Molecular Dynamics of an Embedded-Charge Model of Lysozyme Aqueous Solutions

M. C. Abramo,* C. Caccamo, D. Costa, G. Pellicane, and R. Ruberto

Dipartimento di Fisica, Università degli Studi di Messina and CNISM (Consorzio Nazionale Interuniversitario di Struttura della Materia), Viale F. Stagno d'Alcontres 31, 98166 Messina, Italy

Received: February 22, 2010; Revised Manuscript Received: May 28, 2010

The onset of liquid−vapor separation in an interaction site model of a lysozyme aqueous solution is investigated by means of molecular dynamics (MD). Calculations are performed for a soft-core version of a potential early introduced by Carlsson et al. (*J. Phys. Chem. B* **2001**, *105*, 9040; **2001**, *105*, 12189.) whose liquid−vapor coexistence was studied by Rosch and Errington (*J. Phys. Chem. B* **2007**, *111*, 12591.); our modified model preserves the tailoring onto the experimental lysozyme solution properties embodied by those descriptions. We first show that the structural results obtained by Carlsson et al. at ambient conditions are quite well reproduced by our approach. Then, we cool the system along an isochoric path by monitoring the structural and dynamical properties at various temperatures. We thus show that a fluid−fluid separation takes place at a temperature 15% below the presumed binodal; in particular, we observe that density inhomogeneities develop rather early in the MD run and evolve over tens of nanoseconds into two dense aggregates that eventually merge, after ∼100 ns more, into a single liquid phase separated from a vapor region by a well-defined planar interface. The densities of the two coexisting fluids are compatible with previous determinations of the binodal line. The connections of this work to the overall scenario of phase stability investigations in protein solutions, as well as possible developments based on the use of more refined models, are discussed.

1. Introduction

Protein aggregation has a critical importance in many biological phenomena and is of crucial interest in the preparation and growth of single protein crystals employed for the investigation of the internal macromolecular structure.[1,2] It is known that protein crystallization can be substantially favored in protein solutions that undergo a thermally induced separation into two metastable fluid phases.[3,4] This issue has been addressed by a number of authors through experiments, as well as theoretical or computer simulation approaches (see, e.g., the reviews in refs 1 and 5–7). Among the latter, free energy calculations based on umbrella sampling techniques have indeed shown that, for a central potential model of protein solutions, the crystal nucleation rate increases dramatically in regions close to the metastable fluid−fluid critical point.[8]

Most computer simulation studies on protein solutions have been focused on the characterization of structural properties or phase equilibria. In this respect, a molecular dynamics (MD) approach allows one to investigate the time evolution of such systems and possibly the dynamical processes leading to a phase separation. The success of such a search does not seem obvious *a priori*, as quite long simulations might be necessary to observe the onset, for instance, of liquid−vapor coexistence, which might be hindered by finite-size effects and by the geometry of the simulation cell employed.[8,9] An MD investigation of this type has been performed by two of us[10] for the potential employed in ref 8; herein, we undertake an MD study of a more realistic protein solution model introduced by Carlsson, Malmsten, and Linse (henceforth CML),[11,12] in which the protein−protein interactions are described in terms of distributed interaction sites. Coarse-grained models of this kind fall at an intermediate level between fully atomistic calculations and central potential descriptions in which the macroparticle interaction is usually approximated in terms of simple short-range forces. The former models can determine quite accurately specific properties of the protein system, such as the kinetics of protein folding or free energy surfaces (see, e.g., the recent reviews in refs 13–15), but their use is typically limited to two or a few more proteins in the solvent environment; the latter provide only a qualitative description of the generic phase diagram of protein solutions.[16−26] In this instance, we have shown[27,28] that a DLVO (Derjaguin−Landau−Verwey−Overbeek)-like model[29] of lysozyme and γ-crystallin solutions is able to yield a phase diagram that well reproduces the shape of both the binodal and sublimation experimental lines.

A key feature of coarse-grained models is that they embody the intrinsic anisotropy of protein−protein interactions, a property that also stems from the nonuniform charge density at the molecular surface.[30] Distributed site interaction effects are known to be relevant both in the crystal[31] and in the disordered phase[32] of protein systems. As a result of their existence, one specific region of the protein surface interacts with another one in a mechanism underlying protein recognition, often described in terms of patch−patch interactions;[33] this is the case, for instance, of sickle hemoglobin molecules in which lateral and axial patch−patch interactions are active.[34] Models of this type should be more appropriate for reproducing the detailed shape of the phase diagram as well as more complex phenomena such as self-assembly.[32] Distributed interaction models have been studied by various authors.[11,12,32−49] In particular, a system involving a small number of surface charged sites has been investigated by MD simulations to rationalize the second virial coefficient behavior in protein solutions.[39,40] Models with a few patches have also been studied with a primary attention being paid to gelation/vitrification processes and limited valence effects (see refs 33, 50, and 51 and references therein). A coarse-grained strategy aimed at modeling concentrated protein solutions was systematically exploited in refs 44–46.

* Corresponding author. E-mail: mcabramo@unime.it.

10.1021/jp101590y
Published on Web 06/25/2010

The CML model of lysozyme aqueous solutions[11,12] is based on a tailoring of experimental solution conditions in which the second virial coefficient is fitted, with the lysozymes approximated by spherical macromolecules and the amino acid residues described as a collection of charged sites whose spatial distribution reproduces that of the real protein. CML[12] used Monte Carlo (MC) simulations to study the general structure and clustering properties of such a model, and the results appear to be qualitatively consistent with experimental data. The liquid−vapor coexistence, in a different choice of potential parameters, was later studied by Rosch and Errington,[47] through grand-canonical transition matrix MC simulations, by finding a binodal line narrower than the experimental one; these authors later studied, through the same MC technique, the phase behavior of their model in a confined geometry.[48]

As outlined in the preceding review, most studies on coarse-grained models have been devoted to clarifying generic aspects of protein solution behavior; in fact, aside from the CML case, few of these investigations have focused on specific proteins,[32,34,36,42] and apparently none has encompassed dynamical properties. In such a scenario, we consider it worthwhile to perform an MD study of the CML model, in which a continuous steeply repulsive potential is employed for the central interactions as more suitable for MD simulations. We consider three parametrizations of our model, all characterized by the same distributed charge interactions as in the CML work, but sensibly differing for the excluded-volume and hydrophobic interaction parameters. We first study structural features at ambient conditions and compare our results with those obtained by CML;[12] then, we investigate the evolution of the system upon cooling along an isochoric path, to ascertain whether and under what conditions a fluid−fluid separation occurs.

Brownian dynamics (BD) is another effective simulation technique widely used to investigate the dynamical properties of complex solutions.[54,55] The BD method offers the possibility to take into account implicitly the effects of the solvent environment on the dynamics of the macroparticles, through the random frictional forces appearing in the constituting equations. Diffusion times are thereby correctly described, and because a real solution requires a macroscopic time to evolve a phase separation, this implies that a dynamical investigation of such a process might require a prohibitively long simulation time. On the other hand, the lack of solvent effects in our model prevents a straightforward comparison between its dynamical properties and those corresponding to real protein solutions. Such properties instead provide useful information on the inherent time scales of our model as thermodynamic conditions run and the system approaches the liquid−vapor phase separation. This point will be explicitly discussed in the presentation of results.

This article is organized as follows. Models are introduced in section 2. Methods and simulation details are described in section 3. Results are reported and discussed in section 4, and conclusions are given in section 5.

2. Models

As anticipated in the Introduction, a lysozyme solution at various protein concentrations, pH values, and salt molarities is represented in the CML works[11,12] by an effective one-component fluid, where the interaction among macroparticles is given by two different contributions. The first is constituted by a central interaction that we write in the form

$$U(r) = \begin{cases} \infty & \text{if } r \le \sigma_{\mathrm{CML}} \\ \varepsilon_{\mathrm{CML}}\left(\dfrac{\sigma_{\mathrm{CML}}}{r}\right)^6 & \text{if } r > \sigma_{\mathrm{CML}} \end{cases} \quad (1)$$

where $\sigma_{\mathrm{CML}} = 37.08$ Å is a hard-core diameter and the attractive term—written by CML as ε/r^6, so that their ε corresponds to our $\varepsilon_{\mathrm{CML}}(\sigma_{\mathrm{CML}})^6$—takes into account hydrophobic forces between proteins in the aqueous solutions. The second contribution to the CML potential is given by a certain number of interaction sites representing the amino acid residues, which carry an appreciable amount of electrostatic charges. As illustrated in ref 11, the number of charges (up to 32) depends on the pH of the solution, according to the experimental titration curve;[54] the positions are adjusted so as to lie on a sphere 2 Å beneath the hard-sphere surface. Charges are then treated in a linearized Poisson−Boltzmann framework, giving rise to a (Debye−Hückel) screened interaction

$$U_{\mathrm{DH}}(r) = \frac{Z_i Z_j e^2}{4\pi\varepsilon_0\varepsilon_{\mathrm{r}}}\frac{1}{r}\exp(-\kappa r) \quad (2)$$

where r is the distance between two charged sites, Z_i is the charge of site i, e is the electronic charge, ε_0 is the permittivity of vacuum, and $\varepsilon_{\mathrm{r}} = 78.4$ is the relative permittivity of the solvent. The Debye screening length κ^{-1} is given by

$$\kappa^{-1} = \left[\frac{1}{\varepsilon_0\varepsilon_{\mathrm{r}}k_{\mathrm{B}}T}\sum_m (Z_m e)^2 c_m\right]^{-1/2} \quad (3)$$

where c_m is the bulk concentration of salt species m, related to the ionic strength by $I = 0.5\sum_m Z_m^2 c_m$. According to the procedure described in ref 11, Z_i assumes the values ± 1 and can be switched on/off depending on the pH of the solution. CML[12] determined the strength of the attractive interaction in eq 1 to be $\varepsilon_{\mathrm{CML}} = -3.2$ kJ/mol, according to a best fit of the experimental second virial coefficient of the full potential (given by the sum of contributions in eqs 1 and 2), at temperature $T =$ 298 K, pH ∼4.5, and ionic strength $I = 0.25$ M. The central contribution to the CML potential is shown in Figure 1.

In ref 12, the authors mainly focused on the structural properties of the lysozyme model at $T = 298$ K. More recently, Rosch and Errington (henceforth RE)[47] carried out an investigation of the liquid−vapor phase coexistence of a slightly modified CML model. Specifically, the relative permittivity depends explicitly on the temperature: according to experiments, RE assumed $\varepsilon_{\mathrm{r}} = 86.765 - 0.3232 \times T$ (°C). RE also used a different parametrization of the central interaction given by eq 1: They observed, in fact, that $\varepsilon_{\mathrm{CML}}$ gives rise to a critical temperature of the *bare* central interaction around 225 K; they then deepened the potential well to a value that corresponds to $\varepsilon_{\mathrm{RE}} = -4.27$ kJ/mol, thus obtaining a critical temperature closer to the experimental value of $T_{\mathrm{cr}} \approx 300$ K. The RE central part of total interaction is also shown in Figure 1. Interestingly enough, RE also showed that the phase behavior of the whole model is mainly determined by the central potential, the inclusion of the electrostatic interaction producing only small modifications of the overall shape of the liquid−vapor coexistence curve and of the critical temperature.[47]

In this study, we have modeled the electrostatic interaction through eq 2 as in previous works.[11,12,47] As far as the potential

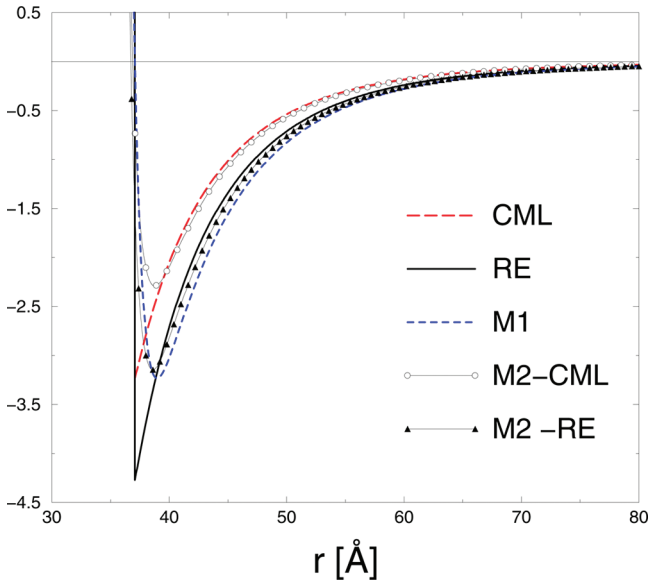

Figure 1. Comparison among the central interactions of various models employed in this work; see eq 4. The CML[12] and RE[47] parametrizations of eq 1 are also shown. The M2-RE curve refers to $T = 298$ K.

in eq 1 is concerned, we have approximated its shape through the soft-core interaction

$$U^{\text{soft}}(r) = \varepsilon\left[\left(\frac{\sigma}{r}\right)^{48} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{4}$$

characterized by a minimum $U^{\text{soft}}_{\text{min}} = -7\varepsilon/(8 \times 2^{3/7})$ at $r_{\text{min}} = \sigma(2)^{1/14}$. The r^{-6} decay gives the same long-range behavior as previous models, whereas the 48-power repulsion sets a reasonable compromise between (i) the most common choice of a 12−6 Lennard-Jones potential, which yields too soft a repulsion, possibly allowing too-close approaches between the charged sites at high temperatures, and (ii) an even steeper repulsive barrier, which would yield too-intense impulsive forces at close contact, thus requiring exceedingly small time steps to integrate the equation of motions.

To fix ε and σ, we followed two different routes, henceforth called M1 and M2, giving rise to different estimates of potential in eq 4. Specifically, in M1, we assume $\sigma_{\text{M1}} = 37.08$ Å, the same hard-core size as used in the CML and RE models. Then, ε_{M1} is fixed in such a way that the M1 and CML attractive potentials have the same well depth of −3.2 kJ/mol, thus obtaining $\varepsilon_{\text{M1}} = -4.93$ kJ/mol. As in the previous studies, ε_{M1} and σ_{M1} are assumed to be independent of the physicochemical conditions of the solution. The central part of potential M1 so obtained is shown in Figure 1. In the more sophisticated procedure M2, we require that the potentials in eqs 1 and 4 have the same second virial coefficients B_2, that is

$$\frac{4}{3}\pi(\sigma_{\text{CML}})^3 + 4\pi\int_{\sigma_{\text{CML}}}^{\infty}\{1 - \exp[-\beta U(r)]\}r^2\,\mathrm{d}r = 4\pi\int_0^{\infty}\{1 - \exp[-\beta U^{\text{soft}}(r)]\}r^2\,\mathrm{d}r \tag{5}$$

where the left-hand side of this equation is B_2 of the original model, with the hard-sphere contribution explicitly reported in the first term of the sum, and the right-hand side is B_2 of model M2. To fix ε and σ, we need a second condition; this is derived through the application of the Weeks−Chandler−Andersen (WCA)[55] approach, usually employed in perturbation theories of liquid systems. In the WCA approach, one aims to determine the hard-core diameter that better fits soft-core repulsion effects. Here, we use this approach in a reverse way to find the soft-core repulsion parameter in eq 4 that better approximates the purely hard-core interaction in eq 1. Specifically, we require

$$\sigma_{\text{CML}} = \int_0^{r_{\text{min}}}(1 - \exp\{-\beta[U^{\text{soft}}(r) + U^{\text{soft}}_{\text{min}}]\})\,\mathrm{d}r \tag{6}$$

This equation above expresses the well-known Barker−Henderson prescription.[56] The quantity $[U^{\text{soft}}(r) + U^{\text{soft}}_{\text{min}}]$ represents the "reference" repulsive part of potential in eq 4 that, according to the WCA splitting procedure,[55] embodies the whole repulsive forces up to r_{min}. Equations 5 and 6 constitute a closed system that can be solved to determine unequivocally the pair ε_{M2} and σ_{M2}. Note that both equations depend explicitly on the temperature, implying that different values of ε_{M2} and σ_{M2} are obtained as the temperature changes.

When the procedure just described is applied to the CML potential (for the unique temperature $T = 298$ K envisaged in ref 12), one obtains $\varepsilon_{\text{M2}} = -3.52$ kJ/mol and $\sigma_{\text{M2}} = 36.88$ Å. The corresponding full potential, given by the sum of eqs 4 and 2, is denoted henceforth as M2-CML; its central part, $U^{\text{soft}}(r)$ is shown in Figure 1. A similar M2 calculation can be performed for the RE parametrization (henceforth the M2-RE model), giving, for instance, at $T = 298$ K, $\varepsilon_{\text{M2}} = -4.84$ kJ/mol and $\sigma_{\text{M2}} = 36.73$ Å. The corresponding $U^{\text{soft}}(r)$ function is also shown in Figure 1. Other values of ε_{M2} and σ_{M2} for different temperatures are listed in Table 1, along with all other ε and σ parameters discussed in this section. It appears from Figure 1 and Table 1 that, at $T = 298$ K, M2-RE and M1 are very close to each other; also, the dependence of ε_{M2} and σ_{M2} on the temperature is rather weak.

TABLE 1: σ and ε Parameters Entering Eqs 1 and 4[a]

model	T (K)	σ (Å)	ε (kJ/mol)
CML	−	37.08	3.2
RE	−	37.08	4.27
M1	−	37.08	4.93
M2-CML	298	36.88	3.52
M2-RE	320	36.76	4.80
	300	36.73	4.83
	280	36.69	4.87
	260	36.66	4.91
	240	36.61	4.95

[a] CML[12] and RE[47] parameters for eq 1 do not depend on the temperature, like the M1 parameters for eq 4. The M2-CML model, eq 4, was studied only at $T = 298$ K, whereas the dependence of the M2-RE model on the temperature is explicitly reported for several values of T.

The mapping we have performed has specific implications for the critical-point temperature, T_{cr}. Indeed, the Noro−Frenkel (NF) law of corresponding states,[57] holding for a large class of central potentials, establishes a direct link between T_{cr} and an appropriate rescaled measure of the second virial coefficient, in terms of an interaction length of a square-well potential. We first recall that, as determined by RE through simulations,[47] the CML and RE potentials have critical temperatures $T_{\text{cr}} = 225$ and 300 K, respectively. As can be verified, both models accurately follow the NF rule. Then, because we have built up the M2 potentials according to eqs 5 and 6 (i.e., precisely through the NF prescriptions), we can reasonably conclude that

the M2-CML and M2-RE potentials will have by construction $T_{cr} = 225$ and 300 K, respectively. The application of the NF rule to potential M1-CML finally yields $T_{cr} = 304$ K. We observe that the almost-equal critical temperatures of the M1-CML and M2-RE models allow us to expect qualitatively similar thermodynamic results in both cases; such an outcome could also be expected on the basis of the close correspondence among the respective potential patterns in Figure 1. For this reason, we investigate along the isochore cooling path only the M2-RE model, implicitly assuming that quite similar results could be obtained for the M1-CML model.

As far as the mass of the lysozyme molecule is concerned (14320 amu), this has been distributed, in a balanced manner, over six point masses of 2000 amu each, placed symmetrically along the x, y, and z axes of the molecular reference frame, at a distance of 14 Å from the center of the macromolecule. The center itself bears the residual mass of 2320 amu. The distance of the six masses from the center was chosen through preliminary simulations so as to achieve, within the equilibration time of the MD runs, an optimized balance between the translational and rotational kinetic energies.

3. Molecular Dynamics Simulations

Using MD simulations, we have studied the static and dynamic properties of the models introduced in the previous section. For this purpose, we used MOLDY,[58] a program dealing with rigid molecules, either monatomic or polyatomic, whose orientations are represented by quaternions. The dynamical equations are integrated through a leapfrog algorithm; the rotational part implements a fast symplectic-reversible integrator, described in ref 59 with good energy and angular momentum conservation properties, especially in long-time simulations. MOLDY implements a Nosé−Poincaré scheme[60] to thermostat the system; the scheme is based on a symplectic discretization method to solve the dynamical equations, quite stable even in the presence of large temperature fluctuations, and preserves the time-reversal symmetry.

We have generally employed a time step of $\Delta t = 5$ fs; this short time step guarantees a good long-term conservation of the total energy of the system (i.e., the energy of the particles in the sample plus the contributions coming from the extra heat reservoir degree of freedom), never exceeding 1−4 parts over 10000. In our simulations, we used samples composed by about 1000 particles (with a total of ∼30000 interaction sites). MOLDY treats such relatively large samples efficiently by means of a linked-list cell algorithm.[61] The interaction cutoff was fixed to 130 Å, roughly corresponding to 3.5 times the protein diameter, with standard long-range corrections; this value is only slightly larger than that employed in previous works,[12,47] namely, 120 Å. Under these conditions, 10^5 simulation steps take 23−27 h, depending on the temperature, on 16 AMD Opteron dual cores at 2.4 GHz. After preliminary tests, the mass parameter for the extra variable representing the heat bath was fixed to 1000 kJ ps^2/mol. This choice allows a tight control of the enforced temperature (within a 1 K precision), the rotational and translational temperatures freely fluctuating around the desired value with a spread of about 2−3 K.

For simulations at ambient conditions, the system is equilibrated at $T = 298$ K during 50-ps runs, with the temperature rescaled at every time step followed by a 500-ps equilibration period, during which the system is left free to evolve subject to the Nosé−Poincaré thermostat. After this initial stage, cumulation runs follow for 2 ns more, divided into blocks of 0.5 ns to calculate the variance over the calculated average values. As for the isochoric path, simulation were carried out in the temperature interval $T = 320-240$ K at intervals of 10 K. For each temperature, the cumulation runs evolved for time intervals ranging from 2 ns at $T = 320-290$ K, to 4−4.5 ns at $T = 280-260$ K, and to 7 ns at $T = 250$ K. At the lowest temperature investigated here, $T = 240$ K, we employed a time step of 50 fs and followed the system evolution over a substantially longer time span, extending to 150 ns. Such long runs were required to observe and properly relax the inhomogeneous phase. At this stage, we preliminarily verified that the evolution of the potential energy obtained with $\Delta t = 5$ fs is highly reproducible with the larger time step employed, even at the level of local behavior. The Hamiltonian conservation properties only slightly worsens, rising to about 7−10 parts over 10 000. On the other hand, this strategy has allowed us to monitor the system over long-time trajectories, otherwise requiring an exceedingly large computational time.

4. Results

We first proceed to a comparison of our soft model predictions with those of the CML hard-core model[12] at $T = 298$ K. We calculate the center−center radial distribution function (rdf) $g(r)$ and its Fourier transform, the structure factor $S(k)$, for the M1 and M2-CML potentials at different protein net charges and ionic strength $I = 0.05$ M. We study solutions with total protein charge $Z = 0$ (pH ∼11, corresponding to the isoelectric point of the solution), as obtained from the balance of 12 positive and 12 negative charges, and $Z = 10$ (i.e., pH ∼4.5), obtained as the sum of 19 positive and 9 negative charges. To such an end, we perform *NVT* simulations of a system composed of $N = 1000$ particles enclosed in a cubic box with a side of 803.6 Å and periodic boundary conditions, corresponding to a lysozyme concentration of $C_p = 3.2$ mM (i.e., ∼46 g/L) or equivalently to a number density of $\rho = 1.93 \times 10^{-6}$ Å^{-3}, corresponding to a packing fractions of $\eta \equiv \pi\rho\sigma^3/6 = 0.0515$ for $\sigma = 37.08$ Å (i.e., for the M1 model) and $\eta = 0.0507$ for $\sigma = 36.88$ Å (i.e., for the M2-CML model).

As is visible from Figure 2, our $g(r)$ functions rise softly from zero to a first maximum at a distance slightly greater than the protein diameter; at $Z = 0$, both models yield a radial distribution function that compares qualitatively well with its counterpart of ref 12. At $Z = 10$, model M1 overestimates the height of the first peak, consistent with the fact that particles feel a stronger central attraction with respect to the original CML prescription; on the other hand, model M2-CML underestimates the same feature but is able to better follow the overall pattern reported in ref 12. A quite satisfactory agreement then emerges from Figure 2 between the present $S(k)$ function and those of CML. In fact, for both protein net charges considered here, the M1 and M2-CML models predict structure factors that faithfully reproduce the main features of their CML counterparts and, in particular, the steep rise in the low-k region for $Z = 0$.

To investigate the occurrence of a phase separation in our model, we perform a sequence of *NVT* simulations along an isochore path between 320 and 240 K. We consider in this case $N = 1029$ lysozyme molecules interacting through the M2-RE site−site potential. Particles are contained in a periodically repeated tetragonal cell, whose sides are $L_x = L_y = 433.5$ Å, $L_z = 3L_x$; these conditions correspond to a solution with a protein concentration $C_p = 100$ g/L ($\rho = 4.21 \times 10^{-6}$ Å^{-3}, corresponding to a mean packing fraction $\eta = 0.1075$ for the M2-RE model; see Table 1). The choice of a noncubic cell favors a phase separation preferentially occurring through the formation of an interface perpendicular to the z axis, an orientation

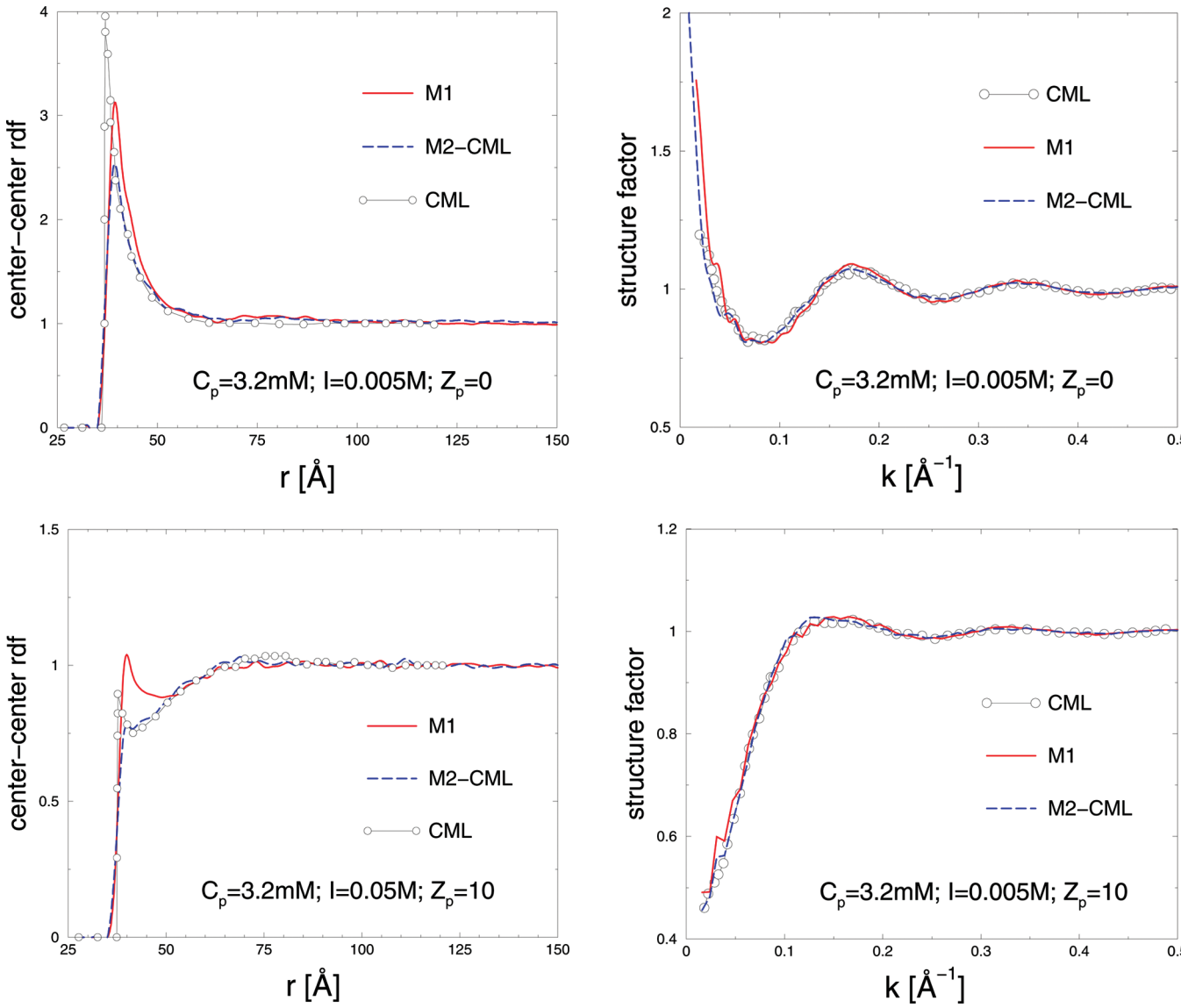

Figure 2. Center−center radial distribution functions (left) and corresponding structure factors (right) for the M1 and M2-CML potentials, at $T =$ 298 K (solution conditions are reported in the panels). Results from the MC simulations of the CML model[12] are also shown.

corresponding to the lowest interfacial area. We then fix a pH ∼4.5, and hence $Z = 10$ (see above), and an added salt concentration of 7%, equivalent to an ionic strength of 1.19 M.

We start by considering, in the left panel of Figure 3, the equilibrium behavior of the potential energy U at fixed temperatures, monitored over the last 2-ns simulation. In the right panel, average values at different temperatures are shown. As is visible, between 260 and 240 K, a marked decrease of the internal energy occurs, with a corresponding change of slope of the U vs T behavior. Moreover, whereas fairly constant values of U are achieved within a time span ranging from 2 to 7 ns for $T \geq 250$ K, at 240 K, the potential energy quite soon exhibits a trend to decrease with time, thus deserving a much longer quest for stable behavior. This is illustrated in Figure 4: It appears that U almost monotonically decreases for ∼140 ns, until it undergoes an abrupt fall, after which it fluctuates around a reasonably stabilized value. Such features clearly signal a substantial internal rearrangement of the system occurring at $T = 240$ K. As we alluded to in the Introduction, our rather detailed account of time scales serves as a useful tool to illustrate how our model relaxes and explores different temporal regimes, as the temperature decreases.

The processes leading up to a fully resolved phase separation are further investigated by the analysis of the density profile along the z axis at decreasing temperatures, reported in the left panel of Figure 5. It can be seen that, at 270 K, sparse fluctuations with respect to the average density are present, whereas at 240 K, two neatly defined distributions are formed, one fairly dense with $C_p \approx 350$ g/L and the other rarefied with $C_p \approx 30$ g/L. The development in time of such a density inhomogeneity is visible in the right panel of Figure 5 and closely follows the corresponding evolution of the internal energy reported in Figure 4: Already after 7 ns, the system exhibits regions of internal densities of ∼200 g/L together with relatively rarefied zones; after 75 ns, two well-defined density peaks are present with an intermediate extended interval of very low density. Finally, at 150 ns, a neat change has occurred, with the appearance of a density peak of ∼350 g/L at the box edges (corresponding to a unique aggregate because of the periodic boundary conditions) and a vast rarefied zone extending for more than 800 Å. It is easy to appreciate the close correspondence of these data with the snapshots of the simulation sample shown in Figure 6. As is visible, early condensed regions are formed at 7 ns, the spatial resolution of which

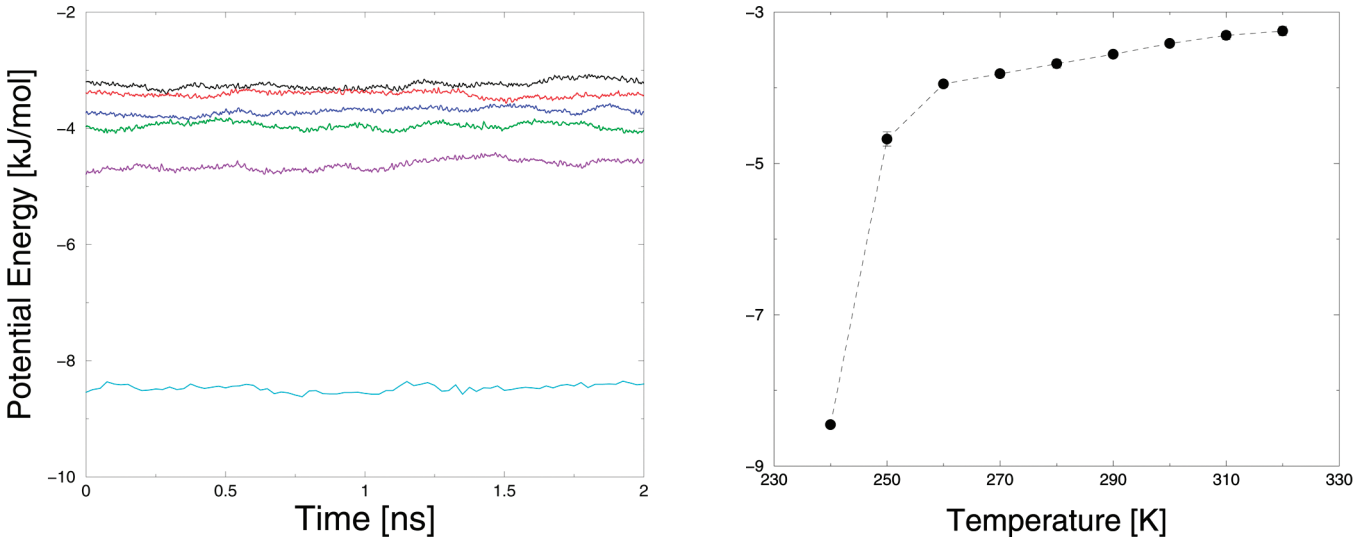

Figure 3. Left: Equilibrium behavior of the potential energy at different temperatures (from top to bottom, $T = 320$, 300, 280, 260, 250, 240 K). Right: Average potential energy as a function of temperature.

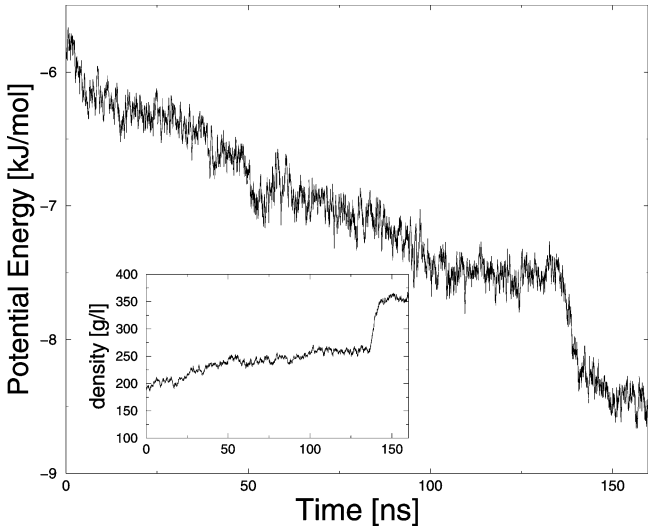

Figure 4. Evolution of potential energy at $T = 240$ K over the time lapse of 150 ns encompassed in our simulations. Inset: Density evolution in the region where a single liquid aggregate is formed after ~140 ns.

appears to be much better defined in the central frame after 75 ns, until eventually, at 150 ns, a single interface is formed. The latter originates from the coalescence of the formations visible in the snapshot at 75 ns; note that, as shown in the inset of Figure 4, the density of the concentrated phase is reached through a rapid increase during the fusion of the two previous aggregates.

The characteristics attained by the system configuration appear to be typical of a liquid−vapor phase separation. This picture can be substantiated on a quantitative basis by an analysis of the center−center $g(r)$ as a function of the temperature, shown in the left panel of Figure 7. It appears that, for $T \leq 270$ K, the $g(r)$ functions decay to 1 from above, a behavior that signals the onset of inhomogeneities in the system. The structure of the dense phase found at $T = 240$ K is shown in the right panel of the same figure. As for the evaluation of correlations in this region, the dense slab extends over ~200 Å (see Figures 5 and 6), and the $g(r)$ function can be calculated over ~600 molecules and estimated up to ~95 Å. As it appears from Figure 7, $g(r)$ has features typical of the liquid state, with a first peak on the order of 4 and well-defined first and second coordination shells.

The dynamical properties of the system as the temperature decreases can be analyzed in terms of the diffusive behavior of macroparticles. For $T \geq 250$ K, the mean square displacement (MSD) can be calculated over all of the particles, because they freely move throughout the simulation box even in the presence of appreciable inhomogeneities. As is visible in the left panel of Figure 8, the MSD patterns tend to flatten fairly regularly as the temperature is reduced. At 240 K, because of the phase separation that occurred, the exchange of particles among the two phases is very slow; then, we calculate the MSD, shown in the right panel of Figure 8, only for those particles that remain confined inside the dense phase. Such estimates were obtained at different times by sampling the system with parallelepiped boxes (dimensions detailed in Table 2), suited to the shapes of the aggregates formed. In this manner, we grant that the diffusional behavior is calculated under fairly constant density conditions (see inset of Figure 8). It appears that the 240 K pattern runs substantially below, and is markedly flatter than, the MSDs at higher temperatures.

The self-diffusion coefficients D are then calculated from the MSD. These are shown in Figure 9; numerical values for $T = 240$ K are reported in Table 2. As is visible, at high temperatures, we find fairly high values (several times 10^{-4} cm^2/s), typical of the gaseous phase, with a relatively monotonic decrease with temperature. Between 250 and 240 K, a marked fall occurs in the diffusion pattern, and we estimate $D \approx (3-5) \times 10^{-5}$cm^2/s (see Table 2), typical of the liquid phase. Such values are one/two order(s) of magnitude greater than the experimental values of lysozyme solutions, namely, 10×10^{-7} cm^2/s^3. This discrepancy is possibly due to the absence of hindrance effects, that would be taken into account by an explicit treatment of the solvent.

To assess our results for the coexisting densities, we assume an equal critical temperature as in RE case (see section 2) and plausibly a fairly similar binodal. At the lowest temperature investigated in ref 47, namely, 275 K, the coexisting densities are ~50 and ~320 g/L on the vapor and liquid sides, respectively; here, at $T = 240$ K, we have average densities of ~30 and 350 g/L, as can be estimated from the density profile (see Figure 5). These values then appear to be consistent with the RE binodal location and shape, albeit no direct comparison can be made for the mismatch of the investigated temperatures. We note that our isochore path would intercept the RE binodal

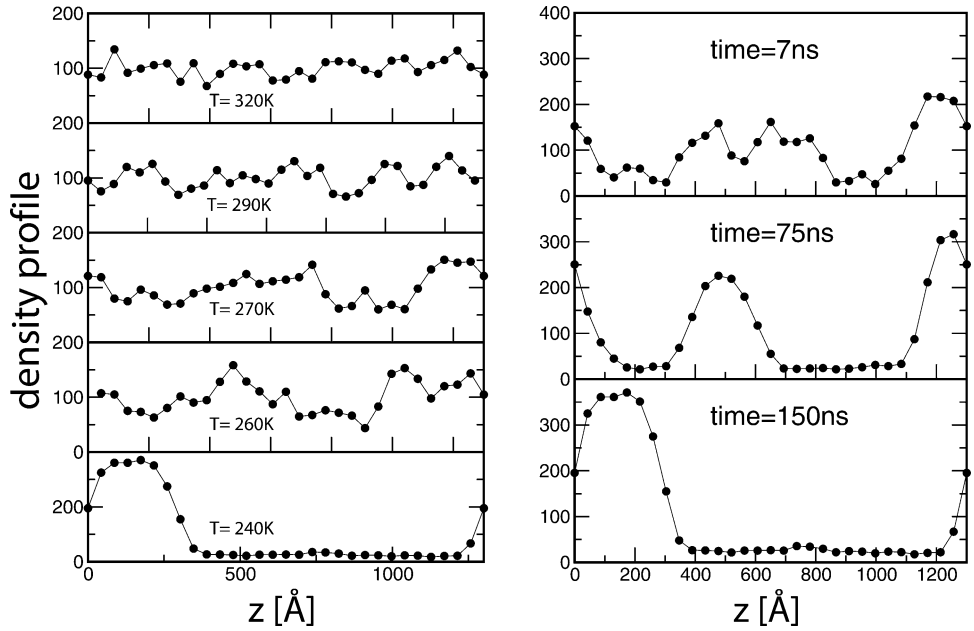

Figure 5. Left: Equilibrium density profile along the z axis of the simulation box at different temperatures. Right: Evolution with time of the density profile at $T = 240$ K.

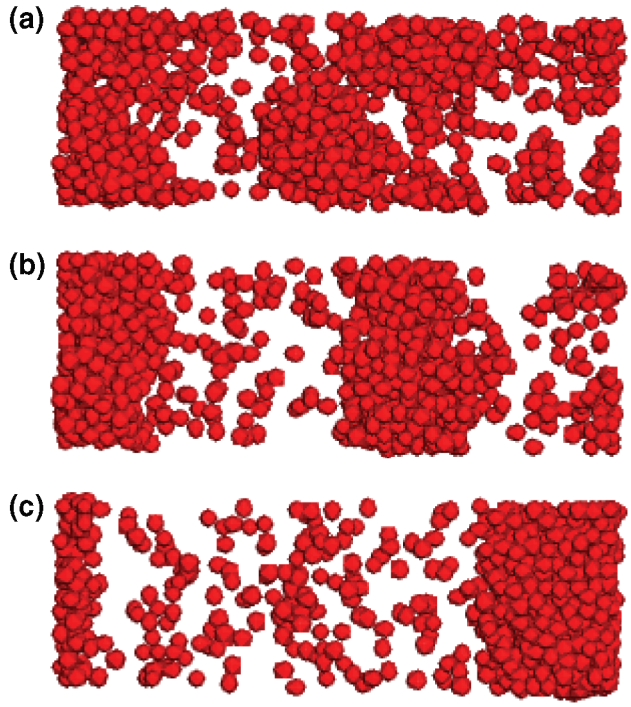

Figure 6. Snapshots of the MD sample at 240 K, taken after (a) 7, (b) 75, and (c) 150 ns.

at ~295 K; thus, our 240 K calculations should correspond to a supercooling of the system by ~45 K, that is, 15% of the binodal temperature. Such a penetration inside the metastable region is necessary to enhance the instability of the system and, so, to overcome well-known free energy barriers to the onset of the phase separation, inherently related to the simulation methodology.[9] The absence of an estimate for the triple-point temperature does not allow us to assess whether, at 240 K, we are below such a temperature and, thus, whether the formed equilibrium is possibly metastable with respect to a vapor−solid transition that might occur on sensibly longer simulation times.

5. Conclusions

We have performed a molecular dynamics (MD) investigation of a one-component, distributed-site model of aqueous lysozyme solution plus added salt. In addition to thermodynamic and structural properties, our study has encompassed the calculation of dynamical properties of such a system, an issue hitherto not addressed for specific bulk protein solutions. We have made reference to a model previously introduced and thoroughly studied by other authors[11,12,47] through Monte Carlo simulations. As more suitable to the MD approach, we have employed a soft-core central potential, rather than the hard-sphere plus attractive short-range interaction conceived by them. We have also devised a strategy to fix the new potential parameters so as to preserve the fit of relevant thermodynamic quantities of the real solutions, as was ensured in the previous model.

We first documented a close reproduction of the radial distribution functions and structure factors of ref 12; this agreement was verified for different values of the protein charge. Then, we examined the evolution of the system along an isochore path by decreasing the temperature from 320 to 240 K, thus bringing the system inside its metastable supercooled region. We found that appreciable density inhomogeneities begin to appear inside the simulation box starting from 270 K. Then, we could detail the onset at 240 K of a well-resolved liquid−vapor phase separation, clearly signaled by thermodynamic and structural functions, by an analysis of dynamical properties and visualized in snapshots of the simulation sample. The coexisting densities appear compatible with previous determinations of the binodal line for the slightly different model of ref 47, in which the lowest temperature investigated was 275 K.

As for our estimates of the diffusion coefficients, they correctly reflect the different dynamical regimes explored by our system. On the other hand, the lack of an explicit solvent description results in a too-fast dynamics, in comparison with experimental data. As is well-known, the inclusion of solvation effects at a microscopic level is currently an open challenge in protein solution description. One could, in principle, imagine extending the present study so as to include the explicit presence

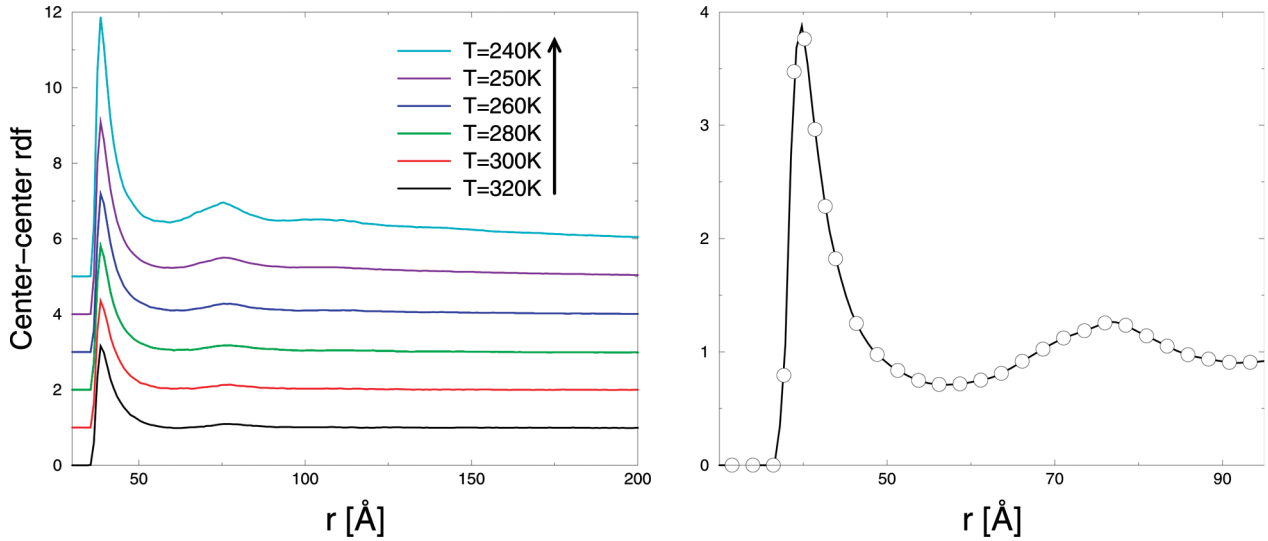

Figure 7. Left: Center−center radial distribution function along the isochore path. The pattern at $T = 240$ K corresponds to the configuration after 7 ns. All curves are progressively shifted by 1 for the clarity sake. Right: $g(r)$ of the dense phase formed after 150 ns at $T = 240$ K.

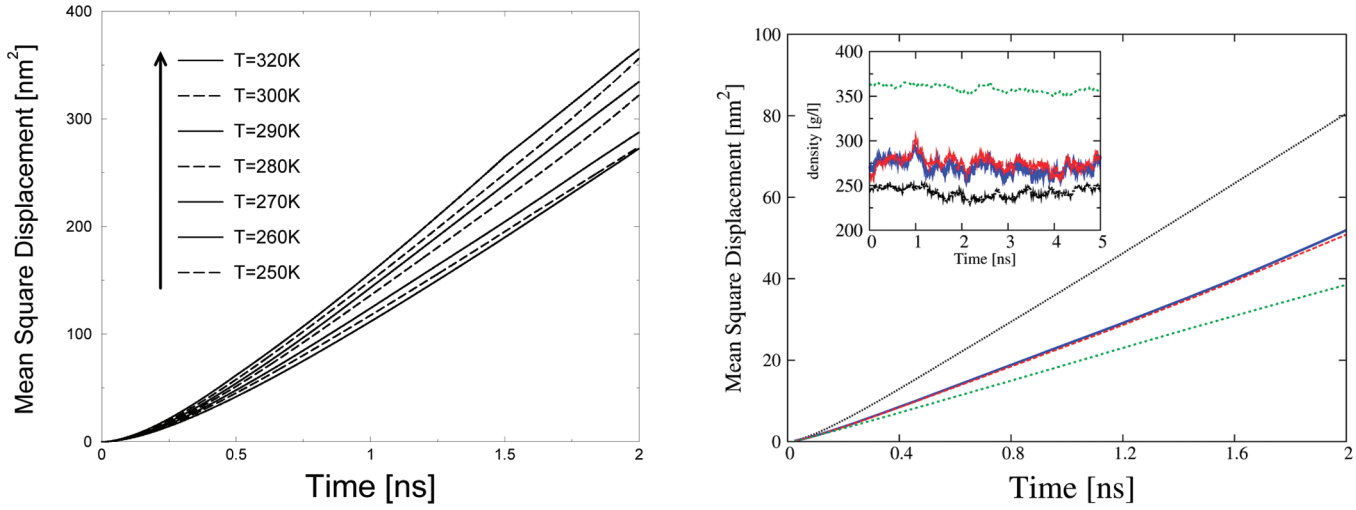

Figure 8. Left: Mean square displacement at various temperatures. Right: MSD and densities (in the inset) in the concentrated regions at 240 K and different times. Black lines: Aggregate on the left of Figure 6b, after 75 ns. Blue and red lines: Aggregate in the center of Figure 6b, after 75 ns (two different samplings; see Table 2). Green lines: Dense slab in Figure 6c after 150 ns.

TABLE 2: Diffusion Coefficients Calculated for Dense-Phase Aggregates at Different Times

localization[a]	box geometry[b]	time window (ns)	D (10^{-5}cm^2/s)
left	$L_x \times L_y \times 6\sigma$	70−75	6.91
central	$L_x \times 7\sigma \times 6\sigma$	70−75	4.31
central	$L_x \times 6\sigma \times 7\sigma$	70−75	4.21
right	$L_x \times L_y \times 6\sigma$	150−155	3.29

[a] Left and central localizations refer to the corresponding positions of aggregates visible in Figure 6b; right refers to Figure 6c. [b] Geometry of the simulation box $\sigma = 37.08$ Å; see text.

of water and salt. However, under ordinary dilution conditions, the number of total sites to consider would rapidly approach 10^5, even for a few tens of proteins involved. Hence, a simulation of the composite systems at a fully molecular level still appears quite demanding. A more realistic perspective seems to be a further implementation of coarse-grained approaches based on effective simplified potentials, optimized through appropriate strategies. In such a context, models of the type envisaged in this work are appealing because they are flexible enough to include, for example, fractional charges, as well as more refined description of hydrophobic effects, and nonspheri-

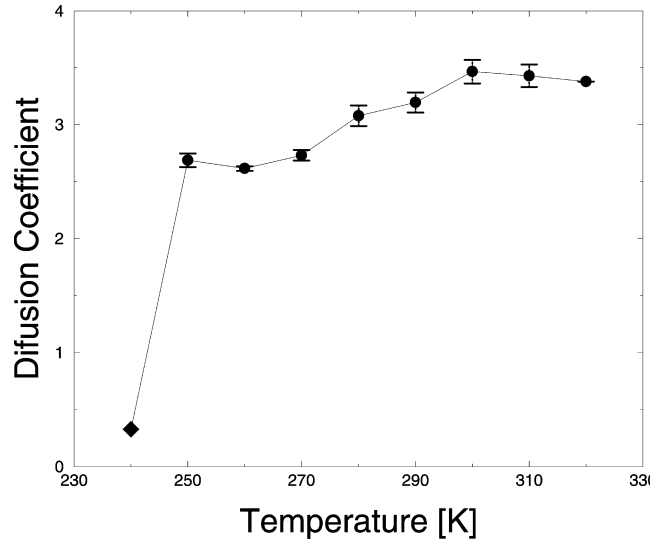

Figure 9. Diffusion coefficient (in units of 10^{-4} cm^2/s) calculated on the basis of the MSD shown in Figure 8. The datum at $T = 240$ K corresponds to the concentrated region at 150 ns (see also Table 2).

cal molecular shapes. We are following a path toward such generalizations as offered by a refined simulation approach recently employed by Pellicane et al.,[62] in which a fully atomistic description of two lysozyme molecules, forming an hydrophobic interface in a discrete molecular solvent, has been undertaken. By averaging over the degrees of freedom of the solvent and of the added electrolyte, one could derive a new effective potential among, for instance, the interaction sites of model proteins and assess other model potentials of current interest.

Acknowledgment. We gratefully acknowledge the computer time and technical support provided by the Centro di Calcolo Elettronico dell'Universita di Messina "A. Villari" and by the PI2S2 Project, managed by the Consorzio COMETA (more information available at http://www.pi2s2.it and http://www.consorzio-cometa.it).

References and Notes

(1) McPherson, A. *Crystallization of Biological Macromolecules*; Cold Spring Harbor Laboratory Press: Woodbury, NY, 1999.

(2) Kelton, K. F. *Crystal Nucleation in Liquids and Glasses* in *Solid State Physics*; Ehrenreich, H., Turnbull, D., Eds.; Academic Press: London, 1991; Vol. 45.

(3) Muschol, M.; Rosenberger, F. Interactions in undersaturated and supersaturated lysozyme solutions: Static and dynamic light scattering results. *J. Chem. Phys.* **1995**, *103*, 10424–10432.

(4) Muschol, M.; Rosenberger, F. Liquid−liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.* **1997**, *107*, 1953–1962.

(5) Wiencek, J. M. New strategies for protein crystal growth. *Annu. Rev. Biomed. Eng.* **1999**, *1*, 505–534.

(6) Kierzek, A. M.; Zielenkiewicz, P. Models of protein crystal growth. *Biophys. Chem.* **2001**, *91*, 1–20.

(7) Sear, R. P. Nucleation: Theory and applications to protein solutions and colloidal suspensions. *J. Phys.: Condens. Matter* **2007**, *19*, 033101/1−28.

(8) ten Wolde, P. R.; Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. *Science* **1997**, *277*, 1975–1978.

(9) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: London, 1996.

(10) Costa, D.; Ballone, P.; Caccamo, C. Kinetics of phase transformations in a model with metastable fluid−fluid separation: A molecular dynamics study. *J. Chem. Phys.* **2002**, *116*, 3327–3338.

(11) Carlsson, F.; Malmsten, M.; Linse, P. Monte Carlo simulations of polyelectrolyte−protein complexation. *J. Phys. Chem. B* **2001**, *105*, 9040–9049.

(12) Carlsson, F.; Malmsten, M.; Linse, P. Monte Carlo Simulations of lysozyme self-association in aqueous solution. *J. Phys. Chem. B* **2001**, *105*, 12189–12195.

(13) Shea, J. M.; Brooks III, C. L. From folding theories to folding proteins: A review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499–535.

(14) Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43–69.

(15) Faisca, P. F. N. The nucleation mechanism of protein folding: A survey of computer simulation studies. *J. Phys.: Condens. Matter* **2009**, *21*, 373102/1–15.

(16) Rosenbaum, D.; Zamora, P. C.; Zukoski, C. F. Phase behavior of small attractive colloidal particles. *Phys. Rev. Lett.* **1996**, *76*, 150–153.

(17) Lomakin, A.; Asherie, N.; Benedek, G. B. Monte Carlo study of phase separation in aqueous protein solutions. *J. Chem. Phys.* **1996**, *104*, 1646–1656.

(18) Malfois, M.; Bonneté, F.; Belloni, L.; Tardieu, A. A model of attractive interactions to account for fluid−fluid phase separation of protein solutions. *J. Chem. Phys.* **1996**, *105*, 3290–3300.

(19) Rosenbaum, D. F.; Kulkarni, A.; Ramakrishnan, S.; Zukoski, C. Protein interactions and phase behavior: Sensitivity to the form of the pair potential. *J. Chem. Phys.* **1999**, *111*, 9882–9890.

(20) Foffi, G.; McCullagh, G. D.; Lawlor, A.; Zaccarelli, E.; Dawson, K. A.; Sciortino, F.; Tartaglia, P.; Pini, D.; Stell, G. Phase equilibria and glass transition in colloidal systems with short-ranged attractive interactions: Application to protein crystallization. *Phys. Rev. E* **2002**, *65*, 031407/1−17.

(21) Zaccarelli, E.; Sciortino, F.; Tartaglia, P.; Foffi, G.; McCullagh, G. D.; Lawlor, A.; Dawson, K. A. Competition between crystallization and glassification for particles with short-ranged attraction. Possible applications to protein crystallization. *Physica A* **2002**, *314*, 539–547.

(22) Petsev, D. N.; Wu, X.; Galkin, O.; Vekilov, P. G. Thermodynamic functions of concentrated protein solutions from phase equilibria. *J. Phys. Chem. B* **2003**, *107*, 3921–3926.

(23) Stradner, A.; Sedgwick, H.; Cardinaux, F.; Poon, W. C. K.; Egelhaaf, S. U.; Schurtenberger, P. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature (London)* **2004**, *432*, 492–495.

(24) Broccio, M.; Costa, D.; Liu, Y.; Chen, S.-H. The structural properties of a two-Yukawa fluid: Simulation and analytical results. *J. Chem. Phys.* **2006**, *124*, 084501/1−9.

(25) Stradner, A.; Foffi, G.; Dorsaz, N.; Thurston, G.; Schurtenberger, P. New insight into cataract formation: Enhanced stability through mutual attraction. *Phys. Rev. Lett.* **2007**, *99*, 198103/1–4.

(26) Cardinaux, F.; Stradner, A.; Schurtenberger, P.; Sciortino, F.; Zaccarelli, E. Modeling equilibrium clusters in lysozyme solutions. *Europhys. Lett.* **2007**, *77*, 48004/1–5.

(27) Pellicane, G.; Costa, D.; Caccamo, C. Theory and simulation of short-range models of globular protein solutions. *J. Phys.: Condens. Matter* **2004**, *16*, S4923–S4936.

(28) Pellicane, G.; Costa, D.; Caccamo, C. Microscopic determination of the phase diagrams of lysozyme and γ-crystallin solutions. *J. Phys. Chem. B* **2004**, *108*, 7538–7541.

(29) Russel, W. B.; Saville, D. A.; Schowalter, W. R. *Colloidal Dispersions*; Cambridge University Press: Cambridge, U.K., 1989.

(30) Weber, G. *Protein Interactions*; Chapman and Hall: New York, 1992.

(31) Janin, J.; Rodier, F. Protein-protein interactions at crystal contacts. *Proteins: Struct., Funct. Genet.* **1995**, *23*, 580–587.

(32) Lomakin, A.; Asherie, N.; Benedek, G. B. Aeolotopic interactions of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9465–9468.

(33) Liu, H.; Kumar, S. K.; Sciortino, F. Vapor-liquid coexistence of patchy models: Relevance to protein phase behavior. *J. Chem. Phys.* **2007**, *127*, 084902/1−5.

(34) Shiryayev, A.; Li, X.; Gunton, J. D. Simple model of sickle hemoglobin. *J. Chem. Phys.* **2006**, *125*, 024902/1−8.

(35) Sear, R. P. Phase behavior of a simple model of globular proteins. *J. Chem. Phys.* **1999**, *111*, 4800–4806.

(36) Hloucha, M.; Lodge, J. F. M.; Lenhoff, A. M.; Sandler, S. I. A patch−antipatch representation of specific protein interactions. *J. Cryst. Growth* **2001**, *232*, 195–203.

(37) Dixit, N. M.; Zukoski, C. F. Crystal nucleation rates for particles experiencing anisotropic interactions. *J. Chem. Phys.* **2002**, *117*, 8540–8550.

(38) Song, X. Role of anisotropic interactions in protein crystallization. *Phys. Rev. E* **2002**, *66*, 011909/1−4.

(39) Allahyarov, E.; Löwen, H.; Hansen, J.-P.; Louis, A. A. Discrete charge patterns, Coulomb correlations and interactions in protein solutions. *Europhys. Lett.* **2002**, *57*, 731–737.

(40) Allahyarov, E.; Löwen, H.; Hansen, J.-P.; Louis, A. A. Non-monotonic variation with salt concentration of the second virial coefficient in protein solutions. *Phys. Rev. E* **2003**, *67*, 051404/1−13.

(41) Kern, N.; Frenkel, D. Fluid−fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J. Chem. Phys.* **2003**, *118*, 9882–9889.

(42) Lund, M.; Jönsson, B. A mesoscopic model for protein−protein interactions in solution. *Biophys. J.* **2003**, *85*, 2940–2947.

(43) Chang, J.; Lenhoff, A. M.; Sandler, S. I. Determination of fluid−solid transitions in model protein solutions using the histogram reweighting method and expanded ensemble simulations. *J. Chem. Phys.* **2004**, *120*, 3003–3014.

(44) Cheung, J. M.; Truskett, T. M. Coarse-grained strategy for modeling protein stability in concentrated solutions. *Biophys. J.* **2005**, *89*, 2372–2384.

(45) Shen, V. K.; Cheung, J. M.; Errington, J. R.; Truskett, T. M. Coarse-grained strategy for modeling protein stability in concentrated solutions. II: Phase behavior. *Biophys. J.* **2006**, *90*, 1949–1960.

(46) Cheung, J. M.; Shen, V. K.; Errington, J. R.; Truskett, T. M. Coarse-grained strategy for modeling protein stability in concentrated solutions. III: Directional protein interactions. *Biophys. J.* **2007**, *92*, 4316–4324.

(47) Rosch, T. W.; Errington, J. R. Investigation of the phase behavior of an embedded charge protein model through molecular simulation. *J. Phys. Chem. B* **2007**, *111*, 12591–12598.

(48) Rosch, T. W.; Errington, J. R. Phase behavior of model confined fluids. Influence of substrate−fluid interaction strength. *J. Phys. Chem. B* **2008**, *112*, 14911–14919.

(49) Li, X.; Gunton, J. D.; Chakrabarti, A. A simple model of directional interactions for proteins. *J. Chem. Phys.* **2009**, *131*, 115101.

(50) Bianchi, E.; Largo, J.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Phase diagram of patchy colloids: Towards empty liquids. *Phys. Rev. Lett.* **2006**, *97*, 168301/1–4.

(51) Russo, J.; Tartaglia, P.; Sciortino, F. Reversible gels of patchy particles: Role of the valence. *J. Chem. Phys.* **2009**, *131*, 014504/1−12.

(52) Chen, J. C.; Kim, A. S. Brownian Dynamics, Molecular Dynamics, and Monte Carlo modeling of colloidal systems. *Adv. Colloid Interface Sci.* **2004**, *112*, 159–173.

(53) McGuffee, S. R.; Elcock, A. H. Atomically detailed simulations of concentrated protein solutions: The effects of salt, pH, point mutations, and protein concentration in simulations of 1000-molecule systems. *J. Am. Chem. Soc.* **2006**, *128*, 12098–12110.

(54) Ramanadham, M.; Sieker, L. C.; Jepsen, L. H. Refinement of triclinic lysozyme: II. The method of stereochemically restrained least squares. *Acta Crystallogr. Sect. B* **1990**, *46*, 63–69.

(55) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.* **1971**, *54*, 5237–5247.

(56) Barker, J. A.; Henderson, D. What is "liquid"? Understanding the states of matter. *Rev. Mod. Phys.* **1976**, *48*, 587–671.

(57) Noro, M. G.; Frenkel, D. Extended corresponding-states behavior for particles with variable range attractions. *J. Chem. Phys.* **2000**, *113*, 2941–2944.

(58) Refson, K. MOLDY: A portable molecular dynamics simulation program for serial and parallel computers. *Comput. Phys. Commun.* **2000**, *126*, 310–329 (information on Moldy available at http://ccpforge.cse.rl.ac.uk/frs/?group_id=34).

(59) Dullweber, A.; Leimkuhler, B.; McLachlan, R. Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.* **1997**, *107*, 5840–5851.

(60) Bond, S. D.; Leimkuhler, B. J.; Laird, B. B. The Nosé−Poincaré method for constant temperature molecular dynamics. *J. Comput. Phys.* **1999**, *151*, 114–134.

(61) Quentrec, B.; Brot, C. New method for searching for neighbors in molecular dynamics computations. *J. Comput. Phys.* **1975**, *13*, 430–432.

(62) Pellicane, G.; Smith, G.; Sarkisov, L. Molecular dynamics characterization of protein crystal contacts in aqueous solutions. *Phys. Rev. Lett.* **2008**, *101*, 248102/1–4.

JP101590Y