

Title

De Novo Engineering of Intracellular Condensates using Artificial Disordered Proteins

Authors:

Michael Dzuricky¹, Bradley A. Rogers², Abdulla Shahid^{3,4}, Paul S. Cremer^{2,5}, Ashutosh Chilkoti^{1†}

Author Addresses:

¹Department of Biomedical Engineering, Duke University, Durham, NC 27708

²Department of Chemistry, Pennsylvania State University, University Park, PA 16802

³Department of Computer Science, Duke University, Durham, NC 27708

⁴Department of Biology, Duke University, Durham, NC 27708

⁵Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802

†Correspondence to chilkoti@duke.edu

Highlights

Artificial IDPs are designed to exhibit specific phase behavior in aqueous solvent.

The two-phase boundary is principally defined by aromatic content and chain length

Intracellular phase behavior is controlled by the same variables as *in vitro*

Artificial IDP condensates can recruit enzymes to increase their catalytic efficiency

Keywords

intrinsically disordered protein, upper critical solution temperature, liquid-liquid phase separation, condensates, puncta, protein polymer, charge balance, aromatic, aliphatic, molecular weight

In Brief

A class of Artificial Intrinsically Disordered Proteins (A-IDPs) that exhibit upper critical solution temperature transition phase behavior to form intracellular condensates in living cells can be rationally designed from simple physio-chemical principles – namely the ratio of aromatic to aliphatic residues, charge balance and molecular weight (MW) of the A-IDP. Using A-IDPs as a minimal condensate scaffold, Dzuricky et. al. engineered intracellular enzymatic condensates that increase β -galactosidase's catalytic efficiency with increasing MW of the A-IDP.

Summary

Phase separation of intrinsically disordered proteins (IDPs) is a remarkable feature of living cells to dynamically control intracellular partitioning of its constituents. Despite the numerous new IDPs that have been identified, progress towards rational engineering of phase separation in cells has been limited. To address this limitation, we systematically scanned the sequence space of native IDPs and designed artificial IDPs (A-IDPs) with different molecular weights (MWs) and aromatic content that exhibit variable condensate saturation concentrations and temperature cloud points *in vitro* and in living cells. These studies identified a set of simple molecular and physical principles that drive liquid-liquid phase separation. Using these principles, we created A-IDP puncta, capable of sequestering an enzyme, whose catalytic efficiency can be manipulated by the MW of the A-IDP. These results provide an engineered platform based on a set of A-IDPs that provide new, phase separation mediated control of biological function in living cells.

Introduction

Intrinsically disordered proteins (IDPs) are receiving significant recognition for their role in various biological (dys)functions (Babu, 2016; Wright and Dyson, 2015). A subset of IDPs, termed biological condensates, physically separate themselves from the cytoplasm to control the accessibility of a variety of macromolecules (Boeynaems et al., 2018; Uversky et al., 2015). While our ability to detect protein disorder has advanced rapidly thanks to sophisticated statistical methods, the ability to predict phase separation has lagged behind (Sickmeier et al., 2007). The prediction of phase separation is non-trivial, as numerous variables influence phase separation. Broadly, they involve: 1) amino acid composition and amino acid patterning of the primary protein sequence (Brangwynne et al., 2015; Lin et al., 2017; Pak et al., 2016); 2) heterotypic interactions with RNA or other macromolecules (Elbaum-Garfinkle et al., 2015; Lin et al., 2015) and 3) solvent quality (Franzmann et al., 2018; Molliex et al., 2015; Nott et al., 2015). There are many studies that note the challenge of predicting IDP phase behavior, but few studies that have directly tackled this problem. Given the recognition of its importance to cellular function, this is now an area of active research and many efforts are ongoing using computational (Best, 2017; Dignon et al., 2018a; Dignon et al., 2018b; Mao et al., 2010) and experimental approaches (Wang et al., 2018). To date, however, most experimental methods to develop a sequence level understanding of IDP phase behavior have relied on mutational strategies of native IDPs with sweeping residue level or domain level mutations (Dzuricky et al., 2018).

We have taken a different and complementary approach to understand how phase behavior is encoded in polypeptides. Analogous to—and inspired by—synthetic polymers that exhibit lower and upper critical solution temperature (LCST/UCST) phase behavior, we began by systematically scanning the sequence space of native IDPs to identify minimal peptide motifs that will confer LCST or UCST phase behavior when polymerized into a macromolecule that consists of many repeats of the peptide motif. (Meyer and Chilkoti, 2004; Quiroz and Chilkoti, 2015). With the greatly reduced sequence complexity of these repetitive polypeptides—compared to native IDPs that exhibit LCST/UST phase behavior—we then made rational changes in the amino acid repeat

motif that systematically propagate along the sequence. These repetitive polypeptides can be rationally designed to exhibit both LCST and UCST phase behavior, and their phase behavior can be systematically modulated by amino acid mutations of the repeat motif (Quiroz and Chilkoti, 2015). These artificial polypeptides also exhibit the same basic principles of phase separation inside cells as native IDPs (Huber et al., 2015; Li et al., 2018; Pastuszka et al., 2012; Shi et al., 2014).

Informed by a heuristic knowledge of factors that drive phase separation in repetitive polypeptides from these studies, herein we set out to create artificial IDPs (A-IDPs) that exhibit phase separation in living cells to impart new functionality to the cell. Our design began with $(G_1-R_2-G_3-D_4-S_5-P_6-Y_7-S_8)_{xx}$ (where xx is the number of repeats between 20 and 80) a sequence inspired by *Drosophila Melanogaster* Rec-1 Resilin, known to exhibit UCST phase behavior. We chose this sequence because it exhibits UCST phase behavior, which appears to be far more common among native IDPs than LCST phase behavior (Elbaum-Garfinkle et al., 2015; Molliex et al., 2015; Nott et al., 2015; Qamar et al., 2018). Using proteomic analysis and design considerations from previous studies, we created a set of 63 A-IDPs consisting of repeats of the parent $(G_1-R_2-G_3-D_4-S_5-P_6-Y_7-S_8)_{xx}$ motif and variants with rational amino acid mutations of the $(G_1-R_2-G_3-D_4-S_5-P_6-Y_7-S_8)_{xx}$ repeat. We characterized the UCST phase behavior for this set of 63 IDPs from which we were able to quantify the effect of various amino acid mutations and modifications to the chain architecture on homotypic liquid-liquid phase separation.

We then used a subset of A-IDPs from this library to engineer intracellular condensates in living cells. The behavior of intracellular condensates for these A-IDPs proved to be surprisingly predictable and tunable, and enabled dynamic control over their cytoplasmic solubility and their interaction with the surrounding environment. Capitalizing on these observations, we created intracellular droplets capable of sequestering an enzyme whose catalytic efficiency within the engineered condensates can be genetically encoded by modulating the MW of the A-IDP.

Results

Identification of a minimal IDP repeat from proteomic analysis and sequence heuristics

We conducted a proteomic analysis of 63 IDPs that form membrane-less organelles to investigate their sequence composition (Uversky et al., 2015). We were particularly interested in categories of amino acids suspected to drive phase behavior *via* intrachain interactions, such as charge-charge, cation- π and hydrogen bonding *via* non-charged polar residues (Balu et al., 2014; Dzuricky et al., 2018; Li et al., 2015; Lin et al., 2016; Pak et al., 2016; Quiroz and Chilkoti, 2015; Shin and Brangwynne, 2017) (Figure 1A). The composition of these 63 proteins is remarkably similar to previously identified repetitive protein polypeptides that exhibit UCST phase behavior and their side-chains groups are chemically similar to synthetic UCST polymers (Balu et al., 2014; Li et al., 2015; Quiroz and Chilkoti, 2015). Using a combination of our previously developed sequence heuristics and insights from this proteomic analysis, we designed an octa-peptide motif that we expected would exhibit robust phase behavior when polymerized into a macromolecule, under physiologically relevant solution conditions.

In order to manage the vast sequence space of all possible mutations of the octapeptide repeat, we classify each amino acid into categories of intrachain interactions that could contribute to UCST phase behavior. N, Q, S, T are classified as polar, uncharged amino acids. R-K and D-E are pairs of positively charged and negatively charged amino acids. G and P are placed into a separate category given their unusual structure and importance in promoting a disordered polypeptide backbone (Figure S1A) (Theillet et al., 2013; Uversky and Dunker, 2010). The remaining amino acids are classified as “hydrophobic” (Kyte and Doolittle, 1982; Wimley and White, 1996). To ensure that we modulate the UCST phase behavior via mutagenesis of the WT repeat, but do not abolish it completely, we only make mutations wherein the mutant maintains the type of interactions and simply modulates the strength of that interaction. For example, R and K are both positively charged under normal physiological pH. Thus, by substituting K for R we maintain the charge neutral state of the polymeric backbone—a parameter known to dramatically affect the observed phase behavior (Das and Pappu, 2013; Lin and Chan, 2017; Mao et al., 2010; Urry et al., 1992). Similarly, N, Q, S and T are all capable of creating hydrogen bonds with water and one another more readily than an aliphatic amino acid such as V. Thus, substituting these four amino acids for one another maintains an equal number of residues per chain capable for forming this particular type of bond.

The wild-type (WT) repeat unit is (G₁-R₂-G₃-D₄-S₅-P₆-Y₇-S₈)₄₀ where 40 refers to the number of repeats. The MW of the A-IDPs was varied between ~15 and ~70 kDa —by varying the number of repeat motifs from 20 to 80— to account for observed differences in MW in the intrinsically disordered regions (IDRs) of naturally occurring IDPs (Figure S1B). The parent sequence is referred to as WT in this paper, and we use a short-hand notation to refer to sequences throughout the text where the bracketed letter refers to a specific point — substitution— mutant. For example, a mutant with a complete substitution of Y₇ in the WT repeat unit with V would result in a notation of “[V₇]-XX”. When a residue is only partially substituted in the A-IDP, we use the notation “[BY_o:ZV_o]" where the B to Z ratio represents the ratio of Y to V ratio in the variant and the subscript *o* is the position of that residue along the repeat unit. For example, [Y₇:V₇]-40 would hence represent 50% of all Y replaced with V, whereas [3Y₇:V₇]-40 would represent a 25% substitution of V for Y. A double mutant, such as 100% substitution of residues at the 5th and 8th position in the octapeptide repeat with Q, would be denoted as [Q_{5,8}]-XX with and fractional substitution at these positions with S and Q would be denoted as [BS_{5,8}:ZQ_{5,8}]-XX where B and Z represent the ratio of S to Q. Full sequence descriptions of common sequences used throughout the paper can be found in Table 1. A full description of all architectures of A-IDPs wherein mutant and WT repeats are mixed along the A-IDP chain can be found in Table S1 and Table S2.

A-IDPs Exhibit Robust and Reversible UCST Phase Behavior in an Aqueous Environment

One advantage of A-IDPs is their minimal interaction with other proteins or biomolecules stemming from their repetitive nature. This feature of A-IDPs combined with their reversible aqueous two-phase separation enables simple column-free purification by UCST phase transition cycling between the one- and two-phase regime of the phase diagram. An example of this purification process is shown in Figure 1B, where the highly expressing A-IDP, [Q_{5,8}]-20, completely phase separates from the soluble fraction of the cell lysate and can be isolated by centrifugation. Subsequent removal of the protein-poor supernatant, dissolution of the protein-rich pellet with urea, and dialysis of the soluble fraction in milli-Q water results in 95-99% pure protein as observed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 1C & Figure S2). The yield of purified A-IDP ranges from 25-300 mg per liter of culture media in shaker flask culture.

A-IDPs [WT]-20 and [Q_{5,8}]-20 exhibit UCST phase behavior *in vitro*. To characterize their phase transition behavior, we employed three different techniques. First, we utilized droplet microfluidics, where monodisperse water droplets are formed in oil containing the A-IDP of interest (Figure 1D). Phase separation can be directly visualized in the spatially limited compartment of water-in-oil microdroplets to observe the types of structures formed as one cools the surrounding medium. These A-IDPs exhibit classic liquid-liquid phase separation where, upon crossing the phase boundary upon cooling from 50 °C to 10 °C, multiple nucleation sites of coacervate condensates are observed (Figure 1D Panel 2). These nucleation sites wet one another and quickly coalesce into a single, spherical A-IDP–dense phase that is in equilibrium with the surrounding A-IDP–poor phase (Figure 1D panel 3 & 3-1/3-2). Upon reheating to 50 °C, the A-IDP–rich phase shrinks in size, as the A-IDP re-solubilizes, re-establishing equilibrium rapidly (Figure 1D Panel 4-1/4-2). A wider field of view of this transition can be found in Figure S3. These data clearly show that these A-IDPs exhibit *reversible* UCST phase separation via coalescence and growth kinetics (Figure 1E).

Second, we employed temperature dependent dynamic light scattering (DLS) to observe the two-phase separation in bulk. A solution of [Q_{5,8}]-20 is heated to 80 °C and DLS data was collected as the solution is cooled to 10 °C. We observe a transition from soluble A-IDP molecules with a hydrodynamic radius (R_h) of 4 nm to aggregates larger than 1 μ m as a function of temperature (Figure 1F and Figure S4). This transition is quite sharp, as it occurs within a 2°C window at ~38°C.

Third, we employed temperature-dependent turbidity measurements at a fixed wavelength of 350 nm to characterize the UCST phase separation while heating and cooling a solution of an A-IDP at a rate of 1°C /min (Figure 1G). Using this technique, we can capture partial phase diagrams of each A-IDP of interest as a function of many different sequence and solution parameters. At dilute volume fractions of [Q_{5,8}]-40 we observed different UCST cloud points that increase as a function of the natural logarithm of A-IDP concentration ($T_i = m \cdot \ln([A-IDP]) + b$). We also confirmed the complete reversibility of the UCST phase behavior of these A-IDPs, with a <1°C difference in UCST T_i after 10 successive heating/cooling ramps (Figure S5).

Arginine Composition, Aromatic to Aliphatic Ratio, Charge Balance and Molecular Weight Define UCST Cloud Point

To understand the effects of a particular residue substitution in the octapeptide repeat on the phase separation for the A-IDP, we created a set of “mutant” A-IDPs ranging from 100% of **a** to 100% of **b** where **a** is the WT repeat unit. The doping scheme wherein the mutant repeat unit **b** is periodically inserted into the WT sequence is visually illustrated by the color-coded schematic in Figure 2A. The mutant repeat is well mixed — distributed along the WT sequence to reduce “blockiness” of the co-polypeptide, which has been shown in LCST polypeptides to lead to nanoscale self-assembly instead of the desired liquid-liquid coacervation (MacEwan et al., 2017). Measuring the UCST phase behavior of these copolymers is analogous to a loss-of-function or gain-of-function screen for UCST cloud point upon substitution of motif **b** for **a** (Figure 2A). Due to experimental limitations, loss of function —phase separation that cannot be detected— is operationally defined as a $T_t < 4^\circ\text{C}$ at volume fractions less than or equal to 0.1 in a 140 mM salinity aqueous buffer. A total length of 40 repeats was chosen for these A-IDPs to approximate the median length (~320 amino acids) of IDRs found in the proteomic analysis of naturally occurring IDPs.

The T_t of the WT and each A-IDP is a linear function of its volume fraction (ϕ) (Figure 2B and C). At a specified ϕ , the T_t is a function of composition ($R^2 = 0.97$), which demonstrates that the behavior of the mutant A-IDPs —block co-polypeptides of **a** and **b**— can be linearly interpolated between that of pure polypeptides of **a** and **b**. The linear behavior of the T_t of these mutant A-IDPs also allows extrapolation of the UCST phase behavior for homopolymers that exhibit a UCST cloud point beyond the experimentally observable range of detection thus putting each point mutation on a single relative scale (Figure S6).

We next tested the effect of fifteen different site-specific substitution mutations of the reference —WT— repeat motif on the saturation concentration (C_{sat}) —defined as is the concentration at which the T_t is 37°C — of the A-IDPs. We found that single residue changes in the octapeptide repeat are capable of changing the C_{sat} — normalized to the degree of substitution defined by the percent change in amino acid composition— of the repeat polypeptide by over by two orders of magnitude at constant molecular weight that ranged from 1-800 μM (Figure

2D,E and F). This can be visualized by normalizing to the saturation concentration of [WT]-40 which is conveniently $\sim 1 \mu\text{M}$ and is shown by the dashed horizontal line in Figure 2F. We do not believe that changes in chain conformation as a result of these mutations is responsible for these effects on the T_t and C_{sat} . Indeed, circular dichroism spectrophotometry shows that the mutant A-IDPs are structurally disordered, consistent with their G- and P-rich composition (Bohicchio et al., 2008; Muiznieks and Keeley, 2010) (Figure S7).

These substitutions present quantitative evidence for the importance of interactions between R and aromatic residues in the repeat motif of the A-IDP. When Y_7 is substituted, we observe dramatic shifts in the UCST cloud point at $\phi = 10^{-3}$, from 66°C to 123°C , 59°C and $\sim 2^\circ\text{C}$ for W, F or H respectively. (Figure S6A). These data indicate that interactions between the cationic side-chain of R and the aromatic side-chain of W, F, Y and H are important driving forces for phase separation, although the strength of these interactions is side-group dependent with $W \gg Y > F \gg H$. Likewise, replacing R for K lowers the UCST cloud point temperature and hence the phase boundary (Figure 2E) and increases the C_{sat} (Figure 2F).

We next looked at the effect of A-IDP MW on phase behavior; we chose A-IDPs with MWs between $\sim 17\text{kDa}$ and $\sim 70\text{kDa}$, as this MW range covers 75% of the IDRs in our proteomic analysis of native IDPs (Figure S1B). Our results indicate that MW exhibits at least as large an effect on UCST cloud point as amino acid substitution (Figure 3A). We observed that the effect of MW on T_t —in the $\sim 17\text{-}70 \text{ kDa}$ range that we studied— can be approximated with a linear fit to the natural logarithm of MW (Figure 3B). By simply doubling the MW of [WT]-40, we were able to create A-IDPs with predicted C_{sat} in the nanomolar regime (Figure 3C), similar to the C_{sat} exhibited by some native IDPs (Molliex et al., 2015; Nott et al., 2015). Notably, by varying both the MW and composition *we can vary the C_{sat} of A-IDPs by over seven orders of magnitude, ranging from 10^{-4} to $10^2 \mu\text{M}$.*

In addition to composition, concentration (ϕ) and MW on T_t , there are several other parameters that have a measurable effect on UCST phase behavior but that do not eliminate UCST phase behavior under physiologically relevant conditions. Uncharged polar substitutions, the ratio of G/P, the syntax of the repeating polypeptide, solution salt content, pH (in the absence of H) and identity of the negatively charged amino acid (E

vs D) all result in smaller changes to the UCST binodal phase boundary than MW, volume fraction, aromatic:aliphatic amino acid ratio and R content (Figure S6B-C). The residue N-terminal to P₆ appears to have a unique impact on the UCST binodal boundary where compositionally identical A-IDPs shifted the UCST binodal lines depending on which polar non-charged residue is located at position 5 of the octapeptide repeat. (Figure S8A). We also produced and tested non-repetitive, but compositionally identical versions of [WT]-20 and observed minimal effects of scrambling the amino acid sequence on the UCST binodal (Figure S8B). Collectively, these results indicate that three parameters — aromatic:aliphatic ratio, volume fraction (ϕ) and MW are the most critical for controlling UCST phase boundaries or C_{sat} *in vitro*.

A-IDPs create dense phase separated condensates at saturation concentrations mediated by amino acid composition

Having observed that C_{sat} and the binodal phase boundary in the dilute regime of the UCST phase diagram of A-IDPs can be modified drastically by amino acid substitution, we were interested in the factors that modulate the high concentration regime of the phase diagram of A-IDPs. Polypeptides [WT]-20 and [Q_{5,8}]-20 express extraordinarily well for recombinant proteins, with yields of ~500 mg L⁻¹ in shaker flask culture, which made it easy to purify over one gram of material to measure the UCST cloud point behavior at high-volume fractions of these A-IDPs directly ($\phi > 0.1$). To minimize the amount of material required, these experiments were performed in a multiplexed linear temperature gradient microfluidic device mounted on an upright light microscope, wherein the T_i could be quantified by the temperature at which phase separation occurs by a visible increase in optical turbidity. These experiments produce binodal phase boundaries similar to optical turbidity measurements that are typically carried out in a UV-vis spectrophotometer (Figure S9) and demonstrate that a ~25°C difference between the two binodal lines of [WT]-20 and [Q_{5,8}]-20 is maintained over the entire range of volume fractions tested. This corresponds to an increase in A-IDP volume fraction in the dense phase (ϕ_2) from $\phi_2 = 0.4$ for [Q_{5,8}]-20 to $\phi_2 = 0.55$ for [WT]-20 at an isotherm of 37°C. In addition to these phase diagram descriptions, phase separation in the presence of low (10kDa) and high (40kDa) MW fluorescently labeled dextran indicate that both [WT]-20 and [Q_{5,8}]-20 droplets are highly exclusionary, as we observed no fluorescence partitioning of dextran into the

dense phase (Figure S10A,B). These data in combination with our ability to easily purify A-IDPs from bacterial cell lysate with phase separation indicate that A-IDPs form highly exclusionary droplets *in vitro* at physiological solution, temperature and pH conditions ($\phi_2 > 0.4$).

A-IDPs have controlled C_{sat} in eukaryotic and prokaryotic cell lines

With a set of A-IDPs that exhibit a range of T_t as a function of concentration, and C_{sat} that vary over seven orders of magnitude we sought to understand: (1) the dynamics of droplet assembly in living cells, and (2) to elucidate if it proceeds *in vivo* similarly to *in vitro*. To explore these two issues, we chose a set of IDPs that have a range of C_{sat} from 1 to 815 μM with MWs of either ~ 17 kDa or ~ 32 kDa. To visualize localization of the A-IDPs within bacterial cells, each A-IDP was genetically fused to a super folder version of green fluorescent protein (sfGFP) (Figure 4A).

Fusion of sfGFP to A-IDPs to [WT]-20, [WT]-40, [3Y7:V7]-40, [Y7:V7]-40 did not eliminate the phase behavior but shifted the phase diagram (Figure 4B & Figure S11). Despite this shift, using confocal fluorescence microscopy, we were able to observe the formation of intracellular droplets of [WT]-20-sfGFP in both transfected human embryonic kidney (HEK) cells and *E. coli* (Figure 4C and Figure 4D, respectively). Interestingly, in the *in vitro* environment of an aqueous droplet in oil, we observed that nucleation occurs at multiple points in the aqueous compartment, but with time all the individual coacervate puncta coalesce into a single, large coacervate puncta. This indicated a lack of any significant energetic barriers to diffusion or coalescence. In HEK cells nucleation of coacervate puncta also occurred at multiple locations throughout the cell. However, unlike the *in vitro* situation, these puncta never coalesced into a single coacervate droplet so that individual coacervate puncta in the 2-4 μm diameter range remained dispersed throughout the cytosol of the HEK cell (Figure S12).

In contrast, phase separation in *E. coli* is significantly different from eukaryotic cells. The initiation of the UCST phase transition in *E. coli* is similar to HEK cells—and *in vitro*—where small densely fluorescent puncta form after the A-IDP concentration in the cell exceeds C_{sat} , that then grow in size over time (Figure 4G). The growth in the size of these puncta, as more A-IDP is expressed with time, is consistent with measurements of sfGFP fluorescence from the bulk *E. coli* population normalized to the absorbance at 600 nm (OD_{600}). The

increase in fluorescence with time indicates that the intracellular concentration of the A-IDP-sfGFP fusion increases with increased protein induction time (Figure S13). Unlike HEK cells however, but similar to *in vitro* experiments, these puncta in *E. coli* coalesce to form a single coacervate droplet per cell (Table S4). This result suggests differences in the diffusivity of the A-IDP between the prokaryotic and eukaryotic cytoplasm and suggest that the barriers to diffusion and coalescence of coacervate droplets in *E. coli* are far lower than in HEK cells. Simultaneously, the residual dilute regime remains at a relatively constant concentration (Figure S14). Together, these results suggest that as the global concentration of the A-IDP within the cell increases with time, the cytoplasmic concentration of the protein is buffered —remains constant— but that the volume of the coacervate increases relative to the size of the cell.

Similar to *in vitro*, the MW and aromatic:aliphatic content affect droplet formation in *E. coli*. Doubling the MW of [WT]-20-sfGFP to [WT]-40-sfGFP decreases C_{sat} enough to cause droplet formation even prior to A-IDP induction, presumably because of leaky transcriptional regulation (Figure 3D). Similarly, increasing the aliphatic content with V at the expense of Y increases the C_{sat} to a concentration that is not measurable in the time course of these experiments (Figure 4E). Although differences in C_{sat} *in vivo* are not as dramatic as predicted by *in vitro* experiments with A-IDP-sfGFP fusions, perhaps because of the effect of intramolecular crowding within the cell, we can modulate the intracellular C_{sat} by at least an order of magnitude using both MW and aromatic:aliphatic ratio of the A-IDP.

A-IDPs exhibit reversible UCST droplet formation in *E. coli*

Just as one can cross a binodal line into the two-phase regime under isothermal conditions by increasing polypeptide volume fraction, this line may be crossed under constant volume fractions by decreasing solvent quality or the chi parameter (χ) (Brangwynne et al., 2015; Rubinstein and Colby, 2003). Experimentally this is most easily accomplished by reducing the temperature of the bulk solution (Rubinstein and Colby, 2003). Similar to the UCST phase behavior of A-IDPs *in vitro*, A-IDPs exhibit reversible UCST phase separation inside cells that is reversible by repeated four cooling and heating cycles (Figure 5A). The phase separation exhibits minimal

hysteresis as the difference in the transition temperature of cooling ($T_{t,C}$) and the transition temperature of heating ($T_{t,H}$) varies by less than 2°C (Figure 5B).

Interestingly, upon multiple heating and cooling cycles, we observed that *E. coli* exhibit spatial phase separation memory, with puncta forming in the same location as the first cycle (Figure S15). Additionally, we observed cooling-triggered phase separation results in a higher number of puncta per cell (Figure S16). The greater number of puncta observed with higher MW species indicates that the number of puncta formed per cell is a function of their diffusion coefficient, consistent with prior studies (Bracha et al., 2018).

Increasing the MW of the A-IDP increases the observed T_t in *E. coli* (Figure 5C). By manipulating the aromatic:aliphatic ratio while keeping MW constant and observing the formation of puncta within individual bacterium at various times post-induction (varying intracellular concentration), we were able to create partial intracellular phase diagrams as a function of T_t and intracellular fluorescence (Figure 5D). This result is important because it ties the observed behavior upon cooling to a specific concentration for a given construct, essentially normalizing the observed cloud point for differing overall levels of protein expression throughout the cell population. Again, with increasing concentration we see an increase in UCST cloud point, although the rate of increase upon increasing concentration does not appear to follow a log-normal dependence.

***De Novo* Design of Functional A-IDP Droplets in Cells**

In order to understand the potential of using spatially confined intracellular coacervate droplets to carry out new functions, we asked the following questions: (1) Can coacervate droplets in cells recruit other molecules, and if so, what, if any, are the size limitations of such molecules? (2) Can these molecules interact with the A-IDP to impart a new function to the droplet?

To answer these questions, we first examined whether a small molecule can diffuse into and react with the A-IDP in a coacervate droplet located within an *E. coli* cell. We designed and expressed an A-IDP — [3Y₇:V₇]-40-UAA — that carries three copies of a unique biorthogonal reactive group — an azide; its primary amino acid sequence is listed in Table S2. After reaching intracellular concentrations greater than the C_{sat} of [3Y₇:V₇]-40-UAA, we incubated live *E. coli* with 1 mg ml⁻¹ dibenzocyclooctyne-dye (DBCO-Alexa488) for 10

min (Figure 6A). After a single wash step to remove excess dye, we observed fluorescent condensates in the cells. These experiments also demonstrate that the ϕ' fraction is also labeled in addition to the condensates. These data clearly show that a small molecule can diffuse from the extracellular environment into preformed condensates created by A-IDP phase separation and react with the A-IDP.

Next, we asked if larger molecules such as proteins are also capable of interacting with an A-IDP puncta. To answer this question, we designed a droplet capture experiment based on split green fluorescent protein (GFP). We first verified if the two components of a split GFP can interact with each other to create a functional GFP molecule if one of the components is fused to an A-IDP. GFP-11-[3Y7:V7]-40-mRuby3 was co-expressed in the presence of GFP-1-10; because the IDP is fused to mRuby3, the A-IDP condensates fluoresce red and can be visualized by fluorescence microscopy within the cell. We see fluorescently active GFP only in the interior of the condensates as seen by the co-localization of green fluorescence with the red fluorescence from the A-IDP condensates, indicating that the fragments GFP bind to each other to create an intact and functional GFP molecules that fluoresces green (Figure S17A). In contrast, in the absence of GFP-1-10 induction, there is minimal green fluorescent inside the red fluorescent condensates (Figure S17B).

These results show that two protein fragments of GFP can find and bind to each other in the cell despite the steric hindrance imposed by an A-IDP and a fluorescent reporter fused to one fragment of the protein. It does not however, prove that a protein can be recruited after an A-IDP condensate has formed, as the protein partners in the previous experiment are co-expressed and could bind in the cytoplasm prior to phase separation that occurs once the intracellular concentration of GFP-11-[3Y7:V7]-40-mRuby3 exceeds its C_{sat} . To directly answer this question, we co-transformed *E. coli* with two plasmids — a Lac operon regulated plasmid that encodes one fragment of GFP (GFP-11) that is fused to [3Y7:V7]-40 and a second plasmid regulated by araBAD operon that encodes the other fragment of GFP (GFP-1-10). Once expression of GFP-11-[3Y7:V7]-40 at 37 °C proceeds long enough that its intracellular concentrations is greater than its C_{sat} , we removed the IPTG induction media, and replaced it with arabinose containing media that induce the expression of the larger GFP fragment (GFP-1-10). We observed that subsequent to arabinose induction, both the ϕ' and ϕ'' fractions of the *E. coli* contained

fluorescently active GFP (Figure 6B). This result suggested that the large GFP fragment is capable of penetrating the preformed condensate in the cell, find its binding partner and form a fully functional molecule, despite the fusion to the A-IDP. Once a fully functional GFP molecule is recruited into the intracellular droplets, it is then possible to dynamically modify the intracellular solubility of the reconstituted GFP–A-IDP by changing the temperature of the bulk (Figure S18).

These experiments clearly show that small molecules and proteins can be recruited into intracellular coacervate droplets in *E. coli* and that a protein can be reconstituted within a coacervate droplet. These results suggested a path for the *de novo* design of intracellular coacervate droplets with new enzymatic function. We chose biocatalysis as the function of interest, because one of the proposed reasons for the evolutionary development of biomolecular condensates is to modulate the kinetics of various biological functions, including enzymatic reactions (Banani et al., 2017; Banjade and Rosen, 2014; Li et al., 2012; Strulson et al., 2012). However, there is little experimental evidence demonstrating how the function of enzymes is modulated by phase separation.

To investigate this, we created an A-IDP fusion that can recruit an enzyme into intracellular droplets to modulate its catalytic activity. We chose β -galactosidase for two reasons: (1) it has a range of small molecule substrates, one of which, Fluorescein Di β -Galactopyranoside (FDG), is colorless but when cleaved by β -galactosidase, will fluoresce green. Thus, using a combination of a red fluorescent protein tagged to our enzyme-A-IDP fusion and fluorescein fluorescence we can track the colocalization of enzymatic reactions with A-IDPs in real time. (2) We had concerns that a large enzyme fused to a large A-IDP would not express at high enough concentrations in *E. coli* and thus not phase separate *in vivo*. To alleviate this concern, we took advantage of the widely used β -galactosidase (LacZ) blue-white screening system, where the so called alpha peptide (α p) complements the mutated enzyme LacZ Δ M15 to create a functional β -galactosidase enzyme. In our system, the α p is fused to a A-IDP-mRuby3 construct, such that enzyme activity is physically linked to the A-IDP which in turn is physically linked to red fluorescence.

Our studies with the DBCO-Alexa488 and split GFP provided the basis for this more complicated experiment. The DBCO-Alexa488 experiment suggested that a small molecule such as an enzyme substrate can penetrate puncta, even if delivered extracellularly (Figure 6A). The split GFP experiment suggested that relatively large proteins can be recruited A-IDP condensates to form a functional protein, suggesting that the same should be possible with the split β -galactosidase system (Figure 6B). This peptide binding system also represents a more ubiquitous, engineered puncta platform as there are a number of split enzyme systems or small protein motifs that have been engineered to bind various intracellular targets.

Thus, we genetically fused the α -peptide (α p) from LacZ β -galactosidase to a A-IDP-mRuby3 construct. Our hypothesis is that the α p-A-IDP-mRuby3 protein can bind and recruit the other fragment of the enzyme — LacZ Δ M15 that has an α -peptide deletion— that is expressed endogenously in genetically modified *E. coli* (KRX, Promega) into intracellular droplets. After protein induction and resulting condensate formation, we deliver the substrate Fluorescein Di β -Galactopyranoside (FDG) to the cell medium where it is trafficked intracellularly, hydrolyzed into fluorescein at the sites of active β -galactosidase, and eventually exported outside the cell (Figure 6C) (Hofmann and Sernetz, 1983; Lewis et al., 1994). By tracking the onset of the green fluorescence of fluorescein with confocal microscopy we can specifically observe where and when enzymatic activity is occurring within the cell and quantitatively track enzyme activity.

In our control experiment – α p-mRuby3 – we observe limited persistence of fluorescence within the cells. It is important to note that the α -peptide itself is known to form inclusion bodies (Broome et al., 2010; Moosmann and Rusconi, 1996), and therefore, even in this control experiment, we observe some puncta inside the bacterial cells. However, upon fusion with [WT]-20, we observe that the fluorescence localizes long enough with the A-IDP condensates to be observed with confocal microscopy (Figure 6D). Despite this increased colocalization, the total fluorescent production over time is not statistically significant from the α p-mRuby3 control (Figure 6E).

When we increase the MW of the A-IDP, and thus decrease C_{sat} , we observe a dose-response effect in the total FDG fluorescence intensity as well as colocalization with the α p-A-IDP-mRuby3 fusion (Figure 6D). α p-[WT]-40-mRuby3 and α p-[WT]-80-mRuby3 have 2.5X and 7.5X greater FDG converted at 20 minutes compared

with the α p-mRuby3 control (Figure 6D and Figure S19). Quantification of the colocalization of green and red fluorescence with Mander's overlap coefficient (Manders et al., 1992) indicates increased colocalization when the α -peptide is fused to A-IDPs compared to the fluorescent reporter alone (Figure S20). To quantify the observed colocalization, we analyzed individual cells within the image frame with green fluorescence that was above the background threshold at each timepoint. Higher MW A-IDPs exhibit higher fluorescence inside the cell normalized to the background at each point in time (Figure 6F). This dose-response effect of MW emphasizes a mechanism of increased persistence of the substrate molecule inside the droplets leading to more efficient green fluorescence conversion.

Quantification of fluorescence production at various substrate concentrations *in vitro* suggests that the mechanism of this effect is a statistically significant increase in the catalytic constant (K_{cat}) of the enzyme with increasing MW. This constant can be interpreted as the "turnover efficiency" of the enzyme or the number of catalytic events that occur per unit time. We observed 1.4X, 1.6X and 4.2X increase in the K_{cat} for α p-[WT]-20-mRuby3, α p-[WT]-40-mRuby3, α p-[WT]-80-mRuby3 compared to the α p-mRuby3 control (Figure S21 and Table S5). Considering our previous observation of increased colocalization of product (fluorescein) and the labeled A-IDP as a function of MW, we propose that the observed increase in fluorescence is caused by increased colocalization of the enzyme and substrate in the condensates, leading to a higher measured K_{cat} . We observed non-significant changes to the Michaelis-Menten constant (K_m) which describes the affinity of the enzyme for the substrate, suggesting that fusion of the A-IDP does not change the binding constant of the enzyme-substrate complex. Using K_{cat} and K_m , we can define a catalytic efficiency which also supports our hypothesis of an increase in the enzymes' efficiency within condensates with increasing MW of the A-IDP. This enhancement in enzymatic efficiency is on the order of magnitude of change observed by various protein engineering techniques used primarily to improve K_{cat} (Duan et al., 2013; Goldsmith and Tawfik, 2017; Nayeem et al., 2009).

We also fused the LacZ alpha peptide to A-IDPs with differing levels of aromatic content at a constant MW (Figure S22A). We hypothesized that differing levels of aromatic content would affect FDG uptake into the droplets and therefore affect overall enzymatic activity. Surprisingly, we observed similar overall levels of

fluorescence between α p-[WT]-40-mRuby3, α p-[3Y7:V7]-40-mRuby3 and α p-[Y7:V7]-40-mRuby3. However, the dynamics of enzymatic activity are different, with A-IDPs of greater aliphatic content allowing for faster uptake into the condensates (Figure S22B). The differences between the ratio of FDG fluorescence inside the cell and outside the cell between α p-A-IDP-mRuby3 fusions with different aliphatic content were insignificant, indicating that MW is the primary driving force for fluorescein and/or FDG persistence inside intracellular droplets (Figure S22C). Completely deleting the aromatic residues from the repeat unit of the A-IDP results in apparently soluble enzymes that do not form intracellular condensates, and whose activity is lower than the enzyme formed by complementation of LacZ Δ M15 with the α p-mRuby3 fusion that has no A-IDP tag (Figure S23A,B).

Discussion

We show herein A-IDPs that consist of repeats of an octapeptide motif inspired by native IDP exhibit reversible UCST phase separation in aqueous solution. Despite the simplicity of their sequence, they recapitulate many of the features seen in more complex, native IDPs. The formation and dynamics of their phase separation into coacervate droplets are controlled by two simple design parameters that are genetically encodable at the sequence level—MW of the A-IDP and the ratio of aromatic:aliphatic residues in the octapeptide repeat. Using these two parameters—aromatic:aliphatic ratio and MW—we were able to produce A-IDPs with C_{sats} ranging from nanomolar to millimolar concentrations. This work supports the growing evidence of R-aromatic interactions that drive phase behavior (Nott et al., 2015; Wang et al., 2018) and also adds additional evidence of the molecular hierarchy that exists between the aromatic groups W, Y, F and H in modulating UCST phase behavior. Although the IDP literature often ignores the importance of MW, our results suggest that MW may be more critical than composition for defining the UCST binodal. We anticipate that these results will inform and dramatically shift the strategy for mutating native IDPs and designing *de novo* IDPs.

These design parameters faithfully translate from *in vitro* to intracellular environments. The A-IDPs phase separate inside cells by the same principles that drive their UCST phase separation *in vitro* indicating that the same thermodynamic driving forces embedded in the sequence and molecular weight also modulate droplet

formation dynamics in isolation. Due to the simplicity of their design, A-IDPs behave *in vivo* as their phase diagrams *in vitro* suggest—as their intracellular concentration increases to a C_{sat} , small phase separating droplets form at individual points in space that continue to grow in size with increasing overall A-IDP concentration inside the cell. This predictable observation has been theorized by previous studies but has not been conclusively demonstrated until now (Holehouse and Pappu, 2018).

Finally, these proteins can be used for the *de novo* design of functional intracellular droplets. We rationally designed intracellular puncta capable of binding and recruiting a β -galactosidase deletion mutant, which could modify the catalytic efficiency of the enzyme-substrate complex—a complex which has not evolved to form intracellular condensates. The catalytic efficiency of the reconstituted enzyme in phase separated coacervate droplets is MW dependent, and increases with the MW of the A-IDP. Higher MW A-IDPs more efficiently sequester the substrate in the enzymatically active, intracellular phase separated puncta, which results in a higher catalytic efficiency as measured by K_{cat} . These proof-of-concept experiments demonstrate that intracellular droplets can be engineered to have non-canonical functions in live cells and provide a new platform for intracellular material manipulation. In summary, with over 60 IDPs synthesized for this study that span a range of C_{sat} , and proof of concept experiments recruiting proteins into coacervate droplets within a cell and thereby modulating protein function, these studies lay the groundwork for the *de novo* design of functional intracellular condensates. We expect that these A-IDPs will be useful as building blocks from which new biological condensates with emergent behaviors can be built within living cells to better study the functional significance of phase separation in living cells and to encode new functions for droplets within cells. We also anticipate that these IDPs will prove useful in other biomedical applications beyond the design of intracellular droplets that can profit from their tunable UCST phase behavior. This marriage of soft material science with biophysical characterization of subcellular materials will continue to be an exciting space for to engineer cells with new or improved function and new biomaterials.

Acknowledgements

A.C. acknowledges support from the National Institutes of Health (NIH) through a MIRA grant R35GM127042 and from the National Science Foundation (NSF) through a DMREF grant DMR-17-29671. P.S.C acknowledges support from the National Science Foundation through CHE-1709735.

Author Contributions

M.D. & A.C. designed the experiments and wrote the manuscript. M.D. performed each measurement and analyzed the data with the exception of microfluidic temperature gradient experiments that were performed and analyzed by B.A.R. A.S. assisted with UCST cloud point determination and non-repetitive A-IDP design. P.S.C. assisted in writing the manuscript.

Declaration of Interests

We report no conflicts of interest.

Supplemental Information

Supplemental Information related to the main text is available free of charge at Cell Publications website.

Table 1 – Amino Acid Sequences of A-IDPs Used in This Study

Protein Name	Full Amino Acid Sequence	Amino Acids (N)	Molecular Weight (Da)
[WT]-20	SKGP-[GRGDSPYS] ₂₀ -GY	166	17004
[WT]-40	SKGP-[GRGDSPYS] ₄₀ -GY	326	33400
[WT]-60	SKGP-[GRGDSPYS] ₆₀ -GY	486	49797
[WT]-80	SKGP-[GRGDSPYS] ₈₀ -GY	646	66193
[Q _{5,8}]-20	SKGP-[GRGDQPYQ] ₂₀ -GY	166	18646
[Q _{5,8}]-40	SKGP-[GRGDQPYQ] ₄₀ -GY	326	36685
[3S _{5,8} :Q _{5,8}]-40	SKGP-[GRGDSPYSGRGDSPYSGRGDSPYSGRGDQPYQ] ₁₀ -GY	326	34221
[S _{5,8} :Q _{5,8}]-40	SKGP-[GRGDSPYSGRGDQPYQ] ₂₀ -GY	326	35042
[S _{5,8} :3Q _{5,8}]-40	SKGP-[GRGDQPYQGRGDQPYQGRGDQPYQGRGDSPYS] ₁₀ -GY	326	35863
[3Y ₇ :V ₇]-40	SKGP-[GRGDSPYSGRGDSPYSGRGDSPYSGRGDSPVS] ₁₀ -GY	326	32760
[Y ₇ :V ₇]-40	SKGP-[GRGDSPYSGRGDSPVS] ₂₀ -GY	326	32119
[V ₇]-40	SKGP-[GRGDSPVS] ₄₀ -GY	326	30839
[3R ₂ :K ₂]-40	SKGP-[GRGDSPYSGRGDSPYSGRGDSPYSGKGDSPYS] ₁₀ -GY	326	33120
[R ₂ :K ₂]-40	SKGP-[GRGDSPYSGKGDSPYS] ₂₀ -GY	326	32840

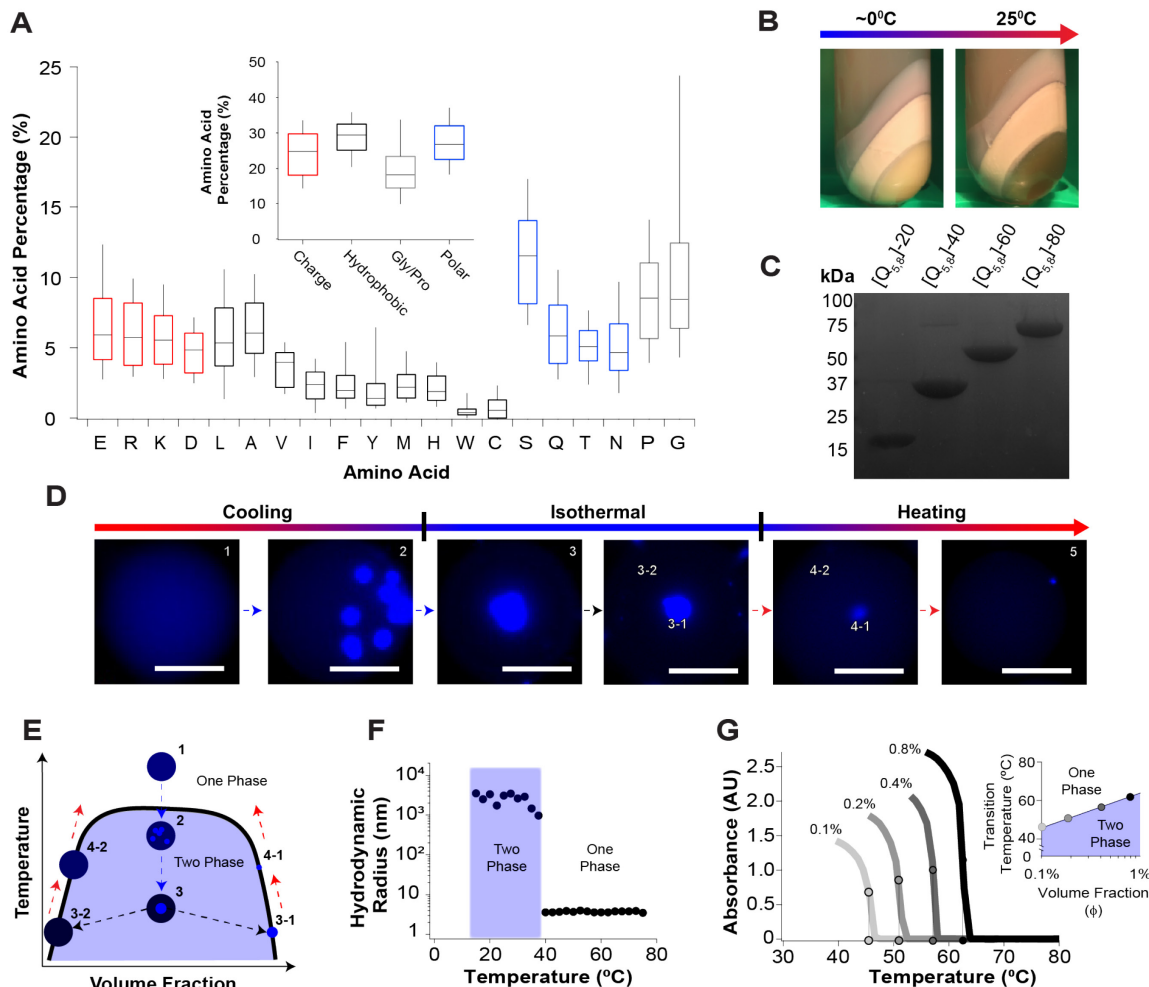


Figure 1 - Artificial Intrinsically Disordered Polypeptides (A-IDPs) Inspired from Native IDPs Exhibit Reversible UCST Phase Behavior. **A.** Proteomic analysis of native IDPs that form biomolecular condensates reveal that they have an abundance of G/P, charged and uncharged polar residues, yet exhibit a balance of overall charge. **B.** An example of a dense, exclusionary phase formed by an UCST exhibiting A-IDP even in the complex medium of bacterial cell lysate. The coacervate shows observes almost complete separation from all other cellular proteins and debris present in the cell lysate after centrifugation, facilitating purification from the insoluble cell lysate fraction without affinity tags. **C.** Example SDS-PAGE gel of a set of A-IDPs — [Q_{5,8}]-20 to [Q_{5,8}]-80 — with conserved sequence but increasing MW that show the high purity of the A-IDPs that is obtained by exploiting their UCST phase behavior without need for any chromatographic purification. **D.** Visualization of UCST phase separation of [Q_{5,8}]-20 in water-in-oil droplets with fluorescence microscopy. Upon cooling, phase separation in a droplet is initiated at multiple sites, the puncta that grow from each site slowly coalesce with one another into a single dense phase. Upon reheating, equilibrium with the surrounding dilute phase is constantly re-established, leading to a higher concentration dilute phase and smaller volume occupied by the dense phase. $\phi = 0.0018$, scale bar = 50 μm . **E.** Schematic UCST phase diagram for a cooling-heating cycle of a UCST polypeptide in a water-in-oil droplet. **F.** Dynamic light scattering data of [Q_{5,8}]-20 demonstrating the change in hydrodynamic radius upon cooling. Upon reaching the cloud point, [Q_{5,8}]-20 transitions from soluble unimeric polypeptides with a radius of hydration of 5-6 nm to micron-sized aggregates. Data collected at $\phi = 0.0043$ in 140 mM PBS, pH 7.4. **G.** UCST cloud points are affected by polypeptide volume fraction in solution. This behavior follows a natural log dependence in the dilute regime ($R^2 = 0.98$).

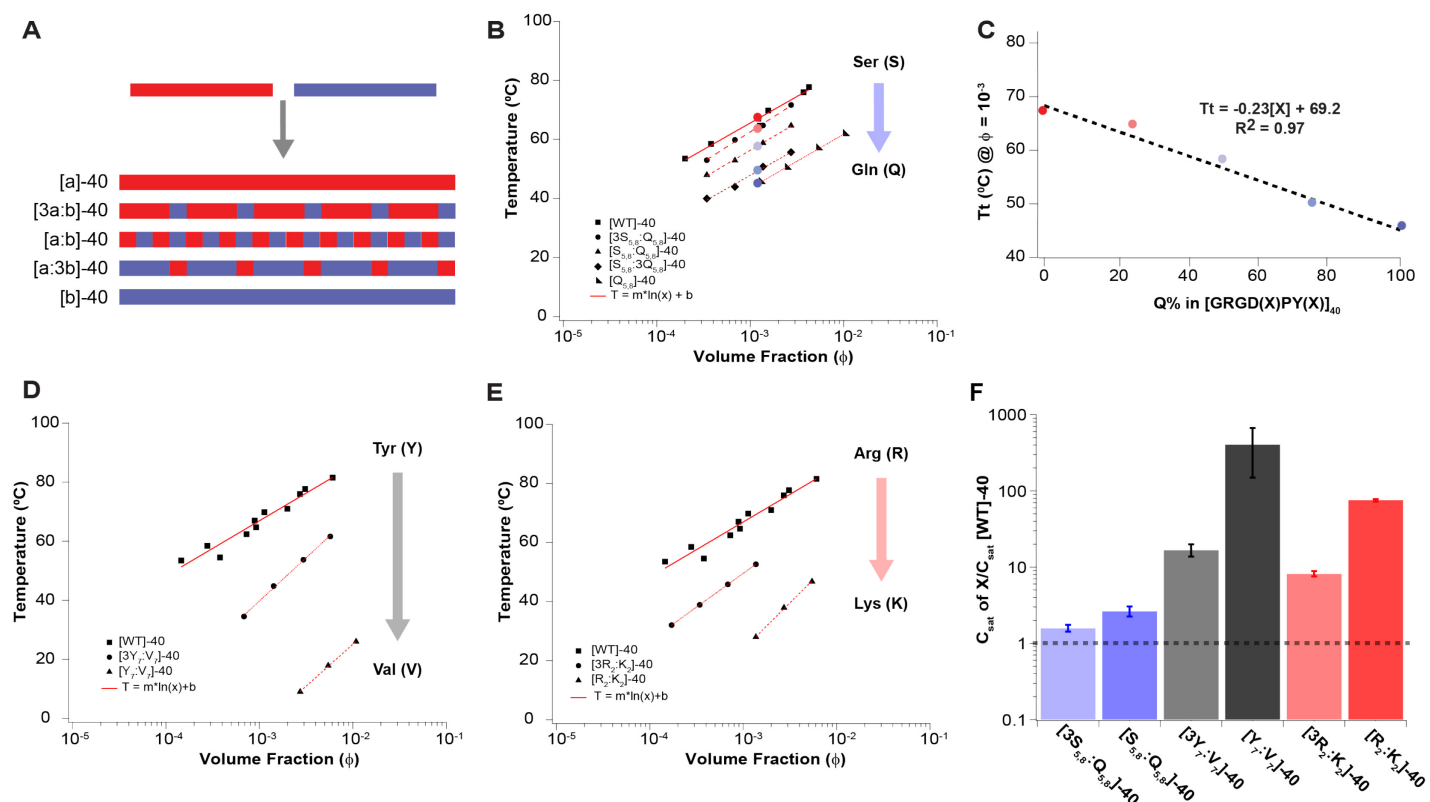


Figure 2 – Control of UCST Cloud Point Using Main Chain Amino Acid Composition. **A.** Schematic describing the methodology for doping repeat unit **b** into a homopolymer of **a**. The WT A-IDP with a high UCST cloud point consisting of 40 repeats of **a**—GRGDSPYS—is doped with increasing fractions of repeat **b**—GRGDQPYS—to probe “loss of function” of UCST phase behavior of polymers of **a**. The doping of **b** into **a** is designed to ensure mixing of the two repeats along the polypeptide chain and minimize blocky behavior. **B.** Doping of **b** into **a** results in mutant IDPs; the UCST cloud point temperature (T_t) of each mutant IDP is a linear function of volume fraction (ϕ) of the A-IDP. **C.** The effect of composition—degree of doping—is a linear function of the degree of substitution of **b** into **a** at constant volume fraction of 10^{-3} ($R^2=0.97$). **D.** Substitution of aromatic Y residues with aliphatic V dramatically reduces T_t . **E.** Substitution of R with K dramatically reduces T_t . A 50% substitution of K for R lowers the T_t by more than 40°C . **F.** The chemical composition can affect the saturation concentration by two orders of magnitude at constant molecular weight ($C_{\text{sat}} @ 37^{\circ}\text{C} = 1\text{--}800 \mu\text{M}$). This can be visualized by normalizing to the saturation concentration of [WT]-40 which is conveniently $\sim 1 \mu\text{M}$ and is shown by the dashed horizontal line.

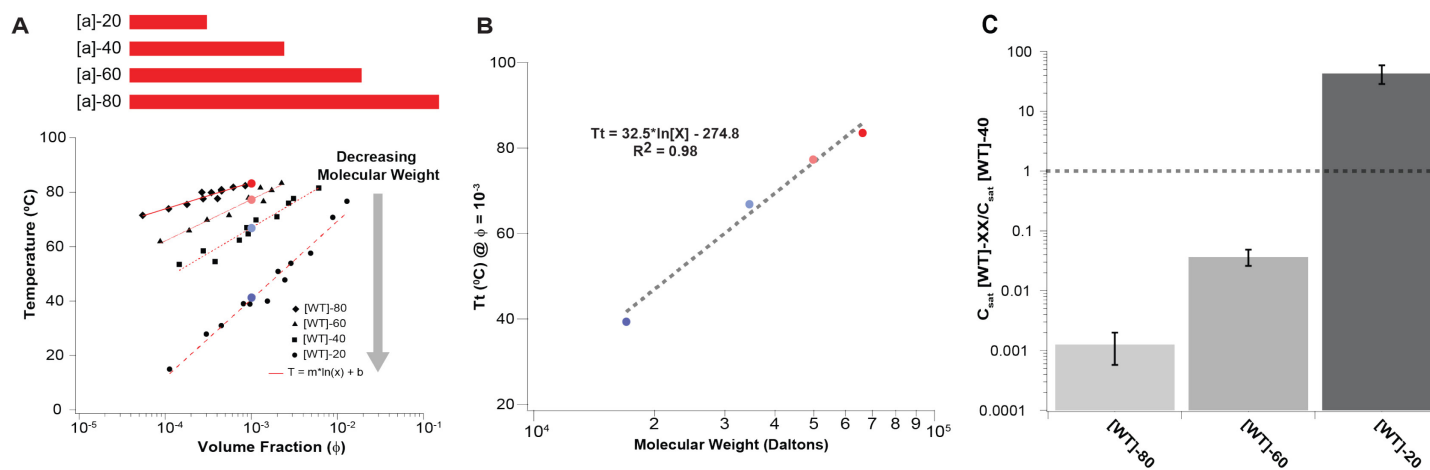


Figure 3 - Control of UCST Cloud Point by Molecular Weight of A-IDP. **A.** The molecular weight of the polypeptide affects the T_t . **B.** The T_t directly scales with the natural log of MW. **C.** At constant chemical composition, it is possible to modulate C_{sat} by over five orders of magnitude simply by changing the MW of the A-IDP (C_{sat} @ 37°C = 1 nM – 400 μ M). [WT]-40 has a C_{sat} of ~1 μ M.

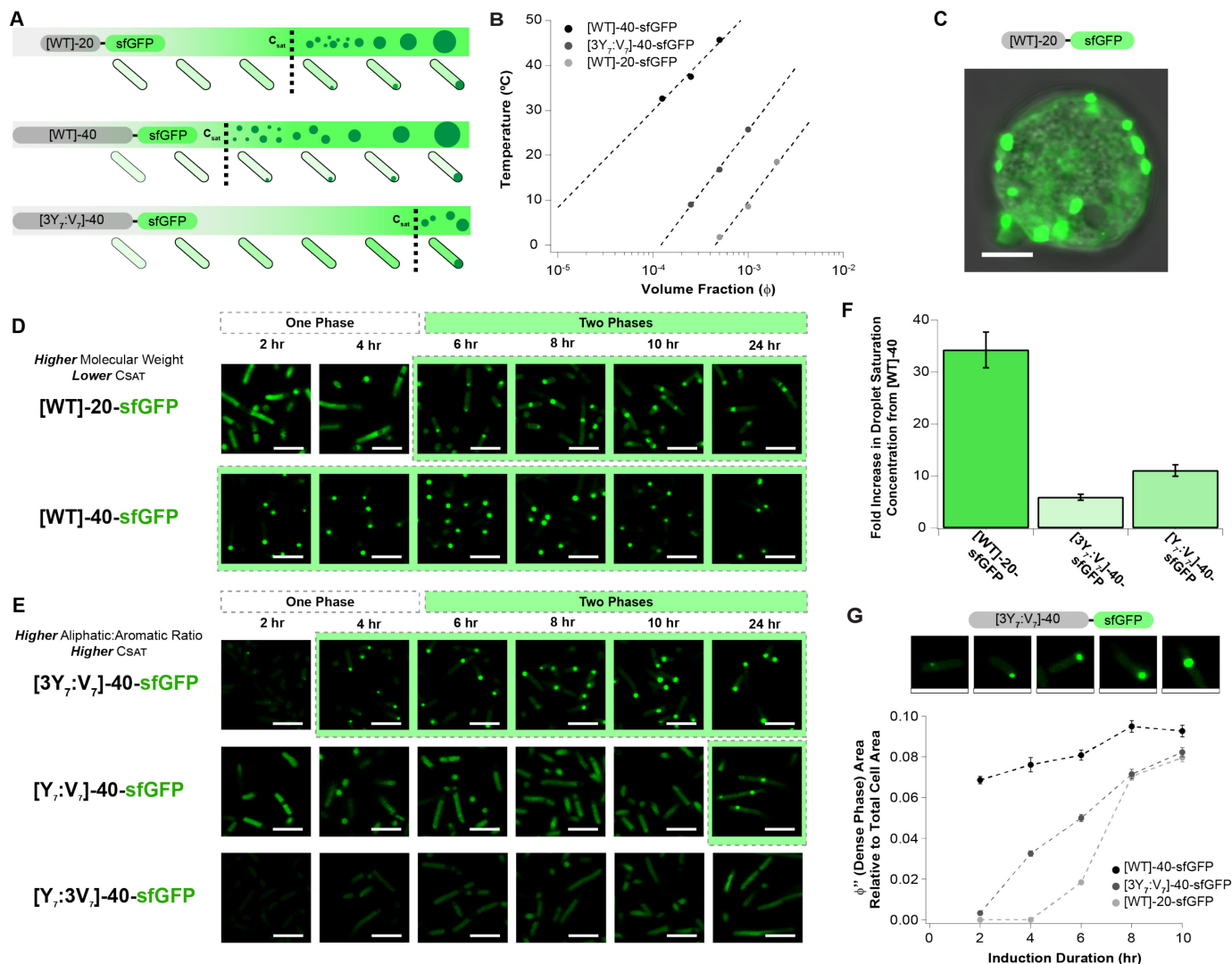


Figure 4 – A-IDPs Exhibit Tunable Intracellular Droplet Formation Based on Molecular Weight and Ratio of Aromatic:Aliphatic Content. All scale bars are 5 μ m. **A.** Schematic describing the use of two key parameters —ratio of aromatic:aliphatic content and molecular weight— to control intracellular droplet formation by modulating C_{sat} . **B.** Partial *in vitro* binodal of A-IDP-sfGFP fusions in the dilute regime in 140 mM PBS, pH 7.4. Similar to A-IDPs, A-IDP-GFP fusion proteins exhibit molecular weight and aromatic content dependent phase behavior. **C.** [WT]-20-sfGFP fusion phase separates in eukaryotic cells (HEK293 cells, Day 5). Instead of forming a single droplet as seen *in vitro* in protocells (see Figure 1C), many distinct droplets are formed indicating either diffusion-limited or arrest-limited coalescence. **D.** Confocal fluorescence images of A-IDP-sfGFP as a function of induction time and molecular weight in *E. coli*. A higher intracellular concentration is required for [WT]-20 versus [WT]-40 to form intracellular droplets. It is noticeable that [WT]-40 has a lower ϕ' —A-IDP poor— soluble phase outside the dense droplet phase compared with [WT]-20. **E.** Reducing the aromatic content increases the C_{sat} in a dose-dependent manner. **F.** A-IDP-sfGFP fusions exhibit a one order of magnitude shift in their C_{sat} as determined by their molecular weight and ratio of aromatic:aliphatic content. **G.** Size of intracellular droplets (ϕ'' or dense phase) grow with induction time. As concentration of the A-IDP-sfGFP increases inside the cell, the soluble concentration outside the droplet does not change (Figure S14) but the size of the intracellular droplets grows relative to the total cell area. Images are individual cells from [3Y₇:V₇]-40-sfGFP cultures at various time points. Error bars represent standard error of the mean.

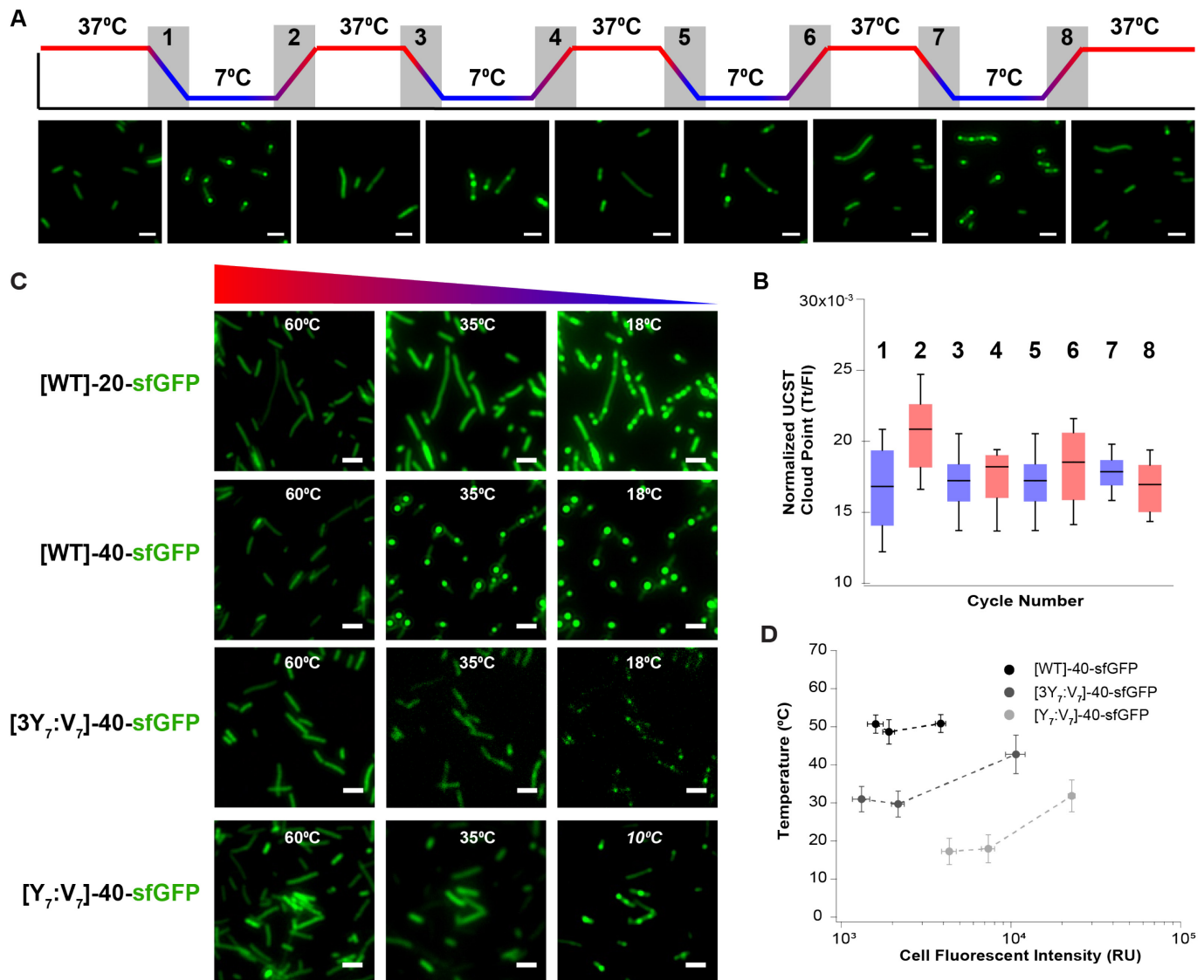


Figure 5 – A-IDs Exhibit Reversible Coacervation in *E. coli* Determined by their Molecular Weight and Aromatic:Aliphatic Ratio. All scale bars are 5 μ m. **A.** Intracellular droplets comprised of [WT]-20-sfGFP can be formed and dissolved reversibly via alternating cooling and heating cycles. This process is completely reversible over four rounds of cooling and heating. Cooling rate = 5°C min⁻¹, induction time of 4 hr. **B.** T_i normalized to the intracellular fluorescence of sfGFP in each individual cell (n = 30) does not change significantly over four heating (red bars) and cooling (blue bars) cycles. Boxes indicate 25th -75th percentile. **C.** The intracellular T_i—similar to *in vitro*— is a function of A-IDP molecular weight and aromatic content. Cooling ramp = 60°C $\rightarrow$ 10°C. Cooling rate = 5°C min⁻¹, A-IDP gene induction time of 8 hr. Whiskers indicate 10th -90th percentile. **D.** Intracellular binodal lines of various A-IDP–sfGFP fusions. T_i increases as a function of cellular fluorescence, a surrogate of A-IDP concentration, and aromatic content of the A-IDP. Data analyzed at 2, 4, 8 hr for [WT]-40-sfGFP and [3Y₇:V₇]-10 and 4, 8, 24 hr for [Y₇:V₇]-40 (n = 30). Error bars indicate standard error of the mean. **F.** Upon reconstitution of sfGFP in the dense phase, the solubility of the reconstituted GFP–A-IDP complex can be modulated with temperature. Data was collected for 36 hr post-IPTG induction and 12 hours post-arabinose induction.

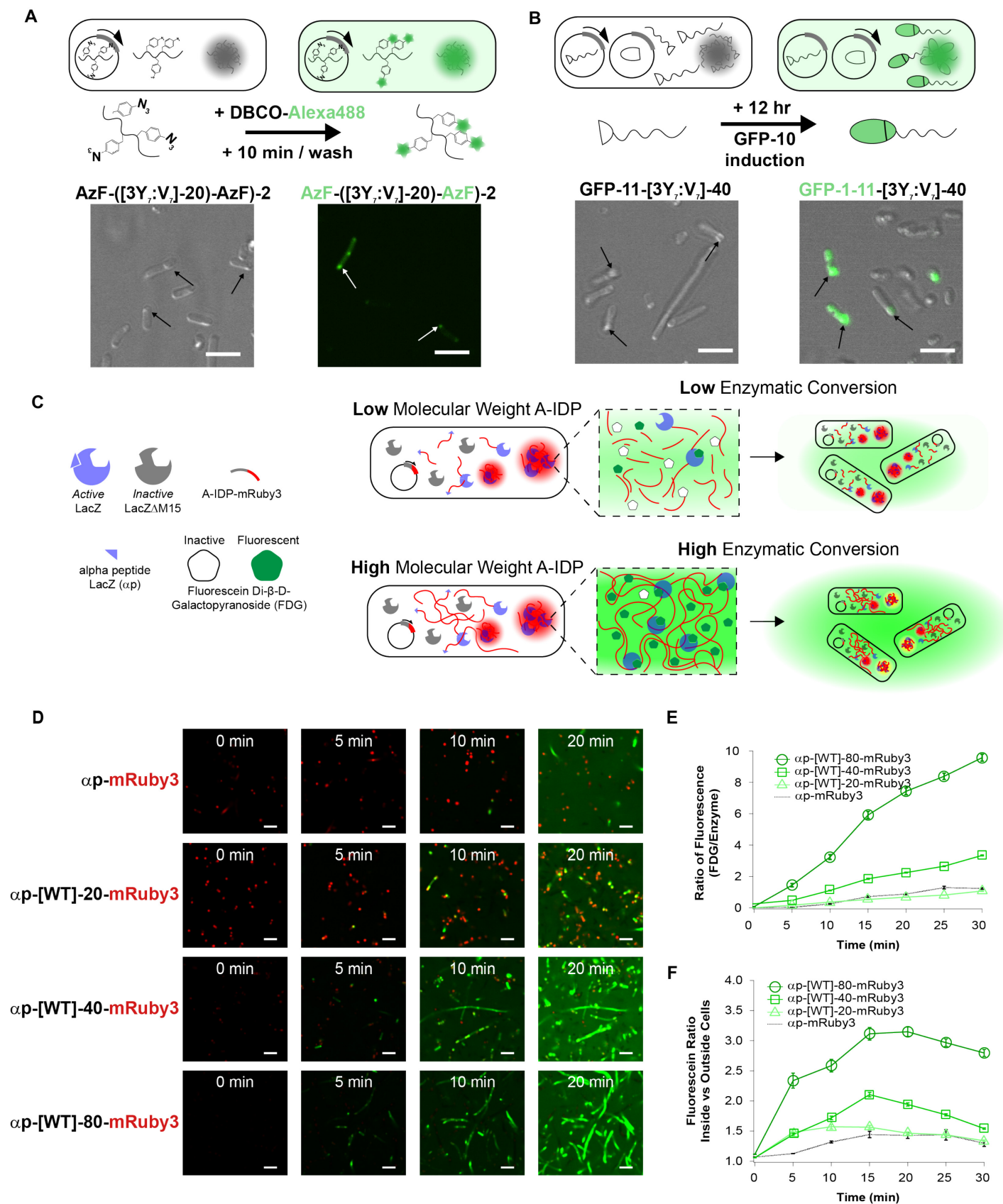


Figure 6 – Engineered Intracellular Droplets with Programmable Function. **A.** Site specific labeling of droplets with a small molecule fluorescent dye. *E. coli* cells containing condensates formed by a [3Y7:V7]-40 variant with azido-phenylalanine (AzF) residues that presents a biorthogonal azide which can be labeled *in situ*

with a Dibenzocyclooctyne-dye conjugate (DBCO-Alexa488). DBCO-Alexa488 mixture can diffuse into cells and into the A-IDP condensates within the cell, labeling the azide groups within 10 min of incubation with live *E. coli*. **B.** Reconstitution of function GFP in condensates by recruitment of a partner from the cytoplasm using a split GFP system. A GFP-11-[3Y7:V7]-40 fusion protein is able to recruit GFP-1-10 from the surrounding cytoplasm into intracellular droplets. Upon formation of intracellular condensates after 24 hr of IPTG induction of GFP-11-[3Y7:V7]-40 (left panel), subsequent induction GFP-1-10 by arabinose induction enables recruitment of GFP-10 into the condensates and reconstitution of functional sfGFP within existing intracellular condensates within 12 hr of GFP-1-10 induction (right panel). **C.** Schematic of enzyme-condensate experiment. The α -peptide (α p) of LacZ is fused to a fluorescent reporter protein (mRuby3) and expressed from an IPTG-inducible gene from a plasmid in the *E. coli* strain KRX that has a deletion mutant of the LacZ gene that produces a truncated, catalytically inactive enzyme lacking the α p. Complementation of Lac Δ M15 by a α p-A-IDP-mRuby3 fusion creates an active enzyme that converts FDG into fluorescein that is then rapidly exported from the intracellular space into the surrounding medium **D.** Confocal microscopy images showing the fluorescent conversion of fluorescein Di- β -D-galactopyranoside (FDG). Note that the puncta-like structures of α p-mRuby3 in the top panel are due to a fraction of the fusion forming inclusion bodies in cells. When the α p is fused to [WT]-20-mRuby3 the fluorescence is first observed at the sites of intracellular phase transition in coacervate droplets, and the fluorescein then diffuses into the cytosol and then out of the cell into the extracellular space. Increasing the molecular weight of the A-IDP leads to increased FDG conversion at earlier timepoints and higher overall conversion after 20 min. Rebalanced images of α p-[WT]-40-mRuby3 and α p-[WT]-80-mRuby3 can be found in Figure S19 for improved visualization of the colocalization of intracellular droplets and converted FDG. **E.** Intracellular concentration of fluorescein produced by catalytic conversion of FDG, normalized to the mRuby3 fluorescence of each individual cell ($n \approx 300$). The catalytic efficiency increases with A-IDP MW, as seen by the greater ratio of green fluorescence resulting from FDG conversion to fluorescein normalized to the red fluorescence of mRuby3 on a molar basis. Both α p-[WT]-40-mRuby3 and α p-[WT]-80-mRuby3 exhibit statistically significant differences from the control (α p-mRuby3). Error bars indicate standard error of the mean. **F.** All α p-A-IDP-mRuby3 fusions exhibit a higher ratio of green fluorescence inside the cell, indicating a greater persistence of fluorescent FDG inside the intracellular space compared to the α p-mRuby3 control. Error bars indicate standard error of the mean. All scale bars are 5 μ m.

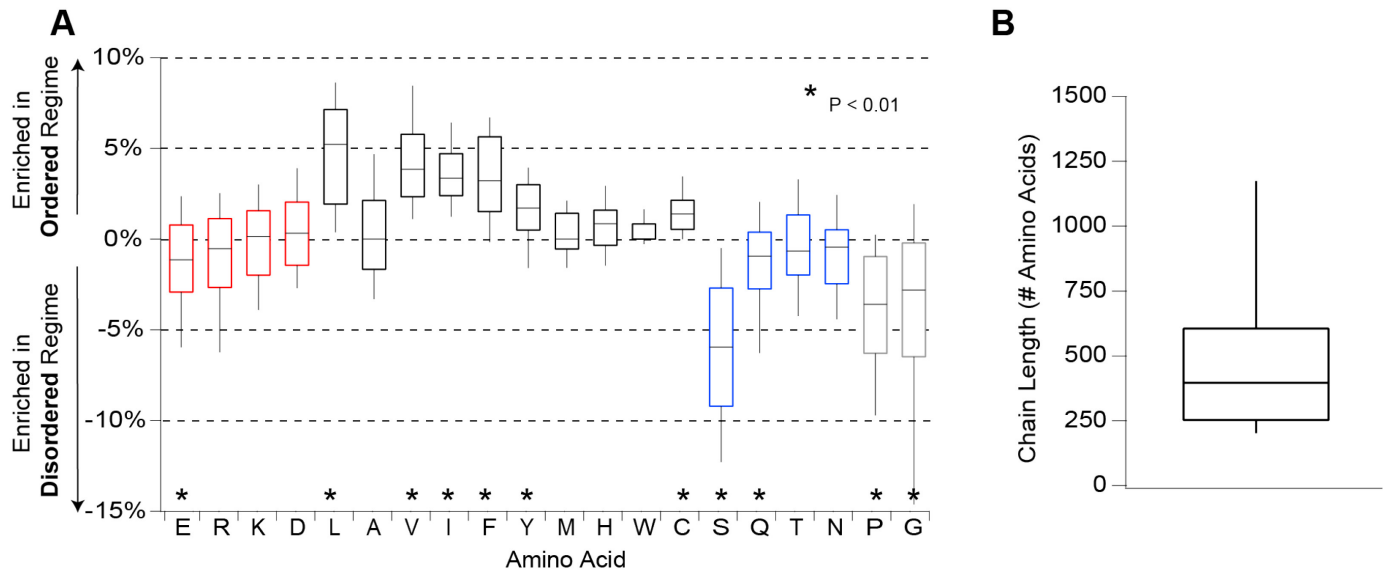


Figure S1 – Additional Proteomic Analysis Related to Figure 1. A. Graph of the difference in amino acid composition between ordered and disordered regions within the same protein. A disordered region was defined as being scored with a >0.5 value using PONDR VSL2, ordered regions with a score of <0.5 . Values were calculated by subtracting the percentage of chain composition in disordered regime from ordered regime. Bars indicate 25th-75th percentiles and whiskers indicate 10th-90th percentiles. Middle line indicates the median of the data set. $N = 63$, * $P < 0.01$ in students t-test between ordered and disordered regions of all proteins sampled. B. Histogram plotting the length of the disordered regions analyzed in this study. Bars indicate 25th-75th percentiles and whiskers indicate 10th-90th percentiles.

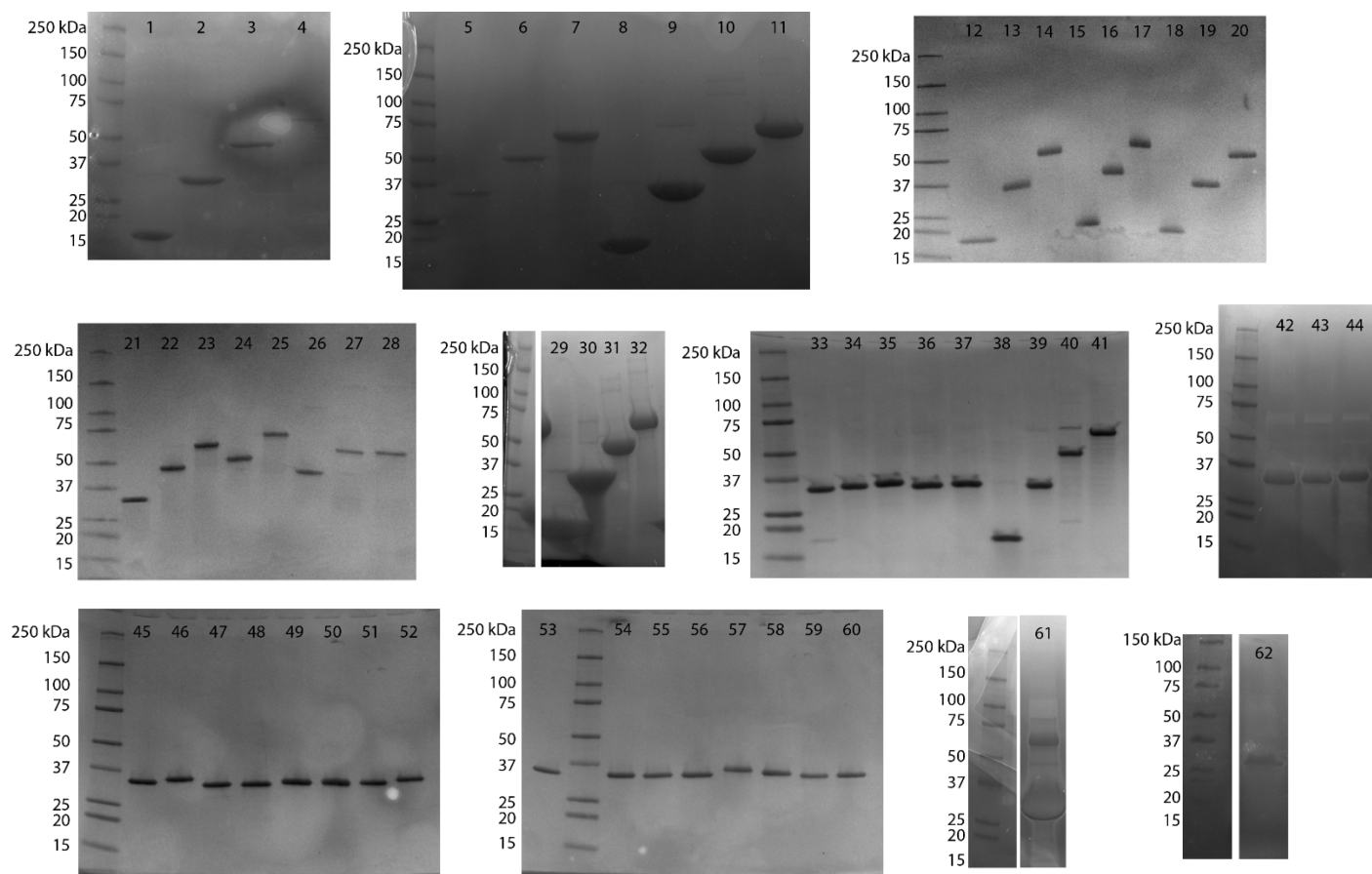
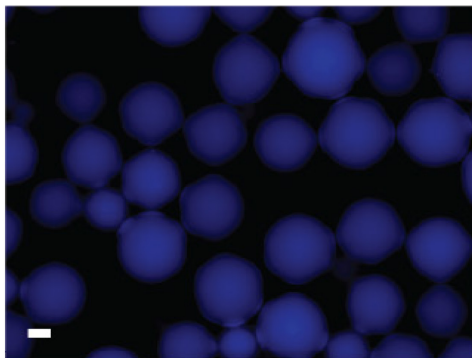
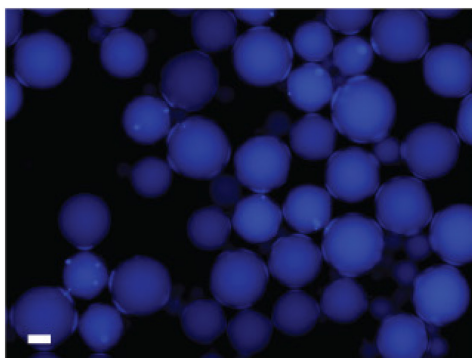


Figure S2 – SDS-PAGE Gels of Purified Proteins Used in this Study Relevant to Figure 1-5 – Lane labels for each protein purified in this study are listed in Table S1 and Table S2.

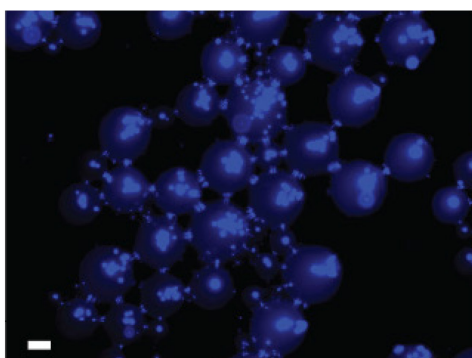
50°C



37°C



24°C



11°C

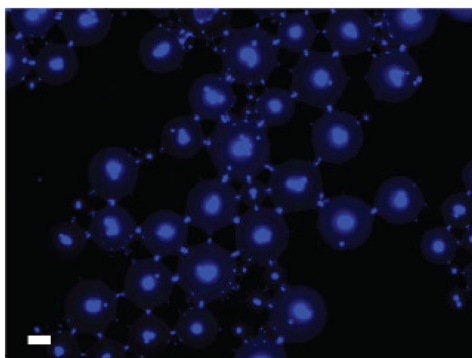


Figure S3 – Wide-field Fluorescent Microscope Images of Fluorescently Labeled [Q_{5,8}]-20 Inside Water-in-Oil Compartments Related to Figure 1. [Q_{5,8}]-20 was labeled with AlexaFluor 350 via NHS chemistry and resuspended in 140 mM PBS at pH 7.4 to a final $\phi = 0.003$. Water-in-oil mixture was transferred to glass slide and cooled from 50°C to 10°C. Scale bar = 20 μ m.

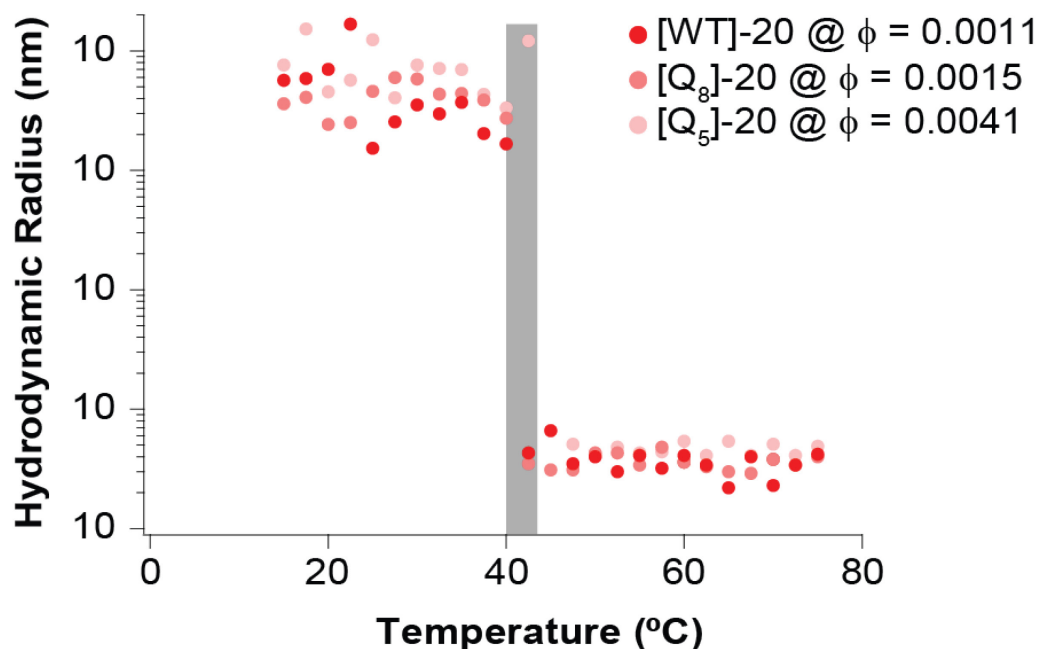


Figure S4 – Additional Dynamic Light Scattering Data Related to Figure 1. Data collected on 20 nm filtered samples at volume fractions that were predicted to exhibit liquid-liquid phase separation at 40°C. Data collected in 140 mM PBS, pH 7.4.

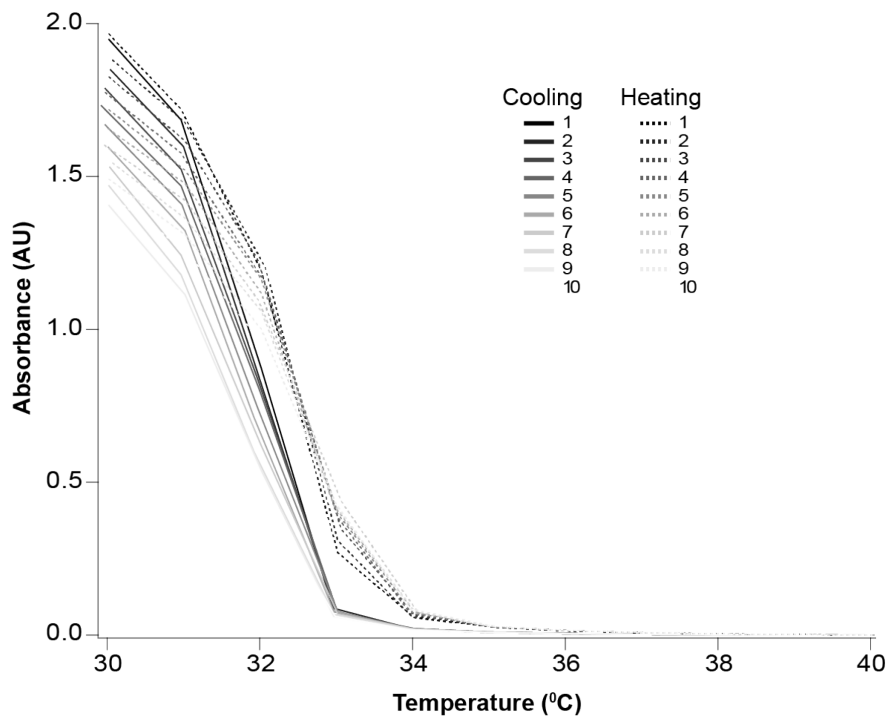


Figure S5 – Cyclic Cooling and Heating Cycles Exhibit Minimal Hysteric Behavior Related to Figure 1. Optical turbidity measured at 350 nm of repeated cooling and heating curves of [Q_{5,8}]-20 @ $\phi = 0.0025$ between 40°C and 30°C.

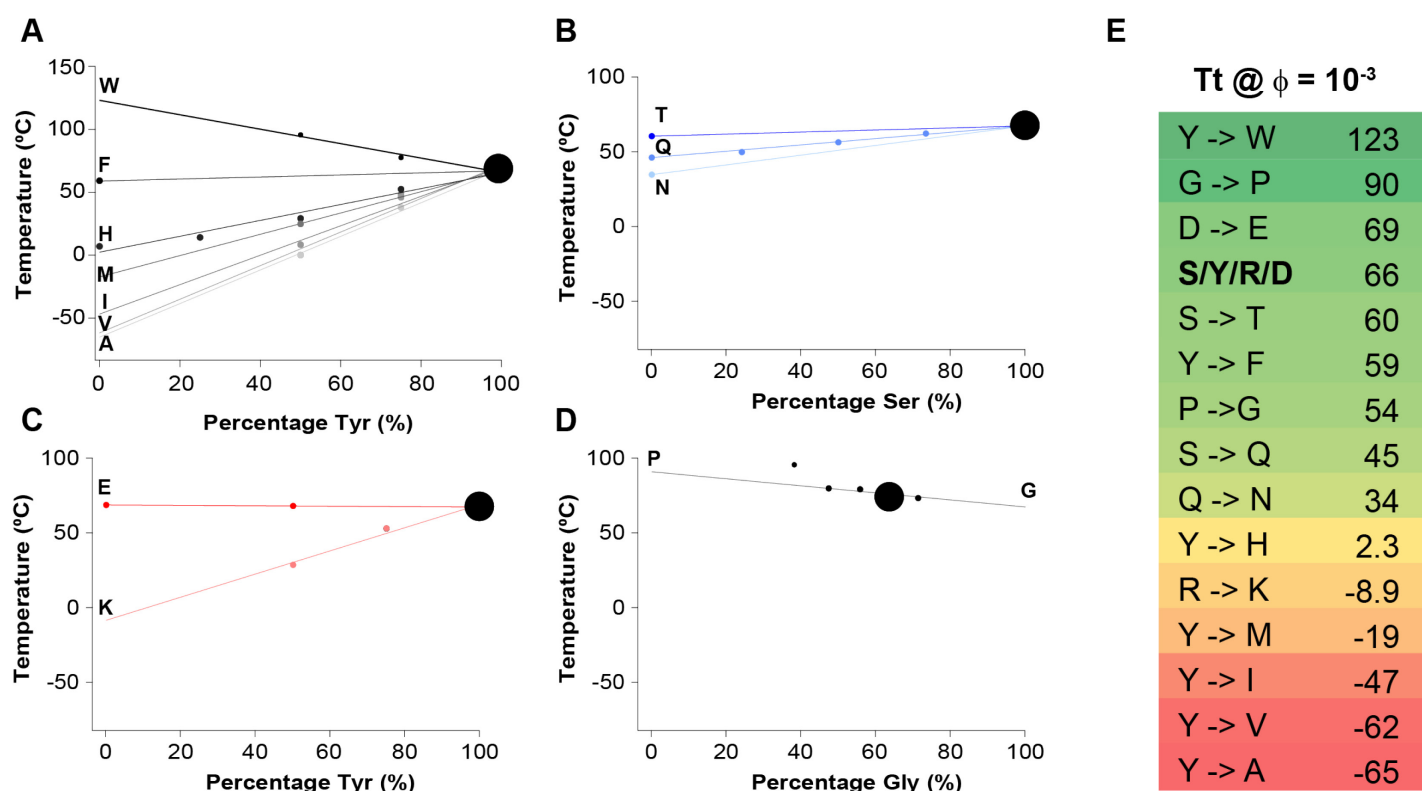


Figure S6 – Effect of Single Amino Acid Substitutions on UCST Cloud Point and New Relative UCST Propensity Scale Related to Figure 2. **A.** Partial binodal phase boundary of well-mixed, di-block polypeptides with varied ratio of aromatic:aliphatic residues. **B.** Partial binodal phase boundary of well-mixed, di-block polypeptides with varied ratio of polar non-charged residues. **C.** Partial binodal phase boundary of well-mixed, di-block polypeptides with varied identity of positively charged and negatively residues. **D.** Partial binodal phase boundary of well-mixed, di-block polypeptides with varied amount of. Data collected under physiologic solution conditions (140 mM PBS, pH 7.4) at $\phi = 10^{-3}$. All polypeptides are 326 amino acids in length. **E.** A relatively scale for UCST propensity based on substitutions made to the [WT] repeating motif. The transition temperatures listed are the UCST cloud point at $\phi = 10^{-3}$ if the left amino acid was replaced with the amino acid to the right of the arrow and the total number of amino acids was 326.

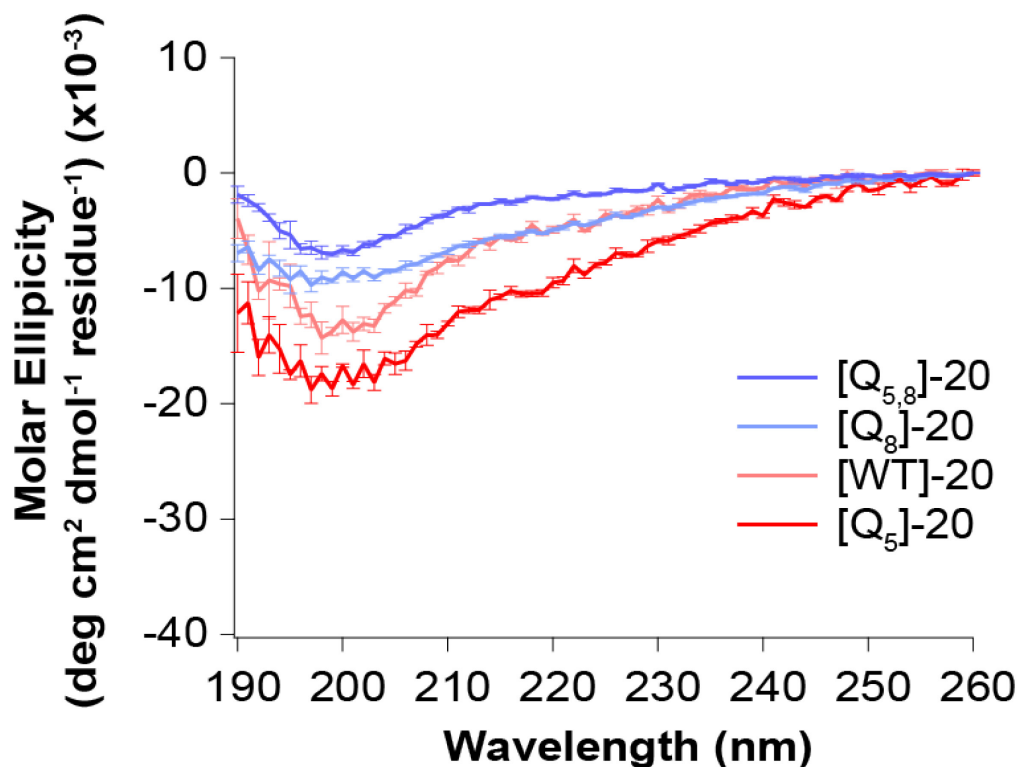


Figure S7 – Analysis of Secondary Structure with Circular Dichroism (CD) Spectroscopy Related to Figure 2. CD spectra of various A-IDPs lack a defined secondary structure curve shape, characteristic of other IDP and other repetitive protein polymers. Data collected at 50°C (soluble chains) at 5 μ M in 5 mM PBS, pH 7.4. Error bars indicate standard deviation of three sequential runs.

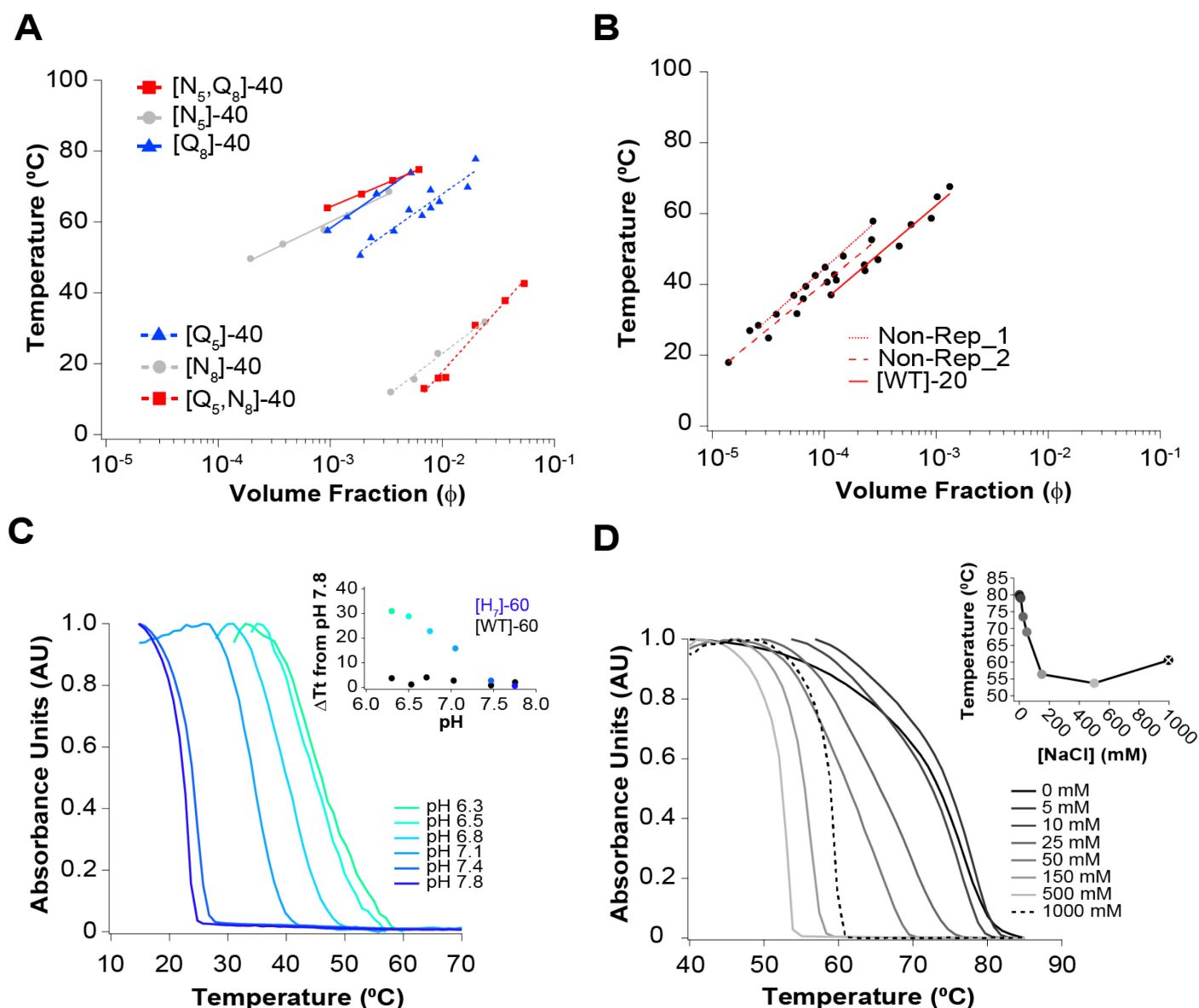


Figure S8 – Minor Effects on UCST Cloud Point in Protein Polypeptides Related to Figure 2. **A.** Partial binodal phase diagram of sequence syntax permutations focused around the Pro residue. Mutations reveal that amino acid mutation site affects the UCST binodal, particularly at the fifth position, but do not eliminate phase behavior. Data collected under physiologic conditions (140 mM PBS, pH 7.4). **B.** Partial binodal phase boundaries of agnostically non-repetitive but compositionally identical versions of [WT]-20. **C.** Turbidity curves of [H₇]-60 in different pH solutions. Decreasing the pH and protonation of the His residues increases and broadens observed UCST phase behavior. This effect centers at ~ pH 7, very close to the predicted pK_a of the imidazole group in H. In contrast to, [WT]-60's UCST cloud point does not change as a function of pH (black dots, graph insert). **D.** Turbidity curves of [Q_{5,8}]-40 in solutions with different concentrations of NaCl. In pure water, [Q_{5,8}]-40 exhibits a broad transition at higher temperatures. Increasing the concentration of NaCl between 0-140 mM reduces and sharpens the UCST cloud point, finally reaching a minimum at ~500 mM. From this point, the protein exhibits a salting-out effect and the transition temperature begins to rise again.

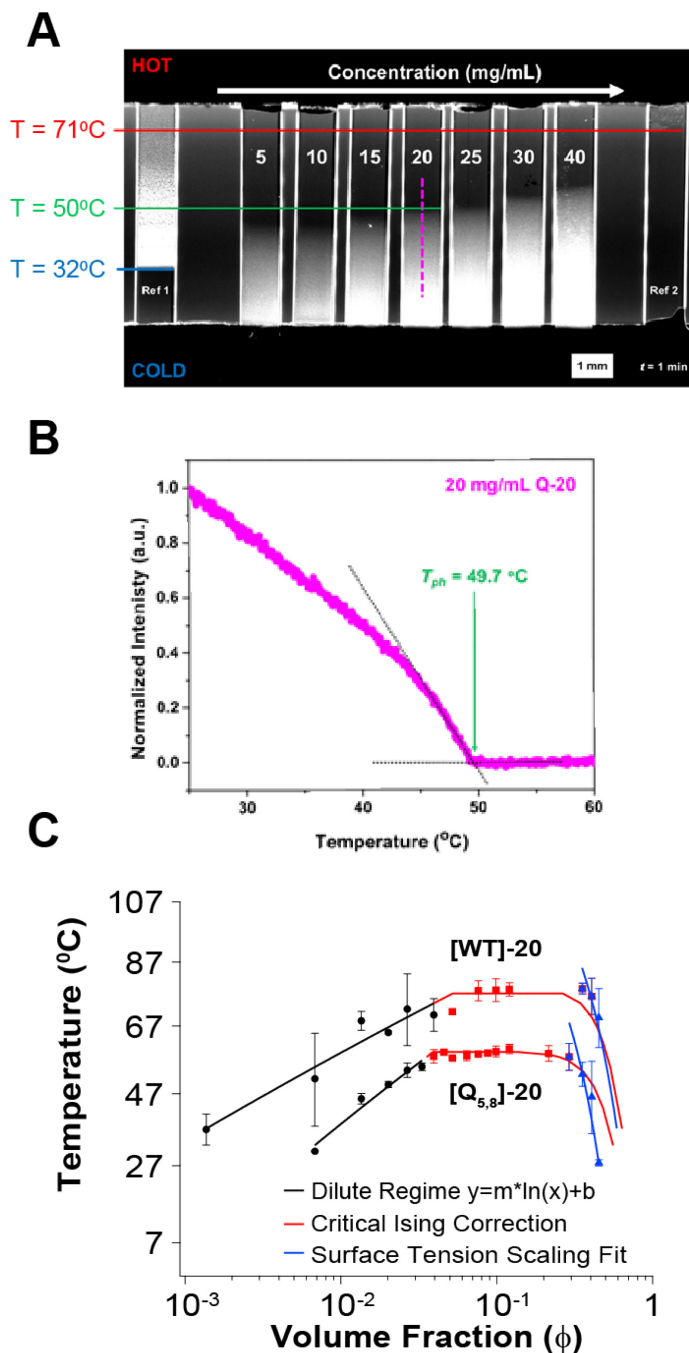


Figure S9 – Mapping Phase Diagrams Using A Temperature Gradient Device Related to Figure 2. A. Representative dark-field image of [Q_{5,8}]-20 solutions on a temperature gradient device. The transition temperatures of the reference solutions (red and blue lines) and the 20 mg ml⁻¹ [Q_{5,8}]-20 solution (green line) are indicated by the horizontal colored lines. The dashed vertical magenta line along the 20 mg ml⁻¹ capillary tube illustrated the region of the image used to measure the line scan. **B.** Line scan of normalized light scattering intensity versus temperature for the 20 mg ml⁻¹ [Q_{5,8}]-20 capillary shown in A. The dashed black lines represent tangent lines for the high temperature baseline and increase in light scattering at lower temperatures. These two lines intersect at T_{ph} , as indicated by the vertical green line. **C.** Final binodal phase lines of [WT]-20 and [Q_{5,8}]-20 using multiple data points from temperature gradient device. A three-piece fit was utilized to fit three regimes that roughly correspond to the dilute, overlap, and semi-dilute regimes of the polypeptide phase diagram. The observed data and subsequent fits demonstrate that polypeptide sequence not only affects UCST cloud point in the dilute regime but over the entire concentration range measured ($\phi \leq 0.5$)

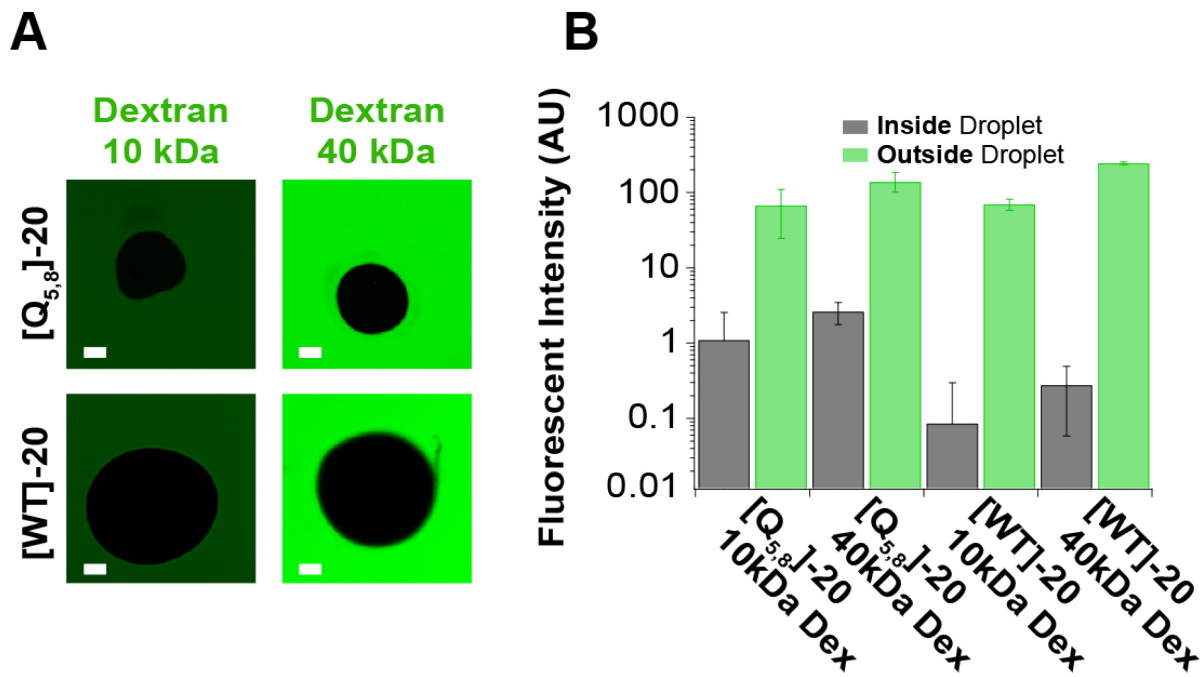


Figure S10 – Quantification of Dextran Uptake During Phase Separation of A-IDPs Related to Figure 1.
A. Fluorescent microscopy images of phase separated droplets in the presence of dextran molecules of different molecular weight (10/40kDa) labeled with Alexa488 (green) fluorophore. Inside the phase separated space (dark circles), there is very little sequestration of the dextran molecules as a function of dextran molecular weight or A-IDP sequence. Scale bar is 20 μ m. B. Quantification of fluorescent signal between the area inside of phase separated droplets and outside.

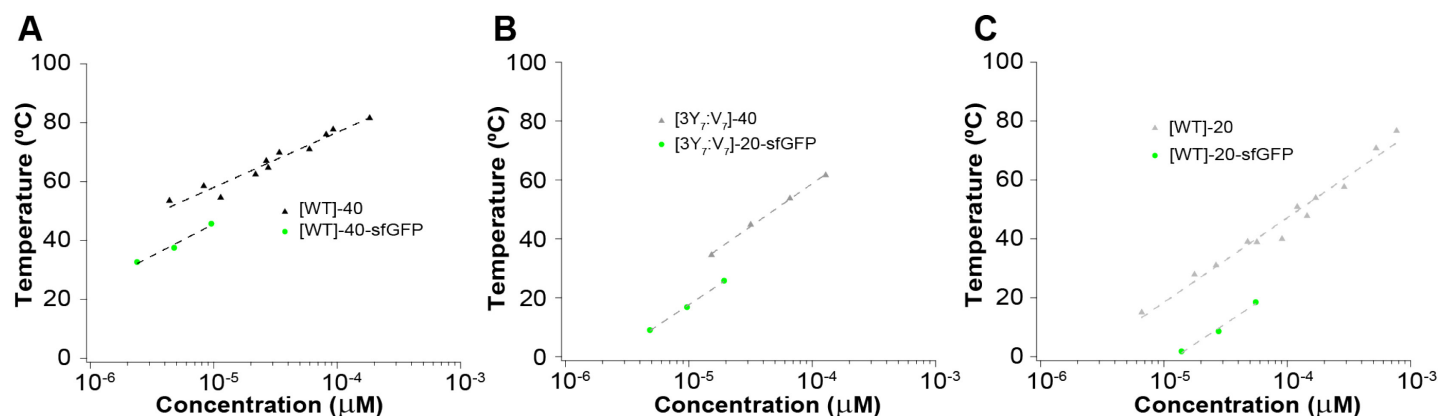


Figure S11 – Comparison of Partial Binodal Phase Diagrams of A-IDP and A-IDP-sfGFP Fusions Related to Figure 4. **A.** Partial binodal phase boundaries of [WT]-40 and [WT]-40-sfGFP. **B.** Partial binodal phase boundaries of [3Y7:V7]-40 and [3Y7:V7]-40-sfGFP. **C.** Partial binodal phase boundaries of [WT]-20 and [WT]-20-sfGFP. The sfGFP fusion lowers the UCST binodal line for all A-IDPs. These data suggest that the larger molecular weight polypeptides are less affected by sfGFP fusion as the observed difference in [WT]-40 and [3Y7:V7]-40 is a only $\sim 10^{\circ}\text{C}$ instead of $\sim 20^{\circ}\text{C}$ for [WT]-20.

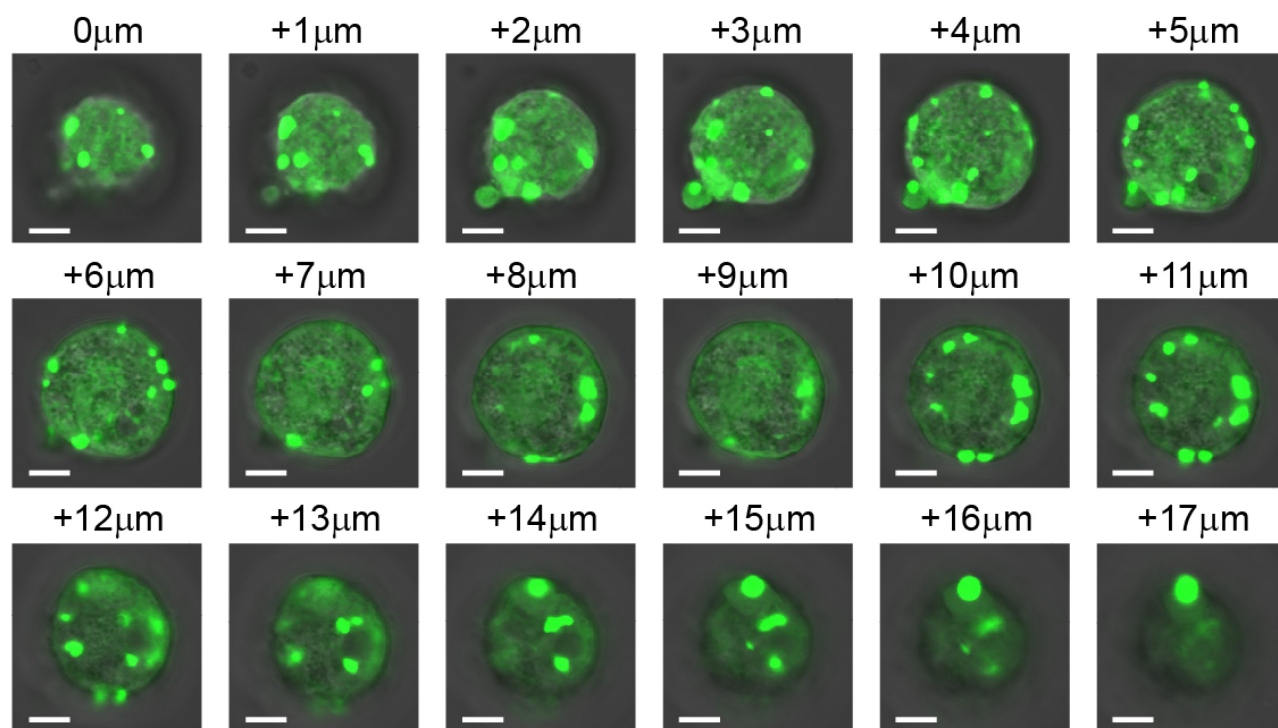


Figure S12 – Additional Confocal Microscopy Images of HEK293 Cells with Transiently Transfected [WT]-20-sfGFP Related to Figure 4. Confocal fluorescence image slices throughout the cell demonstrate that phase separated droplets are formed throughout the cytoplasm without obvious colocalization with other cellular structures. Images taken 24 hours after transfection with 3 μg of pCDNA plasmid that encode [WT]-20-sfGFP. Scale bar = 5 μm .

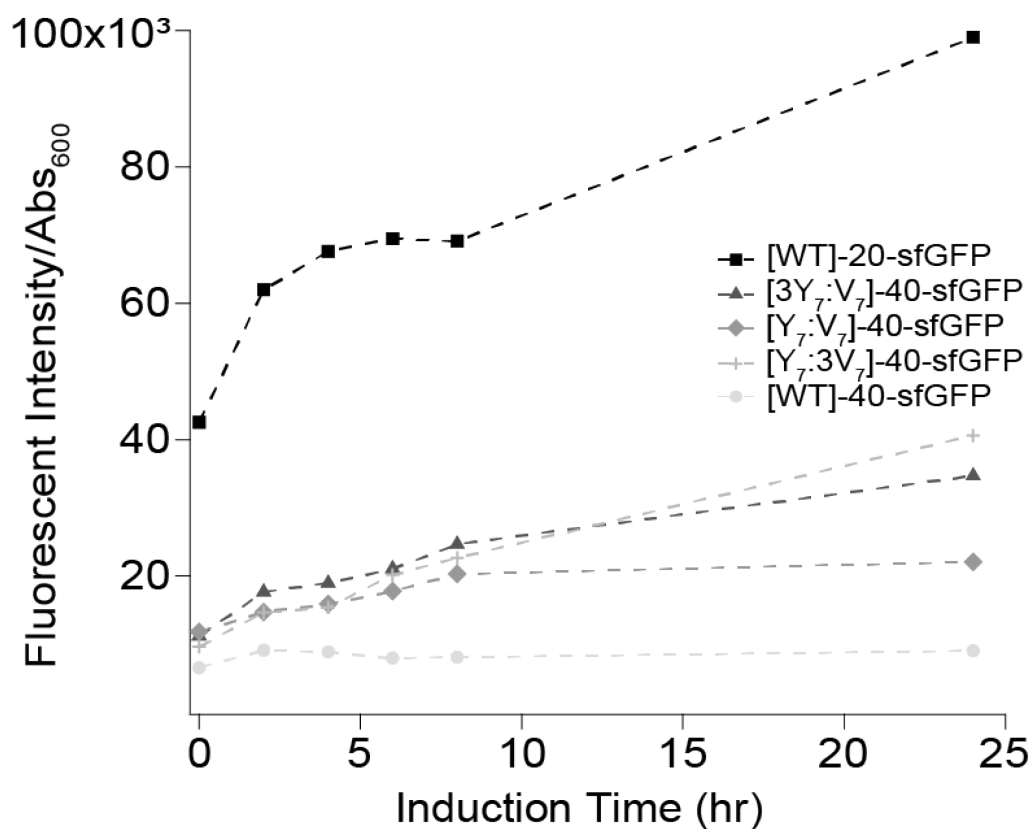


Figure S13 – Measurement of Total Cellular Fluorescence as a Function of Time Post Induction Related to Figure 4. *E. coli* cultures were spun down and resuspended in 140 mM PBS, pH 7.4. The optical turbidity and fluorescence intensity of sfGFP were measured and plotted as a function of time. Data collected at 22 °C.

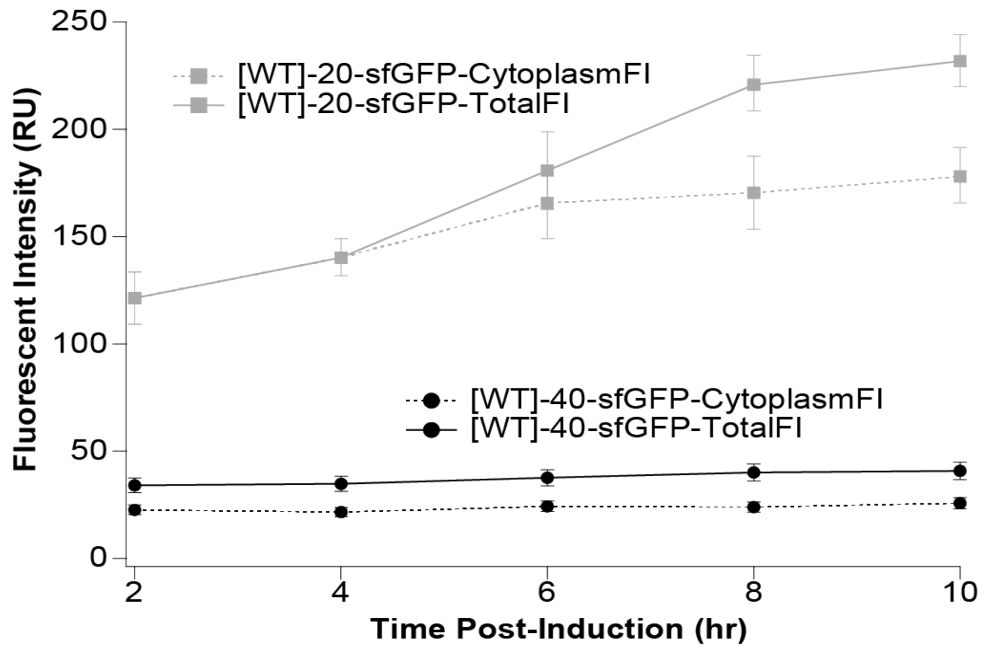


Figure S14 – Measurement of the Cellular Fluorescence at Different Locations Within the Cell Related to Figure 4. Digital partitions were made between the dense phase separated area of the cell and soluble cytoplasmic space using ImageJ. The mean of the total cell fluorescence intensity (solid line) and cytoplasmic fluorescence intensity (dotted line) are plotted as a function of time post-IPTG induction. [WT]-20-sfGFP does not exhibit intracellular droplets until the 6 hr mark. At this point the cytoplasmic fluorescence intensity remains constant but the total fluorescent increases from 6 hr onward. [WT]-40-sfGFP phase transitions prior to the 2-hr timepoint.

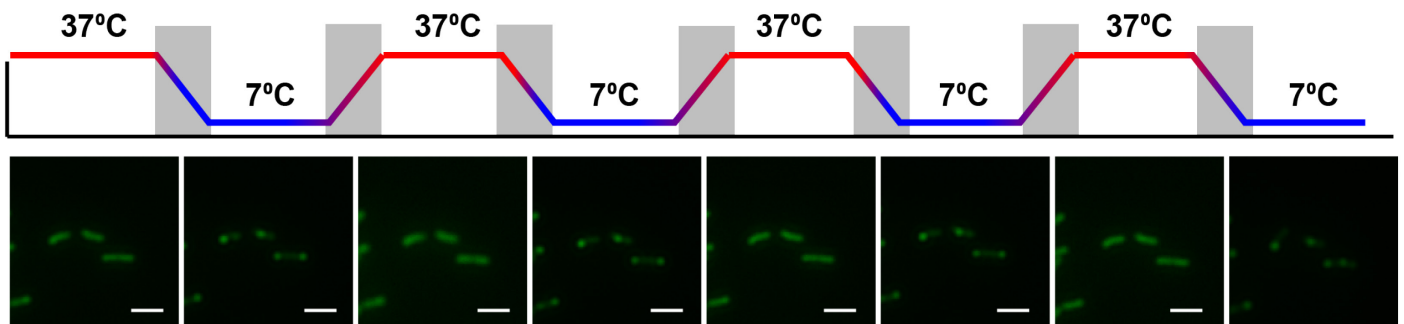


Figure S15 – [WT]-20-sfGFP Exhibits Phase Separation Memory Upon Multiple Cycles of Heating and Cooling Related to Figure 5. Upon multiple heating and cooling cycles, [WT]-20-sfGFP form puncta in the same location as the first cooling cycle. Given the importance of memory, it is critical to note that the observed transition temperature was below room temperature ($\sim 15^{\circ}\text{C}$), suggesting that these cells are naïve to phase separation as they were incubated at 37°C and processed at room temperature. Scale bar indicated $5\ \mu\text{m}$. Cooling and heating rate were set to a constant $5^{\circ}\text{C min}^{-1}$.

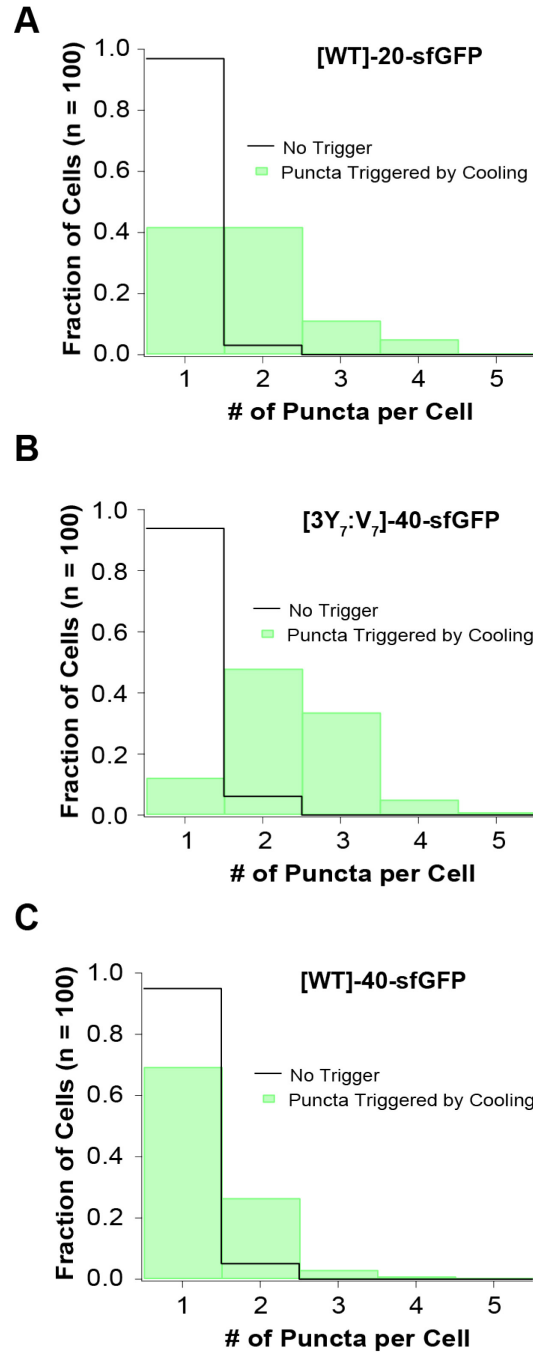


Figure S16 – Image Analysis of the Number of Puncta Formed in Each Cell Related to Figure 5. 100 cells at random were tabulated for each histogram. **A.** Number of intracellular puncta formed in each cell containing [WT]-20-sfGFP during a cooling ramp from 60°C → 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 6 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, where transition temperature (T_i) was between 22°C and 37°C. **B.** Number of intracellular puncta formed in each cell containing [3Y₇:V₇]-40-sfGFP during a cooling ramp from 60°C → 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 4 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, where T_i was between 22°C and 37°C. **C.** Number of intracellular puncta formed in each cell containing [WT]-40-sfGFP during a cooling ramp from 60°C → 10°C (green) and imaged isothermally at 22°C. Isothermal analysis performed at 4 hours post induction, the first timepoint where intracellular puncta were observed. Cooling ramp performed at 4 hours post induction, but the transition observed was >37°C indicating the possibility of memory.

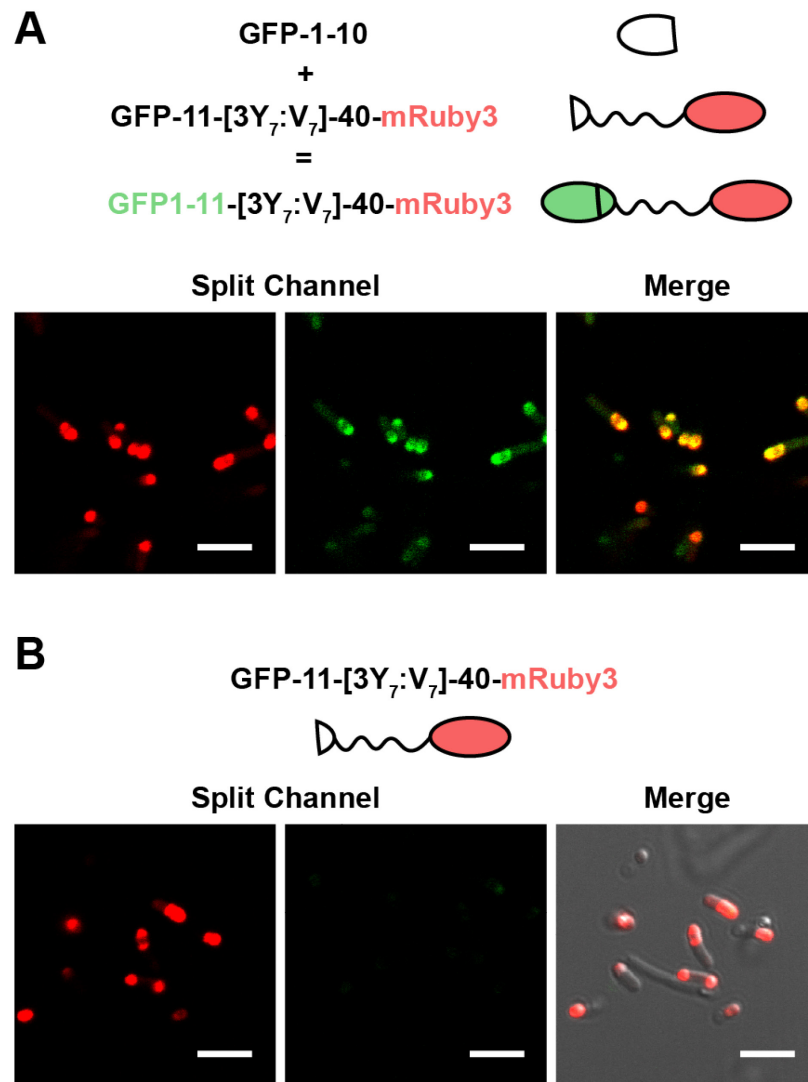


Figure S17 – Confocal Microscope Images of Split GFP Recruitment into Intracellular Droplets Related to Figure 5. A. GFP-11-[3Y₇:V₇]-40-mRuby3 co-expressed in the presence of GFP-1-10 creates fluorescently active GFP only in the interior of the droplet. **B.** In the absence of GFP-1-10 induction, there is little green fluorescent inside the intracellular droplets. Data taken at 22°C. Scale bar = 5 μm.

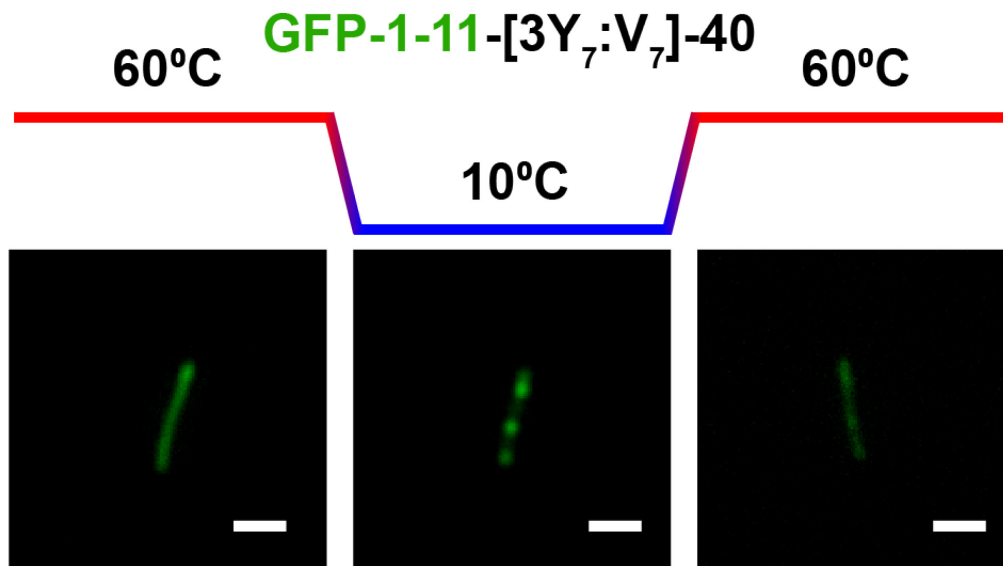


Figure S18 - A-IDPs can Modulate the Solubility of an Endogenously Bound Molecule Related to Figure 6. Upon recruitment of sfGFP into the dense phase, the solubility of the entire complex can be modulated with temperature. Data was collected for 36 hours post-IPTG induction and 12 hours post-arabinose induction. Scale bar = 5 μ m.

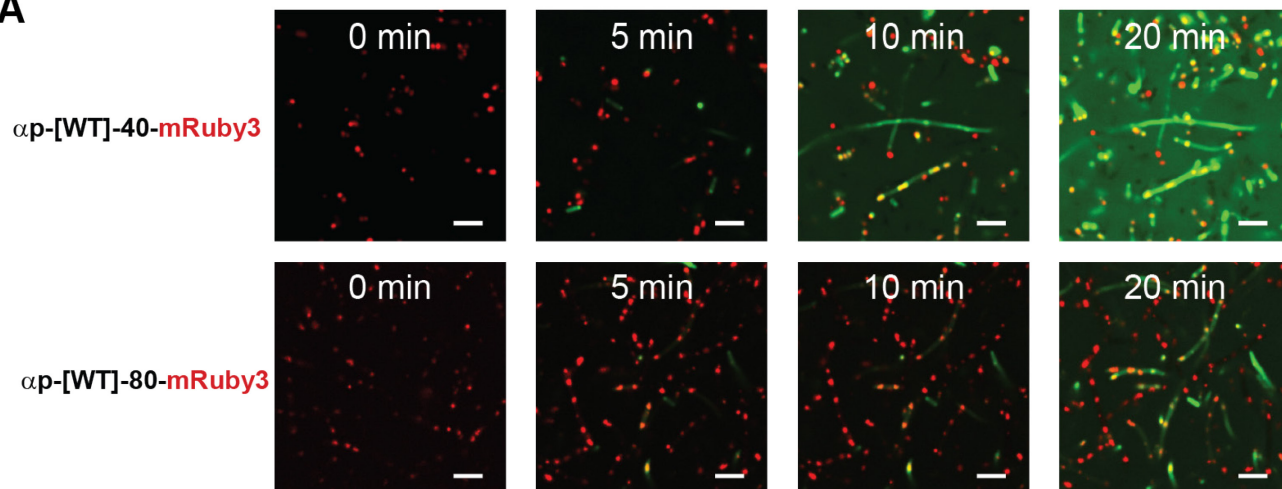
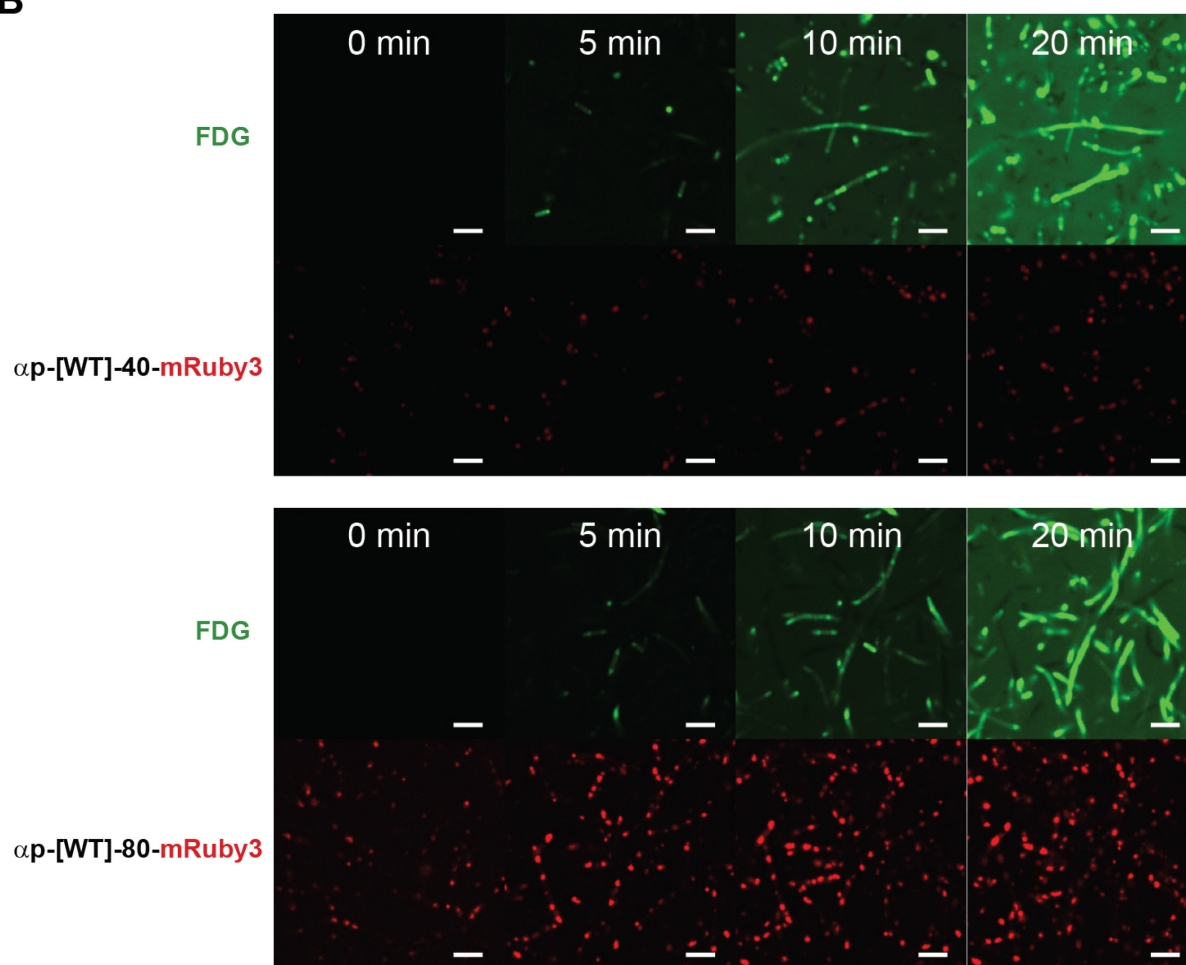
A**B**

Figure S19 – Color Balanced Confocal Microscopy Images of αp -[WT]-40-mRuby3 and αp -[WT]-80-mRuby3 Related to Figure 6. All scale bars are 5 μm . **A.** Color re-balanced images from Figure 6B for improved visualization of the intracellular droplets formed by αp -A-IDP-mRuby3 fusions. **B.** Split channel images of αp -[WT]-40-mRuby3 and αp -[WT]-80-mRuby3.

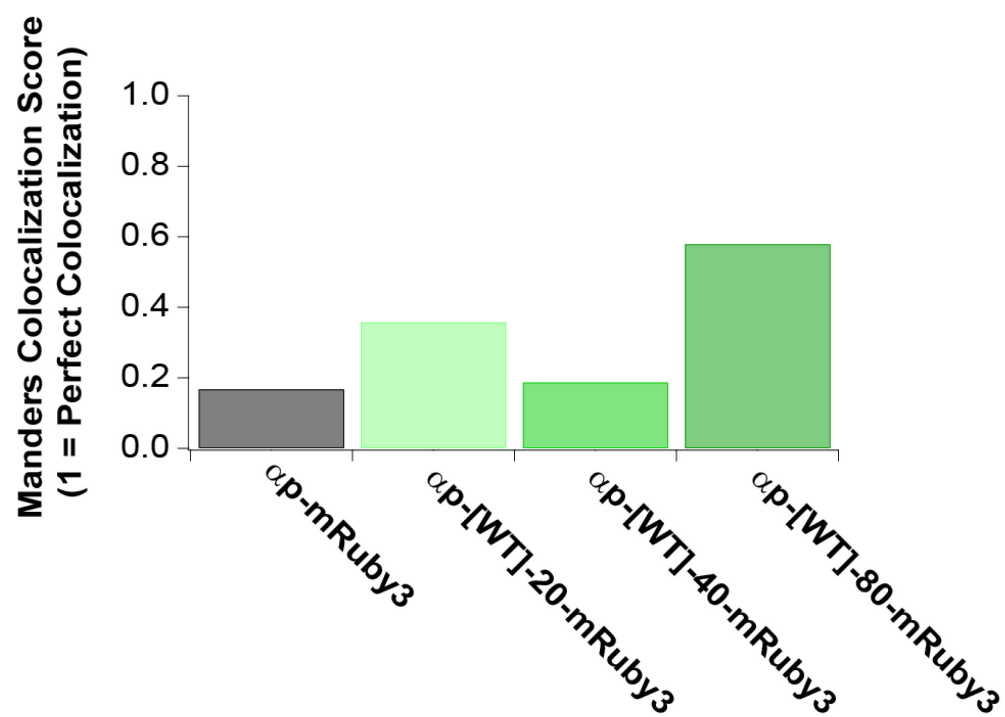


Figure S20 – Mander's Colocalization Score Between Converted FDG and Fluorescent Reporter Related to Figure 6. Data analyzed 30 min after FDG addition. Background threshold was set automatically.

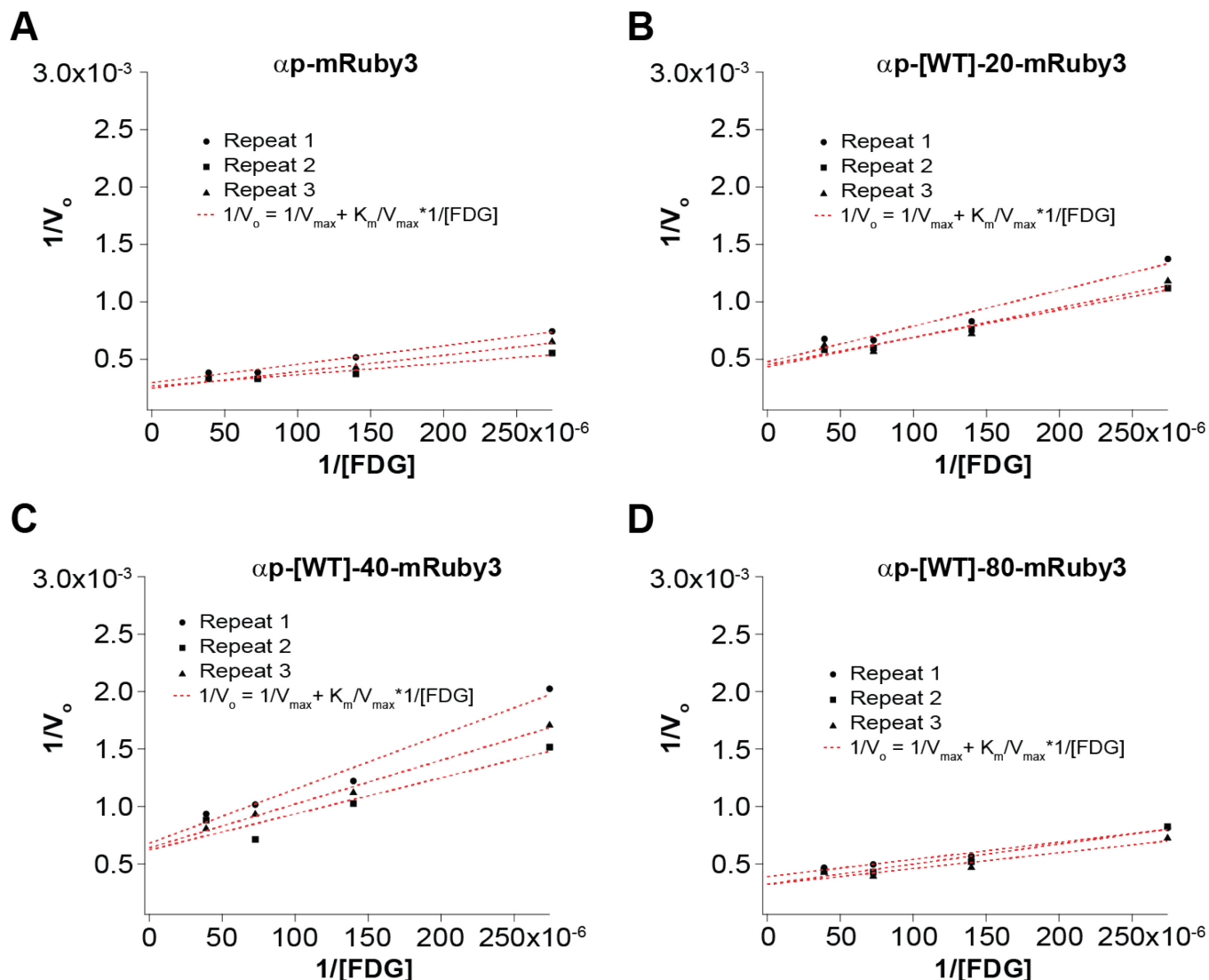


Figure S21 – Lineweaver-Burk Plots for Determining K_m and V_{max} Related to Figure 6. Lineweaver-Burk plots created with variable starting concentrations of FDG for **A.** αp -mRuby3, **B.** αp -[WT]-20-mRuby3, **C.** αp -[WT]-40-mRuby3 and **D.** αp -[WT]-80-mRuby3. Slopes (V_o) were determined from fluorescent generation over the course of 20 minutes. Intercepts and slope were used in the calculation of K_m and V_{max} .

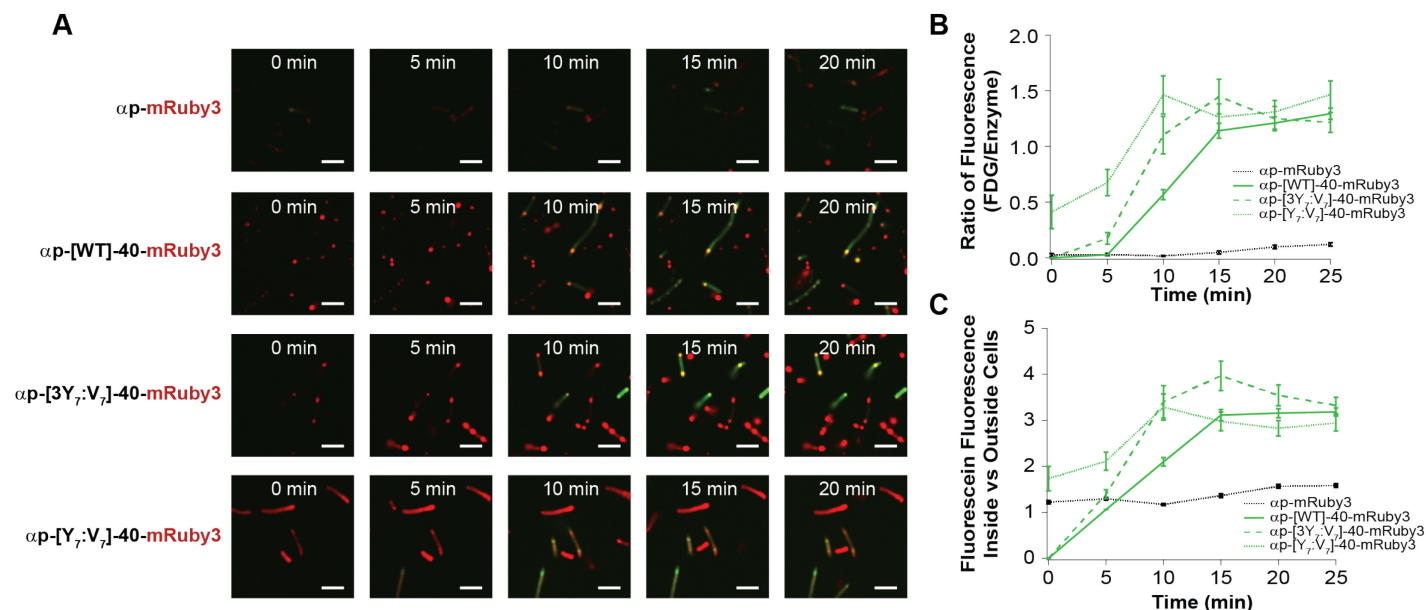


Figure S22 – Enzymatic Droplets Formed with Variable Ratio of Aromatic to Aliphatic Ratio Related to Figure 6. All scale bars are 5 μ m. **A.** Confocal microscopy images observing the fluorescent conversion of Fluorescein Di- β -D-Galactopyranoside (FDG) of αp -mRuby3, αp -[WT]-40-mRuby3, αp -[3Y₇:V₇]-40-mRuby3 and αp -[Y₇:V₇]-40-mRuby3. Decreasing the aromatic:aliphatic ratio does not increase FDG conversion over time but does change the dynamics of uptake with lower aromatic:aliphatic ratio polypeptides observing higher uptake at earlier timepoints after FDG addition. **B.** Quantified amount of converted FDG intracellularly, normalized to the amount of mRuby3 fluorescence. There is little difference between A-IDPs with different ratios of aromatic:aliphatic content. Error bars indicate standard error of the mean. **C.** All αp -A-IDP-mRuby3 fusions exhibit a higher ratio of FDG fluorescence inside the cell, indicating a greater persistence of fluorescent FDG inside the intracellular space compared to the αp -mRuby3 control. There is little difference between A-IDPs with different levels of aromatic content. Error bars indicate standard error of the mean.

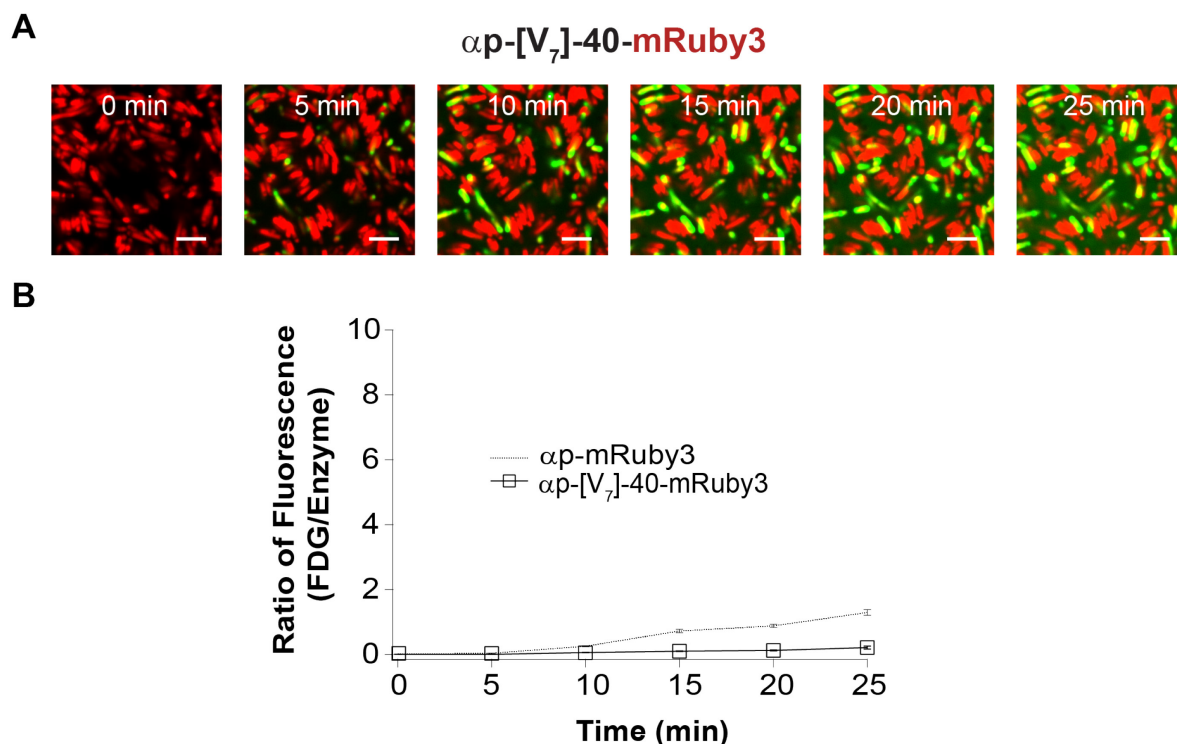


Figure S23 - Enzymatic Activity of αp -[V₇]-40-mRuby3 Related to Figure 6. All scale bars are 5 μ m. **A.** Confocal microscopy images showing the fluorescent conversion of fluorescein Di- β -D-galactopyranoside (FDG) attached to soluble αp -[V₇]-40-mRuby3. **B.** Intracellular concentration of fluorescein produced by catalytic conversion of FDG by αp -[V₇]-40-mRuby3, normalized to the mRuby3 fluorescence of each individual cell ($n \cong 300$). The soluble fusion exhibits a lower level of enzymatic activity than all puncta-forming αp -A-IDP fusions. αp -mRuby3 data from Figure 6 redrawn to show scale. Error bars indicate standard error of the mean.

Table S1 – Amino Acid Sequence of A-IDPs with a Single Repeat Motif.

Protein Name	Full Amino Acid Sequence	Amino Acid Number	Molecular Weight (Da)	μM @ 37C	Image Index
[WT]-20	SKGP-[GRGDSPYS] ₂₀ -GY	166	17004	44.18	1
[WT]-40	SKGP-[GRGDSPYS] ₄₀ -GY	326	33400	0.755	2
[WT]-60	SKGP-[GRGDSPYS] ₆₀ -GY	486	49797	0.044	3
[WT]-80	SKGP-[GRGDSPYS] ₈₀ -GY	646	66193	0.0002	4
[Q _{5,8}]-20	SKGP-[GRGDQPYQ] ₂₀ -GY	166	18646	247.9	8
[Q _{5,8}]-40	SKGP-[GRGDQPYQ] ₄₀ -GY	326	36685	6.593	9
[Q _{5,8}]-60	SKGP-[GRGDQPYQ] ₆₀ -GY	486	54723	1.034	10
[Q _{5,8}]-80	SKGP-[GRGDQPYQ] ₈₀ -GY	646	72762	0.241	11
[T _{5,8}]-40	SKGP-[GRGDTPTY] ₄₀ -GY	326	34522	0.856	60
[N _{5,8}]-20	SKGP-[GRGDNPNY] ₂₀ -GY	166	18085	337.2	12
[N _{5,8}]-40	SKGP-[GRGDNPNY] ₄₀ -GY	326	35562	39.18	13
[N _{5,8}]-60	SKGP-[GRGDNPNY] ₆₀ -GY	486	53040	4.827	14
[H ₇]-40	SKGP-[GRGDSPHS] ₄₀ -GY	326	32359	8376.0	61
[H ₇]-60	SKGP-[GRGDSPHS] ₆₀ -GY	486	48235	286.8	24
[H ₇]-80	SKGP-[GRGDSPHS] ₈₀ -GY	646	64111	327.0	25
[F ₇]-40	SKGP-[GRGDSPFS] ₄₀ -GY	326	32760	1.923	62
[Q ₈]-20	SKGP-[GRGDSPYQ] ₂₀ -GY	166	17825	0.029	38
[Q ₈]-40	SKGP-[GRGDSPYQ] ₄₀ -GY	326	35042	0.135	39
[Q ₈]-60	SKGP-[GRGDSPYQ] ₆₀ -GY	486	52260	3.179	40
[Q ₈]-80	SKGP-[GRGDSPYQ] ₈₀ -GY	646	69478	15.72	41
[Q ₅]-20	SKGP-[GRGDQPYS] ₂₀ -GY	166	17825	149.0	29
[Q ₅]-40	SKGP-[GRGDQPYS] ₄₀ -GY	326	35042	7.941	30
[Q ₅]-60	SKGP-[GRGDQPYS] ₆₀ -GY	486	52260	0.723	31
[Q ₅]-80	SKGP-[GRGDQPYS] ₈₀ -GY	646	69478	0.086	32
[N ₅]-40	SKGP-[GRGDNPYS] ₄₀ -GY	326	34481	0.670	21
[N ₅]-60	SKGP-[GRGDNPYS] ₆₀ -GY	486	51418	0.028	22
[N ₅]-80	SKGP-[GRGDNPYS] ₈₀ -GY	646	68356	0.004	23
[N ₈]-20	SKGP-[GRGDSPYN] ₂₀ -GY	166	17544	484.1	18
[N ₈]-40	SKGP-[GRGDSPYN] ₄₀ -GY	326	34481	16.17	19
[N ₈]-60	SKGP-[GRGDSPYN] ₆₀ -GY	486	51418	3.748	20
[N ₅ ,Q ₈]-40	SKGP-[GRGDNPYQ] ₄₀ -GY	326	36123	2.449	5
[N ₅ ,Q ₈]-60	SKGP-[GRGDNPYQ] ₆₀ -GY	486	53882	0.176	6
[N ₅ ,Q ₈]-80	SKGP-[GRGDNPYQ] ₈₀ -GY	646	71640	0.108	7
[Q ₅ ,N ₈]-20	SKGP-[GRGDQPYN] ₂₀ -GY	166	18365	637.5	15

[Q ₅ ,N ₈]-40	SKGP-[GRGDQPYN] ₄₀ -GY	326	36123	52.62	16
[Q ₅ ,N ₈]-60	SKGP-[GRGDQPYN] ₆₀ -GY	486	53882	13.47	17
[WT]-20-sfGFP	SKGP-[GRGDSPYS]-20-sfGFP	410	44551	216.1	26
[WT]-40-sfGFP	SKGP-[GRGDSPYS]-40-sfGFP	570	60947	3.448	27

Table S2 – Amino Acid Sequence of A-IDPs with Multiple Repeat Motifs.

Protein Name	Full Amino Acid Sequence	Amino Acid Number	Molecular Weight (Da)	μM @ 37C	Image Index
[3S _{5,8} :Q _{5,8}]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD QPYQ] ₁₀ -GY	326	34221	1.600	33
[S _{5,8} :Q _{5,8}]-40	SKGP-[GRGDSFYSGRGDQPYQ] ₂₀ -GY	326	35042	1.907	34
[S _{5,8} :3Q _{5,8}]-40	SKGP-[GRGDQPYQGRGDQPYQGRGDQPYQGRGDSFY] ₁₀ -GY	326	35863	5.062	35
[3Y ₇ :V ₇]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD SPVS] ₁₀ -GY	326	32760	24.54	36
[Y ₇ :V ₇]-40	SKGP-[GRGDSFYSGRGDSPVS] ₂₀ -GY	326	32119	815.5	37
[Y ₇ :3V ₇]-40	SKGP-[GRGDSFVSGRGDSFVSGRGDSFVSGRGD SPYS] ₁₀ -GY	326	31479	Unk	N/A
[V ₇]-40	SKGP-[GRGDSFVS] ₄₀ -GY	326	30839	Unk	N/A
[3Y ₇ :A ₇]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD SPAS] ₁₀ -GY	326	32479	28.46	45
[Y ₇ :A ₇]-40	SKGP-[GRGDSFYSGRGDSPAS] ₂₀ -GY	326	31558	1816	46
[3Y ₇ :I ₇]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD SPIS] ₁₀ -GY	326	32900	12.88	47
[Y ₇ :I ₇]-40	SKGP-[GRGDSFYSGRGDSPIS] ₂₀ -GY	326	32400	214.85 06	48
[3Y ₇ :M ₇]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD SPMS] ₁₀ -GY	326	33080	7.914	49
[Y ₇ :M ₇]-40	SKGP-[GRGDSFYSGRGDSPMS] ₂₀ -GY	326	32761	110.5	50
[3Y ₇ :H ₇]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGRGD SPHS] ₁₀ -GY	326	33140	7.125	42
[Y ₇ :H ₇]-40	SKGP-[GRGDSFYSGRGDSPHS] ₂₀ -GY	326	32880	70.86	43
[Y ₇ :3H ₇]-40	SKGP-[GRGDSFHSRGDSFHSRGDSFHSRGD SPYS] ₁₀ -GY	326	32619	508.5	44
[3R ₂ :K ₂]-40	SKGP-[GRGDSFYSGRGDSFYSGRGDSFYSGKGD SPYS] ₁₀ -GY	326	33120	8.537	51
[R ₂ :K ₂]-40	SKGP-[GRGDSFYSGKGDSPYS] ₂₀ -GY	326	32840	79.88	52
[D ₄ :E ₄]-40	SKGP-[GRGDSFYSGRGESPYS] ₂₀ -GY	326	33681	1.779	53

Table S4 – Number of Phase Separated Domains per *E. coli* as a Function of Induction time (n = 3 images)

Time Post-Induction (hr)	[WT]-20-sfGFP	[WT]-40-sfGFP
2	N/A	1.01 ± 0.02
6	1.03 ± 0.01	1.03 ± 0.03
10	1.13 ± 0.01	1.10 ± 0.09

Table S5 – Michaelis-Menten Enzyme Kinetics Parameters (error is standard error of the mean, n = 3).

Protein	V _{max} (FI _{FDG} *min ⁻¹)	K _m (FI _{FDG})	k _{cat} (min ⁻¹)	k _{cat} /K _m (FI _{FDG} ⁻¹ *min ⁻¹)
αp-mRuby3	3708 ± 183.2	4972 ± 636.3	3.56 ± 0.10	7.43E-04 ± 1.02 E-04
αp-[WT]-20-mRuby3	2165 ± 35.02	5896 ± 380.5	5.10 ± 0.43	8.79E-04 ± 1.16 E-04
αp-[WT]-40-mRuby3	1549 ± 34.75	5988 ± 475.1	5.75 ± 0.13	9.81E-04 ± 9.96 E-05
αp-[WT]-80-mRuby3	2922 ± 174.9	4504 ± 439.7	15.2 ± 0.68	3.45E-03 ± 4.58 E-04

Table S6 – Primers Used for pcDNA5 Cloning of [WT]-20-sfGFP.

pcDNA5-Fwd-[WT]-20-sfGFP	CTCACTATAGGGAGACCCAAGCTGGCTAGCATGAGCAAAGGGCCGGGACGC GGCGATAGT
pcDNA5-Rev-[WT]-20-sfGFP	TTAGCCGTGATGGTGGTGGTGGTGGAGCTCGTTGATTGTCGAGGGCCCTCTA GACTCGAG

References

- Babu, M.M. (2016). The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem Soc Trans* 44, 1185-1200.
- Balu, R., Dutta, N.K., Choudhury, N.R., Elvin, C.M., Lyons, R.E., Knott, R., and Hill, A.J. (2014). An16-resilin: an advanced multi-stimuli-responsive resilin-mimetic protein polymer. *Acta Biomater* 10, 4768-4777.
- Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18, 285-298.
- Banjade, S., and Rosen, M.K. (2014). Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife* 3.
- Best, R.B. (2017). Computational and theoretical advances in studies of intrinsically disordered proteins. *Curr Opin Struct Biol* 42, 147-154.
- Bochicchio, B., Pepe, A., and Tamburro, A.M. (2008). Investigating by CD the molecular mechanism of elasticity of elastomeric proteins. *Chirality* 20, 985-994.
- Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., *et al.* (2018). Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* 28, 420-435.
- Brangwynne, Clifford P., Tompa, P., and Pappu, Rohit V. (2015). Polymer physics of intracellular phase transitions. *Nature Physics* 11, 899-904.
- Broome, A.M., Bhavsar, N., Ramamurthy, G., Newton, G., and Bacion, J.P. (2010). Expanding the utility of beta-galactosidase complementation: piece by piece. *Mol Pharm* 7, 60-74.
- Das, R.K., and Pappu, R.V. (2013). Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci U S A* 110, 13392-13397.
- Dignon, G.L., Zheng, W., Best, R.B., Kim, Y.C., and Mittal, J. (2018a). Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc Natl Acad Sci U S A* 115, 9929-9934.
- Dignon, G.L., Zheng, W., Kim, Y.C., Best, R.B., and Mittal, J. (2018b). Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput Biol* 14, e1005941.

Duan, X., Chen, J., and Wu, J. (2013). Improving the thermostability and catalytic efficiency of *Bacillus deramificans* pullulanase by site-directed mutagenesis. *Appl Environ Microbiol* 79, 4072-4077.

Dzuricky, M., Roberts, S., and Chilkoti, A. (2018). Convergence of Artificial Protein Polymers and Intrinsically Disordered Proteins. *Biochemistry* 57, 2405-2414.

Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C., Eckmann, C.R., Myong, S., and Brangwynne, C.P. (2015). The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* 112, 7189-7194.

Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nuske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R.V., *et al.* (2018). Phase separation of a yeast prion protein promotes cellular fitness. *Science* 359.

Goldsmith, M., and Tawfik, D.S. (2017). Enzyme engineering: reaching the maximal catalytic efficiency peak. *Curr Opin Struct Biol* 47, 140-150.

Hofmann, J., and Sernetz, M. (1983). A kinetic study on the enzymatic hydrolysis of fluoresceindiacetate and fluorescein-di- β -D-galactopyranoside. *Analytical biochemistry* 131, 180-186.

Holehouse, A.S., and Pappu, R.V. (2018). Functional Implications of Intracellular Phase Transitions. *Biochemistry* 57, 2415-2423.

Huber, M.C., Schreiber, A., von Olshausen, P., Varga, B.R., Kretz, O., Joch, B., Barnert, S., Schubert, R., Eimer, S., Kele, P., *et al.* (2015). Designer amphiphilic proteins as building blocks for the intracellular formation of organelle-like compartments. *Nat Mater* 14, 125-132.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157, 105-132.

Lewis, P., Nwoguh, C., Barer, M., Harwood, C., and Errington, J. (1994). Use of digitized video microscopy with a fluorogenic enzyme substrate to demonstrate cell-and compartment-specific gene expression in *Salmonella enteritidis* and *Bacillus subtilis*. *Molecular microbiology* 13, 655-662.

Li, L., Luo, T., and Kiick, K.L. (2015). Temperature-triggered phase separation of a hydrophilic resilin-like polypeptide. *Macromol Rapid Commun* 36, 90-95.

Li, P., Banjade, S., Cheng, H.C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., *et al.* (2012). Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483, 336-340.

Li, Z., Tyrpak, D.R., Lien, C.-L., and MacKay, J.A. (2018). Tunable Assembly of Protein-Microdomains in Living Vertebrate Embryos. *Advanced Biosystems* 2.

Lin, Y., Currie, S.L., and Rosen, M.K. (2017). Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* 292, 19110-19120.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60, 208-219.

Lin, Y.H., and Chan, H.S. (2017). Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. *Biophys J* 112, 2043-2046.

Lin, Y.H., Forman-Kay, J.D., and Chan, H.S. (2016). Sequence-Specific Polyampholyte Phase Separation in Membraneless Organelles. *Phys Rev Lett* 117, 178101.

MacEwan, S.R., Weitzhandler, I., Hoffmann, I., Genzer, J., Gradzielski, M., and Chilkoti, A. (2017). Phase Behavior and Self-Assembly of Perfectly Sequence-Defined and Monodisperse Multiblock Copolypeptides. *Biomacromolecules* 18, 599-609.

Manders, E., Stap, J., Brakenhoff, G., Van Driel, R., and Aten, J. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *Journal of cell science* 103, 857-862.

Mao, A.H., Crick, S.L., Vitalis, A., Chicoine, C.L., and Pappu, R.V. (2010). Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A* 107, 8183-8188.

Meyer, D.E., and Chilkoti, A. (2004). Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules* 5, 846-851.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163, 123-133.

Moosmann, P., and Rusconi, S. (1996). Alpha complementation of LacZ in mammalian cells. *Nucleic Acids Res* 24, 1171-1172.

Muiznieks, L.D., and Keeley, F.W. (2010). Proline periodicity modulates the self-assembly properties of elastin-like polypeptides. *J Biol Chem* 285, 39779-39789.

Nayeem, A., Chiang, S.J., Liu, S.W., Sun, Y., You, L., and Basch, J. (2009). Engineering enzymes for improved catalytic efficiency: a computational study of site mutagenesis in epothilone-B hydroxylase. *Protein Eng Des Sel* 22, 257-266.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., *et al.* (2015). Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57, 936-947.

Pak, C.W., Kosno, M., Holehouse, A.S., Padrick, S.B., Mittal, A., Ali, R., Yunus, A.A., Liu, D.R., Pappu, R.V., and Rosen, M.K. (2016). Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* 63, 72-85.

Pastuszka, M.K., Janib, S.M., Weitzhandler, I., Okamoto, C.T., Hamm-Alvarez, S., and Mackay, J.A. (2012). A tunable and reversible platform for the intracellular formation of genetically engineered protein microdomains. *Biomacromolecules* 13, 3439-3444.

Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F., *et al.* (2018). FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. *Cell* 173, 720-734 e715.

Quiroz, F.G., and Chilkoti, A. (2015). Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat Mater* 14, 1164-1171.

Rubinstein, M., and Colby, R.H. (2003). *Polymer Physics* (Oxford University Press).

Shi, P., Lin, Y.A., Pastuszka, M., Cui, H., and Mackay, J.A. (2014). Triggered sorting and co-assembly of genetically engineered protein microdomains in the cytoplasm. *Adv Mater* 26, 449-454.

Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* 357.

Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N., *et al.* (2007). DisProt: the Database of Disordered Proteins. *Nucleic Acids Res* 35, D786-793.

Strulson, C.A., Molden, R.C., Keating, C.D., and Bevilacqua, P.C. (2012). RNA catalysis through compartmentalization. *Nat Chem* 4, 941-946.

Theillet, F.X., Kalmar, L., Tompa, P., Han, K.H., Selenko, P., Dunker, A.K., Daughdrill, G.W., and Uversky, V.N. (2013). The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord Proteins* 1, e24360.

Urry, D.W., Gowda, D.C., Parker, T.M., Luan, C.H., Reid, M.C., Harris, C.M., Pattanaik, A., and Harris, R.D. (1992). Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* 32, 1243-1250.

Uversky, V.N., and Dunker, A.K. (2010). Understanding protein non-folding. *Biochim Biophys Acta* 1804, 1231-1264.

Uversky, V.N., Kuznetsova, I.M., Turoverov, K.K., and Zaslavsky, B. (2015). Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. *FEBS Lett* 589, 15-22.

Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., *et al.* (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* 174, 688-699 e616.

Wimley, W.C., and White, S.H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Structural & Molecular Biology* 3, 842-848.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16, 18-29.