

## Modeling Charge Penetration Effects in Water–Water Interactions

Tae Hoon Choi

Department of Chemical Engineering Education, Chungnam National University, Daejeon 305-764, Korea  
E-mail: thchoi@cnu.ac.kr

Received May 16, 2014, Accepted June 4, 2014

This report introduces Gaussian electrostatic models (GEMs) to account for charge penetration effects in water–water interactions, allowing electrostatic interactions to be accurately described. Three different Gaussian electrostatic models, GEM-3S, GEM-5S, and GEM-6S are designed with *s*-type Gaussian functions. The coefficients and exponents of the Gaussian functions are optimized using the electrostatic potential (ESP) fitting procedure based on that of the MP2/aug-cc-pVTZ method. The electrostatic energies of ten different water dimers that were calculated with GEM-6S agree well with the results of symmetry-adapted perturbation theory (SAPT), indicating that this designed model can be effectively applied to future water models.

**Key Words** : Gaussian electrostatic models, Charge penetration, Water dimer

### Introduction

Over the last few years, a large number of water models have been continuously introduced due to their fundamental importance for implementing Monte Carlo and molecular dynamics simulations in chemistry and biology.<sup>1–6</sup> The increasing accuracy with respect to the treatment of electrostatic interactions has emerged as a central point in the development of these potentials. Although several popular water models, such as TIP4P,<sup>1</sup> DC (Dang-Chang),<sup>2</sup> TTM (Thole-Type Model),<sup>3</sup> and DPP (Distributed Point Polarizable),<sup>5</sup> successfully described the relative energies of water clusters, none of them considered charge penetration effects<sup>7</sup> because these models used a classical point-charge representation of electrostatics.

The intermolecular electrostatic contribution in the short range is considered by the overlap of the charge densities when two molecules are in close proximity. The electrostatic interaction between these two species is no longer well represented by methods such as Stone's distributed multipolar analysis (DMA)<sup>7</sup> as the nuclei on one molecule are no longer shielded by its own electron density, thus experience a greater attraction for the electron density associated with the other species. The energy resulting from this increased attraction is referred to as charge penetration.

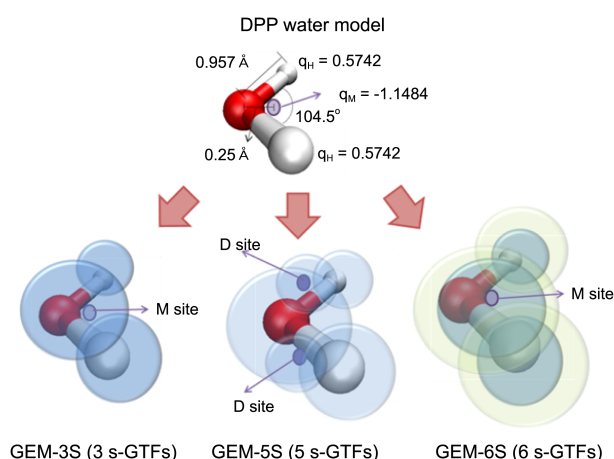
Recently Kumar *et al.* introduced a second generation distributed point polarizable water model (DPP2) by applying explicit terms for charge penetration.<sup>8</sup> This model has been successfully improved for describing interaction energies in small and large water clusters by applying the simple damping scheme for point-charge electrostatics. Introducing a screening function as a pre-factor to the Coulomb potential has been a popular method used to address charge penetration.<sup>9–12</sup> This method does have the advantage of including the short-range electrostatic penetration effect at a very low cost, but it still does not account for all the nuclear-electron interactions. As an alternative method, the Gaussian electrostatic model (GEM) aims to fit electron densities to Gaussian

auxiliary basis sets (ABSs) centered on specific molecular sites, thus can accurately generate the Coulomb interaction energies and explicitly considers nuclear charge and electron densities. This method has been developed and studied in depth by Cisneros and co-workers.<sup>13–17</sup> As a recent work, they introduced distributed multipoles based on the GEM (GEMDM) applied in the AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) force field to improve the treatment of electrostatics.<sup>17</sup> The *ab initio* calculated electron density can be fit with a number of basis functions of differing angular momentum, but ABSs are usually restricted to *s*-type Gaussian functions for the computational efficiency.

In the present study, Gaussian electrostatics models on specific molecular sites have been introduced to describe the charge penetration effects. Three different GEMs have been suggested based on the DPP and SAPT-5S water models:<sup>18</sup> GEM-3S has one *s*-Gaussian type function (*s*-GTF) at each of three sites, GEM-5S has one *s*-GTF located at each of five sites, and GEM-6S has two *s*-GTFs at each of three sites. The coefficients and exponents of the Gaussian functions of the ABSs are fitted to the MP2/aug-ccpVTZ electrostatic potential. The validities of the three models have been examined by calculating the intermolecular Coulomb interactions for several stationary points on the water dimer potential energy surface. These results are compared with those of symmetry-adapted perturbation theory (SAPT), absolutely localized molecular orbitals energy decomposition analysis.<sup>8,26</sup>

### Methodology

**Models.** Figure 1 compares the DPP water model and the three different GEMs considered in this work. GEM-3S locates two *s*-GTFs at the hydrogens and one at the M site, which is displaced along the HOH bisector. Cisneros *et al.* had tested a similar three-site model, where the GTFs were



**Figure 1.** Gaussian electrostatic models: GEM-3S, GEM-5S, and GEM-6S have 3, 5, and 6 s-types.

located at the oxygen and two hydrogen sites.<sup>16</sup> This resulted in a water monomer dipole moment of about 1.9–2.0 Debye, which is much higher than the experimental result (1.855 Debye). Thus, GEM-3S suggested in this study can be applicable as a better descriptor of the dipole moment. GEM-5S employs five sites, one on each atom and two D sites located above and below the HOH plane. The geometry of GEM-5S is adapted from the SAPT-5S model,<sup>18</sup> which is a rigid model based on the vibrational averaged geometry. The positions of the D sites in GEM-5S have been optimized to better fit the electrostatic energy. GEM-6S places two s-GTFs at each of the three sites of GEM-3S for a total of six s-GTFs. Nuclear point charges of +6 and +1 are placed on M and each H, respectively, for both GEM-3S and GEM-6S, whereas those of GEM-5S are located on O and each H.

The electrostatic interaction energy for the classical model using the point-charge (PC) representation is given by

$$E_{PC}^{elec.} = \sum_i \sum_j \frac{Q_{A_i} Q_{B_j}}{r_{A_i B_j}} \quad (1)$$

where  $Q_{A_i}$  and  $Q_{B_j}$  represent the point charges at  $i$  and  $j$  sites of two different monomers A and B, respectively. For the GEMs, a more detailed description of the intermolecular Coulomb energy is required and nucleus–nucleus (N–N), nucleus–electron (N–e), and electron–electron (e–e) contributions must be considered.

$$E_{Intermol}^{elec.} = E_{N-N} + E_{N-e} + E_{e-e}. \quad (2)$$

The following expression gives the intermolecular Coulomb interactions in the GEM model.

$$\begin{aligned} E_{GEM}^{elec.} = & \sum_i \sum_j \frac{Z_{A_i} Z_{B_j}}{r_{A_i B_j}} + \sum_i \int \frac{Z_{A_i} \rho^B(r_2)}{r_{A_i 2}} dr_2 \\ & + \sum_j \int \frac{Z_{B_j} \rho^A(r_1)}{r_{B_j 1}} dr_1 \\ & + \iint \frac{\rho^A(r_1) \rho^B(r_2)}{r_{12}} dr_1 dr_2 \end{aligned} \quad (3)$$

The first term represents the nucleus–nucleus interaction, the second and third terms are the nucleus–electron contributions, and the last term is the electron–electron interaction.  $Z_{A_i}$  and  $Z_{B_j}$  represent the nuclear charges of atom  $i$  on molecule A and atom  $j$  on molecule B, respectively. Only valence nucleus charges are considered, *i.e.*, +6 and +1 are used for O and H, respectively.  $\rho^A$  and  $\rho^B$  represent the Gaussian electron densities of molecules A and B, respectively.

**Fitting Procedure.** Gaussian electron densities can be found through fits to several molecular properties: density, electrostatic potential (ESP), and electric field. In this work, the ESP fitting procedure has been applied based on the reference ESP of the MP2/aug-cc-pVTZ method. The densities in this calculation are based on Z-vector<sup>27</sup> and hence yield multipole moments which are correct analytical derivative energy. The least-squares method is used to optimize the coefficients and exponents of GTFs in addition to locating the D sites by minimizing the following fitting function:

$$\chi^2 = \sum_{\alpha} w_{\alpha} \sum_i W(r_i) (y_{\alpha}(r_i) - y_{\alpha}(r_i, c_k))^2, \quad (4)$$

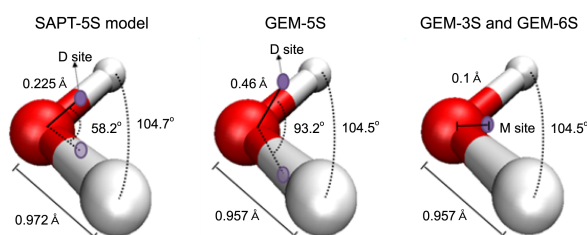
where  $y_{\alpha}(r_i)$  corresponds to the *ab initio* molecular property of interest at point  $i$ ,  $y_{\alpha}(r_i, c_k)$  is the same property evaluated with the  $k$ th ABS element at the same point on the grid, and  $\alpha$  denotes the molecular properties of interest.  $w_{\alpha}$  is the relative weight for property  $\alpha$ , and  $W(r)$  is the weighting function for the point on the grid as suggested by Hu *et al.*<sup>19</sup> and defined as:

$$W(r) = \exp[-\sigma(\log \rho(r) - \log \rho_{ref})^2], \quad (5)$$

in which  $\rho(r)$  is the predefined electron density, and  $\rho_{ref}$  is a reference electron density that was chosen in conjunction with  $\sigma$  to ensure the weighting function behaves as expected. Eq. (4) is minimized by using both the genetic algorithm (GA) and quasi-Newton optimization methods, which employ analytical gradients with respect to the coefficients and exponents of GTFs as well as the position of the D sites. Due to the symmetry of the water monomer, the fitting points span a quarter of the space. As suggested in Reference 16, the minimum cutoff values of 1.5 and 1.7 bohr were used for hydrogen and oxygen, respectively, and a maximum cutoff value of 10 bohr was used with the grid point 0.1 Å. This brings the total number of data points to 157412.

## Results

The D sites of GEM-5S, which are located beyond the HOH plane, were optimized and fitted to ABS, so that the positions of the D sites are different from original locations based on the SAPT-5S model. Figure 2 compares the geometry of the SAPT-5S model with that of GEM-5S. The positions of the D sites for GEM-5S have been optimized, and it resulted in bigger inner angles than those of SAPT-5S. Figure 2 also shows the geometries of GEM-3S and GEM-6S, which include the positions of the M sites. The M sites



**Figure 2.** The original SAPT-5S model and the optimal geometry of GEM-5S, and the geometries of GEM-3S and GEM-6S.

**Table 1.** Optimized coefficients (a.u.) and exponents of GTFs for the electron density and the dipole moments (Debye) of GEM-3S, GEM-5S, and GEM-6S, respectively

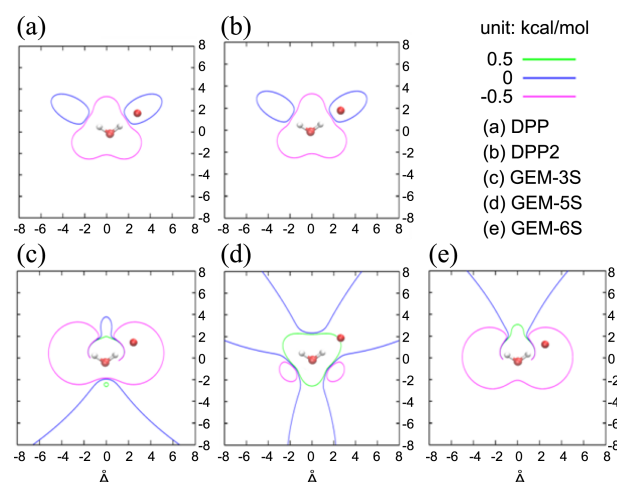
|             | GEM-3S |       | GEM-5S |       | GEM-6S |       |
|-------------|--------|-------|--------|-------|--------|-------|
|             | GTF1   |       | GTF1   |       | GTF1   | GTF2  |
| Coefficient | M:     | 6.776 | O:     | 6.300 | M:     | 4.229 |
|             | 2H:    | 0.612 | 2H:    | 0.463 | 2H:    | 0.167 |
|             |        |       | 2D:    | 0.387 |        | 0.427 |
| Exponent    | M:     | 0.604 | O:     | 0.787 | M:     | 6.209 |
|             | 2H:    | 0.472 | 2H:    | 4.695 | 2H:    | 0.397 |
|             |        |       | 2D:    | 0.193 |        | 0.435 |
| Dipole      | 1.810  |       | 1.849  |       | 1.895  |       |

have been placed at 0.10 Å away from oxygen along the HOH bisector by the trial-and-error method.

The optimized coefficients and exponents of the Gaussian type functions (GTFs) of GEM-3S, GEM-5S, and GEM-6S are shown in Table 1. The dipole moments of the monomer calculated with the GTFs in each model are also indicated in this table. The dipole moment of GEM-5S is very close to the experimental value, 1.855 Debye,<sup>20</sup> whereas those of GEM-3S and GEM-6S are slightly underestimated and overestimated, respectively.

**Electrostatic Potential.** To assess the fitting performance of the three different models for representing the electrostatic potential of each water monomer, I report the difference between the electrostatic potential associated with each model and that calculated at the MP2/aug-cc-pVTZ level of theory as a function of distance. Figure 3 shows the difference contours at the  $\pm 0.5$  kcal/mol level at points in the plane of the water monomer. The  $\pm 0.5$  kcal/mol contours were chosen for display because they represent a sizable portion (10%) of the net binding energy of the water dimer. Also I added the test charge (extra oxygen) in the figure to indicate the location of the second water molecule in the optimized geometry of the water dimer.

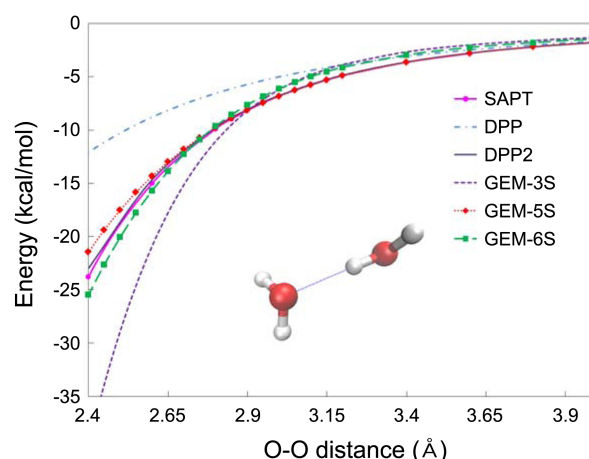
The electrostatic potential differences for the DPP model were greater than 0.5 kcal/mol when the test charge is brought within near 3.5 Å of the center. The difference between the contours of the DPP and DPP2 models was not found to be within  $\pm 0.5$  kcal/mol. The 0.5 kcal/mol contour lines of the ESP differences in GEM-5S were located near  $\pm 3.0$  Å, whereas those in GEM-3S and GEM-6S were found to be near  $\pm 4.5$  Å. Water molecules are usually bound with



**Figure 3.** Electrostatic potential difference plots for the water monomer. In each case the plot gives the difference of the electrostatic potentials calculated using the model and the MP2/aug-cc-pVTZ method. An extra oxygen atom is located in the optimized geometry of the water dimer.

other waters at near  $R_{OO}=3.0$  Å, so this figure indicates GEM-5S shows accurate electrostatics potentials within  $\pm 0.5$  kcal/mol errors when the test charge is located in the region of longer than 3.0 Å whereas GEM-3S and GEM-6S indicate this distance as 4.5 Å. Compared to the DPP and DPP2 water models, only GEM-5S exhibited better performance. GEM-3S and GEM-6S performed worse than the DPP water model during the analysis of the electrostatic potential difference of the water monomer. However, this contour analysis represents the information of only HOH plane.

**Water Dimer.** Figure 4 displays the electrostatic interaction energy of the water dimer as a function of the oxygen-oxygen separation between these two water monomers ( $R_{OO}$ ) for each considered model. This plot was generated while the monomers were kept rigid by fixing the OH bond lengths and HOH angles fixed at the experimentally determined



**Figure 4.** Electrostatic interaction energy for the water dimer, as a function of O-O distance, assuming rigid monomers and fixed flap angles.

values for the gas phase monomer. The flap angles that determine the orientation of these two monomers with respect to the OO axis were frozen at the values they have for the MP2/aug-cc-pVTZ optimized structure. More attractive electrostatic interactions are shown with more negative energies in this figure.

The electrostatic interaction energies of the water dimer are compared with the values obtained from symmetry-adapted perturbation theory (SAPT)<sup>21,22</sup> with the aug-cc-pVTZ basis set as a reference. The energy of the DPP water model is less attractive at short OO distances. Most classical water models, such as TTM2-R, DC, and AMOEBA, give nearly the same electrostatic interaction energies as the DPP water model.<sup>5</sup> These classical models are limited in their ability to express electrostatic energies at short distances since they do not account for charge penetration effects. In the case of the DPP2 model, the electrostatic energies are perfectly matched to the SAPT results because the parameter of the screening function was obtained by a least-square fitting of the electrostatic energies from the SAPT calculations.

The electrostatic energy of GEM-3S is more attractive in the region of  $R_{OO} \leq 2.8$  Å, but less attractive in the region of  $R_{OO} > 2.8$  Å. GEM-5S and GEM-6S provide better electrostatic energy calculations of the dimer than GEM-3S because these models used five and six Gaussian type functions to describe the electron density, respectively. In the region of  $R_{OO} > 2.7$  Å, the electrostatic energies of GEM-5S are almost the same as those of the SAPT calculation.

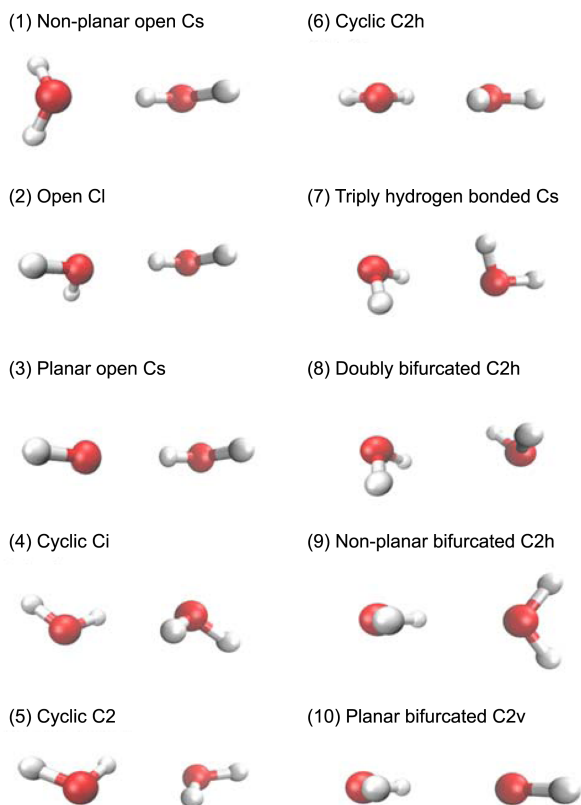


Figure 5. The Smith dimer sets.

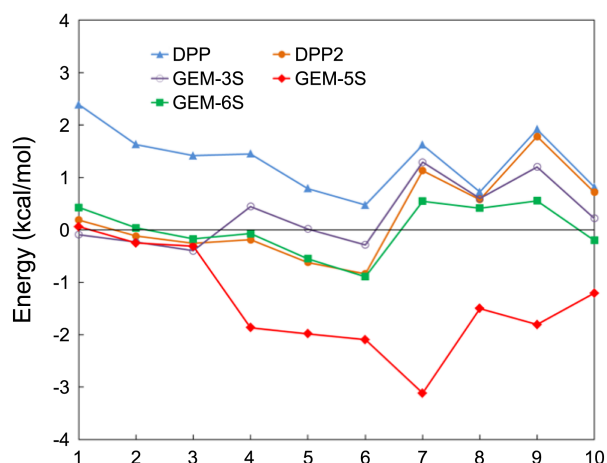


Figure 6. Errors in the electrostatic interaction energies for the ten dimers.

**The Ten Water Dimers.** As a more general test for the water dimer, ten stationary points of the water dimer's potential energy surface (PES) were considered. Figure 5 shows the geometries of the ten stationary points originally found by Smith *et al.*<sup>23</sup> (the so-called Smith dimer set). In this work, they have been characterized by employing second-order Moller-Pleset perturbation theory (MP2), and the OH bond lengths and HOH angles were fixed so that they conform to the rigid geometries of the DPP model.

The reference electrostatic energies of the ten stationary points of the dimer were determined from the SAPT decomposition results.<sup>8,26</sup> Figure 6 shows the errors in intermolecular Coulomb energy for the ten water dimers using the three investigated models, which are compared with those of the DPP and DPP2 water models. The DPP water model has smaller electrostatics energies than the SAPT results because of the lack of charge penetration. The electrostatic energies of GEM-3S, GEM-5S, GEM-6S, and DPP2 models of dimers 1–3 were almost the same as the SAPT results. However, for the stationary points of dimers 4–10, the electrostatic energies of all models greatly deviated from the SAPT results. For dimers 4–10, GEM-5S overestimated the electrostatic energies and gave values with inaccuracies comparable to the DPP water model. GEM-3S produced more accurate results than the DPP water model for all ten dimers, though the errors of the stationary points 7 and 9 were greater than 1 kcal/mol. Compared to the DPP2 model, GEM-3S give better performances for stationary points 5, 6, 9, and 10. GEM-6S reproduced the electrostatic energies for the ten stationary points quite well. It outperformed the other considered models of this study. Quite notably, GEM-6S gave significant improvements for dimers 7, 9, and 10, and the errors for all ten stationary points were less than 1 kcal/mol using this model.

## Conclusion

Three different GEM models have been designed and implemented to reproduce ESP using ABSs. The ABS sites

of GEM-3S, GEM-5S, and GEM-6S were limited to three, five, and six sites with *s*-type Gaussian functions, respectively. The electrostatic energies of the ten dimers were calculated to determine how well these models reproduce charge penetration.

GEM-5S analysis caused the electrostatic energies to be more attractive for seven out of the ten dimers, though it did exhibit the best electrostatic potential among the three tested models based on the difference of the electrostatic potential in the water monomer. The contour diagram of the water monomer only provides accurate information for the electrostatic energies of the HOH plane. Therefore, only three dimers, which have the interactions between the HOH planes, were well described with this model. In contrast to GEM-5S, GEM-6S performed more poorly at describing the electrostatic potential in the contour diagram of the water monomer than the DPP water model. However, the electrostatic energies of the ten dimers calculated with the GEM-6S model were in good agreement with the SAPT results, which indicated better performance than those of even the DPP2 model.

Unexpectedly, GEM-5S did not perform well in the ten water dimers, but GEM-6S agreed well with the SAPT results, which indicates that the M site is a suitable position to put auxiliary basis sets. GEM-3S can also be applied to water models with less computational cost. It is about twice faster than GEM-6S for calculating the electrostatic energy of a certain point.

**Acknowledgments.** This work was supported by the National Research Foundation (NRF) grant funded by the Korea government (NRF-2013R1A1A2008403) and the Supercomputing Center/Korea Institute of Science and Technology Information with supercomputing resources including technical support [No. KSC-2013-C2-009]. The author gratefully acknowledges Prof. Kenneth D. Jordan in University of Pittsburgh for helpful suggestions and comments.

## References

- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, 79, 926.
- Dang, L. X.; Chang, T.-M. *J. Chem. Phys.* **1997**, 106, 8149.
- Bumham, C. J.; Li, J. C.; Xantheas, S. S. *J. Chem. Phys.* **2002**, 116, 1500.
- Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, 107, 5933.
- DeFusco, A.; Schofield, D.; Jordan, K. D. *Mol. Phys.* **2007**, 105, 2681.
- Akin-ojo, O.; Wang, F. *J. Comput. Chem.* **2011**, 32, 453.
- Stone, A.; Alderton, M. *Mol. Phys.* **1985**, 56, 1047.
- Kumar, R.; Wang, F.-F.; Jenness, G. R.; Jordan, K. D. *J. Chem. Phys.* **2010**, 132, 014309.
- Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, 112, 7300.
- Piquemal, J. P.; Gresh, N.; Giessner-Prettre, C. *J. Phys. Chem. A* **2005**, 107, 10353.
- Gresh, N.; Piquemal, J. P.; Krauss, M. *J. Comput. Chem.* **2005**, 26, 1113.
- Cisneros, G. A.; im Tholander, S. N.; Parisel, O.; Darden, T. A.; Elking, D.; Perera, L.; Piquemal, J.-P. *Int. J. Quantum Chem.* **2008**, 108, 1905.
- Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2006**, 125, 184101.
- Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2005**, 123, 044109.
- Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Phys. Chem. B* **2006**, 110, 13682.
- Cisneros, G. A.; Elking, D.; Piquemal, J.-P.; Darden, T. A. *J. Phys. Chem. A* **2007**, 111, 12049.
- Cisneros, G. A. *J. Chem. Theory Comput.* **2012**, 8, 5072.
- Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **2000**, 113, 6687.
- Hu, H.; Lu, Z.; Yang, W. *J. Chem. Theo. Comput.* **2007**, 3, 1004.
- Lovas, F. J. *J. Phys. Chem. Ref. Data* **1978**, 7, 1445.
- Rybak, S.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **1991**, 95, 6576.
- Mas, E. M.; Szalewicz, K.; Bukowski, R.; Jeziorski, B. *J. Chem. Phys.* **1997**, 107, 4207.
- Smith, B. J.; Swanton, D. J.; Pople, J. A.; Schaefer, H. F.; Radom, L. *J. Chem. Phys.* **1990**, 92, 1240.
- Bagus, P. S.; Illas, F. J. *J. Chem. Phys.* **1992**, 96, 8962.
- Piquemal, J.-P.; Marquez, A.; Parisel, O.; Giessner-Prettre, C. *J. Comput. Chem.* **2005**, 26, 1052.
- Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, 111, 8753.
- Dierksen, G. H. F.; Roos, B. O.; Sadlej, A. J. *J. Chem. Phys.* **1981**, 26, 29.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.;