Chapter 4

Application to the intramolecular proton transfer of glycine in aqueous solution

It is well known that glycine, the smallest aminoacid, is found in its neutral form (GLYN) in the gas phase, whereas the zwitterionic form (GLYZ) dominates in crystalline or aqueous media. *Ab initio* calculations using flexible basis sets show that the zwitterionic form does not exist as a stable species in the gas phase. Consequently, if one wants to explore the geometry of the zwitterionic form of glycine, the interactions between GLYZ and the solvent must be included into the calculation using continuum models or cluster models in which glycine and more than two water molecules are included.

On the basis of the equilibrium constant for the GLYZ and GLYN forms in water the relative free energy and enthalpy have been estimated as 7.27 and 10.3 kcal/mol, respectively.[65] Various NMR relaxation techniques have been used to study the kinetics of intramolecular proton transfer from GLYZ to GLYN in aqueous solution and all yielded a free energy of activation of about 14.3 kcal/mol for the proton transfer.[66-68] Slifkin and Ali[68] measured the rate constant for the proton transfer reaction at several temperatures and determined the contributions of enthalpy and $-T\Delta S$ to the free energy of activation of -0.2 and 14.5 kcal/mol, respectively.

The MD techniques have been used to study this intramolecular proton transfer reaction of glycine in aqueous solution. Nagaoka *et al.*[30,31] used an empirical valence bond (EVB) method[32–36] as the force field of the reaction, and calculated the free energy profile of the reaction by using MD simulation and statistical perturbation theory. Their free energy difference between the conformers, which was optimized for the intramolecular proton transfer reaction, agreed with experiments. However, they used the false minimum for GLYZ of *ab initio* calculation in the gas phase, so their agreement with experiment is due to the underestimation of the interaction energy between GLYZ and water molecules.

Tuñón *et al.* used a density functional / molecular mechanics (DF/MM) coupled potential for the MD simulation, and calculated some trajectories for the tautomerization from the GLYN to the GLYZ in aqueous solution. They treated the reacting glycine with the density functional theory (DFT),[37] so they could incorporate dynamically the solvent effects into the electronic state and consider the electron correlation which is very important to calculate the energy profile of the reaction. But they did not carry out detail thermodynamic analysis, because their calculation costs vast computational time and they could not calculate large number of trajectories. Their calculation suggested that the tautomerization occurred easily and the activation energy obtained by experiment comes from the conformational change of the GLYN.

In this chapter, the method proposed in Chapter 2 has been applied to the intramolecular proton transfer reaction of GLY. The EPPFs were determined over the whole reaction path. By using the MC simulation with the determined EPPF, the free energy change of the reaction was calculated and the reaction mechanism was examined in detail.

4.1 Ab initio GB calculation

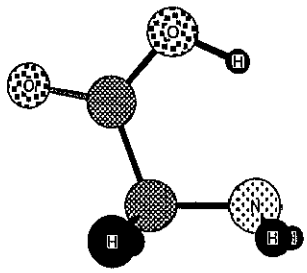
In all *ab initio* calculations presented below, the 6-31++G** basis set was used, and the calculations in solution and in the gas phase mean the calculations of *ab initio* GB method using $\varepsilon = 79$ and 1, respectively. The geometries of GLYZ and GLYN were determined under the restriction of C_s symmetry.

With respect to the GLYN conformer in the gas phase, early *ab initio* calculations predicted $\mathbf{1}_N$ in Figure 4.1 as the energy global minimum,[69] while in spectroscopic experiments $\mathbf{2}_N$ has been detected.[70] It was finally recognised that the apparent contradiction between experiment and theory could be overcome by taking into account the different dipole moment of both conformers[71,72] and that the *ab initio* predictions were correct. Nowaday, the potential energy surface of GLYN in the gas phase has been studied extensively.[73–75] On the other hand, with respect to the GLYZ conformer, *ab initio* calculation with the standard structure predicted $\mathbf{1}_Z$ as the energy minimum in the gas phase which is stabilized by the intramolecular hydrogen bonding. In aqueous solution, $\mathbf{2}_Z$, which may be stabilized by the intermolecular hydrogen bonding,[76] is more stable than $\mathbf{1}_Z$. However, the energy difference between $\mathbf{1}_Z$ and $\mathbf{2}_Z$ is expected to be small.[76]

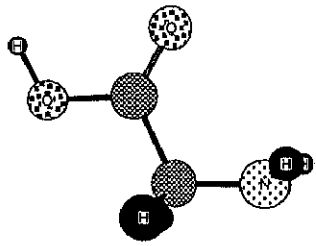
Since the interest in this study is the proton transfer reaction in Figure 4.2, the conformers $\mathbf{1}_Z$ and $\mathbf{1}_N$ were used for GLYZ and GLYN, respectively, in the MC simulation, while $\mathbf{2}_Z$ and $\mathbf{2}_N$ were used in the *ab initio* GB calculations. The relative energies are shown in Table 4.1.

The two conformers of GLYZ have comparable energies; $\mathbf{2}_Z$ lies slightly higher (0.50 kcal/mol) than $\mathbf{1}_Z$ in aqueous solution. In the previous studies of GLYZ in the gas phase by using the standard geometry,[76,78,79] $\mathbf{1}_Z$ is more stable than $\mathbf{2}_Z$ by 1.2-5.1 kcal/mol, so it was found that $\mathbf{1}_Z$ is stabilized by solvation more than $\mathbf{2}_Z$. In our previous MC simulation of GLYZ in aqueous solution which is based on MIDI-4(d) basis set,[76] $\mathbf{2}_Z$ is more stable than $\mathbf{1}_Z$ by 3.3 kcal/mol, because the water

GLYN

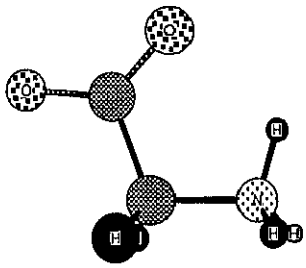


1_N

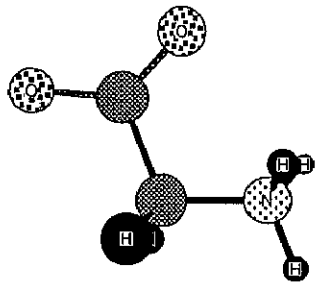


2_N

GLYZ



1_Z



2_Z

Figure 4.1: The conformers of glycine.

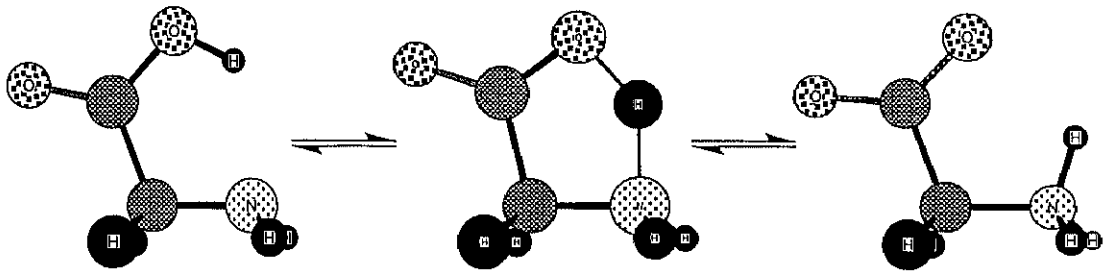


Figure 4.2: Intramolecular proton transfer in glycine.

Table 4.1: Relative energies (in kcal/mol) for conformers of GLYZ and GLYN

	GLYZ		GLYN		Solvent model	Level of calculation
	1_Z	2_Z	1_N	2_N		
This study	0.00	0.02	0.0	0.17	*ab initio* GB	HF/6-31++G**
This study	0.00	0.50	0.0	2.15	*ab initio* GB	MP2/6-31++G**[a]
Ref73	—	—	3.2	0.0	gas phase	HF/6-31G*
Ref73	—	—	1.8	0.0	gas phase	MP2/6-31G*
Ref74	—	—	2.9	0.0	gas phase	HF/TZ2P+f
Ref74	—	—	1.1	0.0	gas phase	CCSD(T)/DZP
Ref77	—	—	2.7	0.0	gas phase	HF/6-31+G**
Ref77	—	—	0.7	0.0	gas phase	MP2/6-31+G**
Ref77	—	—	0.0	0.6	continuum model[b]	HF/6-31+G**
Ref77	—	—	0.0	2.7	continuum model[b]	MP2/6-31+G**
Ref44	0.7	0.0	—	—	*ab initio* GB	HF/6-31G
Ref76	0.0	1.2	—	—	gas phase	HF/6-31++G*
Ref76	1.4	0.0	—	—	*ab initio* GB	HF/MIDI-4(d)
Ref76	0.3	0.0	—	—	*ab initio* GB	HF/6-31++G*
Ref76	3.3	0.0	—	—	MC simulation	HF/MIDI-4(d)
Ref78	0.0	2.3	—	—	gas phase	HF/6-31+G*
Ref78	0.0	4.5	—	—	gas phase	MP4(SDTQ)/6-31+G*
Ref78	0.0	5.1	—	—	gas phase	MP4(SDTQ)/6-31+G*[c]
Ref79	0.0	2.5	—	—	gas phase	HF/4-31G
Ref79	0.7	0.0	—	—	continuum model[d]	HF/4-31G
Ref80	—	—	1.4±0.4	0.0	gas phase	expt

[a] The structure was optimized by using *ab initio* GB HF/6-31++G** calculation. [b] The elipsoidal cavity and multipole expansion model of Rivail *et al.*[81,82] was used. [c] The value included a vibrational zero-point energy. [d] The dielectric PCM model[83] was used.

molecule which is hydrated simultaneously to the amino and carboxyl groups is destabilized by the intramolecular hydrogen-bonding in $\mathbf{1}_Z$. *Ab initio* GB method with same basis set (MIDI-4(d)) predicted that $\mathbf{2}_Z$ is more stable than $\mathbf{1}_Z$ by 1.4 kcal/mol.[76] In *ab initio* GB method, the effect of the destabilization of hydration is not incorporated explicitly, and the destabilization energy is underestimated 1.9 kcal/mol by comparison to the result of MC simulation. By taking this correction into account, the present *ab initio* GB MP2/6-31++G** calculation suggests that $\mathbf{2}_Z$ is more stable than $\mathbf{1}_Z$ by 1.4 kcal/mol.

With respect to the GLYN conformers, they have comparable energies in the HF calculation; $\mathbf{2}_N$ lies slightly more stable (0.17 kcal/mol) than $\mathbf{1}_N$. But inclusion of the electronic correlation makes the energy difference 2.15 kcal/mol.

The geometries of $\mathbf{1}_Z$, $\mathbf{1}_N$, and the transition state (TS) between them were optimized in aqueous solution. The geometrical parameters are shown in Table 4.2. The structure of $\mathbf{1}_Z$ was optimized by using *ab initio* HF/6-31+G* calculation, in which GLYZ has a false minimum structure. The N4-H10 distance is larger in the gas phase than in solution, while the O9-H10 distance is smaller in the gas phase than in solution. The C1-H4-H10, C4-C1-C7, and C1-C7-C9 angles are all smaller in the gas phase, indicating that the N-H hydrogen atom is strongly hydrogen-bonded to the carboxyl group. It is also interesting that GLYZ has the C-O bonds with different lengths of 1.207 and 1.245 Å in the gas phase, while it has similar CO bond lengths, 1.242 and 1.240 Å, in aqueous solution.

The structural differences of $\mathbf{1}_N$ in the gas phase and in solution are relatively small. This reflects the relative magnitudes of the dipole moment, listed in Table 4.3, and the relative magnitudes of the solvation energies.

The TS structure is characterized by the N4-H10 and O9-H10 lengths, 1.335 and 1.179 Å, respectively. These values suggest that the structure of TS is close to that of $\mathbf{1}_N$. This trend agrees well with the fact that the reaction undergoes exothermic

Table 4.2: Structural parameters(lengths, Å; angles, deg) of GLYZ, GLYN and the transition state (TS) between GLYZ and GLYN in aqueous solution and in the gas phase(in parentheses)[a]

	GLYZ[b]	TS	GLYN		GLYZ[b]	TS	GLYN
C1-H2,3	1.081 (1.079)	1.083	1.086 (1.084)	N4-C1-H2,3	109.34 (109.59)	112.17	112.38 (112.00)
C1-N4	1.483 (1.504)	1.466	1.452 (1.454)	C1-N4-H5,6	112.24 (114.07)	113.50	112.82 (113.05)
N4-H5,6	1.011 (1.005)	1.006	1.003 (0.998)	C1-N4-H10	107.20 (98.89)	89.18	82.78 (81.91)
N4-H10	1.011 (1.053)	1.335	1.995 (2.043)	N4-C1-C7	110.92 (105.44)	104.51	111.72 (112.65)
C1-C7	1.526 (1.566)	1.529	1.520 (1.526)	C1-C7-O8	115.14 (115.00)	122.27	122.13 (121.51)
C7-O8	1.240 (1.207)	1.217	1.205 (1.187)	C1-C7-O9	116.44 (111.99)	110.60	115.92 (115.65)
C7-O9	1.242 (1.245)	1.279	1.310 (1.318)	C7-O9-H10	89.90 (94.35)	96.54	107.92 (108.85)
O9-H10	2.022 (1.644)	1.179	0.968 (0.952)				
H10-N4 -C1-H2,3	120.66 (119.82)	118.62	119.68 (120.09)	C7-C1 -N4-H5,6	119.42 (117.13)	117.62	118.42 (118.05)

[a] The structures optimezed by *ab initio* GB HF/6-31++G** calculation, and the number of atoms are in the following figure. [b] The structure of GLYZ in the gas phase was optimized by using *ab initio* HF/6-31+G* calculation.

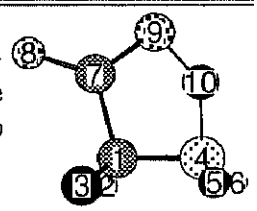


Table 4.3: Dipole moments(in Debye) of GLYZ, GLYN and the transition state (TS) between GLYZ and GLYN in aqueous solution and in the gas phase(in parentheses)[a]

	GLYZ[b]	TS	GLYN
Dipole moment	14.520 (10.750)	9.822	7.146 (6.018)

[a] These values and structures were calculated by using *ab initio* GB HF/6-31++G**. [b] The structure of GLYZ in the gas phase was optimized by using *ab initio* HF/6-31+G* calculation.

from GLYN to GLYZ, as described below.

Table 4.4 shows the Löwdin charges of GLYZ, GLYN and TS. In GLYZ, the localization of the positive charge on NH_3 and negative charge on CO_2 is promoted by solvation of continuum media. The charges of two oxygen atoms in the carboxyl group of GLYZ are non-equivalent in the gas phase.

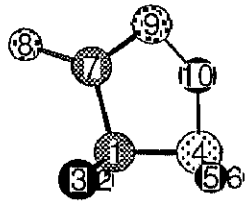
The reaction coordinate of the proton transfer reaction was defined as the distance between the N and H atoms, R_{N-H}. The energy profiles calculated along the reaction coordinate by using the HF and MP2 levels are shown in Figures 4.3 and 4.4, respectively. In each figure, E1(solid line, open circle) and E2(broken line, triangle) are *ab initio* energies in the the gas phase, E_0^{gas}, with the structures optimized in gas phase and in aqueous solution, respectively. E3(dotted line, square) is the energy of the molecule polarized by continuum media, E_0^{sol}, which dose not include the interaction energy between glycine and continuum media. E4(stitch line, black circle) is the energy of the molecule with continuum media, E_{sol}^{SCF}. E3 and E4 were calculated with the structure optimized in aqueous solution. The E4 energies in Figures 4.3 and 4.4 are also shown in Figure 4.5 for detail discussion. In these Figures, the R_{N-H}distances of $\mathbf{1}_Z$, $\mathbf{1}_N$ and the transition state between $\mathbf{1}_Z$ and $\mathbf{1}_N$ in aueous solution are 1.011, 1.995 and 1.335 Å, respectively.

Figures 4.3 and 4.4 indicate that the HF and MP2 calculations give qualitatively same variations for E1, E2, E3 and E4 energies. Let me consider about the result in Figure 4.3. The profile of E1 has the small shoulder at about R_{N-H}= 1.1 Å but is monotonously decreased from GLYZ to GLYN, which supports that the GLYZ is not stable in the gas phase.[78,84,85] The energy maximum at 1.2 Å of E2 and E3 arise from the drastic change of structure in this region. This change may be caused by the wrong definition of a reaction coordinate for this range of R_{N-H}. However, the primary objective in this study is to compare the relative energies among GLYZ, GLYN and GLYTS and this reaction coordinate connecting these three structures

Table 4.4: Löwdin charges of GLYZ, GLYN and the transition state (TS) between GLYZ and GLYN in aqueous solution and in the gas phase(in parentheses)[a]

	GLYZ[b]	TS	GLYN		GLYZ[b]	TS	GLYN
C1	-0.075 (-0.261)	-0.083	-0.090 (-0.107)	C7	0.244 (0.246)	0.284	0.317 (0.295)
H2,3	0.181 (0.237)	0.178	0.174 (0.156)	O8	-0.646 (-0.478)	-0.543	-0.487 (-0.376)
N4	-0.059 (-0.536)	-0.325	-0.482 (-0.483)	O9	-0.643 (-0.592)	-0.496	-0.336 (-0.327)
H5,6	0.270 (0.357)	0.242	0.222 (0.209)	H10	0.276 (0.432)	0.322	0.288 (0.268)
CH_2	0.287 (0.213)	0.273	0.258 (0.205)	NH_3^+	0.757 (0.610)	0.481	— (—)
NH_2	— (—)	0.159	-0.038 (-0.065)	CO_2^-	-1.045 (-0.824)	-0.755	— (—)
CO_2H	— (—)	-0.433	-0.218 (-0.140)				

[a] These values and structures were calculated by using *ab initio* GB HF/6-31++G** level, and the number of atoms are in the following figure. [b] The structure of GLYZ in the gas phase was optimized by using *ab initio* HF/6-31+G* calculation.



is suitable for the present study.

The E1 curve was calculated with the structure optimized in the gas phase ($\varepsilon = 1$) while the E2 curve was calculated with the structure optimized in aqueous solution ($\varepsilon = 79$). Both energies do not include the solvation energy and the difference between E1 and E2 energy curves comes from the structural relaxation effect caused by solvation. This effect is small, 0.57 kcal/mol, for GLYN(R_{N-H}= 2.00Å) but considerably large, 4.58 kcal/mol, for GLYZ(R_{N-H}= 1.00Å), since the structure of GLYZ is affected by solvation more than GLYN.

The difference between E2 and E3 comes from the electronic relaxation effect caused by solvation. These effects are 2.28 and 12.26 kcal/mol for GLYN(R_{N-H}= 2.00Å) and for GLYZ(R_{N-H}= 1.00Å), respectively. In this case, the GLYZ is affected by solvation more than GLYN, and the electronic relaxation effect is larger than the structural one. It should be stressed that this solvation effect supports strongly the use of the polarized electronic states for determination of PF. In the present method, the polarization self-energy, E_{pol}, corresponds to the sum of the energies for the structural and electronic relaxation effects, namely the difference between E1 and E3.

In the results of MP2 calculations shown in Figure 4.4, the E1 curve has no shoulder which appears in the E1 curve in Figure 4.3 and corresponds to the local minimum of GLYZ at HF/6-31+G* level; it disappeared by inclusion of electronic correlation.[78] The structural and electronic relaxation effects and polarization self-energy are 3.59, 18.54 and 22.13 kcal/mol, respectively, for GLYZ(R_{N-H}= 1.00Å). The electronic relaxation effect is significantly enlarged by consideration of electronic correlation in the calculation. This increase is attributed to the enhancement of polarization by solvation and the importance of electronic correlation for such a polarized molecule.

The E4 energies of HF and MP2 calculations are compared in Figure 4.5. It

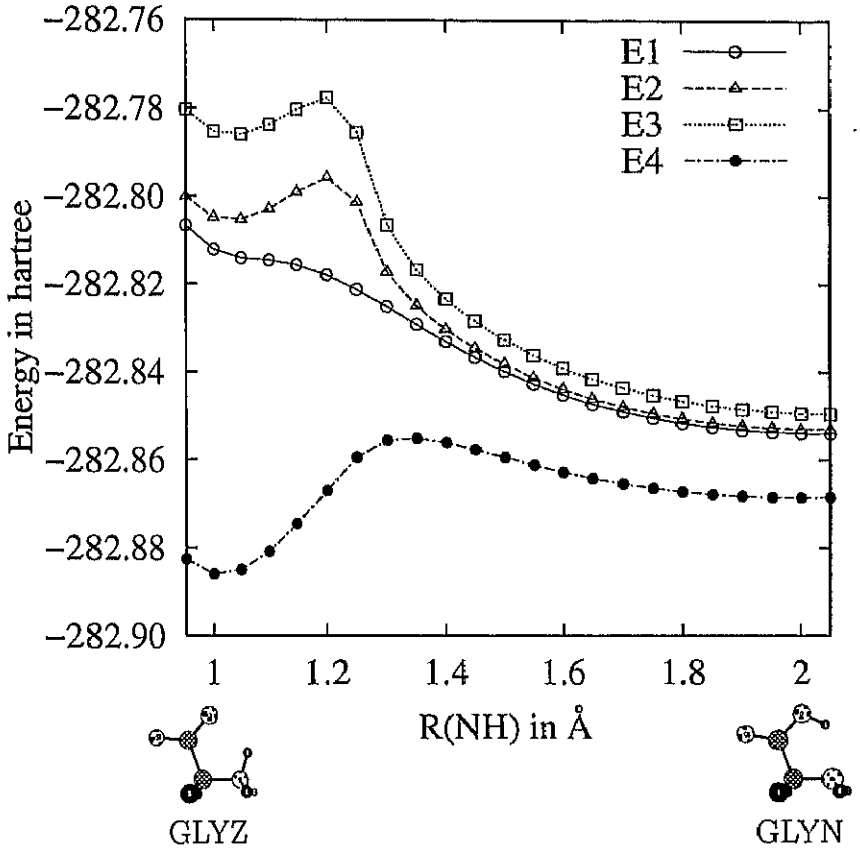


Figure 4.3: HF energy profiles for the proton transfer reaction of glycine: E1(solid line, circle) and E2(broken line, triangle) are the *ab initio* 6-31++G** energy, E_0^{gas}, with the structures optimized in the gas phase and in aqueous solution, respectively. E3(dotted line, square) is the energy of isolated molecule, E_0^{sol}, with the structure optimized in aqueous solution. E4(stitch line, black circle) is the *ab initio* GB 6-31++G** energy, E_{sol}^{SCF}, with the structure optimized in aqueous solution.

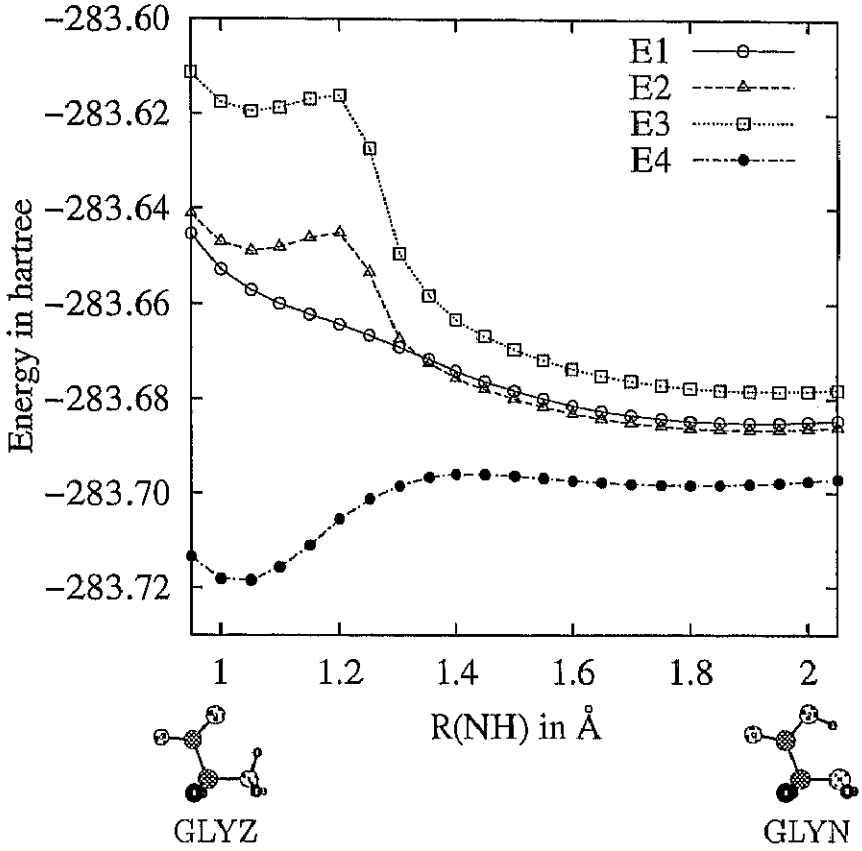


Figure 4.4: MP2 energy profiles for the proton transfer reaction of glycine: E1(solid line, circle) and E2(broken line, triangle) are the *ab initio* 6-31++G** energy, E_0^{gas}, with the structures optimized in the gas phase and in aqueous solution, respectively. E3(dotted line, square) is the energy of isolated molecule, E_0^{sol}, with the structure optimized in aqueous solution. E4(stitch line, black circle) is the *ab initio* GB 6-31++G** energy, E_{sol}^{SCF}, with the structure optimized in aqueous solution.

takes notice that these energies were calculated by using *ab initio* GB method for the proton transfer in Figure 4.2. The main difference in two curves is the activation energies. This result is also shown in Table 4.5. The activation energy considerably decreaces by inclusion of electronic correlation, in particular the activation energy from GLYN to GLYZ is very small. However, it must be noted that the structural and electronic relaxation effects are not self–consistent by MP2 calculation, as in the HF calculation.

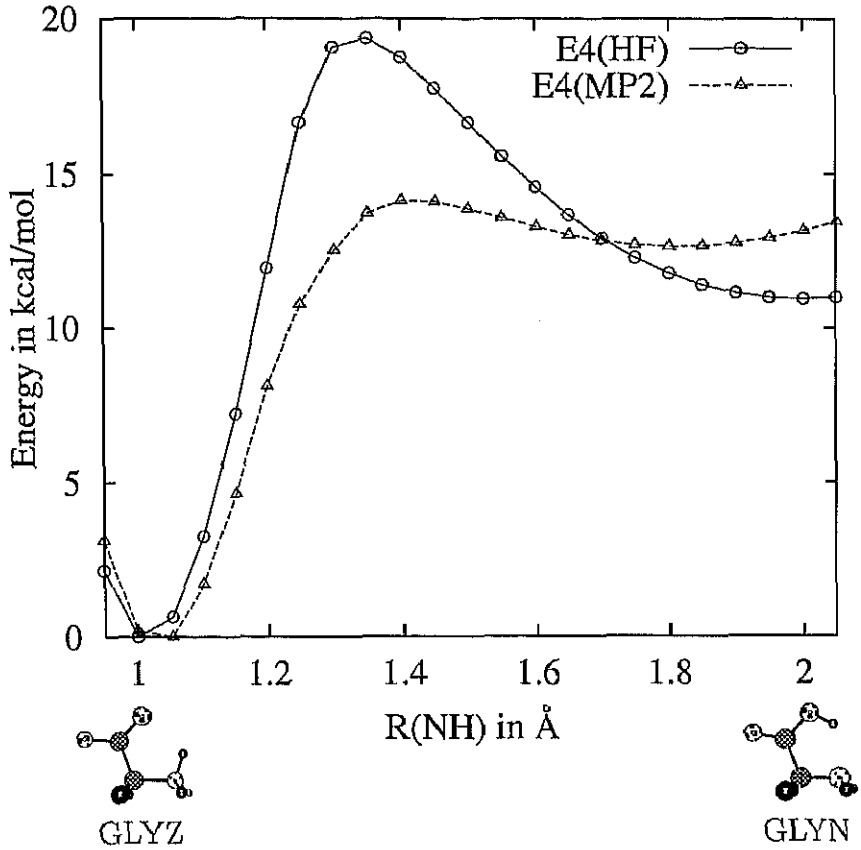


Figure 4.5: *Ab initio* GB energy profiles for the proton transfer reaction of glycine: E4(HF)(solid line, circle) and E4(MP2)(broken line, triangle) are the E4 energies of Figure 4.3 and 4.4, respectively. These energies were calculated with the structure optimized in aqueous solution.

Table 4.5: Relative energies (in kcal/mol) for GLYZ, GLYN and the transition state (TS) between GLYZ and GLYN

	GLYZ	TS	GLYN	Solvent model	Level of calculation
This study	0.00	19.46	11.01	*ab initio* GB	HF/6-31++G**
This study	0.00	13.03	13.02	*ab initio* GB	MP2/6-31++G**[a]
This study	0.00	—	-29.94	gas phase	HF/6-31++G**[a]
This study	0.00	—	-24.23	gas phase	MP2/6-31++G**[a]
Ref78	0.0	0.6	-23.4	gas phase	HF/6-31+G*
Ref78	0.0	-1.2	-15.8	gas phase	MP4/6-31+G*[b]
Ref78	0.0	-1.5[c]	-23.5[c]	gas phase	HF/6-31+G*
Ref78	0.0	-3.3[c]	-15.9[c]	gas phase	MP4/6-31+G*[b]
Ref86	0.0	-1.0	-17.0	gas phase	MP2/DZP++[d]
Ref86	0.0	-1.9[c]	-16.4[c]	gas phase	MP2/DZP++[d]
Ref86	0.0	3.6	-10.4	discrete model[e]	MP2/DZP++[f]
Ref86	0.0	1.0[c]	-11.1[c]	discrete model[e]	MP2/DZP++[f]
Ref86	0.0	6.3	-3.8	discrete model[g]	MP2/DZP++[f]
Ref86	0.0	1.3[c]	-5.0[c]	discrete model[g]	MP2/DZP++[f]
Ref87	0.00[h]	—	-2.16[i]	continuum model[j,k]	HF/6-31G**
Ref87	0.00[h]	—	-0.02[i]	continuum model[j,k]	MP2/6-31G**
Ref87	0.00[h]	—	-1.19[i]	continuum model[j,k]	B3LYP/6-31G**
Ref87	0.00[h]	—	-1.54[i]	continuum model[j,k]	BH&HLYP/6-31G**
Ref87	0.00[h]	—	2.63[i]	continuum model[j]	HF/6-31G**
Ref87	0.00[h]	—	4.02[i]	continuum model[j]	MP2/6-31G**
Ref87	0.00[h]	—	3.09[i]	continuum model[j]	B3LYP/6-31G**
Ref87	0.00[h]	—	2.95[i]	continuum model[j]	BH&HLYP/6-31G**
Ref77	0.00	8.75	-2.29	continuum model[l]	HF/6-31+G**
Ref77	0.00	3.54	1.15	continuum model[l]	MP2/6-31+G**
Ref77	0.00	12.64	-16.40	discrete model[e]	HF/6-31+G**
Ref77	—	15.56	0.00	discrete model[e]	MP2/6-31+G**
Expt ΔH	0.00	-0.2	10.3	aqueous solution	
Expt ΔG	0.00	14.3	7.2	aqueous solution	

[a] The structure was optimized by using *ab initio* GB HF/6-31++G** calculation. [b] The structure was optimized by using *ab initio* HF/6-31+G* calculation. [c] These value included the vibrational zero-point energy. [d] The structure was optimized by using *ab initio* HF/6-31G* calculation. [e] These calculations were carried out by adding a single water molecule. [f] The structure was optimized by using *ab initio* HF/DZP calculation. [g] These calculations were carried out by adding two water molecules. [h] The conformation 2_Z was used for GLYZ. [i] The conformation 2_N was used for GLYN. [j] The generalized conductorlike screening model (GCOSMO)[87] was used. [k] These calculations carried out by using the gas phase geometries. [l] The elipsoidal cavity and multipole expansion model of Rivail *et al.*[81,82] was used.

4.2 Determination of EPPF

The EFFP between glycine and water was determined as the function of the reaction coordinate defined by R_{N-H} by using *ab initio* GB (ϵ=79) / 6-31++G** method. The structure of glycine monomer was optimized at each R_{N-H}(0.95-2.05Å, 23 structures). The form of EPPF employed consists of the Lenard-Jones(12-6), hydrogen-bond(12-10) and Coulomb terms as follows,

$$E_{\alpha\beta}^{int} = \sum_{i\in\alpha}\sum_{j\in\beta}\left(\frac{A_{ij}^{LJ}(R_{NH})}{R_{ij}^{12}} - \frac{C_{ij}^{LJ}(R_{NH})}{R_{ij}^{6}} + \frac{A_{ij}^{HB}(R_{NH})}{R_{ij}^{12}} - \frac{C_{ij}^{HB}(R_{NH})}{R_{ij}^{10}} + \frac{Q_i(R_{NH})Q_j(R_{NH})e^2}{R_{ij}}\right) \quad (4.1)$$

where A_{ij}, C_{ij} and Q_{ij} means the parameters of repulsive, attractive and coulomb terms between i'th and j'th sites, respectively, and LJ and HB denote "Lenard-Jones" and "Hydrogen-Bond", respectively. These parameters depend on R_{N-H}.

These parameters were determined by the following procedure.

(1) The coulomb parameters of glycine, the charges on atoms, were determined so that the parameters reproduce well the electrostatic potential (ESP) around glycine. These coulomb parameters determined at each R_{N-H}value and were representd as the function of R_{N-H}.

(2) The parameters in the LJ and HB terms were determined so that the energies of EPPF in Eqn. 4.1 reproduce the effective interaction energies. For each R_{N-H}value, the 17 types of about 110 geometries of glycine-water complexes were considered. In this procedure, the complexes whose effective interaction energies were larger than 30 kcal/mol were not employed in the parameter fitting.

In the determination of the coulomb parameters, ESPs were calculated at 5000 points/atom in the shell whose radius is 3-4 times of van der Waals radius.[88] The

coulomb parameters determined are shown in Figure 4.6, in which points represent the charges fitted to reproduce the ESPs. The root mean square error (RMS-error) for each structure of glycine is ca. 0.002. Remarkable charges are observed in the changes near R_{N-H}= 1.3 Å, which comes from the drastic change of the glycine structure. These charges are fitted to the following functions,

$$\begin{aligned} Q_i(R) &= \left(a_1R^2 + a_2R + a_3\right) \times T(R) \\ &\quad + \left(a_4R^2 + a_5R + a_6\right) \times (1 - T(R)) \end{aligned} \tag{4.2}$$

where

$$T(R) = \frac{tanh(a_7(R - a_8)) - 1}{2} \tag{4.3}$$

where a_i is an i'th parameter listed in Table 4.6, and $T(R)$ is a switching function to reproduce the correct behavior of the parameters near R_{N-H}= 1.3 Å, and a_7 and a_8 represent the smoothness and position of switching function for R_{N-H}, respectively, and the values of a_8 are certainly 1.2-1.3 in Table 4.6.

The 17 types of glycine-water complexes were used for determination of the PF parameters in LJ and HB terms, and there are 5-8 geometries for each type, which have various interaction distance between glycine and water. In Figures 4.7, 4.8, and 4.9, the all types of complexes for GLYZ (R_{N-H}= 1.05 Å), GLYTS (R_{N-H}= 1.35 Å), and GLYN (R_{N-H}= 1.95 Å), respectively, are shown. The complex-types are divided into 4 groups.

(i) A water molecule interacting with the CH_2 group of glycine.

(ii) A water molecule interacting with the CO_2^- group of glycine.

(iii) A water molecule interacting with the NH_3^+ group of glycine.

(iv) A water molecule interacting with the OH group of GLYN. For the GLYZ, this corresponds to a water simultaneously interacting with CO_2^- and NH_3^+.

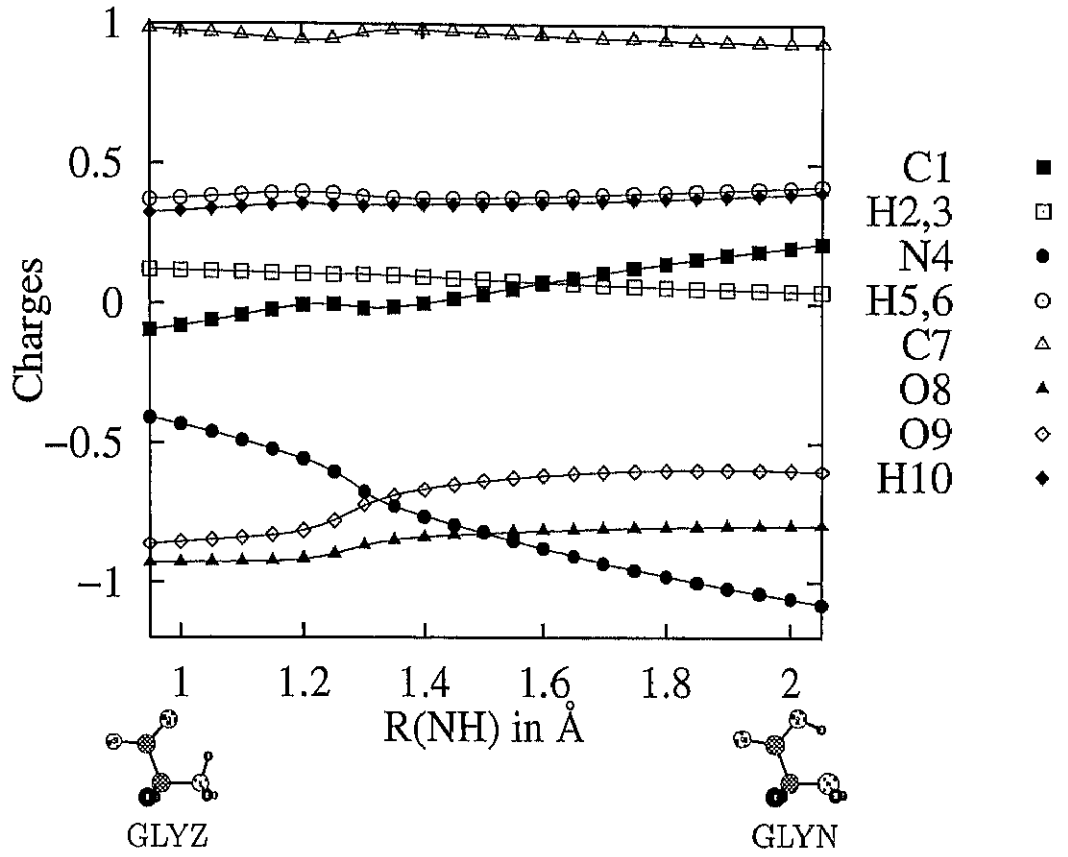


Figure 4.6: ESP charges of glycine and the fitted curves: Solid lines are the fitted functions listed in Table 4.6.

Table 4.6: Potential parameters for coulomb terms

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
Q_{C1}	-0.1232	0.7678	-0.8397	0.5559	-0.7865	0.1505	13.29	1.267
$Q_{H2,3}$	0.044	-0.2311	0.3322	-0.08036	0.0999	0.0983	13.2	1.276
Q_{N4}	0.2011	-1.1804	0.4955	-0.4877	0.4484	-0.3936	23.72	1.289
$Q_{H5,6}$	0.02516	-0.006996	0.3299	0.4064	-0.7208	0.6906	9.86	1.248
Q_{C7}	0.05339	-0.2571	1.2347	-0.2601	0.3747	0.8612	17.46	1.274
Q_{O8}	-0.09447	0.385	-1.1897	-0.05204	0.1264	-1.0024	14.43	1.264
Q_{O9}	-0.2921	1.1002	-1.6303	-0.1306	0.4035	-1.1264	13.75	1.261
Q_{H10}	0.06818	-0.1618	0.4452	0.01976	0.0902	0.2206	34.63	1.213

The effective interaction energies are listed in Table 4.7. These energies were calculated by using the fixed orientationfor each type of interaction and by changing the distances between glycine and water molecules; no geometry optimization was carried out. Since the parameters for the short-range interactions, LJ and HB terms, are determined, the calculation was restricted for the structure with small distances, 1.4-3.1 Å. In some cases there was no energy minimum distance in the calculated range of the interaction distance; they are indicated by parenthesis in Table 4.7.

In the case of effective interaction energies of GLYZ-water complexes, the water-CH_2 interaction energies, -4.6 – -4.7 kcal/mol, are considerably smaller than other ionic group-water interactions, -12.0 – -12.6. For the GLYN-water complexes, the water-CH_2 interaction energies, -2.4 – -3.0 kcal/mol, are comparable with the ionic group-water ones, -3.0 – -6.2 kcal/mol. The interaction energies of Jensen and Gordon,[86] -18.3 – -18.4 and -7.4 – -11.7 kcal/mol for GLYZ and GLYN, respectively, are larger than those in this study, because they took into account the potimization not only for the position of water but also for the glycine structure.

Each parameter of EPPF was expressed as the function of R_{N-H}in Eqn. 4.2, and listed in Table 4.8. In Figure 4.10, the interaction energies obtained by the determined EPPF for glycine-water are compared with the effective interaction energies used to determine the EPPF. The RMS-error of these interaction energies was 1.52 kcal/mol, suggesting that the EPPF reproduce well the effective interaction energies obtained by *ab initio* GB calculations.

The RMS-errors were plotted for the $R_{N-H}$of glycine in Figure 4.11. The RMS-error becomes small as R_{N-H}becames long. The attractive coulomb interactions of GLYZ-water are stronger than that of GLYN-water, corresponding that the repulsive LJ and HB interactions become larger in GLYZ-water interaction energies. The main source of the error is discrepancy between repulsive profiles of EPPF and MO calculation in the very short range, which represents as a deviation in the range of >

Table 4.7: Effective pair interaction energies (in kcal/mol) of glycine-water complexes[a]

	GLYZ (R_{N-H}= 1.05 Å)	GLYTS (R_{N-H}= 1.35 Å)	GLYN (R_{N-H}= 1.95 Å)
A water molecule interacting with CH_2 group of glycine			
1a	-4.6	-3.0	-2.4
1b	-4.7	-3.7	-3.0
1c	(-4.2)	(0.4)	(2.6)
A water molecule interacting with CO_2^- group of glycine			
2a	-12.0	-7.9	-6.1
2b	-10.2	-7.4	-6.2
2c	-7.7	-5.0	-4.1
2d	-6.5	-3.9	-2.9
2e	-9.1	-6.6	-5.2
A water molecule interacting with NH_3^+ group of glycine			
3a	-9.2	(-3.6)	(-0.9)
3b	-12.1	-6.8	-4.8
3c	-9.8	-3.8	-1.4
3d	-9.5	-4.1	(-1.9)
A water molecule interacting with OH group of GLYN or a water simultaneously interacting with CO_2^- and NH_3^+ of GLYZ			
4a	-12.6	(1.3)	(-1.7)
4b	-7.6	-4.9	-3.0
4c	-7.3	(0.7)	(-0.1)
4d	(1.1)	(5.5)	(2.3)
4e	-10.4	-4.9	-2.1

[a] The energies in parentheses are not optimum energies, as described in text.

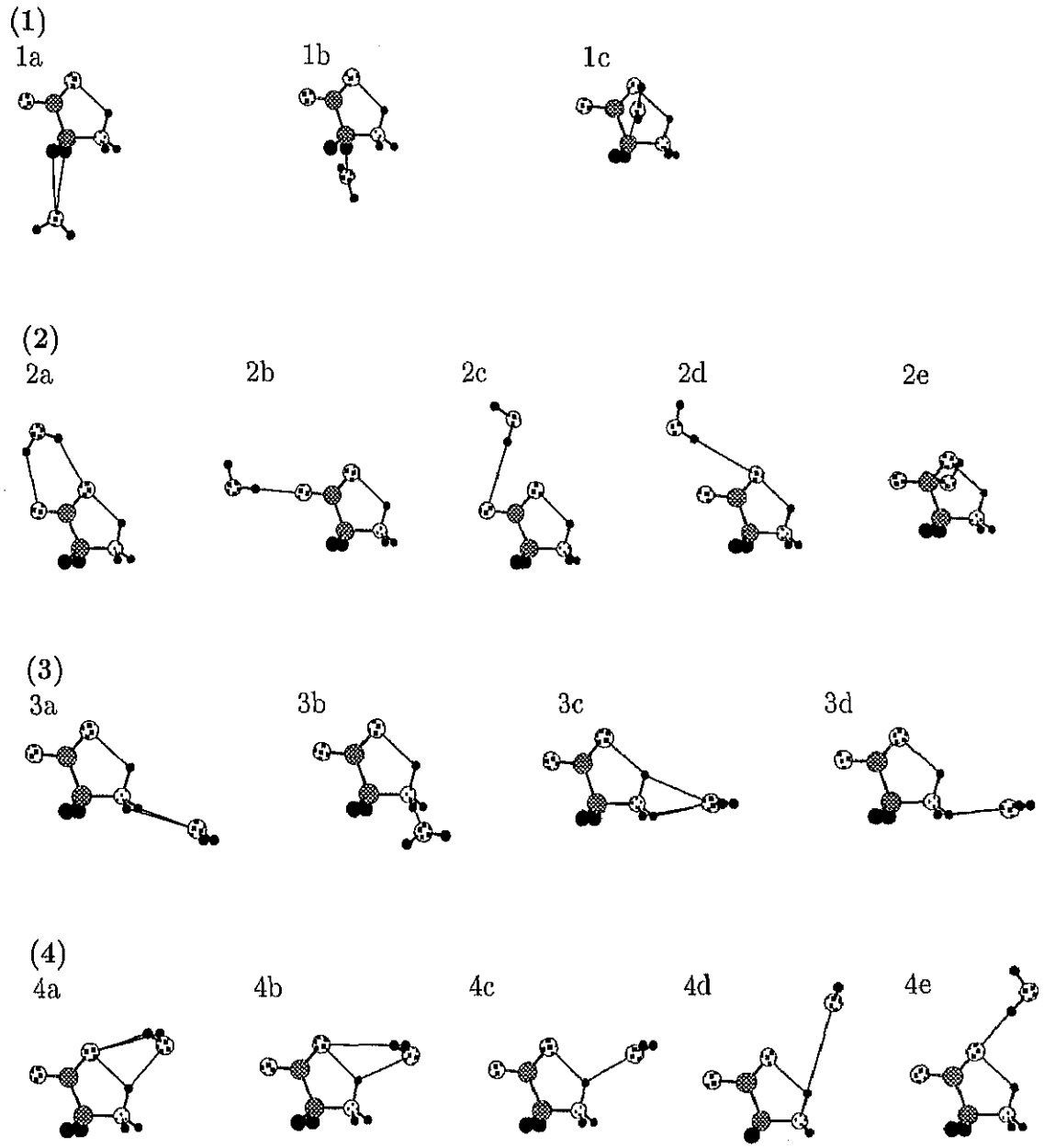


Figure 4.7: The GLYZ(R_{N-H}= 1.05 Å)–water complexes. (**1**)A water molecule interacting with CH_2 group of glycine, (**2**)a water molecule interacting with CO_2^- group of glycine, (**3**)a water molecule interacting with NH_3^+ group of glycine and (4)a water simultaneously interacting with CO_2^- and NH_3^+.

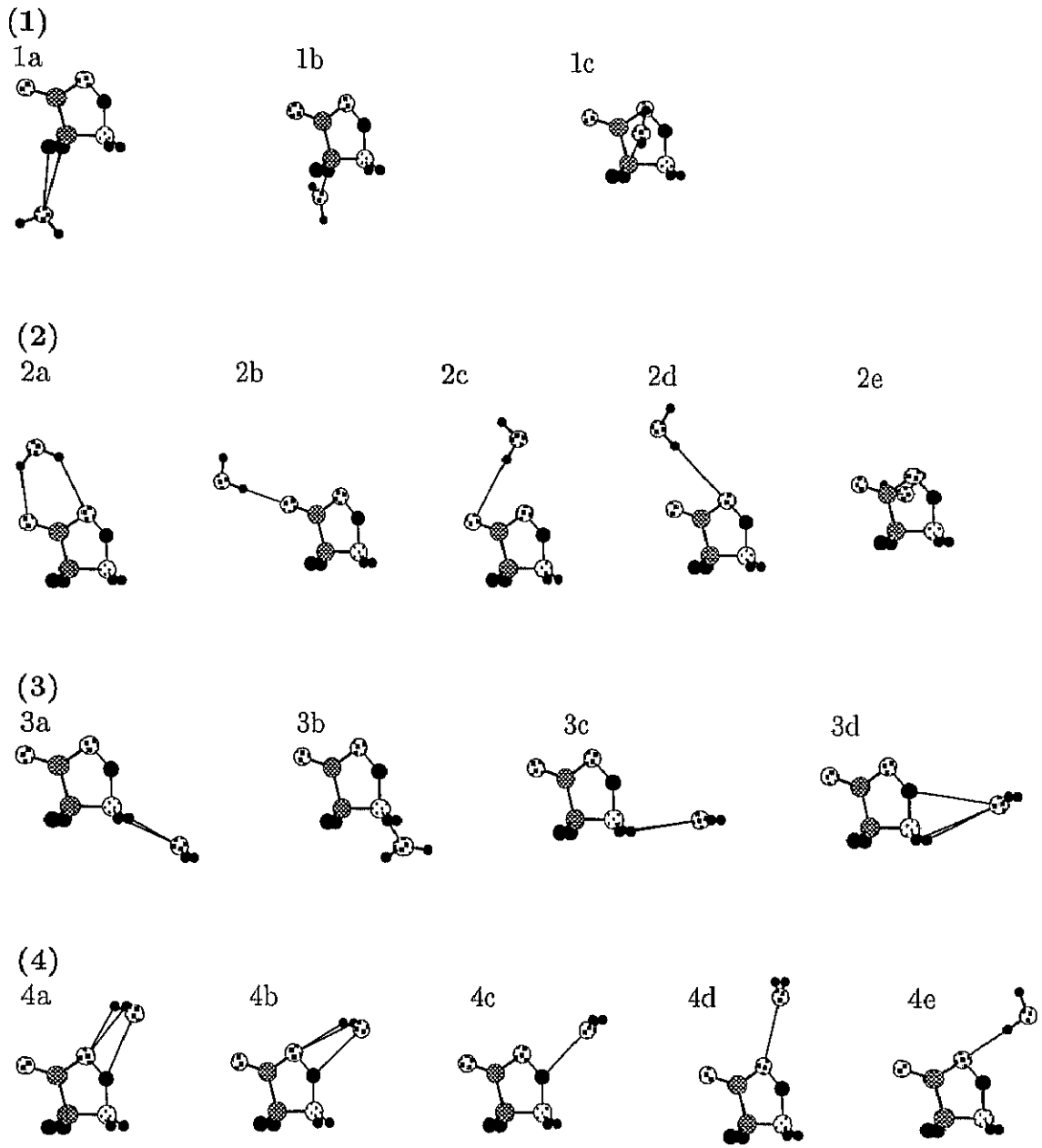


Figure 4.8: The GLYTS(R_{N-H}= 1.35 Å)–water complexes. (1)A water molecule interacting with CH_2 group of glycine, (2)a water molecule interacting with CO_2^- group of glycine, (3)a water molecule interacting with NH_3^+ group of glycine and (4)a water molecule interacting with OH group of glycine.

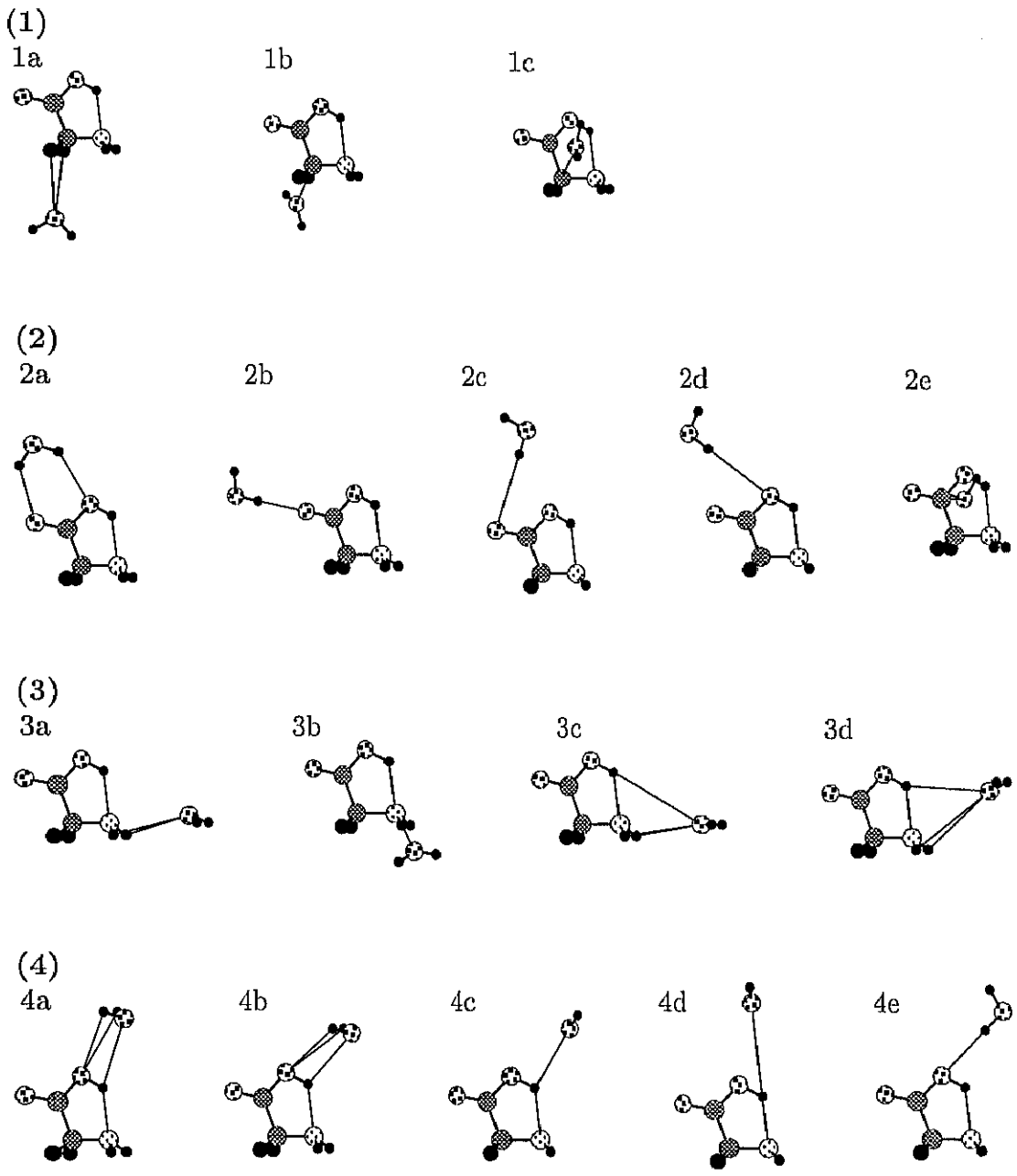


Figure 4.9: The GLYN(R_{N-H}= 1.95 Å)–water complexes. (1)A water molecule interacting with CH_2 group of glycine, (2)a water molecule interacting with CO_2^- group of glycine, (3)a water molecule interacting with NH_3^+ group of glycine and (4)a water molecule interacting with OH group of glycine.

Table 4.8: Potential parameters for Lenard-Jones and hydrogen-bond terms

Site	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
Lenerd-Jones terms								
A_{C1-Ow}	3643000	-7524000	3979000	-509600	1887000	-1491000	-9.85	1.345
C_{C1-Ow}	-22240	45980	-22010	2721	-10650	10980	-10.01	1.348
A_{N4-Ow}	124300	-91820	65230	-319300	1133000	-863200	-14.16	1.329
C_{N4-Ow}	698.5	-2385	2762	2598	-9680	9540	-12.61	1.317
A_{C7-Ow}	-746700	1548000	-626300	116300	-493800	618500	-8.51	1.353
C_{C7-Ow}	16170	-34030	16440	-2429	10300	-10900	-8.70	1.353
A_{O8-Ow}	607500	-1292000	657000	-75900	319700	-320800	-7.72	1.332
C_{O8-Ow}	-10830	22600	-9980	1573	-6453	7723	-8.69	1.341
A_{O9-Ow}	-949500	2139000	-1058000	197700	-670400	660400	-103.01	1.331
C_{O9-Ow}	8309	-16510	7534	-1602	5437	-4369	-14.76	1.299
Hydrogen-bond terms								
$A_{H2,3-Ow}$	3286	-6852	1181	-421	1155	-3100	-9.74	1.306
$C_{H2,3-Ow}$	-543	1134	1174	102.6	-159	1792	-54.99	1.258
$A_{H5,6-Ow}$	2775	-7089	3276	-9178	28050	-23240	-15.51	1.372
$C_{H5,6-Ow}$	1767	-2965	2183	3351	-9696	8558	-12.71	1.443
A_{O8-Hw}	-22890	44160	-24160	3690	-14840	10770	-9.63	1.343
C_{O8-Hw}	12710	-24730	14010	-2003	8017	-5317	-9.60	1.343
A_{O9-Hw}	22250	-30860	410.0	-10890	39090	-40110	-11.41	1.336
C_{O9-Hw}	-10797	14889	645.0	5396	-19343	20432	-11.38	1.335
A_{H10-Ow}	-74450	155700	-83920	77510	-276500	238100	-9.07	1.362
C_{H10-Ow}	35630	-75480	41790	-30820	110800	-95050	-8.90	1.364

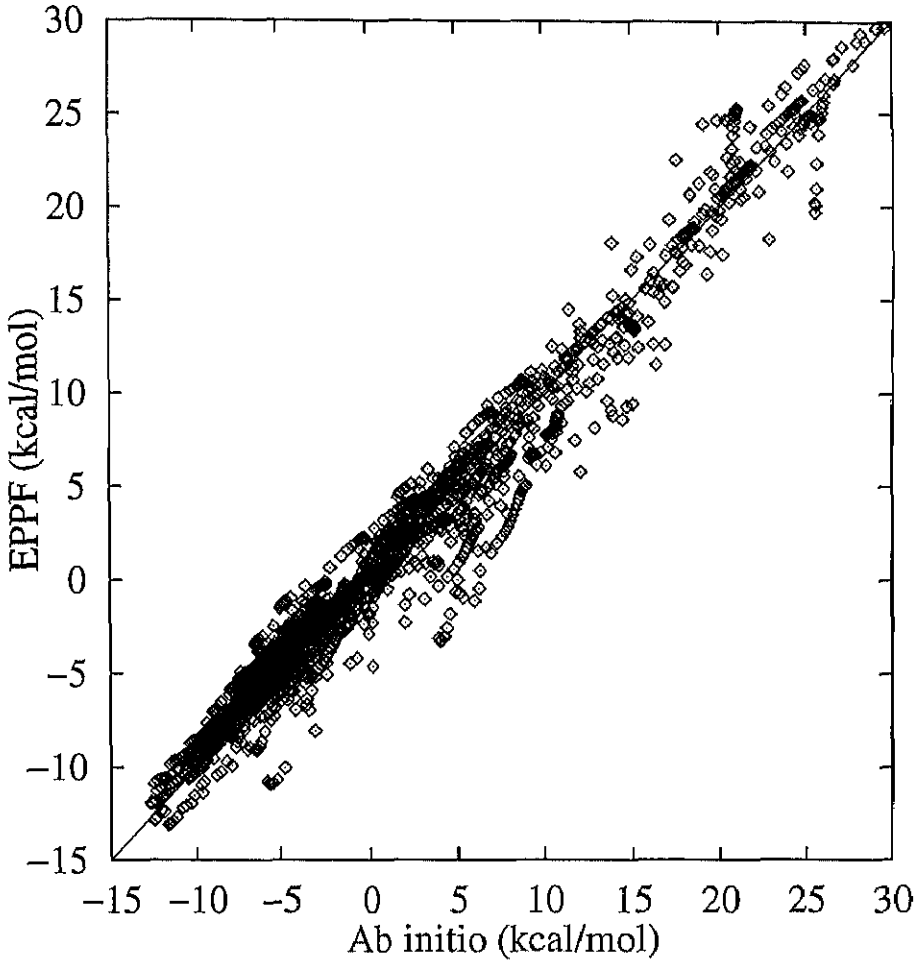


Figure 4.10: The comparison of the interaction energies obtained with the *ab initio* GB calculation and the EPPF determined for the glycine–H_2O configurations.

0 kcal/mol in Figure 3.5. Then the procedure in a least-squares fitting decreases the deviations for unstable complexes but increases the stable ones. The profile of the RMS-error near R_{N-H}= 1.2 Å reflects the drastic change, coming from the drastic change of glycine structure along the reaction coordinate defined by R_{N-H}.

In order to obtain the more accurate fitted EPPF, several improvements can be considered.

(i) The more flexible the functional form for the repulsive term is adopted. For example, an exponential form, $a\ exp(-br)$, has two parameters, a and b, and more flexible than the LJ repulsive term.

(ii) The reaction coordinate is selected more carefully. For the smooth variation of the structure of glycine, an intrinsic reaction coordinate (IRC) or a reaction coordinate define by the difference between $R_{N-H}$and R_{O-H}, $s = R_{N-H} - R_{O-H}$, is available.

The determined parameters appeared in Eqn. 4.1 for GLYZ(R_{N-H}= 1.011 Å), GLYTS(R_{N-H}= 1.335 Å), and GLYN(R_{N-H}= 1.995 Å) are listed in Table 4.9. The repulsive parameters for the O8-Ow LJ term and all HB terms have negative values, and these terms behave as attractive one. But correct correlation between energies of the EPPF and the MO calculation was obtained in the range of -20 – +30 kcal/mol in Figure 4.10, so the role of the repulsive interaction is described correctly by other terms.

For a PF for the water-water interaction, the model determined in Chapter 3 was used because the dipole moment of water, 2.389 D, calculated by using *ab initio* GB HF/6-31++G** level was similar to that for the PF model, 2.424 D.

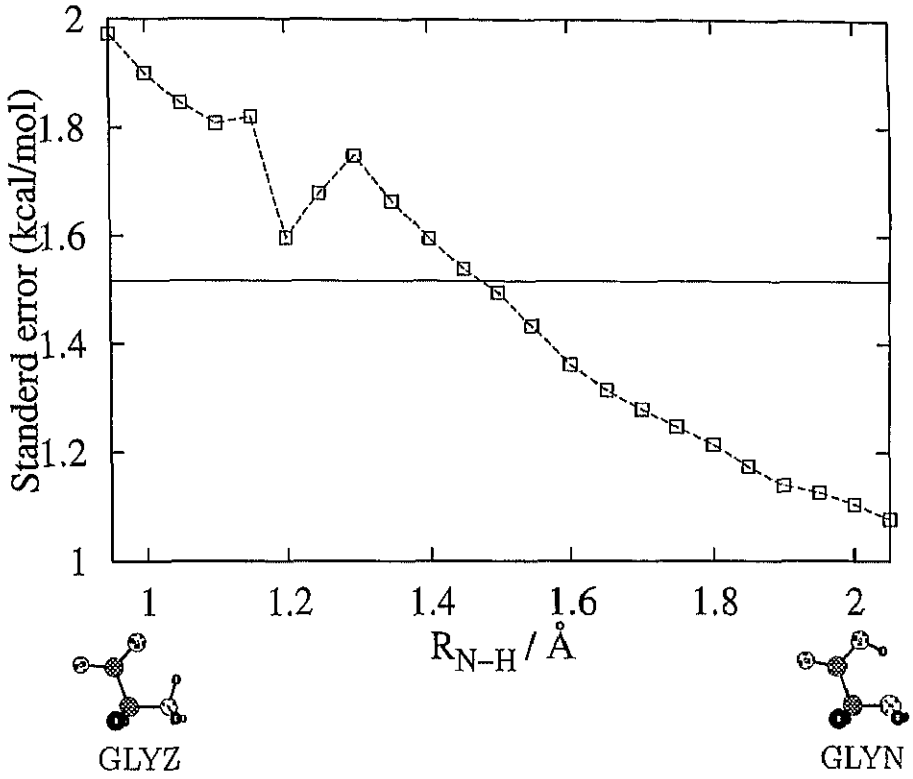


Figure 4.11: The root mean square error (RMS-error) of the interaction energies obtained with *ab initio* GB calculation and the EPPF determined for the GLY–H_2O configurations. The solid line of 1.52 kcal/mol is the RMS-error for all configurations. The broken line and squares are one for configurations of the GLY–H_2O which have corresponding R_{N-H} value.

Table 4.9: Determined potential function for the structures of GLYZ, GLYTS, and GLYN

	GLYZ		GLYTS		GLYN	
Lenerd-Jones terms[a]						
	A	C	A	C	A	C
C1-Ow	96140	1738	289400	546.8	245400	565.1
N4-Ow	99460	1065	118700	1083	126300	571.3
C7-Ow	176000	-1437	133800	-721.9	96130	-25.94
O8-Ow	-28920	1802	-7969	1375	14860	1108
O9-Ow	134900	-667.5	114300	102.7	109700	102.6
Hydrogen-bond terms[b]						
	A	C	A	C	A	C
H2,3-Ow	-2388	1766	-2236	1762	-2472	1882
H5,6-Ow	-1055	990.9	-1464	1386	-3813	2552
O8-Hw	-2909	2001	-4369	2803	-4149	2704
O9-Hw	-8047	4663	-4200	2746	-5455	3321
H10-Ow	-2582	1867	-2738	2080	-4965	3389
Coulomb terms						
		Q		Q		Q
C1		-0.0766		-0.0166		0.2017
H2,3		0.1172		0.0997		0.0463
N4		-0.4388		-0.7161		-1.0590
H5,6		0.3770		0.3787		0.4161
C7		0.9742		0.9772		0.9343
O8		-0.9278		-0.8535		-0.7976
O9		-0.8519		-0.6981		-0.5980
H10		0.3320		0.3507		0.3938

[a] A in $(\text{Å}\cdot\text{kcal/mol})^{12}$, C in $(\text{Å}\cdot\text{kcal/mol})^{6}$. [b] A in $(\text{Å}\cdot\text{kcal/mol})^{12}$, C in $(\text{Å}\cdot\text{kcal/mol})^{10}$.

4.3 Monte Carlo simulation

The free energy differences among GLYZ, GLYTS, and GLYN were calculated by using SPT in MC simulation with EPPF determined above. The MC simulation was carried out under the NPT ensemble (1 glycine and 511 water molecules, 1 atom, and 298 K). The SPT step of ΔR_{N-H}= 0.0125 Å was used in the range of $1.25 \leq R_{N-H} \leq 1.325$ Å, and the SPT step of ΔR_{N-H}= 0.025 Å was used for other range. The calculation of an interaction energy was linearly truncated from 11.4 to 12.0 Å, and long-range coulomb correlation for the result of SPT simulation was considered by using the Onsager model, which is represented by a dipole moment in spherical cavity with the radius of 11.7 Å.

Table 4.10 shows the effective interaction energies of three solution states, GLYZ, GLYN and GLYTS and the SPT free energy differences among them. The difference in the total effective interaction energy between GLYZ and GLYN, 0.097 kcal/mol $\times$ 512 molecules = 49.66 kcal/mol can be divided in two interactions, the solvent-solute and solvent-solvent interaction energies. The interaction energy of the solute molecule with the water molecules was -77.90 kcal/mol for GLYZ and -18.42 kcal/mol for GLYN; the difference is 59.48 kcal/mol. The corresponding difference of the interaction energy among the 511 water molecules is -0.02 kcal/mol $\times$ 511 = -10.22 kcal/mol.

The variation of the volume was small and its effect to the energy is expected to be small; the difference of $P\Delta V$ for GLYTS and GLYZ is 0.004 kcal/mol.

The effective interaction energy for GLYZ is -77.90 kcal/mol. Nagaoka *et al.*obtained the corresponding interaction energy of -66.83 kcal/mol by using the PF determined by *ab initio* calculation in the gas phase; their smaller value may be caused by the small dipole moment derived from the lack of the electronic polarization in the solute molecule.

The SPT free energy differences among the three states were calculated for the

forward and backward samplings. The difference between the forward and backward samplings was 0.27 kcal/mol suggesting that the SPT step employed is small enough to obtain a reliable energy variation.

The free energy diagram for the proton transfer between $\mathbf{1}_Z$ and $\mathbf{1}_N$ is shown in Figure 4.12. The energies cited in Figure 4.12 were calculated by using Eqn. 2.13b with the mean SPT free energy change and E_0^{sol} value calculated at the MP2/6-31++G** level of calculation; they were corrected by using the Onsager model.

The relative energy of $\mathbf{1}_Z$ and $\mathbf{1}_N$ is 0.2 kcal/mol and the activation free energies from $\mathbf{1}_Z$ to $\mathbf{1}_N$ is 5.1 kcal/mol. These values suggest that the tautomerization caused by proton transfer occurs easily between these conformers in which the transferring hydrogen directs to the accepting functional group.

With respect to the conformation of GLYZ, $\mathbf{2}_Z$ is more stable than $\mathbf{1}_Z$ by 1.4 kcal/mol as discussed in section 4.1. According to the *ab initio* calculation for the water cluster of GLYZ, GLYZ interacts strongly with a water molecule through the hydrogen-bonding at the both functional groups, NH_3^+ and CO_2^-,[86] but this glycine-water complex is destabilized in $\mathbf{1}_Z$ because O-H distance of $\mathbf{1}_Z$ is too small to hydrate appropriately with the water molecule.[76]

The conformation of GLYN in the gas phase was examined in detail,[78,84,85] and $\mathbf{2}_N$ was predicted to be the most stable conformation, as described in section 4.1.

The free energy diagram in which the conformational change is taken into account is shown in Figure 4.13. Although there are the various conformers for GLYN, only $\mathbf{2}_N$ is considered, because $\mathbf{2}_N$ is the most stable conformer in the gas phase[73–75] and must be the important conformer in aqueous solution. In this free energy diagram, the GLYZ→GLYN isomerization starts from the conformational change from $\mathbf{2}_Z$ to $\mathbf{1}_Z$, which is followed by an essential proton transfer.

Sheinblatt and Gutowsky[66] measured the mean lifetime between exchanges of protons in the NH_3^+ of GLYZ by using various NMR relaxation techniques, and

Table 4.10: The effective interaction energies, cell lengths and SPT free energy differences of MC simulation for GLYZ, GLYN, GLYTS

R_{N-H}(Å)	GLYZ 1.011	GLYTS 1.335	GLYN 1.995
The effective interaction energy (kcal/mol)			
Total	-12.22	-12.20	-12.12
Glycine	-77.90	-33.28	-18.42
Water	-12.09	-12.16	-12.11
Cell length (Å)	25.00	24.57	25.02
SPT free energy difference (kcal/mol)			
Mean	0.00	28.38	37.17
Forward	0.00	28.40	37.31
Backward	0.00	28.36	37.04

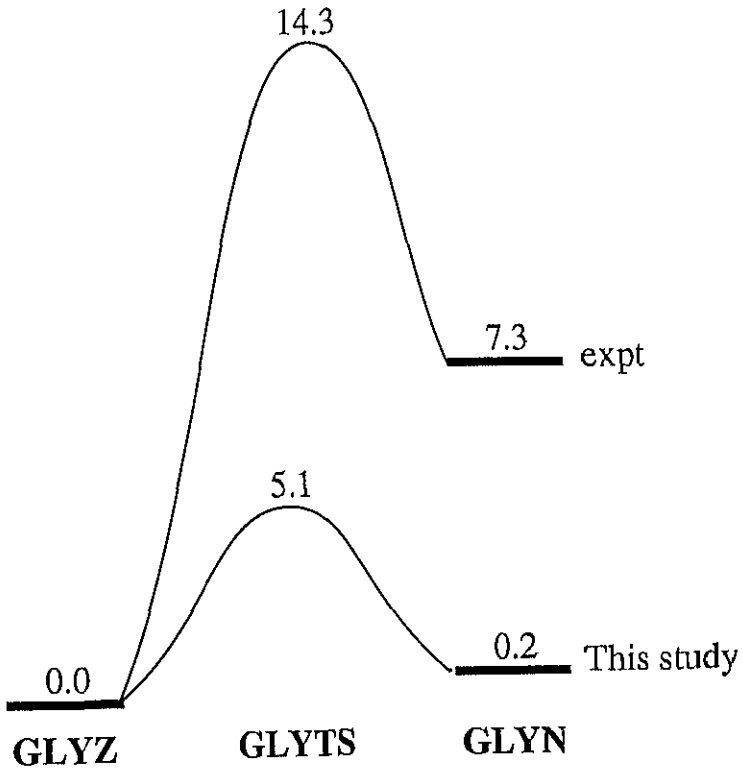


Figure 4.12: The free energy diagram (kcal/mol) for the intramolecular proton transfer from $\mathbf{1}_Z$ to $\mathbf{1}_N$: The experimental values are taken from refs. 66–68.

derived the rate constant for the intramolecular proton transfer in glycine, $\mathbf{1}_Z \rightarrow \mathbf{1}_N$. As described by Sheinblatt and Gutowsky,[66] the reverse process of an original proton with unchanged nuclear spin state, $\mathbf{1}_N \rightarrow \mathbf{1}_Z$, also may occur easily, but this process has no net effect upon the NMR experiment. Thus the rate controlling step may be the conformational change of $\mathbf{2}_N \rightarrow \mathbf{1}_N$, or the proton transfer from $\mathbf{2}_N$ to a water molecule. The conformational change is governed by theby the entropic effect,[73] supporting the experimental fact which the activation free energy and the entropic contribution, $-T\Delta S$, are 14.3 and 14.5 kcal/mol, respectively.

The free energy difference between GLYZ and GLYN, 4.3 kcal/mol, appears to be small. The underestimation of the energy difference may be partially attributed to the lack of the effect of hydrated water molecules, and to the EPPF with RMS-error of 1.52 kcal/mol.

For the lack of the hydrated water molecules, it must be stressed that the crucial contribution of solvent effects to GLYZ has been introduced by the interaction with continuum solvent. But the consideration of specific interactions of water molecules improves the electronic state of the solute molecule, as in the researches for mono-atomic cation.[15–25]

With respect to the second point, the determined EPPF well reproduces the effective interaction energies calculated by *ab initio* GB method, although there are some ways to improve EPPF as described in section 4.2.

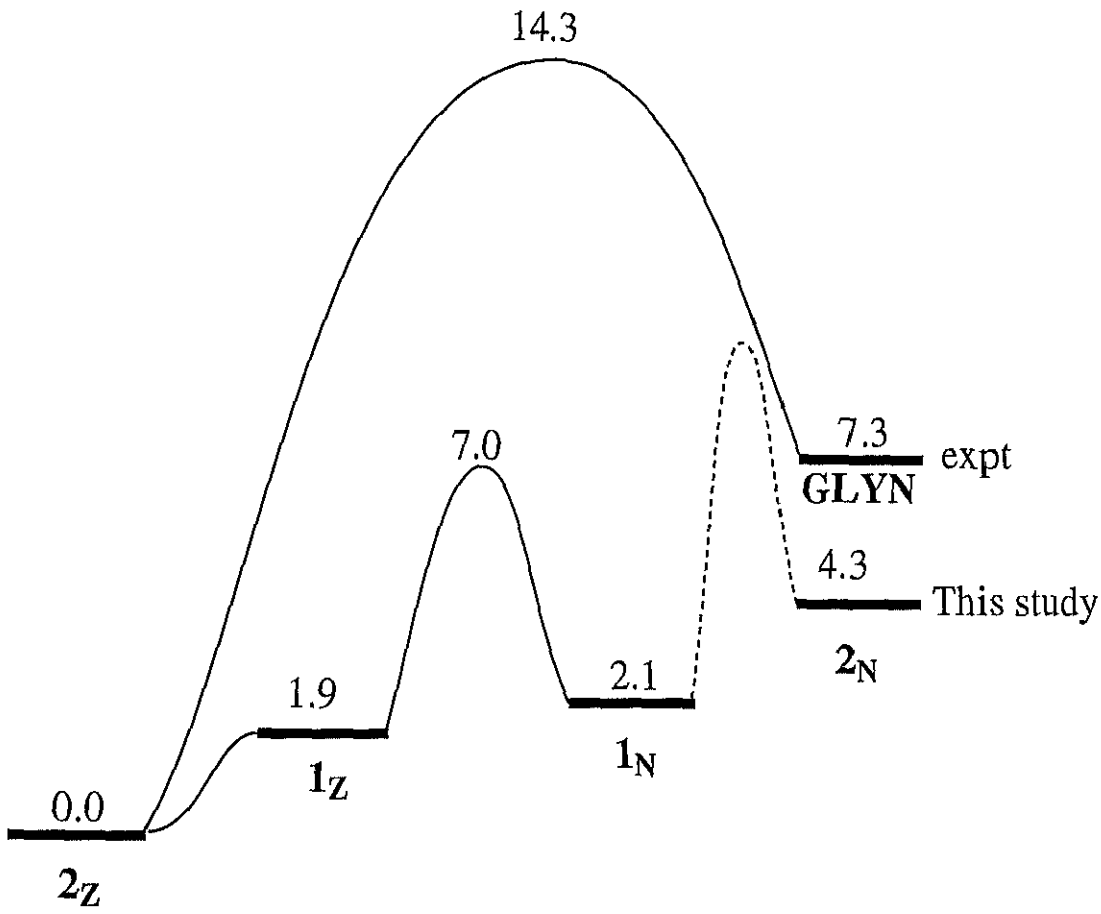


Figure 4.13: The free energy diagram (kcal/mol) for the intramolecular proton transfer from $\mathbf{2}_Z$ to $\mathbf{2}_N$: The experimental values are taken from refs. 66-68.