Efficient Sampling of Ligand Orientations and Conformations in Free Energy Calculations Using the λ-Dynamics Method

Shinichi Banba,[†] Zhuyan Guo,[‡] and Charles L. Brooks, III*

Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037

The recently developed λ-dynamics free-energy based simulation method was used to study the binding of 10 five-member ring heterocycle derivatives to an artificial cavity created by mutagenesis inside cytochrome c peroxidase. Application of λ-dynamics to this system gives a reasonable estimate of the binding affinity of the ligands. This methodology also provides a means to explore the binding orientations and conformations of the ligands inside the binding pocket much better than does conventional MD. This is due to the scaling of forces inherent in the λ-dynamics method, which lowers the barriers separating different binding modes and conformations. Examination of the λ-dynamics trajectory of the ligands revealed alternative binding orientations and conformations not detected by crystallography. Furthermore, a λ-dynamics simulation starting from random initial orientations, in which some ligands take significantly different orientations as compared with those from the X-ray structure, successfully samples the X-ray crystallographic orientations in all ligands. Ligand sampling by conventional MD starting from same initial structures remains trapped in the local minima from which they start. Such efficient sampling of ligand orientations and conformations is expected to diminish the limitation that an initial ligand structure must be close to its true bound orientation in order to yield a reasonable estimate of the binding free energy.

Introduction

The advancements of high throughput screening and combinatorial chemistry techniques have moved drug design and discovery into a high-speed chemistry world. Computational methods that rapidly identify tight binding ligands have become especially appealing. When the target structure is unknown, QSAR is the method of choice.[1,2] Sometimes the target structure is solved either by X-ray or NMR techniques, and structural information from the target protein can be used to assist the design process. The challenge in structure based drug design and discovery is to identify ligands with high binding affinity, or low binding free energy. Several methods, with differing levels of accuracy and computational cost, have been developed and applied to biological systems in recent years.[3−9] Empirical methods such as DOCK[8,9] rank ligands based on shape and chemical complementarity. The simple scoring scheme used in DOCK enables it to rank ligands rapidly. Therefore, it has become a routine tool for rapid 3D-database searches in drug discovery. In a DOCK search, a large "free energy" cutoff value is normally used in order not to miss any promising compounds, and consequently a large number of DOCK hits are generated. To funnel down those hits to a manageable number is still a challenging task in drug discovery, and a more accurate but rapid method is needed to this end.[10]

Recently we have developed the λ-dynamics method for ligand binding free energy calculations.[5,11−13] We have shown previously that this method successfully discriminates the more favored ligands from the less favored ones for benzamidine-based inhibitors complexed with trypsin[5] or five membered ring inhibitors complexed with cytochrome c peroxidase.[13] In the λ-dynamics method, due to the scaling of protein−ligand interactions, as we will present later, the barriers between various binding modes are reduced. This speeds up the exploration of orientational degrees of freedom. For the same reason, conformational sampling of ligands is also enhanced. In drug design applications, the 3-D structure of a protein−ligand complex is often known for only one or two ligands. The initial structures of chemically related putative ligands are prepared by superimposing them with the known compound's complex structure or by using computational methods such as docking. Although many docking algorithms and programs have been recently developed,[8,9,14−17] these approaches do not always give the true binding orientation or conformation.[17−19] Furthermore, the fact that the docking scoring function is often not general results in errors in estimating the binding affinity, making it difficult to identify the correct binding structure out of putative binding conformations detected by these docking methods.[20,21] Theoretically rigorous methods, such as free energy perturbation (FEP)[22] or thermodynamic integration (TI),[3] not only are computationally too intensive to be practical tools in drug design, but they also require that the initial orientation of the ligand is close to its true bound orientation. Furthermore, these methods cannot be applied straightforwardly to ligands that have multiple binding modes. Consequently, there is no viable way to search for a preferred binding orientation or conformation on the basis of free energy. In this application, we show that the λ-dynamics method is capable of addressing these problems and provides a means to explore binding orientations or conformations on a free energy basis. This occurs because the scaled potential used in the λ-dynamics method reduces the potential barriers and

* To whom correspondence should be addressed. E-mail: brooks@scripps.edu, www:http://www.scripps.edu/brooks.
† On leave from Mitsui Chemical Inc., 1144, Togo, Mobara-shi, Chiba 2970-0017, Japan.
‡ Current address: Schering-Plough Research Institute, Kenilworth, NJ 07033.

enhances the sampling efficiency of ligand orientation and conformation.

The rest of the paper is divided as follows. An overview of the method is given first. This is followed by the computational protocols and methodological details. Then the results of our calculations, such as the binding free energy calculations; the binding mode; and the conformational sampling analysis and searching for the optimal binding orientation are presented. Finally, we summarize our work and comment on its future applications in drug design.

Method

The λ-dynamics method is an extension of the FEP method. The details of the method have been presented elsewhere.[5,11-13] For completeness, here we will briefly go through the main ideas and formulas. For a protein and a total of L ligands, a hybrid potential function is constructed as follows:

$$V(X,\{x\},\{\lambda\}) = V_{\text{env}}(X) + \sum_{i=1}^{L}\lambda_i^2(V_i(X,x_i) - F_i) \quad (1)$$

with $\sum_{i=1}^{L}\lambda_i^2 = 1$.

In eq 1, i indicates the ith ligand, X and x_i indicate the coordinates of environmental atoms and those attached to ligand i, respectively, and F_i is a precalculated biasing potential corresponding to the relative free energy of ligand i in the unbound state (relative solvation free energy). This latter term can be rapidly evaluated using continuum solvation models,[23-27] or using free energy perturbation results as we do in this work. Finally, V_{env} is the interaction energy involving the environmental atoms only; $V_i(x)$ is the interaction energy involving ligand i in the protein−ligand complex state; λ_i is the coupling parameter associated with ligand i. The dynamics of the system is generated from the extended Hamiltonian

$$H_0(X,\{x\},\{\lambda\}) = T_x + T_\lambda + V_{\text{env}}(X) + V(\{\lambda\},X,x_i) \quad (2)$$

Here T_x and T_λ are the kinetic energy of the atomic coordinates and λ variables, respectively. The λ's are treated as volumeless particles with mass m_λ. The difference in binding free energy ($\Delta\Delta G$) between ligand i and ligand j can be obtained from

$$\Delta\Delta G_{ij} = -k_{\text{B}}T \ln\frac{P(\lambda_i^2=1,\{\lambda_{m\neq i}^2=0\})}{P(\lambda_j^2=1,\{\lambda_{m\neq j}^2=0\})} \quad (3)$$

$P(\lambda_i^2=1,\{\lambda_{m\neq i}^2=0\})$ is the probability of ligand i having $\lambda_i^2 = 1$. In our λ−dynamics simulations, the probability $P(\lambda_i^2=1,\{\lambda_{m\neq i}^2=0\})$ corresponds to the amount of time ligand i has $\lambda_i^2 = 1$. Practically, we consider $\lambda_i^2 = 1$ if λ_i^2 is greater than a threshold value.

When a multiple topology model is used for the λ-dynamics simulation, additional restraining potentials are required for rapid convergence. The restraining potential is defined as a function of x_i, λ_i, and average coordinates of environmental atoms, X_0.

$$H'(X,x,\lambda) = H_0(X,x,\lambda) + \sum_{i=1}^{L}R_i(X_0,x_i,\lambda_i) \quad (4)$$

Therefore, the restraining potentials do not directly depend on environmental atom coordinates, which are denoted by $X(t)$. We define R_i as follows:

$$R_i = \alpha(V_i(X_0,x_i) - F_i)(1 - \lambda_i^2) \quad \text{and} \quad 0 < \alpha < 1 \quad (5)$$

The newly defined parameter α determines the scaling factor for the force on the ligand when it is noncompetitive ($\lambda_i^2 \sim 0$). The restraining potential disappears when $\lambda_i^2 = 1$. Whereas, when λ_i^2 reaches zero, the ligand is sampled with a force scaled by α, which corresponds to high-temperature T' ($T' = T/\alpha$, where T is the temperature used in the simulation). It is easy to understand the physical meaning of the parameter α. Since the restraining potentials do not depend on $X(t)$ (or $X(t)$ is slowly varying around X_0 as discussed below), the partition function can be expressed as follows when $\lambda_i^2 = 1$.

$$Z'_{\lambda_i^2=1} = \int\exp\{-\beta(V_{\text{env}} + V_i - F_i)\}\,dX\,dx_i \times$$
$$\prod_{k\neq i}^{n}\int\exp\{-\beta R_k(x_k,\lambda_k=0)\}\,dx_k \quad (6)$$

Using this relationship, $\Delta\Delta G$ can be written using two terms.

$$\Delta\Delta G_{j\to i} = -k_{\text{B}}T \ln\frac{Z'_{\lambda_i^2=1}\prod_{k\neq i}^{n}\int\exp\{-\beta R_k(x_k,\lambda_k=0)\}\,dx_k}{Z'_{\lambda_j^2=1}\prod_{l\neq j}^{n}\int\exp\{-\beta R_l(x_l,\lambda_l=0)\}\,dx_l}$$
$$= -k_{\text{B}}T \ln\frac{P'_{\lambda_i^2=1}}{P'_{\lambda_j^2=1}} - k_{\text{B}}T \ln\frac{\int\exp\{-\beta R_i(x_i,\lambda_i=0)\}\,dx_i}{\int\exp\{-\beta R_j(x_j,\lambda_j=0)\}\,dx_j}$$
$$(7)$$

The first term involves the probability of the ligands achieving a dominant state ($\lambda_i^2 = 1$) calculated using the λ-dynamics with the restraining potentials, while the second term is a function of the restraining potential partition functions. The second term can be replaced by the internal energy difference as shown in eq 8, and further reduced if we assume that the entropy term for each ligand, related to the restraining potential, is canceled due to the similarity of the ligands.[13]

$$-k_{\text{B}}T \ln\frac{\int\exp\{-\beta R_i(x_i,\lambda_i=0)\}\,dx_i}{\int\exp\{-\beta R_j(x_j,\lambda_j=0)\}\,dx_j} = (U'_i - U'_j) - (S'_i - S'_j)T$$
$$\approx U'_i - U'_j \quad (8)$$

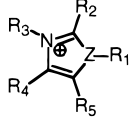The internal energy related to the restraining potential can be estimated from the λ-dynamics simulation using.

$$U'_i = \int R_i(x_i,\lambda_i=0)\exp\{-\beta R_i(x_i,\lambda_i=0)\}\,dx_i$$
$$\approx \frac{1}{n}\sum_{(\lambda_i=0)}^{j=1,n}R_j \quad (9)$$

Furthermore, we have to resolve the question regarding the proper X_0 to use in eq 8. We assume that the environmental atoms move more slowly than the ligand atoms. Consequently, the restraining potential $R_i(X_0,x_i)$ can be replaced by $R_i(X(t),x_i)$.

$$R_i(X_0,x_i) \approx R_i(X(t),x_i) \quad (10)$$

The approximation shown in eq 10 saves considerable computational cost because the interactions between X_0 and $\{x\}$ are eliminated. However, the system no longer strictly obeys Newton's equations of motion since our biasing potential is not

TABLE 1: Structures of Five-member Ring Ligands

name	Z	R_1	R_2	R_3	R_4	R_5
3met	S	-	H	CH_3	H	H
2am4	S	-	NH_2	H	CH_3	H
2am5	S	-	NH_2	H	H	CH_3
34di	S	-	H	CH_3	CH_3	H
234t	S	-	CH_3	CH_3	CH_3	H
345t	S	-	H	CH_3	CH_3	CH_3
nmei	N	CH_3	H	H	H	H
nvi	N	$CH=CH_2$	H	H	H	H
2eti	N	H	CH_2CH_3	H	H	H
dime	N	H	CH_3	CH_3	H	H

"seen" by the environment. As we have shown in a previous study, these approximations appear to be robust.[13] With these approximations, $\Delta\Delta G$ can be estimated using this equation.

$$\Delta\Delta G_{j \to i} \approx -k_B T \ln \frac{P'_{\lambda_i^2 = 1}}{P'_{\lambda_j^2 = 1}} + (U'_i - U'_j) \qquad (11)$$

Simulation Details

As shown in Table 1, the system we have studied is a set of small cationic molecules that bind to an artificial cavity created inside cytochrome c peroxidase (CCP).[28,29] To design an enzyme that oxidizes a specific molecule, mutagenesis was used to create an artificial cavity by replacing the tryptophan by a glycine residue.[28,29] X-ray crystallographic structures of the ligands complexed with the protein indicate that certain molecules bind specifically to the cavity.[28,29] The binding affinity of a series of structurally characterized protein−ligand complexes has been measured using an optical spectroscopy technique.

All computations were performed using the CHARMM molecular dynamics package with the all hydrogen force-field (top_all22_prot/par_all22_prot).[30,31] The missing parameters for the ligands, including bond, angle, dihedral and improper energies, were obtained based on QUANTA parameters for similar atom types and scaled up or down to be consistent with the CHARMM parameter set. The charges of the ligands were obtained from quantum chemical calculations employing a 6-31G** basis set and an ESP fitting scheme (GAUSSIAN 94, calculations are carried out at the U.S. Army Research Laboratories by Drs. S. W. Bunte and G. M. Jensen). The initial coordinates of the complexes were taken from the X-ray crystallographic structures (provided by Professor D. Goodin at The Scripps Research Institute). The protocol for preparing the initial structure has been presented elsewhere.[13] Water molecules were represented by the TIP3P water model of Jorgensen.[32] All bonds containing hydrogen atoms were constrained to their parameter values using the SHAKE algorithm.[33] Nonbonded interactions were treated using a cutoff of 12.4 Å along with van der Waals switching between 8.0 Å and 10.0 Å and an electrostatic shifting function. The temperature of the system was maintained near 300 K by coupling the non-hydrogen atoms to a Langevin heatbath. The λ-dynamics and FEP simulations used a time step of 1 and 1.5 fs, respectively. Molecular dynamics simulation of the solvated protein−ligand complex, which was partitioned using a 12 Å reaction zone and a 3 Å buffer region, was carried out using the stochastic boundary molecular dynamics method (SBMD).[34] The final system contained 1677 protein atoms, one heme group, 21
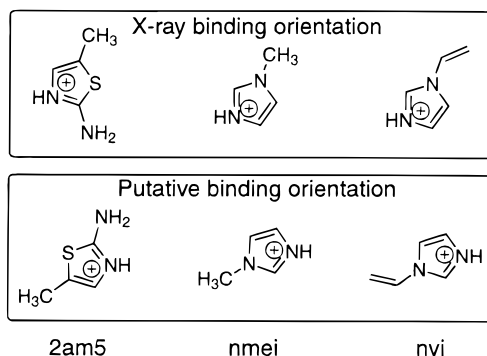

Figure 1. Schematic figures of the X-ray crystallographic binding orientations and putative initial orientations for 2am5, nmei, and nvi.

crystal water molecules, and 88 solvent water molecules. The λ-dynamics calculations were carried out with the solvated protein−ligands system for 300 ps with the initial 30 ps used as an equilibration phase. The ligands included in the simulations are shown in Table 1. The λ trajectories were saved every 15 fs and were used for later analysis.

To explore the efficiency of sampling by the λ-dynamics method, we prepared the initial structures with putative orientations that mimic the actual situation wherein the 3-D structures of the protein−ligand complex were known only for some ligands. Some of the ligands thus were initially modeled with incorrect binding modes. The binding structures of the other ligands were estimated by using the known protein−ligand complex structures. The X-ray crystallographic structure of 234t was chosen as a reference and the structures of the other ligands were generated by superimposition of their five membered rings. Superimposing the ligands was done using QUANTA. Thus, the generated binding orientations of 2am5, nvi, and nmei were very different from their X-ray crystallographic structures, while those of the other ligands were close to the conformation found in their crystallographic structures. Schematic figures of the X-ray crystallographic orientations and putative initial orientations for these ligands are shown in Figure 1. The initial structure with ligands in their putative binding orientations was taken through a series of minimization stages of 400 steps each, during which all the heavy atoms except the bulk water molecules were harmonically restrained to their positions. The initial restraining force constant was 120 (kcal/mol)/Å^2, which was reduced in steps of 10 (kcal/mol)/Å^2 until a force constant of 60 (kcal/mol)/Å^2 was reached. The λ-dynamics calculations were then carried out for 300 ps with the solvated protein−ligand systems including 4, 6, or 10 ligands. The 4-ligand system included four imidazolium derivatives (i.e., 2eti, dime, nmei, and nvi), while the 6-ligand system included the six thiazolium derivatives (i.e., 234t, 2am4, 2am5, 345t, 34di and 3met).

Conventional FEP calculations were also performed on the solvated protein−ligand complex system. To get $\Delta\Delta G$ between all 10 ligands, FEP calculations were carried out for nine ligand pairs.[13] We performed simulations with λ = 0.03125, 0.125, 0.325, 0.5, 0.675, 0.875, and 0.96875, respectively. The free energy change for each simulation was calculated using the double-wide sampling technique.[3] For each λ, a 30 ps equilibration period was followed by 120 ps of data collection. Therefore, a total of 1050 ps of simulation time was used to get one free energy difference between a pair of ligands complexed with the protein.

To perform a λ-dynamics calculation of the relative binding affinity of a set of ligands, the free energy of the ligands in the unbound state (F_i in eq 1) has to be predetermined. This was
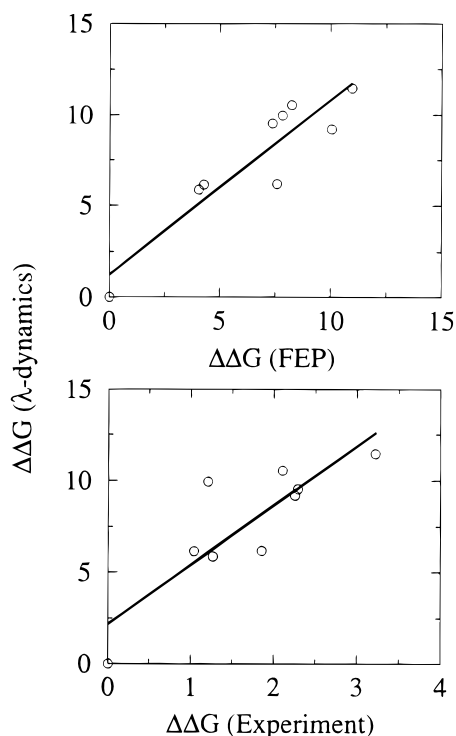
Figure 2. Comparison of relative binding free energy of the ligands between the λ-dynamics and FEP method (a) or experimental results (b) as discussed in ref 13.

Figure 3. Distribution of the dipole orientation projected along a certain direction (X-axis) of 34di and nvi from a 270 ps λ-dynamics trajectory for a 10-ligand system. The dipole moment of the initial X-ray crystallographic structures with 34di and nvi are 0.1 and 2.2, respectively. The dashed line shows the results of conventional MD including only a single ligand. The solid line shows data from the entire λ-dynamics trajectory. The dotted line shows the distributions that include only the bound states ($\lambda^2 > 0.8$) from the λ-dynamics trajectory. There are two dominant orientations with 34di; the X-ray crystallographic orientation and an alternative orientation. Sampling of the ligand orientation by conventional MD was restricted around the initial X-ray crystallographic structure.

also done using the FEP simulation method.[13] In the solvation FEP simulations, the systems consisted of the selected two ligands and 500−503 water molecules in 24.8 Å cubic box with periodic boundary conditions. We performed simulations with $\lambda = 0.125$, 0.5, and 0.875, respectively. For each λ, a 30 ps equilibration period followed by 60 ps of data collection was performed. Therefore, a total of 270 ps simulation time is used for each solvation free energy difference.

Results

Orientational Motion of Ligands Inside the Binding Pocket. As shown in Figure 2, a short λ-dynamics simulation with the 10-ligand system successfully estimated the binding free energy differences as compared with those of FEP simulation and experiment. In this calculation, a value of 0.3 was chosen for α and we chose a threshold of $\lambda^2 = 0.8$ to approximate the $\lambda^2 = 1$ state. The details of the calculated binding free energy differences are discussed elsewhere.[13] In the course of the λ-dynamics trajectory, we noted significant motion of the ligands within the binding pocket. We thus examined the orientational sampling of the ligands in the binding pocket. The efficient sampling of ligand orientations during a simulation is important for a good estimate of ligand binding free energy. Figure 3 gives the distribution of the dipole moment of the ligands projected along a fixed (lab frame) direction. Since the dipole moment is somewhat ill-defined and coordinate system dependent when the total charge of each ligand is not zero, in this analysis, the center of geometry of the selected ligand atom was used as origin for the coordinate system in which the dipole moment was calculated. It is clear from Figure 3 that there are two dominant orientations for the ligand 34di in the bound state ($\lambda^2 > 0.8$), the peak to the left is in alignment with the crystallographic binding orientation, while the one to the right is not. Judging from the relative bound state population of the two orientations, which could be used to determine the
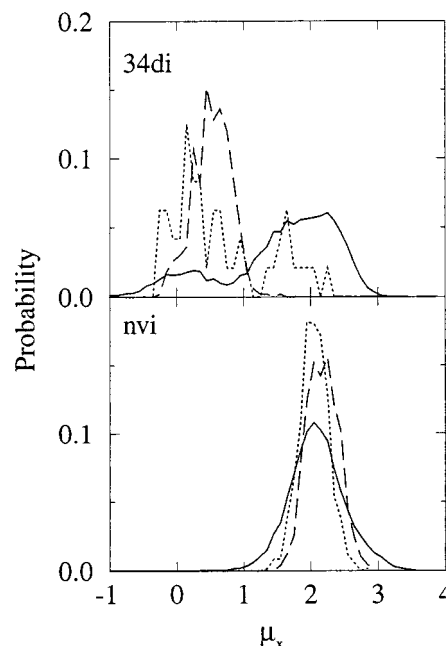
relative binding free energy of the two modes, the crystallographic binding orientation is more stable by about 0.4 kcal/mol than the alternative binding orientation detected in the λ-dynamics simulation. However, the alternative binding orientation is also stable enough to contribute to the binding free energy since 34di reached the λ^2 threshold in this alternative orientation as well. Although the other ligands adopt dominant states ($\lambda^2 > 0.8$) only when their binding orientations are close to X-ray orientations, the λ-dynamics method also explores a larger ligand orientational space than conventional MD as shown in Figure 3b for nvi. Figure 4a,b illustrate snapshots of the X-ray and an alternative binding orientation for 34di, respectively. The coordinates were extracted from the portions of the λ-dynamics trajectory having $\lambda^2 > 0.8$ for this ligand, which corresponds to the ligand in the bound state. The relative orientations of the ligand were obtained by superimposing the protein backbone structures. In the X-ray binding orientation, C−H (ligand) and O−C (Asp-235) form an unconventional hydrogen bond (H··· ·O distance, 1.8−2.2 Å; C−H−O angle, 120−170°). A previous X-ray crystallographic study also observed this unconventional hydrogen bond.[35] The alternative orientation also maintains this type of hydrogen interaction using a different C−H group with a similar distance (1.8−2.2 Å). The C−H···O−C hydrogen bond in the alternative binding orientation is weaker than that in the X-ray orientation due to the different acidity of the two hydrogen atoms involved. The ligand also adopts a parallel orientation with respect to His-175 (see Figure 4a,b). Judging from the average interaction energy calculated from the λ-dynamics trajectory, a slightly stronger ligand−His-175 interaction in the alternative orientation partially compensates for the weaker hydrogen bond between Asp-235 and the ligand. The alternative
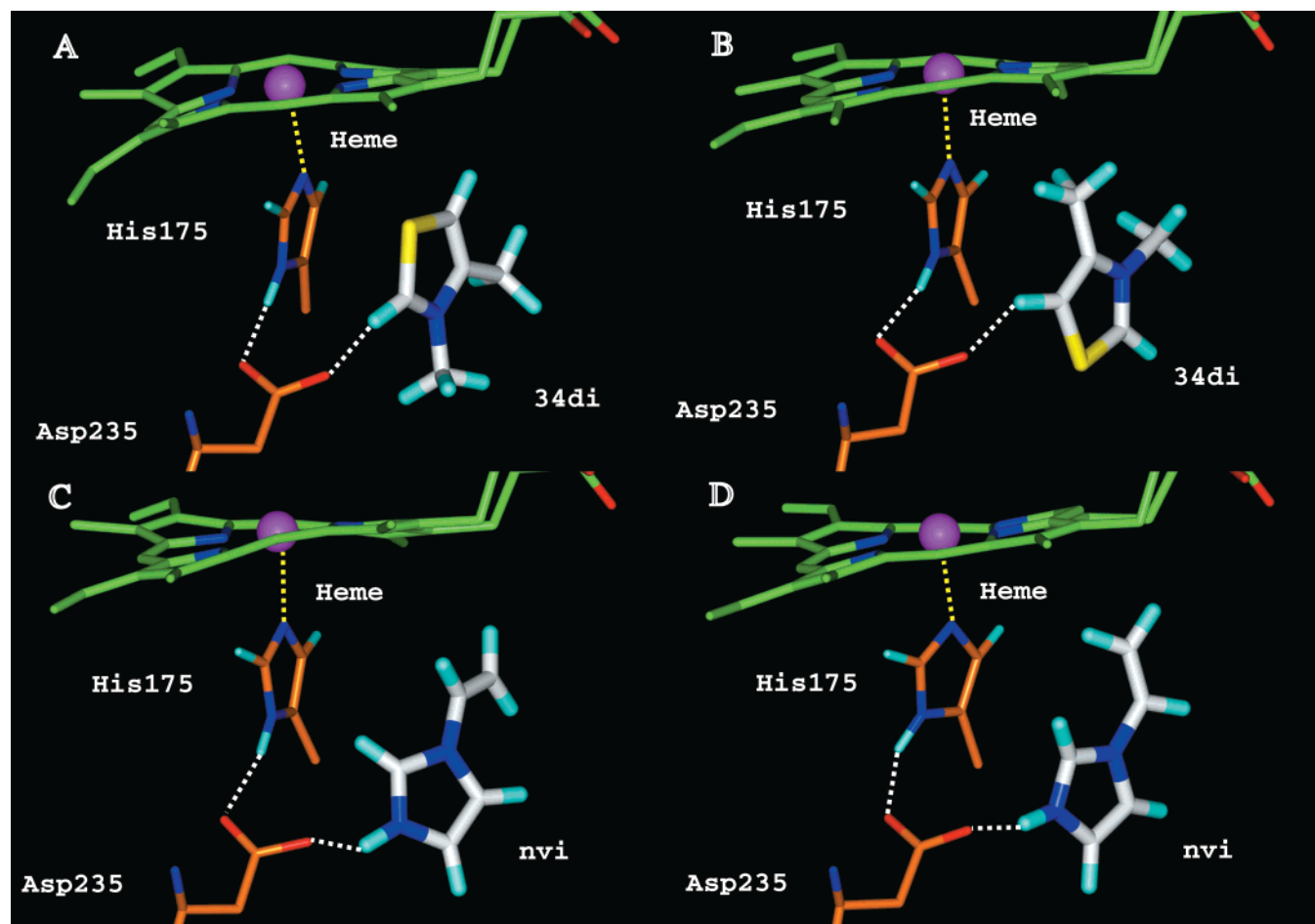
Free Energy Calculations Using λ-Dynamics

J. Phys. Chem. B, Vol. 104, No. 29, 2000 **6907**

Figure 4. Snapshots of the binding modes for 34di and nvi inside the binding pocket taken from a λ-dynamics trajectory with $\lambda^2 > 0.8$, which corresponds to the bound states. The X-ray binding orientation (a) and alternative binding orientation (b) are shown for 34di. (c) and (d) show the X-ray conformation and alternative conformation for nvi. Only the ligands, heme, Asp-235, His-175 are shown for clarity. The carbon atoms for heme, protein, and the ligands are shown in green, orange, and white, respectively. The oxygen, nitrogen, hydrogen are shown by red, blue, cyan, respectively. Iron atom is shown by the sphere.

binding orientation observed in the λ-dynamics trajectory maintains all important interactions with the protein and seems to be a physically reasonable binding orientation. Further investigation by docking studies also revealed a third binding orientation for 34di in which the ligand is inverted about the C−H bond interacting the O−C (Asp-235) in the X-ray orientation. This orientation was also energetically competitive with the two orientations observed in the λ-dynamics simulation. However, this third orientation was not observed in all λ-dynamics simulations, because the transition path from the initial X-ray orientation to this orientation required a large distortion in protein conformation and the λ-dynamics method does not allow such high energy conformations to be sampled for the environmental atoms. The presence of the alternative orientation in 34di proposed here awaits experimental verification.

Conformational Sampling of the Ligands Inside the Binding Pocket. The sampling of ligands inside the protein cavity was also investigated. It should be noted here that we are not addressing the sampling of protein conformations. Although the protein is free to move, we are not focusing on the issue of protein conformational changes upon binding. The protein conformation remains close to its initial conformation during the λ-dynamics simulation (total backbone rmsd ~0.6 Å). Following the same argument as in the previous section, the sampling of ligand conformations is enhanced due to the scaling of forces. Therefore, if a ligand has multiple binding

conformations, the method will be more likely to identify them than regular dynamics simulations. Since the conformation of a molecule is determined by its torsional angles, we chose the torsional sampling of nvi as an example (see Figure 5 for a definition of the torsion). This torsion seems to be appropriate for investigating the efficiency of conformational sampling because it has two local minima and the barrier between them is too high for conventional MD to sample both states within the restricted computational time. Figure 5 shows that this torsional angle sometimes changes between two local minima ($\phi = 0°$ and $180°$) within a 300 ps simulation time, while in conventional MD it stays in one local minimum ($\phi = 0°$) within the same simulation time. These results indicate that the sampling of torsional degrees of freedom is enhanced by the λ-dynamics method. Thus, the λ-dynamics calculation predicts an alternative binding conformation ($\phi = 180°$) for nvi because this ligand reached $\lambda^2 > 0.8$ (see Figure 5) at both the X-ray crystallographic conformation shown in Figure 4c and also $\phi = 180°$. This alternative binding conformation is shown in Figure 4d. It is difficult to estimate the free energy difference between these two conformations due to the restricted sampling, but the alternative conformation seems to be more stable than the X-ray crystallographic conformation by about 0.4 kcal/mol in our force field. The consideration of alternative conformations might lower the $\Delta\Delta G$ of nvi estimated by the λ-dynamics method over that estimated by the FEP method. The alternative conformation also maintains the C−H···O hydrogen bond with
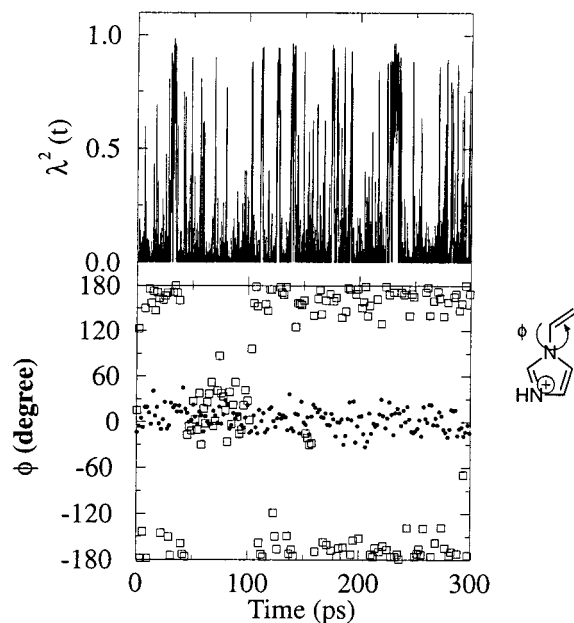
Figure 5. Change of torsion angle (ϕ) for nvi during simulations. The dot and squares show the results of a conventional MD simulation and λ-dynamics simulation, respectively. The torsion angle sometimes makes transitions between two minima (0°and 180°) only in λ-dynamics simulation. The trajectory of λ^2 for nvi is also shown.

Asp-235 with a similar distance. Both orientations have similar interaction energies with the protein environment, where $\lambda^2 > 0.8$ for this ligand. The alternative binding conformation for nvi observed in our λ-dynamics simulation also awaits experimental verification. Enhanced torsional sampling is an interesting feature of the method. Another example of enhanced conformational sampling using a scaled potential function similar to that used in the λ-dynamics (the fluctuating σ-method) comes from a study of folding of a peptide chain.[36] It was found that the use of a fluctuating potential function speeds up the folding of the chain to the final state, while a using rigid, unscaled potential, the chain was trapped in an intermediate state and unable to find the final state within the simulation time.

Exploring Stable Ligand Binding Orientations with λ-Dynamics. λ-dynamics simulations from putative (and incorrect) initial structures were carried out to illustrate its sampling efficiency. To contrast the results from the λ-dynamics method, conventional MD simulations of single ligand−protein complexes were also carried out with three ligands (2am5, nmei, and nvi). The same initial structures were used for both the λ-dynamics simulations and the conventional MD simulations. As shown in Figure 6, conventional MD from our initial structures were all trapped in the local minima near the initial structures and failed to move to the X-ray crystallographic binding structures. On the other hand, in the λ-dynamics simulation with the 6-ligand system, 2am5 was trapped in two local minima but reached the X-ray crystallographic orientation within 300 ps (Figure 7). In a run with a 4-ligand system, which includes the 4 imidazolium derivatives, both nmei and nvi also reached the X-ray crystallographic structures within 300 ps of simulation (Figure 7). A 10-ligand system initiated from the putative binding structures and including both thiazolium and imidazolium derivatives was also tested to confirm sampling efficiency. The three ligands (2am5, nmei, and nvi) successfully reached their crystallographic orientations within 300 ps of the λ-dynamics simulation (Figure 7). Moreover, these ligands reached the bound states ($\lambda^2 > 0.8$) only after they sampled the X-ray crystallographic binding structures with all λ-dynamics
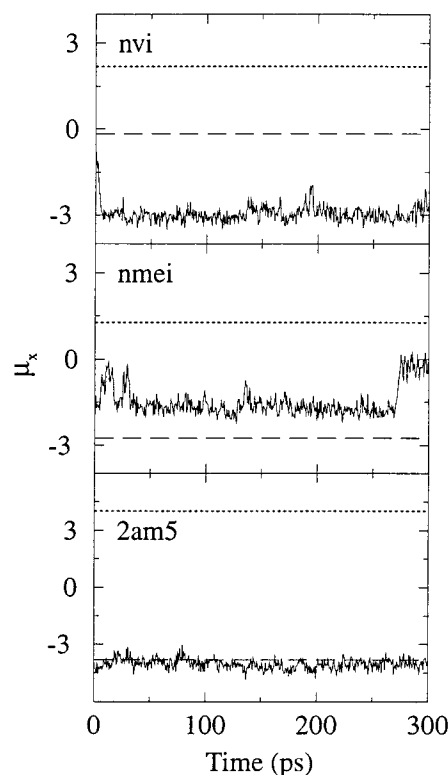

Figure 6. Change of dipole orientation projected along the *x*-direction (μ_x) of 34di, nmei and nvi from conventional MD simulation. Each simulation started from a putative and incorrect initial orientation. The dashed line and dotted line show the dipole orientation of the initial orientation and X-ray orientation, respectively.

simulations started from the incorrect orientations (Figure 7). While 2am5, nmei, or nvi were in incorrect orientations, the other ligands quickly moved to their X-ray binding structures and the probability of their dominant states was roughly proportional to their binding free energies calculated by the FEP method. All ligands took the dominant states only when their orientations were close to X-ray orientations during these three λ-dynamics simulations. We speculate that the alternative binding orientation for 34di detected in the previous simulation was not sampled in these simulations because 34di dominanted the large λ^2 states longer when the strongest binding ligand 2am5 was in an incorrect binding orientation.

These results clearly show that when a ligand adopts λ values near zero rapid exploration of low energy orientations and conformations occur due to the scaling of the potential. Then, at a later time fluctuations in the protein configuration, the ligand conformation, or both, occur to induce a "binding mode" configuration of λ. Furthermore, these results show that the λ-dynamics method can be applied to explore the docking of the ligands on a free energy basis.

Conclusions

We have presented a set of the promising observations for ligand binding, ranking and exploring ligand binding orientations and conformations using the λ-dynamics method. These include the consistency of λ-dynamics calculations with FEP calculations and experimental results, enhanced sampling of orientational and conformational degrees of freedom, and rapid search of binding orientations during a λ-dynamics simulation. For orientational and conformational sampling, this method is much more efficient than an FEP approach. This is because when a ligand is not bound to the protein, small λ^2 values, the internal

Free Energy Calculations Using λ-Dynamics

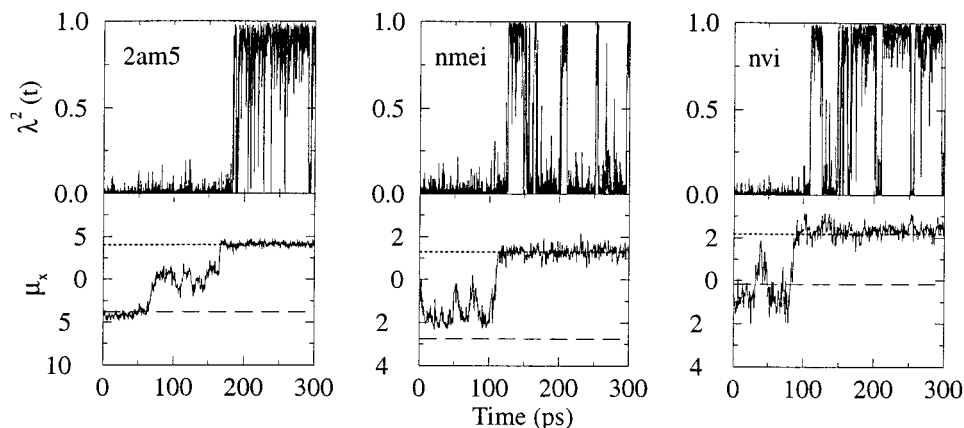*J. Phys. Chem. B, Vol. 104, No. 29, 2000* **6909**

Figure 7. Change of dipole orientation (μ_x) of 34di when (a) 6 ligands are considered and (b) those of nmei and nvi with a 4-ligand system. The λ-dynamics simulations started with incorrect initial orientations for the ligands. The trajectory of the λ^2 values for 34di, nmei and nvi are shown at the same time. The dashed line and dotted line show the dipole orientation of the initial and X-ray orientation, respectively.

energy of the ligand and the protein−ligand interaction energy is scaled. This reduces the barriers between the different ligand orientations and conformations. In other words, when a ligand is not bound to the protein, rapid exploration of low energy binding orientations or conformations occur, enabling the ligand to compete with the dominant ligand at a later time. These results show that the efficient sampling of ligand binding orientations in the λ-dynamics method helps alleviate the restriction that the initial orientation of the ligand inside the binding pocket must be close to its true bound orientation in order to get a reasonable estimate of binding free energy − a prerequisite for other free energy calculation methods. This feature is particularly important in drug lead discovery and optimization when the binding mode is unknown, or the modification of ligands causes a change of binding mode. It is also important in the case where a single ligand could have multiple binding modes. It may be true that enlarging the sampling space delays convergence. However, many ligands compete at the same time in the λ-dynamics method, and most of the time one of them reaches a low energy state and competes with the dominant ligand. It may be argued that the enlarged sampling space obtained by scaling of the potential energies in the λ-dynamics method may contain the ligands in unphysical conformations or orientations. This phenomenon was observed in λ-dynamics without the restraining potential and resulted in slowed convergence.[13] By using this potential and the scaling parameter α (eq 5), one can control the extent to which the sampling space of ligand orientations and conformations is enlarged. These merits make it possible to achieve from a single simulation not only calculations of relatively correct binding free energy differences but also the identification of alternative binding orientations and conformations in some ligands.

The alternative binding orientations detected by the λ-dynamics method were restricted to those that did not include a large conformational change in the protein structure because the interaction energy within environmental atoms was not scaled by λ^2. The restriction of enhanced sampling only for the ligands is one of the merits in the λ-dynamics method. Expansion of the enhanced sampling region to environmental atoms may result in highly nonphysical conformation of the protein structure leading to slow convergence of $\Delta\Delta G$. Although a larger sampling space delays convergence, enhanced sampling of environmental atoms may be possible by redefining some of the environmental atoms as multiple conformations that are scaled by using a second coupling parameter.

The λ-dynamics method may decrease the computational time for the equilibration phase and the dependency of final results on the accuracy of initial ligand orientations and conformations in free energy simulations due to the nature of the extended system hybrid potential (eq 1). Although we have demonstrated in this work that the gain in sampling efficiency outweighed the problem of an enlarged sampling space (when the restraining potential with a proper scaling parameter, α, was used), studies on other systems need to be done to strengthen this point. Currently, we are applying this method to a larger set of compounds with more torsional degrees of freedom.

References and Notes

(1) Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 457−467.

(2) Kubinyi, H. *Drug Discovery Today* **1997**, *2*, 538−546.

(3) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *92*, 18−431.

(4) Fleischman, S. H.; Brooks, C. L., III. *Proteins: Structure, Function, Genetics* **1990**, *7*, 52−61.

(5) Guo, Z.; Brooks, C. L., III. *J. Am. Chem. Soc.* **1998**, *120*, 1920−1921.

(6) Essex, J. W.; Severance, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663−9669.

(7) Novotny, J.; Bruccoleri, R. E.; Davis, M.; Sharp, K. A. *J. Mol. Biol.* **1997**, *268*, 401−411.

(8) Kuntz, I. D.; Blaney, J. M.; Oatley, S.; Langridge, R.; Ferrin, T. *J. Mol. Biol.* **1982**, *161*.

(9) DesJarlais, R. L.; Sheridan, R. P.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. *J. Med. Chem.* **1986**, *29*, 2149−2153.

(10) Zou, X.; Sun, Y.; Kuntz, I. D. *J. Am. Chem. Soc.* **1999**, *121*, 8033−8043.

(11) Kong, X.; Brooks, C. L., III. *J. Chem. Phys.* **1996**, *105*, 2414−2423.

(12) Guo, Z.; Brooks, C. L., III; Kong, X. *J. Phys. Chem. B* **1998**, *102*, 2032−2036.

(13) Banba, S.; Brooks, C. L., III. *J. Chem. Phys.*, in press.

(14) Mattos, C.; Rasmussen, B.; Ding, X.; Petsko, G. A.; Ringe, D. *Nature Struct. Biol.* **1994**, *1*, 55−58.

(15) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293−2304.

(16) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727−748.

(17) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470−489.

(18) Schoichet, B. K.; Stroud, R. M.; Santi, D. V.; Kuntz, I. D.; Perry, K. M. *Science* **1993**, *259*, 1445−1450.

(19) Rutenber, E.; Fauman, E. B.; Keenan, R. J.; Fong, S.; Furth, P. S.; Ortiz de Montellano, P. R.; Meng, E.; Kuntz, I. D.; DeCamp, D. L.; Salto, R.; Rose, J. R.; Craik, C.; Stroud, R. M. *J. Biol. Chem.* **1993**, *268*, 15343–15346.

(20) Vieth, M.; Hirst, J. D.; Kolinski, A.; Brooks, C. L., III. *J. Comput. Chem.* **1998**, *19*, 1612–1622.

(21) Vieth, M.; Hirst, J. D.; Dominy, B. D.; Daigler, H.; Brooks, C. L., III. *J. Comput. Chem.* **1998**, *19*, 1623–1631.

(22) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

(23) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins*: *Structure Function*, *Genetics* **1986**, *1*, 47–59.

(24) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

(25) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122–129.

(26) Schaefer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100*, 1578–1599.

(27) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190–1197.

(28) Fitzgerald, M. M.; McRee, D. E.; Churchill, M. J.; Goodin, D. B. *Biochemistry* **1994**, *33*, 3807–3818.

(29) Fitzgerald, M. M.; Trester, M. L.; Jensen, G. M.; McRee, D. E.; Goodin, D. B. *Protein Sci.* **1995**, *4*, 1844–1850.

(30) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B. W.; Reiher III, E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(31) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(32) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(33) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(34) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843–865.

(35) Musah, R. A.; Jensen, G. M.; Rosenfeld, R. J.; McRee, D. E.; Goodin, D. B.; Bunte, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 9083–9084.

(36) Liu, Z.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 6071–6077.