result from the more polarizable sulfur atom interacting with the polar O−H bond in a dipole−induced dipole sense and from the greater basicity of the sulfur lone pair. The former factor was examined further by calculating the quantity of charge reorganization occurring during complexation, as described below.

Crystallographers have long used deformation density plots to depict the differences in electron density for a molecule relative to its constituent atoms.[78,79] A similar approach can be used to compare the calculated charge density distributions for related molecules. For instance, charge density difference plots have been used to illustrate the reorganization of charge upon electronic excitation,[80] to depict the effect of hydrogen bonding,[81] to visualize and quantify the intramolecular charge transfer that occurs during bond rotation,[82−84] and to understand the nature of bonding itself.[85,86]

The same approach is used here to compute the charge density reorganization that occurs upon complexation of water with methanol, methanethiol, dimethyl ether, dimethyl sulfide, or a second water molecule. The Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) charge density is first computed for the entire complex. From this total density, the independent charge density distributions of the individual molecules comprising the complex are then subtracted. For instance, for the dimer of water with methanol, the distributions for water and methanol are subtracted from that for the dimer.[90] The difference density derived in this manner for water dimer is shown in Figure 5, where the regions enclosed by solid lines represent excess charge density in the dimer relative to the monomers, while the regions enclosed by dashed lines represent regions of depleted charge density. The three-dimensional contours have been drawn at the 0.001 electrons per cubic Bohr surface. Figure 6 depicts the difference density distributions for the hydrogen-bonded complexes of methanol, dimethyl ether, methanethiol, and dimethyl sulfide with water. It is readily apparent by visual inspection that the amount of polarization is similar for oxygen and sulfur.

The extent of electronic reorganization occurring upon complexation was further quantified by direct integration of the difference density distributions appearing in Figures 5 and 6, and the results are provided in Table 12.[88] The total charge reorganization computed in this manner is actually *less* for the cases involving sulfur than for those involving only oxygen. This result suggests that the greater polarizability of sulfur relative to oxygen is *not* the cause of its unexpectedly strong ability to accept a hydrogen bond. Table 12 also provides numerical integrations for the difference densities computed at the MP2 level, and it is readily apparent that the results are virtually identical to those obtained using DFT.

Correlation of Hydrogen Bond Strength with Experimental Gas-Phase Proton Affinity. Experimental gas-phase proton affinities are available for many of the small organic compounds included in this study[89] and are listed in Table 13. Insofar as a hydrogen bond reflects partial transfer of a proton, it would seem logical for the hydrogen bond energy to correlate closely with basicity. Correlations between proton affinity and hydrogen bond geometries[90] and strengths[91,92] have been observed

TABLE 11: Categorization of Complexes by Hydrogen Bond Type

category	structures	$-\Delta E^a$ (kcal/mol)	average $-\Delta E^b$ (kcal/mol)	$-\Delta E$ range[c] (kcal/mol)
DMSO O as acceptor	DMSO1	8.30	7.66	1.29
	DMSO2	7.01		
sp^3 or sp^2 N as acceptor	IM1	6.85	6.54	0.59
	MM1	6.68		
	TA1	6.36		
	PYR1	6.26		
amide carbonyl O as acceptor	FM2	5.94	6.52	1.03
	NMFZ1	6.35		
	NMFE2	6.34		
	NMAZ1	6.68		
	NMAZ2	6.82		
	NMAE2	6.97		
ester carbonyl O as acceptor	MFZ1	4.74	5.22	1.30
	MFZ3	4.99	(5.11)[d]	(0.84)[d]
	MFE1	5.35		
	MFE2	4.71		
	MAZ1	5.28		
	MAZ2	5.55		
	MAE1	5.16		
	MAE2	6.01		
aldehyde and ketone O as acceptor	F1	4.46	5.12	1.27
	A1	5.27		
	A2	5.03		
	AN1	5.73		
acid carbonyl O as acceptor	FAZ2	4.43	4.99	1.36
	FAE2	4.91		
	AAZ2	5.31		
	AAE2	5.79		
	AAE3	4.49		
sp^3 O as acceptor	WD	4.76	4.86	0.50
	M1	5.09		
	M2	4.59		
	DME1	4.98		
nitro O as acceptor	NM1	4.65	4.65	
nitrile sp N as acceptor	CN1	4.57	4.57	
S as acceptor	MT1	3.64	3.88	0.48
	DMS1	4.12		
Cl as acceptor	MC1	2.64	2.64	
acid/ester −O− as acceptor	FAZ3	2.44	2.62	0.47
	AAZ3	2.91		
	MFZ2	2.51		
acid OH as donor	FAE1	7.73	7.31	0.84
	AAE1	6.89		
amide NH as donor	FM3	4.60	4.70	0.91
	NMFZ2	4.29		
	IM2	5.20		
alcohol OH as donor	WD	4.76	4.68	0.17
	M2	4.59		
SH as donor	MT2	1.84	1.84	
bidentate[e]	FAZ1	9.51	9.15	0.87
	FM1	8.64		
	AAZ1	9.46		
	NMFE1	8.95		
	NMAE1	9.18		

[a] Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) energy. [b] Average of energies in previous column for given category. [c] Difference between largest and smallest binding energy for given category. [d] Average excluding the anomalously large value −6.01 kcal/mol for MAE2. [e] Complexes in which water simultaneously acts as a hydrogen bond donor and acceptor.

in the past. Figure 7, which plots the calculated association energies of the molecules in Table 13 against the corresponding gas-phase proton affinities, demonstrates that such a correlation does indeed exist. Not all the calculated complexes are represented in Figure 7; for each compound, only the single most strongly bound complex for which water is clearly the hydrogen bond donor has been plotted. Thus, for instance, the complexes of formamide, acetic acid, and formic acid in which water binds in a bidentate manner have been excluded.

The DFT binding energies yield a correlation coefficient of 0.75, provided the outlying points for DMSO and benzene are omitted. Exclusion of the sulfur bases methanethiol and dimethyl sulfide, which fall roughly 1 kcal/mol above the line, further improves the correlation to 0.87. The MP2 energies yield a similar fit, with a correlation coefficient of 0.85, although the slope is significantly steeper. Despite the slight deviation from the best fit line for sulfur species, the close correlation between binding energy and basicity helps to explain the seemingly anomalous relative strengths of hydrogen bonds to sulfur and oxygen. Sulfur lone pairs can be considered to be more basic than oxygen lone pairs, as shown by the greater proton affinity of methanethiol relative to methanol and of dimethyl sulfide relative to dimethyl ether, and for that reason the complexes in which sulfur accepts a hydrogen bond are quite stable. That benzene does not fall on the line is perhaps not surprising, as the site of protonation does not correspond to

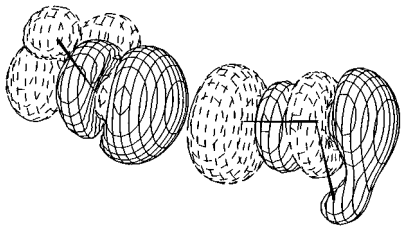


Figure 5. Difference density map showing the charge redistribution occurring upon dimerization of water. The contour level is 0.001 electrons per cubic Bohr. The solid lines represent regions of positive difference density (more electron density in the dimer than in the isolated monomers), while dashed lines represent regions of negative difference density (less electron density in the dimer than in the isolated monomers). Charge densities were calculated at the Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level.

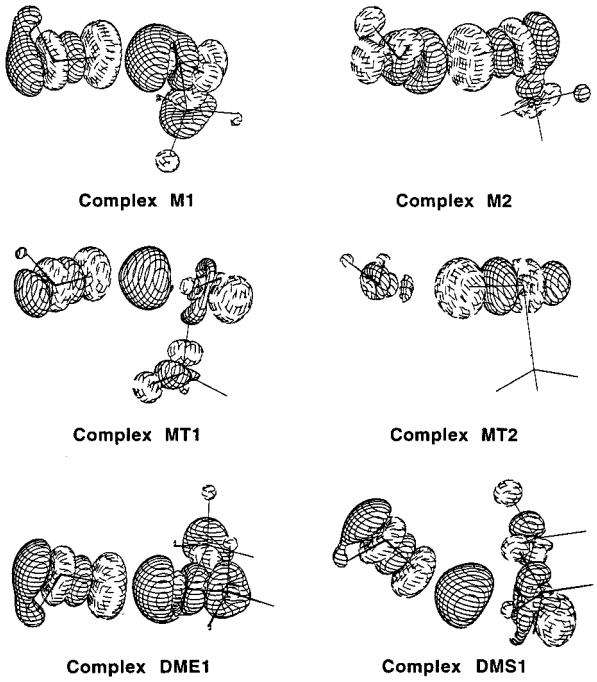


Figure 6. Difference density maps showing the charge redistribution occurring upon formation of the following hydrogen-bonded complexes: (a) methanol−water (water as hydrogen bond donor); (b) methanol−water (water as hydrogen bond acceptor); (c) methanethiol−water (water as hydrogen bond donor); (d) methanethiol−water (water as hydrogen bond acceptor); (e) dimethyl ether−water; (f) dimethyl thioether−water. The contour level is 0.001 electrons per cubic Bohr. The solid lines represent regions of positive difference density (more electron density in the dimer than in the isolated monomers), while dashed lines represent regions of negative difference density (less electron density in the dimer than in the isolated monomers). Charge densities were calculated at the Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level.

TABLE 12: Integration of Charge Difference Densities for Formation of Hydrogen Bonds

complex	name	pt. grp.	integration (Becke3LYP)[a]	integration (MP2)[b]
H_2O-H_2O	WD	C_s	0.0778	0.0783
CH_3OH-H_2O	M1	C_1	0.0947	0.0944
CH_3OH-H_2O	M2	C_s	0.0868	0.0865
CH_3SH-H_2O	MT1	C_1	0.0821	0.0814
CH_3SH-H_2O	MT2	C_s	0.0594	0.0597
$CH_3OCH_3-H_2O$	DME1	C_1	0.1035	0.1031
$CH_3SCH_3-H_2O$	DMS1	C_1	0.0973	0.0955

[a] Based on Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) difference density. [b] Based on MP2/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) difference density.

TABLE 13: Experimental Gas-Phase Proton Affinities and Solution Hydrogen Bond Basicities (kcal/mol)

compound	proton affinity[a]	β_2^H [b]	dipole moment[c]
HF			1.83
H_2O	166.5	0.38	1.86
HCN	171.4		3.09
H_2CO	171.7		2.44
CH_3OH	181.9	0.41	1.69
CH_3NH_2	214.1	0.70	1.34
CH_3Cl	163.0	0.15	2.03
CH_3SH	187.4	0.16	1.59
HCOOH (Z)	178.8		1.57
HCOOH (E)			3.97
CH_3CN	188.4	0.44	4.09
CH_3CHO	186.6	0.40	2.93
$HCONH_2$	198.4		4.01
CH_3OCH_3	192.1	0.43	1.30
CH_3SCH_3	200.6	0.28	1.65
CH_3COOH (Z)	190.2		1.82
CH_3COOH (E)			4.45
$HCOOCH_3$ (Z)	188.9		2.00
$HCOOCH_3$ (E)			4.35
CH_3COCH_3	196.7	0.50	3.13
$(CH_3)_3N$	225.1	0.61	0.54
$HCONHCH_3$ (Z)	205.8		4.05
$HCONHCH_3$ (E)			4.38
CH_3NO_2	179.2	0.25	3.66
CH_3SOCH_3	211.3	0.78	4.19
CH_3COOCH_3 (Z)	197.8		1.95
CH_3COOCH_3 (E)			4.66
$CH_3CONHCH_3$ (Z)			3.89
$CH_3CONHCH_3$ (E)			4.27
imidazole	219.8		3.73
pyridine	220.8	0.62	2.31
benzene	181.3	0.14	0.00

[a] Experimental gas-phase proton affinities taken from Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695 (ref 89). [b] Abraham's experimental solution hydrogen bond basicity parameter, β_2^H, from Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 521−529 (ref 96). [c] Dipole moment in debye units, calculated at the Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level of theory.

a specific atom. Benzene is thus qualitatively different from all the other cases studied here. The reason DMSO yields a complex so much more tightly bound than expected on the basis of gas-phase proton affinity data is not clear. The site of both protonation and hydrogen bond acceptance on DMSO is oxygen, not sulfur.

Numerous experimental scales of solution hydrogen bond acceptor ability have been defined over the years.[93−96] With most of these empirical measures, data are not available for all the compounds studied here. However, many of the scales nonetheless yield fairly good correlations with the calculated complexation energies, at least for the cases where measurements are available. The scale developed by Abraham et al.,[96] for which values are reproduced in Table 13, yields one of the more linear relationships with the calculated binding energies, as shown in Figure 8. Apparently, the relative strength of hydrogen bond donors does not change much on going from the gas phase to nonpolar solution.[97]

Association Energy of the GC and AT Base Pairs. Calculations of the Watson−Crick AT and GC base pairs at the Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level and the corresponding MP2 level were not feasible on the available computer hardware. However, simple Becke3LYP/6-31+G(d(X+),p) optimization was possible, and

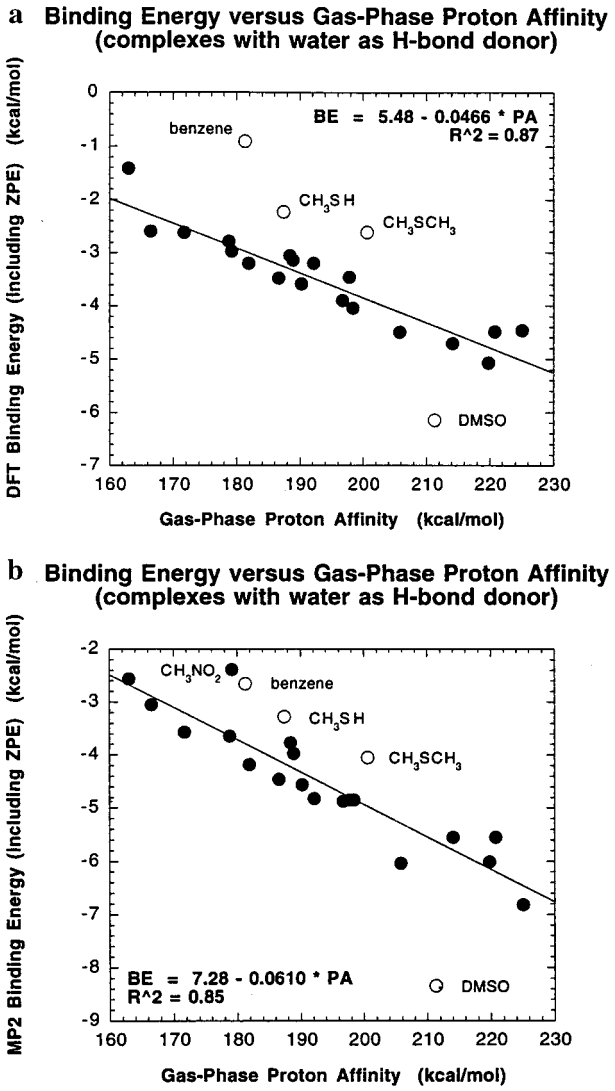


Figure 7. Relationship between binding energies and the experimental gas-phase proton affinities. The points represent the species in Table 13, for which experimental data are available in the literature. For each compound, only the single most strongly bound complex in which water is clearly the hydrogen bond donor has been plotted. (a) Binding energies calculated at the Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level, with the zero-point vibrational energy corrections included. Best fit line: BE = 5.48 − 0.0466 × PA; correlation coefficient = 0.87 (if sulfur bases are included, correlation coefficient = 0.75). (b) Binding energies calculated at the MP2/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) level, with the zero-point vibrational energy corrections included. Best fit line: BE = 7.28 − 0.0610 × PA; correlation coefficient = 0.85. For both (a) and (b), the open circles designate points that were excluded from the linear regression analyses. (Note: BE = binding energy; PA = proton affinity; all quantities in kcal/mol.)

the results are reported in Table 14. Furthermore, we have taken advantage of the extremely close correlation between energies calculated at the Becke3LYP level using the smaller (6-31+G(d(X+),p)) and larger (6-31++G(2d(X+),p)) basis sets to obtain "estimated" association energies at the larger basis set. The estimated energies are also listed in Table 14. The values reported here agree very closely with those obtained by Gould and Kollman at the MP2/6-31G* level[98] and also with those recently obtained by Goddard et al.[99] Santamaria and Vazquez have also computed the stability of the AT and GC base pairs using DFT, with similar results.[100] The final values for the energy of association (including ZPE; i.e., enthalpy at 0 K)

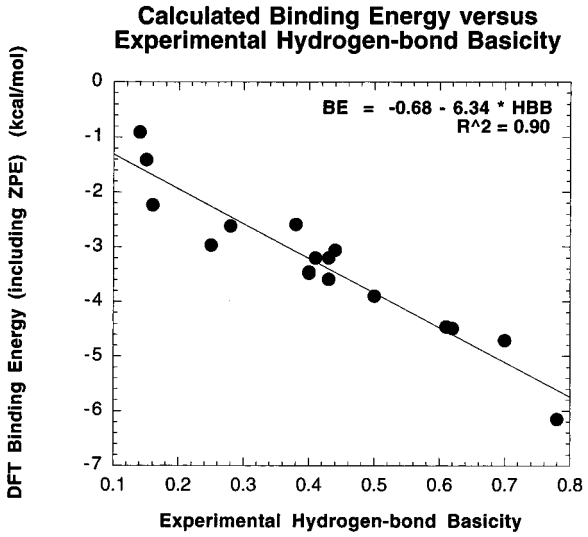


Figure 8. Relationship between binding energies and Abraham's experimental solution hydrogen bond basicity, β_2^H. The points represent the species in Table 13, for which experimental data are available in the literature. For each compound, only the single most strongly bound complex in which water is clearly the hydrogen bond donor has been plotted. Best fit line: BE = −0.68 − 6.34 × HBB; correlation coefficient = 0.90. (Note: BE = binding energy; HBB = hydrogen bond basicity; quantities in kcal/mol.)

TABLE 14: Becke3LYP Calculated Binding Energies of the GC and AT Base Pairs (kcal/mol; without ZPE Correction)

complex	pt. grp.	6-31+G(d(X+),p)	6-31++G(2d(X+),p) (estimated)
AT	C_s	−12.47	−11.99
GC	C_s	−25.41	−24.44

derived by the present method are −12.0 and −24.4 kcal/mol for AT and GC, respectively, compared to the Gould and Kollman values of −12.2 and −25.4 kcal/mol.[101]

Summary

Methodology based on Becke3LYP/6-31++G(2d(X+),p)//Becke3LYP/6-31+G(d(X+),p) ab initio calculations seems to provide an economical and reliable means of determining the strength of hydrogen-bonding interactions. It is capable of reproducing the experimental dipole moment of water in the gas phase with very high accuracy and yields results in agreement with the best calculations available for the association energies of a series of very small hydrogen-bonded complexes. Using this procedure, the energies and structures of 53 hydrogen-bonded complexes of water with various small organic molecules, including alcohols, thiols, ethers, thioethers, carboxylic acids, esters, amines, amides, nitriles, and nitro compounds, were examined systematically. MP2 association energies at the DFT geometries were also computed for comparison purposes and showed close agreement with the DFT values except in the cases of benzene or nitromethane as the hydrogen bond acceptor.

The hydrogen bond geometries were generally linear, and acceptor sites corresponded closely to the positions of lone pairs as predicted by simple hybridization arguments. Structures with sulfur and chlorine atoms showed some deviation from simple expectations and seemed to be largely determined by molecular dipole−dipole interactions.

Categorization of the type of hydrogen bond involved in the various complexes facilitated an ordering of hydrogen bond donor and acceptor abilities for some common functional groups. The best acceptors were the oxygen of DMSO, followed by

nitrogen atoms and the carbonyl oxygen atoms of amides. Other carbonyls were somewhat weaker acceptors and were comparable to nitro and cyano groups, while the −O− oxygen atoms of esters and sulfur atoms were unusually weak. The strongest donors were carboxylic acids. Alcohols, amines, and amides were all somewhat weaker donors than the acids but comparable to each other, while thiols were much weaker.

Strength of association was found to correlate moderately well with experimental gas-phase proton affinity in those cases where water acted unambiguously as the hydrogen bond donor at a single site. Interestingly, sulfur was found to be close to oxygen in hydrogen bond acceptor strength, and the surprisingly strong acceptor ability of sulfur could not be explained in terms of its enhanced polarizability relative to oxygen.

Acknowledgment. The authors thank George Petersson and Joseph Ochterski of Wesleyan University for helpful advice regarding the basis sets necessary for reproducing the association energies of weakly bound complexes accurately. Financial support for this work was provided by a Faculty Start-up Grant for Undergraduate Institutions from the Camille and Henry Dreyfus Foundation, by a Cottrell College Science Award of Research Corporation, and by the National Institutes of Health. Acknowledgment is also made to the donors of the Petroleum Research Fund, administered by the ACS, for partial support of this research.

Supporting Information Available: Tables of energies in hartrees for calculations listed in Tables 1−7, 9, and 15; table of zero-point vibrational energies for complexes; figure showing structures for all optimized complexes (two views for each complex); and full geometric data for all ab initio optimized structures, in Z-matrix form (103 pages). See any current masthead page for ordering and Internet access instructions.

References and Notes

(1) Pimentel, G. C.; McClellan, A. L. *The Hydrogen Bond*; W. H. Freeman: San Francisco, 1960.
(2) Joesten, M. D.; Schaad, L. J. *Hydrogen Bonding*; Marcel Dekker: New York, 1974.
(3) Vinogradov, S. N.; Linnell, R. H. *Hydrogen Bonding*; Van Nostrand Reinhold: New York, 1971.
(4) Legon, A. C.; Millen, D. J. *Chem. Soc. Rev.* **1992**, *21*, 71−78. Legon, A. C. *Chem. Soc. Rev.* **1990**, *19*, 197−237.
(5) Scheiner, S. *Acc. Chem. Res.* **1994**, *27*, 402−408.
(6) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120−126.
(7) Gordon, M. S.; Jensen, J. H. *Acc. Chem. Res.* **1996**, *29*, 536−543.
(8) Hibbert, F.; Emsley, J. *Adv. Phys. Org. Chem.* **1990**, *26*, 255−379.
(9) Li, J. C.; Ross, D. K. *Nature* **1993**, *365*, 327−329.
(10) Franks, F. *Water*; The Royal Society of Chemistry: London, 1983. Frank, H. S.; Wen, W.-Y. *Discuss. Faraday Soc.* **1957**, *24*, 133−140.
(11) Kirby, A. J. *Acc. Chem. Res.* **1997**, *30*, 290−296.
(12) Desirajau, G. R. *Crystal Engineering: The Design of Organic Solids*; Elsevier: Amsterdam, 1989.
(13) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, 1991.
(14) Gerlt, J. A.; Kreevoy, M. M.; Cleland, W. W.; Frey, P. A. *Chem. Biol.* **1997**, *4*, 259−267 and references therein.
(15) Gao, J.; Xia, X.; George, T. F. *J. Phys. Chem.* **1993**, *97*, 9241−9247.
(16) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(17) Del Bene, J. E. *Int. J. Quantum Chem.: Quantum Chem. Symp.* **1992**, *26*, 527−541.
(18) Dill, J. D.; Allen, L. C.; Topp, W. C.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 7220−7226.
(19) Frisch, M. J.; Del Bene, J. E.; Binkley, J. S.; Schaefer, H. F. III *J. Chem. Phys.* **1986**, *84*, 2279−2289.
(20) Jorgensen, W. L. *CHEMTRACTS: Org. Chem.* **1991**, 91−119.
(21) Chakravorty, S. J.; Davidson, E. R. *J. Phys. Chem.* **1993**, *97*, 6373−6383.
(22) Dixon, D. A.; Dobbs, K. D.; Valentini, J. J. *J. Phys. Chem.* **1994**, *98*, 13435−13439.
(23) Pudzianowski, A. T. *J. Chem. Phys.* **1995**, *102*, 8029−8039.
(24) Dimitrova, Y.; Peyerimhoff, S. D. *J. Phys. Chem.* **1993**, *97*, 12731−12736.
(25) Ha, T.-K.; Makarewicz, J.; Bauder, A. *J. Phys. Chem.* **1993**, *97*, 11415−11419.
(26) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273−7287.
(27) Breneman, C. M.; Rhem, M.; Thompson, T. R.; Dung, M. H. *ACS Symp. Ser.* **1994**, *569*, 152−174.
(28) Florián, J.; Johnson, B. G. *J. Phys. Chem.* **1995**, *99*, 5899−5908. Doig, A. J.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 338−343.
(29) Racine, S. C.; Davidson, E. R. *J. Phys. Chem.* **1993**, *97*, 6367−6372.
(30) Rovira, C.; Constans, P.; Whangbo, M. H.; Novoa, J. J. *Int. J. Quantum Chem.* **1994**, *52*, 177−189.
(31) Feyereisen, W. W.; Feller, D.; Dixon, D. A. *J. Phys. Chem.* **1996**, *100*, 2993−2997.
(32) Adalsteinsson, H.; Maulitz, A. H.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 7689−7693.
(33) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.
(34) Hobza, P.; Sauer, J. *Theor. Chim. Acta* **1984**, *65*, 279.
(35) Frisch, M. J.; Pople, J. A.; Del Bene, J. E. *J. Phys. Chem.* **1985**, *89*, 3664−3669.
(36) Novoa, J. J.; Whangbo, M.-H. *J. Am. Chem. Soc.* **1991**, *113*, 9017−9026.
(37) Tse, Y.-C.; Newton, M. D.; Allen, L. C. *Chem. Phys. Lett.* **1980**, *75*, 350.
(38) Herschlag *Science* **1996**, *272*, 97−101.
(39) Nesbitt, D. J. *Chem. Rev.* **1988**, *88*, 843−870.
(40) Tubergen, M. J.; Kuczkowski, R. L. *J. Am. Chem. Soc.* **1993**, *115*, 9263−9266.
(41) Maes, G.; Zeegers-Huyskens, T. *J. Mol. Struct.* **1983**, *100*, 305−315.
(42) Curtiss, L. A.; Blander, M. *Chem. Rev.* **1988**, *88*, 827−841.
(43) Lee, C.; Chen, H.; Fitzgerald, G. *J. Chem. Phys.* **1994**, *101*, 4472−4473.
(44) Sim, F.; St-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391−4400.
(45) Latajka, Z.; Bouteiller, Y. *J. Chem. Phys.* **1994**, *101*, 9793−9799.
(46) Pudzianowski, A. T. *J. Phys. Chem.* **1996**, *100*, 4781−4789.
(47) Gonzalez, L.; Mo, O.; Yanez, M. *J. Comput. Chem.* **1997**, *18*, 1124−1135.
(48) Novoa, J. J.; Sosa, C. *J. Phys. Chem.* **1995**, *99*, 15837−15845.
(49) Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089−10094.
(50) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Phys. Chem.* **1995**, *99*, 10705−10707.
(51) Mijoule, C.; Latajka, Z.; Boris, D. *Chem. Phys. Lett.* **1993**, *208*, 364−368.
(52) Barone, V.; Orlandini, L.; Adamo, C. *Chem. Phys. Lett.* **1994**, *231*, 295−300.
(53) Florián, J.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 3681−3687. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627. Wang, J.; Eriksson, L. A.; Boyd, R. J.; Shi, Z.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 1844−1850.
(54) (a) Wesolowski, T. A.; Parisel, O.; Ellinger, Y.; Weber, J. *J. Phys. Chem. A* **1997**, *101*, 7818−7825. (b) Meijer, E. J.; Sprik, M. *J. Chem. Phys.* **1996**, *105*, 8684−8689. (c) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315−1325. (d) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175−180.
(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94* (Revision C.2); Gaussian, Inc.: Pittsburgh, PA, 1995.
(56) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(57) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(58) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.
(59) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598−2619.
(60) Nyden, M. R.; Petersson, G. A. *J. Chem. Phys.* **1981**, *75*, 1843−1862. Petersson, G. A.; Al-Laham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081−6090. Petersson, G. A.; Tensfeldt, T.; Montgomery, J. A. *J. Chem. Phys.* **1991** *94*, 6091−6101. Montgomery, J. A.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900−5909.
(61) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
(62) Ochterski, J. W. Ph.D. Thesis, Wesleyan University, 1994.
(63) Hadad, C. M.; Rablen, P. R. CASGEN package; Yale University: New Haven, CT, 1992.

(64) van Duijineveldt de Rijdt, J. G. C. M.; van Duijineveldt, F. B. *J. Chem. Phys.* **1992**, *97*, 5019−5030. Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104−6113.

(65) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(66) Curtiss, L. A.; Frurip, D. J.; Blander, M. *J. Chem. Phys.* **1979**, *71*, 2703−2711.

(67) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221−7230.

(68) Ochterski, J. W.; Petersson, G. A.; Wiberg, K. B. *J. Am. Chem. Soc.* **1995**, *117*, 11299−11308.

(69) Rablen, P. R.; Hartwig, J. F. *J. Am. Chem. Soc.* **1996**, *118*, 4648−4653.

(70) Florián, J.; Johnson, B. G. *J. Phys. Chem.* **1994**, *98*, 3681−3687.

(71) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(72) Dobbs, K. W.; Dixon, D. A. *J. Phys. Chem.* **1994**, *98*, 4498−4501.

(73) Murray, C. W.; Laming, G. J.; Handy, N. C.; Amos, R. D. *J. Phys. Chem.* **1993**, *97*, 1868−1871.

(74) Dobado, J. A.; Molina, J. M. *J. Phys. Chem.* **1993**, *97*, 7499−7504. Tao, F.-M.; Pan, Y.-K. *J. Phys. Chem.* **1992**, *96*, 5815−5816. Yang, J.; Kestner, N. R. *J. Phys. Chem.* **1991**, *95*, 9221−9230. Tao, F.-M.; Pan, Y.-K. *J. Phys. Chem.* **1991**, *95*, 3582−3588.

(75) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418−2426. Liu, B.; McLean, A. D. *J. Chem. Phys.* **1989**, *91*, 2348−2359. Novoa, J. J.; Planas, M.; Whangbo, M. H. *Chem. Phys. Lett.* **1994**, *225*, 240−246.

(76) Gotch, A. J.; Zwier, T. S. *J. Chem. Phys.* **1992**, *96*, 3388−3401. Cheney, B. V.; Schulz, M. W. *J. Phys. Chem.* **1990**, *94*, 6268−6273. Karlström, G.; Linse, P.; Wallqvist, A.; Jönsson, B. *J. Am. Chem. Soc.* **1983**, *105*, 3777−3782.

(77) Benzene is actually an even weaker acceptor, with a binding energy of 1.69 kcal/mol, but has not been included in the categorization scheme, since the interaction site on benzene cannot be associated with any one particular atom.

(78) Coppens, P. In *Electron Distributions in the Chemical Bond*; Hall, M. B., Ed.; Plenum Press: New York, 1982; pp 61−92.

(79) Dunitz, J. D. *X-ray Analysis and the Structure of Organic Molecules*; Cornell University Press: Ithaca, NY, 1979.

(80) Walters, V. A.; Hadad, C. M.; Thiel, Y.; Colson, S. D.; Wiberg, K. B.; Johnson, P. M.; Foresman, J. B. *J. Am. Chem. Soc.* **1991**, *113*, 4782−4791. Wiberg, K. B.; Hadad, C. M.; Foresman, J. B.; Chupka, W. A. *J. Phys. Chem.* **1992**, *96*, 10756−10768. Hadad, C. M.; Foresman, J. B.; Wiberg, K. B. *J. Phys. Chem.* **1993**, *97*, 4293−4312.

(81) Gao, J.; Xia, X. *Science* **1992**, *258*, 631−635.

(82) Jorgensen, W. L.; Allen, L. C. *J. Am. Chem. Soc.* **1971**, *93*, 567−574. Jorgensen, W. L.; Allen, L. C. *Chem. Phys. Lett.* **1970**, *7*, 483−485.

(83) Wiberg, K. B.; Rablen, P. R. *J. Am. Chem. Soc.* **1995**, *117*, 2201−2209. Wiberg, K. B.; Rablen, P. R. *J. Am. Chem. Soc.* **1993**, *115*, 9234−9242.

(84) Ransil, B. J.; Sinai, J. J. *J. Chem. Phys.* **1967**, *46*, 4050−4074.

(85) Kollman, P. A.; Liebman, J. F.; Allen, L. C. *J. Am. Chem. Soc.* **1970**, *92*, 1142−1150.

(86) Bader, R. F. W.; Henneker, W. H.; Cade, P. E. *J. Chem. Phys.* **1967**, *46*, 3341−3363. Bader, R. F. W.; Keavenly, I.; Cade, P. E. *J. Chem. Phys.* **1967**, *47*, 3381−3402.

(87) Subtractions were performed while holding the geometries of the separate fragments constant.

(88) To be more precise, the positive and negative regions of the difference density distributions were integrated independently, since the *total* difference density distribution necessarily integrates to zero.

(89) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695−808.

(90) Yamabe, S.; Hirao, K.; Wasada, H. *J. Phys. Chem.* **1992**, *96*, 10261−10264.

(91) Larson, J. W.; McMahon, T. B. *J. Am. Chem. Soc.* **1982**, *104*, 5848−5849.

(92) Herail, M.; Megnassan, E.; Proutiere, A. *J. Phys. Org. Chem.* **1997**, *10*, 167−174.

(93) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; VCH Publishers: New York, 1990; Chapter 7, and references therein.

(94) Chardin, A.; Laurence, C.; Berthelot, M.; Morris, D. G. *J. Chem. Soc., Perkin Trans. 2* **1996** 1047−1051 and references therein. Laurence, C.; Berthelot, M.; Lequestel, J. Y.; Elghomari, M. J. *J. Chem. Soc., Perkin Trans. 2* **1995** 2075−2079 and references therein.

(95) Li, J. J.; Zhang, Y. K.; Hsiu, O. Y.; Carr, P. W. *J. Am. Chem. Soc.* **1992**, *114*, 9813−9828.

(96) Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 521−529.

(97) The Abraham scale is based on hydrogen bond complexation equilibria in tetrachloromethane solution.

(98) Gould, I. R.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 2493−2499.

(99) Brameld, K.; Dasgupta, S.; Goddard, W. A., III *J. Phys. Chem.* **1997**, *101*, 4851−4859.

(100) Santamaria, R.; Vázquez, A. *J. Comput. Chem.* **1994**, *15*, 981−996.

(101) These numbers are very slightly different from those literally reported in the Gould and Kollman paper. The ZPE corrections evident in Table 13 were applied to the ΔE values reported in Table 16 to obtain enthalpies at 0 K for the MP2/6-31G*//MP2/6-31G* level of theory.