

Protein–polymer mixtures in the colloid limit: Aggregation, sedimentation, and crystallization

Cite as: J. Chem. Phys. 155, 114901 (2021); <https://doi.org/10.1063/5.0052122>

Submitted: 29 March 2021 . Accepted: 09 August 2021 . Published Online: 15 September 2021

Rui Cheng, Jingwen Li, Iotzin Ríos de Anda, Thomas W. C. Taylor, Malcolm A. Faers, J. L. Ross Anderson, Annela M. Seddon,  C. Patrick Royall, et al.



[View Online](#)



[Export Citation](#)



[CrossMark](#)

Challenge us.

What are your needs for
periodic signal detection?



Zurich
Instruments

Protein–polymer mixtures in the colloid limit: Aggregation, sedimentation, and crystallization

Cite as: J. Chem. Phys. 155, 114901 (2021); doi: 10.1063/5.0052122

Submitted: 29 March 2021 • Accepted: 9 August 2021 •

Published Online: 15 September 2021



View Online



Export Citation



CrossMark

Rui Cheng,^{1,2} Jingwen Li,^{1,2,3} Iloatzin Ríos de Anda,^{1,4} Thomas W. C. Taylor,^{1,2} Malcolm A. Faers,⁵ J. L. Ross Anderson,^{6,7,8} Annela M. Seddon,^{1,2} and C. Patrick Royall^{1,2,3,9,a)}

AFFILIATIONS

¹ HH Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, United Kingdom

² Bristol Centre for Functional Nanomaterials, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom

³ School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom

⁴ School of Mathematics, University Walk, Bristol BS8 1TW, United Kingdom

⁵ Bayer AG, 40789 Monheim am Rhein, Germany

⁶ School of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom

⁷ School of Cellular and Molecular Medicine, University Walk, Bristol BS8 1TD, United Kingdom

⁸ BrisSynBio Synthetic Biology Research Centre, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, United Kingdom

⁹ Gulliver UMR CNRS 7083, ESPCI Paris, Université PSL, 75005 Paris, France

Note: This paper is part of the JCP Special Topic on Depletion Forces and Asakura–Oosawa Theory.

a) Author to whom correspondence should be addressed: paddy.royall@espci.fr

ABSTRACT

While proteins have been treated as particles with a spherically symmetric interaction, of course in reality, the situation is rather more complex. A simple step toward higher complexity is to treat the proteins as non-spherical particles and that is the approach we pursue here. We investigate the phase behavior of the enhanced green fluorescent protein (eGFP) under the addition of a non-adsorbing polymer, polyethylene glycol. From small angle x-ray scattering, we infer that the eGFP undergoes dimerization and we treat the dimers as spherocylinders with aspect ratio $L/D - 1 = 1.05$. Despite the complex nature of the proteins, we find that the phase behavior is similar to that of hard spherocylinders with an ideal polymer depletant, exhibiting aggregation and, in a small region of the phase diagram, crystallization. By comparing our measurements of the onset of aggregation with predictions for hard colloids and ideal polymers [S. V. Savenko and M. Dijkstra, *J. Chem. Phys.* **124**, 234902 (2006) and Lo Verso *et al.*, *Phys. Rev. E* **73**, 061407 (2006)], we find good agreement, which suggests that the behavior of the eGFP is consistent with that of hard spherocylinders and ideal polymers.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0052122>

I. INTRODUCTION

Protein aggregation and crystallization have important consequences in determining their structure and function and understanding challenges ranging from condensation diseases^{1–3} to the development of new materials.⁴ Controlling their assembly into states in which their functionality can be exploited is crucial to fully realize their potential, and crystallization is fundamental to obtaining protein structure and insights into their function. Despite its importance, crystallization of proteins is one of the most complex and poorly understood topics in biology. Crystallization protocols are mainly based on trial and error assays, with a lack of

standardized approaches. In fact, in average, only 0.04% of crystallization experiments yield good quality crystals.⁸ This is due, in part, to the inherent protein shape and surface complexity as well as the dependence of protein–protein interactions on combinations of pH, temperature, and precipitants (salts and polymers).^{8–13}

Understanding crystallization is often facilitated by reference to the equilibrium phase diagram. In particular, predictions can be made with detailed conditions such as protein concentration, temperature, and pH.^{5–7} Furthermore, recent developments include an improved understanding of the role of clusters in protein crystallization,¹⁴ the effect of polymers in inducing crystallization,^{15–18} the effects of salts on crystallization^{17,19,20} and crystal growth rate,²¹

and the role of temperature in the protein phase diagram.^{22,23} An emphasis has been placed on the role of entropy in contact–contact interactions in proteins.²⁴ Perhaps unsurprisingly, fine manipulation of protein interactions is necessary for self-assembly,^{25,26} and if this can be successfully achieved and coupled with protein engineering, it is possible to manipulate proteins to enable new paths of self-assembly.^{4,27,28}

By contrast, the field of *soft matter* often operates at rather larger lengthscales than the supramolecular lengthscale of proteins. Yet, concepts inspired by soft matter have been applied to proteins with some success.^{8,29–31} Among these are *effective* interactions between the proteins that can be altered by other components, such as added salts and polymers.^{18,32–37} In this way, a soft matter perspective offers some insights to understand and quantify protein interactions and their equilibrium phase diagrams by simplified models, which might provide a systematic way to improve protein crystallization.^{8,29–31,33} Indeed, a parameter that has been used to relate protein phase behavior^{36,38–41} to that of colloidal suspensions^{42–44} is the second viral coefficient, which can be calculated experimentally from osmotic pressure,⁴⁰ static light scattering, dynamic light scattering, SAXS, or small-angle neutron scattering (SANS).^{45–48} While proteins often aggregate at high concentration, some do not, and indeed, interesting glassy behavior reminiscent of hard sphere colloids has been seen for concentrated solutions of eye lens α -crystallin,⁴⁶ which opens the potential for further analogies with colloidal systems.

Examples of the insights gained from this approach of comparing proteins to colloidal systems include the prediction of enhanced nucleation rates in the vicinity of a (metastable) critical point,^{49,50} which have been realized using colloids with a short-ranged attraction^{51,52} and gelation^{32,45,53–57} and so-called liquid–liquid phase separation.⁵⁸ It is also possible to control the pathway of crystallization by manipulating the interactions.⁵⁹ Analogies have also been made between proteins and colloids with so-called mermaid (short-range repulsion and long-range attraction) interactions through the discovery of finite-sized clusters,^{60–68} although the existence of the protein clusters has been questioned.⁶⁹ Meanwhile, the colloidal systems in question have been shown to exhibit much more complex behavior than was originally supposed through a fundamental breakdown in spherical symmetry in the electrostatic repulsions.^{44,70}

Moreover, anisotropy in the shape of the constituent particles is well-known to lead to a markedly different behavior in the formation of liquid crystalline phases. Such phenomena have been explored in biomolecules, in particular, with rod-like viruses.^{37,71} Proteins exhibit anisotropic shape, a non-uniform surface charge, and hydrophobic/hydrophilic patterns.^{8,40,72} This anisotropy is responsible for the directional and localized protein interactions that yield non-close packed crystal structures as well as directed self-assembly.⁷³ This more complex behavior can be captured and reproduced to some extent via *patchy particle* models, where the simulated particles include angular surface directionality of attractive short-range interactions.^{12,74,75} By changing the number, size, and specificity of said patches, the system can be optimized to fully describe the protein behavior.^{28,74,76–80}

In colloidal systems, an effective attraction between the particles can be induced by adding *non-adsorbing* polymers.^{32,37,81} Polymer-induced *protein* precipitation has been investigated

via volume exclusion interactions, i.e., depletion.^{9,82} However, the polymers can interact with the proteins, even in relatively simple “model proteins” such as lysozyme, leading to unexpected behavior.³⁷ For example, polymers can interact with positively charged amino acids (lysine, arginine, and histidine)⁸³ and/or through hydrophobic chemical interactions (for example, with $-\text{CH}_2\text{OCH}_2-$ groups)⁸⁴ both present on the protein surface. Additionally, there can be a preferential formation of hydrogen bonds between the polymer and water, which, in turn, influences protein–protein interactions.⁸⁵ These scenarios can lead to more complex interactions than the non-adsorbing polymer–protein depletion picture.³⁷

Although they are often smaller than colloids, polymers are typically rather larger than proteins,^{35,86} leading to the concept of the *protein limit*,^{87–89} where the polymers are so much larger than the proteins that the relevant lengthscale is the *intra-polymer* persistence length, rather than the polymer radius of gyration that is typically considered in the case of colloid–polymer mixtures. However, here we consider a scenario more akin to colloid–polymer mixtures, where the polymer radius of gyration R_g is *smaller* than or comparable to the protein radius.³⁷ In particular, our system of interest consists of mixtures of *enhanced* Green Fluorescent Protein (eGFP) and poly-ethylene-glycol (PEG). eGFP readily undergoes dimerization^{25,90} such that the proteins resemble short rods. By comparing our results to the colloid–polymer literature,^{92,93} we treat the proteins as spherocylinders (with dimensions deduced from x-ray scattering), in particular, as mixtures of hard spherocylinders and ideal polymers.^{91,92} In this way, we consider a model of the protein–polymer mixture where the only level at which the complexity of the system is treated is via a simplified form for the anisotropy of the proteins, namely, a spherocylinder. We thus treat the protein dimers as hard spherocylinders and the polymers as ideal polymers.

Here, we follow the literature on spherocylinder–polymer mixtures^{91,92} and express the protein aspect ratio as $L/D - 1$ in which the aspect ratio of spheres is then zero, where L is the spherocylinder length and D is the diameter. We interpolate the predictions of polymer fugacity required for polymer-induced demixing between spherocylinders with aspect ratio $L/D - 1 = 5$ ⁹² and spheres.⁹³ Remarkably, given the simplicity of the model, we find good agreement for the geometric parameters of our system with $L/D - 1 = 1.05$.

This paper is organized as follows: In Sec. II, we describe the methods of protein preparation, characterization, as well as estimation of their interactions as spherocylinders. Section III A reports our measurements of phase behavior, including aggregation and crystallization, in salt-screened and salt-free mixtures. We then compare the onset of aggregation with theory and simulation in Sec. III B, where the polymer radius of gyration at different molecular weights is fitted by interpolating predictions from hard colloids and ideal polymers.^{92,93} Finally, a discussion of our findings is presented in Sec. IV.

II. METHODOLOGY

A. Estimation of interactions between proteins

Our system is governed by two control parameters: the protein concentration and the polymer concentration. In the context of treating the system in the spirit of a colloid–polymer mixture, we consider the proteins as spherocylinders and thus the volume

fraction

$$\rho_{\text{eGFP}} = \rho_{\text{eGFP}} \left(\frac{\pi}{6} D^3 + \frac{\pi}{4} D^2 (L - D) \right), \quad (1)$$

where ρ_{eGFP} is the number density of protein dimers. We determine the protein diameter D and length L from x-ray scattering and compare our results to literature values (see Sec. II D). Our choice of spherocylinders is motivated by the literature on colloid–polymer mixtures, for which phase diagrams for hard spherocylinders plus ideal polymers are available.^{91,92,94}

Our second control parameter, polymer concentration, is expressed as the polymer fugacity z_{pol} . We make the significant assumption that the polymers can be treated as an ideal gas and then the fugacity is equal to the polymer number density in a reservoir $z_{\text{pol}} = \rho_{\text{pol}}^{\text{res}}$ in thermodynamic equilibrium with the system. We express the fugacity in dimensionless units of the number of polymer molecules per cubic diameter D^3 .

To compare with predictions from theory and computer simulation, we use the fraction of available volume α to estimate the reservoir polymer number density from that in the experimental system $\rho_{\text{pol}}^{\text{exp}}$, viz., $\rho_{\text{pol}}^{\text{exp}} = \alpha \rho_{\text{pol}}^{\text{res}}$. We use the free-volume approximation for α ,⁹⁴

$$\begin{aligned} \alpha &= (1 - \phi_{\text{eGFP}}) \exp \left[- \left\{ A \left(\frac{\phi_{\text{eGFP}}}{1 - \phi_{\text{eGFP}}} \right) + B \left(\frac{\phi_{\text{eGFP}}}{1 - \phi_{\text{eGFP}}} \right)^2 \right. \right. \\ &\quad \left. \left. + C \left(\frac{\phi_{\text{eGFP}}}{1 - \phi_{\text{eGFP}}} \right)^3 \right\} \right], \\ A &= \frac{6\gamma}{3\gamma - 1} q + \frac{3(\gamma + 1)}{3\gamma - 1} q^2 + \frac{2}{3\gamma - 1} q^3, \\ B &= \frac{1}{2} \left(\frac{6\gamma}{3\gamma - 1} \right)^2 q^2 + \left(\frac{6}{3\gamma - 1} + \frac{6(\gamma - 1)^2}{3\gamma - 1} \right) q^3, \\ C &= \frac{2}{3\gamma - 1} \left(\frac{12\gamma(2\gamma - 1)}{(3\gamma - 1)^2} + \frac{12\gamma(\gamma - 1)^2}{(3\gamma - 1)^2} \right) q^3, \end{aligned} \quad (2)$$

where $\gamma = L/D$ is the length-to-diameter ratio of spherocylinders, q is the polymer–protein size ratio of $2R_g/D$, and R_g is the radius of gyration of polymers. Below, we compare the phase behavior we obtain for our system with literature values for spherocylinder–polymer mixtures.^{91,92,94}

Our proteins carry an electrostatic charge, which we determine below (Sec. II D). To estimate the electrostatic interactions, we used a screened Coulomb (Yukawa) potential. Here, it is convenient to treat the proteins as spheres. We shall see below that although they are not spherical, the electrostatic interactions turn out to be so weak that we believe that to a large extent they can be neglected. Therefore, we estimate their strength with a spherically-symmetric approximation,

$$\beta u_{\text{yuk}}(r) = \begin{cases} \infty & \text{for } r < D, \\ \beta \epsilon_{\text{yuk}} \frac{\exp(-\kappa(r - D))}{r/D} & \text{for } r \geq D, \end{cases} \quad (3)$$

where the contact potential

$$\beta \epsilon_{\text{yuk}} = \frac{Z^2}{(1 + \kappa D/2)^2} \frac{\lambda_B}{D}, \quad (4)$$

and κ is the inverse Debye length,

$$\kappa^2 = 4\pi \lambda_B \sum_i \rho_i^{\text{ion}} (\mathcal{Z}_i^{\text{ion}})^2, \quad (5)$$

where Z is the number of elementary charges on the protein dimers, λ_B is the Bjerrum length, ρ_i^{ion} is the number density of the i th ionic species, and $\mathcal{Z}_i^{\text{ion}}$ is the valency of the i th ionic species. For the system with no added salt, the ionic strength $I = \sum_i \rho_i^{\text{ion}} (\mathcal{Z}_i^{\text{ion}})^2$ was evaluated as the sum of the ion contributions of the weak dissociation of 25 mM HEPES ($\text{pKa} = 7.66$) and the counterion contribution assuming charge neutrality. Thus, by varying protein concentrations, we obtained a range of $I = 1\text{--}4$ mM. For the system where 10 mM NaCl was added, we included to the sum the ion contribution from this salt dissociation, giving a range of $I = 15\text{--}18$ mM. Further details of the Yukawa potential are listed in Table I.

It is important to highlight, as pointed out by Roosen-Runge *et al.*,⁹⁵ that in addition to assuming a spherical shape for the proteins, also an isotropic distribution of ions on their surfaces is assumed. This is not the case for eGFP dimers; thus, the charges calculated should only be considered as effective charges suitable to describe the phenomena observed in our experiments. However, the magnitude of the charge that we determine is sufficiently small that within the DLVO treatment we employ, the electrostatic interactions are found to be very weak, so we believe that at the level of this analysis, a spherical approximation is reasonable.

B. Protein expression and purification

1. Cellular culture for the expression of eGFP

A mini-culture of competent *Escherichia coli* BL21 (DE3) previously transformed with the DNA plasmid-pET45b(+)-eGFP was prepared by inoculating 100 ml of lysogeny broth (LB) and the antibiotic carbenicillin (50 µg/ml) with an isolated *E. coli* colony. The culture was left to grow overnight (16 h) at 37 °C and 180 rpm. 2 ml of this culture was inoculated to a 1 l of LB containing the same antibiotic, which was left to grow under the same previous conditions. The optical density (OD_{600nm}) was monitored until a value of 0.5–0.6 was reached. Then, the production of eGFP was induced by adding 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). After 1 h of induction time, the temperature was changed to 28 °C and was incubated overnight. The cell culture was centrifuged at 4500 g for 15 min at 4 °C. The supernatant obtained was discarded, and the pellet was resuspended in a *lysis buffer* (20 mM imidazole, 300 mM NaCl, and 50 mM potassium phosphate at pH 8.0) and stored at –20 °C.⁹⁶

2. Purification and concentration of eGFP

Cell pellets were thawed and kept on ice, sonicated for three cycles of 30 s (Soniprep 150 plus MSE), and centrifuged at

TABLE I. Effects of adding salts on the Yukawa potential.

| NaCl concentration (mM) | κ^{-1} (nm) | κD | Z | $\beta \epsilon_{\text{yuk}}$ |
|-------------------------|--------------------|------------|------|-------------------------------|
| 0 | 4.65 | 1.51 | 0.95 | 0.0320 |
| 10 | 2.48 | 2.82 | 1.31 | 0.0322 |

18 000 rpm (Sorvall SS34 rotor) at 4 °C for 30 min. The supernatant was recovered and filtered through a 0.22 µm syringe filter (Millipore) and injected to a Ni-NTA (nickel-nitrilotriacetic acid) agarose column (Qiagen) connected to an ÄKTA START purification system (GE Healthcare), previously equilibrated with the lysis buffer mentioned. The bound eGFP was washed with the same lysis buffer to elute the rest of the unbound proteins and eGFP was later eluted with a linear gradient (0%–100%) of a 500 mM imidazole, 300 mM NaCl, and 50 mM potassium phosphate buffer at pH 8.

The recovered proteins were then further purified through size exclusion chromatography to eliminate aggregates and unfolded proteins. A single peak corresponding to a single protein size was collected. eGFP was concentrated to ~3 ml in a 25 mM Tris-Base 150 mM NaCl buffer at pH 7.4. The proteins were applied to a HiLoad Superdex 75 16/600 size exclusion column using the ÄKTA START purification system (GE Healthcare) pre-equilibrated with the same buffer. Protein elution was monitored at 280 nm.

Purified eGFP was filtered through a 0.22 µm syringe filter (Millipore) and concentrated using protein 30 kDa concentrators (ThermoFisher Scientific) at 5000 rpm and 4 °C for the time required to reach the desired volume. The protein concentration was determined by measuring the absorbance at $\lambda_{\text{eGFP}} = 488 \text{ nm}$ with a molar extinction coefficient $\epsilon_{\text{eGFP}} = 56\,000 \text{ M}^{-1} \text{ cm}^{-1}$.⁹⁷

C. Sample preparation

From small-angle x-ray scattering, the purified eGFP was shown to form dimers with a length of 8.2 nm and diameter of 4 nm (see below). Thus, we treated the eGFP molecules as spherocylinders of diameter $D = 4 \text{ nm}$ and length $L = 8.2 \text{ nm}$. We changed the protein buffer to 25 mM HEPES at pH = 7.4. A separate buffer solution with 0 or 200 mM NaCl was used as a stock solution to adjust the final protein and salt concentrations.

We carried out experiments with two different polymer sizes, in particular, polyethylene glycol (PEG, Polymer Laboratories) with molecular weights M_w of 620 and 2000. The polymer radius of gyration R_g was estimated from polymer scaling with an empirical prefactor $R_g^{\text{scale}} = 0.020M_w^{0.58}$,³⁹ leading to R_g^{scale} of 0.83 and 1.64 nm, and a polymer–protein size ratio ($q^{\text{scale}} = 2R_g^{\text{scale}}/D$) of 0.42 and 0.82 for the small and large polymers, respectively.

For each sample, we first mixed different volumes of the protein stock solution (106.4 mg/ml) with different volumes of the HEPES buffer with and without salt to complete a fixed volume of 5 µl, giving a range of protein concentrations of 0.7–30 mg/ml and a constant NaCl concentration of 10 mM (for the samples with salt). To induce effective attractions between the protein molecules, we added different amounts of PEG by weight at room temperature such that we obtained a polymer concentration between 0 and 0.8 g cm⁻³ (fugacity ~1 to 50). The samples were thoroughly shaken for 5 s by a touch-vortexer, immediately imbibed to inside-diameter 0.5 mm capillaries (CM Science) and sealed with optical adhesive (Norland Optical no. 81). Within 5 min, the different phases obtained were characterized through laser scanning confocal microscopy (Leica SP8) at an excitation wavelength of 488 nm and emission wavelength of 509 nm.

D. Characterization

1. SAXS analysis

To characterize the size and shape of the expressed and purified eGFP, we performed SAXS measurements on 25 µl of 10 mg/ml eGFP in 25 mM Tris-Base 150 mM NaCl buffer at pH 7.4 on a SAXSLAB Ganesha 300XL instrument. Samples were loaded into 1.5 mm borosilicate glass capillaries (Capillary Tube Supplies, UK) and sealed with optical adhesive under UV light (Norland 81). The wavevector k range was of 0.006–0.30 Å⁻¹. Background corrections were carried out with both an empty cell and a cell with the buffer only. The obtained data were fitted using the SasView version 4.0 software package.⁹⁸

The results are shown in Fig. 1(a). The scattering intensity, $I(k)$, is given by the product of the form factor $P(k)$ and the static structure factor $S(k)$ via

$$I(k) = \phi_{\text{eGFP}} V_{\text{eGFP}} (\Delta\rho_{\text{scat}})^2 P(k) S(k), \quad (6)$$

where V_{eGFP} is the volume of a protein dimer [Eq. (1)] and $\Delta\rho_{\text{scat}}$ is the difference in scattering length density between the proteins and their solvent.^{99–101} The scattering data were successfully fitted by a cylindrical form factor¹⁰² with a diameter of $4.0 \pm 0.02 \text{ nm}$ and a length of $8.2 \pm 0.08 \text{ nm}$ (see full parameters in Table II). These dimensions are consistent with dimers of eGFP, as illustrated

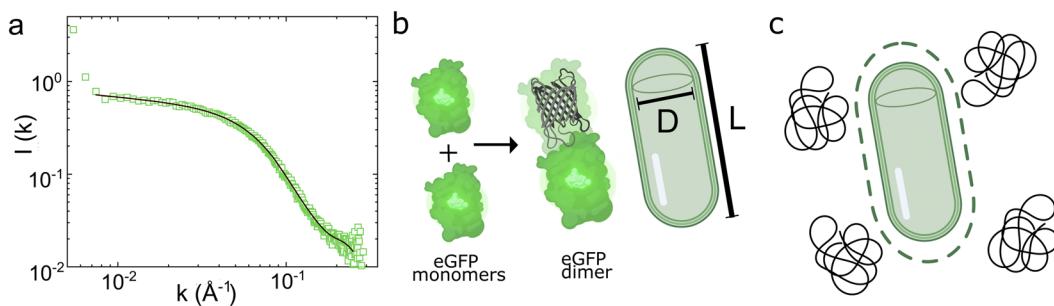


FIG. 1. Determining protein form factor with small angle x-ray scattering (SAXS). (a) SAXS scattering intensity with a cylinder model fitting (line) for 10 mg/ml protein solutions of eGFP. (b) Illustration of the dimer structure of eGFP and the spherocylinder representation used for data analysis, showing the diameter D and the length L . (c) Schematic of the eGFP depletion zone when polymers are present.

in Fig. 1(b). These results are in agreement with previous work on eGFP, where it was found that the protein exists in dimers.¹⁰³

2. Electrophoretic mobility measurement

We performed electrophoretic mobility μ_e measurements on 1 ml protein solutions of 2 mg/ml at 20 °C in a NaCl 10 mM solution using a Zetasizer Nano ZS (Malvern, UK) at a detector angle of 13° and a 4 mW 633 nm laser beam to determine the charge of eGFP following Roosen-Runge *et al.*⁹⁵ Care was taken in order to have the same pH with the buffer used in phase diagram determination. By using electrophoretic light scattering (ELS) via phase analysis light scattering (M3-PALS), the electrophoretic mobility μ_e of eGFP was determined as an external electric field is applied.

From this, we obtained the zeta potential ζ for a spherical particle with diameter D using

$$\mu_e = \frac{2\epsilon\zeta f(\kappa D/2)}{3\eta}, \quad (7)$$

where ϵ_r and η are the dielectric constant and the viscosity of the solvent, respectively, and $f(\kappa D/2)$ is the Henry function evaluated at $\kappa D/2$. The relation between the surface charge density σ and the reduced zeta potential $\tilde{\zeta} = (e\zeta)/(2k_B T)$ is

$$\frac{\sigma e}{k_B T} = 2\epsilon_r \kappa \left[\sinh^2\left(\frac{\tilde{\zeta}}{2}\right) + \frac{4}{\kappa D} \tanh^2\left(\frac{\tilde{\zeta}}{4}\right) \right]^{1/2} + \frac{32}{(\kappa D)^2} \ln\left(\cosh\left(\frac{\tilde{\zeta}}{4}\right)\right). \quad (8)$$

Finally, we obtain the total charge using $\mathcal{Z}e = \pi D^2 \sigma$. The zeta potential value measured was $\zeta = -7.02$ mV, which corresponds to a charge number of $\mathcal{Z} = 0.95$. We list the parameters for the Yukawa potential in Table I.

III. RESULTS

We divide our results as follows: First, a phase diagram is presented for the eGFP-PEG620 system, showing the fluid-aggregation

transition in Sec. III A. To check for any residual effects of protein charges, the comparison between the salt-screened and salt-free system is discussed. We increased the polymer molecular weight to obtain a larger size ratio (using PEG2000), investigating the effects of polymer size on the phase boundary. In Sec. III B, we consider a spherocylinder-sphere system of $L/D - 1 = 1.05$. The polymer radius of gyration is fitted by interpolating between theoretical and computer simulation predictions.^{92,93} Finally, the protein crystallization, formed through depletion attractions with polymer, is discussed in Sec. IV.

A. Phase behavior

1. Salt-screened system

The phase diagram with different states as determined from images from confocal microscopy for the eGFP-PEG620 (small polymer) system with 10 mM of added NaCl salt is shown in Fig. 2. The phase diagram is presented in the plane of protein volume fraction (ϕ_{eGFP}) and polymer fugacity z_{pol} . The phase boundaries are determined by the average between the fluid and aggregated state points. Note that in depletion systems, aggregation and gelation are identified with the liquid-gas phase boundary.^{43,44,54} Thus, while these are non-equilibrium states, comparison with equilibrium phase behavior is nevertheless highly informative. For protein volume fractions below those which we tested, a dotted line is drawn based on the intuition from the literature.^{32,92,104} The smaller the concentration of proteins, the higher the polymer concentration needed for phase separation. As noted above, the protein volume fraction is estimated by assuming that the eGFP molecules are spherocylinders of aspect ratio $L/D - 1 = 1.05$. The protein dimensions determined from SAXS (Sec. II D) and estimated polymer size gave a size ratio $q = (2R_{\text{scale}}^{\text{scale}})/D \sim 0.4$.

As a function of polymer concentration, we first encounter protein solutions where the eGFP appears stable and exhibits no observable aggregation, but instead there is a uniform fluorescent intensity, as the protein dimers are far below the resolution of the microscope [Fig. 2(a)]. Upon increasing the polymer concentration, we see aggregation for polymer fugacity $z_{\text{pol}} = 30.0 \pm 1.0$ (which

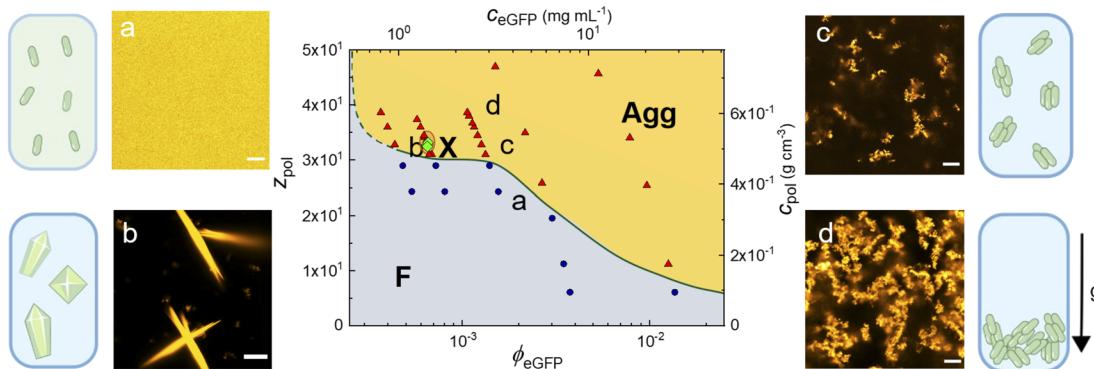


FIG. 2. Phase diagram of the eGFP-PEG system with 10 mM NaCl. Here, the phase behavior is shown in the protein volume fraction–polymer fugacity ($\phi_{\text{eGFP}} - z_{\text{pol}}$) plane, coupled with their concentrations. Fluids, aggregating systems, and crystals are denoted by blue circles, red triangles, and green diamonds, respectively. Confocal microscopy images of different states, along with a schematic representation of the system behavior, are shown in (a)–(d) as follows: (a) fluids, (b) crystals, (c) aggregation, and (d) denser sediments. Here the arrow denotes gravity. The scale bars indicate 10 μm .

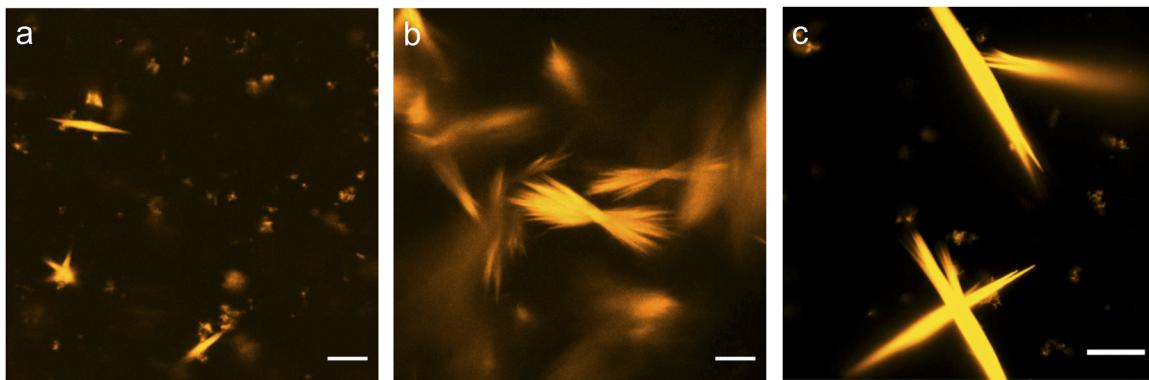


FIG. 3. Crystallization close to the demixing phase boundary: (a) and (c) crystals and aggregates and (b) pure crystals [see Fig. 2(b) for location on the phase diagram]. Upon increasing the polymer volume fraction, crystals were found around $z_{\text{pol}} = 32.5$ and $\phi_{\text{eGFP}} = 0.0011$. Bars = 10 μm .

corresponds to a protein volume fraction around 1.3×10^{-3}) shown in Fig. 2(c). Now, the polymer concentration here is rather high; indeed, the polymer volume fraction $\phi_{\text{pol}} = \rho_{\text{pol}}\pi R_g^3/6$ is of order unity. We return to this point below in Sec. IV.

As the protein concentration is increased, the polymer concentration required for aggregation decreases. Upon further increase in polymer concentration, protein aggregates form quickly and become large enough that considerable quantities sediment to the bottom of the sample where a denser sediment builds up [Fig. 2(d)]. This is reminiscent of aggregation and sedimentation behavior in colloidal systems.¹⁰⁵ In a small region of the phase diagram, we encounter protein crystallization, indicated as green diamonds (see the region denoted as "X") in Fig. 2. We note that there is some lack of smoothness in the phase boundary. Such fluctuations in phase boundaries are well-known in soft matter systems (see, e.g., Ref. 106), and we leave this for further investigation.

Protein crystallization has been related to near-critical behavior.⁴⁹ Here, although the regime of crystallization occurs near the aggregation line (which, by itself, might link it to criticality^{52,107}), the protein volume fraction is vastly lower than any critical isochore that would be expected to occur for this system. Indeed, the volume fraction of the critical isochore for spherical colloids plus polymers with size ratio $q \sim 0.4$ is estimated to be at least $\phi_c \gtrsim 0.25$,⁹³ so it is hard to imagine that critical fluctuations are important here. The lengths of the crystallites that we find are in the range of ~ 4 to 80 μm . Figure 3(b) is pure crystal, while Figs. 3(a) and 3(c) show aggregates that we presume to be amorphous.

2. Salt-free system

To investigate the effect of the (weak) electrostatic interactions between the proteins, we determine the phase behavior in the absence of added salt, as shown in Fig. 4. We find a boundary for aggregation estimated at $z_{\text{pol}} = 30.9 \pm 1.9$ for a protein volume fraction of 1.3×10^{-3} , which is almost indistinguishable to the case with added salt (Fig. 4) at the same protein volume fraction. This is quite consistent with the soft matter inspired analysis of treating the proteins as hard spherocylinders. However, we do not encounter any

crystallization behavior here and return to this in the discussion below.

3. Effects of polymer size

So far, we have discussed the system with the smaller polymer (PEG620), and we now switch to the larger polymer. We chose PEG2000 here because its size is comparable to that of the protein. We therefore expect normal depletion behavior, as described by the Asakura–Oosawa model, unlike the protein limit $q \gg 1$.^{87–89} The phase diagram for the eGFP–PEG2000 system is shown in Fig. 5. The aggregation is found at a much lower fugacity, $z_{\text{pol}} = 1.20 \pm 0.04$, compared with the smaller polymer at the same protein volume fraction around 10^{-2} . This is qualitatively consistent with the literature,^{32,82,104} that the larger the polymer, the lower the fugacity needed for phase separation. Below, we provide a more quantitative comparison. Like the case with no added salt above (Fig. 4), we find no evidence of protein crystallization.

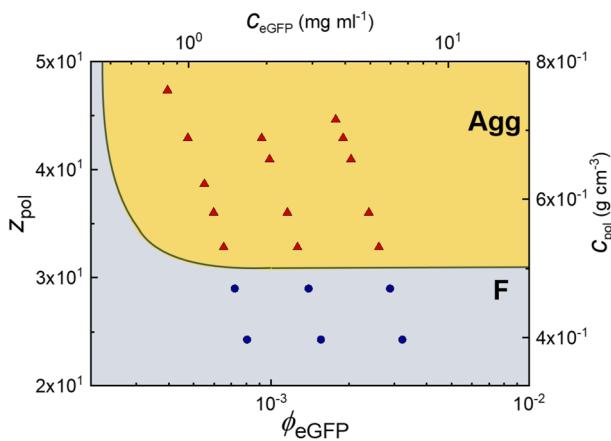


FIG. 4. Phase diagram of the eGFP–PEG620 system with no added salt. Fluids and aggregating systems are denoted by blue circles and red triangles, respectively.

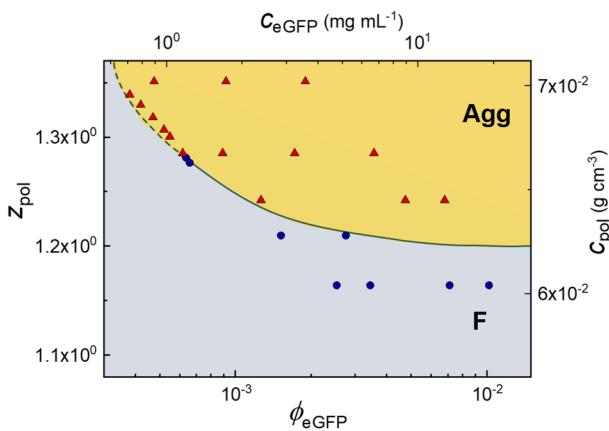


FIG. 5. Phase diagram of eGFP with larger PEG2000 polymer and 10 mM NaCl. Fluids and aggregating systems are denoted by blue circles and red triangles, respectively.

B. Comparison with colloid-polymer mixtures

In order to make a comparison with theoretical and computer simulation predictions, we interpolate between phase boundaries determined for spheres ($L/D - 1 = 0$)⁹³ and spherocylinders of a larger aspect ratio than those we consider here ($L/D - 1 = 5$) (Fig. 6).⁹² It is important to consider the nature of the phase that forms upon demixing. In the case of sphere-polymer mixtures, upon adding polymer at a low colloid volume fraction, the first phase transition that is encountered (for $q \gtrsim 0.3$) is the (colloidal) liquid-gas demixing.^{32,94,104,108} In the case of spherocylinders with aspect ratio $L/D - 1 = 5$, it is fluid-crystal coexistence.^{91,92} Nevertheless, for spheres at $q \approx 0.4\text{--}0.5$, the liquid-gas and fluid-crystal phase

boundaries occur at quite similar values of the polymer fugacity and so here we neglect the difference. We are in any case unaware of any computation of the phase diagram for our parameters, and note that the free volume theory of Lekkerkerker *et al.*¹⁰⁴ is not highly accurate for these parameters.⁹³

We fit the phase boundaries obtained for spheres⁹³ and spherocylinders⁹² by a power law $Z_{pol} = a - b\phi_{eGFP}^c$ at a low value of protein volume fraction $\phi_{eGFP} \sim 10^{-2}$ as a function of q . Here, a , b , and c are fitted constants. The interpolation is done linearly by $z_{pol}^{(0)} - 1.05(z_{pol}^{(0)} - z_{pol}^{(5)})/5$, where $z_{pol}^{(0)}$ is for spheres⁹³ and $z_{pol}^{(5)}$ is for spherocylinders ($L/D - 1 = 5$).⁹² Our interpolation is shown in Fig. 6 where we plot the fitted phase boundaries for fitting data and our interpolation. We interpolate to obtain values of q that are consistent with our measured fugacity for demixing $z_{pol} = 8.63$ (smaller polymer) and $z_{pol} = 1.20$ (larger polymer) at a protein volume fraction around 10^{-2} . We have in any case some uncertainty in determining the size ratio q . As noted above, our estimate for the polymer radius of gyration R_g^{scale} relied on polymer scaling, which may not be accurate for such small polymers. Moreover, there are a variety of other assumptions, such as polymer ideality, and rigidity, which have been addressed in more refined theoretical treatments.^{94,109} We therefore accept some adjustment in our fitted values and take $q^{\text{fit}} = 0.59$ for smaller polymers and $q^{\text{fit}} = 0.90$ for larger polymers, which agree well with our data.

For larger polymers, the fitted polymer radius of gyration R_g^{fit} of 1.80 nm falls close to that obtained from the empirical equation of $R_g^{\text{scale}} = 1.64$ nm (see Sec. II B). In the case of the smaller polymer (PEG620), there is an increase in the fitted size ratio ($q^{\text{fit}} = 0.59$) relative to that obtained from polymer scaling ($q^{\text{scale}} = 0.42$) (see Sec. II B). Now, we consider the assumption that polymers are ideal as in the standard AO model as with $R_g = 1.2$ nm, these are very small polymers to treat as ideal.⁸¹ Dijkstra *et al.*¹⁰⁸ compared additive hard-spheres with ideal polymers using thermodynamic perturbation theory, and they found that for small q and polymer packing fraction ϕ_{pol} , the phase separation is very similar between two models. Here, we have $q = 0.59$, and under these conditions of larger depletants, the behavior of spherical colloids plus ideal polymer and spherical colloids plus hard sphere depletant is rather different, at least at the level of the effective interactions between the larger spheres.¹¹⁰ While we cannot rule out that the polymers may exhibit significant deviations from ideality, given that the phase behavior we find is similar to that of hard spherocylinders and ideal polymer, we note that at the level of our analysis, the polymers appear more likely to be behaving in a manner more akin to ideal polymers than hard spheres.

IV. DISCUSSION

We have seen that the model fluorescent protein-polymer system can, rather surprisingly, be treated in the spirit of a mixture of colloids and non-absorbing polymer where the only additional complexity is an approximate treatment of the anisotropy of the protein dimers. Furthermore, we observe no aggregation for eGFP in the absence of polymer at least to a concentration of 500 mg/ml, corresponding to a volume fraction of $\phi_{eGFP} = 0.48$. At this volume fraction, the protein solution becomes very viscous, consistent

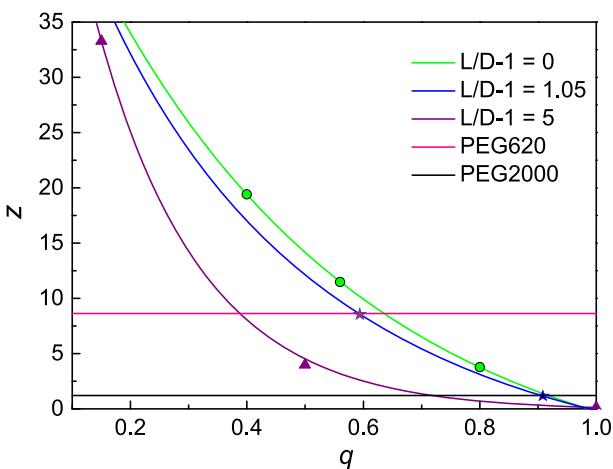


FIG. 6. The interpolation from the data presented by Savenko and Dijkstra⁹² for $L/D - 1 = 5$ (purple triangles) and by Lo Verso *et al.*⁹³ for $L/D - 1 = 0$ (green circles); green and purple lines are fitted by power laws (see text). The blue line is the interpolation, and the pink and black stars are crossovers that denote matching for PEG620 and PEG2000, respectively.

with previous work which found glassy behavior reminiscent of colloidal systems in concentrated eye lens α -crystallin.⁴⁶ Moreover, we found that upon dilution, aggregated protein solutions redissolved, behavior which is compatible with weak, depletion-driven aggregation.

The crystallization behavior in our system re-emphasizes that protein crystals can be produced through the addition of polymer, as noted previously.¹⁰² This is significant because the process is apparently immediate without a fine-tuning of the system. We focus on the low volume fraction regime in this work, and we note that crystals only appeared in a limited region in our phase diagram and then only in the system with the smaller polymer and added salt, not in the case of the larger polymer or without added salt. At first sight, it may seem surprising that we find above (Sec. II A) that the electrostatic interactions are very weak in our system, with or without salt. It is important to highlight that the isoelectric point (pI) of the monomeric unit of eGFP (obtained from its amino acid sequence) is 5.8,¹¹¹ which is close to the pH 7 used in the experiments. This might explain the small values found for the surface charge.

We now enquire as to why not adding salt suppresses the crystallization. The observation of crystallization only in a very limited region of polymer concentration (i.e., attraction strength) is consistent with previous work with (spherical) colloids and polymer mixtures^{32,52,112} and has been interpreted in terms of fluctuation-dissipation theorem violation.¹¹³ Additionally, it has been observed that acidic proteins are more likely to crystallize when the pH of the solution is 0–2.5 units above their pI.¹¹⁴ Our experiments fall within such a range. Thus, only a small amount of salt would be required to overcome small electrostatic repulsions under these favorable conditions. What is perhaps more notable is the limited range of protein concentration in which we see crystallization and the failure of the salt-free system to crystallize. It is quite possible that the region of the phase diagram in which crystallization occurs is so small and is somehow related to more complex behavior than that which we treat here. For example, Fusco *et al.* showed the importance of contacts in the crystallization behavior of rubredoxin.⁷⁴ We speculate that a decrease in the electrostatic repulsions only needs to occur around or in these regions to promote crystal formation, leading to only small amounts of salt required to yield a crystal, in contrast, for example, with isotropic systems. Finally, salts can also affect the hydrophobic protein–protein interactions by increasing the surface tension.⁸⁵ These interactions have shown to be relevant in the formation of a crystal phase and protein solubility,^{115,116} which cannot be discarded in the present study.

Nevertheless, the crystallization that we observe is compatible with the spherocylinder–polymer phase behavior ($L/D - 1 = 5$).^{91,92} It would be most interesting to determine the phase diagram for hard spherocylinders of aspect ratio $L/D - 1 = 1.05$ plus polymer, but for now, we conclude that our finding of protein crystallization is consistent with previous studies of mixtures of hard particles and ideal polymers.^{32,91–94,104}

The polymer volume fractions at which we find aggregation are rather high, of order unity. It is important to enquire whether one can still apply the concept of polymer-induced depletion under these conditions. Accurate computer simulations in which the polymer chain segments are included predict that for the polymer

fractions that we consider here, only small deviations of ideal Asakura–Oosawa behavior are expected.¹¹⁷ While we have treated our eGFP as spherocylinders and this work refers to spherical particles, we are unaware of similar work that pertains to anisotropic particles and thus, in the absence of evidence to the contrary, presume that a simple depletion picture remains reasonably accurate at these polymer concentrations.

While we have suggested that it is possible to account for the behavior of our system by treating the eGFP as hard spherocylinders in a solution of ideal polymers, we can be confident that the situation in reality is much more complex. In addition to an enhancement of hydrophobic interactions from salt addition discussed above, due to the amphiphilic nature of PEG, additional hydrophobic¹¹⁸ and chemical⁸⁴ interactions (via PEG–CH₂OCH₂–groups) between PEG and proteins might also contribute to this phenomenon. Furthermore, PEG molecules can also enhance aggregation and crystallization via effective repulsion since PEG might preferentially form hydrogen bonds with water compared to the proteins.⁸⁵ Finally, we have determined electrostatic interactions between eGFP dimers to be weak if we only consider the *net* charge. Of course, this is a very significant approximation. Monomeric eGFP has a number of charging groups, e.g., 32 acidic residues and so a more sophisticated approach that takes this into account may prove valuable. Such an approach as that noted above for rubredoxin⁷⁴ would be most interesting to pursue here.

In short, further work is needed to explore throughout the metastable region, and then predictions can be validated using the depletion theory. Moreover, the properties of those crystals formed at this low protein concentration and by purely depletion interactions are certainly worth investigating in future research.

V. CONCLUSION

We studied the phase behavior of a model system of fluorescent proteins and polymers (eGFP–PEG) in the “colloid limit” where the polymer depletant is smaller than or comparable in size to the protein. A phase behavior of fluid-aggregation was observed for two polymer sizes, i.e., two polymer–protein size ratios, in addition to a small region of the phase diagram of a system with added salt (NaCl) and small polymers where protein crystallization occurred. At high polymer concentration, protein aggregates were large enough to sediment on the timescale of the experiment and form a sediment whose structure is reminiscent of a gel. In the absence of polymer, solutions of eGFP are stable at least to a concentration of 500 mg/ml (volume fraction at 0.48). This suggests that the eGFP dimers interact rather weakly and that approximating them as hard particles may be reasonable.

Based on the shape of eGFP dimers as deduced from small angle x-ray scattering, we treat them as hard spherocylinders with aspect ratio $L/D - 1 = 1.05$. In the case of the small polymer (PEG 620), the aggregation boundary of polymer fugacity at a protein volume fraction of 1.3×10^{-3} was found to be almost indistinguishable between 30.0 ± 1.0 for the salt-screened system and 30.9 ± 1.9 for the salt-free system. For the larger polymer (PEG2000), aggregation was found at a polymer fugacity of 1.2. Under the assumptions of DLVO theory, the effects of electrostatic interactions between the proteins were found to be weak. Intriguingly, in the case of no added salt, we observed no protein crystallization.

To compare with predictions for hard spherocylinder-ideal polymer mixtures, we interpolated the fugacity for the aggregation phase boundary from existing literature, between $L/D - 1 = 0$ for sphere-polymer mixtures⁹³ and $L/D - 1 = 5$ for spherocylinder-polymer mixtures,⁹² and fitted a polymer radius of gyration of 1.1 nm for PEG620, a fitted polymer-protein size ratio $q^{\text{fit}} = 0.59$. Compared with the empirical estimation of 0.83 nm, this somewhat larger size may be related to some non-ideality in the polymers⁹⁴ (we note that polymer scaling theory is expected to break down for such small polymers in any case). For the larger PEG2000 polymers, we fit a size ratio $q^{\text{fit}} = 0.90$.

The behavior we observed is consistent with the depletion picture of hard spherocylinders and ideal polymers. However, in reality, our system is rather more complex. At our level of analysis and observation, we cannot exclude the possibility that other interactions drive the phenomena that we observe, for example, hydration effects and hydrophobic or electrostatic “patches.” Nevertheless, the fact that in the absence of polymer, the eGFP solution exhibits no aggregation to such high concentrations, at that the aggregates redissolve upon dilution, gives us some optimism that the behavior we observe may be driven by such simple interactions as the excluded volume effects of polymer-induced depletion.

ACKNOWLEDGMENTS

We would like to thank John Russo and Mike Allen for helpful discussions, Richard Stenner for protein expression and purification, and Angélique Coutable-Pennarun for assistance with zeta potential measurements. This work was financially supported by the Bristol Centre for Functional Nanomaterials (BCFN), the Chinese Scholarship Council (CSC), and Bayer AG. I.R.d.A. was funded by the Philip Leverhulme Prize 2018 awarded by the Leverhulme Trust. I.R.d.A., J.L.R.A. and C.P.R. were funded by the Leverhulme Trust grant “Unifying Protein Design and Assembly of Soft Matter for New Materials.” R.C., I.R.d.A., and C.P.R. gratefully acknowledge ERC Grant Agreement No. 617266 NANOPRS for financial support and the Engineering and Physical Sciences Research Council (Grant No. EP/H022333/1). The Ganesha x-ray scattering apparatus used for this research was purchased under EPSRC Grant Atoms to Applications (Grant No. EP/K035746/1). This work benefitted from the SasView software, originally developed by the DANSE project under NSF Award No. DMR-0520547.

APPENDIX: GEOMETRIC PARAMETERS OF eGFP DETERMINED WITH SAXS

Table II shows the parameters obtained from SASView⁹⁸ fitting using a cylinder form factor.

TABLE II. Parameters obtained from SASView⁹⁸ fitting using a cylinder form factor.

| Protein | Radius (Å) | Error radius (Å) | Length (Å) | Error length (Å) | Fitting χ^2 |
|---------|---------------|---------------------|---------------|---------------------|------------------|
| eGFP | 20.5 | 0.08 | 82.3 | 0.7 | 1.19 |

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH image to ImageJ: 25 years of image analysis,” *Nat. Methods* **9**, 671–675 (2012).
- A. M. Morris, M. A. Watzky, and R. G. Finke, “Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature,” *Biochim. Biophys. Acta, Proteins Proteomics* **1794**, 375–397 (2009).
- S. I. A. Cohen, S. Linse, L. M. Luheshi, E. Hellstrand, D. A. White, L. Rajah, D. E. Otzen, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles, “Proliferation of amyloid- β 42 aggregates occurs through a secondary nucleation mechanism,” *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9758–9763 (2013).
- Y. Bai, Q. Luo, and J. Liu, “Protein self-assembly via supramolecular strategies,” *Chem. Soc. Rev.* **45**, 2756–2767 (2016).
- S. Alberti, A. Gladfelter, and T. Mittag, “Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates,” *Cell* **176**, 419–434 (2019).
- Y.-T. Lai, N. P. King, and T. O. Yeates, “Principles for designing ordered protein assemblies,” *Trends Cell Biol.* **22**, 653–661 (2012).
- A. C. Dumetz, R. A. Lewus, A. M. Lenhoff, and E. W. Kaler, “Effects of ammonium sulfate and sodium chloride concentration on PEG/protein liquid–liquid phase separation,” *Langmuir* **24**, 10345–10351 (2008).
- D. Fusco and P. Charbonneau, “Soft matter perspective on protein crystal assembly,” *Colloids Surf., B* **137**, 22–31 (2016).
- R. Hui and A. Edwards, “High-throughput protein crystallization,” *J. Struct. Biol.* **142**, 154–161 (2003).
- Q.-q. Huang, M.-k. Teng, and L.-w. Niu, “Protein crystallization with a combination of hard and soft precipitants,” *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1444–1448 (1999).
- V. Bhamidi, S. Varanasi, and C. A. Schall, “Protein crystal nucleation: Is the pair interaction potential the primary determinant of kinetics?” *Langmuir* **21**, 9044–9050 (2005).
- J. Glaser and S. C. Glotzer, “Looped liquid-liquid coexistence in protein crystallization,” *arXiv:1910.06865* (2019).
- T. Nicolai and D. Durand, “Controlled food protein aggregation for new functionality,” *Curr. Opin. Colloid Interface Sci.* **18**, 249–256 (2013).
- T. Yamazaki, Y. Kimura, P. G. Vekilov, E. Furukawa, M. Shirai, H. Matsumoto, A. E. S. Van Driessche, and K. Tsukamoto, “Two types of amorphous protein particles facilitate crystal nucleation,” *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2154–2159 (2017).
- A. McPherson, “Crystallization of proteins from polyethylene glycol,” *J. Biol. Chem.* **251**, 6300–6303 (1976).
- S. Tanaka and M. Ataka, “Protein crystallization induced by polyethylene glycol: A model study using apoferritin,” *J. Chem. Phys.* **117**, 3504–3510 (2002).
- A. Tardieu, F. Bonneté, S. Finet, and D. Vivarès, “Understanding salt or PEG induced attractive interactions to crystallize biological macromolecules,” *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 1549–1553 (2002).
- J. Bloustein, T. Virmani, G. M. Thurston, and S. Fraden, “Light scattering and phase behavior of lysozyme-poly(ethylene glycol) mixtures,” *Phys. Rev. Lett.* **96**, 087803 (2006).
- M. Yamanaka, K. Inaka, N. Furubayashi, M. Matsushima, S. Takahashi, H. Tanaka, S. Sano, M. Sato, T. Kobayashi, and T. Tanaka, “Optimization of salt concentration in PEG-based crystallization solutions,” *J. Synchrotron Radiat.* **18**, 84–87 (2011).
- O. D. Velev, E. W. Kaler, and A. M. Lenhoff, “Protein interactions in solution characterized by light and neutron scattering: Comparison of lysozyme and chymotrypsinogen,” *Biophys. J.* **75**, 2682–2697 (1998).
- J. D. Schmit and K. Dill, “Growth rates of protein crystals,” *J. Am. Chem. Soc.* **134**, 3934–3937 (2012).
- J.-P. Astier and S. Veesler, “Using temperature to crystallize proteins: A mini-review,” *Cryst. Growth Des.* **8**, 4215–4219 (2008).

- ²³S. Grouazel, F. Bonneté, J.-P. Astier, N. Ferté, J. Perez, and S. Veesler, "Exploring bovine pancreatic trypsin inhibitor phase transitions," *J. Phys. Chem. B* **110**, 19664–19670 (2006).
- ²⁴M. Cieślik and Z. S. Derewenda, "The role of entropy and polarity in intermolecular contacts in protein crystals," *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **65**, 500–509 (2009).
- ²⁵Y. E. Kim, Y. N. Kim, J. A. Kim, H. M. Kim, and Y. Jung, "Green fluorescent protein nanopolygons as monodisperse supramolecular assemblies of functional proteins with defined valency," *Nat. Commun.* **6**, 7134 (2015).
- ²⁶D. J. Mandell and T. Kortemme, "Computer-aided design of functional protein interactions," *Nat. Chem. Biol.* **5**, 797–807 (2009).
- ²⁷Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker, and F. A. Tezcan, "Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals," *Nature* **533**, 369–373 (2016).
- ²⁸S. Zhang, R. G. Alberstein, J. J. De Yoreo, and F. A. Tezcan, "Assembly of a patchy protein into variable 2D lattices via tunable multiscale interactions," *Nat. Commun.* **11**, 3770 (2020).
- ²⁹J. J. McManus, P. Charbonneau, E. Zaccarelli, and N. Asherie, "The physics of protein self-assembly," *Curr. Opin. Colloid Interface Sci.* **22**, 73–79 (2016).
- ³⁰A. Stradner and P. Schurtenberger, "Potential and limits of a colloid approach to protein solutions," *Soft Matter* **16**, 307–323 (2020).
- ³¹I. W. Hamley and V. Castelletto, "Biological soft materials," *Angew. Chem., Int. Ed.* **46**, 4442–4455 (2007).
- ³²W. C. K. Poon, "The physics of a model colloid-polymer mixture," *J. Phys.: Condens. Matter* **14**, R859–R880 (2002).
- ³³N. Dorsaz, L. Filion, F. Smalleenburg, and D. Frenkel, "Spiers memorial lecture: Effect of interaction specificity on the phase behaviour of patchy particles," *Faraday Discuss.* **159**, 9–21 (2012).
- ³⁴F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. A. Skoda, R. M. J. Jacobs, M. Stzucki, and F. Schreiber, "Charge-controlled metastable liquid-liquid phase separation in protein solutions as a universal pathway towards crystallization," *Soft Matter* **8**, 1313–1316 (2012).
- ³⁵F. Zhang, "Nonclassical nucleation pathways in protein crystallization," *J. Phys.: Condens. Matter* **29**, 443002 (2017).
- ³⁶R. Piazza, "Protein interactions and association: An open challenge for colloid science," *Curr. Opin. Colloid Interface Sci.* **8**, 515–522 (2004).
- ³⁷Z. Dogic, K. R. Purdy, E. Grelet, M. Adams, and S. Fraden, "Isotropic-nematic phase transition in suspensions of filamentous virus and the neutral polymer dextran," *Phys. Rev. E* **69**, 051702 (2006).
- ³⁸P. M. Tessier, H. R. Johnson, R. Pazhianur, B. W. Berger, J. L. Prentice, B. J. Bahnsen, S. I. Sandler, and A. M. Lenhoff, "Predictive crystallization of ribonuclease A via rapid screening of osmotic second virial coefficients," *Proteins: Struct., Funct., Bioinf.* **50**, 303–311 (2002).
- ³⁹A. George and W. W. Wilson, "Predicting protein crystallization from a dilute solution property," *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 361–365 (1994).
- ⁴⁰A. H. Elcock and J. A. McCammon, "Calculation of weak protein-protein interactions: The pH dependence of the second virial coefficient," *Biophys. J.* **80**, 613–625 (2001).
- ⁴¹N. Wentzel and J. D. Gunton, "Effect of solvent on the phase diagram of a simple anisotropic model of globular proteins," *J. Phys. Chem. B* **112**, 7803–7809 (2008).
- ⁴²M. G. Noro and D. Frenkel, "Extended corresponding-states behavior for particles with variable range attractions," *J. Chem. Phys.* **113**, 2941–2944 (2000).
- ⁴³P. J. Lu, E. Zaccarelli, F. Ciulla, A. B. Schofield, F. Sciortino, and D. A. Weitz, "Gelation of particles with short-range attraction," *Nature* **453**, 499–503 (2008).
- ⁴⁴C. P. Royall, "Hunting mermaids in real space: Known knowns, known unknowns and unknown unknowns," *Soft Matter* **14**, 4020–4028 (2018).
- ⁴⁵F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, S. Zorn, R. A. Martin, C. M. Martin, G. F. Clark, S. Weggerl, A. Hildebrandt, O. Kohlbacher *et al.*, "Reentrant condensation of proteins in solution induced by multivalent counterions," *Phys. Rev. Lett.* **101**, 148101 (2008).
- ⁴⁶G. Foffi, G. Savin, S. Bucciarelli, N. Dorsaz, G. M. Thurston, A. Stradner, and P. Schurtenberger, "Hard sphere-like glass transition in eye lens α -crystallin solutions," *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16748–16753 (2014).
- ⁴⁷S. J. Kim, C. Dumont, and M. Gruebele, "Simulation-based fitting of protein-protein interaction potentials to SAXS experiments," *Biophys. J.* **94**, 4924–4931 (2008).
- ⁴⁸A. Huang, H. Yao, and B. D. Olsen, "SANS partial structure factor analysis for determining protein-polymer interactions in semidilute solution," *Soft Matter* **15**, 7350–7359 (2019).
- ⁴⁹P. R. ten Wolde and D. Frenkel, "Enhancement of protein crystal nucleation by critical density fluctuations," *Science* **277**, 1975–1978 (1997).
- ⁵⁰G. Pellicane, D. Costa, and C. Caccamo, "Theory and simulation of short-range models of globular protein solutions," *J. Phys.: Condens. Matter* **16**, S4923–S4936 (2004).
- ⁵¹J. R. Savage and A. D. Dinsmore, "Experimental evidence for two-step nucleation in colloidal crystallization," *Phys. Rev. Lett.* **102**, 198302 (2009).
- ⁵²S. L. Taylor, R. Evans, and C. Patrick Royall, "Temperature as an external field for colloid-polymer mixtures: 'Quenching' by heating and 'melting' by cooling," *J. Phys.: Condens. Matter* **24**, 464128 (2012).
- ⁵³E. Zaccarelli, "Colloidal gels: Equilibrium and non-equilibrium routes," *J. Phys.: Condens. Matter* **19**, 323101 (2007).
- ⁵⁴C. P. Royall, M. A. Faers, S. Fussell, and J. Hallett, "Real space analysis of colloidal gels: Triumphs, challenges and future directions," *J. Phys.: Condens. Matter* (submitted) (2021).
- ⁵⁵N. Asherie, A. Lomakin, and G. B. Benedek, "Phase diagram of colloidal solutions," *Phys. Rev. Lett.* **77**, 4832 (1996).
- ⁵⁶F. Cardinaux, T. Gibaud, A. Stradner, and P. Schurtenberger, "Interplay between spinodal decomposition and glass formation in proteins exhibiting short-range attractions," *Phys. Rev. Lett.* **99**, 118301 (2007).
- ⁵⁷A. M. Kulkarni, N. M. Dixit, and C. F. Zukoski, "Ergodic and non-ergodic phase transitions in globular protein suspensions," *Faraday Discuss.* **123**, 37–50 (2003).
- ⁵⁸M. L. Broide, C. R. Berland, J. Pande, O. O. Ogun, and G. B. Benedek, "Binary-liquid phase separation of lens protein solutions," *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5660–5664 (1991).
- ⁵⁹S. Whitelam, "Control of pathways and yields of protein crystallization through the interplay of nonspecific and specific attractions," *Phys. Rev. Lett.* **105**, 088102 (2010).
- ⁶⁰J. Groenewold and W. K. Kegel, "Anomalously large equilibrium clusters of colloids," *J. Phys. Chem. B* **105**, 11702–11709 (2001).
- ⁶¹A. Stradner, H. Sedgwick, F. Cardinaux, W. C. K. Poon, S. U. Egelhaaf, and P. Schurtenberger, "Equilibrium cluster formation in concentrated protein solutions and colloids," *Nature* **432**, 492–495 (2004).
- ⁶²H. Sedgwick, S. U. Egelhaaf, and W. C. K. Poon, "Clusters and gels in systems of sticky particles," *J. Phys.: Condens. Matter* **16**, S4913–S4922 (2004).
- ⁶³A. I. Campbell, V. J. Anderson, J. S. van Duijneveldt, and P. Bartlett, "Dynamical arrest in attractive colloids: The effect of long-range repulsion," *Phys. Rev. Lett.* **94**, 208301 (2005).
- ⁶⁴F. Sciortino, P. Tartaglia, and E. Zaccarelli, "One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion," *J. Phys. Chem. B* **109**, 21942–21953 (2005).
- ⁶⁵A. Malins, S. R. Williams, J. Eggers, H. Tanaka, and C. P. Royall, "The effect of inter-cluster interactions on the structure of colloidal clusters," *J. Non-Cryst. Solids* **357**, 760–766 (2011).
- ⁶⁶C. L. Klix, C. P. Royall, and H. Tanaka, "Structural and dynamical features of multiple metastable glassy states in a colloidal system with competing interactions," *Phys. Rev. Lett.* **104**, 165702 (2010).
- ⁶⁷K. van Gruijthuijsen, M. Obiols-Rabasa, P. Schurtenberger, W. G. Bouwman, and A. Stradner, "The extended law of corresponding states when attractions meet repulsions," *Soft Matter* **14**, 3704–3715 (2018).
- ⁶⁸C. N. Likos, "Effective interactions in soft condensed matter physics," *Phys. Rep.* **348**, 267–439 (2001).
- ⁶⁹A. Shukla, E. Mylonas, E. Di Cola, S. Finet, P. Timmins, T. Narayanan, and D. I. Svergun, "Absence of equilibrium cluster phase in concentrated lysozyme solutions," *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5075–5080 (2008).
- ⁷⁰C. L. Klix, K. Murata, H. Tanaka, S. R. Williams, A. Malins, and C. P. Royall, "Novel kinetic trapping in charged colloidal clusters due to self-induced surface charge organization," *Sci. Rep.* **3**, 2072 (2013).

- ⁷¹M. Adams, Z. Dogic, S. L. Keller, and S. Fraden, "Entropically driven microphase transitions in mixtures of colloidal rods and spheres," *Nature* **393**, 349–352 (1998).
- ⁷²M. Baaden and S. J. Marrink, "Coarse-grain modelling of protein–protein interactions," *Curr. Opin. Struct. Biol.* **23**, 878–886 (2013).
- ⁷³H. Liu, S. K. Kumar, and J. F. Douglas, "Self-assembly-induced protein crystallization," *Phys. Rev. Lett.* **103**, 018101 (2009).
- ⁷⁴D. Fusco, J. J. Headd, A. De Simone, J. Wang, and P. Charbonneau, "Characterizing protein crystal contacts and their role in crystallization: Rubredoxin as a case study," *Soft Matter* **10**, 290–302 (2014).
- ⁷⁵J. P. K. Doye and W. C. K. Poon, "Protein crystallization in vivo," *Curr. Opin. Colloid Interface Sci.* **11**, 40–46 (2006).
- ⁷⁶I. Altan, D. Fusco, P. V. Afonine, and P. Charbonneau, "Learning about biomolecular solvation from water in protein crystals," *J. Phys. Chem. B* **122**, 2475–2486 (2018).
- ⁷⁷S. James, M. K. Quinn, and J. J. McManus, "The self assembly of proteins; probing patchy protein interactions," *Phys. Chem. Chem. Phys.* **17**, 5413–5420 (2015).
- ⁷⁸Y. Li, T. Shi, L. An, and Q. Huang, "Monte Carlo simulation on complex formation of proteins and polysaccharides," *J. Phys. Chem. B* **116**, 3045–3053 (2012).
- ⁷⁹N. Gnan, F. Sciortino, and E. Zaccarelli, "Patchy particle models to understand protein phase behavior," in *Protein Self-Assembly* (Springer, 2019), pp. 187–208.
- ⁸⁰J. Cai and A. M. Sweeney, "The proof is in the *pidar*: Generalizing proteins as patchy particles," *ACS Cent. Sci.* **4**, 840–853 (2018).
- ⁸¹S. Asakura and F. Oosawa, "On interaction between two bodies immersed in a solution of macromolecules," *J. Chem. Phys.* **22**, 1255–1256 (1954).
- ⁸²D. H. Atha and K. C. Ingham, "Mechanism of precipitation of proteins by polyethylene glycols. Analysis in terms of excluded volume," *J. Biol. Chem.* **256**, 12108–12117 (1981).
- ⁸³J. Hašek, "Poly(ethylene glycol) interactions with proteins," *Z. Kristallogr. Suppl.* **2006**, 613–618.
- ⁸⁴I. A. Shkel, D. B. Knowles, and M. T. Record, "Separating chemical and excluded volume interactions of polyethylene glycols with native proteins: Comparison with PEG effects on DNA helix formation: Separating chemical and excluded volume effects of polyethylene glycol," *Biopolymers* **103**, 517–527 (2015).
- ⁸⁵S. D. Durbin and G. Feher, "Protein crystallization," *Annu. Rev. Phys. Chem.* **47**, 171–204 (1996).
- ⁸⁶D. Vivarès, L. Belloni, A. Tardieu, and F. Bonneté, "Catching the PEG-induced attractive interaction between proteins," *Eur. Phys. J. E* **9**, 15–25 (2002).
- ⁸⁷P. G. Bolhuis, E. J. Meijer, and A. A. Louis, "Colloid-polymer mixtures in the protein limit," *Phys. Rev. Lett.* **90**, 068304 (2003).
- ⁸⁸K. J. Mutch, J. S. van Duijneveldt, and J. Eastoe, "Colloid-polymer mixtures in the protein limit," *Soft Matter* **3**, 155–167 (2007).
- ⁸⁹K. J. Mutch, J. S. van Duijneveldt, J. Eastoe, I. Grillo, and R. K. Heenan, "Small-angle neutron scattering study of microemulsion–polymer mixtures in the protein limit," *Langmuir* **24**, 3053–3060 (2008).
- ⁹⁰C. Nakagawa, S. Nishimura, K. Senda-Murata, and K. Sugimoto, "A rapid and simple method of evaluating the dimeric tendency of fluorescent proteins in living cells using a truncated protein of importin α as fusion tag," *Biosci., Biotechnol., Biochem.* **76**, 388–390 (2012).
- ⁹¹P. G. Bolhuis, A. Stroobants, D. Frenkel, and H. N. W. Lekkerkerker, "Numerical study of the phase behavior of rodlike colloids with attractive interactions," *J. Chem. Phys.* **107**, 1551–1564 (1997).
- ⁹²S. V. Savenko and M. Dijkstra, "Phase behavior of a suspension of colloidal hard rods and nonadsorbing polymer," *J. Chem. Phys.* **124**, 234902 (2006).
- ⁹³F. Lo Verso, R. L. Vink, D. Pini, and L. Reatto, "Critical behavior in colloid-polymer mixtures: Theory and simulation," *Phys. Rev. E* **73**, 061407 (2006).
- ⁹⁴H. N. Lekkerkerker and R. Tuinier, *Colloids and the Depletion Interaction*, Lecture Notes in Physics Vol. 833 (Springer Netherlands, Dordrecht, 2011).
- ⁹⁵F. Roosen-Runge, B. S. Heck, F. Zhang, O. Kohlbacher, and F. Schreiber, "Interplay of pH and binding of multivalent metal ions: Charge inversion and reentrant condensation in protein solutions," *J. Phys. Chem. B* **117**, 5777–5787 (2013).
- ⁹⁶T.-Y. D. Tang, D. Cecchi, G. Fracasso, D. Accardi, A. Coutable-Pennarun, S. S. Mansy, A. W. Perriman, J. L. R. Anderson, and S. Mann, "Gene-mediated chemical communication in synthetic protocell communities," *ACS Synth. Biol.* **7**, 339–346 (2018).
- ⁹⁷M. Kaishima, J. Ishii, T. Matsuno, N. Fukuda, and A. Kondo, "Expression of varied GFPs in *Saccharomyces cerevisiae*: Codon optimization yields stronger than expected expression and fluorescence intensity," *Sci. Rep.* **6**, 35932 (2016).
- ⁹⁸M. Doucet, J. H. Cho, G. Alina, J. Bakker, W. Bouwman, P. Butler, K. Campbell, M. Gonzales, R. Heenan, A. Jackson *et al.*, SasView, <https://www.sasview.org/>, 2017.
- ⁹⁹F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, R. A. Martin, C. M. Martin, and F. Schreiber, "Protein interactions studied by SAXS: Effect of ionic strength and protein concentration for BSA in aqueous solutions," *J. Phys. Chem. B* **111**, 251 (2007).
- ¹⁰⁰M. Wolf, F. Roosen-Runge, F. Zhang, R. Roth, M. W. A. Skoda, R. M. J. Jacobs, M. Sztucki, and F. Schreiber, "Effective interactions in protein–salt solutions approaching liquid–liquid phase separation," *J. Mol. Liq.* **200**, 20–27 (2014).
- ¹⁰¹P. Singh, A. Roche, C. F. van der Walle, S. Uddin, J. Du, J. Warwick, A. Pluen, and R. Curtis, "Determination of protein–protein interactions in a mixture of two monoclonal antibodies," *Mol. Pharmaceutics* **16**, 4775–4786 (2019).
- ¹⁰²J. A. J. Arpino, P. J. Rizkallah, and D. D. Jones, "Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222," *PLoS One* **7**, e47132 (2012).
- ¹⁰³D. P. Myatt, L. Hatter, S. E. Rogers, A. E. Terry, and L. A. Clifton, "Monomeric green fluorescent protein as a protein standard for small angle scattering," *Biomed. Spectrosc. Imaging* **6**, 123–134 (2017).
- ¹⁰⁴H. N. W. Lekkerkerker, W. C. K. Poon, P. N. Pusey, A. Stroobants, and P. B. Warren, "Phase behaviour of colloid + polymer mixtures," *Europhys. Lett.* **20**, 559–564 (1992).
- ¹⁰⁵R. Piazza, "Settled and unsettled issues in particle settling," *Rep. Prog. Phys.* **77**, 056602 (2014).
- ¹⁰⁶W. C. K. Poon, E. R. Weeks, and C. P. Royall, "On measuring colloidal volume fractions," *Soft Matter* **8**, 21–30 (2012).
- ¹⁰⁷P.-R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, "Simulation of homogeneous crystal nucleation close to coexistence," *Faraday Discuss.* **104**, 93 (1996).
- ¹⁰⁸M. Dijkstra, J. M. Brader, and R. Evans, "Phase behaviour and structure of model colloid-polymer mixtures," *J. Phys.: Condens. Matter* **11**, 10079–10106 (1999).
- ¹⁰⁹G. J. Fleer and R. Tuinier, "Analytical phase diagrams for colloids and non-adsorbing polymer," *Adv. Colloid Interface Sci.* **143**, 1–47 (2008).
- ¹¹⁰C. P. Royall, A. A. Louis, and H. Tanaka, "Measuring colloidal interactions with confocal microscopy," *J. Chem. Phys.* **127**, 044507 (2007).
- ¹¹¹PepCalc.com: Innovagen peptide property calculator, <http://pepcalc.com/>; accessed September 23, 2015.
- ¹¹²C. P. Royall and A. Malins, "The role of quench rate in colloidal gels," *Faraday Discuss.* **158**, 301–311 (2012).
- ¹¹³D. Klotsa and R. L. Jack, "Predicting the self-assembly of a model colloidal crystal," *Soft Matter* **7**, 6294–6303 (2011).
- ¹¹⁴J. Kirkwood, D. Hargreaves, S. O'Keefe, and J. Wilson, "Using isoelectric point to determine the pH for initial protein crystallization trials," *Bioinformatics* **31**, 1444–1451 (2015).
- ¹¹⁵M. Prevost, S. J. Wodak, B. Tidor, and M. Karplus, "Contribution of the hydrophobic effect to protein stability: Analysis based on simulations of the Ile-96–Ala mutation in barnase," *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10880–10884 (1991).
- ¹¹⁶M. K. Quinn, S. James, and J. J. McManus, "Chemical modification alters protein–protein interactions and can lead to lower protein solubility," *J. Phys. Chem. B* **123**, 4373–4379 (2019).
- ¹¹⁷A. A. Louis, P. G. Bolhuis, E. J. Meijer, and J. P. Hansen, "Polymer induced depletion potentials in polymer-colloid mixtures," *J. Chem. Phys.* **117**, 1893–1907 (2002).
- ¹¹⁸R. A. Curtis and L. Lue, "A molecular approach to bioseparations: Protein–protein and protein–salt interactions," *Chem. Eng. Sci.* **61**, 907–923 (2006).