Multi-condensate state as a functional strategy to optimize the cell signaling output

Aniruddha Chattaraj, Eugene I Shakhnovich

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

Correspondence: shakhnovich@chemistry.harvard.edu

Abstract

The existence of multiple biomolecular condensates inside living cells is a peculiar phenomenon not compatible with the predictions of equilibrium statistical mechanics. In this work, we address the problem of multiple condensates state (MCS) from a functional perspective. We combined Langevin dynamics, reaction-diffusion simulation, and dynamical systems theory to demonstrate that MCS can indeed be a function optimization strategy. Using Arp2/3 mediated actin nucleation pathway as an example, we show that actin polymerization is maximum at an optimal number of condensates. For a fixed amount of Arp2/3, MCS produces a greater response compared to its single condensate counterpart. Our analysis reveals the functional significance of the condensate size distribution which can be mapped to the recent experimental findings. Given the spatial heterogeneity within condensates and non-linear nature of intracellular networks, we envision MCS to be a generic functional solution, so that structures of network motifs may have evolved to accommodate such configurations.

Introduction

Spatiotemporal control of biochemical events enables living cells to respond to changes in the environment and other perturbations. orchestration of cellular biochemistry in time and space is achieved via multiple mechanisms. A key aspect of this organization is the presence of distinct subcellular compartments, such as the nucleus, mitochondria, and Golgi body, each dedicated to specific functions (1). For example, the nucleus is responsible for storing DNA and facilitating mRNA production, while the mitochondria act as the cellular powerhouse, regulating energy metabolism. This task-specific compartmentalization of subcellular space is arguably the most significant step in the evolution of complex lifeforms (1).

Over the last decade, another class of intracellular structures, known as biomolecular condensates (2) that lack a surrounding membrane, has emerged. Clustering of weakly interacting multivalent proteins and nucleic acids underlies the condensate formation (3-7). These proteins have multiple interaction “sites” or “domains” which can form transient physical crosslinks (“bonds”) with their partner molecules. These interaction motifs are usually referred to as “stickers” which are interspersed with “spacer” regions (6, 8). The extent of crosslinking increases with concentration and beyond a system specific threshold concentration, large multi-molecular clusters emerge as phase separated droplets. Theories from polymer science greatly shape our understanding of the underlying physics of biological condensates (8-10). However, one phenomenon remains enigmatic. Classical theory predicts that, after phase transition, the system should form one large droplet coexisting with the soluble phase. But, till now, the majority of cellular and in-vitro experiments revealed a multi-condensate state (MCS) where the liquid droplets dynamically exchange components and never fully coalesce to become one large droplet within the experimental timespan.

We have previously proposed that the MCS is a dynamically arrested metastable state (7) where the interplay of timescales of diffusion and internal relaxation within emerging clusters determine the metastability of condensates and their size distribution. In the current work, we approach the problem of MCS from a functional perspective. We ask whether a state with multiple droplets is functionally more useful than its single droplet counterpart? To answer that question, we consider an experimentally well characterized system of Arp2/3 mediated actin nucleation pathway (4, 11-13).

Condensate forming biomolecules can be broadly classified into two classes – “scaffolds” and “clients” (14). Scaffolds are necessary core components that control the size and composition of

the condensates while clients are selectively recruited to these condensates, presumably serving to establish microenvironments that foster specific biochemical reactions. Furthermore, an emerging paradigm in the field underscores the presence of spatial heterogeneity within these condensates, with distinct molecular species occupying discrete regions - a structural hallmark observed in mesoscopic bodies such as the Nucleolus (15), P-granules (16, 17), Stress granules (18), Paraspeckles (19), and Nanog condensates (20). However, it is not yet clear whether such spatial assembly has any functional relevance. The topological relationship between molecular structures (sequence of sticker and spacer, for example) and the mesoscopic body (spatially heterogenous condensates) they assemble into is an intriguing evolutionary question.

In this study, we consider a minimal-component system that is computationally tractable and experimentally well studied. Our system consists of four upstream components (Nephrin, Nck, NWASP and Arp2/3, Figure 1A) of the actin nucleation pathway (13, 21-24). Polymerization of the cytoskeletal protein actin is an important process that cells utilize to respond to their environment. The dynamic equilibrium between monomeric (Globular, G) and polymeric (Filament, F) actin is regulated by a wide variety of proteins. Arp2/3 complex is a nucleation promoting factor (25) that promotes formation of the branched F-actin network. In resting cells, Arp2/3 remains in an “off” state. Given an extracellular signal, membrane-bound receptor, Nephrin forms cluster and bind to an adapter protein, Nck. Nck recruits NWASP which, in turn, activates the Arp2/3 complex and initiates the F-actin branching. The active Arp2/3 sits on the side of an existing filament (“mother filament”) and creates a branched filament (“daughter filament”) along a 70° angular direction. The extent of branched F-actin network connects the cell’s ability to effectively respond to the external signal. In a series of experiments (11, 12), Rosen and colleagues showed that Nephrin, Nck and NWASP form condensates, both in test-tube and cells. These condensates recruit Arp2/3 and promote local production of F-actin. Therefore, we consider modelling this system as a case study to unravel the functional picture of biological condensates.

Computational models provide useful quantitative and mechanistic insights that, otherwise, may not be accessible via experiments. However, these models are limited to a specific spatiotemporal scale, although any biological problem inherently encompasses multiple scales. To overcome this limitation, we employed a multi-scale simulation strategy. First, we used Langevin Dynamics simulation to probe the structural organization of the condensates. Informed by these finer-resolution simulations, we created a reaction diffusion model to couple the effect

of condensate geometry with protein diffusion. Having the output from these spatial simulations, we built an analytical framework to probe the effect of condensate number on the extent of F-actin production.

Our simulations revealed that, for a fixed amount of Arp2/3, distributing them onto the surfaces of multiple small condensates is a better strategy compared to amassing all the Arp2/3 into a single large droplet. This makes the MCS a desirable solution from the evolutionary perspective. We provide rationale on how similar mechanisms may be operative in other biological condensates to make it a universal function-optimization strategy.

Results

1. Arp2/3 dwells on the surface of hemi-spherical condensates

To probe the structural organization of the condensate, we first consider a simplified bead-spring model of the four-component system (Fig. 1A). Transmembrane receptor protein Nephrin has three phospho-tyrosine (pY) sites which interact with the SH2 domain of protein Nck. Nck also contains three SH3 domains which interact with the proline rich motifs or PRMs that belong to the protein NWASP. The terminal VCA domain of NWASP interacts with one of the two sites on Arp2/3 complex. These binding domains are interspersed with linker beads to confer flexibility. To mimic the receptor diffusion on a membrane, Nephrin is anchored to the XY plane (Fig. 1B). We start the simulation with a uniform distribution of molecules. Nephrins are distributed on the XY surface (cell membrane), while Nck, NWASP and Arp2/3 are uniformly placed inside the simulation volume (cytosol).

We first perform metadynamics simulation to assess the free energy profile of molecular clustering. Since we are interested in how all the proteins assemble near the membrane, we consider a distance metric, as the order parameter, between the membrane bound receptors and freely diffusing proteins in the cytosol. As a consequence of the multivalent clustering, all the proteins condense near the plasma membrane to create a hemi-spherical cluster (Fig. 1B). Fig. S1 shows that the free energy minimum is associated with the clustered state. When the order parameter converges to a steady state (Fig. S1A), we run standard Langevin Dynamics (without bias) to quantify the cluster topology and shape.

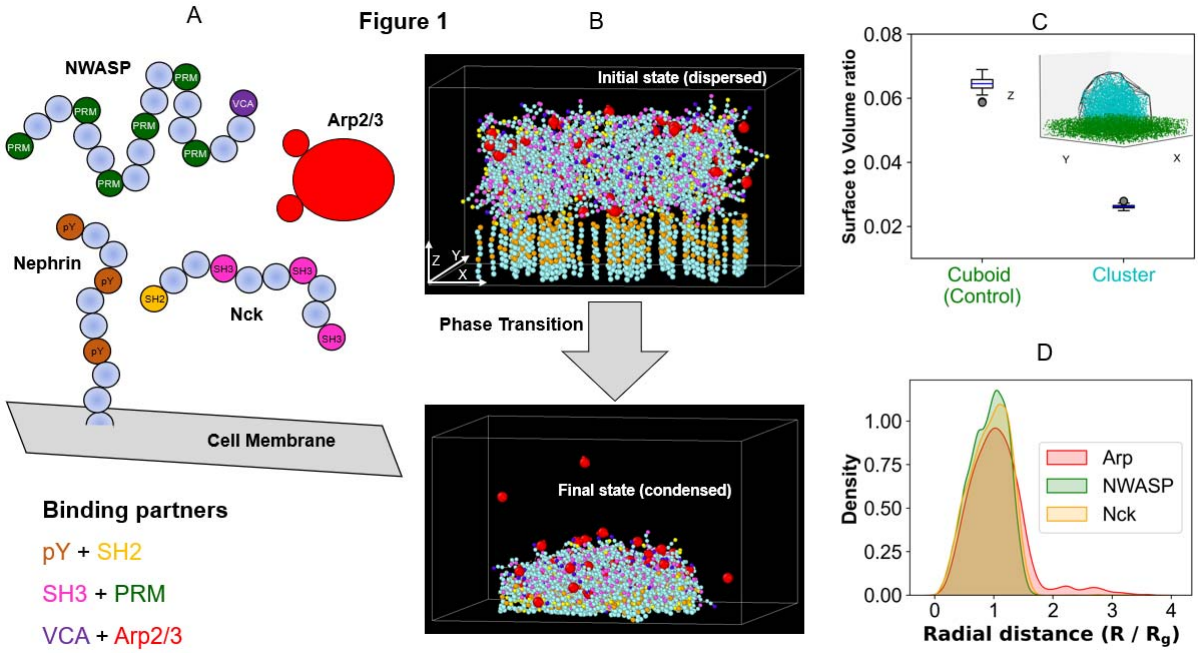

Figure 1: Setup of the Langevin dynamics simulations. (A) Coarse grained representations of the molecular components. Each protein is modelled as a bead-spring polymeric chain. Some beads serve as binding sites while others constituent the linker regions. Membrane bound receptor Nephrin has three phosphotyrosine (pY) sites which could recruit the protein Nck by binding to SH2 domains. Nck also has three SH3 domains which interact with the proline-rich motifs (PRMs) of the protein NWASP. VCA is the terminal domain of NWASP which can bind to one of two domains in the Arp2/3 complex. For Nephrin, Nck and NWASP, light blue beads represent the linker domains. (B) Depiction of the simulation geometry. XY surface acts as the membrane which anchors 80 randomly distributed Nephrin molecules. 240 Nck, 120 NWASP and 60 Arp2/3 molecules are uniformly placed inside the 3D volume. As the consequence of multivalent clustering, the initially dispersed molecules condense into a large cluster. (C) Surface to volume ratio of the membrane-associated cluster (cyan). A hypothetical cuboid (quasi-2D laminar assembly) arrangement is shown (green) for comparison. (D) Spatial location of different molecular types along the cluster radius. Distance (R) of each molecular centroid from the cluster's center of mass is calculated and normalized by the radius of gyration (Rg) of the cluster. 5 stochastic trials are used to compute the distribution in C and D.

Firstly, we compute the volume and surface area of the cluster. We fit all the particle coordinates to a convex hull and compute the surface-to-volume ratio of that topological object (Fig. 1C). We compare our cluster with a quasi-2D laminar assembly or cuboid (green points, Fig. 1C inset) having same volume as the cluster. The surface-to-volume ratio of the cluster is significantly lower than the cuboid reference (Fig. 1C), suggesting a more spherical geometry. Since one side of the cluster is attached to the 2D surface, we conclude that the cluster is hemi-spherical in shape.

Next, we determine the relative locations of different molecular types inside the cluster. To do that, we compute the radial distance (R) between the molecular centroid and the cluster centroid. We then normalize the distance by the radius of gyration (Rg) of the cluster. When $\mathrm{R}/Rg > 1$, the molecule is approximately placed near the periphery. Fig. 1D shows that only Arp2/3 has a long tail along the normalized radial distance, suggesting a peripheral location. Hence, the cluster made of Nephrin, Nck, NWASP and Arp2/3 is hemi-spherical, and Arp2/3 diffuses near the cluster surface.

2. Diffusion mediated local production of F-actin

We next seek to understand the effect of molecular diffusion on actin nucleation. We set up a reaction diffusion model where Arp2/3 diffuses at the surface of a hemi-spherical cluster placed on the XY plane. We include the reversible interconversion between GA (G-actin) and FA (F-actin), as shown in Fig. 2A. Actin in polymeric form has ~10 times slower diffusion than the monomeric form (26). Fig. 2B shows a local production of FA at the cost of local depletion of GA. Since Arp2/3 only diffuses in the vicinity of the cluster and FA has a slower diffusion, we observe an accumulation of FA around the cluster. Upon quantifying the FA concentration near the cluster surface, we see a significant difference between the bulk and the surface (Fig. 2C). It is to be noted that the bulk concentration incorporates all the volume elements including the cluster surface. Interestingly, the surface enrichment of FA depends on the diffusion of FA itself (Fig. 2C). When we double the FA diffusion, the concentration difference of FA between surface and bulk goes down. To probe this trend further, we gradually decrease the FA diffusion, keeping GA diffusion constant. Fig. 2D reveals a pattern of FA surface enrichment as a function of differential diffusion of FA. With slower FA diffusion, local accumulation of FA becomes more prominent. Hence, when Arp2/3 localizes in the cluster and catalyzes the GA to FA conversion,

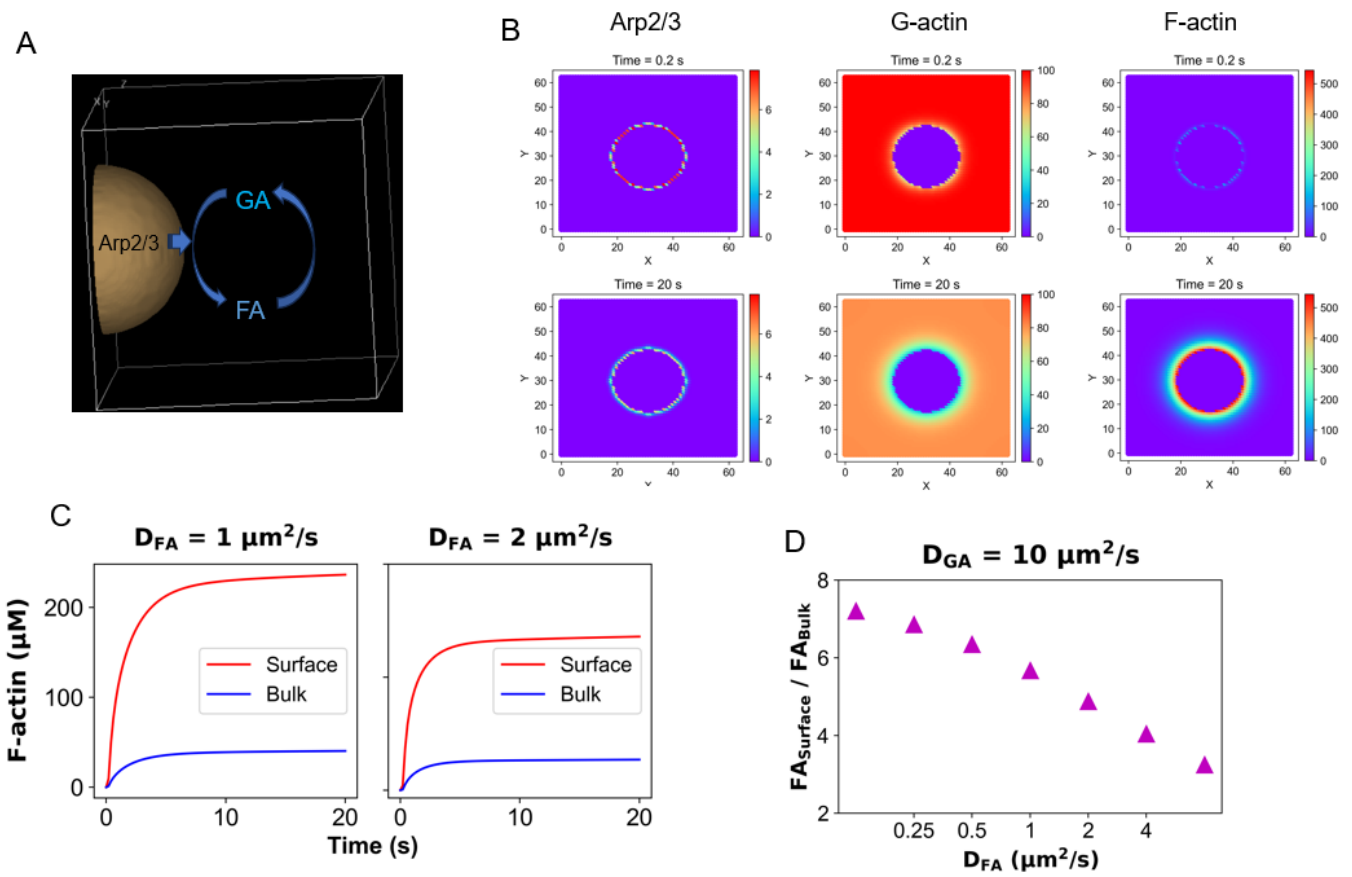

Figure 2: Reaction diffusion model with a single condensate. (A) Illustration of the simulation geometry. The "Virtual Cell" software is used to define the geometry and set up the simulation conditions. The hemispherical condensate is situated on the XY plane. Arp2/3 only diffuses around the condensate surface. Actins (both G and F) can diffuse in the entire region. (B) Concentration profiles of the model components at the beginning of simulations (Time = 0, upper panel) and the steady state (Time = 20s, lower panel). These snapshots represent the plane at Z = 0. D_{GA} = 10 μm^2/s and D_{FA} = 1 μm^2/s. (C) Timecourse of F-actin (FA) concentration at two FA diffusion constants (D_{FA}). To compute the surface concentration, a rectangular region around the condensate is considered, while bulk concentration includes all the volume elements including the condensate surface. (D) Measure of local production of F-actin as a function of differential diffusion of actin in monomeric (D_{GA}) and polymeric (D_{FA}) form. We keep D_{GA} (10 μm^2/s) fixed and vary D_{FA} .

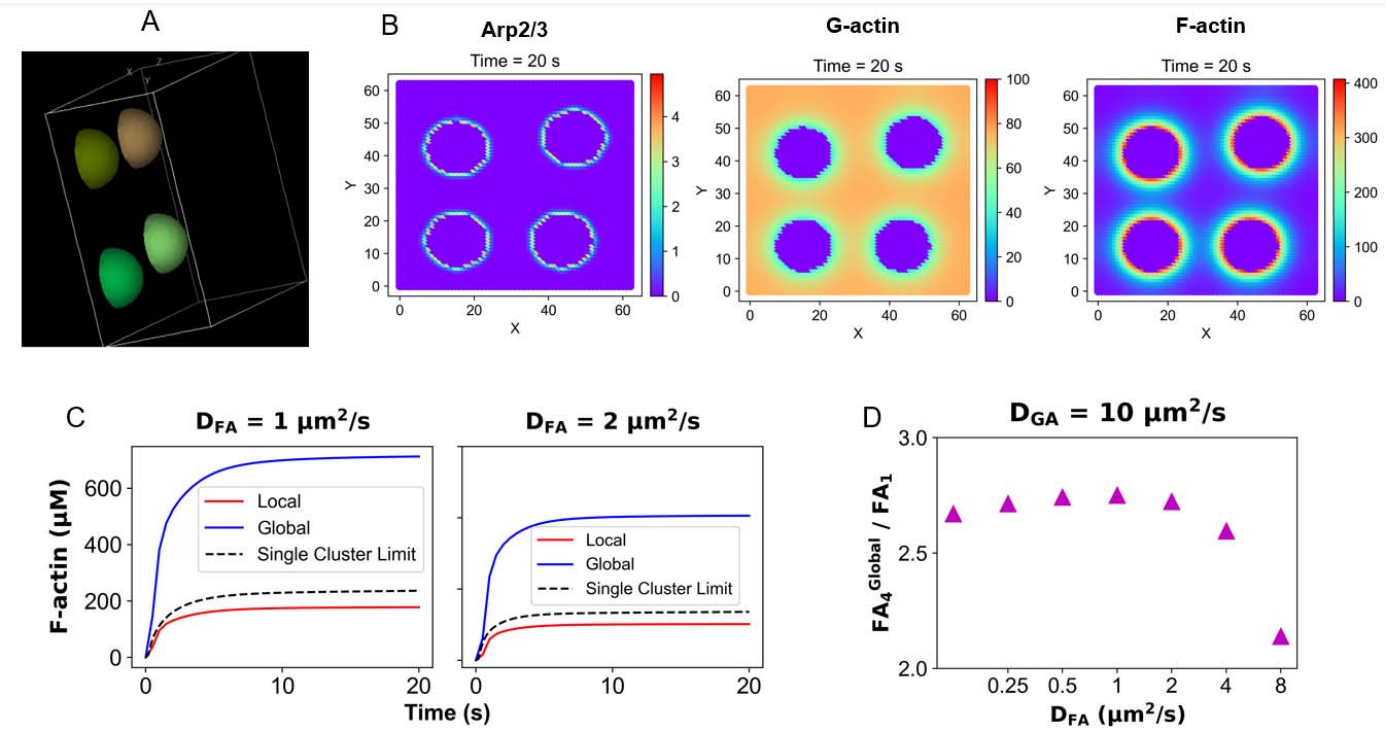

Figure 3: Reaction diffusion model with multiple condensates. (A) Illustration of the simulation setup. This snapshot from the "Virtual Cell" displays four hemispherical condensates situated at the XY plane. Similar to Fig. 2A, Arp2/3 only diffuses around the condensate surface. Actins (both G and F) can diffuse in the entire region. (B) Concentration profiles of the model components at the steady state (Time = 20s, last timepoint). These snapshots represent the plane at Z = 0. D_{GA} = 10 μm^2/s and D_{FA} = 1 μm^2/s. (C) Timecourse of F-actin (FA) concentration at two FA diffusion constants (D_{FA}). "Local" refers to the F-actin (FA) produced at each condensate surface, while "Global" is the cumulative FA produced at four condensate surfaces. To compute the surface concentration, a rectangular region around the condensate is considered. The black dotted line ("single cluster limit") indicates the FA concentration (surface FA from Fig. 2C) when equivalent amount of Arp2/3 is concentrated on the surface of a single large condensate. (D) Quantification of relative gain (in FA production) as a function of differential diffusion of actin in monomeric (D_{GA}) and polymeric (D_{FA}) form. F_4^{Global} indicates the global FA concentration (blue lines in C) in the 4-cluster state while F_1 is the single-cluster limit.

local enrichment of FA is a kinetic phenomenon which is a direct effect of slower diffusion of polymeric actin compared to its monomeric counterpart.

3. Multi-cluster state yields greater F-actin production

We next use our reaction diffusion model to compare a multi-cluster state (MCS) with its single cluster counterpart. Fig. 3A illustrates the simulation set up where we have four hemi-spherical clusters placed on the XY plane. The radius of each cluster is adjusted such that the total volume occupied by four clusters is same as the large cluster considered in Fig. 2A. The number of Arp2/3 on each cluster is also divided by 4 such that the total Arp2/3 count is same as Fig.2A. So, we have a comparison between one-cluster and four-cluster scenarios where total volume of clusters and total Arp2/3 count are the same. The question is which configuration is more efficient in FA production?

Fig. 3B shows that FA production happens near the surface of each cluster. Again, due to Arp2/3 localization in the cluster and slower diffusion of FA, we observe a local FA accumulation around the clusters. When we quantify the FA concentrations, we notice a very interesting trend (Fig. 3C). For the four-cluster configuration, local FA production per cluster is lower (red line) than the single cluster analogue (black dotted line, derived from Fig. 2C). Since there are four clusters present in the system, the global output = 4 * local output. The blue line (Fig. 3C) depicts that the global FA production, for the four-cluster state, is not only greater, but also faster. This behavior is again dependent on FA diffusion constant (two panels, Fig. 3C), which is consistent with the notion that the local FA production is a kinetic phenomenon. If we systematically probe the relative gain in FA production, Fig.3D shows that the global gain in four-cluster state remains almost same for a wider range of FA diffusion. When FA diffusion becomes comparable with GA diffusion, we see a dramatic drop in the relative gain. It is to be noted that the configuration with D_{FA} = 1 $\mu m^2/s$ and D_{GA} = 10 $\mu m^2/s$ is closest to the physiological condition. Hence, slower diffusion mediated local production of FA makes the four-cluster state more efficient compared to the single large cluster.

4. Systems theoretic analysis reveals an optimal cluster count

The reaction diffusion model informs us that (a) FA gets produced locally near the cluster and (b) MCS produces greater FA than its equivalent single cluster analogue. We then seek to probe this trend further with a systems theoretic approach. We first express the n-cluster state by the following ordinary differential equation (ODE):

$$\frac{dF_{local}}{dt} = k1 * G_{local}^2 * \rho(n) * F_{local} - k2 * F_{local} \qquad (1)$$

$$G_{local} + F_{local} = C\ (\text{constant}) \qquad (2)$$

The GA to FA conversion happens locally near the cluster with the constraint of mass conservation (Equation 2), $G_{local} + F_{local}$ = constant. $\rho(n)$ is the surface density of Arp2/3 on each of the n clusters. k1 and k2 are rate constants for polymerization and depolymerization. We assume each cluster to be hemi-spherical such that (appendix 1)

$$\rho(n) = \alpha * n^{-1/3} \quad (3)$$

where α is the surface density of Arp2/3 when n = 1.

We first analyze the single cluster case (n = 1). Fig. 4A displays the bifurcation behavior of the system. We compute the roots of the right-hand side of equation 1 and their stability as a function of α. The upper two panels indicate a stable solution of F = 0, while for the lower two panels, the stable roots vary as we change α. Fig. 4B clearly shows that there is a critical value of α, beyond which the system converges to different non-zero F-actin levels. In other words, surface density of Arp2/3 serves as a tunable bifurcation parameter that switches the system from a non-responsive state to a responsive state. Stability analysis of the roots of Equation 1 yields (appendix 2)

$$\alpha_{critical} = k2/(k1 * C^2) \quad (4)$$

For the parameters (k1 = 0.001, k2 = 1 and C = 100) used in the model, $\alpha_{critical}$ = 0.1, consistent with the bifurcation behavior in Figs. 4A and 4B. For a given dissociation rate parameter (k2 = 1), Fig. 4C shows the state space of $\alpha_{critical}$ as a function of polymerization rate (k1) and total amount of actin pool (C). For a fixed value of k1, if we increase C (move along the column), value of $\alpha_{critical}$ goes down; that is, even lower density of Arp2/3 can trigger polymerization in the presence of a higher amount of actin pool. Similarly, for a given amount of C, higher k1 (move along the row) causes $\alpha_{critical}$ to go down. Overall, Figs. 4A, 4B and 4C show that, to initiate the actin-branching, each cluster needs to recruit a threshold amount of Arp2/3 such that the surface density exceeds $\alpha_{critical}$. The extent of $\alpha_{critical}$ depends on the reaction rates and amount of total actin pool available near the cluster.

We next want to investigate the causal factor underlying the greater response coming from the MCS. Fig. 3 informed us that if we split a large cluster into 4 smaller clusters and distribute N Arp2/3 equally into each cluster (such that each smaller cluster retains N/4), 4-cluster scenario produces higher F-actin response compared to the single large cluster. Although the volumes of clusters are the same in both cases, surface to volume ratio increases as we split one

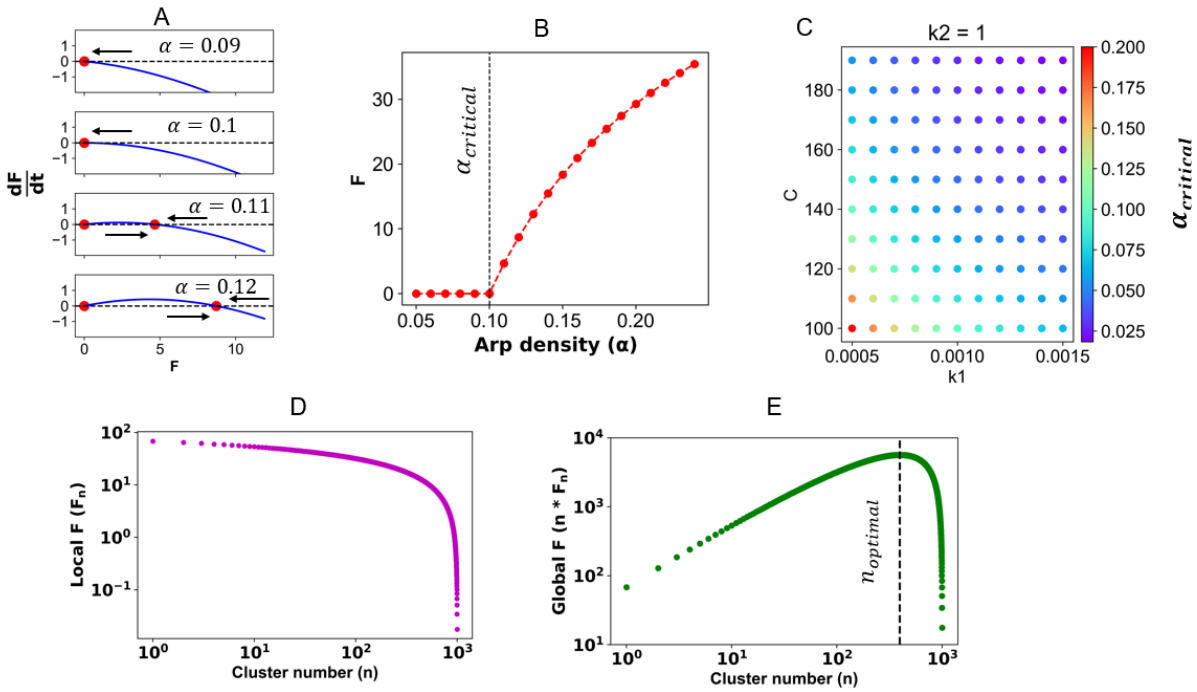

Figure 4: Analytical solution of the ODE model. (A) Bifurcation analysis. Each panel shows the functional value of $\frac{dF}{dt}$ (equation 1) at a given value of Arp density (α). The red circles refer to the "roots" of equation 1, that is, the values of F where $\frac{dF}{dt} = 0$. The arrow represents the direction of "flow"; it is rightward for a region where $\frac{dF}{dt} > 0$ and leftward otherwise. Flow towards a root signifies stability. (B) Solution of F-actin (F) as a function of Arp density. The dotted line indicates the bifurcation point where F production starts to happen. Density of Arp at this point is termed $\alpha_{critical}$. (C) Value of $\alpha_{critical}$ (color-coded) as a function of k1 and C (equation 2). (D) Amount of F produced per cluster in n-cluster system (E) Amount of cumulative F which integrates contributions from all the clusters.

hemisphere into n smaller ones. To probe this effect further, we first derive an expression for local actin production F (F per cluster in the n-cluster system) (appendix 3)

$$F_{local} = C - \sqrt{(1/\alpha)(k2/k1)} * n^{1/6} \qquad (5)$$

Using Equation 5, we compute the local F as a function of cluster number (n). Fig. 4D shows a rather surprising trend. As we increase n, initially the drop in per-cluster F is minimal. As a consequence, on the cumulative global scale (Fig. 4E), we have a gradual increase in F production. We note the logarithm scale along both the axes. As we go to large cluster numbers, beyond a certain value (n = 419), local F starts to drop dramatically which also drags down global production. The differential rate at which local F drops as a function of n yields the most fascinating aspect of the global F trend - an optimal cluster number ($n_{optimal}$). Fig. 4F suggests that there exists a multi-cluster configuration where the F-actin production, on the global or system level, is maximum.

Upon close inspection of Equation 1, we realize that the minimal drop in local F (Fig. 4D) stems from the highly non-linear nature of the F production. Fig. 4B shows that each cluster needs to have a threshold Arp2/3 density ($\alpha_{critical}$) to "turn on" the F production switch. Beyond that point, due to the inherent nonlinearity in production rate, putting more Arp2/3 onto the same cluster is less efficient; rather creating a second reaction center (cluster) with $\alpha_{critical}$ Arp2/3 is a more prudent strategy. Recursively, if we follow the same scheme, creating more clusters produces more F-actin. Since we have a finite amount of Arp2/3 to distribute amongst n clusters, beyond a parameter-specific value of n, Arp2/3 density on each cluster surface falls below the critical level and F production gets "switched off". Hence, the existence of a threshold Arp2/3 density and nonlinear production of F-actin makes the multi-cluster configuration more efficient. Moreover, an optimal number of clusters seems to be the best functional solution.

5. Statistical distribution of cluster sizes yields distinct F-actin production profiles

In the previous section, all the clusters are assumed to have the same size. In reality, we expect a size distribution. So, we next ask what happens to the F-actin profile if we consider a distribution of cluster sizes? We generate distributions with the constraint of volume conservation (Fig. 5A). The recruitment of Arp2/3 to a cluster is proportional to its volume (detailed in the Method section).

We first consider a gaussian distribution. For an n equal-sized-cluster state, if r_{small} is the radius of each cluster, the distribution is created with the mean of r_{small} and a prescribed coefficient of

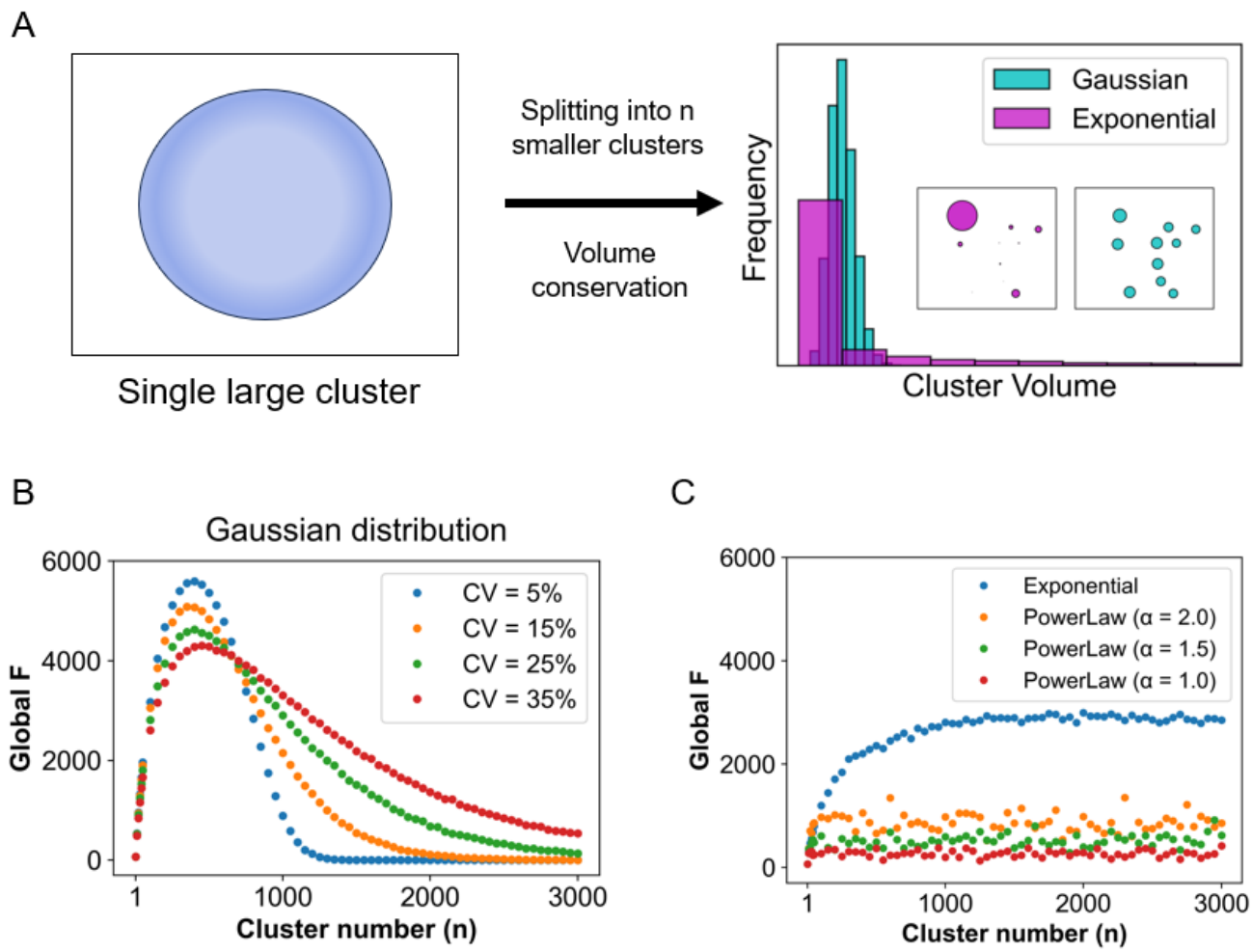

Figure 5: Effect of cluster size distributions on global F-actin production. (A) Illustration of choosing the distribution of cluster sizes. When we split one large cluster (radius = R) into n clusters of equal size, radius of each small cluster, $r_{small} = n^{-0.33} * R$. We create a distribution (mean = r_{small}) of cluster radii in a such way that the total volume is conserved during 1-to-n splitting process. On the right side, gaussian and exponential distributions are shown. Inset displays the single frames drawn from each of the two distributions (n = 10). (B) When cluster size follows a gaussian distribution, trend of global F-actin production at four different coefficients of variation (CVs). $CV = (\sigma/\mu) * 100$, where σ and μ are standard deviation and mean of the gaussian distribution, respectively. CV measures the width of the distribution. $\mu = r_{small}$ for all cases. (C) Trend of global F-actin productions when cluster sizes conform to an exponential or a power law distribution. To create exponential distribution, we need a rate parameter (λ), such that, $f(x, \lambda) = \lambda * exp(-\lambda * x)$. We set the mean, $1/\lambda = r_{small}$. To create a power law distribution, we need a scale (m) and shape (α) parameter, such that, $f(x, \alpha, m) = \frac{\alpha * m^{\alpha}}{x^{\alpha+1}}$. We set m = r_{small}, and vary α (2, 1.5, 1). To compute F-actin production, we generate 10 distributions for each n.

variation (CV). Fig. 5B shows the trend of global F production at three different CVs. In all three cases, the optimal cluster count, where global F reaches maximum, remains similar. However, the shape of the global F curve becomes more asymmetric as we make the size distribution wider (higher CV). Figs. S2A, S2B reveal that the wider gaussian distribution, even at larger n, can sample few points above the bifurcation threshold that produce non-zero F-actin. This results in a longer tail in the global F-actin production (Fig. 5B). From a biological perspective, a wider distribution provides a wider "responsive range", that is, extent of global F is held at a non-zero level for a wider range of cluster sizes.

Lee et al (27) recently quantified the size distributions of natural and synthetic condensates in living cells. They reported that the physiological condensates conform to an exponential distribution, while pathological aggregates follow a power-law like distribution. So, we wanted to find out our model predictions if we prescribed exponential or power-law distribution for the cluster size. Fig. 5C suggests that an exponential distribution produces a distinct F-actin profile which grows initially and beyond a cluster count, it plateaus. If we compare this with the outcome of gaussian distribution, the maximal level of global F is roughly half. However, global F is retained at the maximal level for a really large cluster size range. Furthermore, n = 1000 produces ~3000 times more F than n = 1. For a power-law distribution, the system simply fails to produce a significant response. With larger exponent (α), which signifies smaller tails, the response improves. Figs. S2C and S2D show that the long-tail nature of both exponential and power-law distributions allow sampling of fewer points above the bifurcation threshold even at larger n. However, purely due to the statistical sampling, an exponential distribution can produce higher global F, while the power-law distribution with elongated tail suffers from a lower level of global F.

Discussion

The existence of multiple biological condensates directly contradicts the prediction of polymer physics that states that, at equilibrium, one large droplet should coexist with the soluble phase. Since biological systems operate directly under evolutionary pressure, we investigate the following problem – given a fixed number of molecular components, how to distribute them spatially to achieve maximal output? The output could be system-specific or context dependent. But we seek to uncover a generic principle that may incentivize the stabilization of multi-droplet state. Using a combination of three complementary computational techniques, we have

established that, indeed, multiple droplets are functionally more efficient compared to their single droplet counterpart.

We have used the condensation of signaling proteins (Nephrin, Nck, NWASP and Arp2/3) upstream to actin nucleation pathway as a case study. Langevin dynamics revealed the hemi-spherical shape of the condensate. It also showed that Arp2/3 resides on the surface (Fig. 1). We then used a hemispherical solid geometry as a proxy for the condensate where Arp2/3 diffuses on the surface. The reaction diffusion model predicted a local production of F-actin which depends on the differential diffusion of actin in G and F form (Fig. 2). This local enrichment of F-actin makes a multi-cluster configuration more efficient in producing F-actin (Fig. 3). This diffusion mediated localized F-actin production enabled us to write an ordinary differential equation (Equation 1) where G-actin locally gets converted into F-actin in an Arp2/3 dependent manner. Our theoretical analysis revealed that surface density of Arp2/3 serves as a functional switch that triggers F-actin production beyond a threshold density (Fig. 4). Due to the highly non-linear production rate, having many clusters, each equipped with a threshold amount of Arp2/3, creates an optimal scenario where F-actin production is maximum. Depending on the cluster size distributions, the F-actin production profile, on the system scale, takes distinct shapes (Fig. 5) that may be related to the disease phenotype of cellular systems. This study provides an excellent example of how multiple computational tools can synergistically be employed to gain a holistic understanding of a biological phenomenon. It also demonstrates how to amalgamate scale-specific insights into a multi-scale picture.

Organization of binding sites within a protein sequence determines the mesoscopic architecture of the condensate. In our system, the spatial arrangement of SH2, SH3, PRM and VCA dictates the peripheral location of Arp2/3 which is optimal for the desired downstream effect (F-actin production). If we shuffle the arrangements of binding domains/motifs, the condensate will likely assume a different spatial architecture. It is intriguing to think that for a set of condensate-forming proteins that participate in a biochemical task (signaling pathway, for example), the sequences may have co-evolved to yield a condensate structure that can facilitate downstream processes. Given the spatial heterogeneity of almost all known biomolecular condensates (P granules, Stress granules, Nucleolus etc.), our functional perspective provides a systematic way to think about the core and the interface of these condensates. It might also allow us to rationalize the relative location of the complementary binding sites in the partner molecules that form a condensate.

We have previously proposed that the multi-cluster state or MCS is a kinetically arrested metastable state (7). With recent experimental validation (28), it is becoming evident that spatially heterogeneous multi-phasic nature of condensates as well as the long lifetime of multiple condensates are purely kinetic phenomena. In the current study, we have identified that the downstream reaction module like F-actin production is also controlled by kinetic effects like diffusion. This coupling of kinetic effects in upstream and downstream layers makes multiple condensates an optimum solution. Given the necessity of spatiotemporal regulation of cellular biochemistry, it is not surprising that living cells have evolved to take advantage of the kinetics or timing of events.

Our analysis underscores the functional importance of the size distribution of clusters. The system considered here is responsible for cell signaling. For instance, the rate of actin branching at the leading edge of a moving cell dictates the speed and efficacy of cell migration. Different cluster size distributions essentially produce distinct response patterns (Fig. 5). By creating a more asymmetric distribution (from gaussian to exponential), a “spiked” response becomes more “sustained”. However, there is a limit to this process - creating a fat-tailed distribution like a power-law might trigger a pathological state which simply cannot respond to the external signal. Depending on the spatiotemporal context, different cellular systems can exploit these distributions to achieve specific tasks.

In summary, our multi-scale analysis offers a compelling justification for the presence of multiple condensates within living cells. For a function-optimization incentive, sequence of the individual molecules as well as the structure of the molecular network may have evolved to take advantage of the kinetic effects in maximizing the desired output.

Methods

A. Langevin Dynamics

A.1. Bead-spring representation

We used bead-spring polymers for the coarse-grained representation of the signaling proteins (Fig. 1A). Nephrin, Nck and NWASP all contain intrinsically disordered regions (IDRs) in their sequences (29), but Arp2/3 complex is a large structured protein made with

seven subunits (30). Each component is modelled with a minimal description that includes a known number of stickers (valence) and except for Arp2/3, a pair of spacers between the stickers to mimic the flexibility imparted by the IDRs. For Arp2/3, we have two stickers connected by a bulky spacer to model the steric effects coming from the large seven-subunit protein complex.

A.2. Force Field

Each polymer contains the prescribed number of beads connected by harmonic bonds. Stretching energy of each harmonic bond, $E_{bond} = K_b * (R - R_0)^2$ where K_b is the spring constant and R_0 is the equilibrium bond distance. R measures the distance between the bonded beads at any given time. Each bond is a harmonic distance constraint that enables permanent connectivity within a chain. In our model, R_0 = 10 Å and K_b = 3 $\frac{kcal}{mol * Å^2}$.

To allow chain flexibility, angle (θ) between three successive beads is modelled with a cosine function: $\mathrm{E_{bending}} = \kappa * (1 - \cos\theta)$ where κ decides the bending stiffness. In our model, $\kappa = 2\ \mathrm{kcal.\ mo}l^{-1}$.

Each pair of beads interacts via a non-bonded isotropic interaction, modelled by Lennard-Jones (LJ) potential: $E_{LJ} = 4 * \epsilon * [\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}]$ where σ represents the bead diameter and r is the separation between the beads. ϵ is the depth of energy-well that determines the strength of attractive potential. To achieve computational efficiency, the LJ potential is neglected beyond a cut-off distance (R_{max}). In our model, except for the Arp2/3 spacer, $\sigma = 10$ Å. For the bulky Arp2/3 spacer, $\sigma = 30$ Å. For all pairwise LJ interactions, $R_{max} = 2.5 * \sigma$.

A.3. Modeling specific interactions

To encode "specific" interactions between complementary sticker types, we introduced inter-sticker bonds. When two stickers approach within a cut-off radius (R_{cut}), they can form a "bond" with a probability, p_{on}. The bond can be broken with a probability of p_{off}, at or greater the distance of R_{cut}. These bonds are harmonic distance constraint (equilibrium distance = R_0^s) governed by a spring constant parameter. Energy of these inter-sticker bonds, $E_{Sb} = \mathrm{Ks} * (R - R_0^s)^2$. Since the spring constant (Ks) decides the stiffness of the bond, it also sets the energy cost of breaking the bond. These are "specific" interactions because once a sticker pair is bonded, they can't form another bond with complementary stickers that are

still within R_{cut}. In other words, each sticker has a valency of 1. In our model, $p_{on} = p_{off} = 1$, $R_0^s = 1.122 * \sigma$, $R_{cut} = R_0^s + 1.5$ Å.

We have three types of binding interactions: 1. pY and SH2, 2. SH3 and PRM, 3. VCA and Arp2/3. Bond strength (Ks) for type 1, $Ks = 3 \frac{kcal}{mol * Å^2}$, while for type 2 and type 3, $Ks = 1.8 \frac{kcal}{mol * Å^2}$.

A.4. Modeling confined diffusion

Since Nephrin is a transmembrane receptor protein, our model Nephrin contains a "anchor" bead which is tethered on the XY surface such that it can only diffuse in 2-dimensional plane. This is achieved by setting the z-component of force and velocity (applied on the anchor bead) to zero.

A.5. Simulation Details

We have used the LAMMPS package (31, 32) to perform Langevin dynamics simulations. We performed our simulations in a rectangular cuboid box (X = 600Å, Y = 600Å, Z = 350Å). XY face (and opposite face) has a reflective boundary condition to model the steric hindrance from the membrane, that is, when a particle hits the membrane, it gets bounced off. The other four faces (XZ and YZ) of the box have periodic boundary conditions. Simulation temperature is 310 K. The viscosity of the simulation medium is described with a "damp" parameter which is set to 500 femtoseconds (fs). For the membrane-bound anchor bead, damp = 50 fs since membrane diffusion is slower than the cytosolic diffusion. The damp parameter is inversely proportional to the viscosity of the solvent. Simulation timestep = 25 fs.

We have used metadynamics simulations (33) to facilitate the processes of molecular clustering near the membrane. Metadynamics is an enhanced sampling scheme where the auxiliary gaussian potentials are imposed along a user defined order parameter (also known as collective variables, reaction coordinates etc.) to reconstruct the free energy profile of the process of interest. We used the distance between two groups of chains as our order parameter – one group consists of all the anchor (part of Nephrin) beads and the other group collects the middle beads of all the Nck, NWASP and Arp2/3. This parameter basically monitors how the free energy profile changes (Fig. S1) as the cytosolic molecules gradually

cluster near the membrane. Once the order parameter converges to a steady level, we run standard Brownian dynamics (without bias) to obtain the cluster properties.

A.6. Data Analysis

For metadynamics simulations, we analyzed the time evolution of the order parameter and related potential mean force (PMF) profile. To analyze physical properties of the cluster, we used the configuration files ("restart" files in LAMMPS) containing information of coordinates and topology of the system. The topology information was then converted into a network (cluster) to infer the connected components. We fitted all the (intra-cluster) bead coordinates to a convex hull and computed the volume and surface area of the hull (Fig. 1C). Once we computed the volume of the cluster, to design the cuboid reference (Fig. 1C, inset), we distributed all the beads across a thin layer (green points) such that the total volume remains the same as the cluster.

A.7. Software

Moltemplate: To create the model polymers, we made use of the Moltemplate (34) package which enables the user to create multiple types of chains in a template based manner.

Packmol: The polymers are packed inside the simulation volume using the PACKMOL (35) package.

LAMMPS: The Langevin dynamics simulations are performed using the LAMMPS software package (32). We used the "COLVARS" functionality (36) to perform metadynamics within LAMMPS.

OVITO: We used the OVITO (basic version) software to visualize the particle motions (37).

Python: We used custom python scripts to setup simulations and analyze the data.

We have organized and released example simulations and python code. (https://github.com/achattaraj/OptimalCondensateCount)

Reaction Diffusion Simulations

B.1. Virtual Cell software

We have used the Virtual Cell (VCell) software (38, 39) to build and simulate the reaction diffusion models. VCell can simulate a reaction diffusion system within an arbitrary

geometry. The detailed algorithm, including numerical solvers, boundary conditions etc., can be found here (40).

Briefly, the governing equation is

$$\delta C_i/dt = D_i \nabla^2 C_i + R_i, \quad i = 1,2,3, \dots, n$$

where C_i and D_i are the concentration and diffusion constant of the ith species, respectively. The effect of all the reactions on the ith species is captured by the reaction term R_i, which depends on the concentrations, $C_1, C_2, C_3, \dots, C_n$. If there are m reactions with rate β_j where j = 1,2,3, …,m, then

$$R_i = \sum_j \alpha_{ij} \beta_j$$

α_{ij} is the stoichiometry matrix that represents how many molecules of ith species are consumed or produced during the jth reaction.

B.2. Simulation setup

Since we are modeling the actin branching near the membrane-associated condensates, we have three components – Arp2/3, G-actin (GA) and F-actin (FA). We have used a lumped kinetic expression (similar to (41)) to model the dendritic nucleation of actin branching, described by Equations 1 and 2.

In the single-condensate model, there are two spatial compartments – one is condensate, and the other is cytosol. To resolve and separate out spatial domains, VCell assigns a “membrane” between compartments, which will be called CM or condensate membrane. This should not be confused with the concept of a “cell membrane” which happens to be the XY plane. Our hemispherical condensate is tethered to the XY plane.

GA and FA diffuse in the cytosol. Arp2/3 (“arp_cm”) is initially localized on the CM. By convention, VCell needs reacting species to occupy the same compartment to engage in a reaction. Since Arp2/3 facilitates the GA to FA conversion, we have a translocation process which transfers Arp2/3 from CM to cytosol (“arp_cyt”). The diffusion of arp_cyt is set to be very low. In that way, FA production still happens near the condensate surface and the initial surface density of Arp2/3 determines the reaction propensity.

For our reference system, diffusion constants (in the unit $\mu m^2/s$), GA = 10, FA = 1, arp_cm = 1, arp_cyt = 0.001.

Dimension of the computational domain = (20, 20, 10) µm, which is divided into a cytosol and a hemispherical condensate. Radius of the hemisphere = 4 µm. Hence, volume of cytosol = 3866 μm^3, surface area of the condensate = 100 μm^2.

Initial concentrations of the model components, GA = 100 µM, FA = 0.01 µM, arp_cm = 1000 molecules/μm^2, arp_cyt = 0.

B.3. Model availability

Both the single-condensate (Name: AC_public_OneCondensate_ARP) and four-condensate (Name: AC_public_FourCondensate_ARP) models are publicly available in the VCell database, under the username "Ani". VCell is freely available (https://vcell.org/run-vcell-software) which offer graphical interfaces to define geometry, species, reactions. One can then scan parameters to investigate different biological scenarios.

B.4. Data analysis

Python codes to analyze the surface and bulk concentrations of F-actin can be found here (https://github.com/achattaraj/OptimalCondensateCount)

C. Solution of the Ordinary Differential Equations (ODEs)

C.1. Statistical distribution of cluster sizes

To investigate the effects of unequal cluster sizes, we create cluster size distributions (as detailed in Fig. 5A). To demonstrate the workflow, we will consider a simple case of n = 4, that is, splitting of one large cluster into 4 smaller ones. We sample the 4 clusters in a such a way that the volume is conserved during the process. Let's say, radius of the single large cluster is R (Volume = V_{large}) and total count of Arp2/3 is A molecules. If all the smaller clusters have identical size, then $r_{small} = 4^{-0.33} * R \sim 0.63R$. Now, if want to sample the radii from a gaussian distribution, we draw four radii (r1, r2, r3, r4) from a normal distribution with mean = r_{small} and a CV (coefficient of variation). Total volume of the four clusters,

$$V_{total} = (4/3) * \pi * \sum_i r_i^3$$

To ensure volume conservation, we compute a scale-factor,

$$scalefactor = \left(V_{large} / V_{total}\right)^{1/3}$$

We then multiply the scale-factor to radii of the smaller clusters. Now, the cluster volume after the splitting remains unchanged.

To distribute Arp2/3 amongst the smaller clusters, we assume that extent of Arp2/3 per cluster should be proportional to the volume of the cluster. Since cluster volume decides how much NWASP (and Nck + Nephrin) is in it, and Arp2/3 is recruited by NWASP, Arp2/3 partitioning should be decided by the relative volume fractions of the clusters.

Hence, Arp2/3 count in the ith cluster,

$$a_i = V_i / V_{total}$$

Surface density,

$$S_i = a_i / 4\pi {r_i}^2$$

C.2 Code availability

We have used custom Python scripts to solve the ODEs. All the scripts are available, in the convenient form of Jupyter Notebooks, here (https://github.com/achattaraj/OptimalCondensateCount).

Acknowledgements

We are grateful to Junlang Liu and Srivastav Ranganathan for numerous fruitful discussions. A.C. thanks Junlang for useful discussions regarding the analytical solution presented here. A.C. greatly appreciates all the help with the Virtual Cell software from Profs. Les Loew and Michael Blinov. A.C. thanks Prof. Rohit Pappu for suggesting the analysis with a distribution of cluster sizes. A.C. thanks his colleague Prof. Jon Ditlev for providing constructive feedback on the manuscript. This work was supported by NIH (Grant 5R35GM139571).

Author Contributions

A.C. conceptualized the project, performed simulations, analyzed the data and wrote the paper.

E.I.S. secured the funding, supervised the project, analyzed the data and wrote the paper.

References

1. Cooper G. The Origin and Evolution of Cells. . The Cell: A Molecular Approach 2nd edition Sunderland (MA): Sinauer Associates2000.
2. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18(5):285-98.
3. Falkenberg CV, Blinov ML, Loew LM. Pleomorphic ensembles: formation of large clusters composed of weakly interacting multivalent molecules. Biophys J. 2013;105(11):2451-60.
4. Banjade S, Rosen MK. Phase transitions of multivalent proteins can promote clustering of membrane receptors. Elife. 2014;3.
5. Chattaraj A, Youngstrom M, Loew LM. The Interplay of Structural and Cellular Biophysics Controls Clustering of Multivalent Molecules. Biophys J. 2019;116(3):560-72.
6. Choi J-M, Dar F, Pappu RV. LASSI: A lattice model for simulating phase transitions of multivalent proteins. PLOS Computational Biology. 2019;15(10):e1007028.
7. Ranganathan S, Shakhnovich EI. Dynamic metastable long-living droplets formed by sticker-spacer proteins. Elife. 2020;9.
8. Semenov AN, Rubinstein M. Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics. Macromolecules. 1998;31(4):1373-85.
9. Flory PJ. Thermodynamics of high polymer solutions. The Journal of chemical physics. 1942;10(1):51-61.
10. Huggins ML. Some properties of solutions of long-chain compounds. The Journal of Physical Chemistry. 1942;46(1):151-8.
11. Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. Nature. 2012;483(7389):336-40.
12. Case LB, Zhang X, Ditlev JA, Rosen MK. Stoichiometry controls activity of phase-separated clusters of actin signaling proteins. Science (New York, NY). 2019;363(6431):1093-7.
13. Jones N, Blasutig IM, Eremina V, Ruston JM, Bladt F, Li H, et al. Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. Nature. 2006;440(7085):818-23.
14. Ditlev JA, Case LB, Rosen MK. Who's In and Who's Out-Compositional Control of Biomolecular Condensates. Journal of molecular biology. 2018;430(23):4666-84.
15. Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell. 2016;165(7):1686-97.
16. Wang JT, Smith J, Chen BC, Schmidt H, Rasoloson D, Paix A, et al. Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans. Elife. 2014;3:e04591.
17. Folkmann AW, Putnam A, Lee CF, Seydoux G. Regulation of biomolecular condensates by interfacial protein clusters. Science (New York, NY). 2021;373(6560):1218-24.
18. Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, Parker R. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell. 2016;164(3):487-98.

19. West JA, Mito M, Kurosaka S, Takumi T, Tanegashima C, Chujo T, et al. Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. J Cell Biol. 2016;214(7):817-30.
20. Mizutani A, Tan C, Sugita Y, Takada S. Micelle-like clusters in phase-separated Nanog condensates: A molecular simulation study. PLOS Computational Biology. 2023;19(7):e1011321.
21. Welch MD. The world according to Arp: regulation of actin nucleation by the Arp2/3 complex. Trends in cell biology. 1999;9 11:423-7.
22. Eden S, Rohatgi R, Podtelejnikov AV, Mann M, Kirschner MW. Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. Nature. 2002;418:790-3.
23. Rohatgi R, Nollau P, Ho H-Y, Kirschner MW, Mayer BJ. Nck and Phosphatidylinositol 4,5-Bisphosphate Synergistically Activate Actin Polymerization through the N-WASP-Arp2/3 Pathway*. The Journal of biological chemistry. 2001;276:26448 - 52.
24. Tryggvason K, Pikkarainen T, Patrakka J. Nck Links Nephrin to Actin in Kidney Podocytes. Cell. 2006;125:221-4.
25. Padrick SB, Rosen MK. Physical mechanisms of signal integration by WASP family proteins. Annual review of biochemistry. 2010;79:707-35.
26. Plastino J, Lelidis I, Prost J, Sykes C. The effect of diffusion, depolymerization and nucleation promoting factors on actin gel growth. European Biophysics Journal. 2004;33:310-20.
27. Lee DSW, Choi C-H, Sanders DW, Beckers L, Riback JA, Brangwynne CP, et al. Size distributions of intracellular condensates reflect competition between coalescence and nucleation. Nature Physics. 2023;19(4):586-96.
28. Lin AZ, Ruff KM, Dar F, Jalihal A, King MR, Lalmansingh JM, et al. Dynamical control enables the formation of demixed biomolecular condensates. Nature Communications. 2023;14(1):7678.
29. Pak Chi W, Kosno M, Holehouse Alex S, Padrick Shae B, Mittal A, Ali R, et al. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. Molecular Cell. 2016;63(1):72-85.
30. Goley ED, Welch MD. The ARP2/3 complex: an actin nucleator comes of age. Nature Reviews Molecular Cell Biology. 2006;7(10):713-26.
31. Plimpton S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. Journal of Computational Physics. 1995;117(1):1-19.
32. Thompson AP, Aktulga HM, Berger R, Bolintineanu DS, Brown WM, Crozier PS, et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Computer Physics Communications. 2022;271:108171.
33. Laio A, Gervasio FL. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. Reports on Progress in Physics. 2008;71(12):126601.
34. Jewett AI, Stelter D, Lambert J, Saladi SM, Roscioni OM, Ricci M, et al. Moltemplate: A Tool for Coarse-Grained Modeling of Complex Biological Matter and Soft Condensed Matter Physics. Journal of molecular biology. 2021;433(11):166841.
35. Martínez L, Andrade R, Birgin EG, Martínez JM. PACKMOL: a package for building initial configurations for molecular dynamics simulations. J Comput Chem. 2009;30(13):2157-64.
36. Fiorin G, Klein ML, Henin J. Using collective variables to drive molecular dynamics simulations. Mol Phys. 2013;111(22-23):3345-62.
37. Stukowski A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. Modelling and Simulation in Materials Science and Engineering. 2010;18(1):015012.
38. Schaff J, Fink CC, Slepchenko B, Carson JH, Loew LM. A general computational framework for modeling cellular structure and function. Biophys J. 1997;73(3):1135-46.
39. Cowan AE, Moraru, II, Schaff JC, Slepchenko BM, Loew LM. Spatial modeling of cell signaling networks. Methods Cell Biol. 2012;110:195-221.

40. Schaff JC, Slepchenko BM, Choi Y-S, Wagner J, Resasco D, Loew LM. Analysis of nonlinear dynamics on arbitrary geometries with the Virtual Cell. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2001;11(1):115-31.
41. Ditlev JA, Vacanti NM, Novak IL, Loew LM. An open model of actin dendritic nucleation. Biophys J. 2009;96(9):3529-42.

Supplementary Calculations

Appendix 1: Surface density of Arp2/3 in the n-cluster configuration

Suppose we have a large hemispherical cluster of radius R. We can split the large cluster into n smaller ones, each having a radius of r. Total number of Arp2/3 is A.

Volume (mass) conservation reads: $\frac{2}{3}\pi R^3 = n * \frac{2}{3}\pi r^3 \Rightarrow R = n^{\frac{1}{3}} r$

Surface area of the large cluster, $S = 4\pi R^2$

Surface area of a smaller cluster, $s = 4\pi r^2 = 4\pi R^2 * n^{-\frac{2}{3}} = S * n^{-\frac{2}{3}}$

Surface density of Arp2/3 at the large cluster, $\alpha = \frac{A}{S}$

Surface density of Arp2/3 at a smaller cluster, $\rho = \frac{\frac{A}{n}}{s} = \left(\frac{A}{S}\right) * n^{-\frac{1}{3}} = \alpha * n^{-\frac{1}{3}}$

Appendix 2: Derivation of Arp2/3 threshold density (single cluster case)

For one cluster case,

$$\frac{dF}{dt} = k1 * G^2 * \alpha * F - k2 * F \text{ where } G + F = C \text{ (constant)}$$

$$f = dF/dt = k1 * (C - F)^2 * \alpha * F - k2 * F$$

After differentiating w.r.t. F,

$$f' = (\alpha k1 C^2 - k2) - 4\alpha k1 C * F + 3\alpha k1 * F^2$$

$$f'(F = 0) = \alpha_{critical} k1 C^2 - k2 = 0 \Rightarrow \alpha_{critical} = \frac{k2}{k1 * C^2}$$

Appendix 3: Derivation of local production of F-actin in the n-cluster configuration

First let's consider n = 1 case,

$$\frac{dF}{dt} = k1 * G^2 * \alpha * F - k2 * F \text{ where } G + F = C \text{ (constant)}$$

Steady state, $\frac{dF}{dt} = 0 => k1 * (C - F)^2 * \alpha - k2 = 0$ since $F > 0$

$$F = C - \beta \quad (3a) \qquad \text{where } \beta = \sqrt{\frac{1}{\alpha} * \left(\frac{k2}{k1}\right)}$$

Now for the n-cluster state,

$$\frac{dF_{local}}{dt} = k1 * G_{local}^2 * \rho(n) * F_{local} - k2 * F_{local} \text{ where } G_{local} + F_{local} = C \text{ (constant)}$$

From Appendix 1, $\rho = \alpha * n^{-1/3}$

Following the same scheme as in Equation 3a,

$$F_{local} = C - \beta * n^{1/6} (3b)$$

Supplementary Figures

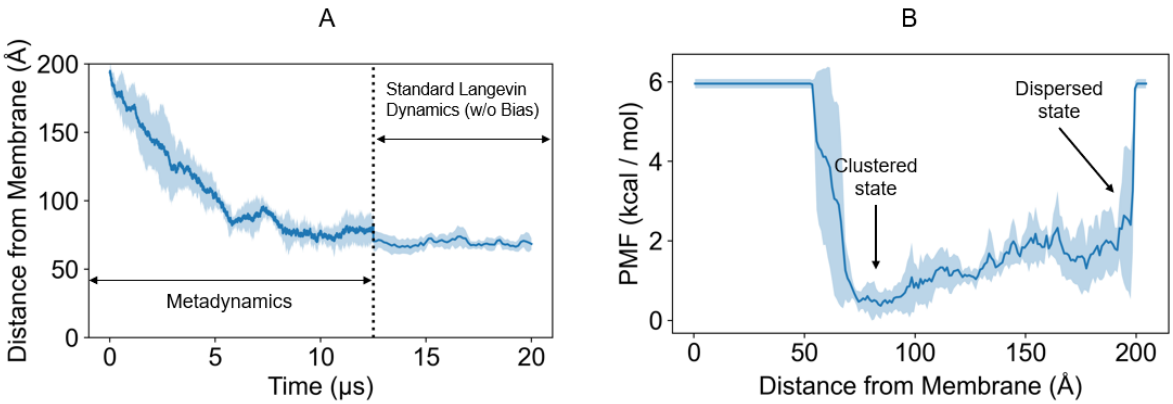

Figure S1: Quantifying the free energy profile of the clustering process by metadynamics simulations. (A) Timecourse of the order parameter. The distance between two groups of beads is computed. Group1 is the Nephrin anchor domains (attached to the 2D surface) and the middle beads of Nck, NWASP and Arp2/3 molecules consist the Group 2. The bias deposition (metadynamics) is stopped once the distance reaches to a plateau. The standard Langevin dynamics trajectory is then used to sample molecular location in the cluster (Fig. 1D). The trajectory (solid line) is an average over 5 stochastic trials where the fluctuation envelope indicates the standard deviation around the mean. (B) Potential mean force (PMF) profile.

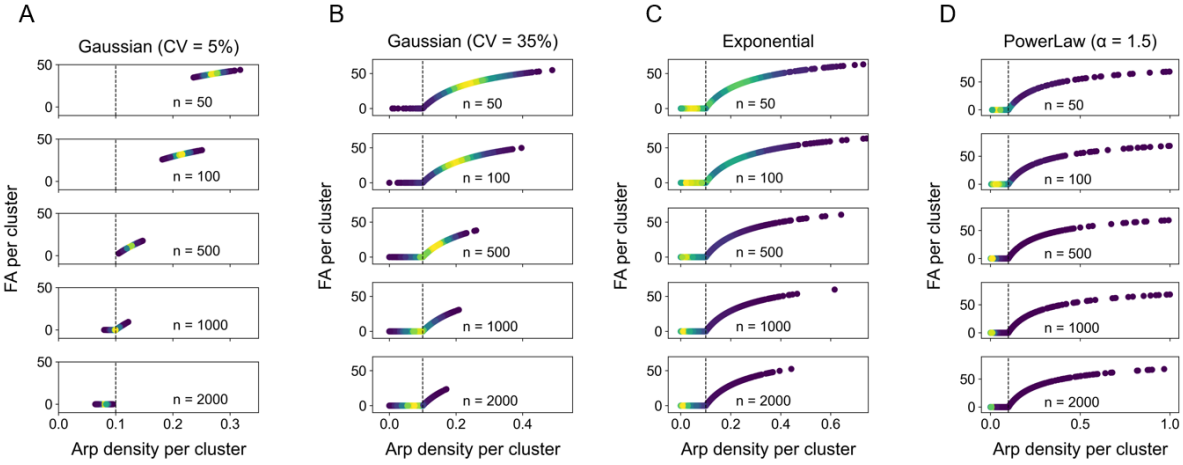

Figure S2: Local F-actin production profiles for different cluster size distributions. Cluster size follows (A) Gaussian distributions with CV = 5%, (B) Gaussian distributions with CV = 35%, (C) Exponential distributions, (D) Power Law distribution with a shape parameter (α) = 1.5. The cluster numbers (n) are indicated for each panel. The color code indicates the frequency. Darker (purple) ones are less frequent while brighter (yellow) ones are more frequent.