Journal of Molecular Liquids 200 (2014) 20–27

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Effective interactions in protein–salt solutions approaching liquid–liquid phase separation

Marcell Wolf [a], Felix Roosen-Runge [a], Fajun Zhang [a,*], Roland Roth [b,*], Maximilian W.A. Skoda [c], Robert M.J. Jacobs [e], Michael Sztucki [d], Frank Schreiber [a]

[a] Institut für Angewandte Physik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany
[b] Institut für Theoretische Physik, Universität Tübingen, Auf der Morgenstelle 14, 72076 Tübingen, Germany
[c] ISIS, Rutherford Appleton Laboratory, Chilton, Didcot OX110QX, UK
[d] Chemistry Research Laboratory, University of Oxford, Oxford OX13QZ, UK
[e] European Synchrotron Radiation Facility, 6 rue Jules Horowitz, F-38043 Grenoble Cedex 9, France

ARTICLE INFO

Article history:
Received 18 November 2013
Received in revised form 22 May 2014
Accepted 4 August 2014
Available online 20 August 2014

Keywords:
Protein–protein interaction
SAXS
Liquid–liquid phase separation
Reduced second virial coefficient

ABSTRACT

We present an experimental study combined with a theoretical discussion of the effective interactions in protein solutions approaching a liquid–liquid phase separation (LLPS) induced by addition of multivalent metal ions. The reduced second virial coefficient, B_2/B_2^{HS}, is used to describe the interaction and discussed with theoretical predictions for colloidal systems. We have determined the salt and protein partitioning in the two coexisting phases, which provides the isothermal binodal of the LLPS in the (c_p, c_s) plane. Two sets of samples, away from and at the LLPS binodal were measured by static light scattering (SLS) and small angle X-ray scattering (SAXS) to determine the second virial coefficient. In all cases, B_2/B_2^{HS} is negative in the condensed regime, and increases by approaching the upper critical point in the (c_p, c_s) plane. The results are compared with a simple colloidal model with isotropic short-ranged attraction and a thermodynamic criterion based on the reduced second virial coefficient. We discuss the application of this theoretical prediction to interpret experimental observations.

1. Introduction

The understanding of protein interactions in aqueous solutions is crucial for many issues in soft and biological matter. While the equilibrium phase diagram of proteins has some similarities with that of model colloidal systems, the physical mechanisms of protein crystallization with its huge importance for structural biology are far from understood, and, in most cases, remain elusive. Proteins, like colloids, in solution can be seen as big atoms that, however, interact by effective interactions resulting from the behavior of the other components in the solution, i.e. the solvent and salt. By changing the solvent conditions it is possible to alter the resulting protein interactions to a large degree, and thereby tune the phase behavior of protein solutions.

A particularly interesting example for phase behavior in protein solution is the metastable liquid–liquid phase separation (LLPS). LLPS in protein solution is a fundamental biophysical phenomenon and provides a mechanism for biological structure formation [1–8] such as a prerequisite for the formation of crystals in cataracts [2,6,8] and fibers responsible for sickle cell anemia and Alzheimer's disease [1,3], and changes on the pathways of protein crystallization [7,9,10].

The physical reason of a metastable LLPS for colloids or proteins in solutions has been demonstrated both experimentally and theoretically to be due to the short-ranged nature of the attractive interactions [1,10–14]. In contrast, in atomic systems, such as argon, where the attractive interaction is long-ranged compared to the molecular size, a stable gas–liquid phase separation exists [12]. Rosenbaum et al. have shown that the crystallization curves for a number of globular protein solutions are similar to those predicted by simulations for a system of hard spheres with a short-ranged attractive Yukawa potential [11,15]. Asherie et al. performed a combined analytical and computational study on the phase diagram of globular colloids [14]. Their study reveals that the interaction range plays a significant role in determining the structure of the phase diagram. A short-ranged attraction, i.e. the interaction range is smaller than ≈25% of the diameter of particles, is prerequisite for the existence of the meta-stable LLPS in protein and colloid systems. Simulations and theoretical studies also support that a short ranged attraction leads to the metastable LLPS [9,10,16]. By comparing existing protein crystallization data with knowledge on a model colloid–polymer mixture, where the attraction range as well as strength between colloids can be tuned by varying the molecular weight and concentration of non-absorbing polymer, Poon suggested a hidden gas–liquid binodal inside the equilibrium fluid–crystal region of the phase diagram [13]. The critical points of LLPS can be nicely described using the thermodynamic criterion based on B_2, which has been used

* Corresponding authors.
E-mail addresses: fajun.zhang@uni-tuebingen.de (F. Zhang), roland.roth@uni-tuebingen.de (R. Roth).

http://dx.doi.org/10.1016/j.molliq.2014.08.006
0167-7322/

as a predictor in protein crystallization. George and Wilson observed that B_2 falls in a narrow range for protein crystallization [17]. Subsequent theoretical work by Vliegenthart and Lekkerkerker demonstrated that B_2 has a nearly constant value at the critical point [16,18] and indeed can be used as a predictor for protein crystallization, i.e. the optimal conditions for crystal growth are either near the critical point where the density fluctuation enhances the nucleation rate [10], or below the critical point but near the protein-poor phase boundary where crystals grow via a two-step procedure [18].

While significant progress has been made in the understanding of the physical mechanism of the metastable LLPS, B_2 as a predictor for the critical point of LLPS has not been tested experimentally. In practice, due to the non-spherical shape and the combination of specific and non-specific interactions, it is not clear how strong the overall attraction has to be to lead to the LLPS in protein solutions.

We have shown that the interactions in protein solutions can be efficiently tuned by the variation of the ionic strength, the nature of ions and the valency of ions [19–25]. A rich phase behavior, including reentrant condensation, metastable liquid–liquid phase separation (LLPS), cluster formation, and crystallization, has been observed when trivalent salts are used [20,21,24–29]. This tunable phase behavior can be used to optimize the conditions for protein crystallization [25–28].

In this work, we present an experimental study combined with a theoretical discussion on the LLPS in protein solutions induced by trivalent cations. We focus on a deeper understanding of the effective protein–protein interactions in the condensed regime as a function of protein and salt concentration, which determines whether macroscopic LLPS or microscopic protein clustering occurs for a given protein solution. We present additional experimental data for the LLPS binodal. In particular, also data close to the critical point in the (c_p, c_s) plane are presented and discussed. The effective protein–protein interactions are studied using static light scattering (SLS) and small angle X-ray scattering (SAXS). The second virial coefficient is used to understand the key question: which attraction strength is needed in order to induce the LLPS. The results are further discussed within the theoretical framework established recently in colloidal systems [16,18,30,31].

2. Theory: second virial coefficient in effective one-component systems

We briefly elaborate on the phenomenon of a liquid–liquid phase separation (LLPS) in a potentially complex mixture of several components such as proteins, the solvent, and added salt. We provide a simple and intuitively transparent argument how LLPS relates to the second virial coefficient.

In many cases, it is natural to focus on the behavior of the largest component of the mixture, the proteins, and treating the remaining parts as a background medium. This can be done in a rigorous way by mapping the Hamiltonian of the mixture onto that of an effective one-component system [30,32] by integrating out the degrees of freedom of the background. This mapping changes the interactions in the system from the bare interactions between all possible species combinations, i.e. protein–water, and protein–ion, to effective interactions between the proteins. At first we assume that the effective interaction potential $V_{eff}(r)$ is spherically symmetric.

For such a system to undergo a liquid–liquid phase separation into a low density fluid phase ("protein–gas") and a high density fluid phase ("protein–liquid"), the effective interaction has to possess a sufficiently strong attractive tail in addition to the repulsion at very short distances. The reason, independent of the precise nature of the effective interaction, is simple to understand: When a low density phase is transformed into a high density phase, the system loses entropy, which has to be compensated by the gain in interaction energy due to the attraction.

A convenient measure for the strength of the attraction is the effective second virial coefficient, which for the assumed spherically symmetric interaction potential, is defined by

$$B_2(T) = 2\pi \int_0^\infty r^2 \left[1 - \exp\left(-\beta V_{eff}(r)\right)\right] dr. \tag{1}$$

If B_2 is positive, the net interaction is repulsive. If it is negative, then the interaction becomes attractive. But how strong has the effective interaction to be in order to drive a phase separation? Vliegenthart and Lekkerkerker made the interesting observation that for various systems the value of the second virial coefficient B_2 at the critical point seems universal [18]:

$$\frac{B_2}{B_2^{HS}} \approx -1.5, \tag{2}$$

where $B_2^{HS} = 16\pi R^3/3$ is the second virial coefficient of a hard sphere of radius R. This observation was also confirmed by Noro and Frenkel [16] and was tested for hard-sphere mixtures [33] using the depletion potential [30].

The sticky hard-sphere model was introduced by Baxter [34] as an example of a system with hard-core repulsion and additional short-ranged attraction, which can undergo fluid–vapor phase separation. Some aspects of the system can be treated analytically within certain approximate closure relations. The interaction potential is given by

$$\beta V_{eff} = \begin{cases} \infty & r < \sigma = 2R \\ -\beta u_0 = \ln\left(\dfrac{12\tau\Delta}{\sigma+\Delta}\right) & \sigma < r < \sigma + \Delta \\ 0 & r > \sigma + \Delta, \end{cases} \tag{3}$$

where usually the limit $\Delta \to 0$ is taken. In this limit the reduced second virial coefficient is given by

$$\lim_{\Delta\to 0} \frac{B_2}{B_2^{HS}} = 1 - \frac{1}{4\tau}. \tag{4}$$

Baxter found that within the Percus–Yevick closure relation the critical point is given by [34]

$$\tau_c = \frac{2-\sqrt{2}}{6} \approx 0.0976, \quad \text{and} \quad \eta_c = \frac{3\sqrt{2}-4}{2} \approx 0.1213, \tag{5}$$

so that for the reduced second virial coefficient, Eq. (4), at the critical point one finds

$$\frac{B_2(\tau = \tau_c)}{B_2^{HS}} = 1 - \frac{1}{4\tau_c} \approx -1.56, \tag{6}$$

which agrees well with the aforementioned criterion, Eq. (2). It should be mentioned that the level of agreement depends on the treatment of the sticky hard spheres. For example, from the energy route (e) the critical point is found at $\tau_c^e \approx 0.1185$ and $\eta_c^e \approx 0.3189$, which results in $B_2(\tau = \tau_c^e)/B_2^{HS} \approx -1.11$ [35]. In computer simulations (s) one finds $\tau_c^s \approx 0.1133$ and $\eta_c^s \approx 0.266$, which results in $B_2(\tau = \tau_c^s)/B_2^{HS} \approx -1.21$ [36]. The agreement of these results with Eq. (2) is still reasonable, even if they are somewhat off.

In order to rationalize the observation by Vliegenthart and Lekkerkerker better we consider phase coexistence between a low density gas phase (density ρ_I) and a high density fluid phase (density ρ_{II}) in more detail. The following argument is kept simple and hence is not indented to be a perfect account for the complex behavior of the system at the critical point. The two phases can coexist at the same temperature T if they are in mechanical and in chemical equilibrium, i.e.

$$P(\rho_I) = P(\rho_{II}) \quad \text{and} \quad \mu(\rho_I) = \mu(\rho_{II}) \tag{7}$$

where P is the pressure and μ the chemical potential of the system under consideration. For mechanical equilibrium close to the critical point to be possible, the pressure at coexistence has to be low, because the pressure of a high density liquid has to be balanced by the low pressure of the coexisting gas. Therefore, at coexistence, it is possible to expand the pressure into a virial series with only few terms with the most prominent contribution being the second virial term:

$$\beta P(\rho) \approx \rho + B_2\rho^2 + B_3\rho^3 + \ldots \tag{8}$$

For our purpose here the virial expansion up to third order is sufficient, as it already allows to describe phase coexistence. For a better quantitative account of the equation of state more virial coefficients would be required. Note that the virial expansion converges rather slowly in general, and leads to additional complications close to the critical region. Only if the attraction is sufficiently strong the pressure of a high density phase can be equally low as that of the low density phase.

The location of the critical point ρ_c and T_c, which is the onset of a fluid–fluid phase separation which is reflected by the onset of a van-der-Waals loop in the pressure, follows from

$$\left.\frac{\partial P(\rho)}{\partial \rho}\right|_{\rho=\rho_c} = 0 = \left.\frac{\partial^2 P(\rho)}{\partial \rho^2}\right|_{\rho=\rho_c}. \tag{9}$$

Note that the first condition in Eq. (9), the vanishing of the first derivative of the pressure P w.r.t. density ρ, also expresses the condition of the vanishing of the inverse compressibility or the bulk modulus. This is important because in the limit of $q \to 0$ the structure factor $S(q)$ is proportional to the compressibility χ_T, i.e. $S(q \to 0) = k_B T \rho \chi_T$. This means that when the derivative of the pressure w.r.t. density, or the inverse compressibility, vanishes, at the critical point or at the spinodal line, the structure factor diverges for $q \to 0$. If one considers a system at a coexisting density, i.e. on the binodal line, then the state is also close to the spinodal, which implies that the compressibility and hence $S(q \to 0)$ are large, but do not diverge. The closer the system is to the critical point, the closer the binodal and spinodal lines are. Therefore close to the critical point, on the binodal (at either the low or the high coexisting density), the compressibility and $S(q \to 0)$ are large, while further away from the critical point, the distance between the binodal and spinodal increases, so that also the compressibility at a coexisting density and the structure factor $S(q \to 0)$ decrease.

By combining Eqs. (8) and (9) we obtain a set of equations which can be solved e.g. for the second and third virial coefficients at the critical point. Here only the result for B_2 is of interest. One finds that

$$\frac{B_2}{B^{HS}} \approx -\frac{1}{4\eta_c} \approx -2.06, \tag{10}$$

where we have inserted the value of the critical packing fraction due to Baxter [34], Eq. (5). While the agreement between Eqs. (2) and (10) is clearly not perfect, this simple argument helps to rationalize the origin of the B_2 criterion based solely on the idea of mechanical equilibrium. Note that the value of the second virial coefficient at the critical point based either on Eq. (2) or on Eq. (10) changes somewhat, if other estimates for the critical packing fraction η_c are employed. For the critical packing fraction from the energy route [35] for example we obtain $B_2/B^{HS} \approx -0.784$ and for the critical packing fraction from computer simulation [36] we obtain $B_2/B^{HS} \approx -0.940$. The magnitude remains of the same order.

The conclusion is that for $\tau < \tau_c$ or $B_2/B_2^{HS} < -1.5$ the effective interaction potential is sufficiently strong to drive a phase separation into a low density and a high density phase.

In general, the effective interaction potential between proteins $V_{eff}(1,2)$ is not spherically symmetric, as assumed so far in this section, but depends on the distance r_{12} between centers of proteins 1 and 2 as well as on their orientations $\hat{\Omega}_1$ and $\hat{\Omega}_2$. For such an effective interaction the second virial coefficient is given by

$$B_2(T) = \frac{1}{2V}\int\int\int\int \left[1 - \exp\left(-\beta V_{eff}(1,2)\right)\right] d^3r_1 \; d\hat{\Omega}_1 \; d^3r_2 \; d\hat{\Omega}_2. \tag{11}$$

The second virial coefficient be calculated analytically only for special cases of $V_{eff}(1,2)$. One interesting example is the Kern–Frenkel potential [37], which accounts for short-ranged square-well attractions of depth $-\epsilon$ between isolated patches on the surface of spheres. Each sphere has n patches, which are contained within a cone, with an apex (of angle $2\,\delta$) at the center of the protein. The fraction of the surface that is covered by patches is given by $\chi = n(1 - \cos(\delta))/2$. The range of the square-well attraction is given by $\lambda\sigma$, where σ is the hard-core diameter. For this model the reduced second virial coefficient is given by [37]

$$\frac{B_2}{B_2^{HS}} = 1 - \chi^2\left(\lambda^3 - 1\right)\left(\exp(\beta\epsilon) - 1\right), \tag{12}$$

which should be compared to Eq. (4).

If the patches cover a sufficiently large fraction of the surface, the reduced second virial coefficient at the critical point in the Kern–Frenkel model is compatible with the observation by Vliegenthart and Lekkerkerker, Eq. (2). However, as the fraction of the surface covered by patches, χ, decreases, the value of reduced second virial coefficient at the critical point can be significantly below -1.5 [38] so that a LLPS requires a strongly increased attraction and eventually becomes impossible.

3. Materials and methods

3.1. Materials and sample preparation

Human serum albumin, HSA and yttrium chloride, YCl_3, were purchased from Sigma-Aldrich and used as received. All samples were prepared at room temperature (22 °C). A series of protein solutions with various salt concentrations were prepared by mixing stock solutions of dissolved protein and salt in degassed Milli-Q water. No buffer was used to avoid the effect of other co-ions. The phase diagram (reentrant condensation and LLPS) was determined by optical transmission and visual inspection. Protein concentrations (c_p) were determined by UV absorption (Cary 50 UV–Visible spectrometer from Varian Inc., California, USA) at a wavelength of 280 nm with a coefficient of 0.51 ml/mg [39]. The protein-poor and protein-rich phases after LLPS were separated by centrifugation. c_p of the protein-poor phase was determined directly by UV absorption, and c_p of the protein-rich phase was calculated from the volume of each phase and the initial protein concentration by the lever rule. Salt partitioning was determined by X-ray absorption as described in detail in Ref. [27].

3.2. Static light scattering (SLS)

For the SLS measurements we assume the complex of the protein and the bound ions as a new “effective particle”. Because of the binding of metal ions to the protein surface, conventional SLS measurements result in a nonlinear relationship between Kc_p/R_θ and c_p. This is reasonable since solutions with constant c_s and varying c_p correspond to different complexes of protein and salt and thus exhibit different interactions throughout the phase diagram, even across the phase boundary. To solve this problem, we use a method described below for our SLS measurement.

First, a series of sample solutions was prepared with a constant c_p (here, 3.1 mg/ml HSA) with various c_s ranging from 0.01 mM to 20 mM across all three regimes. Second, for each sample (2.0 ml), it

was first concentrated up to ≈3 times of its initial c_p using an ultrafiltration tube (Amicon Ultra-15, Millipore Corporation, Billerica, USA, with a molecular weight cut-off of 50 kDa). The salt solution collected in the bottom of the filter was used to dilute the concentrated solution into a series for SLS measurements. By this means, all samples were measured under comparable conditions. Measurements were carried out on a Zetasizer Nano ZS (Malvern Instrument Ltd., Worcestershire, UK) at 25 °C with a fixed angle of 173°. The laser has a wavelength of 632.8 nm and a power of 4 mW. Solutions were filled in a quartz cuvette with a path length of 1 cm.

The second virial coefficient A_2 and the apparent molecular weight M_W were obtained from the Rayleigh equation [40]

$$\frac{K \cdot c_p}{R_\theta} = \frac{1}{M_W} + 2A_2 \cdot c_p, \tag{13}$$

where K is the universal optical constant, and R_θ is the Rayleigh ratio [40].

Since the osmotic pressure in a protein solution can be defined via either the protein concentration c_p or the protein number density ρ, the virial expansion gives two second virial coefficients A_2 and B_2, respectively [1,17,41]. The relation between A_2 and B_2 is given by

$$A_2 = B_2 \cdot \frac{N_A}{M_W^2}. \tag{14}$$

In this work, the molecular weight $M_W = 66$ kDa of a HSA monomer is used in the calculation.

3.3. Small-angle X-ray scattering (SAXS) and data analysis

The SAXS measurements were performed at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, at the beamline ID02 with two sample-to-detector distances of 2 m and 5 m. The energy of the incoming beam was 16.038 keV (wavelength 0.8 Å), with a q-range from 0.007 $Å^{-1}$ to 0.4 $Å^{-1}$. The detector was a fiber optically coupled fast-readout low-noise (FReLoN) CCD based on a Kodak KAF-4320 image sensor in an evacuated flight tube. About 0.1 ml sample was filled into a flow-through quartz capillary. The sample in the scattering volume was exchanged for every exposure. For each sample, 10 exposures of 0.1 s each were measured. The 2D intensity pattern was corrected to an absolute scale and azimuthally averaged to obtain the intensity profiles, and the solvent background was subtracted. More detailed information on data reduction and q-resolution calibration can be found in the literature [23,42].

3.3.1. Data analysis

Small-angle X-ray scattering data can be used to obtain information on the pair interaction potential [43–45]. The scattering intensity, $I(q)$, for a polydisperse or a non-spherical system, can be calculated on the basis of approximation approaches such as the "decoupling approximation" and "average structure factor" approximation [46,47]. Both approaches assume that the particle position is not correlated with its orientation. For the case of non-spherical but monodisperse solutes at a low to intermediate concentration, such as the studied protein solutions, both assumptions give comparable results [19]. Therefore, in this work, the scattering intensity is calculated using the average structure factor approximation, which can be expressed by

$$I(q) = N(\Delta\rho)^2 V^2 P(q)\bar{S}(q) \tag{15}$$

where $q = \frac{4\pi}{\lambda}\sin(2\theta/2)$ is the scattering vector, 2θ is the scattering angle, N is the number of protein molecules per unit volume in the solution, $\Delta\rho = \rho_p - \rho_s$ is the scattering length density difference between the solvent and the solute, and V is the volume of a single protein. $P(q)$ is the form factor of a given protein, i.e. the scattering from a single protein molecule after orientation averaging. A form factor of an oblate ellipsoid with semi-axes a and b is used to model HSA [48]. Using the average structure factor approximation $\bar{S}(q)$ is calculated from a monodisperse spherical system, with an effective sphere diameter. In our case, the protein solution is a monodisperse but a non-spherical system. The effective sphere diameter is calculated from a virtual sphere with the same second virial coefficient as the ellipsoid [22,49]. In the following parts and for simplicity, we use $S(q)$ to denote $\bar{S}(q)$.

4. Results and discussion

4.1. Phase diagram of HSA with YCl_3 in the (c_p, c_s) plane

It has been shown in our previous work that acidic proteins such as HSA in solutions with YCl_3 exhibit a reentrant condensation phase behavior [20,27]. Within the condensed regime, a metastable LLPS exists. Here we show additional experimental results on determining the phase boundary of LLPS, in particular the samples near the upper critical point in the (c_p, c_s) plane. Three sets of solutions with initial protein concentrations of $c_p = 31.1$ mg/ml, 47.8 mg/ml and 74.0 mg/ml were prepared at room temperature as a function of salt concentration in the condensed regime. The sample solutions were initially turbid. Upon centrifugation, a clear protein-poor phase separated from a protein-rich phase. The partitioning of both protein and salt, determined using UV–Vis and X-ray absorption method as described in the Materials and methods section, gives the coexistence curve, the isothermal binodal.

The data are plotted in Fig. 1. The resulting c_p and c_s of the protein-poor phase indicate that the phase boundaries are comparable for all three sets of samples with different initial protein concentrations. The salt concentration in the protein-rich phases cannot be determined easily due to the high viscosity.

The experimental phase diagram on a log-scale shows a closed ellipsoidal area containing the LLPS. For this closed phase boundary, two critical points are expected in the (c_p, c_s) plane — one at low and one at high c_s. Although the precise location of both critical points has not been determined experimentally, possible regions are marked by orange striped areas in Fig. 1. The protein-poor phases for samples with initial protein concentration of $c_p = 47.8$ mg/ml extend much closer

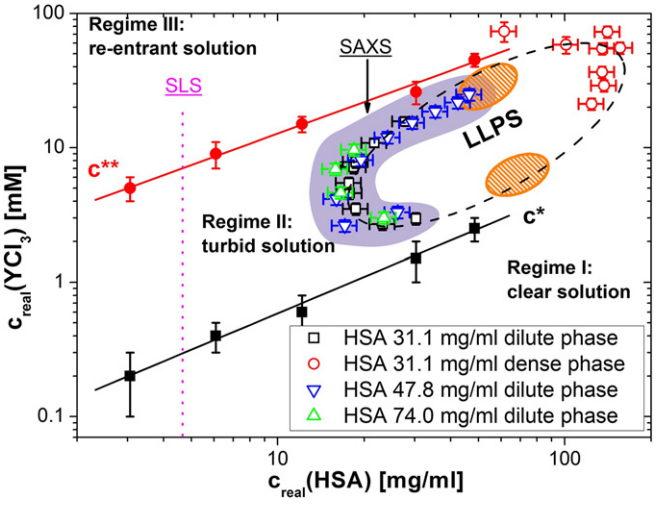

Fig. 1. Plot of reentrant phase diagram with real protein and salt concentrations. The solid black and red symbols (square, circle) and lines correspond to the boundaries of c^* and c^{**}. Data points from the protein-poor and protein-rich phases after LLPS are presented by open symbols. The dashed ellipsoid is a guide to the eyes outlining the LLPS region. The magenta dotted line indicates the samples for SLS measurements. The gray-highlighted area around the LLPS boundary indicates the samples for SAXS. The two orange striped areas indicate the estimated regions for the lower and upper critical points.

to the upper critical point compared to the other two sets of samples. For example, the sample with the initial salt concentration c_s = 32 mM, gives a protein-poor phase with c_p = 46.5 mg/ml which is very close to the initial concentration of c_p = 47.8 mg/ml, indicating proximity to the upper critical point.

Given the closed area of the LLPS at intermediate c_p, the condensed regime of the phase diagram is divided into three regions. On the left-hand side of the LLPS, i.e. with lower protein and salt concentrations, proteins form clusters but no macroscopic LLPS can occur. On the right-hand side of the LLPS, i.e. with much higher protein and salt concentrations, although the state of the system has not been characterized, one would expect an arrested gel or a highly non-equilibrium state.

The following results on the second virial coefficient focus on the understanding of the effective protein–protein interactions in the clustering region and at the binodal of the LLPS. The questions we are interested in are the following: First, how does the interaction potential change with increasing salt concentration for a given protein concentration, such as those along the magenta dotted line in Fig. 1? Second, do the samples with a constant protein to salt ratio, but in different location, share the same interaction potential? For example, the samples located in the clustering region and at the binodal of LLPS. Furthermore, since we have samples located on the binodal of LLPS very close to the critical point, it would be interesting to follow the change of the interaction potential by approaching the critical point and compare it to the value predicted in theory.

We have performed both SLS and SAXS measurements to characterize the effective interactions in the cluster phase as well as at the binodal of LLPS. From that the reduced second virial coefficient has been derived. Samples for SLS and SAXS measurements are labeled by a magenta dotted line and a gray area in Fig. 1, respectively.

4.2. Second virial coefficient determined by static light scattering (SLS)

Using an ultracentrifugation method described in the Materials and methods section, we have performed SLS measurements for a series of samples with a constant initial protein concentration c_p = 3.1 mg/ml and varying salt concentration c_s from 0.1 to 20 mM. As shown in Fig. 1, these samples cover all three regimes. Typical plots of Kc_p/R_θ against c_p are shown in Fig. 2A. In all cases, a clear linear relationship is obtained. From the slope one can determine A_2 using Eq. (13). The linear extrapolation to zero protein concentration for most of the samples points to the value of 66 kDa, corresponding to the molecular weight of a HSA monomer. Therefore, the larger clusters have been effectively removed by filtration or centrifugation, and the residual protein concentration is sufficiently low not to form clusters. For samples with c_s = 1.0 and 7.5 mM, which are near c^* and c^{**}, the extrapolation of Kc/R_θ to $c_p = 0$ is lower, indicating the formation of protein clusters.

Table 1
Second virial coefficient A_2 and the corresponding B_2/B_2^{HS} determined from ultrafiltration SLS measurements for a series of samples with initial c_p = 3.1 mg/ml and various c_s.

YCl_3 [mM]	Y^{3+}/protein	$A_2 \cdot 10^4$ $\left[\frac{mol \cdot ml}{g^2}\right]$	B_2/B_2^{HS} r = 33.5 Å	B_2/B_2^{HS} r = 26.8 Å	B_2/B_2^{HS} r = 40.4 Å
0.1	2.1	13.1/1.7	15.3/1.9	29.8/1.9	8.7/1.9
0.2	4.3	2.8/1.4	3.3/1.6	6.4/1.6	1.9/1.6
0.3	6.4	−0.3/2.9	−0.4/3.3	−0.7/3.3	−0.2/3.3
0.5	10.7	−4.3/0.6	−5.1/0.7	−9.9/0.7	−2.9/0.7
1.0	21.4	−1.5/1.8	−1.8/2.0	−3.5/2.0	−1.0/2.0
2.0	42.9	−0.6/1.8	−0.7/2.1	−1.4/2.1	−0.4/2.1
3.0	64.3	−2.6/1.7	−3.1/1.9	−6.0/1.9	−1.7/1.9
4.0	85.8	0.7/1.6	0.8/6.6	1.6/6.6	0.5/6.6
5.0	107.2	−1.0/2.2	−1.2/2.5	−2.4/2.5	−0.7/2.5
7.5	160.8	−0.6/0.3	−0.7/0.4	−1.3/0.4	−0.4/0.4
9.0	193.0	0.5/0.9	0.5/1.1	1.0/1.1	0.3/1.1
20.0	428.9	1.0/3.4	1.2/3.9	2.4/3.9	0.7/3.9

The values of the second virial coefficient A_2 obtained by fitting the data and the corresponding B_2/B_2^{HS} are listed in Table 1. For the calculation of B_2/B_2^{HS} from A_2, the excluded volume of the single particle is needed. However, precise determination of the excluded volume of a protein in aqueous solution is non-trivial due to the non-spherical shape and the hydration of proteins. As discussed in previous work [22], these two effects affect the effective excluded volume of proteins in solution. It has been shown that the effective radius determined by SAXS provides a good estimation for both effects [22,50,51]. For HSA under physiological conditions, the effective radius determined by SAXS is r = 33.5 Å [22]. In the current study, in the presence of YCl_3, SAXS measurements give a value of r = 40.4 Å (see next section). For comparison, the radius of HSA calculated from the specific volume of the monomer r = 26.8 Å is also used.

The calculated reduced second virial coefficients B_2/B_2^{HS} (r = 40.4 Å) are plotted in Fig. 2B as a function of the number of Y^{3+} per protein. The cyan area shows the theoretical limit of B_2/B_2^{HS} for LLPS ranging from −1.5 (Eq. (2)) to −2.06 (Eq. (10)). From Fig. 2B one can see that the values of B_2/B_2^{HS} are positive at very low salt concentrations and decreases first with increasing c_s, reaching a minimum of −5 at about 10 Y^{3+} per protein, then increases slowly and becomes positive again at very high salt concentrations. This observation is consistent with the reentrant condensation phase behavior of the protein solutions, i.e. in regimes I and III, the interaction is dominated by repulsion, whereas attraction is dominated in regime II. However, due to the systematic

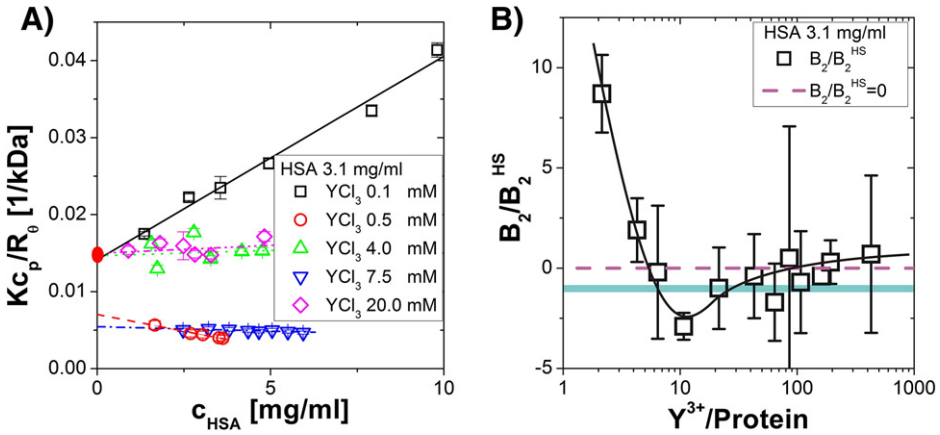

Fig. 2. A) Typical Debye plots of Kc_p/R_θ versus c_p from SLS measurements. The red point of the y-interception corresponds to the inverse molecular weight of a HSA monomer. B) Plot of the reduced second virial coefficient B_2/B_2^{HS} determined by light scattering versus Y^{3+}/protein ratio. The black solid line is a guide to the eyes. The cyan area shows the theoretical limit for B_2/B_2^{HS} with an isotropic interaction potential.

error in the determination of both B_2/B_2^{HS} and the values of c^* and c^{**}, it is not possible to deduce a precise relation between them.

4.3. Effective protein–protein interactions at the LLPS binodal determined by SAXS

The form factor of the protein solutions in the presence of multivalent metal ions has been determined using samples with $c_p = 1.3$ and 6.0 mg/ml with various c_s. All the SAXS curves overlap in the whole q range after normalization by c_p. Fig. 3A shows a typical SAXS profile for a sample with $c_p = 1.3$ mg/ml and $c_s = 8$ mM. Optimization of a model with an ellipsoid form factor gives the dimension of $18 \times 52 \times 52$ Å^3 with an error of 1 Å in each dimension resulting in an effective sphere radius of $r = 40.4$ Å [27]. Compared to the effective radius of HSA ($r = 33.5$ Å) under physiological conditions, the increase in size may be due to the formation of small protein clusters via cation bridging [28]. The dimension of the form factor determined here is used for the following data analysis.

A typical SAXS profile for a sample located at the LLPS binodal with a model fit is shown in Fig. 3B. The SAXS data were collected for a larger q range compared to that shown in our previous work [27]. An upturn is visible in the extended low q region, which can be explained by the protein clustering. To fit the data over the full q range, we use a sum of a power law and a sticky hard sphere (SHS) structure factor. The power law is used to describe the contribution from protein clusters, i.e. the upturn at low q region. We noticed that when leaving the power as a fit parameter, the values obtained for various data are always very close to 3. We therefore fix the power to 3 for all data analysis. By fixing it we avoid artifacts in the fitting procedure for the SHS structure factor which is used to describe the short-ranged attraction between proteins. The volume fraction of the protein was fixed by the value measured by UV–Vis absorption. For the SHS model, Δ was fixed to 0.02σ to prevent artificial coupling with τ. This procedure gives a good fit for the complete scattering curve as shown in Fig. 3B.

Following the data analysis described above, we have fit all SAXS data for samples located at the LLPS binodal. We focus on the samples with an initial $c_p = 47.8$ mg/ml (Fig. 4A) because this series of samples cover a large region of the binodal including the data very close to the upper critical point.

We first discuss the scattering intensities at $q \to 0$, $I(0)$ (Fig. 4B), which provides in a model-free way the effective interactions of the system. $I(0)$ reflects the compressibility χ_T, since $S(q \to 0) = k_B T \rho \chi_T$ [45]. The compressibility χ_T diverges at the spinodal line. The measured samples correspond to the binodal line which approaches the spinodal line once the coexisting densities become closer. Thus, the closer the coexisting densities are, the larger is χ_T and hence $S(q \to 0)$ in the coexisting phases.

The experimental results on the microscopic interactions thus reflect the phase behavior, as can be seen by comparing the SAXS intensity normalized by c_p at $q = 0.01$ Å (Fig. 4B) for samples located at the LLPS binodal. With increasing of c_s, corresponding to an increase of the number of Y^{3+} per protein, above a critical value (around 6 Y^{3+} per protein), the system phase separates. For this c_s the system is close to c^*, which results in large values of χ_T and $S(q \to 0)$. As the c_s is increased further up to around 13 Y^{3+}/protein, the coexistence region broadens, causing χ_T and $S(q \to 0)$ to decrease. In the range between 13 to 40 Y^{3+}/protein the trend is reversed: χ_T and $S(q \to 0)$ increase again, until the solution is mixed again for Y^{3+}/protein above 40. For samples with initial $c_p = 47.8$ mg/ml, as shown in Table 2, further increasing c_s above 20 mM does not change the composition of the protein-poor phase, i.e. the number of Y^{3+}/protein is nearly constant. In addition, from Fig. 1 we know that these samples are approaching the upper critical point, therefore, the normalized scattering intensity shows a different behavior, i.e. decreases at a constant ratio of Y^{3+}/protein in Fig. 4B. As the system needs the minimum attraction to induce LLPS at the critical point as predicted in theory, and $S(q = 0)$ follows the similar trends of $I(q = 0)$ without the contribution of clustering, this explains the decrease of $I(0)$ approaching the critical point.

The corresponding calculated sticky hard sphere structure factors $S(q)$ are shown in Fig. 4C. For reasons of comparison, $S(q)$ has been replotted with a fixed protein volume fraction. As one can see, the $S(q \to 0)$ with the obtained interaction parameters consistently follows the same trend as $I(q \to 0)$ near the critical point, i.e. the overall attraction decreases approaching the critical point.

Finally, we present the effective attraction between proteins obtained from the SAXS data. The values of the stickiness parameter from data fitting are used to calculate the B_2/B_2^{HS} as shown in Fig. 4D. The values of τ, B_2/B_2^{HS} and the depth of the potential u_0 for a complete set of sample are listed in Table 2. The values of τ are generally below $\tau_c = 0.0976$ (Eq. (5)) for all samples located at the binodal of LLPS. The corresponding interaction potential has a depth of $u_0 \sim -4k_BT$ and the value decreases when approaching the critical point. The values of B_2/B_2^{HS} for all samples at the binodal of LLPS are negative. In particular for the samples with the initial c_p of 47.8 mg/ml, the value of the B_2/B_2^{HS} increases towards the critical region predicted in theory, reflecting the vicinity to a critical point. By comparing the B_2/B_2^{HS} obtained by SLS and SAXS, one can see that despite of the large error for the SLS results, the values show a similar trend and range as those from SAXS. This finding

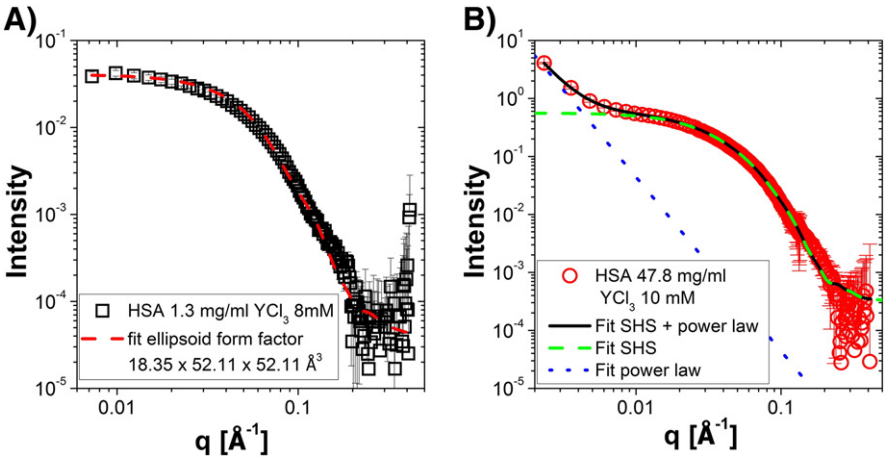

Fig. 3. A) SAXS profile with a model fit using the form factor of an oblate ellipsoid for a dilute protein solution of 1.3 mg/ml HSA with 8 mM YCl_3. B) Typical SAXS profile of a sample located at the LLPS binodal and the corresponding model fit using a sticky hard sphere structure combined with a power law. The sample has an initial $c_p = 47.8$ mg/ml and $c_s = 10$ mM.

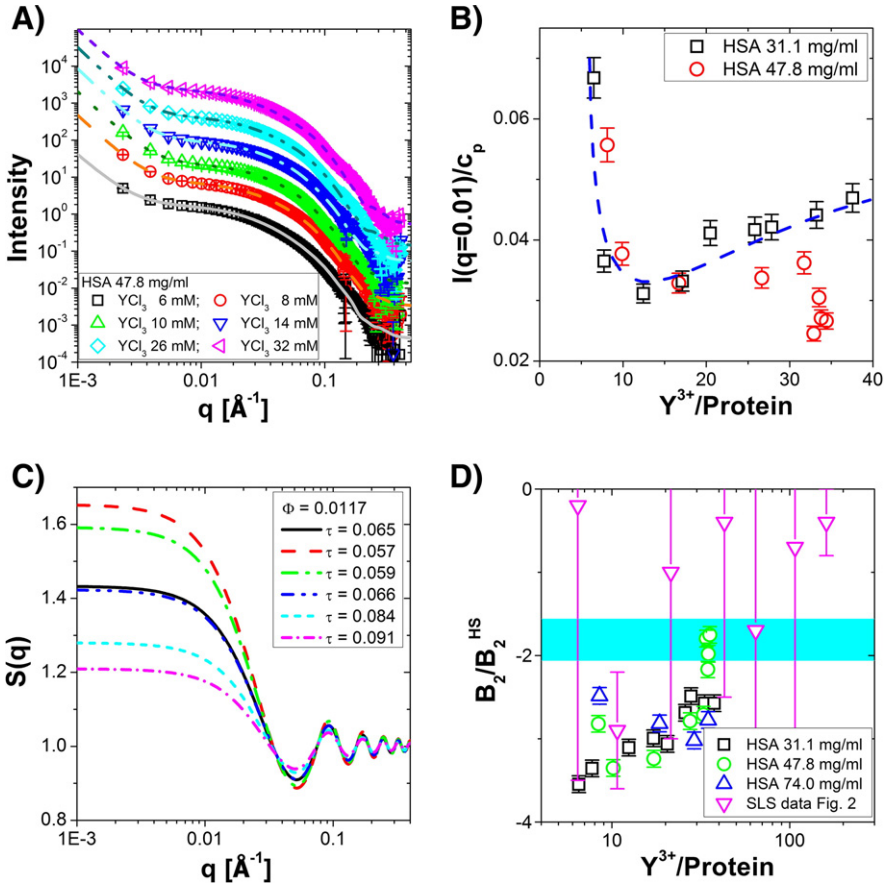

Fig. 4. A) Typical SAXS profiles with curves generated from the model for samples of HSA 47.8 mg/ml at different c_s. The scattering curves are shifted in intensity for clarity. B) The scattering intensity at $q = 0.01$ Å^{-1} normalized by the real c_p after LLPS for samples with initial c_p of 31.1 mg/ml and 47.8 mg/ml. C) Calculated structure factors for a fixed protein volume fraction for different τ values. D) The reduced second virial coefficient B_2/B_2^{HS} as a function of the number of yttrium ions per protein for three different initial c_p of 31.1 mg/ml (black boxes), 47.8 mg/ml (red circles), 74.0 mg/ml (blue triangles). For comparison, the values from light scattering are also plotted as magenta inverted triangles.

indicates a similar interaction potential for proteins in the clustering phase and at the binodal.

5. Conclusion

From the perspective of colloid theory, a metastable liquid–liquid phase separation (LLPS) is caused by a strong attractive potential with a range much smaller than the effective hard sphere particle diameter σ [10,13,14]. Regardless of the precise origin of the short-ranged attraction between the proteins, its presence is essential for the LLPS: the loss of entropy in the high density phase, compared to the corresponding entropy in the low density phase, has to be compensated by the increase in internal energy due to the attraction. Mechanical equilibrium at coexistence implies that the osmotic pressure in the high density phase is equally low as in the low density phase. This can only be achieved by a sufficiently negative value of B_2. In fact it has been predicted theoretically that a reduced second virial coefficient of $B_2/B_2^{HS} < -1.5$ is required for the occurrence of a LLPS [16,18].

Table 2
Reduced second virial coefficient B_2/B_2^{HS} determined from SAXS measurements for samples located at the LLPS binodal.

c_{HSA} [mg/ml]		c_{YCl_3} [mM]		Y^{3+}/protein[b]	τ	B_2/B_2^{HS}[c]	u_0 [k_BT]
Initial c_p	Real c_p[a]	Initial c_s	Real c_s[a]				
47.8	26.2	6.0	3.3	8.4	0.065	−2.82	4.18
47.8	17.1	8.0	2.6	10.2	0.057	−3.35	4.31
47.8	16.0	10.0	4.1	17.2	0.059	−3.24	4.28
47.8	19.5	14.0	8.1	27.5	0.066	−2.79	4.17
47.8	24.2	18.0	11.9	32.8	0.067	−2.71	4.15
47.8	29.4	22.0	15.3	34.6	0.079	−2.17	3.99
47.8	35.4	26.0	18.5	34.8	0.084	−1.98	3.92
47.8	42.5	30.0	21.7	33.9	0.089	−1.80	3.87
47.8	46.5	32.0	24.9	35.6	0.091	−1.75	3.84

[a] Real concentration in the protein-poor phase.
[b] Ratio between real salt and protein concentration in the protein-poor phase.
[c] The error from data fitting is in general below 1%, but we estimate the systematic absolute error to these values is about ±0.10.

Our results from static light scattering (SLS) (Fig. 2B) suggest that while the experimental values for B_2/B_2^{HS} from SLS have to be corrected using the molecular volume of protein, where an effective volume accounting also for hydration and non-sphericity should be considered for the calculation, the final results agree reasonably well with those determined by SAXS. This finding means that the strength of the attractive potential, as measured by B_2/B_2^{HS}, for samples located in the clustering region is similar to the samples at the binodal of LLPS. This result is consistent with the theoretical prediction that the LLPS requires both chemical and mechanical equilibria. While protein solutions with the constant composition share a similar chemical potential, the difference in mechanical pressure, or volume fraction, determines their different states. For most of the experimental results including our previous work [27], the B_2/B_2^{HS} values are determined for conditions near the binodal of the phase boundary. For these samples, values of B_2/B_2^{HS} below the critical value of −1.5 are expected and observed.

In addition, anisotropic interactions can vary the picture considerably. Proteins in our system behave more like as patchy particles since the interaction is dominated by the bridging effect of metal ions [28,52]. The critical value of B_2/B_2^{HS} is not constant when the interaction potential is

changed from isotropic to anisotropic [38]. Even for spherically symmetric potentials there is a variation in the critical value of B_2/B_2^{HS} depending on details of the potential. If the potential is patchy then it is more than likely that the critical value of B_2/B_2^{HS} is lower than −1.5.

In this study, we have successfully studied a series of samples with an initial c_p of 47.8 mg/ml, which extend from binodal to the point very close to the upper critical point in the (c_p, c_s) plane. The evaluated B_2/B_2^{HS} from SAXS shown in Table 2 suggest that, approaching the critical point, the values of B_2/B_2^{HS} become less negative, indicating weaker attraction. The minimum attraction in this series gives B_2/B_2^{HS} of −1.75, slightly lower than the predicted value of −1.5. This result may suggest that either the sample is still not exactly at the critical point, or indeed the anisotropic interaction of our system leads to a lower value of B_2/B_2^{HS} at the critical point.

Acknowledgments

We are grateful for the valuable discussions and comments from T. Narayanan (ESRF, Grenoble, France). We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG). F. R.-R. acknowledges a fellowship of the Studienstiftung des deutschen Volkes.

References

[1] J. Gunton, A. Shiryayev, D. Pagan, Protein Condensation: Kinetic Pathways to Crystallization and Disease, Cambridge University Press, 2007.
[2] O. Annunziata, O. Ogun, G.B. Benedek, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 970–974.
[3] N.R. Asherie, Methods 34 (2004) 266–272.
[4] O. Galkin, K. Chen, R.L. Nagel, R.E. Hirsch, P.G. Vekilov, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 8479–8483.
[5] M. Muschol, F. Rosenberger, J. Chem. Phys. 107 (1997) 1953–1962.
[6] A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. King, G.B. Benedek, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 6116–6120.
[7] P.G. Vekilov, Cryst. Growth Des. 4 (2004) 671–685.
[8] Y. Wang, A. Lomakin, J.J. McManus, O. Ogun, G.B. Benedek, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 13282–13287.
[9] J.F. Lutsko, G. Nicolis, Phys. Rev. Lett. 96 (2006) 046102.
[10] P.R.t. Wolde, D. Frenkel, Science 277 (1997) 1975–1978.
[11] D. Rosenbaum, P. Zamora, C. Zukoski, Phys. Rev. Lett. 76 (1996) 150–153.
[12] V.J. Anderson, H.N.W. Lekkerkerker, Nature 416 (2002) 811–815.
[13] W. Poon, J. Phys. Condens. Matter 14 (2002) R859.
[14] N.R. Asherie, A. Lomakin, G.B. Benedek, Phys. Rev. Lett. 77 (1996) 4832–4835.
[15] M.H.J. Hagen, D. Frenkel, J. Chem. Phys. 101 (1994) 4093–4097.
[16] M.G. Noro, D. Frenkel, J. Chem. Phys. 113 (2000) 2941–2944.
[17] A. George, W.W. Wilson, Acta Crystallogr. D 50 (1994) 361–365.
[18] G.A. Vliegenthart, H.N.W. Lekkerkerker, J. Chem. Phys. 112 (2000) 5364–5369.
[19] F. Zhang, M.W.A. Skoda, R.M.J. Jacobs, R.A. Martin, C.M. Martin, F. Schreiber, J. Phys. Chem. B 111 (2007) 251–259.
[20] F. Zhang, M.W.A. Skoda, R.M.J. Jacobs, S. Zorn, R.A. Martin, C.M. Martin, G.F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher, F. Schreiber, Phys. Rev. Lett. 101 (2008) 148101.
[21] F. Zhang, S. Weggler, M.J. Ziller, L. Ianeselli, B.S. Heck, A. Hildebrandt, O. Kohlbacher, M.W.A. Skoda, R.M.J. Jacobs, F. Schreiber, Proteins 78 (2010) 3450–3457.
[22] F. Zhang, F. Roosen-Runge, M.W.A. Skoda, R.M.J. Jacobs, M. Wolf, P. Callow, H. Frielinghaus, V. Pipich, S. Prevost, F. Schreiber, Phys. Chem. Chem. Phys. 14 (2012) 2483–2493.
[23] L. Ianeselli, F. Zhang, M.W.A. Skoda, R.M.J. Jacobs, R.A. Martin, S. Callow, S. Prévost, F. Schreiber, J. Phys. Chem. B 114 (2010) 3776–3783.
[24] F. Roosen-Runge, B.S. Heck, F. Zhang, O. Kohlbacher, F. Schreiber, J. Phys. Chem. B 117 (2013) 5777–5787.
[25] F. Zhang, F. Roosen-Runge, A. Sauter, M. Wolf, R.M.J. Jacobs, F. Schreiber, Pure Appl. Chem. 86 (2014) 191–202.
[26] F. Zhang, G. Zocher, A. Sauter, T. Stehle, F. Schreiber, J. Appl. Cryst. 44 (2011) 755–762.
[27] F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M.W.A. Skoda, R.M.J. Jacobs, M. Sztucki, F. Schreiber, Soft Matter 8 (2012) 1313–1316.
[28] F. Zhang, F. Roosen-Runge, A. Sauter, R. Roth, M.W.A. Skoda, R. Jacobs, M. Sztucki, F. Schreiber, Faraday Discuss. 159 (2012) 313–325.
[29] D. Soraruf, F. Roosen-Runge, M. Grimaldo, F. Zanini, R. Schweins, T. Seydel, F. Zhang, M. Oettel, F. Schreiber, Soft Matter 10 (2013) 894–902.
[30] R. Roth, R. Evans, A.A. Louis, Phys. Rev. E. 64 (2001) 051202.
[31] D.J. Ashton, N.B. Wilding, R. Roth, R. Evans, Phys. Rev. E. 84 (2011) 061136.
[32] W.G. McMillan Jr., J.E. Mayer, J. Chem. Phys. 13 (1945) 276–305.
[33] M. Dijkstra, R. van Roij, R. Evans, Phys. Rev. E. 59 (1999) 5744–5771.
[34] R.J. Baxter, J. Chem. Phys. 49 (1968) 2770–2774.
[35] R.O. Watts, D. Henderson, R.J. Baxter, Adv. Chem. Phys. 21 (1971) 421–430.
[36] M. Miller, D. Frenkel, Phys. Rev. Lett. 90 (2003) 135702.
[37] N. Kern, D. Frenkel, J. Chem. Phys. 118 (2003) 9882–9889.
[38] G. Foffi, F. Sciortino, J. Phys. Chem. B 111 (2007) 9702–9705.
[39] D.C. Carter, J.X. Ho, Adv. Protein Chem. 45 (1994) 153–203.
[40] W. Brown, Light Scattering: Principles and Development, Monographs on the Physics and Chemistry of Materials, Clarendon Press, 1996.
[41] H. Bajaj, V.K. Sharma, D.S. Kalonia, Biophys. J. 87 (2004) 4048–4055.
[42] T. Narayanan, Soft matter characterization, chap, Synchroton Small-Angle X-Ray Scattering, Springer, 2008. 899ff.
[43] O. Glatter, O. Kratky, Small Angle X-Ray Scattering, Academic Press, 1982.
[44] D.I. Svergun, M.H.J. Koch, Rep. Prog. Phys. 66 (2003) 1735–1782.
[45] J.-P. Hansen, I.R. McDonald, Theory of Simple Liquids, 3rd edn Academic Press, 2006.
[46] J.S. Pedersen, Adv. Colloid Interface Sci. 70 (1997) 171–210.
[47] S.-H. Chen, Annu. Rev. Phys. Chem. 37 (1986) 351–399.
[48] L.A. Feigin, D.I. Svergun, Structure Analysis by Small-Angle X-Ray and Neutron Scattering, Plenum Press, New York, 1987.
[49] A. Isihara, J. Chem. Phys. 18 (1950) 1446–1449.
[50] F. Roosen-Runge, M. Hennig, F. Zhang, R.M.J. Jacobs, M. Sztucki, H. Schober, T. Seydel, F. Schreiber, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 11815–11820.
[51] M. Heinen, F. Zanini, F. Roosen-Runge, D. Fedunova, F. Zhang, M. Hennig, T. Seydel, R. Schweins, M. Sztucki, M. Antalik, F. Schreiber, G. Nägele, Soft Matter 8 (2012) 1404–1419.
[52] F. Roosen-Runge, F. Zhang, F. Schreiber, R. Roth, Proteins as Particles With Attractive Patches, 2014. (under review).