PCCP

Dynamic Article Links

Cite this: *Phys. Chem. Chem. Phys.*, 2012, **14**, 7669–7678

www.rsc.org/pccp

PAPER

Electrostatically embedded many-body method for dipole moments, partial atomic charges, and charge transfer†

Hannah R. Leverentz, Katie A. Maerzke, Samuel J. Keasler, J. Ilja Siepmann and Donald G. Truhlar*

Received 24th December 2011, Accepted 31st January 2012
DOI: 10.1039/c2cp24113g

Fragment methods have been widely studied for computing energies and forces, but less attention has been paid to nonenergetic properties. Here we extend the electrostatically embedded many-body (EE-MB) method to the calculation of cluster dipole moments, dipole moments of molecules in clusters, partial atomic charges, and charge transfer, and we test and validate the method by comparing to results calculated for the entire system without fragmentation. We also compare to calculations carried out by the conventional many-body (MB) method without electrostatic embedding. Systems considered are $NH_3(H_2O)_{11}$, $(NH_3)_2(H_2O)_{14}$, $[Cl(H_2O)_6]^-$, $(HF)_4$, $(HF)_5$, $(HF)_2H_2O$, $(HF)_3H_2O$, and $(HF)_3(H_2O)_2$. With electrostatic embedding, we find that even at the pairwise additive level a quantitatively accurate description of a system's dipole moment and partial charge distribution and a qualitatively accurate description of the amount of intermolecular charge transfer can often be obtained.

1. Introduction

Fragment methods constitute one of the most powerful methods for modeling large and complex chemical systems.[1–34] To date, most fragment models have been developed for the purpose of calculating energies. However there is also a need to calculate other properties, such as the charge distribution. In this article we consider how well the electrostatically embedded many-body (EE-MB) approximation[7,25] can predict charge distributions in extended clusters, in particular cluster dipole moments, dipole moments of individual molecules in clusters, partial atomic charges, and charge transfer. The inclusion of charge transfer in our study is of particular interest because it is usually considered a shortcoming of fragment methods that they do not include charge transfer between fragments (for a recent study showing one way to overcome this restriction, see the work of Isegawa *et al.*[35]).

There has been considerable interest in devising quantum mechanical methods for calculating molecular dipole moments in condensed-phase materials. For example, much work has been done using maximally localized Wannier functions (MLFW) to estimate dipole moments of individual molecules, charge distributions, and bond orders in aqueous solutions and crystalline solids with defects.[36–46] These methods require the user to perform an electronic structure calculation using a plane wave basis set; the final molecular orbitals obtained in the plane wave basis are then transformed into localized functions that can be assigned to a specific molecule or fragment and used to calculate molecular dipole moments and charge distributions. The computational cost of MLWF-based calculations therefore scales asymptotically with system size in the same way that the chosen electronic structure method does. The difficulties of calculating partial atomic charges in condensed phases have also been recently discussed.[47,48]

In this work, we investigate the ability of the EE-MB method to predict charge distributions in molecular clusters. We examine several aspects: (i) modeling the full-system system dipole moments of clusters directly (that is, without using the concept of partial atomic charges charges), (ii) modeling partial atomic charges, (iii) using the partial atomic charges to calculate dipole moments of individual molecules in a cluster, and (iv) using the partial atomic charges to calculate net charges on individual molecules in a cluster. We note that we are working with concepts (partial atomic charges in (ii), subsystem dipole moments in (iii), and subsystem net charges in (iv)) that are not uniquely defined by quantum mechanics, but they are broadly used modeling constructs whose meaning is rather intuitively obvious and is widely understood.

The EE-MB calcuations have the advantage that the asymptotic scaling of their computational cost with system size ultimately is not dependent on the chosen level of electronic structure theory but rather on the level to which one chooses to take the EE-MB expansion.[7] Also, the EE-MB approximation is relatively easy to implement because it does not require

Department of Chemistry, Chemical Theory Center, and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431, USA

† Electronic supplementary information (ESI) available: Tables showing details of the dipole moments, partial atomic charges, and net molecular charges calculated in this study. See DOI: 10.1039/c2cp24113g

minimization procedures and matrix transformations of molecular orbitals; rather, the EE-MB approximation requires only linear combinations of fragment properties. Therefore, it is worthwhile to investigate the ability of various levels of the EE-MB approximation to predict charge distributions in noncovalently interacting molecular clusters.

2. Theory

The EE-MB method has been formulated for predicting energies of an entire system from the energies of its constituent monomers, dimers, and trimers, each embedded in the electrostatic field of the other monomers.[7] The monomers of which the dimers and trimers are composed are nonbonded fragments, *e.g.*, water molecules in a water cluster or water liquid. Here we develop the EE-MB approximations to molecular dipole moments and partial atomic charges. From the latter we can also calculate interfragment charge transfer.

2.1. Dipole moments

The EE-MB approximation to the magnitude of the full-system dipole moment (μ) is calculated as follows: (1) the x, y, and z components of the dipole moment are calculated from the electron densities of the various electrostatically embedded oligomers (monomers, dimers, or trimers) in the system, (2) each of the three components of the full-system dipole moment (μ_p, where $p = x$, y, or z) is calculated as a linear combination of the corresponding components of the oligomer dipoles (the linear combination used is the same as that used for the many-body approximation of the system's potential energy), and (3) the magnitude of the vector resulting from step (2) is taken to be μ. Let N be the total number of monomers into which the system has been divided. Let p be set equal to x, y, or z, let μ_p^i indicate the p-component of the dipole moment vector of electrostatically embedded monomer i, let μ_p^{ij} indicate the p-component of the dipole moment vector of electrostatically embedded dimer ij, and let μ_p^{ijk} indicate the p-component of the dipole moment vector of electrostatically embedded trimer ijk. Let us define $\mu_p^{(1)}$, $\mu_p^{(2)}$, and $\mu_p^{(3)}$ according to eqn (1a)–(1c), respectively:

$$\mu_p^{(1)} \equiv \sum_{i=1}^{N} \mu_p^i \tag{1a}$$

$$\mu_p^{(2)} \equiv \sum_{i<j} (\mu_p^{ij} - \mu_p^i - \mu_p^j) \tag{1b}$$

$$\mu_p^{(3)} \equiv \sum_{i<j<k} [\mu_p^{ijk} - (\mu_p^{ij} - \mu_p^i - \mu_p^j) - (\mu_p^{ik} - \mu_p^i - \mu_p^k) - (\mu_p^{jk} - \mu_p^j - \mu_p^k) - \mu_p^i - \mu_p^j - \mu_p^k] \tag{1c}$$

Then we may write the electrostatically embedded one-body (EE-1B), two-body (or pairwise additive, EE-PA), and three body (EE-3B) approximations to μ_p (the p-component of the full-system dipole moment vector) according to eqn (2a)–(2c), respectively:

$$\mu_p \approx \mu_p^{\text{EE-1B}} = \mu_p^{(1)} \tag{2a}$$

$$\mu_p \approx \mu_p^{\text{EE-PA}} = \mu_p^{(1)} + \mu_p^{(2)} \tag{2b}$$

$$\mu_p \approx \mu_p^{\text{EE-3B}} = \mu_p^{(1)} + \mu_p^{(2)} + \mu_p^{(3)} \tag{2c}$$

Finally, we can approximate the magnitude of the full-system dipole moment vector, μ, using the EE-1B, EE-PA, or EE-3B approximation according to eqn (3a), (3b) or (3c), respectively:

$$\mu \approx \mu^{\text{EE-1B}} = \sqrt{[(\mu_x^{\text{EE-1B}})^2 + (\mu_y^{\text{EE-1B}})^2 + (\mu_z^{\text{EE-1B}})^2]} \tag{3a}$$

$$\mu \approx \mu^{\text{EE-PA}} = \sqrt{[(\mu_x^{\text{EE-PA}})^2 + (\mu_y^{\text{EE-PA}})^2 + (\mu_z^{\text{EE-PA}})^2]} \tag{3b}$$

$$\mu \approx \mu^{\text{EE-3B}} = \sqrt{[(\mu_x^{\text{EE-3B}})^2 + (\mu_y^{\text{EE-3B}})^2 + (\mu_z^{\text{EE-3B}})^2]} \tag{3c}$$

The origin and orientation of each oligomer are kept the same as its origin and orientation in the full cluster to ensure that the EE-MB approximation of the total dipole moment vector retains its intended physical meaning even for charged clusters.

One can also find the "pure" many-body (MB) approximations to the full-system dipole moment using eqn (1)–(3) without electrostatic embedding; in this case one uses the components of the dipole moments of monomers, dimers, and trimers that are not surrounded by some representation of the electrostatic potential of the rest of the system.

2.2. Point charges

Assume that we have a system that has been divided into N monomers. If we had the wave function or electron density of the entire system, we could use one of several available charge analysis methods (such as Mulliken population analysis,[49] CHelpG electrostatic potential fitting,[50] or Natural Population Analysis[51]) to assign a partial charge to each atom in the system. This set of partial charges could in turn be used as a rough approximation of the entire system's electron density. Let us now assume that we do not wish to expend the computational resources necessary to calculate the electron density of the entire system directly but that we do have enough computational resources to calculate the wave functions of the individual monomers, dimers, and trimers in the system. Our objective is to show here that from the wave functions of the electrostatically embedded monomers, dimers, and trimers we may obtain a set of partial charges that represents the electron density of the entire system. As an example, let us focus on the partial charge that we would like to assign to atom A; call this partial charge q_A. Let us say that atom A belongs to monomer i; that is, $A \in i$. We may then approximate q_A (the charge on atom A from the electron density of the entire system) according to the EE-1B, EE-PA, or EE-3B approximations according to eqn (4)–(6), respectively. In the EE-1B approximation we have

$$q_A \approx q_A^{\text{EE-1B}} = q_A^i \tag{4}$$

where q_A^i is the partial charge assigned to atom A from the wave function or electron density of electrostatically embedded monomer i. In the EE-PA approximation we have

$$q_A \approx q_A^{\text{EE-PA}} = q_A^i + \sum_{j \neq i} (q_A^{ij} - q_A^i) \tag{5a}$$

$$= \sum_{j \neq i} q_A^{ij} - (N-2) q_A^i \tag{5b}$$

where q_A^{ij} is the partial charge assigned to atom A from the wave function or electron density of electrostatically embedded dimer ij (which is composed of monomers i and j). The EE-3B approximation yields

$$q_A \approx q_A^{\text{EE-3B}} = q_A^{\text{EE-PA}} + \sum_{j\neq i}^{N}\sum_{\substack{k\neq i\\ k>j}}^{N}[q_A^{ijk} - (q_A^{ij} - q_A^{i}) - (q_A^{ik} - q_A^{i}) - q_A^{i}] \quad (6a)$$

$$= \sum_{j\neq i}^{N}\sum_{\substack{k\neq i\\ k>j}}^{N} q_A^{ijk} - (N-3)\sum_{j\neq i}^{N} q_A^{ij} + \frac{(N-2)(N-3)}{2} q_A^{i} \quad (6b)$$

where q_A^{ijk} is the partial charge assigned to atom A from the wave function or electron density of electrostatically embedded trimer ijk (which is composed of monomers i, j, and k).

As with the dipole moments, one can calculate the pure MB approximations to the atomic partial charge distribution of the entire system by using the partial charge distributions of the isolated monomers, dimers, and trimers.

In addition to calculating the partial atomic charges for their own sake, one can also use them to calculate dipole moments of individual molecules in the cluster.

3. Systems

Several configurations of seven different systems were chosen for this study: forty-four configurations of $(NH_3)(H_2O)_{11}$ (Fig. 1), forty-four configurations of $(NH_3)_2(H_2O)_{14}$ (Fig. 2), six configurations of $[Cl(H_2O)_6]^-$ (Fig. 3 and 4), six configurations of $(HF)_4$ (Fig. 5), ten configurations of $(HF)_5$ (Fig. 5), six configurations of $(HF)_3(H_2O)$ (Fig. 6), and two configurations of $(HF)_3(H_2O)_2$ (Fig. 7). The EE-MB (and in some cases the MB) approximations of the dipole moment of each configuration are compared below to the conventionally

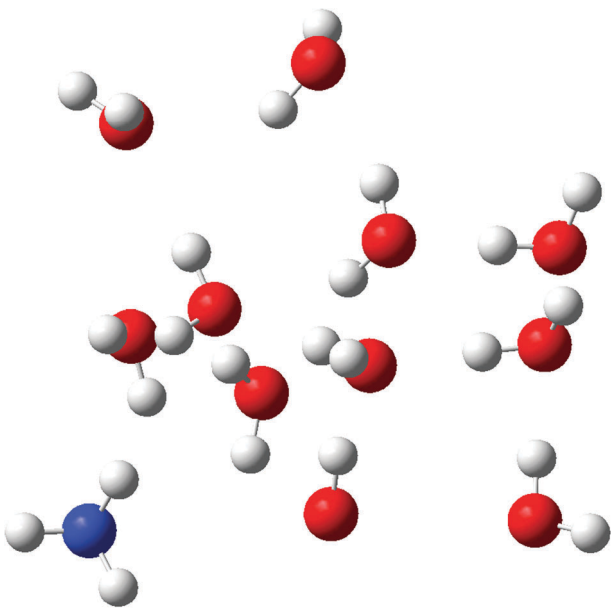

Fig. 1 The starting configuration used for a Monte Carlo simulation of $(NH_3)(H_2O)_{11}$.

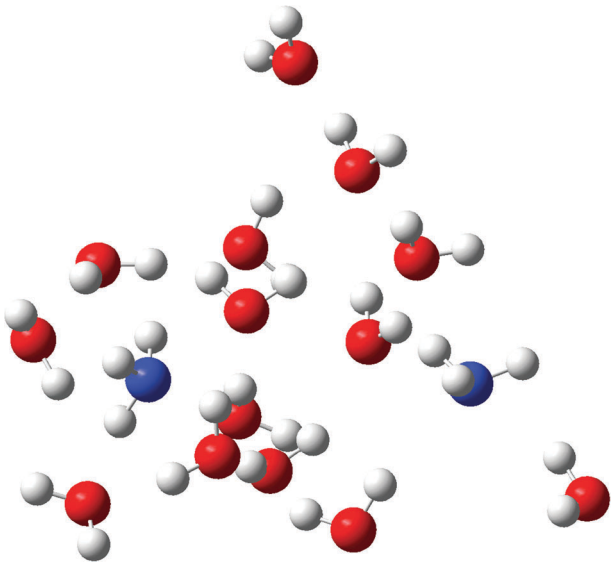

Fig. 2 The starting configuration used for a Monte Carlo simulation of $(NH_3)_2(H_2O)_{14}$.

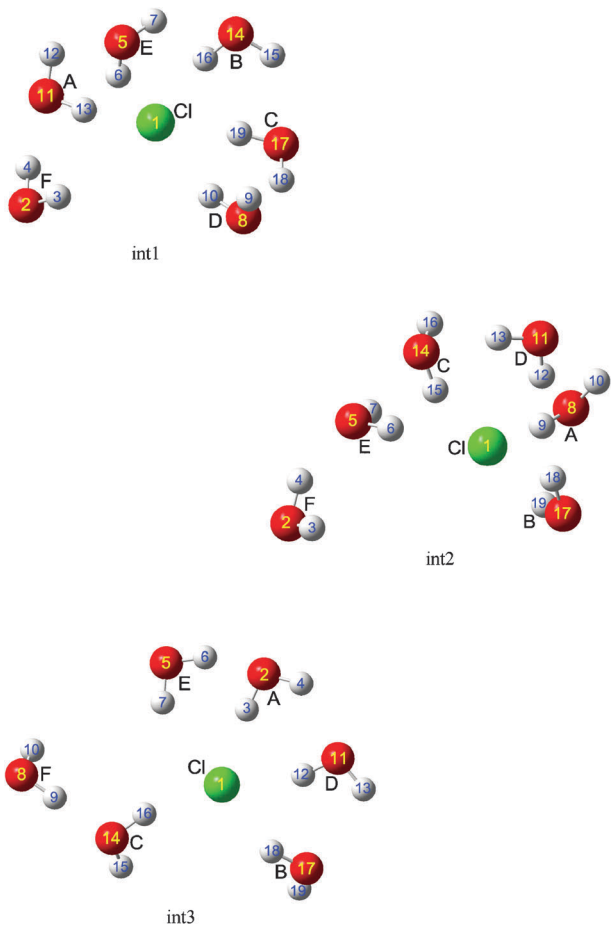

Fig. 3 $[Cl(H_2O)_6]^-$: Interior structures.

calculated dipole moment of that configuration using the M06-2X[52] density functional and the cc-pVTZ+[53] basis set. MB and EE-MB approximations of the full-system M06-2X/cc-pVTZ+ CHelpG[50] point charge distributions

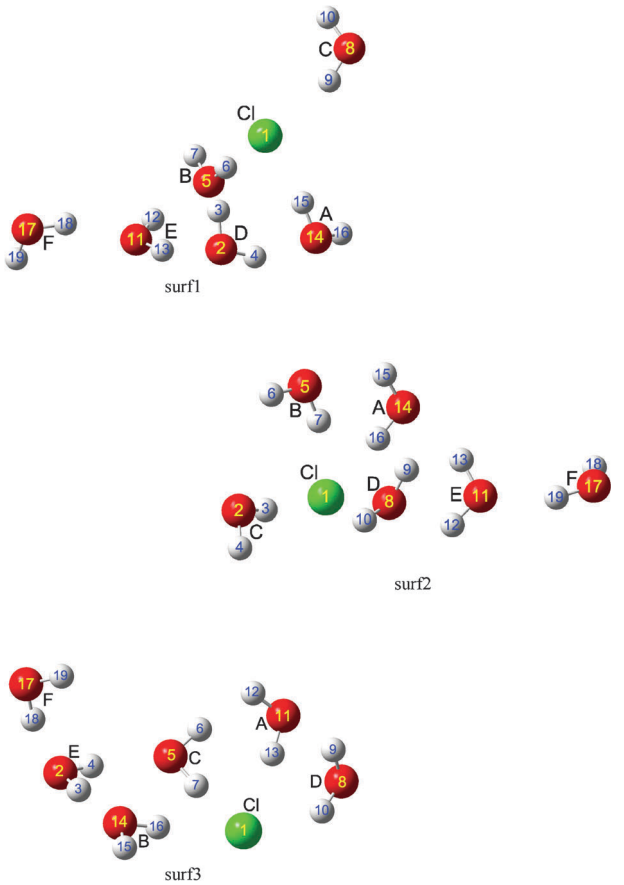

Fig. 4 $[Cl(H_2O)_6]^-$: Surface structures.

are compared below to the conventional M06-2X/cc-pVTZ+ CHelpG point charge distributions of the microhydrated chloride, hydrogen fluoride, and microhydrated hydrogen fluoride systems described in Sections 3.2 and 3.3.

3.1. Aqueous ammonia droplets

The configurations of both of the water–ammonia systems were obtained as follows: a short Monte Carlo (MC) simulation with the SPC/E water model[54] and the OPLS ammonia model[55] was run using the configurations shown in Fig. 1 and 2 as the starting points. The configurations generated during the first 44 steps of these simulations (including the starting configuration and rejected moves as well as accepted ones) were chosen as the test cases of the EE-MB method for dipoles. A conventional (a.k.a. "full-system") M06-2X/cc-pVTZ+ single-point energy and dipole calculation was run on each configuration of each system using *Gaussian 09*.[56] EE-PA and EE-3B dipole moment calculations were performed on each configuration by the M06-2X/cc-pVTZ+ method using a modified version of MBPAC 2007-2.[57] The M06-2X/cc-pVTZ+ method was used to calculate CM4M[58] charges of the equilibrium gas-phase water and ammonia monomers, and these CM4M charges were used as the embedding charges in the EE-MB calculations of the water–ammonia systems.

3.2. Microhydrated chloride ions

The water–chloride systems were also taken from an MC simulation at 250 K with the SPC/E water model and the CHARMM parameters for Cl^-.[59,60] Three of the water–chloride structures (int1, int2, and int3 in Fig. 3) were selected because they represent an "interior" chloride ion in liquid water. Each of the "interior" structures contains at least four hydrogen bonds to the chloride ion where the O–Cl distance is less than 3.7 Å and where the O–H–Cl angle is greater than 150°. The other three water–chloride structures (surf1, surf2, and surf3 in Fig. 4) represent a "surface" chloride ion in liquid water. Each "surface" structure contains fewer than two hydrogen bonds to the chloride ion, where in this case a hydrogen bond is defined as having an O–Cl distance of less than 4.7 Å and an O–H–Cl angle greater than 145°.

Conventional full-system M06-2X/cc-pVTZ+ dipole moment and CHelpG partial atomic charge calculations were run on each configuration of each water–chloride system using *Gaussian 09*. EE-1B, EE-PA, EE-3B, 1B, PA, and 3B dipole moment and CHelpG point charge calculations were also

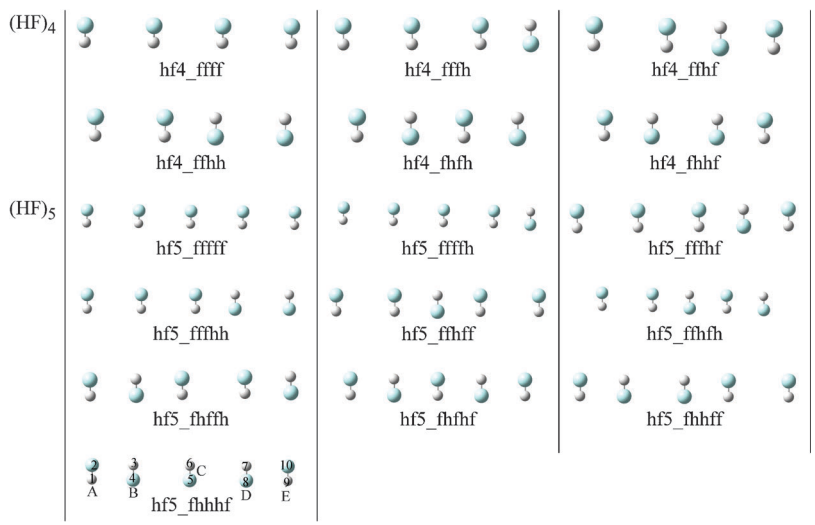

Fig. 5 $(HF)_m$ Clusters, m = 4–5.

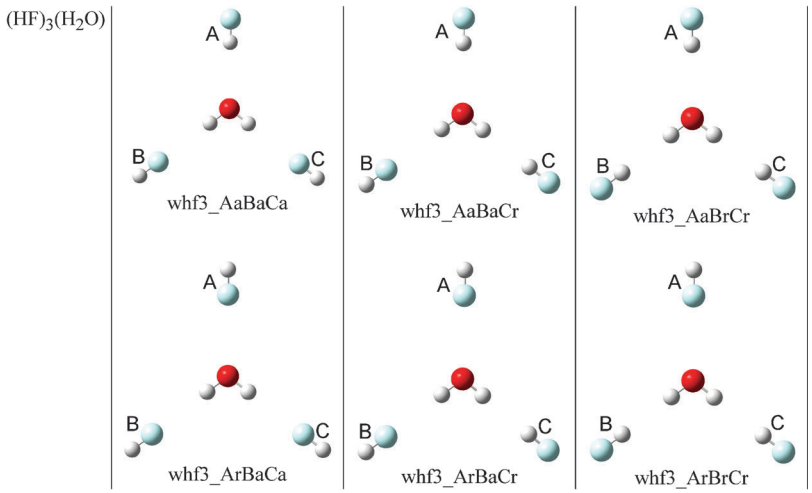

Fig. 6 $(HF)_3(H_2O)$ clusters.

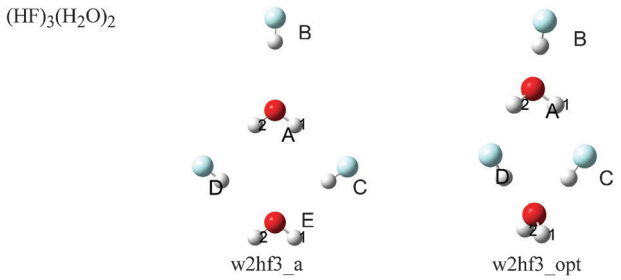

Fig. 7 $(HF)_3(H_2O)_2$ clusters.

performed for each configuration by the M06-2X/cc-pVTZ+ method using a modified version of MBPAC 2007-2.

For the MB and EE-MB calculations of the water–chloride structures, two different fragmentation schemes were used: in fragmentation Scheme 1 (S1), the chloride ion is considered a separate monomer and a point charge of -1 *e* is used as its embedding charge. In fragmentation Scheme 2 (S2), the chloride ion is paired with the water molecule containing the atom nearest to the chloride ion and the resulting $[Cl^{\bullet}H_2O]^-$ group is considered a single monomer. In Fig. 3 and 4, the water molecule that contains the atom closest to the chloride ion is always labeled "A", so that in S2 the fragmentation scheme is always ACl^-, B, C, D, E, F (yielding six monomers) whereas in S1 the fragmentation scheme is always Cl^-, A, B, C, D, E, F (yielding seven monomers). The CHelpG charges of the equilibrium gas-phase water molecule and the equilibrium gas-phase Cl^- (S1) or $[Cl(H_2O)]^-$ (S2) system were used as the embedding charges in the EE-MB calculations of the water–chloride clusters.

3.3. Hydrogen fluoride clusters and microhydrated hydrogen fluoride clusters

The $(HF)_m$ clusters (m = 4–5) shown in Fig. 5 were constructed in the following way: (1) A partial geometry optimization at M06-2X/cc-pV(T+d)Z+ was attempted on each cluster with the constraints that the molecules must lie in a plane and that the molecules retain their initial orientations relative to one another. (2) After performing step 1, it was noted that the geometries of structures that had two fluorine atoms next to each other never converged to a minimum; rather, the molecules with the adjacent fluorine atoms kept drifting farther away from one another. (3) Therefore, a second partial geometry optimization of each $(HF)_m$ cluster was performed with both of the constraints described in step 1 but also with the constraint that the distance between adjacent fluorine atoms be fixed at 3.4 Å. This distance was chosen after performing a quick scan of a slice of the $(HF)_2$ potential energy surface (PES) where the two fluorine atoms are forced to be adjacent to one another. Of the distances used in this scan, 3.4 Å is the distance where the dimer binding energy (the energy of the infinitely separated monomers minus the energy of the dimer) is closest to -1.0 kcal mol^{-1}.

The $(HF)_3(H_2O)$ clusters in Fig. 6 began with the M06-2X/cc-pV(T+d)Z+-optimized isolated gas-phase HF and H_2O molecules; the geometries of these molecules are given in Table 1. In each $(HF)_m(H_2O)$ cluster, the H_2O molecule (frozen in its optimized geometry) is placed in the center and m HF molecules (also frozen in their optimized geometries) are placed around it. HF molecule A is always placed so that the H, the F, and the O (of water) form a straight line perpendicular to the line through the two H atoms of water. If the F atom of A is closest to the O atom of water, then the F atom is placed R_{vdW}(F–O) Å away from the O of water; otherwise, the H atom of A is placed R_{vdW}(H–O) Å away from the O of water, where R_{vdW}(F–O) and R_{vdW}(H–O) are given in Table 2. HF molecules B and C are always placed so that the H, F, the nearest H of water and the O of water form a straight line. If the F atom of B or C is closet to the nearest H atom of water, then F is placed R_{vdW}(F–H) Å away from the nearest

Table 1 Geometric parameters of the M06-2X/cc-pV(T+d)Z+ optimized gas-phase water and hydrogen fluoride molecules

Molecule	Parameter	Value
HF	R_{HF}	0.9194 Å
H_2O	R_{OH}	0.9592 Å
H_2O	θ_{HOH}	105.2°

Table 2 Van der Waals radii of atoms and combined van der Waals radii of atom pairs

Radius	Value (Å)	Ref.
R_{vdW}(F)	1.47	62
R_{vdW}(H)	1.10	62
R_{vdW}(O)	1.52	62
R_{vdW}(F–F)	2.94	R_{vdW}(F) + R_{vdW}(F)
R_{vdW}(F–H)	2.57	R_{vdW}(F) + R_{vdW}(H)
R_{vdW}(F–O)	2.99	R_{vdW}(F) + R_{vdW}(O)
R_{vdW}(H–H)	2.20	R_{vdW}(H) + R_{vdW}(H)
R_{vdW}(H–O)	2.62	R_{vdW}(H) + R_{vdW}(O)
R_{vdW}(O–O)	3.04	R_{vdW}(O) + R_{vdW}(O)

H atom of water. Otherwise, the H atom of B or C is placed R_{vdW}(H–H) Å away from the nearest H atom of water.

The $(HF)_3(H_2O)_2$ clusters are shown in Fig. 7. The structure called w2hf3_a was constructed in exactly the same way as structure whf3_AaBaCa of Fig. 6: the HF and H_2O molecules were kept in their M06-2X/cc-pV(T + d)Z + -optimized isolated gas-phase geometries and were placed at the appropriate R_{vdW} distance away from one another. In w2hf3_a, all molecules lie in the same plane. The w2hf3_a structure was then used as a starting point for an M06-2X/cc-pV(T + d)Z + geometry optimization of the entire cluster. The result of this geometry optimization is the structure named w2hf3_opt. This structure (w2hf3_opt) is a minimum-energy structure on the M06-2X/cc-pV(T + d)Z + $(HF)_3(H_2O)_2$ PES; it contains no imaginary frequencies.

The dipole moment and CHelpG partial charge distribution of each hydrogen fluoride and microhydrated hydrogen fluoride cluster were calculated conventionally and within the EE-1B, EE-PA, EE-3B, 1B, PA, and 3B approximations using the M06-2X density functional and the cc-pV(T + d)Z + basis set. For the EE-MB calculations, the M06-2X/cc-pV(T + d)Z + CHelpG charges of the equilibrium gas-phase hydrogen fluoride or water molecule were used as the embedding charges.

4. Results

In order to assess the accuracy of the EE-MB and MB approximations for the calculation of full-system dipole moments, partial atomic charges, and partial charge transfer between fragments, the EE-MB and MB approximations of these values are compared to the conventionally calculated values at the same level of electronic structure theory (which, for all results shown, is the M06-2X density functional with the cc-pVTZ + basis set). The errors summarized in the tables described in this section are calculated as the EE-MB (or MB) approximate values minus the conventionally calculated values.

Next we describe the results in Tables 2–6; the results will be discussed in Section 5. First we note that Tables 3–5 each include the number, n, of configurations over which each average is taken and the number, N, of monomers into which each configuration of each system is divided.

4.1. Full-system dipole moments

For simplicity, throughout the remainder of this paper we will refer to the magnitude of the dipole moment as the dipole moment.

Table 3 Average cluster dipole moments and mean unsigned errors in cluster dipole moments (in debyes)[a]

				$\langle\lvert\mu^{MB} - \mu\rvert\rangle$		
Ensemble	n	N	$\langle\lvert\mu\rvert\rangle$	1B	PA	3B
EE-MB						
$NH_3(H_2O)_{11}$	44	12	13.89	1.02	0.10	0.08
$(NH_3)_2(H_2O)_{14}$	44	16	17.97	0.50	0.18	0.08
$[Cl(H_2O)_6]^-$	6	7	6.36	0.42	0.10	0.04
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	6.36	0.33	0.07	0.04
$(HF)_{4-5}$	16	4–5	3.198	0.003	0.001	0.000
$(HF)_3(H_2O)_{1-2}$	8	4–5	5.262	0.164	0.009	0.003
MB						
$[Cl(H_2O)_6]^-$	6	7	6.36	1.60	0.17	0.08
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	6.36	1.13	0.17	0.06
$(HF)_{4-5}$	16	4–5	3.198	0.125	0.002	0.000
$(HF)_3(H_2O)_{1-2}$	8	4–5	5.262	0.613	0.023	0.005

[a] n is the number of configurations in the ensemble, N is the number of fragments, μ is the dipole moment from a conventional calculation on the entire system, and μ^{MB} is the dipole moment calculated by the EE-MB or MB method.

Table 4 Mean unsigned errors (MUEs) in partial atomic charges (in atomic units)[a]

				MUE		
Ensemble	n	N	$\langle\lvert q\rvert\rangle$	1B	PA	3B
EE-MB						
$[Cl(H_2O)_6]^-$	6	7	0.56	0.06	0.06	0.04
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	0.56	0.06	0.04	0.03
$(HF)_{4-5}$	16[b]	4–5[c]	0.438	0.004	0.003	0.002
$(HF)_3(H_2O)_{1-2}$	8[d]	4–5[e]	0.452	0.033	0.023	0.005
MB						
$[Cl(H_2O)_6]^-$	6	7	0.56	0.10	0.04	0.04
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	0.56	0.09	0.03	0.03
$(HF)_{4-5}$	16	4–5	0.438	0.014	0.002	0.002
$(HF)_3(H_2O)_{1-2}$	8	4–5	0.452	0.036	0.014	0.009

[a] n is the number of configurations in the ensemble and N is the number of fragments. [b] The average is over 148 partial atomic charges. [c] Six cases with $N = 4$ and ten with $N = 5$. [d] The average is over 78 partial atomic charges. [e] Six cases with $N = 4$ and two with $N = 5$.

Table 5 Average fragment charges and mean unsigned errors in fragment charges (in atomic units)[a]

				$\langle\lvert Q^{MB} - Q\rvert\rangle$		
Ensemble	n	N	$\langle\lvert Q\rvert\rangle$	1B	PA	3B
EE-MB						
$[Cl(H_2O)_6]^-$	6	7	0.15	0.09	0.04	0.03
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	0.15	0.08	0.03	0.03
$(HF)_5$[b] and $(HF)_3(H_2O)_2$	3	5	0.036	0.036	0.023	0.006
MB						
$[Cl(H_2O)_6]^-$	6	7	0.15	0.09	0.03	0.02
$[Cl^{\bullet}H_2O(H_2O)_5]^-$	6	6	0.15	0.07	0.03	0.02
$(HF)_5$[b] and $(HF)_3(H_2O)_2$	3	5	0.036	0.036	0.010	0.012

[a] n is the number of configurations in the ensemble, N is the number of fragments, Q is the net fragment charge from a conventional calculation on the entire system, and Q^{MB} is the net fragment charge calculated by the EE-MB or MB method. [b] Only one configuration of $(HF)_5$ is included in the calculation of these averages: hf5_fhhhf (see Fig. 5).

Table 6 Average[a] water monomer dipole moment (in debyes) calculated from CHelpG charges in $[Cl(H_2O)_6]^-$ clusters[b]

	Conventional	1B	PA	3B	EE-1B	EE-PA	EE-3B
int1	3.030	2.117	3.294	2.747	2.563	2.896	2.701
int2	2.897	2.102	3.266	2.744	2.525	2.955	2.740
int3	2.862	2.085	3.269	2.772	2.541	2.994	2.846
surf1	2.912	2.000	3.152	2.803	2.475	3.008	2.764
surf2	2.865	2.014	3.224	2.631	2.459	3.058	2.726
surf3	3.052	2.113	3.298	2.849	2.641	3.284	2.917
MUE[c]		0.864	0.314	0.178	0.402	0.141	0.154

[a] Averages are taken over the six water monomer dipole moments in each cluster. [b] Fragmentation scheme 1 is used in all MB and EE-MB calculations shown in this table. [c] MUE = mean unsigned error in the average water monomer dipole moment from MB or EE-MB CHelpG charge distributions relative to the average water monomer dipole moment from the conventional CHelpG charge distribution using the M06-2X/cc-pV(T + d)Z + method.

Table 3 shows the mean unsigned errors (MUEs) of the EE-MB calculations of cluster dipole moments (and, in some cases, MB calculations of these dipole moments) over all of the configurations of each type of system studied in this work. Table 3 also shows, to help put these values in perspective, the average, $\langle\mu\rangle$, of the conventionally calculated cluster dipole moments over all configurations of each type of system; note that this is the same as $\langle|\mu|\rangle$.

4.2. Partial atomic charges

Table 4 shows the MUEs in the EE-MB and MB approximations of the CHelpG partial atomic charge distributions of the microhydrated chloride, hydrogen fluoride, and microhydrated hydrogen fluoride clusters used in this work. Table 4 also shows the average, $\langle|q|\rangle$, of the absolute values of the conventionally calculated atomic charges over all atoms in all configurations of each type of system.

4.3. Partial charge transfer

Table 5 shows the MUEs in the EE-MB and MB approximations of the net fragment charges, Q, of the microhydrated chloride, hydrogen fluoride, and microhydrated hydrogen fluoride clusters considered in this work. The net charge on a fragment (in this case, on a monomer in the cluster) is calculated as the sum of the CHelpG charges of the atoms that constitute the given fragment. Additionally, Table 5 shows the average, $\langle|Q|\rangle$, of the absolute values of the conventionally calculated net fragment charges over all configurations of each type of system.

4.4. Molecular dipole moments

The full-system dipole moment is well defined, but fragment dipole moments are model quantities. Despite their model character, molecular dipole moments of molecules in clusters and liquids are widely used for interpretative purposes, and here we show that we can compute them by the EE-MB method. Table 6 shows the average water fragment dipole moments for the $Cl(H_2O)^-$ clusters and also the average deviation from those calculated by CHelpG analysis of full-system density functional calculations.

5. Discussion

5.1. Dipole moments

Table 3 shows that, as expected, the EE-MB full-system dipole moment becomes more accurate as the order of the EE-MB approximation increases. However, even for larger clusters, the EE-PA approximation is able to get the full-system dipole moment quantitatively correct. This is good news because of the more favorable scaling of the cost of the EE-PA approximation with system size than that of the EE-3B approximation. The mean unsigned percentage error (MUPE) over all configurations of all aqueous ammonia droplets is 5.0% in the EE-1B approximation, 0.9% in the EE-PA approximation, and 0.5% in the EE-3B approximation. Table 3 also shows that electrostatic embedding significantly improves the accuracy of the one-body approximation and slightly improves the accuracies of the two- and three-body approximations.

5.2. Partial atomic charges

The $[Cl(H_2O)_6]^-$ system provides an interesting challenge to many-body methods. The charge on Cl^- is of course -1.0 atomic units (a.u.), and the average partial atomic charge and standard deviation on Cl in $Cl^\bullet(H_2O)^-$ in the six configurations (where we always consider the H_2O that contains the atom closest to Cl^- in the overall cluster) is -0.91 ± 0.01 a.u., and when the $Cl^\bullet(H_2O)^-$ fragment is embedded in fifteen embedding charges representing the rest of the cluster, the average charge and its standard deviation on Cl become -0.94 ± 0.04 a.u. However the average charge (and standard deviation) of Cl^- in the entire cluster is -0.70 ± 0.07 a.u. We find that the EE-PA and EE-3B methods account for the reduction of this partial atomic charge from -1.0, -0.91, or -0.94 to -0.70 a.u. quite well. By definition, the EE-1B and 1B approximations cannot predict this reduction in charge because these methods do not allow charge transfer between monomers. However, the EE-PA and EE-3B approximations are able to account for this reduction in the charge assigned to Cl: using the simpler fragmentation scheme where the Cl^- ion is considered a monomer (S1; see Section 3.2), the average charge (and standard deviation) on Cl is -0.78 ± 0.05 a.u. using the EE-PA approximation and is -0.75 ± 0.05 a.u. using the EE-3B approximation. Using the fragmentation scheme where the Cl^- is paired with the closest water molecule to form a single monomer (S2), both the EE-PA and EE-3B approximations do a little better: the average charge (and standard deviation) on Cl is -0.75 ± 0.06 a.u. using the EE-PA approximation and is -0.74 ± 0.07 a.u. using the EE-3B approximation.

Table 4 shows trends in partial atomic charges similar to those noticed with the dipole moments: again, the accuracy of the EE-MB or MB approximation generally increases with the order to which the MB expansion is taken and, again, in most cases (except for the microhydrated hydrogen fluoride clusters) the PA-level approximation is already capable of yielding quantitative accuracy in the partial charge distribution relative to the conventionally calculated CHelpG charge distribution of the entire cluster. However, unlike the trend seen in the dipole moment calculations, the accuracy of the many-body

approximations for the prediction of partial charge distributions depends strongly on whether or not electrostatic embedding is included only in the one-body case; in the pairwise additive and three-body cases, the accuracy of the MB approximation is similar to that of the EE-MB approximation of the same order.

An interesting point to consider is that as the size of a cluster increases, more and more atoms become "buried" within the cluster. Charge analysis methods that are based on finding a set of partial atomic charges that most closely reproduces the electrostatic potential due to the charge density of the entire system (like CHelpG) often have difficulty assigning stable physical charges to buried atoms. Using the EE-PA or the EE-3B method to calculate partial atomic charge distributions of large clusters could be a way to obtain stable physical charges on buried atoms without requiring the imposition of empirical constraints on the magnitudes of the charges assigned to various atoms.

5.3. Partial charge transfer

Partial charge transfer is ubiquitous in liquid-phase systems, nanoparticles, and clusters, but it is notoriously hard to treat by fragment methods.

The net charge on a fragment (calculated as the sum of the partial charges assigned to the atoms that compose the fragment) is an indication of how much charge transfer has occurred to or from that fragment when it moves from being in an isolated state to the state of being a member of a given cluster: for example, an isolated gas-phase water molecule has a net charge of zero a.u., but if it is placed in a cluster containing a negatively-charged ion, its net charge might be −0.20 a.u., which shows that some of the electron density from the anion has been transferred to that water molecule, and molecules with net charges in the range of ±0.2 a.u. were also observed in simulations for liquid water which allowed for inter- and intramolecular charge equilibration.[61] The MUEs in net fragment charges shown in Table 5 are therefore a measure of how well the EE-MB approximations are able to capture the amount charge transfer occurring between the molecules in the clusters studied.

The one-body approximations are incapable of predicting charge transfer between molecules, so for systems containing all neutral fragments the EE-1B and 1B approximations necessarily have errors of 100% in the calculation of net monomer charges. Moving to the PA or EE-PA approximation reduces the error in net monomer charges to between 20 and 30% in most cases. Using the 3B or EE-3B approximation generally reduces the error in net monomer charges to between 13 and 20%. As with the partial charge calculations, a strong dependence on whether or not electrostatic embedding is used is not seen in the accuracy of the MB net monomer charge calculations.

Although neither the EE-MB nor the MB approximations were capable of quantitatively accurate predictions of net monomer charges for the cases chosen in this work, the PA- and 3B-level approximations were able to qualitatively describe the amount of charge transfer between fragments most of the time.

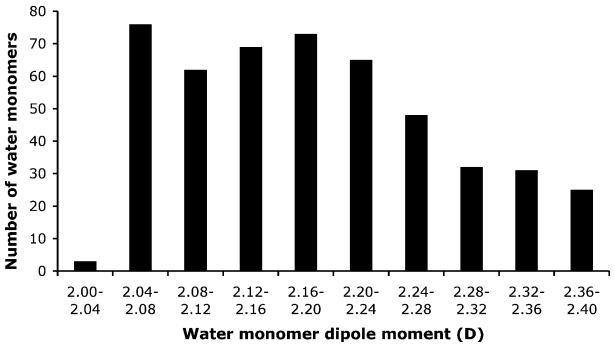

Fig. 8 Distribution of EE-1B water monomer dipole moments in forty-four configurations of $(NH_3)(H_2O)_{11}$.

5.4. Molecular dipole moments

The mean, median, minimum, and maximum molecular dipole moments for water are 2.19, 2.18, 2.03, and 2.34 D respectively, in the $(NH_3)(H_2O)_{11}$ clusters, 2.51, 2.53, 2.12, and 2.76 D, respectively, in the $[Cl(H_2O)_6]^-$ clusters, and 2.47, 2.47, 2.20, and 2.72 D, respectively, in the $[Cl^{\bullet}(H_2O)(H_2O)_5]^-$ clusters. A histogram showing the distribution of the EE-1B water dipole moments over all forty-four $(NH_3)(H_2O)_{11}$clusters is shown in Fig. 8.

As a further illustration of the application of the EE-1B method of the computation of molecular dipole moments, we note that the mean molecular dipole moment of individual water molecules (1B approximation) is 2.07 D in $[Cl(H_2O)_6]^-$ but these values change when one considers the clusters. In fact, if the molecular dipole moments are calculated from the CHelpG charges of the full-system calculations, the mean water monomer dipole moment changes by 46% (the new mean is 2.93 D). Table 6 shows that the EE-PA and EE-3B approximations give similar accuracy for the average molecular dipole moments in the chloride–water clusters, and they both agree better than the 3B approximation with the average molecular dipole moments calculated from full-system density functional calculations. The performance of the EE-PA method is very encouraging.

6. Summary

Our first set of tests using the EE-MB method for predicting charge distributions was concerned with full-system dipole moments. We compared the EE-MB results to the one-body (1B) approximation, in which one simply carries out vectorial addition of the dipole moments of the noninteracting constituents, and to the EE-1B approximation, in which one vectorially adds the dipole moments of the electrostatically embedded monomers. Both of these approximations led to large errors, but the MUEs of EE-1B were reduced by factors of 3 to 18 by the EE-PA approximation and by factors of 6 to 55 by the EE-3B approximation, and the MUEs of the 1B approximation were reduced by factors of 7 to 63 by the PA approximation and by factors of 19 to 123 by the 3B approximation.

The second set of tests using the EE-MB method for predicting charge distributions was concerned with full-system partial atomic charge distributions. The 1B approximation of a particular atom's partial charge is simply that atom's partial charge

when it is in an isolated monomer; the EE-1B approximation of an atom's partial charge is that atom's partial charge when its monomer is surrounded by embedding charges. On average, electrostatic embedding reduces the MUE of the 1B approximation by a factor of two. Including the pairwise additive terms usually reduces the MUEs of the 1B approximation by factors of 3 to 7, and using the 3B approximation also reduces the MUEs by factors of 3 to 7. The EE-PA approximation does not improve upon the EE-1B approximation by as large an amount for this quantity: the MUEs are reduced by factors of 1.0 to 1.5 in going from the EE-1B to the EE-PA approximation. A more significant improvement is seen in going from the EE-1B approximation to the EE-3B approximation: in this case, the MUEs are reduced by factors of 2 to 7.

A more difficult set of tests using the EE-MB method for predicting charge distributions was concerned with the amount of intermolecular charge transfer. The EE-1B and 1B approximations (and many other linear-scaling fragment based methods) do not allow any charge transfer to occur between monomers at all, so even being able to get a qualitative picture of the amount of intermolecular charge transfer occurring in a system using the EE-PA or EE-3B approximations is an improvement. We find that in going from the EE-1B approximation to the EE-PA approximation the MUEs are reduced by factors of 2 to 3 and in going from the EE-1B approximation to the EE-3B approximation the MUEs are reduced by factors of 3 to 6. Similarly, in going from 1B to PA the errors are reduced by factors of 2 to 4 and in going from 1B to 3B the errors are reduced by factors of 3 to 5.

Finally we showed that the EE-MB approximation can also be used to calculate dipole moments of fragments in the cluster.

We conclude that the EE-PA and EE-3B approximations are capable of yielding accurate descriptions of full-system charge distributions through dipole moment and atomic partial charge distribution calculations. The EE-PA and EE-3B approximations are also capable of qualitatively describing the amount of intermolecular charge transfer occurring in a system, which is a difficult task for fragment-based methods.

Acknowledgements

This work was supported in part by the National Science Foundation under Grant Nos. CHE0956776 and CHE1051396.

References

1 N. Ferre and X. Assfeld, *THEOCHEM*, 2003, **632**, 83.
2 T. E. Exner and P. G. Mezey, *Phys. Chem. Chem. Phys.*, 2005, **7**, 4061.
3 D. G. Fedorov and K. Kitaura, *J. Phys. Chem. A*, 2007, **111**, 6904.
4 W. Hua, T. Fang, W. Li, J. Yu and S. Li, *J. Phys. Chem. A*, 2008, **112**, 10864.
5 W. Xie, L. Song, D. G. Truhlar and J. Gao, *J. Chem. Phys.*, 2008, **128**, 234108.
6 M. S. Gordon, J. M. Mullin, S. R. Pruitt, L. B. Roskop, L. V. Slipchenko and J. A. Boatz, *J. Phys. Chem. B*, 2009, **113**, 9646.
7 E. E. Dahlke and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 46.
8 Y. Tong, Y. Mei, Y. L. Li, C. G. Ji and J. Z. H. Zhang, *J. Am. Chem. Soc.*, 2010, **132**, 5137.
9 R. Mata, *Mol. Phys.*, 2010, **108**, 381.
10 S. Hirata, *Mol. Phys.*, 2010, **108**, 3113.
11 H. Stoll and H. Preuss, *Theor. Chem. Acc.*, 1977, **46**, 12.
12 W. Yang, *Phys. Rev. Lett.*, 1991, **66**, 1438.
13 V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret and G. C. Ferenczy, *J. Comput. Chem.*, 1994, **15**, 269.
14 J. Gao, *J. Phys. Chem. B*, 1997, **101**, 657.
15 J. M. Pedulla, K. Kim and K. D. Jordan, *Chem. Phys. Lett.*, 1998, **291**, 78.
16 K. Kitaura, E. Ikeo, T. Asada, T. Nakano and M. Uebayasi, *Chem. Phys. Lett.*, 1999, **313**, 701.
17 C. Amovilli, I. Cacelli, S. Campanile and G. Prampolini, *J. Chem. Phys.*, 2002, **117**, 3003.
18 D. W. Zhang, Y. Xiang and J. Z. H. Zhang, *J. Phys. Chem. B*, 2003, **107**, 12039.
19 S.-I. Sugiki, N. Kurita, Y. Sengoku and H. Sekino, *Chem. Phys. Lett.*, 2003, **382**, 611.
20 S. Li, W. Li and T. Fang, *J. Am. Chem. Soc.*, 2005, **127**, 7215.
21 S. Hirata, M. Valiev, M. Dupuis, S. S. Xantheas, S. Sugiki and H. Sekino, *Mol. Phys.*, 2005, **103**, 2255.
22 N. Jiang, J. Ma and Y. Jiang, *J. Chem. Phys.*, 2006, **124**, 114112.
23 M. A. Collins and V. A. Deev, *J. Chem. Phys.*, 2006, **125**, 104104.
24 R. P. A. Bettens and A. M. Lee, *J. Phys. Chem. A*, 2006, **110**, 8777.
25 E. E. Dahlke and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 1342.
26 E. Suárez, N. Díaz and D. Suarez, *J. Chem. Theory Comput.*, 2009, **5**, 1667.
27 P. Söderhjelm, F. Aquilante and U. Ryde, *J. Phys. Chem. B*, 2009, **113**, 11085.
28 J. O. B. Tempkin, H. R. Leverentz, B. Wang and D. G. Truhlar, *J. Phys. Chem. Lett.*, 2011, **2**, 2141.
29 L. D. Jacobson and J. M. Herbert, *J. Chem. Phys.*, 2011, **134**, 094118.
30 N. J. Mayhall and K. Raghavachari, *J. Chem. Theory Comput.*, 2011, **7**, 1336.
31 D. M. Bates, J. R. Smith, T. Janowski and G. S. Tschumper, *J. Chem. Phys.*, 2011, **135**, 044123.
32 W. Wen and G. J. O. Beran, *J. Chem. Theory Comput.*, 2011, **7**, 3733.
33 M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko, *Chem. Rev.*, 2012, **112**, 632.
34 E. D. Speetzen, H. R. Leverentz, H. Lin and D. G. Truhlar, in *Accurate Condensed-Phase Electronic Structure Theory*, ed. F. R. Manby, CRC Press, Boca Raton, FL, 2011, p. 105.
35 M. Isegawa, J. Gao and D. G. Truhlar, *J. Chem. Phys.*, 2011, **135**, 084107.
36 N. Marzari and D. Vanderbilt, *Phys. Rev. B: Condens. Matter*, 1997, **56**, 12847.
37 P. L. Silvestrelli and M. Parrinello, *J. Chem. Phys.*, 1999, **111**, 3572.
38 G. Berghoid, C. J. Mundy, A. H. Romero, J. Hutter and M. Parrinello, *Phys. Rev. B: Condens. Matter*, 2000, **61**, 10040.
39 M. P. Gaigeot and M. Sprik, *J. Phys. Chem. B*, 2003, **107**, 10344.
40 M. Sharma, R. Resta and R. Car, *Phys. Rev. Lett.*, 2005, **95**, 187401.
41 C. Dellago and M. M. Naor, *Comput. Phys. Commun.*, 2005, **169**, 36.
42 M. Sharma, R. Resta and R. Car, *Phys. Rev. Lett.*, 2007, **98**, 247401.
43 M. J. Mcgrath, J. I. Siepmann, I. F. W. Kuo and C. J. Mundy, *Mol. Phys.*, 2007, **105**, 1411.
44 A. Ferratti, A. Calzolari, B. Bonferroni and R. Di Felice, *J. Phys.: Condens. Matter*, 2007, **19**, 036215.
45 D. Kang, J. Dai and J. Yuan, *J. Chem. Phys.*, 2011, **135**, 024505.
46 B. Kirchner, P. G. Di Dio and J. Hutter, *Top. Curr. Chem.*, 2012, **307**, 109.
47 D.-L. Chen, A. C. Stern, B. Space and J. K. Johnson, *J. Phys. Chem. A*, 2010, **114**, 10225.
48 T. Watanabe, T. A. Manz and D. S. Sholl, *J. Phys. Chem. C*, 2011, **115**, 4824.
49 R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833.
50 C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361.

51 J. P. Foster and F. Weinhold, *J. Am. Chem. Soc*., 1980, **102**, 7211.
52 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc*., 2007, **120**, 215.
53 E. Papajak, H. R. Leverentz, J. Zheng and D. G. Truhlar, *J. Chem. Theory Comput*., 2009, **5**, 1197.
54 H. J. C. Berendsen, J. R. Grigera and T. P Straatsma, *J. Phys. Chem*., 1987, **91**, 6269.
55 R. C. Rizzo and W. L. Jorgensen, *J. Am. Chem. Soc*., 1999, **121**, 4827.
56 M. J. Frisch, *et al*., *GAUSSIAN 09 (Revision A.2)*, Gaussian, Inc., Wallingford, CT, 2009.
57 E. E. Dahlke and D. G. Truhlar, *MBPAC 2007-2*, University of Minnesota, Minneapolis, MN, 2007.
58 R. M. Olson, A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput*., 2007, **3**, 2046.
59 D. Beglov and B. Roux, *J. Chem. Phys*., 1994, **100**, 9050.
60 B. Roux, *Biophys. J*., 1996, **71**, 3177.
61 B. Chen, J. J. Potoff and J. I. Siepmann, *J. Phys. Chem. B*, 2000, **104**, 2378.
62 M. Mantina, R. Valero and D. G. Truhlar, *J. Phys. Chem. A*, 2009, **113**, 5806.

