Synthetic ribonucleoprotein granules regulate translation of a target mRNA in living cells

Seyed Eghtasadi
Duke University
Daniel Shapiro
Duke UNiversity
Sonal Deshpande
Duke University
Junseon Min
Duke University
David Fiflis
Duke University
Dahlia Rohm
Duke University
Yifan Dai
Duke University
Aravind Asokan
Duke University
Charles Gersbach
Duke University https://orcid.org/0000-0003-1478-4013
Ashutosh Chilkoti (chilkoti@duke.edu)
Duke University https://orcid.org/0000-0003-1301-1836

Article

Keywords:

Posted Date: September 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1982007/v1

Synthetic ribonucleoprotein granules regulate translation of a target mRNA in living cells

Seyed Ali Eghtesadi[1*], Daniel Mark Shapiro[1*], Sonal Deshpande[1], Junseon Min[1], David Fiflis[1,2], Dahlia Rohm[1,4], Yifan Dai[1], Aravind Asokan[1,2,3], Charles Gersbach[1,2,4], and Ashutosh Chilkoti[1,#]

[1]Department of Biomedical Engineering, Duke University, Durham, NC 27708, United States
[2]Department of Surgery, Duke University Medical Center, Durham, NC, 27708, United States
[3]Department of Molecular Genetics & Microbiology, Duke University School of Medicine, Durham, NC, 27708, United States
[4]Center for Advanced Genomic Technologies, Duke University, Durham, North Carolina 27708, United States
*co-first authors
[#] corresponding author: chilkoti@duke.edu

Abstract. Biomolecular condensates composed of proteins and RNA, known as ribonucleoprotein granules, are one approach by which cells regulate post-transcriptional gene expression. The formation of ribonucleoprotein granules typically involves the liquid-liquid phase separation of intrinsically disordered proteins with a target mRNA, sequestering the mRNA into the condensate. This sequestration regulates gene expression by inhibiting translation or facilitating RNA processing, such as splicing. Here, we designed a recombinant fusion of the human Pumilio2 homology domain (Pum2) RNA-binding protein and a synthetic intrinsically disordered protein that exhibits liquid-liquid phase separation to create synthetic ribonucleoprotein granules that extrinsically regulate the expression of a target gene. We show that this fusion protein selectively binds an RNA transcript of interest that contains a Pum2-binding RNA sequence at its 3'-end and sequesters the mRNA within a biomolecular condensate. Sequestration of a target mRNA within the condensate largely reduces the translation of the target RNA in protocells and in *E. coli* compared to cells that do not sequester the target mRNA in a synthetic condensate. We demonstrate that the viability of *E. coli* that harbor a cytotoxic protein can be rescued by sequestering the mRNA encoding the cytotoxic protein within a synthetic condensate. Finally, we use RNA-seq to determine that the target RNA of interest is preferentially sequestered in synthetic ribonucleoprotein granules. This approach enables modulation of cell function via the formation of a synthetic biomolecular condensate that spatiotemporally regulates the expression of a target protein.

Introduction

Ribonucleoprotein granules (RNPGs) are biomolecular condensates comprised of proteins and RNA that enable cells to spatiotemporally control numerous biochemical reactions involving RNA. These reactions include RNA splicing[1], RNA transport[2], and regulation of gene expression[3]. RNPG's selectively sequester mRNA into a condensate to upregulate or downregulate certain genes in response to external stimuli such as temperature[4] or stress[5-8]. In many cases, RNPG's

form through liquid-liquid phase separation (LLPS), in which specific proteins and RNA undergo phase separation to form a condensate—a protein-and RNA-rich dense phase—that is spatially distinct from the dilute phase—the cytoplasm[8-13].

RNPGs typically exhibit many properties of liquid droplets, such as fusion with other condensates[14] and shearing and bubbling[15] and exhibit rapid exchange of contents from within the condensate to their surroundings[16]. Most proteins involved in forming RNPGs contain significant stretches of amino acids with no three-dimensional structure called intrinsically disordered regions (IDRs). Proteins with a large proportion of IDRs, known as intrinsically disordered proteins (IDPs), are known to phase separate in solution[17-23]. Furthermore, many disease-associated mutations in IDPs implicated in the formation of RNPGs, such as mutations in the Fused-in-Sarcoma (FUS) protein that are implicated in amyotrophic lateral sclerosis (ALS), are located in the proteins' IDRs[10,24-26]. These mutations have been shown to abolish the liquid-like nature of RNPGs and cause them to exhibit arrested, non-liquid behaviors that are implicated in the pathology of the disease[25,26].

Though nature uses RNPGs to spatiotemporally regulate gene expression and significant effort has been, and continues to be, expended in understanding how native RNPGs are formed and regulated in cells[8,20,27-31], designing synthetic RNPGs to modulate the expression of target proteins has received significantly less attention. We were motivated to address this gap by the belief that engineered RNPGs that enable a target mRNA to be reversibly sequestered within a cell could provide a new tool for synthetic biology that is distinct from and conceptually orthogonal to gene circuits that rely on engineering transcriptional regulatory elements[32-36].

There are several challenges that need to be addressed to engineer functional synthetic RNPGs. First, RNPGs are highly heterogenous, with some constituents that are still unidentified[37]. Second, researchers are currently limited to a few well-characterized IDPs that undergo phase separation to create RNPGs, such as FUS[26,38], HNRNPA1[23,39], or LAF-1[40], and many of these proteins participate in the formation of native condensates. Using these IDPs to create synthetic RNPGs may interfere with native condensates, which may have unanticipated functional consequences on cell function. We hence sought to take an orthogonal approach by using synthetic IDPs to create a synthetic RNPG that specifically sequesters a target mRNA of interest without interfering with normal cell function.

To target a specific mRNA, we chose to use a variant of the Pumilio homology domain previously described as Pum2[41], which is a portion of the human Pumilio2 RNA-binding protein that binds an 8-nucleotide RNA sequence with high specificity and binding affinity[41,42]. The canonical Pum2 recognition sequence (PRS), 5'-$U_1G_2U_3A_4\mathbf{N}_5A_6U_7A_8$-3', was inserted at the 3' end of a gene encoding a target RNA. We chose to use Pum2 to target RNA because its binding to RNA is well characterized[42], it can be expressed in *E. coli* and in eukaryotic cells as a soluble protein[28,43] which is essential for condensate formation via LLPS, and it can be engineered at the molecular level to alter its affinity and specificity for its target RNA[41]. To create condensates of Pum2, we chose elastin-like polypeptides (ELPs), a synthetic IDP[44-48], to fuse to Pum2 and drive the formation of a condensate via LLPS[49]. ELPs consist of repeats of the pentameric amino acid sequence (Val-Pro-Gly-Xaa-Gly)$_n$, where Xaa can be any amino acid except proline, and exhibit well-characterized, tunable, and temperature-responsive LLPS[45].

We show that Pum2-ELP selectively binds RNA transcripts tagged with a PRS at their 3' end. The Pum2-ELP fusion sequesters PRS-tagged RNA in a condensate *in vitro* in protocells containing an *in vitro* transcription-translation (IVTT) mixture. Sequestration of PRS-tagged mRNA encoding mCherry or superfolder green fluorescent protein (sfGFP) into a biomolecular condensate in an IVTT-containing protocell specifically inhibits expression of these proteins even in the presence of other RNA. Building on these *in vitro* results, we demonstrate that the synthetic RNPG can modulate the expression of target genes in *E. coli*. We show that Pum2-ELP forms condensates in *E. coli* that sequester mCherry-PRS encoding mRNA and reduces translation of mCherry compared to cells expressing mCherry mRNA without a PRS. To demonstrate modulation of cell function by the synthetic RNPG, we show that the Pum2-ELP fusion can sequester mRNA with a PRS that encodes a cytotoxic protein, Protein E[50,51], and rescue cell viability compared to cells that express protein E from mRNA without a PRS. Finally, we purify synthetic RNPGs and use RNA-seq to confirm that the synthetic RNPGs preferentially sequester a PRS-tagged RNA of interest. Together, these results demonstrate the creation of a synthetic RNPG that can modulate cell function using only a minimal number of biomolecular components whose properties can be precisely tuned.

Results

RNA binding protein-IDP design. We sought to design RNPGs that can extrinsically control the translation of target mRNA in cell-free expression systems and in *E. coli in the presence of other mRNA species*. These synthetic RNPGs are formed through the sequestration of target mRNA into a coacervate—the protein-rich dense phase that is formed upon LLPS—of an ELP fused to Pum2[49]. ELPs exhibit lower critical solution temperature (LCST) phase behavior, wherein below their cloud point temperature (T_{CP}) or saturation concentration (C_{sat}), the concentration at which the T_{cp} equals the local temperature, typically 37ºC, ELPs are soluble and above the T_{cp} or C_{sat} they exhibit LLPS[45]. As ELPs are genetically encoded, they can be fused to other proteins at the gene level to impart LLPS behavior to a target protein[49]. Furthermore, because ELPs have minimal interactions with other biomolecules[49], we hypothesized that the LLPS of an ELP fused to an RNA binding protein domain would be orthogonal to native cellular processes. We used an ELP with 160 pentapeptide repeats of Val-Pro-Gly-Xaa-Gly, in which the Xaa residue was comprised of 70% alanine and 30% valine (Supplementary Table 1), and which undergoes LLPS at biologically relevant temperatures (37 °C) and concentrations (< 50 μM**)**.

To specifically bind and sequester a target RNA, we created a recombinant fusion of the ELP fused to the C-terminus of Pum2, which binds the RNA sequence 5'-$U_1G_2U_3A_4\mathbf{N}_5A_6U_7A_8$-3'[52,41]. We also synthesized a Pum2-ELP-GFP fusion to track condensate formation by fluorescence microscopy and two controls—Pum2 fused to His_8 (Pum2-His_8) and the ELP. All proteins were recombinantly expressed in *E. coli* and purified to homogeneity, as seen by SDS-PAGE (Supplementary Figure 1). The Pum2-ELP fusion protein exhibits thermally responsive LCST phase behavior, as observed by temperature-programmed turbidimetry (Figure 1A and Supplementary Figure 2-C). Figure 1A shows the cloud point behavior of Pum2-ELP at different concentrations, suggesting that the T_{CP} decreases with increasing protein concentration, which is consistent with the inverse dependence of the T_{cp} on ELP concentration (Supplementary Figures 2A and 2B)[53].

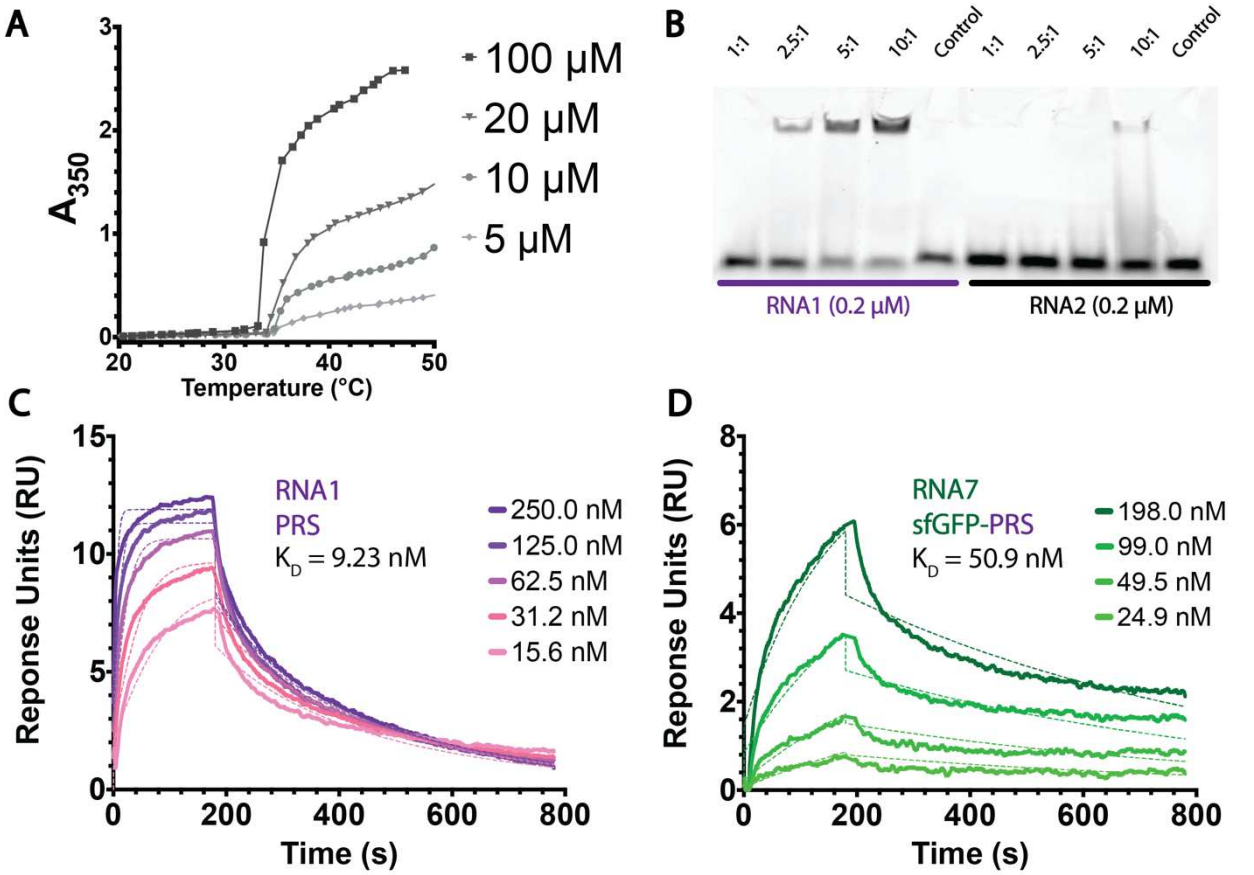

Figure 1. A) T_{CP} as a function of temperature at different concentrations of Pum2-ELP **B)** Electrophoretic mobility shift assay of Pum2-ELP-RNA interaction with RNA1, a 19 bp long RNA that contains a PRS and RNA2, a control that lacks a PRS. The RNA concentration was 0.2 µM and the Pum2-ELP concentration was varied from 1 to 10-fold of the RNA concentration. **C, D)** SPR analysis of Pum2-ELP and RNA interactions. **C)** The association and dissociation curves of RNA 1 binding to Pum2-ELP. **D)** The association and dissociation curves of RNA7, an mRNA that encodes sfGFP with 5 repeats of a PRS at the 3'end (sequence in Supplementary Table 2). The solid lines indicate raw data and the fits to the curves are shown by dashed lines.

Interaction of Pum2-ELP with target RNA. We first examined the interaction of Pum2-ELP with RNA that contains a PRS and a negative control that lacks a PRS by an electrophoretic mobility shift assay. We selected a 19 bp oligonucleotide long RNA containing a PRS—5'-UGUACAUA-3'—at its 3'-end (RNA1) and a negative control (RNA2) that presents a 5'-UUGAUAUA-3' sequence at its 3'-end (Supplementary Table 2)[54]. Pum2-ELP binds RNA1 and retards its migration on a native PAGE gel, indicating strong interactions between Pum2-ELP and RNA1 (Figure1B, lanes labeled in purple). In addition, the extent of RNA1 retardation showed a dependence on the protein:RNA ratio, as the RNA showed greater degree of retardation at a higher protein:RNA ratio. In contrast, Pum2-ELP does not retard the migration of RNA 2 up to a 10:1 ratio of Pum2-ELP to RNA2, where we observed some retardation of RNA2 (Figure1B, lanes labeled in black), suggesting a low level of non-specific interactions between Pum2-ELP and the control RNA at a large molar excess of Pum2-ELP.

To quantify RNA binding to Pum2, we used surface plasmon resonance (SPR) spectroscopy to measure the binding affinity of RNA1 and RNA2 to ELP-Pum2. We found that RNA1 shows strong binding to the immobilized protein with an equilibrium dissociation constant (K_D) of 9.23 nM, a value that is consistent with previous studies[42,54] (Figure 1C), while RNA2 shows minimal interaction with the ELP-Pum2 fusion (Supplementary Figure 3 and Supplementary Table 3).

We next examined the ability of Pum2-ELP to form condensates with PRS-tagged RNA in a protocell that can carry out transcription and translation in vitro. To do so, we first constructed plasmids encoding mCherry-PRS and sfGFP-PRS with five copies of the PRS, or negative controls that encode the fluorescent protein but lack the PRS using recursive directional ligation by plasmid reconstitution (Pre-RDL)[55]. We then used the plasmids as a template to transcribe PRS-tagged mRNA *in vitro* (sequences in Supplementary Table 2). We confirmed by SPR that ELP-Pum2 specifically binds sfGFP-PRS mRNA with a K_D of 50.9 nM (RNA7 in Figure 1D). The slightly lower affinity of RNA7 for ELP-Pum2 compared to RNA1 is likely due to the larger size of RNA7 (~1 kb) compared to RNA1 (19 bp). There were no significant interactions between ELP-Pum2 and purified sfGFP or mCherry mRNA lacking the PRS, as seen by SPR (Supplementary Figure 3, RNA3 and RNA4).

We then used a microfluidic device to fabricate protocells—50 μM diameter aqueous droplets in oil—that contain an *in vitro* transcription-translation (IVTT) mixture[56]. These IVTT protocells provide a well-defined and uniform microenvironment in which we can observe the formation of biomolecular condensates and the expression of target mRNA as a function of LLPS by precisely controlling the environmental temperature[56]. To observe the formation of Pum2-ELP condensates, we fused sfGFP to the C-terminus of the Pum2-ELP fusion protein (Pum2-ELP-sfGFP) and observed the formation of a protein-rich condensate of Pum2-ELP-sfGFP at 37°C by fluorescence microscopy (Supplementary Figure 4).

mCherry RNA and mCherry-PRS mRNA were tagged with FITC, separately mixed with Pum2-ELP and an IVTT mixture and formed into protocells. Protocells were imaged at 37°C by fluorescence microscopy immediately to eliminate any signal from translation of the mCherry RNA. We found that incubation of mCherry-FITC RNA or mCherry-PRS-FITC RNA with Pum2-ELP (without a sfGFP tag) in separate IVTT protocells resulted in spatially localized fluorescence in protocells containing mCherry-PRS-FITC RNA, whereas protocells that contained mCherry-FITC RNA displayed homogeneous fluorescence throughout the protocell, indicating that the Pum2-ELP fusion only binds and sequesters mRNA that is tagged with the PRS (Supplementary Figure 5).

To investigate the sequestration of PRS-tagged mRNA by Pum2-ELP in a more complex, binary mRNA mixture, we mixed Pum2-ELP with Cy3-labeled mCherry mRNA with a PRS tag at its 3' end (mCherry-PRS-Cy3 RNA) and a negative control, FITC-labeled mCherry mRNA (mCherry-FITC RNA) in an IVTT protocell. We raised the temperature to 37ºC to drive condensate formation and observed that the mCherry-PRS-Cy3 RNA was sequestered by Pum2-ELP into condensates while the mCherry-FITC RNA remained soluble throughout the protocell (Figure 2A and Supporting video 1). These results confirmed that mRNA binding by Pum2-ELP and sequestration of the mCherry-PRS-Cy3 RNA in the Pum2-ELP condensates is due to the PRS RNA tag (Figure 2A).

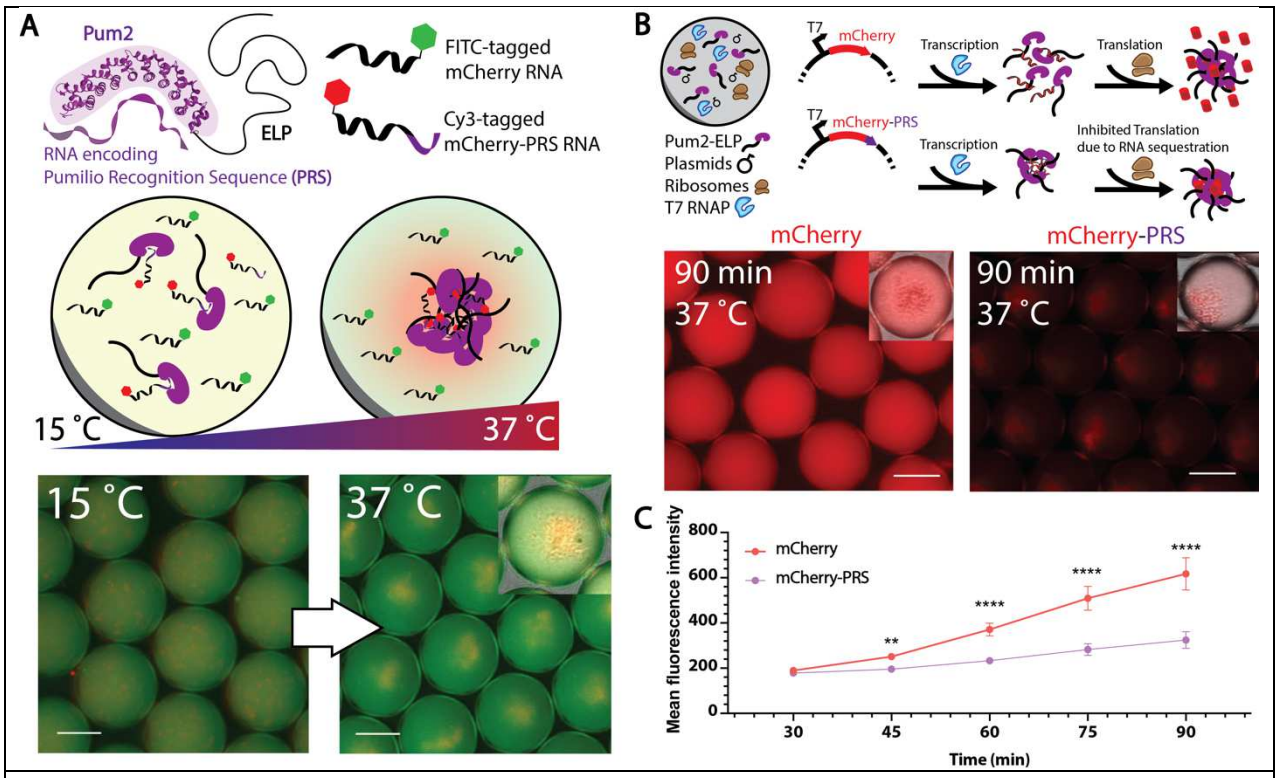

Figure 2: Selective partitioning of target mRNA into synthetic condensates. Fluorescence micrographs for **A)** Selective sequestration of target mRNA (red) by Pum2-ELP fusion in presence of negative control mRNA (green). mCherry encoding mRNA with the PRS sequence was labeled with Cy3 (mCherry-PRS-Cy3 RNA), and mCherry encoding mRNA without the PRS was labeled with FITC (mCherry-FITC RNA). The mixture was incubated with Pum2-ELP in a protocell, and the temperature was raised to 37°C (Supporting video 1). Scale bar in protocell images is 50 µm. **B)** Schematic of the protocell components (top) used to study the effect of mRNA-Pum2-ELP co-localization on expression of mCherry (red) from plasmids encoding mCherry (bottom left) and mCherry-PRS (bottom right) in a protocell containing IVTT mixture after 90 min at 37 ºC. The insets in the fluorescence images in panels A and B show the overlay of mCherry and brightfield channels. Scale bar is 50 µm. **C)** Quantification of mCherry expression by fluorescence image analysis of the protocells with plasmid encoding mCherry and mCherry-PRS in the presence of Pum2-ELP and IVTT mixture at 37 ºC. Mean fluorescence intensity for each protocell was monitored for 90 min. Two-way ANOVA was performed with Sidak's multiple comparison test. **: $p<0.005$, ****<0.0001. Error bars indicate standard deviation for n=10 protocells.

***In vitro* regulation of protein translation in protocells.** Having shown that an mRNA containing a PRS can be captured by a Pum2-ELP fusion and sequestered in a synthetic condensate, we next investigated whether this system could be used to regulate gene expression *in vitro*. To do so, we introduced plasmids encoding mCherry-PRS or mCherry into protocells containing an IVTT mixture and Pum2-ELP. The protocells were incubated at 37°C and the expression level of mCherry was recorded by fluorescence microscopy every 15 min. Visually, protocells containing plasmids encoding mCherry-PRS mRNA showed reduced expression of mCherry protein that was localized to the Pum2-ELP condensate compared to protocells containing plasmids encoding for mCherry mRNA without the PRS, which displayed stronger and uniform mCherry fluorescence throughout the protocell (Figure 2B). We quantified the expression of mCherry by analyzing the

fluorescence intensity of protocells (n=10) over time and found that Pum2-ELP significantly inhibited the expression of mCherry-PRS for up to 90 min (Figure 2C).

We next sought to investigate whether Pum2-ELP can selectively regulate the gene expression of a target mRNA in a more complex milieu that contains another mRNA. To do so, we created plasmids encoding sfGFP, sfGFP-PRS, mCherry and mCherry-PRS (Figure 3A, Supplementary Table 2). We then created protocells containing an IVTT mixture, Pum2-ELP protein, and four binary combinations of these plasmids: mCherry and sfGFP-PRS (protocell I), mCherry-PRS and sfGFP (protocell II), mCherry and sfGFP (protocell III), and mCherry-PRS and sfGFP-PRS (protocell IV) (Supplementary Table 4). The expression of mCherry and sfGFP protein were monitored in the protocells by fluorescence microscopy. We found that the expression of protein from mRNA encoding a PRS was largely limited to the protein-rich condensate, while expression of protein from mRNA lacking a PRS site was homogenous throughout the protocell (Figure 3B). In protocell I, the expression of sfGFP from sfGFP-PRS mRNA was limited to the protein-rich dense phase, while expression of mCherry from mCherry mRNA was uniform throughout the protocell (Supporting video 2). In protocell II, which contains plasmids encoding mCherry-PRS and sfGFP, the reverse is true—expression of mCherry in protocells was localized to the condensate, while the expression of sfGFP was uniform (Supporting video 3). In protocell III, which contains plasmids encoding sfGFP and mCherry mRNA without a PRS, both proteins were expressed homogenously throughout the protocells (Supporting video 4). In protocell IV, the presence of a PRS site in both sfGFP and mCherry mRNA resulted in localization of sfGFP and mCherry protein largely to the condensate (Supporting video 5).

We quantified the expression of the proteins in the protocells by fluorescence image analysis using ImageJ (Figure 3C). We calculated the integrated fluorescence density of protocells, defined as the sum of intensities of pixels across the image, for each channel. We chose to carry out a ratiometric analysis of the integrated fluorescence density for the different protocell types, instead of comparing mean fluorescence intensity of the protocells. This is because the mean fluorescence intensity assigns brighter values for dim pixels and vice-versa, whereas the integrated density provides a more accurate representation of the brightness of each pixel in the protocell. The presence of a PRS in the target mRNA significantly decreased the expression of the respective fluorescent proteins in protocell I (sfGFP), II (mCherry) and IV (mCherry and sfGFP) as compared to protocell III ($p<0.0001$).

We also carried out confocal fluorescence microscopy to validate these findings. The confocal fluorescence images for protocell II (sfGFP and mCherry-PRS) showed localized expression of mCherry in the condensate (inset of Protocell II in Figure 3B and Supplementary Figure 6) whereas the expression of mCherry and sfGFP encoded by mRNA lacking a PRS was homogeneous throughout the protocell (inset of protocell III in Figure 3B). The fluorescence intensity from these 3-D confocal images was integrated to obtain an area under the curve (AUC) metric to quantify the volumetric fluorescence of the protocells (Figure 3D). The ratio of the AUC red:green channels significantly ($p<0.0001$) decreased in protocells that encode for mCherry-PRS and sfGFP mRNA (protocell II) compared to protocells that encode for mCherry and sfGFP mRNA (protocell III) (Figure 3D). This quantitatively confirms the reduced translation of mRNA that is localized in the condensate relative to mRNA that is translated throughout the protocell.

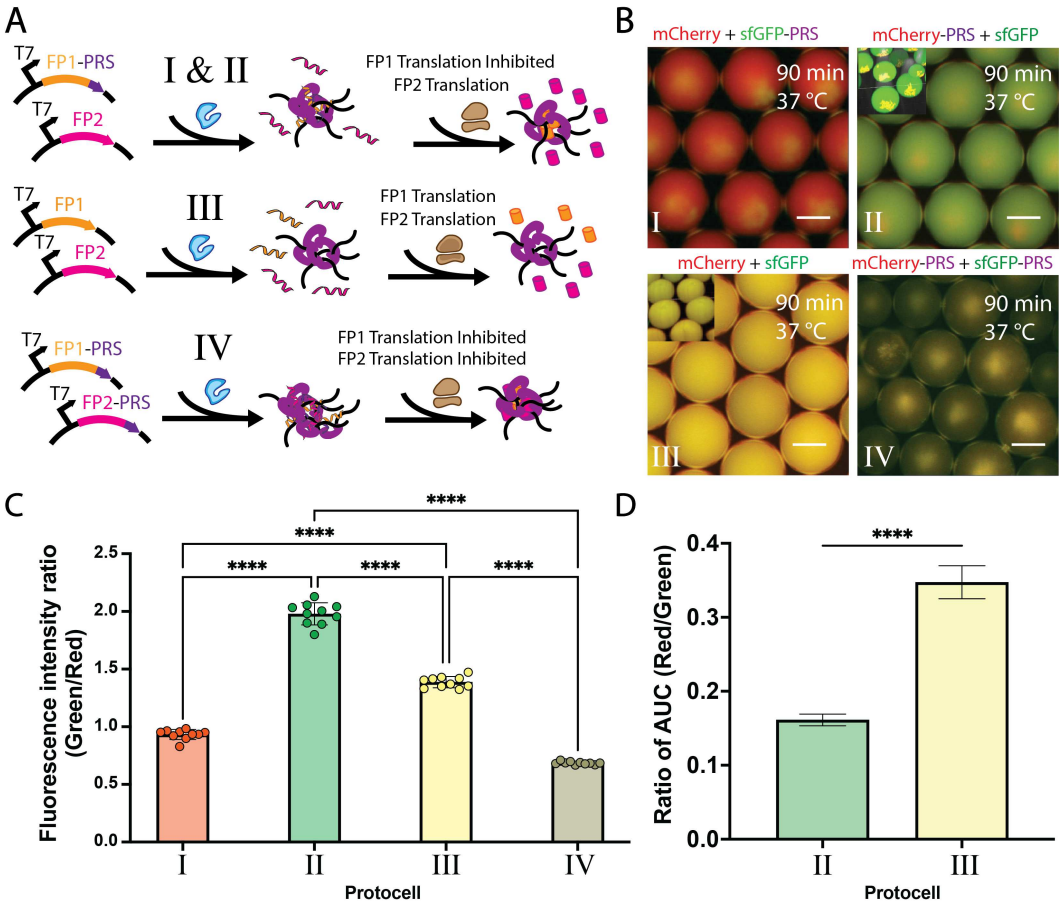

Figure 3: Selective regulation of translation by sequestration of a target RNA within a synthetic RNPG. A) Schematic of regulation of fluorescent protein (FP) expression in synthetic RNPGs. In protocells I and II, the mRNA of one FP contained a PRS and was sequestered, inhibiting translation only of that FP. In protocell III, neither FP mRNA contained a PRS and both proteins were expressed. In protocell IV both FP mRNA had a PRS, leading to sequestration of both mRNA species and inhibition of translation of both proteins. **B)** Fluorescence microscopy images of protocells containing IVTT mixture, Pum2-ELP and plasmids encoding for I: mCherry and sfGFP-PBD, II: sfGFP and mCherry-PBD, III: mCherry and sfGFP and IV: mCherry-PBD and sfGFP-PBD after 90 min at 37ºC. Objective= 20×. Scale bar= 50 µm, (Supporting video 2,3,4 and 5). **C)** Quantification of relative expression of FP in protocells I-IV. Fluorescence images were analyzed by ImageJ to obtain integrated density (mean fluorescence intensity×area) for sfGFP and mCherry. Statistical analysis: Two-way ANOVA with Tukey's multiple comparison test; ****=significantly different from all other protocells ($p<0.0001$). **D)** Ratiometric analysis of 3-D confocal fluorescence microscopy images of protocells; example confocal images are shown in the inset of protocells II and III in panel B. The mean intensity of protocells (n=10) was analyzed along the Z-axis using ImageJ for the green and red channel. The area under the curve (AUC) for each channel was computed by integrating the mean fluorescence intensity along the Z-axis. Data is reported as the ratio of the AUC of the green channel to the red channel (See Supplementary Figure 7 for details) Statistical analysis: Student's t-test; ****:$p<0.0001$. Error bars indicate standard deviation between n=10 protocells.

Finally, to confirm that Pum2 or ELP alone cannot inhibit expression, control experiments were performed with both sfGFP-PRS mRNA and mCherry mRNA lacking a PRS in the presence of ELP or Pum2. We found that both sfGFP and mCherry were expressed throughout the protocells

(Supplementary Figure 7 and Supplementary Table 4, experiments 6 and 7). Collectively, these experiments confirm that binding of a target mRNA by a Pum2-ELP fusion and its sequestration into a condensate are necessary to spatially localize its translation within a condensate and suppress its translation.

Translational regulation in live cells. Having demonstrated that a mRNA bearing a PRS can be selectively sequestered—and its expression suppressed—within a synthetic condensate in vitro, we next sought to investigate the function of this condensate *in vivo*. To do so, we designed a plasmid encoding Pum2-ELP-sfGFP under control of a T7 promoter and co-transformed it into BL21(DE3) *E. coli.* with a plasmid encoding mCherry (Figure 4A(i)) or mCherry-PRS (Figure 4A(ii)) under control of an arabinose-inducible pBAD promoter (Supplementary table 5, cells I and II). We grew the *E. coli* until an OD_{600} of 0.8 was reached, and then induced expression of Pum2-ELP-sfGFP with isopropyl ß-D-1-thiogalactopyranoside (IPTG) and mCherry or mCherry-PRS with arabinose and incubated the cells at 37 °C for another 22 hours. Cells expressing Pum2-ELP-sfGFP with either mCherry (Figure 4A(i)) or mCherry-PRS (Figure 4A(ii)) RNA showed green intracellular puncta, suggesting the formation of an intracellular condensate of Pum2-ELP-sfGFP in both cell populations. The fluorescence intensity of mCherry protein, however, was significantly reduced in cells expressing mCherry-PRS RNA (Figure 4B) compared to cells that expressed mCherry only (Figure 4A(i)), illustrating the role of Pum2-ELP condensates in suppressing the expression of mCherry by sequestering its mRNA.

Control experiments confirmed the expression of ELP-sfGFP condensates in cells without Pum2 (Supplementary table 5, cells III and IV), and showed that ELP-sfGFP condensates do not affect the expression of mCherry protein from mRNA with or without a PRS (Supplementary Figure 8), indicating the need for a condensate-forming Pum2 to suppress translation of a mRNA that contains a PRS. Co-expression of sfGFP-Pum2 protein and mCherry from mRNA with or without the PRS tag led to similar levels of expression of both mCherry (Supplementary table 5, cells V and VI). We also found that translation of mCherry was not affected by translation of Pum2-sfGFP. The translation of Pum2-sfGFP is observed as a homogenous green color throughout the cell (Supplementary Figure 9).

To investigate whether the synthetic RNPG could modulate the function of a protein beyond expression of a fluorescent reporter protein, we investigated whether we could control life and death—cell viability—by a synthetic RNPG. To test this, we replaced the plasmid encoding mCherry with one encoding Protein E with and without a PRS. We chose Protein E as its expression causes cell lysis[50,51,57]. Furthermore, its cytotoxic effect on *E. coli* does not appear to interfere with the cell's transcription or translation machinery, allowing us to precisely determine if cell viability is impaired through the localization of Protein E mRNA in a synthetic condensate and not by a different mechanism.

We designed plasmids encoding Protein E or Protein E-PRS (Supplementary table 1, methods) under control of an arabinose-inducible pBAD promoter and co-transformed each plasmid into BL21(DE3) cells with a plasmid encoding for Pum2-ELP-sfGFP under control of a T7 promoter (Figure 4C, Supplementary table 6). As a positive control, we used *E. coli* BL21(DE3) cells transformed with the plasmid encoding Pum2-ELP-sfGFP only. All cell populations were grown

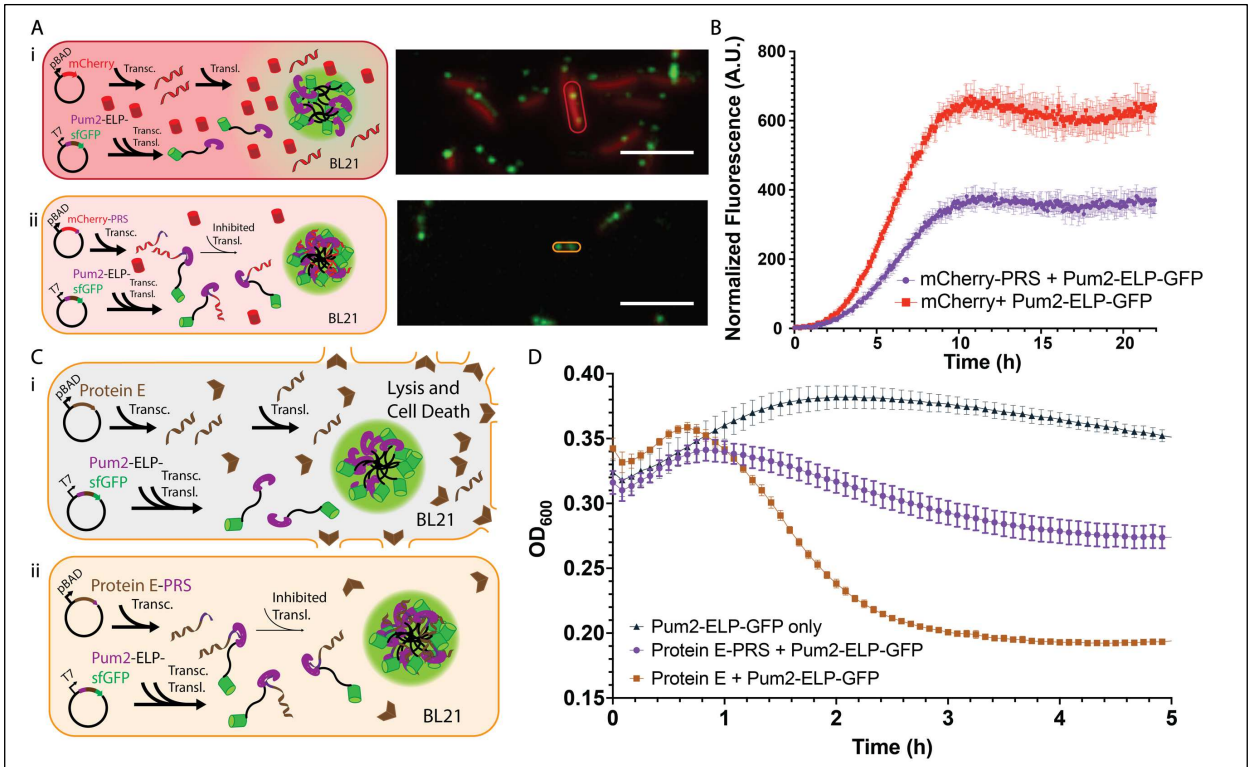

Figure 4: **Intracellular control of translation by synthetic RNPG. A)** Schematic of inhibition of mCherry expression by synthetic RNPG. (i, left) mCherry RNA without a PRS is not sequestered after transcription, and mCherry is expressed throughout the cell. (i, right) Fluorescence microscopy image of cells expressing mCherry and Pum2-ELP-sfGFP. mCherry is seen as a homogeneous red signal throughout the cell, while the Pum2-ELP-sfGFP condensate is seen as a green dot at the poles of the cells. One representative cell is outlined in a red oval. Scale bar: 10 µm. (ii, left) mCherry-PRS RNA is transcribed but is sequestered into the synthetic Pum2-ELP-sfGFP RNPG, inhibiting translation of mCherry. (ii, right) fluorescence microscopy image of cells expressing mCherry-PRS and Pum2-ELP-sfGFP. Inhibited mCherry expression is seen as a very dim red signal throughout the cell, while the Pum2-ELP-sfGFP RNPGs are seen as bright green dots at the poles of the cell. One representative cell is outlined in an orange oval. Scale bar: 10 µm. **B)** mCherry fluorescence, normalized by OD_{600} at each time point, over 22 h after induction of mCherry and Pum2-ELP-sfGFP by arabinose and IPTG, respectively. **C)** Schematic of inhibition of Protein E expression by synthetic RNPG. (i) Protein E mRNA with no PRS is not sequestered after transcription and Protein E is expressed throughout the cell, leading to lysis and cell death. (ii) Protein E-PRS mRNA is sequestered in a Pum2-ELP-sfGFP RNPG, inhibiting translation of Protein E and rescuing cell viability. **D)** OD_{600} of cells expressing Pum2-ELP-sfGFP only (black line) or Pum2-ELP-sfGFP and Protein E-PRS RNA (purple line) or Protein E RNA (brown line) recorded over 5 h after induction of Pum2-ELP-sfGFP with IPTG and Protein E with arabinose.

to an OD_{600} ~0.3 in aerobic conditions, at which point they were induced with 0.2 mM IPTG and 0.02% Arabinose. Growth was then monitored over time for 5 hours at 37 °C for all samples in a plate reader with samples covered with mineral oil to prevent evaporation of media. We found that the OD_{600} for cells expressing only Pum2-ELP-sfGFP increased for 2 hours, and then slightly decreased and plateaued at around 0.35, likely due to the switch from aerobic to anaerobic conditions in the plate reader impacting cell growth (Figure 4D). In stark contrast, the OD_{600} value for cells expressing Protein E with no PRS and Pum2-ELP-sfGFP sharply fell 1 hour after induction due to expression of the cytolytic Protein E. The OD_{600} for cells expressing Protein E-PRS and Pum2-ELP-sfGFP, however, only slightly decreased 1 h after induction and plateaued at

around 0.27. These data suggest that the decreased expression of Protein E from Protein E-PRS mRNA sequestered in the Pum2-ELP-sfGFP synthetic RNPGs partially rescues cell viability.

Quantification of Sequestration of RNA in live cells. To directly confirm that mCherry mRNA transcripts tagged with a PRS are preferentially sequestered in Pum2-ELP-sfGFP condensates inside cells, we identified and quantified the RNA present in synthetic RNPGs produced in BL21(DE3) *E. coli* expressing Pum2-ELP-sfGFP and either mCherry-PRS mRNA or mCherry mRNA with no PRS tag. We isolated synthetic RNPGs using ultracentrifugation (Supplementary Figure 10) and extracted and purified the constituent RNA from Pum2-ELP-sfGFP-containing fractions by TRIzol extraction. We then constructed cDNA libraries from the purified RNA and used Illumina next-generation sequencing to quantify the relative abundance of all the RNA species found in the sample (Figure 5A). We identified the condensate band visually by the fluorescence of GFP in the ultracentrifuge tubes, collected 1 or 1.5 mL fractions down the length of the tube, and ran aliquots of the fractions on a SDS-PAGE gel (Figure 5B, methods). We found that Pum2-ELP-sfGFP was typically found in fraction 11, corresponding to the darkest protein band at 130.23 kDa (Figure 5B), although in some samples fractions 10 and 11 both contained Pum2-ELP-sfGFP (Supplementary Figure 10). In cases where two fractions contained the condensate, RNA was isolated from both fractions and pooled together before constructing a cDNA library. Synthetic RNPGs were isolated from condensate-containing fractions from 3 biological replicates of each cell population — Pum2-ELP-sfGFP + mCherry and Pum2-ELP-sfGFP + mCherry-PRS (Supplementary Figure 10). The cDNA libraries were diluted to 650 pM and pooled together before sequencing (methods). The resulting reads were normalized and a single combined reads per kilobase million (RPKM) value was calculated for each cell population using Rockhopper bacterial RNA-seq software[58]. From the RPKM values we calculated the relative Transcripts per Million (TPM) value for each gene using the formula TPM = 10^6 * (individual transcript RPKM)/(Sum(all RPKM values). TPM allows us to compare the relative abundance of RNA within a sample and, satisfying a few conditions, between samples. Our samples satisfied the conditions for comparing TPM values between samples: the samples were sequenced with the same protocol and RNA strandedness, the same RNA isolation approach was taken for all samples, the fraction of ribosomal RNAs and top expressed transcripts was similar between samples (Supplementary Figure 11), and the total amount of input RNA was similar[59].

We found that mCherry-encoding RNA was 1.8-fold more relatively abundant in the mCherry-PRS sample compared to the mCherry with no PRS sample, with the TPM of mCherry in the mCherry-PRS sample being 185,119 (sense+antisense TPM) and 101,829 (sense+antisense TPM) in the mCherry with no PRS sample. (Figure 5C**,** Supplementary Figure 11, supplementary attachment). This increase in relative abundance of mCherry mRNA coincided with a decrease in the relative abundance of Pum2-ELP-sfGFP mRNA. Surprisingly, the RNA species with the highest relative abundance in both samples was mRNA encoding Pum2-ELP-sfGFP, with a TPM of 383605.7 in the mCherry-PRS sample and 517107.027 in the mCherry with no PRS sample (Figure 5C).

As a control, we investigated the abundance of RNA in condensates from cells expressing just Pum2-ELP-sfGFP without mCherry (Supplementary Figure 12). Pum2-ELP-sfGFP was again the most abundant, with a TPM of 611696.5. The difference in TPM of Pum2-ELP-sfGFP between samples not expressing mCherry at all and those expressing mCherry-PRS, 228,090.8, is close to

the relative abundance of mCherry in the mCherry-PRS sample, 185,119, while the TPM values of other highly abundant genes remained unchanged. These results support our claim that the overabundance of mCherry mRNA present in the mCherry-PRS sample is due to RNA sequestered in the synthetic condensates.

While ribosomal RNA makes up almost 90% of the total RNA in *E. coli* and thus is a common presence in bacterial RNA-seq experiments[60], our *in vitro* experiments showed localized expression of mCherry in Pum2-ELP condensates, suggesting that the Pum2-ELP-sfGFP condensates may sequester ribosomes actively translating mCherry from the sequestered mCherry-PRS RNA. This is consistent with the fact that the abundance of mCherry RNA is on the same level as ribosomal RNA, further supporting our model of RNPG-mediated RNA sequestration.

Our assertion that the ribosomal RNA present in the RNA-seq data is more than just non-specific contamination is further supported by the other proteins potentially present in the ultracentrifuge fractions used for RNA-seq. Indeed, by searching for ribosomal proteins abundant in *E. coli* cytosol[61], we identified a number of ribosomal proteins that may be present in the sample. The SDS-PAGE gel of collected ultracentrifuge fractions (Figure 5B), among other faint bands, shows dark protein bands between 37-50kDA and around 25kDa. The protein Elongation Factor Tu (EF-Tu) is involved in protein synthesis and is 43 kDa, potentially corresponding to the second brightest band in fraction 11. Furthermore, ribosomal proteins 50S L1 (24.6 kDa) and 30S S2 (26.6 kDa) and S4 (23.3 kDa) may make up the dark band at 25 kDa.

Other genes predicted to contain a PRS site in the *E. coli* BL21 (DE3) genome (Supplementary text) are not detected by RNA-seq in any significant amount (Supplementary text), suggesting that the sequestration of a target mRNA into a synthetic RNPG is specific and orthogonal provided that the target mRNA has multiple PRS sites; in our case, mCherry-PRS contained 5 PRS sites. Collectively, the experiments confirm the sequestration of PRS-tagged mRNA within synthetic Pum2-ELP-sfGFP RNPGs.

Discussion

In a previous study, we took the first step toward creating a synthetic RNPG *in vitro* wherein we sequestered a single species of mRNA encoding sfGFP using a promiscuous RNA-binding protein that non-specifically bound RNA[56]. We did so by fusing a promiscuous RNA-binding domain to an ELP and showed that the LLPS of the ELP fusion could modulate expression of a protein from mRNA in IVTT protocells. This study was a necessary, albeit trivial, demonstration of the feasibility of using ELPs to construct a RNPG that only contained a single mRNA and served as a precursor to this more sophisticated study, wherein we sought to create a synthetic RNPG that can selectively and orthogonally target and sequester a single mRNA species out of a complex cellular environment that contains myriad other mRNA molecules.

In this study, we demonstrate that we can extrinsically control the expression of a protein from a target mRNA by a synthetic RNPG even in the presence of other mRNA species. We did so by designing a recombinant fusion of Pum2 and a synIDP, an ELP, that exhibits LLPS. We chose ELPs as the synIDP to drive the formation of synthetic condensates in living cells for several reasons. First ELPs are relatively inert; with a VPGXG repeat motif, they have a low propensity for intermolecular secondary interaction and consequently do not show significant interactions

with other proteins[47]. Second, ELPs have been successfully fused with various protein domains while retaining their solubility and LLPS behavior, suggesting that they are a robust platform for imparting LLPS to diverse RNA-binding protein domains and potentially to other proteins that regulate cellular function beyond control of translation[49]. Third, the LLPS of ELPs has been extensively characterized and their LCST phase behavior can be precisely tuned at the molecular level by control of two orthogonal variables—its sequence and chain length[45,53].

To confer the ELP with the capacity to specifically bind and sequester a target RNA, we examined several RNA binding proteins. Most of these proteins have RNA binding domains that recognize specific sequences in target RNAs[62]. Common RNA binding domains that specifically bind single stranded RNA include K homology domains (KH domains), zinc-finger proteins such as the Com protein[63], RNA recognition motifs (RRMs), the DEAD/DEAH box[64], Pentatricopeptide-repeat (PPR) domains[65], and Pumilio/FBF domains (PUFs)[52]. Other common RNA-binding proteins used in synthetic biology applications include the MS2 and PP7 phage Capsid Proteins (MCP/PCP, respectively) which specifically bind RNA stem loops[34].

We focused our attention on the Pumilio RNA-binding protein family (PUF family), a large family of RNA binding proteins found in eukaryotes[66] that are mainly involved in post-transcriptional regulation of mRNA stability and translation by binding to specific regulatory elements on their mRNA targets[52,67,68]. There are two members of PUM RBPs that are found in humans, Pumilio1 and Pumilio2, which play roles in modulating the function of hundreds of different mRNAs in variety of human tissues[69]. We chose the RNA-binding domain of the Pumilio2 protein, designated Pum2 in this study, for four reasons. First, its interactions with specific RNA sequences are well characterized. Pum2 comprises eight α-helical PUF repeats (36 amino acids in total) and 2 pseudo-repeats at the N and C termini, respectively, which bind to mRNA in an anti-parallel orientation through an inner concave surface[42]. Pum2 binds the sequence 5'-$U_1G_2U_3A_4\mathbf{N}_5A_6U_7A_8$-3', where each of the eight PUF repeats recognizes one nucleotide, located within the 3′UTRs of its target mRNA[43] with nanomolar affinity and high specificity. The interaction of the Pum2 protein with short RNA sequences containing 5'-$U_1G_2U_3A_4\mathbf{N}_5A_6U_7A_8$-3' has been characterized in earlier studies. For instance, Lu et al. studied the RNA:PUM affinities of the RNA-binding domains of Pumilio1 and Pumilio2 (Pum1 and Pum2, respectively) with different $\mathbf{N_5}$ nucleotides in the binding sequence[42]. Cheong et al. have also compared the affinity of Pum RNA-binding domains to various RNA targets (5'-$U_1G_2U_3A_4\mathbf{U}_5A_6U_7A_8$-3') and showed that by changing the characteristic $U_1G_2U_3$ triplet typically found in Pum target RNA to $U_1U_2G_3$, the binding decreases 3300-fold[54]. Second, Pum2 and its analog Pum1 have previously been recombinantly expressed in *E. coli* as soluble proteins, which is a prerequisite to build a synthetic RNPG[28]; our efforts to use common proteins such as MCP or PCP were hindered by insolubility issues. Third, Pum2 is thermally stable between 4°C and 40°C, allowing it to bind RNA across the temperature range of ELP phase behavior. Finally, since Pum2 is derived from a human RNA-binding protein, it is likely to have few interactions in *E. coli* beyond some off-target RNA-binding, allowing us to create biorthogonal synthetic RNPGs. We recognize that one limitation of Pum2 is that its RNA recognition sequence is too short to provide perfect binding specificity and could lead to the off-target capture of native mRNA within the synthetic RNPG in cells, an issue we investigated by purifying the synthetic RNPG from E. coli and carrying out RNA-seq to quantify the type and relative concentration of RNA transcripts that are sequestered within the condensate.

We show that this Pum2-ELP fusion protein selectively binds an RNA transcript of interest that contains a Pum2binding RNA sequence, PRS, at its 3'-end and sequesters the mRNA within a biomolecular condensate. Sequestration of the target mRNA within the synthetic RNPG suppresses translation of the target RNA in IVTT protocells, with a low level of expression observed within the condensate. Similar results were also obtained in *E. coli*, wherein cells that sequestered target mRNA tagged with a PRS in a synthetic RNPG showed 2-fold lower expression of protein from the target mRNA compared to cells that do not form the condensate. To demonstrate the functional utility of this synthetic RNPG, we demonstrated that the viability of *E. coli* that harbor a cytotoxic protein can be partially rescued by sequestering the mRNA encoding the cytotoxic protein within a synthetic RNPG.

This study has several antecedents, though it differs from them in significant ways. An elegant set of studies by Lemke and coworkers showed that a synthetic RNPG[70,71] could successfully recruit a target mRNA to a protein condensate to spatially limit the incorporation of non-canonical amino acids in the protein encoded by the mRNA within the condensate, but they used a naturally occurring IDP with a limited range, and thus tunability, of phase behavior. Navarro et al. used Pum1, a homolog of Pum2, to sequester endogenous mRNA into a synthetic condensate[28]. To form the synthetic condensate, they modified ferritin, a protein with 24 subunits, with a protein partner, F36M-FKBP, that enables weak multivalent intermolecular interactions which drive condensate formation at a critical concentration, and a Pum1 domain to recruit endogenous mRNA to the condensate. In contrast, we built a synthetic RNPG using a synthetic and bio-orthogonal synIDP—an ELP—whose LLPS is tunable at the sequence level over a wide range of temperature and concentration. Furthermore, instead of targeting endogenous mRNA, we instead targeted a specific mRNA that contains a PRS, with the goal of building a biorthogonal condensate to sequester a single mRNA species.

While we saw a lower level of expression of the target in cells with the synthetic RNPG than cells that do not, the functional effect, expression of a reporter protein or effect on cell viability, is only ~2-fold. These results suggest that the Pum2-ELP system is more analogous to a "dimmer" rather an on-off switch of protein translation. This is because of two possible reasons. First, it is possible that not all the mRNA of interest is captured by the Pum2-ELP fusion and sequestered into the condensate, and this depends on the concentrations of the two binding partners, the target mRNA and the Pum2-ELP fusion, and their affinity. Even if this possibility could be avoided by overexpressing the Pum2-ELP and ensuring a high enough affinity of the Pum2-mRNA interaction, it may be intrinsically impossible to completely suppress translation by LLPS as the dilute phase has some concentration, albeit low, of the Pum2-ELP fusion that would presumably bind the target mRNA. The captured mRNA in the dilute phase would then still be able to undergo translation, as seen by the translation of the captured mRNA in the condensate *in vitro*, which despite being bound to the Pum2-ELP fusion successfully undergoes translation.

How then do we turn the dimmer into an on-off switch? Several approaches may help to further suppress translation. First, a synIDP whose phase diagram is close to a square well shape, in which the binodal in the dilute regime is close to the abscissa and has a high negative slope, will ensure that the concentration of the dilute phase (e.g., the cytosol) is as close to zero as possible. This would result in there being little captured mRNA in the cytosol of the cell that could undergo translation. Engineering even simple synIDPs like ELPs to achieve this goal, however, is non-

trivial, as no computational methods yet exist to predict their phase diagram from their sequence with sufficient accuracy. However, empirical methods wherein one synthesizes and screens many synIDPs could potentially identify an IDP with the desired binodal shape. Though this is feasible, it is experimentally intensive given the labor involved in recombinantly synthesizing a large enough set of synIDPs and measuring their binodal in the dilute regime. Engineering the material properties of the condensate is another method to suppress translation This may be possible by designing dynamically arrested condensates that impede the diffusion of tRNA and accessory proteins involved in translation into the condensate.

RNA-seq of purified synthetic Pum2-ELP-sfGFP RNPGs also presents several interesting findings that shed light on the extent of translation repression and its mechanism. First, it revealed that the synthetic RNPG could orthogonally recruit significant levels of target mRNA, and this sequestration of the target mRNA in the condensate explains the reduction in the overall expression of the target gene in the cells. Second, the presence of the ribosome in the condensate fraction is also consistent with the fact that transcription and translation are coupled in *E. coli*, so that as the target mRNA with the PRS is transcribed, it binds to the Pum2-ELP-sfGFP fusion in the cytosol, and some fraction of this bound mRNA also successfully loads onto ribosomes prior to being sequestered in condensates.

Third, RNA-seq showed that the level of *E. coli* mRNA that have a PRS in their sequence was negligible, suggesting that Pum2, despite its short 8-nucleotide long PRS, imparted a high level of specificity in *E. coli* under the experimental conditions of this study. The only exception was the high relative abundance of Pum2-ELP-sfGFP mRNA in the isolated condensate, which may sample may be explained by the presence of a few sequences in its mRNA that are similar to the canonical PRS: 5'-TGTACAGA-3', which is present in the Pum2 gene and differs from the PRS by only one nucleotide, 5'-TGTACAGG-3', which is present in the sfGFP gene and differs from the PRS by two nucleotides, and 5'-TGTATACA-3', which is present in the sfGFP gene and differs from the PRS by one nucleotide. The Pum2-ELP-sfGFP protein may bind these sequences on its own mRNA and sequester it into the synthetic condensate, abetted by the fact that Pum2-ELP-sfGFP is at a high concentration, as it is expressed from a strong promoter on a high copy number plasmid. In our studies on the inhibition of mCherry and Protein E expression by synthetic RNPGs, however, we saw no evidence of inhibition of Pum2-ELP-sfGFP expression over time, which would have been a result of mRNA sequestration in synthetic RNPGs.

The issue of specificity is critical to build synthetic RNPGs that do not impact other cellular function and is especially important as we seek to translate this system to eukaryotic cells, where the number of putative PRS are significantly greater and where Pum2 natively binds a plethora of mRNA[72]. Future work on building synthetic RNPGs in eukaryotic cell using synIDPs as drivers of phase separation could use RNA-binding proteins orthogonal to eukaryotic cells such as the MS2 phase Major Capsid Protein (MCP) or Mu phage Com protein.

Creating an "affinity window", wherein the affinity of a target RNA for the RNA-binding synIDP is much higher than any other native RNA, provides one such solution to this problem. One approach to do so could exploit multivalency to drive specificity, wherein oligomers of Pum2 bind multimers of PRS in the target mRNA. Combining high avidity binding partners and synIDPs with a low C_{sat} may further increase specificity as the highest affinity mRNA —the target mRNA—

will be captured and sequestered into condensates under conditions in the cell where the concentration of the capture agent, the Pum2 oligomer-ELP fusion, will be much lower than the K_D of other RNA in the cell that only contain a single PRS. An alternative approach to generate greater specificity may be possible by incorporating two different RNA motifs in a target RNA which bind orthogonal RNA binding proteins such as Pum2 and MCP[73-75] or Com[63] to also further increase specificity by virtue of the fact that native RNA transcripts are unlikely to have both recognition sites in their sequence.

In conclusion, we believe that synthetic RNPGs provide a new capability in synthetic biology for regulation of translation in *E. Coli.* Fusion of synIDPs to other functional proteins may extend this capability beyond regulation of translation and is the subject of ongoing studies.

Methods

Plasmid constructs: Gene sequences for proteins used in this study were purchased from Integrated DNA Technologies (IDT) and cloned into pET24+ vector that was modified in-house to contain AcuI, BseRI, and BglI cut sites that facilitate the recursive directional ligation by plasmid reconstruction method developed preciously[55]. To construct the highly repetitive ELP sequence, we used the recursive directional ligation by plasmid reconstruction method developed previously Briefly, the vector containing DNA encoding the 20-mer repeat of the ELP was digested with AcuI and BglI (A cut) and BseRI and BglI (B cut) and gel purified using a QIAGEN gel purification kit. The two cuts were ligated to generate oligomers of the 20-mer ELP sequence. This step was repeated multiple times to achieve the desired length of the ELP gene. The same procedure was used for the insertion of the leader and trailer motifs. Plasmids were transformed into the *Escherichia coli* strain BL21(DE3) for recombinant expression.

For experiments with dual protein expression, Pumilio2 genes (Pum2-6xHis, Pum2-ELP, Pum2-ELP-sfGFP) were either expressed from the original pET24+ vector or inserted into a pBAD 33.1 vector under control of an arabinose-inducible promoter as follows: Pum2-ELP genes were isolated from the custom pET24+ vector by digesting them with HindIII and XbaI restriction endonucleases followed by gel purification. The Pum2 genes were then ligated into a similarly cut pBAD 33.1 vector for protein expression induced with arabinose. mCherry and Protein E constructs were inserted into the pET24+ vector under control of an IPTG-inducible T7 promoter or into the pBAD 33.1 vector under control of an arabinose-inducible promoter using Gibson assembly. Cotransformation of plasmids (~1 ng final concentration of each plasmid) was performed on kanamycin/chloramphenicol plates.

Protein expression and purification: BL21(DE3) cells were grown overnight at 37°C in 8 mL of terrific broth (TB) and subsequently inoculated into 1 L of TB. Cells were induced at an OD_{600} of 0.8 with 1 mM isopropyl-β-D-thiogalactoside (IPTG) and were incubated overnight at 16 °C on a shaker incubator.

Cells were harvested by centrifugation (3000 g), resuspended in 20 ml PBS and lysed by sonication for 3 minutes (10 s ON and 40 s OFF, 85% amplitude) on ice (Misonix; Farmingdale, NY). The insoluble fraction of the cell lysates was removed by centrifugation (15000 rpm, 20 min, 4°C) and the supernatant, containing the protein of interest, was retained and kept at 4°C. Nucleic acid

impurities were removed by addition of 10% polyethyleneimine (2 ml per 1 L culture) (MP Biomedicals, Santa Ana, CA) followed by centrifugation (15000 rpm, 20 min, 4°C). Protein impurities were removed by inverse transition cycling (ITC), described elsewhere[76].The purified protein was dialyzed against 1x PBS, 1 mM DTT overnight and concentrated by centrifugal ultrafiltration (Amicon, with a 10 kDa filter).

Purification of Pum2-(His)$_8$: 3 mL of immobilized metal affinity chromatography (IMAC) resin (HisPur TM Ni-NTA Resin (Product No. 88222, Thermo Fisher) was placed in a gravity flow column (3 ml) and was washed with 2–3 column volumes of PBS and PBS containing 10 mM, 20 mM, and 30 mM of imidazole at pH 7.4. Supernatant from cells expressing Pum2-(His)$_8$ was added to the column and allowed to incubate at 4°C for 1 h. The protein was eluted with 1-2 column volumes of PBS containing 300 mM of imidazole at pH 7.4. The eluted protein was dialyzed against 1x PBS, 1 mM DTT twice overnight to remove the imidazole and was concentrated by centrifugal ultrafiltration. (Amicon, with a 10 kDa filter). Protein purity was characterized by 4%–20% gradient tris-HCl (Biorad, Hercules, CA) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and gels were stained with 0.1% w/v Coomassie Blue (Fisher Scientific, Hampton, NH).

Temperature-controlled ultraviolet–visible spectrophotometry: Cloud point temperatures (T_{cp}) were determined by temperature-controlled spectrophotometry on a Cary 300 UV-vis spectrophotometer (Agilent Technologies). Samples containing various concentrations of protein in 1x PBS, 1 mM DTT were heated at 1°C min^{-1} while the absorbance was recorded at 350 nm. The absorbance was normalized to the absorbance at the lowest temperature point in data acquisition. The T_{cp} was determined as the temperature at which the first derivative of the absorbance as a function of temperature was the maximum.

Electrophoretic mobility shift assay (EMSA). RNA 1 (5'-/56FAM/CCAGAAUUGUACAUAU UCG-3') and RNA 2 (5'-/56FAM/CCAGAAUUUGAUAUAUUCG-3'), shown in table 1, were purchased from IDT and used as received. RNA 1 and RNA 2 (19 bp Pum2 target and non-specific RNA, respectively) at 0.2 μM concentration were incubated with increasing concentrations (0.2 μM to 2 μM) of Pum2-ELP protein in PBS supplemented with 1 mM DTT on ice for 10 min. Samples were analyzed on Mini-PROTEAN TBE Precast gels (BioRad). The sample loading buffer was composed of 45 mM Tris base, 45 mM boric acid, 1 mM EDTA and 36% glycerol. Gels were run at 90 V for 25 min in 1x TBE buffer and imaged on a BIORAD gel-doc imager.

Surface Plasmon Resonance (SPR) spectroscopy: A Biacore T200 SPR instrument (GE Healthcare) was used to determine the binding affinity of ELP-Pum2 to the following RNA at 25 °C: Pum2-binding 19 bp RNA (RNA 1) non-binding 19bp RNA (RNA 2), mCherry (RNA 3), sfGFP (RNA 4), and sfGFP-PBD (RNA 7) mRNA. A CM5 sensor chip (GE healthcare) was normalized using 70% glycerol. ELP-Pum2 proteins (VWR) were immobilized using EDC/NHS coupling in flow path 2 of the chip, while no proteins were immobilized in flow path 1 as a control. A solution of 0.5 M EDC/1M NHS was injected at 5 μl/min for 7 min to activate the chip surface of flow paths 1 and 2. In flow path 2, 100 μg/ml of ELP-Pum2 in 10 mM sodium acetate buffer (pH 4.5) was injected at 5 μl/min for 1 min to obtain a final RU (response unit) of around 1,000. Next, in flow paths 1 and 2, 1 M ethanolamine was injected at 5 μl/min for 10 min to inactivate the chip surface. To characterize the binding affinity of ELP-Pum2 to different RNA targets, serial

dilutions of mRNA were injected in both flow paths at a flow rate of 30 µl.min^{-1} for 3 min. Then, the RNA solutions were replaced with a running buffer (1x PBS) for 10 min for to measure the dissociation. After subtraction of the signal of flow path 1 from that of flow path 2, the final sensorgrams were obtained and analyzed using a 1:1 Langmuir binding model in BIAevaluation software (GE healthcare).

mRNA synthesis and purification: To synthesize mRNA, DNA templates were amplified by PCR from a pET24+ vector (30 cycles, 10 s at 98 °C, annealing for 10 s at 62 °C: polymerization for 15 s at 72 ° using GoTaq Green Mix, 2X (Promega) and T7 promoter forward and reverse primers). PCR products were transcribed *in vitro* using the HiScribe T7 Quick High Yield RNA Synthesis Kit (New England Biolabs) according to the manufacturer's protocol. Synthesized RNA was then purified in two steps with Monarch RNA cleanup kit (10 µgr, NEB Inc) and NucAway™ spin columns (Invitrogen, Thermo Fisher Scientific). For fluorescently labeled RNAs, fluorescein-12-UTP (Fluorescein RNA labelling Mix, Sigma-Aldrich) or Cy3-UTP (Jena Biosciences) was added to the transcription reaction (1 mM final concentration).

Fabrication of water-in-oil droplets: Chip microfluidics described elsewhere[56] were used to create water-in-oil emulsion droplets. A dispersed aqueous phase containing protein/plasmids/RNA of interest (or combinations thereof) in NEB Express® Cell-free *E. coli* protein synthesis mixture and an organic continuous phase comprised of 75/5/20% vol/vol TEGOSOFT DEC, ABIL EM 90 and mineral oil were injected into a microfluidic droplet generator at constant flow rates (300 µl h^{-1} for the organic continuous phase and 200 µl h^{-1} for the aqueous phase) using precision syringe pumps. The whole apparatus was kept cold during the production of droplets and the process was monitored with a 5x objective on an inverted microscope (Leica) equipped with a digital microscopy camera (Lumenera Infinity 3-1 charge-coupled device).

Cell-free expression: The NEB Express® Cell-free *E. coli* protein synthesis system was used in this study. Solutions were prepared following the manufacturer's protocol with addition of plasmids/RNA/protein of interest (or their combinations) to reach the desired final concentrations as indicated.

Fluorescence imaging of protocells and E. Coli: Water-in-oil droplets were collected on a concavity glass microscope slide and were imaged on an upright Zeiss Axio Imager D2 microscope with a 20x objective, and the appropriate filter set for GFP (excitation laser 485 nm, emission filter 518 nm), and mCherry (excitation laser 586 nm, emission filter 614 nm). Samples were kept at the desired temperature during imaging using a precision Peltier heating and cooling stage (Linkam LTS120) equipped with a Linkam PE95 digital temperature control unit. The fluorescence was quantified by analyzing images in ImageJ. The protocells (n=10) were manually selected, and the integrated density —defined as the sum of intensities of pixels across the image— was obtained for each channel. As the mean fluorescence intensity assigns brighter values for dim pixels and vice-versa, integrated density was used for true representation of the brightness of each pixel in the protocell. Statistical analysis was performed by two-way ANOVA using Tukey's test in GraphPad Prism 9.

Confocal fluorescence microscopy of protocells: Protocells were imaged on a Andor Dragonfly Spinning Disk confocal microscope equipped with a temperature control chamber using a 20x

objective. Fluorescence signal from sfGFP was acquired using 488 nm and 525/50 nm bandpass filter while mCherry fluorescence was captured using a 561 nm laser with a 600/50 nm bandpass filter at 37 °C. Z-stack images were processed using Imaris software. To analyze the data, the mean fluorescence intensity was obtained along each slice along the Z-axis using NIH's ImageJ. The fluorescence intensity was plotted against Z-distance and the intensity curve for mCherry (red) was normalized with the respective sfGFP (green) fluorescence. The area under curve (AUC) metric was computed from this plot using GraphPad Prism 9 and the ratio of red:green AUC is reported (see Figure 3D). Statistical analysis was performed by Student's t-test in GraphPad Prism 9.

mCherry Expression in E. Coli: E. coli BL21(DE3) cells containing plasmid pPEG (pET24+ derivative encoding Pum2-ELP-sfGFP under T7 promoter) and either plasmid pM5P (pBAD33.1 derivative with mCherry-5xPRS under pBAD promoter) or plasmid pM (pBAD33.1 derivative with mCherry alone under pBAD promoter) were inoculated into 2 mL MOPS EZ Rich Defined Medium (Teknova) supplemented with 40 mg/mL kanamycin and 25 mg/mL chloramphenicol and grown until steady state (about ~16 h) at 37°C in a shaking incubator. 50 µL of cells were added to 2 mL of fresh MOPS EZ Rich Defined Medium supplemented with 40 mg/mL kanamycin and 25 mg/mL chloramphenicol and grown at 37°C in a shaking incubator until the culture reached an OD_{600} of ~0.6 (~3 h). 1 mL of the culture was removed in and Pum2-ELP-sfGFP and the mCherry RNA were induced by adding 0.2 mM IPTG and 0.2% arabinose, respectively. 180 µL of the induced cells were then aliquoted into a sterile Costar 96-well black-walled plate (Corning) with three replicates for each sample and covered with a few drops of mineral oil to prevent evaporation in theplate reader. The mCherry and sfGFP fluorescence and OD_{600} of the sample over time was measured in a Tecan Infinite 200 PRO plate reader (Tecan) for 22 h at 37°C, with measurements every 5 min.

The time-course fluorescence data from the microplate reader was analyzed using a custom R script. Briefly, blank MOPS EZ Rich Defined Medium or 2xYT fluorescence and OD_{600} values were subtracted from all measurements. The mCherry and sfGFP fluorescence values were divided by the OD_{600} values at each time point to determine the OD_{600}-normalized mCherry and sfGFP fluorescence at each time point.

Protein E Expression: E. coli BL21(DE3)+pPEG+pE4P (pBAD33.1 derivative with Protein E-4xPRS under pBAD control) and *E. coli* BL21(DE3)+pPEG+pE (pBAD33.1 derivative with Protein E under pBAD control) were inoculated into 2 mL 2xYT media supplemented with 40 mg/mL kanamycin, 25 mg/mL chloramphenicol, and 1% glucose and grown for 16 h at 37°C in a shaking incubator. 50 µL of the cells were added to 2 mL of fresh 2xYT supplemented with 40 mg/mL kanamycin, 25 mg/mL chloramphenicol, and 1% glucose and grown at 37°C in a shaking incubator until the culture reached an OD_{600} of ~0.6 (~3 h). 1 mL of the culture was removed and Pum2-ELP-sfGFP and the Protein E RNA were induced by adding 0.2 mM IPTG and 0.2% arabinose, respectively. 180 µL of the induced cells were then aliquoted into a sterile Costar 96-well black-walled plate (Corning) with three replicates for each sample and covered with a few drops of mineral oil to prevent evaporation in a plate reader. The OD_{600} of the sample over time was measured in a Tecan Infinite 200 PRO plate reader (Tecan) for 22 h at 37°C, with measurements every 5 min.

Condensate purification: E. coli BL21(DE3) cells containing plasmid pPEG and either plasmid pM5P or plasmid pM were inoculated from frozen stocks into 2 mL of 2xYT media supplemented with 40mg/mL kanamycin and 25 mg/mL and grown in a shaking incubator steady state (~16 hours) at 37°C. 1 mL of *E. coli* BL21(DE3)+pPEG+pM5P or *E. coli* BL21(DE3)+pPEG+pM was added to 50 mL of 2xYT media supplemented with 40 mg/mL kanamycin and 25 mg/mL chloramphenicol and grown in a shaking incubator until the culture reached an OD_{600} of ~0.6 (~3 h). Expression of Pum2-ELP-sfGFP and RNA encoding mCherry-5xPRS or mCherry alone was induced with the addition of 1 mM IPTG and 0.2% arabinose (final concentration). The cultures were then grown at 37°C for 4 h to allow for protein and RNA production. To lyse the cells and release the condensates, 35 mL of the culture was sonicated with a Qsonica Q500 sonicator (Qsonica) on ice at an amplitude of 85%, with cycles of 10s on 40s off to a total time of 2 min, after which the sample was immediately removed and placed in a water bath at 37°C to ensure condensates stayed phase separated. Due to the heat generated by sonication, the temperature of the solution during sonication did not drop below 37°C even while being in an ice bath for the duration of the procedure. 4 mL of each sample was then layered on top of an ultracentrifuge tube pre-layered with the following gradient from bottom to top: (1) 3 mL of 60% Optiprep Density Gradient medium (Sigma Aldrich, shortened to Optiprep) 1 mM $MgCl_2$, 2.5 mM KCl, 0.00025% Phenol Red, (2) 4 mL of 40% Optiprep in 1x Phosphate Buffered Saline (PBS) from 10x PBS (Sigma Aldrich) 1 mM $MgCl_2$, 2.5 mM KCl, (3) 3 mL of 25% Optiprep in 1xPBS from 10x PBS 1 mM $MgCl_2$ 2.5 mM KCl 0.001% phenol red, (4) 4mL of 17% Optiprep in 1xPBS 1mM $MgCl_2$ 2.5mM KCl 1M NaCl. Samples were then centrifuged at 30,000 rpm at 37°C for 18 h in a Sorvall WX+ Ultracentrifuge (Thermo Fisher). Four 1.5mL and seven fractions were collected from the top of the tube and analyzed on an SDS-PAGE gel with procedures described above. Fractions containing Pum2-ELP-sfGFP were then used for RNA isolation.

RNA extraction from Pum2-ELP-sfGFP condensates: RNA was extracted from fractions containing Pum2-ELP-sfGFP using the TRIzol LS reagent (Thermo Fisher) reagent following manufacturer's protocol. In short, 250 μL of the fraction (Pum2-ELP-sfGFP in Optiprep solution) was mixed with 750 μL of TRIzol LS reagent and allowed to incubate for 5 min at room temperature. 200 μL of chloroform was added to the solution and the sample was allowed to incubate for 2 min at room temperature. The samples were then centrifuged at 12,000 x g in a tabletop Eppendorf centrifuge for 15 min at 4ºC. 350 μL of the resulting RNA-containing aqueous phase used for RNA purification. RNA was purified using the PureLink RNA Mini Kit (Thermo Fisher) following the manufacturer's protocol. RNA concentration was measured using a Nanodrop spectrophotometer (Thermo Fisher).

Condensate cDNA library construction: Condensate cDNA libraries were constructed using the KAPA RNA Hyperprep Kit (Roche Sequencing) without depletion of ribosomal RNA according to manufacturer's instructions. Adapters were taken from the KAPA Unique Dual-Indexed Adapter Kit (Roche Sequencing). For condensates from *E. coli* BL21(DE3)+pPEG+pM5P adapters UDI 7, 11, and 14 were used, and for condensates from *E. coli* BL21(DE3)+pPEG+pM adapters UDI 18, 19, and 20 were used. Adapters were first diluted to 1.5 μM using KAPA Adapter Dilution Buffer (Roche Sequencing). In short, the procedure was as follows. Master mixes were created for all steps: 1st strand synthesis MM (1st Strand Synthesis buffer + KAPA Script), 2nd strand synthesis and A-tailing MM (2nd Strand Marking Buffer + 2nd Strand Syntheiss and A-tailing enzyme mix), Adapter ligation MM (Ligation Buffer + DNA Ligase), and Library amplification

MM (KAPA HiFi HotStart ReadyMix + Library Amplification Primer Mix). 10µL of sample RNA was mixed with 10µL of Fragmentation, Priming, and Elution buffer and incubated at 94°C for 6 minutes to create library inserts of 200-300bp. 20µL of fragmented RNA was mixed with 10µL of 1st Strand Synthesis MM and incubated with the following scheme: 25°C for 10 minutes for primer extension, 42°C for 15 minutes for 1st strand synthesis, 70°C for 15 minutes to inactivate enzymes. 30µL of the 1st strand synthesis produce was mixed with 30µL of 2nd Strand Synthesis and A-tailing MM and incubated with the following scheme: 16°C for 30 minutes for 2nd strand synthesis, 62°C for 10 minutes for A-tailing. 60µL of the 2nd strand synthesis produce was mixed with 45µL of the Adapter Ligation MM and 5µL of 1.5µM KAPA Adapter and incubated at room temperature for 15 minutes. Two magnetic bead-based library cleanup steps were conducted with KAPA Pure Beads as per manufacturer's protocol. 20µL of the adapter-ligated, bead-purified DNA was mixed with 30µL of the Library amplification master mix and amplified for 10 PCR cycles with the following scheme: 98°C for 45 seconds for initial denaturation, then 10 cycles of 98°C for 15s for denaturation, 60°C at 30s for annealing, 72°C at 30s for extension, and finally 72°C for 1 min for final extension. Amplified libraries were again cleaned up using the KAPA Pure Beads as per manufacturer's protocol.

Condensate RNA sequencing: Condensate cDNA libraries were first analyzed and quantified using an Agilent TapeStation 4200 with D1000 HS tape (Agilent) to check for cDNA size consistency and a Qubit 4 fluorometer with the dsDNA HS assay (Thermo Fisher) to determine dsDNA concentration as per manufacturer's protocol. Samples were then individually diluted to 2 nM in RSB buffer (Illumina) from the Qubit-measured concentration assuming a 260 bp length as determined by the TapeStation results. 5 µL of each sample was pooled into one master 2 nM library. The master 2nM library was then diluted to 650pM in RSB buffer to a final volume of 22.5 µL, to which 2.5 µL of 1nM PhiX phage DNA was added for library diversity. The entire sample was loaded onto an Illumina NextSeq 2000 using a P2 100 cycle kit cartridge and sequenced using 2x50 bp paired-end sequencing**.**

Condensate RNA sequencing analysis: FASTQ files in a .gz format were downloaded from Illumina Basespace and the reads were mapped to the DNA present in *E. coli* BL21(DE3)+pPEG+pM5P for differential expression analysis using the Rockhopper software package (https://cs.wellesley.edu/~btjaden/Rockhopper/)[58]. For differential expression analysis, sequencing reads from condensates purified from *E. coli* BL21(DE3)+pPEG+pM5P and *E. coli* BL21(DE3)+pPEG+pM were both mapped to the *E. coli* BL21(DE3)+pPEG+pM5P genome as two "experiments" (with or without the 5xPRS) with three replicates each. Both experiments were mapped to the same genome as the only difference between the genomes is the addition of 5xPBD sites on pM5P and all DNA sequences present in *E. coli* BL21(DE3)+pPEG+pM would also be present in *E. coli* BL21(DE3)+pPEG+pM5P. For analysis with Rockhopper, the pPEG plasmid and pM5P plasmid sequences were concatenated into a singular 14kbp "replicon", such that the Rockhopper software aligned the condensate RNA sequencing reads to the *E. coli* BL21(DE3) genome and the pPEG+pM5P replicon plasmid sequences as two "replicons". The parameters used were as follows: allowed mismatches: 0.15. Minimum seed length: 0.33. Minimum expression of UTRs and ncRNAs: 0.5. Minimum reads mapping to a transcript: 20. Minimum transcript length: 50. Minimum count to seed a transcript: 50. Minimum count to extend a transcript: 5. Paired-end reads were used for each replicate. Detected RNA transcripts, reported from transcript start base pair location to finish base pair location by Rockhopper, were automatically assigned to pre-

labeled genes on the *E. coli* BL21 genome by the Rockhopper software and were manually assigned to genes on the pPEG+pM5P plasmid by comparing the transcript start and finish locations to the labeled pPEG+pM5P plasmid. To clean up the reported RNA-seq counts data, duplicate detected transcripts were deleted. Then, a single Reads Per Kilobase Million (RPKM) value for each transcript, taking into account all three replicates, was calculated by Rockhopper. From the determined RPKM values a Transcripts per Million (TPM) value for each gene was determined using the formula TPM = 10^6 * (transcript's individual RPKM)/(Sum(all RPKM values).

Identification of native PRS sites in the E. coli *BL21(DE3) genome:* Each of the four potential PRS sequences (5'-$U_1G_2U_3A_4$**(A/U/G/C)**$_5A_6U_7A_8$-3') was found in an annotated *E. coli* BL21(DE3) genome downloaded from National Center for Biotechnology Information (NCBI) using the Snapgene Viewer v.6.0.2 software. If the PRS was found in an annotated gene on the transcribed coding strand in the correct direction, the PRS was considered to be present in that gene. In addition, if the PRS was found at or less than 100 bp up- or downstream of an annotated gene (presumably in RNA that would be transcribed) the PRS was also considered to be present in that gene.

References

1 Strzelecka, M. *et al.* Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nature Structural & Molecular Biology* **17**, 403-409 (2010).

2 Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* **18**, 285-298 (2017).

3 Henninger, J. E. *et al.* RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell* **184**, 207-225.e224 (2021).

4 Hamada, T. *et al.* Stress granule formation is induced by a threshold temperature rather than a temperature difference in Arabidopsis. *Journal of Cell Science* **131**, jcs216051 (2018).

5 Buchan, J. R., Muhlrad, D. & Parker, R. P bodies promote stress granule assembly in Saccharomyces cerevisiae. *Journal of Cell Biology* **183**, 441-455 (2008).

6 Corbet, G. A. & Parker, R. RNP Granule Formation: Lessons from P-Bodies and Stress Granules. *Cold Spring Harb Symp Quant Biol*, 040329 (2020).

7 Guillén-Boixet, J. *et al.* RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell* **181**, 346-361.e317 (2020).

8 Treeck, B. V. *et al.* RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *PNAS* **115**, 2734-2739 (2018).

9 Buchan, J. R. mRNP granules. *RNA Biol* **11**, 1019-1030 (2014).

10 Murthy, A. C. *et al.* Molecular interactions underlying liquid−liquid phase separation of the FUS low-complexity domain. *Nature Structural & Molecular Biology* **26**, 637-648 (2019).

11 Shapiro, D. M., Ney, M., Eghtesadi, S. A. & Chilkoti, A. Protein Phase Separation Arising from Intrinsic Disorder: First-Principles to Bespoke Applications. *J. Phys. Chem. B* (2021).

12 Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).

13 Berry, J., Weber, S. C., Vaidya, N., Haataja, M. & Brangwynne, C. P. RNA transcription modulates phase transition-driven nuclear body assembly. *PNAS* **112**, E5237 (2015).

14 Rey, T. *et al.* Mitochondrial RNA granules are fluid condensates positioned by membrane dynamics. *Nature Cell Biology* **22**, 1180-1186 (2020).

15 Brangwynne, C. P. *et al.* Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **324**, 1729-1732 (2009).

16 Handwerger, K. E., Cordero, J. A. & Gall, J. G. Cajal Bodies, Nucleoli, and Speckles in the Xenopus Oocyte Nucleus Have a Low-Density, Sponge-like Structure. *Molecular Biology of the Cell* **16**, 202-211 (2005).

17 Jakob, U., Kriwacki, R. & Uversky, V. N. Conditionally and Transiently Disordered Proteins: Awakening Cryptic Disorder To Regulate Protein Function. *Chem. Rev.* **114**, 6779-6805 (2014).

18 Uversky, V. N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Current Opinion in Structural Biology* **44**, 18-30 (2017).

19 Vernon, R. M. & Forman-Kay, J. D. First-generation predictors of biological protein phase separation. *Current Opinion in Structural Biology* **58**, 88-96 (2019).

20 Babinchak, W. M. *et al.* The role of liquid–liquid phase separation in aggregation of the TDP-43 low-complexity domain. *J. Biol. Chem.* **294**, 6306-6317 (2019).
21 Kato, M. *et al.* Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **149**, 753-767 (2012).
22 Martin, E. W. & Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **57**, 2478-2487 (2018).
23 Molliex, A. *et al.* Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123-133 (2015).
24 Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Molecular Cell* **60**, 231-241 (2015).
25 Conicella, Alexander E., Zerze, Gül H., Mittal, J. & Fawzi, Nicolas L. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537-1549 (2016).
26 Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066-1077 (2015).
27 Alshareedah, I. *et al.* Interplay between Short-Range Attraction and Long-Range Repulsion Controls Reentrant Liquid Condensation of Ribonucleoprotein–RNA Complexes. *J. Am. Chem. Soc.* **141**, 14593-14602 (2019).
28 Garcia-Jove Navarro, M. *et al.* RNA is a critical element for the sizing and the composition of phase-separated RNA–protein condensates. *Nature Communications* **10**, 3230 (2019).
29 Hennig, S. *et al.* Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol* **210**, 529-539 (2015).
30 Sanchez-Burgos, I., Espinosa, J. R., Joseph, J. A. & Collepardo-Guevara, R. RNA length has a non-trivial effect in the stability of biomolecular condensates formed by RNA-binding proteins. *PLOS Computational Biology* **18**, e1009810 (2022).
31 Bauer, K. E. *et al.* RNA supply drives physiological granule assembly in neurons. *Nature Communications* **13** (2022).
32 Potvin-Trottier, L., Lord, N. D., Vinnicombe, G. & Paulsson, J. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**, 514-517 (2016).
33 Panteli, J. T. & Forbes, N. S. Engineered bacteria detect spatial profiles in glucose concentration within solid tumor cell masses. *Biotechnol. Bioeng.* **113**, 2474-2484 (2016).
34 Zalatan, Jesse G. *et al.* Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* **160**, 339-350 (2015).
35 Dong, C., Fontana, J., Patel, A., Carothers, J. M. & Zalatan, J. G. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nature Communications* **9**, 2489 (2018).
36 Xie, M. & Fussenegger, M. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nature Reviews Molecular Cell Biology* **19**, 507-525 (2018).
37 Roden, C. & Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nature Reviews Molecular Cell Biology*, 1-13 (2020).
38 Shin, Y. *et al.* Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171.e114 (2017).
39 Lin, Y., Protter, David S. W., Rosen, Michael K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell* **60**, 208-219 (2015).

40 Elbaum-Garfinkle, S. *et al.* The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *PNAS* **112**, 7189-7194 (2015).
41 Adamala, K. P., Martin-Alarcon, D. A. & Boyden, E. S. Programmable RNA-binding protein composed of repeats of a single modular unit. *PNAS* **113**, E2579-E2588 (2016).
42 Lu, G. & Hall, T. M. T. Alternate modes of cognate RNA recognition by human PUMILIO proteins. *Structure (London, England : 1993)* **19**, 361-367 (2011).
43 Wang, X., Zamore, P. D. & Hall, T. M. T. Crystal Structure of a Pumilio Homology Domain. *Molecular Cell* **7**, 855-865 (2001).
44 Mozhdehi, D. *et al.* Genetically encoded lipid–polypeptide hybrid biomaterials that exhibit temperature-triggered hierarchical self-assembly. *Nature chemistry* **10**, 496-505 (2018).
45 Christensen, T., Hassouneh, W., Trabbic-Carlson, K. & Chilkoti, A. Predicting Transition Temperatures of Elastin-Like Polypeptide Fusion Proteins. *Biomacromolecules* **14**, 1514-1519 (2013).
46 Li, N. K., Quiroz, F. G., Hall, C. K., Chilkoti, A. & Yingling, Y. G. Molecular Description of the LCST Behavior of an Elastin-Like Polypeptide. *Biomacromolecules* **15**, 3522-3530 (2014).
47 Roberts, S., Dzuricky, M. & Chilkoti, A. Elastin-like Polypeptides as Models of Intrinsically Disordered Proteins. *FEBS letters* **589**, 2477-2486 (2015).
48 Quiroz, F. G. *et al.* Intrinsically disordered proteins access a range of hysteretic phase separation behaviors. *Science Advances* **5**, eaax5177 (2019).
49 Hassouneh, W., Christensen, T. & Chilkoti, A. Elastin-like Polypeptides as a Purification Tag for Recombinant Proteins. *Curr Protoc Protein Sci* **CHAPTER**, Unit-6.11 (2010).
50 Altman, E., Young, K., Garrett, J., Altman, R. & Young, R. Subcellular localization of lethal lysis proteins of bacteriophages lambda and phiX174. *Journal of Virology* **53**, 1008-1011 (1985).
51 Witte, A., Wanner, G., Sulzner, M. & Lubitz, W. Dynamics of PhiX174 protein E-mediated lysis of Escherichia coli. *Archives of Microbiology* **157**, 381-388 (1992).
52 Wang, X., McLachlan, J., Zamore, P. D. & Hall, T. M. T. Modular Recognition of RNA by a Human Pumilio-Homology Domain. *Cell* **110**, 501-512 (2002).
53 Meyer, D. E. & Chilkoti, A. Quantification of the Effects of Chain Length and Concentration on the Thermal Behavior of Elastin-like Polypeptides. *Biomacromolecules* **5**, 846-851 (2004).
54 Cheong, C.-G. & Hall, T. M. T. Engineering RNA sequence specificity of Pumilio repeats. *PNAS* **103**, 13635-13639 (2006).
55 McDaniel, J. R., Mackay, J. A., Quiroz, F. G. & Chilkoti, A. Recursive Directional Ligation by Plasmid Reconstruction Allows Rapid and Seamless Cloning of Oligomeric Genes. *Biomacromolecules* **11**, 944-952 (2010).
56 Simon, J. R., Eghtesadi, S. A., Dzuricky, M., You, L. & Chilkoti, A. Engineered Ribonucleoprotein Granules Inhibit Translation in Protocells. *Molecular Cell* **75**, 66-75.e65 (2019).
57 Buckley, K. J. & Hayashi, M. Lytic activity localized to membrane-spanning region of ϕX174 E protein. *Molecular and General Genetics MGG* **204**, 120-125 (1986).
58 McClure, R. *et al.* Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Research* **41**, e140-e140 (2013).

59 Zhao, S., Ye, Z. & Stanton, R. Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *RNA* **26**, 903-909 (2020).
60 Peano, C. *et al.* An efficient rRNA removal method for RNA sequencing in GC-rich bacteria. *Microbial Informatics and Experimentation* **3**, 1 (2013).
61 Ishihama, Y. *et al.* Protein abundance profiling of the Escherichia coli cytosol. *BMC Genomics* **9**, 102 (2008).
62 Auweter, S. D. *et al.* Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *Embo j* **25**, 163-173 (2006).
63 Nowacka, M. *et al.* Specific interaction of zinc finger protein Com with RNA and the crystal structure of a self-complementary RNA duplex recognized by Com. *PLoS One* **14**, e0214481 (2019).
64 Perčulija, V. & Ouyang, S. in *Helicases from All Domains of Life* (ed Renu Tuteja) 141-171 (Academic Press, 2019).
65 Manna, S. An overview of pentatricopeptide repeat proteins and their applications. *Biochimie* **113**, 93-99 (2015).
66 Zhang, C. & Muench, D. G. A Nucleolar PUF RNA-binding Protein with Specificity for a Unique RNA Sequence. *J. Biol. Chem.* **290**, 30108-30118 (2015).
67 Wickens, M., Bernstein, D. S., Kimble, J. & Parker, R. A PUF family portrait: 3′UTR regulation as a way of life. *Trends in Genetics* **18**, 150-157 (2002).
68 Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141 (2010).
69 Goldstrohm, A. C., Hall, T. M. T. & McKenney, K. M. Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. *Trends in Genetics* **34**, 972-990 (2018).
70 Reinkemeier, C. D., Girona, G. E. & Lemke, E. A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. *Science* **363**, eaaw2644 (2019).
71 Reinkemeier, C. D. & Lemke, E. A. Dual film-like organelles enable spatial separation of orthogonal eukaryotic translation. *Cell* **184**, 4886-4903.e4821 (2021).
72 Bohn, J. A. *et al.* Identification of diverse target RNAs that are functionally regulated by human Pumilio proteins. *Nucleic Acids Research* **46**, 362-386 (2018).
73 Ni, C.-Z. *et al.* Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure* **3**, 255-263 (1995).
74 van den Worm, S. H. E. *et al.* Crystal structures of MS2 coat protein mutants in complex with wild-type RNA operator fragments. *Nucleic Acids Research* **26**, 1345-1351 (1998).
75 Yoon, J.-H., Srikantan, S. & Gorospe, M. MS2-TRAP (MS2-tagged RNA affinity purification): Tagging RNA to identify associated miRNAs. *Methods* **58**, 81-87 (2012).
76 Meyer, D. E. & Chilkoti, A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature Biotechnology* **17**, 1112-1115 (1999).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SyntheticRNPGChilkotiSI.pdf
- Supportingvideo1.mp4
- Supportingvideo2.mp4
- Supportingvideo32.mp4
- Supportingvideo4.mp4
- Supportingvideo5.mp4