Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion

Tadashi Ando and Jeffrey Skolnick[1]

Center for the Study of Systems Biology, School of Biology, Georgia Institute of Technology, 250 14th Street North West, Atlanta, GA 30318-5304

Edited by Michael L. Klein, Temple University, Philadelphia, PA, and approved August 31, 2010 (received for review July 30, 2010)

To begin to elucidate the principles of intermolecular dynamics in the crowded environment of cells, employing Brownian dynamics (BD) simulations, we examined possible mechanism(s) responsible for the great reduction in diffusion constants of macromolecules in vivo from that at infinite dilution. In an *Escherichia coli* cytoplasm model comprised of 15 different macromolecule types at physiological concentrations, BD simulations of molecular-shaped and equivalent sphere representations were performed with a soft repulsive potential. At cellular concentrations, the calculated diffusion constant of GFP is much larger than experiment, with no significant shape dependence. Next, using the equivalent sphere system, hydrodynamic interactions (HI) were considered. Without adjustable parameters, the in vivo experimental GFP diffusion constant was reproduced. Finally, the effects of nonspecific attractive interactions were examined. The reduction in diffusivity is very sensitive to macromolecular radius with the motion of the largest macromolecules dramatically slowed down; this is not seen if HI dominate. In addition, long-lived clusters involving the largest macromolecules form if attractions dominate, whereas HI give rise to significant, size independent intermolecular dynamic correlations. These qualitative differences provide a testable means of differentiating the importance of HI vs. nonspecific attractive interactions on macromolecular motion in cells.

Brownian dynamics | correlated motion

One of the most characteristic features of the interiors of cells is the high total concentration of biological macromolecules. Typically, 20%–40% of the cytoplasmic volume is occupied by proteins, nucleic acids, and other macromolecules (1–3). Under these conditions, although the molar concentration of each protein ranges from nM to μM, the distance between neighboring proteins is comparable to the size of the proteins. Therefore, simulating the crowded intracellular environment is crucial to understanding the nature of living systems.

Macromolecular crowding exerts surprisingly large effects on the thermodynamics and kinetics of processes such as macromolecular association, protein stability, and enzyme activity (1, 4, 5). The diffusion and partitioning of macromolecules are highly restricted by intermolecular steric repulsions as well as nonspecific attractive interactions. Consequently, the in vivo and in vitro rates and equilibria of biological reactions can differ by orders of magnitude. While there have been several models of metabolic networks or signaling pathways designed to elucidate the relationship between molecular and cellular behavior (6), effects of macromolecular crowding are at best only partially considered (7).

Diffusion is one of the most important physical parameters that describe motions of molecules in a fluid. The diffusion constants of macromolecules in the cytoplasm as well as in membranes have been measured by various techniques including "single-particle tracking" (8), "fluorescence recovery after photobleaching (FRAP)" (9), and "fluorescence correlation spectroscopy" (10). All experiments show that the in vivo diffusion of proteins is greatly reduced compared to dilute conditions. Recently, Elowitz et al. (11) and Konopka et al. (12) applied FRAP to measure the diffusion coefficient of GFP in the *Escherichia coli* (*E. coli*) cytoplasm. Both groups reported that the diffusion coefficient of GFP in vivo is about 10 times less than that at infinite dilution in water. What is responsible for this reduction? To date, there have only been a few reports on the simulation of crowded, cytosol-like systems to analyze macromolecule motion. The first study was reported by Bicount and Field (13), where the cytoplasm was modeled as a mixture of three different types of spheres representing ribosomes, average proteins, and tRNAs, and Langevin dynamics simulations were performed using the DLVO (Derjaguin, Landau, Verwey, and Overbeek) potential. Ridgway et al. (14) and Roberts et al. (15) also modeled twelve different kinds of macromolecules as repulsive spheres and performed dynamic simulations. Although excluded volume effects significantly affected molecular motions, these interactions could not account for the factor of ∼10 reduction in GFP's diffusion constant. Recently, McGuffee and Elcock performed Brownian dynamics (BD) simulations using atomically detailed macromolecule models with electrostatic potentials (16). Despite inclusion of electrostatics, McGuffee and Elcock had to adjust the strength of the van der Waals interactions to reproduce GFP's in vivo diffusion constant (16). We note that the strength of attractive interactions can be tuned to set the magnitude of the diffusion constant anywhere from the dilute solution value to zero. Therefore, the mechanisms responsible for the large slow down of diffusion remain uncertain.

In this study, as a necessary first step towards whole cell modeling, we performed BD simulations of the *E. coli* cytoplasm to address the following issues in crowded, heterogeneous environments: (*i*) the effect of macromolecular shape on diffusion, (*ii*) the importance of hydrodynamic interactions (HI) on diffusion, and (*iii*) the differences in dynamic behavior when HI or attractive interactions dominate. To date, the effect of macromolecular shape has been ignored; rather, the single sphere per molecule model was used in simulations of crowded systems (13, 14, 17). While intermolecular HI play an important role in determining the dynamics of concentrated particles (18, 19), it has been neglected in the most simulations of biological macromolecules due to its long-range nature and high computational cost. Finally, how to differentiate the slow down in dynamics due to HI from that due to nonspecific attractive interactions has not been addressed.

The importance of HI on the dynamics of colloidal suspensions has been well studied using computational methods such as Stokesian dynamics (18), lattice-Boltzmann approaches (20), multiparticle collision dynamics (21), and dissipative particle dynamics (22). Stokesian dynamics, despite its computational cost, is advantageous because it is straightforward to compare simulations with and without HI, a lattice grid is unnecessary, and it reproduces the dynamic properties of homogeneous, monodisperse colloidal suspensions even at a volume fraction

Author contributions: J.S. designed research; T.A. performed research; T.A. and J.S. analyzed data; and T.A. and J.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: skolnick@gatech.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1011354107/-/DCSupplemental.

of ~0.5. Here, we apply Stokesian dynamics to simulate the diffusion of a polydisperse collection of macromolecules in crowded, heterogeneous intracellular three-dimensional environments.

Results

Estimation of Diffusion Tensor of a Macromolecule from Atomic Structure. To estimate diffusion tensors of macromolecules from their atomic structures, the rigid-particle formalism was used (23–25). As described in *Methods*, proteins are represented by their Cα beads, with a bead radius of 6.1 Å (the only optimized parameter) giving the best fit to the experimental translational and rotational diffusion coefficients of macromolecules at infinite dilution (see Table S1 and Fig. S1). For example, the method provides a translational diffusion coefficient for GFP of 8.9 $Å^2$/ns at 293 K, in excellent agreement with experiment, 8.7 $Å^2$/ns (26). To simulate the inside of cells, we also include nucleic acids. Using the same bead radius as proteins, the calculated translational diffusion coefficient of tRNA is 7.6 $Å^2$/ns at 293 K, consistent with the experimental value of 7.8 $Å^2$/ns (27).

Effect of Macromolecular Shape on Diffusion. What is the effect of explicit macromolecular shape on diffusion in crowded environments? Demonstration of a minor shape effect would enable HI to be calculated in a computationally tractable manner. We performed BD simulations with only steric repulsions for systems that consider either the explicit shape or the equivalent Stokes sphere radius. To compare the diffusivity of macromolecules in the two molecular representations as well as with experiment, we focus on translational diffusion. Rotational diffusion is also likely to play an important role in intracellular dynamics. We will analyze rotational diffusion in future work. Hereafter, "diffusion" refers to translational diffusion.

The *E. coli* protocell in the two different representations is shown in Fig. 1 (see also Fig. S2). As an example, at a total concentration of 300 mg/mL, Fig. S3 shows the energies, mean square displacements (MSD), and diffusion coefficients for several macromolecules as a function of time. From the point of view of energy, the systems equilibrate quickly (Fig. S3 *A* and *B*). As in other simulation studies on cytosol-like systems, crossover from anomalous to normal diffusion is observed for all molecules at short times <1 μs (Fig. S3 *C–F*) (14–16). Subsequently, a linear relationship between time and MSD is observed, with rapid convergence of the respective diffusion constants.

Trajectories of the explicit molecular-shaped and equivalent sphere systems are shown in Movie S1. In Fig. 2, the ratio of the long-time translational diffusion coefficients, D^L, observed in the virtual cytoplasm system to that estimated in dilute solution, D_0, as a function of Stokes radius is shown. Qualitatively consistent with other simulations (14–16) and experimental results on eukaryotic cells (2), for both systems, at three different concentrations, D^L/D_0 decreases with increasing molecular radius. The results of explicit molecular-shaped and equivalent sphere systems are very close over the entire range of radii for concentrations of 250 and 300 mg/mL. At 350 mg/mL, the long-time diffusion constants of the sphere system are slightly less than in the molecular-shaped system. These results suggest that at a macromolecular concentration of 300 mg/mL or less, effects of macromolecular shape on molecular diffusion in crowded environments are small and that the equivalent sphere per macromolecule model is a reasonable approximation for the analysis of in vivo diffusion.

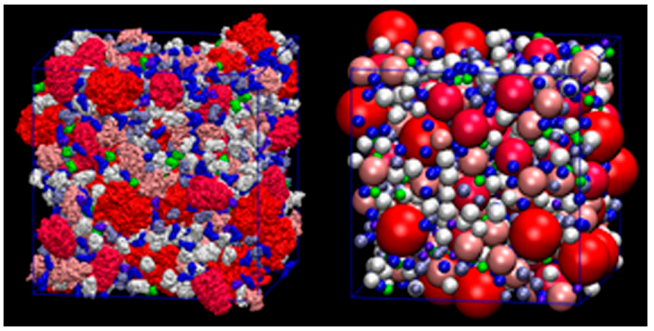

Fig. 1. Molecular-shaped (*left*) and sphere (*right*) systems at 300 mg/mL. Macromolecules are represented in different colors. Figures were generated by VMD (56).

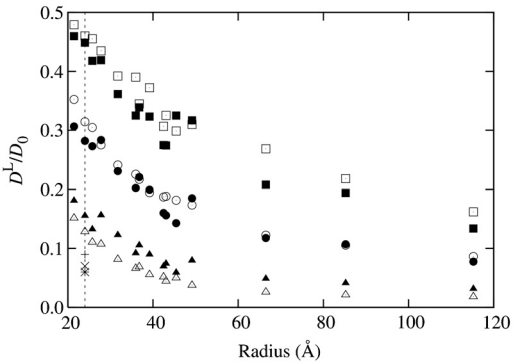

Fig. 2. Long-time diffusion constant ratio as a function of macromolecule radius in the sphere (open symbols) and molecular-shaped systems (filled symbols) with steric repulsion. Squares, circles, and triangles are at 250, 300, and 350 mg/mL, respectively. The reductions in diffusion constant of GFP measured in vivo of DH5α (11), BL21(DE3) (12), and K-12 *E. coli* (12) are shown by plus, cross, and asterisk, respectively. Dashed line is GFP's radius.

In in vivo experiments, D^L/D_0 of GFP is 0.06–0.09 (11, 12) (see Fig. 2). On the other hand, D^L/D_0 values of GFP in the simulated equivalent sphere system at 250 mg/mL and 300 mg/mL are 0.46 and 0.31, this is more than three times larger than experiment. Even for the molecular-shaped system at 350 mg/mL, the simulated diffusion constant of GFP is still two times larger than experiment. These results indicate that (consistent with other studies (14–16)), although excluded volume effects reduce macromolecular diffusion in intracellular environments, they cannot explain the factor of ~10–16 reduction observed in vivo.

Effect of Hydrodynamic Interactions on Diffusion. Next, we performed BD simulations with HI using equivalent sphere systems to evaluate the effects of HI on diffusion in crowded environments. In addition to the Rotne-Prager-Yamakawa (28, 29) interaction tensor, (widely used in biomolecular simulations to incorporate the long-range effects of HI), we also account for the lubrication forces that play a crucial role in the short-range interactions that are especially important in dense systems (18, 30). Because HI calculations are computationally expensive (computational costs with and without HI are roughly $O(N^3)$ and $O(N)$, respectively, with N the number of particles), we considered somewhat smaller systems containing ~400 macromolecules compared to the sphere system with just repulsive interactions (Table 1); without HI, the calculated diffusion constants are insensitive to whether the larger or smaller system is considered.

A representative trajectory is shown in Movie S1. Fig. S4*A*, where the normalized short-time diffusion coefficients, D^S/D_0, of GFP, GAPDH, and the ribosome (small, medium, and large macromolecules) at 300 mg/mL as a function of simulation time up to 15 μs are shown. It is clear that all macromolecules have reached their long-time limit. The MSD and diffusion coefficients of GFP, GAPDH, and the ribosome as a function of simulation time are shown in Fig. S4 *B*, *C*. As also seen in simulations without HI, crossover from anomalous to normal diffusion is observed at short times (<1 μs). MSD values after 5 μs were used to estimate the long-time diffusion coefficients of all macromolecules.

Table 1. Properties of simulated systems

	250 mg/mL		300 mg/mL		350 mg/mL	
Box size	$(100\ nm)^3$	$(75\ nm)^3$	$(100\ nm)^3$	$(71\ nm)^3$	$(100\ nm)^3$	$(67\ nm)^3$
Total number of molecules	1,000	423	1,152	412	1,295	390
Volume fraction in molecular-shaped system*	0.38	0.38	0.45	0.44	0.52	0.52
Volume fraction in sphere system	0.43	0.43	0.51	0.50	0.59	0.60

*Volumes of molecules were calculated using a bead radius of 6.1 Å.

Fig. 3 shows D^S/D_0 and D^L/D_0 values at three different concentrations as a function of Stokes radius. D^L/D_0 obtained from the BD simulations of sphere systems with just steric repulsions are also shown. Similar to the simulations without HI, D^S/D_0 and D^L/D_0 decrease with increasing radius. For D^S, HI greatly reduces the diffusion constants of all particles; in contrast, D^S is always equal to D_0 when HI are ignored; the reduction in short-time diffusion coefficient is a purely hydrodynamic property; D^S equals D_0 when HI are absent (31).

Interestingly, as shown in Fig 3*B*, BD simulations with HI for GFP in the 300 mg/mL system (a reasonable estimate of the macromolecular concentration in *E. coli*) gives D^L/D_0 of 0.08, in good agreement with the observed experimental values of 0.09 in DH5α (11), 0.07 in BL21(DE3) (12), and 0.06 in the K-12 strain (12) grown in a rich medium (Fig. 3*B*). These results indicate that steric crowding and HI are two major factors responsible for the reduction in diffusion of macromolecules in intracellular environments. Indeed, without any other assumptions, these two effects quantitatively reproduce the experimentally observed diffusion constant of GFP in vivo.

Effect of Nonspecific Attractive Interactions on Diffusion. Because nonspecific, attractive interactions are another plausible cause of the large reduction of in vivo diffusivity (4), we next examined their effects. We consider each macromolecule to be a sphere having a rough surface, which is filled with small van der Waals (vdW) particles with a 3 Å diameter (see *SI Text*: *Eqs*. **40**–**42**, with ϵ_{LJ} of 0.37 kcal/mol). Because the 300 mg/mL system equilibrated slowly from the point of view of energy due to the slow relaxation of interacting clusters (Fig. S5), simulations were performed up to 50 μs, and the first 30 μs were ignored in the calculation of diffusion constants.

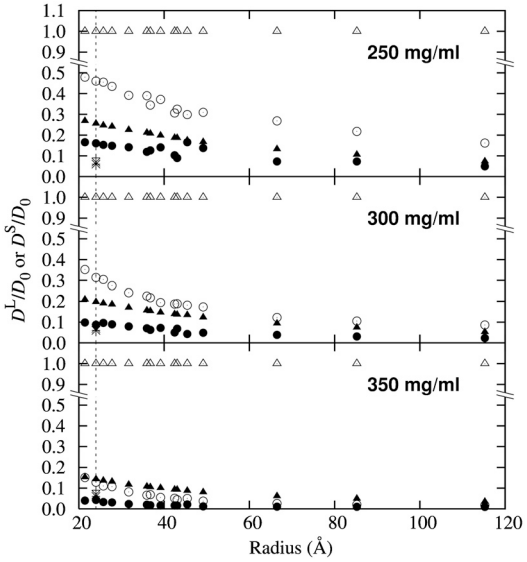

Fig. 3. Reduction in diffusivity as a function of radius for the system with HI at three different concentrations. Triangles and circles represent D^S/D_0 and D^L/D_0, respectively. Open (filled) symbols are values in the sphere model with repulsive (HI) interactions. Plus, cross, and asterisk symbols are as in Fig. 2. Dashed line is GFP's radius.

As shown in Fig. 4, in this model, D^L is very strongly dependent on macromolecular radius compared to the HI model. This same qualitative behavior is seen if the explicit molecular shape system with a Lennard-Jones potential is used or if we consider the rough sphere system with electrostatic interactions (see *SI Text*: *Eq*. **43** and Fig. S6); thus, the results are quite robust and invariant to the specific form of the potential. We would expect that if nonspecific attractive interactions dominate in vivo diffusion, the strength of this reduction would be very sensitive to molecular radius, whereas if HI dominate, the reduction in mobility is far less sensitive. Moreover, D^L for molecules whose radii exceed GFP have greatly reduced mobility and are almost immobile due to the formation of long-lived clusters (see Movie S1). Because there is a dramatic qualitative difference in predicted behavior, which mechanism dominates could be determined experimentally.

Large-Distance and Long-Time Intermolecular Correlations. Finally, the dynamical correlations in space and time between macromolecules were examined. Such effects are expected to be present when HI are included. To analyze the correlation between particles, we calculated a normalized pair correlation function, C_{ij} (see *SI Text*: *Eq*. **47**). C_{ij} ranges from −1 to 1. When two particles are positively correlated, $C_{ij} > 0$, and when they are negatively correlated, $C_{ij} < 0$.

Representative C_{ij} of large and small particle pairs up to 100 ns in time and 10 Å in space for the sphere system with and without HI as well as with the nonspecific attractive interaction model are shown in Fig. 5. In the model without HI, where $D^L \sim 3$ times larger than in the HI model, (Fig 5 top) $C_{ij} < 0.1$ even at short times (<30 ns) for both pairs. In contrast, for both pairs of molecules, a significant positive intermolecular dynamic correlation for the simulation with HI is evident, though these are on average weak, $C_{ij} < 0.3$. Due to short-range attractive interactions in the nonspecific binding model, for large molecules a positive correlation is seen for short distances (<5 Å), but this correlation rapidly decays in space but not time. This same qualitative behavior is seen when electrostatic interactions are allowed between all macromolecular pairs with the notable exception of ribosome-ribosome and ribosome-GroEL/ES pairs. For these molecules, due to their large negative charges, dynamical correlations are weak over the entire range of time and space.

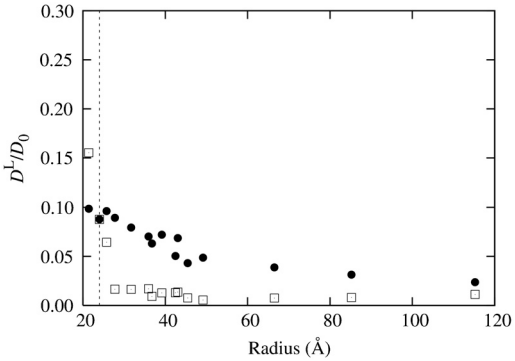

Fig. 4. At 300 mg/mL, long-time diffusion constant ratio, D^L/D_0, as a function of radius in the nonspecific, van der Waals interaction (HI) model is represented by squares (filled circles). Dashed line is GFP's radius.

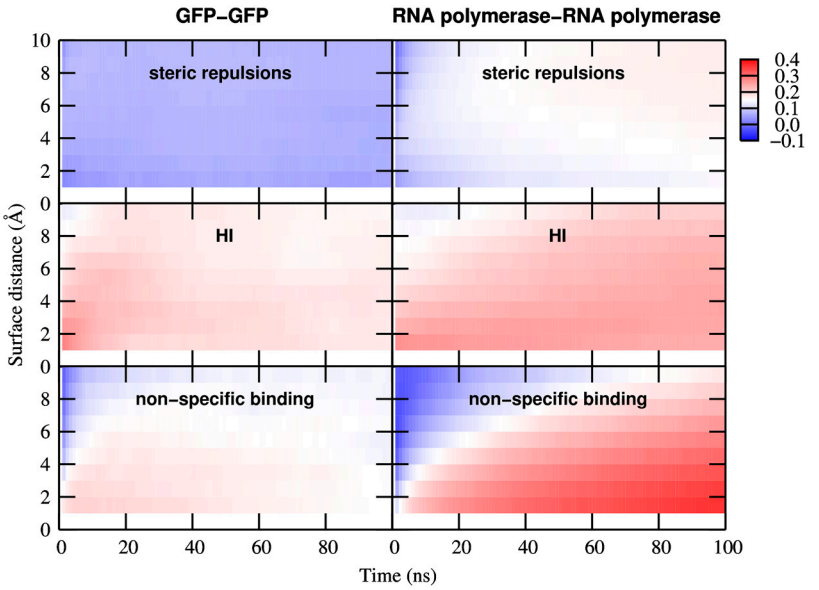

Fig. 5. Normalized pair correlation function, C_{ij}, averaged over GFP-GFP (*left*) and RNA polymerase-RNA polymerase (*right*) pairs for three different simulation models at 300 mg/mL. The Stokes radii of GFP and RNA polymerase are 24.0 and 66.5 Å, respectively.

When nonspecific attractive interactions dominate, consistent with the very small values of D^L, long-lived clusters at small intermolecular distances involving large macromolecules are seen. These results suggest that the dynamics in systems with HI and with nonspecific attractive interactions are qualitatively different. When HI are included, quite long distance and time dynamical correlations between particles of all sizes are observed but tight clusters of large particles are absent.

Discussion

The goal of this study is to evaluate the role of molecular shape, HI, and nonspecific attractive interactions in the reduction of macromolecular diffusion in intracellular environments. To ascertain the importance of shape details where only soft repulsions are allowed, we considered two molecular representations: a molecular-shaped system and an equivalent sphere per macromolecule. An important result is that the equivalent sphere per macromolecule model well describes the diffusion of macromolecules in intracellular environments. We next examined whether HI exert a significant effect on macromolecular dynamics. For the heterogeneous, crowded environment typical of cells, these BD simulations take into account not only far-field HI but also near-field lubrication effects in three dimensions. The key finding of this work is that the two factors, excluded volume effects and HI, are sufficient to explain the large reduction in diffusion of macromolecules observed in vivo. Indeed, the diffusion constant of GFP in vivo can be quantitatively predicted by the inclusion of HI *without any adjustable parameters or other* ad hoc *assumptions*.

Nonspecific attractive interactions (e.g., vdW or contact potentials) represent another plausible mechanism that could cause the large reduction of diffusivity in vivo. However, by tuning the magnitude of intermolecular attractive interactions, we can scale the diffusion constant between the infinite dilution value and zero. How can one experimentally differentiate if nonspecific, attractive interactions or HI dominate macromolecular diffusion? One of the biggest differences between vdW and HI is their interaction range: HI are long-range interactions which decay $\sim 1/r$, while vdW are short-range interactions that decay rapidly as $1/r^6$ on the atomic scale (although their cumulative effects act at long-range with a large sensitivity to particle size). The other big difference is in the direction of the forces: vdW is always attractive between two particles except when particles overlap, while the hydrodynamic force is repulsive for particles approaching each other and attractive for particles moving apart. Thus, the following qualitative differences in dynamic behavior result: First, the short-time diffusion constant, D^S, is the same as the infinite dilution, D_0, value when attractive interactions dominate, whereas if HI dominate, D^S should differ significantly from D_0, with a difference that is size dependent. Second, for nonspecific attractions, the reduction in long-time diffusion constant should be strongly size dependent, with long-lived clusters forming with larger macromolecules that dramatically reduces their diffusion constant. If HI dominate, D^L should exhibit a much weaker dependence on particle radius. Third, if HI are important, significant correlations in intermolecular dynamics for all size macromolecules should be observed in both space and time. These correlated motions are entirely absent in purely repulsive systems that ignore HI. With attractions, long-lived short distance correlations resulting from long-lived cluster formation should be seen for the larger macromolecules; this acts to dramatically reduce the diffusion constant of large molecules. We suspect that this large reduction in diffusivity is a nonphysical effect, and possibly precludes nonspecific intermolecular interactions as exerting a dominant effect on the dynamics.

A number of other factors can also affect intracellular diffusion: (*i*) *Electrostatic interactions between molecules*. In principle, electrostatic interactions are long-ranged. However, the salt concentration inside cells is ~150 mM, so that they are well screened with a Debye length of ~8 Å. McGuffee and Elcock recently simulated a bacterial cytoplasm model where electrostatic interactions were treated by using Poisson-Boltzmann equations (16). However, the diffusion coefficient of GFP is just slightly smaller than that obtained without electrostatic interactions; both values were 3–4 times larger than experiment. Moreover as shown in *SI Text*, inclusion of electrostatic interactions at the level of a linearized Poisson-Boltzmann equation using a single charge per one molecule does not modify the qualitative conclusions. Heterogeneous charge distributions on molecular surface, like in real biomolecules, may affect macromolecular motions. However, because electrostatic interactions are highly screened due to the short Debye length found in physiological conditions, we believe that our conclusions would not qualitatively change. (*ii*) *Viscosity of the cytoplasm*. In our simulations, the viscosity of the cytoplasm equals the value in water. The in vivo cytoplasm viscosity has been

measured and is not significantly larger than bulk water, i.e., it is less than 2 centipoise (2, 14, 32). (*iii*) *GFP dimerization*. It is well known that GFP tends to dimerize in solutions of low (<100 mM) ionic strength (33). All of these physical factors will decrease macromolecular diffusivity in vivo. However, based on other work (16) and our results, which show that the reduction in diffusion of GFP in the simulation with HI is very close to experiment, we expect the contribution of these three factors to be small.

Our results have demonstrated the likely importance of HI in macromolecular diffusion in vivo. However, there are a few possible limitations: First, the properties of a fluid on the nanometer scale are different from the bulk (34–36). HI determined by solving Stokes equations may not fully describe the molecular situation. (But the ability of Stokesian dynamics to describe the diffusive behavior of macromolecular colloid solutions suggests that such discrepancies in practice are minor). In order to fully validate the continuum limit assumption, molecular dynamics simulations with explicit solvent models would be necessary. Second, HI were considered only for the equivalent sphere system where the detailed molecular shape is ignored. Without HI, we demonstrated that this representation is a very good approximation, but we have not explicitly shown this for the system with HI. Recently, an analytical formula that estimates the crossover time from anisotropic to isotropic diffusion of an arbitrarily shaped object in three dimensions using its $6N \times 6N$ diffusion tensor matrix was introduced (37). The longest crossover time of molecule in our simulation system estimated by using this formula is 1.7 μs for ribosome. Therefore, the effect of shape and diffusion anisotropy on the analysis of long-time translational diffusion is expected to be small. Third, our method is computationally very expensive. For much longer and larger simulations, further improvements in computational efficiency are essential. Banchio and Brady introduced a unique idea into BD simulations with HI to reduce the computational cost from $O(N^3)$ to $O(N^{1.25} \ln N)$, where a mean field approximation for far-field HI and a Chebyshev polynomial instead of Cholesky decomposition are applied (38). Using or developing an approximate method such as local-density dependent HI (39, 40), is another possibility. However, the aforementioned methods were developed for homogeneous colloidal suspensions. Therefore, to develop a new method for heterogeneous cellular systems, a series of simulations based purely on hydrodynamic theory, like this study, must be performed. These simulations can serve as the reference system to assess approximate methods.

Diffusion in concentrated colloidal dispersions is well studied experimentally as well as theoretically. For monodisperse, hard sphere spherical particles, analytical equations that describe the concentration dependence of the short-time and long-time self-diffusion coefficients have been derived by Tokuyama and Oppenheim (41, 42). These theoretical values are in good agreement with experimental results on diffusivity even for condensed systems. Average values of short-time and long-time diffusion coefficients obtained from our simulations at three different concentrations decrease with increasing volume fraction, which are qualitatively similar to Tokuyama and Oppenheim's theoretical analysis. However, the short-time diffusion coefficients are smaller and the long-time diffusion coefficients are larger than the theoretical values with deviation less than 0.1. These small deviations may reflect differences between the monodisperse systems assumed by the theory and the heterogeneous systems under consideration. Additionally, crowding effects on the macromolecular motions are usually analyzed in the presence of a single type of crowding agent. A recent simulation study showed that the helix-coil transition temperature of a flexible helical homopolymer depends on the size of the crowding particles even at an identical volume fraction (43). This dependence may imply that the heterogeneity of the environment is important for understanding crowding effects in vivo. From the point of this view, our simulation method would be a good tool to analyze the in vivo thermodynamics and kinetics of macromolecular dynamics.

In conclusion, genome-sequencing has provided a detailed "parts list" for life (44). Recently, the proteome wide prediction of protein structure and function has also become practical (44–50). The next frontier in biophysics is to integrate this information and construct *in silico* cells that not only can describe the behavior of living systems in terms of individual biomolecules but which also can elucidate new biological principles describing their collective behavior. Until now, little attention has been paid to the biophysical properties of the crowded, heterogeneous environments found in cells, which have a great impact on the biological processes taking place. Therefore, modeling these crowding effects is an important first step towards whole cell simulation.

In that spirit, by conducting a series of BD simulations, the following conclusions were obtained: First, excluded volume effects and HI are likely the two major factors that likely account for the large reduction of diffusion of macromolecules in vivo. Second, representing a molecule by one sphere is a reasonable approximation for analyzing in vivo macromolecular diffusion. Third, the effects of HI can be experimentally demonstrated by examining the reduction in short-time and long-time diffusion constants over the infinite dilution values, and by exploring the presence of intermolecular dynamic correlations at times scales on the hundreds of nanoseconds. If validated, this experimental result will provide insight into the hitherto unsuspected importance of HI on intracellular dynamics.

Methods

The detailed description of the methodology employed here is described in *SI Text* with the salient points presented below.

Estimation of Diffusion Tensor of a Macromolecule from Its Atomic Structure. Proteins and nucleic acids are represented by their Cα and P, C4′, N1, and N9 beads, respectively. To account for the shape anisotropy of macromolecules, their infinite dilution diffusion tensors were calculated using rigid-particle theory (23–25). In this theory, the radius of the beads is the only parameter needed as input. The bead radius was optimized using twelve different proteins whose molecular mass ranges from 6 kDa to 230 kDa, (24) (Table S1 and Fig. S1) to reproduce their diffusion constants in dilute conditions.

Construction of the Intracellular Environment. We built systems mimicking the *E. coli* cytoplasm environment based on the data reported by Ridgway et al (14) and the CyberCell database of the physical properties of *E. coli* (51). The concentration of macromolecules in the *E. coli* cytoplasm is estimated to be in the range of 300–400 mg/mL (3). Therefore, intracellular models at three different concentrations, 250, 300, and 350 mg/mL were constructed that contain 15 different types of macromolecules abundant in the *E. coli* cytoplasm (52). These macromolecules include ribosomes, chaperonin GroEL/ES, RNA polymerase, glycolytic and some other enzymes, tRNA, and GFP (Fig. S2). For molecular-shaped systems (see Fig 1, left), molecules were represented by strings of beads described in the previous section. The number and types of macromolecules are summarized in Table 1 and Table S3, with their estimated hydrodynamic properties presented in Table S2.

To investigate the effects of molecular shape on diffusion, we constructed a system (see Fig 1, right) where each macromolecule i was represented by an equivalent sphere Stokes radius a_i, is given by $6\pi\eta a_i = k_B T / D_{0,i}$, in which $D_{0,i}$ is the calculated translational diffusion coefficient of molecule i (see Table S2) at infinite dilution.

Hydrodynamic Interactions. In order to include not only the far-field HI but also the many-body and near-field HI in simulations, the Durlofsky-Brady-Bossis approach to Stokesian dynamics was used (18, 53).

Brownian Dynamics Algorithm. For a BD simulation without HI, the integration scheme developed by Ermak and McCammon (54) was used. On including HI, in principle, the diffusion tensor of each molecule depends on the instantaneous configuration of the entire system (but we would expect screening to emerge at higher concentrations). To explore the role of HI, the modified

midpoint BD algorithm introduced by Banchio and Brady (38), and based on Fixman's idea (55) was used.

Potential Functions. Repulsive interactions between intermolecular particles in BD simulations without HI are represented by a soft-sphere, harmonic potential, (see *SI Text: Eq.* **38**). In BD simulations with HI, we do not implement explicit repulsive forces between particles, because lubrication forces prevent particle overlap. For the nonspecific binding model, attractive interactions were described by Lennard-Jones potential function (see also *SI Text: Eqs.* **40–42**). In this model, the surface roughness of macromolecule was also considered by including a correction factor. The potential was then optimized to reproduce the in vivo mobility of GFP.

Simulation Conditions and Analysis. All simulations were performed at 298 K with periodic boundary conditions. For all systems, ten independent simulations were run, each with different, randomly generated initial configurations. For BD simulations of repulsive and nonspecific, attractive binding models without HI, 30 and 50 μs simulations were performed with a time step of 0.5 and 0.1 ps, respectively. For BD simulations with HI, we ran 15 μs simulations with a time step of 2 ps. For simulations with nonspecific binding and HI, we used a smaller system than that used for the strictly repulsive model. The first 5, 30, and 5 μs of simulations of repulsive, nonspecific binding models, and HI models were ignored in our analysis, respectively. To estimate the long-time diffusion coefficients of all the macromolecules, MSD values after a time interval of 5 μs were used.

ACKNOWLEDGMENTS. This work was supported in part by NIH Grant No GM-37408. J.S. and T.A. conceived the project, T.A. performed the algorithm development and calculations.

1. Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* 26:597–604.
2. Luby-Phelps K (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int Rev Cytol* 192:189–221.
3. Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J Mol Biol* 222:599–620.
4. Zhou HX, Rivas GN, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Ann Rev Biophys* 37:375–397.
5. Zimmerman SB, Minton AP (1993) Macromolecular crowding—biochemical, biophysical, and physiological consequences. *Annu Rev Bioph Biom* 22:27–65.
6. Di Ventura B, Lemerle C, Michalodimitrakis K, Serrano L (2006) From in vivo to in silico biology and back. *Nature* 443:527–533.
7. Takahashi K, Arjunan SNV, Tomita M (2005) Space in systems biology of signaling pathways—towards intracellular molecular crowding in silico. *Febs Lett* 579:1783–1788.
8. Jin S, Verkman AS (2007) Single particle tracking of complex diffusion in membranes: simulation and detection of barrier, raft, and interaction phenomena. *J Phys Chem B* 111:3625–3632.
9. Verkman AS (2003) Diffusion in cells measured by fluorescence recovery after photobleaching. *Method Enzymol* 360:635–648.
10. Bacia K, Kim SA, Schwille P (2006) Fluorescence cross-correlation spectroscopy in living cells. *Nat Methods* 3:83–89.
11. Elowitz MB, Surette MG, Wolf PE, Stock JB, Leibler S (1999) Protein mobility in the cytoplasm of *Escherichia coli*. *J Bacteriol* 181:197–203.
12. Konopka MC, Shkel IA, Cayley S, Record MT, Weisshaar JC (2006) Crowding and confinement effects on protein diffusion in vivo. *J Bacteriol* 188:6115–6123.
13. Bicout DJ, Field MJ (1996) Stochastic dynamics simulations of macromolecular diffusion in a model of the cytoplasm of *Escherichia coli*. *J Phys Chem* 100:2489–2497.
14. Ridgway D, et al. (2008) Coarse-grained molecular simulation of diffusion and reaction kinetics in a crowded virtual cytoplasm. *Biophysical Journal* 94:3748–3759.
15. Roberts E, Stone JE, Sepulveda L, Hwu W-MW, Luthey-Schulten Z (2009) Long time-scale simulations of in vivo diffusion using GPU hardware. *Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing* (IEEE Computer Society, Washington, DC), pp 1–8.
16. McGuffee SR, Elcock AH (2010) Diffusion, crowding, and protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput Biol* 6:e1000694.
17. Dix JA, Verkman AS (2008) Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys* 37:247–263.
18. Brady JF, Bossis G (1988) Stokesian dynamics. *Annu Rev Fluid Mech* 20:111–157.
19. Russel WB, Saville DA, Schowalter WR (1989) *Colloidal dispersions* (Cambridge University Press, Cambridge, New York), pp xvii, 525 p, 521 leaf of plates.
20. Ladd AJC (1993) Short-time motion of colloidal particles—numerical-simulation via a fluctuating lattice-boltzmann equation. *Phys Rev Lett* 70:1339–1342.
21. Malevanets A, Kapral R (1999) Mesoscopic model for solvent dynamics. *J Chem Phys* 110:8605–8613.
22. Groot RD, Warren PB (1997) Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J Chem Phys* 107:4423–4435.
23. Carrasco B, Garcia de la Torre J (1999) Hydrodynamic properties of rigid particles: comparison of different modeling and computational procedures. *Biophys J* 76:3044–3057.
24. Garcia De La Torre J, Huertas ML, Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* 78:719–730.
25. Garcia De La Torre J, Jimenez A, Freire JJ (1982) Monte-Carlo calculation of hydrodynamic properties of freely jointed, freely rotating, and real polymethylene chains. *Macromolecules* 15:148–154.
26. Terry BR, Matthews EK, Haseloff J (1995) Molecular characterization of recombinant green fluorescent protein by fluorescence correlation microscopy. *Biochem Bioph Res Co* 217:21–27.
27. Potts R, Fournier MJ, Ford NC (1977) Effect of aminoacylation on conformation of yeast phenylalanine transfer-Rna. *Nature* 268:563–564.
28. Rotne J, Prager S (1969) Variational treatment of hydrodynamic interaction in polymers. *J Chem Phys* 50:4831–4837.
29. Yamakawa H (1970) Transport properties of polymer chains in dilute solution—hydrodynamic interaction. *J Chem Phys* 53:436–443.
30. Phillips RJ, Brady JF, Bossis G (1988) Hydrodynamic transport-properties of hard-sphere dispersions .1. suspensions of freely mobile particles. *Phys Fluids* 31:3462–3472.
31. Brady JF (1994) The long-time self-diffusivity in concentrated colloidal dispersions. *J Fluid Mech* 272:109–133.
32. Verkman AS (2002) Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci* 27:27–33.
33. Yang F, Moss LG, Phillips GN (1996) The molecular structure of green fluorescent protein. *Nat Biotechnol* 14:1246–1251.
34. Benz M, Chen NH, Jay G, Israelachvili JI (2005) Static forces, structure and flow properties of complex fluids in highly confined geometries. *Ann Biomed Eng* 33:39–51.
35. Bhushan B, Israelachvili JN, Landman U (1995) Nanotribology—friction, wear and lubrication at the atomic-scale. *Nature* 374:607–616.
36. Maali A, Bhushan B (2008) Nanorheology and boundary slip in confined liquids using atomic force microscopy. *J Phys-Condens Mat* 20:315201.
37. Schluttig J, Korn CB, Schwarz US (2010) Role of anisotropy for protein-protein encounter. *Phys Rev E* 81:030902.
38. Banchio AJ, Brady JF (2003) Accelerated Stokesian dynamics: Brownian motion. *J Chem Phys* 118:10323–10332.
39. Heyes DM (1996) Brownian dynamics simulations of self and collective diffusion of near hard sphere colloidal liquids: inclusion of many-body hydrodynamics. *Mol Phys* 87:287–297.
40. Urbina-Villalba G, Garcia-Sucre M, Toro-Mendoza J (2003) Average hydrodynamic correction for the Brownian dynamics calculation of flocculation rates in concentrated dispersions. *Phys Rev E* 68:061408.
41. Tokuyama M, Oppenheim I (1995) On the theory of concentrated hard-sphere suspensions. *Physica A* 216:85–119.
42. Tokuyama M, Oppenheim I (1995) Dynamics of self-diffusion process in concentrated hard-sphere suspensions of interacting Brownian particles. *J Korean Phys Soc* 28:S327–S332.
43. Kudlay A, Cheung MS, Thirumalai D (2009) Crowding effects on the structural transitions in a flexible helical homopolymer. *Phys Rev Lett* 102:118101.
44. Skolnick J, Fetrow JS (2000) From genes to protein structure and function: novel applications of computational approaches in the genomic era. *Trends Biotechnol* 18:34–39.
45. Arakaki AK, Huang Y, Skolnick J (2009) EFICAz2: enzyme function inference by a combined approach enhanced by machine learning. *BMC Bioinformatics* 10:107.
46. Baker D, Sali A (2001) Protein structure prediction and structural genomics. *Science* 294:93–96.
47. Brylinski M, Skolnick J (2010) Comparison of structure-based and threading-based approaches to protein functional annotation. *Proteins* 78:118–134.
48. Gao M, Skolnick J (2008) DBD-Hunter: a knowledge-based method for the prediction of DNA-protein interactions. *Nucleic Acids Res* 36:3978–3992.
49. Pandit SB, et al. (2010) PSiFR: an integrated resource for prediction of protein structure and function. *Bioinformatics* 26:687–688.
50. Zhang Y, Skolnick J (2004) Automated structure prediction of weakly homologous proteins on a genomic scale. *Proc Natl Acad Sci USA* 101:7594–7599.
51. Sundararaj S, et al. (2004) The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of *Escherichia coli*. *Nucleic Acids Res* 32:D293–D295.
52. Ishihama Y, et al. (2008) Protein abundance profiling of the *Escherichia coli* cytosol. *Bmc Genomics* 9:102.
53. Durlofsky L, Brady JF, Bossis G (1987) Dynamic simulation of hydrodynamically interacting particles. *J Fluid Mech* 180:21–49.
54. Ermak DL, Mccammon JA (1978) Brownian Dynamics with hydrodynamic interactions. *J Chem Phys* 69:1352–1360.
55. Fixman M (1978) Simulation of polymer dynamics .1. general theory. *J Chem Phys* 69:1527–1537.
56. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graphics* 14:33–38.