

Understanding Supramolecular Assembly of Supercharged Proteins

Michael I. Jacobs, Prateek Bansal, Diwakar Shukla, and Charles M. Schroeder*

Cite This: *ACS Cent. Sci.* 2022, 8, 1350–1361

Read Online

ACCESS |



Metrics & More

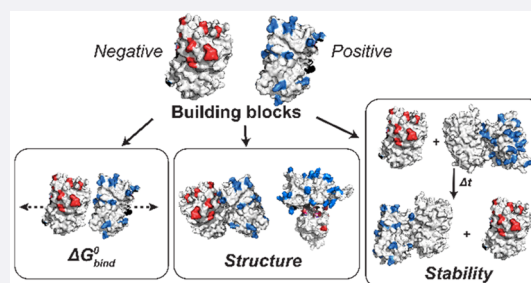


Article Recommendations



Supporting Information

ABSTRACT: Ordered supramolecular assemblies have recently been created using electrostatic interactions between oppositely charged proteins. Despite recent progress, the fundamental mechanisms governing the assembly of oppositely supercharged proteins are not fully understood. Here, we use a combination of experiments and computational modeling to systematically study the supramolecular assembly process for a series of oppositely supercharged green fluorescent protein variants. We show that net charge is a sufficient molecular descriptor to predict the interaction fate of oppositely charged proteins under a given set of solution conditions (e.g., ionic strength), but the assembled supramolecular structures critically depend on surface charge distributions. Interestingly, our results show that a large excess of charge is necessary to nucleate assembly and that charged residues not directly involved in interprotein interactions contribute to a substantial fraction (~30%) of the interaction energy between oppositely charged proteins via long-range electrostatic interactions. Dynamic subunit exchange experiments further show that relatively small, 16-subunit assemblies of oppositely charged proteins have kinetic lifetimes on the order of ~10–40 min, which is governed by protein composition and solution conditions. Broadly, our results inform how protein supercharging can be used to create different ordered supramolecular assemblies from a single parent protein building block.



INTRODUCTION

The assembly of biologically encoded molecules into ordered supramolecular structures gives rise to unique properties in natural biological systems, such as increased stability and complex functionality of macromolecular assemblies in cells.^{1–5} Inspired by nature, researchers have used biomolecular assembly strategies to develop several new biotechnologies for a range of applications, including drug delivery, sensors, and facilitated charge transport.^{6–9} In recent years, significant experimental and computational efforts have been directed at understanding the structure–function relationships of new materials constructed from assembled biological building blocks.¹⁰ However, despite recent progress, the fundamental mechanisms governing biomolecular assembly are complex and not yet fully understood, which has led to the construction of synthetic biomolecular assemblies using informed design strategies and trial and error experimentation. Overall, the development of new functional biomaterials would greatly benefit by achieving a detailed, fundamental understanding of the protein assembly process.

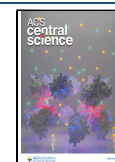
Supramolecular assembly of proteins can be driven by several different strategies including receptor–ligand, metal–ligand, electrostatic, and hydrophobic interactions.¹¹ Using *de novo* engineering of synthetic proteins, different assembled protein architectures have been created,¹² including one-dimensional (1D) nanowires,^{13–15} nanorings and nanotubes,^{16–19} and 2D and 3D protein crystals.^{14,20–24} However, protein engineering strategies for supramolecular assembly

generally require precise and targeted design of complex molecular interfaces due to the highly specific nature of the governing intermolecular interactions. Given the complexity of biomolecular interfaces and interactions, many supramolecular assembly strategies are not easily generalizable across different classes of proteins, which limits their potential use in developing new functional biological materials.

Electrostatic interactions between oppositely charged proteins are known to drive hierarchical assembly of both folded and unfolded proteins.²⁵ Electrostatic interactions between folded proteins have been used to create binary protein crystals from highly symmetric cage proteins^{26–30} as well as to localize charged proteins in Matryoshka-like cages³¹ and protein capsules.^{32–35} Supramolecular assembly of proteins via electrostatic interactions can be extended to initially uncharged proteins by supercharging protein surfaces to include charged residues.^{36,37} Recent work has shown that oppositely supercharged green fluorescent protein (GFP) variants readily assemble into an organized 16-subunit protomer, suggesting that highly symmetrical building blocks may not be necessary to drive hierarchical assembly into

Received: June 21, 2022

Published: September 13, 2022



ordered structures.³⁸ Although electrostatic interactions provide a promising method of protein supramolecular assembly that could yield synthetic biological assemblies and materials with functional properties, we lack a complete understanding of the underlying assembly process and the molecular design rules governing electrostatic-mediated protein–protein interactions.

Complexities in supercharged protein assembly arise due to the nature of multicharge and multibody interactions, which includes the role of surface charge distributions. In general, protein supercharging does not yield a uniform surface charge distribution, but rather incorporates localized charged regions dispersed over the entire protein surface. Prior computational and experimental studies on spherical and polyhedral colloidal particles have shown that both shape complementarity³⁹ as well as patchy attractive interactions^{40,41} give rise to symmetric hierarchical assembly. By changing the distribution of attractive patches on the surface of a protein, it is thought that proteins can assemble into a variety of different quaternary structures. For example, ferritin cage proteins are known to assemble into different hierarchical structures by changing the location of the attractive patches or the assembly conditions.^{20,22,23,30,42} In addition, it was reported that changing the distribution of charges on protein surfaces affects their ability to form complex coacervates with strong polyelectrolytes.^{43,44} We therefore hypothesized that different hierarchical assembled structures can be obtained from a single parent protein building block by changing the net charge and surface charge distributions on oppositely supercharged protein partner pairs.

In this work, we use a combination of experiments and computational modeling to understand the assembly mechanisms of supercharged proteins. Using a series of oppositely supercharged proteins (Figure 1), we explore how net charge and surface charge distributions on a common parent protein building block (GFP) affect intermolecular interactions and assembled structures. We further demonstrate that supercharged proteins with similar net charges but different surface charge distributions assemble into a variety of different structures with distinct properties and kinetic stability. Overall, our work provides an improved understanding of the fundamental mechanisms by which oppositely supercharged proteins assemble, and these results will be useful in guiding the rational design of new synthetic biological assemblies.

RESULTS AND DISCUSSION

Noninteracting Charged Residues Affect Supercharged Protein Assembly. We began by studying the role of noninteracting charged residues in supramolecular protein assemblies. Prior work reported the structure of an assembled 16-mer protein (two stacked octamers) composed of oppositely supercharged GFPs (denoted as GFP-17 and Ceru+32o in this work) solved at 3.47 Å resolution using cryo-electron microscopy.³⁸ At this spatial resolution, 176 specific interprotein interactions (e.g., hydrogen bonds and salt bridges) were identified as stabilizing the structure. Although 126 (72%) of the interactions involved a mutated charged amino acid, only 29% of the mutated charged amino acids were involved in stabilizing interprotein interactions in the ordered assembly.³⁸ A natural question then arises: what is the role of noninteracting, outward-facing mutated charged residues in promoting supercharged assembly?

To understand the role of noninteracting charged residues in supramolecular assembly of supercharged proteins, we

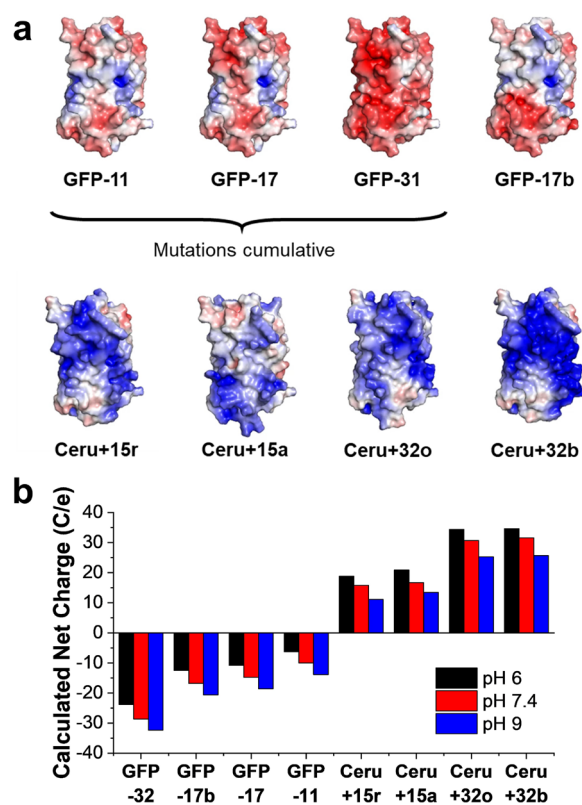


Figure 1. Supercharged GFP variants considered in this work. (a) Electrostatic surface potential representations of the solvent accessible surface area for GFP variants calculated using the linearized Poisson–Boltzmann equation with Adaptive Poisson–Boltzmann Solver (ABPS) ($\pm 10 k_B T/e$; blue for positive and red for negative). (b) Nominal net charge of the GFP and Ceru variants calculated using PROPKA at each of the pH values used here.⁵⁸

designed and expressed two minimally mutated GFP and Cerulean variants that contain only the mutations participating in interprotein interactions and do not contain any of the externally facing, noninteracting charged amino acids identified by Simon et al.³⁸ The minimally mutated variant of GFP-17 (named GFP-min) has 4 of the original 11 mutations and a nominal net charge of -8 at pH 7.4. The minimally mutated variant of Ceru+32o (named Ceru-min) has 6 of the original 24 mutations and a nominal net charge of $+3$ at pH 7.4. Clearly, the minimally mutated variants have significantly smaller net charges than the fully supercharged variants, but they retain the requisite set of mutated amino acids involved in stabilizing the interprotein interfaces in the 16-mer assembled structure.

We began by determining the Förster resonance energy transfer (FRET) efficiency between the fully supercharged proteins (Ceru+32o/GFP-17) and minimally mutated protein pairs (Ceru-min/GFP-min) as a function of ionic strength (Figure 2a). As ionic strength increases, FRET efficiency between GFP-17 and Ceru+32o decreases as the electrostatic interactions between the proteins are more effectively screened. Based on the spectral overlap between unmodified Cerulean⁴⁵ and superfolder GFP,⁴⁶ the Förster distance R_0 for the protein pair is calculated to be 5.5 nm.⁴⁷ Thus, for FRET efficiency >0.5 , the average spacing between Ceru+32o and GFP-17 is expected to be less than approximately 5.5 nm. Our results show that the FRET efficiencies between Ceru-min and GFP-min are significantly lower than those between the fully

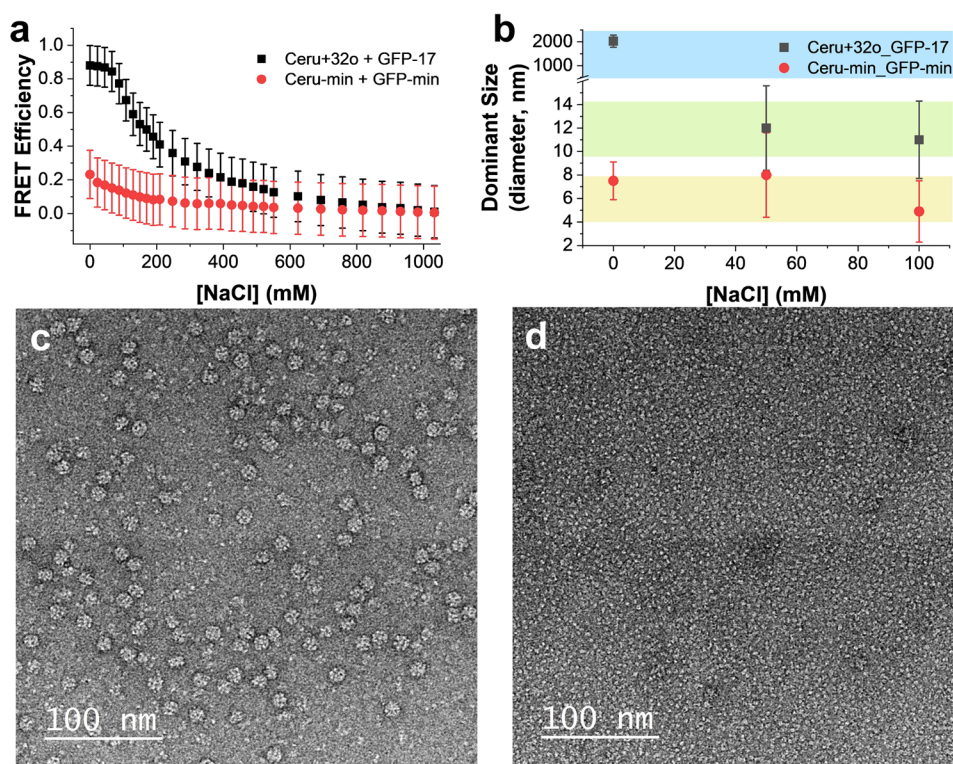


Figure 2. Characterization of fully charged and minimally charged GFP mutants. (a) FRET efficiencies between fully charged Ceru+32o/GFP-17 (black squares) and minimally mutated Ceru-min/GFP-min (red circles) at different salt concentrations. (b) Dominant assembly size as measured by DLS for fully charged and minimally mutated variants at different concentrations of NaCl. (c) Representative negative-stain TEM image of assemblies from fully charged Ceru+32o/GFP-17 at 75 mM NaCl showing protomer formation. (d) Representative negative-stain TEM image of minimally mutated Ceru-min/GFP-min at 0 mM NaCl showing no assembly or interactions between the proteins.

supercharged variants at all NaCl concentrations and are <0.5 at all salt concentrations, indicating weakly associating proteins. Ultimately, these results suggest that minimally mutated variants do not associate as strongly as the supercharged protein pairs.

Dynamic light scattering (DLS) was used to determine the average size (hydrodynamic radius) of assembled protein structures as a function of ionic strength (Figure 2b). As previously reported for assembly between the fully supercharged proteins (Cer+32o/GFP-17),³⁸ large aggregates (>1000 nm in diameter) form at low salt concentration. As salt concentration increases, intermediate particle sizes (~ 12 nm) are observed, indicating that the proteins assemble into a protomer structure. Compared to the fully supercharged variants, assemblies formed from the minimally mutated variants (Cer+min/GFP-min) have average sizes matching the monomeric species even at low salt concentrations. For example, the average diameter of particles at 0 M NaCl is 7.5 nm, which is not significantly different than the monomeric species (approximately 5–7 nm in diameter). These results are consistent with FRET measurements and suggest that Cer+min and GFP-min do not strongly associate at the range of salt concentrations considered here.

Transmission electron microscopy (TEM) was further used to directly image protein assemblies formed from oppositely charged GFP variants. Figure 2c shows a TEM micrograph of a negatively stained protein assembly between Cer+32o and GFP-17 prepared at 75 mM NaCl. An octameric ring structure is clearly present, suggesting that the proteins assemble into an ordered 16-mer, as previously reported.³⁸ However, only

unassembled proteins are observed between Cer+min and GFP-min even with no added salt (Figure 2d). Thus, our results show that the minimally mutated variants are incapable of forming the ordered 16-mer structure, and the externally facing, noninteracting charged amino acids in the fully charged variants are necessary to promote assembly.

MD Simulations and Binding Energy of Charged Protein Dimers. Based on the lack of assembly between the minimally mutated variants, we hypothesized that the noninteracting mutated charged amino acids promote assembly by two potential mechanisms: (1) the noninteracting surface charges contribute stabilizing long-range electrostatic interactions within the assembled structure, or (2) the noninteracting surface charges are required to nucleate protomer formation. To test the validity of these hypotheses, atomistic MD simulations were performed to determine the standard free energy of binding (ΔG^0_{bind}) between supercharged protein dimers and minimally mutated protein dimers. The assembled 16-mer is composed of four distinct protein–protein interfaces:³⁸ two intraplanar interfaces (named “Cer+32 Clockwise” and “GFP-17 Clockwise”, Figure 3a) and two interplanar interfaces (named “Inter-GFP-17” and “Inter-ring”, Figure 3b), where the plane is the toroidal plane of the 16-mer, normal to the axis of the torus. Figures 3c–f show the dimeric systems that were investigated: dimers with the fully supercharged and minimally mutated proteins oriented to obtain the four primary and distinct protein–protein interfaces in the 16-mer. ΔG^0_{bind} quantifies the free energy released during the association process and depends on the protein–

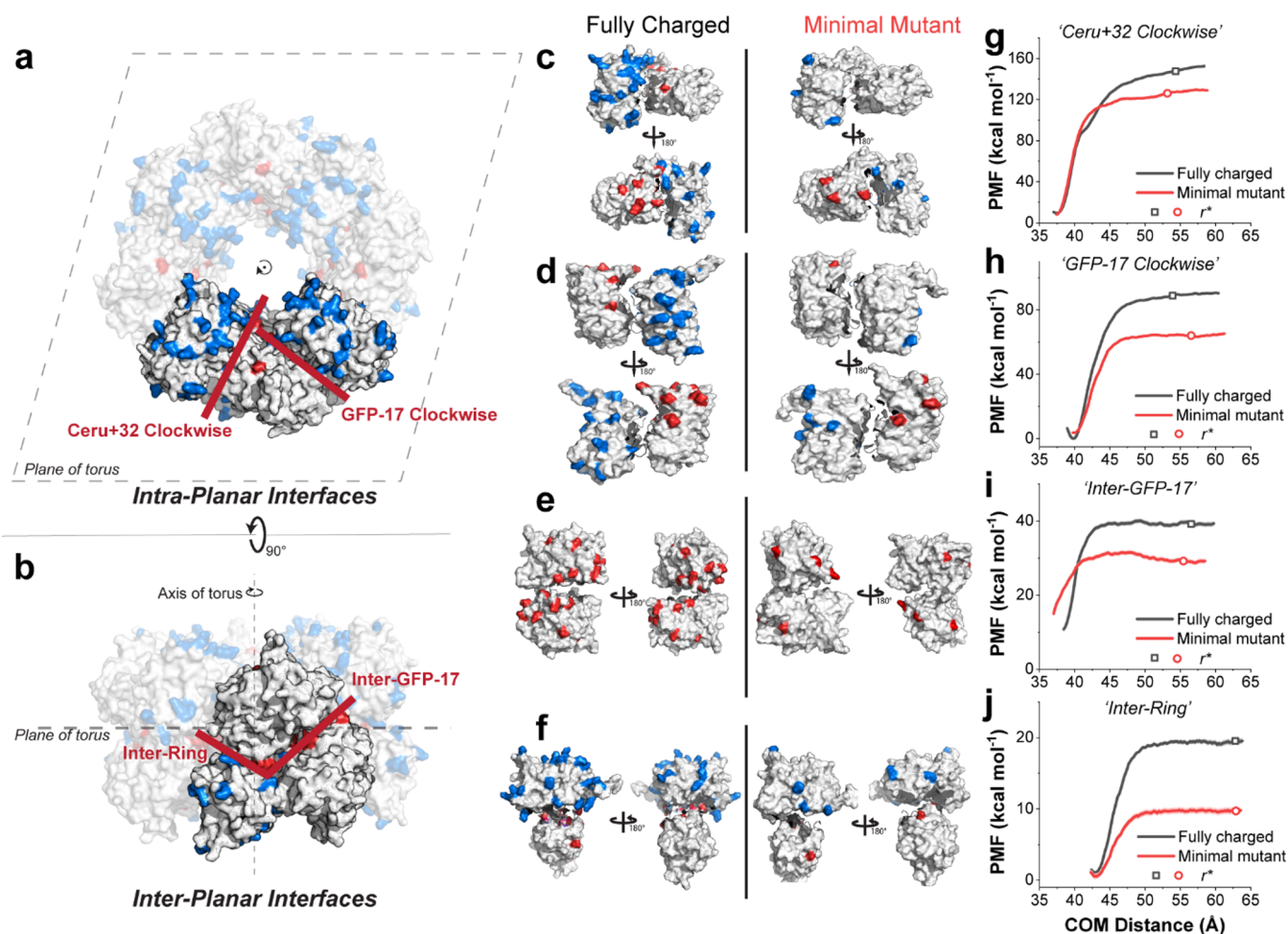


Figure 3. Binding free energy calculations for the protomeric interfaces for the fully charged and minimally mutated proteins. The four primary protein–protein interfaces comprising the assembled 16-mer identified by Simon et al.³⁸ are shown. (a) “CerU+32 Clockwise” and “GFP-17 Clockwise” are defined as intraplanar interfaces, and (b) “Inter-GFP-17” and “Inter-Ring” are defined as cross-planar interfaces, with the plane being the toroidal plane of the 16-mer, normal to the axis of the torus. The positively- and negatively charged amino acid mutations are shown in blue and red, respectively. Dimer systems that represent each of the four interfaces are shown in (c) “CerU+32 Clockwise”, (d) “GFP-17 Clockwise”, (e) “Inter-GFP-17”, and (f) “Inter-Ring”. The separation PMF values as a function of center-of-mass (COM) distance are shown for each of the interfaces (g–j).

protein interface, number of contacts, and separation distance between the proteins.

MD simulations were performed using a previously reported method to determine the absolute binding free energies using biased simulations (Figure 3 and Table 1).^{48–53} To calculate ΔG_{bind}^0 , two proteins in each system are first separated from their initial bound pose until their respective center of masses (COM) has moved by at least 25 Å and the separation potential of mean force (PMF) has plateaued, indicating that the proteins are unbound and noninteracting (Figure 3g–j). The separation PMFs for each dimeric interaction follow a

Table 1. Computed ΔG_{bind}^0 Values for Each Interface

interface	fully charged ΔG_{bind}^0 (kcal mol ⁻¹)	minimally mutated ΔG_{bind}^0 (kcal mol ⁻¹)
Ceru+32 Clockwise	-84 ± 4	-56 ± 3
GFP-17 Clockwise	-52 ± 1	-34 ± 2
Inter-GFP-17	-13.4 ± 0.4	-7.1 ± 0.7
Inter-Ring	-7.5 ± 0.6	-5.6 ± 0.6

similar trajectory as the separation distance increases, such that a nearly linear increase from the equilibrium position to a plateau value is observed. The PMF is determined as a product of the pulling force and separation distance, which allows for a threshold separation r^* to be defined where the contributions of interprotein, nonbonded interactions taper out (i.e., only long-range electrostatic interactions remain), and the force required to pull the proteins apart decreases. At separation distances larger than r^* , the PMF is observed to remain nearly constant. Separation simulations were performed under various conformational, orientational, and positional constraints using steered MD, and the contributions of the constraints to ΔG_{bind}^0 were explored using umbrella sampling and adaptive biasing force (ABF) methods (Supporting Information). In all cases, the separation PMF is found to be the largest contributor to ΔG_{bind}^0 .

The values of ΔG_{bind}^0 and separation PMFs for eight different protein dimer systems show several interesting trends. First, interfaces with more contacts in the bound state have larger values of ΔG_{bind}^0 and separation PMF (Figure S1). For example, “CerU+32 Clockwise”, an intraplanar interface with 8+ interprotein contacts, has the highest PMF at r^* and the

largest ΔG_{bind}^0 (-84 ± 4 kcal mol $^{-1}$). This is followed by the intraplanar interface “GFP-17 Clockwise” (6+ interprotein contacts, $\Delta G_{\text{bind}}^0 = -52 \pm 1$ kcal mol $^{-1}$), interplanar Inter-GFP-17 (3+ interprotein contacts, $\Delta G_{\text{bind}}^0 = -13.4 \pm 0.4$ kcal mol $^{-1}$), and finally interplanar Inter-Ring (2 interprotein contacts, $\Delta G_{\text{bind}}^0 = -7.5 \pm 0.6$ kcal mol $^{-1}$). These results suggest that the intraplanar interfaces dominate the free-energy of binding in the assembled 16-mer structure, and mutations that disrupt these interactions are expected to inhibit supramolecular assembly.

Our results show that the values of ΔG_{bind}^0 for the intraplanar interfaces are significantly larger than previously reported values for protein–protein interfaces (which are typically 10–20 kcal mol $^{-1}$).⁵⁴ This is consistent with the notion that interactions between supercharged protein mutants arise due to strong electrostatic interactions that are absent in parent proteins. However, ΔG_{bind}^0 is also a function of the size of the interface,^{55,56} so equivalently large values of ΔG_{bind}^0 are expected for larger endogenous protein–protein interfaces. Nevertheless, the ΔG_{bind}^0 values for interactions between minimally charged proteins are consistently 20–30% smaller than those between the fully charged proteins. MD simulations were further used to understand the contribution of each constraint imposed on the proteins during separation to the calculated ΔG_{bind}^0 values. Interestingly, we find that the equilibrium values for the conformational, positional, and orientational constraints are similar (typically $\pm 5\%$) for the fully charged and minimally mutated systems (Figures S2–S5 and Tables S1–S4). Therefore, these results suggest that the fully charged and minimally charged protein variants preferentially adopt similar equilibrium bound configurations. Thus, the difference in ΔG_{bind}^0 between the fully charged and minimally charged proteins is attributed to the absence of long-range electrostatic interactions in the minimally mutated variants. Due to the logarithmic nature of free energy calculations ($\Delta G_{\text{bind}} = -k_{\text{b}}T \ln K_{\text{bind}}$), small changes in ΔG_{bind}^0 can have dramatic effects on the equilibrium binding constant K_{bind} . For example, the K_{bind} values for the fully charged “CerU+32 Clockwise” interface is a factor of $\sim 10^{12}$ larger than that of the minimally mutated protein interface, which suggests that long-range electrostatic interactions significantly increase the probability of forming an interface and promoting assembly. Overall, these results suggest that identification of specific protein–protein contacts from high-resolution protein structures alone is insufficient to fully understand the interactions between supercharged proteins.

Net Charge on Supercharged Proteins Predicts Protein–Protein Interactions. We next studied the role of net charge and surface charge distribution on the assembly of supercharged proteins. Here, we designed and expressed a series of positively and negatively charged GFP variants (Figure 1). Differences in surface charge distributions are visualized from electrostatic surface potential representations (Figure 1a).⁵⁷ At pH 7.4, the negatively charged GFP variants had nominal net charges of -11 , -17 (two variants: GFP-17 and GFP-17b), and -32 , and the positively charged CerU variants had nominal net charges of $+15$ (two variants: CerU+15r and CerU+15a) and $+32$ (two variants: CerU+32o and CerU+32b). For variants with the same net charge, different sets of mutations were used to generate different surface charge distributions. By varying the solution pH, the net charges on the proteins (and surface charge distributions) were varied to

produce more negatively charged (lower pH) or positively charged (higher pH) monomeric proteins (Figure 1b).^{57,58}

Figure 4a shows FRET efficiency when CerU+32o is mixed in equimolar amounts with each of the negatively charged GFP

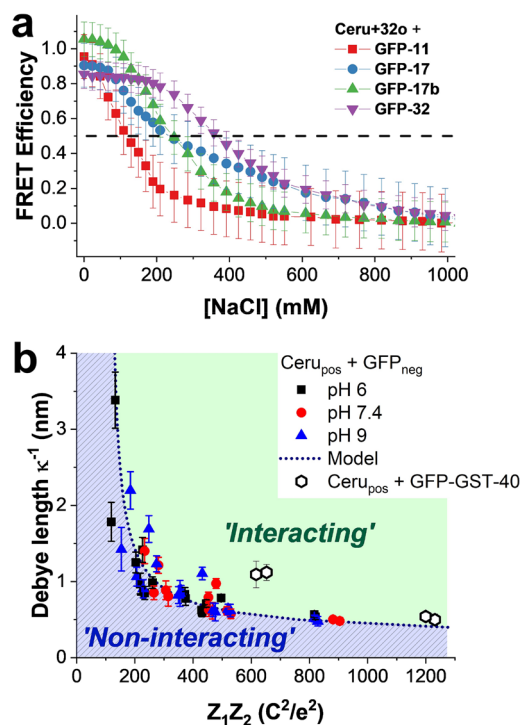


Figure 4. Assembly of supercharged proteins with different net charges. (a) Measured FRET efficiency between CerU+32o/GFP-11 (red square), GFP-17 (blue circle), GFP-17b (green triangle), and GFP-32 (purple triangle) at different salt concentrations. The dashed horizontal black line represents the critical FRET efficiency where proteins are either interacting (FRET > 0.5) or noninteracting (FRET < 0.5). (b) The critical Debye screening length κ^{-1} (calculated from the critical NaCl concentration using from eq 1) vs Coulombic attraction ($Z_1 Z_2$) for all supercharged GFP combinations at pH 6 (black squares), pH 7.4 (red circles), and pH 9 (blue triangles). The dashed line represents the fit to these data using eq 2. Critical Debye screening lengths for interactions between positive CerU variants and GFP-GST-40 (open hexagons) are not included in this fit. The model represents a phase diagram: for Debye screening lengths κ^{-1} greater than or less than that predicted, oppositely supercharged proteins are expected to be macroscopically “interacting” (green shaded area) or “noninteracting” (blue shaded area), respectively.

variants. For all mixtures, the FRET efficiency decreases with increasing NaCl concentration, indicating that the average distance between CerU+32o and the negatively charged GFP variants increases with increasing salt concentration. Disassembly of charged protein structures has been previously observed with increasing ionic strength and arises from screened electrostatic interactions.^{26–29,38,43,59,60} At high salt concentrations, electrostatic interactions are more effectively screened (i.e., shorter Debye screening length), and proteins are expected to exist in monomeric, noninteracting states. For more highly negatively charged protein variants, larger ionic strengths are required to induce disassembly. For example, the assembly between CerU+32o and GFP-11 had a FRET efficiency = 0.5 at 110 mM NaCl, whereas the assembly between CerU+32o and GFP-32 had a FRET efficiency = 0.5 at 350 mM NaCl. The FRET efficiencies between CerU+15r,

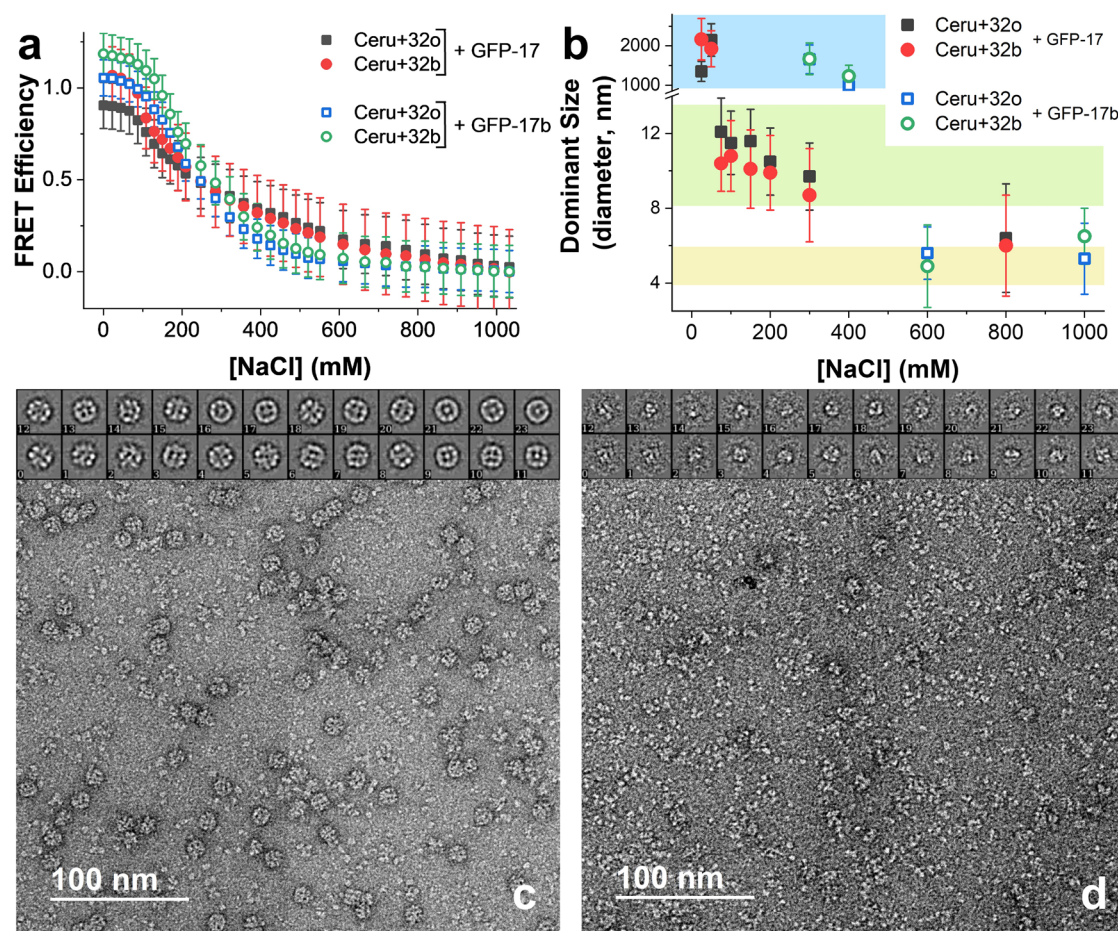


Figure 5. Effect of protein surface charge distribution on assembly. (a) FRET efficiencies between Ceru+32o/GFP-17 (black filled square), Ceru+32b/GFP-17 (red filled circle), Ceru+32o/GFP-17b (blue open square), Ceru+32b/GFP-17b (green open circle) at different concentrations of NaCl. (b) Dominant assembly size as measured by DLS for each protein combination at different concentrations of NaCl. (c) Representative TEM image of protomers from Ceru+32o/GFP-17 at 75 mM NaCl. The 24 class averages ($n \approx 2000$ particles) are shown at the top and show a highly homogeneous structure. (d) Representative TEM image of protomers from Ceru+32b/GFP-17 at 75 mM NaCl. The 24 class averages ($n \approx 800$ particles) are shown at the top and show a heterogeneous structure.

Ceru+15a, and Ceru+32b and each of the negatively charged GFP variants at different NaCl concentrations and pH values are shown in Figures S6–S8, and the effect of pH on the assembled structures for Ceru+32o/GFP-17 is shown in Figures S9 and S10.

Our results generally show that more highly charged proteins robustly assemble into supramolecular structures at higher ionic strengths, which is attributed to the larger Coulombic attraction (Z_1Z_2) between oppositely charged proteins. To understand the role of net protein charge on assembly, we determined the critical salt concentration at which FRET efficiency = 0.5 (Table S5). The critical salt concentration was used to determine a corresponding Debye screening length κ^{-1} :

$$\kappa^{-1} = \sqrt{\frac{\epsilon_0 \epsilon_w k_b T}{e^2 \sum_i c_i z_i^2}} \quad (1)$$

where ϵ_0 is vacuum permittivity, ϵ_w is the dielectric constant of water, k_b is Boltzmann's constant, T is absolute temperature, e is the elementary charge, and c_i and z_i are the number densities and valencies of the electrolyte ions (here NaCl). As shown in Figure 4b, the critical Debye screening length $\kappa_{\text{crit}}^{-1}$ is inversely related to the product of net charge on each protein (Z_1Z_2).

Upon increasing Coulombic attraction between oppositely charged proteins, a higher salt concentration (and shorter Debye screening length) is required to disrupt assembly. A similar trend is also observed at pH 6 and 9 for slightly different net charges (Figure 1b). Similar values of critical ionic strength were determined using different monovalent and divalent salts (KCl, KNO₃, CaCl₂, and MgSO₄) (Figure S11), suggesting the observed trend is generalizable to different solution compositions.

We further used a simple electrostatic model to understand the observed inverse relationship between Coulombic attraction Z_1Z_2 and critical Debye screening length $\kappa_{\text{crit}}^{-1}$. The electrical double-layer force F_{el} between two small spherical charged ions with diameter σ separated by distance r in solution is given by⁶¹

$$F_{\text{el}} = \frac{+Z_1Z_2e^2}{4\pi\epsilon_0\epsilon_w r^2} \frac{1 + \kappa r}{1 + \kappa\sigma} e^{-\kappa(r-\sigma)} \quad (2)$$

At a FRET efficiency = 0.5, the average positively charged Ceru variants are separated from negatively charged GFP variants by the Förster radius $R_0 = 5.5$ nm. Here, we assume that a constant electrical double-layer force between oppositely charged proteins is required to maintain this separation

distance. This model yields the dashed line in Figure 4b and shows an inverse relationship between Coulombic attraction Z_1Z_2 and critical Debye screening length $\kappa_{\text{crit}}^{-1}$. The data were best fit with an ion diameter $\sigma = 4.4 \pm 0.1$ nm and an electric double-layer force $F_{\text{el}} = 10.8 \pm 0.8$ pN (eq 2). Interestingly, the average center-of-mass separation between GFP-17 and Ceru+32o in the 16-mer was reported as 3.5–4.1 nm from the high-resolution cryo-EM structure,³⁸ which is consistent with the electrostatic model. Moreover, the average electrical double-layer force between oppositely charged proteins with ≈ 1 nm separation is on the same order of magnitude as F_{el} measured between charged colloidal particles.⁶² We note that this simple model assumes an isotropic charge distribution around a spherical ion, though we expect that anisotropic charge distributions contribute to differences in $\kappa_{\text{crit}}^{-1}$ observed in the experimental data.

To assess the generality of the electrical double-layer relationship, FRET efficiencies were determined for mixtures of the positively charged Ceru variants and a fusion protein consisting of negatively supercharged glutathione S-transferase (GST-40) fused with neutrally charged GFP (named GFP-GST-40, Figure S12). GST-40 is a supercharged dimeric protein with a nominal net charge of -39 at pH 7.4.³⁶ The critical Debye screening lengths for GFP-GST-40 and each positively charged Ceru variant are shown in Figure 4b, and the critical NaCl concentrations are provided in Table S5. Although an inverse relationship between Debye length and Coulombic attraction is observed, the critical Debye lengths for the fusion proteins are slightly larger than those predicted by the simple electrostatic model, which suggests that electrostatic attractions are more readily screened, possibly arising from electrostatic shielding due to the neutrally charged GFP.

We further used the electrical double-layer analysis to develop an approximate phase diagram for supercharged assembly (Figure 4b). If the Debye screening length of the solution is longer than that predicted by $\kappa_{\text{crit}}^{-1}$, then oppositely supercharged proteins are expected to exist in an interacting state (defined when the average separation distance is less than the Förster distance R_0 for a characteristic protein pair), and assembly is expected. Conversely, if the Debye length is shorter than $\kappa_{\text{crit}}^{-1}$, then supercharged proteins are expected to exist in noninteracting states. Overall, net charge provides a simple metric to predict whether supercharged GFP variants will interact and assemble in solution, but the precise phase diagram will change depending on protein concentrations or ratio of positively to negatively charged protein.²⁶

Assembled Structures Depend on Precise Interprotein Interactions. We further explored how changing surface charge distribution (at a constant net charge) affects the assembly of proteins. Figure 5a shows the FRET efficiency between negatively charged GFP variants (GFP-17 or GFP-17b) and positively charged Ceru variants (Ceru+32o or Ceru+32b) at different salt concentrations at pH 7.4. Because the net charge and Coulombic attraction are similar for these protein pairs, the FRET efficiency traces are similar, and nearly identical critical salt concentrations are determined. However, bulk average FRET measurements alone cannot provide additional detailed information on the assembly process for supercharged proteins.

To understand the role of surface charge distributions, we studied supramolecular assembly of charged protein variants with different surface charge distributions using DLS, confocal microscopy, and negative stain TEM. Figure 5b shows the

dominant assembly sizes determined from DLS. At low salt concentration, all of the oppositely charged protein pair combinations assemble into large (>1000 nm) aggregates with fractal-like structures (Figure S13).³⁸ Upon increasing salt concentration, the average diameter of the dominant assembled structures between GFP-17 and either Ceru+32o or Ceru+32b was 10–12 nm, which is consistent with the aggregate size previously observed between Ceru+32o and GFP-17, suggesting a stable protomeric structure.³⁸ On the other hand, the dominant assembled structures between GFP-17b and either Ceru+32o or Ceru+32b consisted of large aggregates (>1000 nm) for NaCl concentrations <400 mM. Interestingly, no intermediate-sized protomeric structures were observed for protein pairs involving GFP-17b, which only differs from GFP-17 in the distribution of surface charges. At high NaCl concentrations (>600 mM), the dominant structure in solution from DLS matched the monomeric protein size for all protein combinations, indicating electrostatic-driven assembly was effectively screened.

Negative-stain TEM was used to image protein combinations showing evidence of protomer structure formation. Ceru+32o and Ceru+32b formed a 10–12 nm assembled structures with GFP-17 at 50 mM NaCl. As expected, Ceru+32o and GFP-17 formed distinct, homogeneous supramolecular ring-like assemblies, and class averages of selected particles ($n \approx 2000$) were consistent with prior reports (Figure 5c).³⁸ However, assembled structures between Ceru+32b and GFP-17 were heterogeneous, and class averages of selected particles ($n \approx 800$) did not yield evidence of an ordered structure (Figure 5d). Overall, these results highlight the role of surface charge distribution on facilitating and stabilizing specific intermolecular interactions promoting ordered assembly.

Stability of Assembled Structures Depends on Size and Specific Interprotein Interactions. We next assessed the stability of supramolecular complexes by performing a series of dynamic subunit exchange measurements. First, the concentration dependence of supramolecular assemblies was assessed using dilution experiments to qualitatively assess the kinetic stability protein assemblies. In these experiments, the assembled protein structures were diluted with buffered solution, and equilibrium FRET ratios were measured after ~ 15 min (Figure S14). We note that these dilution experiments were only used to identify potential candidates for dynamic subunit exchange experiments, as described below. If proteins rapidly exchange with nearby protein partners in solution (which has been reported for electrostatically driven complex coacervate formation),^{63,64} then the dominant assembled structure is expected to depend on solution composition and protein concentration. For assemblies between proteins with relatively small net charges (e.g., those involving Ceru+15 variants), our results show that the FRET ratio decreased upon dilution, suggesting rapid dynamic subunit exchange in solution. However, for supramolecular assemblies consisting of proteins with larger net charges, the FRET ratio did not significantly change upon dilution, suggesting slow dynamic exchange in solution.

To understand the kinetic stability of assembled protein structures that are relatively insensitive to dilution, we performed a series of dynamic subunit exchange measurements using nonfluorescent analogues of the supercharged fluorescent proteins. Nonfluorescent analogues of the negatively charged GFP variants (containing a Gly-Gly-Gly construct in place of the chromophore)⁶⁵ were designed and expressed. As

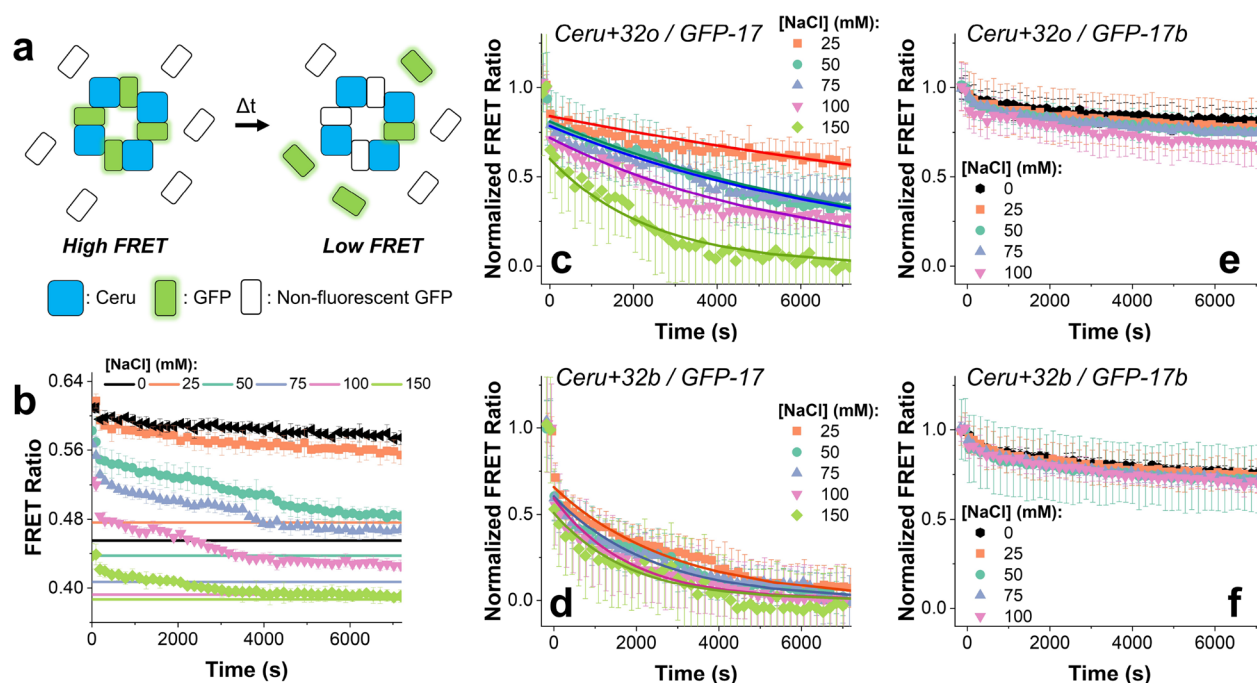


Figure 6. Dynamic subunit exchange experiments. (a) Schematic of dynamic subunit exchange. As nonfluorescent GFP analogue is incorporated into assembled structures, the FRET ratio decreases. (b) Transient FRET ratio between Ceru+32o/GFP-17 after the nonfluorescent negatively charged analogue (GFP-17nf) is added to solution at different concentrations of NaCl. The solid lines represent the final FRET ratio with all GFP-17nf incorporated at each salt concentration. (c–f) Normalized transient FRET ratios using the initial (before nonfluorescent negatively charged analogs were added) and final FRET ratios at different salt concentrations for (c) Ceru+32o/GFP-17, (d) Ceru+32b/GFP-17, (e) Ceru+32o/GFP-17b, and (f) Ceru+32b/GFP-17b. Kinetic traces from modeled dynamic subunit exchange of protomers from Ceru+32o/GFP-17 and Ceru+32b/GFP-17 are shown as solid lines.

expected, the negatively charged GFP variants and their corresponding nonfluorescent analogues (designated as “nf”) were observed to assemble with positively charged Ceru variants with identical average sizes and shapes as determined by DLS and TEM imaging (Figure S15). Dynamic exchange experiments were then performed by adding nonfluorescent proteins to a solution containing preassembled structures from oppositely supercharged fluorescent proteins, and changes in FRET were measured as the nonfluorescent analogues dynamically exchanged with GFP in the supramolecular assembled structures (Figure 6a). The time-dependent FRET ratio for GFP-17nf exchange with Ceru+32o/GFP-17 is shown in Figure 6b for a series of different NaCl concentrations. In all cases, the transient FRET ratio was found to decrease as a function of time when GFP-17nf is added to solution (at $t = 0$ min). The FRET ratio was not observed to change over the time scale of the measurement when buffer was added instead of GFP-17nf (Figure S16), suggesting that exchange of nonfluorescent GFP-17nf into the assembled structure (rather than photobleaching) is responsible for observed changes in fluorescence. In addition, the final FRET ratio at each salt concentration was measured by premixing GFP-17 and GFP-17nf before triggering assembly with Ceru+32o (solid lines, Figure 6b). The transient FRET ratio approaches the final FRET ratio more rapidly at higher salt concentrations, indicating faster uptake of GFP-17nf and a more dynamic assembled structure. A series of dynamic exchange measurements was also performed with Ceru+32b/GFP-17, Ceru+32o/GFP-17b, and Ceru+32b/GFP-17b as a function of ionic strength (Figure S17).

Transient normalized FRET ratios after addition of nonfluorescent negatively charged GFP analogues are shown

in Figure 6c–f. The normalized FRET ratio for protomer-forming assemblies (i.e., assembly between GFP-17 and either Ceru+32o or Ceru+32b) decreases significantly after addition of GFP-17nf, and the rate of change varies with salt concentration (Figure 6c,d). In some cases, there was an initial rapid decrease in the transient FRET ratio, which was observed even for samples where large aggregates were removed by centrifugation (Figure S18). Based on these results, we hypothesize that the initial rapid decrease in FRET ratio arises from displacement of weakly associating proteins from the protomers, and the slower decrease in FRET ratio arises from dynamic subunit exchange within the protomer. On the other hand, large aggregate forming assemblies (i.e., assembly between GFP-17b and either Ceru+32o or Ceru+32b) show much slower change in normalized FRET ratio at all salt concentrations measured (Figure 6e,f). The difference in behavior between these two systems likely arises due to structural differences that affect the ability of proteins to exchange with nearby partners in the surrounding solution.

Kinetic Model for Dynamic Exchange. To quantitatively describe the kinetic stability experiments, we developed a simple kinetic model to understand the dynamic exchange of protein subunits for complexes between Ceru+32o/GFP-17 and Ceru+32b/GFP-17 (Table S6).^{66–68} In this model, monomers dissociate from the protomer sequentially according to a unimolecular rate constant k_d , which is assumed to be constant for all proteins in the protomer. The probability of a protomer dissociating to yield a free fluorescent monomer of GFP-17 (G) or free nonfluorescent analogue (A) is determined by the composition of the assembled protein structure. In terms of association, monomeric GFP-17 (G) or nonfluorescent analogue (A) proteins form protomers with a

bimolecular rate constant k_a . Because of the strong binding energy between oppositely charged proteins, association reactions are assumed to be fast, and dissociation reactions are rate limiting (Figure S19). Similar exchange kinetics were observed when different concentrations of nonfluorescent analogues are added to solution, which supports the assumption that dissociation is rate limiting (Figure S20). A detailed description of the kinetic model is provided in Supporting Information.

Results from the kinetic model are shown as solid lines in Figure 6c,d. The incorporation of nonfluorescent analogue (A) into the protomers is simulated by numerically solving the coupled differential rate equations while matching experimental conditions ($[G_8A_0]_{\text{initial}} = 0.05 \mu\text{M}$, $[G]_{\text{initial}} = 0 \mu\text{M}$, $[A]_{\text{initial}} = 4 \mu\text{M}$).⁶⁹ Here, k_d was varied to directly compare the simulated fraction of GFP-17 remaining in the protomer to the experimentally determined normalized FRET ratio. For protomer assembly between Ceru+32o and GFP-17 (Figure 6c), k_d increases as the salt concentration increases (Figure S21), ranging from 0.0004 s^{-1} to 0.003 s^{-1} at 25 mM and 150 mM NaCl, respectively. These values of k_d suggest kinetic lifetimes ranging from 7 to 40 min depending on the solution composition. For protomer assembly between Ceru+32b and GFP-17 (Figure 6d), exchange rates are less sensitive to salt concentration; k_d ranges from 0.0025 to 0.004 s^{-1} at 25 and 150 mM NaCl, respectively. Overall, the protomers formed from Ceru+32b/GFP-17 are observed to be less stable compared to protomers formed from Ceru+32o/GFP-17, which could arise from the heterogeneous structures of Ceru+32b/GFP-17 assemblies (Figure Sd) and lack of cooperative interactions between subunits. In addition, faster exchange kinetics could arise from protomers with fewer total subunits (Figure S22).

In general, the kinetic lifetimes for dynamic subunit exchange measurements of protomers are consistent with those observed in endogenous biological protein complexes,⁷⁰ which range from seconds (e.g., RNA polymerase)⁷¹ to days (e.g., nuclear pore complex).⁷² From this view, assembly between oppositely supercharged proteins provides a mechanism to drive long-lasting association between two otherwise noninteracting proteins. However, the stability of the assembled structure strongly depends on the precise intermolecular interactions and solution conditions. In addition, assembly between oppositely supercharged proteins can be used to generate extremely stable biomaterials with transport-limited exchange leading to kinetic lifetimes >1 day (i.e., large aggregate forming assemblies).

CONCLUSIONS

In this work, we study the assembly process of oppositely supercharged proteins using a combination of experiments and computational modeling. Our results highlight the role of surface charge distributions on the supramolecular assembly process for oppositely supercharged proteins. Using minimally mutated variants of supercharged proteins (GFP-min and Ceru-min) that retain only the mutations ostensibly involved in stabilizing the assemblies, we systematically explored the role noninteracting charges play on assembly. Minimally charged variants were not observed to assemble into ordered hierarchical structures. Results from MD simulations show that the values of ΔG^0_{bind} are relatively large for both the minimally charged and fully charged variants compared to naturally occurring protein–protein interfaces.⁵⁴ Unexpectedly, ΔG^0_{bind}

for the minimally charged protein variants are significantly smaller ($\sim 30\%$) than those of the fully charged variants even though the interfaces retain all specific local attractions (e.g., H-bonding and salt bridge interactions). Such differences in binding energy are attributed to long-range electrostatic interactions between the fully charged variants. From this view, specific interprotein interactions contribute to the bulk of the binding energy between oppositely supercharged proteins, but noninteracting excess charges are required to stabilize the assembled protein complexes.

Furthermore, our results suggest that identification of specific interprotein interactions from high-resolution structures alone is insufficient to fully describe the interactions and assembly process between oppositely supercharged proteins. Our results show that assembly of supercharged proteins is governed by both these local attractions as well as long-range electrostatic interactions. Therefore, these findings will be useful for informing the rational design of new hierarchical structures based on predictions of complementary surfaces.^{73–77} In particular, rational design methods for supercharged proteins should take into consideration long-range electrostatic interactions away from the interface for stabilizing protein complexes, which are known to play a role in kinetic assembly and disassembly processes involving supramolecular protein complexes.

Our results elucidate how net charge and average Coulombic attraction can predict whether oppositely supercharged proteins will interact as a function of ionic strength conditions. We found that the supramolecular structure strongly depends on the precise distribution of charges on the proteins surface. Prior work has shown that different assembled structures can be obtained from a common parent protein building block by changing the supramolecular assembly strategy (e.g., ferritin cage proteins).^{20,22,23,30,42} Importantly, our results show that proteins with similar net charges but different surface charge distributions give rise to different assembled structures. Surprisingly, the kinetic stabilities of these different supramolecular structures were found to depend on surface charge distribution, suggesting that assembled structures with different functional properties can be formed from a common building block. Overall, our results show that supercharging proteins provides an efficient method to create stable, long-lived protein assemblies while further highlighting the role of surface charge distribution and precise stabilizing interprotein interactions on the structure of assembled supercharged proteins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscentsci.2c00730>.

Description of materials, experimental methods, and computational methods including amino acid sequences, denaturing protein gels, and UV–vis and fluorescence spectra; Depiction of simulated protein–protein interfaces; PMFs of orientational, positional, and conformation constraints for bound and unbound states; Plots of FRET efficiency versus NaCl for protein assemblies at pH 6, 7.4, and 9; Plot of soluble fraction of GFP-17/Ceru+32o assemblies versus NaCl concentration; FRET efficiency versus salt concentration for protein assemblies for different salt identities; TEM images of GFP-17/Ceru+32o at different pH values and at different

ratios; Confocal microscopy images of protein assembly; Plot of FRET ratio versus protein concentration; Comparison of assemblies from GFP-17 and non-fluorescent analogue; Plots of transient FRET ratios after buffer addition, after nonfluorescent analogue addition, and with/without large aggregates; Plot of final FRET ratios versus nonfluorescent analogue concentrations; Simulated exchange kinetics with different association rate constants k_a and different protomer sizes; Plot of measured and simulated exchange kinetics for different concentrations of nonfluorescent analogue; Plot of modeled dissociation rate constants k_d ; Tables containing contributing terms to ΔG_{bind}^0 for each interface; Table with critical salt concentrations and protein charges for different assemblies; Table with dynamic subunit exchange kinetic model; Table with relevant MD simulation parameters (PDF)

Transparent Peer Review report available (PDF)

AUTHOR INFORMATION

Corresponding Author

Charles M. Schroeder – Beckman Institute for Advanced Science and Technology, Department of Chemical and Biomolecular Engineering, and Department of Materials Science and Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, United States; orcid.org/0000-0001-6023-2274; Email: cms@illinois.edu

Authors

Michael I. Jacobs – Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, United States

Prateek Bansal – Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, United States; orcid.org/0000-0002-1887-2320

Diwakar Shukla – Beckman Institute for Advanced Science and Technology and Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, United States; orcid.org/0000-0003-4079-5381

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acscentsci.2c00730>

Notes

The authors declare no competing financial interest.

¹Department of Chemistry and Biochemistry, Texas State University, San Marcos, Texas 78666, United States

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0022035. M.I.J. acknowledges a Beckman Institute Postdoctoral Fellowship and a Mistletoe Fellowship. D.S. and P.D.B. acknowledge support from NSF MCB-1845606. We thank Prof. Andrew Ellington for providing plasmids encoding for GFP-11, GFP-17, GFP-32, and Ceru+32o, and we thank Prof. Catherine Murphy for use of the Malvern Zetasizer Nano ZS. P.D.B. thanks Balaji Selvam, Soumajit Dutta, and Jiming Chen for useful discussions and The Blue Waters Petascale Computing Facility and National Center for Supercomputing

Applications (NCSA), which are supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois.

REFERENCES

- (1) Ahnert, S. E.; Marsh, J. A.; Hernández, H.; Robinson, C. V.; Teichmann, S. A. Principles of Assembly Reveal a Periodic Table of Protein Complexes. *Science* (80-) **2015**, *350*, aaa2245.
- (2) Garcia-Seisdedos, H.; Empereur-Mot, C.; Elad, N.; Levy, E. D. Proteins Evolve on the Edge of Supramolecular Self-Assembly. *Nature* **2017**, *548* (7666), 244–247.
- (3) Goodsell, D. S.; Olson, A. J. Structural Symmetry and Protein Function. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105–153.
- (4) Levy, E. D.; Erba, E. B.; Robinson, C. V.; Teichmann, S. A. Assembly Reflects Evolution of Protein Complexes. *Nature* **2008**, *453* (June), 1262–1265.
- (5) Chiesa, G.; Kiriakov, S.; Khalil, A. S. Protein Assembly Systems in Natural and Synthetic Biology. *BMC Biol.* **2020**, *18* (1), 1–18.
- (6) Liu, X.; Fu, T.; Ward, J.; Gao, H.; Yin, B.; Woodard, T.; Lovley, D. R.; Yao, J. Multifunctional Protein Nanowire Humidity Sensors for Green Wearable Electronics. *Adv. Electron. Mater.* **2020**, *6* (2000721), 1–7.
- (7) Fu, T.; Liu, X.; Gao, H.; Ward, J. E.; Liu, X.; Yin, B.; Wang, Z.; Zhuo, Y.; Walker, D. J. F.; Yang, J. J.; et al. Bioinspired Bio-Voltage Memristors. *Nat. Commun.* **2020**, *11* (1861), 1–10.
- (8) Wang, M.; Zuris, J. A.; Meng, F.; Rees, H.; Sun, S.; Deng, P.; Han, Y.; Gao, X.; Pouli, D. Efficient Delivery of Genome-Editing Proteins Using Bioreducible Lipid Nanoparticles. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (11), 2868–2873.
- (9) Stevens, C. A.; Kaur, K.; Klok, H. A. Self-Assembly of Protein-Polymer Conjugates for Drug Delivery. *Adv. Drug Delivery Rev.* **2021**, *174*, 447–460.
- (10) Zhu, J.; Avakyan, N.; Kakkis, A.; Ho, A. M.; Han, K.; Li, Y.; Zhang, Z.; Choi, T. S.; Na, Y.; Yu, C.; et al. Protein Assembly by Design. *Chem. Rev.* **2021**, *121*, 13701–13796.
- (11) Bai, Y.; Luo, Q.; Liu, J. Protein Self-Assembly via Supramolecular Strategies. *Chem. Soc. Rev.* **2016**, *45* (10), 2756–2767.
- (12) Hamley, I. W. Protein Assemblies: Nature-Inspired and Designed Nanostructures. *Biomacromolecules* **2019**, *20* (5), 1829–1848.
- (13) Zhang, W.; Luo, Q.; Miao, L.; Hou, C.; Bai, Y.; Dong, Z.; Xu, J.; Liu, J. Self-Assembly of Glutathione S-Transferase into Nanowires. *Nanoscale* **2012**, *4* (19), 5847–5851.
- (14) Yang, G.; Ding, H. M.; Kochovski, Z.; Hu, R.; Lu, Y.; Ma, Y. Q.; Chen, G.; Jiang, M. Highly Ordered Self-Assembly of Native Proteins into 1D, 2D, and 3D Structures Modulated by the Tether Length of Assembly-Inducing Ligands. *Angew. Chemie - Int. Ed.* **2017**, *56* (36), 10691–10695.
- (15) Oohora, K.; Burazerovic, S.; Onoda, A.; Wilson, Y. M.; Ward, T. R.; Hayashi, T. Chemically Programmed Supramolecular Assembly of Hemoprotein and Streptavidin with Alternating Alignment. *Angew. Chemie - Int. Ed.* **2012**, *51* (16), 3818–3821.
- (16) Sendai, T.; Biswas, S.; Aida, T. Photoreconfigurable Supramolecular Nanotube. *J. Am. Chem. Soc.* **2013**, *135* (31), 11509–11512.
- (17) Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y.; Dai, B.; Sakai, F.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X.; et al. Precise and Reversible Protein-Microtubule-Like Structure with Helicity Driven by Dual Supramolecular Interactions. *J. Am. Chem. Soc.* **2016**, *138* (6), 1932–1937.
- (18) Carlson, J. C. T.; Jena, S. S.; Flenniken, M.; Chou, T. F.; Siegel, R. A.; Wagner, C. R. Chemically Controlled Self-Assembly of Protein Nanorings. *J. Am. Chem. Soc.* **2006**, *128* (23), 7630–7638.
- (19) Bai, Y.; Luo, Q.; Zhang, W.; Miao, L.; Xu, J.; Li, H.; Liu, J. Highly Ordered Protein Nanorings Designed by Accurate Control of Glutathione S-Transferase Self-Assembly. *J. Am. Chem. Soc.* **2013**, *135* (30), 10966–10969.

- (20) Zhou, K.; Zang, J.; Chen, H.; Wang, W.; Wang, H.; Zhao, G. On-Axis Alignment of Protein Nanocage Assemblies from 2D to 3D through the Aromatic Stacking Interactions of Amino Acid Residues. *ACS Nano* **2018**, *12* (11), 11323–11332.
- (21) Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. Metal-Directed, Chemically Tunable Assembly of One-, Two- and Three-Dimensional Crystalline Protein Arrays. *Nat. Chem.* **2012**, *4* (5), 375–382.
- (22) Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A. A Metal Organic Framework with Spherical Protein Nodes: Rational Chemical Design of 3D Protein Crystals. *J. Am. Chem. Soc.* **2015**, *137* (36), 11598–11601.
- (23) Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A. Synthetic Modularity of Protein-Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2017**, *139* (24), 8160–8166.
- (24) Brodin, J. D.; Carr, J. R.; Sontz, P. A.; Tezcan, F. A. Exceptionally Stable, Redox-Active Supramolecular Protein Assemblies with Emergent Properties. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (8), 2897–2902.
- (25) Ma, C.; Malessa, A.; Boersma, A. J.; Liu, K.; Herrmann, A. Supercharged Proteins and Polypeptides. *Adv. Mater.* **2020**, *32*, 1905309.
- (26) Kostiaainen, M. A.; Hiekkataipale, P.; Laiho, A.; Lemieux, V.; Seitsonen, J.; Ruokolainen, J.; Ceci, P. Electrostatic Assembly of Binary Nanoparticle Superlattices Using Protein Cages. *Nat. Nanotechnol.* **2013**, *8* (1), 52–56.
- (27) Liljeström, V.; Mikkilä, J.; Kostiaainen, M. A. Self-Assembly and Modular Functionalization of Three-Dimensional Crystals from Oppositely Charged Proteins. *Nat. Commun.* **2014**, *5*, 4445.
- (28) Liljeström, V.; Seitsonen, J.; Kostiaainen, M. A. Electrostatic Self-Assembly of Soft Matter Nanoparticle Cocrystals with Tunable Lattice Parameters. *ACS Nano* **2015**, *9* (11), 11278–11285.
- (29) Korpi, A.; Ma, C.; Liu, K.; Nonappa; Herrmann, A.; Ikkala, O.; Kostiaainen, M. A. Self-Assembly of Electrostatic Cocrystals from Supercharged Fusion Peptides and Protein Cages. *ACS Macro Lett.* **2018**, *7* (3), 318–323.
- (30) Künzle, M.; Eckert, T.; Beck, T. Binary Protein Crystals for the Assembly of Inorganic Nanoparticle Superlattices. *J. Am. Chem. Soc.* **2016**, *138* (39), 12731–12734.
- (31) Beck, T.; Tetter, S.; Künzle, M.; Hilvert, D. Construction of Matryoshka-Type Structures from Supercharged Protein Nanocages. *Angew. Chemie - Int. Ed.* **2015**, *54* (3), 937–940.
- (32) Seebeck, F. P.; Woycechowsky, K. J.; Zhuang, W.; Rabe, J. P.; Hilvert, D. A Simple Tagging System for Protein Encapsulation. *J. Am. Chem. Soc.* **2006**, *128* (14), 4516–4517.
- (33) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D. Directed Evolution of a Protein Container. *Science (80-)*. **2011**, *331* (6017), 589–592.
- (34) Azuma, Y.; Zschoche, R.; Tinzl, M.; Hilvert, D. Quantitative Packaging of Active Enzymes into a Protein Cage. *Angew. Chemie - Int. Ed.* **2016**, *55* (4), 1531–1534.
- (35) Glasgow, J. E.; Capehart, S. L.; Francis, M. B.; Tullman-Ercek, D. Osmolyte-Mediated Encapsulation of Proteins inside MS2 Viral Capsids. *ACS Nano* **2012**, *6* (10), 8658–8664.
- (36) Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *J. Am. Chem. Soc.* **2007**, *129* (33), 10110–10112.
- (37) Der, B. S.; Kluwe, C.; Miklos, A. E.; Jacak, R.; Lyskov, S.; Gray, J. J.; Georgiou, G.; Ellington, A. D.; Kuhlman, B. Alternative Computational Protocols for Supercharging Protein Surfaces for Reversible Unfolding and Retention of Stability. *PLoS One* **2013**, *8* (5), e64363.
- (38) Simon, A. J.; Zhou, Y.; Ramasubramani, V.; Glaser, J.; Pothukuchy, A.; Gollihar, J.; Gerberich, J. C.; Leggere, J. C.; Morrow, B. R.; Jung, C.; et al. Supercharging Enables Organized Assembly of Synthetic Biomolecules. *Nat. Chem.* **2019**, *11* (3), 204–212.
- (39) Damasceno, P. F.; Engel, M.; Glotzer, S. C. Predictive Self-Assembly of Polyhedra into Complex Structures. *Science (80-)* **2012**, *337* (6093), 453–457.
- (40) Zhang, Z.; Glotzer, S. C. Self-Assembly of Patchy Particles. *Nano Lett.* **2004**, *4* (8), 1407–1413.
- (41) Pawar, A. B.; Kretzschmar, I. Fabrication, Assembly, and Application of Patchy Particles. *Macromol. Rapid Commun.* **2010**, *31* (2), 150–168.
- (42) Künzle, M.; Eckert, T.; Beck, T. Metal-Assisted Assembly of Protein Containers Loaded with Inorganic Nanoparticles. *Inorg. Chem.* **2018**, *57* (21), 13431–13436.
- (43) Kapelner, R. A.; Obermeyer, A. C. Ionic Polypeptide Tags for Protein Phase Separation. *Chem. Sci.* **2019**, *10* (9), 2700–2707.
- (44) Zervoudis, N. A.; Obermeyer, A. C. The Effects of Protein Charge Patterning on Complex Coacervation. *Soft Matter* **2021**, *17* (27), 6637–6645.
- (45) Rizzo, M. A.; Springer, G. H.; Granada, B.; Piston, D. W. An Improved Cyan Fluorescent Protein Variant Useful for FRET. *Nat. Biotechnol.* **2004**, *22* (4), 445–449.
- (46) Pedelacq, J.-D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S. Engineering and Characterization of a Superfolder Green Fluorescent Protein. *Nat. Biotechnol.* **2006**, *24* (1), 79–89.
- (47) Wu, P.; Brand, L. Resonance Energy Transfer: Methods and Applications. *Anal. Biochem.* **1994**, *218*, 1–13.
- (48) Zhao, C.; Shukla, D. Structural Basis for Negative Regulation of ABA Signaling by ROP11. *bioRxiv*, May 22, 2020. DOI: 10.1101/2020.05.20.107185.
- (49) Zhao, C.; Shukla, D. Molecular Basis of the Activation and Dissociation of Dimeric PYL2 Receptor in Abscisic Acid Signaling. *Phys. Chem. Chem. Phys.* **2022**, *24* (2), 724–734.
- (50) Woo, H. J.; Roux, B. Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6825–6830.
- (51) Gumbart, J. C.; Roux, B.; Chipot, C. Efficient Determination of Protein-Protein Standard Binding Free Energies from First Principles. *J. Chem. Theory Comput.* **2013**, *9* (8), 3789–3798.
- (52) Moffett, A. S.; Shukla, D. How Do Brassinosteroids Activate Their Receptors? *bioRxiv*, May 7, 2019. DOI: 10.1101/630640.
- (53) Plattner, N.; Doerr, S.; De Fabritius, G.; Noé, F. Complete Protein-Protein Association Kinetics in Atomic Detail Revealed by Molecular Dynamics Simulations and Markov Modelling. *Nat. Chem.* **2017**, *9* (10), 1005–1011.
- (54) Kastriitis, P. L.; Moal, I. H.; Hwang, H.; Weng, Z.; Bates, P. A.; Bonvin, A. M. J. J.; Janin, J. A Structure-Based Benchmark for Protein-Protein Binding Affinity. *Protein Sci.* **2011**, *20* (3), 482–491.
- (55) Li, X.; Keskin, O.; Ma, B.; Nussinov, R.; Liang, J. Protein-Protein Interactions: Hot Spots and Structurally Conserved Residues Often Locate in Complemented Pockets That Pre-Organized in the Unbound States: Implications for Docking. *J. Mol. Biol.* **2004**, *344* (3), 781–795.
- (56) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot Spots—A Review of the Protein-Protein Interface Determinant Amino-Acid Residues. *Proteins* **2007**, *68* (4), 803–812.
- (57) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS Biomolecular Solvation Software Suite. *Protein Sci.* **2018**, *27*, 112–128.
- (58) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537.
- (59) Cummings, C. S.; Obermeyer, A. C. Phase Separation Behavior of Supercharged Proteins and Polyelectrolytes. *Biochemistry* **2018**, *57* (3), 314–323.
- (60) Spruijt, E.; Bakker, H. E.; Kodger, T. E.; Sprakel, J.; Cohen Stuart, M. A.; Van Der Gucht, J. Reversible Assembly of Oppositely Charged Hairy Colloids in Water. *Soft Matter* **2011**, *7* (18), 8281–8290.

- (61) Israelachvili, J. N. *Electrostatic Forces between Surfaces in Liquids* **2011**, 291.
- (62) Sinha, P.; Szilagyi, I.; Montes Ruiz-Cabello, F. J.; Maroni, P.; Borkovec, M. Attractive Forces between Charged Colloidal Particles Induced by Multivalent Ions Revealed by Confronting Aggregation and Direct Force Measurements. *J. Phys. Chem. Lett.* **2013**, *4* (4), 648–652.
- (63) Jia, T. Z.; Hentrich, C.; Szostak, J. W. Rapid RNA Exchange in Aqueous Two-Phase System and Coacervate Droplets. *Orig. Life Evol. Biosph.* **2014**, *44* (1), 1–12.
- (64) Fares, H. M.; Marras, A. E.; Keating, C. D.; Ting, J. M.; Tirrell, M. V. Impact of Wet-Dry Cycling on the Phase Behavior and Compartmentalization Properties of Complex Coacervates. *Nat. Commun.* **2020**, *11* (5423), 1–13.
- (65) Barondeau, D. P.; Putnam, C. D.; Kassmann, C. J.; Tainer, J. A.; Getzoff, E. D. Mechanism and Energetics of Green Fluorescent Protein Chromophore Synthesis Revealed by Trapped Intermediate Structures. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12111–12116.
- (66) Endres, D.; Zlotnick, A. Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers. *Biophys. J.* **2002**, *83* (2), 1217–1230.
- (67) Chen, Y.; Erickson, H. P. Rapid in Vitro Assembly Dynamics and Subunit Turnover of FtsZ Demonstrated by Fluorescence Resonance Energy Transfer. *J. Biol. Chem.* **2005**, *280* (23), 22549–22554.
- (68) Keetch, C. A.; Bromley, E. H. C.; Mccammon, M. G.; Wang, N.; Christodoulou, J.; Robinson, C. V. L55P Transthyretin Accelerates Subunit Exchange and Leads to Rapid Formation of Hybrid Tetramers. *J. Biol. Chem.* **2005**, *280* (50), 41667–41674.
- (69) Ianni, J. C. *Kintecus*; www.kintecus.com. Windows Version 6.01, 2017.
- (70) Tusk, S. E.; Delalez, N. J.; Berry, R. M. Subunit Exchange in Protein Complexes. *J. Mol. Biol.* **2018**, *430* (22), 4557–4579.
- (71) Stasevich, T. J.; McNally, J. G. Assembly of the Transcription Machinery: Ordered and Stable, Random and Dynamic, or Both? *Chromosoma* **2011**, *120*, 533–545.
- (72) Rabut, G.; Doye, V.; Ellenberg, J. Mapping the Dynamic Organization of the Nuclear Pore Complex inside Single Living Cells. *Nat. Cell Biol.* **2004**, *6* (11), 1114–1121.
- (73) Matsuura, K. Rational Design of Self-Assembled Proteins and Peptides for Nano- and Micro-Sized Architectures. *RSC Adv.* **2014**, *4* (6), 2942–2953.
- (74) McCafferty, C. L.; Marcotte, E. M.; Taylor, D. W. Simplified Geometric Representations of Protein Structures Identify Complementary Interaction Interfaces. *Proteins Struct. Funct. Bioinforma.* **2021**, *89* (3), 348–360.
- (75) Rego, N. B.; Xi, E.; Patel, A. J. Identifying Hydrophobic Protein Patches to Inform Protein Interaction Interfaces. *PNAS* **2020**, *118*, e2018234118.
- (76) King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science (80-)* **2012**, *336*, 1171–1175. Design
- (77) Boyken, S. E.; Chen, Z.; Groves, B.; Langan, R. A.; Oberdorfer, G.; Ford, A.; Gilmore, J. M.; Xu, C.; DiMaio, F.; Pereira, J. H.; et al. De Novo Design of Protein Homo-Oligomers with Modular Hydrogen-Bond Network-Mediated Specificity. *Science (80-)* **2016**, *352* (6286), 680–687.



CAS BIOFINDER DISCOVERY PLATFORM™

CAS BIOFINDER HELPS YOU FIND YOUR NEXT BREAKTHROUGH FASTER

Navigate pathways, targets, and
diseases with precision

Explore CAS BioFinder

