Structure, dynamics, and energetics of water at the surface of a small globular protein: A molecular dynamics simulation

Shubhra Ghosh Dastidar and Chaitali Mukhopadhyay*
Department of Chemistry, University of Calcutta, 92. A.P.C. Road, Kolkata, 700 009, India
(Received 30 May 2003; published 28 August 2003)

The dynamics of water around a biomolecular surface has attracted a lot of attention recently. We report here protein-solvent simulation studies of the small globular protein ubiquitin (human). The simulations are run unconstrained, without freezing the bonds. The mean square displacements of the water oxygen atoms show a sublinear trend with time. The diffusion coefficient data indicate that the water in the first hydration layer behaves like water at a temperature that is roughly 12 °C lower than the average temperature of the system (27 °C). Both the dipolar second-rank relaxation and the survival time correlation function of the water layers show two decay constants, indicating contributions from fast and slow dynamics. A calculation of the interaction energy between the water layers and protein indicates that the interaction energy sharply decreases beyond 4 Å from the protein surface.

DOI: 10.1103/PhysRevE.68.021921 PACS number(s): 87.15.−v, 66.10.−x

I. INTRODUCTION

The water molecules in the hydration shell of a biological macromolecule, e.g., protein and DNA, play a crucial role in determining their structure, function, and dynamics [1,2]. The dynamical behavior of water molecules in the immediate vicinity of the protein surface might become important for enzyme-substrate recognition processes. It has been reported that the mobility of water molecules around hydrophobic and hydrophilic sites is important for the activity of the enzymes [1,3,4]. The energetics and dynamics of water desolvation have been suggested to be a determining factor in the process of protein-ligand recognition [5]. It has been widely recognized that a minimum amount of water is needed for the biological functionality of a protein [6–8]. Moreover, a decrease in hydration level leads in general to an inhibition of the protein mobility [6,8]. On the other hand, it was realized long ago that proteins influence both the spatial and dynamical arrangement of their neighboring water layers [9–11]. It has been established that the properties of the water molecules in the vicinity of a biomolecule differ significantly from those of the bulk water, and these water molecules are called ''biological water'' [12]. This distinction has been made clear by Nandi and Bagchi [13] in relation to dielectric relaxation.

Understanding the structural organization and dynamical properties of the water molecules around a protein is important in predicting the structural and functional alterations of protein upon changing the environment. Because of important role in biological functionality, there have been many studies, both theoretical and experimental [14–23] showing the differences in dynamical behavior of biological water and bulk water. Molecular dynamics (MD) simulation is a very useful tool to perform an experiment *in silico* to show the distinct behavior of biological water. A number of studies have been made using MD simulations to investigate the dynamical characteristics of the water molecules at the protein-solvent interface (for a recent review, see Ref. [22]). The dynamical properties of water in shells around the protein are found to depend on the distance from the biomolecular surface [9,24–27]. In particular, the mean square displacement (MSD) evaluated by MD simulations of water molecules that move in the region close to the protein surface was found to be sublinear with time [24–27]; and the results were confirmed by neutron scattering measurements [28]. Dipolar first- and second-rank relaxations as well as the survival function correlation have been found to follow a stretched exponential decay [10,24,29], indicative of complex dynamics of the hydration shell water around the protein surface.

Although numerous studies have been performed on a variety of protein-water systems, several issues still remain to be addressed. Since the water molecules in the first hydration shell form hydrogen bonds with the protein surface, bond vibrations are likely to play an important role in determining the stability and lifetime of the hydrogen bonds. The OH-stretch vibration, in particular, in liquid water is important because of its coupling to the hydrogen bond [30]. Extensive investigations [31] have shown that the vibrational relaxation takes place with a time constant of 740 ± 25 fs in liquids (e.g., water), which could affect the observed short time (or fast) dynamics of the hydration water. In general, the MD simulations are carried out in the constrained mode, freezing the bond vibrations. This may lead to some artifact in the reported fast dynamics. Apart from that, only limited attempts have been made to quantify the energetics of the protein-water interaction.

In this context, we report here the observations obtained from MD simulations of the small globular protein ubiquitin in an explicit water environment. In order to study the effect of undamped bond vibrations on the dynamics of the protein-water system, no constraints on the bond lengths were added and a short integration time step (0.5 fs) was used. As expected, we observed differences in the dynamic properties of

*Author to whom correspondence should be addressed. FAX: 91-33-2351-9755. Email address: chaitali@cucc.ernet.in, chaitalicu@yahoo.com

the hydration water and bulk water. The value of the diffusion coefficient is in excellent agreement with experimental values [32]. A previous report of a constrained MD simulation of ubiquitin in a water box, by Abseher *et al.* [10], showed that both the dipolar first- and second-rank relaxation follows a single stretched-exponential time law. We report here a double stretched-exponential decay for the dipolar second-rank relaxation, indicating the presence of both fast and slow dynamic components, and show how the relative contributions of the faster and slower components of decay as well as the relaxation times vary for water molecules at different distances from the protein surface. The residence times for the water molecules residing in different layers around the protein also show a bimodal decay. We have made an attempt to measure the protein-solvent interaction energy and show how this interaction energy changes in magnitude for water molecules in different water layers. The calculation shows that there is a sharp difference between water molecules residing within 4 Å of the protein surface and molecules that reside beyond that region. This might be the reason for the observed anomalous behavior of water molecules near the protein surface.

II. METHODS

The initial structure of the protein (human ubiquitin) was collected from the protein data bank (PDB code 1UBQ); it contained 58 crystallization water molecules. It was solvated in a cubic water box (TIP3P model) after energy minimization of the initial crystal structure. The box size was 47.5 Å×47.5 Å×47.5 Å, containing a total of 2940 water molecules, including the crystallization water. Setting a periodic boundary to the system, the entire set was first energy minimized and was heated to 300 K and equilibrated. Six independent trajectories were prepared for performing MD simulation runs, including four trajectories each of 100 ps, one trajectory of 120 ps, and one trajectory of 1 ns. No constraint on the bond vibration was imposed. The time step used was 0.5 fs and the nonbonding cutoff value was 12 Å. The nonbonded lists were updated after each 25 steps. The simulation was performed using CHARMM (version 28). Different types of analysis need different types of sampling intervals and trajectory lengths. The coordinates were usually saved after each 125 fs. In addition to that, in each (100 ps or 120 ps) trajectory from 80 ps to 100 ps (for the 1 ns trajectory it was the last 20 ps), the coordinates were saved after each 12.5 fs, which gave sufficient resolution to study the short time dynamics [33]. This part (i.e., 80–100 ps) of the trajectory was used to analyze the diffusion properties of the system. We started with a water box of size 47.5 Å ×47.5 Å×47.5 Å. The water box was initially optimized at the temperature 300 K and the density was 0.9 g/cm^3. A similar size of water box was prepared with pure water using the TIP3P model. It was simulated up to 100 ps at 300 K, saving the coordinates after each 12.5 fs. This trajectory was used to calculate the diffusion coefficient of pure water. We performed our simulation on a PIII IBM server at our department. MICROCAL ORIGIN 5.0 was used for plotting the data and drawing the figures and for giving the different types of fits (e.g., stretched-exponential decay) to the data.

III. RESULTS

The dynamical quantities discussed here have been evaluated in water shells around the whole protein, involving increasing distances from the protein surface. A number of water layers of different thickness R have been defined. The first includes solvent molecules moving within 4 Å of the protein surface. The others concern the water molecules that move within larger distances (e.g., $R=8$ Å, 14 Å, etc.), with the widest region including almost all the water molecules in the water box. Since the water molecules can migrate from one shell to another during the simulation run, checking the position of a water molecule in a particular layer only at the beginning of the time interval investigated could lead to an incorrect evaluation of the averages of the various types of function calculated. Therefore the position of each water molecule at each MD configuration step was classified and the water trajectory for the subsequent time step was followed; in other words, a dynamic checking of the position of the water molecules was ensured during the analysis. The scheme is akin to the scheme mentioned in previous studies [33,34]. The results have been checked in different independent trajectories. In each case the last part of the trajectory was used for analysis.

A. Distribution of solvent molecules around the protein

The picture of the solvent distribution was found from a plot of the radial distribution function, which is defined as

$$g_\alpha(r)=\frac{\langle \Delta N_\alpha(r)\rangle}{4\pi N_w \rho \Delta r} \tag{1}$$

where $\langle \Delta N_\alpha(r)\rangle$ is the number of water molecules averaged over time, within a distance $r\pm\Delta r/2$ of a hydration site α, ρ is the density of the bulk water, and N_w is the total number of water molecules in the system [33]. Δr was chosen as 0.1 Å. Instead of calculating the radial distribution for a specific hydration site, it is possible to calculate $\langle \Delta N(r)\rangle$, which is actually the number of water molecules within a shell of thickness Δr at a distance r from the nearest protein atom [22]. The radial distribution function was calculated here by averaging over trajectory lengths of 100 ps and 400 ps separately. The data are been plotted in Fig. 1. The two different calculations show almost the same curve, indicating that the system has achieved a stage of equilibrium within a very short time. The first peak at about 1.9 Å arises from a strong interaction between the water oxygen and the hydrogen bond acceptor group on the protein surface and the second peak, whose location is around 2.7 Å, is due to the interaction between the water molecules and the nonhydrogen atoms of the protein [22]. The radial distribution function is very useful to get a better insight in defining the hydration shells around the protein. For further analysis of the dynamical properties we select the water layer of thickness 4 Å measured from the protein surface as the first shell of hydration.

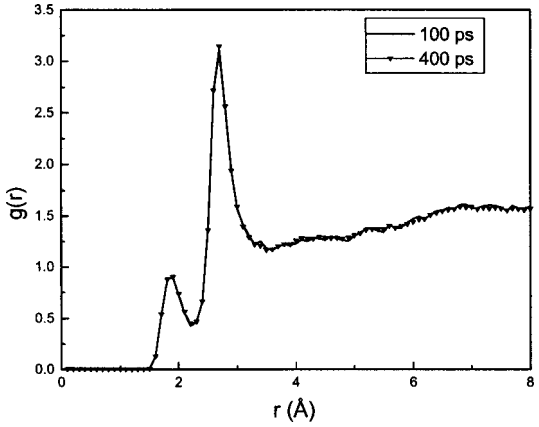

FIG. 1. Water-protein radial distribution around ubiquitin as a function of distance between water oxygen atoms and the nearest protein atoms, including hydrogen atoms. The distribution was averaged over 100 ps (continuous line) and 400 ps (line containing triangles) separately and is shown by two different lines in the figure.

This cutoff value includes all the water molecules that are responsible for the first two peaks in the radial distribution function.

B. Velocity autocorrelation of the solvent molecules

The velocity autocorrelation function $C_{vv}(t)$ of the water molecules (Fig. 2) plotted against time clearly shows that $C_{vv}(t)$ becomes almost zero within a very short time (within a fraction of 1 ps) and a diffusive regime is established. $C_{vv}(t)$ is calculated as follows:

$$C_{vv}(t)=\frac{\langle \mathbf{v}(0)\cdot\mathbf{v}(t)\rangle}{\langle v(0)^2\rangle}, \tag{2}$$

where $\mathbf{v}(t)$ and $\mathbf{v}(0)$ are the velocities of water molecules at time t and at time $t=0$, respectively. $C_{vv}(t)$ was calculated using a trajectory length of 1.2 ps with 12.5 fs resolution. The angular brackets indicate the average over both the time and the solvent molecules. The data show that the diffusive regime is established within a fraction of a picosecond, which is very usual for such systems [22]. This information is needed during calculation of the diffusion coefficient of solvent molecules.

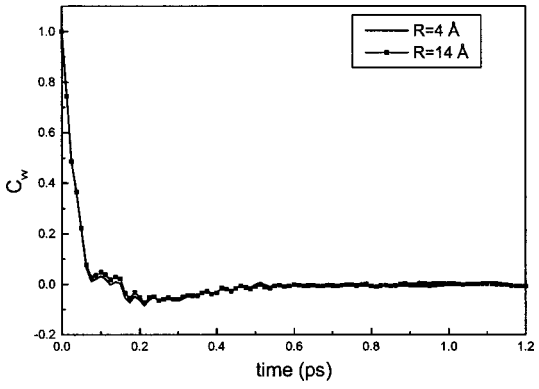

FIG. 2. Translational velocity autocorrelation function for water molecules belonging to the two water layers: continuous line, $R=4$ Å, and line containing squares, $R=14$ Å, where R is measured from the protein surface.

TABLE I. Diffusion coefficient values for the water in different layers around the protein, calculated from the plot of MSD versus time (Fig. 3) according to Eq. (3). The α values of Eq. (4) were calculated from the slope of the best fitting straight line for the data from 2 ps to 10 ps. WATCR and WATIP indicate the data for the crystallographic water and pure water, respectively, using the TIP3P water model. WATXP indicates the experimental values at different temperatures [31].

Water layer	$D/10^{-5}$ (cm^2/s)	α
$R=0-4$ Å	1.44	0.78
$R=0-6$ Å	1.90	0.82
$R=0-8$ Å	2.13	0.85
$R=0-14$ Å	2.48	0.88
WATCR ($R=0-4$ Å)	1.13	0.77
WATIP	2.90	0.91
WATXP	1.3 (278 K)	
	1.8 (288 K)	
	2.3 (298 K)	

C. Diffusion properties of the water molecules

The water mobility in the proximity of the protein surface exhibits a wide range of dynamical behavior, from very tightly bound water to extremely mobile water diffusing on the protein surface. A good reporter of this mobility is represented by the self-diffusion coefficient which is widely used in both spectroscopic investigation [35] and MD simulation approaches for liquids [36]. The solvent mobility is most conveniently described by the diffusion coefficient D related to the slope of the molecular MSD by the Einstein relationship, which in d dimensions is [36]

$$D=\frac{1}{2d}\lim_{\Delta t\to\infty}\frac{\langle|\mathbf{r}_i(t)-\mathbf{r}_i(0)|\rangle}{\Delta t}=\frac{1}{2d}\lim_{\Delta t\to\infty}\frac{\langle\Delta r^2\rangle}{\Delta t}, \tag{3}$$

where $\mathbf{r}_i(t)$ and $\mathbf{r}_i(0)$ are the position vectors of the ith solvent molecule at time t and at time $t=0$ respectively. The angular brackets indicate averaging over both the time and the solvent molecules. This method of calculation of D requires storing of the coordinates with a higher frequency during the simulation run [37]. The time interval Δt has to be large compared to the correlation time of the velocity autocorrelation function, so that any dynamical coherence in the motion of the molecules disappears [38,39]. Figure 2 shows that the velocity autocorrelation reaches a value of zero within a fraction of 1 ps, indicating the establishment of the diffusive regime. The diffusion coefficients for the water in different layers have been calculated (Table I) from the slope of the linear fit of the plot of the MSD of water oxygen atoms vs time during the last 8 ps of a total 10 ps trajectory

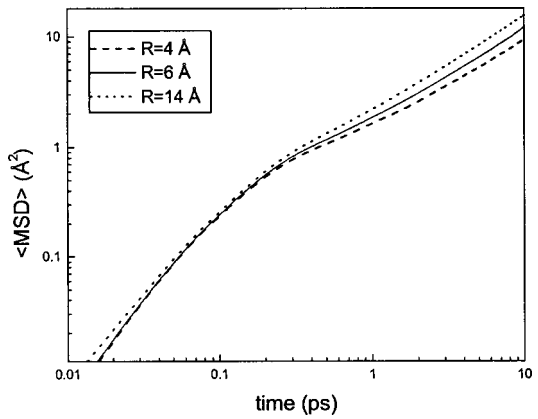

FIG. 3. Mean square displacements of water molecules versus time around fully hydrated ubiquitin obtained by restricting the analysis to the water molecules moving within layers characterized by different distances R measured from the protein surface: dashed line, $R=4$ Å, continuous line, $R=6$ Å, and dots, $R=14$ Å. Each curve was obtained by averaging over ten different time regines and the corresponding water ensemble. The values of the diffusion coefficients and other parameters are reported in Table I.

length with a resolution of 12.5 fs (the Δt being 8 ps). Usually, at very short times (less than approximately 0.2 ps), before the diffusive regime is established, the MSD's follow a ballistic regime ($\Delta r^2 \propto \Delta t^2$), followed by a transient period, after which the MSD curves seem to exhibit a linear trend as a function of time [24,40]. The Einstein relationship [Eq. (3)] for the determination of the diffusion coefficient of a solvent presumes a linear increase of the diffusing particle MSD with time. This condition, usually satisfied for most homogeneous isotropic three-dimensional liquids on time scales longer than a few picoseconds, does not hold for water molecules diffusing around a protein. After the break from the ballistic regime, the $\langle \Delta r^2 \rangle$ values follow the law [24,25]

$$\langle \Delta r^2 \rangle \propto \Delta t^{\alpha}. \tag{4}$$

The value of the exponent α is 1 in the ideal case, but any deviation from unity (usually $\alpha<1$) indicates the presence of anomalous diffusion. Any additional perturbation that restricts the motion of the solvent molecules in an inhomogeneous solution compared to the pure liquids causes a decrease of the value of α. It is easy to visualize the phenomenon and also to obtain the value of α using a log-log plot of MSD vs time (Fig. 3). The values of α for water moving within layers of different thickness measured from the protein surface have been calculated from the plot shown in Fig. 3. Different time origins have been used to label the water molecules traveling within a specified distance R from the protein surface and subsequent 10 ps trajectories have been used for the analysis. The values of α and D are shown in Table I.

The data show that the water within a distance of 4 Å from the protein shows the lowest value of diffusion coefficient. When the same calculation is done at 14 Å, D and α become close to the values obtained after simulating the pure water (WATIP) system using the TIP3P model for water at 300 K. The diffusion coefficient obtained for WATIP is very close to the experimental values [32] for water (WATXP). A value less than 1 for α is indicative of retarded motion with respect to the Brownian behavior, which actually occurs for the hydration layers [41]. The greater deviation of α values near the surface of the protein indicate that anomalous diffusion is taking place for water molecules moving in the proximity of the protein. The value increases as the water molecules move away from the protein surface. It is to be noted that the crystallographic waters show the lowest value of diffusion constant, which is an indication that they are most retarded.

D. Rotational diffusion

Information about the influence of the interaction between the protein and the solvent on the diffusive properties of the hydration water can be obtained from the study of the rotational diffusion of the water electrical dipole. The reorientational dynamics of the water electrical dipole μ can be analyzed by means of the autocorrelation function Γ_l defined as [10,33,39,41,42]:

$$\Gamma_l(t) = \langle P_l(\boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t)) \rangle, \tag{5}$$

where P_l is the lth-order polynomial and $\boldsymbol{\mu}(t)$ is the unit vector along the molecular axis at time t; the angular brackets indicate a time average. The first- and second-order Legendre polynomials are usually investigated. The first-order properties can be derived from infrared spectroscopy, while the second-order polynomial reflects the quadrupolar properties, which can be investigated by NMR [10]. The relaxation of the rotational correlation function of protein hydration water can be fitted to a biexponential function [38,43]:

$$\Gamma_l(t) = Ae^{-(t/\tau_s)} + Be^{-(t/\tau_l)}, \tag{6}$$

where τ_s and τ_l are the relaxation orientation times. This trend may be interpreted as arising from two processes, a fast one accounting for spatially restricted motion due to the librational mode and a slow one involving rearrangement of the neighboring molecules [10,43]. In particular, this relaxation in the proximity of the protein surface can be more accurately described by a stretched exponential, given by

$$\Gamma_l(t) = Ae^{-(t/\tau_s)^{\beta_s}} + Be^{-(t/\tau_l)^{\beta_l}}, \tag{7}$$

where β_l and β_s are two stretching parameters. This type of time behavior is termed Kohlrausch-Williams-Watts (KWW) relaxation [44,45]. Table II shows the two sets of decay constants, stretching parameters, and preexponential factors for water molecules belonging to the layers of different thickness described by R, measured from the protein surface. The subscript s or l indicates that the values are responsible for the short and long components, respectively, of the decay constants. The rotational relaxation of the solvent molecules with time is shown in Figs. 4 and 5 and the data in Table II have been obtained from these plots. The relative proportions of the shorter and longer components are given by the con-

TABLE II. Fitting parameters for the data obtained for the rotational reorientation function calculated according to Eq. (5). For P_1, the best fit was obtained with a single stretched-exponential function, for P_2, with a sum of two stretched-exponential functions [see Eq. (7)].

R (Å)	Method	A	τ_s (ps)	β_s	B	τ_l (ps)	β_l
4.0	P_1				1.0	5.67	0.67
8.0	P_1				1.0	4.71	0.71
14.0	P_1				1.0	4.35	0.75
4.0[a]	P_1				1.0	14.65	0.51
Pure water[b]	P_1				1.0	4.08	0.76
4.0	P_2	0.47	0.48	0.35	0.53	2.19	0.83
8.0	P_2	0.51	0.41	0.39	0.49	2.07	0.91
14.0	P_2	0.46	0.29	0.40	0.54	1.97	0.90
4.0[a]	P_2				1.0	2.87	0.44
Pure water[b]	P_2	0.43	0.30	0.48	0.57	1.85	0.90

[a]Crystallographic water molecules.
[b]Simulated at 300 K using the TIP3P model.

stants A and B [Eq. (7)], where A shows the contribution from the shorter component (i.e., fast decay) and B is for the longer one (i.e., slow decay). For both Γ_1 and Γ_2 the relaxation becomes fast as we proceed from the protein surface to the bulk water and that trend is indicated by the decreasing values of the decay constants responsible for different layers of water molecules. The decrease in deviation from exponentiality is observed from the increasing values of the stretching parameter β with increasing values of R. The best fit for Γ_1 was obtained with a single stretched-exponential function whereas for Γ_2 it was obtained with a sum of two stretched-exponential functions. In previous work with the same protein [10], these relaxations were reported with a single stretched-exponential function, which fails to show the two different types of decay constant (shorter and longer) that we report here.

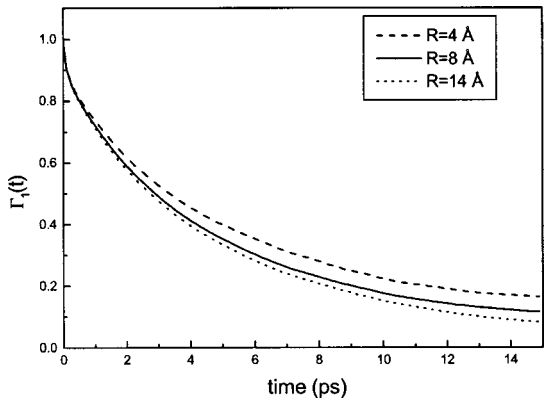

FIG. 4. Rotational reorientation of the water molecule dipole direction for $l=1$ [see Eq. (5)]. The fitting parameters [see Eq. (7)] are reported in Table II. Data for three progressively thicker layers have been plotted: dashed, $R=4$ Å, continuous line, $R=8$ Å, and dots, $R=14$ Å.

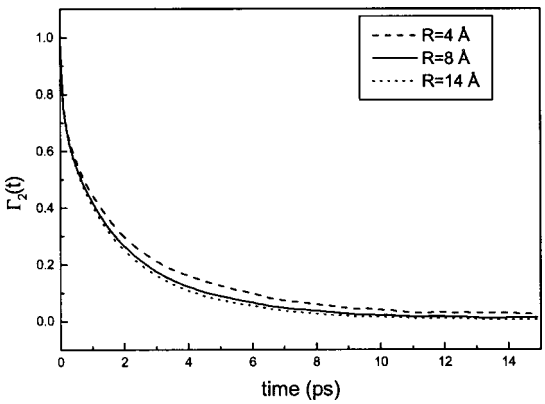

FIG. 5. Rotational reorientation of the water molecule dipole direction for $l=2$ [see Eq. (5)]. The fitting parameters [see Eq. (7)] extracted from the plot are reported in Table II. Data for three progressively thicker layers defined in Fig. 4 have been plotted.

E. Residence time analysis

Water residence times could provide useful insights into the structural and translational dynamical behavior of interfacial water in the first or successive hydration shells of protein atoms exposed to the solvent. Commonly, the residence time is evaluated from the survival time correlation function $C_R(t)$ [46–48] describing the relaxation of the hydration shells of a protein atom (or even a layer) around a macromolecular body [33]. The ''layer survival time correlation function'' can be defined as

$$C_R(t)=\frac{1}{N_w}\sum_{j=1}^{N_w}\frac{\langle P_{Rj}(0)P_{Rj}(t)\rangle}{\langle P_{Rj}(0)^2\rangle}, \tag{8}$$

where P_{Rj} is a binary function that takes the value of 1 if the jth water molecule stays in the layer of thickness R for a time t without getting out during this interval and of zero otherwise [33]. This quantity $C_R(t)$ measures the probability that a water molecule remains in a given layer for a certain time t, without having ever exchanged with the bulk water. The relaxation trend of $C_R(t)$ provides information about the local dynamics of the hydration water molecules. It can be approximated by a single exponential function [47,48]

$$C_R(t)=Ae^{-(t/\tau_s)}. \tag{9}$$

Fitting the data with the above equation provides the relaxation time τ, which represents the mean residence time of the water molecules within a specified distance R from the protein surface. According to the KWW law, a stretched-exponential fit, given by

$$C_R(t)=Ae^{-(t/\tau_s)^\beta}, \tag{10}$$

is used to show the deviation from exponential behavior, which is reflected by the deviation from unity of the value of the stretching parameter β. In the present study, as for the dipolar second-order relaxation, a double stretched-exponential fit was better instead of using a single stretched exponential:

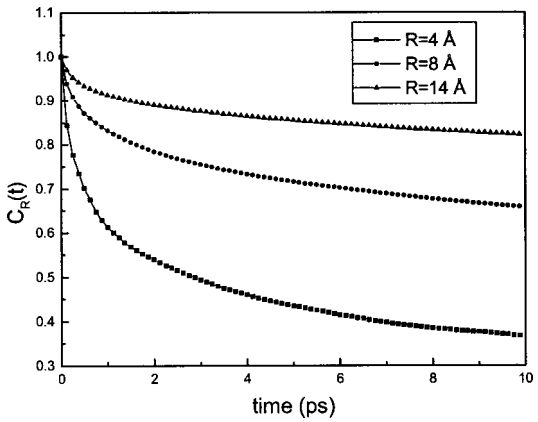

FIG. 6. Plot of survival time correlation function $C_R(t)$ for three water layers defined by R measured from the protein surface: squares, $R=4$ Å, circles, $R=8$ Å, and triangles, $R=14$ Å. The fitting parameters [see Eq. (11)] extracted from the plot are reported in Table III.

$$C_R(t)=Ae^{-(t/\tau_s)^{\beta_s}}+Be^{-(t/\tau_l)^{\beta_l}}, \tag{11}$$

where τ_s and τ_l are the shorter and longer components of the decay constants. These decays correspond to the solvent molecules that stay in the hydration layer for a prolonged period of time or enter and then immediately leave. The decay of the survival time correlation function is shown in Fig. 6. The values of the decay constants along with the stretching parameter of Eq. (11) are shown in Table III. The coefficients A and B of Eq. (11) give the relative proportions of the shorter and longer components, where A is responsible for the shorter component (i.e., fast decay) and B is for the longer one (i.e., slow decay). It can be inferred from the data that, as a hydration layer of larger thickness (measured from the protein surface) is taken into consideration, the probability of exchange with the bulk water for the water molecules within the hydration layer becomes less. The stretching parameter β increases as R increases from 4 Å to 14 Å and the decay of the survival function approaches exponentiality. Comparison of the relaxation times observed for the shells 0–4 Å ($\tau_s=0.39$ ps, $\tau_l=18.0$ ps) and 14–18 Å ($\tau_s=0.28$ ps, $\tau_l=2.5$ ps) clearly indicates that water molecules in the first layer are less mobile compared to the bulk water.

F. Interaction energy of the diffusing water molecules

The deviation of the behavior of the solvent molecules near the surface from their behavior in bulk is due to the strong interaction between the protein surface and the water molecules. Figure 7 shows the variation in the interaction energy per molecule of H_2O with time, averaged over the water molecules whose oxygen atoms are within a layer specified by R, labeled at time $t=0$. The plot, hence, shows the variation of protein-solvent interaction energy due to diffusion of water molecules with time. As we consider layers at larger distances from the protein surface, the magnitude of the average interaction energy at time $t=0$ decreases and varies less rapidly with time. It is interesting to see that the interaction energy (at $t=0$) between the water layer and protein sharply decreases (in magnitude, i.e., neglecting the sign) as R becomes more than 2 Å. This shows that the anomalous behavior of the hydration water, which is mainly due to the stronger interaction of the protein surface with the solvent molecules that reside within a distance of 2 Å from the protein surface. The interaction energy gradually increases (decreases in magnitude, if the sign is neglected) as the molecules diffuse out of the hydration layer. The layers beyond 4 Å have an almost constant interaction with the protein molecule. If, instead of marking the water molecules at $t=0$ as belonging to a particular layer, at each dynamic trajectory frame they are reselected as the water molecules residing (at that instant) within a layer specified by R_d, the plot (Fig. 8) shows that the average interaction energy per molecule of H_2O remains almost invariant. This is the fingerprint of dynamic equilibrium present between the water layers. The average interaction energy between the protein and water within 2 Å is ~ -16 kcal/mole, whereas that within 2–4 Å is ~ -6.2 kcal/mole. Thus water molecules within 2 Å get some additional stability due to interaction with the protein surface, and this stability decreases with time as they diffuse to the bulk water. The absolute values of the energies might be dependent on the force field used (CHARMM22).

TABLE III. Fitting parameters for the data obtained for survival time correlation function $C_R(t)$ calculated according to Eq. (8). The data were fitted to a curve according to Eq. (11).

R (Å)	A	τ_s (ps)	β_s	B	τ_l (ps)	β_l
0–4	0.4	0.39	0.67	0.6	18.0	0.85
0–8	0.2	0.53	0.69	0.8	46.9	0.89
0–14	0.1	0.68	0.70	0.9	137.2	0.92
14–18	0.4	0.28	0.70	0.6	2.5	0.90

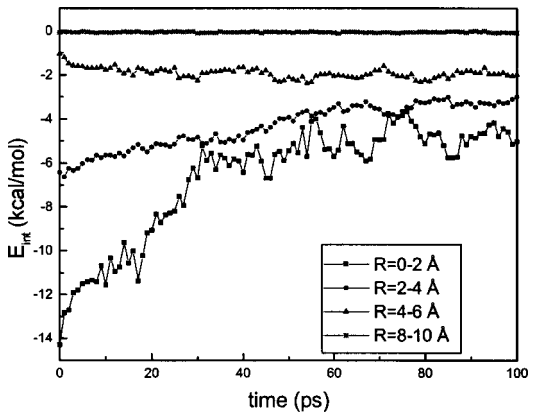

FIG. 7. Variation of solvent-protein interaction energy E_{int} with time, restricting the calculation to the solvent molecules present within a region specified by R: squares, $R=0-2$ Å, circles, $R=2-4$ Å, triangles, $R=4-6$ Å, and stars, $R=8-10$ Å. The molecules are labeled at time $t=0$. The energy is expressed in kilocalories per mole of solvent molecules.

IV. DISCUSSION

We report here the anomalous behavior of water molecules near the surface of the small globular protein ubiq-

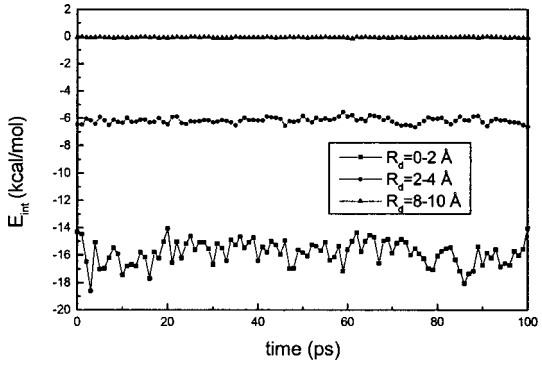

FIG. 8. Variation of solvent-protein interaction energy E_{int} with time, restricting the calculation to the solvent molecules present within a region specified by R_d: squares, $R_d=0-2$ Å, circles. $R_d=2-4$ Å, and triangles, $R_d=8-10$ Å. R_d indicates that during the analysis the molecules were selected after each 1 ps as those molecules that are present within the specified region at that moment.

uitin. It has been established both by experiments and theory that the dynamics of the hydration water can be separated into fast and slow components [49]. It has also been established that OH vibrational relaxations are coupled to the hydrogen bond stability and dynamics [20,30]. The conventional way of performing MD simulations is by freezing bond vibrations and thereby increasing the time step of numerical integration. We have not used any constraints on bond flexibility and have used a time step of 0.5 fs. To study the fast dynamics of solvent molecules the intervals of saving the coordinates also must be reasonably short to have sufficient resolution for analysis. A part of the analysis was done using coordinates saved after each 125 fs, but the diffusion properties were analyzed using coordinates saved after each 12.5 fs. These time intervals are sufficient to provide data accurate enough to study the short time dynamics [33]. Since the definition of the hydration shell is always arbitrary to some extent, we have used a radial distribution plot (Fig. 1) to get insight into the extent of the first hydration shell. The radial distribution plot shows that there is a group of water molecules up to a distance of around 3.0 Å from the protein surface. Thus, using a cutoff radius of 4 Å for the first shell, we ensured that we included in this subset of solvent molecules only those that have a greater chance of interacting directly with the protein surface. Any data obtained from the water within the range $4<R<11$ Å give information about the structural organization beyond the first hydration layer [9]. The translational motion of the water molecules has been analyzed by quantifying their diffusive characteristics. A comparison of the diffusion coefficients (see Table I) obtained from the simulation and experimental data shows that the water molecules within a distance of 4 Å from the protein surface behave like water at a temperature at least $12-15\,^{\circ}$C lower than the average temperature of the system. The extent of the retardation influence of the protein surface on the diffusion properties of the water is best reflected by the α values of Eq. (4), which are obtained from the log-log plot (Fig. 3). The MSD vs time plot shows that before the diffusive regime is established $\langle \Delta r^2 \rangle$ follows a ballistic regime, giving a value of roughly 2 for α, and in the figure the lines for the different water layers are very similar. After the diffusive regime is established the lines for the water molecules residing within different ranges of distances split up, and the different slopes indicate different α values (less than 1). As we proceed from the bulk water to the protein surface, the value of α decreases. The lowest value of α is recorded for the water molecules within the first hydration shell, which is a clear indication that the most restricted motion of the water molecules occurs for those that are directly interacting with the protein surface. The α value and the diffusion coefficients recorded for $R=14$ Å are very close ($\alpha=0.88$) to the values obtained from the simulation of pure water ($\alpha=0.91$), and the value of D is very close to the experimental values for pure water. The diffusion coefficient and α values recorded for the crystallographic waters show their distinct dynamical behavior.

The shell averaged rotational reorientational function was fitted with a sum of stretched-exponential functions. A single stretched-exponential function was needed to obtain the best fit for Γ_1, whereas a sum of two stretched-exponential functions gave the best fit for Γ_2. The values of the stretching parameter β shown in Table II indicate that the deviation from exponentiality increases on going from bulk water to the water in close proximity to the protein. This is indicative of complex dynamics. The multi-stretched-exponential fitting for Γ_2 is useful to reveal the shorter and longer components of the decay constants for this type of relaxation in the system. The constants A and B show the relative contributions of the two types of component. The values in Table II show that the contributions from the two types of decay are almost equal ($\sim 50\%$) in all layers. Interestingly, less and less retardation effect on the rotational motion is evident from the decreasing values of the relaxation times with increasing distance from the protein surface. The data obtained are quite consistent and unambiguous. When a similar fitting was done with the survival time correlation function according to Eq. (11), a slower exchange (according to the increasing values of the decay constants, shown in Table III) of water molecules with the bulk water is observed as the thickness of the water layer measured from the protein surface increases. Here also the stretching parameter β approaches the value of 1 on going from the protein surface to the bulk. To compare the residence times for the water molecules from different water layers of similar thickness, a comparison of the decay constants for the layers 0–4 Å and 14–18 Å is very effective. The value of 18.0 ps for the longer component in the layer 0–4 Å becomes 2.5 ps in the layer 14–18 Å. This shows that the water molecules at a larger distance move more freely than those at the protein surface. The residence time of water at the protein surface is a signature of its mobility and binding. In a recent experiment using femtosecond time resolution, Pal *et al.* [12] reported a bimodal decay for the hydration correlation function with two primary relaxation times: ultrafast, typically 1 ps or less, and longer, typically 15–40 ps. Another recent report from the same group on hydration at the surface of the protein monellin [23] also using femtosecond resolution has shown that the hydration correlation function, which decays due to rotational and translational motion of the water at the protein surface and in

bulk, exhibits a bimodal behavior with time constants of 1.3 and 16 ps, mirroring relaxation of the free or quasifree water molecules and the surface-bound water layer (minimum binding energy of 1–2 kcal/mol). Our observed residence times are well within the range of experimentally observed values.

The appearance of a shorter component could be due to the fact that water molecules vibrate and librate inside a microscopic cage formed by their nearest neighbors, before escaping from such structures [33]. Water molecules close to the boundary of a solvent layer might cross forward and backward across its surface during such motion, and this might give rise to the fast initial decay. Previous work with ubiquitin [10] did not show the bimodal nature of the decay of the correlation of the survival function. The reported values of decay constants and stretching parameters also differ from ours. The differences may be due to the simulation scheme, definition of hydration layers, and finally the less effective single stretched-exponential fitting of the data.

We have also attempted to calculate the amount of energy involved in this interaction, which puts such a control and retardation on the system that the residence times become almost seven times larger at the protein surface than in the bulk water. The quantification of this interaction energy should give an insight into the slow dynamics at the protein surface. The protein-solvent interaction energy per monomer of solvent averaged over the ensemble is shown in Figs. 7 and 8. A larger negative value of the interaction energy indicates more stability. The figures show that the interaction energy for water molecules residing within a distance of 4 Å and labeled at $t=0$ varies with time. This is because these molecules lose H bonds and other strong interactions with the protein surface as they diffuse with time into the bulk water.

All the results reported here have been checked over several independent trajectories of different simulation time spans, and the observed results were consistent. This indicates that the system reaches dynamic equilibrium within a very short time, and a few hundred picoseconds of MD simulation is good enough to study the short time dynamics of systems consisting of a small globular protein like ubiquitin. Our work reveals how a difference in MD simulation procedure can influence the results of studies with the same protein and shows to what extent the time steps and resolution are important.

V. CONCLUSION

It is well accepted now that water play an important role in determining protein structure, and it is also true that the influence of a protein on the dynamical behavior of water is not negligible. State-of-the-art experimental measurements of water dynamics at femtosecond resolution are possible. Theoretical techniques involving computer simulation are used today more to explain and verify the experimental results as well as to suggest new experiments. The opening of this possibility requires scientists to refine the simulation results of existing data and forces them to search for the suitable simulation schemes and methods of analysis which can give results closer to the experimental values and are reproducible. Improvements in results by modifying the simulation schemes and methods of analysis should lead us to progress. Unless the comparison of data for similar types of system but with more advanced simulation schemes becomes common, no effective conclusions can be drawn in any direction. Here, we have shown that MD simulation without freezing the vibrations of bonds with a 0.5 fs integration time step and a maximum of 12.5 fs data collection resolution for analysis produces diffusion properties close to the experimental results, and the data are much better than those previously reported [10]. The use of the TIP3P model for the water box also shares the credit for this improvement.

We report a structured radial distribution function and from the interaction energy studies we clearly show the equilibration between bound and free water molecules. Our scheme of analysis is also appropriate to describe the dynamical phenomenona occurring at different time scales. The presence of dynamics with different time scales has been proved very recently from femtosecond experiments [49]. In our report the fast and slow components of translational and rotational dynamics have clearly been identified, quantified, and discussed. The present study thus provides a path for improving the simulation results.

ACKNOWLEDGMENT

Shubhra Ghosh Dastidar is thankful to CSIR, India for financial support through CSIR-NET.

[1] *Protein-Solvent Interactions*, edited by R. B. Gregory (Dekker, New York, 1995).
[2] M. M. Teeter, Annu. Rev. Biophys. Biophys. Chem. **20**, 577 (1991).
[3] M. F. Colombo, D. C. Rau, and V. A. Parsegian, Science **256**, 655 (1992).
[4] V. Launnas and B. M. Pettit, Proteins: Struct., Funct., Genet. **18**, 148 (1994).
[5] C. J. Camacho, Z. Weng, S. Vajda, and C. Delisi, Biophys. J. **76**, 1166 (1999).
[6] P. L. Poole and J. L. Finley, Biopolymers **22**, 255 (1983).
[7] I. D. Kuntz and W. Kauzmann, Adv. Protein Chem. **28**, 239 (1974).
[8] J. L. Rupley, E. Gratton, and G. Careri, Trends Biochem. Sci. **8**, 18 (1983).
[9] V. Lounnas, B. M. Pettil, and G. N. Philips, Jr., Biophys. J. **66**, 601 (1994).
[10] R. Abseher, H. Schreiber, and O. Steinhauser, Proteins **25**, 366 (1996).
[11] C. L. Brooks III and M. Karplus, J. Mol. Biol. **208**, 159 (1989).
[12] S. K. Pal, J. Peon, and A. H. Zewail, Proc. Natl. Acad. Sci. U.S.A. **99**, 1763 (2002).
[13] N. Nandi and B. Bagchi, J. Phys. Chem. B **101**, 10 954 (1997).
[14] F. T. Burling, W. I. Weis, K. M. Flaherty, and A. T. Brunger, Science **271**, 72 (1996).

[15] D. I. Severgun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai, Proc. Natl. Acad. Sci. U.S.A. **95**, 2267 (1998).
[16] M. Gerstein and C. Chothia, Proc. Natl. Acad. Sci. U.S.A. **93**, 10 167 (1996).
[17] V. A. Makarov, M. Feig, B. K. Andrews, and M. Pettit, Biophys. J. **75**, 150 (1998).
[18] X. Cheng and B. P. Schoenborn, J. Mol. Biol. **220**, 381 (1991).
[19] W. Gu and B. P. Schoenborn, Proteins **22**, 20 (1995).
[20] E. H. Grant, V. E. R. McClean, N. R. V. Nightingale, R. J. Scheppard, and M. J. Chapman, Bioelectromagnetics (N.Y.) **7**, 151 (1986).
[21] O. Gottfried, L. Edvards, and K. Wüthrich, Science **254**, 974 (1991).
[22] A. R. Bizzarri and S. Cannistraro, J. Phys. Chem. B **106**, 6617 (2002).
[23] S. K. Pal, J. Peon, and A. H. Zewail, Proc. Natl. Acad. Sci. U.S.A. **99**, 10 964 (2002).
[24] A. R. Bizzarri and S. Cannistraro, Phys. Rev. E **53**, 3040 (1996).
[25] A. R. Bizzarri and S. Cannistraro, Europhys. Lett. **37**, 201 (1997).
[26] M. Levitt and R. Sharon, Proc. Natl. Acad. Sci. U.S.A. **85**, 7557 (1988).
[27] C. Rocchi, A. R. Bizzarri, and S. Cannistraro, Chem. Phys. Lett. **263**, 559 (1996).
[28] M. Settles and W. Doster, Faraday Discuss. **103**, 269 (1996).
[29] M. Marchi, F. Sterpone, and Matteo Ceccarelli, J. Am. Chem. Soc. **124**, 6787 (2002).
[30] S. Woutersen, U. Emmerichs, H. K. Nienhuys, and H. J. Bakker, Phys. Rev. Lett. **81**, 1106 (1998).
[31] S. Woutersen and H. J. Bakker, Nature (London) **402**, 507 (1999).
[32] R. Mills, J. Phys. Chem. **77**, 685 (1973).
[33] C. Rocchi, A. R. Bizzarri, and S. Cannistraro, Phys. Rev. E **57**, 3315 (1998).
[34] C. Rocchi, A. R. Bizzarri, and S. Cannistraro, Chem. Phys. **214**, 261 (1997).
[35] R. Kimmich, T. Gneiting, K. Kotitschke, and G. Schnur, Biophys. J. **58**, 1183 (1990).
[36] M. P. Allen and D. J. Tildesley, *Computer Simulation of Molecular Liquids* (Clarendon Press, Oxford, 1987).
[37] R. Chitra and S. Yashonath, J. Phys. Chem. B **101**, 5437 (1997).
[38] H. J. Ber. Bensen-Ges-Hertz, Phys. Chem. **75**, 183 (1971).
[39] D. A. McQuarrie, *StatMech* (Harper and Row, New York, 1976).
[40] P. Ahlstörm, O. Teleman, and B. Jönsson, J. Am. Chem. Soc. **110**, 4198 (1998).
[41] J. Klafter, G. Zumofen, and A. Blumen, Chem. Phys. **177**, 821 (1993).
[42] D. A. Zichi and P. J. Rosky, J. Chem. Phys. **84**, 2814 (1986).
[43] M. Norin, F. Haeffner, K. Hult, and O. Edholns, Biophys. J. **67**, 548 (1994).
[44] G. Williams and D. C. Watts, Discuss. Faraday Soc. **66**, 80 (1970).
[45] R. Kohlrausch, Ann. Phys. (Leipzig) **12**, 353 (1947).
[46] R. M. Brunne, E. Liepinsh, G. Otting, K. Wuthrich, and W. F. Gunstern, J. Mol. Biol. **231**, 1040 (1993).
[47] A. E. Garcia and L. Stiller, J. Comput. Chem. **14**, 1396 (1993).
[48] R. W. Impey, P. A. Madden, and I. R. McDonald, J. Phys. Chem. **87**, 5071 (1983).
[49] S. K. Pal, J. Peon, B. Bagchi, and A. H. Zewail, J. Phys. Chem. **106**, 12 376 (2002).