Formation of synthetic RNP granules using engineered phage-coat-protein - RNA complexes

Authors: Naor Granik[1], Noa Katz[2], Sarah Goldberg[2], and Roee Amit[2,3*]

Affiliations:

[1] Department of Applied Mathematics, Technion - Israel Institute of Technology, Haifa 32000, Israel.

[2] Department of Biotechnology and Food Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel.

[3] The Russell Berrie Nanotechnology Institute, Technion - Israel Institute of Technology, Haifa 32000, Israel.

***** To whom correspondence should be addressed: roeeamit@technion.ac.il.

Introductory paragraph: synthetic long non-coding RNA molecules encoded with multiple phage coat protein binding sites form phase separated RNP granules when co-expressed with their cognate binding proteins.

Abstract: Observations of RNA-protein (RNP) granules within living cells have generated important insights on intracellular processes in recent years. Here, we demonstrate that PP7-coat protein (PCP) and Qβ-coat protein (QCP) co-expressed with synthetic long non-coding RNA (slncRNA) encoded with multiple CP binding sites self-assemble into phase-separated synthetic RNP (SRNP) granules within *E. coli* cells and *in vitro*. SNRP granules exhibit hallmarks of RNP granule formation including phase separation from a semi-dilute liquid *in vitro*, and retention of structure after cell lysis. *In vivo* real-time tracking of SRNP granules reveals that they are found in cell poles of exponentially growing cells, and that their structure exhibits gel-like properties by the slow stochastic insertion and shedding of CP-slncRNA complexes. We propose a transcriptional model that takes into account the cytosolic structure of E. *coli* (i.e., the cytosol is divided into a biomolecularly dense domain in the nucleoid-dominated region and a dilute liquid domain in the polar regions), and use it to provide a theoretical framework for explaining the increase in intracellular CP-titer as a function slncRNA structure, and the appearance of SRNP granules in non-polar regions of stationary phase cells. Consequently, the SRNP-granule data and resultant model has implications for various bacterial transcriptional and translational processes, and provides an alternative explanation for the super-Poisson kinetics attributed to transcriptional bursts.

Introduction

Liquid – liquid phase separation (LLPS), the process by which a homogeneous solution separates into molecularly dense and dilute liquid phases, has been connected to a wide range of natural cellular processes in virtually all forms of life (Banani et al., 2017; Boeynaems et al., 2018; Hyman et al., 2014; Langdon and Gladfelter, 2018; Polymenidou, 2018). In cells, LLPS results in the formation of membrane-

less compartments containing a high-concentration mix of biomolecules (e.g., proteins, RNA, etc.), which is surrounded by a low-concentration solution. An important sub-class of this phenomenon, which has garnered increased attention from the scientific community in recent years, are the ribonucleoprotein (RNP) granules which are ubiquitous in many cell types. RNP granules are typically formed from long noncoding RNAs (lncRNAs) that are bound by various RNA binding proteins (RBPs) (Bond and Fox, 2009; Clemson et al., 2009; Polymenidou, 2018). These nucleoprotein complexes are evolutionary conserved and have been shown to play an important regulatory role. Examples of such compartments include paraspeckles, stress granules, P-bodies, and nuclear speckles among others (Banani et al., 2017; Mao et al., 2011). RNP granules display various elastic, viscosity, and other rheological properties which makes it difficult to classify them as strictly "liquid", "solid", "gel-like", or "glass-like" particles. The particular properties of each particle are thought to depend on many parameters including RNA structure and sequence, various internal protein domains. (e.g., intrinsically disordered domains (IDDs), protein-protein interaction motifs, etc.), pH, ion concentrations, temperature, cytosol content etc. Consequently, the dependence on this wide spectrum of parameters as well as the importance of these molecular condensates to various cellular functions have made these complexes an active area of modern research.

Given the ubiquity and importance of these compartments to cell function, we hypothesized that synthetic RNP-granules made from fully controlled and programmable components can be informative on the underlying relationship between various parameters and the resultant particle-nature or phase-behavior of this class of objects. Synthetic RNA-protein constructs have already been in use for nearly two decades. Perhaps the most ubiquitous system is composed of RNA molecules that are encoded with multiple or repeat hairpin motifs that can bind the phage coat proteins of PP7 or MS2. Using this system to label the 5' or 3' end of the transcript has become commonplace in the last two decades (Bertrand et al., 1998; Fusco et al., 2003; Golding et al., 2005; Jones and Elf, 2018; Larson et al., 2009; Tutucci et al., 2018), thus allowing one to visualize RNA transcripts when the coat proteins are co-expressed. This approach, originally introduced by Singer and others (Bertrand et al., 1998; Larson et al., 2009; Tutucci et al., 2018), was devised for the purpose of probing the dynamics of transcription and other RNA-related processes, irrespective of cell-type. When co-expressed, the coat-protein-bound RNA molecules yield bright puncta, which are similar in appearance to natural RNP granules. Consequently, we wondered whether the CP-puncta observed in various systems were, in fact, synthetic RNP granules, and thus their optical properties reflect granular dynamics rather than transcriptional ones.

In this paper, we examine this question by relying on our previous works (Katz et al., 2021, 2019, 2018) to design and synthesize a variety of CP binding synthetic long non-coding RNA (slncRNA) molecules, which form bright puncta when bound by either tdPP7 coat protein (PCP) or Qβ-coat protein (QCP). We show that *in vivo* CP-puncta only form in bacterial cell poles. The puncta survive cell lysis,

and *in vitro* a mix of the same RNA and proteins form similar puncta that are consistent with current understanding of RNP granule formation and phase transition models (Tian et al., 2020). By tracking puncta fluorescent signals *in vivo*, we demonstrate that for all slncRNA molecules used in our experiment, the various puncta emitted similar signals characterized by bursts of increasing or decreasing fluorescent intensity. Using these observations and the fact that a similar "burst" signal was also observed from puncta *in vitro*, we concluded that "fluorescent-bursts" corresponded to addition or shedding events of slncRNA-CP nucleoprotein complexes occurring at slow rates that are consistent with the puncta being a phase-separated gel or glass-like RNP granules. Consequently, we show that CP-puncta can be considered synthetic RNP granules, which has important consequences for both the RNP granule research community, as well as the way by which they are currently being used in cell-function studies.

Results

slncRNA-CP complexes form phase separated synthetic RNP granules *in-vivo* and *in-vitro*

We designed two synthetic long non-coding RNA (slncRNA) binding-site cassettes using our binding site resource (Katz et al., 2021, 2019, 2018). The first slncRNA, Qβ-5x_ PP7-4x, consisted of five native Qβ and four native PP7 binding sites, in an interlaced manner. The second slncRNA, Qβ-10x, consisted of ten high-affinity (Katz et al., 2021) Qβ binding sites (Figure. 1a). The new slncRNA cassettes and the PP7-24x cassette (Hocine et al., 2013) were each cloned downstream to a pT7 promoter on a single copy plasmid and transformed into BL21-DE3 *E. coli* cells, together with a plasmid encoding for either Qβ-mCherry or tdPP7-mCherry fusion proteins from an inducible promoter. Single cells expressing the cassettes and CPs were imaged every 10 seconds for 60 minutes under constant conditions on an epi-fluorescent microscope. For all cassettes used in the experiment, the images revealed formation of various puncta at the majority of cell poles (Figure. 1b). Quantifying the fraction of cells that display at least one punctum revealed a dependence on the number of binding sites, in accordance with the multivalency model of LLPS formation (Klosin et al., 2020; Li et al., 2012) (Figure. 1c).

To provide further evidence that these puncta are phase-separated molecular condensates, we lysed cells expressing the Qβ-mCherry fusion protein only, and cells expressing both the fusion protein and the binding site cassette consisting of ten Qβ binding sites. Previous studies have shown that RNP granules survive lysis and require special treatment post-lysis in order to properly extract the RNA (Begovich and Wilhelm, 2020; Lee et al., 2020; Putnam and Seydoux, 2021). To show that CP-puncta maintain their structure post-lysis, we first measured the turbidity of the cell lysates. The results (Figure 2a) show a 1.7-fold increase in turbidity (measured at OD600), a known signature of a liquid suspension containing phase separated droplets (Feng et al., 2019; Wang et al., 2019). We further examined the cell lysates via microscopy (Fig. 2b) and flow cytometer (Fig. 2c). Both the flow cytometry and microscopy

data support an interpretation that a population of fluorescent and diffraction limited particle-like objects exist within the lysates of only the cells expressing the Qβ-10x slncRNA and its QCP binding partner.

We next attempted to create CP-puncta *in vitro* using an *in vitr*o transcribed slncRNA containing 14 and 15 high affinity PP7 and MS2 binding sites arranged in an interlaced manner together with a purified tandem dimer PCP fused to mCherry (tdPCP-mCherry) in standard buffer conditions (see methods and Fig. 3a). To do so we mixed different titers of slncRNA and tdPCP-mCherry protein, each varied over two orders of magnitude. CP-puncta like structures were detected only in slncRNA and proteins concentrations of 100 ng/μl and above (Fig. 3b). The example images display bright puncta that are embedded within a filamentous structure surrounded by a non-fluorescent solution. Quantifying the average intensity of the puncta in the various images obtained (Fig. 3c) reveals a fluorescent intensity distribution which declines by two order of magnitude (i.e., from ~10^5 to ~10^3) in a step-like function as the RNA concentration is reduced from 100 to 10 ng/μl. Likewise, the intensity distribution of the puncta declines in a more gradual fashion as the protein concentration is reduced, but overall a similar disappearance of puncta is observed. Together the *in vitro* slncRNA/protein titration experiments, *in vivo* images, and lysis results are consistent with an interpretation that the CP puncta are synthetic RNP granules, based on phase-based RNP granule formation models which have been described elsewhere (Rhine et al., 2020; Treeck et al., 2018; Wheeler et al., 2016).

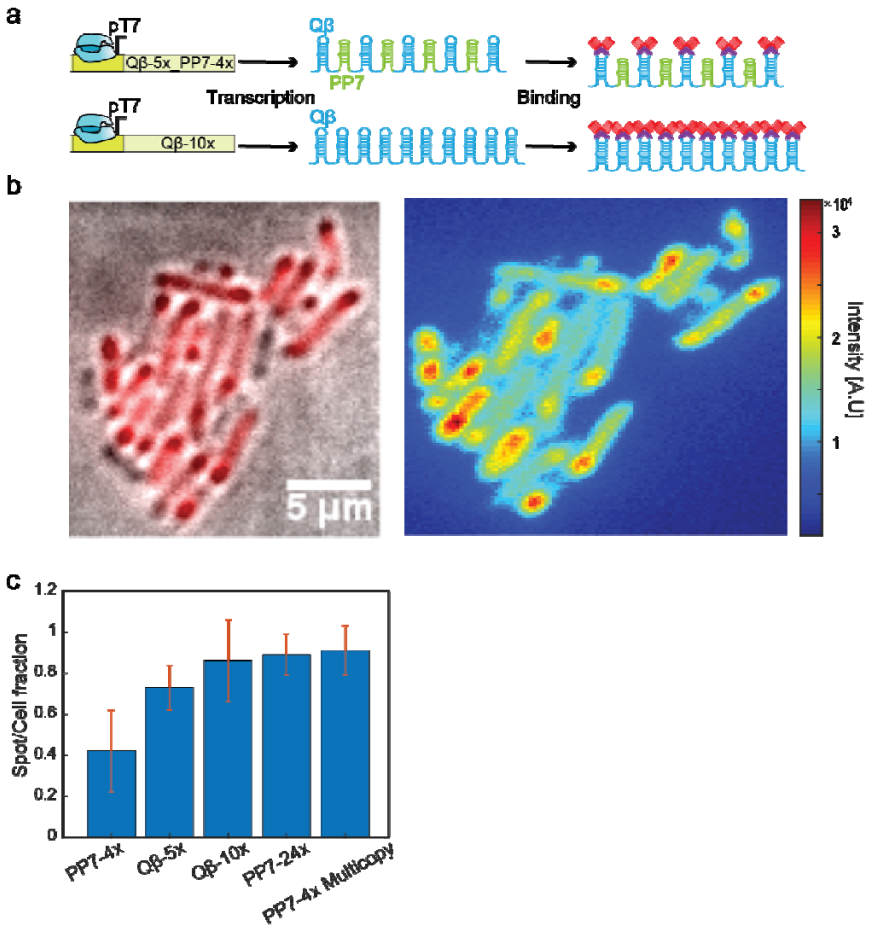


Figure. 1: Synthetic liquid-liquid phase separated droplets within bacterial cells. a, Construct diagram depicting pT7 expression of the two new slncRNA cassettes used in this study, in the presence of Qβ-mCherry. **b,** (left) Fluorescent image of cell expressing the Qβ-5x-PP7-4x slncRNA together with Qβ-mCherry. (right) Heatmap depiction of the image on left showing puncta within cells. **c,** Cell fraction showing number of puncta as a function of cassette-type. Note, PP7-4x and Qβ-5x indicate the Qβ-5x-PP7-4x cassette expressed together with tdPP7-mCherry or Qβ-mCherry, respectively. Error bars indicate standard deviation.

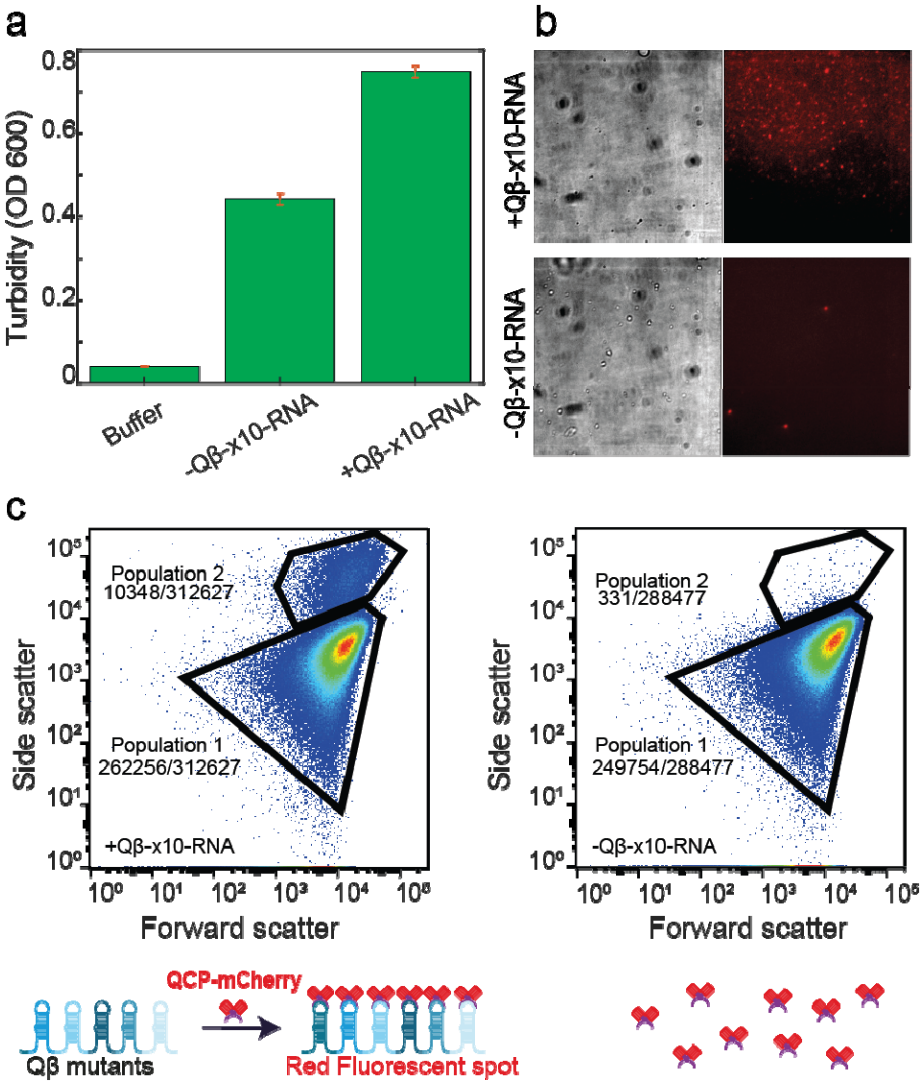


Figure 2: E. coli cell lysates showing a suspension of granules embedded within a dilute liquid. **a**, Turbidity (absorption) measurements of cell lysates expressing QCP-mCherry that also either express the Qβ-10x slncRNA (right) or not (middle). **b,** Microscopy images depicting *E. coli* cell lysates containing Qβ-mCherry with Qβ-10x slncRNA (top) and without (bottom). The lysate containing the slncRNA shows a clear DIC slide and a fluorescent image depicting a dense layer of sub-micron resolution puncta, while the lysate containing only Qβ-mCherry shows a handful of puncta. **c**, Flow cytometry side scatter vs forward scatter plots for the *E. coli* cell lysates. The lysate containing the slncRNAs (left) shows a second population at high side-scatter values that are consistent with denser particles. The same population does not appear in the lysate lacking the slncRNAs (right).

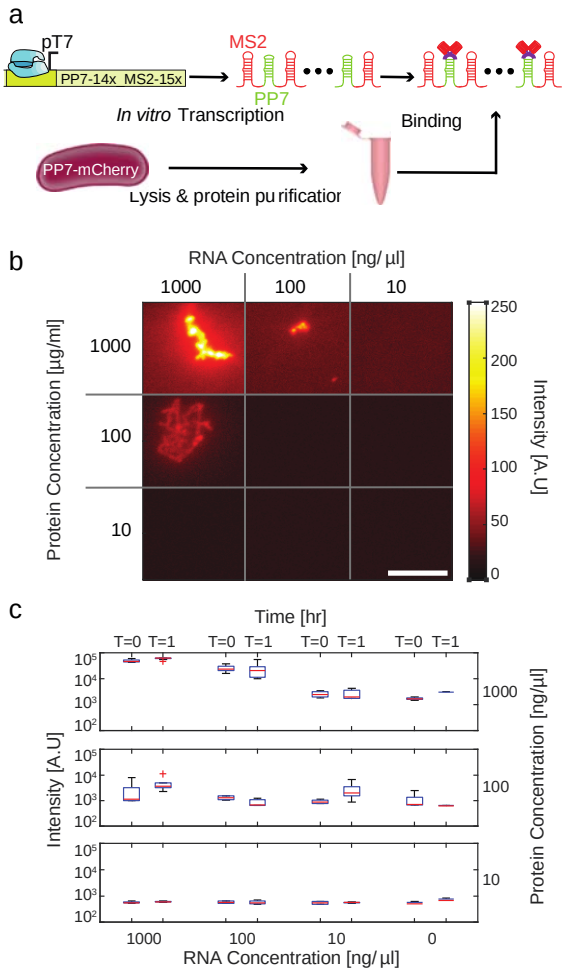


Figure. 3: *In vitro* granule formation. a, Construct diagram depicting pT7 *in vitro* expression of the slncRNA cassettes used, in the presence of tdPP7-mCherry fusion protein purified from bacterial lysate. **b,** Microscopy images for serial dilutions of reaction components taken at T = 1hr after reaction setup. Stock concentrations show the formation of highly fluorescent filamentous structures as seen in the top left image. Lower RNA concentrations result in shorter structures while lower protein concentration result in weaker fluorescence. Scale bar is 10 µm **c,** Maximal observed intensity values for each reaction condition.

Synthetic RNP granules are gel-like particles

A hallmark of liquid-liquid phase separation (LLPS) (Feng et al., 2019; Hyman et al., 2014) is the non-correlated exchange of molecules between one liquid phase and the surrounding phase-separated solution. These exchange events are predicted to occur independently of one another at a rapid rate which depends on the transient concentration of the molecules in the surrounding solution, the average size of the molecules, and their diffusion rate (Klosin et al., 2020). If the RNP-granules are liquid-like, they should exchange with the surrounding solution fluorescent CPs, slncRNA, or CP-bound slncRNA complexes. Such exchanges should be reflected by sudden changes in granule fluorescent intensity. Therefore, to test whether the synthetic RNP granules display liquid-like characteristics, we tracked the fluorescence intensity of each in a field-of-view over 60 minutes. We then analyzed the signal brightness of each diffraction limited granule for every time point using a customized analysis algorithm (see Supplementary Methods and Appendices 1-3). In Figure. 4a, we plot representative intensity vs time signals for the Qβ-5x-PP7-4x-based granules together with Qβ-mCherry (denoted Qβ-5x), obtained from multiple granules tracked in different fields of view on separate days (40 repetitions in total). In Figure 4b, we plot a similar signal from the PP7-14x-MS2-15x slncRNA together with tdPCP-mCherry granule obtained *in vitro* showing a similar temporal behavior to the *in vivo* granules. For both *in vitro* and *in vivo* synthetic RNP granules, the signals are either decreasing or increasing in overall intensity and dispersed within them are sharp variations in brightness, that are also either increasing or decreasing, which we termed "signal bursts". Next, we employed a statistical threshold which flagged these signal variation events whose amplitude was determined to not be part of the underlying signal noise (p-value<1e-3) (See Methods and appendices for definitions of bursts, algorithm details, and relevant numerical controls). These events were classified as either increasing signal bursts (green), decreasing signal bursts (red), or non-classified segments (blue) (Figure. 4a, b). In Figure. 4c we plot the distributions of amplitude (ΔI) for all three event types, obtained from ~300 punctum traces for the Qβ-5x data. The plot shows the distributions of the three separated populations of non-classified, increasing, and decreasing signal bursts, with the number of positive and negative burst events being approximately equal. Moreover, a similarly symmetric burst distribution is recorded for the PP7-4x, Qβ-10x, and PP7-24x cassettes (Figure. 4 –figure supplement 1). However, for the *in vitro experiment*, a skewed distribution towards negative bursts is observed (Fig. 4d). This is consistent with not having active transcription of slncRNA within the *in vitro* set-up leading to a net degradation of the slncRNAs over time, which implies that *in vivo* the balance between positive and negative intensity bursts is reflective of steady state in the levels of slncRNAs within the cells.

Next, we examined whether the occurrence of positive and negative fluorescent intensity burst events are independent of one another. To do so, we checked whether there was a bias for one type of

burst or the other after a non-classified period that lasted more than 2.5 minutes (see Supplementary methods). The results (Figure. 5a) show that no such bias seems to exist, i.e., either a positive or negative burst seems to occur after non-classified events with equal probability for all four cassette types *in vivo*. We then measured the amplitudes of the bursts for all four *in vivo* cassette-RBP pairings and found that both positive and negative amplitudes are proportional to the number of binding sites within the encoded cassette (Figure. 5b), reaching an apparent signal saturation for the Singer PP7-24x slncRNA. Finally, we measured the average lag-time between burst events. To do so, we collected all lag-time measurements and computed the frequency that a particular lag-time interval occurs. The lag-time frequency distribution behaved as an exponential distribution (Fig. 5c), which is consistent with a simple Poisson or Markov process as expected from a non-correlated exchange process of molecules between two phases. This then allowed us to extract the burst event time constant for each granule distribution, which yielded a value of ~20-30 min average lag-time between fluorescent intensity burst events (Fig. 5d). Together, the data provided in Figs. 4 and 5, support an interpretation that fluorescent burst events correspond to insertion and shedding events of O(1) slncRNA-CP complexes into or from the synthetic RNP granules. However, the slow rate (i.e. $O(10^3$ sec)) by which these events are occurring is inconsistent with expectations from an LLPS molecular dynamics (MD) simulation published recently (Yagasaki et al., 2019). Therefore, the slow non-correlated exchange rate, the association of bursts with exchange events of slncRNA-CP complexes (rather than just CPs), and the filament-like structure appearing in the *in vitro* images suggest that the synthetic RNP granules seem to exhibit gel or glass-like properties, rather than classic liquid or solid-like characteristics.

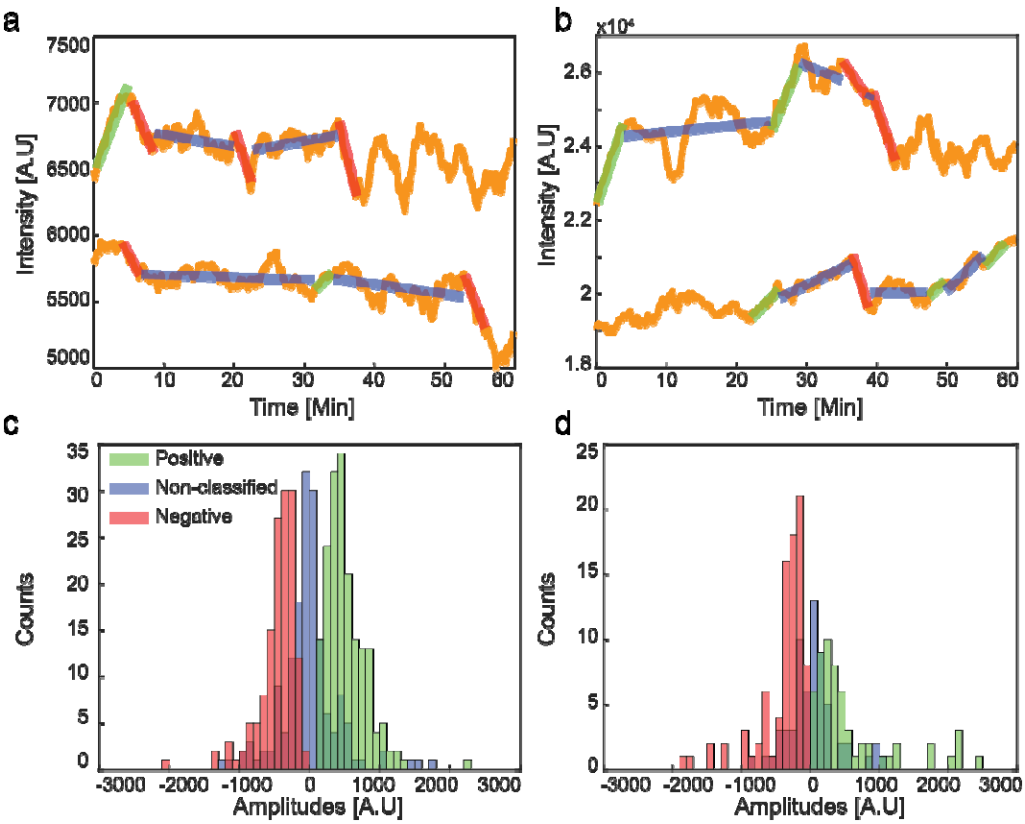


Figure. 4 Synthetic granules are characterized by insertion and shedding events of RNA-RBP complexes. a, b, Sample traces with annotations of puncta signal for the Qβ-5x cassette measured *in vivo* (**a**) and for PP7-14x-MS2-15x slncRNA *in vitro* (**b**). Annotations represent increasing intensity burst events (green), decreasing intensity burst events (red), and non-classified signal (blue), respectively. **c, d,** Empirical amplitude distributions gathered from ~300 traces *in vivo* (**c**) and *in vitro* (**d**). Green, blue, and red correspond to increasing, decreasing and non-classified signal segments.

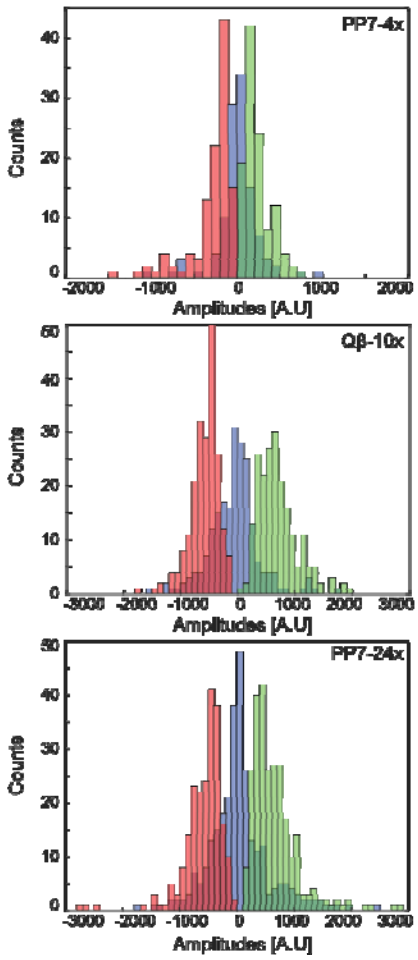


Figure. 4 – figure supplement 1: Amplitude distributions of different binding sites cassettes. Top – PP7-4x, collected from 256 traces, center – Qβ-10x, collected from 430 traces, bottom – PP7-24x, collected from 390 traces.

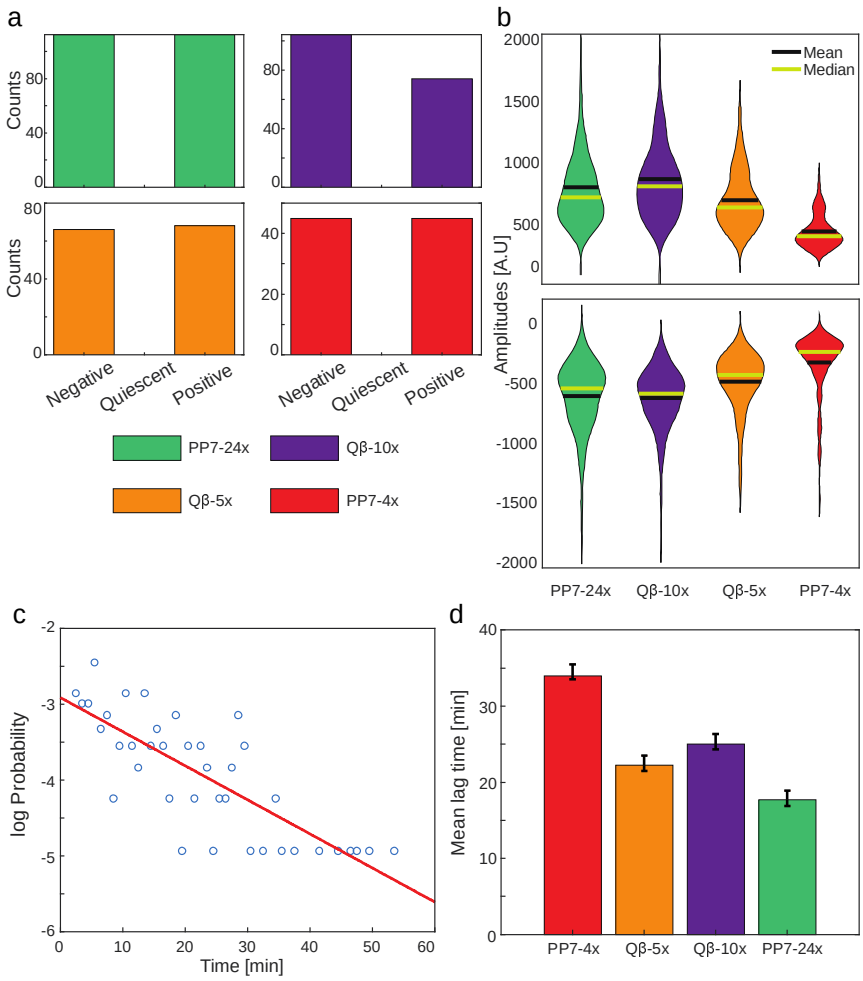


Figure. 5: Synthetic granules exhibit gel-like properties. a, Bar-graph showing the number of events for both negative and positive bursts immediately following a long (>2.5 min) non-classified event. From top-left, in clockwise direction: PP7-24x, Qβ-10x, PP7-4x, Qβ-5x. **b,** Violin plots showing amplitude distribution as a function of cassette type for both increasing (top) and decreasing (bottom) bursts. **c,** Lag-time distribution between insertion events for Qβ-5x r-square of fit is 0.63. **d,** Bar plot showing extracted mean lag times for all four cassette-RBP pairings. Error bars indicate 95% confidence intervals.

Synthetic RNP granule characteristics depend on slncRNA size and CP-type

In order to further characterize the shedding and insertion dynamics occurring between the SRNP granule and the surrounding dilute phase, we estimated the number of slncRNA-CP complexes that exist within the gel-like RNP granule phase. To do so, we fitted each shedding and insertion event amplitude distribution to a Poisson model which is justified by the uncorrelated occurrence of insertion and shedding events as a function of time (Fig. 5a). Figure. 6a-b present a sample fit for the PP7-4x burst amplitude distribution data, with three Poisson functions for λ = 1 (red), 2 (green), and 3 (black), corresponding to a mean of 1, 2, and 3 slncRNA-CP complexes per burst, respectively. The fits show that while the λ=3 distribution provides the best fit to the data (corresponding to a mean of three slncRNAs per burst), the λ=1 distribution provides the best fit to the tail of the distribution but fails at lower amplitude values. This may be due to our analysis threshold that treats many of the small amplitude events as unclassified. Higher values of λ provide a progressively worse fit.

We repeated this analysis for the three additional cassette configurations and computed the estimated intensity per slncRNA-CP complex (K_0) for each slncRNA-type (Figure 6 – figure supplement 1). Both the Poisson fits (Fig. 6c) and empirical distribution analysis (Fig. 5b) suggest that at least for the range of 4-10 binding sites, the number of sites in a cassette can be determined by the amplitude distribution at a resolution as low as a single binding site with a fluorescence signature that can be estimated to be ~40-60 A.U. Furthermore, there seem to be important differences between fitted insertion and shedding amplitudes for PCP and QCP granules. For the former, the fitted insertion and shedding amplitudes are nearly identical for the PP7-4x and PP7-24x, while for the latter the shedding amplitudes are smaller by ~12% for Qβ-5x and ~25% for Qβ-10x respectively. This indicates that for QCP granules slncRNA-CP complexes that are shedded may be either partially degraded or have reduced QCP occupancy as compared with insertion events.

Finally, using the single molecule intensity estimate obtained from the λ=1 approximation, we computed an estimate for the number of slncRNA-RBP complexes within each granule, averaged over the duration of the trace. We then plotted the distribution of the average number of complexes per granule, for each cassette-RBP pairing (Fig. 6d). The results show that for the Qβ-5x, Qβ-10x, and PP7-24x slncRNA cassettes granules are estimated to contain a similar number of ~10-30 slncRNA-RBP complexes, while the granules for the PP7-4x cassette seem to be comprised of about half this amount. Consequently, the SRNP granules seem to be microgel-like particles that are composed of several tens of RNA molecules and more than a hundred CP-mCherry proteins, and whose particular structure depends on slncRNA and CP-type.

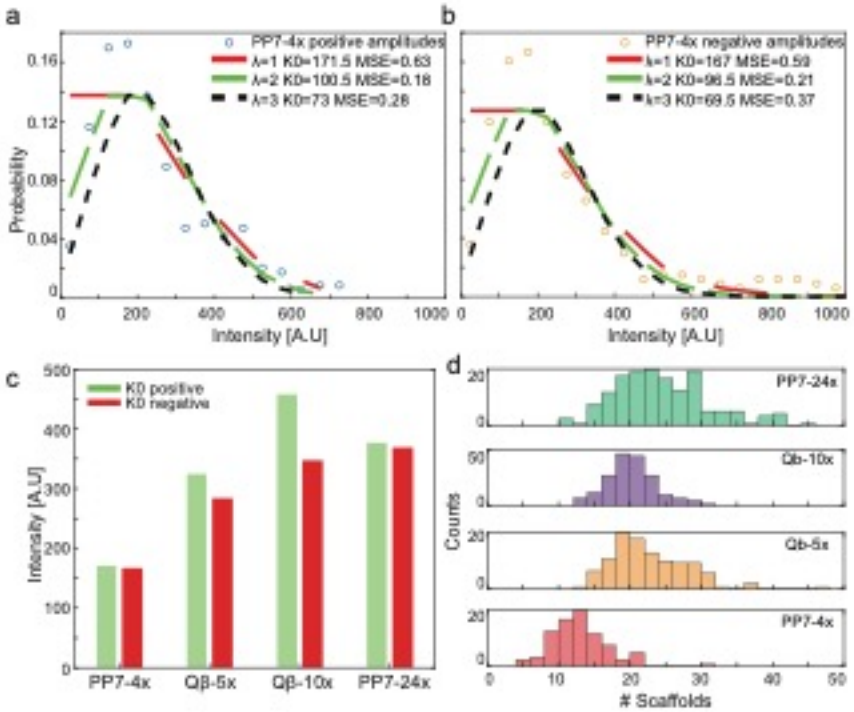


Figure. 6: Insertion and shedding event analysis suggest that granule structure depends on slncRNA and CP-type a,b, Poisson functions fits for the amplitude distribution of insertion assuming 1, 2, or 3 mean insertion (**a**) and shedding (**b**) events. **c,** Extracted burst amplitude for a single slncRNA-RBP complex, assuming a Poisson distribution with λ=1. (Green) and (red) correspond to insertion and shedding amplitudes, respectively. **d,** Distribution corresponding to the number of slncRNAs per puncta, assuming the value of K0 shown in panel (**c**).

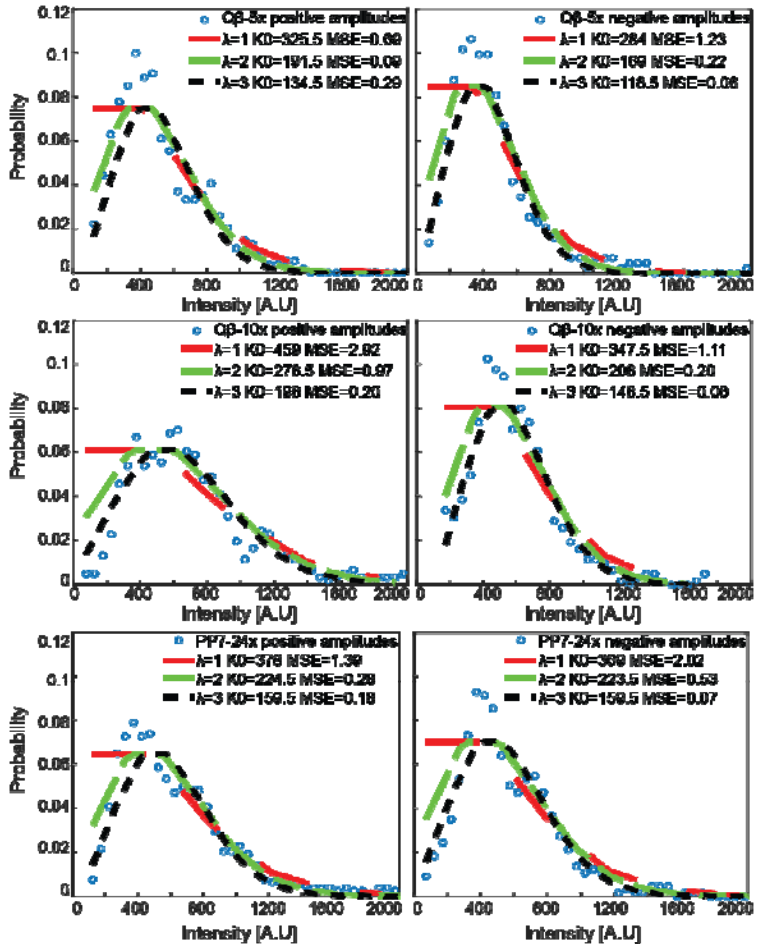


Figure 6 – figure supplement 1: Fitting of amplitude data to Poisson distributions. Top Row – Qβ-5x, middle row – Qβ-10x, bottom row – PP7-24x. Left column – insertion event amplitudes, right column – shedding event amplitudes.

Cytosolic density variation controls location of CP-slncRNA granule formation

We next wanted to explore the underlying cause for the propensity of SRNP granules to form only in the cell poles of exponentially growing cells. We hypothesized that this was due to cytosolic density variation within the *E. coli* bacterium. Specifically, it is assumed that the *E.coli* cytosol consists of a dense molecular region in the central portion of the cell consistent with the location of the nucleoid (Fisher et al., 2013), and a semi-dilute molecular phase in the polar regions. Given the fact that RNP granules are known to form via liquid-liquid phase separation and need to be enveloped by a semi-dilute liquid, it is therefore not surprising that a molecular dense region like the nucleoid does not facilitate this transition. To provide support for the density variation hypothesis, we grew the SRNP-granule expressing cells in starvation conditions for several hours, triggering a transition to stationary phase (see methods). In stationary phase the nucleoid is known to condense (Janissen et al., 2018; Kim et al., 2004; Wolf et al., 1999), thus increasing the amount of cellular volume which is likely to be molecularly dilute. This, in theory, generates a much larger accessible cellular volume for synthetic RNP-granule formation, which should lead to a different presentation of the RNP granule phenomena as compared with exponentially growing cells. In Figure 7, we show typical images of stationary-phase bacteria displaying 'bridging' (the formation of a high intensity streak between the spots) of puncta (Fig. 7a), whereby RNP granules seem to fill out the available dilute volume, and the emergence of a third granule at the center of the cell (Fig. 7b). Both behaviors are observed in ~40% of stationary-phase cells and are morphologically different than the granules appearing under normal growth conditions (Fig. 7c).

A potential consequence of the fact CP-slncRNA complexes cannot form granules in the dense-nucleoid region of the cell is that this region likely serves as a reservoir of CP-slncRNA complexes. This assumption is based on two tenets: first, the nucleoid region is where new slncRNA and CP proteins are being synthesized. Second, the lack of membrane separating the nucleoid from the semi-dilute polar regions implies that there must also be free exchange of CP, slncRNA, and CP-slncRNA complexes between both regions. The synthesis and exchange processes maintain slncRNA and CP-titers at steady-state leading to the symmetric burst distribution observed in our *in vivo* experiments. We then wondered how an increase in the amount of steady state slncRNA molecules would affect the structure of the granule, over-all cell fluorescence, and burst distributions. To do so, we expressed the PP7-4x on a multicopy plasmid and compared microscopy observations to the ones obtained for the single copy PP7-4x encoded on the single-copy BAC. As Fig. 7d-e shows, we observed an increase in both the background signal (Fig. 7d), and in the average number of estimated CP-slncRNA complexes within the granule (Fig. 7e), while maintaining the symmetric burst distribution which is a hallmark of all our *in vivo* experiments. Consequently, the exchange of slncRNA-CP complexes between both domains must occur at some rate,

which controls both the average size of the RNP granules and the level of the background fluorescence reflecting the concentration of free CP-slncRNA complexes within the nucleoid region.

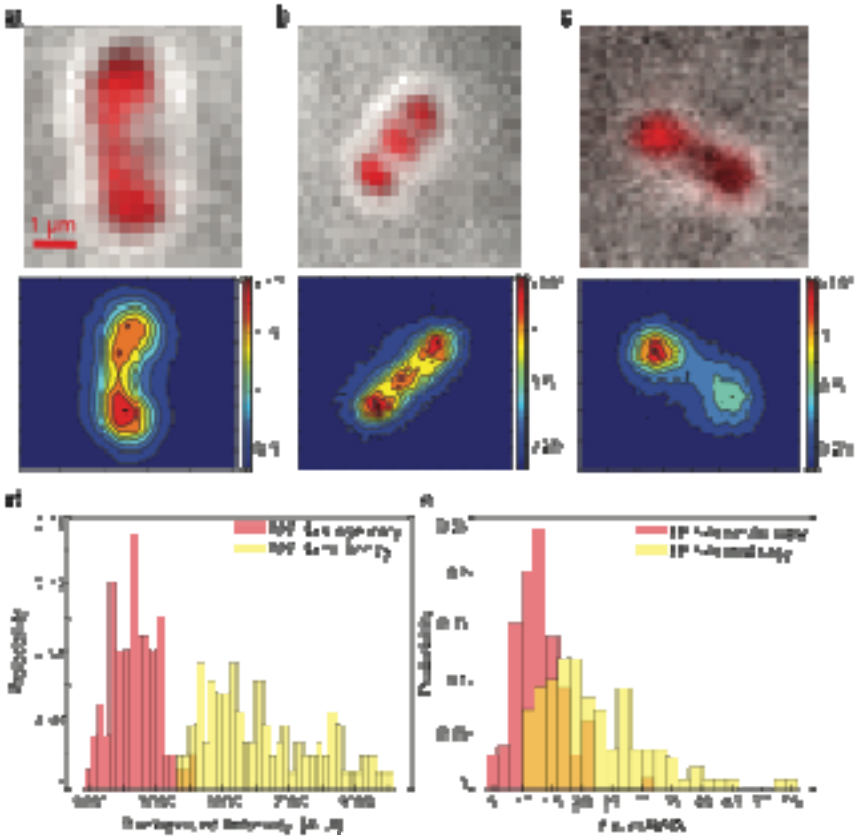


Figure. 7: Polar location Synthetic RNP granules reflects a biphasic structure for cytosol. **a-b,** Typical images of fluorescent bacteria in stationary phase, which are very different than the 2-granule image obtained for exponentially growing cells **(c)**. A close examination shows "bridging" or spreading of granules **(a)**, and emergence of an additional granule in the middle of the cell **(b)**. **d,** Background fluorescence signal for the PP7-4x slncRNA expressed from a multi-copy plasmid (yellow) and single-copy plasmid (red). **e**, Distribution of the number of slncRNA-RBP complexes within the granule for each case.

Granule-based gene-expression model

To place our findings in a broader theoretical perspective, we constructed a kinetic model for steady state slncRNA and CP expression in the presence of existent SRNP granules based on the work of (Klosin et al., 2020) (see schematic in Fig. 8a). In the model, we assume that upon slncRNA transcription and subsequent CP-slncRNA complex assembly within the nucleoid region, complexes diffuse in and out

from the nucleoid region and into the semi-dilute polar regions at some rate. In addition, another kinetic process controls the absorption and shedding events of the CP-slncRNA complexes into and out of the existent SRNP granules. The model assumes for the sake of generality different degradation rates for the slncRNA-CP complexes within the nucleoid, semi-dilute, and SRNP granule domains (γ_n, γ_+, and γ_- respectively). This in turn can lead to the increased levels of steady-state CP-mCherry as a function of slncRNA structure as γ_- in particular can potentially depend on the structure of the slncRNA component of the granule.

Evaluating the model (see model section in methods) even without taking into consideration different degradation rates inside and outside the granules leads to a prediction that the frequency of absorption events should occur at a rate that is proportional to the CP-slncRNA complex's volume fraction within the nucleoid phase. To test this prediction, we plot in Fig. 8b-left the mean fluorescent background intensity as a function of mean lag-time between events measured for each slncRNA type used in the experiment. The plot shows a linear relationship for all the single copy systems (i.e., cells where the slncRNAs are expressed from a single copy plasmid) as predicted by the model. In addition, background CP-mCherry fluorescence is significantly higher for all constructs as compared with an *E. coli* cell expressing CP-mCherry without a cognate slncRNA target (Fig. 8b-right) and depends strongly on the number of binding sites encoded within the slncRNA. This observation is consistent with a having reduced slncRNA-CP degradation rate, which in turn leads to a higher steady state level of the CPs. Thus, SRNP granules can allow for increased expression of desirable proteins beyond the generic transcriptional, translational, and degradation rates of these proteins when they are not bound in complex.

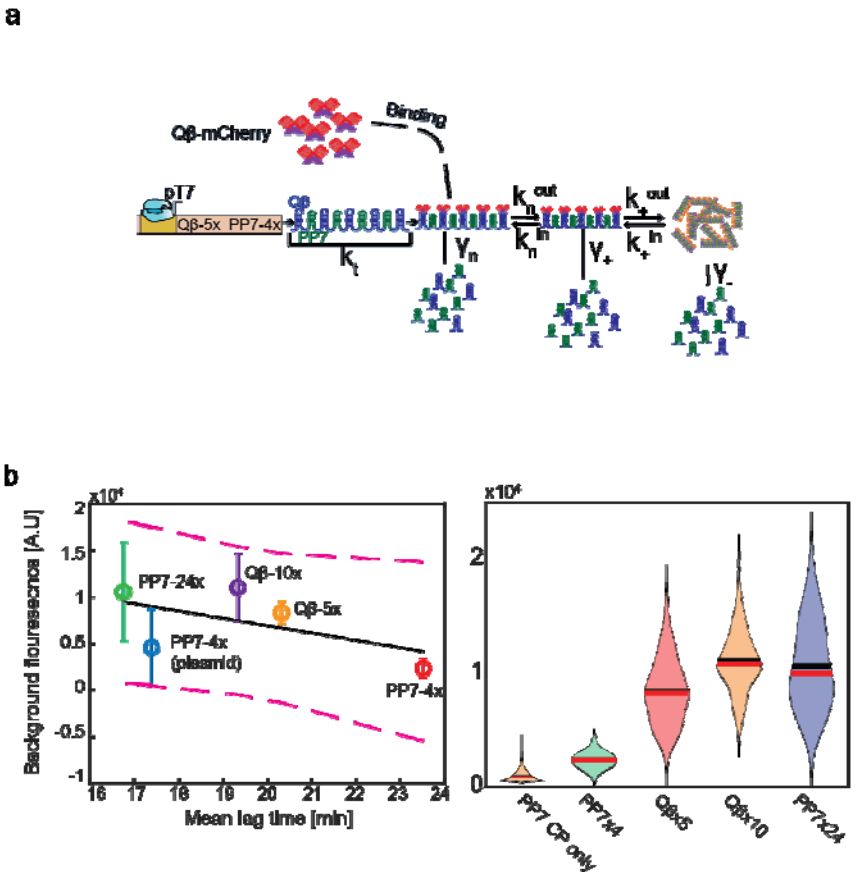


Figure. 8: Polar location Synthetic RNP granules reflects a biphasic structure for cytosol. **a**, Model showing the effects of the biphasic hypothesis on insertion and shedding of a slncRNA. Parameters: and are the slncRNA transcriptional and degradation rates, , correspond to the rates by which the slncRNA-RBP complexes leave/re-enter the nucleoid phase, and correspond to the insertion/shedding rates of the slncRNA-RBP complexes from the dilute to the droplet phase. The biphasic model is an extension of the simple rate-equation gene expression model and leads to a Super-Poisson distribution of RNA for any RNA species (Appendix 4). **b**, (left) scatter plot of background fluorescence vs. mean lag time between insertion events for all cassette types. The experimental systems containing a single copy plasmid show a linear relationship as depicted by the black line. The multi-copy plasmid experimental system shows a deviation from that model in regard to background fluorescence intensity. (right) Violin distribution of background intensity for each construct used in the experiment.

Discussion

In this study, we show that synthetic RNP granules can be designed and assembled using phage coat proteins and RNA molecules that encode multiple CP binding sites. We demonstrate granule formation *in vivo*, *in vitro*, and via cell lysis. In addition, using fluorescent signal analysis we show that

synthetic CP-based RNP granules are composed of a concentrated RNA-protein ensemble which exhibit gel-like characteristics rather than classic liquid or solid. It is assumed that in order to form RNP granules there must be an attractive cross interaction either between the protein themselves, or between the various RNA molecules. Such an interaction is thought to trigger the nucleation of granules (Wheeler et al., 2016). In the case of the CP-based SRNP granules used here, the protein component used was either tdPCP or QCP. For both cases, we used CP mutants that are incapable of forming capsids, and thus are assumed to be non-interacting (Chao et al., 2008). This is supported by the fact that we did not observe inclusion bodies, fluorescent granules, or residual capsid formation in any experiment *in vitro* and *in vivo* that only included high titer of the CP-proteins here or in previous works (Katz et al., 2021, 2019, 2018). However, since quantitative differences were observed between the two granule types (i.e., reduced shedding amplitude for QCP-granules as compared with tdPCP-granules- Fig. 6c), it is possible that a residual weak attractive interaction was nevertheless maintained. While this weak interaction is insufficient for the assembly of capsids or for granules made of RNA with three binding sites or less, it may be sufficiently strong to allow for the assembly of granules when the slncRNA contains four or more binding sites. Alternatively, it is also possible that due to the ubiquity of granules in nature, the default state for any RNA bound by multiple non-charged proteins will be to nucleate granule formation due to simple steric considerations, and that our observations may be part of a general trend.

We also demonstrated that the synthetic granules can be used as a novel synthetic biology tool for applied and basic research. As a basic research tool, we relied on the granule tendency to only nucleate in a semi-dilute liquid environment to infer that *E. coli* has at least two cytosolic domains, which are differentiated by biomolecular density. The dense domain is mostly associated with the nucleoid, while the semi-dilute domain occupies the non-nucleoid regions of the cell. The immediate consequence of this model is that RNP granules can only form in polar regions of exponentially growing cells. Therefore, RNA that is found in the poles is not likely to require an active mode of localization as for membrane or nucleoid targeting mRNAs (Kannaiah et al., 2019). On the applicative side, we show that cells co-expressing CP-mCherry and slncRNAs display a stronger fluorescent signal than cell expressing CP-mCherry alone. The amount of increased expression depends on the number of binding sites encoded on the slncRNA. This indicates that these cells have a higher steady-state levels of CP-mCherry proteins, and that the slncRNAs and/or the granules provide protection for the proteins from degradation. This increase in protein titer has the potential to provide a new method for increasing the capacity of protein production, which is needed in the biotechnological, food, and pharmaceutical industries.

Finally, the various set of granule related phenomenon have inspired us to formulate a generic transcriptional model, which takes into account kinetic transitions between the phases within the bacterial cell. In its simplest form (i.e. without CP-proteins present, and thus no granule formation), our model

proposes that in addition to transcription and degradation, a typical bacterial mRNA will also be subject to a kinetic process of molecular exchange across the interface between the dense nucleoid and semi-dilute cytosolic domains. In this simplified form, the model is nearly identical (see methods) to the one analyzed by Sanchez *et al*.(Sanchez *et al*., 2011) , which yielded a prediction of Super-Poisson distribution for bacterial mRNA. This prediction has been verified by various experimental studies in recent years (Chong et al., 2014; Golding et al., 2005; Jones and Elf, 2018). Prior to this work, these Super-Poisson mRNA distributions have been attributed to sporadic bursts of mRNA transcription that are spaced by periods of promoter quiescence due to an hypothesized stochastic promoter switching mechanism (Lionnet and Singer, 2012; Paulsson, 2005; So et al., 2011). By contrast, in our model a Super-Poisson distribution for gene expression emerges as a mechanism-independent and parameter-free prediction (see Figure. 8a and appendix 4), which is a direct consequence of the interface that differentiates between the two biomolecular domains across the bacterial cell. Therefore, our model corresponds to an Occam's razor explanation for the Super-Poisson distributions and if proven to be true by additional experimental studies is likely to have far-reaching implications for our understanding bacterial gene expression and possibly for Eukaryotic cells as well.

Materials and Methods

Bacterial strains

E. coli BL21-DE3 cells which encode the gene for T7 RNAP downstream from an inducible pLac/Ara promoter was used for all reported experiments. *E. coli* TOP10 (Invitrogen, Life Technologies, Cergy-Pontoise) was used for cloning procedures.

Addgene plasmids

- pCR4-24XPP7SL was a gift from Robert Singer (Addgene plasmid # 31864; http://n2t.net/addgene:31864; RRID: Addgene_31864)
- pBAC-lacZ was a gift from Keith Joung (Addgene plasmid # 13422; http://n2t.net/addgene:13422; RRID: Addgene_13422)

Construction of the binding sites cassettes plasmids

The cassette sequence containing 5 PP7-wt and 4 Qβ-wt binding sites with randomized spacer sequences:

cctaggcgattatgacgttattctactttgattgtgatgcatgtctaagacagcatcgcctgctggtcgtgactaaggagtttatatggaaacccttacgagacaatgctacctt
accggtcgggcccacttgtttttacccatgatgcatgtctaagacagcatcgcctgctggtcgtgactaaggagtttatatggaaacccttagaaacagccgtcgccttgaa

gccgagaacaatgcatgtctaagacagcatatggattgcctgtctgttaaggagtttatatggaaacccttacatcaggcttcgcagtatgcaacgcttgcgatgcatgtcta
agacagcatttcaccgctttcctaagtaaggagtttatatggaaacccttagtactaactcgcagatgcatgtctaagacagcatcagaaacgtcacgtcctggc
(Qβ and PP7 binding sites marked in red and green respectively), was ordered from GenScript, Inc. (Piscataway, NJ), as part of a Puc57 plasmid, flanked by EcoRI and HindIII restriction sites.

pBAC-lacZ backbone plasmid was obtained from Addgene (plasmid #13422). Both insert and vector were digested using the above restriction sites and ligated to form BAC-Qβ-5x-PP7-4x.

The Qβ-10x cassette was designed using the RNA binding sites resource (Katz et al., 2021). The binding site sequence:

gaattcttacaaaggaactgtaacagtccttctcgtgctgatcgtgacttggatgtccaagacaccaacgagacaatgctaccttaccgtcggcccacttgtttttacccatga
catgacgagatactcgcatgtcgcctgctggtcgtgacatgcatgtctaagacagcatgaaacagccgtcgccttgaagccgagaacattgcatgtcgaagacagcaaat
ggattgggtctccaattcctgtctgtttccatgactaagtcaggaacatcaggcttcgcagtatgcaacgcttgcgatgcattgcaaagcaagcatttcaccgctttcctaaga
aggatagtaatgactaccttgtactaactcgcagatcgaactctaagagtcgatcagaaacgtcacgtcctggcaaccatgtcagggacaggtttggaagaattc

(Qβ binding sites marked in red), was ordered from Twist Bioscience (San Francisco, CA), flanked by BamHI restriction sites. Insert and pSMART BAC (Lucigen, Middleton, WI) vector were digested with BamHI and ligated to form BAC-Qβ-10x.

The sequence containing the PP7-14x_MS2-15x binding sites cassette:

Gagaaacgtttcgacattat
atggaatgcgaaagtggaacgtaatggacatgaagacgattacgcttcacacggaggatgcgggaaacatgaagatcacccatgttcgcttaaccatggat
agggatcacccatgttgcggtggtgcgtcaaccagagatttcatatgggaaactctgggacacgctgtatttatacatgaggatcaccatgtgtgcttaaatatgggtaagtt
gaccattaggcaactgtaagatgctccggttaattccagtttatatggaaacggaattgatgtaccgttgagcaagaacacgattacggttcttcgattagatatggtttaaac
gacaatatatggattgcgtttggggcacgccgtctggagaagaccattaggcttcttactgcgaccgcaataaaggagattatatgaaatccctttgcgcgcaaccggg
tagaagatcaccattagggatctgtaactcacggcgctattacgagtcaatatggtgaccgtaaagctagggcatgtgccagaagagcattagccttctccttggtgggga
agcgataagcacattataaggaatggcttaaagtggtgcggcgggggacttgaccattaggcaagtgtgctagaccctggtctttcgagaaaatatggtttccgaaagaa
ctatacgaagtgacatgcgaggattacccgcatatggtgcaaatgggagaattggagtaaatatggttacccaataggctagagcatgacggcagtgagaattatccact
ggttagcgggttaccgagattgcacattatatggaatggcaattgattcatgccggtcgtttgtgaggagtacccacaaaatgggagggtgctatataaccaggttatatgc
aaccggttaggccgttgtgttagtttagttcagcattagcgaactgtgcaagacccggtggctaaggagtttatatggaaacccttagccctcgagcatgctaacatgagga
ttacccatgtgatgggtttgaaacgtgcaattatttgatatggcaaaaattggtagggcatggctgcagcgtgagaattatccacgctaggctagagcatggcggtattga
gttcgggtttgagatgaccattaggcatctgtgctagagcatgcgaaaacgacataatatggtatgcgttttggggtgctagtataccacatgatgagcacacatgtatgg
(MS2 and PP7 binding sites marked in blue and red respectively), was ordered from Integrated DNA Technologies (Coralville, IA), together with a T7 promoter sequence as a gblock, and cloned into a vector for amplification.

Design and Construction of Fusion-RBP Plasmids

Fusion-RBP plasmids were constructed as previously reported (Katz et al., 2018). Briefly, RBP sequences lacking a stop codon were amplified via PCR off either Addgene or custom-ordered templates. All RBPs presented (PCP, and QCP) were cloned into the RBP plasmid between restriction sites KpnI

and AgeI, immediately upstream of an mCerulean gene lacking a start codon, under the so-called RhlR promoter containing the rhlAB las box(Medina et al., 2003) and induced by N-butyryl-L-homoserine lactone (C4-HSL) (Cayman Chemicals, Ann Arbor, Michigan). The backbone contained either an Ampicillin (Amp) or Kanamycin (Kan) resistance gene depending on experiment. mCerulean gene was replaced by mCherry using restriction cloning between sites XbaI and AgeI.

Sample preparation

BL21-DE3 cells expressing the two plasmid system (single copy plasmid containing the binding sites array, and a multicopy plasmid containing the fluorescent protein fused to an RNA binding protein) were grown overnight in 5 ml Luria Broth (LB), in 37□ with appropriate antibiotics (CM, AMP), and in the presence of two inducers – 1.6 ul Isopropyl β-D-1-thiogalactopyranoside (IPTG) (final concentration 1 mM), and 2.5 ul C4-HSL (final concentration 60 μM) to induce expression of T7 RNA polymerase and the RBP-FP respectively. Overnight culture was diluted 1:100 into 3 ml solution of BioAssay (BA)-LB (95%-5% v:v) with appropriate antibiotics and induced with 1 μl IPTG (final concentration 1 mM) and 1.5 μl C4-HSL (final concentration 60 μM). For stationary phase tests, cells were diluted into 3 ml Dulbecco's Phosphate-Buffered Saline (PBS) (Biological Industries, Israel) with similar quantities of induction and antibiotics. Culture was shaken for 3 hours in 37□ before being applied to a gel slide (3 ml PBSx1, mixed with 0.045g SeaPlaque low melting Agarose (Lonza, Switzerland), heated for 20 seconds and allowed to cool for 25 minutes). 1.5 μl cell culture was deposited on a gel slide and allowed to settle for an additional 30 minutes before imaging.

Cell lysis and extract analysis

Two strains of BL21-DE3 cells, one expressing both the Qβ-mCherry fusion protein and the Qβ-10x binding sites cassette, and the other expressing only the fusion protein, were grown overnight in 10 ml LB with appropriate antibiotics in 37° C. Following overnight growth cultures were diluted 1/100 into two vials of 500 ml Terrific Broth (TB), with appropriate antibiotics and full induction (150 μl IPTG and 250 μl C4-HSL) and grown in 37° C to OD_{600} > 10. Cells were harvested, resuspended in 45 ml of buffer (50 mM Tris-HCl pH 7.0, 100 mM NaCl and 0.02% NaN_3), disrupted by four passages through an EmulsiFlex-C3 homogenizer (Avestin Inc., Ottawa, Canada), and centrifuged (13,300 *RPM* for 30 min) to obtain a soluble extract. Turbidity was measured using a plate reader (Tecan, F200) at $OD_{600.}$ Flow cytometry measurements were done using MACSQuant VYB flow cytometer (Miltenyi Biotec, Auburn, CA).

Microscopy

Gel slide was kept at 37° inside an Okolab microscope incubator (Okolab, Italy). A time lapse experiment was carried out by tracking a field of view for 60 minutes on Nikon Eclipse Ti-E epifluorescent microscope (Nikon, Japan) using the Andor iXon Ultra EMCCD camera at 6 frames-per-minute with a 250 msec exposure time per frame to avoid photo-bleaching and sufficient recovery of fluorescence signal. Excitation was performed at 585 [nm] (mCherry) wavelengths by a CooLED (Andover, UK) PE excitation system.

Quantification of the fraction of cells presenting puncta was done by taking 10-15 snapshots of different fields of view (FOV) containing cells. The number of cells showing puncta and the total number of fluorescent cells in the FOV were counted manually.

In vitro transcription of RNA binding sites cassette

The vector containing the PP7-14x_MS2-15x cassette was digested with a restriction site immediately following the desired sequence. Purified digestion products were used as template for in vitro transcription using HiScribe™ T7 High Yield RNA Synthesis Kit (NEB, New England Biolabs). RNA products were purified using Monarch RNA Cleanup Kit (NEB) and stored in -80□.

Protein expression and purification

E. coli cells expressing both the tdPP7-mCherry fusion protein and the PP7-14x_MS2-15x binding sites cassette were grown overnight in 10 ml LB with appropriate antibiotics in 37□ C. Following overnight growth cultures were diluted 1/100 into two vials of 500 ml Terrific Broth (TB), with appropriate antibiotics and induction (100 μl C4-HSL) and grown in 37□ C to OD600 > 10. Cells were harvested, resuspended in 30 ml of buffer (50 mM Tris-HCl pH 7.0, 100 mM NaCl and 0.02% NaN3), disrupted by four passages through an EmulsiFlex-C3 homogenizer (Avestin Inc., Ottawa, Canada), and centrifuged (13,300 RPM for 30 min) to obtain a soluble extract. Fusion protein was purified using the HisLink Protein purification resin (Promega) according to the manufacturer's instructions and eluted in 1xPBS.

In vitro granule assembly assay

In vitro experiments were performed in granule buffer (750 mM NaCl, 1 mM MgCl2, 10% PEG4000). Reactions were set up as follows: 8 ul granule buffer, 1 ul protein, 1 ul RNA, 0.5 ul Ribolock RNase Inhibitor (Thermo Fisher Scientific). 1 ul from the reaction was then deposited on a glass slide and imaged in a Nikon Eclipse Ti-E epifluorescent microscope (Nikon, Japan) using the Andor iXon Ultra EMCCD camera with a 200 msec exposure time. Excitation was performed at 585 [nm] (mCherry)

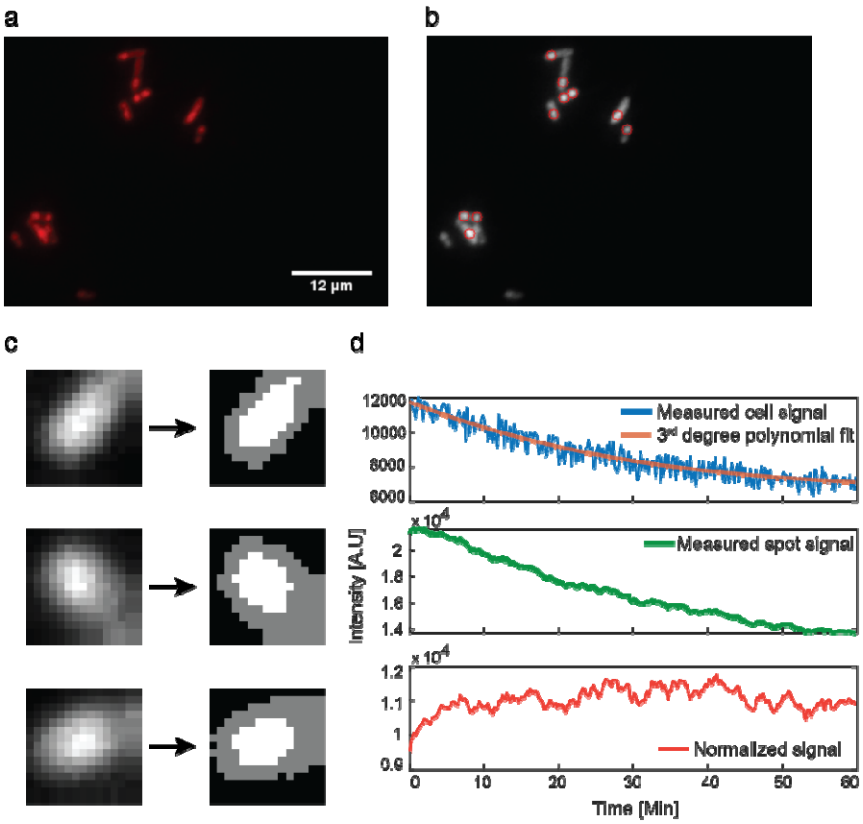
wavelengths by a CooLED (Andover, UK) PE excitation system. Images were taken immediately after reaction setup (time T=0) and after 1 hour (time T=1).

Appendix 1

Image Analysis

The brightest spots (top 10%) in the field of view were tracked over time and space via the imageJ MosaicSuite plugin (Sbalzarini and Koumoutsakos, 2005; Schindelin et al., 2012; Schneider et al., 2012). A typical field of view usually contained dozens of cells, a portion of which were not fluorescent while others presented distinct bright speckles, localized at the cell poles (Appendix 1 – figure 1a,b).

The tracking data, (x,y,t coordinates of the bright spots centroids), together with the raw microscopy images were fed to a custom built Matlab (The Mathworks, Natick, MA) script designed to normalize the relevant spot data. Normalization was carried out as follows: for each bright spot, a 14-pixel wide sub-frame was extracted from the field of view, with the spot at its center. Each pixel in the sub-frame was classified to one of three categories according to its intensity value. The brightest pixels were classified as ‘spot region’ and would usually appear in a cluster, corresponding to the spot itself. The dimmest pixels were classified as ‘dark background’, corresponding to an empty region in the field of view. Lastly, values in between were classified as ‘cell background’ (Appendix 1 – figure 1c). Classification was done automatically using Otsu’s method (Otsu, 1979). From each sub-frame, two values were extracted, the mean of the ‘spot region’ pixels and the mean of the ‘cell background’ pixels, corresponding to spot intensity value and cell intensity value. This was repeated for each spot from each frame in the data, resulting in sequences of intensity vs. time for the spot itself and for the cell background. (Appendix 1 – figure 1d)



Appendix 1 – figure 1: **Image processing scheme**. **a**, Raw microscopy image showing bacterial cells containing bright spots. **b**, Bright spots are identified and their position over time and space is recorded. **c**, The environment of each spot is classified into 3 regions, based on intensity values. The brightest pixels are classified as 'spot' (marked in white), the darkest pixels are classified as 'dark background' indicating empty space, and pixels with intermediate values are classified as cell background (marked in gray). **d**, The mean values of the spot pixels and cell background pixels are recorded over time resulting in the spot signal (green) and cell signal (blue). The spot signal is then normalized to remove photobleaching and global background effects (red).

Signal Analysis

We assume a noise model comprised of both additive and exponential components, corresponding to fluorescent proteins (bound or unbound) not relating to the spot itself, and photobleaching. This can be described as follows:

$$y(t) = (S(t) + c(t)) \cdot f(t) \tag{1.1}$$

$$c(t) = c_0(t) \cdot f(t) \tag{1.2}$$

where $y(t)$ is the observed spot signal, $S(t)$ is the underlying spot signal which we try to extract, $c(t)$ is the observed cell background signal, $c_0(t)$ is the underlying background signal and $f(t)$ is the photobleaching component.

To find $S(t)$, we assume:

$$c_0(t) \approx c_0 = const \tag{1.3}$$

This leads to:

$$\frac{y(t)}{c(t)} = \frac{S(t) + c_0}{c_0} \tag{1.4}$$

$$S(t) = c_0\left(\frac{y(t)}{c(t)}\right) - c_0 = c_0\left(\frac{y(t)}{c(t)} - 1\right) \tag{1.5}$$

To get $y(t)$, we filter the measured spot signal with a moving average of span 13, in order to remove high frequency noise effects, and smooth out fluctuations (see section – Identifying burst events). To get $c(t)$, we fit the measured cell background signal to a 3rd degree polynomial (fitting to higher degree polynomials did not change the results). This is done to capture the general trend of the signal while completely eliminating fluctuations due to random noise.

Identifying burst events

We assume the total fluorescence is comprised of three distinct signal processes: RNP granule fluorescence, background fluorescence and noise. We further assume that background fluorescence is slowly changing, as compared with granule fluorescence which depends on the dynamic and frequent insertion and shedding events occurring in the droplet. Finally, we consider noise to be a symmetric, memory-less process. Based on these assumptions, we define a “signal-burst” event as a change or shift in the level of signal intensity leading to either a higher or lower new sustainable signal intensity level. To

identify such shifts in the base-line fluorescence intensity, we use a moving-average filter of 13 points (i.e., 2 minutes) to smooth the data. The effect of such an operation is to bias the fluctuations of the smoothed noisy signal in the immediate vicinity of the bursts towards either a gradual increase or decrease in the signal (Appendix 1 – figure 2a). Random single fluctuations, which do not settle on a new baseline level are not expected to generate a gradual and continuous increase or decrease over multiple time-points in a smoothed signal. Following this, we search for contiguous segments of gradual increase or decrease and record only those whose probability for occurrence is 1 in 1000 or less given a Null hypothesis of randomly fluctuating noise.

To translate this probability to a computational threshold, we first compute the intensity difference distribution for every trace separately. This distribution is computed by collecting all the instantaneous differences in signal ($\Delta S(t_i)= S(t_i)- S(t_{i-1})$) and binning them (Appendix 1 – figure 2b). Given a particular trace the likelihood for observing an instantaneous signal increase event in a time-point (t_i) can therefore be computed as follows:

$$p_{inc} = \frac{N(\Delta S(t_i) > 0)}{N_{tot}} \tag{1.6}$$

where $N(\Delta S(t_i)>0)$ and N_{tot} correspond to the number of increasing instantaneous events and total number of events in a trace respectively. Likewise, the number of decreasing instantaneous events is defined as:

$$p_{dec} = \frac{N(\Delta S(t_i) < 0)}{N_{tot}} \tag{1.7}$$

This in turn allows us to compute the number of consecutive instantaneous signal increase events (m) to satisfy our 1 in 1000 threshold for a significant signal increase burst event m as follows:

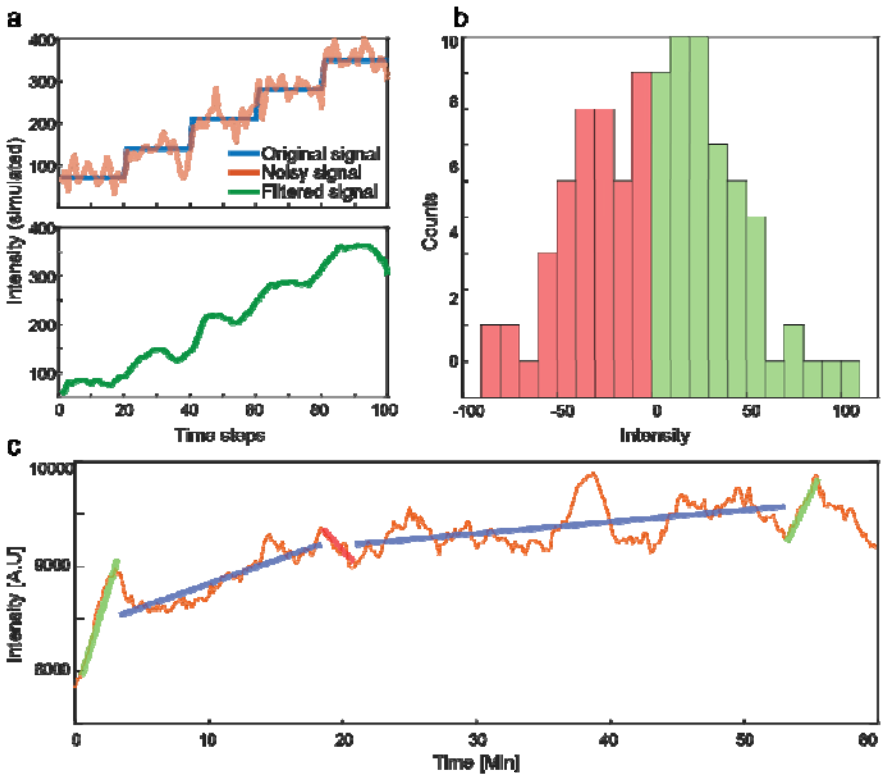
$$p_{inc}^{m} = \frac{1}{2^{10}} \Rightarrow m\log_2(p_{inc}) = -10 \Rightarrow m = \frac{-10}{\log_2(p_{inc})} \tag{1.8}$$

The threshold is calculated for each signal separately and is usually in the range of 7-13 time points. An analogous threshold is calculated for decrements in the signal and is typically in the range $[m-1, m+1]$.

To account for the presence of the occasional strong instantaneous noise fluctuations appearing in experimental signals, we allow isolated reversals in the signal directionality (e.g., an isolated one time point decrease in an otherwise continuous signal increase environment). Furthermore, since the moving average filter itself can induce correlations in the signal, we determined that the minimum allowed

threshold is the moving average window span. This means that any calculated threshold lower than the moving average size is increased to this bare minimum.

We mark each trace with the number of events whose duration exceeds the threshold and define those as bursts. Segments within the signal that are not classified as either a negative or positive burst event are considered unclassified. Unclassified segments are typically signal elements whose noise profile does not allow us to make a classification into one or the other event-type. For each identified segment we record the amplitude (ΔI), and duration (Δt). In Appendix 1 – figure 2c we mark the classifications on a sample trace with positive "burst", negative "burst", and non-classified events in green, red, and blue, respectively. We confine our segment analysis between the first and last significant segments identified in each signal, since we cannot correctly classify signal sections that extend beyond the observed trace.



Appendix 1 – figure 2: **Identification of burst events**. **a**, Top: simulated step signal (blue) with added white Gaussian noise (orange). Bottom: noisy signal after moving average filter. **b**, Intensity difference distribution for the signal presented in panel A. **c**, Sample experimental signal (orange) overlaid with markers indicating identified segments in green, blue, and red, corresponding to positive bursts, quiescent segments, and negative bursts.

Estimating the signal amount per slncRNA-RBP complex

Given the fact that we cannot directly infer the fluorescence intensity associated with a single RNA-RBP complex, we fitted the distributions with a modified Poisson function of the form:

$$p(I) = \frac{\lambda^{\frac{I}{k_0}} e^{-\lambda}}{\left(\frac{I}{k_0}\right)!} \tag{1.9}$$

where I is the experimental fluorescence amplitude, λ is the Poisson parameter (rate), and k_0 is a fitting parameter whose value corresponds to the amplitude associated with a single RBP-bound slncRNA molecule within the burst. For each rate we chose the fit to k_0 that minimizes the deviation (MSE) from the experimental data.

Appendix 2

Numerical simulations of signal types

To check that our analysis is consistent with an underlying random burst signal, we simulated three types of base signals with added noise components. For each simulation type, 1000 signals of 360 time-points were simulated and analyzed using the same data analysis process described in the methods section.

We simulated flat constant signals, gradually ascending signals, and signals containing multiple burst events. Two noise components were added to all signals, based on our noise model. White Gaussian noise of magnitude 40 [A.U] peak-to-peak amplitude, matching the value estimated from experimental traces, and an exponential component, simulating photobleaching (Appendix 2 – figure 1).

We then applied our burst-detection algorithm described above and found that for the flat signal (Appendix 2 – figure 1a) positive and negative bursts (green and red respectively) and non-classified events are detected. However, a closer examination of the results reveals that the burst amplitude width is smaller by a factor of ~5-10 as compared with the experimental data bursts, and the total number of events observed (458 positive, 452 negative, and 298 non-classified segments found) is significantly smaller than the experimental data, indicating roughly 1 event per signal, as expected from our base assumption that a rare noise event occurs once in a thousand time points. For the gradually increasing signal with additional noise, (Appendix 2 – figure 1c) a negligible number of negative burst-like events was detected by our algorithm, with a pronounced bias towards positive events (1111 positive, 9 negative and 467 non-classified). The scarcity of events can be explained by the positive bias in the signal which results in a steep increase in the statistical threshold for event identification. Similar simulations with a decreasing signal show a mirror image of amplitude distribution (data not shown).

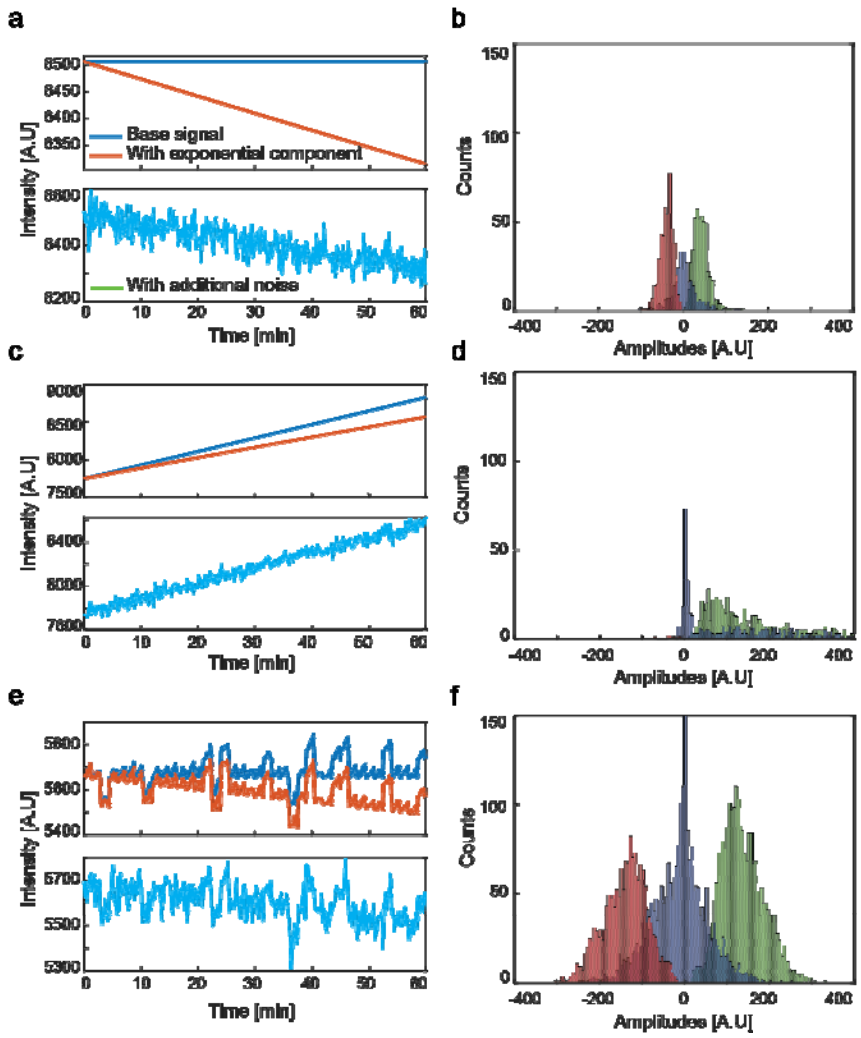
Finally, a signal designed to mimic our interpretation of the experimental data containing randomly distributed instantaneous bursts, both increasing and decreasing with multiple possible amplitudes was analyzed (Appendix 2 – figure 1e). Our simulated signals resulted in a symmetric amplitude distribution, comprising of non-Gaussian or skewed amplitude distributions. Additionally, the range of amplitudes observed is 2-3x larger as compared with the case for the constant signal, with the non-classified amplitudes presenting a wider distribution. A total of 2298 positive, 1831 negative and 2489 non-classified segments were found.

Estimating statistical significance of burst events in all traces recorded

To compute whether the number of burst events identified via our algorithm is statistically significant, we simulated a constant base-line intensity amplitude with overlaid white Gaussian noise. For each numerical trace, we simulated 360 times points (corresponding to a ~60-minute experimental trace) and identified the total number of "increasing" and "decreasing" burst events in accordance with the algorithm described in detailed above. Here, we used m = 10 (see eqn. 1.8) consecutive increasing or decreasing instantaneous signal difference events as our threshold. We identified 458 and 298 increasing and decreasing burst events respectively in 1000 simulated traces with constant baseline. By comparison, we found 2298 and 1831 increasing and decreasing burst events respectively in 1000 simulated traces containing bursts, which using Fisher's test yield a p-value of 4e-309 and 2e-310 for the significance of the increasing and decreasing burst findings.

We repeated this statistical test for experimental data, comparing the PP7-4x data against traces measured from cell containing only tdPP7-mCherry with no expression of our RNA cassettes, using the latter as a baseline akin to the constant signal simulations. We identified 7 increasing and 6 decreasing burst events in 150 traces gathered from the cells lacking RNA binding sites, while for the PP7-4x data we identified 112 increasing and decreasing burst events in 255 experimental traces, which using Fisher's test yields a p-value of 2e-13.

To verify that burst events that occur after a non-classified period lasting 2.5 minutes or longer are not biased, we performed a statistical test for randomness where the null hypothesis is that events are in random order. The tests yielded p-values of 0.7 for PP7-4x, 0.03 for Qβ-5x, 0.4 for Qβ-10x, and 0.5 for PP7-24x. Indicating that the burst events do appear at random at the 1% significance level.



Appendix 2 – figure 1: **Signal type simulations**. **a**, Simulated constant signal (blue), with photobleaching (orange), and added noise (cyan). **b**, Amplitude distributions of burst events identified from 1000 constant signals. **c**, Simulated signal with slope (blue), with photobleaching (orange), and added noise (cyan). **d**, Amplitude distributions of burst events identified from 1000 sloped signals. **e**, Simulated signal with burst events (blue), with photobleaching (orange), and added noise (cyan). **f**, Amplitude distributions of burst events identified from 1000 bursty signals.

Appendix 3
Signal analysis parameter selection

Subframe length

As part of the analysis process, the immediate surroundings of each discovered bright spot are recorded as a sub-frame containing the spot at its center, from this sub-frame the mean spot intensity and mean background intensity are calculated. The selection of the sub-frame length used to calculate the background intensity is an important parameter in the analysis process that might bring about unwanted noise into the resulting statistics. A large sub-frame might include other cells, with possibly different bright spots of themselves, inserting a bias into both the cell background intensity, and spot intensity signals. On the other hand, a small sub-frame might not have a sufficient spot-to-background area ratio, resulting in an underestimated cell background signal.

To select the appropriate sub-frame length, we analyzed the Qβ-10x data with sub-frames of different lengths – 10, 14, 20, and 30 pixels. Appendix 3 – figure 1a shows an example of this where the orange, yellow and red squares correspond to sub-frames of 10,14 and 20 pixels in length, and the panel itself constitutes a 30-pixel wide sub-frame. The criteria for this selection process are the mean ratio between cell area to spot area; percentage of frames where this ratio is less than one; and the ratio between the spot mean intensity to the cell mean intensity without any filtering or fitting. These criteria are designed to find the length that does not cause an overestimation of cell background against spot or vice versa (as could be the case where more than one bright spot fall inside the sub-frame). From these tests we learned that lengths of 10 and 14 pixels result in a mean ratio of less than two (i.e., on average the sizes of the bright spot and of its surrounding environment are equal) (Appendix 3 – figure 1b). However, a sub-frame length of 10 pixels results in nearly a fifth of frames where the cell background is less than one and thus potentially underestimated (Appendix 3 – figure 1c). Finally, the intensity ratios show that the mean ratio does not vary much between the different options, however the spread is more conserved for lengths of 10 and 14 pixels (Appendix 3 – figure 1d). Following these tests, we chose a sub-frame length of 14 pixels for our analysis process.

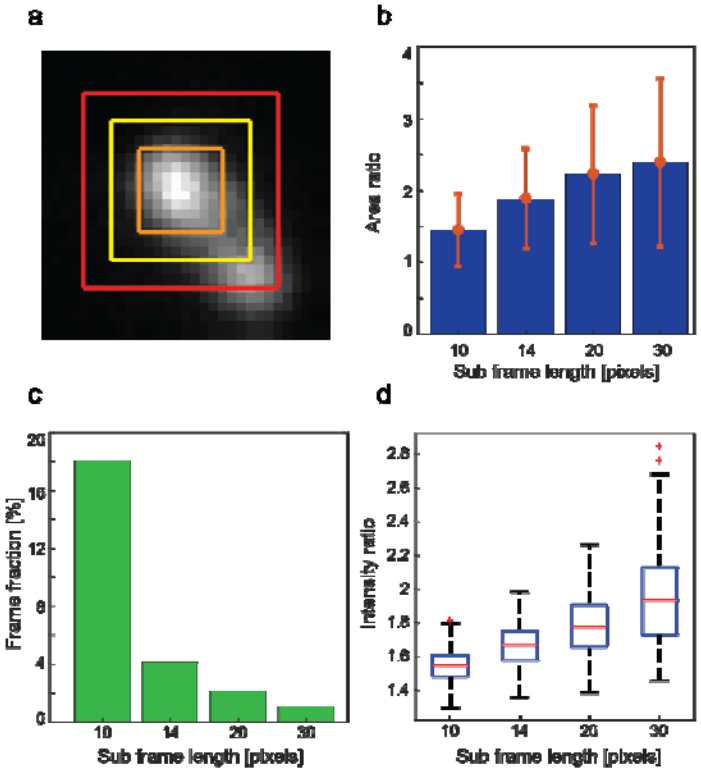
Moving average span

The moving average window span is a critical component in the signal analysis process. It is used both as a noise reduction filter, and as a means to bias sharp signal jumps (See materials and methods). The filter span plays another significant role, as it is the minimal allowed length for a burst duration. Choosing a small value might introduce false positives into the statistics, while a large value would cause many actual burst events to be discarded. To find the optimal span length we compare the number of

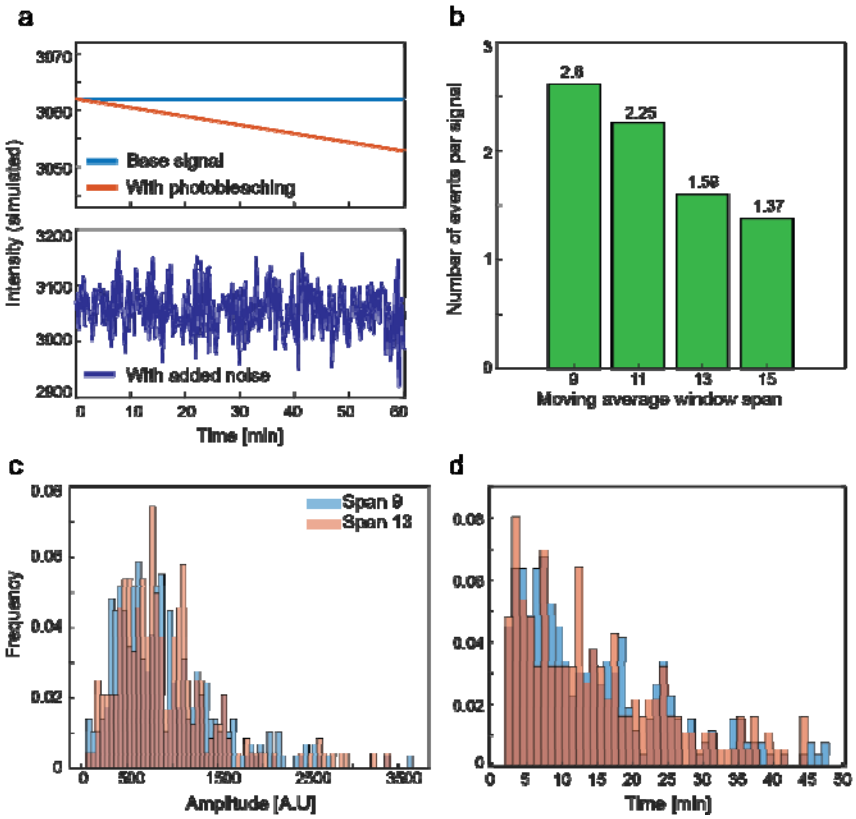
events found in a simulated flat signal, such a signal should not produce any bursts under noise-less conditions. For this we simulated 1000 constant signals, 360 time points each, with an added white Gaussian noise and an exponential component and applied our data analysis procedure. (Appendix 3 – figure 2a). An ideal result for this test would be less than one event of each type, i.e. positive and negative bursts, per signal (Appendix 3 – figure 2b).

We further show that using intermediate span length values (9-13 time points), has little effect on the qualitative nature of the results Appendix 3 – figure 2c, d).

Following these tests, we decided on a span of 13 time points. This value results in one event or less of each type per simulated signal, while still allowing us to record the statistical nature of the experimental signals.



Appendix 3 – figure 1: Subframe length selection. **a**, Example of different sub-frame lengths. Image is a sub-frame with length of 30 pixels. Red, yellow, and orange squares correspond to sub-frames of length 20, 14 and 10 pixels accordingly. **b**, Ratio between cell background area to spot area (both are in number of pixels). **c**, Percentage of cells where the area ratio presented in (b) is less than one, indicating probable underestimation of the cell background. **d**, Ratio between spot mean intensities to cell background mean intensities (i.e. each spot is divided by its corresponding cell background). Horizontal lines represent 25 and 75 percentiles.



Appendix 3 – figure 2: Moving average span length selection. **a**, Sample simulated signal used for testing. Blue line is the underlying constant signal, orange line represents the same signal with an added photobleaching component. Cyan signal is the orange signal with added white Gaussian noise. **b**, Total number of identified events of any kind per simulated signal. **c**, Positive amplitude histograms of Qβx10 data analyzed using a moving average filter of 9 time points (blue) and 13 time points (orange). **d**, Duration of quiescent segments of Qβx10 data analyzed using a moving average filter of 9 time points (blue) and 13 time points (orange).

Appendix 4

Theoretical Model

Liquid-liquid phase separation has been recently modelled by Klosin et al. (Klosin et al., 2020). In this section, we will expand the Klosin model to a case where the bacterial cell has initially a dense-nucleoid and dilute phase, and the RNA is transcribed within the nucleoid phase. If the RNA is sufficiently multivalent, a droplet forms within the dilute phase background. Our model will describe the rates by which RNA is transcribed, exchanged between the nucleoid and dilute phases, and at which conditions it will form an RNP granule within the dilute phase.

Thermodynamic Model Assumptions

We assume a cell containing two phases: a dense nucleoid phase and dilute cytosolic phase. The nucleoid phase fills ~75% of the cell volume and the dilute phase occupies mostly the cell pole regions. We then express a synthetic and multivalent long non-coding RNA molecule (slncRNA) containing multiple binding sites for an RNA-binding protein (RBP), which is co-expressed as a fusion with an mCherry fluorescent protein. Given these assumptions, we can now write a free energy as follows (an expansion of the Klosin free energy):

$$F = V_n f_n(\phi_n) + V_+ f_+(\phi_+) + V_- f_-(\phi_-) + \Gamma_n A_n + \Gamma_- A_- \tag{1.10}$$

Where, following Klosin's notations, V_n, V_+, V_- correspond to the volume of the nucleoid, dilute and droplet phases. Similarly, ϕ_n, ϕ_+, and ϕ_- correspond to the volume fractions of each phase, and f_n, f_+, f_- correspond to the free energy density of each phase. Γ_n and Γ_- are the nucleoid and droplet phase surface tensions with corresponding area A_n and A_-. In addition, we note that the total slncRNA-RBP complex present in the system at steady state is:

$$N_T = N_n + N_+ + N_- \tag{1.11}$$

Where N_n, N_+, and N_- correspond to the number of molecular complexes in the nucleoid, dilute, and droplet phases, respectively.

Kinetic Model

In the following, we will derive the kinetic model describing such a system according to the schematic presented in Figure. 8a. Here we assume that a single promoter located within the nucleoid phase encodes the slncRNA, which immediately leads to the formation of the slncRNA-RBP complex. Molecular complexes then diffuse around the nucleoid phase and are transported out of the nucleoid phase into the dilute cytosolic phase at a rate proportional to their diffusion coefficient times their volume fraction defined according to the Klosin model as follows:

$$k_n^{out} = \frac{6 D_n V_n^3}{\upsilon} \tag{1.12}$$

Where υ corresponds to the unit volume, i.e., the volume of a single molecule, and D_n corresponds to the diffusion constant within the nucleoid phase, which is assumed to be different than the one in the

cytosolic or dilute phase. Note, this is due to the Stokes-Einstein equation which to a first approximation defines the diffusion coefficient as:

$$D = \frac{k_B T}{6\pi\eta r} \tag{1.13}$$

Where η , the dynamic viscosity, is expected to vary for the dilute and nucleoid phases.

Given the above definitions, the goal of the model analyzed below is to estimate the rate of increasing signal bursts, which corresponds to k_+^{out} in the schematic of Figure. 4a.

<u>Evaluating the model</u>

To evaluate the model, we assume that the all three liquid phases are permeable and allow exchange of slncRNA-RBP molecular complexes. This implies that each phase can be modeled as a state within a Master equation context, with rates controlling the transition between each state. Given this assumption, we can now write a Master equation model for the kinetics of this multiphasic system in accordance with the schematic shown in Figure. 8a.

$$\begin{pmatrix} \partial_t p_n(N) \\ \partial_t p_+(N) \\ \partial_t p_-(N) \end{pmatrix} = \begin{pmatrix} -(N\gamma_n + k_t + k_n^{out})p_n(N) + k_n^{in}P_+(N) + k_t p_n(N-1) + (N+1)\gamma_n P_n(N+1) \\ -(N\gamma_+ + k_n^{in} + k_+^{out})p_+(N) + k_n^{out}p_n(N) + k_+^{in}p_-(N) + (N+1)\gamma_+ p_+(N+1) \\ -(N\gamma_- + k_+^{in})p_-(N) + k_+^{out}p_+(N) + (N+1)\gamma_- p_-(N+1) \end{pmatrix} \tag{1.14}$$

Which can be written in vector form as follows:

$$\frac{d}{dt}\vec{p}(N) = [\hat{K} - \hat{R} - N\hat{\Gamma}]\vec{p}(N) + \hat{R}\vec{p}(N-1) + (N+1)\hat{\Gamma}\vec{p}(N+1) \tag{1.15}$$

Where,

$$\hat{K} = \begin{pmatrix} -k_n^{out} & k_n^{in} & 0 \\ k_n^{out} & -(k_n^{in} + k_+^{out}) & k_+^{in} \\ 0 & k_+^{out} & -k_+^{in} \end{pmatrix}, \hat{R} = \begin{pmatrix} k_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \hat{\Gamma} = \begin{pmatrix} \gamma_n & 0 & 0 \\ 0 & \gamma_+ & 0 \\ 0 & 0 & \gamma_- \end{pmatrix} \tag{1.16}$$

In order to determine k_+^{out} , we evaluate the zeroth moment of the master equation as follows:

$$\overrightarrow{M_0} = \begin{pmatrix} M_0^n \\ M_0^+ \\ M_0^- \end{pmatrix} = \begin{pmatrix} \sum_{N=1}^{\infty} p_n(N) \\ \sum_{N=1}^{\infty} p_+(N) \\ \sum_{N=1}^{\infty} p_-(N) \end{pmatrix} \tag{1.17}$$

with the following condition:

$$\vec{u} \bullet \overrightarrow{M_0} = 1 \tag{1.18}$$

ensuring that the total probability for the kinetic system to be in one of the states adds up to 1.

Next, we evaluate the zeroth moment in steady state, which allows us to use the following assumptions:

$$\begin{aligned} k^t &= 0 \\ \gamma_n &= \gamma_+ = \gamma_- = 0 \\ k_n^{in} &= k_+^{out} \end{aligned} \tag{1.19}$$

Where the last equation implies that the rate of exit from the dilute phase is the same, regardless of direction.
Plugging these to the following equation:

$$0 = [\overline{K} - N\hat{\Gamma}]\overrightarrow{M_0} + [\overline{R} + N\hat{\Gamma}]\overrightarrow{M_0} \tag{1.20}$$

$$\overline{K}\overrightarrow{M_0} = 0 \tag{1.21}$$

This then allows us to write the following set of equations:

$$\begin{aligned} -k_n^{out} M_0^n + k_n^{in} M_0^+ &= 0 \\ k_n^{out} M_0^n - (k_n^{in} + k_+^{out}) M_0^+ + k_+^{in} M_0^- &= 0 \\ k_+^{out} M_0^+ - k_+^{in} M_0^- &= 0 \\ M_0^n + M_0^+ + M_0^- &= 1 \end{aligned} \tag{1.22}$$

which allows us to solve for k_+^{out} as follows:

$$M_0^+ = \frac{k_+^{in}}{k_+^{out}} M_0^- = \frac{k_n^{out}}{k_n^{in}} M_0^n \tag{1.23}$$

Plugging in our third assumption:

$$k_+^{in} = k_n^{out} \frac{M_0^n}{M_0^-} \tag{1.24}$$

$$k_{+}^{out} = k_{n}^{out}\left(\frac{M_0^n}{M_0^+}\right) = \frac{6D_n V_n^{1/3}}{\upsilon}\phi_n\left(\frac{M_0^n}{M_0^+}\right) \tag{1.25}$$

Showing that the burst of signal increases should occur at a rate that is proportional to the complex's volume fraction within the nucleoid phase.

Implication of bi-phasic cellular model to transcription

Given the bi-phasic model, we wanted to compute the Fano factor for a general mRNA that does not necessarily phase separate in the dilute cytosol phase to a third droplet phase. In this case equation 2.7 is simplified as follows:

$$\hat{K} = \begin{pmatrix} -k_n^{out} & k_n^{in} \\ k_n^{out} & -k_n^{in} \end{pmatrix},\ \hat{R} = \begin{pmatrix} k_t & 0 \\ 0 & 0 \end{pmatrix},\ \hat{\Gamma} = \begin{pmatrix} \gamma_n & 0 \\ 0 & \gamma_+ \end{pmatrix} \tag{1.26}$$

Here, we assume that each phase is characterized by a different degradation rate. Degradation is a process by which an RNAase is assumed to diffuse around until it finds its target. If we accept the assumption that each phase is characterized by a different diffusion coefficient, then the rate of degradation should also vary in accordance. However, for the sake of simplicity, we will assume a constant degradation rate across the cell, and thus we get:

$$\hat{K} = \begin{pmatrix} -k_n^{out} & k_n^{in} \\ k_n^{out} & -k_n^{in} \end{pmatrix},\ \hat{R} = \begin{pmatrix} k_t & 0 \\ 0 & 0 \end{pmatrix},\ \hat{\Gamma} = \begin{pmatrix} \gamma & 0 \\ 0 & \gamma \end{pmatrix} \tag{1.27}$$

Evaluating the zeroth moment

In this case, the zeroth moment is defined as follows:

$$\overrightarrow{M_0} = \begin{pmatrix} M_0^n \\ M_0^+ \end{pmatrix} \equiv \begin{pmatrix} \sum_{N=1}^{\infty} p_n(N) \\ \sum_{N=1}^{\infty} p_+(N) \end{pmatrix} \tag{1.28}$$

leading to the following equations

$$\begin{gathered} -k_n^{out} M_0^n + k_n^{in} M_0^+ = 0 \\ M_0^n + M_0^+ = 1 \end{gathered} \tag{1.29}$$

which allows us to solve for the different components:

$$M_0^+ = \frac{k_n^{out}}{k_n^{out} + k_n^{in}}$$
$$M_0^n = \frac{k_n^{in}}{k_n^{out} + k_n^{in}} \tag{1.30}$$

Evaluating the first moment

The first moment is defined as follows:

$$\overrightarrow{M_1} = \begin{pmatrix} M_1^n \\ M_1^+ \end{pmatrix} \equiv \begin{pmatrix} \sum_{N=1}^{\infty} N p_n(N) \\ \sum_{N=1}^{\infty} N p_+(N) \end{pmatrix} \tag{1.31}$$

from which we can calculate the mean number of molecules per cell as follows:

$$\langle N \rangle = \vec{u} \bullet \overrightarrow{M_1} = M_1^n + M_1^+ \tag{1.32}$$

Next, we evaluate the Master equation for the first moment in steady state as follows:

$$0 = (\hat{K} - \hat{\Gamma} + \hat{R})\overrightarrow{M_1} + \hat{R}\overrightarrow{M_0} \tag{1.33}$$

To obtain an expression for the mean, we multiply equation 20 by the unitary vector to obtain:

$$\bar{k} = \frac{k_t k_n^{in}}{k_n^{out} + k_n^{in}} = \vec{u} \bullet \hat{\Gamma} \bullet \overrightarrow{M_1} = \gamma \langle N \rangle \tag{1.34}$$

Evaluating the second moment and the Fano factor

The second moment is defined as follows:

$$\overrightarrow{M_2} = \begin{pmatrix} M_2^n \\ M_2^+ \end{pmatrix} \equiv \begin{pmatrix} \sum_{N=1}^{\infty} N^2 p_n(N) \\ \sum_{N=1}^{\infty} N^2 p_+(N) \end{pmatrix} \tag{1.35}$$

where,

$$\vec{u} \bullet \overrightarrow{M_2} = M_2^n + M_2^+ = \langle N^2 \rangle \tag{1.36}$$

Using Sanchez et al. (Sanchez et al., 2011), in steady state we get the following matrix equation:

$$0 = 2\vec{u} \bullet \overline{R} \bullet \overrightarrow{M_1} + \vec{u} \bullet \overline{R} \bullet \overrightarrow{M_0} - 2\vec{u} \bullet \hat{\Gamma} \bullet \overrightarrow{M_2} + \vec{u} \bullet \hat{\Gamma} \bullet \overrightarrow{M_1} \tag{1.37}$$

which reduces to:

$$\vec{u} \bullet \hat{\Gamma} \bullet \overrightarrow{M_2} = \vec{u} \bullet \overline{R} \bullet \overrightarrow{M_1} + \overline{k} \tag{1.38}$$

$$\left\langle N^2 \right\rangle = \frac{\vec{u} \bullet \overline{R} \bullet \overrightarrow{M_1}}{\gamma} + \left\langle N \right\rangle \tag{1.39}$$

This then allows us to define a Fano factor as follows:

$$F_n = \frac{\left\langle N^2 \right\rangle - \left\langle N \right\rangle^2}{\left\langle N \right\rangle} = 1 + \frac{1}{\left\langle N \right\rangle}\left(\frac{\vec{u} \bullet \overline{R} \bullet \overrightarrow{M_1}}{\gamma} - \left\langle N \right\rangle^2 \right) \tag{1.40}$$

which after further evaluation reduces to (see also Sanchez et al. – eq. 19):

$$F_n = 1 + \left\langle N \right\rangle \left(\frac{k_n^{out}}{k_n^{in}} \right)\left(\frac{\gamma}{\gamma + k_n^{out} + k_n^{in}} \right) = 1 + \left(\frac{k_t k_n^{out}}{\gamma + k_n^{out} + k_n^{in}} \right) \tag{1.41}$$

Which is a signature of a super-Poisson distribution as was observed experimentally in bacteria. Therefore, even if we assume nothing additional about the standard biological dogma, having two phases which exchange molecules between them is sufficient for generating the deviation from Poisson behavior that was previously attributed to transcriptional bursting. As a result, if we accept the experimental evidence for the existence of these two phases, the super-Poisson distributions of mRNA that was previously observed is an immediate consequence of this physical state. Note, this result is consistent with the model shown in Klosin et al. (Klosin et al., 2020), where the total noise strength behaves in a Super-Poisson fashion.

References and Notes:

Banani SF, Lee HO, Hyman AA, Rosen MK. 2017. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**:285–298. doi:10.1038/nrm.2017.7

Begovich K, Wilhelm JE. 2020. An In Vitro Assembly System Identifies Roles for RNA Nucleation and ATP in Yeast Stress Granule Formation. *Molecular Cell* **79**:991-1007.e4. doi:10.1016/j.molcel.2020.07.017

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. 1998. Localization of ASH1 mRNA particles in living yeast. *Molecular cell* **2**:437–445.

Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, Tompa P, Fuxreiter M. 2018. Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**:420–435. doi:10.1016/j.tcb.2018.02.004

Bond CS, Fox AH. 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. *J Cell Biol* **186**:637–644. doi:10.1083/jcb.200906113

Chao JA, Patskovsky Y, Almo SC, Singer RH. 2008. Structural basis for the coevolution of a viral RNA–protein complex. *Nature Structural & Molecular Biology* **15**:103–105. doi:10.1038/nsmb1327

Chong S, Chen C, Ge H, Xie XS. 2014. Mechanism of Transcriptional Bursting in Bacteria. *Cell* **158**:314–326. doi:10.1016/j.cell.2014.05.038

Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33**:717–726. doi:10.1016/j.molcel.2009.01.026

Feng Z, Chen X, Wu X, Zhang M. 2019. Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications. *J Biol Chem* **294**:14823–14835. doi:10.1074/jbc.REV119.007895

Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, Kleckner N. 2013. Four-Dimensional Imaging of E. coli Nucleoid Organization and Dynamics in Living Cells. *Cell* **153**:882–895. doi:10.1016/j.cell.2013.04.006

Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard J-M, Singer RH, Bertrand E. 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Current Biology* **13**:161–167.

Golding I, Paulsson J, Zawilski SM, Cox EC. 2005. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **123**:1025–1036. doi:10.1016/j.cell.2005.09.031

Hocine S, Raymond P, Zenklusen D, Chao JA, Singer RH. 2013. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nature Methods* **10**:119–121. doi:10.1038/nmeth.2305

Hyman AA, Weber CA, Jülicher F. 2014. Liquid-Liquid Phase Separation in Biology. *Annual Review of Cell and Developmental Biology* **30**:39–58. doi:10.1146/annurev-cellbio-100913-013325

Janissen R, Arens MMA, Vtyurina NN, Rivai Z, Sunday ND, Eslami-Mossallam B, Gritsenko AA, Laan L, de Ridder D, Artsimovitch I, Dekker NH, Abbondanzieri EA, Meyer AS. 2018. Global DNA Compaction in Stationary-Phase Bacteria Does Not Affect Transcription. *Cell* **174**:1188-1199.e14. doi:10.1016/j.cell.2018.06.049

Jones D, Elf J. 2018. Bursting onto the scene? Exploring stochastic mRNA production in bacteria. *Current Opinion in Microbiology* **45**:124–130. doi:10.1016/j.mib.2018.04.001

Kannaiah S, Livny J, Amster-Choder O. 2019. Spatiotemporal Organization of the E. coli Transcriptome: Translation Independence and Engagement in Regulation. *Molecular Cell* **76**:574-589.e7. doi:10.1016/j.molcel.2019.08.013

Katz N, Cohen R, Solomon O, Kaufmann B, Atar O, Yakhini Z, Goldberg S, Amit R. 2019. Synthetic 5′ UTRs Can Either Up- or Downregulate Expression upon RNA-Binding Protein Binding. *Cell Systems* **9**:93-106.e8. doi:10.1016/j.cels.2019.04.007

Katz N, Cohen R, Solomon O, Kaufmann B, Atar O, Yakhini Z, Goldberg S, Amit R. 2018. An in Vivo Binding Assay for RNA-Binding Proteins Based on Repression of a Reporter Gene. *ACS Synth Biol* **7**:2765–2774. doi:10.1021/acssynbio.8b00378

Katz N, Tripto E, Granik N, Goldberg S, Atar O, Yakhini Z, Orenstein Y, Amit R. 2021. Overcoming the design, build, test (DBT) bottleneck for synthesis of nonrepetitive protein-RNA binding cassettes for RNA applications. *Nature Communications*.

Kim J, Yoshimura SH, Hizume K, Ohniwa RL, Ishihama A, Takeyasu K. 2004. Fundamental structural units of the Escherichia coli nucleoid revealed by atomic force microscopy. *Nucleic Acids Res* **32**:1982–1992. doi:10.1093/nar/gkh512

Klosin A, Oltsch F, Harmon T, Honigmann A, Jülicher F, Hyman AA, Zechner C. 2020. Phase separation provides a mechanism to reduce noise in cells. *Science* **367**:464–468. doi:10.1126/science.aav6691

Langdon EM, Gladfelter AS. 2018. A New Lens for RNA Localization: Liquid-Liquid Phase Separation. *Annual Review of Microbiology* **72**:255–271. doi:10.1146/annurev-micro-090817-062814

Larson DR, Singer RH, Zenklusen D. 2009. A Single Molecule View of Gene Expression. *Trends Cell Biol* **19**:630–637. doi:10.1016/j.tcb.2009.08.008

Lee C-YS, Putnam A, Lu T, He S, Ouyang JPT, Seydoux G. 2020. Recruitment of mRNAs to P granules by condensation with intrinsically-disordered proteins. *eLife* **9**:e52896. doi:10.7554/eLife.52896

Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF, Russo PS, Jiang Q-X, Nixon BT, Rosen MK. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**:336–340. doi:10.1038/nature10879

Lionnet T, Singer RH. 2012. Transcription goes digital. *EMBO Rep* **13**:313–321. doi:10.1038/embor.2012.31

Mao YS, Zhang B, Spector DL. 2011. Biogenesis and function of nuclear bodies. *Trends Genet* **27**:295–306. doi:10.1016/j.tig.2011.05.006

Medina G, Juárez K, Valderrama B, Soberón-Chávez G. 2003. Mechanism of Pseudomonas aeruginosa RhlR Transcriptional Regulation of the rhlAB Promoter. *Journal of Bacteriology* **185**:5976–5983. doi:10.1128/JB.185.20.5976-5983.2003

Otsu N. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* **9**:62–66. doi:10.1109/TSMC.1979.4310076

Paulsson J. 2005. Models of stochastic gene expression. *Physics of Life Reviews* **2**:157–175. doi:10.1016/j.plrev.2005.03.003

Polymenidou M. 2018. The RNA face of phase separation. *Science* **360**:859–860. doi:10.1126/science.aat8028

Putnam A, Seydoux G. 2021. Chapter Four - Cell-free reconstitution of multi-condensate assemblies In: Keating CD, editor. Methods in Enzymology, Liquid-Liquid Phase Coexistence and Membraneless Organelles. Academic Press. pp. 83–113. doi:10.1016/bs.mie.2020.07.004

Rhine K, Vidaurre V, Myong S. 2020. RNA Droplets. *Annual Review of Biophysics* **49**:247–265. doi:10.1146/annurev-biophys-052118-115508

Sanchez A, Garcia HG, Jones D, Phillips R, Kondev J. 2011. Effect of Promoter Architecture on the Cell-to-Cell Variability in Gene Expression. *PLoS Computational Biology* **7**:e1001100. doi:10.1371/journal.pcbi.1001100

Sbalzarini IF, Koumoutsakos P. 2005. Feature point tracking and trajectory analysis for video imaging in cell biology. *J Struct Biol* **151**:182–195. doi:10.1016/j.jsb.2005.06.002

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**:676–682. doi:10.1038/nmeth.2019

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**:671–675. doi:10.1038/nmeth.2089

So L, Ghosh A, Zong C, Sepúlveda LA, Segev R, Golding I. 2011. General properties of transcriptional time series in Escherichia coli. *Nature Genetics* **43**:554–560. doi:10.1038/ng.821

Tian S, Curnutte HA, Trcek T. 2020. RNA Granules: A View from the RNA Perspective. *Molecules* **25**:3130. doi:10.3390/molecules25143130

Treeck BV, Protter DSW, Matheny T, Khong A, Link CD, Parker R. 2018. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *PNAS* **115**:2734–2739. doi:10.1073/pnas.1800038115

Tutucci E, Vera M, Biswas J, Garcia J, Parker R, Singer RH. 2018. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nature Methods* **15**:81–89. doi:10.1038/nmeth.4502

Wang Z, Zhang G, Zhang H. 2019. Protocol for analyzing protein liquid–liquid phase separation. *Biophys Rep* **5**:1–9. doi:10.1007/s41048-018-0078-7

Wheeler JR, Matheny T, Jain S, Abrisch R, Parker R. 2016. Distinct stages in stress granule assembly and disassembly. *eLife* **5**:e18413. doi:10.7554/eLife.18413

Wolf SG, Frenkiel D, Arad T, Finkel SE, Kolter R, Minsky A. 1999. DNA protection by stress-induced biocrystallization. *Nature* **400**:83–85. doi:10.1038/21918

Yagasaki T, Matsumoto M, Tanaka H. 2019. Liquid-liquid separation of aqueous solutions: A molecular dynamics study. *J Chem Phys* **150**:214506. doi:10.1063/1.5096429

Acknowledgments: The authors would like to acknowledge the Technion's LS&E staff (Tal Katz-Ezov and Anastasia Diviatis) for help with sequencing and some imaging experiments. In addition the authors would like to thank professor Yoav Schectman from the Department of Biomedical Engineering in the Technion for help with image and signal analysis.

Funding: This project received funding from the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (Grant No. 152/11); Marie Curie Reintegration Grant No. PCIG11-GA-2012-321675; and from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement 664918 (MRG-Grammar) and 851065 (CARBP).

Author contributions: NG designed and carried out the microscopy experiments and analysis for all data. NK designed and synthesized the PP7-4x-Qβ-5x and Qβ-10x cassettes. SG assisted and guided the experiments and image analysis. RA supervised the study. NG and RA wrote the manuscript.

Competing interests: The authors declare no competing interests.

Data and materials availability: All datasets and original code used in this paper are available from: https://github.com/naorgk/slncRNA_Analysis. All bacterial plasmids constructed for this work are available upon request.