JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

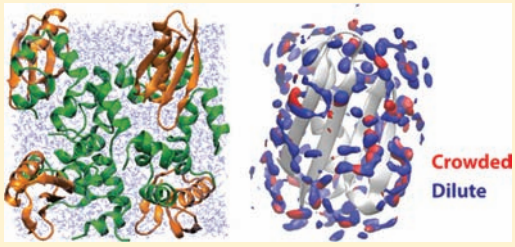
Protein Crowding Affects Hydration Structure and Dynamics

Ryuhei Harada,[†] Yuji Sugita,[†,#,⊥] and Michael Feig*[,#,‡]

[†]RIKEN Advanced Institute for Computational Science, 7-1-26 minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047 Japan
[#]RIKEN Quantitative Biology Center, 7-1-26 minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[⊥]RIKEN Advanced Science Institute, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan
[‡]Department of Biochemistry & Molecular Biology and Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States

Supporting Information

ABSTRACT: The effect of protein crowding on the structure and dynamics of water was examined from explicit solvent molecular dynamics simulations of a series of protein G and protein G/villin systems at different protein concentrations. Hydration structure was analyzed in terms of radial distribution functions, three-dimensional hydration sites, and preservation of tetrahedral coordination. Analysis of hydration dynamics focused on self-diffusion rates and dielectric constants as a function of crowding. The results show significant changes in both structure and dynamics of water under highly crowded conditions. The structure of water is altered mostly beyond the first solvation shell. Diffusion rates and dielectric constants are significantly reduced following linear trends as a function of crowding reflecting highly constrained water in crowded environments. The reduced dynamics of diffusion is expected to be strongly related to hydrodynamic properties of crowded cellular environments while the reduced dielectric constant under crowded conditions has implications for the stability of biomolecules in crowded environments. The results from this study suggest a prescription for modeling solvation in simulations of cellular environments.



INTRODUCTION

The cytoplasm contains up to 400 g/L of macromolecules occupying 5% to 40% of the total volume of the cell.[1] Macromolecular crowding has been shown to alter biomolecular structure and dynamics compared to dilute environments.[2−4] Past studies of crowding have focused largely on volume exclusion by crowder molecules[4] with the main conclusion that such crowding effects lead to more compact states.[3,4] However, when more realistic environments with protein crowders are considered, the effect of crowding tends to destabilize rather than stabilize native states.[5−7] Most discussions of crowding so far have focused on the nature of protein−crowder interactions, while relatively little is known about the effect of crowding on the structure and dynamics of water. In highly crowded environments, the fraction of interfacial water is between 30% and 70% of the total water in the cell.[8] Since physical properties of interfacial waters differ from those of bulk water,[9] the overall properties of water are expected to be significantly different in the presence of high concentrations of macromolecular crowders. In particular, highly crowded environments may alter hydration structure due to strong interactions with nearby macromolecules. Furthermore, diffusion rates and the dielectric response may be lowered because of constraints on water mobility. Altered hydration properties as a result of crowding may in turn affect the structure, dynamics, and function of biomolecules. Especially, a reduced dielectric constant would enhance charge−charge interactions, stabilize secondary structure through strengthened hydrogen bonding, and diminish the advantage of sequestering hydrophobic components away from solvent.[10] Therefore, the effects of crowding on water properties are a key part in developing a full understanding of the behavior of biomolecules in cellular environments. A good understanding of solvation in cellular environments is also essential for the development of effective mean-field models of cellular environments[10] that allow studies of biomolecular structure and dynamics on cellular scales.[11,12]

There is little direct insight into the effect of crowding on hydration to date, but a number of relevant studies have examined water properties near surfaces and under confinement. Water molecules near protein surfaces diffuse more slowly,[13] are more ordered,[14] and generally have a lower density over bulk water in confined spaces.[15] Moreover, water molecules near hydrophobic surfaces behave significantly different than bulk water because of a lack of hydrogen bond donors and acceptors[15,16] while a slow-down in water dynamics is also observed near hydrophilic surfaces.[17−19]

Narrow confinement is expected to further alter the structure of water and significantly perturb the hydrogen-bond network. Indeed, a previous study has reported a reduction of the average coordination number from about 3.6 in the bulk to about 2.2.[15] The confinement-imposed constraints on water molecules reduce the possibility for hydrogen-bond exchange

Received: November 27, 2011
Published: February 21, 2012

and hinder reorientation of water dipoles. This reduces the effective dielectric response and leads to local polarization of interfacial waters that is strong enough to induce an effective long-range attraction between macromolecules.[20,21] According to modeling studies water molecules under hydrophobic confinement move about an order of magnitude slower than those in bulk.[20] An NMR study of water in large confined spaces (300 to 4000 nm) furthermore suggests that even in those environments the molecular mobility of water can be significantly below that of the bulk.[22] Force microscopy experiments suggest that water confined between two hydrophilic surfaces that are less than about 2 nm apart exhibits an increased viscosity by several orders of magnitude greater compared to bulk solvent.[17−19] In addition to a reduction in self-diffusion rates, the dielectric constant is also found to be significantly reduced in confined environments.[20] Further evidence for a reduced dielectric response in cellular environments comes from a comparison of calculated and experimental solubility. A reduced dielectric constant of 55 for the water in the protein crystal,[23] where macromolecular densities are similar to those of crowded cellular environments, resulted in the best agreement between theory and experiment.

To obtain a better understanding of water in crowded biological environments, we performed a series of fully atomistic molecular dynamics (MD) simulations of highly concentrated protein solutions to mimic biological crowding environments. We found that the structure of water remained largely unaffected for low concentrations of crowder proteins but radial distributions of water were affected significantly for highly crowded systems with reduced densities past the first solvation peak. In contrast, diffusion rates and the dielectric constants of water decreased linearly as a function of crowding to about a third of bulk values under most crowded conditions. The observed changes in water properties for highly crowded environments are presumed to have a profound impact on biological macromolecules in cellular environments. Furthermore, our results provide a prescription for developing better mean-field models of crowded cellular environments.

In the remainder of the paper we will describe our results in more detail. The paper is organized as follows: First, the methodology of our simulations is explained, then, results are presented and discussed.

METHODS

To mimic crowded protein environments, two types of systems were set up. For the choice of protein crowders, we considered two simple globular proteins, segment B1 of streptococcal protein G (protein G)[24] and chicken villin headpiece subdomain (villin). Protein G consists of a four-stranded β-sheet against which an α-helix is packed. Villin is a three bundle helix of an F-actin-binding domain located on the far C-terminus of the super villin.[25] These mini-proteins were selected because their small size allows for long simulation times, because they are known to be well-behaved in computer simulations, stably maintaining their native states, and because they are not known to aggregate to a significant extent with each other. To study crowding effects, we prepared two different systems (Protein G/Protein G and Villin) to cover proteins with different secondary structures (α-helical villin vs mostly β-sheet) and compare a single protein crowder type vs system with mixed crowders within the limitations of the computer resources available to us. The first system consisted of eight protein G molecules (56 residues, 864 atoms, PDB entry 1PGB[24]) (PG1−PG4; cf. Table 1); the second system consisted of four protein G molecules and eight villin headpiece subdomains (36 residues, 597 atoms, PDB entry 1YRF[25]) (PGVH1−PGVH5; cf. Table 1). Both systems were solvated in explicit solvent under periodic boundary conditions.

Table 1. Simulation Conditions (298K, 1 bar)

system	box[a] (Å)	N_{water}	protein vol (%)	concn [g/L] protein-G	concn [g/L] villin[b]	length [ns]
PG0	60.77	7122	3[c]	45.8[c]		300
PG1	82.88	16100	10	144.4		300
PG2	74.88	11130	14	195.8		300
PG3	66.88	7195	20	274.9		300
PG4	58.88	4036	30	402.9		300
PGVH1	81.88	14815	12	74.9	101.44	300
PGVH2	73.88	9989	17	101.9	138.09	300
PGVH3	65.88	6303	25	143.8	194.75	300
PGVH4	57.88	3456	37	212.0	287.18	300
PGVH5	53.88	2610	43	262.9	356.01	300

[a]Length of cubic box. [b]Villin headpiece. [c]Effectively at infinite dilution; PG0: protein G × 1; PG1, PG2, PG3, and PG4: protein G × 8; PGVH1, PGVH2, PGVH3, and PGVH4: protein G × 4 + villin headpiece ×8.

Furthermore, to examine the effect of concentration, five different box sizes were considered for each system (cf. Table 1). The protein concentrations ranged from 144 g/L for PG1 to 619 g/L for the mixed protein G/villin system PGVH5 corresponding to protein crowder volume fractions between 10% and 43%. For reference, we also ran a system with a single copy of protein G (PG0) that effectively represents the case of infinite dilution. Equilibrium snapshots of all

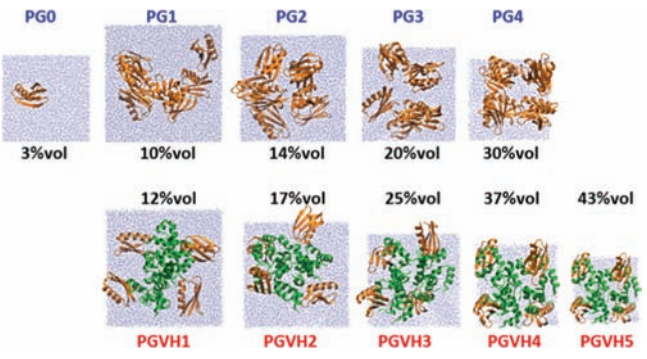


Figure 1. Simulated systems after 100 ns (cf. Table 1). Protein G is shown in brown and villin headpiece subdomain in green.

simulated systems are depicted in Figure 1. The initial systems were set up by placing the protein G and villin molecules in random orientations at the corners of cubic boxes. More specifically, in the PG1-PG4 systems, protein G was placed with the center of mass at (±13 Å, ±13 Å, ±13 Å). In the PGVH1−PGVH5 systems the four protein G molecules were placed at (−18 Å, −18 Å −18 Å), (18 Å, 18 Å, −18 Å), (−18 Å, 18 Å, 18 Å), and (18 Å, −18 Å, 18 Å) and the eight villin molecules were placed at (12 Å, 12 Å, 12 Å), (−12 Å, 12 Å, −12 Å), (−12 Å, −12 Å, 12 Å), (12 Å, −12 Å, −12 Å), (22 Å, 0, 0), (−22 Å, 0, 0), (0, −22 Å, 0), and (0, 0, 22 Å). All the systems were then solvated with explicit TIP3P water molecules (transferable interaction potential- 3 point)[26] in cubic boxes with different lengths ranging from 54 to 83 Å (cf. Table 1). PG1−PG4 were neutralized by 32 sodium ions, placed randomly in the solvent box, PG0 was neutralized with 4 sodium ions. The initial systems were minimized and subsequently heated to 298 K with short simulations at increasing temperature (4 ps at 50 K, 4 ps at 100 K, 4 ps at 200 K, 4 ps at 250 K, at 10 ps at 298 K). Production simulations were then continued in the NPT (constant number of particles, pressure, and temperature) ensemble at a temperature of 298 K and a pressure of 1 bar. All simulations were performed by using the molecular dynamics simulation package NAMD version 2.7b2.[27] The CHARMM 27 all-

dx.doi.org/10.1021/ja211115q | J. Am. Chem. Soc. 2012, 134, 4842−4849

atom force field (Chemistry at Harvard molecular mechanics)[28] was used in combination with the CMAP (cross-correlation map) correction term.[29] The SHAKE algorithm[30] was used to constrain bond lengths involving hydrogen atoms so that a longer integration time step of 2.0 fs could be used. For temperature and pressure control, a Langevin thermostat and piston were used with a friction coefficient of 5 ps^{-1} and a collision period of 0.2 ps, respectively. The particle mesh Ewald (PME) method was employed for the calculation of electrostatic interactions.[31] The direct sum and Lennard-Jones interactions were smoothly truncated at a cutoff distance of 12 Å, using a switching function becoming effective at 10 Å. Each simulation was carried out for 300 ns with coordinates recorded every 0.2 ps. Analysis of the simulations was carried out with the MMTSB Tool set (multiscale modeling tools in structural biology),[32] CHARMM,[33] and custom-written analysis scripts. VMD[34] was used to visualize and generate molecular graphics.

RESULTS

The effect of protein crowding on water structure and dynamics was analyzed from molecular dynamics simulations of protein G and protein G/villin solutions at different concentrations. In the following, we will first describe the water structure and then discuss diffusion and dielectric properties as a function of crowding.

Water Structure. To describe the effect of crowding on water structure we first analyzed the water-accessible volume and radial distribution functions of water with respect to the protein crowders. Furthermore, we describe water structure through three-dimensional density distributions and hydrogen-bonding order parameter calculations.

Accessible Volume. The accessible volume is a function of distance from the closest solute heavy atom was calculated to determine what fraction of water molecules is far enough from any solute to retain bulk properties (cf. Figure 2a,b, and S1 in the Supporting Information). The accessible volume is defined as the volume in which water molecules could exist at a given distance from the closest crowder heavy atom. The accessible volume decreases with distance in the crowded cases compared to the increase in volume of spherical shells in the noncrowded case. Furthermore, for distances over 3 Å, the accessible volume decreases more rapidly with increasing amount of crowding. In the most crowded cases, the accessible volume reduces to zero at 10 Å with 30% vol protein G crowders (PG4) and at 8 Å with 43% vol protein G/villin crowders (PGVH5). To further quantify this point, Table 2 compares the volume within the first two solvation shells vs bulk volume (>7 Å from the protein) as a function of crowding. For crowder volume fractions of 30% and more, the bulk volume is reduced to below 10% while most of the accessible volume is found within the first solvation shell (≤4 Å) from the protein. This means that in the most crowded cases there is theoretically almost no room for bulk water and overall hydration properties are therefore expected to be altered significantly.

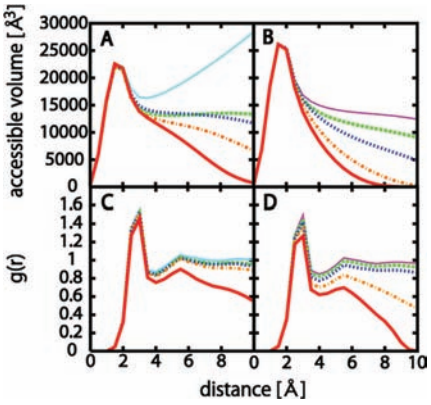


Figure 2. Theoretically accessible volume (A, B) and radial distribution functions (C, D) for water molecules as a function of distance from nearest protein crowder heavy atoms in protein G (A, C) and protein G/villin (B, D) systems. The accessible volume was calculated by integrating grid volume elements with 0.5 Å spacing according to the nearest distance of a given volume element to a crowder heavy atom. Accessible volume and radial distribution functions were averaged over 300 ns. Radial distribution functions were normalized by the accessible volumes and standard bulk water number density of 0.033 $Å^{-3}$. Line colors indicate results from different concentrations as follows: PG0: light blue; PGVH1: magenta; PG1/PGVH2: green; PG2/PGVH3: blue; PG3/PGVH4: orange; PG4/PGVH5: red.

Table 2. Accessible Volume within First and Second Solvation Shell vs Bulk

system	protein [vol %]	1st shell ($r \leq 4$ Å) [%]	2nd shell (4 Å < $r \leq$ 7 Å) [%]	bulk ($r >$ 7 Å) [%]
PG1	10	22.4	14.4	63.3
PG2	14	31.7	20.4	47.9
PG3	20	44.0	24.6	31.5
PG4	30	65.6	26.7	7.7
PGVH1	12	28.1	16.3	55.6
PGVH2	17	37.7	19.1	43.3
PGVH3	25	53.4	22.2	24.5
PGVH4	37	76.2	19.6	4.2
PGVH5	43	86.4	13.0	0.6

Radial Distribution Functions. The radial distribution functions were calculated for the distances between water oxygen atoms and the nearest heavy atoms of any of the protein crowder molecules (cf. Figure 2c,d, and S1 in the Supporting Information). Radial distribution functions $d(r)$ were obtained in a histogram as a function of distance r and then divided by the bulk water density (0.034 $Å^{-3}$) and the theoretically accessible volumes $V(r)$ from parts a and b of Figure 2 to give the normalized functions $g(r)$ shown in parts c and d of Figure 2 according to eq 1:

$$g(r) = \frac{d(r)}{V(r)\cdot 0.034\ Å^{-3}} \quad (1)$$

It can be seen that the first solvation peak remains largely unaffected even under highly crowded conditions. The second solvation peak is also present in all cases but for highly crowded systems with volume fractions of 30% and above the water density is reduced significantly relative to that for bulk densities with increasing distance from the solute. We also separately analyzed accessible volume and radial distribution functions for different residue types of the closest protein atom (grouped into hydrophobic, polar, or charged residues) (cf. Figure S3, Supporting Information) but did not come to significantly different conclusions than for the overall distribution functions. The reduction in bulk density is presumed to be a result of an increasing number of small solvent cavities at highly crowded conditions that are too small to accommodate water molecules. Interestingly this effect is negligible for less than 30% vol but rapidly dominates for crowder fractions of more than 30% vol.

dx.doi.org/10.1021/ja211115q | J. Am. Chem. Soc. 2012, 134, 4842−4849

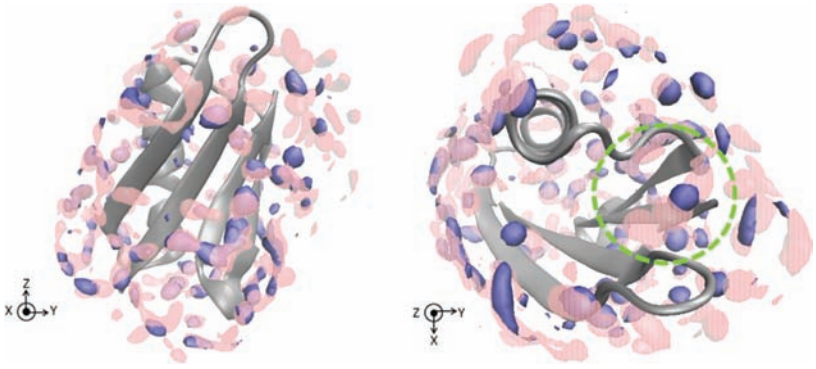


Figure 3. 3D water densities around protein G for the highly crowded PG4 system (pink) compared to densities from noncrowded PG0 system (blue). The densities were averaged over 300 ns from sampling of water around each of the crowder proteins after superposition to a common reference structure. Density contours are shown at a level of 0.18 Å^{-3}. The dashed green circle indicates a prominent solvent-exposed hydrophobic region between residues T25 on the helix and T2 on the sheet, respectively. The left and right figures show front and top views.

Three-Dimensional Water Densities. The three-dimensional water densities were calculated to further compare the structure of water between noncrowded and crowded systems (cf. Figure 3). Most of the high density sites in the first solvation shell appear to be conserved even at highly crowded conditions. However, there are also some regions where the first solvation shell sites are significantly different. One such region is a hydrophobic patch near the surface (marked with a dashed circle in Figure 3) where hydration in the noncrowded case appears to be more extensive and involve a different water structure than in the crowded case. On the other hand, water sites near polar and charged residues were relatively well conserved. In the second solvation shell, the differences between the noncrowded and crowded case were more pronounced. In particular, many of the ordered sites present in the noncrowded case appear to be missing in the crowded simulation. These results suggest that although the radial distribution of water molecules in the first solvation shell is largely unaffected, crowding may alter specific hydration sites in a crowder-dependent fashion.

Hydrogen-Bonding Order Parameters. The hydrogen-bonding order parameters were calculated to examine the effect of crowding on the tetrahedral coordination of water molecules. As a metric we calculated the following order parameter for water molecule i:

$$q_i = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\left(\cos\psi_{ijk} + \frac{1}{3}\right)^2 \tag{2}$$

where ψ_{ijk} is the angle formed between a central oxygen atom i and two of the nearest four oxygen atoms j and k.[35] This summation is over the six possible O−O−O angles involving the central molecule i and pairs of its four neighbors. For a perfect tetrahedral arrangement such as hexagonal ice, $q = 1$ since the values of cos ψ_{ijk} are equal to $-1/3$ with $\psi_{ijk} = 109.5°$. On the other hand, when the relative arrangement of the central atom is completely uncorrelated with its neighbors, the mean value of q vanishes. This corresponds to the ideal gas limit. The resulting distributions are shown in parts a and b of Figure 4. Compared to the noncrowded system, the distributions of the crowded systems were shifted to lower values of q. This indicates that protein crowding disrupts the partial tetrahedral arrangement of water in solution while random orientations ($q = 0$) are slightly increased. This effect

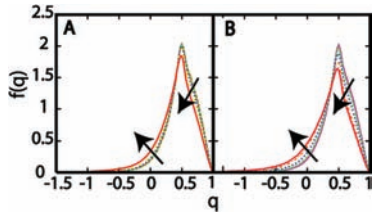


Figure 4. The distribution of the orientational order parameter q of protein G systems (PG0−PG4) (A) and protein G/villin systems (PGVH1−PGVH5) (B) from the first 100 ns of the trajectories. The fraction of water molecules with q-values between $q + dq/2$ and $q - dq/2$ is $f(q)dq$. The arrows indicate increased crowding. Line colors are the same as in Figure 2.

appears to be most pronounced for the PGVH5 system where the protein volume fraction is above 40% and is likely due to restricted mobility of water molecules within cavities that disrupt tetrahedral coordination. However, the overall effect is relatively minor even for the most crowded case.

Water Dynamics. We will now turn to an analysis of dynamic features, in particular water diffusion and dielectric properties.

Diffusion Coefficients. The diffusion coefficients were calculated from mean-square displacements (MSD) of water oxygen atoms:

$$\mathrm{MSD}(t) = \langle |r(t' + t) - r(t')|^2 \rangle_{t'} \tag{3}$$

In normal diffusion, the diffusion coefficient is calculated from the MSD according to

$$D = \lim_{t\to\infty} \frac{1}{6t}\mathrm{MSD}(t) \tag{4}$$

where t is time. Finite periodic box size effects were corrected according to

$$D_S = D_{PBC} + \frac{k_B T \xi}{6\pi\eta L} \tag{5}$$

where D_{PBC} is the uncorrected diffusion coefficient, T is the temperature, L is the box length, η is the shear viscosity of pure water, and the correction factor $\xi = 2.83729$.[36] The MSD time evolution indicates a slowdown in diffusion as a function of crowding even for moderate amounts of crowding (cf. Figure 5a,b). Diffusion rates calculated from the long-time slope of MSD vs time linearly decrease with protein volume fraction (cf.

dx.doi.org/10.1021/ja211115q | J. Am. Chem. Soc. 2012, 134, 4842−4849

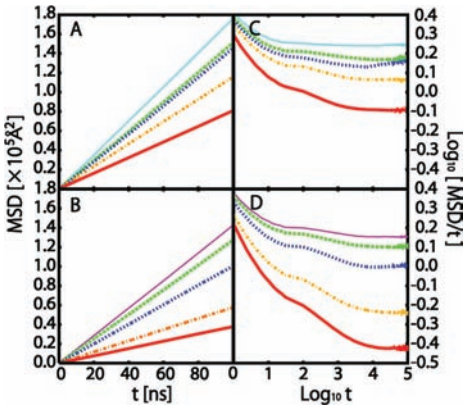


Figure 5. Mean square displacement (MSD) as a function of time for protein G systems (PG0−PG4; A, C) and protein G/villin systems (PGVH1−PGVH5; B, D) calculated from the first 4 ns trajectories for short time scale (0 to 0.03 ns) and from the total 300 ns trajectories for long time scale (0.03 to 100 ns). Line colors are the same as in Figure 2.

Figure 6). The results do not seem to depend on the crowder proteins since the results from PG and PGVH systems fall on

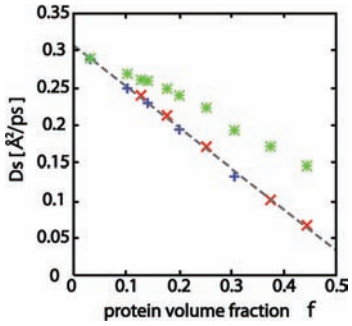


Figure 6. Water diffusion constants as a function of crowder protein volume fraction. Blue and red points correspond to PG0−PG4 and PGVH1−PGVH5 systems, respectively. Diffusion constants were calculated from MSD values during 80 to 100 ns (cf. Figure 5A,B). The dashed line is a linear fit with $D_S = -0.569f + 0.308$ [Å²/ps]. The green points are theoretical predictions of overall diffusion rates from MD simulations around a single protein and based on the fraction of waters close to a protein surface where diffusion is known to be reduced significantly.[13] Details of how the theoretical predictions were obtained for the crowded system are given in the Supporting Information.

the same line. This indicates that water diffusion rates may depend on protein crowding in a universal manner. Theoretical extrapolation suggests that diffusion reaches zero for ∼55% vol when there would be no more movable water molecules.

In normal diffusion, MSD is strictly a linear function of time. However, previous studies have suggested that anomalous diffusion may occur in crowded environments.[37] In anomalous diffusion, MSD follows the following law:

$$\mathrm{MSD}(t) \sim t^{2/d_w} \tag{6}$$

where d_w is the anomalous diffusion exponent. If $d_w = 2$, normal diffusion is recovered. To judge whether anomalous diffusion occurs, plots of $\log[\mathrm{MSD}(t)/t]$ vs $\log t$ are convenient (cf. Figure 5c,d). Anomalous diffusion is present at short time scales (<10 ns) where $\log[\mathrm{MSD}(t)/t]$ varies as a function of $\log t$. It appears that the degree of anomalous diffusion increases with crowding but at long time scales (>10 ns) diffusion becomes normal in all cases. The presence of anomalous diffusion is likely due to a cage effect in highly crowded environments where water molecules can move freely on short time scales but are more restricted from making long-range displacements over longer time scales. In this model, the change from anomalous to normal diffusion essentially depends on the time required for escaping the cage imposed by the crowder molecules.

Dielectric Constants. The dielectric constants ε were calculated for water only in each of the simulations from the box dipole fluctuations. For a finite cubic system with periodic boundary conditions, ε is obtained in general from the following equation:

$$\varepsilon - 1 = 4\pi \frac{\langle M^2 \rangle - \langle M \rangle^2}{3Vk_BT} \tag{7}$$

where M is the total box dipole moment, V is the volume of the system, T is the temperature, and k_B is Boltzmann's constant.[38] Equation 7 applies directly to a homogeneous system where all molecules contribute to the box dipole moment. To calculate the dielectric constant of only the water component, the dipole moment due to water as well as a cross-term between the water dipole and the dipole of the rest of the system have to be considered.[39] Furthermore, the volume in eq 7 has to be scaled by the fraction occupied by the water molecules instead of the total box volume. This then results in the following modified expression for calculating the dielectric constant of water in the context of the crowder molecules:

$$\varepsilon - 1 = 4\pi \frac{\left\langle M_{water}^2 \right\rangle - \left\langle M_{water} \right\rangle^2 + \left(\left\langle M_{water} \cdot M_{rest} \right\rangle - \left\langle M_{water} \right\rangle \left\langle M_{rest} \right\rangle \right)}{3V(1-f)k_BT} \tag{8}$$

where M_{water} is the box dipole moment due to all of the water molecules, M_{rest} is the dipole moment due to the rest of the system (proteins and counterions), and f is the protein volume fraction. The resulting dielectric constants (assuming T = 298 K) with and without the correlation term are shown in Figure 7. Without the correlation term, the dielectric constant decreases as a function of the crowder volume fraction to

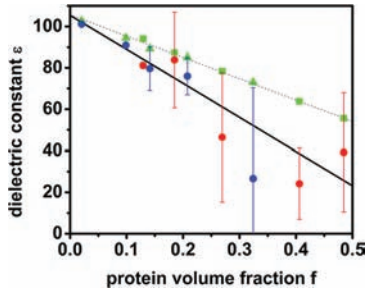


Figure 7. Dielectric constants of water only as a function of protein volume fraction for each system from averages over 300 ns. Blue and red circles correspond to values calculated according to eq 8 for PG0−PG4 and PGVH1−PGVH5 systems, respectively. Errors estimated by comparing results for 0−150 and 150−300 ns intervals are shown as vertical lines. Green symbols (triangles for PG0−PG4, squares for PGVH1−PGVH5) indicate calculated values according to eq 7 without the water−protein/ion correlation term. The data points (including the uncertainties as weights) were fitted to the linear functions $\varepsilon = -103.4f + 105.7$ (dashed line) and $\varepsilon = -164.5f + 105.4$ (solid line).

dx.doi.org/10.1021/ja211115q | J. Am. Chem. Soc. 2012, 134, 4842−4849

below 60 for a volume fraction of 0.5. As for the diffusion rates, the data from both sets of simulations fall onto the same line (cf. Figure 7) suggesting that the decrease is universal. Inclusion of the correlation has a negligible effect for the infinite dilution case, consistent with previous studies of a single biomolecule in a box of water.[39] As the protein crowder fraction increases, the correlation appears to become increasingly important, although, due to sampling limitations, the uncertainty in M_{rest} is much larger than that for M_{water} (see errorbars in Figure 7). All data points suggest that the water−crowder cross-term further reduces the dielectric constant of water. More specifically, a weighted fit that takes the uncertainties into account suggests that the dielectric constant of water in crowded environments is reduced to below 30 at a volume fraction of 0.5 (cf. Figure 7).

DISCUSSION

Previous simulations of crowded biological environments have relied mostly on coarse-grained models of solutes and crowders and even atomistically detailed models of protein solutes were combined with implicit solvent.[12,40] In our study, we have represented both protein crowders and solvent in full atomistic detail to obtain a full picture of hydration in crowded environments. The results from our simulations presented here suggest that macromolecular crowding has a modest effect on water properties for low levels of crowding (<30% vol) but significantly affects hydration structure and dynamics for highly crowded environments (>30% vol). Thus, altered hydration properties are expected to be present in cellular environments where protein volume fractions range between 20% and 40%. For the most crowded environments, water structure beyond the first solvation shell is significantly altered and, even within the first solvation shell, different hydration patterns arise as a result of crowding. Furthermore, diffusion rates are significantly decreased and the dielectric constant is reduced substantially due to constrained water mobility in highly crowded environments,

Reduced diffusion in crowded environments is well-known for biomolecular solutes.[4] We show here that self-diffusion of water is also significantly reduced as a result of crowding. It is interesting to compare these results with the previously described reduction of diffusion near biomolecular surfaces in noncrowded environments.[13] In that case, water diffusion rates were found to decrease smoothly from bulk values to about half at the protein surface. Convoluting that profile (cf. Figure 2 in that work[13]) with radial water density distribution from our simulations in crowded environments (cf. details in the Supporting Information) yields an estimate of how much diffusion rates would be expected to decrease simply because fewer waters are sufficiently far enough from any protein surface to retain bulk properties. The corresponding data in Figure 6 show that the actual decrease in diffusion rates upon crowding is similar but larger than these theoretical estimates. Hence, diffusion is reduced in part simply because a larger fraction of water molecules are near a protein surface. However, there is an additional reduction due to crowding, presumably because of interactions with multiple nearby proteins.

Reduced diffusion translates into increased viscosity according to the Einstein−Stokes equation

$$D_S = \frac{k_B T}{3\pi\eta d} \tag{9}$$

where d is a radius of the particle. This means that the experimentally observed increase in viscosity in cellular environments[41] may be partly also due to increased solvent viscosity. Such an increase of viscosity would be directly related to the constraints on water dynamics by the nearby protein crowders that limit their mobility due to both restricted space and protein−water interactions.

The significant decrease in dielectric response of water in highly crowded environments is probably the most important finding from our study. It appears that the decrease is independent of the particular type of crowder protein and follows a linear function. Hence, even at low concentrations of crowders there is already a modest decrease in the dielectric constant while the water dielectric constant appears to be reduced to ε = 30−50 for crowder volume fractions between 0.3 and 0.4.

As the reduction in diffusional dynamics upon crowding the decrease in dielectric constant is apparently also due to constrained mobility of the water dipoles that limits their ability to respond to changes in electric fields. To further support this point we calculated the water dipole rotational autocorrelation function (cf. Figure 8). The data clearly show an increased

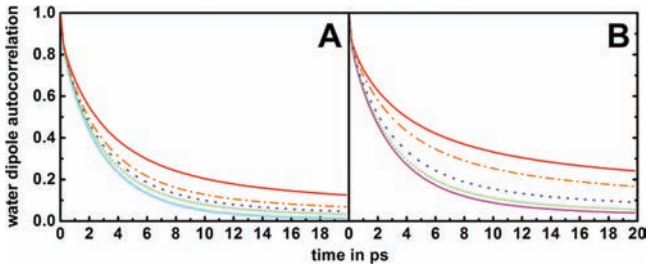


Figure 8. Water dipole autocorrelation function $C(t) = \langle \mathrm{d}(t)\cdot\mathrm{d}(\tau + t)\rangle_\tau$ for water dipoles $\mathrm{d}(t)$ as a function of time t for each system from averages over the last 20 ns for protein G systems (PG0−PG4; A) and protein G/villin systems (PGVH1−PGVH5; B). Line colors are the same as in Figure 2.

correlation as a function of crowding that is consistent with a reduced dielectric response. A reduced dielectric constant in biological environments has been suggested previously from experiments with model hydrophobes.[20] Furthermore, previous solubility calculations suggest a dielectric constant of ε = 55 for modeling the condensed crystalline phase with a protein volume fraction of 0.3.[23] According to the fit in Figure 7 we estimate a dielectric constant of 56 for a protein volume fraction of 0.3, which is surprisingly close to the previous prediction.

We predict a dielectric constant of water in the range of 30−50 for highly crowded environments. The overall polarization response of a crowded cellular environment is assumed to be further reduced due to the presence of cosolvents[42] and nearby proteins (with a dielectric constant between 5 and 20[43]). Such a low effective dielectric constant is expected to profoundly impact the energetics of biomolecules in cellular environments. A reduced dielectric response of the environment would diminish the hydrophobic effect and increase the strength of hydrogen bonds and salt bridges. As a result, we would expect a competing effect of protein tertiary structure destabilization vs secondary structure stabilization. This effect was illustrated in a previous study of melittin in different dielectric environments,[10] where high dielectric environments favored a "folded" conformation with a mini-hydrophobic core while intermediate

dx.doi.org/10.1021/ja211115q | J. Am. Chem. Soc. 2012, 134, 4842−4849

dielectric constants shifted the equilibrium to a more extended helical conformation where backbone hydrogen bonding was maximized at the cost of exposing hydrophobic side chains to solvent. In particular, the reduced dielectric response of the environment may at least partially explain the recently observed destabilization of proteins in highly concentrated protein solutions[5] and biological cells.[6]

The present study offers detailed insight into the structure and dynamics of water in concentrated protein solutions that is not available from experiments. However, the computational approach taken here faces certain limitations due to the choice of force fields, small system sizes, and the amount of conformational sampling that could be obtained. One particular issue is the choice of water model, in this case TIP3P, that is necessitated by the CHARMM protein force field used here. TIP3P is known to overestimate dynamic properties. Both the dielectric constant and self-diffusion rates of water around protein G in dilute solvent are overestimated compared to the experimental values for pure solvent at T = 298 K (ε = 102 vs 78.[44] and D_S = 0.3 vs 0.23[45]) but agree with previous studies.[13] However, under the assumption that diffusion rates and dielectric constants are overestimated to a similar extent at different degrees of crowding, the calculated values could be corrected by scaling with a common factor so that the values for dilute solvent match the experimental values.

Limitations in conformational sampling are apparent in the uncertainties of the calculated box dipole moment for the protein/counterion components. Although crowder proteins diffuse significantly on the order of the box size and exchange partners frequently on nanosecond time scales based on visual inspection of the trajectories, it is clear that protein−protein configurations of the most crowded systems are not sampled fully on the submicrosecond time scales reported here. However, most of the quantities reported here primarily depend on water relaxation properties and are most certainly converged on the time scales of the simulations presented here as indicated, e.g., by the small error bars of the water box dipole moments.

Another caveat concerns the relevance of the model systems studied here for describing real biological environments. Crowded cellular environments contain a more diverse set of proteins, most of them larger than the proteins used here, as well as additional solutes such as nucleic acids and metabolites that are neglected here. It is unclear how a more diverse set of proteins would affect the results presented here, but the small size of the crowder proteins used here may introduce systematic errors. There are likely fewer and larger cavities as crowders become larger at constant crowder volume fraction. As a result, the accessible volume and radial distributions of water far away from a protein surface may decrease less significantly in actual cellular environments as a function of crowding than reported here. Furthermore, if a larger fraction of waters retain more bulk-like properties in the presence of larger crowders, the diffusion rates and dielectric constants may also be reduced to a lesser extent as a function of crowding. It is difficult to fully assess the crowder size effect on hydration properties without carrying out additional simulations with larger crowder molecules, but it appears likely that the effect is small for typical cellular proteins with 200−300 residues.

CONCLUDING REMARKS

In this study, we investigated physical properties of water in the presence of protein crowders at different concentrations. We find that under highly crowded conditions, hydration properties change significantly from bulk solvent. Both hydration structure and dynamics are altered as a function of crowding. The most dramatic change is a significant decrease in self-diffusion and dielectric response. The reduced diffusion rates are expected to affect hydrodynamic properties in cellular environments while a reduced dielectric response alters the thermodynamics of folded proteins. While the focus of this study has been solely on the structure and dynamics of water, future studies will investigate the effect of altered hydration properties on biomolecular solutes in more detail. The results from this work suggest a prescription for developing mean-field models of solvation in cellular environments, for example, by developing implicit descriptions of cellular environments by using the reduced dielectric response reported here. It is our hope that such models will facilitate physically realistic studies of biomolecular dynamics on cellular scales.

ASSOCIATED CONTENT

Supporting Information

Additional methodological details and an additional water distribution profiles. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: feig@msu.edu. Phone: +1-517-432-7439

Author Contributions

The manuscript was written through contributions of all authors.

Funding

NIH GM092949 and GM084953, RIKEN-QBIC, MEXT SPIRE

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Suyong Re for providing his program for calculating order parameter of water structure. Computational resources were provided by RICC (RIKEN Integrated Cluster of Clusters). Funding from NIH GM092949, NIH GM084953, and RIKEN-QBiC (to M.F.) is acknowledged. A part of this research has been funded by MEXT SPIRE Supercomputational Life Science (to YS).

REFERENCES

(1) Ellis, R. J.; Minton, A. P. *Nature* **2003**, *425*, 27.
(2) Despa, F. *Ann. N.Y. Acad. Sci.* **2005**, *1066*, 1.
(3) Zhou, H. X.; Rivas, G. N.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375.
(4) Minton, A. P.; Wilf, J. *Biochemistry* **1981**, *20*, 4821.
(5) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. *J. Am. Chem. Soc.* **2011**, *133*, 7116.
(6) Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; et al. *Nature* **2009**, *458*, 106.
(7) Feig, M.; Sugita, Y. *J Phys Chem B* 2011.
(8) Ball, P. *Chem. Rev* **2008**, *108*, 74.
(9) Despa, F.; Orgill, D. P.; Lee, R. C. *Ann. N.Y. Acad. Sci.* **2005**, *1066*, 54.
(10) Tanizaki, S.; Clifford, J. W.; Connelly, B. D.; Feig, M. *Biophys. J.* **2008**, *94*, 747.
(11) Elcock, A. H. *Curr. Opin. Struct. Biol.* **2010**, *20*, 196.

(12) McGuffee, S. R.; Elcock, A. H. *PLoS Comput. Biol.* **2010**, *6*, xxx.
(13) Makarov, V. A.; Feig, M.; Pettitt, B. M. *Biophys. J.* **1998**, *75*, 150.
(14) Feig, M.; Pettitt, B. M. *Struct. Fold Des.* **1998**, *6*, 1351.
(15) Thompson, H.; Soper, A. K.; Ricci, M. A.; Bruni, F.; Skipper, N. T. *J. Phys. Chem. B* **2007**, *111*, 5610.
(16) Chandler, D. *Nature* **2002**, *417*, 491.
(17) Golding, I.; Cox, E. C. *Phys. Rev. Lett.* **2006**, *96*, 098102.
(18) Major, R. C.; Houston, J. E.; McGrath, M. J.; Siepmann, J. I.; Zhu, X. Y. *Phys. Rev. Lett.* **2006**, *96*, 177803.
(19) Li, T. D.; Gao, J. P.; Szoszkiewicz, R.; Landman, U.; Riedo, E. *Phys. Rev. B* **2007**, *75*, 115415.
(20) Despa, F.; Fernandez, A.; Berry, R. S. *Phys. Rev. Lett.* **2004**, *93*, 228104.
(21) Jacobson, B. *Nature* **1953**, *172*, 666.
(22) Tsukahara, T.; Hibara, A.; Ikeda, Y.; Kitamori, T. *Angew. Chem., Int. Ed.* **2007**, *46*, 1180.
(23) Tjong, H.; Zhou, H. X. *Biophys. J.* **2008**, *95*, 2601.
(24) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657.
(25) Davies, D. R.; Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7517.
(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(27) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283.
(28) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586.
(29) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400.
(30) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(31) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(32) Feig, M.; Karanicolas, J.; Brooks, C. L. III *J. Mol. Graphics Modell.* **2004**, *22*, 377.
(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(34) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(35) Errington, J. R.; Debenedetti, P. G. *Nature* **2001**, *409*, 318.
(36) Yeh, I. C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873.
(37) Saxton, M. J. *Biophys. J.* **1994**, *66*, 394.
(38) Mountain, R. D.; Thirumalai, D. *Comput. Phys. Commun.* **1991**, *62*, 352.
(39) Yang, L.; Weerasinghe, S.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **1995**, *69*, 1519.
(40) Elcock, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2340.
(41) Kuimova, M. K.; Botchway, S. W.; Parker, A. W.; Balaz, M.; Collins, H. A.; Anderson, H. L.; Suhling, K.; Ogilby, P. R. *Nat. Chem.* **2009**, *1*, 69.
(42) Paruta, A. N. *J. Pharm. Sci.* **1964**, *53*, 1252.
(43) Honig, B.; Gilson, M.; Fine, R.; Rashin, A. *Biophys. J.* **1984**, *45*, A129.
(44) Fernandez, D. P.; Mulev, Y.; Goodwin, A. R. H.; Sengers, J. M. H. L. *J. Phys. Chem. Ref. Data* **1995**, *24*, 33.
(45) Hertz, H. G. Nuclear magnetic relaxation spectroscopy. In *Water: A Comprehensive Treatise*; Franks, F., Ed.; Plenum Press: New York, 1973; Vol. 3; p 301.