The Filling Potential Method: A Method for Estimating the Free Energy Surface for Protein−Ligand Docking

Yoshifumi Fukunishi,*,† Yoshiaki Mikami,†,‡ and Haruki Nakamura†,§

Japan Biological Information Research Center (JBIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6, Aomi, Koto-ku, Tokyo 135-0064, Japan, Japan Biological Informatics Consortium (JBIC), 2-41-6, Aomi, Koto-ku, Tokyo 135-0064, Japan, and Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

Received: May 27, 2003; In Final Form: September 11, 2003

We developed a molecular simulation method suitable for estimation of binding free energy, called the filling potential method, based on the concept of the Taboo search, which is a type of self-avoiding random walk consisting of a cycle of local-minimum searches and transition state searches. The filling potential method is an umbrella potential sampling method, and enables the ligand molecule to drift from its local minima automatically. In the case of the filling potential method, the umbrella potential is a combination of Gaussian-type repulsive potentials, which are located on the trajectory of the ligand. Without setting the reaction coordinates a priori, this method searches for and determines the suitable reaction coordinates by successive generation of umbrella potentials based on its trajectory analysis. The weighted histogram analysis for these trajectories gives the binding free energy of the ligand to the receptor protein. It was applied to the complex of thermolysin and its inhibitors in explicit water, and the free energy surfaces with the stable binding state and the energy barrier were examined. The calculated binding free energies agreed well with the experimental results.

I. Introduction

Binding free energy (ΔG) is the most important measure of ligand binding activity. Precise estimation of ΔG based on statistical physics is one of the most challenging problems in computational chemistry. To meet this challenge, many functions and algorithms have been developed for assessing and rationalizing the protein−ligand interactions. In addition to a knowledge-based scoring system for estimating ΔG, several force fields with implicit water models have been widely used for various simulations, including rigid and flexible docking simulations.[1] For these simulations, ΔG has been evaluated by the semiempirical method by using the solvent-accessible surface area (SA) and the continuum theory, in which water molecules are treated as a continuum with a high dielectric constant.[2−6] The results of recent studies with the generalized Born (GB)/SA method have shown good agreement with the results of the more time-consuming method to solve the Poisson−Boltzmann equations.

Problems with the implicit water model begin to arise when a solute molecule approaches in close proximity to another solute molecule.[7−9] One such problem is a change of the effective dielectric constant due to a small cavity between the two solutes. This affects the hydrophilic interaction. Another problem is the additional cavity formation energy due to the small cavity between two solute molecules. The additional energy arises from the failure of the SA approximation. The scaled particle theory, which is one of the basic theories underlying the SA method, assumes that the solute is in a spherical cavity. This problem affects hydrophobic interaction. The correction for these two effects is not straightforward, because there is no pairwise potential that can represent both of them.

The application of precise simulation with an explicit water model is restricted to evaluations of the difference of ΔG for different compounds with high similarity. The popular methods for this purpose are free energy perturbation (FEP) and thermodynamic integration (TI).[10−15] A recently developed method allows estimation of the free energy difference ($\Delta\Delta G$) of several ligands in a single MD simulation. In addition, FEP and TI have been applied to calculate the potential of mean force (PMF) of small solute molecules,[16,17] of which the energy difference between binding and dissociated states gives ΔG. To calculate PMF, we must know the reaction path a priori. Some recent studies have used MD simulation to investigate the protein−ligand adsorption/dissociation process,[18,19] but in most cases, the reaction paths of protein−ligand adsorption or dissociation processes remain unknown. Grubmüller et al. developed a force probe MD simulation that mimics a single molecule atomic force microscopy.[20−22] Using their force probe simulation, they studied the dissociation path of a small ligand molecule from a receptor protein in explicit water. Even in this method, the direction of dissociation must be given a priori.

The difficulty in the conventional MD simulation (classical dynamics) is the poor sampling, which is an intrinsic feature of the Boltzmann−Gibbs ensemble. There have been many reports on the efficient application of non-Boltzmann sampling methods to perform highly effective sampling on artificial potential surfaces and/or artificial dynamics. The multicanonical MD first developed by Berg et al. is a powerful method for conformational sampling,[23,24] and Nakajima et al. extended the original Monte Carlo simulation to an equation of motion for an MD simulation,[25] an extension that rendered McMD suitable for

* Address correspondence to this author. E-mail: y-fukunishi@jbirc.aist.go.jp.
† National Institute of Advanced Industrial Science and Technology.
‡ Japan Biological Informatics Consortium.
§ Osaka University.

10.1021/jp035478e CCC: $25.00
Published on Web 11/01/2003

application to a biological system. They then applied McMD to a docking study of a short proline-rich peptide to an Src homology-3 (SH3) domain system. Straub et al. proposed an MD on a Tsallis potential,[26,27] which realized the Tsallis-type effective potential energy. This method was also applied to a docking of the streptavidin/biotin system.[26,27] The disadvantage of McMD is the length of the iteration procedure required to obtain the multicanonical ensemble. In some cases, this method requires several-ten times iteration of several tens nanoseconds MDs. Both the McMD and Tsallis potential methods are MD simulations on a scaled potential surface, and thus the only potential energy distribution follows the multicanonical or Tsallis ensemble. Recently, Fukuda et al. proposed a Tsallis dynamics in which both the potential and kinetic energy distributions can satisfy the Tsallis ensemble.[28]

As an alternative means of solving this problem, we here present a method utilizing umbrella sampling, the filling potential (FP) method, which is based on the concept of the Taboo search, a type of self-avoiding random walk, and which repeats a search for the local-minimum and the transition state. The main purpose of this method is to calculate ΔG starting from a given protein−ligand complex structure. To calculate ΔG, an MD trajectory or a set of ensembles of the atom coordinates must connect the binding state with the dissociated state. The method for deriving the reaction path is based on the concept of the Taboo search.

The Taboo search is a global search algorithm that consists of sequential searches of the local minimum and the saddle points.[29] Starting from an initial coordinate, the system reaches a local minimum that is determined by means of a conventional search algorithm. Then, the system climbs the potential surface in the direction of the most tolerant gradient. The system overcomes the energy barrier and goes down to the next minimum by the local minimum search. These steps are repeated until there is no new local minimum left. The point is that the system remembers the previous trajectory, and the trajectory becomes self-avoiding. That is, when the system visits a local minimum that it has visited before, it climbs in a direction that is different from that followed previously. To preserve memory, the system remembers the trajectory of the selected information. Starting from a global minimum, the system visits other local minima, including the dissociated state, if there exists an activation energy for adsorption.

The concept of the Taboo search has previously been applied to the MD simulation to create the so-called local elevation method.[30] The local elevation method is a type of umbrella potential method that consists of the sum of many Gaussian-type repulsive potentials that represent the memory of the trajectory. This method has been successfully applied to pentane and cyclosporin by using the repulsive potential on main-chain dihedral angles.

A related concept is known as the filling function method. When the system is trapped in a local minimum, this method uses a Gaussian-type function to fill up the potential surface, and then the system can escape from the minimum. To use this method, the depth of each minimum must be known a priori.[31] In the usual MD simulation, the system follows the free energy surface, which is not known a priori and is what we wish to determine. Kumar et al. proposed an iterative method that utilizes weighted histogram analysis to generate the distributions iteratively and thereby flatten the free energy surface.[32−34] The energy surface is flattened by subtraction of the predicted free energy surface, which is calculated from the previous simulation. This method succeeded in calculating the PMF of an alanine peptide. An MD simulation with the WHAM method has also been applied to a ligand−protein docking analysis with empirical potential function.[33,34] Grubmüller et al. proposed a successive generation of the umbrella potentials based on the principal component analysis for an MD trajectory.[35] The umbrella potential consists of the sum of many Gaussian-type functions that represent the PMF. This method was applied to a small protein system, and it succeeded in realizing a global conformation search of the small protein. Voter proposed an accelerated molecular dynamics method based on transition state theory.[36,37] This method utilizes a bias potential as an umbrella potential to accelerate the state-to-state transition rate constant. The effect of the bias potential is measured by "boosted time", which shows how many times the simulation is accelerated. This method was applied to examine the diffusion process of metal atoms on a metal surface.

Global conformation search is the main purpose of these sampling methods, while we are interested in the free energy difference between the protein−ligand binding state and dissociated state. For this purpose, we only need local sampling around the dissociation reaction path. In this work, we present a method suitable for estimation of binding free energy, the FP method, and apply it to two protein−ligand complexes in explicit water: the Bacillus Thermoproteolyticus hydrolase (thermolysin) with either of its two inhibitors.

II. Method

The FP method is an umbrella potential sampling method that enables the ligand molecule to drift from its local minima automatically. The umbrella potential is a combination of Gaussian-type repulsive potentials, which are located on the trajectory of a ligand. Without setting a reaction coordinate a priori, this method searches and determines the suitable reaction coordinates by the successive generation of umbrella potentials based on its trajectory analysis. The method for obtaining the reaction path is based on the concept of the Taboo search.

To mimic a self-avoiding trajectory, the FP method sets a repulsive potential on the previous position of the ligand in the MD simulation. If the repulsive potential is not high enough to avoid a visit by the ligand, the ligand can visit the same position again. An additional repulsive potential is then added to the previous repulsive potential, until the ligand can no longer visit the same position. Since the FP method constructs the potential based on a set of the previous positions of the ligand, for the longer simulation, a larger allocation of memory is required. To reduce the memory requirement, the method does not use the coordinates of all atoms, but rather uses the coordinates of only one or a few atoms to construct the potential. The selected atoms are designated as "landmarks". Usually a landmark atom is an atom of the ligand molecule.

Our idea is similar to that of conformational flooding and local elevation.[30,35] The difference is that conformational flooding and local elevation is a global structure search, whereas our goal is to calculate the PMF along a dissociation path. Thus, in our method, the trajectory of the ligand molecule is restricted along the reaction path and WHAM is used to reduce the computational error.

The umbrella potential for FP (V_{FP}) consists of the repulsive potentials (V_{rep}) and a position-restraint potential (V_{res}). The umbrella potential for the nth simulation (V_{FP}^{n}) is

$$V_{FP}^{n} = \sum_{k=1}^{n-1} V_{rep}^{k} + V_{res}^{n} \quad (1)$$

where V_{rep}^k is a repulsive potential constructed from the kth simulation, and V_{res}^n is a position restraint potential that restrains the ligand around the initial coordinates of the nth simulation. Both V_{rep}^k and V_{res}^n act on the landmark atom(s) only.

Let $X_i(t)_{landmark} = \{x_i(t), y_i(t), z_i(t)\}_{landmark}$ be a set of x, y, z coordinates at time t of the ith landmark atom(s). Let $X_i^k = (x_i^k, y_i^k, z_i^k)$ be the final coordinates of the ith landmark atom of the kth simulation. V_{rep}^k is a Gaussian-type function defined as follows:

$$V_{rep}^k(X(t)_{landmark}) = c(k) \exp\left[\sum_{i=1}^{landmark} \{(X_i(t) - X_i^k)A^t(X_i(t) - X_i^k)\}\right] \quad (2)$$

where $c(k)$ and 3×3 real-symmetric matrix $\mathbf{A}$ determine the height and the anisotropic width of the Gaussian potential. In our simulation, the matrix $\mathbf{A}$ is simplified as $a_x E$, where a_x is a constant and E is a 3×3 unit matrix.

V_{res}^n works to localize the ligand around a specific position. Suitable choices of V_{res}^n can retain the overlap of trajectories between the nth and the $(n+1)$th simulations. The overlap of trajectories is essentially important to determine the free energy surface afterward. Let $D_i^k(t)$ be a distance of the ith landmark atom between that at time t and the final position of the kth simulation.

$$D_i^k(t) = \sqrt{(x_i(t) - x_i^k)^2 + (y_i(t) - y_i^k)^2 + (z_i(t) - z_i^k)^2} \quad (3)$$

V_{res}^n is a kind of CAP-type potential defined as follows:

$$V_{res}^n(X(t)_{landmark}) = \sum_{i=1}^{landmark} \frac{1}{2} k_i^n D_i^{n-1}(t); \quad \sum_{i=1}^{landmark} D_i^{n-1}(t) \geq D_0 = 0; \quad \sum_{i=1}^{landmark} D_i^{n-1}(t) < D_0 \quad (4)$$

where k_i^n is a force constant and D_0 is a radius of this half harmonic potential. For $n = 1$, x_i^k, y_i^k, and z_i^k are set to the initial coordinates of the ith atom in the first simulation.

The first simulation is performed on the umbrella potential $V_{FP}^1 = V_{res}^1$. The ligand molecule moves around the initial coordinate. The second simulation is performed on $V_{FP}^2 = V_{rep}^1 + V_{res}^2$. The ligand again moves around the final coordinate of the first simulation. The third simulation is performed on $V_{FP}^3 = V_{rep}^1 + V_{rep}^2 + V_{res}^3$. The ligand moves around the final coordinate of the second simulation and feels a repulsive potential from the first coordinate of the first and the second simulations. This iterative procedure is performed until the ligand reaches the dissociated state. The parameter a_x of eq 2 corresponds to the width of the umbrella potential. The width of the umbrella potential must be smaller than the system size. For example, a constant umbrella potential, which covers the whole system, is useless. The height of the umbrella potential must be small to retain the overlap of trajectories of the different states.

WHAM is a trajectory analysis method that combines several trajectories of different umbrella potentials at different temperatures.[32] Let L be the total number of iterations. Let β be the Boltzmann factor. In the FP method, all simulations are performed at the same temperature. Let n_m be the number of the samples in the mth simulation. Let $N_k(X)$ be the number of the samples at coordinate X in the kth simulation. f_m is the nondimensional free energy of the state in the mth simulation.

WHAM minimizes the expected computational error by solving eqs 5 and 6 simultaneously.

$$P_{j,\beta}(X) = \frac{\sum_{k=1}^{R} N_k(X) \exp\{-\beta V_{FP}^j(X)\}}{\sum_{m=1}^{R} n_m \exp\{f_m - \beta V_{FP}^m(X)\}} \quad (5)$$

$$\exp(-f_j) = \sum_{\xi} P_{j,\beta}(X) \quad (6)$$

To maintain a high level of accuracy, the following equation is used in place of eqs 5 and 6, as described previously.[32]

$$\exp(-f_i) = \sum_{k=1}^{R}\sum_{t=1}^{nk} \frac{\exp(-\beta V_{FP}^i\{X^k(t)\})}{\sum_{m=1}^{R} n_m \exp(f_m - \beta V_{FP}^m\{X^k(t)\})} \quad (7)$$

What we are interested in is the free energy along the reaction path. The free energy F_i around the ith landmark $\{x_i, y_i, z_i\}$ is calculated from the average value of $P_{0,\beta}(X)$ of eq 5 in a sphere region around the ith landmark. When the 0th and the Nth states stand for the binding state and the dissociated state, then the free energy difference $(F_0 - F_N)$ corresponds to the binding free energy.

A new MD simulation program system, prestoX, was developed. All the MD simulations were performed by an MD module, cosgene, which is an extended version of PRESTO.[38] The FP method is available in this program system. All MD calculations were carried out on a Compaq ES40 cluster system.

In the following applications, we applied the parm99 force field to the protein model, and the general AMBER force field (gaff) to the ligand molecules.[39−41] The atomic charges of ligand molecules are determined by the restricted electrostatic point charge (RESP) procedure with use of HF/6-31G* level quantum chemical calculations.[39] We used GAMESS and Gaussian98 to perform these quantum chemical calculations.[42,43]

III. Results and Discussion

A. CH_4 and CH_4 Pair in Water Systems. We applied the FP method to a methane molecule in water and a methane pair in water. In the former system, one methane molecule is in a cube containing 215 TIP3P water models.[45] In the latter system, two methane molecules are in the same cube containing 214 TIP3P water models. Each edge of the cube is 18.64775 Å, and a periodic boundary condition is applied to both cases. The AMBER parm99 force field was used with an 8 Å residue-base cutoff for the nonbonded interactions. All simulations were performed at 300 K and the unit time step was 2.0 fs with the SHAKE constraint for hydrogen atoms.[46] In the methane pair system, a position restraint potential restricts one methane molecule at the center of the cube.

The FP procedure was as follows.

Step 1: Preparatory MD simulation of 60 ps at 300 K was performed for equilibration of the system.

Step 2: The trajectory of landmark atom(s) was sampled by the 100 ps MD simulation at 300 K with the umbrella potential V_{FP}. For the first MD, $V_{rep} = 0$.

Step 3: The parameters of V_{FP} were determined by the last coordinate of the landmark atom.

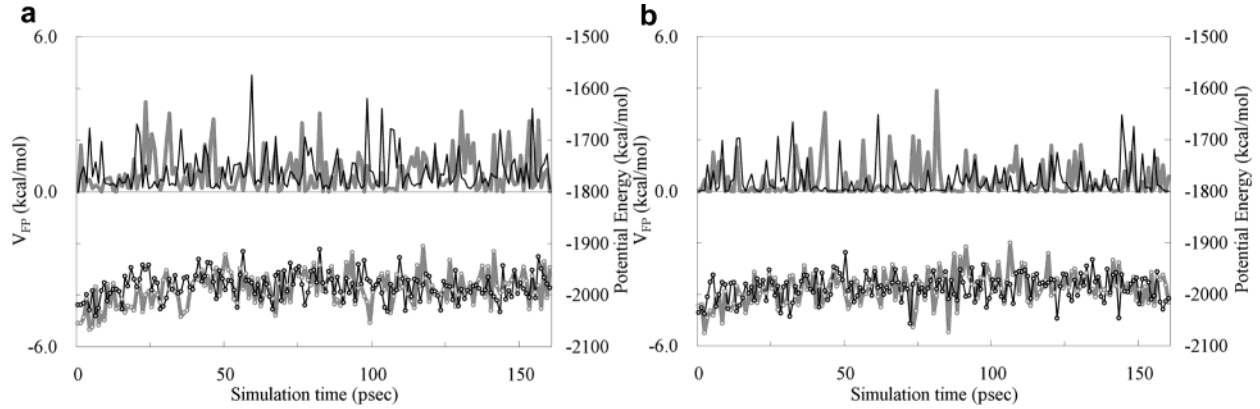

Figure 1. Energy trajectories of methane in water systems of the 1st and the 20th simulations. Circles and solid lines represent the total energy and the umbrella potential energy (V_{FP}), respectively. The black and gray lines represent the results of the 1st and the 20th simulations, respectively.

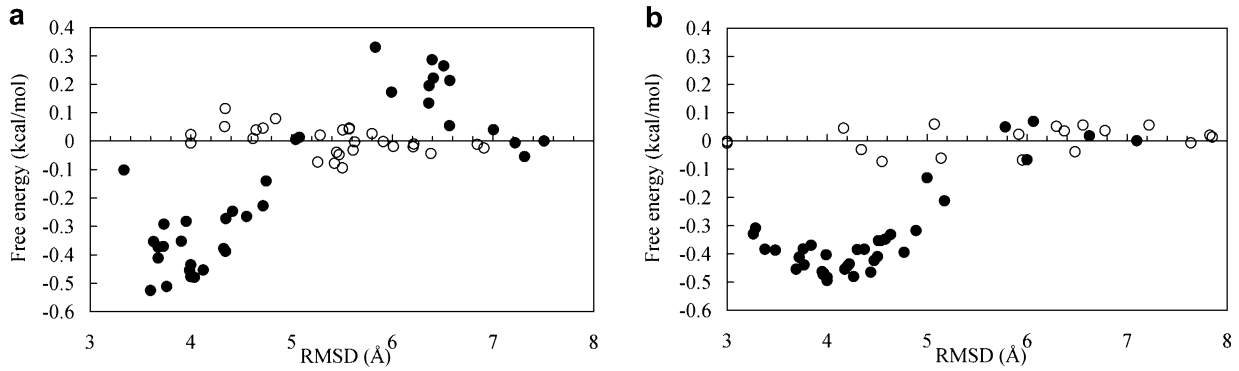

Figure 2. PMFs of single methane and a methane pair in water. Open and filled circles represent the PMFs of a single methane molecule and a methane pair, respectively. The RMSD for the single methane system represents the RMSD from the initial coordinate of the methane. The same axis represents the carbon−carbon distance for the methane pair system.

Step 4: V_{FP} was reconstructed. Then, we returned to step 1. This procedure was repeated until the ligand molecule reached the dissociated state.

Step 5: After steps 1−4 were completed, we applied WHAM to all of the trajectories.

This procedure required 20−30 iterations until the ligand reached the dissociated state. The landmark atom was a carbon atom of the methane molecule. The WHAM procedure converged after 30−50 iterations, and the number of necessary iterations depended on the number of trajectories.

Panels a and b in Figure 1 show the energy trajectories of the 1st and 20th simulations with different parameters. The parameters of Figure 1a were $a_x = 10.0$ Å^{-2}, $c = 0.5$ kcal/mol, k of eq 4 = 5.0 kcal/mol/Å^2, and D_0 of eq 4 = 0.2 Å. Those of Figure 1b were $a_x = 10.0$ Å^{-2}, $c = 0.5$ kcal/mol, k of eq 4 = 10.0 kcal/mol/Å^2, and D_0 of eq 4 = 1.0 Å. Since the perturbation by V_{FP} is much smaller than the energy fluctuation, we consider that the systems were close to equilibrium within 60 ps. Panels a and b in Figure 2 show the PMFs of one and two methane molecules in water; the parameter sets for panels a and b in Figure 2 are the same as those for Figure 1, panels a and b, respectively. Free energy F_i is calculated from the P of eq 5 averaged in a sphere region of 0.2 Å radius. Open and filled circles represent the results of one and two methane molecule systems, respectively. The PMFs of one methane molecule were flat in panels a and b in Figure 2, and these results were considered satisfactory. The standard deviation of PMF was about 0.1 kcal/mol, which corresponds to the computational discrepancy of this procedure. The PMFs of the methane pair show that the equilibrium CH_4-CH_4 distance is about 4 Å, and the ΔG of this system is about 0.5 kcal/mol. There is a low energy barrier for the dissociation and the adsorption around RMSD = 6 Å. The calculated binding free energy does not depend on these parameters. These results agree with the results of previous studies.[45]

Figure 2a,b shows the distribution of F_i vs D_i^k in eq 3, and so the definition of the axis is slightly different from that of the previous study.[46] The trajectory of each state distributes in the range of 0.1−0.3 Å, and thus we can discuss the binding free energy and the rough shape of the PMF, allowing 0.3 Å error in RMSD.

B. Ligand and Protein Model System. We applied the FP method to the Bacillus Thermoproteolyticus hydrolase protein (thermolysin) with the inhibitor benzyloxycarbonyl-D/L-threonine (z-D-Thr/ z-L-Thr) complex system.

Figure 3 shows the model complex system composed of thermolysin with the inhibitor z-D-Thr. The model protein structure for the MD simulation consists of 136 residues, from Gly109 to Ser245. Three asparatic acids, Asp124, Asp126, and Asp226, and one glutamic acid, Glu160, were replaced by their protonated states to neutralize the total charge of the protein model, since the Cα carbons of these amino acids are more than 11 Å away from the zinc ion in the docking pocket. We set water molecules in a spherical region of radius 22 Å around the docking pocket. The total number of atoms of these systems was 4645 and 4636 for the z-D-Thr and z-L-Thr systems, respectively.

Figure 4 is a schematic representation of the ligand-binding pocket, in which the ligand molecule interacts with the residues around the zinc ion. The dashed lines represent the intermolecular interactions. His142, His146, and Glu166 hold the zinc ion, which binds the carboxyl group of the ligand. There

Figure 3. Black stick models at the left and right of the figure represent the initial and final coordinates of z-D-Thr, respectively. The gray ribbon represents the portion of the thermolysin used for this simulation. The black sphere represents the zinc ion, which is located in the reaction center. The small gray sticks represent water molecules.

are three different kinds of interactions between the ligand and the protein. They are the Coulombic interaction between the zinc and the carboxyl group, the hydrogen bond between the amide NH of the ligand and the amide oxygen of the Ala113 backbone, and the hydrophobic interaction between the phenyl group and several hydrophobic residues, Phe130, Val139, Ile188, and Leu202. The binding mode of z-L-Thr is qualitatively the same as that of z-D-Thr.

We applied the position restraint to the backbone atoms of the protein during the MD simulations. Because our model protein is only a part of the whole protein, some hydrophobic residues, which form a hydrophobic core in the original protein, may be located on the surface. In addition, the protonation of Asps and Glu could make the native protein structure unstable.

A CAP boundary condition with 22 Å radius was applied to the ligand and TIP3P water molecules. All MD simulations utilized a 12 Å residue-base cutoff procedure for the nonbond interaction and the SHAKE procedure for all hydrogen atoms, and the unit time step was 2.0 fs. The temperature was controlled by the Hoover−Evans Gaussian restraint method. All complex structures were relaxed by 500 ps canonical simulation at 310 K.

The RESP charges are summarized in Table 1, and the atomic charges of z-D-Thr are exactly the same as those of z-L-Thr; the chiral center of the z-Thr is indicated by an asterisk in Figure 4. The RESP charges of z-Thr were determined based on the single molecule in vacuo. We determined the atomic charges of the zinc ion and its surrounding residues, since the zinc ion can form a covalent bond with its surrounding residues. We adopted the RESP charge of zinc and its surrounding residues without the inhibitor. The RESP charges of His142, His146, Glu166 ,and the zinc ion were determined based on a super molecule consisting of Ace-His-NH_2, Ace-His-NH_2, Ace-Glu-NH_2, and the zinc. These atoms are superposed onto the crystal structure of thermolysin.

To check the reliability of the force field and our model, we compared the structures of the modeled region of the thermolysin with the z-L-Thr complex and with the original structure determined by X-ray, and the RMSDs of the backbone atoms and all heavy atoms including side chains were 0.13 and 0.13 Å, respectively. From the RMSD, our model is considered to be reasonable.

C. MD Simulation of the Thermolysin System. We applied the FP procedure to z-D/L-Thr with the two different parameter sets. The FP procedure was as follows.

Step 1: Preparatory MD simulation for equilibration of the system was performed. First, 50 ps MD simulation was given at 350 K, followed by 50 ps MD simulation at 310 K.

Step 2: The trajectory of landmark atom(s) was sampled by the 200 ps MD simulation at 310 K with the umbrella potential V_{FP}. For the first MD, $V_{rep} = 0$.

Step 3: The center of V_{FP} was shifted 0.3 Å to the last coordinate of the landmark atom. If the distance between the current coordinate and the last coordinate was less than 0.3 Å,

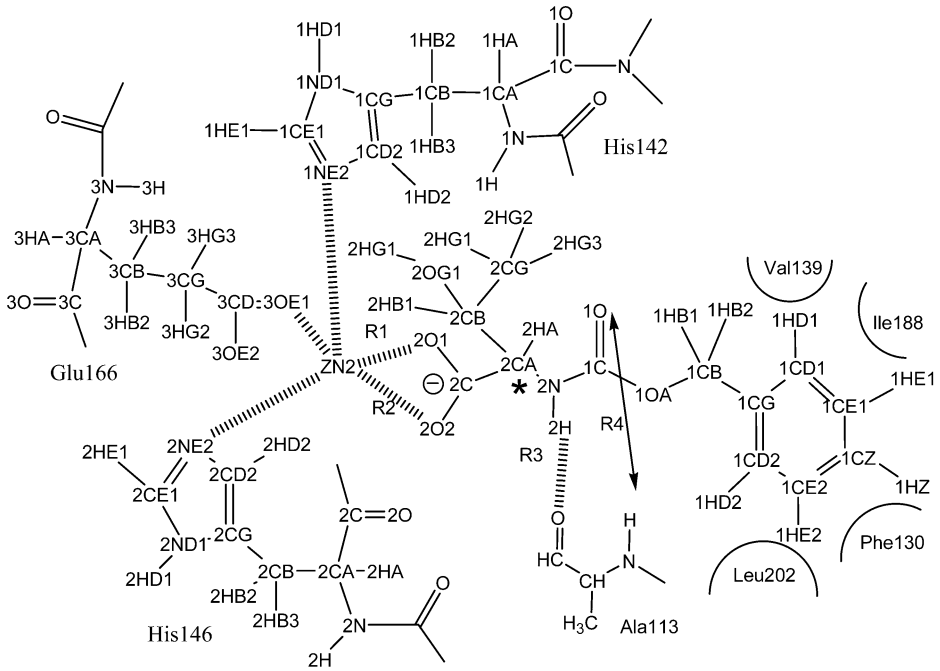

Figure 4. Ligand binding pocket and the inhibitor, z-D/L-Thr. Dashed lines represent the intermolecular interactions. The asterisk indicates the chiral center of z-D/L-Thr. R1 to R4 represent the interatomic distances between the interacting atom pairs.

TABLE 1: Atomic Charges of the Protein Model and z-D/L-Thr

residue	atom	charge	residue	atom	charge	residue	atom	charge
His142	N	−0.2214	His146	HE1	0.1337	z-Thr	C7	0.0766
His142	H	0.1998	His146	NE2	−0.6535	z-Thr	O8	−0.3909
His142	CA	−0.2664	His146	CD2	−0.0894	z-Thr	C9	0.9628
His142	HA	0.1560	His146	HD2	0.2361	z-Thr	O10	−0.6504
His142	CB	−0.0263	His146	C	0.7369	z-Thr	N11	−0.7759
His142	HB2	0.1018	His146	O	−0.5855	z-Thr	C12	0.0768
His142	HB3	0.0618	Glu166	N	−0.3944	z-Thr	C13	0.8670
His142	CG	0.0194	Glu166	H	0.3121	z-Thr	O14	−0.7861
His142	ND1	−0.2811	Glu166	CA	−0.2378	z-Thr	O15	−0.7861
His142	HD1	0.3831	Glu166	HA	0.0444	z-Thr	C16	0.3319
His142	CE1	0.2939	Glu166	CB	−0.0827	z-Thr	C17	−0.1471
His142	HE1	0.1838	Glu166	HB2	0.0593	z-Thr	O18	−0.6871
His142	NE2	−0.7674	Glu166	HB3	0.0242	z-Thr	H19	0.1425
His142	CD2	0.1658	Glu166	CG	0.0999	z-Thr	H20	0.1208
His142	HD2	0.0725	Glu166	HG2	−0.0412	z-Thr	H21	0.1208
His142	C	0.7611	Glu166	HG3	0.0274	z-Thr	H22	0.1425
His142	O	−0.6694	Glu166	CD	0.8433	z-Thr	H23	0.1239
His146	N	−0.3013	Glu166	OE1	−0.7043	z-Thr	H24	0.0444
His146	H	0.3054	Glu166	OE2	−0.8920	z-Thr	H25	0.0444
His146	CA	−0.2640	Glu166	C	0.8091	z-Thr	H26	0.3750
His146	HA	0.1675	Glu166	O	−0.6526	z-Thr	H27	0.0294
His146	CB	−0.4628	Zinc	ZN2	1.4609	z-Thr	H28	−0.0144
His146	HB2	0.1546	z-Thr	C1	−0.1833	z-Thr	H29	0.0245
His146	HB3	0.1890	z-Thr	C2	−0.1342	z-Thr	H30	0.0245
His146	CG	0.2919	z-Thr	C3	0.0908	z-Thr	H31	0.0245
His146	ND1	−0.2208	z-Thr	C4	−0.1342	z-Thr	H32	0.3848
His146	HD1	0.3846	z-Thr	C5	−0.1833			
His146	CE1	0.1349	z-Thr	C6	−0.1349			

the center of V_{FP} was set to the last coordinate of the landmark atom.

Step 4: V_{FP} was reconstructed. We then returned to step 1. This procedure was repeated until the ligand molecule reached the dissociated state. We switched the height of E_{rep} when the ligand dissociated. When the RMSD was <2.5 Å, c was 3 kcal/mol. When the RMSD was >2.5 Å and when the RMSD was <6 Å, c was 2.0 kcal/mol. Otherwise, c was 0.5 kcal/mol, k was 5.0 kcal/mol/Å^2, and D_0 was 0.3 Å. Two parameter sets were used for $a_x = 3.0$ and 2.0 Å^{-2}. The simulations with the former and the latter parameter sets were named Data1 and Data2, respectively.

Step 5: After steps 1−4 were completed, we applied WHAM to all of the trajectories.

This procedure required 100−150 iterations until the ligand reached the dissociated state. The landmark atom was an amide nitrogen atom of the ligand molecule, since the nitrogen atom is close to the center of mass of the ligand and the ligand includes only one nitrogen atom. The WHAM procedure converged after 500−1000 iterations, and the number of necessary iterations depended on the number of trajectories. Free energy F_i was calculated from the P of eq 5 averaged in a spherical region of 0.2 Å radius.

Panels a an db in Figure 5 show the energy trajectories of the 1st and 120th simulations of z-D-Thr and z-L-Thr. The temperature change from 350 to 310 K in Step 1 causes the potential energy jump at 50 ps. Neither the total potential energy nor the V_{FP} drift after 100 ps. Since the perturbation by V_{FP} is much smaller than the energy fluctuation, we consider that the systems were close to equilibrium within 100 ps.

The FP procedure constructed two dissociation reaction paths for each ligand molecule, and this method generated the PMFs for both the z-D-Thr and z-L-Thr molecules. Panels a−d in Figure 6 show the distribution of the trajectories of the landmark atom and the reaction paths of z-D-Thr and z-L-Thr in the 3D space. The initial coordinate (binding state) is around (X, Y, Z) = (8 Å, 2 Å, 6 Å) and the final dissociated state is around (X, Y, Z) = (16 Å, 5 Å, 10 Å). The reaction paths follow the first and the last coordinates of each simulation. As expected, the reaction paths appear to describe a random walk. The distribu-

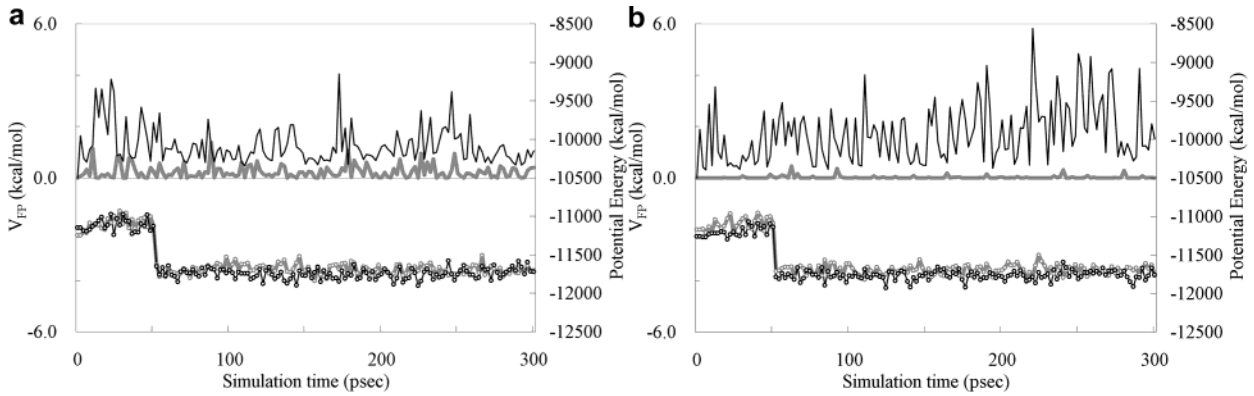

Figure 5. Energy trajectories of z-D-Thr (a) with the parameter set Data1 and z-L-Thr (b) with the parameter set Data1. Open circles and solid lines represent the total energy and the umbrella potential energy (V_{FP}), respectively. The black and gray lines represent the results of the 1st and the 120th simulation, respectively.

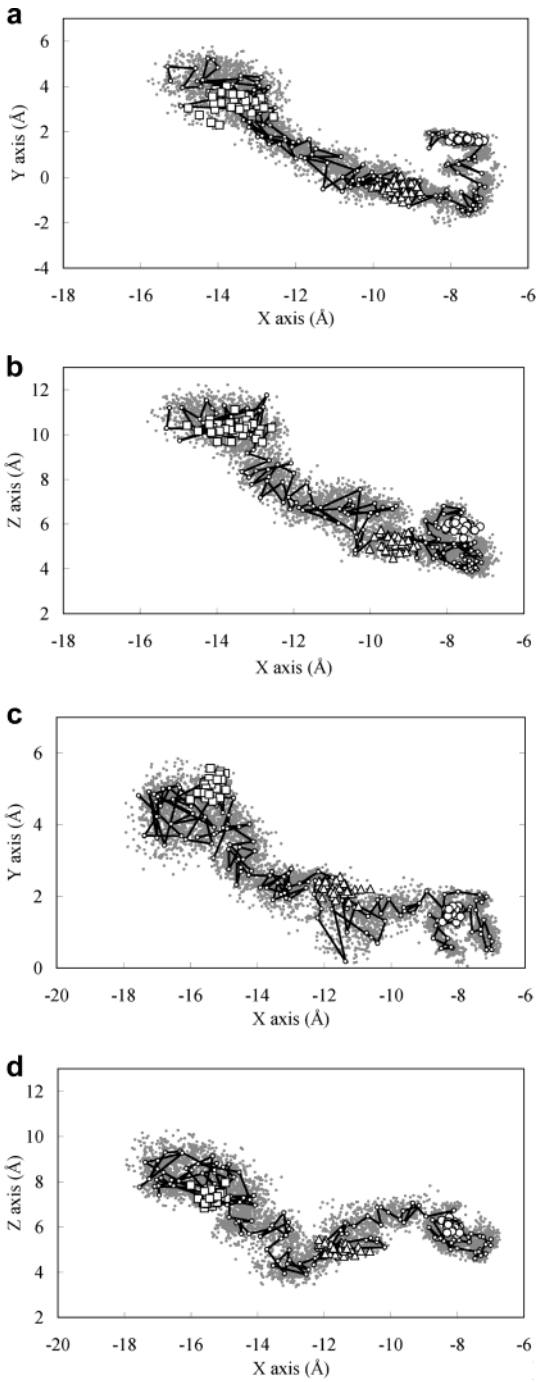

Figure 6. Trajectories and the reaction paths of z-D-Thr and z-L-Thr in 3D space. Axes are the 3D coordinates of the laboratory system. Gray dots, large open circles, triangles, and squares represent the total trajectories and the 1st, the 60th, and the 120th trajectories, respectively. Only 1% of the sampled points are depicted. The small circles connected by solid lines represent the reaction paths: (a and b) z-D-Thr with the Data1 parameter set; (c and d) z-L-Thr with the Data1 parameter set.

tions of trajectories cover the reaction paths. In Figure 6a, the overlap of the trajectories is relatively poor around $(X, Y) = (7$ Å, 0 Å). The number of actual data points is 100 times greater than the depicted data points, so that the overlap of trajectories is sufficient to converge the WHAM procedure. The distributions of the 1st, 60th, and 120th trajectories are localized in a small area of radius of ~1 Å. This is an advantage of the FP method over other global search methods, since the individual simulation times needed to realize the equilibrium can remain constant.

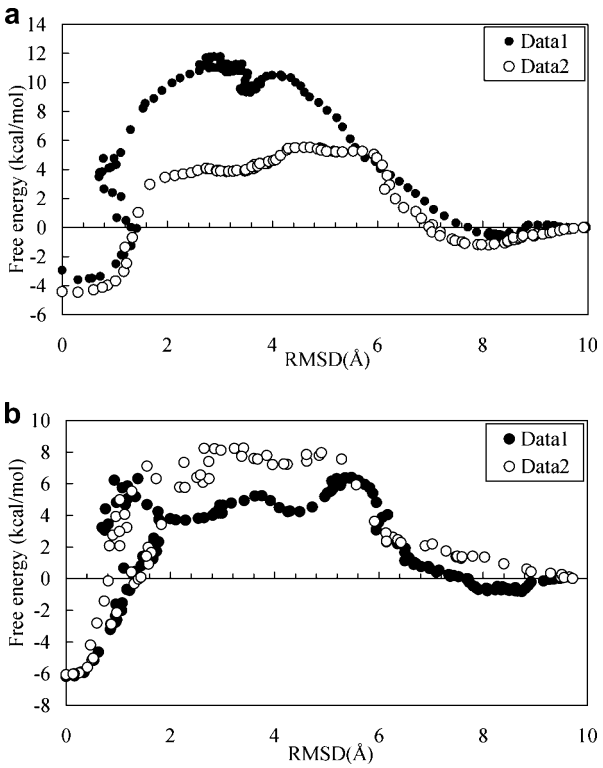

Figure 7. PMFs of z-D-Thr (a) and z-L-Thr (b). The filled and open circles correspond to the results obtained with the parameter sets Data1 and Data2, respectively (see text). RMSD (Å) is the distance between the nitrogen atom of the binding state and that of the final coordinate of each trajectory.

TABLE 2: Experimental and Calculated ΔG and $\Delta\Delta G$ in kcal/mol

	exptl	Data1	Data2	calcd (av)
z-D-Thr	−3.2	−2.9	−4.4	−3.7
z-L-Thr	−5.9	−6.1	−6.2	−6.1
$\Delta\Delta G$(L−D)	−2.7			−2.4

Panels a and b in Figure 7 show the PMFs of z-D-Thr and z-L-Thr, respectively. The free energy around the reaction path (F_i) is plotted against the RMSD. The RMSD is the distance between the landmark atom of the binding state and that of the final structure of each trajectory. The energy of the maximum RMSD was set to zero for each PMF. The initial structure (RMSD = 0 Å) was the energy minimum, which was the binding state. There was a high-energy barrier at RMSD = 2−6 Å, and the barrier heights were 10−16 kcal/mol for dissociation and 6−12 kcal/mol for adsorption. The shapes of the PMFs differed, since each of their reaction paths were different.

Table 2 summarizes the binding free energy of each system. The energy difference between the binding state and the dissociated state gives the binding free energy. The pressure and the volume of this system are unclear, since the CAP boundary condition is applied. We therefore assumed that the pressure dependence of ΔG is small and the calculated binding free energy corresponds to ΔG. In this case, the state around RMSD ~ 10 Å was defined as the dissociated state. The binding free energies agreed well with the experimental values, while the PMF profiles differed. The discrepancy between the calculated and the experimental values was less than 1 kcal/mol.

Around RMSD = 0−2 Å, the free energy was distributed over a wide energy range. For example, the energy was 2−10 kcal/mol at RMSD = 1 Å, as shown in Figure 7b. The free energy values at RMSD > 8 Å were distributed in the range of

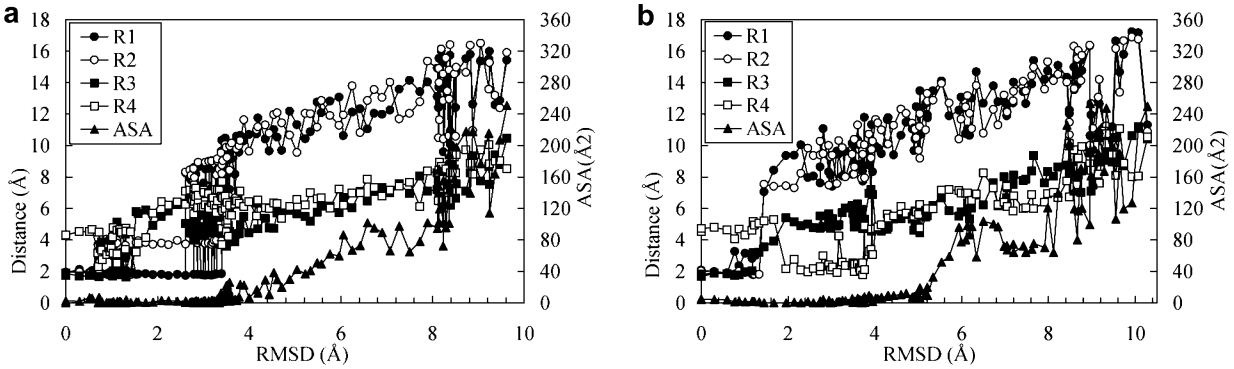

Figure 8. Change of interactions between thermolysin and z-D-Thr. Panels a and b correspond to the results obtained with the parameter sets Data1 and Data2, respectively. R1−R4 are defined as in Figure 3, and RMSD is defined as in Figure 4.

<1 kcal/mol. The reaction coordinates are the 3D coordinates of the landmark atom, so that the RMSD value cannot determine the complete 3D coordinates of all the ligand atoms. Thus, the relationship of the free energy and the RMSD is not a one-to-one correspondence. At the dissociated state, the free energy does not depend on its coordinate. Thus, ΔG is clearly determined by the energy difference between the binding state and the dissociated state, while the free energy of the intermediate state cannot be uniquely determined by RMSD values.

Using the FP procedure, we constructed two qualitatively different reaction paths for each ligand molecule. We monitored several ligand−protein interatomic distances and the surface area of the ligand to observe the intermolecular interactions in the dissociation process. In Figures 4, 6, and 7, the monitored distances were the distance between Zn^{2+} and an oxygen atom (2O1) of the ligand (R1), that between Zn^{2+} and the other oxygen atom (2O2) of the ligand (R2), that between the main chain oxygen atom of Ala113 and an amide hydrogen atom (2H) of the ligand (R3), and that between a main chain hydrogen atom of Ala113 and a carbonyl oxygen atom (1O) of the ligand (R4). The changes in these distances represent the formation and braking of the hydrophilic interactions. The observed surface area is the accessible surface area of the phenyl group and one CH_2 group (1CB, 1HB1, and 1HB2) of the ligand molecule. The change of this accessible surface area corresponds to the change of the hydrophobic interaction between the ligand and the protein.

Panels a and b in Figure 8 show the dissociation paths of Data1 and Data2 of z-D-Thr, respectively. In Figure 8a, R2 and R3 were broken at RMSD ∼ 1 Å, and then R1 began to break at RMSD ∼ 3 Å. After RMSD > 4 Å, the phenyl group of the ligand began to dissociate. The hydrophobic accessible surface increased by 249.66 Å^2. In Figure 8b, R1, R2, and R3 were broken at RMSD ∼ 1 Å, and the system formed a new bond R4 at RMSD = 2 Å. R4 was broken at RMSD ∼ 4 Å, and the phenyl group of the ligand dissociated after RMSD > 5 Å. The hydrophobic accessible surface increased by 250.94 Å^2. In Figure 8a, the PMFs derived by using the parameter sets Data1 and Data2 are shown, and their reaction paths are qualitatively different. We examined the reaction paths of z-L-Thr, and the hydrophilic interactions were broken first in both cases. The phenyl group was dissociated last. The reaction path depends on the parameters of the FP procedure, so we should note that these reaction paths are artificial. Under natural conditions, the ligand follows not one but numerous reaction paths.[20−22]

There are several problems in our calculation. The first is the limitation of size of our model system. Our model is a part of thermolysin with the CAP boundary condition, and we applied a position restraint on the protein backbone atoms. Therefore, our model cannot treat a large protein structural change by induced fitting. The second problem concerns the difference of solvent between our model and the experiment. Our solvent was pure water, while that of the experiment contained small ions. The third problem is the definition of the dissociated state in our simulation, i.e., RMSD = 10 Å is not the infinite distance. The fourth is the cutoff of 1−5 Coulomb interaction. And finally, there is a problem in regard to the classical treatment for the zinc-carboxyl group. In our model, the zinc−carboxyl interactions were Coulombic and van der Waals interactions. Under natural conditions, a covalent bond may occur between the zinc ion and the carboxyl group, which is neglected in our simulation. We will examine these potential drawbacks one by one.

In regard to the limitation of system size, we assumed no structural change of the backbone atoms of the protein model according to the ligand docking/dissociation process in this simulation. We compared the structures of the modeled region of the thermolysin with z-D/L-Thr complexes to the modeled region of the thermolysin without the ligand molecule (PDB code: 1L3F[47]), and the RMSDs of the backbone atoms were 0.48 and 0.59 Å, respectively. The backbone RMSD between the free thermolysin (1L3F) and the final coordinates of our simulations were 0.47−0.48 Å, and the backbone RMSD between the initial and final coordinates of our simulations were 0.12−0.13 Å. Thus, the structural change around the ligand pocket, which was modeled, is negligible.

With regard to the difference of the solvation structure, the definition of the calculated ΔG was different from that of the experimental ΔG. In the experimental condition, the ligands were dissolved in H_2O containing 10 mM $CaCl_2$ and 10 μM $ZnSO_4$, with a pH of 6.5. The current simulation did not precisely account for the pH and the density of solvent ions. In our calculation, the solvent was pure water. A specific zinc−Cl^- ion interaction may occur in the binding pocket when the ligand is absent. We tried a MD simulation for the protein model system using a Cl^- ion in place of z-Thr. We found that the zinc ion and Arg203 bind the Cl^- ion, and the Zn−Cl^- distance and the distance between the Cl^- and CZ atom of Arg203 were 4.76 and 4.48 Å, respectively. A water molecule was bridged between Zn^{2+} and Cl^-. We applied the FP method for a single Cl^- ion instead of a ligand, and we found that the Cl^- complex form was in a meta-stable state and the ΔG for Cl^- was approximately 1 kcal/mol. This value is much smaller than that of the ΔG of z-D/L-Thr. This also means that the hydrophobic interaction is dominant in ΔG of z-D/L-Thr rather than the electrostatic interaction.

We applied a cutoff of 1−5 Coulomb interaction, since the simulation costs 3−4 weeks when using a 4 CPUs MPI parallel

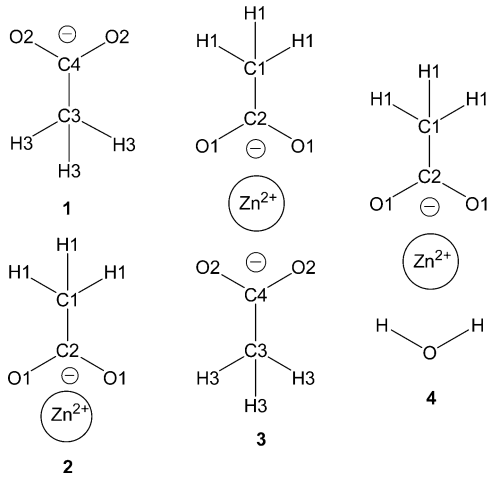

Figure 9. Model molecules for estimation of the Zn-carboxyl covalent interaction. Molecule **1** is an isolated CH_3COO^- molecule. Molecule **2** is a $Zn^{2+}(CH_3COO^-)$ super molecule. Molecule **3** is a $Zn^{2+}(CH_3COO^-)_2$ complex. Molecule **4** is a $Zn^{2+}(CH_3COO^-)H_2O$ complex.

calculation. Without cutoff, the cost is 10 times longer than that of the 12 Å cutoff simulation. Therefore, to validate our methodology, we must accept the artifact, which is derived from the cutoff. In our modeling, only 136 out of a total of 316 residues of the whole protein were used with the CAP water, which covers only a part of the protein model. Thus, in the case of our model, the perturbation of modeling itself could be larger than the artifact of the cutoff and the protonation of Asps and Glu. The solution structure of small ions, such as Na^+Cl^- in water, could be unreliable with a 1−5 Coulomb interaction cutoff. Therefore, the partial protonation of some Asps and Glu may be preferable to the use of counterions to neutralize the protein model with the Coulomb interaction cutoff.

We next examined the mode of the charge fitting and the potential-surface precision with respect to the Zn−O interaction. The Zn^{2+} ion can form a covalent bond with the carboxyl group of the z-Thr molecule in our thermolysin system. Also, the Zn^{2+} ion can form a covalent bond with an oxygen atom of a water molecule when the ligand dissociates from the binding site. Thus, the actual effect of the covalent bond on ΔG is the energy difference between the binding energy of Zn^{2+}−O of the ligand and that of Zn^{2+}−O of the water molecule.

The model systems were $Zn^{2+}(CH_3COO^-)_2$ and $Zn^{2+}(CH_3COO^-)H_2O$, and Figure 9 shows the schematic representation of these systems. We calculated the binding energy (ΔE) between the CH_3COO^- (molecule **1**) and $Zn^{2+}CH_3COO^-$ complex (molecule **2**) and that between H_2O and molecule **2**. Molecule **2** represents the zinc ion and Glu166, His142, and His146 in Figure 4. Molecule **1** represents the ligand molecule. This system is much simpler than the actual reaction center in Figure 4, which allowed us to make a more precise calculation.

ΔE values are defined as the energy difference between the energy of **3** and the sum of the energies of **1** and **2,** and also as the energy difference between the energy of **4** and the sum of the energies of **2** and H_2O. ΔE was estimated by both the quantum chemical calculation and the RESP charges. The discrepancy between ΔE by the quantum chemical calculation and ΔE by the RESP charges could be the covalent bond interaction. To determine a suitable method for deriving the RESP charge, we performed the derivation using several different strategies.

The RESP charges were determined based on both the Hartree−Fock (HF)/6-31G* level and the MP2/6-311+G** level.[43] Table 3 summarizes the RESP charges; the RESP1 and -2 charge sets are based on the HF wave function, and the RESP3 and -4 charge sets are based on the MP2 wave function. In the RESP1 and -3 sets, the charge of Zn was set to +2, and the two CH_3COO^- carried the same RESP charges for molecule **1**. The RESP2 and -4 charges were determined based on molecules **1** and **2**. The charge of H_2O was −0.7842 for the oxygen atom at the MP2 level. The charges of the TIP3P model were also used.[45]

TABLE 3: RESP Charges and ΔE in kcal/mol

	HF charge		MP2 charge	
	RESP1	RESP2	RESP3	RESP4
C1	−0.3321	−0.2124	−0.2928	−0.1921
H1	0.0359	0.1251	0.0191	0.1332
C2	0.9246	0.8331	0.9795	0.6932
O1	−0.8501	−0.7886	−0.8720	−0.6665
C3	−0.3321	−0.3321	−0.2928	−0.2928
H3	0.0359	0.0359	0.0191	0.0191
C4	0.9246	0.9246	0.9795	0.9795
O2	−0.8501	−0.8501	−0.8720	−0.8720
Zn	2.0000	1.5812	2.0000	1.4323
ΔE	255.00	203.20	255.71	192.42

We next estimated the binding energy (ΔE) by using energy minimization based on the AMBER parm99 force field with RESP1−6 charge sets. Table 3 summarizes the RESP charges and calculated ΔE values. In contrast, we estimated the quantum chemical ΔE by MP2/6-311++G**//MP2/6-311G**. ΔE determined by the MP2 method was 214.93 kcal/mol for molecule **3**. In any case, the electron donation from COO^- to Zn^{2+} forms a covalent bond between them, and this effect decreases the charge of the Zn^{2+} ion. Thus, the RESP1 and the RESP3 charge sets without electron transfer overestimated ΔE. The RESP2 charge set showed the best result. Also, ΔE from the MP2 method was 51.85 kcal/mol for molecule **4**. The ΔE values of molecule **4** from the AMBER force field with RESP charge were 41.64 kcal/mol from the RESP charge at the MP2 level and 48.86 kcal/mol from the HF level RESP charge with the TIP3P model.

The charge-fitting mode of the AMBER force field is based on the HF wave function. Thus it is reasonable to adopt the RESP2 charge-fitting mode on our thermolysin system; we used the RESP charges of the protein model with the zinc system and the RESP charges of isolated z-D/L-Thr, respectively. The total reaction was molecule **3** + H_2O → molecule **1** + molecule **4**, and the total reaction energy from the MP2 method was 163.08 kcal/mol. The total reaction energy from the RESP2 set with TIP3P was 154.33 kcal/mol. Thus, the discrepancy between the MP2 and the classical force field calculations was 8.74 kcal/mol for ΔE, which is 5% of the total reaction energy. In the water system, ΔG of the thermolysin system was 3∼6 kcal/mol, suggesting that the solvent effect reduces the reaction energy. While the HF charge is not very accurate, we can expect that the solvent effect reduces the inaccuracy of the force field as well as the inaccuracy of ΔG.

The hydrophobic accessible surface area increased by ∼250 Å^2 in the dissociation. When the atomic solvation parameter of carbon was ∼16 cal/Å^2/mol,[4,5] the free energy change by this effect was ca. −8 kcal/mol. This means that the hydrophobic interaction made the major contribution to ΔG, and the hydrophilic interaction was much smaller than the hydrophobic interaction.

These errors in the ΔG of z-D-Thr are similar to those in the ΔG of z-L-Thr. Since the solvation of the D compound must be the same as that of the L compound, the binding mode of z-D-Thr is qualitatively the same as that of z-L-Thr, and the force

field parameter of z-D-Thr is exactly the same as that of z-L-Thr. The energy difference between the L and D compounds ($\Delta\Delta G$) does not depend on these errors, and so the calculated $\Delta\Delta G$ can be compared with the experimental $\Delta\Delta G$. The simulated $\Delta\Delta G$ was in good agreement with the experimental values.

IV. Conclusion

We developed a method suitable for binding free energy estimation, called the filling potential (FP) method, based on the concept of the Taboo search. The FP method is an umbrella potential sampling method that enables the ligand molecule to drift from its local minima automatically. The umbrella potential is a combination of Gaussian-type repulsive potentials, which are located on the trajectory of the ligand. Without setting the reaction coordinates a priori, this method searches for and determines the suitable reaction coordinates by the successive generation of umbrella potentials based on its trajectory analysis. The WHAM of these trajectories gives the binding free energy.

This method was applied to a simple model of methane in water and a complex of thermolysin and its two inhibitors, z-D-Thr and z-L-Thr, in explicit water models. The FP method generated the reaction paths for the dissociation. There were several different paths, and the dissociation of the phenyl group of the ligand corresponded to the transition state. The calculated PMF for each ligand showed the free energy difference between the binding state and the dissociated state. The simulated binding free energy was in good agreement with the experimental ΔG. The computational discrepancy of $\Delta\Delta G$ was less than 1 kcal/mol.

Acknowledgment. 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. We thank Drs. M. Senda and T. Senda (JBIRC, AIST, Japan) for providing the crystal structures of thermolysin-inhibitor complexes, and Dr. S. Kidokoro (Nagaoka University of Technology, Japan) for the thermodynamic data. We thank Dr. Masaya Orita and Dr. Shigeo Fujita (Yamanouchi Pharmaceutical Co., Ltd.) for their helpful suggestions. Ms. Naoki Ihara assisted with one of the molecular topologies. We thank Mr. Takashi Kurosawa (Hitachi East Japan Solutions, Ltd.) for his technical support.

References and Notes

(1) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(2) Simonson, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243−252.
(3) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717−726.
(4) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199−203.
(5) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086−3091.
(6) Calimet, N.; Schaefer, M.; Simonson, T. *Proteins*: *Struct., Funct., Genet.* **2001**, *45*, 144−158.
(7) Fukunishi, Y.; Suzuki, M. *J. Phys. Chem.* **1996**, *100*, 5634−5636.
(8) Hummer, G.; Garde, S.; Garcia, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951−8956.
(9) Rashin, A. A. *J. Phys. Chem.* **1989**, *93*, 4664−4669.
(10) Bash, P.; Singh, C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574−576.
(11) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564−568.
(12) Sneddon, S. F.; Tobias, D. J.; Brooks, C. L., III *J. Mol. Biol.* **1989**, *216*, 817−820.
(13) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.
(14) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *92*, 18−431.
(15) Kong, X.; Brooks, C. L., III *J. Chem. Phys.* **1996**, *105*, 2414−2423.
(16) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: New York, 1987.
(17) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins*; Wiley: New York, 1988.
(18) Mangoni, R.; Roccatano, D.; Di Nola, A. *Proteins* **1999**, *35*, 153−162.
(19) Di Nola, A.; Roccatano, D.; Berendsen, H. J. C. *Proteins* **1994**, *19*, 174−182.
(20) Heymann, B.; Grubmüller, H. *Biophys. J.* **2001**, *61*, 1295−1313.
(21) Heymann, B.; Grubmüller, H. *Phys. Rev. Lett.* **2000**, *84*, 6126−6129.
(22) Grubmüller, H.; Heymann, B.; Tavan, P. *Science* **1996**, *271*, 997−999.
(23) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249−253.
(24) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9−12.
(25) Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. *Chem. Phys. Lett.* **1997**, *278*, 297−301.
(26) Andricioaei, I.; Straub, J. E. *Phys. Rev. E* **1996**, *53*, 3055−3058.
(27) Pak, Y.; Wang, S. *J. Phys. Chem. B* **2000**, *104*, 354−359.
(28) Fukuda, I.; Nakamura, H. *Phys. Rev. E* **2002**, *65*, 26105−26109.
(29) Cvijovic, D.; Klinowski, J. *Science* **1995**, *267*, 664−666.
(30) Huber, T.; Torda, A. E.; van Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **1994**, 8, 695−708.
(31) Ge, R. P.; Qin, Y. F. *JOTA* **1987**, *54*, 241−252.
(32) Kumar, S.; Bouzida, D.; Swndsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Compt. Chem.* **1992**, *13*, 1011−1021.
(33) Kumar, S.; Payne, P. W.; Vasquez, M. *J. Compt. Chem.* **1996**, *17*, 1269−1275.
(34) Verkhivker, G. M.; Rejto, P. A.; Bouzida, D.; Arthurs, S.; Colson, A. B.; Freer, S. T.; Gehlhaar, D. K.; Larson, V.; Luty, B. A.; Marrone, T.; Rose, P. W. *J. Mol. Recognit.* **1999**, *12*, 371−389.
(35) Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893−2906.
(36) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908−3911.
(37) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665−4677.
(38) Morikami, K.; Nakai, T.; Kidera, A.; Saito, M.; Nakamura, H. *Comput. Chem.* **1992**, *16*, 243−248.
(39) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(40) Àqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(41) Darden, T.; Pearlman, D.; Pedersen, L. G. *J. Chem. Phys.*, **1998**, *109*, 10921−10935.
(42) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. J.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(43) 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. M.; 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.; 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.
(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(45) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(46) Payne, V. A.; Matubayasi, N.; Murphy, L. R.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 2054−2060.
(47) Hausrath, A. C.; Matthews, B. W. *Acta Crystallogr., Sect. D* **2002**, *58*, 1002−1007.