Molecular geometry-dependent atomic charge calculation with modified charge equilibration method

Kozo Aoki[1], Shigenori Tanaka[2], Tatsuya Nakano[3*]

[1] *Collaborative Research Center of Frontier Simulation Software for Industrial Science (FSIS), the University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan*
[2] *Graduate School of Science and Technology,Kobe University, 1-1, Rokkodai, Nada-ku, Kobe 657-8501, Japan*
[3] *Division of Medicinal Safety Science, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*

*E-mail: nakano@nihs.go.jp

Abstract

We have improved a modified charge equilibration (MQEq) method for calculating the geometry-dependent distribution of atomic charges. In this paper, Ohno-Klopman, Ohno and DasGupta-Huzinaga equations are adopted to express the shielding effect, and the calculated atomic charges with these MQEq methods are in good agreement with those by the HF/6-31G(d,p) calculations for several organic molecules. These MQEq methods would be useful to estimate the charge distribution for large molecules.

1. Introduction

In the classical force field approach, a non-bonded interaction is represented by the following formula:

$$E_{non-bonded} = E_{short+middle} + E_{es} \qquad (1)$$

A serious problem is that the conventional molecular mechanics and the molecular dynamics simulations use the fixed charges that cannot represent the relaxation of charge distribution. The atomic charges depend upon varied molecular geometry, which is essential to evaluate the appropriate electrostatic energies, , for biomolecular simulation such as receptor-ligand docking.

In 2001, Nakano et al. proposed the modified charge equilibration (MQEq) method [1] for calculating the geometry-dependent distribution of atomic charges with the aid of Nishimoto-Mataga equation [2]. It is not necessary in their method to iterate simultaneous equations for evaluating charge equilibration in contrast to the original QEq method [3]. The MQEq charge is in good agreement with the Mulliken charge obtained by the HF/STO-3G level of theory [1]. On the other hand, Oda and Hirono introduced an atom type for a QEq/PD method [4], and appropriate atomic charges were obtained for various molecules [5,6] with the use of Ohno-Klopman equation [7, 8]. In this study, we compare the Ohno-Klopman [7, 8], Ohno [7], and DasGupta-Huzinaga [9] equations to calculate the Coulomb shielding effect with the Nishimoto-Mataga [2] equation that was employed in the previous paper [1]. We apply various MQEq methods to several organic molecules and some polypeptides, and compare the results with the atomic charges obtained through ab initio MO calculations. Through these calculations, we assess the validity of various MQEq methods.

2. Method

In the original QEq method [3], the electrostatic energy is expressed as a sum of intra-atomic contributions and inter-atomic interactions between pairs of atoms:

$$E_{es} = \sum_{i=1}^{N} \left(E_i^0 + \chi_i^0 q_i + \frac{1}{2} J_{ii}^0 q_i^2 \right) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} J_{ij} q_i q_j , \qquad (2)$$

where N is the number of atoms, q_i is an atomic charge and E_i^0 is the energy of the isolated neutral atom at the ground state. In this work, the Ohno-Klopman (OK) (Eqs. (3) and (4)), Ohno (O) (Eqs. (3) and (5)), DasGupta-Huzinaga (DH) (Eqs. (6) and (7)) equations are used to calculate the Coulomb integrals instead of the Nishimoto-Mataga (NM) (Eqs. (6) and (8)) equation that was used in the previous paper [1],

$$J_{ij} = \frac{1}{\sqrt{R_{ij}^2 + \gamma_{ij}}} \qquad (3)$$

$$\gamma_{ij} = \frac{1}{4}\left(\frac{1}{J_{ii}^0} + \frac{1}{J_{jj}^0}\right)^2 \tag{4}$$

$$\gamma_{ij} = 4\left(\frac{1}{J_{ii}^0 + J_{jj}^0}\right)^2 \tag{5}$$

$$J_{ij} = \frac{1}{R_{ij} + \gamma_{ij}} \tag{6}$$

$$\gamma_{ij} = \frac{2}{J_{ii}^o e^{k_i R_{ij}} + J_{jj}^o e^{k_j R_{ij}}} \qquad (k_i, k_j = 0.4) \tag{7}$$

$$\gamma_{ij} = \frac{2}{J_{ii}^0 + J_{jj}^0} \tag{8}$$

$$\chi_i^0 = \frac{1}{2}(IP + EA) \tag{9}$$

$$J_{ii}^0 = IP - EA, \tag{10}$$

where IP is the ionization potential, EA is the electron affinity, and R_{ij} is the distance between the i-th atom and the j-th atom. In the following calculations, we used the atomic parameters χ_A^0 and J_{AA}^0 in the literature [3] that are independent of molecules.

We find a set of atomic charges giving a stationary point regarding the electrostatic energy with constraints for the total charge of the system and the subsystems.

$$Q_{total} = \sum_{i=1}^{N} q_i = \sum_{I=1}^{N_{mol}} Q_{total}^I \tag{11}$$

$$Q_{total}^I = \sum_{i=N_{I-1}+1}^{N_I} q_i, \tag{12}$$

where N_{mol} is the number of the subsystems and $N_0 = 0$. Without loss of generality, the I-th subsystem consists of atoms $N_{I-1} + 1$ to N_I. We consider the following functional L with multipliers λ_I,

$$L = E_{es} + \sum_{I=1}^{N_{mol}} \lambda_I\left(Q_{total}^I - \sum_{i=N_{I-1}+1}^{N_I} q_i\right). \tag{13}$$

The derivative of the L to q_i is

$$\frac{\partial L}{\partial q_i} = \frac{\partial E_{es}}{\partial q_i} - \lambda_I = \chi_i^0 + J_{ii}^0 q_i + \sum_{j \neq i} J_{ij} q_j - \lambda_I = 0 \quad for\ i \in I \tag{14}$$

Eqs. (12) and (14) give the dense asymmetric simultaneous linear equations:

$$\mathbf{Cq} = \mathbf{D}, \tag{15}$$

where

$$C_{(N_{I-1}+1)j} = \begin{cases} 1 & for \ j \in I \\ 0 & for \ j \notin I \end{cases} \tag{16}$$

$$C_{ij} = J_{(N_{I-1}+1)j} - J_{ij} \quad for \ i \in I \ and \ i \neq N_{I-1}+1$$

and

$$D_{N_{I-1}+1} = Q^I_{total} \tag{17}$$

$$D_i = \chi^0_i - \chi^0_{N_{I-1}+1} \quad for \ i \in I \ and \ i \neq N_{I-1}+1$$

Since the Ohno-Klopman, Ohno, DasGupta-Huzinaga and Nishimoto-Mataga equations do not contain the atomic charges, q_i, it is not necessary to iterate the simultaneous equations for evaluating a charge equilibration.

Atomic charge calculations were carried out for the following molecules: formaldehyde, acetaldehyde, formamide, formic acid, acetic acid, propane, methylamine, methanol, ethanol, toluene, phenol, and three (α-helix, β-strand and extended) conformers of polyalanine with five residues, (Ala)$_5$. Molecular structures of small molecules were taken from the HF/6-31G(d,p) optimized geometries. Three conformers of (Ala)$_5$ were modeled with the Accelrys InsightII molecular modeling system [10]. The Hartree-Fock-Roothaan [11] calculation with STO-3G [12] and 6-31G(d,p) [13-15] basis sets were performed for those organic molecules and the three conformers of (Ala)$_5$. The Mulliken [16], the CHELPG [17], the Merz-Singh-Kollman (MSK) [18, 19] and the original QEq charges were calculated using the Gaussian98 program package [20].

3. Results and Discussion

Table 1 shows the standard deviation of the QEq and MQEq atomic charges from some reference charges of the small organic molecules including the H, C, N, and O atoms. The MQEq (NM) atomic charges are in good agreement with the Mulliken charges obtained by the HF/STO-3G calculations, while the MQEq(OK, O and DH) charges are close to the Mulliken, the CHELPG and the MSK charges at the HF/6-31G(d,p) level of theory.

Milner-White [21] discussed the electronic character of formamide that is a model of peptide bond. The classical formula of peptide bond has the two resonance formulas as shown in Figure 1, in one of which the nitrogen atom has a positive charge (b).
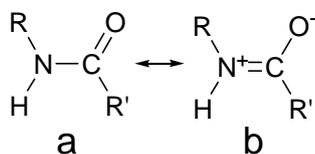

Figure 1. Resonance structure of peptide bond.

It is shown that the partial charge of the nitrogen of amides is negative rather than positive from the resonance forms by the quantum mechanical calculations [22]. As shown in Table 2, the MQEq(NM) charges for formamide are in good agreement with the HF/STO-3G charges. The three MQEq(OK,O,DH) charges are very close to the CHELPG and the MSK charges with the strong polar character at the HF/6-31G(d,p) level of theory, but the HF/STO-3G charge does not describe the polar amide bond.

Table 1. Standard deviation of QEq and MQEq atomic charges of several organic molecules from the reference charges.

Formaldehyde	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.11	0.10	0.24	0.20
QEq	0.16	0.07	0.21	0.17
MQEq(OK)	0.21	0.04	0.12	0.08
MQEq(O)	0.22	0.06	0.10	0.07
MQEq(DH)	0.24	0.07	0.09	0.06

Acetaldehyde	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.09	0.18	0.25	0.27
QEq	0.14	0.07	0.20	0.19
MQEq(OK)	0.24	0.07	0.13	0.09
MQEq(O)	0.27	0.11	0.15	0.11
MQEq(DH)	0.34	0.18	0.19	0.14

Formamide	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.08	0.30	0.47	0.45
QEq	0.15	0.13	0.30	0.28
MQEq(OK)	0.36	0.12	0.07	0.05
MQEq(O)	0.38	0.14	0.05	0.03
MQEq(DH)	0.39	0.15	0.05	0.06

Formic acid	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.10	0.21	0.32	0.30
QEq	0.27	0.09	0.18	0.16
MQEq(OK)	0.55	0.29	0.19	0.20
MQEq(O)	0.59	0.34	0.22	0.24
MQEq(DH)	0.57	0.32	0.20	0.22

Acetic acid	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.09	0.24	0.31	0.32
QEq	0.20	0.11	0.16	0.17
MQEq(OK)	0.46	0.22	0.18	0.16
MQEq(O)	0.49	0.26	0.21	0.19
MQEq(DH)	0.54	0.30	0.27	0.24

Table 1. (*continued*)

Propane	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.02	0.11	0.14	0.15
QEq	0.18	0.09	0.28	0.25
MQEq(OK)	0.11	0.04	0.19	0.16
MQEq(O)	0.15	0.07	0.21	0.18
MQEq(DH)	0.15	0.08	0.20	0.16

Methylamine	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.02	0.17	0.39	0.37
QEq	0.21	0.12	0.42	0.39
MQEq(OK)	0.19	0.04	0.29	0.26
MQEq(O)	0.20	0.05	0.31	0.28
MQEq(DH)	0.17	0.02	0.31	0.28

Methanol	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.09	0.09	0.21	0.20
QEq	0.24	0.11	0.28	0.26
MQEq(OK)	0.29	0.12	0.17	0.15
MQEq(O)	0.29	0.12	0.17	0.15
MQEq(DH)	0.26	0.08	0.17	0.14

Ethanol (trans)	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.07	0.12	0.22	0.20
QEq	0.20	0.10	0.26	0.23
MQEq(OK)	0.26	0.11	0.15	0.12
MQEq(O)	0.28	0.13	0.16	0.13
MQEq(DH)	0.27	0.13	0.16	0.13

Toluene	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.02	0.10	0.13	0.19
QEq	0.08	0.04	0.11	0.13
MQEq(OK)	0.09	0.06	0.09	0.11
MQEq(O)	0.11	0.07	0.10	0.10
MQEq(DH)	0.13	0.10	0.10	0.08

Table 1. (*continued*)

Phenol	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MSK
MQEq(NM)	0.06	0.14	0.21	0.23
QEq	0.15	0.08	0.16	0.18
MQEq(OK)	0.23	0.10	0.11	0.12
MQEq(O)	0.24	0.10	0.11	0.12
MQEq(DH)	0.25	0.11	0.08	0.09

Table 2. Calculated atomic charges of formamide.

	MQEq				QEq	HF/STO-3G	HF/6-31G(d,p)		
	OK	O	DH	NM		Mulliken	Mulliken	CHELPG	MSK
O	-0.56	-0.57	-0.60	-0.34	-0.42	-0.24	-0.56	-0.62	-0.59
C	0.67	0.71	0.78	0.16	0.33	0.23	0.56	0.79	0.72
N	-0.91	-0.94	-0.91	-0.29	-0.63	-0.41	-0.73	-0.99	-0.99
H	0.01	-0.01	-0.02	0.13	0.16	0.06	0.10	-0.02	0.01
H	0.43	0.44	0.41	0.19	0.30	0.18	0.32	0.44	0.44
H	0.36	0.37	0.33	0.15	0.25	0.18	0.31	0.40	0.41

To confirm the capability of the MQEq(OK,O,DH) method for the polypeptide system, we calculated the charge distribution of the three (α-helix, β-strand and extended) conformers of (Ala)$_5$. Table 3 shows the standard deviation of atomic charges of the three conformers of (Ala)$_5$. It is found that the peptide bond's heavy atoms (N and O) are very polar from the HF/6-31G(d,p) calculations. The MQEq(NM) charges are in good agreement with the HF/STO-3G charges for the peptide. Since the HF/STO-3G charge does not describe the strong polar character, the MQEq(NM) method should not be adopted for the peptide molecule. On the other hand, the MQEq(OK) charges are close to the Mulliken, the CHELPG and the MSK charges at the HF/6-31G(d,p) level of theory. Thus the MQEq(OK) method is appropriate for atomic charge calculations of biomolecules.

Table 3. Standard deviation of QEq and MQEq atomic charges from the reference charges on the α-helix, β-strand, and extended conformers of (Ala)$_5$.

α-helix	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MKS
MQEq(NM)	0.06	0.28	0.28	0.26
QEq	0.16	0.13	0.18	0.17
MQEq(OK)	0.25	0.06	0.15	0.19
MQEq(O)	0.27	0.07	0.16	0.20
MQEq(DH)	0.29	0.08	0.17	0.22

β-strand	HF/STO-3G	HF/6-31G(d,p)		
	Mulliken	Mulliken	CHELPG	MKS
MQEq(NM)	0.06	0.27	0.31	0.32
QEq	0.15	0.14	0.20	0.19
MQEq(OK)	0.24	0.07	0.14	0.12
MQEq(O)	0.26	0.08	0.15	0.12
MQEq(DH)	0.29	0.09	0.15	0.12

Table 3. (*continued*)

| extended | HF/STO-3G | HF/6-31G(d,p) | | |
	Mulliken	Mulliken	CHELPG	MKS
MQEq(NM)	0.07	0.28	0.33	0.33
QEq	0.15	0.15	0.23	0.22
MQEq(OK)	0.25	0.07	0.16	0.15
MQEq(O)	0.27	0.08	0.17	0.15
MQEq(DH)	0.29	0.08	0.17	0.16

Figure 2 shows shielded Coulomb potentials between two carbon atoms (J_{C-C}). As shown in Figure 2, the screening effect strengthens in order of OK, DH, and NM. A similar tendency was seen about nitrogen and oxygen atom pairs (Figure 3 and 4). The difference of this screening effect is one of the reasons for the above results.
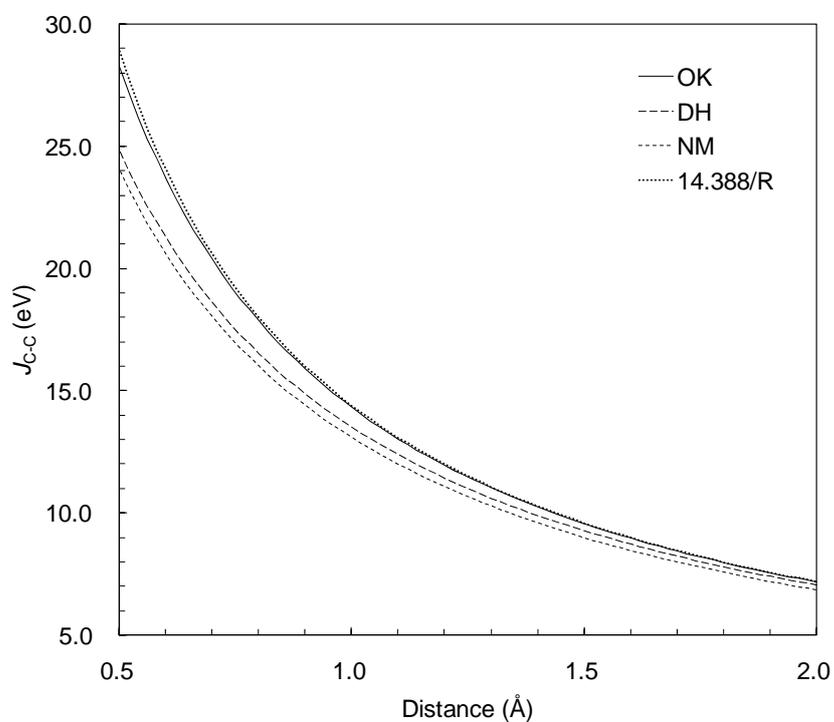
Figure 2. Shielded potentials between two carbon atoms. Abbreviation is as follows; OK: Ohno-Klopman, DH: DasGuputa-Huzinaga, NM: Nishimoto-Mataga, 14.338/R: unshielded Coulomb potential.
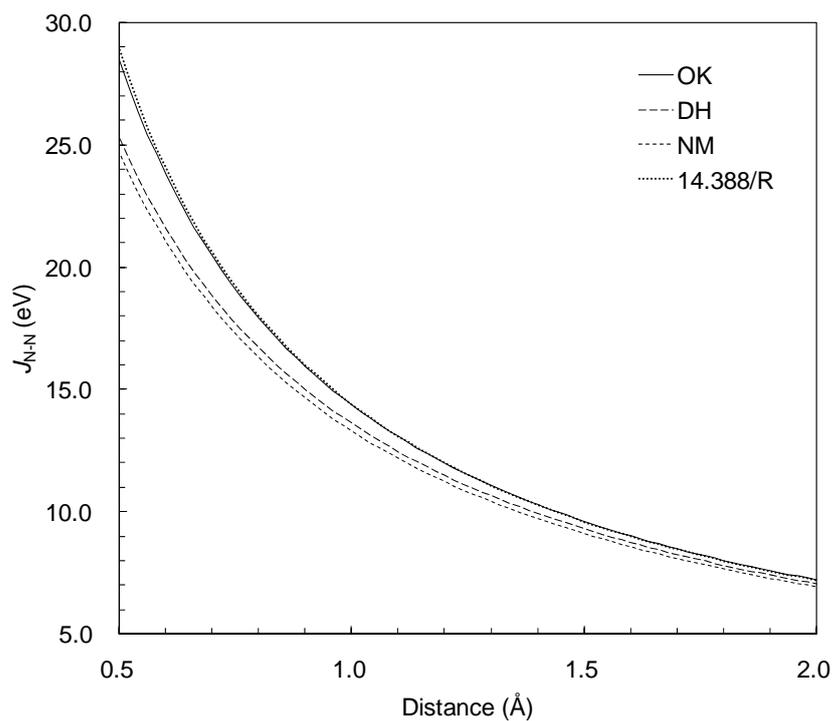
Figure 3. Shielded potentials between two nitrogen atoms. Abbreviation is as follows; OK: Ohno-Klopman, DH: DasGuputa-Huzinaga, NM: Nishimoto-Mataga, 14.338/R: unshielded Coulomb potential.
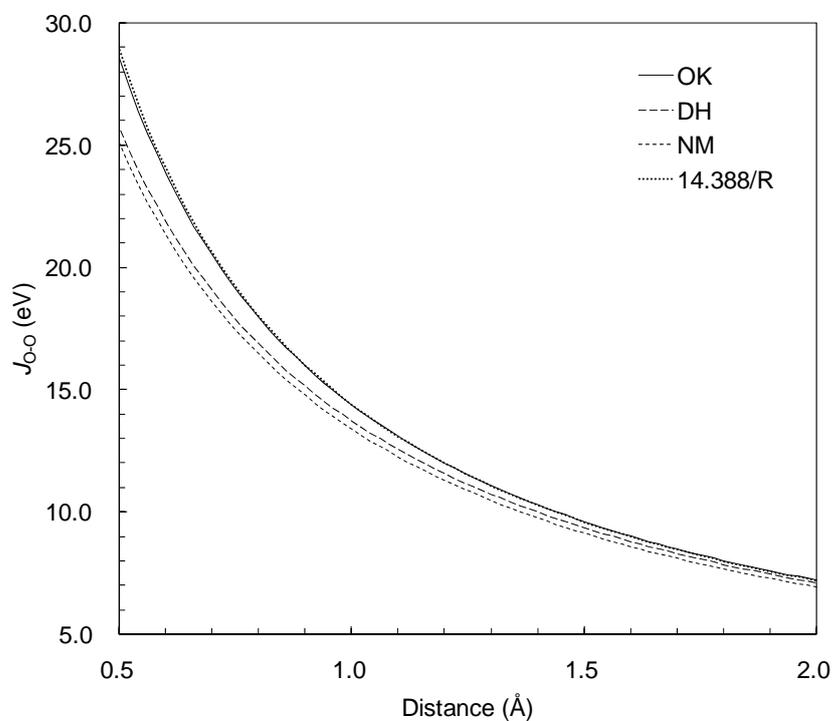

Figure 4. Shielded potentials between two oxygen atoms. Abbreviation is as follows; OK: Ohno-Klopman, DH: DasGuputa-Huzinaga, NM: Nishimoto-Mataga, 14.338/R: unshielded Coulomb potential.

4. Concluding Remarks

Since the calculated atomic charges with the MQEq (Ohno-Klopman, Ohno, and DasGupta-Huzinaga) methods are in good agreement with those using the HF/6-31G (d,p) calculations for the small organic molecules and peptides, it would be useful to calculate the protein's electrostatic potential and an electric field with these methods, appropriately accounting for polarization effects. In addition, the MQEq calculation is faster than the original QEq procedure since the simultaneous linear equations do not need to be solved iteratively. It is also remarked that analytical gradient of electrostatic energy which can be calculated easily in the MQEq method is crucial to perform the geometry optimization and molecular dynamics simulations. Furthermore, by introducing explicit atom types and the parameter optimization of χ_A^0 and J_{AA}^o with MSK charges obtained by MP2 calculations, the MQEq method is expected to allow the more accurate calculation of atomic charges. Such studies are in progress.

References

[1] [T. Nakano, T. Kaminuma, M. Uebayasi, and Y. Nakata, 3D Structure Based Atomic Charge Calculation for Molecular Mechanics and Molecular Dynamics Simulations, *CBIJ*, **1**, 35-40 (2001).

[2] K. Nishimoto, A MO Theoretical Study of Organic Dyes I. Effect of Chemical Softness on the Electronic Spectra, *Bull. Chem. Soc. Jpn.*, **66**, 1876-1880 (1993) and references therein.

[3] A. K. Rappé and W. A. Goddard III, Charge Equilibration for Molecular Dynamics Simulations, *J. Phys. Chem.*, **95**, 3358-3363 (1991).

[4] D. Bakowies and W. Thiel, Semiempirical Treatment of Electrostatic Potentials and Partial Charges in Combined Quantum Mechanical and Molecular Mechanical Approaches, *J. Comput. Chem.*, **17**, 87-108 (1996).

[5] A. Oda and S. Hirono, The Introduction of Atom Type and Calculations of New Parameters for Charge Equilibration Method, *J. Comput. Chem. Jpn.*, **2**, 17-26 (2003).

[6] A. Oda and S. Hirono, Geometry-dependent atomic charge calculations using charge equilibration method with empirical two-center Coulombic terms, *J. Mol. Struct. (THEOCHEM)*, **634**, 159-170 (2003).

[7] K. Ohno, Some Remarks on the Pariser-Parr-Pople Method, *Theoret. Chim. Acta (Berl.)*, **2**, 219-227 (1964).

[8] G. Klopman, A Semiempirical Treatment of Molecular Structures. III. Equipotential Orbitals for Polyatomic Systems, *J. Am. Chem. Soc.*, **87**, 3300-3303 (1965).

[9] A. DasGupta and S. Huzinaga, New Developments in CNDO Molecular Orbital Theory, *Theoret. Chim. Acta (Berl.)*, **35**, 329-340 (1974).

[10] InsightII Version 98.0 (Accelrys, San Diego, 1998).

[11] C. C. J. Roothaan, New Developments in Molecular Orbital Theory, *Rev. Mod. Phys.*, **23**,

69-89 (1951).

[12] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals, *J. Chem. Phys.*, **51**, 2657-2664 (1969).

[13] R. Ditchfield, W. J. Hehre, and J. A. Pople, Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules, *J. Chem. Phys.*, **54**, 724-728 (1971).

[14] W. J. Hehre, R. Ditchfield, and J. A. Pople, Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules, *J. Chem. Phys.*, **56**, 2257-2261 (1972).

[15] P. C. Hariharan and J. A. Pople, The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies, *Theoret. Chim. Acta (Berl.)*, **28**, 213-222 (1973).

[16] R. S. Mulliken, Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I, *J. Chem. Phys.*, **23**, 1833-1840 (1955).

[17] C. M. Breneman and K. B. Wiberg, Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis, *J. Comput. Chem.*, **11**, 361-373 (1990).

[18] B. H. Besler, K. M. Merz, Jr. and P. A. Kollman, Atomic Charges Derived from Semiempirical Methods, *J. Comput. Chem.*, **11**, 431-439 (1990).

[19] U. C. Singh and P. A. Kollman, An Approach to Computing Electrostatic Charges for Molecules, *J. Comput. Chem.*, **5**, 129-145 (1984).

[20] M. J. Frisch et al., Gaussian 98 (Revision A.9), Gaussian, Inc., Pittsburgh, PA, 1998.

[21] E. J. Milner-White, The partial charge of the nitrogen atom in peptide bonds, *Protein Sci.*, **6**, 2477-2482 (1997).

[22] K. B. Wiberg, C. M. Hadad, P. R. Rablen, and J. Cioslowski, Substituent Effects. 4. Nature of Substituent Effects at Carbonyl Groups, *J. Am. Chem. Soc.*, **114**, 8644-8654 (1992).