Soft Matter

Cite this: *Soft Matter*, 2012, **8**, 1313

www.rsc.org/softmatter

COMMUNICATION

Charge-controlled metastable liquid–liquid phase separation in protein solutions as a universal pathway towards crystallization†

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

Received 20th October 2011, Accepted 1st December 2011
DOI: 10.1039/c2sm07008a

We demonstrate that a metastable liquid–liquid phase separation (LLPS) in protein aqueous solutions can be induced by multivalent metal ions at room temperature. We determine the salt and protein partitioning in the two coexisting phases. The structure factor obtained by small angle X-ray scattering provides direct evidence for a short-ranged attraction, which leads to the metastability of the LLPS. An extended phase diagram with three control parameters (temperature, protein and salt concentration) provides a conclusive physical picture consistent with a criterion for the second virial coefficient. The presented isothermal control mechanism of the phase behavior opens new perspectives for the understanding of controlled phase behavior in nature. Furthermore, we discuss the application of this framework in predicting and optimizing conditions for protein crystallization.

Metastable liquid–liquid phase separation (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[2–4] and fibers responsible for sickle cell anemia and Alzheimer's disease,[2,5] and changes on the pathways of protein crystallization.[7,9,10] Although claimed to be universal, a metastable LLPS has so far been observed only in a few protein systems due to the rather small accessible temperature window in conventional approaches.[1,2,6] Proteins, like colloids, in solution often interact *via* effective interactions caused by other components of the solution, *i.e.* the solvent, salt, *etc.* Changing the solvent conditions alters the resulting interactions to a large extent. Usually the range of the effective attraction for colloids or proteins in solutions is shorter than the size of the particles, which leads to a metastable LLPS (corresponding to the gas-liquid phase transition in one-component systems).[1,9,11–14] Theory, simulations, and experiments have predicted that density fluctuations near LLPS can significantly reduce the energy barrier of protein condensation including crystallization, *i.e.* in addition to the classical nucleation mechanism, a two-step nucleation mechanism has been proposed.[7,9–11,14] Experimental and theoretical studies of colloid-polymer systems provide insight into the phase behavior of a system with short-ranged attractions, where the range and the strength of attraction can be tuned by the molecular weight and concentration of the polymer.[13,15] While for conventional colloids it is relatively easy to induce attractions with polymer, in protein solutions, where charge effects are ubiquitous, it seems natural to employ charges to tune the interactions. It is not obvious *a priori*, though, whether the predictions for conventional colloids apply to proteins with their discrete charge pattern, non-spherical shape and softness.

In this Letter, we demonstrate that trivalent salts can be used to optimize crystallization conditions for globular proteins along general physical arguments. We first present the phase diagram of human serum albumin (HSA), a medium size globular protein very abundant in blood with a molecular weight of 67 kDa, as a function of the concentration of protein, c_p, and the concentration of the trivalent salt YCl_3, c_s, at 20 °C (Fig. 1a). Trivalent counterions have been demonstrated to be an efficient method for tuning interactions in protein solutions by the binding of cations to acidic residues (Asp and Glu) on the protein surface, causing a charge-inversion of the protein,[16,17] whereas mono- and divalent ions do not cause such phase behavior in general. In the experimental range (3 to 80 mg mL^{-1}), the (c_p, c_s) plane is divided into three regimes, with regime I and III corresponding to a single-phase solution and regime II a two-phase state.[16] Here, we focus on the condensed phase in regime II, which is initially turbid (Fig. 1a). Closer examination by optical microscopy (Fig. 1b) reveals tiny droplets of the protein-rich phase suspended in solution. These can merge and grow, indicating a LLPS. A similar LLPS induced by YCl_3 is observed in bovine serum albumin (BSA) and β-lactoglobulin solutions.[19] Thus, LLPS is not specific for HSA, but a universal phenomenon for negatively charged proteins in the presence of trivalent cations, which makes the approach discussed here applicable to, in principle, 46% of the entire protein family.[17]

For 31 mg mL^{-1} HSA solutions at 20 °C, the LLPS occurs at a salt concentration c_s between 3 and 20 mM (Fig. 1a). The partitioning of both protein and salt into two coexisting phases is determined by X-ray and ultra-violet light absorption.[20] In Fig. 1a, the pairs of

[a] *Universität Tübingen, Institut für Angewandte Physik, D-72076 Tübingen, Germany*
[b] *Institut für Theoretische Physik, Universität Erlangen-Nürnberg, D-91058 Erlangen, Germany*
[c] *ISIS, Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK*
[d] *Chemistry Research Laboratory, University of Oxford, Oxford, OX1 3QZ, UK*
[e] *European Synchrotron Radiation Facility, F-38043 Grenoble Cedex 9, France*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2sm07008a

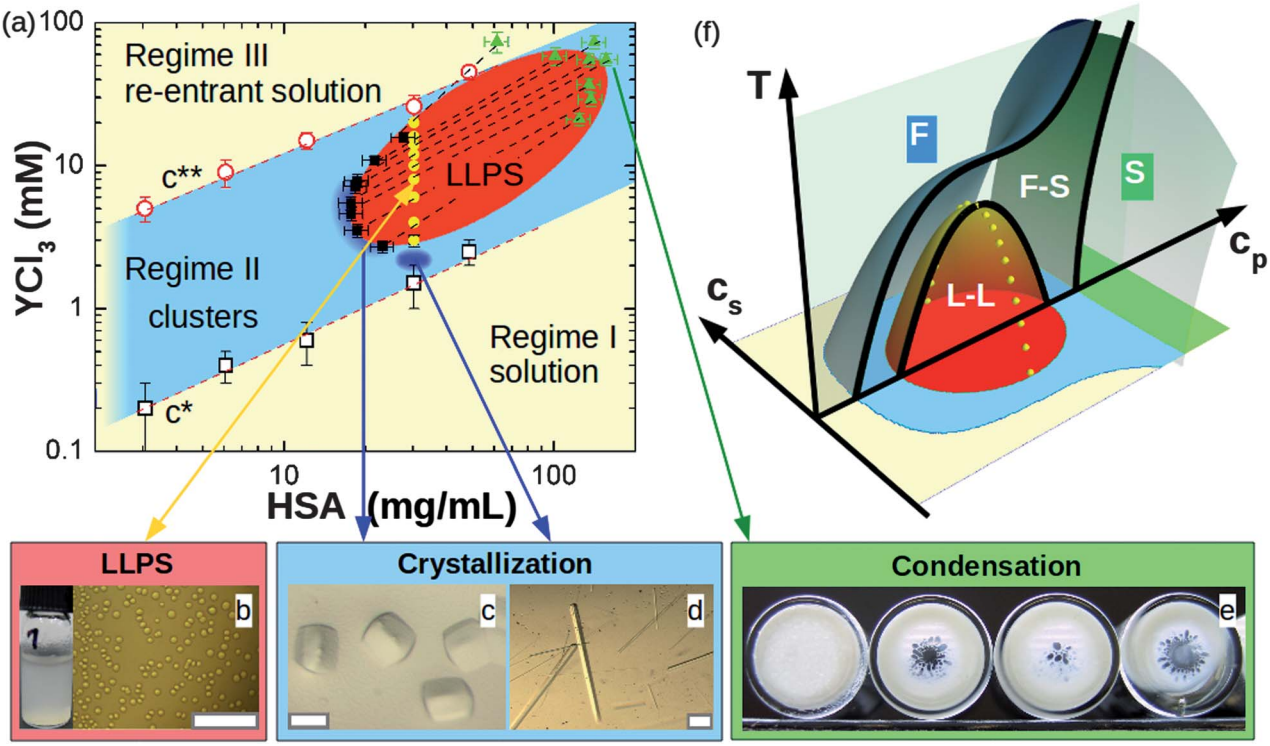

Fig. 1 Tuning interactions in protein solutions towards controlled protein crystallization (all scale bars correspond to 0.1 mm): (a) Phase diagram of HSA controlled by YCl_3 in the (c_p, c_s) plane. The open symbols with error bar represent the boundary between the regimes as determined by optical transmission, while the solid symbols (red area) denote coexisting liquid phases. (b) A typical optical microscopy image of a freshly prepared sample with 31.0 mg mL^{-1} HSA and 4.0 mM YCl_3 shows small droplets of protein-rich phases, which coalesce proving that a non-arrested LLPS occurs. (c + d) Crystallization with different growth mechanisms is observed in the dilute coexisting phase (c) and in the region slighly below the LLPS boundary (c_s = 2.0 mM) (d). (e) Amorphous aggregation in the protein-rich coexisting phase after storage at 20 °C for two weeks indicates the general metastability of the regime II to aggregation (pictures for initial HSA concentration 31.0 mg mL^{-1} with 4, 6, 10 and 15 mM YCl_3). (f) Sketch of a phase diagram with three control parameters: temperature, T, protein and salt concentration, c_p and c_s. The phase behavior can be connected to the phase diagram established in colloidal systems with a short-range attraction. Importantly, the two crystallization areas (c and d) at the same temperature are consistent with predictions for optimized crystallization conditions around the LLPS critical point (d) and at the dilute coexisting phase at a suitable attraction strength (c).[18] The presented phase diagram with T and c_s as control parameter in addition to c_p thus proves the metastable LLPS to be a universal pathway towards crystallization.

coexisting phases ($c_p^{(1)}$, $c_s^{(1)}$) and ($c_p^{(2)}$, $c_s^{(2)}$) (connected by dotted tie-lines) define the coexistence region (red ellipsoidal area). The LLPS occurs within a closed area in the $c_p \sim c_s$ plane within Regime II. While inside the coexistence we find LLPS, the condensed phase of Regime II outside the area is populated with amorphous clusters.

Also the stability of the L–L coexistence phase, which is controlled by the effective attraction range, supports this picture as follows. The freshly prepared protein-rich phase is transparent with light yellow color. After several days up to weeks of storage at 20 °C, the dense liquid eventually becomes gel-like (Fig. 1e). For the protein-poor phase, crystals can grow from these solutions (Fig. 1c). These observations indicate that the LLPS is ultimately metastable: the free energy of the dense liquid phase is lower than that of the initial solution, whereas it is higher than that of the solid state (aggregates or gel).[7] The metastability of the LLPS induced by multivalent counterions is similar to the thermally-driven LLPS in concentrated protein solutions, such as lysozyme, where the protein-rich phase decays into a solid phase over time.[8,21] These observations are consistent with the theoretical predictions that short-ranged attractions in protein solutions cause the L–L coexistence curve to move below the gas–solid curve and the LLPS becomes metastable.[8,9,22]

From the perspective of colloid theory, a metastable LLPS is caused by a strong attractive potential with a range much smaller than the effective hard sphere particle diameter.[9,13,14] 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.

Using the virial expansion of the osmotic pressure, one obtains the second viral coefficient B_2 as a measure of the integrated strength of the interaction ($\beta = 1/(k_B T)$):

$$\beta P(\rho) \approx \rho + B_2\rho^2 + B_3\rho^3 + \ldots \quad (1)$$

$$B_2 = 2\pi \int_0^\infty \mathrm{d}r\, r^2 \left[1 - \exp\left(-\beta V_{\mathrm{eff}}(r)\right)\right]. \quad (2)$$

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 was observed that a reduced second virial coefficient of $B_2/B_2^{HS} < -1.5$ is required for the occurence of a LLPS.[18,20,23]

In order to further understand the multivalent cation induced LLPS on the molecular level, the effective protein-protein interaction has been investigated by SAXS.[16,20,24–26] Typical SAXS curves of the protein-poor phase after the LLPS are shown in Fig. 2. Data fitting of a series of sample solutions with $c_p = 6.0$ mg mL^{-1} gives a form factor of $1.7 \times 5.3 \times 5.3$ nm^3[20] resulting in an effective sphere diameter of $\sigma = 8.1$ nm. To further quantify the attractive potential, a short-ranged square well structure factor was used to fit the data (Fig. 2),[25,27] accounting for steric and short-range interaction:

$$\beta V_{\mathrm{eff}}(r) = \begin{cases} \infty & 0 < r < \sigma \\ \ln\left(12\ \tau\ \dfrac{\Delta}{\sigma+\Delta}\right) & \sigma < r < \sigma + \Delta \\ 0 & \sigma + \Delta < r \end{cases} \qquad (3)$$

where Δ is the width of the attractive well. In the limit $\Delta \to 0$ the well-known case of sticky hard spheres (SHS) is recovered and only the stickiness parameter τ is required to account for the attraction and relates directly to the virial coefficient

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

Within the SHS model, the critical point of the LLPS occurs at a stickiness parameter of $\tau_c = (2-\sqrt{2})/6 \approx 0.0976$,[28] which corresponds to $B_2(\tau = \tau_c)/B_2^{HS} \approx -1.56$ and thus satisfies the criterion for the second virial coefficient.[18,23]

For the SAXS data fitting, Δ was fixed to 0.02 σ to prevent artefactial coupling with τ. The obtained τ values (see inset of Fig. 2) are generally below τ_c for YCl_3 concentrations between 3 and 20 mM, which is consistent with the observed LLPS. The related interaction potential has a depth of $\sim 4\ k_BT$. Approaching the upper and lower LLPS boundary for the salt concentration, τ increases towards τ_c, reflecting the vicinity to a critical point as expected from Fig.1f.

The absolute scattering at low q, $I(0)$, adds to this consistent physical picture in a model-free way. $I(0)$ is determined by the compressibility χ_T, since $S(q \to 0) = k_BT\rho\ \chi_T$.[29] The compressibility χ_T diverges at the spinodal line. Thus, the closer the coexisting densities are, the closer binodal and spinodal lines are, and 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 easily by comparing the SAXS intensity for small values of q (Fig. 2, Inset) with the LLPS phase behavior shown in Fig. 1. As the YCl_3 concentration, for a fixed c_p of 31.0 mg mL^{-1}, is increased up to 3 mM, the system phase-separates. For this c_s the system is close to the lower critical point of the LLPS, which results in large values of χ_T and $S(q \to 0)$. As the YCl_3 concentration is increased further up to 10 mM, the coexistence region broadens, causing χ_T and $S(q \to 0)$ to decrease. In the range between 10 to 30 mM of YCl_3 the trend is reversed: χ_T and $S(q \to 0)$ increase again, until the system mixes for YCl_3 concentrations above 30 mM.

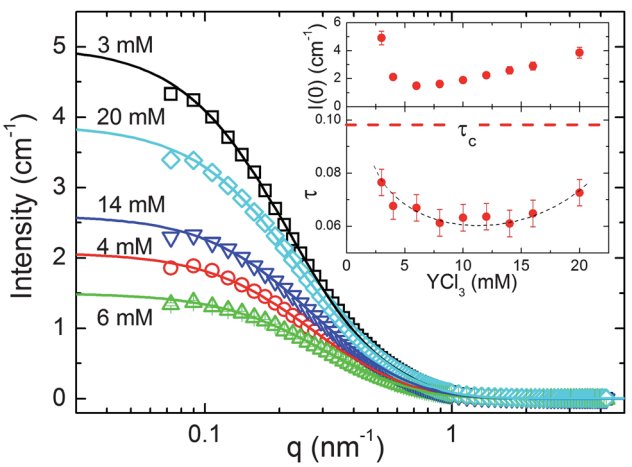

Fig. 2 SAXS data with model fitting for the protein-poor phases from sample solutions with initial c_p of 31.0 mg mL^{-1} after LLPS. Only every second data point is plotted for clarity. The insets show τ and $I(0)$ as a function of c_s.

The phase behavior shown in Fig. 1a can now be understood along the usual phase diagram for colloids with short-range attraction, when taking into account that protein and bound Y^{3+} ions build effective complexes with a tunable (short-range) attraction. Importantly, this binding mechanism is generally present in protein systems, rendering the presented approach a universal pathway towards controlling phase behavior in protein solutions.

In Fig. 1f, the black lines present a typical phase diagram in the (T, c_p) plane for such a protein-ion complex. By changing the ratio between protein and bound Y^{3+} ions, the interaction strength varies, which is reflected in a shift of the critical point of the LLPS (visualized for selected complexes by the yellow points). Regarding an isothermal plane (c_s, c_p) leads back to the phase diagram presented in Fig. 1a. Note that the two coexistence lines for the solid-fluid transition at the dilute limit should merge and for high protein concentration, the phase behavior becomes more complex.[30]

The critical points of LLPS can be nicely described using the thermodynamic criterion based on B_2, which has been used as a predictor in protein crystallization. George and Wilson observed that B_2 falls in a narrow range for protein crystallization.[31] Subsequent theoretical work by Vliegenthart and Lekkerkerker demonstrated that B_2 has a nearly constant value at the critical point 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,[9] or below the critical point but near the protein-poor phase boundary where crystals grow *via* a two-step procedure.[18]

The growth of high quality protein single crystals (Fig. 1c and d) supports these theoretical predictions. Importantly, by using multivalent ions, both mechanisms can be observed at the same temperature: Fig. 1c corresponds to the case of a two-step nucleation and Fig. 1d represents nucleation close to the LLPS critical point. By varying temperature, crystallization conditions can be further optimized. While a detailed crystallographic study is beyond the scope of this Letter, several proteins have been previously crystallized using YCl_3 as a crucial additive without clarifying its exact role.[32,33] Fig. 1 now provides a context for the physical understanding.

The exact physical origin of the short-range attraction remains to be identified – van der Waals, solvation and structural forces, which have a range in the order of 1.1 σ (ref. 11), can contribute to the short-range attraction. Effective spheres as representation of proteins has proved successful for data modeling and the description of the dynamics of protein solutions.[34–36] However, protein-Y^{3+} complexes with a non-spherical shape, nonhomogeneous surface charge pattern and complex hydration properties are expected to show highly directional, short-ranged interactions. Thus, a more detailed theoretical analysis of this system could be achieved within the framework of patchy particles,[37–39] which could also explain the rather low volume fractions at the LLPS critical point below 10%.[40]

However, even without the precise knowledge of the microscopic mechanism, multivalent metal ions can be successfully used to control the macroscopic phase behavior of globular proteins. The observed phase behavior can be understood and described even quantitatively based on fundamental thermodynamic principles and colloid theory. In summary, we have presented a theoretically consistent and comprehensive understanding of the controlled optimization of protein crystallization around the metastable LLPS in protein solutions solely induced by multivalent ions. The results suggest a universal applicability of the approach and open the field for further systematic studies of protein phase behavior.

The authors profited from discussions with C.P. Royall (Bristol, UK), S. Dietrich (Stuttgart, Germany) and T. Narayanan (ESRF, Grenoble, France). We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG).

References

1 J. Gunton, A. Shiryayev and D. Pagan, *Protein condensation: kinetic pathways to crystallization and disease*, Cambridge Univ Pr, 2007.
2 O. Annunziata, O. Ogun and G. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 970.
3 A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. King and G. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 6116.
4 Y. Wang, A. Lomakin, J. J. McManus, O. Ogun and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 13282–13287.
5 O. Galkin, K. Chen, R. Nagel, R. Hirsch and P. Vekilov, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 8479.
6 N. Asherie, *Methods*, 2004, **34**, 266–272.
7 P. Vekilov, *Cryst. Growth Des.*, 2004, **4**, 671–685.
8 M. Muschol and F. Rosenberger, *J. Chem. Phys.*, 1997, **107**, 1953.
9 P. ten Wolde and D. Frenkel, *Science*, 1997, **277**, 1975.
10 J. Lutsko and G. Nicolis, *Phys. Rev. Lett.*, 2006, **96**, 46102.
11 D. Rosenbaum, P. Zamora and C. Zukoski, *Phys. Rev. Lett.*, 1996, **76**, 150–153.
12 V. Anderson and H. Lekkerkerker, *Nature*, 2002, **416**, 811–815.
13 W. Poon, *J. Phys.: Condens. Matter*, 2002, **14**, R859.
14 N. Asherie, A. Lomakin and G. Benedek, *Phys. Rev. Lett.*, 1996, **77**, 4832–4835.
15 L. Belloni, *J. Phys.: Condens. Matter*, 2000, **12**, R549–R587.
16 F. Zhang, M. Skoda, R. Jacobs, S. Zorn, R. Martin, C. Martin, G. Clark, S. Weggler, A. Hildebrandt and O. Kohlbacher, *et al.*, *Phys. Rev. Lett.*, 2008, **101**, 148101.
17 F. Zhang, S. Weggler, M. J. Ziller, L. Ianeselli, B. S. Heck, A. Hildebrandt, O. Kohlbacher, M. W. A. Skoda, R. M. J. Jacobs and F. Schreiber, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 3450–3457.
18 G. Vliegenthart and H. Lekkerkerker, *J. Chem. Phys.*, 2000, **112**, 5364.
19 F. Zhang, G. Zocher, A. Sauter, T. Stehle and F. Schreiber, *J. Appl. Crystallogr.*, 2011, **44**, 755–762.
20 supplementary material, online available.
21 M. Shah, O. Galkin and P. Vekilov, *J. Chem. Phys.*, 2004, **121**, 7505.
22 D. Petsev, X. Wu, O. Galkin and P. Vekilov, *J. Phys. Chem. B*, 2003, **107**, 3921–3926.
23 M. Noro and D. Frenkel, *J. Chem. Phys.*, 2000, **113**, 2941.
24 F. Zhang, M. Skoda, R. Jacobs, R. Martin, C. Martin and F. Schreiber, *J. Phys. Chem. B*, 2007, **111**, 251–259.
25 S. Kline, *J. Appl. Crystallogr.*, 2006, **39**, 895–900.
26 L. Ianeselli, F. Zhang, M. Skoda, R. Jacobs, R. Martin, S. Callow, S. Pre-vost and F. Schreiber, *J. Phys. Chem. B*, 2010, **114**, 3776.
27 S. Menon, C. Manohar and K. Rao, *J. Chem. Phys.*, 1991, **95**, 9186.
28 R. Baxter, *J. Chem. Phys.*, 1968, **49**, 2770.
29 J.-P. Hansen and I. McDonald, *Theory of simple liquids*, Academic Press, 1986.
30 E. Zaccarelli, *J. Phys.: Condens. Matter*, 2007, **19**, 323101.
31 A. George and W. W. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 361–365.
32 S. Bouyain, P. Longo, S. Li, K. Ferguson and D. Leahy, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15024.
33 H. Xu, M. Lambert, V. Montana, K. Plunket, L. Moore, J. Collins, J. Oplinger, S. Kliewer, R. Gampe and D. McKee, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 13919.
34 B. R. Jennings and K. Parslow, *Proc. R. Soc. London, Ser. A*, 1988, **419**, 137–149.
35 F. Roosen-Runge, M. Hennig, F. Zhang, R. M. J. Jacobs, M. Sztucki, H. Schober, T. Seydel and F. Schreiber, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 11815–11820.
36 M. Heinen, F. Zanini, F. Roosen-Runge, D. Fedunová, F. Zhang, M. Hennig, T. Seydel, R. Schweins, M. Sztucki and M. Antalík, *et al.*, *Soft Matter*, 2012, DOI: 10.1039/c1sm06242e, online first .
37 M. S. Wertheim, *J. Stat. Phys.*, 1984, **35**, 19–34.
38 G. Foffi and F. Sciortino, *J. Phys. Chem. B*, 2007, **111**, 9702–9705.
39 A. B. Pawar and I. Kretzschmar, *Macromol. Rapid Commun.*, 2010, **31**, 150–168.
40 E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli and F. Sciortino, *Phys. Rev. Lett.*, 2006, **97**, 168301.

