THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Atomic Detail Brownian Dynamics Simulations of Concentrated Protein Solutions with a Mean Field Treatment of Hydrodynamic Interactions

Paolo Mereghetti*[,†,‡] and Rebecca C. Wade*[,†,§]

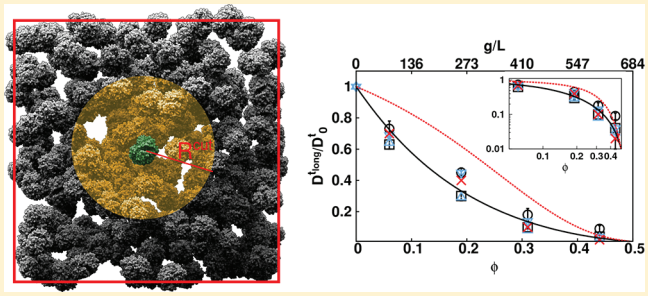
[†]Heidelberg Institute for Theoretical Studies (HITS) gGmbH, Schloß-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany

[‡]Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

[§]Center for Molecular Biology (ZMBH), University of Heidelberg, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany

Supporting Information

ABSTRACT: High macromolecular concentrations are a distinguishing feature of living organisms. Understanding how the high concentration of solutes affects the dynamic properties of biological macromolecules is fundamental for the comprehension of biological processes in living systems. In this paper, we describe the implementation of mean field models of translational and rotational hydrodynamic interactions into an atomically detailed many-protein Brownian dynamics simulation method. Concentrated solutions (30−40% volume fraction) of myoglobin, hemoglobin A, and sickle cell hemoglobin S were simulated, and static structure factors, oligomer formation, and translational and rotational self-diffusion coefficients were computed. Good agreement of computed properties with available experimental data was obtained. The results show the importance of both solvent mediated interactions and weak protein−protein interactions for accurately describing the dynamics and the association properties of concentrated protein solutions. Specifically, they show a qualitative difference in the translational and rotational dynamics of the systems studied. Although the translational diffusion coefficient is controlled by macromolecular shape and hydrodynamic interactions, the rotational diffusion coefficient is affected by macromolecular shape, direct intermolecular interactions, and both translational and rotational hydrodynamic interactions.



INTRODUCTION

Among several types of macromolecular motion found in living systems, free diffusion represents the simplest and most ubiquitous way by which molecules move. This motion is essential for processes like signal transmission and the transport of matter.[1] Several factors can affect free macromolecular diffusion in the cell or in the extracellular matrix. Among these, crowding due to the presence of many other mobile macromolecular solutes and confinement effects due to more static steric barriers such as membranes or cytoskeletal structures are the two main factors. Here, we restrict our attention to the effects of macromolecular crowding due to mobile molecules whose size is similar to the size of the protein of interest. Several studies have highlighted how macromolecular crowding affects the behavior of proteins in terms of stability, association, and dynamics; for reviews, see refs 2−4. For protein diffusion and association, macromolecular crowding leads to a combination of effects, some of which oppose each other and make net effect difficult to predict.

The excluded volume effect, i.e., the mutual impenetrability of all solute molecules, has been considered as the leading characteristic of macromolecular crowding.[4,5] Recently, increasing attention has been paid to the effects of macromolecular crowding on protein−protein interactions (PPI) and hydrodynamic interactions (HI). The role of PPI in macromolecular crowding has been investigated computationally and experimentally by several groups.[6−11] Even weak protein−protein interactions (K_d > 1 μM), which lead to transient association, have been found to be crucial for the stability and dynamics of proteins.[6−8] PPI strongly depend on the details of the individual proteins, such as the net charge, the charge distribution, the presence of hydrophobic patches, and hydrogen bonding capabilities.

A more universal effect, which is less dependent on the chemical composition of the macromolecule, is due to hydrodynamic interactions. HI have been studied for many years in colloidal solutions and polymer physics,[12,13] but only

Special Issue: Macromolecular Systems Understood through Multiscale and Enhanced Sampling Techniques

Received: December 28, 2011
Revised: May 15, 2012
Published: May 17, 2012

   dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

recently has attention been given to the influence of HI on protein motion in biological cells or in cell-like environments.[14−20] HI arise due to the forces exerted on solute particles from the velocity fields generated by other solute particles moving through a fluid.[21] Computationally, this effect can be treated by modeling solvent atoms explicitly as is done in classcal molecular dynamics (MD) simulations. However, as MD simulation algorithms scale as $O(n^2)$, where n is the total number of atoms in the simulations (solute plus solvent), the associated computational cost becomes prohibitive as the system grows. For example, for a protein of the size of the hemoglobin (∼60 kD), the number of solvent atoms typically used would be around 10 000. Then, the number of atoms in a simulated hemoglobin solution, which requires hundreds of proteins in a periodic box, is of the order of several million. Since solution properties are generally determined on the microsecond to millisecond time scale, atomically detailed simulations become challenging. To overcome this problem, the solvent can be taken into account implicitly via the diffusion tensor, a $3N \times 3N$ matrix (where N is the number of solute particles) which accounts for the solvent-mediated correlation among the macromolecules. Implicit solvent algorithms, such as Stokesian, Langevin, or Brownian dynamics (BD) methods, require Cholesky factorization of the diffusion tensor, an algorithm which scales as $O(N^3)$, to be done every time step. For this reason, the calculation of the correlated random displacements due to HI is still a bottleneck for implicit solvent many-particle simulations.[22] A possible solution is to extensively use CPU-GPU parallel algorithms as, for example, is done in the BD simulation package BD_BOX.[23] Another way to tackle this problem is to introduce approximations which reduce the computational cost while keeping the accuracy as high as possible. Several methods have been developed in this direction. One approach is to use Fixman's Chebyshev polynomials approximation, which scales as $O(N^{2.5})$.[24] Alternatively, in a recent publication,[22] Geyer and Winter described an approximation based on a first order expansion of the hydrodynamic multibody interactions which scales as $O(N^2)$. This method however cannot be applied to highly concentrated systems where HI have a dominant role because of the close contact between particles.

A third approach, which is employed in this work, is to perform BD simulations which include an average hydrodynamic effect computed by using a mean field approach. This method, which was originally described by Heyes,[25] was subsequently used by Urbina−Villalba et al.[26] to study flocculation rates in concentrated dispersions and by Sun and Weinstein[27] for the investigation of crowding effects on diffusion and association in dense hard-sphere solutions. The use of this mean field approach results in runtimes that scale as $O(N^2)$ with only a small overhead compared to BD simulations without HI. For example, for a solution of 250 hemoglobin tetramers in a periodic box, the simulation times on 16 CPU cores with and without HI are 0.21 and 0.25 μs/day, respectively. The drawback of this method is that only the average effect is captured and nonequilibrium dynamic properties are missing. However, it has the main advantage that it can be easily and efficiently implemented in implicit solvent simulations to accurately describe equilibrium hydrodynamic properties even for highly dense systems.

We recently developed an atomically detailed many-protein Brownian dynamics simulation method which we applied to dilute protein solutions.[28] The method showed good agreement with experimental data even though HI were neglected. Moving to a more dense cell-like environment, the effects of HI cannot be neglected any more. In this paper, we describe the implementation of rotational and translational HI, computed with a mean field model, into our BD simulation software. The method is applied to investigate the behavior of oxy-myoglobin (Mb), oxy- and deoxy-hemoglobin A (HbA and Deoxy-HbA), and deoxy-hemoglobin S (HbS) solutions at different concentrations, from dilute to highly concentrated solutions (30−40% volume fraction). These systems, which are well studied from the theoretical and experimental point of view, represent particularly well suited cases for assessing the accuracy of the method. Moreover, understanding the properties of concentrated solutions of these proteins is of significant biological and biomedical relevance.

METHODS

Interaction Energies and Forces. The forces were computed as finite-difference derivatives of the pairwise free energies of interaction between proteins. For each pair of proteins, the interaction free energy, ΔG^{1-2}, was defined as

$$\begin{aligned}\Delta G^{1-2} = {} & \frac{1}{2}\sum_{i_2}\Phi_{el_1}(\mathbf{r}_{i_2})\cdot q_{i_2} + \frac{1}{2}\sum_{j_1}\Phi_{el_2}(\mathbf{r}_{j_1})\cdot q_{j_1} && \text{[electrostatic interaction]}\\ & + \sum_{i_2}\Phi_{edesolv_1}(\mathbf{r}_{i_2})\cdot q_{i_2}^2 + \sum_{j_1}\Phi_{edesolv_2}(\mathbf{r}_{j_1})\cdot q_{j_1}^2 && \text{[electrostatic desolvation]}\\ & + \sum_{m_2}\Phi_{npdesolv_1}(\mathbf{r}_{m_2})\cdot SASA_{m_2} + \sum_{n_1}\Phi_{npdesolv_2}(\mathbf{r}_{n_1})\cdot SASA_{n_1} && \text{[non-polar desolvation]}\\ & + \sum_{m_2}E_{softcore_1}(\mathbf{r}_{m_2}) + \sum_{n_1}E_{softcore_2}(\mathbf{r}_{n_1}) && \text{[soft-core repulsion]}\end{aligned} \tag{1}$$

A detailed description and parametrization of eq 1 can be found in refs 28 and 29. In eq 1, Φ are interaction potentials, q are effective charges,[30] SASA are solvent accessible surface areas, E_{softcore} are soft-core repulsion energies, and $\mathbf{r}$ are atomic coordinates. For computational efficiency, all interaction potentials, Φ, were mapped onto grids centered on the simulated proteins.

Brownian Dynamics Simulations. The positions and orientations of the particles were propagated using the Ermak-McCammon[31] algorithm. Translational motion is described by

$$\mathbf{r}_i(t_1) = \mathbf{r}_i(t_0) + \sum_j \frac{\partial D_{ij}^t}{\partial \mathbf{r}_j(t_0)}\Delta t + \sum_j \frac{D_{ij}^t}{kT}\mathbf{F}_j(t_0)\Delta t + \mathbf{R}_i \tag{2}$$

where $\mathbf{r}_i$ is the position of the center of geometry of protein i, $\Delta t = (t_1 - t_0)$, $\mathbf{F}_j(t_0)$ is the force acting on particle j at time t_0, and $\mathbf{R}_i$ is a random displacement arising from the collision of the protein i with the solvent molecules and is defined by a Gaussian distribution with mean $\langle \mathbf{R}_i \rangle = 0$ and covariance $\langle \mathbf{R}_i \mathbf{R}_j \rangle = 2D_{ij}^t \Delta t$. D_{ij}^t is the translational diffusion tensor.

An analogous equation is used to propagate the rotation of the proteins

$$\mathbf{w}_i(t_1) = \mathbf{w}_i(t_0) + \sum_j \frac{\partial D_{ij}^r}{\partial \mathbf{w}_j(t_0)}\Delta t + \sum_j \frac{D_{ij}^r}{kT}\mathbf{T}_j(t_0)\Delta t + \mathbf{W}_i \tag{3}$$

where $\mathbf{w}_i$ is the rotation of protein i, $\mathbf{T}_j(t_0)$ is the torque acting on protein j at time t_0, $\mathbf{W}_i$ is a stochastic rotation of protein i with mean $\langle \mathbf{W}_i \rangle = 0$ and covariance $\langle \mathbf{W}_i \mathbf{W}_j \rangle = 2D_{ij}^r \Delta t$, and D_{ij}^r is the rotational diffusion tensor.

BD simulations were carried out using 250 proteins that were initially randomly positioned (avoiding overlaps) in a

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

rectangular box with periodic boundary conditions. By using 250 proteins, we guarantee that each protein is surrounded by at least 2 proteins in each direction. The dimensions of the box were varied according to the concentration of the protein solution.

Each system was subjected to 6 μs of simulation at 300 K. Equilibration was assessed by monitoring the convergence of the radial distribution function and the stabilization of the energies. In all cases, 1 μs was sufficient to obtain an equilibrated system according to these criteria and the remaining 5 μs were used for the analysis. For HbA, Deoxy-HbA, and HbS at 0.25 volume fraction, the simulations were extended for an additional 5 μs to have better statistics for the oligomer analysis. Simulations done to compute the short-time rotational diffusion coefficient only (i.e., simulations with direct interactions plus translational and rotational hydrodynamic interactions) were performed for 2 microseconds only because of the fast convergence of the short-time rotational diffusion coefficient.

The integration time step was 0.5 ps. The positions and orientations of the proteins were recorded along with energy values every 1.0 ns.

BD simulations were performed with SDAMM,[28] a parallelized program based on the SDA software[32] capable of handling many proteins treated as rigid bodies in atomic detail.

For further details on SDAMM, see ref 28.

Hydrodynamic Interaction. Translational (tHI) and rotational (rHI) hydrodynamic interactions, were implemented in SDAMM using a mean field method.[25−27] In this method, isotropic short-time translational and rotational diffusion coefficients ($D^{t_{short}}$, $D^{r_{short}}$) for a protein of a given type are assigned and rescaled based on the local volume fraction (ϕ_i) occupied by proteins. The rescaled value is then used in eqs 2 and 3 for updating the particle positions and orientations . Since we consider each protein to have an isotropic diffusion coefficient, the diffusion tensor is diagonal and eqs 2 and 3 simplify to

$$\mathbf{r}_i(t_1) = \mathbf{r}_i(t_0) + \frac{D^{t_{short}}(\phi_i)}{kT}\mathbf{F}_i(t_0)\Delta t + \mathbf{R}_i \quad (4)$$

$$\mathbf{w}_i(t_1) = \mathbf{w}_i(t_0) + \frac{D^{r_{short}}(\phi_i)}{kT}\mathbf{T}_i(t_0)\Delta t + \mathbf{W}_i \quad (5)$$

For a diagonal diffusion tensor, the divergence term simplifies to $(\partial D_i(\phi)/\partial r_i)\Delta t$. During a time step Δt of 0.5 ps, each particle is assumed to move in a constant environment and thus have a constant self-diffusion coefficient. Since in our simulations the term $(\partial D_i(\phi)/\partial r_i)\Delta t$ is within the interval $[10^{-6},10^{-7}]$ Å and the average displacement per time step is $\sim 10^{-2}$ Å, we considered the divergence term negligible.

We define the local volume, V_i, as the volume of the sphere of radius R^{cut} centered on protein i. The local volume fraction ϕ_i for the protein i is obtained by dividing the sum of the volumes of the proteins within R^{cut} by the local volume V_i.[26] The volume of a protein, v, is computed by approximating the proteins as spheres having a radius equal to the hydrodynamic radius (σ^{stokes}) estimated using HYDROPRO.[33] The cutoff for the local volume, R^{cut}, is set to four times the hydrodynamic radius of the central protein. A protein j is totally included in the local volume when the center-to-center distance d_{ij} between the central protein i and protein j is less than $R_{cut} - \sigma_j^{stokes}$. When a protein k is only partially included within R^{cut}, that is, when $R^{cut} - \sigma_j^{stokes} < d_{ik} < R^{cut} + \sigma_k^{stokes}$, we account for that portion of protein volume derived by the sphere−sphere intersection, as illustrated in Figure 1.

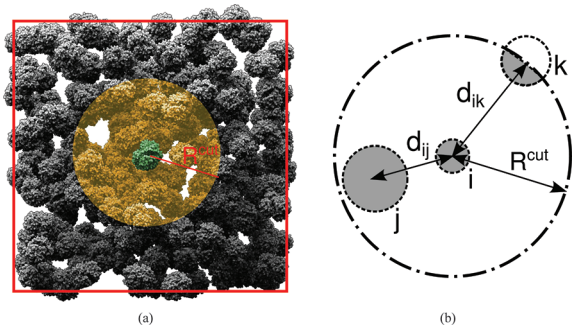


Figure 1. Schematic representation of a simulation box and the model used for the calculation of local volume fraction. (a) As an example of a simulation box we show here a snapshot (time = 5 μs) of the HbA solution at 0.11 volume fraction. Taking as representative the green protein in the center, the spherical volume of radius R_{cut} used to compute the local volume fraction is indicated in yellow. (b) For the calculation of the local volume fraction, proteins, which can be of different sizes, are described with their hydrodynamic radius.

The local volume fraction is thus defined as

$$\phi_i = \frac{v_i + \sum_j v_j + \sum_k v_k}{V_i} \quad (6)$$

where i, j, and k denote the central protein, the proteins totally included, and those partially included in R^{cut}, respectively.

The local volume fraction is then used to compute the short-time translational ($D^{t_{short}}(\phi_i)$) and rotational ($D^{r_{short}}(\phi_i)$) self-diffusion coefficients. The short-time translational diffusion coefficient is obtained using the Tokuyama model,[14,34,35] derived for a concentrated hard-sphere suspension of particles interacting with both direct and hydrodynamic interactions.

$$D^{t_{short}}(\phi_i) = D_0^t/[1 + H(\phi_i)] \quad (7)$$

where D_0^t is the infinite dilution value of the translational diffusion coefficient of protein i and the scalar function $H(\phi_i)$ is the contribution due to the "static" hydrodynamic interactions valid in the short-time regime.[35]

$$H(\phi_i) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)} \quad (8)$$

where $b = [(9/8)\phi_i]^{1/2}$ and $c = (11/16)\phi_i$.

We took into account rotational HI by rescaling the short-time rotational diffusion coefficient, using the model derived by Cichocki et al. which includes lubrication forces as well as two- and three-body expansions of the mobility functions[36]

$$D^{r_{short}}(\phi_i)/D_0^r = 1 - 0.631\phi_i - 0.726\phi_i^2 \quad (9)$$

where the isotropic infinite dilution value D_0^r is estimated using HYDROPRO.[33]

The local volume fraction dependent diffusion coefficients are recomputed at every time step of the Brownian dynamics simulations.

System and Protein Preparation. Polar hydrogen atoms were added to the protein structures according to the specified pH and ionic strength using the H++ software.[37]

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

Atomic partial charges and radii were assigned to all of the atoms from the OPLS united atom force field.[38] Heme parameters were assigned from the GROMOS96 force field.[39] Electrostatic potential grids, Φ, were computed by solving the linearized Poisson−Boltzmann equation using the program UHBD.[40] The ionic strength was set to 150 mM for HbA, Deoxy-HbA and HbS and 5 mM for Mb with a Stern layer of 1.5 Å. The relative dielectric permittivity, ε_r was set to 2 for the protein and 78.5 for the solvent. The protein van der Waals surface was used to define the dielectric boundary.

Myoglobin. The crystal structure of horse heart myoglobin (Mb) in the oxy state was taken from the Protein Data Bank: 1DWR.[41] The structure was solved by X-ray crystallography at 1.45 Å resolution.[41] The electrostatic potential grid size was set to 120 × 120 × 120 Å^3, and the electrostatic desolvation grids were set to 70 × 70 × 70 Å^3. The nonpolar desolvation grid dimension was set to 65 × 65 × 65 Å^3, and the soft-core repulsion grid was set to 60 × 60 × 60 Å^3. The grid spacing was set to 1.0 Å for all grids. Mb simulations were performed at pH 7 and an ionic strength of 5 mM. Under these conditions, the computed net charge of Mb was +2*e*. Simulations were performed at volume fractions of 0.06, 0.19, 0.31, and 0.44.

Hemoglobin A. Simulations were performed using the X-ray structure of the tetrameric oxy-hemoglobin A (HbA) solved at 1.25 Å, PDB code 2DN1.[42] The size of the electrostatic potential grid was set to 200 × 200 × 200 Å^3. The electrostatic desolvation grid was set to 120 × 120 × 120 Å^3. The nonpolar desolvation and soft-core repulsion grids were both set to 100 × 100 × 100 Å^3. As for Mb, the grid spacing was set to 1.0 Å for all grids. HbA simulations were performed at pH 7 and an ionic strength of 150 mM. The computed net charge of the HbA tetramer in these conditions was +2*e*. Simulations were performed at volume fractions of 0.04, 0.11, 0.16, and 0.25. HbA was always considered to diffuse as a tetramer. In considering HbA's oligomers, we always refer to oligomeric forms of the tetramer. For comparison one simulation of the deoxygenated form of hemoglobin A (Deoxy-HbA) was performed using the crystal structure solved at 1.74 Å, PDB code 2HHB,[43] and a volume fraction of 0.25.

Hemoglobin S. The crystal structure of human hemoglobin S in the deoxy state (HbS) was taken from the Protein Data Bank: 2HBS.[44] The structure was solved by X-ray crystallography at 2.05 Å resolution.[44] Interaction potential grids were set as for hemoglobin A. HbS simulations was performed at pH 7 and an ionic strength of 150 mM. Under these conditions, the computed net charge of the HbS tetramer was +4*e*. One simulation at a volume fraction of 0.25 was performed.

■ ANALYSIS

Small Angle Scattering Structure Factors. Structure factors, $S(q)$, were computed by Fourier transformation of the radial distribution function, $g(r)$,[45] as follows

$$S(q) = 1 + 4\pi\rho \int_0^\infty h(r)\frac{\sin(rq)}{rq} r^2\, dr$$

where ρ is the number density, r is the center-to-center distance, q is the magnitude of the scattering vector given by $q = 4\pi\lambda^{-1}\sin(\theta/2)$ (where θ is the total scattering angle), and $h(r)$ is the total correlation function which is given by $h(r) = g(r) - 1$.

Translational Self-Diffusion Coefficient. The time dependent time-ensemble averaged mean squared translational displacement ($\langle \mathrm{msd}^t(\tau)\rangle_{te}$) was computed using

$$\langle \mathrm{msd}^t(\tau)\rangle_{te} = \frac{1}{N'}\sum_{i=1}^{N'}\frac{1}{t^{end}-\tau+1}\sum_{t=0}^{t^{end}-\tau}[\mathbf{r}_i(t+\tau)-\mathbf{r}_i(t)]^2 \quad (10)$$

where N' is the number of proteins of the same type out of the total number N of proteins in the simulation and τ is the time interval.

A straight line fit was performed on a closed interval $[\tau_1,\tau_2]$, and $D^{t_{long}}$, the long time diffusion coefficient, was computed from the slope of this straight line using the Einstein relation

$$D^{t_{long}} = \frac{\langle \mathrm{msd}^t(\tau)\rangle_{te}}{6\delta t} = \frac{\langle \mathrm{msd}^t(\tau_2)\rangle_{te} - \langle \mathrm{msd}^t(\tau_1)\rangle_{te}}{6(\tau_2-\tau_1)} \quad (11)$$

Five different $D^{t_{long}}$ values were obtained for five consecutive nonoverlapping intervals of equal size ($t^{end}/5$) (corresponding to 1 μs in our simulations), which span all of the postequilibration simulation time. Since no sub- or superlinear behavior was detected for $\tau > 1$ μs, an average value and a standard deviation were computed by averaging the diffusion coefficients obtained in the last four intervals, i.e., $\tau = 1-5$ μs.

For comparison, two theoretical models which describe the concentration dependence of the normalized long-time translational self-diffusion coefficient were also used. One model, which is based on scaled particle theory (SPT), was derived by Han and Herzfeld[46] and expresses $D^{t_{long}}(\phi)/D_0^t$ as

$$\frac{D^{t_{long}}(\phi)}{D_0^t} = \exp\left\{-\frac{2}{3}\left(3\frac{\phi}{1-\phi} + \frac{9}{2}\frac{\phi^2}{(1-\phi)^2} + \frac{9}{4}\frac{\phi^3}{(1-\phi)^3}\right)\right\} \quad (12)$$

A more sophisticated model, which includes HI as well as direct interactions, was derived by Tokuyama[14,34] using the same theoretical framework as for the derivation of the short-time translational diffusion described above. According to this model

$$\frac{D^{t_{long}}(\phi)}{D_0^t} = \frac{\frac{D^{t_{short}}(\phi)}{D_0^t}}{1+\kappa\frac{D^{t_{short}}(\phi)}{D_0^t}\left(\frac{\phi}{\phi_c}\right)\left(1-\frac{\phi}{\phi_c}\right)^{-2}} \quad (13)$$

where the short-time self-diffusion coefficient, $D^{t_{short}}(\phi)$, is defined by eq 7 in the section "Hydrodynamic Interaction", $\phi_c = (4/3)^3/(7\ln 3 - 8\ln 2 + 2)$, and κ was set to 2 according to ref 14.

Rotational Diffusion Coefficient. The rotational diffusion coefficient was obtained by first computing the time-ensemble averaged autocorrelation function of the protein orientation vectors $\mathbf{e}(t)$ and fitting it with a single exponential of the form $\langle \mathbf{e}(t)\mathbf{e}(t+\tau)\rangle \propto Ae^{-t/\tau_{rel}}$. In all cases the fitting gave a correlation coefficient greater than 0.9999. The rotational diffusion coefficient was then obtained from the relaxation time τ_{rel} as $D^r = 1/2\tau_{rel}$. To obtain an average value and a standard deviation, the postequilibration trajectory was split into five consecutive time intervals, a D^r was computed for each interval and then averaged. The value of $\langle \mathbf{e}(t)\mathbf{e}(t+\tau)\rangle$ decays to zero

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

within the first few hundred nanoseconds, so the computed D^r is considered to be the short-time rotational diffusion coefficient $D^{r_{short}}$.

Alternatively, as described in ref 47, the time-ensemble averaged rotational mean squared displacement, $\langle msd^r(\tau)\rangle_{te}$, was obtained, and an average value and a standard deviation were computed using the same procedure as for the translational diffusion coefficient and the relation $D^r = \langle msd^r(\tau)\rangle_{te}/4\delta t$. As for the translational mean square displacement, in all cases, we observed a linear time dependence of the rotational msd for $\tau > 1\mu s$. The D^r estimated from the rotational msd is assumed to be representative of the long-time rotational diffusion $D^{r_{long}}$.

We compared the short-time rotational diffusion coefficient obtained from the simulations with the analytical model derived by Cichocki et al.[36] (eq 9).

Oligomer Analysis. The average fractions of oligomeric species (dimer, trimer, etc.) were computed by recording the occurrence of the oligomeric states at each step of the simulation and then averaging over the total number of steps. An oligomer is defined as a group of two or more proteins which are in contact with each other independently of the oligomer lifetime. A contact is defined following the "atomic contact criterion" for the definition of encounter complexes described in ref.[48] Namely, an encounter complex is formed when at least N^{ind} independent contacts between two proteins occur. A contact is established when the centers of two atoms (one from each protein) are closer than a given cutoff, d_c. The independence of the contacts is ensured by considering only atoms within a protein that are further from each other than a distance, d_{min}. In the analysis, we considered C_α atoms and we set $N^{ind} = 2$, $d_c = 8.5$ Å, and $d_{min} = 6.0$ Å. The parameters were set to be quite stringent to avoid picking up false contacts only determined by close protein packing in the high volume fraction simulations.

In addition to describing the tendency of each residue to be involved in a contact with another protein, the number of times a C_α atom i was found within $d_c = 8.5$ Å of a C_α atom of another protein was counted (n^i_{cont}).

To measure the stability of a given oligomeric species, we monitored all cases where an oligomer was destroyed by the loss or addition of one or more monomeric units. The autocorrelation function of the lifetime of the oligomers was then fitted to a double-exponential function. The faster component was due to the rapid recrossing of the 4.0 Å cutoff distance used in the contact criterion. The slow component, with period τ_{off}, was considered to be representative of a true oligomer dissociation event and could be used to estimate the lifetime of the oligomer.

RESULTS AND DISCUSSION

In the following subsections, we first discuss dynamic properties, i.e., translational and rotational diffusion coefficients, and then we analyze thermodynamic properties including oligomer formation and structure factors. Finally, a concluding discussion is given.

Excluded Volume and tHI Affect the Translational Diffusion Coefficient. The long-time translational self-diffusion coefficients ($D^{t_{long}}$) of Mb and HbA were computed from BD simulations at several concentrations and compared to experimental and theoretical data. The long-time regime is generally assumed to occur for time scales longer than the structural relaxation time τ_0^I. The structural relaxation times were estimated for different concentrations according to the relation $\tau_0^I = r_{cc}^2/D_0^t$,[49] where r_{cc} is the average interparticle center-to-center distance and D_0^t is the infinite dilution translational self-diffusion coefficient. Following ref,[16] the average interparticle distance was estimated using $r_{cc} = (\phi_{csp}/\phi)^{1/3}2a$ where $\phi_{csp} = \pi/3\sqrt{2} = 0.74$ is the closest spherical packing and a is the hydrodynamic radius of the protein. The values of τ_0^I are given in Table 1.

Table 1. Structural Relaxation Times, τ_0^I (μs), for Mb, HbA, Deoxy-HbA, and HbS at Different Concentrations

ϕ	Mb	HbA	deoxy-HbA	HbS
0.1	0.6	2.4	2.3	2.3
0.2	0.4	1.5	1.4	1.4
0.3	0.3	1.1	1.1	1.1

For both Mb and HbA, BD simulations in the absence of translational hydrodynamic interactions (tHI) result in a slight overestimation of the translational self-diffusion coefficients and the relative deviation from the experimental value is larger at higher concentrations. See Figure 2.

The computed values with only soft-core repulsion and without tHI are below those of the Han and Herzfeld model[46] which describes the dependence of the $D^{t_{long}}(\phi_i)$ on the volume fraction based on scaled particle theory[50] and therefore only includes excluded volume interactions. This means that in our simulations, in addition to the excluded volume effect, shape effects play a role in the reduction of $D^{t_{long}}(\phi_i)$ compared to the Han and Herzfeld model. It is worth noting that the HbA diffusion coefficients for the soft-core simulations without translational hydrodynamic interactions, are slightly closer to the Han model than in the myoglobin case. This indicates a less pronounced shape effect for HbA compared to Mb. The role of the shape will become clearer later for rotational diffusion. For the Brownian dynamics simulations which include tHI (with or without the full interaction terms), we observe a quantitative agreement with the experimental values and the analytical Tokuyama model[14,34] as indicated by Figure 2. Until now, we have considered translational hydrodynamic interactions only. When the rotational hydrodynamic interactions are switched on, a negligible effect is observed on the translational diffusion coefficient. Hence, data obtained from simulations which include tHI and rHI were not shown on the plot.

In the simulations, besides the soft-core repulsion, three other terms contribute to the interactions forces: electrostatic interaction, electrostatic desolvation, and nonpolar desolvation terms. Since the net charge of both Mb and HbA is low, $+2e$, the long-range self-repulsion will be weak, especially for the 150 mM ionic strength simulations of HbA. Hence, we do not expect significant electrostatic effects on the dynamics of our proteins. Indeed, the average electrostatic interaction energy per protein, in the range of concentrations simulated, is of the order of 1 kT for HbA and Mb. Proteins can easily come close to each other where short-range interactions become important. The short-range interactions are described by an electrostatic desolvation term, which represents the free energy penalty associated with the displacement of water from polar regions and a nonpolar desolvation term, which describes the favorable removal of water from the surface. Although the electrostatic desolvation term shows only a modest unfavorable contribution to the total interaction energy, +2 kT on average in both HbA and Mb, the nonpolar term has a dominant role.

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

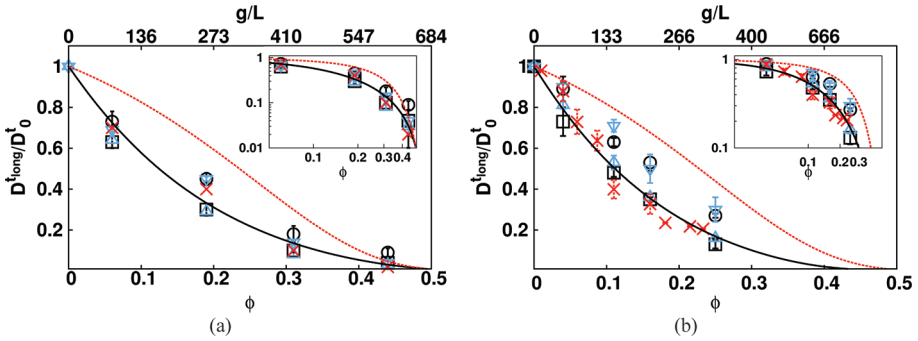


Figure 2. Normalized long-time translational self-diffusion coefficient of Mb (a) and HbA (b). Experimental values (red crosses) and BD simulations with (black squares) and without (black circles) translational hydrodynamic interactions. BD simulations with soft-core interactions only with (blue triangle pointing up) and without (blue triangle pointing down) translational hydrodynamic interactions. Han and Herzfeld[46] and Tokuyama[14] theoretical models are shown by dashed red and black continuous lines, respectively. The inset is a log−log plot of the same data.

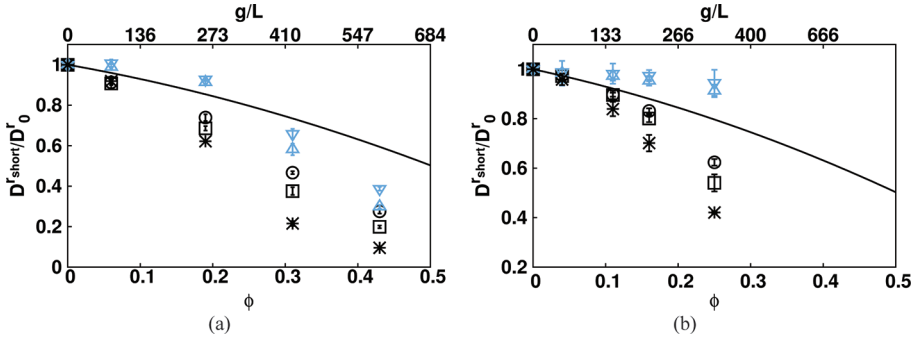


Figure 3. Normalized rotational diffusion coefficient of Mb (a) and HbA (b) estimated from the correlation function of protein orientations. BD simulations with (black squares) and without (black circles) translational hydrodynamic interactions. BD simulations with soft-core only with (blue triangle pointing up) and without (blue triangle pointing down) translational hydrodynamic interactions. BD simulations with direct interactions plus translational and rotational HI are shown with black stars. The black line shows the theoretical prediction using the model derived by Cichocki et al.[36]

The average nonpolar desolvation energy per protein is [−4,−8] and [−6,−15] kT for HbA and Mb, respectively. The minimum and maximum values in brackets correspond to the values for the lowest and the highest concentration simulations. This short-range attraction does not however affect the long-time translational self-diffusion coefficient, as the simulations with and without the full interaction terms give very similar results, see Figure 2. This implies that, at long-times, the translational dynamics are controlled essentially by excluded volume interactions and translational hydrodynamic effects.[15,16] A minor contribution comes also from the shape while direct interactions effects are negligible. These results are in line with the observation of Ando and Skolnick that excluded volume and tHI are the major factors in determining the large reduction of translational diffusion coefficient of macromolecules in vivo.[15] It should be pointed out that just from the agreement between the experimental long-time translational diffusion coefficient and the Tokuyama model one cannot directly conclude that the long-time diffusion coefficient is controlled only by excluded volume and HI. Indeed, the Tokuyama model does not consider explicitly electrostatic interactions, desolvation terms, and shape. It is also possible that the experimental data agree with the Tokuyama model due to compensatory effects. By means of BD simulations, we could directly test the effects of direct interactions, tHI, and rHI on the long-time diffusion coefficient.

Rotational Diffusion Coefficient Depends on Shape and Direct Interactions. A different situation arises when the behavior of the rotational diffusion coefficient is considered. In Figure 3, the rotational diffusion coefficient obtained from BD simulations is compared with an analytical model derived by Cichocki[36] (see eq 9).

Considering the simulations done with only the soft-core term (with or without HI), we observe that the rotational diffusion coefficient of the HbA is almost unaffected by the protein concentration whereas the Mb rotational diffusion coefficient decreases quite significantly with concentration. Since only shape is accounted in these simulations, these differences can only result from the more spherical shape of the HbA tetramer compared to Mb. The difference in shape between Mb and HbA can be quantified using the globularity index which is the ratio between the smallest and the largest eigenvalues of the covariance matrix of atomic coordinates.[51] A value of 1 indicates a perfect sphere while a value of 0 indicates a one or two-dimensional object. The globularity index is equal to 0.6 for HbA and 0.3 for Mb. The flatter shape of Mb leads to a stronger effect on the rotational diffusion coefficient. Hydrodynamic interactions applied to the translational diffusion coefficient have an effect on rotational diffusion in the simulations; that is, BD simulations with tHI show a stronger reduction of the rotational diffusion coefficient with increasing concentration. Including rotational hydrodynamic interactions leads to an even greater reduction of the rotational diffusion coefficient, as shown by the black stars in Figure 3.

Switching on direct interactions, contrary to what was observed for the translational diffusion coefficient, a strong

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

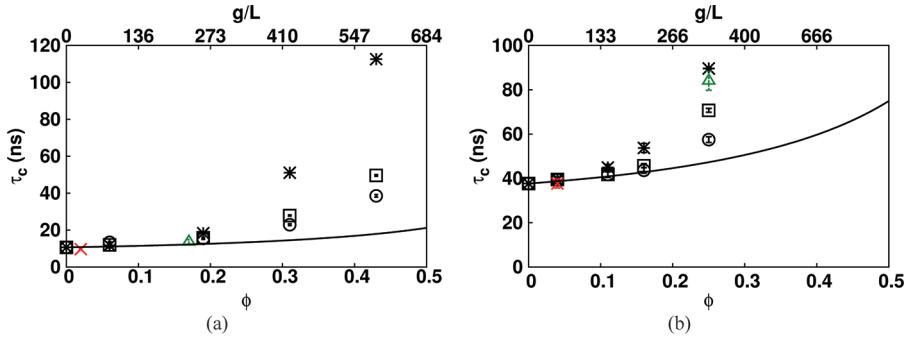


Figure 4. (a) Mb and (b) HbA rotational correlation times. In the cell (green triangle) and in solution (red cross) NMR data.[52] BD simulations with direct interactions and translational HI (black squares), with translational and rotational HI (black stars), and without HI (black circles). The black line shows the theoretical prediction using the model derived by Cichocki et al.[36]

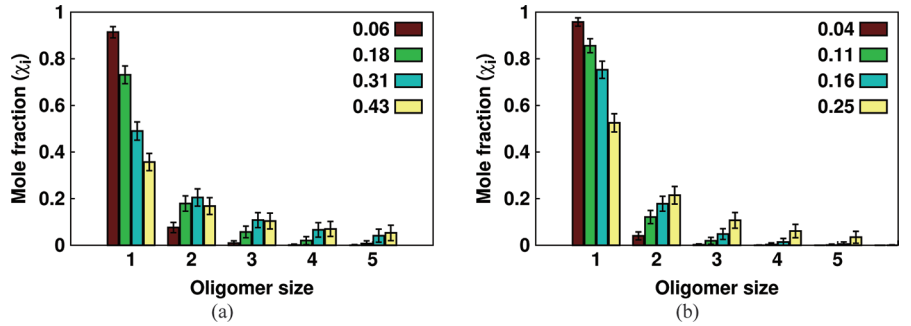


Figure 5. Average mole fractions of Mb (a) and HbA (b) oligomers at different volume fractions given by different bars (see figure legend). Oligomeric state whose molar fraction is less than 0.02 for any concentration does not appear in the plot.

effect on the rotational diffusion coefficient is observed; see Figure 3. Both the simulations with and without HI show a lower rotational diffusion coefficient compared to the analytical model. This finding establishes a primary role of weak short-range attractions in the rotational dynamics of the proteins which, in turn, influences association rates.

The rotational diffusion coefficients given in Figure 3 were estimated from the autocorrelation function of the orientational vectors (see Methods). Very similar results, and hence the same conclusions, can be obtained by determining the rotational diffusion coefficient from the rotational mean square displacement (see Methods) as can be seen in Figure S1 in the Supporting Information. The equivalence between the two ways of computing D^r implies that the short-time rotational diffusion coefficient converges to its long-time value within the first few hundreds of nanoseconds.

From the rotational diffusion coefficient, the rotational correlation time can be computed as $\tau_c = 1/6D^r$. The rotational correlation times of HbA have been obtained experimentally by Wang et al.[52] In their work, they measured τ_c for hemoglobin in red blood cells (RBCs) and in dilute solution (<30 g/L) using NMR. In the same work, Wang et al. measured the rotational diffusion coefficient of myoglobin in perfused rat myocardium and in dilute solution using NMR.[52] In Figure 4, we compare our computed values for HbA and Mb with the experimental data, assuming that the concentration in RBCs and in myocardial cells is about 300[53] and 200 g/L, respectively. The dilute solution experimental values were compared to our 10 g/L simulation. The value of 200 g/L in myocardial cells was estimated by choosing the concentration of myoglobin solutions at which the translational diffusion coefficient (obtained from simulations) corresponds to the translational diffusion coefficient measured in myocardial cells.[54] That is, the translational diffusion diffusion coefficient of 4.24×10^{-7} cm^2 s^{-1} estimated by Lin and colleagues by NMR at 25°C in rat myocardium[54,55] corresponds to the value obtained in our 0.19 volume fraction (200 g/L) Mb solution simulation.

In the Mb case, the experimental rotational correlation time in myocardial cells has been estimated to be 1.4 times larger than that in the dilute solution.[52] Computationally, we found an almost quantitative agreement: the τ_c from the BD simulation with direct interactions plus tHI and rHI at 200 g/L is 1.8 times larger than the value for the 10 g/L simulation; see Figure 4. Moreover, this value agrees with the increase (1.35 times) in the rotational correlation time of apo-myoglobin monomers from dilute to 225 g/L solutions of RNase A measured by fluorescence correlation spectroscopy (FCS).[56]

In RBCs, the experimental rotational correlation time for hemoglobin is 2.2 ± 0.2 times higher than in dilute solution.[52] Deriving τ_c from our simulations which include direct interactions plus tHI and rHI, we obtain a value of τ_c at 333 g/L, 2.31 times larger than the dilute solution value; see Figure 4. These data show that direct interactions and translational and rotational hydrodynamic interactions are required for correctly describing the rotational diffusion coefficients.

Formation of Dynamics Oligomers in Solution. The strong concentration effect on the rotational diffusion in the simulations with the full interaction terms can be explained by the formation of short lifetime oligomers determined by weak short-range attraction. Indeed, it has been shown in other cases that transient protein−protein associations lead to a slowing down of protein tumbling due to a change in the effective molecular size.[6,8,57] To assess this hypothesis, we computed the molar fraction and the lifetime of oligomers in the simulations

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

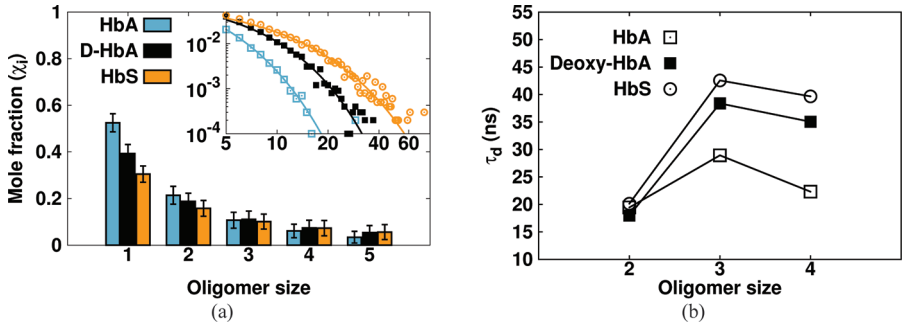


Figure 6. (a) Average mole fractions of HbA, Deoxy-HbA, and HbS oligomers up to pentamers at 0.25 volume fraction. Mole fractions of oligomers larger than or equal to pentamers are shown in the inset on a log−log scale. Continuous lines shows stretched exponential fits. (b) Oligomer dissociation times for HbA, Deoxy-HbA, and HbS.

for different concentrations. First, we recall the distinction between transient, dynamic and permanent oligomers defined by Liu et al.[57] If the lifetime of an oligomer is much smaller than the time that it needs to diffuse through one protein diameter, τ_0^D, the oligomer is defined as transient, and the dynamics will be essentially determined by the monomers. If the lifetime is long enough for it to diffuse more than one particle diameter, the oligomer is defined as dynamic and the system's behavior is expected to be dominated by oligomers for the short-time behavior, although its long time dynamic behavior will be determined by monomer motions. For a permanent oligomer, all monomeric units inside it move together. Its lifetime is expected to be longer than most of the experimental time needed for measuring macroscopic quantities.[57] The diffusional relaxation time τ_0^D, similarly to the previously defined structural relaxation time, τ^I, can be estimated as $\tau_0^D = a^2/D_0^t$, where a is the hydrodynamic radius of a monomer and D_0^t is its diffusion coefficient. For Mb and HbA, we assumed hydrodynamic radii of 21.0 Å and 31.9 Å respectively (estimated using HYDROPRO[33]), which give a τ_0^D of 43 ns for Mb and 160 ns for HbA.

Figure 5 shows the molar fractions of the different oligomeric species (monomers to pentamers) found in the myoglobin and hemoglobin simulations with full interaction terms and HI.

The molar fraction of the oligomers increases monotonically with increasing concentration. This concentration dependence suggests the presence of an activation enthalpy barrier, e.g., small electrostatic repulsion, which has to be overcome for oligomerization.

Considering myoglobin solutions, we observe that above 0.31 volume fraction a plateau value is reached for all oligomeric species. This trend was not seen for HbA for which the maximum concentration studied is 0.25. However, the 0.19 ϕ Mb simulation and the 0.16 ϕ simulation of HbA show a comparable oligomeric pattern, suggesting that the tendency to form oligomers in HbA and Mb simulations is similar. The formation of oligomers in Mb simulations is consistent with the experimental finding of the formation of stable dimers of Apomyoglobin observed by Zorrilla and colleagues in concentrated RNase A solutions.[56]

Considering the HbA simulations, the presence of oligomers in the simulations agrees with the fact that a small amount of oligomers had to be assumed to fit the scattering curves of concentrated HbA solutions.[19,58,59] However, contrary to Stadler et al.,[60] we do not detect fractal aggregates or very large clusters in our simulations. A reason for this could be the absence of internal structural flexibility which, as already mention above, has been suggested to be the cause of the formation of large aggregates due to the exposure of hydrophobic patches.[60,61] In Figure 6a we compare the molar fraction of oligomers found in the 0.25 volume fraction HbA, Deoxy-HbA, and Deoxy-HbS simulations. Higher fraction of oligomeric species is observed in the Deoxy-HbA and HbS simulations than in the HbA.

Taking HbA as the reference, we observe an increase in the oligomerization of Deoxy-HbA which can be attributed to the differences in the electrostatic potential due to the quaternary structure rearrangement of the deoxy form (see Figure S2 in the Supporting Information). Indeed, the rearrangement of the quaternary structure leads to a decrease of the dipole moment and an increase in the quadrupole moment in the deoxy form compared to the oxy form (see Table S1 in the Supporting Information). The more quadrupolar electrostatic potential of the deoxy state facilitates the formation of higher order oligomers, whereas the more dipolar shape of the electrostatic potential in the oxy form favors the formation of dimers. For the oligomeric fractions of HbS, we observe a further increase of the fraction of oligomers (see Figure 6a). Hence, the Glu 6 to Val mutation constitutes an effect additional to the oxy-deoxy structural rearrangement, which facilitates the formation of oligomers. The increased contact propensity of Val 6 compared to Glu 6 is clearly visible from the number of residue contacts, n_{cont}^i, shown in Figure S3 in the Supporting Information. Large oligomers (tens of monomers) are more abundant for HbS than Deoxy-HbA or HbA (see inset in Figure 6a). Small fractions of very large transient oligomers (60−70 monomers) can be identified the simulations of HbS. These very large oligomers have a branched and disordered shape, examples are shown in Figure S4 in the Supporting Information. The trend observed in the oligomer analysis is confirmed by comparison of the radial distribution functions, $g(r)$, of HbA, Deoxy-HbA, and HbS (Figure 7). The Deoxy forms show a more pronounced peak at a distance corresponding to the hydrodynamic diameter of the tetramer compared to the oxy form. The $g(r)$ of HbS has an additional peak at a closer distance (53.6 Å), attributable to the hydrophobic patch introduced by the valine mutation. These results are consistent with a two-step mechanism of fiber formation,[62−64] where the formation of metastable dense dynamic clusters is a precursor step to the formation of ordered nuclei of the HbS polymer. Moreover, they agree with the higher homogeneous nucleation rate observed for HbS compared to HbA.[62]

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523−8533

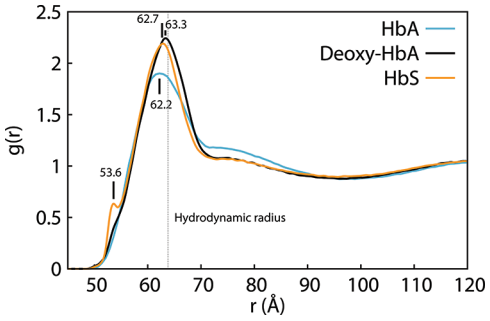


Figure 7. Radial distribution functions for HbA, Deoxy-HbA and HbS computed from simulations at 0.25 volume fraction.

To investigate the nature of the oligomeric species found in the simulations, we examined their lifetimes; see Figure 8.

Considering the highest concentrations for the Mb and HbA solutions, the diffusional relaxation time τ_0^D is about 4 and 8 times greater than the average lifetime of the oligomers of the Mb and HbA, respectively.

This implies that we are in the transient oligomer range. That is, the long-time dynamics is dominated by the monomers, which explains why the long-time translational diffusion coefficients agree with the diffusion coefficients of soft-sphere suspensions[14] as shown in Figure 2. On the other hand, the formation of these transient clusters affects the rotational dynamics, as previously shown (see Figure 3).

Concerning the comparison of HbA, Deoxy-HbA, and HbS, as shown in Figure 6, the stabilities of the oligomers follow the same trend as the oligomeric fractions; that is, they are higher in the deoxy forms and the highest stabilities are found for the HbS oligomers.

The oligomers found in these simulations are not present simply due to local density fluctuations. This was checked by performing an oligomer analysis on the simulations which only include the soft-core repulsion term, using the same contact parameters applied to the analysis of the full interaction term simulations. From this test, we found out that at all concentrations soft-core simulations result in 99% monomers, indicating that the oligomers found in the full interaction term simulations depend on the formation of "true" contacts.

We should also mention that the oligomers formed in the simulations have to be considered as encounter complexes. The absence of internal flexibility and a specific treatment of hydrogen bonds in the simulations hinder the formation of fully bound oligomers.

Accuracy of the Protein–Protein Interaction Model. To assess the accuracy of the force field used in BD simulations, the static structure factors, $S(q)$, of HbA and Mb were computed and compared to the available experimental curves.

The $S(q)$ of HbA in red blood cells (RBCs) was obtained experimentally by Krüger.[19,59] In the RBCs, the ionic strength is nearly 200 mM[65] and the HbA concentration is estimated to be around 300 g/L.[66] We therefore compared the experimental $S(q)$ with the computed structure factor for the simulation at the highest concentration (333 g/L). For the myoglobin case, the experimental $S(q)$ values are available for different volume fractions in salt-free conditions.[20]

As shown in Figure 9 the computed and experimental $S(q)$ curves agree quite well confirming the validity of the description of interprotein interactions used in our BD simulations.

We do not expect an exact quantitative agreement between the experimental and the BD derived $S(q)$. The experimental structure factors, indeed, are extremely sensitive to environmental conditions and their computation relies on an interparticle interaction potential and the solution of the Ornstein–Zernike equation.[67] Several parameters which affect the shape of $S(q)$ have to be tuned to achieve the best fit of the experimental intensities.[67] On the computational side, the simulation conditions do not exactly correspond to the experimental ones and the approximation used in the force-field may also represent a potential source of discrepancy. Nevertheless, a reasonably good agreement is found between the experimental and computed $S(q)$ profiles and particularly, the concentration effect on the shape of $S(q)$ in the myoglobin case is properly reproduced.

Methodological Considerations. From the computational point of view, models for the simulations of dense many protein systems, have been developed by Elcock[10] and Skolnick[15] with the aim of reproducing the cell cytoplasm of *E. coli*. In the first model,[10] based on BD simulations, proteins were treated in atomic detail with particular attention directed to modeling protein–protein interactions which were calibrated to reproduce in vivo GFP translational diffusion coefficients. Elcock and colleagues were able to reproduce in vivo experimental thermodynamic stabilities of few proteins. However in their model, hydrodynamic interactions were neglected.

Ando and Skolnick used Stokesian dynamics[68] with proteins simplified to effective soft-spheres while hydrodynamic interactions were carefully treated by considering far-field HI, as well as many-body and near-field HI. These authors conclude that excluded volume effects and HI are the two

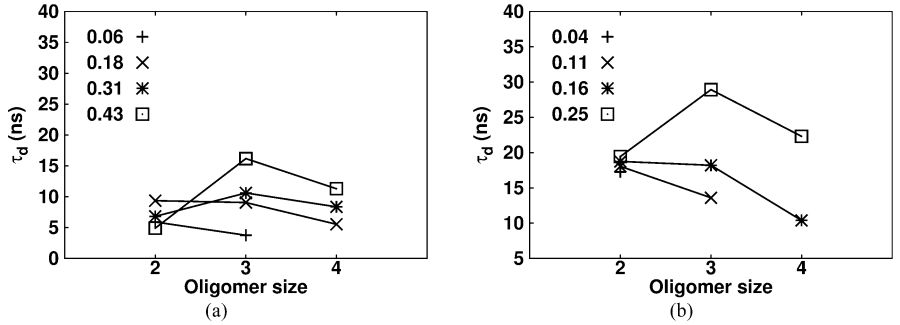


Figure 8. Mb (a) and HbA (b) oligomers dissociation times at different volume fractions. If a value is not plotted, e.g., the tetramers at the lowest concentrations, it means that they form in too low a concentration to obtain reliable dissociation times.

dx.doi.org/10.1021/jp212532h | J. Phys. Chem. B 2012, 116, 8523–8533

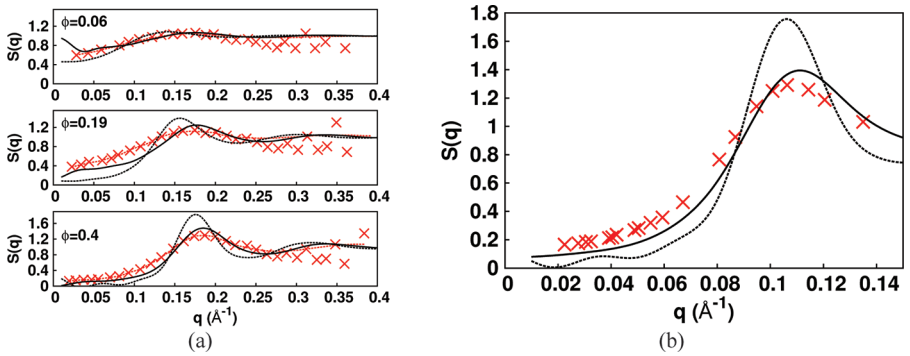


Figure 9. (a) Mb structure factors, $S(q)$, at different concentrations (given by ϕ). Experimental values[20] (red crosses), computed curves from BD simulations with full interaction terms (black continuous line) and with the soft-core term only (black dashed line), and theoretical fit to experimental data[20] (red dashed line). (b) HbA experimental $S(q)$[19,59] (red crosses) and computed curve from BD simulations with full interactions (black continuous line) and with the soft-core term only (black dashed line).

main factors that account for the large reduction of translational diffusion coefficients of macromolecules in vivo. Moreover, they found that the proteins can be modeled as single spheres and protein−protein interactions play a minor role in the diffusional dynamics.

In our approach, in addition to an accurate description of protein−protein interactions,[28,30,69] we add a model for hydrodynamic interactions which accurately reproduces equilibrium properties of concentrated protein solutions. The two principal factors which are still missing in this method are the internal protein flexibility and the solvent correlation effects. Despite these limitations, we found that the shape of the molecule, protein−protein interactions, and HI all contribute, at different time scales, to the dynamic of proteins in solution. These observations, for homogeneous solutions, can be extended to more complex systems like the cell cytoplasm where the highly heterogeneous environment makes it even more important to model protein−protein interactions and solvent effects well.

CONCLUSIONS

In this paper, we have described the implementation of mean-field HI in atomically detailed many-protein BD simulations and the application of this method to the study of concentrated Mb, HbA, and HbS solutions.

Two main conclusions can be drawn. First, our results clearly indicate that excluded volume, shape, protein−protein interactions, and hydrodynamic interactions should all be taken into account to properly describe the diffusion and association of proteins in a dense protein solution. In particular, we found a qualitative difference in the concentration dependence of the translational and rotational diffusion coefficients. That is, while the translational diffusion coefficient is essentially affected by excluded volume and translational hydrodynamic interactions, the rotational diffusion coefficient is influenced by the shape of the molecule, direct interactions and both rotational and translational hydrodynamic interactions.

Second, from comparison of the HbA, Deoxy-HbA, and HbS simulations, we found that oligomers form in the simulations in a concentration dependent manner with higher propensity to aggregate in the HbS cases compared to the oxygenated HbA and deoxygenated HbA. These results are consistent with a two-step mechanism proposed for the homogeneous nucleation of the HbS polymers which involves the formation of metastable dense clusters as precursors in the fiber formation process.

ASSOCIATED CONTENT

Supporting Information

Long-time rotational diffusion coefficient of Mb and HbA (Figure S1); dipolar and quadrupolar moments of HbA, Deoxy-HbA, and HbS (Table S1); electrostatic isopotential surfaces for HbA, Deoxy-HbA, and HbS (Figure S2); residue contact propensity of HbA, Deoxy-HbA, and HbS (Figure S3); examples of large oligomers identified in HbS 0.25 volume fraction simulations (Figure S4). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: paolo.mereghetti@h-its.org; rebecca.wade@h-its.org.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Klaus Tschira Foundation and the Center for Modelling and Simulation in the Biosciences (BIOMS) in Heidelberg and a grant for supercomputing time at the Environmental Molecular Science Laboratory (Grant No. 30994).

REFERENCES

(1) Mika, J. T.; Poolman, B. *Curr. Opin. Struct. Biol.* **2011**, *22*, 117−126.
(2) Elcock, A. H. *Curr. Opin. Struct. Biol.* **2010**, *20*, 196−206.
(3) Gershenson, A.; Gierasch, L. M. *Curr. Opin. Struct. Biol.* **2010**, *21*, 32−41.
(4) Minton, A. P. *Curr. Biol.* **2002**, *16*, R269−R271.
(5) Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597−604.
(6) Wang, Q.; Zhuravleva, A.; Gierasch, L. M. *Biochemistry* **2011**, *50*, 9225−9236.
(7) Rosen, J.; Kim, Y. C.; Mittal, J. *J. Phys. Chem. B* **2011**, *115*, 2683−2689.
(8) Wang, Y.; Li, C.; Pielak, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 9392−9397.
(9) Li, X.; Moal, I. H.; Bates, P. *Proteins* **2010**, *78*, 3189−3196.
(10) McGuffee, S. R.; Elcock, A. H. *PLoS Comput. Biol.* **2010**, *6*, e1000694.
(11) Zhang, J.; Liu, X. Y. *J. Chem. Phys.* **2003**, *119*, 10972−10976.
(12) Rotne, J.; Prager, S. *J. Chem. Phys.* **1969**, *50*, 4831−4837.

(13) Yamakawa, H. *J. Chem. Phys.* **1970**, *53*, 436−443.
(14) Tokuyama, M.; Moriki, T.; Kimura, Y. *Phys. Rev. E* **2011**, *83*, 1−8.
(15) Ando, T.; Skolnick, J. *Proc. Natl. Acad. Sci.* **2010**, *107*, 18457−18462.
(16) Roosen-Runge, F.; Hennig, M.; Seydel, T.; Zhang, F.; Skoda, M. W.; Zorn, S.; Jacobs, R. M. J.; Maccarini, M.; Fouquet, P.; Schreiber, F. *Biochim. Biophys. Acta* **2010**, *1804*, 68−75.
(17) Frembgen-Kesner, T.; Elcock, A. H. *Biophys. J.* **2010**, *99*, L75−7.
(18) Frembgen-Kesner, T.; Elcock, A. H. *J. Chem. Theory Comput.* **2009**, *5*, 242−256.
(19) Doster, W.; Longeville, S. *Biophys. J.* **2007**, *93*, 1360−1368.
(20) Longeville, S.; Doster, W.; Kali, G. *J. Chem. Phys.* **2003**, *292*, 413−424.
(21) Nägele, G. Brownian Dynamics simulations. In *Computational Condensed Matter Physics*; 37th IFF Spring School 2006; Forschungszentrum Jülich Publishing, 2006; Vol. *32*.
(22) Geyer, T.; Winter, U. *J. Chem. Phys.* **2009**, *130*, 114905.
(23) Dlugosz, M.; Zielinski, P.; Trylska, J. *J. Comput. Chem.* **2011**, *32*, 2734−2744.
(24) Fixman, M. *Macromolecules* **1986**, 1204−1207.
(25) Heyes, D. *J. Non-Newtonian Fluid Mech.* **1997**, *68*, 101−124.
(26) Urbina-Villalba, G.; Garcia-Sucre, M.; Toro-Mendoza, J. *Phys. Rev. E* **2003**, *68*, 061408.
(27) Sun, J.; Weinstein, H. *J. Chem. Phys.* **2007**, *127*, 155105.
(28) Mereghetti, P.; Gabdoulline, R. R.; Wade, R. C. *Biophys. J.* **2010**, *99*, 3782−3791.
(29) Gabdoulline, R. R.; Wade, R. C. *J. Am. Chem. Soc.* **2009**, *131*, 9230−9238.
(30) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868−3878.
(31) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352−1360.
(32) Gabdoulline, R. R.; Wade, R. C. *Biophys. J.* **1997**, *72*, 1917−1929.
(33) de la Torre, J. G.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719−730.
(34) Tokuyama, M.; Oppenheim, I. *Phys. Rev. E* **1994**, *50*, R16−R18.
(35) Tokuyama, M.; Oppenheim, I. *Physica A* **1995**, *216*, 85−119.
(36) Cichocki, B.; Ekiel-Jezewska, M. L.; Wajnryb, E. *J. Chem. Phys.* **1999**, *111*, 3265−3273.
(37) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368−W371.
(38) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(39) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, 1996.
(40) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Andrew, I.; Antosiewicz, J.; Gilsong, M. K.; Bagheri, B.; Scott, L.; Ridgway McCammon, J. A. *Comput. Phys. Commun.* **1995**, *91*, 57−95.
(41) Sherwood, C.; Mauk, A.; Brayer, G. D. *J. Mol. Biol.* **1987**, *193*, 227.
(42) Park, S.-Y.; Yokoyama, T.; Shibayama, N.; Shiro, Y.; Tame, J. R. H. *J. Mol. Biol.* **2006**, *360*, 690−701.
(43) Fermi, G.; Perutz, M.; Shaanan, B.; Fourme, R. *J. Mol. Biol.* **1984**, *175*, 159−174.
(44) Harrington, D. J.; Adachi, K.; Royer, W. E. *J. Mol. Biol.* **1997**, *272*, 398−407.
(45) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: New York, 1989.
(46) Han, J.; Herzfeld, J. *Biophys. J.* **1993**, *65*, 1155−1161.
(47) Mazza, M.; Giovambattista, N.; Stanley, H.; Starr, F. *Phys. Rev. E* **2007**, *76*, 031203.
(48) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329−341.
(49) Koenderink, G. H.; Zhang, H.; Aarts, D. G. a. L.; Lettinga, M. P.; Philipse, A. P.; Nägele, G. *Farady Discuss.* **2003**, *123*, 335−354.
(50) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. *J. Chem. Phys.* **1959**, *31*, 369−380.
(51) Narsinghani, T.; Chaturvedi, S. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 461−468.
(52) Wang, D.; Kreutzer, U.; Chung, Y.; Jue, T. *Biophys. J.* **1997**, *73*, 2764−2770.
(53) Zuckerman, K. S. Approach to the anemias. In *Cecil Medicine*, 23rd ed.; Saunders Elsevier: Philadelphia, 2007.
(54) Lin, P.-C.; Kreutzer, U.; Jue, T. *J. Physiol.* **2007**, *578*, 595−603.
(55) Gros, G.; Wittenberg, B. a.; Jue, T. *J. Exp. Biol.* **2010**, *213*, 2713−2725.
(56) Zorrilla, S.; Rivas, G.; Acuna, A. U.; Lillo, M. P. *Protein Sci.* **2004**, *13*, 2960−2969.
(57) Liu, Y.; Porcar, L.; Chen, J.; Chen, W.-R.; Falus, P.; Faraone, A.; Fratini, E.; Hong, K.; Baglioni, P. *J. Phys. Chem. B* **2011**, *115*, 7238−7247.
(58) Krueger, S.; Chen, S. H.; Hofrichter, J.; Nossal, R. *Biophys. J.* **1990**, *58*, 745−757.
(59) Krueger, S.; Nossal, R. *Biophys. J.* **1988**, *53*, 97−105.
(60) Stadler, A. M.; Schweins, R.; Zaccai, G.; Lindner, P. *J. Phys. Chem. Lett.* **2010**, *1*, 1805−1808.
(61) Stadler, A. M.; Digel, I.; Artmann, G. M.; Embs, J. P.; Zaccai, G.; Büldt, G. *Biophys. J.* **2008**, *95*, 5449−5461.
(62) Ivanova, M.; Jasuja, R.; Kwong, S.; Briehl, R. W.; Ferrone, F. a. *Biophys. J.* **2000**, *79*, 1016−22.
(63) Galkin, O.; Pan, W.; Filobelo, L.; Hirsch, R. E.; Nagel, R. L.; Vekilov, P. G. *Biophys. J.* **2007**, *93*, 902−913.
(64) Pan, W.; Galkin, O.; Filobelo, L.; Nagel, R. L.; Vekilov, P. G. *Biophys. J.* **2007**, *92*, 267−277.
(65) Mouat, M. F.; Manchester, K. L. *Compar. Haematol. Int.* **1998**, *8*, 58−60.
(66) Dong, H.; Qin, S.; Zhou, H.-X. *PLoS Comput. Biol.* **2010**, *6*, e1000833.
(67) Fukasawa, T.; Sato, T. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3187−3196.
(68) Brady, J. F.; Bossis, G. *Ann. Rev. Fluid. Mech.* **1988**, *20*, 111−157.
(69) Mereghetti, P.; Wade, R. C. *BMC Biophys.* **2011**, *4*, 2−11.