

Structure Optimization of Solute Molecules via Free Energy Gradient Method

Masataka Nagaoka

Graduate School of Human Informatics, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan

Received March 5, 2003

Fundamental ideas of the free energy gradient method are briefly reviewed with three applications: the stable structures of glycine and ammonia-water molecule pair in aqueous solution and the transition state (TS) structure of a Menshutkin reaction $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$ in aqueous solution, which is the first example of full TS optimization of all internal degrees of freedom.

Key Words : Free energy gradient, Structure optimization, Transition state

Introduction

In a lot of chemical, biological and environmental phenomena, the chemical reaction dynamics *in solution* or *across interface* plays a very important role, where the microscopic solvation structures of solute molecules offer essential and inevitable information.¹ To obtain a stationary point on a multi-dimensional free energy surface (FES),¹ e.g., stable state (SS) or transition state (TS) in solution chemical reaction, a new TS optimization method, the free energy gradient (FEG) method, has been developed.²⁻⁹ Analogous to the method for the Born-Oppenheimer (BO) potential energy surface using *ab initio* molecular orbital (MO) calculation, the FEG method utilizes force and hessian on the FES, which can be calculated by molecular dynamics (MD) method and the free energy perturbation (FEP) theory.^{10,11} Furthermore, on the basis of the method, we have proposed the definition of the intrinsic reaction coordinate (IRC) on the FES.⁴

In this review article, the theoretical idea of FEG method is first introduced in the next section (Theoretical Treatments) together with the algorithmic procedure of structure optimization. Next in the section (Application), three applications are summarized: first for the glycine in water, second for the ammonia-water molecule pair in water, and third for the TS optimization of a Menshutkin reaction $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$ in aqueous solution, which is the first example of full TS optimization on the multidimensional FES.

Theoretical Treatments

A. Structure Optimization. In MD simulation, forces acting on each atom of the solute molecule by all solvent molecules are always calculated every time step. By time-averaging these forces, the “force” on FES, *i.e.*, a minus of FEG, is obtained as a function of the solute molecular structure \mathbf{q}^s as follows,^{4,7-9}

$$\mathbf{F}^{\text{FE}}(\mathbf{q}^s) = - \frac{\partial G(\mathbf{q}^s)}{\partial \mathbf{q}^s} = - \left\langle \frac{\partial V_{\text{RS}}(\mathbf{q}^s)}{\partial \mathbf{q}^s} \right\rangle, \quad (1)$$

where $G(\mathbf{q}^s)$ is the FE function and $V_{\text{RS}}(\mathbf{q}^s)$ is the sum of the solute intramolecular potential energy and the solute-solvent interaction energy at \mathbf{q}^s . The brackets $\langle \dots \rangle$ in Eq. (1) denote the time average that is equal to the equilibrium ensemble average,

$$\langle \dots \rangle = \frac{\int d\mathbf{q}^B (\dots) \exp(-\beta V)}{\int d\mathbf{q}^B \exp(-\beta V)}, \quad (2)$$

where V is the whole potential of solution, under the condition that the solute molecular structure is restricted to be \mathbf{q}^s and \mathbf{q}^B denotes the solvent coordinates as a whole.

For structure optimization, in the FEG method with the steepest-descent-path following procedure, the i +1th reactant structure \mathbf{q}_{i+1}^s is taken to be

$$\mathbf{q}_{i+1}^s = \mathbf{q}_i^s + \Delta \mathbf{q}_i^s, \quad (3)$$

where an adaptive displacement vector $\Delta \mathbf{q}_i^s$ is defined as

$$\Delta \mathbf{q}_i^s \equiv c_i \cdot \mathbf{M}^{-1} \cdot \mathbf{F}_i^{\text{FE}} \quad (4)$$

by multiplying

$$\mathbf{F}_i^{\text{FE}} \equiv \mathbf{F}^{\text{FE}}(\mathbf{q}_i^s) = - \left\langle \frac{\partial V_{\text{RS}}(\mathbf{q}^s)}{\partial \mathbf{q}^s} \right\rangle_i \quad (5)$$

by an adaptive constant c_i of dimension T^2 and the inverse of the mass matrix,

$$\mathbf{M} = \begin{bmatrix} m_1 & & & & & \\ & m_1 & & & & \\ & & m_1 & & & \\ & & & \ddots & & \\ & & & & m_N & \\ & & & & & m_N \\ \mathbf{0} & & & & & & m_N \end{bmatrix} \quad (6)$$

where m_j is the j th atomic mass of N -atomic molecule involved.

To obtain the stable molecular structure in solution, the

*Tel&Fax: +81-52-789-5623; <http://frontier.ncube.human.nagoya-u.ac.jp/> E-mail: mnagaoka@info.human.nagoya-u.ac.jp

procedure in the FEG method is summarized as the following processes:

- [P1] Start with the geometry \mathbf{q}_k^s , $k = 0$.
 [P2] For \mathbf{q}_k^s , calculate the force on the FES \mathbf{F}_k^{FE} .
 [P3] Find the stationary point using the force:

$$\Delta \mathbf{q}_k^s = c_k \cdot \mathbf{M}^{-1} \cdot \mathbf{F}_k^{\text{FE}}. \quad (7)$$

If the force \mathbf{F}_k^{FE} is small enough within the tolerance of convergence and/or the predicted change in the geometry $\Delta \mathbf{q}_k^s$ is small enough to be satisfied with the condition

$$\left\langle \frac{\partial V_{\text{RS}}(\mathbf{q}^s)}{\partial \mathbf{q}^s} \right\rangle \approx 0; \text{ zero-gradient condition}, \quad (8)$$

then, stop, setting the step number n_{fin} .

- [P4] Set $\mathbf{q}_{k+1}^s = \mathbf{q}_k^s + \Delta \mathbf{q}_k^s$, $k = k + 1$ and return to P2.

B. Free Energy Optimization. During the structure optimization procedure, as the solute molecular geometry \mathbf{q}^s varies from \mathbf{q}_i^s to \mathbf{q}_{i+1}^s (Eq. (3)), the corresponding FE change ΔG_i is provided with by the FEP theory^{10,11} and is obtained by

$$\Delta G_i = G_{i+1} - G_i = -k_B T \ln \langle \exp[-\beta \{V_{\text{RS}}(\mathbf{q}_{i+1}^s) - V_{\text{RS}}(\mathbf{q}_i^s)\}] \rangle_i, \quad (9)$$

where the subscript i in the average $\langle \dots \rangle_i$ means that it is taken over the sampling at \mathbf{q}_i^s . Thus, the FE profile can be drawn by connecting these differences from the initial structure \mathbf{q}_0^s to the final one $\mathbf{q}_{\text{opt}}^s$ at optimization. Then, the FE change of stabilization ΔG_{stabi} results in

$$\Delta G_{\text{stabi}} = \sum_{i=0}^{n_{\text{fin}}} \Delta G_i. \quad (10)$$

Application

A. Glycine in Water. First, the FEG method was applied to the multidimensional geometry optimization of glycine zwitterion (ZW) in aqueous solution in order not only to demonstrate its applicability but also to examine its efficiency. The solute potential energy $V_{\text{RS}}(\mathbf{q}^s)$, *i.e.*, the sum

of the ZW intramolecular potential energy and the interaction between ZW and a solvent water molecule (TIP4P) was described by a fitted potential energy function obtained previously by the empirical valence bond (EVB) method.^{5,6,12-14} Then, the most stable ZW structure in aqueous solution was obtained within the tolerance assumed, and it was found that the FE and enthalpy changes of stabilization from the initial geometry optimized in the gas phase are -0.9 and -3.5 kcal/mol, respectively, and the amino and carboxyl groups are spatially separated each other due to their solvating with water molecules (Figure 1).⁷ Comparing the contributions of enthalpy and entropy to FE, the former is attributed to the main origin of the FE stabilization during the optimization procedure, and it was found that the solvation entropy prevents water molecules from solvating the ZW more strongly.⁷

B. $\text{NH}_3 + \text{H}_2\text{O} \rightarrow \text{NH}_4^+ + \text{OH}^-$. Second, corresponding to the reactant state of the ionization reaction $\text{NH}_3 + \text{H}_2\text{O} \rightarrow \text{NH}_4^+ + \text{OH}^-$ ($K_a = 2.24 \times 10^{-5}$ (25 °C) (Expt.)), the hydrated structure of ammonia molecule in aqueous solution was theoretically optimized as an ammonia-water molecule pair ($\text{H}_3\text{N} \cdots \text{H}_2\text{O}$) by the FEG method (Figure 2).⁹ The $\text{H}_3\text{N} \cdots \text{H}_2\text{O}$ potential energy, *i.e.*, the sum of the intrapair potential energy and the interaction between the pair and a solvent water molecule (TIP3P) was described by a hybrid quantum mechanical and molecular mechanical (QM/MM) method combined with a semiempirical MO method at the PM3 level of theory.⁹ In the QM/MM formalism, $V_{\text{RS}}(\mathbf{q}^s)$ can be represented, for an instantaneous solvent configuration \mathbf{q}^B , as

$$V_{\text{RS}}(\mathbf{q}^s) = \langle \Psi | \hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}} | \Psi \rangle, \quad (11)$$

where Ψ denotes the SCF wave function of electrons of the solute molecules (\mathbf{q}^s) in solution. In Figure 3, both $\Delta G(300 \text{ K})$ and ΔV_{R} are shown as functions of the step number, which denote the FE change at 300 K and the corresponding change in the solute potential energy $V_{\text{R}} (= \langle \Psi | \hat{H}_{\text{QM}} | \Psi \rangle)$, respectively. At the first step of optimization, ΔG decreases suddenly to almost -0.3 kcal/mol, while it becomes almost flat up to 9. Presently, judging from the calculated root mean square values,⁹ the geometry at the step number 8 is taken to be the most stable structure of the ammonia-water molecule

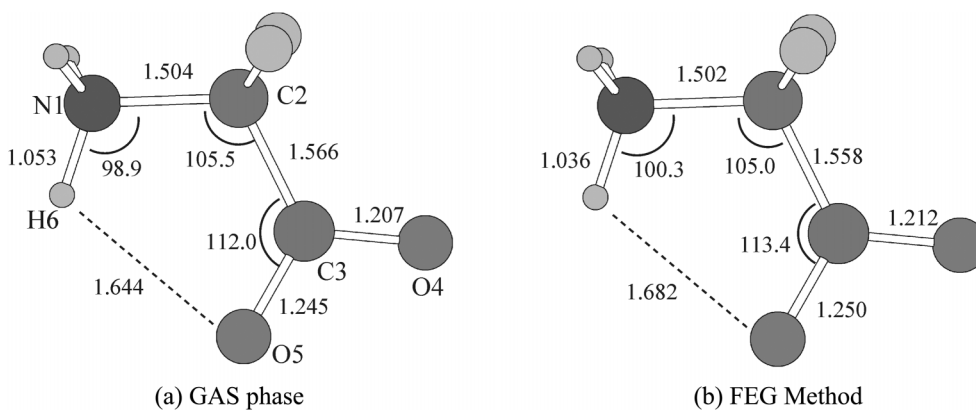


Figure 1. Stable-state structures optimized (a) in the gas phase by the HF level of theory and (b) in the aqueous solution obtained by the FEG method. Bond lengths are in Å and bond angles are in degree.

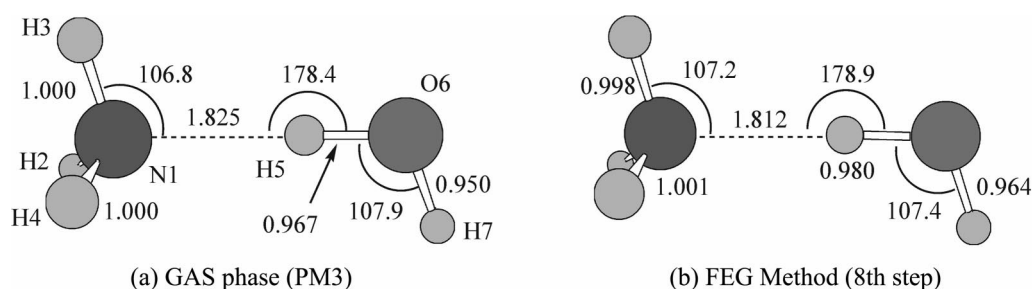


Figure 2. Stable-state structures optimized (a) in the gas phase by the PM3 and (b) in the aqueous solution obtained by the FEG method. Bond lengths are in Å and bond angles are in degree.

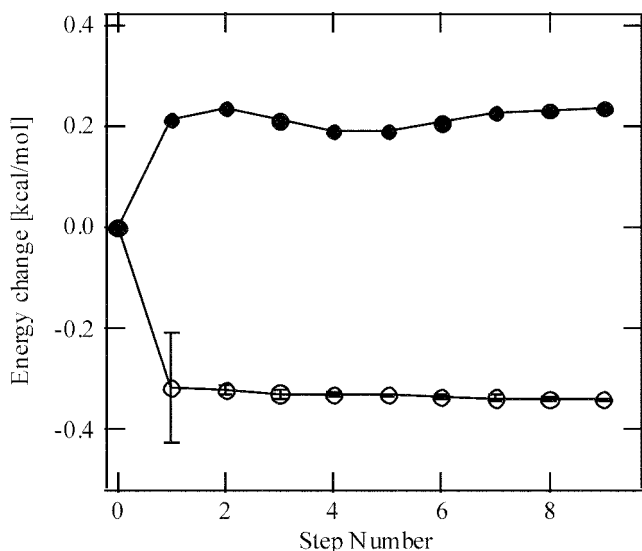


Figure 3. Free energy change (open circles with an error bar) and solute potential energy change (closed circles) of the stabilization of ammonia-water 1 : 1 pair in aqueous solution using the FEG method combined with the QM/MM-MD method at the PM3 level of theory.

pair, and then the FE change of stabilization was estimated -0.3 kcal/mol from the FE for the same structure as that of the cluster in gas phase.

The optimized structure was found to be almost the same as that in the gas phase except a longer OH bond length of the water molecule (Figure 2). However, its realization in solution is accomplished by virtue of subtle fulfillment of both “zero gradient” condition (Eq. (8)) and “force balance” condition:

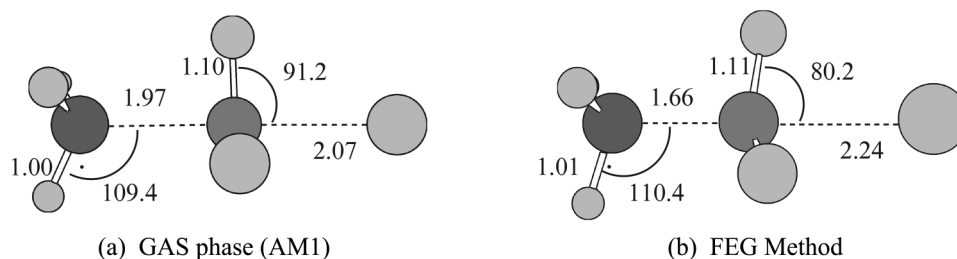


Figure 4. Transition-state structures optimized (a) in the gas phase by the AM1 and (b) in aqueous solution obtained by the FEG method. Bond lengths are in Å and bond angles are in degree.

$$\left\langle \frac{\partial \langle \Psi | \hat{H}_{\text{QM}} | \Psi \rangle}{\partial \mathbf{q}^s} \right\rangle = - \left\langle \frac{\partial \langle \Psi | \hat{H}_{\text{QM/MM}} | \Psi \rangle}{\partial \mathbf{q}^s} \right\rangle, \quad (12)$$

meaning the compensation for the balance between the solute intramolecular potential energy gradient and the forces acting on each solute atom due to solvation.⁹

C. $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$ ---The First Example of Full TS-optimization. Recently, the TS for the Menshutkin reaction $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$ in aqueous solution was located on the FES with respect to all internal degrees of freedom of the reactant molecules (Figure 4).⁸ As explained in the subsection III-B, the solute-solvent system was described by a hybrid QM/MM method.⁸ The TS thus obtained (Figure 4(b)) was found earlier than that in the gas phase (Figure 4(a)), which is in accord with the Hammond postulate. As a result, the reaction path in water was found to deviate largely from that in the gas phase.⁸ One can understand therefore that, in aqueous solution, the reaction should proceed accompanied with such a structural deformation to enhance the charge separation. It is concluded that, in such a reaction including charge separation, TS structure optimization on an FES is inevitable for obtaining valid information about a TS in polar solution.⁸

Finally, it is concluded that the present FEG method works quite well in spite of its simple optimization scheme equipped with only the adaptive displacement vector.

Acknowledgements. The author thanks Prof. K. S. Kim for his kindly inviting me to contribute to the special issue of Bulletin of Korean Chemical Society for the Proceedings of the 10th Korea-Japan joint symposium. This work was supported partly by a Grant-in-Aid for Science Research

from the Ministry of Education, Science and Culture in Japan.

References

1. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Warshel, A.; Wiley: New York, 1991.
 2. Nagaoka, M.; Okuno, Y.; Yoshida, N.; Yamabe, T. *Int. J. Quant. Chem.* **1994**, *51*, 519-527.
 3. Nagaoka, M.; Okuno, Y.; Yoshida, N.; Yamabe, T. In *The Proceeding of the Third China-Japan Symposium on Theoretical Chemistry*; Xiamen, 1994; p 31.
 4. Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. *Int. J. Quantum Chem.* **1998**, *70*, 95-103.
 5. Nagaoka, M.; Okuyama-Yoshida, N.; Yamabe, T. *J. Phys. Chem. A* **1998**, *102*, 8202-8208.
 6. Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. *ibid.* **1998**, *102*, 285-292.
 7. Okuyama-Yoshida, N.; Kataoka, K.; Nagaoka, M.; Yamabe, T. *J. Chem. Phys.* **2000**, *113*, 3519-3524.
 8. Hirao, H.; Nagae, Y.; Nagaoka, M. *Chem. Phys. Lett.* **2001**, *348*, 350-356.
 9. Nagae, Y.; Ohishi, Y.; Naruse, N.; Nagaoka, M. *submitted for publication*.
 10. Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420-1426.
 11. Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607-1604.
 12. Warshel, A.; Weiss, R. M. *J. Am. Chem. Soc.* **1980**, *102*, 6218-6226.
 13. Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523-2544.
 14. Chang, Y.-T.; Miller, W. H. *J. Phys. Chem.* **1990**, *94*, 5884-5888.
-