JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Protein Simulations in Fluids: Coupling the OPEP Coarse-Grained Force Field with Hydrodynamics

Fabio Sterpone,*[,†] Philippe Derreumaux,[†,‡] and Simone Melchionna[§]

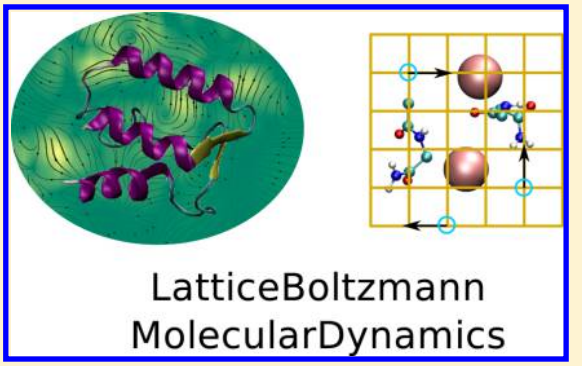
[†]Laboratoire de Biochimie Théorique, IBPC, CNRS UPR9080,Univ. Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005 Paris, France

[‡]IUF, Institut Universitaire de France, Boulevard Saint Michel, 75005 Paris, France

[§]CNR-IPCF, Consiglio Nazionale delle Ricerche, Rome, Italy

Supporting Information

ABSTRACT: A novel simulation framework that integrates the OPEP coarse-grained (CG) model for proteins with the Lattice Boltzmann (LB) methodology to account for the fluid solvent at mesoscale level is presented. OPEP is a very efficient, water-free and electrostatic-free force field that reproduces at quasi-atomistic detail processes like peptide folding, structural rearrangements, and aggregation dynamics. The LB method is based on the kinetic description of the solvent in order to solve the fluid mechanics under a wide range of conditions, with the further advantage of being highly scalable on parallel architectures. The capabilities of the approach are presented, and it is shown that the strategy is effective in exploring the role of hydrodynamics on protein relaxation and peptide aggregation. The end result is a strategy for modeling systems of thousands of proteins, such as in the case of dense protein suspensions. The future perspectives of the multiscale approach are also discussed.



1. INTRODUCTION

The study of proteins in cell-like environments is a challenging frontier for modern biophysics and molecular biology.[1,2] In fact, protein mobility,[3,4] stability,[5] and function in the extremely heterogeneous and crowded space as the interior of cells differ from those in dilute solution conditions. Just to say, in a eukariotic cell, depending on the organelle or subcellular location, 5 to 40% of the total volume is occupied by proteins, nucleic acids, and other large molecules. Because of this molecular crowding, excluded volume is expected to affect protein relaxation, mobility, and stability.[2] For example, protein motion restrained in a crowded disordered medium is predicted to be slower and subdiffusive.[4] Retardation has been confirmed by several investigations based on single particle tracking,[6] fluorescence correlation spectroscopy,[7] fluorescence-recovery-after-photobleaching,[8] and quasi elastic neutron scattering.[9] At an occupied volume fraction approaching 30−40%, protein diffusion reduces by about 1 order of magnitude or more with respect to the dilute solution.[9−11] Moreover, excluded volume stabilizes proteins by disfavoring highly elongated entropic states.[2] However, the close packing and the action of specific or unspecific interactions, e.g., of electrostatic origin, could favor misfolding and aggregation. Experiments have reported both stabilization and destabilization depending on the nature of the protein and of the crowders.[5,12,13]

Despite the growing interest and the increasing experimental investigations, the net effect of crowding is still not completely understood. In fact, experimental studies are particularly challenging since the intracellular environment is difficult to manipulate and control.[7,13−15] Therefore, to overcome some of the experimental limitations and obtain complementary information at high spatial and temporal resolution, a computational approach is necessary.[16] Following seminal work,[17,18] important efforts are currently under way.[19−24] Unfortunately, a fully atomistic molecular resolution for describing a cell-like environment is still prohibitive and is currently limited to simulating only small systems or short time scales. Therefore, a coarse-grained (CG) molecular model removing solvent degrees of freedom is a preferred strategy but requires including the computationally expensive hydrodynamic effects in order to account for solvent mediated interactions.[21,25−27]

Ando and Skolnick[21] have clearly shown that hydrodynamic interactions (HI) are essential to describe protein mobility under crowding and therefore to reproduce the experimental behavior. They used Stokesian dynamics,[28] Brownian Dynamics[29] with near and far-field HI, applied to a model of *E. coli* cellular environment where proteins were treated as spheres of different sizes. Their model was used to show that only by including HI the mobility of a target protein (GFP) slowed down by a factor 10 as compared to the HI-free case, being then in agreement with experiments.[10] Very recently, Bandyopahyay and co-workers[24] also presented a grained

Received: November 14, 2014

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

model of the *E. coli* cytoplasm with proteins described by one or few connected beads, thus including a simplified intraprotein flexibility. The model was used to investigate protein mobility by applying Brownian Dynamics including approximated HI (mean-field). An attempt to provide a more realistic description of proteins in cells was done by McGuffee and Elcock;[19] their *E. coli* environment was made up of about 1000 proteins at atomistic resolution but moving as rigid bodies. Unfortunately for systems of such resolution, HI could not be included in the Brownian Dynamics scheme. In fact, HI come at a high computational cost ($O(N^3)$), and for methodologies based on Green's function, the nonlocal character of HI makes it necessary to use drastic approximations when the number of particles becomes too high; see discussion in ref 30.

Simulations based on CG models and including HI find wide application in several contexts. HI crucially affect the protein mobility in dilute solution,[22,26] the protein folding/unfolding kinetics[26,31,32] or the kinetics of lipid aggregation,[33] and protein association.[34] In this work, we present in detail an original multiscale method that couples the CG model for proteins OPEP[35−38] with an explicit representation of the solvent dynamics. At variance with previous schemes, the explicit and on-the-fly solution of the fluid kinetics is done via the Lattice Boltzmann (LB) methodology[39,40] and allows including the HI in a natural way.[41] One of the appealing features of the approach is its amenability to parallel computing that makes it particularly efficient in simulating systems composed by many proteins. During the Gordon Bell Prize 2013 contest, the scheme was used to simulate about 18,000 fully flexible proteins for tenths of nanoseconds on about 18,000 GPUs. This unprecedented simulation established the capability of the framework, and an overview of the results was presented in a recent review paper on the OPEP force field.[38]

In the following, we first describe the computational method, namely, the CG model for proteins OPEP and its coupling with the Lattice Boltzmann methodology. We then apply the technique to test the stability of isolated globular proteins in the hydrodynamic fluid, the effect of HI on the aggregation process of amyloid peptides, and the intrinsic multiscale potentialities of the method for treating systems such as protein suspensions.

2. METHODS

2.1. Force Field OPEP. The CG model OPEP[38,42] represents each amino acid by reserving an atomistic description to the backbone (5 centers of force for N, HN, C_ω, C, and O atoms), while one bead is used for the protein side chain. An exception is made by proline that is defined by all its heavy atoms. The OPEP force field is defined as a sum of local, nonbonded, and hydrogen bonding terms. The local terms describe bond stretching, angle bending, and proper and improper dihedral rotations using functional form as in common to all-atoms force fields. The nonbonded terms describe attractive and repulsive interactions between the center of forces. OPEP is water and electrostatic free; however, specific potentials mimic the backbone hydrogen-bond propensity along with cooperativity and the salt-bridge between ionic side chains. A complete and detailed description of the whole Hamiltonian and the main changes introduced in the latest versions of the model[36,37] is given in the Supporting Information of ref 38.

In the past, OPEP has been used in combination with several computational techniques, such as Monte Carlo,[43] molecular dynamics,[44] replica exchange molecular dynamics (REMD),[45] and simulated tempering[46] and successfully applied to study peptide folding[47] and amyloid aggregation.[48−50] At variance with other simplified force fields, OPEP is able to maintain a protein stable at long time scale without recurring to external constraints that "freeze" the secondary structures. Moreover, the force field is not biased toward the native configuration of proteins, so folding just emerges from the basic elementary forces of the model. Finally, we mention that even if, as for any other CG scheme, the energy scale of the force field is somehow arbitrary, the effective tuning of the relative weights of the potential energy terms results in protein melting temperatures ranging from 280 to 400 K, with deviation from experimental ones by about 20 K for the systems tested so far.[38] As a consequence of this quality, recently the model was successfully applied to study the different thermal stability of a pair of homologues proteins of relatively large size, ∼200 amino acids.[51] When compared to other CG models, a limitation of OPEP is that the atomistic description of the protein backbone requires a short integration time step to account for the high frequency intramolecular dynamics, a value that does not exceed 1.5−2 fs. In the simulations presented in this work, we used a time step of 1.5 fs.

2.2. Coupling OPEP to Lattice Boltzmann. *2.2.1. Lattice Boltzmann Method.* The motion of protein atoms has a direct effect on the surrounding aqueous environment. HI arise from the motion of atoms that generate a velocity field in the surrounding space. The resulting flow acts on other protein atoms and therefore produces effective, solvent-mediated interactions. Our purpose is to represent explicitly the dynamics of the solvent rather than employing an effective, analytical expression for HI. The solvent is described by considering the fluid as a continuum in a probabilistic sense.

This is the realm of kinetic theory, as envisioned by Boltzmann, where the fluid molecules are described by the single particle distribution function, $f(\mathbf{r},\mathbf{v},t)$, that encodes the probability of finding at time t a group of molecules at position $\mathbf{r}$ and carrying a velocity $\mathbf{v}$. While the Boltzmann description holds down to the molecular scale, collisional events taking place among the molecules complete the description. However, it is particularly difficult to determine the precise form of the collisions in the liquid phase, and such events can be accurately modeled only in special cases, as for the hard sphere fluid.[52]

The fluid is intrinsically characterized by the Knudsen number, $Kn = \lambda/L_H$, that regulates the deviation of the distribution from the local equilibrium state, where λ is the water mean free path ($\simeq$ 5 Å[53]), and L_H is a characteristic hydrodynamic scale, ranging from a few intermolecular distances to the protein length scale (10−100 Å), so that $0.05 < Kn < 0.5$. Such a range implies that in the liquid state water typically features small departures from local equilibrium. In the macroscopic limit, that is, after several molecular collisions and over a few molecular diameters, kinetic theory recovers the Navier−Stokes equations of fluid mechanics.

It is crucial to note that the kinetic description is richer than the macroscopic one, as it contains informations regarding the full probabilistic character of matter via a given statistical distribution, while fluid mechanics only deals with the first two moments of such distribution, namely, the solvent density field, $\rho(\mathbf{r},t)$, and the velocity field, $\mathbf{u}(\mathbf{r},t)$. These are related by the continuity equation $(\partial\rho)/(\partial t) + \nabla\cdot(\rho\mathbf{u}) = 0$, and the Navier−Stokes equation describes their macroscopic time evolution as follows:

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} = -\frac{1}{\rho}\nabla P + \nu\nabla^2\mathbf{u} + \frac{\mathbf{G}}{m} \quad (1)$$

Here, P is the pressure, accounting for the thermodynamic forces mutually exerted between fluid elements, ν is the kinematic viscosity, accounting for dissipation arising from gradients in the velocity field, and $\mathbf{G}$ is the external force exerted by the protein atoms at given locations, exerted over the fluid molecules of mass m. In the following, for the sake of simplicity we initially neglect from the treatment the presence of spontaneous microscopic fluctuations in water, and it will be explicitly introduced in section 2.2.3.

The kinetic evolution of the fluid solvent is formulated in numerical terms by the LB method. In essence, the LB method is intrinsically based on the assumption of small departures from local equilibrium, a condition that is obeyed by liquid water under standard conditions. Given a spatial discretization of the distribution, the $f(\mathbf{r},\mathbf{v},t)$ is represented over a Cartesian mesh with mesh points collectively indicated as $\{\mathbf{x}\}$ and mesh spacing Δx.

In the LB method, the velocity space is represented by a small set of discrete speeds and the dynamics evolves in timesteps of duration Δt. In essence, $f(\mathbf{r},\mathbf{v},t)$ is replaced by the simpler quantity, $f_p(\mathbf{x},t)$, where the index p indicates a discrete velocity vector. The kinetic moments of the populations $f_p(\mathbf{x},t)$ provide the local mass density $\rho(\mathbf{x},t) = \sum_p f_p(\mathbf{x},t)$ and velocity $\mathbf{u}(\mathbf{x},t) = 1/\rho\sum_p \mathbf{c}_p f_p(\mathbf{x},t)$. The stress tensor $\vec{P}(\mathbf{x},t) = 1/\rho \sum_p \mathbf{c}_p \mathbf{c}_p f_p(\mathbf{x},t)$ has diagonal components related to the fluid pressure, while the off-diagonal terms relate to the shear stress. At the thermodynamic level, the solvent pressure obeys the ideal gas equation of state, a simple, yet effective description such that only hydrodynamic interactions and internal viscous forces are accounted for.

The fluid populations then evolve via the following updating scheme

$$f_p(\mathbf{x} + \mathbf{c}_p\Delta t, t + \Delta t) = f_p(\mathbf{x}, t) - \omega\Delta t(f_p - f_p^{eq})(\mathbf{x}, t) + g_p(\mathbf{x}, t) \quad (2)$$

The right-hand side of eq 2 contains the effect of fluid−fluid molecular collisions that relax the populations toward a Maxwellian equilibrium, f_p^{eq}. This is the Bhatnagar−Gross−Krook collisional operator, local in space and time, and being proportional to the nonequilibrium component of the distribution, $(f_p - f_p^{eq})$. The Maxwellian equilibrium is expressed by a second-order expansion in the macroscopic velocity $\mathbf{u}$, $f_p^{eq} = w_p\rho[1 + \mathbf{u}\cdot\mathbf{c}_p/c_s^2 + \mathbf{uu}:(\mathbf{c}_p\mathbf{c}_p - c_s^2\,\mathbf{I})/2c_s^4]$. In addition, $c_s = 1/(3)^{1/2}$ is the speed of sound, w_p is a set of weights normalized to unity, and $\mathbf{I}$ is the unit tensor in Cartesian space. The Chapman−Enskog multiscale analysis shows that the relaxation frequency ω is related to the fluid kinematic viscosity via the expression $\nu = c_s^2(1/\omega - \Delta t/2)$.[40]

The $\{w_p\}$ form a set of normalized weights, and we follow the implementation provided by the D3Q19 lattice with $p = 0,\ldots,18$. The discrete speeds span the crystal structure of a face centered cubic lattice, so that $\sum_p w_p\mathbf{c}_p\mathbf{c}_p/c^2 = \vec{1}$, where the double arrow denotes a rank-two tensor $\vec{A}=A_{ab}$, a dyadic product $\mathbf{uv} = u_a v_b$ and $\vec{1} = \delta_{ab}$. This choice preserves isotropy with respect to dyadic products, since $1/(c^4)\sum_p w_p \sum_{bc}\overline{(c_p c_p)}_{ab}\,\overline{(c_p c_p)}_{bc}\,u_c u_d = u_a u_d$, with the overbar denoting the traceless tensor. For the D3Q19 scheme, the weights w_p are equal to 1/3 for the population corresponding to the null discrete speed, 1/18 for the ones connecting first mesh neighbors, and 1/36 for the ones connecting second mesh neighbors. Finally, the last term in eq 2 represents the coupling between the solvent and the atoms.

2.2.2. Coupling Molecules and LB. As stated above, the evolution of the fluid is perturbed by the presence of solute particles via the action of the term $g_p(\mathbf{x},t)$. An accurate expression is given by[54,55]

$$g_p(\mathbf{x}, t) = -w_p\Delta t\left[\frac{\mathbf{G}\cdot\mathbf{c}_p}{c_s^2} + \frac{(\mathbf{G}\cdot\mathbf{c}_p)(\mathbf{u}\cdot\mathbf{c}_p) - c_s^2\mathbf{G}\cdot\mathbf{u}}{c_s^4}\right]$$

The force $\mathbf{G}$ contains any external force and the exchange of momentum induced by N moving atoms. This term, for example, includes the drag force and extra contributions as the random noise encoded in the molecular dynamics. In order to have a consistent treatment of the external forces, the fluid velocity is redefined as $\mathbf{u} = 1/\rho\ \sum_p \mathbf{c}_p f_p + (\Delta t/2)\mathbf{G}$.[54]

The drag force exerted between the i-th particle and the fluid is modeled as[41,56]

$$\mathbf{F}_i^D(\mathrm{R}_i) = -\gamma(\mathbf{V}_i - \tilde{\mathbf{u}}(\mathrm{R}_i)) \quad (4)$$

where $\{\mathbf{V}_i\}$ is the atom velocities, and $\tilde{\mathbf{u}}$ indicates the fluid velocity field distributed over the region occupied by the atom. The constant γ represents the translational coupling coefficients that regulate the exchange of momentun between particles and fluid and is related to the particle diffusion coefficient at infinite dilution via the fluctuation−dissipation relationship $D = kT/m\gamma$. At the macroscopic level, the coupling constant γ accounts for Stokes law, such that a sphere of radius a moving in a quiescent fluid experiences a drag force $\mathbf{F}^D = -6\pi\eta a\mathbf{V}$, with ρ the fluid density, and dynamic viscosity $\eta = \rho\nu$. In the simulations discussed later on, the parameter γ was set to 0.1 fs^{-1}, the density $\rho = 10^3$ $(\mathrm{kg/m}^3)$, and the kinematic viscosity $\nu = 0.166$ $(\mathrm{Å}^2/\mathrm{fs})$, a value that mimics liquid water, whose kinematic viscosity varies between 0.18 and 0.08 $\mathrm{Å}^2/\mathrm{fs}$ in the temperature interval 0−30 °C. The case of different values of these parameters is explicitly stated in the text.

As we intend to treat proteins at the microscopic level, Stokes law is not generally applicable, and the finite extension of particles is modeled by the lumped form, eq 4. Accounting for the particle extension sets the relative length scale between the fluid and the particle, which is the fluid velocity distributed over the region centered on the particle position. A first choice is to take the fluid velocity at the nearest grid point to the particle position. However, such a choice does not allow altering the relative length scale between the fluid and particle and could cause small discontinuities in the drag force as the particle moves to different locations. A more flexible scheme is such that the particle velocity is distributed over multiple lattice points by a smooth version of the Dirac delta function, and the fluid velocity is obtained by interpolation of velocity from its neighboring lattice points. The smooth delta function acts as an interpolation kernel and is given by the area-weighted average function:

$$\delta_h(\mathbf{x} - \mathbf{R}_i) = d_h(x - X_i)d_h(y - Y_i)d_h(z - Z_i) \quad (5)$$

where the hat function is defined as

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$d_h(r) = \begin{cases} (1 - |r|/h) & \text{for } |r| \le h \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

At the smallest scale, h is chosen as the lattice spacing Δx with an effective range of $2\Delta x$ but a larger support given when multiples of Δx are adopted. The hat function consists of a trilinear interpolation scheme, but one can also smear according to a higher order scheme, by taking

$$d_h(r) = \begin{cases} \frac{\Delta x}{2h}\left[1 + \cos\left(\frac{\pi |r|}{h}\right)\right] & \text{for } |r| \le h \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Again, depending on the parameter h the lattice support varies. In the simulations presented in the next sections, we used this latter scheme with a value $h = 2\Delta x$ and a spacing $\Delta x = 1$ Å. The effect of several resolution and interpolation schemes is discussed in the last part of the article. It is a simple matter to verify that the chosen smearing function sums over the lattice to unity. In addition, it guarantees the invariance over continuous translations.

For a given choice of the interpolation kernel, the fluid velocity at the particle position is taken to be the weighted velocity:

$$\tilde{\mathbf{u}}(\mathbf{R}_i) = \sum_{\mathrm{x}} \mathbf{u}(\mathrm{x}) \delta_h(\mathbf{x} - \mathbf{R}_i) \quad (8)$$

Clearly, the wider is the support the heavier is the computational effort, with the resulting discrete convolutions being proportional to 8, 64, ... points for $h = \Delta x, 2\Delta x, \ldots$.

The particle dynamics is given by the following equations of motion:

$$\dot{\mathbf{R}}_i = \mathbf{V}_i$$
$$\dot{\mathbf{V}}_i = \frac{\mathbf{F}_i}{M_i} - \gamma(\mathbf{V}_i - \tilde{\mathbf{u}}_i) + \mu_i \quad (9)$$

μ_i is white noise with $\langle \mu_i(t) \rangle = 0$ and $\langle \mu_i(t)\mu_i(t') \rangle = 2\gamma k_B T/M_i \delta(t-t')\mathbf{I}$, where k_B is the Boltzmann constant and T the temperature. Equation 9 is integrated in time over the time step Δt_{MD} according to the symplectic position Verlet algorithm[57] and, if $\Delta t_{MD} = \Delta t$, the particle and LB dynamics are updated in a synchronous way. It is worth noting that by setting $\tilde{\mathbf{u}} = 0$, the standard Langevin dynamics is recovered.

In order to observe hydrodynamic behavior down to the Δx scale, it is usually considered that the fluid mean free path should not exceed Δx.[58] In our case, we lift this requirement by allowing for a subnanometric mesh spacing so that hydrodynamic behavior will spontaneously emerge over distances larger than Δx. This strategy has proven to be accurate when the LB method is used to simulate nanofluids.[59,60] At the same time, as shown in previous literature,[41,61] given the smooth BGK collisional operator, HI between pairs of particles are reproduced when Δx is comparable to the interparticle distance. This occurrence should be considered as an artifact of the model since the hydrodynamic scale should be set only for distances larger than a few (typically >2) molecular diameters. Consistently, choosing the size of Δx, as compared for instance to the atomic size, is a matter of trade-off between including finer, although somehow overhydrodynamic, versus coarser interactions and the computational effort.

In the temporal domain, the LB time step Δt can be chosen to be comparable (or synchronous) to Δt_{MD}. However, the speed of sound of the LB fluid cannot be modified, and by taking $\Delta x = 1$ Å and $\Delta t = 1$ fs, it results to be about 100-fold faster than that for real water, that is, the LB solvent is much less compressibile than what water is in reality. Such a difference would be minimized by considering a much larger LB time step but at the cost of reducing numerical robustness. For our needs, however, an artificially faster speed of sound does not spoil the accuracy of HI.

A schematic view of the coupling between LB methods and the molecular dynamics technique is given in Figure 1. The coupling has been implemented in the in-house developed code for multiphysics simulation, MUPHY.[62] A release of the coupled engine is planned for the future.

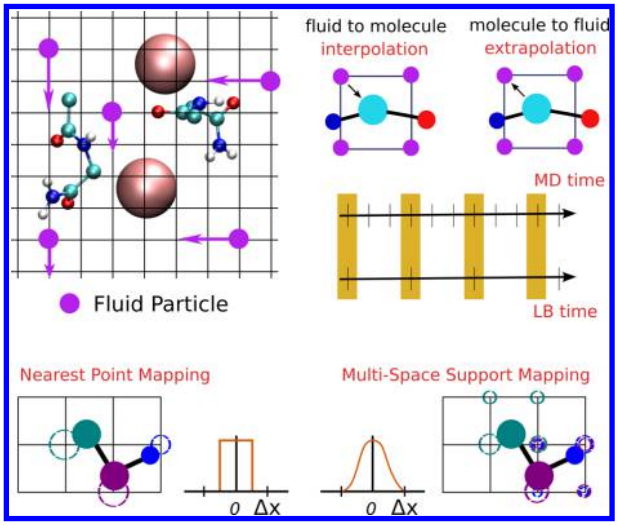


Figure 1. Schematic representation of the coupling between the molecular system and the fluid. The fluid "particles" are associated with a set of discrete probability distributions $f_p(x,t)$ where p indicates a discrete velocity along one of the lattice directions. The drag force F_i^D on the molecular particle i is obtained by interpolating the lattice velocity $\tilde{\mathbf{u}}$ at the particle position. Similarly, the drag force $\mathbf{F}_i^D$ and the random force μ_i act on the fluid, and their values are extrapolated at the lattice point. Several schemes are available to map particle positions and lattice points; see the text for details.

2.2.3. Thermal Noise in the Fluid. Finally, in order to be consistent with the microscopic treatment of the protein particles and the fluid, the effect of thermal noise needs to be taken into consideration also at the fluid mesoscopic level. The stochastic version of the LB is given by the following modification of eq 2:[61]

$$f_p(\mathbf{x} + \mathbf{c}_p\Delta t, t + \Delta t) = f_p(\mathbf{x}, t) - \omega\Delta t(f_p - f_p^{eq})(\mathbf{x}, t) + g_p(\mathbf{x}, t) + \xi_p(\mathbf{x}, t) \quad (10)$$

where the fluctuating term $\xi_p(\mathbf{x},t)$ acts at the level of the stress tensor and nonhydrodynamic modes and reads

$$\xi_p = \sqrt{\frac{\rho k_B T \omega \Delta t(2 - \omega\Delta t)}{\Delta x^3 c_S^2}} \sum_{k=4}^{18} w_p \chi_{kp} \mathcal{N}_k$$

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Summary of the Simulations Performed to Test the LBMD Methodology and Values of the Most Important Parameters Used in the Simulations[a]

system	f.f.	time [ns]	n particles	time step [fs]	L_x [Å]	Δx [Å]	ν [Å^2/fs]	γ [fs^{-1}]
1CLB	v.4	90/100	447	1.5	50	1.0	0.1666	0.1
1E0L	v.4	90/100	221	1.5	40	1.0	0.1666	0.1
1FCL	v.4	90/100	332	1.5	40	1.0	0.1666	0.1
1SHG	v.4	90/100	342	1.5	40	1.0	0.1666	0.1
2DA1	v.4	90/100	418	1.5	50	1.0	0.1666	0.1
1E0L*	v.4	30	221	1.5	40	1.0	0.1666	0.1/0.01/0.001
$A\beta_{16-22}$	v.5	110/150	864	1.5	65	1.0	0.1666	0.1
$A\beta_{16-22}$ *	v.5	50	864	1/1.5	65	1.3/1.44/1.857/2.6	0.1666	0.1

[a]The asterisk indicates tests where parameters have been modified.

Table 2. Set of Proteins Investigated Using OPEP v.4 Coupled with Hydrodynamic Interactions[a]

system	size [a.a]	rigid core	time [ns]	rmsd [Å]	n_t	Q
1CLB	75	3−14, 20−39, 47−53, 60−73	100/90	3.5/2.7	0.71/0.81	0.68/0.79
1E0L	37	7−33	100/90	3.0/2.8	0.67/0.70	0.65/0.67
1FCL	56	3−55	100/90	3.6/3.5	0.75/0.76	0.70/0.65
1SHG	57	8−11, 29−61	100/90	4.3/4.0	0.67/0.72	0.51/0.64
2DA1	70	16−63	100/90	3.2/2.7	0.74/0.79	0.58/0.75

[a]Simulations were performed at T = 300 K. The residues forming the rigid core of each protein are reported in the third column. For the simulation time and the three collective variables, rmsd, fraction of native torsion (n_t), and fraction of native contacts (Q), we report data from simulations without and with hydrodynamic coupling, left and right data, respectively, in each column. The rigid core residues are indicated according to the PDB numbering.

$\{\chi_k\}_{k=0,18}$ is a set of 18 adimensional lattice eigenvectors and is orthonormal according to the scalar product

$$\sum_{\rho=0}^{18} \omega_p \chi_{kp} \chi_{lp} = \delta_{kl} \tag{11}$$

$$\sum_{k=0}^{18} \omega_p \chi_{kp} \chi_{kp'} = \delta_{pp'} \tag{12}$$

and $\mathcal{N}_k$ is a set of 15 random numbers extracted from a Gaussian distribution of zero mean and unit variance. The eigenvectors $\{\chi_k\}$ are constructed based on the LB scheme in use (as for D3Q19); the interested reader can find the technical details in ref 61. They correspond to hydrodynamic moments: k = 0 is relative to mass density, k = 1−3 to mass current, and k = 4−9 on the (symmetric) momentum flux tensor, while the remaining k = 10−18 correspond to nonhydrodynamic modes. The kinetic eq 10 reproduces the fluctuating Navier−Stokes equation,[63,64] $\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} = -1/\rho \nabla P + \nu \nabla^2 \mathbf{u} + 1/\rho \nabla \vec{\sigma} + \mathbf{G}/m$, where $\vec{\sigma}$ is the stochastic component of the stress tensor. The stochastic force acts at the level of the smallest length scale Δx and produces consistent fluctuations up to the largest wavelength of the simulation box.[61]

2.3. Multitime and Space Resolution. The coupled lattice Boltzmann and molecular dynamics (LBMD) techniques provide an intrinsically multiphysics multiscale framework for molecular simulations. The further use of the OPEP force field specializes the technique to proteins in solution. The multiscaling can be designed at different levels, concerning both the molecular or fluid resolution. First, the molecular motion can be evolved using a multiple time step technique for integrating separately bonded and nonbonded interactions of the OPEP Hamiltonian. Moreover, the molecular representation based on the OPEP force field can be further grained by simplifying the intramolecular interactions and reducing them to an effective elastic network or to a Gō-like model. However, the lattice resolution underlying the fluid dynamics can be changed in order to reduce the cost of accounting for hydrodynamics. Finally, the coupling can be designed to be synchronous, with the fluid and the molecular motion propagated at equal times, or asynchronous, e.g. with the fluid evolved only every n molecular steps. A detailed discussion of the multiscaling strategies is reserved to the final section of this work.

3. RESULTS

In the following, we present a detailed testing of the methodology. At first, we have investigated the capabilities of our implementation and inquired the long-time stability of a protein in the hydrodynamic fluid. Then, we have explored the effect of HI on the kinetics of amyloid aggregation, a relevant biological process. Finally, we have benchmarked the multiscale features of the LBMD technique. For the sake of clarity, in Table 1 we summarize the body of simulations performed here with the relevant parameters used. Details about the simulation performances are reported in Supporting Information, Table S.1.

3.1. Stability of Model Proteins. We have considered five small proteins previously used to assess the quality of the latest versions (v.4[36] and v.5[37]) of the OPEP force field. The size of these proteins ranges from 37 to 75 amino acids; see Table 2. For each protein, we performed simulations of about 90/100 ns with and without hydrodynamics. Our simulations show that OPEP is capable of maintaining the native fold of the proteins stable within 100 ns, in the absence and also in the presence of the surrounding fluid. In Table 2, we report the average values of a set of collective variables (CV) monitoring the stability of the proteins, namely, the C_α root-mean-square deviation (rmsd) calculated with respect to the initial configuration (from NMR or X-ray), the fraction of native contact (Q), and the fraction of torsional angle (n_t). The formal definition of these variables[65] is given in Supporting Information; here, we just mention that the rmsd was computed only for the rigid

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

core of the protein, the well-defined region by NMR as reported in Table 2. For all of the CVs, the average was computed using the trajectory length and is reported in Table 2. In each column, the first value refers to the solvent free MD and the second one to the simulations including the LB evolution (LBMD).

By visual inspection, we first note that the protein's relaxed configurations look very similar in both simulations. This can be also appreciated by looking at the final steady values of rmsd and (n_t) in Figure 2. This result is not surprising since

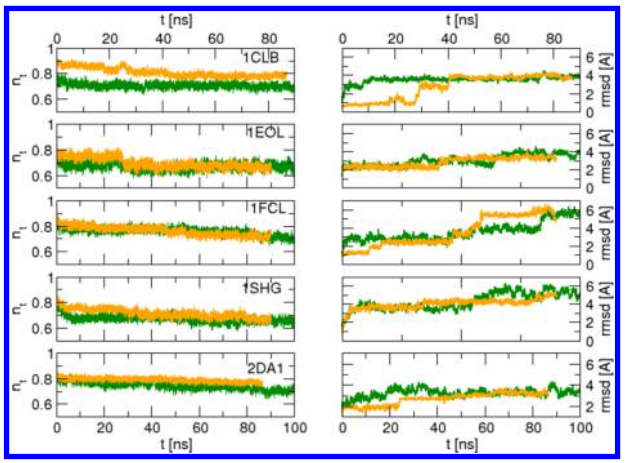


Figure 2. Time evolution of the fraction of native torsion n_t and RMSD for the set of proteins. Data in green refer to simple Langevin dynamics, and data in yellow refer to simulation with HI (LBMD).

hydrodynamics does not affect the thermodynamics of the system; hence, the states at equilibrium (and their distribution) should be the same. Averaged over the full trajectories, the five proteins display an average rmsd varying between 3.0 (1E0L) and 4.3 (1SHG) in solvent free MD and between 2.7 (1CLB) and 4.0 (1SHG) including hydrodynamics. At the end of the simulations, three proteins over five exhibit a very low rmsd, about 3/4 Å; in two other cases (1FCL and 1SHG proteins), the deviation from the initial state is larger, but the drift is mainly due to the rearrangement of the secondary structures with no systematic unfolding observed; see details in Supporting Information. However, when looking at the relaxation process in itself (Figure 2) it appears that the protein evolution is somehow softened by HI, and we observe longer lived intermediate states. This is also mirrored by the different average values for each CVs when computed along all the trajectories; see Table 2.

Since the coupling between protein dynamics and the fluid is encoded in the friction coefficient γ (see eq 4), we tested its effect on protein structural relaxation. From the final equilibrated configuration of the protein 1E0L, we restarted a series of simulations with different values for the parameter γ and reassigned the particle velocities. For the pure Langevin dynamics, γ relates to the memory-less motion, but when HI are included, the coupling accounts for both the fluid viscosity and the effective hydrodynamic radius of the molecular center of force, i.e., according to Stokes law $\gamma = 6\pi\eta a/m$ with η, a, and m the dynamic viscosity of the fluid and the size and the mass of the particle, respectively. In Figure 3, the time evolution of the fraction of native torsion n_t is reported for the two type of simulations and for several values of γ. During the solvent free MD, the protein tends to drift away from the initial structure,

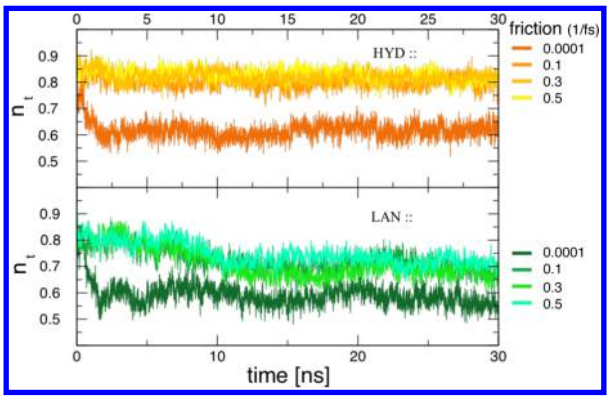


Figure 3. Time line of the fraction of native torsions for 1E0L. It was restarted from the equilibrated structure, using different values of the friction γ for the Langevin dynamics (bottom) and LBMD (top).

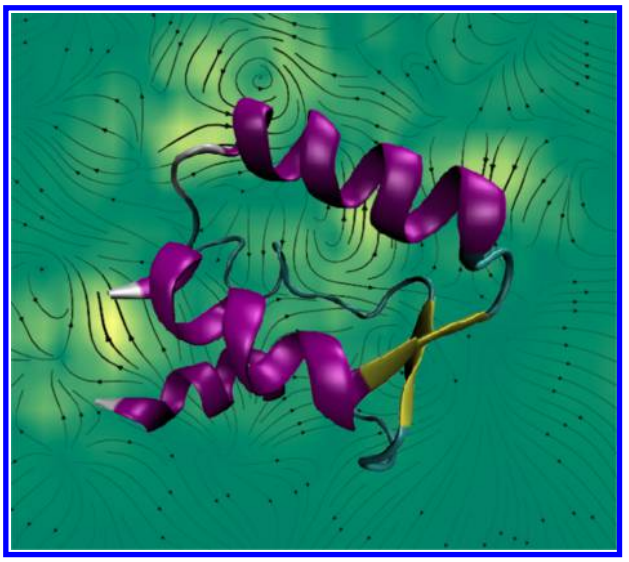


Figure 4. Protein 1E0L embedded in the hydrodynamic fluid: pictorial representation of the fluid streamline pattern around the protein.

occurring for all values of γ considered. Diffusion of a single Langevin particle is inversely proportional to the friction coefficient, $D = kT/m\gamma$ so that in the weak coupling regime particle motion is faster. On the contrary, when HI are included, the system shows a long-lived memory and tightly fluctuates around the starting configuration. Only in the limit of vanishing coupling, the relaxation drives the protein toward new configurational states. This finding shows how HI drain or pump energy in the protein by altering the relaxation process from one state to another. Interestingly, while HI have been reported to accelerate folding events in a Gō-like model,[26,31] by favoring the collapse toward the compact structure, no author has yet studied how HI affect the internal relaxation of a folded protein. Our preliminary results indicate that for the compact folded state the system may survive longer in conformational substates when HI are included. A pictorial representation of the fluid organization around the protein is given in Figure 4.

The possibility to control the coupling between the fluid evolution and the protein internal dynamics, for example, by tuning either the friction γ or the viscosity η, makes our framework a powerful tool to investigate how protein motion is influenced by the solvent viscosity without an explicit representation of the latter. A related, interesting debate

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 3. Traslational and Rotational Diffusion Constants Calculated from Simulations Including or Excluding HI (Left and Right Values in the Columns, Respectively)[a]

system	vectors	D_T [Å^{2}ns^{-1}]	D_R [ns^{-1}]	HYDROPRO D_T [Å^2 ns^{-1}]	HYDROPRO D_R [ns^{-1}]
1CLB	3−13, 27−35	2.8/5.5	$0.8/0.6\cdot10^{-2}$	13.1	$3.4\cdot10^{-2}$
1E0L	9−12, 16−21	4.5/8.4	$1.2/1.4\cdot10^{-2}$	15.6	$5.6\cdot10^{-2}$
1FCL	3−7, 24−36	3.3/5.0	$0.4/1.2\cdot10^{-2}$	14.2	$4.3\cdot10^{-2}$
1SHG	24−29, 44−48	4.5/6.5	$0.5/1.0\cdot10^{-2}$	14.0	$4.4\cdot10^{-2}$
2DA1	35−43, 50−62	2.0/4.1	$0.4/1.5\cdot10^{-2}$	11.2	$2.0\cdot10^{-2}$
1E0L*	9−12, 16−21	−/17.0	$-/4.6\cdot10^{-2}$	15.6	$5.6\cdot10^{-2}$

[a]In the last two columns, the "experimental" values estimated by the HYDROPRO software[71] are reported for comparison. In the last row, we report the values computed for a LBMD simulation using friction $\gamma = 0.05$ fs^{-1}.

concerns the effect of solvent viscosity on protein folding kinetics.[66−68] A weak dependence on solvent viscosity has been reported experimentally and rationalized as being due to protein internal friction on dominating the folding kinetics.[67,68] Moreover, we also mention that the coupling with an effective solvent also opens the possibility to explore energy transfer processes from solvent to protein and vice versa[69] in the context of water-free CG models.

In the current implementation, the coupling parameter γ is the same for all molecular centers of force, but heterogeneous, per-particle values depending on the amino-acid nature or local topology of the protein/solvent interface can be easily incorporated. In fact, recently all atom simulations clearly probed the heterogeneity of the solvent viscosity in the proximity of biological interfaces.[70]

3.2. Transport Properties. We now turn our attention to the transport properties of the proteins. From the performed simulations, we extracted the translational and rotational diffusion constants. For the former, we performed a linear fit of the mean square displacement (MSD) at long time scale (<40 ns) and extracted the diffusion constant D_T using the Einstein relationship, $\langle \delta R^2 \rangle = 6D_T t$. The rotational diffusion constant was calculated by considering the second order rotational relaxation of selected rigid vectors of the protein structures. Namely, for each molecule we have individuated two vectors associated with the orientation of two well-defined rigid secondary structure elements. The fragment of the proteins defining these vectors are reported in Table 3. The decay ($t < 30\ ns$) of the time-correlation functions of the vector orientation, $C_2(t) = \langle 1/2(3cos(\theta(t) - 1)) \rangle$, were then fitted with an exponential function, $C_2(t) = e^{-6D_R t}$. The estimated values of D_T and D_R are reported in Table 3 for all the simulations. Following the framework of ref 26, we have also estimated for our systems the experimental diffusion constants by using the software HYDROPRO.[71]

As already reported by other authors,[26,29,33] the HI systematically enhance the diffusivity of the molecules in the solution. For example, when HI are included in the Brownian Dynamics of the Gō-like model of proteins, the diffusion coefficient for folded and unfolded proteins[26] is a factor of 20 larger as compared to that in the absence of HI and are essential to correctly reproduce the experimental data. In our case, the speed up is about a factor of two for the translational motion. For the LBMD simulations, the D_T is comprised in the range 4−8 Å^2/ns, and the results are about 2−3 times smaller than the experimental data. The same behavior is recovered for all but one system when the rotational diffusion is considered, D_R being 1.2 to 4 times larger when HI are included in the simulation than in the pure Langevin dynamics. The only exception is the 1CLB system for which the values for D_R are close to each other, but HI slightly reduce rotational motion. The values calculated from the LBMD trajectories are, as for translational motion, 1.2−4 times smaller than the experimental estimates. However, as already anticipated, the mobility of the protein depends on the parameters characterizing our fluid. In particular, we have probed for one protein (1E0L) that a fine-tuning of the friction γ results in a very good agreement with the experimental HYDROPRO value. In fact, when using $\gamma = 0.05$ fs^{-1} we found that the diffusion constants are $D_T = 17$ Å^2/ns and $D_R = 4.6\cdot10^{-2}$ ns^{-1}.

As already reported in ref 38, the LBMD scheme based on the OPEP force field was delivered to investigated the effect of crowding on protein mobility, showing that for high occupied volume fraction, ∼ 30%, diffusion at the nanosecond time-scale slows down about 1 order of magnitude with respect to the dilute solution limit, in good agreement with data from neutron scattering.[9]

3.3. Amyloid Aggregation. When applying the computational scheme to the problem of amyloid aggregation, the inclusion of HI in CG simulations is expected to crucially influence the kinetics of molecular aggregation.[22,33,34] We consider a system composed of 18 monomers of the peptide Aβ_{16-22} capped by Ace and NH_2 groups and placed in a box of dimension 65× 65 × 65 Å^3, corresponding to a concentration of 100 mM. The system contains 864 particles and is simulated for about 100/150 ns depending on the simulation scheme. The multiscale simulation is compared to the standard Langevin dynamics. The aggregation process is quantified by calculating the gyration radius of the system, R_g, the time evolution of which is reported in Figure 5. Our results clearly show that when OPEP is coupled to the solvent, the initial aggregation process speeds up by about 15%. In the presence of HI, the Aβ_{16-22} monomers collapse quickly and a first steady state is reached in about 20 ns, with the size of the aggregate being $R_g \simeq$ 15.7 Å. A second minor structural reorganization is observed at a longer time, 50 ns, making the aggregate more compact as a result of optimized packing. This state persists for the rest of the simulation. A similar two-step mechanism is also observed for the pure MD, but the size of the intermediate and final states are rather larger than the ones observed in the active fluid. The first aggregation takes place in about 20 ns, and the metastable structure has an $R_g \simeq$ 18 Å. The final collapse occurs at about 70 ns and is characterized by slow kinetics and leads to an aggregate of $R_g \simeq$ 16 Å. A pictorial representation of the aggregation is given in Figure 7.

This finding clearly shows how solvent mediated interactions facilitate the encounter and structural shaping of Aβ_{16-22} monomers, with the final collapse being accelerated by about 40%. Our result agrees with recent simulations reported by Ando and Skolnik,[33] where, using a sort of specialized dumbbell

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

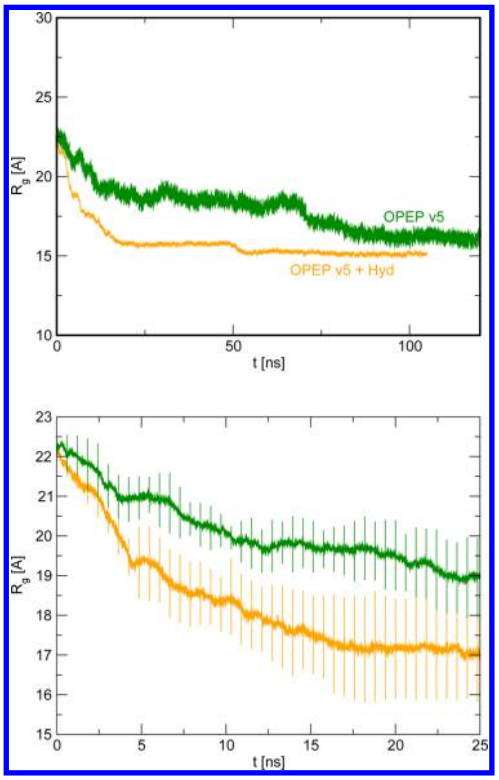


Figure 5. Top: Effect of hydrodynamic interactions on amyloid aggregation as monitored by the time evolution of the radius of gyration, R_g. Orange and green curves refer to the simulations with and without hydrodynamics, respectively. The exponential fit of the initial collapse is reported in solid lines (red and light green). The characteristic time τ of the initial collapse extracted by the exponential fit has a value of $\tau = 7$ and $= 8$ ns, for the simulation including HI and the one without. Bottom: Initial aggregation of the peptides as it results from the average of the time decay of R_g over independent trajectories. The error bars indicate the variability of the data computed as the (Max-Min)/2.

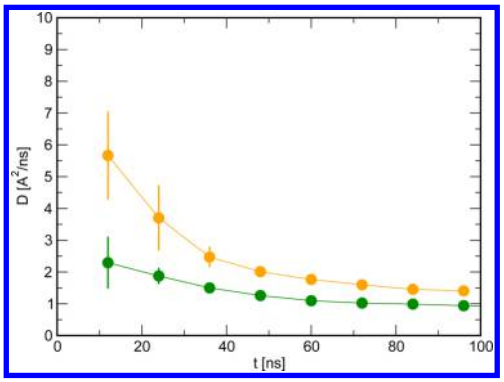


Figure 6. Average diffusion coefficient D calculated in several time windows of 12 ns each by fitting the mean square displacement of the peptides centers of mass. Orange and green symbols refer to simulations including or not including hydrodynamics interactions, respectively.

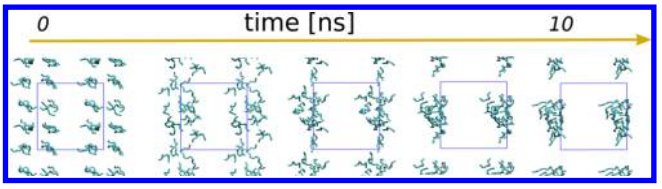


Figure 7. Molecular representation of the amyloid aggregation process from the LBMD simulation. In the initial state, the peptides are placed on a regular lattice, and in about 10 ns, they collapse to form a compact aggregate.

model, it was proved that HI substantially increase the rate of lipid aggregation. The faster collapse of lipids was rationalized in terms of a reaction-diffusion model[72] and by observing that for polymeric entities modeled as a series of N linked beads, the encounter rate is favored by HI due to the different scaling of the diffusion constant with the polymer size as compared to the free draining case. With HI, the diffusion coefficient scales as $D^{HI} \approx D_0 N^{-\nu}$ with $0 < \nu < 1$, whereas in the free draining limit, $D^{FD} \approx D_0 N^{-1}$.

We therefore compute the diffusion constant for the monomers in our system as a function of time. This allows detecting changes in monomer diffusivity during the aggregation process. The trajectories are separated in time windows of 12 ns each. For each time window, we calculate the *MSD* of the monomer center of mass and then average over all peptides. The average diffusion constant is extracted from the linear fit of the $MSD(t)$ for each window. The obtained values are reported in Figure 6.

We first notice that the mobility of the monomers is enhanced by HI by a factor 2.4 in the initial time window where the solution can be considered rather dispersed. As the aggregation process proceeds, the speed-up of monomer mobility reduces to a factor 1.5 with respect to the solvent free MD. An accurate investigation of the diffusion of amyloid proteins such as Aβ in a hydrodynamic fluid along with its dependence on fluid friction and viscosity is reserved for a forthcoming work; here, we just mention that with HI the initial diffusion coefficient is about a factor four smaller than the experimental value $D = 2.3 \times 10^{-10}$ m^2/s (23 Å^2/ns).[73] This difference is caused by two factors. First, as reported in the previous paragraph, with the initial choice of our parameters protein diffusion is 2−3 times slower than that in experiments. Second, because of the high concentration even in the initial time window, the monomers do not diffuse as in the very dilute solution.

From the single long simulations we just presented, it is not possible to extract statistically meaningful information about the kinetics of the collapse. To add extra weight to these preliminary results, we have generated five initial states, all sharing an initial value for the gyration radius of the system, $R_g \simeq 22-23$ Å. The time evolution of R_g was then average over the independent trajectories. The results plotted in the bottom panel of Figure 5 confirms that HI speed up the initial encounter of the peptides and their collapse.

A systematic test of how the diffusivity of amyloid species depends on the essential parameters of our simulations, i.e., the friction γ and the fluid viscosity ν, and on system concentration is reserved to a further investigation. However, our results confirm that also for the correct modeling of molecular aggregation in aqueous solution it is essential to include HI.[33]

3.4. Toward Crowding. In this final section, we unfold the potential of the LBMD technique to tackle the problem of protein crowding. The inclusion of HI in simulations of systems composed by tens, hundreds, and ultimately thousands of proteins has a heavy computational cost. Indeed, in terms of the LB, the computational effort scales cubically with the number of linear lattice points. In the test cases presented so far, we used a high resolution scheme for the LB part with a lattice spacing Δx

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

= 1 Å. In these conditions, depending on the size of the molecular system, the LB component and the LB−MD coupling contribute 30%−60% of the total elapsed time so that it is desirable to reduce the computational effort of handling the solvent.

Since OPEP, as other CG models, casts together multiple atoms to model the amino acid side chains (SC) and thus it associates to each SC center of force a characteristic radius of several Angstroms, the resolution of the HI lattice can be decreased without compromising significantly the accuracy of the overall method. We have tested this strategy on the amyloid system. In Figure 8, we report the elapsed time for a parallel

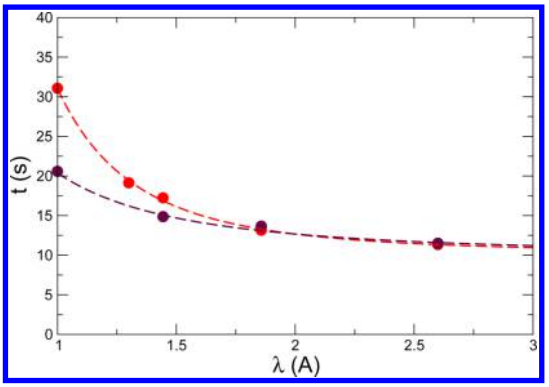


Figure 8. Benchmark of the multiresolution scheme. The time to produce 1000 MD steps on 8 cores for the system composed of 18 Aβ monomers is reported as a function of grid spacing. Red dots refer to synchronous coupling between the MD and the LB. Maroon dots refer to asynchronous coupling, with LB integration every two MD steps. The dashed lines indicate a fit; see the text.

simulation of 1000 MD steps allocated on 8 cores (Intel xeon 5660) for several resolutions of the LB mesh, Δx. The red dots refer to a synchronous coupling between MD and LB; thus, HI are evaluated at each MD step. As expected, the simulation time scales approximately as the cube of the mesh size and the benchmark points can be fitted with a function $t = A\cdot(\Delta x^{-\alpha} + t_0)$. For synchronous time-stepping, $\alpha = 3.12$, while for asynchronous time-stepping (HI evaluated each 2 MD steps), $\alpha = 1.95$. The value $At_0 \sim 10$ s is the same for both calculations, signaling the irreducible time due to MD plus the particle/fluid coupling. In fact, the coupling weighs about 60% (~6 s), and a standard 1000 MD steps costs about 4 s. For a selected protein, the 1E0L, we have proved that small changes in the lattice resolution have a very negligible impact on both protein stability and diffusion. For instance, with a spacing $\Delta x = 1.33$ and 2.0 Å, the average RMSD over a trajectory of 100 ns is 2.7 and 3.5 Å, respectively. Moreover, the effect on the diffusion constant is less than 10%. How lattice resolution and appropriate scaling of viscosity influences the kinetics of complex processes like amyloid aggregation will be discussed in a separate work.

As we increase the LB mesh spacing, the computational gain can be further improved by using a less expensive interpolation scheme to map the particle positions on the mesh points. The data presented in Figure 8 are obtained by using the interpolation scheme according to eq 6 and spanning 4 lattice spaces, as described in the Methods section. Using a simpler scheme such as the nearest point interpolation, we observed that the overall simulation cost further drops down by about 30%.

The performance of the simulations can be further improved by tuning the dynamics of the molecular systems. When the dynamics of the protein is governed by the complete OPEP Hamiltonian, the time step for the simulations is indeed confined to very small values, practically the typical values used in standard all atom simulations, 1−2 fs. However, when dealing with a protein suspension composed of many molecules, a fast initial sampling of the molecular packing does not require the finest details of the molecular motion. In this respect, several strategies can be adopted for enlarging the integration step, such as to define a harmonic elastic network to restrain the protein motion or a simplified Gō-like model based on the OPEP force field. The associated grained protein motion would be coupled to a grained fluid motion as discussed above. The finest resolution can be recovered to investigate in detail selected states of the suspension.

When molecular dynamics is evolved with large timesteps, the coupling with HI must be carefully scaled in order to maintain a physical control of the parallel-time evolution. As discussed in the Methods section, the intrinsic time scale of the fluid is controlled by the dynamic viscosity η (or ν), while the intrinsic time scale of the molecular system is encoded by the physical Hamiltonian and reflected by the integration time step. In synchronous LBMD, the fluid relaxation is evaluated at each MD step; thus, when the molecular system evolves with a larger time step, the fluid relaxation occurs at a slower rate. In order to preserve the same physical viscosity of water and its internal dynamics, the numerical viscosity should be scaled in such a way that $\nu = \nu_0\cdot(\Delta t/\Delta t_0)$, where t_0 and ν_0 are reference values. We have probed the scaling behavior by considering the zero temperature velocity relaxation of an amyloid peptide Aβ_{16-22} in fluid. We have assigned initial velocities (10^{-4} Å/fs) to each particle of the peptide and followed the time relaxation of the velocity of the peptide center of mass. When HI are included, three different time scales are expected.[74] As a reference, we have considered the relaxation process governed by a MD integration time step $\Delta t_0 = 1$ fs and a fluid viscosity $\nu_0 = 0.166$ (Å^2/fs). The associated data are reported in Figure 9. The relaxation was also monitored using a different time step $\tau = 0.5$ and 1.5 fs, and using either a viscosity $\nu = \nu_0$ (solid red and orange lines) or a scaled one, $\nu = \nu_0\cdot(\Delta t/\Delta t_0)$ (red and orange

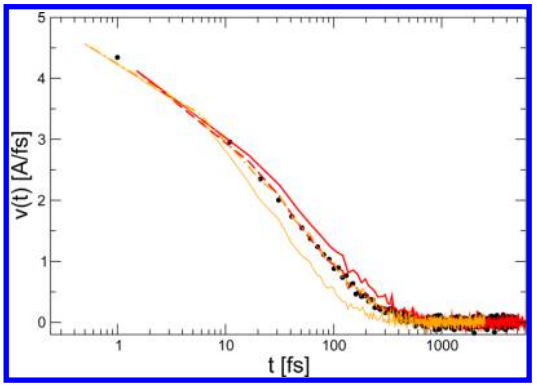


Figure 9. Time-relaxation of the CM velocity at zero temperature in synchronous MD and LB dynamics. Black circles refer to the reference tandem, molecular time step $\Delta t_0 = 1$ fs and fluid viscosity $\nu_0 = 0.166$(Å^2/fs). Solid red and orange lines refer to relaxation processes sampled by MD time step $\Delta t = 1.5$ fs and $\Delta t = 0.5$ fs, and viscosity ν_0; dashed lines refer to the same process sampled using a rescaled fluid viscosity, $\nu = \nu_0\cdot(\Delta t/\Delta t_0)$.

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

dashed lines). The time relaxations obtained via the scaled approach accurately reproduce the target relaxation process.

4. CONCLUSIONS

The coupling of a water-free, coarse-grained model with hydrodynamic interactions is an effective strategy to study protein suspensions and provides complementary information to all atom simulations, which necessarily accesses a smaller number of proteins and shorter time scales. A few research groups are active in this field and have presented challenging applications so far.[16,19,21,22,24,26,33,34] In the present work, we described a novel implementation that couples the OPEP force field for proteins with the Lattice Boltzmann methodology to resolve explicit and on-the-fly the solvent dynamics, leading to HI.

The framework combines the excellent performances of OPEP in reproducing the protein motion at quasi-atomistic resolution, with the excellent scalability of the LB technique on parallel architectures, making the scheme a powerful tool for inquiring protein structure and dynamics under crowding conditions.[38] We tested the simulation technique and showed that it is effective in monitoring protein fluctuations and mobility in the the folded state as well as in peptide aggregation.

In perspective, the proposed modeling of the solvent can be improved along the mesoscopic Lattice Boltzmann model for water-like fluids described in refs 75 and 76. By water-like fluids, we imply that the fluid may include generic features of the hydrogen bonds, particularly the tendency to form ordered domains, with selected bond angles, against the disordering action of thermal noise. The possibility to merge the OPEP force field with the water-like treatment of the solvent opens new ways to describe the fluid solvent with accurate thermodynamics yet with a minimal set of degrees of freedom. As a result, the intrinsic multiscale nature of the methodology would fully unfold and find application in those biological problems where such level of detail is required.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Definition of collective variables and description of protein stabilities; and benchmark data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: fabio.sterpone@ibpc.fr.

Funding

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) Grant Agreement no. 258748. A part of this work was performed using HPC resources from GENCI [CINES and TGCC] (Grant x201376818 and x2014076818). We acknowledge the financial support for infrastructure from ANR-11-LABX-0011-01.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thanks M. Chiricotto, M. Bernaschi, and M. Bisson for useful discussions.

■ REFERENCES

(1) Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597−604.
(2) Zhou, H. X.; Rivas, G.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375−97.
(3) Dix, J. A.; Verkman, A. *Annu. Rev. Biophys.* **2008**, *37*, 247−263.
(4) Hofling, F.; Franosch, T. *Rep. Prog. Phys.* **2013**, *76*, 046602.
(5) Wang, Y.; Mohona, S.; Smith, A. E.; Krois, A. S.; Pielak, G. J. *J. Am. Chem. Soc.* **2012**, *134*, 16614−16618.
(6) Jin, S.; Verkman, A. *J. Phys. Chem. B* **2007**, *111*, 3625−3632.
(7) Malchus, N.; Weiss, M. *J. Fluoresc.* **2010**, *20*, 19−26.
(8) Verkman, A. *Methods Enzymol.* **2003**, *360*, 635−648.
(9) Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R. M. J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11815−11820.
(10) Konopka, M. C.; Shkel, I. A.; Record, S. C. M.; Weisshaar, J. *J. Bacteriol.* **2006**, *188*, 6115−6123.
(11) Wang, Y.; Li, C.; Pielak, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 9392−9397.
(12) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. *J. Am. Chem. Soc.* **2011**, *133*, 7116−7120.
(13) Monteith, W. B.; Pielak, G. J. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *31*, 11335−11340.
(14) Sakakibara, D.; Sasaki, A.; Ikeya, T.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; Walchli, M.; Smith, B. O.; Shirakawa, M.; Guntert, P.; Ito, Y. *Nature* **2009**, *458*, 102−105.
(15) Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; Shirakawa, M. *Nature* **2009**, *458*, 106−110.
(16) Frembgen-Kesner, T.; Elcock, A. H. *Biophys. Rev.* **2013**, *5*, 109−119.
(17) Bicout, D. J.; Field, M. J. *J. Phys. Chem.* **1996**, *100*, 2489−2497.
(18) Ridgway, D.; Broderick, G.; Lopez-Campistrous, A.; Ruaini, M.; Winter, P.; Hamilton, M.; Boulanger, P.; Kovalenko, A.; Ellison, M. J. *Biophys. J.* **2008**, *94*, 3748−3759.
(19) McGuffee, S. R.; Elcock, A. H. *PLoS Comput. Biol.* **2010**, *6*, e1000694.
(20) Harada, R.; Tochio, N.; Kigawa, T.; Sugita, Y.; Feig, M. *J. Am. Chem. Soc.* **2013**, *135*, 3696−3701.
(21) Ando, T.; Skolnick, J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 18457−18462.
(22) Mereghetti, P.; Wade, R. C. *J. Phys. Chem. B* **2012**, *116*, 8523−8533.
(23) Feig, M.; Sugita, Y. *J. Mol. Graphics Modell.* **2013**, *45*, 144−156.
(24) Hasnain, S.; McClendon, C.; Hsu, M.; Jacobson, M.; Bandyopadhyay, P. *PLoS One* **2014**, *9*, e106466.
(25) Gruebele, M.; Thirumalai, D. *J. Chem. Phys.* **2013**, *139*, 121701.
(26) Frembgen-Kesner, T.; Elcock, A. H. *J. Chem. Theory Comput.* **2009**, *5*, 242−256.
(27) Szymczak, P.; Cieplak, M. *J. Phys.: Condens. Matter* **2011**, *23*, 033102.
(28) Durlofsky, L.; Brady, J. F.; Bossis, G. *J. Fluid Mech.* **1987**, *180*, 21−49.
(29) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352−1360.
(30) Ando, T.; Chow, E.; Skolnick, J. *J. Chem. Phys.* **2013**, *139*, 121922.
(31) Cieplak, M.; Niewieczerza, S. *J. Chem. Phys.* **2009**, *130*, 124906.
(32) Kamata, K.; Araki, T.; Tanaka, H. *Phys. Rev. Lett.* **2009**, *102*, 108303.
(33) Ando, T.; Skolnick, J. *Biophys. J.* **2013**, *104*, 96−105.
(34) Frembgen-Kesner, T.; Elcock, A. H. *Biophys. J.* **2010**, *99*, L75−L77.
(35) Song, W.; Wei, G.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2008**, *112*, 4410−4418.
(36) Chebaro, Y.; Pasquali, S.; Derreumaux, P. *J. Phys. Chem. B* **2012**, *116*, 8741−8752.
(37) Sterpone, F.; Nguyen, P.; Kalimeri, M.; Derreumaux, P. *J. Chem. Theory. Comput.* **2013**, *9*, 4574−4584.

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(38) Sterpone, F.; Melchionna, S.; Tuffery, P.; Pasquali, S.; Mousseau, N.; Cragnolini, T.; Chebaro, Y.; St-Pierre, J.-F.; Kalimeri, M.; Barducci, A.; Laurin, Y.; Tek, A.; Baaden, M.; Nguyen, P. H.; Derreumaux, P. *Chem. Soc. Rev.* **2014**, *43*, 4871−4893.
(39) Benzi, R.; Succi, S.; Vergassola, M. *Phys. Rep.* **1992**, *222*, 145−197.
(40) Chen, S.; Doolen, G. D. *Annu. Rev. Fluid Mech.* **1998**, *30*, 329−364.
(41) Ahlrichs, P.; Duenweg, B. *J. Chem. Phys.* **1999**, *111*, 8225.
(42) Maupetit, J.; Tuffery, P.; Derreumaux, P. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 394−408.
(43) Derreumaux, P. *J. Chem. Phys.* **1997**, *106*, 5260−5270.
(44) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, 025101.
(45) Spill, Y. G.; Pasquali, S.; Derreumaux, P. J. Chem. *Theory Comput* **2011**, *7*, 1502−1510.
(46) Nguyen, P. H.; Okamoto, Y.; Derreumaux, P. *J. Chem. Phys.* **2013**, *138*, 061102.
(47) Barducci, A.; Bonomi, M.; Derreumaux, P. *J. Chem. Theory Comput.* **2011**, *7*, 1928−1934.
(48) Chebaro, Y.; Derreumaux, P. *Proteins* **2009**, *75*, 442−452.
(49) Chebaro, Y.; Jiang, P.; Zang, T.; Mu, Y.; Nguyen, P. H.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2012**, *116*, 8412−8422.
(50) Nasica-Labouze, J.; Meli, M.; Derreumaux, P.; Colombo, G.; Mousseau, N. *PLoS Comput. Biol.* **2011**, *7*, e1002051.
(51) Kalimeri, M.; Derreumaux, P.; Sterpone, F. *J. Non-Cryst. Solids, Conf.* **2015**, *407*, 494−501.
(52) Marconi, U. M. B.; Melchionna, S. *J. Chem. Phys.* **2009**, *131*, 014105.
(53) Miller, A. A. *J. Chem. Phys.* **1963**, *38*, 1568−1571.
(54) Guo, Z.; Zhen, C.; Shi, B. *Phys. Rev. E* **2002**, *65*, 046308.
(55) Shan, X.; Yuan, X.; Chen, H. *J. Fluid Mech.* **2006**, *550*, 413.
(56) Fyta, M.; Kaxiras, E.; Melchionna, S.; Succi, S. *Comput. Sci. Eng.* **2008**, *3*, 10.
(57) Melchionna, S. *J. Chem. Phys.* **2007**, *127*, 044108.
(58) Succi, S. *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*; Clarendon Press: Oxford, U.K., 2001.
(59) Melchionna, S.; Marconi, U. M. B. *Europhys. Lett.* **2008**, *81*, 34001.
(60) Horbach, J.; Succi, S. *Phys. Rev. Lett.* **2006**, *96*, 224503.
(61) Duenweg, B.; Ladd, A. *Adv. Polym. Sci.* **2008**, *221*, 89.
(62) Bernaschi, M.; Melchionna, S.; Succi, S.; Fyta, M.; Kaxiras, E.; Sircar, J. *Comput. Phys. Commun.* **2009**, *180*, 1495−1502.
(63) Landau, L. D.; Lifshitz, E. M. *Fluid Mechanics*; Pergamon Press: London, 1963.
(64) Guyon, E.; Hulin, J.-P.; Petit, L.; Mitescu, C. D. Physical Hydrodynamics. Oxford University Press: Oxford, U.K., 2007.
(65) Kalimeri, M.; Rahaman, O.; Melchionna, S.; Sterpone, F. *J. Phys. Chem. B* **2013**, *117*, 13775−13785.
(66) Schulz, J. C. F.; Schmidt, L.; Best, R. B.; Dzubiella, J.; Netz, R. R. *J. Am. Chem. Soc.* **2012**, *134*, 6273−6279.
(67) Chung, H. S.; Eaton, W. A. *Nature* **2013**, *502*, 685−688.
(68) de Sancho, S.; Sirur, A.; Best, R. B. *Nature Comm.* **2014**, *5*, 4307.
(69) Moritsugu, K.; Kidera, A.; Smith, J. *J. Phys. Chem. B* **2014**, *118*, 8559−8565.
(70) Pronk, S.; Lindahl, E.; Kasson, P. M. *Nature Comm.* **2014**, *5*, 3034.
(71) Torre, G. D. L.; Huertas, J. M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719−730.
(72) Northrup, S. H.; Hynes, J. T. *J. Chem. Phys.* **1979**, *71*, 871−883.
(73) Danielsson, J.; Jarvet, J.; Damberg, P.; Gräslund, A. *Magn. Reson. Chem.* **2002**, *40*, S89−S97.
(74) Hinch, E. J. *J. Fluid Mech.* **1975**, *72*, 499−511.
(75) Mazzitelli, I.; Venturoli, M.; Melchionna, S.; Succi, S. *J. Chem. Phys.* **2011**, *135*, 124902.
(76) Moradi, N.; Greiner, A.; Melchionna, S.; Rao, F.; Succi, S. *Phys. Chem. Chem. Phys.* **2014**, *16*, 15510−15518.

DOI: 10.1021/ct501015h
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX