Transferable ab Initio Intermolecular Potentials. 1. Derivation from Methanol Dimer and Trimer Calculations

Wijnand T. M. Mooij,*,† Frans B. van Duijneveldt,‡
Jeanne G. C. M. van Duijneveldt-van de Rijdt,‡ and Bouke P. van Eijck†

Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands, and Theoretical Chemistry Group, Debye Institute, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands

Received: May 19, 1999; In Final Form: September 7, 1999

An intermolecular potential was derived from ab initio calculations on methanol dimers and trimers. Ninety-four methanol dimer geometries and seventeen trimer geometries have been studied using an interaction-optimized basis set. Because we aimed at transferability, all terms of the model were fitted separately, in order to obtain a potential in which these energy terms have a well-defined physical meaning. Electrostatic interactions were described by atomic multipole moments obtained by fitting to the monomer electrostatic potential. The polarization energy was modeled using atomic dipole polarizabilities. Scaled empirical polarizabilities reproduced the energy nonadditivity in methanol trimers very accurately. The dispersion energy was described by a damped r^{-6} atom−atom potential, which was fitted to separately calculated dispersion energies. Exchange energy and remaining short-ranged terms were modeled by an exponential repulsion term, including some anisotropic features. This repulsion model was fitted to the total SCF+MP2 interaction energies of the dimers. The maximum deviation for near-equilibrium geometries was ∼0.2 kcal/mol. The calculated interaction energies for dimers containing methane, water, and dimethyl ether were in similar agreement, both with our own ab initio calculations and with the best values available in the literature. In addition, trends in hydrogen-bond distances in these dimers and in methanol trimers were well reproduced by our model. So, the potential was seen to be transferable to related systems.

1. Introduction

Computer modeling of organic solids is most commonly performed using empirical atom−atom potential energy functions.[1] This approach is often rather successful in reproducing geometrical features of experimental crystal structures. Considering the approximate nature of these potentials, this may seem remarkable. However, one should keep in mind that most models for nonbonded interactions are generally parameterized on this type of structural information. Data on energetics used in the parameterization is usually limited to a number of sublimation energies, whose comparison to calculated lattice energies is not without uncertainties.[1] Therefore the accuracy of the calculated lattice energies is much more doubtful.

In crystal structure prediction one demands a very high accuracy on just these lattice energies, as it is usually assumed that the experimentally observed structure corresponds to the one with the lowest calculated energy.[2] Because differences in energy between hypothetical structures can be extremely small, very accurate intermolecular potentials are needed for a reliable energy ranking. In this approach thermodynamic and kinetic effects are neglected. This is questionable,[3,4] and accurate energy values will be a prerequisite to assess the importance of these neglected effects.

For the electrostatic interactions, the use of atomic multipole moments has proven to be a large improvement over atom-centered partial charges.[5−7] In those studies, ab initio derived multipole moments were added to an empirical 6-exp potential. In the present work we go beyond that, and a complete intermolecular potential is determined from ab initio calculations. Our goal is to develop a general potential that can be used for calculating the interactions between molecules such as alkanes, ethers, alcohols, and even carbohydrates. Because it is computationally expensive to calculate accurate interaction energies, we chose the smallest possible model system, i.e., the methanol dimer. We require our potential to be transferable. To this end, we construct an atom−atom potential that separately reproduces different components of the interaction energy (e.g., electrostatic, repulsion, dispersion), instead of total interaction energies alone. At the accuracy we aim at, atomic multipole moments are not transferable from one molecule to the other (see, e.g., ref 8). Therefore, for each molecule of interest a new set of atomic multipole moments has to be calculated.

Other approaches to derive ab initio intermolecular potentials have also pursued a separation of the interaction energy in physically distinct contributions. The anisotropic site potential for water of Millot and Stone[9,10] is parameterized on data from intermolecular perturbation theory. It is complicated, involving quadrupole polarizabilities and a very elaborate anisotropic repulsion model. In the latest revision[10] an explicit charge-transfer term was added. The model involves, however, a single-site dispersion term based on experimental data, which prohibits transferability to other molecules. Moreover, the complexity of such a potential hinders its application to large numbers of hypothetical crystal structures, while it does not necessarily increase its accuracy.

---

* Author to whom correspondence should be addressed.

† Department of Crystal and Structural Chemistry. E-mail: w.t.m.mooij@chem.uu.nl, b.p.vaneijck@chem.uu.nl.

‡ Theoretical Chemistry Group. E-mail: f.b.vanduijneveldt@chem.uu.nl, j.g.c.m.vanderijdt@chem.uu.nl.

10.1021/jp991641n CCC: $18.00

Other approaches, such as NEMO[11,12] and the systematic potential by Wheatley,[13,14] are mainly based on monomer properties and on the assumption that the repulsion energy is proportional to the overlap of the monomer wavefunctions or charge distributions. Constructing a potential for a new molecule of interest is then computationally less expensive than in a true ab initio approach. The price one pays, however, is that the accuracy of the repulsion model is uncertain. Since at near-equilibrium geometries the repulsion term is large in comparison with the total interaction energy, small relative errors in this term may spoil the accuracy of the final interaction energies. We therefore prefer to use as accurate a repulsion energy as is attainable. We will show that the resulting model is transferable; therefore, we need not derive a new potential for another molecule of interest at all.

We developed our model on the most accurate ab initio data we could afford, performing counterpoise-corrected supermolecular calculations at the SCF+MP2 level on methanol dimers in a newly designed interaction-optimized basis set. We do not share the pessimistic view on the accuracy of ab initio interaction energies.[9,10,13] Admittedly, it is hard to obtain by ab initio methods interaction energies accurate to within a few tenths of a kcal/mol, but in the end we shall be interested in energy differences more than in absolute energies. Due to compensation of errors, the former can usually be estimated more accurately. In order to enhance this compensation of errors, our ab initio data must be of equal accuracy for all geometries of interest.

Deriving a potential from dimer calculations means deriving a gas-phase potential. In order to obtain transferability of our model to the condensed phase, it is essential to allow for nonadditivity in the model. This is accomplished through the use of atomic dipole polarizabilities which we parameterize on methanol trimer calculations.

In this article we first report the methods, basis set, and geometries used in the ab initio calculations. Secondly, we describe the parameterization of the electrostatic, polarization, dispersion, and repulsion parts of our potential together with the ab initio data that was used for fitting each energy term. Then we assess the accuracy of the model and test the assumed transferability. To this end we compare results of our model with ab initio data for dimers involving water, methane and dimethylether. In a companion article[15] we show that our model is suitable for modeling of crystals. There we also report the application of the potential to the prediction of the crystal structures of methanol, ethanol, dioxane, and propane.

2. Ab Initio Calculations

2.1. Method. Interaction energies were calculated in the supermolecule approach, using the counterpoise procedure[16] to avoid the basis set superposition error (BSSE). The interaction energy at a certain geometry (α) is then obtained as

$$\Delta E_{\mathrm{AB}}(\alpha) = E_{\mathrm{AB}}(\alpha) - E_{\mathrm{A,DCBS}}(\alpha) - E_{\mathrm{B,DCBS}}(\alpha)$$

for a dimer AB, and as

$$\Delta E_{\mathrm{ABC}}(\alpha) = E_{\mathrm{ABC}}(\alpha) - E_{\mathrm{A,TCBS}}(\alpha) - E_{\mathrm{B,TCBS}}(\alpha) - E_{\mathrm{C,TCBS}}(\alpha)$$

for a trimer ABC. $E_{\mathrm{A,DCBS}}$ and $E_{\mathrm{A,TCBS}}$ are the energies of monomer A, calculated in the complete basis set of the dimer or the trimer, respectively (DCBS = dimer-centered basis set; TCBS = trimer-centered basis set).

TABLE 1: Interaction-Optimized Basis Set for Methanol (IOM)

	C	O	H_O	H_C
	EZ[24,25]	EZ[24,25]	DZ[26]	DZ[26]
α_P			0.195/0.78[27]	0.15/0.6
α_d	0.15/0.6	0.1667/0.5/1.5[27]		
α_f	0.3	0.3125/1.25		
α_g	0.3	0.75		

Interaction energies were calculated at the SCF+MP2 level. For neutral hydrogen-bonded dimers higher-order correlation effects tend to cancel, making MP2 the method of choice for the calculation of a potential energy surface.[17] Moreover, relative energies are even less influenced than absolute energies. For example, MP2 relative energies for a number of water trimer configurations differ by less then 0.06 kcal/mol from their CCSD(T) counterparts.[18]

The total MP2 interaction energy can be partitioned in the following contributions:

$$\Delta E_{\mathrm{MP2}}^{\mathrm{tot}} = \Delta E_{\mathrm{SCF}}^{(1)} + \Delta E_{\mathrm{SCF}}^{(2)} + \Delta E_{\mathrm{MP2}}^{\mathrm{corr}}$$

$\Delta E_{\mathrm{SCF}}^{(1)}$ is the first-order SCF interaction energy, which contains the coulombic and exchange−repulsion interactions between the unperturbed monomers. $\Delta E_{\mathrm{SCF}}^{(2)}$ is the energy subsequently gained in the SCF process, which contains polarization and charge-transfer effects. The MP2 correlation energy ($\Delta E_{\mathrm{MP2}}^{\mathrm{corr}}$) contains the dispersion energy, in addition to corrections on $\Delta E_{\mathrm{SCF}}^{(1)}$ and $\Delta E_{\mathrm{SCF}}^{(2)}$. The most important correction originates from the reduction of the dipole moment of the monomer by correlation. For hydrogen-bonded geometries, this results in an electrostatic interaction energy that is smaller than at the SCF level and consequently in a positive contribution to $\Delta E_{\mathrm{MP2}}^{\mathrm{corr}}$.

Dimer and trimer calculations were calculated using the ATMOL[19] system of programs and its local extensions SERVEC[20] and INTACAT,[21] using spherical harmonic d, f, and g functions.

2.2. Basis Set and Accuracy. The calculation of interaction energies by ab initio methods always suffers from the basis set incompleteness error, and large basis sets are needed to obtain accurate results. We chose to use an interaction-optimized basis set, because such a set can be much smaller than a “standard” set at the same level of accuracy.[22] For example, a basis set of only 249 functions for water dimer can achieve equivalent results to a basis set of 1046 functions.[23] This IO249 set employs bond functions, which can be very effective in saturating the dispersion energy.[22] However, in many geometries that we wanted to consider it is far from obvious where to place such bond functions. The final goal of this study is to obtain the relative energies for different crystal packings as accurately as possible. Relative energies are not hindered by some underestimation of the total dispersion energy, as long as the underestimation is equally large for all packings. This would be endangered if bond functions would lower the dispersion energy in certain geometries, but not in others. Therefore, we decided to refrain from the use of bond functions.

In the appendix we report the optimization of a basis set on interaction energies of a methanol dimer. The resulting interaction-optimized set (referred to as IOM) is given in Table 1. In order to illustrate the accuracy achieved with this basis set, we calculated the MP2(frozen core) interaction energy for water dimer (Table 2). This value is ∼0.2 kcal/mol above the MP2 limit for frozen monomer geometries.[23,29,30] This places the IOM basis at the accuracy of the cc-pVQZ and the aug-cc-pVTZ basis sets.[31,32] The latter are, however, substantially larger than our

TABLE 2: Interaction Energy (kcal/mol) for Water Dimer in the Feller−Frisch Geometry[28] Using Different Basis Sets

basis	N[a]	SCF	MP2	total
cc-pVTZ	116	−3.55	−0.85	−4.40[b]
IOM (this work)	136	−3.58	−1.10	−4.68
aug-cc-pVTZ	184	−3.54	−1.10	−4.64[b]
cc-pVQZ	230	−3.57	−1.10	−4.67[b]
limit (IO249)	249	−3.58	−1.29	−4.87[c]

[a] *N* denotes the total number of contracted functions in the basis. [b] Results taken from ref 28. [c] Result taken from ref 23.

basis set, so the use of an interaction-optimized basis set proved to be advantageous. Higher-order correlation[29] and core correlation[23,30] result in a limiting value for the binding energy of water dimer of ∼5.0 kcal/mol, so our calculations underestimate this by ∼0.3 kcal/mol. This absolute error is mainly caused by an underestimation of the dispersion energy.

What can one infer from these results on water dimer regarding the absolute accuracy of the interaction energies for methanol dimer? One could argue that $\Delta E^{\mathrm{corr}}_{\mathrm{MP2}}$ is larger for methanol dimer than for water dimer (i.e., by a factor of 1.8 in near-equilibrium geometries), and therefore that its error will be equally larger. One could also argue that the underestimation of the dispersion energy is not caused by the leading term in the dispersion-energy series (r^{-6}), since the quality of this term is dependent upon the dipole polarizabilities of the monomers. The deficiency will be mostly in the higher powers of r^{-1}, so it will be most serious at short range. This is supported by the fact that the use of bond functions proves to be very efficient to saturate the dispersion energy. Since in their optimal geometries, methanol dimer contains the same short-range interactions as water dimer, the error in the interaction energy might be not much larger. We suppose that the underestimation in the interaction energy is somewhere in between both lines of reasoning, around 0.4 kcal/mol. As argued above, the relative energies of different geometries and different crystal packings will be more accurate due to compensation of errors.

2.3. Geometries. Ninety-four different methanol dimer geometries were constructed, which are summarized in Figure 1. Emphasis was put on geometries with close contacts, because at long range the interaction energies are dominated by electrostatics, which can be parameterized from monomer properties. Care was taken to include close contacts between all different atom types, in order to provide enough information on their repulsion energy for the fitting of parameters. Dimers A−F probe different O−H···O hydrogen bond geometries. Dimers G−J contain C−H···O contacts, K−N sample CH_3···

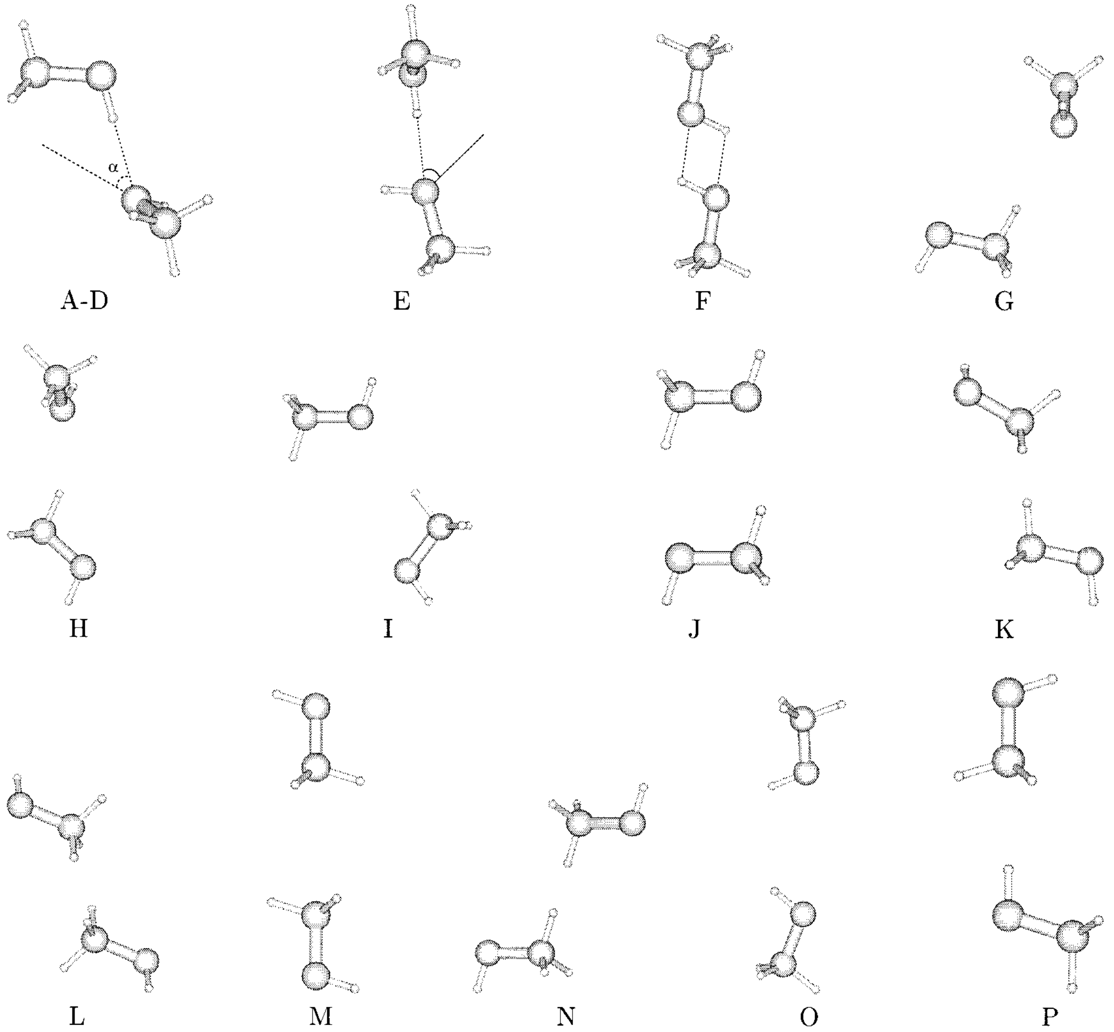

Figure 1. Overview of the dimer geometries. A: R_{OO} 2.5−5.0 Å, acceptor angle 50°. The acceptor angle (α) is defined as the angle between the bisector of the COH angle of the hydrogen-bond acceptor and the O−O line. The donor molecule lies in the bisector plane. B and C: acceptor angle varied (−75 to 100°), R_{OO} 2.85 Å and 3.30 Å. D: the O−H···O angle varied, R_{OO} 2.70 Å. E: angle in COH plane varied, R_{OO} 2.85 Å. O−H of donor molecule lies in the plane of the acceptor molecule. F: C_i, C_2, and C_{2h} structures in analogy to water dimer.[33] G: acceptor angle varied (−90 to 90°), R_{CO} = 3.5 Å. H: C−H···O angle varied at R_{CO} = 3.25 Å. I−P: R_{CC} and R_{CO} varied in the range 3.0−4.5 Å.

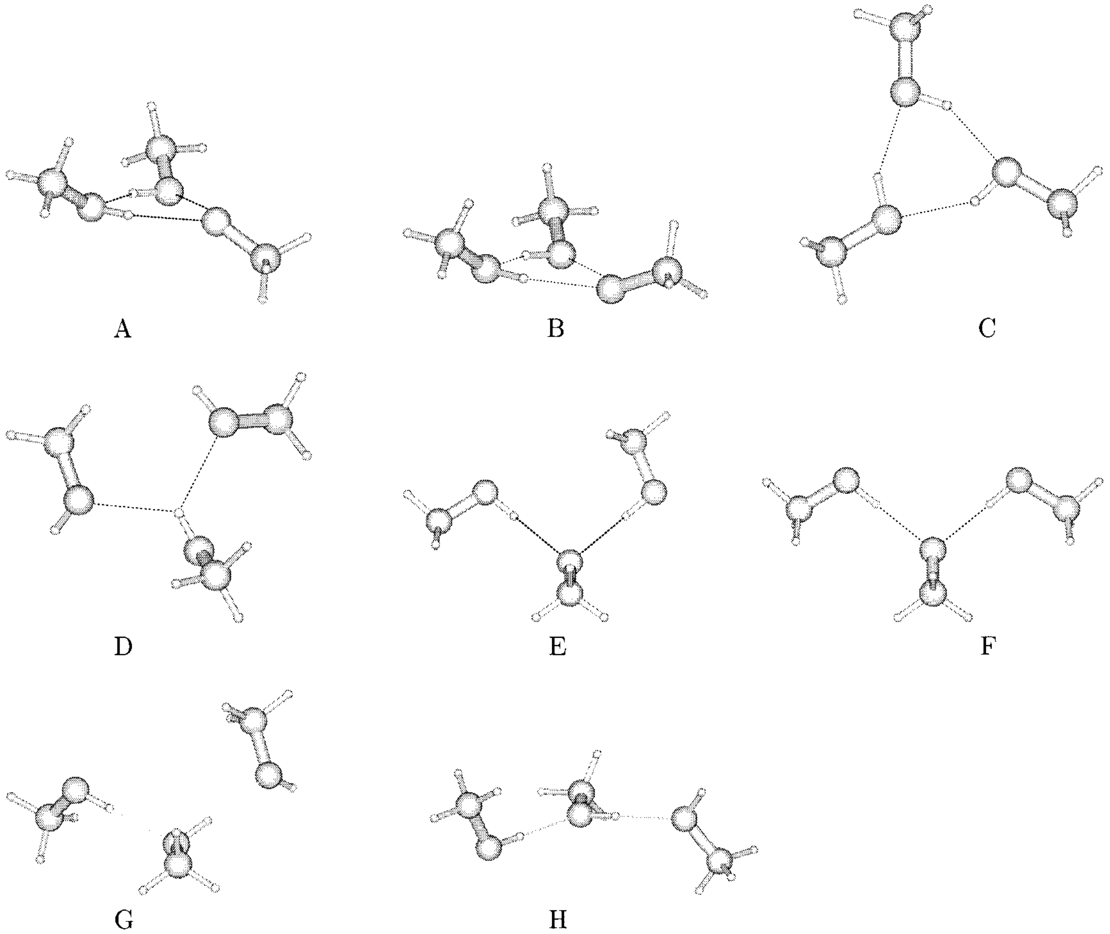

Figure 2. Overview of the trimer geometries. A−C: R_{OH} 1.8−2.2 Å. D: R_{OH} 2.0−2.5 Å. E,F: R_{OO} 2.85 Å. G,H: R_{OO} 2.5−3.0 Å, both acceptor angles 0−50°.

CH_3 orientations, and O and P include O···O and CH_3···H−O contacts. Seventeen trimer geometries were constructed, which are given in Figure 2. The intramolecular parameters were r_s values from microwave spectroscopy[34] and were kept fixed. All geometries together with the calculated ab initio interaction energies are available as Supporting Information.

3. Constructing the Intermolecular Potential

In order to obtain transferability to similar molecules, we aim at deriving a potential in which the separate energy terms have a well-defined physical meaning. To this end, we need to decompose the total interaction energy in physically distinct contributions. At long range the total interaction energy (E_{inter}) can be described by the electrostatic (E_{elec}), polarization (E_{pol}), and dispersion (E_{disp}) contributions. In our model we treat all of these terms separately. This leaves out the short-range, overlap-dependent terms, such as exchange-repulsion, charge-transfer, and penetration energies. In our model these were all treated together in a single exponential repulsion term (E_{rep}). In the subsequent sections we describe the parameterizations of each of these four terms.

3.1. Electrostatic Energy. The (penetration-free) electrostatic interaction energy $E_{\mathrm{elec}}^{\mathrm{abinitio}}$ can be calculated from a multicentre multipole expansion. Such an expansion can be derived directly from the wavefunction by a distributed multipole analysis (DMA).[35] The resulting multipole expansion gives the exact electrostatic potential, provided it is taken to high order. If the expansion is truncated, convergence becomes uncertain. For the model we wished to limit ourselves to atomic expansion sites, with moments up to quadrupole. Although the DMA procedure is designed for optimal convergence, truncation at this level of a DMA-derived multipole expansion is reported to produce errors of ∼10% in the electrostatic potential (ESP).[36] Indeed, we found that for methanol an atom-centered DMA-derived expansion with moments up to quadrupole results in a root-mean-square deviation (rmsd) of 7.4% in the ESP. The DMA-procedure employs only the wavefunction, and no information on the electrostatic potential is used. Better results are obtained by explicit fitting of atomic multipole moments (AMMs) to the ESP. By definition, this approach ensures the best description of the electrostatic potential at the desired level of truncation of the multipole series. It is commonly used for atomic partial charges,[37,38] but has been used for multipolar models as well.[37] We used a model that consisted of monopoles and dipoles on hydrogens, and monopoles, dipoles, and quadrupoles on carbon and oxygen, which resulted in an excellent fit (0.6%).

The electrostatic contribution to the dimer interaction energy (E_{elec}) is calculated with these ESP-fitted AMMs. Considering the very low rmsd of the fit, these electrostatic interaction energies will not deviate significantly from the $E_{\mathrm{elec}}^{\mathrm{abinitio}}$, which we did not calculate. The derivation of these AMMs is, however, computationally somewhat costly, as they are derived from an MP2/IOM calculation. For methanol this is not a problem, but for future use of this potential on larger molecules[15] we also looked for a simpler approach capable of giving electrostatic interaction energies close to the E_{elec} values. To this end we calculated the electrostatic energy for all dimer geometries using sets of ESP-fitted AMMs derived using different wavefunctions.

We noted a rather small effect of any polarization functions on carbon and hydrogen, a general reduction of the interaction energy upon reducing the basis set, and an overestimation of the interaction energy at the SCF level. In this way we arrived at the SCF/DZ(2d^{O}) level (only d functions on oxygen) as an acceptable method for calculating the AMMs. The rmsd of these electrostatic energies with E_{elec} was only 0.16 kcal/mol. The largest deviation was an underestimation of the electrostatic repulsion in dimer N (by 0.3 kcal/mol at $R_{\mathrm{CC}} = 3.75$ Å, 0.8 kcal/mol at 3.25 Å).

For all ESP fittings monomer wavefunctions were calculated using the program GAUSSIAN94,[39] and the program MOLDEN[40] was used to calculate the electrostatic potential on a series of Connolly surfaces.[41] Eight surfaces were used, at 1.2−2.6 times the van der Waals radii. The (penetration-free) electrostatic potential was calculated from a DMA-derived expansion including bond centres as additional expansion sites and moments up to octopole. A routine was written to fit any combination of atomic monopoles, dipoles or quadrupoles to this well-converged ESP. Fitting was performed by singular value decompositioning with constraints treated by elimination,[42] employing the SVD routine from the SLATEC library. This procedure is now available within MOLDEN.[40] Only the charge neutrality was imposed, and no rank deficiencies were encountered at a singular value tolerance of 0.001.

3.2. Dispersion Energy. Dispersion energies for all dimers were calculated using London's sum-over-state second-order perturbation theory with Møller−Plesset partitioning ($E_{\mathrm{disp}}^{\mathrm{abinitio}}$). The monomers were described in their own basis (MCBS), because such a description of dispersion energy is most closely related to our dispersion model. Other dispersion-energy contributions that are present in $\Delta E_{\mathrm{MP2}}^{\mathrm{corr}}$ (exchange dispersion and charge-transfer correlation) are overlap dependent, therefore cannot be fitted with a r^{-6} term and are better absorbed in the repulsion part of the potential.

These $E_{\mathrm{disp}}^{\mathrm{abinitio}}$ values were used to determine the parameters of the dispersion part of our potential, which was taken to be the sum of interatomic pair interactions modeled by a damped r^{-6} formula:

$$E_{\mathrm{disp}}(r_{ij}) = f(r_{ij})\frac{\sqrt{C_i C_j}}{r_{ij}^6}$$

in which the C's are the dispersion coefficient of two atoms, and r_{ij} is their interatomic distance. A simple damping function like the one that is part of the HFD (Hartree−Fock + damped dispersion) model[43] seemed preferable. However, this function introduces an unwanted discontinuity in the second derivative of the potential. Therefore, a similar function was constructed, with a continuous second derivative:

$$f(r_{ij}) = \left(1 - \exp\left[-\left(\frac{c_{\mathrm{damp}} r_{ij}}{R_i^{\mathrm{vdw}} + R_j^{\mathrm{vdw}}}\right)^3\right]\right)^2$$

With $c_{\mathrm{damp}} \approx 1.93$ this function behaves similarly to the HFD damping.

All geometries were used in the fit with a uniform weight. Parameters were optimized using a modified version of the XTALFIT program within the TINKER modeling package,[44] which incorporates a Levenberg−Marquardt nonlinear least-squares method. Standard van der Waals radii[45] were first used in the damping function. The fit was improved considerably by allowing the van der Waals radius of polar hydrogen to be optimized. The rms error of the fit was 0.08 kcal/mol. The resulting parameters are given in Table 3.

TABLE 3: Dispersion and Polarization Parameters[a]

	C	O	H_C	H_O
C_i	657.13	460.49	19.23	6.07
R_i^{vdw}	1.7	1.52	1.2	0.8
α_i	1.227	0.737	0.448	0.448

[a] Dispersion coefficients (C_i) in kcal mol^{-1} Å^6, van der Waals radii (R_i^{vdw}) in Å, $c_{\mathrm{damp}} = 1.93$, atomic polarizabilities (α_i) in Å^3

They seem to be physically reasonable, following the trend in atomic polarizability (carbon > oxygen > hydrogen). Comparison to r^{-6} parameters in empirical atom−atom potentials is not very useful, because such potentials are mostly fitted as a whole, and no physical meaning should be attributed to the individual terms. More sensible comparison is possible to the dispersion model for $(H_2O)_2$ proposed by Szczęśniak et al.[46] They report $C_{\mathrm{OO}} = 444.8$ kcal/mol, $C_{\mathrm{HH}} = 6.06$, and $C_{\mathrm{OH}} = 74.3$ (all in kcal mol^{-1} Å^6) for a damped r^{-6} formula. The similarity with our values is clear, although detailed comparison is hindered by differences in the damping. Still, it gives confidence that our parameters are also individually reasonable and not only valid in this specific combination.

3.3. Polarization Energy. One could have hoped to be able to use $\Delta E_{\mathrm{SCF}}^{(2)}$ for fitting of the polarization energy term. This term contains overlap-dependent contributions, but at larger separations these effects vanish and $\Delta E_{\mathrm{SCF}}^{(2)}$ equals the classical electrostatic polarization. So, by excluding geometries with close contacts, one could use $\Delta E_{\mathrm{SCF}}^{(2)}$ to fit a polarization model. Obviously, the distinction between long range and short range is not clear, and it turned out that short-range points could only be fitted at the expense of long-range points. As a consequence, the fit was seen to be severely dependent upon the exact choice of geometries.

Therefore we used another approach, and calculated the energy nonadditivity in methanol trimers. Nonadditivity in a trimer is defined as

$$E_{\mathrm{nonad}} = \Delta E_{\mathrm{ABC}} - \Delta E_{\mathrm{AB}} - \Delta E_{\mathrm{BC}} - \Delta E_{\mathrm{AC}}$$

By fitting a polarization model directly on this nonadditivity one explicitly fits the quantity that provides the main reason for including a polarization term in the model. Inherently, the energies of methanol trimers will be better reproduced than when $E_{\mathrm{pol}}^{\mathrm{abinitio}}$ values from perturbational theory on dimers are used. This is thought to increase the transferability to larger clusters and eventually condensed phases.

Interaction energies for the seventeen trimers were calculated at the SCF level, since in a study on water trimer nonadditive contributions at the MP2 level were seen to be negligible.[18,47] For computational reasons, a smaller basis set had to be used, but from the basis set optimizations it was concluded that the addition of f and g functions to the heavy atoms has a small influence on both $\Delta E_{\mathrm{SCF}}^{(1)}$ and $\Delta E_{\mathrm{SCF}}^{(2)}$ (Appendix, Tables 6 and 7). So, trimer calculations were performed in the IOM basis without any f or g functions. All interaction energies were calculated in the TCBS. As all ΔE's are partitioned, the nonadditivity can also be partitioned in a $\Delta E_{\mathrm{SCF}}^{(1)}$ and a $\Delta E_{\mathrm{SCF}}^{(2)}$ part.

Nonadditivity was often significant, varying from −2.8 to +0.7 kcal/mol, and $\Delta E_{\mathrm{SCF}}^{(2)}$ is the major source of this nonadditivity, similar to results on water trimers.[47,48] The nonadditivity in $\Delta E_{\mathrm{SCF}}^{(1)}$ was seen to be negligible for most geometries (<0.05 kcal/mol), apart from some geometries involving short contacts

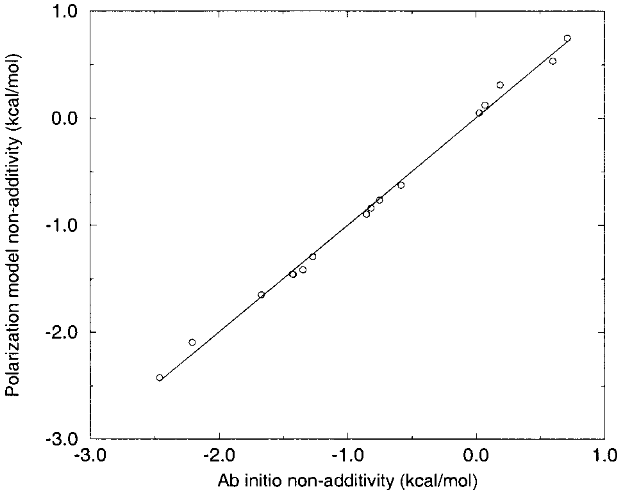

Figure 3. Ab initio nonadditivities in $\Delta E^{(2)}_{\mathrm{SCF}}$ versus model nonadditivities.

between all three oxygens in bifurcated (D) and cyclic geometries (A−C). It was always cooperative, ranging up to −0.4 kcal/mol, and in the bifurcated geometries it outweighs the small anticooperativity in $\Delta E^{(2)}_{\mathrm{SCF}}$.

The dominance of $\Delta E^{(2)}_{\mathrm{SCF}}$ in the nonadditivity supports our choice to include polarization energy as the only nonadditive term, because $\Delta E^{(2)}_{\mathrm{SCF}}$ contains this contribution. The nonadditivities in $\Delta E^{(2)}_{\mathrm{SCF}}$ were used to derive the atomic polarizabilities in a dipole polarization model, which uses only intermolecular polarization. The polarization energy for an atom was calculated using

$$E_{\mathrm{pol}} = -{}^{1}/_{2} \propto E_0^2$$

where α is an isotropic atomic polarizability, and E_0 is the damped electrostatic field at the location of this atom due to the permanent AMMs on all atoms not within the same molecule. A short-range damping of the electrostatic field was applied using the square root of the dispersion damping function, an approach previously used by others.[9,13] To be consistent, we calculated the electrostatic field E_0 with AMMs obtained at the level of theory that was used in the trimer calculations.

We first evaluated the polarization energies in all trimers using atomic polarizabilities from the literature. Two common empirical parameter sets are those due to Miller[49] and Thole.[50] Our model does not involve mutual polarization of induced dipoles, which means that we assume the polarizability of a molecule to consist of additive atomic polarizabilities. Therefore we started from Miller's additive "ahp" values. The correlation between the model and the $\Delta E^{(2)}_{\mathrm{SCF}}$ nonadditivities turned out to be remarkably good (0.998). Because we did not think that the accuracy of both the model and the ab initio calculations allowed for any significant improvement by adjusting the individual polarizabilities, we only determined an optimal overall scaling for the Miller polarizabilities. After scaling Miller's polarizabilities by 1.157 the rms deviation from the $\Delta E^{(2)}_{\mathrm{SCF}}$ nonadditivity was 0.05 kcal/mol, the maximum deviation 0.12 kcal/mol (see Figure 3). The resulting parameters are given in Table 3.

Both the cooperative and the anticooperative effects are well reproduced by this rather simple model. In recent work on the fluctuating charge (FQ) model,[51] problems were reported in describing the anticooperativity present in geometries in which one oxygen accepts two hydrogen bonds (like trimers E and F). In the FQ model polarization is modeled by charge flow between the atoms within a molecule. The authors stated that one can expect that accepting a second hydrogen bond is cooperative, as upon accepting the first hydrogen bond the oxygen becomes more negatively charged. They conclude that in such geometries anticooperativity results from short-ranged repulsion effects caused by unfavorable molecular-orbital interactions. This explanation is contradicted by the fact that in our calculations on similar geometries $\Delta E^{(1)}_{\mathrm{SCF}}$ is nearly additive or even slightly cooperative. The anticooperativity results from $\Delta E^{(2)}_{\mathrm{SCF}}$, which in our model is explained solely on the grounds of electrostatic polarization. We think this is not unexpected, because in such a geometry the electrostatic field caused by the second donating water molecule cancels the field of the first water molecule to a large extent. The failure of the FQ model in this respect may be a result of the lack of out-of-plane polarizability in a FQ water model using only atomic sites, since in these geometries the components of the electrostatic field perpendicular to the plane of the acceptor molecule cancel.

3.4. Repulsion Energy. The repulsion term of our potential serves to model all remaining, short-ranged contributions to the interaction energy, most importantly the exchange−repulsion energy, but also the charge-transfer, penetration, and exchange−dispersion contributions. So, the repulsion model was fitted to $\Delta E^{\mathrm{tot}}_{\mathrm{MP2}} - E_{\mathrm{elec}} - E_{\mathrm{pol}} - E_{\mathrm{disp}}$. In dimer geometries B and C we observed a preference for an acceptor angle of 0° in this remaining part of the energy. This preference can be interpreted as anisotropy of the oxygen atom: one does not expect the variation in for example the C···O distances to give rise to this variation in repulsion energy, especially not in dimer C, where such repulsions are expected to be very small. Therefore anisotropy in the oxygen repulsion was introduced, based on a model proposed by Price.[52] The repulsion energy was calculated using

$$E_{\mathrm{rep}}(r_{ij}) = A_{ij}\exp[-B_{ij}(r_{ij} - D_i(\boldsymbol{n_i}\cdot\boldsymbol{d_{ij}})^2 - D_j(\boldsymbol{n_j}\cdot\boldsymbol{d_{ij}})^2)]$$

where $\boldsymbol{n}$ is a unit vector perpendicular to the C−O−H plane, and $\boldsymbol{d_{ij}}$ is a unit vector in the direction of r_{ij}. The anisotropy parameter D is only nonzero for oxygen. So, for interactions not involving oxygen, this formula reduces to the normal exponential repulsion formula.

The fitting turned out to be not very straightforward. There was no unique solution, as many models could be constructed that all fitted the data more or less equally well in a least-squares sense. Strong correlations exist between the repulsion parameters. Not only are the A and B parameters for a single atom-type correlated, also correlation between parameters for different atoms can be quite large, since repulsion energy can often be described either on one or on the other atom. This occurred most severely for carbon and hydrogen in the methyl group. Many variations of the fit were tried, applying combination rules for the parameters A_{ij} and B_{ij} for all, some or none of the combinations of atoms.

Most fits described the hydrogen bonded geometries equally well, but it turned out to be difficult to model all methyl contacts accurately. When dimer M was correctly modeled, geometries K and N were too attractive; when the latter were satisfactory, M geometries were too repulsive. In order to improve this, the parameters for C and H_C were refined on about 20 methyl-contact geometries. Many fits were tried, but one of the best fits could be obtained by a model involving only repulsion on the hydrogen atoms, with the repulsion center shifted inwards by ~15%. Such a model was recently proposed by Fraschini

TABLE 4: Repulsion Parameter Set

i	*j*	$A_{ij}{}^{a}$	$B_{ij}{}^{a}$
C	C, H_C, H_O	0	0
C	O	81304	3.778
O	O	92607	4.049
O	H_C	9733	3.575
O	H_O	3903	3.671
H_C	H_C	4460	3.571
H_C	H_O	1228	3.744
H_O	H_O	338	3.918
D_O (Å)		0.075	
C−H bond shortening		85.5%	
RMSD^b		0.25/0.31/0.9	

[a] A_{ij} in kcal/mol, B_{ij} in Å^{-1}. H_C−H_O parameters taken from combination rules: $A_{ij} = \sqrt{A_{ii}A_{jj}}$, $B_{ij} = (B_{ii} + B_{jj})/2$. [b] Root-mean-square deviations for dimers with an interaction energy of less than 0 (71 dimers), less than 5 kcal/mol (89 dimers), and for all 94 dimers, in this order.

and Stone[53] for modeling the exchange−repulsion energy in methane dimer. There it was seen to be only slightly less accurate than much more elaborate models, incorporating many anisotropy parameters. Considering the very different origin of these parameterizations, and the fact that our repulsion model also describes many other short-range effects, the similarity between the models is remarkable.

Subsequently, all other parameters were refined using the following type of weighting:

$$w_i = \exp(-(E_i - E_{\min,i})/c)$$

In this formula $E_{\min,i}$ means the lowest energy within the dimer set for the type of interaction present in that dimer, e.g., −5.42 kcal/mol for OH···O bonded geometries and −1.17 kcal/mol for CH···O bonded geometries. *c* was taken to be 4 kcal/mol. This somewhat arbitrary weighting function is meant to focus the fit towards local minima in the potential rather than to the global minimum only, in order to ensure a balance in the importance of all different interactions in the fit. The cyclic CH···O bonded geometry (dimer J) remained problematic: when dimer J was well reproduced, all other CH···O geometries (dimers G−I) were generally too attractive. Because the latter were thought to be more representative for contacts occurring in crystals, we biased the fit to the dimers G−I by reducing the weight on dimers J with C···C distances less than 3.75 Å. The final parameters are reported in Table 4.

4. Evaluation of the Potential

For dimers A and B the potential is plotted together with the ab initio data. Figure 4 illustrates the anisotropy in the repulsion energy around the oxygen atom. In Figure 5 one can see that E_{pol} deviates from $\Delta E^{(2)}_{\text{SCF}}$ at short range. This is as expected, since the latter does not only contain polarization energy but also overlap-dependent contributions, such as the charge-transfer energy. Note that the difference between $\Delta E^{(2)}_{\text{SCF}}$ and E_{pol} is absorbed in the repulsion potential. This difference can be very significant: for methanol dimer in a near-equilibrium geometry E_{pol} equals −1.5 kcal/mol, while $\Delta E^{(2)}_{\text{SCF}}$ amounts to −3.0 kcal/mol. So, our polarization model does not describe $\Delta E^{(2)}_{\text{SCF}}$ itself, while at the same time it describes the nonadditivity in $\Delta E^{(2)}_{\text{SCF}}$ accurately. We can only conclude that the remaining terms in $\Delta E^{(2)}_{\text{SCF}}$ are approximately additive in the trimers that we have considered.

For all dimers and trimers minima in the potential energy were determined by rigid-body optimization of their geometries, using a local version of the TINKER package[44] which incorporates some new energy and derivative routines.

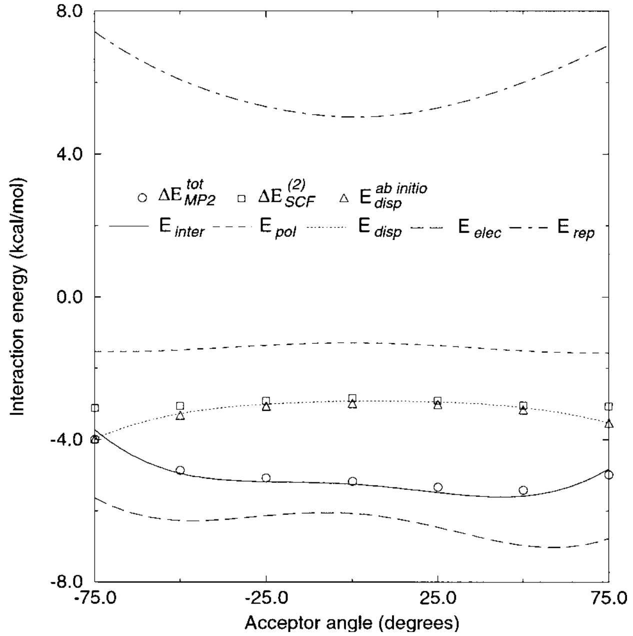

Figure 4. Interaction energies for dimer B (O···O distance 2.85 Å), including both ab initio data and those calculated from the model.

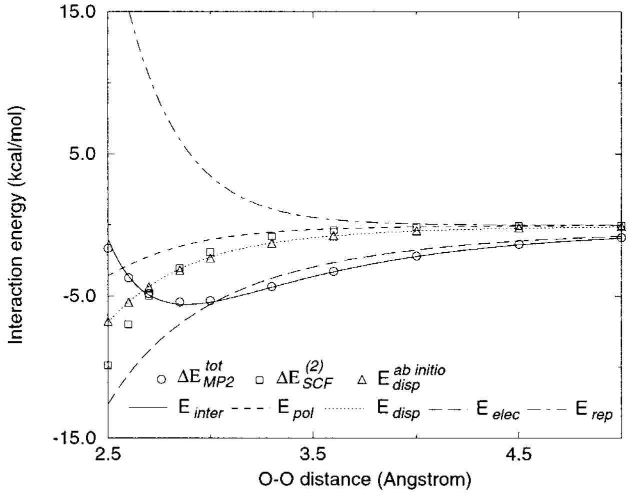

Figure 5. Interaction energies for dimer A (α = 50°), including both ab initio data and those calculated from the model.

4.1. Dimers. Interpolation of the ab initio data for dimer A (α = 50°) resulted in an optimal O···O distance of 2.89 Å. The global minimum of the model has an O···O distance of 2.87 Å at an acceptor angle of 44° and a nearly linear hydrogen bond (179°). The total energy in this geometry is −5.63 kcal/mol. The ab initio interaction energy in this geometry was calculated to be −5.44, to be compared with −5.30 kcal/mol from previous (less accurate) ab initio work at the counterpoise-corrected MP2 level.[54,55] The fact that our potential gives a somewhat too deep minimum compared to the ab initio data can be an advantage, since the limiting ab initio value may be 0.4 kcal/mol lower still, as discussed above. Adding this estimated basis set incompleteness error, brings our estimate for the binding energy of methanol dimer to ∼5.8 kcal/mol.

Most geometries optimized either to this global minimum or to the cyclic C−H···O bonded dimer J (−2.18 kcal/mol, R_{CO} = 3.51 Å). This minimum lies above the (interpolated) ab initio one of −2.41 kcal/mol. This is due to the problems that were encountered in fitting these geometries. Only three other

TABLE 5: Comparison of ab Initio and Model Interaction Energies for Various Dimers. Energies in kcal/mol

	geometry	R (Å)[a]	α[a]	β[a]	$E_{\mathrm{AB}}^{\text{ab initio}}$	$E_{\mathrm{AB}}^{\text{model}}$	$E_{\mathrm{disp}}^{\text{ab initio}}$	$E_{\mathrm{disp}}^{\text{model}}$
$(CH_3OH)_2$	*b*	2.87	44°	179°	−5.44	−5.63	−2.95	−2.87
$(H_2O)_2$	*b*	2.93	59°	174°	−4.68	−4.77	−1.97	−1.78
	Feller−Frisch	2.91	57°	173°	−4.69	−4.77	−2.03	−1.84
H_2O,CH_3OH	H_2O donor[b]	2.88	45°	178°	−5.25	−5.37	−2.51	−2.37
	H_2O acceptor[b]	2.93	58°	176°	−4.76	−4.91	−2.26	−2.08
$(CH_3)_2O,H_2O$	*b*	2.84	25°	179°	−5.43	−5.62	−2.99	−2.90
$(CH_4)_2$	*b*	3.61			−0.43	−0.45	−1.06	−0.93
		3.70			−0.44	−0.44	−0.91	−0.80
CH_4,H_2O	F−H[b]	3.45			−0.76	−0.91	−1.23	−1.07
	F−H	3.70			−0.78	−0.81	−0.76	−0.66
	V−O[b]	3.73	7.5°	180°	−0.56	−0.59	−0.69	−0.58
	V−O	3.70	0°	180°	−0.56	−0.58	−0.71	−0.60
	F−O	3.70			−0.16	−0.14	−0.43	−0.41

[a] R is the smallest R_{OO}, R_{CO}, or R_{CC}. α as defined in Figure 1; β is the O−H···O angle. [b] Geometry optimized in the model.

stationary points were found, starting from dimers M, F, and P. For those geometries the model and ab initio energies agree within 0.05 kcal/mol.

4.2. Trimers. For the cyclic trimer most ab initio work observed a shortening of the O···O distance relative to the dimer,[56−59] whereas Wheatley's systematic potential[14] resulted in a lengthening of the O···O distance. Our equilibrium geometry for the trimer, resembling trimer A, was seen to be in line with the ab initio work: the three O···O distances were 2.801, 2.802, and 2.815 Å, which amounts to an average R_{OO} shortening of ∼0.07 Å. This shortening is slightly underestimated compared to the ab initio studies, where values ranging from 0.08 to 0.10 Å were reported.

Our model gives an interaction energy of −16.85 kcal/mol for the equilibrium trimer geometry, compared with −15.23 kcal/mol at the counterpoise-corrected MP2/6-311+G(2d,2p) level.[57] This difference can be rationalized by noting that even using a slightly larger basis (i.e., 6-311+G(2df,2p)) only −5.15 kcal/mol has been obtained for the dimer whereas our model yields −5.63 kcal/mol. Since the trimer contains three such dimer interactions this accounts for the bulk of the difference.

Minima resembling trimers B and C were 0.78 kcal/mol and 2.31 kcal/mol less favorable. Recent DFT results at the B3LYP/6-311++G(3d2f,2p) level report values of 0.80 and 1.46 kcal/mol for these differences.[56] Within our potential the planar trimer C contained ∼2.5 kcal/mol less dispersion energy compared to the other two geometries, which only differ by 0.1 kcal/mol in that respect. Lack of dispersion energy at the theoretical (DFT) level could be part of the explanation of the larger difference in relative energy for this geometry C.

The minimum for the double acceptor geometry in trimer E (−10.52 kcal/mol) showed a lengthening of the O···O distance to 2.91 Å, illustrating the anticooperative effects present in this geometry. We did not find a stable linear trimer within our model, which is again in agreement with ab initio results.[56]

4.3. Other Molecules: Water, Methane, Dimethyl Ether. In order to assess to what extent our model is transferable to similar molecules, we studied some dimers containing water, methane, and dimethyl ether. For each of these molecules AMMs were obtained by ESP-fitting at the MP2/IOM level. First we calculated the optimal geometry using our potential. Then we calculated the ab initio MP2/IOM interaction energies, together with the dispersion energy, in this optimal geometry. In order to compare our ab initio results to previous work we also considered some other geometries. For dimers between methane and water we used some of geometries calculated by Szczęśniak et al.:[60] the H_3CH···O hydrogen bond (V(ertex)−O), the HCH_3···HO hydrogen bond (F(ace)−H), and HCH_3···O (F(ace)−O).

The model interaction energies are generally in excellent agreement with the ab initio data (Table 5). Therefore, we conclude that the assumed transferability to similar molecules holds. A large relative deviation occurs for the OH···H_3C hydrogen bond in methane−water dimer (F−H). The repulsion energy is apparently somewhat underestimated in the model. This is, however, a rather unimportant geometry from the point of view of crystal structure modeling. Another point is that for all dimers the model results in somewhat underestimated dispersion energies compared to the ab initio data. Because this underestimation occurs for methane, water, and mixed dimers alike, there is no straightforward explanation to be found in an imbalance of the individual parameters (e.g., too much C−C and too little O−O dispersion). Still, subtle changes of the parameters might improve the dispersion model for these systems as well, but for the moment we are rather pleased with the degree of transferability observed. The total interaction energy is always closer to the ab initio value then one would expect based on the deviations in the dispersion energy, which must be due to a fortuitous compensation of errors. For most dimers, the minimum in the potential is somewhat deeper than the ab initio value. This can be advantageous, since the latter is somewhat underestimated.

For some systems we can also compare our results to previous ab initio work. These comparisons are somewhat hindered by the fact that such work differs in methodology and accuracy. Water dimer has been the subject of many ab initio studies, and limiting values at the MP2(frozen-core) level have been established.[23,29,30] So, for this system we can conclude that our potential is ∼0.2 kcal/mol too shallow, while the optimal O···O distance is rather well reproduced (i.e., 0.02 Å too long). For other systems, the limiting interaction energies are not known, and the previous ab initio work is not more accurate than our own ab initio calculations. Therefore we cannot arrive at such precise conclusions about the absolute accuracy of our model. Still, we can obtain valuable information from relative energies and trends in geometry.

For dimers involving methanol, water, and dimethyl ether we can compare our results to counterpoise-corrected calculations at the MP2/6-311+G(2df,2p) level.[57] All our interaction energies are lower than the values obtained at this level, which is not surprising: the interaction energy for water dimer at this level was reported to be only −4.47 kcal/mol. However, relative stabilities are in good agreement with our results, and trends in the O···O distances are excellently reproduced. For example, at this ab initio level the O···O distance for methanol dimer is reported to be 0.06 Å shorter than for water dimer, and in the methanol−water mixed dimers the O···O distance is 0.05 Å shorter when water acts as a hydrogen-bond donor compared to water as an acceptor.

For methane we have obtained a minimum energy that is lower than that calculated by Szczęśniak et al.,[61] who reported −0.33 kcal/mol at 3.70 Å and −0.35 kcal/mol at 3.96 Å. It is hard to say whether our optimal C···C distance is too short, because the $(CH_4)_2$ potential as a function of R is very flat and improving the basis set can have a significant influence on the position of the minimum. For methane−water previous ab initio work[25,60] is in close agreement with the present ab initio results. Our optimal C···O distances for the V−O geometry (3.73 Å) is somewhat shorter than estimated by Szczęśniak et al.[60] (3.91 Å), which may be due to the larger basis set employed in our ab initio work.

5. Conclusions

In this work we have described the derivation of a general transferable potential for organic molecules such as alkanes, ethers, and alcohols. The potential was derived from high-quality ab initio interaction energies for methanol, which it accurately reproduces. The assumed transferability to other molecules was seen to hold, as the model proved to be equally accurate for related systems. An important factor that contributes to this transferability is that we fitted the electrostatic, dispersion, polarization, and repulsion energies all separately, leading to a potential where these terms have a well-defined physical meaning. Of course, if one wants to study a specific system (e.g., water) one could use the methodology presented in this article to derive a more accurate potential for that system, at the expense of transferability.

Electrostatic interactions are a very important contribution for hydrogen-bonded systems. Therefore it is essential to use a highly accurate model for this term, including proper treatment of anisotropic features. An atomic multipole model was derived by fitting to the electrostatic potential of the monomer. This was found to have superior convergence properties compared to a DMA-derived multipole model.

Energy nonadditivity is significant in methanol trimers, varying from −2.8 to +0.7 kcal/mol, nearly all present in the SCF deformation energy ($\Delta E^{(2)}_{SCF}$). This justifies that the only nonadditive term in the potential is the polarization energy. Atomic dipole polarizabilities were shown to provide a simple but accurate way to model the nonadditive effects. Our polarization model was fitted to reproduce the nonadditivity in $\Delta E^{(2)}_{SCF}$, but gives polarization energies that, at short range, are significantly less than the $\Delta E^{(2)}_{SCF}$ values themselves. Apparently, all overlap-dependent contributions to $\Delta E^{(2)}_{SCF}$, such as charge-transfer, are approximately additive.

The dispersion energies from perturbational theory were fitted by a damped r^{-6} formula. Finally, the repulsion term in our potential serves to model all remaining, short-ranged interactions, and includes anisotropy of the oxygen atom. Since it was the most difficult term to parameterize, the deviation from the ab initio data is usually larger at short range. For minimum-energy geometries the maximum deviation between the model and the ab initio data is around 0.2 kcal/mol. Mostly the potential gives the deeper minimum, which may be advantageous because it compensates for the underestimation due to the basis set incompleteness error. The model was seen to be of similar accuracy for dimers involving methane, water, and dimethyl ether. Trends in hydrogen-bond distances and relative energies in these dimers as well as in methanol trimers are in line with ab initio studies.

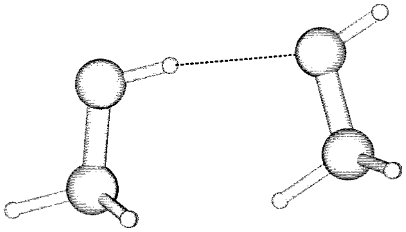

Figure 6. Geometry used for the optimization of the exponents of the polarization functions. R_{OO} = 2.95 Å, R_{CC} = 3.54 Å.

Since we are interested in modeling crystal structures, we now have to assess whether our model is appropriate for simulation of condensed phases. The way to do this is by performing energy minimization for experimental crystal structures. In a companion article[15] we describe the validation of this potential by such energy minimizations as well as its application to crystal structure prediction.

Acknowledgment. This work was supported by the Council for Chemical Sciences of the Netherlands Organization for Scientific Research (CW-NWO), in the framework of the PPM/CMS-crystallization project. We thank Mr. G. Schaftenaar, for the integration of the multipole fitting routine in the MOLDEN program, and Dr. J. H. van Lenthe, for assistance with the various ab initio programs.

Appendix

A. Basis Set Optimization. The basis set used in this study was developed specifically for methanol dimer. It is an interaction-optimized basis designed to give *uniform accuracy*[23] in ΔE_{AB}. This means that the truncation error in ΔE_{AB} due to basis set incompleteness is kept about the same for each of the various polarization function types ϕ_i, at all sites in the system. In practice this is achieved in an iterative series of ΔE_{AB} calculations in which the number of ϕ_i sets (d, f, g, ..., on O and C, etc.) and their exponents are varied systematically. The overall ΔE_{AB} truncation error we aimed at was 0.2 kcal/mol, which amounts to an admissible error of about 0.02 kcal/mol for a given ϕ_i-type. In practice a ϕ_i-set contributing more than about 0.04 kcal/mol to ΔE_{AB} should then be kept in the final basis.

For the isotropic part of the basis we needed a set that combines moderate size with an accurate description of the atomic outer region. Therefore the EZ ("extended zeta")[24,25] basis was used for carbon and oxygen. This set is a (10,6) primitive set of GTO's contracted to a [5,3] basis. For hydrogen the DZ[26] basis was used, which is a (4s) primitive set of GTO's contracted to a [2s] basis. The EZ basis on C and O was augmented with sets of d, f, and g polarization functions, while p functions were added to the DZ basis at all H's. The exponents of these polarization functions were optimized in this work, with the exception of the α_d^O and $\alpha_p^{H_O}$, which were taken from previous work on water dimer.[27] When we enlarged the basis set by using two sets of functions of the same angular momentum we used $\alpha_1 = 2\alpha$, $\alpha_2 = \alpha/2$; and $\alpha_1 = 3\alpha$, $\alpha_2 = \alpha$, $\alpha_3 = \alpha/3$ when we used three sets.

The determination of the exponents for the remaining polarization functions was performed on a hydrogen-bonded methanol dimer (Figure 6). A geometry was chosen in which not only the oxygen but also the carbon atoms are in close contact, in order to ensure that variation of the various exponents would have a significant effect on the interaction energies. The intramolecular geometry was kept fixed throughout the calculations.

TABLE 6: Components of the Interaction Energy (kcal/mol) for Various f and g Exponents on Carbon

	a	$\alpha_f^C(0.15/0.6^a)$	$\alpha_f^C(0.15/0.6)^a$ $\alpha_g^C(0.3)$	$\alpha_f^C(0.1875/0.75)^a$ $\alpha_g^C(0.45)$	$\alpha_f^C(0.125/0.5)^a$ $\alpha_g^C(0.2)$	$\alpha_f^C(0.3)^a$ $\alpha_g^C(0.3)$	$\alpha_f^C(0.3)^b$ $\alpha_g^C(0.3)$	$\alpha_f^C(0.3)^c$ $\alpha_g^C(0.3)$
ΔE_{MP2}^{tot}	−3.567	−3.635	−3.683	−3.677	−3.682	−3.679	−3.648	−3.686
ΔE_{MP2}^{corr}	−2.667	−2.734	−2.778	−2.773	−2.777	−2.770	−2.748	−2.781
ΔE_{SCF}^{1}	1.091	1.093	1.093	1.093	1.093	1.088	1.096	1.096
ΔE_{SCF}^{2}	−1.992	−1.995	−1.998	−1.998	−1.998	−1.997	−1.995	−2.001

[a] 3d(0.9/0.3/0.1) on C + p (0.3) on H_C. [b] 2d (0.6/0.15) on C + p (0.3) on H_C. [c] 2d (0.6/0.15) on C + 2p (0.15/0.6) on H_C.

TABLE 7: Components of the Interaction Energy (kcal/mol) for Various f and g Exponents on Oxygen

		$\alpha_f^O(0.3125/1.25)$	$\alpha_f^O(0.3125/1.25)$ $\alpha_g^O(0.75)$	$\alpha_f^O(0.25/1.0)$ $\alpha_g^O(0.5)$	$\alpha_f^O(0.40625/1.625)$ $\alpha_g^O(1.125)$
ΔE_{MP2}^{tot}	−3.479	−3.639	−3.683	−3.684	−3.690
ΔE_{MP2}^{corr}	−2.592	−2.724	−2.778	−2.774	−2.778
ΔE_{SCF}^{1}	1.093	1.082	1.093	1.091	1.085
ΔE_{SCF}^{2}	−1.983	−1.997	−1.998	−2.001	−1.997

A.1. Optimization. A.1.1. Carbon. The exponent of a single d-function on carbon was varied from 0.05 to 4.5, while the exponents of the polarization functions on oxygen and hydrogen were kept fixed ($\alpha_p^H = 0.39$ and $\alpha_d^O = 0.5$, no f or g functions). There was hardly any effect on ΔE_{SCF}. The influence on ΔE_{MP2}^{corr} was substantially larger, at maximum −0.15 kcal/mol for $\alpha_d^C = 0.25$ and 0.3.

Subsequently the exponents of the f and the g functions on carbon were optimized simultaneously. The optimization was done using a (2dfg) polarization function set, with $\alpha_d^C = 0.6/0.15$. The exponent α_g^C was varied from 0.03 to 2.5, and for each choice the f exponent was taken to be the average of α_d^C and α_g^C. The SCF energies were hardly influenced (<0.01 kcal/mol); the largest effect on the MP2 energy occurred for $\alpha_g^C =$ 0.3 and 0.22. After completion of all optimizations, α_g^C was further checked by varying it with EZ(3d2fg) on C and O and DZ(p) on hydrogen. The components of the interaction energy are given in Table 6. The influence on all energy terms was very small and 0.3 was taken to be the final value, resulting in $\alpha_d^C = \alpha_f^C = \alpha_g^C = 0.3$.

A.1.2. Oxygen. The exponents of the f and the g functions on oxygen were optimized simultaneously. The optimization was done using a (2dfg) polarization function set, with $\alpha_d^O =$ 1.0/0.25. The α_g^O was varied from 0.2 to 1.5, and for each choice the f exponent was taken to be the average of α_d^O and α_g^O. SCF energies were only slightly influenced (∼0.03 kcal/mol) by these f and g functions. The largest effect on ΔE_{MP2}^{corr} was −0.13 kcal/mol for $\alpha_g^O = 0.36$. However, ΔE_{MP2}^{corr} was rather insensitive to changes of α_g^O up to 1.5. After completion of all optimizations the g exponent on oxygen was varied with EZ(3d2fg) on C and O and DZ(p) on hydrogen. Components of the interaction energies are given in Table 7. The influence of varying the g exponent around $\alpha_g^O = 0.75$ on all energy terms was very small, and 0.75 was taken to be the final value.

A.1.3. Hydrogen. The exponent of the p function on hydrogens bonded to carbon was varied from ∼0.1 to ∼1.5. $\Delta E_{SCF}^{(1)}$ was only slightly influenced (∼0.02 kcal/mol). The largest effect on the MP2 energy contribution was −0.3 kcal/mol, around $\alpha_p^H = 0.3$, which was taken to be the final value.

A.2. Final Choice of the Basis Set. Tables 6 and 7 show that both on oxygen and on carbon the g functions contribute in the order of 0.05 kcal/mol to the interaction energy. As discussed above, these g functions must therefore be kept in the final basis, and the same threshold governs the choice of extra p, d, or f sets. The f functions were seen to be more important for oxygen (0.16 kcal) than for carbon (0.068). Moreover, addition of the second set of f functions on carbon had a very small effect (0.004 kcal/mol) and one set would suffice. We also tested whether changing from DZ(p) on H_C to DZ(2p) would be an improvement. However, program limitations of 255 basis functions forced us to decrease the number of functions on another atom in order to achieve this. Considering the smaller effects of the f functions on carbon, a (2dfg) set was taken for that element. From Table 6 it can be seen that the energy that was gained by adding the extra p functions on H_C (∼0.04 kcal/mol) equals the amount lost by switching from 3d to 2d on carbon. Hence, a (3dfg) set on carbon combined with a 2p set on all hydrogens would be preferable. However, with 258 functions this set is just too large. For use in our further calculations we chose (2dfg) on carbon with (2p) on hydrogen, as the use of two p functions on hydrogen was thought to be more balanced with respect to the d, f, and g functions for C and O.

Our exponents for C and O compare well with the exponents of the aug-cc-pVTZ and aug-cc-pVQZ sets for these atoms.[31,32] They cover the low-exponent (3d2fg) subset of the (4d3f2g) aug-cc-pVQZ set, and they tend to be somewhat lower than the aug-cc-pVTZ exponents. This is as expected, since we optimized on interaction energy and low exponents are more important for an accurate dispersion energy. So, adding the mentioned subset of the aug-cc-pVQZ polarization functions to an EZ basis may be a general recipe to construct a medium-sized basis set for the calculation of interaction energies.

Supporting Information Available: Geometries of all methanol dimers and trimers, together with ab initio and model interaction energies. IOM and DZ($2d^O$) basis sets. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Gavezzotti, A.; Filippini, G. In *Theoretical Aspects and Computer Modeling of the Molecular Solid State*; Gavezzotti, A., Ed.; John Wiley & Sons: New York, 1997; Chapter 3.

(2) Verwer, P.; Leusen, F. J. J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 1998; Vol. 12, Chapter 7.

(3) Gavezzotti, A. *Acc. Chem. Res.* **1994**, *27*, 309.

(4) Gdanitz, R. J. *Curr. Opin. Solid State Mater. Sci.* **1998**, *3*, 414.

(5) Price, S. L.; Wibley, K. S. *J. Phys. Chem.* **1997**, *A101*, 2198.

(6) Coombes, D. S.; Nagi, G. K.; Price, S. L. *Chem. Phys. Lett.* **1997**, *265*, 532.

(7) Mooij, W. T. M.; van Eijck, B. P.; Price, S. L.; Verwer, P.; Kroon, J. *J. Comput. Chem.* **1998**, *19*, 459.

(8) Price, S. L.; Stone, A. J. *J. Chem. Soc., Faraday Trans.* **1992**, *88*, 1755.

(9) Millot, C.; Stone, A. J. *Mol. Phys.* **1992**, *77*, 439.

(10) Millot, C.; Soetens, J.-C.; Martins-Costa, M. T. C.; Hodges, M. P.; Stone, A. J. *J. Phys. Chem.* **1998**, *A102*, 754.

(11) Wallqvist, A.; Ahlström, P.; Karlström, G. *J. Phys. Chem.* **1990**, *94*, 1649.

(12) Åstrand, P.-O.; Wallqvist, A.; Karlström, G.; Linse, P. *J. Chem. Phys.* **1991**, *95*, 8419.

(13) Wheatley, R. J. *Mol. Phys.* **1996**, *87*, 1083.

(14) Buck, U.; Siebers, J.-G.; Wheatley, R. J. *J. Chem. Phys.* **1998**, *108*, 20.

(15) Mooij, W. T. M.; van Eijck, B. P.; Kroon, J. *J. Phys. Chem. A* **1999**, *103*, 9883.

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

(17) Del Bene, J. E.; Shavitt, I. In *Molecular Interactions: From van der Waals to Strongly Bound Complexes*; Scheiner, S., Ed.; John Wiley and Sons: New York, 1997; Chapter 5.

(18) Klopper, W.; Schütz, M.; Lüthi, H. P.; Leutwyler, S. *J. Chem. Phys.* **1995**, *103*, 1085.

(19) Saunders, V.; Guest, M. F. *ATMOL Program Package*; SERC Daresbury Laboratory: Great Britain.

(20) van Lenthe, J. H. *SERVEC, ATMOL Vector Service Program*; Utrecht University: The Netherlands.

(21) van Lenthe, J. H. *INTACAT Program Package*; Utrecht University: The Netherlands.

(22) van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. In *Theoretical Treatments of Hydrogen Bonding*; Hadži, D., Ed.; John Wiley and Sons: New York, 1997; Chapter 2

(23) van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *J. Chem. Phys.* **1999**, *111*, 3812.

(24) van Lenthe, J. H.; van Duijneveldt, F. B. *J. Chem. Phys.* **1984**, *81*, 3168.

(25) van Mourik, T. Correlated ab initio calculations on weakly bonded systems. PhD thesis, Utrecht University, Utrecht, 1994.

(26) van Duijneveldt, F. B. *IBM Research Report, RJ945*; IBM: San Jose, 1971.

(27) van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *J. Chem. Phys.* **1992**, *97*, 5019.

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

(29) Halkier, A.; Koch, H.; Jørgenson, P.; Christiansen, O.; Beck Nielsen, I. M.; Helgaker, T. *Theor. Chem. Acc.* **1997**, *97*, 150.

(30) Schütz, M.; Brdarski, S.; Widmark, P.-O.; Lindh, R.; Karlström, G. *J. Chem. Phys.* **1997**, *107*, 4597.

(31) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(32) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6769.

(33) Smith, B. J.; Swanton, D. J.; Pople, J. A.; Schaeffer, H. F., III; Radom, L. *J. Chem. Phys.* **1990**, *92*, 1240.

(34) Gerry, M. C. L.; Lees, R. M.; Winnewisser, G. *J. Mol. Spectrosc.* **1976**, *61*, 231.

(35) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047.

(36) Ángyán, J. G.; Chipot, C. *Int. J. Quantum Chem.* **1994**, *52*, 17.

(37) Williams, D. E. *J. Comput. Chem.* **1988**, *9*, 745.

(38) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Johnson, P. M. W. G. 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. *Gaussian94*, Revision B.2; Gaussian Inc.: Pittsburgh, PA, 1995.

(40) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Design*. In press. MOLDEN is available from: http://www.caos.kun.nl/~schaft/molden/molden.html.

(41) Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548.

(42) Hinsen, K.; Roux, B. *J. Comput. Chem.* **1997**, *18*, 368.

(43) Ahlrichs, R.; Penco, R.; Scoles, G. *Chem. Phys.* **1977**, *19*, 119.

(44) Ponder, J. W. *TINKER: Software Tools for Molecular Design*, Version 3.6; Washington University of Medicine: St. Louis, 1998. Available from: http://dasher.wustl.edu/tinker/.

(45) Rowland, R. S.; Taylor, R. *J. Phys. Chem.* **1996**, *100*, 7384.

(46) Szczęśniak, M. M.; Brenstein, R. J.; Cybulski, S. M.; Scheiner, S. *J. Phys. Chem.* **1990**, *94*, 1781.

(47) Szczęśniak, M. M.; Chałasiński, G. *J. Mol. Struct. (Theochem)* **1992**, *261*, 37.

(48) Clementi, E.; Kolos, W.; Lie, G. C.; Ranghino, G. *Int. J. Quantum Chem.* **1980**, *17*, 377.

(49) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.

(50) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(51) Liu, Y.-P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.

(52) Price, S. L. *Mol. Phys.* **1986**, *58*, 651.

(53) Fraschini, E.; Stone, A. J. *J. Comput. Chem.* **1998**, *19*, 847.

(54) Bizzarri, A.; Stolte, S.; Reuss, J.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *Chem. Phys.* **1990**, *143*, 423.

(55) Bleiber, A.; Sauer, J. *Chem. Phys. Lett.* **1995**, *238*, 243.

(56) Mó, O.; Yáñez, M.; Elguero, J. *J. Chem. Phys.* **1997**, *107*, 3592.

(57) Masella, M.; Flament, J. P. *J. Chem. Phys.* **1998**, *108*, 7141.

(58) Hagemeister, F. C.; Gruenloh, C. J.; Zwier, T. S. *J. Phys. Chem.* **1998**, *A102*, 82.

(59) Provencal, R. A.; Paul, J. B.; Roth, K.; Chapo, C.; Casaes, R. N.; Saykally, R. J.; Tschumper, G. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1999**, *110*, 4258.

(60) Szczęśniak, M. M.; Chałasiński, G.; Cybulski, S. M.; Cieplak, P. *J. Chem. Phys.* **1993**, *98*, 3078.

(61) Szczęśniak, M. M.; Chałasiński, G.; Cybulski, S. M.; Scheiner, S. *J. Chem. Phys.* **1990**, *93*, 4243.