Supplementary Materials for Volume buffering in multi-component phase separation

Logan de Monchaux-Irons[1,2]†, Nina Han[1,2]†, Benjamin Frühbauer[1,2]†,
Jianning Jiang[1,2], Frédéric Allain[1,2], Madhav Jagannathan[1,2,*],
Leonidas Emmanouilidis[1,2,*], Thomas C.T. Michaels[1,2,*]
[*]Corresponding author. Email: madhav.jagannathan@bc.biol.ethz.ch,
leonidas@bc.biol.ethz.ch, thomas.michaels@bc.biol.ethz.ch
[†]These authors contributed equally to this work.

This PDF file includes:

Materials and Methods
Supplementary Text
Figures S1 to S5

Materials and Methods

In vitro experiments

Protein expression and preparation

Gb1-His_6 tagged FUS and meGFP-FUS were expressed in Escherichia coli strain BL21 (DE3) in LB medium supplemented with 50 μM/mL kanamycin. Once the optical density reached 0.6 – 0.7, overnight protein expression was induced by the addition of 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C. Cells were lysed by sonication in the lysis buffer containing 50 mM HEPES at pH 7.5, 150 mM NaCl, and 1 tablet of protease inhibitor cocktail (Roche). Lysates were clarified by 20 minutes of centrifugation at 15,000 rpm. Pellets were then collected and underwent further solubilization via homogenization with buffer A, which consists of 8 M urea, 1 M NaCl, 50 mM HEPES, and 2 mM DTT at pH 7.5. The solubilized protein was isolated through 25 minutes of centrifugation at 17,000 rpm. Ni-NTA chromatography was used to purify the protein. Once the protein was loaded, Ni-NTA beads (Qiagen) were washed with buffer A1, which has 1 M urea, 150 mM NaCl, 50 mM HEPES, and 2 mM DTT at pH 7.5. Tagged protein was eluted using elution buffer containing 250 mM imidazole, 1 M urea, 150 mM NaCl, 50 mM HEPES, and 2 mM DTT at pH 7.5. The Gb1-His_6 tag was removed by TEV protease (in-house purified) cleavage through overnight dialysis in buffer B, which has 1 M urea, 150 mM NaCl, 5 mM β-mercaptoethanol, 2 mM DTT, and 50 mM HEPES at pH 7.5. Before a second time Ni-NTA chromatography, the protein was resolubilized at -80 °C in buffer C with 6 M urea, 150 mM NaCl, 50 mM HEPES, and 2 mM DTT at pH 7.5. The untagged protein was collected in the flow-through of the second-time Ni-NTA chromatography and concentrated before storing at -80 °C.

RNA preparation

The U1snRNA SL34 was cloned into pUC19 plasmid and was linearized via BsaI digestion. Utilizing in-house produced T7 RNA polymerase, in-vitro transcribed RNA was purified by a high-performance liquid chromatography (HPLC) system (Gilson) under denaturing conditions with an anion-exchange column (preparative Dionex DNA Pac PA-100 column, 22 250 mm). After butanol extraction to remove the urea from the purification, U1snRNA SL34 was refolded at 95 °C for 5

minutes. For U1snRNA SL34-Cy5, post-labeling reaction was conducted by transferring pCp-Cy5 (Cytidine-5'-phosphate-3'-(6-aminohexyl) phosphate) (Jena Bioscience) onto 3'-hydroxyl group of U1snRNA SL34 utilizing T4 RNA ligase (NEB).

Induction of FUS liquid droplets

All FUS droplet samples studied in this work were induced by the ten-time dilution of protein stocks in the droplet buffer containing 5 mM phosphate, 1 mM TCEP and 5 μM $ZnCl_2$ at pH 6.0. As a result, the final buffer in the FUS biphasic sample was 0.6 M urea, 15 mM NaCl, 5 mM HEPES, 5 mM phosphate, 1 mM TCEP, and 5 μM $ZnCl_2$ at pH 7.0. For samples containing U1snRNA SL34, RNA was mixed with the droplet buffer prior to droplet induction.

Nanodrop measurements

Followed by centrifugation of 20 μL biphasic samples at 10,000 x g for 5 minutes, dilute phase samples were obtained by isolating the supernatant into a new Eppendorf tube. The protein concentration was then determined by measuring the absorbance at 280 nm via Nanodrop (NanoDrop OneC ThermoFisher). Each data point contains 3 independent replicates.

Fluorescent microscopy

Prior to droplet induction, meGFP-FUS was premixed with unlabeled FUS in 1:50 molar ratio and U1snRNA SL34-Cy5 was pre-mixed with unlabeled U1snRNA SL34 in 1:10 molar ratio. Upon droplet induction, samples were immediately loaded onto the 8-well glass bottom slides (VI 0.4 ibiTreat, Ibidi) and imaged on a Nikon Spinning Disk SoRa confocal microscope (Nikon Eclipse Ti2) with 100X magnification oil objective. 488-nm laser line and 640-nm laser were used to acquire meGFP-FUS and U1snRNA SL34-Cy5 respectively. Images were obtained with 100-ms exposure time. Fiji (*39*) was used for all subsequent image processing.

NMR 1D DOSY experiments in agarose gel

NMR agarose buffer, which consists of 0.5% agarose (weight/volume), 5 mM phosphate, 1 mM TCEP, and 5 μM $ZnCl_2$ at pH 6.0, was prepared by bringing the mixture to boil using a microwave.

Before sample preparation, the NMR agarose buffer was equilibrated to 55 °C. NMR samples were prepared by quickly mixing the FUS stock, the U1snRNA SL34 stock, 10% D2O, and the NMR agarose buffer. Then using a glass pipette, the samples were immediately transferred into 3 mm NMR tubes. All NMR 1D DOSY experiments were recorded at room temperature using Avance III HD at 900-MHz with cryogenic probes. The pulse sequence utilized for the 1D DOSY experiment was described in the previous work (*21*). All spectra were processed with Topspin 3.2. (Fig. S1).

In vivo experiments

Cell culture and microscopy

NIH-3T3 cells were cultured in high glucose DMEM with pyruvate supplemented with 10% FBS (Life Technologies Limited, UK). The HMGA1 coding sequence was cloned into a pCDNA3 vector downstream of a mNeonGreen (mNG) fluorescent tag. Cells were grown for 24 hours on round 12 mm coverslips to 50% confluency, followed by transfection with Lipofectamine™ 2000 Transfection Reagent (Invitrogen, USA), according to manufacturer's instructions. Cells were fixed in 3.6% formaldehyde 48h post-transfection and permeabilized for 5 minutes with 1x PBS containing 0.4% Triton-X 100. Cells were mounted with VECTASHIELD® Antifade Mounting Medium with DAPI (Vector Laboratories, USA). The mounting medium was additionally supplemented with 1:1000 Sphero™ Fluorescent Yellow Particles 0.5 um (Spherotech, USA) to enable estimation of expression levels across samples (Fig. S2 D). Imaging of nuclei and fluorescent beads was done on a Leica SP8 AOBS laser scanning confocal microscope using a 63x oil immersion objective (NA = 1.4). Experiments consisted of 3-4 independent replicate samples; 4 images of fluorescent beads in different regions of the sample were recorded per sample.

Image segmentation and analysis

Nuclei and chromocenters were segmented on the DAPI channel (Fig. S2 C, excitation at 405 nm), and fluorescent beads were segmented on the mNG channel (excitation at 488 nm) using custom-built CellProfiler (version 4.2.6) pipelines accessible via the repository (10.5281/zenodo.13711525), and fluorescent intensities were measured within the segmented objects. Normalized concentrations of mNG and mNG-HMGA1 were determined by normalization of fluorescent

intensity to the median upper-quantile intensity of fluorescent beads (Fig. S2 D). Image data was filtered to exclude wrongly segmented nuclei, any nuclei above an empirically determined size threshold, as well as nuclei containing more than 65 chromocenters, pre-processing script accessible via (10.5281/zenodo.13711525). Data processing and statistical analysis and visualization were done in Python and R. Representative microscopy images were prepared with ImageJ/FIJI.

Supplementary Text

Volume line construction

Consider a multi-component system comprising n solutes and a solvent that phase separates into a dense phase coexisting with a dilute phase. Let ϕ_i^{I} denote the concentration of component i in the dense phase, and ϕ_i^{II} the concentration of species i in the dilute phase. The volume V^{I} of the droplet phase can be calculated using the equations

$$\overline{\phi_i} = \frac{V^{\mathrm{I}}}{V}\phi_i^{\mathrm{I}} + \frac{V^{\mathrm{II}}}{V}\phi_i^{\mathrm{II}}, \quad (i = 1, \cdots, n) \tag{S1}$$

$$V = V^{\mathrm{I}} + V^{\mathrm{II}}, \tag{S2}$$

where $\overline{\phi_i}$ is the total (average) concentration of component i. These equations are solved to yield

$$\frac{V^{\mathrm{I}}}{V} = \frac{\overline{\phi_i} - \phi_i^{\mathrm{II}}}{\phi_i^{\mathrm{I}} - \phi_i^{\mathrm{II}}}. \tag{S3}$$

To construct the volume line, we take a sample average concentration of the form

$$\overline{\phi_i} = \alpha\, \phi_i^{\mathrm{I}} + (1-\alpha)\, \phi_i^{\mathrm{II}} \quad (0 \le \alpha \le 1). \tag{S4}$$

Indeed, inserting Eq. (S4) into the definition of the volume of the droplet phase Eq. (S3), we obtain:

$$\frac{V^{\mathrm{I}}}{V} = \frac{\alpha\, \phi_i^{\mathrm{I}} + (1-\alpha)\, \phi_i^{\mathrm{II}} - \phi_i^{\mathrm{II}}}{\phi_i^{\mathrm{I}} - \phi_i^{\mathrm{II}}} = \alpha. \tag{S5}$$

Thus, the volume of the droplet phase is constant and equal to α.

Flory-Huggins for two-component phase separation

In the following, we consider a two-component mixture comprising two solutes and a solvent. The two solutes of total concentrations $(\overline{\phi_1}, \overline{\phi_2})$ phase separate into two different phases described by

the concentrations in phase I: $(\phi_1^{\mathrm{I}}, \phi_2^{\mathrm{I}})$ and phase II: $(\phi_1^{\mathrm{II}}, \phi_2^{\mathrm{II}})$, as described by the Flory-Huggins free-energy density

$$\begin{aligned}\frac{f(\phi_1,\phi_2)}{k_BT} =& \frac{\phi_1}{N_1}\ln(\phi_1) + \frac{\phi_2}{N_2}\ln(\phi_2) + (1-\phi_1-\phi_2)\ln(1-\phi_1-\phi_2) \\ &- \frac{\chi_{11}}{2}\phi_1^2 - \frac{\chi_{22}}{2}\phi_2^2 - \chi_{12}\phi_1\phi_2\,,\end{aligned} \tag{S6}$$

where ϕ_1 and ϕ_2 are the volume fractions of the solutes, and N_1, N_2 are the numbers of persistence lengths (number of lattice sites occupied by each solute). χ_{11} and χ_{22} are the self-attractive or homotypic interaction parameters of the solutes, while χ_{12} is the heterotypic (attraction between solutes) interaction energy parameters.

By selecting sufficient homotypic and heterotypic interaction energies, the free energy landscape over ϕ_1, ϕ_2 can contain different coexisting equilibria points. The equilibria points are connected by concentration tie lines where picking an average composition $\bar{\phi}_i$ on these lines causes demixing into the equilibria points thus forming coexisting phases. Selecting strong homotypic and weak heterotypic interactions yields segregative LLPS and selecting strong heterotypic and weak homotypic interactions leads to associative LLPS.

A method of calculating the locations of minima is by taking the convex hull of the free-energy landscape (*40*). Our computational method involved making a discretized free-energy manifold over the ϕ_1 and ϕ_2 space and then taking the convex hull construction. The simplexes formed by the convex hull are stretched between minima and are analogous to concentration tie lines. Solutions from this method for segregative and associative LLPS are shown in the main text in Fig. 2 C. The phase diagrams show the gradient of the concentration tie lines connecting the equilibria points depending on the type of LLPS (segregative or associative).

Determination of the slope of volume lines

In this appendix, we obtain an expression for the slope of the volume lines in dependence on the interaction parameters between components using the Flory-Huggins model described previously. To this end, we parameterize the binodal line via the arc-length coordinate s:

$$\Phi^i(s) = \begin{pmatrix}\phi_1^i(s)\\ \phi_2^i(s)\end{pmatrix} \tag{S7}$$

where i = I, II. To determine the slope of the volume lines, we consider the tangent vector

$$\partial_s \Phi^i(s) = \begin{pmatrix} \partial_s \phi_1^i(s) \\ \partial_s \phi_2^i(s) \end{pmatrix} \tag{S8}$$

where ∂_s denotes the derivative with respect to s. The volume lines are defined by the equations

$$\overline{\phi_j} = \alpha\, \phi_j^{\mathrm{I}} + (1-\alpha)\, \phi_j^{\mathrm{II}} \tag{S9}$$

where $j = 1, 2$. The local slope of the volume lines, k, is then given by

$$k = \frac{\partial_s \overline{\phi_2}}{\partial_s \overline{\phi_1}} = \frac{\alpha\, \partial_s \phi_2^I + (1-\alpha)\, \partial_s \phi_2^I}{\alpha\, \partial_s \phi_1^I + (1-\alpha)\, \partial_s \phi_1^I}\,. \tag{S10}$$

The conditions for phase equilibrium (binodal line) are defined as

$$\mu_1^I = \mu_1^{II}\,, \tag{S11}$$

$$\mu_2^I = \mu_2^{II}\,, \tag{S12}$$

$$\Pi^I = \Pi^{II}\,, \tag{S13}$$

where

$$\mu_j = \nu_j \frac{\partial f}{\partial \phi_j}$$

is the chemical potential of component j and

$$\Pi = -f + \sum_j \phi_j \frac{\partial f}{\partial \phi_j}$$

is the osmotic pressure. These conditions (S11) and (S13) can be rewritten as

$$\mathrm{Hess}(\Phi^I)\, \partial_s \Phi^I = \mathrm{Hess}(\Phi^{II})\, \partial_s \Phi^{II} \tag{S14}$$

$$\langle \Phi^I, \mathrm{Hess}(\Phi^I)\, \partial_s \Phi^I \rangle = \langle \Phi^{II}, \mathrm{Hess}(\Phi^{II})\, \partial_s \Phi^{II} \rangle \tag{S15}$$

where

$$\mathrm{Hess} = \begin{pmatrix} \frac{\partial^2 f}{\partial \phi_1^2} & \frac{\partial^2 f}{\partial \phi_1 \partial \phi_2} \\ \frac{\partial^2 f}{\partial \phi_1 \partial \phi_2} & \frac{\partial^2 f}{\partial \phi_2^2} \end{pmatrix} \tag{S16}$$

is the Hessian matrix of the free energy density f. Combining Eqs. (S14) and (S15), we find

$$\langle \Phi^I - \Phi^{II}, \mathrm{Hess}(\Phi^I)\, \partial_s \Phi^I \rangle = 0\,. \tag{S17}$$

The tie line connecting coexisting phases can be defined as the vector $(\Phi^I - \Phi^{II})$. Since the Hessian is a symmetric matrix, we can write:

$$\langle \mathrm{Hess}(\Phi^I)(\Phi^I - \Phi^{II}), \partial_s \Phi^I \rangle = 0\,. \tag{S18}$$

To approximate the tie-line direction, we use the eigenvectors of the Hessian corresponding to the lowest eigenvalue λ_{min} (*41*):

$$\mathrm{Hess}(\Phi^I)(\Phi^I - \Phi^{II}) = \lambda_{min}(\Phi^I - \Phi^{II})$$

leading to

$$\langle (\Phi^I - \Phi^{II}), \partial_s \Phi^I \rangle = 0 \tag{S19}$$

Therefore,

$$\partial_s \Phi^I \in (\Phi^I - \Phi^{II})^{\perp} \tag{S20}$$

where $X^{\perp}$ denotes the orthogonal complement of X. The same result can be found for $\partial_s \Phi^{II}$. Condition S20 implied that if we write

$$(\Phi^I - \Phi^{II}) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{S21}$$

then

$$\partial_s \Phi^I = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix} \tag{S22}$$

Combined with (8) we have

$$k \approx -\frac{v_1}{v_2} = -\frac{1}{\text{slope of tie line}} \tag{S23}$$

By using the approximation of the slope of the tie lines from (*41*) for the limit where $\phi_1 \ll \phi_2 \ll 1$, we obtain the slope of the volume lines as:

$$k \approx (1 + 2\chi^{\Delta}) N_1 \phi_1 \tag{S24}$$

where $\chi^{\Delta} = (\chi_{ij} - \chi_{ii} - \chi_{jj})/2$.

Mass line construction

The in vitro experiment suggested that the amount/mass of protein in the dense phase could remain constant despite changing the total concentration of protein. A similar construction to the volume lines can be constructed when we consider the mass of a component in a phase to be equal to the concentration multiplied by the volume of that phase

$$M_j^I = \frac{\phi_j^{\mathrm{I}}(\overline{\phi_j} - \phi_j^{\mathrm{II}})}{\overline{\phi_j}(\phi_j^{\mathrm{I}} - \phi_j^{\mathrm{II}})}, \quad j = 1, 2\,. \tag{S25}$$

Setting that equal to a constant mass β we can solve for the input concentrations that result in a constant mass

$$\overline{\phi_j} = \frac{\phi_j^{\mathrm{I}}\phi_j^{\mathrm{II}}}{\beta\,\phi_j^{\mathrm{II}} + (1-\beta)\,\phi_j^{\mathrm{I}}} \quad (0 \le \beta \le 1)\,. \tag{S26}$$

Figures S1-S5

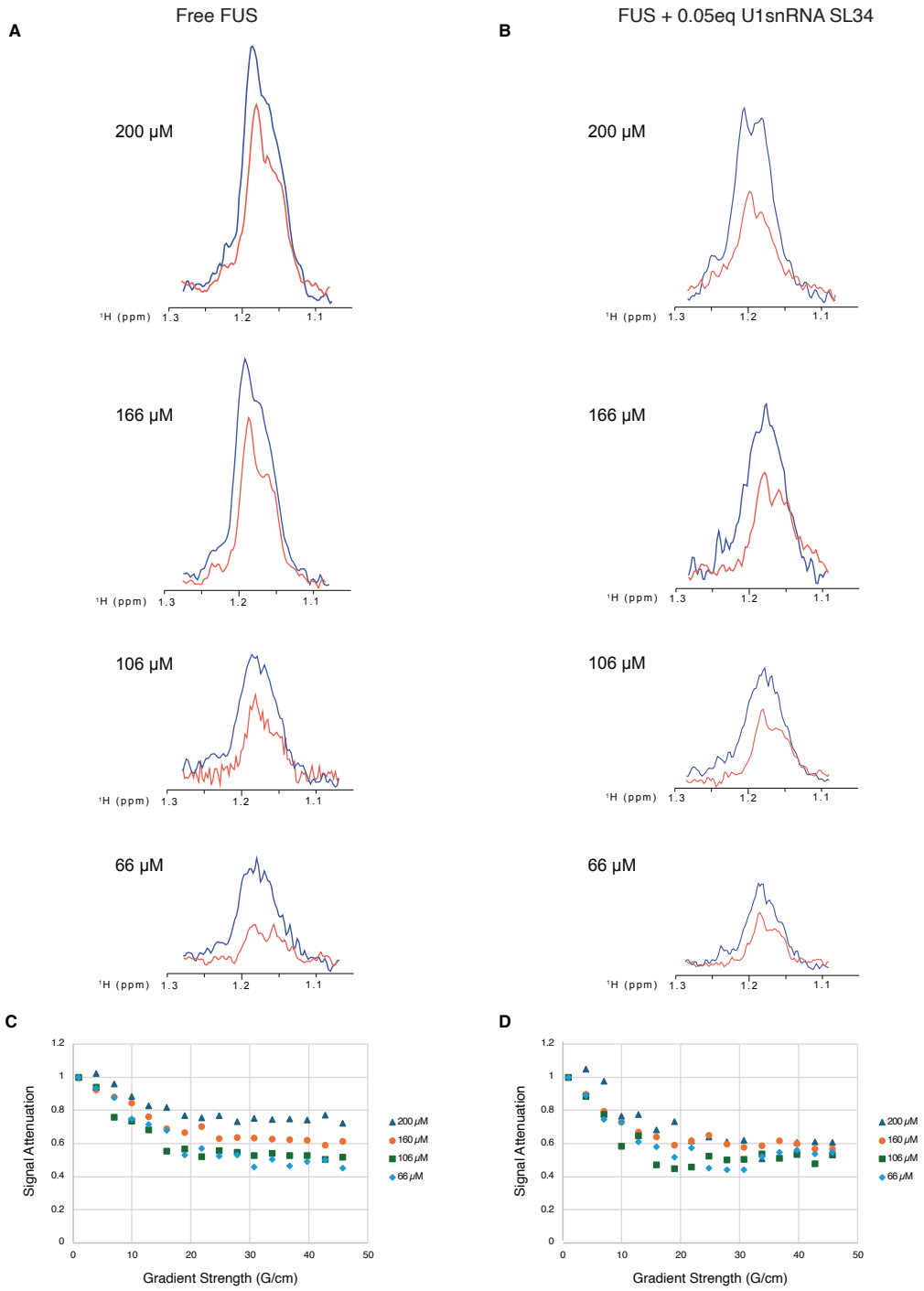

Figure S1: **DOSY spectra of FUS.** (**A-B**) Overlay of 1H 1D DOSY spectra acquired at the lowest gradient strength (0.96 G/cm, in blue) and highest gradient strength (45.72 G/cm, in red) for (A) FUS only and (B) FUS-U1snRNA SL34 complex. (C-D) Integral of selected spectral region shown in (A-B) at different FUS input concentrations as a function of applied gradient strength for (C) FUS only and (B) FUS-U1snRNA complex.

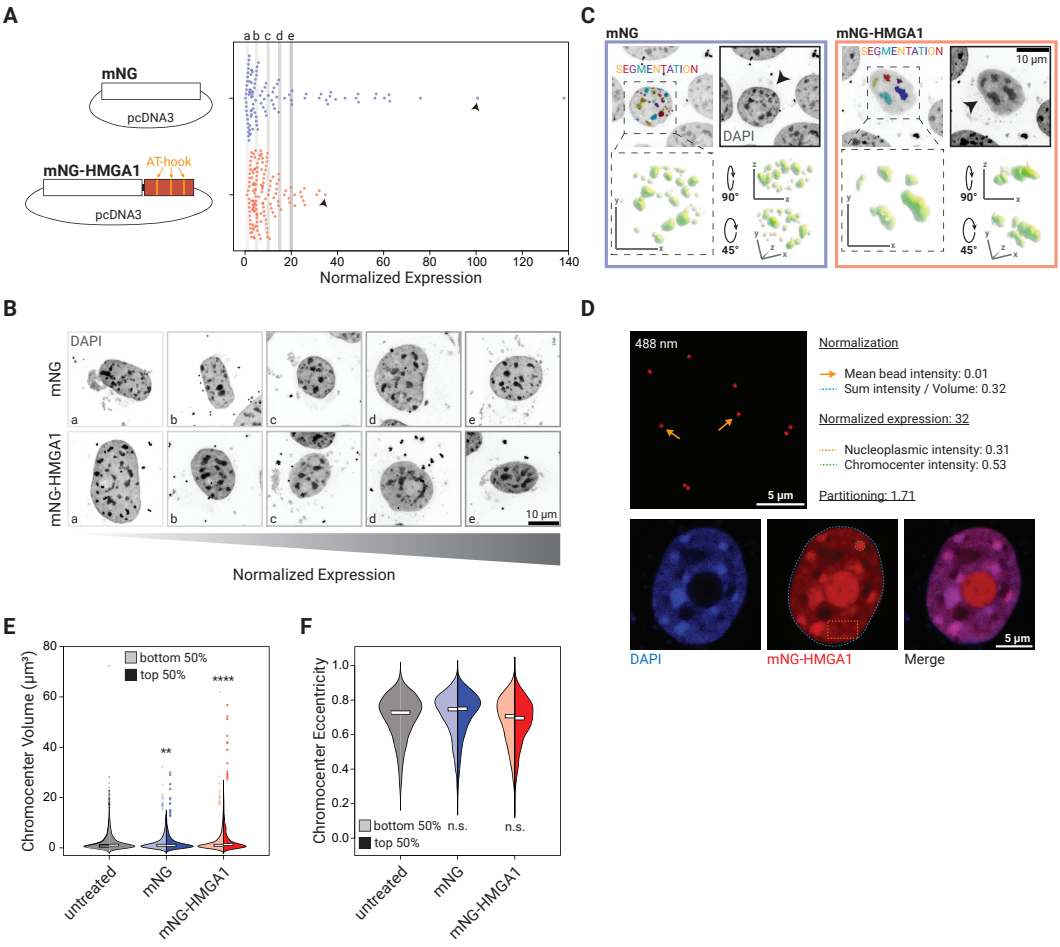

Figure S2: **Expression of HMGA1** (A) Schematic representation and expression level distribution of the mNG and mNG-HMGA1 constructs used in this study. Positions of the DNA-binding AT-hook motifs are highlighted in the schematic. Grey lines and letters in the graph correspond to images shown in panel (B), arrowheads point to the cells shown in panel (C). (B) Representative images of cells expressing different levels of mNG and mNG-HMGA1. 3D images are flattened onto a single plane via maximum intensity projection. (C) Illustration of chromocenter segmentation for cells expressing high levels of mNG and mNG-HMGA1. Arrowheads correspond to the graph in panel (A). (D) Schematic overview of the intensity-normalization procedure used in this study. (E) Volume distribution of individual chromocenters, corresponding to Fig. 1 D. 1% biggest chromocenters are shown as points. "**" represents $p < 0.01$, "****" represents $p < 0.0001$; Student's t-test for independent samples. (F) Eccentricity of individual chromocenters. "n.s." represents $p > 0.05$; Student's t-test for independent samples.

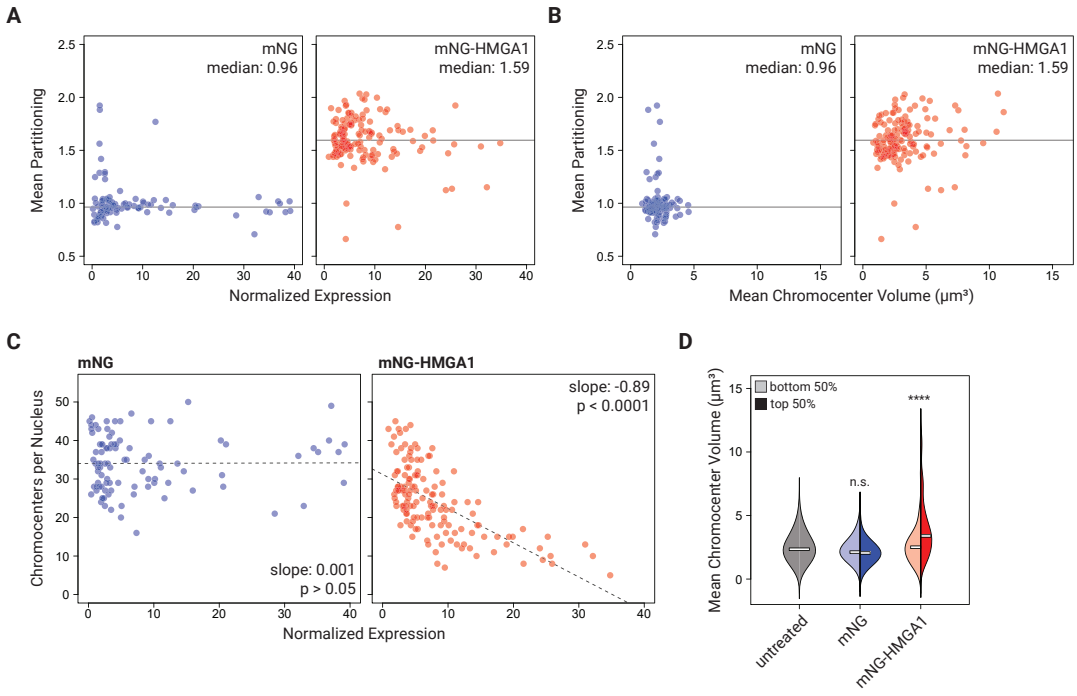

Figure S3: **Mean partitioning of mNG and mNG-HMGA1.** Mean partitioning of mNG and mNG-HMGA1 fluorescent signal into chromocenters by (A) overall level of expression and (B) mean chromocenter volume. Horizontal lines represent median values. (C) Quantification of chromocenter number per nucleus per cell in relation to differing levels of mNG and mNG-HMGA1 expression. Dashed line represents linear fit. (D) Mean chromocenter volume in bottom 50% and top 50% of mNG or mNG-HMGA1 expressing cells. "****" represents $p < 0.0001$, "n.s." represents $p > 0.05$; Student's t-test for independent samples.

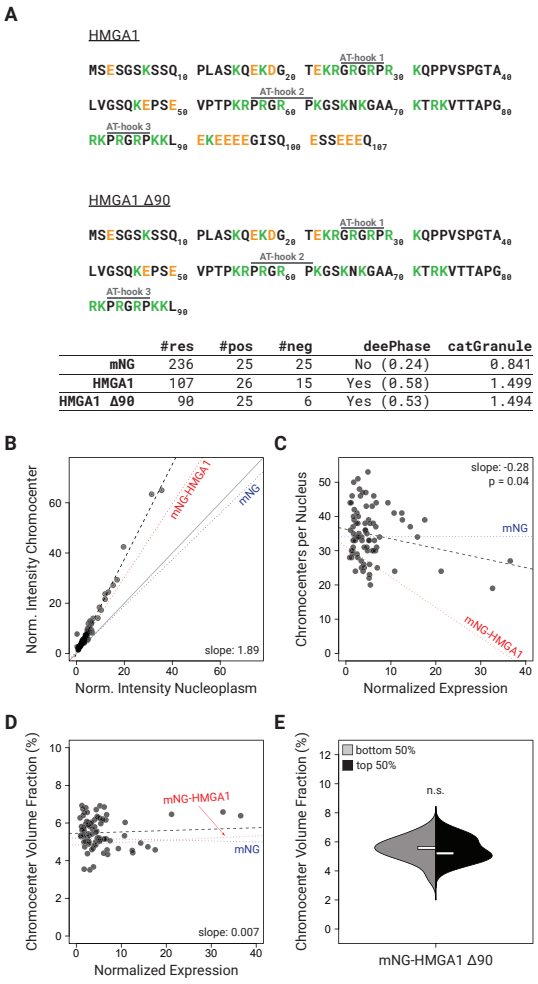

Figure S4: **Sequence and partitioning of HMGA1 and HMGA1 Δ90.** (A) Sequence representation of HMGA1 and HMGA1 Δ90, as well as phase-separation propensity predictions by deePhase (https://deephase.ch.cam.ac.uk/) and catGranule (http://www.tartaglialab.com/) for mNG, HMGA1 and HMGA1 Δ90. Charged residues are highlighted in green (positive) and orange (negative). Positions of DNA-binding AT-hook motifs are indicated above the sequence. (B) Quantification of partitioning of HMGA1 Δ90 (n = 84 cells) into chromocenters, corresponding to Fig. 1 B. Dashed line represents linear fit, dotted lines show comparison-fits for mNG (blue) and mNG-HMGA1 (red). (C) Quantification of chromocenter number per nucleus in relation to mNG-HMGA1 Δ90 expression level. (D) Chromocenter volume fraction in relation to HMGA1 Δ90 expression. (E) Chromocenter volume fraction in bottom 50% and top 50% of mNG-HMGA1 Δ90 expressing cells. "n.s." represents $p > 0.05$, Student's t-test for independent samples.

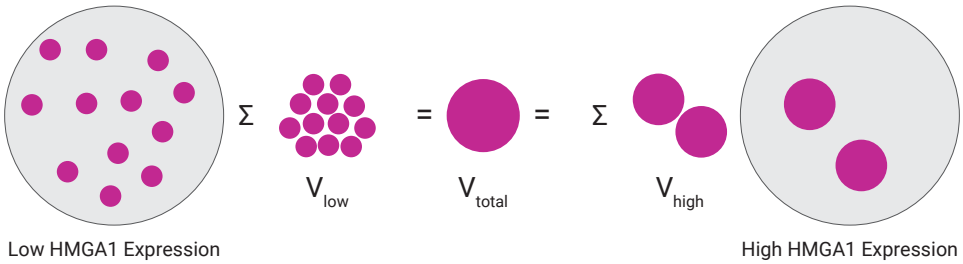

Figure S5: **Model of HMGA1-mediated chromocenter clustering and its effect on the total chromocenter volume.**