JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Prediction of Protein−Ligand Binding Structures by Replica-Exchange Umbrella Sampling Simulations: Application to Kinase Systems

Hironori Kokubo,*[,†] Toshimasa Tanaka,[†] and Yuko Okamoto[‡,§,⊥,¶]

[†]Medicinal Chemistry Research Laboratories, Pharmaceutical Research Division, Takeda Pharmaceutical, 26-1 Muraoka-Higashi 2-Chome, Fujisawa, Kanagawa 251-8585, Japan

[‡]Department of Physics, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan

[§]Structural Biology Research Center, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan

[⊥]Center for Computational Science, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603, Japan

[¶]Information Technology Center, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

ABSTRACT: We have applied our prediction method, which is based on the replica-exchange umbrella sampling for protein−ligand binding structures, to two kinase systems (p38 and JNK3) with two different ligand molecules for each kinase. Starting from configurations in which the protein and the ligand are far away from each other, our method predicted the ligand binding structures in excellent agreement with the experimental data from PDB in all four cases, which suggests the general applicability of our method to kinase systems. In addition, the protein flexibility was shown to be essential to predict the correct binding structure for one of the systems, where dihydroquinolinone was bound to p38 alpha kinase (PDB ID: 1OVE).

INTRODUCTION

A large number of protein−ligand complex structures are registered in the Protein Data Bank (PDB), and there are more than 18 000 such structures presently.[1] The information on protein−ligand binding structures makes it possible to understand the interactions between a ligand and a protein in atomistic detail, and it is very useful for drug design in pharmaceutical research. However, there are many classes of proteins for which crystallization is difficult, such as proteins with highly flexible loops, intrinsically disordered proteins, and membrane proteins.

Protein−ligand complex structures are also helpful for binding affinity predictions. Prediction of the binding affinity of a ligand is now widely used in computational biophysics and chemistry.[2] Various methods are used, such as simple empirical scoring methods (docking scores), the more sophisticated MM/PBSA method,[3] and the most rigorous but computationally expensive double decoupling method.[4,5] Accurate binding structures of the ligand to proteins are necessary as input information for affinity prediction by these methods. When a particular water molecule plays an important role in protein−ligand binding by bridging hydrogen bonds between a ligand and a protein or metal ions, an explicit water molecule model should be used. However, it is currently unavailable by most docking software.

A molecular dynamics (MD) method is a powerful tool to investigate various biological processes in atomistic detail, but the application to the prediction of protein−ligand binding structures has been believed to be difficult due to the required huge computational cost. However, large-scale and long-time simulations have been getting more practical recently by the advent of supercomputers such as TSUBAME, K computer, Sequoia,[6] and a special-purpose computer ANTON.[7] Dror et al. performed long-time MD simulations of the ligand binding to G protein-coupled receptors and examined the mechanism by using ANTON.[8]

We have recently developed a prediction method[9] of protein−ligand binding structures by replica-exchange umbrella sampling method (REUS),[10] which is one of the generalized-ensemble algorithms that greatly enhances conformational sampling (for reviews, see refs 11−13, for example). Different from most existing docking software, our method takes into account protein flexibility, water molecules in atomistic detail, and entropy and predicts a binding structure from the global

Received: May 27, 2013
Published: September 3, 2013

   dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

minimum free energy state. Although our method was hitherto successfully applied to five protein−ligand systems,[9] the applicability was not fully investigated.

In this article, we investigate how our approach to predict ligand binding structures can be applied to kinase systems. The kinase family is inarguably one of the most attractive drug targets because it regulates various biological reactions. Most protein kinase inhibitors target ATP binding sites because they generally exist in deep pockets and it is easy to acquire high potency. It is well known that flexibility plays an essential role in protein−ligand binding.[14,15] A kinase domain takes ON (active) and OFF (inactive) conformations by the open-close motion of N-terminal and C-terminal lobes. Available protein−ligand complex structures from the PDB also show that the flexibility of the A-loop, C-loop, and DFG motif is important for protein−ligand binding. We have performed the REUS simulations for kinase p38 and JNK3 with different ligands (dihydroquinolinone and 4-azaindole for p38; dihydroanthrapyrazole and fluorobenzamide for JNK3). The importance of the protein flexibility for protein−ligand binding is also examined by performing the comparative REUS simulation with the restraints on the heavy atoms of the receptor and by performing the docking by one of the typical docking software.

The rest of the article is organized as follows. In the second section, we summarize the details of our method. In the third section, we provide the computational details. In the fourth section, we present the results and discussion about the applications of our method to the two kinase systems as test cases. The fifth section is devoted to conclusions.

METHODS

We first briefly review our method for predicting protein−ligand binding structures (see ref 9 for details). Our methods consist of the following three steps:

In the first step, we first determine the charge and other force field parameters for the ligand molecules. The details of this force field parameter determination will be given in the next section. Next, we immerse each protein−ligand complex in a water box; each system is neutralized by adding the appropriate ions for charged complexes. Here, we define a reaction coordinate ξ as the distance between the center of mass of the ligand molecule and the center of mass of the backbone heavy atoms of two selected residues of the protein. Two residues of proteins near the ligand binding sites are arbitrarily selected. We then perform an REUS simulation. Although the information of the ligand binding sites is necessary for our method, it is often known for kinase targets by experiments and theoretical considerations. If there are multiple possible binding sites as in, for example, the adenine binding site and back pocket for a ligand, however, multiple independent REUS simulations by assuming different putative binding sites should be performed.

We now briefly describe the REUS method for completeness, focusing on our application, although the general formalism of REUS has already been explained elsewhere.[10] We consider M noninteracting replicas of the original system with N atoms. We can write the Hamiltonian for the ith replica with atomic coordinates $q^{[i]}=(\boldsymbol{q}^{[i]}_1, ..., \boldsymbol{q}^{[i]}_N)$ and momenta $p^{[i]}=(\boldsymbol{p}^{[i]}_1, ..., \boldsymbol{p}^{[i]}_N)$ as

$$H_m(q^{[i]}, p^{[i]}) = K(p^{[i]}) + E_{\lambda_m}(q^{[i]}) \tag{1}$$

where K is the kinetic energy and the potential energy E_{λ_m} with the original potential E_0 and umbrella potential $V_m(q^{[i]})$ can be written as

$$E_{\lambda_m}(q^{[i]}) = E_0(q^{[i]}) + V_m(q^{[i]}) \qquad \text{for } (m = 1, ..., M) \tag{2}$$

We use the following harmonic, restraining potential (referred to as umbrella potential) for $V_m(q^{[i]})$ by using the reaction coordinate ξ, which represents the distance between the ligand and the binding site:

$$V_m(q) = k_m(\xi(q) - d_m)^2 \tag{3}$$

where d_m is the midpoint distance and k_m is the strength of the restraining potentials. Without loss of generality, we can assume that the d_m values are in an increasing order, i.e., $d_1 < d_2 < ...< d_M$. The replicas are chosen so that the potential energy for each replica includes exactly one umbrella potential, namely replica i ($i = 1, ..., M$), and umbrella potential $V_m(q^{[i]})$, where ($m = 1, ..., M$), are in one-to-one correspondence. Because the replicas are noninteracting, the weight factor for state X in this generalized ensemble is given by the product of the Boltzmann factors for each replica

$$\begin{aligned} W_{\mathrm{Mrem}}(X) &= \prod_{i=1}^{M} \exp\{-\beta H_{m(i)}(q^{[i]}, p^{[i]})\} \\ &= \exp\{-\sum_{i=1}^{M} \beta H_{m(i)}(q^{[i]}, p^{[i]})\} \end{aligned} \tag{4}$$

where $m(i)$ is the permutation function that connects replica number i to label m of the umbrella potential.

Suppose that replicas i and j have the neighboring umbrella potentials V_m and V_{m+1}, respectively, in state X. We have the following transition probability of replica exchange from state X to state X' (w), where umbrella potentials V_m and V_{m+1} are exchanged between replicas i and j in state X':

$$w(X \to X') \equiv w(x^{[i]}_m, x^{[j]}_{m+1}) = \begin{cases} 1 & \text{for } \Delta \le 0, \\ \exp(-\Delta) & \text{for } \Delta > 0, \end{cases} \tag{5}$$

where

$$\Delta = \beta(V_m(q^{[j]}) - V_m(q^{[i]}) - V_{m+1}(q^{[j]}) + V_{m+1}(q^{[i]})) \tag{6}$$

Note that we need to newly evaluate the umbrella potential energy $V_{m+1}(q^{[i]})$ and $V_m(q^{[j]})$ with the exchanged coordinates.

We remark that when a sufficient random walk in the reaction coordinate space cannot be realized, we can further enhance conformational sampling by performing a two-dimensional REUS simulation in which the replica exchange in both temperature and umbrella potential is performed.[10] Achieving the thermal tempering and umbrella exchanging in one-dimensional space may be also useful.[16]

In the second step, we obtain the value of the reaction coordinate ξ that corresponds to the global minimum in potential of mean force (PMF). The weighted histogram analysis method (WHAM)[17] or its variant, the multistate Bennett acceptance ratio estimator (MBAR),[18,19] is suitable for obtaining the canonical distributions. Here, we use MBAR. Suppose that we have made a single run of the REUS simulation with M replicas that correspond to M different umbrella potentials V_m ($m = 1, ..., M$). Let β and N_m be $1/k_BT$,

dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

Table 1. Properties of the Four Ligand Molecules Examined in the Present Study.[a]

ligand	weight	NumrotB	Numheavy	HBacc	HBdon	MOE log*P*	PDB ID	affinity	protein
dihydroquinolinone	488.4	3	33	1	0	6.72	1OVE	0.74	p38
4-azaindole	305.3	2	23	3	1	3.50	1OZ1	6.5	p38
dihydroanthrapyrazole	220.2	0	17	3	2	3.40	1PMV	150	JNK3
fluorobenzamide	300.4	4	21	2	1	3.46	2O2U	3000	JNK3

[a]Weight, molecular weight; NumrotB, number of rotatable bonds; Numheavy, number of heavy atoms; HBacc, number of hydrogen bond acceptor sites; HBdon, number of hydrogen bond donor sites; MOE log*P*, calculated log octanol/water partition coefficient (*P*) by MOE; PDB ID, Protein Data Bank ID; affinity, experimental binding affinity [nM]; protein, name of the protein that the ligand binds to.

where k_B is the Boltzmann constant and T is absolute temperature, and the total number of samples obtained with the mth umbrella potential, respectively. The MBAR equation can then be written as

$$f_m = -\ln\left[\sum_{m'=1}^{M}\sum_{k=1}^{n_{m'}} \frac{\exp(-\beta E_m(q_{m'}(k)))}{\sum_{m''=1}^{M} n_{m''}\exp(f_{m''} - \beta E_{m''}(q_{m'}(k)))}\right] \quad (7)$$

where $q_{m'}(k)$ is the kth coordinate data in the MD trajectory obtained with $V_{m'}$. The dimensionless Helmholtz free energy f_m in the above equation are solved self-consistently by iteration.[18,19] For instance, all f_m are set to zero first and inserted into the right-hand side of eq 7 to obtain new f_m. Next, these new f_m are substituted again into the right-hand side of eq 7, and f_m are further updated, and so on. This iteration is continued until all the values of f_m converge. The PMF of the original, unbiased system without the umbrella potential is given by

$$W_{\lambda=\{0\}}(\xi) = -k_B T \ln[P_{\lambda=\{0\}}(\xi)] \quad (8)$$

where

$$P_{\lambda=\{0\}}(\xi) = \sum_{m=1}^{M}\sum_{k=1}^{n_m} \frac{\exp(-\beta E_0(q_m(k)))}{\sum_{m'=1}^{M} n_{m'}\exp(f_{m'} - \beta E_{m'}(q_m(k)))} = \sum_{m=1}^{M}\sum_{k=1}^{n_m} \frac{1}{\sum_{m'=1}^{M} n_{m'}\exp(f_{m'} - \beta V_{m'}(q_m(k)))} \quad (9)$$

Here, $\lambda = \{0\}$ means that the system is unbiased, and the summation over k is taken only for the coordinate $q_m(k)$ with the reaction coordinate value ξ (namely, $q_m(k)$ contributes to the summation only when $\xi \le \xi(q_m(k)) < \xi + \delta\xi$ for a small interval $\delta\xi$). We can then determine the distance ξ that corresponds to the global minimum of the PMF $W_{\lambda=\{0\}}(\xi)$.

We also used the standard WHAM for comparison. MBAR has the advantage over WHAM, especially when we need to treat multidimensional space, because histograms of energy distributions are not required for MBAR. We confirmed that the PMF obtained by MBAR and WHAM was essentially identical, although the number of required iterations in MBAR was much smaller than that of WHAM.

In the third step, we obtain the predicted binding structures of the ligand molecules employing the principal component analysis (PCA).[20−25] PCA is an effective method for classifying the conformations of a molecule. First, n conformations, which have a reaction coordinate ξ that is close to the value corresponding to the global minimum of the PMF $W_{\lambda=\{0\}}(\xi)$ and which have the potential energy (excluding the umbrella potential) close to the average (unbiased) potential energy, are extracted from the REUS simulations. We calculated the following variance-covariance matrix C_{ij}:

$$C_{ij} = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle \quad (10)$$

where $(q_1, q_2, q_3, ..., q_{3N_L-2}, q_{3N_L-1}, q_{3N_L}) = (x_1, y_1, z_1, ..., x_{N_L}, y_{N_L}, z_{N_L})$, $\langle q_i \rangle = \sum_{i=1}^{n} q_i/n$, x_i, y_i, and z_i are the Cartesian coordinates of the ith atom, N_L is the total number of ligand atoms, and n is the total number of MD trajectory samples. Note that we do not superimpose each structure onto a reference structure because our purpose is to predict ligand binding structures and we need to analyze differences not only in the ligand conformations but also in the ligand positions and orientations. This symmetric matrix is diagonalized, and the eigenvalues and eigenvectors are obtained. The first and the second principal component axes are defined as the eigenvectors with the largest and the second-largest eigenvalues, respectively. The ith principal component μ_i of each sampled structure is defined by the inner product

$$\mu_i = \vec{v}_i \cdot (\vec{q} - \langle \vec{q} \rangle) \quad (11)$$

where $\vec{v}_i$ is the ith eigenvector, $\vec{q} = (q_1, q_2, q_3, ..., q_{3N_L-2}, q_{3N_L-1}, q_{3N_L})$, and $\langle \vec{q} \rangle = \sum_{i=1}^{n} \vec{q}/n$.

We analyzed the free energy landscapes for the first and the second principal component axes and used the following equation to calculate the free energy as a function of two principal components, μ_1 and μ_2:

$$F(\mu_1, \mu_2) = -k_B T \ln P(\mu_1, \mu_2) \quad (12)$$

where $P(\mu_1, \mu_2)$ is the probability of finding the structure with the principal component values μ_1 and μ_2. We selected a typical structure that exists in the global minimum free energy state on the free energy landscape. This is our predicted binding structure.

COMPUTATIONAL DETAILS

We studied two kinase proteins, p38 and JNK3, with two ligands for each protein. We selected four protein−ligand complexes from the Protein Data Bank (PDB): 1OVE, 1OZ1, 1PMV, and 2O2U. Table 1 lists basic properties of the four ligand molecules with the information for the receptor proteins. We see that these molecules generally have typical drug-like properties and obey the well known rule of five;[26] that is, most orally administered drugs have a molecular weight of 500 or less, an octanol−water partition coefficient log *P* no higher than 5 (except for the case of 1OVE), 5 or fewer hydrogen bond donor sites, and 10 or fewer hydrogen donor sites. Figure 1 shows the structures of the four ligand molecules complexed to proteins, as registered in the PDB.

The charge parameters of the ligand molecules were determined by the RESP method.[28,29] These parameters

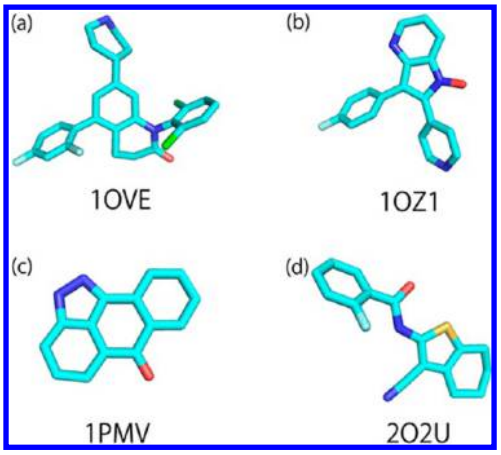

Figure 1. Ligand molecule conformations complexed to proteins (from PDB). PDB IDs are (a) 1OVE, (b) 1OZ1, (c) 1PMV, and (d) 2O2U. The figures were created using PyMOL.[26]

generally depend on the input ligand conformations. The important factor for molecular simulations is the extent to which the ligand free energy landscapes are sensitively influenced. We have recently examined the dependency of ligand free energy landscapes on charge parameters and solvent models in detail.[30] Our results have shown that ligand free energy landscapes do not depend on input conformations very much in solution, although they do depend strongly on them under a vacuum condition. Hence, the input ligand conformations for charge determination were taken from the Protein Data Bank (PDB)[1] files. The RESP charges were calculated by an ab initio molecular orbital program at the HF/6-31G* level and RESP utilities in the antechamber program suite.[31] The molecular mechanics parameters by GAFF were employed for bond, angle, torsion, and Lennard-Jones parameters for ligand molecules.[31,32] The missing force field parameters were assigned with the *parmcheck* utility in AMBER.[33] We employed AMBER *ff99SB* force field parameters[34] for proteins and the TIP3P model[35] for water.

Next, we immersed each protein−ligand complex in a TIP3P water rectangular parallelepiped box with a solvation of at least 10.0 Å from the protein surface. The total number of water molecules was 4634 for the 1OVE system, 4705 for the 1OZ1 system, 4113 for the1PMV system, and 4657 for the 2O2U system. The systems were neutralized by adding sodium or chloride ions for the charged complexes. That is, for the 1OVE system, we added no ions; for the 1OZ1 system, we added one sodium ion; for the 1PMV system, we added three chloride ions; for the 2O2U system, we added three chloride ions. The periodic boundary conditions and the particle mesh Ewald method[36] were employed for all the simulations.

We performed energy minimizations of the protein−ligand complexes with PDB conformations by the steepest descent and conjugate gradient method (in 500 steps for each minimization) with all the heavy atoms of the protein and the ligand subjected to harmonic restraints with a bond constant of 10 kcal/(mol Å^2). We then performed energy minimization without any restraints for 1000 steps. Next, we carried out equilibration MD simulations for 1.2 ns at temperature 300 K and pressure 1.0 atm. We used harmonic restraints on heavy atoms of the protein and the ligand only for the initial 0.1 ns of these equilibration runs. The Berendsen thermostat was employed in the simulations for temperature

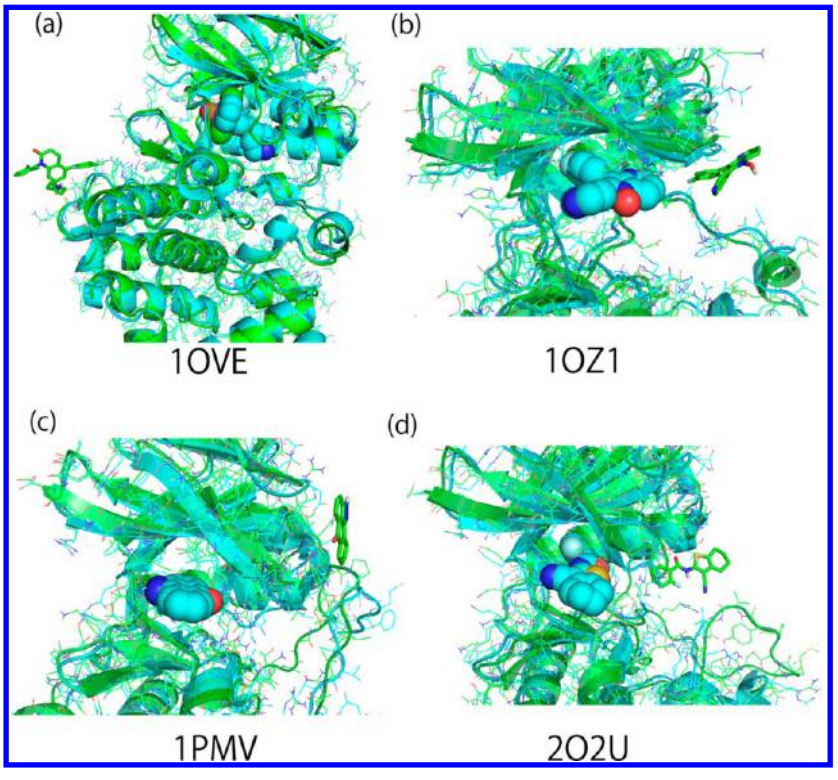

Figure 2. Snapshots from the umbrella sampling simulations with the largest midpoint distance value of d_m = 25 Å (green) superimposed on the experimental structures from PDB (blue). PDB IDs are shown below each figure. The ligands are in stick representation. The space-filled ligand molecules from PDB are also shown to indicate the correct ligand binding positions. Water molecules are not shown so that the protein−ligand conformations can be seen clearly. The figures were created using PyMOL.[27]

dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

control at 300 K, and the pressure was controlled by the weak-coupling method at 1.0 atm.[37] The SHAKE method[38] was employed to constrain bond lengths involving hydrogen atoms, and the time step was 2.0 fs. We used the equilibrated structure of each system as the reference structure to define the harmonic restraints on the heavy atoms in the protein that were far away from the ligand. Namely, the heavy atoms that were more than 12 Å away from the ligand in the equilibrated structure were under the harmonic restraints with the strength k = 1.0 kcal/(mol Å^2). These restraints were used for all the simulations discussed below. Note that we allowed protein structures to be flexible for the heavy atoms that were within 12 Å from the ligands in the equilibrated structure. The ligand binding structures were stable during the equilibration runs for all four systems.

The reaction coordinate ξ for each system was defined as the distance between the center of mass of the ligand molecule and the center of mass of the backbone heavy atoms of two selected residues of the protein. Two residues of proteins around the ligand binding pockets were arbitrarily selected: Met109 and Gly110 for 1OVE, Met109 and Gly110 for 1OZ1, Leu148 and Met149 for 1PMV, and Leu148 and Met149 for 2O2U. The REUS method enhanced the sampling along the defined reaction coordinate, and hence, the existence of the correct binding position on the path often sampled during the REUS simulations was required. Therefore, we selected two residues that exist around the binding pocket for each system. Note that the ligand does not necessarily need to make close contact with selected two residues. The same predicted binding structure should be obtained as far as the correct binding structure exists along the ligand-approaching paths.

The following 26 umbrella potentials were prepared: $k_m(\xi(q)-d_m)^2$, where the midpoint distance values were d_m = 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 10.5, 11.0, 11.5, 12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.5, 22.0, 23.5, and 25.0 Å and the strength of the restraining potentials was k_m = 1.0 kcal/(mol Å^2) for $d_m \leq$ 13.0 Å and k_m = 0.5 kcal/(mol Å^2) for $d_m >$ 13.0 Å. These umbrella potentials were used not only in the REUS production runs but also in the preparation umbrella sampling[39] and REUS simulations. These parameters were chosen so that we have sufficient replica exchanges. The midpoint distance of the first umbrella potential was 5.0 Å, which is sufficiently small when the Lennard-Jones repulsion volume is considered. The same number of replicas (26) was used for all the REUS simulations. We observed that the midpoint distance of 25.0 Å for the 26th umbrella potential was sufficiently large to ensure that ligands were away from the binding pocket and are in solvent regions as shown in Figure 2. The data for the first 20 ns of the REUS run was not used for the analyses because of the dependency on the initial structures. The dependency was judged by looking at the time series of the root-mean-square distance (RMSD) of the ligand molecules from the correct binding structures at the same umbrella potential. Note that we here calculated RMSD without translations and rotations.

The initial conformations for the 26 replicas in the REUS simulations were prepared as follows. Starting from the equilibrated ligand-bound structures, which were obtained by the equilibration simulations of the PDB structures described above, the ligands were gradually pushed away from the protein pocket along the reaction coordinate ξ until ξ became sufficiently large, at which distance the ligands were thoroughly surrounded by the solvent water. This process was achieved by using the umbrella sampling method[36] with the harmonic umbrella potentials as defined above. The umbrella potential, whose midpoint distance was the closest to the value of the reaction coordinate obtained after the equilibration run, was first selected as V_{m_0}, and the umbrella sampling simulation using V_{m_0} was performed. We then sequentially performed the umbrella sampling simulations by using the umbrella potential with the neighboring midpoint (i.e., V_{m_0+1}). We continued this process until the midpoint of the umbrella potential became sufficiently large.

Figure 2 shows the typical snapshots from the umbrella sampling simulations with the largest midpoint value of d_m = 25 Å. We see that ligand molecules were away from proteins and were completely outside the protein pockets. As references, the correct binding positions from PDB are also shown as space-filled molecules, which did not actually exist in the simulations. We also made sure that the ligand never went beyond the periodic boundary box during the simulations by using the umbrella potential with the largest midpoint. It is because (1) our unit cells of the simulations were set so that we have the water solvation of at least 10 Å from the protein edge on each direction, (2) the ligand binding sites existed relatively deeply inside the protein (typically 10 Å inside from the protein edge), and (3) the ligand had the steric repulsion from the proteins until at least ξ = 5 Å (see Figure 5 below). Therefore, the reaction coordinates could exist in the unit cell up to roughly 25 Å (= 10 + 10 + 5 Å).

Next, we sequentially performed the umbrella sampling simulations in the reverse order starting from the structures in Figure 2. That is, the reaction coordinate was gradually decreased by using umbrella potentials with relatively small midpoints. Each umbrella sampling simulation was performed for 100 ps, and the 26 final conformations obtained from the consecutive umbrella sampling simulations, starting from the restraining potential with d_m = 25 Å to that with d_m = 5 Å, were used as the initial structures for the 26 replicas in the REUS simulations. We found that the correct binding structures were rarely found in this process even though the simulations included the correct reaction coordinate distances.

The production runs were performed for 220−240 ns for each replica, and the data from the last 100 ns were used for the analyses. The data of the first 20 ns data were not used for the analyses because of the dependency on the initial structures. The dependency was judged by looking at the time series of the RMSD without translations and rotations of the ligand molecules from the correct binding structures at the same umbrella potential. For 1OVE and 1OZ1, 55 000 conformations and for 1PMV and 2O2U, 60 000 conformations at each umbrella potential were collected for the analyses of PMF, which were extracted from the trajectory files.

In order to obtain the predicted binding structures of ligand molecules from PMF, we used principal component analysis (PCA). PCA is an effective method to classify the conformations of a molecule. Around 10 000 conformations for each system, which have the reaction coordinates close to the global minimum PMF (with the range of ±0.2 Å) and the potential energy (excluding umbrella potential) close to the average potential energy (with the range of ±10.0 kcal/mol), were extracted from the REUS simulations. We predicted the ligand binding structures from the typical structures that existed in the bin of the global minimum free energy state on the free energy landscapes.

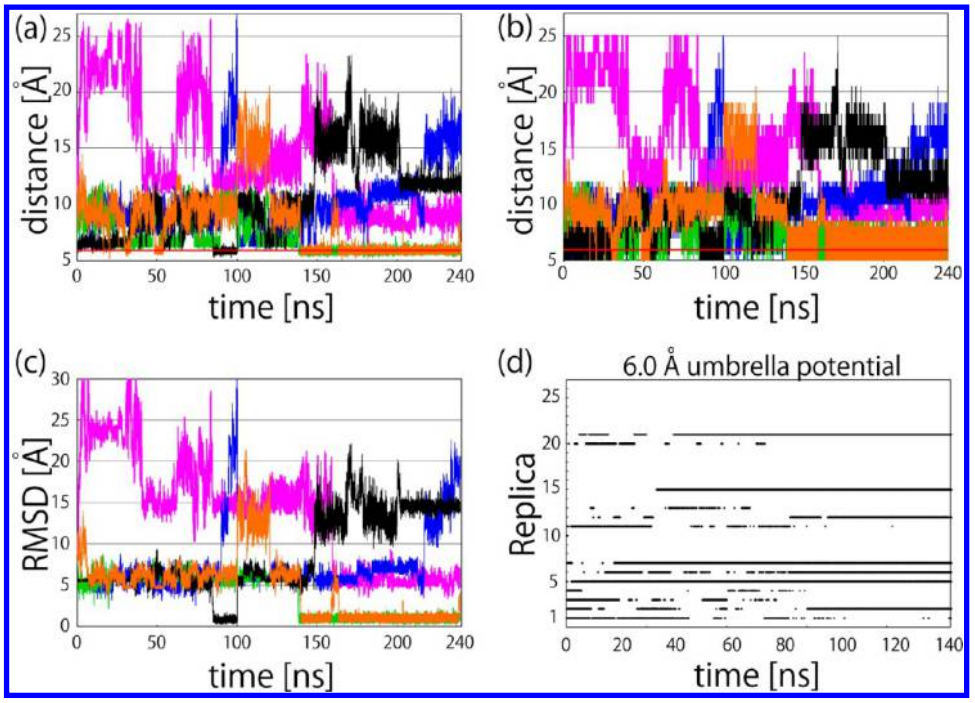

Figure 3. Time series of (a) the reaction coordinate ξ (the distances from the protein pocket), (b) the midpoint of umbrella potential from the protein pocket, (c) RMSD (in Å) from the correct binding mode (a blue ligand in Figure 6c below), and (d) replica exchange for the third umbrella potential (k_3 = 1.0 kcal/(mol Å^2) and d_3 = 6.0 Å) obtained from the REUS simulation of the 1PMV system. Time series of replicas 3 (black), 4 (blue), 6 (green), 12 (orange), and 16 (pink) are shown for (a), (b), and (c) as representatives. The bold horizontal red lines at 5.9 Å in (a) and (b) show the correct binding distance from the PDB.

We additionally used one of the typical docking software *GLIDE*[40] for the comparison. The initial structures were prepared using Protein Preparation Wizard by the default setting. OPLS2005 force field parameter was used, and the minimization was performed with the restraints on all the heavy atoms. The obtained structures were used for the docking templates. First, we used the SP mode to obtain 20 docking modes. Second, the obtained 20 modes were further refined by the XP mode. The binding modes were sorted according to the XP scores. Both self- and cross- docking were performed. In self-docking, the receptor conformation binding to the ligand was used as the docking template, and on the other hand, in cross-docking the receptor conformation binding to another ligand was used as the docking template. We confirmed that the docking software successfully reproduced the experimental binding modes for self-docking for all the systems, though it failed in some cases for cross-docking as shown in Figure 10.

■ RESULTS AND DISCUSSION

We performed the REUS simulations of 220−240 ns per replica for the four protein−ligand systems to predict the ligand binding structures. The simulation lengths were determined by examining the time series of the RMSD without translations and rotations of the ligand at the same umbrella potential. For example, the correct experimental binding distance from PDB is 7.3 Å in the case of the 2O2U system. Thus, we examined the time series of RMSD at the umbrella potential with the midpoint of 7.5 Å. It was found that the binding structures similar to the experimental ones were found for the first time after approximately 20 ns. Therefore, we performed 240 ns simulations for this system in order to check the convergence for more than 200 ns. The first 20 ns of data was not used for analysis because the time series of RMSD indicated the influence by the initial structures in the first 20 ns of the REUS simulation. In the prospective study, we predict the ligand binding structure from the global minimum free energy state by following our prediction protocol described in the Methods section, where we use the data excluding the initial 20 ns. We then calculate the RMSD from the predicted binding structure and can estimate the dependency on the initial structures similarly.

Figure 3 shows the time series of various quantities that we obtained from the REUS simulation of 1PMV as a representative system. Figure 3a shows the time series of the reaction coordinate (the distance of the ligand from the pocket) through the umbrella potential exchange for some of the replicas. We observed that the reaction coordinate oscillated between the low value (the pocket region) and the high value (the solvent region) and that sufficient sampling was achieved. Other replicas behaved similarly.

Figure 3b shows the time series of the values of the midpoints of the umbrella potentials corresponding to Figure 3a. When the midpoint was large, the reaction coordinate was large, and when the former was small, so was the latter. There is a strong correlation between Figures 3a and 3b, as is expected. The red horizontal lines at 6 Å in Figures 3a and 3b show the correct binding distances in PDB. When we compare (a) and (b) in detail, we observe that the actual distances in (a) do not necessarily follow the midpoints of the umbrella potential in (b), especially in the case of short distances. This is reasonable because the ligand has discrete stable regions inside the pocket. In addition, the fact that a ligand never approaches the distance of 5.0 Å suggests the steric exclusion. On the other hand, we observed that ligands at long distances tended to follow the umbrella potentials very closely for all the systems.

dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

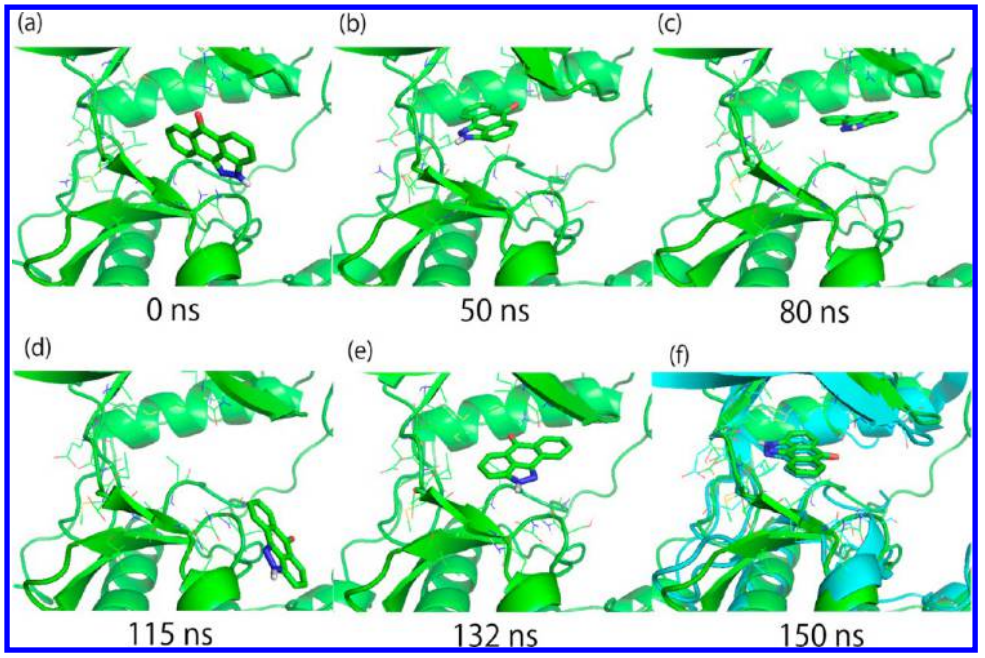

Figure 4. Typical snapshots of replica 12 from the REUS simulation of the 1PMV system. The configurations were taken at (a) 0, (b) 50, (c) 80, (d) 115, (e) 132, and (f) 150 ns. The RMSD from the correct ligand binding mode from PDB (1PMV) is (a) 9.1, (b) 4.8, (c) 7.3, (d) 13.3, (e) 5.9, and (f) 0.8 Å. Water molecules are suppressed to show the ligand positions clearly. In (f) the PDB conformation of the ligand (blue) was superimposed (see Figure 6c below for an enlarged picture). The figures were created using PyMOL.[27]

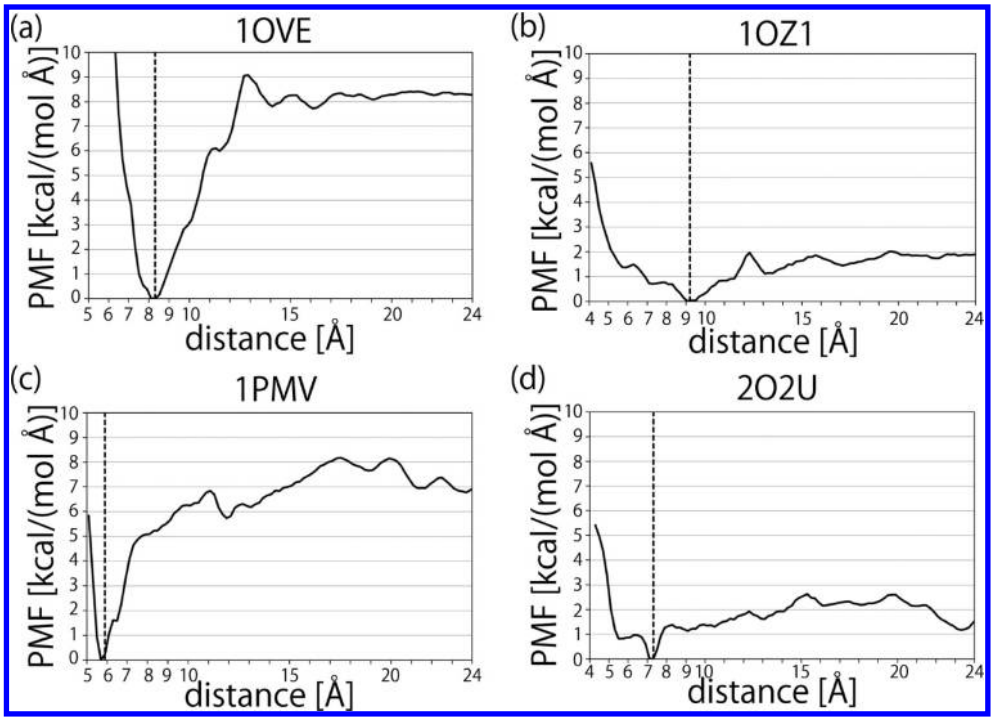

Figure 5. Potential of mean force profiles along the reaction coordinate ξ (the distance from the protein pocket) for all of the four systems, (a) 1OVE, (b) 1OZ1, (c) 1PMV, and (d) 2O2U. The vertical dotted lines show the distances of the experimental structures from PDB: (a) 8.3, (b) 9.2, (c) 5.9, and (d) 7.3 Å.

Figure 3c shows the time series of RMSD of the ligand from the PDB structure. Here, we calculated the RMSD of ligand-heavy atoms between the structure from the simulation and that from PDB without translations and rotations. We see that the RMSD oscillated between the lowest and the highest values found and that a wide conformational space along the reaction coordinate was indeed sampled during the simulations. By comparing Figures 3a and 3c, we see that Replicas 1 (black), 6 (green), and 12 (orange) experienced the distances close to the correct binding distance within the initial 50 ns, but they actually had very different conformations and large RMSD from the experimental binding structure in PDB. However, we observe that the correct binding structures were found

dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

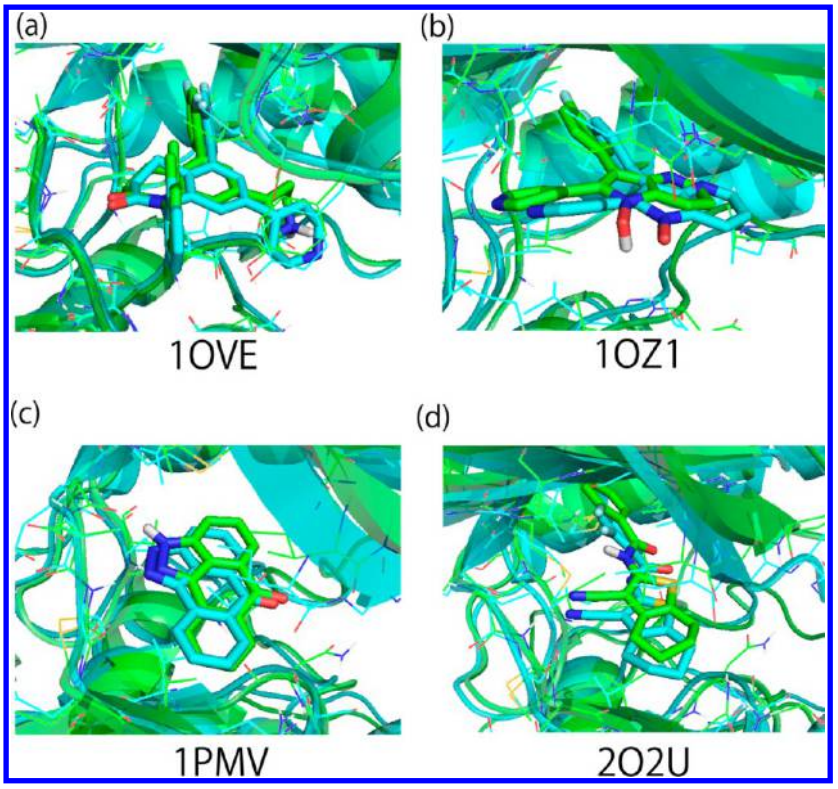

Figure 6. Comparisons of the predicted binding modes by the REUS simulations (green) with the experimental ligand binding modes from PDB (blue).The figures were created using PyMOL.[26]

independently by different replicas as the REUS simulation proceeded.

Figure 3d is the time series of replica exchange. The time series of the replica number at the umbrella potential with the midpoint distance of 6.0 Å is shown. We see that many replicas reached the midpoint value of 6.0 Å. These figures show that the replica-exchange simulations were appropriately performed. Thus far, we examined how well the REUS simulation worked in the case of a 1PMV system. We also confirmed that the REUS simulations were appropriately performed in all other systems.

In Figure 4, we show typical snapshots from the REUS simulation for one of the replicas (Replica 12) in the case of the 1PMV system. The ligand approached the pocket from (a) to (b) with an incorrect ligand orientation that did not fit the pocket shape very well. We see that the ligand binding mode in (b) is the configuration that is rotated by 180° from the correct one in (f). Thus, the ligand gradually moved away from the pocket in (c) and (d), and it approached the pocket again through many replica exchanges in (e). However, the ligand binding structure in (e) looks very different from the correct binding mode. Both the ligand and the protein pocket changed the conformations from (e) to (f) and finally found the correct binding structure in (f). This change of the protein pocket shape is called an "induced fit." In (f), we also show the correct binding structure from PDB by the molecule in blue color, which is in excellent agreement with the simulation result. During the process from (a) to (f), we observed large fluctuations of protein structures, especially for side chains and loop regions. We observed that the REUS simulation realized random walks not only in the reaction coordinate space but also in the conformational space, and they did not get trapped in one of a huge number of local-minimum-energy states.

Figure 5 shows the PMF $W_{\lambda=\{0\}}(\xi)$ for the four systems that we examined in this study. Our unit cells of the simulations were set so that we have a water solvation of at least 10 Å from the protein edge on each direction. This condition allowed the reaction coordinates of PMF, namely the distances from the protein pockets, to range up to about 25 Å, staying within the unit cells. Thus, roughly speaking, the ranges of the receptor pockets were about 15 Å. We collected approximately 10 000 ligand structures with the global minimum PMF inside the pocket for each system. The global−minimum PMF inside the pocket was also the global minimum PMF over the entire distance range for (a) 1OVE, (b) 1OZ1, (c) 1PMV, and (d) 2O2U systems. These collected structures were then analyzed by PCA. The contributions of the sum of the eigenvalues of the first two principal components to the total were between 48% and 94% for the four systems, which implies that these principal components were dominant.

The free energy landscapes of the four systems were analyzed by applying PCA (see eq 12 in the Methods section) to the collected structures with the global minimum PMF in Figure 5. We found only one stable structure, namely, one predicted binding structure, on the free energy landscape for each ligand in all of the systems.

Figure 6 compares the predicted binding modes (green) with the experimental ones (blue) from PDB. It appears that the predicted binding modes are in excellent agreement with the experimental binding modes for all the four systems when the fluctuations of both the protein and the ligand structures are considered. We quantified the agreements of the predicted ligand structures with the ones of PDB in Figure 6 in the following way. We first superimposed the coordinates with

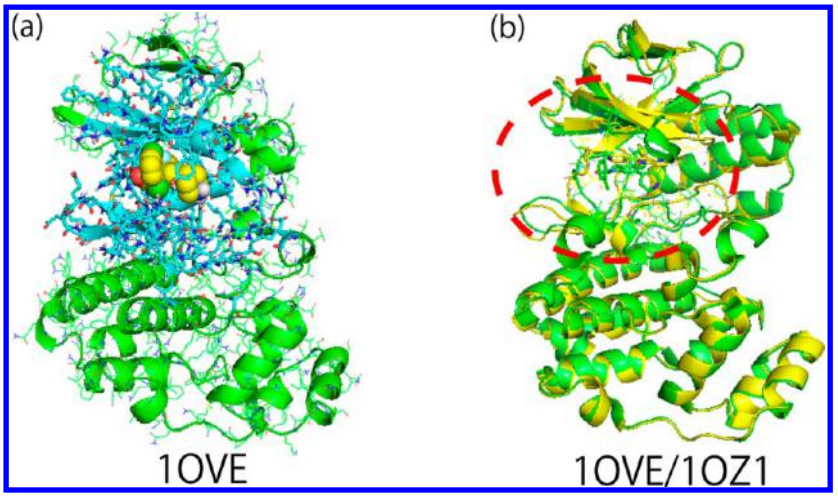

Figure 7. (a) The flexible region (blue,) which is within 12 Å from the ligand and the restrained region (green) in the case of 1OVE. (b) The superposition of 1OVE (green) and 1OZ1 (yellow). The dashed circle in red shows the flexible region in the simulations. The figures were created using PyMOL.[27]

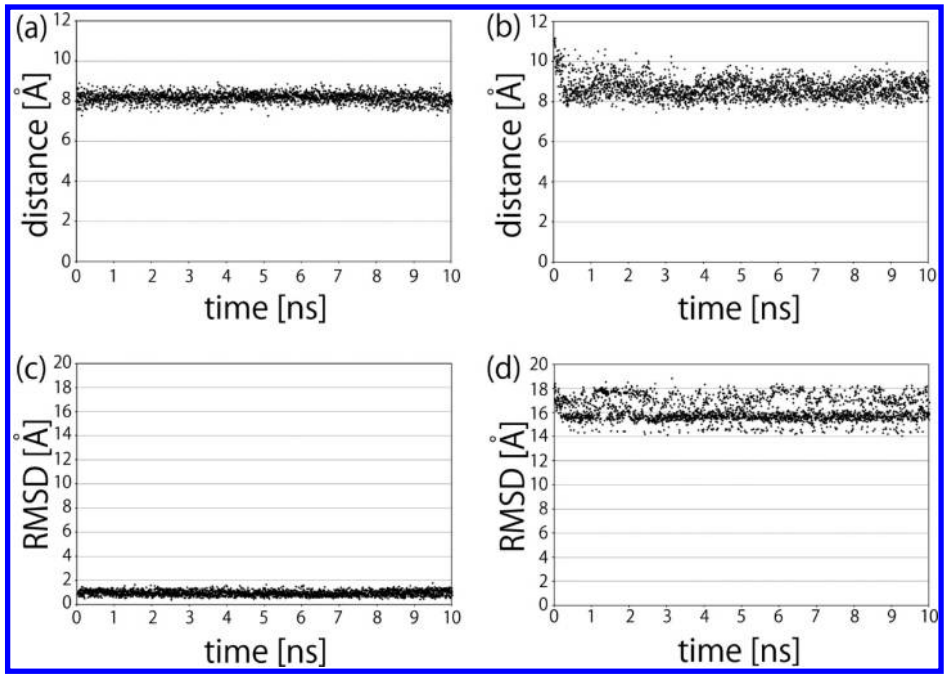

Figure 8. Time series of the reaction coordinate ξ (the distances from the protein pocket) at the umbrella potential with the midpoint of 8.0 Å obtained from the REUS simulation of the 1OVE system (a) without any restraints on all the atoms within 12 Å around a ligand and (b) with the restraints (k = 1.0 kcal/(mol Å^2)) on all the heavy atoms of a protein. (c) and (d) are the time series RMSD of a ligand molecule from the correct binding mode (a blue ligand in Figure 6a), corresponding to (a) and (b), respectively. The correct binding distance from PDB is 8.3 Å.

respect to all the protein C$^\alpha$ atoms by translations and rotations. These coordinate transformations gave the coordinates of the ligands (one predicted and the other PDB). We then calculated the RMSD of the two ligand structures without translations and rotations. The calculated RMSD values were (a) 0.99, (b) 1.61, (c) 0.71, and (d) 1.56 Å. The RMSD values in (b) and (d) are relatively large, but we think that the predicted binding modes are essentially the same, when the apparent translation of the atoms around the pockets is considered, because the structures were superimposed with respect to all of the protein C$^\alpha$ atoms. In order to examine the difference of the ligand internal conformations, we also calculated the RMSD values of the ligand conformations after the superposition by translations and rotations, which are: (a) 0.46, (b) 0.26, (c) 0.10, and (d) 0.38 Å. We see that the ligand internal conformations were very similar in all the systems.

Figure 7a shows the flexible region that is within 12 Å from the ligand and the restrained region in the case of 1OVE. We see that the large region including the surrounding loop region belongs to the flexible region. It is important to make the region that might be related to the ligand binding flexible. Figure 7b shows the superimposed structures of two p38 proteins (1OVE and 1OZ1). When we see Figure 7b in detail, the region within 12 Å from the ligand is different from each other, but the region apart from the ligand is very similar. Therefore, we used 12 Å for the threshold distance in our simulations. We think that 12 Å will be suitable for typical kinase targets, although it should be checked depending on the

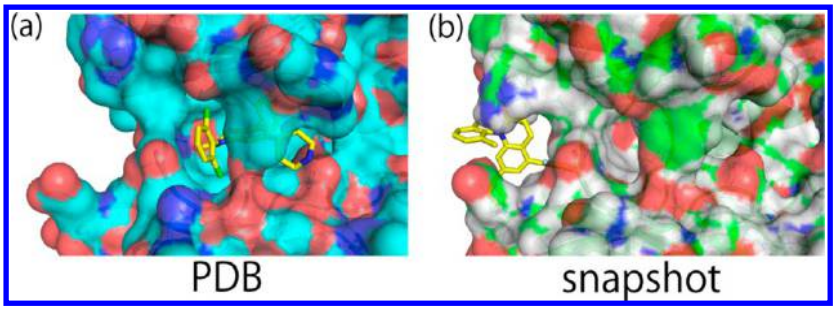

Figure 9. (a) PDB structure of 1OVE and (b) a snapshot from the REUS simulation with the restraints on all the heavy atoms of the protein for the replica with the umbrella potential of the midpoint of 8.0 Å (replica 12) at 10 ns. The figures were created using PyMOL.[27]

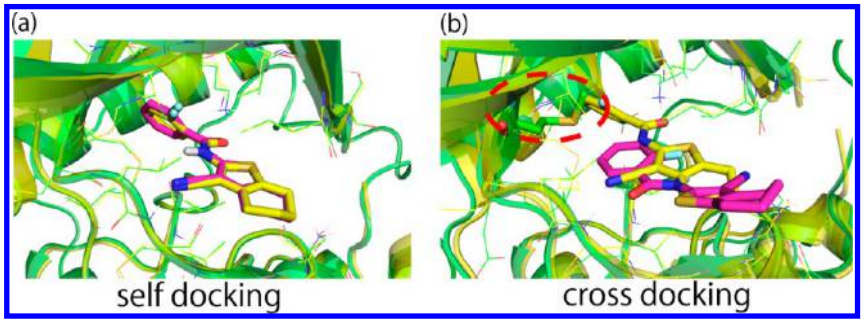

Figure 10. Comparisons of the predicted biding modes by the docking software *GLIDE*[36] (pink) with the experimental ligand binding modes from PDB (yellow). (a) Self-docking, which used the receptor conformation of 2O2U (PDB ID) for docking of the ligand of 2O2U, and (b) cross-docking, which used the receptor conformation of 1PMV for docking of the ligand of 2O2U. The residue encircled by the red dotted line is Met146, by which the correct binding mode (yellow) seemed to be sterically hindered. The figures were created using PyMOL.[27]

system. Figure 2a shows that the loop conformation around the binding pocket of the initial structure of our simulation (green) is very different from that of the PDB one (blue), probably because the receptor was in an apo state (a ligand was in solvent). However, Figure 6a shows that this loop conformation was properly changed to that similar to the PDB conformation after the ligand binding during our simulation. These observations suggest that our simulations were not biased toward the initial conformations, although the cross-docking test should be performed as a more rigorous check.

To investigate the need for protein flexibility for the protein−ligand association from a different point of view, we have performed the comparative REUS simulations with the restraints on the heavy atoms in the protein for one of the systems (PDB ID: 1OVE). All of the heavy atoms were constrained harmonically with a bond constant of 1.0 kcal/(mol Å^2) to the reference structure after the equilibration run, whereas the heavy atoms near the pocket were not constrained in the original conditions. The pocket shape of the reference structure should be suitable for the ligand because the reference structure was prepared by the equilibration run with the protein−ligand complex correctly binding to each other. The initial structures of the REUS simulations were prepared similarly as written in Computational Details section, but with the harmonic restraints of the heavy atoms in the protein.

Figure 8 shows the time series of the reaction coordinate ξ and RMSD that we obtained from the REUS simulations of 1OVE with and without restraints on the heavy atoms in the protein. Figures 8a and 8b are the time series of the reaction coordinate ξ (the distance of the ligand from the pocket) at the umbrella potential with the midpoint of 8.0 Å experienced by different replicas (a) with the original conditions (without restraints near the pocket) and (b) with the restraints on all the heavy atoms of the protein, respectively. We observe that both simulations properly sampled the structures with the distance specified by the umbrella potential. However, it is shown that the structures sampled by these two REUS simulations were very different.

Figure 8 panels c and d show the time series of RMSD of a ligand molecule from the correct binding structure from PDB (a blue ligand in Figure 6a) corresponding to Figure 8a and Figure 8b, respectively. We see that the simulation under the original conditions in Figure 8c sampled the structures close to the experimental one correctly. On the other hand, the simulation under the restraint condition on all the heavy atoms in Figure 8d did not find the correct binding structure in PDB. This is a reasonable result when we compare the PDB structure with structures from typical snapshots of the simulation corresponding to Figure 8d.

Figure 9 compares the PDB structure of 1OVE and a typical snapshot from the REUS simulation with the restraints on all the heavy atoms of the protein, which could not find the correct binding structure. We see that the ligand in PDB is surrounded by the protein atoms and the gateway seems to be too narrow if the heavy atoms in the protein are fixed. Therefore, the ligand could not approach the correct binding region by the REUS simulation with the restraints in (b). However, when the protein atoms around the pocket region were flexible, the ligand successfully found the correct binding structure by the REUS simulation (see Figure 6a). These results imply the importance of the protein flexibility for a ligand binding to a protein.

It is also interesting to investigate whether one of the typical docking software can predict the correct binding mode. We thus performed self- and cross-docking for all of the four systems we examined by using *GLIDE*.[40] In self-docking, the

dx.doi.org/10.1021/ct4004383 | J. Chem. Theory Comput. 2013, 9, 4660−4671

receptor conformation binding to the ligand was used as the docking template; on the other hand, in cross-docking the receptor conformation binding to another ligand was used as the docking template. We confirmed that the docking successfully reproduced the experimental binding modes for self-docking for all the systems. Figure 10a shows one of the results of self-docking for the system of 2O2U. We see that the docking actually predicted the correct binding mode very accurately. However, the performance of the cross-docking deteriorated very much. Figure 10b shows the result of cross-docking for the ligand of 2O2U by using the template of 1PMV. We see that the predicted binding mode was very different from the experimental one. No binding mode similar to the experimental one was included in 20 predicted modes by *GLIDE*. When we see Figure 10b in detail, we can presume why the docking software failed. We see that the side chain of Met146 in the template conformation of 1PMV has steric hindrance with the correct binding mode of the ligand of 2O2U. We think that this is one of the main reasons that docking did not work very well. Note that it can be due to the insufficiency of the force field parameters or the treatment of solvent effects in docking. We also found that the cross-docking of the ligand of 1PMV to the template of 2O2U failed similarly. It is well known that the conformational change of a protein called "induced-fit" is a common limitation of the existing docking software. Although the induced-fit tool is available for some docking software, it is limited to small conformational changes, such as the rotation of hydroxyl group or scaling van der Waals radii. On the other hand, we found that during our REUS simulations, the side chain of Met146 changed the conformations flexibly depending on the ligand and identified the correct binding modes as shown in Figure 6.

CONCLUSIONS

In this article, we have applied our prediction method for protein−ligand binding structures, which is based on the replica-exchange umbrella sampling, to two kinase systems (p38 and JNK3) with two different ligand molecules for each kinase. We showed that starting from the configurations in which the protein and the ligand are far away from each other, our method predicted the ligand binding structures in excellent agreement with the experimental data from PDB, which suggests the general applicability of our method to kinase systems. In addition, the protein flexibility was shown to be essential to predict the correct binding structure for one of the systems (PDB ID: 1OVE) by performing the comparative REUS simulation with the restraints on the heavy atoms of the receptor. We also examined the performance of one of the docking software and found that it performed very well when the template protein conformation of the docking was suitable to the ligand. However, it failed when the template conformation was not suitable for the ligand, which we usually cannot know in advance. These results suggest the importance of the protein flexibility for protein−ligand binding.

We expect that our method can relieve computational chemists of the considerable efforts devoted to develop a reasonable binding model by using various scoring functions and experimental structure−activity relationship information (SAR). By taking into account protein flexibility by MD simulations, one may not need to consider multiple templates of receptor structures, which are suitable for each class of ligand, and multiple docking conditions for docking software. In addition, one may not need to consider the rescoring of multiple docking structures for each ligand because our method predicts only one binding structure as the global minimum free energy state in most cases, though one or several binding structures may be also predicted as local-minimum free energy states in some cases.

Future work will involve the applications of the present method to simulations of more complex systems such as the DFG motif of kinase, which requires conformational fluctuations for ligand binding, and membrane proteins. Further improvements of the method are also possible. Two-dimensional replica exchange by the introduction of additional dimensions to the REUS, such as the van der Waals replica exchange, and the replica exchange with solute tempering,[42−46] may be effective for the enhancement of the ligand binding events.

AUTHOR INFORMATION

Corresponding Author

*H. Kokubo. E-mail: hironori.kokubo@takeda.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The simulations and computations were performed on the TSUBAME Grid Cluster at the Global Scientific Information and Computing Center of Tokyo Institute of Technology, supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Open Advanced Research Facilities Initiative. Y.O. was supported, in part, by the collaboration funds from Takeda Pharmaceutical Company and Grants-in-Aid for Scientific Research (A) (No. 25247071), for Scientific Research on Innovative Areas ("Fluctuations and Biological Functions" and "Dynamical Ordering & Integrated Functions"), and for Computational Materials Science Initiative and for High Performance Computing Infrastructure from MEXT, Japan.

REFERENCES

(1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(2) Homeyer, N.; Gohlke, H. *J. Comput. Chem.* **2013**, *34*, 965−973.
(3) Kuhn, B.; Kollman, P. A. *J. Med. Chem.* **2000**, *43*, 3786−3791.
(4) Gilson, M. K.; Given, J. A.; Bush, B.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(5) Gilson, M. K.; Zhou, H.-X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21−42.
(6) Top 500 Supercomputer Sites. http://www.top500.org/ (accessed May 26, 2013).
(7) Shaw, D. E.; et al. *Commun. ACM* **2008**, *51*, 91−97.
(8) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118−13123.
(9) Kokubo, H.; Tanaka, T.; Okamoto, Y. *J. Comput. Chem.* **2011**, *32*, 2810−2821.
(10) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(11) Hansmann, U. H. E; Okamoto, Y. *Curr. Opin. Struct. Biol.* **1999**, *9*, 177−183.
(12) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96−123.
(13) Kokubo, H.; Okamoto, Y. *Mol. Simul.* **2006**, *32*, 791−801.
(14) Durrant, J. D.; McCammon, J. A. *Curr. Opin. Pharmacol.* **2010**, *10*, 770−774.

(15) Grant, B. J.; Gorfe, A. A.; McCammon, J. A. *Curr. Opin. Struct. Biol.* **2010**, *20*, 142−147.
(16) Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1937−1942.
(17) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(18) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(19) Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40−57.
(20) Teeter, M. M.; Case, D. A. *J. Phys. Chem.* **1990**, *94*, 8091−8097.
(21) Kitao, A.; Hirata, F.; Go, N. *Chem. Phys.* **1991**, *158*, 447−472.
(22) Garcia, A. E. *Phys. Rev. Lett.* **1992**, *68*, 2696−2699.
(23) Abagyan, R.; Argos, P. *J. Mol. Biol.* **1992**, *225*, 519−532.
(24) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins: Struct.. Funct., Genet.* **1993**, *17*, 412−425.
(25) Kitao, A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164−169.
(26) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.
(27) Delano, W. L. *The PyMOL Molecular Graphics System*, Version 1.4.1; Schrödinger: Portland, OR.
(28) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(29) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357−1377.
(30) Okamoto, Y.; Tanaka, T.; Kokubo, H. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 699−712.
(31) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(32) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(33) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(34) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(36) Darden, T.; York, D.; Pederson, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(38) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(39) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(40) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J. Med. Chem.* **2006**, *49*, 6177−6196.
(41) Itoh, S. G.; Okumura, H.; Okamoto, Y. *J. Chem. Phys.* **2010**, *132*, 134105.
(42) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13749−13754.
(43) Affentranger, R.; Tavernelli, I.; Di Iorio, E. E. *J. Chem. Theory Comput.* **2006**, *2*, 217−228.
(44) Wang, L.; Friesner, R. A.; Berne, B. J. *J. Phys. Chem. B* **2011**, *115*, 9431−9438.
(45) Terakawa, T.; Kameda, T.; Takada, S. *J. Comput. Chem.* **2011**, *32*, 1228−1234.
(46) Moors, S. L. C.; Michielssens, S.; Ceulemans, A. *J. Chem. Theory Comput.* **2011**, *7*, 231−237.