# CALCULATION OF PROTEIN–LIGAND BINDING FREE ENERGY USING SMOOTH REACTION PATH GENERATION (SRPG) METHOD; A COMPARISON OF THE EXPLICIT WATER MODEL, GB/SA MODEL AND DOCKING SCORE FUNCTION

DAISUKE MITOMO[1]
d-mitomo@aist.go.jp

YOSHIFUMI FUKUNISHI[2,3]
y-fukunishi@aist.go.jp

JUNICHI HIGO[4]
higo@protein.osaka-u.ac.jp

HARUKI NAKAMURA[2,5]
harukin@protein.osaka-u.ac.jp

[1] *Japan Biological Informatics Consortium (JBIC), 2-41-6, Aomi, Koto-ku, Tokyo 135-0064, Japan*
[2] *Biomedicinal Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6, Aomi, Koto-ku, Tokyo 135-0064, Japan*
[3] *Pharmaceutical Innovation Value Chain, BioGrid Center Kansai, 1-4-2 Shinsenri-Higashimachi, Toyonaka, Osaka 560-0082, Japan*
[4] *The Center for Advanced Medical Engineering and Informatics, Osaka University, Open Laboratories for Advanced Bioscience and Biotechnology, 6-2-3, Furuedai, Suita, Osaka 565-0874, Japan*
[5] *Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan*

We compared the protein–ligand binding free energies (ΔG) obtained by the explicit water model, the MM-GB/SA (molecular-mechanics generalized Born surface area) model, and the docking scoring function. The free energies by the explicit water model and the MM-GB/SA model were calculated by the previously developed Smooth Reaction Path Generation (SRPG) method. In the SRPG method, a smooth reaction path was generated by linking two coordinates, one a bound state and the other an unbound state. The free energy surface along the path was calculated by a molecular dynamics (MD) simulation, and the binding free energy was estimated from the free energy surface. We applied these methods to the streptavidin-and-biotin system. The ΔG value by the explicit water model was close to the experimental value. The ΔG value by the MM-GB/SA model was overestimated and that by the scoring function was underestimated. The free energy surface by the explicit water model was close to that by the GB/SA model around the bound state (distances of < 6 Å), but the discrepancy appears at distances of > 6 Å. Thus, the difference in long-range Coulomb interaction should cause the error in ΔG. The scoring function cannot take into account the entropy change of the protein. Thus, the error of ΔG could depend on the target protein.

*Keywords*: protein–ligand binding free energy; MD; MM-GB/SA; docking score

## 1. Introduction

Various types of solvent models have been used in molecular dynamics (MD) simulations. There are two major types of solvent model. One is the explicit solvent model, which treats water molecules explicitly. When using the explicit model, we can calculate solute-solvent interactions precisely, but the simulation is time-consuming. The

other type comprises implicit solvent models. Among the implicit models, the generalized Born (GB)/ surface area (SA) model is widely used [22, 27, 30]. An alternative way to estimate the ΔG value is a docking scoring function. There have been many kinds of protein–ligand docking programs and algorithms reported [2, 3, 9, 13, 17, 19, 23]. The calculations of ΔG by the scoring functions are much faster than those by the explicit water model and the MM-GB/SA model. The scoring function is not precise. Namely, the average error is about 2.5 kcal/mol.

In the GB method, hydrophilic interactions between solute and solvent are approximately calculated. The solvent is treated as a continuum, and a high dielectric constant is used. One problem with the GB method is the calculation of the cavity effect. When a cavity is generated between the solutes, the calculated electrostatic interactions in the GB model are underestimated because of the continuum solvent model [11, 15, 24]. In the SA method, the hydrophobic interaction between solute and solvent is calculated only approximately. The hydrophobic interaction is not precisely proportional to the ASA. To calculate the protein–ligand binding free-energy (ΔG), the molecular-mechanics GB/SA model (MM-GB/SA) has been used [12, 18]. In the MM-GB/SA method, part of the entropy change associated with the ligand binding is estimated by normal mode analysis [5, 26]. In normal mode analysis, the linearity of the protein structure fluctuation is assumed, and the entropy change due to non-linear motion cannot be estimated precisely.

To calculate ΔG, we developed a new method, the Smooth Reaction Path Generation (SRPG) method [8]. In the SRPG method, first, a smooth reaction path that links a bound state and an unbound state is calculated. Second, the potential of mean force (PMF) along the path is calculated by an MD simulation. Third, the PMF around the bound state is calculated to estimate the probability of the bound state. Finally, the ΔG value between the bound and the unbound states is calculated. In the SRPG method, the entropy change of protein and ligand could be estimated precisely. Before the current study, we calculated the ΔG of the streptavidin–biotin complex in explicit water using the SRPG method. The estimated ΔG was very close to the experimental value.

In the current study, we calculated the ΔG of the streptavidin–biotin complex using the MM-GB/SA model with the SRPG method and the docking score program, Sievgene [9, 33]. These results were compared with the result obtained by the explicit water model.

## 2. Methods and Materials

We calculated the protein–ligand binding free energy, ΔG, using three methods: the SRPG calculation with explicit water, the SRPG calculation with the GB/SA model, and the docking score of Sievgene/myPresto.

When using the SRPG method to calculate the protein–ligand binding free energy, first, the ligand dissociation path from a protein binding site to an unbound state is calculated. Second, the smooth reaction path is calculated from the dissociation path. Third, the PMF along the smooth reaction path is calculated. Fourth, the free energy surface around the bound state is calculated to determine the probability of the bound

state. Finally, the ΔG is calculated from the ratio of the bound and unbound state probabilities.

### 2.1. *ΔG Calculation*

The ΔG is calculated from the ratio of the bound and unbound state probabilities. The ΔG is

$$\Delta G = -k_B T \ln \frac{P_B}{P_U}, \tag{1}$$

where $k_B$, T, $P_B$ and $P_U$ are the Boltzmann constant, temperature, and the probabilities of the bound and unbound state, respectively. Each probability is calculated by integrating an energy function at each state. The probabilities of the bound and unbound state are

$$P_B = \int_{R_B} \exp(-\beta G(r)) dr \tag{2}$$

$$P_U = \int_{R_U} \exp(-\beta G(r)) dr , \tag{3}$$

where $R_B$, $R_U$, β and G(**r**) are the bound and unbound state region, $1/k_BT$ and the free energy at position r, respectively. The free energy around the bound state is approximated using harmonic potentials. The free energy minimum around the bound state is at position $\mathbf{r}_0$. The free energy around the bound state is

$$G(\vec{r}) = G(\vec{r_0}) + (\frac{k_x}{2}\Delta x^2 + \frac{k_y}{2}\Delta y^2 + \frac{k_z}{2}\Delta z^2), \tag{4}$$

where

$$\vec{r} = \vec{r_0} + \overrightarrow{\Delta r} \tag{5}$$

$$\overrightarrow{\Delta r} = (\Delta x, \Delta y, \Delta z) . \tag{6}$$

The xyz axes in Eq. (4) are the principle axes of ligand distribution around the binding site. The force constants, $k_x$, $k_y$, and $k_z$ in Eq. (4) are calculated by the least square fitting of the free energy surface along each axis. The probability of the bound state is

$$P_B = \int_{V_B} \exp(-\beta(G(r_0) + \frac{k_x}{2}x^2 + \frac{k_y}{2}y^2 + \frac{k_z}{2}z^2)) dxdydz , \tag{7}$$

where $V_B$ is the volume of the bound state. The free energy surface around the unbound state is flat. The probability of the unbound state is

$$P_U = \int_0^R 4\pi r^2 \exp(-\beta G(r_\infty))dr = \frac{4\pi}{3} R^3 \exp(-\beta G(r_\infty)) = V_0 \exp(-\beta G(r_\infty)), \quad (8)$$

where R, $G(r_\infty)$ and $V_0$ are the region of the unbound state, the free energy of the unbound state at position $r_\infty$ and the volume around the unbound state. The $r_\infty$ is a coordinate at unbound state, which is obtained from the ligand dissociation path. If the concentration of the ligand around the unbound state is 1M, the volume, $V_0$, is 1661 Å$^3$.

### 2.2. *Ligand Dissociation Path*

The ligand dissociation path is necessary to calculate the smooth reaction path which links bound and unbound states. To obtain the ligand dissociation path, an MD simulation is performed with a starting conformation of the protein–ligand complex at high temperature in vacuum. The filling potential (FP) method is used in the MD simulation, which enables the ligand to drift from its local minima automatically, because the free energy minimum is very deep around the protein–ligand bound state [10]. In addition, the ligand atoms have no atomic charge to reduce the attractive interactions between the protein and the ligand atoms. One of the ligand atoms is selected as a landmark atom to represent the ligand coordinates and trajectories.

### 2.3. *Smooth Reaction Path*

It is necessary to obtain a smooth reaction path to calculate the PMF using the thermodynamics integration (TI) method [4, 28]. In the current study, a smooth reaction path linking the bound and unbound state is calculated from the dissociation path using the Legendre function. The initial and final coordinates are selected from the ligand trajectories, which represent the bound and the unbound states. The Legendre function is complete, so that a linear combination of the Legendre function can represent any kind of function. The Legendre function technique generates various kinds of paths that link the initial and the final coordinates. An appropriate curve is selected as a smooth reaction path. Fig. 1 shows a schematic representation of a smooth reaction path.

### 2.4. *PMF Calculation*

The TI method was applied to the calculation of the potential of mean force (G(R)). The G(R) was calculated by integrating the average force acting on the landmark atom as follows:

$$G(R) = \int_0^R < \vec{F}(\vec{r}) > \cdot d\vec{r}, \quad (9)$$

where $<\mathbf{F}(\mathbf{r})>$ is the average force acting on the landmark atom. The average force was calculated from the MD trajectories. The ligand landmark atom was restricted at a discrete point by using an umbrella potential along the reaction path, and the average

force was calculated by removing the effect of the umbrella potential. The distance between the discrete points was set to 0.1–0.2 Å (Fig. 1).

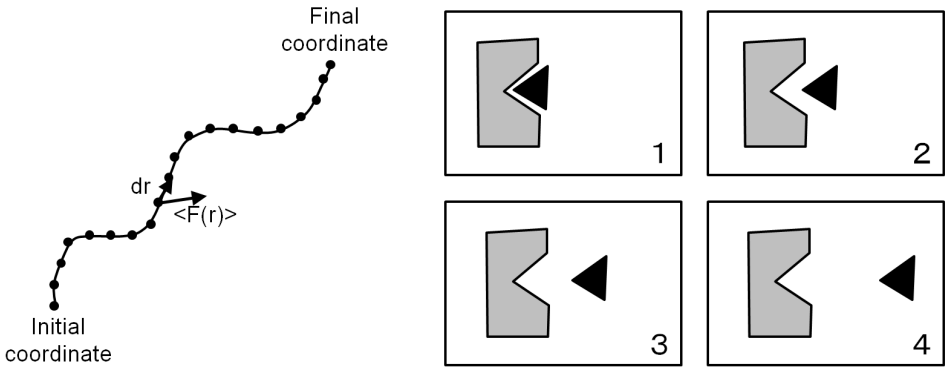


Figure 1. A schematic representation of a smooth reaction path (left) and protein–ligand systems generated along the reaction path (right). Each black point along the path corresponds to the position of a ligand landmark atom. The initial and the final coordinates correspond to bound and unbound states. In the right four boxes, the large gray objects represent proteins, and the black triangles represent ligands. Each box corresponds to a state used in the MD calculation.

### 2.5. *Computational Models*

We adopted a streptavidin–biotin complex system (PDB ID: 1stp) as an example system [31]. We followed the modeling procedure of previous works [10, 32]. The topology files of streptavidin and biotin were created using tplgene and tplgeneL in myPresto [33]. For the streptavidin, all of the Asp, Glu, Arg, and Lys residues were treated as being charged. In the explicit water model, the water molecules were set spherically around the protein–ligand binding site. The radius of the water sphere was 25 Å. Streptavidin and biotin consist of 1744 and 31 atoms, respectively. When using the explicit water model, the system contained 4 $Cl^-$, 7 $Na^+$, and 4857 water atoms. We constrained bond lengths between heavy and hydrogen atoms using the SHAKE algorithm [25]. In the explicit water model, 12 Å residue-base cutoffs were used for 1 to 5 electrostatic interactions. The atomic charges of biotin were determined by the restricted electrostatic point charge (RESP) procedure using the HF/6-31G*-level quantum chemical calculations with the program Gaussian 98 [7]. The Amber force field parm99 was used for streptavidin [6]. The general Amber force filed (GAFF) was used for biotin [29]. The TIP3P water model was applied to the water molecules [16] as same as the previous works.

We also used the GB/SA solvent model. The force field and the atomic charges were exactly the same as those of the explicit water model. The ligand dissociation path was the same as the explicit water model. In the GB model, the dielectric constant of the protein ($\varepsilon_p$) was set to 1. The dielectric constant of the solvent water was set to 78.3. In the SA model, the probe radius was 1.4 Å and the atomic solvation parameter was set to 10 cal/mol/Å$^2$. The SHAKE algorithm was applied to all hydrogen atoms as in the explicit water model. Cutoffs of 25 Å were used for 1 to 5 electrostatic interactions.

The ΔG value was also estimated by the docking simulation program, Sievgene/myPresto. The atomic charges were exactly the same as those of the explicit water model. The ligand dissociation path was the same as that of the explicit water model. We calculated the ΔG value of the protein–ligand complexes using Sievgene along the smooth reaction path.

## 3. Results

We calculated a ligand dissociation path using the MD simulation with the FP method. First, we performed energy minimization of the crystal structure with position restraint at the protein heavy atoms and the ligand landmark atom. The time step was 1.5fs. Then, we performed an MD simulation in a vacuum for the ligand dissociation from the protein–ligand bound state using the FP method. The temperature was set to 700K. In this dissociation calculation, the ligand atomic charges were set to zero to accelerate the ligand dissociation by reducing the attractive interaction. Using the ligand trajectory, we calculated a smooth dissociation path by the Legendre fitting method. A smooth dissociation path linking the bound state and unbound state is shown in Fig. 2. We prepared 112 protein–ligand complex systems along the dissociation path. The distances between neighboring ligand landmark atoms along the dissociation path were less than 0.3 Å.

The PMF obtained by the explicit water model is shown as a curve with filled circles in Fig. 3. The water molecules were assumed to form a sphere (radius=25 Å) around the bound state with a CAP potential. The counter ions, $Cl^-$ and $Na^+$, were also added to neutralize the total charge of the system. Each system was energy-minimized with a position restraint at the protein heavy atoms and the ligand landmark atom.

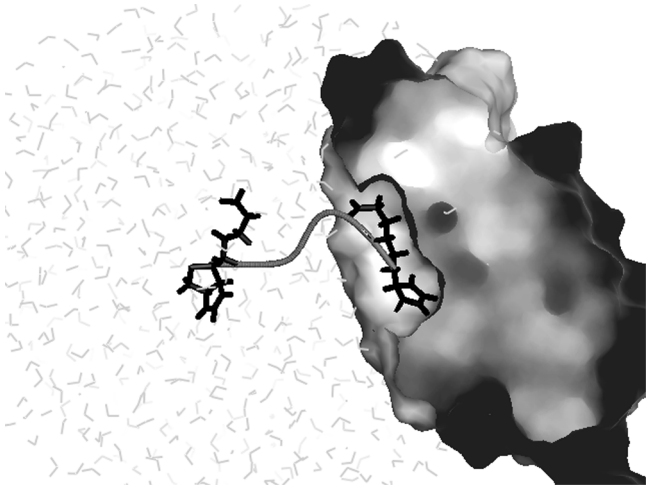

Figure 2. Smooth reaction path (gray curve), streptavidin (surface) and biotins (black stick) with CAP water.

We performed an MD simulation of 1.75ns for equilibration with an umbrella potential at the ligand landmark atom. The time step was set to 1.5fs. The temperature was set to 300K. Then, we performed an MD simulation of 750ps for data sampling. The PMF has a deep minimum at RMSD=0 Å and an energy barrier at RMSD=6 Å. The PMF is continuous. This means that the reaction path is smooth enough to calculate the PMF. The PMF becomes flat when the RMSD > 8 Å. Thus, there is no interaction between the protein and the ligand, and the ligand coordinates at RMSD=15 Å should be in the unbound state.

The PMF obtained by the MM-GB/SA model is shown as a curve with open circles in Fig. 3. We performed an MD simulation with the GB/SA model; the equilibration time was 300 ps, and the data sampling time was 750ps. The time step was set to 1.5fs. The temperature was set to 300K. The PMF had a deep minimum at RMSD=0 Å, and the profile of PMF was similar to that by the explicit water model in the region of RMSD < 6 Å. However, the energy barrier at RMSD=6 Å disappeared. The PMF value converged to a constant as the RMSD value increased; still, the PMF did not become flat when the RMSD > 15 Å. The long-range interaction between the protein and the ligand was not sufficiently shielded by the GB method.

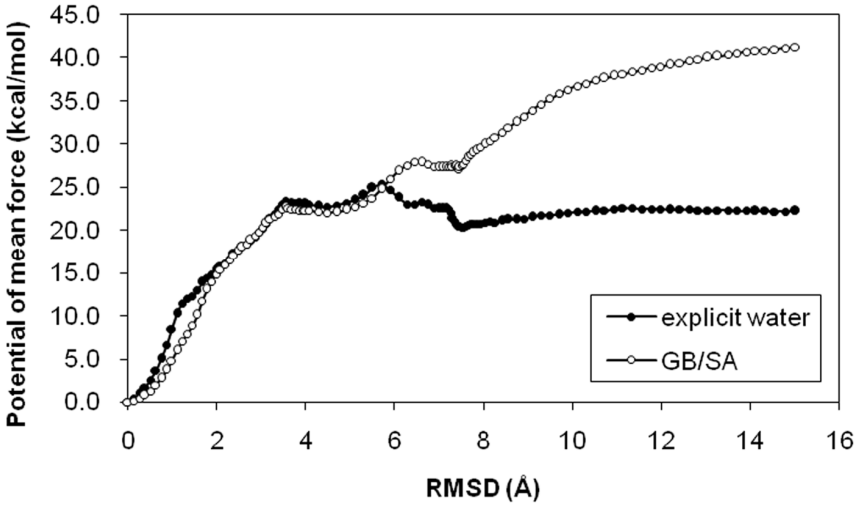


Figure 3. PMF calculated along the smooth reaction path in the explicit and GB/SA solvent models.

The free energy surface obtained by the docking score is shown in Fig. 4. We calculated the ΔG along the dissociation path using the docking program, Sievgene. Fig. 4 is not the PMF, since the docking score corresponds to the ΔG. In Sievgene, the electrostatic potential is estimated by a pair-wise Coulombic potential with distance-dependent dielectric constant (ε = 4 R, where R is the atomic distance in Å unit). The energy profile does not have an energy barrier, as with the PMF by the MM-GB/SA model.

To calculate the probability of the bound state, the free energy surface around the bound state was calculated and fitted to a harmonic potential by the least square method. The PMF along the x, y, and z axes were calculated using the TI method. In this PMF calculation, the ligand was set with an umbrella potential every 0.1 Å from the coordinate of the energy minimum in the bound state. The force constants of the explicit water model, kx, ky and kz were 8.28, 7.15 and 8.27 kcal/mol/Å$^2$. The free energies G($r_0$) and G($r_\infty$) were 0 kcal/mol and 22.1 kcal/mol. The probabilities of the bound and unbound states were calculated from Eq. (7) and Eq. (8). The ΔG was calculated from Eq. (1). The ΔG was -16.5 kcal/mol and the experimental one was -18.3 kcal/mol. The ΔG value obtained by the explicit water model was quite close to the experimental value.

The force constants of the GB/SA model, the kx, ky and kz values were 7.22, 5.63 and 6.08 kcal/mol/Å$^2$. The free energies G($r_0$) and G($r_\infty$) were 0 kcal/mol and 41.2 kcal/mol. The calculated ΔG value was -35.7 kcal/mol. The GB/SA model obviously overestimated the ΔG value, while the same reaction path, the same atomic-charge values, and the same method (SRPG method) were used. The force constants of the GB/SA model were close to those of the explicit water model. This means that the free energy surface around the bound state of the GB/SA model is similar to that of the explicit model. However, the energy difference of G($r_0$)-G($r_\infty$) by the GB/SA model was two times greater than that by the explicit water model. Thus, the estimation of the long-range interaction should cause this error.

The ΔG value by the docking score was -8.74kcal/mol. The docking program underestimated the ΔG value. The error was about 10kcal/mol, much larger than the average error of this docking score of 2.5 kcal/mol. The ΔG values obtained by these methods and their errors are summarized in Table 1.

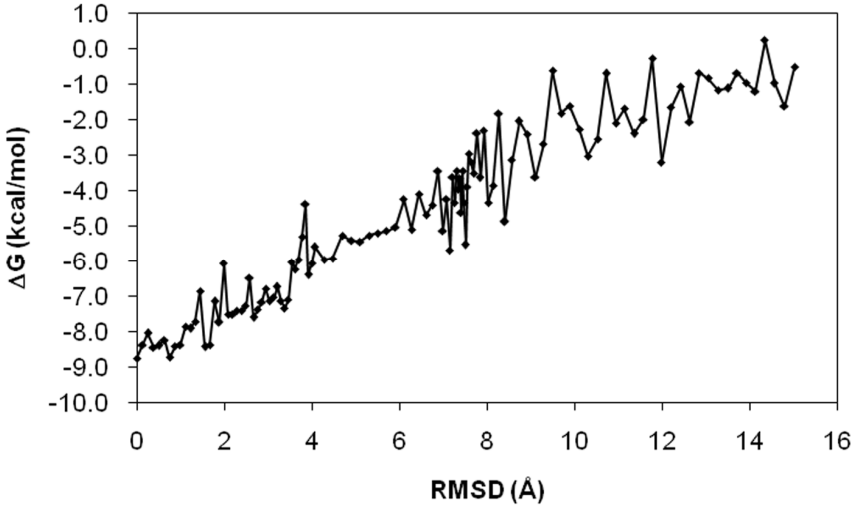


Figure 4. ΔG value along the smooth reaction path obtained by the docking program, Sievgene.

Table 1: ΔG values obtained by various methods and their errors

| Method | ΔG (kcal/mol) | Error (kcal/mol) |
|---|---|---|
| Exptl | -18.3 | – |
| explicit water model | -16.5 | -1.8 |
| MM-GB/SA | -35.7 | 17.4 |
| Docking score (sievgene) | -8.74 | -9.56 |

## 4. Discussion

The PMF obtained by the explicit water model was almost the same as that by the GB/SA model in the region of RMSD < 6 Å. When RMSD < 6 Å, the ligand was at the binding site. In the region of RMSD > 7 Å, the ligand was outside of the binding site. The PMF by GB/SA gradually increased after the dissociation. At RMSD=15 Å, the free energy difference between the explicit and the GB/SA models was about 20 kcal/mol. This value is very large. The counter ions of the explicit water model can shield the electrostatic interaction between the protein and the ligand. In the GB/SA model, the dielectric constant of the solvent was set to 78.3 without counter ions. This result suggests that the counter ions could be considered to affect the ΔG and PMF of this system.

The ΔG value obtained by the docking score was -8.74 kcal/mol. The ΔG value was very different from the experimental value of -18.3 kcal/mol. The ΔG value of streptavidin–biotin is very large compared to the ΔG values for many other protein–ligand complexes. Namely, in common drug molecules the ligand efficiency (LE), which is a ΔG per one heavy atom of a ligand, is 0.2-0.5 [1, 21]. In these cases, the docking score function can give appropriate ΔG values with the average error of 2.5 kcal/mol. On the contrary, in the case of streptavidin–biotin, LE=1.14, which is much higher than usual. The calculated LE value obtained by the docking score was 0.55. In general, the entropy change of protein is large value comparing to the ΔG value. The ΔG value of soft protein should be a small value and that of hard protein should be a large value. The streptavidin–biotin system is hard protein, so that the docking method should underestimate the ΔG value. The score function of Sievgene is based on the vdW, ASA, hydrogen bond and Coulomb interactions. The scoring function itself is not so different from the other scoring functions qualitatively. The rigidness of the protein should cause the high LE, since the target protein showed high rigidness. Thus, the current result should be qualitatively similar to that obtained by the other scoring function.

The ΔG value obtained by the SRPG method does not depend on the reaction path by its definition. While the reaction path is not well-defined such as intrinsic reaction coordinate, the RMSD values in the current study is rough indication. Thus, we must note that the profile of PMF itself depends on the reaction path.

## 5. Conclusion

We compared the protein–ligand binding free energies (ΔG) obtained by the explicit water model, the MM-GB/SA model, and the docking scoring function. The free energies by the explicit water model and the MM-GB/SA model were calculated by the previously developed SRPG method. We applied these methods to the streptavidin-and-biotin system. The ΔG value by the explicit water model (ΔG=-16.5 kcal/mol) was close to the experimental value (ΔG=-18.3 kcal/mol). The ΔG value by the MM-GB/SA model (ΔG=-35.7 kcal/mol) was overestimated and that by the scoring function (ΔG=-8.7 kcal/mol) was underestimated. In the current study, the explicit water model was shown to be the most precise among these three methods.

The free energy surface by the explicit water model was close to that by the GB/SA model around the bound state, while the discrepancy was observed at the distance (RMSD) > 6 Å. The PMF obtained by the explicit water model became flat in the region of RMSD > 6 Å. On the contrary, the PMFs obtained by the MM-GB/SA model and the scoring function were not flat in the region of RMSD > 6 Å. The SRPG method can precisely take into account the entropy change of the protein and the ligand. Thus, the difference in the long-range Coulomb interaction should cause the error in ΔG. The explicit water model (with the counter ions) can strongly shield the electrostatic interaction, but the GB/SA model should weakly shield this electrostatic interaction. The GB/SA model adopted in the current study cannot take into consideration the concentration of counter ions. Consideration of the counter ions could improve the accuracy of the MM-GB/SA method. The ΔG value obtained by the scoring function was very different from the experimental value. The scoring function cannot take into account the entropy change of protein. Thus, the error of ΔG value could depend on the target protein. This result suggests that the accuracy of the scoring function could be improved by taking into account the entropy change of the target protein.

## Acknowledgments

This work was supported by grants from the New Energy and Industrial Technology Development Organization of Japan (NEDO) and the Ministry of Economy, Trade, and Industry (METI) of Japan.

## References

[1] Abad-Zapatero, C. and Metz, J.T., Ligand efficiency indices as guideposts for drug discovery, *Drug Discov. Today*, 10:464-9, 2005.
[2] Abagyan, R., Totrov, M., and Kuznetsov, D., ICM: a new method for structure modeling and design: application to docking and structure prediction from the disordered native conformation, *J Compt. Chem.,* 15:488-506, 1994.

[3] Baxter, C.A., Murray, C.W., Clark, D.E., Westhead, D.R. and Eldridge, M.D., Flexible docking using tabu search and an empirical estimate of binding affinity, *Proteins: Structure, Function, and Genetics*, 33:367-82, 1998.

[4] Beveridge, D.L. and DiCapua, F.M., Free energy via molecular simulation: applications to chemical and biomolecular systems, *Annu. ReV. Biophys. Biophys. Chem.*, 18:431-92, 1989.

[5] Brooks, B. and Karplus, M., Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor, *Proc. Natl. Acad. Sci. U S A*, 80:6571–75, 1983.

[6] Case, D.A., Darden, T.A., Cheatham, T.E. III., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Merz, K.M., Wang, B., Pearlman, D.A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Caldwell, J.W., Ross, W.S. and Kollman, P.A., *AMBER 8*, UCSF, 2004.

[7] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Zakrzewski, V.G., Montgomery, J.A., Stratmann, R.E. Jr., Burant, J.C., Dapprich, S., Millam, J.M., Daniels, A.D., Kudin, K.N., Strain, M.C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G.A.; Ayala, P.Y.; Cui, Q.; Morokuma, K.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B., Cioslowski, J., Ortiz, J.V., Baboul, A.G., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Andres, J.L., Gonzalez, C., Head-Gordon, M., Replogle, E.S., Pople, J.A., *Gaussian 98, Revision A.9*, Gaussian, Inc., Pittsburgh, PA, 1998.

[8] Fukunishi, Y., Mitomo, D. and Nakamura, H., Protein-ligand binding free energy calculation by the smooth reaction path generation (SRPG) method, *J. Chem. Info. Model.*, 49:1944-51, 2009.

[9] Fukunishi, Y., Mikami, Y. and Nakamura, H., Similarities among receptor pockets and among compounds: Analysis and application to *in silico* ligand screening, *J. Mol. Graph. and Model.*, 24:34-45, 2005.

[10] Fukunishi, Y., Mikami, Y. and Nakamura, H., The filling potential method: A method for estimating the free energy surface for protein-ligand docking, *J. Phys. Chem. B.*, 107:13201-10, 2003.

[11] Fukunishi, Y. and Suzuki, M., Potential of mean force calculation of solute molecules in water by a modified solvent-accessible surface method, *J. Phys. Chem.*, 100:5634-36, 1996.

[12] Gohlke, H. and Case, D.A., Converging free energy estimates: MM-PB-(GB)SA studies on the protein-protein complex Ras-Raf, *J. Comput. Chem.*, 25:238-50, 2004.

[13] Goodsell, D.S. and Olson, A.J., Automated Docking of Substrates to Proteins by Simulated Annealing, *Proteins: Structure, Function and Genetics*, 8:195-202, 1990.

[14] Hawkins, D.G., Cramer, J.C. and Truhlar, G.D., Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium, *J. Phys. Chem.*, 100:19824-39, 1996.

[15] Hummer, G., Garde, S., Garcia, A. E., Pohorille, A. and Pratt, L.R., An information theory model of hydrophobic interactions, *Proc. Natl. Acad. Sci. USA*, 93:8951–55, 1996.
[16] Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L., Comparison of simple potential functions for simulating lipid water, *J. Chem. Phys.*, 79:926-35, 1983.
[17] Jones, G., Willet, P., Glen, R.C., Leach, A.R. and Taylor, R., Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol.*, 267:727-48,1997.
[18] Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T.. Duan, Y., Wang, W., Doninni, O., Cieplak, P., Srinivasan, J., Case, D.A. and Cheatham III, T.E., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models, *Acc. Chem. Res.*, 33:889-97, 2000.
[19] Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., A Geometric approach to macromolecule-ligand interactions, *J. Mol. Biol.*, 161:269-88, 1982.
[20] Mitomo, D., Watanabe, Y.S., Kamiya, N., and Higo, J., Explicit and GB/SA solvents: Each with two different force fields in multicanonical conformational sampling of a 25-residue polypeptide, *Chem. Phys. Lett.*, 427:399-403, 2006.
[21] Orita, M., Ohno, K. and Niimi, T., Two 'Golden Ratio' indices in fragment-based drug discovery, *Drug Discov. Today*, 14:321-8, 2009.
[22] Qiu, D., Shenkin, P.S., Hollinger, F.P. and Still, W.C., The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii, *J.Phys.Chem.A*, 101:3005-14, 1997.
[23] Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., A fast flexible docking method using an incremental construction algorithm, *J. Mol. Biol.* 261:470-89, 1996.
[24] Rashin, A.A., Electrostatics of Ion-Ion Interactions in Solution, *J. Phys . chem.*, 93: 4664-9, 1989.
[25] Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C., Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comp. Phys.*, 23:327-41, 1977.
[26] Srinivasan, J., Cheatham, T.E., Cieplak, P., Kollman, P.A. and Case, D.A., Continuum solvent studies of the stability of DNA, RNA and phosphoramidate-DNA helices, *J. Am. Chem. Soc.*, 120:9401-9, 1998.
[27] Still, W.A., Tempczyk, A., Hawley, R.C. and Hendrickson, T., Surface Concentrations and Residence Times of Intermediatof Methane, *J. Am. Chem. Soc.*, 112:6127-9, 1990.
[28] Straatsma, T.P. and McCammon, J.A., Computational Alchemy, *Annu. ReV. Phys. Chem.*, 43:407-35, 1992.
[29] Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A., Development and testing of a general amber force field, *J. Compt. Chem.*, 25:1157-74, 2004.
[30] Watanabe, Y.S., Kim, J.G., Fukunishi, Y. and Nakamura, H., Free energy landscapes of small peptides in an implicit solvent model determined by force-

biased multicanonical molecular dynamics simulation, *Chem. Phys. Lett.*, 400:258-263, 2004.

[31] Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. and Salemme, F.R., Structural origins of high-affinity biotin binding to streptavidin, *Science* 243:85-8, 1989.

[32] Woo, H.J. and Roux, B., Calculation of absolute protein-ligand binding free energy from computer simulations, *Proc. Natl. Acad. Sci. USA*, 102:6825–30, 2005.

[33] http://medals.jp/myPresto/index.html