

Numerical study of a binary Yukawa model in regimes characteristic of globular proteins in solutions

Achille Giacometti and Domenico Gazzillo

Istituto Nazionale per la Fisica della Materia and Dipartimento di Chimica Fisica, Università di Venezia, S. Marta DD 2137, I-30123 Venezia, Italy

Giorgio Pastore

Dipartimento di Fisica Teorica, Università di Trieste and INFN-DEMOCRITOS, National Simulation Center, Strada Costiera 11, Miramare, I-34100 Trieste, Italy

Tushar Kanti Das*

ICTP, Diploma Course, Strada Costiera 11, Miramare P.O. Box 586, I-34100 Trieste, Italy

(Received 28 June 2004; published 21 March 2005)

The main goal of this paper is to assess the limits of validity, in the regime of low concentration and strong Coulomb coupling (high molecular charges), of a simple perturbative approximation to the radial distribution functions (RDF's), based upon a low-density expansion of the potential of mean force and proposed to describe protein-protein interactions in a recent small-angle-scattering (SAS) experimental study. A highly simplified Yukawa (screened Coulomb) model of monomers and dimers of a charged globular protein (β -lactoglobulin) in solution is considered. We test the accuracy of the RDF approximation, as a necessary complementary part of the previous experimental investigation, by comparison with the fluid structure predicted by approximate integral equations and exact Monte Carlo (MC) simulations. In the MC calculations, an Ewald construction for Yukawa potentials has been used to take into account the long-range part of the interactions in the weakly screened cases. Our results confirm that the perturbative first-order approximation is valid for this system even at strong Coulomb coupling, provided that the screening is not too weak (i.e., for Debye length smaller than monomer radius). A comparison of the MC results with integral equation calculations shows that both the hypernetted-chain (HNC) and Percus-Yevick closures have a satisfactory behavior under these regimes, with the HNC being superior throughout. The relevance of our findings for interpreting SAS results is also discussed.

DOI: 10.1103/PhysRevE.71.031108

PACS number(s): 05.20.Jj, 61.10.Eq, 61.12.Ex

I. INTRODUCTION

In spite of the large effort devoted in the last decades, a clear understanding of the interactions of macromolecules in solution is still far from being achieved [1,2]. In particular, this is true in the case of globular proteins, which share with colloidal systems a number of common properties [3].

From the experimental point of view, there exist several biophysical techniques for obtaining quantitative data on protein-protein interactions under physiologically relevant conditions. Small-angle scattering (SAS), for instance, is currently believed to provide very reliable information, under very different experimental conditions (pH , ionic strength, temperature, etc.). If the particle form factors are known, dividing the SAS intensity by the average form factor yields the experimental average structure factor, which is related to the radial distribution functions (RDF) $g_{ij}(r)$ (i and j are species indexes). A recent experiment [4] reported small-angle x-ray scattering (SAXS) measurements on structural properties of a particular globular protein, the β -lactoglobulin (β LG), in acidic solutions ($pH=2.3$), at sev-

eral values of ionic strength in the range 7–507 mM. For this protein there is a clear evidence of a monomer-dimer equilibrium affected by the ionic strength of the solution [5], and the authors of Ref. [4] were able to achieve a good fit of the experimental data by using a highly simplified “two-component macroion model” (mimicking monomers and dimers of β LG), with effective forces represented by hard-sphere (HS) terms plus the repulsive Yukawa (screened Coulomb) part of the well-known Derjaguin-Landau-Verwey-Overbeek (DLVO) potential [6]. One important novelty of that study, compared with previous ones, is the proposal of a relatively simple, improved approximation to the RDF's, suitable for best-fit programs and not restricted to the particular model but equally well applicable to different spherically symmetric potentials.

From the theoretical point of view, information on intermolecular forces can be extracted from the experimental average structure factor by comparison with a theoretical one, whose calculation requires the choice not only of an interaction model but also of a recipe for deriving the RDF's from the intermolecular potentials. At present, the most accurate techniques for evaluating RDF's are the “exact” computer simulations—Monte Carlo (MC) and molecular dynamics (MD)—and the “approximate” integral equations (IE's) from the statistical mechanical theory of classical fluids [7]. Un-

*Present address: Department of Physics, Shahjalal University of Science & Technology, Sylhet-3100, Bangladesh.

fortunately, such complex methods can hardly be included in a best-fit program for analyzing experimental data. In fact, MC or MD simulations require long computational times and become difficult in regimes characteristic of globular proteins in solution (i.e., low concentration, high charges, asymmetry in size and charge among the components of the mixture). On the other hand, only for a very limited number of simple potentials and within an even more limited number of approximate “closures” do IE’s of liquid theory admit analytical solutions, providing closed-form expressions to be inserted into best-fit codes [8]. In all other cases, an iterative numerical procedure is necessary, and this poses a major drawback to any fitting scheme. Moreover, numerical solution of the IE closures tends to become unstable or does not converge in the region of our interest.

In order to simplify the problem, most analyses of SAS data for highly dilute solutions employ the crude approximation of neglecting all intermolecular forces, assuming either large interparticle separations or weak interactions. In this case, $g_{ij}(r)=1$, the average structure factor equals unity and the SAS intensity depends only upon the average form factor. A common first improvement over the previous choice then corresponds to approximating the RDF’s with their zero-density limit, given by the Boltzmann factor—i.e., $g_{ij}(r)=\exp[-\beta\phi_{ij}(r)]$, where $\phi_{ij}(r)$ is the pair potential and $\beta=(k_B T)^{-1}$ the inverse of the thermal energy (absolute temperature multiplied by Boltzmann’s constant). However, this zero-density approximation becomes insufficient at moderate concentrations or in regimes of colloidal or protein solutions when electrostatic interactions are strong—i.e., at low ionic strength.

Motivated by this scenario, Ref. [4] proposed a more accurate representation of $g_{ij}(r)$ that takes into account, according to a perturbative scheme, terms up to the first order in the density expansion of the potential of mean force, $W_{ij}(r)=-\beta^{-1}\ln g_{ij}(r)$ [9] [note that the $W_{ij}(r)$ expansion should not be confused with the RDF one, since these two expansions differ even at the first order in density]. While the satisfactory best-fit results of Ref. [4] seem to indicate that a first-order approximation to $W_{ij}(r)$ (W1 approximation) is sufficiently accurate for low concentrations such as the experimental conditions under study, there is no way, *a priori*, to tell *where* this approximation breaks down, in the absence of some “exact results” to compare with. On the other hand, as the experimental conditions present in the analysis of Ref. [4] are fairly typical in the context of proteins in solution, we feel that it would be interesting to make such a comparison. Thus the main subject of the present paper, which complements the methodological part of the work of Ref. [4], is not the proposal of a new potential model for β LG, but a test of the W1 approximation against more accurate (MC and IE) structural results.

We perform MC simulations, at constant volume, temperature, and total number of macroparticles, for the *same* HS-Yukawa-DLVO binary model, representing monomers and dimers of β LG, investigated in Ref. [4]. Various values of the screening parameter are considered, and the MC results for $g_{ij}(r)$ are compared with the corresponding ones predicted by the aforesaid W1 approximation as well as by

some commonly used IE’s. In order to ensure always a good accuracy, the MC calculations are carried out with and without a suitable Ewald construction [10–12], which is expected to play a major role in the cases of strong long-range interactions (weak screening). Although the theoretical framework for the Ewald construction, well known for unscreened Coulomb forces, has already been extended to Yukawa interactions in recent Refs. [11,12], this work represents, to the best of our knowledge, the first MC detailed analysis of its implementation and performance for the repulsive Yukawa case [13].

Our calculations allow a rather precise determination of the limits of validity for the W1 expansion. They also show clearly the degree of reliability of some typical IE’s under these frequently encountered, demanding, regimes. It is worthwhile stressing that our results are in fact rather general, as there exists a large variety of physical phenomena which can be described by Yukawa potentials [14]. The existence of “exact” computer simulations for a binary model with these potentials would then prove to be useful within a much more general context than the one treated here.

II. PROTEIN-PROTEIN INTERACTION POTENTIAL

When mesoscopic (colloidal or protein) particles with ionizable surface groups are put into a microscopic polar solvent (like water), most of the charged surface groups dissociate into the solvent and form microscopic *counterions*, usually carrying one or two elementary charges. Consequently, the big particles acquire high charges of opposite sign and are called *macroions* or *polyions*. At equilibrium the counterions are located around the charged macroions, forming an electric double layer. The counterion distribution tends to screen the repulsions between macroions, which have charges of the same sign. The result is a screened Coulomb (Yukawa) repulsion between macroions, which ensures the stability of the solution (charge stabilization) with respect to a possible irreversible flocculation. An important feature of such repulsions is that they can be tuned by adding a suitable amount of a simple electrolyte to the solution. In fact, such a salt provides additional free *microions* (coions, with same charge sign as the macroions, as well as other counterions), which increase the degree of screening and thus reduce the macroion-macroion repulsions [15,17].

A β LG solution thus consists of many components: two different forms of macroions (protein monomers and dimers), counterions, coions, and the solvent. At neutral pH, the structure of the β LG protein is dimeric, while at acidic pH (a condition more similar to the physiological one) a partial dissociation into two monomers takes place. The monomer-dimer equilibrium, which determines the molar fractions of both macroion species, depends upon the ionic strength of the solution. At low ionic strength, the screening is weak and the electrostatic repulsions predominate over the attractive forces responsible for the formation of dimers; as a consequence, most of the macroions are monomers. On the contrary, at high ionic strength a strong screening reduces the monomer-monomer repulsions in such a way that a large fraction of dimers can form.

As in Ref. [4], we represent such a β LG multicomponent solution at a highly simplified, “primitive model,” level of description, using an effective “two-component macroion model,” which takes into account only protein particles [15–17]. In fact, the solvent is regarded as a uniform dielectric continuum, all microions are treated as pointlike particles, and macroions (both monomers and dimers) are assumed to be charged hard spheres, with different diameters. The presence of both solvent and microions appears only in the macroion-macroion *effective* potentials. In the spirit of the DLVO theory [6], we shall then describe the protein-protein interactions with the simple effective potential

$$\phi_{ij}(r) = \phi_{ij}^{\text{HS}}(r) + \phi_{ij}^{\text{Y}}(r) \quad (1)$$

($i, j=1, 2$, with species 1 and 2 corresponding to monomers and dimers, respectively). Here the hard-sphere term accounts for excluded volume effects

$$\phi_{ij}^{\text{HS}}(r) = \begin{cases} +\infty, & 0 < r < \sigma_{ij}, \\ 0, & r > \sigma_{ij}, \end{cases} \quad (2)$$

where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ is the distance of closest approach between two macroparticles of species i and j . On the other hand, the *renormalized* Yukawa term

$$\phi_{ij}^{\text{Y}}(r) = \frac{Z_i Z_j e^2}{\varepsilon(1 + \kappa_D \sigma_j/2)(1 + \kappa_D \sigma_i/2)} \frac{\exp[-\kappa_D(r - \sigma_{ij})]}{r} \quad (3)$$

represents an *effective* screened Coulomb repulsion between two *isolated* macroions in a sea of *microions* and has the same Yukawa form as in the Debye-Hückel theory of electrolytes, but with coupling coefficients of DLVO type [6]. Here, e is the elementary charge, ε the dielectric constant of the solvent, Z_i the valency of species i , and κ_D the inverse Debye screening length due *only* to *microions*, given by

$$\kappa_D = \sqrt{\frac{8\pi\beta e^2}{\varepsilon} \frac{N_A}{1000} (I_c + I_s)}. \quad (4)$$

N_A is the Avogadro number, and $I_c = (1/2)c_c Z_c^2$ denotes the ionic strength of the counterions originated from the ionization of the protein macromolecules (the molar concentration c_c of these counterions is related to the macroion concentrations through the electroneutrality condition, $c_c |Z_c| = c_1 |Z_1| + c_2 |Z_2|$), while $I_s = (1/2)\sum_i c_i^{\text{micro}} (Z_i^{\text{micro}})^2$ is the ionic strength of all microions (cations and anions) generated by added salts. Clearly, κ_D^{-1} depends on temperature and represents an indication of the range of the screened Coulomb interactions, with $\kappa_D \rightarrow 0$ corresponding to pure Coulomb potentials, whereas $\kappa_D \rightarrow \infty$ yields the opposite HS limit. While in real experiments κ_D is fixed by the chemical conditions of the solution (namely, I_c and I_s), in this work we shall not consider I_s as an independent variable, but in view of our methodological purpose, we shall regard $\kappa_D \sigma_1 \equiv \zeta$ as an independent reduced screening parameter.

A measure of the concentration of the two-macroion effective mixture will be given by the volume fraction

$$\eta = \frac{\pi}{6} \sum_{i=1}^2 \rho_i \sigma_i^3, \quad (5)$$

where ρ_i is the partial number density of the i th macroion species (pointlike microions and solvent do not appear here). The definition of the model is then completed by providing one of the two molar fractions $x_i = \rho_i/\rho$ ($i=1, 2$), where $\rho = \sum_i \rho_i$ is the total density.

Now, following partly Ref. [4], we add three remarks about some assumptions involved in the choice of the model potential.

(i) At first glance one might suspect that reducing dimers to equivalent spheres (with a volume twice as large as the monomer); i.e., neglecting the asymmetry of the dimer molecular shape may seem a too drastic simplification. In order to clarify this point, it is to be stressed that in Refs. [4,5] two different levels of description for the dimer were used in the two factors which contribute to the SAS intensity. The coherent scattering intensity $I(q)$ was written as

$$I(q) \propto \sum_{i,j} (\rho_i \rho_j)^{1/2} F_i^*(q) F_j(q) S_{ij}(q), \quad (6)$$

where q is the magnitude of the scattering vector, $F_i(q)$ the angular average of the form factor of species i , and the Ashcroft-Langreth partial structure factors (for spherically symmetric intermolecular potentials) are defined by

$$S_{ij}(q) = \delta_{ij} + 4\pi(\rho_i \rho_j)^{1/2} \int_0^\infty r^2 h_{ij}(r) \frac{\sin(qr)}{qr} dr \quad (7)$$

in terms of the three-dimensional Fourier transform of $h_{ij}(r) = g_{ij}(r) - 1$. A very accurate procedure was used to calculate numerically both macroion form factors $F_1(q)$ and $F_2(q)$ from crystallographic data, taking into account, in particular, the *exact* elongated shape and structure of the dimer—i.e., its distribution of scattering matter [4,5]. Thus the approximation of spherical dimers was restricted only to the calculation of $S_{ij}(q)$, which is related, through $g_{ij}(r)$, to the intermolecular potentials. At low protein concentrations, the choice of spherically symmetric hard-core potentials can indeed be justified. As in such regimes the average distance among particles is large, intermolecular forces are dominated by the long-range electrostatic interactions, whereas the details of the short-range repulsions (i.e., the excluded volume effects) are irrelevant.

(ii) Our potentials are purely repulsive. We have not included the attractive van der Waals part of the DLVO potential for charged colloidal suspensions (the so-called Hamaker term [6]), as it has already been shown to be unnecessary for this system in previous work [4]. The basic reason is that van der Waals attractions may be fully masked by the electrostatic repulsions when the latter are strong and are also negligible for moderately charged particles with a diameter smaller than 50 nm [16]. Moreover, the Hamaker term diverges at contact, so that, to circumvent this singularity, the inclusion of the attractive term would require the addition of a Stern layer of counterions (with finite size) condensed on the macroion surface [5].

(iii) Given that the specific protein forms dimers, it appears that the β LG necessarily has a short-range monomer-monomer attraction (related to the surface groups), which causes the aggregation into dimers and determines the monomer molar fraction x_1 . One expects this attractive term (possibly including hydrogen bonding) to be rather complex and non spherically symmetric. If such a contribution were clearly understood and easily tractable, one could start from a more fundamental viewpoint, choosing a model which considers only monomers and includes the aforementioned attraction into their pair potential. One could then monitor the dimerization fraction within this *one-component* system. However, this analysis may be a project on its own right and goes beyond the aims of the present study.

More simply, in order to avoid poorly known and angular-dependent potentials, the authors of Refs. [4,5] adopted the viewpoint of using a *binary* (monomer-dimer) rather than a one-component model, and the required attraction was accounted only indirectly, by using a chemical association equilibrium to evaluate x_1 [4,5].

While the dependence of x_1 upon the added salt (i.e., upon I_s) must be taken into account in any best-fit analysis with the binary model [4,5], in the present work, for the sake of simplicity, we shall consider x_1 as an independent parameter. Most of our calculations will be performed at equal molar fractions $x_1=x_2$, but in the last part of the paper we shall also address the effect of changing the molar fractions.

III. LOW-DENSITY EXPANSION OF THE MEAN-FORCE POTENTIAL

As discussed in the Introduction, one of the most commonly used procedures to compute RDF's $g_{ij}(r)$ for a given pair potential $\phi_{ij}(r)$ goes through the solution of the Ornstein-Zernike (OZ) IE's from the liquid state theory, within some approximate closure relation. This can typically be done only numerically, with the exception of few simple cases (for some potentials and peculiar closures), where the solution can be worked out analytically [7].

Note that, for HS-Yukawa potentials, the OZ equations do admit analytical solution [18–20], within the so-called “mean spherical approximation” (MSA), to be discussed further below. On the other hand, under the experimental regime which we are interested in [4]—namely, low density and strong electrostatic repulsions (weak screening)—the MSA is well known to display a serious drawback since RDF's may assume unphysical negative values close to contact distance σ_{ij} , for particles i and j which repel each other. To overcome this shortcoming for repulsive Yukawa models, it would be possible to utilize an analytical “rescaled MSA” [16,21,22] (this possibility will not be investigated in the present paper) or to resort to different closures.

In general, then, only numerical solutions are feasible, and thus IE algorithms can hardly be included into best-fit programs for the analysis of SAS results.

The use of analytical solutions or simple approximations requiring only a minor computational effort is clearly much more advantageous when fitting experimental data. This can be done by resorting to the exact, albeit formal, relation

$$g_{ij}(r) = \exp[-\beta W_{ij}(r)], \quad (8)$$

$$-\beta W_{ij}(r) = -\beta \phi_{ij}(r) + \omega_{ij}(r), \quad (9)$$

where $W_{ij}(r)$ is the potential of mean force, which includes the direct pair potential $\phi_{ij}(r)$ as well as $-\beta^{-1}\omega_{ij}(r)$ —i.e., the indirect interaction between i and j due to their interactions with all remaining macroparticles of the fluid. In the density expansion of $W_{ij}(r)$,

$$-\beta W_{ij}(r) = \ln g_{ij}(r) = -\beta \phi_{ij}(r) + \omega_{ij}^{(1)}(r)\rho + \omega_{ij}^{(2)}(r)\rho^2 + \dots, \quad (10)$$

the *exact* power coefficients $\omega_{ij}^{(k)}(r)$ ($k=1, 2, \dots$) can be computed by using standard diagrammatic techniques [9], which yield the results, in terms of multidimensional integrals of products of Mayer functions,

$$f_{ij}(r) = \exp[-\beta \phi_{ij}(r)] - 1. \quad (11)$$

In the zero-density limit, $\omega_{ij}(r)$ vanishes and $g_{ij}(r)$ reduces to the Boltzmann factor—i.e.,

$$g_{ij}(r) = \exp[-\beta \phi_{ij}(r)] \quad \text{as } \rho \rightarrow 0, \quad (12)$$

which represents a zeroth-order (W0) approximation, frequently used in the analysis of experimental scattering data. The W0 approximation avoids the problem of solving the OZ equations, but is largely inaccurate except, perhaps, at extremely low densities. We then consider the first-order perturbative correction (W1 approximation) [4]

$$g_{ij}(r) = \exp[-\beta \phi_{ij}(r) + \omega_{ij}^{(1)}(r)\rho]. \quad (13)$$

By construction, this expression is never negative, thus overcoming the major drawback of MSA. The explicit expression of $\omega_{ij}^{(1)}(r)$ reads

$$\omega_{ij}^{(1)}(r) = \sum_k x_k \gamma_{ij,k}^{(1)}(r) = \sum_k x_k \int d\mathbf{r}' f_{ik}(r') f_{kj}(|\mathbf{r} - \mathbf{r}'|). \quad (14)$$

The evaluation of the convolution integral $\gamma_{ij,k}^{(1)}(r)$ is most easily carried out in bipolar coordinates. After an integration over angle variables $\gamma_{ij,k}^{(1)}(r)$ reduces to

$$\gamma_{ij,k}^{(1)}(r) = \frac{2\pi}{r} \int_0^\infty dx [x f_{ik}(x)] \int_{|x-r|}^{x+r} dy [y f_{kj}(y)]. \quad (15)$$

Of course, the use of the W1 approximation is not restricted to the model of this paper, and the proposed calculation scheme can be equally well applied to different spherically symmetric potentials. While it was shown in Ref. [4] how this first-order correction largely improves the fit of experimental scattering data, over the W0 one and under those experimental conditions, little could be said on the limits of validity of the W1 approximation with respect to an (hypothetical) exact calculation. This is the reason why we tackle this task here by a comparison with MC simulations for a binary HS-Yukawa-DLVO system.

IV. MC SIMULATIONS AND EWALD SUM FOR YUKAWA FLUIDS

The difficulties involved in Monte Carlo calculations dealing with pure Coulomb potentials are well known [23]. It is now widely appreciated the usefulness of the so-called Ewald sum for long-range electrostatic interactions [10,23]. On the other hand, a similar construction for Yukawa potentials has appeared in the literature quite recently [11,12]. We now briefly recall the procedure detailed in Refs. [11,12]. In order to keep notation as simple as possible, we shall restrict ourselves to simple Yukawa potentials, the extension to our actual potential [Eqs. (1)–(3)] being obvious. The basic idea is to start with the total potential energy

$$U = \frac{1}{2} \sum_{\alpha\beta=1}^N q_\alpha q_\beta \frac{e^{-\kappa_D r_{\alpha\beta}}}{r_{\alpha\beta}}, \quad (16)$$

where N is the total number of macroparticles, $r_{\alpha\beta} = |\mathbf{r}_\alpha - \mathbf{r}_\beta|$, and $q_\alpha = Z_\alpha e$, $q_\beta = Z_\beta e$ are the charges. This term is then split into a sum of two contributions, one evaluated in real space, while the other is calculated in momentum space on wave vectors given by $\mathbf{k} = 2\pi\mathbf{n}/V^{1/3}$ (V is the volume of the system and \mathbf{n} a unit vector of integer components). To this aim an auxiliary continuous Gaussian charge distribution

$$\rho_q(r) = \left(\frac{\lambda^2}{\pi}\right)^{3/2} e^{-\lambda^2 r^2} \quad (17)$$

is exploited. For λ values such that the real-space contribution is limited to particles in the basic simulation cell, the final result reads

$$\begin{aligned} U = & \frac{1}{2} \sum_{\alpha\beta=1}^N q_\alpha q_\beta \\ & \times \frac{\text{erfc}(\lambda r_{\alpha\beta} + \kappa_D/2\lambda) e^{\kappa_D r_{\alpha\beta}} + \text{erfc}(\lambda r_{\alpha\beta} - \kappa_D/2\lambda) e^{-\kappa_D r_{\alpha\beta}}}{2r_{\alpha\beta}} \\ & + \sum_{\alpha\beta=1}^N \frac{1}{V} \sum_{\mathbf{k}} q_\alpha q_\beta \frac{4\pi}{k^2 + \kappa_D^2} \exp\left(\frac{-(k^2 + \kappa_D^2)}{4\lambda^2}\right) \\ & \times \cos(\mathbf{k}_{\alpha\beta} \cdot \mathbf{r}_{\alpha\beta}) + \sum_{\alpha} q_\alpha^2 \left[-\frac{2\lambda}{\sqrt{\pi}} \exp\left(\frac{-\kappa_D^2}{4\lambda^2}\right) \right. \\ & \left. + \kappa_D \text{erfc}\left(\frac{\kappa_D}{2\lambda}\right) \right], \quad (18) \end{aligned}$$

where in the first sum we exclude the terms with equal indices and we have introduced the complementary error function

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} dz e^{-z^2}. \quad (19)$$

The first two terms in Eq. (18) represent the real- and momentum-space summations, respectively, while the last two contributions refer to the self-energy [11,12]. In the limit $\kappa_D \rightarrow 0$, the above equation reduces to the Coulomb case [23], as it should. Equation (18) contains λ as an adjustable parameter, and we have performed a detailed analysis for its

optimal choice, so that the original potential (16) is recovered for the range of κ_D values of interest under the experimental conditions of Ref. [4], without using too many terms in the reciprocal-space summation. Our results indicate $\lambda \sim 6.5/L$ (with L being the side length of the cubic simulation box) to be the optimal choice, which is of the same order of magnitude of the one typically used in the Coulomb case.

V. INTEGRAL EQUATIONS

Our next task is to test the performance of some IE's under the experimental conditions of Ref. [4]. This will strengthen the usefulness of the W1 approximation, in view of its simplicity compared to a typical IE calculation for a binary mixture. The OZ IE's of the liquid state theory for p -component mixtures with spherically symmetric interactions read [7]

$$h_{ij}(r) = c_{ij}(r) + \rho \sum_{l=1}^p x_l \int d\mathbf{r}' c_{il}(r') h_{lj}(|\mathbf{r} - \mathbf{r}'|), \quad (20)$$

and their solution can be accomplished only in the presence of an additional approximate relation (closure) between the direct correlation function (DCF) $c_{ij}(r)$ and the total correlation function $h_{ij}(r) = g_{ij}(r) - 1$ ($p=2$ in the present case). The most known among these approximations are [7] (1) the Percus-Yevick (PY) closure

$$c_{ij}(r) = [e^{-\beta\phi_{ij}(r)} - 1][1 + \gamma_{ij}(r)], \quad (21)$$

where $\gamma_{ij}(r) = h_{ij}(r) - c_{ij}(r)$, (2) the hypernetted chain (HNC) closure

$$c_{ij}(r) = e^{-\beta\phi_{ij}(r) + \gamma_{ij}(r)} - 1 - \gamma_{ij}(r), \quad (22)$$

and (3) the MSA, much simpler than the above two, with the DCF being related only to the potential outside the core

$$c_{ij}(r) = -\beta\phi_{ij}(r) \quad r \geq \sigma_{ij}, \quad (23)$$

complemented by the condition of excluded volume, $g_{ij}(r) = 0$ inside the hard cores.

Other possible more refined closures, which can be regarded as a combination of the above three, will be also briefly addressed in this work.

VI. NUMERICAL RESULTS

A β LG monomer is composed of 162 amino acid residues; 20 of these are basic, so that at $pH=2.3$ the monomer is expected to be positively charged, with about 20 proton charges. In our calculations we fix all parameters close to their best-fit ‘‘experimental’’ values [4], $\sigma_1=40 \text{ \AA}$, $\sigma_2=2^{1/3}\sigma_1 \approx 50.40 \text{ \AA}$, $Z_1=20$, $Z_2=40$, $T=298.15 \text{ K}$, and $\varepsilon=78.5$ (strictly speaking, in Ref. [4], $T=293.15$, $\sigma_1=38.30 \text{ \AA}$, and the ratio Z_2/Z_1 was about 1.8, since 2 of the 20 amino acids of the monomer are at the monomer-monomer interface in the dimer).

The packing fraction $\eta=0.01$ is also very close to that determined from the experimental protein concentration ($\eta=0.0096$) [4]. We then vary the dimensionless screening parameter $\zeta = \kappa_D \sigma_1$ in the range $\zeta \sim 1-10$, roughly equivalent

to the range of ionic strength I_s (from 7 to 507 mM) examined in the aforesaid SAS measurements for β LG [4,5] [where $\zeta=1.41$ when $I_s=7$ mM (weak screening, monomer molar fraction $x_1=0.85$) and $\zeta=9.08$ when $I_s=507$ mM (strong screening, $x_1=0.05$)]. Note that an increase of ζ has the effect of reducing not only the range of the HS-Yukawa-DLVO potentials but also their amplitudes, as described by Eq. (3).

In order to obtain the W1 approximation to the RDF's, we have evaluated all the convolution terms $\gamma_{ij,k}^{(1)}(r)$, given by Eq. (15), at the grid points $r_i=i\Delta r$ ($i=1, \dots, 500$), with $\Delta r=1$ Å. At each r_i value, the double integral, Eq. (15), has been carried out numerically by using the trapezoidal rule for both x and y integrations. For the x integration, we have chosen as upper limit the value $x_{\max}=\max(x_{\text{cut}}, \sigma_2+r)$, with $x_{\text{cut}}=\sigma_2+12/\kappa_D$, and as grid size $\Delta x=x_{\text{cut}}/400$. For the y integration, $\Delta y=\Delta x$.

The MC simulations have been performed at constant N, V, T , with and without the Ewald procedure for a correct treatment of the long-range electrostatic interactions. Most calculations refer to a total number of particles $N=216$, divided in monomers and dimers according to the fixed monomer molar fraction x_1 . Although the sample size may seem rather small with respect to present-day standards, one has to take into account that the Ewald construction takes a great computational effort with increasing N . In any case, we have carried out some additional calculations with a larger number of particles in order to check for possible finite-size effects and found no significant differences in the results. Hence, we shall use this value of N throughout, with one exception which will be described later on. The simulation starts from an appropriate lattice distribution of molecules. We have typically employed 10^5 equilibration steps to eliminate any memory of the initial configuration artificially introduced into the fluid. Then 5×10^5 additional steps have been used to collect sufficient information for the statistical averages required to calculate the RDF's.

With the same parameters we have also solved the OZ integral equations numerically, by means of an efficient algorithm proposed by Labik *et al.* [24] employing 1024 grid points, with a mesh size $\Delta r=0.01\sigma_1$, and 20 basis functions. The PY, HNC, and MSA closures have been employed. As expected, the MSA results (not shown in our figures) poorly describe the MC data and exhibit the above-mentioned drawbacks of the MSA closure in regimes with strong coupling at high dilution [16]. We have explicitly checked that other, more sophisticated, approximations, such as the Rogers-Young (RY) closure [25] or the Zerah-Hansen (HMSA) one [26], which attempt to achieve thermodynamic consistency of compressibility and virial pressures by interpolating between two of the above closures (PY-HNC and MSA-HNC, respectively), are of no use here and such a thermodynamic consistency is never achieved, presumably because of the combined effect of low densities and strong long-range repulsions [27].

Finally, both MC and IE calculations for $g_{ij}(r)$ have been compared with the corresponding results from the first-order W1 approximation, with the aim to assess the limits of validity where the expression given by Eq. (13) can be safely

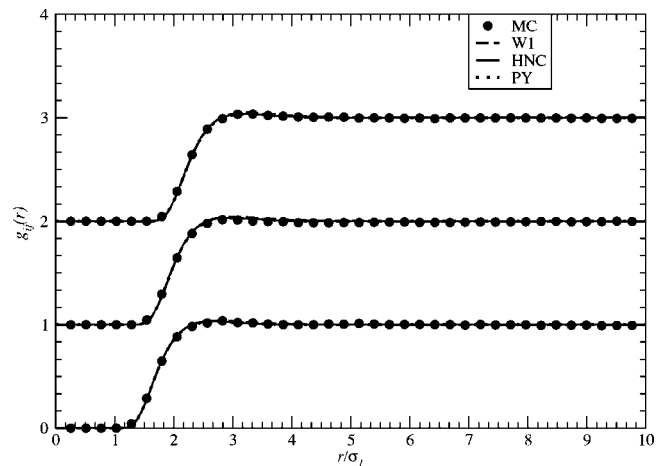


FIG. 1. Partial correlation functions $g_{11}(r)$, $g_{12}(r)$, and $g_{22}(r)$ (in order from bottom to top) as a function of the rescaled distance r/σ_{ij} for $\zeta=3$ and $x_1=0.5$. Circles correspond to MC calculations, solid lines to HNC, dotted lines to PY, and dashed lines to the first-order W1 approximation. Here and in the following the components 12 and 22 have been shifted upwards by one and two units, respectively.

exploited, under conditions typical of proteins in solution. As further elaborated below, we find that for values $\zeta \gtrsim 2$ (i.e., $\kappa_D^{-1} \lesssim \sigma_1/2$) the W1 approximation well describes the behavior of the RDF's.

When ζ is large (in the range $\zeta \sim 5-10$) the Yukawa interactions are strongly screened, and the RDF's essentially reduce to the typical HS ones, with the first maximum corresponding to the contact distance σ_{ij} .

Figure 1 depicts the comparison between the MC results and the W1 approximation for $\zeta=3$ (corresponding to a moderately weak screening) and $x_1=0.5$, which is when both monomers and dimers are present in equal measure. Note that these conditions are close to one of the experimental cases reported in Ref. [4], where $I_s=47$ mM corresponds to $\zeta=2.8$ and $x_1=0.48$. On the other hand, as the ionic strength I_s is lowered from 507 mM to 7 mM, the experimental system switches from a fluid almost completely made up of dimers ($x_1=0.05$) to one almost completely made up of monomers ($x_1=0.85$). This rather peculiar feature is specific to the β LG and will also be considered further on. Here, however, our main aim is to test the W1 approximation under the simple, symmetric, condition of equal molar fractions, since we already know, from Ref. [4], that the first-order approximation well describes the β LG experimental data, which display, in particular, a lowering in the scattering intensity at small angles, with a progressive development of an interference peak at low ionic strengths. In Fig. 1 we also report the results from the HNC and PY IE's (solid and dotted lines), which are practically indistinguishable on the employed scale. It is apparent that in the case of Fig. 1 the W1-RDF's $g_{11}(r)$, $g_{12}(r)$, and $g_{22}(r)$ are in excellent agreement with their MC, HNC, and PY counterparts. Note that, for all three RDF's, $g_{ij}(r)$ remains zero even in a region outside the hard core, while the position of the peak lies at a distance larger than σ_{ij} , as a consequence of the strong Yukawa repulsions.

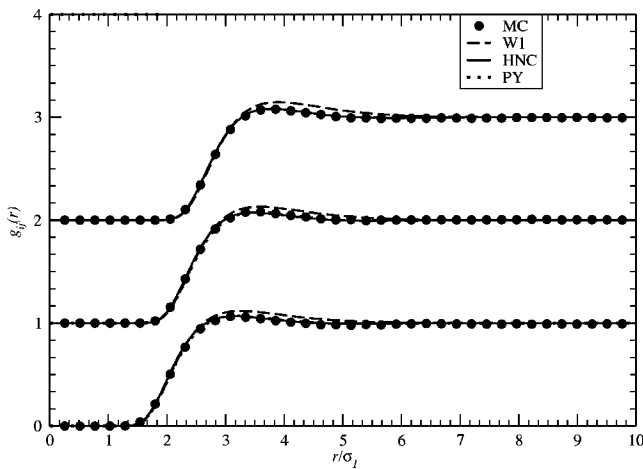


FIG. 2. Same as above with $\zeta=2$ and $x_1=0.5$.

A departure of the first-order W1 approximation from the MC results can be observed for smaller values of the screening parameter ζ , where higher-order terms in the density expansion of $W_{ij}(r)$, Eq. (10), begin to have a non-negligible effect. This is indicated in Fig. 2 for the case $\zeta=2$, which corresponds to $\kappa_D^{-1}=\sigma_1/2$, with the Debye screening length being equal to the monomer radius (among the experimental data of Ref. [4] we find $\zeta=2$ and $x_1=0.73$ when $I_s=17$ mM). Again the HNC and PY RDF's are nearly identical with each other and with MC data. On the other hand, the W1 approximation predicts peaks nearly at the same positions as the PY and HNC closures, while its peak heights are slightly overestimated. However, the agreement between W1 and MC results can still be regarded as rather good.

In regimes with weaker screening the discrepancies become more and more pronounced. The breakdown of all the considered approximations can be clearly appreciated in Fig. 3 for $\zeta=1$ (note that the case with the weakest screening in Ref. [4] corresponds to $I_s=7$ mM, $\zeta=1.41$, and $x_1=0.85$). The W1 results are not reported in this figure, since they are

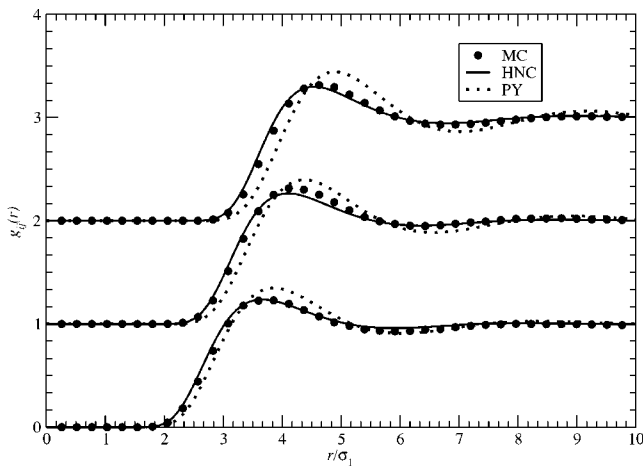


FIG. 3. Comparison of the results from HNC, PY, and MC in the calculation of the partial radial distributions functions for $\zeta=1$ and $x_1=0.5$. The first-order W1 approximation is not depicted as it overshoots the MC results roughly by a factor of 2.

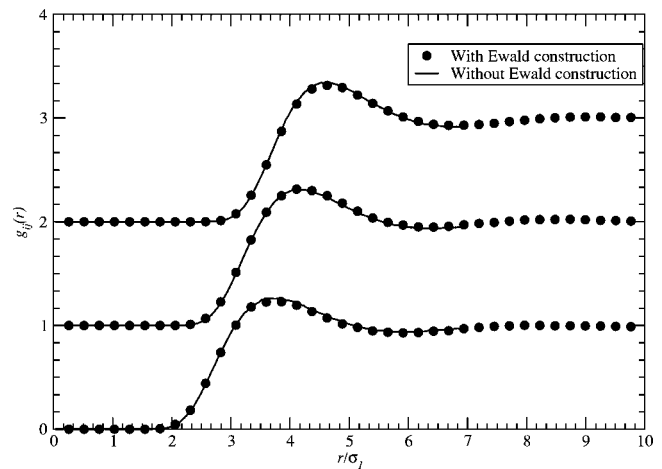


FIG. 4. Partial correlation functions $g_{11}(r)$, $g_{12}(r)$, and $g_{22}(r)$ (in order from bottom to top) as a function of the rescaled distance r/σ_1 , as computed with (circles) and without (solid line) the Ewald construction, for $\zeta=1$ and $x_1=0.5$.

way off from the MC data (with an overestimation of about a factor of 2). On the other hand, even the results from the PY approximation are significantly displaced from the MC RDF's. The difference between the HNC and PY results is apparent, particularly for the latter, as expected. The PY approximation overestimates both the heights and positions of the peaks, compared to the HNC ones. Overall the PY approximation fails to describe the MC calculation for $\zeta < 2$, whereas the HNC closure is consistently in good agreement with the MC data. Such a good performance of the HNC closure closely resembles the good agreement between HNC and MC, even at strong Coulomb coupling, for the *one-component* fluid of *point* charges (electron gas or plasma, with $\zeta=0$) (OCP) in a uniform neutralizing background [28]. However, the results for our binary model with screening at packing fraction $\eta=0.01$ can hardly be compared with the available MC simulations for *one-component* charged hard spheres (OCCS, with $\zeta=0$) in a uniform neutralizing background, at $\eta=0.3-0.4$ [29]. Moreover, it is known the inadequacy of the HNC for high charges at low concentrations (for instance, in the dilute regime of 2-2 aqueous electrolytes [30,31], where bridge diagrams become non-negligible for like-charge RDF's). On the other hand, despite the large number of comparisons among PY, HNC, and MC predictions carried out over the years, we are not aware of a similar detailed RDF investigation under regimes characteristic of globular proteins in solution, for HS-Yukawa-DLVO binary models.

Next we consider the effect of taking into proper account the long-range nature of the interactions (in the weakly screened case) with the use of the Ewald construction. This is illustrated in Figs. 4 and 5, where the RDF's computed with and without the Ewald construction are compared at $\zeta=1$ and $\zeta=0.25$, respectively. Clearly, very little difference is detected between these two calculations when $\zeta=1$ (and when $\zeta=0.5$, not shown). We find that the presence of the Ewald construction begins to be important for very low values of the screening parameter ($\zeta \lesssim 0.25$ —i.e., $\kappa_D^{-1} \gtrsim 4\sigma_1$), as shown in Fig. 5. Supplementary calculations, not reported

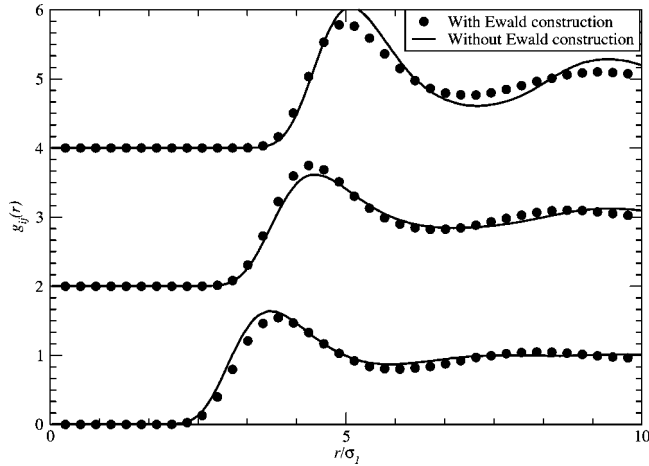


FIG. 5. Same as above with $\zeta=0.25$ and $x_1=0.5$. Note that the scale has been changed with respect to previous figures. Accordingly, here components 12 and 22 have been shifted upward by two and four units, respectively.

here, confirm that this is true even for lower values of protein charges—that is, for weaker Coulomb coupling.

Finally, we consider the effect of varying the molar fractions. While the exact conditions reported in the β LG experiment pose a very difficult challenge to an accurate MC calculation in view of the particular combination of strong asymmetry and repulsions, we can nevertheless easily account for the general trend. This is depicted in Fig. 6, where we have assumed $\zeta=2$ and $x_1=0.75$, in closer analogy with a β LG experimental case, $\zeta=2$ and $x_1=0.73$, when $I_s=17$ mM. It is apparent how the performance of the first-order W1 approximation is comparable to the corresponding symmetric case, $\zeta=2$ and $x_1=0.5$.

Figure 7 refers to the asymmetric case with the weakest screening in Ref. [4]—i.e., $\zeta=1.41$ and $x_1=0.85$ (corresponding to the lowest value of ionic strength, $I_s=7$ mM). Again, the HNC and PY results are in good agreement with the MC ones, and even the performance of the W1 approximation can be regarded as acceptable, in agreement with the results of Ref. [4]. We note that, in view of the low molar fraction of

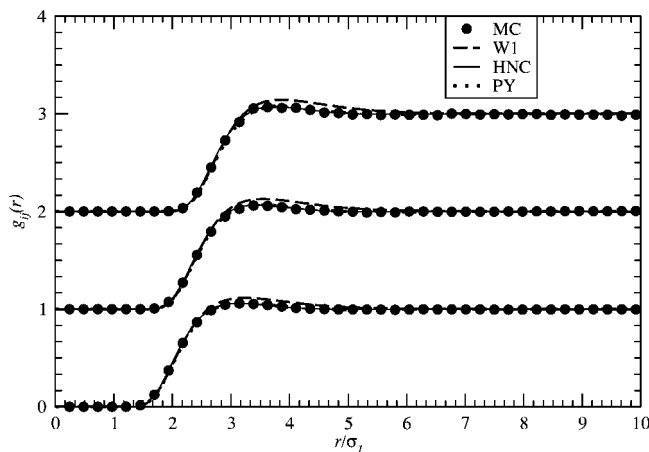


FIG. 6. Asymmetric case, corresponding to Fig. 2, with $\zeta=2$ and $x_1=0.75$.

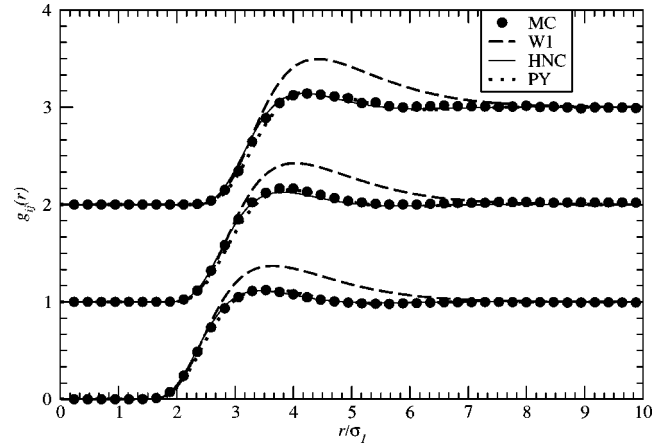


FIG. 7. Asymmetric case with $\zeta=1.41$ and $x_1=0.85$, corresponding to the lowest value of ionic strength $I_s=7$ mM (i.e., the weakest screening) investigated in Ref. [4].

species 2 (dimers), the results of Fig. 7 refer to a higher number of particles ($N=512$).

VII. CONCLUSIVE REMARKS

This work represents a necessary verification of the best-fit analysis of SAS experimental data, for solutions of β -lactoglobulin, presented in Ref. [4]. In the present paper we have assessed the limits of validity of the W1 approximation, exploited in that work to calculate, in a simple way, the RDF's in regimes typical of a large class of globular proteins in solution—that is, low concentrations and high macroion charges. This task has been accomplished by considering the *same* highly simplified model proposed in Ref. [4] (i.e., a *binary* mixture of monomers and dimers of the protein, with HS-Yukawa-DLVO effective potentials) and comparing the corresponding $g_{ij}(r)$ obtained by three different methods: the first-order density expansion of the potential of mean force (W1 approximation), “exact” MC simulations, and approximate IE's. All results reported here refer to $\eta=0.01$ and high macroion charges, $Z_1=20$ and $Z_2=40$. For the MC simulations we have implemented an Ewald construction for Yukawa potentials, which ensures a proper treatment of the long-range part of the interactions, and we have tested its relevance as a function of the screening parameter ζ . In the IE calculations simple closures (PY, HNC, and MSA) as well as more elaborated ones (RY and HMSA) have been considered.

We can summarize the obtained results as follows.

(i) The first-order W1 approximation can be considered reliable in regimes with low concentration ($\eta=0.01$) even for strong Coulomb coupling [up to charges of $(10-20)e$ on macroions with diameters of $40-50$ Å], provided that the screening is strong enough—i.e., when $\zeta \geq 2$ or, equivalently, $\kappa_D^{-1} \leq \sigma_1/2$ (Debye length smaller than monomer radius). This finding demonstrates that the previous usage of the W1 approximation in Ref. [4] was fully legitimate, for all considered cases including those with the lowest ionic strength ($I_s=7$ mM, $x_1=0.85$, $\zeta=1.41$), which lies near the borderline

of the reliability region. For weaker screening (lower values of ζ or larger κ_D^{-1}) at least second-order terms in the density expansion should be taken into account. However, the resulting W2 approximation would require a much higher computational effort and thus could not be conveniently included into a best-fit program for analyzing SAS experimental data.

(ii) In the MC simulations the Ewald construction for Yukawa potentials starts to be important for weak screening corresponding to $\zeta \lesssim 0.25$ ($\kappa_D^{-1} \gtrsim 4\sigma_1$), and this is true even for lower values of the protein charges.

(iii) Both the HNC and PY IE's yield sufficiently accurate values of the RDF's, as long as $\zeta \gtrsim 2$. For lower values of ζ HNC is still accurate, whereas PY starts to deviate as expected. The MSA predictions, on the other hand, are very poor even in those regimes where the W1 approximation can be considered reliable. Under these conditions both the RY and HMSA closures are found not to achieve thermodynamic consistency between compressibility and virial pressures.

(iv) The sufficient accuracy of the W1 approximation in the regimes of our interest (tested in this paper against "exact" MC results), together with its success (shown in Ref. [4]) in reproducing the main features of the experimental SAS intensity curves for the examined β LG solutions, confirms the good performance of the highly idealized two-macroion model, which includes spherically symmetric HS-Yukawa-DLVO repulsions, a monomer-dimer chemical equilibrium, and the "exact" form factors, evaluated by taking into account the real nonspherical structure of the dimer.

Clearly, all complex characteristics of the interactions between globular proteins cannot be explained by the "primitive" level of description adopted in Ref. [4] and here. We have followed the generally accepted philosophy of exploiting the simplest possible description of the system, which yet can provide useful information on the basic underlying interaction mechanism. The determination of the "true" protein-protein potentials thus remains an open problem.

Our choice of *purely repulsive* interactions illustrates the minimal assumptions allowing a satisfactory reproduction of the SAS data for β LG. In many studies on colloidal or protein solutions, satisfactory results were obtained from very simplified models. The use of sophisticated potentials, with a large number of different contributions, is often unnecessary at the first stages. Moreover, a high level of description for potentials would be in striking contrast with the poor level of approximation to the RDF's (W0 approximation) commonly adopted in many analyses of experimental data.

As regards the approximation of *spherical symmetry*, used for the protein-protein interactions (but not in the calculation

of the form factors), we remark that it represents a common simplifying choice. In particular, it is worth recalling a very recent study by Pellicane *et al.* [32], which reports evidence that the phase diagram of prototype globular protein solutions (lysozyme and γ -crystallin in water and added salt) can be reasonably reproduced by a spherically symmetric representation of macromolecular interactions. These authors employed a HS-Yukawa-DLVO one-component potential, including the Hamaker attractive part.

Evidently, in addition to the molecular granularity of the solvent and the finite sizes of all microions, a highly refined model description of protein solutions should embody the asymmetry of the molecular shape as well as the heterogeneity of the macroion surface charge distribution. The presence of different charged surface groups may produce "charge patches" that have a sign opposite to that of the net macroion charge. The importance of non-spherically-symmetric models with an inhomogeneous distribution of positively and negatively charged groups was recently investigated in a MC study of the electrostatic complexation of flexible polyelectrolytes with α -lactalbumin and β -lactoglobulin [33].

As a final remark to the present paper, it is worth pointing out that we are not aware of any previous investigations of this type within the HS-Yukawa-DLVO binary model and in regimes typical of globular proteins in solution. Our results and the methodological approach based upon the W1 approximation are expected to be useful in the analysis of SAS experiments. It would be rather interesting to pursue a similar study on the thermodynamic predictions of the first-order approximation. This could be easily carried out, as all thermodynamic quantities can be inferred either directly or through knowledge of the RDF's. Another interesting issue, within the present framework, involves an increase of the asymmetry between the two considered molecular sizes, which is known to lead to possible depletion effects [34]. We plan to perform such investigations in a future publication.

ACKNOWLEDGMENTS

We are particularly grateful to Francesco Spinozzi, Flavio Carsughi, and Paolo Mariani for enlightening discussions and ongoing collaboration on the subject reported in this work. T.K.D. thanks Professor S.R. Shenoy and the Abdus Salam ICTP Trieste for some support. The Italian MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) through a PRIN-COFIN project and the INFN (Istituto Nazionale di Fisica della Materia) are gratefully acknowledged for partial financial support.

-
- [1] D. G. Grier, E. R. Dufresne, and S. H. Behrens, *Interactions in Colloidal Suspensions* (The University of Chicago Press, Chicago, 2000).
- [2] D. M. Bloor and E. Wyn-Jones, *The Structure, Dynamics and Equilibrium Properties of Colloidal Systems* (Kluwer Academic, Dordrecht, 1990).

- [3] R. Piazza, *Curr. Opin. Colloid Interface Sci.* **5**, 38 (2000).
- [4] F. Spinozzi, D. Gazzillo, A. Giacometti, P. Mariani, and F. Carsughi, *Biophys. J.* **82**, 2165 (2002).
- [5] G. Baldini, S. Beretta, G. Chirico, H. Franz, E. Maccioni, P. Mariani, and F. Spinozzi, *Macromolecules* **32**, 6128 (1999).
- [6] E. J. Verwey and J. Th. G. Overbeek, *Theory of the Stability of*

- Lyophobic Colloids* (Elsevier, Amsterdam, 1948).
- [7] J. P. Hansen and I. R. McDonald, *The Theory of Simple Liquids* (Academic Press, London, 1986).
- [8] F. Carsughi, A. Giacometti, and D. Gazzillo, *Comput. Phys. Commun.* **133**, 66 (2002); D. Gazzillo and A. Giacometti, *J. Chem. Phys.* **113**, 9837 (2000); **120**, 4742 (2004).
- [9] E. Meeron, *J. Chem. Phys.* **28**, 630 (1958).
- [10] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).
- [11] Y. Rosenfeld, *Mol. Phys.* **88**, 1357 (1996).
- [12] G. Salin and J. M. Caillol, *J. Chem. Phys.* **113**, 10459 (2000).
- [13] In a recent paper [E. Schöll-Paschinger, D. Levesque, J. J. Weiss, and G. Kahl, *J. Chem. Phys.* **122**, 024507 (2004)], the same Ewald construction has been used in a different regime for *attractive* hard-sphere Yukawa mixtures.
- [14] C. Rey, L. J. Gallego, L. E. González, and D. J. González, *J. Chem. Phys.* **97**, 5121 (1992).
- [15] H. Löwen, *Phys. Rep.* **237**, 249 (1994).
- [16] G. Nägele, *Phys. Rep.* **272**, 215 (1996).
- [17] J. P. Hansen and H. Löwen, *Annu. Rev. Phys. Chem.* **51**, 209 (2000).
- [18] L. Blum and J. S. Hoye, *J. Stat. Phys.* **19**, 317 (1978).
- [19] M. Ginoza, *Mol. Phys.* **71**, 145 (1990).
- [20] J. B. Hayter and J. Penfold, *Mol. Phys.* **42**, 109 (1981).
- [21] J. P. Hansen and J. B. Hayter, *Mol. Phys.* **46**, 651 (1982).
- [22] H. Ruiz-Estrada, M. Medina-Noyola, and G. Nägele, *Physica A* **168**, 919 (1990).
- [23] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Academic Press, New York, 1987).
- [24] S. Labik, A. Malijevsky, and P. Vonka, *Mol. Phys.* **56**, 709 (1985).
- [25] F. J. Rogers and D. A. Young, *Phys. Rev. A* **30**, 999 (1984).
- [26] G. Zerah and J. P. Hansen, *J. Chem. Phys.* **84**, 2336 (1986).
- [27] Note that the possibility of a lack of thermodynamic consistency has been already noted earlier both for RY [G. Nägele, *Phys. Rep.* **272**, 215 (1996)] and HMSA [G. Kahl and G. Pastore, *Europhys. Lett.* **7**, 37 (1988); F. Ould-Kaddour and G. Pastore, *Mol. Phys.* **81**, 1011 (1994)].
- [28] K. G. Ng, *J. Chem. Phys.* **61**, 2680 (1974).
- [29] J. P. Hansen and J. J. Weis, *Mol. Phys.* **33**, 1379 (1977).
- [30] P. J. Rossky, J. B. Dudowicz, B. L. Tembe, and H. L. Friedman, *J. Chem. Phys.* **73**, 3372 (1980).
- [31] S. Ciccariello and D. Gazzillo, *J. Chem. Phys.* **76**, 1181 (1982).
- [32] G. Pellicane, D. Costa, and C. Caccamo, *J. Phys. Chem. B* **108**, 7538 (2004).
- [33] R. de Vries, *J. Chem. Phys.* **120**, 3475 (2004).
- [34] See, e.g., C. Likos, *Phys. Rep.* **348**, 267 (2001).