Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

Biophysical Journal
Article

Biophysical Society

A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein

Tse-Chiang Huang,[1] Robert Levenson,[2] Youli Li,[3] Phillip Kohl,[3] Daniel E. Morse,[4] M. Scott Shell,[1,*] and Matthew E. Helgeson[1,*]

[1]Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, California; [2]Life Sciences, Soka University of America, Aliso Viejo, California; [3]Materials Research Laboratory, University of California Santa Barbara, Santa Barbara, California; and [4]Department of Molecular, Cellular, and Developmental Biology, University of California Santa Barbara, Santa Barbara, California

ABSTRACT Reflectin is an intrinsically disordered protein known for its ability to modulate the biophotonic camouflage of cephalopods based on its assembly-induced osmotic properties. Its reversible self-assembly into discrete, size-controlled clusters and condensed droplets are known to depend sensitively on the net protein charge, making reflectin stimuli-responsive to pH, phosphorylation, and electric fields. Despite considerable efforts to characterize this behavior, the detailed physical mechanisms of reflectin's assembly are not yet fully understood. Here, we pursue a coarse-grained molecular understanding of reflectin assembly using a combination of experiments and simulations. We hypothesize that reflectin assembly and phase behavior can be explained from a remarkably simple colloidal model whereby individual protein monomers effectively interact via a short-range attractive and long-range repulsive (SA-LR) pair potential. We parameterize a coarse-grained SA-LR interaction potential for reflectin A1 from small-angle x-ray scattering measurements, and then extend it to a range of pH values using Gouy-Chapman theory to model monomer-monomer electrostatic interactions. The pH-dependent SA-LR interaction is then used in molecular dynamics simulations of reflectin assembly, which successfully capture a number of qualitative features of reflectin, including pH-dependent formation of discrete-sized nanoclusters and liquid-liquid phase separation at high pH, resulting in a putative phase diagram for reflectin. Importantly, we find that at low pH size-controlled reflectin clusters are equilibrium assemblies, which dynamically exchange protein monomers to maintain an equilibrium size distribution. These findings provide a mechanistic understanding of the equilibrium assembly of reflectin, and suggest that colloidal-scale models capture key driving forces and interactions to explain thermodynamic aspects of native reflectin behavior. Furthermore, the success of SA-LR interactions presented in this study demonstrates the potential of a colloidal interpretation of interactions and phenomena in a range of intrinsically disordered proteins.

SIGNIFICANCE The self-assembly of intrinsically disordered proteins (IDPs) is now widely appreciated for its functional biological role and is emerging as an attractive platform for new stimuli-responsive biomaterials. This work provides new insights into the pH-modulated self-assembly of the IDP reflectin, a particularly compelling example of precisely controlled IDP assembly into clusters of well-defined size. Despite the complexity of reflectin's sequence and its conformational diversity, we demonstrate that its clustering and phase behavior are well described by a remarkably simple, experimentally derived colloidal pair interaction potential involving short-range attractions and pH-modulated, long-range electrostatic repulsions. The ability of such a colloidal model to capture this type of complex assembly suggests a broader strategy for understanding and ultimately engineering a wider range of IDPs.

INTRODUCTION

Protein-based materials have emerged as sustainable replacements in conventional polymeric formulations due to their sequence modularity (1–3), stimuli responsiveness (4–7), and diverse chemical functionality (8). Recent efforts have particularly sought to understand intrinsically disordered proteins (IDPs) as a new class of important biomaterials (8–10). Compared with globular, fibrous, or membrane proteins, IDPs are either fully or partially unstructured and, through biological triggers, can accomplish unique biological functions by modulating their configurations and

Submitted March 8, 2024, and accepted for publication July 1, 2024.

*Correspondence:. shell@ucsb.edu or helgeson@ucsb.edu

Editor: Jianhan Chan.

https://doi.org/10.1016/j.bpj.2024.07.004

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

assembly through highly cooperative and dynamic multivalent interactions (11).

The reflectin protein is an exceptional example of IDP assembly because of its unique ability to modulate the color-changing camouflage of cephalopods based on its aggregation-induced osmotic response (12–14). Among all of the known reflectin proteins, reflectin A1 (here simply termed reflectin) is the most abundant protein existing in the iridescent iridocyte cells, located in the dermis of the skin of squid. Iridocytes are known to be crucial for modulating the refractive index and spacing of their subcellular Brag lamellae, and hence the color of the skin of squid (15). The activation by phosphorylation of reflectin protein within iridocytes is triggered by a signal transduction cascade from the neuronal signal through the putative neuromodulator acetylcholine (14–16). Upon phosphorylation, reflectin assembles into discrete-sized clusters with sizes of 10–100 nm in diameter (12,17). This dynamically tunes the osmotic pressure of the fluid between the highly folded lamellae of the iridocyte outer membrane, which in turn controls the interlamellar spacing, thereby modulating the wavelength-selective diffraction of light from the iridocyte (14,18).

The dynamically controlled assembly and function of reflectin emanate from its distinctive sequence. The primary structure of reflectin comprises two distinct parts: the evolutionarily conserved domains and the linker domains between them (19). Detailed sequencing of reflectin (20,21) established that most of the histidine and other positively charged groups are located in the linkers, making these domains potentially much more positively charged and pH responsive (since histidine has a $pK_a = 6.0$)(22). Therefore, upon phosphorylation, the linker domain charges become almost completely neutralized, which is thought to dramatically decrease the electrostatic repulsion between reflectin monomers, facilitating their self-assembly (17,18). Specifically, as the net protein charge decreases, experiments show that reflectin self-assembles into size-controlled, spherical clusters, and under sufficient charge neutralization can exhibit features of liquid-liquid phase separation (LLPS) (18). As reflectin transitions between a monomeric state to clusters or LLPS, the assembly process induces a significant change in osmotic pressure, which in its native biological state reduces the interlamellar spacing of highly folded membranes constituting the lamellae of tunably reflective Bragg stacks within iridocytes (12,14,16,23).

In vivo, the charge state of reflectin is tuned biologically through phosphorylation (16). However, it has been shown that reflectin's charge state can also be regulated in vitro through pH (12) and external electric fields (24,25) that act as in vitro surrogates for charge neutralization of the cationic protein by phosphorylation in vivo. This transferability has made reflectin of significant technological interest for engineered photonics applications in optical bioelectronics (26–28), clothing (29,30), and biomedical technologies (31). More broadly, reflectin's unique responsive self-assembly into a variety of aggregated states makes it an excellent model IDP system to learn from and to capture design principles for engineering broader classes of stimuli-responsive biomolecular materials.

Despite efforts over the past 20 years to characterize reflectin's in vitro assembly behavior, a complete mechanistic picture of the physical driving forces of reflectin assembly has yet to emerge. Previously, it was hypothesized that phosphorylation of positively charged reflectin decreases intermolecular Coulombic repulsion between protein monomers, thereby enabling self-assembly of into clusters via intermolecular attraction. To validate this hypothesis, earlier studies used pH modulation (12), low-voltage electroreduction (24), and genetic engineering (18) to mimic phosphorylation in vitro and characterize the reversible assembly of reflectin (12). These efforts suggested that the size of assembled reflectin clusters is controlled predominately by the charge state of the protein (17,24).

Despite these insights, the exact molecular origins of intermolecular attraction that provide the driving force for assembly are still unknown. Moreover, the nature of the dynamic balance of attractive and repulsive forces underlying size-controlled clustering is not well understood. Several lines of computational and experimental evidence (21,32–34) suggest that nonspecific hydrophobic interactions between conserved domains are the primary initial source of intermolecular attractions. In addition, Morse and co-workers examined the amino acid composition of the evolutionarily conserved domain, suggesting that a high frequency of cation-pi and sulfur-pi cross-links facilitate the formation of stable β sheet structures (18). On the other hand, Gorodetsky and co-workers performed molecular simulations of interacting reflectin fragments, and found multiple sulfur-pi cross-links in a self-assembled, truncated reflectin variant (21,32). Regardless, the manner by which attractive forces (hydrophobic, cation-pi, or sulfur-pi) interact with electrostatic repulsions to promote stable, pH-dependent equilibrium cluster sizes is unclear.

In this work, we hypothesize that reflectin self-assembly can be explained through a model well studied for interactions between synthetic colloids and globular proteins, the short-range attraction and the long-range repulsion (SA-LR) model (35–39). In 2004, Mossa et al. demonstrated how SA-LR colloidal interactions alone produce ground-state energetics that prefer the formation of equilibrium, discrete-sized clusters driven by a balance between the attractive and repulsive parts of the interaction potential (35). It was found that the repulsive part of the potential primarily determines the size and the shape of the clusters; specifically, weaker repulsion produces larger clusters. Subsequently, the SA-LR interaction model has been used to describe the interaction potential for a number of globular and highly folded proteins and highly folded proteins including lysozyme (40), γ-globulin (41), monoclonal antibodies (42), and ribonucleoprotein-RNA complex (43).

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

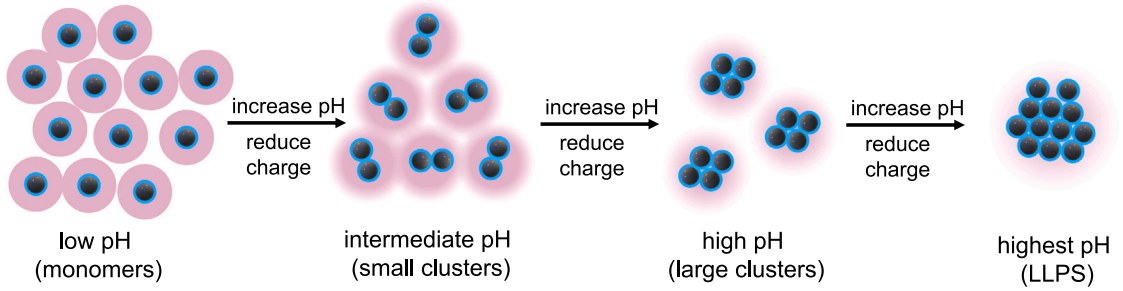

FIGURE 1 Illustration of reflectin self-assembly depicted as spherical colloids with short-range attractive and long-range repulsive interactions. Short-range attractions (*blue*) facilitate assembly of monomers into clusters, but are balanced by longer-range electrostatic repulsion (*red*) that thermodynamically limits the cluster size. Self-assembly of reflectin is triggered as the repulsive electrostatic forces are reduced in strength, proceeding from monomers to discrete-sized clusters and eventually a state of liquid-liquid phase separation (LLPS).

Despite its previous use in describing highly folded and globular protein assembly, the SA-LR model has yet to be applied to understand or explain the assembly of IDPs. This is not surprising because the idea of representing IDPs as effective, spherical colloidal-like particles neglects many complex features of their interactions—IDPs exhibit dynamic, disordered, and often coil-like conformational fluctuations. As a result, the effective pairwise interaction between two IDPs averages over many conformations and atomic interactions (electrostatic, hydrophobic, hydrogen bonding, etc.) and is likely to be soft because their polymeric structures can interpenetrate. It is therefore not obvious whether a spherical model with an isotropic SALR pair potential can appropriately capture the basic interactions between the conformationally dynamic IDPs that drive their larger-scale assembly. Nevertheless, it is encouraging that previous studies show that the self-assembled clusters of reflectin are found to be spherical and size controlled (16,23) in a manner that is similar to what has been observed in lysozyme and γ-globulin, where the phase behavior and intermolecular interactions of these globular proteins were successfully captured by an SA-LR model (41,43–45).

We hypothesize that the SA-LR model can capture the basic physics of reflectin interactions that are responsible for their assembly into size-controlled clusters (Fig. 1). From previous studies, we surmise that the long-range repulsion stems from electrostatic forces between like-charged reflectin monomers, while the short-range attraction might emerge from hydrophobic or cation-π interactions between the residues (presumably of the conserved domains, as originally suggested by Morse and co-workers (18) and Gorodetsky and co-workers (21,32)). We hypothesize that the self-assembly of reflectin is controlled through the balance of repulsion and attraction by fine-tuning the repulsion. With increasing pH, or decreasing net charges, the self-assembly of reflectin will start from monomers, progress to equilibrium size-controlled clusters, and eventually enter LLPS (46,47) as the repulsion can no longer limit the growth of the clusters (Fig. 1).

In this study, we propose a workflow that combines small-angle x-ray scattering (SAXS) and molecular dynamics simulations to test the relevance of the SA-LR model to reflectin equilibrium assembly and phase behavior (Fig. 2). To extract a putative SA-LR potential for reflectin as a function of pH, here we first employ SAXS on relatively dilute protein solutions. SAXS measurements in conjunction with liquid-state theory have been widely used to determine effective pairwise interaction potentials of systems ranging from structured proteins (48,49) to colloidal particles (50). Using SAXS modeling, we extract the SA-LR interaction potential from structure factors encoded in the scattering intensity at low pH where reflectin exists predominantly in the monomer state. We then extrapolate the interaction potential to a range of pH based on the known net charge of reflectin using the Gouy-Chapman theory for screened electrostatic interactions. We use molecular dynamics simulations of the SA-LR model to evaluate the pH-dependent equilibrium formation of colloidal reflectin assemblies, determine their size and shape, and characterize their reversibility. With the interaction potential, we are able to evaluate the equilibrium and nonequilibrium assembly behavior across a range of charge states, as tuned by pH. The resulting model is able to replicate the experimentally observed shape, size, and discrete size distribution of reflectin assembly. Furthermore, with perturbation theory and the model pH-dependent interaction potential, we use simulations and liquid-state theory to create a putative colloidal phase diagram for reflectin from the SA-LR model. This suggests that the SA-LR pair potential, which has been applied previously to globular proteins and highly folded protein systems, may also well-describe effective interactions and behaviors among IDPs such as reflectin.

METHODS

SA-LR interaction potential

We test the hypothesis that reflectin proteins can be considered highly coarse-grained particles with simple effective pairwise interactions that govern their basic assembly and phase behavior. This involves several simplifying assumptions. First, we consider reflectin proteins as spherical colloidal "particles" that interact via a SA-LR pair interaction potential. Specifically, we use the two-Yukawa model for the potential—one of the simplest nontrivial models of this family of interactions that has been

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

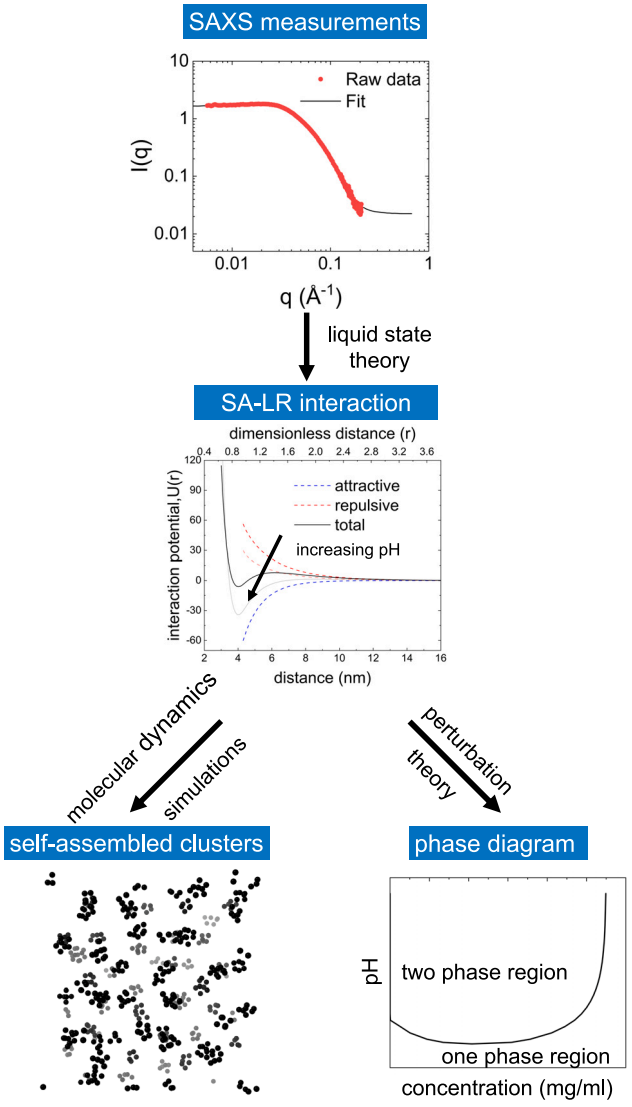

FIGURE 2 Schematic of overall modeling workflow, involving inverse fitting of SAXS data and molecular dynamics simulations. We first extract the structure factor, S(q), from experimentally measured SAXS data. The structure factor is used to estimate an interaction potential in the form of short-range attractive and long-range repulsive (SA-LR) potentials at different pH. With the pH-dependent interaction potential, molecular dynamics simulations are performed to capture the self-assembly of reflectin protein. With the same interaction potential, the phase diagram of reflectin is calculated using liquid-state theory.

successfully applied to protein systems (40,44,45,51)—to allow for arbitrary strength and range,

$$U(r) = -K_1\frac{1}{r}e^{-Z_1(r-1)} + K_2\frac{1}{r}e^{-Z_2(r-1)} \text{ for } r > 1 \qquad (1)$$

where the parameters K_i and Z_i control the strength and the range of the interactions, respectively; the first term ($i = 1$) corresponds to the attractive, while the second term ($i = 2$) corresponds to the repulsive contribution. The dimensionless (center-to-center) separation distance r is defined in units of the particle diameter (d), taken to be twice the radius of gyration of reflectin monomers. While the choice of K_1, K_2, Z_1, Z_2 allows for a wide range of potential forms, a major task in this study is to identify pH-dependent values of these parameters that effectively describe pair interactions between reflectin monomers. Specifically, here we limit the interaction parameters to $Z_1 > Z_2$ at any particular pH, such that the repulsion is always longer ranged than the attraction (i.e., an SA-LR potential).

The SA-LR model only defines the interaction when the colloidal molecules are not at contact, i.e., the dimensionless separation distance (r) is greater than the effective particle diameter (d). For smaller separations, an additional contribution to the interaction is needed to describe the short range excluded volume repulsion when two particles approach and potentially overlap. Conventionally, in liquid- and solid-state physics of small molecules and colloids, a hard sphere repulsion might be applied:

$$U(r) = \left\{ \begin{array}{ll} \infty & \text{for } r < 1 \\ 0 & \text{for } r \geq 1 \end{array} \right\} \qquad (2)$$

However, because reflectin is an IDP and behaves more like a polymer or polyelectrolyte, a softer repulsion is more suitable. As such, we use the Gaussian core model (52) for the short-range excluded volume repulsions, which provides an approximate interaction between idealized excluded volume polymer chains (53).

$$U(r) = \epsilon \times e^{\frac{r^2}{\sigma^2}} \qquad (3)$$

Here, ϵ characterizes the strength of the repulsion and σ is a dimensionless equivalent hard sphere radius. We approximate the values of these variables through liquid-state perturbation theory. Note that, at contact ($r = 0$), the potential attains a finite energy due to the fact that the centers of mass of two polymer chains can overlap without atomic overlap. The overall interaction potential is thus a piecewise combination of Eqs. 1 and 3. A linear interpolation of Eq. 1 is applied when $r \leq 1$ to match the derivative of the two-Yukawa potentials at $r = 1$, thereby ensuring piecewise continuous forces between a pair of chains. The contribution of this interpolation becomes negligible when the excluded volume repulsion represented by Eq. 3 becomes significant. With the combination of the various contributions to the potential (Eqs. 1 and 3 and the matching interpolation), we find a value of $\sigma = 0.35$. Lower values of σ would strongly affect the location and the strength of the attractive well, while higher σ would change the strength and the location of the repulsive barrier. We then calculate ϵ by matching the intermolecular radial distribution function, $g(r)$ (which at infinite dilution is related to the interaction potential by $g(r) = \exp[-\beta U(r)]$), to an equivalent system of particles with hard sphere interactions. This approximation of hard sphere repulsion with a finite and smooth potential shifts the location of the attractive minimum in the SA-LR contribution by less than 5%. This approximation is numerically tractable in molecular dynamics simulations.

SAXS

SAXS spectra of the native reflectin A1 protein were reported previously (17). Briefly, SAXS measurements were performed at beamline 4-2 of the Stanford Synchrotron Radiation Lightsource at 9 keV (1 × 0.2 mm beam size) using custom-made stopped-flow sample holder consisting of a mounted x-ray quartz capillary connected to a syringe pump. Data from 2D scattering images were obtained using a Pilatus3 X 1M 2D-detector 3.5 m from the sample (q-range: 0.0027–0.21 Å^{-1}). A flight tube under vacuum was inserted between the sample and the detector to minimize air scattering and x-ray absorption. To increase counting statistics without causing protein denaturation, 10 1-s stopped-flow experiments were performed for each sample and the resulting 1D scattering profiles were subsequently averaged. Data reduction and processing was performed using the Nika software package (54).

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

Estimation of the interaction potential from SAXS measurements

To obtain reflectin-matched parameters for the potential in Eq. 1, we use an established algorithm (50) with SAXS measurements at low pH and concentration, where reflectin is expected to exist in the monomeric, nonassembled state and thus clustering is negligible. The measured SAXS scattering intensity, $I(q)$, at wave vector q is modeled by:

$$I(q) = c\varphi_p V_p P(q) S(q) + I_{bkg} \tag{4}$$

where c is a coefficient that is determined by the instrument configuration and the scattering contrast of the material, V_p and φ_p are the molecular volume and volume fraction of particles, respectively, and I_{bkg} is the incoherent background scattering intensity. The form factor, $P(q)$, is related to the Fourier transform of the intramolecular pair distance correlation function, $p(r)$, and encodes intramolecular structure. The structure factor, $S(q)$, is related to the Fourier transform of the intermolecular radial distribution function, $g(r)$, and encodes the structure of a solution of particles, which depends sensitively on the pair interaction potential $U(r)$.

To model $P(q)$ for reflectin A1, we assume that the colloidal embodiment of a single protein monomer can be approximated as a sphere with uniform scattering length density and polydisperse radius. The inclusion of radius polydispersity is meant to capture the fluctuating conformation of the molecule, which should be significant for an IDP. First, we applied a Guinier analysis to estimate the radius of gyration, R_g, of reflectin monomers, whereby

$$\lim_{q\to 0} \ln I(q) = \ln I(q) - \frac{q^2 R_g^2}{3} \tag{5}$$

The value of R_g is later held fixed during the fitting of $P(q)$. To initially evaluate candidate models for $P(q)$, we assessed two models in which the size distribution is represented either as a Gaussian or log-normal. We found that the log-normal size distribution provides a better fit to the experimental data. R_g is assumed to be the radius (d/2) of the reflectin colloid particle in later calculations. With fixed volume fraction φ_p (determined from the protein concentration) and radius of gyration R_g (determined from the Guinier analysis), we determine a best-fit value of the polydispersity of reflectin for different concentrations (shown in supporting material, Table S1).

The resulting best fit values for $P(q)$ are then used to estimate the structure factor $S(q)$ by $S(q) = [I(q) - I_{bkg}]/P(q)$. With S(q) estimated from experiments, the radial distribution function, $g(r)$, is then given by its Fourier transform. This permits determination of the underlying pairwise interaction potential, $U(r)$, since there is a unique albeit complex relationship between the two, which is reflected in the concentration dependence of $S(q)$ (55). To perform the inverse fitting to determine $U(r)$, we use the NIST Center for Neutron Research scattering data reduction and analysis software (56,57), which implements an integral equation theory of liquids based on the Ornstein-Zernike decomposition and MSA closure, such that, for any trial potential $U(r)$, the corresponding $g(r)$ can be approximated. This algorithm iteratively refines the form of $U(r)$—here determined by the four parameters of the two-Yukawa potential $\{K_1,Z_1,K_2,Z_2\}$—until the predicted $g(r)$ matches the one extracted from the SAXS experiments.

The procedure just described estimates colloidal pair interaction potentials for reflectin from SAXS data at pH 4.5. At higher pH values, reflectin begins to cluster (12), and it becomes difficult to model the measured $S(q)$ using perturbation theory as the latter does not account for the nonmonomeric assemblies, which contribute a second peak at low q values due to cluster-cluster interactions (45). We therefore use solution electrolyte theory to extrapolate the interaction potential to higher pH, based on the expected net charge of the protein at different pH values. Specifically, we use as a baseline the potential determined at lower pH where clustering is negligible, and then assume that only the repulsive part of the potential is affected by pH and follows a Gouy-Chapman form governed by screened electrostatic interactions:

$$U(r) = \psi_0 \frac{1}{r} e^{-\kappa(r-1)} \tag{6}$$

$$\lambda_D = \kappa^{-1} = \left(\frac{\varepsilon k_b T}{e^2 \sum_i c_i z_i^2} \right)^{1/2} \tag{7}$$

$$\overline{\sigma} = \frac{2kT\varepsilon}{z_i e} \sinh \frac{z_i e \psi_0}{2kT} \tag{8}$$

where κ^{-1} is the Debye length, ε is the dielectric permittivity, k_b is the Boltzmann constant, c_i is the concentration ions of type I with charge z_i, and e is the elementary charge. The surface potential ψ_0 is related to the surface charge density ($\overline{\sigma}$). Note that by comparison to the two-Yukawa potential (Eq. 1), we have the exact relation $K_2 = \psi_0$ and $Z_2 = \kappa$. With Eqs. 6, 7, and 8, this enables us to extrapolate the value of K_2 determined at low pH to higher pH values, where the surface charge $\overline{\sigma}$ is assumed to be proportional to the net charge of the protein, with the latter reported in a previous study (17). This scheme implicitly assumes that changes in pH do not modify the attractive contributions to protein-protein interactions; we evaluate this assumption briefly when comparing the results of the model with experiments.

Colloidal reflectin molecular dynamics simulations

Using the OpenMM molecular dynamics package, we simulate colloidal models of reflectin where each protein particle interacts through the determined SA-LR pair potential. In all cases, we simulate a system with 500 particles, using periodic boundary conditions and box sizes that span the range of experimental concentrations. The simulations use a Verlet integrator (58) with the Andersen thermostat (59) to maintain the temperature at 300 K. The particles are first randomly distributed inside the simulation box, and the system is subsequently energy minimized to remove core overlaps. The simulations are then run for 2.4 billion steps, 0.4 billion steps for equilibration and 2 billion steps for production.

To address slow clustering dynamics in the simulations observed at higher pH values, we introduce a sampling methodology based on Hamiltonian replica exchange that accelerates sampling and facilitates attainment of equilibrium configurations (60,61). In this approach, we simulate in parallel many replicas of the same system that differ in pH and hence effective interaction potential. Periodically, instantaneous configurations of systems adjacent in pH are swapped via a Monte Carlo approach with the following acceptance criterion based on detailed balance:

$$P_{12,Swap}^{Acc} = \min\left[1, e^{\beta(U_1(r_1)+U_2(r_2)-U_1(r_2)-U_2(r_1))}\right] \tag{9}$$

With this method, replicas at high pH with larger energy barriers and slower sampling are able to stochastically visit lower pH states where particles more easily rearrange. This not only increases the efficiency and attainment of equilibrium, but also produces data for all pH values through a single, highly parallel simulation.

Determination of cluster sizes

To compare the simulation result with experimental dynamic light scattering (DLS) measurements, we first determine the cluster distribution by counting numbers of bonded monomers within a cluster. The bonding between two particles is defined when the surface-surface distance between them two particles is within 0.1 (0.55 nm), corresponding to the

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

approximate location of the attractive well in the potential. This choice is supported by the fact that the average distance between monomers in the clusters is around 1.05 (5.78 nm). Then, we applied two different methods to quantitatively calculate the radius. In the first method, the hydrodynamic radius is determined by assuming the packing density of the monomers is 0.64, as reported for random sphere packing (62), with the radius of the monomer equal to the volume-average radius from SAXS fitting (2.754 nm). In the second method, the radius of gyration of each cluster is calculated and the hydrodynamic radius is approximated through the relationship proposed in previous study (63).

RESULTS AND DISCUSSION

Because reflectin self-assembly is highly sensitive to its charge state, as manipulated through phosphorylation, pH (12,18), and external electric fields (24,25), we hypothesize that the assembly can be described by a model whereby a system of individual colloidal particles representing reflectin monomers interact through SA-LR interactions, with the charge state of reflectin controlling the strength of the repulsive forces. To investigate the relevance of the SA-LR model, we first use SAXS data to extract SA-LR potential parameters and then pursue molecular dynamics simulations to quantify the clustering and phase behavior implied by the model, as illustrated in Fig. 2.

Determination of a colloidal SA-LR potential for solution-phase reflectin

We use SAXS measurements to fit the four parameters of the two-Yukawa potential describing the SA-LR interaction (K_1, Z_1, K_2, Z_2). Informed by DLS(17), at pH 4.5 reflectin is intrinsically disordered and shows little self-assembly or cluster formation, existing primarily in the monomeric state (12,17,18). Therefore, we first extract the interaction potential from this low pH condition where binary interactions dominate intermolecular structure in solution, using an analytical solution (56,57) based on liquid-state theory to identify an interaction that predicts the SAXS results. To facilitate the fitting algorithm, we initialize a value for Z_2 equal to the Debye length for the buffer conditions of the sample (see Eqs. 1 and 4). A Guinier analysis is used to estimate the monomer radius of gyration, R_g, whereby a linearized plot of $\ln I(q)$ vs. q^2 can be approximated by the relation $\ln I(q) = \ln I(q=0) - \frac{q^2R_g^2}{3}$, yielding $R_g = 2.139 \pm 0.289$ nm for a reflectin monomer. This allows us to estimate the protein volume fraction, φ_p, from the measured concentration of the proteins, using $I(q=0) = \frac{4}{3}\pi R_g^3 \varphi_p (\Delta\rho)^2$, where $\Delta\rho$ is the difference in scattering length density between the protein and solvent. These data serve to both facilitate initial guesses of the adjustable fitting parameters and to assess the self-consistency of the model fitting results.

To obtain the two-Yukawa interaction potential at pH 4.5, we first fit S(q) separately at four different concentrations to generate initial values for the interaction potential parameters K_1, Z_1, K_2, and Z_2. Then, these values are used as initial guesses for performing a simultaneous, global fit (shown in Table S1) of $S(q)$ for all four concentrations, as shown in Fig. 3 *a*. The fits (*black line*) are in good agreement with the experimental data—including the low-q plateau related to $S(q = 0)$ as well as the location and magnitude of the primary correlation peak (e.g., at q $\sim$ 0.03 Å^{-1} for 200 μM). An exception is for the lower concentrations at small q values (20 and 50 μM), where the scattering is easily compromised by instrumental contributions related to beam-stop misalignment when the sample scattering is weak. As such, we avoid fitting I(q) for $q < 0.008$ Å^{-1} for these samples.

The resulting best-fit structure factors (Fig. 3 *b*) are used to generate an overall two-Yukawa interaction potential for this pH that can describe the entire concentration range. Because the default fitting algorithm based on SAXS and USAXS data reduction models $S(q)$ by a two-Yukawa potential with a hard-sphere repulsion by default (Fig. 3 *c*, *dash line*), we subsequently replace the hard-sphere repulsion with a Gaussian core model suitable to capturing effective polymer-polymer interactions (53) with parameters tuned using liquid-state perturbation theory (64) as described in the Methods, SA-LR interaction potential section. This modification affects the depth and the location of the attractive well by approximately 0.75 and 5%, respectively. The secondary repulsive barrier, which largely regulates pH-responsive assembly behavior (35,45), remains the same (Fig. 3 *c*, *red line*). This augmented soft repulsion potential is used for all subsequent calculations.

The resulting best-fit potential (Fig. 3 *c*, *red line*) shows an attractive well with a depth of $\sim$6.9 k_bT near a 4.1 nm separation distance ($\sim 1.9R_g$), and a longer-ranged repulsive barrier of strength $\sim$7.5 k_bT around 6.2 nm ($\sim 2.9R_g$). The overall magnitude of the interaction at short and long range is stronger than what has been reported for lysozyme (51), which might result from reflectin's higher net-charge (+45.47 vs. +11.06 at pH 4.5, based on the Henderson-Hasselbalch equation and reported pK_a values of ionizable groups from a previous study (65)). We hypothesize that, at this condition, because the repulsive interaction outweighs the attraction, reflectin prefers to remain monomeric due to the interprotein, long-range electrostatic repulsion.

With a pairwise interaction determined at pH 4.5, we estimate the potential at higher pH from Eqs. 6 and 8; this theoretical extrapolation is necessary since SAXS data at these conditions would contain significant contributions from clustering and higher-order interactions, as shown in a previous study (45). This would limit the use of inverse fitting procedures in this study, which are not compatible with cluster-cluster interactions. If we compare Eqs. 1 and 6, K_2 is proportional to ψ_0, which can be estimated by Eq. 8. Although the exact ψ_0 is unknown, we estimate K_2 at higher pH using the net charge of the protein, approximated

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

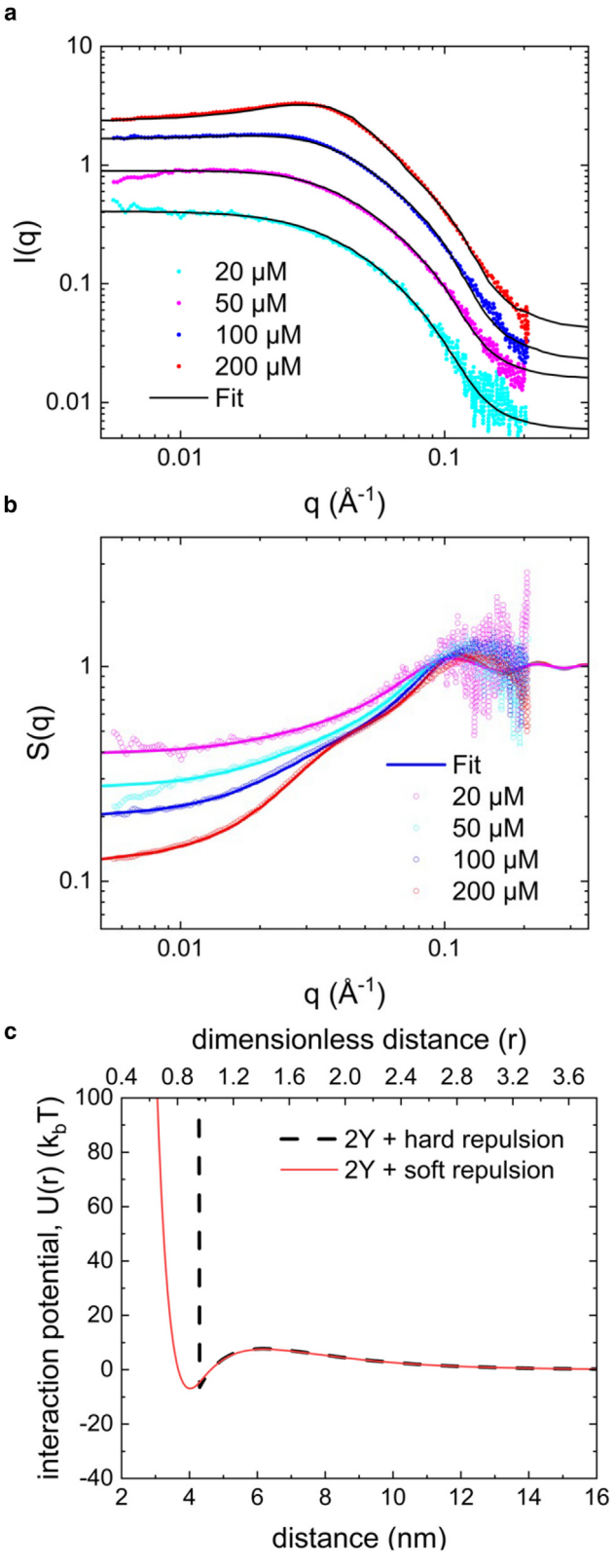

FIGURE 3 Estimation of colloidal interaction potentials from SAXS measurements. (*a*) SAXS data and fitting based on a liquid-state theoretical model of reflectin A1 at pH 4.5 with concentrations of 20, 50, 100, and 200 μM. (*b*) Structure factor of (*a*) fit with two-Yukawa potential (100 μM). (*c*) Best-fit interaction potential resulting from (*b*) (*dotted black line*), and interaction potential after the hard-sphere repulsion is replaced with a soft repulsion (*red line*).

from its amino acid composition and the Henderson-Hasselbalch equation (12,18), as shown in Eq. 10.

$$K_2 = \frac{2kT}{ze}\sinh^{-1}\left\{\frac{\overline{\sigma}}{\overline{\sigma}_{pH4.5}}\sinh\left(\frac{zeK_{2,pH\ 4.5}}{2kT}\right)\right\} \quad (10)$$

Fig. 4 *a* shows the estimated net charge of reflectin in the pH range 4.5–7.5 and the corresponding K_2 based on Eq. 10. Both values follow the same trend with increasing pH, decreasing roughly linearly from pH 4.5 to 6.5, and then decreasing more slowly around pH 7.0. Moreover, Fig. 4 *b* shows the resulting fitted and extrapolated potentials as a function of pH for reflectin interactions consisting of the two-Yukawa SA-LR model with a Gaussian core soft repulsion (data shown in Table S1). As pH increases, the depth of the attractive well grows and the height of the repulsive barrier decreases, consistent with the tendency for reflectin to form discrete clusters at elevated pH. It is the balance between these attractive and repulsive contributions to the SA-LR potential that ultimately enables the formation of equilibrium clusters of precisely tuned size. Specifically, the attractive contribution to the potential can be thought to contribute a surface tension, facilitating the formation of clusters to reduce the surface energy, whereas the repulsive potential provides an energy penalty that limits the size to which clusters can grow due to long-range electrostatic repulsion.

The depth of the attractive well increases from 6.9 k_bT to almost 55 k_bT, while the height of the repulsive barrier decreases from 7.5 k_bT to almost 0. At pH 4.5, the strength of repulsion and attraction is almost the same while the attractive region is quite narrow (~1.0 nm). However, at pH 5.5, the strength of attraction (25 k_bT) is 10 times stronger than the repulsion (2.6 k_bT), while the attractive region grows 2 times wider (~2.0 nm). As pH grows higher than 6.0, the repulsive barrier almost vanishes, such that there is no longer a mechanism to limit cluster size; at this state, we might expect reflectin to liquid-liquid phase separate.

Formation of size-controlled clusters due to SA-LR interactions

To examine the implications of the fitted SA-LR potential models for reflectin self-assembly, we use molecular dynamics simulations to interrogate the predicted equilibrium behavior of the colloidal reflectin model. We test the predictions of the simulations against experimentally observed features of reflectin assembly, including the formation of clusters with discrete size and low dispersity, the dependence of the cluster size on the pH/charge state, and the dynamic exchange of protein monomers between clusters. Importantly, we quantify the size of the clusters as a

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

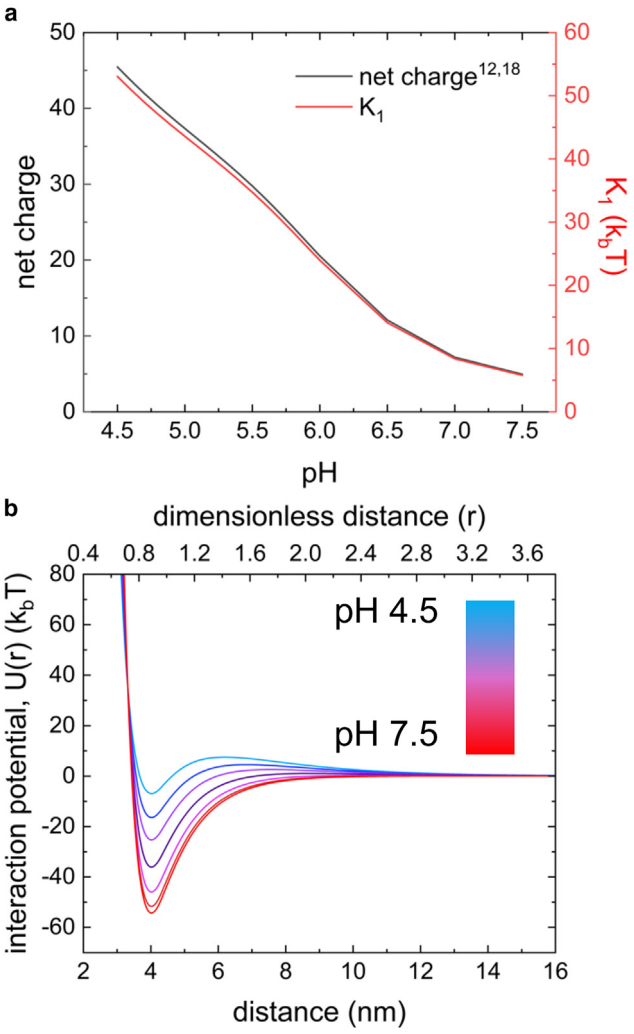

FIGURE 4 Estimated pH-dependence of colloidal interaction potential for reflectin A1 from Gouy-Chapman theory. (*a*) Net charge (12,18) of reflectin (left axis, *black*) and estimated K_2 value (right axis, *red*) in the pH range 4.5–7.5. (*b*) The resulting estimated reflectin pair interaction potential as a function of pH based on a two-Yukawa form with a Gaussian-core repulsion. With increasing pH, the attractive well increases and the repulsive barrier decreases.

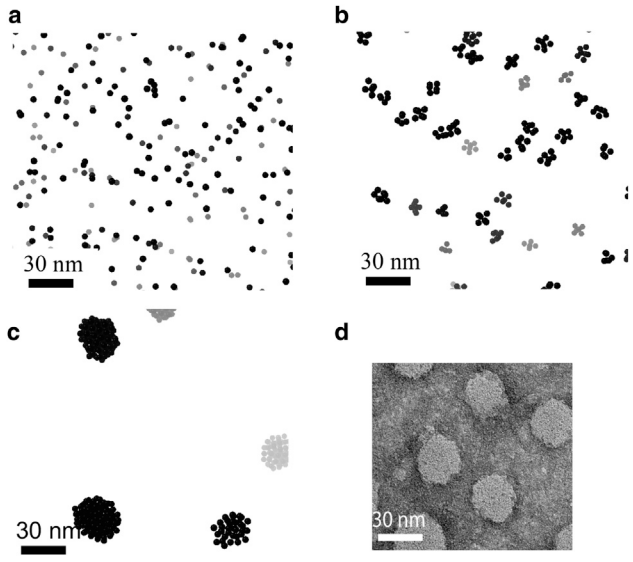

FIGURE 5 Snapshots from colloidal molecular dynamics simulations (100 μM) of reflectin A1 assembly at (*a*) pH 4.5, (*b*) pH 5.2, and (*c*) pH 7.0. (*d*) TEM of reflectin A1 clusters at pH 7.0 (12).

function of pH and the location of the phase transition toward bulk LLPS at higher pH.

We first demonstrate the influence of the growing attractive well and decreasing repulsive barrier with increasing pH on cluster formation and assembly, using simulation snapshots as shown in Fig. 5. We generally find three qualitative behaviors of the SA-LR model of reflectin, illustrated in Fig. 5, *a*, *b*, and *c*, respectively. At low pH (~4.5–4.85), almost all colloidal reflectin particles are in the free monomer state, with limited formation of clusters: ~5–10% of the monomers transiently form dimers and small clusters (typically containing 3–4 monomers). At intermediate pH (~4.85–5.5), clusters are stable and size controlled with low polydispersity. At pH > 5.5, the simulated cluster assemblies are larger (Fig. 5 *c*), and mimic the observed size and spheroidal morphology of those observed in transmission electron microscopy experiments (12) (Fig. 5 *d*).

Fig. 6 compares the average radii of the clusters obtained from simulations and experiments using DLS (17). In DLS, the reported size of protein clusters represents a scattering intensity-weighted average, where the weighting factor is proportional to the scattering intensity, I_i, of a cluster with hydrodynamic radius $R_{h,i}$, which in the Rayleigh scattering limit scales as $I_i \propto R_{h,i}^6$ (66). To compare the simulations with the experimental results, we therefore average the cluster size distributions using two different methods to quantitatively compare with the DLS measurements. In the first method (represented by the lower bound of the error area in Fig. 6), the hydrodynamic radius is determined by assuming the packing density of the monomers is 0.64, as reported for random sphere packing (62), with the radius of the monomer equal to the volume-average radius from SAXS fitting (2.754 nm). In the second method (represented by the upper bound of the error area in Fig. 6), the radius of gyration of each cluster is calculated and the hydrodynamic radius is approximated through the relationship $R_h/R_g \approx 1.29$ proposed in previous study (63).

An important consideration for both the simulations and experiments is whether the measured average cluster size (Fig. 6) represents the thermodynamic equilibrium state of the system. In simulations, this was tested by performing simulations from two different initial conditions—one in which all particles in the simulation volume initially exist as monomers (*solid line*), and another in which the volume initially contains a single cluster of all particles (*dashed line*). We find that, for pH < 5.2, these two simulations produce indistinguishable average cluster sizes, indicating that

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

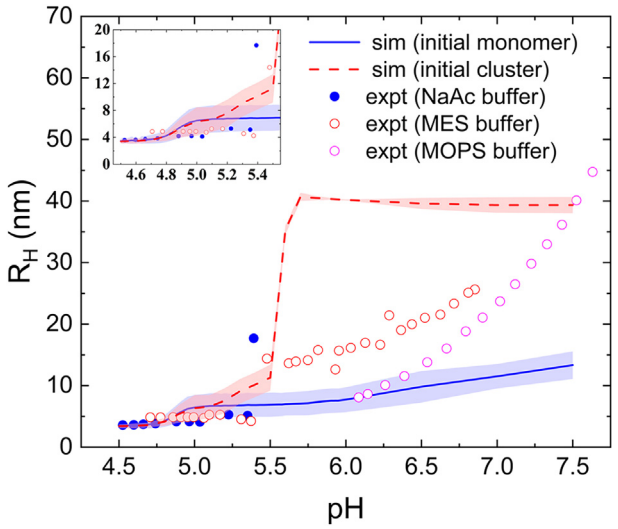

FIGURE 6 Hydrodynamic radius of assembled reflectin clusters. Lines represent simulation results, whereas points show experimental dynamic light scattering measurements (17). The blue filled points represent experiments in 20 mM sodium acetate (NaAc) buffer, which was used for SAXS measurements to inform the extracted interaction potential; empty points represent measurements in two other buffers, red for 2-(*N*-morpholino) ethane sulfonic acid (MES) and pink for 3-(*N*-morpholino) propane sulfonic acid (MOPS). The added salt concentration is maintained constant at 20 mM both for these experiments, and the concentration of reflectin A1 is 10 μM, as in simulations. The two sets of simulation results correspond to different initial structures, involving either randomly dispersed monomers (*solid line*) or a single cluster (*dashed line*). Error bars on the lines represent the possible hydrodynamic radius of the clusters between two different methods of approximating hydrodynamic radius. The upper bound represents the hydrodynamic radius calculated from radius of gyration from the simulations, using a methodology from a previous study (63). The lower bound gives the hydrodynamic radius calculated through an assumed packing density of the monomer in the clusters. The lines display the average of the two different methods.

the data in Fig. 6 represent the fully equilibrated system under these conditions. By contrast, for pH $>$ 5.2, we observe a significant discrepancy in the simulated average cluster size, suggesting that they are not able to attain equilibrium for the given run length. For cases initialized in the monomer state, the attractive well limits the exchange of monomers and therefore slows the formation of larger clusters by monomer exchange. Similarly, the attractive well limits the speed of cluster breakup for the simulations initialized with a single cluster. Therefore, we hypothesize that at moderate pH the true equilibrium size of reflectin clusters lies between the solid and the dashed lines if the equilibrium size is small so that the system size effects can be neglected.

For pH $>$ 5.6, the initialized single cluster remains intact for the extent of the simulated time (2400 ns), pinned to the system size. In these simulations, we assessed whether this result is sensitive to system size by considering simulations with 500–2000 particles, and in all cases found that the initial cluster remains stable above pH 5.5. Therefore, the hydrodynamic radius shown after pH 5.5 in Fig. 6, *dashed line*) does not correspond with the equilibrium size of a cluster. Interestingly, at pH 5.5, both simulations with a single cluster initial structure and experiments in 3-(*N*-morpholino) propane sulfonic acid (MOPS) buffer show a large step increase in the size of the assemblies. One potential explanation for these findings is that, at sufficiently high pH, the system undergoes LLPS. We test this idea later using predictions from liquid-state theory of the equilibrium phase behavior of the system.

Despite the difficulties in obtaining fully equilibrated clusters at sufficiently high pH, the comparison of these two initialization conditions successfully bounds the cluster size distribution observed in experiments (Fig. 6), where blue solid points represent experiments with sodium acetate buffer (pH 4.5–5.5), the same case used to extract the interaction potential through SAXS data. The hollow points represent other buffer conditions: red for 2-(*N*-morpholino) ethane sulfonic acid (pH 5.5–6.5) and pink for MOPS (pH 6.0–7.5).

Overall, the simulations involving initial monomer states for the protein predict cluster sizes that are quantitatively similar to those observed in experiments over most of the investigated pH range (18). For pH $<$ 4.90, both simulations and experiments produce results consistent with reflectin proteins in the monomer state (dispersed fluid). In this region, the repulsive barrier is still quite strong ($\sim$5–7 k_bT), which limits the growth of clusters that is larger than dimers. Even if a trimer is formed, the thermal energy ($\sim$1.5 k_bT) can easily destabilize such a cluster.

The small difference in R_H observed between experiments and simulation in this pH range results from differences in the monomer size extracted from SAXS (R_g = 2.15 nm) and from DLS (R_H = 3.64 nm). This difference likely arises from two factors. First, the hydrodynamic radius R_H (which characterizes the radius of an equivalent sphere with identical hydrodynamic drag coefficient) is in general larger than the radius of gyration (which characterizes the breadth of the molecular density distribution) for proteins. Second, typical DLS instrumentation usually results in imprecision in measured sizes in the single nm range due to instrumental limitations as well as the corrupting influence of any large contaminant particles.

When 4.9 $<$ pH $<$ 5.5, both simulations and experiments suggest that reflectin forms discrete clusters of narrow size distribution, which can be observed as a small step increase in hydrodynamic radius near pH $\sim$ 4.9 (12,18). In this state (which is commonly termed a "cluster fluid" in SA-LR systems), the number of free monomers decreases, and vanishes after pH 5.15 (Fig. 5 *b*). Indeed, at this condition, the repulsive barrier has a strength of 2–4 k_bT, while the attraction can be an order of magnitude stronger ($\sim$15–20 k_bT). Therefore, reflectin monomers are able to self-assemble to an equilibrium size whereby the integrated repulsive energy over the cluster balances the shorter-ranged attractive energy, resulting in size-controlled clusters. For pH $>$ 5.2, the attraction becomes so strong that reflectin acts as a

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

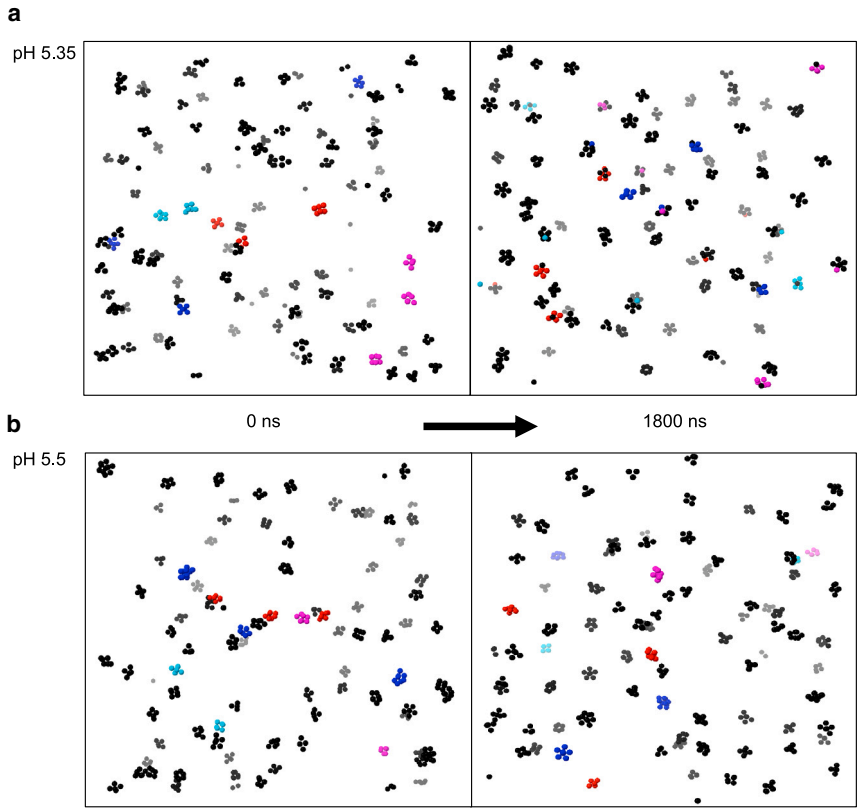

FIGURE 7 Kinetic behavior of cluster formation and monomer exchange at 100 μM for (*a*) equilibrium clustering (pH $<$ 5.35) and (*b*) kinetically limited clustering (pH $>$ 5.35). The left panels show a snapshot of the simulation at the beginning of production run, where monomers in the same clusters are labeled with identical color. The right panels show a snapshot of the simulation at the end of production run (1.8 μs). Mixing of colors between clusters over time illustrates the process of monomer exchange.

sticky particle (67) and the energy barrier for a particle to escape a cluster is so large, around 20–55 k_bT, that the kinetics of dissociation slow dramatically, which we discuss next.

Kinetics of cluster formation and monomer exchange

The evolution of the average cluster size can be understood in terms of the kinetics of assembly and disassembly of individual clusters under the action of thermal fluctuations, and their influence on the resulting cluster size distribution. To illustrate this, Fig. 7 shows representative snapshots of simulations with monomers initially labeled with different colors based on cluster membership, and presents similar snapshots a significant time later (1800 ns) to assess monomer exchange between clusters. For a protein concentration of 10 μM, the clusters are too dilute to effectively demonstrate the exchange of the monomers between different clusters since they are too infrequent or small to observe. Therefore, to better demonstrate monomer exchange, we perform simulations at 100 μM (Fig. 7). Indeed, we expect a concentration and pH dependence of the rate of monomer exchange and so we assess the exchange process for different pH and concentrations. For pH $<$ 5.20 (or 5.35 for 100 μM) (Fig. 7 *a*), the colored monomers move between different clusters, demonstrating dynamic exchange, even though the size distribution of clusters remains stable in time. We thus conclude that clustering of reflectin by SA-LR interactions is an *equilibrium* phenomenon at these conditions, whereby an equilibrium distribution of cluster sizes is achieved by a kinetically reversible process of monomer exchange between clusters.

By contrast, for pH $>$ 5.20 (or 5.35 in 100 μM) (Fig. 7 *b*), no such exchange of monomers between clusters is observed over the accessible time course of the simulations (1.8 μs). This pH range also corresponds to the conditions where the observed cluster sizes begin to show a dependence on the initial simulation configuration (monomers versus single cluster). Both findings indicate that the kinetics of the assembly and disassembly process slows as the strength of attraction is increased, such that simulations in this pH range are not able to reach equilibrium in the given simulation time.

The effects of exchange kinetics on the assembly process become more evident when examining the full distribution of cluster sizes predicted by the simulations. Experimentally, the characteristic size of self-assembled reflectin clusters is precisely regulated by pH, with a narrow polydispersity evident by electron microscopy (12). As seen in Fig. 8, the simulations initialized in the monomer state capture this low dispersity as a function of pH. For pH $<$ 4.8, the predicted size distribution is dominated by the monomer population ($R_H \sim 4$ nm). For pH 4.8–5.0,

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

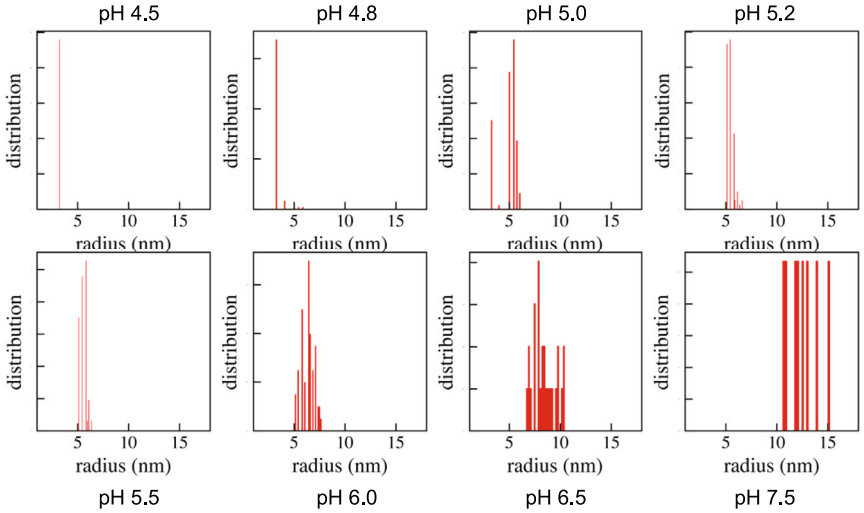

FIGURE 8 Representative size distributions of reflectin clusters at different pH based on molecular dynamics simulations, with a system size of 500 particles corresponding to 10 μM reflectin protein.

the predicted size distributions of the clusters narrows, with some monomers coexisting with a small population of relatively monodisperse clusters. For pH $>$ 5.0, the system enters the region of equilibrium clustering, where size is controlled by the height of the repulsive barrier, leading to a specific dominant cluster size and low polydispersity. While the predicted size polydispersity increases with increasing pH, this can be explained by the kinetic and size limitations of the simulation: the simulations likely do not reach equilibrated configurations, with the ultimate cluster sizes larger than those reached on the simulation timescale, or the true equilibrium cluster sizes may exceed the total number of particles in the simulations. Indeed, with 500 particles, the largest radius of the cluster is around 21.5 nm, smaller than the experimental value at pH 7.5.

We note that the pH at which we begin to observe these effects of slowed exchange kinetics on the cluster size distribution is concentration dependent, and occurs at a higher pH at elevated concentration. This might come from two different mechanisms: first, given that a higher protein concentration will in general produce smaller average distances between clusters, and therefore accelerate monomer exchange kinetics through both decreased diffusion lengths of individual "free" monomers as well as collisions between clusters, and, secondly, the possibility of different sizes and shapes of clusters can also lead to changes in the exchange kinetics.

When pH $>$ 5.5, recall that the experimentally observed cluster sizes begin to differ significantly from the simulations (either when initialized with dispersed monomers or with a large single cluster). The results just described provide three potential explanations for these deviations. First, for pH $\sim$5.5, as mentioned above, the systems are not fully equilibrated due to the slowed kinetics of monomer exchange at elevated pH, which itself results from the stronger attractive potential. In other words, protein monomers experience more durable "bonds" with neighboring monomers, and therefore require longer time to escape from a cluster, leading to slower kinetics of cluster equilibration.

A second potential explanation stems from the assumptions made to extrapolate the colloidal SA-LR model to higher pH values, in which it was assumed that only the repulsive (electrostatic) contribution to the potential changes with increasing pH as the net charge on the protein is reduced. This assumption may be inappropriate, especially for pH values approaching the point of zero charge of reflectin (pH $\sim$8) (16,18). Specifically, for pH values near the point of zero charge, one might anticipate that the attractive potential would also likely change due to differences in the chemistry of various amino acid residues in different ionization states. For example, as histidine becomes deprotonated at higher pH, it becomes relatively hydrophobic, which may contribute to an increase in the strength of the attraction. This additional attractive force is not considered in the present colloidal interaction model, and increased attraction would also lead to slowed equilibration kinetics. In addition, without the electrostatic repulsion keeping the molecule in an extended and intrinsically disordered state, interactions involving internal degrees of freedom might become important for determining the structures of the assemblies. For example, additional attractive forces from specific cation-pi and sulfur-pi cross-links shown (13,34) or other correlations between intramolecular and intermolecular structure observed in previous studies (34) may begin to dominate the interactions at elevated pH.

Finally, it is possible that, as the repulsive contribution to the potential vanishes for sufficiently high pH, the repulsive energy is no longer capable of stabilizing discrete-sized clusters, and a homogeneous protein solution may become unstable, resulting in LLPS of a condensed protein phase. The possibility of LLPS will be discussed further in the next section.

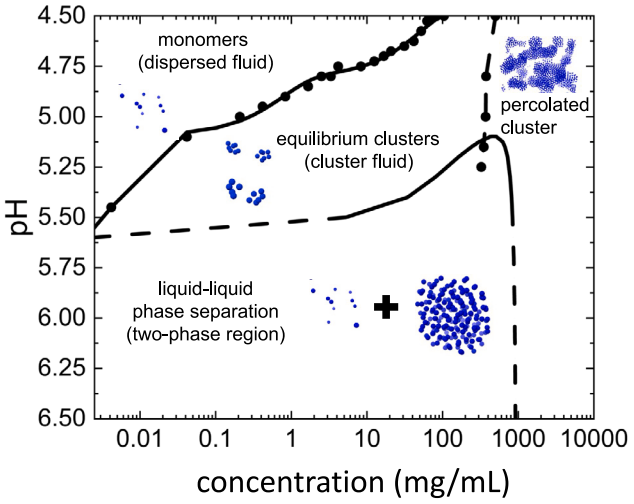

FIGURE 9 Equilibrium phase diagram predicted for the SA-LR colloidal model of reflectin in a space of pH and concentration (in mg/mL). Four different regimes are observed: the dispersed fluid, cluster fluid, percolated cluster, and two-phase region. Both the boundaries between the dispersed fluid and the cluster fluid, and between the cluster fluid and the percolated cluster, are determined by molecular dynamics simulations at different pH and concentrations (shown as *black dots*). Particles in the dispersed fluid region primarily exist as monomers, while the cluster fluid region is defined when 80% of particles participate in clusters. In the percolated cluster region, the clustering is such that a continuous path of bonded monomers exists for more than 50% of the simulation time. Lastly, the boundary of the two-phase region is calculated through liquid-state perturbation theory (supporting material, section 3). The dashed portions of the phase boundary suggest limits of the theory, which extrapolates the phase boundary to lower concentrations where the ability to distinguish two phases in the liquid-state theory falls below the numerical resolution of the calculations.

Phase diagram and LLPS driven by SA-LR interactions

Other protein systems involving SA-LR interactions are known to undergo LLPS when the strength of attraction becomes sufficiently large relative to that of the repulsion, resulting in thermodynamic instability of a homogeneous solution of clusters with respect to the formation of a macroscopic protein-rich condensate phase (46,47). To assess this possibility for the reflectin system, we use the interaction potential model obtained from SAXS (Fig. 4) with liquid-state perturbation theory (supporting material, section 3) to calculate an equilibrium phase diagram for the system (Fig. 9) in the space of pH and concentration (mg/mL). A primary assumption of these calculations involves a reference system of hard spheres to capture interparticle correlations that allow us to calculate the system free energy; while we do not expect reflectin to interact specifically like hard spheres at short range, we use this approximate model to construct the phase diagram for simplicity and the fact that it is well studied and is readily numerically calculated (64,68,69).

More generally, previous studies of the SA-LR model established that it predicts a number of distinct colloidal states (46,47). When the SA-LR potential possesses a sufficiently strong repulsive barrier, the phase diagram does not contain a region of LLPS but rather four distinct one-phase regions: a dispersed fluid (dominated by the presence of free monomers), a cluster fluid (dominated by the presence of discrete size-controlled clusters), a cluster percolated state (in which a single cluster spans the simulated volume, and a random percolated state). To test for these states, we also conducted simulations at fixed pH and increasing protein concentration to estimate the boundary where reflectin transitions from a monomer-dominant to cluster-dominant equilibrium state (following the analysis of Fig. 8), as well as the boundary where percolation of protein clusters is observed; i.e., where the clusters are sufficiently concentrated to allow the formation of a sample-spanning network. To approximate the percolation transition in the latter case, we determine the pair percolation threshold whereby a cluster contains a continuous path of "bonds" between protein monomers that spans the walls of the periodic box. The percolation state is identified only when such continuous paths exist for more than 50% of the simulation time.

Regions of LLPS typically occur only when there is either no repulsive barrier, or the repulsion is sufficiently short range (47). To quantitatively compare the present work with earlier studies, we define a relevant effective temperature on which the strength of attraction is scaled (Fig. S2). For the reflectin system, we use pH in place of an effective temperature variable that sets the strength of monomer-monomer attractions, since increases in pH lead to a decrease in the repulsive contribution and therefore an increase in the strength of attraction.

According to previous classifications on proteins modeled by SA-LR interactions (46,47), the reflectin model in this study can be classified as a "type III" SA-LR system, which is distinguished by a longer-range attraction (>20% of the core diameter). This would be, to our knowledge, the first experimental system that fits this type of SA-LR potential. Overall, however, we find that the predicted phase diagram we obtain for reflectin (Fig. 9) resembles the phase diagram of a "type II" SA-LR(47) model, including a dispersed fluid phase, an equilibrium cluster fluid at relatively low pH, a cluster percolated state at high concentration, and lastly a region of LLPS, representing coexistence between a protein-depleted monomer-rich phase and a dense liquid cluster phase. Interesting structures including so-called void phases were previously predicted in simulations of a type III SA-LR model (70). Although the present simulations were not conducted to sufficiently high concentration to test for such states, such simulation may help elucidate structures in the dense phases developed during LLPS.

The putative phase diagram provides another interpretation of the simulations of Fig. 6, which were performed at a concentration of 10 μM (0.436 mg/mL). At this concentration, reflectins exist as monomers for pH $<$ 4.9, transitioning to size-controlled equilibrium clusters at larger pH, and crossing the equilibrium LLPS phase boundary near pH $\sim$5.5–5.6.

Please cite this article in press as: Huang et al., A colloidal model for the equilibrium assembly and liquid-liquid phase separation of the reflectin A1 protein, Biophysical Journal (2024), https://doi.org/10.1016/j.bpj.2024.07.004

Interesting, the LLPS boundary is precisely where a step increase in cluster size is observed in experiments, and where the simulations initialized with a single cluster begin to show that it remains stable within the simulated runtime.

The computed phase diagram also provides a potential explanation for the differences between experiments involving different buffers (*red* and *pink points* in Fig. 6). Experimentally, the repulsive potential is controlled by the pH and the buffer concentration. The pH determines the charge state of reflectin, which affects the "bare" strength of the repulsive contribution of the potential. By contrast, increased salt (buffer) concentration will screen both the range and the strength of the repulsion, thereby reducing its contribution to the interaction potential. Therefore, as one approaches the equilibrium liquid-liquid phase boundary, the solution behavior of reflectin will become very sensitive to the specific buffer conditions. Small deviations in pH or buffer concentration will thus affect the self-assembly dramatically. For example, a slightly higher salt concentration could shift the phase transition to lower pH.

CONCLUSION

We developed a colloidal model for protein-protein interactions of the IDP reflectin involving pH-dependent SA-LR pair interactions. When simulated in colloidal molecular dynamics, the SA-LR model accurately predicts a number of experimentally observed aspects of reflectin self-assembly, and offers a number of insights into the physical mechanisms of its effective colloidal behavior. For example, the results suggest that cluster formation is dominated by an interplay of charge-regulated, screened electrostatic interactions, and nonspecific colloidal attractions between reflectin monomers. Furthermore, the observation of dynamic monomer exchange between nanoscale clusters suggests that such clusters are equilibrium assemblies with a thermodynamically preferred size and shape, as opposed to dynamically arrested structures. This dynamic exchange could potentially provide an explanation for the fast, reversible kinetics of assembly observed in vivo (13,16,23), as exchange of monomers through free solution could render reflectin more readily susceptible to dephosphorylation relative to more crowded assemblies.

Our results also provide a potential explanation for the relatively sharp change in cluster sizes at a particular pH in vitro, as the computational model predicts a transition to slowed monomer exchange and eventually LLPS at this condition, suggesting that the cluster size distribution may persist or even become trapped in an out-of-equilibrium state for extended times. The prediction of a more comprehensive phase diagram for reflectin assembly, and its relation to the underlying SA-LR model interactions, could be used to guide the design and interpretation of future experiments aimed at understanding and controlling reflectin assembly in vitro for both fundamental studies and bio-inspired applications.

Beyond reflectin, the present study demonstrates the successful transferability of the SA-LR model to IDPs. This outcome is not entirely expected given the conformational flexibility of IDPs as opposed to the more globular and conformationally rigid proteins to which it has been successfully applied previously. The reason for this success may derive from the specific nature of the intramolecular structures that underly the colloidal interactions by which reflectin assembles. Nevertheless, we find it remarkable that such a simple coarse-grained model of an IDP, involving spherical, thermodynamically averaged interactions without any intramolecular detail, is able to provide such an accurate description of reflectin assembly. Based on these results, we expect that such an approach can be generalized to other IDPs involving nonspecific charge-regulated assembly. Our results thus call for a broader exploration of the ability for the SA-LR colloidal model to predict assembly of other IDPs, to establish the range of protein structures more firmly where it can be applied to understand the function and engineering of protein assemblies.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2024.07.004.

AUTHOR CONTRIBUTIONS

T.-C.H., D.E.M., M.S.S., and M.E.H. designed the study. R.L. prepared protein samples, Y.L. and P.K. performed SAXS experiments, and T.-C.H. performed data modeling. T.-C.H. constructed the pH-dependent interaction potentials and performed molecular dynamics simulations and thermodynamic calculations. All authors wrote and revised the manuscript.

ACKNOWLEDGMENTS

This research was supported in part by the Institute for Collaborative Biotechnologies through grant W911-NF-23-1-0330 from the U.S. Army Research Office. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. This work was also supported by the BioPACIFIC Materials Innovation Platform of the National Science Foundation under award no. DMR-1933487. M.S.S. gratefully acknowledges funding support from the NSF through award no. CHEM-1800344.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Brangwynne, C. P., P. Tompa, and R. V. Pappu. 2015. Polymer physics of intracellular phase transitions. *Nat. Phys.* 11:899–904.
2. Martin, E. W., and T. Mittag. 2018. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry.* 57:2478–2487.

3. Weber, S. C. 2017. Sequence-encoded material properties dictate the structure and function of nuclear bodies. *Curr. Opin. Cell Biol.* 46:62–71.
4. Sreekumar, A., M. K. Nyati, …, A. M. Chinnaiyan. 2001. Profiling of cancer cells using protein microarrays: Discovery of novel radiation-regulated proteins. *Cancer Res.* 61:7585–7593.
5. Song, J., R. Levenson, …, D. E. Morse. 2020. Reflectin Proteins Bind and Reorganize Synthetic Phospholipid Vesicles. *Langmuir.* 36:2673–2682.
6. Naughton, K. L., L. Phan, …, A. A. Gorodetsky. 2016. Self-Assembly of the Cephalopod Protein Reflectin. *Adv. Mater.* 28:8405–8412.
7. Nagase, K., S. F. Yuk, …, T. Okano. 2011. Thermo-responsive protein adsorbing materials for purifying pharmaceuticalprotein on exposed charging surface. *J. Mater. Chem.* 21:2590–2593.
8. Keul, N. D., K. Oruganty, …, Z. A. Wood. 2018. The entropic force generated by intrinsically disordered segments tunes protein function. *Nature.* 563:584–588.
9. Santos, J., V. Iglesias, …, S. Ventura. 2020. pH-Dependent Aggregation in Intrinsically Disordered Proteins Is Determined by Charge and Lipophilicity. *Cells.* 9:145.
10. Bari, K. J., and D. D. Prakashchand. 2021. Fundamental Challenges and Outlook in Simulating Liquid-Liquid Phase Separation of Intrinsically Disordered Proteins. *J. Phys. Chem. Lett.* 12:1644–1656.
11. Uversky, V. N., J. R. Gillespie, and A. L. Fink. 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins.* 41:415–427.
12. Levenson, R., C. Bracken, …, D. E. Morse. 2016. Cyclable Condensation and Hierarchical Assembly of Metastable Reflectin Proteins, the Drivers of Tunable Biophotonics. *J. Biol. Chem.* 291:4058–4068.
13. Morse, D. E., and E. Taxon. 2020. Reflectin needs its intensity amplifier: Realizing the potential of tunable structural biophotonics. *Appl. Phys. Lett.* 117:220501.
14. DeMartini, D. G., D. V. Krogstad, and D. E. Morse. 2013. Membrane invaginations facilitate reversible water flux driving tunable iridescence in a dynamic biophotonic system. *Proc. Natl. Acad. Sci. USA.* 110:2552–2556.
15. Cooper, K. M., and R. T. Hanlon. 1986. Correlation of iridescence with changes in iridophore platelet ultrastructure in the squid Lolliguncula brevis. *J. Exp. Biol.* 121:451–455.
16. Izumi, M., A. M. Sweeney, …, D. E. Morse. 2010. Changes in reflectin protein phosphorylation are associated with dynamic iridescence in squid. *J. R. Soc. Interface.* 7:549–560.
17. Levenson, R., B. Malady, …, D. E. Morse. 2021. Protein Charge Neutralization is the Proximate Driver Dynamically Tuning a Nanoscale Bragg Reflector. Preprint at bioRxiv. https://doi.org/10.1101/2021.04.23.441158.
18. Levenson, R., C. Bracken, …, D. E. Morse. 2019. Calibration between trigger and color: Neutralization of a genetically encoded coulombic switch and dynamic arrest precisely tune reflectin assembly. *J. Biol. Chem.* 294:16804–16815.
19. Crookes, W. J., L. L. Ding, …, M. J. McFall-Ngai. 2004. Reflectins: the unusual proteins of squid reflective tissues. *Science.* 303:235–238.
20. DeMartini, D. G., M. Izumi, …, D. E. Morse. 2015. Structures, Organization, and Function of Reflectin Proteins in Dynamically Tunable Reflective Cells. *J. Biol. Chem.* 290:15238–15249.
21. Umerani, M. J., P. Pratakshya, …, A. A. Gorodetsky. 2020. Structure, self-assembly, and properties of a truncated reflectin variant. *Proc. Natl. Acad. Sci. USA.* 117:32891–32901.
22. Li, S., and M. Hong. 2011. Protonation, tautomerization, and rotameric structure of histidine: a comprehensive study by magic-angle-spinning solid-state NMR. *J. Am. Chem. Soc.* 133:1534–1544.
23. Tao, A. R., D. G. DeMartini, …, D. E. Morse. 2010. The role of protein assembly in dynamically tunable bio-optical tissues. *Biomaterials.* 31:793–801.
24. Liang, S. P., R. Levenson, …, L. Sepunaru. 2020. Electrochemistry as a surrogate for protein phosphorylation: voltage-controlled assembly of reflectin A1. *J. R. Soc. Interface.* 17:20200774. https://doi.org/10.1098/rsif.2020.0774.
25. Lin, Y. C., E. Masquelier, …, D. E. Morse. 2023. Voltage-calibrated, finely tunable protein assembly. *J. R. Soc. Interface.* 20:20230183.
26. Ordinario, D. D., L. Phan, …, A. A. Gorodetsky. 2014. Bulk protonic conductivity in a cephalopod structural protein. *Nat. Chem.* 6:596–602.
27. Qin, G., P. B. Dennis, …, D. L. Kaplan. 2013. Recombinant reflectin-based optical materials. *J. Polym. Sci. B Polym. Phys.* 51:254–264. https://doi.org/10.1002/polb.23204.
28. Torculas, M., J. Medina, …, X. Hu. 2016. Protein-Based Bioelectronics. *ACS Biomater. Sci. Eng.* 2:1211–1223.
29. Kramer, R. M., W. J. Crookes-Goodson, and R. R. Naik. 2007. The self-organizing properties of squid reflectin protein. *Nat. Mater.* 6:533–538.
30. Phan, L., R. Kautz, …, A. A. Gorodetsky. 2016. Dynamic Materials Inspired by Cephalopods. *Chem. Mater.* 28:6804–6816.
31. Hu, X., S. Ricci, …, P. Gawason. 2021. Protein and Polysaccharide-Based Electroactive and Conductive Materials for Biomedical Applications. *Molecules.* 26:4499.
32. Chatterjee, A., P. Pratakshya, …, A. A. Gorodetsky. 2023. Squid Skin Cell-Inspired Refractive Index Mapping of Cells, Vesicles, and Nanostructures. *ACS Biomater. Sci. Eng.* 9:978–990.
33. Cai, T., K. Han, …, C. Xie. 2019. Reconstruction of Dynamic and Reversible Color Change using Reflectin Protein. *Sci. Rep.* 9:5201.
34. Levenson, R., D. G. DeMartini, and D. E. Morse. 2017. Molecular mechanism of reflectin's tunable biophotonic control: Opportunities and limitations for new optoelectronics. *Apl. Mater.* 5:104801. https://doi.org/10.1063/1.4985758.
35. Mossa, S., F. Sciortino, …, E. Zaccarelli. 2004. Ground-state clusters for short-range attractive and long-range repulsive potentials. *Langmuir.* 20:10756–10763.
36. Cardinaux, F., A. Stradner, …, E. Zaccarelli. 2007. Modeling equilibrium clusters in lysozyme solutions. *Europhys. Lett.* 77:48004.
37. Stradner, A., H. Sedgwick, …, P. Schurtenberger. 2004. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature.* 432:492–495.
38. Li, W., B. A. Persson, …, M. Zackrisson Oskolkova. 2016. Concentration-Induced Association in a Protein System Caused by a Highly Directional Patch Attraction. *J. Phys. Chem. B.* 120:8953–8959.
39. Groenewold, J., and W. K. Kegel. 2001. Anomalously Large Equilibrium Clusters of Colloids. *J. Phys. Chem. B.* 105:11702–11709.
40. Liu, Y., L. Porcar, …, P. Baglioni. 2011. Lysozyme protein solution with an intermediate range order structure. *J. Phys. Chem. B.* 115:7238–7247.
41. Balbo, J., P. Mereghetti, …, R. C. Wade. 2013. The shape of protein crowders is a major determinant of protein diffusion. *Biophys. J.* 104:1576–1584.
42. Yearley, E. J., P. D. Godfrin, …, Y. Liu. 2014. Observation of small cluster formation in concentrated monoclonal antibody solutions and its implications to solution viscosity. *Biophys. J.* 106:1763–1770.
43. Alshareedah, I., T. Kaur, …, P. R. Banerjee. 2019. Interplay between Short-Range Attraction and Long-Range Repulsion Controls Reentrant Liquid Condensation of Ribonucleoprotein-RNA Complexes. *J. Am. Chem. Soc.* 141:14593–14602.
44. Phan-Xuan, T., E. Bogdanova, …, V. Kocherbitov. 2020. Hydration-Induced Structural Changes in the Solid State of Protein: A SAXS/WAXS Study on Lysozyme. *Mol. Pharm.* 17:3246–3258.
45. Stradner, A., F. Cardinaux, and P. Schurtenberger. 2006. A small-angle scattering study on equilibrium clusters in lysozyme solutions. *J. Phys. Chem. B.* 110:21222–21231.
46. Godfrin, P. D., N. E. Valadez-Perez, …, Y. Liu. 2014. Generalized phase behavior of cluster formation in colloidal dispersions with competing interactions. *Soft Matter.* 10:5061–5071.
47. Liu, Y., and Y. Xi. 2019. Colloidal systems with a short-range attraction and long-range repulsion: Phase diagrams, structures, and dynamics. *Curr. Opin. Colloid Interface Sci.* 39

48. Zhang, F., M. W. Skoda, …, F. Schreiber. 2007. Protein interactions studied by SAXS: effect of ionic strength and protein concentration for BSA in aqueous solutions. *J. Phys. Chem. B.* 111:251–259.

49. Julius, K., J. Weine, …, R. Winter. 2019. Impact of Macromolecular Crowding and Compression on Protein–Protein Interactions and Liquid–Liquid Phase Separation Phenomena. *Macromolecules.* 52:1772–1784.

50. Ryu, B. K., S. M. Fenton, …, R. N. Zia. 2022. Modeling colloidal interactions that predict equilibrium and non-equilibrium states. *J. Chem. Phys.* 156:224101.

51. Porcar, L., P. Falus, …, Y. Liu. 2009. Formation of the Dynamic Clusters in Concentrated Lysozyme Protein Solutions. *J. Phys. Chem. Lett.* 1:126–129.

52. Stillinger, F. H. 1976. Phase transitions in the Gaussian core system. *J. Chem. Phys.* 65:3968–3974.

53. Flory, P. J. 2008. Theoretical predictions on the configurations of polymer chains in the amorphous state. *J. Macromol. Sci., Part B.* 12:1–11.

54. Ilavsky, J. 2012. Nika: software for two-dimensional data reduction. *J. Appl. Crystallogr.* 45:324–328.

55. Henderson, R. L. 1974. A uniqueness theorem for fluid pair correlation functions. *Phys. Lett.* 49:197–198.

56. Kline, S. R. 2006. Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Crystallogr.* 39:895–900.

57. Liu, Y., W. R. Chen, and S. H. Chen. 2005. Cluster formation in two-Yukawa fluids. *J. Chem. Phys.* 122:44507.

58. Verlet, L. 1967. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* 159:98–103.

59. Andersen, H. C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* 72:2384–2393.

60. Affentranger, R., I. Tavernelli, and E. E. Di Iorio. 2006. A Novel Hamiltonian Replica Exchange MD Protocol to Enhance Protein Conformational Space Sampling. *J. Chem. Theor. Comput.* 2:217–228.

61. Nymeyer, H. 2008. How Efficient Is Replica Exchange Molecular Dynamics? An Analytic Approach. *J. Chem. Theor. Comput.* 4:626–636.

62. Jaeger, H. M., and S. R. Nagel. 1992. Physics of the granular state. *Science.* 255:1523–1531.

63. Nygaard, M., B. B. Kragelund, …, K. Lindorff-Larsen. 2017. An Efficient Method for Estimating the Hydrodynamic Radius of Disordered Protein Conformations. *Biophys. J.* 113:550–557.

64. Levesque, D., and L. Verlet. 1969. Perturbation Theory and Equation of State for Fluids. *Phys. Rev.* 182:307–316.

65. Thurlkill, R. L., G. R. Grimsley, …, C. N. Pace. 2006. pK values of the ionizable groups of proteins. *Protein Sci.* 15:1214–1218.

66. Berne, B. J., and R. Pecora. 2000. Dynamic light scattering: with applications to chemistry, biology, and physics. *J. Chem. Educ.* 54:A430.

67. Sedgwick, H., S. U. Egelhaaf, and W. C. K. Poon. 2004. Clusters and gels in systems of sticky particles. *J. Phys. Condens. Matter.* 16:S4913–S4922.

68. Barker, J. A., and D. Henderson. 1967. Perturbation Theory and Equation of State for Fluids. II. A Successful Theory of Liquids. *J. Chem. Phys.* 47:4714–4721.

69. Trokhymchuk, A., I. Nezbeda, …, D. Henderson. 2005. Hard-sphere radial distribution function again. *J. Chem. Phys.* 123:24501.

70. Archer, A. J., C. Ionescu, …, L. Reatto. 2008. Theory for the phase behaviour of a colloidal fluid with competing interactions. *J. Phys. Condens. Matter.* 20:415106.