Title:

Nanometer condensate organization in live cells derived from partitioning measurements

Authors:

Christina Dollinger[1], Evdokiia Potolitsyna[1], Abigail G. Martin[1], Archish Anand[1], Gandhar K. Datar[1,2], Jeremy D. Schmit[3], Joshua A. Riback[1]†

[1]Department of Molecular and Cellular Biology, Baylor College of Medicine; Houston, TX 77030, United States of America

[2]Medical Scientist Training Program, Baylor College of Medicine; Houston, TX 77030, United States of America

[3]Department of Physics, Kansas State University, Manhattan, KS 66506, United States of America.

†Email: Josh.Riback@bcm.edu

Abstract:

Biomolecules associate, forming condensates that house essential biochemical processes, including ribosome biogenesis. Unraveling how condensates shape macromolecular assembly and transport requires cellular measurements of nanoscale structure. Here, we determine the organization around and between specific proteins at nanometer resolution within condensates, deploying thermodynamic principles to interpret partitioning measurements of designed protein probes. When applied to the nucleolus as a proof of principle, the data reveals considerable inhomogeneity, deviating from that expected within a liquid-like phase. The inhomogeneity can be attributed to ribosome biogenesis, with the local meshwork weakening as biogenesis progresses, facilitating transport. Beyond introducing an innovative modality for biophysical interrogation, our results suggest condensates are far from uniform, simple liquids, a property we conjecture enables regulation and proofreading.

Introduction

Organized molecular transport mitigates the stochasticity of Brownian motion, thereby enabling function. In the nucleus, the biogenesis of diverse RNA-protein complexes around chromatin occurs via bidirectional trafficking within biomolecular condensates. Condensates, being membrane-less organelles, require a spatially continuous internal network of interactions. An ongoing issue is the precise properties of this network and its role in function. Whereas proteins and nucleic acids undergo liquid-liquid phase separation (LLPS) in the test tube, forming spherical droplets that coalesce and exhibit rapid internal dynamics (*1*–*3*), condensates in cells often lack the features expected from simple LLPS (*4*–*14*), including a nanometer internal structure observable by electron microscopy (*15*). Determining the organization of biomolecules within condensates is critical for understanding how condensates impact biochemical processes during physiology and disease. While more detailed frameworks beyond simple LLPS are being proposed (*10*), elucidating the complexity of endogenous condensates requires precise cellular measurements.

Classical methods for resolving the structure of materials are poorly suited for condensates in cells. Small-angle scattering and electron microscopy, considered the most powerful and direct ways to study liquid-like organization, cannot resolve specific proteins within cells, which limits them to highly simplified model condensates (*14*–*18*). In contrast, microrheological methods, such as FRAP, fluctuation spectroscopies, and single particle bead tracking, can measure apparent viscosities and infer meshworks but are challenging to interpret (*10*, *19*–*21*). Spatial proteomics methods provide information on protein abundances and localization, not precise distances (*22*). To address these deficiencies, we develop strategies to elucidate the internal organization within a condensate at nanometer resolution in live cells.

Extracting partitioning energies in cells with microscopy

Rather than directly observing the internal structure of condensates, we will deduce them in cells using partitioning measurements paired with precisely designed protein constructs sensitive to local condensate structural attributes. We drew inspiration from experiments developed to assess the stability and folding of soluble and membrane proteins (*23*–*27*). This includes measurements quantifying the partitioning of peptides into different solvents that provided insight into the relationship between amino acid sequence and stability (e.g. hydrophobicity). In contrast to those idealized partitioning measurements, condensates in cells are highly composition-dependent (*7*), requiring partitioning measurements at near-zero probe expression to investigate unaltered endogenous properties (**Fig 1A**). To proceed, we focused on the granular component (GC) of the nucleolus as a strong proof-of-principle system for our measurements as 1) the GC is well established as a condensate, 2) it has a well-characterized proteome and transcriptome that are strongly associated with its primary function in ribosome biogenesis, and 3) it has well-documented deviations from simple LLPS, including EM-observable substructure that can aid validation of our methods (*28*, *29*). Notably, we have previously found that ribosomal intermediates diffuse on average ~5,000 fold slower than

nucleolar proteins (*13*); however, it remains unclear what structural properties imbue such a substantial difference in transport.

To enable quantitative microscopy, we introduced a homozygous knock-in of mCherry at the C-terminal end of NPM1, the most abundant GC protein that contributes to nucleolar formation (*2*, *7*, *30*). In this line, the NPM1-mCherry signal allows for demarcation of the nucleoplasmic (Np) and nucleolar (GC) regions that enable quantification of the ratio (or slope) between the concentrations (C) in compartments known as the, partition coefficient, K (**Fig 1B-C**). This includes not only for the NPM1 channel but also for exogenously introduced fluorescently tagged protein fusions, referred to here as 'probes,' (**Fig. 1A-C**). Following quantification of ~30 cells (median) for each of the 148 probes employed throughout this manuscript, we determined that the concentration dependence of the probe essentially always fits empirically to a hyperbola, yielding two quantitative values, K° and C^{max}, being the partition coefficient extrapolated to zero probe expression and max probe concentration in the GC, respectively (**Fig. S1**). The partition coefficient, K°, can be converted into a measure of the strength of the probe for the GC, the apparent standard-state free energy of transfer or $\Delta G^{\circ tr}$, through the relationship, $\Delta G^{\circ tr}$ = -RTln(K°). Excitingly, $\Delta G^{\circ tr}$ is the strength extrapolated to zero probe overexpression, here denoted by the "∘" (methods), thus mitigating well-known consequences of overexpression (*7*, *11*, *31*, *32*), including changes seen at higher probe levels that weaken endogenous NPM1 partitioning into the nucleolus (**Fig 1C, top**).

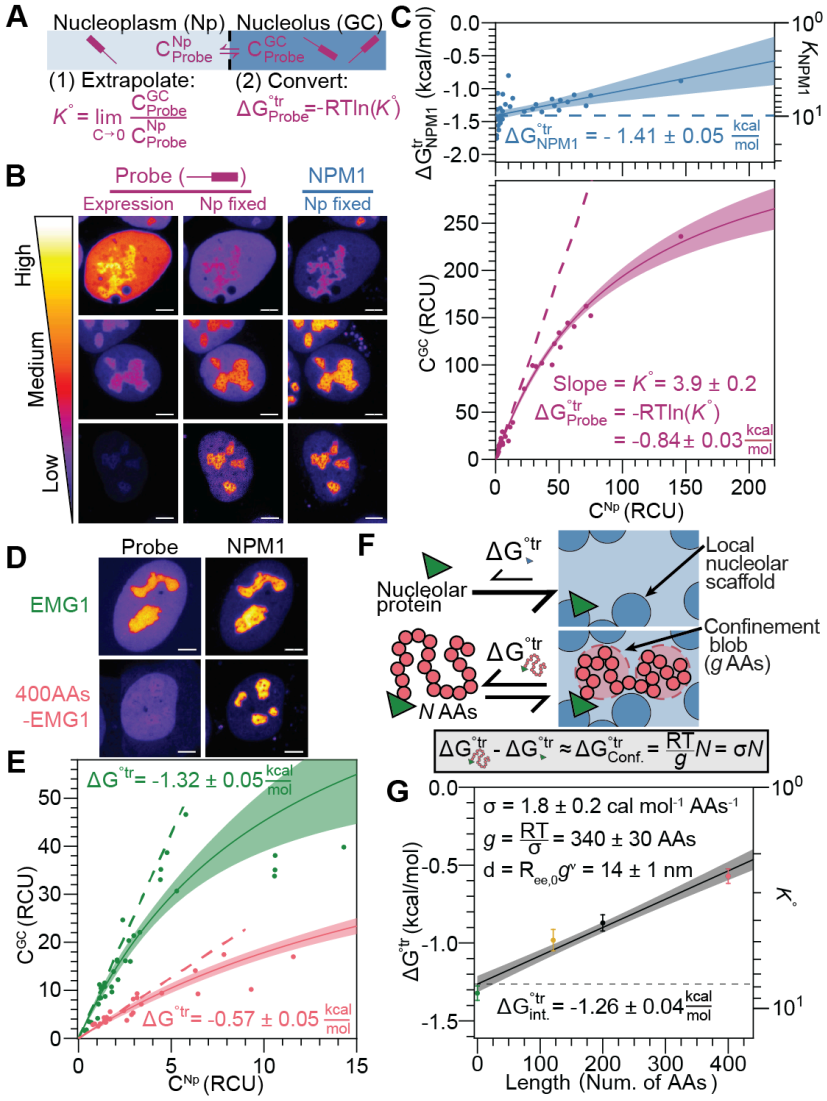

Figure 1: **Local size exclusion measures the mesh size around a specific protein within a condensate.** (**A**) Standard state transfer free energy measurement requires the determination of the concentration, *C*, of a probe molecule between compartments. (**B**) Overexpression of probe, shown NPM1 C-terminal truncation with mGreenLantern (mGL), in U2OS with endogenous NPM1-mCherry. Np fixed sets the value of the nucleoplasm at the same LUTs dictated by a max partitioning of all images. (**C**) Quantification of the probe levels (bottom) and the endogenous NPM1 ΔG^{tr} (top) between the GC and Np phases. Dashed lines show trends extrapolated to zero probe overexpression. Relative concentration units (RCUs) are the number of photons emitted at reference settings (methods). (**D**) Cells with low overexpression of indicated constructs. (**E**) Quantification of indicated EMG1 probes. (**F**) Transfer free energy measurement for a nucleolar protein without and with an *N* residue chain. (**G**) $\Delta G^{\circ tr}$ dependence on the number of chain residues added to the EMG1 probe. Calculation of the apparent mesh size shown (inset). Unless indicated, all images are Np fixed, and scale bars are 5 μm.

Measuring mesh size with *Local Size Exclusion* (*LSE*)

Equipped with the ability to measure the partitioning strength ($\Delta G^{\circ tr}$) of a probe, we sought to exploit it to determine the local organization around a protein. Notably, condensates have mesh sizes due to networks of interactions (*21*). Thus, we wondered if the addition of a chain, i.e. an inert intrinsically disordered region to a nucleolar protein, would result in *Local Size Exclusion* (*LSE*), decreasing its enrichment into the nucleolus. Indeed, the GC contains large ribosomal intermediates and multivalent scaffold proteins, such as NPM1, (*7*, *13*, *30*, *33*) that could form a collective network or meshwork that would locally confine the chain. To test this principle, we employed EMG1, an essential accessory factor of ribosome biogenesis, and expressed it alone or as a C-terminal fusion with a chain of 400 random non-hydrophobic amino acids (AAs). The addition of the 400AA chain results in a clear drop in its enrichment in the GC, weakening the K° and $\Delta G^{\circ tr}$ of EMG1 from 8.5 ± 0.6 and -1.3 ± 0.1 kcal/mol to 2.5 ± 0.2 and -0.6 ± 0.1 kcal/mol (**Fig. 1D-E**). This change can provide a quantifiable metric of the mesh size (see below).

The degree of LSE can be quantified utilizing principles from polymer physics, wherein the mesh size is defined by the region encompassing an average chain of *g* length, known as the confinement blob (*34*, *35*) (**Fig 1F**). Quantitatively, the confinement free energy ($\Delta G^{\circ tr}_{Conf.}$) for a chain of length *N* is proportional to the number of confinement blobs, *N/g*, yielding $\Delta G^{\circ tr}_{Conf.} = \frac{RT}{g}N$; with the slope defined as the σ-value ($\sigma = \frac{RT}{g}$). Thus, the σ-value allows us to calculate *g*, and, by employing scaling laws of unfolded proteins in water (*36*), calculate the approximate mesh size diameter, *d* (i.e., $d \propto g^{\nu}$). Furthermore, the intercept, $\Delta G^{\circ tr}_{int.}$, provides information about the differences in affinity between the phases, including changes in chemical miscibility and binding interactions. To corroborate this formula, we conducted additional measurements with two EMG1 probes, one fused with only half of the 400AA chain and another fused with the disordered region of NPM1 (121AAs). These additional values are consistent with this repulsive and linear free energy of confinement with respect to *N* (**Fig 1G**, methods). From the fit σ-value, we calculate an approximate mesh size around EMG1 as 14 ± 1 nm (**Fig 1G**). Indeed, the observation of a meshwork around EMG1 in the GC is consistent with the notion of the nucleolus as a condensate (*28*).

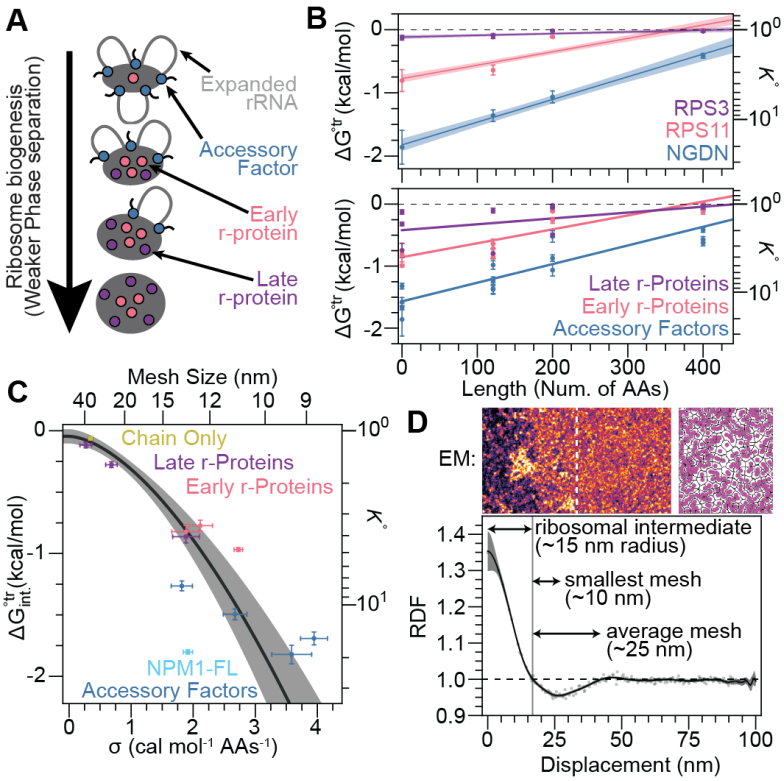

Figure 2: **LSE reveals that the local mesh size increases during ribosome biogenesis.** (**A**) Schematic of ribosome biogenesis and the categories of proteins studied. (**B**) LSE for individual (top) and categorically grouped proteins (bottom, **Fig S3**). (**C**) Relationship between σ and $\Delta G^{\circ tr}_{int.}$. Fit assumes the $\Delta G^{\circ tr}_{int.}$ is proportional to the mesh contact number density ($\propto \sigma^{3\nu}$, methods). (**D**) EM data showing the Np and GC separated by the dashed line (fire mode, top left). Image of segmented grains from EM image (top right). Graph of radial distribution function (RDF) from the center of the GC grains in EM images (bottom, **Fig S5**).

To examine the generality of the EMG1 results as well as the robustness of LSE, we applied LSE to nucleolar proteins, focusing on proteins at different stages of ribosome biogenesis that occur in the nucleolus (*37*): (1) additional accessory factors NGDN, IMP3, and RCL1, (2) early ribosomal proteins RPL7L1 (uL30), RPS14 (uS11), and RPS11 (uS17), and (3) late ribosomal proteins RPL10A (uL1), RPS3 (uS3), and RPL12 (uL11) (**Fig 2A, S3A**). Expression of these proteins alone yielded a significant and expected (*7*, *33*) correspondence between the $\Delta G^{\circ tr}$ and the stage of ribosome biogenesis ($p<0.001$, **Fig 2B-C, S3B**). Additionally, like EMG1, adding chains resulted in LSE, decreasing GC partitioning (increase in $\Delta G^{\circ tr}$) for all proteins (**Fig 2B, S2**). For a homogenous liquid, the mesh size and, therefore, the σ-values are expected to be nearly constant (*21*). Unexpectedly, the sensitivity to chain length varied with σ-values (slope) ranging from 0.1 to 4 cal/mol/length, corresponding to approximate mesh sizes ranging from 10 to 50 nm, and with a significant correlation to ribosome biogenesis ($p<0.01$, **Fig 2C**). The $\Delta G^{\circ tr}_{int.}$ and σ-values were significantly correlated ($p<0.001$), consistent with a model whereby the

valence of the ribosomal intermediate (inferred vis-a-vi $\Delta G^{\circ tr}_{int.}$) dictates the number of mesh contacts ($\propto \sigma^{3\nu}$) and vis-versa (fit in **Fig. 2C**, methods, also **Fig S4**). These data are inconsistent with the GC being a homogenous liquid and instead suggest a gradual change in the local environment dictated by the progression of ribosome biogenesis.

Surprised, we sought to further interrogate the inhomogeneity of the GC. First, we focused on NPM1, a generic scaffold protein with rapid dynamics (*13*) and without a direct role in ribosome biogenesis, measuring a σ-value of 1.9 ± 0.1 kcal/mol/length or 13.2 ± 0.3 nm in line with the local meshwork observed in accessory factors and early ribosomal proteins (**Fig 2C, light blue)**. When we employed the chains alone, utilizing additional random chains at each length to avoid systematic issues of charge and AA composition (**Tabel S3**), we measured a σ-value of 0.35 ± 0.03 kcal/mol/length or 34 ± 2 nm in line with the late ribosomal proteins and substantially different from EMG1 and the other accessory factors (**Fig 2C, gray)**. Indeed, the discrepancy between the mesh size of chains and that of NPM1, further exhibits inhomogeneity from 10-50 nm consistent with that seen in the ribosome biogenesis factors.

Our measured mesh sizes are large (10-50 nm) in contrast to those measured in reconstituted droplets (2-4 nm) (*21*, *38*). Thus, we sought to orthogonally validate them. As such, we turned to electron microscopy (EM) as the GC, or “granular component”, is named for EM-observed granules that correspond to the ribosomal intermediates (*39*). Utilizing available high-resolution 4 nm isotropic EM data (*40*), we identified the GC, segmented granules, and computed the radial distribution function (RDF, (*13*)), revealing a broad distribution of sizes that are consistent with our range (**Fig 2E**). Similar analyses of the electron cross-correlation and the histogram of the diameter between the surface of the granules confirmed the RDF (**Fig S5**). Indeed, previous studies have identified the existence of proteinaceous interstitial regions interspersed within the GC that are depleted for granules and RNA (*41*). All together, the agreement of these orthogonal lines of evidence, provides validation to our method and its ability to measure the local mesh size around specific proteins.

Distance-dependent preferences between proteins inferred with *Length-Dependent Spatial Coupling* (*LDSC*)

Given the degree of organization implied by our LSE results, we sought to determine how much condensate structure contributes to protein localization. Again, we turned to NPM1 for its rapid dynamics. To proceed, we linearly decomposed the transfer free energy of NPM1, separating the energetic contributions by its domains (**Fig. 3A, S6A**). This allows the quantification of the intrinsic role of each domain and the influence of couplings between different domains, as done in protein folding for domains or other terms (*27*, *42*). We determined the value for these terms by measuring the $\Delta G^{\circ tr}$ of all NPM1 truncations and performing the appropriate linear calculations (**Fig. 3B,** methods**, S6B**). As expected, the analysis reveals that the pentamerization domain has the strongest role in NPM1 recruitment to the GC as it includes oligomerization with the four other full-length NPM1 molecules (**Fig 3B, blue-left**). We also

observed that the C-terminal RNA binding domain, i.e., the nucleolar localization sequence (RBD/NoLS), contributes favorably to GC partitioning (**Fig 3B, blue-right**). The repulsive energetic terms include the nearest-neighbor couplings, being the impact of NPM1's IDR fused to NPM1's other domains, and the intrinsic domain contribution of NPM1's IDR; these are consistent with the aforementioned role of the free energy of confinement of an AA chain (**Fig 3B, blue-middle, green**). Strikingly, the spatial coupling ($\Delta G^{\circ tr}_{Sp.Co.}$) between the N and C terminal domains promotes NPM1 localization to the GC slightly more than the domain contribution of the NoLS/RBD (**Fig 3B, magenta vs. blue-right**). Indeed, this shows that the structure of the condensate should not be ignored when considering the energetic determinants that drive proteins to partition into condensates.

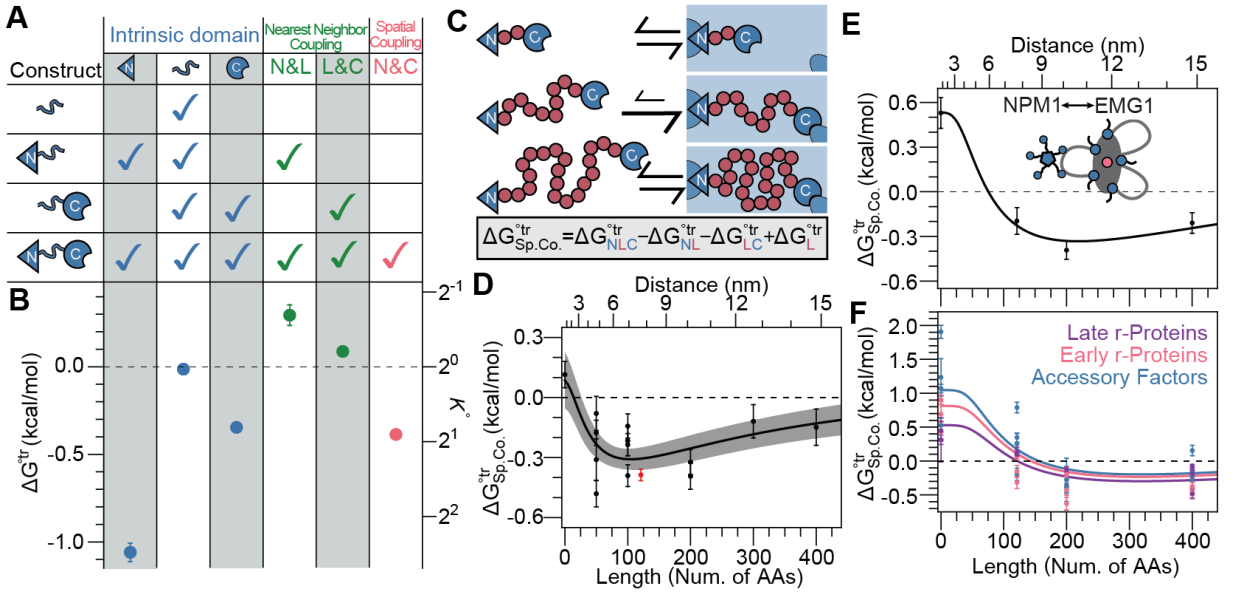

Figure 3: LDSC measures the distance between proteins. (**A**) Linear breakdown of the $\Delta G^{\circ tr}$ for each NPM1 construct by its domains into energies as indicated. (**B**) Results of linear decomposition of the $\Delta G^{\circ tr}$ (**Fig S6**). (**C**) Schematic depicting the spatial coupling (Sp.Co.) as a function of chain length (top) and its calculation (bottom). (**D**) Length-dependent Sp.Co. (LDSC) between the N and C terminal domains of NPM1. Chain (red) of NPM1; not used in fit. (**E**) LDSC between EMG1 and NPM1. (**F**) LDSC between NPM1 and proteins grouped by role in ribosome biogenesis. For analytical form in D-F see methods.

Inspired by the strong spatial coupling between NPM1's termini, we developed the *Length-Dependent Spatial Coupling* (*LDSC*) approach, which can measure the preferential distance-dependent organization between two proteins (or protein domains) within a condensate (i.e., g(r), methods). Indeed, if the spatial coupling is due to a preferred organization of biomolecules around NPM1 then it should decrease both if the number of residues in NPM1's IDR is increased or decreased due to the elastic nature of a polymeric chain (**Fig 3C,** methods). To test this, we measured the spatial coupling between NPM1's terminal domains by replacing NPM1's IDR with different polymeric chains, employing those utilized for LSE above (**Fig 3D**). Fitting yields a preferred organization at a chain length of 130 ± 13 residues and an end-to-end

distance of 8.0 ± 0.4 nm. This preferential length is within error of the 121 residues in NPM1's IDR (**Fig 3D,** red), suggesting that the length of NPM1's IDR is optimized for the location of RNA around it. Furthermore, this preferential RNA organization is similar to the 5 ± 1 nm difference in the mesh size obtained with LSE between the NPM1 full length and N-terminus alone (**Fig 2C**). This demonstrates the ability of LDSC to measure local preferential structure between domains.

Next, we employed the LDSC approach to determine how proteins involved in ribosome biogenesis are organized around (full-length) NPM1. As such, we determined the spatial coupling between NPM1 and the ribosome biogenesis proteins fused with different chains. Starting with EMG1, we found that the spatial coupling with NPM1 was repulsive without a chain but became more attractive with one (**Fig 3E**). For each probe we tested, we found that having a zero residue chain was highly unfavorable, and the degree of repulsion significantly correlated with ribosome biogenesis ($p<0.001$, **Fig 3F, S6B-C**, 1.1, 0.8, and 0.3 ± 0.1 kcal/mol for the three protein types). While at the inferred ~7.5 nm, the spatial coupling was barely repulsive (0.13 ± 0.02 kcal/mol), at ~10 nm and ~15 nm, it became favorable (-0.26 ± 0.02 and -0.22 ± 0.02 kcal/mol, respectively). Fitting globally yielded a preferential distance at a chain length of 340 ± 30 residues, corresponding to an end-to-end distance of 14 ± 1 nm (**Fig S6D**). Additionally, we measured a repulsive spatial coupling between NPM1's N-terminus and all proteins, suggesting that the organization is preferentially away from the center of the NPM1 pentamer (**Fig S6**). These data suggest that NPM1 does not directly solvate accessory factors and ribosomal proteins, consistent with our previous models that NPM1 participates in GC formation through heterotypic interactions with rRNA not proteins (*7*, *13*). Furthermore, the higher exclusion of NPM1 during early ribosome biogenesis stages suggests that NPM1 is specifically excluded from the step-wise assembly process, consistent with a general nucleolar scaffolding role that is in the spirit of phase separation. Regardless, these data indicate that the GC has clear organization and, more broadly, that the LDSC technique enables the elucidation of nanometer-level information within a condensate between two specific proteins.

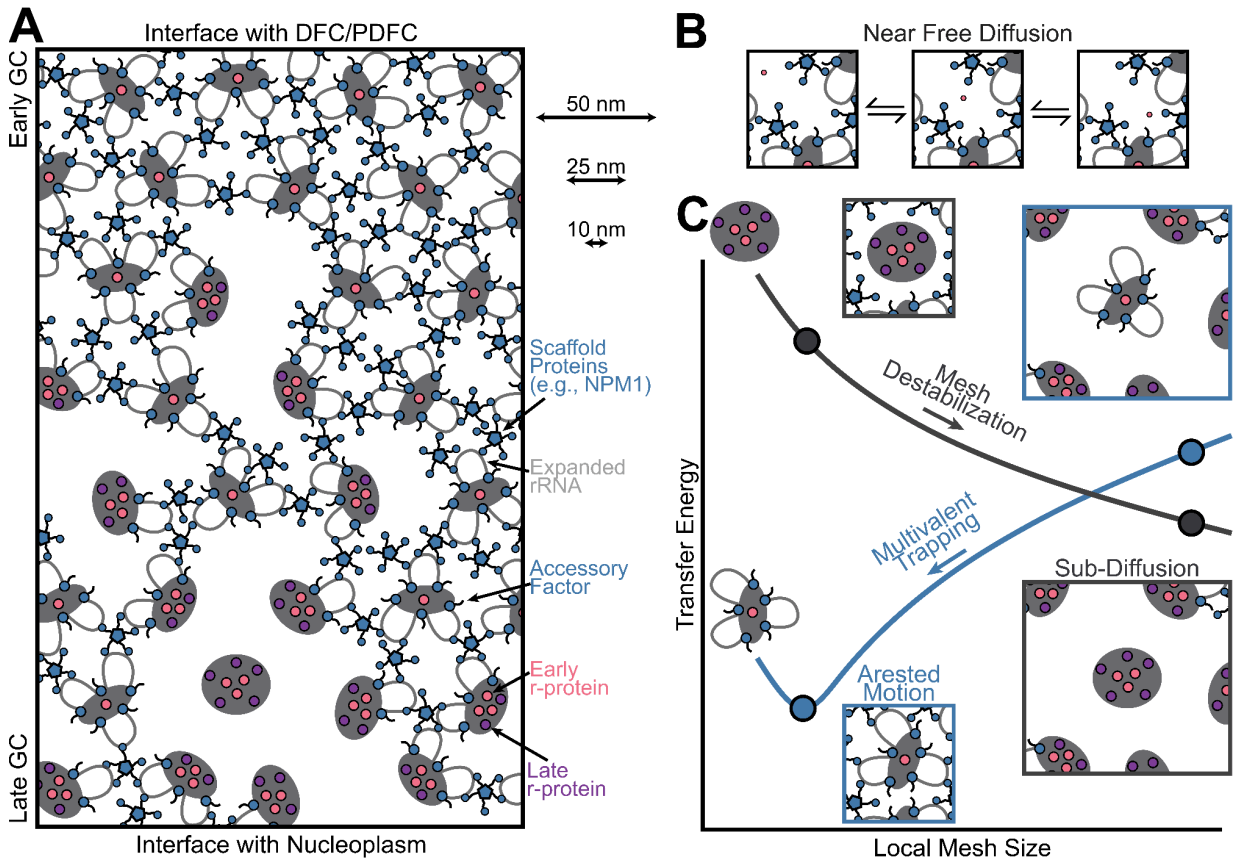

Figure 4: LSE and LDSC provide a nanometer picture of molecular interactions and transport during ribosome biogenesis. (**A**) Gradient of ribosome biogenesis and mesh sizes within the GC. Note that client molecules, including scaffold proteins that are in client-like states, are omitted to highlight mesh. (**B**) Depiction of free early r-protein transport within the GC meshwork. (**C**) Local mesh size dependence on transfer energy and transport for an early ribosomal intermediate (blue) and the completed ribosome (gray) due to local changes in the number of interactions (**Fig S4**, methods).

Our measurements enable an unprecedented structural picture of the nucleolus and its interplay with ribosome biogenesis (**Fig 4**). The local mesh size around earlier intermediates, inferred here via measurements on accessory factors and early ribosomal proteins (**Fig 2B-C, 4A**), is between 10-15 nm. Not only is this value substantially higher than those observed in reconstituted systems (~3 nm, (*21*, *38*)), but it lies between that of a folded protein domain (~3-5 nm) and the diameter of ribosomal intermediates (~30 nm, **Fig 2D**). This implies that, in contrast to ribosomal intermediates, proteins can easily transport through the mesh via free diffusion or reptation for single and multidomain proteins, respectively (**Fig 4B**), consistent with existing data (*13*, *43*). Ribosomal intermediates initially form a structured core surrounded by multiple expanded rRNA stretches that progressively collapse during biogenesis in the nucleolus (**Fig 2A, 4A**) (*33*); here we determine that this step-wise process results in the sequential destabilization of the energetic preferences of the local meshwork (**Fig S4, 4C**), eventually deforming it to a measured size of around 40-50 nm - and thus, larger than the ribosome. This mesh destabilization switches the transport from arrested, due to polyvalent trapping likely from

adhesive (heterotypic) interactions between scaffold proteins (e.g., with NPM1) and expanded rRNA segments (**Fig 3**), to sub-diffusive (**Fig 4C**), analogous to other systems (*44*). These structural inferences provide insight and are consistent with the disconnect between the dynamics of proteins and ribosomal intermediates in the GC (*13*), the existence of constrained (rRNA) territories around individual rDNA-containing chromosomes (*45*), and the inhomogeneity of GC organization observed in previous EM studies as so-called interstitial voids (*41*). Finally, our data suggests heterotypic interactions are quite distal (~14 nm) from the location of ongoing assembly (**Fig 3E-F**), providing a mechanistic blueprint for minimizing competition between phase separation and ribosome biogenesis.

These inferred principles yield a general nanometer-scale structure-function relationship for the nucleolus that imbues *selective transport* for proteins and completed ribosomal subunits, not ribosomal intermediates (**Fig. 4**). Similar to the “selective transport model” in the nuclear pore (*5*, *44*), our model ties the transport to deformation of the mesh underlying the interactions that drive condensate formation. Additionally, in cells, the selective transport model would imply a similar inhomogeneity of mesh sizes due to the nuclear transport receptors deforming the mesh to facilitate transport. However, differences in selectivity are clearly focused on functional differences between the nuclear pore and the nucleolus, including mesh sizes that exclude and include most proteins, respectively. Thus, we anticipate that condensate properties have been selected to couple condensate structural features with functional roles.

The complexity of biological condensates *in situ* has made elucidating their properties difficult and controversial. Indeed, there has been much debate about how to prove that something qualifies as a condensate or forms through LLPS (*46*, *47*) and what interactions drive formation or composition (*48*, *49*). This includes warranted criticisms of measurements on droplets induced by oligomerization, overexpression in cells, or LLPS in test-tube reconstitution from only a few components (*5*, *50*, *51*); our approaches, however, aim to mitigate such concerns by characterization in live cells and extrapolating to zero expression of the protein probe of interest. Ultimately, our methodologies have the potential to transcend the investigation of condensates, offering invaluable insights into the energetics of other cellular microenvironments.

The framework presented here lays the groundwork for a quantitative understanding of condensates by elucidating the energetic determinants that drive proteins into condensates. Indeed, the specific applications of LSE and LDSC detail how designed probes with quantitative microscopy can be employed toward determining structure-function relationships, i.e., how the nanometer structure and organization of biomolecules within condensates and the interplay with molecular biology and biochemistry within them. As suggested by EM of many condensates (*15*), we anticipate that other condensates beyond the GC phase of the nucleolus will be inhomogeneous and with local organization and structure. Such structure may help with regulation and proofreading, facilitating the trapping of substrates and intermediates and specific release of products and would have implications for changes in the context of disease where condensates are compromised (*28*, *52*).

Acknowledgments: We thank members of the Riback Lab for discussions and feedback on the manuscript. We thank Steven Boeynaems, Kyle Eagen, Alex Holehouse, Paulo Onuchic, and Tobin Sosnick for feedback on the manuscript. We thank Margaret Goodell for assistance and advice with CRISPR editing, including access to equipment. J.A.R. is a CPRIT Scholar in Cancer Research. This work was supported by funding from the Cancer Prevention and Research Institute of Texas (CPRIT) New Investigator Grant RR210040 (J.A.R.), the Ted Nash Long Life Foundation (J.A.R.), the Leukemia Research Foundation (J.A.R.), the Searle Scholars Program (J.A.R.), NIH F30CA268725 (G.K.D.), NIH CA183252 (Goodell PI), NIH R01GM141235 (J.D.S). This work was also supported by the Cytometry and Cell Sorting Core with funding from CPRIT and the NIH (RP180672; CA125123, RR024574).

Author Contributions: Conceptualization, J.A.R.; Investigation, C.D., E.P., A.M., A.A., and G.D.; Methodology, C.D., E.P., A.M., A.A., and G.D.; Data curation, C.D. and J.A.R.; Formal analysis, J.D.S. and J.A.R.; Writing - original draft, C.D., E.P., and J.A.R.; Writing - review & editing: C.D., E.P., A.M., A.A., G.D., J.D.S., and J.A.R.; Supervision, J.A.R.

Competing interests: None

Data and materials availability: All data can be found at the lab's github (https://github.com/CellularPhysicalChemistryGroup).

Materials and Methods:

Cell culture

U2OS were cultured in sterile filtered McCoy's 5A Modified Medium (1X) with L-Glutamine (Gibco 16600-082) and supplemented with 10% FBS (Avantor 89510-186) and 1% Penicillin–Streptomycin (Gibco 15140-122). HEK293 cells were cultured with DMEM (1X) with 4.5 g/L D-Glucose and L-GLutamine (Gibco 11965-092). Cells were kept in a humidified incubator (VWR, PHCbi MCO-170AICUVDL-PA) with 5% CO_2 and at 37°C. CO_2 levels were calibrated every six month sterilization cycle. Temperature was checked every week using a thermometer. For routine passaging, cells were washed with PBS pH7.2 (1X) (Gibco 20012-043), dissociated with trypsin (Trypsin-EDTA 0.25%, Gibco 25200072) then plated into a T25 cm^2 flask. For imaging, cells were plated into an 8-well (Ibidi 80807) or 96-well plate (Cellvis, P96-1.5H-N). Mycoplasma testing on all cell cultures was conducted every two weeks using ABM's Mycoplasma Test Kit and protocol (Applied Biological Materials G238).

CRISPR protocol

We used CRISPR to knock in mCherry at the C terminus of NPM1 in U2OS. CRISPR synthetic guide RNAs (sgRNAs) were ordered to target the NPM1 (Synthego), and a mCherry HDR template was purchased from Twist. 100-200 bp homology arms were added to the mCherry template via PCR with Kappa HiFi HotStart Ready Mix (Roche 501965217) (**Table S1**). The resulting product was purified using Ampure XP beads, according to the vendor's protocol. U2OS cells were then prepared in Buffer R (Thermofisher BR5) for electroporation. Approximately 6 X 10^4 cells were mixed with 1µg of Cas9 (PNA Bio CNP1) and sgRNA and then incubated at room temperature for 20 minutes. Subsequently, 1µg of amplified homology template was added. The cells then underwent electroporation using the Neon Transfection System (Thermofisher NEON18S) at 1350 V, 35 ms, and 1 pulse. After recovery, the cells were sorted and then validated using primers designed for genotyping (**Table S1**). Genomic DNA was extracted using a NEB kit (T3010L), and target sequences were amplified via PCR using genotyping primers. The PCR products were then analyzed on an agarose gel. The results showed a homozygous edit with both NPM1 wild type tagged with mCherry at a band of approximately 800 bp.

Construct production

To facilitate the production of constructs in the manuscript, we developed a custom vector editing pTwist Lenti SSFV, replacing the standard multicloning site with a sequence containing three cloning sites for Gibson assembly with a monomeric fluorescent protein, mGreenLantern (mGL, (*53*)) located directly prior to the second site (**Table S2**). The three sites (N-terminal, Chain, and C-terminal), have standard overlaps for each site such that an insert can be put at a specific site in any vector where that site is still free (i.e., insertion of a sequence in a site does not impact overlaps for other sites). Vectors cut with either Nhe1 (NEB R3131L), Afe1 (NEB R3552S), or Spe1 (NEB R3133S) are mixed with inserts containing appropriate overlaps with reagents for Gibson Assembly (NEB E2621S). Single colonies were picked, and constructs were validated by full plasmid sequencing with Plasmidsaurus. Typically, inserts were added to the N-terminal and/or the C-terminal site prior to the chain site due to the higher efficiency and

lower background of the chain site. This cloning scheme allows for modular insertion at multiple sites, accelerating production of the 148 constructs used and built upon in this manuscript.

Chains were random, generated from an equal probability of all of the AAs except aliphatic and aromatics (histidine, cysteine, alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine, tryptophan) (**Table S3**). DNA for four of these randomized inert 100 residue sequences were ordered from Twist Biosciences and cloned into the vector into the chain site. After sequence validation, chains were transfected into U2OS cells to validate expression and confirm the absence of substantial nucleolar enrichment or depletion. From the 100AA chains all of the others were derived except for NPM1's IDR, which was used as a model for a common nucleolar IDR (**Table S3**).

The ribosome biogenesis proteins employed were identified for their experimentally determined roles in this process (accessory factors, early, and late ribosomal proteins) (*37*). These proteins were chosen due to their localization in the GC and whether they would fit into the custom vector cloning site framework. The proteins were also chosen for their tag site in previous literature (e.g. C-terminal site for all except for NPM1) (**Table S4**). Primers for each were created using a reference genome. Sequences of canonical isoforms were amplified from the cDNA of U2OS and HEK293 cells.

High-resolution live-cell microscopy

All images were taken on a microscope setup containing a Nikon's Ti2E, a VisiTech instant SIM microscope (VT-iSIM), a Hamamatsu ORCA Quest qCMOS camera, and a CFI60 Plan Apochromat Lambda D 100X oil immersion objective lens. Cells were maintained in a 5% CO_2 at 37°C chamber and imaged with immersion Oil Type B 37°C (Nikon 77005) during all imaging sessions. Imaging of green (mGL) and red (NPM1-mCherry) was done sequentially via excitation with 505 nm and 561 nm laser, respectively. No detectable bleed-through was observed.

For transient transfection, cells were plated at 20-25,000 per cm^2 in an 8-well plate. When transfecting, 0.25-5µg plasmids (500-5000 ng/µL), 1 µL Lipofectomine P3000 reagent, 0.75 µL Lipofectomine (Thermofisher L3000008), and 12.5 µL Optimem 1X (Gibco 31985-070) was used per well. Transfection was performed following the vendor's standard protocol. Cells were imaged 48-72 hours after transfection.

Power meter measurements were taken at the focal plane for both channels before imaging to verify the proper alignment and performance (e.g., sufficient warming up). Dye standards were also used approximately once a month to determine the no-emission background (due to factors such as scattering of the excitation light) and vignetting needed for flat field correction for all lasers at all power settings. Additionally, this procedure determined the digital level (DL) offset and the conversion between DL and photoelectrons (via fitting low light photon counts histogram) for all camera settings. The camera linearity was verified manually.

Determination of ROIs and calculation of RCU

All images acquired were analyzed using custom plugins employed in Micro-Manger and imageJ/Fiji. This plugin facilitated regions of interest (ROIs) selection for measurements of the granular component (GC) of the nucleolus and the nucleoplasm (Np). As we have done previously (*7*, *13*). Np ROIs were selected manually using the 561 and 505 nm channels as

circles or polygons (to avoid out-of-plane nucleoli) and the GC of the nucleolus was selected automatically by finding the center of the 5 pixel gaussian blur max and using a circle ROI with a radius of 5 pixels in the endogenous NPM1-mCherry (561 nm channel) for analysis. As much as possible (e.g. except when confirming probe expression at the desired levels), this step was blinded to the probe channel. Relative concentration units (RCUs) were then determined by averaging the DLs in the ROIs, subtracting no-emission/no-sample background and dividing by the average of the flat field calibration. Subsequently, a sample-specific background was also subtracted. RCU values were sufficiently reproducible day to day - with some values being measured more than a year apart.

Electron Microscopy analysis

Electron microscopy data of an interphase HeLa cell was analyzed from (*40*). A single plane was identified where the boundary between the nucleolus and the nucleoplasm could be determined. Centers and RDF analysis was done similarly as published in (*13*). Briefly, a dilated binary of the image was created by binarizing the 2-pixel blurred nucleolar image at the halfway value between the 0.005 and 0.995 quantiles and then applying a two-pixel disk dilation. Centers were identified by finding the MaxDetect of the two-pixel blurred distance transform of the dilated binary. For segmentation (**Fig S5**), the centers were dilated by the average ribosome intermediate diameter and a skeleton transformation was done. The CDF was fit with 20 Berstein bases filtered by using the most substantial 97.5% error

A note of caution on the use of the term steady-state transfer free energy.

We choose to employ the ◦ in both K° and $\Delta G^{\circ tr}$ to represent the partitioning and transfer free energy for the constructs as they exist in the unperturbed dilute and dense phases (i.e. nucleoplasm and nucleolus). Here we intend this to refer to that without overexpression in the average interphase cellular state. Additionally, we designate this "◦" the "apparent standard state". Application of these terms from physical chemistry should be considered with some caution, especially in light of previous debates on the use of such terms (*54*, *55*). Indeed the notion that the regions studied are true-phases and in steady-state are likely both incorrect to some extent (*5*, *13*). While corrections and details may be later employed to make these better thermodynamic values, we hope the totality of the approaches presented in the manuscript show the value such frameworks provide.

Extraction of the steady-state transfer free energy.

To determine the steady-state $\Delta G^{\circ tr}$ and ΔG^{cmax} for a specific probe, single cell concentrations (C_{Np} and C_{GC} , as measured above) were empirically fit to a hyperbola:

$$C_{GC} = \frac{C_{Np}}{\frac{1}{\exp(\frac{-1}{RT}\Delta G^{\circ tr})} + \frac{C_{Np}}{\exp(\frac{-1}{RT}\Delta G^{cmax})}}$$

$$\frac{1}{\exp(\frac{-1}{RT}\Delta G^{\circ tr})} + \frac{C_{Np}}{\exp(\frac{-1}{RT}\Delta G^{cmax})}\frac{-1}{RT}\Delta G^{\circ tr}$$

This equation was chosen as it has the following limits:

$$\lim_{C_{Np}\to 0} C_{GC} = C_{Np}\exp(\frac{-1}{RT}\Delta G^{\circ tr})$$

and

$$\lim_{C_{Np}\to \infty} C_{GC} = \exp(\frac{-1}{RT}\Delta G^{cmax})$$

The fit was done with an iterative routine to eliminate outliers that may bias the fit to various non-gaussian noises. Specifically, each cell was weighted by the inverse of the total error being the deviation from the previous iteration fit (sum of squares including both X and Y deviations equally). Deviations higher than a 1% probability from the median error (~2.6 fold higher) were down-weighted by the square root of the inverse of the fold error. At double that value, they were down-weighted by simply the inverse. In 8 of 148 cases, it was manually determined that the deviation was still too large and fitting was adjusted by focusing the fit on a specific range. Errors were approximated using default settings in Mathematica. Partitioning (K°) and *Cmax* were solved by dividing by *-RT* and applying the exponential to $\Delta G^{\circ tr}$ and ΔG^{cmax}, respectively.

Note that here, steady-state is meant in the most general way to encompass multiple states, non-equilibrium effects, and other complications that may complicate the equilibrium nature and standard assumptions assumed in the test tube. Future works will aim to remove additional complexities of cells; despite limitations, we stress that our underlying assumption is that the errors from these cellular complexities will be less than the errors by reconstitution of facsimiles in the test tube or optogenetic manipulation in cells.

Extraction of sigma and the approximate mesh size.

To extract sigma (σ) and intercept ($\Delta G^{\circ tr}_{int}$) for each protein in LSE, a linear fit to the $\Delta G^{\circ tr}$ as a function of chain length was employed where the VarienceEstimaterFunction is the median variance of the individual data points. We show the standard deviation on these values when plotting (e.g. set the Confidence level to ~0.7). After converting to 'g' AAs as described, $d = Ree0\, g^{\nu}$ is employed where *Ree0* is 0.55 nm and ν is 0.55 as we have found previously for proteins in water (*36*).

Complexities for the confinement free energy employed here should be noted. We have treated the meshwork in the nucleoplasm ($\sigma_{Np} \approx 0$) as far greater than the nucleolus (*15*). The values reported here are only in fact the average, thus:

$$\Delta G^{\circ tr}_{Conf.} = -RT\ln\left[\int_0^{\infty} P(\sigma) e^{\frac{-\sigma N}{RT}} d\sigma\right]$$

$$= \left(\int_0^{\infty} \sigma\, P(\sigma)\, d\sigma\right) N + \left(\left(\int_0^{\infty} \sigma\, P(\sigma)\, d\sigma\right)^2 - \left(\int_0^{\infty} \sigma^2\, P(\sigma)\, d\sigma\right)\right) \frac{N^2}{2RT} + \ldots$$

by taking the taylor expansion around the chain length *N* and where $P(\sigma)$ is the normalized distribution of local confinement around the protein with a *N*=0 length chain. Thus, here we assume that the first term dominates. Note this has implications for *g* and *d* as a similar derivation will note that the average g is $\frac{1}{\int_0^\infty \frac{1}{g} P(g)\, dg}$ (harmonic mean).

Relationships between sigma, the density of mesh contacts, and the transfer free energy.

To understand the principles governing the relationship between (σ) and the transfer free energy of a protein ($\Delta G^{\circ tr}_{int}$), we took a mean-field approach. Specifically, the $\Delta G^{\circ tr}_{int}$ should be proportional to the average density of local mesh contacts (δ). Using simple scaling arguments, we can determine the relationship between δ and σ :

$$\delta \propto d^{-3} \propto g^{-3\nu} \propto \sigma^{3\nu}$$

using the relationships that density is diameter cubed, diameter is chain length (*g*) to the ν , and g is inversely proportional to σ. Thus, a linear fit of $\Delta G^{\circ tr}_{int}$ as a function of $\sigma^{3\nu}$ using the sum of squares error with 50 iterations to allow for proper propagation of errors using the best fit from previous iterations. The standard deviation of the fit is shown by using a confidence level of 0.7.

Note after including necessary conversions, σ can also be converted into a rough scaffold contact concentration with the relation $0.25\,\sigma^{3\nu}$ having units of millimolar, and yielding values of 1.5, 1, and 0.2 mM for the accessory factors, early r-Proteins, and late r-Proteins respectively. Including proper conversions yields the slope $-1.1 \pm 0.2\ \frac{kcal}{mol} mM$ for the relation between $\Delta G^{\circ tr}_{int}$ and $\sigma^{3\nu}$.

Linear decomposition of transfer free energy by domains.

To elucidate the energetics driving NPM1 and constructs employed in the LDSC approach, we employ the following formalism. Notably, we focus on the 3-domain structure which we will abbreviate “N” - “L” - “C” as done in the schematics (**Fig 3, S6**). To proceed we split the free energy of the experimental constructs into three domain-wise terms: 1) intrinsic free energies ($\Delta G^{\circ tr}_{intr.}$), 2) nearest-neighbor coupling free energies ($\Delta G^{\circ tr}_{NN.Co.}$), and 3) next-nearest-neighbor couplings that we call spatial couplings free energies ($\Delta G^{\circ tr}_{Sp.Co.}$).

We use at most six constructs (here designated by the inclusion of the N,L, or C domains where the fluorescent protein (FP) is as depicted in **Fig S6**):

$$\Delta G^{\circ tr}_{FP} = \Delta G^{\circ tr}_{intr.\,FP}$$
$$\Delta G^{\circ tr}_{L} = \Delta G^{\circ tr}_{intr.\,L} + \Delta G^{\circ tr}_{intr.\,FP}$$
$$\Delta G^{\circ tr}_{N} = \Delta G^{\circ tr}_{intr.\,N} + \Delta G^{\circ tr}_{intr.\,FP}$$
$$\Delta G^{\circ tr}_{C} = \Delta G^{\circ tr}_{intr.\,C} + \Delta G^{\circ tr}_{intr.\,FP}$$

$$\Delta G^{\circ tr}_{NL} = \Delta G^{\circ tr}_{intr.\,N} + \Delta G^{\circ tr}_{intr.\,L} + \Delta G^{\circ tr}_{intr.\,FP} + \Delta G^{\circ tr}_{NN.Co.\,N\&L}$$
$$\Delta G^{\circ tr}_{LC} = \Delta G^{\circ tr}_{intr.\,L} + \Delta G^{\circ tr}_{intr.\,C} + \Delta G^{\circ tr}_{intr.\,FP} + \Delta G^{\circ tr}_{NN.Co.\,L\&C}$$
$$\Delta G^{\circ tr}_{NLC} = \Delta G^{\circ tr}_{intr.\,N} + \Delta G^{\circ tr}_{intr.\,L} + \Delta G^{\circ tr}_{intr.\,C} + \Delta G^{\circ tr}_{intr.\,FP} + \Delta G^{\circ tr}_{NN.Co.\,N\&L} + \Delta G^{\circ tr}_{NN.Co.\,L\&C} + \Delta G^{\circ tr}_{Sp.Co.}$$

this allows us to solve the solution for each of the three terms in the linear decomposition:

$$\Delta G^{\circ tr}_{intr.\,FP} = \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{intr.\,C} = \Delta G^{\circ tr}_{C} - \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{intr.\,N} = \Delta G^{\circ tr}_{N} - \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{intr.\,L} = \Delta G^{\circ tr}_{L} - \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{NN.Co.\,N\&L} = \Delta G^{\circ tr}_{NL} - \Delta G^{\circ tr}_{N} - \Delta G^{\circ tr}_{L} + \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{NN.Co.\,L\&C} = \Delta G^{\circ tr}_{LC} - \Delta G^{\circ tr}_{L} - \Delta G^{\circ tr}_{C} + \Delta G^{\circ tr}_{FP}$$
$$\Delta G^{\circ tr}_{Sp.Co.} = \Delta G^{\circ tr}_{NLC} - \Delta G^{\circ tr}_{NL} - \Delta G^{\circ tr}_{LC} + \Delta G^{\circ tr}_{L}$$

Note that we have ignored the couplings between the fluorescent protein and other domains for experimental practicality. Furthermore, in the spatial coupling these terms cancel out making the LDSC maximally robust to complexities of the fluorescent tag.

Fitting the LDSC.

The structure of liquids is quantified in statistical mechanics by the radial distribution function often called the pair-correlation function denoted $g_{AB}(r)$ or simply $g(r)$, a term related to the potential of mean force. $g(r)$ is the probability of finding a component A at a radial distance r away from a component B. Here we conjecture that the spatial coupling can be approximated in terms of $g(r)$ specifically:

$$\Delta G^{\circ tr}_{Sp.Co.} \approx -RT\ln\left[\int_0^\infty P(r|N)g(r)dr\right]$$

where $P(r|N)$ is the end-to-end distribution for a chain of *N* residues that are fused between A and B.

Practically, to fit $\Delta G^{\circ tr}_{Sp.Co.}$, we focus on two terms, one a distance independent term (related to either A and B bound but not both) and a chain length dependent term with a preferential binding at a specific distance. Thus:

$$\Delta G^{\circ tr}_{Sp.Co.} = -RT\ln\left[e^{-\frac{\Delta G_{indep}}{RT}} + e^{-\frac{\Delta G_{dep}(N)}{RT}}\right]$$

with

$$\Delta G_{dep}(N) = \Delta G_{dep}(0) + RT\left(\frac{3\lambda}{2N} + \frac{3}{2}\ln[N] - \ln[\lambda]\right)$$

where λ is the length of AAs at the preferred distance and *N* is the length of the chain between A and B of the construct. Note this dependence is derived assuming a gaussian end-to-end distribution between the two linkers (*34*, *35*). Additionally, to account for a few residues of linker and the fluorescent protein, the equation is fit with *N* values increased by 25 residues. When globally fitting for NPM1 LDSC we add a $\sigma^{3\nu}$ dependent term to ΔG_{indep}.

Model for local energetics depending on mesh size for ribosomal intermediate.

To approximate the local environment and ribosomal intermediate dependence on the transfer free energy, we decomposed the energy by the number of contacts between the ribosomal intermediate and the nearby mesh molecules (N_{Cont}), the number of unsatisfied or free contacts of the ribosomal intermediate (N_{RI}^{Free}), and the mesh molecules (N_{Mesh}^{Free}) yielding the following equation:

$$\Delta G^{\circ tr}(N_{Cont}, N_{RI}^{Free}, N_{Mesh}^{Free}) \approx \Delta\epsilon_{Cont} N_{Cont} + \Delta\epsilon_{RI}^{Free} N_{RI}^{Free} + \Delta\epsilon_{Mesh}^{Free} N_{Mesh}^{Free}$$

Intuitively, our picture of the nucleolus implies that contacts between mesh and expanded rRNA segments are favorable while unsatisfied interactions are unfavorable ($\Delta\epsilon_{Cont} < 0$, $\Delta\epsilon_{RI}^{Free} > 0$, and $\Delta\epsilon_{Mesh}^{Free} > 0$). Now, to simplify, we employ the relationships $N_{RI}^{Free} = N_{RI}^{Max} - N_{Cont}$ and $N_{Mesh}^{Free} = N_{Mesh}^{Max} - N_{Cont}$ and, by assuming that the number of contacts is always maximized locally, $N_{Cont} = min(N_{RI}^{Max}, N_{Mesh}^{Max})$. Thus simplifying:

$$\Delta G^{\circ tr}(N_{RI}^{Max}, N_{Mesh}^{Max}) \approx \Delta\epsilon_{RI}^{Free} N_{RI}^{Max} + \Delta\epsilon_{Mesh}^{Free} N_{Mesh}^{Max} + (\Delta\epsilon_{Cont} - \Delta\epsilon_{RI}^{Free} - \Delta\epsilon_{Mesh}^{Free})\left(min(N_{RI}^{Max}, N_{Mesh}^{Max})\right)$$

To model this within our schematic (**Fig 2A**), we employ N_{RI}^{Max} values from 0 to 3 (for the valence of expanded rRNA segments in the different stages of ribosome biogenesis) and make N_{Mesh}^{Max} continuous, and empirically connect it to mesh size. Additionally, we use $\Delta\epsilon_{Cont} = -RT$, $\Delta\epsilon_{Mesh}^{Free} = RT$, and $\Delta\epsilon_{RI}^{Free} = \frac{1}{2}RT$, yielding schematic in **Fig S4**.

Supplementary Figures:

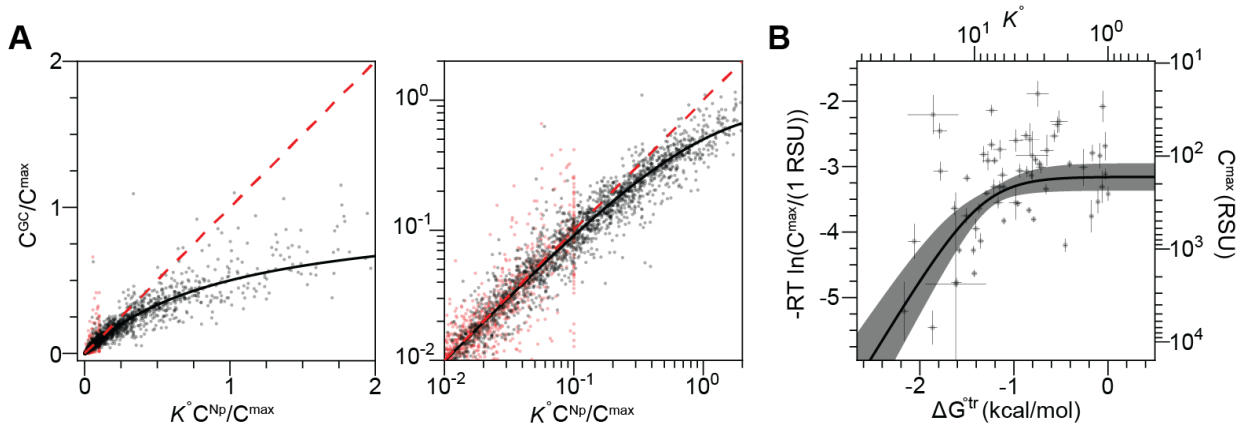

Figure S1: Expression of probes has a hyperbolic dependence. (**A**) Linear-Linear and Log-Log dimensionless overlay of all 148 datasets (5113 data points) where x and y axis are normalized by Cmax/K and Cmax, respectively. For those where Cmax cannot be confidently determined (86 datasets), here distinguished by no measured nucleoplasmic value greater than 0.1 Cmax/K, those are shown with the max x value normalized to 0.1 and all points red. (**B**) Relationship between partitioning and the max concentration, highlighting that there is a minimal amount of probe accommodated in the GC except at higher partitioning, presumably due to an increase in the number of interactions facilitating phase separation including entry into local environments with smaller mesh sizes.

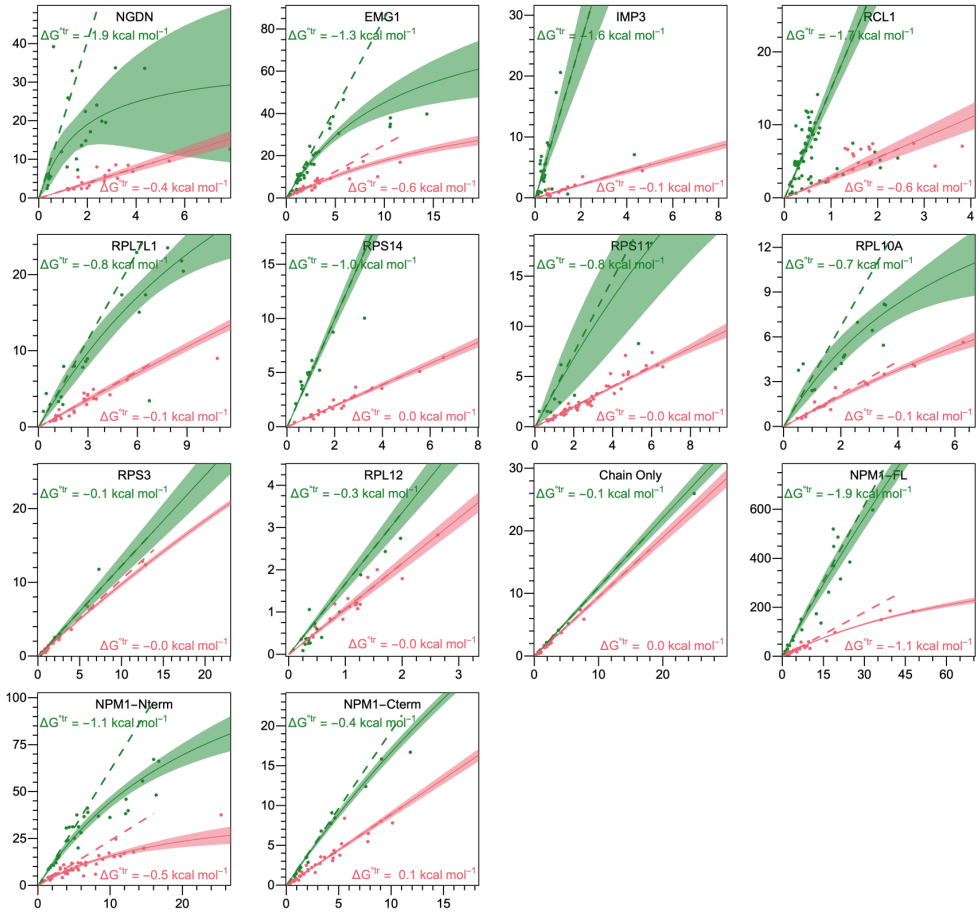

Figure S2: Adding chains to individual constructs always results in LSE. Each construct studied comparing partitioning when fused to no chain (green) and a 400AA chain (red). X axis: C^{Np} (RCU), Y axis: C^{GC} (RCU).

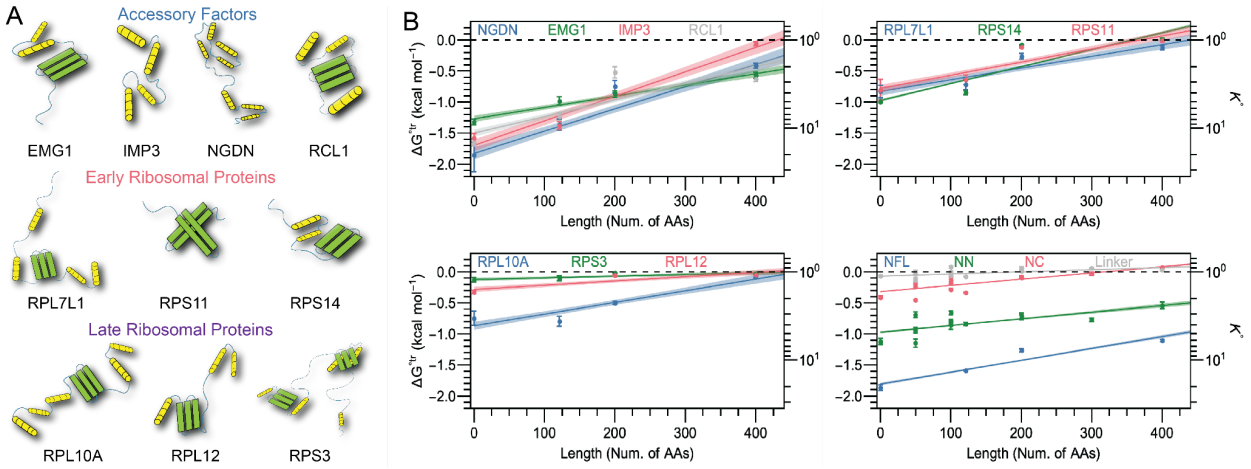

Figure S3: LSE on all constructs grouped by their roles in ribosome biogenesis. (**A**) Cartoon of proteins tested in this manuscript, showing individual structure of each protein. (**B**) Individual $\Delta G^{\circ tr}$ energies for different chain lengths in accessory factors (top left), early ribosomal proteins (top right), late ribosomal proteins (bottom left), and NPM1 and its truncations (bottom right).

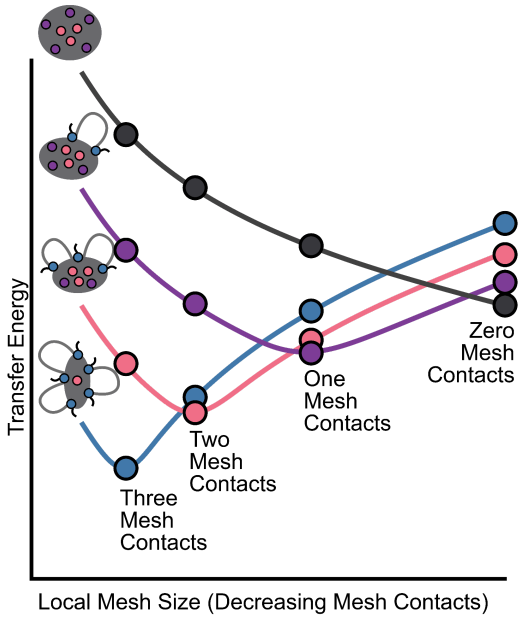

Figure S4: Transfer energy dependence on local mesh size and stage of ribosome biogenesis. The amount of transfer energy needed for contact to be made throughout ribosome biogenesis (blue: accessory factors, pink: early ribosomal proteins, purple: late ribosomal proteins, black: completed ribosome). See methods for full details.

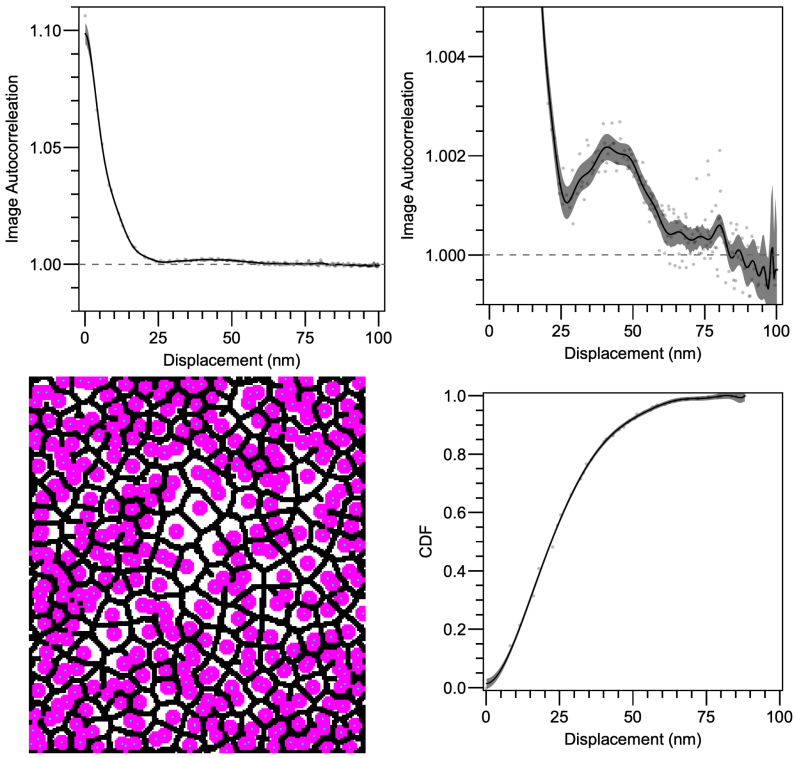

Figure S5: EM data confirms mesh sizes of ribosomal intermediates. Image autocorrelation function (top) showing similar curvature to the RDF (**Fig 2D**). Centers (cyan), inferred location of intermediates (magenta) mesh sizes from EM data (top left) and distribution of mesh sizes using probe data (top right). Segmented grains from EM data, distances measured from the center of each grain (teal dots in purple circles, bottom left) and location of the center of confinement blobs (i.e. skeleton transform between the grains, black in bottom left) used to calculating the cumulative distribution function (CDF) of the mesh sizes (bottom right).

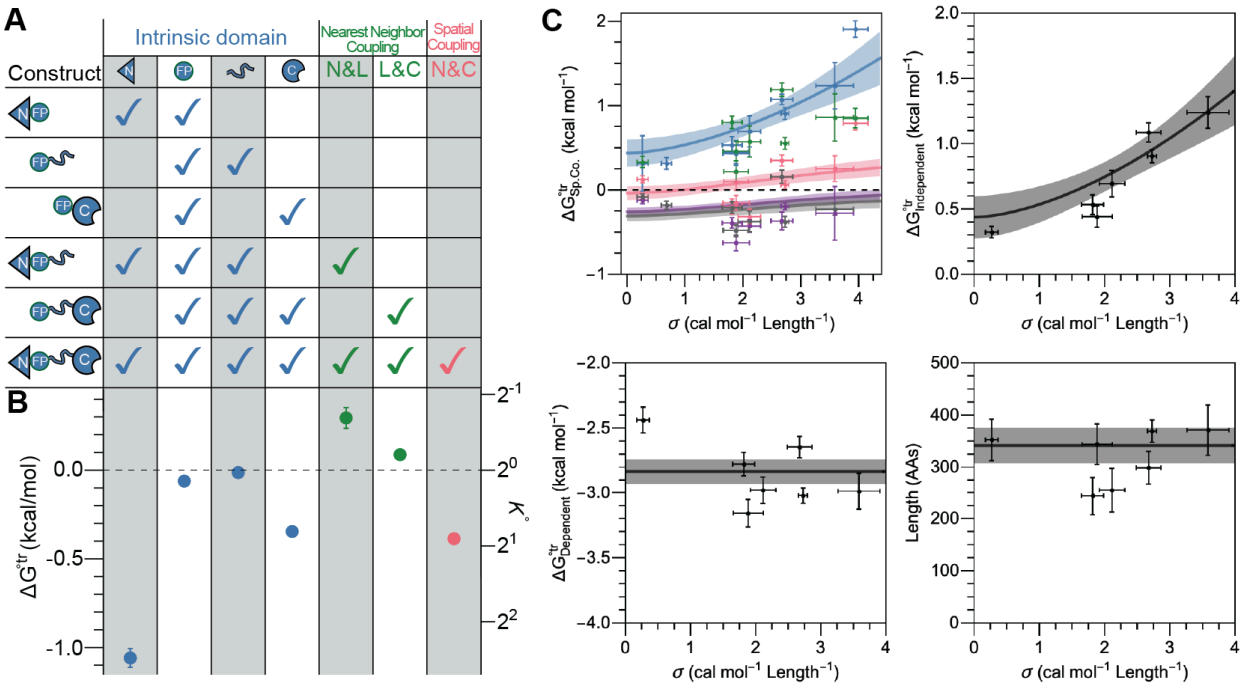

Figure S6: Full LDSC derivation and fit validation. (**A**) Linear breakdown of the $\Delta G^{\circ tr}$ for each NPM1 construct by its domains, including the fluorescent protein's contributions, into energies as indicated. (**B**) Results of linear decomposition of the $\Delta G^{\circ tr}$. (**C,** top left) Spatial coupling for NPM1 and ribosome biogenesis factors with chains of length 0 (blue), 121 (red), 200 (gray), and 400 (purple). Additionally, NPM1-N-terminus spatial coupling with 0 chain (green). Individually fit curves shown (top right, and bottom) for specified parameters. Global fit and dependencies are shown in all plots in C.

Supplementary Tables:

Table S1: Sequences of HDR template/sgRNA. Nucleotide sequences for CRISPR procedure and validating cell lines after electroporation. Red text is mCherry sequence.

Oligo	Sequence	Note
NPM1 wild type/mCherry	GTTCACATTTTTATGACTGATTAAAGTGTTTGGAATTAAATTACATCTGAGTATAAATTTTCT TGGAGTCATATCTTTATCTAGAGTTAACTCTCTGGTGGTAGAATGAAAAATAGATGTTGAA CTATGCAAAGAGACATTTAATTTATTGATGTCTATGAAGTGTTGTGGTTCCTTAACCACATT TCTTTTCTTTTTTTTCCAGGCCATTCAGGACCTTTGGCAATGGAGAAAATCACTAGGAAG CGGAGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTC AAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGC GAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGC CCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTA CGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTC AAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGAC TCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCC CCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGC GGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGA AGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCG TGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGA GGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGG CATGGACGAGCTGTACAAGTAGGAAAATAGTTTAAACAATTTGTTAAAAAATTTTCCGTCT TATTTCATTTCTGTAACAGTTGATATCTGGCTGTCCTTTTTATAATGCAGAGTGAGAA	HDR template
sgNPM1 wild type	UCCAGGCUAUUCAAGAUCUC	sgRNA
NPM1 wild type/mCherry	FW: TGATGTCTATGAAGTGTTGTGGT RV: ACACGGTAGGGAAAGTTCTCA	Genotyping primers

Table S2: DNA sequence of vector. Nucleotide sequence and features added to the multicloning vector site. Red text is mGL sequence. Italicized bases are the overlap regions and the bold bases are the digest sites.

Changed vector feature	Sequence	Note
Cloning site	ACCAATCAGCCTGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTTCCCGAG CTCTATAAAAGAGCTCACAACCCCTCACTCGGCGCGCCAGTCCTCCGAT TGACTGAGTCGCCCTGATCATTGTCGATCCTACCATCCACTCGACACAC CCGCCAGG*ATCCTGACGCGTAT***GGCTAGC**GGTAGTGGTAGCGGCAGTG TGTCTAAGGGGGAAGAGTTATTCACCGGAGTAGTGCCTATCTTAGTGGA GCTCGACGGAGACGTGAATGGTCACAAGTTTTCCGTTAGGGGGGAGGG TGAGGGAGATGCCACAAACGGAAAATTGACCCTCAAGTTTATTTGTACA ACGGGGAAGCTGCCAGTGCCATGGCCAACTCTGGTGACTACACTGGG GTACGGTGTCGCTTGCTTCGCCAGGTATCCTGACCACATGAAGCAGCAT GACTTTTTCAAGAGCGCCATGCCAGAGGGCTATGTACAAGAACGAACCA TTTCATTCAAAGACGATGGGACATATAAAACCCGGGCTGAGGTCAAGTT TGAAGGCGACACACTCGTAAATAGGATAGTGCTGAAAGGTATCGACTTC AAAGAGGATGGAAATATACTGGGACATAAACTGGAATACAACTTTAACAG CCATAAGGTCTACATCACTGCAGACAAACAGAAGAACGGCATTAAGGCA AATTTTAAAACCCGCCACAACGTCGAGGATGGTGGCGTACAATTGGCAG ACCACTACCAGCAGAATACGCCAATTGGAGATGGCCCAGTGTTGTTACC TGATAATCATTATCTGAGTCATCAAAGCAAGCTCAGCAAGGATCCCAATG AGAAGAGGGACCACATGGTGCTTAAGGAACGTGTGACCGCTGCTGGGA	

	TCACCCATGA*TATGGATGAACTGTACA**AGCGCT**CTGGTAGCGGCAG**TAC TAGT***TGACTCGAGGGCCCGCATCCACCATCGCAGACTTATCATCTTATTC CAGATGCGTGCGGATGGAATTCGAGCTCGGTACCTTTAAGACCAATGAC TTACAAGGCAGCTGTAGATCTTAGCCACTTTT	

Table S3: AA Chain sequences. Amino acid sequences for each chain construct. Last column is how longer chains were generated.

Chain	Sequence	Generation from:
NPM1 IDR (121AAs)	EEDAESEDEEEEDVKLLSISGKRSAPGGGSKVPQKKVKLAADEDDDDDDEED DDEDDDDDDFDDEEAEEKAPVKKSIRDTPAKNAQKSNQNGKDSKPSSTPRSK GQESFKKQEKTPKTPKG	NPM1 full length 120-241
L50A	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ	First 50AA of L100A
L50B	PRKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRT	First 50AA of L100B
L50C	TPQPDGGEDQNGSDRKPSGTQRTKQNTTGGTEQNGGPRDPTKPSSNENGR	First 50AA of L100C
L50D	RSKPKQKEKGDKNKQTNDPRDDDERQEEDSTGPTTNDTPTKQPDKEKSQR	First 50AA of L100D
L100A	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ TGGPDTRGQSNQEESEESPSKSDTGERRNKQNEQKPNSNNNRTDTTNKKD	
L100B	PRKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRT PGENTDGGDPEGGEERSTQDETDKSPKRPQDPTRTKDEGGNDSDKRPGRK	
L100C	TPQPDGGEDQNGSDRKPSGTQRTKQNTTGGTEQNGGPRDPTKPSSNENGR ERNKTSQSRQSRPDSTPDEDTPPEGTKGDNPQGRKPGDGDRKEGESDESN	
L100D	RSKPKQKEKGDKNKQTNDPRDDDERQEEDSTGPTTNDTPTKQPDKEKSQRK QNDNKKGPPQNERPNQTRDDDQRKEDEERQQDNKQTRETGNTDNEDETK	
L150	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ TGGPDTRGQSNQEESEESPSKSDTGERRNKQNEQKPNSNNNRTDTTNKKDP RKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRT	L100A + L50B
L200AB	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ TGGPDTRGQSNQEESEESPSKSDTGERRNKQNEQKPNSNNNRTDTTNKKDP RKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRTP GENTDGGDPEGGEERSTQDETDKSPKRPQDPTRTKDEGGNDSDKRPGRK	L100A + L100B
L200CD	TPQPDGGEDQNGSDRKPSGTQRTKQNTTGGTEQNGGPRDPTKPSSNENGR ERNKTSQSRQSRPDSTPDEDTPPEGTKGDNPQGRKPGDGDRKEGESDESN RSKPKQKEKGDKNKQTNDPRDDDERQEEDSTGPTTNDTPTKQPDKEKSQRK QNDNKKGPPQNERPNQTRDDDQRKEDEERQQDNKQTRETGNTDNEDETK	L100C + L100D
L300	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ TGGPDTRGQSNQEESEESPSKSDTGERRNKQNEQKPNSNNNRTDTTNKKDP RKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRTP GENTDGGDPEGGEERSTQDETDKSPKRPQDPTRTKDEGGNDSDKRPGRKT	L100A + L100B + L100C

	PQPDGGEDQNGSDRKPSGTQRTKQNTTGGTEQNGGPRDPTKPSSNENGRE RNKTSQSRQSRPDSTPDEDTPPEGTKGDNPQGRKPGDGDRKEGESDESN	
L400	ERDTDDPSDREPKEEENQESDQSRKRNDENKNDNQNGKEKQNGEDRREKQ TGGPDTRGQSNQEESEESPSKSDTGERRNKQNEQKPNSNNNRTDTTNKKDP RKRTRDEDRNTNDQTPTGNSQNPGENPKKQNGKRQEKTSERNPNTGQRTP GENTDGGDPEGGEERSTQDETDKSPKRPQDPTRTKDEGGNDSDKRPGRKT PQPDGGEDQNGSDRKPSGTQRTKQNTTGGTEQNGGPRDPTKPSSNENGRE RNKTSQSRQSRPDSTPDEDTPPEGTKGDNPQGRKPGDGDRKEGESDESNR SKPKQKEKGDKNKQTNDPRDDDERQEEDSTGPTTNDTPTKQPDKEKSQRKQ NDNKKGPPQNERPNQTRDDDQRKEDEERQQDNKQTRETGNTDNEDETK	L100A + L100B + L100C + L100D

Table S4: NPM1 domain truncations. Amino acid sequences of all proteins used in manuscript. Specific CCDS numbers and ribosome biogenesis roles included for clarity. The last column gives the position of the fluorescent protein and reference of precedent if applicable. For example, “N-terminal” for NPM1 is NPM1-mGL.

Protein	**Protein length**	**Sequence**	**Placement of protein (precedent)**
NPM1 full length	294	MEDSMDMDMSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPVV LRLKCGSGPVHISGQHLVAVEEDAESEDEEEEDVKLLSISGKRSAPGGGS KVPQKKVKLAADEDDDDDEEDDDEDDDDDDFDDEEAEEKAPVKKSIRD TPAKNAQKSNQNGKDSKPSSTPRSKGQESFKKQEKTPKTPKGPSSVEDI KAKMQASIEKGGSLPKVEAKFINYVKNCFRMTDQEAIQDLWQWRKSL	N-terminal
NPM1 N-terminus	1-119	MEDSMDMDMSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPVV LRLKCGSGPVHISGQHLVAV	N-terminal
NPM1 IDR	120-241	EEDAESEDEEEEDVKLLSISGKRSAPGGGSKVPQKKVKLAADEDDDDDD EEDDDEDDDDDDFDDEEAEEKAPVKKSIRDTPAKNAQKSNQNGKDSKPS STPRSKGQESFKKQEKTPKTPKG	Chain
NPM1 C-terminus terminus	242-294	PSSVEDIKAKMQASIEKGGSLPKVEAKFINYVKNCFRMTDQEAIQDLWQW RKSL	C-terminus
NGDN (Accessory factor) CCDS41926 .1	315	MAALGVLESDLPSAVTLLKNLQEQVMAVTAQVKSLTQKVQAGAYPTEKGL SFLEVKDQLLLMYLMDLTHLILDKASGGSLQGHDAVLRLVEIRTVLEKLRPL DQKLKYQIDKLIKTAVTGSLSENDPLRFKPHPSNMMSKLSSEDEEEDEAED DQSEASGKKSVKGVSKKYVPPRLVPVHYDETEAEREKKRLERAKRRALS SSVIRELKEQYSDAPEEIRDARHPHVTRQSQEDQHRINYEESMMVRLSVS KREKGRRKRANVMSSQLHSLTHFSDISALTGGTVHLDEDQNPIKKRKKIPQ KGRKKKGFRRRR	C-terminus (*56*)
IMP3 (Accessory factor) CCDS 10282.1	184	MVRKLKFHEQKLLKQVDFLNWEVTDHNLHELRVLRRYRLQRREDYTRYN QLSRAVRELARRLRDLPERDQFRVRASAALLDKLYALGLVPTRGSLELCDF VTASSFCRRRLPTVLLKLRMAQHLQAAVAFVEQGHVRVGPDVVTDPAFLV TRSMEDFVTWVDSSKIKRHVLEYNEERDDFDLEA	C-terminus (*57*)
EMG1 (Accessory factor) CCDS	244	MAAPSDGFKPRERSGGEQAQDWDALPPKRPRLGAGNKIGGRRLIVVLEG ASLETVKVGKTYELLNCDKHKSILLKNGRDPGEARPDITHQSLLMLMDSPL NRAGLLQVYIHTQKNVLIEVNPQTRIPRTFDRFCGLMVQLLHKLSVRAADG PQKLLKVIKNPVSDHFPVGCMKVGTSFSIPVVSDVRELVPSSDPIVFVVGA	C-terminus (*57*)

73430.1		FAHGKVSVEYTEKMVSISNYPLSAALTCAKLTTAFEEVWGVI	
RCL1 (Accessory factor) CCDS 6456.1	373	MATQAHSLSYAGCNFLRQRLVLSTLSGRPVKIRKIRARDDNPGLRDFEASF IRLLDKITNGSRIEINQTGTTLYYQPGLLYGGSVEHDCSVLRGIGYYLESLLC LAPFMKHPLKIVLRGVTNDQVDPSVDVLKATALPLLKQFGIDGESFELKIVR RGMPPGGGGEVVFSCPVRKVLKPIQLTDPGKIKRIRGMAYSVRVSPQMA NRIVDSARSILNKFIPDIYIYTDHMKGVNSGKSPGFGLSLVAETTSGTFLSA ELASNPQGQGAAVLPEDLGRNCARLLLEEIYRGGCVDSTNQSLALLLMTL GQQDVSKVLLGPLSPYTIEFLRHLKSFFQIMFKIETKPCGEELKGGDKVLM TCVGIGFSNLSKTLK	C-terminus (*58*)
RPL7L1 (uL30) (Early ribosomal protein) CCDS 4873.2	255	MISSCTTRKMAEQEQRKIPLVPENLLKKRKAYQALKATQAKQALLAKKEQK KGKGLRFKRLESFLHDSWRQKRDKVRLRRLEVKPHALELPDKHSLAFVV RIERIDGVSLLVQRTIARLRLKKIFSGVFVKVTPQNLKMLRIVEPYVTWGFP NLKSVRELILKRGQAKVKNKTIPLTDNTVIEEHLGKFGVICLEDLIHEIAFPG KHFQEISWFLCPFHLSVARHATKNRVGFLKEMGTPGYRGERINQLIRQLN	C-terminus
RPS14 (uS11) (Early ribosomal protein) CCDS 4307.1	151	MAPRKGKEKKEEQVISLGPQVAEGENVFGVCHIFASFNDTFVHVTDLSGK ETICRVTGGMKVKADRDESSPYAAMLAAQDVAQRCKELGITALHIKLRATG GNRTKTPGPGAQSALRALARSGMKIGRIEDVTPIPSDSTRRKGGRRGRRL	C-terminus (*57*)
RPS11 (uS17) (Early ribosomal protein) CCDS 12769.1	158	MADIQTERAYQKQPTIFQNKKRVLLGETGKEKLPRYYKNIGLGFKTPKEAI EGTYIDKKCPFTGNVSIRGRILSGVVTKMKMQRTIVIRRDYLHYIRKYNRFE KRHKNMSVHLSPCFRDVQIGDIVTVGECRPLSKTVRFNVLKVTKAAGTKK QFQKF	C-terminus (*57*)
RPL10A (uL1) (Late ribosomal protein) CCDS 4806.1	217	MSSKVSRDTLYEAVREVLHGNQRKRRKFLETVELQISLKNYDPQKDKRFS GTVRLKSTPRPKFSVCVLGDQQHCDEAKAVDIPHMDIEALKKLNKNKKLV KKLAKKYDAFLASESLIKQIPRILGPGLNKAGKFPSLLTHNENMVAKVDEVK STIKFQMKKVLCLAVAVGHVKMTDDELVYNIHLAVNFLVSLLKKNWQNVRA LYIKSTMGKPQRLY	C-terminus (*57*)
RPS3 (uS3) (Late ribosomal protein) CCDS 58161.1	243	MAVQISKKRKFVADGIFKAELNEFLTRELAEDGYSGVEVRVTPTRTEIIILAT RTQNVLGEKGRRIRELTAVVQKRFGFPEGSVELYAEKVATRGLCAIAQAES LRYKLLGGLAVRRACYGVLRFIMESGAKGCEVVVSGKLRGQRAKSMKFV DGLMIHSGDPVNYYVDTAVRHVLLRQGVLGIKVKIMLPWDPTGKIGPKKPL PDHVSIVEPKDEILPTTPISEQKGGKPEPPAMPQPVPTA	C-terminus (*59*)
RPL12 (uL11) (Late ribosomal protein) CCDS 6872.1	165	MPPKFDPNEIKVVYLRCTGGEVGATSALAPKIGPLGLSPKKVGDDIAKATG DWKGLRITVKLTIQNRQAQIEVVPSASALIIKALKEPPRDRKKQKNIKHSGNI TFDEIVNIARQMRHRSLARELSGTIKEILGTAQSVGCNVDGRHPHDIIDDIN SGAVECPAS	C-terminus

References:

1. A. Molliex, J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, H. J. Kim, T. Mittag, J. P. Taylor, Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123–133 (2015).

2. M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, C. P. Brangwynne, Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).

3. C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, A. A. Hyman, Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).

4. T. Mittag, R. V. Pappu, A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell*, doi: 10.1016/j.molcel.2022.05.018 (2022).

5. A. Musacchio, On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J.* **41**, e109952 (2022).

6. F. Erdel, K. Rippe, Formation of Chromatin Subcompartments by Phase Separation. *Biophys. J.* **114**, 2262–2270 (2018).

7. J. A. Riback, L. Zhu, M. C. Ferrolino, M. Tolbert, D. M. Mitrea, D. W. Sanders, M.-T. Wei, R. W. Kriwacki, C. P. Brangwynne, Composition-dependent thermodynamics of intracellular phase separation. *Nature* **581**, 209–214 (2020).

8. J. A. Riback, C. D. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. R. Sosnick, D. A. Drummond, Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **168**, 1028–1040.e19 (2017).

9. D. T. McSwiggen, A. S. Hansen, S. S. Teves, H. Marie-Nelly, Y. Hao, A. B. Heckert, K. K. Umemoto, C. Dugast-Darzacq, R. Tjian, X. Darzacq, Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. *Elife* **8** (2019).

10. R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, M. Kar, Phase Transitions of Associative Biomacromolecules. *Chem. Rev.* **123**, 8945–8987 (2023).

11. P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon, M. K. Rosen, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).

12. A. Chattaraj, M. L. Blinov, L. M. Loew, The solubility product extends the buffering concept to heterotypic biomolecular condensates. *Elife* **10** (2021).

13. J. A. Riback, J. M. Eeftens, D. S. W. Lee, S. A. Quinodoz, A. Donlic, N. Orlovsky, L. Wiesner, L. Beckers, L. A. Becker, A. R. Strom, U. Rana, M. Tolbert, B. W. Purse, R. Kleiner, R. Kriwacki, C. P. Brangwynne, Viscoelasticity and advective flow of RNA underlies nucleolar form and function. *Mol. Cell* **83**, 3095–3107.e9 (2023).

14. F. Dar, S. R. Cohen, D. M. Mitrea, A. H. Phillips, G. Nagy, W. C. Leite, C. B. Stanley, J.-M.

Choi, R. W. Kriwacki, R. V. Pappu, Biomolecular condensates form spatially inhomogeneous network fluids. *Nat. Commun.* **15**, 3413 (2024).

15. A. S. Belmont, Nuclear Compartments: An Incomplete Primer to Nuclear Compartments, Bodies, and Genome Organization Relative to Nuclear Architecture. *Cold Spring Harb. Perspect. Biol.*, a041268 (2021).

16. R. G. Larson, *The Structure and Rheology of Complex Fluids* (Oxford university press New York, 1999)vol. 150.

17. H. Zhou, J. Hutchings, M. Shiozaki, X. Zhao, L. K. Doolittle, S. Yang, R. Yan, N. Jean, M. Riggi, Z. Yu, E. Villa, M. K. Rosen, Quantitative spatial analysis of chromatin biomolecular condensates using cryo-electron tomography, *bioRxivorg* (2024). https://doi.org/10.1101/2024.12.01.626131.

18. D. Scholl, T. Boyd, A. P. Latham, A. Salazar, A. Khan, S. Boeynaems, A. S. Holehouse, G. C. Lander, A. Sali, D. Park, A. A. Deniz, K. Lasker, Cellular function of a biomolecular condensate is determined by its ultrastructure, *bioRxiv* (2024). https://doi.org/10.1101/2024.12.27.630454.

19. M. Baum, F. Erdel, M. Wachsmuth, K. Rippe, Retrieving the intracellular topology from multi-scale protein mobility mapping in living cells. *Nat. Commun.* **5**, 4494 (2014).

20. N. O. Taylor, M.-T. Wei, H. A. Stone, C. P. Brangwynne, Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophys. J.* **117**, 1285–1300 (2019).

21. M.-T. Wei, S. Elbaum-Garfinkle, A. S. Holehouse, C. C.-H. Chen, M. Feric, C. B. Arnold, R. D. Priestley, R. V. Pappu, C. P. Brangwynne, Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.*, nchem.2803 (2017).

22. J. V. Oakley, B. F. Buksh, D. F. Fernández, D. G. Oblinsky, C. P. Seath, J. B. Geri, G. D. Scholes, D. W. C. MacMillan, Radius measurement via super-resolution microscopy enables the development of a variable radii proximity labeling platform. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2203027119 (2022).

23. W. C. Wimley, S. H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **3**, 842–848 (1996).

24. C. Tanford, Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J. Am. Chem. Soc.* **84**, 4240–4247 (1962).

25. C. P. Moon, K. G. Fleming, Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10174–10177 (2011).

26. B. A. Krantz, T. R. Sosnick, Engineered metal binding sites map the heterogeneous folding landscape of a coiled coil. *Nat. Struct. Biol.* **8**, 1042–1047 (2001).

27. M. Petersen, D. Barrick, Analysis of Tandem Repeat Protein Folding Using Nearest-Neighbor Models. *Annu. Rev. Biophys.* **50**, 245–265 (2021).

28. D. L. J. Lafontaine, J. A. Riback, R. Bascetin, C. P. Brangwynne, The nucleolus as a

multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22**, 165–182 (2021).

29. T. Pederson, The nucleolus. *Cold Spring Harb. Perspect. Biol.* **3** (2011).

30. D. M. Mitrea, J. A. Cika, C. B. Stanley, A. Nourse, P. L. Onuchic, P. R. Banerjee, A. H. Phillips, C.-G. Park, A. A. Deniz, R. W. Kriwacki, Self-interaction of NPM1 modulates multiple mechanisms of liquid–liquid phase separation. *Nat. Commun.* **9**, 842 (2018).

31. D. W. Sanders, N. Kedersha, D. S. W. Lee, A. R. Strom, V. Drake, J. A. Riback, D. Bracha, J. M. Eeftens, A. Iwanicki, A. Wang, M.-T. Wei, G. Whitney, S. M. Lyons, P. Anderson, W. M. Jacobs, P. Ivanov, C. P. Brangwynne, Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* **181**, 306–324.e28 (2020).

32. S. F. Banani, A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker, M. K. Rosen, Compositional Control of Phase-Separated Cellular Bodies. *Cell* **166**, 651–663 (2016).

33. A. J. LaPeruta, J. Micic, J. L. Woolford Jr, Additional principles that govern the release of pre-ribosomes from the nucleolus into the nucleoplasm in yeast. *Nucleic Acids Res.*, doi: 10.1093/nar/gkac430 (2022).

34. M. Rubinstein, R. H. Colby, Others, *Polymer Physics* (Oxford university press New York, 2003)vol. 23.

35. I. Teraoka, *Polymer Solutions* (John Wiley & Sons, Nashville, TN, ed. 1, 2004).

36. J. A. Riback, M. A. Bowman, A. M. Zmyslowski, C. R. Knoverek, J. M. Jumper, J. R. Hinshaw, E. B. Kaye, K. F. Freed, P. L. Clark, T. R. Sosnick, Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science* **358**, 238–241 (2017).

37. E. Nicolas, P. Parisot, C. Pinto-Monteiro, R. de Walque, C. De Vleeschouwer, D. L. J. Lafontaine, Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nat. Commun.* **7**, 11390 (2016).

38. N. Galvanetto, M. T. Ivanović, A. Chowdhury, A. Sottini, M. F. Nüesch, D. Nettels, R. B. Best, B. Schuler, Extreme dynamics in a biomolecular condensate. *Nature* **619**, 876–883 (2023).

39. R. L. Ochs, Methods used to study structure and function of the nucleolus. *Methods Cell Biol.* **53**, 303–321 (1998).

40. L. Heinrich, D. Bennett, D. Ackerman, W. Park, J. Bogovic, N. Eckstein, A. Petruncio, J. Clements, S. Pang, C. S. Xu, J. Funke, W. Korff, H. F. Hess, J. Lippincott-Schwartz, S. Saalfeld, A. V. Weigel, COSEM Project Team, Whole-cell organelle segmentation in volume electron microscopy. *Nature* **599**, 141–146 (2021).

41. J. C. R. Politz, I. Polena, I. Trask, D. P. Bazett-Jones, T. Pederson, A nonribosomal landscape in the nucleolus revealed by the stem cell protein nucleostemin. *Mol. Biol. Cell* **16**, 3401–3410 (2005).

42. S. H. White, W. C. Wimley, Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).

43. L. Zhu, T. M. Richardson, L. Wacheul, M.-T. Wei, M. Feric, G. Whitney, D. L. J. Lafontaine, C. P. Brangwynne, Controlling the material properties and rRNA processing function of the nucleolus using light. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 17330–17335 (2019).

44. J. Witten, K. Ribbeck, The particle in the spider's web: transport through biological hydrogels. *Nanoscale* **9**, 8080–8095 (2017).

45. H. Mangan, B. McStay, Human nucleoli comprise multiple constrained territories, tethered to individual chromosomes. *Genes Dev.* **35**, 483–488 (2021).

46. S. Chong, T. G. W. Graham, C. Dugast-Darzacq, G. M. Dailey, X. Darzacq, R. Tjian, Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol. Cell*, doi: 10.1016/j.molcel.2022.04.007 (2022).

47. F. Erdel, A. Rademacher, R. Vlijm, J. Tünnermann, L. Frank, R. Weinmann, E. Schweigert, K. Yserentant, J. Hummert, C. Bauer, S. Schumacher, A. Al Alwash, C. Normand, D.-P. Herten, J. Engelhardt, K. Rippe, Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Mol. Cell* **0** (2020).

48. J. D. Schmit, M. Feric, M. Dundr, How hierarchical interactions make membraneless organelles tick like clockwork. *Trends Biochem. Sci.* **46**, 525–534 (2021).

49. A. S. Holehouse, S. Alberti, Molecular determinants of condensate composition. *Mol. Cell* **85**, 290–308 (2025).

50. A. Tartakoff, P. DiMario, E. Hurt, B. McStay, V. G. Panse, D. Tollervey, The dual nature of the nucleolus. *Genes Dev.* **36**, 765–769 (2022).

51. D. T. McSwiggen, M. Mir, X. Darzacq, R. Tjian, Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.*, doi: 10.1101/gad.331520.119 (2019).

52. C. Mathieu, R. V. Pappu, J. Paul Taylor, Beyond aggregation: Pathological phase transitions in neurodegenerative disease. *Science* **370**, 56–60 (2020).

53. B. C. Campbell, E. M. Nabel, M. H. Murdock, C. Lao-Peregrin, P. Tsoulfas, M. G. Blackmore, F. S. Lee, C. Liston, H. Morishita, G. A. Petsko, mGreenLantern: a bright monomeric fluorescent protein with rapid expression and cell filling properties for neuronal imaging. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 30710–30721 (2020).

54. H. S. Chan, K. A. Dill, Solvation: Effects of molecular size and shape. *J. Chem. Phys.* **101**, 7007–7026 (1994).

55. A. Ben-Naim, Standard thermodynamics of transfer. Uses and misuses. *J. Phys. Chem.* **82**, 792–803 (1978).

56. C.-R. Sihn, Y.-S. Lee, J.-S. Jeong, K. Park, S. H. Kim, CANu1, a novel nucleolar protein, accumulated on centromere in response to DNA damage. *Genes Cells* **13**, 787–796 (2008).

57. N. H. Cho, K. C. Cheveralls, A.-D. Brunner, K. Kim, A. C. Michaelis, P. Raghavan, H. Kobayashi, L. Savy, J. Y. Li, H. Canaj, J. Y. S. Kim, E. M. Stewart, C. Gnann, F. McCarthy, J. P. Cabrera, R. M. Brunetti, B. B. Chhun, G. Dingle, M. Y. Hein, B. Huang, S. B. Mehta, J.

S. Weissman, R. Gómez-Sjöberg, D. N. Itzhak, L. A. Royer, M. Mann, M. D. Leonetti, OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* **375**, eabi6983 (2022).

58. D. J. Leary, M. P. Terns, S. Huang, Components of U3 snoRNA-containing complexes shuttle between nuclei and the cytoplasm and differentially localize in nucleoli: implications for assembly and function. *Mol. Biol. Cell* **15**, 281–293 (2004).

59. S. B. Lee, I.-S. Kwon, J. Park, K.-H. Lee, Y. Ahn, C. Lee, J. Kim, S. Y. Choi, S.-W. Cho, J.-Y. Ahn, Ribosomal protein S3, a new substrate of Akt, serves as a signal mediator between neuronal apoptosis and DNA repair. *J. Biol. Chem.* **285**, 29457–29468 (2010).