Published in final edited form as:
Quantum Bioinform V (2011). 2013 March ; 30: 375–387. doi:10.1142/9789814460026_0032.

IMPORTANCE OF EXCLUDED VOLUME AND HYDRODYNAMIC INTERACTIONS ON MACROMOLECULAR DIFFUSION *IN VIVO*

TADASHI ANDO and
Center for the Study of Systems Biology, School of Biology, Georgia Institute of Technology 250 14th Street NW, Atlanta, GA 30318-5304, USA

JEFFREY SKOLNICK
Center for the Study of Systems Biology, School of Biology, Georgia Institute of Technology 250 14th Street NW, Atlanta, GA 30318-5304, USA

Abstract

The interiors of all living cells are highly crowded with macromolecules, which results in a considerable difference between the thermodynamics and kinetics of biological reactions *in vivo* from that *in vitro*. 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 *E. coli* cytoplasm modelcomprised of 15 different macromolecule types at physiological concentrations, where macromolecules were represented by spheres with their Stokes radii, BD simulations were performed with and without hydrodynamic interactions (HI). Without HI, the calculated diffusion constant of green fluorescent protein (GFP) is much larger than experiment. On the other hand, when HI were considered, the *in vivo* experimental GFP diffusion constant is almost reproduced without adjustable parameters. In addition, HI give rise to significant, size independent intermolecular dynamic correlations. These results suggest that HI play an important role on macromolecular dynamics *in vivo*.

1. Introduction

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.

Diffusion is one of the most important physical parameters that describe motions of molecules in a fluid. Recently, Elowitz et al. [4] and Konopka et al. [5] applied fluorescence recovery after photobleaching to measure the diffusion coefficient of green fluorescent protein (GFP) in the *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? In this study, as a necessary first step towards whole cell modeling, we performed Brownian dynamics (BD) simulations of the *E. coli* cytoplasm to address the

importance of hydrodynamic interactions on diffusion. While intermolecular HI play an important role in determining the dynamics of concentrated particles [6, 7], it has been neglected in the most simulations of biological macromolecules due to its long-range nature and high computational cost. Here, we apply Stokesian dynamics to simulate the diffusion of a polydisperse collection of macromolecules in crowded, heterogeneous intracellular 3D environments.

2. Methods

2.1. Hydrodynamic Interactions

The Rotne-Prager-Yamakawa (RPY) tensor is commonly used to take into account the HI in dynamic simulations of biomolecules because this tensor retains positive definiteness even when particles overlap [8-10]. Actually, the RPY tensor contains the two-body long-range or far-field contributions to particle mobility. However, in concentrated systems, e.g. the inside of cells, far-field many-body HI as well as near-field HI, so-called “lubrication forces”, play important roles in determining mobility (see Fig. 4 in Ref. [11]). 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 [6, 12]. Their approach can reproduce the properties of dense monodisperse suspensions whose volume fraction is up to 0.50 [13, 14]. Here, we briefly outline this method. For details, please refer to the original reference [12].

For N particles system in a Newtonian fluid and in absence of an external shear flow, the hydrodynamic forces acting on particles, **F**, are related to the particle velocities, **U**, through the Stoke equation:

$$\mathbf{F}=\mathbf{R}\cdot\mathbf{U}, \quad (1)$$

where **R** is the resistance matrix and is the inverse of the mobility matrix, **M**:

$$\mathbf{R}=\mathbf{M}^{-1}. \quad (2)$$

When torque-angular velocity is not considered, the so-called “F version” in Ref. [12], **F** and **U** are $3N \times 1$ vectors and **R** and **M** are $3N \times 3N$ matrices. Then, the diffusion matrix of the system is simply given by

$$\mathbf{D}=k_{\mathrm{B}}T\mathbf{M}. \quad (3)$$

Here, k_{B} is Boltzmann constant and T is the temperature. The resistance tensor **R**, which contains both near-field lubrication effects and far-field many-body interactions, is calculated as

$$\mathbf{R}=(\mathbf{M}^{\infty})^{-1}+\mathbf{R}_{2\mathrm{B}}-\mathbf{R}_{2\mathrm{B}}^{\infty}. \quad (4)$$

The first term, $(\mathbf{z}^{\infty})^{-1}$, represents the contribution of many-body, far-field interactions. The second term, $\mathbf{R}_{2\mathrm{B}}$, represents the exact two-body HI, which includes both near-field and far-

field interactions. The third term $\mathbf{R}_{2B}{}^{\infty}$ is the resistance tensor that represents two-body far-field interactions. The far-field part has already been included on $(\mathbf{M}^{\infty})^{-1}$. Thus, in order not to count these interactions twice, we must subtract off the two-body interactions. This is the standard method to correct for the lubrication effects in the resistance tensor.

$\mathbf{M}^{\infty}$ can be estimated by the Ewald summation of the RPY tensor developed by Beenakker for the case of periodic boundary conditions [15]. Due to the long range nature of HI (which decays as $1/r$, with r the intermolecular distance between particles), use of the Ewald summation technique is necessary not only for accuracy but also for obtaining positive definite matrices in the calculation of the mobility tensor under periodic boundary conditions [14]. In addition, inverting $\mathbf{M}^{\infty}$ corresponds to including many-body interactions [12]. $\mathbf{R}_{2B}$ is calculated by the exact two-body solution of Jeffrey and Onishi [16]. $\mathbf{R}_{2B}{}^{\infty}$ is obtained by simply inverting a two-body mobility matrix containing terms to the same order in $1/r$ as $\mathbf{M}^{\infty}$.

When lubrication effects were added to the resistance tensor, the correction method developed by Cichocki et al. [17] is used to prevent a divergence of the translational self-diffusion coefficient for certain configurations. Collective motions are separated out from the standard lubrication correction as follows [17]:

$$\mathbf{R}=(\mathbf{M}^{\infty})^{-1}+\mathbf{s}_{2B}-\mathbf{s}_{2B}^{\infty}, \quad (5)$$

where

$$\mathbf{s}_{2B}=\sum_{\alpha}^{N}\sum_{\beta>\alpha}^{N}\mathbf{q}^{T}\cdot\mathbf{R}_{2B}(\alpha,\beta)\cdot\mathbf{q}, \quad (6)$$

$$\mathbf{s}_{2B}^{\infty}=\sum_{\alpha}^{N}\sum_{\beta>\alpha}^{N}\mathbf{q}^{T}\cdot\mathbf{R}_{2B}^{\infty}(\alpha,\beta)\cdot\mathbf{q}. \quad (7)$$

Here, α and β represents indices of particles, and the matrix $\mathbf{q}$ is given by

$$\mathbf{q}=\frac{1}{2}\begin{pmatrix}\mathbf{I} & -\mathbf{I}\\ -\mathbf{I} & \mathbf{I}\end{pmatrix}. \quad (8)$$

This modified correction method was essential to prevent a divergence of the translational self-diffusion coefficient for our crowded and heterogeneous system.

2.2. Brownian Dynamics Algorithm

When HI are considered, the diffusion tensor macromolecule depends in principle on the configuration of the entire system and varies over time. We can write the propagation equation for such Brownian particles as

$$\begin{aligned}\mathbf{r} = & \mathbf{r}^0 + (\nabla \cdot \mathbf{D})\,\Delta t + \frac{\mathbf{D}\cdot\mathbf{F}^{\mathrm{p}}}{k_{\mathrm{B}}T}\Delta t + \mathbf{G}(\Delta t) \\ = & \mathbf{r}^0 + k_{\mathrm{B}}T(\nabla \cdot \mathbf{M})\,\Delta t + (\mathbf{M}\cdot\mathbf{F}^{\mathrm{p}})\,\Delta t + \mathbf{G}(\Delta t),\end{aligned} \quad (9)$$

where **r** is the particle's position vector and $\mathbf{F}^{\mathrm{p}}$ is the deterministic conservative force acting on the particle. **G**(Δ*t*) is the random displacement due to Brownian motion, which has the following properties:

$$\langle \mathbf{G}(\Delta t)\rangle = 0, \langle \mathbf{G}(\Delta t)\,\mathbf{G}(\Delta t)\rangle = 2k_{\mathrm{B}}T\mathbf{M}\Delta t. \quad (10)$$

In contrast to a BD algorithm with constant diffusion tensors, we need to evaluate the spatial gradient of the mobility tensor in the BD simulation with HI, in which the explicit computation of ∇. **M** is a O(N^3) task. To avoid this expensive calculation, we used a method introduced by Banchio and Brady [18], which is based on Fixman's idea [19], the so-called “mid-point scheme”. In this method, Eqs. 9 and 10 are re -written as follows:

$$\mathbf{r} = \mathbf{r}^0 + k_{\mathrm{B}}T(\nabla \cdot \mathbf{M})\,\Delta t + \mathbf{M}\cdot\left(\mathbf{F}^{\mathrm{p}} + \mathbf{F}^{\mathrm{B}}\right)\Delta t, \quad (11)$$

$$\left\langle \mathbf{F}^{\mathrm{B}}\right\rangle = 0, \left\langle \mathbf{F}^{\mathrm{B}}(0)\,\mathbf{F}^{\mathrm{B}}(t)\right\rangle = 2k_{\mathrm{B}}T\mathbf{R}/\Delta t. \quad (12)$$

Here, $\mathbf{F}^{\mathrm{B}}$ is the Brownian force, obtained through the Cholesky decomposition [20]. The procedure for this mid-point algorithm is the following:

1. Compute the velocity $\mathbf{U}^0$ using an initial configuration $\mathbf{r}^0$

$$\mathbf{U}^0 = \left(\mathbf{R}^0\right)^{-1}\cdot\left(\mathbf{F}^{\mathrm{p},0} + \mathbf{F}^{\mathrm{B},0}\right). \quad (13)$$

   Here, superscript 0 represents the value evaluated at position $\mathbf{r}^0$.

2. Move the particles to intermediate positions **r**’ by a small fraction of a time step, Δ*t*/*m*

$$\mathbf{r}' = \mathbf{r}^0 + \frac{\Delta t}{m}\mathbf{U}^0. \quad (14)$$

   In this study, an *m* of 100 was used.

3. Calculate a new velocity **U**’ at the intermediate positions using the forces evaluated at $\mathbf{r}^0$

$$\mathbf{U}' = \left(\mathbf{R}'\right)^{-1}\cdot\left(\mathbf{F}^{\mathrm{p},0} + \mathbf{F}^{\mathbf{B},0}\right). \quad (15)$$

4. Calculate the drift velocity, $\mathbf{U}^{\mathrm{drift}}$,

$$\mathbf{U}^{drift} = \frac{m}{2}\left(\mathbf{U}' - \mathbf{U}^0\right). \quad (16)$$

5. Finally, update the positions of the particles for time step Δ*t*.

$$\mathbf{r}=\mathbf{r}^0+\left(\mathbf{U}^0+\mathbf{U}^{drift}\right)\Delta t. \quad (17)$$

For the simulations without HI, Ermak and McCammon algorithm described by Eqs. 9 and 10 was used [20, 21].

2.3. Simulation System

A simulation system having over 1,000 macromolecules consisting of 15 different kinds of proteins and tRNA was constructed in a 100 nm × 100 nm × 100 nm box based on the data reported by Ridgway et al [22] and the CyberCell database of the physical properties of *E. coli* [23], where each macromolecule was represented by an equivalent sphere with their Stokes radii (Table 1, Fig. 1). Stokes radii were estimated by using rigid-particle theory [21, 24]. The macromolecular concentration was set to 300 mg/ml, which is a reasonable estimate of the macromolecular concentration in *E. coli* (300−340 mg/ml) [3]. Volume occupancies calculated with the Stokes radii and radii of gyration of the macromolecules are 51% and 22%, respectively.

2.4. Repulsive Interaction Model

Repulsive interactions between intermolecular particles in BD simulations without HI were represented by a soft-sphere potential described by

$$V_{ss}(r_{ij})=\begin{cases} k_{ss}(r_{ij}-r_{\mathrm{m}})^2 & \text{if} \quad r_{ij}\le r_{\mathrm{m}} \\ 0 & \text{if} \quad r_{ij}>r_{\mathrm{m}}, \end{cases} \quad (18)$$

where r_{ij} is the distance between particles *i* and *j*, and k_{ss} is a force constant. r_{m} is $r_{\mathrm{c}} + \Delta_{ss}$, in which r_{c} is sum of radii of particles *i* and *j*, a_i and a_j, and Δ_{ss} is an arbitrary parameter representing a buffer distance between particles. In this study, a Δ_{ss} of 2 Å and k_{ss} of $5k_{\mathrm{B}}T/\Delta^2_{ss}$ were used, which means $V_{ss} = 5k_{\mathrm{B}}T$ at the distance r_{c}. In BD simulations with lubrication forces, we do not use any repulsive forces between particles, since the lubrication forces prevent particles from overlapping.

2.5. Simulation Conditions and Analysis

All simulations were performed at 298 K with periodic boundary conditions. For all simulation systems, ten different initial configurations for each system were randomly generated without significant overlaps to insure representative conformational sampling. For BD simulations of the repulsive model without HI, 30 μs simulations were performed with a time step of 0.5 ps. For simulations with HI, 12 μs simulations were performed with a time step of 2 ps. Trajectories for the first 5 μs of simulations were discarded for analysis. In calculating lubrication forces, particle pairs having $s < 4$ were evaluated every time step, where $s = 2r_{ij}/(a_i + a_j)$. On the other hand, since far-field HI are insensitive to small configuration changes, $\mathbf{M}^{\infty}$ was computed every 500 steps.

In simulations that include HI, the short-time self diffusion coefficient for translational motion of particle type i, D_1^s, is defined by [11]

$$D_i^s = \frac{1}{3N_i} \sum_{\alpha \in i}^{N_i} tr\left(\mathbf{D}^{\alpha\alpha}\right). \quad (19)$$

Here, N_i is the number of type *i* particles in the system and $\mathbf{D}^{aa}$ is the 3 × 3 matrix of the self part of the diffusion tensor. The long-time self diffusion coefficient is defined as

$$D^{\mathrm{L}} = \lim_{r \to \infty} \frac{1}{6t} \left\langle \left| \mathbf{r}(t) - \mathbf{r}(0) \right|^2 \right\rangle. \quad (20)$$

Finally, to analyze the correlations between particles in time and space, we calculate the normalized pair correlation function, C_{ij}, given by

$$C_{ij}(d_0, \tau) = \frac{\sum \left[\left(\Delta \mathbf{r}_i(\tau) \cdot \Delta \mathbf{r}_j(\tau) \right) \delta\left(d_0 - d_{ij}\right) \right]}{\sqrt{\sum \left| \Delta \mathbf{r}_i(\tau) \delta\left(d_0 - d_{ij}\right) \right|^2} \sqrt{\sum \left| \Delta \mathbf{r}_j(\tau) \delta\left(d_0 - d_{ij}\right) \right|^2}}, \quad (21)$$

where d_0 is a specified the surface distance between particles *i* and *j*, and τ is the time interval. $\delta(d_0 - d_{ij})$ is the Dirac delta function. d_{ij} is the surface distance between particles *i* and *j* at time *t* given by

$$d_{ij} = \left| \mathbf{r}_i(t) - \mathbf{r}_j(t) \right| - a_i - a_j. \quad (22)$$

The summation in Eq. 21 is over all time points *t* and all independent simulations.

3. Results

3.1. Effect of Excluded Volume and Hydrodynamic Interactions on Diffusion

We performed BD simulations with and without HI to evaluate the effects of HI on diffusion in crowded environments. In addition to the RPY [25, 26] 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 [6, 11].

Fig. 2 shows D^S/D_0 and D^L/D_0 values at three different concentrations as a function of Stokes radius, where D_0 is the diffusion coefficient in infinite dilution. 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 [13].

In *in vivo* experiments, D^L/D_0 of GFP is 0.06–0.09 [4, 5] (see Fig. 2). On the other hand, D^L/D_0 value of GFP in the simulated equivalent sphere system without HI is 0.31; this is more than 3 times larger than experiment. This result indicates that although excluded

volume effects reduce macromolecular diffusion in intracellular environments, they cannot explain the factor of ~10-16 reduction observed *in vivo*. However, when HI are considered, BD simulations give D^L/D_0 of 0.14 for at 300 mg/ml, which is close to the observed experimental values (Fig. 2). Because the concentration of macromolecules in the E. coli cytoplasm is estimated to be in range of 300−340 mg/ml [3], the concentration of 300 mg/ml used in this study is a lower limit of physiological conditions. Therefore, diffusivity of GFP obtained from the simulation at higher concentrations get even closer to experiment. 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 well reproduce the experimentally observed diffusion constant of GFP *in vivo*.

3.2. Large-distance and Long Time Intermolecular Correlations

Next, 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}. 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 non-specific attractive interaction model are shown in Fig. 3. In the model without HI, where D^L ~3 times larger than in the HI model, (Fig 3 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. These results clearly show that HI give rise to quite long distance and time dynamical correlations between particles of all sizes even in the high crowded environment.

4. Discussion

The goal of this study is to evaluate the role of HI in the reduction of macromolecular diffusion in intracellular environments. To ascertain the importance of HI, we performed BD simulations that 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 almost quantitatively predicted by the inclusion of HI without any adjustable parameters or other ad hoc assumptions.

A number of other factors can also affect intracellular diffusion: 1) *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 [27]. 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. Heterogeneous

charge distributions on molecular surface, like in real biomolecules, may affect macromolecular motions. However, since electrostatic interactions are highly screened due to the short Debye length found in physiological conditions, we believe that our conclusions would not qualitatively change. 2) *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 cP [2, 22, 28]. 3) *GFP dimerization*. It is well known that GFP tends to dimerize in solutions of low (< 100 mM) ionic strength [29]. All of these physical factors will decrease macromolecular diffusivity *in vivo*. However, based on other work [27] and our results, which show that the reduction in diffusion of GFP in the simulation with HI is 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 [30-32]. HI determined by solving Stokes equations may not fully describe the molecular situation. 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 is a very good approximation [21, 24], 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 [33]. The longest crossover time of molecule in our simulation system estimated by using this formula is 1.7 μs for the ribosome. Therefore, the effect of shape and diffusion anisotropy on the analysis of long-time translational diffusion is expected to be small.

Genome-sequencing has provided a detailed “parts list” for life [34]. Recently, the proteome wide prediction of protein structure and function has also become practical [34-40]. 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. The simulation method developed in this study could be a good tool to analyze the *in vivo* thermodynamics and kinetics of macromolecular dynamics.

Acknowledgements

This work was supported in part by grant No. GM-37408 of the Division of General Medical Sciences of the National Institutes of Health.

References

1. Ellis RJ. Trends Biochem Sci. 2001; 26:597. [PubMed: 11590012]
2. Luby-Phelps K. Int Rev Cytol. 2000; 192:189. [PubMed: 10553280]

3. Zimmerman SB, Trach SO. J Mol Biol. 1991; 222:599. [PubMed: 1748995]
4. Elowitz MB, Surette MG, Wolf PE, Stock JB, Leibler S. J Bacteriol. 1999; 181:197. [PubMed: 9864330]
5. Konopka MC, Shkel IA, Cayley S, Record MT, Weisshaar JC. J Bacteriol. 2006; 188:6115. [PubMed: 16923878]
6. Brady JF, Bossis G. Annu Rev Fluid Mech. 1988; 20:111.
7. Russel, WB.; Saville, DA.; Schowalter, WR. Cambridge monographs on mechanics and applied mathematics. Cambridge University Press; Cambridge ; New York: 1989. Colloidal dispersions.; p. xviip. 5251 leaf of plates
8. Antosiewicz J, McCammon JA. Biophys J. 1995; 69:57. [PubMed: 7669910]
9. Arya G, Zhang Q, Schlick T. Biophys J. 2006; 91:133. [PubMed: 16603492]
10. Jian H, Schlick T, Vologodskii A. J Mol Biol. 1998; 284:287. [PubMed: 9813118]
11. Phillips RJ, Brady JF, Bossis G. Phys Fluids. 1988; 31:3462.
12. Durlofsky L, Brady JF, Bossis G. J Fluid Mech. 1987; 180:21.
13. Brady JF. J Fluid Mech. 1994; 272:109.
14. Brady JF, Phillips RJ, Lester JC, Bossis G. J Fluid Mech. 1988; 195:257.
15. Beenakker CWJ. J Chem Phys. 1986; 85:1581.
16. Jeffrey DJ, Onishi Y. J Fluid Mech. 1984; 139:261.
17. Cichocki B, Ekiel-Jezewska ML, Wajnryb E. J Chem Phys. 1999; 111:3265.
18. Banchio AJ, Brady JF. J Chem Phys. 2003; 118:10323.
19. Fixman M. J Chem Phys. 1978; 69:1527.
20. Ermak DL, McCammon JA. J Chem Phys. 1978; 69:1352.
21. Ando, T.; Skolnick, J. QUANTUM BIO-INFORMATICS IV. Accardi, L.; Freudenberg, W.; Ohya, M., editors. Vol. 28. World Scientific Publishing Co. Pte. Ltd.; Singapore: 2011. p. 413-426.
22. Ridgway D, et al. Biophysical Journal. 2008; 94:3748. [PubMed: 18234819]
23. Sundararaj S, et al. Nucleic Acids Res. 2004; 32:D293. [PubMed: 14681416]
24. Ando T, Skolnick J. P Natl Acad Sci USA. 2010; 107:18457.
25. Rotne J, Prager S. J Chem Phys. 1969; 50:4831.
26. Yamakawa H. J Chem Phys. 1970; 53:436.
27. McGuffee SR, Elcock AH. PLoS Comput Biol. 2010; 6:e1000694. [PubMed: 20221255]
28. Verkman AS. Trends Biochem Sci. 2002; 27:27. [PubMed: 11796221]
29. Yang F, Moss LG, Phillips GN. Nat Biotechnol. 1996; 14:1246. [PubMed: 9631087]
30. Benz M, Chen NH, Jay G, Israelachvili JI. Ann Biomed Eng. 2005; 33:39. [PubMed: 15709704]
31. Bhushan B, Israelachvili JN, Landman U. Nature. 1995; 374:607.
32. Maali A, Bhushan B. J Phys-Condens Mat. 2008; 20:315201.
33. Schluttig J, Korn CB, Schwarz US. Phys Rev E. 2010; 81:030902.
34. Skolnick J, Fetrow JS. Trends Biotechnol. 2000; 18:34. [PubMed: 10631780]
35. Arakaki AK, Huang Y, Skolnick J. BMC Bioinformatics. 2009; 10:107. [PubMed: 19361344]
36. Baker D, Sali A. Science. 2001; 294:93. [PubMed: 11588250]
37. Brylinski M, Skolnick J. Proteins. 2010; 78:118. [PubMed: 19731377]
38. Gao M, Skolnick J. Nucleic Acids Res. 2008; 36:3978. [PubMed: 18515839]
39. Pandit SB, et al. Bioinformatics. 2010; 26:687. [PubMed: 20080513]
40. Zhang Y, Skolnick J. Proc Natl Acad Sci U S A. 2004; 101:7594. [PubMed: 15126668]

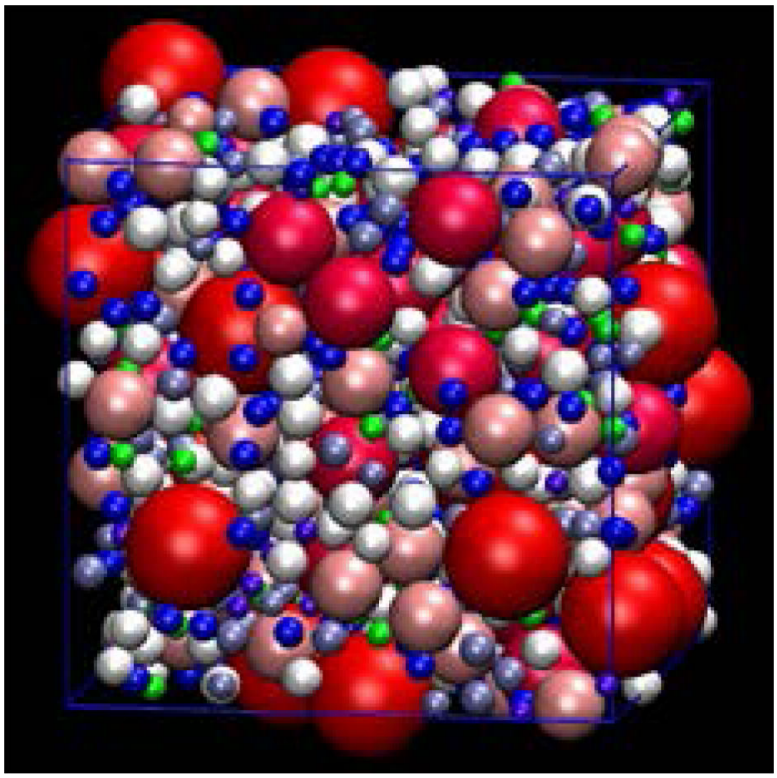

Figure. 1.
Simulation system in 100 nm × 100 nm × 100 nm box. Macromolecules are represented in different colors. Radii of molecules correspond to their Stokes radii.

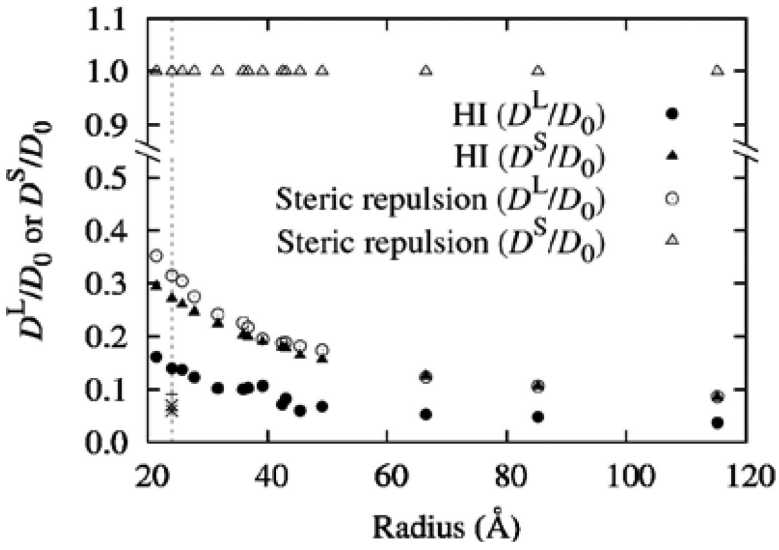

Figure 2.

Reduction in diffusivity as a function of Stokes radius. Plus, cross, and asterisk are diffusivity of GFP measured *in vivo* of DH5α [4], BL21(DE3) [5], and K-12 [5] *E. coli*, respectively.

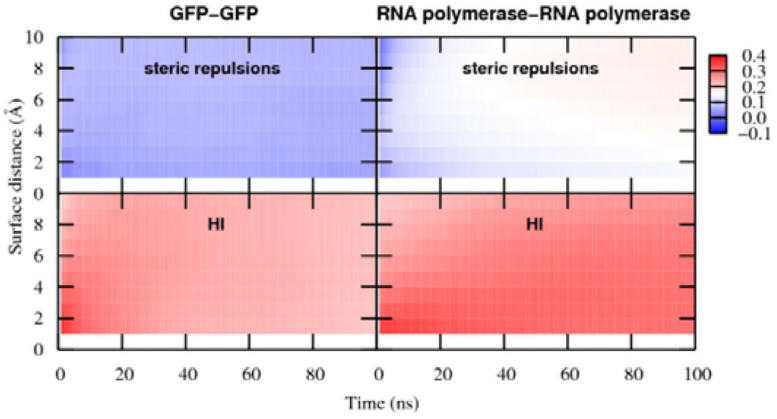

Figure 3.

Normalized pair correlation function, C_{ij}, averaged over GFP-GFP (left) and RNA polymerase-RNA polymerase (right) pairs. The Stokes radii of GFP and RNA polymerase are 24.0 and 66.5 Å, respectively.

Table 1

Simulated macromolecules, their physical properties, and number in 100 nm × 100 nm × 100 nm simulation box.

Name	PDB ID	Molecular weight (kDa)	Molecular weight range (kDa)*	Radius of gyration (Å)	Stokes radius (Å)†	D_0 (Å²/ns)‡	Number of molecules in the box
Ribonuclease HI	1JL1	17.5	0-20	14.9	21.4	11.4	70
GFP	1W7S	26.9	20-40	16.9	24.0	10.2	113
Malonyl CoA-acyl carrier protein transacylase	1MLA	32.4	20-40	18.5	25.7	9.52	94
Triosephosphate isomerase	1TRE	54.0	40-60	24.5	31.7	7.73	126
Fructose 1-6 bisphosphate aldolase	1DOS	78.2	60-80	28.5	36.8	6.67	91
Enolase 6-phosphogluconate dehydrogenase	1E9I	91.3	80-100	26.5	35.9	6.82	92
	2ZYA	106.6	100-120	29.4	39.2	6.26	31
Phosphoenolpyruvate-protein phosphotransferase	2HWG	129.3	120-140	31.8	42.5	5.77	34
Glyceraldehyde-3-P dehydrogenase	1S7C	145.0	140-160	31.6	43.1	5.69	42
Cystathionine gamma-synthase	1CS1	167.5	160-180	33.2	45.4	5.40	6
Phosphoglycerate dehydrogenase	1YBA	182.1	180-200	40.1	49.2	4.99	12
RNA polymerase	1IW7	432.9	200+	50.3	66.5	3.69	76
GroEL/ES	1AON	877.6	200+	68.4	85.2	2.88	37
70S Ribosome	3I1Q, 3I1R	2,155.2	Ribosomes	84.3	115.2	2.13	29
Initial tRNA	3CW5	24.8	tRNAs	22.7	27.8	8.83	299

*This molecular weight range corresponds to the class used in Ref. [22].

†Stokes radii were calculated by $6\pi\eta a = k_B T/D_0$, where η is the viscosity of water, D_0 is the translational diffusion coefficient in dilute solution estimated by the rigid-particle theory [21, 24], and a is the Stokes radius.

‡The diffusion constants are at 298 K.