



**Towards Predictive Control of Reversible Nanoparticle
Assembly with Solid-Binding Proteins**

Journal:	<i>Soft Matter</i>
Manuscript ID	SM-ART-01-2024-000094
Article Type:	Paper
Date Submitted by the Author:	22-Jan-2024
Complete List of Authors:	Cai, Yifeng; University of Washington, Chemical Engineering Qi, Xin; University of Washington, Chemical Engineering Boese, Julia; University of Washington, Chemical Engineering Zhao, Yundi; University of Washington, Chemical Engineering Hellner, Brittney; University of Washington, Chemical Engineering Chun, Jaehun; Pacific Northwest National Laboratory, Experimental & Computational Engineering Mundy, Christopher; Pacific Northwest National Laboratory, Baneyx, François; University of Washington, Chemical Engineering

Towards Predictive Control of Reversible Nanoparticle Assembly with Solid-Binding Proteins

Yifeng Cai,^a Xin Qi,^a Julia Boese,^a Yundi Zhao,^a Brittney Hellner,^a Jaehun Chun,^{b,c} Christopher J. Mundy,^{a,b} and François Baneyx*^a

^aDepartment of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States. ^bPhysical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington 99354, United States. ^cLevich Institute and Department of Chemical Engineering, CUNY City College of New York, New York, New York 10031, United States

Keywords: materials-binding peptides; dynamic materials; active matter; reconfiguration; stimuli-responsiveness; size control; multiscale modeling

Abstract – Although a broad range of ligand-functionalized nanoparticles and physico-chemical triggers have been exploited to create stimuli-responsive colloidal systems, little attention has been paid to the reversible assembly of unmodified nanoparticles with non-covalently bound proteins. Previously, we reported that a derivative of green fluorescent protein engineered with oppositely located silica-binding peptides mediates the repeated assembly and disassembly of 10-nm silica nanoparticles when pH is toggled between 7.5 and 8.5. We captured the subtle interplay between interparticle electrostatic repulsion and their protein-mediated short-range attraction with a multiscale model energetically benchmarked to collective system behavior captured by scattering experiments. Here, we show that both solution conditions (pH and ionic strength) and protein engineering (sequence and position of engineered silica-binding peptides) provide pathways for reversible control over growth and fragmentation, leading to clusters ranging in size from 25 nm protein-coated particles to micrometer-size aggregate. We further find that the higher electrolyte environment associated with successive cycles of base addition eventually eliminates reversibility. Our model accurately predicts these multiple length scales phenomena. The underpinning concepts provide design principles for the dynamic control of other protein- and particle-based nanocomposites.

Introduction

Despite significant progress over the past decades, the self-assembly of nanoscale inorganic building blocks into dynamic and non-equilibrium structures that will enable the materials and technologies of tomorrow remains difficult and inefficient.¹⁻⁷ Nanoparticle clustering and dispersion have been controlled using spontaneous crystallization,⁸ patterned substrates,^{9,10} emulsions,¹¹ light,¹² and magnetic fields.^{13,14} However, the most common approach to induce self-assembly is to functionalize nanoparticles with interacting chemical or biological ligands.¹⁵ Among the universe of available ligands, polymers have been popular to manipulate nanoparticle clustering with temperature,¹⁶⁻¹⁸ while DNA has proven valuable to create topologically complex architectures and organize nanoparticles with nanoscale precision and programmability owing to the specificity of Watson-Crick base pairing and the advent of DNA origami technologies.¹⁹⁻²² Nevertheless, the approach remains limited by the number and type of available interactions, the strength of these interactions, and the length scales at which they operate. Furthermore, reconfiguration may be challenging, especially for large systems of particles.

With their high degree of chemical and structural complexity, a broad range of functions, and amenability to engineering and design, proteins hold great promise for the construction of hierarchical and hybrid materials.²⁵ Most proteins can be endowed with the ability to bind one or more arbitrary material by genetic fusion or insertion of solid-binding peptides (SBPs) within their framework.²⁶ Some SBPs originate from biomineralizing proteins (e.g., the R5 peptide from silaffins).²⁷ However, most are 7 to 12 residues-long sequences selected by combinatorial display techniques for their affinity for a target material.²⁸ While solid-binding peptides and proteins have been extensively used to study biotic-abiotic interactions, control biomimetic mineralization, and connect inorganic and synthetic components in nanobiotechnology and biomedical applications,^{26,29-32} their potential to drive colloidal

reconfiguration in response to external triggers remains largely unexplored.

Previously, we reported that a derivative of superfolder green fluorescent proteins (sfGFP) genetically modified with two oppositely located silica-binding peptides called Car9³³ supports the repeated assembly (at pH 7.5) and disassembly (at pH 8.5) of silica nanoparticles (SiNPs; **Fig. 1**). We accurately captured the system's dynamical behavior using a multiscale framework that combines colloidal theory with molecular simulations and derives its energetics from the matching of X-ray scattering experiments (**Fig. S1**).³⁴ The model energetics explicitly or implicitly incorporate pH, ionic strength, and binding interactions as tunable input parameters, providing an avenue for predictive control of colloidal assembly and reconfiguration by engineering of solution conditions. Here, we experimentally verify the model's underlying assumption and validate predictability by demonstrating that a fine control over the mean size of protein-assembled clusters can be achieved through both solution and protein engineering. We also identify limitations to reversibility and discuss the potential of our framework for the predictive manipulation of the assembly state of other stimuli-responsive, protein-based nanomaterials.

Experimental Section

DNA manipulations and protein purification

Plasmids pET24a(+)-sfGFP-Car9, pET24a(+)-sfGFP::Car9, and pET24a(+)-sfGFP::Car9-Car9, which encode derivatives of sfGFP modified with a Car9 sequence at the C-terminus, within loop 9 of the protein, or at both locations, respectively, have been described.³³ Plasmid pET24a(+)-sfGFP::Car9-K8AK11A, which encodes a sfGFP derivative with a native Car9 sequence in loop 9 and the K8AK11A mutations in the C-terminal Car9 sequence was constructed by ligating a DNA cassette encoding sfGFP::Car9 into pET24a(+)-sfGFP-K8K11A digested with *KpnI* and *BsrGI*. Plasmid pET24a(+)-sfGFP::K8AK11A-Car9 which encodes the dual-tagged protein sfGFP::K8AK11A-Car9 was constructed by site-directed mutagenesis of the Car9 sequence in loop 9 of sfGFP::Car9-Car9 to one containing the K8AK11A substitutions. Plasmid pET24a(+)-Car9-sfGFP::Car9 which encodes the dual-tagged protein Car9-sfGFP::Car9, was constructed by ligating a DNA cassette encoding Car9-sfGFP into pET24a(+)-sfGFP::Car9 digested with *NdeI* and *BsrGI*.³⁵ All constructs were verified by DNA sequencing and plasmids were introduced into *E. coli* BL21(DE3). Cultures were grown and induced, and proteins purified by silica affinity purification as described.^{33,36} Purity was greater than 95% as judged by SDS-PAGE analysis of purified proteins.

Protein-Mediated Assembly and Disassembly of Silica Nanoparticles

Experiments were performed essentially as described.³⁴ Briefly, 10 nm nominal diameter SiNPs that encapsulate rhodamine dye and retain a native silica surface chemistry were purchased from Micromod Partikeltechnologie GmbH (Rostock, Germany). Dynamic light scattering (DLS) characterization of these particles revealed a hydrodynamic diameter (D_h) of approximately 17 nm at both pH 7.5 and 8.5,

and a zeta potential of approximately -25 mV at pH 7.5 that increased to about -35 mV at pH 8.5 (**Fig. S2**). Proteins were mixed with SiNPs at a 5 to 1 molar equivalent using proteins at a 5 μM concentration in 1 mL of 20 mM Tris-HCl, pH 7.5. We have previously determined that this molar equivalent is optimal to produce micrometer size aggregates that take days to flocculate at pH 7.5 but can be rapidly, and near quantitatively, disassembled into protein-coated nanoparticles by increasing the pH to 8.5.³⁴ Tubes were wrapped in aluminum foil and rotated at room temperature on a Dynabeads sample mixer (Invitrogen) for 30 minutes at 30 rpm before acquisition of DLS data. **Fig. S3** shows that whereas wild type sfGFP does not bind to SiNPs at pH 8.5, both the monofunctional sfGFP-Car9, and the bifunctional sfGFP::Car9-Car9 proteins do as evidenced by a progressive increase in the mean D_h of the size distributions.

For pH cycling experiments, samples were prepared in 1 mL of pH 7.5 buffer, as above, and supplied with 1M NaOH or HCl in 1 μL increments. The tube was inverted 10 times after addition of base or acid and the pH was measured with a Mettler Toledo microtip pH electrode. Once the solution reached the indicated pH (which required 8 μL of 1M NaOH or HCl for the first cycle, and 10-15 μL for subsequent cycles), the tube was rotated for 30 minutes as above and DLS data acquired. Each cycle of pH increase adds an average of 10 mM Na^+ to the system. To study the influence of cations, samples were prepared in 1mL of pH 7.5 buffer as above and supplied with 8 μL of 1M NaOH or KOH to raise the pH to 8.5. NaCl or KCl was added from 1M stock solutions to reach the indicated concentrations of Na^+ or K^+ .

Size distribution measurement using dynamic light scattering (DLS)

A Zetasizer Nano ZS (Malvern Instruments) was used to monitor the size and formation of protein-NP

aggregates. Three independent measurements comprising of 12-20 cycles were performed to collect hydrodynamic diameter (D_h) or zeta potential data under each set of experimental conditions. Full width at half maximum (FWHM) values were obtained from the distribution profiles.

Rigid body simulations

Rigid body (RB) molecular dynamics (MD) simulations were performed to obtain the structural ensemble of clusters obtained when sfGFP::Car9-Car9 is mixed with 10 nm SiNP under different solution conditions (pH and Na^+ concentration). In this highly coarse-grained model, we assume that five sfGFP::Car9-Car9 molecules are irreversibly attached to each SiNP via their C-terminal Car9 segments owing to the high affinity it confers for silica (~ 50 kJ/mol).³⁴ Long- and short-range interactions are calculated as before.³⁴ Briefly, long-range interactions consider colloidal forces between SiNPs which consist of an attractive dispersion interaction, $W_{SS}(r)$, and a repulsive electrostatic interaction using a mean-field approach,³⁷ $V_{SS}(r)$:

$$U_{SS}(r) = W_{SS}(r) + V_{SS}(r) = -\frac{AR_S}{12(r-2R_S)} + \epsilon \left(\frac{k_B T}{e}\right) \gamma^2 \frac{R_S^2}{r^2} \ln(1 + e^{-\kappa(r-2R_S)}) \quad \text{Eq. 1}$$

where A is the Hamaker constant for silica, ϵ is the dielectric constant of the solution, k_B is the Boltzmann constant, T is the temperature, e is the elementary charge, κ is the Debye screening parameter, γ is the silica surface potential which depends on solution pH, r is the center-to-center distance between two SiNPs, and R_S is the radius of the SiNPs. The short-range interaction considers the interactions of the free silica-binding peptide in a bifunctional solid-binding protein bound to a SiNP via another SBP. An expanded Lennard-Jones potential approximates the curvature-corrected binding-free energy profile obtained from a combination of metadynamics and surface element integration. Previously,³⁴ we identified $11 < \epsilon < 16$ kJ/mol as a relevant range for the attractive depth of the short-

range protein-SiNP interaction potential, and continue working with this range of ϵ here.

To investigate the influence of NaCl, we tune the relevant parameters in **Eq. 1** to reflect experimental condition. We use $R_S = 5$ nm for 10 nm diameter SiNPs and incorporate NaCl concentration effects through κ :

$$\kappa = \left(\frac{2\rho e^2}{\epsilon k_B T} \sum x_i z_i \right)^{1/2} \quad \text{Eq. 2}$$

where ρ is the concentration of NaCl.

The pH values in the RB simulations are reflected by the surface potential of amorphous silica, which is represented as y in **Eq. 1**. Since this property is challenging to measure experimentally, we used the modified Grahame equation, which relates surface charge density and surface potential via:

$$q = \frac{2\epsilon k_B T}{ez} \kappa \left(2 \sinh \left(\frac{\Psi_s}{2} \right) + \frac{4}{\kappa R} \tanh \left(\frac{\Psi_s}{4} \right) \right) \quad \text{Eq. 3}$$

where $\Psi_s = \frac{ezy}{2k_B T}$.

The surface charge density of amorphous silica at different solution pH is obtained from the work of Emami, et al.³⁸ For pH 8.0, we use a y value of -185 mV obtained by linear interpolation of the y values at pH 7.5 (-170 mV) and 8.5 (-200 mV) as input for **Eq. 1**. SiNP-SiNP interactions calculated based on **Eq. 1** were tabulated for the subsequent MD simulations.

All RB simulations were performed using the LAMMPS package, version 19Sep19.³⁹ The initial configuration for simulations contained 100 (or 800 for Na⁺ effects) protein-decorated SiNPs positioned in a 5x5x5 (or 10x10x10) cubic lattice in a cubic box with each side being 3214 Å (or 6429 Å). The pattern of surface proteins was completely randomized for each SiNP. We adopted a timestep of 5 fs and each simulation was run for 1000 to 3000 ns to ensure equilibration. We justify that the sampling time (t_{sim}) is sufficient by comparing it to the time for singlets to diffuse over the characteristic length of the largest cluster that we can observe ($t_D = 0.034$ ns) using $t_D = a^2/D$, where a is the diameter of the

maximum size obtained in MD simulation ($a = 100$ nm, see **Fig. 2A**), and D is the singlet diffusivity, $2.95 \text{ cm}^2/\text{s}$.³⁴ The last 500 frames of the equilibrated system with an interval of 1 ns were selected for data analysis. Simulations were run under the NVT ensemble using the Nosé-Hoover chain thermostat at 300 K to preserve the dynamics in all translational degree of freedom.⁴⁰

Results and discussion

Predicting the effects of the solution pH on the size of protein-assembled clusters

We have reported elsewhere that sfGFP::Car9-Car9, a derivative of superfolder green fluorescent protein (sfGFP) equipped with two oppositely located Car9 silica-binding sequences – one engineered within loop 9 of the protein and the other fused to its C-terminus (**Fig. 1A**) – assembles 10 nm silica nanoparticles (SiNPs) into micrometer size clusters when supplied at a five to one molar equivalent over particles at pH 7.5.³⁴ Remarkably, raising the solution pH by a single unit causes these aggregates to rapidly break down into individual protein-coated nanoparticles that can be repeatedly cycled between dispersed and assembled states by toggling the solution pH between 8.5 and 7.5 (**Fig. 1B**, also see **Fig. S4**).

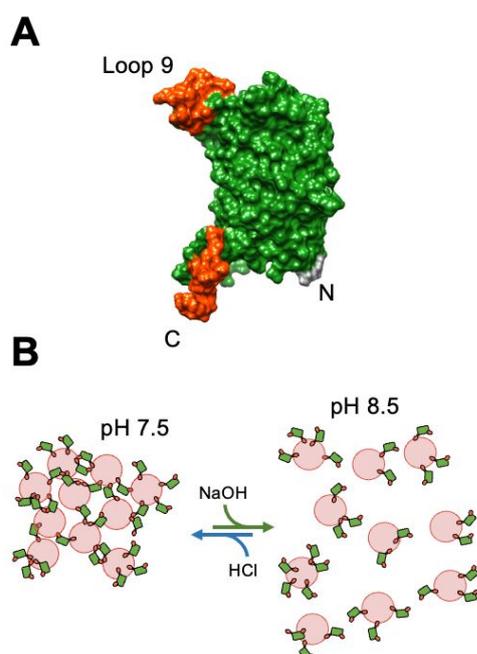


Fig. 1 pH-dependent assembly and disassembly of silica nanoparticles (SiNPs) by sfGFP::Car9-Car9. **(A)** Molecular surface of the bifunctional solid-binding protein. The sfGFP framework is colored in green and the Car9 silica-binding segments of amino acid sequence DSARGFKKPGKR are colored in orange. The location of permissive loop 9 and that of the protein's amino (N), and carboxyl (C) termini are indicated. **(B)** Schematic illustration of the reversible dispersion and assembly of SiNPs (pink spheres) when the pH is switched between 7.5 and 8.5 by addition of NaOH or HCl.

The surface of SiNPs is negatively charged above pH 2.5-3.0 and becomes progressively more negative as the pH increases due to the deprotonation of surface silanols.⁴¹ This pH-dependent change in surface chemistry affects both the long-range SiNP-SiNP electrostatic repulsion and the short-range attractive protein-SiNP interactions that are dominated by the electrostatic interaction of the Car9 peptide basic residues with SiO₂.⁴² Using a multiscale theoretical framework that characterizes the aggregation ensemble by integrating the long-range colloidal forces derived from the DLVO theory and short-range protein-silica interactions approximated from atomistic MD simulations, we captured the reversible association-dissociation behavior of the system which is rooted in a subtle interplay between SiNP-SiNP electrostatic repulsion and protein-SiNP attraction (**Fig. S1**).³⁴ Starting with dispersed protein-coated SiNPs, the model uses rigid body (RB) simulations to provide a measure of protein-mediated nanoparticle assembly as a function of the attractive depth of the protein-SiNP potential ε . The assembled state is captured by N_{\max}/N_0 , where N_{\max} is the number of particles in the largest cluster when the assembly becomes stable, and N_0 refers to monomeric, protein-decorated SiNP ($N_0 = 1$). Using sfGFP::Car9-Car9 at a 5-fold molar equivalent over 10 nm SiNP and 100 nanoparticles in the simulation box, we previously found that N_{\max}/N_0 sharply transitions from 1 to 100 (at pH 7.5) and from 100 to 1 (at pH 8.5) over a narrow band of ε (from 9.5 to 16 kJ/mol) (see ref. ³⁴ and **Fig. S5A**). Importantly, we obtained the *precise* values of ε at which these transitions occur (15 kJ/mol for pH 7.5 and 13 kJ/mol for pH 8.5), by identifying the ε values that best match ultra-small angle X-ray scattering patterns collected at 1:1, 3:1 and 5:1 protein to SiNP molar equivalents, a proxy for reduced protein-SiNP interactions which leads to the formation of smaller clusters.³⁴

While the resulting energetically benchmarked model does not explicitly include pH as an adjustable parameter, it holds potential for predicting how the assembled state changes with this variable

through the effect of pH on the surface potential of the SiNPs (*via* parameter y in **Eq. 1**). To explore this possibility, we conducted RB simulations at an intermediate pH of 8.0 and over the working range of ϵ values (**Fig. S5**). Recognizing that N_{\max}/N_0 does not provide information on size distributions and seeking to enable a more direct comparison with experimental DLS data, we extracted both the number (**Fig. S5B**) and average diameter (**Fig. 2A**) of the monomers and clusters populating the last 500 frames (250 ns) of our 1 to 3 μs RB simulations, depending on the time to reach an equilibrated structure. At all pH investigated, the hybrid system evolves from protein-decorated SiNPs that are completely separated from one another at low ϵ (i.e., 100 clusters each 20 nm in diameter) to a fully aggregated state at high ϵ (i.e., a single cluster 110 nm in diameter). However, the onset of the transition occurs at

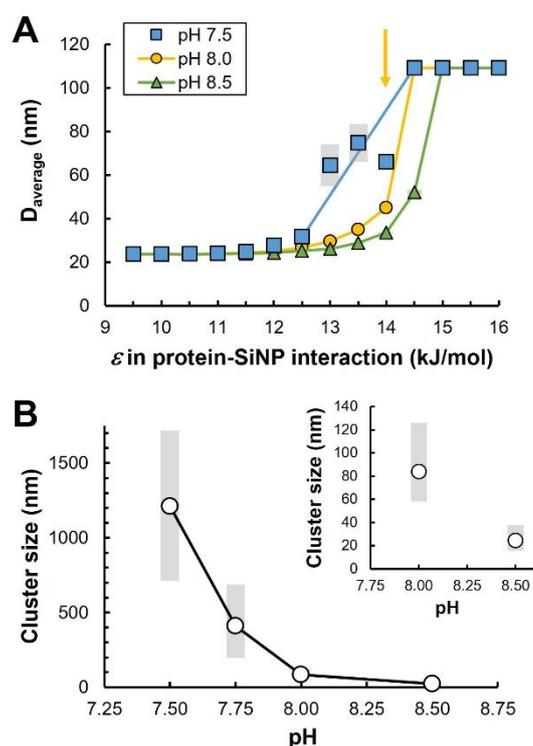


Fig. 2. Predicting and measuring the size of protein-assembled SiNP clusters at various pH. **(A)** The average diameter of clusters populating the last 500 frames of RB simulations conducted at a 5:1 molar equivalent of protein to particles and at the indicated pH are plotted vs. the depth of the protein-SiNP attractive potential. All clusters are assumed to be spherical. Bars correspond to the standard error of the mean for the frames analyzed. The yellow arrow indicates the interpolated “precise” value of ϵ at pH 8.0. **(B)** Mean D_h of clusters obtained by mixing 10 nm SiNP with a 5 molar equivalent of sfGFP::Car9-Car9 at the indicated pH. Bars correspond to the full width at half maximum (FWHM) of the size distributions.

different ϵ for each pH values (**Fig. 2A** yellow symbols, and **Fig. S5B**). Taking 14 kJ/mol as the “precise” ϵ value at pH 8.0 (a reasonable assumption considering an ϵ of 15 kJ/mol at pH 7.5 and 13 kJ/mol at pH 8.5), we find that the simulation box is mostly populated by ~ 45 nm clusters at equilibrium, and that the diameter of these clusters sharply increases as we approach pH 7.5 and the corresponding ϵ of 15 kJ/mol (**Fig. 2A** and **Fig. S5B**).

To validate these predictions, we used DLS to measure the mean hydrodynamic diameter of protein-assembled clusters in the $7.5 \leq \text{pH} \leq 8.5$ range. In good agreement with the model output, clusters have a D_h of 84 ± 34 nm at pH 8.0 and abruptly grow to hundreds of nanometers, and ultimately micrometers, as the pH approaches 7.5. We however note that little credence should be placed in the diameters that the model offers under highly aggregated conditions ($\text{pH} < 8.0$) owing to the fact that only 100 particles are used in the simulation. In short, our multiscale framework effectively captures the nonlinear influence of the solution pH on sfGFP::Car9-Car9-mediated SiNP assembly and provides an estimate of cluster size at $\text{pH} \geq 8.0$.

Predicting ionic effects and chemical fatigue

The subtle interplay between long-range SiNP-SiNP repulsion and short-range protein-SiNP attraction endows our system with a technologically desirable feature: the ability to repeatedly assemble SiNPs into large aggregates and to resolve them into protein-coated particles by switching the pH between 7.5 and 8.5. This is experimentally achieved by adding concentrated NaOH and HCl to the solution (see Experimental Section).³⁴ However, the accumulation of millimolar concentrations of Na^+ as pH cycles are repeated should affect the Debye screening parameter, weakening electrostatic repulsion between SiNPs, and shifting the system to a new balance between long- and short-range interactions.

To investigate the issue in more detail, we used DLS to collect size distributions over successive cycles of pH increase and decrease (from pH 7.5 to 8.5, and back to pH 7.5). Screening effects were observed as early as the first cycle and manifested themselves by an about 20% increase in the mean D_h of clusters produced upon return to pH 7.5 (**Fig. 3A** and **3C**). Over the second cycle, increasing the pH to 8.5 caused efficient cluster disassembly. However, a small peak persisted at 1100 nm (**Fig. 3B**) and return to pH 7.5 yielded clusters whose mean D_h had increased to ~ 2000 nm (**Fig. 3A**). These trends amplified over the third cycle, but it was during the fourth cycle – and at a cumulative carried over sodium concentration of 40 mM – that reversibility failed, with micrometer size aggregates observed at both pH (**Fig. 3C**).

The chemical fatigue associated with sodium accumulation upon pH cycling provided us with an opportunity to further test the predictive power of our model. To this end, we emulated the effect of base addition by calculating N_{\max}/N_0 values at pH 8.5 and increasing Na^+ concentrations. We increased the total number of particles to 800 to ensure sufficient statistics under aggregated conditions. At an ε of 13 kJ/mol and pH 8.5 (green arrow), NaCl and the model predicted quantitative assembly of protein-decorated SiNPs above 35 mM NaCl (**Fig. 4A**). To verify these predictions, we formed micrometer size

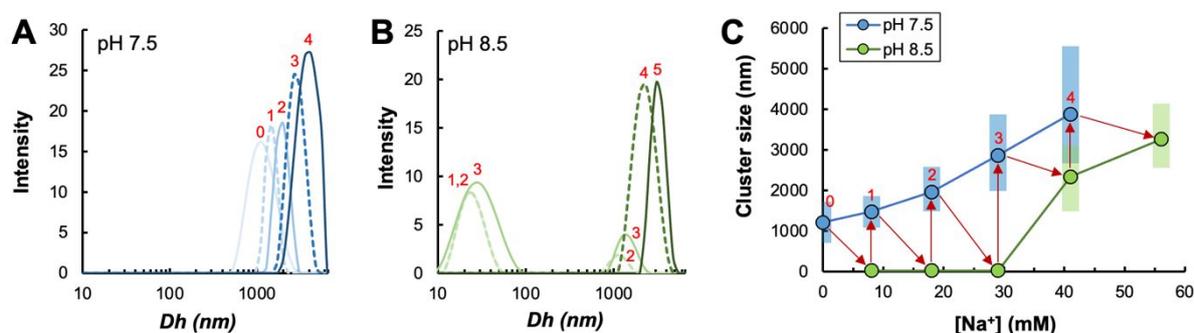


Fig. 3. Multiple cycles of chemical pH switching leads to loss of reversibility. **(A)** Size distributions of clusters obtained by mixing 10 nm SiNP with a 5 molar equivalent of sfGFP::Car9-Car9 at pH 7.5 (trace 0) and after successive cycles of pH increase to 8.5 via NaOH addition, and return to pH 7.5 by HCl addition (traces 1 to 4). **(B)** Corresponding size distributions at pH 8.5. Note that bimodal distributions are observed at pH 8.5 during the second and third cycles. **(C)** The mean D_h is plotted as a function of the calculated concentration of carried-over Na^+ associated with NaOH addition. Minor peaks at pH 8.5 are ignored. Bars represent the FWHM of the distributions.

clusters by mixing 10 nm SiNPs with a 5-fold molar equivalent of sfGFP::Car9-Car9 at pH 7.5 and adjusted the pH to 8.5 and the NaCl concentration between 10 and 40 mM. The size distributions obtained by DLS were in excellent agreement with model predictions, showing a progressive increase in mean D_h to about 65 nm in the $10 < [\text{NaCl}] < 25$ mM range, the onset of larger cluster (~ 150 nm) formation at 30 mM, a sharp transition to the micrometer size regime at 35 mM NaCl, and the production of large aggregates at higher salt concentration (**Fig. 4B-C**). Similar results were obtained when KCl was substituted for NaCl (**Table S1**), indicating that the observed effects are driven by long-range electrostatics associated with bare regions of protein-decorated SiNPs rather than molecular details. Thus, our multiscale framework is broadly useful as both a predictive and diagnosis tool.

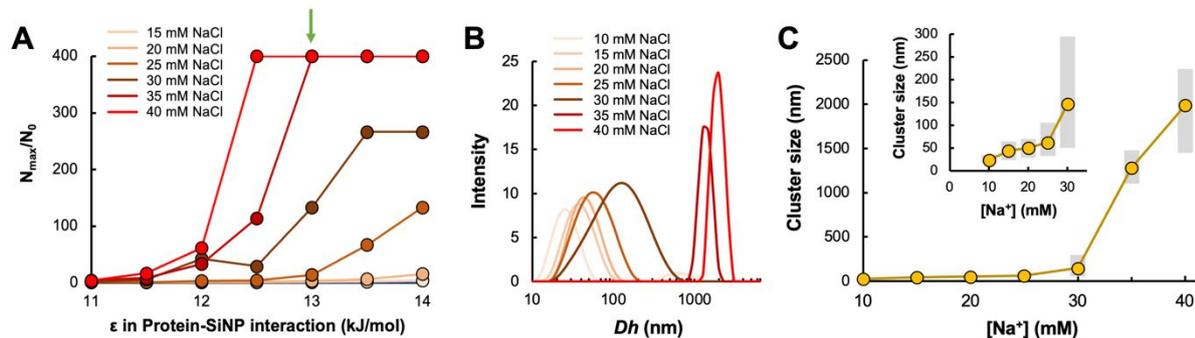


Fig. 4. Na^+ concentrations above 30 mM trigger the formation large protein-SiNP clusters at pH 8.5. **(A)** The assembled state N_{mean}/N_0 is plotted as a function of the depth of the protein-SiNP attractive potential ϵ for a 5:1 molar equivalent of sfGFP::Car9-Car9 (5 μM) to 10 nm SiNPs and at the indicated concentrations of Na^+ . A larger number of particles ($N = 800$) was used than in ref. 29 to achieve statistical significance. The maximum value of N_{mean}/N_0 is 400 because two large clusters containing 400 SiNPs rather than a single cluster containing 800 SiNPs form over the length of our simulation. The green arrow show the precise value of ϵ at pH 8.5 **(B)** Size distributions of clusters obtained by mixing 10 nm SiNP with a 5 molar equivalent of sfGFP::Car9-Car9 at pH 7.5 and changing the pH to 8.5 and NaCl concentration to the indicated values. Adjusting the pH with NaOH leads to an initial Na^+ concentration of 10 mM. **(C)** Mean D_h and FWHM (gray bars) of the size distributions shown in panel B plotted as a function of the total sodium concentration.

Manipulating SiNP assembly outcomes via protein engineering

Although the two Car9 silica-binding peptides engineered within the sfGFP framework have the same amino acid sequence, one is constrained within an internal loop of the protein while the other is fused to its C-terminus and projects into the solvent (**Fig. 1A** and **5A**). These distinct structural environments influence silica binding affinity. Indeed, Surface Plasmon Resonance (SPR) measurements supplemented by MD simulations have revealed that the “loop” Car9 binds more weakly to silica than the C-terminal Car9 at both pH 7.5 and 8.5.³⁴ Based on this observation, and on DLS data showing that a protein shell surrounds SiNPs at both pH, our model assumes that each sfGFP::Car9-Car9 molecule stably binds SiNPs through its C-terminal Car9 extension while making use of the loop Car9 to bridge primary particles together. Thus, under “standard” conditions of a 5:1 molar equivalent of proteins to nanoparticles, the average SiNP is decorated with five C-terminally bound proteins whose loop Car9 segments bind to bare silica regions on other protein-decorated SiNPs, and do so more effectively as electrostatic repulsion between SiNPs decreases with the solution pH.

To explore the potential of affinity modulation for cluster size control, we replaced Lys-8 (K8) and Lys-11 (K11) in the C-terminal Car9 extension of sfGFP::Car9-Car9 by Ala (A) residues. These substitutions lead to a substantial decrease in silica binding affinity owing to the replacement of two positively charged “anchoring” residues by neutral ones.⁴² When the resulting protein (sfGFP::Car9-K8AK11A, **Fig. 5B**) was mixed with SiNPs at a 5 to 1 molar equivalent, we observed 150 ± 60 nm clusters at pH 7.5 by DLS. These assemblies are 8-fold smaller than those observed with wild type sfGFP::Car9-Car9, underscoring the critical role of the C-terminal Car9 peptide in the formation of large aggregates (**Fig. 5E**, top panel, compare filled and open circles; also see **Table S2**).

Next, we introduced the same substitutions in the loop Car9 sequence, anticipating that the resulting sfGFP::K8AK11A-Car9 variant (**Fig. 5C**) would be fully functional for the decoration of

SiNPs via its unmodified C-terminal Car9 extension, but that it would be less efficient at bridging vicinal particles due to the lower affinity of the mutated loop Car9 for silica. Consistent with this expectation, the clusters produced at pH 7.5 were about 3-fold smaller ($D_h = 450 \pm 170$ nm) than those obtained with wild type sfGFP::Car9-Car9 (Fig. 5E, top panel, filled circles and triangles).

Finally, we relocated the C-terminal Car9 extension to the N-terminus of sfGFP::Car9. In this protein (Car9-sfGFP::Car9), the location of the fusion joint between silica-binding peptide and sfGFP framework is flipped. As result, the N-terminus of Car9 projects into the solution while its C-terminus

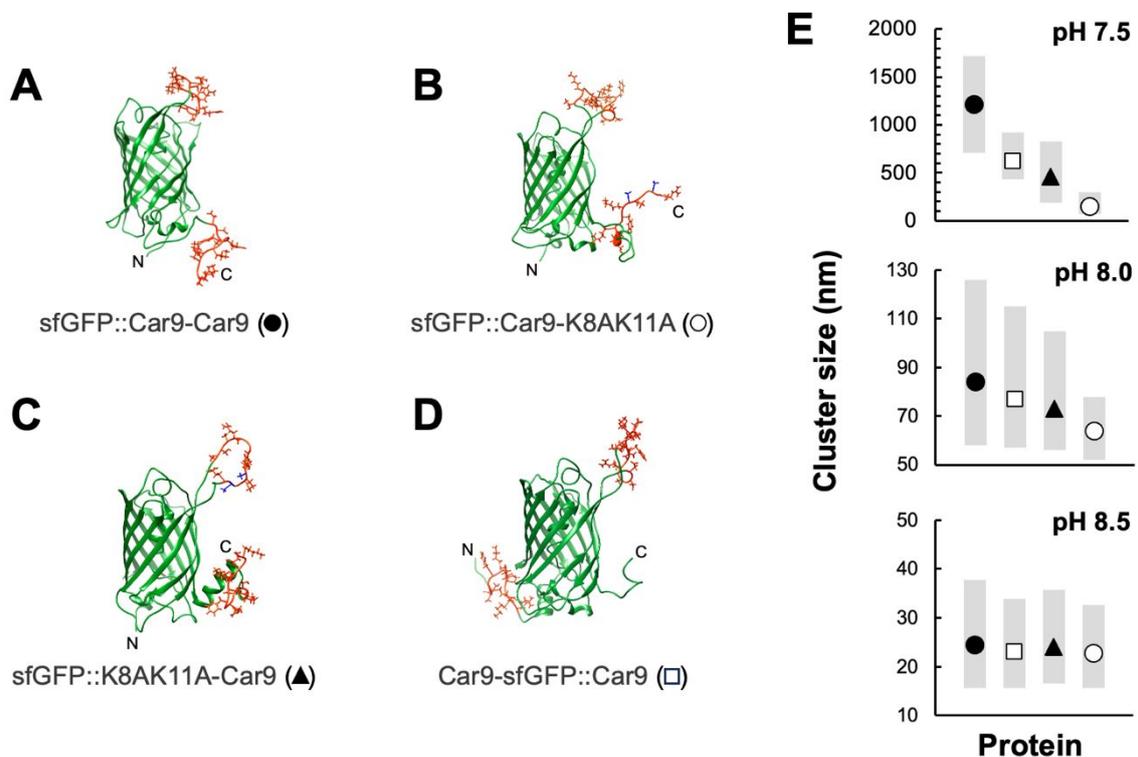


Fig. 5 Influence of SBP sequence, SBP position, and solution pH on the outcomes of protein-mediated SiNP assembly. (A-D) Ribbon structures of the bifunctional silica-binding proteins used in this study. Initial configurations were predicted using AlphaFold2 with MMseqs2 (ColabFold) and structures were relaxed by running MD simulations in aqueous solutions. The sfGFP framework is in green, the Car9 silica-binding segments in orange, and the mutated residues in blue. The location the amino (N) and carboxyl (C) termini are indicated. (E) Mean D_h and FWHM (bars) of the size distributions obtained upon mixing 5 μ M of sfGFP::Car9-Car9 (filled circles), sfGFP::Car9-K8AK11A (open circles), sfGFP::K8AK11A-Car9 (filled triangles), or Car9-sfGFP::Car9 (open squares) with SiNPs at a 5 to 1 molar equivalent and a pH of 7.5 (top panel). Aggregates disassemble into \sim 24 nm protein coated SiNPs when the pH is raised to 8.5 (bottom panel). Clusters in the 60-90 nm size range are obtained at an intermediate pH of 8.0 (middle panel).

is connected to the protein chain (**Fig. 5D**). To determine how this operation would impact silica-binding, we flowed monofunctional derivatives of sfGFP equipped with Car9 sequences at loop, C- or N-terminal locations on silica-coated SPR chips. We found that the equilibrium SPR shift – a measure of silica-binding affinity – decreased by about a third when Car9 was moved from the C- to the N-terminus of sfGFP, and that the Car9-sfGFP protein had a similar affinity for silica as the sfGFP::Car9 loop variant (**Fig. S6**). Therefore, we expected that Car9-sfGFP::Car9 should bind SiNPs equally well to SiNP through either of its Car9 segments and that this lack of directional binding coupled with a lower affinity for silica would yield clusters of intermediate size. At a mean D_h of 620 ± 250 nm at pH 7.5 (**Fig. 5E**, top panel, open squares), the aggregates obtained were indeed half the size of those obtained with sfGFP::Car9-Car9 but about 40% larger than those produced by sfGFP-K8AK11A-Car9.

For all above experiments, size distributions were sharper and narrower when scattering intensities were converted to particle numbers, indicating that a small number of large aggregates are the main contributors to polydispersity (**Fig. S7A-B** and **Table S2**). Additionally, and for all proteins examined, raising the solution pH to 8.5 led to near quantitative aggregate disassembly into protein-decorated nanoparticles of $D_h \sim 24$ nm (**Fig. 5E**, bottom panel), although a few isolated clusters persisted, as evidenced by DLS and SEM (**Fig S7C** and **Fig. S8**). Finally, and as expected from **Fig. 2**, the use of an intermediate pH of 8.0 stabilized clusters ranging in size from 84 ± 34 nm (sfGFP::Car9-Car9) to 64 ± 13 nm (sfGFP::Car9-K8AK11A), with the other two proteins yielding intermediate size assemblies of mean D_h of 77 ± 29 nm (Car9-sfGFP-Car9) and 73 ± 24 nm (sfGFP::K8AK11A-Car9) (**Fig. 5E**, middle panel).

Taken together, the above results are fully consistent with our model assumption that stable binding of SiNP by the high affinity C-terminal Car9 extension of sfGFP::Car9-Car9 is necessary for

the formation of micrometer size aggregates. Our data also demonstrate that the mean size of protein-assembled clusters can be tuned over nearly an order of magnitude by combining protein engineering approaches with solution pH adjustments in the 7.5 to 8.5 window.

Conclusions

Inspired by nature's ability to synthesize hierarchical materials that respond to environmental challenges by reconfiguration or self-healing,⁴³⁻⁴⁵ we have used bifunctional solid-binding proteins to build a dynamical protein-nanoparticle system in which the mean size of assembled clusters and assembly-disassembly behavior can be controlled by both protein design and solution conditions. Several salient features are worth highlighting. First, and unlike many stimuli-responsive systems, the surface of the nanoparticles is not chemically modified with ligands that confer desirable physico-chemical characteristics. Similarly, the SBPs that are genetically installed within the sfGFP framework do not form covalent bonds with the nanoparticles. Rather, they recognize the inorganic surface with an affinity that is modulated by both sequence and structural context. Second, from protein-decorated SiNPs to micrometer size aggregates, cluster size can be precisely tuned via protein engineering, protein to particle molar ratio, pH, and salinity. This sensitivity to protein and solution chemistry is rooted in a modulation of the relative contributions of attractive protein-SiNP interactions and repulsive SiNP-SiNP interactions over a narrow band of energy. Third, while the system can be repeatedly switched between assembled and dispersed states, the use of NaOH to raise the pH limits the numbers of productive cycles to three owing to the buildup of sodium ions and associated charge screening effects. Such chemical fatigue effects are not uncommon with stimuli-responsive systems that accumulate waste,² and could be addressed in the future by making use of photoacids and photobases to toggle the

solution pH with light.^{46–48} Finally, our energetically-benchmarked model was shown to accurately predict the influence of both pH and cations on the outcomes of assembly.

Together with prior insights, the experimental validation of the model's predictive power and the improved understanding of the sfGFP::Car9-Car9 – SiNP system garnered in this study sets the stage for the expansion of the concept to charged nanoparticles beyond silica (e.g., common and complex metal oxides, nitrides and carbides) and solid-binding proteins beyond sfGFP::Car9-Car9, including those selected or designed for self-assembly²⁵ and responsiveness to orthogonal stimuli.⁴⁹ Unlike with traditional nanoparticle functionalization schemes, the goal will not be to maximize ligand coverage but rather to leave enough of the particle surface solvent-accessible so as to enable sufficient repulsive interactions to finely balance protein-SiNP interactions that are themselves dependent on the distance separating SBPs and their relative affinity for the particle surface. We finally note that the use of anisotropic particles, the introduction of additional constraints (e.g., steric interactions),^{50,51} and the delicate coupling to translational and orientational particle dynamics,^{15,52} should provide a rich playground for the fabrication of reconfigurable protein-based materials with applications in catalysis, opto-electronics, and biomedicine.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences (BES), as part of the Energy Frontier Research Centers program: CSSAS, The Center for the Science of Synthesis Across Scales under Award Number DE-SC0019288. Theoretical analyses were partially supported by the DOE-BES Division of Materials Science and Engineering, Synthesis and Processing Sciences Program FWP 67554. PNNL is a multi-program

national laboratory operated for the Department of Energy by Battelle under Contracts No. DE-AC05-76RL01830.

Conflicts of interest

FB has a financial interest in Proteios Technology which commercializes the Car9 technology.

References

- 1 F. Li, D. P. Josephson and A. Stein, *Angew. Chem. Int. Ed.*, 2011, **50**, 360–388.
- 2 M. Grzelczak, L. M. Liz-Marzán and R. Klajn, *Chem. Soc. Rev.*, 2019, **48**, 1342–1361.
- 3 T. Bian, Z. Chu and R. Klajn, *Adv. Mater.*, 2020, **32**, 1905866.
- 4 S. Mann, *Nat. Mater.*, 2009, **8**, 781–792.
- 5 J. S. Kahn and O. Gang, *Angew. Chem. Int. Ed.*, 2022, **61**, e202105678.
- 6 Z. Li, Q. Fan and Y. Yin, *Chem. Rev.*, 2022, **122**, 4976–5067.
- 7 S. C. Glotzer and M. J. Solomon, *Nat. Mater.*, 2007, **6**, 557–562.
- 8 T. Palberg, *J. Phys. Condens. Matter*, 2014, **26**, 333101.
- 9 G. A. Ozin and S. M. Yang, *Adv. Funct. Mater.*, 2001, **11**, 95–104.
- 10 A. van Blaaderen, R. Ruel and P. Wiltzius, *Nature*, 1997, **385**, 321–324.
- 11 H. Jiang, Y. Sheng and T. Ngai, *Curr. Opin. Colloid Interface Sci.*, 2020, **49**, 1–15.
- 12 X. Chen, L. Lin, Z. Li and H.-B. Sun, *Adv. Funct. Mater.*, 2022, **32**, 2104649.
- 13 A. F. Demirörs, P. P. Pillai, B. Kowalczyk and B. A. Grzybowski, *Nature*, 2013, **503**, 99–103.
- 14 L. He, M. Wang, J. Ge and Y. Yin, *Acc. Chem. Res.*, 2012, **45**, 1431–1440.
- 15 D. Li, Q. Chen, J. Chun, K. Fichthorn, J. De Yoreo and H. Zheng, *Chem. Rev.*, 2023, **123**, 3127–3159.
- 16 R. Shenhar, T. B. Norsten and V. M. Rotello, *Adv. Mater.*, 2005, **17**, 657–669.
- 17 Y. Li, N. Wang, X. Huang, F. Li, T. P. Davis, R. Qiao and D. Ling, *ACS Appl. Bio Mater.*, 2020, **3**, 121–142.
- 18 C. Yi, Y. Yang, B. Liu, J. He and Z. Nie, *Chem. Soc. Rev.*, 2020, **49**, 465–508.
- 19 S. Y. Park, A. K. R. Lytton-Jean, B. Lee, S. Weigand, G. C. Schatz and C. A. Mirkin, *Nature*, 2008, **451**, 553–556.
- 20 R. J. Macfarlane, B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin, *Science*, 2011, **334**, 204–208.
- 21 X. Lan, Z. Su, Y. Zhou, T. Meyer, Y. Ke, Q. Wang, W. Chiu, N. Liu, S. Zou, H. Yan and Y. Liu, *Angew. Chem.*, 2017, **129**, 14824–14828.
- 22 Y. Tian, J. R. Lhermitte, L. Bai, T. Vo, H. L. Xin, H. Li, R. Li, M. Fukuto, K. G. Yager, J. S. Kahn, Y. Xiong, B. Minevich, S. K. Kumar and O. Gang, *Nat. Mater.*, 2020, **19**, 789–796.
- 23 C. Shen, X. Lan, C. Zhu, W. Zhang, L. Wang and Q. Wang, *Adv. Mater.*, 2017, **29**, 1606533.
- 24 Y. Liu, L. Ma, S. Jiang, C. Han, P. Tang, H. Yang, X. Duan, N. Liu, H. Yan and X. Lan, *ACS Nano*, 2021, **15**, 16664–16672.

- 25 L. Shao, J. Ma, J. L. Prelesnik, Y. Zhou, M. Nguyen, M. Zhao, S. A. Jenekhe, S. V. Kalinin, A. L. Ferguson, J. Pfaendtner, C. J. Mundy, J. J. De Yoreo, F. Baneyx and C.-L. Chen, *Chem. Rev.*, 2022, **122**, 17397–17478.
- 26 K. Pushpavanam, J. Ma, Y. Cai, N. Y. Naser and F. Baneyx, *Annu. Rev. Chem. Biomol. Eng.*, 2021, **12**, 333–357.
- 27 N. Kröger, R. Deutzmann and M. Sumper, *Science*, 1999, **286**, 1129–1132.
- 28 F. Baneyx and D. T. Schwartz, *Curr. Opin. Biotechnol.*, 2007, **18**, 312–317.
- 29 M. Sarikaya, C. Tamerler, A. K.-Y. Jen, K. Schulten and F. Baneyx, *Nat. Mater.*, 2003, **2**, 577–585.
- 30 M. B. Dickerson, K. H. Sandhage and R. R. Naik, *Chem. Rev.*, 2008, **108**, 4935–4978.
- 31 N. Alvisi and R. de Vries, *Mater. Today Bio*, 2023, **19**, 100580.
- 32 T. R. Walsh and M. R. Knecht, *Chem. Rev.*, 2017, **117**, 12641–12704.
- 33 B. L. Coyle and F. Baneyx, *Biotechnol. Bioeng.*, 2014, **111**, 2019–2026.
- 34 X. Qi, Y. Zhao, K. Lachowski, J. Boese, Y. Cai, O. Dollar, B. Hellner, L. Pozzo, J. Pfaendtner, J. Chun, F. Baneyx and C. J. Mundy, *ACS Nano*, 2022, **16**, 1919–1928.
- 35 B. Hellner, S. B. Lee, A. Subramaniam, V. R. Subramanian and F. Baneyx, *Langmuir*, 2019, **35**, 5013–5020.
- 36 J. Soto-Rodríguez, B. L. Coyle, A. Samuelson, K. Aravagiri and F. Baneyx, *Protein Expr. Purif.*, 2017, **135**, 70–77.
- 37 J. E. Sader, S. L. Carnie and D. Y. C. Chan, *J. Colloid Interface Sci.*, 1995, **171**, 46–54.
- 38 F. S. Emami, V. Puddu, R. J. Berry, V. Varshney, S. V. Patwardhan, C. C. Perry and H. Heinz, *Chem. Mater.*, 2014, **26**, 5725–5734.
- 39 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- 40 G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
- 41 S. K. Parida, S. Dash, S. Patel and B. K. Mishra, *Adv. Colloid Interface Sci.*, 2006, **121**, 77–110.
- 42 B. Hellner, S. Alamdari, H. Pyles, S. Zhang, A. Prakash, K. G. Sprenger, J. J. De Yoreo, D. Baker, J. Pfaendtner and F. Baneyx, *J. Am. Chem. Soc.*, 2020, **142**, 2355–2363.
- 43 A. McDougal, B. Miller, M. Singh and M. Kolle, *J. Opt.*, 2019, **21**, 073001.
- 44 J. Ren, Y. Wang, Y. Yao, Y. Wang, X. Fei, P. Qi, S. Lin, D. L. Kaplan, M. J. Buehler and S. Ling, *Chem. Rev.*, 2019, **119**, 12279–12336.
- 45 J. C. Cremaldi and B. Bhushan, *Beilstein J. Nanotechnol.*, 2018, **9**, 907–935.
- 46 P. K. Kundu, D. Samanta, R. Leizrowice, B. Margulis, H. Zhao, M. Börner, T. Udayabhaskararao, D. Manna and R. Klajn, *Nat. Chem.*, 2015, **7**, 646–652.
- 47 A. Yucknovsky, S. Mondal, A. Burnstine-Townley, M. Foqara and N. Amdursky, *Nano Lett.*, 2019, **19**, 3804–3810.
- 48 L. Wimberger, S. K. K. Prasad, M. D. Peeks, J. Andréasson, T. W. Schmidt and J. E. Beves, *J. Am. Chem. Soc.*, 2021, **143**, 20758–20768.
- 49 N. Du, F. Ye, J. Sun and K. Liu, *ChemBioChem*, 2022, **23**, e202100416.
- 50 Y. Min, M. Akbulut, K. Kristiansen, Y. Golan and J. Israelachvili, *Nat. Mater.*, 2008, **7**, 527–538.
- 51 K. Larson-Smith and D. C. Pozzo, *Soft Matter*, 2011, **7**, 5339–5347.
- 52 J. J. D. Yoreo, E. Nakouzi, B. Jin, J. Chun and C. J. Mundy, *Faraday Discuss.*, 2022, **235**, 9–35.