Rotational and Translational Diffusion of Proteins as a Function of Concentration

Zahedeh Bashardanesh,† Johan Elf,† Haiyang Zhang,‡ and David van der Spoel*†‡

1Uppsala Center for Computational Chemistry, Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-75124 Uppsala, Sweden
2Department of Biological Science and Engineering, School of Chemistry and Biological Engineering, University of Science and Technology Beijing, 100083 Beijing, China

ABSTRACT: Atomistic simulations of three different proteins at different concentrations are performed to obtain insight into protein mobility as a function of protein concentration. We report on simulations of proteins from diluted to the physiological water concentration (about 70% of the mass). First, the viscosity was computed and found to increase by a factor of 7−9 going from pure water to the highest protein concentration, in excellent agreement with in vivo nuclear magnetic resonance results. At a physiological concentration of proteins, the translational diffusion is found to be slowed down to about 30% of the in vitro values. The slow-down of diffusion found here using atomistic models is slightly more than that of a hard sphere model that neglects the electrostatic interactions. Interestingly, rotational diffusion of proteins is slowed down somewhat more (by about 80−95% compared to in vitro values) than translational diffusion, in line with experimental findings and consistent with the increased viscosity. The finding that rotation is retarded more than translation is attributed to solvent-separated clustering. No direct interactions between the proteins are found, and the clustering can likely be attributed to dispersion interactions that are stronger between proteins than between protein and water. Based on these simulations, we can also conclude that the internal dynamics of the proteins in our study are affected only marginally under crowding conditions, and the proteins become somewhat more stable at higher concentrations. Simulations were performed using a force field that was tuned for dealing with crowding conditions by strengthening the protein−water potential.

INTRODUCTION

The intracellular environment is densely packed with macromolecules like nucleic acids, proteins, and sugars. The macromolecule concentration in, for instance, Escherichia coli is ≈300−400 mg/mL,1 corresponding to ≈40% of the intracellular volume.2,3 It has been shown that a high concentration of macromolecules influences the thermodynamics and kinetics of cellular processes.3−7 Most in vitro biochemical investigations are performed using dilute solutions of up to 10 mg/mL. Despite this obvious difference, it is commonly assumed, due to the lack of available data, that biochemical quantities such as protein−ligand binding constants or enzymatic reaction rates in vivo are comparable to the in vitro values. Under in vivo conditions, steric effects and nonspecific interactions are significant, however, and these constraints likely affect cellular processes. Dynamics of biomolecules have been studied extensively by nuclear magnetic resonance (NMR) experiments,8,9 and this has led to models describing the internal protein dynamics as a combination of intramolecular vibrations and overall tumbling.10 The typical rotational tumbling time for a small protein of a few nanoseconds, which has emerged from NMR experiments,8,11 is not the same in a crowded environment.12 Therefore, an in-depth understanding of the diffusive properties of proteins under conditions mimicking the interior of a cell is of general interest for understanding cellular processes.

Diffusion of macromolecules in the cytoplasm can be measured through a variety of experimental techniques, including direct or indirect optical techniques such as fluorescence correlation spectroscopy (FCS),13 single-particle tracking (SPT),14,15 fluorescence recovery after photobleaching (FRAP),16,17 or spectroscopic techniques such as pulse field gradient (PFG) NMR via diffusion-ordered spectroscopy (DOSY).18 For example, the translational diffusion coefficient of green fluorescent protein (GFP) in E. coli was measured by FRAP to be reduced to 10% of the value in water.19,20 Theoretical models based on hard sphere (HS) models allow investigation of the steric effects on HS properties, due to
increasing “macromolecular” concentration. Such models explain the excluded volume effect for simplified physical systems and can be used to simulate diffusion-limited, reversible association reactions. Smooth hard spheres with only repulsive interactions are likely not an accurate approximation of biomolecules, however, since they do not capture effects, e.g., due to interactions between cellular components or the role of water. More complex modeling studies have been applied to study the effects of crowding, although these studies were still performed without explicit water.

A number of full atomistic simulations of proteins under crowding conditions with explicit water have been published. Spiga et al. focused on the effect of crowding by high concentrations of glucose on protein dynamics and found significantly retarded protein dynamics and protein dehydration. From this work, they find that rotational as well as translational diffusion decrease in a similar manner due to crowding. Water dynamics in a crowded environment was addressed by the Feig group in multiple studies. In one study, the structure of water around proteins was investigated, while another focused on water dynamics by studying proteins of different sizes. One further simulation study of Villin, a 36-residue protein, under crowding conditions, showed that protein–protein interactions led to protein rotational motion being retarded more than translational motion. von Bülow and co-workers have presented large simulations of concentrated protein solutions to analyze dynamic properties. The findings from these papers will be discussed below.

An issue that has recently come to the light is that traditional additive force fields are not well suited to study proteins at high concentrations or disordered proteins. Three solutions to these problems have been proposed: to strengthen the interaction between proteins and water by selectively strengthening the Lennard-Jones (LJ) interactions, to strengthen the water–water dispersion interaction or to weaken the Lennard-Jones interaction between the protein atoms. Best et al. proposed to replace the water model to weaken the LJ interaction between the protein and water. For ubiquitin, a 76-residue protein, under crowding conditions, showed that protein–protein interactions led to protein rotational motion being retarded more than translational motion. von Bülow and co-workers have presented large simulations of concentrated protein solutions to analyze dynamic properties. The findings from these papers will be discussed below.

Table 1. Content of the Simulation Boxes as a Function of the Number of Proteins

| Number of Proteins | one protein | two proteins | four proteins | eight proteins |
|--------------------|-------------|--------------|---------------|----------------|
| SOL Na+/Cl− C      | SOL Na+/Cl− C                         |
| ubiquitin          | 16 422      | 46/46        | 3.3           | 16 042         | 46/46 | 6.6 |
| 2k57               | 10 871      | 46/46        | 3.3           | 12 318         | 46/46 | 26.9 |
| 2kim               | 23 549      | 46/46        | 3.3           | 23 029         | 46/46 | 26.9 |

*SOL refers to the number of waters and C the protein concentration (mM).
For all simulations, the Amber99SB-ws force field was used\textsuperscript{36} in combination with the TIP4P/2005 water model.\textsuperscript{39} Electrostatic interactions were treated using the particle mesh Ewald algorithm\textsuperscript{55} and Lennard-Jones interactions with a cutoff of 1 nm. All covalent bonds were constrained at their equilibrium length using the LINCS algorithm,\textsuperscript{56} allowing an integration time step of 2 fs. The temperature was controlled at 310 K using the \textit{v}-rescale algorithm\textsuperscript{57} and a coupling time of 0.5 ps. The pressure was controlled at 1 bar using the Parrinello–Rahman algorithm\textsuperscript{58} with a time constant of 10 ps. Thirty-nine simulations of 1 \(\mu\)s each were performed on parallel computers using the GROMACS software.\textsuperscript{50,51}

\textbf{Hard Sphere Conformational Sampling.} Conformations of nonoverlapping “ubiquitin-like” hard spheres (HSs) were generated for \(N = 2, 4, \) or 8 particles. Two radii of gyration \(R_g = 1.2\) and 1 nm were used as the sphere size, the average sizes of the simulation boxes were 7.96, 7.94, and 7.91 nm (box edge, respectively). Then, random coordinates for \(N\) HS particles were generated in cubic periodic boxes with the edge indicated. This was done 100 000 times to generate a reasonably sized sample. Distance analysis was performed as described below for the MD simulations.

\textbf{Translational Diffusion Coefficient.} A mean square displacement (MSD) analysis was performed to calculate the translational diffusion coefficient.\textsuperscript{39} The diffusion coefficients were extracted by a linear fit to the MSD by averaging blocks with a length of 100 ns. Errors in the diffusion coefficients were estimated by bootstrapping. The diffusion coefficient needs to be corrected for finite-size effects\textsuperscript{60} through eq 1

\[
D_{\text{PBC}} = D_0 - \frac{2.837k_B T}{6\eta L} \approx D_0 - 6.44 \times 10^{-22}/(\eta L) \tag{1}
\]

where \(\eta\) is the viscosity, \(k_B = 1.38 \times 10^{-23}\) m\(^2\) kg/(s\(^2\) K), \(T = 310\) K, and \(L = 7, 8, 9 \times 10^{-9}\) m, respectively, for 2k57, ubiquitin, and 2kim. Here, the shear viscosity is derived for each simulation through the Einstein relation\textsuperscript{59,61}

\[
\eta = \lim_{t \to \infty} \frac{1}{2k_B T} \frac{d}{dt} \frac{1}{6} \sum_{i=1}^{N} \sum_{\beta=1, \alpha \neq \beta}^{N} \left( \int_{t_i}^{t_{i+1}} P_{\beta\alpha}(t) \, dt \right)^2 q_0
\]

where \(V\) is the volume of the simulation box and \(P_{\beta\alpha}\) is the off-diagonal components of the pressure. Statistics is improved by taking the average over all six off-diagonal pressure tensor components.

\textbf{Rotational Correlation Time and Internal Dynamics.} To estimate the rotational correlation time (global tumbling), each trajectory of a protein was fitted rotationally and translationally to the initial structure. From these trajectories, the autocorrelation functions \(C_i(t)\) were calculated for all backbone amide (NH) vectors \(\mathbf{r}_i\) using

\[
C_i(t) = \int_0^\infty P_2(\mathbf{r}_i(t) \cdot \mathbf{r}_i(0)) \, dt
\]

where \(P_2\) is the second-order Legendre polynomial, \(P_2(x) = (1/2)(3x^2 - 1)\). Subsequently, the order parameter \(\langle S^2\rangle\) and the internal correlation time \(\tau_e\) were derived by fitting to the function\textsuperscript{10}

\[
C_i(t) = \langle S^2 \rangle + (1 + \langle S^2 \rangle) \exp(-t/\tau_e) \tag{4}
\]

for each NH vector \(\mathbf{r}_i\). These two values \((\langle S^2 \rangle, \tau_e)\) were averaged over all NH vectors to obtain \(\langle S^2 \rangle\) and \(\langle \tau_e \rangle\), and these were used in turn as the initial guess for calculating the overall tumbling time \(\tau_M\) by fitting the averaged correlation function

\[
\langle C(t) \rangle = \langle (S^2 + 1 + (S^2)) \exp(-t/\langle \tau_e \rangle) \exp(-t/\tau_M) \rangle \tag{5}
\]

to the autocorrelation function of the original trajectory averaged over the whole protein. This decomposition assumes that the overall tumbling is isotropic, which need not be the case; however, this assumption is applied commonly.\textsuperscript{10,62} The estimated overall tumbling time \(\tau_M\) together with order parameters \(S^2\) and internal correlation times \(\tau_e\) obtained initially were then used to minimize \(\chi^2\) defined as

\[
\chi^2 = \sum_{i=1}^{N} \left[ C_i(t) - \left( (S^2_i + 1 + S^2_i) \exp(-t/\tau_e) \right) \exp(-t/\tau_M) \right]^2
\]

with respect to the \(2N + 1\) parameters \(S^2_i\), \(\tau_e\) and \(\tau_M\). The tumbling time that is proportional to the rotational diffusion coefficient through \(D_{\text{rot}} = 1/6\tau_M\) needs to be corrected for finite-size effects\textsuperscript{55} through eq 7

\[
D_{\text{PBC}} = D_0 - \frac{k_B T}{6\eta V} \approx D_0 - 7.13 \times 10^{-22}/(\eta V) \tag{7}
\]

\textbf{RESULTS}

\textbf{Protein Stability.} The mean square fluctuation of the proteins was analyzed at the residue level to see the effect of crowding on the stability of proteins (Figure 2). In the case of 2k57 and ubiquitin, the structures of proteins are intact and...
also the internal dynamics of proteins are not affected by crowding (Figure 2A,2B). However, for 2kim, the region from residue 30 to 50, containing a loop and an α-helix, unfolds to some extent in the simulations at a low protein concentration (Figure 2C). Upon comparing the simulation of 2kim using the amber99sb-ws force field with a simulation using the standard Amber ff99SB-ILDN force field, we note that no unfolding occurs with the original force field (Figure 2C,D) as visualized in Figure 3. This may indicate that the water−protein interactions were strengthened too much in the flm−ws force field. This notion is consistent with the finding that there is less unfolding as the water concentration decreases (Figure 2C).

**Internal Protein Dynamics.** The internal dynamics of the proteins was analyzed based on the motions of the backbone NH vectors using the Lipari and Szabo analysis (eq 4). Figure 4 shows the average order parameters \( S_2 \) for the three proteins as a function of concentration. In conjunction with this, Figure 5 shows the internal relaxation times \( \tau_e \). The plots show the typical shape where residues in loop regions are more flexible with low \( S_2 \) (Figure 4) and high \( \tau_e \) (Figure 5) than those in stable secondary structure elements. For ubiquitin, the region around residue 50 is more labile at low protein concentration than at high protein concentration (Figures 4B and 5B), this is, however, not apparent in the root mean square fluctuations (RMSFs) (Figure 2B) that are low at all ubiquitin concentrations. The same region, around residue 30−50, in 2kim that showed large RMSF (Figure 2C) also has low order parameters (Figure 4C) and slow internal tumbling (Figure 5C).

**Overall Mobility.** The shear viscosity \( \eta \), the average translational diffusion constants \( D_T \) for the proteins, and the overall tumbling times \( \tau_M \) for each system are tabulated in Table 2. The viscosity in the single protein simulations, at low concentrations, is very similar as expected, at around \( 1.3 \times 10^{-3} \) kg/(m s). For reference, the viscosity for TIP4P/2005 has been estimated to be \( 0.855 \times 10^{-3} \) kg/(m s), and the addition of protein and salt increases the viscosity rather drastically to about a factor of 7−9 going from pure water to physiological protein concentration (Table 2). The viscosity is notoriously difficult to compute because the pressure fluctuations converge slowly. As a result, the computed \( \eta \) for two ubiquitins and for two or four 2kim proteins are too low despite averaging over three replicas of 1 μs. This has an impact on the protein diffusion constants for these systems through eq 1 (Table 2), but the trends in Figure 6A are clear anyway.
To be able to compare the dynamics for different proteins, translational diffusion constants $D_P$ were normalized to be one at dilute conditions. $D_P/D_{P,0}$ is displayed as a function of protein volume fraction in Figure 6B. At a biomolecular mass fraction of more than 20%, the diffusion coefficient is reduced to 25—38% of dilute conditions (25% for 2k57, 38% for ubiquitin, and 35% for 2kim). For comparison, the gray line in Figure 6B shows the estimated decline in the diffusion coefficient for hard spheres (HSs) modeled according to the Enskog theory$^{66}$

$$D_{HS} = D_{HS,0}(1 - \phi)^3/(1 - 0.5\phi^2)$$

where $\phi$ is the protein volume fraction.

The average rotational tumbling times $\tau_M/\tau_{M,0}$ and corresponding uncertainties were obtained from all of the proteins in each simulation and are displayed as a function of biomolecular mass fraction in Figure 6C. At the highest concentration, we find that the global tumbling time $\tau_M$ increases by a factor of 21.9 ± 4.1 for 2k57, 11.5 ± 3.4 for
ubiquitin, and 6.6 ± 1.5 for 2kim. We note that the spread in the results, which is different from the protein translational diffusion, is compatible with NMR experiments. Mean square displacement plots of simulations of 2, 4, 8, and 64 ubiquitins are provided in the Supporting Information. Simulations of 8 and 64 proteins have the same protein concentration, and the fact that the results of both simulations are similar shows that the size of the systems is sufficient.

Water Diffusion. The water diffusion coefficient \( D_w \) is reduced by about 40% as the protein volume fraction increases to that corresponding to physiological conditions (Figure 6D). The HS model overestimates the reduction of the diffusion coefficient for water under crowding conditions slightly.

Protein–Protein and Protein–Water Interactions. Protein–protein hydrogen bonds were counted in all simulations with more than one protein, to test for clustering (Table 3). A geometric criterion was used where a hydrogen bond was counted when the donor–acceptor distance was less than or equal to 30°. Mean square displacement plots of simulations of 2, 4, 8, and 64 ubiquitins coincide with a reduced number of intramolecular hydrogen bonds (Table 4). Radial distribution functions (RDFs) are provided in Figures S6–S8. The plots give center-of-mass distance RDFs for simulations of two and eight proteins. Except for one simulation of two proteins of 2kim (Figure S8), the probability of finding proteins close to each other is larger at the highest concentration (eight proteins) than at the lower concentration (two proteins), which is logical from the perspective of reduced volume available to the proteins at higher concentrations.

Distance Analysis. We analyzed distances between protein heavy atom at different concentrations for different cut-off distances equal to 0.27, 0.57, and 0.87 nm in the same manner as Nawrocki et al.70 We considered distances at 0.27 nm to correspond to direct interactions and 0.57 and 0.87 nm for the first and second hydration layers, respectively. Increasing the cut-off leads to a significantly larger fraction of clustered proteins. A comparison with a ubiquitin-like hard sphere model (Figure 7) shows that the clustering cannot be explained by excluded volume as the largest clusters in the MD simulation are populated to much larger fractions than the HS model. We note that the radius used for the hard spheres was taken somewhat ad hoc as the radius of gyration. Using a smaller radius, corresponding to a lower concentration, leads to smaller and fewer protein clusters (Figure 7). The analysis at the shortest distance, 0.27 nm, corresponds to a strong hydrogen bond or salt-bridge. Apparently, this underestimated the clustering significantly, and therefore the hydrogen bonding analysis (Table 3) is not sufficient to draw conclusion about clustering behavior. Rather than direct interactions, long-range dispersion interactions between the proteins78,79 or simply the hydrophobic effect80 may influence the clustering. Figures S17–S19 show that the clustering occurs rapidly during the time of simulations and is reproducible in the three replicas of the protein simulations.

### DISCUSSION

It has been stated that water determines the structure and dynamics of proteins.25 However, this picture is biased by the fact that most biochemical experiments are performed under dilute conditions.70 Protein properties depend on their environment, for instance, through electrostatic interactions with other constituents of that environment75,76 and the charge distribution on biomolecular surfaces has been suggested to affect macromolecular dynamics.90 Macromolecular crowding has also been reported to alter the translational and rotational diffusion of biomolecules in experimental studies.12,75,76 In this work, we employed molecular dynamics simulations to study proteins at high concentrations to mimic crowding effects. It is well established that MD simulations allow us to probe protein–water interactions in spatial and temporal details (see also shown in Figures 2C and 3. The high number of protein–water hydrogen bonds in single-molecule 2kim simulations coincides with a reduced number of intramolecular hydrogen bonds (Table 4).

#### Table 3. Number of Intermolecular Protein–Protein Hydrogen Bonds Averaged over the Last 500 ns of the Simulations

| protein   | two copies | four copies | eight copies |
|-----------|------------|-------------|--------------|
| 2k57      | 0.4 ± 0.1  | 0.5 ± 0.4   | 0.5 ± 0.7    |
| ubiquitin | 0.7 ± 0.5  | 0.5 ± 0.5   | 0.6 ± 0.8    |
| 2kim      | 1.1 ± 1.0  | 0.6 ± 0.4   | 0.6 ± 0.6    |

“Average and standard deviation over three replicas.

#### Table 4. Number of Intramolecular Protein–Protein Hydrogen Bonds Averaged over the Last 500 ns of the Simulations

| protein   | one copy | two copies | four copies | eight copies |
|-----------|----------|------------|-------------|--------------|
| 2k57      | 37.0 ± 1 | 36.4 ± 2   | 35.2 ± 2    | 36.7 ± 1     |
| ubiquitin | 55.5 ± 2 | 55.1 ± 2   | 55.2 ± 2    | 54.8 ± 2     |
| 2kim      | 53.1 ± 6 | 58.1 ± 4   | 55.7 ± 5    | 57.2 ± 4     |

“Average and standard deviation over three replicas.

Table 5. Number of Protein–water Hydrogen Bonds Averaged over the Last 500 ns of the Simulations

| protein   | one copy | two copies | four copies | eight copies |
|-----------|----------|------------|-------------|--------------|
| 2k57      | 167 ± 2  | 167 ± 3    | 168 ± 4     | 163 ± 2      |
| ubiquitin | 186 ± 3  | 186 ± 3    | 185 ± 2     | 183 ± 4      |
| 2kim      | 257 ± 15 | 242 ± 11   | 249 ± 12    | 242 ± 9      |

“Average and standard deviation over three replicas.
e.g., ref 77). Furthermore, MD simulations allow us to analyze the internal dynamics78 as well as the diffusive properties of proteins62,79 in comparison with, in particular, nuclear magnetic resonance (NMR) experiments. The choice of proteins was guided here by the availability of the experimental data for protein tumbling.

Simulations of highly concentrated systems with conventional force fields have been reported to lead to unphysical aggregation of biomolecules due to strong protein–protein interactions.32,33 Here, we used the force field by Best et al.36 that has been tuned for simulations of high concentrations of macromolecules, by strengthening the protein–water Lennard-Jones interactions. The force field is used in conjunction with the TIP4P/2005 water model 39 to obtain better kinetic properties. The combination was validated in two additional papers, one on hydration free energy of amino acids and diffusion of amino acids42 and one comparing the Best et al. force field36 to other approached for shifting the relative strength of interactions toward stronger protein–water interactions.38 As noted in the introduction, alternative combinations of water models and protein force fields have been applied as well.28,31

The translational mobility of three small proteins, 2k57 (55 AA), 1ubq (76 AA), and 2kim (102 AA), are found here to be 6 times slower for 2k57, 3.1 times slower for ubiquitin, and 3.0 times slower for 2kim at a physiologically relevant protein volume fraction compared to the dilute conditions (Figure 6B). In other words, the slow-down of translational diffusion is more than the factor of two reported for the mini-protein Villin by Nawrocki et al.30 We do find a slow-down of rotational tumbling (Figure 6C) comparable to that reported by Nawrocki et al. for Villin.30 Protein concentration is an important factor and our simulations were performed at a somewhat higher concentration than those by Nawrocki30 or von Bülow,31 and therefore we find stronger retardation of mobility. Experimentally, a range of 4–17 times slow-down of translational diffusion has been reported.17,19,20,80,81 The fact that our simulation results are at the lower end of that range could be because there are only weak protein–protein interactions in our simulations (Figure 7). The clustering observed may explain that the rotational tumbling is retarded more than translational diffusion because direct interactions will hinder rotational motion. A cellular environment is crowded by molecules in a wide range of sizes and charge distributions, and transient interactions between cellular components could affect diffusional properties,76 in line with the result on the Villin protein.30 The slow-down in diffusion found in this work is correlated to the increase in viscosity (Figure 6A), although it is known that the Stokes–Einstein relation between diffusion and viscosity holds approximately only82,83 since the particle size that enters the equation is not an exact quantity. However, recent work by von Bülow suggests that the Stokes–Einstein relation may be applied to proteins in solution as well.31

The rotational diffusion slow-down for the different proteins (Figure 6C) has a large spread, a factor of 22 for 2k57, 11 for ubiquitin, and 7 for 2kim. These results are compatible with the NMR data that show that translational- and rotational diffusion may or may not be coupled, depending on the protein.67 For 2k57, the large increase in \( \tau_M \) (Figure 6C and Table 2) can likely not be explained by the somewhat reduced hydration (Table 5), as was noted for protein mobility in a crowded glucose solution,26 but rather by weak interactions as discussed above. It has been noted before that the hydration properties may differ between proteins of different sizes, leading to different diffusion properties.28,84 Harada et al.

![Figure 7. Distance-based cluster analysis of three proteins, 2k57, ub, and 2kim, at three different concentrations for three different distances. In addition, a similar analysis of a ubiquitin-like hard sphere (HS-ub1.2 and HS-ub1, see Methods) was performed. Colors indicate the distance criterion used. Blue: 0.27 nm, red: 0.57 nm, and green: 0.87 nm.](https://example.com/image.png)
suggest a relationship between protein–water interactions and viscosity, based on the Stokes–Einstein equation.28 Our results suggest that the increase of viscosity with protein concentration is very similar to the three proteins studied here (Figure 6A). Indeed, the increase of the viscosity by a factor of 7–9 is very similar to the factor of 8 reported from in vivo experiments.26 This suggests that the slow-down in rotation has another cause, e.g., the charge distribution on the surface or weak dispersion interactions. We find that the rotational diffusion of the proteins in our study seems to be size-dependent, unlike the translational diffusion and viscosity (Figure 6A–C). At the highest concentration of proteins, small proteins are slowed down more than big proteins for an unknown reason. The effect of protein–protein interactions on translational and rotational diffusion has previously been investigated by the addition of high concentrations of either synthetic polymers or protein crowders to dilute protein solutions as an attempt to mimic the cellular environment.12 These results showed that synthetic polymers led to translational diffusion being retarded more than rotational diffusion. However, using protein crowders or cell lysate, the rotational diffusion was retarded more than translational diffusion,12 these results seem to agree with the findings in this work (Figure 6B,C).

The differences between our simulations and the HS model predictions for diffusion are small for both proteins (Figure 6B) and water (Figure 6D). The reasonable correspondence between full atomistic MD and HS models suggests that the excluded volume explains the drop in diffusion to some extent;24 however, interactions are important as well (Figure 7). Indeed, Brownian dynamics simulations, where only steric effects are taken into account, form a simple tool to test Enskog predictions for the effects of crowding. It has been shown that the diffusion of simple hard sphere molecules slows down as a function of molecular volume occupation in agreement with Enskog theory.85 Hard sphere Brownian dynamics simulation coupled with hydrodynamic interactions can in principle be tuned to give a good estimate on slow-down of translational diffusion as a function of crowding.83,86,92 However, these techniques cannot give insight into the role of explicit waters or into the rotational mobility of biomolecules.

McGuie and Elcock studied the effects of crowding in a cytoplasmic model system of E. coli containing the 50 most abundant proteins, but without explicit water.25 Steric forces with and without electrostatic interactions cause different dynamics as shown by the difference in the diffusion coefficient of GFP as one of the constituents of their model, however, the GFP diffusion coefficient predicted by that model25 was still not correct when compared to measurements.7,19,20 The addition of an intramolecular interaction by Lennard–Jones potential was used to obtain a correct diffusion coefficient of GFP.25 However, for this, the Lennard–Jones parameter had to be exaggerated since the role of water was neglected.87 Skolnick highlights in a review 88 the importance of hydrodynamic interactions for understanding biomolecular mobility under crowding conditions. However, these models predict, for instance, that the decay of translational diffusion is size-dependent, something that seems to be at odds with our findings (Figure 6B) although the range of protein sizes studied here is limited and crowders of different sizes and chemical composition can have different effects. It should be noted that, in general, hydrodynamic interactions need to be considered only when no explicit solvent is present.86

It has been shown that proteins become more stable under crowding conditions effectuated through inert synthetic polymers, where only the steric effects are important.89–92 However, when doing the measurement in a cellular-like environment, it was found that electrostatic interactions destabilize protein structures when opposite charges on the surface attract each other.93 In some cases, electrostatic interactions may stabilize the structures due to repulsive forces.3,94,95 These findings emphasize that the effect of direct intermolecular interactions on stability is not negligible. Few intermolecular protein–protein hydrogen bonds were found in this study (Table 3), but solvent-separated interactions are important (Figure 7), and there is an effect of protein concentration on protein stability (Figure 2). It cannot be excluded that due to the force field used for this study, the solute–solvent interactions are strengthened so much that the proteins have the tendency to bind to water more than to the other proteins and, as a result, it seems that the region containing a loop and α-helix (residue 30–50) of the 2kim protein is destabilized (Figure 2C). This behavior was not observed when the original Amber ff99SB-ILDN was used (Figures 2D and 3). Interestingly, the stability increases when the protein concentration goes up for both 2kim (Figure 2C) and ubiquitin (Figure 4B), in line with the experimental studies mentioned.99–92

Without a doubt, the relationship between protein stability and the environment or protein dynamics and the environment is very complex.28,31,70 Although MD simulations are ideally suited to study large and complex systems of biomolecules, be it virus particles or crowding effects in atomistic detail, it will take considerable effort to tune force fields to become completely transferable between different chemical environments.100,101 It may be necessary to consider long-range dispersion interactions explicitly7,27,28 or to include higher-order terms of the London dispersion forces to obtain more accurate simulations.102 However, to study the effect of crowding without altering the van der Waals parameters, using a model with higher-order dispersion coefficients or a polarizable force field might give results closer to experimental values, even though conventional force field improvement is still on-going as well.103

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.9b02835.

Mean square displacement of proteins as a function of time (Figures S1–S13); protein–protein radial distribution functions (Figures S14–S16); and maximum protein–protein cluster size as a function of time (Figures S17–S19) (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

E-mail: david.vanderspoel@icm.uu.se. Phone: +46 18 4714205.

**ORCID**

Haiyang Zhang: 0000-0002-2410-7078

David van der Spoel: 0000-0002-7659-8526

DOI: 10.1021/acsomega.9b02835

ACS Omega 2019, 4, 20664–20664
Funding
The Swedish research council is acknowledged for financial support to DvdS (grant 2013-5947) and for a grant of computer time (SNIC2017-12-41) through the parallel computer center, Stockholm, Sweden. H.Z. acknowledges support from the National Natural Science Foundation of China (grant 21606016).

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Zimmerman, S. B.; Trach, S. O. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J. Mol. Biol. 1991, 222, 599–620.
(2) Luby-Phelps, K. Cytoarchitectural and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int. Rev. Cytol. 1999, 192, 189–221.
(3) Ellis, R. J. Macromolecular crowding: an important but neglected aspect of the intracellular environment. Curr. Opin. Struct. Biol. 2001, 11, 114–119.
(4) Zimmerman, S. B.; Minton, A. P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu. Rev. Biophys. Biomol. Struct. 1993, 22, 27–65.
(5) Sarkar, M.; Li, C.; Pielak, G. J. Soft interactions and crowding. Biophys. Rev. 2013, 5, 187–194.
(6) Mourão, M. A.; Hakim, J. B.; Schnell, S. Connecting the dots: the effects of macromolecular crowding on cell physiology. Biophys. J. 2014, 107, 2761–2766.
(7) Rivas, G.; Minton, A. P. Macromolecular crowding in vitro, in vivo, and in between. Trends Biochem. Sci. 2016, 41, 970–981.
(8) Kay, L. E.; Torchia, D. A.; Bax, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. Biochemistry 1989, 28, 8972–8979.
(9) Palmer, A. G., III Probing molecular motion by NMR. Curr. Opin. Struct. Biol. 1997, 7, 732–737.
(10) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. I. Theory and Range of Validity. J. Am. Chem. Soc. 1982, 104, 4546–4559.
(11) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. J. Mol. Biol. 1999, 283, 1071–1081.
(12) Wang, Y.; Li, C.; Pielak, G. J. Effects of proteins on protein diffusion. J. Am. Chem. Soc. 2010, 132, 9392–9397.
(13) Krichevsky, O.; Bonnet, G. Fluorescence correlation spectroscopy: the technique and its applications. Rep. Prog. Phys. 2002, 65, 251–297.
(14) English, B. P.; Hauryliuk, V.; Sanamrad, A.; Tankov, S.; Dekker, N. H.; Elf, J. Single-molecule investigations of the stringent response in living bacterial cells. Proc. Natl. Acad. Sci. U.S.A. 2011, 108, E365–E373.
(15) Monnier, N.; Barry, Z.; Park, H. Y.; Su, K.-C.; Katz, Z.; English, B. P.; Dey, A.; Pan, K.; Cheeseman, I. M.; Singer, R. H. Inferring transient particle transport dynamics in live cells. Nat. Methods 2015, 12, 838–840.
(16) Ellenberg, J.; Sigga, E. D.; Moreira, J. E.; Smith, C. L.; Presley, J. F.; Woodman, H. J.; Lippincott-Schwartz, J. Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. J. Cell Biol. 1997, 138, 1193–1206.
(17) Mullineaux, C. W.; Nenninger, A.; Ray, N.; Robinson, C. Diffusion of green fluorescent protein in three cell environments in Escherichia coli. J. Bacteriol. 2006, 188, 3442–3448.
(18) Johnson, C. Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications. Prog. Nucl. Magn. Reson. Spectrosc. 1999, 34, 203–256.
(19) Elovitz, M. B.; Surette, M. G.; Wolf, P.-E.; Stock, J. B.; Leibler, S. Protein Mobility in the Cytoplasm of Escherichia coli. J. Bacteriol. 1999, 181, 197–203.
(20) Konopka, M. C.; Shkel, I. A.; Cayley, S.; Record, M. T.; Weisshaar, J. C. Crowding and confinement effects on protein diffusion in vivo. J. Bacteriol. 2006, 188, 6115–6123.
(21) Bashardasteh, Z.; Lötstedt, P. Efficient Green’s Function Reaction Dynamics (GFRD) simulations for diffusion-limited, reversible reactions. J. Comput. Phys. 2018, 357, 78–89.
(22) Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Hrench, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E. Water Determines the Structure and Dynamics of Proteins. Chem. Rev. 2016, 116, 7673–7697.
(23) Singer, P. M.; Asthagiri, D.; Chapman, W. G.; Hirasaki, G. J. Molecular dynamics simulations of NMR relaxation and diffusion of bulk hydrocarbons and water. J. Magn. Reson. 2017, 277, 15–24.
(24) Ando, T.; Skolnick, J. Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. Proc. Natl. Acad. Sci. U.S.A. 2010, 107, 18457–18462.
(25) McGuffee, S. R.; Elcock, A. H. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS Comput. Biol. 2010, 6, No. e1000694.
(26) Spiga, E.; Abriata, L. A.; Piazza, F.; Dal Peraro, M. Dissecting the effects of concentrated carbohydrate solutions on protein diffusion, hydration, and internal dynamics. J. Phys. Chem. B 2014, 118, 5310–5321.
(27) Yu, I.; Morì, T.; Ando, T.; Harada, R.; Jung, J.; Sugita, Y.; Feig, M. Biomolecular interactions modulate macromolecular structure and dynamics in atmospheric model of a bacterial cytoplasm. eLife 2016, 5, No. e19274.
(28) Harada, R.; Sugita, Y.; Feig, M. Protein crowding affects hydration structure and dynamics. J. Am. Chem. Soc. 2012, 134, 4842–4849.
(29) Wang, P.-H.; Yu, I.; Feig, M.; Sugita, Y. Influence of protein crowder size on hydration structure and dynamics in macromolecular crowding. Chem. Phys. Lett. 2017, 671, 65–70.
(30) Nawrocki, G.; Wang, P.-H.; Yu, I.; Sugita, Y.; Feig, M. Slowdown in diffusion in crowded protein solutions correlates with transient cluster formation. J. Phys. Chem. B 2017, 121, 11072–11084.
(31) von Bülow, S.; Sigga, M.; Linke, M.; Hummer, G. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. Proc. Natl. Acad. Sci. U.S.A. 2019, 116, 9843–9852.
(32) Petrov, D.; Zagrovic, B. Are current atomic force fields accurate enough to study proteins in crowded environments. PLoS Comput. Biol. 2014, 10, No. e1003638.
(33) Abriata, L. A.; Dal Peraro, M. Assessing the potential of atomistic molecular dynamics simulations to probe reversible protein-protein recognition and binding. Sci. Rep. 2015, 5, No. 10549.
(34) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J. Phys. Chem. B 2015, 119, S113–S123.
(35) Piana, S.; Klepeis, J. L.; Shaw, D. E. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. Curr. Opin. Struct. Biol. 2014, 24, 98–105.
(36) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association. J. Chem. Theory Comput. 2014, 10, 5113–5124.
(37) Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D. Properties of liquids in Molecular Dynamics Simulations with explicit long-range Lennard-Jones interactions. J. Chem. Theory Comput. 2015, 11, 2938–2944.
(38) Bashardanesh, Z.; van der Spoel, D. Impact of Dispersion Coefficient on Simulations of Proteins and Organic Liquids. J. Phys. Chem. B 2018, 122, 8018–8027.

(39) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. J. Chem. Phys. 2005, 123, No. 234505.

(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys. 1983, 79, 926–935.

(41) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A systematic study of water models for molecular simulation. J. Chem. Phys. 1998, 108, 10220–10230.

(42) Zhang, H.; Yin, C.; Jiang, Y.; van der Spoel, D. Force Field Benchmark of Amino acids: I. Hydration and Diffusion in Different Water Models. J. Chem. Inf. Model. 2018, 58, 1037–1052.

(43) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Struct., Funct., Bioinf. 2010, 78, 1950–1958.

(44) Best, R. B.; De Sancho, D.; Mittal, J. Residue-specific α-helix propensities from molecular simulation. Biophys. J. 2012, 102, 1462–1467.

(45) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Struct., Funct., Genet. 2006, 65, 712–725.

(46) Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J. Phys. Chem. B 2009, 113, 9004–9015.

(47) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Compressibility, Expansion Coefficient and Dielectric Constant. J. Chem. Theory Comput. 2012, 8, 61–74.

(48) Zhang, J.; Tuguldur, B.; van der Spoel, D. Force field benchmark II: Gibbs energy of solvation of organic molecules in liquids. J. Chem. Inf. Model. 2015, 55, 1192–1201.

(49) Zhang, J.; Tuguldur, B.; van der Spoel, D. Correction to Force field benchmark II: Gibbs energy of solvation of organic molecules in liquids. J. Chem. Inf. Model. 2016, 56, 819–820.

(50) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. B.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel benchmark II: Gibbs energy of solvation of organic molecules in liquids. J. Chem. Inf. Model. 2016, 56, 819–820.

(51) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015, 1, 2–15.

(52) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of Ubiquitin Refined at 1.8 Å Resolution. J. Mol. Biol. 1987, 194, 531–544.

(53) Rossi, P.; Swapna, G. V. T.; Huang, Y. J.; Aramini, J. M.; Anklint, C.; Conover, K.; Hamilton, K.; Xiao, R.; Acton, T. B.; Ertelink, A.; et al. A microscope protein sample screening pipeline. J. Biomol. NRMR 2010, 46, 11–22.

(54) Aramini, J. NMR Determined Rotational Correlation Time, 2010. http://www.nmr2.buffalo.edu/nswg.wiki/NMR_determined_Rotational_correlation_time.

(55) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. J. Chem. Phys. 1995, 103, 8577–8592.

(56) Hess, B; P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. J. Chem. Theory Comput. 2008, 4, 116–122.

(57) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. J. Chem. Phys. 2007, 126, No. 041101.

(58) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. J. Appl. Phys. 1981, 52, 7182–7190.
(82) Walser, R.; Hess, B.; Mark, A. E.; van Gunsteren, W. F. Further investigation on the validity of Stokes-Einstein behaviour at the molecular level. *Chem. Phys. Lett.* 2001, 334, 337–342.

(83) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic Properties of Water/Alcohol Mixtures Studied by Computer Simulation. *J. Chem. Phys.* 2003, 119, 7308–7317.

(84) Makarov, V. A.; Feig, M.; Andrews, B. K.; Pettitt, B. M. Diffusion of solvent around biomolecular solutes: a molecular dynamics simulation study. *Biophys. J.* 1998, 75, 150–158.

(85) Kim, J. S.; Yethiraj, A. Effect of macromolecular crowding on reaction rates: a computational and theoretical study. *Biophys. J.* 2009, 96, 1333–1340.

(86) Frembgen-Kesner, T.; Elcock, A. H. Striking Effects of Hydrodynamic Interactions on the Simulated Diffusion and Folding of Proteins. *J. Chem. Theory Comput.* 2009, 5, 242–256.

(87) Elcock, A. H.; Gabdoulline, R. R.; Wade, R. C.; McMannon, J. A. Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. *J. Mol. Biol.* 1999, 291, 149–162.

(88) Skolnick, J. Perspective: On the importance of hydrodynamic interactions in the subcellular dynamics of macromolecules. *J. Chem. Phys.* 2016, 145, No. 100901.

(89) Cheung, M. S.; Klimov, D.; Thirumalai, D. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* 2005, 102, 4753–4758.

(90) Miklos, A. C.; Li, C.; Sharaf, N. G.; Pielak, G. J. Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry* 2010, 49, 6984–6991.

(91) Charlton, L. M.; Barnes, C. O.; Li, C.; Orans, J.; Young, G. B.; Pielak, G. J. Residue-level interrogation of macromolecular crowding effects on protein stability. *J. Am. Chem. Soc.* 2008, 130, 6826–6830.

(92) Kilburn, D.; Rob, J. H.; Guo, L.; Briber, R. M.; Woodson, S. A. Molecular crowding stabilizes folded RNA structure by the excluded volume effect. *J. Am. Chem. Soc.* 2010, 132, 8690–8696.

(93) Sarkar, M.; Smith, A. E.; Pielak, G. J. Impact of reconstituted cytosol on protein stability. *Proc. Natl. Acad. Sci. U.S.A.* 2013, 110, 19342–19347.

(94) Tyrrell, J.; McGinnis, J. L.; Weeks, K. M.; Pielak, G. J. The cellular environment stabilizes adenine riboswitch RNA structure. *Biochemistry* 2013, 52, 8777–8785.

(95) Monteith, W. B.; Pielak, G. J. Residue level quantification of protein stability in living cells. *Proc. Natl. Acad. Sci. U.S.A.* 2014, 111, 11335–11340.

(96) Freddolino, P. L.; Arkhipov, A. S.; Larson, S. B.; McPherson, A.; Schulten, K. Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure* 2006, 14, 437–449.

(97) Larsson, D. S. D.; Liljas, L.; van der Spoel, D. Virus capsid dissolution studied by microsecond molecular dynamics simulations. *PLoS Comput. Biol.* 2012, 8, No. e1002502.

(98) Larsson, D. S. D.; van der Spoel, D. Screening for the location of RNA using the chloride ion distribution in simulations of virus capsids. *J. Chem. Theory Comput.* 2012, 8, 2474–2483.

(99) de Oliveira dos Santos Soares, R.; Bortot, L. O.; van der Spoel, D.; Caliri, A. Membrane vesiculation induced by proteins of the dengue virus envelope studied by molecular dynamics simulations. *J. Phys.: Condens. Matter.* 2017, 29, No. 504002.

(100) Ghahremanpour, M. M.; van Maaren, P. J.; Caleman, C.; Hutchison, G. R.; van der Spoel, D. Polarizable drude model with s-type Gaussian or slater charge density for general molecular mechanics force fields. *J. Chem. Theory Comput.* 2018, 14, 5553–5566.

(101) Walz, M. M.; Ghahremanpour, M. M.; van Maaren, P. J.; van der Spoel, D. Phase-transferable force field for alkali halides. *J. Chem. Theory Comput.* 2018, 14, 5933–5948.

(102) Wennberg, C. L.; Murtola, T.; Hess, B.; Lindahl, E. Lennard-Jones Lattice Summation in Bilayer Simulations Has Critical Effects on Surface Tension and Lipid Properties. *J. Chem. Theory Comput.* 2013, 9, 3527–3537.

(103) Walters, E.; Mohebifar, M.; Johnson, E. R.; Rowley, C. N. Evaluating the London Dispersion Coefficients of Protein Force Fields Using the Exchange-Hole Dipole Moment Model. *J. Phys. Chem. B* 2018, 122, 6690–6701.

(104) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Moley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* 2010, 114, 2549–2564.

(105) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* 2013, 9, 5430–5449.

(106) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* 2018, E4758–E4766.