ACS SyntheticBiology

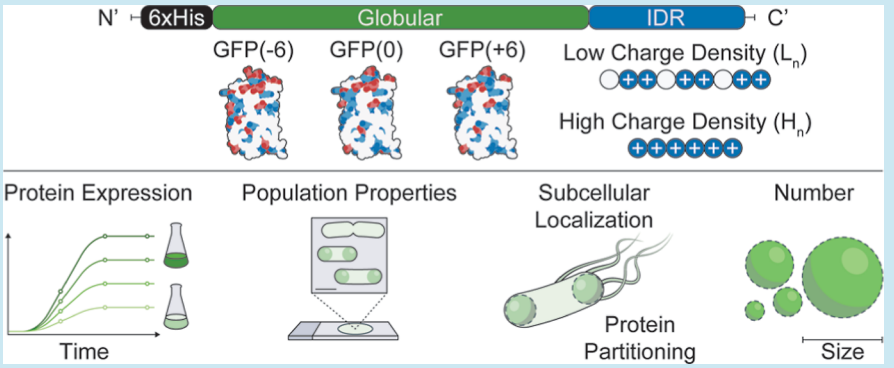
pubs.acs.org/synthbio Research Article

Charge-Patterned Disordered Peptides Tune Intracellular Phase Separation in Bacteria

Jane Liao, Vivian Yeong, and Allie C. Obermeyer*

Cite This: *ACS Synth. Biol.* 2024, 13, 598–612 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Subcellular phase-separated compartments, known as biomolecular condensates, play an important role in the spatiotemporal organization of cells. To understand the sequence-determinants of phase separation in bacteria, we engineered protein-based condensates in *Escherichia coli* using electrostatic interactions as the main driving force. Minimal cationic disordered peptides were used to supercharge negative, neutral, and positive globular model proteins, enabling their phase separation with anionic biomacromolecules in the cell. The phase behavior was governed by the interaction strength between the cationic proteins and anionic biopolymers, in addition to the protein concentration. The interaction strength primarily depended on the overall net charge of the protein, but the distribution of charge between the globular and disordered domains also had an impact. Notably, the protein charge distribution between domains could tune mesoscale attributes such as the size, number, and subcellular localization of condensates within *E. coli* cells. The length and charge density of the disordered peptides had significant effects on protein expression levels, ultimately influencing the formation of condensates. Taken together, charge-patterned disordered peptides provide a platform for understanding the molecular grammar underlying phase separation in bacteria.

KEYWORDS: *biomolecular condensate, membraneless organelle, complex coacervation, protein phase separation*

INTRODUCTION

The compartmentalization of biomolecules is crucial for the spatiotemporal organization of cells. In addition to the well-known membrane-bound organelles that serve diverse functions in eukaryotic cells, a host of naturally occurring membraneless intracellular compartments have been discovered and characterized in eukaryotes.[1] Commonly referred to as biomolecular condensates, these intracellular compartments often concentrate proteins and nucleic acids and play important roles in biological processes spanning molecular- to cellular-length scales.[2] New insights into the mechanism of their formation gained over the last few decades have shed light on numerous endogenous condensates in eukaryotes, including the nucleolus, Cajal bodies, nuclear speckles, paraspeckles, stress granules, and P granules.[3−9] These phase-separated assemblies have been implicated in a variety of essential cellular processes, ranging from regulating cell signaling to redirecting metabolic flux and adapting to stresses.

Recent work suggests that condensates also serve an equally important role in the organization of the bacterial cytoplasm.[10−15] However, despite significant progress toward understanding the formation, composition, material properties, and biological functions of condensates in eukaryotic cells, such insights are still largely lacking when it comes to bacterial cells.

The formation mechanism of bacterial condensates can be elucidated by borrowing concepts from eukaryotic counterparts, where weak multivalent interactions between constituent biomolecules are central to driving condensate assembly.[16]

Received: September 11, 2023
Revised: December 18, 2023
Accepted: January 15, 2024
Published: February 3, 2024

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598–612

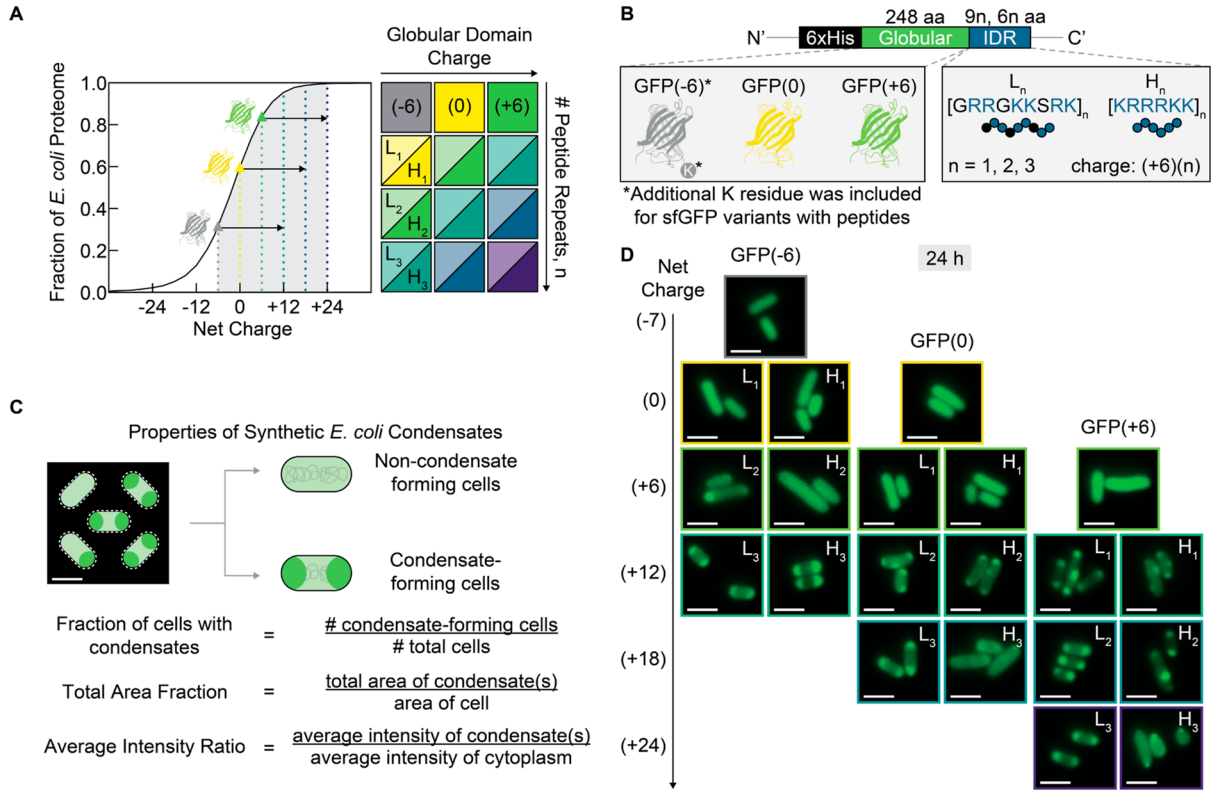


Figure 1. Design of disordered cationic peptides with varying charge density to create synthetic biomolecular condensates in *E. coli*. (A) Cumulative distribution of proteins in the *E. coli* proteome (UP000002032) by expected charge. Triangular markers indicate the predicted charge of engineered isotropic proteins and arrows indicate the net charge increase from disordered cationic peptides. (B) Schematic for the design of globular and C-terminal charge-patterned disordered peptide domains. (C) Cell population, cell, and condensate properties assessed from fluorescence microscopy images. (D) Representative microscopy images of cells at 24 h post-induction shown according to overall net charge. GFP(−6)-L_2 and all protein variants with net charge ≥+12 form foci within the cell. Scale bars are 2 μm.

Here, valency refers to the number of interaction sites, which can be surface regions on a folded protein or amino acid motifs in an intrinsically disordered region or protein (IDR or IDP). A simple yet powerful framework for describing multivalent biomolecules is the stickers-and-spacers model from polymer physics, where stickers are defined as specific interaction sites that form physical cross-links through noncovalent interactions such as hydrophobic, cation−π, aromatic, and electrostatic.[17−19] Importantly, multivalent biomolecules above a threshold concentration, c_{perc}, can undergo a networking transition known as bond percolation. Similarly, above a threshold concentration c_{sat}, biomolecules can undergo a density transition whereby specific biomolecules demix into coexisting dilute and dense phases. Accordingly, phase separation-aided bond percolation transitions have been proposed as a potential mechanism for condensate formation and dissolution in eukaryotic and bacterial cells.[20]

Compositionally, condensate constituents broadly comprise scaffolds that drive phase separation and clients that are recruited through interactions with the scaffold biomolecules. Studies of both endogenous and synthetic eukaryotic condensates have primarily focused on IDRs as condensate scaffolds.[21] In contrast to folded globular domains, the hallmark of IDRs is a high degree of conformation heterogeneity, which enables more flexible interaction modes. A subset of IDRs known as low-complexity domains show compositional bias toward a small set of amino acid stickers such as aromatic or charged residues.[22] Although IDRs are ubiquitous in eukaryotic proteomes, comprising around 33% of the proteome, they are relatively scarce in bacteria and make up around 4% of bacterial proteomes.[23] Despite this, IDRs have been identified as drivers of phase separation in prominent examples of endogenous bacterial condensates in *Caulobacter crescentus*, including the intrinsically disordered domain of Polar Organizing Protein Z (PopZ)[24] and the disordered C-terminal domain of RNase E, an essential endoribonuclease involved in RNA degradation.[25,26]

Given the importance of IDRs and biomolecular condensates for molecular functions and cellular processes in bacteria, several studies have engineered synthetic IDP-based systems, often with simplified repetitive sequence motifs, to uncover the sequence-determinants and molecular grammar of bacterial condensates. Examples in *Escherichia coli* (*E. coli*) include the recombinant overexpression of IDPs such as elastin-like polypeptides, spider silk, and resilin to form membraneless compartments,[27,28] the de novo engineering of condensates using artificial IDPs with different molecular weights and aromatic content,[29] and the use of synthetic resilin-like polypeptides fused to functional domains to increase gene transcription by enriching plasmids and key components of the transcriptional machinery within condensates.[30]

Although significant progress has been made in understanding the sequence heuristics of IDP-driven phase separation in bacterial cells, most studies have predominantly focused on IDPs or IDRs while overlooking contributions from globular protein domains to the overall phase behavior. Previous work has shown that increasing the surface charge

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

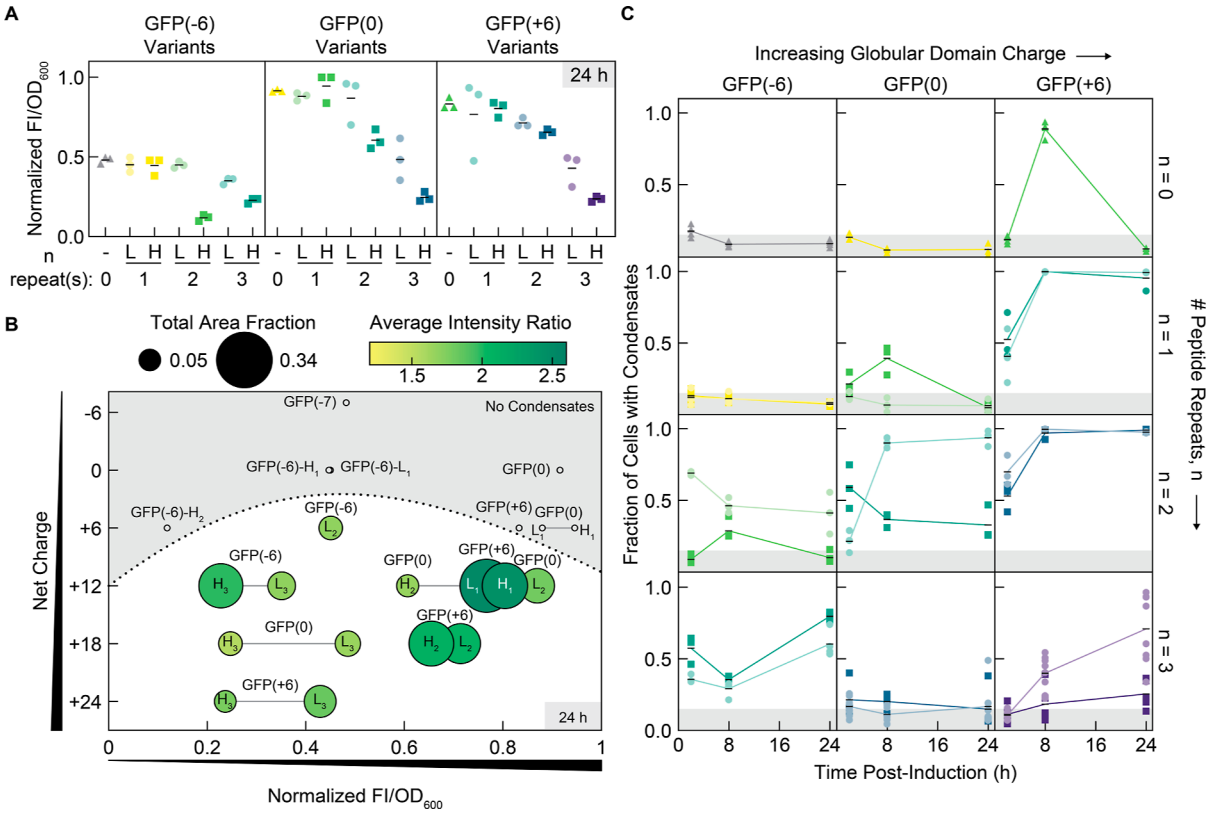


Figure 2. Overall net charge, charge distribution, peptide charge-patterning, and protein concentration determine the phase separation behavior of GFP variants. (A) The ratio of fluorescence intensity (FI) to cell density (OD_{600}) was used as a proxy for intracellular GFP concentration. At 24 h post-induction, the normalized FI/OD_{600} decreases with both peptide charge density and peptide length. Data was normalized to GFP(0)-H_1 at 24 h as the maximum. Three biological replicates and the respective means are shown. Triangles, circles, and squares indicate isotropic variants, low-charge density peptide variants, and high-charge density peptide variants, respectively. (B) At 24 h post-induction, the total area fraction and median intensity ratio of condensates depend on the net charge and expression of the protein. Variants with high net charge and FI/OD_{600} tend to form larger and brighter condensates. Marker size and shading indicate the median total area fraction and median average intensity ratio in cells with condensates. Gray lines show the pairwise comparisons between peptides with different charge densities on the same globular domain. Gray circles mark the FI/OD_{600} for variants that do not form condensates at 24 h post-induction. (C) Variants with a net charge of −7 or 0 do not form condensates at 2, 8, or 24 h post-induction. Some variants with a net charge of +6 exhibit reversible condensate formation over time. The mean and a minimum of three biological replicates for each variant are shown. Four biological replicates are shown for sfGFP, GFP(−6)-H_2, GFP(−6)-L_3, GFP(0)-H_3, and GFP(+6); five biological replicates are shown for GFP(+6)-H_3; and eight biological replicates are shown for GFP(0)-L_3 and GFP(+6)-L_3. At least 83 cells were analyzed per biological replicate. Triangles, circles, and squares indicate isotropic variants, low-charge density peptide variants, and high-charge density peptide variants, respectively.

of globular proteins can enable phase separation with nucleic acids in bacteria via electrostatic interactions, even in the absence of a disordered domain.[31] It was subsequently shown that achieving this high net charge (>+6) via a charged peptide domain also resulted in protein condensation.[32] Building on that prior work, here, we engineered modular proteins with a globular domain and a C-terminal charge-patterned disordered peptide domain, given that many endogenous condensate scaffolds consist of IDRs tethered to folded functional domains. The use of minimal cationic peptides allowed us to preserve the relative and contextual importance of the globular domain in de novo condensates. We explored the interplay of contributions from both domains by systematically constructing a panel of proteins from three globular green fluorescent proteins (GFP) derived from superfolder GFP (sfGFP) with different surface charges, as has been done previously,[32] and, in this case, with two cationic disordered peptide motifs with varied charge-patterning.

Specifically, we explored the charge, charge-patterning, and concentration dependencies of intracellular phase separation by overexpressing this panel of engineered proteins in *E. coli* to form de novo condensates. We show that the formation of electrostatically driven condensates is primarily dependent on the net charge and in vivo concentration of the protein scaffold. Important sequence-specific determinants, such as amino acid composition, charge-patterning, and peptide length, altered the interaction strength and influenced protein expression levels. Refinement of reported image analysis approaches enabled analysis of subcellular condensate features and revealed that sequence-encoded features could tune mesoscale attributes, such as the size, number, and subcellular localization of condensates within *E. coli* cells, alluding to intricate structure−form−function relationships that underlie endogenous condensates. Taken together, the use of short disordered peptides is a versatile approach to expanding our fundamental understanding of the sequence parameters governing biomolecular condensates in bacteria.

RESULTS AND DISCUSSION

Design of Disordered Cationic Peptides with Varying Charge Densities. Along with an abundance of anionic biomacromolecules such as RNA and DNA, the *E. coli*

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

proteome is also composed of negatively charged proteins (Figure 1A). This allows us to leverage electrostatic attraction and entropic gains from counterion release to drive the partitioning of cationic proteins of interest in complex coacervates. It was previously demonstrated that the negative globular GFP could be supercharged with a C-terminal cationic peptide to form phase-separated intracellular condensates without the need to introduce additional anionic partners.[32] Recent mounting evidence has also revealed the importance of the arrangement of residues in the primary sequence in modulating the strength of the weak, multivalent interactions responsible for intracellular phase separation.[32−35] To explore the role of charge-patterning in the complex coacervation of proteins in cells, we designed two cationic peptides with differing charge densities (Figures 1B and S1). The low-charge density peptide (L_n) has the amino acid sequence of GRRGKKSRK, with neutral glycine and serine residues interspaced between every 2 cationic residues, while the high-charge density peptide (H_n) has the sequence KRRRKK and only consists of cationic residues. Repeats of the base peptide sequences were also studied to investigate the effects of increasing peptide length, which alters both the net protein charge and the balance of charge between the disordered peptide domain and the globular protein domain. Here, we also probed the relative importance of electrostatic interactions from both folded and disordered domains by appending the low- and high-charge density cationic peptides to engineered GFP variants based on sfGFP, with anionic [GFP(−6)], neutral [GFP(0)], or cationic [GFP(+6)] net charge. In addition to the presence of phase separation within a cell population, we aimed to elucidate the parameters governing the number, size, and protein partitioning of engineered condensates (Figure 1C). As a more conservative estimate of protein partitioning, we calculated the average fluorescence intensity of condensates relative to the cytoplasm rather than comparing the maximum intensity within condensates to the cytoplasm. We hypothesized that both the condensate size and strength of protein partitioning would depend on the net charge and the intracellular protein concentration. We also postulated that proteins with equivalent net charge could have different propensities for homotypic and heterotypic interactions based on whether the charge resides primarily on the folded domain or the disordered domain.

Role of Charge-Patterning in Promoting Bacterial Condensate Formation. In cases in which the protein interaction strength is sufficient for associative heterotypic phase separation, condensates should form if cellular levels of the protein are greater than the saturation concentration, c_{sat}. Given the hypothesized importance of the intracellular protein concentration in enabling phase separation, the fluorescence intensity normalized by the cell density (FI/OD_{600}) was used as a proxy to estimate the mean cell protein concentration. At 24 h post-induction, both the OD_{600} and the FI/OD_{600} were found to decrease with increasing peptide charge density and length, suggesting that long charge-dense disordered peptides have a significant effect on the host cell growth and protein expression levels (Figures 2A and S2A,B).

To investigate the effects of charge-patterning on intracellular phase separation facilitated by these short, disordered peptides, each GFP variant was expressed in *E. coli* cells and monitored by fluorescence microscopy for the presence of bacterial condensates (Figure 1D). Consistent with previous observations, the net charge was found to be a main governing parameter for condensate formation, and all variants with net charge ≥+12 showed observable condensates at 24 h post-induction.[32] Interestingly, GFP(−6)-L_2 also demonstrated intracellular phase separation at 24 h post-induction despite having a net charge of +6. The existence of a critical net charge for phase separation was also evident in in vitro turbidity assays using purified protein and total RNA from torula yeast (Figure S2F,G). Analysis of the size of condensates and the protein partitioning in condensates according to the net charge and intracellular protein concentration revealed only minor differences with respect to the peptide sequence at 24 h post-induction for variants with the GFP(+6) domain (Figure 2B). This suggests that small sequence modifications on the disordered peptide domain have minor effects at long expression times if the globular domain also facilitates phase separation.

Although differences due to the charge density of the peptide were largely overshadowed when appended to a positive globular domain, the charge-patterning of the peptide had noticeable effects when tethered to negative or neutral globular domains. Notably, there were differences in the size and fluorescence intensities between GFP(−6)-L_3 and GFP(−6)-H_3 condensates, as well as between GFP(0)-L_2 and GFP(0)-H_2 condensates. As neither the GFP(−7) nor the GFP(0) globular domains phase separated without disordered cationic peptides, the peptide was likely the main driver of phase separation for these proteins by facilitating multivalent interactions with anionic macromolecules. Consequently, the primary sequence of the peptide is particularly important for negative or neutral globular domains. In contrast, the peptide charge-patterning is less important when both the globular and peptide domains have the potential for complex coacervation, as with the GFP(+6) variants fused to cationic peptides. Irrespective of the peptide charge density, GFP(+6) protein variants formed larger and brighter condensates at 24 h post-induction when compared to GFP(−6) or GFP(0) variants with equivalent net charge.

Quantitative image analysis further revealed that the fraction of condensate-forming and noncondensate-forming cells changes in the late-log, early stationary, and late-stationary phases of cell growth at 2, 8, and 24 h post-induction, respectively (Figure 2C). Variants with high-charge density peptides at the net charge threshold of +6 formed reversible condensates, whereby the fraction of cells with condensates increased from 2 to 8 h and decreased from 8 to 24 h post-induction. Condensate reversibility was previously demonstrated with the isotropic GFP(+6) variant, and these findings provide additional evidence that anisotropic proteins with disordered peptides can also undergo similar reversible condensate formation. In general, longer peptide lengths appeared to enable a moderate fraction of the cell population to form condensates even as early as 2 h post-induction, as with GFP(−6)-L_2, GFP(−6)-L_3, GFP(−6)-H_3, GFP(0)-H_2, GFP(+6)-L_2, and GFP(+6)-H_2. We hypothesized that this may be due to an increase in the protein charge patchiness and the radii over which the charge extends. For increasingly longer peptide lengths, the reduction in protein expression likely counteracted the increase in the interaction affinity. This is shown for GFP(0) and GFP(+6) with 3 repeated motifs, which have lower propensity to form condensates compared to respective variants with 2 repeated motifs. Among the complete charge matrix of variants that were able to form condensates, we chose to examine variants with net charges of

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

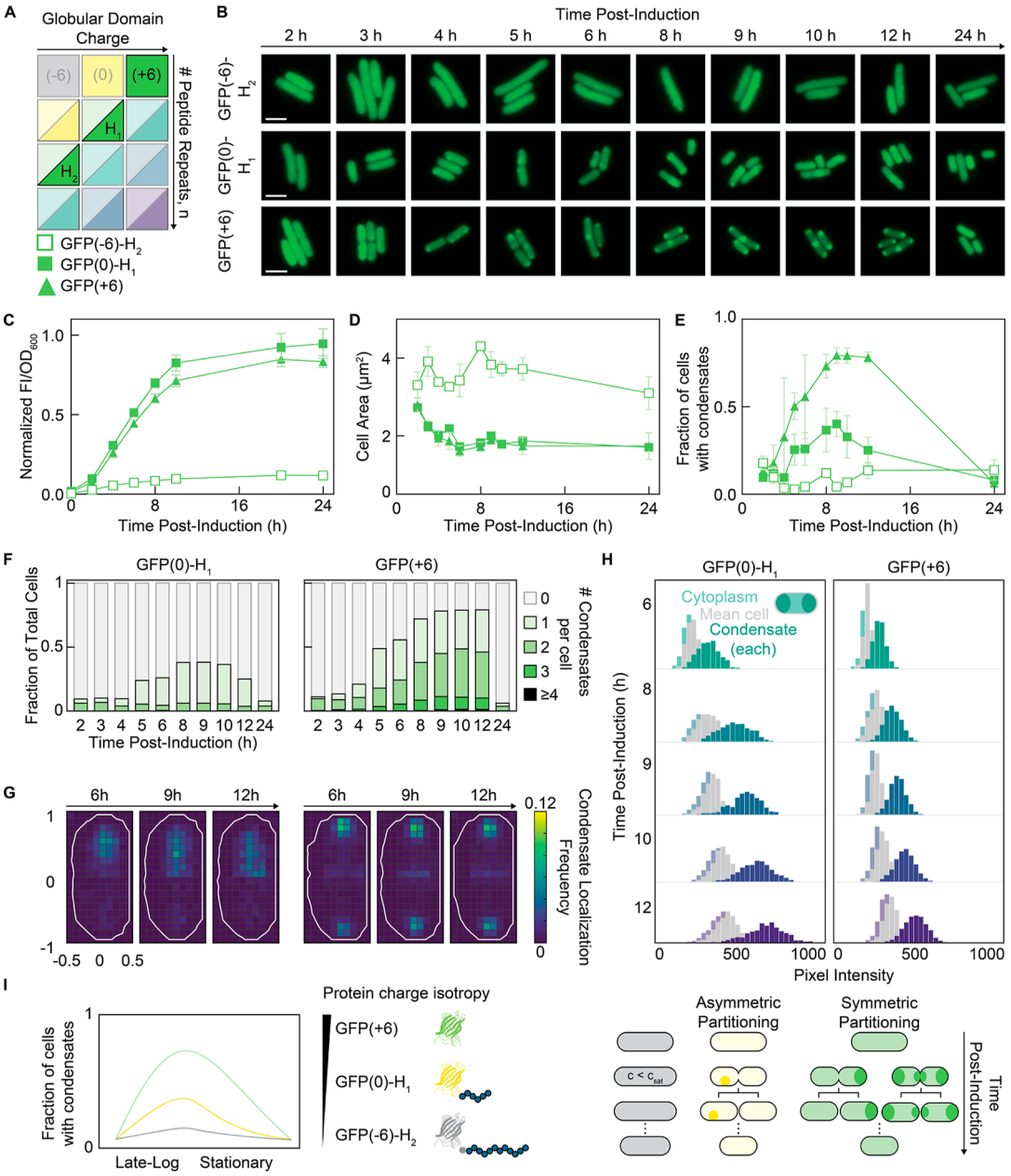


Figure 3. Variants with a net charge of +6 are at the net charge threshold for condensate formation and exhibit condensate reversibility at the population level over time. The spatiotemporal properties of condensates vary with the charge distribution between the globular protein and disordered peptide domains. (A) Schematic of the experimental charge matrix with the +6 net charge high-charge density peptide variants highlighted. (B) Representative time-course microscopy images of cells from 2 to 24 h post-induction for GFP(−6)-H_2, GFP(0)-H_1, and GFP(+6). All three variants show the formation and dissolution of condensates over time. Scale bars are 2 μm. (C) Normalized FI/OD_{600} over time for GFP(−6)-H_2, GFP(0)-H_1, and GFP(+6). Data points and error bars indicate the mean and standard deviation of the three biological replicates. Lines connecting the means of three biological replicates are shown. (D) Cells expressing GFP(−6)-H_2 are approximately twice as large as cells expressing GFP(0)-H_1 or GFP(+6). The mean and standard deviation of three biological replicates are shown. (E) The peak fraction of cells with condensates is higher, and the duration of condensate formation increases for more isotropic charge distributions. At least 134 cells were analyzed per replicate and time point. The mean and standard deviation of three biological replicates are shown. (F) Among condensate-forming cells, a majority of GFP(0)-H_1 cells form 1 condensate per cell, while a majority of GFP(+6) cells form ≥2 condensates per cell. Data from all biological replicates were consolidated. (G) As shown by binned, normalized, and aligned 1-pixel local maxima heat maps of all condensate-forming cells, GFP(0)-H_1 condensates (left) do not have a strong spatial localization preference, while GFP(+6) condensates (right) preferentially localize at the cell poles in shorter cells and at the cell poles and midcell in longer cells. Outlines depict an approximate cell contour. (H) The pixel intensity distribution of GFP(0)-H_1 condensates shows a larger spread over time when compared to that of GFP(+6) condensates, suggesting greater cell-to-cell variability of condensates when the short peptide domain is the primary driver of phase separation. Data is shown for the condensate-forming subpopulation of consolidated biological replicates with a histogram bin width of 30. (I) Volume and shape changes during the cell cycle may influence the formation and dissolution of condensates by concentrating or diluting the protein within the cell. The initial cell size increase of GFP(−6)-H_2 may hinder phase separation by decreasing the protein concentration below the saturation concentration. The number and localization of condensates determine whether condensates will be asymmetrically or symmetrically inherited from the mother cell by the daughter cell.

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

+6 and +12 in more detail to understand the effects of charge distribution and peptide charge density at and above the charge threshold for intracellular phase separation in *E. coli*.

Reversible Condensate Formation at the Charge Threshold Is Affected by Charge Isotropy. It was previously observed that at a net charge of +6, both isotropic variants and variants with a cationic disordered peptide had the propensity to undergo reversible phase separation. This behavior was shown to correlate with the growth phases, with condensates forming during the late-log phase and disassembling during the stationary phase. Here, fluorescence microscopy revealed similar condensate reversibility at the bulk population level among the variants with high-charge density peptide domains (Figure 3A,B). The fraction of cells with condensates increased throughout the late log-phase until 8 or 9 h post-induction (Figure 3E). The maximum fraction of cells with condensates differed among the variants according to the protein charge distribution. A very low fraction of cells expressing GFP(−6)-H_2 formed condensates by 8 h post-induction, while around half of all cells expressing GFP(0)-H_1 formed condensates by 9 h, and nearly all cells expressing isotropic GFP(+6) formed condensates by 8 h. The lower prevalence of condensates among cells expressing GFP(−6)-H_2 may be attributed to significantly lower protein expression compared to cells expressing GFP(0)-H_1 and GFP(+6) (Figures 3C and S3). Interestingly, cells expressing GFP(−6)-H_2 initially increased in cell area during the late-log phase and were approximately twice as long as cells with GFP(0)-H_1 or GFP(+6) in the stationary phase (Figure 3D). As the late-log phase was shown to be a critical period for initial condensate formation, we hypothesized that the dilution of the protein via cell elongation could have further prevented the formation of GFP(−6)-H_2 condensates.

To examine this further, we characterized condensate properties such as number, localization, and partitioning for GFP(0)-H_1 and GFP(+6). These two variants had very similar protein expression levels and cell areas over time, enabling a direct comparison with fewer confounding variables such as intracellular protein concentration (Figure 3C,D). Interestingly, among condensate-forming cells, GFP(0)-H_1 preferentially formed one condensate per cell, while GFP(+6) had a tendency to form two or more condensates per cell (Figure 3F). Given the similar intracellular protein concentrations between these two variants, the greater number of condensates per cell for GFP(+6) may be attributed to more potential nucleation sites and a wider phase separation window. It is known that a higher connectivity between multivalent molecules is key to network formation and phase separation.[33,34] A positive isotropic globular domain potentially provides more discrete nodes to initiate and support system-spanning networks compared to a short, localized, charge-dense peptide domain. However, both of these proteins also contained a His6 peptide tag at the N-terminus that could potentially account for differences between the two proteins. While the His tag is unlikely to be positively charged at physiological pH, the protonation state may vary between the two proteins and could vary as a function of the growth phase. In addition to differences in the number of condensates per cell, GFP(0)-H_1 condensates also did not exhibit preferential subcellular localization within a given cell compared to GFP(+6), which strongly localized to both cell poles. In longer cells, GFP(+6) also had a strong preference for a third foci localization at the midcell (Figure 3G).

The number and localization of condensates can influence the inheritance of condensates from mother cells to daughter cells, which dictates the persistence of condensate-forming cells in a population. In prior work, the selective spatial localization of biomolecules has been shown to modulate the differential cell fate of daughter cells and population-level cellular behaviors.[24,30,35,36] For example, engineering the asymmetric segregation of plasmid DNA to a single position within *E. coli* can lead to cells with distinct differentiated states.[35] Similarly, we observed that single polarly localized condensates were typically asymmetrically inherited during cell division, whereby only one of the daughter cells retained the condensate (Figure 3I). This limited the maximum fraction of condensate-containing cells in the population to roughly half of the cell population, as with GFP(0)-H_1 at 8 h post-induction. Conversely, cells expressing GFP(+6) preferentially formed condensates at both cell poles, resulting in a symmetric inheritance. It was noted that dividing cells expressing GFP(+6) also had a tendency to form a third condensate at the midcell position in addition to condensates at the poles. In this case, both daughter cells could inherit two condensates. Regardless, GFP(+6) condensates were both more prevalent and persistent in the cell population, suggesting that simply altering the distribution of the net charge of the protein scaffold could have broader population-level consequences.

In addition to differences in the number and localization of condensates, the degree of protein partitioning between the condensed phase and the surrounding cytoplasm was also investigated for the GFP(0)-H_1 and GFP(+6) variants. In agreement with the FI/OD_{600} data, the mean cell fluorescence, as calculated by the total cell fluorescence normalized by the cell area, increased slightly from 6 to 12 h post-induction (Figure 3H). The distributions of the mean cytoplasm intensity and mean condensate intensity also increased as a function of this. Although the average intensity ratio, comparing the condensate intensity to the cytoplasm intensity, did not change significantly from 6 to 12 h post-induction, being ∼1.6−1.7 for GFP(0)-H_1 and ∼1.5−1.6 for GFP(+6), greater cell-to-cell variability was seen in the broadening condensate intensity distribution of GFP(0)-H_1. This suggests that the earlier disassembly of GFP(0)-H_1 condensates is a result of exiting a narrower phase separation window.

Overall, the results suggest that even at similar protein expression levels, the phase separation behavior of a protein depends on whether the charge is isotropically distributed on the globular domain or resides entirely on a highly charge-dense peptide domain. Growth phase-specific phase separation implies that changes in the biomolecule concentrations, metabolism-dependent material properties, and overall chemical environment within the cell cytoplasm can significantly influence the phase separation of cationic proteins.[37−39] In a batch culture, the total RNA-to-protein ratio has been shown to decrease with lower growth rates in the nutrient-limited stationary phase, while ribosomes become inactive and undergo ribosome hibernation.[37,40] It is well known that many other biophysical changes accompany the transition from the log-phase to the stationary phase of cell growth, including cell volume decrease, cell shape change, nucleoid compaction, and cell wall thickening.[41,42] On a single-cell level, proteins may also be diluted by cell elongation or concentrated by cell division as cells undergo rounds of cell division. All of these growth phase-dependent processes can influence the phase separation of proteins.

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

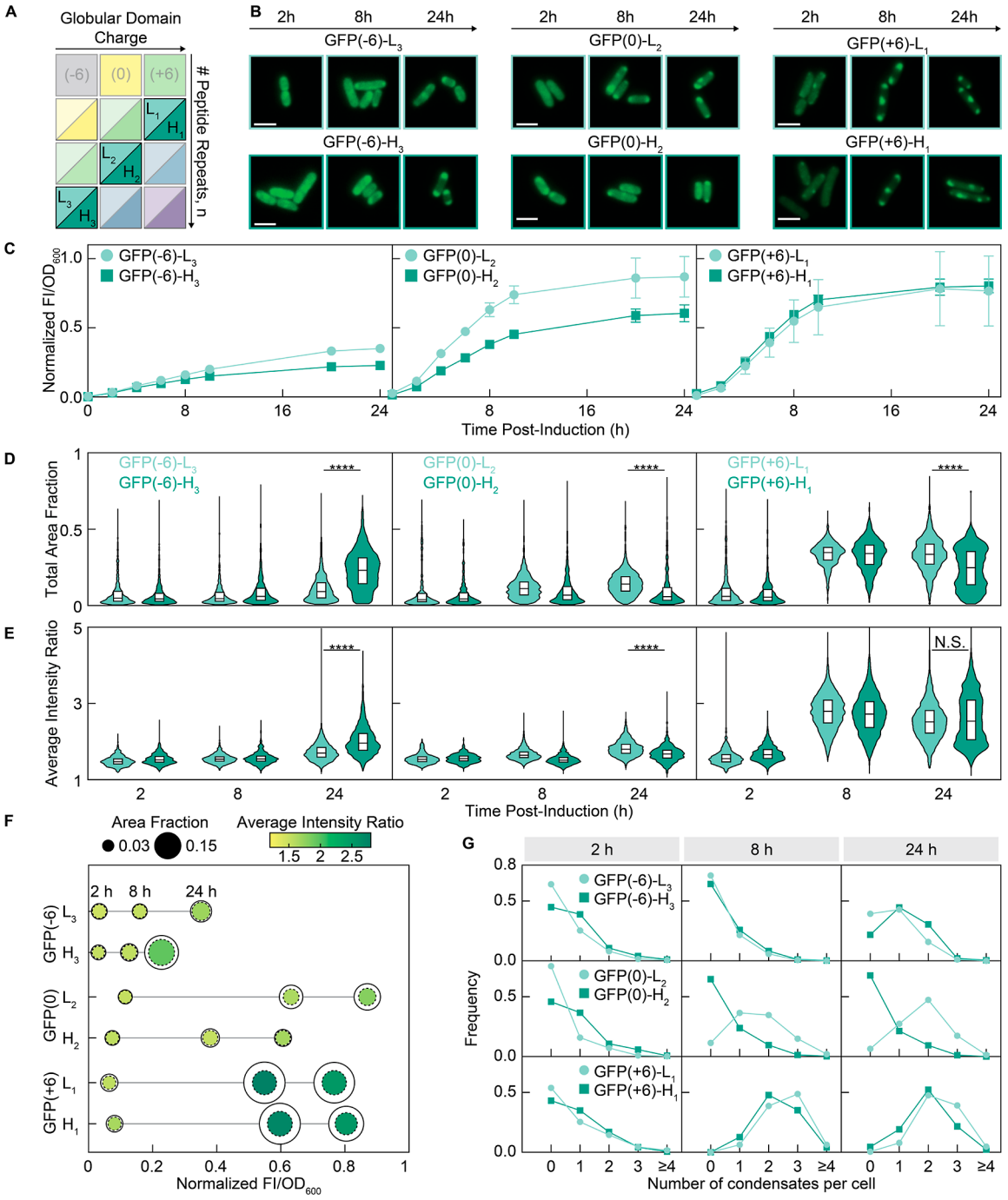


Figure 4. For variants with a net charge of +12, the peptide charge density and charge distribution between domains affect interaction strength and protein expression and, consequently, the condensate area, intensity, and number of condensates per cell. (A) Schematic of the experimental charge matrix with the +12 net charge variants highlighted. (B) Representative microscopy images of cells at 2, 8, and 24 h post-induction for each variant. All variants with net charge ≥+12 still have cells with foci at 24 h post-induction. Scale bars are 2 μm. (C) For negative GFP(−6) and neutral GFP(0) globular domains, appending a high-charge density polypeptide results in lower normalized FI/OD_{600}. Data was normalized to GFP(0)-L_2 at 24 h post-induction as the maximum. Lines connecting the means of three biological replicates (data points) are shown. (D) The total area fraction of condensates increases over time. The total area fractions of condensates formed with proteins that have cationic charge on both the globular and disordered peptide domains are larger than those for protein variants that only have cationic charge on the peptide domain. (E) The intensity ratio of condensates was used as a proxy for the degree of protein partitioning. On a negative globular domain, a high-charge density peptide results in more protein partitioning in the condensate. The converse is true on a neutral domain, where a low-charge density peptide led to more protein partitioning. On a positive globular domain, the partitioning is less affected by the charge density of the disordered peptide. GFP(+6) variants show higher protein partitioning. For (D,E), data from condensate-forming cells from all biological replicates were consolidated and each condensate was analyzed separately. At least 725 condensates were analyzed for each distribution. L peptide variants are light teal and H peptide variants are dark teal. Differences between variants at 24 h post-induction were evaluated by the Mann−Whitney test, with **** indicating $p <$ 0.0001. (F) Summary data of the median area fractions and median intensity ratios of condensates as a function of the normalized FI/OD_{600} at 2, 8, and 24 h post-induction. The lower FI/OD_{600} of GFP(−6) variants may explain the smaller total area fraction and lower intensity ratios of the condensates at 8 h post-induction. Although the FI/OD_{600} of GFP(0) variants is similar to that of GFP(+6) variants, a neutral globular domain may

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

Figure 4. continued

have lower interaction strength compared to a positive globular domain, resulting in smaller and less bright condensates overall. Markers for each variant are in the order of time post-induction (2, 8, and 24 h). Dashed circles indicate the median area fraction of individual condensates, and the solid circles indicate the median total area fraction of all condensates in a given cell. Although the total area fraction increases over time, the area fraction of individual condensates appears to have an upper bound. (G) The FI/OD_{600} may affect the number of condensates per cell. GFP(−6) variants have lower FI/OD_{600} at 24 h post-induction, and the majority of condensate-forming cells have 1 condensate. GFP(0)-L_2 has slightly higher FI/OD_{600} at 24 h post-induction and preferentially contains 2 condensates per cell, while GFP(0)-H_2 has lower FI/OD_{600} and a majority do not form condensates. GFP(+6) variants have sufficient interaction strength and FI/OD_{600} to preferentially form 2 or 3 condensates per cell. Data from all biological replicates were consolidated.

Protein Scaffolds with Higher Net Charge Maintain Condensates at Later Time Points. In addition to protein scaffolds with a net charge of +6, protein variants with a higher net charge of +12 and longer disordered peptide domains were also studied to elucidate the effects of charge and charge distribution on condensate properties (Figure 4A). Microscopy images revealed that condensates were present at 2, 8, and 24 h post-induction (Figure 4B). In contrast with most +6 variants that did not have condensates at 24 h post-induction, all +12 variants still had a moderate-to-high fraction of cells with condensates at longer time points, which may be attributed to greater overall net charge or longer peptide lengths, resulting in a broader two-phase region.

The +12 variants could be categorized by the approximate intracellular protein concentration (Figures 4C and S4). Specifically, GFP(−6)-L_3, GFP(−6)-H_3, and GFP(0)-H_2 all had relatively lower protein expression levels. In the case of the GFP(−6)-L_3/H_3 variants, this was likely due to the difficulty in expressing highly anisotropic proteins with long stretches of cationic residues. In the case of GFP(0)-H_2, low expression may be due to the moderately long but highly charge dense H_2 domain as a similar reduction in expression was seen previously with GFP(−6)-H_2 (Figure 3C). The remaining variants, GFP(0)-L_2, GFP(+6)-L_1, and GFP(+6)-H_1, did not exhibit significant reduction in protein expression when compared to the respective unmodified globular domains at 2, 8, or 24 h post-induction (Figure S2A).

Several condensate properties were consistent with FI/OD_{600} or intracellular protein concentration. Variants with a higher FI/OD_{600} had condensates at 24 h post-induction that altogether occupied a greater fraction of the cell (Figure 4D). Interestingly, although the total area fraction of condensates increased substantially over time for variants with higher FI/OD_{600}, the area fraction of individual condensates did not increase significantly (Figure 4F). Instead, the increase in the total area fraction occupied by condensates was a result of more condensates being formed per cell (Figure 4G). Using the total area fraction as a proxy for the total volume fraction of condensates, this suggests that although there is a driving force for increasing the total volume fraction of condensates, steric occlusion or other effects may prevent the formation of larger individual condensates or the fusion of multiple condensates.

The average partitioning of proteins in condensates generally increased over time for all variants (Figure 4E). The effects of the peptide charge density were most noticeable on the negative GFP globular domain, with the high-charge density H_3 peptide leading to more protein partitioning in the condensed phase. The converse was true on a neutral domain, where a low-charge density peptide led to more protein partitioning. On a positive globular domain, the average intensity ratios were less impacted by the charge density of the peptide. This supports the idea of a hierarchical relationship among the different sequence parameters, where the overall net charge and globular domain charge are important when the peptide length is short, while the peptide charge density is more relevant when the peptide is adequately long and serves as the primary interaction domain.

Modular Disordered Cationic Peptides Sequester Multiple Proteins in Engineered Condensates. In addition to encapsulating a single protein of interest in phase-separated condensates, the modular nature of the cationic peptides can be leveraged to colocalize multiple proteins within the same condensed phase. This allowed us to explore the sequence-determinants of multiprotein condensates that are more representative of endogenous systems. Here, we demonstrate that multiprotein condensates consisting of GFP and another globular fluorescent protein, mCherry, can be engineered in *E. coli* cells with the low- and high-charge density peptide domains. Specifically, we coexpressed GFP(−6)-H_2 and mCherry(−6)-L_n, where n represents the number of peptide motif repeats as before (Figure 5A). As a control, we first expressed mCherry(−6)-L_n alone and confirmed that like GFP(−6)-L_n, a net charge greater than +6 was required for mCherry condensate formation (Figures S5E and 2C). Minor differences in the phase separation behavior between GFP(−6)-L_2/L_3 and mCherry(−6)-L_2/L_3 may be attributed to differences in the protein expression or protein maturation times. Additionally, sfGFP-derived GFP variants can weakly dimerize and may contribute to the overall valency compared to mCherry, which remains monomeric.

We further hypothesized that the characterized phase separation behavior of GFP and mCherry would largely remain the same even if they were present in multiprotein assemblies (Figure 5B). As such, we expected GFP(−6)-H_2 to still undergo reversible phase separation over time when coexpressed (Figure 2C). Consistent with our expectations, GFP(−6)-H_2 formed reversible condensates regardless of the net charge of mCherry, although a much higher fraction of cells contained GFP(−6)-H_2 condensates at 8 h post-induction (Figures 2C and 5C). Given that RNA length has been shown to modulate condensates,[43,44] a potential explanation may be that the longer mRNA transcripts can facilitate more associative interactions with GFP(−6)-H_2 to promote phase separation. In the case of coexpression with mCherry(−6)-L_2 or mCherry(−6)-L_3, the synergistic effects of expressing two proteins that are independently capable of phase separation may also help promote GFP(−6)-H_2 condensation.

Similarly, mCherry(−6), mCherry(−6)-L_1, and mCherry(−6)-L_2 also behaved in a predictable manner when coexpressed with GFP(−6)-H_2 (Figures S5E and 5D). In the case of mCherry(−6) and mCherry(−6)-L_1, neither protein had sufficient charge to phase separate when expressed alone or when coexpressed with GFP(−6)-H_2, as shown by the average intensity ratio along the medial axis of the cell (Figure

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

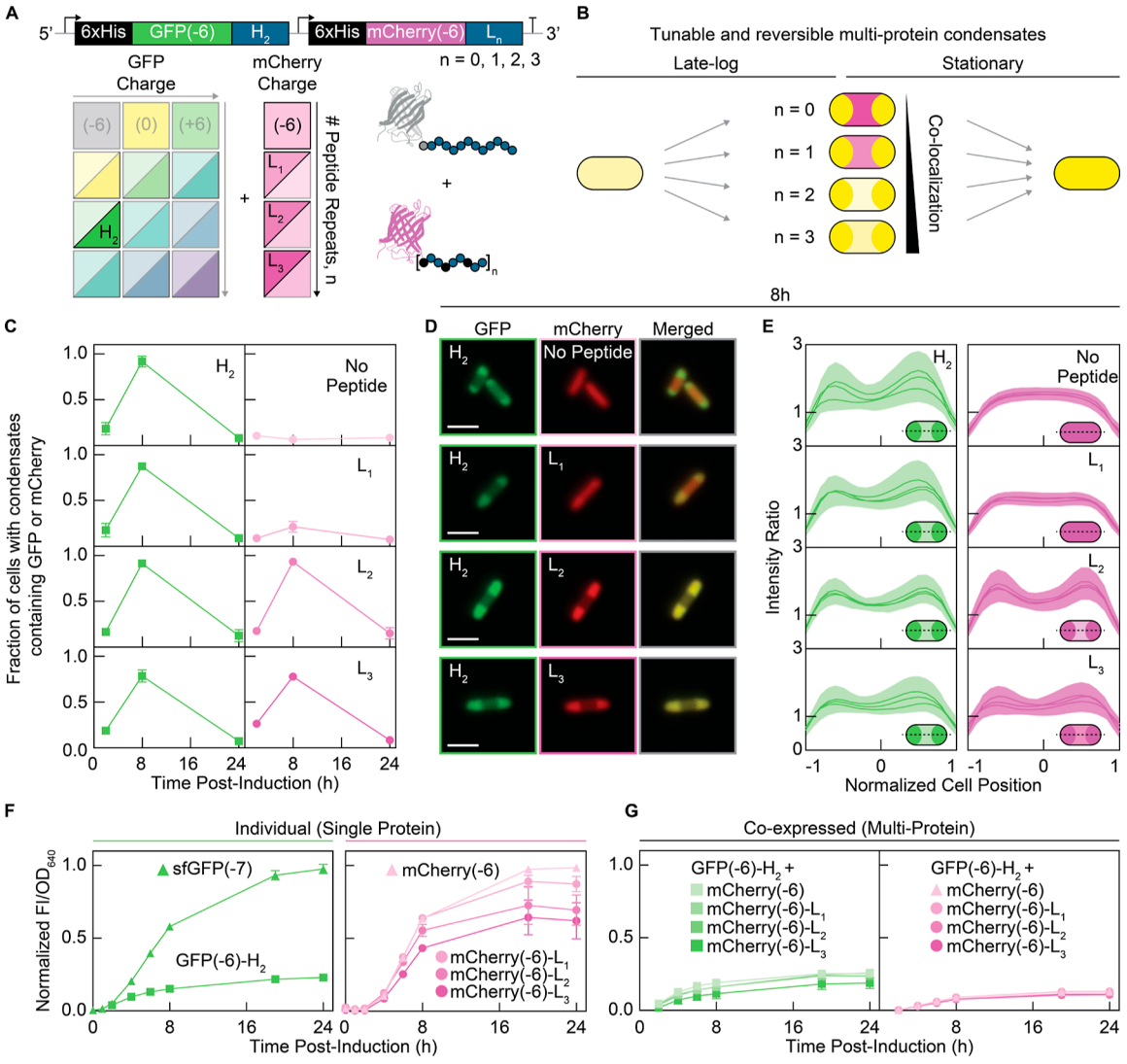


Figure 5. Modular disordered cationic peptides can be used to engineer reversible multiprotein condensates with GFP and mCherry. The degree of mCherry partitioning can be tuned with the length of the peptide. (A) GFP(−6)-H_2 and mCherry-L_n were coexpressed from a single transcript by separate T7 promoters and ribosome binding sites. Experimental charge matrix shows the peptides used for each protein. (B) Schematic of the reversibility and tunability of multiprotein GFP and mCherry condensates. (C) The fraction of cells with condensates containing GFP(−6)-H_2 is similar despite different coexpression partners, suggesting that GFP(−6)-H_2 acts as a scaffold. The fraction of cells with condensates containing mCherry-L_n is tunable and increases with the number of peptide motif repeats, *n*, at 8 h post-induction. Microscopy images from each fluorescence channel were analyzed independently. The fraction of cells with condensates from each channel was calculated separately. Mean and standard deviation of three biological replicates are shown. (D) Representative microscopy images at 8 h post-induction when nearly all cells form condensates with at least one protein. GFP(−6)-H_2 and mCherry-L_n show stronger colocalization with increasing cationic peptide length on mCherry. Scale bar is 2 μm. (E) The intensity ratio along the medial axis shows preferential localization of GFP(−6)-H_2 at the cell poles. mCherry-L_n colocalizes with GFP(−6)-H_2 when it is charge-equivalent or has slightly higher charge than GFP(−6)-H_2. The intensity ratios along the medial axis were normalized to the cell position and averaged. For cells containing one condensate, the medial profile was aligned with the condensate on the positive end of the medial axis. Lines and shading indicate the mean and standard deviation of three biological replicates. Microscopy images from each fluorescence channel were analyzed independently. (F) FI/OD_{640} for cells expressing a single protein show that the high-charge density H_2 peptide significantly reduced the protein expression of GFP(−6)-H_2 compared to sfGFP. Increasing repeats of the low-charge density peptide sequence moderately reduced the expression of mCherry-L_n when compared to mCherry. Data for GFP and mCherry variants were normalized to sfGFP(−7) at 24 h and mCherry(−6) at 24 h as the maximum. The mean and standard deviation for three biological replicates are shown. GFP and mCherry fluorescence were measured with an excitation/emission of 470/510 nm and 570/610 nm, respectively. OD was measured at 640 nm. (G) When coexpressed, both GFP(−6)-H_2 and mCherry-L_n show similar expression levels to that of GFP(−6)-H_2 when expressed alone. Data for GFP and mCherry variants were normalized to sfGFP(−7) at 24 h and mCherry(−6) at 24 h as the maximum. The mean and standard deviation for three biological replicates are shown.

5E). Coexpression of charge-equivalent mCherry(−6)-L_2 and GFP(−6)-H_2 revealed that mCherry(−6)-L_2 behaved nearly identically to GFP(−6)-H_2 in terms of condensate prevalence among the cell population, localization to the cell poles, and formation and disassembly over time. This likely indicates that the two proteins are forming a single condensed phase that contains both proteins; however, due to the resolution of traditional fluorescence microscopy, we cannot rule out the possibility that the two proteins form heterogeneous, demixed condensates that colocalize at the cell poles. In support of a

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

single, well-mixed condensed phase, in vitro experiments with GFP(−6)-H_2 and mCherry(−6)-L_2 also showed that the two proteins formed homogeneous liquid-like droplets with RNA at varying charge fractions (Figure S6). Interestingly, coexpression with GFP(−6)-H_2 appeared to slightly modify the phase separation behavior of mCherry(−6)-L_3. Although when expressed alone, mCherry(−6)-L_3 condensates were present after 24 h post-induction, mCherry(−6)-L_3 condensates seemed to disassemble earlier along with GFP(−6)-H_2 condensates when coexpressed. Taken together, both the invariable phase behavior of GFP(−6)-H_2 despite different coexpression partners and the modulating effect it has on the phase behavior of mCherry(−6)-L_3 suggest that it acts as a scaffold, potentially due to the earlier upstream expression of GFP(−6)-H_2, the higher charge density of the H_2 peptide domain, or a slightly higher expression level (Figure 5G).

CONCLUSIONS

We previously demonstrated that condensate formation in *E. coli* is dependent on the overall protein charge and that appending repeats of the short peptide sequence, GSSKKRKKR, onto various protein globular domains promoted the protein phase separation. Here, we expand upon our previous study and investigate additional cationic peptide sequences of varying charge densities. Using quantitative image analysis, we obtained further information about mesoscale features, such as the number, size, localization, and partitioning in synthetic condensates. Additionally, we showed that these design principles can largely be extended to engineer the colocalization of multiple proteins in synthetic condensates.

Consistent with previous results, the overall net charge and intracellular protein concentration were shown to be important parameters governing the heterotypic phase separation of cationic protein scaffolds with anionic biomolecules in the cell. Proteins with negative or neutral net charge did not phase separate in vivo, while most proteins with a net charge ≥+6 had sufficient valence to form condensates. At the phase separation net charge threshold of +6, variants had the propensity to undergo reversible phase separation at the bulk population level, whereby condensates formed during the late-log phase and disassembled in the stationary phase of cell growth. Protein variants with a higher net charge of +12 formed condensates that could be maintained at longer time points, corresponding to phase separation over a broader range of protein concentrations. In the context of synthetic biology, these short cationic peptides can be easily appended to other globular proteins to control not only condensate formation but also the reversibility or persistence of condensates in different growth phases.

Furthermore, we found that the patterning of charged residues in the disordered peptide domain had a significant effect on protein expression and a more charge-dense motif noticeably reduced the intracellular protein concentration. When the disordered peptide domain was the sole contributor of the net cationic charge, as with GFP(−6)-L_3 and GFP(−6)-H_3, the peptide charge density had a noticeable effect on the size and partitioning of proteins in the condensates. Specifically, the higher-charge density KRRRKK peptide motif led to slightly larger condensates and higher protein partitioning than that of the GRRGKKSRK peptide motif. Taken together, this suggests that when the peptide domain is the primary driver of phase separation, highly charge-dense sequences may increase the affinity of the protein toward anionic biomolecules to facilitate phase separation. However, this is balanced by a counteracting effect whereby highly charge-dense sequences negatively impact cell physiology and protein expression levels and consequently impede phase separation. This suggests that although highly charge-dense peptides can be used to provide additional interaction strength for the phase separation of negative proteins, the metabolic burden of expressing the engineered protein is a physiological constraint that must be considered.

Although understanding the design parameters governing the intracellular phase separation of a single protein is invaluable, endogenous biomolecular condensates in bacteria are multicomponent systems consisting of numerous proteins and nucleic acids. As such, we also used the modular nature of cationic peptides to elucidate the sequence-determinants of the phase behavior and partitioning of two globular proteins, GFP and mCherry, in multiprotein systems. We specifically studied cases in which the cationic peptide was the main driver of phase separation by coexpressing GFP(−6)-H_2 with mCherry(−6)-L_n. We varied the valence or net charge of mCherry by changing the length of the attached disordered peptide. Consistent with our previous observations, the net charge dictated the degree of mCherry colocalization with GFP(−6)-H_2 in condensates. When both proteins were equivalent in net charge, they exhibited almost identical partitioning, localization, and condensate prevalence among a given cell population. Other properties such as the overall reversibility of the multiprotein condensate could be attributed to GFP(−6)-H_2, which demonstrated condensate reversibility in single-protein condensates, although to a lesser extent. Overall, the same sequence-determinants governing single-protein systems can be extended to multiprotein systems, laying the groundwork to use cationic peptides as a tool to recruit multiple enzymes within a pathway for metabolic engineering in synthetic condensates. When the use of these cationic peptides for applications in synthetic biology is designed, it is important to note that the peptides can have some modest impacts on cell growth (Figure S2A). This is in line with recent reports of the cellular toxicity of cationic peptides and proteins.[45,46] Interestingly, the peptide sequences studied here also bear a close resemblance with antimicrobial peptides, which are typically cationic and 10−25 amino acid-long. Such peptides have been shown to inhibit transcription and translation in vitro and form condensate-like clusters with nucleic acids in bacteria.[46]

Given that proteins implicated in endogenous bacterial condensates, such as *C. crescentus* PopZ and RNase E, often have modular architectures consisting of folded globular domains tethered to hypervariable disordered regions, it is crucial to understand the effects of valence contributions from both domains when designing synthetic systems.[15] In particular, blocky charge patterns in the disordered C-terminal domain of RNase E appear to be key in enhancing driving forces for phase separation.[47] Our results also show that the patterning of the charge on the disordered domain can affect the size of condensates and protein partitioning within the condensed phase. However, it should be noted that the disordered domains of PopZ and RNase E are significantly longer, with 75 and 450 amino acids, respectively, which may alter the relative role of the disordered domain. Similarly, other designed IDPs that promote phase separation in bacteria have typically been longer, ranging from 160 to 640 amino acids in

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

length, such that contributions from the globular domain were largely ignored.[27−30] However, we and others have shown that short terminal disordered peptides can still affect intracellular protein phase separation in a broad context. For example, commonly used epitope tags such as FLAG, HA, and Myc tags were shown to alter the phase separation propensity of proteins when placed on the N-termini, whereby tyrosine residues promoted phase separation, while negative residues inhibited phase separation.[48] We envision that a comprehensive understanding of sequence-determinants such as the patterning and strength of sticker-type interactions, such as the cationic residues studied here, on globular and disordered domains will help shed light on the functional significance of bacterial condensates in key cellular processes.

METHODS

Material Availability. Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Allie Obermeyer (aco2134@columbia.edu). Plasmids generated in this study have been deposited to Addgene (#205027−#205052). Raw microscopy images are deposited on Figshare (doi: 10.6084/m9.figshare.23289650), and custom MATLAB codes are publicly available on GitHub (https://github.com/Obermeyer-Group) upon publication.

Plasmid Construction. The two different charge density cationic peptides were designed by repeating the base peptide motif, GRRGKKSRK or KRRRKK, 1−3 times for the L and H peptides, respectively. These peptides were appended to sfGFP variants or mCherry using a combination of PCR, restriction enzyme digestion, and T4 ligation. An additional lysine residue was included between the C-terminus of sfGFP and the disordered peptide domain to establish charge equivalence between sfGFP and mCherry. All variants contained an N-terminal 6× His tag with either a GGA or GG linker for sfGFP and mCherry, respectively. Primers (Integrated DNA Technologies) with engineered restriction enzyme sites were used to clone cationic peptides onto the C-terminus of sfGFP. For the sfGFP variants, the forward and reverse primers contained NcoI and XhoI cut sites, respectively. For the mCherry variants, the forward and reverse primers contained XbaI and XhoI cut sites, respectively. Primers were diluted to 10 μM in Milli-Q water. PCR reactions were performed using Phusion polymerase as instructed by New England Biolabs (NEB). Purified PCR products and the pETDuet vector backbone were digested with restriction enzymes, NcoI and XhoI for the sfGFP variants or XbaI and XhoI for the mCherry variants. DpnI was added to the digestion reaction of the PCR amplified sfGFP inserts. Reactions were run on a 1% agarose gel (TopVision agarose), and bands were excised and purified using a QIAquick gel extraction kit (Qiagen). The final plasmid was assembled using T4 ligase (NEB) using a 5:1 molar ratio of insert to vector for the GFP variants and a 3:1 molar ratio of insert to vector for the mCherry variants. 2 μL of the ligation reaction was used to transform chemically competent NEB5α cells. Following sequence verification by Genewiz, the plasmids were transformed into NiCo21(DE3) cells (NEB). The bicistronic GFP(−6)-H_2+mCherry-L_n variant was constructed using HiFi Assembly. Briefly, PCR reactions were performed to add overlapping regions. PCR products were purified using a QIAquick PCR purification kit (Qiagen). Purified fragments were assembled using NEBuilder HiFi DNA Assembly Master Mix and annealed at 50 °C for 15 min to insert the mCherry gene downstream from GFP. The assembled plasmid was transformed into NEB5α cells. After sequence verification by Oxford Nanopore sequencing (Plasmidsaurus), the plasmid was transformed to NiCo21-(DE3) cells.

Strains and Growth Conditions. Glycerol stocks of NiCo21(DE3) cells (NEB) transformed with each variant were streaked onto LB agar plates with ampicillin (100 μg/mL, Gold Biotechnology). A single colony was inoculated into sterilized LB medium supplemented with ampicillin (100 μg/mL) and grown in an incubator (Thermo Fisher Scientific MaxQ 6000) for 16−18 h at 37 °C while being shaken at 225 rpm. The overnight cultures were back-diluted to OD_{600} ∼0.1 in 25 mL of LB media supplemented with ampicillin (100 μg/mL) in sterile 125 mL Erlenmeyer flasks. Cultures were grown at 37 °C, with shaking at 225 rpm for 2−3 h. At OD_{600} = 0.7−1.0, cultures were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Gold Biotechnology) and then maintained at 25 °C with shaking at 225 rpm for 24 h. Twenty μL aliquots were taken from the cultures at various time points after induction for imaging via optical microscopy.

Cell Growth and Protein Expression Assays. The cellular growth of NiCo21(DE3) *E. coli* and fluorescent protein production were monitored for the GFP variants studied here using a plate-based growth assay. Three colonies containing the plasmid for each variant were selected from a freshly grown LB agar plate and grown to saturation overnight at 37 °C with shaking at 225 rpm in 1 mL of LB supplemented with ampicillin in a 24-well plate. The optical density (OD_{600}) was measured and each replicate was back-diluted to an OD_{600} of ∼0.1 with LB with added ampicillin, for a total volume of 1 mL per culture. Three wells of 1 mL of LB were included to measure the background media fluorescence. After 2−3 h, when the cultures reached a corrected OD_{600} between 0.7 and 1.0, the cultures were induced with 1 mM IPTG. The plate was then incubated at 25 °C. At 0, 2, 4, 6, 8, 10, 20, and 24 h post-induction, the OD_{600} and GFP fluorescence (λ_{ex} = 488 nm, λ_{em} = 530 nm) were measured on an Infinite M200 Pro microplate reader (Tecan). The data were analyzed by dividing the background-corrected GFP fluorescence values by the background-corrected OD_{600} values. For variants coexpressing mCherry, the OD at 640 nm was measured instead to avoid interference with the protein absorbance. Time point measurements were taken at 0, 1, 2, 4, 6, 8, 19, and 24 h for the single-protein controls and at 2, 4, 6, 8, 19, and 24 h for the coexpressed variants. To account for the instrument bandwidth and ensure minimal crosstalk between proteins, GFP fluorescence was measured at λ_{ex} = 470 and λ_{em} = 510 nm and mCherry fluorescence was measured at λ_{ex} = 570 and λ_{em} = 610 nm.

Fluorescence and Optical Microscopy. Cell samples were applied to agarose pads for imaging. The agarose pads were made by preparing 1% (w/v) agarose (TopVision) in Milli-Q water. 50 μL of melted agarose was then pipetted onto a 25 mm × 75 mm microscope slide and immediately covered by an 18 mm × 18 mm #1.5 coverslip (Thermo Fisher Scientific). After the agarose solidified, the coverslip was slowly removed, and 2 μL of cell culture from the 20 μL aliquot was added on top of the agarose pad. The agarose pad with the sample was gently covered with a coverslip and sealed on all four sides with clear nail polish. Cells were imaged on agarose pads at 2, 8, and 24 h post-induction, or 2, 3, 4, 5, 6, 8, 9, 10, 12, and 24 h post-induction for the time course experiment

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

(Figure 2). Images were taken using a 100× oil 1.40 NA UPlanSAPO objective (Olympus) with illumination from the EVOS GFP light cube (λ_{ex} = 470/22 nm; λ_{em} = 525/50 nm), EVOS Texas Red light cube (λ_{ex} = 585/29 nm; λ_{em} = 628/32 nm), and brightfield channels on an EVOS FL Auto 2 inverted fluorescent microscope. Approximately 10 fields of view were taken for each sample to acquire images with a sufficient number of cells per strain (>96 cells per biological replicate). At least 3 biological replicates were performed for each strain at each time point.

Protein Purification for In Vitro State Diagrams. An overnight culture was inoculated into 1 L of LB medium supplemented with ampicillin (100 μg/mL). The same growth and protein expression conditions as those described above were used. Following protein expression, cells were harvested by centrifugation (Thermo Fisher Scientific Sorvall Legend XTR) in a swinging bucket rotor (Thermo Fisher Scientific TX-750) at 4000 rpm for 15 min and resuspended in lysis buffer (50 mM NaH_2PO_4, 1 M NaCl, pH 8.0). The resuspended cell pellet was then subjected to one freeze−thaw cycle. RNase A (225 μg/L of culture) and DNase I (200 μg/L of culture) were added to the thawed resuspended cells prior to lysis by sonication (cycle: 2 s on, 4 s off) at 40% amplitude using a 1/8 in. probe for 10 min (Thermo Fisher Scientific). Soluble proteins were separated from cell debris by centrifugation in a fixed angle rotor (Thermo Fisher Scientific, Fiberlite F15−8 × 50cy) at 10,000 rpm for 30 min at room temperature. Proteins in the soluble fraction were collected and purified by using immobilized metal affinity chromatography. Briefly, 10−15 mL of His-Pur Ni-NTA resin (Thermo Fisher Scientific) was used per 1 L of the cell culture. The resin was initially equilibrated in 3 column volumes of lysis buffer before incubation with the soluble cell lysate. Fractions of unbound lysate, wash (lysis buffer containing 50 mM imidazole), and elution (lysis buffer containing 250 mM imidazole) were collected. Fractions were analyzed on a Bolt 4−12% Bis−Tris Plus gel (Invitrogen) to determine the protein purity. Samples were prepared in 4× LDS loading buffer and incubated at 95 °C for 10 min. 10 μL of each sample was run at 200 V for 22 min. Gels were stained following the SimplyBlue SafeStain protocol and imaged (Figure S2C−E). Pure fractions were then dialyzed against a 50 mM HEPES-NaOH, pH 7.4, buffer for at least 3 h per equilibration with a total of 7 exchanges of buffer. Dialyzed protein samples were concentrated using Amicon Ultra centrifugal filter units with a 10 kDa molecular weight cutoff (Sigma-Aldrich).

Liquid Chromatography−Mass Spectrometry. LC−MS was performed on select variants after purification. Protein stock solutions were diluted to 0.1 mg/mL in Milli-Q water. Reversed phase chromatography was performed using an Agilent 1260 HPLC system equipped with a 1 mm × 75 mm InfinityLab Poroshell 300Extend-C18 column (671750-902, Agilent Technologies) with 0.1% formic acid in Milli-Q water and acetonitrile as solvents A and B, respectively. The flow rate was 0.4 mL/min, and the column temperature was set at 25 °C. The column was equilibrated with 3% B for 10 min, and a sample volume of 5 μL was injected onto the column. The gradient elution profile was from 3% to 90% B for 4 min, followed by washing with 97% B for 1 min. Full mass data were acquired with a range of 100−3000 m/z and an acquisition rate of 1 spectra/s on an Agilent 6230 LC/TOF system. The instrument was operated in the positive mode for intact protein analysis. Mass spectrum was deconvoluted using Agilent MassHunter BioConfirm software with the built-in Maximum Entropy algorithm.

Turbidimetry Assay for In Vitro State Diagrams. Protein concentrations were obtained by measuring absorbance at 488 nm on a Cary 60 UV−vis spectrophotometer (Agilent Technologies) and calculating concentrations using Beer's Law. The extinction coefficient for sfGFP (ϵ = 8.33 × 10^5 M^{-1} cm^{-1}) was used to calculate the concentration for all supercharged GFP variants. Stock solutions of fluorescent protein (20 and 0.2 mg/mL) and total RNA from torula yeast type VI (2 and 0.02 mg/mL) (Sigma-Aldrich #R6625-25G) were prepared in 50 mM HEPES-NaOH, pH 7.4. The pH of the RNA stock was adjusted to 7.4 with 10 M NaOH immediately after dissolution. Protein and RNA stock solutions were filtered using a 0.22 μm SFCA Thermo Fisher Scientific Nalgene 25 mm syringe filter. State diagrams were constructed by mixing buffer, protein, and RNA at various mass ratios to a total volume of 50 μL in a tissue culture-treated polystyrene 96-well half-area plate (Corning, #3697). The absorbance (A) of each sample was measured at 600 nm immediately after mixing on a plate reader (Tecan Infinite M200 Pro) at 25 °C and converted to turbidity using the following equation: Turbidity = 100 − 10^{2-A}. Heatmaps of the turbidity values at each of the tested macromolecule concentrations were generated in MATLAB. For GFP(−6)-L_2, a filled contour plot of turbidity isolines was also generated in MATLAB. For in vitro experiments with GFP(−6)-H_2, mCherry-L_2, and RNA, stock solutions of 8 mg/mL protein and 1 mg/mL RNA were mixed to a total volume of 30 μL in 10 mM HEPES-NaOH, pH 7.4 (Figure S6). The mass concentration ratio of GFP(−6)-H_2 to mCherry-L_2 was 1:1. The total protein and RNA concentration was fixed at 5 mg/mL, and the positive charge fractions was varied (0.5, 0.6, and 0.8).

Electrostatic Maps of GFP Variants. The PDB file for sfGFP (PDB ID: 2B3P) was obtained from the RCSB protein data bank and loaded in PyMOL v2.5.5 (Schrödinger, LLC). Using the PDB file for sfGFP as a template, GFP(0) and GFP(+6) were made by introducing residue substitutions in PyMOL.[31] PDBs of the minimized peptide structures were obtained using PEP-FOLD3.[49] The C-terminus of the globular domain was appended to the N-terminus of the disordered peptides in PyMOL, and electrostatic surface maps were generated by using the ABPS Electrostatics plugin.

Image Preprocessing and Cell Segmentation. Fluorescence microscopy images were preprocessed using Fiji with a plug-in. Briefly, images were compiled into stacks and were subjected to a rolling ball background subtraction with a 100-pixel radius. Stacks of images were individually thresholded using the Li method in MicrobeJ, and cells that met specific area, length, width, circularity, and angularity constraints were identified and considered for analysis. The following constraints were used for all image stacks: area = 100−2000 $pixels^2$, length = 20−max pixels, width = 0−20 pixels, circularity = 0.25−max, curvature = 0−max, sinuosity = 0−max, solidity = 0−max, and intensity = 0−max. Bacteria were detected using the fit shape, rod-shaped mode. Images of individual cells identified by MicrobeJ were loaded into MATLAB for the processing and identification of condensates. Cell attributes, including the cell area, mean width, and length, were also imported from MicrobeJ. A custom MATLAB script was used to identify condensates and determine the fraction of

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612

condensate-containing cells, the condensate area fraction, the condensate intensity, and the condensate localization.

Condensate Thresholding within Cells. The 2D interpolated cell arrays from MicrobeJ were stored. Any remaining background pixels in the 2D array with an intensity lower than the minimum intensity from the cell contour were removed. The pixel intensity was sampled at three locations along the medial axis of each cell to approximate the intensity of a noncondensate area within the cell. In each of the three locations, the intensities of nine adjacent pixels were averaged, and the lowest of the three averages was taken as the approximate average pixel intensity of the cytoplasm. All regions in a given cell with 20% higher intensity than the cell's approximated average intensity and with at least 5 contiguous pixels at this higher intensity were classified as condensates. All other pixels were classified as the cytoplasm. Pixels from the cytoplasm and identified condensates were stored as 2D arrays along with the parent cell.

Fraction of Cells with Condensates. All cells were classified as condensate-forming cells or noncondensate-forming cells. The fraction of condensate-containing cells for each replicate was calculated by dividing the number of condensate-forming cells by the total number of cells.

Cell and Condensate Dimensions. Cell dimensions, including the width, length, and area, were obtained from MicrobeJ. The area of each condensate within a cell was determined by MATLAB. Condensates were analyzed separately for cells containing multiple condensates. The area fraction of each condensate was calculated by dividing the area of the condensate by the area of the cell. For cells with multiple condensates, the area fractions occupied by all condensates were summed up to obtain the total area fraction of condensates per cell.

Cell and Condensate Pixel Intensities. The mean cell intensity within each cell contour was obtained from MicrobeJ. The mean intensities of the condensate and cytoplasm for each cell were calculated with MATLAB. For cells containing multiple condensates, the intensity was calculated separately for each condensate. The intensity ratio was calculated as the mean intensity of the condensate divided by the mean intensity of the cytoplasm. Data from MATLAB was transferred to GraphPad Prism (v. 9.5.1) for plotting and visualization.

Condensate Localization Heat Maps. Localization heat maps were generated by binning the normalized cell position of the 1-pixel maxima from all condensates into a 20 × 10 array. Cells containing one condensate were aligned with the condensate at the positive end of the medial axis. A representative cell contour was generated by binning background pixels into a 20 × 10 array and generating a contour that represents 90% of all cells.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.3c00564.

Analysis of the disorder and charge-patterning of C-terminal cationic peptides, cell growth and protein expression data for variants with different cationic peptide lengths, cell growth and protein expression data over time for all +6 net charge variants, cell growth and protein expression data over time for all +12 net charge variants, cell growth and protein expression data over time for all GFP and mCherry coexpressed variants, LC−MS and microscopy data for the in vitro phase separation of GFP(−6)-H_2 and mCherry-L_2 with total yeast RNA, key resources, and primer sequences for the construction of L- and H-peptide variants (PDF)

Time-lapse of the partial dissolution of GFP(+6) condensates in single cells from 12 to 13 h post-induction (AVI)

■ AUTHOR INFORMATION

Corresponding Author

Allie C. Obermeyer − *Department of Chemical Engineering, Columbia University, New York, New York 10027, United States;* orcid.org/0000-0003-2412-2021; Email: aco2134@columbia.edu

Authors

Jane Liao − *Department of Chemical Engineering, Columbia University, New York, New York 10027, United States*

Vivian Yeong − *Department of Chemical Engineering, Columbia University, New York, New York 10027, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.3c00564

Author Contributions

Conceptualization, J.L., V.Y., and A.C.O.; Investigation, J.L. and V.Y.; Data analysis and writing—original draft and review and editing, J.L., V.Y., and A.C.O.

Notes

The authors declare the following competing financial interest(s): A.C.O is a co-founder of Werewool, a company that is engaged in the development of performance textiles that incorporate engineered proteins. J.L. declares no competing interests.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (DMR: 1848388) (J.L., V.Y., and A.C.O.). J.L. also acknowledges support from the Blavatnik Fund for Engineering Innovations in Health.

■ REFERENCES

(1) Hirose, T.; Ninomiya, K.; Nakagawa, S.; Yamazaki, T. A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol.* **2023**, *24*, 288−304.

(2) Lyon, A. S.; Peeples, W. B.; Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 215−235.

(3) Pederson, T. The plurifunctional nucleolus. *Nucleic Acids Res.* **1998**, *26*, 3871−3876.

(4) Lafontaine, D. L. J.; Riback, J. A.; Bascetin, R.; Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 165−182.

(5) Cajal, S. R. Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Trab. Lab. Invest. Biol.* **1903**, *2*, 129−221.

(6) Saitoh, N.; Spahr, C. S.; Patterson, S. D.; Bubulya, P.; Neuwald, A. F.; Spector, D. L. Proteomic Analysis of Interchromatin Granule Clusters. *Mol. Biol. Cell* **2004**, *15*, 3876−3890.

(7) Fox, A. H.; Lam, Y. W.; Leung, A. K.; Lyon, C. E.; Andersen, J.; Mann, M.; Lamond, A. I. Paraspeckles. *Curr. Biol.* **2002**, *12*, 13−25.

(8) Gilks, N.; Kedersha, N.; Ayodele, M.; Shen, L.; Stoecklin, G.; Dember, L. M.; Anderson, P. Stress Granule Assembly Is Mediated by Prion-like Aggregation of TIA-1. *Mol. Biol. Cell* **2004**, *15*, 5383−5398.
(9) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A. A. Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **2009**, *324*, 1729−1732.
(10) Gao, Z.; Zhang, W.; Chang, R.; Zhang, S.; Yang, G.; Zhao, G. Liquid-Liquid Phase Separation: Unraveling the Enigma of Biomolecular Condensates in Microbial Cells. *Front. Microbiol.* **2021**, *12*, 751880.
(11) Azaldegui, C. A.; Vecchiarelli, A. G.; Biteen, J. S. The emergence of phase separation as an organizing principle in bacteria. *Biophys. J.* **2021**, *120*, 1123−1138.
(12) Abbondanzieri, E. A.; Meyer, A. S. More than just a phase: the search for membraneless organelles in the bacterial cytoplasm. *Curr. Genet.* **2019**, *65*, 691−694.
(13) Muthunayake, N. S.; Tomares, D. T.; Childers, W. S.; Schrader, J. M. Phase-separated bacterial ribonucleoprotein bodies organize mRNA decay. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, No. e1599.
(14) Greening, C.; Lithgow, T. Formation and function of bacterial organelles. *Nat. Rev. Microbiol.* **2020**, *18*, 677−689.
(15) Cohan, M. C.; Pappu, R. V. Making the Case for Disordered Proteins and Biomolecular Condensates in Bacteria. *Trends Biochem. Sci.* **2020**, *45*, 668−680.
(16) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 285−298.
(17) Cates, M. E.; Witten, T. A. Chain conformation and solubility of associating polymers. *Macromolecules* **1986**, *19*, 732−739.
(18) Semenov, A. N.; Rubinstein, M. Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics. *Macromolecules* **1998**, *31*, 1373−1385.
(19) Ginell, G. M.; Holehouse, A. S.. In *In Phase-Separated Biomolecular Condensates*; Zhou, H.-X., Spille, J.-H., Banerjee, P. R., Eds.; *Methods in Molecular Biology*; Springer US: New York, NY, 2023; Vol. *2563*, pp 95−116.
(20) Mittag, T.; Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **2022**, *82*, 2201−2214.
(21) Schuster, B. S.; Reed, E. H.; Parthasarathy, R.; Jahnke, C. N.; Caldwell, R. M.; Bermudez, J. G.; Ramage, H.; Good, M. C.; Hammer, D. A. Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat. Commun.* **2018**, *9*, 2985.
(22) Simon, J. R.; Carroll, N. J.; Rubinstein, M.; Chilkoti, A.; López, G. P. Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. *Nat. Chem.* **2017**, *9*, 509−515.
(23) Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; et al. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014**, *114*, 6589−6631.
(24) Lasker, K.; Von Diezmann, L.; Zhou, X.; Ahrens, D. G.; Mann, T. H.; Moerner, W. E.; Shapiro, L. Selective sequestration of signalling proteins in a membraneless organelle reinforces the spatial regulation of asymmetry in Caulobacter crescentus. *Nat. Microbiol.* **2020**, *5*, 418−429.
(25) Al-Husini, N.; Tomares, D. T.; Bitar, O.; Childers, W. S.; Schrader, J. M. α-Proteobacterial RNA Degradosomes Assemble Liquid-Liquid Phase-Separated RNP Bodies. *Mol. Cell* **2018**, *71*, 1027−1039.e14.
(26) Al-Husini, N.; Tomares, D. T.; Pfaffenberger, Z. J.; Muthunayake, N. S.; Samad, M. A.; Zuo, T.; Bitar, O.; Aretakis, J. R.; Bharmal, M.-H. M.; Gega, A.; Biteen, J. S.; Childers, W. S.; Schrader, J. M. BR-Bodies Provide Selectively Permeable Condensates that Stimulate mRNA Decay and Prevent Release of Decay Intermediates. *Mol. Cell* **2020**, *78*, 670−682.e8.
(27) Ge, X.; Conley, A. J.; Brandle, J. E.; Truant, R.; Filipe, C. D. M. In Vivo Formation of Protein Based Aqueous Microcompartments. *J. Am. Chem. Soc.* **2009**, *131*, 9094−9099.
(28) Wei, S.-P.; Qian, Z.-G.; Hu, C.-F.; Pan, F.; Chen, M.-T.; Lee, S. Y.; Xia, X.-X. Formation and functionalization of membraneless compartments in Escherichia coli. *Nat. Chem. Biol.* **2020**, *16*, 1143−1148.
(29) Dzuricky, M.; Rogers, B. A.; Shahid, A.; Cremer, P. S.; Chilkoti, A. De novo engineering of intracellular condensates using artificial disordered proteins. *Nat. Chem.* **2020**, *12*, 814−825.
(30) Dai, Y.; Farag, M.; Lee, D.; Zeng, X.; Kim, K.; Son, H.-i.; Guo, X.; Su, J.; Peterson, N.; Mohammed, J.; Ney, M.; Shapiro, D. M.; Pappu, R. V.; Chilkoti, A.; You, L. Programmable synthetic biomolecular condensates for cellular control. *Nat. Chem. Biol.* **2023**, *19*, 518−528.
(31) Yeong, V.; Werth, E. G.; Brown, L. M.; Obermeyer, A. C. Formation of Biomolecular Condensates in Bacteria by Tuning Protein Electrostatics. *ACS Cent. Sci.* **2020**, *6*, 2301−2310.
(32) Yeong, V.; Wang, J.-w.; Horn, J. M.; Obermeyer, A. C. Intracellular phase separation of globular proteins facilitated by short cationic peptides. *Nat. Commun.* **2022**, *13*, 7882.
(33) Espinosa, J. R.; Joseph, J. A.; Sanchez-Burgos, I.; Garaizar, A.; Frenkel, D.; Collepardo-Guevara, R. Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 13238−13247.
(34) Dai, Y.; You, L.; Chilkoti, A. Engineering synthetic biomolecular condensates. *Nat. Rev. Bioeng.* **2023**, *1*, 466−480.
(35) Molinari, S.; Shis, D. L.; Bhakta, S. P.; Chappell, J.; Igoshin, O. A.; Bennett, M. R. A synthetic system for asymmetric cell division in Escherichia coli. *Nat. Chem. Biol.* **2019**, *15*, 917−924.
(36) Lin, D.-W.; Liu, Y.; Lee, Y.-Q.; Yang, P.-J.; Ho, C.-T.; Hong, J.-C.; Hsiao, J.-C.; Liao, D.-C.; Liang, A.-J.; Hung, T.-C.; Chen, Y.-C.; Tu, H.-L.; Hsu, C.-P.; Huang, H.-C. Construction of intracellular asymmetry and asymmetric division in Escherichia coli. *Nat. Commun.* **2021**, *12*, 888.
(37) Hu, X.-P.; Dourado, H.; Schubert, P.; Lercher, M. J. The protein translation machinery is expressed for maximal efficiency in Escherichia coli. *Nat. Commun.* **2020**, *11*, 5260.
(38) Parry, B. R.; Surovtsev, I. V.; Cabeen, M. T.; O'Hern, C. S.; Dufresne, E. R.; Jacobs-Wagner, C. The Bacterial Cytoplasm Has Glass-like Properties and Is Fluidized by Metabolic Activity. *Cell* **2014**, *156*, 183−194.
(39) Zhu, Y.; Mustafi, M.; Weisshaar, J. C. Biophysical Properties of Escherichia coli Cytoplasm in Stationary Phase by Superresolution Fluorescence Microscopy. *mBio* **2020**, *11*. DOI: 10.1128/mbio.00143-2.
(40) Prossliner, T.; Skovbo Winther, K.; Sørensen, M. A.; Gerdes, K. Ribosome Hibernation. *Annu. Rev. Genet.* **2018**, *52*, 321−348.
(41) Mengin-Lecreulx, D.; van Heijenoort, J. Effect of growth conditions on peptidoglycan content and cytoplasmic steps of its biosynthesis in Escherichia coli. *J. Bacteriol.* **1985**, *163*, 208−212.
(42) Navarro Llorens, J. M.; Tormo, A.; Martínez-García, E. Stationary phase in gram-negative bacteria. *FEMS Microbiol. Rev.* **2010**, *34*, 476−495.
(43) Aumiller, W. M.; Pir Cakmak, F.; Davis, B. W.; Keating, C. D. RNA-Based Coacervates as a Model for Membraneless Organelles: Formation, Properties, and Interfacial Liposome Assembly. *Langmuir* **2016**, *32*, 10042−10053.
(44) Wollny, D.; Vernot, B.; Wang, J.; Hondele, M.; Safrastyan, A.; Aron, F.; Micheel, J.; He, Z.; Hyman, A.; Weis, K.; Camp, J. G.; Tang, T. D.; Treutlein, B. Characterization of RNA content in individual phase-separated coacervate microdroplets. *Nat. Commun.* **2022**, *13*, 2626.
(45) Boeynaems, S.; et al. Aberrant Phase Separation Is a Common Killing Strategy of Positively Charged Peptides in Biology and Human Disease, 2023. Pages: 2023.03.09.531820 Section: New Results. https://www.biorxiv.org/content/10.1101/2023.03.09.531820vweb1 (accessed 12/18/2023).

(46) Sneideris, T.; Erkamp, N. A.; Ausserwöger, H.; Saar, K. L.; Welsh, T. J.; Qian, D.; Katsuya-Gaviria, K.; Johncock, M. L. L. Y.; Krainer, G.; Borodavka, A.; Knowles, T. P. J. Targeting nucleic acid phase transitions as a mechanism of action for antimicrobial peptides. *Nat. Commun.* **2023**, *14*, 7170.
(47) Lin, Y.-H.; Forman-Kay, J. D.; Chan, H. S. Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Phys. Rev. Lett.* **2016**, *117*, 178101.
(48) Dao, T. P.; Rajendran, A.; Galagedera, S. K.; Haws, W.; Castañeda, C. A. Short disordered N-termini & proline-rich domain are major regulators of UBQLN1/2/4 phase separation. *Biophys. J.* **2023**, DOI: 10.1016/j.bpj.2023.11.3401.
(49) Lamiable, A.; Thévenet, P.; Rey, J.; Vavrusa, M.; Derreumaux, P.; Tufféry, P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res.* **2016**, *44*, W449−W454.

https://doi.org/10.1021/acssynbio.3c00564
ACS Synth. Biol. 2024, 13, 598−612