学位論文

Theoretical study on solvation effect by means of multistage strategy:

Combination of low level sampling and high level calculation

多段階計算による溶媒和効果についての理論化学的研究：

低精度計算によるサンプリングと高精度計算の組み合わせ

２013 年

広島大学大学院理学研究科

化学専攻

宮本　秀範

目　次

１．主論文

Theoretical study on solvation effect by means of multistage strategy:

Combination of low level sampling and high level calculation

多段階計算による溶媒和効果についての理論化学的研究：

低精度計算によるサンプリングと高精度計算の組み合わせ

２．公表論文

(1) Ab Initio QM/MM-MC Study on Hydrogen Transfer of Glycine Tautomerization in Aqueous Solution: Helmholtz Energy Changes along Water-mediated and Direct Processes.

H. Miyamoto, M. Aida, *Chemistry Letters*, **2013**, 42, 598.

(2) Helmholtz Energy Change between Neutral and Zwitterionic Forms of Glycine in Aqueous Solution using Ab Initio Expanded QM/MM-MC with QM Solvent.

H. Miyamoto, M. Aida, *Chemistry Letters*, **2013**, 42, 1010.

主論文

Theoretical study on solvation effect by means of multistage strategy: Combination of low level sampling and high level calculation

Hidenori Miyamoto

Department of Chemistry, Graduate School of Science,

Hiroshima University

Acknowledgement

I am deeply grateful to Professor Misako Aida for kindly guidance and a lot of encouragement throughout my studies.

I am deeply grateful to Professor Taka-aki Ishibashi in Tsukuba university for helpful comments and valuable discussions about my research.

I am deeply grateful to Dr. Yukiteru Katsumoto, Assistant Professor, for helpful suggestions.

I am grateful to Professor Takayuki Ebata, Professor Katsuyoshi Yamasaki, and all of the professors and staffs in the department of chemistry in Hiroshima University.

I am grateful to Professor Suehiro Iwata in Keio University, and Dr. Michel Dupuis in Pacific Northwest National Laboratory.

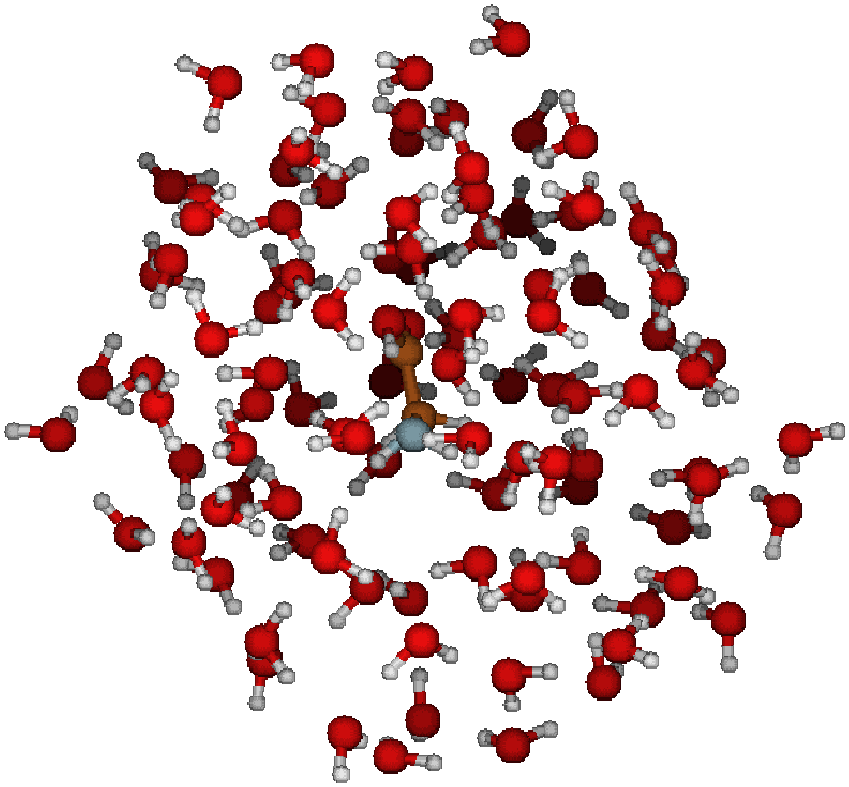
I want to thank Dr. Masayuki Ohisa, Mr. Hideo Doi, Mr. Dai Akase for valuable discussions and comments.

I also want to thank Mr. Akihiro Maeda, Dr. Shunsuke Mieda, Mr. Hiroshi Ando, and all the members of quantum chemistry research group.

Contents

Chapter 1

General Introduction



1.1 Introduction

Most of the chemical reactions occur in solution. Especially, chemical reactions in aqueous solution are very important because relatively large polar effect of water interacts with solute molecules. Hence, their environment is different from the that in the gas phase. And chemical reactions in aqueous solution are also important for not only chemist but also biologist. Thus, chemical reactions in aqueous solution have been studied for the long time.

Ab initio MO method can routinely be used to calculate the energy of molecules in the gas phase and at the ground state. In aqueous solution, however, energy calculation methods are under development yet. That is because there are a lot of configurations of water molecules at finite temperature, and it is essential to take into account those effects. In other words, we have to take into account not only energy but also free energy of the system at finite temperature.

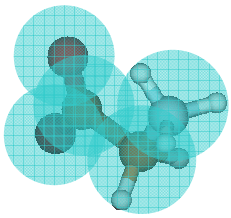
Polarizable continuum model (PCM) has been widely used to calculate molecular properties in the aqueous solution. However, PCM cannot consider the direction of solvation such as a hydrogen bonding and a solvent structure originating from solvent-solvent interactions. Water cluster systems have been also used to consider the hydration effect. In such a model, artifactual strong hydrogen bonds tend to be made up between the solute and water, and/or water and water. Those hydrogen bonds tend to make a special hydrogen bonding pattern, which is different from the pattern in solution. Hence, we adopted a water droplet system followed by Monte Carlo simulation.

In this study, we used two type of simulation, MD simulation and MC simulation. (MD simulation results are omitted in this thesis.) MD simulation provides time evolution of the system. We can obtain IR spectrum of the system by using Fourier transformation of dipole moment changes. MC simulation provides configuration distribution of the system. Hence, we can obtain free energy change of the system by using free energy perturbation theory. In the gas phase, NF glycine is more stable than ZW. However, glycine exists as a zwitterion (NH_3+CH_2COOP, GLYZW) in aqueous solution and in the crystal phase. Tautomerization reaction of glycine is a good model reaction to estimate solvation effect.

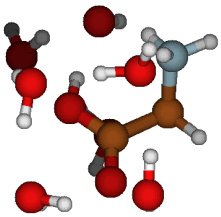
There are many routes of tautomerization reaction. A ZW glycine molecule is a compound with an intramolecular hydrogen bond. Hence, we can assume that glycine molecule gets converted from NF to ZW via intramolecular/intermolecular hydrogen transfer in aqueous solution. In addition, free energy perturbation theory is a state function. Then resulting free energy difference does not depend on reaction routes.

Though there is a slight minimum on the potential energy surface with HF/6-31G(d), there is no minimum on the gas phase potential energy surface corresponding to a glycine zwitterion with basis set which contains polarization functions on hydrogen atoms (e.g. 6-31G(d, p)) or higher level of theory (e.g. MP2 level of theory).

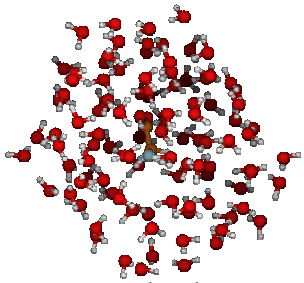
Energy calculation method



isolate molecule
(Polarizable continuum model)



water cluster
(full *ab initio* MO)

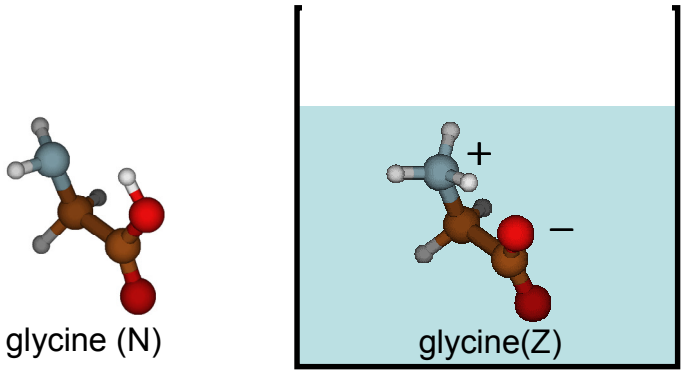


water droplet
(QM/MM)

Simulation method

Monte Carlo simulation,
Configuration distribution
Free energy

Molecular Dynamics simulation
Time evolution
IR spectrum



MC simulation

Chapter 2: The route of tautomerization reaction

Chapter 3: High level calculation of free energy

1.2 Framework of ab Initio QM/MM Method

A mixed Hamiltonian QM/MM model with the solute treated as QM and the solvent molecules treated as MM is used. The Hamiltonian for the total system consists of three terms:

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{QM/MM} + \hat{H}_{MM} \quad , \qquad (1)$$

where the first term is the vacuum Hamiltonian of the QM sub-system, the second term is the QM/MM interaction Hamiltonian and the last term corresponds to the energy of the MM sub-system. The second term of Eq. 1 is given by

$$\hat{H}_{QM/MM} = \hat{H}^{el}_{QM/MM} + \hat{H}^{vdW}_{QM/MM} \quad , \qquad (2)$$

where the first term describes the electrostatic interaction between the QM sub-system and the charge distribution of MM sub-system and the second term is the van der Waals interaction between the atoms in QM and MM sub-systems. The last term of Eq. 1 is given by

$$\hat{H}_{MM} = \hat{H}_{MM/MM} + \hat{H}^{vib}_{MM} \quad , \qquad (3)$$

where the first term describes the interaction between the atoms in the MM sub-system, and the second term describes the vibrational energy of the molecules in the MM sub-system.

The total energy of the QM/MM system is comprised of five terms;

$$E^{QM/MM} = E_{qm} + E^{el}_{qm/mm} + E^{vdW}_{qm/mm} + E_{mm/mm} + E_{vib} \quad . \qquad (4)$$

Here, the first term on the right-hand side is the total energy of the QM sub-system that is calculated with ab initio MO at a selected level of theory. The second term is the electrostatic energy between the QM sub-system and the MM charge distribution. The potential operator acting on the electron

density due to the MM charges is included in the one-electron Hamiltonian operator in ab initio MO method of the QM sub-system. The third term is the van der Waals energy between the QM sub-system and the MM sub-system, where the Lennard-Jones parameters [1] are assigned for the atoms of the QM sub-system. The fourth term is the interaction energy in the MM sub-system. The last term is the vibrational energy of the MM sub-system, which is calculated from the quartic potential [2], The calculation level of HF/6-31G(d) and MP2/6-31G(d) were used for the QM sub-system, and the TIP3P parameter set for water was used for the MM sub-system.

1.3 Free energy perturbation theory

The thermodynamics perturbation theory was used to calculate the free energy differences in 101 water molecules. The difference of free energy between solute structures i and j was obtained by: Zwanzig's equation.

$$\Delta A = -kT \ln[< \exp\{-(Ej - Ei)/kT\} >_{(i)}]$$

The bracket denotes the ensemble average of the difference in total QM/MM energy Ei and Ej of the system with solute in i and j configurations respectively, and with the solvent molecules in configurations originating from a Monte Carlo random walk on the condition that the energy of solute is based on the i configuration.

1.4 Multistage strategy

1.4.1 Simulated Annealing of the System

Since many solvent molecules are included in the system, there will be a lot of local minima on the potential energy surface. To find a more stable solvent configuration, we take a procedure of the simulated annealing in combination with the energy optimization. The Monte Carlo algorithm is used for the simulated annealing process, in which a high temperature (298 K) is given to the solvent sub-system with the fixed solute geometry, allowing the solvent sub-system to equilibrate around the solute and then cooling by every 14.9 K until the temperature is 14.9K. In the Monte Carlo method, 505,000 configurations are generated randomly at each temperature. In the simulated annealing process, the positions and the directions of the MM molecules are equilibrated and the QM part (the solute) is fixed: therefore, the first term (QM energy term) and the last term (MM vibrational energy term) in Eq. 4 are eliminated during the simulated annealing calculations. Thus, in the simulated annealing process, the total energy of the system is comprised of three terms:

$$E^{MM/MM} = E^{el}_{qm/mm} + E^{vdW}_{qm/mm} + E_{mm/mm} \quad . \qquad (5)$$

The energy functions are identical to the corresponding terms in Eq. 4 except for the electrostatic interaction term (the first term of the right-hand side of Eq. 5, which is replaced by the Coulomb interactions between the CHELPG charges [3] by HF/6-31G(d) of the QM atoms and the partial charges of the MM atoms.

1.4.2 Double perturbation approach

The double perturbation approach of Dupuis et al. [4] was adopted. In this approach, MM/MM level of theory was used for the random walk, with the solute subsystem treated at the MM level of theory, with the same Lennard-Jones parameters as those used in the QM/MM model and with atomic partial charges extracted from a fit to the electrostatic potential of the solute in the gas phase (CHELPG). In the present work, for each solute configuration i, 20200000 solvent configurations were generated at the MM/MM level of theory, at constant temperature T = 298 K and constant volume (NVT ensemble). Following the MM/MM random walk and energy averaging, 2020 configurations were selected randomly among the random walk configurations, the total QM/MM energies of the system with solute structures i, j and the same solvent configurations were calculated at the QM/MM level of theory.

1.4.3 Three layer free energy calculation

We treated the solvent molecules at the MM level of theory using the TIP3P water potential [5]. Water molecules which solvate a solute molecule directly are very important for solute-solvent interaction. Hence we have to treat those water molecules as QM molecules. To perform this idea, we implemented three layer free energy calculation, and performed to evaluate hydration free energy. Core layer which is reaction center is calculated by means of high level of theory. Intermediate layer which directly contacts reaction layer is calculated by means of intermediate level of theory. Outside

layer is calculated by means of low level of theory.

Free energy perturbation theory was carried out by means of QM/MM method on condition that MM water molecules within 2.45 Å of the glycine were treated as QM molecule ((gly + $n_{Q\text{-}W}$w)(QM) + (101-$n_{Q\text{-}W}$)w (MM)). Besides, MP2 level was adopted to calculate reaction center, because the activation barrier and the N-Z energy difference are strongly sensitive to electron correlation effects.

Resulting hydrogen free energy is "gly(QM: MP2/6-31G(d)) + $n_{Q\text{-}W}$w(QM: HF/6-31G(d)) + (101-$n_{Q\text{-}W}$)w(MM)": the solute molecule is calculated using MP2/6-31G(d)//HF/6-31G(d) level of theory, solvent molecules within 2.45 Å of the glycine are treated at HF/6-31G(d) level of theory, and other solvent molecules are treated at TIP3P.

1.5 References and notes

1. Freindorf, M.; Gao, J. *Journal of Computational Chemistry*. **1996**, *17*, 386–395.

2. Bartlett, R. J.; Shavitt, I.; Purvis, G. D. *The Journal of Chemical Physics*. **1979**, *71*, 281.

3. Breneman, C. M.; Wiberg, K. B. *Journal of Computational Chemistry*. **1990**, *11*, 361–373.

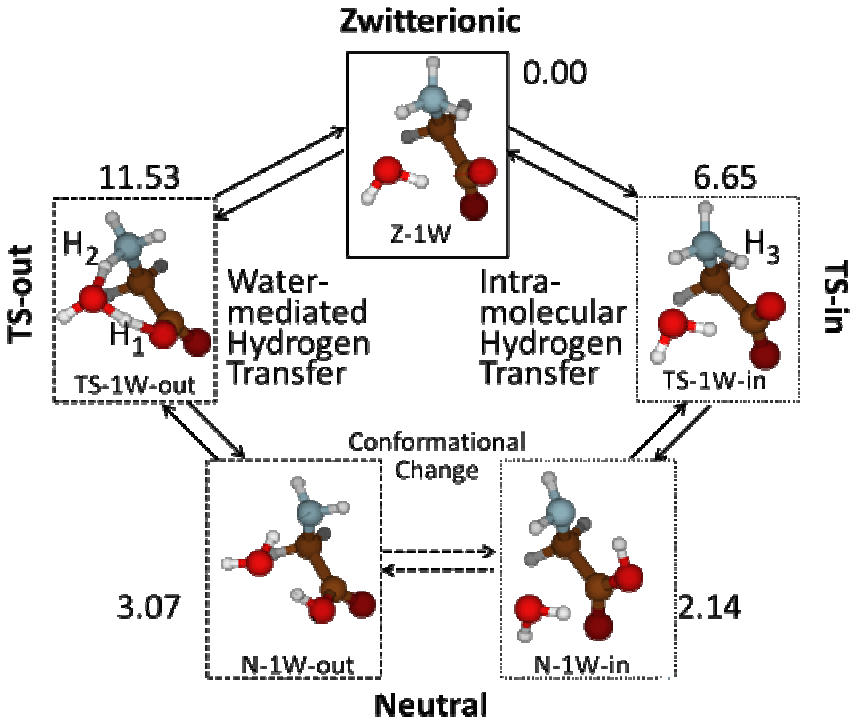
4. Dupuis, M. *Journal of Molecular Structure: THEOCHEM*. **2003**, *632*, 173–183.

5. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, J. Chem. Phys. 1983, 79, 926.

Chapter 2

Comparison of reaction routes for tautomerization of glycine

Ab Initio QM/MM-MC Study on Hydrogen Transfer of Glycine Tautomerization in Aqueous Solution: Helmholtz Energy Changes along Water-mediated and Direct Processes



2.1 Introduction

Glycine is the smallest amino acid: it exists as neutral (N) form in the gas phase, and as zwitterionic (Z) form in aqueous and in crystalline states. Although tautomerization of glycine between N and Z forms has been extensively studied, the process of hydrogen transfer in glycine tautomerization is not completely understood yet.[1-7] For micro-hydrated glycine complexes with 2 to 5 water molecules, it was shown that the energy barrier for hydrogen transfer by direct path was lower than those by water-mediated paths.[7] Ab initio QM/MM-FEG method was applied to glycine tautomerization via direct hydrogen transfer surrounded by MM water molecules,[4] where QM energy calculations were done at the theoretical level of MP2/6-31+G(d) with glycine structural parameters optimized at HF/6-31+G(d), and it was demonstrated that the computational free energy difference between N and Z forms compared well with experimentally observed values. On the other hand, ReaxFF (reactive force field) MD simulation was applied to glycine tautomerization (direct as well as mediated by one or two water molecules), and it was suggested that hydrogen transfer of glycine was mediated by a single water molecule in aqueous environment.[5] Thus, hydrogen transfer path in glycine tautomerization is still controversial.

2.2 Results and discussion

2.2.1 Reaction routes

In this thesis, we aim at presenting which route is probable starting from the same complex in

aqueous solution: one is a single water-mediated hydrogen transfer path and the other is a direct hydrogen transfer path. To this end, we treat a complex of glycine with one water molecule, gly-1W, at QM level (ab initio molecular orbital (MO) method with MP2/6-31G(d)). Helmholtz energy differences along the tautomerization routes are calculated based on the thermodynamic perturbation theory with Monte Carlo (MC) sampling of 100 MM water molecules surrounding the QM complex. Furthermore, we decompose the processes of single- and double-hydrogen transfers to neutral, transition state (TS) and zwitterionic regions based on the AIM analysis, and we show that it is not 'proton' but 'hydrogen atom' that transfers.

First, ab initio MO method was employed for the complex gly-1W in the gas phase. We obtained the stationary structures of intra- and inter-molecular hydrogen transfer reactions using MP2/6-31G(d) level of theory. Since it is probable that the N-Z energy differences, the TS structures and the barrier heights are strongly sensitive to the electron correlation effect, we adopted MP2 level of theory for ab initio MO calculations throughout this study. All stationary structures were confirmed by means of vibrational frequency calculation. We found one Z form structure (Z-1W) and two N forms (N-1W-out and N-1W-in) and two TS structures (TS-1W-out and TS-1W-in). The structures are shown in Figure 1(A). The coordinates are given in Supporting Information.[21] Intrinsic reaction coordinates (IRC) along two paths (water-mediated and direct paths) were calculated at the same level of theory, and we confirmed that TS-1W-out is a TS connecting Z-1W

and N-1W-out, and TS-1W-in is a TS connecting Z-1W and N-1W-in. Ab initio MO calculations were performed using

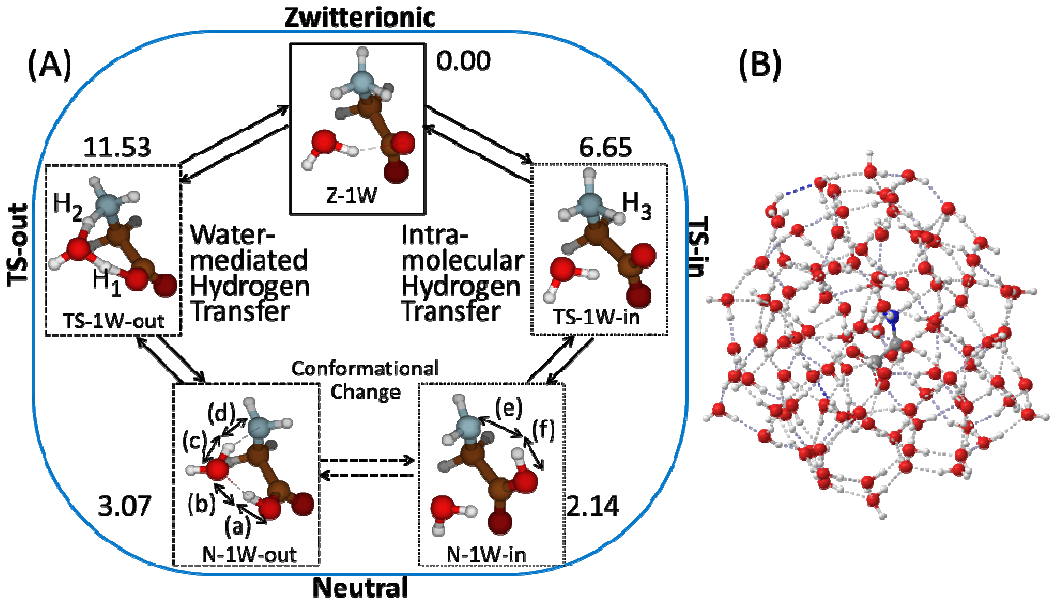


Figure 1. (A) Routes of hydrogen transfer of glycine with one water (QM complex). Helmholtz energy changes (in kcal/mol) are indicated relative to Z-1W. (B) 100 MM water molecules surrounding the QM complex.

Gaussian 03 [8] and GAMESS.[9]

Along two IRC paths, we performed the analysis of atoms-in-molecules (AIM)[10] of electron density of MP2 level, using the program packages of Gaussian 09[11] and AIM2000.[12] For each of some selected points along two IRC paths, the electron density and its Laplacian at each of bond critical points (BCPs) on (a) $O_{carboxyl} - H_1$, (b) $H_1 - O_{water}$, (c) $O_{water} - H_2$, (d) $H_2 - N$, (e) $N - H_3$, and (f) $H_3 - O_{carboxyl}$ were calculated. Those plots are shown in Figure 2. The numbering of hydrogen atoms and the bonds considered are indicated in Figure 1(A).

2.2.2 Electron density analysis along the reaction route

In the AIM analysis, the hydrogen bonding interaction is characterized by a positive value of the density Laplacian at BCP of the bonding, and a covalent bond by a negative value.[10] Along each IRC path of hydrogen transfer, the formation and the cleavage of a covalent bond and hydrogen bonding are clearly seen in Figure 2. We used the change in the density Laplacian at BCP of a hydrogen bond to define the borders between TS and N regions, and between TS and Z regions. We consider a point as a border where the density Laplacian at BCP of a hydrogen bond turns to decrease, which indicates that a hydrogen bond starts to change toward a covalent bond and that a hydrogen atom starts to transfer. In Figure 2, four borders between the regions are indicated.

Along the water-mediated tautomerization process, two hydrogen atoms (H_1 and H_2) transfer. As shown in Figure 2(A), in the transfer process of H_1, the change in the density at BCP on (a) is

concerted with that at BCP on (b), and the crossing of two changes occur near TS-1W-out. In the transfer process of H_2, the change in the density at BCP on (c) is concerted with that at BCP on (d), and the crossing of two changes occur near TS-1W-out. It is noteworthy that two crossings (one for (a) and (b), and the other for (c) and (d)) occur near TS-1W-out. The same characteristics are found for the changes in the density Laplacian (Figure 2(B)). This coincidence indicates that two hydrogen atoms (H_1 and H_2) transfer simultaneously.

Along the direct tautomerization process, one hydrogen atom (H_3) transfers. In the transfer process of H_3, the change in the density (Figure 2(A)) at BCP on (e) is concerted with that at BCP on (f), and the crossing of two changes occur near TS-1W-in. The change in the density Laplacian (Figure 2(B)) at BCP on (e) is concerted with that at BCP on (f), and the crossing of two changes occur near TS-1W-in. Two crossings (one for the density and the other for the density Laplacian) occur near TS-1W-in.

2.2.3 Hydration Helmholtz energy change along the reaction route

Secondly, we put the complex gly-1W in aqueous environment. For aqueous solution, we used a mixed Hamiltonian QM/MM model, with the complex gly-1W treated at QM level of theory (MP2/6-31G(d)) and other 100 solvent water molecules treated at MM level of theory (TIP3P[13]). We used the QM/MM protocol which is the same as used before.[14, 15] In our protocol, the potential operator acting on the electron density of QM part due to the effective charges of solvent MM

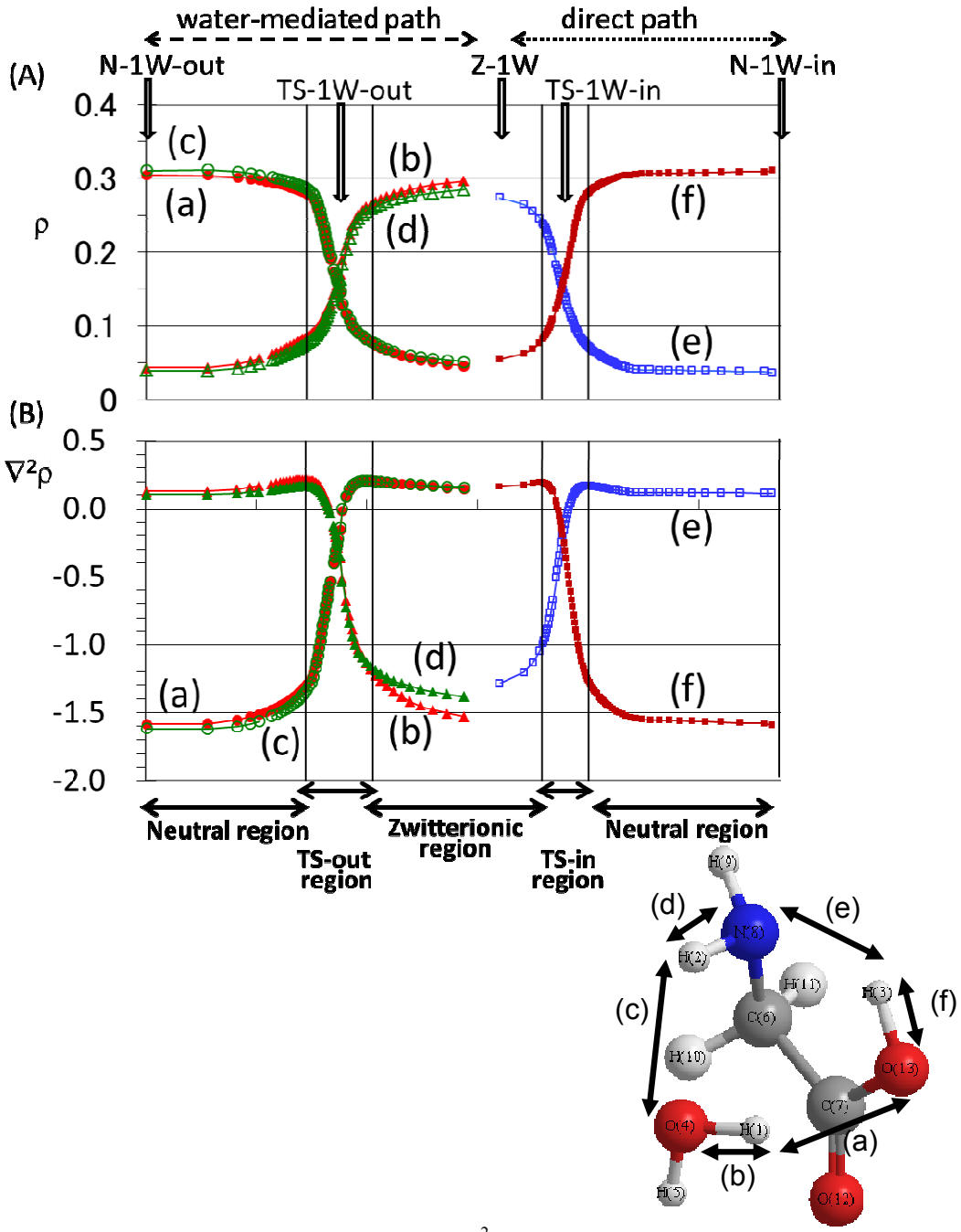


Figure 2. (A) Electron density ρ and (B) its Laplacian $\nabla^2\rho$ of bond critical points of the bonds (a) – (f) along two IRC paths of gly-1W tautomerization. See Figure 1(A) for the bond indices

molecules is included in the one-electron Hamiltonian operator of the QM subsystem. Total QM/MM energy of a system can be expressed as follows.

$$E^{\mathrm{QM/MM}} = E_{\mathrm{QM}} + E_{\mathrm{QM\text{-}MM}} + E_{\mathrm{MM}} \qquad (1)$$

E_{QM} is the energy of QM subsystem (*i.e.*, gly-1W complex) in the presence of the effective charges of solvent MM molecules. $E_{\mathrm{QM\text{-}MM}}$ is the interaction energy between QM subsystem and MM molecules, including electrostatic and van der Waals terms. E_{MM} is the interaction energy between MM molecules. The set of parameters[16] used for Z and N forms in the QM/MM calculation is listed in Supporting Information.[21] The parameters of N and O of the Z form are different from those of the N form. Along the hydrogen transfer processes, the parameters of N and O of glycine must change. Thus, we used a different value which changes linearly according to the variation in the distance between transferring H and N for the parameter of N or between transferring H and O for the parameter of O in glycine.

We calculated Helmholtz energy differences in aqueous QM/MM system along the selected points on the IRC paths of gly-1W by means of the thermodynamic perturbation theory. The difference in Helmholtz energy between gly-1W structures *i* and *j* in aqueous QM/MM system is given by:

$$\Delta A_{ij} = -kT[\ln\left\langle \exp\left\{-(E_j^{\mathrm{QM/MM}} - E_i^{\mathrm{QM/MM}})/kT\right\}\right\rangle_{(i)}] \qquad (2)$$

The bracket $\langle\ldots\rangle_{(i)}$ denotes the ensemble average of the difference in total QM/MM energies of gly-1W (in the structures *i* and *j*) with the solvent molecules in configurations originating from a

Monte Carlo random walk around gly-1W in the structure *i* in a solvation sphere with a radius of 9.32 Å (Figure 1(B)), in which the water density is 1 g cm^{-3}.

To sample solvent configurations, we adopted the double perturbation approach of Dupuis et al.[17] In this approach, an intermediate MM/MM level of theory is used for the random walk, with the gly-1W subsystem treated at the MM level of theory, with the same Lennard-Jones parameters as those used in the QM/MM model and with atomic partial charges extracted from a fit to the electrostatic potential of the solute in the gas phase (CHELPG charges).[18] In the present work, for each gly-1W structure *i* along the selected IRC points, 20000000 solvent configurations were generated at the MM/MM level of theory, at constant temperature T=298 K and constant volume (NVT ensemble), of which 2000 configurations were selected randomly to be used for QM/MM calculations. The total QM/MM energies of gly-1W structures *i* and *j* with the same solvent configuration based on gly-1W structure *i* were calculated, and their differences were averaged. The numbers of the points used for calculations along IRC paths were 43 between N-1W-out and TS-1W-out, 18 between TS-1W-out and Z-1W, 10 between Z-1W and TS-1W-in, and 69 between TS-1W-in and N-1W-in. The free energy calculations were performed using HONDO.[19]

It is informative to express the change in Helmholtz energy ΔA of a given QM/MM aqueous system relative to a reference state and as the sum of two terms,[20] one being the change in QM energy ΔE_{QM} and the other being the change in solvation free energy $\Delta\Delta A_{solv}$ as follows.

$$\Delta A = \Delta E_{\mathrm{QM}} + \Delta\Delta A_{\mathrm{solv}} \quad (3)$$

From the free energy perturbation theory, it is the sequence of ΔA that is calculated. Since we know ΔE_{QM} values along IRC paths, we can obtain the sequence of solvation free energy $\Delta\Delta A_{\mathrm{solv}}$. Note that these values are relative values. Any point can be taken as a reference state. In Figure 3, the change in Helmholtz energy ΔA is plotted relative to Z-1W along the IRC steps in black full circle. The change in solvation free energy $\Delta\Delta A_{\mathrm{solv}}$ is plotted in blue open circle relative to N-1W-in. The change in QM energy ΔE_{QM} is plotted in red open square such that the relation of Eq. 3 is satisfied using the values of ΔA and $\Delta\Delta A_{\mathrm{solv}}$ at each step.

As shown in Figure 3, Z-1W is the most stable in aqueous solution from the viewpoint of Helmholtz energy. In Z region, QM energy is high, while it is partly compensated by the solvation free energy. Along hydrogen transfer, solvation free energy changes drastically. The stabilization by the solvation free energy is larger in the water-mediated double hydrogen transfer path than that in the direct hydrogen transfer path, while the barrier height of the QM part is higher in the double hydrogen transfer path than that in the direct transfer path. As a result, the total Helmholtz energy barrier height for direct transfer path is lower than that for water-mediated transfer path. The QM calculations for micro-hydrated glycine complexes had shown that the energy barrier increased with the increasing number of water molecules in the transfer bridge.[7] It is probable that the increase in the number of water molecules in the transfer path does not result in the decrease in the free energy

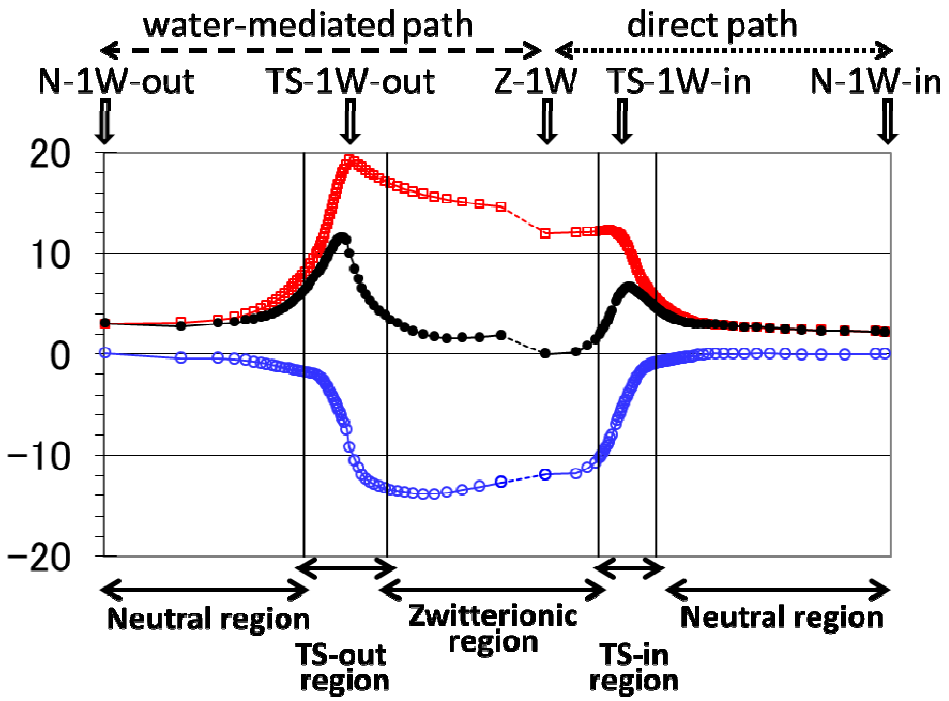


Figure 3. Changes in Helmholtz energy ΔA (black full circle), QM energy ΔE_{QM} (red open square) and solvation free energy $\Delta\Delta A_{solv}$ (blue open circle) are plotted along IRC steps. The values are in kcal/mol.

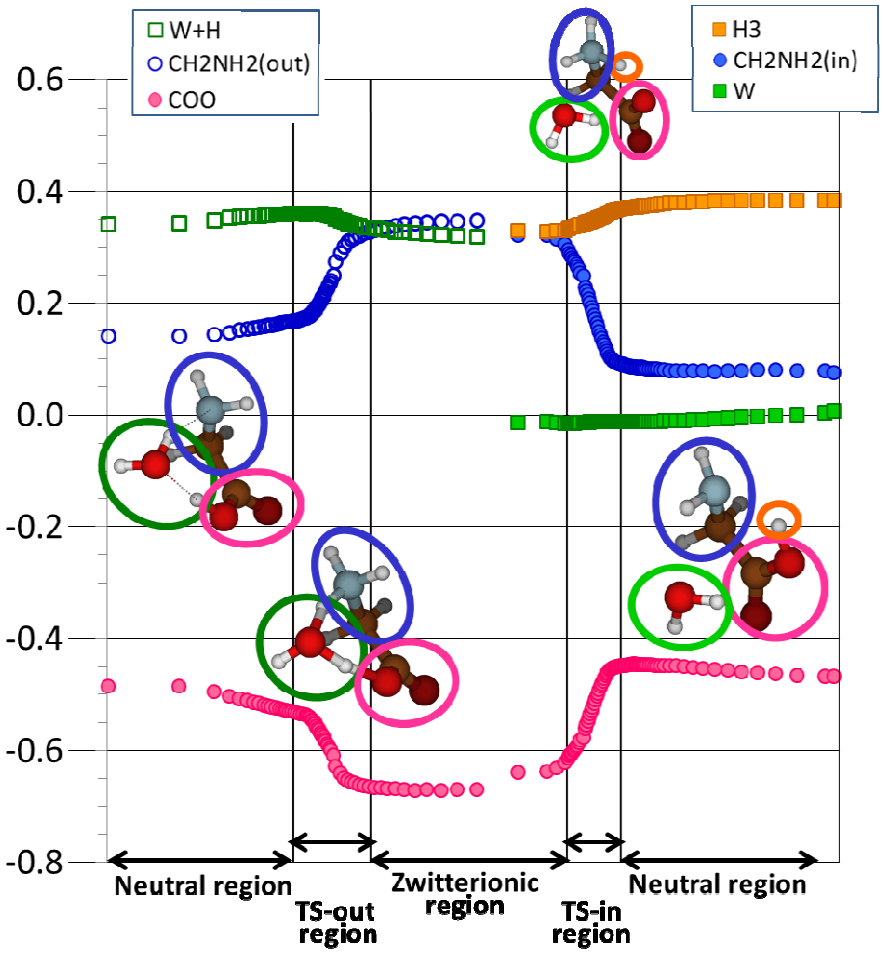


Figure 4. Changes of group charges along two IRC paths.

barrier height in aqueous environment.

For the direct transfer path, the free energy barrier height from Z to N form in the current work is 6.65 kcal/mol and it is lower than that of the previous work,[4] where HF/6-31+G(d) was used for the optimization of the TS structure.

Along both tautomerization routes, the changes in the group charges (from CHELPG atomic charges of the QM part) are plotted in Figure 4. The CH_2NH_2 and COO group charges change drastically in the TS regions. In the zwitterionic region, the CH_2NH_2 group charge is highly positive, and the COO group charge is highly negative. This results in the large stabilization of the solvation free energy in Z region as shown in Figure 3. It should be noted that the group charge of the transferring group (W+H in the double transfer path, and H_3 in the direct path) is almost constant along either IRC path, and that it is 'hydrogen atom' that transfers. This characteristic that the charge of the transferring atom in the tautomerization is almost constant is also seen in the AIM charges (see Appendix A).

2.3 Conclusion

Starting from the QM complex of glycine with one water in aqueous solution, we found that the Helmholtz energy barrier of the water-mediated hydrogen transfer path is higher than that of the direct path. We conclude that the tautomerization of glycine occurs in a direct path in aqueous environment.

2.4 References and Notes

1. M. Nagaoka, N. O.-Yoshida, T. Yamabe, *J. Phys. Chem. A* **1998**, 102, 8202.

2. B. Balta, V. Aviyente, *J. Comput. Chem.* **2003**, 24, 1789.

3. B. Balta, V. Aviyente, *J. Comput. Chem*. **2004**, 25, 690.

4. N. Takenaka, Y. Kitamura, Y. Koyano, T. Asada, M. Nagaoka, *Theor. Chem. Acc.* **2011**, 130, 215.

5. Rahaman, A. C. T. van Duin, W. A. Goddard III, D. J. Doren, *J. Phys. Chem. B*, **2011**, 115, 249.

6. S. Tolosa, A. Hidalgo, J. A. Sansón, *J. Phys. Chem. B*, **2012**, 116, 13033.

7. X.-J. Meng, H.-L. Zhao, X.-S. Ju, *Comput. Theoret. Chem.* **2012**, 1001, 26.

8. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B.

B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, 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, C. Gonzalez, and J. A. Pople., *Gaussian 03 (Revision A.02)*, Gaussian, Inc., Wallingford CT, **2004**.

9. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, 14, 1347.

10. R. F. W. Bader, *Chem. Rev.* **1991**, 91, 893.

11. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman,

J. V. Ortiz, J. Cioslowski, and D. J. Fox., *Gaussian 09 (Revision C.01)*, Gaussian, Inc., Wallingford CT, **2009**.

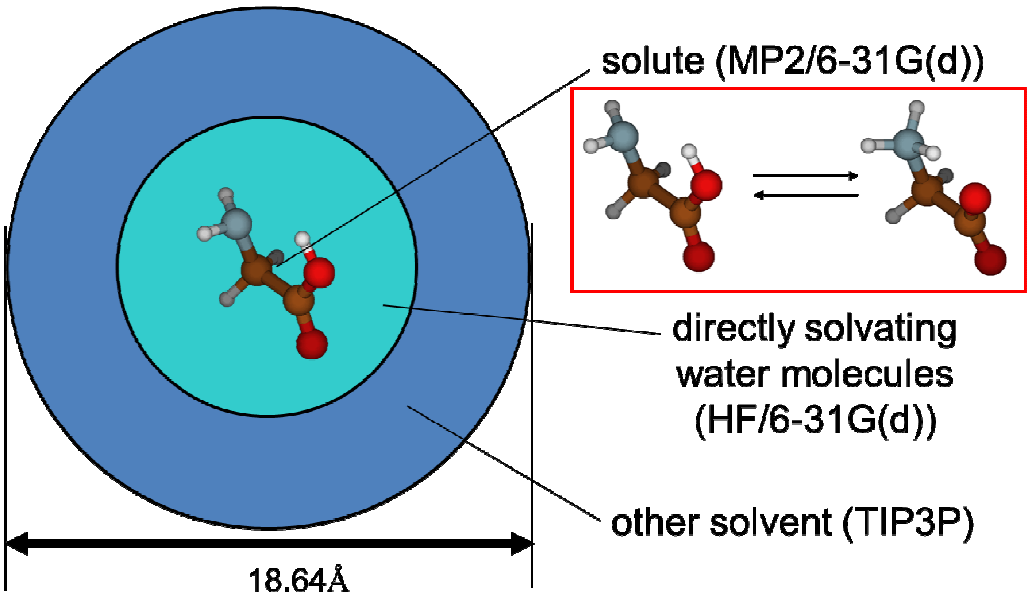
12. F. Biegler-Konig, J. Schonbohm, D. Bayles, *J. Comput. Chem.* **2001,** 22, 545.
13. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, 79, 926.
14. M. Aida, H. Yamataka, M. Dupuis, *Int. J. Quantum Chem.* **2000**, 77, 199.
15. M. Dupuis, M. Aida, Y. Kawashima, K. Hirao, *J. Chem. Phys*. **2002**, 117, 1242.
16. M. Freindorf, J. Gao, *J. Comput. Chem*. **1996**, 17, 386.
17. M. Dupuis, G. K. Schenter, B. G. Garrett, E. E. Arcia, *J. Mol. Struct.* (*Theochem*) **2003**, 632, 173.
18. C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, 11, 361.
19. M. Dupuis, *HONDO2004*, based on HONDO-95 available from the Quantum Chemistry Program Exchange, Indiana University.
20. M. Ohisa, H. Yamataka, M. Dupuis, M. Aida, *Phys. Chem. Chem. Phys*. **2008**, 10, 844.
21. Appendix A

Chapter 3

Quantitative estimation of solvation free energy

Helmholtz Energy Change between Neutral and Zwitterionic Forms of Glycine in Aqueous Solution

using Ab Initio Expanded QM/MM-MC with QM Solvent.



3.1 Introduction

Glycine is a molecule that can form an intramolecular hydrogen bond, and may convert from a neutral (N) form to a zwitterionic (Z) form in aqueous solution. Since the conversion between N and Z of glycine can be considered as a prototype of the tautomerization of amino acid, the process and the energy profile of the N-Z conversion have been extensively studied.[1-7] Experimentally, the free energy of Z form was estimated to be 7.27 kcal/mol lower than that of N form,[8] and the free energy barrier height from Z to N was estimated to be 14.6 kcal/mol.[9] Recently, we have shown in the quantum mechanical/molecular mechanical framework with Metropolis Monte Carlo method (QM/MM-MC method) that the free energy barrier height of a water-mediated tautomerization path is higher than that of the direct tautomerization path in aqueous environment, and that it is not 'proton' but 'hydrogen atom' that transfers.[7]

3.2 Computational method

In the current work, we aim at performing the quantitative calculation of the N-Z free energy difference of glycine in aqueous solution. To this end, we present an extended QM/MM-MC method to calculate the Helmholtz energy change. Solvation effect cannot be elucidated simply by taking account of only a few solvent molecules in a few selected configurations of the solvent, even if high level of theory is used. It is necessary to treat many solvent molecules and many solvent configurations explicitly. This is why several methods such as QM/MM-MC method have been

adopted. If MC simulations are to be performed to sample solvent configurations satisfactorily, a lower level of theory must be used. If a lower level of theory is used throughout for free energy calculations, however, a resultant computational value may not be reliable. In this Letter, the following three points are addressed to evaluate quantitatively the free energy change in aqueous solution. The first is how to obtain a solute structure in aqueous solution which may be largely different from that in the gas phase. The second is how to include close surrounding water molecules in QM part in free energy calculations. We present a new methodology here, which we call the "expanded QM/MM-MC method." The last is how to treat the solute with higher level QM in the free energy evaluation. Thus, we present the multistage evaluation of Helmholtz energy change.

At the beginning, we calculated the potential energy surface of a glycine molecule in the gas phase at the HF/6-31G(d) level of theory (Figure 1). By means of vibrational frequency calculations, all the stationary structures were confirmed. The IRC pathway connecting N, TS and Z structures were calculated. In Figure 1, N and Z structures are indicated by red full circles and TS structure by a red full triangle. Some points along the IRC pathway are indicated by black full squares. Note that the hydrogen transfer proceeds in a very narrow ravine. Ab initio MO calculations were performed using Gaussian 03.[10]

To obtain the optimized N and Z structures in aqueous solution, we put glycine in aqueous environment with 101 water molecules in a sphere such that the density of the solvent is 1 g/cm^3

with a radius of 9.32Å. Since many solvent molecules are included in the system, an abundance of

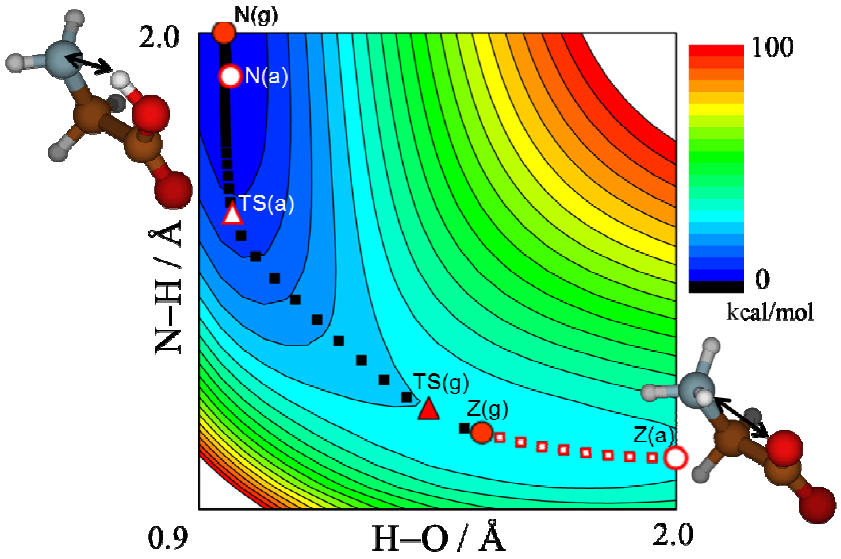


Figure 1. Potential energy surface of glycine at the HF/6-31G(d) level. The red full circles (N(g) and Z(g)) and the red full triangle (TS(g)) indicate N, Z and TS in the gas phase, respectively. The white circles (N(a) and Z(a)) and the white triangle (TS(a)) indicate N, Z and TS in aqueous solution, respectively.

local minima may exist. To find a more stable structure, we take the procedure of the combination of the simulated annealing[11] and the QM/MM geometry optimization, as follows. For the geometry optimization in aqueous environment, we use the QM/MM-vib level. Here, a glycine molecule is treated at the QM level of theory (HF/6-31G(d)) and other 101 solvent water molecules treated at the MM level of theory (TIP3P[12]) incorporated with the vibrational potential.[15] Total QM/MM energy of a system is expressed as follows.[13, 14]

$$E^{\mathrm{QM/MM}} = E_{\mathrm{QM}} + E_{\mathrm{QM\text{-}MM}} + E_{\mathrm{MM\text{-}MM}} + E_{\mathrm{vib}}. \quad (1)$$

Here, E_{QM} is the energy of QM subsystem in the presence of the effective charges of MM solvent molecules. $E_{QM\text{-}MM}$ is the interaction energy between QM subsystem and MM molecules. $E_{MM\text{-}MM}$ is the interaction energy between MM molecules. Evib is the total of the vibrational energy of each MM water molecule. The set of parameters used for the QM/MM calculation is same as those used in our previous work,[7] except for the vibrational energy, which is calculated from the quartic force field.[15]

The Metropolis Monte Carlo (MC) algorithm was used for the simulated annealing process, in which a high temperature (298K, in this work) was given to the system, allowing it to equilibrate and then cooling by every 14.9 K until the temperature is zero, and 505000 configurations were generated randomly at each temperature. In an annealing process, a solute structure was fixed, and the second and the third terms of Eq. (1) were used for energy calculations, in which the electrostatic

interaction term between solute and solvent was calculated by the Coulomb interactions between the CHELPG[16] using HF/6-31G(d) of the solute atoms and the partial charges of the solvent atoms.

In the combination procedure of the simulated annealing and the QM/MM geometry optimization, with a fixed geometry of glycine, several cycles of the simulated annealing process were performed, and then, all the positions of the molecules (glycine and 101 MM water molecules) were optimized in the framework of QM/MM-vib level. This procedure was repeated several times. Then we obtained an optimized geometry of glycine in 101 water molecules. Thus obtained N and Z structures in aqueous solution are plotted in Figure 1, with the white circles. Note that the Z structure in aqueous solution is very much different from that in the gas phase (see Figure S1[17]). The XYZ coordinates of these structures are listed in Tables S1 and S2.[17] In the gas phase, one of H atoms of NH_3 group is strongly hydrogen-bonded to one of O atoms of OOC group. In aqueous solution, they are separated so that both groups are exposed to solvent molecules. The H…O distance, 2.05Å, in aqueous solution is longer than that reported previously.[4] This may be because we performed the combination procedure of the simulated annealing and geometry optimization several times to obtain the Z form in aqueous solution.

The routes between the structures (N or Z) in the gas phase and the corresponding structures in aqueous solution were created by dividing the differences in the coordinates linearly. White squares in Figure 1 indicate the route from the Z structure in the gas phase to that in aqueous solution. Since

the reaction path does not affect the N-Z Helmholtz energy difference in the current case, the quality of the reaction path does not matter so much. We use the intermediate steps (plots in Figure 1) to evaluate the Helmholtz energy difference between the N and Z structures in aqueous solution, by means of the free energy perturbation method.

Since the solvation effect is the target of the current work and solvent configurational contributions must be important, we assume the system obeys classical statistical mechanics. The difference ΔAij in the classical Helmholtz energy of the system between glycine structures i and j is given by:

$$\Delta A_{ij} = -kT[\ln\left\langle \exp\left\{-(E_j^{\mathrm{QM/MM}} - E_i^{\mathrm{QM/MM}})/kT\right\}\right\rangle_{(i)}] . \quad (2)$$

The bracket <…>(i) denotes the ensemble average of the difference in total QM/MM energies of glycine (in the structures i and j) with the solvent molecules in configurations originating from the Metropolis sampling around glycine in the structure i in a solvation sphere with a radius of 9.32 Å .

To sample solvent configurations, we adopted the double perturbation approach.[18] In this approach, at a solute structure, 20200000 solvent configurations were generated (NVT ensemble at T=298K) using the second and the third terms of Eq. (1) for energy calculations, of which 2020 configurations were selected randomly to be used for the expanded QM/MM calculations. The total QM/MM energies of the solute structures i and j with the same solvent configuration based on the solute structure i were calculated, where the solute and directly solvating water molecules were treated as QM (HF/6-31G(d)). We regarded a water molecule as 'directly solvating,' when an atomic distance

between the solute and the water molecule is less than 2.45Å. A snapshot of solvent configuration around Z form of glycine is shown in Figure 2. The QM/MM and free energy calculations were performed using HONDO.[19]

Thus obtained Helmholtz energy change of the system can be divided[20] into the free energy change of the solute and the solvation free energy change as follows.

$$\Delta A_{system}^{HF/HF/MM} = \Delta A_{solute}^{HF} + \Delta\Delta A_{solvation}^{HF/MM} . \qquad (3)$$

The superscript indicates the level of theory used for the evaluation for each term. Since our free energy difference calculations are performed with a fixed geometry of the solute for each of the tautomerization route, the free energy change of the solute is equal to the potential energy change of the solute along the tautomerization. Then we obtain

$$\Delta A_{system}^{HF/HF/MM} = \Delta E_{solute}^{HF} + \Delta\Delta A_{solvation}^{HF/MM} . \qquad (4)$$

The potential energy change of the solute is obtained by performing ab initio MO calculations at the theoretical level of HF/6-31G(d). Thus, the solvation free energy change, i.e., the second term of the right-hand side of Eq.(4), is obtained as the difference between the left-hand side and the first term of the right-hand side of Eq.(4). The obtained solvation free energy change is plotted with blue diamond in Figure 4(A).

The following is the last stage of our evaluation of free energy change. Since it is probable that the energy differences among N, TS and Z structures of the solute are strongly sensitive to the electron

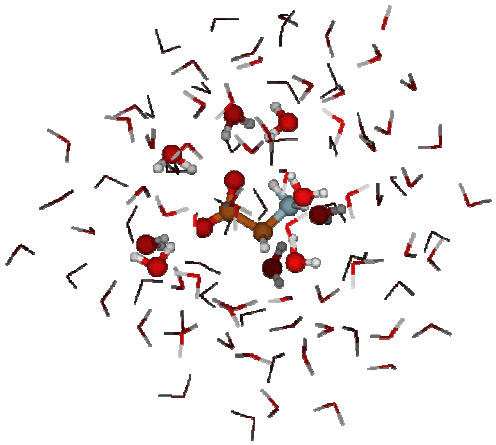


Figure 2. A snapshot of solvent configuration. Every water molecule is replaced by QM, when an atomic distance between the solute and the water molecule is less than 2.45Å. The molecules included in the QM sub-system are drawn with the ball-and-stick model.

correlation effect, we perform MP2/6-31G(d) calculations for the solute, and we replace the HF potential energy change with the MP2 energy change. Then we obtain the Helmholtz energy change, in which the solute is treated with MP2/6-31G(d), and the interaction energy between the solute and directly surrounded water molecules is treated with HF/6-31G(d) and the rest of water molecules by TIP/3P; namely,

$$\Delta A_{system}^{MP2/HF/MM} = \Delta E_{solute}^{MP2} + \Delta\Delta A_{solvation}^{HF/MM} . \qquad (5)$$

This is our final expression for Helmholtz energy change in this thesis. Schematic representation is shown in Figure 3.

3.3 Results and discussion

We used 52 points along the route of direct hydrogen transfer of glycine in aqueous solution, for each of which 2020 expanded QM/MM-MC calculations were performed, where a glycine molecule with directly solvating water molecules were calculated at HF/6-31G(d) surrounded by the rest of water molecules of TIP3P. Then the replacement of QM energy of HF/6-31G(d) with that of MP2/6-31G(d) led us to the Helmholtz energy change according to Eq.(5). The computational result is shown in Figure 4. The standard deviations for each step are indicated for Helmholtz energy change (black) and solvation free energy change (blue).

TS structure in aqueous solution, which corresponds to the highest point on the Helmholtz energy change, is plotted in the solute potential energy surface with the white triangle (Figure 1). It is

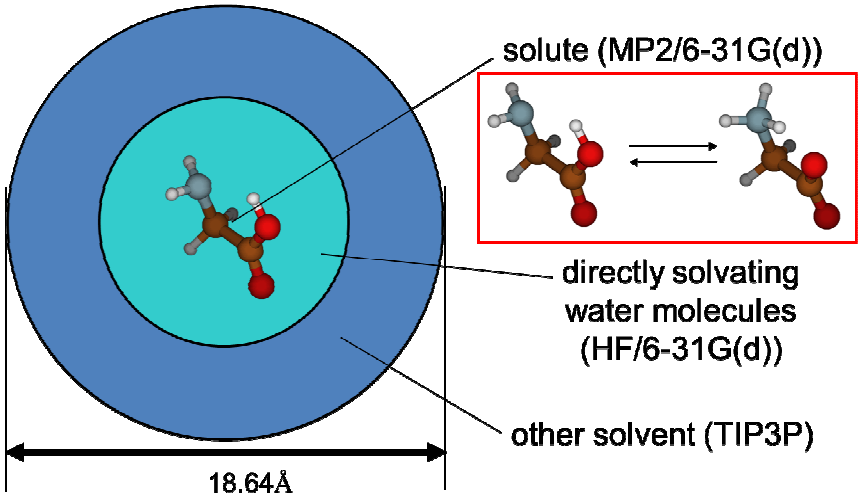


Figure 3. Schematic representation of the theoretical levels which are used to calculate the Helmholtz energy change of glycine tautomerization in the expanded QM/MM-MC method.

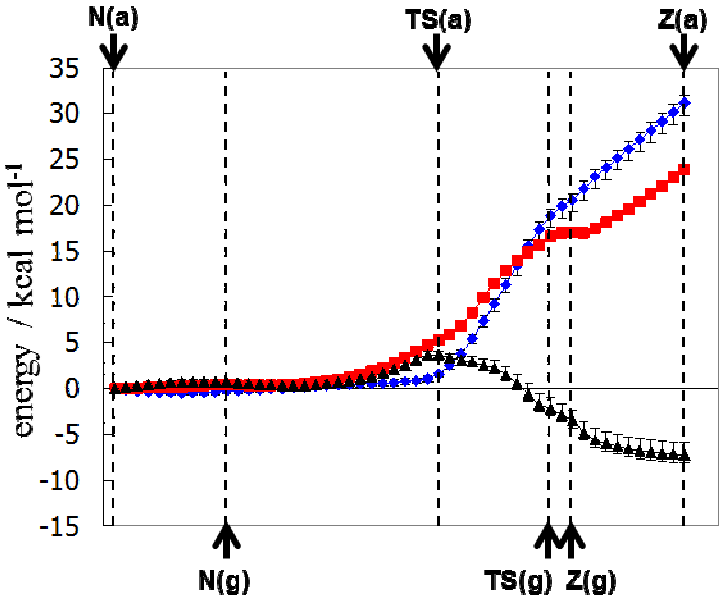


Figure 4. Helmholtz energy change (black full triangles), solute potential energy change (red full squares) and solvation free energy change (blue full diamond) along the hydrogen transfer route of glycine in aqueous solution, according to Eq.(5). The solvation free energy change is plotted with the reversed sign. N(a), TS(a) and Z(a) indicate N, TS and Z forms in aqueous solution, respectively. N(g), TS(g) and Z(g) indicate those in the gas phase. The standard deviations for each step are indicated for Helmholtz energy change (black) and solvation free energy change (blue).

interesting to note that the TS structure is closer to the N form in aqueous solution than in the gas phase. Along the hydrogen transfer route, the changes in the group charges (from CHELPG atomic charges of the QM part at HF/6-31G(d)) are plotted in Figure S2(B).[17] During the hydrogen transfer, the charge of the transferring atom is almost constant, and the charge redistribution occurs in the rest of the molecule. Starting from the N form, when the CH_2NH_2 and COO group charges start to change drastically, the solvation free energy starts to change. In the current work, we treated water molecules around the solute (when they are within 2.45Å of the solute) as QM (HF/6-31G(d)), which resulted in larger stabilization in solvation free energy than that of the previous work[7] where all solvent water molecules were treated as MM. When a solute is ionized, water molecules surrounding it must be polarized, and it is required to treat them as QM.

Regarding the tautomerization route, there may be other paths than the one considered in the current work. Furthermore, there may be various states corresponding to the N form and Z form of glycine in aqueous solution. In the current work, we do not extend our calculations to those other routes or other states. Concerning the N states, it has been suggested that the free energy differences between those states are rather small.[21]

3.4 Conclusion

In aqueous solution, our computational Helmholtz energy of the Z form of glycine is lower than

that of the N form by 7.25 kcal/mol, and it compares well with the experimental value (7.27 kcal/mol).[8] This is a coincidence. Nonetheless, the multistage strategy, which we have presented in this Letter, is a promising way to evaluate the Helmholtz energy change quantitatively; solvent configurations are generated at MM level, solvation free energy is evaluated treating directly surrounding water molecules at HF level, and the solute energy change is evaluated at MP2 level.

3.5 References and Notes

1 M. Nagaoka, N. O.-Yoshida, T. Yamabe, J. Phys. Chem. A 1998, 102, 8202.

2 K. Leung, S. B. Rempe, J. Chem. Phys. 2005, 122, 184506.

3 C. M. Aikens, M. S. Gordon, J. Am. Chem. Soc. 2006, 128, 12835.

4 N. Takenaka, Y. Kitamura, Y. Koyano, T. Asada, M. Nagaoka, Theor. Chem. Acc. 2011, 130, 215.

5 C.H. Choi, S. Re, M. Feig, Y. Sugita, Chem. Phys. Lett. 2012, 539-540, 218.

6 N. Takenaka, Y. Kitamura, Y. Koyano, M. Nagaoka, J. Chem. Phys. 2012, 137, 024501.

7 H. Miyamoto, M. Aida, Chem. Lett. 2013, in press.

8 G. Wada, E. Tamura, M. Okina, M. Nakamura, Bull. Chem. Soc. Jpn. 1982, 55, 3064.

9 M. A. Slifkin, S. M. Ali, J.Mol. Liq. 1984, 28, 215.

10 M. J. Frisch, et al., Gaussian 03 (Revision A.02), Gaussian, Inc., Wallingford CT, 2004.

11 S. Kirkpatrick, C. D. Gelatt, Jr., M. P. Vecchi, Science, 1983, 220, 671

12 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, J. Chem. Phys. 1983, 79, 926.

13 M. Aida, H. Yamataka, M. Dupuis, Int. J. Quantum Chem. 2000, 77, 199.

14 M. Dupuis, M. Aida, Y. Kawashima, K. Hirao, J. Chem. Phys. 2002, 117, 1242.

15 R.J. Bartlett, I. Shavitt, G.D. Purvis, J. Chem. Phys. 1979, 71, 281.

16 C. M. Breneman, K. B. Wiberg, J. Comput. Chem. 1990, 11, 361.

17 Appendix B.

18 M. Dupuis, G. K. Schenter, B. G. Garrett, E. E. Arcia, J. Mol. Struct. (Theochem) 2003, 632, 173.

19 M. Dupuis, HONDO2004, based on HONDO-95 available from the Quantum Chemistry Program Exchange, Indiana University.

20 M. Ohisa, H. Yamataka, M. Dupuis, M. Aida, Phys. Chem. Chem. Phys. 2008, 10, 844.

21 Y. Kitamura, N. Takenaka, Y. Koyano, M. Nagaoka, Chem. Phys. Lett., 2011, 514, 261

Chapter 4

General Conclusion

In this thesis, I showed how to represent solvation effect correctly.

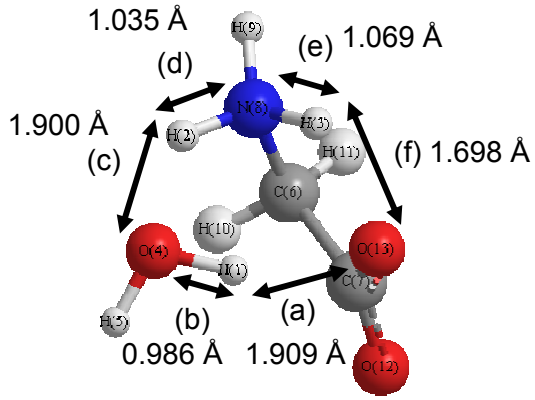
Free energy calculation in the aqueous solution is generally performed by low level of theory because there are a lot of molecules and configurations of solvent. The Multistage strategy method can calculate free energy by means of high level of theory. That is because multistage method can separate sampling method into lower level of sampling and higher level recalculation, besides, can higher level of recalculation by means of separating the system into three layers, bulk solvents (low level of theory), direct hydrating solvent (middle level of theory), and reaction center (high level of theory). Especially, in the case of polar solute molecule, we found it is important to treat solvent molecules around solute as QM solvent for calculating solvation free energy correctly, because polar molecule interact other molecules strongly.

Appendix

Appendix A

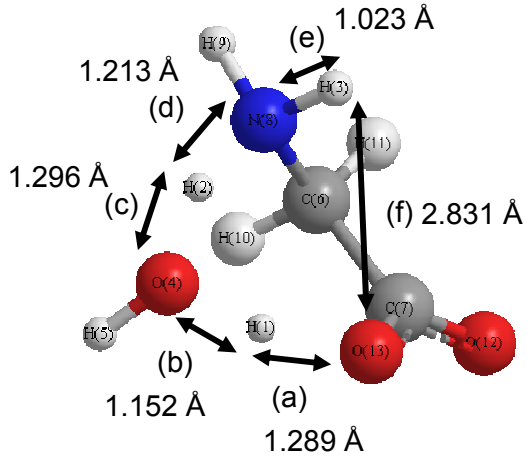
s.1.1 Coordinates of QM complexes of Figure 1

Z-1W



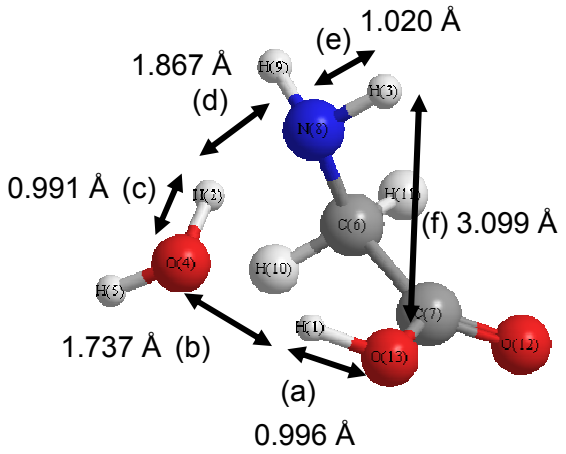
H(1)	-1.262621	-1.234252	0.212065
H(2)	-1.587967	0.839701	-0.257021
H(3)	-0.657080	0.767535	1.095604
O(4)	-2.012624	-1.007139	-0.386604
H(5)	-1.949275	-1.638676	-1.119376
C(6)	0.466445	1.122157	-0.474200
C(7)	0.996937	-0.233677	0.104458
N(8)	-0.832756	1.359745	0.223476
H(9)	-1.080820	2.333671	0.400755
H(10)	0.320388	1.064758	-1.551583
H(11)	1.155543	1.934207	-0.238457
O(12)	1.970235	-0.726842	-0.466524
O(13)	0.306243	-0.630524	1.117145

TS-1W-med



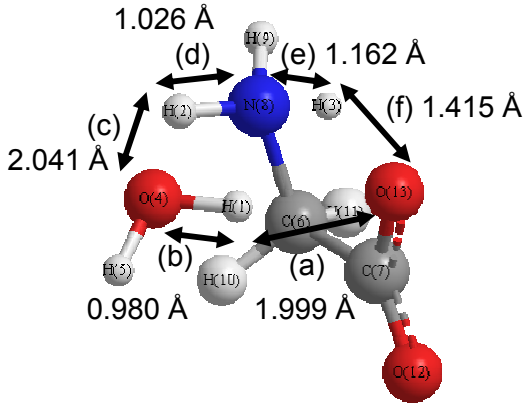
H(1)	0.979743	-1.163382	-0.038685
H(2)	1.662183	0.377255	-0.034735
H(3)	0.775452	1.198087	-1.364813
O(4)	2.062283	-0.836409	0.179613
H(5)	2.254968	-1.034666	1.111601
C(6)	-0.361693	1.087303	0.368458
C(7)	-1.046975	-0.254037	-0.005537
N(8)	0.927704	1.286938	-0.357070
H(9)	1.326905	2.211035	-0.180812
H(10)	-0.136475	1.099522	1.438887
H(11)	-1.055633	1.898484	0.148912
O(12)	-2.274631	-0.258217	0.076583
O(13)	-0.268786	-1.229687	-0.350996

N-1W-med



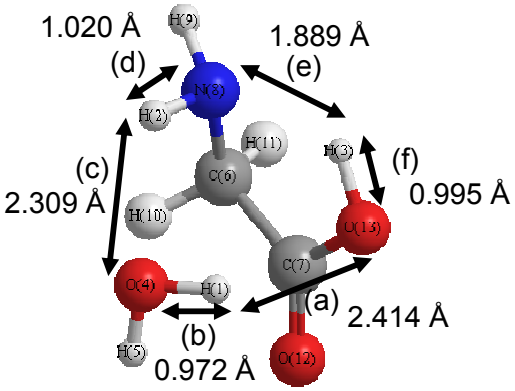
H(1)	0.679367	-1.148319	-0.057707
H(2)	2.067222	0.173793	-0.158139
H(3)	0.354716	1.478689	-1.371487
O(4)	2.363743	-0.744259	0.066908
H(5)	2.735047	-0.694993	0.961398
C(6)	-0.401946	1.027194	0.482672
C(7)	-1.075202	-0.272852	0.046081
N(8)	0.701646	1.420191	-0.413807
H(9)	1.006979	2.362180	-0.170140
H(10)	0.013183	0.884967	1.487262
H(11)	-1.201845	1.770426	0.548177
O(12)	-2.288185	-0.343927	-0.041925
O(13)	-0.288470	-1.323580	-0.214387

TS-1W-in



H(1)	-1.274701	-1.261995	0.261636
H(2)	-1.638680	0.970715	-0.207401
H(3)	-0.510576	0.624983	1.064909
O(4)	-1.995517	-1.032402	-0.361564
H(5)	-1.998691	-1.747448	-1.015613
C(6)	0.412302	1.074627	-0.579542
C(7)	0.993971	-0.211035	0.088055
N(8)	-0.800615	1.374822	0.225604
H(9)	-0.934076	2.351739	0.482264
H(10)	0.168370	0.892041	-1.624892
H(11)	1.124927	1.897220	-0.507914
O(12)	1.939457	-0.780225	-0.447246
O(13)	0.334821	-0.503943	1.173398

N-1W-in



H(1)	-0.460576	1.908992	0.668289
H(2)	1.775186	0.764979	-0.043416
H(3)	0.468693	-0.401430	1.485559
O(4)	-0.054152	2.168199	-0.176052
H(5)	-0.778632	2.050208	-0.812407
C(6)	0.714231	-0.838484	-0.691432
C(7)	-0.626005	-0.614283	-0.004156
N(8)	1.769668	-0.240756	0.125884
H(9)	2.684006	-0.616073	-0.111930
H(10)	0.631866	-0.453043	-1.712757
H(11)	0.875669	-1.919974	-0.753177
O(12)	-1.695998	-0.621819	-0.579956
O(13)	-0.514005	-0.412850	1.327531

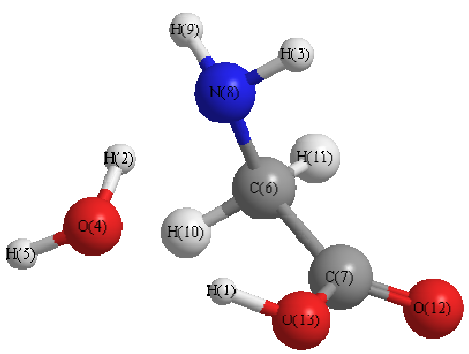
s.1.2 Lennard-Jones parameters used in the QM/MM calculation

$$E^{vdW} = \sum_{i>j}^{M}\sum 4\varepsilon_{m_i m_j}\left[\left(\frac{\sigma_{m_i m_j}}{R_{m_i m_j}}\right)^{12} - \left(\frac{\sigma_{m_i m_j}}{R_{m_i m_j}}\right)^{6}\right]$$

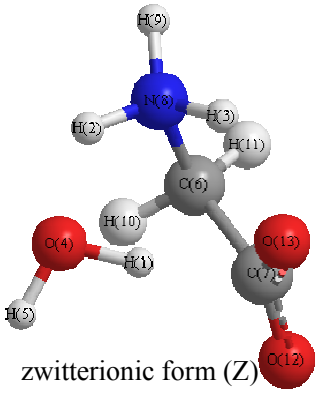
$$\sigma_{m_i m_j} = \left(\sigma_{m_i}\sigma_{m_j}\right)^{1/2}$$

$$\varepsilon_{m_i m_j} = \left(\varepsilon_{m_i}\varepsilon_{m_j}\right)^{1/2}$$

Here, M is the total number of the atoms of MM level, and m_i and m_j are the indices for the atoms.



neutral form (N)



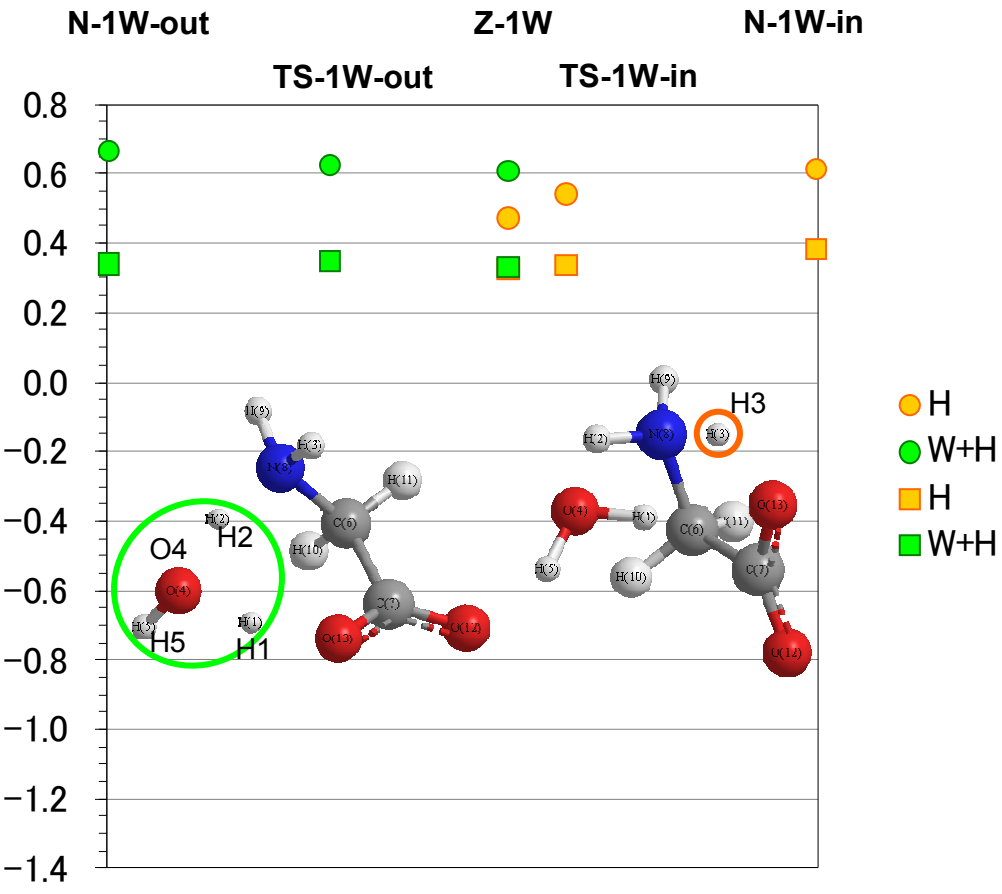
zwitterionic form (Z)

The parameters are taken from ref.15.

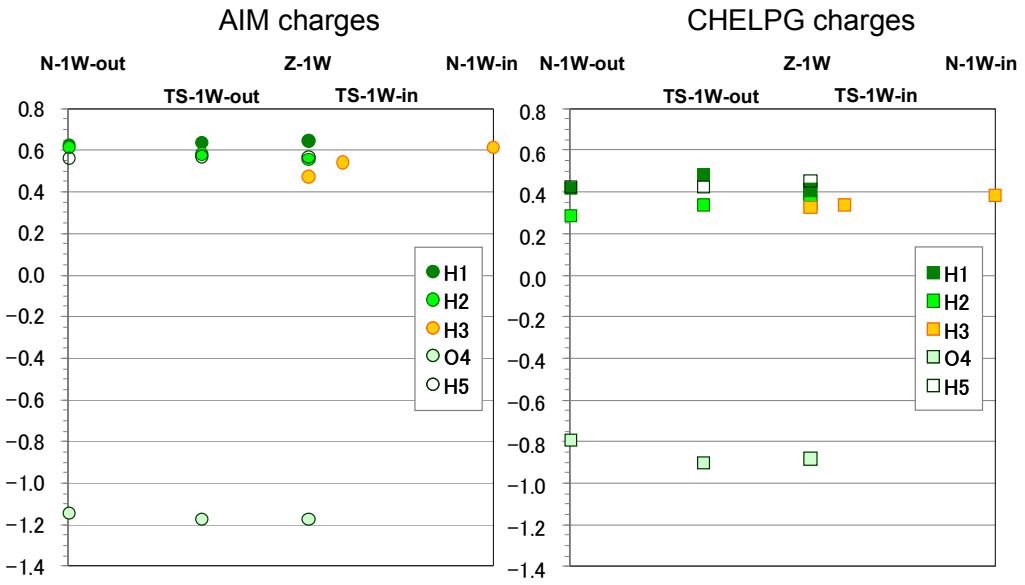
neutral form			zwitterionic form		
	σ(Å)	ε(kcal/mol)		σ(Å)	ε(kcal/mol)
H(1)	1.3	0.1	H(1)	1.3	0.1
H(2)	1.3	0.1	H(2)	1.3	0.1
H(3)	1.3	0.1	H(3)	1.3	0.1
O(4)	3.6	0.15	O(4)	3.6	0.15
H(5)	1.3	0.1	H(5)	1.3	0.1
C(6)	3.8	0.08	C(6)	3.8	0.08
C(7)	3.8	0.08	C(7)	3.8	0.08
N(8)	3.9	0.2	N(8)	3.35	0.15
H(9)	1.3	0.1	H(9)	1.3	0.1
H(10)	2.6	0.08	H(10)	2.6	0.08
H(11)	2.6	0.08	H(11)	2.6	0.08
O(12)	3.2	0.15	O(12)	3.2	0.15
O(13)	3.6	0.15	O(13)	3.2	0.15

s.1.3 AIM and CHELPG charges of transferring hydrogen atom of glycine

AIM group charges and CHELPG group charges are plotted in respect to N-1W-out, TS-1W-out, Z-1W, TS-1W-in, and N-1W-in respectively. AIM group charges are plotted in green and orange circle. CHELPG charges are plotted in green and orange square.



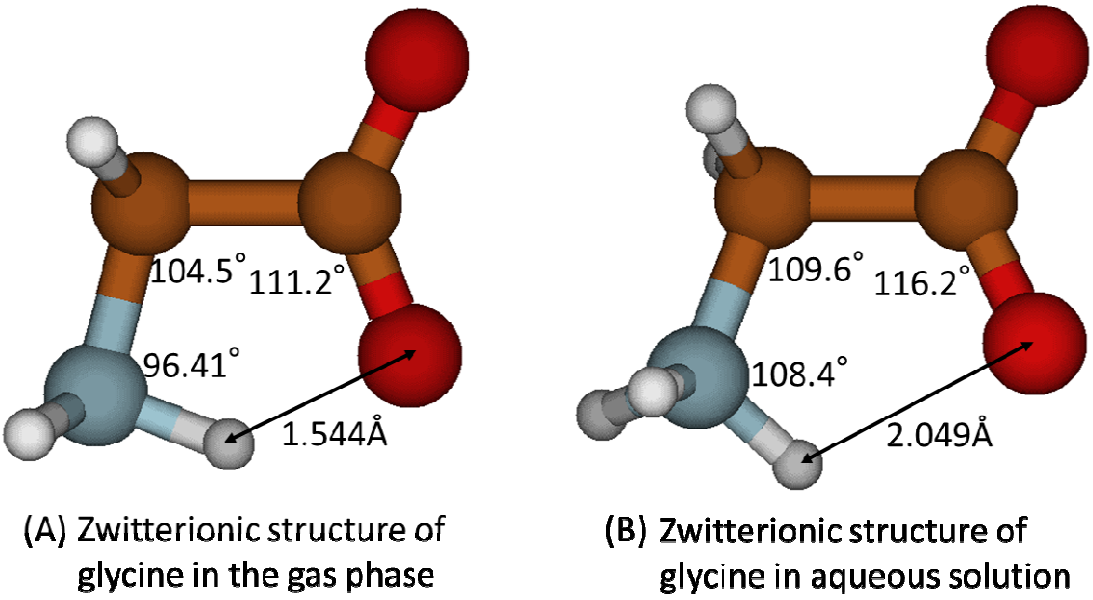
AIM atomic charges and CHELPG atomic charges are plotted in respect to N-1W-out, TS-1W-out, Z-1W, TS-1W-in, and N-1W-in respectively. .



Appendix B

s.2.1 Structure of Zwitterionic forms of glycine in the gas phase and in aqueous solution

Figure S1. Zwitterionic forms of glycine in the gas phase (A) and in aqueous solution (B). Angles are in degrees and lengths are in Å.

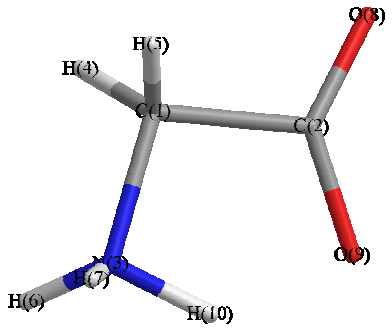


(A) Zwitterionic structure of glycine in the gas phase

(B) Zwitterionic structure of glycine in aqueous solution

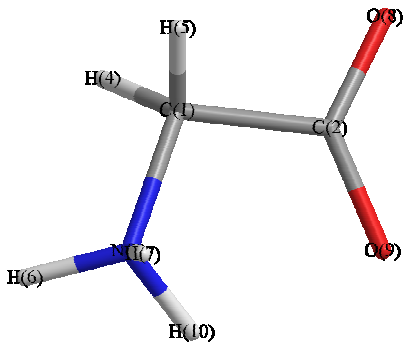
s.2.2 Coordinates of zwitterionic forms of glycine in the gas phase and aqueous solution

Table S1. XYZ coordinates (in Å) of zwitterionic structure of glycine in the gas phase



C(1)	-0.8208251000	-0.6024053000	0.0000000000
C(2)	0.0073968000	0.7304470000	0.0000000000
N(3)	0.2147741000	-1.6885116000	0.0000000000
H(4)	-1.4379211000	-0.6820973000	-0.8811160000
H(5)	-1.4379211000	-0.6820973000	0.8811160000
H(6)	0.2007042000	-2.2729017000	-0.8169766000
H(7)	0.2007042000	-2.2729017000	0.8169766000
O(8)	-0.6381631000	1.7453481000	0.0000000000
O(9)	1.2341133000	0.4997356000	0.0000000000
H(10)	1.0728063000	-1.0365891000	0.0000000000

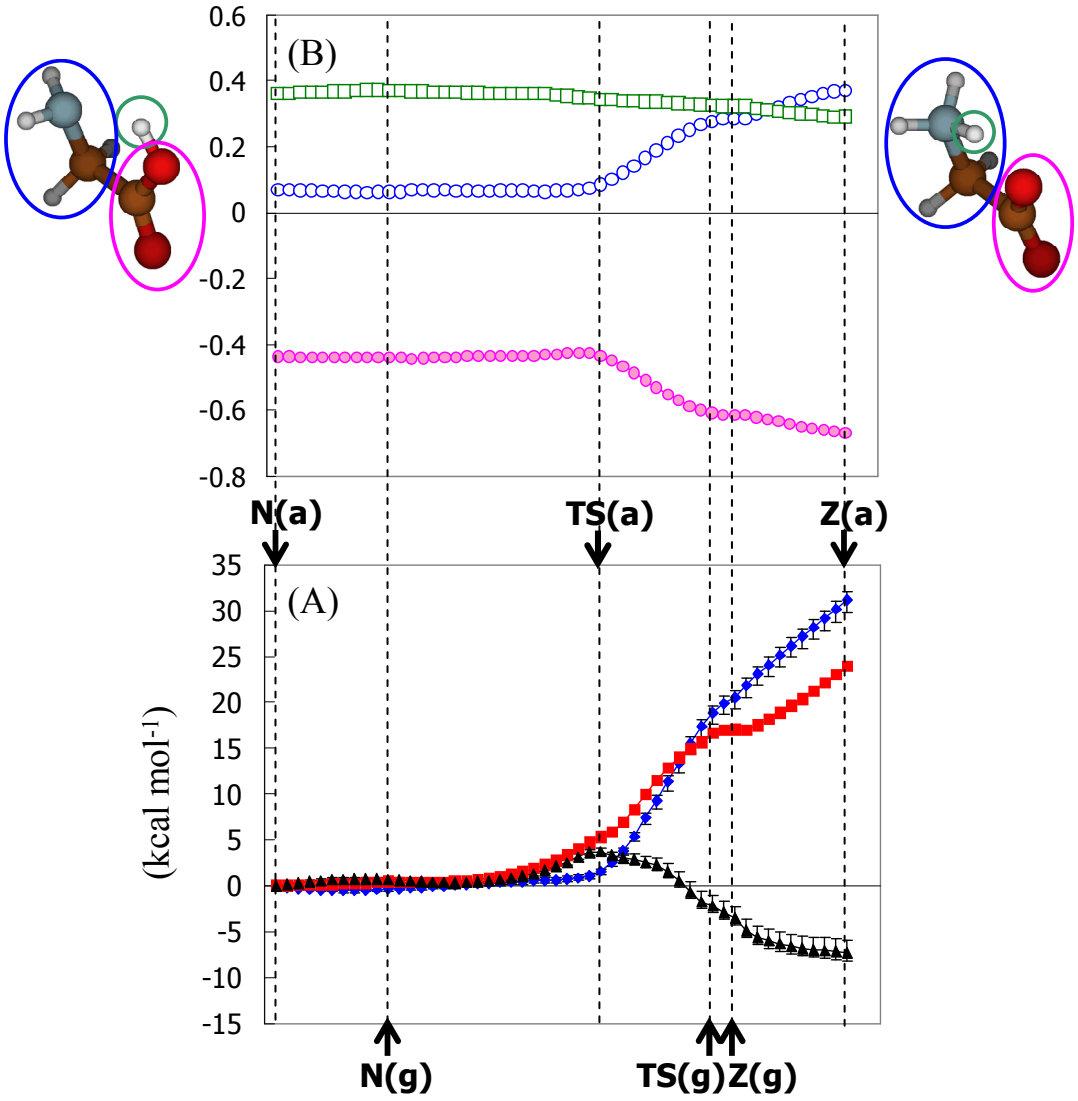
Table S2. XYZ coordinates (in Å) of zwitterionic structure of glycine in aqueous solution



C(1)	-0.7158772638	-0.5496393957	-0.0658149055
C(2)	0.0697890220	0.7627276650	0.0169425425
N(3)	0.2302701527	-1.7018602845	-0.0434949575
H(4)	-1.2806869677	-0.5791272566	-0.9837189375
H(5)	-1.4013165841	-0.6292230816	0.7644435300
H(6)	0.0141106849	-2.4280634291	-0.7157565870
H(7)	0.3017067332	-2.1204562990	0.8808630660
O(8)	-0.6139211620	1.7865661266	0.0793806848
O(9)	1.2994641607	0.6611982506	0.0328289998
H(10)	1.1596115283	-1.3587413494	-0.2827810676

s.2.3 Group charge changes along reaction coordinate

Figure S2. (A) Helmholtz free energy change (black full triangles), solute potential energy change (red full squares) and solvation free energy change (blue full diamond) along the hydrogen transfer route of glycine in aqueous solution, according to Eq.(5). The solvation free energy change is plotted with the reversed sign. N(a), TS(a) and Z(a) indicate N, TS and Z forms in aqueous solution, respectively. N(g), TS(g) and Z(g) indicate those in the gas phase. The standard deviations for each step are indicated for Helmholtz free energy change (black) and solvation free energy change (blue). (B) Changes in group charges of the solute along the transfer route.



公表論文

(1) Ab Initio QM/MM-MC Study on Hydrogen Transfer of Glycine Tautomerization in Aqueous Solution: Helmholtz Energy Changes along Water-mediated and Direct Processes.

H. Miyamoto, M. Aida, *Chemistry Letters*, **2013**, 42, 598.

(2) Helmholtz Energy Change between Neutral and Zwitterionic Forms of Glycine in Aqueous Solution using Ab Initio Expanded QM/MM-MC with QM Solvent.

H. Miyamoto, M. Aida, *Chemistry Letters*, **2013**, 42, 1010.