BioChem Press

Inter*net* Electronic Journal of
Molecular Design

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Haruo Hosoya on the occasion of the 65th birthday
Part 4

Guest Editor: Jun–ichi Aihara

Incorporation of Solvent Effects into Charge Equilibration Approach by Using Generalized Born Formula

Naofumi Nakayama[1] and Umpei Nagashima[2]

[1] Best Systems Inc., 2–1–6 Sengen, Tsukuba, Ibaraki 305–0047, Japan
[2] National Institute of Advanced Industrial Science and Technology, 1–1–1 Higashi, Tsukuba, Ibaraki 305–8562, Japan

Citation of the article:
N. Nakayama and U. Nagashima, Incorporation of Solvent Effects into Charge Equilibration Approach by Using Generalized Born Formula, *Internet Electron. J. Mol. Des.* **2002**, *1*, 659–667, http://www.biochempress.com.

Incorporation of Solvent Effects into Charge Equilibration Approach by Using Generalized Born Formula[#]

Naofumi Nakayama[1],* and Umpei Nagashima[2]

[1] Best Systems Inc., 2–1–6 Sengen, Tsukuba, Ibaraki 305–0047, Japan
[2] National Institute of Advanced Industrial Science and Technology, 1–1–1 Higashi, Tsukuba, Ibaraki 305–8562, Japan

Abstract
Motivation. The charge equilibrium (QEq) approach can treat isolated systems, but no effects from external fields are considered. However, in molecular simulations, it is often desirable to include electrostatic interactions with solvents, especially simulations for biological molecules. Therefore, the QEq approach including solvent effects is effective for molecular simulations.
Method. We developed a QEq approach that incorporates solvent effects by using the generalized Born (GB) model, and call it the QEq–GB approach.
Results. The QEq–GB approach indicates that dipole moments for certain organic compounds increase with increasing dielectric constant, which is the expected behavior.
Conclusions. The QEq–GB approach can be used in molecular–mechanics and molecular–dynamics calculations to calculate the charge distribution of molecular systems in solvents.
Keywords. Charge equilibration approach; solvent effect; generalized Born formula; dipole moment.

Abbreviations and notations	
QEq, charge equilibration	GB, generalized Born
PM3, parametric method 3	

1 INTRODUCTION

The charge distribution within molecules must be ascertained for determining the electrostatic energies in molecular mechanics and molecular dynamics calculations. In conventional molecular simulations, however, changes in molecular structure or in electrostatic environment prevent accurate representation of the relaxation of the charge distribution.

Rappé and Goddard proposed a charge equilibration (QEq) approach for predicting charge distributions that depend on the molecular geometry [1]. They calculated charge distributions and

dipole moments of organic, inorganic, biological, and polymer systems. Their calculated dipole moments agree well with measured dipole moments [1].

The QEq approach described in [1] can treat isolated systems, but no effects from external fields are considered. However, in molecular simulations, it is often desirable to include electrostatic interactions with solvents, especially for simulations including biological molecules. Therefore, for molecular simulations, a QEq approach that includes solvent effects will be effective.

To improve the applicability of the QEq approach, we developed a QEq approach that incorporates solvent effects. We used the GB model to represent the solvation free–energy because the energy expressed by the generalized Born (GB) model (a continuum solvation model) is only dependent on the partial charges of molecules [2–4]. We therefore call this the QEq–GB approach. Here, we show its formulation and application to calculate dipole moments and hydration free–energies of organic compounds, such as aldehydes, acids, amides, amines, and alcohols.

2 COMPUTATIONAL METHODS

2.1 Original QEq Approach

In the original QEq approach, the total electrostatic energy is represented as:

$$E^{\text{vac}}(Q_1 \cdots Q_N) = \sum_A E_A(Q_A) + \sum_{A<B} Q_A Q_B J_{AB} \tag{1}$$

where $E_A(Q_A)$ is the energy of atom A, and can be expressed as:

$$E_A(Q_A) = E_{A0} + \chi_A^0 Q_A + \frac{1}{2} J_{AA}^0 Q_A^2 \tag{2}$$

$$\chi_A^0 = \frac{1}{2}(IP_A + EA_A) \tag{3}$$

$$J_{AA}^0 = IP_A - EA_A \tag{4}$$

where *IP* and *EA* represent the ionization potential and electron affinity, respectively. χ_A^0 and J_{AA}^0 represent the electronegativity and self–Coulomb repulsion, respectively, and their values are the same as those given in [1]. J_{AB} is the Coulomb interaction between unit charges A and B.

To calculate J_{AB}, Rappé and Goddard described the atomic density in terms of a single s–type Slater orbital to correct for the shielding effect [1]. In addition, in their QEq model only the orbital exponent of H is allowed to be charge–dependent [1]. Therefore, to evaluate the charge equilibration, Slater–type overlap integrals must be calculated, and simultaneous equations must be solved iteratively when H atoms are included in a system.

Recently, Nakano *et al.* reported dipole moments of organic compounds calculated by using a

QEq approach [5] that uses the Nishimoto–Mataga expression [6] for evaluating Coulomb integrals as follows:

$$J_{AB} = \frac{1}{R_{AB} + \dfrac{2}{J_{AA}^0 + J_{BB}^0}} \text{ (a.u.)} \tag{5}$$

where R_{AB} is the distance between atoms A and B. Because Eq. (5) does not contain information on the atomic charges, to evaluate the charge equilibration, simultaneous equations are solved directly without iterating. In addition, the correlation between measured dipole moments and dipole moments calculated by using Nakano's method was nearly the same as the correlation between measured dipole moments and dipole moments calculated by using the Gaussian98 program [1,7] (see Sec.3.1). In this study, we used Eq. (5) for evaluating Coulomb integrals.

Differentiating Eq. (1) with respect to Q_A leads to an expression for the atomic–scale chemical potential:

$$\chi_A^{\text{vac}}(Q_1 \cdots Q_N) = \frac{\partial E^{\text{vac}}}{\partial Q_A} = \chi_A^0 + \sum_B J_{AB} Q_B \tag{6}$$

At equilibrium, the requirement is that the atomic chemical potentials be equal, thus leading to the following $N - 1$ conditions:

$$\chi_1 = \chi_2 = \cdots = \chi_N \tag{7}$$

The condition on the total charge:

$$Q_{\text{total}} = \sum_{i=1}^N Q_i \tag{8}$$

thus leads to a total of N simultaneous equations:

$$\mathbf{C}^{\text{vac}} \mathbf{Q} = \mathbf{D}^{\text{vac}} \tag{9}$$

where

$$C_{1j}^{\text{vac}} = 1 \tag{10}$$

$$C_{ij}^{\text{vac}} = J_{1j} - J_{ij} \text{ for } i \geq 2 \tag{11}$$

and

$$D_1^{\text{vac}} = Q_{\text{total}} \tag{12}$$

$$D_i^{\text{vac}} = \chi_i^0 - \chi_1^0 \text{ for } i \geq 2 \tag{13}$$

2.2 Incorporation of Solvent Effects into the QEq Approach

In our QEq–GB approach, we used the generalized Born model to incorporate solvent effects into the QEq approach. In the formulation of the QEq–GB approach, the total energy in solution is expressed as:

$$E^{\text{sol}}(Q_1 \cdots Q_N) = E^{\text{vac}}(Q_1 \cdots Q_N) + \Delta G_{\text{sol}} \tag{14}$$

The solvation free–energy of a molecule can be expressed as:

$$\Delta G_{\text{sol}} = -\frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\sum_A \sum_B Q_A Q_B \Gamma_{AB} \tag{15}$$

where ε is the dielectric constant of the solvent, A and B represent atoms in the molecule, and Q_A and Q_B are the partial charges on atoms A and B, respectively. Γ_{AA} can be expressed as

$$\Gamma_{AA} = \frac{1}{r_A} \tag{16}$$

where r_A is the atomic radius of A. Γ_{AB} is the two–center repulsive energy between unit charges on atoms A and B, and can be evaluated by using Ohno–Klopman–type expression:

$$\Gamma_{AB} = \frac{1}{\sqrt{R_{AB}^2 + \left\{\frac{1}{2}(r_A + r_B)\right\}^2}} \tag{17}$$

where R_{AB} is the distance between atoms A and B. The atomic radius r_A can be expressed as a function of Q_A as:

$$r_A = \alpha_A \exp(-\beta_A Q_A) \tag{18}$$

In this study, the values of α and β determined by Takahashi *et al.* [4] were used to calculate r_A for H, C, N, and O atoms (Table 1).

Table 1. Parameters used for calculating atomic radius r_A.

Atom	α (Å)	β
H	1.36	0.303
C	1.85	0.125
N	1.53	0.101
O	1.46	0.086

Differentiating E^{sol} with respect to Q_A leads to similar forms of Eq. (6) for the solvent:

$$\chi_A^{\text{sol}}(Q_1 \cdots Q_N) = \frac{\partial E^{\text{sol}}}{\partial Q_A} = \chi_A^0 + \sum_B J_{AB} Q_B + \frac{\partial \Delta G_{\text{sol}}}{\partial Q_A} \tag{19}$$

$$\frac{\partial \Delta G_{\text{sol}}}{\partial Q_A} = -\left(1 - \frac{1}{\varepsilon}\right)\left(\sum_B Q_B \Gamma_{AB} + \sum_B Q_A Q_B \frac{\partial \Gamma_{AB}}{\partial Q_A} - \frac{1}{2}Q_A^2 \frac{\partial \Gamma_{AA}}{\partial Q_A}\right) \tag{20}$$

$$\frac{\partial \Gamma_{AA}}{\partial Q_A} = \beta_A \Gamma_{AA} \tag{21}$$

$$\frac{\partial \Gamma_{AB}}{\partial Q_A} = \frac{1}{4} \Gamma_{AB}^3 (r_A + r_B) \beta_A r_A \tag{22}$$

At equilibrium, the constraints on Eqs. (19) to (22) are similar to those on Eqs. (7) and (8). Therefore,

$$\mathbf{C}^{\text{sol}} \mathbf{Q} = \mathbf{D}^{\text{sol}} \tag{23}$$

where

$$C_{1j}^{\text{sol}} = C_{1j}^{\text{vac}} \tag{24}$$

$$C_{ij}^{\text{sol}} = C_{ij}^{\text{vac}} - \left(1 - \frac{1}{\varepsilon}\right)\left(\Gamma_{1j} - \Gamma_{ij} + Q_1 \frac{\partial \Gamma_{1j}}{\partial Q_1} - Q_i \frac{\partial \Gamma_{ij}}{\partial Q_i}\right) \text{ for } i \geq 2 \tag{25}$$

and

$$D_1^{\text{sol}} = D_1^{\text{vac}} \tag{26}$$

$$D_i^{\text{sol}} = D_i^{\text{vac}} - \frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\left(Q_1^2 \frac{\partial \Gamma_{11}}{Q_1} - Q_i^2 \frac{\partial \Gamma_{ii}}{\partial Q_i}\right) \text{ for } i \geq 2 \tag{27}$$

In addition, it is necessary to account for reduced interaction of charged atoms in a molecule with the atoms in a solvent compared to the interaction that is expected when the atom is regarded as a mono–atomic ion. Therefore, in the QEq–GB approach, we also use steric factors similar to those described in Eqs. (16) and (17):

$$\Gamma_{AA} = \frac{f_A}{r_A} \tag{28}$$

$$\Gamma_{AB} = \frac{1}{\sqrt{R_{AB}^2 + \left\{\frac{1}{2}\left(\frac{r_A}{f_A} + \frac{r_B}{f_B}\right)\right\}^2}} \tag{29}$$

where f_A is

$$f_A = \frac{S_A}{4\pi r_A^2} \tag{30}$$

where S_A is the exposed surface area of atom A in the molecule.

---

http://www.biochempress.com

2.3 Iteration Procedure and Molecular Geometry

In the QEq–GB approach, an iterative calculation procedure is necessary because the elements of the $\mathbf{C}^{\text{sol}}$ matrix and the $\mathbf{D}^{\text{sol}}$ vector contain information on atomic charges. The initial values of the elements of the $\mathbf{Q}$ vector are calculated by using Eq. (9):

$$\mathbf{Q}^{(0)} = \left(\mathbf{C}^{\text{vac}}\right)^{-1} \mathbf{D}^{\text{vac}} \tag{31}$$

Within each iteration, the differences in the elements of $\mathbf{Q}$ between steps n and $n-1$ are calculated as:

$$\left(\mathbf{Q}^{\text{diff}}\right)_i = \left|\left(\mathbf{Q}^{(n)}\right)_i - \left(\mathbf{Q}^{(n-1)}\right)_i\right| \tag{32}$$

The iterative procedure is considered converged when $\mathbf{Q}^{\text{diff}} < 5.0 \times 10^{-4}$ for all elements.

The molecular geometries of all compounds are calculated by using the PM3 method with the Gaussian98 program package [7].

3 RESULTS AND DISCUSSION

3.1 Comparing the Results Obtained by QEq on Gaussian98 with those Obtained by Nakano's Method

To compare with the implementation of QEq by Rappé and Goddard in Gaussian98 package and with Nakano's method, we calculated dipole moments for the 11 molecules listed in Table 2. Figure 1 shows the correlation between measured [8] and calculated values. The correlation between measured dipole moments and dipole moments calculated by using Nishimoto–Mataga's expression [6] for evaluating the Coulomb integrals was similar to the correlation between measured values and values calculated by using the method by Rappé and Goddard [1].

Table 2. Measured and calculated dipole moments (Debye) and computational time (sec) to calculate dipole moments.

Compound	Dipole moment (Debye)			Computational time (sec)	
	Measured	Gaussian98	Nakano	Gaussian98	Nakano
Formaldehyde	2.3315	3.2580	2.6878	1.5	0.048
Acetaldehyde	2.750	3.5734	3.3561	1.5	0.048
Formic acid	1.4214	1.3664	1.2424	1.5	0.047
Acetic acid	1.70	2.6302	2.8291	1.6	0.047
Formamide	3.711	3.0712	2.8833	1.5	0.047
Propane	0.0841	0.0002	0.0391	1.6	0.048
Methylamine	1.27	1.4102	0.7894	1.5	0.047
Methanol	1.66	2.7530	1.9416	1.5	0.049
Ethanol	1.441	2.5527	1.9607	1.6	0.047
Toluene	0.375	0.0985	0.3235	1.6	0.048
Phenol	1.224	2.7887	2.3414	1.6	0.048

Table 2 shows the CPU time required for calculating dipole moments with an SGI Octane (300 MHz CPU, 2.5 GB RAM). Nakano's method was more than 30 times faster than the Gaussian98

program. Because both methods yield similar values, Nakano's method is therefore superior for calculating dipole moments because of its higher computational efficiency. We therefore used Nakano's method in our QEq–GB approach.
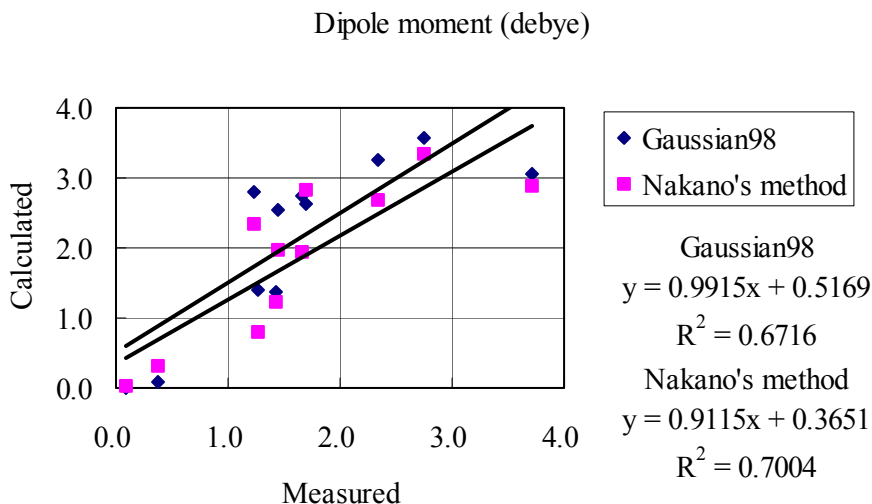
Dipole moment (debye)

Figure 1. Correlation between measured and calculated dipole moments.

3.2 Dipole Moments in Solution

Table 3 shows calculated dipole moments of some organic compounds for ε = 1.0, 2.0, 5.0, 10.0, 50.0, and 80.0. The dipole moments increase with increasing ε for all compounds except methylamine. In methylamine, however, because the charge distribution for ε = 80 is more polarized than that in vacuo, the polarization of methylamine is achieved in CH_3 and NH_2 groups.

Table 3. Calculated dipole moments (Debye) for various compounds.

Compound	Dielectric constant (ε)					
	1.0	2.0	5.0	10.0	50.0	80.0
Formaldehyde	2.6878	2.8424	2.9441	2.9795	3.0085	3.0112
Acetaldehyde	3.3561	3.7028	3.9455	4.0333	4.1060	4.1127
Formic acid	1.2424	1.3472	1.4198	1.4461	1.4679	1.4699
Acetic acid	2.8291	3.2001	3.4773	3.5810	3.6695	3.6772
Formamide	2.8833	3.2839	3.5840	3.6963	3.7928	3.8015
Methylamine	0.7894	0.7771	0.7697	0.7676	0.7662	0.7662
Methanol	1.9416	1.9890	2.0194	2.0298	2.0382	2.0390
Ethanol	1.9607	2.0462	2.1040	2.1244	2.1413	2.1429
Phenol	2.3414	2.6567	2.8957	2.9867	3.0638	3.0713

Increasing dipole moment with increasing ε is expected because the absolute value of the solvation free–energy of molecules given by Eq. (15) increases with increasing molecular polarization.

Table 4. Charge distribution of methylamine (see Figure 2).

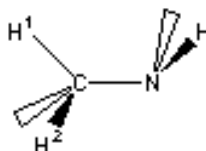| Atom	In vacuo	$\varepsilon = 80$
C	−0.091	−0.116
N	−0.340	−0.374
H^1(–C)	+0.056	+0.061
H^2(–C)	+0.065	+0.067
H(–N)	+0.123	+0.147
CH_3	+0.095	+0.080

Figure 2. Numbering of atoms in methylamine.

In all cases, our calculated dipole moments by using the QEq–GB approach converged within 4 iterations. Therefore, the QEq–GB approach is computationally efficient for molecular–mechanics and molecular–dynamics calculations to obtain charge distributions of molecular systems in solvents.

Acknowledgment

4 REFERENCES

[1]   A. K. Rappé and W. A. Goddard III, Charge Equilibration for Molecular Dynamics Simulations, *J. Phys. Chem.* **1991**, 95, 3358–3363.
[2]   C. J. Cramer and D. G. Truhlar, Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics, *Chem. Rev.* **1999**, 99, 2161–2200.
[3]   O. Kikuchi, T. Matsuoka, H. Sawahata, and O. Takahashi, Ab Initio Molecular Orbital Calculations Including Solvent Effect by Generalized Born Formula. Conformation of Zwitterionic Forms of Glycine, Alanine, and Serine in Water, *J. Mol. Struct.* (*Theochem*) **1994**, 305, 79–87.
[4]   O. Takahashi, H. Sawahata, Y. Ogawa, O. Kikuchi, *J. Mol. Struct.* (*Theochem*) **1997**, 393, 141–150.
[5]   T. Nakano, T. Kaminuma, M. Uebayashi, and Y. Nakata, 3D Structure Based Atomic Charge Calculation for Molecular Mechanics and Molecular Dynamics Simulations, *Chem–Bio Informatics J.* **2001**, 1, 35–40.
[6]   K. Nishimoto and N. Mataga, *Z. Physik. Chem. Neue Folge* **1957**, 12, 335–338.
[7]   M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al–Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head–Gordon, E. S. Replogle, and J. A. Pople, Gaussian 98, Revision A.9, Gaussian, Inc., Pittsburgh PA, 1998.
[8]   The Chemical Society of Japan, Kagaku Binran (Chemical Handbook), 4th ed., Maruzen, Tokyo, 1993, II–575.

Biographies

Naofumi Nakayama is a research scientist at Best Systems Inc., Tsukuba, Japan. After obtaining a Ph.D. in physical chemistry from the University of Tsukuba, Dr. Nakayama began developing theories and computer programs for doing semi–empirical and ab–initio MO calculations for a computer–aided molecular design system, MolWorks (http://www.molworks.com). MolWorks is shareware software and not only does MO calculations but also property estimations. He is a major contributor to MolWorks.

Umpei Nagashima is a principal researcher at Grid Technology Research Center of the National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan, and is also a professor of physical chemistry at the University of Tsukuba. Prof. Nagashima obtained a Ph.D. from Hokkaido University in 1983 and became an assistant professor of the Institute for Molecular Science, Okazaki National Research Institutes, Okazaki, Japan. More recently, his major research project is the development of applications of MO calculations based on grid computing technology.