Molecular Dynamics Free Energy Calculations to Assess the Possibility of Water Existence in Protein Nonpolar Cavities

Masataka Oikawa[†‡*] and Yoshiteru Yonetani[†*]
[†]Computational Biology Group, Quantum Beam Science Directorate, Japan Atomic Energy Agency, Kyoto, Japan; and [‡]Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan

ABSTRACT Are protein nonpolar cavities filled with water molecules? Although many experimental and theoretical investigations have been done, particularly for the nonpolar cavity of IL-1β, the results are still conflicting. To study this problem from the thermodynamic point of view, we calculated hydration free energies of four protein nonpolar cavities by means of the molecular dynamics thermodynamic integration method. In addition to the IL-1β cavity (69 Å^3), we selected the three largest nonpolar cavities of AvrPphB (81 Å^3), Trp repressor (87 Å^3), and hemoglobin (108 Å^3) from the structural database, in view of the simulation result from another study that showed larger nonpolar cavities are more likely to be hydrated. The calculations were performed with flexible and rigid protein models. The calculated free energy changes were all positive; hydration of the nonpolar cavities was energetically unfavorable for all four cases. Because hydration of smaller cavities should happen more rarely, we conclude that existing protein nonpolar cavities are not likely to be hydrated. Although a possibility remains for much larger nonpolar cavities, such cases are not found experimentally. We present a hypothesis to explain this: hydrated nonpolar cavities are quite unstable and the conformation could not be maintained.

INTRODUCTION

Interiors of proteins are not completely filled with protein atoms, but have several cavities (1). In some cases, protein cavities contain water molecules (1–3), and such internal water molecules can play important roles in many biological processes, e.g., in redox reaction (4), electron (5,6) and proton transfer (7,8) in proteins, and catalysis of substrates (9,10). Internal water molecules also affect the thermodynamic stability of proteins. A systematic mutation analysis of Ile to Ala (11) showed that inserting one water molecule into cavities created by mutation generally produces a free energy gain of 1–2 kcal/mol. There are many other experimental (12–14) and theoretical (12,15–24) studies on internal water molecules that indicate the biological importance of water.

Despite many characterization studies on protein cavities performed so far, little is known about the cases in which no polar sites (i.e., atoms serving as a hydrogen-bond donor or acceptor) appear on the cavity surface. Whether such particular cavities, referred to as nonpolar cavities, contain water molecules remains uncertain. An analysis of 121 protein structures that included high-resolution data indicated that no water molecules were found for any of the cavities that lacked surface polarity (3). However, the result has to be carefully understood that even if no water molecules are observed crystallographically, another possibility remains. Water molecules in nonpolar cavities may be in a dynamic, disordered state without any directional interactions such as hydrogen bonds (25,26). It will be difficult to detect such water molecules from conventional x-ray or neutron crystal analyses.

Interleukin 1β (IL-1β) is well known for the large nonpolar cavity ~69 Å^3 (note that all cavity volumes shown hereafter are based on the Connolly surface calculations that we performed; see Materials and Methods). Several experiments have been performed to measure the state of this cavity, but the results are not consistent with each other; different experiments led to different conclusions on the existence of water in the cavity. In 1995, Ernst et al. (26) reported that several signals were detected by water-selective two-dimensional nuclear magnetic resonance (NMR) spectroscopy, which implied the existence of water molecules. In 1999, Yu et al. (25) maintained (based on an electron density analysis of the cavity) that two water molecules reside there. However, Quillin et al. (27), in 2006, reached the opposite conclusion: independent analysis of the electron density showed that the nonpolar cavity of IL-1β was not hydrated. And in other x-ray (25,27–31) and NMR (32) structure data, no water molecules were observed in the cavity. Thus, it remains controversial whether the nonpolar cavity of IL-1β is hydrated.

Because it is currently difficult to draw a definite conclusion from experiments, a theoretical approach by molecular simulation will be useful. Thermodynamic stability of cavity hydration can be theoretically evaluated by molecular simulation techniques, which have been applied to various protein cavities, i.e., BPTI (22), Barnase (22), tetrabrachion (16), and T4-lysozyme (12). However, all of these cavities were polar. Nonpolar cavities have not been studied extensively so far. Somani et al. (33) and Zhang and Hermans (34) carried out molecular dynamics (MD) simulations on

Submitted July 27, 2009, and accepted for publication January 14, 2010.

*Correspondence: oikawa@cp.kyoto-u.ac.jp or yonetani.yoshiteru@jaea.go.jp

Editor: Kathleen B. Hall.

0006-3495/10/06/2974/10 $2.00

doi: 10.1016/j.bpj.2010.01.029

the nonpolar cavity of IL-1β. However, these studies cannot resolve the controversies about the nonpolar cavity either, because their results were different from each other. Somani et al. showed that four water molecules stably exist within the cavity, whereas Zhang and Hermans showed that the cavity is empty. Vaitheeswaran et al. (35) performed Monte Carlo simulations to calculate hydration free energies of an ideal nonpolar cavity of spherical shape and fullerenes C_{140} and C_{180}. Their result showed that hydration of the nonpolar sphere becomes stable only when its diameter is >10 Å (we calculated the volume to be ~137 Å^3; see Materials and Methods). These sufficiently large cavities are able to contain at least three-to-four water molecules, so the water molecules can stably exist within the cavity by forming a hydrogen-bonded cluster.

Size dependence of the hydration free energy reported by Vaitheeswaran et al. (35) has provided valuable insight into the possibility of hydration of nonpolar cavities. According to their result of size dependence, hydration of the IL-1β nonpolar cavity with the volume ~69 Å^3 is judged to be unstable. However, their calculation was performed using an ideal, solid spherical cavity, and such calculation did not consider the nature of protein cavities, i.e., their anisotropy and flexibility. Hydration of nonpolar cavities in real proteins should be carefully addressed, in order to permit consideration of these effects.

Here, to answer the question of whether hydrated nonpolar cavities exist in proteins, we calculated the hydration free energies of protein nonpolar cavities by using MD simulations that were themselves based on a model that took the structural anisotropy and flexibility of protein cavities into account. In addition to the cavity of IL-1β, the three largest protein nonpolar cavities were chosen from the structural database: *Pseudomonas* avirulence protein (AvrPphB), Trp repressor, and trematode hemoglobin. These cavities have the volumes 81, 87, and 108 Å^3, respectively, which belong to the largest class of existing protein nonpolar cavities. As shown by the result of Vaitheeswaran et al. (35), larger cavities are more likely to be hydrated, although even our largest cavity (108 Å^3) is insufficient for stable hydration.

In this study, we performed MD simulations using two protein models, namely, flexible and rigid. The flexible model allows protein atoms around the cavity to fluctuate and move, but the rigid one does not. From our results for free energies, we were able to clarify the effects of anisotropy and flexibility on the cavity, and found that these effects were not so large as to cause any essential change of the hydration thermodynamics. This means hydration of the four largest cavities studied here are energetically unfavorable. For smaller cavities, hydration is even less probable. For larger nonpolar cavities, a possibility remains; however, such hydration is not greatly expected, as discussed later. With this conclusion in mind, we again discuss the unsolved problem with the IL-1β nonpolar cavity.

MATERIALS AND METHODS

Selection of proteins with a large nonpolar cavity

Four nonpolar cavities of IL-1β, AvrPphB, Trp repressor, and hemoglobin were chosen as follows. To begin, we prepared a data set of 1371 protein structures that included 100 structures selected by Word et al. (36) and a second data set of 1271 structures from the PISCES server (37) (results obtained July 21, 2007). The first data set was used because all the structures were measured at high resolution (<1.7 Å). The second data set of structures was selected to accord with the criteria of 20% sequence identity, 1.6 Å resolution, and *R*-factor 0.25. Because 12 structures were found to be in both data sets, a total of 1359 nonredundant structures were actually used.

Next, the 1359 structures were analyzed by the Molecular Surface Package (MSP) (38) to search for the cavities. Note that a "protein cavity" was defined as a space that is enclosed by the Connolly surface (39) and is separated from the protein exterior. The Connolly surface was calculated using a probe sphere of a radius 1.4 Å. Atomic radii 1.54, 1.45, and 1.57 Å were used for hydroxyl and carbonyl oxygen atoms, carboxyl oxygen atoms, and peptide nitrogen atoms, respectively. For other atoms, a default setting (38) was used. Before executing the MSP program, protein hydrogen atoms and any molecules other than proteins, such as crystallographic water, were removed.

From the MSP analysis, we found 5948 cavities; among them, 859 had no heavy atoms other than carbon on the surface. We defined them as "nonpolar cavities". Volumes of these nonpolar cavities were 12–108 Å^3 (Fig. 1). We chose the three largest cavities of AvrPphB (81 Å^3, 1UKF (40)), Trp repressor (87 Å^3, 2WRP (41)), and hemoglobin (108 Å^3, 1H97 (42)), considering that larger cavities are more expected to be hydrated (35). In addition, the large nonpolar cavity of IL-1β (69 Å^3) was selected

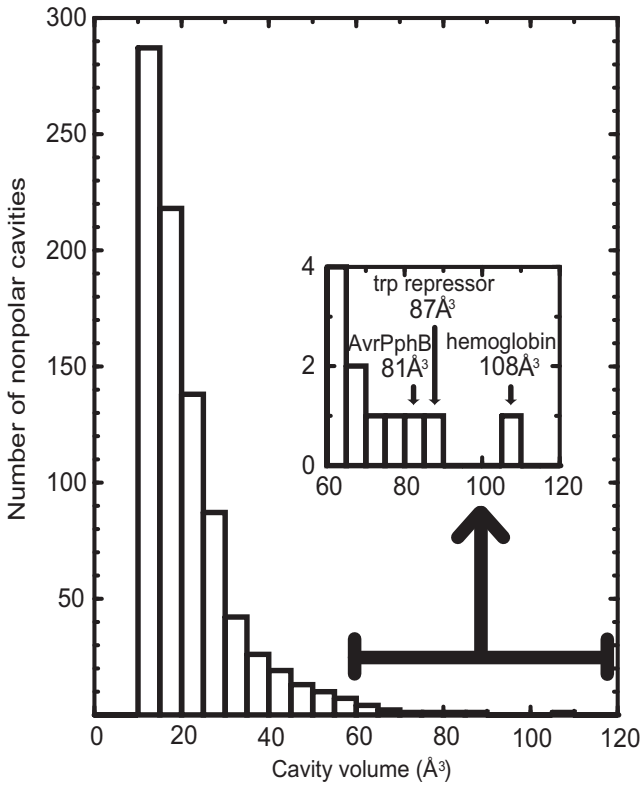

FIGURE 1 Distribution of volumes of 859 nonpolar cavities found using the MSP (38) (see Materials and Methods). (*Inset*) Enlarged view for the range from 60 to 120 Å^3.

because of many reports about this cavity (25–27). Several structural data of IL-1β are deposited in PDB. We used the structure of 9ILB (25) in accordance with the previous theoretical work (33). We carefully checked the neighbor regions of these cavities, and confirmed that all polar sites are at least 1.9 Å away from the cavity surface. Thus, internal water molecules cannot make a hydrogen bond unless the cavity changes structurally (shown later in Results).

Shapes and locations of the four selected nonpolar cavities are shown in Fig. 2. The nonpolar cavity of IL-1β is located at the center of the protein. The other cavities are located near the protein surface. The hemoglobin cavity has a dumbbell shape, while the others are nearly spherical. The cavity of the Trp repressor is located in the interface of the homodimer, so it has a twofold rotational symmetry.

Method for the free energy calculation

Hydration free energy of protein cavities was evaluated by the statistical mechanics formulation proposed by Roux et al. (15). We assume a protein cavity that would be large enough to contain n water molecules. Using the Roux's formulation, we can calculate the free energy change of transferring n water molecules from a bulk region to the protein cavity, $\Delta G_{\mathrm{hyd}}^{0\to n}$, as

$$\Delta G_{\mathrm{hyd}}^{0\to n} = \Delta G_{\mathrm{cavity}}^{0\to n} - n\Delta G_{\mathrm{bulk}}^{0\to 1} - nk_{\mathrm{B}}T\ln\left[\rho_{\mathrm{bulk}}(2\pi k_{\mathrm{B}}T/k_{\mathrm{harm}})^{3/2}\right] + k_{\mathrm{B}}T\ln n!, \tag{1}$$

where k_{B} is the Boltzmann constant, T is the temperature, and ρ_{bulk} is the density of water in the bulk phase. The value $\Delta G_{\mathrm{cavity}}^{0\to n}$ is the free energy difference between following two protein-water systems. One system is given by the potential

$$U_0 + u^{\mathrm{res}},$$

in which n water molecules have no interactions with the rest of the system, but are weakly restrained around an internal cavity site $\mathbf{r}^*$ with the harmonic potential acting on the position of water oxygen $\mathbf{r}_{\mathrm{i}}$, as

$$u^{\mathrm{res}} = \sum_{i=1}^{n}\left(1/2\right)k_{\mathrm{harm}}(\mathbf{r}_{\mathrm{i}} - \mathbf{r}^*)^2.$$

Another system is given by a potential U_1, in which n water molecules interact with the rest of the system and reside within the cavity. Using these potentials, $\Delta G_{\mathrm{cavity}}^{0\to n}$ is written by

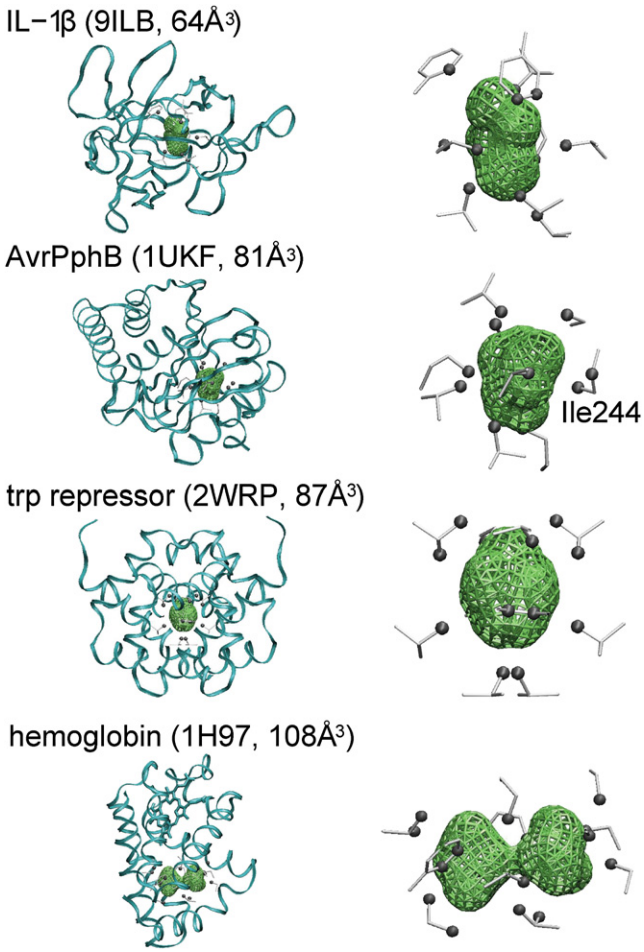

FIGURE 2 Cavity locations of the four selected proteins (*left*) and the shapes of nonpolar cavities (*right*). Cavity surfaces are shown by green mesh. Side chains of residues, which construct the cavities, are shown by gray sticks. Among them, atoms exposed to the cavity surface are shown by gray spheres. The figures were created by using VMD (62).

$$\Delta G_{\mathrm{cavity}}^{0\to n} = -k_{\mathrm{B}}T\ln\frac{\int_{\mathrm{cavity}}\mathrm{d}\mathbf{r}_1\cdots\int_{\mathrm{cavity}}\mathrm{d}\mathbf{r}_n\int_{\mathrm{bulk}}\mathrm{d}\mathbf{r}_{n+1}\cdots\int_{\mathrm{bulk}}\mathrm{d}\mathbf{r}_{\mathrm{N}}\int d\Gamma e^{-U_1/k_{\mathrm{B}}\mathrm{T}}}{\int\mathrm{d}\mathbf{r}_1\cdots\int\mathrm{d}\mathbf{r}_n\int_{\mathrm{bulk}}\mathrm{d}\mathbf{r}_{n+1}\cdots\int_{\mathrm{bulk}}\mathrm{d}\mathbf{r}_N\int d\Gamma e^{-\left[U_0+\sum_1^n(1/2)k_{\mathrm{harm}}(\mathbf{r}_{\mathrm{i}}-\mathbf{r}^*)^2\right]/k_{\mathrm{B}}\mathrm{T}}}, \tag{2}$$

where N is the total number of water molecules in the cavity and bulk regions, and $\int d\Gamma$ denotes the integral over the configurations of atoms other than water (i.e., protein and ions).

The value $\Delta G_{\mathrm{cavity}}^{0\to n}$ was calculated by the thermodynamic integration passing through the intermediate systems between the systems U_1 and $U_0+ u^{\mathrm{res}}$. In this study, we change a system from U_1 to $U_0 + u^{\mathrm{res}}$ by following three processes:

$$\left(U_1^{\mathrm{ele}}, U_1^{\mathrm{vdW}}, 0\right)\to\left(U_1^{\mathrm{ele}}, U_1^{\mathrm{vdW}}, u^{\mathrm{res}}\right)\to\left(U_0^{\mathrm{ele}}, U_1^{\mathrm{vdW}}, u^{\mathrm{res}}\right)\to\left(U_0^{\mathrm{ele}}, U_0^{\mathrm{vdW}}, u^{\mathrm{res}}\right). \tag{3}$$

This is achieved by changing three coupling parameters (λ^{ele}, λ^{vdW}, and λ^{res}) for electrostatic, van der Waals, and restraining potentials to be

$$(0,0,0)\to(0,0,1)\to(1,0,1)\to(1,1,1). \tag{4}$$

In the first and second processes, intermediate systems for

$$\lambda^{\mathrm{res}} = \{0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

and

$$\lambda^{\mathrm{ele}} = \{0, 0.25, 0.5, 0.75, 1\}$$

were generated by linearly combining the two end-point systems. In the third process, the nonlinear combination with the soft-core potential (43) was used

to avoid a singularity appearing near $\lambda^{\mathrm{vdW}} = 1$ (44). Intermediate systems were generated for

$$\lambda^{\mathrm{vdW}} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

The value $\Delta G_{\mathrm{bulk}}^{0\to 1}$ in Eq. 1 is the free energy difference between the two bulk water systems; one system requires that all water molecules interact with each other, and the other system requires that one water molecule, out of all of them, does not interact. The water molecule whose interaction changes between the two systems has a fixed center of mass. The value $\Delta G_{\mathrm{bulk}}^{0\to 1}$ was evaluated using the thermodynamic integration as well, but in this case, the following two processes were employed:

$$\left(U_1^{\mathrm{ele}}, U_1^{\mathrm{vdW}}\right) \to \left(U_0^{\mathrm{ele}}, U_1^{\mathrm{vdW}}\right) \to \left(U_0^{\mathrm{ele}}, U_0^{\mathrm{vdW}}\right) \qquad (5)$$

The coupling parameters (λ^{ele}, λ^{vdW}) change as follows:

$$(0,0) \to (1,0) \to (1,1). \qquad (6)$$

Intermediate systems for the $\Delta G_{\mathrm{bulk}}^{0\to 1}$ calculation were set in the same way as in $\Delta G_{\mathrm{cavity}}^{0\to n}$.

The third and last terms in the right-hand side of Eq. 1 are the correction terms. The former is the compensation for introducing the restraining potential (15). The latter considers the configuration overlap of n noninteracting water molecules.

In principle, the resultant free energy $\Delta G_{\mathrm{hyd}}^{0\to n}$ (Eq. 1) does not depend on the restraining potentials (15). However, sampling efficiency for the configurations of restrained water (see Eq. 2) highly depends on the parameters r^* and k_{harm}. In this study, r^* was set to the center of the cavity. The k_{harm} values, 0.23, 0.28, 0.35, and 0.11 kcal/mol Å^2, were used for IL-1β, AvrPphB, Trp repressor, and hemoglobin, respectively. These values were obtained from the relation (15)

$$V = (2\pi k_{\mathrm{B}}T/k_{\mathrm{harm}})^{\frac{3}{2}},$$

so that the whole space of the cavity could be sampled when interactions of internal water molecules were turned off and the restraining potential was applied. Here, V is the volume of a sphere that envelops the cavity.

System preparations and MD simulations

We carried out MD simulations for protein-water and bulk water systems to obtain the free energy changes, $\Delta G_{\mathrm{cavity}}^{0\to n}$ and $\Delta G_{\mathrm{bulk}}^{0\to 1}$ (Eq. 1).

Protein-water systems

For each of the IL-1β, AvrPphB, Trp repressor, and hemoglobin configurations, we prepared four systems that differed in the number of water molecules in the cavity (i.e., $n = 1$, 2, 3, and 4). Crystallographic water molecules were removed from the crystal structures, except in polar cavities. Hydrogen atoms were generated according to the protonation state at pH 7, and missing heavy atoms were modeled by the AMBER7 Leap module (45).

Because no water molecules were assigned for the nonpolar cavities considered here, we incorporated them into the cavities as follows. The first water molecule was placed at the center of the cavity. Further water molecules, ≤ 4, were sequentially inserted near the first one. Then, energy minimization was performed on each arrangement of water.

Sodium counterions, i.e., one for IL-1β, nine for AvrPphB, six for Trp repressor, and five for hemoglobin, were added for system neutralization. Approximately 6000–8000 water molecules were placed around the proteins to maintain at least 8 Å hydration layers, and then a three-dimensional periodic boundary condition was imposed.

The systems of IL-1β, AvrPphB, and Trp repressor were modeled with the OPLS-AA (46,47) and TIP3P parameters (48). The van der Waals interactions were calculated with cutoff of 9 Å, and the electrostatic interactions were calculated by the particle-mesh Ewald (49) with real-space cutoff of 9 Å. Only the hemoglobin was modeled with the GROMOS 53A6 parameters (50) because the heme parameters were not included in the OPLS-AA. This system was simulated in the GROMOS default setting (50), with simple point charge (SPC) water (51) and twin-range cutoff of 10 and 14 Å, for van der Waals force evaluation.

For each system, simulations using both rigid and flexible models were performed. All protein atoms were fixed in the rigid model, whereas in the flexible model they were allowed to move. However, even in the flexible case, weak harmonic potentials with a force constant of 2 kcal/mol Å^2 were applied to the protein backbone atoms (which are >15 Å apart from the center of the cavities) to enable us to avoid translational displacement of the proteins.

All MD and free energy calculations were performed using the GROMACS 3.3.1 package (52,53). MD simulations were started from the system where n internal water molecules have full interactions, i.e.,

$$\left(\lambda^{\mathrm{ele}}, \lambda^{\mathrm{vdW}}, \lambda^{\mathrm{res}}\right) = (0,0,0).$$

After steepest-descent minimization of 2000 steps, 110-ps MD simulations were carried out at 300 K and 1 atm. Subsequently, the system was updated along the path of λ (see Eqs. 3 and 4) by gradually reducing the interactions of n water. At each stage of λ, 10-ps equilibration and 100-ps production runs were carried out. We confirmed the validity of this trajectory length by additionally performing the calculation with different settings, 5-ps equilibration + 50-ps production and 20-ps equilibration + 200-ps production (data not shown). At the final stage of (λ^{ele}, λ^{vdW}, λ^{res}) = (1,1,1), the interactions of n water were removed completely. In this way, $\Delta G_{\mathrm{cavity}}^{0\to n}$ for any case of $n = 1$–4 was directly calculated by reducing the interactions of n water molecules simultaneously, rather than one by one.

Atomic trajectories were generated by numerically solving the equation of motion with the 2-fs time step. Temperature and pressure were controlled by the Langevin thermostat and the Berendsen barostat (54), respectively. However, it should be noted that in the GROMACS 3.3.1 package, options of fixed atoms and pressure control cannot be used at the same time. Thus, the rigid MD simulations were performed with the volume fixed at a value obtained from the corresponding flexible MD simulation in which the pressure was controlled at 1 atm.

Bulk water systems

$\Delta G_{\mathrm{bulk}}^{0\to 1}$ (Eq. 1) was evaluated for two cases of TIP3P and SPC. The setting for calculating intermolecular interactions was the same as in the protein-water systems. The TIP3P and SPC water systems were constructed using 512 and 1000 water molecules, respectively. A three-dimensional periodic boundary condition was imposed on each system. The larger number of SPC water molecules is due to the usage of the long distance twin-range cutoff. After the simulation at (λ^{ele}, λ^{vdW}) = (0, 0), the system was updated by changing λ (see Eq. 5). At each stage of λ, 1-ns equilibration and 10-ns production runs were performed at 300 K and 1 atm. Other MD settings such as those for force evaluation and temperature and pressure controls, were the same as in the protein-water simulations.

RESULTS

Rigid-model calculation: comparison with the previous results with a spherical cavity

Cavity hydration free energies obtained from the rigid-model calculation are shown in Fig. 3 (see the legend for several missing points, e.g., Trp repressor). For comparison, previous results by Vaitheeswaran et al. (35) are also shown. The previous calculations were performed on an ideal, rigid nonpolar sphere, i.e., interaction sites are uniformly distributed on the sphere. Thus, the main difference between the

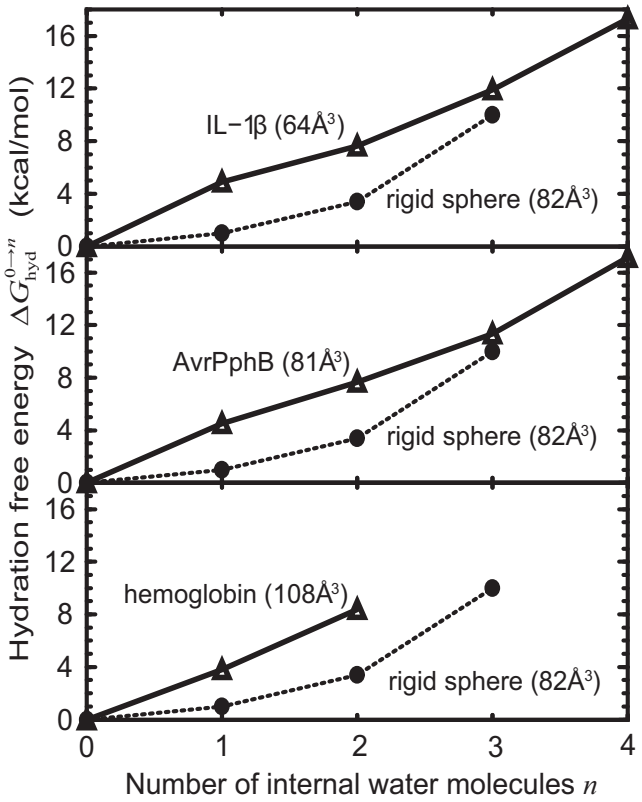

FIGURE 3 Cavity hydration free energies from the rigid-model calculations. For comparison, previous data from a spherical model (taken from Fig. 1 in (35)) are shown (*solid circles*). Results for Trp repressor are not shown, because water molecules got out of the cavity during the MD simulation. For the same reason, data for hemoglobin ($n = 3$ and 4) are missing from the figure. Details of the free energies are given in Table S1 in the Supporting Material.

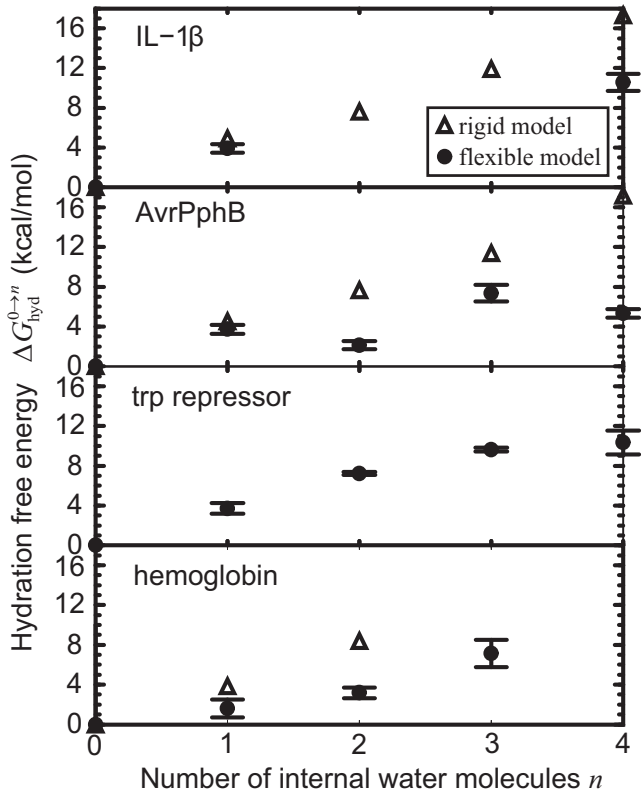

FIGURE 4 Cavity hydration free energies: comparison between the rigid- and flexible-model results. Data for IL-1β ($n = 3$ and 4) could not be evaluated because the water molecules got out of the cavity during the MD simulations. Errors of the free energies were estimated from three independent simulations performed with different random seed numbers for the Langevin thermostat. The errors were <0.3 kcal/mol in any case of the rigid-model results. Details of the free energies are given in Table S2 in the Supporting Material.

present and previous models is the shape of the cavities. As shown in Fig. 3, any insertion of water molecules increases the free energy monotonically, showing that hydration of nonpolar cavities is not favorable. Results from both models qualitatively agree with each other; the previous sphere model (35) is enough to make rough estimates of cavity hydration thermodynamics despite the simplicity. Small differences remain between both of the results, which may be due, in part, to the difference in the force field used.

Flexible-model calculation: comparison between the flexible and rigid models

Cavity hydration free energies from the flexible-model calculations are shown in Fig. 4 (some points of the data are not shown as in the rigid case; see Fig. 4 legend), together with the rigid-model results shown above (Fig. 3). The flexible model yielded lower free energies than the rigid model; however, they were all positive, which indicated that the hydration of the nonpolar cavities is unfavorable. The distinct feature from the rigid case is that the flexible model exhibits irregular changes (e.g., AvrPphB). The free energy does not monotonically increase as the number of water molecules increase. Note that these results were reproducible, so it is not due to calculation error. Independently performed simulations with different length of trajectories and different initial conditions provided almost the same results. (Protein conformational changes may occur on much longer nanosecond-to-millisecond timescales (55,56). Strictly speaking, the free energy should be obtained by averaging over all these conformations. We should mention, however, that our free energies in Fig. 4, which were calculated from only conformations near the starting x-ray structure, do not include contributions of this slow fluctuation.) We will mention the reason for the nonmonotonic increase in the last part of this section.

As expected, the flexible model induces some changes in interactions between water and cavity wall, leading to suppression of the increase in free energy. To check the free energy contribution of interactions between water and cavity, we calculated the potential energy difference $\Delta E^{(\mathrm{flexible})} - \Delta E^{(\mathrm{rigid})}$ from 1-ns MD trajectories, where ΔE is the potential

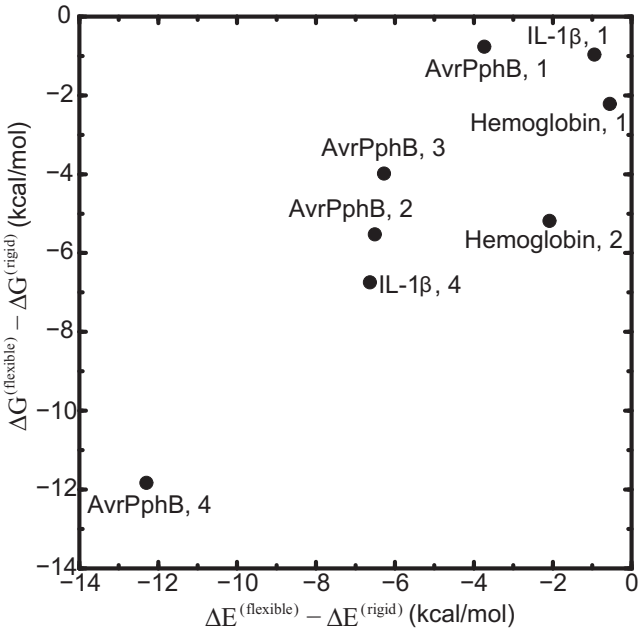

FIGURE 5 Energy differences between flexible and rigid models. ΔG is the cavity hydration free energy (Figs. 3 and 4) and ΔE is in the potential energy due to the water-cavity wall interactions. Each point has a label for the protein name and the number of internal water molecules n. Data for Trp repressor were not obtained (see the caption in Fig. 3).

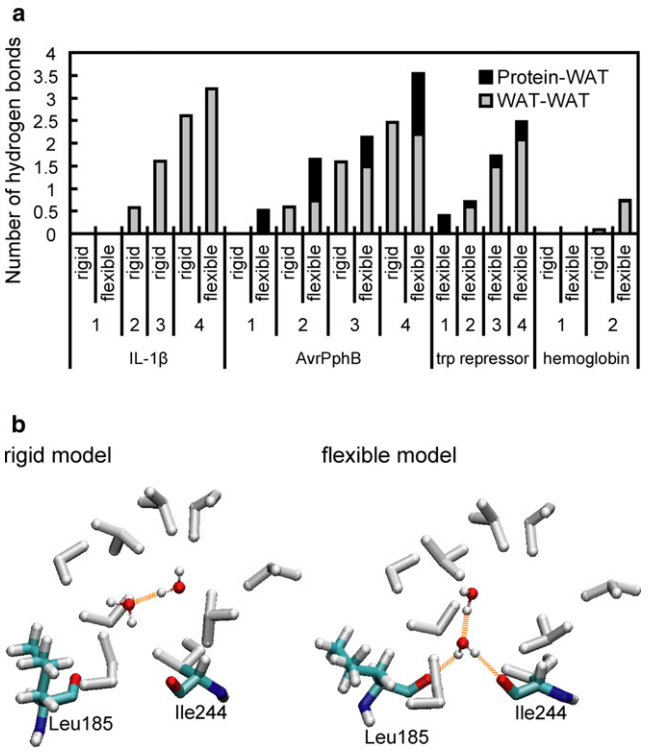

FIGURE 6 (*a*) Number of hydrogen bonds associated with the internal water molecules. Water-water and protein-water hydrogen bonds are shaded and solid, respectively. These were obtained by averaging over 1-ns trajectories. (*b*) Typical snapshots from the rigid (*left*) and flexible (*right*) simulations of AvrPphB ($n = 2$). Hydrogen bonds are represented by orange dashed lines. Protein atoms near the cavity surface are shown. Main-chain carbonyl oxygen atoms of Leu^{185} and Ile^{244} form a hydrogen bond to the water molecule in the cavity (*right*). The figures were created by using VMD (62).

energy due to the interactions between internal water molecules and the surrounding residues. Residues were included in this analysis if any atom existed within 8 Å of the center of the cavity in x-ray crystal conformation. In Fig. 5, we show the $\Delta E^{(\text{flexible})} - \Delta E^{(\text{rigid})}$ as well as $\Delta G^{(\text{flexible})} - \Delta G^{(\text{rigid})}$ (here, $\Delta G^{(\text{flexible})}$ and $\Delta G^{(\text{rigid})}$ denote the flexible and rigid $\Delta G_{\text{hyd}}^{0\rightarrow n}$, respectively). We can see that both values are almost the same amount, e.g., for IL-1β ($n = 4$), $\Delta E^{(\text{flexible})} - \Delta E^{(\text{rigid})} = -6.63$ kcal/mol, while $\Delta G^{(\text{flexible})} - \Delta G^{(\text{rigid})} = -6.75$ kcal/mol. This component of potential energy substantially contributes to the free energy difference between the flexible and rigid models.

We further found that the above observation comes from positional displacement of cavity-forming residues. Some parts of the cavity wall were replaced by polar atoms, and internal water molecules often made hydrogen bonds to the exposed atoms. This means that nonpolarity of the cavity cannot be maintained. We show the formation of the hydrogen bonds in Fig. 6 *a*, where the average numbers of protein-water and water-water hydrogen bonds calculated using 1-ns trajectories are given. No protein-water hydrogen bonds were found in rigid cases, whereas they were frequently formed in flexible cases. Thus, the flexible cavities had more hydrogen bonds than did the rigid cavities. We give a graphical example in Fig. 6 *b*. Only a water-water hydrogen bond was formed in the rigid cavity (Fig. 6 *b*, *left*). However, in the flexible cavity, not only the water-water hydrogen bond but also two hydrogen bonds were formed between the water and the cavity (Fig. 6 *b*, *right*). These additional hydrogen bonds are thought to help the reduction of free energy in the flexible model.

The nonmonotonic increase of free energy, which was prominently observed in the case of AvrPphB (Fig. 4), can be also explained in terms of the number of protein-water hydrogen bonds (Fig. 6 *a*). AvrPphB with $n = 2$ and 4 had approximately one more protein-water hydrogen bond than those with $n = 1$ and 3. It is probable that these additional hydrogen bonds reduce the free energy for $n = 2$ and 4, which results in the obtained unusual free energy change noted above.

DISCUSSION

Hydration free energies for the nonpolar cavities of IL-1β (69 $Å^3$), AvrPphB (81 $Å^3$), Trp repressor (87 $Å^3$), and hemoglobin (108 $Å^3$) were evaluated to assess the possibility of hydration of nonpolar cavities. The rigid-model calculation showed that changes in free energy were positive for all

the cavities examined, indicating that these nonpolar cavities are not hydrated stably. This result is essentially consistent with the previous result from an ideal, spherical cavity (35), as it is shown in the previous study that nonpolar cavities of these sizes (69–108 Å^3) are not expected. By the flexible model, positive values decreased, but the overall tendency did not change; the free energy changes were still positive. Therefore, the probability of water molecules residing in the protein nonpolar cavities is low, regardless of the cavity models. These four nonpolar cavities were the largest ones among existing protein nonpolar cavities. For smaller nonpolar cavities, hydration becomes less probable (35). Therefore, hydration of protein nonpolar cavities is unlikely in general.

Suppose there were still-larger nonpolar cavities than those studied here. Could hydration then occur? Vaitheeswaran et al. (35) predicted that very large nonpolar cavities (>~137 Å^3) can be hydrated by the making of a hydrogen-bonded water cluster. However, according to our studies, nonpolar cavities of this size are not found experimentally (see Fig. 1). We suppose the reason is as follows. If the cavity was rigid, hydration could be achieved by making a water cluster, as suggested previously (35). However, hydrated protein nonpolar cavities will be quite unstable, as our MD simulations indicated; the nonpolar cavities often changed their shape and became polar. In other words, if hydrated nonpolar cavities were given, they could not be maintained owing to their flexibility. This scenario reasonably explains why water-clustering hydration is not found in nonpolar cavities of real proteins. Our hypothesis is strengthened by the fact that so many hydrated polar cavities, of such large size, are found (see Fig. 7 *a*). If our hypothesis is correct, such hydration will never be found even if the structural database is updated in the future.

Water molecules in cavities of crystal structures

In this study, we have shown that any water insertion to nonpolar cavities is thermodynamically unfavorable, which suggests a low probability of water residence in protein nonpolar cavities. This conclusion is consistent with our statistical analysis (Fig. 7), which shows the distribution of polar/nonpolar cavities with/without water molecules in 1359 proteins (see Materials and Methods). Water molecules are very often found in polar cavities, whereas they are scarcely observed in nonpolar cavities. Exceptionally, 10 nonpolar cavities had a single water molecule, but these did not turn out to be strictly nonpolar; by checking these cases in detail, we found that polar atoms exist in sites deeply buried below the cavity surface. The distances from the internal water were <~3.8 Å, and thus, hydrogen bonding is possible. These 10 cases were found to be ''nonpolar'' under the present definition that is based only on the atoms on the cavity surface (see Materials and Methods); however, they actually did possess polarity.

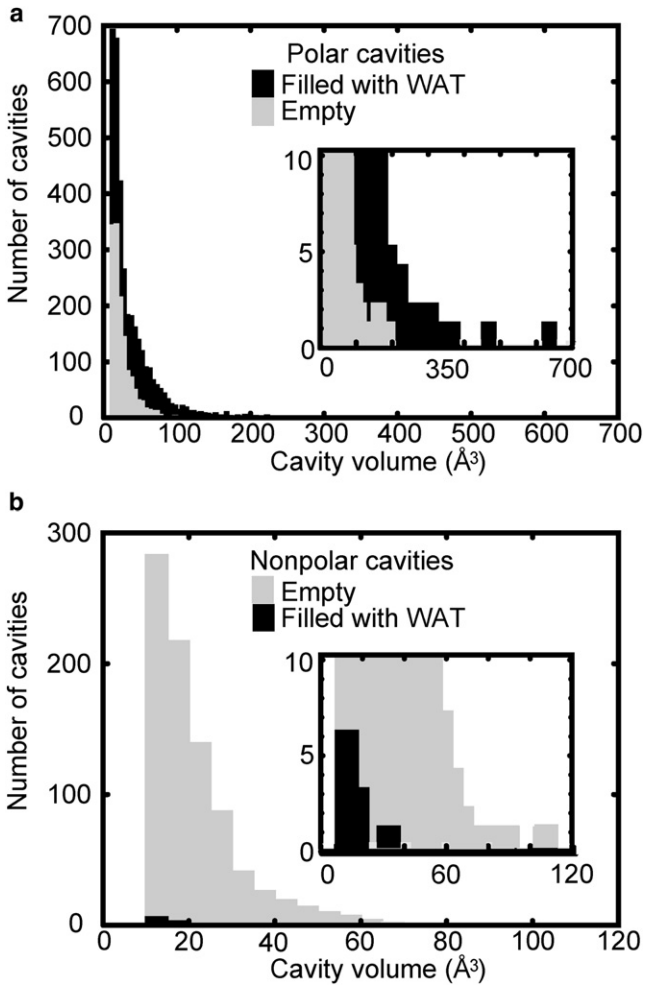

FIGURE 7 Distribution of the volume of 5948 protein cavities; (*a*) polar and (*b*) nonpolar cavities. Cavities with and without water molecules are solid and shaded, respectively. One-hundred-eighty-three polar cavities, which were found to be filled with molecules other than water, are not shown. In the insets, ordinates are enlarged.

Reconsideration about the nonpolar cavity of IL-1β: Is it hydrated or not?

The nonpolar cavity of IL-1β has been most frequently examined among the four cavities studied here; however, there is still controversy about the state within the cavity. Quillin et al. (27) analyzed electron density within the cavity and indicated that the nonpolar cavity is not hydrated. Indeed, in atomic coordinates solved by other x-ray (25,27–31) and NMR (32) experiments, no water molecules are assigned for the nonpolar cavity. Our calculations also lead to the same conclusion. Contrary to this, two results have been reported. By water-selective two-dimensional NMR spectroscopy, Ernst et al. (26) reported several signals that would imply that the cavity is hydrated. Yu et al. (25) also maintained that, by integrating electron density within the cavity, the cavity is hydrated. It is, however, pointed out by Quillin et al. (27) and Matthews et al. (57) that some problematic

points are involved in these interpretations. In the analysis of Yu et al. (25), a spherical region of radius 6 Å was considered as the cavity, but the volume of 905 $Å^3$ was too large compared to the actual cavity size of 69 $Å^3$. Thus, the number of electrons within the cavity may be overestimated (27). As for the NMR signals detected by Ernst et al. (26), we cannot deny the possibility that water molecules at other sites are attributed to the signals (57).

It may be important to note the following point. As shown by the results of free energies, stable hydration is not allowed energetically, but water molecules can visit the cavity at least transiently. Nevertheless, this consideration does not explain the current discrepancy about the water existence, as the probability of water residence is still very low. Our estimated free energy cost of hydrating the IL-1β cavity was >~4 kcal/mol (Figs. 3 and 4), so the probability of water residence is <0.1% under physiological conditions, although the water existence supposed by Ernst et al. (26) and Yu et al. (25) was meant for a stable residence such as probability ~70% (25).

The state of the IL-1β cavity has been studied theoretically as well. Somani et al. (33) calculated cavity hydration free energies by using an approximate representation based on the cumulant expansion. The free energy changes were 7.8, 5.4, 9.0, and −3.6 kcal/mol for 1–4 water molecules. From the negative value of −3.6 kcal/mol, Somani et al. insisted that the nonpolar cavity of IL-1β is filled with four water molecules. This finding may have an important implication for future studies of nonpolar hydration, but currently we must be careful about the interpretation, as, contrarily, the present calculations provided large positive values for this hydration. Somani et al. (33) estimated the contribution of the fourth water molecule to be −12.6 (= −3.6–9.0) kcal/mol. It seems, however, that this value is too large for the contribution of a single water molecule. According to the free energy data obtained by using a more rigorous method (17), free energy change of adding a single water molecule to a protein is, at most, −11.2 ± 0.5 kcal/mol. The approximation adopted by Somani et al., i.e., that potential energies of internal water molecules have Gaussian distribution, may be a reason for this discrepancy. However, such an assumption was not employed in our free energy calculations.

Another theoretical study on the nonpolar cavity of IL-1β was done by Zhang and Hermans (34). They calculated potential energy changes with water insertion, and suggested that the hydration is not favorable.

A remaining possibility of transient visit of water molecules in concert with protein motion

Proteins structurally fluctuate and sometimes experience a large conformational change. Water molecules in proteins are also dynamic; several experimental (55,56,58,59) and simulation studies (60) show that water molecules in a polar cavity sometimes escape from the cavity and again enter it, coupled with the protein motion. Therefore, our last concern is about a transient, metastable state that appears as the result of a protein conformational change. Suppose that a cavity was usually nonpolar, but transiently achieved some polarity with the conformational change. Could water molecules then visit the temporary space? Although such reversible switching between the stable and metastable states was not obtained in our simulations, we cannot make any clear assessment on this possibility. This is because even if there is a possibility, it is difficult to sample protein conformations structurally far from the starting one with MD simulations. Such an event might be obtained from a simulation with a much longer time. Another possible choice is the strategy of overcharging (8,61), which can drive slow events by introducing artificial changes to atomic charges of a residue. If we can generate a conformational change suitably by this method, energy contributions of associated electrostatic changes will be evaluated.

CONCLUSIONS

To investigate whether hydration happens in protein nonpolar cavities, we calculated hydration free energies of four protein cavities that are the largest among the existing protein nonpolar cavities. The resultant free energies showed that hydration of these nonpolar cavities is not plausible for any case of flexible or rigid models. Together with lower expectation for smaller nonpolar cavities (35), stable hydration is not expected for any existing protein nonpolar cavities. There remains a possibility that water molecules visit a nonpolar cavity transiently, particularly when conformational changes occur in the protein.

If there were still larger nonpolar cavities, does hydration occur? To our knowledge, such large nonpolar cavities have not been found experimentally. The possibility that very large nonpolar cavities are filled with clustered water molecules has been suggested from a simulation using an ideal, solid sphere (35). However, we speculate that as for protein cases, such hydration is not achieved. We observed that nonpolar cavities filled with water underwent deformation and finally became polar, i.e., conformations of the nonpolar cavities are quite unstable. This may be one reason why hydrated nonpolar cavities do not appear in protein interiors.

As for the nonpolar cavity of IL-1β, two opposite notions have been proposed. Our results revealed that, when viewed from the thermodynamic point of view, hydration of the nonpolar cavity is not favored.

SUPPORTING MATERIAL

Two tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00173-6.

We thank Dr. Hidetoshi Kono and Profs. Takeshi Kawabata, Kei Yura, and Nobuhiro Gō for valuable comments and discussion. M.O. acknowledges much support from Profs. Shin Ishii and Takeshi Kawabata, which enabled

him to carry out this work during the PhD program at the Nara Institute of Science and Technology.

M.O. was supported by a fellowship from the Japan Atomic Energy Agency. This work was supported in part by a Grant-in-Aid for Scientific Research to Y.Y. (No. 21107532) from the Ministry of Education, Culture, Sports, Science and Technology in Japan.

REFERENCES

1. Rashin, A. A., M. Iofin, and B. Honig. 1986. Internal cavities and buried waters in globular proteins. *Biochemistry.* 25:3619–3625.
2. Williams, M. A., J. M. Goodfellow, and J. M. Thornton. 1994. Buried waters and internal cavities in monomeric proteins. *Protein Sci.* 3: 1224–1235.
3. Hubbard, S. J., K. H. Gross, and P. Argos. 1994. Intramolecular cavities in globular proteins. *Protein Eng.* 7:613–626.
4. Langen, R., G. M. Jensen, …, A. Warshel. 1992. Protein control of iron-sulfur cluster redox potentials. *J. Biol. Chem.* 267:25625–25627.
5. de la Lande, A., S. Martí, …, V. Moliner. 2007. Long distance electron-transfer mechanism in peptidylglycine α-hydroxylating monooxygenase: a perfect fitting for a water bridge. *J. Am. Chem. Soc.* 129: 11700–11707.
6. Muegge, I., P. X. Qi, …, A. Warshel. 1997. The reorganization energy of cytochrome *c* revisited. *J. Phys. Chem. B.* 101:825–836.
7. Schobert, B., L. S. Brown, and J. K. Lanyi. 2003. Crystallographic structures of the M and N intermediates of bacteriorhodopsin: assembly of a hydrogen-bonded chain of water molecules between Asp-96 and the retinal Schiff base. *J. Mol. Biol.* 330:553–570.
8. Pisliakov, A. V., P. K. Sharma, …, A. Warshel. 2008. Electrostatic basis for the unidirectionality of the primary proton transfer in cytochrome *c* oxidase. *Proc. Natl. Acad. Sci. USA.* 105:7726–7731.
9. Schlichting, I., J. Berendzen, …, S. G. Sligar. 2000. The catalytic pathway of cytochrome p450cam at atomic resolution. *Science.* 287:1615–1622.
10. Warshel, A., F. Sussman, and J. K. Hwang. 1988. Evaluation of catalytic free energies in genetically modified proteins. *J. Mol. Biol.* 201:139–159.
11. Takano, K., Y. Yamagata, and K. Yutani. 2003. Buried water molecules contribute to the conformational stability of a protein. *Protein Eng.* 16:5–9.
12. Collins, M. D., G. Hummer, …, S. M. Gruner. 2005. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. *Proc. Natl. Acad. Sci. USA.* 102: 16668–16671.
13. Ando, N., B. Barstow, …, S. M. Gruner. 2008. Structural and thermodynamic characterization of T4 lysozyme mutants and the contribution of internal cavities to pressure denaturation. *Biochemistry.* 47: 11097–11109.
14. Hayashi, I., H. Mizuno, …, M. Ikura. 2007. Crystallographic evidence for water-assisted photo-induced peptide cleavage in the stony coral fluorescent protein Kaede. *J. Mol. Biol.* 372:918–926.
15. Roux, B., M. Nina, …, J. C. Smith. 1996. Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study. *Biophys. J.* 71:670–681.
16. Yin, H., G. Hummer, and J. C. Rasaiah. 2007. Metastable water clusters in the nonpolar cavities of the thermostable protein tetrabrachion. *J. Am. Chem. Soc.* 129:7369–7377.
17. Barillari, C., J. Taylor, …, J. W. Essex. 2007. Classification of water molecules in protein binding sites. *J. Am. Chem. Soc.* 129:2577–2587.
18. Fischer, S., and C. S. Verma. 1999. Binding of buried structural water increases the flexibility of proteins. *Proc. Natl. Acad. Sci. USA.* 96:9613–9615.
19. Fischer, S., J. C. Smith, and C. S. Verma. 2001. Dissecting the vibrational entropy change on protein/ligand binding: burial of a water molecule in bovine pancreatic trypsin inhibitor. *J. Phys. Chem. B.* 105:8050–8055.
20. Verma, C. S., and S. Fischer. 2005. Protein stability and ligand binding: new paradigms from in-silico experiments. *Biophys. Chem.* 115:295–302.
21. Helms, V., and R. C. Wade. 1998. Hydration energy landscape of the active site cavity in cytochrome P450cam. *Proteins.* 32:381–396.
22. Olano, L. R., and S. W. Rick. 2004. Hydration free energies and entropies for water in protein interiors. *J. Am. Chem. Soc.* 126:7991–8000.
23. Damjanović, A., B. García-Moreno, …, A. E. García. 2005. Molecular dynamics study of water penetration in staphylococcal nuclease. *Proteins.* 60:433–449.
24. Park, S., and J. G. Saven. 2005. Statistical and molecular dynamics studies of buried waters in globular proteins. *Proteins.* 60:450–463.
25. Yu, B., M. Blaber, …, D. L. Caspar. 1999. Disordered water within a hydrophobic protein cavity visualized by x-ray crystallography. *Proc. Natl. Acad. Sci. USA.* 96:103–108.
26. Ernst, J. A., R. T. Clubb, …, G. M. Clore. 1995. Demonstration of positionally disordered water within a protein hydrophobic cavity by NMR. *Science.* 267:1813–1817.
27. Quillin, M. L., P. T. Wingfield, and B. W. Matthews. 2006. Determination of solvent content in cavities in IL-1β using experimentally phased electron density. *Proc. Natl. Acad. Sci. USA.* 103:19749–19753.
28. Finzel, B. C., L. L. Clancy, …, H. M. Einspahr. 1989. Crystal structure of recombinant human interleukin-1 β at 2.0 Å resolution. *J. Mol. Biol.* 209:779–791.
29. Shaanan, B., A. M. Gronenborn, …, G. M. Clore. 1992. Combining experimental information from crystal and solution studies: joint x-ray and NMR refinement. *Science.* 257:961–964.
30. Priestle, J. P., H. P. Schär, and M. G. Grütter. 1989. Crystallographic refinement of interleukin 1 β at 2.0 Å resolution. *Proc. Natl. Acad. Sci. USA.* 86:9667–9671.
31. Veerapandian, B., G. L. Gilliland, …, T. L. Poulos. 1992. Functional implications of interleukin-1 β based on the three-dimensional structure. *Proteins.* 12:10–23.
32. Clore, G. M., P. T. Wingfield, and A. M. Gronenborn. 1991. High-resolution three-dimensional structure of interleukin 1 β in solution by three- and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry.* 30:2315–2323.
33. Somani, S., C. P. Chng, and C. S. Verma. 2007. Hydration of a hydrophobic cavity and its functional role: a simulation study of human interleukin-1β. *Proteins.* 67:868–885.
34. Zhang, L., and J. Hermans. 1996. Hydrophilicity of cavities in proteins. *Proteins.* 24:433–438.
35. Vaitheeswaran, S., H. Yin, …, G. Hummer. 2004. Water clusters in nonpolar cavities. *Proc. Natl. Acad. Sci. USA.* 101:17002–17005.
36. Word, J. M., S. C. Lovell, …, D. C. Richardson. 1999. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* 285:1711–1733.
37. Wang, G., and R. L. Dunbrack, Jr. 2005. PISCES: recent improvements to a PDB sequence culling server. *Nucleic Acids Res.* 33:W94–W98.
38. Connolly, M. L. 1993. The molecular surface package. *J. Mol. Graph.* 11:139–141.
39. Connolly, M. L. 1983. Analytical molecular surface calculation. *J. Appl. Cryst.* 16:548–558.
40. Zhu, M., F. Shao, …, Z. Xu. 2004. The crystal structure of *Pseudomonas* avirulence protein AvrPphB: a papain-like fold with a distinct substrate-binding site. *Proc. Natl. Acad. Sci. USA.* 101:302–307.
41. Lawson, C. L., R. G. Zhang, …, P. B. Sigler. 1988. Flexibility of the DNA-binding domains of Trp repressor. *Proteins.* 3:18–31.
42. Pesce, A., S. Dewilde, …, M. Bolognesi. 2001. Very high-resolution structure of a trematode hemoglobin displaying a TyrB10-TyrE7 heme distal residue pair and high oxygen affinity. *J. Mol. Biol.* 309:1153–1164.
43. Mobley, D. L., J. D. Chodera, and K. A. Dill. 2006. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* 125:084902–084916.

44. Beutler, T. C., A. E. Mark, …, W. F. van Gunsteren. 1994. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* 222:529–539.

45. Case, D. A., T. E. Cheatham 3rd, …, R. J. Woods. 2005. The AMBER biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.

46. Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–11236.

47. Kaminski, G. A., R. A. Friesner, …, W. L. Jorgensen. 2001. Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B.* 105:6474–6487.

48. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

49. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

50. Oostenbrink, C., A. Villa, …, W. F. van Gunsteren. 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* 25:1656–1676.

51. Berendsen, H. J. C., J. P. M. Postma, …, J. Hermans. 1981. Intermolecular forces. *In* Proceedings of the 14th Jerusalem Symposium on Quantum Chemistry and Biochemistry. B. Pullman, editor. Reidel, Dordrecht, The Netherlands.

52. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7:306–317.

53. Van Der Spoel, D., E. Lindahl, …, H. J. Berendsen. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.

54. Berendsen, H. J. C., J. P. M. Postma, …, J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81: 3684–3690.

55. Denisov, V. P., and B. Halle. 1995. Protein hydration dynamics in aqueous solution: a comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. *J. Mol. Biol.* 245:682–697.

56. Denisov, V. P., B. Halle, …, H. D. Hörlein. 1995. Residence times of the buried water molecules in bovine pancreatic trypsin inhibitor and its G36S mutant. *Biochemistry.* 34:9046–9051.

57. Matthews, B. W., A. G. Morton, and F. W. Dahlquist. 1995. Use of NMR to detect water within nonpolar protein cavities. *Science.* 270:1847–1849.

58. Denisov, V. P., J. Peters, …, B. Halle. 2004. Accelerated exchange of a buried water molecule in selectively disulfide-reduced bovine pancreatic trypsin inhibitor. *Biochemistry.* 43:12020–12027.

59. Denisov, V. P., J. Peters, …, B. Halle. 1996. Using buried water molecules to explore the energy landscape of proteins. *Nat. Struct. Biol.* 3:505–509.

60. García, A. E., and G. Hummer. 2000. Water penetration and escape in proteins. *Proteins.* 38:261–272.

61. Kato, M., and A. Warshel. 2006. Using a charging coordinate in studies of ionization induced partial unfolding. *J. Phys. Chem. B.* 110: 11566–11570.

62. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.