Electronic Supplementary Material (ESI) for Soft Matter.

Supporting Information

The effects of protein charge patterning on complex coacervation

*Nicholas A. Zervoudis, Allie C. Obermeyer**

Department of Chemical Engineering, Columbia University, New York, NY 10027

Table of Contents.

1. Materials and Methods

All enzymes for molecular biology and competent cells [NiCo21 (DE3), BL21 (DE3), NEB-5α] were purchased from New England Biolabs (Ipswich, MA). All primers were purchased from Integrated DNA Technologies (Coralville, IA). All chemicals and media components were purchased from Fisher Scientific (Pittsburgh, PA) or Sigma Aldrich (St. Louis, MO) and were used as received. Poly(4-vinyl-*N*-methylpyridinium iodide), qP4VP (M_n = 28,000, $Đ$ = 1.2), was purchased from Polymer Source (Montreal, QC). qP4VP was dissolved in 10 mM Tris to a concentration of 1 mg mL^{-1}. The cationic polymer solution was adjusted to a pH of 7.4. Both protein and polymer solutions were stored at 4 °C until use. Unless otherwise noted, all experiments were conducted using 1 mg mL^{-1} working solutions of protein and polymer in 10 mM Tris (pH 7.4).

2. Cloning

The sfGFP plasmid was a gift from the Banta lab (Columbia University). Anionic GFP mutants were cloned using sfGFP (N-terminal, 6xHis) as a template. Forward and reverse primers were designed specifically for each amino acid tag sequence using NEBuilder (https://nebuilder.neb.com). The tag sequences were appended to the 3' end (protein C-terminus) of the sfGFP gene using restriction enzyme cloning or HiFi assembly. For both methods, the insert was amplified by PCR to include the mutations at the 3' end of the gene. Briefly, template DNA, dNTPs (final conc. 200 μM), primers (final conc. 0.5 μM), Phusion DNA polymerase, and HF Buffer were combined as per the PCR protocol for Phusion High-Fidelity DNA polymerase with a total sample volume of 50 μL. PCR reactions were denatured (98 °C), annealed, and extended (72 °C). PCR was performed for a total of 35 cycles at annealing temperatures of 52 °C for restriction enzyme cloning or as recommended by NEBuilder for HiFi assembly. PCR products were purified using a Qiagen PCR purification kit. For restriction enzyme mediated cloning prior to PCR purification, the PCR reaction was treated with Dpn1 to digest the template DNA and NcoI and XhoI to digest the PCR amplified inserts. Insert DNA and vector DNA were purified via the Qiagen PCR purification kit and agarose gel electrophoresis (followed by Qiagen Gel Purification kit), respectively. Otherwise, HiFi assembly was performed following the HiFi DNA Assembly Reaction protocol using a 2:1 molar ratio of insert DNA to vector and incubating for 15 minutes at 50 °C. Ligated DNA was transformed into NEB-5α cells and the sequence of the resulting plasmid was confirmed by Sanger sequencing (Genewiz).

Amino Acid Sequences of sfGFP mutants:
Tag sequence is in bold, anionic residues in the tag in red, <u>thrombin cleavage site</u> is underlined

<u>sfGFP (-7)</u>
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYK

τ 1-1 (-9)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDEE

τ 2-1 (-9)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDGDDGES

τ 6-1 (-9)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDDEGGS

τ 1-2 (-12)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDEEEDD

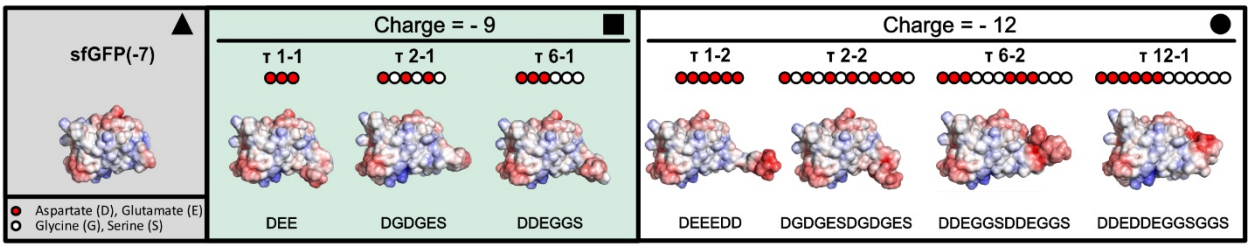
τ 2-2 (-12)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDGDDGESDGDDGES

τ 6-2 (-12)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGSDDEGGSDDEGGS

τ 12-1 (-12)
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQKNGIKAN

FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLE
FVTAAGITHGMDELYKLVPRGS**DDEDDE**GGSGGS

Summary of protein mutants:



Displayed electrostatic maps are approximate structures meant to illustrate differences between the charge distribution of protein mutants. Electrostatic maps were generated using the optimized sfGFP structure (PDB). Minimized polypeptide tag sequence structures were determined using PEP-FOLD3 (http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms). Tag structures were imported into Pymol and manually appended to the C-terminus of sfGFP. PQR files were generated at pH 7.4 (http://nbcr-222.ucsd.edu/pdb2pqr_2.1.1/) and the solvent-accessible surface was visualized using the APBS plugin in Pymol.

GFP(-12) – Isotropic Control, **Mutations from sfGFP** in bold/underlined[1]
MGHHHHHHGGASKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDAT**E**GKLTLKFI
CTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTISFKDDG
TYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNFNSHNVYITADKQ**E**NGIKAN
FKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNE**D**RDHMVLLE
FVTAAGITHGMDELYK

---

3. Protein Expression, Purification, and Preparation

Protein Expression. Anionic GFP mutants were expressed in NiCo21 (DE3) or BL21 (DE3) cells in 1 L cultures of LB media supplemented with 100 μg mL^{-1} ampicillin. Cells were grown at 37 °C, with shaking at 250 rpm to an OD_{600} between 0.6 and 0.8. Subsequently, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cultures were grown for an additional 16-20 h after induction at 37 °C with shaking at 250 rpm.

Protein Purification. Cells were harvested by centrifugation (4000 rpm for 20 min), resuspended in 15-20 mL of cell lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0), and stored at -20 °C. Immediately prior to cell lysis by sonication, 200 μL of DMSO solubilized Protease Inhibitor Cocktail (Sigma Aldrich, P8849) was added to resuspended cells. Cells were lysed in a -20 °C aluminum bead bath using probe tip sonication for 10 min (2 s pulse on, 4 s pulse off). Desired soluble components were separated from cell debris by centrifugation (10,000 rpm for 30 min). GFP was isolated from other soluble components via Ni-NTA affinity chromatography. Wash buffer consisted of 50 mM NaH_2PO_4, 300 mM NaCl, and 35 mM imidazole (pH 8.0), and elution buffer consisted of 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole (pH 8.0). Relative volumes of wash and elution buffers were similar to the manufacturer's instructions with slight modifications made dependent on the mutant. Fractions of the flow through, wash, and

elution were collected and analyzed by SDS-PAGE. Pure fractions were combined and concentrated by centrifugal ultrafiltration with a molecular weight cutoff (MWCO) of 10 kDa. Buffer was exchanged into 10 mM Tris (pH 7.4) with 1 mM EDTA by dialysis using regenerated cellulose dialysis membranes (MWCO of 3.5 kDa) at 4 °C in the dark. At least seven buffer changes were performed over a minimum of 21 h (3 h per exchange).

Sample Preparation. GFP mutants were characterized by MALDI-TOF mass spectrometry to confirm the mass and purity of each protein. Samples were prepared using a 10 mg mL^{-1} sinapinic acid matrix (7:3 acetonitrile to H_2O with 0.1% trifluoroacetic acid) and calibrated with Protein Standard II (Bruker). Matrix (60%) and protein sample (40%) were premixed and 1 µL was spotted for analysis.

The concentration of GFP was determined by measuring absorbance at 488 nm (ε = 83,300 M^{-1} cm^{-1})[2] using a Cary 60 UV-Vis spectrophotometer. GFP working solutions were prepared at 1 mg mL^{-1} in 10 mM Tris (pH 7.4).

4. Cell Growth Kinetics and Protein Expression

Cell growth kinetics and protein expression for different protein mutants were quantified in a Tecan Infinite M200 Pro plate reader. Glycerol stocks were streaked on a LB agar plate containing ampicillin; a single colony was selected and grown in a 5 mL culture (LB media supplemented with 100 µg mL^{-1} ampicillin) overnight to an OD_{600} of approximately 2. This culture was back-diluted to an OD of 0.60 and 1 mM IPTG was added to quantify cell growth kinetics (Figure S3) and protein expression (Figure S4). All dilutions were done by using the OD_{600} of the overnight culture to calculate the respective volumes of overnight culture and LB media needed for a final volume of 1 mL. Each dilution was split into 6 wells (100 µL each) in flat, clear bottom, black polystyrene 96-well plates (Corning). Absorbance at 600 nm was used to measure optical density. The measured absorbance was corrected to that of a 1 cm pathlength using a measured path correction.[3] Fluorescence of GFP mutants (λ_{ex} = 488 nm, λ_{em} = 530 nm, Gain = 50) was used to measure relative protein concentration.[4] The assay was run at 37 °C for 15 h with orbital shaking, with measurements taken every 20 min. To verify that the direct back-dilution to mid-log phase (OD = 0.6) did not impact the induced protein expression, a similar experiment was performed with a different back-dilution. An overnight culture was grown to an OD_{600} of approximately 2. The overnight culture was back-diluted to 1 mL samples (n = 3) with an OD of 0.10. Replicates were then grown to an OD of 0.60 and IPTG (1 mM final conc.) was added to induce protein expression. Each method resulted in nearly identical protein yields and cell growth kinetics.

5. Complex Coacervation Assays

Results from these assays are summarized in Table S1.

Turbidimetric Titration. Protein and polymer solutions were mixed at protein mass fractions ranging from 0 to 1 in increments of 0.04 with a total sample volume of 50 µL. Each mass fraction corresponded to a specific negative charge fraction, calculated using the expected polymer and protein charge. The expected protein charge was calculated using the Henderson-Hasselbalch equation and pK_a values of the isolated amino acids

(Asp = 3.65; Glu = 4.25; Arg = 12.48; Lys = 10.53).[5] At the pH used in these studies these are the only amino acids that contribute significantly to the protein net charge. Inclusion of His (pK_a = 6.00) in this analysis resulted in modest changes to the predicted protein charge (+0.6) and even more minor changes to the calculated charge fraction (~1-3% change). Samples at each mass fraction were prepared in triplicate in tissue culture-treated polystyrene 96-well half-area plates (Corning). Absorbance (λ = 600 nm) was used to evaluate phase separation of the protein/polycation mixtures and was measured using a Tecan Infinite M200 Pro plate reader after 10 s of orbital shaking. The turbidity (τ) was calculated from the absorbance (A) using the formula $\tau = 100 - 10^{(2 - A)}$. Turbidity values are plotted as a function of negative charge fraction $f^- = M^-/(M^- + M^+)$, where M^- and M^+ correspond to the charge per mass of the negative (protein) and positive (polymer) species, respectively (Figure S5A).

Encapsulation Efficiency. From the turbidity data, five specific mass fractions were selected for each GFP mutant to analyze the partitioning of GFP in the dense coacervate phase (Table S1 and Figure S5B). At these ratios, protein and polymer were mixed in triplicate in 1.5 mL tubes at a total sample volume of 100 µL. After spontaneous phase separation and 10 min incubation, the samples were microcentrifuged at 13,000 rpm for 25 min to physically separate the dense and dilute phases. The dilute phase was removed via pipette and diluted 2-fold and 10-fold with 10 mM Tris (pH 7.4) in flat, clear bottom, black polystyrene 96-well plates (Corning). The absorbance (λ = 488 nm) and fluorescence (λ_{ex} = 488 nm, λ_{em} = 530 nm, Gain = 47, 55) of the dilute phase was measured using a Tecan Infinite M200 Pro plate reader. The concentration of GFP (mg mL^{-1}) in the dilute phase was calculated using calibration curves previously determined from serial dilutions of GFP from 0.25 µg mL^{-1} to 0.4 mg mL^{-1}.

Salt (NaCl) Titration. The ionic strength dependence of GFP/polycation phase separation was determined by titrating NaCl into protein/polymer mixtures (Figure S5C). Macromolecule composition was optimally selected from turbidity and encapsulation data at the ratio with the highest protein encapsulation. Microscopy images were considered in this optimization process on the basis of condensate morphology, where selected mixing ratios appeared to phase separate into two distinct liquid states. Mixtures were prepared in 4 mL poly(methyl methacrylate) cuvettes at a total sample volume of 1 mL. The sample was stirred at 850 rpm and the percent transmittance (%T) of the phase separated mixture was measured using a Cary 60 UV-Vis spectrophotometer. 5 M NaCl (in 10 mM Tris, pH 7.4) was added to the mixture in 1 µL increments. Transmittance measurements were taken 5, 10, and 15 s after each incremental addition of salt to allow for sample equilibration. NaCl was added following this procedure until %T stabilized and remained unchanged following 4-5 subsequent additions of salt. Turbidity (τ) was calculated from percent transmittance (%T) using the formula τ = 100 – %T.

Phase Portraits. Protein and polymer solutions were prepared at 2 mg mL^{-1}. A salt solution of 1 M NaCl in 10 mM Tris (pH 7.4) was prepared and used to investigate the effect of ionic strength on the partitioning of macromolecules in the dilute and coacervate phases. Protein and polymer were mixed at the maximum encapsulation ratio, as described above, reported in Table S1, at various salt concentrations with a total sample volume of 50 µL. Samples were prepared in triplicate in 96-well round-bottom plates

(Nunc) and incubated for approximately 10 min. Phases were separated by centrifugation at 3,000 rpm for 30 min. The supernatant was extracted and diluted 10-fold (1 M NaCl, 10 mM Tris, pH 7.4). The remaining coacervate pellet was imaged using a Gel-Doc XR+ (Bio-Rad), thresholded using ImageJ, and computationally analyzed to determine the volume using a calibration curve of known volumes (Figure S7, see Matlab script). The coacervate pellet was dissolved and diluted in 350 µL of 1 M NaCl, 10 mM Tris (pH 7.4). The absorbance (λ = 488 nm) and fluorescence (λ_{ex} = 488 nm, λ_{em} = 530 nm, Gain = 47, 55) of the coacervate and dilute phases were measured in flat, clear bottom, black polystyrene 96-well plates (Corning) using a Tecan Infinite M200 Pro plate reader. To determine the binodal phase boundary, the concentration of each anionic GFP mutant (mg mL^{-1}) at a range of ionic strengths (25 mM to 200 mM) was calculated for both the dilute and coacervate phase using calibration curves (Figure 2B).

Equilibration Experiments. Phase portraits were also investigated for multiple equilibration conditions. Figure S8 represents this equilibration data for variant τ 6-2. In each of these cases, the equilibration step was performed and immediately followed by the standard binodal composition assay described above. These included longer equilibration times prior to centrifugation (24 h, 1 week) and thermal equilibration (20 °C → 50 °C → 20 °C) in a thermal cycler. This temperature ramping was well below the thermal denaturation temperature of sfGFP.[6,7]

Microscopy. Microscopy images were collected using an EVOS FL Auto 2 inverted fluorescence microscope (Invitrogen). From the turbidity data, five specific mass fractions were selected for each GFP mutant for analysis via microscopy. GFP mutants and polycation were mixed together in black 384-well plates with optically-clear bottoms (Nunc) at a total sample volume of 30 µL. Samples were incubated at room temperature for 3 h. Images were taken using a 20X, long working distance objective (NA 0.4) in both GFP fluorescence (λ_{ex} = 470 nm, λ_{em} = 525 nm) and phase-contrast channels.

6. Isothermal Titration Calorimetry.

Data Acquisition. Isothermal Titration Calorimetry was performed in a Malvern MicroCal Auto-iTC200 in triplicate for a subset of protein mutants to compare coacervation thermodynamics. This subset included the proteins: sfGFP, τ 2-2, τ 6-2, and isotropic GFP(-12). This was done with 1 µL injections of polymer into a 200 µL cell containing protein at a temperature of 25 °C, waiting 300 s between injections. Additionally, multiple blanks were performed using 10 mM Tris buffer as the titrant and each respective protein in the cell. This blank was subtracted from collected data as a minor correction accounting for heat of mixing. Similarly, titrations of polymer into buffer were performed to account for heat of dilution. Both of these contributions to measured enthalpy changes were relatively minor. qP4VP was used at a concentration of 4.2 mg mL^{-1} (0.150 mM) and proteins were used at a concentration of 1 mg mL^{-1} (~0.034 mM). Data were exported in a CSV file, corrected, and parameterized in MATLAB. Rigorous cleaning with detergent and salt-in-salt titrations with PBS were performed after experimental runs to ensure dissolution of the coacervate and to prevent cross contamination between trials.

ITC Data Analysis. Data analysis was performed via the Two Step Model described by Priftis et al.[8,9] This model includes 6 parameters in the parameter space: ion-pairing

stoichiometry (n_{IP}), coacervate stoichiometry (n_{coac}), characteristic enthalpies (ΔH_{IP}, ΔH_{coac}), affinity constant (K_A), and a constant proportional to the width of a gaussian curve approximating the range where coacervation occurs (α). Parameter estimates are provided with the corresponding standard error measurements. The MATLAB code used to fit the data is also included in SI Section 16.

7. Fig. S1 - MALDI-TOF of Protein Mutants

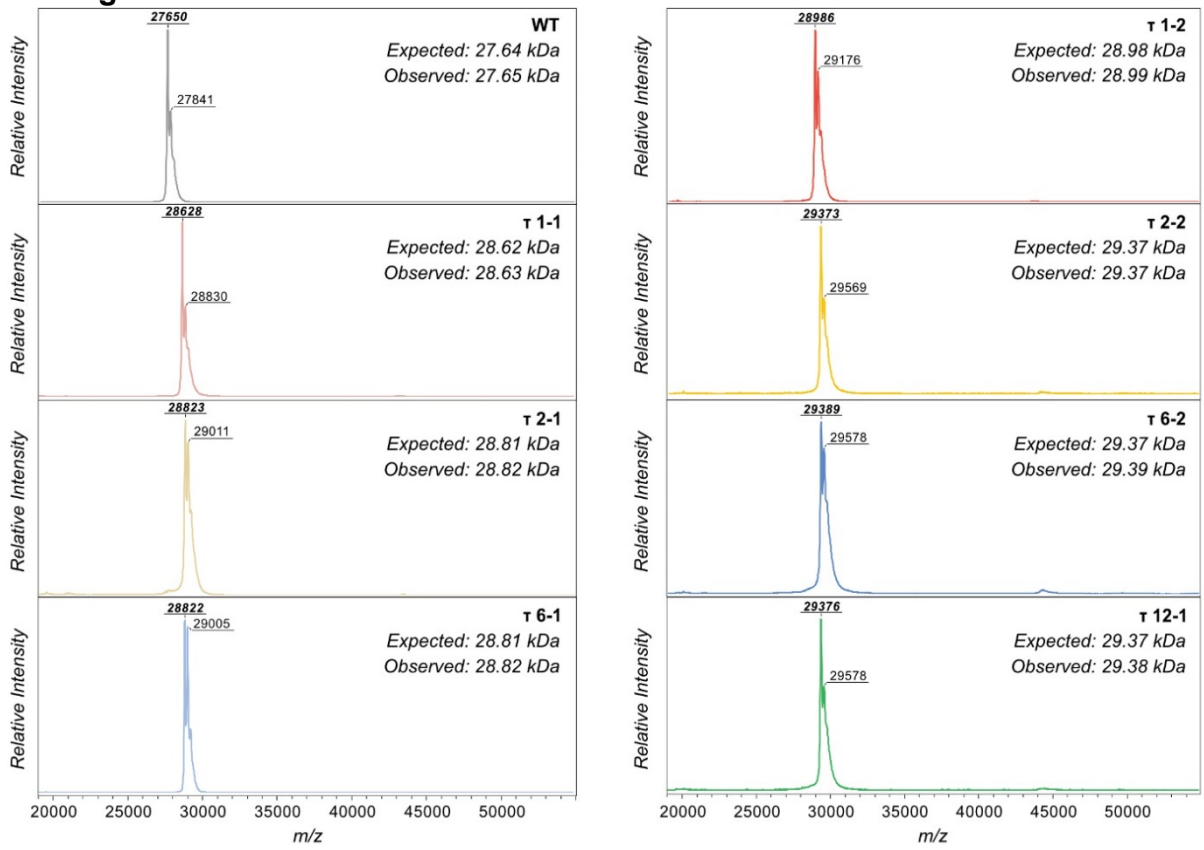


Figure S1. MALDI-TOF MS characterization of purified proteins. Expected molar mass refers to the adjusted theoretical mass, accounting for methionine cleavage and chromophore formation. Observed molar mass refers to the corresponding methionine-adjusted peak (bold, underlined peak). The higher mass peak corresponds to the protein with methionine.

8. Fig. S2 - Electrophoresis Analysis of Engineered Proteins

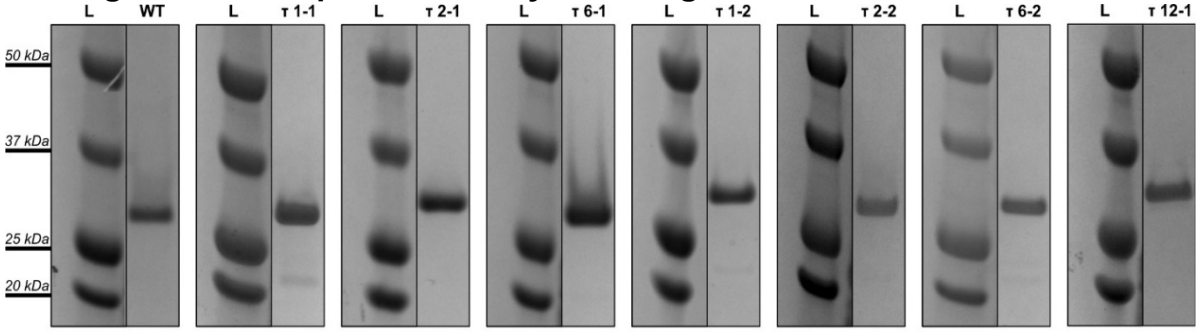


Figure S2. SDS-PAGE analysis of sfGFP and tagged GFP mutants. For each case, bands corresponding to the ladder and the GFP mutant were cropped from the same gel. 4-12% Bis-Tris Bolt gels (Thermofisher) were run at 200 V for 30 minutes.

9. Fig. S3 - *E. Coli* Cell Growth Kinetics

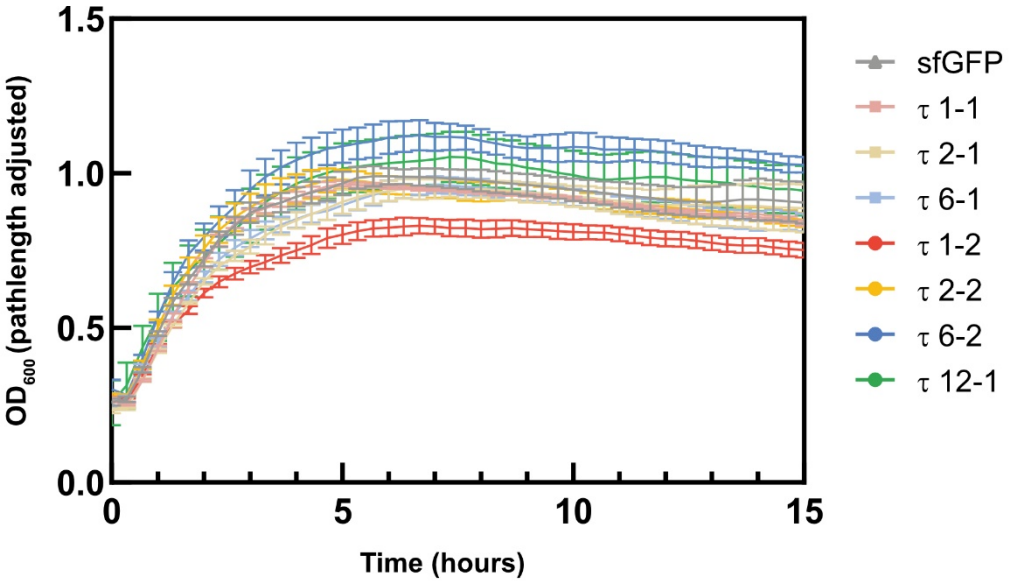


Figure S3. Cell growth kinetics of *E. coli* expressing each member of the protein library were analyzed using whole-cell measurements. Overnight cultures of *E. coli* cells expressing each protein mutant were grown to an optical density between 1 and 2 and back-diluted to OD_{600} = 0.60. IPTG was added and back-dilutions were plated with 6 replicates and the absorbance at 600 nm was recorded over 15 h. Absorbance was converted to OD_{600} using a pathlength correction for the 100 µL sample volume. Error bars represent the standard deviation for absorbance measurements in each sample group at that time, scaled according to the pathlength correction.

10. Fig. S4 - Whole Cell Fluorescence Measurements in E. Coli

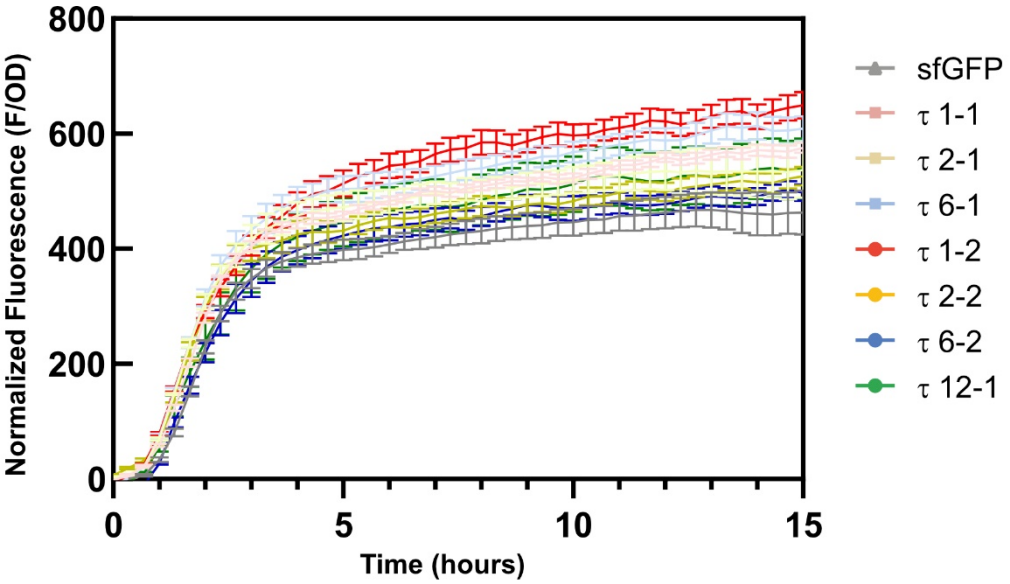


Figure S4. Whole-cell fluorescence measurements of *E. coli* cells (induced with IPTG) expressing each protein mutant. Fluorescence measurements of the induced sample group are normalized to the optical density of the induced sample group. Overnight cultures of *E. coli* cells expressing each protein mutant were grown to an optical density between 1 and 2 and back-diluted to OD_{600} = 0.60. Back-dilutions were plated with 6 replicates and OD (600 nm) and GFP fluorescence (λ_{ex} = 488 nm, λ_{em} = 530 nm, Gain = 50) were recorded over 15 h. Error bars represent the error propagation of absorbance and fluorescence measurements in each sample group, computed via the standard method of adding in quadrature for measurement errors assumed to be governed by independent Gaussian distributions (sqrt. of the sum of squared fractional uncertainty)

11. Table S1 - Summary of Protein and Phase Separation Parameters

Table S1. Summary of complex coacervation behavior of GFP mutants used in this study. Five charge fractions were selected from turbidimetric assays for protein encapsulation studies. **Bold fractions** indicate the relative amount of protein/polymer used in subsequent analysis (salt titrations, phase portraits). These fractions are highlighted in the microscopy images of Figure S6 with a black border. GFP(-12), an isotropically charged sfGFP variant, was used as a control in this study. Complex coacervation parameters for this protein were previously reported by Kapelner.[1]

	Expected Charge	*Expected Molecular Weight (kDa)*	*Negative Charge Fraction (f-)*	*Protein Mass Fraction*	*Turbidity*	*Fraction Encapsulated*
WT sfGFP	-7	27.64	0.092 0.128 **0.153** 0.230 0.396	0.64 0.72 **0.76** 0.84 0.92	49.81 78.99 82.65 80.14 67.65	0.790 0.921 0.971 0.992 0.590
τ 1-1	-9	28.62	0.114 0.133 **0.144*** 0.224 0.275	0.64 0.68 **0.70** 0.80 0.84	88.09 90.98 90.47* 77.67 84.87	0.935 0.986 0.989 0.986 0.874
τ 2-1	-9	28.82	0.097 0.132 **0.156** 0.223 0.345	0.60 0.68 **0.72** 0.80 0.88	74.93 86.84 87.09 85.76 79.70	0.868 0.968 0.999 0.977 0.685
τ 6-1	-9	28.82	0.113 0.132 **0.156** 0.223 0.274	0.64 0.68 **0.72** 0.80 0.84	73.41 84.72 90.76 75.62 79.23	0.907 0.962 0.989 0.901 0.804
τ 1-2	-12	28.99	0.147 0.171 **0.184*** 0.279 0.337	0.64 0.68 **0.70** 0.80 0.84	88.03 91.39 89.71* 89.83 81.96	0.974 0.996 0.998 0.966 0.898
τ 2-2	-12	29.37	0.145 0.169 **0.197** 0.276 0.334	0.64 0.68 **0.72** 0.80 0.84	76.29 87.89 90.79 92.87 82.05	0.952 0.994 0.998 0.970 0.872
τ 6-2	-12	29.37	0.145 **0.169** 0.197 0.276 0.334	0.64 **0.68** 0.72 0.80 0.84	91.19 90.29 87.26 93.62 83.79	0.993 0.998 0.998 0.923 0.809
τ 12-1	-12	29.37	0.125 **0.169** 0.232 0.334 0.412	0.60 **0.68** 0.76 0.84 0.88	65.10 89.60 67.92 81.06 48.23	0.887 0.989 0.980 0.853 0.654
GFP(-12)	-12	27.62	**0.207**‡	**0.72**	87.41‡	0.862‡

** denotes charge fractions that were not investigated in turbidimetric assays. In each case, the turbidity value listed is an average of the adjacent points (0.68 and 0.72).*

‡ *values and parameters reported by Kapelner.*[1]

12. Fig. S5 - Summary of Complex Coacervation Assays with sfGFP Mutants and qP4VP

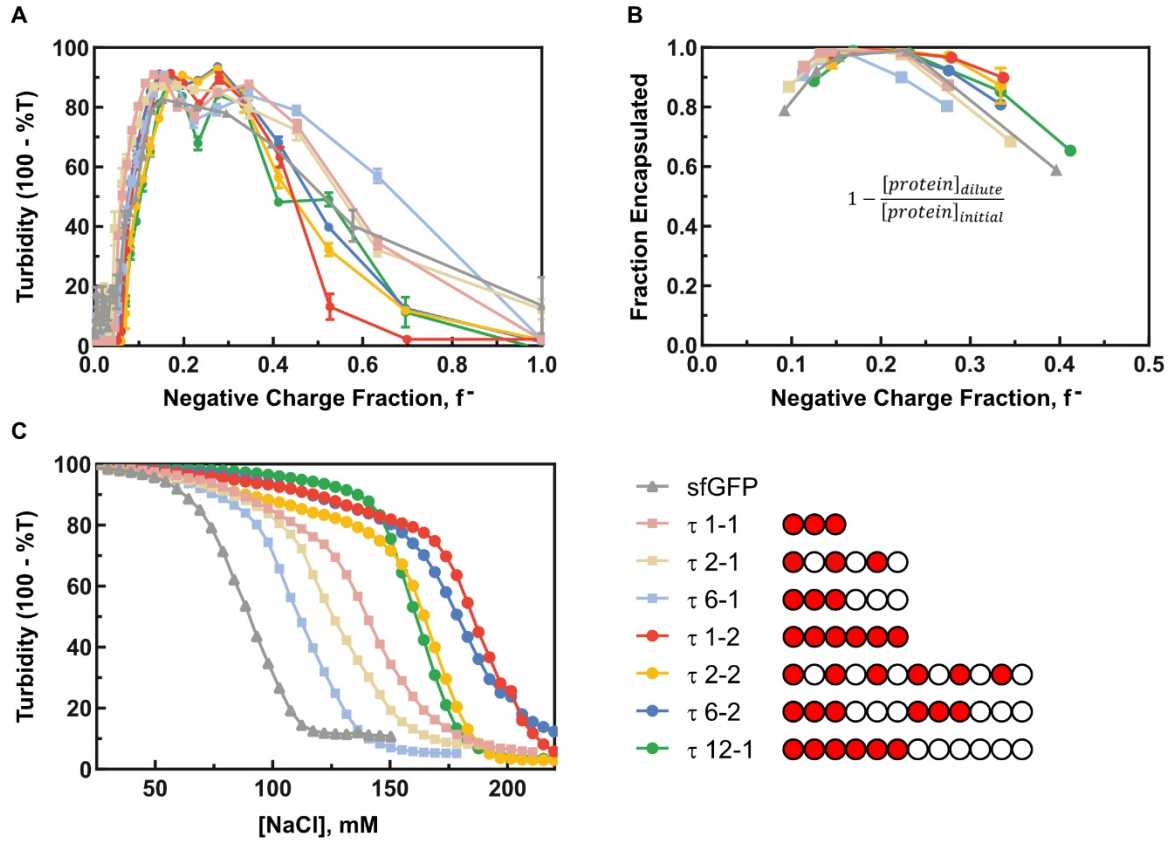
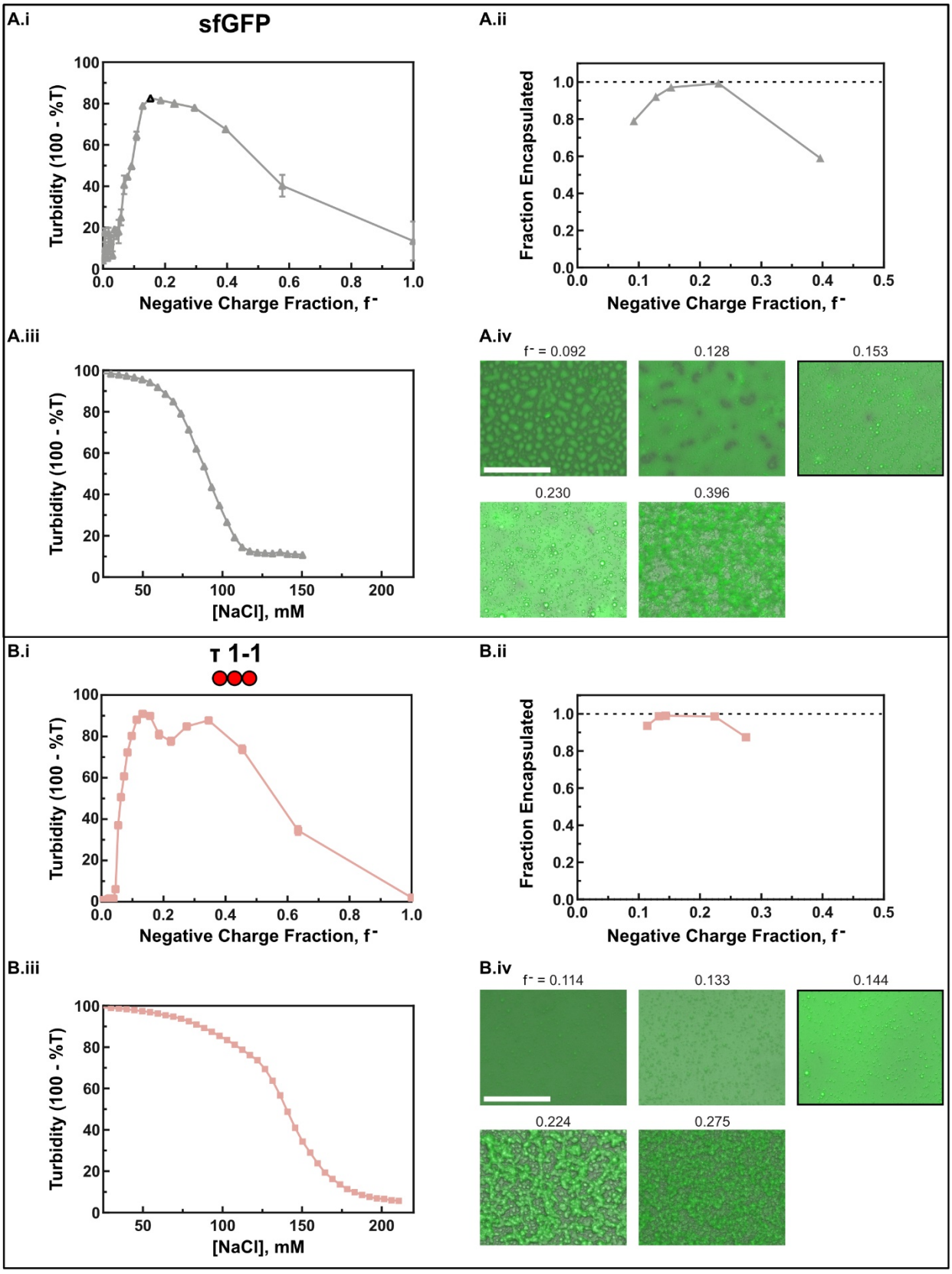
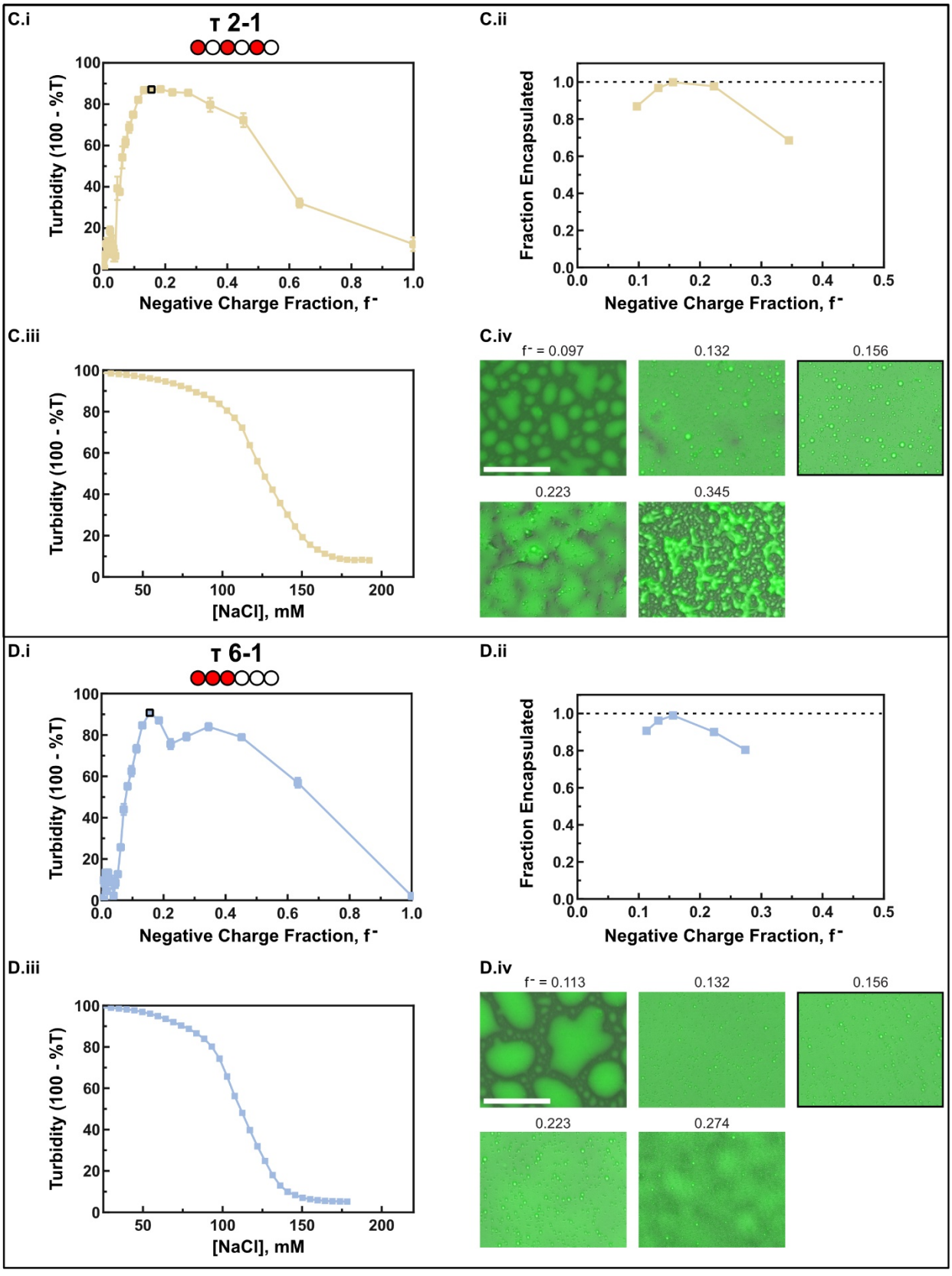
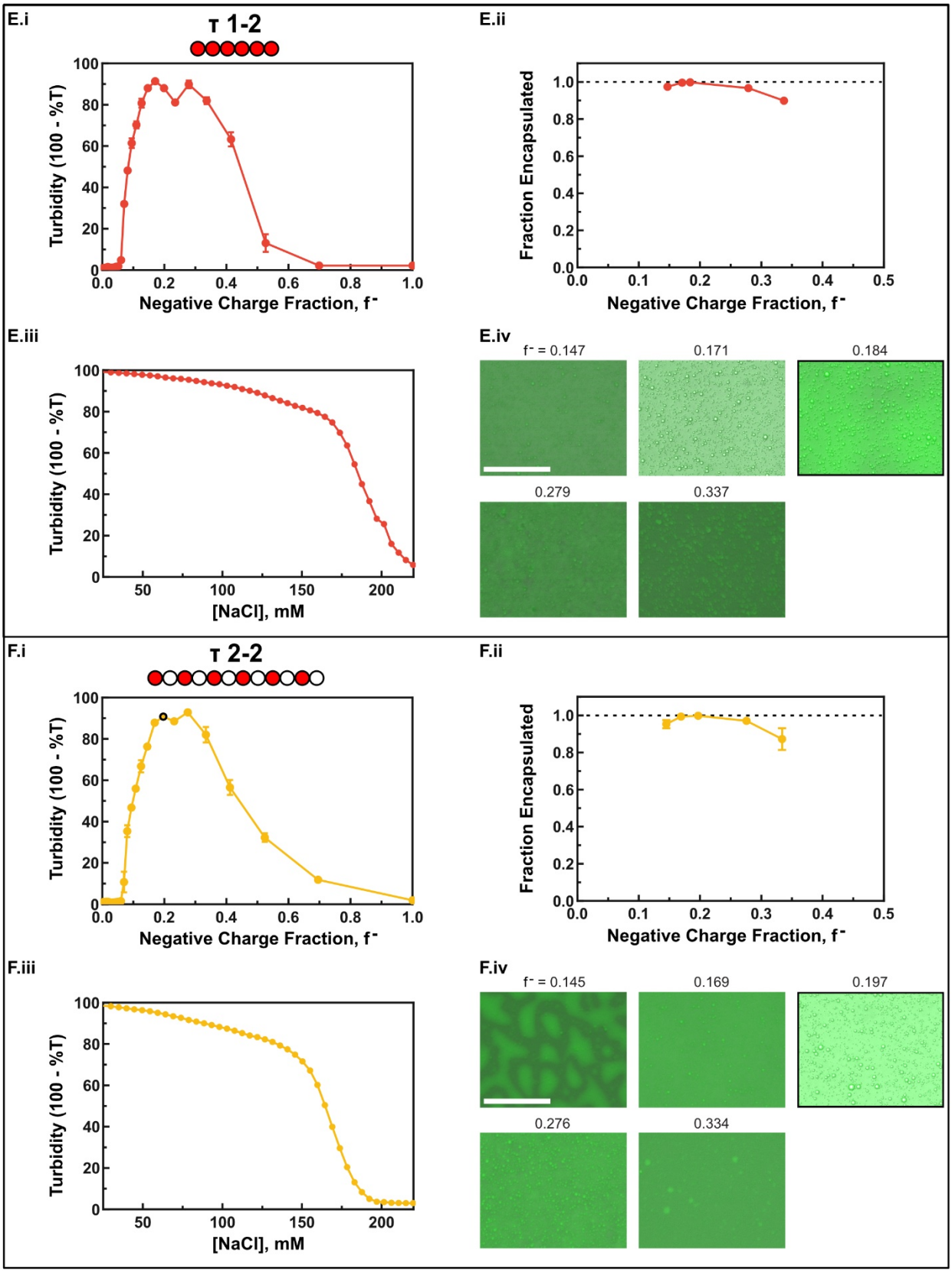


Figure S5. Summary of phase separation assays performed with sfGFP and tagged mutants with qP4VP. All assays were performed at a total macromolecule concentration of 1 mg mL^{-1} in 10 mM Tris, pH 7.4. (A) Turbidity measurements for different ratios of protein and polymer are shown as a function of negative charge fraction. (B) Encapsulation efficiency of tagged GFP mutants compared to sfGFP for 5 macromolecule ratios selected by initial turbidity studies (Table S1). The fraction of protein encapsulated was quantified by measuring the protein concentration in the dilute phase after centrifugation where, *fraction encapsulated* = $1 - [protein]_{\text{dilute phase}} / [protein]_{\text{initial}}$. (C) GFP mutant and qP4VP complex dissolution as a function of NaCl concentration was evaluated via titration studies. The macromolecule ratios analyzed were selected on the basis of encapsulation efficiency and microscopy data and were subsequently the ratios investigated in phase portraits, as indicated in Table S1. Transmittance (%T) was measured after 1 µL additions of 5 M NaCl.

13. Fig. S6 - Detailed Experimental Results for sfGFP and Tagged Protein Mutants




C.i
T 2-1
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
C.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
C.iii
Turbidity (100 - %T)
[NaCl], mM
C.iv
f⁻ = 0.097
0.132
0.156
0.223
0.345
D.i
T 6-1
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
D.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
D.iii
Turbidity (100 - %T)
[NaCl], mM
D.iv
f⁻ = 0.113
0.132
0.156
0.223
0.274


E.i
T 1-2
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
E.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
E.iii
Turbidity (100 - %T)
[NaCl], mM
E.iv
f⁻ = 0.147
0.171
0.184
0.279
0.337
F.i
T 2-2
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
F.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
F.iii
Turbidity (100 - %T)
[NaCl], mM
F.iv
f⁻ = 0.145
0.169
0.197
0.276
0.334

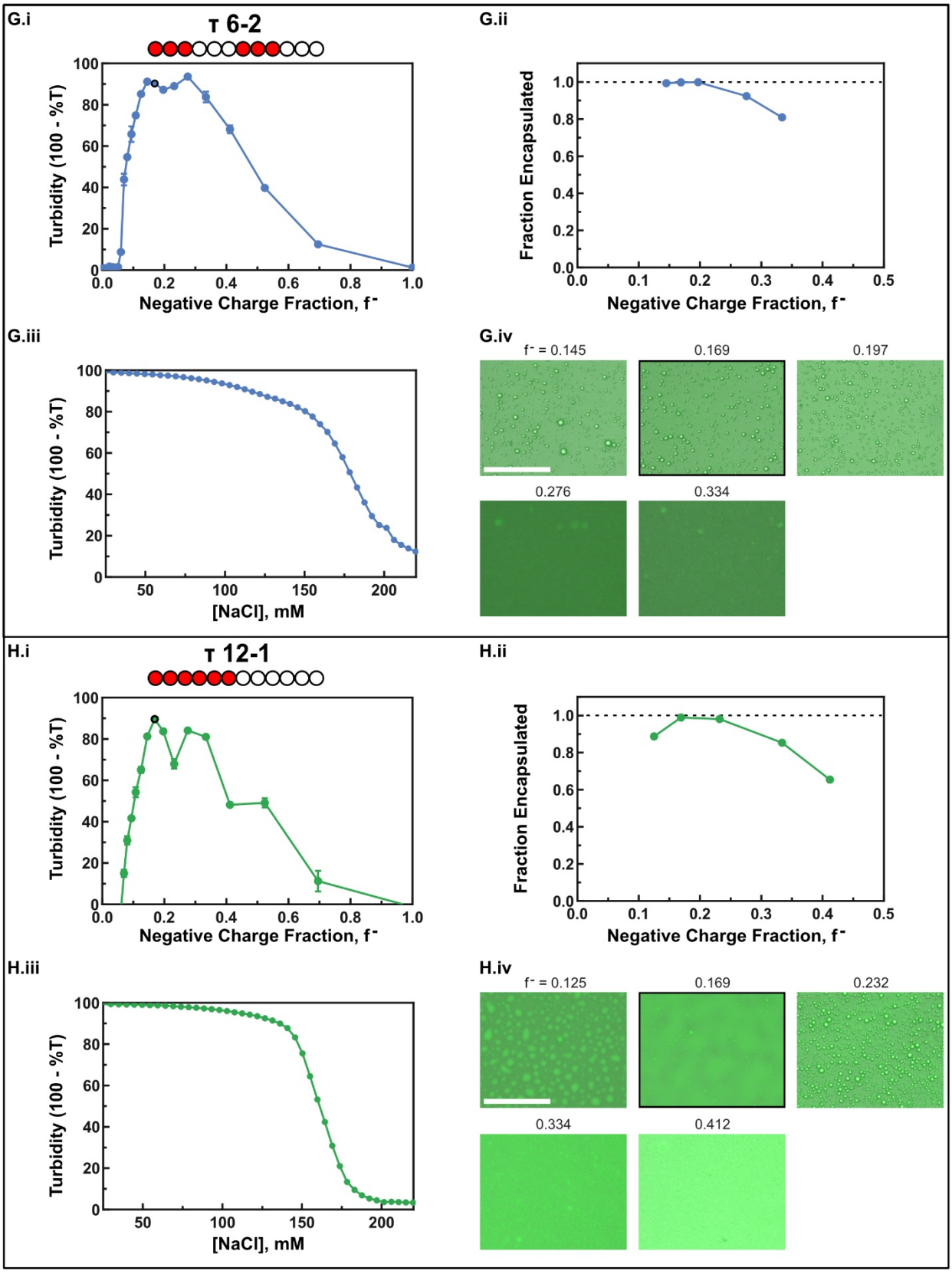
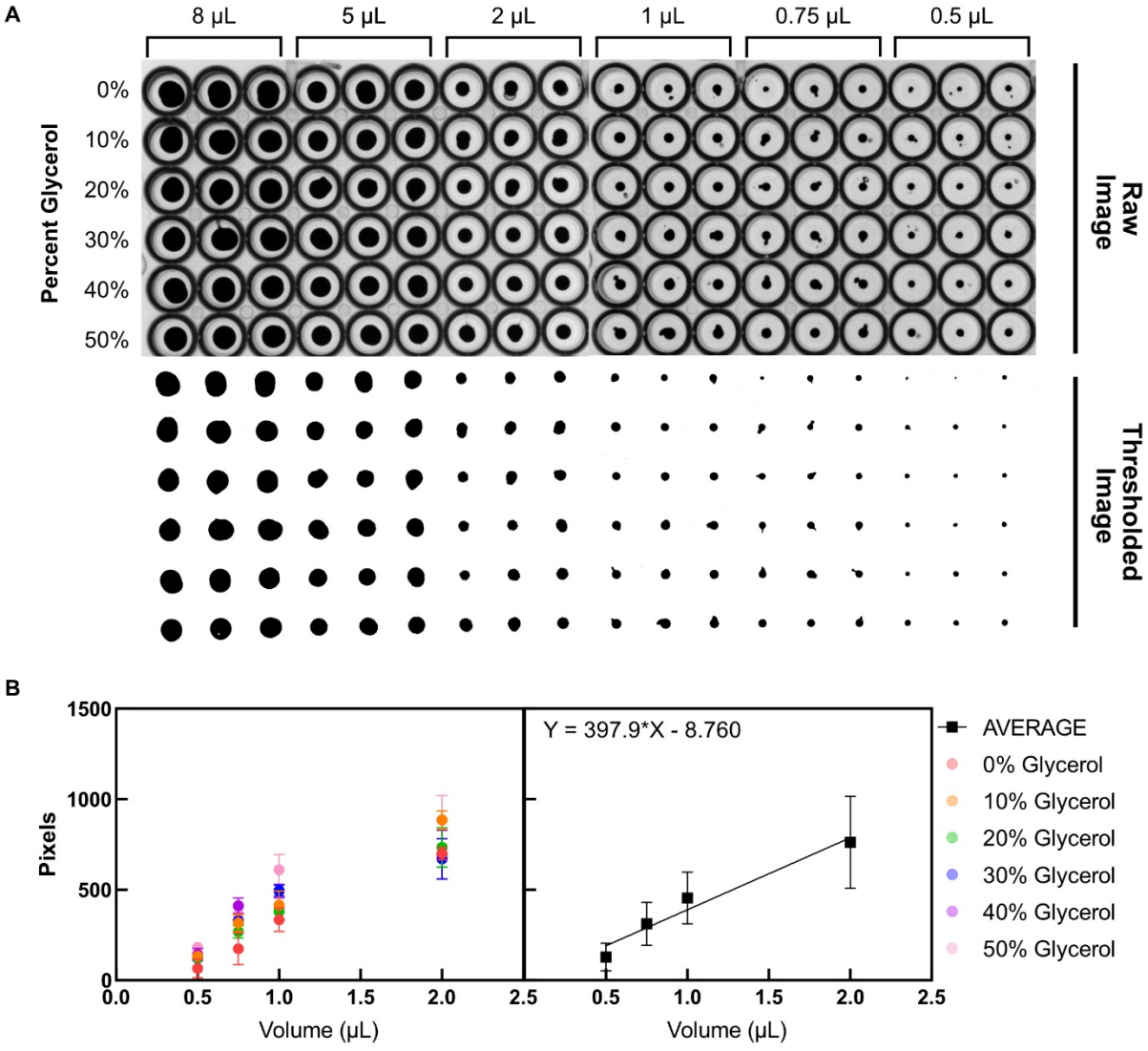
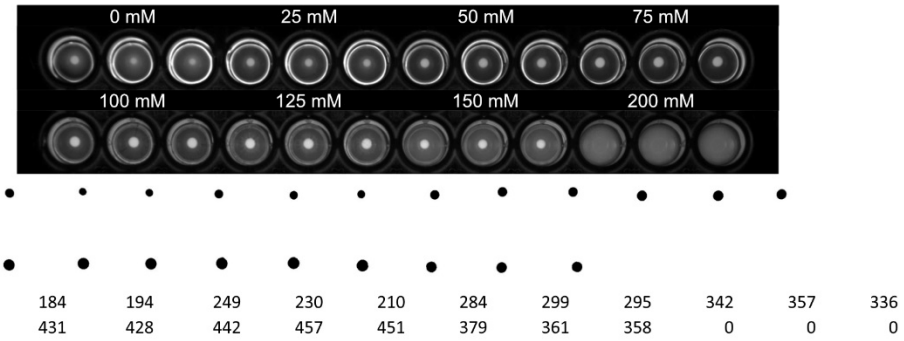

G.i
T 6-2
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
G.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
G.iii
Turbidity (100 - %T)
[NaCl], mM
G.iv
f⁻ = 0.145
0.169
0.197
0.276
0.334
H.i
T 12-1
Turbidity (100 - %T)
Negative Charge Fraction, f⁻
H.ii
Fraction Encapsulated
Negative Charge Fraction, f⁻
H.iii
Turbidity (100 - %T)
[NaCl], mM
H.iv
f⁻ = 0.125
0.169
0.232
0.334
0.412

Figure S6. Detailed experimental results for sfGFP and tagged protein mutants: (A) sfGFP, (B) τ 1-1, (C) τ 2-1, (D) τ 6-1, (E) τ 1-2, (F) τ 2-2, (G) τ 6-2, (H) τ 12-1. All experiments were conducted with GFP and qP4VP at a total macromolecule concentration of 1 mg mL^{-1} in 10 mM Tris, pH 7.4. (i) Individual turbidimetric assays for each anionic protein with cationic polymer, qP4VP across a range of negative charge fractions. (ii) Encapsulation efficiency for each protein mutant was approximated for 5 macromolecule ratios (See Table S1). These ratios were selected on the basis of turbidity results. (iii) Salt titrations with NaCl to investigate complex dissolution for the GFP and qP4VP system. NaCl (5 M) was added in 1 µL increments while stirring. Transmittance (%T) was measured in a UV-Vis spectrophotometer to approximate the critical salt concentration of protein/polymer complexes. (iv) Microscopy images of protein/polymer complexes after approximately 3 h of incubation. Merged images are displayed of the phase contrast and GFP channels. The same 5 macromolecule ratios were studied and are displayed in increasing negative charge fraction from left to right. The macromolecule ratio investigated in titrations and phase portraits is emphasized with a black border in both turbidity graphs (when ratio was part of turbidimetric assay) and microscopy images. Scale bar = 125 µm.

14. Fig. S7 - Measuring Coacervate Volumes: Calibration and Analysis



C



D

NaCl Concentration (mM)	Average Droplet Volume (µL)	Standard Deviation
0	-	-
25	0.60	0.05
50	0.76	0.02
75	0.89	0.03
100	1.09	0.02
125	1.15	0.02
150	0.94	0.03
200	-	-

Figure S7. Outline and example of the method used to determine coacervate droplet volume. Protein and polymer mixtures were centrifuged in round bottom plates and overhead images were used to determine volume using a calibration. (A) Calibration curves were generated using mixtures of water, glycerol, and a small amount of Coomassie Blue. This range of glycerol concentrations spans a density range of 1.0 - 1.13 g mL^{-1}, compared to an approximate coacervate density range of 1.06 - 1.15 g mL^{-1}. Small volume droplets were submerged in a droplet of mineral oil to prevent evaporation. Raw images (top) were thresholded (bottom) and analyzed to create a calibration for droplet volume using the number of pixels detected. (B) Calibration curves for each glycerol concentration were generated spanning a volume of 0.5 - 2.0 µL (left). Due to varying amounts of salt, protein, and polymer in the coacervate phase a range of coacervate densities was expected. An average trendline (right) was used to compute the coacervate volume (X) from the number of pixels (Y). (C) Example overhead UV images of GFP and qP4VP coacervates (in triplicate) after centrifrugation: raw (top), thresholded (middle), and pixels detected for each well (bottom). (D) Calculated average droplet volumes using the trendline in (B) at each NaCl concentration. Volumes at 0 mM were omitted due to spreading and an inability to threshold in many cases.

15. Fig. S8 - Testing for Equilibration in Phase Portraits for GFP mutant τ 6-2

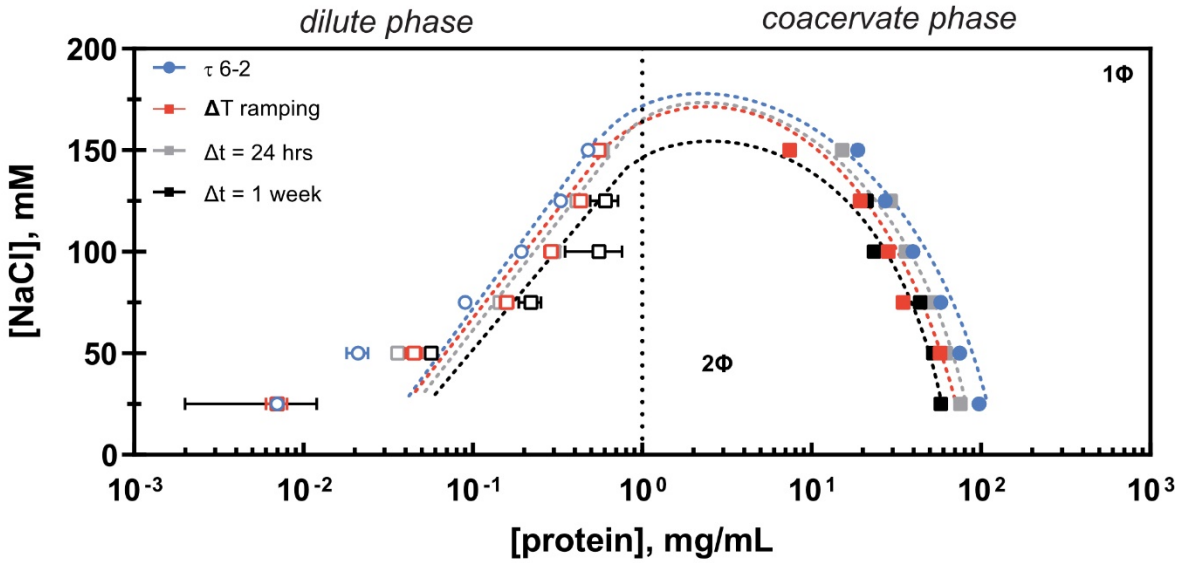


Figure S8. Phase portraits for GFP mutant τ 6-2 at discrete NaCl concentrations (0*, 25, 50, 75, 100, 125, 150, 200 mM) under different incubation conditions. Phase portraits were made for sample groups that underwent longer equilibration times (24 h, 1 wk) compared to the standard incubation time of 10 min. Thermal equilibration was investigated using a thermal cycler, where protein, polymer, and salt were mixed according to protocol, heated to 50 °C (below denaturation temperature[5,6]) and slowly ramped (0.1 °C/s) back to room temperature. Protein concentrations in the dilute phase (left) and coacervate phase (right) were determined using fluorescence measurements, using a calibration curve of overhead UV Images to determine the volume of the bulk coacervate phase (Figure S7). Dotted lines approximating the binodal phase boundary are simply a visual aid and do not represent theoretical modeling or fitting. The initial protein concentration of 1 mg mL^{-1} has been indicated using a dotted vertical line. *Protein concentrations at 0 mM NaCl have been omitted due to the formation of less-hydrated films, where the volume of the coacervate could not be determined.

16. Fig. S9 - Summary of ITC Results and Parameters

A

	n_{IP}	n_{coac}	ΔH_{IP}	ΔH_{coac}	K_A	α
WT	0.270 ± 0.004	0.102 ± 0.005	3.99 ± 0.01	0.198 ± 0.040	54.7 ± 0.5	0.068 ± 0.006
τ 2-2	0.271 ± 0.004	0.165 ± 0.013	4.00 ± 0.01	0.682 ± 0.052	76.2 ± 1.3	0.086 ± 0.002
τ 6-2	0.303 ± 0.060	0.174 ± 0.005	4.00 ± 0.01	0.760 ± 0.095	117.5 ± 22.5	0.093 ± 0.005
(-12)	0.329 ± 0.023	0.138 ± 0.011	3.98 ± 0.01	0.446 ± 0.077	67.5 ± 5.0	0.110 ± 0.016

B

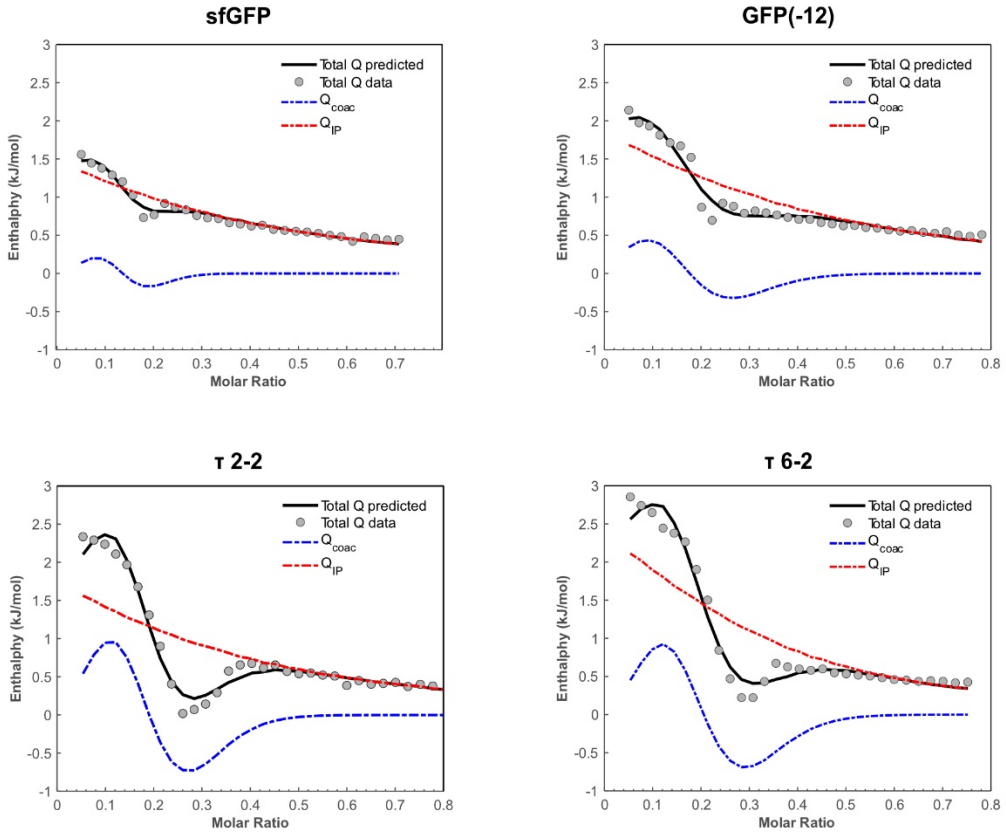


Figure S9. Summary and visualization of ITC fits for selected subset of protein mutants (A) Parameter values for ITC fits determined by a parameter fitting script written in MATLAB using the model described previously by Priftis et al.[7,8] (B) MATLAB generated ITC fits using the Two-Step Binding Model. Error in parameter values represents the standard deviation between trials (N = 3).

17. MATLAB Code.

Coacervate Droplet Volume Analysis

```
I = imread('Droplet_Image_File.tif');
[J, rect] = imcrop(I);
I2 = imcrop(I,rect);

% Get the dimensions of the image.  numberOfColorBands should be = 3.
[rows columns numberOfColorBands] = size(I2)

blockSizeR = 145; % Rows in block.
blockSizeC = 145; % Columns in block.

% Figure out the size of each block in rows.
% Most will be blockSizeR but there may be a remainder amount of less than that.
wholeBlockRows = floor(rows / blockSizeR);
blockVectorR = [blockSizeR * ones(1, wholeBlockRows), rem(rows, blockSizeR)];
% Figure out the size of each block in columns.
wholeBlockCols = floor(columns / blockSizeC);
blockVectorC = [blockSizeC * ones(1, wholeBlockCols), rem(columns, blockSizeC)];

% Now display all the blocks.
plotIndex = 1;
numPlotsR = size(ca, 1);
numPlotsC = size(ca, 2);
for r = 1 : numPlotsR;
    for c = 1 : numPlotsC;
        fprintf('plotindex = %d,   c=%d, r=%d\n', plotIndex, c, r);
        % Specify the location for display of the image.
        subplot(numPlotsR, numPlotsC, plotIndex);
        rgbBlock = ca{r,c};
        imshow(rgbBlock); % Could call imshow(ca{r,c}) if you wanted to.
        [rowsB columnsB numberOfColorBandsB] = size(rgbBlock);
        caption = sprintf('Block #%d', ...
            plotIndex, numPlotsR*numPlotsC, rowsB, columnsB);
%         title(caption);
        drawnow;
        % Increment the subplot to the next location.
        plotIndex = plotIndex + 1;
    end
end


A=nan(numPlotsR,numPlotsC); % A is a matrix of pixels
for n=1:numPlotsR
    for m=1:numPlotsC
        BW= imbinarize(ca{n,m});
        BW1=double(BW);
        [ymax,xmax] = size(ca{n,m});
        White_pix=0;
        Black_pix=0;
        for j=1:(xmax)-1
            for i=1:(ymax)-1
                if BW1(i,j)==1
                    White_pix = White_pix+1;
                else
                    Black_pix=Black_pix+1;
                end
            end
        end
        A(n,m)=Black_pix;
    end
end
```

ITC Fitting Analysis

Script

```
% Script to fit ITC data using a method by Sarah Perry's Group featured in Nature Communication 2017
% Sequence and entropy based contol...

% Fitting for complexation (ion pairing) and complex coacervation will be
% separately done



% This method of analysis will include the fitting of 6 parameters:
% n_IP, n_coac, DH_IP, DH_coac, K_assoc, and alpha

C_Inj = 0.150; % mM
Vo = 200; % uL
```

Load Data

```
Data = importdata("t221.xlsx");
%InjV = Data(:,2); % uL
Xt = Data(:,3); % mM
Mt = Data(:,4); % mM
Ratio = Data(:,5);

Enthalpy_Exp = Data(:,9); % kJ/mol

% Theta_IP = 0.5.*(1 + (Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))-
% sqrt((1+(Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))).^2-(4.*Xt./(n_IP.*Mt))));
```

Optimization

```
params0 = [1,1,1,2,5,0.10];
lb = [0,0,0,0,0,0];
ub = [1,1,4,10,200,1];
options = optimoptions('fmincon','Algorithm','interior-point',...
    'MaxFunctionEvaluations',1e5,'MaxIterations',1e5,'TolX', 1e-10);

 Z = fmincon(@(params)predictor(params,Data),params0,[],[],[],[],...
     lb,ub,[],options);


n_IP = Z(1);
n_coac = Z(2);
DH_IP = Z(3);
DH_coac = Z(4);
K_assoc = Z(5);
alpha = Z(6);

InjV = Data(:,2); % uL
Xt = Data(:,3); % mM
Mt = Data(:,4); % mM
Ratio = Data(:,5);
Enthalpy_Exp = Data(:,9); % kJ/mol

Theta_IP = 0.5.*(1 + (Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))- ...
    sqrt((1+(Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))).^2-...
    (4.*Xt./(n_IP.*Mt))));
```

```
Q_IP = (n_IP.*Theta_IP.*Mt.*DH_IP*Vo)./(C_Inj.*InjV);

DQ_IP = Q_IP(2:end) - Q_IP(1:end-1) + InjV(1:end-1).*(Q_IP(2:end) +...
    Q_IP(1:end-1))/(2*Vo);

f = Xt./(Xt+Mt);
f_coac = n_coac/(1+n_coac);
Theta_coac = exp(-(f-f_coac).^2/alpha.^2);

Q_coac = (n_coac.*Theta_coac.*Mt.*DH_coac*Vo)./(C_Inj.*InjV);

DQ_coac = Q_coac(2:end) - Q_coac(1:end-1) + InjV(1:end-1).*(Q_coac(2:end) +...
    Q_coac(1:end-1))/(2*Vo);

DQtot = DQ_coac + DQ_IP;

err = sum((Enthalpy_Exp(2:end) - DQtot).^2);
```

Predictor Function

```
function err = predictor(params,Data)
```

```
C_Inj = 0.150; % mM
Vo = 200; % uL


n_IP = params(1);
n_coac = params(2);
DH_IP = params(3);
DH_coac = params(4);
K_assoc = params(5);
alpha = params(6);

Enthalpy_Exp = Data(:,9); % kJ/mol
```

IP

```
Theta_IP = 0.5.*(1 + (Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))- ...
    sqrt((1+(Xt./(n_IP.*Mt))+(1./(K_assoc.*n_IP.*Mt))).^2-(4.*Xt./(n_IP.*Mt))));

Q_IP = (n_IP.*Theta_IP.*Mt.*DH_IP*Vo)./(C_Inj.*InjV);

DQ_IP = Q_IP(2:end) - Q_IP(1:end-1) + InjV(1:end-1).*(Q_IP(2:end) + Q_IP(1:end-1))/(2*Vo);
```

Coac

```
f = Xt./(Xt+Mt);
f_coac = n_coac/(1+n_coac);
Theta_coac = exp(-(f-f_coac).^2/alpha.^2);

Q_coac = (n_coac.*Theta_coac.*Mt.*DH_coac*Vo)./(C_Inj.*InjV);

DQ_coac = Q_coac(2:end) - Q_coac(1:end-1) + InjV(1:end-1).*(Q_coac(2:end) + Q_coac(1:end-1))/(2*Vo);

DQtot = DQ_coac + DQ_IP;

err = sum((Enthalpy_Exp(2:end) - DQtot).^2);
```

18. References.

1. Kapelner, R. A. & Obermeyer, A. C. Ionic polypeptide tags for protein phase separation. *Chem. Sci.* **10**, 2700–2707 (2019).
2. Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nat Methods* **16**, 277–278 (2019).
3. Lampinen, J., Raitio, M., Perälä, A., Oranen, H. & Harinen, R.-R. Microplate Based Pathlength Correction Method for Photometric DNA Quantification Assay. 7.
4. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24**, 79–88 (2006).
5. Amino Acids Reference Chart. *Sigma-Aldrich Metabolomics* https://www.sigmaaldrich.com/US/en/technical-documents/technical-article/protein-biology/protein-structural-analysis/amino-acid-reference-chart#prop
6. Lawrence, M. S., Phillips, K. J. & Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. *J. Am. Chem. Soc.* **129**, 10110–10112 (2007).
7. Saeed, I. A. & Ashraf, S. S. Denaturation studies reveal significant differences between GFP and blue fluorescent protein. *International Journal of Biological Macromolecules* **45**, 236–241 (2009).
8. Priftis, D., Laugel, N. & Tirrell, M. Thermodynamic Characterization of Polypeptide Complex Coacervation. *Langmuir* **28**, 15947–15957 (2012).
9. Chang, L.-W. *et al.* Sequence and entropy-based control of complex coacervates. *Nat Commun* **8**, 1273 (2017).