The Charge Response Kernel with Modified Electrostatic Potential Charge Model

Akihiro Morita* and Shigeki Kato

Department of Chemistry, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

Received: November 9, 2001; In Final Form: January 25, 2002

The charge response kernel (CRK), $\partial Q_a/\partial V_b$, was reformulated using a modified definition of the partial charge Q_a. The modifications could effectively prevent numerical instability of the partial charges based on the electrostatic potential, and thus eliminate spurious components of CRK due to the fitting problem. The modified CRK was directly derived via the coupled-perturbed Hartree−Fock equations. The performance of this modified CRK model was examined by applying it to some test molecules including ethanol, DMSO, chloroform, and trimethylamine.

1. Introduction

It has been recognized in solution chemistry for a few decades that electronic polarization has a crucial role in accurate modeling of the solvation structure and dynamics in solutions. By properly taking account of the electronic polarization, some experimental facts have been elucidated via molecular dynamics (MD) simulations, including the anomalously slow diffusion coefficients of some aromatic radicals,[1,2] or the very fast vibrational relaxation of the azide ion.[3] A key quantity to represent the electronic polarization throughout those works has been the charge response kernel (CRK), which was recently proposed via ab initio molecular orbital (MO) formulations.[1] The CRK is defined in terms of the intramolecular site representation as $\partial Q_a/\partial V_b$, where Q_a is the partial charge at the site a and V_b the electrostatic potential at the site b. This quantity is utilized to represent the partial charge redistribution among the molecular sites with respect to the external electrostatic perturbation. A similar fluctuating charge model with empirical formulation has been also proposed by Berne and co-workers.[4,5]

The CRK has been shown to be a quite promising method of general polarizable molecular models, for a number of reasons listed below. First, the formulation is fully based on the ab initio MO theory at any level of accuracy, and thus no empirical adjustment of the parameters is required. Second, it is capable of describing any nonlocal charge distribution among the sites, whereas some other polarizable models where each site carries a usual polarizability can describe only the local polarization at each site. Third, in implementing the CRK to MD codes, usual interaction site molecular models could be readily extended.

Despite these potential advantages, we were aware of some technical problems that remain to be resolved, to establish the general usefulness of this theoretical method. One of the most crucial problems is the definition of the partial charges Q_a. In the original formulation of ref 1, the partial charges were assigned to the intramolecular sites through the least-squares fitting to the ambient electrostatic potential outside the molecule. Although the partial charges thus defined, generally called ESP (electrostatic potential) charges,[6,7] have an advantage to be optimized to describe the intermolecular electrostatic interaction, it is well-known that the charge assignment often becomes problematic, particularly when those of buried atoms are involved. The partial charge distribution of the buried atoms is not well defined from the electrostatic potential in the outer region, because the buried partial charges are shielded and readily offset by other surrounding charges, and thus have merely indirect influence to the electrostatic potential in the outer region. The ill-defined behavior of the ESP charges may become even more serious in the CRK than the partial charges themselves, since the intramolecular charge redistribution could be easily contaminated by the instability of the charge definition. Therefore, to obtain the well-defined CRK, we should distinguish the real charge flow induced by the external field from the spurious one due to the uncertainty of the charge assignment. Though the spurious flow may little affect the intermolecular electrostatic interaction, the CRK involving the spurious charge flow could facilitate the unphysical divergence of the polarization,[2,8,9] when it is applied to charge polarizable MD simulations.

In this paper we introduce a modified fitting procedure of the electrostatic potential and extend the CRK formulation in accord with the model. The present philosophy is analogous to the restrained electrostatic potential (RESP) method developed by Kollman and co-workers.[10−13] The RESP model was devised to avoid this problem of the charge assignment with little increase of the computational cost, and has been shown to be almost as accurate as the original ESP model in the intermolecular electrostatic interaction.[10] While the RESP model is quite convenient and successful, the extension to the CRK poses a problem, since the charge-fitting procedure of the RESP model involves a nonlinear equation which has to be solved iteratively. Accordingly the analytical definition of CRK based on the RESP charge should involve complicated nonlinear equations, and the usual coupled-perturbed Hartree−Fock (CPHF) equations[14−16] could not be utilized. Therefore, we introduce an alternative definition of the partial charges based on the electrostatic potential. This partial charge model is as robust and effective to avoid the uncertainty problem of the charges as the RESP model, and in fact it gives similar partial charges to the RESP model, whereas the extension to the CRK has no problem with this model via the CPHF equations. When we use a CRK model with the virtue of the RESP charges, the present model could be recommended.

* To whom correspondence should be addressed. E-mail: morita@kuchem.kyoto-u.ac.jp.

10.1021/jp014114o CCC: $22.00
Published on Web 03/23/2002

The remainder of this paper is as follows. In section 2, the definition of the partial charges and the CRK based on this model are described. The new model is applied to some test molecules in section 3, i.e., ethanol, dimethyl sulfoxide (DMSO), chloroform, and trimethylamine, to distinguish the spurious charge flow in the original CRK and to determine the parameter involved in the new model. A brief conclusion and future prospect are given in section 4.

2. Formulation

The original definitions of the partial charges and the CRK were given elsewhere,[1] and thus the modifications are emphasized here. In the usual least-squares fitting procedure for a set of the ESP charges $\{Q_a\}$, the sum of the squared deviation L is minimized,

$$L = \sum_n^{\text{grids}} w_n \left[\sum_a^{\text{sites}} \frac{Q_a}{|\mathbf{R}_s(a) - \mathbf{R}_g(n)|} + \sum_{p,q}^{\text{AO}} D_{pq} \left\langle p \left| \frac{1}{|\mathbf{r} - \mathbf{R}_g(n)|} \right| q \right\rangle - \sum_c^{\text{nuclei}} \frac{Z_c}{|\mathbf{R}_n(c) - \mathbf{R}_g(n)|} \right]^2 \quad (1)$$

where $\mathbf{R}_s(a)$, $R_n(c)$, $\mathbf{R}_g(n)$, and $\mathbf{r}$ denote the coordinates of the site a, the atomic nucleus c, the grid point n and electron, respectively. (Note that the coordinates of the sites and atomic nuclei are not necessarily identical.) D_{pq} is the electronic density matrix, and p, q refer to the atomic orbitals as the basis functions. Z_c is the nuclear charge. w_n is the weighting factor in the least-squares fitting. The equation to be solved is

$$\frac{\partial}{\partial Q_a}[L - 2\lambda(\sum_a^{\text{sites}} Q_a - Q) - 2\boldsymbol{\xi}\cdot(\sum_a^{\text{sites}} Q_a \mathbf{R}_s(a) - \boldsymbol{\mu})] = 0 \quad (2)$$

where 2λ and $2\boldsymbol{\xi} = (2\xi_x, 2\xi_y, 2\xi_z)$ are the Lagrange multipliers to impose constraint on the total charge Q and the dipole moment $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$. Equation 2 leads to the following linear equation for each a to determine the set of partial charges $\{Q_a\}$,

$$\sum_b^{\text{sites}} A_{ab} Q_b - B_a - \lambda - \boldsymbol{\xi}\cdot\mathbf{R}_s(a) = 0 \quad (3)$$

where

$$A_{ab} = \sum_n^{\text{grids}} w_n \frac{1}{|\mathbf{R}_s(a) - \mathbf{R}_g(n)|} \cdot \frac{1}{|\mathbf{R}_s(b) - \mathbf{R}_g(n)|} \quad (4)$$

$$B_a = \sum_n^{\text{grids}} w_n \left\{ -\sum_{p,q}^{\text{AO}} D_{pq} \left\langle p \left| \frac{1}{|\mathbf{r} - \mathbf{R}_g(n)|} \right| q \right\rangle + \sum_n^{\text{nuclei}} \frac{Z_c}{|\mathbf{R}_n(c) - \mathbf{R}_g(n)|} \right\} \cdot \frac{1}{|\mathbf{R}_s(a) - \mathbf{R}_g(n)|} \quad (5)$$

Note that the matrix $\mathbf{A}$ is symmetric and positive definite when the weighting factor w_n is positive, which is almost always the case. The partial charge Q_a is thus given as

$$Q_a = \sum_b^{\text{sites}} [\mathbf{A}^{-1}]_{ab}(B_b + \lambda + \boldsymbol{\xi}\cdot\mathbf{R}_s(b)) \quad (6)$$

where $\mathbf{A}^{-1}$ is the inverse matrix of $\mathbf{A}$. The explicit expressions of λ and $\boldsymbol{\xi}$ are given in eq 17 of Appendix A.

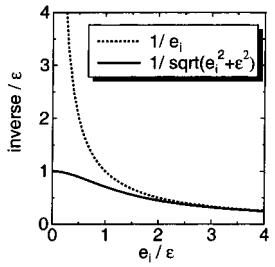

Figure 1. The damping effect of the inverse eigenvalue in the modified fitting procedure. The dotted curve refers to the original inverse eigenvalue $1/e_i$, and the solid curve to the modified one $1/\sqrt{e_i^2+\epsilon^2}$.

In the above least-squares fitting procedure, the problem of the charge fitting may occur in solving eq 3, when the matrix $\mathbf{A}$ has a very small eigenvalue. Thus a small displacement of $\{Q_a\}$ along the eigenvector little affects the accuracy of the least-squares fitting, and often unphysically large displacement along this mode results from the fitting procedure. Hence we modify the $\mathbf{A}$ matrix by augmenting such small eigenvalues. Suppose the $\mathbf{A}$ matrix is diagonalized as follows:

$$\mathbf{A} = \mathbf{P}^{\mathbf{T}}\cdot\mathbf{E}\cdot\mathbf{P}, \quad \mathbf{E} = \begin{bmatrix} e_1 & & \\ & \ddots & \\ & & e_N \end{bmatrix} (e_i > 0) \quad (7)$$

then the eigenvalue e_i is replaced by $\sqrt{e_i^2+\epsilon^2}$, where ϵ is a certain parameter. The modified inverse matrix $(\mathbf{A}^{\text{mod}})^{-1}$ is subsequently defined as

$$(\mathbf{A}^{\text{mod}})^{-1} = \mathbf{P}^{\mathbf{T}}\cdot(\mathbf{E}^{\text{mod}})^{-1}\cdot\mathbf{P},$$
$$(\mathbf{E}^{\text{mod}})^{-1} = \begin{bmatrix} 1/\sqrt{e_1^2+\epsilon^2} & & \\ & \ddots & \\ & & 1/\sqrt{e_N^2+\epsilon^2} \end{bmatrix} \quad (8)$$

Equation 8 is employed in eq 6 (or eqs 16 and 19 in Appendix A) to obtain the modified partial charges.

The effect of this modification is illustrated in Figure 1. This figure implies that the unstable behavior is suppressed when the eigenvalue e_i is equal to or less than ϵ, while the inverse eigenvalue $1/e_i$ is little influenced by this modification at a larger e_i. The parameter ϵ will be defined in the next section through examining some test molecules.

The derivation of the charge response kernel $\partial Q_a/\partial V_b$ is now straightforward via the CPHF equations. When the electrostatic perturbation Hamiltonian H' is given in the site representation as

$$H' = \sum_a^{\text{sites}} \hat{Q}_a V_a \quad (9)$$

where $\hat{Q}_a$ is the partial charge operator at the site a. $\hat{Q}_a$ consists of the electronic and nuclear parts as

$$\hat{Q}_a = \sum_i^{\text{electrons}} \hat{q}_a(i) + Q_a^{\text{nuc}} \quad (10)$$

where the explicit expressions of $\hat{q}_a$ and Q_a^{nuc} are given in eqs 26 and 27 of Appendix A. $\hat{Q}_a$ derives the partial charge Q_a of eq 6 via $Q_a = \langle\Psi|\hat{Q}_a|\Psi\rangle$ with the electronic wave function Ψ.

The modified partial charge operator is defined in the same way by replacing $\mathbf{A}^{-1}$ in eqs 16, 26, and 27 with $(\mathbf{A}^{\mathrm{mod}})^{-1}$ derived in eq 8.

When Ψ is a closed-shell restricted Hartree−Fock (RHF) wave function, the CPHF equation for each site a is as follows:

$$(\epsilon_l - \epsilon_i)U_{li}^a + \sum_j^{\mathrm{occ}} \sum_k^{\mathrm{vir}} H_{likj} U_{kj}^a = -Q_{li}^{\mathrm{a}} \tag{11}$$

where the suffixes $i-l$ denote canonical MOs; i and j stand for the occupied MOs, and k and l for the virtual, respectively. ϵ_i is the ith canonical orbital energy, $H_{likj} = 4(li|kj) - (lk|ij) - (lj|ik)$, and $Q_{li}^{\mathrm{a}} = \langle l|\hat{q}_a|i\rangle$. Using U obtained from eq 11, the CRK is given as

$$\frac{\partial Q_a}{\partial V_b} = \frac{\partial^2 E}{\partial V_a \partial V_b} = \sum_i^{\mathrm{occ}} \sum_k^{\mathrm{vir}} 4 Q_{ik}^a U_{ki}^b \tag{12}$$

where E is the total energy, and i and k stand for the occupied and virtual MOs, respectively.

Finally, we make a few comments on the nature of the CRK. As discussed in section 1, the CRK represents the general response of the intramolecular charge distribution to the external field on the basis of the site representation, and therefore the usual molecular polarizability is expressed as a special case of the perturbation field. The molecular polarizability $\boldsymbol{\alpha}$ is derived by the CRK as

$$\boldsymbol{\alpha} = -\sum_{a,b}^{\mathrm{sites}} \frac{\partial Q_a}{\partial V_b} \mathbf{R}_s(a) \otimes \mathbf{R}_s(b) \tag{13}$$

where $\otimes$ denotes the tensor product. The value of $\boldsymbol{\alpha}$ in eq 13 will be confirmed to be consistent to that of the direct ab initio calculations in the next section.

We note the asymptotic behavior of the partial charges and CRK with very large ϵ. Equation 8 implies that $(\mathbf{A}^{\mathrm{mod}})^{-1}$ is asymptotically scaled with $1/\epsilon$ when ϵ is sufficiently large. Subsequently, the partial charge Q_a is also scaled with $1/\epsilon$ as indicated by eq 6, and hence the linearity of the CPHF eq 11 implies that U is scaled in the same way. Therefore, the CRK in eq 12 is asymptotically scaled with $1/\epsilon^2$. The scaling factors of the partial charges and CRK are consistent with the fact that the partial charges and CRK are the first and second derivatives of the total energy; $Q_a = \partial E/\partial V_a$ and $\partial Q_a/\partial V_b = \partial^2 E/\partial V_a \partial V_b$.

3. Results and Discussion

The modified partial charges and CRK described in section 2 is now applied to some test molecules, i.e., ethanol, DMSO, chloroform, and trimethylamine, to examine the validity of the present model and to determine the damping parameter involved, ϵ.

The grid points to evaluate the electrostatic potential, $\mathbf{R}_g(n)$ in eq 1, are distributed on the envelope of the united spheres surrounding the molecule. The center of each sphere is located at an atomic nucleus and its radius is the standard atomic van der Waals radius by Bondi[17] plus a 1.5 Å margin. This margin approximately corresponds to the radius of a contact atomic site of another molecule. The grid distribution adheres to the point group symmetry of the molecule. Each sphere would have 1646 grid points evenly distributed on the 4π steradian area if the atom were completely exposed. The total numbers of grid points N_{grid} are 3584 for ethanol, 3821 for DMSO, 3664 for chloroform, and 4204 for trimethylamine. The weighting factor, w_n in eq 1, is set to be unity for all the grid points. We employ all atoms for the sites throughout; the site locations $\mathbf{R}_s$ are identical to the atomic nuclei $\mathbf{R}_n$.

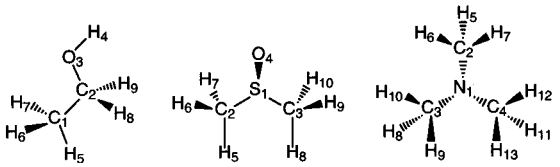

Figure 2. Schematic pictures of ethanol (left part), DMSO (center), and trimethylamine (right) with the serial numbers of the atomic sites. Note that ethanol and DMSO have C_s symmetry and trimethylamine has C_{3v} at their optimized geometries.

For the present discussion, ab initio MO calculations are employed in the HF/6-31G* level. This is because the RESP charges have been determined at this level of calculations and we could make use of those results to optimize the damping parameter ϵ. The molecular geometries are also optimized at the same level of calculations. The electronic structure calculations were carried out via the HONDO package[18] complemented with the program code developed by us for the partial charges and the CRK calculations.

3.1. Ethanol and DMSO. Both ethanol and DMSO have two buried atoms of carbon, i.e., C_1 and C_2 for ethanol and C_2 and C_3 for DMSO in Figure 2. Due to these buried sites, the molecules could suffer from the fitting problem discussed in section 2, which is certainly the case as shown below. The optimized molecular geometry retains the C_s symmetry in either case. In the following, we first discuss the ethanol case in details as an example, and the parallel discussion also holds in the DMSO case as shown later.

Table 1 shows the eigenvalues and eigenvectors of the original **A** matrix of ethanol in eq 4. The modes 1 (with the symmetry A′) and 2 (A′) have particularly small eigenvalues, 0.003 and 0.005, and accordingly these modes could cause the fitting problem. The eigenvectors of these modes 1 and 2 essentially consist of the buried carbons, C_1 and C_2 respectively, with the largest amplitude and the adjacent atoms (C_2, and H_5-H_7 for the mode 1, C_1, O_3, and H_8-H_9 for the mode 2) with the opposite phase. These eigenvectors correspond to the charge displacement from the buried carbon sites, C_1 and C_2, to the adjacent sites surrounding the buried carbons, or vice versa, which is not well defined from the electrostatic potential of the outer region. The partial charges of the buried sites are determined to be −0.258 and 0.497 via the original ESP fitting with $\epsilon = 0$ in Table 2, which substantially overestimate the RESP charges, −0.099 and 0.312, determined from the same level of wave function HF/6-31G*.[12]

Table 2 also displays the eigenvalue analysis of the CRK with $\epsilon = 0$, which apparently indicates that two of the CRK modes have particularly large negative eigenvalues, −42.4 and −38.4. We notice in Table 2 that the eigenvectors of these CRK modes are quite analogous with the modes 1 and 2 of the **A** matrix in Table 1. The strong correlation suggests that these soft CRK modes 1 and 2 in Table 2 might be a direct consequence of the particularly small eigenvalues of the modes 1 and 2 in Table 1, implying the problem of charge fitting.

This problem is further investigated via the dependence of the CRK modes and charges on the damping parameter ϵ. Figure 3 demonstrates that the two CRK modes with the largest negative eigenvalues at $\epsilon = 0$ are readily damped with increasing ϵ to approach the asymptotic behavior $\sim 1/\epsilon^2$, whereas the eigenvalues of the other modes are nearly invariant until $\epsilon \sim 10^{-1}$. One observes many avoided crossings among the modes

TABLE 1: Eigenvalues and Eigenvectors of the Original A Matrix (Equation 4) of Ethanol[a]

	1(A′) 0.003	2(A′) 0.005	3(A′) 0.179	4(A″) 0.672	5(A′) 1.102	6(A″) 4.350	7(A′) 5.890	8(A′) 15.98	9(A′) 557.2
1 C	0.88	0.20	−0.04	0.00	0.01	0.00	−0.06	0.27	0.33
2 C	−0.27	0.85	−0.16	0.00	−0.15	0.00	0.16	−0.12	0.33
3 O	−0.03	−0.25	−0.76	0.00	0.03	0.00	−0.37	−0.33	0.33
4 H	0.04	0.04	0.54	0.00	0.43	0.00	−0.30	−0.57	0.33
5 H	−0.24	−0.10	−0.12	0.00	0.72	0.00	0.31	0.44	0.33
6 H	−0.22	−0.10	0.19	−0.50	−0.26	−0.50	−0.34	0.33	0.33
7 H	−0.22	−0.10	0.19	0.50	−0.26	0.50	−0.34	0.33	0.33
8 H	0.03	−0.27	0.08	0.50	−0.26	−0.50	0.46	−0.18	0.34
9 H	0.03	−0.27	0.08	−0.50	−0.26	0.50	0.46	−0.18	0.34

[a] Parentheses in the first row indicate the symmetry of the modes, and the second row the eigenvalues in au. The serial numbers of the atomic sites are shown in Figure 2.

TABLE 2: The ESP Charges Q_a and the Eigenvalues/Vectors of the CRK of Ethanol at $\epsilon = 0$

		eigenvalues/vectors of CRK								
	Q_a	1(A′) −42.4	2(A′) −38.4	3(A′) −4.93	4(A″) −2.23	5(A″) −2.23	6(A′) −2.05	7(A′) −1.12	8(A′) −0.52	9(A′) 0.00
1 C	−0.258	0.85	0.30	0.02	0.00	0.00	−0.01	−0.08	0.27	0.33
2 C	0.497	−0.37	0.82	−0.12	0.00	0.00	−0.03	0.23	−0.09	0.33
3 O	−0.719	0.06	−0.21	−0.80	0.00	0.00	−0.08	−0.22	−0.38	0.33
4 H	0.415	0.00	0.01	0.54	0.00	0.00	0.10	−0.48	−0.60	0.33
5 H	0.041	−0.20	−0.12	−0.06	0.00	0.00	0.82	−0.14	0.37	0.33
6 H	0.079	−0.22	−0.14	0.09	0.28	−0.65	−0.39	−0.19	0.35	0.33
7 H	0.079	−0.22	−0.14	0.09	−0.28	0.65	−0.39	−0.19	0.35	0.33
8 H	−0.067	0.05	−0.26	0.12	−0.65	−0.28	0.00	0.54	−0.13	0.33
9 H	−0.067	0.05	−0.26	0.12	0.65	0.28	0.00	0.54	−0.13	0.33

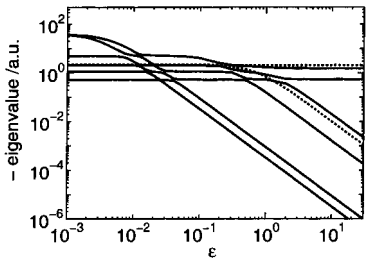

Figure 3. The eigenvalues of the CRK modes of ethanol (cf. Table 2 at $\epsilon = 0$) as a function of ϵ. The solid lines denote the A′ modes, and the dotted lines the A″ modes. The ordinate refers to the absolute eigenvalue in the log scale.

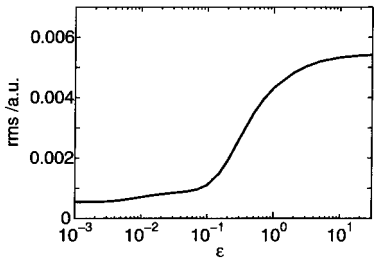

Figure 4. The root-mean-square deviation $\sqrt{L/N_{\text{grid}}}$ of the electrostatic potential reproduced via the ethanol ESP charges as a function of ϵ.

with the same symmetry in Figure 3, but the mode characters of the "diabatic states" are preserved through the avoided crossings, and the two softest modes at $\epsilon = 0$ are diabatically correlated with the two hardest modes (with smallest negative eigenvalues) at sufficiently large $\epsilon \gtrsim 0.05$.

Figure 4 deals with the root-mean-square deviation of the ESP fitting procedure as an index of the fitting accuracy, $\sqrt{L/N_{\text{grid}}}$, where L is defined in eq 1 and N_{grid} is the number of the grid points. While the root-mean-square deviation increases with ϵ, the fitting accuracy is not so deteriorated with increasing ϵ until $\epsilon \sim 10^{-1}$. In fact, the root-mean-square deviation

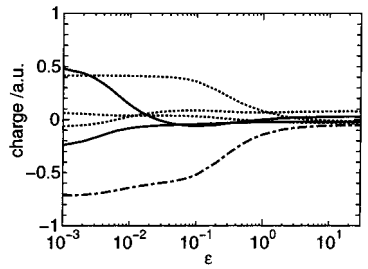

Figure 5. The partial charges of the ethanol sites as a function of ϵ. The two solid lines denote the carbon sites C_1 and C_2, the dash−dotted line the oxygen O_3, and the three dotted lines the hydrogens H_4, H_5−H_7, and H_8−H_9. The charges of the equivalent methyl hydrogens H_5−H_7 are averaged. (See Q_a in Table 2 for further assignment.)

$\sqrt{L/N_{\text{grid}}}$ is only 1.65 times enhanced at $\epsilon = 0.05$ compared to that at $\epsilon = 0$. The eigenvalues of the CRK modes 1 and 2 are already quite suppressed to be −0.34 and −0.13 at $\epsilon = 0.05$ from −42.4 and −38.4 at $\epsilon = 0$, which implies that the charge polarization along these modes should be less significant than those along the other modes.

The partial charges are displayed in Figure 5 as a function of ϵ, which shows that even in the range $\epsilon = 0-10^{-1}$ where the root-mean-square deviation does not vary much, the magnitude of the ESP charges of the buried sites is dramatically reduced.

The polarizability derived from the CRK via eq 13 was examined. The CRK-derived polarizability always agreed with the polarizability directly obtained from ab initio calculations within 0.04% over all range of ϵ considered. This excellent agreement is attributed to the constraint on the dipole moment in the charge fitting procedure, regardless of ϵ.

In the DMSO case, essentially the same story as ethanol can be applied, as shown in Figure 6. The **A** matrix of DMSO has two particularly small eigenvalues, 0.0029 (with the symmetry A′) and 0.0030 (A″). These two eigenvalues are far smaller than the third smallest one, 0.238, and the two modes correspond to

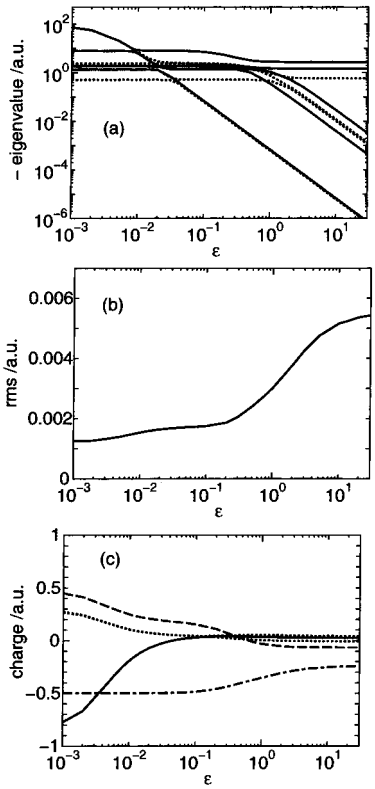

Figure 6. DMSO. (a) The eigenvalues of the CRK modes. The solid lines denote the A′ modes, and the dotted lines A″. (b) The root-mean-square deviation $\sqrt{L/N_{\mathrm{grid}}}$ for the ESP charges. (c) The site partial charges. The dashed line denotes the sulfur site S_1, the solid line the two equivalent carbon sites C_2 and C_3, the dash−dotted line the oxygen O_4, and the dotted line the methyl hydrogens H_5-H_{10}. The partial charges of the equivalent methyl hydrogens H_5-H_{10} are averaged.

the charge displacement from the two buried carbon sites to the adjacent sites or vice versa. Subsequently, the CRK at $\epsilon = 0$ has two particularly large negative eigenvalues, relevant to the two buried sites. Figure 6a clearly shows that the two CRK modes with large negative eigenvalues at $\epsilon = 0$ are readily suppressed with increasing ϵ.

Next we provide an optimum ϵ for the partial charges and CRK calculations, on the basis of the above considerations. Figures 4 and 6b imply that any value of ϵ in the range of $\epsilon = 0-10^{-1}$ could be employed in terms of the fitting accuracy. The derived ESP charges are displayed in Figures 5 and 6c as a function of ϵ, which indicate that even in the allowable range $\epsilon = 0-10^{-1}$, the ESP charges, particularly those of the buried carbon sites, do vary considerably. Therefore, besides the root-mean-square deviation, we make use of another criterion to optimize ϵ. Fox and Kollman[12] determined the RESP charges of ethanol and DMSO using the same 6-31G* basis set as shown in Table 3. Accordingly we determined the optimum ϵ in the allowable range so that the partial charges best match the RESP charges. The recommended value of ϵ is $\epsilon = 0.006$. This value should be reasonable, since it effectively augments the smallest eigenvalues of the **A** matrix (about 0.003−0.005) associated to the buried sites, while it little affects the other, larger eigenvalues of A ($\gtrsim$0.1) as discussed in Figure 1. The partial charges and CRK of ethanol and DMSO with $\epsilon = 0.006$ are listed in Table 3. This table also shows the averaged values among the methyl hydrogens in the parentheses, since it would be convenient to regard the methyl hydrogens as equivalent sites, when the internal rotation of the methyl moiety is considered in actual simulations.

3.2. Chloroform. Chloroform is a simple example having one buried atom. One might expect the similar problem of charge fitting, but this system has an exceptional behavior as discussed below.

Figure 7 summarizes the results of the ϵ dependence. We see in Figure 7a that one of the CRK modes with the largest negative eigenvalue at $\epsilon = 0$ is readily suppressed as ϵ increases, whereas the other CRK modes keep almost invariant eigenvalues over ϵ. The former mode corresponds to the charge displacement from the central carbon site to the others or vice versa, and this character is retained "diabatically" through the avoided crossing at $\epsilon \approx 0.07$ in Figure 7a.

However, a remarkable feature is seen in Figure 7 (b) that the root-mean-square deviation of the ESP fitting is nearly constant over ϵ. On the other hand, Figure 7c reveals a transition region at $\epsilon = 10^{-2}-10^{-1}$, where the negative charge on the carbon site shifts to the three equivalent chlorine sites. This transition behavior can be understood using eq 6 or 19, indicating that the partial charge Q_a consists of the first term and the rest involving the Lagrange multipliers; the former results from the least-squares fitting, and the rest from the constraint on the total charge and dipole moment. When ϵ is sufficiently large, the first term becomes negligible and consequently the Lagrange constraint takes over the partial charge distribution.

A peculiar feature of this molecule is that the two sets of the partial charge distribution, before and after the transition, have equivalent performance in terms of the root-mean-square deviation. In the present system, there remains only one net degree of freedom for the charge distribution, since the three partial charges on the carbon, hydrogen, and three equivalent chlorine sites are restricted by the constraint on the total charge and the dipole moment along the molecular axis. The only remaining degree of freedom, corresponding to the charge displacement from the central carbon to the chlorines, happens to have little influence on the root-mean-square deviation. This insensitivity of the root-mean-square deviation to ϵ arises from very few degrees of freedom in the partial charge distribution. Nonetheless, we note that the partial charges determined with the recommended $\epsilon = 0.006$ in the previous subsection are $q(\mathrm{C}) = -0.3529$, $q(\mathrm{H}) = 0.2925$, and $q(\mathrm{Cl}) = 0.0201$, which are consistent to the RESP charges using the same 6-31G* basis set: $q(\mathrm{C}) = -0.3460$, $q(\mathrm{H}) = 0.3097$, and $q(\mathrm{Cl}) = 0.0121$.[12]

3.3. Trimethylamine. Trimethylamine, $N(CH_3)_3$, has three buried carbon sites. The optimized geometry includes a pyramidal nitrogen site with the overall C_{3v} symmetry, as schematically illustrated in Figure 2. The **A** matrix of trimethylamine has three particularly small eigenvalues, 0.0019 with the degenerate E symmetry and 0.0021 with A_1; the fourth smallest eigenvalue is 0.210, much larger in two orders of magnitude. Therefore, it is expected that the recommended value of $\epsilon = 0.006$ should also be reasonably applied to this molecule.

The ϵ dependence of (a) the eigenvalues of the CRK modes, (b) the root-mean-square deviation of the charge fitting, and (c) the partial charges is displayed in Figure 8. Part a reveals three CRK modes, i.e., one A_1 mode and two degenerate E modes, are quite sensitive to the variation of ϵ. These three modes are characterized as the charge distribution among the three carbon sites and their adjacent sites, essentially the same character of the unstable modes discussed in section 3.1. The root-mean-square deviation in part b indicates a transition region at $\epsilon = 10^{-1}-10^{1}$, where the root-mean-square deviation

TABLE 3: The Partial Charges Q_a and CRK of Ethanol and DMSO Determined by HF/6-31G* with $\epsilon = 0.006$[a]

(a) Ethanol

	C1	C2	O3	H4	H5	H6	H7	H8	H9
Q_a	−0.110	0.252	−0.669	0.415	0.012	0.053	0.053	−0.002	−0.002
					(0.039)	(0.039)	(0.039)		
Q_a^{RESP}	−0.099	0.312	−0.672	0.414	0.035	0.035	0.035	−0.029	−0.029
1	−6.590								
2	−0.831	−12.135							
3	−0.211	2.329	−3.922						
4	0.434	0.401	1.918	−1.915					
5	1.534	0.991	−0.153	−0.067	−1.901				
	(1.866)	(1.178)	(0.068)	(−0.206)	(−2.137)				
6	2.032	1.272	0.178	−0.275	0.028	−2.255			
	(1.866)	(1.178)	(0.068)	(−0.206)	(−0.036)	(−2.137)			
7	2.032	1.272	0.178	−0.275	0.028	−0.165	−2.255		
	(1.866)	(1.178)	(0.068)	(−0.206)	(−0.036)	(−0.036)	(−2.137)		
8	0.800	3.351	−0.158	−0.110	−0.229	−0.378	−0.437	−2.521	
					(−0.348)	(−0.348)	(−0.348)		
9	0.800	3.351	−0.158	−0.110	−0.229	−0.437	−0.378	−0.317	−2.521
					(−0.348)	(−0.348)	(−0.348)		

(b) DMSO

	S1	C2	C3	O4	H5	H6	H7	H8	H9	H10
Q_a	0.296	−0.329	−0.329	−0.498	0.121	0.147	0.162	0.121	0.147	0.162
					(0.143)	(0.143)	(0.143)	(0.143)	(0.143)	(0.143)
Q_a^{RESP}	0.316	−0.324	−0.324	−0.521	0.142	0.142	0.142	0.142	0.142	0.142
1	−7.593									
2	2.541	−12.782								
3	2.541	0.476	−12.782							
4	2.621	0.320	0.320	−2.494						
5	−0.064	3.590	−0.322	−0.186	−2.634					
	(−0.018)	(3.268)	(−0.120)	(−0.128)	(−2.460)					
6	−0.042	3.669	−0.462	−0.096	−0.531	−2.730				
	(−0.018)	(3.268)	(−0.120)	(−0.128)	(−0.328)	(−2.460)				
7	0.051	2.544	0.425	−0.102	−0.192	−0.259	−2.016			
	(−0.018)	(3.268)	(−0.120)	(−0.128)	(−0.328)	(−0.328)	(−2.460)			
8	−0.064	−0.322	3.590	−0.186	0.374	0.114	−0.148	−2.634		
	(−0.018)	(−0.120)	(3.268)	(−0.128)	(0.038)	(0.038)	(0.038)	(−2.460)		
9	−0.042	−0.462	3.669	−0.096	0.114	0.442	−0.104	−0.531	−2.730	
	(−0.018)	(−0.120)	(3.268)	(−0.128)	(0.038)	(0.038)	(0.038)	(−0.328)	(−2.460)	
10	0.051	0.425	2.544	−0.102	−0.148	−0.104	−0.198	−0.192	−0.259	−2.016
	(−0.018)	(−0.120)	(3.268)	(−0.128)	(0.038)	(0.038)	(0.038)	(−0.328)	(−0.328)	(−2.460)

[a] Unit: au. The serial numbers of the atomic sites are shown in Figure 2. The values in parentheses are the averaged ones among the equivalent methyl hydrogens, i.e., H_5-H_7 for Ethanol, H_5-H_7, and H_8-H_{10} for DMSO. The RESP charges Q_a^{RESP} by Fox and Kollman[12] are also shown for comparison.

noticeably increases. This figure implies that we could use any value of ϵ before the transition, $\epsilon < 10^{-1}$, in terms of the fitting accuracy measured by the root-mean-square deviation. The above discussion in the trimethylamine case is in quite parallel with that in the case of ethanol or DMSO in section 3.1.

The partial charge distribution in part (c) shows a characteristic behavior in the nitrogen site. At $\epsilon = 0$, the nitrogen site has a small partial charge in magnitude, −0.039, while the carbon sites have a much larger (more negative) charge, −0.570. This tendency is contrary to our chemical intuition based on the electronegativity[19,20] or the Mulliken charges. As ϵ increases, the carbon charge is substantially reduced while the nitrogen charge negatively *increases* to exceed the carbon charge at $\epsilon = 0.0048$ and show a negative maximum −0.387 at $\epsilon = 0.05$. The partial charges at the recommended $\epsilon = 0.006$ are $q(N_1) = -0.289$, $q(C_2-C_4) = -0.223$, and $q(H_5-H_{13}) = 0.107$, where the equivalent methyl hydrogens are averaged. The corresponding average RESP charges are −0.330, −0.165, and 0.092, respectively. The appropriate ϵ leads to a well-behaved set of the partial charges with a larger negative charge on the nitrogen, and at the same time substantially suppresses the eigenvalues of the three CRK modes as shown in part a.

4. Concluding Remarks

In the ab initio formulation of the charge response kernel $\partial Q_a/\partial V_b$, a useful quantity to represent the intramolecular electronic polarization, the partial charge at site a, Q_a, is least-squares determined from the electrostatic potential of the outer region of the molecule. However, it is widely known that partial charges thus determined are often not well behaved, particularly when some buried sites are involved. This problem of unstable partial charge fitting could also affect the definition of CRK, because of the contamination by spurious intramolecular charge redistribution. Therefore in this paper, we proposed a modified formulation of the CRK to resolve this problem. The partial charges are redefined using the modified **A** matrix ($\mathbf{A}^{mod}$), which makes the new charge definition as robust as that of the RESP model. The modified charge assignment readily leads to the corresponding CRK formulation using the CPHF equations. The CRK accurately reproduces the polarizability obtained directly via the ab initio MO calculations.

The present formulation is applied to ethanol, DMSO, chloroform and trimethylamine. All these molecules have one or more buried sites, and thus are expected to suffer from the problem of the partial charge fitting. The spurious charge flow

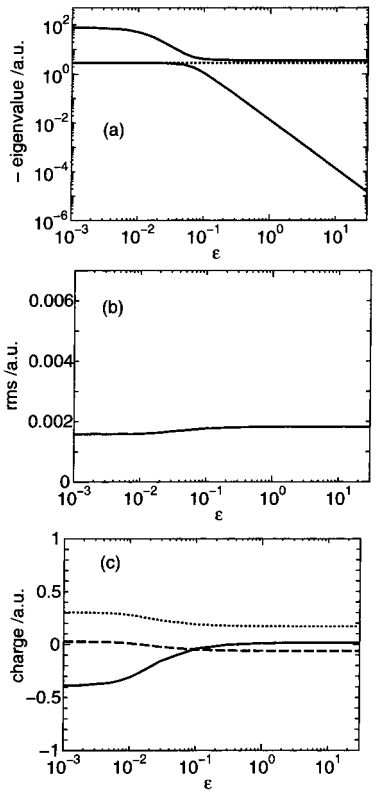

Figure 7. Chloroform. (a) The eigenvalues of the CRK modes. The solid lines denote the A_1 modes, and the dotted line the degenerate E modes. (b) The root-mean-square deviation $\sqrt{L/N_{\text{grid}}}$ for the ESP charges. (c) The site partial charges. The solid line denotes the carbon site, the dash line the three equivalent chlorine sites, and the dotted line the hydrogen.

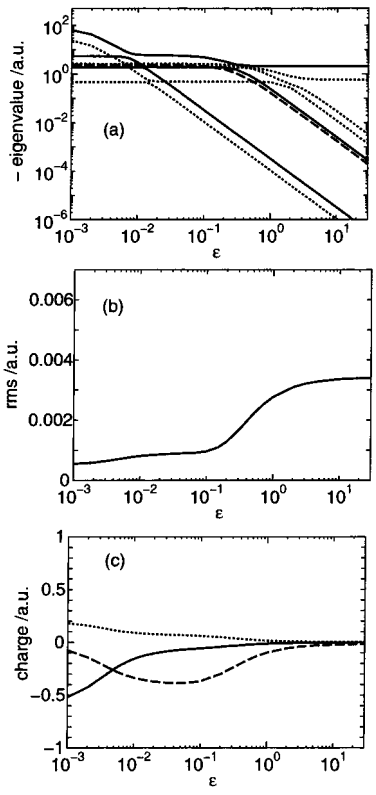

Figure 8. Trimethylamine. (a) The eigenvalues of the CRK modes. The solid lines denote the A_1 modes, the dashed line the A_2 mode, and the dotted lines the degenerate E modes. (b) The root-mean-square deviation $\sqrt{L/N_{\text{grid}}}$ for the ESP charges. (c) The site partial charges. The dash line denotes the nitrogen site N_1, the solid line the three equivalent carbon sites C_2-C_4, and the dotted line the hydrogen sites H_5-H_{13}. The partial charges of the equivalent methyl hydrogens H_5-H_{13} are averaged.

originating from the buried sites is readily distinguished by systematically changing the damping parameter ϵ. The fitting accuracy of the modified charges is almost as good as that of the original charges in the range of $\epsilon = 0-10^{-1}$, and a recommended damping parameter ϵ is then determined to be $\epsilon = 0.006$ at little expense of the fitting accuracy, so that the partial charges are consistent with those of the RESP model.

We note that the recommended damping parameter $\epsilon = 0.006$ should be further examined to establish the transferability to other molecules. In application to another system, the damping parameter ϵ could also be optimized to the system through the eigenvalue analysis of the **A** matrix.

We think that the present work to improve the CRK model is an important step toward general usefulness of the CRK model. We will further make use of this modified CRK model for polarizable MD simulations. In certain cases, the conformational dependence of the CRK should also be considered in the MD application.

5. Acknowledgment

This work was supported by the grant-in-aid from the Ministry of Education and Science, Japan.

Appendix A: Partial Charge Expression

This Appendix supplements the definition of the partial charge Q_a and the operator $\hat{Q}_a$ discussed in section 2. The partial charge fitting in section 2 is accompanied with the constraint on the total charge Q and the dipole moment $\boldsymbol{\mu}$, which are given using eq 6 as follows:

$$Q = \sum_a^{\text{sites}} Q_a = \sum_{a,b}^{\text{sites}} [\mathbf{A}^{-1}]_{ab}\{B_b + \lambda + \boldsymbol{\xi}\cdot\mathbf{R}_s(b)\} \quad (14)$$

$$\boldsymbol{\mu} = \sum_a^{\text{sites}} Q_a \mathbf{R}_s(a) = \sum_{a,b}^{\text{sites}} [\mathbf{A}^{-1}]_{ab}\mathbf{R}_s(a)\{B_b + \lambda + \boldsymbol{\xi}\cdot\mathbf{R}_s(b)\} \quad (15)$$

Now the following two notations **C** and **F** are introduced as

$$C(b) = \sum_a^{\text{sites}} [\mathbf{A}^{-1}]_{ab}, \quad \mathbf{F}(b) = \sum_a^{\text{sites}} [\mathbf{A}^{-1}]_{ab}\mathbf{R}_s(a) \quad (16)$$

Hereafter in this section the summation over b will be employed for the sites. Then the Lagrange multipliers λ, $\xi_x-\xi_z$ in eqs 14 and 15 are expressed in a matrix form

$$\begin{bmatrix} \lambda \\ \xi_x \\ \xi_y \\ \xi_z \end{bmatrix} = \mathbf{G}^{-1} \begin{bmatrix} Q - \sum_b C(b)B_b \\ \mu_x - \sum_b F_x(b)B_b \\ \mu_y - \sum_b F_y(b)B_b \\ \mu_z - \sum_b F_z(b)B_b \end{bmatrix} \quad (17)$$

where

$$\mathbf{G} = \begin{bmatrix} \sum_b C(b) & \sum_b F_x(b) & \sum_b F_y(b) & \sum_b F_z(b) \\ \sum_b F_x(b) & \sum_b F_x(b)R_{sx}(b) & \sum_b F_x(b)R_{sy}(b) & \sum_b F_x(b)R_{sz}(b) \\ \sum_b F_y(b) & \sum_b F_y(b)R_{sx}(b) & \sum_b F_y(b)R_{sy}(b) & \sum_b F_y(b)R_{sz}(b) \\ \sum_b F_z(b) & \sum_b F_z(b)R_{sx}(b) & \sum_b F_z(b)R_{sz}(b) & \sum_b F_z(b)R_{sz}(b) \end{bmatrix} \quad (18)$$

Therefore, the partial charge Q_a in eq 6 is expressed as

$$Q_a = \sum_b [\mathbf{A}^{-1}]_{ab} B_b + [C(a)\, F_x(a)\, F_y(a)\, F_z(a)]\mathbf{G}^{-1} \begin{bmatrix} Q - \sum_b C(b)B_b \\ \mu_x - \sum_b F_x(b)B_b \\ \mu_y - \sum_b F_y(b)B_b \\ \mu_z - \sum_b F_z(b)B_b \end{bmatrix} \quad (19)$$

Next, the partial charge operator $\hat{Q}_a$ in eq 10 is given, which consists of the electronic part $\hat{q}$ and the nuclear part Q^{nuc}. We could also decompose B_b, Q, and $\boldsymbol{\mu}$ in eq 19 into the electronic and nuclear parts as follows:

$$B_b = -\sum_{p,q}^{\mathrm{AO}} D_{pq}\langle p|\hat{b}_b|q\rangle + b_b^{\mathrm{nuc}} \quad (20)$$

$$Q = -\sum_{p,q}^{\mathrm{AO}} D_{pq}\langle p|q\rangle + N^{\mathrm{nuc}} \quad (21)$$

$$\boldsymbol{\mu} = -\sum_{p,q}^{\mathrm{AO}} D_{pq}\langle p|\hat{\mathbf{r}}|q\rangle + \boldsymbol{\mu}^{\mathrm{nuc}} \quad (22)$$

where

$$\hat{b}_b = \sum_n^{\mathrm{grids}} w_n \frac{1}{|\hat{\mathbf{r}} - \mathbf{R}_g(n)|} \cdot \frac{1}{|\mathbf{R}_s(b) - \mathbf{R}_g(n)|} \quad (23)$$

$$b_b^{\mathrm{nuc}} = \sum_n^{\mathrm{grids}} w_n \sum_c^{\mathrm{nuclei}} Z_c \frac{1}{|\mathbf{R}_n(c) - \mathbf{R}_g(n)|} \cdot \frac{1}{|\mathbf{R}_s(b) - \mathbf{R}_g(n)|} \quad (24)$$

and $\hat{\mathbf{r}}$ in eqs 22 and 23 is the electronic position operator. N^{nuc} and $\boldsymbol{\mu}^{\mathrm{nuc}}$ in eqs 21 and 22 are the total charge and the dipole moment of the nuclei; i.e.,

$$N^{\mathrm{nuc}} = \sum_c^{\mathrm{nuclei}} Z_c, \quad \boldsymbol{\mu}^{\mathrm{nuc}} = \sum_c^{\mathrm{nuclei}} Z_c \mathbf{R}_n(c) \quad (25)$$

Thus the electronic part $\hat{q}_a$ and nuclear part Q_a^{nuc} of eq 10 are represented as follows:

$$\hat{q}_a(i) = -\sum_b [\mathbf{A}^{-1}]_{ab}\hat{b}_b(i) - [C(a)\, F_x(a)\, F_y(a)\, F_z(a)]\mathbf{G}^{-1} \begin{bmatrix} 1 - \sum_b C(b)\hat{b}_b(i) \\ \hat{x}(i) - \sum_b F_x(b)\hat{b}_b(i) \\ \hat{y}(i) - \sum_b F_y(b)\hat{b}_b(i) \\ \hat{z}(i) - \sum_b F_z(b)\hat{b}_b(i) \end{bmatrix} \quad (26)$$

$$Q_a^{\mathrm{nuc}} = \sum_b [\mathbf{A}^{-1}]_{ab} b_b^{\mathrm{nuc}} + \{C(a)\, F_x(a)\, F_y(a)\, F_z(a)]\mathbf{G}^{-1} \begin{bmatrix} N^{\mathrm{nuc}} - \sum_b C(b)b_b^{\mathrm{nuc}} \\ \mu_x^{\mathrm{nuc}} - \sum_b F_x(b)b_b^{\mathrm{nuc}} \\ \mu_y^{\mathrm{nuc}} - \sum_b F_y(b)b_b^{\mathrm{nuc}} \\ \mu_z^{\mathrm{nuc}} - \sum_b F_z(b)b_b^{\mathrm{nuc}} \end{bmatrix} \quad (27)$$

References and Notes

(1) Morita, A.; Kato, S. *J. Am. Chem. Soc.* **1997**, *119*, 4021−4032.
(2) Morita, A.; Kato, S. *J. Chem. Phys.* **1998**, *108*, 6809−6818.
(3) Morita, A.; Kato, S. *J. Chem. Phys.* **1998**, *109*, 5511−5523.
(4) Rick, S.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(5) Rick, S.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672−679.
(6) Momany, F. A. *J. Phys. Chem.* **1978**, *82*, 592−601.
(7) Cox, S. R.; Williams, D. E. *J. Comput. Chem.* **1981**, *2*, 304−601.
(8) Sprik, M.; Klein, M. L. *J. Chem. Phys.* **1988**, *89*, 7556−7560.
(9) Bernardo, D. N.; Ding, Y.; K.-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180−4187.
(10) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(11) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(12) Fox, T.; Kollman, P. A. *J. Phys. Chem. B* **1998**, *102*, 8070−8079.
(13) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(14) Gerratt, J.; Mills, I. M. *J. Chem. Phys.* **1968**, *49*, 1719.
(15) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225−241.
(16) Pulay, P. *Adv. Chem. Phys.* **1987**, *69*, 241−286.
(17) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.
(18) Dupuis, M.; Watts, J. D.; Villar, H. O.; Hurst, G. J. B. *HONDO*, version 7.0. *QCPE 544* **1987**.
(19) Pauling, L., Ed.; *The Nature of the Chemical Bond;* Cornell University Press: Ithaca, NY, 1960.
(20) Mulliken, R. S. *J. Chem. Phys.* **1934**, *2*, 782.