

# Mathematical model of nuclear speckle morphology

Shingo Wakao<sup>1</sup>, Noriko Saitoh<sup>2</sup>, Akinori Awazu<sup>1,3</sup>

<sup>1</sup> *Graduate School of Integrated Sciences for Life, Hiroshima University, Higashihiroshima, Hiroshima 739-8526, Japan*

<sup>2</sup> *Division of Cancer Biology, The Cancer Institute of Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan*

<sup>3</sup> *Research Center for the Mathematics on Chromatin Live Dynamics, Hiroshima University, Higashihiroshima, Hiroshima 739-8526, Japan*

**Nuclear speckles are nuclear bodies consisting of populations of small and irregularly shaped droplet-like molecular condensates that contain various splicing factors. Recent experiments have shown the following morphological features of nuclear speckles: (I) Each molecular condensate contains SON and SRRM2 proteins, and MALAT1 non-coding RNA surrounds these condensates; (II) In the normal interphase of the cell cycle, these condensates are broadly distributed throughout the nucleus in multicellular organisms. In contrast, the fusion of condensates leads to the formation of strongly condensed spherical droplets when cell transcription is suppressed; (III) SON is dispersed spatially in MALAT1 knocked-down cells, whereas MALAT1 is dispersed in SON knocked-down cells by the collapse of nuclear speckles. However, the detailed interactions among molecules that reveal the mechanisms of this rich variety of morphologies remain unknown. In this study, a coarse-grained molecular dynamics model of the nuclear speckle was developed considering the dynamics of SON, SRRM2 or SRSF2, MALAT1, and pre-mRNA as representative components of condensates. The simulations reproduced the abovementioned morphological changes, by which the interaction strength among the representative components of the condensates was predicted.**

## Introduction

Multicellular organisms contain various non-membranous nuclear bodies in the cell nucleus. Their dynamics and biophysical properties are intimately linked to physiological activities, such as gene regulation [1,2]. Typical nuclear bodies, such as the nucleolus and PML bodies, form a spherical shape during interphase, but they change their morphologies under certain cellular or experimental conditions and in diseased cells [1,3].

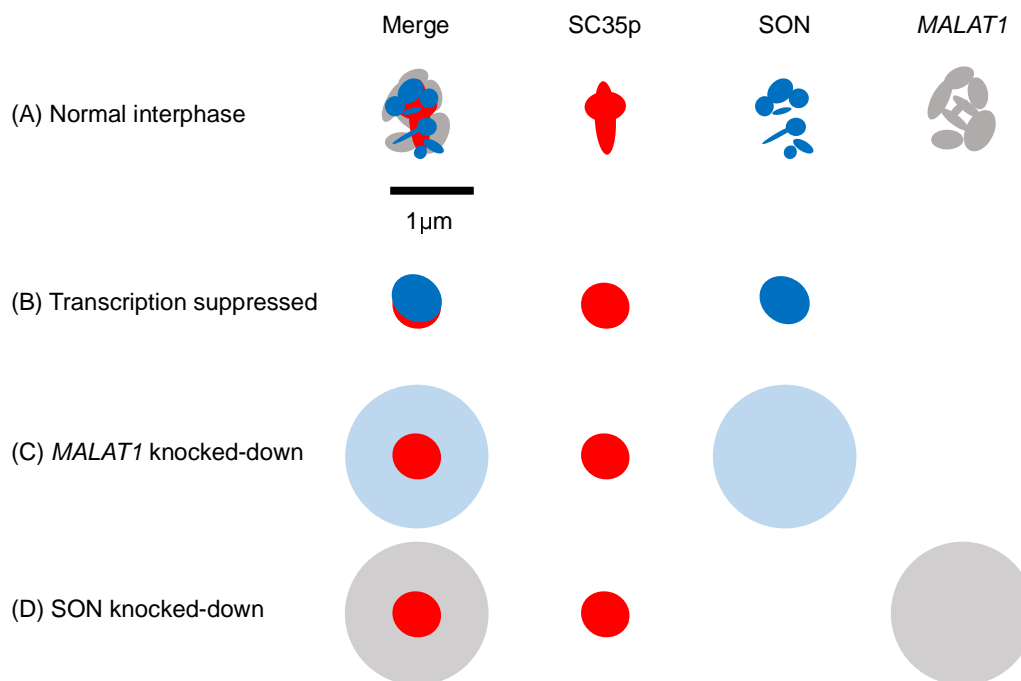
Nuclear speckles are one of the major nuclear bodies formed by over a hundred protein species, including various pre-mRNA splicing factors [4,5,6] and noncoding RNAs, like *MALAT1* [7,8]. Nuclear speckles are populations of small irregular-shaped droplet-like molecular condensates distributed throughout the nucleus, enriched in the inter-chromosome compartments [5,9,10,11,12,13,14,15,16]. Transcriptionally active genes are often found at the nuclear speckle periphery [11], and their transcripts are associated with or transit through nuclear speckles [17,18].

Nuclear speckles were originally identified via immunofluorescence using the monoclonal antibody mAb SC35 that was raised against biochemically isolated mammalian spliceosomes [19]. This antibody had been long considered to be based on immunoblots, and that it recognizes a pre-mRNA splicing factor of 35 kDa of SRSF2/SC35, with possible cross-reactivity to other splicing factors harboring repetitive serine and arginine (SR) sequences. A recent study re-characterized mAb SC35, revealing that its main target may be SRRM2 (formally known as SRm300), a spliceosome-associated protein [15]. Because the antigens remain somewhat elusive, we herein refer proteins probed with mAb SC35 as SC35p, which stands for mAb SC35-recognizing protein candidates. Among these, we have focused on SRRM2 and SRSF2. Recent studies involving high-resolution fluorescence imaging reported the following morphological features of nuclear speckles: I) During interphase, the core region of each small molecular condensate of nuclear speckles contained SON and SC35p, and *MALAT1* tended to surround them like a shell (Figure 1A) [20]; II) When all gene transcriptions were suppressed, the molecular condensates fused with each other and formed dense spherical condensates with stronger fluorescent intensity than those in normal interphase (Figure 1B) [13,21,22]; III) Conversely, when *MALAT1* was knocked-down, SC35p was sustained in the condensates but SON dispersed (Figure 1C) [20]; IV) Similarly, when SON was knocked-down, SC35p was sustained in the condensates but *MALAT1* dispersed (Figure 1D) [8,20,21]. In the latter two cases, nuclear speckles were collapsed.

The representative molecules regulating nuclear speckle morphology were identified via the aforementioned experiments. However, the detailed interactions among these molecules that could reveal the underlying mechanisms remain unclear. For other nuclear bodies, such as the nucleolus and paraspeckle, the development of mathematical models provided deeper insights into the relationships between their morphologies and molecular chemo-mechanical features [3,23,24]. Recently, a mathematical model of the molecular condensates of nuclear speckles was developed that accounted for the interactions among SON, SC35p, and *MALAT1* and reproduced their hierarchical condensations [20]. However, this model had some flaws. For example, it was easily expected that the dispersion of SON by *MALAT1* knockdown could not be reproduced because this model assumed the SON-SON attractive forces as well as SON-SC35p attractive

forces.

In this study, we aimed to reveal the mechanism and role played by each molecule of the nuclear speckle in determining its morphological features. To that end, we developed a coarse-grained molecular dynamics model of the local region of nuclear speckles with a few molecular condensates. SON, SC35p, *MALAT1*, and pre-mRNA were considered as the representative components of droplets. Simulations of the proposed model with appropriate assumptions of the effective interactions among these molecules reproduced all the above-mentioned morphological changes in the local region of the nuclear speckle. The results also suggested that SRRM2 and *MALAT1* were regarded as the core and shell components of core-shell structures in nuclear speckles.



**Figure 1** Illustrations of molecular distributions in the local region of nuclear speckles observed using fluorescence microscopy imaging under various conditions. Distributions of SC35p, SON, and *MALAT1* in (A) normal interphase, (B) transcriptional repression state, (C) *MALAT1* knocked-down state, and (D) SON knocked-down state [8,13,20,21,22].

## Materials and Methods

### Candidates of Proteins Probed with mAb SC35 (SC35p): SRSF2 and SRRM2

mAb SC35 has been used to observe nuclear speckles as this antibody recognizes SRSF2/SC35. However, a recent re-characterization study revealed that a major target for this antibody was actually SRRM2 [15]. Therefore, in this study, SRSF2 and SRRM2 were assumed as SC35p for their individual models. The model for assuming SC35 as SRSF2 was called the “SF2 model”, and that for assuming SC35 as SRRM2 was called the “RM2-model”.

### **Coarse-Grained Models of SON, SRRM2, and SRSF2**

All three focused proteins, SON, SRRM2, and SRSF2, are intrinsically disordered proteins that involve partially folded regions, and their intrinsically disordered regions (IDRs) occupy a large ratio of their amino acid sequences (Figure 2A, B) [4,25,26]. This indicates that they do not possess unique stable folding structures but change their shapes like random polymers. Therefore, each of these proteins was described using one soft spherical particle with their respective radii  $r^p$  ( $p$ : SON, SRRM2, SRSF2) for simplicity. Here, the particles representing SON, SRRM2, and SRSF2 were named SON, SRSF2, and SRRM2 particle, respectively. Each particle indicates the spatial region that the chain on the protein tends to occupy. Therefore, the excluded volume effects among these particles were assumed to be so weak that a small ratio of the regions in the two contacting particles was allowed to overlap with each other.

The radii  $r^p$  were assumed according to the arguments of random polymers with excluded

volume as  $r^p = [\text{ACd}] \left( \frac{L^p}{[\text{ACd}]} \right)^{\frac{3}{5}}$ , where  $[\text{ACd}] = 0.36$  (nm) indicates the average distance

between the edge of the amino group and that of the carboxyl group in each amino acid, and  $L^p$  indicates the effective length of proteins. In the present study, for the SON and SRSF2 models,  $L^p$  was defined as the edge-to-edge distance of proteins containing secondary structures along the amino acid sequence as follows:  $L^p = [\text{sum of all distances between N-terminals and C-terminals of partially folded regions (nm)}] + [\text{ACd}] \cdot [\text{Number of amino acids belonging to IDR}]$  (Figure 2A, B, Supplementary Figure S1). The three-dimensional coordinates of atoms and positions of folded regions in SON and SRSF2 were obtained in PDB format files from the UniProt database [<https://www.uniprot.org>] ID Q01130 and P18583 [27], where these features were inferred via Alpha-Fold2 [28,29]. Based on these estimates,  $r^{\text{SON}} = \sim 18.4$  nm and  $r^{\text{SRSF2}} = \sim 3.4$  nm were obtained.

In this study, SRRM2 was assumed to be a simple unfolded random polymer of amino acids because most of the regions were IDR, and no public data of the three-dimensional coordinates of atoms and positions of folded regions in SRRM2 were available (Figure 2C). Therefore,  $L^p = [\text{ACd}] \cdot [\text{Number of amino acids belonging to IDR}]$  was assumed, which gave  $r^{\text{SRRM2}} = \sim 20.8$  nm.

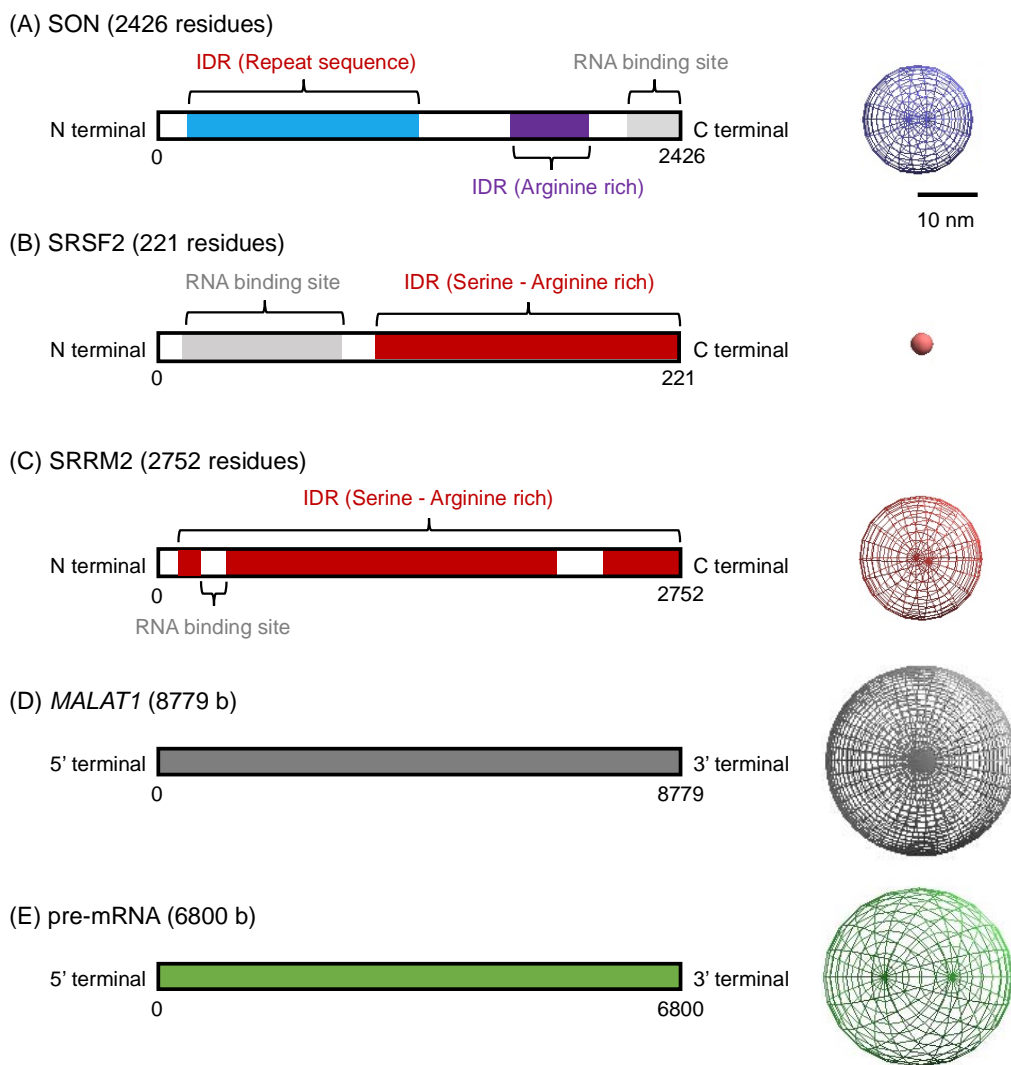
In the present cases, the rigidity of the random polymer is known to decrease with an increase in

the number of segments corresponding to  $\frac{L^p}{[ACd]}$ . The focused molecules were expected to be more rigid than random polymers in their central regions but more similar to them in their outer regions. Therefore, in this model, the rigidity parameters of these particles,  $q^p$ , were assumed to be proportional to the estimated conventional arguments of the rubber elasticity of ideal chains, as follows:  $q^p = \frac{k_B T}{L^p [ACd]^2}$ . Recent molecular dynamics simulations showed that the elastic stress-strain relations of random polymers with excluded volumes could be approximated by those of ideal chains when the polymer concentrations were not as high as those of solids [30]. Here,  $q^{SON} = \sim 1.43 \times 10^{-5} \text{ kg} \cdot \text{s}^{-2}$ ,  $q^{SRRM2} = \sim 1.16 \times 10^{-5} \text{ kg} \cdot \text{s}^{-2}$ , and  $q^{SRSF2} = \sim 2.35 \times 10^{-4} \text{ kg} \cdot \text{s}^{-2}$  were obtained.

### Coarse-Grained Models of *MALAT1* and Pre-mRNAs

In the present model, RNAs were assumed to be unfolded random polymers, and each RNA chain was described as one soft spherical particle with their respective radii  $r^R$  ( $R$ : *MALAT1* or pre-mRNA). Here, the particles representing *MALAT1* and pre-mRNA were respectively called *MALAT1* and pre-mRNA particles. Similar to protein models, each particle indicated the spatial region that the chain on the protein tended to occupy (Figure 2D, E); therefore, the excluded volume effects among these particles were assumed to be so weak that a small ratio of the regions in the two contacting particles were allowed to overlap with each other.

The radii  $r^R$  were assumed as  $r^p = [5' - 3' \text{ distance}][\text{number of nucleotides}]^{\frac{3}{5}}$ , where  $[5' - 3' \text{ distance}]$  was assumed as 0.3 nm, and the number of nucleotides for *MALAT1* was assumed as 8779 bp. In the present study, the number of pre-mRNAs was assumed to be a unique value of 6.8 kb that was the same as the average base pairs of gene-transcribed regions in the human genome, estimated based on the Refseq data [31]. The rigidity parameters of these particles,  $q^R$ , were also estimated in the same way as the above-mentioned cases of proteins, where  $q^R = \frac{k_B T}{[\text{number of nucleotides}][5' - 3' \text{ distance}]^2}$  was assumed. Further,  $r^{MALAT1} = \sim 33.0 \text{ nm}$ ,  $r^{\text{pre-mRNA}} = \sim 29.9 \text{ nm}$ ,  $q^{MALAT1} = \sim 5.75 \times 10^{-6} \text{ kg} \cdot \text{s}^{-2}$ , and  $q^{\text{pre-mRNA}} = \sim 6.76 \times 10^{-6} \text{ kg} \cdot \text{s}^{-2}$  were obtained.



**Figure 2** Construction of a coarse-grained model of each molecule. Lengths and local regions in proteins and RNAs (left) and soft spherical particle models (right) of (A) SON, (B) SRSF2, (C) SRRM2, (D) *MALAT1*, and (E) pre-mRNA. IDR, intrinsically disordered region.

### Motion of Each Molecule

Because each biomolecule was expected to exhibit three-dimensional Brownian motion in the nucleus, the position of the  $i$ -th particle with radius  $r_i$  in the  $(x, y, z)$  three-dimensional space,  $\mathbf{x}_i = (x_i, y_i, z_i)$ , was assumed to obey the Langevin equation as follows:

$$\gamma_i \frac{d\mathbf{x}_i}{dt} = -\frac{\partial}{\partial \mathbf{x}_i} V + \mathbf{R}_i(t) \quad (1)$$

where  $\gamma_i$  and  $\mathbf{R}_i(t)$  are the coefficients of drag force and Gaussian white noise, respectively, playing the role of the random force from the nucleoplasm to  $i$ -th particle, obeying  $\langle \mathbf{R}_i(t) \rangle = 0$  and  $\langle \mathbf{R}_i(t) \mathbf{R}_{i'}(t) \rangle = 6\gamma_i k_B T \delta_{ii'} \delta(t-s)$ , with Boltzmann constant  $k_B$  and temperature  $T$ , where  $k_B T = 4.141947 \times 10^{-21} \text{ kg m}^2 \text{ s}^{-2}$  ( $T = 300 \text{ K}$ ). Here,  $\delta_{ij}$  indicates the Kronecker delta and  $\delta(\dots)$  indicates the Dirac delta function. The term  $V$  indicates that the potential of the system involves the potential of forces among the particles and effect of the boundary conditions. The drag coefficient  $\gamma_i$  for each particle was assumed as  $6\pi\eta m r_i$ , with the viscosity of the nucleoplasm  $\eta = 0.64 \text{ kg m}^{-1} \text{ s}^{-1}$  [32].

Particle motion was assumed to be restricted to a spherical region with a radius of  $500 \text{ nm}$ . Note that  $500 \text{ nm}$  was slightly larger than the characteristic length of the inter-chromosome components in the cell nucleus. Such a wide space was employed in the simulations to reveal the mechanism of the droplet morphologies from the self-assembling behaviors of nuclear speckle components without the influence of other factors, such as chromatin territories and other nuclear bodies.

### Model of Effective Interactions among Molecules

The potential of system  $V$  to provide the forces working on and among the particles is given as:

$$V = V_{\text{collision}} + V_{\text{attract}} + V_{\text{boundary}} \quad (2)$$

where  $V_{\text{col}}$  is the potential of the collisional interactions among the particles owing to their excluded volumes,  $V_{\text{attract}}$  is the potential of the attractive interactions among the particles owing to the electrostatic interactions or biochemically specific affinities among molecules, and  $V_{\text{boundary}}$  is the potential of the effect of the boundary wall to restrict particles in the above-mentioned spherical region.

The potential  $V_{\text{col}}$  was assumed via the following:

$$V_{\text{collision}} = \sum_{i < i'} \theta((r_i + r_{i'}) - |\mathbf{x}_i - \mathbf{x}_{i'}|) C_{i,i'} k_{i,i'}^c \left[ -2 \left( \frac{r_i + r_{i'}}{|\mathbf{x}_i - \mathbf{x}_{i'}|} \right)^6 + \left( \frac{r_i + r_{i'}}{|\mathbf{x}_i - \mathbf{x}_{i'}|} \right)^{12} + 1 \right] \quad (3)$$

where  $k_{i,i'}^c = k_0^c \sqrt{q_i q_{i'}}$  represents the repulsion strength,  $C_{i,i'}$  is the value of  $C_{i,i'} = 1$  ( $0$ ) when repulsion works (does not work) in between  $i$ -th and  $i'$ -th particles.

$\theta$  is the Heaviside step function, defined as follows:

$$\theta(y) = \begin{cases} 1 & (y \geq 0) \\ 0 & (y < 0) \end{cases} \quad (4)$$

The potential  $V_{attract}$  was assumed by the following equation:

$$V_{attract} = \sum_{i < i'} A_{i,i'} k_{i,i'}^a \left[ -2 \left( \frac{(r_i + r_{i'}) - d_{i,i'}}{|x_i - x_{i'}|} \right)^6 + \left( \frac{(r_i + r_{i'}) - d_{i,i'}}{|x_i - x_{i'}|} \right)^{12} \right] \quad (5)$$

where  $k_{i,i'}^a$  represents the strength of the attractive force, and  $A_{i,i'}$  is the value that  $A_{i,i'} = 1$  (0) when attractive force works (does not work) in between  $i$ -th and  $i'$ -th particles.  $d_{i,i'}$  was assumed to be smaller than  $r_i + r_{i'}$  because the particles representing the space occupied by polymers were assumed to be so soft that the stable distance between two of them was expected to be nearer than the sum of their radii when attractive forces work in between them. In this study,  $d_{i,i'}$  was chosen to satisfy the following relation:

$$\theta \left( (r_i + r_{i'}) - d_{i,i'} \right) k_{i,i'}^c \left[ -2 \left( \frac{r_i + r_{i'}}{(r_i + r_{i'}) - d_{i,i'}} \right)^6 + \left( \frac{r_i + r_{i'}}{(r_i + r_{i'}) - d_{i,i'}} \right)^{12} + 1 \right] = \frac{k_B T}{2} \quad (6)$$

where  $d_{i,i'}$  provided the typical length of the region where two particles could overlap when only the repulsion represented by Eq. (3) was assumed to work in between them (Supplementary Figure S2).

The potential  $V_{boundary}$  was assumed by the following equation:

$$V_{boundary} = \sum_{i < i'} \theta \left( (R - r_i) - |x_i| \right) k^w q_i \left[ -2 \left( \frac{R - r_i}{|x_i|} \right)^6 + \left( \frac{R - r_i}{|x_i|} \right)^{12} + 1 \right] \quad (7)$$

where  $k^w$  represents the strength of the restriction from the boundary wall, and  $R$  represents the radius of the simulation space, assumed as 500 nm.

### Simulation and Data Analysis Method

The time integral of Langevin Eq. (1) was calculated numerically using the Euler-Maruyama method with a unit step of  $3.16 \times 10^{-7}$  (sec). For the initial condition (time = 0 (sec)), the particles were placed randomly in the sphere within a radius of 250 nm. The radial distributions of distances between two particles are defined as  $R^{X,Y}(dist) = [\text{Probability that the distance between X and Y particles} = dist]$ . Note that the accumulation or dispersion features of the X and Y particle populations were provided by the profile, in particular the height, width, and tail of the first peak nearest from  $dist = 0$  of  $R^{X,Y}(dist)$ .

Recent experimental studies suggested that SC53p always condensates independently of the cell state [8,13,20,21,22]. Therefore, to evaluate the morphological features of particle condensates,



the radial distributions of SC53p (SRRM2 or SRSF2), SON, *MALAT1*, and pre-mRNA particles from SC53p (SRRM2 or SRSF2) particles were measured; the radial distribution of X particle (X = SON, SRSF2, SRRM2, *MALAT1*, or pre-mRNA) was defined as  $R^X(dist) = [\text{Probability that the distance between X and SC53p (SRRM2 or SRSF2) particles} = dist]$ . Additionally, the radial distributions of SON particles from other SON particles and those of *MALAT1* particles from other *MALAT1* particles is defined as  $R^{SS}(dist) = [\text{Probability that the distance between a SON and other SON particles} = dist]$  and  $R^{MM}(dist) = [\text{Probability that the distance between a } *MALAT1* \text{ and other } *MALAT1* \text{ particles} = dist]$ , which were also measured to evaluate their condensations or dispersions. These distributions were estimated using the particle configurations between the time interval 403–474 (*sec*), where the radial distribution of the typical state was obtained without the influence of the initial particle configuration.

## Results

### Simulation of RM2 Model using Appropriately Designed Parameters of Specific Molecular Interactions Reproducing Droplet Morphologies in Nuclear Speckles

The RM2 model with the following specific intermolecular interactions was simulated (Figure 3A, Supplementary Figure S2A), which could reproduce the experimentally observed morphological features of droplets of nuclear speckles (Figure 1):

- I) Repulsive forces caused by the excluded volumes of molecules between two particles representing the same molecular species were assumed as  $C_{i,ir} = 1$  and  $A_{i,ir} = 0$ . The parameter providing the strength of the repulsive forces,  $k_0^c$ , was given as  $1.4 \times 10^3 k_B T$ .
- II) A specific attractive force between SRRM2 and SON particles and between SRRM2 and pre-mRNA particles was assumed as  $C_{i,ir} = 0$ ,  $A_{i,ir} = 1$ . The attractive strength parameter was assumed uniformly as  $k_{i,ir}^a = 15 k_B T$ .
- III) The specific attractive force between SRRM2 and *MALAT1* particles and between SON and *MALAT1* particles was assumed as  $C_{i,ir} = 0$ ,  $A_{i,ir} = 1$ . The attractive strength parameter was assumed uniformly as  $k_{i,ir}^a = 10 k_B T$ .
- IV) The interaction between SON and pre-mRNA particles and that between *MALAT1* and pre-mRNA particles were assumed to be repulsive, as  $C_{i,ir} = 1$ ,  $A_{i,ir} = 0$ , and  $k_0^c = 1.4 \times 10^3 k_B T$ .

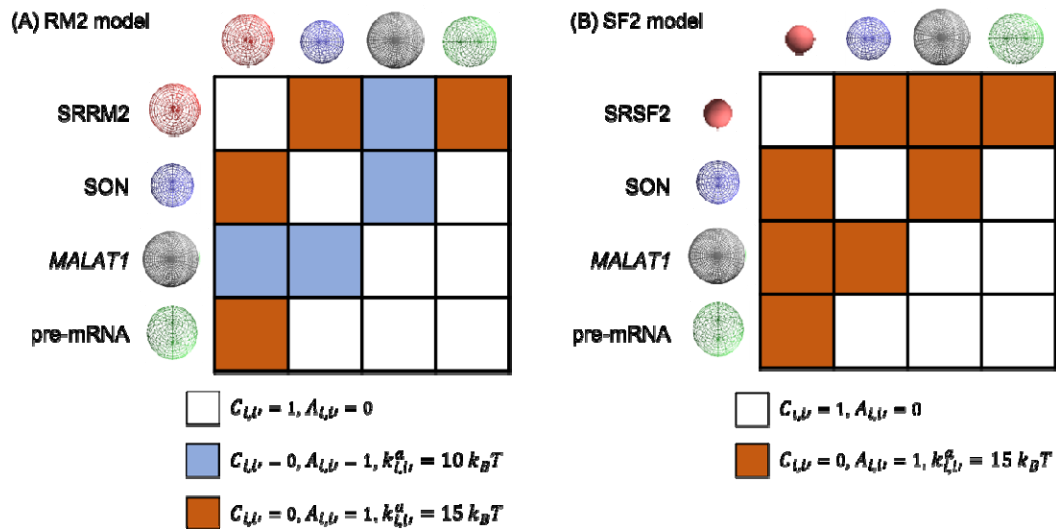
First, the numbers of SRRM2, SON, *MALAT1*, and pre-mRNA particles were 60, 60, 60, and 420, respectively, and was simulated to consider the behavior of droplets of nuclear speckles under normal conditions during interphase. The present model exhibited broadly distributed particle condensates (Figure 4A). Further,  $R^{SON}(dist)$ ,  $R^{SRRM2}(dist)$ , and  $R^{MALAT1}(dist)$  exhibited a

peak at the similar near , but involved a lower peak and longer tail than and . This fact indicated that the condensates of SRSF2 and SON particles were surrounded by *MALAT1* particles similar to the core-shell structure observed in experiments (Figure 1A) [14,20].

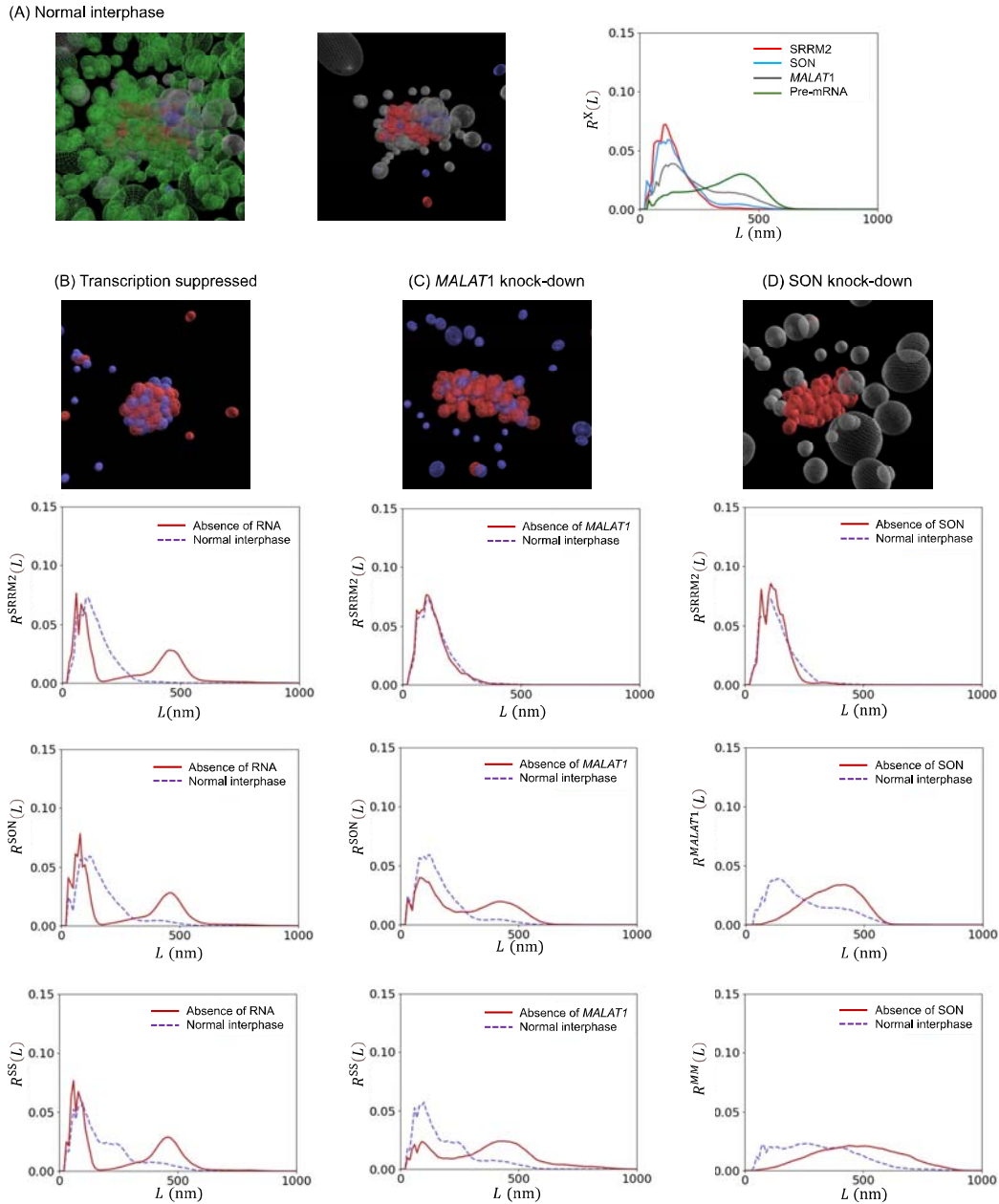
Second, to consider the situation after transcription activity was suppressed in the entire nucleus, the numbers of SRRM2, SON, *MALAT1*, and pre-mRNA particles were 60, 60, 0, and 0, respectively. The positions of the large peaks that were the nearest from of , and had shifted to close to , and their tails became shorter than the previous case (Figure 4B). This indicates that SRRM2 and SON had formed a denser condensate than the previous case, in which denser droplets were formed owing to the transcription suppression observed in the experiments (Figure 1B) [13,21,22].

Third, to consider the *MALAT1* knocked-down state, the numbers of SRRM2, SON, *MALAT1*, and pre-mRNA particles were 60, 60, 0, and 420, respectively. In this case, the qualitative features of were almost unchanged from those in the first case. However, and were broader, and the peak nearest to was much smaller than that in the first case (Figure 4C). This indicates that SON was spatially dispersed, as observed in *MALAT1* knockdown experiments (Figure 1C) [20].

Finally, to consider the SON knocked-down state, the numbers of SRRM2, SON, *MALAT1*, and pre-mRNA particles were 60, 0, 60, and 420, respectively. In this case, the qualitative features of were almost unchanged from those in the first case. However, and were broader, and the peak nearest to was much smaller than that in the first case (Figure 4D). This indicates that *MALAT1* was spatially dispersed, as observed in the SON knockdown experiments (Figure 1D) [8,20,21].



**Figure 3** Matrices of interaction networks among particles in (A) RM2 and (B) SF2 models with appropriately designed parameters. White boxes indicate repulsion, blue boxes indicate weak attraction, and brown boxes indicate strong attraction between a pair of particles. Each profile of interaction potential between each pair of particles is shown in Supplementary Figure S2.



**Figure 4** Simulation results of the RM2 model with appropriately designed parameters. (A) Typical snapshots of the simulations (left), the same snapshots but without pre-mRNA (middle), and radial distributions of particles,  $R^{SRRM2}(L)$ ,  $R^{SON}(L)$ ,  $R^{MALAT1}(L)$ , and  $R^{pre-mRNA}(L)$

(right) in the case of normal interphase. (B–D) Typical snapshots of the simulations without pre-mRNA (first),  $R^{SRRM2}(L)$  (second),  $R^{SON}(L)$  or  $R^{MALAT1}(L)$  (third), and  $R^{SON}(L)$  or  $R^{MALAT1}(L)$  and  $R^{SS}(L)$  or  $R^{MM}(L)$  (fourth), in the case that (B) transcription was suppressed, (C) *MALAT1* was knocked-down, and (D) SON was knocked-down. Each particle color indicates molecular species as explained in Figure 2.

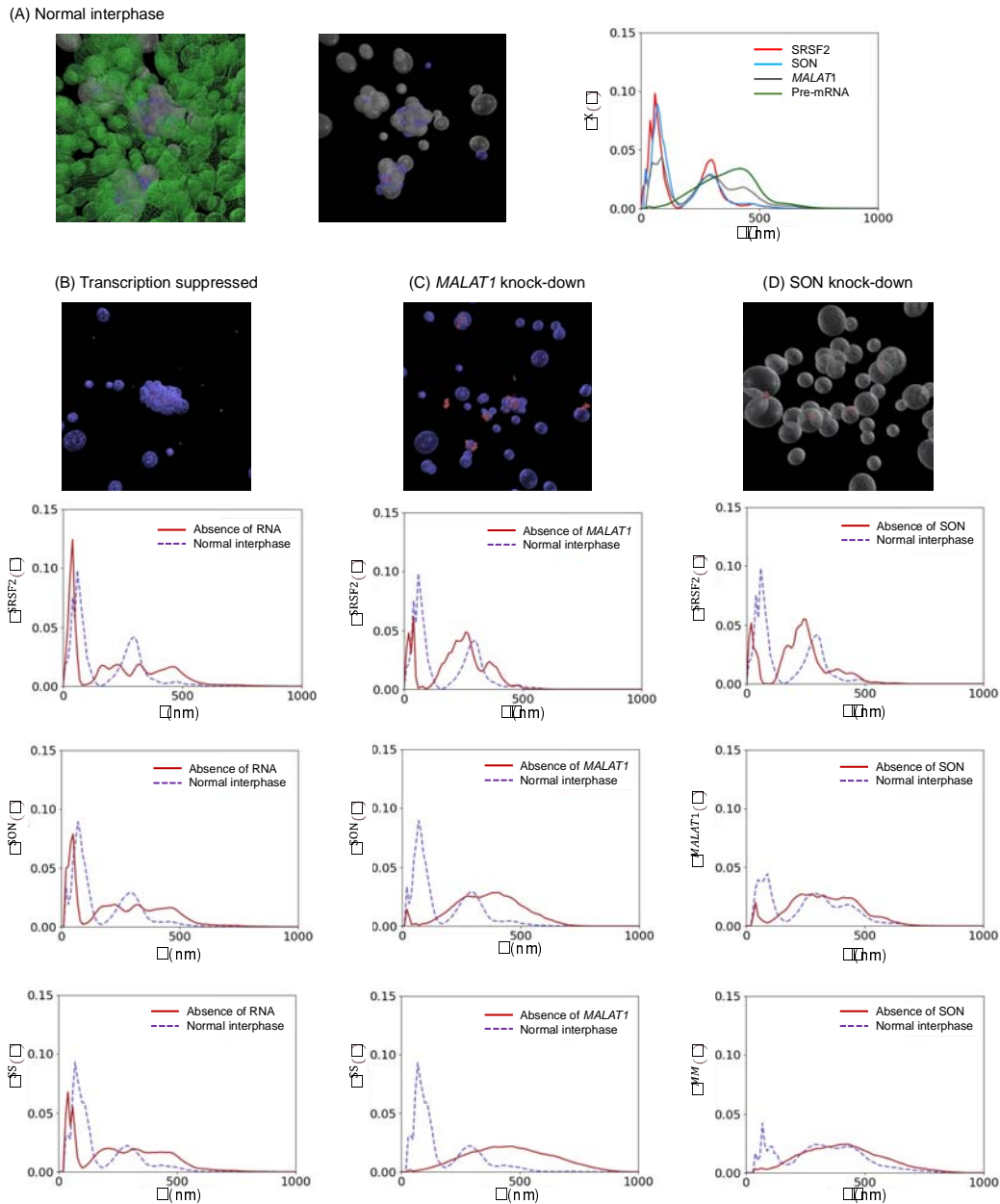
### **Simulation of SF2 Model using Appropriately Designed Parameters of Specific Molecular Interactions Reproducing Droplet Morphologies in Nuclear Speckles**

The SF2 model could also exhibit the following morphological features of particle condensates similar to the experimentally observed droplets of nuclear speckles (Figure 1) when the appropriate specific intermolecular interactions were simulated (Figure 3B, Supplementary Figure S2B):

- I) The interactions among SON, SRSF2, and pre-mRNA were assumed to be the same as those among SON, SRRM2, and pre-mRNA particles assumed in the previous RM2 model. The interaction between *MALAT1* and pre-mRNA particles was assumed to be the same as in the previous SF2 model.
- II) Specific attractive forces were assumed between SRSF2 and *MALAT1* particles and between SON and *MALAT1* particles as  $C_{i,i'} = 0$ ,  $A_{i,i'} = 1$ , and  $k_{i,i'}^a = 15 k_B T$ .

Note that the appropriate interaction strengths between *MALAT1* particles and other particles seemed to change slightly, which may have been owed to different particle sizes between SRSF2 and SRRM2 particles. Additionally, the number of SRSF2 particles was assumed as twice that of SON and *MALAT1* particles as referred to the previously proposed mathematical model [20], where the number of SRSF2, SON, *MALAT1*, and pre-mRNA particles were assumed as 120, 60, 60, and 420, respectively. With these assumptions, the SF2 model could qualitatively exhibit similar behaviors of SON, SC35p, and *MALAT1* particles to those obtained in the previous RM2 model except for the following facts (Figure 5, Supplementary Figure S3).

- i) In the normal interphase simulation,  $R^{SON}(dist)$ ,  $R^{SRSF2}(dist)$ , and  $R^{MALAT1}(dist)$  exhibited similar profiles near  $dist = 0$ , and  $R^{MALAT1}(dist)$  did not show a longer tail (Figure 5A) than in the RM2 model (Figure 4A). Furthermore, SRSF2 and SON particles were not surrounded by *MALAT1* particles, which differed from the core-shell structure observed in experiments [14,20].
- ii) In the *MALAT1* knockdown and SON knockdown simulations, the large peaks of  $R^{SRSF2}(dist)$  near  $dist = 0$  appeared but their heights became lower (Figure 5C–D), which differed from those in the RM2 model (Figure 4C–D). These facts indicated that the SRSF2 condensation was also weakened because of *MALAT1* or SON knockdown, which differed from past experiments [8,20,21].



**Figure 5** Simulation results of the SF2 model with appropriately designed parameters. (A) Typical snapshots of the simulations (left), the same snapshots but without pre-mRNA (middle), and radial distributions of particles,  $R^{\text{SRSF2}}(L)$ ,  $R^{\text{SON}}(L)$ ,  $R^{\text{MALAT1}}(L)$ , and  $R^{\text{pre-mRNA}}(L)$  (right) in the case of normal interphase. (B–D) Typical snapshots of the simulations without pre-mRNA (first),  $R^{\text{SRSF2}}(L)$  (second),  $R^{\text{SON}}(L)$  or  $R^{\text{MALAT1}}(L)$  (third), and  $R^{\text{SON}}(L)$  or  $R^{\text{MALAT1}}(L)$  and  $R^{\text{SS}}(L)$  or  $R^{\text{MM}}(L)$  (fourth), in the case that (B) transcription was suppressed, (C) *MALAT1* was knocked-down, and (D) SON was knocked-down. Each particle color indicates molecular species as explained in Figure 2.

## Simulations using Interaction Parameters Assumed by Electrostatic Properties of Molecules

Regarding the electrostatic properties of molecules, the interaction parameters between each pair of particles can be assumed as follows:

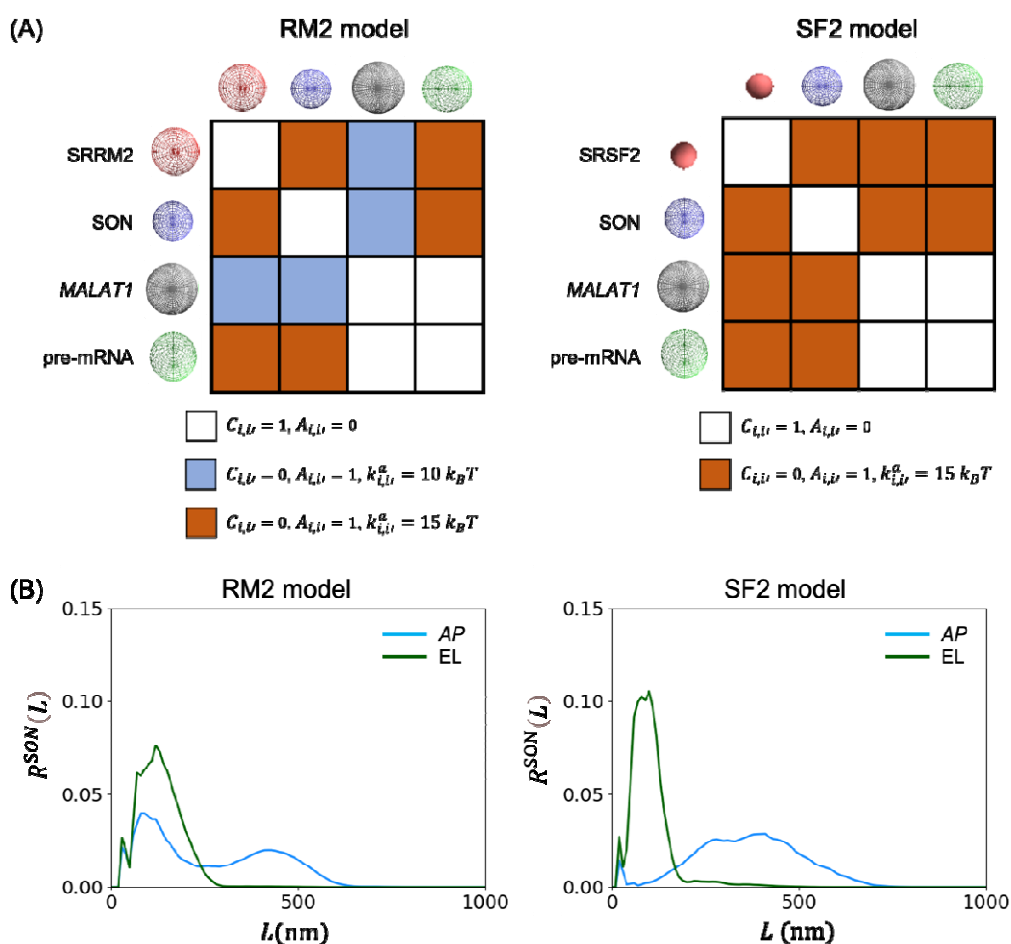
The IDRs of SRRM2 contained 642 serine and 468 arginine residues, and those of SRSF2 contained 42 serine and 35 arginine residues. Therefore, SC35p was expected to exhibit a negative charge. Conversely, SON was expected to exhibit a positive charge because the serine-arginine-rich IDR contains more number of arginine residues ( $n = 87$ ) than serine residues ( $n = 67$ ). Therefore,  $C_{i,ir} = 0$  and  $A_{i,ir} = 1$  may be assumed between the SC35p and SON particles because the electrostatic attractive forces may work in between these particles (Figure 6A, B). Conversely,  $C_{i,ir} = 1$  and  $A_{i,ir} = 0$  may be assumed between two SRRM2 particles, two SRSF2 particles, and two SON particles because both particles in each pair have the same charges (Figure 6A, B).

RNAs are negatively charged polymers. Therefore,  $C_{i,ir} = 1$  and  $A_{i,ir} = 0$  were assumed between two *MALAT1* particles, two pre-mRNA particles, and *MALAT1* and pre-mRNA particles (Figure 6A).

Both SRRM2 and SRSF2 were expected to be involved in RNA-binding sites [27]. The binding of two molecules through such specific binding sites is generally expected to be stronger than electrostatic forces. Therefore, although the total charges of SRRM2 and SRSF2 were assumed to be negative,  $C_{i,ir} = 0$  and  $A_{i,ir} = 1$  could be assumed between SC35p and RNA (*MALAT1* and pre-mRNA) particles (Figure 6A, B).

Similar to SC35p, SON also contains RNA-binding sites [27,33]. Additionally, the total charge of SON was expected to be positive. Therefore, an attractive force between the SON and RNA particles was naturally assumed as  $C_{i,ir} = 0$  and  $A_{i,ir} = 1$ . Here, the attraction strength between SON and RNA particles was expected to be smaller than that between SON and SC35p, because RNAs were expected to have a stronger negative charge than SC35p.

Notably, the assumptions of the interactions among particles in the previous section were consistent with the above-mentioned intuitive molecular interactions based on electrostatic forces, except for the interactions between SON and pre-mRNA. Therefore, the RM2 and SF2 model simulations were performed using the same interaction parameters among molecules as in the previous arguments, except for those between SON and pre-mRNA, where  $C_{i,ir} = 0$ ,  $A_{i,ir} = 1$ , and  $k_{i,ir}^a = 15 k_B T$  were assumed. In these cases, SON dispersion did not occur even when the model was simulated without *MALAT1* (Figure 6C, D), indicating that the spatial dispersion of SON due to *MALAT1* knockdown (Figure 1C) could not be reproduced in either model.



**Figure 6** Simulation results of RM2 and SF2 models with parameters based on electrostatic properties. (A) Matrices of interaction networks among particles. White boxes indicate repulsion, blue boxes indicate weak attraction, and brown boxes indicate strong attraction between a pair of particles. (B) Radial distributions of SON particles in RM2 and SF2 models with appropriately designed parameters (AP) and electrostatic property-based parameters (EL) in the *MALAT1* knockdown state.

## Discussion

In this study, a coarse-grained molecular dynamics model, with SON, SC35p (SRRM2 or SRSF2), *MALAT1*, and pre-mRNA as the representative components of condensates, was developed to simulate the morphological behavior of droplet-like condensates of nuclear speckles. Simulations of the model assuming SRRM2 as SC35p reproduced various experimentally observed cell state-dependent droplet morphologies. Simulations of the model

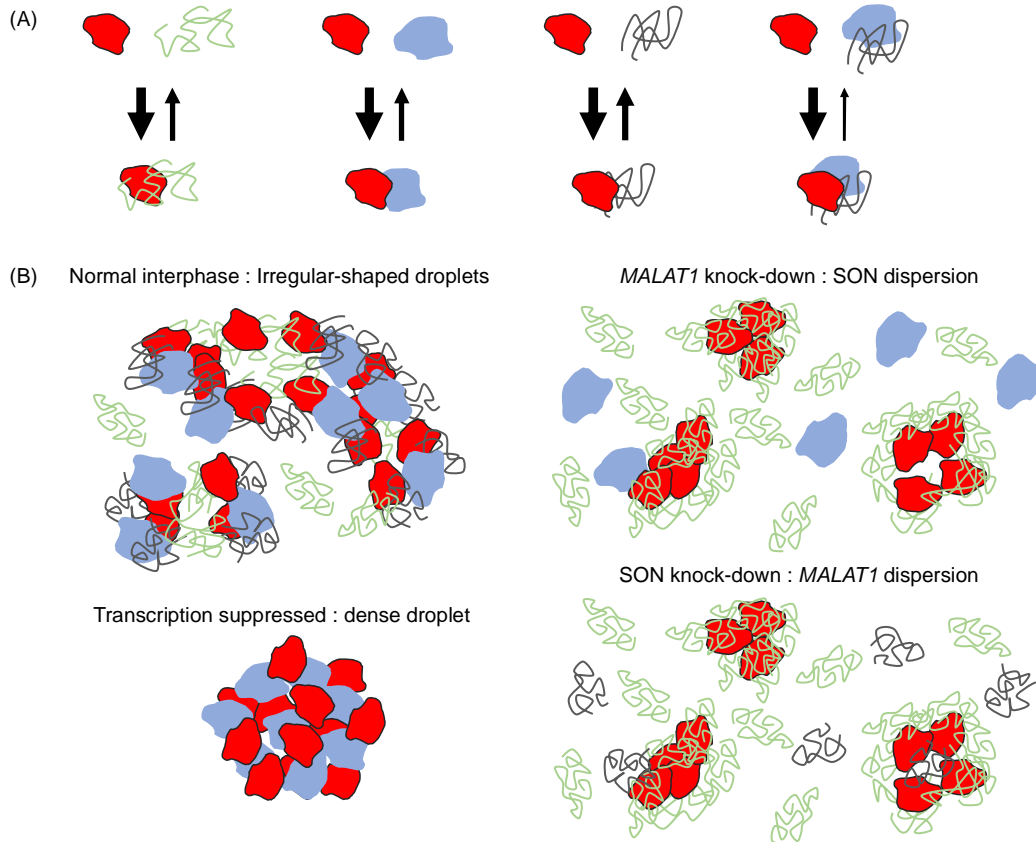
assuming SRSF2 as SC35p also exhibited similar behaviors but some features were inconsistent with experimentally observed phenomena. These facts suggested that SC35 should be SRRM2 rather than SRSF2, supporting a recent proposal [15]. Additionally, the effective interactions and their strengths among the representative components of nuclear speckles were predicted through these simulations.

Our findings suggest the following mechanism for the cell state-dependent morphology of nuclear speckles as well as the roles played by the representative molecules (Figure 7):

- I) SRRM2 exhibited attractive interactions with all other representative molecules of nuclear speckle components; SRRM2-SON interaction was assumed similar (Figure 3–4) or slightly weaker than SRRM2-pre-mRNA interaction (Supplementary Figure S4), SRRM2-MALAT1 interaction was assumed slightly weaker than SRRM2-SON (Figure 7A), and attractive interactions also occurred between SON and *MALAT1*.
- II) In the normal interphase of the cell nucleus, pre-mRNAs and SON-*MALAT1* complexes are considered to cluster around SRRM2. However, they were expected to spread spatially owing to their mutual repulsive forces. Notably, the attractive interaction between SRRM2 and SON or *MALAT1* was similar or slightly weaker than that between SRRM2 and pre-mRNA. However, to dissociate from the SON-MALAT1 complex, SRRM2 must dissociate from SON and MALAT1 almost simultaneously, so that the attractive interaction between SRRM2 and SON-MALAT1 complex could be more effective than that between SRRM2 and pre-mRNA (Figure 7A). Therefore, SRRM2 distributed to connect them, which is why the nuclear speckle component was broadly distributed (Figure 7). In such a broad molecular assembly, SON tended to be distributed nearer to *MALAT1*, because the attractive force between SON and SRRM2 was stronger than that between SRRM2 and *MALAT1*.
- III) When transcription of the entire nucleus was suppressed, few RNAs existed in the nucleus. Therefore, SON and SRRM2 formed larger dense condensates owing to the mutual attractive interaction between them, without competition between SON and pre-mRNA (Figure 7).
- IV) When *MALAT1* was knocked down, SON and pre-mRNA competed for binding with SRRM2. However, more pre-mRNA could bind to SRRM2 than SON even if the affinity of SON to SRRM2 was similar or slightly weaker than that of pre-mRNA. This is because the volume fraction of pre-mRNA was much larger than that of SON; hence, SRRM2 tended to interact with pre-mRNA more frequently than SON. The results showed that most of the SRRM2 was bound to pre-mRNA, while SON was excluded; therefore, SRRM2 sustained its condensate but SON dispersed spatially.



Similarly, SRRM2 sustained its condensate, but *MALAT1* dispersed spatially when SON was knocked down (Figure 7).



**Figure 7** (A) Illustrations of affinities among SON (blue), SRRM2 (red), MALAT1 (black), and pre-mRNA (green). (B) Mechanism and molecular roles of morphological changes in molecular accumulations in local region of nuclear speckles (B). Each molecule color indicates molecular species as explained in Figure 2, and arrow width represents transition rates.

In the present model, RNAs are assumed to be random polymers. However, RNAs may exhibit various metastable partial folding structures containing many swinging double-strand branches. Additionally, the coupling of some proteins with RNAs may also modify the shape and effective size of RNAs. In these cases, the RNA model can be modified to include slightly smaller and more rigid particles than the presented model. However, the collisional behaviors among such particles are expected to be qualitatively similar to those among large and soft particles in simulations with an appropriate range of  $k_B T$ . Therefore, the results are expected to remain unchanged even after such modifications to the model.

In this study, the effective interactions among the representative components of condensates in

nuclear speckles were conjectured to be different from those expected from electrostatic interactions. Thus, these components were expected to form a specific affinity interaction network stronger than electrostatic interactions. However, the chemo-mechanical origins of these interactions cannot be clarified by present arguments. Therefore, further experimental studies are warranted to elucidate the specific molecular binding sites and binding manner of proteins. Further, the influences of protein modifications including serine phosphorylation of SR proteins and the possibility of interactions should be clarified via other molecules. Moreover, it should be noted that we only simulated a small part of the nuclear speckle with a few molecular condensates. Based on the proposed model with more particles, the simulation of whole nuclear morphological behaviors of nuclear speckles using higher performance computations is a crucial future issue.

### **Conclusion**

Overall, a coarse-grained molecular dynamics model of nuclear speckle was developed, and the interactions among representative molecules reproduced a rich variety of experimentally observed morphological changes in the droplet-like condensates of this nuclear body. The application of such arguments is expected to provide important insights into the mechanisms underlying the cell state- and cell cycle-dependent behaviors of other nuclear bodies.

### **Conflict of Interest**

The authors declare that they have no conflict of interests.

### **Author Contributions**

S.W., N.S., and A.A. conceived and designed the study; S.W. and A.A. conducted the mathematical model construction and simulations; S.W. and A.A. analyzed the data; S.W., N.S., and A.A. wrote the manuscript.

### **Data Availability**

The evidence data generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

### **Acknowledgments**

We thank Masashi Fujii for his guidance. This work was supported by Grants-in-Aid for Scientific Research (KAKENHI) [Grant Number 21K06124 (A.A.), 18H05531 (N.S.)] from the Japan Society for Promotion of Science. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics.

## References

- [1] Pederson, T. Dynamics and genome-centricity of interchromatin domains in the nucleus. *Nature Cell Biology*. 4(12), E287–E291 (2002).
- [2] Mao, Y. S., Zhang, B., Spector, D. L. Biogenesis and function of nuclear bodies. In *Trends in Genetics*. 27(8), 295–306 (2011). <https://doi.org/10.1016/j.tig.2011.05.006>
- [3] Matsumori, H., Watanabe, K., Tachiwana, H., Fujita, T., Ito, Y., Tokunaga, M., et al. Ribosomal protein L5 facilitates rDNA-bundled condensate and nucleolar assembly. *Life Science Alliance*. 5(7) (2022). <https://doi.org/10.26508/lsa.202101045>
- [4] Saitoh, N., Spahr, C. S., Patterson, S. D., Bubulya, P., Neuwald, A. F., Spector, D. L. Proteomic Analysis of Interchromatin Granule Clusters. *Molecular Biology of the Cell*. 15(8), 3876–3890 (2004). <https://doi.org/10.1091/mbc.E04>
- [5] Dopie, J., Sweredoski, M. J., Moradian, A., Belmont, A. S. Tyramide signal amplification mass spectrometry (TSA-MS) ratio identifies nuclear speckle proteins. *Journal of Cell Biology*. 219(9) (2020). <https://doi.org/10.1083/jcb.201910207>
- [6] Spector, D. L., Lamond, A. I. Nuclear speckles. *Cold Spring Harbor Perspectives in Biology*. 3(2), 1–12 (2011). <https://doi.org/10.1101/cshperspect.a000646>
- [7] Smith, K. P., Hall, L. L., & Lawrence, J. B. Nuclear hubs built on RNAs and clustered organization of the genome. In *Current Opinion in Cell Biology*. 64, 67–76 (2020). <https://doi.org/10.1016/j.ceb.2020.02.015>
- [8] Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*. 39(6), 925–938 (2010). <https://doi.org/10.1016/j.molcel.2010.08.011>
- [9] Thiry, M. Behavior of interchromatin granules during the cell cycle. In *Eur. J. Cell Biol.* 68, 14–24 (1995).
- [10] Lamond, A. I., Spector, D. L. Nuclear speckles: A model for nuclear organelles. In *Nature Reviews Molecular Cell Biology*. 4(8), 605–612 (2003). <https://doi.org/10.1038/nrm1172>
- [11] Hall, L. L., Smith, K. P., Byron, M., Lawrence, J. B. Molecular anatomy of a speckle. *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology*. 288(7), 664–675 (2006). <https://doi.org/10.1002/ar.a.20336>
- [12] Cajal, S. R. El núcleo de las células piramidales del cerebrohumano y de algunos mamíferos. *Trab. Lab. Invest. Biol.* 8, 27–62 (1910).
- [13] Kim, J., Han, K. Y., Khanna, N., Ha, T., Belmont, A. S. Nuclear speckle fusion via long-range directional motion regulates speckle morphology after transcriptional inhibition.

Journal of Cell Science. 132(8) (2019). <https://doi.org/10.1242/jcs.226563>

[14] Chen, Y., Belmont, A. S. Genome organization around nuclear speckles. In *Current Opinion in Genetics and Development*. 55, 91–99 (2019). Elsevier Ltd.

<https://doi.org/10.1016/j.gde.2019.06.008>

[15] Ilk, İ. A., Malszycki, M., Lübke, A. K., Schade, C., Meierhofer, D., Aktaş, T. Son and SRRM2 are essential for nuclear speckle formation. *Elife*. 9 (2020).

<https://doi.org/10.7554/eLife.60579>

[16] Zhang, Q., Kota, K. P., Alam, S. G., Nickerson, J. A., Dickinson, R. B., Lele, T. P.

Coordinated Dynamics of RNA Splicing Speckles in the Nucleus. *Journal of Cellular Physiology*. 231(6), 1269–1275 (2016). <https://doi.org/10.1002/jcp.25224>

[17] Johnson, C., Primorac, D., Mckinstry, M., Mcneil, J., Rowe, D., Lawrence, J. B. Tracking COL1A1 RNA in Osteogenesis Imperfecta: Splice-defective Transcripts Initiate Transport from the Gene but Are Retained within the SC35 Domain. *The Journal of Cell Biology*. 150(3), 417–432 (2020).

[18] Wang, K., Wang, L., Wang, J., Chen, S., Shi, M., Cheng, H. Intronless mRNAs transit through nuclear speckles to gain export competence. *Journal of Cell Biology*. 217(11), 3912–3929 (2018). <https://doi.org/10.1083/jcb.201801184>

[19] Fu, X. – D., Maniatis, T. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature*. 343(6257), 437–441 (1990).

[20] Fei, J., Jadaliha, M., Harmon, T. S., Li, I. T. S., Hua, B., Hao, Q., et al. Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *Journal of Cell Science*. 130(24), 4180–4192 (2017). <https://doi.org/10.1242/jcs.206854>

[21] Sharma, A., Takata, H., Shibahara, K., Bubulya, A., Bubulya, P. A. Son Is Essential for Nuclear Speckle Organization and Cell Cycle Progression. *Molecular Biology of the Cell*. 21, 650–663 (2010). <https://doi.org/10.1091/mbc.E09>

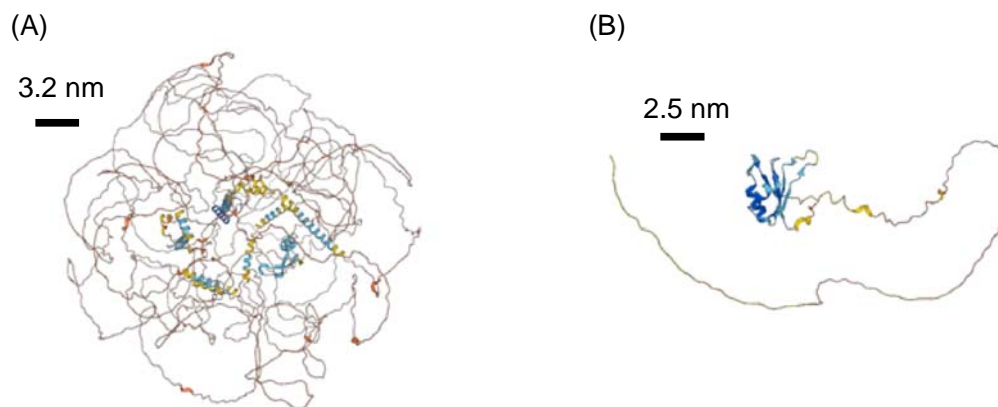
[22] O’Keefe, R. T., Mayeda, A., Sadowski, C. L., Krainer, A. R., Spector, D. L. Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. *Journal of Cell Biology*. 124(3), 249–260 (1994). <https://doi.org/10.1083/jcb.124.3.249>

[23] Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell*. 165(7), 1686–1697 (2016). <https://doi.org/10.1016/j.cell.2016.04.047>

[24] Yamazaki, T., Yamamoto, T., Yoshino, H., Souquere, S., Nakagawa, S., Pierron, G., et al. Paraspeckles are constructed as block copolymer micelles. *The EMBO Journal*. 40(12) (2021). <https://doi.org/10.15252/embj.2020107270>

[25] Fu, X.- D., Maniatis, T. Isolation of a Complementary DNA That Encodes the Mammalian Splicing Factor SC35. *Science*. 256(5056), 535–538 (1992).

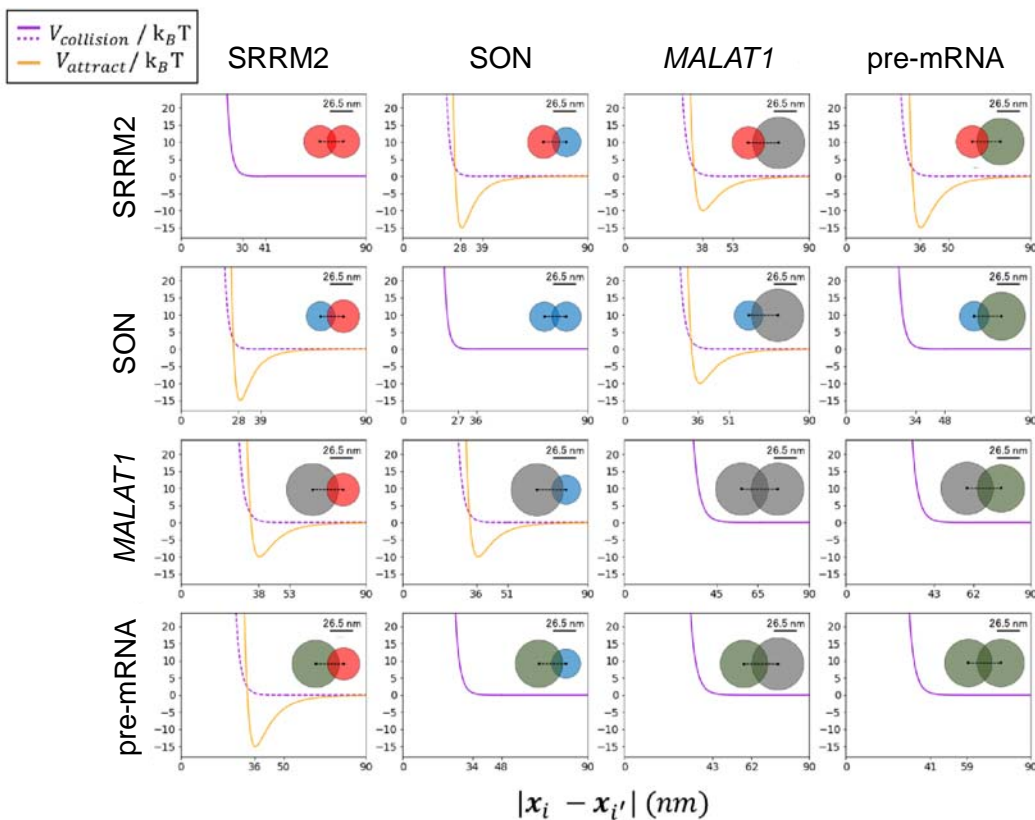
- [26] Blencowe, B. J., Baurén, G., Eldridge, A. G., Issner, R., Nickerson, J. A., Rosonina, E., et al. The SRm160/300 splicing coactivator subunits. *Rna*. 6(1), 111–120 (2000).
- [27] Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Ahmad, S., Alpi, E., et al. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*. (2022). <https://doi.org/10.1093/nar/gkac1052>
- [28] Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*. 50(D1), D439–D444 (2022). <https://doi.org/10.1093/nar/gkab1061>
- [29] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596(7873), 583–589 (2021). <https://doi.org/10.1038/s41586-021-03819-2>
- [30] Gao, J., Weiner, J. H. Excluded-Volume Effects in Rubber Elasticity. 2. Ideal Chain Assumption. 1. *Macromolecules*. 20(10), 2525–2531 (1987).
- [31] Lander, S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. Initial sequencing and analysis of the human genome. *Nature*. 409(6822), 860–921 (2001).
- [32] Lin, Y. T., Frömberg, D., Huang, W., Delivani, P., Chacón, M., Tolić, I. M., et al. Pulled Polymer Loops as a Model for the Alignment of Meiotic Chromosomes. *Physical Review Letters*. 115(20) (2015). <https://doi.org/10.1103/PhysRevLett.115.208102>
- [33] Ahn, E. Y., DeKolver, R. C., Lo, M. C., Nguyen, T. A., Matsuura, S., Boyapati, A., et al. SON Controls Cell-Cycle Progression by Coordinated Regulation of RNA Splicing. *Molecular Cell*. 42(2), 185–198 (2011). <https://doi.org/10.1016/j.molcel.2011.03.014>



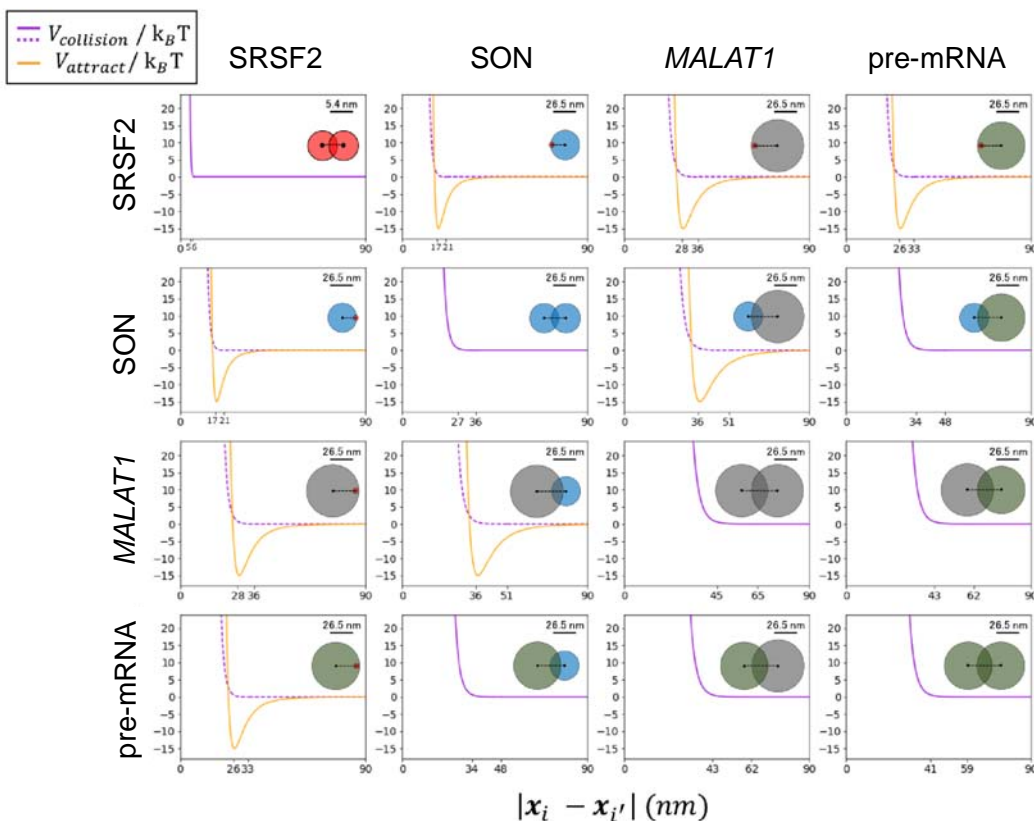
**Supplementary Figure S1** Ribbon descriptions of rough molecular shapes of (A) SON and (B) SRSF2

obtained from UniProt database [27]: ID P18583 and Q01130. The placement of secondary structures on the amino acid sequence is derived from these data. It would be natural to consider that the IDR is distributed more spatially than in these figures.

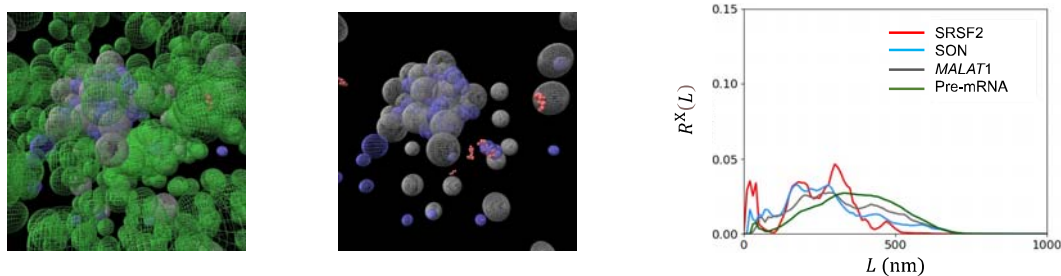
(A) RM2 model



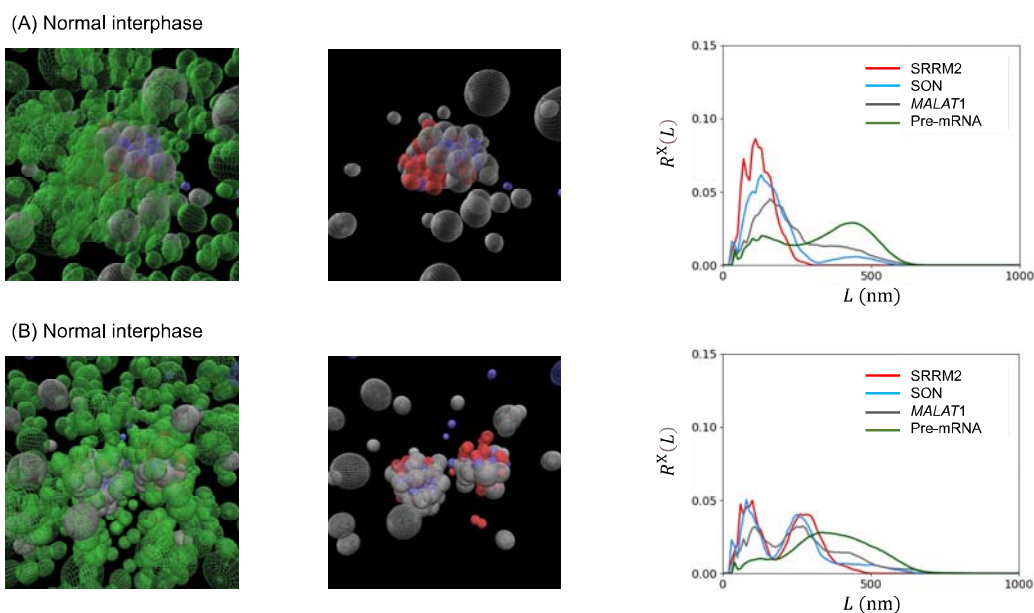
(B) SF2 model



**Supplementary Figure S2** Profiles of interaction potential between each pair of particles in (A) RM2 model and (B) SF2 model that correspond to the interaction network shown in Figure 3. Solid curve indicates each potential profile and broken curve indicates the function  $U(|\mathbf{x}_i - \mathbf{x}_{i'}|) = \theta((r_i + r_{i'}) - |\mathbf{x}_i - \mathbf{x}_{i'}|) k_{i,i'}^c \left[ -2 \left( \frac{r_i + r_{i'}}{|\mathbf{x}_i - \mathbf{x}_{i'}|} \right)^6 + \left( \frac{r_i + r_{i'}}{|\mathbf{x}_i - \mathbf{x}_{i'}|} \right)^{12} + 1 \right]$ . Accordingly,  $d_{i,i'} = |\mathbf{x}_i - \mathbf{x}_{i'}|$ , satisfying  $U(|\mathbf{x}_i - \mathbf{x}_{i'}|) = \frac{k_B T}{2}$  as eqn. (6) in the text. Each pair of two overlapped circles indicates the position of two particles when  $d_{i,i'} = |\mathbf{x}_i - \mathbf{x}_{i'}|$ . Color of each particle indicates molecular species as explained in Figure 2.



**Supplementary Figure S3** Simulation results of the SF2 model where the attractive strength parameters between SRSF2 and *MALAT1* particles and between SON and *MALAT1* particles was modified uniformly as  $k_{i,i}^a = 10 k_B T$ . Matrices of interaction networks among particles was given as the same as Figure 3A except that SRRM2 was replaced to SRSF2. Snapshots of the simulations (left), the same snapshots but without pre-mRNA (middle), and radial distributions of particles,  $R^{\text{SRSF2}}(L)$ ,  $R^{\text{SON}}(L)$ ,  $R^{\text{MALAT1}}(L)$ , and  $R^{\text{pre-mRNA}}(L)$  (right) in the case of normal interphase. Each particle color indicates molecular species as explained in Figure 2. In this case, particles tend to be divided into SRSF2-pre-mRNA particles condensates and SON-*MALAT1* particles condensates with pre-mRNA differently from the results in Figure 4–5.



**Supplementary Figure S4** Simulation results of the RM2 model with appropriately designed parameters with a modification that the attractive strength parameter between SRRM2 and pre-mRNA was assumed  $k_{i,i}^a = 16 k_B T$ . In this case, the attractive force between SRRM2 and SON was slightly weaker than that between SRRM2 and pre-mRNA. (A–B) Simulation results from different initial conditions.



Snapshots of the simulations (left) and the same snapshots but without pre-mRNA (middle) in the case of normal interphase. Each particle color indicates molecular species as explained in Figure 2. Qualitatively the same results as in Figure 3 were obtained in vast cases (A), but two particle condensates forming core-shell like structures were also observed in some cases (B). In latter case, the radial distributions of particles,  $R^{SRRM2}(L)$ ,  $R^{SON}(L)$ ,  $R^{MALAT1}(L)$ , and  $R^{\text{pre-mRNA}}(L)$  (right) exhibited two peaks where  $R^{SON}(dist)$ ,  $R^{SRRM2}(dist)$ , and  $R^{MALAT1}(dist)$  exhibited peaks at the similar values of  $dist$ , but  $R^{MALAT1}(dist)$  involved lower peaks and longer tail than  $R^{SON}(dist)$  and  $R^{SRRM2}(dist)$ .